

Tópicos de Resolução - Exame ER (2025)

(20) 1. Das afirmações seguintes, assinale com uma cruz (X) as que são verdadeiras e as que são falsas.

		V	F
1	Uma estimativa é uma variável aleatória, função da amostra.		X
2	A função de verosimilhança não depende de nenhum parâmetro desconhecido.		X
3	A região de rejeição num teste de hipóteses resulta quando fixamos o erro de 1ª espécie .	X	
4	Um intervalo aleatório (IA) cobre o parâmetro θ com determinada probabilidade, enquanto que um intervalo de confiança é uma concretização deste IA para uma determinada amostra.	X	
5	Num teste de hipóteses, um valor- p superior ao erro de 1ª espécie indica que devemos rejeitar a hipótese nula.		X
6	Considere o modelo $y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$. Nesse caso, a variável residual tem valor esperado nulo.		X
7	Considere o modelo $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$, com $\beta_2 > 0$. Suponha que, em alternativa, foi estimado o modelo $y_i = \delta_0 + \delta_1 x_{1i} + \delta_2 x_{2i} + \delta_3 x_{3i} + e_i$. No segundo caso definitivamente $\delta_2 > 0$.		X
8	Num teste de hipóteses para amostras emparelhadas, a região de rejeição pode ser bilateral.	X	
9	Considere o modelo $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$ com $R^2=0.7$. Neste caso o modelo explica 30% da variação no y .		X
10	Num teste de hipóteses sobre igualdade de variâncias entre duas populações normais, a estatística-teste tem distribuição F-snedecor.	X	

RESPOSTA ABERTA

2. Seja X uma variável aleatória com a seguinte função densidade:

$$f(x | \theta) = \theta x^{\theta-1}, \quad 0 < x < 1; \theta > 0$$

Recolheu-se uma amostra casual desta população, tendo-se observado: (0.11, 0.25, 0.03, 0.15).

- (20) (a) Mostre que o estimador de máxima verosimilhança para θ é dado por: $\hat{\theta} = -n / \sum_{i=1}^n \ln(X_i)$.

RESPOSTA:

$$L(\theta | x_1, \dots, x_n) = \prod_{i=1}^n (\theta x_i^{\theta-1}) = \theta^n \left(\prod_{i=1}^n x_i \right)^{\theta-1}, \quad \theta > 0$$

$$l(\theta | x_1, \dots, x_n) = n \ln \theta + (\theta - 1) \sum_{i=1}^n \ln x_i, \quad \theta > 0$$

$$\begin{aligned} l'(\theta | x_1, \dots, x_n) = 0 &\Leftrightarrow \frac{n}{\theta} + \sum_{i=1}^n \ln x_i = 0 \Leftrightarrow \frac{n + \theta \sum_{i=1}^n \ln x_i}{\theta} = 0 \Leftrightarrow n + \theta \sum_{i=1}^n \ln x_i = 0 \Leftrightarrow \\ &\Leftrightarrow \theta^* = -\frac{n}{\sum_{i=1}^n \ln x_i} \end{aligned}$$

Como $l''(\theta | x_1, \dots, x_n) = -n/\theta^2 < 0$, então conclui-se que o EMV para θ é dado por:

$$\hat{\theta} = -\frac{n}{\sum_{i=1}^n \ln X_i}$$

- (10) (b) Sabendo que $E(X) = \theta/(\theta + 1)$, use os dados da amostra para calcular a estimativa de máxima verosimilhança para a média desta população. Justifique o seu cálculo.

RESPOSTA:

Note-se que $E(X)$ é uma função biunívoca de θ . Logo, pela propriedade da invariância dos EMV, tem-se:

$$\hat{E}(X) = \widehat{\left(\frac{\theta}{\theta + 1} \right)} = \frac{\hat{\theta}}{\hat{\theta} + 1}$$

Como a estimativa MV para θ é: $\hat{\theta} = -4/(\ln 0.11 + \ln 0.25 + \ln 0.03 + \ln 0.15) \approx 0.44458$, então:

$$\hat{E}(X) \approx \frac{0.44458}{0.44458 + 1} \approx 0.30776$$

- (10) (c) O limite inferior de Fréchet-Cramér-Rao é θ^2/n . Relacione este resultado com a variância dos estimadores para θ .

RESPOSTA:

Com esta informação sabe-se que a variância dos estimadores centrados para θ não pode ser inferior a θ^2/n , pela desigualdade de Fréchet-Cramér-Rao. Logo, a existir estimador mais eficiente de θ , este terá de ter variância igual a θ^2/n .

3. Num estudo sobre agroturismo, questionaram-se aleatoriamente 250 quintas da região norte e 300 quintas da região sul. Das quintas da região norte, 160 referiram ter atividades de agroturismo, enquanto na região sul apenas 120 quintas referiram ter tais atividades. De acordo com esta informação, construa um intervalo de confiança a 95% para a diferença de proporções entre as quintas do norte e sul que fornecem atividades de agroturismo. Interprete o intervalo obtido e conclua.

