GET00182 – Estatística II – Prova 1 – 26/04/2017 – 1/2017

1. Seja $X_1, X_2, \cdots, X_n$ uma amostra aleatória simples de uma população $X \sim Bern(p)$. Determine o valor da constante a que torna o estimador

$$\Gamma = a\sum_{i=1}^{n}(X_i + X_i^2)$$

não viesado para estimar p.

Solução:

$$X \sim Bern(p)\left\{\Rightarrow \begin{array}{l} \mathrm{E}(X) = p \\ \mathrm{E}(X^2) = p \\ \mathrm{Var}(X) = p(1-p) \end{array}\right.$$

Temos que ter $\mathrm{E}(\Gamma) = p$.

$$\mathrm{E}(\Gamma) = p \Leftrightarrow a\sum_{i=1}^{n}\left[\mathrm{E}(X_i) + \mathrm{E}(X_i^2)\right] = p \Leftrightarrow a\sum_{i=1}^{n}(p+p) = p \Leftrightarrow a2np = p \Leftrightarrow a = \frac{1}{2n}$$

---

2. Seja $X_1, X_2, \cdots, X_n$ uma amostra aleatória simples de uma população $X \sim \exp(\lambda)$.[1] Sabe-se que $T = 2\lambda\sum_{i=1}^{n} X_i$ tem distribuição qui-quadrado com $2n$ graus de liberdade.

(a) Use as informações dadas para obter a expressão de um intervalo de confiança para λ com nível de confiança $1-\alpha$.

Solução:

Consideremos os dois valores críticos da distribuição qui-quadrado com $2n$ graus de liberdade $\chi^2_{2n;\alpha/2}$ e $\chi^2_{2n;1-\alpha/2}$ que deixam probabilidade $\alpha/2$ e $1-\alpha/2$ acima deles. Então

$$\mathrm{P}\left(\chi^2_{2n;1-\alpha/2} < \chi^2_{2n} < \chi^2_{2n;\alpha/2}\right) = 1-\alpha \Rightarrow$$
$$\mathrm{P}\left(\chi^2_{2n;1-\alpha/2} < 2\lambda\sum_{i=1}^{n} X_i < \chi^2_{2n;\alpha/2}\right) = 1-\alpha \Rightarrow$$
$$\mathrm{P}\left(\frac{\chi^2_{2n;1-\alpha/2}}{2\sum_{i=1}^{n} X_i} < \lambda < \frac{\chi^2_{2n;\alpha/2}}{2\sum_{i=1}^{n} X_i}\right) = 1-\alpha$$

Logo, o intervalo de confiança de nível $1-\alpha$ para λ é

$$\left(\frac{\chi^2_{2n;1-\alpha/2}}{2\sum_{i=1}^{n} X_i}\,;\ \frac{\chi^2_{2n;\alpha/2}}{2\sum_{i=1}^{n} X_i}\right)$$

[1] $f_X(x) = \lambda e^{-\lambda x}\ x > 0 \qquad \mathrm{E}(X) = \frac{1}{\lambda}$

(b) O tempo para falha de determinado componente eletrônico pode ser descrito por uma distribuição exponencial com parâmetro λ desconhecido. Uma amostra aleatória simples de 10 desses componentes foi colocada em teste até que todos falhassem. Os tempos (em horas) para falha observados foram:

607,5 1947,0 37,6 129,9 409,5 529,5 109,0 582,4 499,0 188,1

Obtenha um intervalo de confiança de 90% para o *tempo médio para falha* do componente. (Obs.: $\sum_{i=1}^{10} X_i = 5039,5$)

Solução:

Os valores críticos da qui-quadrado com 20 graus de liberdade são

$$\chi^2_{20;0,95} = 10,851 \qquad \chi^2_{20;0,05} = 31,410$$

e o intervalo de confiança para λ é

$$\left(\frac{10,851}{2 \cdot 5039,5} \; \frac{31,410}{2 \cdot 5039,5} \right) = (0,001077\,;\, 0,003116)$$

Para o tempo de vida médio $\frac{1}{\lambda}$, o intervalo de confiança é

$$\left(\frac{2 \cdot 5039,5}{31,410} \,;\, \frac{2 \cdot 5039,5}{10,851} \right) = (320,885\,;\, 928,854)$$

---

3. Um professor solicitou a Joãozinho que construísse um intervalo de confiança para a média de uma população normal com variância 16. Para tal, o professor forneceu os resultados obtidos a partir de uma amostra de tamanho 25. Como faltavam apenas 5 minutos para começar o jogo do Flamengo, seu time de coração, Joãozinho fez a tarefa correndo e entregou o seguinte resultado para o professor:

[1,032 ; 4,168]

(a) Qual é o valor do estimador pontual da média populacional?

Solução:

$$\overline{x} = \frac{1,032 + 4,168}{2} = 2,6$$

(b) Qual foi o nível de confiança utilizado por Joãozinho?

