Universidade Federal Fluminense

Instituto de Matemática e Estatística

Apostila de Estatística II

Ana Maria Lima de Farias

Departamento de Estatística

Agosto 2017

Conteúdo

Parte I

Inferência para uma população

Capítulo 1

Inferência estatística – Conceitos básicos

1.1 Introdução

A análise de um conjunto de dados por meio de técnicas descritivas (numéricas e gráficas) proporciona uma boa ideia da distribuição desses. Em particular, a distribuição de frequências é um instrumento bastante importante para avaliarmos a variabilidade das observações de um fenômeno aleatório. A partir dessas frequências, podemos calcular medidas de posição e variabilidade como, por exemplo, média, mediana, moda, desvio padrão etc. Tais frequências e medidas calculadas a partir dos dados são, em geral, estimativas de quantidades desconhecidas, associadas a populações das quais os dados foram extraídos na forma de amostras. As frequências relativas, por exemplo, são estimativas de probabilidades de ocorrência de certos eventos de interesse.

Quando realizamos uma análise de dados, é bastante razoável buscarmos alguma forma de regularidade/padrão (ou um modelo) presente nas observações. Com suposições adequadas, e sem observarmos diretamente o fenômeno aleatório de interesse, podemos criar modelos (matemáticos) teóricos capazes de reproduzir, de maneira satisfatória, a distribuição de frequências associada a um fenômeno aleatório diretamente observado. Tais modelos teóricos são chamados *modelos probabilísticos* e são objeto de estudo nas disciplinas de Teoria das Probabilidades.

Os modelos probabilísticos são, então, utilizados para medir a variabilidade de fenômenos aleatórios de acordo com as suas distribuições de probabilidades, que podem ser referentes a variáveis aleatórias discretas ou contínuas. Na prática, é comum o pesquisador ter alguma ideia sobre a forma da distribuição, mas não dos valores exatos dos parâmetros que a especificam. Surge, assim, a necessidade de descobrirmos (ou estimarmos) os parâmetros da distribuição para a sua posterior utilização.

EXEMPLO 1.1 Altura de adultos

Em um estudo antropométrico em nível nacional, uma amostra de 5000 adultos é selecionada dentre os adultos brasileiros e um dos objetivos é estimar a altura média dos adultos brasileiros.

- Neste exemplo, a população é o conjunto de todos os brasileiros adultos. No entanto, o interesse (um deles, pelo menos) está na altura dos brasileiros. Assim, nesse estudo, a cada sujeito da população associamos um número correspondente à sua altura. Como vimos, essa é a definição de variável aleatória: uma função que associa a cada ponto do espaço amostral (conjunto de todos os brasileiros) um número real. Dessa forma, a nossa população pode ser representada pela variável aleatória X = "altura do adulto brasileiro". Como essa é uma variável aleatória contínua, a ela está associada uma função densidade de probabilidade f e da literatura, sabemos que é razoável supor que essa seja a densidade normal. Assim, nossa população, nesse caso, é representada por uma variável aleatória $X \sim N\left(\mu; \sigma^2\right)$. Conhecendo os valores de μ e σ teremos informações completas sobre a nossa população.

- Uma forma de obtermos os valores de μ e σ é medindo as alturas de todos os brasileiros adultos. Mas esse seria um procedimento caro e demorado. Uma solução, então, é retirar uma amostra (subconjunto) da população e estudar essa amostra. Suponhamos que essa amostra seja retirada com reposição e que os sorteios sejam feitos de forma independente, isto é, o resultado de cada extração não altera o resultado das demais extrações. Ao sortearmos o primeiro elemento, estamos realizando um experimento que dá origem à variável aleatória X_1 ="altura do primeiro elemento"; o segundo elemento dá origem à variável aleatória X_2 ="altura do segundo elemento" e assim por diante. Como as extrações são feitas com reposição, todas as variáveis aleatórias $X_1, X_2, \ldots$ têm a mesma distribuição, que reflete a distribuição da altura de todos os brasileiros adultos. Para uma amostra específica, temos os valores observados $x_1, x_2, \ldots$ dessas variáveis aleatórias.

♦♦

EXEMPLO 1.2

Consideremos, agora, uma pesquisa eleitoral, em que estamos interessados no resultado do segundo turno de uma eleição presidencial brasileira. O interesse final é saber a proporção de votos em um e outro candidato (vamos simplificar a situação ignorando votos nulos, indecisos etc.).

- Mais uma vez, nossos sujeitos de pesquisa são pessoas com 16 anos ou mais, aptas a votar. O interesse final é saber a proporção de votos em cada um dos candidatos. Então, cada sujeito de pesquisa dá origem a uma variável aleatória binária, isto é, uma variável aleatória que assume apenas dois valores. Como visto, podemos representar esses valores por 1 (candidato A) e 0 (candidato B), o que define uma variável aleatória de Bernoulli, ou seja, essa população pode ser representada pela variável aleatória $X \sim Bern(p)$. O parâmetro p representa a probabilidade de um sujeito dessa população votar no candidato A. Uma outra interpretação é que p representa a proporção populacional de votantes no candidato A.

- Como não é viável entrevistar todos os eleitores, utiliza-se uma amostra de eleitores para se obter informação sobre p e cada sujeito de pesquisa indica o candidato em que vai votar (A ou B). Como antes, vamos supor que essa amostra seja retirada com reposição. Ao sortearmos o primeiro elemento, estamos realizando um experimento que dá origem à variável aleatória X_1 = "voto do primeiro elemento"; o segundo elemento dá origem à variável aleatória X_2 = "voto do segundo elemento" e assim por diante. Como as extrações são feitas com reposição, todas as variáveis aleatórias $X_1, X_2, \ldots$ têm a mesma distribuição de Bernoulli populacional, isto é, $X_i \sim Bern(p), i = 1, 2, \ldots$ e são independentes.

♦♦

EXEMPLO 1.3 Duração de lâmpadas

Suponha que estejamos interessados em estudar o tempo de vida, medido em horas, das lâmpadas produzidas por uma determinada empresa. Para esse tipo de teste, é necessário deixar as lâmpadas acesas até que se queimem.

- Neste exemplo, a população alvo é formada por todas as lâmpadas fabricadas ou que venham a ser fabricadas pela empresa, o que caracteriza uma população teoricamente infinita. Um modelo teórico (probabilístico) possível para a distribuição da variável populacional "tempo de vida" é a distribuição exponencial com parâmetro λ.

- Aqui é impossível trabalharmos com amostragem com reposição, uma vez que o experimento termina com a lâmpada selecionada queimada. No entanto, como a população é muito grande (infinita), extrações sem reposição dos elementos de uma amostra de tamanho n finito e, em geral, pequeno em relação ao tamanho da população, podem ser tratadas como extrações com reposição (você deve ter visto tal situação no estudo das distribuições binomial e hipergeométrica). Assim, nossa amostra pode ser vista, aproximadamente, como um conjunto de variáveis aleatórias $X_1, X_2, \ldots, X_n$ independentes e identicamente distribuídas com $X_i \sim \exp(\lambda)$.

♦♦

1.2 População

Como ilustrado nos exemplos acima, a inferência estatística trata do problema de se obter informação sobre uma *população* a partir de uma *amostra*. Embora a população real possa ser constituída de pessoas, empresas, animais etc., as pesquisas estatísticas buscam informações sobre determinadas características dos sujeitos, características essas que podem ser representadas por números. Sendo assim, a cada sujeito da população está associado um número, o que nos permite apresentar a seguinte definição.

DEFINIÇÃO **População**

A população de uma pesquisa estatística é uma variável aleatória X, com sua função de probabilidade, que descreve a característica de interesse.

Os métodos de inferência nos permitirão obter estimativas dos parâmetros da distribuição de probabilidade de tal variável aleatória, que pode ser contínua ou discreta.

1.3 Amostra aleatória simples

Como já dito, é bastante comum o emprego da amostragem em pesquisas estatísticas. Nas pesquisas por amostragem, uma amostra é selecionada da população de interesse e todas as conclusões serão baseadas apenas nessa amostra. Para que seja possível inferir resultados para a população a partir da amostra, é necessário que esta seja "representativa" da população.

Embora existam vários métodos de seleção de amostras, vamos nos concentrar aqui no caso mais simples, que é a *amostragem aleatória simples*. Segundo tal método, toda amostra de mesmo tamanho n tem igual chance (probabilidade) de ser sorteada. É possível extrair amostras aleatórias simples *com* e *sem* reposição. Quando estudamos as distribuições binomial e hipergeométrica, vimos que a distribuição binomial correspondia a extrações com reposição e a distribuição hipergeométrica correspondia a extrações sem reposição. No entanto, para populações grandes – ou infinitas – extrações com e sem reposição não levam a resultados muito diferentes. Assim, no estudo da Inferência Estatística, lidaremos sempre com amostragem aleatória simples *com* reposição. Este método de seleção atribui a cada elemento da população a mesma probabilidade de ser selecionado e esta probabilidade se mantém constante ao longo do processo de seleção da amostra (se as extrações fossem sem reposição isso não aconteceria). No restante desse curso omitiremos a expressão "com reposição", ou seja, o termo amostragem (ou amostra) aleatória simples sempre se referirá à amostragem com reposição. Por simplicidade, muitas vezes abreviaremos o termo amostra aleatória simples por aas.

Uma forma de se obter uma amostra aleatória simples é escrever os números ou nomes dos elementos da população em cartões iguais, colocar estes cartões em uma urna misturando-os bem e fazer os sorteios necessários, tendo o cuidado de colocar cada cartão sorteado na urna antes do próximo sorteio. Na prática, em geral são usados programas de computador, uma vez que as populações tendem a ser muito grandes.

Agora vamos formalizar o processo de seleção de uma amostra aleatória simples, de forma a relacioná-lo com os problemas de inferência estatística que iremos estudar.

Seja uma população representada por uma variável aleatória X. De tal população será sorteada uma amostra aleatória simples com reposição de tamanho n. Como visto nos exemplos anteriores, cada sorteio dá origem a uma variável aleatória X_i e, como os sorteios são com reposição, todas essas variáveis são independentes e têm a mesma distribuição de X. Isso nos leva à seguinte definição.

DEFINIÇÃO **Amostra aleatória simples**

Uma **amostra aleatória simples** (aas) de tamanho n de uma variável aleatória X (população) com distribuição de probabilidade f é um conjunto de n variáveis aleatórias $X_1, X_2, ..., X_n$ independentes e identicamente distribuídas (iid) com $X_i \sim f$.

É interessante notar a convenção usual: o valor observado de uma variável aleatória X é representado pela letra minúscula correspondente. Assim, depois do sorteio de uma amostra aleatória simples de tamanho n, temos valores observados $x_1, x_2, \ldots, x_n$ das respectivas variáveis aleatórias.

1.4 Estatísticas e parâmetros

Obtida uma amostra aleatória simples, é possível calcular diversas características desta amostra, como, por exemplo, a média, a mediana, a variância etc. Qualquer uma destas características é uma função de $X_1, X_2, ..., X_n$ e, portanto, é também uma variável aleatória (o seu valor depende da amostra sorteada). Por exemplo, a média amostral é a variável aleatória definida por

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Temos, então, a seguinte definição:

DEFINIÇÃO **Estatística amostral**

Uma **estatística amostral** (ou simplesmente estatística) T é qualquer função da amostra $X_1, X_2, ..., X_n$ que não dependa de parâmetros desconhecidos, isto é,

$$T = g(X_1, X_2, ..., X_n)$$

em que g é uma função qualquer que não depende de parâmetros desconhecidos.

Algumas estatísticas amostrais são

- média amostral

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \tag{1.1}$$

- variância amostral

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \tag{1.2}$$

- mínimo amostral

$$Y_{(1)} = \min\{X_1, X_2, \dots, X_n\}$$

- máximo amostral

$$Y_{(n)} = \max\{X_1, X_2, \dots, X_n\}$$

- amplitude amostral

$$W = Y_{(n)} - Y_{(1)}$$

Note que nenhuma das funções acima depende de qualquer parâmetro desconhecido. Por exemplo, a função $Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ não é uma estatística, pois depende dos parâmetros desconhecidos μ e σ.

É comum usar o termo **estimador** no lugar de estatística. Note que, sendo as estatísticas variáveis aleatórias, elas são representadas por letras maiúsculas: X, Y, Z etc. Para uma amostra específica, o valor obtido para o estimador será denominado **estimativa** e será representada por letras minúsculas. Por exemplo, temos as seguintes notações correspondentes à média e à variância amostrais:

- Estimadores: $\overline{X}$ e S^2
- Estimativas: $\overline{x}$ e s^2

De forma análoga, temos as características de interesse da população. No entanto, para diferenciar entre as duas situações (população e amostra), atribuímos nomes diferentes.

DEFINIÇÃO **Parâmetro**

Um **parâmetro** é uma característica da população.

Assim, se a população é representada pela variável aleatória X, alguns parâmetros são a esperança $E(X)$ (média) e a variância $\text{Var}(X)$ de X.

Com relação às características mais usuais, vamos usar a seguinte notação:

Característica	Parâmetro (população)	Estatística (amostra)
Média	μ	$\overline{X}$
Variância	σ^2	S^2
Número de elementos	N	n

Lembre-se que, para uma variável aleatória discreta (finita) uniforme,

$$\mu = \mathrm{E}(X) = \frac{1}{N}\sum_{i=1}^{N} X_i$$

$$\mathrm{Var}(X) = \frac{1}{N}\sum_{i=1}^{N}[X_i - \mathrm{E}(X)]^2 = \frac{1}{N}\sum_{i=1}^{N}[X_i - \mu]^2 = \frac{1}{N}\sum_{i=1}^{N} X_i^2 - \mu^2$$

1.5 Distribuições amostrais

Nos problemas de inferência, estamos interessados em estimar um parâmetro θ da população (por exemplo, a média populacional) através de uma amostra aleatória simples $X_1, X_2, ..., X_n$. Para isso, usamos uma estatística T (por exemplo, a média amostral) e, com base no valor obtido para T a partir de uma particular amostra, iremos tomar as decisões que o problema exige. Já foi dito que T é uma variável aleatória, uma vez que depende da amostra sorteada; amostras diferentes fornecerão diferentes valores para T.

EXEMPLO 1.4

Considere a população é $\{1, 3, 4, 8\}$, isto é, este é o conjunto dos valores da característica de interesse da população em estudo. Assim, para esta população, ou seja, para essa variável aleatória X temos

$$\mathrm{E}(X) = \mu = \frac{1}{4}(1 + 3 + 4 + 8) = 4$$

$$\begin{aligned}\mathrm{Var}(X) &= \sigma^2 = \frac{1}{4}\left[(1-4)^2 + (3-4)^2 + (4-4)^2 + (8-4)^2\right] \\ &= 6,5\end{aligned}$$

Suponha que dessa população iremos extrair uma amostra aleatória simples de tamanho 2 e a estatística que iremos calcular é a média amostral. Algumas possibilidades de amostra são 1,1, 1,3, 4,8, para as quais os valores da média amostral são 1, 2 e 6, respectivamente. Podemos ver, então, que há uma variabilidade nos valores da estatística e, assim, seria interessante que conhecêssemos tal variabilidade. Conhecendo tal variabilidade, temos condições de saber "quão infelizes" podemos ser no sorteio da amostra. No exemplo acima, as amostras 1,1 e 8,8 são as que têm média amostral mais afastada da verdadeira média populacional. Se esses valores tiverem chance muito mais alta do que os valores mais próximos de $\mathrm{E}(X)$, podemos ter sérios problemas.

Para conhecer o comportamento da média amostral, teríamos que conhecer todos os possíveis valores de $\overline{X}$, o que equivaleria a obter todas as possíveis amostras de tamanho 2 de tal população. Nesse exemplo, como só temos 4 elementos na população, a obtenção de todas as amostras aleatórias simples de tamanho 2 não é difícil.

Lembre-se do estudo de análise combinatória: como o sorteio é feito com reposição, em cada um dos sorteios temos 4 possibilidades. Logo, o número total de amostras aleatórias simples é $4 \times 4 = 16$. Por outro lado, em cada sorteio, cada elemento da população tem a mesma chance de ser sorteado; como são 4 elementos, cada elemento tem probabilidade 1/4 de ser sorteado. Finalmente, como os sorteios são independentes, para obter a probabilidade de um par de elementos pertencer à amostra basta multiplicar as probabilidades (lembre-se que $\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B)$ quando A e B são independentes). Na Tabela 1.1 a seguir listamos todas as possíveis amostras, com suas respectivas probabilidades e para cada uma delas, apresentamos o valor da média amostral.

Tabela 1.1 – Distribuição amostral da média amostral

Amostra	Probabilidade	Média amostral $\overline{x}$
$(1,1)$	$(1/4)\times(1/4)=1/16$	$(1+1)/2=1$
$(1,3)$	$(1/4)\times(1/4)=1/16$	$(1+3)/2=2$
$(1,4)$	$(1/4)\times(1/4)=1/16$	$(1+4)/2=2,5$
$(1,8)$	$(1/4)\times(1/4)=1/16$	$(1+8)/2=4,5$
$(3,1)$	$(1/4)\times(1/4)=1/16$	$(3+1)/2=2$
$(3,3)$	$(1/4)\times(1/4)=1/16$	$(3+3)/2=3$
$(3,4)$	$(1/4)\times(1/4)=1/16$	$(3+4)/2=3,5$
$(3,8)$	$(1/4)\times(1/4)=1/16$	$(3+8)/2=5,5$
$(4,1)$	$(1/4)\times(1/4)=1/16$	$(4+1)/2=2,5$
$(4,3)$	$(1/4)\times(1/4)=1/16$	$(4+3)/2=3,5$
$(4,4)$	$(1/4)\times(1/4)=1/16$	$(4+4)/2=4$
$(4,8)$	$(1/4)\times(1/4)=1/16$	$(4+8)/2=6$
$(8,1)$	$(1/4)\times(1/4)=1/16$	$(8+1)/2=4,5$
$(8,3)$	$(1/4)\times(1/4)=1/16$	$(8+3)/2=5,5$
$(8,4)$	$(1/4)\times(1/4)=1/16$	$(8+4)/2=6$
$(8,8)$	$(1/4)\times(1/4)=1/16$	$(8+8)/2=8$

Analisando esta tabela, podemos ver que os possíveis valores de $\overline{X}$ são 1; 2; 2,5; 3; 3,5; 4; 4,5; 5,5; 6; 8 e podemos construir a sua função de de probabilidade, notando, por exemplo, que o valor 2 pode ser obtido através de duas amostras: (1,3) ou (3,1). Como essas amostras correspondem a eventos mutuamente exclusivos, a probabilidade de se obter uma média amostral igual a 2 é

$$P\left(\overline{X}=2\right)=P\left[(1,3)\cup(3,1)\right]=P\left[(1,3)\right]+P\left[(3,1)\right]=\frac{1}{16}+\frac{1}{16}=\frac{2}{16}$$

Com o mesmo raciocínio, obtemos a seguinte função de probabilidade para $\overline{X}$:

$\overline{x}$	1	2	2,5	3	3,5	4	4,5	5,5	6	8
$P(\overline{X}=\overline{x})$	1/16	2/16	2/16	1/16	2/16	1/16	2/16	2/16	2/16	1/16

Note que a variável aleatória de interesse aqui é $\overline{X}$! Daí segue que

$$\begin{aligned}
E(\overline{X}) &= 1\times\frac{1}{16}+2\times\frac{2}{16}+2,5\times\frac{2}{16}+3\times\frac{1}{16}+3,5\times\frac{2}{16}+\\
&\quad 4\times\frac{1}{16}+4,5\times\frac{2}{16}+5,5\times\frac{2}{16}+6\times\frac{2}{16}+8\times\frac{1}{16}\\
&= 4,0=\mu
\end{aligned}$$

$$\begin{aligned}
E(\overline{X}^2) &= 1^2\times\frac{1}{16}+2^2\times\frac{2}{16}+2,5^2\times\frac{2}{16}+3^2\times\frac{1}{16}+3,5^2\times\frac{2}{16}+\\
&\quad 4^2\times\frac{1}{16}+4,5^2\times\frac{2}{16}+5,5^2\times\frac{2}{16}+6^2\times\frac{2}{16}+8^2\times\frac{1}{16}=19,25\\
\mathrm{Var}(\overline{X}) &= 19,25-4^2=3,25=\frac{6,5}{2}=\frac{\sigma^2}{2}=\frac{\sigma^2}{n}
\end{aligned}$$

Neste exemplo podemos ver que $E(\overline{X})=\mu$ e $\mathrm{Var}(\overline{X})=\sigma^2/2$, onde 2 é o tamanho da amostra. Esses resultados estão nos dizendo que, em média (esperança), a estatística $\overline{X}$ é igual à média da população e que sua variância é igual à variância da população dividida pelo tamanho da amostra. Nas Figuras 1.1

e 1.2 temos os gráficos da função de probabilidade de X (população) e de $\overline{X}$ (amostra), respectivamente. Podemos ver que a distribuição de $\overline{X}$ tem menor dispersão em torno da média $\mu = 4$. Note que essa média e essa variância são calculadas ao longo de todas as possíveis amostras aleatórias simples de tamanho 2.

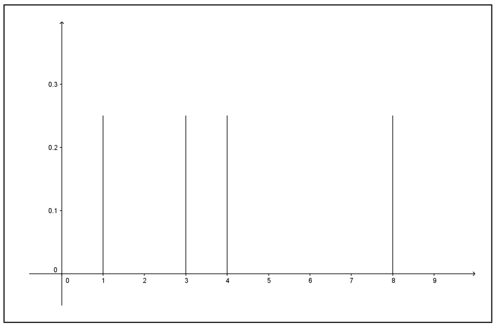


Figura 1.1 – Distribuição de probabilidade de X (população)

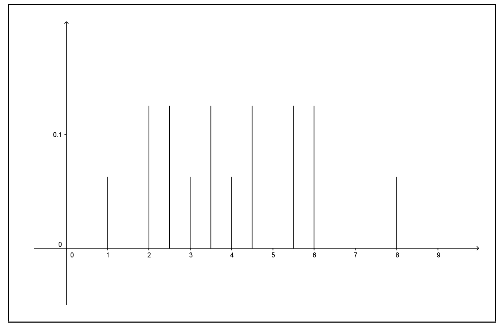


Figura 1.2 – Distribuição de probabilidade de $\overline{X}$ – ($n = 2$)

Consideremos, agora, a mesma situação, só que, em vez de estudarmos a média amostral, uma medida de posição, vamos estudar a dispersão. Como visto, a variância populacional é $\text{Var}(X) = 6,5$. Para a amostra, vamos trabalhar com dois estimadores. Um deles será S^2, definido na Equação (1.2) e o outro,

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2 \tag{1.3}$$

Da mesma forma que fizemos para a média amostral, vamos calcular o valor dessas estatísticas para cada uma das amostras. Na Tabela 1.2 temos os resultados parciais e globais de interesse.

Tabela 1.2 – Distribuição amostral de 2 estimadores da variância

Amostra	$\overline{x}$	$(x_1 - \overline{x})^2$	$(x_2 - \overline{x})^2$	$\sum_{i=1}^{2}(x_i - \overline{x})^2$	S^2	$\widehat{\sigma}^2$
$(1,1)$	1	$(1-1)^2$	$(1-1)^2$	0	0	0
$(1,3)$	2	$(1-2)^2$	$(3-2)^2$	2	2	1
$(1,4)$	$2,5$	$(1-2,5)^2$	$(4-2,5)^2$	$4,5$	$4,5$	$2,25$
$(1,8)$	$4,5$	$(1-4,5)^2$	$(8-4,5)^2$	$24,5$	$24,5$	$12,25$
$(3,1)$	2	$(3-2)^2$	$(1-2)^2$	2	2	1
$(3,3)$	3	$(3-3)^2$	$(3-3)^2$	0	0	0
$(3,4)$	$3,5$	$(3-3,5)^2$	$(4-3,5)^2$	$0,5$	$0,5$	$0,25$
$(3,8)$	$5,5$	$(3-5,5)^2$	$(8-5,5)^2$	$12,5$	$12,5$	$6,25$
$(4,1)$	$2,5$	$(4-2,5)^2$	$(1-2,5)^2$	$4,5$	$4,5$	$2,25$
$(4,3)$	$3,5$	$(4-3,5)^2$	$(3-3,5)^2$	$0,5$	$0,5$	$0,25$
$(4,4)$	4	$(4-4)^2$	$(4-4)^2$	0	0	0
$(4,8)$	6	$(4-6)^2$	$(8-6)^2$	8	8	4
$(8,1)$	$4,5$	$(8-4,5)^2$	$(1-4,5)^2$	$24,5$	$24,5$	$12,25$
$(8,3)$	$5,5$	$(8-5,5)^2$	$(3-5,5)^2$	$12,5$	$12,5$	$6,25$
$(8,4)$	6	$(8-6)^2$	$(4-6)^2$	8	8	4
$(8,8)$	8	$(8-8)^2$	$(8-8)^2$	0	0	0

Podemos ver que a função de probabilidade de S^2 é

s^2	0	$0,5$	2	$4,5$	8	$12,5$	$24,5$
$P(S^2 = s^2)$	4/16	2/16	2/16	2/16	2/16	2/16	2/16

e a função de probabilidade de $\widehat{\sigma}^2$ é

k	0	0,25	1	2,25	4	6,25	12,25
$P(\widehat{\sigma}^2 = k)$	4/16	2/16	2/16	4/16	2/16	2/16	2/16

Para essas distribuições temos:

$$E(S^2) = 0 \times \frac{4}{16} + \frac{2}{16}(0,5 + 2 + 4,5 + 8 + 12,5 + 24,5) = \frac{104}{16} = 6,5 = \sigma^2 = \mathrm{Var}(X)$$

e

$$E(\widehat{\sigma}^2) = 0 \times \frac{4}{16} + \frac{2}{16}(0,25 + 1 + 2,25 + 4 + 6,25 + 12,25) = \frac{52}{16} = 3,25$$

Vemos que, em média, S^2 é igual à variância populacional, o que não ocorre com $\widehat{\sigma}^2$.

♦♦

Este exemplo ilustra o fato de que qualquer estatística amostral T é uma variável aleatória, que assume diferentes valores para cada uma das diferentes amostras e tais valores, juntamente com a probabilidade de cada amostra, nos forneceriam a função de probabilidades de T, caso fosse possível obter todas as amostras aleatórias simples de tamanho n da população. Isso nos leva à seguinte definição, que é um conceito central na Inferência Estatística.

DEFINIÇÃO **Distribuição amostral de um estimador**

A **distribuição amostral** de uma estatística T é a função de probabilidades de T ao longo de todas as possíveis amostras aleatórias simples de tamanho n.

Podemos ver que a obtenção da distribuição amostral de qualquer estatística T é um processo tão ou mais complicado do que trabalhar com a população inteira. Na prática, o que temos é uma única amostra e é com essa única amostra que temos que tomar as decisões pertinentes ao problema em estudo. Esta tomada de decisão, no entanto, será facilitada se conhecermos resultados teóricos sobre o comportamento da distribuição amostral.

1.6 Propriedades de estimadores

No exemplo anterior, relativo à variância amostral, vimos que $E(S^2) = \sigma^2$ e $E(\widehat{\sigma}^2) \neq \sigma^2$. Analogamente, vimos também que $E(\overline{X}) = \mu$. Vamos explorar um pouco mais o significado desses resultados antes de passar a uma definição formal da propriedade envolvida.

Dada uma população, existem muitas e muitas amostras aleatórias simples de tamanho n que podem ser sorteadas. Cada uma dessas amostras resulta em um valor diferente da estatística de interesse ($\overline{X}$ e S^2, por exemplo). O que esses resultados estão mostrando é como esses diferentes valores se comportam em relação ao verdadeiro (mas desconhecido) valor do parâmetro.

Considere a Figura 1.3, em que o alvo representa o valor do parâmetro e os "tiros", indicados pelo símbolo x, representam os diferentes valores amostrais da estatística de interesse.

Nas partes (a) e (b) da figura, os tiros estão em torno do alvo, enquanto nas partes (c) e (d) isso não acontece. Comparando as partes (a) e (b), podemos ver que na parte (b) os tiros estão mais concentrados

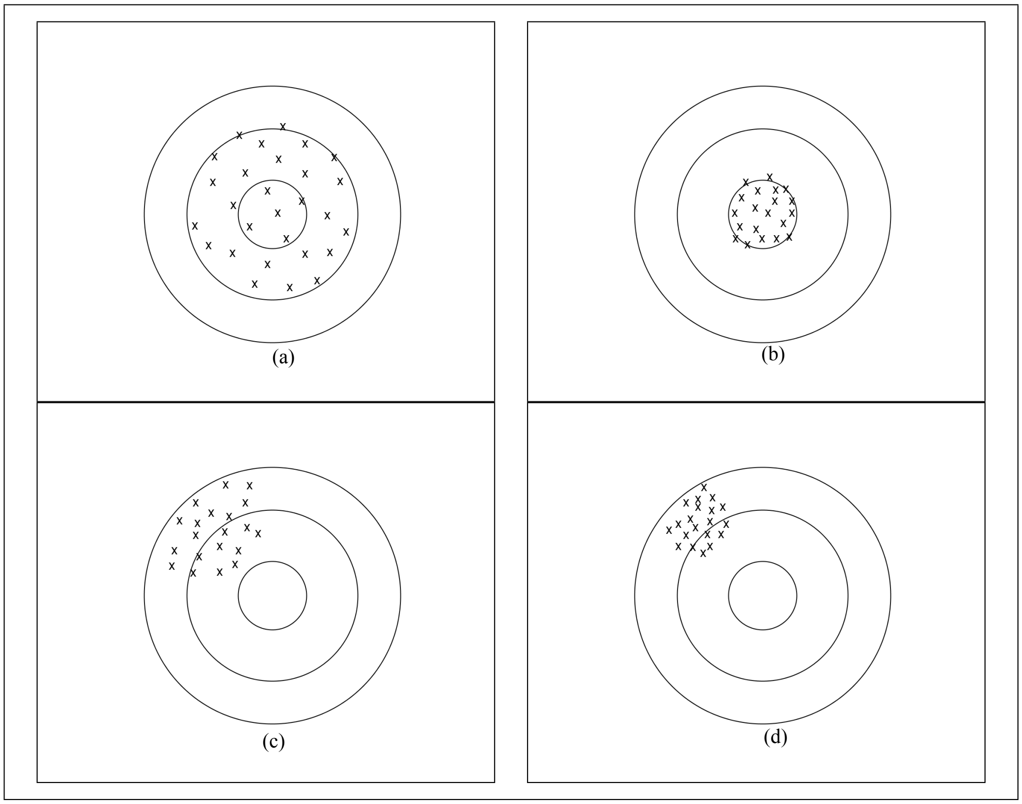


Figura 1.3 – Propriedades de estimadores

em torno do alvo, isto é, têm menor dispersão. Isso refletiria uma pontaria mais certeira do atirador em (b). Analogamente, nas partes (c) e (d), embora ambos os atiradores estejam com a mira deslocada, os tiros do atirador (d) estão mais concentrados em torno de um alvo; o deslocamento poderia até ser resultado de um desalinhamento da arma. Já o atirador (c), além de estar com o alvo deslocado, ele tem os tiros mais espalhados, o que reflete menor precisão.

- Nas partes (a) e (b), temos dois estimadores que fornecem estimativas centradas em torno do verdadeiro valor do parâmetro, ou seja, as diferentes amostras fornecem valores distribuídos em torno do verdadeiro valor do parâmetro. A diferença é que em (a) esses valores estão mais dispersos e, assim, temos mais chance de obter uma amostra "infeliz", ou seja, uma amostra que forneça um resultado muito afastado do valor do parâmetro. Essas duas propriedades estão associadas à esperança e à variância do estimador, que são medidas de centro e dispersão, respectivamente.

- Nas partes (c) e (d), as estimativas estão centradas em torno de um valor diferente do parâmetro de interesse e, na parte (c), a dispersão é maior.

Temos, assim, ilustrados os seguintes conceitos.

DEFINIÇÃO Viés de um estimador

Seja $X_1, X_2, \cdots, X_n$ uma amostra aleatória simples de uma população X, cuja lei de probabilidade depende de um parâmetro θ. Se T é um estimador de θ, definimos seu **viés** ou **vício** como

$$\mathrm{B}(T) = \mathrm{E}(T) - \theta \tag{1.4}$$

Se $\mathrm{B}(T) = 0$ então $\mathrm{E}(T) = \theta$ e dizemos que T é um **estimador não-viesado** de θ.

Como nos exemplos vistos, a esperança $\mathrm{E}(T)$ é calculada ao longo de todas as possíveis amostras, ou seja, é a esperança da distribuição amostral de T. Nas partes (a) e (b) da Figura 1.3 os estimadores são não-viesados e nas partes (c) e (d), os estimadores são viesados.

Com relação aos estimadores $\overline{X}$, S^2 e $\widehat{\sigma}^2$, provaremos, no próximo capítulo, que os dois primeiros são não-viesados para estimar a média e a variância populacionais, respectivamente, enquanto $\widehat{\sigma}^2$ é viesado para estimar a variância populacional.

DEFINIÇÃO Eficiência de um estimador

Se T_1 e T_2 são dois estimadores não-viesados do parâmetro θ, diz-se que T_1 é **mais eficiente** que T_2, se $\mathrm{Var}(T_1) < \mathrm{Var}(T_2)$.

Na Figura 1.3, o estimador da parte (b) é mais eficiente que o estimador da parte (a).

É interessante observar que o conceito de eficiência, que envolve a variabilidade de um estimador, está associado a estimadores não-viesados. Para analisar estimadores viesados, podemos usar o erro quadrático médio, definido a seguir.

DEFINIÇÃO Erro quadrático médio

Seja $X_1, X_2, \cdots, X_n$ uma amostra aleatória simples de uma população X, cuja lei de probabilidade depende de um parâmetro θ. Se T é um estimador de θ, definimos seu **erro quadrático médio** como

$$\mathrm{EQM}(T) = \mathrm{E}\,(T - \theta)^2 \tag{1.5}$$

Lembrando que a esperança de uma constante é a própria constante, podemos ver que

$$\begin{aligned}
\mathrm{EQM}(T) &= \mathrm{E}\,(T-\theta)^2 = \mathrm{E}\,[(T-\mathrm{E}(T)+\mathrm{E}(T)-\theta)]^2 \\
&= \mathrm{E}\,[T-\mathrm{E}(T)]^2 + \mathrm{E}\,[\mathrm{E}(T)-\theta]^2 + 2\,\mathrm{E}\,\{[T-\mathrm{E}(T)][\mathrm{E}(T)-\theta]\} \\
&= \mathrm{Var}(T) + [\mathrm{E}(T)-\theta]^2 + 2\,[\mathrm{E}(T)-\theta]\,\mathrm{E}\,[T-\mathrm{E}(T)] \\
&= \mathrm{Var}(T) + [\mathrm{B}(T)]^2 + 2\,[\mathrm{E}(T)-\theta]\,\cancel{[\mathrm{E}(T)-\mathrm{E}(T)]} \\
&= \mathrm{Var}(T) + [\mathrm{B}(T)]^2 \qquad (1.6)
\end{aligned}$$

A equação (1.6) decompõe o erro quadrático médio em termos da variância e do quadrado do vício do estimador. Para estimadores não-viesados, resulta que $\mathrm{EQM}(T) = \mathrm{Var}(T)$. Estimadores viesados podem ser uma opção interessante para estimar um parâmetro se seu erro quadrático médio for pequeno.

EXEMPLO 1.5 $X \sim \exp(\lambda)$ (Larson (1982))

Seja X_1, X_2 uma amostra aleatória simples de uma população $X \sim \exp(\lambda)$; logo, a função densidade de X é

$$f(x) = \lambda e^{-\lambda x} \qquad x > 0$$

e $\mathrm{E}(X) = \mu = \dfrac{1}{\lambda}$ e $\mathrm{Var}(X) = \sigma^2 = \dfrac{1}{\lambda^2}$. Considerando a média amostral como estimador de μ, temos que

$$\mathrm{E}(\overline{X}) = \frac{1}{2}\,[\mathrm{E}(X_1) + \mathrm{E}(X_2)] = \frac{1}{2}\left(\frac{1}{\lambda} + \frac{1}{\lambda}\right) = \frac{1}{\lambda}$$

$$\mathrm{Var}(\overline{X}) = \frac{1}{4}\,[\mathrm{Var}(X_1) + \mathrm{Var}(X_2)] = \frac{1}{4}\left(\frac{1}{\lambda^2} + \frac{1}{\lambda^2}\right) = \frac{1}{2\lambda^2}$$

Logo, $\overline{X}$ é não-viesado para estimar μ e, portanto, $\mathrm{EQM}(\overline{X}) = \mathrm{Var}(\overline{X})$.

Consideremos, agora, o estimador dado pela média geométrica de X_1, X_2, ou seja,

$$\widehat{\mu}_1 = \sqrt{X_1 X_2}$$

Para calcular o erro quadrático médio de $\widehat{\mu}_1$, precisamos dos seguintes resultados sobre a função gama:

$$\begin{aligned}
\Gamma(\alpha) &= \int_0^{+\infty} x^{\alpha-1} e^{-x} \mathrm{d}x \\
\Gamma\left(\frac{1}{2}\right) &= \sqrt{\pi} \\
\Gamma(\alpha) &= (\alpha-1)\Gamma(\alpha-1) \Rightarrow \Gamma\left(\frac{3}{2}\right) = \left(\frac{3}{2}-1\right)\Gamma\left(\frac{3}{2}-1\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{1}{2}\sqrt{\pi}
\end{aligned}$$

$$\mathrm{E}(\sqrt{X}) = \int_0^{\infty} \sqrt{x}\lambda e^{-\lambda x} \mathrm{d}x \underbrace{=}_{t=\lambda x} \int_0^{\infty} \left(\frac{t}{\lambda}\right)^{1/2} \lambda e^{-t} \frac{\mathrm{d}t}{\lambda} = \int_0^{\infty} \frac{1}{\sqrt{\lambda}} t^{1/2} e^{-t} \mathrm{d}t = \frac{1}{\sqrt{\lambda}} \int_0^{\infty} t^{3/2-1} e^{-t} dt = \frac{\Gamma\left(\frac{3}{2}\right)}{\sqrt{\lambda}} = \frac{1}{2}\sqrt{\frac{\pi}{\lambda}}$$

$$
\begin{aligned}
\mathrm{E}(\widehat{\mu}_1) &= \mathrm{E}\left(\sqrt{X_1X_2}\right) \underbrace{=}_{\text{indep.}} \mathrm{E}\left(\sqrt{X_1}\right)\mathrm{E}\left(\sqrt{X_2}\right) = \left(\frac{1}{2}\sqrt{\frac{\pi}{\lambda}}\right)^2 = \frac{\pi}{4\lambda} \\
\mathrm{E}\left(\widehat{\mu}_1^2\right) &= \mathrm{E}\left(\sqrt{X_1X_2}\right)^2 = \mathrm{E}(X_1X_2) = \mathrm{E}(X_1)\,\mathrm{E}(X_2) = \frac{1}{\lambda^2} \\
\mathrm{Var}(\widehat{\mu}_1) &= \mathrm{E}\left(\widehat{\mu}_1^2\right) - [\mathrm{E}(\widehat{\mu}_1)]^2 = \frac{1}{\lambda^2} - \left(\frac{\pi}{4\lambda}\right)^2 = \frac{16-\pi^2}{16\lambda^2} \\
\mathrm{B}(\widehat{\mu}_1) &= \mathrm{E}(\widehat{\mu}_1) - \frac{1}{\lambda} = \frac{\pi}{4\lambda} - \frac{1}{\lambda} = \frac{\pi-4}{4\lambda} \\
\mathrm{EQM}(\widehat{\mu}_1) &= \mathrm{Var}(\widehat{\mu}_1) + [\mathrm{B}(\widehat{\mu}_1)]^2 = \frac{16-\pi^2}{16\lambda^2} + \left(\frac{\pi-4}{4\lambda}\right)^2 = (4-\pi)\cdot\frac{1}{2\lambda^2} = (4-\pi)\cdot\mathrm{Var}(\overline{X}) \underbrace{<}_{4-\pi<1} \mathrm{Var}(\overline{X})
\end{aligned}
$$

Embora $\widehat{\mu}_1$ seja um estimador viesado, seu erro quadrático médio é menor do que o de $\overline{X}$, um estimador não-viesado.

♦♦

Uma outra propriedade dos estimadores está relacionada à ideia bastante intuitiva de que, à medida que se aumenta o tamanho da amostra, mais perto devemos ficar do verdadeiro valor do parâmetro.

DEFINIÇÃO Consistência

Uma sequência $\{T_n\}$ de estimadores de um parâmetro θ é consistente se, para todo $\varepsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}\{|T_n - \theta| > \varepsilon\} = 0$$

De maneira informal, o que essa definição nos diz é que a probabilidade de a distância entre T_n e θ, dada por $|T_n - \theta|$, ser"grande" ($> \epsilon$) tende a zero quando o tamanho da amostra cresce.

Uma maneira alternativa de verificar se uma sequência de estimadores é consistente é dada a seguir.

TEOREMA 1.1

Uma sequência $\{T_n\}$ de estimadores de um parâmetro θ é consistente se

$$
\begin{aligned}
\lim_{n\to\infty} E(T_n) &= \theta \\
\lim_{n\to\infty} Var(T_n) &= 0
\end{aligned}
$$

▼

EXEMPLO 1.6 A média amostral

Vimos, através de exemplos, que $\mathrm{E}(\overline{X}) = \mu$ e $\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}$. Resulta, pelo Teorema 1.1, que $\overline{X}_n$ é consistente para estimar a média populacional μ.

♦♦

EXEMPLO 1.7

Amostras de tamanho 3 serão retiradas da população $\{1, 2, 4, 6, 8\}$. Considere os seguintes estimadores para a média da população:

- média amostral:
$$\overline{X} = \frac{X_1 + X_2 + X_3}{3}$$
- média amostral ponderada:
$$\overline{X}_p = \frac{X_1 + 2X_2 + X_3}{4}$$
- ponto médio
$$\Delta = \frac{\min(X_1, X_2, X_3) + \max(X_1, X_2, X_3)}{2}$$

Mostre que

(a) $\overline{X}$ e $\overline{X}_p$ são não-viesados;

(b) $\overline{X}$ é mais eficiente que $\overline{X}_p$;

(c) Δ é viesado, mas sua variância é menor que a variância de $\overline{X}$ e de $\overline{X}_p$.

Solução

Para a população temos que

$$\begin{aligned}
\mu &= \frac{1+2+4+6+8}{5} = 4,2 \\
\sigma^2 &= \frac{1^2+2^2+4^2+6^2+8^2}{5} - (4,2)^2 = 6,56
\end{aligned}$$

Pelo princípio da multiplicação, o número total de amostras é $5 \times 5 \times 5 = 125$ e cada uma dessas amostras tem probabilidade $\frac{1}{5} \times \frac{1}{5} \times \frac{1}{5} = \frac{1}{125}$.

Na Tabela 1.3 estão listadas as 125 amostras possíveis, juntamente com os valores de cada um dos estimadores.

Tabela 1.3 – Distribuição amostral da média amostral

X_1	X_2	X_3	$\overline{X}$	$\overline{X}_p$	Δ	X_1	X_2	X_3	$\overline{X}$	$\overline{X}_p$	Δ	X_1	X_2	X_3	$\overline{X}$	$\overline{X}_p$	Δ
1	1	1	1	1	1	2	8	1	11/3	4,75	4,5	6	6	1	13/3	4,75	3,5
1	1	2	4/3	1,25	1,5	2	8	2	4	5	5	6	6	2	14/3	5	4
1	1	4	2	1,75	2,5	2	8	4	14/3	5,5	5	6	6	4	16/3	5,5	5
1	1	6	8/3	2,25	3,5	2	8	6	16/3	6	5	6	6	6	6	6	6
1	1	8	10/3	2,75	4,5	2	8	8	6	6,5	5	6	6	8	20/3	6,5	7
1	2	1	4/3	1,5	1,5	4	1	1	2	1,75	2,5	6	8	1	5	5,75	4,5
1	2	2	5/3	1,75	1,5	4	1	2	7/3	2	2,5	6	8	2	16/3	6	5
1	2	4	7/3	2,25	2,5	4	1	4	3	2,5	2,5	6	8	4	6	6,5	6
1	2	6	3	2,75	3,5	4	1	6	11/3	3	3,5	6	8	6	20/3	7	7
1	2	8	11/3	3,25	4,5	4	1	8	13/3	3,5	4,5	6	8	8	22/3	7,5	7
1	4	1	2	2,5	2,5	4	2	1	7/3	2,25	2,5	8	1	1	10/3	2,75	4,5
1	4	2	7/3	2,75	2,5	4	2	2	8/3	2,5	3	8	1	2	11/3	3	4,5
1	4	4	3	3,25	2,5	4	2	4	10/3	3	3	8	1	4	13/3	3,5	4,5
1	4	6	11/3	3,75	3,5	4	2	6	4	3,5	4	8	1	6	5	4	4,5
1	4	8	13/3	4,25	4,5	4	2	8	14/3	4	5	8	1	8	17/3	4,5	4,5
1	6	1	8/3	3,5	3,5	4	4	1	3	3,25	2,5	8	2	1	11/3	3,25	4,5
1	6	2	3	3,75	3,5	4	4	2	10/3	3,5	3	8	2	2	4	3,5	5
1	6	4	11/3	4,25	3,5	4	4	4	4	4	4	8	2	4	14/3	4	5
1	6	6	13/3	4,75	3,5	4	4	6	14/3	4,5	5	8	2	6	16/3	4,5	5
1	6	8	5	5,25	4,5	4	4	8	16/3	5	6	8	2	8	6	5	5
1	8	1	10/3	4,5	4,5	4	6	1	11/3	4,25	3,5	8	4	1	13/3	4,25	4,5
1	8	2	11/3	4,75	4,5	4	6	2	4	4,5	4	8	4	2	14/3	4,5	5
1	8	4	13/3	5,25	4,5	4	6	4	14/3	5	5	8	4	4	16/3	5	6
1	8	6	5	5,75	4,5	4	6	6	16/3	5,5	5	8	4	6	6	5,5	6
1	8	8	17/3	6,25	4,5	4	6	8	6	6	6	8	4	8	20/3	6	6
2	1	1	4/3	1,25	1,5	4	8	1	13/3	5,25	4,5	8	6	1	5	5,25	4,5
2	1	2	5/3	1,5	1,5	4	8	2	14/3	5,5	5	8	6	2	16/3	5,5	5
2	1	4	7/3	2	2,5	4	8	4	16/3	6	6	8	6	4	6	6	6
2	1	6	3	2,5	3,5	4	8	6	6	6,5	6	8	6	6	20/3	6,5	7
2	1	8	11/3	3	4,5	4	8	8	20/3	7	6	8	6	8	22/3	7	7
2	2	1	5/3	1,75	1,5	6	1	1	8/3	2,25	3,5	8	8	1	17/3	6,25	4,5
2	2	2	2	2	2	6	1	2	3	2,5	3,5	8	8	2	6	6,5	5
2	2	4	8/3	2,5	3	6	1	4	11/3	3	3,5	8	8	4	20/3	7	6
2	2	6	10/3	3	4	6	1	6	13/3	3,5	3,5	8	8	6	22/3	7,5	7
2	2	8	4	3,5	5	6	1	8	5	4	4,5	8	8	8	8	8	8
2	4	1	7/3	2,75	2,5	6	2	1	3	2,75	3,5						
2	4	2	8/3	3	3	6	2	2	10/3	3	4						
2	4	4	10/3	3,5	3	6	2	4	4	3,5	4						
2	4	6	4	4	4	6	2	6	14/3	4	4						
2	4	8	14/3	4,5	5	6	2	8	16/3	4,5	5						
2	6	1	3	3,75	3,5	6	4	1	11/3	3,75	3,5						
2	6	2	10/3	4	4	6	4	2	4	4	4						
2	6	4	4	4,5	4	6	4	4	14/3	4,5	5						
2	6	6	14/3	5	4	6	4	6	16/3	5	5						
2	6	8	16/3	5,5	5	6	4	8	6	5,5	6						

A partir daí, obtemos as seguintes distribuições amostrais:

x	$P(\overline{X} = x)$
1	1/125
4/3	3/125
5/3	3/125
2	4/125
7/3	6/125
8/3	6/125
3	9/125
10/3	9/125
11/3	12/125
4	10/125
13/3	9/125
14/3	12/125
5	6/125
16/3	12/125
17/3	3/125
6	10/125
20/3	6/125
22/3	3/125
8	1/125

$$E(\overline{X}) = 1\times\frac{1}{125}+\frac{4}{3}\times\frac{3}{125}+\ldots+8\times\frac{1}{125} = 4,2 = \mu$$

$$E(\overline{X}^2) = 1^2\times\frac{1}{125}+\left(\frac{4}{3}\right)^2\times\frac{3}{125}+\ldots+8^2\times\frac{1}{125} = \frac{22375}{1125}$$

$$Var(\overline{X}) = \frac{22305}{1125}-(4,2)^2 = 2,186667 = \frac{6,56}{3} = \frac{\sigma^2}{3}$$

x	$P(\overline{X}_p = x)$
4/4	1/125
5/4	2/125
6/4	2/125
7/4	4/125
8/4	3/125
9/4	4/125
10/4	6/125
11/4	6/125
12/4	8/125
13/4	4/125
14/4	10/125
15/4	4/125
16/4	9/125
17/4	4/125
18/4	10/125
19/4	4/125
20/4	8/125
21/4	4/125
22/4	8/125
23/4	2/125
24/4	7/125
25/4	2/125
26/4	6/125
28/4	4/125
30/4	2/125
32/4	1/125

$$E(\overline{X}_p) = 1\times 1125+\frac{5}{4}\times\frac{2}{125}+\ldots+8\times\frac{1}{125} = 4,2 = \mu$$

$$E(\overline{X}_p^2) = 1^2\times 1125+\left(\frac{5}{4}\right)^2\times\frac{2}{125}+\ldots+8^2\times\frac{1}{125} = \frac{40200}{2000} = 20,1$$

$$Var(\overline{X}_p) = 20,1-(4,2)^2 = 2,46$$

Δ	$P(\Delta = x)$
2/2	1/125
3/2	6/125
4/2	1/125
5/2	12/125
6/2	6/125
7/2	18/125
8/8	13/125
9/2	24/125
10/2	24/125
12/2	13/125
14/2	6/125
16/2	1/125

$$E(\Delta) = 1 \times 1125 + \frac{3}{2} \times \frac{6}{125} + \ldots + 8 \times \frac{1}{125} = \frac{1062}{250} = 4,248$$

$$E(\Delta^2) = 1^2 \times 1125 + \left(\frac{3}{2}\right)^2 \times \frac{6}{125} + \ldots + 8^2 \times \frac{1}{125} = \frac{9952}{500} = 19,904$$

$$\mathrm{Var}(\Delta) = 19,904 - (4,248)^2 = 1,858496$$

Na tabela a seguir apresentamos o resumo dos resultados obtidos.

Estimadores de $\mu = 4,200$				
Estimador	Esperança	Variância	Viés	EQM
$\overline{X}$	4,200	2,186667	0	2,186667
$\overline{X}_p$	4,200	2,460000	0	2,460000
Δ	4,248	1,858496	0,048	1,860800

Conclui-se que $\overline{X}$ e $\overline{X}_p$ são estimadores não-viesados de μ e que $\overline{X}$ é mais eficiente que $\overline{X}_p$, uma vez que $\mathrm{Var}(\overline{X}) < \mathrm{Var}(\overline{X}_p)$.

O estimador Δ é viesado, pois $E(\Delta) \neq \mu$. No entanto, o erro quadrático médio desse estimador é menor que os EQM's dos dois estimadores não-viesados.

♦♦

1.7 Alguns métodos de obtenção de estimadores

Definidas as propriedades desejáveis de um estimador, a questão que se coloca é: como conseguir estimadores? Neste curso vamos ver 2 métodos, que, no entanto, não esgotam as possibilidades. Por exemplo, na disciplina de Inferência, você estudará o método da máxima verossimilhança, que não será abordado aqui e na disciplina de Modelos Lineares você aprofundará o estudo do método dos mínimos quadrados.

O contexto geral é o seguinte: de uma população representada pela variável aleatória X extrai-se uma amostra aleatória simples $X_1, X_2, \ldots, X_n$ com o objetivo de se estimar um parâmetro $\theta = (\theta_1, \theta_2, \ldots, \theta_r)$. A distribuição de probabilidade f da variável X depende de tal parâmetro, o que representaremos por $f(x; \theta)$.

1.7.1 Método dos momentos

A ideia geral do método dos momentos é a seguinte: o estimador $\widehat{\theta}$ será obtido como solução das equações que igualam os momentos populacionais aos momentos amostrais. Por exemplo, se queremos estimar a média populacional (momento de ordem 1), o método dos momentos estabelece, como estimador, a média amostral. Vamos recordar as definições pertinentes.

DEFINIÇÃO Momento de uma variável aleatória

O momento μ_k de ordem k de uma variável aleatória X é definido como

$$\mu_k = \mathrm{E}(X^k)$$

Se X é contínua, temos que

$$\mu_k = \int x^k f(x;\theta)dx$$

e para o caso discreto

$$\mu_k = \sum_x x^k f(x;\theta)$$

DEFINIÇÃO Momento amostral

Dada uma amostra aleatória simples $X_1, X_2, \dots, X_n$ de uma população X, o momento amostral m_k de ordem k é definido como

$$m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

DEFINIÇÃO Método dos momentos

$\widehat{\theta}$ é o estimador para θ obtido pelo método dos momentos se ele for solução das equações

$$m_k = \mu_k \qquad k = 1, 2, \dots, r$$

EXEMPLO 1.8 Distribuição de Poisson

Seja $X \sim Poi(\lambda)$. Vamos obter o estimador pelo métodos dos momentos para λ. A função de probabilidade de X é

$$\mathrm{P}(X = x) = \frac{\lambda^x}{x!}e^{-\lambda}$$

e foi visto que $\mathrm{E}(X) = \lambda = \mu_1$. Igualando

$$\mu_1 = m_1 \Rightarrow \widehat{\lambda} = \overline{X}$$

EXEMPLO 1.9 Distribuição Exponencial

Seja $X \sim \exp(\beta)$. Então

$$f(x;\beta) = \frac{1}{\beta}e^{-x/\beta}$$

e $E(X) = \beta$. Como na Poisson, o estimador pelo método dos momentos será $\widehat{\beta} = \overline{X}$. Com a outra parametrização

$$f(x;\lambda) = \lambda e^{-\lambda x}$$

temos que $E(X) = \dfrac{1}{\lambda}$ e o estimador pelo método dos momentos de λ é $\widehat{\lambda} = \dfrac{1}{\overline{X}}$.

EXEMPLO 1.10 Distribuição Normal

Se $X \sim N(\mu;\sigma^2)$, temos que

$$\begin{aligned} E(X) &= \mu \Rightarrow \mu_1 = \mu \\ \text{Var}(X) &= \sigma^2 \Rightarrow E(X^2) - [E(X)]^2 = \sigma^2 \Rightarrow \mu_2 - (\mu_1)^2 = \sigma^2 \end{aligned}$$

Resulta que os estimadores pelo método dos momentos são

$$\begin{aligned} \widehat{\mu} &= \overline{X} \\ \widehat{\sigma}^2 &= m_2 - m_1^2 = \frac{1}{n}\sum X_i^2 - \overline{X}^2 = \frac{1}{n}\sum \left(X_i - \overline{X}\right)^2 \end{aligned}$$

1.7.2 Método dos mínimos quadrados

Vamos apresentar o método dos mínimos quadrados através de exemplos para fornecer a ideia geral do método, que pode ser usado para ajustar qualquer modelo a um conjunto de dados. No entanto, tais modelos só farão sentido como "representantes" de uma população (variável aleatória) se tivermos condições de avaliar suas propriedades, a qualidade do ajuste etc. Essa é a abordagem formal, que será dada em disciplinas posteriores, como a de Modelos Lineares.

EXEMPLO 1.11 Altura média de alunos

Suponha que nosso interesse esteja em estimar a altura média dos alunos do curso de Estatística da UFF. Para isso, seleciona-se uma amostra de 10 alunos e mede-se cuidadosamente a altura de cada um, em centímetros. Na Figura 1.4 apresenta-se o diagrama de pontos das alturas dos 10 alunos da amostra, juntamente com um segmento que representa a verdadeira e desconhecida altura média μ de todos os alunos de Estatística. Qualquer que seja μ, sempre haverá alunos com altura abaixo e acima de tal altura média. Dessa forma, podemos pensar na altura X_i de cada aluno i como sendo a média μ mais um "erro":

$$X_i = \mu + \epsilon_i \Longleftrightarrow X_i - \mu = \epsilon_i \tag{1.7}$$

Na Figura 1.5 ilustram-se essas diferenças como segmentos horizontais.

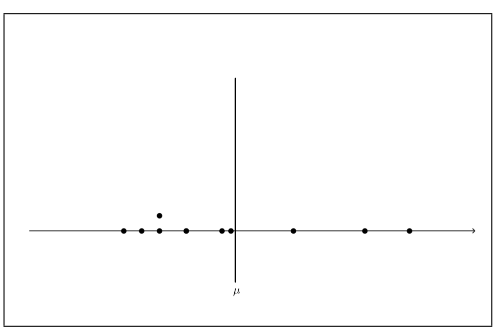


Figura 1.4 – Alturas de 10 alunos

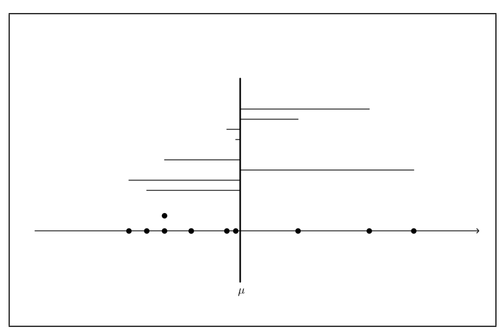


Figura 1.5 – Desvio das alturas em relação à média

O método dos mínimos quadrados estima μ por um estimador $\widehat{\mu}$ que minimize a soma dos quadrados de tais erros. O motivo para tomarmos o quadrado é o mesmo pelo qual definimos a variância como média dos desvios quadráticos: a soma desses erros será sempre nula, pois temos alturas acima e abaixo da média. Então, para encontrar o estimador de mínimos quadrados para μ, temos que minimizar a soma de quadrados dos erros, ou seja, o estimador $\widehat{\mu}$ é solução de

$$\min \sum_{i=1}^{10} (X_i - \mu)^2 \tag{1.8}$$

Seja, então, $g(\mu) = \sum_{i=1}^{10} (X_i - \mu)^2$. Para encontrar o valor de μ que minimiza essa função, temos que calcular sua derivada, igualar a zero e verificar se o ponto crítico encontrado é de máximo ou minimo através do estudo da derivada segunda.

$$g'(\mu) = -2 \sum_{i=1}^{10} (X_i - \mu) = -2 \sum_{i=1}^{10} X_i + 2 \cdot 10\mu$$

$$g''(\widehat{\mu}) = 2 \cdot 10$$

$$g'(\widehat{\mu}) = 0 \Leftrightarrow -2 \sum_{i=1}^{10} X_i + 2 \cdot 10\widehat{\mu} = 0 \Leftrightarrow \widehat{\mu} = \frac{1}{10} \sum_{i=1}^{10} X_i = \overline{X}$$

$$g''(\widehat{\mu}) = 20 > 0 \Rightarrow \text{ponto de míimo}$$

Assim como no método dos momentos, o estimador de mínimos quadrados para a média populacional é a média amostral.

♦♦

EXEMPLO 1.12 Peso e altura de alunos

Continuando com o exemplo anterior, suponha que nosso interesse agora seja estudar o peso de cada aluno em função da sua altura. Para cada um dos 10 alunos, além da altura, foi obtido também o seu peso, em quilos. Será que existe alguma relação entre peso e altura? Será que alunos mais altos tendem a pesar mais? Na Figura 1.5 apresenta-se o diagrama de dispersão, com o peso representado no eixo vertical (y) e a altura no eixo horizontal (x). Analisando esse gráfico, podemos ver que há uma tendência aproximadamente linear entre as duas variáveis e podemos, então, pensar em ajustar um modelo do tipo $Y = a + bX$, em que Y = peso e X = altura. Na Figura 1.7 exibe-se uma reta que passa pelo "meio" dos dados. Como no exemplo anterior, há uma variação em torno da reta, com pontos acima e abaixo dela. Assim, podemos expressar nosso modelo como

$$Y = a + bX + \epsilon \tag{1.9}$$

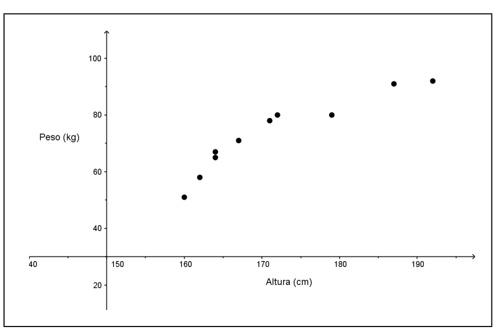


Figura 1.6 – Altura e peso de 10 alunos

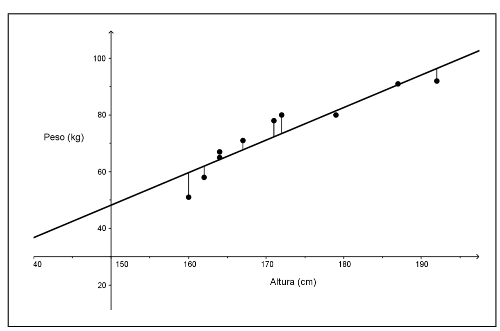


Figura 1.7 – Desvio dos pesos observados em relação ao peso estimado

Para encontrar a "melhor" reta, no sentido dos mínimos quadrados, temos que analisar a diferença entre o y observado (peso) e o y estimado pelo modelo linear, para cada x. Algumas dessas diferenças estão ilustradas como segmentos verticais na Figura 1.7. Como a reta depende de dois parâmetros (intercepto e inclinação), a função a minimizar agora é uma função de duas variáveis e teremos que determinar suas derivadas parciais:

$$
\begin{aligned}
g(a,b) &= \sum_{i=1}^{n}(Y_i - a - bX_i)^2 \\
\frac{\partial g(a,b)}{\partial a} &= -2\sum_{i=1}^{n}(Y_i - a - bX_i) = -2\left(\sum_{i=1}^{n} Y_i - na - b\sum_{i=1}^{n} X_i\right) \\
\frac{\partial g(a,b)}{\partial b} &= -2\sum_{i=1}^{n}(Y_i - a - bX_i)X_i = -2\left(\sum_{i=1}^{n} X_iY_i - a\sum_{i=1}^{n} X_i - b\sum_{i=1}^{n} X_i^2\right) \\
\frac{\partial^2 g(a,b)}{\partial a^2} &= 2n > 0 \\
\frac{\partial^2 g(a,b)}{\partial b^2} &= 2\sum_{i=1}^{n} X_i^2 > 0
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial g(a,b)}{\partial a} = 0 \Rightarrow \sum_{i=1}^{n} Y_i - n\widehat{a} - \widehat{b}\sum_{i=1}^{n} X_i = 0 \Rightarrow \widehat{a} = \frac{1}{n}\sum_{i=1}^{n} Y_i - \widehat{b}\frac{1}{n}\sum_{i=1}^{n} X_i \Rightarrow \\
&\widehat{a} = \overline{Y} - \widehat{b}\overline{X}
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial g(a,b)}{\partial b} = 0 \Rightarrow \sum_{i=1}^{n} X_iY_i - \widehat{a}\sum_{i=1}^{n} X_i - \widehat{b}\sum_{i=1}^{n} X_i^2 = 0 \Rightarrow \sum_{i=1}^{n} X_iY_i - n\widehat{a}\overline{X} - \widehat{b}\sum_{i=1}^{n} X_i^2 = 0 \Rightarrow \\
&\sum_{i=1}^{n} X_iY_i - n\overline{X}(\overline{Y} - \widehat{b}\overline{X}) - \widehat{b}\sum_{i=1}^{n} X_i^2 = 0 \Rightarrow \left(\sum_{i=1}^{n} X_iY_i - n\overline{X}\,\overline{Y}\right) - \widehat{b}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right) = 0 \Rightarrow \\
&\widehat{b} = \frac{\sum_{i=1}^{n} X_iY_i - n\overline{XY}}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X}^2)}
\end{aligned}
$$

Os estimadores de mínimos quadrados para o intercepto e a inclinação da reta são

$$
\widehat{a} = \overline{Y} - \widehat{b}\overline{X} \tag{1.10}
$$

$$
\widehat{b} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X}^2)} \tag{1.11}
$$

Para o nosso exemplo, temos o seguinte:

	X_i	Y_i	X_iY_i	X_i^2	Y_i^2
	160	51	8160	25600	2601
	162	58	9396	26244	3364
	164	65	10660	26896	4225
	164	67	10988	26896	4489
	167	71	11857	27889	5041
	171	78	13338	29241	6084
	172	80	13760	29584	6400
	179	80	14320	32041	6400
	187	91	17017	34969	8281
	192	92	17664	36864	8464
Soma	1718	733	127160	296224	55349

e isso resulta em

$$\overline{X} = 171,8 \qquad \overline{Y} = 73,3$$

$$\widehat{b} = \frac{127160 - 10 \cdot 171,8 \cdot 73,3}{296224 - 10 \cdot 171,8^2} = \frac{1230,6}{1071,6} = 1,1484$$

$$\widehat{a} = 73,3 - 1,1484 \cdot 171,8 = -123,9951$$

A reta de mínimos quadrados, exibida na Figura 1.7, é dada por

$$\widehat{\text{Peso}} = -123,9951 + 1,1484 \cdot \text{Altura}$$

♦♦

Em geral, os estimadores de mínimos quadrados são obtidos minimizando-se a soma dos quadrados dos erros $Y - \widehat{Y}$.

1.8 Exercícios propostos

1. Em cada uma das afirmativas seguintes, identifique o número em negrito como o valor de um parâmetro populacional ou de uma estatística amostral. Justifique sua resposta.

   (a) Um pesquisador realizou um estudo para investigar a prevalência do mal de Alzheimer em pessoas com idade acima de 85 anos. Os resultados indicaram que **47,2%** dos pacientes estudados mostraram sintomas consistentes com a doença.

   (b) A equipe de uma revista de suporte ao consumidor testou uma amostra aleatória de 19 chás verdes em relação ao sabor e aos benefícios à saúde. O preço médio por xícara de chá foi relatado como **2,80** reais.

   (c) Durante um verão recente, o mês de janeiro foi particularmente quente no Rio de Janeiro. A companhia estadual de energia relatou que o número médio de quilowatts usados por cada consumidor durante esse mês de janeiro foi **850**.

   (d) Um fabricante de computadores afirma que a vida média das baterias de laptops é **6,7** horas.

2. Considere, novamente, a população descrita no Exemplo 1.4. Obtenha a distribuição amostral da mediana amostral e verifique se ela é um estimador não-viesado para a mediana populacional.

3. Seja $X_1, X_2, \cdots, X_n$ uma amostra aleatória simples de uma população $X \sim Bern(p)$. Calcule o erro quadrático médio do seguinte estimador de p:

$$\widetilde{P} = \frac{1}{n+1}\sum_{i=1}^{n} X_i$$

4. (Larson (1982)) Seja $X_1, X_2, \cdots, X_6$ uma amostra aleatória de uma população $X \sim N(\mu; \sigma^2)$. Se

$$\widetilde{\sigma}^2 = C[(X_1 - X_2)^2 + (X_3 - X_4)^2 + (X_5 - X_6)^2]$$

determine o valor de C para que $\widetilde{\sigma}^2$ seja um estimador não-viesado para σ^2.

5. (Kokoska (2011)) Previsões de ondas de águas profundas (> 300m) são importantes para grandes navios de carga. Um método de predição sugere que a velocidade do vento (x, em nós) esteja relacionada linearmente com a altura da onda (y, em pés). Obteve-se uma amostra aleatória de boias, e mediu-se a velocidade do vento e a altura da onda em cada uma. Encontre a reta de mínimos quadrados que explicita essa relação com base nos dados a seguir.

Vento	9	11	10	10	11	9	9	6	9	5	8	9
Onda	2,9	1,4	1,7	0,9	1,2	1,0	1,5	0,7	1,9	0,1	2,0	2,6

Vento	12	9	12	9	12	8	7	13	9	8	6	8
Onda	3,0	1,7	2,1	1,5	3,1	2,7	0,4	2,5	1,7	0,6	0,7	1,4

Capítulo 2

Algumas distribuições amostrais

Vimos, no capítulo anterior, que uma população estatística é representada por uma variável aleatória X e, assim, a média (esperança) μ e a variância σ^2 de tal variável são parâmetros de interesse. Vimos, também, através de exemplos, que a média amostral $\overline{X}$ e S^2 são estimadores não-viesados de μ e σ^2, respectivamente. Lembre-se que tais propriedades são definidas a partir das distribuições amostrais de $\overline{X}$ e S^2, que são as distribuições de probabilidade ao longo de todas as possíveis amostras aleatórias simples de tamanho n. Neste capítulo apresentaremos resultados formais sobre as distribuições amostrais de $\overline{X}$ e S^2.

2.1 Distribuição amostral da média amostral $\overline{X}$

2.1.1 Média e variância de $\overline{X}$

TEOREMA 2.1

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória simples de tamanho n de uma população representada pela variável aleatória X com média μ e variância σ^2. Então,

$$E(\overline{X}) = \mu \tag{2.1}$$

$$Var(\overline{X}) = \frac{\sigma^2}{n} \tag{2.2}$$

Demonstração

Por definição de amostra aleatória simples, as X_i são independentes e todas têm a mesma distribuição de X; logo, $\mathrm{E}(X_i) = \mu$ e $\mathrm{Var}(X_i) = \sigma^2$. Da independência resulta que $\mathrm{Cov}(X_i, X_j) = 0 \; \forall i \neq j$.

$$\mathrm{E}(\overline{X}) = \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}n\mu = \mu$$

ou seja, $\overline{X}$ é um estimador não-viesado da média populacional μ, qualquer que seja a população.

$$\begin{aligned}\text{Var}(\overline{X}) &= \text{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\text{Var}\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\left[\sum_{i=1}^{n}\text{Var}(X_i) + \sum_{i\neq j}\text{Cov}(X_i, X_j)\right] \\ &= \frac{1}{n^2}\left(\sum_{i=1}^{n}\sigma^2 + 0\right) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}\end{aligned}$$

Note que a variância tende a zero quando $n \to \infty$.

■

É importante notar que esse resultado se refere a qualquer população X. O que ele estabelece é que as médias amostrais das diferentes amostras aleatórias simples de tamanho n tendem a "acertar o alvo"da média populacional μ, conforme ilustrado na Figura 1.3, partes (a) e (b). Além disso, à medida que o tamanho amostral n aumenta, a dispersão em torno do alvo, medida por $\text{Var}(\overline{X})$, vai diminuindo e tende a zero quando $n \to \infty$.

Note que a variância de $\overline{X}$ depende diretamente da variância populacional e inversamente do tamanho da amostra. A título de ilustração, vamos considerar duas populações: $X_1 \sim \exp(0,5)$ e $X_2 \sim \exp(1,0)$, cujas funções densidade são exibidas na Figura 2.1. Resulta que

$$\begin{aligned}\text{E}(X_1) = 2,0 \quad \text{Var}(X_1) = 4,0 \\ \text{E}(X_2) = 1,0 \quad \text{Var}(X_2) = 1,0\end{aligned}$$

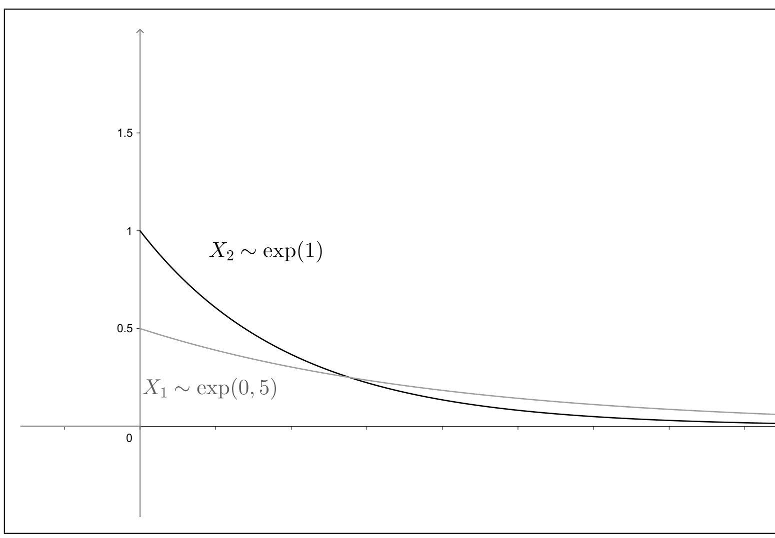
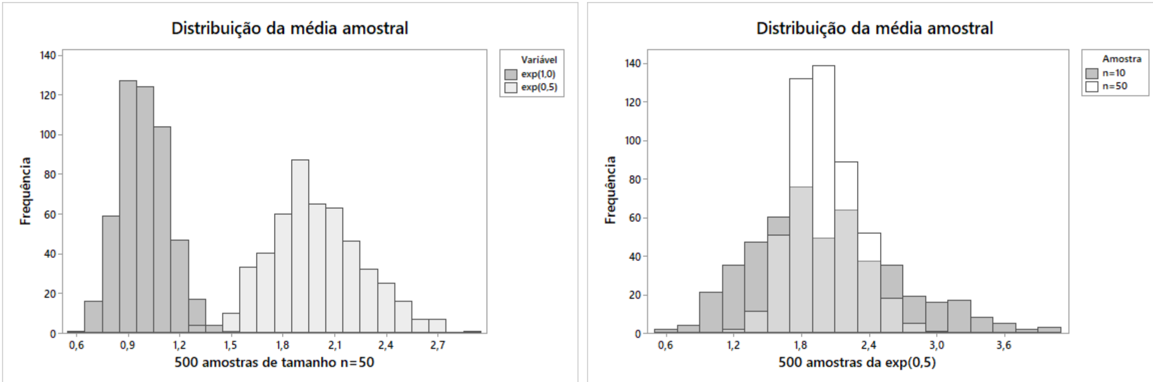


Figura 2.1 – Duas distribuições exponenciais

Para ilustrar o efeito da variância populacional, foram calculadas 500 médias amostrais obtidas de amostras de tamanho 50 de cada uma das populações. Os histogramas que aproximam a distribuição amostral de $\overline{X}$ estão na Figura 2.2a, que confirma o resultado: o histograma em cinza escuro, correspondente à exponencial de parâmetro 1, apresenta menor variabilidade. Para ilustrar o efeito do tamanho da amostra, foram extraídas 500 amostras de tamanho 10 e 500 amostras de tamanho 50 da população $X_1 \sim \exp(0,5)$. Os histogramas das distribuições amostrais das médias amostrais desses dois conjuntos de amostras estão na Figura 2.2b, onde podemos ver que, para amostras de tamanho 50, a dispersão da distribuição amostral é menor. (Note que a parte cinza claro corresponde à sobreposição do cinza escuro e do branco)



(a) Efeito de σ^2 **(b)** Efeito de n

Figura 2.2 – Variância da média amostral

O desvio padrão da distribuição amostral de qualquer estatística é usualmente chamado de *erro padrão*. Então, o erro padrão da média amostral é $\text{EP}(\overline{X}) = \dfrac{\sigma}{\sqrt{n}}$.

2.1.2 Populações normais

Na prática estatística, várias populações podem ser descritas, pelo menos aproximadamente, por uma distribuição normal. Obviamente, o teorema anterior continua valendo no caso de uma população normal, mas temos uma característica a mais da distribuição amostral da média: ela é também normal.

TEOREMA 2.2

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória simples de tamanho n de uma ***população normal****, isto é, uma população representada por uma variável aleatória normal $X \sim N(\mu; \sigma^2)$. Então, a distribuição amostral da média amostral $\overline{X}$ é normal com média μ e variância σ^2/n, ou seja,*

$$X \sim N\left(\mu; \sigma^2\right) \Rightarrow \overline{X} \sim N\left(\mu; \frac{\sigma^2}{n}\right) \tag{2.3}$$

Demonstração

Para completar a prova do teorema, temos que provar que a distribuição de $\overline{X}$ é normal. Para isso, vamos usar a função geradora de momentos da $N(\mu, \sigma^2)$, que é dada por

$$m_X(t) = \exp\left(\mu t + \sigma^2 \frac{t^2}{2}\right)$$

Por definição, temos que

$$\begin{aligned} m_{\overline{X}}(t) &= \text{E}\left(e^{t\overline{X}}\right) = \text{E}\left\{\exp\left[t\frac{1}{n}(X_1 + X_2 + \ldots + X_n)\right]\right\} \\ &= \text{E}\left[\exp\left(\frac{t}{n}X_1\right) \cdot \exp\left(\frac{t}{n}X_2\right) \cdots \exp\left(\frac{t}{n}X_n\right)\right] \end{aligned}$$

Como as X_i's são independentes e identicamente distribuídas, resulta

$$\begin{aligned} m_{\overline{X}}(t) &= \mathrm{E}\left[\exp\left(\frac{t}{n}X_1\right)\right]\cdot \mathrm{E}\left[\exp\left(\frac{t}{n}X_2\right)\right]\cdots \mathrm{E}\left[\exp\left(\frac{t}{n}X_n\right)\right] \\ &= m_{X_1}\left(\frac{t}{n}\right)\cdot m_{X_2}\left(\frac{t}{n}\right)\cdots m_{X_n}\left(\frac{t}{n}\right) = \left[\exp\left(\mu\frac{t}{n}+\frac{\sigma^2}{2}\frac{t^2}{n^2}\right)\right]^n \\ &= \exp\left[n\left(\mu\frac{t}{n}+\frac{\sigma^2}{2}\frac{t^2}{n^2}\right)\right] = \exp\left(\mu t+\frac{\sigma^2}{n}\frac{t^2}{2}\right) \end{aligned}$$

Mas essa é a função geradora de momentos de uma normal com média μ e variância $\frac{\sigma^2}{n}$, o que completa a prova do teorema.

■

Na Figura 2.3 ilustra-se o comportamento da distribuição amostral de $\overline{X}$ com base em amostras de tamanho $n = 4$ retiradas de uma população $X \sim N(2; 3^2)$. A título de comparação, apresenta-se também a distribuição populacional. Podemos ver que ela é mais dispersa que a distribuição amostral de $\overline{X}$, mas ambas estão centradas no verdadeiro valor populacional $\mu = 2$

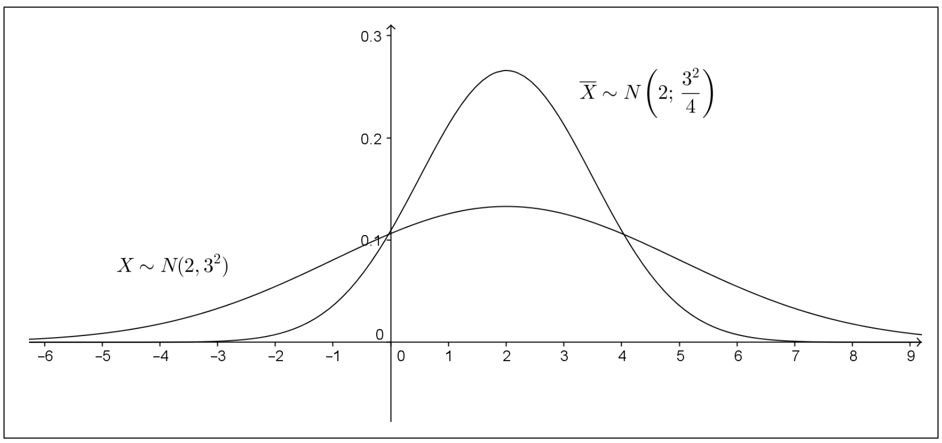


Figura 2.3 – Distribuição amostral de $\overline{X}$ com base em aas de tamanho $n = 4$ de uma população $N(2; 9)$.

EXEMPLO 2.1 Carga de elevador

A capacidade máxima de um elevador é de 500 kg. Se a distribuição dos pesos dos usuários pode ser aproximada por uma $N(70; 100)$, qual é a probabilidade de que sete pessoas ultrapassem este limite? E seis pessoas?

Solução

Podemos considerar os sete passageiros como uma amostra aleatória simples da população de todos os usuários, representada pela v.a. $X \sim N(70; 100)$. Seja, então, $X_1, \ldots, X_7$ uma aas de tamanho $n = 7$. Se o peso máximo é 500kg, para que sete pessoas ultrapassem o limite de segurança temos de ter

$$\sum_{i=1}^{7} X_i > 500 \Rightarrow \frac{1}{7}\sum_{i=1}^{7} X_i > \frac{500}{7} \Rightarrow \overline{X} > 71,4286$$

Mas, por (2.2), sabemos que

$$\overline{X} \sim N\left(70; \frac{100}{7}\right)$$

Logo,

$$
\begin{aligned}
P(\overline{X} > 71,4286) &= P\left(\frac{\overline{X}-70}{\sqrt{\frac{100}{7}}} > \frac{71,4286-70}{\sqrt{\frac{100}{7}}}\right) \\
&= P(Z > 0,38) = 0,5 - P(0 \leq Z \leq 0,38) \\
&= 0,5 - 0,1480 = 0,3520
\end{aligned}
$$

Com seis pessoas teríamos que ter

$$
\begin{aligned}
P\left(\overline{X} > \frac{500}{6}\right) &= P\left(Z > \frac{83,333-70}{\sqrt{\frac{100}{6}}}\right) \\
&= P(Z > 3,26) = 0,5 - \text{tab}(3,26) \\
&= 0,5 - 0,4994 = 0,0006
\end{aligned}
$$

Podemos ver que existe uma probabilidade alta (0,352 ou 35,2% de chance) de sete pessoas ultrapassarem o limite de segurança. Já com seis pessoas, essa probabilidade é bastante pequena (0,0006 ou 0,06%). Assim, o número máximo de pessoas no elevador deve ser estabelecido como seis ou menos. ♦♦

EXEMPLO 2.2

Considere uma população representada por $X \sim N(5, 2^2)$.

(a) Calcule $P(3 < X < 7)$.

(b) Se $\overline{X}$ é a média de uma amostra aleatória simples de 16 elementos retirados dessa população, calcule $P(3 < \overline{X} < 7)$.

(c) Construa, em um único sistema de coordenadas, os gráficos das distribuições de X e $\overline{X}$.

(d) Que tamanho deveria ter a amostra para que $P(3 < \overline{X} < 7) = 0,95$?

Solução

(a)

$$
\begin{aligned}
P(3 < X < 7) &= P\left(\frac{3-5}{2} < Z < \frac{7-5}{2}\right) = P(-1 < Z < 1) \\
&= 2 \times P(0 < Z < 1) = 2 \times 0,3413 = 0,6826
\end{aligned}
$$

(b) Com $n = 16$, resulta que $\overline{X} \sim N\left(5; \frac{4}{16}\right)$

$$
\begin{aligned}
P(3 < \overline{X} < 7) &= P\left(\frac{3-5}{\sqrt{\frac{4}{16}}} < Z < \frac{7-5}{\sqrt{\frac{4}{16}}}\right) \\
&= P(-4 < Z < 4) = 2 \times P(0 < Z < 4) \approx 1,00
\end{aligned}
$$

(c) Veja a Figura 2.4. Como visto, a distribuição amostral com $n = 16$ é menos dispersa que a distribuição populacional e, então, podemos ver que, entre 3 e 7, temos concentrada praticamente toda a distribuição de $\overline{X}$, ou seja, praticamente todos os possíveis valores de $\overline{X}$ estão entre 3 e 7.

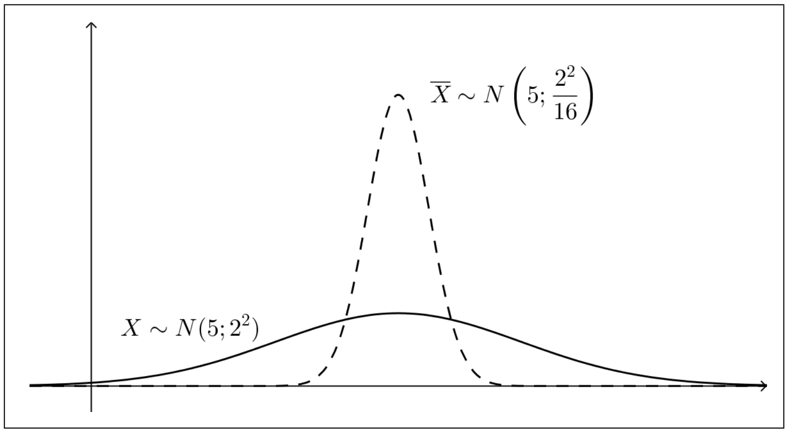


Figura 2.4 – Distribuição amostral de $\overline{X}$ com base em aas de tamanho $n = 16$ de uma população $N(5; 4)$.

(d)

$$P(3 < \overline{X} < 7) = 0,95 \Leftrightarrow P\left(\frac{3-5}{\sqrt{\frac{4}{n}}} < Z < \frac{7-5}{\sqrt{\frac{4}{n}}}\right) = 0,95 \Leftrightarrow$$

$$P(-\sqrt{n} < Z < \sqrt{n}) = 0,95 \Leftrightarrow 2 \times P(0 < Z < \sqrt{n}) = 0,95 \Leftrightarrow P(0 < Z < \sqrt{n}) = 0,475 \Leftrightarrow$$
$$\sqrt{n} = 1,96 \Leftrightarrow n \approx 4$$

A título de ilustração, apresentam-se na Figura 2.5 as distribuições amostrais de $\overline{X}$ para $n = 16$ e $n = 4$, juntamente com a distribuição populacional. Veja com cuidado o efeito do tamanho da amostra sobre a dispersão da distribuição amostral!

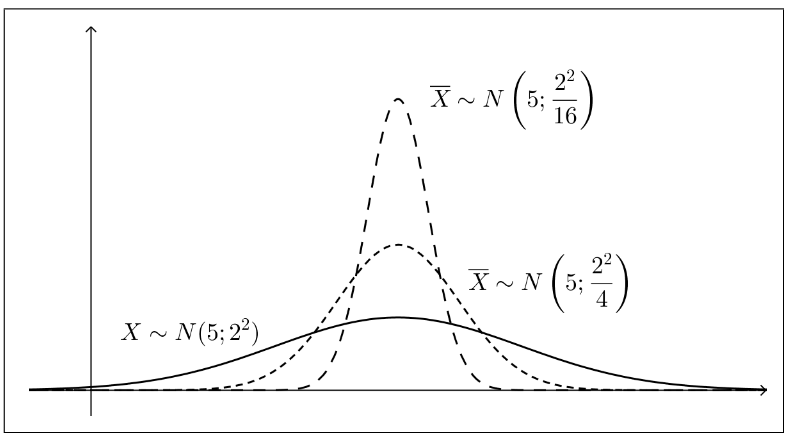


Figura 2.5 – Distribuição amostral de $\overline{X}$ com base em aas de tamanho $n = 16$ de uma população $N(5; 4)$.

♦♦

EXEMPLO 2.3 Regulagem de máquinas

A máquina de empacotar um determinado produto o faz segundo uma distribuição normal, com média μ e desvio padrão 20g.

(a) Em quanto deve ser regulado o peso médio μ para que apenas 10% dos pacotes tenham menos do que 1000g?

(b) Com a máquina assim regulada, qual é a probabilidade de que o peso total de quatro pacotes escolhidos ao acaso seja inferior a 4 kg?

Solução

(a) Seja X a variável aleatória que representa o peso (em gramas) dos pacotes. Sabemos, então, que $X \sim N(\mu; 20^2)$. Queremos que

$$\mathrm{P}(X < 1000) = 0,10 \Leftrightarrow \mathrm{P}\left(\frac{X-\mu}{20} < \frac{1000-\mu}{20}\right) = 0,10 \Leftrightarrow \mathrm{P}\left(Z < \frac{1000-\mu}{20}\right) = 0,10$$

Então, na densidade normal padrão, à esquerda da abscissa $\frac{1000-\mu}{20}$ temos que ter uma probabilidade de 0,10. Logo, essa abscissa tem que ser negativa. Usando a simetria da densidade normal, temos as seguintes equivalências:

$$\mathrm{P}\left(Z < \frac{1000-\mu}{20}\right) = 0,10 \Leftrightarrow \mathrm{P}\left(Z > -\frac{1000-\mu}{20}\right) = 0,10 \Leftrightarrow$$
$$\mathrm{P}\left(0 \leq Z \leq \frac{\mu-1000}{20}\right) = 0,40 \Leftrightarrow \frac{\mu-1000}{20} = 1,28 \Leftrightarrow \mu = 1025,6 \text{ g}$$

Veja a Figura 2.6 onde são ilustradas essas equivalências.

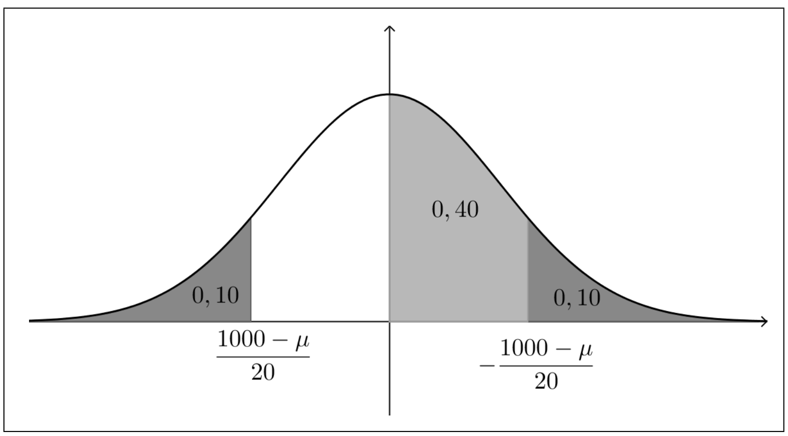


Figura 2.6 – Solução do Exemplo 2.3

(b) Sejam X_1, X_2, X_3, X_4 os pesos dos 4 pacotes da amostra. Queremos que $\sum_{i=1}^{4} X_i < 4000$g. Isso é equivalente a $\overline{X} < 1000$. Logo,

$$\begin{aligned}
\mathrm{P}(\overline{X} < 1000) &= \mathrm{P}\left(\frac{\overline{X}-1025,6}{\sqrt{\frac{400}{4}}} < \frac{1000-1025,6}{\sqrt{\frac{400}{4}}}\right) = \mathrm{P}(Z < -2,56) = \mathrm{P}(Z > 2,56) \\
&= 0,5 - \mathrm{P}(0 \leq Z \leq 2,56) = 0,5 - 0,4948 = 0,0052
\end{aligned}$$

Com a máquina regulada para 1025,6 g, há uma probabilidade de 0,0052 de que uma amostra de 4 pacotes apresente peso médio inferior a 1000g. Note que com um pacote apenas, essa probabilidade é de 10%. Por isso, as inspeções de controle de qualidade são sempre feitas com base em amostras de tamanho $n > 1$.

♦♦

EXEMPLO 2.4 Regulagem de máquinas – continuação

Volte ao Exemplo 2.3. Depois de regulada a máquina, prepara-se uma carta de controle de qualidade. Uma amostra de 4 pacotes será sorteada a cada hora. Se a média da amostra for inferior a 994g ou superior a 1040g, a produção deve ser interrompida para ajuste da máquina, ou seja, o peso médio deve ser ajustado.

(a) Qual é a probabilidade de uma parada desnecessária?

(b) Se a máquina se desregulou para $\mu = 1000$g, qual é a probabilidade de se continuar a produção fora dos padrões desejados?

Solução

Com a máquina regulada, temos que $X \sim N(1025,6; 400)$

(a) Parada desnecessária: amostra indica que o processo está fora de controle ($\overline{X} < 994$ ou $\overline{X} > 1040$), quando, na verdade, o processo está ajustado ($\mu = 1025,6$). Neste caso, podemos usar a notação de probabilidade condicional para auxiliar na solução do exercício. Queremos calcular

$$
\begin{aligned}
& \mathrm{P}\left[\left(\overline{X} < 994\right) \cup \left(\overline{X} > 1040\right) \mid \overline{X} \sim N\left(1025,6; \tfrac{400}{4}\right)\right] \\
= & \mathrm{P}\left[\overline{X} < 994 \mid \overline{X} \sim N(1025,6; 100)\right] + \mathrm{P}\left[\overline{X} > 1040 \mid \overline{X} \sim N(1025,6; 100)\right] \\
= & \mathrm{P}\left(Z < \frac{994 - 1025,6}{10}\right) + \mathrm{P}\left(Z > \frac{1040 - 1025,6}{10}\right) \\
= & \mathrm{P}(Z < -3,16) + \mathrm{P}(Z > 1,44) \\
= & \mathrm{P}(Z > 3,16) + \mathrm{P}(Z > 1,44) \\
= & [0,5 - \mathrm{P}(0 < Z < 3,16)] + [0,5 - \mathrm{P}(0 < Z < 1,44)] \\
= & 1,0 - 0,4992 - 0,4251 \\
= & 0,0757
\end{aligned}
$$

(b) Agora queremos

$$
\begin{aligned}
& \mathrm{P}\left[994 \leq \overline{X} \leq 1040 \mid \overline{X} \sim N(1000; 100)\right] = \mathrm{P}\left(\frac{994 - 1000}{10} \leq Z \leq \frac{1040 - 1000}{10}\right) \\
= & \mathrm{P}(-0,6 \leq Z \leq 4) = \mathrm{P}(-0,6 \leq Z < 0) + \mathrm{P}(0 \leq Z \leq 4) \\
= & \mathrm{P}(0 \leq Z \leq 0,6) + \mathrm{P}(0 \leq Z \leq 4) = 0,2257 + 0,5 = 0,7257
\end{aligned}
$$

Note que a probabilidade de uma parada desnecessária é pequena, à custa de uma alta probabilidade de se operar fora de controle, muito embora a diferença de 25,6 g represente um grande desvio em relação ao ponto de controle.

2.1.3 Teorema Limite Central

Os resultados vistos anteriormente são válidos para populações normais, isto é, se uma população é normal com média μ e variância σ^2, então a distribuição amostral de $\overline{X}$ é também normal com média μ e variância σ^2/n, onde n é o tamanho da amostra. O Teorema Limite Central nos fornece um resultado análogo para qualquer distribuição populacional, desde que o tamanho da amostra seja suficientemente grande. A seguir apresentamos uma versão informal de tal teorema, que será estudado com mais detalhes e rigor probabilístico na disciplina de Inferência.

TEOREMA 2.3 ***Teorema Limite Central***

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória simples de uma população X tal que $E(X) = \mu$ e $Var(X) = \sigma^2 < \infty$. Então

$$\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n \to \infty]{} N(0,1) \tag{2.4}$$

▼

A interpretação prática do Teorema Limite Central é a seguinte: para amostras "grandes" de qualquer população, podemos aproximar a distribuição amostral de $\overline{X}$ por uma distribuição normal com a mesma média populacional e variância igual à variância populacional dividida pelo tamanho da amostra. Quão grande deve ser a amostra para se obter uma boa aproximação depende das características da distribuição populacional. Se a distribuição populacional não se afastar muito de uma distribuição normal, a aproximação será boa, mesmo para tamanhos pequenos de amostra. Na Figura 2.7 ilustra-se esse teorema para uma distribuição exponencial com parâmetro $\lambda = 1$, cujo gráfico pode ser visto na Figura 2.1. Os histogramas representam a distribuição amostral de $\overline{X}$ ao longo de 5.000 amostras de tamanhos 10, 50, 100 e 5000. Assim, podemos ver que, embora a população seja completamente diferente da normal, a distribuição amostral de $\overline{X}$ vai se tornando cada vez mais próxima da normal à medida que n aumenta.

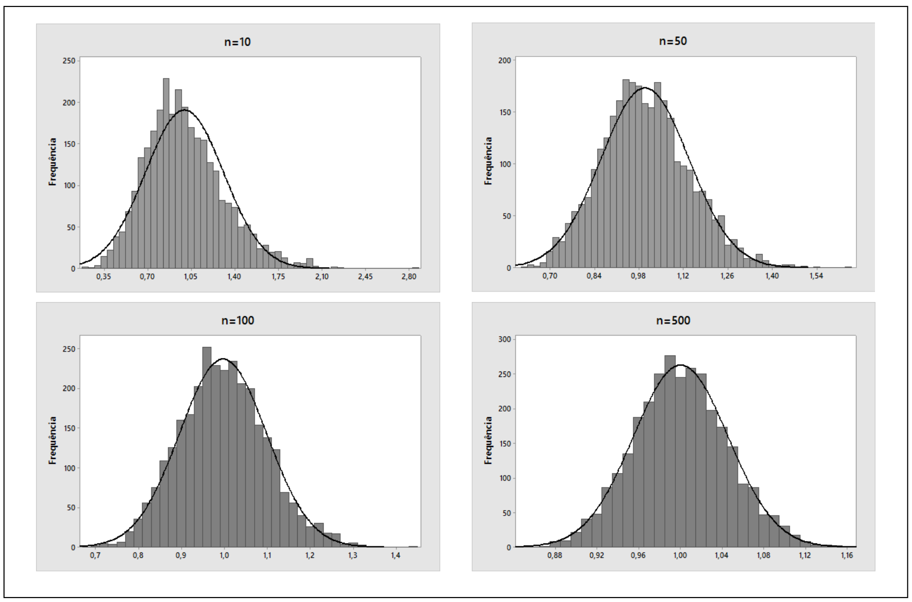


Figura 2.7 – Ilustração do Teorema Limite Central para uma população $X \sim exp(1)$

Em termos práticos, esse teorema é de extrema importância e por isso é chamado teorema central. Em geral, amostras de tamanho $n > 30$ já fornecem uma aproximação razoável.

EXEMPLO 2.5 Garrafas de refrigerante

A divisão de inspeção do Departamento de Pesos e Medidas de uma determinada cidade está interessada em calcular a real quantidade de refrigerante que é colocada em garrafas de dois litros, no setor de engarrafamento de uma grande empresa de refrigerantes. O gerente do setor de engarrafamento informou à divisão de inspeção que o desvio padrão para garrafas de dois litros é de 0,05 litro. Uma amostra aleatória de 100 garrafas de dois litros, obtida deste setor de engarrafamento, indica uma média de 1,985 litro. Qual é a probabilidade de se obter uma média amostral de 1,985 ou menos, caso a afirmativa do gerente esteja certa? O que se pode concluir?

Solução

Note que são dados apenas a média e o desvio-padrão da população, sem se especifica sua distribuição, ou seja, as afirmativas do gerente são: $\mu = 2$ e $\sigma = 0,05$. Como $n = 100$, podemos usar o Teorema Limite Central. Logo, se a afirmativa do gerente for correta, $\overline{X} \approx N\left(2; \frac{0,05^2}{100}\right)$.

$$\begin{aligned}
\mathrm{P}(\overline{X} \le 1,985) &= \mathrm{P}\left(Z \le \frac{1,985 - 2}{\frac{0,05}{10}}\right) \\
&= \mathrm{P}(Z \le -3,0) = \mathrm{P}(Z \ge 3,0) \\
&= 0,5 - \mathrm{P}(0 \le Z \le 3,0) = 0,5 - 0,49865 = 0,00135
\end{aligned}$$

A probabilidade de se obter esse valor nas condições dadas pelo gerente é muito pequena, o que pode nos fazer suspeitar da veracidade das afirmativas. É provável que ou a média não seja 2 (e, sim, menor que 2), ou o desvio padrão não seja 0,05 (e, sim, maior que 0,05).

♦♦

2.1.4 Aproximação normal da binomial

O Teorema Limite Central nos diz que, se X é uma população com média μ e variância σ^2, então a distribuição amostral da média de uma amostra aleatória simples de tamanho n se aproxima de uma distribuição normal com média μ e variância $\frac{\sigma^2}{n}$ quando $n \to \infty$.

Usando as propriedades da média e da variância, podemos estabelecer esse teorema em termos de $S_n = \sum_{i=1}^{n} X_i$, em vez de $\overline{X}$. Como $S_n = n\overline{X}$, então $\mathrm{E}(S_n) = n\,\mathrm{E}(\overline{X}) = n\mu$ e $\mathrm{Var}(S_n) = n^2\,\mathrm{Var}(\overline{X}) = n^2\frac{\sigma^2}{n} = n\sigma^2$; isso nos dá o seguinte resultado:

TEOREMA 2.4 ***Teorema Limite Central***

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória simples de uma população X tal que $E(X) = \mu$ e $Var(X) = \sigma^2$. Então

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \xrightarrow[n\to\infty]{} N(0,1)$$

▼

A variável aleatória binomial foi definida como "número de sucessos em n repetições independentes de um experimento de Bernoulli com parâmetro p". Então, uma variável binomial é a soma de n variáveis independentes $Bern(p)$. Pelo teorema acima e usando o fato de que, se $X \sim Bern(p)$, então $E(X) = p$ e $Var(X) = p(1-p)$, podemos dizer que a distribuição binomial com parâmetros n e p se aproxima de uma normal com média np e variância $np(1-p)$ à medida que n aumenta.

Alguns cuidados devem ser tomados na aproximação da binomial pela normal. Um fato importante a observar é que a distribuição binomial é discreta, enquanto a variável normal é contínua. Por exemplo, para uma variável binomial faz sentido calcular $\mathrm{P}(X = k)$, mas na normal, essa probabilidade é nula, qualquer que seja k. Assim, para usar a aproximação normal para calcular probabilidades da binomial, usamos a *correção de continuidade* que consiste em substituir cada valor k possível na binomial pelo intervalo de valores $(k - 0,5; k + 0,5)$ na normal aproximadora. Note que o centro desse intervalo é o valor k da binomial. Por exemplo, para calcular $\mathrm{P}(X = 5)$, usando a variável normal aproximadora Y, calculamos $\mathrm{P}(4,5 < Y < 5,5)$.

Vamos ver alguns outros exemplos, considerando a variável binomial $X \sim bin(50; 0,2)$.

- $$\begin{aligned} P(X > 40) &= P\left[(X = 41) \cup (X = 42) \cup \cdots \cup (X = 50)\right] \\ &\approx P\left[(40,5 \leq Y \leq 41,5) \cup (41,5 \leq Y \leq 42,5) \cup \cdots \cup (49,5 \leq Y \leq 50,5)\right] \\ &\approx P(Y \geq 40,5) \end{aligned}$$

  Como 50 é o maior valor possível da binomial, na normal aproximadora também tomamos o maior valor possível , ou seja, $+\infty$.

- $$\begin{aligned} P(X \geq 40) &= P\left[(X = 40) \cup (X = 41) \cup (X = 42) \cup \cdots \cup (X = 50)\right] \\ &\approx P\left[(39,5 \leq Y \leq 40,5) \cup (40,5 \leq Y \leq 41,5) \cup \cdots \cup (49,5 \leq Y \leq 50,5)\right] \\ &\approx P(Y \geq 39,5) \end{aligned}$$

- $$\begin{aligned} P(X < 16) &= P\left[(X = 15) \cup (X = 14) \cup (X = 13) \cup \cdots \cup (X = 0)\right] \\ &\approx P\left[(14,5 \leq Y \leq 15,5) \cup (13,5 \leq Y \leq 14,5) \cup \cdots \cup (-0,5 \leq Y \leq 0,5)\right] \\ &\approx P(Y \leq 15,5) \end{aligned}$$

- $$\begin{aligned} P(X \leq 16) &= P\left[(X = 16) \cup (X = 15) \cup (X = 14) \cup \cdots \cup (X = 0)\right] \\ &\approx P\left[(15,5 \leq Y \leq 16,5) \cup (14,5 \leq Y \leq 15,5) \cup \cdots \cup (-0,5 \leq Y \leq 0,5)\right] \\ &\approx P(Y \leq 16,5) \end{aligned}$$

Na tabela a seguir, resumimos o procedimento de aproximação da binomial pela normal, que está ilustrado na Figura 2.8; em ambos, temos $X \sim bin(n; p)$ e $Y \sim N[np; np(1-p)]$.

Binomial	Aproximação Normal
$P(X = k)$	$P(k - 0,5 \leq Y \leq k + 0,5)$
$P(X \leq k)$	$P(Y \leq k + 0,5)$
$P(X < k) = P(X \leq k - 1)$	$P(Y \leq (k-1) + 0,5) = P(Y \leq k - 0,5)$
$P(X \geq k)$	$P(Y \geq k - 0,5)$
$P(X > k) = P(X \geq k + 1))$	$P(Y \geq (k+1) - 0,5) = P(Y \geq k + 0,5)$

A aproximação dada pelo Teorema Limite Central é melhor para valores grandes de n. Existe a seguinte regra empírica para nos ajudar a decidir o que é "grande":

! A distribuição binomial com parâmetros n e p pode ser aproximada por uma distribuição normal com média $\mu = np$ e variância $\sigma^2 = np(1-p)$ se são satisfeitas as seguintes condições:

1. $n \geq 30$
2. $np \geq 5$
3. $n(1-p) \geq 5$

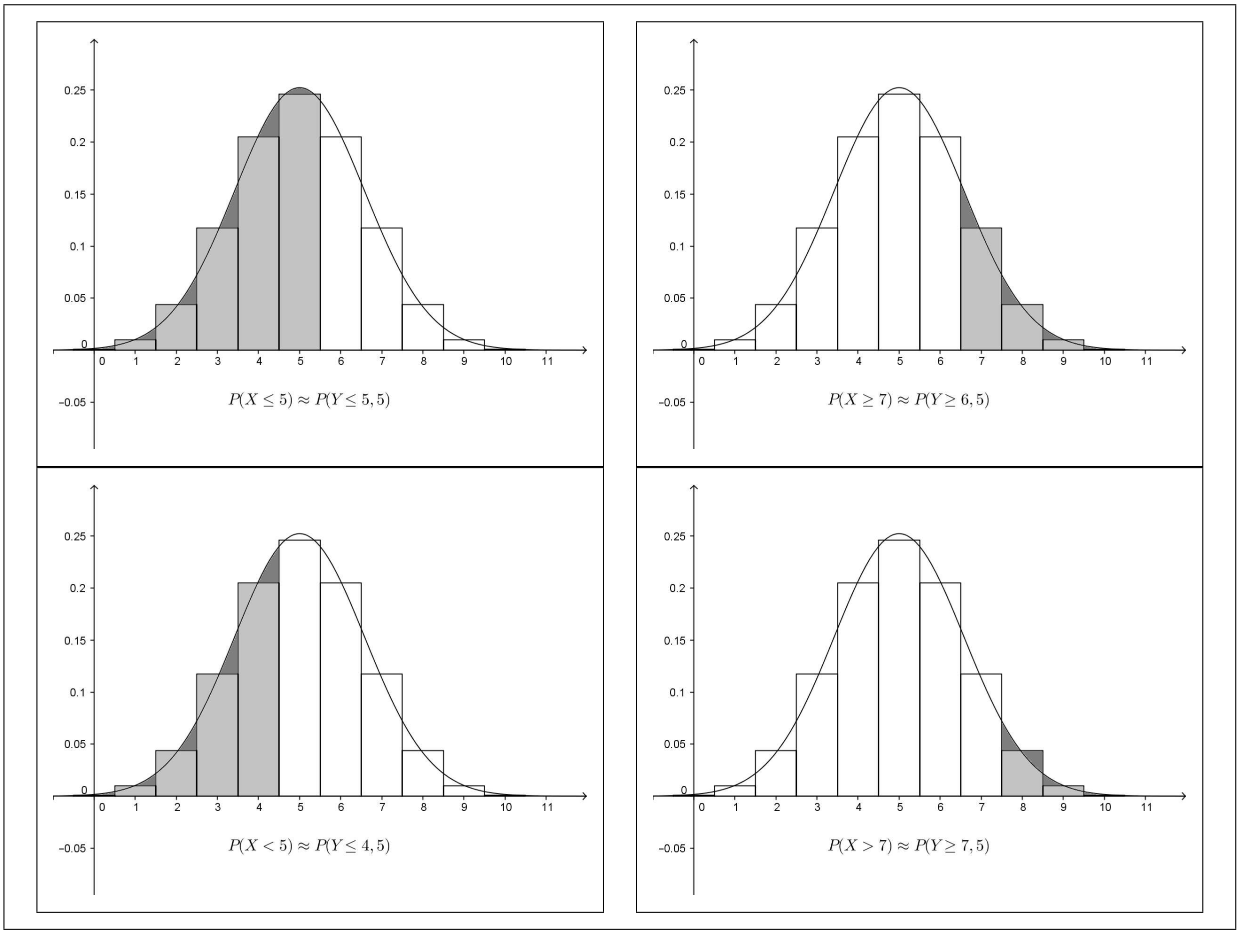


Figura 2.8 – Aproximação da distribuição $Bin(10; 0,5)$ pela $N(5; 2,5)$

EXEMPLO 2.6 Honestidade de uma moeda

Uma moeda é lançada 50 vezes, com o objetivo de se verificar sua honestidade. Se ocorrem 36 caras nos 50 lançamentos, o que podemos concluir?

Solução

Suponhamos que a moeda seja honesta, isto é, que $p = 1/2$. Neste caso, a variável X = "número de caras em 50 lançamentos" segue uma distribuição binomial com parâmetros $n = 50$ e $p = 1/2$. Como n é grande, podemos usar a aproximação normal dada pelo Teorema Limite Central. Assim, poderíamos calcular a probabilidade de obtermos 36 caras em 50 lançamentos como $\mathrm{P}(35,5 \leq X \leq 36,5)$, em que $X \sim N\left(25; \frac{50}{4}\right)$. No entanto, nosso objetivo é verificar se a moeda é, ou não, honesta. Sob tal ponto de vista, se o resultado 36 caras for indício de que a moeda não é honesta, qualquer outro valor acima também o será. Dessa forma, é mais razoável responder à pergunta do problema calculando a probabilidade de obtermos 36 ou mais caras em 50 lançamentos. Se essa probabilidade for pequena, desconfiaremos da honestidade da moeda.

O Teorema Limite Central nos diz que, se $X \sim Bin(50; 1/2)$ então $X \approx N(25; 50/4)$. Logo,

$$\mathrm{P}(X \geq 36) \approx \mathrm{P}\left(Z \geq \frac{35,5-25}{\sqrt{50/4}}\right) = \mathrm{P}(Z \geq 2,97) = 0,5 - 0,4985 = 0,0015$$

Note que essa probabilidade é bastante pequena, ou seja, há uma pequena probabilidade de obtermos 36 ou mais caras em 50 lançamentos de uma moeda honesta. Isso pode nos levar a suspeitar sobre a honestidade da moeda!

♦♦

2.2 Distribuição amostral da proporção

Consideremos, agora, uma população em que cada elemento é classificado de acordo com a presença ou ausência de determinada característica. Por exemplo, podemos pensar em eleitores escolhendo entre dois candidatos, pessoas classificadas de acordo com o sexo, trabalhadores classificados como trabalhador com carteira assinada ou não, e assim por diante. Em termos de variável aleatória, essa população é representada por uma variável de Bernoulli, isto é:

$$X = \begin{cases} 1, & \text{se elemento possui a característica de interesse} \\ 0, & \text{se elemento não possui a caracaterística de interesse} \end{cases}$$

Vamos denotar por p a proporção de elementos da população que possuem a característica de interesse. Então,

$$\begin{aligned} \mathrm{P}(X = 1) &= p \\ \mathrm{E}(X) &= p \\ \mathrm{Var}(X) &= p(1-p) \end{aligned}$$

Em geral, o parâmetro p é desconhecido e precisamos estimá-lo a partir de uma amostra, da mesma forma como fizemos no caso da média de uma população normal. Então, seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória simples de uma população $X \sim Bern(p)$. Sabemos que

$$\begin{aligned} \mathrm{E}(\overline{X}) &= \mathrm{E}(X) = p \\ \mathrm{Var}(\overline{X}) &= \frac{\mathrm{Var}(X)}{n} = \frac{p(1-p)}{n} \end{aligned}$$

Mas, note que $\overline{X}$ nada mais é que a proporção dos elementos da amostra que possuem a característica de interesse, ou seja, $\overline{X}$ é a proporção amostral, que denotaremos por $\widehat{P}$. Resulta, então, que $\widehat{P}$ é um estimador não-viesado para a proporção populacional p.

Pelo Teorema Limite Central, se n for suficientemente grande, então

$$\widehat{P} \approx N\left(p; \frac{p(1-p)}{n}\right)$$

A aproximação dada pelo Teorema Limite Central será melhor para valores grandes de n e a mesma regra empírica usada para a aproximação da binomial pela normal se aplica aqui, conforme explicado a seguir.

! Distribuição amostral da proporção amostral

Seja $X_1, X_2, \cdots, X_n$ uma amostra aleatória simples de uma população $X \sim Bern(p)$. Para n suficientemente grande, a distribuição da proporção amostral pode ser aproximada pela distribuição normal com média $\mu = p$ e variância $\sigma^2 = \frac{p(1-p)}{n}$, isto é,

$$\widehat{P} \approx N\left(p; \frac{p(1-p)}{n}\right) \tag{2.5}$$

Essa aproximação pode ser usada se as seguintes condições forem satisfeitas:

1. $n \geq 30$
2. $np \geq 5$
3. $n(1-p) \geq 5$

EXEMPLO 2.7 Itens defeituosos num lote

De um grande lote de produtos manufaturados, extrai-se uma amostra aleatória simples de 200 itens. Se 10% dos itens do lote são defeituosos, calcule a probabilidade de serem sorteados no máximo 24 itens defeituosos.

Solução

As condições para utilização da aproximação normal são válidas, pois com $n = 200$ e $p = 0,1$ temos que:

$$\begin{aligned} 200 \times 0,1 &= 20 > 10 \\ 200 \times 0,9 &= 180 > 10 \end{aligned}$$

Ter no máximo 24 itens defeituosos na amostra equivale a ter uma proporção amostral de, no máximo, 0,12. Então, o problema pede

$$\begin{aligned} \mathrm{P}(\widehat{P} \leq 0,12) &= \mathrm{P}\left(\frac{\widehat{P} - 0,1}{\sqrt{\frac{0,1 \times 0,9}{200}}} \leq \frac{0,12 - 0,1}{\sqrt{\frac{0,1 \times 0,9}{200}}}\right) \\ &\approx \mathrm{P}(Z \leq 0,9428) = 0,5 + tab(0,94) = 0,8264 \end{aligned}$$

O valor exato é $\mathrm{P}(X \leq 24) = 0,855106$.

♦♦

EXEMPLO 2.8 Confiabilidade de componentes

A confiabilidade de um componente é a probabilidade de que ele funcione sob as condições desejadas. Uma amostra aleatória simples de 2500 desses componentes é extraída e cada componente testado. Calcule a probabilidade de obtermos pelo menos 75 itens defeituosos supondo que a confiabilidade do item seja:

(a) 0,995

(b) 0,85

Solução

Ter pelo menos 75 defeituosos é equivalente a ter uma proporção de defeituosos de pelo menos 0,03.

(a) Se a confiabilidade é 0,995, então a probabilidade de o item ser defeituoso é 0,005. Note que $2500 \times 0,005 = 12,5$ e $2500 \times 0,995 = 2487,5$ de modo que podemos usar a aproximação normal.

$$\begin{aligned} P(\widehat{P} \geq 0,03) &= P\left(\frac{\widehat{P} - 0,005}{\sqrt{\frac{0,005 \times 0,9995}{2500}}} \geq \frac{0,03 - 0,005}{\sqrt{\frac{0,005 \times 0,995}{25000}}}\right) \\ &\approx P(Z \geq 17,7) \approx 0 \end{aligned}$$

(b) Se a confiabilidade é 0,85, então a probabilidade de o item ser defeituoso é 0,15. Note que $2500 \times 0,15 = 375$ e $1.000 \times 0,85 = 2125$, de modo que podemos usar a aproximação normal.

$$\begin{aligned} P(\widehat{P} \geq 0,03) &= P\left(\frac{\widehat{P} - 0,15}{\sqrt{\frac{0,15 \times 0,85}{2500}}} \geq \frac{0,03 - 0,15}{\sqrt{\frac{0,15 \times 0,85}{2500}}}\right) \\ &\approx P(Z \geq -16,8) \approx 1 \end{aligned}$$

♦♦

2.3 Estimadores da variância

No capítulo anterior, consideramos dois estimadores para a variância: S^2 e $\widehat{\sigma}^2$. Através de um exemplo, vimos que $\widehat{\sigma}^2$ é um estimador viesado. Vamos demonstrar agora que S^2 é não-viesado para estimar a variância de uma população qualquer.

TEOREMA 2.5

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória simples extraída de uma população com N elementos e variância populacional

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2$$

em que $\mu = \frac{1}{N}\sum\limits_{i=1}^{N} X_i$ é a média (esperança) populacional. Então

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

é um estimador não-viesado para estimar σ^2.

Demonstração

$$
\begin{aligned}
\sum_{i=1}^{n}(X_i-\overline{X})^2 &= \sum_{i=1}^{n}(X_i-\mu+\mu-\overline{X})^2 = \sum_{i=1}^{n}\left[(X_i-\mu)-\left(\overline{X}-\mu\right)\right]^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2+\sum_{i=1}^{n}\left(\overline{X}-\mu\right)^2-2\sum_{i=1}^{n}(X_i-\mu)\left(\overline{X}-\mu\right) \\
&= \sum_{i=1}^{n}(X_i-\mu)^2+n\left(\overline{X}-\mu\right)^2-2\left(\overline{X}-\mu\right)\sum_{i=1}^{n}(X_i-\mu) \\
&= \sum_{i=1}^{n}(X_i-\mu)^2+n\left(\overline{X}-\mu\right)^2-2\left(\overline{X}-\mu\right)\left(\sum_{i=1}^{n}X_i-n\mu\right) \\
&= \sum_{i=1}^{n}(X_i-\mu)^2+n\left(\overline{X}-\mu\right)^2-2\left(\overline{X}-\mu\right)\left(n\overline{X}-n\mu\right) \\
&= \sum_{i=1}^{n}(X_i-\mu)^2-n\left(\overline{X}-\mu\right)^2
\end{aligned}
$$

Daí segue que

$$
\begin{aligned}
E(S^2) &= E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\right] = \frac{1}{n-1}E\left[\sum_{i=1}^{n}(X_i-\mu)^2-n\left(\overline{X}-\mu\right)^2\right] \\
&= \frac{1}{n-1}\left[\sum_{i=1}^{n}E\,(X_i-\mu)^2-nE\left(\overline{X}-\mu\right)^2\right]
\end{aligned}
$$

Mas como $\mu = E(X_i) = E(\overline{X})$ e $E\,(X_i-\mu)^2 = Var(X_i) = \sigma^2$ e $E\left(\overline{X}-\mu\right)^2 = Var(\overline{X})$ resulta que

$$
\begin{aligned}
E(S^2) &= \frac{1}{n-1}\left[\sum_{i=1}^{n}Var(X_i)-nVar(\overline{X})\right] \\
&= \frac{1}{n-1}\left(\sum_{i=1}^{n}\sigma^2-n\frac{\sigma^2}{n}\right) \\
&= \frac{1}{n-1}\left(n\sigma^2-\sigma^2\right) \\
&= \sigma^2
\end{aligned}
$$

e isso completa a prova. ■

Note que

$$
\begin{aligned}
S^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2 \\
\widehat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2
\end{aligned}
$$

Logo,

$$
\frac{\widehat{\sigma}^2}{S^2} = \frac{n-1}{n} \Rightarrow \widehat{\sigma}^2 = \frac{n-1}{n}S^2 \Rightarrow \mathrm{E}(\widehat{\sigma}^2) = \frac{n-1}{n}\sigma^2
$$

ou seja, $\widehat{\sigma}^2$ é viesado para estimar σ^2. Mas notem que o fator de viés, dado por $\frac{n-1}{n}$, tende a 1 quando $n \to +\infty$.

2.4 Exercícios propostos

1. Uma amostra de tamanho $n = 18$ é extraída de uma população normal com média 15 e desvio padrão 2,5. Calcule a probabilidade de que a média amostral

   (a) esteja entre 14,5 e 16,0;

   (b) seja maior que 16,1.

2. Uma empresa produz parafusos em duas máquinas. O comprimento dos parafusos produzidos em ambas é aproximadamente normal com média de 20 mm na primeira máquina e 25 mm na segunda máquina e desvio padrão comum de 4 mm. Uma caixa com 16 parafusos, sem identificação, é encontrada e o gerente de produção determina que, se o comprimento médio for maior que 23 mm, então a caixa será identificada como produzida pela máquina 2. Especifique os possíveis erros nessa decisão e calcule as suas probabilidades.

3. Definimos a variável $e = \overline{X} - \mu$ como sendo o *erro amostral* da média, onde $\overline{X}$ é a média de uma amostra aleatória simples de tamanho n de uma população com média μ e desvio padrão σ.

   (a) Determine $E(e)$ e $Var(e)$.

   (b) Se a população é normal com $\sigma = 20$, que proporção das amostras de tamanho 100 terá erro amostral absoluto maior do que 2 unidades?

   (c) Neste caso, qual deve ser o valor de δ para que $\mathrm{P}(|\,e\,| > \delta) = 0,01$?

   (d) Qual deve ser o tamanho da amostra para que 95% dos erros amostrais absolutos sejam inferiores a 1 unidade?

4. Uma fábrica produz parafusos especiais, para atender um determinado cliente, que devem ter comprimento de 8,5 cm. Como os parafusos grandes podem ser reaproveitados a um custo muito baixo, a fábrica precisa controlar apenas a proporção de parafusos pequenos. Para que o processo de produção atinja o lucro mínimo desejável, é necessário que a proporção de parafusos pequenos seja no máximo de 5%.

   (a) Supondo que a máquina que produz os parafusos o faça de modo que os comprimentos tenham distribuição normal com média μ e desvio padrão de 1,0 cm, em quanto deve ser regulada a máquina para satisfazer as condições de lucratividade da empresa?

   (b) Para manter o processo sob controle, é programada uma carta de qualidade. A cada hora será sorteada uma amostra de 4 parafusos e, se o comprimento médio dessa amostra for menor que 9,0 cm, o processo de produção é interrompido para uma nova regulagem da máquina. Qual é a probabilidade de uma parada desnecessária?

   (c) Se a máquina se desregulou de modo que o comprimento médio passou a ser 9,5 cm, qual é a probabilidade de se continuar o processo de produção fora dos padrões desejados?

5. A divisão de inspeção do Departamento de Pesos e Medidas de uma determinada cidade está interessada em calcular a real quantidade de refrigerante que é colocada em garrafas de 2 litros, no setor de engarrafamento de uma grande empresa de refrigerantes. O gerente do setor de engarrafamento informou à divisão de inspeção que o desvio padrão para garrafas de 2 litros é de 0,05 litro. Uma amostra aleatória de 100 garrafas de 2 litros, obtida deste setor de engarrafamento, indica uma média de 1,985 litro. Qual é a probabilidade de se obter uma média amostral de 1,985 ou menos, caso a afirmativa do gerente esteja certa? O que se pode concluir?

6. Em cada um dos exercícios abaixo, verifique que as condições para aproximação da binomial pela normal são satisfeitas e calcule a probabilidade pedida usando a aproximação normal.

   (a) $X \sim bin(18; 0,4)$ $\quad$ $\mathrm{P}(X \geq 15)$ e $\mathrm{P}(X < 2)$

   (b) $X \sim bin(40; 0,3)$ $\quad$ $\mathrm{P}(X < 10)$ e $\mathrm{P}(25 < X < 28)$

(c) $X \sim bin(65; 0,9)$ $P(X = 58)$ e $P(60 < X \leq 63)$
(d) $X \sim bin(100; 0,2)$ $P(25 \leq X \leq 35)$
(e) $X \sim bin(50; 0,2)$ $P(X > 26)$ e $P(5 \leq X < 10)$
(f) $X \sim bin(50; 0,7)$ $P(X \leq 25)$
(g) $X \sim bin(100; 0,5)$ $P(42 < X \leq 56)$
(h) $X \sim bin(100; 0,5)$ $P(X > 60)$
(i) $X \sim bin(20; 0,4)$ $P(X = 5)$
(j) $X \sim bin(30; 0,3)$ $P(X \geq 12)$
(k) $X \sim bin(80; 0,1)$ $P(9 < X < 11)$
(l) $X \sim bin(30; 0,2)$ $P(12 \leq X \leq 16)$
(m) $X \sim bin(50; 0,3)$ $P(X > 18)$
(n) $X \sim bin(28; 0,2)$ $P(X = 6)$
(o) $X \sim bin(95; 0,4)$ $P(30 \leq X < 48)$

7. Em uma sondagem, perguntou-se a 1002 membros de determinado sindicato se eles haviam votado na última eleição para a direção do sindicato e 701 responderam afirmativamente. Os registros oficiais obtidos depois da eleição mostram que 61% dos membros aptos a votar de fato votaram. Calcule a probabilidade de que, dentre 1002 membros selecionados aleatoriamente, no mínimo 701 tenham votado, considerando que a verdadeira taxa de votantes seja de 61%. O que o resultado sugere?

8. Supondo que meninos e meninas sejam igualmente prováveis, qual é a probabilidade de nascerem 36 meninas em 64 partos? Em geral, um resultado é considerado não-usual se a sua probabilidade de ocorrência é pequena, digamos, menor que 0,05. é não-usual nascerem 36 meninas em 64 partos?

9. Com base em dados históricos, uma companhia aérea estima em 15% a taxa de desistência entre seus clientes, isto é, 15% dos passageiros com reserva não aparecem na hora do vôo. Para otimizar a ocupação de suas aeronaves, essa companhia decide aceitar 400 reservas para os vôos em aeronaves que comportam apenas 350 passageiros. Calcule a probabilidade de que essa companhia não tenha assentos suficientes em um desses vôos. Essa probabilidade é alta o suficiente para a companhia rever sua política de reserva?

10. No controle de qualidade de produtos, uma técnica comumente utilizada é a *amostragem de aceitação*. Segundo essa técnica, um lote inteiro é rejeitado se contiver mais do que um número determinado de itens defeituosos. A companhia X compra parafusos de uma fábrica em lotes de 5000 e rejeita o lote se uma amostra aleatória simples de 20 parafusos contiver pelo menos 2 defeituosos. Se o processo de fabricação tem uma taxa de 10% de defeituosos, qual é a probabilidade de um lote ser rejeitado pela companhia X?

Capítulo 3

Intervalos de confiança baseados na distribuição normal

3.1 Ideias básicas sobre intervalos de confiança

O objetivo central da Inferência Estatística é obter informações sobre uma população a partir do conhecimento de uma única amostra. Em geral, a população é representada por uma variável aleatória X, com uma lei de probabilidade f_X, que pode ser discreta ou contínua. Dessa população, então, extrai-se uma amostra aleatória simples com reposição, que dá origem a um conjunto $X_1, X_2, \ldots, X_n$ de n variáveis aleatórias independentes e identicamente distribuídas, todas com a mesma distribuição f_X. Se f_X depende de um ou mais parâmetros, temos de usar a informação obtida a partir da amostra para estimar esses parâmetros, de forma a conhecermos a distribuição.

Nos capítulos anteriores, por exemplo, vimos que a média amostral $\overline{X}$ é um bom estimador da média populacional μ, no sentido de que ela tende a "acertar o alvo" da verdadeira média populacional, ou seja, $\overline{X}$ é um estimador não-viesado para μ. Mas vimos, também, que existe uma variabilidade nos valores de $\overline{X}$, ou seja, cada possível amostra dá origem a um valor diferente do estimador.

Na prática, temos apenas uma amostra e, assim, é importante que se dê alguma informação sobre essa possível variabilidade do estimador. Ou seja, é importante informar o valor do estimador $\widehat{\theta}$ obtido com uma amostra específica, mas é importante informar também que o verdadeiro valor do parâmetro θ poderia estar em um determinado intervalo, digamos, no intervalo $[\widehat{\theta}_1, \widehat{\theta}_2]$. Dessa forma, informamos a nossa *margem de erro* no processo de estimação; tal margem de erro é consequência do processo de seleção aleatória da amostra.

O que vamos estudar agora é como obter esse intervalo, de modo a "acertar na maioria das vezes", isto é, queremos um procedimento que garanta que, na maioria das vezes – das amostras possíveis –, o intervalo obtido conterá o verdadeiro valor do parâmetro. A expressão "na maioria das vezes" será traduzida como "probabilidade alta". Veja a Figura 3.1: aí os intervalos são representados pelas linhas horizontais e podemos ver que 2 deles, em preto, não "acertam o alvo", no sentido de não conterem o verdadeiro valor do parâmetro θ, representado pela linha vertical preta tracejada.

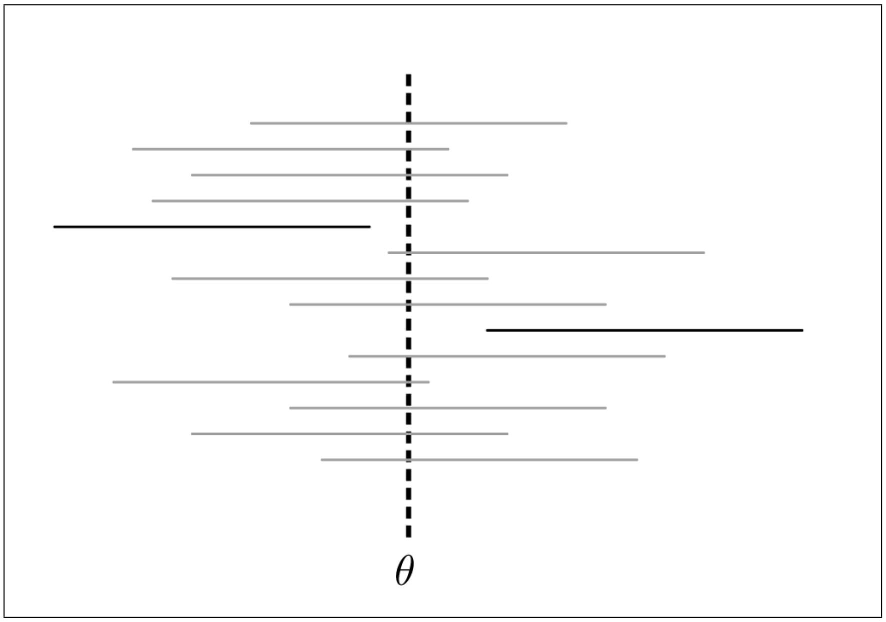


Figura 3.1 – Interpretação dos intervalos de confiança

DEFINIÇÃO Intervalo de confiança

Seja uma população representada pela variável aleatória X cuja lei de probabilidade depende de um parâmetro desconhecido θ. Se $X_1, X_2, \cdots X_n$ é uma amostra aleatória simples de tal população, então as estatísticas $\widehat{\theta}_1$ e $\widehat{\theta}_2$, $\widehat{\theta}_1 < \widehat{\theta}_2$, formam um intervalo de confiança de $100(1-\alpha)\%$ para θ se

$$\mathrm{P}\left(\widehat{\theta}_1 \leq \theta \leq \widehat{\theta}_2\right) \geq 1-\alpha \tag{3.1}$$

qualquer que seja o valor de θ. $1-\alpha$ é chamado *nível de confiança* do intervalo.

Note que, em (3.1), no máximo um dos limites $\widehat{\theta}_1$ e $\widehat{\theta}_2$ pode ser uma constante, incluindo $\pm\infty$. Sendo assim, o intervalo de confiança $\left(\widehat{\theta}_1, \widehat{\theta}_2\right)$ é, na verdade, um intervalo aleatório cujos limites dependem da amostra sorteada, ou seja, os limites do intervalo de confiança são *variáveis aleatórias*. Cada amostra dá origem a um intervalo diferente, mas o procedimento de obtenção dos intervalos garante probabilidade $1-\alpha$ de "acerto". Para um intervalo específico, obtido a partir de uma amostra específica, ou ele contém, ou ele não contém o verdadeiro parâmetro e não temos condições de saber qual dessas duas opções é verdadeira. Só sabemos que o *método* de obtenção do intervalo garante probabilidade $1-\alpha$ de acerto.

EXEMPLO 3.1 Interpretando um intervalo de confiança

Em um estudo sobre o Índice de Massa Corporal (IMC), foi reportado o seguinte intervalo de confiança de 95% para o IMC médio μ de determinada população, com base em uma amostra de 650 mulheres: $[26,2; 27,4]$. O que podemos dizer e o que não podemos dizer com base nesse intervalo?

Solução

O que definitivamente *não podemos dizer* é que há uma probabilidade de 0,95 de μ, o verdadeiro IMC médio populacional, estar no intervalo dado. Note que o intervalo dado é um único intervalo – ou μ está no intervalo ou μ não está no intervalo e não temos como saber qual é verdade. O que interessa é que apenas uma dessas afirmativas é verdadeira com probabilidade 1 e a outra, portanto, não pode acontecer (probabilidade 0).

O que *podemos dizer* sobre o intervalo dado é que ele foi gerado a partir de uma amostra específica com um método que tem 95% de chance de gerar intervalos, baseados em outras amostras, que conterão o parâmetro populacional μ. Isso nos leva a dizer que estamos 95% confiantes em que nosso intervalo obtido a partir dessa amostra específica contém o verdadeiro parâmetro μ.

♦♦

3.2 Intervalo de confiança para a média de uma população normal, σ^2 conhecida

Vamos agora, introduzir os métodos para obtenção do intervalo de confiança para a média de uma população. Como visto, a média populacional é um parâmetro importante, que pode ser muito bem estimado pela média amostral $\overline{X}$. Para apresentar as ideias básicas, vamos considerar um contexto que é pouco frequente na prática. O motivo para isso é que, em termos didáticos, a apresentação é bastante simples. Como o fundamento é o mesmo para contextos mais gerais, essa abordagem se justifica.

Consideremos, então, uma população descrita por uma variável aleatória normal com média μ e variância $\sigma^2 : X \sim N(\mu; \sigma^2)$. Vamos supor que o valor de σ^2 seja conhecido e que nosso interesse seja estimar a média μ a partir de uma amostra aleatória simples $X_1, X_2, \dots, X_n$. Como visto no Capítulo 2, a distribuição amostral de $\overline{X}$ é normal com média μ e variância $\frac{\sigma^2}{n}$, ou seja

$$X \sim N\left(\mu; \sigma^2\right) \Longrightarrow \overline{X} \sim N\left(\mu; \frac{\sigma^2}{n}\right)$$

Da definição de distribuição amostral, isso significa que os diferentes valores de $\overline{X}$ obtidos a partir das diferentes possíveis amostras se distribuem normalmente em torno de μ com variância $\frac{\sigma^2}{n}$.

Das propriedades da distribuição normal, resulta que

$$Z = \frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0; 1)$$

ou equivalentemente,

$$Z = \sqrt{n}\frac{\overline{X} - \mu}{\sigma} \sim N(0; 1) \tag{3.2}$$

Para completar a construção do intervalo de confiança, vamos apresentar a seguinte definição, ilustrada na Figura 3.2:

DEFINIÇÃO Valor crítico da normal

O **valor crítico** de $Z \sim N(0;1)$ associado à probabilidade α é a abscissa z_α tal que

$$P(Z > z_\alpha) = \alpha \tag{3.3}$$

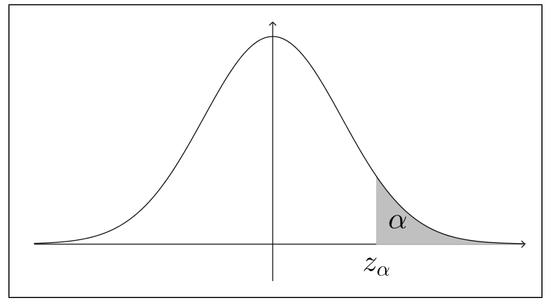


Figura 3.2 – Valor crítico z_α da $N(0;1)$

Se considerarmos, agora, o valor crítico $z_{\alpha/2}$, conforme ilustrado na Figura 3.3, resulta que, se $Z \sim N(0;1)$, então

$$P\left(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}\right) = 1 - \alpha \tag{3.4}$$

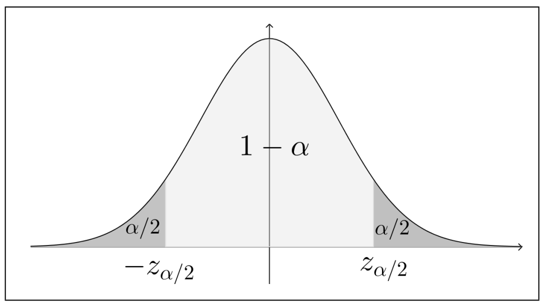


Figura 3.3 – Valor crítico $z_{\alpha/2}$ da $N(0;1)$

Mas isso vale para a distribuição normal padrão, em geral. Então, usando os resultados das Equações 3.2 e 3.4, obtemos que

$$P\left(-z_{\alpha/2} \leq \sqrt{n}\frac{\overline{X}-\mu}{\sigma} \leq z_{\alpha/2}\right) = 1-\alpha$$

o que é equivalente a

$$\begin{aligned}
P\left(-z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \overline{X}-\mu \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) &= 1-\alpha \Leftrightarrow \\
P\left(-\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq -\mu \leq -\overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) &= 1-\alpha \Leftrightarrow \\
P\left(\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) &= 1-\alpha \qquad (3.5)
\end{aligned}$$

Observe a expressão (3.5); ela nos diz que

$$P\left(\mu \in \left[\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\,;\,\overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]\right) = 1-\alpha$$

Mas essa é exatamente a forma geral de um intervalo de confiança, conforme explicitado na Equação 3.1 (note que os limites são variáveis aleatórias!). Temos, então, a seguinte conclusão:

! Intervalo de confiança para a média de uma população normal, σ^2 conhecida

Seja $X \sim N(\mu; \sigma^2)$ uma população, cuja variância σ^2 é conhecida. Se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória simples dessa população, então o intervalo de confiança de nível de confiança $1 - \alpha$ para a média populacional μ é dado por

$$\left[\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \,;\, \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] \tag{3.6}$$

em que $z_{\alpha/2}$ é o valor crítico da distribuição normal correspondente à probabilidade $\alpha/2$.

O intervalo de confiança para μ pode ser escrito na forma $[\overline{X} - \epsilon; \overline{X} + \epsilon]$, onde $\epsilon = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ é a *margem de erro*. Como visto, essa margem de erro está associada ao fato de que diferentes amostras fornecem diferentes valores de $\overline{X}$. As diferentes amostras fornecem diferentes intervalos de confiança, mas uma proporção de $100 \times (1 - \alpha)\%$ desses intervalos irá conter o verdadeiro valor de μ. Note que aqui é fundamental a interpretação de probabilidade como frequência relativa: estamos considerando os diferentes intervalos que seriam obtidos, caso sorteássemos todas as possíveis amostras. Assim, o nível de confiança está associado à confiabilidade do processo de obtenção do intervalo: esse processo é tal que acertamos (isto é, o intervalo contém μ) em $100 \times (1 - \alpha)\%$ das vezes. Na Figura 3.4 ilustra-se essa interpretação dos intervalos de confiança para uma população normal com variância 4 e tamanho de amostra $n = 16$. A distribuição normal padrão representa a distribuição de probabilidade dos valores de $\sqrt{16}\frac{\overline{X} - \mu}{2}$. Valores extremos de tal estatística levam a intervalos de confiança que não contêm o verdadeiro parâmetro, representados pelos intervalos em preto. Os valores centrais, que têm alta probabilidade $(1-\alpha)$ de ocorrência levam a intervalos que contêm o verdadeiro valor do parâmetro (intervalos em cinza).

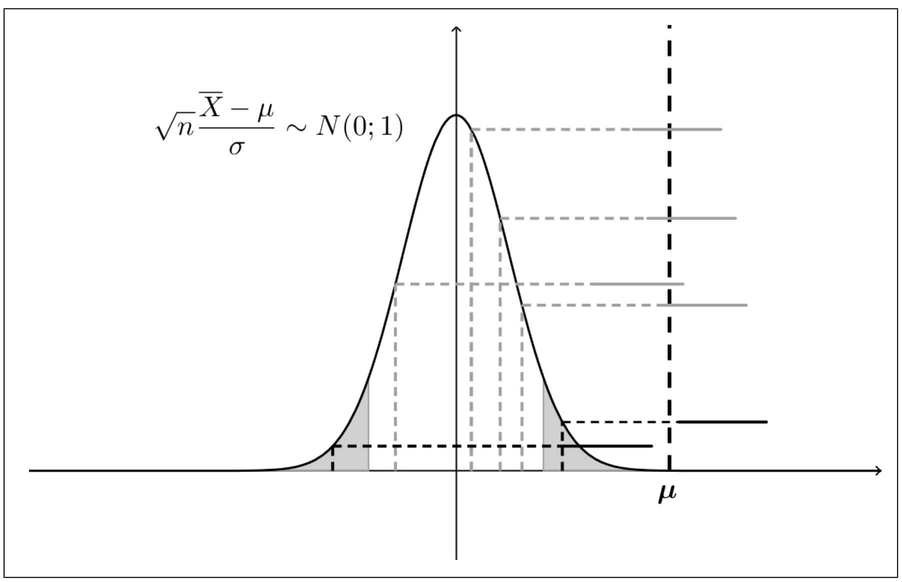


Figura 3.4 – Interpretação do IC para a média da $N(\mu; \sigma^2)$

Na prática, temos apenas *uma* amostra e o intervalo obtido com essa amostra específica, ou contém ou não contém o verdadeiro valor de μ. A afirmativa

$$\mathrm{P}\left(\mu \in \left[\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \,;\, \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]\right) = 1 - \alpha$$

é válida porque ela envolve a variável aleatória $\overline{X}$, que assume diferentes valores para as diferentes amostras. Quando substituímos o estimador $\overline{X}$ por uma estimativa específica $\overline{x}$ obtida a partir de uma amostra particular, temos apenas um intervalo e não faz mais sentido falar em probabilidade.

É interessante notar que alguns autores trabalham com os quantis da distribuição normal no lugar dos valores críticos e às vezes usam a mesma notação, ou seja, denotam por z_α o quantil de ordem α. Essa prática se justifica quando pensamos em termos de programas computacionais, que têm rotinas para cálculos com a função de distribuição da normal. Denotando por Φ a função de distribuição da normal padrão, temos, segundo essa notação:

$$\Phi(z_\alpha) = \alpha \Leftrightarrow z_\alpha = \Phi^{-1}(\alpha)$$

Em termos dos quantis, o intervalo de confiança de nível de confiança $1 - \alpha$ é dado por

$$\left[\overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\,;\,\overline{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

Note que, pela simetria da densidade normal, esses 2 quantis são simétricos, ou seja, um deles será sempre negativo! Nessa apostila trabalharemos com o valor crítico das distribuições, e não os quantis.

EXEMPLO 3.2 Pesos de homens adultos

Em determinada população, o peso dos homens adultos é distribuído normalmente com um desvio padrão de 16kg. Uma amostra aleatória simples de 36 homens adultos é sorteada desta população, obtendo-se um peso médio de 78,2kg. Construa um intervalo de confiança de nível de confiança 0,95 para o peso médio de todos os homens adultos dessa população.

Solução

Vamos inicialmente determinar o valor crítico associado ao nível de confiança de 0,95. Como $1-\alpha = 0,95$, resulta que $\alpha = 0,05$ e $\alpha/2 = 0,025$.

Analisando a Figura 3.3, vemos que a probabilidade nas duas caudas da distribuição normal padrão é de 0,05; logo, em cada cauda, a probabilidade é 0,025. Em termos da Tabela I da distribuição normal padrão no Apêndice A, isso significa que a probabilidade entre 0 e $z_{0,025}$ é $(0,50-0,025) = 0,475$ e, assim, devemos procurar no corpo da tabela o valor de 0,475 para determinar a abscissa $z_{0,025}$. Veja a Figura 3.5.

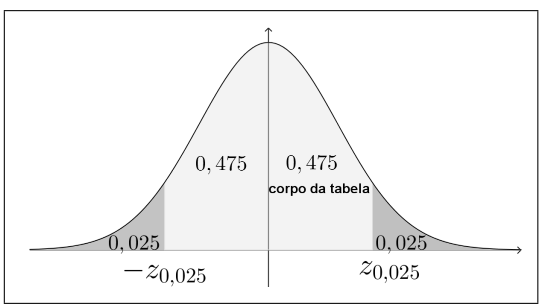


Figura 3.5 – Valor crítico $z_{0,025}$ da $N(0;1)$

Procurando no corpo da tabela da distribuição normal padrão, vemos que o valor 0,475 corresponde à abscissa $z_{0,025} = 1,96$. Logo, nosso intervalo de confiança é

$$\left[78,2 - 1,96 \times \frac{16}{\sqrt{36}}\,;\;\; 78,2 + 1,96 \times \frac{16}{\sqrt{36}}\right] = [72,9733\,;\;\; 83,4267]$$

Esse intervalo contém ou não o verdadeiro valor de μ, mas o procedimento utilizado para sua obtenção nos garante que há 95% de chance de estarmos certos, isto é, 95% dos intervalos construídos com esse métido conteriam o verdadeiro valor de μ.

♦♦

3.2.1 Margem de erro

Vamos, agora, analisar a margem de erro do intervalo de confiança para a média de uma população normal com variância conhecida. Ela é dada por

$$\epsilon = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \tag{3.7}$$

Lembrando que o erro padrão é o desvio padrão do estimador, podemos escrever

$$\epsilon = z_{\alpha/2}\, \mathrm{EP}_{\overline{X}} \tag{3.8}$$

Analisando a equação (3.7), vemos que a margem de erro depende diretamente do valor crítico e do desvio padrão populacional e é inversamente proporcional à raiz quadrado do tamanho da amostra.

Na Figura 3.6 ilustra-se a relação de dependência da margem de erro com o desvio padrão populacional σ. Temos duas distribuições amostrais centradas na mesma média e baseadas em amostras de mesmo tamanho: $\overline{X}_1 \sim N\left(\mu; \dfrac{\sigma_1^2}{n}\right)$ e $\overline{X}_2 \sim N\left(\mu; \dfrac{\sigma_2^2}{n}\right)$ com $\sigma_1^2 < \sigma_2^2$. Nas duas distribuições, a área total das caudas sombreadas é α, de modo que os intervalos limitados pelas linhas verticais são os intervalos de confiança de nível $1-\alpha$, ou seja, a área central em ambas distribuições é $1-\alpha$. Para a distribuição mais dispersa, isto é, com σ maior, o comprimento do intervalo é maior. Esse resultado deve ser intuitivo: se há mais variabilidade na população, a nossa margem de erro para estimação da média populacional tem que ser maior, mantidas fixas as outras condições (tamanho de amostra e nível de confiança).

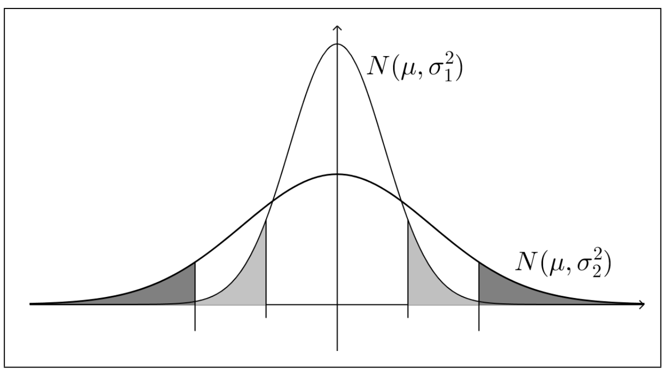


Figura 3.6 – Margem de erro versus dispersão populacional: $\sigma_1 < \sigma_2 \Rightarrow \epsilon_1 < \epsilon_2$

Por outro lado, se mantivermos fixos o tamanho da amostra e o desvio padrão populacional, é razoável, também, que a margem de erro seja maior para um nível de confiança maior. Ou seja, se queremos aumentar a probabilidade de acerto, é razoável que o intervalo seja maior. Aumentar a probabilidade de acerto significa aumentar o nível de confiança, o que acarreta em um valor crítico $z_{\alpha/2}$ maior. Veja a **Figura 3.7**, onde ilustra-se o intervalo de confiança para dois níveis de confiança diferentes: $1-\alpha_1 > 1-\alpha_2$. O primeiro intervalo é maior, refletindo o maior grau de confiança, ou seja, o preço que se paga por um nível de confiança maior é que o comprimento do intervalo de confiança também será maior.

Finalmente, mantidos o mesmo desvio padrão populacional e o mesmo nível de confiança, quanto maior o tamanho da amostra, menor será a margem de erro, mas a redução da margem de erro depende de $\sqrt{n}$; assim, para reduzir a margem de erro pela metade, teremos que quadruplicar o tamanho da amostra:

$$\epsilon' = \frac{\epsilon}{2} \Rightarrow \frac{1}{\sqrt{n'}} = \frac{1}{2}\frac{1}{\sqrt{n}} \Rightarrow \sqrt{n'} = 2\sqrt{n} \Rightarrow n' = 4n$$

EXEMPLO 3.3 Resultados de pesquisa

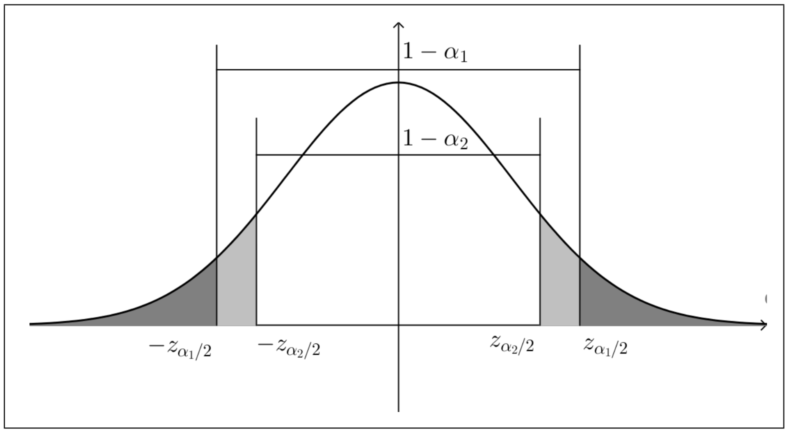


Figura 3.7 – Margem de erro versus nível de confiança: $\alpha_1 < \alpha_2 \Rightarrow (1-\alpha_1) > (1-\alpha_2) \Rightarrow \epsilon_1 > \epsilon_2$

Na divulgação dos resultados de uma pesquisa, publicou-se o seguinte texto (dados fictícios):

Com o objetivo de se estimar a média de uma população, estudou-se uma amostra de tamanho $n = 45$. De estudos anteriores, sabe-se que essa população é muito bem aproximada por uma distribuição normal com desvio padrão 3, mas acredita-se que a média tenha mudado desde esse último estudo. Com os dados amostrais obteve-se o intervalo de confiança $[1,79; 3,01]$.

Quais são o as informações importantes que não foram divulgadas? Como podemos obtê-las?

Solução

Quando se divulga um intervalo de confiança para um certo parâmetro, é costume publicar também a estimativa pontual. Nesse caso, temos que informar a média amostral $\overline{x}$, que pode ser achada observando-se que o intervalo de confiança é simétrico em torno de $\overline{x}$. Logo, $\overline{x}$ é o ponto médio do intervalo:

$$\overline{x} = \frac{1,79 + 3,01}{2} = 2,4$$

Daí conclui-se que a margem de erro é $\epsilon = 2,4 - 1,79 = 0,61$. Outra informação importante é o nível de confiança, que deve ser encontrado a partir da abscissa $z_{\alpha/2}$ na margem de erro:

$$0,61 = z_{\alpha/2} \times \frac{3}{\sqrt{45}} \Rightarrow z_{\alpha/2} = \frac{0,61 \times \sqrt{45}}{3} = 1,36$$

Consultando a tabela da distribuição normal, vemos que $P(0 \leq Z \leq 1,36) = 0,4131$. Logo, o nível de confiança é $2 \times 0,4131 = 0,8262 \approx 0,83$. Veja a Figura 3.8.

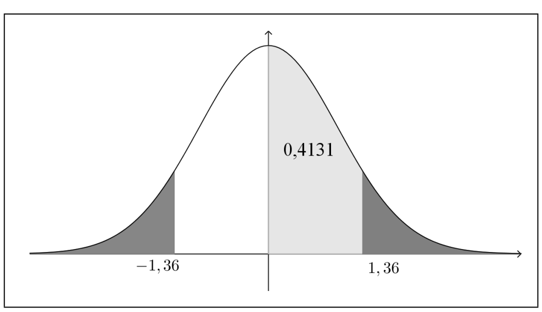


Figura 3.8 – Determinação do nível de confiança

♦♦

3.2.2 Determinação do tamanho da amostra

No planejamento de pesquisas, é importante ter-se uma ideia do tamanho de amostra necessário. Analisando a equação (3.7), pode-se observar que, na estimação da média de uma população normal com variância conhecida, temos

$$\epsilon = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \Longrightarrow \sqrt{n} = z_{\alpha/2}\frac{\sigma}{\epsilon} \Longrightarrow n = \left(z_{\alpha/2}\frac{\sigma}{\epsilon}\right)^2 \tag{3.9}$$

Assim, podemos determinar o tamanho da amostra necessário para valores pré estabelecidos da margem de erro e do nível de confiança. Note a relação entre o tamanho da amostra n e as três grandezas envolvidas: variância populacional, nível de confiança e margem de erro.

EXEMPLO 3.4 Tamanho de amostra

Deseja-se estimar a média de uma população normal com nível de confiança de 90% e margem de erro máxima de 0,08. Qual deve ser o tamanho da amostra se a variância populacional conhecida é

(a) $\sigma^2 = 4$

(b) $\sigma^2 = 16$

Solução

(a) $n_{\sigma=2} = \left(z_{0,05}\frac{2}{0,08}\right)^2 = \left(\frac{1,64 \cdot 2}{0,08}\right)^2 = 1681$

(b) $n_{\sigma=4} = \left(z_{0,05}\frac{4}{0,08}\right)^2 = \left(\frac{1,64 \cdot 4}{0,08}\right)^2 = 6724$

Note que a razão entre as variâncias populacionais é 4 e o mesmo ocorre com os tamanhos amostrais. ♦♦

3.2.3 Intervalos de confiança unilaterais

Para o intervalo de confiança apresentado, ambos limites, inferior e superior, são variáveis aleatórias. Em algumas situações, pode ser de interesse que apenas um desses limites seja aleatório. Eis alguns exemplos:

- Um psicólogo deseja estabelecer um limite de confiança superior de 90% para o verdadeiro tempo de reação a um determinado estímulo, ou seja, ele deseja um limite ℓ_S tal que 90% de todos os tempos de reação serão menores que ℓ_S.
- Um engenheiro deseja estabelecer um limite de confiança inferior de 90% para o verdadeiro tempo de vida médio de determinado tipo de componente eletrônico, ou seja, ele deseja um limite ℓ_I tal que 90% de todos os tempos de vida serão maiores que ℓ_I.

Intervalo de confiança unilateral superior

Consideremos uma população $X \sim N(\mu; \sigma^2)$ cuja variância é conhecida. Nosso objetivo é determinar um intervalo $(-\infty, \widehat{L}_S)$ tal que

$$\mathrm{P}[\mu \in (-\infty, \widehat{L}_S)] = 1 - \alpha$$

Note que isso equivale a

$$P(\mu < \widehat{L}_S) = 1 - \alpha$$

ou seja, estamos determinando o limite de confiança superior (exemplo do psicólogo). Como antes, $\widehat{L}_S$, o limite do intervalo, é uma variável aleatória e daí faz sentido falar em probabilidade. Note também que essa última expressão envolve uma probabilidade $1-\alpha > 0,5$ à direita ($\widehat{L}_S > \mu$); em termos de distribuição normal, isso equivale a uma abscissa menor que a média.

Sabemos que

$$P(Z \geq z_{1-\alpha}) = P(Z \geq -z_\alpha) = 1 - \alpha \Rightarrow P\left(\sqrt{n}\frac{\overline{X} - \mu}{\sigma} \geq -z_\alpha\right) = 1 - \alpha \Rightarrow P\left(\mu \leq \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

e, assim, o intervalo de confiança unilateral superior para μ é

$$\left(-\infty\,;\, \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right) \tag{3.10}$$

A interpretação é a mesma: se obtivéssemos várias e várias amostras e para cada uma construíssemos o intervalo de confiança de acordo com esse método, o intervalo conteria o verdadeiro parâmetro em $1 - \alpha$ das vezes. Veja a Figura 3.9: para os intervalos em preto, $\mu > \ell_S$ ($100 \times \alpha$% dos intervalos) e para os intervalos em cinza, $\mu \leq \ell_S$ ($100 \times (1 - \alpha)$% dos intervalos).

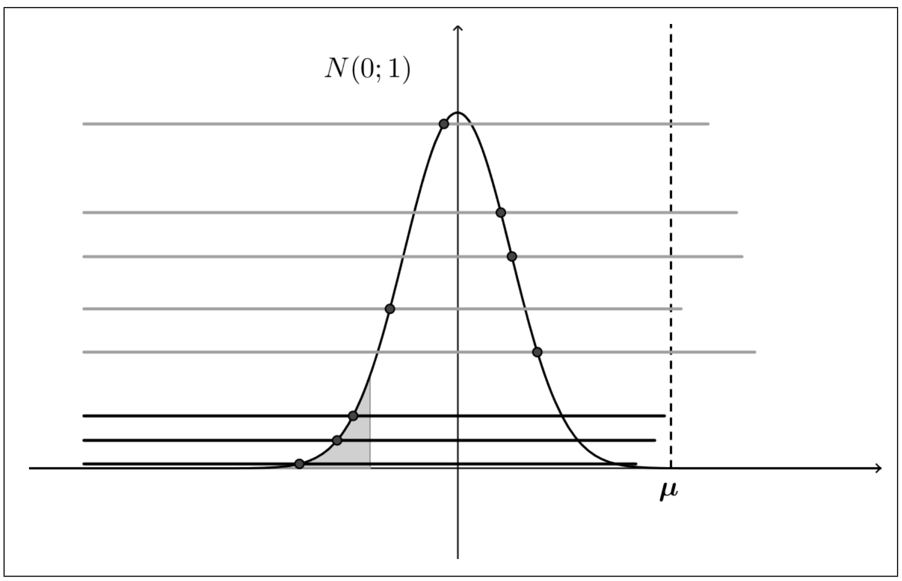


Figura 3.9 – Interpretação do IC unilateral superior para a média da $N(\mu;\sigma^2)$

Intervalo de confiança unilateral inferior

Consideremos novamente uma população $X \sim N(\mu;\sigma^2)$ cuja variância é conhecida. O objetivo agora é determinar um intervalo $(\widehat{L}_I, +\infty)$ tal que

$$P[\mu \in (\widehat{L}_I, +\infty)] = 1 - \alpha$$

Note que isso equivale a

$$P(\mu > \widehat{L}_I) = 1 - \alpha$$

ou seja, estamos determinando o limite de confiança inferior (exemplo do engenheiro). Como antes, $\widehat{L}_I$, o limite do intervalo, é uma variável aleatória, o que nos permite falar em probabilidade.

Sabemos que

$$\mathrm{P}(Z \leq z_\alpha) = 1 - \alpha \Rightarrow \mathrm{P}\left(\sqrt{n}\frac{\overline{X} - \mu}{\sigma} \leq z_\alpha\right) = 1 - \alpha \Rightarrow \mathrm{P}\left(\mu \geq \overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

e, assim, o intervalo de confiança unilateral inferior para μ é

$$\left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}\,;\,\infty\right) \tag{3.11}$$

A interpretação é a mesma: se obtivéssemos várias e várias amostras e para cada uma construíssemos o intervalo de confiança de acordo com esse método, o intervalo conteria o verdadeiro parâmetro em $1 - \alpha$ das vezes. Veja a Figura 3.10: para os intervalos em preto, $\mu < \ell_I$ (α% dos intervalos) e para os intervalos em cinza, $\mu \geq \ell_I$ ($(1 - \alpha)$% dos intervalos).

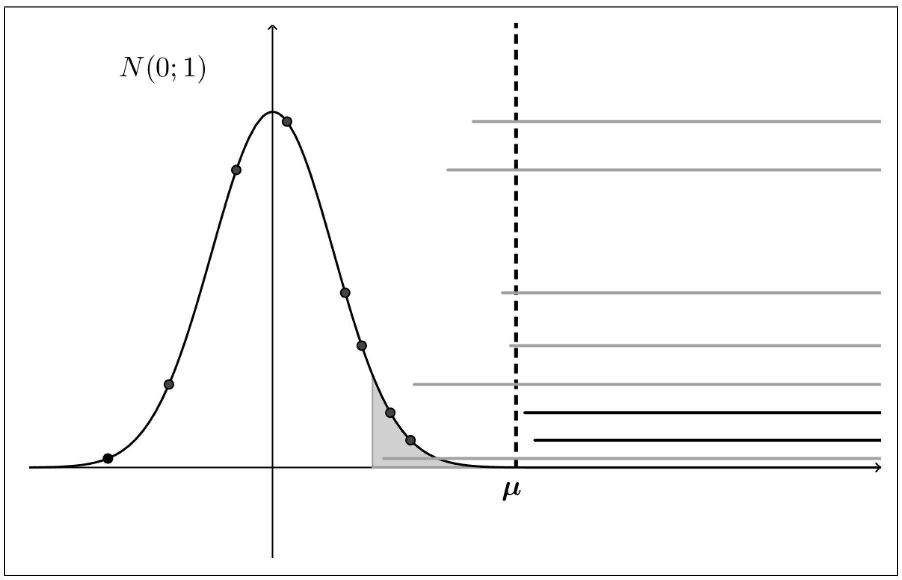


Figura 3.10 – Interpretação do IC unilateral inferior para a média da $N(\mu;\sigma^2)$

3.3 Intervalo de confiança para uma proporção

O procedimento de construção do intervalo de confiança para a proporção populacional é totalmente análogo ao do intervalo de confiança para a média de uma população normal com variância conhecida, visto anteriormente.

No Capítulo 2, vimos que, para amostras grandes,

$$\widehat{P} \approx N\left(p; \frac{p(1-p)}{n}\right) \tag{3.12}$$

Sendo assim, é verdade que

$$\mathrm{P}\left(-z_{\alpha/2} \leq \frac{\widehat{P} - p}{\sqrt{\frac{p(1-p)}{n}}} \leq z_{\alpha/2}\right) \approx 1 - \alpha$$

e, portanto

$$\begin{aligned}
\mathrm{P}\left(-z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \le \widehat{P} - p \le z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) &= 1-\alpha \Rightarrow \\
\mathrm{P}\left(-\widehat{P} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \le -p \le -\widehat{P} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) &= 1-\alpha \Rightarrow \\
\mathrm{P}\left(\widehat{P} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \le p \le \widehat{P} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) &= 1-\alpha
\end{aligned}$$

Como no caso da média, chegamos a uma expressão do seguinte tipo:

$$\mathrm{P}\left(\widehat{P} - \epsilon \le p \le \widehat{P} + \epsilon\right) = 1-\alpha$$

que é a expressão de um intervalo de confiança de nível de confiança $1-\alpha$ para a proporção populacional. Mas note que a margem de erro, neste caso, é

$$\epsilon = z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

e depende de p, o parâmetro de interesse. Sendo assim, temos que obter alguma estimativa de p para estimarmos a margem de erro e podermos construir o intervalo de confiança; essa estimativa pode vir de estudos anteriores, de informações de especialistas ou, então, da própria amostra usada para construir o intervalo de confiança. Vamos denotar essa estimativa por $\widehat{p}_0$.

! Intervalo de confiança para uma proporção populacional

Seja $X_1, X_2, \dots, X_n$ uma amostra aleatória simples de uma população $X \sim Bern(p)$. Para n suficientemente grande, o intervalo de confiança aproximado para p de nível de confiança $1-\alpha$ é dado por

$$\left[\widehat{P} - z_{\alpha/2}\sqrt{\frac{\widehat{p}_0(1-\widehat{p}_0)}{n}} \quad ; \quad \widehat{P} + z_{\alpha/2}\sqrt{\frac{\widehat{p}_0(1-\widehat{p}_0)}{n}}\right]$$

onde $\widehat{p}_0$ é alguma estimativa de p e $z_{\alpha/2}$ é o valor crítico da distribuição normal.

Uma outra abordagem para se obter uma estimativa prévia $\widehat{p}_0$ é utilizar o pior cenário, ou seja, o valor de p que resulte na maior variância possível do estimador. Da equação 3.3, vemos que essa variância depende $p(1-p)$. Na Figura 3.11 temos o gráfico da função $f(p) = p(1-p)$ para $p \in [0,1]$. Note que o máximo da função é atingido quando $p = 0,5$; então, mantidos o nível de confiança e o tamanho da amostra fixos, a margem de erro será máxima quando $p = 0,5$ e, nesse caso, teremos o maior intervalo de confiança possível. Essa é uma abordagem conservadora, que pode ser usada quando não se tem qualquer conhecimento sobre o valor p para gerar uma estimativa razoável $\widehat{p}_0$.

EXEMPLO 3.5 Linha de produção

Um gerente de produção deseja estimar a proporção de peças defeituosas em uma de suas linhas de produção. Para isso, ele seleciona uma amostra aleatória simples de 100 peças dessa linha de produção,

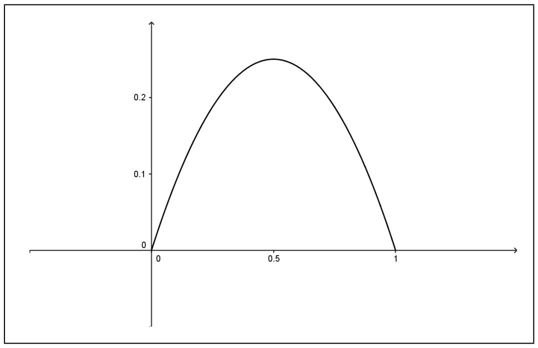


Figura 3.11 – Gráfico de $f(p) = p(1-p)$

obtendo 30 defeituosas. Determine o intervalo de confiança para a verdadeira proporção de peças defeituosas nessa linha de produção a um nível de confiança de 95%.

Solução

O primeiro fato a observar é que a amostra é grande, com sucessos e fracassos suficientes. Com um nível de confiança é $1-\alpha = 0,95$, temos que procurar na tabela da normal o valor crítico $z_{\alpha/2} = z_{0,025} = 1,96$. Como não temos estimativa prévia da proporção de defeituosas p, podemos usar a proporção amostral $\widehat{p} = 0,30$. Assim, a margem de erro estimada é

$$\epsilon = 1,96 \times \sqrt{\frac{0,3 \times 0,7}{100}} = 0,0898$$

e o intervalo de confiança é

$$[0,30 - 0,0898; 0,30 + 0,0898] = [0,2102; 0,3898]$$

Se usarmos a estimativa conservadora, a margem de erro será

$$\epsilon = 1,96 \times \sqrt{\frac{0,5 \times 0,5}{100}} = 0,098$$

que resulta no intervalo $[0,202\,;\,0,398]$. ♦♦

3.3.1 Margem de erro

Assim como no caso da média da população normal, a margem de erro do intervalo de confiança para uma proporção populacional pode ser escrita como

$$\epsilon = z_{\alpha/2}\,\mathrm{EP}_{\widehat{P}}$$

A diferença fundamental aqui é que temos de estimar o erro padrão como

$$\widehat{\mathrm{EP}}_{\widehat{P}} = \sqrt{\frac{\widehat{p}_0(1-\widehat{p}_0)}{n}}$$

enquanto no contexto da população normal com variância conhecida, o erro padrão era conhecido e igual a

$$\mathrm{EP}_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

Mas em ambos os casos, a margem de erro será maior para níveis de confiança maiores e para populações mais dispersas e pode ser diminuída aumentando-se o tamanho da amostra, mantendo-se constantes as demais condições.

3.3.2 Determinação do tamanho da amostra

Para estimação de uma proporção populacional, temos

$$\epsilon = z_{\alpha/2}\frac{\sqrt{\widehat{p}_0(1-\widehat{p}_0)}}{\sqrt{n}} \Longrightarrow \sqrt{n} = z_{\alpha/2}\frac{\sqrt{\widehat{p}_0(1-\widehat{p}_0)}}{\epsilon} \Longrightarrow n = \left(z_{\alpha/2}\frac{\sqrt{\widehat{p}_0(1-\widehat{p}_0)}}{\epsilon}\right)^2 \tag{3.13}$$

Trabalhando com o pior cenário, isto é, com $p = 0,5$, essa fórmula, que resulta no maior tamanho de amostra possível para ϵ e $1-\alpha$ fixos, se simplifica para

$$n = \left(z_{\alpha/2}\frac{0,5}{\epsilon}\right)^2 \Longrightarrow n = \left(\frac{z_{\alpha/2}}{2\epsilon}\right)^2 \tag{3.14}$$

EXEMPLO 3.6 Lançamento de um novo produto

Para estudar a viabilidade de lançamento de um novo produto no mercado, o gerente de uma grande empresa contrata uma firma de consultoria estatística para estudar a aceitação do produto entre os clientes potenciais. O gerente deseja obter uma estimativa com erro máximo de 1% com probabilidade de 80% e pede ao consultor estatístico que forneça o tamanho de amostra necessário.

(a) De posse das informações dadas, o consultor calcula o tamanho da amostra necessário no pior cenário. O que significa "pior cenário" nesse caso? Qual o tamanho de amostra obtido pelo consultor?

(b) O gerente acha que o custo de tal amostra seria muito alto e autoriza o consultor a realizar um estudo piloto com uma amostra de 100 pessoas para obter uma estimativa da verdadeira proporção. O resultado desse estudo piloto é uma estimativa $\widehat{p} = 0,76$ de aceitação do novo produto. Com base nessa estimativa, o consultor recalcula o tamanho da amostra necessário. Qual é esse tamanho?

(c) Selecionada a amostra com o tamanho obtido no item anterior, obteve-se uma proporção de 72% de clientes favoráveis ao produto. Construa um intervalo de confiança para a verdadeira proporção com nível de confiança de 90%.

Solução

(a) O pior cenário é quando a população está dividida meio-a-meio em suas preferências, ou seja, quando $p = 0,5$. Com nível de confiança de 80%, obtemos $z_{0,10} = 1,28$ – essa abscissa deixa 10% em cada cauda da distribuição normal padrão. Nesse caso,

$$0,01 = 1,28 \times \sqrt{\frac{0,5 \times 0,5}{n}} \Longrightarrow n = \left(\frac{1,28}{0,01}\right)^2 \times 0,25 = 4096$$

(b) Vamos agora utilizar $\widehat{p} = 0,76$:

$$0,01 \quad = \quad 1,28 \times \sqrt{\frac{0,76 \times 0,24}{n}} \Longrightarrow n = \left(\frac{1,28}{0,01}\right)^2 \times 0,76 \times 0,24 = 2988,4$$

ou seja, $n = 2989$.

(c) $1-\alpha = 0,90 \Longrightarrow z_{0,05} = 1,64$ – essa abscissa deixa 5% em cada cauda da distribuição normal padrão.

$$\epsilon = 1,64 \times \sqrt{\frac{0,72 \times 0,28}{2989}} = 0,0135$$

e o intervalo de confiança é

$$[0,72 - 0,0135; 0,72 + 0,0135] = [0,7065; 0,7335]$$

♦♦

3.4 Exercícios propostos

1. Encontre os valores críticos da normal padrão correspondentes aos seguintes níveis de confiança $1 - \alpha = 0,90; 0,99; 0,80$.

2. Encontre o nível de confiança correspondente aos seguintes valores críticos $z_{\alpha/2} = 1,28; 1,80$.

3. De uma população normal com desvio padrão 2, extrai-se uma amostra aleatória simples de tamanho 36, que fornece o seguinte resultado: $\sum_{i=1}^{36} x_i = 1236$. Calcule o intervalo de confiança para a média populacional μ, utilizando o nível de confiança de 98%.

4. Considere os dois intervalos de confiança a seguir, obtidos a partir de uma mesma amostra de uma população $N(\mu; 16)$. Sem fazer qualquer cálculo, identifique para qual deles o nível de confiança é maior.

$$[13,04; 16,96]$$
$$[12,42; 17,58]$$

5. Obtido um intervalo de confiança para a média de uma $N(\mu; 25)$, o que deve ser feito para se reduzir a margem de erro pela metade se não devemos alterar o nível de confiança?

6. De uma população $N(\mu; 9)$ extrai-se uma amostra aleatória simples de tamanho 25, obtendo-se $\sum_{i=1}^{25} x_i = 60$. Desenvolva detalhadamente o intervalo de confiança de nível de confiança 99% para a média da população.

7. Determine o tamanho da amostra necessário para se estimar a média de uma população normal com $\sigma = 4,2$ para que, com confiança de 95%, o erro máximo de estimação seja $\pm 0,05$.

8. O peso X de um certo artigo é descrito aproximadamente por uma distribuição normal com $\sigma = 0,58$. Uma amostra de tamanho $n = 25$ resultou em $\overline{x} = 2,8$. Desenvolva detalhadamente o intervalo de confiança de nível de confiança 0,90.

9. De uma população normal com $\sigma = 5$, retira-se uma amostra aleatória simples de tamanho 50, obtendo-se $\overline{x} = 42$.
   (a) Obtenha o intervalo de confiança para a média com nível de significância de 5%.
   (b) Qual é o erro de estimação?
   (c) Para que o erro seja ≤ 1, com probabilidade de acerto de 95%, qual deverá ser o tamanho da amostra?

10. Os valores da venda mensal de determinado artigo têm distribuição aproximadamente normal com desvio padrão de R$500,00. O gerente da loja afirma vender, em média, R$34.700,00. O dono da loja, querendo verificar a veracidade de tal afirmativa, seleciona uma amostra aleatória das vendas em determinado mês, obtendo os seguintes valores:

33840, 00	32960, 00	41815, 00	35060, 00	35050, 00
32940, 00	32115, 00	32740, 00	33590, 00	33010, 00

   (a) Obtenha o intervalo de confiança para a venda média mensal com nível de significância de 5%.
   (b) Obtenha o intervalo de confiança para a venda média mensal com nível de significância de 1%.
   (c) Em qual dos dois níveis de confiança podemos afirmar que o gerente se baseou para fazer a afirmativa?

11. Construa um intervalo de confiança para a proporção populacional para cada um dos casos listados a seguir:

(a) $n = 600 \qquad \alpha = 2\%$
Número de "sucessos" na amostra = 128

(b) $n = 1200 \qquad \alpha = 10\%$
Número de "sucessos" na amostra = 710
Estimativa prévia $\widehat{p}_0 = 55\%$

12. Uma amostra de 300 habitantes de uma grande cidade revelou que 180 desejavam a fluoração da água. Encontre o intervalo de confiança para a verdadeira proporção dos que não desejam a fluoração da água para

    (a) um nível de significância de 5%;

    (b) um nível de confiança de 96%.

13. Querendo estimar a proporção de peças defeituosas em uma linha de produção, examinou-se uma amostra de 100 peças, encontrando-se 32 defeituosas. Sabe-se que o estimador $\widehat{P}$ para esse tamanho de amostra tem desvio padrão de 3%. Calcule o intervalo de confiança ao nível de significância de 3%.

14. Em uma pesquisa de mercado, 57 das 150 pessoas entrevistadas afirmaram que comprariam determinado produto sendo lançado por uma empresa. Essa amostra é suficiente para se estimar a verdadeira proporção de futuros compradores, com uma precisão de 0,08 e uma confiança de 90%? Em caso negativo, calcule o tamanho de amostra necessário.

15. Uma amostra aleatória simples de 400 itens forneceu 100 itens correspondentes ao evento Sucesso.

    (a) Qual é a estimativa pontual $\widehat{p}$ para a verdadeira proporção de Sucessos na população?

    (b) Qual é o erro padrão estimado de $\widehat{p}$?

    (c) Calcule o intervalo de confiança para a verdadeira proporção de Sucessos na população ao nível de confiança de 80%.

16. Em uma sondagem, uma estimativa preliminar de "Sucessos" em uma população é de 0,35. Que tamanho deve ter uma amostra para fornecer um intervalo de confiança de 95% com uma margem de erro de 0,05?

Capítulo 4

Mais sobre intervalos de confiança para parâmetros da $N(\mu; \sigma^2)$

No capítulo anterior, introduzimos o conceito de intervalo de confiança, aplicando-o no contexto pouco realista de uma população normal com variância conhecida. Amostragem de populações normais é um tópico importante, que leva a propriedades interessantes de estatísticas amostrais. Mas é necessário ampliar o contexto inicial, lidando com populações normais com média e variância desconhecidas. Esse é o assunto deste capítulo.

4.1 Amostragem de populações normais

O foco deste capítulo é a estimação da média e da variância de uma população normal. Assim, iremos considerar uma amostra aleatória simples $X_1, X_2, \cdots, X_n$ de uma população $X \sim N(\mu; \sigma^2)$.

4.1.1 A distribuição qui-quadrado

Uma das distribuições que surgirão no estudo de estimadores para os parâmetros da normal é a distribuição qui-quadrado, que é um caso particular da distribuição gama, cuja função densidade é dada a seguir:

$$f_X(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)} x^{\alpha-1} \lambda^\alpha e^{-\lambda x} & \text{, se } x > 0 \\ 0 & \text{, caso contrário.} \end{cases} \tag{4.1}$$

Note que essa densidade depende de dois parâmetros: α, chamado parâmetro de forma, e λ, chamado parâmetro de escala. Com essa parametrização, temos que

$$\mathrm{E}(X) = \frac{\alpha}{\lambda} \qquad \mathrm{Var}\,(X) = \frac{\alpha}{\lambda^2} \qquad M_X(t) = \frac{\lambda^\alpha}{(\lambda - t)^\alpha}, \quad t < \lambda.$$

A distribuição qui-quadrado é um caso particular da distribuição gama, em que o parâmetro de forma α é igual a $\frac{n}{2}$, com n inteiro positivo, e o parâmetro de escala λ é $\frac{1}{2}$. A função densidade resultante é

$$f_X(x) = \frac{1}{\Gamma(n/2)} x^{(n/2)-1} \left(\frac{1}{2}\right)^{n/2} e^{-(1/2)x} = \frac{1}{\Gamma(n/2) 2^{n/2}} x^{(n/2)-1} e^{-x/2}.$$

que depende apenas do parâmetro n, chamado *graus de liberdade* (gl). Usaremos a notação $X \sim \chi^2_n$ para indicar que a variável aleatória X tem distribuição qui-quadrado com n graus de liberdade. Então, se $X \sim \chi^2_n$, sua função densidade é

$$f_X(x) = \begin{cases} \dfrac{1}{\Gamma(n/2)2^{n/2}} x^{(n/2)-1} e^{-x/2} & \text{, se } x > 0 \\ 0 & \text{, caso contrário.} \end{cases} \tag{4.2}$$

e

$$\mathrm{E}(X) = \frac{n/2}{1/2} = n \qquad \mathrm{Var}(X) = \frac{n/2}{(1/2)^2} = 2n \qquad M_X(t) = \frac{1}{(1-2t)^{n/2}}, \quad t < \frac{1}{2}$$

Com relação aos momentos da distribuição qui-quadrado, temos que, se $X \sim \chi^2_n$, então

$$\mathrm{E}(X^k) = \begin{cases} 2^k \dfrac{\Gamma\left(k+\frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} & k > -\dfrac{n}{2} \\ +\infty & k \leq -\dfrac{n}{2} \end{cases}$$

Na Figura 4.1 são apresentados gráficos das distribuições χ^2 com 1, 2 e 3 graus de liberdade.