Solução:

$$\epsilon = 2,6 - 1,032 = 1,568 = z_{\alpha/2} \cdot \frac{4}{5} \Rightarrow z_{\alpha/2} = 1,568 \cdot 1,25 = 1,96 \Rightarrow 1 - \alpha = 0,95$$

(c) O professor, na verdade, já havia solicitado que Joãozinho utilizasse um nível de confiança de 90%, mas Joãozinho não escutou, porque só pensava no Flamengo. Se ele tivesse feito com o nível de confiança correto, o intervalo teria comprimento menor ou maior? Responda sem fazer qualquer cálculo adicional.

Solução:

Como o professor pediu um nível de confiança menor (90%) do que o utilizado por Joãozinho (95%), o intervalo de confiança correto teria comprimento menor.

VIRE!

4. O gerente de RH de um hospital está preocupado com a demora no atendimento no serviço de triagem e também com as diferenças entre funcionários no tempo de atendimento. Parte dessa demora é devida a processos burocráticos envolvendo planos de saúde. Decide, então, realizar um treinamento para os funcionários. Depois do treinamento, ele seleciona uma amostra de 16 tempos de atendimento que resultam numa média de 9,3 minutos e desvio padrão de 2,83 minutos.

   (a) Construa intervalos de confiança de 90% para a média e a variância do novo tempo de atendimento. Assuma que a distribuição do tempo de atendimento possa ser bem aproximada por uma distribuição normal.

   Solução:

   T =tempo de atendimento após o treinamento:

$$T \sim N(\mu; \sigma^2) \Rightarrow \begin{cases} \sqrt{n}\dfrac{\overline{X}-\mu}{S} \sim t_{n-1} \\ \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1} \end{cases}$$

   I.C. para a média μ

$$\epsilon = t_{15;0,05}\frac{2,83}{\sqrt{16}} = 1,753 \cdot 0,7075 = 1,238$$

$$[9,3 - 1,238\,;\ 9,3 + 1,238] = [8,062\,;\ 10,538]$$

   I.C. para σ^2

$$\chi^2_{15;0,05} = 24,996 \quad \chi^2_{15;0,95} = 7,261$$

$$\left[\frac{15 \cdot 2,83^2}{24,996}\,;\ \frac{15 \cdot 2,83^2}{7,261}\right] = [4,806\,;\ 16,545]$$

   (b) Antes do treinamento, o tempo médio de atendimento era de 12 minutos. Há algum indício nos seus resultados de que que houve melhora?

   Solução:

   O intervalo de confiança para a média está totalmente à esquerda de 12, o que é um indício de edução do tempo médio.

---

5. Determine o que está errado em cada uma das afirmativas a seguir, justificando sua resposta:

   (a) A hipótese nula para um teste de hipótese sobre a média de uma população é $H_0 : \overline{X} = 25$.

   Solução:

   A hipótese nula é especificada em termos de um parâmetro, e não em termos de uma estatística amostra. O correto seria $H_0 : \mu = 25$.

   (b) O erro quadrático médio de um estimador é sempre menor que a variância do estimador.

   Solução:

   Sabemos que

$$EQM(\widehat{\theta}) = \text{Var}(\widehat{\theta}) + [B(\widehat{\theta})]^2$$

   Logo, o erro quadrático médio de um estimador é sempre maior ou igual à sua variância.

(c) Tendo obtido o intervalo de confiança de 90% $[0,265\,;\,0,328]$ para uma proporção p, podemos afirmar que a probabilidade de esse intervalo conter o verdadeiro valor de p é 0,90.

Solução:

A probabilidade de 0,90 refere-se à probabilidade de "acerto" do método de obtenção dos intervalos de confiança, isto é, se sorteássemos muitas e muitas amostras e para c ada uma construíssemos o respectivo intervalo de confiança, 90% desses intervalos conteriam o verdadeiro valor de p. Um intervalo obtido com uma amostra específica ou contém ou não contém p.

(d) Região crítica e valor P para um teste de hipótese referem-se, ambos, a alguma probabilidade.

Solução:

Valor P é uma probabilidade, mas região crítica é um conjunto de valores de uma estatística amostral que levam à rejeição da hipótese nula de um teste.

---

6. 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 uma amostra aleatória simples de tamanho $n = 16$:

$$RC : \overline{X} > 48,75$$

(a) Calcule a probabilidade do erro tipo I.

Solução:

$$\begin{aligned} \alpha &= \mathrm{P}(\overline{X} > 48,75 | \overline{X} \sim N(45; 36/16)) = \mathrm{P}\left(Z > \frac{48,75 - 45}{1,5}\right) = \mathrm{P}(Z > 2,5) \\ &= 0,5 - 0,4938 = 0,0062 \end{aligned}$$

(b) Calcule a probabilidade do erro tipo II se $\mu = 47$.

Solução:

$$\begin{aligned} \beta &= \mathrm{P}(\overline{X} \leq 48,75 | \overline{X} \sim N(47; 36/16)) = \mathrm{P}\left(Z \leq \frac{48,75 - 47}{1,5}\right) = \mathrm{P}(Z \leq 1,17) \\ &= 0,5 + 0,3790 = 0,8790 \end{aligned}$$