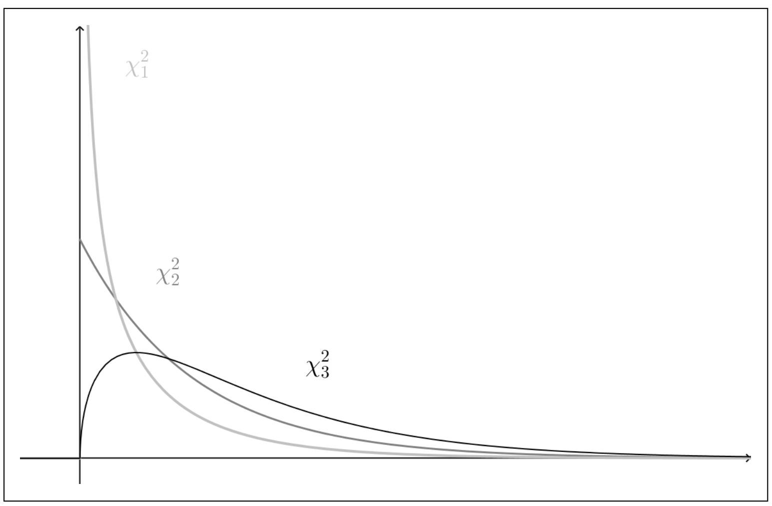


Figura 4.1 – Distribuição qui-quadrado com 1, 2 e 3 graus de liberdade

Outros fatos importantes sobre a distribuição qui-quadrado são dados no lema a seguir.

LEMA 4.1 ***Fatos sobre a distribuição qui-quadrado***

a. Se $Z \sim N(0;1)$, então $Z^2 \sim \chi^2_1$.

b. Se $X_1, X_2, \cdots, X_n$ são variáveis aleatórias independentes com $X_i \sim \chi^2_{k_i}$, então $X_1 + X_2 + \cdots + X_n \sim \chi^2_{k_1+k_2+\cdots+k_n}$.

Demonstração

a. Temos que $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$, $-\infty < z < +\infty$. Seja $Y = Z^2$. Para $y < 0$, $F_Y(y) = 0$ e para $y > 0$

$$\begin{aligned}
F_Y(y) &= \mathrm{P}(Y \le y) = \mathrm{P}(Z^2 \le y) = \mathrm{P}(-\sqrt{y} \le Z \le \sqrt{y}) \underbrace{=}_{\text{cont.}} F_Z(\sqrt{y}) - F_Z(-\sqrt{y}) \Longrightarrow \\
F_Y'(y) &= \frac{1}{2\sqrt{y}} F_Z'(\sqrt{y}) + \frac{1}{2\sqrt{y}} F_Z'(-\sqrt{y}) \Longrightarrow f_Y(y) = \frac{1}{2\sqrt{y}} \left[f_Z(\sqrt{y}) + f_Z(-\sqrt{y}) \right] \Longrightarrow \\
f_Y(y) &= \frac{1}{2\sqrt{y}} \left[\frac{1}{\sqrt{2\pi}} e^{-y/2} + \frac{1}{\sqrt{2\pi}} e^{-y/2} \right] = \frac{1}{\sqrt{2}\sqrt{\pi}} y^{-1/2} e^{-y/2} = \frac{1}{2^{1/2}\Gamma\left(\frac{1}{2}\right)} y^{-1/2} e^{-y/2}
\end{aligned}$$

Comparando essa última expressão com a densidade dada em (4.2), vemos que $Y = Z^2 \sim \chi_1^2$.

b. Seja $W = \sum_{i=1}^{n} X_i$. Então, sua função geradora de momentos é

$$m_Y(t) = \mathrm{E}\left[\exp\left(t \sum_{i=1}^{n} X_i \right) \right] = \mathrm{E}\left(e^{tX_1} e^{tX_2} \cdots e^{tX_n} \right) \underbrace{=}_{\text{indep.}} m_{X_1}(t) m_{X_2}(t) \cdots m_{X_n}(t) = \frac{1}{(1-2t)^{(k_1+k_2+\cdots+k_n)/2}}$$

Essa é a função geradora de momentos de uma distribuição qui-quadrado com $gl = k_1 + k_2 + \cdots + k_n$, o que completa a prova.

■

Na Tabela III do Apêndice A são dados alguns valores críticos (ver Figura 4.2) da distribuição qui-quadrado com graus de liberdade variando de 1 a 35. Assim como no caso da distribuição normal, o valor crítico $\chi^2_{n;\alpha}$ é tal que $\mathrm{P}\left(\chi^2_n \ge \chi^2_{n;\alpha}\right) = \alpha$.

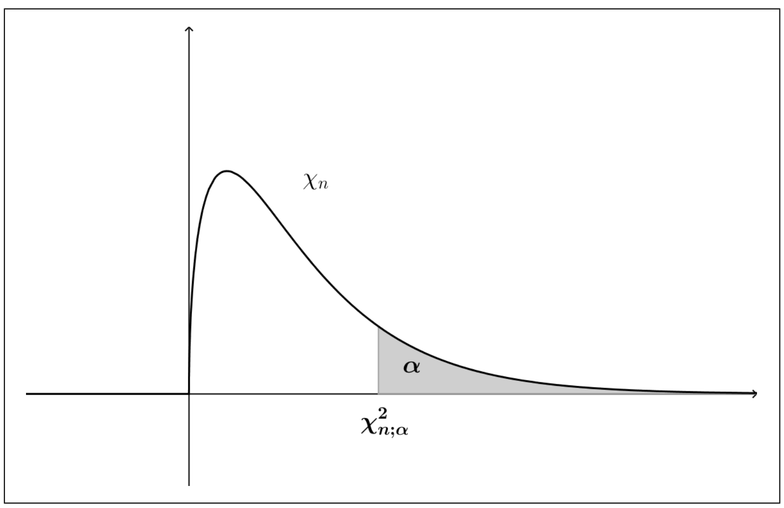


Figura 4.2 – Valor crítico da χ^2_n

4.1.2 A distribuição $t-$Student

Outra distribuição que desempenha papel central na inferência para populações normais é a distribuição t de Student, obtida por William Gosset (1876-1937), que trabalhava na Cervejaria Guinness, na Irlanda. Como a cervejaria não permitia a publicação de resultados de pesquisa obtidos por seus funcionários, Gosset publicou, sob o pseudônimo Student, o artigo "The probable error of a mean", na revista Biometrika, vol. 6, no. 1. Essa distribuição é apresentada no teorema a seguir, cuja demonstração pode ser vista no Apêndice C.

TEOREMA 4.1 ***A distribuição*** t

Sejam Z e Y duas variáveis aleatórias independentes, tais que $Z \sim N(0;1)$ e $Y \sim \chi^2_n$. Então, a variável aleatória

$$T = \frac{Z}{\sqrt{\frac{Y}{n}}} \tag{4.3}$$

tem função densidade dada por

$$f_T(t) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \tag{4.4}$$

chamada densidade t com n graus de liberdade.

▼

Usaremos a seguinte notação para representar uma variável T com distribuição t com n graus de liberdade: $T \sim t_n$.

COROLÁRIO 4.1 ***Média e variância da distribuição*** t

Se $T \sim t_n$, então

$$E(T) = \begin{cases} 0 & n > 1 \\ +\infty & 0 < n \leq 1 \end{cases} \tag{4.5}$$

$$Var(T) = \begin{cases} \dfrac{n}{n-2} & n > 2 \\ +\infty & 0 < n \leq 2 \end{cases} \tag{4.6}$$

Demonstração

Da definição da variável T, resulta que

$$\mathrm{E}(T) = \sqrt{n}\,\mathrm{E}(Z)\,\mathrm{E}\left(Y^{-1/2}\right)$$

Obviamente, $\mathrm{E}(Z) = 0$, mas $\mathrm{E}\left(Y^{-1/2}\right)$ é o momento de ordem $-1/2$ de uma variável qui-quadrado com n graus de liberdade e esse momento será finito apenas se $-\frac{1}{2} > -\frac{n}{2}$, ou equivalentemente, se $n > 1$. Logo, se $0 < n \leq 1$, $\mathrm{E}(T) = +\infty$ e se $n > 1$, $\mathrm{E}(T) = 0$.

Analogamente,

$$\mathrm{E}(T^2) = n\,\mathrm{E}(Z^2)\,\mathrm{E}\left(Y^{-1}\right)$$

Temos que $\mathrm{E}(Z^2) = \mathrm{Var}(Z) = 1$, mas $\mathrm{E}\left(Y^{-1}\right)$ é o momento de ordem -1 de uma variável qui-quadrado com n graus de liberdade e esse momento será finito apenas se $-1 > -\frac{n}{2}$, ou equivalentemente, se $n > 2$. Logo, se $0 < n \leq 2$, $\mathrm{E}(Y^2) = +\infty$ e se $n > 2$,

$$\mathrm{E}(Y^{-1}) = 2^{-1}\frac{\Gamma\left(-1+\frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} = \frac{1}{2}\frac{\Gamma\left(\frac{n}{2}-1\right)}{\left(\frac{n}{2}-1\right)\Gamma\left(\frac{n}{2}-1\right)} = \frac{1}{n-2}$$

e, portanto, se $n > 2$, $\mathrm{E}(T^2) = \mathrm{Var}(T) = \dfrac{n}{n-2}$, o que completa a demonstração.

■

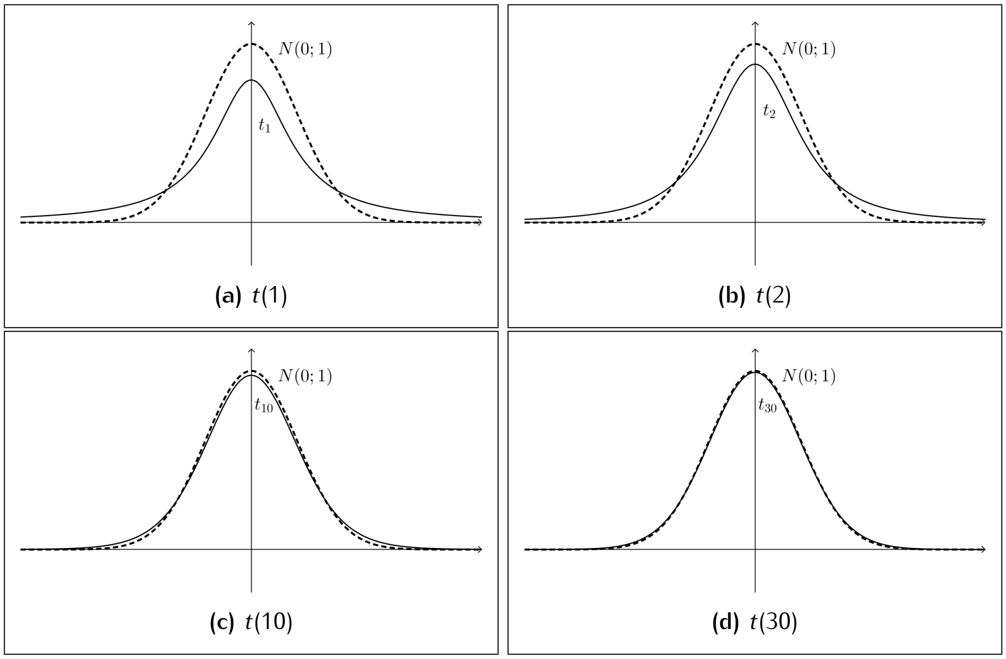


Figura 4.3 – Normal padrão *versus* t-Student

Na Figura 4.3 são apresentados gráficos das distribuições t com 1, 2, 10 e 30 graus de liberdade. A título de comparação, ilustra-se também a distribuição normal padrão. Podemos ver que, à medida que aumenta o número de graus de liberdade, a distribuição t se aproxima da $N(0;1)$.

Na Tabela IV do Apêndice A são dados alguns valores críticos (veja Figura 4.4) da distribuição t com graus de liberdade variando de 1 a 35. Como antes, o valor crítico $t_{n;\alpha}$ é tal que $\mathrm{P}\left(t_n \geq t_{n;\alpha}(n)\right) = \alpha$.

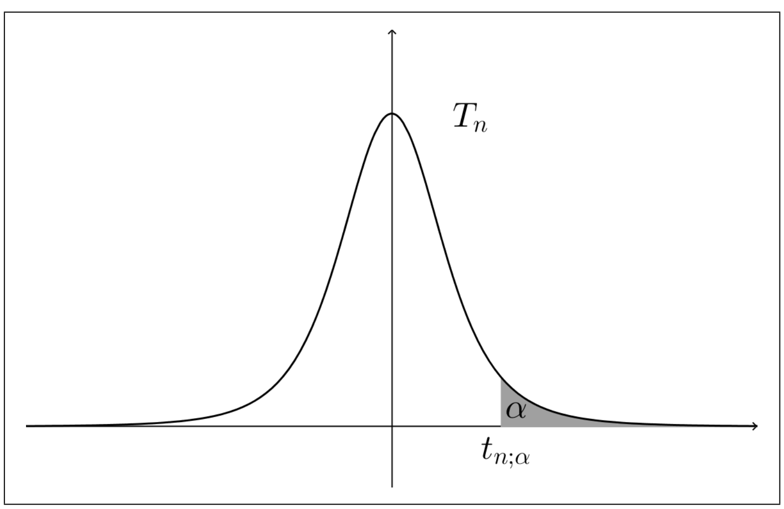


Figura 4.4 – Valor crítico da t_n

4.1.3 Fórmulas recursivas para cálculo da média e da variância amostrais

Vamos estabelecer, agora, fórmulas recursivas para o cálculo de $\overline{X}$ e S^2, ou seja, fórmulas que permitem atualização dos valores a partir do acréscimo à amostra de uma nova observação. Para isso, vamos usar a seguinte notação: $\overline{X}_n$ e S_n^2 representam a média e a variância amostrais calculadas com base em n observações.

$$\begin{aligned}
\overline{X}_{n+1} &= \frac{1}{n+1}\sum_{i=1}^{n+1} X_i = \frac{1}{n+1}\sum_{i=1}^{n} X_i + \frac{1}{n+1}X_{n+1} = \frac{n}{n+1}\frac{1}{n}\sum_{i=1}^{n} X_i + \frac{1}{n+1}X_{n+1} \\
&= \frac{n}{n+1}\overline{X}_n + \frac{1}{n+1}X_{n+1} \Rightarrow \overline{X}_{n+1} = \frac{n\overline{X}_n + X_{n+1}}{n+1} \qquad (4.7)
\end{aligned}$$

$$\overline{X}_{n+1} - \overline{X}_n = \frac{n\overline{X}_n + X_{n+1}}{n+1} - \overline{X}_n = \frac{n\overline{X}_n + X_{n+1} - (n+1)\overline{X}_n}{n+1} = \frac{X_{n+1} - \overline{X}_n}{n+1} \qquad (4.8)$$

$$\begin{aligned}
nS_{n+1}^2 &= \sum_{i=1}^{n+1}\left(X_i - \overline{X}_{n+1}\right)^2 = \sum_{i=1}^{n+1}\left(X_i - \overline{X}_n + \overline{X}_n - \overline{X}_{n+1}\right)^2 \\
&= \sum_{i=1}^{n+1}\left(X_i - \overline{X}_n\right)^2 + \sum_{i=1}^{n+1}\left(\overline{X}_n - \overline{X}_{n+1}\right)^2 + 2\sum_{i=1}^{n+1}\left(X_i - \overline{X}_n\right)\left(\overline{X}_n - \overline{X}_{n+1}\right) \\
&= \sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2 + \left(X_{n+1} - \overline{X}_n\right)^2 + (n+1)\left(\overline{X}_{n+1} - \overline{X}_n\right)^2 - 2\left(\overline{X}_{n+1} - \overline{X}_n\right)\sum_{i=1}^{n+1}\left(X_i - \overline{X}_n\right) \\
&\underbrace{=}_{(4.8)} (n-1)S_n^2 + \left(X_{n+1} - \overline{X}_n\right)^2 + (n+1)\left(\frac{X_{n+1} - \overline{X}_n}{n+1}\right)^2 - 2\frac{X_{n+1} - \overline{X}_n}{n+1}\sum_{i=1}^{n+1}\left(X_i - \overline{X}_n\right) \\
&= (n-1)S_n^2 + \left(X_{n+1} - \overline{X}_n\right)^2 + \frac{\left(X_{n+1} - \overline{X}_n\right)^2}{n+1} - 2\left(\frac{X_{n+1} - \overline{X}_n}{n+1}\right)\left[\cancel{\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)} + (X_{n+1} - \overline{X}_n)\right] \\
&= (n-1)S_n^2 + \left(X_{n+1} - \overline{X}_n\right)^2 - \frac{\left(X_{n+1} - \overline{X}_n\right)^2}{n+1} \\
&= (n-1)S_n^2 + \frac{n}{n+1}\left(X_{n+1} - \overline{X}_n\right)^2 \qquad (4.9)
\end{aligned}$$

4.1.4 Distribuição amostral de S^2

Para demonstrar um teorema que estabelece a distribuição de S^2, apresentada por Stigler (1984), faremos uso do seguinte:

RESULTADO 4.1 ***Fatos sobre a distribuição normal multivariada***

Se U e V são variáveis aleatórias independentes, cada uma com distribuição normal, e se

$$\begin{aligned}
X &= aU + bV \\
Y &= cU + dV
\end{aligned}$$

então (X,Y) tem distribuição normal bivariada e, portanto, se X e Y forem não correlacionadas, elas também serão independentes.

TEOREMA 4.2 ***Distribuição amostral de*** S^2

Seja $X_1, X_2, \cdots X_n$ *uma amostra aleatória simples de uma população* $X \sim N(\mu; \sigma^2)$.

(a) $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$;

(b) $\overline{X}$ *e* S^2 *são independentes.*

Demonstração

A demonstração será feita por indução em n, o tamanho da amostra, com uso do Teorema 2.2 que estabelece que $\overline{X} \sim N\left(\mu; \dfrac{\sigma^2}{n}\right)$. Usaremos, mais uma vez, a notação $\overline{X}_n$ e S^2_n definida anteriormente.

- $n = 2$

$$
\begin{aligned}
S_2^2 &= \frac{1}{2-1}\sum_{i=1}^{2}\left(X_i - \overline{X}_2\right)^2 = \left(X_1 - \frac{X_1+X_2}{2}\right)^2 + \left(X_2 - \frac{X_1+X_2}{2}\right)^2 = \left(\frac{X_1-X_2}{2}\right)^2 + \left(\frac{X_2-X_1}{2}\right)^2 \Rightarrow \\
S_2^2 &= \frac{(X_1-X_2)^2}{2} \qquad (4.10)
\end{aligned}
$$

X_1 e X_2 são independentes, cada uma com distribuição $N(\mu; \sigma^2)$; logo,

$$
X_1 - X_2 \sim N\left(0; 2\sigma^2\right) \Rightarrow \frac{X_1 - X_2}{\sqrt{2}\sigma} \sim N(0;1) \Rightarrow \frac{(X_1-X_2)^2}{2\sigma^2} \sim \chi^2_1 \Rightarrow (2-1)\frac{S_2^2}{\sigma^2} \sim \chi^2_1
$$

e isso prova a parte (a) para $n = 2$.

X_1 e X_2 são variáveis normais independentes; logo, pelo Resultado 4.1, $X_1 + X_2$ e $X_1 - X_2$ têm distribuição normal bivariada. Mas

$$
\mathrm{Cov}(X_1 - X_2, X_1 + X_2) = \mathrm{Var}(X_1) - \mathrm{Var}(X_2) = 0
$$

Logo, $X_1 + X_2$ e $X_1 - X_2$ são independentes o que prova que $\overline{X}_2 = \dfrac{X_1+X_2}{2}$ e $S_2^2 = \dfrac{(X_1-X_2)^2}{2}$ são independentes.

- Provamos que o teorema vale para $n = 2$. Suponhamos, agora,que ele seja válido para n, ou seja,

$$
(n-1)\frac{S_n^2}{\sigma^2} \sim \chi^2_{n-1} \qquad (4.11)
$$

$$
\overline{X}_n \text{ e } S_n^2 \text{ são independentes} \qquad (4.12)
$$

Vamos provar que é válido para $n + 1$.

- $n + 1$

Usando o resultado dado em (4.9), obtemos que

$$
n\frac{S_{n+1}^2}{\sigma^2} = (n-1)\frac{S_n^2}{\sigma^2} + \frac{n}{n+1}\left(\frac{X_{n+1} - \overline{X}_n}{\sigma}\right)^2
$$

⋆ X_{n+1} é independente de S_n^2, que só depende de $X_1, \cdots, X_n$

⋆ $\overline{X}_n$ é independente de S_n^2 pela hipótese de indução (4.12)

⋆ Logo,

$$(n-1)\frac{S_n^2}{\sigma^2} \text{ é independente de } \frac{n}{n+1}\left(\frac{X_{n+1}-\overline{X}_n}{\sigma}\right)^2 \tag{4.13}$$

$$\overline{X}_{n+1} = \frac{n\overline{X}_n + X_{n+1}}{n+1} \text{ é independente de } S_n^2 \tag{4.14}$$

⋆

$$X_{n+1}-\overline{X}_n \sim N\left(0;\sigma^2+\frac{\sigma^2}{n}\right) \Rightarrow \frac{n}{n+1}\left(\frac{X_{n+1}-\overline{X}_n}{\sigma}\right)^2 \sim \chi^2(1) \tag{4.15}$$

$$(n-1)\frac{S_n^2}{\sigma^2} \sim \chi^2(n-1) \text{ pela hipótese de indução (4.12} \tag{4.16}$$

⋆ Logo, $n\dfrac{S_{n+1}^2}{\sigma^2}$ é a soma de duas variáveis qui-quadrado independentes e, pelo Lema 4.1, resulta que

$$n\frac{S_{n+1}^2}{\sigma^2} \sim \chi^2_{n-1+1}$$

o que prova a parte (a) do teorema para $n+1$.

⋆ X_{n+1} e $\overline{X}_n$ são variáveis aleatórias normais independentes. Logo, $X_{n+1}-\overline{X}_n$ e $X_{n+1}+n\overline{X}_n$ têm distribuição normal bivariada pelo Resultado 4.1. Mas

$$\text{Cov}\left(X_{n+1}-\overline{X}_n, X_{n+1}+n\overline{X}_n\right) = \text{Var}(X_{n+1}) - n\,\text{Var}(\overline{X}_n) = 0$$

o que significa que $X_{n+1}-\overline{X}_n$ e $X_{n+1}+n\overline{X}_n$ são independentes, ou seja, $X_{n+1}-\overline{X}_n$ é independente de $\overline{X}_{n+1}$.

⋆ Temos, então:

$$S_{n+1}^2 = \frac{n-1}{n}S_n^2 + \frac{1}{n+1}\left(X_{n+1}-\overline{X}_n\right)^2$$

$$S_n^2 \text{ é independente de } \overline{X}_{n+1}$$

$$X_{n+1}-\overline{X}_n \text{ é independente de } \overline{X}_{n+1}$$

e, portanto, S_{n+1}^2 é independente de X_{n+1}, o que completa a demonstração. ■

4.1.5 Distribuição de $\overline{X}$

Se $X_1, X_2, \cdots, X_n$ é uma amostra aleatória simples de uma população $X \sim N(\mu;\sigma^2)$, segue, dos Teoremas 2.2 e 4.2, que

$$\sqrt{n}\frac{\overline{X}-\mu}{\sigma} \sim N(0;1)$$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$\overline{X} \text{ e } S^2 \text{ são independentes}$$

Logo, pelo Teorema 4.1,

$$\frac{\sqrt{n}\dfrac{\overline{X}-\mu}{\sigma}}{\sqrt{\dfrac{\dfrac{(n-1)S^2}{\sigma^2}}{n-1}}} \sim t_{n-1}$$

ou, equivalentemente,

$$\sqrt{n}\frac{\overline{X}-\mu}{S} \sim t_{n-1} \tag{4.17}$$

4.2 Intervalo de confiança para a variância σ^2

Na Figura 4.5 temos ilustrados os valores críticos da χ^2_{n-1} que deixam probabilidade $\alpha/2$ em cada cauda, ou seja, para uma variável aleatória χ^2_{n-1}, temos

$$P\left(\chi^2_{n-1;1-\alpha/2} \leq \chi^2_{n-1} \leq \chi^2_{n-1;\alpha/2}\right) = 1-\alpha$$

Embora a distribuição não seja simétrica, vamos considerar probabilidades iguais nas duas caudas.

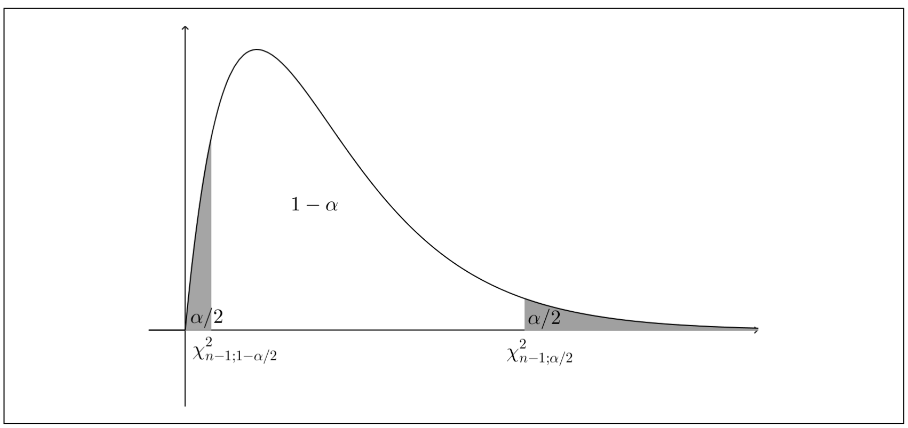


Figura 4.5 – Valores críticos da χ^2_n para intervalos de confiança

Mas, pelo Teorema 4.2, sabemos que $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$. Logo,

$$\begin{aligned}
P\left(\chi^2_{n-1;1-\alpha/2} \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{n-1;\alpha/2}\right) &= 1-\alpha \Rightarrow \\
P\left(\frac{\chi^2_{n-1;1-\alpha/2}}{(n-1)S^2} \leq \frac{1}{\sigma^2} \leq \frac{\chi^2_{n-1;\alpha/2}}{(n-1)S^2}\right) &= 1-\alpha \Rightarrow \\
P\left(\frac{(n-1)S^2}{\chi^2_{n-1;\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{n-1;1-\alpha/2}}\right) &= 1-\alpha
\end{aligned}$$

Nessa última expressão, temos um intervalo cujos limites são variáveis aleatórias, pois dependem da amostra sorteada. Além disso, esses limites não dependem do parâmetro σ^2. Assim, essa última expressão nos fornece um intervalo de confiança para a variância de uma população normal com nível de confiança $1 - \alpha$.

! Intervalo de Confiança para a Variância da $N(\mu; \sigma^2)$

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória simples de uma população $X \sim N\left(\mu; \sigma^2\right)$. O intervalo de confiança para σ^2 de nível de confiança $1 - \alpha$ é

$$\left[\frac{(n-1)S^2}{\chi^2_{n-1;\alpha/2}} \,;\, \frac{(n-1)S^2}{\chi^2_{n-1;1-\alpha/2}}\right] \tag{4.18}$$

em que $\chi^2_{n;\,\alpha}$ é o valor crítico da distribuição qui-quadrado com n graus de liberdade que deixa área α acima dele.

A interpretação do intervalo de confiança para σ^2 é a mesma: se construíssemos os intervalos de confiança pela fórmula 4.18 para todas as possíveis amostras, $100 \times (1 - \alpha)\%$ dos intervalos conteriam o verdadeiro valor de σ^2. Na Figura 4.6 ilustra-se essa interpretação dos intervalos de confiança para a variância de uma população normal com amostras de tamanho $n = 17$. A distribuição qui-quadrado com 16 gl representa a distribuição de probabilidade dos valores de $(n-1)S^2/\sigma^2$. Valores extremos de tal estatística levam a intervalos de confiança que não contêm o verdadeiro parâmetro, representados pelos intervalos em preto. Os valores centrais, que têm alta probabilidade $(1 - \alpha)$ de ocorrência, levam a intervalos que contêm o verdadeiro valor do parâmetro (intervalos em cinza).

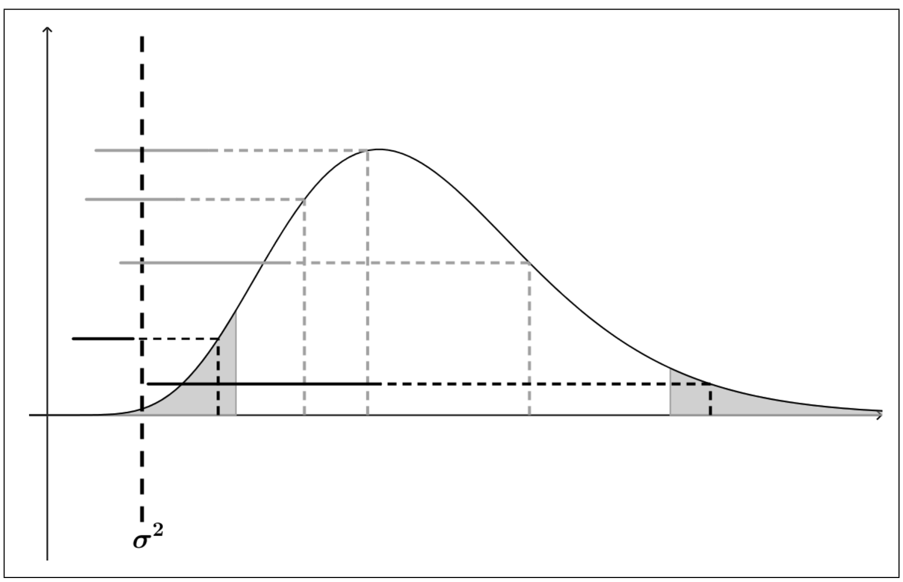


Figura 4.6 – Interpretação do IC para a variância da $N(\mu; \sigma^2)$

Observe que os comprimentos dos intervalos não são constantes, como no caso dos intervalos de confiança para a média de uma normal com variância conhecida. Aqui, o comprimento do intervalo é

$$\delta = (n-1)S^2 \left(\frac{1}{\chi^2_{n-1;1-\alpha/2}} - \frac{1}{\chi^2_{n-1;\alpha/2}}\right)$$

ou seja, quanto maior a variância amostral, maior o comprimento do intervalo de confiança baseado numa amostra de tamanho n e nível de confiança $1 - \alpha$.

EXEMPLO 4.1

De uma população normal com média e variância desconhecidas, extrai-se uma amostra de tamanho 15 obtendo-se $\overline{x} = 12$ e $s^2 = 49$. Obtenha um intervalo de confiança para a variância populacional, utilizando o nível de confiança de 95%.

Solução

O requisito para o IC para σ^2 é satisfeito, uma vez que a população é normal. Temos que usar a distribuição χ^2 com $n-1 = 14$ graus de liberdade. Como o nível de confiança é de 95%, em cada cauda da distribuição temos que ter 2,5%. Assim, para a cauda superior, devemos usar o valor crítico $\chi^2_{14;0,025}$ procurando na linha correspondente a 14 graus de liberdade e na coluna correspondente à probabilidade de 0,025. Encontramos que $\chi^2_{14;0,025} = 26,119$.

Para a cauda inferior, devemos usar o valor crítico $\chi^2_{14;0,975}$, procurando na linha correspondente a 14 graus de liberdade e na coluna correspondente à probabilidade de 0,975. Encontramos que $\chi^2_{14;0,975} = 5,629$. Logo, o intervalo de confiança é

$$\left[\frac{14\times 49}{26,119};\frac{14\times 49}{5,629}\right] = [26,26;121,87]$$

♦♦

4.3 Intervalo de confiança para a média μ

Na Figura 4.7 temos ilustrados os valores críticos da t_{n-1} que deixam probabilidade $\alpha/2$ em cada cauda, ou seja, para uma variável aleatória t_{n-1}, temos

$$\mathrm{P}\left(-t_{n-1;\alpha/2} \leq t_{n-1} \leq t_{n-1;\alpha/2}\right) = 1-\alpha$$

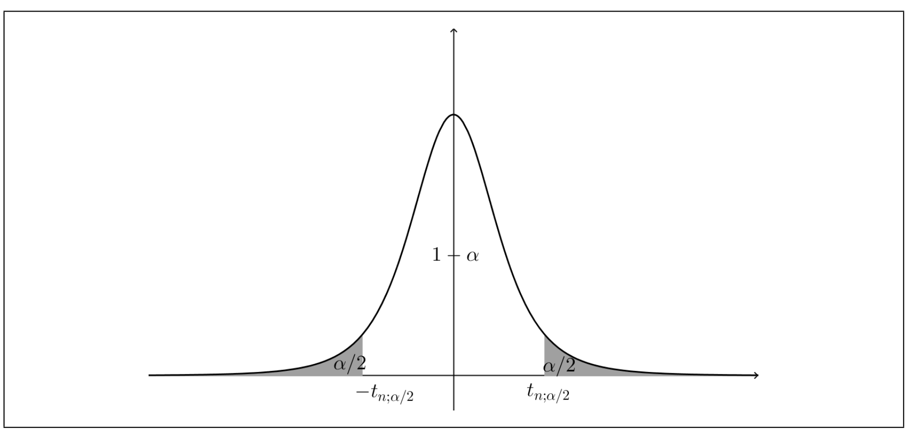


Figura 4.7 – Valores críticos da χ^2_n para intervalos de confiança

Mas, pelo resultado dado em (4.1.5), sabemos que $\sqrt{n}\dfrac{\overline{X}-\mu}{S} \sim t_{n-1}$. Logo,

$$
\begin{aligned}
&\mathrm{P}\left(-t_{n-1;\,\alpha/2} \le \sqrt{n}\frac{\overline{X}-\mu}{S} \le t_{n-1;\,\alpha/2}\right) = 1-\alpha \Longrightarrow \\
&\mathrm{P}\left(-t_{n-1;\,\alpha/2}\frac{S}{\sqrt{n}} \le \overline{X}-\mu \le t_{n-1;\,\alpha/2}\frac{S}{\sqrt{n}}\right) = 1-\alpha \Longrightarrow \\
&\mathrm{P}\left(\overline{X} - t_{n-1;\,\alpha/2}\frac{S}{\sqrt{n}} \le \mu \le \overline{X} + t_{n-1;\,\alpha/2}\frac{S}{\sqrt{n}}\right) = 1-\alpha
\end{aligned}
$$

Essa última expressão é o intervalo de confiança para a média μ de uma população normal com variância desconhecida.

> **! Intervalo de Confiança para a Média da $N(\mu;\sigma^2)$ – σ^2 Desconhecida**
>
> Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória simples de uma população $X \sim N\left(\mu;\sigma^2\right)$.
> O intervalo de confiança para μ de nível de confiança $1-\alpha$ é
>
> $$\left[\overline{X} - t_{n-1;\,\alpha/2}\frac{S}{\sqrt{n}}\,;\,\overline{X} + t_{n-1;\,\alpha/2}\frac{S}{\sqrt{n}}\right]$$
>
> onde $t_{n-1;\,\alpha/2}$ é o valor crítico da distribuição t-Student com $n-1$ graus de liberdade que deixa área $\alpha/2$ acima dele.

4.3.1 Margem de erro

Note, mais uma vez, a forma do intervalo de confiança:

$$\overline{X} \pm \epsilon$$

onde a margem de erro ϵ, agora, é definida em termos do valor crítico da distribuição t e do erro- padrão estimado de $\overline{X}$:

$$\epsilon = t_{n-1;\,\alpha/2}\frac{S}{\sqrt{n}} \doteq t_{n-1;\,\alpha/2}\widehat{EP}(\overline{X}) \tag{4.19}$$

onde

$$\widehat{EP}(\overline{X}) = \frac{S}{\sqrt{n}} \tag{4.20}$$

4.3.2 Amostras grandes

Vimos que, para populações normais, a distribuição exata da estatística $T = \sqrt{n}\dfrac{\overline{X}-\mu}{S}$ é $t(n-1)$. Mas vimos também que, quando o número de graus de liberdade é grande, a diferença entre as distribuições t e $N(0;1)$ torna-se desprezível.

Por outro lado, se a população não é normal, mas tem média μ e variância σ^2, o Teorema Limite Central nos diz que a distribuição de $\sqrt{n}\dfrac{\overline{X}-\mu}{\sigma}$ se aproxima de uma $N(0;1)$ à medida que n aumenta. Pode-se mostrar que esse resultado continua valendo se substituirmos σ por seu estimador S.

A conclusão dessas duas observações é a seguinte:

! Intervalo de confiança para a média baseado em grandes amostras

Dada uma amostra aleatória simples $X_1, X_2, \ldots, X_n$ de uma população X com média μ e variância σ^2, então

$$\sqrt{n}\frac{\overline{X}-\mu}{S} \approx N(0;1)$$

para n suficientemente grande. Nesse caso, o intervalo de confiança aproximado de nível de confiança $1-\alpha$ para μ é

$$\left[\overline{X} - z_{\alpha/2}\frac{S}{\sqrt{n}}\,;\,\overline{X} + z_{\alpha/2}\frac{S}{\sqrt{n}}\right]$$

EXEMPLO 4.2

De uma população normal com média e variância desconhecidas, extrai-se uma amostra de tamanho 15 obtendo-se $\overline{x} = 12$ e $s^2 = 49$. Obtenha um intervalo de confiança para a verdadeira média populacional, utilizando o nível de confiança de 95%.

Solução

Os seguintes requisitos para o IC para μ são satisfeitos: a população é normal e a amostra é pequena. Dessa forma, temos que usar a distribuição t com $n-1 = 14$ graus de liberdade. Como o nível de confiança é de 95%, em cada cauda da distribuição temos que ter 2,5%. Assim, devemos procurar a abscissa $t_{14;0,025}$ procurando na linha correspondente a 14 graus de liberdade e na coluna correspondente à área de 0,025. Encontramos

$$t_{14;0,025} = 2,145$$

A margem de erro é

$$\epsilon = 2,145 \times \frac{7}{\sqrt{15}} = 3,8769$$

e o intervalo de confiança, $[12-3,8769; 12+3,8769] = [8,1231; 15,8769]$

♦♦

EXEMPLO 4.3

A seguinte amostra foi extraída de uma população normal: $6, 6, 7, 8, 9, 9, 10, 11, 12$. Construa o intervalo de confiança para a média populacional, com nível de confiança de 90%.

Solução

Como antes, temos uma amostra pequena de uma população normal; logo, temos que usar a distribuição t-Student. Como $n = 9$, $gl = n-1 = 8$.

A média amostral é

$$\begin{aligned}\overline{x} &= \frac{\sum x_i}{n} \\ &= \frac{6+6+7+8+9+9+10+11+12}{9} = \frac{78}{9} = 8,667\end{aligned}$$

e a variância amostral é

$$\begin{aligned} S^2 &= \frac{1}{n-1}\sum(x_i-\overline{x})^2 = \frac{1}{n-1}\left[\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right] = \\ &= \frac{1}{8}\left[6^2+6^2+7^2+8^2+9^2+9^2+10^2+11^2+12^2-\frac{78^2}{9}\right] \\ &= \frac{1}{8}\left[712-\frac{6084}{9}\right] = \frac{36}{8} = 4,5 \end{aligned}$$

Como o nível de significância é $\alpha = 10\%$, o nível de confiança é $1-\alpha = 90\%$. Em cada cauda da distribuição $t(8)$ temos que ter área igual a 5%. Assim, temos que procurar na linha correspondente a 8 graus de liberdade a abscissa relativa à área superior de 0,05. Obtemos $t_{8;0,05} = 1,860$. A margem de erro é

$$\epsilon = 1,860 \times \sqrt{\frac{4,5}{9}} = 1,315$$

e o intervalo de confiança é $[8,667-1,315; 8,667+1,315] = [7,352; 9,982]$

♦♦

EXEMPLO 4.4

A partir de uma amostra aleatória simples de tamanho $n = 100$, os seguintes valores foram obtidos: $\overline{x} = 12,36$ e $S^2 = 132,56$. Obtenha um intervalo de confiança de nível de confiança 90% para a média populacional μ.

Solução

Como o tamanho amostral é grande, podemos usar a aproximação normal. Como $1-\alpha = 0,90$, em cada cauda temos que ter 5% e, assim, devemos procurar no corpo da tabela da distribuição normal o valor mais próximo de 0,45. Resulta que $z_{0,05} = 1,64$, o que nos dá a seguinte margem de erro:

$$\epsilon = 1.64 \times \sqrt{\frac{132.56}{100}} = 1,8882$$

O intervalo de confiança de 90% de confiança é $[12.36-1.8882\,;\,12.36+1.8882] = [10.472\,;\,14.248]$

♦♦

4.4 Exercícios propostos

1. Seja $X \sim \chi^2_{17}$ uma variável aleatória com distribuição qui-quadrado com 17 graus de liberdade. Encontre o valor da abscissa k tal que:

   (a) $\mathrm{P}(X > k) = 0,02$

   (b) $\mathrm{P}(X < k) = 0,02$

   (c) $\mathrm{P}(X < k) = 0,90$

2. Para uma distribuição t de Student com 12 graus de liberdade, encontre a probabilidade de cada uma das seguintes regiões (esboce um gráfico para auxiliar na solução do exercício):

   (a) à esquerda de $1,782$;

   (b) à direita de $-1,356$;

   (c) à direita de $2,681$;

   (d) entre $1,083$ e $3,055$;

(e) entre $-1,356$ e $2,179$.

3. Encontre os seguintes valores críticos da distribuição t de Student:

   (a) $t_{15;0,05}$
   (b) $t_{18;0,90}$
   (c) $t_{25;0,975}$

4. Os tempos gastos por quinze funcionários em uma das tarefas de um programa de treinamento estão listados abaixo. é razoável supor, nesse caso, que essa seja uma amostra aleatória simples de uma população normal, ou seja, é razoável supor que a população de todos os tempos de funcionários submetidos a esse treinamento seja aproximadamente normal. Obtenha o intervalo de confiança de nível de confiança de 95% para

   (a) o tempo médio populacional;
   (b) a variância populacional.

52	44	55	44	45	59	50	54
62	46	54	58	60	62	63	

5. Uma amostra aleatória simples de uma população normal apresenta as seguintes características:

$$n = 25 \qquad \overline{x} = 500 \qquad s^2 = 900$$

   Construa um intervalo de confiança de nível de confiança de 98% para

   (a) a média populacional;
   (b) a variância populacional.

6. Em uma fábrica, uma amostra de 30 parafusos apresentou os seguintes diâmetros (em mm):

10	13	14	11	13	14	11	13	14	15
12	14	15	13	14	12	12	11	15	16
13	15	14	14	15	15	16	12	10	15

   Supondo que os diâmetros possam ser descritos por uma variável normal, obtenha um intervalo de confiança

   (a) para o diâmetro médio de todos os parafusos produzidos nessa fábrica, usando o nível de confiança de 98%;
   (b) para a variância do diâmetro de todos os parafusos produzidos nessa fábrica.

   Obs.: Para facilitar a solução do exercício, você pode usar os seguintes resultados:

$$\sum_{i=1}^{30} x_i = 401 \qquad \sum_{i=1}^{30} x_i^2 = 5443$$

7. Repita o exercício anterior, supondo que $n = 100$.

Capítulo 5

Testes de hipóteses – Conceitos básicos

5.1 Introdução

Na teoria de estimação, vimos que é possível, por meio de estatísticas amostrais adequadas, estimar parâmetros de uma população, dentro de um certo intervalo de confiança.

Nos testes de hipóteses, em vez de se construir um intervalo de confiança no qual se espera que o parâmetro da população esteja contido, testa-se a validade de uma afirmação sobre um parâmetro da população. Então, em um teste de hipótese, procura-se tomar decisões a respeito de uma população com base em informações obtidas de uma amostra dessa população.

Vamos trabalhar com alguns exemplos para ilustrar os conceitos básicos necessários para a construção de testes de hipóteses estatísticos.

EXEMPLO 5.1 Amostra de anéis de vedação - parte 1

Uma empresa compra anéis de vedação de dois fabricantes. Segundo informações dos fabricantes, os anéis do fabricante 1 têm diâmetro médio de 14 cm com desvio padrão de 1,2 cm e os anéis do fabricante 2 têm diâmetro médio de 15 cm com desvio padrão de 2,0 cm. Ambos os processos de produção geram anéis com diâmetros cuja distribuição é aproximadamente normal.

Uma caixa com 16 anéis sem identificação é encontrada pelo gerente do almoxarifado. Embora ele suspeite que a caixa seja oriunda do fabricante 1, decide fazer uma medição dos anéis e basear sua decisão no diâmetro médio da amostra: se o diâmetro médio for maior que 14,5 cm, ele identificará a caixa como oriunda do fabricante 2; caso contrário, ele identificará a caixa como oriunda do fabricante 1.

Esse é um problema típico de decisão empresarial. Vamos analisá-lo sob o ponto de vista estatístico, estudando os possíveis erros e suas probabilidades de ocorrência. Para isso, precisamos formular uma *hipótese nula*, que é a afirmação de interesse sobre um parâmetro da população e uma *hipótese alternativa*, que é a hipótese que será considerada, caso haja evidências contra a hipótese nula.

A hipótese nula, normalmente designada por H_0, é uma afirmação que é estabelecida com o objetivo de ser testada; ela pode ser rejeitada ou não e a regra de decisão será baseada nessa hipótese nula. Geralmente, a hipótese nula é formulada de tal forma que o objetivo é rejeitá-la.[1]

Neste exemplo, existem apenas duas possibilidades para a origem dos anéis de vedação. Como o gerente suspeita que a caixa venha do fabricante 1, vamos estabelecer a hipótese nula de forma que o

[1] Mais adiante, veremos um procedimento objetivo para estabelecimento das hipóteses nula e alternativa em contextos mais complexos.

resultado desejado seja rejeitá-la. Definimos, então, a hipótese nula como sendo

$$H_0 : \text{anéis vêm do fabricante 2}$$

e, como só há dois fabricantes, a hipótese alternativa será

$$H_1 : \text{anéis vêm do fabricante 1}$$

Se denotamos por X a variável aleatória que representa o diâmetro dos anéis, essas hipóteses se traduzem como

$$\begin{aligned} H_0 &: X \sim N(15; 2,0^2) \\ H_1 &: X \sim N(14; 1,2^2) \end{aligned}$$

A regra de decisão do gerente é baseada na média amostral observada para os 16 anéis encontrados. Como dito, nossa decisão deve ser expressa sempre em termos de H_0. Logo, a regra de decisão é

$$\begin{aligned} \overline{X} &\leq 14,5 \Longrightarrow \text{rejeito } H_0 \\ \overline{X} &> 14,5 \Longrightarrow \text{não rejeito } H_0 \end{aligned}$$

Os erros associados a essa regra de decisão são:

Erro I: rejeitar H_0 quando H_0 é verdadeira
Erro II: não rejeitar H_0 quando H_0 é falsa

Se H_0 é verdadeira, a amostra vem de uma população normal com média 15 e desvio padrão 2,0. Nesse caso, a média amostral com base em uma amostra de tamanho 16 é também normal com média 15 e desvio padrão $\frac{2,0}{\sqrt{16}}$.

Se H_0 é falsa, a amostra vem de uma população normal com média 14 e desvio padrão 1,2. Nesse caso, a média amostral com base em amostra de tamanho 16 é também normal com média 14 e desvio padrão $\frac{1,2}{\sqrt{16}}$.

Então, as probabilidades associadas aos erros podem ser expressas em termos de probabilidade condicional:

$$\begin{aligned} \text{P(Erro I)} &= \text{P}\left[\overline{X} \leq 14,5 | \overline{X} \sim N\left(15; \frac{2,0^2}{16}\right)\right] \\ \text{P(Erro II)} &= \text{P}\left[\overline{X} > 14,5 | \overline{X} \sim N\left(14; \frac{1,2^2}{16}\right)\right] \end{aligned}$$

Na Figura 5.1, a probabilidade associada ao erro I corresponde à área sombreada de cinza escuro, enquanto a área sombreada de cinza claro corresponde à probabilidade do erro tipo II.

Vamos calcular essas probabilidades. Em geral, a probabilidade do erro tipo I é denotada por α e a probabilidade do erro tipo II por β. Assim,

$$\begin{aligned} \alpha &= \text{P(Erro I)} = \text{P}\left[\overline{X} \leq 14,5 | \overline{X} \sim N\left(15; \frac{2,0^2}{16}\right)\right] = \text{P}\left(Z \leq \frac{14,5-15}{\frac{2}{4}}\right) \\ &= \text{P}(Z \leq -1,00) = \text{P}(Z \geq 1,00) = 0,5 - tab(1,00) = 0,5 - 0,34134 = 0,15866 \end{aligned}$$

$$\begin{aligned} \beta &= \text{P(Erro II)} = \text{P}\left[\overline{X} > 14,5 | \overline{X} \sim N\left(14; \frac{1,2^2}{16}\right)\right] = \text{P}\left(Z > \frac{14,5-14}{\frac{1.2}{4}}\right) \\ &= \text{P}(Z > 1,67) = 0,5 - tab(1,67) = 0,04746 \end{aligned}$$

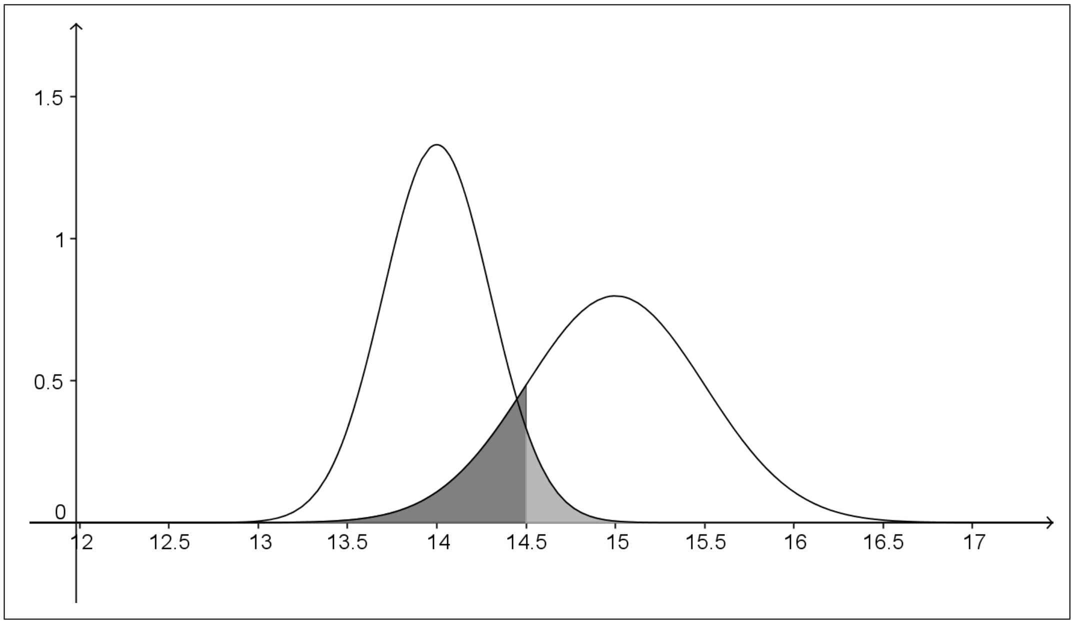


Figura 5.1 – Probabilidades dos Erros tipo I e II para o Exemplo 5.1

♦♦

É importante que se entenda a sutileza da notação, lembrando que letras maiúsculas são usadas para representar variáveis aleatórias e letras minúsculas para representar o valor observado de uma variável aleatória. Quando falamos da probabilidade do erro ou mesmo da regra de decisão em termos gerais, estamos considerando o procedimento decisório geral. Como esse procedimento depende da amostra sorteada, temos de expressar as probabilidades dos erros e a regra de decisão levando em conta as possíveis amostras, ou seja, temos de levar em conta a variável aleatória $\overline{X}$ que descreve a média amostral de uma possível amostra aleatória simples de tamanho n. A decisão do gerente será tomada em função do resultado amostral observado de uma amostra específica.

No exemplo, a regra de decisão geral é: se $\overline{X} > 14,5$, o gerente classifica como produção do fabricante 2. Assim, se a caixa em questão tiver uma média de, por exemplo, 14,4, o gerente classificará a caixa como produzida pelo fabricante 1.

EXEMPLO 5.2 Amostra de anéis de vedação - parte 2

Para resumir os resultados do exemplo anterior, podemos construir o seguinte quadro:

		Decisão do Gerente	
		Fabricante 1 (H_1)	Fabricante 2 (H_0)
Fabricante	1 (H_1)	OK	Erro II ($\beta = 0,04746$)
Verdadeiro	2 (H_0)	Erro I ($\alpha = 0,15866$)	OK

Vemos aí que a probabilidade do erro tipo I é maior. Analisando a **Figura 5.1**, podemos ver também que, se mudarmos a regra de decisão escolhendo um valor de corte diferente de 14,5, essas probabilidades se alterarão. Aumentando α, diminui β e vice-versa.

Vamos, agora, estabelecer uma nova regra de decisão de modo que a probabilidade do erro tipo I passe a ser 0,05. A nossa região de rejeição, ou *região crítica*, continua tendo a forma $\overline{X} \leq k$. Pela **Figura**

5.1, vemos que k tem de ser menor que 14,5. O que queremos é

$$
\begin{aligned}
&\alpha = 0,05 \Leftrightarrow \mathrm{P}\left[\overline{X} \leq k \,|\, \overline{X} \sim N\left(15; \frac{2,0^2}{16}\right)\right] = 0,05 \Leftrightarrow \\
&\mathrm{P}\left(Z \leq \frac{k-15}{\frac{2}{4}}\right) = 0,05 \Leftrightarrow \mathrm{P}\left(Z \geq -\frac{k-15}{0,5}\right) = 0,05 \Leftrightarrow \\
&0,5 - tab\left(-\frac{k-15}{0,5}\right) = 0,05 \Leftrightarrow tab\left(-\frac{k-15}{0,5}\right) = 0,45 \Leftrightarrow \\
&-\frac{k-15}{0,5} = 1,64 \Leftrightarrow k = 14,18
\end{aligned}
$$

Com essa nova regra, rejeitamos H_0 se $\overline{X} \leq 14,18$ e o erro tipo II passa a ter probabilidade

$$
\begin{aligned}
\beta &= \mathrm{P(Erro\ II)} = \mathrm{P}\left[\overline{X} > 14,18 | \overline{X} \sim N\left(14; \frac{1,2^2}{16}\right)\right] \\
&= \mathrm{P}\left(Z > \frac{14,18-14}{\frac{1,2}{4}}\right) = \mathrm{P}(Z > 0,6) \\
&= 0,5 - tab(0,6) = 0,27425
\end{aligned}
$$

♦♦

EXEMPLO 5.3 Amostra de anéis de vedação - parte 3

Suponha, agora, que o gerente queira igualar as probabilidades de erro. Qual deve ser a regra de decisão?

$$
\begin{aligned}
&\alpha = \beta \Leftrightarrow \\
&\mathrm{P}\left[\overline{X} \leq k \mid \overline{X} \sim N\left(15; \frac{2,0^2}{16}\right)\right] = \mathrm{P}\left[\overline{X} > k \,|\, \overline{X} \sim N\left(14; \frac{1,2^2}{16}\right)\right] \Leftrightarrow \\
&\mathrm{P}\left(Z \leq \frac{k-15}{\frac{2.0}{4}}\right) = \mathrm{P}\left(Z > \frac{k-14}{\frac{1.2}{4}}\right) \Longleftrightarrow \frac{k-15}{0,5} = -\frac{k-14}{0,3} \Leftrightarrow \\
&0,3k - 4,5 = -0,5k + 7 \Leftrightarrow 0,8k = 11,5 \Leftrightarrow k = 14,375
\end{aligned}
$$

Neste caso, as probabilidades dos erros tipo I e II são

$$
\begin{aligned}
\alpha &= \beta = \mathrm{P}\left[\overline{X} \leq 14,375 \,|\, \overline{X} \sim N\left(15; \frac{2,0^2}{16}\right)\right] \\
&= \mathrm{P}\left(Z \leq \frac{14,375-15}{0,5}\right) \\
&= \mathrm{P}(Z \leq -1,25) = \mathrm{P}(Z \geq 1,25) = 0,5 - tab(1,25) = 0,10565
\end{aligned}
$$

♦♦

EXEMPLO 5.4 Amostra de anéis de vedação - parte 4

O procedimento de se fixar a probabilidade α do erro tipo I é o mais utilizado pois, em geral, na prática a situação não é tão simples como a escolha entre duas decisões.

Suponha, nos dois exemplos anteriores, que a empresa compre anéis de diversos fabricantes mas, pelas características de produção do fabricante 2, os anéis produzidos por ele sejam especiais para a empresa. Assim, é importante identificar corretamente a origem, caso eles sejam oriundos do fabricante 2. Nesta situação, nossas hipóteses passariam a ser:

$$\begin{aligned} H_0 &: \text{ anéis são produzidos pelo fabricante 2} \\ H_1 &: \text{ anéis não são produzidos pelo fabricante 2} \end{aligned}$$

Queremos que a probabilidade α seja pequena; assim, podemos fixar α como 0,05 ou mesmo 0,01. De posse do valor dessa probabilidade, poderíamos estabelecer a região crítica ou região de rejeição. A diferença fundamental aqui está no cálculo da probabilidade do erro tipo II: não existe um único valor de β, já que, sob H_1, a distribuição pode ter qualquer média.

♦♦

5.2 Conceitos básicos

O contexto em que se baseia a teoria de teste de hipótese é basicamente o mesmo da teoria de estimação por intervalo de confiança. Temos uma população representada por uma variável aleatória X cuja distribuição de probabilidade depende de algum parâmetro θ. O interesse agora está em testar a veracidade de alguma afirmativa sobre θ.

5.2.1 Hipóteses nula e alternativa

A hipótese nula, representada por H_0, é a hipótese básica que queremos testar. Nesse texto consideraremos apenas hipóteses nulas simples, isto é, hipóteses que estabelecem que o parâmetro de interesse é *igual* a um determinado valor. A forma geral é:

$$H_0 : \theta = \theta_0$$

Alguns exemplos são:

$$H_0 : \mu = 6 \qquad H_0 : p = 0,5 \qquad H_0 : \sigma^2 = 25$$

O procedimento de teste de hipótese resultará em uma *regra de decisão* que nos permitirá *rejeitar* ou *não rejeitar* H_0.

A hipótese alternativa, representada por H_1, é a hipótese que devemos considerar no caso de rejeição da hipótese nula. A forma mais geral de H_1 é a hipótese *bilateral*

$$H_1 : \theta \neq \theta_0$$

Em algumas situações, podemos ter informação que nos permita restringir o domínio da hipótese alternativa. Por exemplo, se uma empresa farmacêutica está testando um novo medicamento para enxaqueca no intuito de reduzir o tempo entre a ingestão do medicamento e o alívio dos sintomas, uma possível hipótese alternativa é

$$H_1 : \mu < 10$$

Temos, então, hipóteses unilaterais à esquerda

$$H_1 : \theta < \theta_0$$

e hipóteses unilaterais à direita:

$$H_1 : \theta > \theta_0$$

A escolha entre essas formas de hipótese alternativa se faz com base no conhecimento sobre o problema sendo considerado e deve ser feita *antes* de se ter o resultado da amostra.

Nesse texto consideraremos o seguinte procedimento prático para determinação das hipóteses nula e alternativa.

> "Traduza" a afirmação do problema como uma desigualdade. Faça o mesmo para a afirmação que é a sua negação. A desigualdade que *não* envolve o sinal de = será a hipótese alternativa e a hipótese nula é sempre do tipo $\theta = \theta_0$.

EXEMPLO 5.5 Determinação de H_0 e H_1

Considerando as seguintes afirmativas como parte de um problema de teste de hipóteses, determine as hipóteses nula e alternativa apropriadas.

(a) O tempo médio é de, no máximo, 15 minutos

(b) Há, em média, pelo menos 15 clientes.

(c) A proporção de clientes tem de ser pelo menos 60%.

(d) A proporção de defeituosos tem de ser menor que 5%.

(e) O comprimento médio tem de ser 10cm.

Solução

(a) Afirmativa dada: $\mu \leq 15$
Complementar: $\mu > 15$

A desigualdade que não contém o sinal de = ($\mu > 15$) torna-se a hipótese alternativa:

$$H_0 : \mu = 15$$
$$H_1 : \mu > 15$$

(b) Afirmativa dada: $\mu \geq 15$
Complementar: $\mu < 15$

$$H_0 : \mu = 15$$
$$H_1 : \mu < 15$$

(c) Afirmativa dada: $p \geq 60\%$
Complementar: $p < 60\%$

$$H_0 : p = 0,6$$
$$H_1 : p < 0,6$$

(d) Afirmativa dada: $p < 5\%$
Complementar: $p \geq 5\%$

$$H_0 : p = 0,05$$
$$H_1 : p < 0,05$$

(e) Afirmativa dada: $\mu = 10$
Complementar: $\mu \neq 10$

$$H_0 : \mu = 10$$
$$H_1 : \mu \neq 10$$

♦♦

5.2.2 Estatística de teste, erros e regra de decisão

Assim como na construção dos intervalos de confiança, usaremos uma estatística amostral apropriada para construir o nosso teste de hipótese, e, nesse contexto, essa estatística é chamada *estatística de teste*. As estatísticas de teste naturalmente dependem do parâmetro envolvido no teste e nesse texto consideraremos os parâmetros média, variância e proporção (que também é uma média).

O procedimento de decisão será definido em termos da hipótese nula H_0, com duas decisões possíveis: (i) rejeitar H_0 ou (ii) não rejeitar H_0. No quadro a seguir, resumimos as situações possíveis.

		Decisão	
		Rejeitar H_0	Não rejeitar H_0
Possibilidades	H_0 verdadeira	Erro I	OK
	H_0 falsa	OK	Erro II

Vemos, aí, que existem duas possibilidades de erro:

Erro tipo I: rejeitar H_0 quando H_0 é verdadeira
Erro tipo II: não rejeitar H_0 quando H_0 é falsa

A decisão sobre a hipótese nula é tomada com base em uma regra que estabelece um conjunto de valores, chamado *região crítica* ou *região de rejeição*, de modo que, se o valor observado da estatística amostral cair nessa região, rejeitaremos H_0; caso contrário, não rejeitaremos H_0. Vamos denotar por RC a região crítica.

5.2.3 Região crítica e nível de significância

Em geral, a definição da região crítica é feita da seguinte forma: RC é o conjunto de valores cuja probabilidade de ocorrência é *pequena* sob a hipótese de veracidade de H_0. Sendo assim, a região crítica é construída com base na suposição de que H_0 é verdadeira.

A definição de "probabilidade pequena" se faz por meio da escolha da probabilidade α do erro tipo I, chamada *nível de significância* ou *tamanho do teste*, isto é:

$$\alpha = \text{P(erro tipo I)} = \text{P(rejeitar } H_0 \,|\, H_0 \text{ é verdadeira)}$$

Em geral, o valor de α é pequeno e as escolhas mais comuns são $\alpha = 0,05$ e $\alpha = 0,01$.

Definido o nível de significância α, podemos estabelecer a região crítica usando a distribuição amostral da estatística de teste.

EXEMPLO 5.6 Honestidade de uma moeda

Considere uma situação em que estamos interessados em verificar se uma moeda é honesta, isto é, $H_0 : p = 0,5$. Como não temos qualquer informação sobre o possível tipo de viés, nossa regra de decisão se baseará no número de coroas obtidas em 10 lançamentos. Se o número de coroas for muito pequeno ou muito grande, rejeitaremos a hipótese de honestidade da moeda. Para definir o que é grande ou pequeno, vamos estabelecer um nível de significância máximo de 1%, ou seja, a probabilidade de rejeitarmos a hipótese nula de honestidade da moeda quando, na verdade, ela é honesta tem de ser no máximo 0,01. Na tabela a seguir temos as probabilidades de ocorrência de cada um dos resultados possíveis, supondo que a moeda seja honesta. Nesse caso, se X é o número de coroas em 10 lançamentos, então $X \sim bin(10; 0,5)$.

Número de coroas x	P($X = x$)
0	0,0009766
1	0,0097656
2	0,0439453
3	0,1171875
4	0,2050781
5	0,2460938
6	0,2050781
7	0,1171875
8	0,0439453
9	0,0097656
10	0,0009766

A probabilidade de obtermos 0 ou 10 coroas com uma moeda honesta é $2 \times 0,0009766 = 0,0019531$ e se acrescentarmos os resultados 1 coroa ou 9 coroas, a soma das probabilidades é 0,021484, que é maior que 0,01. Assim, nossa regra de decisão deve ser "rejeitar H_0 se saírem 0 ou 10 coroas" e, nesse caso, a probabilidade do erro I é $\alpha = 0,0019531$.

♦♦

5.3 Exercícios propostos

1. Estabeleça as hipóteses nula e alternativa para as seguintes situações:

   (a) Depois de uma pane geral no sistema de informação de uma empresa, o gerente administrativo deseja saber se houve alteração no tempo de processamento de determinada atividade. Antes da pane, o tempo de processamento podia ser aproximado por uma variável aleatória normal com média de 100 minutos e desvio padrão de 10 minutos. O gerente acredita que a pane não tenha alterado a variabilidade do processo.

   (b) O dono de uma média empresa decide investigar a alegação de seus empregados de que o salário médio na sua empresa é menor que o salário médio nacional, que é de 900 reais.

   (c) Uma empresa fabricante de balas afirma que o peso médio de suas balas é de pelo menos 2 gramas.

2. Considere uma população normal com variância 225, da qual se extrai uma amostra aleatória simples de tamanho 25. Deseja-se testar as seguintes hipóteses:

$$\begin{aligned} H_0 &: \mu = 40 \\ H_1 &: \mu = 45 \end{aligned}$$

   (a) Se a região crítica é $RC : \overline{X} > 43$ calcule as probabilidades dos erros tipo I e II.

   (b) Determine a região crítica da forma $\overline{X} > k$ tal que a probabilidade do erro tipo I seja 0,10. Nesse caso, qual é a probabilidade do erro tipo II?

3. Considere uma população normal com variância 225, da qual se extrai uma amostra aleatória simples de tamanho 25. Deseja-se testar as seguintes hipóteses:

$$\begin{aligned} H_0 &: \mu = 40 \\ H_1 &: \mu \neq 40 \end{aligned}$$

   e para isso define-se a seguinte região crítica:

$$RC : \overline{X} > 46 \text{ ou } \overline{X} < 34$$

   (a) Calcule a probabilidade do erro tipo I.

   (b) Calcule a probabilidade do erro tipo II se $\mu = 36$.

4. Considere uma população normal com variância 64, da qual se extrai uma amostra aleatória simples de tamanho 16. Deseja-se testar as seguintes hipóteses:

$$\begin{aligned} H_0 &: \mu = 23 \\ H_1 &: \mu = 28 \end{aligned}$$

   (a) Se a região crítica é $RC : \overline{X} > 25,5$ calcule as probabilidades dos erros tipo I e II.

   (b) Determine a região crítica da forma $\overline{X} > k$ tal que a probabilidade do erro tipo I seja 0,05. Nesse caso, qual é a probabilidade do erro tipo II?

5. Desejando-se testar as hipóteses

$$\begin{aligned} H_0 &: \mu = 45 \\ H_1 &: \mu < 45 \end{aligned}$$

   sobre a média μ de uma população normal com variância 36, estabeleceu-se a seguinte região crítica com base em amostra aleatória simples de tamanho $n = 16$:

$$RC : \overline{X} < 41,25$$

   (a) Calcule a probabilidade do erro tipo I.

   (b) Calcule a probabilidade do erro tipo II se $\mu = 43$.

Capítulo 6

Testes de hipóteses baseados na distribuição normal

6.1 Introdução

Neste capítulo, aplicaremos os conceitos básicos sobre a teoria de teste de hipótese à situação específica em que a estatística de teste tem, pelo menos aproximadamente, distribuição normal. Veremos inicialmente testes para a média de uma população normal e, depois, testes para uma proporção populacional baseados em grandes amostras.

Vamos apresentar, inicialmente, alguns exemplos que ilustrarão diversas possibilidades que podem surgir na prática.

EXEMPLO 6.1 Tempo de processamento - parte 1

Depois de uma pane geral no sistema de informação de uma empresa, o gerente administrativo deseja saber se houve alteração no tempo de processamento de determinada atividade. Antes da pane, o tempo de processamento podia ser aproximado por uma variável aleatória normal com média de 100 minutos e desvio padrão de 10 minutos. O gerente acredita que a pane não tenha alterado a variabilidade do processo. Uma amostra de 16 tempos de processamento após a pane revela uma média de 105,5 minutos. Ao nível de significância de 5%, qual é a conclusão sobre a alteração do tempo médio de processamento?

Solução

Seja T a variável aleatória que representa o tempo de processamento. Do enunciado, sabemos que $T \sim N(\mu, 10^2)$ e sabemos, também, que antes da pane, $\mu = 10$.

- Hipóteses Nula e Alternativa

  O interesse do gerente é comparar os tempos antes e depois da pane. Antes da pane, o tempo médio de processamento era de 100 minutos. Como ele não sabe o tipo de alteração que pode ter ocorrido, precisa saber se o tempo médio depois da pane é diferente do tempo anterior. Temos, assim, as seguintes afirmativas $\mu = 100$ e $\mu \neq 100$, que nos levam às seguintes hipóteses nula e alternativa:

$$
\begin{aligned}
H_0 &: \mu = 100 \\
H_1 &: \mu \neq 100
\end{aligned}
$$

- Estatística de teste

Como a população é normal, sabemos que a distribuição da média amostral também é normal, e como não deve ter havido alteração na variabilidade do processo, resulta que o desvio padrão é de 10 minutos em qualquer situação.

Logo,

$$\overline{X} \sim N\left(\mu; \frac{100}{16}\right) \Leftrightarrow \frac{\overline{X} - \mu}{2,5} \sim N(0;1)$$

e nossa estatística de teste será

$$Z = \frac{\overline{X} - \mu}{2,5} \sim N(0;1)$$

- Nível de significância e região crítica

  Pelo enunciado do problema, o nível de significância é de 5%. Isso significa que a probabilidade de erro tipo I é 0,05. Como visto, o erro tipo I consiste em rejeitar a hipótese nula quando ela é verdadeira. Logo,

$$\alpha = \text{P}(\text{rejeitar } H_0 \,|\, H_0 \text{ verdadeira}) = 0,05$$

  Quando H_0 é verdadeira, $\mu = 100$ e, portanto,

$$H_0 \text{ verdadeira} \Longrightarrow Z_0 = \frac{\overline{X} - 100}{2,5} \sim N(0;1)$$

  A lógica do processo de decisão em um teste de hipótese é a seguinte: temos a distribuição da estatística de teste, supondo H_0 verdadeira. Nesse caso, nossa estatística de teste é Z_0 e a distribuição sob H_0 é a normal padrão. Valores observados de Z_0 com pequena probabilidade de ocorrência sob essa hipótese são indicativos de que a hipótese não é verdadeira. Assim, a região crítica consiste nos valores de Z_0 nas caudas da distribuição $N(0,1)$, que são as regiões de pequena probabilidade. Para delimitar essas regiões de pequena probabilidade, usamos o nível de significância e a hipótese alternativa. Como nesse exemplo a hipótese alternativa é bilateral, temos que tomar valores nas duas caudas da distribuição, distribuindo igualmente a probabilidade de erro, que é 5%. Veja a Figura 6.1:

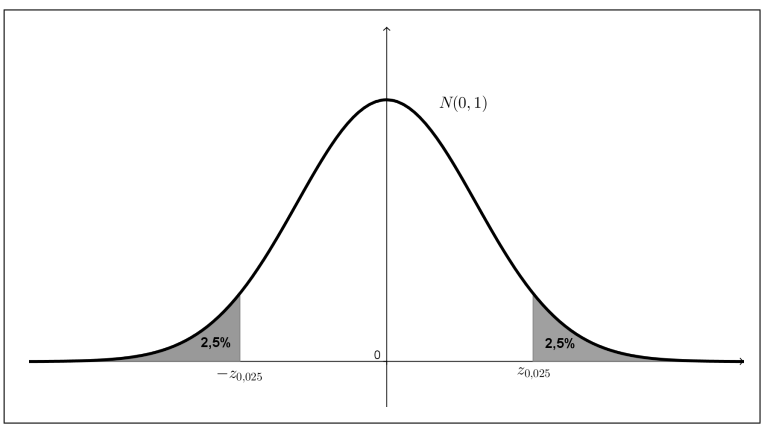


Figura 6.1 – Região crítica para o Exemplo 6.1

  Então, nossa região crítica consiste em valores observados da estatística de teste Z_0 que caem na área sombreada da Figura 6.1. Essa área sombreada é delimitada pelo valor crítico da $N(0,1)$ que deixa 2,5% acima dele, ou seja,

$$RC: \qquad Z_0 > z_{0,025} \qquad \text{ou} \qquad Z_0 < -z_{0,025}$$

  Olhando na tabela da distribuição normal, resulta

$$RC: \qquad Z_0 > 1,96 \qquad \text{ou} \qquad Z_0 < -1,96$$

ou equivalentemente,

$$RC: \qquad |Z_0| > 1,96$$

- Decisão e conclusão

  Os dados observados fornecem o valor $\overline{x} = 105,5$ minutos, que resulta no seguinte valor da estatística de teste:

$$z_0 = \frac{105,5 - 100}{2,5} = 2,2 > 1,96$$

  Como o valor da estatística de teste para a amostra observada está na região crítica, devemos rejeitar a hipótese nula, ou seja, as evidências amostrais indicam uma alteração do tempo de processamento da tarefa após a pane.

- Observação sobre a região crítica

  Vimos que a região crítica é $|Z_0| > 1,96$, ou seja

$$\left|\frac{\overline{X} - 100}{2,5}\right| > 1,96 \Leftrightarrow$$
$$\overline{X} > 100 + 1,96 \cdot 2,5 \qquad \text{ou} \qquad \overline{X} > 100 - 1,96 \cdot 2,5$$

  Assim, rejeitamos H_0 para valores de $\overline{X}$ distantes do valor 100 especificado em H_0. Como o teste é bilateral, "distante" pode ser acima ou abaixo de 100. ♦♦

EXEMPLO 6.2 Tempo de processamento - parte 2

Na mesma situação do exemplo anterior, é bastante razoável supor que o gerente esteja interessado apenas no caso de aumento do tempo de processamento. Afinal, se o tempo diminuir, isso significa que a tarefa vai ser executada mais rapidamente, o que representa um ganho.

Solução

- Hipóteses Nula e Alternativa

  As duas possibilidades são:

$$\begin{array}{lll} \mu & \leq & 100 \qquad \text{OK!} \\ \mu & > & 100 \qquad \text{Problema!} \end{array}$$

  Seguindo nosso procedimento, temos a seguinte situação:

$$\begin{array}{lll} H_0 & : & \mu = 100 \\ H_1 & : & \mu > 100 \end{array}$$

- estatística de teste

  A estatística de teste continua sendo

$$Z_0 = \frac{\overline{X} - 100}{2,5} \sim N(0;1)$$

- Nível de significância e região crítica

  O nível de significância é, ainda, 5%. Como antes, valores observados de Z_0 com pequena probabilidade de ocorrência sob H_0 são indicativos de que a hipótese não é verdadeira. Assim, a região crítica consiste nos valores de Z_0 na cauda da distribuição $N(0,1)$, na direção da hipótese alternativa. Agora, a hipótese alternativa é *unilateral à direita* e, portanto, a região crítica consiste

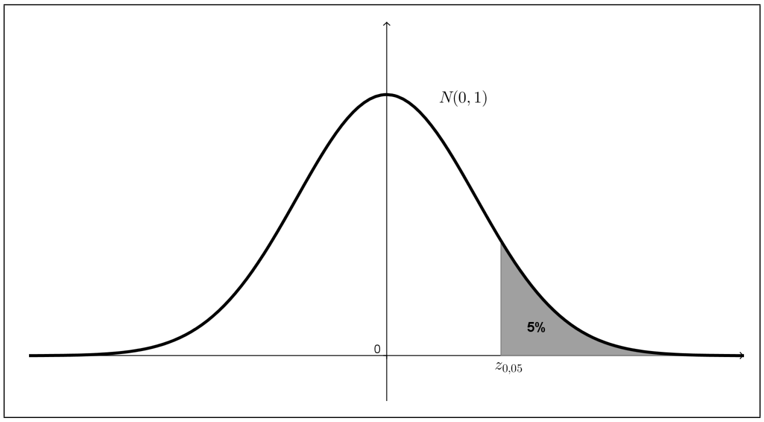


Figura 6.2 – Região crítica para o Exemplo 6.2

nos valores na cauda superior que respondem pela probabilidade de 5% do erro tipo I. Veja a Figura 6.2:

Então, nossa região crítica consiste em valores observados da estatística de teste Z_0 que caem na área sombreada da Figura 6.2. Essa área sombreada é delimitada pelo valor crítico da $N(0,1)$ que deixa 5% acima dele, ou seja,

$$RC: \qquad Z_0 > z_{0,05}$$

Olhando na tabela da distribuição normal, resulta

$$RC: \qquad Z_0 > 1,64$$

que é equivalente a $\overline{X} > 100 + 1,64 \cdot 2,5$, "distante" do valor 100 na direção da hipótese alternativa.

- Decisão e conclusão

  O valor da estatística de teste não se altera:

$$z_0 = \frac{105,5 - 100}{2,5} = 2,2 > 1,64$$

e como antes, devemos rejeitar a hipótese nula, ou seja, as evidências amostrais indicam um aumento do tempo de processamento da tarefa após a pane. ♦♦

EXEMPLO 6.3 Proporção de alunos

Uma pesquisa foi realizada com alunos da UFF visando, entre outras coisas, estimar a proporção dos alunos que têm conhecimento do Regulamento dos Cursos de Graduação dessa universidade (dados fictícios). Foram entrevistados 952 alunos, selecionados aleatoriamente, dos quais 132 afirmaram ter lido o Regulamento dos Cursos de Graduação. Suponha que a universidade decida lançar uma campanha de esclarecimento se a verdadeira proporção de alunos que conhecem o regulamento for inferior a 15%. Há razão para se lançar essa campanha? Justifique sua resposta através de um teste de hipótese com nível de significância de 5%.

Solução

Nosso problema agora é fazer um teste de hipótese sobre uma *proporção populacional*. Vimos que a proporção amostral é um bom estimador da proporção populacional e, para amostras grandes,

$$\widehat{P} \approx N\left(p, \frac{p(1-p)}{n}\right)$$

- Hipóteses nula e alternativa

  Afirmativa dada: $p < 0,15$
  Complementar: $p \geq 0,15$

  Isso nos leva às seguintes hipóteses:

$$\begin{array}{ll} H_0: & p = 0,15 \\ H_1: & p < 0,15 \end{array}$$

- Estatística de teste

  Sob a hipótese de que H_0 é verdadeira,

$$Z_0 = \frac{\widehat{P} - 0,15}{\sqrt{\frac{0,15\times(1-0,15)}{952}}} \approx N(0,1)$$

- Nível de significância e região crítica

  O nível de significância é 5%. Como antes, valores observados de Z_0 com pequena probabilidade de ocorrência sob H_0 são indicativos de que a hipótese não é verdadeira. Assim, a região crítica consiste nos valores de Z_0 na cauda da distribuição $N(0,1)$, na direção da hipótese alternativa. Agora, a hipótese alternativa é *unilateral à esquerda* e, portanto, a região crítica consiste nos valores na cauda inferior que respondem pela probabilidade de 5% do erro tipo I. Veja a Figura 6.3:

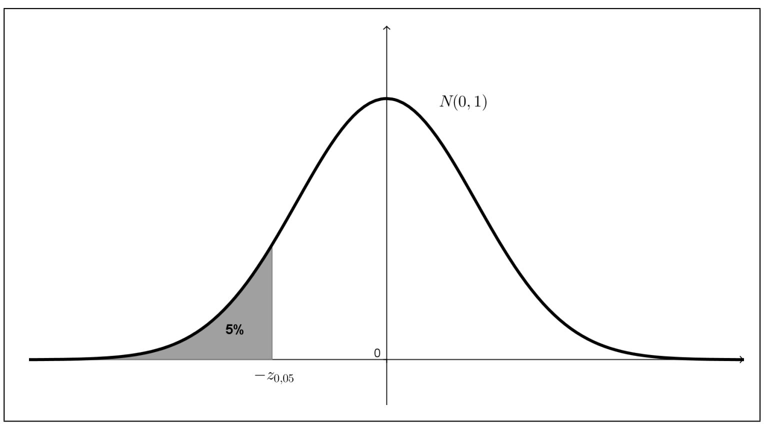


Figura 6.3 – Região crítica para o Exemplo 6.3

Então, nossa região crítica consiste em valores observados da estatística de teste Z_0 que caem na área sombreada da Figura 6.2. Essa área sombreada é delimitada pelo valor crítico da $N(0,1)$ que deixa 5% abaixo dele, ou seja,

$$RC: \qquad Z_0 < -z_{0,05}$$

Olhando na tabela da distribuição normal, resulta

$$RC: \qquad Z_0 < -1,64$$

que é equivalente a

$$\widehat{P} < 0,15 - 1,64 \cdot \sqrt{\frac{0,15 \times (1-0,15)}{952}}$$

- Decisão e conclusão

  O valor da estatística de teste é

$$z_0 = \frac{\frac{132}{952} - 0,15}{\sqrt{\frac{0,15\times(1-0,15)}{952}}} = -0,9803 \not< -1,64$$

O valor observado da estatística de teste não está na região crítica; logo, deixamos de rejeitar a hipótese nula, ou seja, não há razão para se lançar a campanha de esclarecimento. ♦♦

6.2 Teste de hipótese sobre a média de uma $N(\mu; \sigma^2)$ – σ^2 conhecida

Os dois primeiros exemplos anteriores ilustram o procedimento para construção de um teste de hipótese sobre a média de uma população normal com variância conhecida. De posse de uma amostra aleatória simples $X_1, X_2, \dots, X_n$ extraída de uma população $X \sim N(\mu; \sigma^2)$, nosso interesse está em testar a hipótese nula

$$H_0 : \mu = \mu_0$$

a um nível de significância α.

Dependendo do conhecimento sobre o problema, a hipótese alternativa pode tomar uma das três formas:

$$H_1 : \mu \neq \mu_0 \qquad H_1 : \mu > \mu_0 \qquad H_1 : \mu < \mu_0$$

Em qualquer dos casos, a estatística de teste baseia-se na média amostral; se a variância σ^2 é conhecida, sabemos que

$$Z = \frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0, 1)$$

A região crítica é estabelecida em função do nível de significância, que é a probabilidade α do erro tipo I:

$$\alpha = \mathrm{P}(\text{rejeitar } H_0 \,|\, H_0 \text{ verdadeira})$$

Quando H_0 é verdadeira, $\mu = \mu_0$ e, portanto, nossa estatística de teste é

$$Z_0 = \frac{\overline{X} - \mu_0}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0, 1)$$

Valores observados de Z_0 com pequena probabilidade de ocorrência são indicativos de que a hipótese não é verdadeira. Assim, a região crítica consiste nos valores de Z_0 na(s) cauda(s) da distribuição $N(0, 1)$, na direção da hipótese alternativa.

A seguir apresentamos os resultados para cada uma das possíveis hipóteses alternativas.

- Hipótese nula e estatística de teste

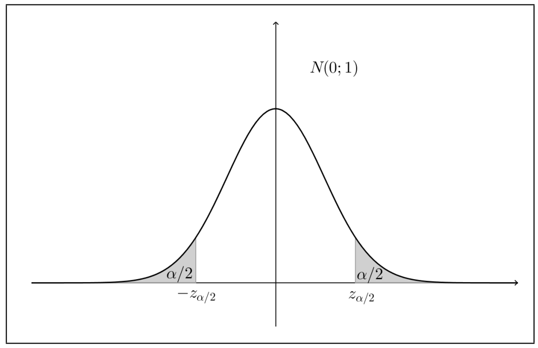
$$H_0 : \mu = \mu_0$$

$$Z_0 = \sqrt{n}\frac{\overline{X} - \mu_0}{\sigma} \underbrace{\sim}_{\text{sob } H_0} N(0, 1)$$

- Teste bilateral

$H_1 : \mu \neq \mu_0$

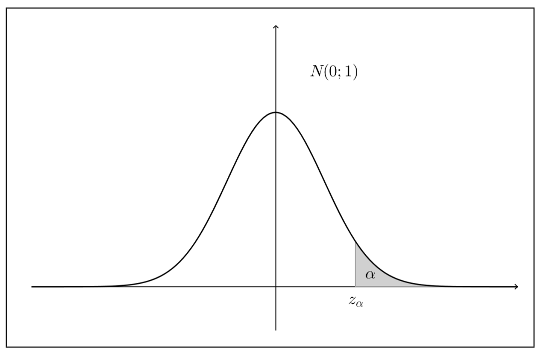
Região crítica:
$Z_0 < -z_{\alpha/2}$ ou $Z_0 > z_{\alpha/2}$



- Teste unilateral à direita

$H_1 : \mu > \mu_0$

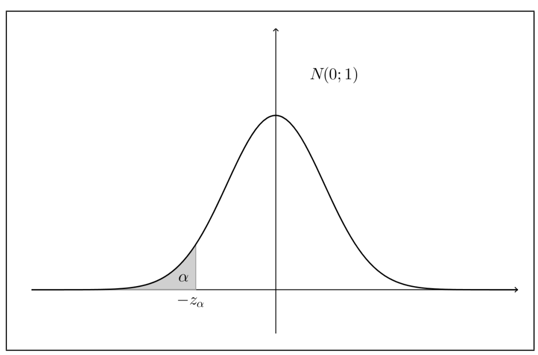
Região crítica:
$Z_0 > z_\alpha$



- Teste unilateral à esquerda

$H_1 : \mu < \mu_0$

Região crítica:
$Z_0 < -z_\alpha$



6.3 Teste de hipótese sobre uma proporção populacional

O Exemplo 6.3 ilustra o procedimento para construção de um teste de hipótese sobre uma proporção populacional p. De posse de uma *grande* amostra aleatória simples $X_1, X_2, \ldots, X_n$ extraída de uma população $X \sim Bern(p)$, nosso interesse está em testar a hipótese nula

$$H_0 : p = p_0$$

a um nível de significância α.

Dependendo do conhecimento sobre o problema, a hipótese alternativa pode tomar uma das três formas:

$$H_1 : p \neq p_0 \qquad H_1 : p > p_0 \qquad H_1 : p < p_0$$

Em qualquer dos casos, a estatística de teste baseia-se na proporção amostral; para *grandes amostras*, sabemos que

$$Z = \frac{\widehat{P} - p}{\sqrt{\frac{p(1-p)}{n}}} \approx N(0,1)$$

A região crítica é estabelecida em função do nível de significância, que é a probabilidade α do erro tipo I:

$$\alpha = \text{P}(\text{rejeitar } H_0 \,|\, H_0 \text{ verdadeira})$$

Quando H_0 é verdadeira, $p = p_0$ e, portanto, nossa estatística de teste é

$$Z_0 = \frac{\widehat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \approx N(0,1)$$

Valores observados de Z_0 com pequena probabilidade de ocorrência são indicativos de que a hipótese não é verdadeira. Assim, a região crítica consiste nos valores de Z_0 na(s) cauda(s) da distribuição $N(0,1)$, na direção da hipótese alternativa.

A seguir apresentamos os resultados para cada uma das possíveis hipóteses alternativas.

- Hipótese nula e estatística de teste

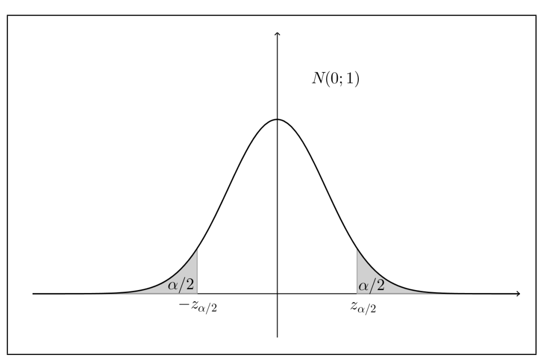
$$H_0 : p = p_0$$

$$Z_0 = \sqrt{n}\frac{\widehat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \underbrace{\approx}_{\text{sob } H_0} N(0,1)$$

- Teste bilateral

$H_1 : p \neq p_0$

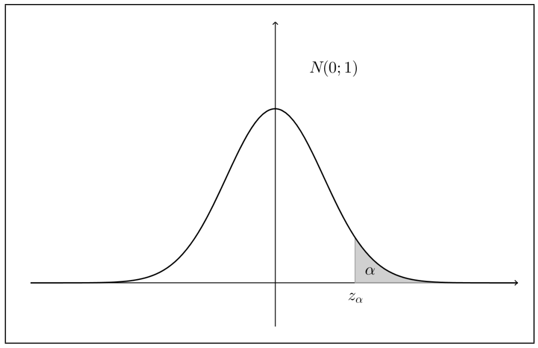
Região crítica:
$Z_0 < -z_{\alpha/2}$ ou $Z_0 > z_{\alpha/2}$



- Teste unilateral à direita

$H_1 : p > p_0$

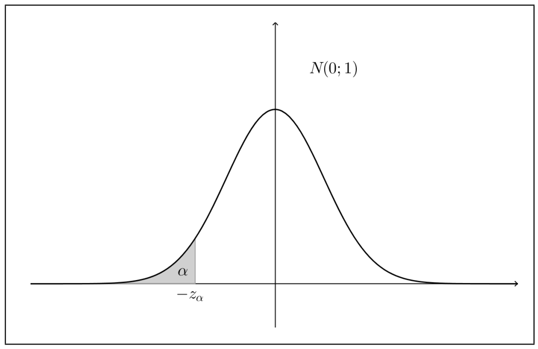
Região crítica:
$Z_0 > z_\alpha$



- Teste unilateral à esquerda

$H_1 : p < p_0$

Região crítica:
$Z_0 < -z_\alpha$



6.4 Valor P

Nos exemplos anteriores, a determinação da região crítica foi feita com base no nível de significância, isto é, fixado o nível de significância, encontramos o valor crítico que define os limites entre valores prováveis (aqueles que *não levam* à rejeição de H_0) e pouco prováveis (aqueles que *levam* à rejeição de H_0) sob a hipótese de veracidade de H_0.

Um outro procedimento bastante usual, especialmente quando são utilizados programas computacionais, consiste em calcular a probabilidade de se obter um valor da estatística de teste tão ou mais extremo que o valor observado, se H_0 for verdadeira. Um valor pequeno para tal probabilidade é indício de que H_0 não seja verdadeira. "Tão ou mais extremo" é sempre no sentido da hipótese alternativa, ou seja, no sentido de se rejeitar a hipótese nula. Temos, assim, a seguinte definição.

DEFINIÇÃO **Valor P ou probabilidade de significância**

O **valor** P é a probabilidade de se obter um valor da estatística de teste tão ou mais extremo que o valor observado, supondo-se H_0 verdadeira.

Vamos ilustrar esse conceito considerando novamente os três exemplos anteriores.

EXEMPLO 6.4 Valor P para o Exemplo 6.1

O valor observado da estatística de teste é $z_0 = 2,2$ e a hipótese alternativa é bilateral. Então, consideramos igualmente extremo o valor simétrico $-2,2$, ou seja, tão ou mais extremo significa ser maior que $2,2$, ou menor que $-2,2$ e o valor P é

$$P = \mathrm{P}(Z > 2,2) + \mathrm{P}(Z < -2,2) = 2 \times \mathrm{P}(Z > 2,2) = 2 \times [0,5 - tab(2,2)] = 0,0278$$

Na Figura 6.4 ilustra-se esse valor. O que esse resultado está nos dizendo é o seguinte: se H_0 for verdadeira, a probabilidade de obtermos um valor tão extremo quanto 2,2 na direção da hipótese alternativa, ou seja, em qualquer direção, já que H_1 é bilateral, é $0,0278$. Essa é uma probabilidade pequena, o que significa que é pouco provável obtermos um valor tão extremo quando H_0 é verdadeira. Logo, é razoável supormos que a hipótese nula não seja verdadeira, a mesma conclusão obtida ao trabalharmos com o nível de significância de 5%.

Na verdade, rejeitaríamos a hipótese nula para qualquer nível de significância maior que 0,0278. Note que tais níveis de significância implicariam em valores críticos menores do que o valor observado z_0 e, portanto, levariam à rejeição de H_0. Assim, o valor P é o menor nível de significância que leva à rejeição de H_0.

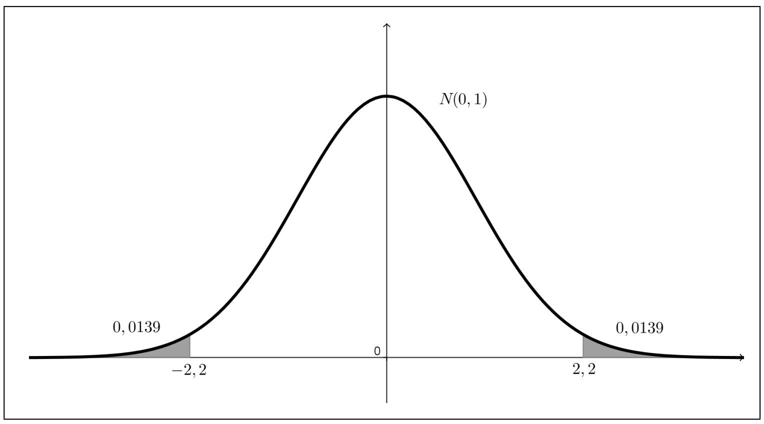


Figura 6.4 – Valor P para o Exemplo 6.1

♦♦

EXEMPLO 6.5 Valor P para o Exemplo 6.2

Como antes, o valor observado da estatística de teste é $z_0 = 2,2$, mas agora a hipótese alternativa é unilateral à direita. Então, valores tão ou mais extremos são aqueles maiores que $2,2$ e o valor P é

$$P = \mathrm{P}(Z > 2,2) = 0,5 - tab(2,2) = 0,0139$$

Na Figura 6.5 ilustra-se esse valor. O que esse resultado está nos dizendo é o seguinte: se H_0 for verdadeira, a probabilidade de obtermos um valor tão ou mais extremo que 2,2 é $0,0139$. Novamente, essa é uma probabilidade pequena, o que significa que é pouco provável obtermos um valor tão extremo quando H_0 é verdadeira. Logo, é razoável supormos que a hipótese nula não seja verdadeira, a mesma conclusão obtida ao trabalharmos com o nível de significância de 5%. Como antes, rejeitaríamos a hipótese nula para qualquer nível de significância maior que 0,0139.

♦♦

EXEMPLO 6.6 Valor P para o Exemplo 6.3

O valor observado da estatística de teste é $z_0 = -0,9803$, e a hipótese alternativa é unilateral à esquerda. Então, valores tão ou mais extremos são aqueles menores que $-0,9803$ e o valor P é

$$P = \mathrm{P}(Z < -0,9803) = \mathrm{P}(Z > 0,9803) = 0,5 - tab(0,98) = 0,1635$$

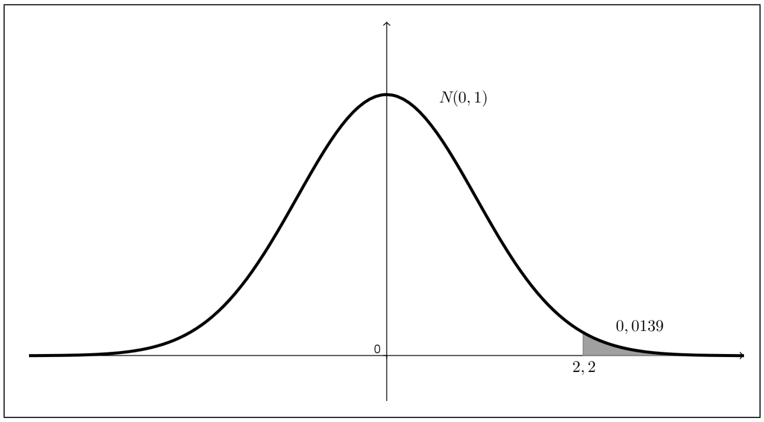


Figura 6.5 – Valor P para o Exemplo 6.2

Na Figura 6.6 ilustra-se esse valor. O que esse resultado está nos dizendo é o seguinte: se H_0 for verdadeira, há uma probabilidade alta de obtermos um valor tão ou mais extremo que $-0,9803$. Assim, não há evidência que indique que H_0 seja falsa.

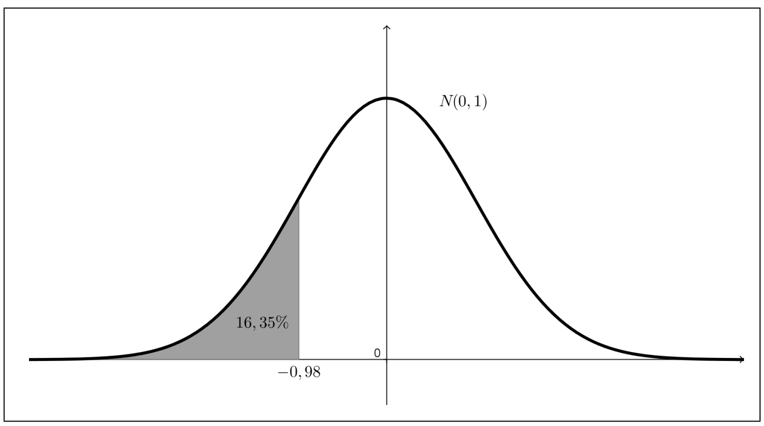


Figura 6.6 – Valor P para o Exemplo 6.3

♦♦

6.4.1 Procedimento geral para obtenção do valor P

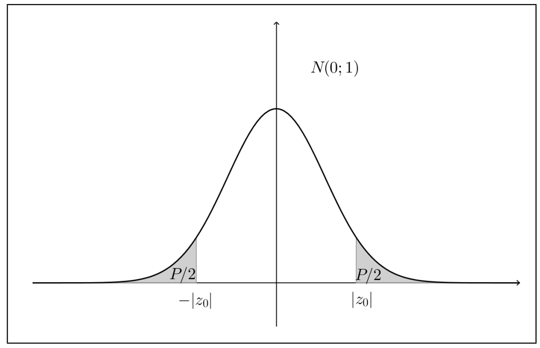
Os exemplos acima ilustram o procedimento para obtenção do valor P quando a estatística de teste tem distribuição normal. Como essa é uma distribuição simétrica, podemos sempre calcular o valor P trabalhando na cauda superior da distribuição normal padrão; para isso, basta usar o valor absoluto $|z_0|$ do valor observado da estatística de teste.

- Teste bilateral

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

$$P = \mathrm{P}(Z < -|z_0|) + \mathrm{P}(Z > |z_0|)$$

$$P = 2 \times \mathrm{P}(Z > |z_0|)$$



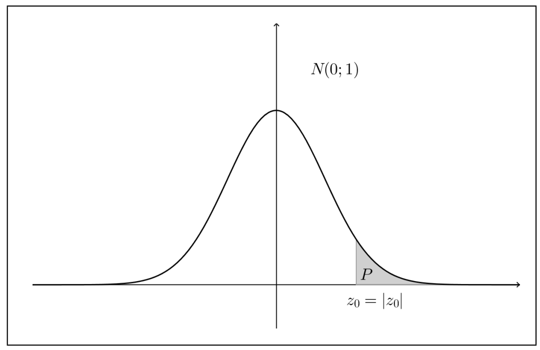
- Teste unilateral à direita

  Podemos supor que $z_0 > 0$. Caso contrário, o valor P será maior que $0,5$, o que leva à não rejeição de H_0 para qualquer nível de significância razoável.

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0$$

$$P = \mathrm{P}(Z > z_0)$$

$$P = \mathrm{P}(Z > |z_0|)$$



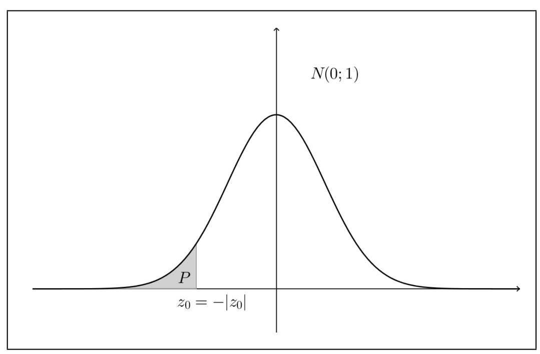
- Teste unilateral à esquerda

  Podemos supor que $z_0 < 0$. Caso contrário, o valor P será maior que $0,5$, o que leva à não rejeição de H_0 para qualquer nível de significância razoável.

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

$$P = \mathrm{P}(Z < z_0) = \mathrm{P}(Z < -|z_0|)$$

$$P = \mathrm{P}(Z > |z_0|)$$



6.4.2 Valor P e nível de significância

Vimos que o nível de significância α é a probabilidade do erro tipo I e o valor crítico correspondente delimita a região de rejeição, ou seja, valores da estatística de teste que caem na região crítica levam à rejeição de H_0. O valor P, por sua vez, é a probabilidade de se obter valores tão extremos quanto o observado e essa probabilidade, sendo pequena, leva à rejeição da hipótese nula.

Como podemos, então, relacionar o valor P e o nível de significância α em termos do processo decisório? Veja a Figura 6.7, onde ilustramos a situação para um teste unilateral à direita. Qualquer valor

z_0 maior que z_α leva à rejeição de H_0. Mas tais valores correspondem a probabilidades menores na cauda da distribuição, ou seja, correspondem a valores P menores que α. Isso nos leva à seguinte observação:

> **!** Valor P *versus* nível de significância
>
> O valor P é o menor nível de significância para o qual a hipótese nula H_0 é rejeitada, ou seja,
>
> $$\text{rejeitamos } H_0 \Leftrightarrow P \leq \alpha$$

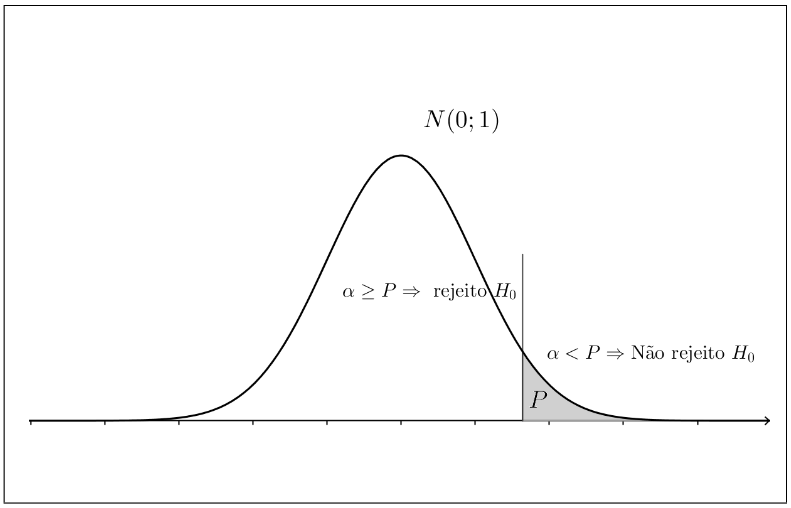


Figura 6.7 – Valor P *versus* nível de significância

EXEMPLO 6.7 Peso de bala

Uma empresa fabricante de balas afirma que o peso médio de suas balas é de pelo menos 2 gramas. Pela descrição do processo de produção, sabe-se que o peso das balas distribui-se normalmente com desvio padrão de 0,5 grama. Uma amostra de 25 balas apresenta peso médio de 1,81 gramas. O que se pode concluir sobre a afirmação do fabricante? Estabeleça sua conclusão usando um nível de significância de 5% e também o valor P.

Solução

Seja X a variável aleatória que representa o peso das balas. Então, $X \sim N(\mu; 0,25)$. Como $n = 25$, resulta que

$$Z = \frac{\overline{X} - \mu}{\sqrt{\frac{0,25}{25}}} \sim N(0,1)$$

A afirmativa do fabricante é $\mu \geq 2$. Logo, a negação de tal afirmação é $\mu < 2$. Como essa última expressão não contém o sinal de igualdade, ela se torna a hipótese alternativa. Então, nossas hipóteses são:

$$\begin{aligned} H_0 &: \mu = 2 \\ H_1 &: \mu < 2 \end{aligned}$$

Para $\alpha = 0,05$, a região crítica é

$$RC : Z_0 < -z_{0,05} = -1,64$$

O valor observado da estatística de teste é

$$z_0 = \frac{1,81 - 2,00}{\sqrt{\frac{0,25}{25}}} = -1,9 < -1,64$$

Como o valor observado da estatística de teste está na região crítica, rejeita-se a hipótese nula, ou seja, há evidência de que o peso médio seja menor que 2 gramas.

Temos também que

$$P = \mathrm{P}(Z > |-1,9|) = 0,5 - tab(1,9) = 0,0287$$

Assim, rejeitaríamos H_0 para qualquer nível de significância maior que 2,87%, o que inclui 5%. ♦♦

6.5 Função Característica de Operação e Poder do Teste

No procedimento de teste de hipótese, as decisões possíveis são rejeitar ou não rejeitar H_0. Definem-se, então, as seguintes funções em termos das probabilidades de cada uma delas: função característica de operação e função poder.

DEFINIÇÃO Função característica de operação e função poder

A **função característica de operação** $\beta(\theta)$ é definida como

$$\beta(\theta) = \mathrm{P}(\text{não rejeitar } H_0 \mid \theta)$$

A **função poder do teste** é definida como

$$\pi(\theta) = \mathrm{P}(\text{rejeitar } H_0 \mid \theta) = 1 - \beta(\theta)$$

Note que essas são funções de θ, o verdadeiro e desconhecido valor do parâmetro θ. Se este valor estiver no conjunto de valores definidos pela hipótese alternativa, então $\pi(\theta)$ corresponde a uma probabilidade de acerto: ela mede a probabilidade de se rejeitar H_0 quando H_0 é falsa. Por outro lado, se a hipótese nula é $H_0 : \theta = \theta_0$, então

$$\begin{aligned}
\pi(\theta_0) &= 1 - \beta(\theta_0) \\
&= 1 - \mathrm{P}(\text{não rejeitar } H_0 \mid \theta_0) \\
&= 1 - \mathrm{P}(\text{não rejeitar } H_0 \mid H_0 \text{ verdadeira}) \\
&= \mathrm{P}(\text{rejeitar } H_0 \mid H_0 \text{ verdadeira}) \\
&= \alpha
\end{aligned}$$

EXEMPLO 6.8 Honestidade da moeda, continuação

Voltando ao exemplo sobre a honestidade da moeda, temos que, se H_0 é verdadeira, então $X \sim bin(10; 0,5)$, em que X = "número de coroas". Nossa regra de decisão é rejeitar H_0 se $X = 0$ ou $X = 10$. Nesse caso, a probabilidade do erro tipo I é:

$$\alpha = \mathrm{P}\left[X = 0|X \sim bin\left(10; \frac{1}{2}\right)\right] + \mathrm{P}\left[X = 0|X \sim bin\left(10; \frac{1}{2}\right)\right] = 2 \times \left(\frac{1}{2}\right)^1 0 = 0,0019531$$

A função poder do teste é

$$\pi(p) = \mathrm{P}(\text{rejeitar } H_0 \,|\, p) = \mathrm{P}(\{X = 0\} \cup \{X = 10\} \,|\,) = p^{10} + (1-p)^{10}$$

e a função característica de operação é

$$\beta(p) = \mathrm{P}(\text{não rejeitar } H_0 \,|\, p) = \mathrm{P}(0 < X < 10 \,|\, p) = 1 - p^{10} - (1-p)^{10}$$

Na Figura 6.8 são apresentados os gráficos dessas duas funções.

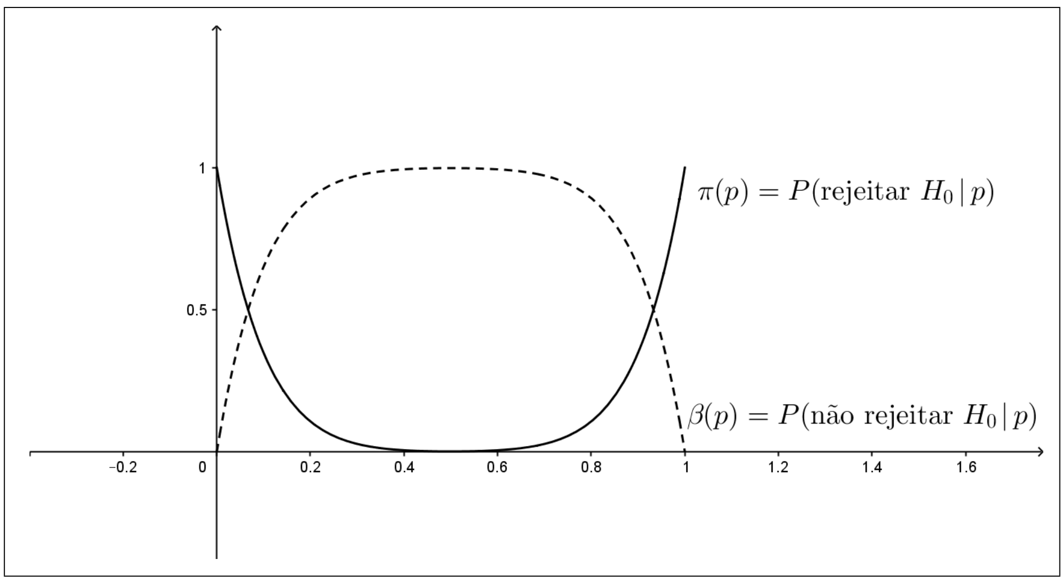


Figura 6.8 – Probabilidades dos Erros tipo I e II para o Exemplo 5.6

♦♦

6.5.1 Poder do teste Z bilateral

Consideremos o teste bilateral para a média μ de uma população normal com variância σ^2 conhecida:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

A hipótese alternativa pode ser escrita em função da diferença $\Delta = \mu - \mu_0$ como

$$H_1 : \Delta = \mu - \mu_0 \neq 0$$

Vimos que, para um nível de significância α, a regra de decisão é rejeitar H_0 se $\left|\sqrt{n}\dfrac{\overline{X}-\mu_0}{\sigma}\right| > z_{\alpha/2}$.[1] Assim, a função poder do teste é

$$\begin{aligned}
\pi(\mu) &= \mathrm{P}(\text{rejeitar } H_0 \mid \mu) = 1 - \mathrm{P}(\text{não rejeitar } H_0 \mid \mu) \\
&= 1 - \mathrm{P}\left(-z_{\alpha/2} \le \sqrt{n}\frac{\overline{X}-\mu_0}{\sigma} \le z_{\alpha/2} \mid \mu\right) \\
&= 1 - \mathrm{P}\left(-z_{\alpha/2} \le \sqrt{n}\frac{\overline{X}-(\mu-\Delta)}{\sigma} \le z_{\alpha/2} \mid \mu\right) \\
&= 1 - \mathrm{P}\left(-z_{\alpha/2} \le \sqrt{n}\frac{\overline{X}-\mu}{\sigma} + \sqrt{n}\frac{\Delta}{\sigma} \le z_{\alpha/2} \mid \mu\right) \\
&= 1 - \mathrm{P}\left(-z_{\alpha/2} - \sqrt{n}\frac{\Delta}{\sigma} \le Z \le z_{\alpha/2} - \sqrt{n}\frac{\Delta}{\sigma}\right)
\end{aligned}$$

ou seja, em termos da diferença Δ, temos

$$\pi(\Delta) = 1 - \left[\Phi\left(z_{\alpha/2} - \sqrt{n}\frac{\Delta}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \sqrt{n}\frac{\Delta}{\sigma}\right)\right] \tag{6.1}$$

Na Figura 6.9 ilustra-se o poder do teste bilateral para a média de uma população normal com $\sigma = 20$ e nível de significância $\alpha = 0,05$. Note que quando $\Delta = 0$, isto é, $\mu = \mu_0$, $\pi(\mu_)) = \alpha = 0,05$. Além disso, quanto maior Δ, maior a probabilidade de rejeitarmos H_0. Esse resultado deve ser intuitivo: quanto mais afastado μ estiver de μ_0, mais fácil é identificar que $\mu \neq \mu_0$. Valores próximos de μ_0 têm baixo poder e muitas vezes, estipula-se o poder desejado para determinada diferença como uma condição a ser atendida e isso define o tamanho de amostra que será necessário, conforme veremos a seguir.

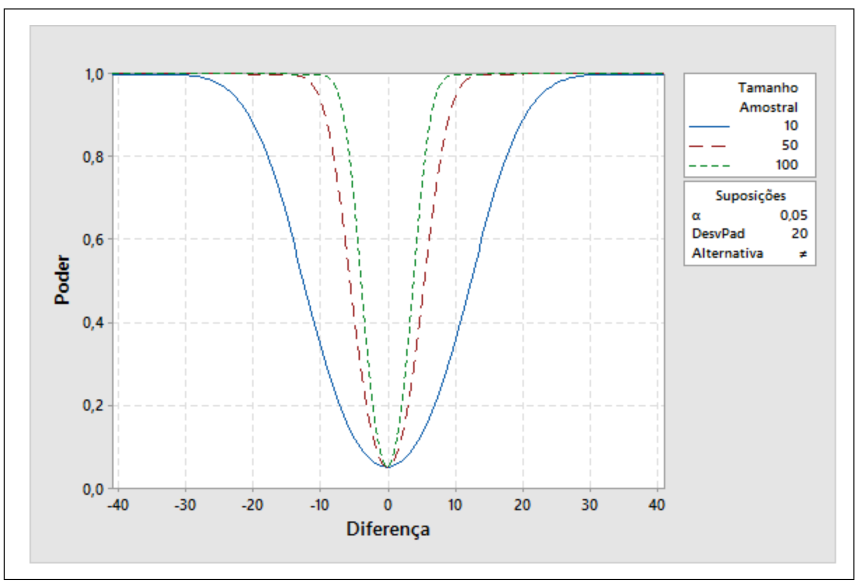


Figura 6.9 – Poder do teste Z bilateral em função de $\Delta = \mu - \mu_0$

Tamanho da amostra

Podemos ver também, na figura anterior, que, quanto maior n, maior o poder do teste. Uma questão que se coloca, então, é: quão grande deve ser n para que tenhamos um determinado poder, ou seja, se

[1] z_α é o valor crítico da $N(0;1)$ tal que $\mathrm{P}(Z > z_\alpha) = \alpha$

queremos $\pi(\Delta^*) = \pi^*$, qual deve ser o tamanho n da amostra?

$$\begin{aligned}
&\pi(\Delta^*) = \pi^* \Rightarrow \pi^* = 1 - \left[\Phi\left(z_{\alpha/2} - \sqrt{n}\frac{\Delta^*}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \sqrt{n}\frac{\Delta^*}{\sigma}\right)\right] \Rightarrow \\
&1 - \pi^* = \Phi\left(z_{\alpha/2} - \sqrt{n}\frac{\Delta^*}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \sqrt{n}\frac{\Delta^*}{\sigma}\right)
\end{aligned}$$

Suponhamos, inicialmente, que $\Delta^* > 0$. Na Figura 6.10a ilustram-se as abcissas envolvidas e aí podemos ver que $\Phi\left(-z_{\alpha/2} - \sqrt{n}\frac{\Delta^*}{\sigma}\right) \approx 0$. Logo,

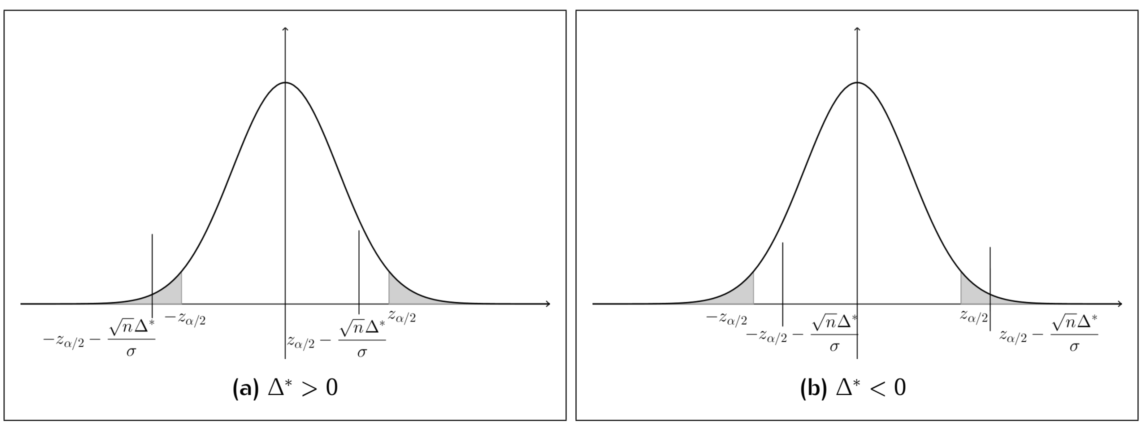
$$\begin{aligned}
&1 - \pi^* \approx \Phi\left(z_{\alpha/2} - \sqrt{n}\frac{\Delta^*}{\sigma}\right) \Rightarrow \pi^* \approx \mathrm{P}\left(Z > z_{\alpha/2} - \sqrt{n}\frac{\Delta^*}{\sigma}\right) \Rightarrow \\
&z_{\alpha/2} - \sqrt{n}\frac{\Delta^*}{\sigma} \approx z_{\pi^*} \Rightarrow \sqrt{n} \approx \frac{(z_{\alpha/2} - z_{\pi^*}) \cdot \sigma}{\Delta^*}
\end{aligned}$$

Analogamente, se $\Delta^* < 0$, podemos ver na Figura 6.10b que $\Phi\left(z_{\alpha/2} - \sqrt{n}\frac{\Delta^*}{\sigma}\right) \approx 1$ e, portanto

$$\begin{aligned}
&\pi^* \approx \Phi\left(-z_{\alpha/2} - \sqrt{n}\frac{\Delta^*}{\sigma}\right) = \mathrm{P}\left(Z \le -z_{\alpha/2} - \sqrt{n}\frac{\Delta^*}{\sigma}\right) = \mathrm{P}\left(Z \ge z_{\alpha/2} + \sqrt{n}\frac{\Delta^*}{\sigma}\right) \Rightarrow \\
&z_{\alpha/2} + \sqrt{n}\frac{\Delta^*}{\sigma} \approx z_{\pi^*} \Rightarrow \sqrt{n} \approx \frac{(z_{\pi^*} - z_{\alpha/2}) \cdot \sigma}{\Delta^*}
\end{aligned}$$

Em ambos os casos resulta

$$n \approx \left(\frac{(z_{\alpha/2} - z_{\pi^*}) \cdot \sigma}{\Delta^*}\right)^2 \tag{6.2}$$



(a) $\Delta^* > 0$ **(b)** $\Delta^* < 0$

Figura 6.10 – Abcissas no cálculo de $\pi(\Delta^*$

6.5.2 Poder do teste Z unilateral

Teste unilateral à direita

Consideremos inicialmente o teste para

$$\begin{aligned}
H_0 &: \mu = \mu_0 \\
H_1 &: \mu > \mu_0
\end{aligned}$$

Como antes, podemos escrever a hipótese alternativa em termos da diferença $\Delta = \mu - \mu_0$ como

$$H_1 : \Delta > 0$$

Para um nível de significância α, a regra de decisão é rejeitar H_0 se $\sqrt{n}\dfrac{\overline{X} - \mu_0}{\sigma} > z_\alpha$. Assim, a função poder do teste é

$$\begin{aligned}
\pi(\mu) &= \mathrm{P}(\text{rejeitar } H_0 \,|\, \mu) = \mathrm{P}\left(\sqrt{n}\frac{\overline{X} - \mu_0}{\sigma} > z_\alpha \,|\, \mu\right) \\
&= \mathrm{P}\left(\sqrt{n}\frac{\overline{X} - (\mu - \Delta)}{\sigma} > z_\alpha \,|\, \mu\right) \\
&= \mathrm{P}\left(\sqrt{n}\frac{\overline{X} - \mu}{\sigma} + \sqrt{n}\frac{\Delta}{\sigma} > z_\alpha \,|\, \mu\right) \\
&= \mathrm{P}\left(Z > z_\alpha - \sqrt{n}\frac{\Delta}{\sigma}\right)
\end{aligned}$$

ou seja, em termos da diferença Δ, temos

$$\pi(\Delta) = 1 - \Phi\left(z_\alpha - \sqrt{n}\frac{\Delta}{\sigma}\right) \qquad \Delta > 0 \tag{6.3}$$

Na Figura 6.11a ilustra-se a função poder do teste unilateral à direita para a média de uma população normal com $\sigma = 20$ e nível de significância $\alpha = 0,05$, para diferentes tamanhos de amostra. Note que quando $\Delta = 0$, isto é, $\mu = \mu_0$, $\pi(\mu_0) = \alpha = 0,05$.

Teste unilateral à esquerda

Consideremos inicialmente o teste para

$$\begin{aligned} H_0 &: \mu = \mu_0 \\ H_1 &: \mu < \mu_0 \end{aligned}$$

Em termos da diferença $\Delta = \mu - \mu_0$ temos

$$H_1 : \Delta < 0$$

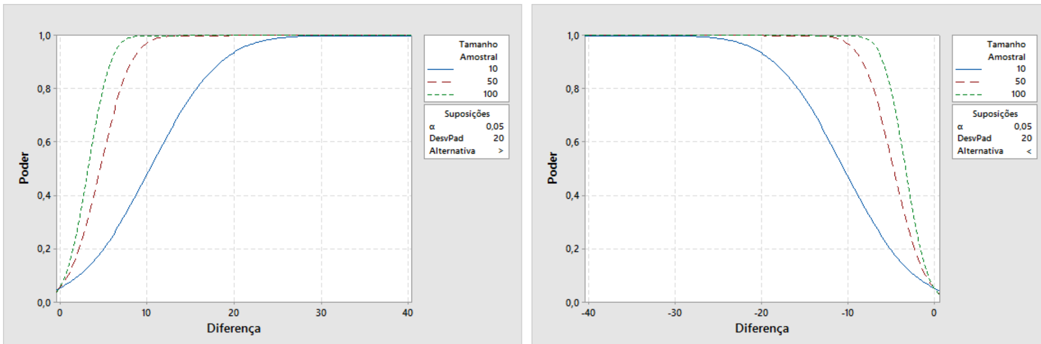
Para um nível de significância α, a regra de decisão é rejeitar H_0 se $\sqrt{n}\dfrac{\overline{X} - \mu_0}{\sigma} < -z_\alpha$. Assim, a função poder do teste é

$$\begin{aligned}
\pi(\mu) &= \mathrm{P}(\text{rejeitar } H_0 \,|\, \mu) = \mathrm{P}\left(\sqrt{n}\frac{\overline{X} - \mu_0}{\sigma} < -z_\alpha \,|\, \mu\right) \\
&= \mathrm{P}\left(\sqrt{n}\frac{\overline{X} - (\mu - \Delta)}{\sigma} < -z_\alpha \,|\, \mu\right) \\
&= \mathrm{P}\left(\sqrt{n}\frac{\overline{X} - \mu}{\sigma} + \sqrt{n}\frac{\Delta}{\sigma} < -z_\alpha \,|\, \mu\right) \\
&= \mathrm{P}\left(Z < -z_\alpha - \sqrt{n}\frac{\Delta}{\sigma}\right) \\
&= \mathrm{P}\left(Z > z_\alpha + \sqrt{n}\frac{\Delta}{\sigma}\right)
\end{aligned}$$

ou seja, em termos da diferença Δ, temos

$$\pi(\Delta) = 1 - \Phi\left(z_\alpha + \sqrt{n}\frac{\Delta}{\sigma}\right) \qquad \Delta < 0 \tag{6.4}$$

Na Figura 6.11b ilustra-se a função poder do teste unilateral à direita para a média de uma população normal com $\sigma = 20$ e nível de significância $\alpha = 0,05$, para diferentes tamanhos de amostra o poder do . Note que quando $\Delta = 0$, isto é, $\mu = \mu_0$, $\pi(\mu_0) = \alpha = 0,05$.



(a) Unilateral à direita **(b)** Unilateral à esquerda

Figura 6.11 – Poder do teste Z unilateral em função de $\Delta = \mu - \mu_0$

Tamanho da amostra

Se queremos $\pi(\Delta^*) = \pi^*$ em um teste unilateral à direita, podemos ver, de (6.3), que

$$\pi(\Delta^*) = \pi^* = 1 - \Phi\left(z_\alpha - \sqrt{n}\frac{\Delta^*}{\sigma}\right) = \mathrm{P}\left(Z > z_\alpha - \sqrt{n}\frac{\Delta^*}{\sigma}\right) \Rightarrow$$
$$z_\alpha - \sqrt{n}\frac{\Delta^*}{\sigma} = z_{\pi^*} \Rightarrow \sqrt{n} = \frac{(z_\alpha - z_{\pi^*})\cdot\sigma}{\Delta^*}$$

De forma análoga, para o teste unilateral à esquerda,

$$\pi(\Delta^*) = \pi^* = 1 - \Phi\left(z_\alpha + \sqrt{n}\frac{\Delta^*}{\sigma}\right) = \mathrm{P}\left(Z > z_\alpha + \sqrt{n}\frac{\Delta^*}{\sigma}\right) \Rightarrow$$
$$z_\alpha + \sqrt{n}\frac{\Delta^*}{\sigma} = z_{\pi^*} \Rightarrow \sqrt{n} = \frac{(z_{\pi^*} - z_\alpha)\cdot\sigma}{\Delta^*}$$

Para ambos os testes unilaterais, obtemos o mesmo tamanho de amostra dado por

$$n = \left[\frac{(z_\alpha - z_{\pi^*})\cdot\sigma}{\Delta^*}\right]^2 \tag{6.5}$$

Esse é um resultado razoável tendo em vista a simetria dos dois testes.

EXEMPLO 6.9

Consideremos uma população representada por uma variável aleatória normal com média μ e variância 400. Deseja-se testar

$$\begin{aligned} H_0 &: \mu = 100 \\ H_1 &: \mu \neq 100 \end{aligned}$$

com base em uma amostra aleatória simples de tamanho $n = 16$. Para tal, define-se a seguinte região crítica:

$$RC : \overline{X} < 85 \text{ ou } \overline{X} > 115$$

1. Calcule a probabilidade do erro tipo I.

2. Calcule a função poder do teste para os seguintes valores de μ: 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125.

3. Quanto vale a função poder do teste quando $\mu = 100$?

Solução

Como a população é normal com média μ e variância 400, sabemos que $\overline{X}$ também é normal com média μ e variância $\frac{400}{16} = 25$.

1. Sob a hipótese nula, $\mu = 100$. Então,

$$\begin{aligned}
\alpha &= \text{P(rejeitar } H_0 \,|\, H_0 \text{ verdadeira)} \\
&= \text{P}\left[\left\{\overline{X} < 85\right\} \cup \left\{\overline{X} > 115\right\} \mid \overline{X} \sim N(100; 25)\right] \\
&= \text{P}\left[\overline{X} < 85 \,|\, \overline{X} \sim N(100; 25)\right] + \text{P}\left[\overline{X} > 115 \,|\, \overline{X} \sim N(100; 25)\right] \\
&= \text{P}\left(Z < \frac{85 - 100}{5}\right) + \text{P}\left(Z > \frac{115 - 100}{5}\right) \\
&= \text{P}(Z < -3) + \text{Pr}(Z > 3) \\
&= 2 \times \text{Pr}(Z > 3) \\
&= 2 \times [0,5 - \text{P}(0 \leq Z \leq 3)] \\
&= 0,0027
\end{aligned}$$

2. A função poder é dada por

$$\begin{aligned}
1 - \beta(\mu) &= 1 - \text{P(não rejeitar } H_0 \,|\, \mu) \\
&= 1 - \text{P}(85 \leq \overline{X} \leq 115 \,|\, \mu) \\
&= 1 - \text{P}\left[85 \leq \overline{X} \leq 115 \,|\, \overline{X} \sim N(\mu; 25)\right] \\
&= 1 - \text{P}\left(\frac{85 - \mu}{5} \leq Z \leq \frac{115 - \mu}{5}\right)
\end{aligned}$$

Vamos ilustrar o cálculo para $\mu = 75$:

$$\begin{aligned}
1 - \beta(75) &= 1 - \text{P}(2 \leq Z \leq 8) \\
&= 1 - [\text{P}(0 \leq Z \leq 8) - \text{P}(0 \leq Z \leq 2) =] \\
&= 0,97725
\end{aligned}$$

De forma análoga, obtemos a seguinte tabela:

μ	$1-\beta(\mu)$
75	0,97725
80	0,84134
85	0,50000
90	0,15866
95	0,02278
100	0,00270
105	0,02278
110	0,15866
115	0,50000
120	0,84134
125	0,97725

Observe que, para $\mu = 100$, valor da hipótese nula, a função poder é igual à probabilidade do erro tipo I (nível de significância).

É interessante notar também que quanto mais distante do valor $\mu_0 = 100$, maior o poder do teste, ou seja, há uma probabilidade mais alta de se rejeitar H_0 quando o valor alternativo μ está bem distante de μ_0.

EXEMPLO 6.10 Tamanho de amostra

Suponha que o nível de colesterol para homens com idades entre 20 e 24 anos seja normalmente distribuído com média de 180 mg/100ml e desvio padrão de 46 mg/100ml. Estamos interessados em testar, ao nível de significância de 0,01, se, para um determinado grupo, o nível de colesterol é mais alto que 180 mg/100ml. Mas é importante, também, que tenhamos uma chance de no máximo 5% de deixar de rejeitar a hipótese nula no caso de o verdadeiro nível de colesterol ser de pelo menos 210 mg/100 ml. Qual tamanho de amostra devemos utilizar?

Solução

Os dados do problema nos dão o seguinte:

$$H_0 : \mu = 180$$
$$H_1 : \mu > 180$$

$\alpha = 0,01 \Rightarrow z_{0,01} = 2,33$

$\text{P}(\text{não rejeitar } H_0|\mu = 210) = 0,05 \Rightarrow \text{P}(\text{rejeitar } H_0|\mu = 210) = 0,95 \Leftrightarrow \pi(210) = 0,95$

$\pi^* = 0,95 \Rightarrow z_{\pi^*} = -1,645$ $\qquad \Delta^* = 210 - 180 = 30$

Logo,

$$n = \left[\frac{(2,33 - (-1,645)) \cdot 46}{30}\right]^2 \Rightarrow n \geq 38$$

6.5.3 Poder do teste Z bilateral para proporções

Para o teste bilateral sobre uma proporção baseado em grandes amostras

$$H_0 : p = p_0$$
$$H_1 : p \neq p_0$$

a regra de decisão é rejeitar H_0 se $\left|\sqrt{n}\frac{\widehat{P}-p_0}{\sqrt{p_0(1-p_0)}}\right| > z_{\alpha/2}$.[2] Assim, a função poder do teste é

$$
\begin{aligned}
\pi(p) &= \mathrm{P}(\text{rejeitar } H_0 \,|\, p) = 1 - \mathrm{P}(\text{não rejeitar } H_0 \,|\, p) \\
&= 1 - \mathrm{P}\left(-z_{\alpha/2} \le \frac{\widehat{P}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \le z_{\alpha/2} \,|\, p\right) \\
&= 1 - \mathrm{P}\left(-z_{\alpha/2} \le \frac{\widehat{P}-p+p-p_0}{\sqrt{\frac{p(1-p)}{n}}}\frac{\sqrt{\frac{p(1-p)}{n}}}{\sqrt{\frac{p_0(1-p_0)}{n}}} \le z_{\alpha/2} \,|\, p\right) \\
&= 1 - \mathrm{P}\left(-z_{\alpha/2}\frac{\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}} \le \frac{\widehat{P}-p+p-p_0}{\sqrt{\frac{p(1-p)}{n}}} \le z_{\alpha/2}\frac{\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}} \,|\, p\right) \\
&= 1 - \mathrm{P}\left(-z_{\alpha/2}\frac{\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}} \le \frac{\widehat{P}-p}{\sqrt{\frac{p(1-p)}{n}}} + \frac{p-p_0}{\sqrt{\frac{p(1-p)}{n}}} \le z_{\alpha/2}\frac{\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}} \,|\, p\right) \\
&= 1 - \mathrm{P}\left(-z_{\alpha/2}\frac{\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}} - \frac{p-p_0}{\sqrt{\frac{p(1-p)}{n}}} \le \frac{\widehat{P}-p}{\sqrt{\frac{p(1-p)}{n}}} \le z_{\alpha/2}\frac{\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}} - \frac{p-p_0}{\sqrt{\frac{p(1-p)}{n}}} \,|\, p\right) \\
&= 1 - \mathrm{P}\left(-z_{\alpha/2}\frac{\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}} - \frac{p-p_0}{\sqrt{\frac{p(1-p)}{n}}} \le Z \le z_{\alpha/2}\frac{\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}} - \frac{p-p_0}{\sqrt{\frac{p(1-p)}{n}}} \,|\, p\right)
\end{aligned}
$$

ou seja,

$$
\pi(p) = 1 - \left[\Phi\left(\frac{z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}} - (p-p_0)}{\sqrt{\frac{p(1-p)}{n}}}\right) - \Phi\left(\frac{-z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}} - (p-p_0)}{\sqrt{\frac{p(1-p)}{n}}}\right)\right] \tag{6.6}
$$

Tamanho da amostra

Para se ter um poder $\pi^* = \pi(p^*)$, o tamanho necessário da amostra pode ser obtido notando que

$$
\pi^* = \pi(p^*) \Rightarrow 1 - \pi^* = \Phi\left(\frac{z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}} - (p^*-p_0)}{\sqrt{\frac{p^*(1-p^*)}{n}}}\right) - \Phi\left(\frac{-z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}} - (p^*-p_0)}{\sqrt{\frac{p^*(1-p^*)}{n}}}\right) \tag{6.7}
$$

[2] z_α é o valor crítico da $N(0;1)$ tal que $\mathrm{P}(Z > z_\alpha) = \alpha$

Se $p^* - p_0 > 0$, $\Phi\left(\dfrac{-z_{\alpha/2}\sqrt{\dfrac{p_0(1-p_0)}{n}} - (p^* - p_0)}{\sqrt{\dfrac{p^*(1-p^*)}{n}}}\right) \approx 0$ e, portanto,

$$1 - \pi^* \approx \Phi\left(\frac{z_{\alpha/2}\sqrt{\dfrac{p_0(1-p_0)}{n}} - (p^* - p_0)}{\sqrt{\dfrac{p^*(1-p^*)}{n}}}\right) \Rightarrow$$

$$\pi^* \approx \mathrm{P}\left(Z > \frac{z_{\alpha/2}\sqrt{\dfrac{p_0(1-p_0)}{n}} - (p^* - p_0)}{\sqrt{\dfrac{p^*(1-p^*)}{n}}}\right) = \mathrm{P}\left(Z > z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{p^*(1-p^*)}} - \frac{p^* - p_0}{\sqrt{\dfrac{p^*(1-p^*)}{n}}}\right)$$

Daí resulta que

$$z_{\pi^*} = z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{p^*(1-p^*)}} - \frac{p^* - p_0}{\sqrt{\dfrac{p^*(1-p^*)}{n}}} \Rightarrow$$

$$\sqrt{n}\frac{p^* - p_0}{\sqrt{p^*(1-p*)}} = z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{p^*(1-p^*)}} - z_{\pi^*} \Rightarrow$$

$$\sqrt{n} = z_{\alpha/2}\frac{\sqrt{p_0(1-p_0)}}{p^* - p_0} - z_{\pi^*}\frac{\sqrt{p^*(1-p^*)}}{p^* - p_0}$$

e, portanto

$$n \approx \left(\frac{z_{\alpha/2}\sqrt{p_0(1-p_0)} - z_{\pi^*}\sqrt{p^*(1-p^*)}}{p^* - p_0}\right)^2 \tag{6.8}$$

Voltando à equação 6.7, se $p^* - p_0 < 0$, $\Phi\left(\dfrac{z_{\alpha/2}\sqrt{\dfrac{p_0(1-p_0)}{n}} - (p^* - p_0)}{\sqrt{\dfrac{p^*(1-p^*)}{n}}}\right) \approx 1$ o que implica que

$$\pi^* \approx \Phi\left(\frac{-z_{\alpha/2}\sqrt{\dfrac{p_0(1-p_0)}{n}} - (p^* - p_0)}{\sqrt{\dfrac{p^*(1-p^*)}{n}}}\right) = \mathrm{P}\left(Z > \frac{z_{\alpha/2}\sqrt{\dfrac{p_0(1-p_0)}{n}} + (p^* - p_0)}{\sqrt{\dfrac{p^*(1-p^*)}{n}}}\right) \Rightarrow$$

$$z_{\pi^*} = \frac{z_{\alpha/2}\sqrt{\dfrac{p_0(1-p_0)}{n}} + (p^* - p_0)}{\sqrt{\dfrac{p^*(1-p^*)}{n}}} = z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{p^*(1-p^*)}} + \sqrt{n}\frac{p^* - p_0}{\sqrt{p^*(1-p^*)}} \Rightarrow$$

$$\sqrt{n} = z_{\pi^*}\frac{\sqrt{p^*(1-p^*)}}{p^* - p_0} - z_{\alpha/2}\frac{\sqrt{\dfrac{p_0(1-p_0)}{p^*(1-p^*)}}}{\dfrac{p^* - p_0}{\sqrt{p^*(1-p^*)}}} = z_{\pi^*}\frac{\sqrt{p^*(1-p^*)}}{p^* - p_0} - z_{\alpha/2}\frac{\sqrt{p_0(1-p_0)}}{p^* - p_0}$$

e isso resulta em

$$n \approx \left(\frac{z_{\pi^*}\sqrt{p^*(1-p^*)} - z_{\alpha/2}\sqrt{p_0(1-p_0)}}{p^* - p_0} \right)^2 \tag{6.9}$$

mesma expressão obtida anteriormente.

6.5.4 Poder do teste Z unilateral para proporções

Teste unilateral à direita

Consideremos inicialmente o teste para

$$\begin{aligned} H_0 &: p = p_0 \\ H_1 &: p > p_0 \end{aligned}$$

Para um nível de significância α, a função poder do teste é

$$\begin{aligned} \pi(p) &= \mathrm{P}\left(\frac{\widehat{P} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} > z_\alpha | p \right) \\ &= \mathrm{P}\left(\frac{\widehat{P} - p + p - p_0}{\sqrt{\dfrac{p(1-p)}{n}}} \frac{\sqrt{\dfrac{p(1-p)}{n}}}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} > z_\alpha \mid p \right) \\ &= \mathrm{P}\left(\frac{\widehat{P} - p + p - p_0}{\sqrt{\dfrac{p(1-p)}{n}}} > z_\alpha \frac{\sqrt{\dfrac{p_0(1-p_0)}{n}}}{\sqrt{\dfrac{p(1-p)}{n}}} \mid p \right) \\ &= \mathrm{P}\left(Z > \frac{z_\alpha\sqrt{\dfrac{p_0(1-p_0)}{n}} - (p - p_0)}{\sqrt{\dfrac{p(1-p)}{n}}} \right) \end{aligned}$$

ou seja,

$$\pi(p) = 1 - \Phi\left(Z > \frac{z_\alpha\sqrt{\dfrac{p_0(1-p_0)}{n}} - (p - p_0)}{\sqrt{\dfrac{p(1-p)}{n}}} \right) \tag{6.10}$$

Teste unilateral à esquerda

Para o teste

$$\begin{aligned} H_0 &: p = p_0 \\ H_1 &: p < p_0 \end{aligned}$$

a função poder do teste é

$$
\begin{aligned}
\pi(p) &= \mathrm{P}\left(\frac{\widehat{P}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} < -z_\alpha \,|\, p\right) \\
&= \mathrm{P}\left(\frac{\widehat{P}-p+p-p_0}{\sqrt{\frac{p(1-p)}{n}}}\frac{\sqrt{\frac{p(1-p)}{n}}}{\sqrt{\frac{p_0(1-p_0)}{n}}} < -z_\alpha \,|\, p\right) \\
&= \mathrm{P}\left(\frac{\widehat{P}-p+p-p_0}{\sqrt{\frac{p(1-p)}{n}}} < -z_\alpha\frac{\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}} \,|\, p\right) \\
&= \mathrm{P}\left(Z < -\frac{z_\alpha\sqrt{\frac{p_0(1-p_0)}{n}}-(p-p_0)}{\sqrt{\frac{p(1-p)}{n}}}\right) \\
&= \mathrm{P}\left(Z > \frac{z_\alpha\sqrt{\frac{p_0(1-p_0)}{n}}+(p-p_0)}{\sqrt{\frac{p(1-p)}{n}}}\right)
\end{aligned}
$$

ou seja,

$$
\pi(p) = 1-\Phi\left(\frac{z_\alpha\sqrt{\frac{p_0(1-p_0)}{n}}+(p-p_0)}{\sqrt{\frac{p(1-p)}{n}}}\right) \tag{6.11}
$$

Tamanho da amostra

Se queremos $\pi(p^*) = \pi^*$ em um teste unilateral à direita, podemos ver, de (6.10), que

$$
\begin{aligned}
&\pi(p^*) = \pi^* = \mathrm{P}\left(Z > \frac{z_\alpha\sqrt{\frac{p_0(1-p_0)}{n}}-(p^*-p_0)}{\sqrt{\frac{p^*(1-p^*)}{n}}}\right) \Rightarrow \\
&z_{\pi^*} = z_\alpha\sqrt{\frac{p_0(1-p_0)}{p^*(1-p^*)}} - \frac{p^*-p_0}{\sqrt{\frac{p^*(1-p^*)}{n}}} \Rightarrow \\
&\sqrt{n}\frac{p^*-p_0}{\sqrt{p^*(1-p^*)}} = z_\alpha\sqrt{\frac{p_0(1-p_0)}{p^*(1-p^*)}} - z_{\pi^*} \Rightarrow \\
&\sqrt{n} = z_\alpha\frac{\sqrt{p_0(1-p_0)}}{p^*-p_0} - z_{\pi^*}\frac{\sqrt{p^*(1-p^*)}}{p^*-p_0}
\end{aligned}
$$

ou seja,

$$
n = \left(\frac{z_\alpha\sqrt{p_0(1-p_0)} - z_{\pi^*}\sqrt{p^*(1-p^*)}}{p^*-p_0}\right)^2 \tag{6.12}
$$

Se queremos $\pi(p^*) = \pi^*$ em um teste unilateral à esquerda, podemos ver, de (6.11), que

$$
\begin{aligned}
\pi(p^*) = \pi^* = \mathrm{P}\left(Z > \frac{z_\alpha \sqrt{\dfrac{p_0(1-p_0)}{n}} + (p^* - p_0)}{\sqrt{\dfrac{p^*(1-p^*)}{n}}} \right) \Rightarrow \\
z_{\pi^*} = z_\alpha \sqrt{\frac{p_0(1-p_0)}{p^*(1-p^*)}} + \frac{p^* - p_0}{\sqrt{\dfrac{p^*(1-p^*)}{n}}} \Rightarrow \\
\sqrt{n}\frac{p^* - p_0}{\sqrt{p^*(1-p^*)}} = z_{\pi^*} - z_\alpha \sqrt{\frac{p_0(1-p_0)}{p^*(1-p^*)}} - \Rightarrow \\
\sqrt{n} = z_{\pi^*}\frac{\sqrt{p^*(1-p^*)}}{p^* - p_0} - z_\alpha \frac{\sqrt{p_0(1-p_0)}}{p^* - p_0}
\end{aligned}
$$

ou seja,

$$
n = \left(\frac{z_{\pi^*}\sqrt{p^*(1-p^*)} - z_\alpha \sqrt{p_0(1-p_0)}}{p^* - p_0} \right)^2 \tag{6.13}
$$

Note que essa é a mesma expressão obtida para o teste unilateral à direita. Esse é um resultado razoável tendo em vista a simetria dos dois testes.

6.6 Intervalo de confiança e teste de hipótese

Embora os objetivos dos testes de hipóteses e dos intervalos de confiança sejam diferentes, é possível estabelecer uma relação entre eles no sentido de se tomar decisão de rejeitar ou não H_0 com base em intervalos de confiança. Mas é importante notar a seguinte diferença: nos intervalos de confiança, dada uma estimativa do parâmetro, obtemos um intervalo para os possíveis valores do parâmetro. Nos testes de hipóteses, dado um valor θ_0 do parâmetro, obtemos uma faixa de valores (região crítica), para a estatística de teste, que leva à rejeição da hipótese nula.

Para estudar a relação entre testes de hipóteses e intervalos de confiança, vamos considerar o caso de inferência sobre a média de uma população normal com variância conhecida.

Vimos que o intervalo de confiança superior tem a forma $\left(-\infty, \overline{X} + z_\alpha \dfrac{\sigma}{\sqrt{n}}\right)$. Então, podemos dizer que

$$
\mathrm{P}\left(\mu \in \left(-\infty, \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}} \right) \right) = 1 - \alpha
$$

ou seja, para $100 \times (1 - \alpha)\%$ dos intervalos assim construídos a partir de todas as possíveis amostras o verdadeiro parâmetro μ estará contido no intervalo. Mas isso é equivalente a

$$
\mathrm{P}\left(\mu \geq \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}} \right) = \alpha
$$

que é equivalente a

$$
\mathrm{P}\left(\overline{X} \leq \mu - z_\alpha \frac{\sigma}{\sqrt{n}} \right) = \alpha
$$

Mas essa é a região crítica do teste de hipótese unilateral à esquerda! Assim, se μ_0 não pertencer ao intervalo de confiança superior, rejeitamos $H_0 : \mu = \mu_0$ em favor de $H_1 : \mu < \mu_0$.

Analogamente, para o intervalo de confiança inferior $\left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}} \; +\infty\right)$, podemos dizer que

$$
\begin{aligned}
&\mathrm{P}\left(\mu \in \left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}\,, +\infty\right)\right) = 1 - \alpha \Leftrightarrow \\
&\mathrm{P}\left(\mu \leq \overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}\right) = \alpha \Leftrightarrow \\
&\mathrm{P}\left(\overline{X} \geq \mu + z_\alpha \frac{\sigma}{\sqrt{n}}\right) = \alpha
\end{aligned}
$$

e essa é a região crítica do teste de hipótese unilateral à direita! Assim, se μ_0 não pertencer ao intervalo de confiança inferior, rejeitamos $H_0 : \mu = \mu_0$ em favor de $H_1 : \mu > \mu_0$.

Para o intervalo de confiança bilateral, temos

$$
\begin{aligned}
&\mathrm{P}\left(\mu \in \left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}\,, \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right)\right) = 1 - \alpha \Leftrightarrow \\
&\mathrm{P}\left[\left(\mu \leq \overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}\right) \cup \left(\mu \geq \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right)\right] = \alpha \Leftrightarrow \\
&\mathrm{P}\left[\left(\overline{X} \geq \mu + z_\alpha \frac{\sigma}{\sqrt{n}}\right) \cup \left(\overline{X} \leq \mu - z_\alpha \frac{\sigma}{\sqrt{n}}\right)\right] = \alpha
\end{aligned}
$$

e essa é a região crítica do teste de hipótese bilateral! Assim, se μ_0 não pertencer ao intervalo de confiança bilateral, rejeitamos $H_0 : \mu = \mu_0$ em favor de $H_1 : \mu \neq \mu_0$.

Resumindo, temos as seguintes associações:

Intervalo de Confiança (I.C.)		Se $\mu_0 \notin$ I.C., rejeito $H_0 : \mu = \mu_0$ em favor de
Superior:	$\left(-\infty\,, \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right)$	$H_1 : \mu < \mu_0$
Inferior:	$\left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}\,, +\infty\right)$	$H_1 : \mu > \mu_0$
Bilateral:	$\left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}\,, \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right)$	$H_1 : \mu \neq \mu_0$

EXEMPLO 6.11

Vamos considerar uma população normal com desvio padrão $\sigma = 2$, da qual se extrai uma aas de tamanho $n = 16$. Suponhamos que o valor observado da média amostral seja $\overline{x} = 2$. Então, o intervalo de confiança superior de nível $1 - \alpha = 0,95$ é

$$
\left(-\infty\,;\, 2 + 1,64 \times \frac{2}{\sqrt{16}}\right) = (-\infty\,;\, 2,82)
$$

que está representado pela linha vermelha horizontal na Figura 6.12 e o intervalo de confiança inferior é

$$\left(2 - 1,64 \times \frac{2}{\sqrt{16}}\,;\, +\infty\right) = (1,18\,;\, +\infty)$$

que está representado pela linha vermelha horizontal na Figura 6.13.

Em ambas as figuras, as diferentes curvas normais representam a distribuição amostral da estatística de teste $Z_0 = \sqrt{n}\frac{\overline{X}-\mu_0}{\sigma}$ para diferentes valores de μ_0 e a média amostral observada $\overline{x} = 2$ está indicada pela linha pontilhada em vermelho.

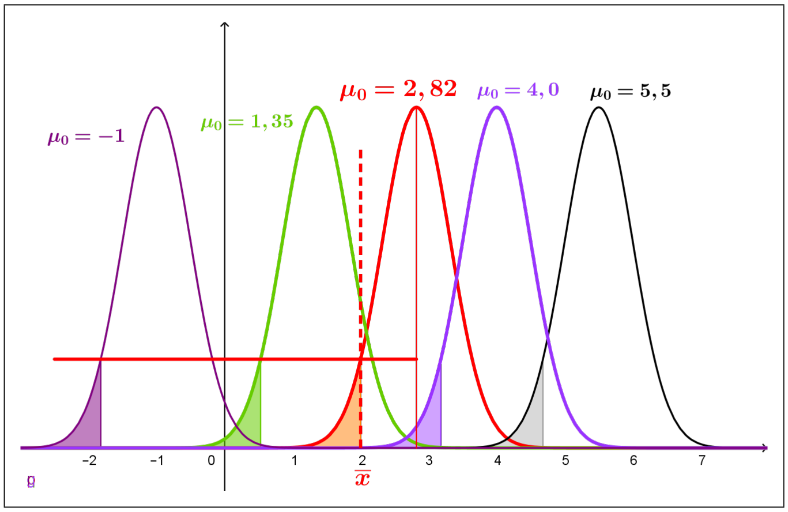


Figura 6.12 – Intervalo de confiança unilateral superior e teste de hipótese unilateral à esquerda

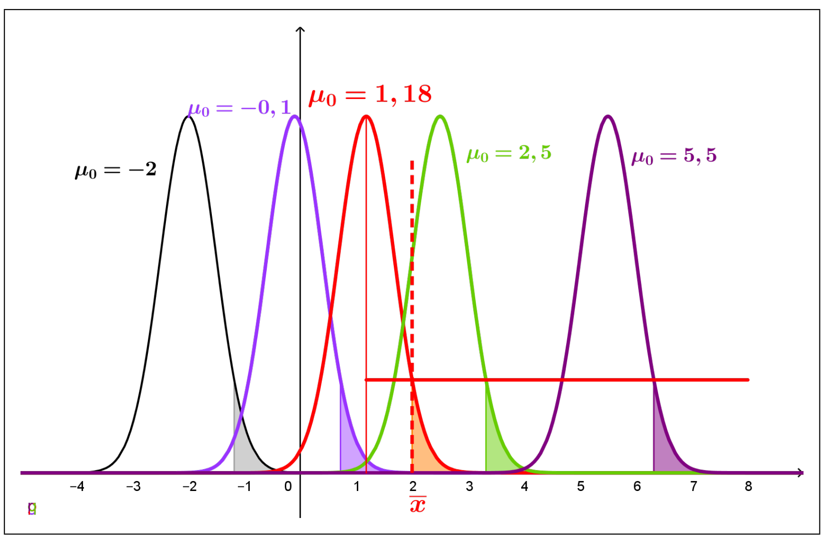


Figura 6.13 – Intervalo de confiança unilateral inferior e teste de hipótese unilateral à direita

Vamos analisar a Figura 6.12, onde temos o intervalo de confiança superior. A área sombreada nas caudas das diferentes curvas normais representa a região crítica para o teste com $H_1 : \mu < \mu_0$. Para valores de μ_0 fora do intervalo de confiança ($\mu_0 = 5,5$ e $\mu_0 = 4,0$), o valor de $\overline{x}$ está na região crítica, ou seja, rejeitamos $H_0 : \mu = \mu_0$ em favor de $H_1 : \mu < \mu_0$. Para valores de μ_0 dentro do intervalo de confiança, ($\mu_0 = 1,35$ e $\mu_0 = -1$), o valor de $\overline{x}$ não está na região crítica, ou seja, não rejeitamos $H_0 : \mu + \mu_0$. O ponto de "corte" (curva normal em vermelho) é exatamente

$$\mu_0 = 2,82 = 2,0 + 1,64 \times \frac{2}{\sqrt{16}}$$

.

Vamos analisar a Figura 6.13, onde temos o intervalo de confiança inferior. A área sombreada nas caudas das diferentes curvas normais representa a região crítica para o teste com $H_1 : \mu > \mu_0$. Para valores de μ_0 fora do intervalo de confiança ($\mu_0 = -2$ e $\mu_0 = -0,1$), o valor de $\overline{x}$ está na região crítica, ou seja, rejeitamos $H_0 : \mu = \mu_0$ em favor de $H_1 : \mu > \mu_0$. Para valores de μ_0 dentro do intervalo de confiança, ($\mu_0 = 2,55$ e $\mu_0 = 5,5$), o valor de $\overline{x}$ não está na região crítica, ou seja, não rejeitamos $H_0 : \mu + \mu_0$. O ponto de "corte" (curva normal em vermelho) é exatamente

$$\mu_0 = 1,18 = 2,0 - 1,64 \times \frac{2}{\sqrt{16}}$$

.

6.7 Exercícios propostos

1. Uma amostra aleatória simples de tamanho $n = 9$ extraída de uma população normal com desvio padrão 3,1 apresentou média igual a $\overline{x} = 13,35$. Deseja-se testar

$$\begin{aligned} H_0 &: \mu = 12,8 \\ H_1 &: \mu \neq 12,8 \end{aligned}$$

   (a) Determine a região crítica correspondente ao nível de significância $\alpha = 0,02$.

   (b) Com base na região crítica encontrada no item anterior, estabeleça a conclusão, tendo o cuidado de usar um fraseado que não seja puramente técnico.

   (c) Calcule o valor P e interprete o resultado obtido.

2. Em uma linha de produção, peças são produzidas de modo que o comprimento seja normalmente distribuído com desvio padrão de 0,5 cm. Ajustes periódicos são feitos na máquina para garantir que as peças tenham comprimento apropriado de 15 cm, pois as peças muito curtas não podem ser aproveitadas (as peças longas podem ser cortadas). A cada hora são extraídas 9 peças da produção, medindo-se seu comprimento. Estabeleça uma regra de decisão para definir se o processo está operando adequadamente. Use o nível de significância de 0,1%.

3. Depois de desenvolver um algoritmo para acelerar a execução de determinada tarefa rotineira em um escritório de contabilidade, o analista de sistema analisa uma amostra de 25 tempos, obtendo uma média 46,5 segundos. Dos dados passados, ele sabe que o tempo de execução é aproximadamente normal com média de 48,5 segundos e desvio padrão de 5 segundos. Use o método do valor P para decidir se o algoritmo do analista realmente melhorou o desempenho do sistema.

4. Uma propaganda afirma que o consumo médio de gasolina de determinada marca de automóvel é de 12 litros por 100 quilômetros rodados, com desvio padrão de 1,0 litro. Um teste com 36 automóveis desta marca acusa um consumo médio de 12,4 litros por 100 quilômetros rodados. O que se pode concluir sobre a propaganda?

5. Considere uma população normal com variância 225, da qual se extrai uma amostra aleatória simples de tamanho 25. Deseja-se testar as seguintes hipóteses:

$$\begin{aligned} H_0 &: \mu = 40 \\ H_1 &: \mu \neq 40 \end{aligned}$$

   e para isso define-se a seguinte região crítica:

$$RC : \overline{X} > 46 \text{ ou } \overline{X} < 34.$$

(a) Calcule a probabilidade do erro tipo I.

(b) Obtenha a expressão geral para a função poder do teste.

(c) Calcule o poder do teste para os seguintes valores de $\mu : 20, 22, 24, \ldots, 56, 58, 60$.

(d) Esboce o gráfico da função poder.

6. Em uma pesquisa com 800 estudantes universitários, 385 afirmaram possuir computador. Teste a hipótese de que pelo menos 50% dos estudantes universitários possuem computador. Use $\alpha = 0,10$.

7. Uma pesquisa entre 700 trabalhadores revela que 12,3% obtiveram seus empregos através de indicações de amigos ou parentes. Teste a hipótese de que mais de 10% dos trabalhadores conseguem seus empregos por indicação de amigos ou parentes, utilizando 5% como nível de significância.

8. O nível de aprovação da qualidade das refeições servidas em um restaurante universitário era de 20%, quando houve uma movimentação geral dos estudantes que forçou a direção do restaurante a fazer mudanças. Feitas essas mudanças, sorteia-se uma amostra de 64 estudantes usuários do restaurante e 25 aprovam a qualidade da comida. Você diria, ao nível de significância de 5%, que as mudanças surtiram efeito?

9. Deseja-se testar a honestidade de uma moeda. Para isso, lança-se a moeda 200 vezes, obtendo-se 115 caras. Qual é a sua conclusão sobre a honestidade da moeda? Para responder a essa questão, calcule e interprete o valor P.

10. A direção de um grande jornal nacional afirma que 25% dos seus leitores são da classe A. Se, em uma amostra de 740 leitores, encontramos 156 da classe A, qual é a conclusão que tiramos sobre a afirmativa da direção do jornal?

Capítulo 7

Mais sobre testes de hipóteses para parâmetros da $N(\mu; \sigma^2)$

Assim como na construção de intervalos de confiança para parâmetros de uma população normal, faremos uso dos resultados sobre as distribuições amostrais de $\overline{X}$ e S^2 vistos no Capítulo para construir testes de hipóteses sobre a média μ e a variância σ^2 de uma população normal. Vimos que, se $X_1, X_2, \cdots, X_n$ é uma amostra aleatória simples de uma população $X \sim N(\mu; \sigma^2)$, então

$$\frac{(n-1)\,S^2}{\sigma^2} \sim \chi^2_{n-1} \tag{7.1}$$

$$\sqrt{n}\frac{\overline{X}-\mu}{S} \sim t_{n-1} \tag{7.2}$$

Essas serão nossas estatísticas de teste e, como antes, a região crítica será formada pelos valores pouco prováveis dessas estatísticas de teste sob a hipótese de veracidade de H_0.

7.1 Teste de hipótese sobre a variância σ^2

A hipótese nula que iremos considerar é

$$H_0 : \sigma^2 = \sigma_0^2$$

e a hipótese alternativa pode tomar uma das três formas:

$$H_1 : \sigma^2 \neq \sigma_0^2 \qquad H_1 : \sigma^2 > \sigma_0^2 \qquad H_1 : \sigma^2 < \sigma_0^2$$

Como antes, a escolha entre essas três possibilidades se faz com base no conhecimento do problema e deve ser feita antes de se conhecer o resultado da amostra. Se não temos informação alguma sobre a alternativa, temos que usar um teste bilateral.

A regra de decisão consiste em definir a região crítica RC como o conjunto de valores cuja probabilidade de ocorrência é *pequena* sob a hipótese de veracidade de H_0 e a estatística de teste é

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

Os valores com pequena probabilidade de ocorrência estão nas caudas da distribuição e isso nos leva às seguintes regras de decisão.

- Hipótese nula e estatística de teste

$$H_0 : \sigma^2 = \sigma_0^2$$

$$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2} \underbrace{\sim}_{\text{sob } H_0} \chi^2(n-1)$$

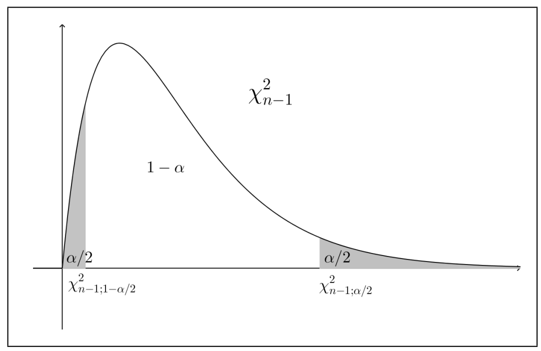
$$c_0^2 = \text{ valor observado de } \chi_0^2$$

- Teste bilateral

$H_1 : \sigma^2 \neq \sigma_0^2$

Região crítica:
$\chi_0^2 < \chi^2_{n-1;1-\alpha/2}$ ou $X_0^2 > \chi^2_{n-1,\alpha/2}$

Valor P :
$P = 2 \cdot \min\left[\mathrm{P}(\chi^2_{n-1} > c_0^2), \mathrm{P}(\chi^2_{n-1} < c_0^2)\right]$

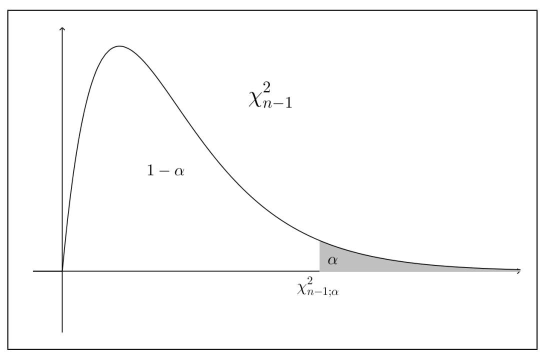


- Teste unilateral à direita

$H_1 : \sigma^2 > \sigma_0^2$

Região crítica:
$\chi_0^2 > \chi^2_{n-1,\alpha}$

Valor P :
$P = \mathrm{P}(\chi^2_{n-1} > c_0^2)$

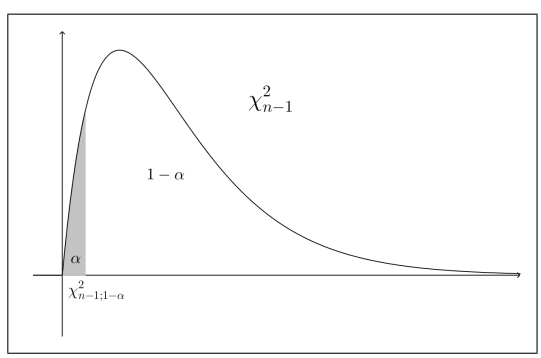


- Teste unilateral à esquerda

$H_1 : \sigma^2 < \sigma_0^2$

Região crítica:
$\chi_0^2 < \chi^2_{n-1,1-\alpha}$

Valor P :
$P = \mathrm{P}(\chi^2_{n-1} < c_0^2)$



Como antes, a definição do valor P é a probabilidade de se obter um valor tão ou mais extremo como o observado na amostra em estudo, supondo-se H_0 verdadeira. Para o cálculo exato do valor P de um teste de hipótese sobre a variância de uma população normal, é necessário um programa computacional. A partir da Tabela III do Apêndice A, podemos obter apenas limites para o valor P, conforme ilustrado nos exemplos a seguir.

EXEMPLO 7.1

Uma amostra aleatória simples de tamanho $n = 16$ foi retirada de uma população normal, obtendo-se $s^2 = 32,1$. Ao nível de significância de 5% pode-se dizer que $\sigma^2 \neq 20$?

Solução

As hipóteses são

$$\begin{array}{rcl} H_0 & : & \sigma^2 = 20 \\ H_1 & : & \sigma^2 \neq 20 \end{array}$$

Com 15 graus de liberdade, teste bilateral e nível de significância de 5%, os valores críticos necessários são

$$\begin{array}{rcl} \chi^2_{15;0,975} & = & 6,262 \\ \chi^2_{15;0,025} & = & 27,488 \end{array}$$

e a região crítica é

$$\chi_0^2 > 27,488 \qquad \text{ou} \qquad \chi_0^2 < 6,262$$

O valor observado da estatística de teste é

$$c_0^2 = \frac{15 \times 32,1}{20} = 24,075$$

que não pertence à região crítica. Logo, não se rejeita a hipótese nula, ou seja, não podemos afirmar que $\sigma^2 \neq 20$.

Olhando na Tabela III, na linha correspondente a 15 graus de liberdade, vemos que 24,075 está entre os valores 22,307 e 24,996, que correspondem às probabilidades 0,10 e 0,05, respectivamente. Logo, o valor $P/2$ é tal que $0,05 < P/2 < 0,10$ e, portanto, $0,10 < P < 0,20$.

Com auxílio do Minitab, obtemos que $\mathrm{P}(\chi^2_{15} \leq 24,075) = 0,936169$. Logo, o valor P é

$$P = 2 \cdot \min(0,936169, 1 - 0,936169) = 2 \cdot (1 - 0,936169) = 0,127662.$$

EXEMPLO 7.2

O gerente de um posto de abastecimento de combustível muito utilizado por caminhoneiros realiza uma pesquisa entre esses clientes com o objetivo de planejar esquemas de trabalho e de suprimento de diesel. Relatórios do sindicato nacional indicam que a quantidade média de diesel comprada por semana é de 1310 litros, com desvio padrão de 89,4 litros. Para uma amostra de 20 caminhoneiros, o gerente obteve os seguintes dados sobre a quantidade de diesel comprada semanalmente:

1283	1317	1226	1298	1382	1344	1314	1298	1298	1355
1242	1234	1298	1355	1287	1253	1234	1344	1295	1321

(a) Há alguma evidência que sugira que a verdadeira variância populacional no combustível diesel comprado por semana nesse posto seja diferente de 7900 l^2? Suponha normalidade e use $\alpha = 0,01$.

(b) Ache limites para o valor P associado a esse teste de hipótese.

Solução

(a) As hipótese são

$$H_0 : \sigma^2 = 7900$$
$$H_1 : \sigma^2 \neq 7900$$

Os dados fornecem

$$\sum_{i=1}^{20} x_i = 25978 \qquad \sum_{i=1}^{20} x_i^2 = 33780632$$

Logo,

$$s^2 = \frac{1}{19}\left(33780632 - \frac{25978^2}{20}\right) = 1989,884211$$

Os valores críticos para o teste de hipótese são $\chi^2_{19;0,005} = 38,582$ e $\chi^2_{19;0,995} = 6,844$ e o valor observado da estatística de teste é

$$c_0^2 = \frac{19 \times 1989,884211}{7900} = 4,7858$$

Como esse valor está na região crítica, rejeita-se H_0, ou seja, há evidências de que a variância seja diferente de 7900 l^2.

(b) Da Tabela IV do Apêndice A, vemos que

$$4,7858 < 4,912$$

Logo, $P/2 < 0,0005$, ou seja, $P < 0,0010$.

Com auxílio do Minitab, obtemos que $\mathrm{P}(\chi^2_{19} \leq 4,7858) = 0,0004126$. Logo, o valor P é

$$P = 2 \cdot \min(0,0004126, 1 - 0,0004126) = 2 \cdot 00004126 = 0,0008252.$$

7.1.1 Poder do teste qui-quadrado para σ^2

Consideremos o teste bilateral

$$\begin{aligned} H_0 &: \sigma^2 = \sigma_0^2 \\ H_1 &: \sigma^2 \neq \sigma_0^2 \end{aligned}$$

O poder do teste é

$$\begin{aligned}
\pi(\sigma^2) &= \mathrm{P}(\text{rejeitar } H_0 \,|\, \sigma^2) \\
&= \mathrm{P}\left(\frac{(n-1)S^2}{\sigma_0^2} < \chi^2_{n-1;1-\alpha/2} \,|\, \sigma^2\right) + \mathrm{P}\left(\frac{(n-1)S^2}{\sigma_0^2} > \chi^2_{n-1;\alpha/2} \,|\, \sigma^2\right) \\
&= \mathrm{P}\left(\frac{(n-1)S^2}{\sigma^2 \cdot \dfrac{\sigma_0^2}{\sigma^2}} < \chi^2_{n-1;1-\alpha/2} \,|\, \sigma^2\right) + \mathrm{P}\left(\frac{(n-1)S^2}{\sigma^2 \cdot \dfrac{\sigma_0^2}{\sigma^2}} > \chi^2_{n-1;\alpha/2} \,|\, \sigma^2\right) \Longrightarrow
\end{aligned}$$

$$\pi(\sigma^2) = \mathrm{P}\left(\chi^2_{n-1} < \frac{\sigma_0^2}{\sigma^2} \cdot \chi^2_{n-1;1-\alpha/2}\right) + \mathrm{P}\left(\chi^2_{n-1} > \frac{\sigma_0^2}{\sigma^2} \cdot \chi^2_{n-1;\alpha/2}\right) \tag{7.3}$$

Para o teste unilateral $H_1 : \sigma^2 > \sigma_0^2$, temos

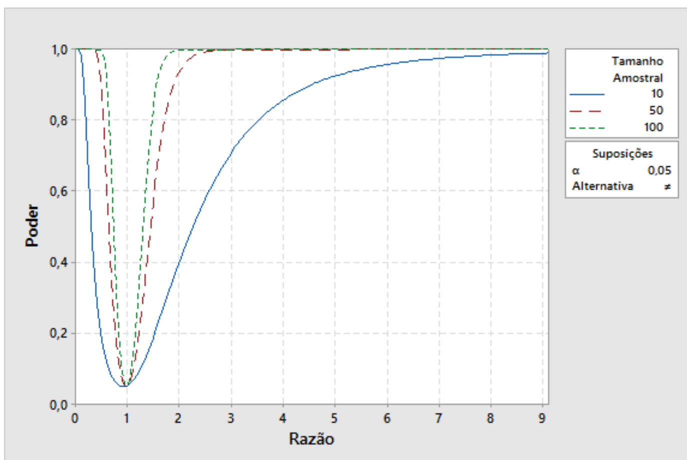
$$\pi(\sigma^2) \quad = \quad \mathrm{P}\left(\frac{(n-1)S^2}{\sigma_0^2} > \chi^2_{n-1;\alpha} \mid \sigma^2\right) = \mathrm{P}\left(\frac{(n-1)S^2}{\sigma^2 \cdot \dfrac{\sigma_0^2}{\sigma^2}} > \chi^2_{n-1;\alpha} \mid \sigma^2\right) \Longrightarrow$$

$$\pi(\sigma^2) = \mathrm{P}\left(\chi^2_{n-1} > \frac{\sigma_0^2}{\sigma^2} \cdot \chi^2_{n-1;\alpha}\right) \tag{7.4}$$

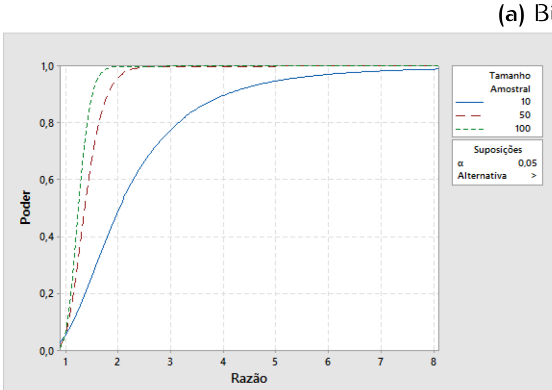
Para o teste unilateral à esquerda, raciocínio análogo leva a

$$\pi(\sigma^2) = \mathrm{P}\left(\chi^2_{n-1} < \frac{\sigma_0^2}{\sigma^2} \cdot \chi^2_{n-1;1-\alpha}\right) \tag{7.5}$$

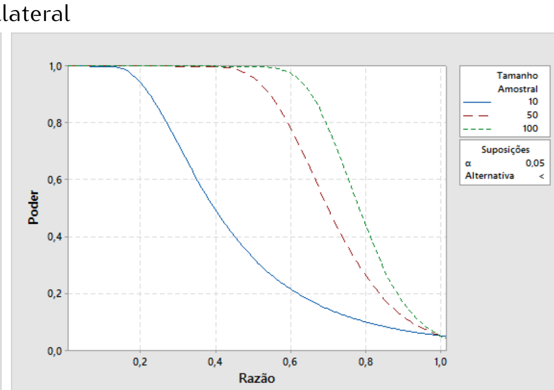
Na Figura 7.1 ilustram-se as funções poder dos testes qui-quadrado bi e unilaterais.



(a) Bilateral



(b) Unilateral à direita



(c) Unilateral à esquerda

Figura 7.1 – Poder do teste χ^2 unilateral em função de $\Delta = \dfrac{\sigma^2}{\sigma_0^2}$

7.2 Teste de hipótese sobre a média μ

Agora vamos generalizar os resultados da Seção 6.2, considerando uma população $X \sim N(\mu; \sigma^2)$ em que a variância σ^2 não é conhecida. O procedimento de teste de hipóteses sobre a média de uma população normal quando a variância não é conhecida é absolutamente análogo ao caso em que conhecemos σ^2. A mudança diz respeito à estatística de teste e sua distribuição, que agora passam a ser

$$T = \sqrt{n}\frac{\overline{X} - \mu}{S} \sim t(n-1)$$

A seguir apresentamos os resultados pertinentes para cada um dos tipos de hipótese alternativa.

- Hipótese nula e estatística de teste

$$H_0 : \mu = \mu_0$$

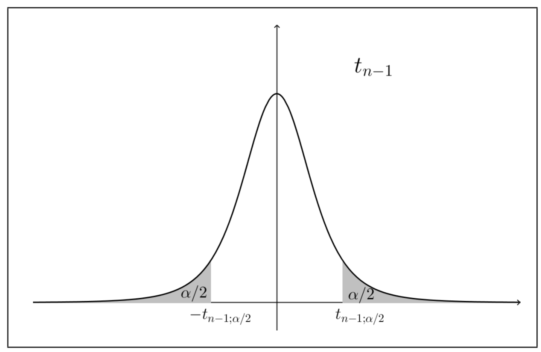
$$T_0 = \sqrt{n}\frac{\overline{X} - \mu_0}{S} \underbrace{\sim}_{\text{sob } H_0} t(n-1)$$

$$t_0 = \text{ valor observado de } T_0$$

- Teste bilateral

$H_1 : \mu \neq \mu_0$

Região crítica:
$T_0 < -t_{n-1,\alpha/2}$ ou $T_0 > t_{n-1,\alpha/2}$
Valor P :
$P = 2 \cdot \mathrm{P}(T_{n-1} > |t_0|)$

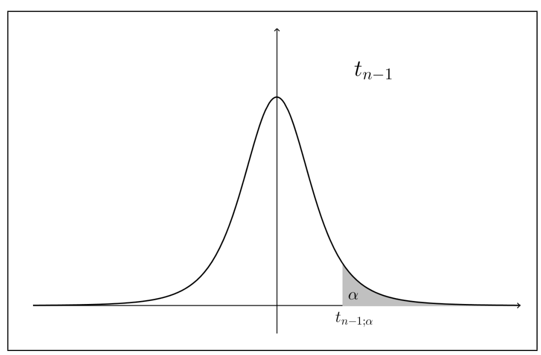


- Teste unilateral à direita

$H_1 : \mu > \mu_0$

Região crítica:
$T_0 > t_{n-1,\alpha}$

Valor P :
$P = \mathrm{P}(T_{n-1} > t_0)$

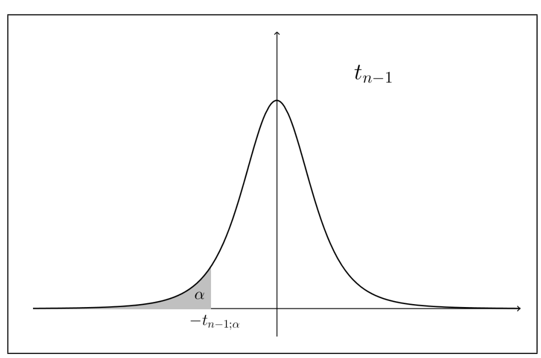


- Teste unilateral à esquerda

$H_1 : \mu < \mu_0$

Região crítica:
$T_0 < -t_{n-1,\alpha}$

Valor P :
$P = \mathrm{P}(T_{n-1} < t_0)$



A definição do valor P é exatamente a mesma, mas para o cálculo exato é necessário um programa computacional. A partir da Tabela III do Apêndice A, podemos obter apenas limites para o valor P, conforme ilustrado nos exemplos a seguir.

EXEMPLO 7.3

Depois de uma pane geral no sistema de informação de uma empresa, o gerente administrativo deseja saber se houve alteração no tempo de processamento de determinada atividade. Antes da pane, o tempo de processamento podia ser aproximado por uma variável aleatória normal com média de 100 minutos. Uma amostra de 16 tempos de processamento após a pane revela uma média $\overline{x} = 105,5$ minutos e um desvio padrão $s = 10$ minutos. Ao nível de significância de 5%, qual é a conclusão sobre a alteração do tempo médio de processamento?

Solução

Como visto, as hipóteses do problema são

$$\begin{aligned} \mu &= 100 \\ \mu &\neq 100 \end{aligned}$$

Como a segunda expressão não envolve o sinal de igualdade, ela se torna a hipótese alternativa:

$$\begin{aligned} H_0 &: \mu = 100 \\ H_1 &: \mu \neq 100 \end{aligned}$$

Como a população é normal e a variância não é conhecida, temos que usar a distribuição t de Student com $n-1 = 16-1 = 15$ graus de liberdade. Para um teste bilateral com nível de significância de 5%, a abscissa de interesse é aquela que deixa área de 0,025 acima. Consultando a Tabela IV do Apêndice A, resulta

$$t_{15;\,0,025} = 2,131$$

A estatística de teste é

$$T_0 = \frac{\overline{X} - 100}{\frac{10}{\sqrt{16}}} \sim t(15)$$

e a região crítica é

$$T_0 > 2,131 \text{ ou } T_0 < -2,131$$

O valor observado da estatística de teste é

$$t_0 = \frac{105,5 - 100}{\frac{10}{\sqrt{16}}} = 2,2$$

Como esse valor pertence à região crítica, rejeitamos a hipótese nula e concluímos que houve alteração no tempo de processamento após a pane.

O valor P é, por definição,

$$P = 2 \times \mathrm{P}(t_{15} > 2,2)$$

Olhando na tabela na linha correspondente a 15 graus de liberdade, vemos que o valor 2,2 está entre 2,131 e 2,602, que correspondem às probabilidades 0,025 e 0,01. Logo

$$0,01 < \mathrm{P}(t_{15} > 2,2) < 0,025 \Rightarrow 0,02 < P < 0,05$$

O valor exato é $P/2 \underbrace{=}_{\text{Excel}}$ `1-dist.t(2,2;15;1)=0,021948` ou $P = 0{,}043896$.

♦♦

EXEMPLO 7.4

Na mesma situação do exemplo anterior, vamos considerar o caso em que o gerente esteja interessado apenas no aumento do tempo de processamento. Neste caso, as hipóteses são:

$$\begin{array}{lll} \mu & \leq & 100 \quad \text{OK!} \\ \mu & > & 100 \quad \text{Problema!} \end{array}$$

Para definir qual é a hipótese nula, vamos usar o mesmo procedimento. Em um teste unilateral, a hipótese alternativa deve ser aquela que não envolve o sinal de igualdade. No nosso exemplo, essa é a hipótese $\mu > 100$. A hipótese nula, tendo que ser uma hipótese simples, passa a ser $\mu = 100$, ou seja:

$$\begin{array}{lll} H_0 & : & \mu = 100 \\ H_1 & : & \mu > 100 \end{array}$$

Como antes, a estatística de teste é

$$T_0 = \frac{\overline{X} - 100}{\frac{10}{\sqrt{16}}} \sim t(15)$$

mas a região crítica passa a ser

$$T_0 > t_{15;\,0,05}$$

Consultando a tabela da distribuição t, resulta que

$$t_{15;\,0,05} = 1,753$$

o que nos leva à região crítica

$$T_0 > 1,753$$

Novamente rejeitamos a hipótese nula, ou seja, as evidências amostrais indicam um aumento do tempo de processamento da tarefa após a pane.

O valor P é, agora

$$P = \mathrm{P}(t_{15} > 2,2)$$

e, portanto

$$0,01 < P < 0,025$$

O valor exato é $P = 0,0219$.

♦♦

EXEMPLO 7.5

O dono de uma média empresa decide investigar a alegação de seus empregados de que o salário médio na sua empresa é menor que o salário médio nacional. Para isso, ele analisa uma amostra de 25 salários, obtendo uma média de 894,53 reais e desvio padrão de 32 reais. De informações obtidas junto ao sindicato patronal, ele sabe que, em nível nacional, o salário médio é de 900 reais. Supondo que seja razoável aproximar a distribuição dos salários por uma distribuição normal, vamos construir um teste de hipótese apropriado, com um nível de significância de 10%.

Solução

O problema aqui consiste em decidir se os salários são menores ou não do que a média nacional de 900 reais, ou seja, as situações de interesse são

$$\begin{array}{lll} \mu & < & 900 \\ \mu & \geq & 900 \end{array}$$

Logo, nossas hipóteses são:

$$\begin{aligned} H_0 &: \mu = 900 \\ H_1 &: \mu < 900 \end{aligned}$$

A região crítica é definida em termos da estatística de teste

$$T_0 = \frac{\overline{X} - 900}{\frac{32}{\sqrt{25}}} \sim t(24)$$

como

$$T_0 < -t_{24;\,0,10}$$

Com nível de significância de 10%, a abscissa de interesse é aquela que deixa área de 10% acima dela em uma distribuição t com 24 graus de liberdade:

$$t_{24;0,10} = 1,318$$

Logo, a região crítica é

$$T_0 < -1,318$$

O valor observado da estatística de teste é

$$t_0 = \frac{894,53 - 900}{\frac{32}{\sqrt{25}}} = -0,8547$$

que não está na região crítica. Logo, não rejeitamos H_0, ou seja, as evidências amostrais apontam que os salários da empresa não são menores que a média nacional.

O valor P é

$$P = \mathrm{P}(t_{24} < -0,8547) = \mathrm{P}(t_{24} > 0,8547)$$

e, pela Tabela IV, podemos dizer apenas que $P > 0,15$. O valor exato é $P = 0,20058$.

♦♦

7.2.1 Poder do teste t para a média μ

Consideremos o teste t para

$$\begin{aligned} H_0 &: \mu = \mu_0 \\ H_1 &: \mu > \mu_0 \end{aligned}$$

Como antes, podemos escrever a hipótese alternativa em termos da diferença $\Delta = \mu - \mu_0$ como

$$H_1 : \Delta > 0$$

Para um nível de significância α, a regra de decisão é rejeitar H_0 se $\sqrt{n}\dfrac{\overline{X} - \mu_0}{S} > t_{n-1;\alpha}$. Assim, a função poder do teste é

$$\begin{aligned} \pi(\mu) &= \mathrm{P}(\text{rejeitar } H_0 \,|\, \mu) = \mathrm{P}\left(\sqrt{n}\frac{\overline{X} - \mu_0}{S} > t_{n-1;\alpha} \,|\, \mu\right) \\ &= \mathrm{P}\left(\sqrt{n}\frac{\overline{X} - (\mu - \Delta)}{S} > t_{n-1;\alpha} \,|\, \mu\right) \end{aligned} \tag{7.6}$$

Note que, se a média é μ, $\dfrac{\overline{X}-\mu+\Delta}{\frac{\sigma}{\sqrt{n}}} \sim N\left(\dfrac{\sqrt{n}\Delta}{\sigma};1\right)$ e, portanto,

$$\sqrt{n}\frac{\overline{X}-(\mu-\Delta)}{S} = \frac{\dfrac{\overline{X}-\mu+\Delta}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\dfrac{\frac{(n-1)S^2}{\sigma^2}}{n-1}}}$$

tem distribuição t não central com $n-1$ graus de liberdade e parâmetro de não centralidade $\dfrac{\Delta}{\frac{\sigma}{\sqrt{n}}}$, conforme definição a seguir:

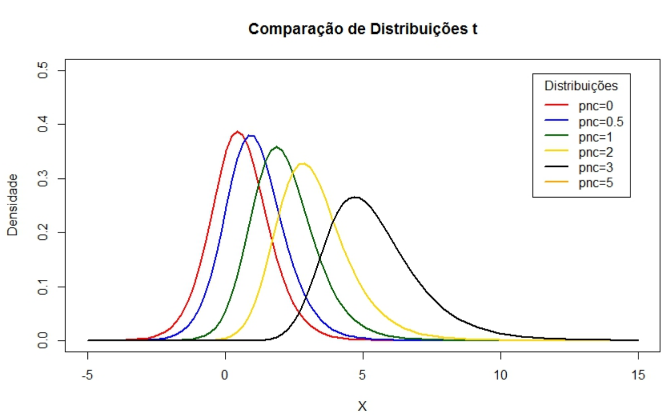
DEFINIÇÃO Densidade t não central

Sejam $Z \sim N(0;1)$ e $W \sim \chi^2_n$ variáveis aleatórias independentes e seja δ uma constante real qualquer. Então

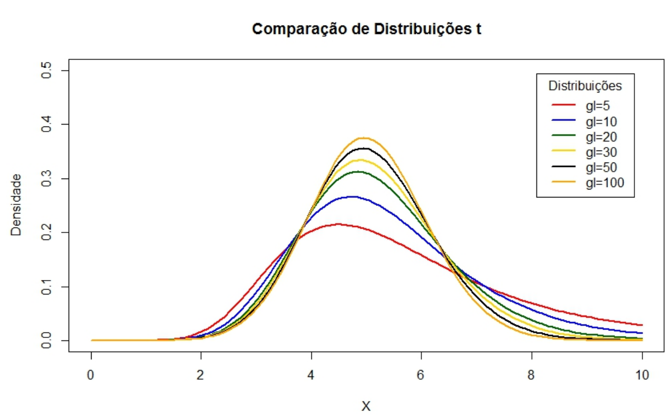
$$T = \frac{Z+\delta}{\sqrt{W/n}}$$

tem distribuição t não central com n graus de liberdade e parâmetro de não centralidade δ. Notação: $T \sim t_{n,\delta}$.

Note que, se $\delta = 0$, então $T \sim t_n$. Na Figura 7.2a ilustra-se o efeito do parâmetro de não centralidade sobre a forma da função densidade: quanto maior o valor do parâmetro de não centralidade, mais assimétrica fica a função de densidade, afastando-se da t_n. Na Figura 7.2b, vemos o efeito do número de graus de liberdade: à medida que o número de graus de liberdade aumenta, a densidade se aproxima da $N(\delta;1)$.



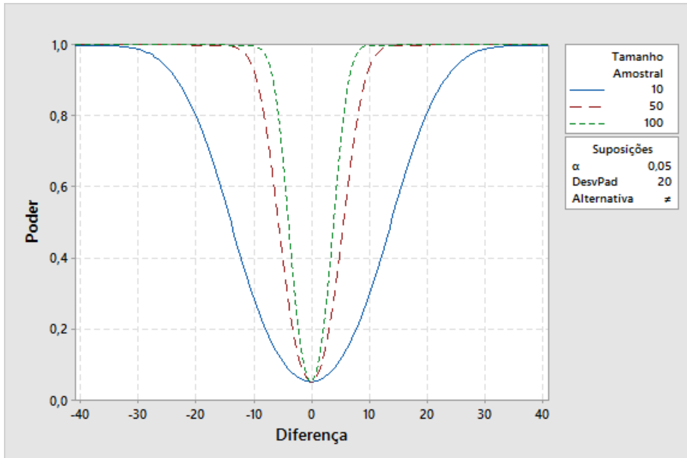
(a) Efeito do parâmetro de não centralidade



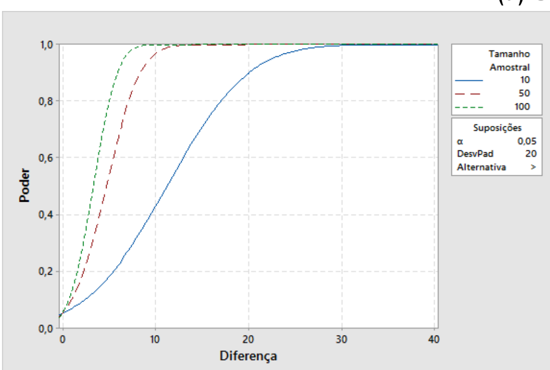
(b) Efeito dos graus de liberdade

Figura 7.2 – Distribuição t não central

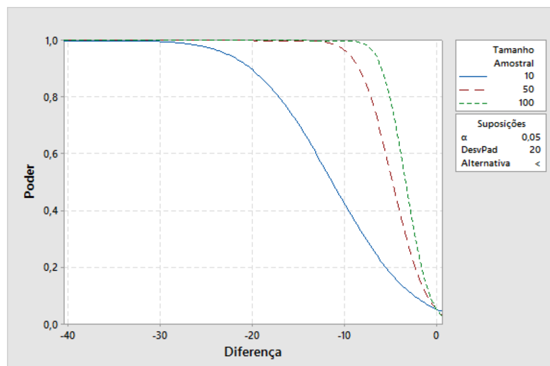
Voltando à expressão (7.6), vemos que o poder do teste t depende da distribuição acumulada de uma variável t não central e, mais do que antes, é necessário o uso de softwares para o seu cálculo. Na Figura 7.3 temos os gráficos da função poder para o teste t com desvio padrão amostral $s = 20$ e tamanhos amostrais $n = 10, 50, 100$.



(a) Bilateral



(b) Unilateral à direita



(c) Unilateral à esquerda

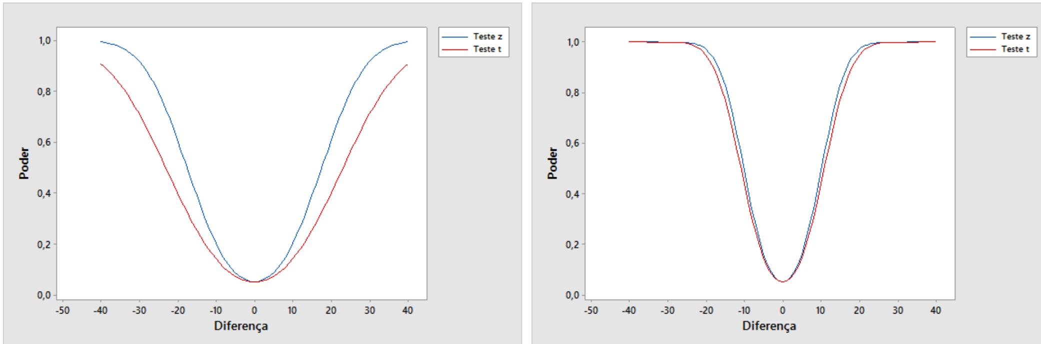
Figura 7.3 – Poder do teste t em função de $\Delta = \mu - \mu_0$

Nas Figuras 7.4a e 7.4b comparam-se as funções poder dos testes z e t, de tamanho $\alpha = 0,05$, para $n = 5$ e $n = 15$, com desvio padrão 20.

7.3 Exercícios propostos

1. Uma empresa fabricante de balas afirma que o peso médio de suas balas é de pelo menos 2 gramas. Pela descrição do processo de produção, sabe-se que o peso das balas distribui-se normalmente. Uma amostra de 25 balas apresenta peso médio de 1,98 gramas e um desvio padrão de 0,5 grama. O que se pode concluir sobre a afirmação do fabricante? Use um nível de significância de 5%.

2. Em uma linha de produção, peças são produzidas de modo que o comprimento seja normalmente distribuído. Ajustes periódicos são feitos na máquina para garantir que as peças tenham comprimento apropriado de 15 cm, pois as peças muito curtas não podem ser aproveitadas (as peças longas podem ser cortadas). A cada hora são extraídas 9 peças da produção, medindo-se seu comprimento. Uma dessas amostras apresenta comprimento médio de 14,5 cm e desvio padrão de 0,5 cm. Use o nível de significância de 0,1% para testar a hipótese de que o processo está operando adequadamente.



(a) $n = 5$ (b) $n = 15$

Figura 7.4 – Poder dos testes z e t bilaterais em função de $\Delta = \mu - \mu_0$

3. Depois de desenvolver um algoritmo para acelerar a execução de determinada tarefa rotineira em um escritório de contabilidade, o analista de sistema analisa uma amostra de 25 tempos, obtendo uma média 46,5 segundos e desvio padrão de 5 segundos. Dos dados passados, ele sabe que o tempo de execução é aproximadamente normal com média de 48,5 segundos. Use o nível de significância de 5% para decidir se o algoritmo do analista realmente melhorou o desempenho do sistema.

4. Uma propaganda afirma que o consumo médio de gasolina de determinada marca de automóvel é de 12 litros por 100 quilômetros rodados. Um teste com 36 automóveis desta marca acusa um consumo médio de 12,4 litros por 100 quilômetros rodados com desvio padrão de 1 litro por quilômetro rodado. O que se pode concluir sobre a propaganda? Use o nível de significância de 10%.

5. Uma amostra aleatória simples de tamanho $n = 9$ extraída de uma população normal apresentou média igual a $\overline{x} = 13,35$ e desvio padrão $s = 3,1$. Deseja-se testar

$$\begin{aligned} H_0 &: \mu = 12,8 \\ H_1 &: \mu \neq 12,8 \end{aligned}$$

   (a) Determine a região crítica correspondente ao nível de significância $\alpha = 0,02$.
   (b) Com base na região crítica encontrada no item anterior, estabeleça a conclusão, tendo o cuidado de usar um fraseado que não seja puramente técnico.

6. Os dados a seguir são oriundos de uma população normal com variância σ^2 supostamente igual a 36,8.

233,1	226,1	220,3	247,6	232,9	232,8	235,9	232,4
249,4	207,4	231,8	232,1	220,7	229,6	242,5	229,3

   (a) Realize um teste de hipótese apropriado para verificar a veracidade da origem dos dados. Use $\alpha = 0,05$.
   (b) Ache limites para o valor P associado a esse teste de hipótese.

7. Dados históricos indicam que a variância na taxa de câmbio do iene japonês contra o dólar americano é aproximadamente 1,56. Obteve-se uma amostra aleatória de 30 taxas de câmbio de fechamento, que acusou uma variância $s^2 == 2,2$.

   (a) Realize um teste de hipótese para verificar se houve mudança na variância na taxa de câmbio.
   (b) Ache limites para o valor P associado a esse teste de hipótese.

8. O diretor geral de um grande escritório de contabilidade está preocupado com a demora na execução de determinada tarefa e também com a variabilidade dos tempos de execução, uma vez que essa tarefa é executada por diferentes funcionários. Dados históricos revelam que o tempo médio tem sido de 40 minutos, com desvio padrão de 6 minutos. Depois de um intenso treinamento, uma amostra de 14 tempos acusa média de 35,6 minutos e desvio padrão de 3,4 minutos.

(a) As evidências amostrais indicam que o treinamento foi bem sucedido? Responda a essa pergunta construindo testes de hipóteses apropriados com nível de significância $\alpha = 2,5\%$. Certifique-se de indicar todas as etapas do processo: hipóteses nula e alternativa, estatística de teste e região crítica, limites para os valores P, conclusão em linguagem não técnica e também as suposições teóricas para resolver o problema.

(b) Construa intervalos de confiança de 95% para os parâmetros populacionais de interesse para refletir a situação depois do treinamento.

Capítulo 8

Testes para Normalidade

Os métodos de inferência vistos até agora baseiam-se fortemente na hipótese de normalidade da população, ou seja, todas as estatísticas utilizadas na construção de intervalos de confiança e testes de hipóteses partiam da suposição de que $X \sim N(\mu; \sigma^2)$. Sendo assim, é necessário termos ferramentas para verificar se tal suposição é razoável para determinada amostra. Veremos agora alguns dos testes mais comumente utilizados para tal.

8.1 Função de distribuição empírica e quantis

Alguns métodos gráficos e não paramétricos baseiam-se na função de distribuição empírica, que fornece a proporção de observações da amostra que são menores ou iguais a determinado valor x. A definição formal é dada a seguir.

DEFINIÇÃO Função de distribuição empírica

Seja $X_1, X_2, \cdots, X_n$ uma amostra aleatória simples de uma população X com função de distribuição F_X. Dada uma amostra observada $x_1, x_2, \cdots, x_n$, a função de distribuição empírica é definida como

$$\widehat{F}_n(x) = \frac{1}{n} \#\{x_i \leq x\} \qquad -\infty < x < +\infty$$

Pode-se provar que $\widehat{F}_n(x) \xrightarrow{n \to \infty} F(x)$ e esse resultado leva a alguns testes sobre a verdadeira distribuição populacional F, baseados em "distâncias" entre $\widehat{F}_n$ e F.

EXEMPLO 8.1

Considere a seguinte amostra de tamanho $n = 5$: $2, 3, 8, 13, 6$. Vamos calcular sua função de distribuição empírica. Para isso, é conveniente ordenar a amostra: $2, 3, 6, 8, 13$. Vemos, então, que

$$\widehat{F}_n(x) = \begin{cases} 0 & \text{, se } x < 2 \\ 0,2 & \text{, se } 2 \leq x < 3 \\ 0,4 & \text{, se } 3 \leq x < 6 \\ 0,6 & \text{, se } 6 \leq x < 8 \\ 0,8 & \text{, se } 8 \leq x < 13 \\ 1 & \text{, se } x \geq 13 \end{cases}$$

e seu gráfico é apresentado na Figura 8.1.

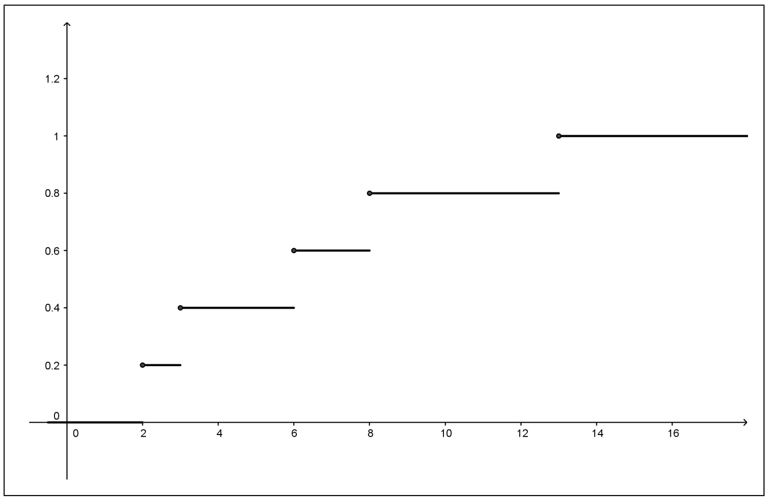


Figura 8.1 – Função de distribuição empírica para o Exemplo 8.1

♦♦

Os quantis também desempenham papel importante nos testes de normalidade. Note na função de distribuição empírica do exemplo acima que 60% das observações são menores ou iguais a 6; logo, 6 é o quantil de ordem 0,6 e $\widehat{F}_5(6) = 0,6$. No entanto, não existe valor x tal que que $\widehat{F}_5(x) = 0,5$, ou seja, não conseguimos definir a mediana, por exemplo, a partir da função de distribuição empírica. Esse problema é contornado trabalhando-se com uma "suavização" de $\widehat{F}_n$. Veja a Figura 8.2.

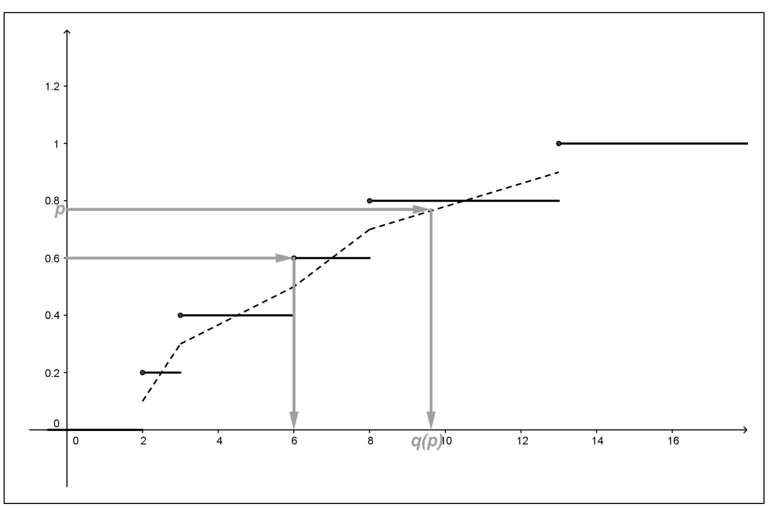


Figura 8.2 – Função de distribuição empírica suavizada para o Exemplo 8.1

A curva contínua é formada por segmentos de reta que ligam os pontos $(x_{(i)}, p_i)$ com $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ e

$$p_i = \frac{i - 0,5}{n} \tag{8.1}$$

e os quantis podem ser obtidos tomando-se a função inversa. No caso de p não coincidir com qualquer p_i, o quantil é obtido observando-se que o ponto $(q(p), p)$ está sobre o segmento de reta que passa pelos pontos $(q(p_i), p_i)$ e $(q(p_{i+1}), p_{i+1})$. Veja as linhas em forma de setas na Figura 8.2. Dessa forma, temos a seguinte definição.

DEFINIÇÃO **Quantil de ordem p**

O quantil de ordem p, para qualquer $0 < p < 1$, é definido como

$$q(p) = \begin{cases} x_{(i)} & \text{, se } p = p_i = \dfrac{i - 0,5}{n}, \ i = 1, 2, \cdots, n \\ (1 - f_i) \cdot q(p_i) + f_i \cdot q(p_{i+1}) & \text{, se } p_i < p < p_{i+1} \\ x_{(1)} & \text{, se } p < p_1 \\ x_{(n)} & \text{, se } p > P_n \end{cases}$$

sendo $f_i = \dfrac{p - p_i}{p_{i+1} - p_i}$

Outras definições possíveis envolvem diferentes valores para p_i, tais como

$$p_i = \frac{i}{n+1} \qquad\qquad p_i = \frac{i - 0,375}{n + 0,25} \tag{8.2}$$

sendo essa última adotada nas rotinas do R.

8.2 Gráfico dos quantis normais

O gráfico dos quantis normais é utilizado para se verificar visualmente se determinado conjunto de dados (amostra) é oriundo de uma distribuição normal. A ideia subjacente é: se meus dados vêm de uma população $X \sim N(\mu; \sigma^2)$, então $\text{P}(X \leq x) = \text{P}\left(Z \leq \dfrac{x - \mu}{\sigma}\right)$. Essa comparação será feita nos valores observados de X através da função de distribuição empírica.

Considerando a amostra ordenada $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ e as respectivas probabilidades associadas definidas por 8.1 ou 8.2, temos que $\widehat{F}_n(x_{(i)}) = p_i$ e, portanto,

$$F_X(x_{(i)}) = p_i \Rightarrow \Phi\left(\frac{x_{(i)} - \mu}{\sigma}\right) = p_i \Rightarrow \frac{x_{(i)} - \mu}{\sigma} = z_{1-p_i} \tag{8.3}$$

sendo z_{1-p_i} o valor crítico da normal padrão tal que $\text{P}(Z > z_{1-p_i}) = 1 - p_i$ e, portanto, $\text{P}(Z \leq z_{1-p_i}) = p_i$.

O gráfico é construído plotando-se o quantil de ordem p_i da distribuição normal padrão versus $x_{(i)}$, ou seja, os pontos no gráfico são $(z_{1-p_i}, x_{(i)})$. Se os dados são oriundos de uma normal com média μ e desvio padrão σ, então, de acordo com (8.3), o padrão dos pontos deve ser aproximadamente linear. Dessa forma, temos que buscar padrões de linearidade nos gráficos de quantis normais. Grandes afastamentos de uma relação linear levam à rejeição da hipótese de normalidade dos dados. Nas Figuras 8.3 e 8.4 apresentamos alguns gráficos gerados a partir de dados simulados de uma $N(4; 2^2)$ com tamanhos de amostra $n = 10$ e $n = 50$. Esses gráficos foram gerados pelo Minitab. A reta central é o padrão ideal, ou seja, o comportamento esperado para uma distribuição exatamente normal (as curvas externas representam o intervalo de confiança para a função de distribuição). Nesses gráficos podemos ver o efeito do tamanho da amostra; como a função de distribuição empírica converge para a verdadeira função de distribuição, o padrão linear será mais nítido para amostras maiores. Em geral, para amostras pequenas, é difícil analisar o padrão de linearidade em um gráfico de quantis normais.

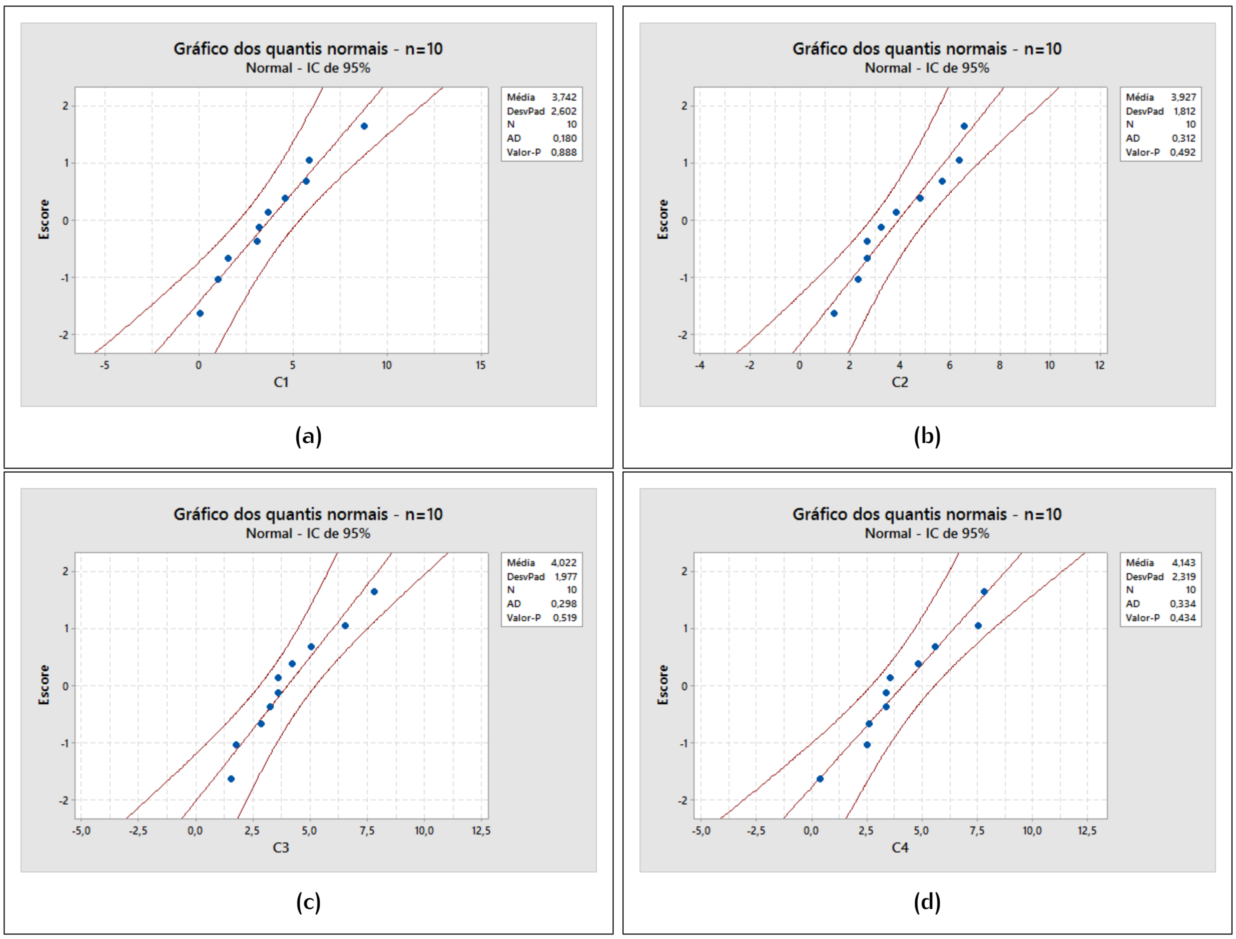


Figura 8.3 – Amostras da Normal padrão – $n = 10$

8.3 Testes de normalidade baseados na distribuição empírica

A análise de um padrão linear em um gráfico de quantis normais é, além de difícil, subjetiva. Assim, é necessário algum método mais objetivo para verificar a hipótese de normalidade dos dados. Veremos, agora, alguns desses métodos, que são testes das hipóteses gerais:

$$
\begin{aligned}
H_0 &: \text{Os dados vêm de uma distribuição normal} \\
H_1 &: \text{Os dados não vêm de uma distribuição normal}
\end{aligned}
$$

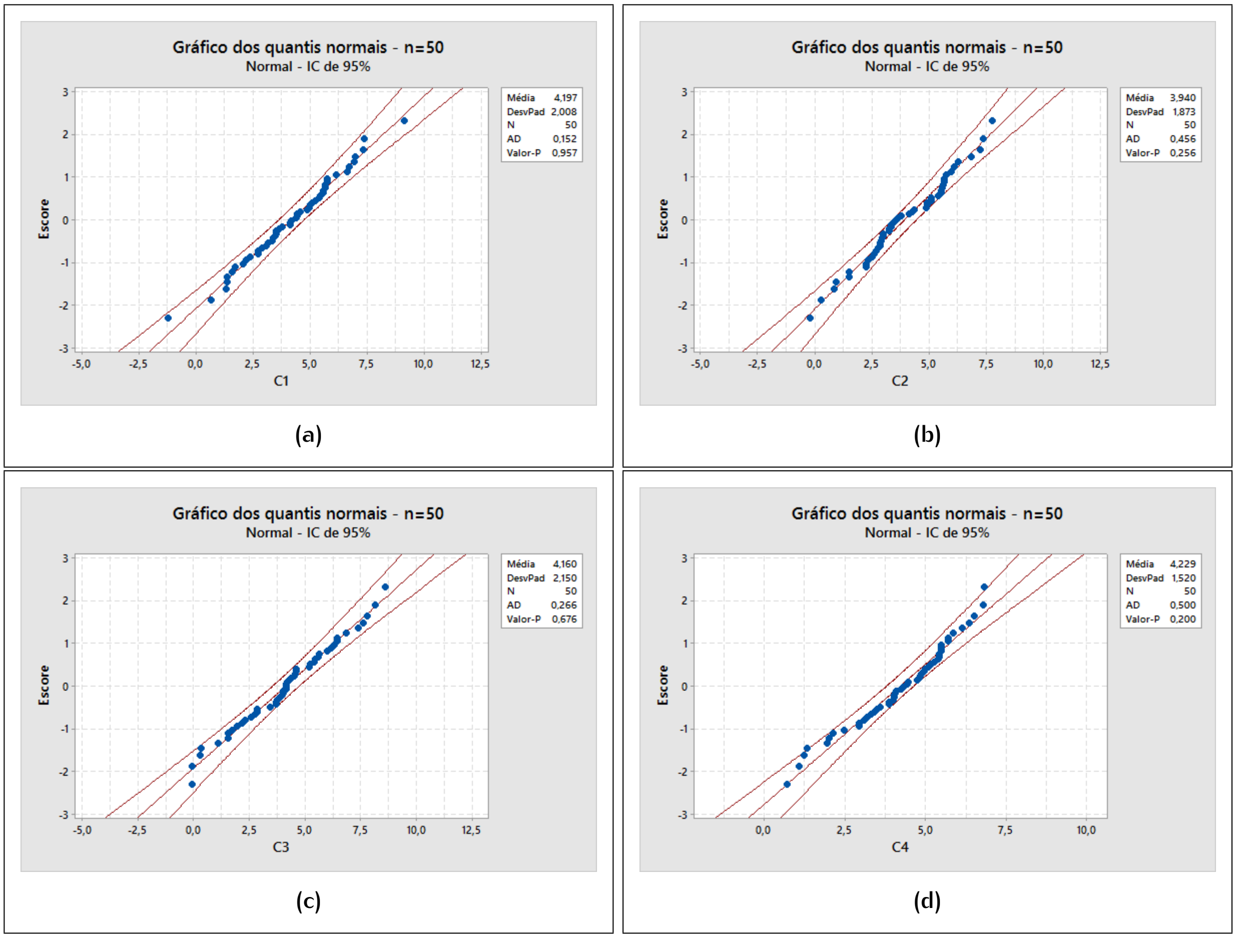


Figura 8.4 – Amostras da Normal padrão – $n = 50$

Todos eles se baseiam na função de distribuição empírica $\widehat{F}_n(x)$ que, como dito, converge para $F_X(x) = \Phi(x)$ quando $n \to \infty$. Mas como não conhecemos a distribuição específica da qual a amostra foi retirada, mesmo sob a hipótese de veracidade de H_0 (μ e σ são desconhecidos), temos que usar algum estimador para $F(x)$ e o estimador natural é $\widehat{F}(x) = \Phi\left(\dfrac{x - \overline{X}}{S}\right)$. Diferentes testes usam diferentes métricas de discrepância para comparar $\widehat{F}(x)$ e $\widehat{F}_n(x)$. Se as diferenças forem "grandes", rejeita-se a hipótese nula. O cálculo de valores críticos e valores P para os diferentes testes requer uso de software.

8.3.1 Teste de Lilliefors (Kolmogorov-Smirnov)

A estatística de teste para os testes de Kolmogorov-Smirnov e Lilliefors mede a discrepância local máxima:

$$D = \max_x \left\{ \left| \widehat{F}_n(x) - \Phi\left(\frac{x - \overline{X}}{S}\right) \right| \right\} \tag{8.4}$$

e sua fórmula de cálculo é

$$D = \max\left[\max\left\{\frac{i}{n} - \Phi\left(\frac{x(i)-\overline{x}}{s}\right)\right\}, \max\left\{\Phi\left(\frac{x(i)-\overline{x}}{s}\right) - \frac{i-1}{n}\right\}\right] \tag{8.5}$$

A diferença entre os dois testes se dá quando a hipótese nula não especifica os valores de μ e de σ da distribuição normal. Usar a média e o desvio padrão amostrais faz com que a distribuição de $\widehat{F}$ se torne mais próxima da distribuição empírica, ou seja, as discrepâncias tendem a ser menores do que seriam se H_0 especificasse uma distribuição normal bem definida. Isso requer, então, um ajuste nos valores críticos e nos valores P.

EXEMPLO 8.2

Considere os seguintes dados (já ordenados), que foram simulados a partir de uma distribuição normal com média 5 e desvio padrão 2:

$$2,374018 \quad 3,586709 \quad 3,896683 \quad 4,333655 \quad 6,029997$$

. A média amostral é $\overline{x} = 4,044212$ e o desvio padrão amostral é $s = 1,327379$. Na tabela a seguir temos as informações necessárias para o cálculo da estatística de teste de Kolmogorov-Smirnov e na Figura 8.5 ilustram-se as discrepâncias entre $\widehat{F}$ e $\widehat{F}_5$ como os segmentos verticais pontilhados:

$x_{(i)}$	$p_i = \widehat{F}_5(x_{(i)})$	$z_i = \frac{x_{(i)}-\overline{x}}{s}$	$\Phi(z_i) = \widehat{F}(x_{(i)})$	$\lvert\widehat{F}_5(x_{(i)}) - \widehat{F}(x_{(i)})$
2,374018	0,2	-1,258265	0,104148	0,095852
3,586709	0,4	-0,344667	0,365173	0,034827
3,896683	0,6	-0,111143	0,455751	0,144249
4,333655	0,8	0,218056	0,586307	**0,213693**
6,029997	1,0	1,496019	0,932676	0,067324

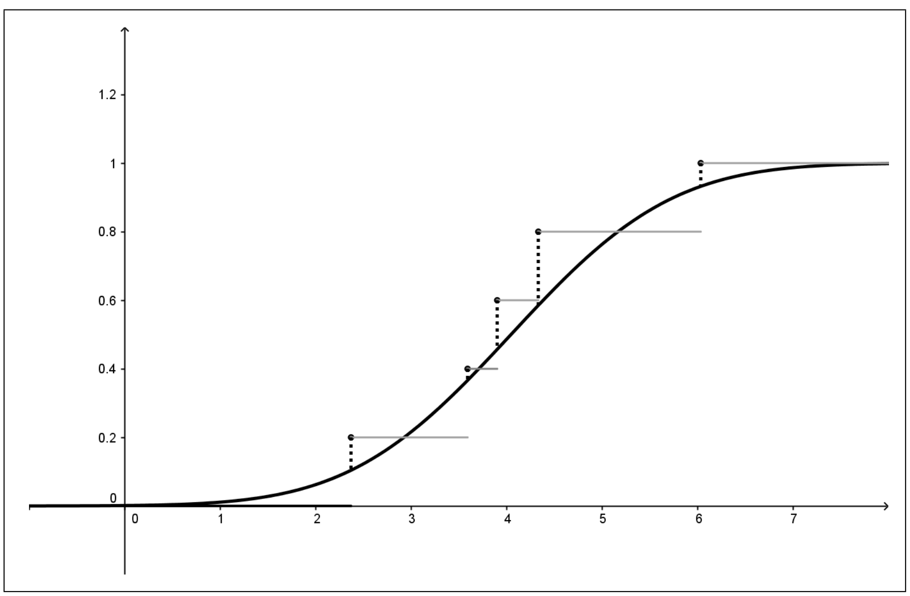


Figura 8.5 – Cálculo da estatística de teste de Lilliefors

Da tabela dada no artigo original de Lilliefors, conclui-se que o valor $P > 0,20$, ou seja, *não* rejeitamos a hipótese nula de normalidade dos dados.

8.3.2 Teste de Anderson-Darling

A estatística de teste de Anderson-Darling é definida como

$$A^2 = \int_{-\infty}^{\infty} \frac{\left[\widehat{F}_n(x) - \Phi\left(\frac{x-\overline{x}}{s}\right)\right]^2}{\Phi\left(\frac{x-\overline{x}}{s}\right)\left[1 - \Phi\left(\frac{x-\overline{x}}{s}\right)\right]} \frac{1}{s}\phi\left(\frac{x-\overline{x}}{s}\right) dx \tag{8.6}$$

onde $\phi(x) = \Phi'(x)$. Como antes, estamos analisando a probabilidade da diferença, agora ao quadrado, entre $\widehat{F}_n$ e $\widehat{F}$, mas agora relativa às probabilidades nas caudas. Dessa forma, o teste de Anderson-Darling tende a ser mais eficaz para detectar desvios nas laterais da distribuição.

A fórmula de cálculo para a estatística de Anderson-Darling é

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}\left[(2i-1)\log\left(\Phi\left(\frac{x_{(i)}-\overline{x}}{s}\right)\right) + (2n+1-2i)\log\left(1-\Phi\left(\frac{x_{(i)}-\overline{x}}{s}\right)\right)\right] \tag{8.7}$$

e na tabela a seguir mostramos os cálculos para os dados do Exemplo 8.2.

i	$x_{(i)}$	$z_i = \frac{x_{(i)}-\overline{x}}{s}$	$\Phi(z_i)$	$\ln[\Phi(z_i)]$	$\ln[1-\Phi(z_i)]$	$(2i-1)\ln[\Phi(z_i)]$	$(2n-2i+1)\ln[1-\Phi(z_i)]$
1	2,374018	-1,258265	0,104148	-2,261942	-0,10998	-2,261942	-0,989820
2	3,586709	-0,344667	0,365173	-1,007385	-0,454402	-3,022156	-3,180814
3	3,896683	-0,111143	0,455751	-0,785808	-0,608349	-3,92904	-3,041745
4	4,333655	0,218056	0,586307	-0,533911	-0,882632	-3,73738	-2,647895
5	6,029997	1,496019	0,932676	-0,069698	-2,698233	-0,62728	-2,698233
					Soma	-13,577798	-12,558507

Obtemos, então, que

$$A^2 = -5 - \frac{1}{5}(-13,577798 - 12,558507) = 0,227261$$

e o Minitab fornece o valor $P = 0,639$. Novamente, *não* rejeitamos a hipótese de normalidade dos dados.

8.4 Teste de Shapiro-Wilk

A estatística de teste de Shapiro-Wilk é obtida dividindo-se o quadrado de uma combinação linear das estatísticas de ordem amostrais pela estimativa usual da variância, S^2. Sem entrar em detalhes que estão além do nível deste curso, a estatística de teste é

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \tag{8.8}$$

onde as constantes a_i dependem das estatísticas de ordem de uma amostra de tamanho n da população normal padrão. Com antes, programas computacionais são necessários para a análise. Tais programas fornecem o valor da estatística e do respectivo valor P.

8.5 Exercícios propostos

Utilize software para verificar se os dados apresentados nos exercícios abaixo relacionados são provenientes de uma população normal.

- Capítulo 3 – Exercício 10
- Capítulo 4 – Exercícios 4, 6

Parte II

Inferência para duas populações

Capítulo 9

Inferência com Amostras Independentes

Vamos considerar, inicialmente, o caso de termos amostras *independentes* de duas populações normais. Como no caso de uma população, vamos começar com a situação simplificada em que as variâncias populacionais são conhecidas e nosso interesse está na inferência sobre as médias populacionais. Em seguida, faremos inferência sobre as variâncias populacionais para, finalmente, considerarmos o caso geral em que tanto as médias como as variâncias populacionais são desconhecidas.

9.1 Introdução

É muito comum encontrarmos, na prática, situações em que o objetivo é comparar dois grupos diferentes. Será que meninos e meninas do ensino médio gastam o mesmo número de horas por semana navegando na internet? O desempenho de alunos em um exame nacional melhora depois que eles realizam um curso preparatório? Irmãos gêmeos respondem da mesma forma a um determinado estímulo? A proporção de pessoas favoráveis a determinado projeto de um governo estadual é a mesma na zona urbana e na zona rural?

Em todos esses exemplos, queremos comparar duas populações: meninos e meninas, alunos antes e depois do curso preparatório, irmãos gêmeos, zona urbana e zona rural. Como no caso de uma população, tomaremos nossas decisões com base em amostras aleatórias simples retiradas dessas populações. Mas podemos ver, nesses exemplos, duas situações diferentes: quando comparamos irmãos gêmeos ou alunos antes e depois de um curso preparatório e quando comparamos meninos e meninas do ensino médio ou zonas urbana e rual. No primeiro caso, há uma dependência entre as amostras, o que não ocorre no segundo caso.

9.2 Definições e notação

No estudo da inferência para uma população, vimos que a suposição de termos uma amostra aleatória simples da população de interesse era fundamental para o desenvolvimento dos estimadores e estatísticas de teste. Nos problemas de inferência a partir de duas amostras não é diferente, mas temos que nos preocupar também com possíveis relações entre as populações e/ou amostras.

Sejam duas populações representadas pelas variáveis aleatórias X_1 e X_2 e sejam $X_{11}, X_{12}, \ldots, X_{1n_1}$ e $X_{21}, X_{22}, \ldots, X_{2n_2}$ amostras aleatórias simples de tamanhos n_1 e n_2 retiradas dessas populações. Neste capítulo, estudaremos apenas amostras *independentes*, cuja definição é dada a seguir..

DEFINIÇÃO Amostras Independentes

As amostras são *independentes* se o processo de seleção dos indivíduos ou objetos na amostra 1 não tem qualquer efeito sobre, ou qualquer relação com, a seleção dos indivíduos ou objetos na amostra 2. Se as amostras não são independentes, elas são *dependentes*.

Na Tabela 9.1 apresentamos a notação que será utilizada referente aos parâmetros dessas populações e seus respectivos estimadores. Assim como no caso de uma população, vamos assumir que as populações sejam normais.

Tabela 9.1 – Notação: Inferência para duas populações

População	Parâmetros		Amostra	Estatística Amostral		Valores observados	
	Média	Variância		Média	Variância	Média	Variância
$X_1 \sim N(\mu_1; \sigma_1^2)$	μ_1	σ_1^2	$X_{11}, X_{12}, \ldots, X_{1n_1}$	$\overline{X}_1$	S_1^2	$\overline{x}_1$	s_1^2
$X_2 \sim N(\mu_2; \sigma_2^2)$	μ_2	σ_2^2	$X_{21}, X_{22}, \ldots, X_{2n_2}$	$\overline{X}_2$	S_2^2	$\overline{x}_2$	s_2^2

9.3 Inferência sobre médias de duas populações normais com variâncias conhecidas

Como no caso de uma população, vamos considerar, inicialmente, o caso de amostras *independentes* de duas populações normais cujas variâncias são conhecidas. Assim, nosso interesse está na inferência sobre as médias populacionais e tal inferência se baseará nas médias amostrais, sobre as quais sabemos que

$$\overline{X}_1 \sim N\left(\mu_1; \frac{\sigma_1^2}{n_1}\right) \qquad \overline{X}_2 \sim N\left(\mu_2; \frac{\sigma_2^2}{n_2}\right)$$

Como estamos supondo que as amostras são independentes, resulta que $\overline{X}_1$ e $\overline{X}_2$ são independentes. Logo,

$$\overline{X}_1 - \overline{X}_2 \sim N\left(\mu_1 - \mu_2; \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right) \tag{9.1}$$

ou equivalentemente,

$$\frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0; 1) \tag{9.2}$$

9.3.1 Intervalo de confiança para $\mu_1 - \mu_2$

Temos, a partir de (9.2), que

$$P\left(-z_{\alpha/2} < \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < z_{\alpha/2}\right) = 1 - \alpha$$

Logo,

$$P\left((\overline{X}_1 - \overline{X}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < (\mu_1 - \mu_2) < (\overline{X}_1 - \overline{X}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) = 1 - \alpha$$

o que nos dá o seguinte intervalo de confiança de $100(1-\alpha)\%$ para $\mu_1 - \mu_2$.

$$\left((\overline{X}_1 - \overline{X}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}; (\overline{X}_1 - \overline{X}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) \tag{9.3}$$

Note que os limites são variáveis aleatórias e, assim, faz sentido falar que a probabilidade é $1 - \alpha$. Isso significa que se repetirmos várias vezes o processo de amostragem e subsequente construção do intervalo de confiança correspondente, "acertaremos" em $100(1-\alpha)\%$ das vezes, ou seja, o intervalo conterá o verdadeiro parâmetro $\mu_1 - \mu_2$.

Para uma amostra específica, o intervalo de confiança é

$$\left((\overline{x}_1 - \overline{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}; (\overline{x}_1 - \overline{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) \tag{9.4}$$

e esse intervalo ou contém, ou não contém o verdadeiro parâmetro. Não faz sentido dizer que o intervalo dado em (9.4) contém o parâmetro com probabilidade $1 - \alpha$. O que podemos dizer é que o método de obtenção dos intervalos garante uma probabilidade de acerto de $1 - \alpha$.

9.3.2 Teste de hipótese sobre $\mu_1 - \mu_2$

Vamos, agora, estabelecer os procedimentos para o teste de hipóteses referentes à diferença entre as médias, de maneira análoga ao caso de uma população. Nosso objetivo, então, é testar [1]

$$H_0 : \mu_1 - \mu_2 = 0 \tag{9.5}$$

As hipóteses alternativas possíveis são:

$$H_1 : \mu_1 - \mu_2 \neq 0 \qquad H_1 : \mu_1 - \mu_2 < 0 \qquad H_1 : \mu_1 - \mu_2 > 0$$

O procedimento de teste consiste em rejeitar a veracidade da hipótese nula H_0 sempre que obtivermos valores da estatística de teste com pequenas probabilidades de ocorrência sob H_0. Para o teste de hipótese sobre a diferença de médias de populações normais com variâncias conhecidas, a estatística de teste é dada em (9.2) e probabilidades pequenas correspondem à cauda da distribuição normal. Lembrando que sob H_0, $\mu_1 - \mu_2 = 0$, isso nos leva às seguintes regras de decisão para um nível de significância α (lembre-se que $\alpha = P(\text{rejeitar } H_0 | H_0 \text{ é verdadeira})$:

[1] É possível trabalhar com hipóteses mais gerais envolvendo uma diferença Δ_0, mas os procedimentos são análogos.

- Hipótese nula e estatística de teste

$$H_0 : \mu_1 - \mu_2 = 0$$

$$Z_0 = \frac{(\overline{X}_1 - \overline{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \underbrace{\sim}_{\text{sob } H_0} N(0,1)$$

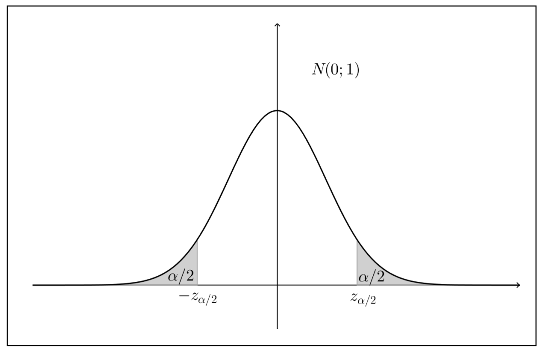
$$z_0 = \text{valor observado de } Z_0$$

- Teste bilateral

$H_1 : \mu_1 - \mu_2 \neq 0$

Região crítica:
$|Z_0| > z_{\alpha/2}$

Valor P :
$P = 2 \cdot \mathrm{P}(Z > |z_0|)$

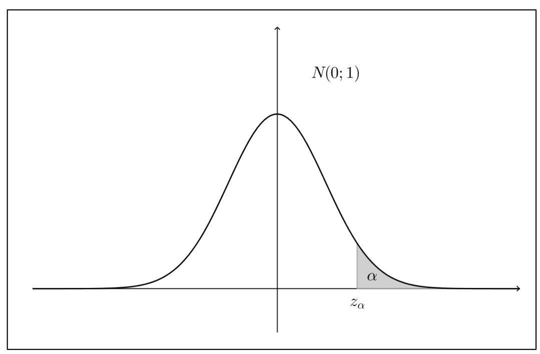


- Teste unilateral à direita

$H_1 : \mu_1 - \mu_2 > 0$

Região crítica:
$Z_0 > z_\alpha$

Valor P :
$P = \mathrm{P}(Z > z_0)$

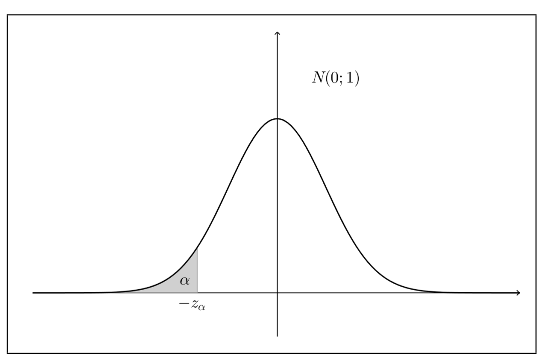


- Teste unilateral à esquerda

$H_1 : \mu_1 - \mu_2 < 0$

Região crítica:
$Z_0 < -z_\alpha$

Valor P :
$P = \mathrm{P}(Z > |z_0|)$



9.3.3 Poder do teste e tamanho de amostra

No mesmo contexto da subseção anterior, vamos considerar o teste da hipótese de igualdade de médias, que é equivalente a

$$H_0 : \mu_1 - \mu_2 = 0 \tag{9.6}$$

e calcular o seu poder quando

$$\mu_1 - \mu_2 = \Delta$$

Teste bilateral

Considerando a regra de decisão vista acima, o poder do teste pode ser escrito em função de Δ como

$$\begin{aligned}
\pi(\Delta) &= \mathrm{P}(\text{rejeitar } H_0 \,|\, \Delta) = 1 - \mathrm{P}\left(-z_{\alpha/2} \le \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \le z_{\alpha/2} \,|\, \Delta \right) \\
&= 1 - \mathrm{P}\left(-z_{\alpha/2} \le \frac{\overline{X_1} - \overline{X_2} - \Delta + \Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \le z_{\alpha/2} \,|\, \Delta \right) \\
&= 1 - \mathrm{P}\left(-z_{\alpha/2} - \frac{\Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \le Z \le z_{\alpha/2} - \frac{\Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \right) \Longrightarrow
\end{aligned}$$

$$\pi(\Delta) = 1 - \left[\Phi\left(z_{\alpha/2} - \frac{\Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \right) - \Phi\left(-z_{\alpha/2} - \frac{\Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \right) \right] \tag{9.7}$$

Assim como no caso de uma população, é possível determinar o tamanho da amostra necessário para se ter um poder π^* na detecção de uma diferença Δ^*, desde que se trabalhe com amostras de *mesmo* tamanho n. Nesse caso, temos

$$\pi(\Delta) = 1 - \left[\Phi\left(z_{\alpha/2} - \frac{\Delta}{\sqrt{\dfrac{\sigma_1^2 + \sigma_2^2}{n}}} \right) - \Phi\left(-z_{\alpha/2} - \frac{\Delta}{\sqrt{\dfrac{\sigma_1^2 + \sigma_2^2}{n}}} \right) \right]$$

Se queremos $\pi(\Delta^*) = \pi^*$ com $\Delta^* > 0$, então $\Phi\left(-z_{\alpha/2} - \dfrac{\Delta^*}{\sqrt{\dfrac{\sigma_1^2\sigma_2^2}{n}}}\right) \approx 0$, o que resulta em

$$\pi(\Delta^*) = \pi^* \approx 1 - \Phi\left(z_{\alpha/2} - \frac{\Delta^*}{\sqrt{\dfrac{\sigma_1^2+\sigma_2^2}{n}}}\right) \Rightarrow \pi^* \approx \mathrm{P}\left(Z > z_{\alpha/2} - \frac{\Delta^*}{\sqrt{\dfrac{\sigma_1^2+\sigma_2^2}{n}}}\right) \Longrightarrow$$

$$z_{\pi^*} \approx z_{\alpha/2} - \frac{\Delta^*}{\sqrt{\dfrac{\sigma_1^2+\sigma_2^2}{n}}} \Longrightarrow \sqrt{n} \approx \frac{z_{\alpha/2} - z_{\pi^*}}{\Delta^*}\sqrt{\sigma_1^2+\sigma_2^2} \Longrightarrow$$

$$n \approx \frac{\left(z_{\alpha/2} - z_{\pi^*}\right)^2\left(\sigma_1^2+\sigma_2^2\right)}{(\Delta^*)^2} \tag{9.8}$$

Se $\Delta^* < 0$, então $\Phi\left(z_{\alpha/2} - \dfrac{\Delta^*}{\sqrt{\dfrac{\sigma_1^2\sigma_2^2}{n}}}\right) \approx 1$, o que resulta em

$$\pi(\Delta^*) = \pi^* \approx \Phi\left(-z_{\alpha/2} - \frac{\Delta^*}{\sqrt{\dfrac{\sigma_1^2+\sigma_2^2}{n}}}\right) = \mathrm{P}\left(Z > z_{\alpha/2} + \frac{\Delta^*}{\sqrt{\dfrac{\sigma_1^2+\sigma_2^2}{n}}}\right) \Longrightarrow$$

$$z_{\pi^*} \approx z_{\alpha/2} + \frac{\Delta^*}{\sqrt{\dfrac{\sigma_1^2+\sigma_2^2}{n}}} \Longrightarrow \sqrt{n} \approx \frac{z_{\pi^*} - z_{\alpha/2}}{\Delta^*}\sqrt{\sigma_1^2+\sigma_2^2}$$

e isso nos leva à mesma expressão para n dada em (9.8).

Teste unilateral à direita

Considerando a regra de decisão vista acima, o poder do teste pode ser escrito em função de $\Delta > 0$ como

$$\begin{aligned}
\pi(\Delta) &= \mathrm{P}(\text{rejeitar } H_0 \,|\, \Delta) = \mathrm{P}\left(\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} > z_\alpha \,|\, \Delta\right) \\
&= \mathrm{P}\left(\frac{\overline{X}_1 - \overline{X}_2 - \Delta + \Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} > z_\alpha \,|\, \Delta\right) = \mathrm{P}\left(Z > z_\alpha - \frac{\Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}\right) \Longrightarrow
\end{aligned}$$

$$\pi(\Delta) = 1 - \Phi\left(z_\alpha - \frac{\Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}\right) \tag{9.9}$$

Se queremos determinar o tamanho comum das amostras necessário para se ter um poder π^* na detecção de uma diferença $\Delta^* > 0$, então temos que ter

$$\pi(\Delta^*) = \pi^* = \mathrm{P}\left(Z > z_\alpha - \frac{\Delta^*}{\sqrt{\dfrac{\sigma_1^2 + \sigma_2^2}{n}}}\right) \Longrightarrow z_{\pi^*} = z_\alpha - \frac{\Delta^*}{\sqrt{\dfrac{\sigma_1^2 + \sigma_2^2}{n}}}$$

o que resulta em

$$n \approx \frac{(z_\alpha - z_{\pi^*})^2 \left(\sigma_1^2 + \sigma_2^2\right)}{(\Delta^*)^2} \tag{9.10}$$

Teste unilateral à esquerda

O poder do teste pode ser escrito em função de $\Delta < 0$ como

$$\begin{aligned}
\pi(\Delta) &= \mathrm{P}(\text{rejeitar } H_0 \,|\, \Delta) = \mathrm{P}\left(\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} < -z_\alpha \,|\, \Delta\right) \\
&= \mathrm{P}\left(\frac{\overline{X}_1 - \overline{X}_2 - \Delta + \Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} < -z_\alpha \,|\, \Delta\right) = \mathrm{P}\left(Z < -z_\alpha - \frac{\Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}\right) \\
&= \mathrm{P}\left(Z > z_\alpha + \frac{\Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}\right) \Longrightarrow
\end{aligned}$$

$$\pi(\Delta) = 1 - \Phi\left(z_{\alpha/2} + \frac{\Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}\right) \tag{9.11}$$

Se queremos determinar o tamanho comum das amostras necessário para se ter um poder π^* na detecção de uma diferença $\Delta^* > 0$, então temos que ter

$$\pi(\Delta^*) = \pi^* = \mathrm{P}\left(Z > z_\alpha + \frac{\Delta^*}{\sqrt{\dfrac{\sigma_1^2 + \sigma_2^2}{n}}}\right) \Longrightarrow z_{\pi^*} = z_\alpha + \frac{\Delta^*}{\sqrt{\dfrac{\sigma_1^2 + \sigma_2^2}{n}}}$$

o que resulta em

$$n \approx \frac{(z_{\pi^*} - z_\alpha)^2 \left(\sigma_1^2 + \sigma_2^2\right)}{(\Delta^*)^2} \tag{9.12}$$

mesma expressão obtida para o teste unilateral à direita.

9.4 Inferência sobre duas proporções - amostras grandes

Consideremos, agora,o caso em que nosso interesse está na comparação de duas proporções. Isso significa que nossas duas populações são descritas por variáveis Bernoulli, ou seja, $X_1 \sim Bern(p_1)$ e $X_2 \sim Bern(p_2)$ respectivamente. Se $X_{11}, X_{12}, \cdots, X_{1n_1}$ e $X_{21}, X_{22}, \cdots, X_{2n_2}$ são amostras grandes dessas populações , o Teorema Limite Central nos dá que

$$\overline{X}_1 = \widehat{P}_1 \approx N\left(p_1; \frac{p_1(1-p_1)}{n_1}\right) \qquad \widehat{P}_2 \approx N\left(p_2; \frac{p_2(1-p_2)}{n_2}\right)$$

Se as amostras forem independentes, resulta que

$$\widehat{P}_1 - \widehat{P}_2 \approx N\left(p_1 - p_2; \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right) \tag{9.13}$$

sendo necessárias as condições $n_1 > 30$, $n_1 p_1 \geq 5$, $n_1(1-p_1) \geq 5$, $n_2 > 30$, $n_2 p_2 \geq 5$ e $n_2(1-p_2) \geq 5$.

9.4.1 Intervalo de confiança para $p_1 - p_2$

Satisfeitas as condições de grandes amostras e simetria, resulta de 9.13 que

$$\mathrm{P}\left(-z_{\alpha/2} < \frac{(\widehat{P}_1 - \widehat{P}_2) - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}} < z_{\alpha/2}\right) \cong 1 - \alpha$$

Logo,

$$\mathrm{P}\left((\widehat{P}_1 - \widehat{P}_2) - z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} < (p_1 - p_2) < (\widehat{P}_1 - \widehat{P}_2) + z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}\right) \cong 1-\alpha$$

Como no caso univariado, estimamos a variância a partir da amostra, o que nos dá o seguinte intervalo de confiança de nível de confiança aproximadamente igual a $1 - \alpha$:

$$\left(\widehat{P}_1 - \widehat{P}_2 - z_{\alpha/2}\sqrt{\frac{\widehat{P}_1(1-\widehat{P}_1)}{n_1} + \frac{\widehat{P}_2(1-\widehat{P}_2)}{n_2}}\,;\ \widehat{P}_1 - \widehat{P}_2 + z_{\alpha/2}\sqrt{\frac{\widehat{P}_1(1-\widehat{P}_1)}{n_1} + \frac{\widehat{P}_2(1-\widehat{P}_2)}{n_2}}\right)$$

9.4.2 Teste de hipótese sobre $p_1 - p_2$

Consideremos o teste da hipótese da hipótese nula

$$H_0 : p_1 = p_2$$

contra uma das alternativas

$$H_1 : p_1 \neq p_2 \qquad H_1 : p_1 > p_2 \qquad H_1 : p_1 < p_2$$

O teste de grandes amostras se baseia na estatística dada em (9.13). Mas, sob H_0, $p_1 = p_2 = p$ e a variância de $\widehat{P}_1 - \widehat{P}_2$ nesse caso é

$$\sigma^2_{\widehat{P}_1 - \widehat{P}_2} = \frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2} = p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

e, sob H_0,

$$Z_0 = \frac{(\widehat{P}_1 - \widehat{P}_2)}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \approx N(0;1) \tag{9.14}$$

Como as amostras vêm de populações com a mesma proporção p, podemos usar as amostras combinadas para estimar p, definindo o estimador combinado

$$\widehat{P}_C = \frac{\text{número total de sucessos nas 2 amostras}}{n_1 + n_2} = \frac{U_1 + U_2}{n_1 + n_2} \tag{9.15}$$

em que $U_i = \sum_{j=1}^{n_i} X_{ij}$ é o número de sucessos na amostra i.

Esse estimador pode ser reescrito como

$$\widehat{P}_C = \frac{U_1 + U_2}{n_1 + n_2} = \frac{n_1\widehat{P}_1 + n_2\widehat{P}_2}{n_1 + n_2} = \frac{n_1}{n_1 + n_2}\widehat{P}_1 + \frac{n_2}{n_1 + n_2}\widehat{P}_2 \tag{9.16}$$

que é uma média ponderada das proporções amostrais $\widehat{P}_1$ e $\widehat{P}_2$.

Usando (9.14) e (9.16), chegamos às seguintes regras de decisão para um nível de significância α:

- Hipótese nula e estatística de teste

$$H_0 : p_1 - p_2 = 0$$

$$Z_0 = \frac{(\widehat{P}_1 - \widehat{P}_2)}{\sqrt{\widehat{P}_C(1 - \widehat{P}_C)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \approx N(0;1)$$

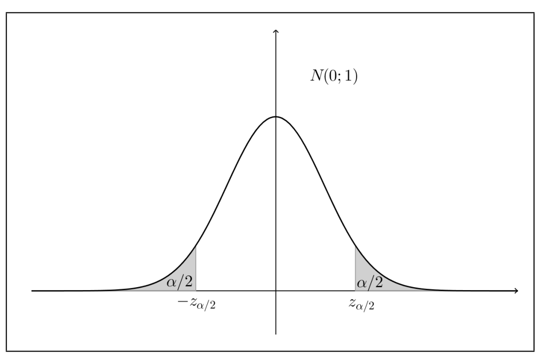
$$z_0 = \text{valor observado de } Z_0$$

- Teste bilateral

$H_1 : p_1 - p_2 \neq 0$

Região crítica:
$|Z_0| > z_{\alpha/2}$

Valor P :
$P = 2 \cdot \text{P}(Z > |z_0|)$

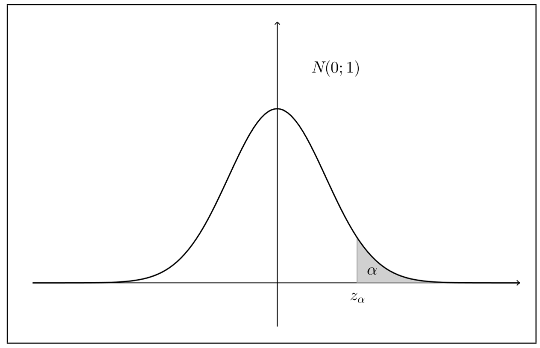


- Teste unilateral à direita

$H_1 : p_1 - p_2 > 0$

Região crítica:
$Z_0 > z_\alpha$

Valor P :
$P = \mathrm{P}(Z > z_0)$

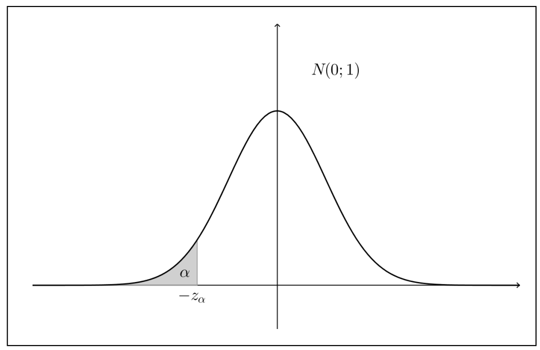


- Teste unilateral à esquerda

$H_1 : p_1 - p_2 < 0$

Região crítica:
$Z_0 < -z_\alpha$

Valor P :
$P = \mathrm{P}(Z > |z_0|)$



Para o caso geral em que $H_0 : p_1 - p_2 = \Delta_0$, com $\Delta_0 \neq 0$, o teste se baseia na estatística

$$Z_0 = \frac{(\widehat{P}_1 - \widehat{P}_2) - \Delta_0}{\sqrt{\dfrac{\widehat{P}_1\left(1 - \widehat{P}_1\right)}{n_1} + \dfrac{\widehat{P}_2\left(1 - \widehat{P}_2\right)}{n_2}}} \approx N(0;1)$$

que leva à região crítica e ao valor P análogos aos vistos anteriormente. A diferença aqui é que não podemos usar o estimador combinado.

9.5 Inferência sobre variâncias de duas populações normais

9.5.1 A Distribuição F

No estudo comparativo de variâncias de populações normais, faremos uso da distribuição F, assim denominada em homenagem ao estatístico Ronald Fisher (1890-1962), e cuja função densidade é

$$f(x) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \frac{x^{\frac{\nu_1-2}{\nu_2}}}{\left(1+\dfrac{\nu_1 x}{\nu_2}\right)^{\frac{\nu_1+\nu_2}{2}}} \qquad x > 0 \tag{9.17}$$

Essa distribuição depende dos dois parâmetros, ν_1 e ν_2 e usaremos a notação $X \sim F_{\nu_1,\nu_2}$ para indicar que a variável aleatória X tem distribuição F com parâmetros ν_1 e ν_2. Os parâmetros da distribuição F

são chamados graus de liberdade do numerador e do denominador, respectivamente. Temos os seguintes resultados sobre a média e a variância dessa distribuição:

$$X \sim F_{\nu_1,\nu_2} \Longrightarrow \begin{cases} \mathrm{E}(X) = \dfrac{\nu_2}{\nu_2 - 2} & \text{se } \nu_2 > 2 \\ \\ \mathrm{Var}(X) = \dfrac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)} & \text{se } \nu_2 > 4 \end{cases} \tag{9.18}$$

O teorema a seguir será fundamental na inferência sobre variâncias de duas populações normais.

TEOREMA 9.1 ***Sejam U e V duas variáveis aleatórias independentes tais que $U \sim \chi^2_n$ e $V \sim \chi^2_m$. Então***

$$W = \frac{U/n}{V/m} \sim F_{n,m} \tag{9.19}$$

Assim, os graus de liberdade da distribuição F referem-se aos graus de liberdade das duas variáveis qui-quadrado.

▼

Na Figura 9.1 temos o gráfico da distribuição F com $\nu_1 = 5$ graus de liberdade no numerador e $\nu_2 = 5, 10, 40$ graus de liberdade no denominador. Na Figura 9.2 fixa-se $\nu_2 = 5$ e varia-se $n_1 = 5, 10, 40$.

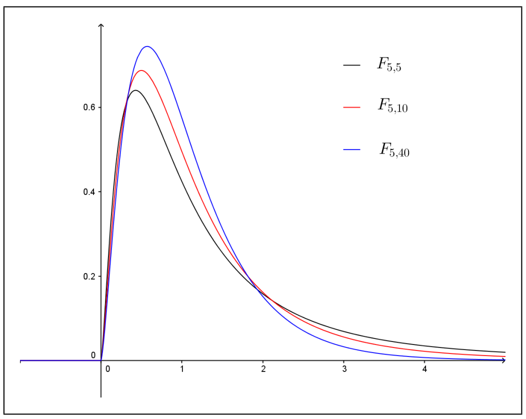


Figura 9.1 – $F_{\nu_1,\nu_2} : \nu_1 = 5; \nu_2 = 5, 10, 40$

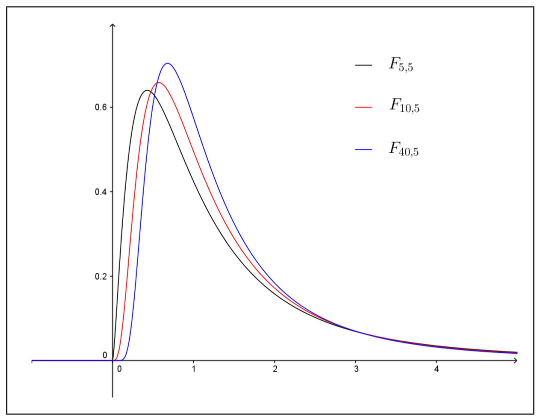


Figura 9.2 – $F_{\nu_1,\nu_2} : \nu_1 = 5, 10, 40; \nu_2 = 5$

Vamos denotar por $F_{n,m;\alpha}$ o valor crítico da distribuição $F_{n,m}$, isto é

$$\mathrm{P}(F_{n,m} \geq F_{n,m;\alpha}) = \alpha \tag{9.20}$$

Na Figura 9.3 ilustra-se o conceito do valor crítico de uma distribuição F e nas Tabelas V, VI e VII do Apêndice A exibem-se tais valores críticos para $\alpha = 0,05$, $\alpha = 0,025$ e $\alpha = 0,01$ com graus de liberdade específicos. Programas estatísticos devem ser usados para se calcular um valor crítico qualquer.

Uma propriedade importante dos valores críticos da distribuição F é a seguinte. Seja $k = F_{n,m;\alpha}$; logo,

$$\mathrm{P}(F_{n,m} > k) = \alpha \Leftrightarrow \mathrm{P}\left(\frac{1}{F_{n,m}} < \frac{1}{k}\right) = \alpha \Leftrightarrow \mathrm{P}\left(\frac{1}{F_{n,m}} \geq \frac{1}{k}\right) = 1 - \alpha \Leftrightarrow$$

$$\mathrm{P}\left(\frac{1}{\dfrac{\chi^2_n/n}{\chi^2_m/m}} \geq \frac{1}{k}\right) = 1 - \alpha \Leftrightarrow \mathrm{P}\left(\frac{\chi^2_m/m}{\chi^2_n/n} \geq \frac{1}{k}\right) = 1 - \alpha \Leftrightarrow \mathrm{P}\left(F_{m,n} \geq \frac{1}{k}\right) = 1 - \alpha \Leftrightarrow \frac{1}{k} = F_{m,n;1-\alpha}$$

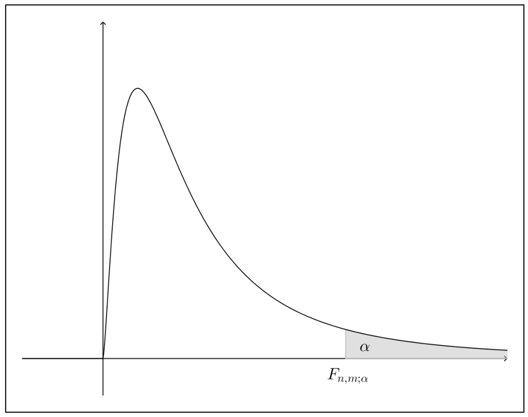


Figura 9.3 – Valor crítico da distribuição $F_{n,m}$

Resulta, assim, que

$$F_{n,m;\alpha} = \frac{1}{F_{m,n;1-\alpha}} \tag{9.21}$$

EXEMPLO 9.1

Vamos determinar k tal que $\mathrm{P}(F_{5,10} < k) = 0,05$.

Solução

$$\mathrm{P}(F_{5,10} < k) = 0,05 \Leftrightarrow \mathrm{P}(F_{5,10} \geq k) = 0,95 \Leftrightarrow k = F_{5,10;0,95} = \frac{1}{F_{10,5;0,05}} = \frac{1}{4,735} = 0,21119$$

♦♦

9.5.2 Comparação das variâncias de duas populações normais

Consideremos, agora, a situação em que nosso interesse está na comparação das variâncias σ_1^2 e σ_2^2 de duas populações normais. Os estimadores não viesados para essas variâncias são S_1^2 e S_2^2 em que

$$S_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_{1i} - \overline{X}_1)^2 \qquad S_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(X_{2i} - \overline{X}_2)^2$$

Se as populações são normais, isto é, $X_1 \sim N(\mu_1; \sigma_1^2)$ e $X_2 \sim N(\mu_2; \sigma_2^2)$ então

$$U_1 = \frac{(n_1 - 1)S_1^2}{\sigma_1^2} \sim \chi^2_{n_1-1} \qquad U_2 = \frac{(n_2 - 1)S_2^2}{\sigma_2^2} \sim \chi^2_{n_2-1}$$

Como estamos supondo que as amostras são independentes, as duas variáveis aleatórias U_1 e U_2 acima também são independentes e, portanto, pelo Teorema 9.1, resulta que

$$\frac{\dfrac{U_1}{n_1 - 1}}{\dfrac{U_2}{n_2 - 1}} \sim F_{n_1-1,n_2-1}$$

ou seja,

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F_{n_1-1,n_2-1} \tag{9.22}$$

9.5.3 Intervalo de confiança para σ_1^2/σ_2^2

Usando (9.22) e os valores críticos da distribuição F_{n_1-1,n_2-1} obtemos que

$$\begin{aligned}
P\left(F_{n_1-1,n_2-1;1-\alpha/2} < \frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} < F_{n_1-1,n_2-1;\alpha/2}\right) &= 1-\alpha \Rightarrow \\
P\left(\frac{F_{n_1-1,n_2-1;1-\alpha/2}}{S_1^2/S_2^2} < \frac{1}{\sigma_1^2/\sigma_2^2} < \frac{F_{n_1-1,n_2-1;\alpha/2}}{S_1^2/S_2^2}\right) &= 1-\alpha \Rightarrow \\
P\left(\frac{S_1^2/S_2^2}{F_{n_1-1,n_2-1;\alpha/2}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2/S_2^2}{F_{n_1-1,n_2-1;1-\alpha/2}}\right) &= 1-\alpha
\end{aligned}$$

e, portanto, o intervalo de confiança para $\dfrac{\sigma_1^2}{\sigma_2^2}$ é

$$\left(\frac{S_1^2/S_2^2}{F_{n_1-1,n_2-1;\alpha/2}}; \frac{S_1^2/S_2^2}{F_{n_1-1,n_2-1;1-\alpha/2}}\right) \tag{9.23}$$

ou

$$\left(\frac{S_1^2/S_2^2}{F_{n_1-1,n_2-1;\alpha/2}}; \frac{S_1^2/S_2^2}{\dfrac{1}{F_{n_2-1,n_1-1;\alpha/2}}}\right) \tag{9.24}$$

9.5.4 Teste de hipótese sobre σ_1^2/σ_2^2

Vamos, agora, estabelecer os procedimentos para o teste de hipóteses referentes à razão das variâncias de duas populações normais. Nosso objetivo, então, é testar

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1 \ (\sigma_1^2 = \sigma_2^2)$$

contra uma das alternativas possíveis

$$H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1 \ (\sigma_1^2 \neq \sigma_2^2) \qquad H_1 : \frac{\sigma_1^2}{\sigma_2^2} > 1 \ (\sigma_1^2 > \sigma_2^2) \qquad H_1 : \frac{\sigma_1^2}{\sigma_2^2} < 1 \ (\sigma_1^2 < \sigma_2^2)$$

Lembrando que, sob H_0, $\sigma_1 = \sigma_2$, temos as seguintes regras de decisão para um nível de significância α:

- Hipótese nula e estatística de teste

$$H_0 : \sigma_1 = \sigma_2$$

$$F_0 = \frac{S_1^2}{S_2^2} \sim F_{n_1-1,n_2-1}$$

$$f_0 = \text{valor observado de } F_0$$

- Teste bilateral

$$H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1 \ (\sigma_1^2 \neq \sigma_2^2)$$

Região crítica:

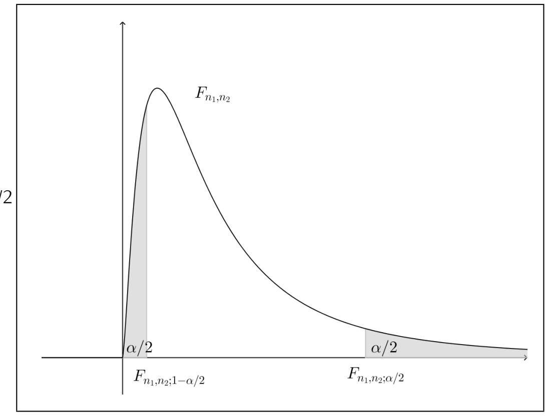
$F_0 < F_{n_1-1,n_2-1;1-\alpha/2}$ ou $F_0 > F_{n_1-1,n_2-1;\alpha/2}$

Valor P :

$P = 2 \cdot \mathrm{P}(F_{n_1-1,n_2-1} > f_0)$ se $f_0 > 1$

ou

$P = 2 \cdot \mathrm{P}(F_{n_1-1,n_2-1} < f_0)$ se $f_0 < 1$



- Teste unilateral à direita

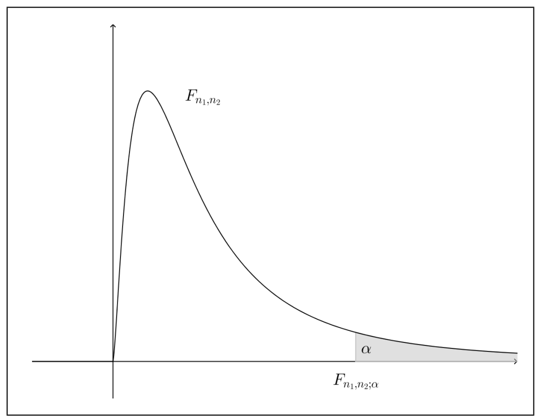
$$H_1 : \frac{\sigma_1^2}{\sigma_2^2} > 1 \ (\sigma_1^2 > \sigma_2^2)$$

Região crítica:

$F_0 > F_{n_1-1,n_2-1;\alpha}$

Valor P :

$P = \mathrm{P}(F > f_0)$



- Teste unilateral à esquerda

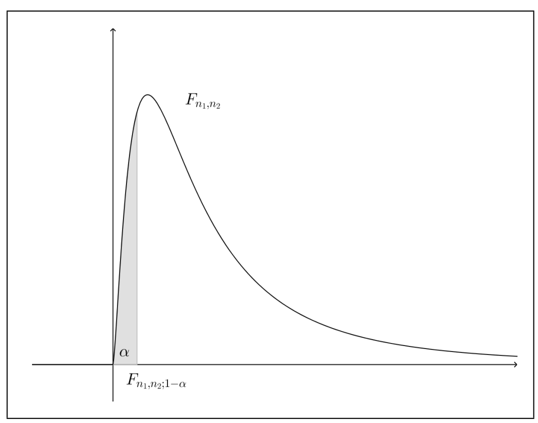
$$H_1 : \frac{\sigma_1^2}{\sigma_2^2} < 1 \ (\sigma_1^2 < \sigma_2^2)$$

Região crítica:

$$F_0 < F_{n_1-1,n_2-1;1-\alpha} = \frac{1}{F_{n_2-1,n_1-1;\alpha}}$$

Valor P :

$P = \mathrm{P}(F < f_0)$



9.6 Inferência sobre médias de duas populações normais com variâncias desconhecidas

9.6.1 Variâncias populacionais iguais

Suponhamos, agora, que $X_1 \sim N(\mu_1; \sigma^2)$ e $X_2 \sim N(\mu_2; \sigma^2)$ e que σ^2, a variância comum, seja desconhecida. Dos resultados já vistos sobre as distribuições da média e da variância amostrais de uma população normal, temos que

$$\overline{X}_1 \sim N\left(\mu_1; \frac{\sigma^2}{n_1}\right) \qquad \frac{(n_1-1)S_1^2}{\sigma^2} \sim \chi^2_{n_1-1} \qquad \overline{X}_1\ ,\ S_1^2 \text{ independentes} \tag{9.25}$$

$$\overline{X}_2 \sim N\left(\mu_2; \frac{\sigma^2}{n_2}\right) \qquad \frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2_{n_2-1} \qquad \overline{X}_2\ ,\ S_2^2 \text{ independentes} \tag{9.26}$$

Como estamos supondo que as amostras são independentes, resulta que

$$\overline{X}_1 - \overline{X}_2 \sim N\left(\mu_1 - \mu_2; \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}\right) \tag{9.27}$$

$$\frac{(n_1-1)S_1^2}{\sigma^2} + \frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2_{n_1+n_2-2} \tag{9.28}$$

$$\overline{X}_1 \text{ independente de } \overline{X}_2, S_2^2 \tag{9.29}$$

$$S_1^2 \text{ independente de } \overline{X}_2, S_2^2 \tag{9.30}$$

De (9.25) e (9.29) resulta que $\overline{X}_1$ é independente de S_1^2 e S_2^2 e de (9.26) e (9.30) resulta que $\overline{X}_2$ é independente de S_1^2 e S_2^2. Obtemos, então, que as variáveis dadas em (9.27) e 9.28) são independentes e, portanto

$$T = \frac{\dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}}{\sqrt{\dfrac{\dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{\sigma^2}}{n_1+n_2-2}}} \sim t_{n_1+n_2-2}$$

ou seja

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_C^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} \sim t_{n_1+n_2-2} \tag{9.31}$$

em que

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \tag{9.32}$$

é um estimador não viesado para σ^2, obtido como média ponderada dos estimadores não viesados S_1^2 e S_2^2, com os pesos sendo definidos pelos graus de liberdade.

A estatística dada em (9.31) será utilizada na construção de testes de hipóteses e intervalos de confiança para a diferença entre as médias.

Intervalo de confiança para $\mu_1 - \mu_2$

Temos, a partir de (9.31), que

$$\mathrm{P}\left(-t_{n_1+n_2-2;\alpha/2} < \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} < t_{n_1+n_2-2;\alpha/2}\right) = 1 - \alpha$$

Logo,

$$
\begin{aligned}
& P\left((\overline{X}_1 - \overline{X}_2) - t_{n_1+n_2-2;\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} < (\mu_1 - \mu_2)\right. \\
< \quad & \left.(\overline{X}_1 - \overline{X}_2) + t_{n_1+n_2-2;\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right) = 1 - \alpha
\end{aligned}
$$

o que nos dá o seguinte intervalo de confiança de $100(1-\alpha)\%$ para $\mu_1 - \mu_2$.

$$
\left((\overline{X}_1 - \overline{X}_2) - t_{n_1+n_2-2;\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}; (\overline{X}_1 - \overline{X}_2) + t_{n_1+n_2-2;\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right) \tag{9.33}
$$

Como antes, os limites são variáveis aleatórias e, assim, faz sentido falar que a probabilidade é $1-\alpha$. Isso significa que se repetirmos várias vezes o processo de amostragem e subsequente construção do intervalo de confiança correspondente, "acertaremos" em $100(1-\alpha)\%$ das vezes, ou seja, o intervalo conterá o verdadeiro parâmetro $\mu_1 - \mu_2$.

Para uma amostra específica, o intervalo de confiança é

$$
\left((\overline{x}_1 - \overline{x}_2) - t_{n_1+n_2-2;\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}; (\overline{x}_1 - \overline{x}_2) + t_{n_1+n_2-2;\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right) \tag{9.34}
$$

e esse intervalo ou contém, ou não contém o verdadeiro parâmetro. O que podemos dizer é que o método de obtenção dos intervalos garante uma probabilidade de acerto de $100(1-\alpha)\%$.

Teste de hipótese sobre $\mu_1 - \mu_2$

Vamos estabelecer, agora, os procedimentos para o teste de hipóteses referentes à diferença entre as médias, ou seja, nosso objetivo é testar

$$
H_0 : \mu_1 = \mu_2 \qquad \text{ou} \qquad H_0 : \mu_1 - \mu_2 = \Delta_0
$$

Como antes, o procedimento de teste consiste em rejeitar a veracidade da hipótese nula H_0 sempre que obtivermos valores da estatística de teste com pequenas probabilidades de ocorrência sob H_0. Na distribuição t, assim como na normal, probabilidades pequenas correspondem às caudas da distribuição e isso nos leva aos seguintes procedimentos.

- Hipótese nula e estatística de teste

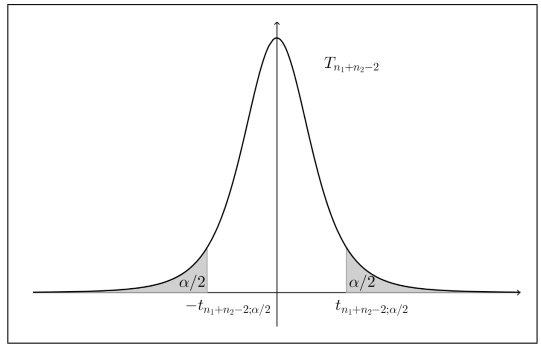
$$
\begin{aligned}
& H_0 : \mu_1 - \mu_2 = \Delta_0 \\
& T_0 = \frac{(\overline{X}_1 - \overline{X}_2) - \Delta_0}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2} \\
& t_0 = \text{ valor observado de } T_0
\end{aligned}
$$

- Teste bilateral

$H_1 : \mu_1 - \mu_2 \neq \Delta_0$

Região crítica:
$|T_0| > t_{n_1+n_2-2;\alpha/2}$

Valor P :
$P = 2 \cdot \mathrm{P}(T_{n_1+n_2-2} > |t_0|)$

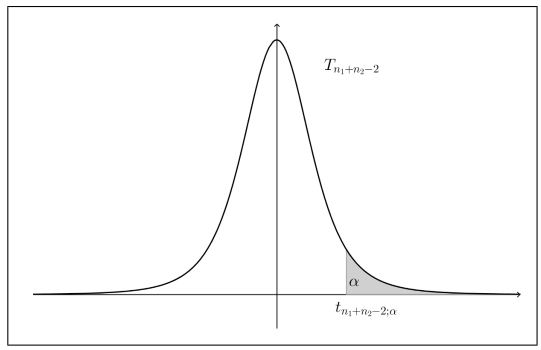


- Teste unilateral à direita

$H_1 : \mu_1 - \mu_2 > \Delta_0$

Região crítica:
$T_0 > t_{n_1+n_2-2;\alpha}$

Valor P :
$P = \mathrm{P}(T_{n_1+n_2-2} > t_0)$

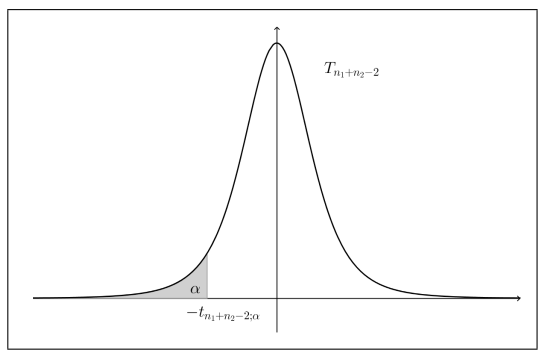


- Teste unilateral à esquerda

$H_1 : \mu_1 - \mu_2 < \Delta_0$

Região crítica:
$T_0 < -t_{n_1+n_2-2;\alpha}$

Valor P :
$P = \mathrm{P}(T_{n_1+n_2-2} > |t_0|)$



9.6.2 Variâncias populacionais diferentes

Consideremos, agora, o caso mais geral em que $X_1 \sim N(\mu_1; \sigma_1^2)$ e $X_2 \sim N(\mu_2; \sigma_2^2)$ e que ambas as variâncias são desconhecidas. Cada uma das variâncias amostrais S_1^2 e S_2^2 estima a variância populacional correspondente, mas a padronização fornece apenas uma estatística de teste aproximadamente distribuída como uma t de Student. Mais precisamente

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \approx t_\nu \tag{9.35}$$

em que o número de graus de liberdade ν é dado por

$$\nu = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}} \tag{9.36}$$

Uma expressão alternativa, que é mais fácil e fornece resultados mais precisos numericamente, é

$$\nu = \frac{(n_1 - 1)(n_2 - 1)}{(n_2 - 1)C^2 + (n_1 - 1)(1 - C)^2} \qquad \text{em que} \qquad C = \frac{\dfrac{s_1^2}{n_1}}{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}} \tag{9.37}$$

Uma abordagem conservadora é considerar

$$\nu = \min(n_1 - 1, n_2 - 1) \tag{9.38}$$

Em qualquer dos casos, a construção de testes de hipóteses e intervalos de confiança se faz de maneira análoga com base na distribuição t_ν.

Intervalo de confiança para $\mu_1 - \mu_2$

Temos, a partir de (9.31), que

$$P\left(-t_{\nu;\alpha/2} < \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} < t_{\nu;\alpha/2}\right) \approx 1 - \alpha$$

Logo,

$$P\left((\overline{X}_1 - \overline{X}_2) - t_{\nu;\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < (\mu_1 - \mu_2) < (\overline{X}_1 - \overline{X}_2) + t_{\nu;\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}\right) = 1 - \alpha$$

o que nos dá o seguinte intervalo de confiança de $100(1 - \alpha)\%$ para $\mu_1 - \mu_2$.

$$\left((\overline{X}_1 - \overline{X}_2) - t_{\nu;\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}; (\overline{X}_1 - \overline{X}_2) + t_{\nu;\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}\right) \tag{9.39}$$

Para uma amostra específica, o intervalo de confiança é

$$\left((\overline{x}_1 - \overline{x}_2) - t_{\nu;\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}; (\overline{x}_1 - \overline{x}_2) + t_{\nu;\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right) \tag{9.40}$$

e esse intervalo ou contém, ou não contém o verdadeiro parâmetro. O que podemos dizer é que o método de obtenção dos intervalos garante uma probabilidade de acerto de $100(1 - \alpha)\%$.

Observação: Quando ambos os tamanhos amostrais n_1 e n_2 são grandes, $\overline{X}_1 - \overline{X}_2$ é aproximadamente normal e podemos usar a distribuição normal como aproximação das distribuições amostrais acima.

Teste de hipótese sobre $\mu_1 - \mu_2$

- Hipótese nula e estatística de teste

$$H_0 : \mu_1 - \mu_2 = \Delta_0$$

$$T_0 = \frac{(\overline{X}_1 - \overline{X}_2) - \Delta_0}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \approx t_\nu$$

$$\nu = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}} \quad \text{ou} \quad \nu = \min(n_1 - 1, n_2 - 1)$$

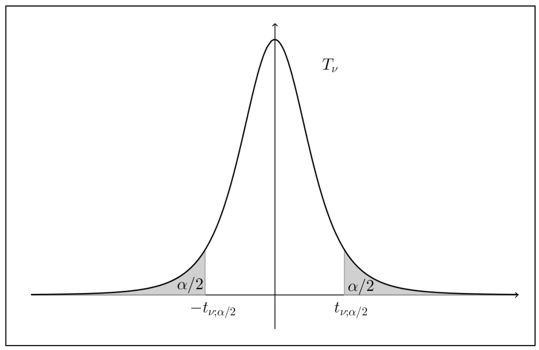
$$t_0 = \text{ valor observado de } T_0$$

- Teste bilateral

$H_1 : \mu_1 - \mu_2 \neq \Delta_0$

Região crítica:
$|T_0| > t_{\nu;\alpha/2}$

Valor P :
$P = 2 \cdot \mathrm{P}(T_\nu > |t_0|)$

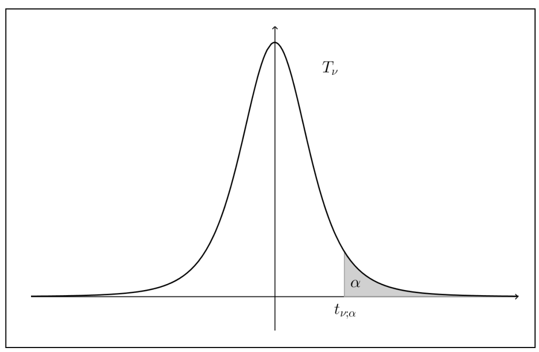


- Teste unilateral à direita

$H_1 : \mu_1 - \mu_2 > \Delta_0$

Região crítica:
$T_0 > t_{\nu;\alpha}$

Valor P :
$P = \mathrm{P}(T_\nu > t_0)$

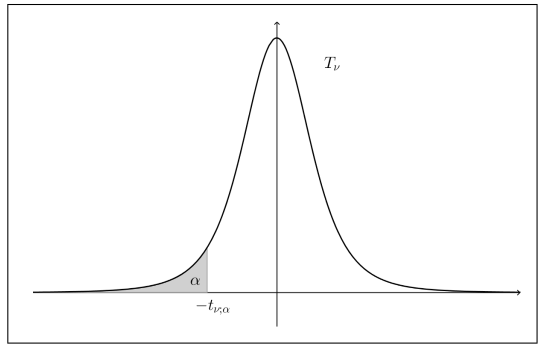


- Teste unilateral à esquerda

$H_1 : \mu_1 - \mu_2 < \Delta_0$

Região crítica:
$T_0 < -t_{\nu;\alpha}$

Valor P :
$P = \mathrm{P}(T_\nu > |t_0|)$



EXEMPLO 9.2

Em muitos estados americanos, os advogados são encorajados a realizar trabalhos *pro bono*, tanto por suas firmas quanto pelos conselhos de assessoramento judicial. No entanto, em anos recentes, os advogados têm dedicado mais tempo a clientes particulares e menos tempo a ajuda legal *pro bono*. Obtiveram-se amostras aleatórias independentes de advogados de duas grandes firmas, e o número de horas *pro bono* durante o ano anterior foi registrado para cada advogado. As estatísticas-resumo são apresentadas na tabela que segue. Admita que as populações subjacentes sejam normais.

Firma	Dados da amostra		
	Tamanho	Média	Variância
A	18	75,1	5,92
B	14	80,9	5,65

(a) Há alguma evidência que sugira que o número médio de horas anuais *pro bono* seja diferente nessas duas firmas de advocacia? Use $\alpha = 0,05$.

(b) Construa um intervalo de confiança de 95% para a diferença nas horas médias *pro bono*. Esse intervalo de confiança apoia sua conclusão da parte (a)? Explique.

Solução

(a) Para decidir qual teste usar para a comparação das médias, temos, inicialmente, que testar a igualdade das variâncias.

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

$$\begin{aligned} F_0 &= \frac{S_1^2}{S_2^2} \sim F_{17,13} \\ F_{17,13;0,025} &= 2,983 \\ F_{17,13;0,975} &= \frac{1}{F_{13,17;0,025}} = \frac{1}{2,753} = 0,3632 \end{aligned}$$

A região crítica é $F_0 > 2,983$ ou $F_0 < 0,3632$. O valor observado da estatística de teste é $f_0 = \dfrac{5,92}{5,65} = 1,048$, que não pertence à região crítica. Logo, não rejeitamos a hipótese de igualdade das variâncias e passamos a realizar o teste para comparação de médias com base na hipótese de igualdade de variâncias populacionais.

O estimador combinado da variância é

$$S_p^2 = \frac{n_1 - 1}{n_1 + n_2 - 2} S_1^2 + \frac{n_2 - 1}{n_1 + n_2 - 2} S_2^2 = \frac{17}{30} \times 5,92 + \frac{13}{30} \times 5,65 = 5,803$$

Queremos testar

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

e a estatística de teste é

$$T_0 = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \sim t_{30}$$

O valor crítico é $t_{30,0,025} = 2,042$ e rejeitamos H_0 se $|t_0| > 2,042$. Os dados fornecem

$$t_0 = \frac{75,1 - 80,9}{\sqrt{5,803 \left(\frac{1}{18} + \frac{1}{14} \right)}} = -6,7566$$

que está na região crítica. Assim, rejeita-se a hipótese de que o número médio de horas *pro bono* seja o mesmo para as duas firmas.

Na Tabela IV, vemos que, para 30 graus de liberdade, a maior abscissa é 3,385, que é menor que $|t_0| = 6,7566$. Assim, podemos afirmar que o valor P é menor que 0,001.

(b) O intervalo de confiança é

$$(75,1 - 80,9) \pm 2,042 \sqrt{5,803 \left(\frac{1}{18} + \frac{1}{14} \right)} = (-7,5529\,;\, -4,0471)$$

que não contém o 0, o que corrobora a rejeição da hipótese de igualdade das médias no item anterior.

♦♦

9.7 Exercícios Propostos

1. O peso total (com a caixa) de uma máquina de costura portátil é uma consideração importante. Suponha que Singer afirme ter a máquina mais leve, com menos 5 libras (2,77 kg) em relação às demais. Obtiveram-se amostras aleatórias independentes de uma Singer e de uma máquina comparável Simplicity, e os pesos (em libras) de cada uma foi registrado. As estatísticas-resumo e as variâncias conhecidas são apresentadas na tabela que segue.

Máquina de costura	Tamanho amostral	Média amostral	Variância populacional
Simplicity	42	17,99	2,89
Singer	38	13,26	2,25

   (a) Há alguma evidência para se refutar a afirmativa da Singer? Use $\alpha = 0,01$.

   (b) Ache o valor P associado a esse teste.

   (c) A hipótese de normalidade é necessária nesse problema? Por que ou por que não?

2. Magnésio é usado por todas as células de seu corpo e ajuda os nervos a funcionarem adequadamente. De acordo com a Base de Dados sobre Nutrientes do Departamento de Agricultura dos Estados Unidos, meia xícara de feijão vegetariano cozido e uma batata média cozida sem a casca têm a mesma quantidade de magnésio (40 miligramas). Para verificar essa afirmação, obtiveram-se amostras aleatórias independentes de feijão cozido e batatas cozidas, e mediu-se a quantidade de magnésio em cada porção (em miligramas). As estatísticas-resumo e as variâncias conhecidas são apresentadas na tabela que segue.

Alimento	Tamanho amostral	Média amostral	Variância populacional
Feijão vegetariano cozido	18	39,58	2,47
Batata cozida	18	40,12	0,87

   (a) Admita que as distribuições subjacentes sejam normais. Há alguma evidência para se refutar a afirmativa? Use $\alpha = 0,01$. Ache o valor P para esse teste de hipótese.

   (b) Suponha que os tamanhos amostrais sejam $n_1 = n_2 = 38$. Agora, há alguma evidência para se refutar a afirmativa? Ache o valor P para esse teste de hipótese.

   (c) Quais devem ser os tamanhos amostrais ($n_1 = n_2$) para que o teste de hipótese seja significante ao nível $\alpha = 0,01$?

3. Latas de alumínio são feitas a partir de grandes lingotes sólidos, prensados sob rolos de alta pressão e cortados como biscoitos a partir de folhas finas. O alumínio é ideal para latas porque é leve, forte e reciclável. Uma companhia afirma que um novo processo de fabricação diminui a quantidade de alumínio necessária para se fazer uma lata e, portanto, diminui o peso. Obtiveram-se amostras aleatórias independentes de latas de alumínio feitas pelos processos velho e novo, e o peso (em onças) de cada uma é dado na tabela que segue.

Processo antigo (1)										
0,52	0,49	0,47	0,47	0,48	0,52	0,55	0,49	0,52	0,50	0,50
0,50	0,51	0,51	0,50	0,53	0,49	0,51	0,52	0,51	0,51	
Processo novo (2)										
0,51	0,51	0,50	0,48	0,47	0,49	0,46	0,46	0,52	0,50	0,48
0,51	0,50	0,48	0,51	0,44	0,48	0,47	0,50	0,51	0,48	

   Há alguma evidência de que as latas de alumínio feitas pelo novo processo tenham um peso médio populacional menor? Admita que as populações sejam normais, com variâncias iguais, e use $\alpha = 0,01$. Por que você acha que um nível de significância pequeno seja importante aqui?

4. Para ajudar os lojistas em seu planejamento, a cada ano se realiza um estudo para se determinar quanto as pessoas pretendem gastar com presentes nas festas de fim de ano. Em uma pesquisa de novembro de 2008, obteve-se uma amostra de compradores e lhes foi pedido que estimassem a quantia que pretendiam gastar (em dólares) com presentes. A média amostral dos gastos antecipada foi relatada por gênero, grupo de idade, e nível de renda. Considere as estatísticas-resumo dadas na tabela que segue.

Grupo de	Tamanho amostral	Média amostral	Desvio padrão amostral
Homens	21	784,00	37,50
Mulheres	19	652,00	17,01

   Historicamente, os homens relatam gastos maiores do que os das mulheres. Com base nos dados de 2008, há alguma evidência que sugira que a quantidade média que os homens pretendem gastar seja maior do que a quantidade média que as mulheres pretendem gastar? Use $\alpha = 0,10$, e admita que as populações sejam normais.

5. Em 2008, Minnesota e Carolina do Norte criaram a maioria dos perus dos Estados Unidos, aproximadamente 49 milhões e 39 milhões de aves, respectivamente. Obtiveram-se amostras aleatórias independentes de perus congelados de cada estado, e cada peru foi pesado. Os dados resultantes (em libras) são apresentados nas tabela que segue.

Minnesota							
10,1	11,5	17,1	13,4	15,9	17,9	14,9	9,5
14,5	12,5	14,2	16,8	13,7	16,0	19,4	11,4
Carolina do Norte							
19,9	14,0	19,9	12,3	17,0	25,2	23,9	7,8
15,8	21,2	13,8	7,4	15,1	10,1	3,2	17,2

Há alguma evidência que sugira que haja maior variabilidade no peso de perus congelados da Carolina do Norte do que de Minnesota? Use $\alpha = 0,01$ e admita normalidade.

Capítulo 10

Inferência com Amostras Dependentes

Em vários estudos de duas populações, não é possível assumir independência entre as amostras. Experimentos que envolvem medidas de cada indivíduo ou objeto antes e depois de algum evento resultam em dados emparelhados – cada observação antes está associada a, ou emparelhada com uma observação depois. Aqui as variáveis X_1 e X_2 são medidas no mesmo sujeito. Outros exemplos podem envolver marido/mulher ou irmãos gêmeos, casos em que temos n pares de indivíduos e as variáveis X_1 e X_2 são medidas em cada um dos indivíduos do par. Embora as variáveis se refiram a sujeitos diferentes, não é razoável supor independência.

DEFINIÇÃO **Amostras Dependentes ou Emparelhadas**

Em um conjunto de dados emparelhados, cada indivíduo ou objeto na amostra 1 está associado com um indivíduo ou objeto semelhante na amostra 2. *Semelhante* significa que os indivíduos ou objetos compartilham alguma característica fundamental, comum, podendo, ou não, ser o mesmo indivíduo ou objeto.

Consideremos, agora, o caso de amostras emparelhadas, em que duas variáveis X_1 e X_2 são medidas em um mesmo indivíduo da amostra ou nos indivíduos de cada par da amostra. Como antes, vamos supor que $X_1 \sim N(\mu_1; \sigma_1^2)$ e $X_2 \sim N(\mu_2; \sigma_2^2)$. Nosso interesse continua sendo a diferença $\mu_1 - \mu_2$. O que muda agora é que $\overline{X}_1$ e $\overline{X}_2$ não são mais independentes.

Podemos, então, pensar na nossa amostra como sendo uma amostra de pares $(X_{11}, X_{21}), (X_{12}, X_{22}), \dots, (X_{1n}, X_{2n})$ de observações retiradas das populações X_1 e X_2, respectivamente. Para cada par definimos uma nova variável

$$D_i = X_{1i} - X_{2i}$$

Segue que $D_1, D_2, \dots, D_n$ formam uma amostra aleatória simples da variável $D = X_1 - X_2$, pois cada uma delas se refere a um indivíduo/objeto ou par diferente e, portanto, são independentes. Como X_1 e X_2 têm distribuição normal, D também tem distribuição normal com média $\mu_1 - \mu_2$. A variância de D pode ser estimada a partir da amostra $D_1, D_2, \ldots, D_n$ como

$$S_D^2 = \frac{1}{n-1}\sum_{i=1}^{n}(D_i - \overline{D})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} D_i^2 - n\overline{D}^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n} D_i^2 - \frac{\left(\sum_{i=1}^{n} D_i\right)^2}{n}\right) \tag{10.1}$$

em que

$$\overline{D} = \frac{1}{n}\sum_{i=1}^{n} D_i = \overline{X}_1 - \overline{X}_2 \tag{10.2}$$

10.1 Intervalo de confiança para $\mu_1 - \mu_2$

Sabemos que

$$\sqrt{n}\frac{\overline{D} - \mu_D}{S_D} \sim t_{n-1}$$

e, portanto, o intervalo de confiança para $\mu_1 - \mu_2$ é

$$\left((\overline{X}_1 - \overline{X}_2) - t_{n-1;\alpha/2}\frac{S_D}{\sqrt{n}}\,;\,(\overline{X}_1 - \overline{X}_2) + t_{n-1;\alpha/2}\frac{S_D}{\sqrt{n}} \right) \tag{10.3}$$

10.2 Teste de hipótese sobre $\mu_1 - \mu_2$

Os procedimentos de inferência sobre $\mu_1 - \mu_2$ se basearão na distribuição t de Student com $n-1$ graus de liberdade.

- Hipótese nula e estatística de teste

$$H_0 : \mu_D = \mu_1 - \mu_2 = \Delta_0$$

$$T_{D_0} = \frac{\overline{D} - \Delta_0}{S_D/\sqrt{n}} \sim t_{n-1} \text{ sob } H_0$$

$$t_{D_0} = \text{ valor observado de } T_{D_0}$$

- Teste bilateral

$$H_1 : \mu_D \neq \Delta_0$$

$$\text{Região crítica :} \quad |T_{D_0}| > t_{n-1;\alpha/2}$$

$$\text{Valor } P: \quad P = 2 \cdot \mathrm{P}(T_{n-1} > |t_{D_0}|)$$

- Teste unilateral à direita

$$H_1 : \mu_D > \Delta_0$$

$$\text{Região crítica :} \quad T_{D_0} > t_{n-1;\alpha}$$

$$\text{Valor } P: \quad P = \mathrm{P}(T_{n-1} > t_{D_0})$$

- Teste unilateral à esquerda

$$H_1 : \mu_D < \Delta_0$$

$$\text{Região crítica} \quad T_{D_0} < -t_{n-1;\alpha}$$

$$\text{Valor } P: \quad P = \mathrm{P}(T_{n-1} > |t_{D_0}|)$$

EXEMPLO 10.1

O gerente de uma loja de conveniências loja está considerando colocar novas caixas registradoras visando aumentar a precisão e diminuir o tempo de saída. Reuniu-se uma amostra aleatória de sete compras típicas de itens da loja. Cada sacola de compras dos itens foi totalizada por um operador de caixa usando a máquina antiga e, depois, pelo mesmo operador usando a máquina nova. Os tempos (em segundos) são apresentados na tabela que segue. Há alguma evidência que sugira que os tempos de saída sejam diferentes com as duas registradoras? Use $\alpha = 0,01$, e encontre limites para o valor P associado a esse teste de hipótese.

Sacola de compras	1	2	3	4	5	6	7
Registradora antiga	45	83	62	65	39	66	62
Registradora nova	42	55	45	44	17	66	69

Solução

Note que, nesse exemplo, a característica comum, que emparelha os dados, é o operador de caixa. Vamos tomar os tempos associados à registradora antiga como a população 1 e os tempos associados à registradora nova como população 2. Nas Figuras 10.2 e 10.1 temos a saida do Minitab para o teste de Anderson-Darling, que mostra que evidência de normalidade dos dados. Podemos, então, seguir para realizar o teste t de dados emparelhados.

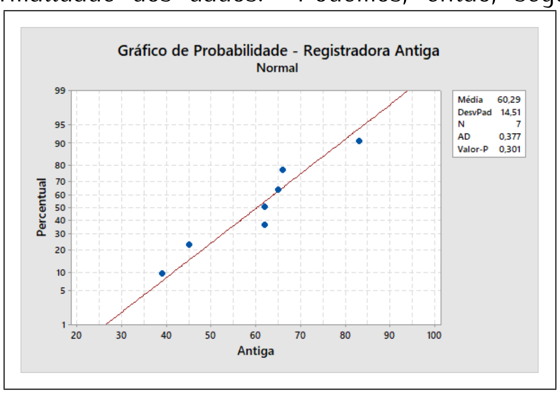


Figura 10.1 – Registradora Antiga

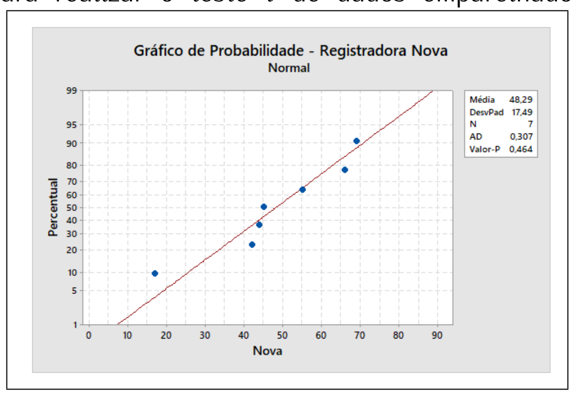


Figura 10.2 – Registradora Nova

Sacola de compras	1	2	3	4	5	6	7
Registradora antiga X_1	45	83	62	65	39	66	62
Registradora nova X_2	42	55	45	44	17	66	69
Diferença $D = X_1 - X_2$	3	28	17	21	22	0	-7

Queremos testar

$$H_0 : \mu_1 - \mu_2 = 0 \quad (\mu_1 = \mu_2)$$
$$H_1 : \mu_1 - \mu_2 \neq 0 \quad (\mu_1 \neq \mu_2)$$

A estatística de teste é

$$T = \frac{\overline{D}}{\frac{S_D}{\sqrt{n}}} \sim t_6$$

e o valor crítico é $t_{6;0,005} = 3,707$. Os dados nos fornecem a média amostral

$$\overline{d} = \frac{3 + 28 + 17 + 21 + 22 + 0 - 7}{7} = \frac{84}{7} = 12$$

e a variância amostral (note a fórmula de cálculo mais simples)

$$\begin{aligned} s_D^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \overline{d})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} d_i^2 - n\overline{d}^2\right) \\ &= \frac{1}{6}(9 + 784 + 289 + 441 + 484 + 0 + 49 - 7 \times 144) = \frac{1048}{6} = 174,6667 \end{aligned}$$

Logo, o valor observado da estatística de teste é

$$t_0 = \frac{12}{\sqrt{\frac{174,6667}{7}}} = 2,4023 < 3,707$$

Sendo assim, não se rejeita a hipótese nula, ou seja, os dados indicam que os tempos gastos com as registradoras novas não são, em média, diferentes dos tempos gastos com as registradoras antigas.

Olhando na Tabela IV do Apêndice A, na linha correspondente a 6 graus de liberdade, vemos que o valor observado $t_0 = 2,4023$ está entre os valores 2,313 e 2,612, que correspondem às probabilidades 0,03 e 0,02. Assim, podemos dizer que a probabilidade na cauda direita, acima de 2,4023, está no intervalo $(0,02; 0,03)$. Como o teste é bilateral, o valor P estará no intervalo $(0,04; 0,06)$. Minitab nos fornece $P = 2 \times 0,0265641 = 0,0531282$.

♦♦

10.3 Exercícios Propostos

1. Vinte e um programadores de computador de firmas TI (Tecnologia da Informação) em todo o país foram selecionados aleatoriamente. Pediu-se a cada um para escrever um código em C++ e em Java para uma aplicação específica. O tempo de execução (em segundos) de cada programa, por linguagem de computador, é dado a seguir.

C++										
44,4	43,0	46,6	41,2	44,6	44,3	47,3	49,5	46,5	46,2	42,8
47,5	43,3	41,5	45,3	45,2	48,7	47,1	44,1	43,6	45,1	
Java										
52,0	52,6	41,8	51,4	64,3	62,1	41,2	58,0	49,9	51,1	50,6
54,9	50,6	54,0	44,1	59,4	56,0	31,3	49,6	49,4	48,8	

   (a) Qual é a característica comum que torna esses dados emparelhados?
   (b) Admita normalidade. Realize o teste de hipótese apropriado para determinar se há alguma evidência de que o tempo médio de execução seja maior para programas Java do que para programas C++. Use $\alpha = 0,001$.
   (c) Ache limites para o valor P associado a esse teste de hipótese.

2. Um consultor que trabalha para o quartel da Polícia Estadual afirma que as armas de serviço dispararão com uma velocidade de boca maior se o cano estiver adequadamente limpo. Obteve-se uma amostra aleatória de armas de 9 mm, e mediu-se a velocidade de boca (em pés por segundo) de um único tiro de cada arma. Cada arma foi profissionalmente limpa e a velocidade de boca de um segundo tiro (com o mesmo tipo de bala) foi medida. Os dados são apresentados na tabela que segue.

Arma	1	2	3	4	5	6
Antes	1505	1419	1504	1494	1510	1506
Depois	1625	1511	1459	1441	1472	1521

(a) Qual é a característica comum que torna esses dados emparelhados?

(b) Admita normalidade. Realize o teste de hipótese apropriado para determinar se há alguma evidência de que uma arma limpa dispara com velocidade de boca maior. Use $\alpha = 0,01$.

(c) Ache limites para o valor P associado a esse teste de hipótese.

Parte III

Análise de variância de um fator

Capítulo 11

Análise de variância de um fator

11.1 Conceitos Básicos

No estudo da inferência para duas populações, vimos como estimar a diferença entre médias populacionais de variáveis quantitativas a partir de amostras independentes retiradas de *duas* populações. Mas, e se quisermos comparar mais de duas populações? Estudaremos, agora, o método da análise de variância ou ANOVA (do inglês <u>AN</u>alysis <u>O</u>f <u>VA</u>riance), que permite comparar médias de várias populações representadas por variáveis quantitativas. Assim como no caso de duas populações, algumas hipóteses sobre os dados – ou populações – devem ser satisfeitas. Eis algumas situações que podem se de interesse:

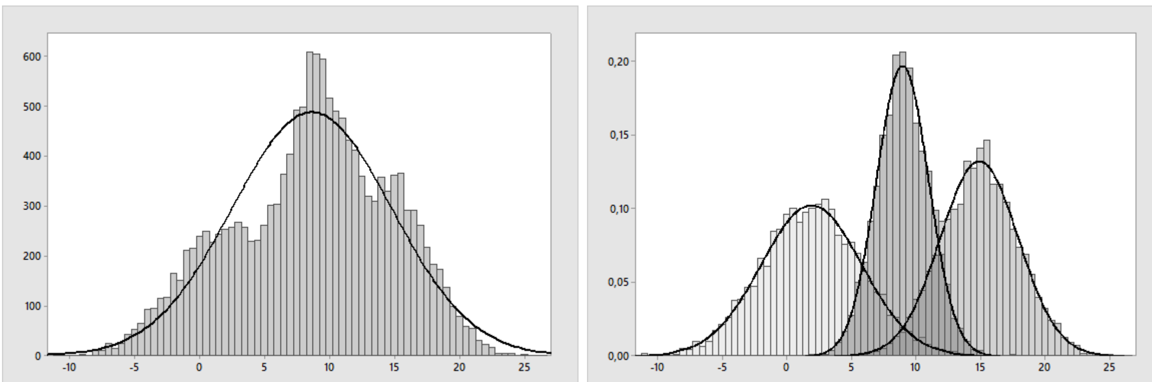
- Dois processos novos de produção de chaves devem ser comparados com o processo tradicional no intuito de se comparar o peso médio das chaves.
- Deseja-se comparar a acidez, medida pelo pH, da água de riachos em quatro grandes parques nacionais.
- Com o intuito de obter a maior produtividade, compara-se a produção de milho em lotes plantados sob quatro diferentes níveis de concentração de fertilizante.

Em todos esses exemplos, temos uma população (chaves e seu peso, riachos e seu pH, lotes e produção de milho) categorizada segundo um fator (tipo de processo no caso das chaves, os parques nacionais no caso do pH da água e os níveis de concentração de fertilizante no caso da produção de milho). Esse é o contexto da **análise da variância de um fator**: uma população, representada por uma variável quantitativa X, categorizada por um único fator com k níveis. Amostras aleatórias simples *independentes* são retiradas para cada um dos níveis do fator. Os níveis do fator muitas vezes são chamados de tratamentos, grupos ou sub-populações.

Na Tabela 11.1 apresentamos um esquema das informações básicas para uma análise de variância. Temos amostras independentes dos diferentes tratamentos e nosso objetivo é determinar se essas observações vêm de uma única população (Figura 11.1a) ou de populações distintas (Figura 11.1b).

Tabela 11.1 – Dados típicos para uma Análise de Variância de um fator

Tratamento		Amostra				Estatísticas amostrais		
1	$X_1 \sim (\mu_1; \sigma_1^2)$	X_{11}	X_{12}	$\cdots$	X_{1n_1}	$\overline{X}_{1\cdot}$	S_1^2	n_1
2	$X_2 \sim (\mu_2; \sigma_2^2)$	X_{21}	X_{22}	$\cdots$	X_{2n_2}	$\overline{X}_{2\cdot}$	S_2^2	n_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	$X_i \sim (\mu_i; \sigma_i^2)$	X_{i1}	X_{i2}	$\cdots$	X_{in_i}	$\overline{X}_{i\cdot}$	S_i^2	n_i
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	$X_k \sim (\mu_k; \sigma_k^2)$	X_{k1}	X_{k2}	$\cdots$	X_{kn_k}	$\overline{X}_{k\cdot}$	S_k^2	n_k



(a) Uma única população? **(b)** Ou 3 populações?

Figura 11.1 – Problema típico da análise de variância

11.1.1 Definições e propriedades básicas

Vamos, agora, estabelecer definições e propriedades básicas.

- Tamanho amostral total

$$n = n_1 + n_2 + \cdots + n_k \tag{11.1}$$

- Total e média amostrais para o tratamento i

$$X_{i\cdot} = \sum_{j=1}^{n_i} X_{ij} \tag{11.2}$$

$$\overline{X}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij} = \frac{1}{n_i} X_{i\cdot} \tag{11.3}$$

O primeiro subscrito em X_{ij} indica o tratamento e o segundo subscrito, o elemento da respectiva amostra. O ponto em $X_{i\cdot}$ indica que estamos somando ao longo de todos os elementos do tratamento i.

- Total e média amostrais gerais

$$X_{\cdot\cdot} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} X_{ij} = \sum_{i=1}^{k} X_{i\cdot} \tag{11.4}$$

$$\overline{X}_{\cdot\cdot} = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} X_{ij} = \frac{1}{n} X_{\cdot\cdot} \tag{11.5}$$

Note que aqui estamos tratando todos os dados conjuntamente, como se fossem amostra de uma única população (Figura 11.1a).

- Desvios em torno de médias
  - ⋆ Observações em relação à média geral

$$X_{ij} - \overline{X}_{..} \qquad \text{(segmentos em verde na Figura 11.2)} \tag{11.6}$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_{..}) = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij} - \sum_{i=1}^{k}\sum_{j=1}^{n_i} \overline{X}_{..} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij} - n\overline{X}_{..} = 0 \tag{11.7}$$

  - ⋆ Médias dos tratamentos em relação à média geral

$$\overline{X}_{i\cdot} - \overline{X}_{..} \qquad \text{(segmentos em vermelho na Figura 11.2)} \tag{11.8}$$

$$\sum_{i=1}^{k} n_i(\overline{X}_{i\cdot} - \overline{X}_{..}) = \sum_{i=1}^{k} n_i\overline{X}_{i\cdot} - \overline{X}_{..}\sum_{i=1}^{k} n_i = \sum_{i=1}^{k} n_i\frac{X_{i\cdot}}{n_i} - \frac{X_{..}}{n}n = \sum_{i=1}^{k} X_{i\cdot} - X_{..} = 0 \tag{11.9}$$

  Note que aqui levamos em consideração o número de observações em cada tratamento.

  - ⋆ Observações em relação à média do respectivo tratamento

$$X_{ij} - \overline{X}_{i\cdot} \qquad \text{(segmentos em azul na Figura 11.2)} \tag{11.10}$$

$$\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_{i\cdot}) = \sum_{j=1}^{n_i} X_{ij} - \sum_{j=1}^{n_i} \overline{X}_{i\cdot} = \sum_{j=1}^{n_i} X_{ij} - n_i\overline{X}_{i\cdot} = 0 \tag{11.11}$$

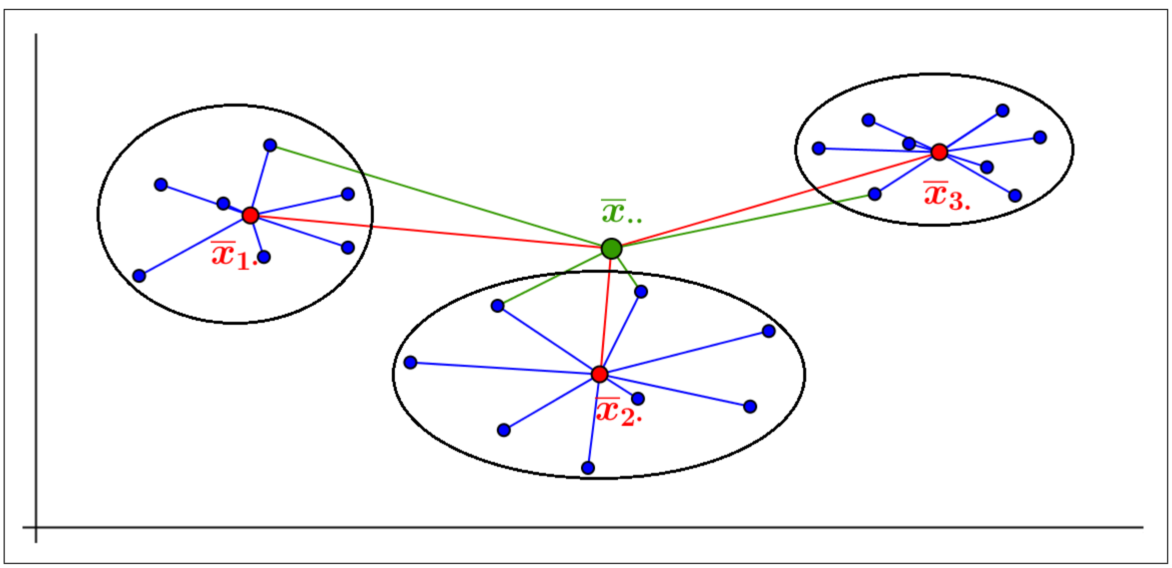


Figura 11.2 – Ilustração dos desvios em torno das médias

11.1.2 Decomposição da soma dos quadrados total

O método da análise de variância se baseia nas somas desses desvios, mas elevados ao quadrado – uma medida de variabilidade.

Note que é válida a seguinte igualdade:

$$X_{ij} - \overline{X}_{..} = X_{ij} - \overline{X}_{i\cdot} + \overline{X}_{i\cdot} - \overline{X}_{..}$$

Elevando ambos os lados ao quadrado e somando ao longo de todas as observações, obtemos

$$\begin{aligned}
\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{\cdot\cdot})^2 &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{i\cdot}+\overline{X}_{i\cdot}-\overline{X}_{\cdot\cdot})^2 \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{i\cdot})^2+\sum_{i=1}^{k}\sum_{j=1}^{n_i}(\overline{X}_{i\cdot}-\overline{X}_{\cdot\cdot})^2+2\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{i\cdot})(\overline{X}_{i\cdot}-\overline{X}_{\cdot\cdot}) \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{i\cdot})^2+\sum_{i=1}^{k}(\overline{X}_{i\cdot}-\overline{X}_{\cdot\cdot})^2\sum_{j=1}^{n_i}1+2\sum_{i=1}^{k}(\overline{X}_{i\cdot}-\overline{X}_{\cdot\cdot})\underbrace{\cancel{\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{i\cdot})}}_{0 \text{ por } 11.11}
\end{aligned}$$

Resulta que

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{\cdot\cdot})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{i\cdot})^2+\sum_{i=1}^{k}n_i(\overline{X}_{i\cdot}-\overline{X}_{\cdot\cdot})^2$$

ou equivalentemente

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{\cdot\cdot})^2 = \sum_{i=1}^{k}n_i(\overline{X}_{i\cdot}-\overline{X}_{\cdot\cdot})^2+\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{i\cdot})^2 \tag{11.12}$$

O membro no lado esquerdo de (11.12) é a soma dos quadrados dos desvios de todas as observações em torno da média geral, uma medida de variabilidade geral dos dados, que é chamada *soma dos quadrados total* (SQT), isto é:

$$\text{SQT} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{\cdot\cdot})^2 \tag{11.13}$$

O primeiro somatório no lado direito de (11.12) é uma medida de variação entre as médias dos tratamentos e a média geral (segmentos em vermelho na Figura 11.2); sendo assim, ele é chamado de **soma de quadrados devida ao tratamento** ou **soma de quadrados entre grupos**. Vamos denotá-la por SQG.

$$\text{SQG} = \sum_{i=1}^{k}n_i(\overline{X}_{i\cdot}-\overline{X}_{\cdot\cdot})^2 \tag{11.14}$$

No segundo somatório no lado direito de (11.12) temos uma medida de variação entre elementos do mesmo grupo, uma vez que são considerados os desvios de cada elemento e a média do seu grupo (segmentos em azul na Figura 11.2). Essa soma representa o que deixou de ser explicado pelo fator A e representa uma variabilidade dentro dos grupos. Assim, é chamada de **soma de quadrados dos erros** ou **soma de quadrado dentro dos grupos** e a denotaremos por SQE.

$$\text{SQE} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{i\cdot})^2 = \sum_{i=1}^{k}(n_i-1)S_i^2 \tag{11.15}$$

Temos, assim, a decomposição

$$\text{SQT} = \text{SQG} + \text{SQE} \tag{11.16}$$

11.1.3 Graus de liberdade

A cada soma de quadrados está associado um número de graus de liberdade, que pode ser pensado como o número de parcelas independentes no somatório.

- Soma de quadrados total

  Sabemos, por (11.7), que a soma de todos os desvios em torno da média geral é 0; sendo assim, se conhecermos $n-1$ dos n desvios $X_{ij} - \overline{X}_{..}$, o n-ésimo fica determinado, ou seja, há $n-1$ parcelas independentes no somatório. Logo, a soma de quadrados total tem $n-1$ graus de liberdade:

  $$\text{SQT} \to \text{gl} = n - 1$$

  Podemos pensar nos graus de liberdade da seguinte forma também: temos n observações para estimar a média geral; sendo assim, "sobram" $n-1$ graus de liberdade.

- Soma de quadrados devida ao tratamento

  De maneira análoga, sabemos, por (11.9), que a soma das k parcelas $n_i(\overline{X}_{i.} - \overline{X}_{..})$ é zero, e assim, há $k-1$ parcelas independentes, o que leva a $k-1$ graus de liberdade.

  $$\text{SQG} \to \text{gl} = k - 1$$

- Soma de quadrados dos erros

  Podemos ver, por (11.11) que, para cada tratamento i, a soma dos desvios das observações em torno da média do respectivo grupo é 0; sendo assim, se conhecermos $n_i - 1$ dos n_i desvios $X_{ij} - \overline{X}_{i.}$, o n_i-ésimo fica determinado. Logo, cada soma de quadrados $\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_{i.})^2$ tem $n_i - 1$ graus de liberdade e, portanto, a soma de quadrados devida aos erros tem $(n_1 - 1) + (n_2 - 1) + \cdots + (n_k - 1) = n - k$ graus de liberdade.

  $$\text{SQE} \to \text{gl} = n - k$$

  Podemos pensar nos graus de liberdade da seguinte forma também: temos n observações para estimar k médias; sendo assim, "sobram" $n-k$ graus de liberdade.

Note que a igualdade das somas de quadrados vale também para o número de graus de liberdade:

$$\text{SQT} = \text{SQG} + \text{SQE} \Rightarrow \text{gl}_{\text{SQT}} = \text{gl}_{\text{SQG}} + \text{gl}_{\text{SQE}}$$

11.1.4 Médias quadráticas

A divisão de uma soma de quadrados pelo seu número de graus de liberdade resulta em uma *média quadrática*. Sendo assim, temos

$$\text{MQG} = \frac{\text{SQG}}{k-1} \tag{11.17}$$

$$\text{MQE} = \frac{\text{SQE}}{n-k} \tag{11.18}$$

Note que a média quadrática total nada mais é que a variância S_X^2 dos dados; sendo assim, ela não recebe um outro nome especial. COm relação à MQE, temos

$$\text{MQE} = \frac{1}{n-k}\sum_{i=1}^{k}(n_i - 1)S_i^2 = S_p^2 \tag{11.19}$$

ou seja, a MQE nada mais é que a média ponderada (pelos graus de liberdade) das variâncias dos k grupos, resultado análogo ao visto no caso de duas populações com variâncias iguais..

11.1.5 Tabela da ANOVA

As informações acima costumam ser resumidas em uma tabela, chamada de tabela da ANOVA, cuja forma geral é

Tabela 11.2 – Tabela da ANOVA de um fator

Fonte de variação	SQ	GL	MQ
Fator ou Grupo	SQG	$k-1$	MQG
Erro	SQE	$n-k$	MQE
Total	SQT	$n-1$	

Na próxima seção veremos como usar essas informações para testar a hipótese de igualdade das médias.

11.1.6 Fórmulas computacionais

Assim como no cálculo da variância S^2, vamos apresentar fórmulas alternativas que, além de serem numericamente mais precisas, são mais fáceis de serem obtidas em cálculos manuais. Tais fórmulas são completamente análogas à fórmula já vista para S^2.

- SQT

$$\begin{aligned}
\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{..})^2 &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}^2-2X_{ij}\overline{X}_{..}+\overline{X}_{..}^2)\\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^2-2\overline{X}_{..}\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}+\overline{X}_{..}^2\sum_{i=1}^{k}\sum_{j=1}^{n_i}1\\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^2-2n\overline{X}_{..}^2+n\overline{X}_{..}^2\\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^2-n\overline{X}_{..}^2
\end{aligned}$$

$$\text{SQT}=\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_{..})^2=\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^2-n\overline{X}_{..}^2=\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^2-\frac{X_{..}^2}{n} \tag{11.20}$$

- SQG

$$\begin{aligned}
\sum_{i=1}^{k} n_i(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2 &= \sum_{i=1}^{k} n_i(\overline{X}_{i\cdot}^2 - 2\overline{X}_{i\cdot}\overline{X}_{\cdot\cdot} + \overline{X}_{\cdot\cdot}^2) \\
&= \sum_{i=1}^{k} n_i\overline{X}_{i\cdot}^2 - 2\overline{X}_{\cdot\cdot}\sum_{i=1}^{k} n_i\overline{X}_{i\cdot} + \overline{X}_{\cdot\cdot}^2\sum_{i=1}^{k} n_i \\
&= \sum_{i=1}^{k} \frac{X_{i\cdot}^2}{n_i} - 2\overline{X}_{\cdot\cdot}\sum_{i=1}^{k} X_{i\cdot} + n\overline{X}_{\cdot\cdot}^2 \\
&= \sum_{i=1}^{k} \frac{X_{i\cdot}^2}{n_i} - 2\overline{X}_{\cdot\cdot}n\overline{X}_{\cdot\cdot} + n\overline{X}_{\cdot\cdot}^2 \\
&= \sum_{i=1}^{k} \frac{X_{i\cdot}^2}{n_i} - n\overline{X}_{\cdot\cdot}^2
\end{aligned}$$

Logo,

$$\text{SQG} = \sum_{i=1}^{k} n_i(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2 = \sum_{i=1}^{k} \frac{X_{i\cdot}^2}{n_i} - n\overline{X}_{\cdot\cdot}^2 = \sum_{i=1}^{k} \frac{X_{i\cdot}^2}{n_i} - \frac{X_{\cdot\cdot}^2}{n} \tag{11.21}$$

- SQE

$$\text{SQE} = \text{SQT} - \text{SQG} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}^2 - \sum_{i=1}^{k} \frac{X_{i\cdot}^2}{n_i} \tag{11.22}$$

Com essas fórmulas mais simples, a tabela da ANOVA se torna

Tabela 11.3 – Tabela da ANOVA de um fator com fórmulas computacionais

Fonte de variação	SQ	GL	MQ
Fator ou Grupo	$\text{SQG} = \sum_{i=1}^{k} \frac{X_{i\cdot}^2}{n_i} - \frac{X_{\cdot\cdot}^2}{n}$	$k-1$	$\text{MQG} = \frac{\text{SQG}}{k-1}$
Erro	$\text{SQE} = \text{SQT} - \text{SQG}$	$n-k$	$\text{MQE} = \frac{\text{SQE}}{n-k}$
Total	$\text{SQT} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}^2 - \frac{X_{\cdot\cdot}^2}{n}$	$n-1$	

EXEMPLO 11.1

Considere os seguintes dados no contexto da ANOVA:

Grupo 1	33	27	27	32	27	31	23	26	34
Grupo 2	27	35	32	28	35	39	33		
Grupo 3	30	36	33	35	33	28			

Construa a tabela da ANOVA, identificando todos os tamanhos amostrais.

Solução

Organizando os cálculos em uma tabela para cálculos manuais, temos:

Obs.	Grupo 1		Grupo 2		Grupo 3		3 grupos juntos	
	x_{1j}	x_{1j}^2	x_{2j}	x_{2j}^2	x_{3j}	x_{3j}^2		
1	33	1089	27	729	30	900		
2	27	729	35	1225	36	1296		
3	27	729	32	1024	33	1089		
4	32	1024	28	784	35	1225		
5	27	729	35	1225	33	1089		
6	31	961	39	1521	28	784		
7	23	529	33	1089				
8	26	676						
9	34	1156						
Soma	260	7622	229	7597	195	6383	684	21602

$$n_1 = 9 \qquad n_2 = 7 \qquad n_3 = 6 \qquad\qquad n = 9 + 7 + 6 = 22$$

$$x_{1.} = 260 \Rightarrow x_{1.}^2 = 260^2 = 67600$$

$$x_{2.} = 229 \Rightarrow x_{2.}^2 = 229^2 = 52441$$

$$x_{3.} = 195 \Rightarrow x_{3.}^2 = 195^2 = 38025$$

$$x_{..} = 684 \Rightarrow x_{..}^2 = 684^2 = 467856$$

$$\sum_{j=1}^{n_1} x_{1j}^2 = 7622 \qquad \sum_{j=1}^{n_2} x_{2j}^2 = 7597 \qquad \sum_{j=1}^{n_2} x_{3j}^2 = 6383$$

$$\sum_{i=1}^{3}\sum_{j=1}^{n_i} x_{ij}^2 = 21602$$

$$\text{SQG} = \frac{67600}{9} + \frac{52441}{7} + \frac{38025}{6} - \frac{467856}{22} = 74,0007$$

$$\text{SQT} = 21602 - \frac{467856}{22} = 335,8182$$

$$\text{SQE} = 335,8182 - 74,0007 = 261,8175$$

A tabela da ANOVA é

Fonte de variação	SQ	GL	MQ
Fator A	$74,0007$	2	37,00035
Erro	$261,8175$	19	13,7799
Total	$335,8182$	21	

♦♦

11.2 O modelo da ANOVA de um fator

Assim como no caso do teste t para comparação de duas médias, o modelo da ANOVA exige que as populações $X_1, X_2, \cdots, X_k$ sejam normais e, além disso, as variâncias devem ser iguais. Assim, $X_i \sim N(\mu_i, \sigma^2)$, $i = 1, 2, \cdots, k$ e amostras aleatórias simples independentes de tamanhos $n_1, n_2, \cdots, n_k$ são retiradas dessas populações. Esses pressupostos podem ser resumidos através do seguinte modelo

$$X_{ij} = \mu_i + \epsilon_{ij} \qquad \epsilon_{ij} \sim N(0; \sigma^2) \text{ iid} \tag{11.23}$$

A hipótese de variâncias iguais é a hipótese de *homoscedasticidade* (mesma variação).

11.2.1 O teste da ANOVA

As hipóteses

A hipótese de interesse na análise da variância é se as médias são iguais, ou seja

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \tag{11.24}$$

com hipótese alternativa dada por

$$H_1 : \mu_i \neq \mu_j \text{ para algum } i \neq j \tag{11.25}$$

Note que aqui não temos uma hipótese alternativa simples; não temos como dizer se é um teste bilateral ou unilateral!

A estatística de teste

Pode-se mostrar que, se H_0 é verdadeira, então

$$F_0 = \frac{\text{MQG}}{\text{MQE}} \sim F_{k-1,n-k} \tag{11.26}$$

A região crítica

Para definir a região crítica, vamos calcular o valor esperado de MQG e MQE.

- MQG

  Por (11.21), temos que

$$
\begin{aligned}
E(SQG) &= \sum_{i=1}^{k} \frac{E(X_{i\cdot}^2)}{n_i} - \frac{1}{n} E(X_{\cdot\cdot}^2) \\
&= \sum_{i=1}^{k} \frac{\mathrm{Var}(X_{i\cdot}) + [E(X_{i\cdot})]^2}{n_i} - \frac{1}{n}\left\{\mathrm{Var}(X_{\cdot\cdot}) + [E(X_{\cdot\cdot})]^2\right] \\
&= \sum_{i=1}^{k} \frac{n_i\sigma^2 + n_i^2\mu_i^2}{n_i} - \frac{1}{n} n\sigma^2 - \frac{(n_1\mu_1 + n_2\mu_2 + \cdots + n_k\mu_k)^2}{n} \\
&= k\sigma^2 + \sum_{i=1}^{k} n_i\mu_i^2 - \sigma^2 - \frac{\left[\sum_{i=1}^{k} n_i\mu_i\right]^2}{n} \\
&= (k-1)\sigma^2 + \sum_{i=1}^{k} n_i\mu_i^2 - n\left[\sum_{i=1}^{k} \frac{n_i\mu_i}{n}\right]^2
\end{aligned}
$$

O somatório dentro dos colchetes é uma média ponderada das médias populacionais dos k tratamentos. Assim, vamos denotá-la por $\mu_\cdot$, isto é

$$
\mu_\cdot = \frac{1}{n} \sum_{i=1}^{k} n_i\mu_i \tag{11.27}
$$

Temos, então, que

$$
E(SQG) = (k-1)\sigma^2 + \sum_{i=1}^{k} n_i\mu_i^2 - n\mu_\cdot^2
$$

Mas

$$
\begin{aligned}
\sum_{i=1}^{k} n_i(\mu_i - \mu_\cdot)^2 &= \sum_{i=1}^{k} n_i(\mu_i^2 - 2\mu_i\mu_\cdot + \mu_\cdot^2) \\
&= \sum_{i=1}^{k} n_i\mu_i^2 - 2\mu_\cdot \sum_{i=1}^{k} n_i\mu_i + \mu_\cdot^2 \sum_{i=1}^{k} n_i \\
&= \sum_{i=1}^{k} n_i\mu_i^2 - 2n\mu_\cdot^2 + n\mu_\cdot^2 \\
&= \sum_{i=1}^{k} n_i\mu_i^2 - n\mu_\cdot^2
\end{aligned}
$$

Logo,

$$
E(SQG) = (k-1)\sigma^2 + \sum_{i=1}^{k} n_i(\mu_i - \mu_\cdot)^2
$$

e, portanto,

$$
E(MQG) = \sigma^2 + \frac{1}{k-1} \sum_{i=1}^{k} n_i(\mu_i - \mu_\cdot)^2 \tag{11.28}
$$

- MQE

  Por (11.22), temos que

$$
\begin{aligned}
\mathrm{E(SQE)} &= \sum_{i=1}^{k}\sum_{j=1}^{n_i} \mathrm{E}(X_{ij}^2) - \sum_{i=1}^{k} \frac{\mathrm{E}(X_{i\cdot}^2)}{n_i} \\
&= \sum_{i=1}^{k}\left[\sum_{j=1}^{n_i} \mathrm{Var}(X_{ij}) + [\mathrm{E}(X_{ij})]^2\right] - \sum_{i=1}^{k} \frac{\mathrm{Var}(X_{i\cdot}) + [\mathrm{E}(X_{i\cdot})]^2}{n_i} \\
&= n\sigma^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i} \mu_i^2 - \sum_{i=1}^{k} \frac{n_i\sigma^2 + (n_i\mu_i)^2}{n_i} \\
&= n\sigma^2 + \cancel{\sum_{i=1}^{k} n_i\mu_i^2} - k\sigma^2 - \cancel{\sum_{i=1}^{k} n_i\mu_i^2} \\
&= (n-k)\sigma^2
\end{aligned}
$$

  e, portanto,

$$
\mathrm{E(MQE)} = \sigma^2 \tag{11.29}
$$

- A região crítica

  De (11.29) podemos ver que a média quadrática dos erros é um estimador não viesado para a variância comum σ^2. Por outro lado, se H_0 for verdadeira, MQG também é um estimador não viesado de σ^2, mas, em geral, $\mathrm{E(MQG)} > \mathrm{E(MQE)} = \sigma^2$. Logo, sob a hipótese alternativa H_1, o valor esperado do numerador da estatística de teste (11.26) será maior que o valor esperado do denominador. Sendo assim, rejeitaremos H_0 para valores grandes da estatística de teste, ou seja, o *teste F da ANOVA é um teste unilateral à direita* cuja região crítica é

$$
F_0 = \frac{\mathrm{MQG}}{\mathrm{MQE}} > F_{k-1,n-k;\alpha} \tag{11.30}
$$

11.2.2 Estimação das médias

O estimador pontual da média μ_i é $\overline{X}_{i\cdot}$. Para construir o intervalo de confiança, usamos MQE como estimador da variância σ^2 e a distribuição amostral será a $t-$Student com $n-k$ graus de liberdade, que é o número de graus de liberdade da SQE. Assim, o intervalo de confiança de nível $1-\alpha$ para μ_i é

$$
\left[\overline{X}_{i\cdot} - t_{n-k;\alpha/2}\sqrt{\frac{\mathrm{MQE}}{n_i}}\ ;\ \overline{X}_{i\cdot} + t_{n-k;\alpha/2}\sqrt{\frac{\mathrm{MQE}}{n_i}}\right] \tag{11.31}
$$

EXEMPLO 11.2

Vamos completar o Exemplo 11.1, fazendo o teste da hipótese de igualdade das três médias. Para isso, completamos a tabela da ANOVA acrescentando uma coluna com o valor da estatística F e outra coluna com o valor P.

Fonte de variação	SQ	GL	MQ	F	Valor P
Fator A	74,0007	2	37,00035	2,685	0,093979
Erro	261,8175	19	13,7799		
Total	335,8182	21			

A estatística F foi calculada como

$$F = \frac{37,00035}{13,7799} = 2,685$$

e o valor P foi calculado com auxílio do Minitab como

$$P = \mathrm{P}(F_{2,19} > 2,685) = 0,093979$$

O valor P é razoavelmente grande; assim, não rejeitaríamos a hipótese de de igualdade das médias para níveis de significância usuais como 5%, por exemplo.

Vamos, agora, calcular os intervalos de confiança de 95% para as médias.

$t_{19;0,025} = 2,09302 \qquad 2,09032\sqrt{\mathrm{MQE}} = 2,09032\sqrt{13,7799} = 7,7595$

$$\overline{x}_{1.} = \frac{260}{9} = 28,8889 \qquad \overline{x}_{2.} = \frac{229}{7} = 32,7143 \qquad \overline{x}_{3.} = \frac{195}{6} = 32,5000$$

- Intervalo de confiança para μ_1

$$\left[28,8889 - \frac{7,7595}{\sqrt{9}}\,;\, 28,8889 + \frac{7,7595}{\sqrt{9}}\right] = [26,3024\,;\, 31,4754]$$

- Intervalo de confiança para μ_2

$$\left[32,7143 - \frac{7,7595}{\sqrt{7}}\,;\, 32,7143 + \frac{7,7595}{\sqrt{7}}\right] = [29,7815\,;\, 35,6471]$$

- Intervalo de confiança para μ_3

$$\left[32,5 - \frac{7,7595}{\sqrt{6}}\,;\, 32,5 + \frac{7,7595}{\sqrt{6}}\right] = [29,3322\,;\, 35,6678]$$

Na Figura 11.3 apresenta-se a saída do Minitab para os dados deste exemplo. Com exceção do Sumário do Modelo, todas as outras informações foram calculadas nos exemplos. O desvio padrão combinado (última linha) é simplesmente a raiz quadrada da MQE. Nas Figuras 11.4 e 11.5 são exibidos os boxplots dos dados e os intervalos de confiança para as médias, respectivamente.

ANOVA com um fator: G1; G2; G3

```
Método

Hipótese nula            Todas as médias são iguais
Hipótese alternativa     No mínimo uma média é diferente
Nível de significância   α = 0,05

Assumiu-se igualdade de variâncias para a análise


Informações dos Fatores

Fator  Níveis  Valores
Fator       3  G1; G2; G3


Análise de Variância

Fonte  GL  SQ (Aj.)  QM (Aj.)  Valor F  Valor-P
Fator   2     74,00     37,00     2,69    0,094
Erro   19    261,82     13,78
Total  21    335,82


Sumário do Modelo

      S     R2  R2(aj)  R2(pred)
3,71212  22,04%  13,83%     0,00%


Médias

Fator  N  Média  DesvPad     IC de 95%
G1     9  28,89     3,72  (26,30; 31,48)
G2     7  32,71     4,19  (29,78; 35,65)
G3     6  32,50     3,02  (29,33; 35,67)

DesvPad Combinado = 3,71212
```

Figura 11.3 – Saída do Minitab para o Exemplo 11.1

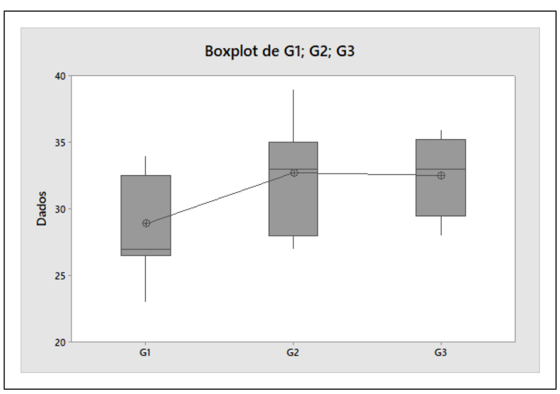


Figura 11.4 – Boxplot dos dados do Exemplo 11.1

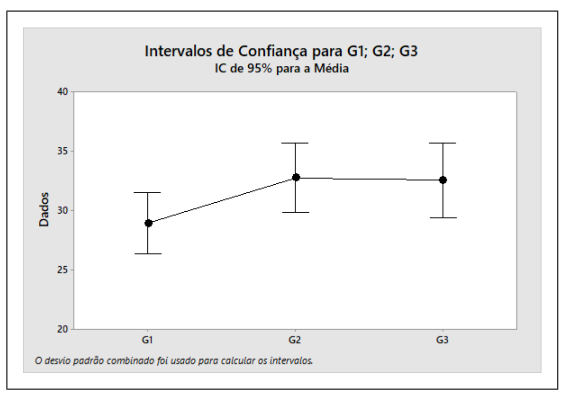


Figura 11.5 – IC para as médias do Exemplo 11.1

11.3 Verificação das hipóteses do modelo

O modelo da ANOVA se baseia em três hipóteses fundamentais:

1. Independência
2. Normalidade
3. Homogeneidade de variâncias

Gráficos de resíduos são uma importante ferramenta na análise da independência. Se as hipóteses do modelo são satisfeitas, os resíduos (valor observado - valor ajustado) não devem apresentar qualquer tipo de estrutura. No caso da ANOVA de um fator, o valor ajustado é simplesmente a média amostral do grupo.

11.3.1 Independência

Essa hipótese estabelece que deve haver independência entre as observações dentro de cada grupo e entre grupos. No planejamento do experimento é fundamental que a obtenção dos dados seja feita de forma apropriada, pois a violação da hipótese de independência é um problema sério, difícil de se corrigir. A aleatorização do experimento é um passo importante para obtenção da independência.

11.3.2 Normalidade

Os testes de normalidade já vistos devem ser aplicados a cada um dos grupos.

11.3.3 Homogeneidade de variâncias

A hipótese de homoscedasticidade pode ser verificada com alguns testes de

$$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2 \tag{11.32}$$

contra a alternativa de que nem todas as variâncias são iguais:

$$H_1 : \sigma_i \neq \sigma_j \quad \text{para algum } i \neq j \text{ , } j = 1, 2, \cdots, k \tag{11.33}$$

Veremos, aqui, dois desses testes.

Teste de Bartlett

O teste de Bartlett se baseia em uma estatística que é distribuída aproximadamente como uma qui-quadrado com $k-1$ graus de liberdade. No entanto, esse teste é bastante sensível à hipótese de normalidade. A estatística de teste é

$$X^2 = \frac{(n-k)\ln(\text{MQE}) - \sum_{i=1}^{k}(n_i-1)\ln(S_i^2)}{1 + \frac{1}{3(k-1)}\left[\sum_{i=1}^{k}\left(\frac{1}{n_i-1}\right) - \frac{1}{n-k}\right]} \tag{11.34}$$

e, sob a hipótese nula de igualdade de variâncias, $X^2 \approx \chi^2_{k-1}$. Esse é um teste unilateral superior, ou seja, rejeita-se H_0 para valores grandes de X^2, ou seja, rejeita-se a hipótese nula se

$$X^2 > \chi^2_{k-1;\alpha}$$

Teste de Levene

O teste de Levene é mais robusto contra falta de normalidade dos dados e sua estatística segue uma distribuição F sob H_0:

$$L = \frac{n-k}{k-1} \frac{\sum_{i=1}^{k} n_i (\overline{Z}_{i\cdot} - \overline{Z}_{\cdot\cdot})^2}{\sum_{i=1}^{k} \sum_{j=1}^{n_i} (Z_{ij} - \overline{Z}_{i\cdot})^2} \underbrace{=}_{\text{sob } H_0} F_{k-1,n-k} \tag{11.35}$$

em que

$$Z_{ij} = \left| X_{ij} - \overline{X}_{i\cdot} \right| \qquad \text{desvio absoluto dos } X_{ij} \text{ em relação à média do grupo} \tag{11.36}$$

$$\overline{Z}_{i\cdot} = \sum_{j=1}^{n_i} Z_{ij} \qquad \text{média dos } Z_{ij} \text{ no grupo } i \tag{11.37}$$

$$\overline{Z}_{\cdot\cdot} = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} Z_{ij} \qquad \text{média geral dos } Z_{ij} \tag{11.38}$$

Esse também é um teste unilateral superior, ou seja, a região crítica é

$$L > F_{k-1,n-k;\alpha}$$

11.4 Exercícios propostos

1. O tempo de resposta (em milissegundos) foi determinado para três diferentes tipos de circuitos usados em uma calculadora eletrônica. Os dados são os seguintes:

Tipo de circuito	Tempo de resposta				
1	19	22	20	18	25
2	20	21	33	27	40
3	16	15	18	26	17

   (a) Teste a hipótese de que os três tipos de circuito têm o mesmo tempo médio de resposta. Use $\alpha = 0,05$.

   (b) Calcule intervalos de confiança de 95% para os três tempos médios de resposta.

2. Breitling vende pulseiras para relógios masculinos em ouro, prata e titânio. Obteve-se uma amostra aleatória de cada tipo (em estilos semelhantes), e o peso de cada pulseira (em gramas) foi registrado. Os dados constam da tabela que segue.

Pulseira	Peso (g)							
Ouro	7,9	7,2	7,8	8,1	7,9	8,3	9,9	
Prata	9,5	7,0	8,7	7,6	7,5	9,3	7,3	6,9
Titânio	6,7	7,1	6,5	7,1	5,5	6,7	4,9	3,9

(a) Realize um teste de análise de variância para determinar se há alguma evidência de que os pesos médios de algum par de tipos de pulseira sejam diferentes. Inclua uma tabela ANOVA. Use $\alpha = 0,05$.

(b) Calcule o peso médio amostral para cada amostra e calcule os intervalos de confiança de 95% para cada um dos pesos populacionais. Dada sua conclusão na parte (a), quais pares de médias populacionais você acha que sejam diferentes?

3. Realizou-se um estudo para se comparar a quantidade de sal em batatas fritas. Obtiveram-se amostras aleatórias de quatro variedades e registrou-se a quantidade de sal em cada porção de 1 onça (em mg de sódio). Os dados são apresentados na tabela que segue.

Marca	Sódio (mg)					
A	338	155	239	184	185	261
B	235	238	251	229	233	232
C	164	197	135	214	148	230
D	290	343	294	373	306	357

Realize um teste de análise de variância para determinar se há alguma evidência de que a quantidade populacional média de sal por porção seja diferente para, pelo menos, duas variedades. Use $\alpha = 0,05$.

Capítulo 12

Análise de acompanhamento

12.1 Introdução

Quando o teste F acusa diferença significativa entre as médias dos k tratamentos, não há informação de qual, ou quais, são diferentes. Sendo assim, é necessária uma análise de acompanhamento (*follow up*) para identificar aonde está a diferença. Note que essa análise só faz sentido se o teste F foi significante.

Como estamos comparando várias médias, tal análise envolve múltiplas comparações de pares de médias. Uma possível solução seria analisar individualmente cada par possível de médias através de um teste t com nível de significância *individual* α. Em um teste de igualdade de várias médias, ainda queremos manter pequena a probabilidade de erro global e assim define-se a *taxa de erro global* (em inglês, family-wise error rate) como sendo a probabilidade de se cometer pelo menos um erro tipo I entre todas as comparações aos pares. Suponhamos que haja 4 grupos; então, existem $(4 \cdot 3)/2 = 6$ pares de médias a comparar. Se fizermos as 6 comparações através de testes t independentes com $\alpha = 0,05$, a probabilidade de obtermos *pelo menos um* teste significante (dentre os 6) quando H_0 é verdadeira será $1 - 0,95^6 = 0,265$, bem maior que 0,05!

Em geral, se há m pares de médias a comparar, a taxa de erro global é

$$\overline{\alpha} = 1 - (1 - \alpha)^m \tag{12.1}$$

Vamos denotar por E_i o evento "rejeitar $H_0 \,|\, H_0$ é verdadeira no teste i". Então, se o nível de significância individual é α, resulta que

$$\overline{\alpha} = \mathrm{P}(E_1 \cup E_2 \cup \cdots \cup E_m) \leq \sum_{i=1}^{m} \mathrm{P}(E_i) = m\alpha \tag{12.2}$$

Na Figura 12.1 ilustra-se a variação da taxa de erro global em função do número de testes individuais sendo feitos, para um nível de significância individual $\alpha = 0,05$. Podemos ver que a taxa de erro global cresce rapidamente à medida que aumento o número de comparações sendo feitas. Há várias propostas para tratar a comparação simultânea de várias médias, de forma a controlar a taxa de erro global, mas não há consenso sobre qual é o "melhor". Apresentaremos agora alguns desses métodos.

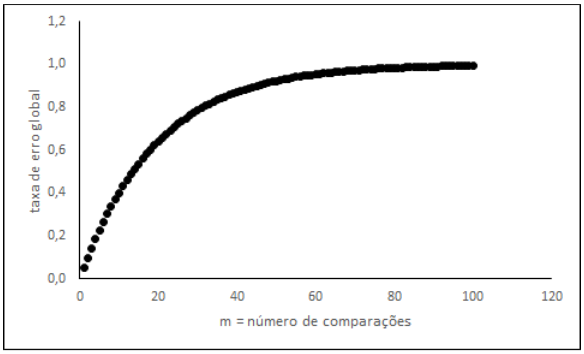


Figura 12.1 – Taxa de erro global para nível de significância individual $\alpha = 0,05$

12.2 Procedimento de comparações múltiplas de Bonferroni

Suponha que temos k grupos e, portanto, $c = \binom{k}{2} = \frac{k(k-1)}{2}$ pares de médias a comparar. Cada par de médias será comparado através de um teste t, com a variância sendo estimada pela MQE. Mas, o nível de significância de cada teste individual será o nível de significância global dividido por c. Essa é uma forma de "corrigir" a taxa de erro global, tendo em conta o resultado dado em (12.2).

O intervalo de confiança de Bonferroni de nível $1 - alpha$ para comparação das médias das populações i_1 e i_2 tem, então, limites dados por

$$\left(\overline{X}_{i_1\cdot} - \overline{X}_{i_2\cdot}\right) \pm t_{n-k;\alpha/(2c)} \sqrt{\text{MQE}\left(\frac{1}{n_{i_1}} + \frac{1}{n_{i_2}}\right)} \tag{12.3}$$

Os graus de liberdade da $t-$Student vêm da média quadrática dos erros. O nível de significância correspondente a cada intervalo individual é ajustado para o número de comparações: note que $\alpha/(2c) = (\alpha/2)/c$.

Com cada um desses intervalos testa-se a hipótese

$$H_0 : \overline{\mu}_{i_1} = \overline{\mu}_{i_2}$$

e rejeita-se H_0 se o 0 não estiver contido no intervalo de confiança.

EXEMPLO 12.1 Pulseiras de relógios (Kokoska)

Pulseiras para relógios masculinos são feitas em ouro, prata e titânio. Obteve-se uma amostra aleatória de cada tipo (em estilos semelhantes), e o peso de cada pulseira (em gramas) foi registrado. Os dados constam da tabela que segue.

Pulseira	Peso (g)							
Ouro	7,9	7,2	7,8	8,1	7,9	8,3	9,9	
Prata	9,5	7,0	8,7	7,6	7,5	9,3	7,3	6,9
Titânio	6,7	7,1	6,5	7,1	5,5	6,7	4,9	3,9

Na Figura 12.2 apresenta-se a saída do Minitab da ANOVA.

```
ANOVA com um fator: Ouro; Prata; Titânio

Método
Hipótese nula              Todas as médias são iguais
Hipótese alternativa       No mínimo uma média é diferente
Nível de significância     α = 0,05

Assumiu-se igualdade de variâncias para a análise

Informações dos Fatores
Fator  Níveis  Valores
Fator       3  Ouro; Prata; Titânio

Análise de Variância
Fonte   GL  SQ (Aj.)  QM (Aj.)  Valor F  Valor-P
Fator    2     21,20    10,601     9,97    0,001
Erro    20     21,27     1,064
Total   22     42,47

Sumário do Modelo
      S     R2  R2(aj)  R2(pred)
1,03131  49,92%  44,91%    34,04%

Médias
Fator    N  Média  DesvPad      IC de 95%
Ouro     7  8,157    0,840  (7,344; 8,970)
Prata    8  7,975    1,038  (7,214; 8,736)
Titânio  8  6,050    1,165  (5,289; 6,811)

DesvPad Combinado = 1,03131
```

Figura 12.2 – Saída do Minitab para o Exemplo 12.1

Para um nível de significância global de $\alpha = 0,05$, o nível de significância individual deverá ser $0,05/3$ e, assim, como a MQ tem 20 graus de liberdade, o valor crítico é[1]

$$t_{20;0,05/6} = 2,61277$$

e, portanto,

$$t_{20;0,016667/2} \cdot \sqrt{\text{MQE}} = 2,61277 \cdot 1,03131 = 2,694576$$

Os intervalos de confiança são:

- Ouro – Prata

$$(8,157 - 7,975) \pm 2,694576\sqrt{\frac{1}{7} + \frac{1}{8}} = 0,182 \pm 1,394575 = (-1,212575\,;\ 1,576575)$$

- Ouro – Titânio

$$(8,157 - 6,050) \pm 2,694576\sqrt{\frac{1}{7} + \frac{1}{8}} = 2,107 \pm 1,394575 = (0,712425\,;\ 3,501575)$$

- Prata – Titânio

$$(7,975 - 6,050) \pm 2,694576\sqrt{\frac{1}{8} + \frac{1}{8}} = 1,925 \pm 0,463628 = (0,577712\,;\ 3,272288)$$

Analisando os intervalos, vemos que há diferença significante entre os pesos das pulseiras de ouro e titânio e das pulseiras de prata e titânio; os pesos das pulseiras de titânio são significantemente diferentes (menores) que os pesos das pulseiras de ouro ou prata. ♦♦

[1] obtido com o Minitab

12.3 A diferença mínima significante de Fisher

A ideia central subjacente ao teste da diferença mínima significante proposto por Fisher em 1935 é calcular a menor diferença significante (DMS) como se fosse a única diferença a ser comparada – com um teste t. Cada diferença, em módulo, será declarada significante se for maior que DMS (em inglês, least significant difference – LSD). O cálculo de DMS é feito da seguinte forma:

- $n_1 = n_2 = \cdots = n_k = n^*$

$$\text{DMS} = t_{n-k;\alpha/2}\sqrt{\frac{2}{n^*}\cdot \text{MQE}} \tag{12.4}$$

- nem todos os n_i's iguais

$$\text{DMS} = t_{n-k;\alpha/2}\sqrt{\text{MQE}\left(\frac{1}{n_{i_1}}+\frac{1}{n_{i_2}}\right)} \tag{12.5}$$

em que i_1 e i_2 são as médias sendo comparadas.

Rejeita-se $H_0 : \mu_{i_1} = \mu_{i_2}$ se

$$\left|\overline{X}_{i_1\cdot} - \overline{X}_{i_2\cdot}\right| > \text{DMS}$$

Note que quando os n_i's não são todos iguais, é necessário calcular DMS para cada par de médias sendo comparadas.

EXEMPLO 12.2 Pulseiras de relógios (continuação)

Para aplicar o teste da DMS, observamos, primeiro, que

$$t_{20;0,025} = 2,08596$$

Logo,

$$t_{20;0,025}\sqrt{\text{MQE}} = 2,08596\sqrt{1,064} = 2,151676$$

- Ouro − Prata

$$(8,157 - 7,975) \pm 2,151676\sqrt{\frac{1}{7}+\frac{1}{8}} = 0,182 \pm 1,113598 = (-0,931598\,;\ 1,295598)$$

- Ouro − Titânio

$$(8,157 - 6,050) \pm 2,151676\sqrt{\frac{1}{7}+\frac{1}{8}} = 2,107 \pm 1,113598 = (0,993492\,;\ 3,220598)$$

- Prata − Titânio

$$(7,975 - 6,050) \pm 2,151676\sqrt{\frac{1}{8}+\frac{1}{8}} = 1,925 \pm 1,075838 = (0,849162\,;\ 3,000838)$$

Na Figura 12.3 temos a saída do Minitab para os intervalos de confiança baseados na DMS de Fisher; note que as diferenças foram tomadas ao contrário das nossas, daí os sinais invertidos dos limites dos IC.

```
Informações de Agrupamento Usando o Método LSD de Fisher e Confiança de 95%

Fator    N  Média  Agrupamento
Ouro     7  8,157  A
Prata    8  7,975  A
Titânio  8  6,050    B

Médias que não compartilham uma letra são significativamente diferentes.


Testes Individuais de Fisher para as Diferenças de Médias

                     Diferença      EP da                          Valor-P
Diferença de Níveis  de Médias  Diferença      IC de 95%    Valor-T  Ajustado
Prata - Ouro            -0,182      0,534  (-1,296;  0,931)   -0,34     0,736
Titânio - Ouro          -2,107      0,534  (-3,221; -0,994)   -3,95     0,001
Titânio - Prata         -1,925      0,516  (-3,001; -0,849)   -3,73     0,001

Nível de confiança simultâneo = 88,18%
```

Figura 12.3 – Saída do Minitab para o Exemplo 12.1 - Intervalos de confiança da DMS de Fisher

♦♦

12.4 A diferença honestamente significante de Tukey

A ideia principal subjacente ao teste da diferença honestamente significante (DHS) proposto por Tukey é a comparação de todas as diferenças aos pares usando a mesma distribuição amostral utilizada para a maior diferença, o que torna o teste de Tukey bastante conservador. A distribuição para a maior diferença se baseia na distribuição da amplitude studentizada descoberta por William Gosset. Essa distribuição refere-se à estatística

$$q = \frac{\max(x_1, x_2, \cdots, x_n) - \min(x_1, x_2, \cdots, x_n}{s}$$

e depende do número n de observações (ou grupos) e do número de graus de liberdade do estimador da variância comum σ^2.

O teste da DHS de Tukey rejeita $H_0 : \mu_{i_1} = \mu_{i_2}$ se

$$\left|\overline{X}_{i_1 \cdot} - \overline{X}_{i_2 \cdot}\right| > \text{DHS}$$

em que DHS é calculada da seguinte forma:

- $n_1 = n_2 = \cdots = n_k = n^*$

$$\text{DHS} = q_{k,n-k;\alpha}\sqrt{\frac{\text{MQE}}{n^*}} \tag{12.6}$$

- nem todos os n_i's iguais

$$\text{DHS} = q_{k,n-k;\alpha}\sqrt{\frac{\text{MQE}}{2}\left(\frac{1}{n_{i_1}} + \frac{1}{n_{i_2}}\right)} \tag{12.7}$$

  em que i_1 e i_2 são as médias sendo comparadas.

O teste de Tukey, ao considerar a maior diferença, preocupa-se apenas com o tamanho da diferença. Sendo assim, é um teste unilateral à direita.

Embora haja semelhança com a estatística t para duas amostras, note que as médias sendo comparadas são escolhidas a posteriori, ou seja, depois de observados os dados. Assim, a distribuição não é mais a t e, sim, a da amplitude studentizada, cuja tabela de valores críticos para $\alpha = 0,05$ e $\alpha = 0,01$ pode ser encontrada no Apêndice A.

EXEMPLO 12.3 Pulseiras de relógios - continuação

No Exemplo 12.1 temos 3 grupos: ouro com $n_O = 7$, prata com $n_P = 8$ e titânio com $n_T = 8$. Usando a função `qtukey` do R, obtemos

```
qtukey(p = 0.95, nmeans = 3, df = 20)=3.577935
```

- Ouro — Prata

$$\begin{aligned} \text{DHS} &= 3,577935\sqrt{\frac{1,064}{2}\left(\frac{1}{7}+\frac{1}{8}\right)} = 1,351 \\ \overline{x}_O - \overline{x}_P &= 8,157 - 7,975 = 0,182 < 1,3506 \end{aligned}$$

- Ouro — Titânio

$$\begin{aligned} \text{DHS} &= 3,577935\sqrt{\frac{1,064}{2}\left(\frac{1}{7}+\frac{1}{8}\right)} = 1,351 \\ \overline{x}_O - \overline{x}_T &= 8,157 - 6,050 = 2,107 > 1,3506 \end{aligned}$$

- Prata — Titânio

$$\begin{aligned} \text{DHS} &= 3,577935\sqrt{\frac{1,064}{2}\left(\frac{1}{8}+\frac{1}{8}\right)} = 1,305 \\ \overline{x}_P - \overline{x}_T &= 7,975 - 6,050 = 1,925 > 1,305 \end{aligned}$$

Na Figura 12.4 temos a saída do Minitab. Note a forma de apresentar o resultado do teste: médias que não compartilham uma letra são significantemente diferentes. Vemos, então, que titânio é diferente tanto do ouro quanto da prata. Observe, também, que embora o rótulo seja "IC de 95%", os intervalos são construídos com base no nível de confiança individual de 98,01%; 95% refere-se ao nível de confiança global.

Comparações Emparelhadas de Tukey

Informações de Agrupamento Usando Método de Tukey e Confiança de 95%

Fator	N	Média	Agrupamento
Ouro	7	8,157	A
Prata	8	7,975	A
Titânio	8	6,050	B

Médias que não compartilham uma letra são significativamente diferentes.

Testes Simultâneos de Tukey para as Diferenças de Médias

Diferença de Níveis	Diferença de Médias	EP da Diferença	IC de 95%	Valor-T	Valor-P Ajustado
Prata - Ouro	-0,182	0,534	(-1,533; 1,169)	-0,34	0,938
Titânio - Ouro	-2,107	0,534	(-3,458; -0,756)	-3,95	0,002
Titânio - Prata	-1,925	0,516	(-3,230; -0,620)	-3,73	0,004

Nível de confiança individual = 98,01%

Figura 12.4 – Saída do Minitab para o Exemplo 12.1 - Teste e Intervalos de confiança da DHS de Tukey

♦♦

12.5 Teste de Duncan

As médias dos k tratamentos são arranjadas em ordem crescente e o erro padrão de cada média é determinado como

$$S_{\overline{x}} = \sqrt{\frac{\text{MQE}}{n^*}} \tag{12.8}$$

se $n_1 = n_2 = \cdots = n_k = n_*$ e por

$$S_{\overline{x}} = \sqrt{\frac{\text{MQE}}{n_H}} \tag{12.9}$$

se nem todos os n_i's são iguais sendo

$$n_H = \frac{k}{\displaystyle\sum_{i=1}^{k} \left(\frac{1}{n_i}\right)} \tag{12.10}$$

a média harmônica dos n_i's.

As diferenças observadas entre as médias são comparadas com valores da tabela de amplitudes significantes de Duncan. Essa tabela depende de dois parâmetros: ν, o número de graus de liberdade da MQE, e p, o número de médias no intervalo de comparação. O esquema da sequência geral de comparações é o seguinte:

- k comparações da maior média com (Figura 12.5)
  - ⋆ a menor média – $p = k$
  - ⋆ a segunda menor média – $p = k - 1$
  - ⋆ $\vdots$
  - ⋆ a $(k-1)$-ésima menor média – $p = k - (k-1) = 1$
- Segunda maior média com
  - ⋆ a menor média – $p = k - 1$
  - ⋆ a segunda menor média – $p = k - 2$
  - ⋆ $\vdots$
  - ⋆ a $(k-2)$-ésima menor média – $p = k - 1 - (k-2) = 1$

O processo continua até que os $\dfrac{k(k-1)}{2}$ pares de médias tenham sido comparados.

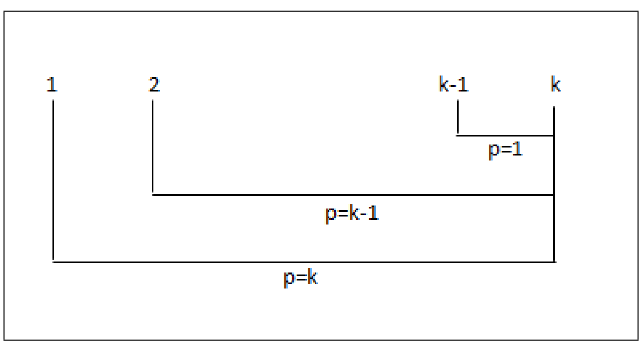


Figura 12.5 – Esquema de comparação da maior média com as demais – Teste de Duncan

Os valores críticos para comparação das diferenças de médias são definidos por

$$R_p = r_{p,\nu;\alpha} \cdot S_{\overline{x}} \tag{12.11}$$

com $r_{p,\nu;\alpha}$ dado nas Tabelas X e XI do Apêndice A.

Por ser um teste bem trabalhoso, o uso de software é absolutamente necessário aqui. O teste de Duncan não está implementado no Minitab.

EXEMPLO 12.4 Teste de Duncan

Considere as informações de uma análise de variância dadas na Figura 12.6. O teste F é significante ao nível $\alpha = 0,01$. Vamos aplicar o teste de Duncan a esses dados. A média harmônica dos tamanhos amostrais é

$$n_H = \frac{4}{\frac{1}{4} + \frac{1}{8} + \frac{1}{6} + \frac{1}{6}} = 5,6471$$

ANOVA com um fator: A; B; C; D

```
Método

Hipótese nula             Todas as médias são iguais
Hipótese alternativa      No mínimo uma média é diferente
Nível de significância    α = 0,01

Assumiu-se igualdade de variâncias para a análise


Informações dos Fatores

Fator  Níveis  Valores
Fator       4  A; B; C; D


Análise de Variância

Fonte  GL  SQ (Aj.)  QM (Aj.)  Valor F  Valor-P
Fator   3     49,93    16,643    13,05    0,000
Erro   20     25,50     1,275
Total  23     75,43


Sumário do Modelo

      S      R2  R2(aj)  R2(pred)
1,12923  66,19%  61,12%    50,89%


Médias

Fator  N   Média  DesvPad       IC de 99%
A      4  55,150    1,261  (53,543; 56,757)
B      8  56,900    1,175  (55,764; 58,036)
C      6  53,100    1,161  (51,788; 54,412)
D      6  55,533    0,931  (54,222; 56,845)

DesvPad Combinado = 1,12923
```

Figura 12.6 – Esquema de comparação da maior média com as demais – Teste de Duncan

A estimativa do erro padrão da média é

$$\sqrt{s_{\overline{x}}} = \sqrt{\frac{1,275}{5,6471}} = 0,4752$$

Da tabela das diferenças significantes de Duncan com $\alpha = 0,01$ e $\nu = 24 - 4 = 20$ obtemos

- $p = 2$ $\quad r_2 = 4,024 \Rightarrow R_2 = 0,4752 \times 4,024 = 1,9122$
- $p = 3$ $\quad r_3 = 4,197 \Rightarrow R_3 = 0,4752 \times 4,197 = 1,944$
- $p = 4$ $\quad r_4 = 4,312 \Rightarrow R_4 = 0,4752 \times 4,312 = 2,0491$

As médias ordenadas são

C	A	D	B
53,100	55,150	55,533	56,900

Comparações

- $B - C$ $\quad 56,900 - 53,100 = 3,800 > R_4$
- $B - A$ $\quad 56,900 - 55,150 = 1,750 < R_3$
- $B - D$ $\quad 56,900 - 55,533 = 1,367 < R_2$
- $D - C$ $\quad 55,533 - 53,100 = 2,433 > R_3$
- $D - A$ $\quad 55,533 - 55,150 = 0,383 < R_2$
- $A - C$ $\quad 55,150 - 53,100 = 2,050 > R_2$

Na Figura 12.7 ilustram-se essas comparações, com as médias "iguais" unidas por segmentos.

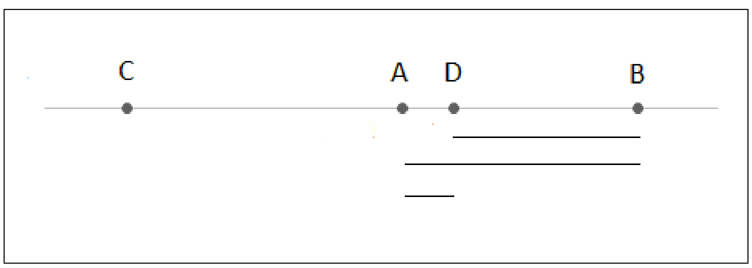


Figura 12.7 – Comparação das médias para o Exemplo 12.4

Parte IV

Análise de dados categóricos

Capítulo 13

Análise de dados categóricos

13.1 Introdução

Vamos, agora, estudar algumas técnicas de inferências para dados qualitativos. Assim como no estudo anterior de populações normais, vamos considerar os casos de uma população, em que cada indivíduo é classificado segundo alguma variável qualitativa, e duas populações, das quais amostras retiradas podem ser independentes ou dependentes (dados emparelhados) e para cada elemento da amostra são observadas duas variáveis qualitativas. Os objetivos da inferência variam para as diferentes situações. A título de ilustração, vamos considerar os seguintes exemplos.

EXEMPLO 13.1 Meio de transporte

Em uma pesquisa com alunos da UFF em Niterói, perguntou-se a cada um deles o meio de transporte que utilizavam no trajeto de casa para a universidade. Cada aluno escolhia entre uma das seguintes possibilidades: só ônibus, só barca, ônibus e barca, carro, caminhada/bicicleta. As proporções amostrais $\widehat{P}_1$, $\widehat{P}_2$, $\widehat{P}_3$, $\widehat{P}_4$ e $\widehat{P}_5$ são estimadores das verdadeiras proporções populacionais p_1, p_2, p_3, p_4 e p_5 de usuários de cada meio de transporte e um interesse poderia ser testar se $p_1 = 0,60$, $p_2 = 0,10$, $p_3 = 0,20$, $p_4 = 0,05$ e $p_5 = 0,05$. Esse é um *teste de aderência*, ou seja, estamos testando se nossos dados são compatíveis com (aderem a) determinada distribuição.

EXEMPLO 13.2 Meio de transporte e gênero

Ainda no estudo sobre meio de transporte, um interesse poderia ser estudar se há diferença entre homens e mulheres no meio de transporte utilizado. Para isso, amostras *independentes* de homens e mulheres seriam retiradas da população de todos os estudantes da UFF em Niterói e para os indivíduos de cada amostra seria registrada a variável meio de transporte. Note que a segunda variável – gênero – definiu as populações. O interesse aqui seria testar se homens e mulheres usam igualmente os diferentes meios de transporte, ou seja, queremos testar $p_{1H} = p_{1M}$, $p_{2H} = p_{2M}$, $p_{3H} = p_{3M}$, $p_{4H} = p_{4M}$ e $p_{5H} = p_{5M}$, em que p_{iH} e p_{iM} representam as proporções de homens e mulheres que utilizam o meio de transporte i. Este é um exemplo de um *teste de homogeneidade*, ou seja, estamos testando se as populações de homens e mulheres são homogêneas em relação à variável meio de transporte. Note que em cada população a variável de interesse – meio de transporte – segue uma distribuição multinomial. Sendo assim, nosso interesse é testar se as duas distribuições multinomiais são iguais (ou homogêneas).

EXEMPLO 13.3 Meio de transporte e uso do bandejão

Continuando com o estudo sobre alunos da UFF em Niterói, outro interesse poderia ser a relação entre meio de transporte utilizado e uso do bandejão da UFF. Assim, para cada aluno da amostra seriam registradas duas variáveis: meio de transporte utilizado e uso do bandejão (Sim ou Não). Agora temos dados emparelhados (amostras dependentes) e nosso interesse é um *teste de independência* entre as variáveis.

Os três testes citados acima se basearão numa estatística que compara as frequências *observadas* O com as frequências *esperadas* E sob a suposição de veracidade da hipótese nula. Grandes diferenças entre tais frequências são indicativo de que a hipótese nula não é verdadeira. Mas para termos uma medida de distância, consideraremos as diferenças ao quadrado, isto é, $(O-E)^2$. Além disso, iremos considerar distâncias *relativas*: uma distância $(O-E)^2 = 50$ será mais relevante para um valor esperado de 30 do que para um valor esperado de 100. Dessa forma, nossa estatística de teste terá a forma final dada por

$$\sum_i \frac{(O_i - E_i)^2}{E_i} \tag{13.1}$$

e essa estatística terá (aproximadamente) uma distribuição qui-quadrado com número de graus de liberdade que varia de acordo com o teste considerado. Note que a base de comparação é o valor esperado sob H_0, um valor fixo, bem determinado, que não depende da amostra sorteada.

Vamos, agora, detalhar cada um dos testes.

13.2 Dados univariados: Teste de aderência

Consideremos uma população descrita por uma variável categórica X que assume k valores (no Exemplo 13.1, X é o meio de transporte com $k = 5$ categorias). Sejam $p_1, p_2, \cdots, p_k$ as proporções populacionais das categorias $1, 2, \cdots, k$, respectivamente. Queremos testar

$$H_0 : p_1 = p_{10}\,,\; p_2 = p_{20}\,,\; \cdots\,,\; p_k = p_{k0} \tag{13.2}$$

$$H_1 : p_i \neq p_{i0} \text{ para pelo menos um } i \tag{13.3}$$

em que $p_{i0}, i = 1, 2, \cdots, k$ são as proporções hipotéticas que devem satisfazer $\sum_{i=1}^{k} p_{i0} = 1$.

Suponha que uma amostra de tamanho n seja selecionada de tal população. Seja N_i o número de observações da amostra pertencentes à categoria i, $i = 1, 2, \cdots, k$ (esses números mudam ao longo de todas as possíveis amostras de tamanho n). Se H_0 é verdadeira, o número esperado de observações na categoria i é

$$e_i = np_{i0} \tag{13.4}$$

Na Tabela 13.1 temos o resumo dessa situação.

Tabela 13.1 – Teste de aderência - esquema dos dados

Categoria	1	2	$\cdots$	k	Soma
Proporção populacional	p_1	p_2	$\cdots$	p_k	1
Proporção populacional sob H_0	p_{10}	p_{20}	$\cdots$	p_{k0}	1
Número observado	N_1	N_2	$\cdots$	N_k	n
Número esperado sob H_0	np_{10}	np_{20}	$\cdots$	np_{k0}	n

A estatística de teste é

$$X^2 = \sum_{i=1}^{k} \frac{(N_i - np_{i0})^2}{np_{i0}} \underbrace{\approx}_{\text{sob } H_0} \chi^2_{k-1} \tag{13.5}$$

Tal aproximação é boa se $e_i = np_{i0} \geq 5$ para todo $i = 1, 2, \ldots, k$. Valores grandes indicam grandes afastamentos entre os valores observados e esperados; assim, a região crítica para um nível de significância α é

$$X^2 > \chi^2_{k-1;\alpha} \tag{13.6}$$

É interessante observar as seguintes propriedades sobre os valores observados e esperados.

- $\sum_{i=1}^{k} e_i = \sum_{i=1}^{k} N_i = n$ (soma dos valores esperados)

  De fato: $\sum_{i=1}^{k} e_i = \sum_{i=1}^{k} np_{i0} = n \sum_{i=1}^{k} p_{i0} = n \cdot 1 = n$

- $\sum_{i=1}^{k} (N_i - e_i) = 0$ (desvios em relação aos esperados)

  De fato: $\sum_{i=1}^{k} (N_i - e_i) = \sum_{i=1}^{k} N_i - \sum_{i=1}^{k} e_i = n - n = 0$

EXEMPLO 13.4 Meio de transporte – continuação

Uma amostra de 200 estudantes da UFF de Niterói mostrou que 65 usam apenas ônibus como meio de transporte, 7 usam apenas barca, 20 usam ônibus e barca, 5 usam carro e 2 caminham ou vão de bicicleta até a universidade. Vamos testar

$$H_0 : p_1 = 0,60\,,\ p_2 = 0,10\,, p_3 = 0,20\,,\ p_4 = p_5 = 0,05$$

Na tabela a seguir ilustram-se os cálculos necessários.

Categoria i	Valor observado O_i	Valor esperado E_i	Parcela de X^2
1=Só ônibus	124	$200 \cdot 0,6 = 120$	$\frac{(124-120)^2}{120}$
2=Só barca	22	$200 \cdot 0,1 = 20$	$\frac{(22-20)^2}{20}$
3=Ônibus e barca	35	$200 \cdot 0,2 = 40$	$\frac{(35-40)^2}{40}$
4=Carro	9	$200 \cdot 0,05 = 10$	$\frac{(9-10)^2}{10}$
5=Caminhada/bicicleta	10	$200 \cdot 0,05 = 10$	$\frac{(10-10)^2}{10}$

Todos os valores esperados são maiores que 5 e, assim, podemos usar a aproximação qui-quadrado com 4 graus de liberdade. A região crítica é $X^2 > 9,4877$ para um nível de significância $\alpha = 0,05$. O valor observado da estatística de teste é

$$x_0^2 = \frac{16}{120} + \frac{4}{20} + \frac{25}{40} + \frac{1}{10} + \frac{0}{10} = \frac{16 + 24 + 75 + 12 + 0}{120} = \frac{129}{120} = 1,075$$

Como $1,075 < 9,4877$, não rejeitamos H_0, ou seja, não há evidências de que as proporções populacionais sejam diferentes das hipotéticas.

♦♦

EXEMPLO 13.5 Duas categorias

No estudo da inferência para uma população $X \sim Bern(p_1)$, vimos que a estatística de teste para $H_0 : p_1 = p_{10}$ é

$$Z = \frac{\widehat{P} - p_0}{\sqrt{p_{10}(1-p_{10})}} \underbrace{\approx}_{\text{sob } H_0} N(0;1)$$

Como são apenas duas categorias, basta trabalhar com uma delas e a outra estará determinada. Se denotarmos por N_i o número de observações na categoria i ($i = 1$ - sucessos; $i = 2$ - fracassos) e por p_{i0} a proporção de observações na categoria i para sermos coerentes com a notação anterior em que $k \geq 2$, a aproximação normal para a binomial nos dá que

$$Z = \frac{N_1 - np_{10}}{\sqrt{np_{10}(1-p_{10})}} \underbrace{\approx}_{\text{sob } H_0} N(0;1)$$

Logo,

$$Q = Z^2 = \frac{(N_1 - np_{10})^2}{np_{10}(1-p_{10})} \underbrace{\approx}_{\text{sob } H_0} X_1^2$$

Notando que $(1 - p_{10}) + p_{10} = 1$, podemos reescrever Q como

$$\begin{aligned}
Q = Z^2 &= \frac{(N_1 - np_{10})^2}{np_{10}(1-p_{10})}[(1-p_{10}) + p_{10}] \\
&= \frac{(N_1 - np_{10})^2}{np_{10}} + \frac{(N_1 - np_{10})^2}{n(1-p_{10})} \\
&= \frac{(N_1 - np_{10})^2}{np_{10}} + \frac{[(n - N_2) - n(1-p_{20}]^2}{np_{20})} \\
&= \frac{(N_1 - np_{10})^2}{np_{10}} + \frac{[-N_2 + np_{20}]^2}{np_{20})} \\
&= \frac{(N_1 - np_{10})^2}{np_{10}} + \frac{[N_2 - np_{20}]^2}{np_{20})}
\end{aligned}$$

Vemos, assim, que há uma equivalência entre as estatísticas Z e χ^2 quando $k = 2$.

♦♦

13.3 Dados bivariados

Consideremos um exemplo em que as variáveis categóricas $X =$ "opinião sobre a legalização do uso medicinal da maconha" com 3 níveis (a favor, indiferente e contra) e $Y =$ "faixa etária " com 4 níveis (16 a 18, 19 a 25, 26 a 40 e mais de 40) serão estudadas em uma pesquisa por amostragem entre os alunos de uma grande universidade. O que veremos agora é que a forma de se coletarem os dados é fundamental para se definir o tipo de análise pertinente.

Suponhamos, inicialmente, que sejam selecionadas amostras independentes de tamanhos $n_1, \cdots, n_4$ das quatro faixas etárias e cada elemento de cada uma das 4 amostras indique sua opinião sobre a legalização do uso medicinal da maconha. O tipo de análise que podemos fazer aqui consiste em ver se as proporções de pessoas em cada categoria são as mesmas para as quatro faixas etárias. Ou seja, estamos comparando quatro distribuições multinomiais e isso leva ao *teste de homogeneidade*.

Suponhamos, agora, que uma amostra de n alunos seja retirada e cada aluno seja classificado de acordo com sua faixa etária ***e*** sua opinião sobre o uso medicinal da maconha. Temos, agora, dados emparelhados, ou seja, as amostras são dependentes e o objetivo aqui é ver se há dependência entre as duas variáveis. Isso nos leva ao *teste de independência*.

Em ambos os casos, podemos resumir as informações através de uma tabela de contingência, ou tabela de dupla entrada com as J categorias de uma variável sendo exibidas nas colunas e as I categorias da outra variável nas linhas. Na tabela a seguir ilustra-se a notação a ser utilizada nos testes. O ponto no subscrito indica soma dos valores ao longo da respectiva dimensão.

Tabela 13.2 – Tabela de dados bivariados

		Variável coluna				
		1	2	$\cdots$	J	Total de linha
Variável linha	1	n_{11}	n_{12}	$\cdots$	n_{1J}	$n_{1\cdot}$
	2	n_{21}	n_{22}	$\cdots$	n_{2J}	$n_{2\cdot}$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	I	n_{I1}	n_{I2}	$\cdots$	n_{IJ}	$n_{I\cdot}$
Total de coluna		$n_{\cdot 1}$	$n_{\cdot 2}$	$\cdots$	$n_{\cdot J}$	$n = n_{\cdot\cdot}$

$$n_{ij} = \text{frequência observada na cela } (i, j)$$

$$n_{i\cdot} = n_{i1} + n_{i2} + \cdots + n_{iJ} = \sum_{j=1}^{J} n_{ij} \quad \text{(total da linha } i\text{)}$$

$$n_{\cdot j} = n_{1j} + n_{2j} + \cdots + n_{Ij} = \sum_{i=1}^{I} n_{ij} \quad \text{(total da coluna } j\text{)}$$

$$n_{\cdot\cdot} = n = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \quad \text{(total de observações na amostra)}$$

13.3.1 Amostras independentes: Teste de homogeneidade

Como visto, um dos contextos que origina dados bivariados é quando amostras independentes são retiradas de várias populações e cada indivíduo ou objeto é classificado de acordo com uma variável qualitativa. Neste caso, uma das variáveis é a que identifica a população e a outra é a que classifica os sujeitos. Em tal contexto, nosso interesse é testar se as proporções em cada categoria são as mesmas em todas as populações. No Exemplo 13.2, as populações são formadas pelos homens e pelas mulheres e a variável de classificação é o meio de transporte. O interesse é testar se as proporções de homens e mulheres que usam cada tipo de transporte são iguais. Na primeira situação do exemplo no início da seção, idade é a variável que identifica a população e a opinião sobre o uso medicinal da maconha é a variável de interesse, que classifica os sujeitos. O objetivo é testar se a proporção de pessoas com cada opinião é a mesma nas quatro faixas etárias.

Na Tabela 13.2, suponhamos que haja J populações com a variável de classificação tendo I categorias. Nosso interesse é testar

$$H_0 : p_{i1} = p_{i2} = \cdots = p_{iJ} \quad \forall i \tag{13.7}$$

Sob H_0, temos, então, que

$$\frac{n_{i1}}{n_{\cdot 1}} = \frac{n_{i2}}{n_{\cdot 2}} = \cdots = \frac{n_{iJ}}{n_{\cdot J}} = \frac{n_{i1} + n_{i2} + \cdots + n_{iJ}}{n_{\cdot 1} + n_{\cdot 2} + \cdots + n_{\cdot J}} = \frac{n_{i\cdot}}{n} = p_{i\cdot}$$

em que $p_{i\cdot}$ é a proporção de elementos na amostra que pertencem à categoria i.

Assim, sob H_0, a proporção na categoria i em cada população deve ser igual à proporção da categoria i na amostra toda. Como há $n_{\cdot j}$ elementos na população j, esperamos, então, que uma proporção $p_{i\cdot}$ desses elementos esteja na categoria i, ou seja,

$$\frac{e_{ij}}{n_{\cdot j}} = p_{i\cdot} = \frac{n_{i\cdot}}{n}$$

Isso nos dá que

$$e_{ij} = \frac{n_{i\cdot} n_{\cdot j}}{n} \tag{13.8}$$

Como antes, a estatística de teste é uma medida de distância relativa entre frequências observadas e frequências esperadas:

$$X^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \tag{13.9}$$

e se $e_{ij} \geq 5 \quad \forall\, i, j$

$$X^2 \approx \chi^2_{(I-1)(J-1)} \tag{13.10}$$

Valores grandes da estatística levam à rejeição da hipótese nula de igualdade das proporções em cada categoria ao longo das populações.

As mesmas propriedades sobre os valores observados e esperados continuam valendo.

- $\sum_{i=1}^{I} \sum_{j=1}^{J} e_{ij} = n$

  De fato:

$$\sum_{i=1}^{I} \sum_{j=1}^{J} e_{ij} = \frac{1}{n} \sum_{i=1}^{I} n_{i\cdot} \sum_{j=1}^{J} n_{\cdot j} = \frac{1}{n} n \cdot n = n$$

- $\sum_{i=1}^{I} \sum_{j=1}^{J} (n_{ij} - e_{ij}) = 0$

  De fato:

$$\sum_{i=1}^{I} \sum_{j=1}^{J} (n_{ij} - e_{ij}) = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} - \sum_{i=1}^{I} \sum_{j=1}^{J} e_{ij} = n - n = 0$$

EXEMPLO 13.6 Meio de transporte e gênero – continuação

Com relação ao Exemplo 13.2, suponhamos que amostras independentes de 300 homens e de 250 mulheres tenham revelado a seguinte distribuição:

Frequências observadas			
Meio de transporte	Homens	Mulheres	Total
Ônibus	77	116	193
Barca	48	19	67
Ônibus e barca	47	54	101
Carro	96	28	124
Caminhada/bicicleta	32	33	65
Total	300	250	550

Vamos realizar o teste de homogeneidade, ou seja, vamos testar, ao nível de significância $\alpha = 0,05$, se as proporções em cada categoria de meio de transporte são iguais entre homens e mulheres. Para isso, precisamos calcular as frequências esperadas em cada cela. Para a cela (1,1) (Homens que vão de ônibus) temos

$$e_{11} = \frac{n_{1\cdot}n_{\cdot 1}}{n} = \frac{193 \cdot 300}{550} = 105,27273$$

De forma análoga obtemos as frequências esperadas para as outras celas, que estão na tabela a seguir:

Frequências esperadas			
Meio de transporte	Homens	Mulheres	Total
Ônibus	105,27273	87,72727	193
Barca	36,54545	30,45455	67
Ônibus e barca	55,09091	45,90909	101
Carro	67,63636	56,36364	124
Caminhada/bicicleta	35,45455	29,54545	65
Total	300	250	550

Como todas as frequências esperadas são maiores que 5, podemos usar a aproximação qui-quadrado, ou seja, $X^2 \approx \chi^2_4$. Note que o número de graus de liberdade é $(5-1)(2-1) = 4$ e $\chi^2_{4;0,05} = 9,4877$. A parcela da cela (1,1) na estatística qui-quadrado é

$$x_{11} = \frac{(77 - 105,27273)^2}{105,27273} = 7,59311$$

Na tabela a seguir, apresentamos as contribuições de cada cela:

Contribuição para o X^2		
Meio de transporte	Homens	Mulheres
Ônibus	7,59311	9,11173
Barca	3,59023	4,30828
Ônibus e barca	1,18827	1,42592
Carro	11,89443	14,27331
Caminhada/bicicleta	0,3366	0,40392

Somando todas essas celas, obtemos $X^2 = 54,1258 > 9,4877$. Dessa forma, rejeitamos a hipótese nula, ou seja, as proporções em cada categoria de meio de transporte não são iguais entre homens e mulheres. Analisando a tabela das contribuições, podemos ver que as maiores diferenças estão nas categorias de ônibus e carro. Há mais mulheres utilizando ônibus e mais homens utilizando carros do que se esperaria.

Na Figura 13.1 temos a saída do programa Minitab para essa análise.

♦♦

Estatísticas Tabuladas: MTransp; Gênero

```
Linhas: MTransp   Colunas: Gênero

                       Homens  Mulheres  Todos

Ônibus                     77       116    193
                       105,27     87,73
                        7,593     9,112

Barca                      48        19     67
                        36,55     30,45
                        3,590     4,308

Ônibus&Barca               47        54    101
                        55,09     45,91
                        1,188     1,426

Carro                      96        28    124
                        67,64     56,36
                       11,894    14,273

Caminhada/bicicleta        32        33     65
                        35,45     29,55
                        0,337     0,404

Todos                     300       250    550

Conteúdo da Célula:      Contagem
                         Contagem esperada
                         Contribuição para Qui-Quadrado


Qui-Quadrado de Pearson = 54,126; GL = 4; Valor-P = 0,000
```

Figura 13.1 – saída do Minitab para o Exemplo 13.6

13.3.2 Amostras dependentes: Teste de independência

Consideremos, agora, o caso de dados emparelhados, ou seja, cada indivíduo da amostra é classificado segundo duas variáveis qualitativas X e Y. Embora a representação tabular dos dados amostrais seja a mesma apresentada na Tabela 13.2, a interpretação e uso dos dados é completamente diferente. Em cada cela, a contagem n_{ij} nos dá o número de sujeitos para os quais $X = i$ e $Y = j$; nosso interesse é determinar se as variáveis X e Y são independentes. Sabemos que, se X e Y são independentes, então

$$\mathrm{P}(X = i, Y = j) = \mathrm{P}(X = i)\,\mathrm{P}(Y = j)$$

Então, sob a veracidade de H_0, temos que ter

$$\frac{e_{ij}}{n} = \frac{n_{i\cdot}}{n} \cdot \frac{n_{\cdot j}}{n}$$

ou equivalentemente

$$e_{ij} = \frac{n_{i\cdot} n_{\cdot j}}{n} \tag{13.11}$$

Se todas as frequências esperadas forem de pelo menos 5, pode-se mostrar que

$$X^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \approx \chi^2_{(I-1)(J-1)} \tag{13.12}$$

EXEMPLO 13.7 Meio de transporte e uso do bandejão – continuação

A título de ilustração da diferença de interpretação da estatística qui-quadrado no contexto de dados emparelhados, vamos considerar os mesmos dados do exemplo anterior, mas agora representando uma amostra de 550 alunos, dos quais 300 usam o bandejão e 250 não usam, de acordo com a seguinte distribuição:

Meio de transporte	Usa bandejão	Não usa bandejão	Total
Ônibus	77	116	193
Barca	48	19	67
Ônibus e barca	47	54	101
Carro	96	28	124
Caminhada/bicicleta	32	33	65
Total	300	250	550

O cálculo das frequências esperadas e da estatística de teste é o mesmo e concluímos, ao nível de significância $\alpha = 0,05$, que as variáveis "meio de transporte" e "uso do bandejão" não são independentes.

13.4 Exercícios propostos

1. Em recente estudo, obteve-se uma amostra aleatória de proprietários de pequenas empresas, e pediu-se a cada um que apontasse o maior problema enfrentado por sua empresa. Os resultados são apresentados na seguinte tabela de frequências de uma entrada.

Problema	**Frequência**
Custo do seguro saúde	430
Seguro de responsabilidade	145
Indenização ao trabalhador	135
Custo de combustíveis	90

   Em um relatório econômico, as verdadeiras proporções para cada categoria foram dadas como 0,50, 0,20, 0,20 e 0,10. Esses dados fornecem alguma evidência que contradiga o relatório econômico? Ache limites para o valor P associado a esse teste.

2. Os escritórios de admissão dos campi da Universidade da Califórnia mantêm registros históricos cuidadosos dos candidatos. Em 2008, as proporções de estudantes que se candidataram para as faculdades do sistema universitário por localização na Califórnia foram as seguintes: Los Angeles (LA), 29,2%; San Francisco (SF), 26,1%; Orange County (OC), 9,9%; Riverside/San Bernardino (RS), 7,7%; todas as outras (O), 27,1%.1 Suponha que se tenha obtido uma amostra aleatória de candidatos em 2009 para a qual foram obtidas as seguintes frequências para as localizações.

Localização	LA	SF	OC	RS	O
Frequência	125	96	45	44	128

   Há alguma evidência de mudança na proporção de candidatos por localização na Califórnia? Use $\alpha = 0,005$.

3. Embora a harmonização de alimento e vinho seja subjetiva e uma ciência não exata, tradicionalmente diz-se que vinho tinto combina com carne vermelha, e vinho branco combina com peixe e aves. Obteve-se uma amostra aleatória de jantares em restaurantes quatro estrelas, e cada jantar foi classificado de acordo com a comida e o vinho pedido. Eis a tabela de frequências de dupla entrada resultante.

		Vinho	
		Tinto	Branco
Comida	Carne vermelha	86	46
	Peixe ou ave	50	64

Há alguma evidência de que a comida e o vinho sejam dependentes? Teste a hipótese relevante com $\alpha = 0,005$. Esses dados sugerem que os jantares ainda estejam seguindo as harmonizações tradicionais de comida e vinho?

4. Obtiveram-se amostras aleatórias de apostadores em quatro cassinos de Las Vegas, e perguntou-se a cada apostador qual jogo jogava mais. Os resultados são apresentados na tabela de frequências de dupla entrada que segue.

		Jogo			
		Blackjack	Pôquer	Roleta	Caça-níquel
Cassino	Bellagio	22	20	38	66
	Caesar's	30	38	22	68
	Golden Nugget	28	25	21	81
	Harrah's	38	25	29	84

Realize um teste para homogeneidade de populações. Há alguma evidência que sugira que a verdadeira proporção de apostadores em cada jogo não seja a mesma para todos os cassinos? Use $\alpha = 0,05$.

5. Foram obtidas amostras aleatórias de clientes de duas lojas diferentes de material para escritório, e perguntou-se a cada cliente qual tipo de recurso para escrita preferiam. Os resultados estão resumidos na seguinte tabela de frequências de dupla entrada.

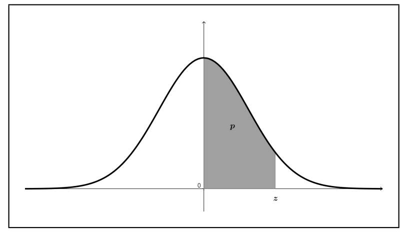
		Recurso para escrita		
		Lápis	Lapiseira	Caneta
Loja	Casa Cruz	183	164	480
	Kalunga	130	202	420

Realize um teste para homogeneidade de populações. Use $\alpha = 0,01$. Estabeleça sua conclusão, justifique sua resposta e ache limites para o valor P associado a esse teste.

Apêndice A

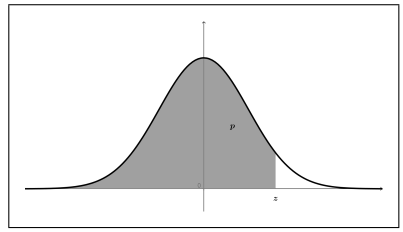
Tabelas

Tabela I
Distribuição normal padrão
$p = \mathrm{P}(0 \leq Z \leq z)$



Casa inteira e 1a.decimal	2a. casa decimal									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4,0	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Tabela II
Distribuição acumulada da normal padrão
$p = \mathrm{P}(Z \leq z)$

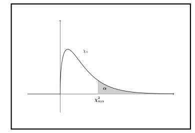


Casa inteira	2a. casa decimal									
e 1a.decimal	0	1	2	3	4	5	6	7	8	9
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,7	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
4,0	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

Tabela III

Valores críticos $\chi^2_{n,\alpha}$ da qui-quadrado

$P(\chi^2_n > \chi^2_{n,\alpha}) = \alpha$



gl	Probabilidade α na cauda superior															
n	0,999	0,995	0,990	0,980	0,975	0,950	0,900	0,800	0,200	0,100	0,050	0,025	0,020	0,010	0,005	0,001
1	0,000	0,000	0,000	0,001	0,001	0,004	0,016	0,064	1,642	2,706	3,841	5,024	5,412	6,635	7,879	10,828
2	0,002	0,010	0,020	0,040	0,051	0,103	0,211	0,446	3,219	4,605	5,991	7,378	7,824	9,210	10,597	13,816
3	0,024	0,072	0,115	0,185	0,216	0,352	0,584	1,005	4,642	6,251	7,815	9,348	9,837	11,345	12,838	16,266
4	0,091	0,207	0,297	0,429	0,484	0,711	1,064	1,649	5,989	7,779	9,488	11,143	11,668	13,277	14,860	18,467
5	0,210	0,412	0,554	0,752	0,831	1,145	1,610	2,343	7,289	9,236	11,070	12,833	13,388	15,086	16,750	20,515
6	0,381	0,676	0,872	1,134	1,237	1,635	2,204	3,070	8,558	10,645	12,592	14,449	15,033	16,812	18,548	22,458
7	0,598	0,989	1,239	1,564	1,690	2,167	2,833	3,822	9,803	12,017	14,067	16,013	16,622	18,475	20,278	24,322
8	0,857	1,344	1,646	2,032	2,180	2,733	3,490	4,594	11,030	13,362	15,507	17,535	18,168	20,090	21,955	26,124
9	1,152	1,735	2,088	2,532	2,700	3,325	4,168	5,380	12,242	14,684	16,919	19,023	19,679	21,666	23,589	27,877
10	1,479	2,156	2,558	3,059	3,247	3,940	4,865	6,179	13,442	15,987	18,307	20,483	21,161	23,209	25,188	29,588
11	1,834	2,603	3,053	3,609	3,816	4,575	5,578	6,989	14,631	17,275	19,675	21,920	22,618	24,725	26,757	31,264
12	2,214	3,074	3,571	4,178	4,404	5,226	6,304	7,807	15,812	18,549	21,026	23,337	24,054	26,217	28,300	32,909
13	2,617	3,565	4,107	4,765	5,009	5,892	7,042	8,634	16,985	19,812	22,362	24,736	25,472	27,688	29,819	34,528
14	3,041	4,075	4,660	5,368	5,629	6,571	7,790	9,467	18,151	21,064	23,685	26,119	26,873	29,141	31,319	36,123
15	3,483	4,601	5,229	5,985	6,262	7,261	8,547	10,307	19,311	22,307	24,996	27,488	28,259	30,578	32,801	37,697
16	3,942	5,142	5,812	6,614	6,908	7,962	9,312	11,152	20,465	23,542	26,296	28,845	29,633	32,000	34,267	39,252
17	4,416	5,697	6,408	7,255	7,564	8,672	10,085	12,002	21,615	24,769	27,587	30,191	30,995	33,409	35,718	40,790
18	4,905	6,265	7,015	7,906	8,231	9,390	10,865	12,857	22,760	25,989	28,869	31,526	32,346	34,805	37,156	42,312
19	5,407	6,844	7,633	8,567	8,907	10,117	11,651	13,716	23,900	27,204	30,144	32,852	33,687	36,191	38,582	43,820
20	5,921	7,434	8,260	9,237	9,591	10,851	12,443	14,578	25,038	28,412	31,410	34,170	35,020	37,566	39,997	45,315
21	6,447	8,034	8,897	9,915	10,283	11,591	13,240	15,445	26,171	29,615	32,671	35,479	36,343	38,932	41,401	46,797
22	6,983	8,643	9,542	10,600	10,982	12,338	14,041	16,314	27,301	30,813	33,924	36,781	37,659	40,289	42,796	48,268
23	7,529	9,260	10,196	11,293	11,689	13,091	14,848	17,187	28,429	32,007	35,172	38,076	38,968	41,638	44,181	49,728
24	8,085	9,886	10,856	11,992	12,401	13,848	15,659	18,062	29,553	33,196	36,415	39,364	40,270	42,980	45,559	51,179
25	8,649	10,520	11,524	12,697	13,120	14,611	16,473	18,940	30,675	34,382	37,652	40,646	41,566	44,314	46,928	52,620
26	9,222	11,160	12,198	13,409	13,844	15,379	17,292	19,820	31,795	35,563	38,885	41,923	42,856	45,642	48,290	54,052
27	9,803	11,808	12,879	14,125	14,573	16,151	18,114	20,703	32,912	36,741	40,113	43,195	44,140	46,963	49,645	55,476
28	10,391	12,461	13,565	14,847	15,308	16,928	18,939	21,588	34,027	37,916	41,337	44,461	45,419	48,278	50,993	56,892
29	10,986	13,121	14,256	15,574	16,047	17,708	19,768	22,475	35,139	39,087	42,557	45,722	46,693	49,588	52,336	58,301
30	11,588	13,787	14,953	16,306	16,791	18,493	20,599	23,364	36,250	40,256	43,773	46,979	47,962	50,892	53,672	59,703

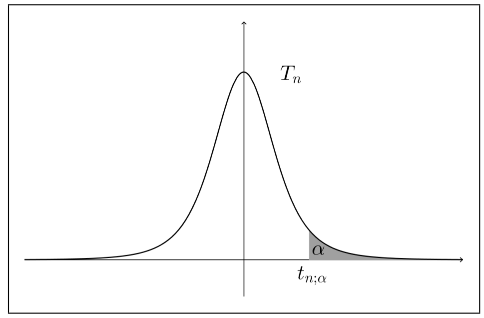
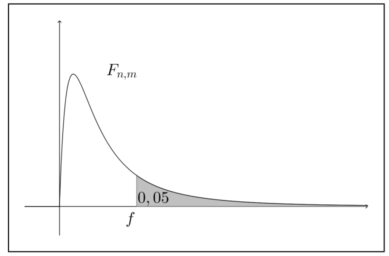


Tabela IV
Valores críticos $t_{n;\alpha}$ da t-Student
$P(T_n > t_{n;\alpha}) = \alpha$

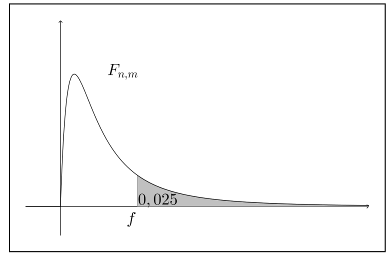
gl	Probabilidade α na cauda superior												
n	0,150	0,100	0,060	0,050	0,040	0,030	0,025	0,020	0,010	0,005	0,0025	0,002	0,001
1	1,963	3,078	5,242	6,314	7,916	10,579	12,706	15,895	31,821	63,657	127,321	159,153	318,309
2	1,386	1,886	2,620	2,920	3,320	3,896	4,303	4,849	6,965	9,925	14,089	15,764	22,327
3	1,250	1,638	2,156	2,353	2,605	2,951	3,182	3,482	4,541	5,841	7,453	8,053	10,215
4	1,190	1,533	1,971	2,132	2,333	2,601	2,776	2,999	3,747	4,604	5,598	5,951	7,173
5	1,156	1,476	1,873	2,015	2,191	2,422	2,571	2,757	3,365	4,032	4,773	5,030	5,893
6	1,134	1,440	1,812	1,943	2,104	2,313	2,447	2,612	3,143	3,707	4,317	4,524	5,208
7	1,119	1,415	1,770	1,895	2,046	2,241	2,365	2,517	2,998	3,499	4,029	4,207	4,785
8	1,108	1,397	1,740	1,860	2,004	2,189	2,306	2,449	2,896	3,355	3,833	3,991	4,501
9	1,100	1,383	1,718	1,833	1,973	2,150	2,262	2,398	2,821	3,250	3,690	3,835	4,297
10	1,093	1,372	1,700	1,812	1,948	2,120	2,228	2,359	2,764	3,169	3,581	3,716	4,144
11	1,088	1,363	1,686	1,796	1,928	2,096	2,201	2,328	2,718	3,106	3,497	3,624	4,025
12	1,083	1,356	1,674	1,782	1,912	2,076	2,179	2,303	2,681	3,055	3,428	3,550	3,930
13	1,079	1,350	1,664	1,771	1,899	2,060	2,160	2,282	2,650	3,012	3,372	3,489	3,852
14	1,076	1,345	1,656	1,761	1,887	2,046	2,145	2,264	2,624	2,977	3,326	3,438	3,787
15	1,074	1,341	1,649	1,753	1,878	2,034	2,131	2,249	2,602	2,947	3,286	3,395	3,733
16	1,071	1,337	1,642	1,746	1,869	2,024	2,120	2,235	2,583	2,921	3,252	3,358	3,686
17	1,069	1,333	1,637	1,740	1,862	2,015	2,110	2,224	2,567	2,898	3,222	3,326	3,646
18	1,067	1,330	1,632	1,734	1,855	2,007	2,101	2,214	2,552	2,878	3,197	3,298	3,610
19	1,066	1,328	1,628	1,729	1,850	2,000	2,093	2,205	2,539	2,861	3,174	3,273	3,579
20	1,064	1,325	1,624	1,725	1,844	1,994	2,086	2,197	2,528	2,845	3,153	3,251	3,552
21	1,063	1,323	1,621	1,721	1,840	1,988	2,080	2,189	2,518	2,831	3,135	3,231	3,527
22	1,061	1,321	1,618	1,717	1,835	1,983	2,074	2,183	2,508	2,819	3,119	3,214	3,505
23	1,060	1,319	1,615	1,714	1,832	1,978	2,069	2,177	2,500	2,807	3,104	3,198	3,485
24	1,059	1,318	1,612	1,711	1,828	1,974	2,064	2,172	2,492	2,797	3,091	3,183	3,467
25	1,058	1,316	1,610	1,708	1,825	1,970	2,060	2,167	2,485	2,787	3,078	3,170	3,450
26	1,058	1,315	1,608	1,706	1,822	1,967	2,056	2,162	2,479	2,779	3,067	3,158	3,435
27	1,057	1,314	1,606	1,703	1,819	1,963	2,052	2,158	2,473	2,771	3,057	3,147	3,421
28	1,056	1,313	1,604	1,701	1,817	1,960	2,048	2,154	2,467	2,763	3,047	3,136	3,408
29	1,055	1,311	1,602	1,699	1,814	1,957	2,045	2,150	2,462	2,756	3,038	3,127	3,396
30	1,055	1,310	1,600	1,697	1,812	1,955	2,042	2,147	2,457	2,750	3,030	3,118	3,385
31	1,054	1,309	1,599	1,696	1,810	1,952	2,040	2,144	2,453	2,744	3,022	3,109	3,375
32	1,054	1,309	1,597	1,694	1,808	1,950	2,037	2,141	2,449	2,738	3,015	3,102	3,365
33	1,053	1,308	1,596	1,692	1,806	1,948	2,035	2,138	2,445	2,733	3,008	3,094	3,356
34	1,052	1,307	1,595	1,691	1,805	1,946	2,032	2,136	2,441	2,728	3,002	3,088	3,348
35	1,052	1,306	1,594	1,690	1,803	1,944	2,030	2,133	2,438	2,724	2,996	3,081	3,340

Tabela V
Valores críticos f da distribuição $F_{n,m}$
$P(F_{n,m} > f) = 0,05$



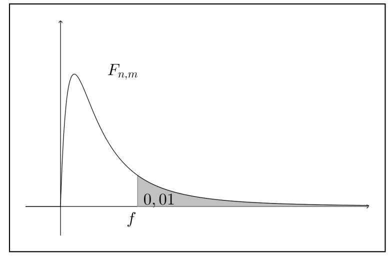
		GL numerador																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	1	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54	241,88	242,98	243,91	244,69	245,36	245,95	246,46	246,92	247,32	247,69	248,01
	2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,40	19,41	19,42	19,42	19,43	19,43	19,44	19,44	19,44	19,45
	3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74	8,73	8,71	8,70	8,69	8,68	8,67	8,67	8,66
	4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94	5,91	5,89	5,87	5,86	5,84	5,83	5,82	5,81	5,80
G	5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68	4,66	4,64	4,62	4,60	4,59	4,58	4,57	4,56
L	6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,98	3,96	3,94	3,92	3,91	3,90	3,88	3,87
	7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57	3,55	3,53	3,51	3,49	3,48	3,47	3,46	3,44
d	8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28	3,26	3,24	3,22	3,20	3,19	3,17	3,16	3,15
e	9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07	3,05	3,03	3,01	2,99	2,97	2,96	2,95	2,94
n	10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91	2,89	2,86	2,85	2,83	2,81	2,80	2,79	2,77
o	11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79	2,76	2,74	2,72	2,70	2,69	2,67	2,66	2,65
m	12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69	2,66	2,64	2,62	2,60	2,58	2,57	2,56	2,54
i	13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,60	2,58	2,55	2,53	2,51	2,50	2,48	2,47	2,46
n	14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,57	2,53	2,51	2,48	2,46	2,44	2,43	2,41	2,40	2,39
a	15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48	2,45	2,42	2,40	2,38	2,37	2,35	2,34	2,33
d	16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,46	2,42	2,40	2,37	2,35	2,33	2,32	2,30	2,29	2,28
o	17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41	2,38	2,35	2,33	2,31	2,29	2,27	2,26	2,24	2,23
r	18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,31	2,29	2,27	2,25	2,23	2,22	2,20	2,19
	19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34	2,31	2,28	2,26	2,23	2,21	2,20	2,18	2,17	2,16
	20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31	2,28	2,25	2,22	2,20	2,18	2,17	2,15	2,14	2,12
	21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25	2,22	2,20	2,18	2,16	2,14	2,12	2,11	2,10
	22	4,3	3,44	3,05	2,82	2,66	2,55	2,46	2,4	2,34	2,3	2,26	2,23	2,20	2,17	2,15	2,13	2,11	2,10	2,08	2,07
	23	4,28	3,42	3,03	2,8	2,64	2,53	2,44	2,37	2,32	2,27	2,24	2,2	2,18	2,15	2,13	2,11	2,09	2,08	2,06	2,05
	24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,22	2,18	2,15	2,13	2,11	2,09	2,07	2,05	2,04	2,03
	25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,20	2,16	2,14	2,11	2,09	2,07	2,05	2,04	2,02	2,01

Tabela VI
Valores críticos f da distribuição $F_{n,m}$
$P(F_{n,m} > f) = 0,025$



GL denominador	GL numerador 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	647,79	799,50	864,16	899,58	921,85	937,11	948,22	956,66	963,28	968,63	973,03	976,71	979,84	982,53	984,87	986,92	988,73	990,35	991,80	993,10
2	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,41	39,41	39,42	39,43	39,43	39,44	39,44	39,44	39,45	39,45
3	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,37	14,34	14,30	14,28	14,25	14,23	14,21	14,20	14,18	14,17
4	12,22	10,65	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,79	8,75	8,71	8,68	8,66	8,63	8,61	8,59	8,58	8,56
5	10,01	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,57	6,52	6,49	6,46	6,43	6,40	6,38	6,36	6,34	6,33
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46	5,41	5,37	5,33	5,30	5,27	5,24	5,22	5,20	5,18	5,17
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82	4,76	4,71	4,67	4,63	4,60	4,57	4,54	4,52	4,50	4,48	4,47
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30	4,24	4,20	4,16	4,13	4,10	4,08	4,05	4,03	4,02	4,00
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96	3,91	3,87	3,83	3,80	3,77	3,74	3,72	3,70	3,68	3,67
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72	3,66	3,62	3,58	3,55	3,52	3,50	3,47	3,45	3,44	3,42
11	6,72	5,26	4,63	4,28	4,04	3,88	3,76	3,66	3,59	3,53	3,47	3,43	3,39	3,36	3,33	3,30	3,28	3,26	3,24	3,23
12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37	3,32	3,28	3,24	3,21	3,18	3,15	3,13	3,11	3,09	3,07
13	6,41	4,97	4,35	4,00	3,77	3,60	3,48	3,39	3,31	3,25	3,20	3,15	3,12	3,08	3,05	3,03	3,00	2,98	2,96	2,95
14	6,30	4,86	4,24	3,89	3,66	3,50	3,38	3,29	3,21	3,15	3,09	3,05	3,01	2,98	2,95	2,92	2,90	2,88	2,86	2,84
15	6,20	4,77	4,15	3,80	3,58	3,41	3,29	3,20	3,12	3,06	3,01	2,96	2,92	2,89	2,86	2,84	2,81	2,79	2,77	2,76
16	6,12	4,69	4,08	3,73	3,50	3,34	3,22	3,12	3,05	2,99	2,93	2,89	2,85	2,82	2,79	2,76	2,74	2,72	2,70	2,68
17	6,04	4,62	4,01	3,66	3,44	3,28	3,16	3,06	2,98	2,92	2,87	2,82	2,79	2,75	2,72	2,70	2,67	2,65	2,63	2,62
18	5,98	4,56	3,95	3,61	3,38	3,22	3,10	3,01	2,93	2,87	2,81	2,77	2,73	2,70	2,67	2,64	2,62	2,60	2,58	2,56
19	5,92	4,51	3,90	3,56	3,33	3,17	3,05	2,96	2,88	2,82	2,76	2,72	2,68	2,65	2,62	2,59	2,57	2,55	2,53	2,51
20	5,87	4,46	3,86	3,51	3,29	3,13	3,01	2,91	2,84	2,77	2,72	2,68	2,64	2,60	2,57	2,55	2,52	2,5	2,48	2,46
21	5,83	4,42	3,82	3,48	3,25	3,09	2,97	2,87	2,80	2,73	2,68	2,64	2,60	2,56	2,53	2,51	2,48	2,46	2,44	2,42
22	5,79	4,38	3,78	3,44	3,22	3,05	2,93	2,84	2,76	2,70	2,65	2,60	2,56	2,53	2,50	2,47	2,45	2,43	2,41	2,39
23	5,75	4,35	3,75	3,41	3,18	3,02	2,90	2,81	2,73	2,67	2,62	2,57	2,53	2,50	2,47	2,44	2,42	2,39	2,37	2,36
24	5,72	4,32	3,72	3,38	3,15	2,99	2,87	2,78	2,70	2,64	2,59	2,54	2,50	2,47	2,44	2,41	2,39	2,36	2,35	2,33
25	5,69	4,29	3,69	3,35	3,13	2,97	2,85	2,75	2,68	2,61	2,56	2,51	2,48	2,44	2,41	2,38	2,36	2,34	2,32	2,30

Tabela VII
Valores críticos f da distribuição $F_{n,m}$
$P(F_{n,m} > f) = 0,01$



		GL numerador																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	1	4052,18	4999,50	5403,35	5624,58	5763,65	5858,99	5928,36	5981,07	6022,47	6055,85	6083,32	6106,32	6125,86	6142,67	6157,28	6170,10	6181,43	6191,53	6200,58	6208,73
	2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,41	99,42	99,42	99,43	99,43	99,44	99,44	99,44	99,45	99,45
	3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,13	27,05	26,98	26,92	26,87	26,83	26,79	26,75	26,72	26,69
	4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,45	14,37	14,31	14,25	14,20	14,15	14,11	14,08	14,05	14,02
G	5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,96	9,89	9,82	9,77	9,72	9,68	9,64	9,61	9,58	9,55
L	6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72	7,66	7,60	7,56	7,52	7,48	7,45	7,42	7,40
	7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,54	6,47	6,41	6,36	6,31	6,28	6,24	6,21	6,18	6,16
d	8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,73	5,67	5,61	5,56	5,52	5,48	5,44	5,41	5,38	5,36
e	9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,18	5,11	5,05	5,01	4,96	4,92	4,89	4,86	4,83	4,81
n	10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,77	4,71	4,65	4,60	4,56	4,52	4,49	4,46	4,43	4,41
o	11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,46	4,40	4,34	4,29	4,25	4,21	4,18	4,15	4,12	4,10
m	12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,22	4,16	4,10	4,05	4,01	3,97	3,94	3,91	3,88	3,86
i	13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96	3,91	3,86	3,82	3,78	3,75	3,72	3,69	3,66
n	14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80	3,75	3,70	3,66	3,62	3,59	3,56	3,53	3,51
a	15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67	3,61	3,56	3,52	3,49	3,45	3,42	3,40	3,37
d	16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,62	3,55	3,50	3,45	3,41	3,37	3,34	3,31	3,28	3,26
o	17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,46	3,40	3,35	3,31	3,27	3,24	3,21	3,19	3,16
r	18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,43	3,37	3,32	3,27	3,23	3,19	3,16	3,13	3,10	3,08
	19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,36	3,30	3,24	3,19	3,15	3,12	3,08	3,05	3,03	3,00
	20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,29	3,23	3,18	3,13	3,09	3,05	3,02	2,99	2,96	2,94
	21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,24	3,17	3,12	3,07	3,03	2,99	2,96	2,93	2,90	2,88
	22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,18	3,12	3,07	3,02	2,98	2,94	2,91	2,88	2,85	2,83
	23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,14	3,07	3,02	2,97	2,93	2,89	2,86	2,83	2,80	2,78
	24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,09	3,03	2,98	2,93	2,89	2,85	2,82	2,79	2,76	2,74
	25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	3,06	2,99	2,94	2,89	2,85	2,81	2,78	2,75	2,72	2,70

Tabela VIII
Valores críticos $q_{p,\nu}$ da amplitude studentized – $\alpha = 0,05$

ν	p										
	2	3	4	5	6	7	8	9	10	11	12
2	6,080	8,331	9,799	10,881	11,734	12,435	13,028	13,542	13,994	14,396	14,759
3	4,501	5,910	6,825	7,502	8,037	8,478	8,852	9,177	9,462	9,717	9,946
4	3,927	5,040	5,757	6,287	6,706	7,053	7,347	7,602	7,826	8,027	8,208
5	3,635	4,602	5,218	5,673	6,033	6,330	6,582	6,801	6,995	7,167	7,323
6	3,460	4,339	4,896	5,305	5,628	5,895	6,122	6,319	6,493	6,649	6,789
7	3,344	4,165	4,681	5,060	5,359	5,606	5,815	5,997	6,158	6,302	6,431
8	3,261	4,041	4,529	4,886	5,167	5,399	5,596	5,767	5,918	6,053	6,175
9	3,199	3,948	4,415	4,755	5,024	5,244	5,432	5,595	5,738	5,867	5,983
10	3,151	3,877	4,327	4,654	4,912	5,124	5,304	5,460	5,598	5,722	5,833
11	3,113	3,820	4,256	4,574	4,823	5,028	5,202	5,353	5,486	5,605	5,713
12	3,081	3,773	4,199	4,508	4,750	4,950	5,119	5,265	5,395	5,510	5,615
13	3,055	3,734	4,151	4,453	4,690	4,884	5,049	5,192	5,318	5,431	5,533
14	3,033	3,701	4,111	4,407	4,639	4,829	4,990	5,130	5,253	5,364	5,463
15	3,014	3,673	4,076	4,367	4,595	4,782	4,940	5,077	5,198	5,306	5,403
16	2,998	3,649	4,046	4,333	4,557	4,741	4,896	5,031	5,150	5,256	5,352
17	2,984	3,628	4,020	4,303	4,524	4,705	4,858	4,991	5,108	5,212	5,306
18	2,971	3,609	3,997	4,276	4,494	4,673	4,824	4,955	5,071	5,173	5,266
19	2,960	3,593	3,977	4,253	4,468	4,645	4,794	4,924	5,037	5,139	5,231
20	2,950	3,578	3,958	4,232	4,445	4,620	4,768	4,895	5,008	5,108	5,199
21	2,941	3,565	3,942	4,213	4,424	4,597	4,743	4,870	4,981	5,081	5,170
22	2,933	3,553	3,927	4,196	4,405	4,577	4,722	4,847	4,957	5,056	5,144
23	2,926	3,542	3,914	4,180	4,388	4,558	4,702	4,826	4,935	5,033	5,121
24	2,919	3,532	3,901	4,166	4,373	4,541	4,684	4,807	4,915	5,012	5,099
25	2,913	3,523	3,890	4,153	4,358	4,526	4,667	4,789	4,897	4,993	5,079
26	2,907	3,514	3,880	4,141	4,345	4,511	4,652	4,773	4,880	4,975	5,061
27	2,902	3,506	3,870	4,130	4,333	4,498	4,638	4,758	4,864	4,959	5,044
28	2,784	3,332	3,655	3,883	4,058	4,200	4,319	4,421	4,511	4,590	4,662
29	2,892	3,493	3,853	4,111	4,311	4,475	4,613	4,732	4,837	4,930	5,014
30	2,897	3,499	3,861	4,120	4,322	4,486	4,625	4,745	4,850	4,944	5,029
31	2,892	3,493	3,853	4,111	4,311	4,475	4,613	4,732	4,837	4,930	5,014
32	2,888	3,486	3,845	4,102	4,301	4,464	4,601	4,720	4,824	4,917	5,001
33	2,884	3,481	3,838	4,094	4,292	4,454	4,591	4,709	4,812	4,905	4,988
34	2,881	3,475	3,832	4,086	4,284	4,445	4,581	4,698	4,802	4,894	4,976
35	2,871	3,461	3,814	4,066	4,261	4,421	4,555	4,671	4,773	4,863	4,945
36	2,868	3,457	3,809	4,060	4,255	4,414	4,547	4,663	4,764	4,855	4,936
37	2,865	3,453	3,804	4,054	4,249	4,407	4,540	4,655	4,756	4,846	4,927
38	2,863	3,449	3,799	4,049	4,243	4,400	4,533	4,648	4,749	4,838	4,919
39	2,861	3,445	3,795	4,044	4,237	4,394	4,527	4,641	4,741	4,831	4,911
40	2,858	3,442	3,791	4,039	4,232	4,388	4,521	4,634	4,735	4,824	4,904
50	2,841	3,416	3,758	4,002	4,190	4,344	4,473	4,584	4,681	4,768	4,846
60	2,829	3,399	3,737	3,977	4,163	4,314	4,441	4,550	4,646	4,732	4,808
70	2,821	3,386	3,722	3,960	4,144	4,293	4,419	4,527	4,621	4,706	4,781
80	2,814	3,377	3,711	3,947	4,129	4,277	4,402	4,509	4,603	4,686	4,761
90	2,810	3,370	3,702	3,937	4,118	4,265	4,389	4,495	4,588	4,671	4,746
100	2,806	3,365	3,695	3,929	4,109	4,256	4,379	4,484	4,577	4,659	4,733

Fonte: Valores gerados com a função `ptukey` do R

Tabela IX

Valores críticos $q_{p,\nu}$ da amplitude studentized – $\alpha = 0,01$

ν	p										
	2	3	4	5	6	7	8	9	10	11	12
2	13,902	19,016	22,564	25,372	27,757	29,856	31,730	33,412	34,926	36,293	37,533
3	8,260	10,620	12,170	13,322	14,239	14,998	15,646	16,212	16,713	17,164	17,573
4	6,511	8,120	9,173	9,958	10,583	11,101	11,542	11,925	12,263	12,565	12,839
5	5,702	6,976	7,804	8,421	8,913	9,321	9,669	9,971	10,239	10,479	10,696
6	5,243	6,331	7,033	7,556	7,972	8,318	8,612	8,869	9,097	9,300	9,485
7	4,949	5,919	6,542	7,005	7,373	7,678	7,939	8,166	8,367	8,548	8,711
8	4,745	5,635	6,204	6,625	6,959	7,237	7,474	7,680	7,863	8,027	8,176
9	4,596	5,428	5,957	6,347	6,657	6,915	7,134	7,325	7,494	7,646	7,784
10	4,482	5,270	5,769	6,136	6,428	6,669	6,875	7,054	7,213	7,356	7,485
11	4,392	5,146	5,621	5,970	6,247	6,476	6,671	6,841	6,992	7,127	7,250
12	4,320	5,046	5,502	5,836	6,101	6,320	6,507	6,670	6,814	6,943	7,060
13	4,260	4,964	5,404	5,726	5,981	6,192	6,372	6,528	6,666	6,791	6,903
14	4,210	4,895	5,322	5,634	5,881	6,085	6,258	6,409	6,543	6,663	6,772
15	4,167	4,836	5,252	5,556	5,796	5,994	6,162	6,309	6,438	6,555	6,660
16	4,131	4,786	5,192	5,489	5,722	5,915	6,079	6,222	6,348	6,461	6,564
17	4,099	4,742	5,140	5,430	5,659	5,847	6,007	6,147	6,270	6,380	6,480
18	4,071	4,703	5,094	5,379	5,603	5,787	5,944	6,081	6,201	6,309	6,407
19	4,046	4,669	5,054	5,334	5,553	5,735	5,889	6,022	6,141	6,246	6,342
20	4,024	4,639	5,018	5,293	5,510	5,688	5,839	5,970	6,086	6,190	6,285
21	4,004	4,612	4,986	5,257	5,470	5,646	5,794	5,924	6,038	6,140	6,233
22	3,986	4,588	4,957	5,225	5,435	5,608	5,754	5,882	5,994	6,095	6,186
23	3,970	4,566	4,931	5,195	5,403	5,573	5,718	5,844	5,955	6,054	6,144
24	3,955	4,546	4,907	5,168	5,373	5,542	5,685	5,809	5,919	6,017	6,105
25	3,942	4,527	4,885	5,144	5,347	5,513	5,655	5,778	5,886	5,983	6,070
26	3,930	4,510	4,865	5,121	5,322	5,487	5,627	5,749	5,856	5,951	6,038
27	3,918	4,495	4,847	5,101	5,300	5,463	5,602	5,722	5,828	5,923	6,008
28	3,908	4,481	4,830	5,082	5,279	5,441	5,578	5,697	5,802	5,896	5,981
29	3,898	4,467	4,814	5,064	5,260	5,420	5,556	5,674	5,778	5,871	5,955
30	3,889	4,455	4,799	5,048	5,242	5,401	5,536	5,653	5,756	5,848	5,932
31	3,881	4,443	4,786	5,032	5,225	5,383	5,517	5,633	5,736	5,827	5,910
32	3,873	4,433	4,773	5,018	5,210	5,367	5,500	5,615	5,716	5,807	5,889
33	3,865	4,423	4,761	5,005	5,195	5,351	5,483	5,598	5,698	5,789	5,870
34	3,859	4,413	4,750	4,992	5,181	5,336	5,468	5,581	5,682	5,771	5,852
35	3,852	4,404	4,739	4,980	5,169	5,323	5,453	5,566	5,666	5,755	5,835
36	3,846	4,396	4,729	4,969	5,156	5,310	5,439	5,552	5,651	5,739	5,819
37	3,840	4,388	4,720	4,959	5,145	5,298	5,427	5,538	5,637	5,725	5,804
38	3,835	4,381	4,711	4,949	5,134	5,286	5,414	5,526	5,623	5,711	5,790
39	3,830	4,374	4,703	4,940	5,124	5,275	5,403	5,513	5,611	5,698	5,776
40	3,825	4,367	4,695	4,931	5,114	5,265	5,392	5,502	5,599	5,685	5,764
50	3,787	4,316	4,634	4,863	5,040	5,185	5,308	5,414	5,507	5,590	5,665
60	3,762	4,282	4,594	4,818	4,991	5,133	5,253	5,356	5,447	5,528	5,601
70	3,745	4,258	4,566	4,786	4,957	5,096	5,214	5,315	5,404	5,483	5,555
80	3,732	4,241	4,545	4,763	4,931	5,069	5,185	5,284	5,372	5,451	5,521
90	3,722	4,227	4,529	4,745	4,911	5,048	5,162	5,261	5,348	5,425	5,495
100	3,714	4,216	4,516	4,730	4,896	5,031	5,144	5,242	5,328	5,405	5,474

Fonte: Valores gerados com a função `ptukey` do R

Tabela X

Valores críticos $r_{p,v;0,05}$ para o teste de Duncan – $\alpha = 0,05$

gl	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	17,969	17,969	17,969	17,969	17,969	17,969	17,969	17,969	17,969	17,969	17,969	17,969	17,969	17,969	17,969	17,969	17,969	17,969	17,969
2	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085
3	4,501	4,516	4,516	4,516	4,516	4,516	4,516	4,516	4,516	4,516	4,516	4,516	4,516	4,516	4,516	4,516	4,516	4,516	4,516
4	3,926	4,013	4,033	4,033	4,033	4,033	4,033	4,033	4,033	4,033	4,033	4,033	4,033	4,033	4,033	4,033	4,033	4,033	4,033
5	3,635	3,749	3,796	3,814	3,814	3,814	3,814	3,814	3,814	3,814	3,814	3,814	3,814	3,814	3,814	3,814	3,814	3,814	3,814
6	3,460	3,586	3,649	3,680	3,694	3,697	3,697	3,697	3,697	3,697	3,697	3,697	3,697	3,697	3,697	3,697	3,697	3,697	3,697
7	3,344	3,477	3,548	3,588	3,611	3,622	3,625	3,625	3,625	3,625	3,625	3,625	3,625	3,625	3,625	3,625	3,625	3,625	3,625
8	3,261	3,398	3,475	3,521	3,549	3,566	3,575	3,579	3,579	3,579	3,579	3,579	3,579	3,579	3,579	3,579	3,579	3,579	3,579
9	3,199	3,339	3,420	3,470	3,502	3,523	3,536	3,544	3,547	3,547	3,547	3,547	3,547	3,547	3,547	3,547	3,547	3,547	3,547
10	3,151	3,293	3,376	3,430	3,465	3,489	3,505	3,516	3,522	3,525	3,525	3,525	3,525	3,525	3,525	3,525	3,525	3,525	3,525
11	3,113	3,256	3,341	3,397	3,435	3,462	3,480	3,493	3,501	3,506	3,509	3,510	3,510	3,510	3,510	3,510	3,510	3,510	3,510
12	3,081	3,225	3,312	3,370	3,410	3,439	3,459	3,474	3,484	3,491	3,495	3,498	3,498	3,498	3,498	3,498	3,498	3,498	3,498
13	3,055	3,200	3,288	3,348	3,389	3,419	3,441	3,458	3,470	3,478	3,484	3,488	3,490	3,490	3,490	3,490	3,490	3,490	3,490
14	3,033	3,178	3,268	3,328	3,371	3,403	3,426	3,444	3,457	3,467	3,474	3,479	3,482	3,484	3,484	3,484	3,484	3,484	3,484
15	3,014	3,160	3,250	3,312	3,356	3,389	3,413	3,432	3,446	3,457	3,465	3,471	3,476	3,478	3,480	3,480	3,480	3,480	3,480
16	2,998	3,144	3,235	3,297	3,343	3,376	3,402	3,422	3,437	3,449	3,458	3,465	3,470	3,473	3,476	3,477	3,477	3,477	3,477
17	2,984	3,130	3,222	3,285	3,331	3,365	3,392	3,412	3,429	3,441	3,451	3,459	3,465	3,469	3,472	3,474	3,475	3,475	3,475
18	2,971	3,117	3,210	3,274	3,320	3,356	3,383	3,404	3,421	3,435	3,445	3,454	3,460	3,465	3,469	3,472	3,473	3,474	3,474
19	2,960	3,106	3,199	3,264	3,311	3,347	3,375	3,397	3,415	3,429	3,440	3,449	3,456	3,462	3,466	3,469	3,472	3,473	3,474
20	2,950	3,097	3,190	3,255	3,303	3,339	3,368	3,390	3,409	3,423	3,435	3,445	3,452	3,459	3,463	3,467	3,470	3,472	3,473
21	2,941	3,088	3,181	3,247	3,295	3,332	3,361	3,385	3,403	3,418	3,431	3,441	3,449	3,456	3,461	3,465	3,469	3,471	3,473
22	2,933	3,080	3,173	3,239	3,288	3,326	3,355	3,379	3,398	3,414	3,427	3,437	3,446	3,453	3,459	3,464	3,467	3,470	3,472
23	2,926	3,072	3,166	3,233	3,282	3,320	3,350	3,374	3,394	3,410	3,423	3,434	3,443	3,451	3,457	3,462	3,466	3,469	3,472
24	2,919	3,066	3,160	3,226	3,276	3,315	3,345	3,370	3,390	3,406	3,420	3,431	3,441	3,449	3,455	3,461	3,465	3,469	3,472
25	2,913	3,059	3,154	3,221	3,271	3,310	3,341	3,366	3,386	3,403	3,417	3,429	3,439	3,447	3,454	3,459	3,464	3,468	3,471
26	2,907	3,054	3,149	3,216	3,266	3,305	3,336	3,362	3,382	3,400	3,414	3,426	3,436	3,445	3,452	3,458	3,463	3,468	3,471
27	2,902	3,049	3,144	3,211	3,262	3,301	3,332	3,358	3,379	3,397	3,412	3,424	3,434	3,443	3,451	3,457	3,463	3,467	3,471
28	2,897	3,044	3,139	3,206	3,257	3,297	3,329	3,355	3,376	3,394	3,409	3,422	3,433	3,442	3,450	3,456	3,462	3,467	3,470
29	2,892	3,039	3,135	3,202	3,253	3,293	3,326	3,352	3,373	3,392	3,407	3,420	3,431	3,440	3,448	3,455	3,461	3,466	3,470
30	2,888	3,035	3,131	3,199	3,250	3,290	3,322	3,349	3,371	3,389	3,405	3,418	3,429	3,439	3,447	3,454	3,460	3,466	3,470
35	2,871	3,018	3,114	3,183	3,235	3,276	3,309	3,337	3,360	3,379	3,396	3,410	3,423	3,433	3,443	3,451	3,458	3,464	3,469
40	2,858	3,005	3,102	3,171	3,224	3,266	3,300	3,328	3,352	3,372	3,389	3,404	3,418	3,429	3,439	3,448	3,456	3,463	3,469
60	2,829	2,976	3,073	3,143	3,198	3,241	3,277	3,307	3,333	3,355	3,374	3,391	3,406	3,419	3,431	3,441	3,451	3,460	3,468
80	2,814	2,961	3,059	3,130	3,185	3,229	3,266	3,297	3,323	3,346	3,366	3,384	3,400	3,414	3,427	3,438	3,449	3,458	3,467
120	2,800	2,947	3,045	3,116	3,172	3,217	3,254	3,286	3,313	3,337	3,358	3,377	3,394	3,409	3,423	3,435	3,446	3,457	3,466
240	2,786	2,933	3,031	3,103	3,159	3,205	3,243	3,276	3,304	3,329	3,350	3,370	3,388	3,404	3,418	3,432	3,444	3,455	3,466
∞	2,772	2,918	3,017	3,089	3,146	3,193	3,232	3,265	3,294	3,320	3,343	3,363	3,382	3,399	3,414	3,428	3,442	3,454	3,466

Fonte: http://www2.accsnet.ne.jp/ miwa/probcalc/duncan/index.html

Tabela XI

Valores críticos $r_{p,v;0,05}$ para o teste de Duncan – $\alpha = 0,01$

gl	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	90,024	90,024	90,024	90,024	90,024	90,024	90,024	90,024	90,024	90,024	90,024	90,024	90,024	90,024	90,024	90,024	90,024	90,024	90,024
2	14,036	14,036	14,036	14,036	14,036	14,036	14,036	14,036	14,036	14,036	14,036	14,036	14,036	14,036	14,036	14,036	14,036	14,036	14,036
3	8,260	8,321	8,321	8,321	8,321	8,321	8,321	8,321	8,321	8,321	8,321	8,321	8,321	8,321	8,321	8,321	8,321	8,321	8,321
4	6,511	6,677	6,740	6,755	6,755	6,755	6,755	6,755	6,755	6,755	6,755	6,755	6,755	6,755	6,755	6,755	6,755	6,755	6,755
5	5,702	5,893	5,989	6,040	6,065	6,074	6,074	6,074	6,074	6,074	6,074	6,074	6,074	6,074	6,074	6,074	6,074	6,074	6,074
6	5,243	5,439	5,549	5,614	5,655	5,680	5,694	5,701	5,703	5,703	5,703	5,703	5,703	5,703	5,703	5,703	5,703	5,703	5,703
7	4,949	5,145	5,260	5,333	5,383	5,416	5,439	5,454	5,464	5,470	5,472	5,472	5,472	5,472	5,472	5,472	5,472	5,472	5,472
8	4,745	4,939	5,056	5,134	5,189	5,227	5,256	5,276	5,291	5,302	5,309	5,313	5,316	5,317	5,317	5,317	5,317	5,317	5,317
9	4,596	4,787	4,906	4,986	5,043	5,086	5,117	5,142	5,160	5,174	5,185	5,193	5,199	5,202	5,205	5,206	5,206	5,206	5,206
10	4,482	4,671	4,789	4,871	4,931	4,975	5,010	5,036	5,058	5,074	5,087	5,098	5,106	5,112	5,117	5,120	5,122	5,123	5,124
11	4,392	4,579	4,697	4,780	4,841	4,887	4,923	4,952	4,975	4,994	5,009	5,021	5,031	5,039	5,045	5,050	5,054	5,057	5,059
12	4,320	4,504	4,622	4,705	4,767	4,815	4,852	4,882	4,907	4,927	4,944	4,957	4,969	4,978	4,986	4,993	4,998	5,002	5,005
13	4,260	4,442	4,560	4,643	4,706	4,754	4,793	4,824	4,850	4,871	4,889	4,904	4,917	4,927	4,936	4,944	4,950	4,955	4,960
14	4,210	4,391	4,508	4,591	4,654	4,703	4,743	4,775	4,802	4,824	4,843	4,859	4,872	4,884	4,894	4,902	4,909	4,916	4,921
15	4,167	4,346	4,463	4,547	4,610	4,660	4,700	4,733	4,760	4,783	4,803	4,820	4,834	4,846	4,857	4,866	4,874	4,881	4,887
16	4,131	4,308	4,425	4,508	4,572	4,622	4,662	4,696	4,724	4,748	4,768	4,785	4,800	4,813	4,825	4,835	4,843	4,851	4,858
17	4,099	4,275	4,391	4,474	4,538	4,589	4,630	4,664	4,692	4,717	4,737	4,755	4,771	4,785	4,797	4,807	4,816	4,824	4,832
18	4,071	4,246	4,361	4,445	4,509	4,559	4,601	4,635	4,664	4,689	4,710	4,729	4,745	4,759	4,771	4,782	4,792	4,801	4,808
19	4,046	4,220	4,335	4,418	4,483	4,533	4,575	4,610	4,639	4,664	4,686	4,705	4,722	4,736	4,749	4,760	4,771	4,780	4,788
20	4,024	4,197	4,312	4,395	4,459	4,510	4,552	4,587	4,617	4,642	4,664	4,684	4,701	4,716	4,729	4,741	4,751	4,761	4,769
21	4,004	4,177	4,291	4,374	4,438	4,489	4,531	4,567	4,597	4,622	4,645	4,664	4,682	4,697	4,711	4,723	4,734	4,743	4,752
22	3,986	4,158	4,272	4,355	4,419	4,470	4,513	4,548	4,578	4,604	4,627	4,647	4,664	4,680	4,694	4,706	4,718	4,728	4,737
23	3,970	4,141	4,254	4,337	4,402	4,453	4,496	4,531	4,562	4,588	4,611	4,631	4,649	4,665	4,679	4,692	4,703	4,713	4,723
24	3,955	4,126	4,239	4,322	4,386	4,437	4,480	4,516	4,546	4,573	4,596	4,616	4,634	4,651	4,665	4,678	4,690	4,700	4,710
25	3,942	4,112	4,224	4,307	4,371	4,423	4,466	4,502	4,532	4,559	4,582	4,603	4,621	4,638	4,652	4,665	4,677	4,688	4,698
26	3,930	4,099	4,211	4,294	4,358	4,410	4,452	4,489	4,520	4,546	4,570	4,591	4,609	4,626	4,640	4,654	4,666	4,677	4,687
27	3,918	4,087	4,199	4,282	4,346	4,397	4,440	4,477	4,508	4,535	4,558	4,579	4,598	4,615	4,630	4,643	4,655	4,667	4,677
28	3,908	4,076	4,188	4,270	4,334	4,386	4,429	4,465	4,497	4,524	4,548	4,569	4,587	4,604	4,619	4,633	4,646	4,657	4,667
29	3,898	4,065	4,177	4,260	4,324	4,376	4,419	4,455	4,486	4,514	4,538	4,559	4,578	4,595	4,610	4,624	4,637	4,648	4,659
30	3,889	4,056	4,168	4,250	4,314	4,366	4,409	4,445	4,477	4,504	4,528	4,550	4,569	4,586	4,601	4,615	4,628	4,640	4,650
35	3,852	4,017	4,128	4,210	4,273	4,325	4,369	4,406	4,437	4,465	4,490	4,511	4,531	4,549	4,565	4,579	4,593	4,605	4,616
40	3,825	3,988	4,098	4,180	4,243	4,295	4,339	4,376	4,408	4,436	4,461	4,483	4,503	4,521	4,537	4,552	4,566	4,579	4,591
60	3,762	3,922	4,030	4,111	4,174	4,226	4,270	4,307	4,340	4,368	4,394	4,417	4,437	4,456	4,474	4,489	4,504	4,518	4,530
80	3,732	3,890	3,997	4,077	4,140	4,192	4,236	4,273	4,306	4,335	4,360	4,384	4,405	4,424	4,442	4,458	4,473	4,487	4,500
120	3,702	3,858	3,964	4,044	4,107	4,158	4,202	4,239	4,272	4,301	4,327	4,351	4,372	4,392	4,410	4,426	4,442	4,456	4,469
240	3,672	3,827	3,932	4,011	4,073	4,125	4,168	4,206	4,239	4,268	4,294	4,318	4,339	4,359	4,378	4,394	4,410	4,425	4,439
∞	3,643	3,796	3,900	3,978	4,040	4,091	4,135	4,172	4,205	4,235	4,261	4,285	4,307	4,327	4,345	4,363	4,379	4,394	4,408

Fonte: http://www2.accsnet.ne.jp/ miwa/probcalc/duncan/index.html

Apêndice B

Solução dos Exercícios

B.1 Exercícios do capítulo 1

1. (a) Como o valor se refere aos pacientes estudados, e não a *todos* os pacientes, esse é o valor de uma estatística amostral.
   (b) Estatística amostral - foi testada uma amostra
   (c) Parâmetro populacional - a companhia realizou os cálculos com base em todos os clientes
   (d) Parâmetro populacional - o fabricante está se referindo à população de todas as baterias. Embora ele não tenha condições de testar todas as baterias, ele está fazendo uma hipótese sobre o parâmetro populacional. No estudo de teste de hipóteses, o procedimento se baseia na hipótese que se faz sobre um parâmetro populacional.
2. Como estamos trabalhando com amostras de tamanho 2, a média e a mediana amostrais são as mesmas, ou seja, a distribuição amostral da mediana amostral $\widehat{Q}_2$ é igual à distribuição da média amostral. Logo, $\mathrm{E}\left(\widehat{Q}_2\right) = 4,0$, que é diferente da mediana populacional, que é 3,5.
3.

$$X \sim Bern(p) \Rightarrow \sum_{i=1}^{n} X_i \sim Bin\,(n;p) \Rightarrow \qquad \mathrm{E}\left(\sum_{i=1}^{n} X_i\right) = np \qquad \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = np(1-p)$$

$$\begin{aligned}
\mathrm{E}(\tilde{P}) &= \frac{n}{n+1}p \\
\mathrm{B}(\tilde{P}) &= \frac{n}{n+1}p - p = \frac{-p}{n+1} \\
\mathrm{Var}(\tilde{P}) &= \frac{np(1-p)}{(n+1)^2} \\
\mathrm{EQM}(\tilde{P}) &= \frac{np(1-p)}{(n+1)^2} + \frac{p^2}{(n+1)^2} = \frac{np(1-p)+p^2}{(n+1)^2}
\end{aligned}$$

   Note que $\tilde{P}$ é viesado para estimar p e para um tamanho n fixo de amostra, o EQM máximo ocorre quando $p = \dfrac{1}{2}\dfrac{n}{n-1}$.
4.

$$\mathrm{E}(\widetilde{\sigma}^2) = \sigma^2 \Leftrightarrow C\left[\mathrm{E}(X_1 - X_2)^2 + \mathrm{E}(X_3 - X_4)^2 + \mathrm{E}(X_{-}X_6)^2\right] = \sigma^2$$

   Como as X_i's são iid, segue que

$$\begin{aligned}
\mathrm{E}(X_i - X_{i+1})^2 &= \mathrm{E}(X_i^2) + \mathrm{E}(X_{i+1}^2) - 2\,\mathrm{E}(X_i X_{i+1}) \\
&= \mathrm{Var}(X_i) + [\mathrm{E}(X_i)]^2 + \mathrm{Var}(X_{i+1}) + [\mathrm{E}(X_{i+1})]^2 - 2\,\mathrm{E}(X_i)\,\mathrm{E}(X_{i+1}) \\
&= \sigma^2 + \mu^2 + \sigma^2 + \mu^2 - 2\mu\cdot\mu = 2\sigma^2
\end{aligned}$$

Logo,

$$E(\widetilde{\sigma}^2) = \sigma^2 \Leftrightarrow C\left[3 \cdot 2\sigma^2\right] = \sigma^2 \Leftrightarrow C = \frac{1}{6}$$

5. .

	Vento (X)	Onda (Y)	XY	X^2
	9	2,9	26,1	81
	11	1,4	15,4	121
	10	1,7	17,0	100
	10	0,9	9,0	100
	11	1,2	13,2	121
	9	1,0	9,0	81
	9	1,5	13,5	81
	6	0,7	4,2	36
	9	1,9	17,1	81
	5	0,1	0,5	25
	8	2,0	16,0	64
	9	2,6	23,4	81
	12	3,0	36,0	144
	9	1,7	15,3	81
	12	2,1	25,2	144
	9	1,5	13,5	81
	12	3,1	37,2	144
	8	2,7	21,6	64
	7	0,4	2,8	49
	13	2,5	32,5	169
	9	1,7	15,3	81
	8	0,6	4,8	64
	6	0,7	4,2	36
	8	1,4	11,2	64
Soma	219	39,3	384	2093

$$\begin{aligned} \widehat{b} &= \frac{384 - \dfrac{219 \cdot 39,3}{24}}{2093 - \dfrac{219^2}{24}} = 0,268296 \\ \widehat{a} &= \frac{39,3}{24} - 0,268296 \cdot \frac{219}{24} = -0,8107 \end{aligned}$$

e a equação da reta é

$$\widehat{\text{Onda}} = -0,8107 + 0,268296 \cdot \text{Vento}$$

B.2 Capítulo 2

1. $\overline{X} \sim N\left(15; \frac{2,5^2}{18}\right)$

(a)

$$\begin{aligned} P(14,5 \le \overline{X} \le 16) &= P\left(\frac{14,5-15}{\sqrt{\frac{2,5^2}{18}}} \le Z \le \frac{16-15}{\sqrt{\frac{2,5^2}{18}}}\right) = P(-0,85 \le Z \le 1,70) \\ &\underbrace{=}_{\text{Tabela 1}} P(-0,85 \le Z \le 0) + P(0 < Z \le 1,70) = P(0 \le Z \le 0,85) + P(0 \le Z \le 1,70) \\ &= 0,3023 + 0,4554 = 0,7577 \\ &\underbrace{=}_{\text{Tabela 2}} \Phi(1,70) - \Phi(-0,85) = \Phi(1,70) - [1 - \Phi(0,85)] = 0,9554 - (1 - 0,8023) = 0,7577 \end{aligned}$$

(b)

$$\mathrm{P}(\overline{X} > 16,1) = \mathrm{P}\left(Z > \frac{16,1-15}{\sqrt{\frac{2,5^2}{18}}}\right) = \mathrm{P}(Z > 1,87) \underbrace{=}_{\text{Tabela 2}} 1 - \Phi(1,87) = 1 - 0,9693 = 0,0307$$
$$\underbrace{=}_{\text{Tabela 2}} 0,5 - \mathrm{P}(0 \leq Z \leq 1,87) = 0,5 - 0,4693 = 0,0307$$

2. Os erros são:

E_1 : estabelecer que são da máquina 1, quando na verdade foram produzidos pela máquina 2 ou E_2 : estabelecer que são da máquina 2, quando na verdade foram produzidos pela máquina 1. A regra de decisão é a seguinte:

$$\begin{aligned} \overline{X} &> 23 \Longrightarrow \text{ máquina 2} \\ \overline{X} &\leq 23 \Longrightarrow \text{ máquina 1} \end{aligned}$$

Na máquina 1 o comprimento é $N(20;16)$ e na máquina 2, $N(25;16)$.

$$\begin{aligned} \mathrm{P}(E_1) &= \mathrm{P}\left[\overline{X} \leq 23 \,|\, \overline{X} \sim N\left(25; \frac{16}{16}\right)\right] = \mathrm{P}\left(Z \leq \frac{23-25}{1}\right) = \mathrm{P}(Z \leq -2) \\ &\underbrace{=}_{\text{Tabela 2}} 1 - \Phi(2,00) = 1 - 0,9772 \underbrace{=}_{\text{Tabela 1}} 0,5 - \mathrm{P}(0 \leq Z \leq 2) = 0,5 - 0,4772 = 0,0228 \end{aligned}$$

$$\begin{aligned} \mathrm{P}(E_2) &= \mathrm{P}\left[\overline{X} > 23 \,|\, \overline{X} \sim N\left(20; \frac{16}{16}\right)\right] = \mathrm{P}\left(Z > \frac{23-20}{1}\right) = \mathrm{P}(Z > 3) \\ &\underbrace{=}_{\text{Tabela 2}} 1 - \Phi(3,00) = 1 - 0,9987 \underbrace{=}_{\text{Tabela 1}} 0,5 - \mathrm{P}(0 \leq Z \leq 3) = 0,5 - 0,4987 = 0,0013 \end{aligned}$$

3. Note que e é igual a $\overline{X}$ menos uma constante e sabemos que $E(\overline{X}) = \mu$ e $Var(\overline{X}) = \dfrac{\sigma^2}{n}$.

(a) Das propriedades da média e da variância, resulta que

$$\begin{aligned} \mathrm{E}(e) &= \mathrm{E}(\overline{X}) - \mu = \mu - \mu = 0 \\ \mathrm{Var}(e) &= \mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n} \end{aligned}$$

(b) $X \sim N(\mu; 20^2)$ e $n = 100$. Queremos

$$\begin{aligned} \mathrm{P}(|e| > 2) &= \mathrm{P}(e < -2) + \mathrm{P}(e > 2) = \mathrm{P}(\overline{X} - \mu < -2) + \mathrm{P}(\overline{X} - \mu > 2) \\ &= \mathrm{P}\left(\frac{\overline{X}-\mu}{\frac{20}{10}} < -\frac{2}{\frac{20}{10}}\right) + \mathrm{P}\left(\frac{\overline{X}-\mu}{\frac{20}{10}} > \frac{2}{\frac{20}{10}}\right) = \mathrm{P}(Z < -1) + \mathrm{P}(Z > 1) = 2 \times \mathrm{P}(Z > 1) \\ &\underbrace{=}_{\text{Tabela 1}} 2 \times [0,5 - \mathrm{P}(0 \leq Z \leq 1)] = 2 \times (0,5 - 0,3413) = 0,3174 \\ &\underbrace{=}_{\text{Tabela 2}} 2\,[1 - \Phi(1)] = 2\,(1,0 - 0,8413) = 0,3174 \end{aligned}$$

(c)

$$\begin{aligned} &\mathrm{P}(|e| > \delta) = 0,01 \Leftrightarrow \mathrm{P}(e < -\delta) + \mathrm{P}(e > \delta) = 0,01 \Leftrightarrow \mathrm{P}(\overline{X} - \mu < -\delta) + \mathrm{P}(\overline{X} - \mu > \delta) = 0,01 \Leftrightarrow \\ &\mathrm{P}\left(\frac{\overline{X}-\mu}{\frac{20}{10}} < -\frac{\delta}{\frac{20}{10}}\right) + \mathrm{P}\left(\frac{\overline{X}-\mu}{\frac{20}{10}} > \frac{\delta}{\frac{20}{10}}\right) = 0,01 \Leftrightarrow \mathrm{P}\left(Z < -\frac{\delta}{2}\right) + \mathrm{P}\left(Z > \frac{\delta}{2}\right) = 0,01 \Leftrightarrow \\ &2 \times \mathrm{P}\left(Z > \frac{\delta}{2}\right) = 0,01 \Leftrightarrow \mathrm{P}\left(Z > \frac{\delta}{2}\right) = 0,005 \Leftrightarrow 0,5 - \mathrm{P}\left(0 \leq Z \leq \frac{\delta}{2}\right) = 0,005 \\ \Leftrightarrow \quad &\underbrace{\mathrm{P}\left(0 \leq Z \leq \frac{\delta}{2}\right) = 0,495}_{\text{corpo da Tabela 1}} \Leftrightarrow \frac{\delta}{2} = 2,58 \Leftrightarrow \delta = 5,16 \end{aligned}$$

(d)

$$\mathrm{P}\left(|e|<1\right)=0,95 \Leftrightarrow \mathrm{P}(-1<\overline{X}-\mu<1)=0,95 \Leftrightarrow \mathrm{P}\left(-\frac{1}{\frac{20}{\sqrt{n}}}<Z<\frac{1}{\frac{20}{\sqrt{n}}}\right)=0,95 \Leftrightarrow$$
$$\mathrm{P}\left(-\frac{1}{\frac{20}{\sqrt{n}}}<Z<0\right)+\mathrm{P}\left(0\le Z<\frac{1}{\frac{20}{\sqrt{n}}}\right)=0,95 \Leftrightarrow 2\times\mathrm{P}\left(0\le Z<\frac{1}{\frac{20}{\sqrt{n}}}\right)=0,95 \Leftrightarrow$$
$$\underbrace{\mathrm{P}\left(0\le Z<\frac{1}{\frac{20}{\sqrt{n}}}\right)}_{\text{corpo da Tabela 1}}=0,475 \Leftrightarrow \frac{\sqrt{n}}{20}=1,96 \Leftrightarrow \sqrt{n}=39,2 \Leftrightarrow n\approx 1537$$

4. Parafusos pequenos: $X<8,5$, onde X é o comprimento do parafuso.

   (a) $X\sim N(\mu;1)$. Como $\mathrm{P}(X<8,5)=0,05$, resulta que 8,5 tem que ser menor que μ, ou seja, a abscissa $\frac{8,5-\mu}{1}$ tem que estar no lado negativo da escala da normal padronizada.

$$\mathrm{P}(X<8,5)=0,05 \Leftrightarrow \mathrm{P}\left(Z<\frac{8,5-\mu}{1}\right)=0,05 \Leftrightarrow \mathrm{P}\left(Z>-\frac{8,5-\mu}{1}\right)=0,05 \Leftrightarrow$$
$$\underbrace{\mathrm{P}(0\le Z\le\mu-8,5)=0,45}_{\text{Corpo da Tabela 1}} \Leftrightarrow \mu-8,5=1,64 \Leftrightarrow \mu=10,14$$

   (b) Parada desnecessária: amostra indica processo fora de controle ($\overline{X}<9$), quando, na verdade, o processo está sob controle ($\mu=10,14$).

$$\begin{aligned}\mathrm{P}\left[\overline{X}<9\,|\,\overline{X}\sim N\left(10,14;\frac{1}{4}\right)\right] &= \mathrm{P}\left(Z<\frac{9-10,14}{0,5}\right)=\mathrm{P}(Z<-2,28)=\mathrm{P}(Z>2,28)\\ &= 0,5-\underbrace{\mathrm{P}(0\le Z\le 2,28)}_{\text{Tabela 1}}=0,5-0,4887=0,0113\end{aligned}$$

   (c) Máquina desregulada: $\overline{X}>9$; processo operando sem ajuste: $X\sim N(9,5;1)$

$$\begin{aligned}\mathrm{P}\left[\overline{X}>9\,|\,\overline{X}\sim N\left(9,5;\frac{1}{4}\right)\right] &= \mathrm{P}\left(Z>\frac{9-9,5}{0,5}\right)=\mathrm{P}(Z>-1)=\mathrm{P}(-1<Z<0)+\mathrm{P}(Z\ge 0)\\ &= \underbrace{\mathrm{P}(0<Z<1)}_{\text{Tabela 1}}+\mathrm{P}(Z\ge 0)=0,3413+0,5=0,8413\end{aligned}$$

5. Afirmativa do gerente: $\mu=2$ e $\sigma=0,05$. Como $n=100$, podemos usar o teorema limite central . Logo, $\overline{X}\approx N\left(2;\frac{0,05^2}{100}\right)$.

$$\begin{aligned}\mathrm{P}(\overline{X}\le 1,985) &= \mathrm{P}\left(Z\le\frac{1,985-2}{\frac{0,05}{10}}\right)=\mathrm{P}(Z\le -3,0)=\mathrm{P}(Z\ge 3,0)=0,5-\underbrace{\mathrm{P}(0\le Z<3,0)}_{\text{Tabela 1}}\\ &= 0,5-0,4987=0,0013\end{aligned}$$

   A probabilidade de se obter esse valor nas condições dadas pelo gerente é muito pequena, o que pode nos fazer suspeitar da veracidade das afirmativas. É provável que, ou a média não seja 2 (e, sim, menor que 2), ou o desvio padrão não seja 0,05 (e, sim, maior que 0,05). Esboce gráficos da normal para compreender melhor esse comentário!

6. (a) $18\times 0,4=7,2>5$ $\qquad 18\times 0,6=10,8>5$ $\qquad X\approx N(7,2;4,32)$

$$\mathrm{P}(X\ge 15) \approx \mathrm{P}\left(Z\ge\frac{14,5-7,2}{\sqrt{4,32}}\right)=\mathrm{P}(Z\ge 3,51)=0,5-0,4998=0,0002$$
$$\mathrm{P}(X<2) \approx \mathrm{P}\left(Z\le\frac{1,5-7,2}{\sqrt{4,32}}\right)=\mathrm{P}(Z\le -2,74)=\mathrm{P}(Z\ge 2,74)=0,5-0,4969=0,0031$$

   (b) $40\times 0,3=12>5$ $\qquad 40\times 0,7=28>5$ $\qquad X\approx N(12;\,8,4)$

$$\mathrm{P}(X<10) \approx \mathrm{P}\left(Z\le\frac{9,5-12}{\sqrt{8,4}}\right)=\mathrm{P}(Z\le -0,86)=\mathrm{P}(Z\ge 0,86)=0,5-0,3051=0,1949$$
$$\mathrm{P}(25<X<28) \approx \mathrm{P}\left(\frac{25,5-12}{\sqrt{8,4}}\le Z\le\frac{27,5-12}{\sqrt{8,4}}\right)=\mathrm{P}(4,66\le Z\le 5,35)\approx 0$$

(c) $65 \times 0,9 = 58,5 > 5$ $\qquad 65 \times 0,1 = 6,5 > 5$ $\qquad X \approx N(58,5; 5,85)$

$$P(X = 58) \approx P\left(\frac{57,5 - 58,5}{\sqrt{5,85}} \leq Z \leq \frac{58,5 - 58,5}{\sqrt{5,85}}\right) = P(-0,41 \leq Z \leq 0) = P(0 \leq Z \leq 0,41) = 0,1591$$

$$P(60 < X \leq 63) \approx P\left(\frac{60,5 - 58,5}{\sqrt{5,85}} \leq Z \leq \frac{63,5 - 58,5}{\sqrt{5,85}}\right) = P(0,83 \leq Z \leq 2,07) = 0,4808 - 0,2967 = 0,1841$$

(d) $100 \times 0,2 = 20,0 > 5$ $\qquad 100 \times 0,8 = 80,0 > 5$ $\qquad X \approx N(20; 16)$

$$P(25 \leq X \leq 35) \approx P\left(\frac{24,5 - 20}{4} \leq Z \leq \frac{35,5 - 20}{4}\right) = P(1,13 \leq Z \leq 3,88) = 0,4999 - 0,3708 = 0,1291$$

(e) $50 \times 0,2 = 10,0 > 5$ $\qquad 50 \times 0,8 = 40,0 > 5$ $\qquad X \approx N(10; 8)$

$$P(X > 26) \approx P\left(Z \geq \frac{26,5 - 10}{\sqrt{8}}\right) = P(Z \geq 5,83) \approx 0$$

$$\begin{aligned} P(5 \leq X < 10) &\approx P\left(\frac{4,5 - 10}{\sqrt{8}} \leq Z \leq \frac{9,5 - 10}{\sqrt{8}}\right) = P(-1,94 \leq Z \leq -0,18) = P(0,18 \leq Z \leq 1,94) \\ &= 0,4738 - 0,0714 = 0,4024 \end{aligned}$$

(f) $np = 35 \qquad n(1-p) = 15 \qquad X \approx N(35; 10,5)$

$$P(X \leq 25) \approx P\left(Z \leq \frac{25,5 - 35}{\sqrt{10,5}}\right) = P(Z \leq -2,93) = 0,5 - 0,4983 = 0,0017$$

(g) $np = 50 \qquad n(1-p) = 50 \qquad X \approx N(50; 25)$

$$P(42 < X \leq 56) \approx P\left(\frac{42,5 - 50}{5} \leq Z \leq \frac{56,5 - 50}{5}\right) = P(-1,5 \leq Z \leq 1,3) = 0,4332 + 0,4032 = 0,8364$$

(h) $np = 50 \qquad n(1-p) = 50 \qquad X \approx N(50; 25)$

$$P(X > 60) \approx P\left(Z \geq \frac{60,5 - 50}{5}\right) = P(Z \geq 2,1) = 0,5 - 0,4821 = 0,0179$$

(i) $np = 8 \qquad n(1-p) = 12 \qquad X \approx N(8; 4,8)$

$$\begin{aligned} P(X = 5) &\approx P\left(\frac{4,5 - 8}{\sqrt{4,8}} \leq Z \leq \frac{5,5 - 8}{\sqrt{4,8}}\right) = P(-1,60 \leq Z \leq -1,14) = P(1,14 \leq Z \leq 1,60) \\ &= 0,44520 - 0,37286 = 0,07234 \end{aligned}$$

(j) $np = 9 \qquad n(1-p) = 21 \qquad X \approx N(9; 6,3)$

$$P(X \geq 12) \approx P\left(Z \geq \frac{11,5 - 9}{\sqrt{6,3}}\right) = P(Z \geq 1) = 0,5 - 0,3413 = 0,1587$$

(k) $np = 8 \qquad n(1-p) = 72 \qquad X \approx N(8; 7,2)$

$$P(9 < X < 11) \approx P\left(\frac{9,5 - 8}{\sqrt{7,2}} \leq Z \leq \frac{10,5 - 8}{\sqrt{7,2}}\right) = P(0,56 \leq Z \leq 0,93) = 0,3238 - 0,2123 = 0,1115$$

(l) $np = 6 \qquad n(1-p) = 24 \qquad X \approx N(8; 4,8)$

$$P(12 \leq X \leq 16) \approx P\left(\frac{11,5 - 8}{\sqrt{4,8}} \leq Z \leq \frac{16,5 - 8}{\sqrt{4,8}}\right) = P(1,60 \leq Z \leq 3,88) = 0,4999 - 0,4452 = 0,0547$$

(m) $np = 15 \qquad n(1-p) = 35 \qquad X \approx N(15; 10,5)$

$$P(X > 18) \approx P\left(Z \geq \frac{18,5 - 15}{\sqrt{10,5}}\right) = P(Z \geq 1,08) = 0,5 - 0,3599 = 0,1401$$

(n) $np = 5,6 \qquad n(1-p) = 22,4 \qquad X \approx N(5,6; 4,48)$

$$P(X = 6) \approx P\left(\frac{5,5 - 5,6}{\sqrt{4,48}} \leq Z \leq \frac{6,5 - 5,6}{\sqrt{4,48}}\right) = P(-0,05 \leq Z \leq 0,43) = 0,0199 + 0,1664 = 0,1863$$

(o) $np = 38 \qquad n(1-p) = 57 \qquad X \approx N(38; 22,8)$

$$P(30 \leq X < 48) \quad \approx \quad P\left(\frac{29,5-38}{\sqrt{22,8}} \leq Z \leq \frac{47,5-38}{\sqrt{22,8}}\right) = P(-1,78 \leq Z \leq 1,99) = 0,4767 + 0,4625 = 0,9392$$

7. X = "número de pessoas que votaram". Então $X \sim bin(1002; 0,61)$ e $X \approx N(611,22; 238,3758)$

$$P(X \geq 701) \approx P\left(Z \geq \frac{700,5 - 611,22}{\sqrt{238.3758)}}\right) = P(Z \geq 5,78) = 0$$

Se a proporção de votantes é de 61%, a probabilidade de encontrarmos 701 ou mais votantes em uma amostra aleatória simples de 1002 é muito baixa. Talvez as pessoas entrevistadas não estejam sendo sinceras, com vergonha de dizer que não votaram...

8. X = "número de meninas em 64 partos"; $X \sim bin(64; 0,5)$ e $X \approx N(32; 16)$

$$P(X \geq 36) \quad \approx \quad P\left(Z \geq \frac{35,5-32}{4}\right) = P(Z \geq 0,875) = 0,5 - 0,3106 = 0,1894$$

Esse é um resultado que pode ocorrer por mero acaso, ou seja, não é um resultado não-usual.

9. X = "número de passageiros que se apresentam para o voo em questão". $X \sim bin(400; 0,85)$ e $X \approx N(340; 51)$.

$$P(X > 350) \quad \approx \quad P\left(Z \geq \frac{350,5-340}{\sqrt{51}}\right) = P(Z \geq 1,47) = 0,5 - 0,42922 = 0,07078$$

Essa é uma probabilidade um pouco alta; talvez valha a pena a companhia rever a política de reservas e aceitar menos que 400 reservas.

10. X = "número de defeituosos na amostra"; $X \sim bin(20; 0,1)$. A taxa de falhas do processo é constante e igual a 0,1. Embora a amostragem seja feita sem reposição, podemos usar a aproximação binomial, uma vez que temos uma população razoavelmente grande e o tamanho da amostra é bem menor do que o tamanho da população. Note que aqui não podemos usar a aproximação normal, uma vez que $20 \times 0,1 = 2 < 5$. Queremos

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - [P(X=0) + P(X=1)] \\ &= 1 - \binom{20}{0}(0,1)^0(0,9)^{20} - \binom{20}{1}(0,1)(0,9)^{19} = \\ &= 1 - 0,39175 = 0,60825 \end{aligned}$$

B.3 Capítulo 3

1. $1 - \alpha = 0,90 \Longrightarrow z_{0,05} = 1,64$
   $1 - \alpha = 0,99 \Longrightarrow z_{0,005} = 2,58$
   $1 - \alpha = 0,80 \Longrightarrow z_{0,10} = 1,28$

2. $P(0 \leq Z \leq 1,28) = 0,3997 \Longrightarrow (0,5 - \alpha/2) = 0,3997 \Rightarrow \alpha/2 = 0,5 - 0,3997 = 0,1003 \Rightarrow\Rightarrow \alpha = 0,2006 \Rightarrow 1 - \alpha \approx 0,80$ ou 80%
   $P(0 \leq Z \leq 1,80) = 0,4641 \Longrightarrow (0,5 - \alpha/2) = 0,4641 \Rightarrow \alpha/2 = 0,5 - 0,4641 \Rightarrow \alpha/2 = 0,0359 \Rightarrow \alpha = 0,0718 \Rightarrow 1 - \alpha = 0,9282 \approx 0,93$ ou 93%

3. $1 - \alpha = 0,98 \Longrightarrow \alpha/2 = 0,01 - -z_{0,01} = 2,33$

$$\epsilon = 2,33 \times \frac{2}{\sqrt{36}} = 0,7767$$

Como a média amostral observada é $\overline{x} = \frac{1236}{36} = 34,333$, o intervalo de confiança é

$$[34,333 - 0,7767; 34,333 + 0,7767] = [33,556 \,;\; 35,110]$$

4. Como a amostra é a mesma, isso significa que a população é a mesma, bem como o tamanho de amostra, ou seja, σ e n são os mesmos. Vimos que um nível de confiança maior resulta em um intervalo de confiança maior; logo, o segundo intervalo foi construído com base em um nível de confiança maior do que o utilizado na construção do primeiro.

5. Mantidos fixos o nível de confiança e o desvio padrão populacional, vimos que a margem de erro é inversamente proporcional à raiz quadrada de n. Assim, para reduzir pela metade a margem de erro, temos que dobrar $\sqrt{n}$, ou seja, temos que quadruplicar o tamanho amostral n.

6. É dado que $X \sim N(\mu; 9)$. Como $n = 25$, sabemos que

$$\overline{X} \sim N\left(\mu; \frac{9}{25}\right)$$

Com $1-\alpha = 0,99$, temos que $\alpha = 0,01$ e $\alpha/2 = 0,005$. Assim, temos que procurar no corpo da tabela a abscissa correspondente ao valor $0,5 - 0,005 = 0,495$,o que nos dá $z_{0,005} = 2,58$. Então

$$\begin{aligned}
&P(-2,58 \leq Z \leq 2,58) = 0,99 \Rightarrow \\
&P\left(-2,58 \leq \frac{\overline{X} - \mu}{\sqrt{\frac{9}{25}}} \leq 2,58\right) = 0,99 \Rightarrow \\
&P\left(-2,58 \times \sqrt{\frac{9}{25}} \leq \overline{X} - \mu \leq 2,58 \times \sqrt{\frac{9}{25}}\right) = 0,99 \Rightarrow \\
&P(-1,548 \leq \overline{X} - \mu \leq 1,548) = 0,99 \Rightarrow \\
&P(\overline{X} - 1,548 \leq \mu \leq \overline{X} + 1,548) = 0,99
\end{aligned}$$

Como a média amostral obtida é $\overline{x} = \frac{60}{25} = 2,4$ o intervalo de confiança de 99% de confiança é

$$[2,4 - 1,548\,;\, 2,4 + 1,548] = [0,852\,;\, 3,948]$$

7. Queremos $|\epsilon| \leq 0,05$, com $\sigma = 4,2$ e $1 - \alpha = 0,95$.

$$1 - \alpha = 0,95 \Rightarrow z_{\alpha/2} = 1,96$$

Então

$$\begin{aligned}
1,96 \times \frac{4,2}{\sqrt{n}} &\leq 0,05 \Rightarrow \\
\sqrt{n} &\geq \frac{1,96 \times 4,2}{0,05} = 164,64 \Rightarrow \\
n &\geq 27106,3296
\end{aligned}$$

Logo, o tamanho mínimo necessário é $n = 27107$.

8. é dado que $X \sim N(\mu; 0,58^2)$. Como $n = 25$, sabemos que

$$\overline{X} \sim N\left(\mu; \frac{0,58^2}{25}\right)$$

Com $1 - \alpha = 0,90$, temos que $\alpha = 0,10$ e $\alpha/2 = 0,05$. Assim, temos que procurar no corpo da tabela a abscissa correspondente ao valor $0,5 - 0,05 = 0,45$,o que nos dá $z_{0,05} = 1,64$. Então

$$\begin{aligned}
&P(-1,64 \leq Z \leq 1,64) = 0,90 \Rightarrow \\
&P\left(-1,64 \leq \frac{\overline{X} - \mu}{\sqrt{\frac{0,58^2}{25}}} \leq 1,64\right) = 0,90 \Rightarrow \\
&P\left(-1,64 \times \frac{0,58}{5} \leq \overline{X} - \mu \leq 1,64 \times \frac{0,58}{5}\right) = 0,90 \Rightarrow \\
&P(-0,19024 \leq \overline{X} - \mu \leq 0,19024) = 0,90 \Rightarrow \\
&P(\overline{X} - 0,19024 \leq \mu \leq \overline{X} + 0,19024) = 0,90
\end{aligned}$$

Como a média amostral obtida é $\overline{x} = 2,8$ o intervalo de confiança de nível de confiança 99% é

$$[2,8 - 0,19024\,;\, 2,8 + 0,19024] = [2,60976\,;\, 2,99024]$$

9. $\alpha = 0,05 \Rightarrow 1 - \alpha = 0,95 \Rightarrow z_{0,025} = 1,96$

(a) A margem de erro é

$$\epsilon = 1,96 \times \frac{5}{\sqrt{50}} = 1,3859$$

Logo, o intervalo de confiança de nível de confiança 0,95 é

$$[42 - 1,\ 385\,9\,;\ 42 + 1,3859] = [40,\ 6141\,;\ 43,3859]$$

(b) Como visto em (a) a margem de erro é $\epsilon = 1,3859$.

(c) Temos que reduzir a margem de erro; logo, o tamanho da amostra terá que ser maior que 50.

$$\begin{aligned} \epsilon &= 1,96 \times \frac{5}{\sqrt{n}} \leq 1 \Rightarrow \\ \sqrt{n} &\geq 1,96 \times 5 = 9,8 \Rightarrow \\ n &\geq 9,8^2 = 96,04 \end{aligned}$$

Logo, n deve ser no mínimo igual a 97.

10. A média amostral é $\overline{x} = \frac{343120}{10} = 34312$.

(a) A margem de erro é

$$\epsilon = 1,96 \times \frac{500}{\sqrt{10}} = 309,9$$

Logo, o intervalo de confiança de nível de confiança 95% é

$$[34312 - 309,9\,;\ 34312 + 309,9] = [34002,1\,;\ 34621,9]$$

(b) A margem de erro é

$$\epsilon = 2,58 \times \frac{500}{\sqrt{10}} = 407,93$$

Logo, o intervalo de confiança de nível de confiança 95% é

$$[34312 - 407,93\,;\ 34312 + 407,93] = [33904,07\,;\ 34719,93]$$

(c) O gerente deve estar usando o nível de confiança de 99%).

11. (a) $\alpha = 2\% \Rightarrow 1 - \alpha = 98\% \Rightarrow z_{0,01} = 2,33$
$\widehat{p} = \frac{128}{600} = 0,213\,3$

$$\epsilon = 2,33 \times \sqrt{\frac{0,213\,3(1 - 0,2133)}{600}} = 0,03897$$

e o intervalo de confiança é

$$[0,2133 - 0,03897; 0,2133 + 0,03897] = [0,17433; 0,25227]$$

(b) $\alpha = 10\% \Rightarrow 1 - \alpha = 90\% \Rightarrow z_{0,05} = 1,64$
$\widehat{p} = \frac{710}{1200} = 0,59167 =$

$$\epsilon = 1,64 \times \sqrt{\frac{0,55 \times 0,45}{1200}} = 0,02355$$

e o intervalo de confiança é

$$[0,59167 - 0,02355; 0,59167 + 0,02355] = [0,56812; 0,61522]$$

12. O problema pede a estimativa para a proporção dos que não querem a fluoração; logo, $\widehat{p} = \frac{120}{300} = 0,4$

(a) $\alpha = 5\% \Rightarrow 1 - \alpha = 95\% \Rightarrow z_{0,025} = 1,96$

$$\epsilon = 1,96 \times \sqrt{\frac{0,4 \times 0,6}{300}} = 0,05544$$

e o intervalo de confiança é

$$[0,4 - 0,05544; 0,4 + 0,05544] = [0,34456; 0,045544]$$

(b) $1-\alpha = 96\% \Rightarrow z_{0,02} = 2,05$

$$\epsilon = 2,05 \times \sqrt{\frac{0,4 \times 0,6}{300}} = 0,05798$$

e o intervalo de confiança é

$$[0,4 - 0,05798; 0,4 + 0,05798] = [0,34202; 0,045798]$$

13. É dado que $n = 100$, $\widehat{p} = 0,32$ e $EP(\widehat{P}) = 0,03$.
$\alpha = 3\% \Rightarrow z_{0,015} = 2,17$

$$\epsilon = 2,17 \times 0,03 = 0,0651$$

$$[0,32 - 0,0651; 0,32 + 0,0651] = [0,2549; 0,3851]$$

14. $\widehat{p} = \frac{57}{150} = 0,38$. Para uma margem de erro de 0,08 e um nível de confiança de 90%, o tamanho da amostra teria que ser

$$n \geq \left(\frac{1,64}{0,08}\right)^2 \times 0,38 \times 0,62 = 99,011$$

Como o tamanho da amostra é 150, essa amostra é suficiente.

15. (a) $\widehat{p} = \frac{100}{400} = 0,25$

(b) $EP(\widehat{P}) = \sqrt{\frac{0,25 \times 0,75}{400}} = 0,02651$

(c) $1-\alpha = 0,80 \Rightarrow z_{0,1} = 1,28$

$$[0,25 - 1,28 \times 0,021651; 0,25 + 1,28 \times 0,021651] = [0,22229; 0,27771]$$

16. $\widehat{p}_0 = 0,35$

$$n \geq \left(\frac{1,96}{0,05}\right)^2 \times 0,35 \times 0,65 = 349,59$$

Logo, $n \geq 350$

B.4 Capítulo 4

1. Na linha correspondente a 17 graus de liberdade, devem ser consultadas as seguintes colunas:

(a) $\alpha = 0,02 \Longrightarrow k = 30,992$

(b) $\alpha = 0,98 \Longrightarrow k = 7,255$

(c) $\alpha = 0,1 \Longrightarrow k = 24,769$

2. Temos que usar a Tabela 2, concentrando-nos na linha correspondente a 12 graus de liberdade. Os valores dados podem ser encontrados no corpo da tabela nesta linha.

(a) à direita de $1,782$ temos uma área de $0,05$; logo, à esquerda de $1,782$ a área é de $0,95$.

(b) A área abaixo de $-1,356$ é igual à área acima de $1,356$, que é de $0,10$. Logo, à esquerda de $-1,356$ temos uma área de $0,10$ e à direita de $-1,356$ temos uma área de $0,90$.

(c) à direita de $2,681$ a área é $0,01$.

(d) à direita de $1,083$ a área é $0,15$; à direita de $3,055$ a área é de $0,005$. Logo, a área entre $1,083$ e $3,055$ é $0,15 - 0,005 = 0,145$.

(e) Como visto no item (b), a área à direita de $-1,356$ é $0,90$. A área à direita de $2,179$ é $0,025$. Logo, a área entre $-1,356$ e $2,179$ é $0,90 - 0,025 = 0,875$

3. (a) $t_{15;0,05} = 1,753$

(b) O primeiro fato a observar é que $t_{18;0,90}$ tem que ser negativo, pois à direita dele a área é de $0,90 > 0,50$. Se à direita a área é 0,90, a área à esquerda é 0,10. Pela simetria da curva, $t_{18;0,90} = -t_{18;0,10}$. Resulta que

$$t_{18;0,90} = -t_{18;0,10} = -1,33$$

(c) Analogamente encontra-se que $t_{25;0,975} = -2,060$

4. Contexto: População normal e amostra pequena; distribuição envolvida: t-Student
$n = 15 \qquad 1-\alpha = 0,95 \Rightarrow t_{14;0,025} = 2,145$

$$\begin{aligned}\overline{x} &= \frac{808}{15} = 53,8667 \\ s^2 &= \frac{1}{14}\left[44176 - \frac{808^2}{15}\right] = 46,5524\end{aligned}$$

$$\epsilon = 2,145 \times \sqrt{\frac{46,5524}{15}} = 3,7788$$

O intervalo de confiança é

$$[53,8667 - 3,7788; 53,8667 + 3,7788] = [50,088; 57,6455]$$

5. Contexto: População normal e amostra pequena; distribuição envolvida: t-Student
$t_{24;0,01} = 2,492$

$$\left[500 - 2,492 \times \sqrt{\frac{900}{25}}; 500 + 2,492 \times \sqrt{\frac{900}{25}}\right] = [485,05; 514,95]$$

6. Contexto: População normal e amostra pequena; distribuição envolvida: t-Student
$\alpha = 2\% \Rightarrow t_{29;0,01} = 2,462$

$$\begin{aligned}\overline{x} &= \frac{401}{30} = 13,367 \\ s^2 &= \frac{1}{29}\left[5443 - \frac{401^2}{30}\right] = 2,861\end{aligned}$$

O intervalo de confiança é

$$\left[13,367 - 2,462 \times \sqrt{\frac{2,861}{30}}; 13,367 + 2,462 \times \sqrt{\frac{2,861}{30}}\right] = [12,607; 14,127]$$

7. Como n é grande, podemos usar a abscissa da distribuição normal $z_{0,01} = 2,33$ (o valor exato é $t_{99;0,01} = 2,3646$),

$$\left[13,78 - 2,33 \times \sqrt{\frac{2,865}{100}}; 13,78 + 2,33 \times \sqrt{\frac{2,865}{100}}\right] = [13,386; 14,174]$$

B.5 Capítulo 5

1. (a) Antes da pane: $T \sim N(100; 100)$
Depois da pane: $T \sim N(\mu; 100)$ - afirmativa dada: $\mu \neq 100$

$$\begin{aligned}H_0 &: \mu = 100 \\ H_1 &: \mu \neq 100\end{aligned}$$

(b) Afirmativa dada: $\mu < 900$

$$\begin{aligned}H_0 &: \mu = 900 \\ H_1 &: \mu < 900\end{aligned}$$

(c) Afirmativa dada: $\mu \geq 2$

$$\begin{aligned}H_0 &: \mu = 2 \\ H_1 &: \mu < 2\end{aligned}$$

2. $$\left.\begin{array}{l} X \sim N(\mu; 225) \\ n = 25 \end{array}\right\} \Rightarrow \overline{X} \sim N\left(\mu; \frac{225}{25}\right) \text{ ou } \overline{X} \sim N(\mu; 9)$$

(a)

$$\alpha = \mathrm{P}(\overline{X} > 43 \,|\, \overline{X} \sim N(40; 9)) = \mathrm{P}\left(Z > \frac{43-40}{3}\right) = \mathrm{P}(Z > 1,0) = 0,5 - 1,0) = 0,1587$$

$$\beta = \mathrm{P}(\overline{X} \leq 43 \,|\, \overline{X} \sim N(45; 9) = \mathrm{P}\left(Z \leq \frac{43-45}{3}\right) = \mathrm{P}(Z \leq -0,67) = 0,5 - \mathrm{P}(0 \leq Z \leq 0,67) = 0,2514$$

(b)

$$\alpha = 0,10 \Leftrightarrow \mathrm{P}\left[\overline{X} > k \,|\, \overline{X} \sim N(40; 9)\right] = 0,10 \Leftrightarrow \mathrm{P}\left(Z > \frac{k-40}{3}\right) = 0,10 \Leftrightarrow$$
$$tab\left(\frac{k-40}{3}\right) = 0,40 \Leftrightarrow \frac{k-40}{3} = 1,28 \Leftrightarrow k = 43,84$$

$$\begin{aligned}\beta &= \mathrm{P}(\overline{X} \leq 43,84 \,|\, \overline{X} \sim N(45; 9) = \mathrm{P}\left(Z \leq \frac{43.84-45}{3}\right) = \mathrm{P}(Z \leq -0,39) \\ &= 0,5 - \mathrm{P}(0 \leq Z \leq 0,39) = 0,34827\end{aligned}$$

3. $\left.\begin{array}{l} X \sim N(\mu; 225) \\ n = 25 \end{array}\right\} \Rightarrow \overline{X} \sim N\left(\mu; \frac{225}{25}\right)$ ou $\overline{X} \sim N(\mu; 9)$

(a)

$$\begin{aligned}\alpha &= \mathrm{P}\left[\overline{X} < 34 \,|\, \overline{X} \sim N(40; 9)\right] + \mathrm{P}\left[\overline{X} > 46 \,|\, \overline{X} \sim N(40; 9)\right] \\ &= \mathrm{P}\left(Z < \frac{34-40}{3}\right) + \mathrm{P}\left(Z > \frac{46-40}{3}\right) \\ &= \mathrm{P}(Z < -2) + \mathrm{P}(Z > 2) = 2 \times \Pr(Z > 2) = 2 \times [0,5 - \mathrm{P}(0 \leq Z \leq 2,)] = 0,0455\end{aligned}$$

(b)

$$\begin{aligned}\beta &= \mathrm{P}(\text{não rejeitar } H_0 \,|\, \mu = 36) = \mathrm{P}\left[34 \leq \overline{X} \leq 46 \,|\, \overline{X} \sim N(36; 9)\right] \\ &= \mathrm{P}\left(\frac{34-36}{3} \leq Z \leq \frac{46-36}{3}\right) = \mathrm{P}(-0,67 \leq Z \leq 3,33) = \mathrm{P}(0 \leq Z \leq 3,33) + \mathrm{P}(0 \leq Z \leq 0,67) = 0,745\end{aligned}$$

4. $\left.\begin{array}{l} X \sim N(\mu; 64) \\ n = 16 \end{array}\right\} \Rightarrow \overline{X} \sim N\left(\mu; \frac{64}{16}\right)$ ou $\overline{X} \sim N(\mu; 4)$

(a)

$$\begin{aligned}\alpha &= \mathrm{P}\left[\overline{X} > 25,5 \,|\, \overline{X} \sim N(23; 4)\right] = \mathrm{P}\left(Z > \frac{25,5-23}{2}\right) = \mathrm{P}(Z > 1,25) \\ &= 0,5 - \mathrm{P}(0 \leq Z \leq 1,25) = 0,1056\end{aligned}$$

$$\beta = \mathrm{P}(\overline{X} \leq 25,5 \,|\, \overline{X} \sim N(28; 4) = \mathrm{P}\left(Z \leq \frac{25,5-28}{2}\right) = \mathrm{P}(Z \leq -1,25) = \mathrm{P}(Z > 1,25) = 0,1056$$

(b)

$$\alpha = 0,05 \Leftrightarrow \mathrm{P}\left[\overline{X} > k \,|\, \overline{X} \sim N(23; 4)\right] = 0,05 \Leftrightarrow \mathrm{P}\left(Z > \frac{k-23}{2}\right) = 0,05 \Leftrightarrow$$
$$tab\left(\frac{k-23}{2}\right) = 0,45 \Leftrightarrow \frac{k-23}{2} = 1,64 \Leftrightarrow k = 26,28$$

$$\begin{aligned}\beta &= \mathrm{P}(\overline{X} \leq 26,28 \,|\, \overline{X} \sim N(28; 4) = \mathrm{P}\left(Z \leq \frac{26,28-28}{2}\right) = \mathrm{P}(Z \leq -0,86) \\ &= \mathrm{P}(Z \geq 0,86) = 0,5 - \mathrm{P}(0 \leq Z \leq 0,86) = 0,19489\end{aligned}$$

5. $\left.\begin{array}{l} X \sim N(\mu; 36) \\ n = 16 \end{array}\right\} \Rightarrow \overline{X} \sim N\left(\mu; \frac{36}{16}\right)$ ou $\overline{X} \sim N\left(\mu; 1,5^2\right)$

   (a)

$$\begin{aligned} \alpha &= \mathrm{P}\left[\overline{X} < 41,25 \,|\, \overline{X} \sim N(45; 1,5^2)\right] = \mathrm{P}\left(Z < \frac{41,25-45}{1,5}\right) = \mathrm{P}(Z < -2,5) \\ &= 0,5 - \mathrm{P}(0 \leq Z \leq 2,5) = 0,0062 \end{aligned}$$

   (b)

$$\begin{aligned} \beta &= \mathrm{P}(\text{não rejeitar } H_0 \,|\, \mu = 43) = \mathrm{P}\left[\overline{X} \geq 41,25 \,|\, \overline{X} \sim N(43; 1,5^2)\right] \\ &= \mathrm{P}\left(Z \geq \frac{41,25-43}{1,5}\right) = \mathrm{P}(Z \geq -1,17) = 0,5 + \mathrm{P}(0 \leq Z \leq 1,17) = 0,8790 \end{aligned}$$

B.6 Capítulo 6

1. (a) $X \sim N(\mu; 3,1^2)$ $\quad n = 9 \quad$ $\overline{x} = 13,35$
   $\alpha = 0,02 \Longrightarrow z_{0,01} = 2,33$
   Estatística de teste: $Z_0 = \dfrac{\overline{X} - 12,8}{\frac{3,1}{3}} \sim N(0;1)$ sob H_0.

$$RC : Z_0 < -2,33 \qquad \text{ou} \qquad Z_0 > 2,33$$

   (b) O valor observado da estatística de teste é

$$z_0 = \frac{13,35 - 12,8}{\frac{3,1}{3}} = 0,532$$

   que não pertence à região crítica. Logo, não há evidência amostral suficiente para rejeitarmos a hipótese de que a média da população seja 12,8.

   (c)

$$P = 2 \times \mathrm{P}(Z > 0,532) = 2 \times [0,5 - \mathrm{P}(0 \leq Z \leq 0,53)] = 0,5962$$

   O valor P é bastante alto; logo a hipótese nula só seria rejeitada para níveis de significância maiores que 0,5962. Isso é evidência de que não se pode rejeitar a hipótese nula em qualquer nível de significância razoável.

2. O problema na produção surge quando $\mu < 15$. Logo, nossas hipóteses são:

$$\begin{aligned} H_0 &: \mu = 15 \\ H_1 &: \mu < 15 \end{aligned}$$

   Estatística de teste: $Z_0 = \dfrac{\overline{X} - 15}{\frac{0,5}{3}} \sim N(0,1)$ sob H_0.

$$RC : Z_0 < -z_{0,001} \qquad \text{ou} \qquad Z_0 < -3,09$$

   Escrevendo a região crítica em termos da média amostral temos

$$\frac{\overline{X} - 15}{\frac{0,5}{3}} < -3,09 \Rightarrow \overline{X} < 15 - 3,09 \times \frac{0,5}{3} \Rightarrow \overline{X} < 14,485$$

   Então se $\overline{X} < 14,485$ o processo deve ser interrompido para um novo ajuste.

3. A intenção do analista é reduzir o tempo; logo, o interesse dele é que $\mu < 48,5$. A negação dessa afirmativa é $\mu \geq 48,5$. Logo, nossas hipóteses são:

$$\begin{aligned} H_0 &: \mu = 48,5 \\ H_1 &: \mu < 48,5 \end{aligned}$$

A estatística de teste é

$$Z_0 = \frac{\overline{X} - 48,5}{\frac{5}{5}} \sim N(0,1) \text{ sob } H_0.$$

e o valor observado é $z_0 = 46,5 - 48,5 = -2$, que resulta no seguinte valor P :

$$P = \mathrm{P}(Z < -2,0) = \Pr(Z > 2,0) = 0,5 - \mathrm{P}(0 \leq Z \leq 2,0) = 0,02275$$

Podemos afirmar que o tempo de execução reduziu, a qualquer nível de significância inferior 2,275%. Note que rejeitamos a hipótese nula ao nível de significância de 5%, mas não a 1%!

4. Se o consumo for menor ou igual a 12 litros por 100 km, não há problema com a propaganda. O problema surge se o consumo for superior. Logo, nossas hipóteses são:

$$\begin{aligned} H_0 &: \mu = 12 \\ H_1 &: \mu > 12 \end{aligned}$$

Supondo que o consumo X possa ser aproximado por uma distribuição normal, nossa estatistica de teste é

$$Z_0 = \frac{\overline{X} - 12}{\frac{1}{6}} \sim N(0,1) \text{ sob } H_0.$$

O valor observado é

$$z_0 = \frac{12,4 - 12}{\frac{1}{6}} = 2,4$$

e o valor P é

$$P = \mathrm{P}(Z > 2,4) = 0,5 - \mathrm{P}(0 \leq Z \leq 2,4) = 0,0082$$

A propaganda parece ser enganosa, pois a probabilidade de se obter um consumo médio de 12,4 litros por 100 km é pequena se o consumo realmente for de 12 litros por 100 km. Note que H_0 é rejeitada para qualquer nível de significância $\alpha \geq 0,82\%$, o que inclui os níveis de significância usuais de 1% e 5%.

5. (a)

$$\begin{aligned}
\alpha &= \mathrm{P}(\text{rejeitar } H_0 \,|\, H_0 \text{ verdadeira}) \\
&= \mathrm{P}\left[\left\{\overline{X} < 34\right\} \cup \left\{\overline{X} > 46\right\} \mid \overline{X} \sim N\left(40; \frac{225}{25}\right)\right] \\
&= \mathrm{P}\left[\overline{X} < 34 \,|\, \overline{X} \sim N(40;9)\right] + \mathrm{P}\left[\overline{X} > 46 \,|\, \overline{X} \sim N(40;9)\right] \\
&= \mathrm{P}\left(Z < \frac{34-40}{3}\right) + \mathrm{P}\left(Z > \frac{46-40}{3}\right) \\
&= \Pr(Z < -2) + \Pr(Z > 2) \\
&= 2 \times \Pr(Z > 2) \\
&= 0,0456
\end{aligned}$$

(b) A função poder é dada por

$$\begin{aligned}
\pi(\mu) &= 1 - \mathrm{P}(\text{não rejeitar } H_0 \,|\, \mu) \\
&= 1 - \mathrm{P}(34 \leq \overline{X} \leq 46 \,|\, \mu) \\
&= 1 - \Pr\left[34 \leq \overline{X} \leq 46 \,|\, \overline{X} \sim N(\mu;9)\right] \\
&= 1 - \mathrm{P}\left(\frac{34-\mu}{3} \leq Z \leq \frac{46-\mu}{3}\right) \\
&= 1 - \Phi\left(\frac{46-\mu}{3}\right) + \Phi\left(\frac{34-\mu}{3}\right)
\end{aligned}$$

(c) Note que $\mu = 20, 22, \cdots, 58, 60$ corresponde a $\Delta = -20, -18, \cdots, 18, 20$.

Δ	$\pi(\Delta)$	Δ	$\pi(\Delta)$
-20	0,97730	2	0,06312
-18	0,94530	4	0,11781
-16	0,88511	6	0,21281
-14	0,78841	8	0,34508
-12	0,65577	10	0,50040
-10	0,50040	12	0,65577
-8	0,34508	14	0,78841
-6	0,21281	16	0,88511
-4	0,11781	18	0,94530
-2	0,06312	20	0,97730
0	0,04560		

Observe que, para $\Delta = 0$, valor da hipótese nula, a função poder é igual à probabilidade do erro tipo I (nível de significância).

é interessante notar também que quanto mais distante do valor $\mu_0 = 100$, ou equivalentemente, quanto maior $|\delta|$, maior o poder do teste, ou seja, há uma probabilidade mais alta de se rejeitar H_0 quando o valor alternativo μ está bem distante de μ_0.

(d) Veja a Figura B.1 a seguir.

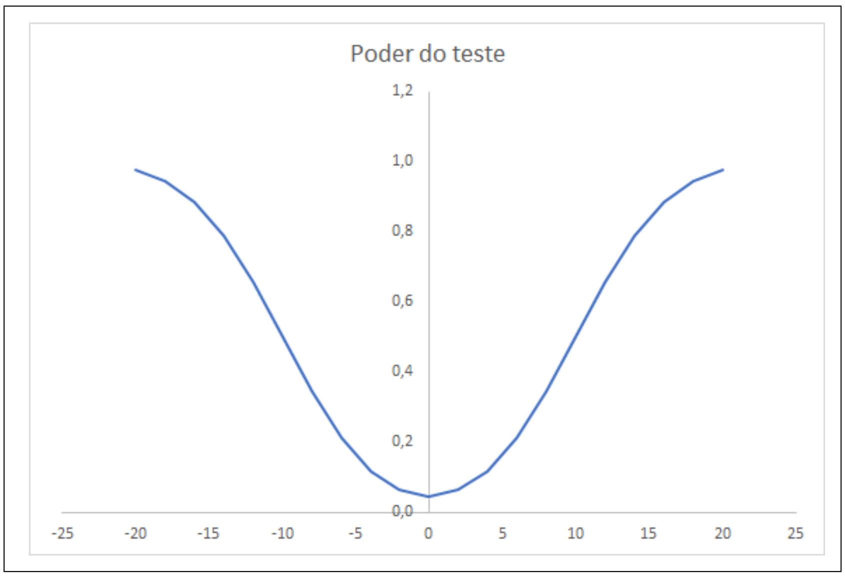


Figura B.1 – Poder do teste para o Exercício 5c

6. $\widehat{p} = \frac{385}{800} = 0,48125$
A afirmativa de interesse é “pelo menos 50% dos estudantes possuem computador”, ou seja, $p \geq 0,5$. Logo, as hipóteses são

$$\begin{array}{lcl} H_0 & : & p = 0,50 \\ H_1 & : & p < 0,50 \end{array}$$

$$\alpha = 0,10 \Longrightarrow z_{0,1} = 1,28$$

A estatística de teste é

$$Z_0 = \frac{\hat{P} - 0,5}{\frac{0,5}{\sqrt{800}}} \approx N(0,1) \text{ sob } H_0$$

e a região crítica é

$$Z_0 < -1,28$$

O valor observado da estatística de teste é

$$z_0 = \frac{0,48125 - 0,5}{\frac{0,5}{\sqrt{800}}} = -1,0607$$

Como o valor observado não pertence à região crítica, não podemos rejeitar a hipótese nula. Ou seja, os dados trazem evidência de que a proporção de estudantes que possuem computador é de pelo menos 50%.

7. A afirmativa de interesse é "mais de 10% dos trabalhadores conseguem seus empregos por indicação de amigos ou parentes", ou seja, $p > 0,10$, cuja negativa é $p \leq 0,10$. Logo, as hipóteses são

$$\begin{aligned} H_0 &: \quad p = 0,10 \\ H_1 &: \quad p > 0,10 \end{aligned}$$

Com $\alpha = 5\%$ e um teste unilateral, $z_{0,05} = 1,64$.
A estatística de teste é

$$Z_0 = \frac{\hat{P} - 0,1}{\sqrt{\frac{0,1 \times 0,9}{700}}} \approx N(0,1) \text{ sob } H_0$$

e a região crítica é

$$Z_0 > 1,64$$

O valor observado da estatística de teste é

$$z_0 = \frac{0,123 - 0,1}{\sqrt{\frac{0,1 \times 0,9}{700}}} = 2,0284$$

Como o valor observado da estatística de teste pertence à região crítica, rejeita-se a hipótese nula, ou seja, os dados dâo evidência de que mais 10% dos trabalhadores conseguem seus empregos por indicação de parentes ou amigos.

8. O interesse é verificar se $p > 0,20$. Logo,

$$\begin{aligned} H_0 &: \quad p = 0,20 \\ H_1 &: \quad p > 0,20 \end{aligned}$$

A estatística de teste é

$$Z_0 = \frac{\hat{P} - 0,2}{\sqrt{\frac{0,2 \times 0,8}{64}}} \approx N(0,1) \text{ sob } H_0$$

Como $\alpha = 5\%$ e o teste é unilateral, resulta que $z_{0,05} = 1,64$ e a região crítica é

$$Z_0 > 1,64$$

O valor observado da estatística de teste é

$$z_0 = \frac{\frac{25}{64} - 0,20}{\sqrt{\frac{0,2 \times 0,8}{64}}} = 3,8125$$

que está na região crítica; logo, rejeita-se a hipótese nula, ou seja, as evidências amostrais indicam que houve melhora com as mudanças.

9. As hipóteses são

$$\begin{aligned} H_0 &: \quad p = 0,5 \\ H_1 &: \quad p \neq 0,5 \end{aligned}$$

e a estatística de teste é

$$Z_0 = \frac{\hat{P} - 0,5}{\sqrt{\frac{0,5 \times 0,5}{200}}} \approx N(0,1) \text{ sob } H_0$$

O valor observado da estatística de teste é

$$z_0 = \frac{\frac{115}{200} - 0,5}{\sqrt{\frac{0,5 \times 0,5}{200}}} = 2,1213 \approx 2,12$$

e o valor P para o teste bilateral é

$$P = 2 \times \mathrm{P}(Z_0 > 2,12) = 2 \times (0,5 - \mathrm{P}(0 \leq Z \leq 2,12)) = 0,034$$

Como o valor P é pequeno, a probabilidade de obtermos 115 caras em 200 lançamentos de uma moeda honesta é pequena, o que nos leva a suspeitar da honestidade da moeda. A hipótese nula seria rejeitada para qualquer nível de significância $\alpha \geq 3,4\%$. Isso inclui $\alpha = 5\%$, mas não $\alpha = 1\%$.

10. Com as informações disponíveis, nossas hipóteses são:

$$\begin{aligned} H_0 &: p = 0,25 \\ H_1 &: p \neq 0,25 \end{aligned}$$

e a estatística de teste é

$$Z_0 = \frac{\widehat{P} - 0,25}{\sqrt{\frac{0,25 \times 0,75}{740}}} \approx N(0,1) \text{ sob } H_0$$

O valor observado da estatística de teste é

$$z_0 = \frac{\frac{156}{740} - 0,25}{\sqrt{\frac{0,25 \times 0,75}{740}}} = -2,46$$

e o valor P para o teste bilateral é

$$P = 2 \times \mathrm{P}(Z_0 > 2,46) = 2 \times (0,5 - \mathrm{P}(0 \leq Z \leq 2,46)) = 0,0139$$

Como o valor P é bastante pequeno, devemos rejeitar a hipótese nula de que a proporção de leitores da classe A é igual a 25%.

B.7 Capítulo 7

1. A afirmativa do fabricante é $\mu \geq 2$. Logo, a negação de tal afirmação é $\mu < 2$. Como essa última expressão não contém o sinal de igualdade, ela se torna a hipótese alternativa. Então, nossas hipóteses são:

$$\begin{aligned} H_0 &: \mu = 2 \\ H_1 &: \mu < 2 \end{aligned}$$

$X \sim N(\mu, \sigma^2)$ $\quad n = 25$ $\quad \overline{x} = 1,98$ $\quad s = 0,5$
A estatística de teste é

$$T_0 = \frac{\overline{X} - 2}{\frac{0,5}{5}} \sim t_{24} \text{ sob } H_0$$

$n = 25; \alpha = 0,05 \Longrightarrow t_{24;\,0,05} = 1,711$. Logo, a região crítica é

$$T_0 < -1,711$$

O valor observado da estatística de teste é

$$t_0 = \frac{1,98 - 2,0}{\frac{0,5}{5}} = -0,2$$

que não pertence à região crítica; logo, não podemos rejeitar H_0, ou seja, as evidências amostrais indicam que as balas pesam pelo menos 2 gramas.

2. O problema na produção surge quando $\mu < 15$. Logo, nossas hipóteses são:

$$\begin{aligned} H_0 &: \mu = 15 \\ H_1 &: \mu < 15 \end{aligned}$$

$X \sim N(\mu, \sigma^2)$ $\quad n = 9$ $\quad \overline{x} = 14,5$ $\quad s = 0,5$
A estatística de teste é

$$T_0 = \frac{\overline{X} - 15}{\frac{0,5}{3}} \sim t_8 \text{ sob } H_0$$

$n = 9, \alpha = 0,001 \Longrightarrow t_{8;\,0,001} = 4,501$. A região crítica é

$$T_0 < -4,501$$

O valor observado da estatística de teste é

$$t_0 = \frac{14,5 - 15}{\frac{0,5}{3}} = -3,0$$

que não está na região crítica. Logo, não podemos rejeitar H_0, ou seja, as evidências amostrais indicam que o processo está operando adequadamente.

3. A intenção do analista é reduzir o tempo; logo, o interesse dele é que $\mu < 48,5$. A negação dessa afirmativa é $\mu \geq 48,5$. Logo, nossas hipóteses são:

$$\begin{aligned} H_0 &: \mu = 48,5 \\ H_1 &: \mu < 48,5 \end{aligned}$$

$X \sim N(\mu, \sigma^2) \qquad n = 25 \qquad \overline{x} = 46,5 \qquad s = 5$
A estatística de teste é

$$T_0 = \frac{\overline{X} - 48,5}{\frac{5}{5}} \sim t_{24} \text{ sob } H_0$$

$n = 25, \alpha = 0,05 \Longrightarrow t_{24;\,0,05} = 1,711$. Logo, a região crítica é

$$T_0 < -1,711$$

O valor observado desta estatística é

$$t_0 = \frac{46,5 - 48,5}{\frac{5}{5}} = -2,0$$

Como o valor observado $t_0 = -2,0$ pertence à região crítica, devemos rejeitar H_0, ou seja, as evidências amostrais indicam que o analista foi bem-sucedido em reduzir o tempo de execução.

4. Se o consumo for menor ou igual a 12 litros por 100 km, não há problema com a propaganda. O problema surge se o consumo for superior. Logo, nossas hipóteses são:

$$\begin{aligned} H_0 &: \mu = 12 \\ H_1 &: \mu > 12 \end{aligned}$$

Supondo que o consumo X possa ser aproximado por uma distribuição normal, temos:
$X \approx N(\mu, \sigma^2) \qquad n = 36 \qquad \overline{x} = 12,4 \qquad s = 1$
A estatística de teste é

$$T_0 = \frac{\overline{X} - 12}{\frac{1}{6}} \sim t_{35} \text{ sob } H_0$$

$n = 36, \alpha = 10\% \Rightarrow t_{35;\,0,10} = 1,306$ e a região crítica é

$$T_0 > 1,306$$

O valor observado da estatística de teste é

$$t_0 = \frac{12,4 - 12}{\frac{1}{6}} = 2,4$$

Como o valor observado $t_0 = 2,4$ está na região crítica, devemos rejeitar H_0, ou seja, a propaganda parece ser enganosa.

5. (a) $X \sim N(\mu, \sigma^2) \qquad n = 9 \qquad \overline{x} = 13,35 \qquad s = 3,1$
   A estatística de teste é

$$T_0 = \frac{\overline{X} - 12,8}{\frac{3,1}{3}} \sim t_8 \text{ sob } H_0$$

   $n = 9, \alpha = 0,02 \Rightarrow t_{8;\,0,01} = 2,896$. Logo, a região crítica é

$$T_0 > +2,896 \text{ ou } T_0 < -2,896$$

(b) O valor observado da estatística de teste é

$$t_0 = \frac{13,35 - 12,8}{\frac{3,1}{3}} = 0,53226$$

que não pertence à região crítica; logo, não podemos rejeitar H_0.

6. A afirmativa é que temos amostra de uma população normal com $\sigma^2 = 36,8$, cuja negativa é $\sigma^2 \neq 36,8$. Logo, nossas hipóteses são

$$\begin{array}{lll} H_0 & : & \sigma^2 = 36,8 \\ H_1 & : & \sigma^2 \neq 36,8 \end{array}$$

A estatística de teste é

$$\chi_0^2 = \frac{15\,S^2}{36,8} \sim \chi^2_{15} \text{ sob } H_0$$

.

$$n = 16 \qquad \alpha = 0,05 \Rightarrow \begin{cases} \chi^2_{15;0,975} = 6,262 \\ \chi^2_{15;0,025} = 27,488 \end{cases}$$

A região crítica é

$$\chi_0^2 < 6,232 \qquad \text{ou} \qquad \chi_0^2 > 27,488$$

.

Para esses dados, temos que

$$\sum_{i=1}^{16} x_i = 3703,9 \qquad\qquad \sum_{i=1}^{16} x_i^2 = 859017,65$$

Logo, a variância amostral é

$$s^2 = \frac{1}{15}\left(859017,65 - \frac{3703,9^2}{16}\right) = 105,8632917$$

e o valor observado da estatística de teste é

$$\chi_0^2 = \frac{15 \times 105,8632917}{36,8} = 43,15$$

que pertence à região crítica. Logo, os dados dão evidência de que a variância é diferente de 36,8.

7. (a) Supondo normalidade, nossas hipóteses são

$$\begin{array}{lll} H_0 & : & \sigma^2 = 1,56 \\ H_1 & : & \sigma^2 \neq 1,56 \end{array}$$

A estatística de teste é

$$\chi_0^2 = \frac{29\,S^2}{1,56} \sim \chi^2_{29} \text{ sob } H_0$$

.

$$n = 30 \qquad \alpha = 0,05 \Rightarrow \begin{cases} \chi^2_{29;0,975} = 16,047 \\ \chi^2_{15;0,025} = 45,722 \end{cases}$$

A região crítica é

$$\chi_0^2 < 16,047 \qquad \text{ou} \qquad \chi_0^2 > 45,722$$

.

O valor observado da estatística de teste é

$$\chi_0^2 = \frac{29 \times 2,2}{1,56} = 40,897$$

que não pertence à região crítica. Logo, os dados dão evidência de que não houve alteração variância da taxa de câmbio.

(b) Olhando na linha correspondente a 29 graus de liberdade, vemos que $39,087 < 40,897 < 42,557$. A abscissa 39,087 deixa probabilidade 0,1 acima dela e a abscissa 42,557 deixa probabilidade 0,05 acima dela, Logo, a probabilidade acima do valor observado 40,897 está entre 0,05 e 0,1. Como o teste é bilateral, o valor P está entre 0,1 e 0,2.

8. Temos afirmativas sobre a média e sobre a variância do tempo de execução da tarefa. O objetivo do treinamento é reduzir tanto o tempo médio, quanto a variabilidade (variância). Logo, temos dois testes a realizar.

Teste sobre a média:

$$\begin{aligned} H_0 &: \mu = 40 \\ H_1 &: \mu < 40 \end{aligned}$$

Supondo que o tempo de execução X possa ser aproximado por uma distribuição normal, temos:

$X \approx N(\mu, \sigma^2) \qquad n = 14 \qquad \overline{x} = 35,6 \qquad s = 3,4$

A estatística de teste é

$$T_0 = \frac{\overline{X} - 40}{\frac{3,4}{\sqrt{14}}} \sim t_{13} \text{ sob } H_0$$

$n = 14, \alpha = 2,5\% \Rightarrow t_{13;\,0,025} = 2,160$ e a região crítica é

$$T_0 < -2,160$$

O valor observado da estatística de teste é

$$t_0 = \frac{35,6 - 40}{\frac{3,4}{\sqrt{14}}} = -4,84$$

Como o valor observado $t_0 = -4,84$ está na região crítica, devemos rejeitar H_0, ou seja, há evidências de que o treinamento conseguiu reduzir o tempo médio de execução.

B.8 Capítulo 8

Dados do Exercício 3.10 - Os 3 testes rejeitam a hipótese de normalidade (valores P pequenos); note que há um valor discrepante!

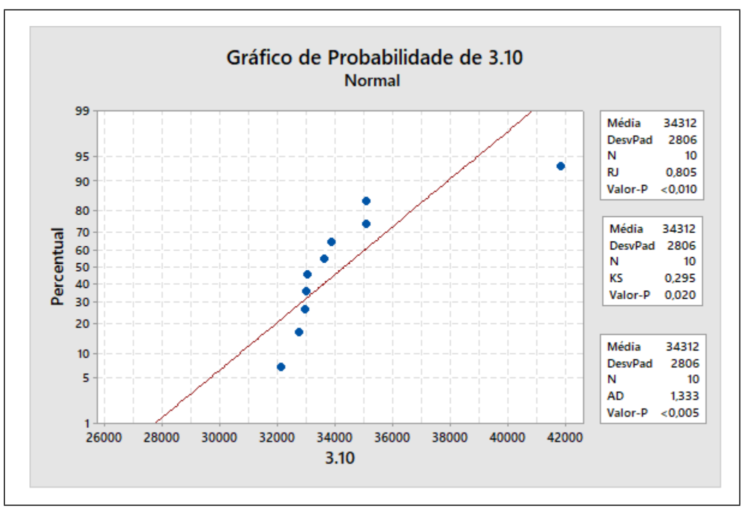


Figura B.2 – Testes de Normalidade para dados do Exercício 3.10

Dados do Exercício 4.4 - Os 3 testes não rejeitam a hipótese de normalidade (valores P grandes).

Dados do Exercício 3.10 - valores repetidos influenciam na normalidade. Apenas o teste de Ryan-Joiner não rejeita a hipótese de normalidade.

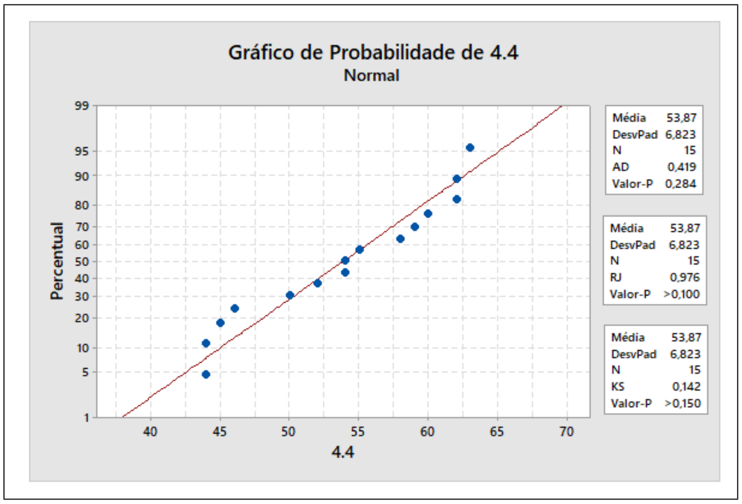


Figura B.3 – Testes de Normalidade para dados do Exercício 4.4

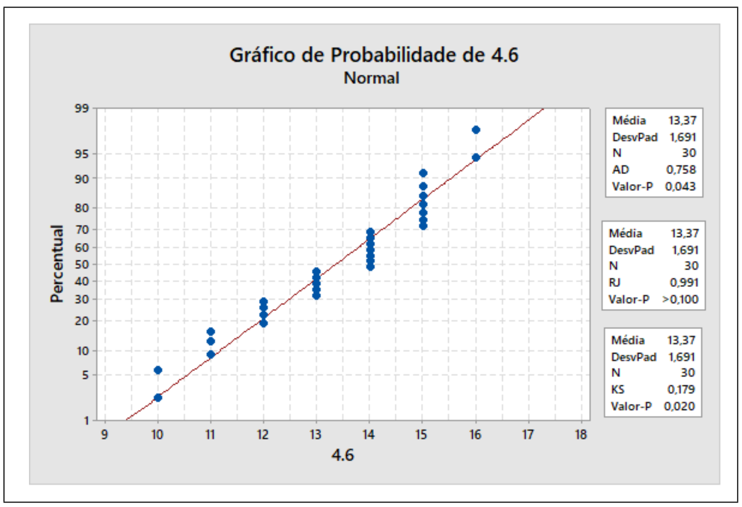


Figura B.4 – Testes de Normalidade para dados do Exercício 4.6

B.9 Capítulo 9

1. Contexto do problema: amostras independentes de duas populações com variâncias conhecidas. Ambas as amostras são grandes, o que nos permite aplicar o Teorema Limite Central.

   (a)
   - População 1: Simplicity (μ_P)
   - População 2: Singer (μ_G)
   - Afirmativa da Singer: $\mu_P - \mu_G < 5$
   - Negação: $\mu_P - \mu_G \geq 5$

$$H_0 : \mu_P - \mu_G = 5$$
$$H_1 : \mu_P - \mu_G < 5$$

Teste unilateral à esquerda – $\alpha = 0,01 \Rightarrow z_{0,01} = -2,33$

Estatística de teste:

$$Z_0 = \frac{\overline{X}_P - \overline{X}_G - \Delta_0}{\sqrt{\frac{\sigma_P^2}{n_P} + \frac{\sigma_G^2}{n_G}}} \approx N(0;1) \quad \text{sob } H_0$$

Região crítica: $Z_0 < -2,33$
Valor observado da estatística de teste:

$$z_0 = \frac{(17,99 - 13,26) - 5}{\sqrt{\frac{2,89}{42} + \frac{2,25}{38}}} = -0,75461 > -2,33$$

Não se rejeita H_0; não há evidências de que as máquinas da Singer sejam mais leves por 5 kg.

(b) $P = \mathrm{P}(Z \leq -0,75461) = 0,5 - 0,2734 = 0,2266$

(c) Como as amostras são grandes, podemos usar a aproximação normal dada pelo Teorema Limite Central.

2. Contexto do problema: amostras independentes de duas populações com variâncias conhecidas. Pelo enunciado, podemos supor que as populações sejam normais.

(a)
- População 1: Feijão (μ_F)
- População 2: Batata (μ_B)
- Afirmativa dada: $\mu_F = \mu_B$
- Negação: $\mu_F \neq \mu_B$

$$H_0 : \mu_F = \mu_B$$
$$H_1 : \mu_F \neq \mu_B$$

Teste bilateral – $\alpha = 0,01 \Rightarrow z_{0,005} = 2,58$
Estatística de teste:

$$Z_0 = \frac{\overline{X}_F - \overline{X}_B}{\sqrt{\frac{\sigma_F^2 + \sigma_B^2}{n}}} \approx N(0;1) \quad \text{sob } H_0$$

Região crítica: $|Z_0| > 2,58$
Valor observado da estatística de teste:

$$z_0 = \frac{39,58 - 40,12}{\sqrt{\frac{2,47 + 0,87}{18}}} = -1,2536$$

Não se rejeita H_0, ou seja não há evidências para se refutar a afirmativa do Departamento de Agricultura.

(b) Valor observado da estatística de teste:

$$z_0 = \frac{39,58 - 40,12}{\sqrt{\frac{2,47 + 0,87}{38}}} = -1,82143$$

Ainda não se rejeita H_0.

(c) Para se rejeitar H_0 temos que ter

$$z_0 = \frac{39,58 - 40,12}{\sqrt{\frac{2,47 + 0,87}{n}}} < -2,58 \Rightarrow \sqrt{n} \cdot (-0,29547) < -2,58 \Rightarrow \sqrt{n} > \frac{2,58}{0,29547} = 8,73171 \Rightarrow n \geq 77$$

3. Embora não tenha sido solicitado, vamos verificar se as suposições feitas no enunciado são válidas.

- Normalidade – Nas Figuras B.5 e B.6 temos os gráficos de probabilidade normal e os valores P do teste de Anderson-Darling. Para as duas populações, não se rejeita a hipótese de normalidade.

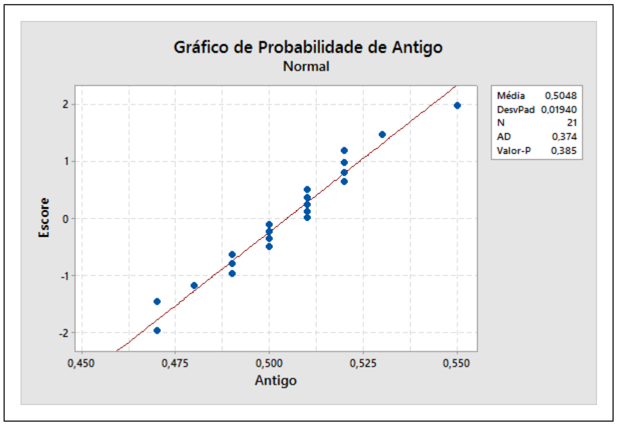


Figura B.5 – Processo antigo

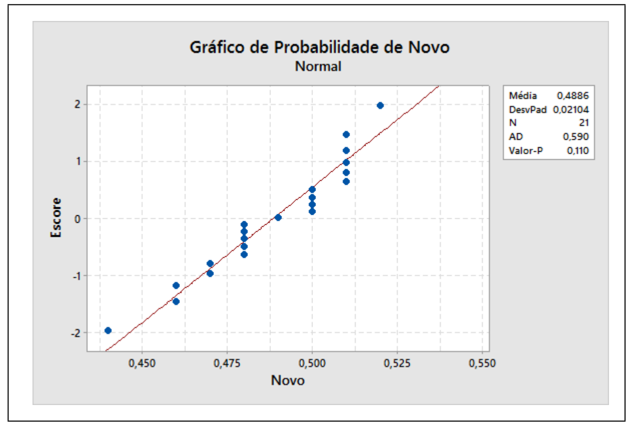


Figura B.6 – Processo novo

- Variâncias iguais

  Sob a hipótese de variâncias iguais de populações normais, $F_0 = \frac{S_1^2}{S_2^2} \sim F_{20,20}$

  $S_1^2 = 0,00037619 \qquad S_2^2 = 0,00044286$

  $P = 2.\,\mathrm{P}\left(F_{20,20} < \frac{0,00037619}{0,00044286}\right) = 2 \cdot 0,359389 = 0,719$

  Não se rejeita a hipótese de variâncias iguais. Veja a saída do Minitab na Figura B.7.

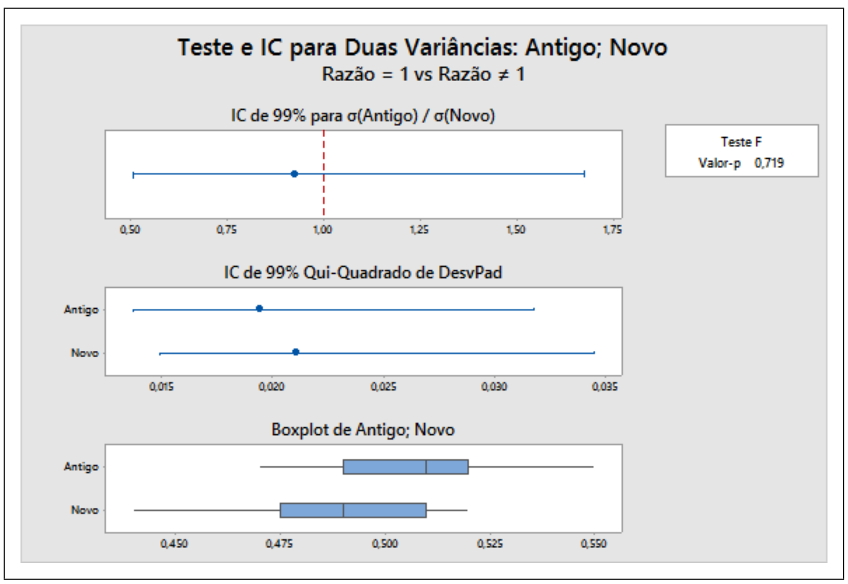


Figura B.7 – Comparação das duas variâncias

Seguimos, agora, com o teste para comparação das médias, baseado na hipótese de populações normais com a mesma variância. Como os tamanhos amostrais são iguais, o estimador combinado da variância será a média aritmética das variâncias amostrais:

$$S_p^2 = \frac{0,00037619 + 0,00044286}{2} = 0,00040953$$

$$H_0 : \mu_N = \mu_A$$
$$H_0 : \mu_N < \mu_A$$

Estatística de teste $T_0 = \dfrac{\mu_N - \mu_A}{\sqrt{S_p^2\left(\dfrac{1}{21}+\dfrac{1}{21}\right)}} \sim t_{21+21-2}$ sob H_0

Região crítica $T_0 < -t_{40;0,01} = -2,42326$ com software ou $-2,33$ pela aproximação normal.

Valor observado da estatística de teste $t_0 = \dfrac{\dfrac{10,26-10,6}{21}}{\sqrt{0,00040953\left(\dfrac{1}{21}+\dfrac{1}{21}\right)}} = -2,5925 < -2,42326$

Rejeita-se H_0; há evidências de que o processo novo resulta em peso menor.

4. Com base nas informações dadas, temos amostras pequenas e independentes de duas populações normais com variâncias desconhecidas quaisquer.

$$H_0 : \mu_H = \mu_M$$
$$H_0 : \mu_H > \mu_M$$

Estatística de teste
$T_0 = \dfrac{\mu_H - \mu_M}{\sqrt{\dfrac{S_H^2}{n_H}+\dfrac{S_M^2}{n_M}}} \approx t_\nu$ sob H_0
Graus de liberdade
$$\nu = \frac{\left(\dfrac{37,50^2}{21}+\dfrac{17,01^2}{19}\right)^2}{\dfrac{\left(\dfrac{37,50^2}{21}\right)^2}{20}+\dfrac{\left(\dfrac{17,01^2}{19}\right)^2}{18}} = 28,49347041 \approx 29$$
Abordagem conservadora: $\nu = \min(21-1, 19-1) = 18$
Região crítica
$T_0 > t_{29;0,10} = 1,311$ ou $T_0 > t_{18;0,10} = 1,067$
Valor observado da estatística de teste
$$t_0 = \frac{784-652}{\sqrt{\dfrac{37,50^2}{21}+\dfrac{17,01^2}{19}}} = 14,56$$
Rejeita-se H_0; há evidências de que os homens relatam gastos maiores do que as mulheres.

5. Veja a saída do Minitab para o teste de Anderson-Darling para normalidade das duas populações (Figuras B.8 e B.9). Não se rejeita a hipótese de normalidade das populações.

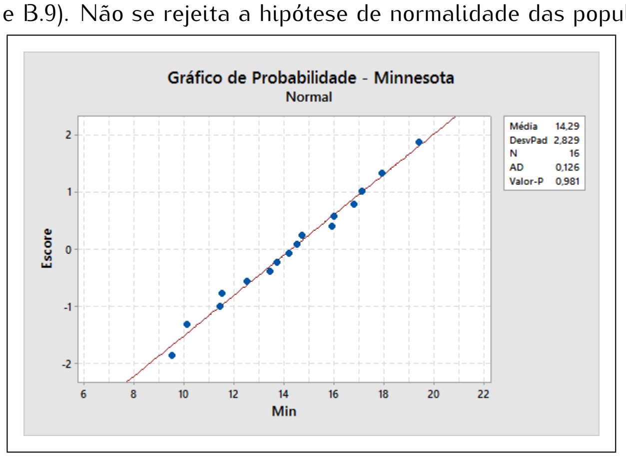


Figura B.8 – Minnesota

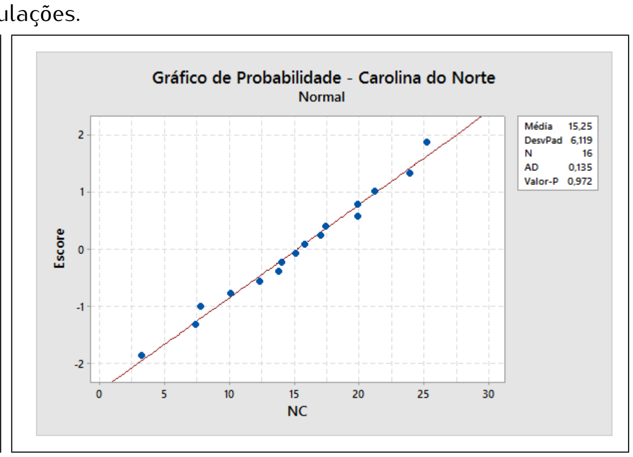


Figura B.9 – Carolina do Norte

$$H_0 : \sigma_M^2 = \sigma_C^2$$
$$H_1 : \sigma_M^2 < \sigma_C^2$$

ou

$$H_0 : \frac{\sigma_C^2}{\sigma_M^2} = 1$$
$$H_1 : \frac{\sigma_C^2}{\sigma_M^2} > 1$$

Estatística de teste

$F_0 = \dfrac{S_{CN}^2}{S_{Min}^2} \sim F_{15,15}$ sob H_0

Região crítica: $F_0 > F_{15,15;0,01} = 3,52219$ (por software).

Valor da estatística de teste: $f_0 = \dfrac{s_C^2}{s_M^2} = \dfrac{37,44667}{8,00383} = 4,67859$

Rejeita-se H_0; há evidência de maior variabilidade no peso dos perus da Carolina do Norte.

A título de exercício, vamos construir o intervalo de confiança de 98% para a razão $\dfrac{\sigma_C^2}{\sigma_M^2}$.

$F_{15,15;0,01} = 3,52219 \Rightarrow F_{15,15;0,99} = \dfrac{1}{3,52219} = 0,28391$

Sabemos que

$$\frac{\dfrac{S_C^2}{\sigma_C^2}}{\dfrac{S_M^2}{\sigma_M^2}} \sim F_{15,15}$$

.

Logo,

$$P\left(0,28391 \le \frac{\dfrac{S_C^2}{\sigma_C^2}}{\dfrac{S_M^2}{\sigma_M^2}} \le 3,52219\right) = 0,98 \Rightarrow$$
$$P\left(0,28391 \le \frac{S_C^2}{S_M^2} \cdot \frac{\sigma_M^2}{\sigma_C^2} \le 3,52219\right) = 0,98 \Rightarrow$$
$$P\left(0,28391 \cdot \frac{S_M^2}{S_C^2} \le \frac{\sigma_M^2}{\sigma_C^2} \le 3,52219 \cdot \frac{S_M^2}{S_C^2}\right) = 0,98 \Rightarrow$$
$$P\left(\frac{1}{3,52219} \cdot \frac{S_C^2}{S_M^2} \le \frac{\sigma_C^2}{\sigma_M^2} \le \frac{1}{0,2839171} \cdot \frac{S_C^2}{S_M^2}\right) = 0,98$$

O intervalo de confiança é

$$\left(\frac{4,67859}{3,52219}\,;\,\frac{4,67859}{0,28391}\right) = (1,32832\,;\,16,47913)$$

Note que ambos os limites são maiores que 1, comprovando que há maior variabilidade na Carolina do Norte.

B.10 Capítulo 10

1. Nas Figuras B.10 e B.11 temos a saída o Minitab para o teste de Anderson-Darling. Não se rejeita a hipótese de normalidade.

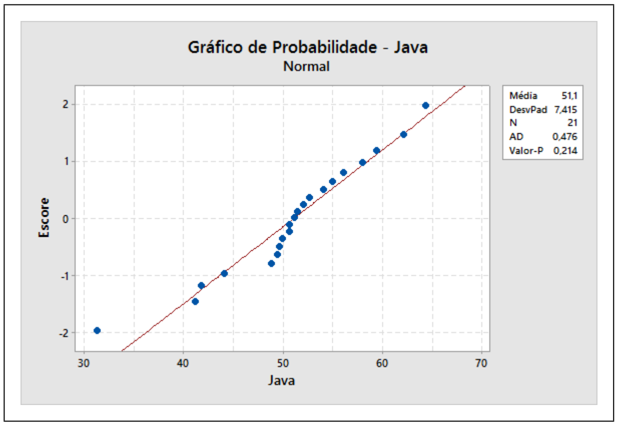


Figura B.10 – Java

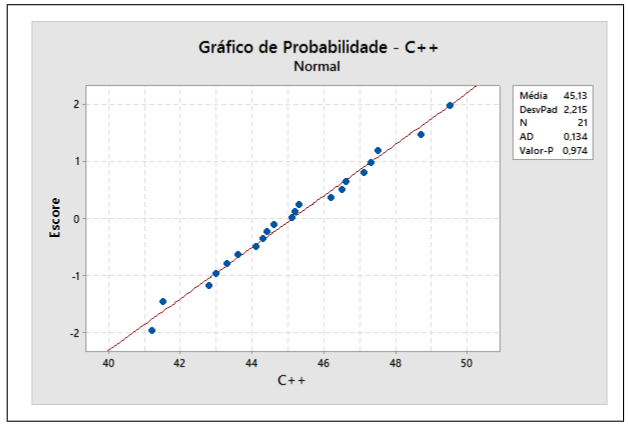


Figura B.11 – C++

(a) Os dados são emparelhados porque o *mesmo programador* fez os dois programas.

(b)

$$H_0 : \mu_J = \mu_C$$
$$H_0 : \mu_J > \mu_C$$

As diferenças $d_i = x_{iJ} - x_{iC}$ são dadas na tabela a seguir:

7,6	9,6	-4,8	10,2	19,7	17,8	-6,1	8,5	3,4	4,9	7,8
7,4	7,3	12,5	-1,2	14,2	7,3	-15,8	5,5	5,8	3,7	

$$\sum_{i=1}^{21} d_i = 135,3 \qquad \sum_{i=1}^{21} d_i^2 = 2039,69$$

$$\overline{d} = 6,44286 \qquad s_d^2 = \frac{1}{20}\left(2039,69 - \frac{135,3^2}{21}\right) = 58,39857$$

Estatística de teste: $T_0 = \dfrac{\overline{X}_J - \overline{X}_C}{\sqrt{\dfrac{S_d^2}{n}}} \sim t_{20}$ sob H_0

Região crítica: $T_0 > t_{20;0,001} = 3,552$

Valor da estatística de teste: $t_0 = \dfrac{6,44286}{\sqrt{\dfrac{58,39857}{21}}} = 3,8635 > 3,552$

Rejeita-se H_0; há evidência de que o tempo de execução com Java é maior do que o tempo de execução com C++.

(c) $P = \mathrm{P}(t_{20} > 3,8635) = 0,000484$ (com software)

Pela tabela, podemos dizer apenas que $P < 0,001$, uma vez que 3,552 é o maior valor disponível para $gl = 20$ e corresponde a $\alpha = 0,001$.

2. (a) Os dados são emparelhados porque a *mesma arma* é usada para dar os 2 tiros, antes e depois da limpeza.

(b)

$$H_0 : \mu_D = \mu_A$$
$$H_0 : \mu_D > \mu_A$$

As diferenças $d_i = x_{iD} - x_{iA}$ são dadas na tabela a seguir:

120	92	-45	-53	-38	15

$$\sum_{i=1}^{21} d_i = 91 \qquad \sum_{i=1}^{21} d_i^2 = 29367$$

$$\overline{d} = 15,16667 \qquad s_d^2 = \frac{1}{5}\left(29367 - \frac{91^2}{6}\right) = 5597,36667$$

Estatística de teste: $T_0 = \dfrac{\overline{X}_D - \overline{X}_A}{\sqrt{\dfrac{S_d^2}{6}}} \sim t_5 \quad \text{sob } H_0$

Região crítica: $T_0 > t_{5;0,01} = 3,365$

Valor da estatística de teste: $t_0 = \dfrac{15,16667}{\sqrt{\dfrac{5597,36667}{6}}} = 0,4966$

Não se rejeita H_0; há evidências de que a limpeza da arma não aumenta a velocidade de boca.

(c) $P = \mathrm{P}(t_5 > 0,4966) = 0,320265$ (com software)

Pela tabela, podemos dizer apenas que $P > 0,15$, uma vez que 1,156 é o menor valor disponível para $gl = 5$ e corresponde a $\alpha = 0,15$.

B.11 Capítulo 13

1. Teste de aderência

$p_{10} = 0,5 \qquad p_{20} = 0,2 \qquad p_{30} = 0,2 \qquad p_{40} = 0,1$

$$
\begin{aligned}
&H_0 : p_i = p_{i0} \qquad \forall i\\
&H_1 : \text{pelo menos um } p_i \neq p_{i0}
\end{aligned}
$$

Sob H_0, as frequências esperadas são

$$
\begin{array}{ll}
E_1 = 0,5 \times 800 = 400 & \qquad E_2 = 0,2 \times 800 = 160\\
E_3 = 0,2 \times 800 = 160 & \qquad E_4 = 0,1 \times 800 = 80
\end{array}
$$

$E_i \geq 5 \quad \forall i$ – aproximação qui-quadrado OK! $\qquad k = 4 \rightarrow \chi^2_{3;0,05} = 7,815$

Valor observado da estatística de teste:

$$
\chi_0^2 = \frac{(430-400)^2}{400} + \frac{(145-160)^2}{160} + \frac{(135-160)^2}{160} + \frac{(90-80)^2}{80} = 8,8125
$$

Rejeita-se H_0; há evidências, ao nível de 5%, de que pelo menos uma das proporções p_i é diferente do valor dado no relatório. Veja saída do Minitab na Figura B.12.

Teste Qui-Quadrado de Qualidade de Ajuste para Contagens Observadas na Variável: Problema

Categoria	Observado	Teste de Proporção	Esperado	Contribuição para Qui-Quadrado
SeguroSaúde	430	0,5	400	2,25000
SeguroResp	145	0,2	160	1,40625
Indenização	135	0,2	160	3,90625
Combustíveis	90	0,1	80	1,25000

N	GL	Qui-Quadrado	Valor-p
800	3	8,8125	0,032

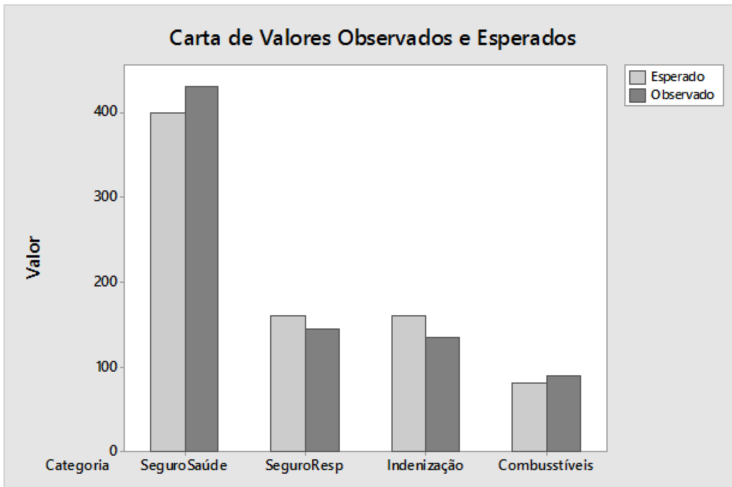


Figura B.12 – Saída do Minitab para o Exercício 1

2. Teste de aderência

$p_{LA,0} = 0,292 \qquad p_{SF,0} = 0,261 \qquad p_{OC,0} = 0,099 \qquad p_{RS,0} = 0,077 \qquad p_{O,0} = 0,271$

$$H_0 : p_i = p_{i0} \qquad \forall i$$
$$H_1 : \text{pelo menos um } p_i \neq p_{i0}$$

Sob H_0, as frequências esperadas são

$$E_{LA} = 0,292 \times 438 = 127,896 \qquad E_{SF} = 0,261 \times 438 = 114,318$$
$$E_{OC} = 0,099 \times 438 = 43,362 \qquad E_{RS} = 0,077 \times 438 = 33,726 \qquad E_O = 0,271 \times 438 = 118,698$$

$E_i \geq 5 \quad \forall i$ – aproximação qui-quadrado OK! $\qquad k = 5 \rightarrow \chi^2_{4;0,005} = 14,860$

Valor observado da estatística de teste:

$$\chi^2_0 = \frac{(125-127,896)^2}{127,896} + \frac{(96-114,318)^2}{114,318} + \frac{(45-43,362)^2}{43,362} + \frac{(44-33,726)^2}{33,726} + \frac{(128-118,698)^2}{118,698} = 6,92143$$

Não se rejeita H_0; não há evidências, ao nível de 0,5%, de que as proporções p_i em 2009 sejam diferentes das proporções em 2008. Veja saída do Minitab na Figura B.13.

Teste Qui-Quadrado de Qualidade de Ajuste para Contagens Observadas na Variável: Local

Categoria	Observado	Teste de Proporção	Esperado	Contribuição para Qui-Quadrado
LA	125	0,292	127,896	0,06558
SF	96	0,261	114,318	2,93523
OC	45	0,099	43,362	0,06188
RS	44	0,077	33,726	3,12978
O	128	0,271	118,698	0,72897

N	GL	Qui-Quadrado	Valor-p
438	4	6,92143	0,140

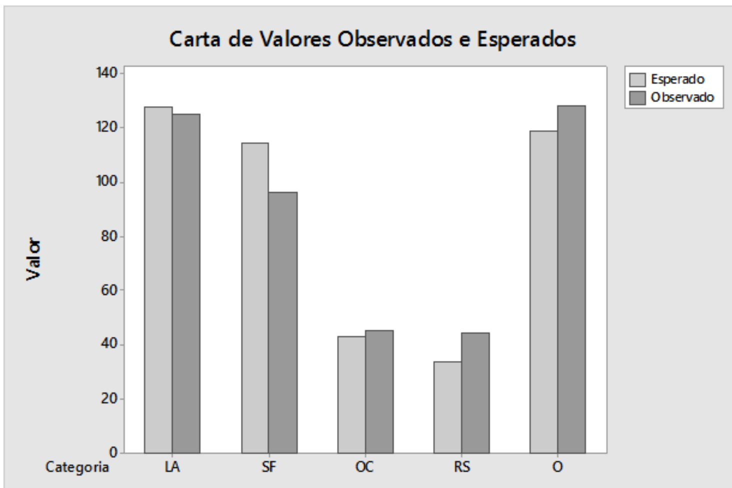


Figura B.13 – Saída do Minitab para o Exercício 2

3. Dados emparelhados - teste de independência

$$H_0 : \text{vinho e carne são independentes}$$
$$H_1 : \text{vinho e carne não são independentes}$$

Sob H_0, as frequências esperadas são

$$e_{11} = \frac{132 \cdot 136}{246} = 72,9756 \qquad e_{12} = \frac{132 \cdot 110}{246} = 59,0244$$
$$e_{21} = \frac{114 \cdot 136}{246} = 63,0244 \qquad e_{12} = \frac{114 \cdot 110}{246} = 50,9756$$

$E_i \geq 5 \quad \forall i$ – aproximação qui-quadrado OK! $\quad gl = (2-1)(2-1) = 1 \rightarrow \chi^2_{1;0,005} = 7,879$ Valor observado da estatística de teste:

$$\chi^2_0 = \frac{(86-72,9756)^2}{72,9756} + \frac{(46-59,0244)^2}{59,0244} + \frac{(50-63,0244)^2}{63,0244} + \frac{(64-50,9756)^2}{50,9756} = 11,2178$$

Rejeita-se H_0; há evidências, ao nível de 0,5%, de que as variáveis “escolha do vinho” e “escolha da carne” não são independentes. Veja saída do Minitab na Figura B.14.

Teste Qui-Quadrado para Independência: C13; Vinho

```
Linhas: Carne  Colunas: Vinho

               Tinto   Branco  Todos

Vermelha          86     46    132
               72,98  59,02
               2,325  2,874

Branca            50     64    114
               63,02  50,98
               2,692  3,328

Todos            136    110    246

Conteúdo da Célula:       Contagem
                          Contagem esperada
                          Contribuição para Qui-Quadrado


Qui-Quadrado de Pearson = 11,218; GL = 1; Valor-P = 0,001
Qui-Quadrado da Razão de Verossimilhanças = 11,284; GL = 1; Valor-P = 0,001
```

Figura B.14 – Saída do Minitab para o Exercício 3

4. Teste de homogeneidade

O objetivo é testar se as proporções de usuários de cda tipo de recurso de escrita são as mesmas para as 2 lojas. Na Figura B.16 temos a saída o Minitab. Aí podemos ver que todas as frequências esperadas são maiores que 5, o que permite o uso da aproximação qui-quadrado. O número de graus de liberdade é $gl = (2-1)(3-1) = 2$ e o valor crítico para $\alpha = 0,01$ é 9,210. O valor da estatística de teste é 13,388, o que nos leva à rejeição da hipótese nula, ou seja, as proporções de usuários dos diferentes tipos de recursos não são as mesmas nas 2 lojas. Pela tabela, temos que o valor P está entre 0,001 e 0,005. O Minitab dá $P = 0,001$, mas o valor mais preciso é 0,001238.

5. Teste de homogeneidade

O objetivo é testar se as proporções de jogadores em cada um dos jogos são as mesmas nos 4 cassinos. Na Figura B.15 temos a saída o Minitab. Aí podemos ver que todas as frequências esperadas são maiores que 5, o que permite o uso da aproximação qui-quadrado. O número de graus de liberdade é $gl = (4-1)(4-1) = 9$ e o valor crítico para $\alpha = 0,05$ é 16,919. O valor da estatística de teste é 18,801, o que nos leva à rejeição da hipótese nula, ou seja, as proporções de jogadores nos diferentes jogos não são as mesmas em todos os cassinos. Pela tabela, temos que o valor P está entre 0,025 e 0,05. O Minitab dá $P = 0,027$.

Teste Qui-Quadrado para Homogeneidade: Jogo; Cassino

```
Linhas: Jogo   Colunas: Cassino

               Bellagio  Caesar  GoldenNugget  Harrahs  Todos

Blackjack            22      30            28       38    118
                  27,00   29,22         28,67    33,11
                 0,9270  0,0207        0,0156   0,7233

Poquer               20      38            25       25    108
                  24,71   26,75         26,24    30,30
                 0,8994  4,7353        0,0584   0,9274

Roleta               38      22            21       29    110
                  25,17   27,24         26,72    30,86
                 6,5368  1,0085        1,2261   0,1123

CacaNiquel           66      68            81       87    302
                  69,11   74,79         73,37    84,73
                 0,1399  0,6164        0,7935   0,0608

Todos               146     158           155      179    638

Conteúdo da Célula:       Contagem
                          Contagem esperada
                          Contribuição para Qui-Quadrado


Qui-Quadrado de Pearson = 18,801; GL = 9; Valor-P = 0,027
Qui-Quadrado da Razão de Verossimilhanças = 17,678; GL = 9; Valor-P = 0,039
```

Figura B.15 – Saída do Minitab para o Exercício 4

Teste Qui-Quadrado para Homogeneidade: Recurso; Loja

```
Linhas: Recurso   Colunas: Loja

             CCruz  Kalunga  Todos

Lápis          183      130    313
             163,9    149,1
             2,218    2,439

Lapiseira      164      202    366
             191,7    174,3
             4,000    4,399

Caneta         480      420    900
             471,4    428,6
             0,158    0,174

Todos          827      752   1579

Conteúdo da Célula:       Contagem
                          Contagem esperada
                          Contribuição para Qui-Quadrado


Qui-Quadrado de Pearson = 13,388; GL = 2; Valor-P = 0,001
Qui-Quadrado da Razão de Verossimilhanças = 13,410; GL = 2; Valor-P = 0,001
```

Figura B.16 – Saída do Minitab para o Exercício 5

Apêndice C

Algumas demonstrações

C.1 Transformação de variáveis aleatórias contínuas

Para demonstrar o Teorema 4.1, faremos uso do seguinte resultado:

RESULTADO C.1 *Sejam X, Y variáveis aleatórias contínuas com densidade conjunta $f_{X,Y}(x, y) > 0$ numa região $A \subset \mathbb{R}^2$. Suponha que (i) $u = h_1(x, y)$ e $v = h_2(x, y)$ definam uma transformação biunívoca de $A \subset \mathbb{R}^2$ em $B \subset \mathbb{R}^2$; (ii) as derivadas parciais de $x = h_1^{-1}(u, v)$ e $y = h_2^{-1}(u, v)$ sejam contínuas em $B \subset \mathbb{R}^2$; (iii) o Jacobiano da transformação seja diferente de zero para $(u, v) \in B$. Então, a densidade conjunta de $U = h_1(X, Y)$ e $V = h_2(X, Y)$ é dada por*

$$f_{U,V}(u, v) = |J|\, f_{X,Y}\left(h_1^{-1}(u, v), h_2^{-1}(u, v)\right)$$

onde $|J|$ é o módulo do Jacobiano da transformação definido pelo seguinte determinante:

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

C.2 Demonstração do Teorema 4.1

Sejam $Z \sim N(0; 1)$ e $Y \sim \chi^2(n)$ variáveis aleatórias independentes. Vamos obter a densidade da variável aleatória

$$T = \frac{Z}{\sqrt{Y/n}}$$

Da hipótese de independência segue que

$$f_{Z,Y}(z, y) = f_Z(z) f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{n/2}} y^{n/2-1} e^{-y/2}$$

Vamos completar a transformação definindo

$$\begin{aligned} t &= \frac{z}{\sqrt{y/n}} \\ v &= y \end{aligned}$$

Então a região

$$A = \{(z, y) : -\infty < z < \infty; 0 < y < \infty\}$$

é transformada na região

$$B = \{(t, v) : -\infty < t < \infty; 0 < v < \infty\}$$

e temos que

$$\begin{aligned} z &= t\sqrt{v/n} \\ y &= v \end{aligned}$$

Logo, o jacobiano da transformação é

$$J = \begin{vmatrix} \frac{\partial z}{\partial t} & \frac{\partial z}{\partial v} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \sqrt{v/n} & \frac{t}{\sqrt{n}}\frac{1}{2\sqrt{v}} \\ 0 & 1 \end{vmatrix} = \sqrt{v/n}$$

Assim, a densidade conjunta de (T, V) é

$$\begin{aligned} f_{T,V}(t,v) &= \sqrt{v/n}\frac{1}{\sqrt{2\pi}}e^{-\left(t^2 v/n\right)/2}\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{n/2}}v^{n/2-1}e^{-v/2} \\ &= \frac{v^{1/2}}{\sqrt{2n\pi}}\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{n/2}}v^{n/2-1}e^{-\frac{1}{2}\left(\frac{t^2}{n}+1\right)v} \\ &= \frac{1}{\sqrt{2n\pi}}\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{n/2}}v^{\frac{n+1}{2}-1}e^{-\frac{1}{2}\left(\frac{t^2}{n}+1\right)v} \end{aligned}$$

e a densidade marginal de T é

$$\begin{aligned} f_T(t) &= \int f_{T,V}(t,v)dv \\ &= \frac{1}{\sqrt{2n\pi}}\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{n/2}}\int_0^\infty v^{\frac{n+1}{2}-1}e^{-\frac{1}{2}\left(\frac{t^2}{n}+1\right)v}dv \end{aligned}$$

Façamos a seguinte mudança de variável

$$w = \frac{1}{2}\left(\frac{t^2}{n}+1\right)v$$

Então

$$\begin{aligned} f_T(t) &= \frac{1}{\sqrt{2n\pi}}\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{n/2}}\int_0^\infty \left[\frac{w}{\frac{1}{2}\left(\frac{t^2}{n}+1\right)}\right]^{\frac{n+1}{2}-1} e^{-w}\frac{dw}{\frac{1}{2}\left(\frac{t^2}{n}+1\right)} \\ &= \frac{1}{\sqrt{2n\pi}}\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{n/2}}\frac{1}{\left[\frac{1}{2}\left(\frac{t^2}{n}+1\right)\right]^{\frac{n+1}{2}}}\int_0^\infty w^{\frac{n+1}{2}-1}e^{-w}dw \\ &= \frac{1}{\sqrt{n\pi}}\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{1/2}2^{n/2}\cdot\left(\frac{1}{2}\right)^{\frac{n+1}{2}}}\frac{1}{\left(\frac{t^2}{n}+1\right)^{\frac{n+1}{2}}}\Gamma\left(\frac{n+1}{2}\right) \\ &= \frac{1}{\sqrt{n\pi}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} \end{aligned}$$

Essa é a densidade da t_n, conforme afirmado no Teorema 4.1.

Bibliografia

Kokoska, S. (2011) *Introductory Statistics: A problem-solving approach*. W.H.Freeman.

Larson, H. J. (1982) *Introduction to Probability Theory and Statistical Inference*. Wiley, 3a. ed.

Moore, D. S., Notz, W. I. e Fligner, M. A. (2017) *A Estatística Básica e Sua Prática*. LTC Editora, 7 ed.

Shapiro, S. S. e Wilk, M. (1965) An analysis of variance test for normality (complete samples). *Biometrika*, **52**, 591–611.

Stigler, S. M. (1984) Kruskal's proof of the joint distribution of $\overline{X}$ e s^2. *The American Statistician*, **2**, 134–135.

Triola, M. F. (2014) *Elementary Statistics*. Pearson, 12 ed.