RESPOSTA:

Populações de Bernoulli

$$\mathbf{VF:} Z = \frac{\bar{X}_1 - \bar{X}_2 - (\theta_1 - \theta_2)}{\sqrt{\frac{\bar{X}_1(1 - \bar{X}_1)}{n_1} + \frac{\bar{X}_2(1 - \bar{X}_2)}{n_2}}} \stackrel{a}{\sim} N(0, 1)$$

Como $\bar{x}_1 = 160/250 = 0.64$ e $\bar{x}_2 = 120/300 = 0.4$ e $z_{0.025} = 1.96$, o IC a 95% para $\theta_1 - \theta_2$ vem:

$$\left(\bar{x}_1 - \bar{x}_2 \mp z_{0.025} \sqrt{\frac{\bar{x}_1(1 - \bar{x}_1)}{n_1} + \frac{\bar{x}_2(1 - \bar{x}_2)}{n_2}} \right) = \left(0.64 - 0.4 \mp 1.96 \sqrt{\frac{0.64 \times 0.36}{250} + \frac{0.4 \times 0.6}{300}} \right) =$$

$$= (0.158675, 0.321325)$$

Podemos afirmar com uma confiança de 95% que a diferença de proporções entre as quintas do norte e do sul que fornecem atividades de agroturismo situa-se entre cerca de 0.159 e 0.321. Como o intervalo contém valores estritamente positivos, existe evidência estatística de que a proporção de quintas que fornecem atividades de agroturismo é superior nas quintas do norte.

- (40) 4. Dois artigos cosmeticos para o cabelo, A e B, foram testados em 12 mulheres tendo-se obtido os seguintes resultados para os respectivos efeitos na hidratação do cabelo:

Participante	Efeito de A	Efeito de B
1	10.5	5.2
2	2.3	1.4
3	8.2	7.5
4	13.5	8.7
5	4.5	3.9
6	8.1	6.5
7	11.5	12.2
8	5.6	5.9
9	7.4	8.5
10	9.5	10.2
11	1.3	2.4
12	9.2	9.5

Com base nesses dados e num teste estatístico adequado e considerando uma dimensão de 0.05, verifique se ha uma diferença entre a eficacia dos dois produtos.

RESPOSTA:

Trata-se de um teste para diferença de medias em **amostras emparelhadas**.

Z = diferenças =

Efeito de A – Efeito de B = [5.3, 0.9, 0.7, 4.8, 0.6, 1.6, -0.7, -0.3, -1.1, -0.7, -1.1, -0.3]

$$\bar{Z} = 0.81$$

$$S_Z'^2 = \sum_{i=1}^{12} (Z_i - \bar{Z})^2 / 11 = 4.64$$

$$S_Z' = \sqrt{4.64} = 2.15$$

$$T = \frac{\bar{Z}}{S_Z' / \sqrt{n}} \sim t(n - 1)$$

$$T_{obs} = \frac{\bar{Z}}{2.15 / \sqrt{12}} = 1.30$$

t(n-1) 11 graus de liberdade

$$t_{crit} = 2.201$$

$$W = \{t : t < -2.201 \text{ ou } t > 2.201\}$$

$$t_{obs} < 2.201$$

Não rejeitamos H_0 . Os produtos tem o mesmo efeito em media.

5. Considere o seguinte modelo para estudar o salário por hora dos trabalhadores portugueses:

$$salar\acute{o} = \beta_0 + \beta_1 exper + \beta_2 exper^2 + \beta_3 educ + u, \quad (1)$$

onde, para cada trabalhador, *salar\acute{o}* corresponde ao salário mensal bruto em milhares de euros, *exper* corresponde ao número de anos de experiência profissional, e *educ* corresponde ao número de anos de escolaridade. Assuma que o erro, u , tem média nula.

Com base numa amostra de 200 trabalhadores, recolhida em 2024, estimou-se o seguinte modelo (com os erros-padrão dentro de parêntesis) com recurso ao método dos Mínimos Quadrados Ordinários (OLS):

$$\widehat{salar\acute{o}} = \underset{(0.120)}{1.101} + \underset{(0.015)}{0.054} exper - \underset{(0.001)}{0.001} exper^2 + \underset{(0.025)}{0.080} educ, \quad (2)$$

$$SSR = 190, R^2 = 0.48$$

- (10) (a) Interprete a estimativa $\hat{\beta}_3$ e teste a sua significância estatística considerando $\alpha = 0.1$.

Resposta: Cada ano adicional de escolaridade está associado, em média, a um aumento de 80 euros no salário mensal bruto, ceteris paribus.

Teste de significância estatística:

$$H_0 : \beta_3 = 0 \quad \text{contra} \quad H_1 : \beta_3 \neq 0$$

$$T = \frac{\hat{\beta}_3 - \beta_3}{\hat{\sigma}_{\hat{\beta}_3}} = \frac{\hat{\beta}_3}{\hat{\sigma}_{\hat{\beta}_3}} \sim t_{(200-3-1)} \xrightarrow{d} N(0, 1)$$

$$W_{0.05} = \{t : t < -1.645 \text{ ou } t > 1.645\}$$

$$t_{obs} = \frac{0.080}{0.025} = 3.2$$

Logo, rejeita-se a hipótese nula ao nível de significância de 10%. A variável *educ* tem relevância estatística.

- (10) (b) Construa um intervalo de confiança a 95% para β_2 . O que pode concluir sobre a significância estatística deste coeficiente?

Resposta: O intervalo de confiança com 95% de confiança para β_2 é dado por:

$$\left(\hat{\beta}_2 - z_{0.025} \times \hat{\sigma}_{\hat{\beta}_2}; \hat{\beta}_2 + z_{0.025} \times \hat{\sigma}_{\hat{\beta}_2} \right) = (-0.001 - 1.96 \times 0.001; -0.001 + 1.96 \times 0.001)$$

$$= (-0.00296; 0.00096)$$

Como o intervalo de confiança inclui o zero, o coeficiente não é estatisticamente significativo ao nível de 5% .

- (20) (c) Calcule a previsão em média para *salario* sabendo que o trabalhador tem 12 anos de escolaridade e 10 anos de experiência profissional. Proponha uma reparametrização do modelo (1) que permita construir um intervalo de confiança para essa previsão.

Resposta:

$$\widehat{E(salario|exper = 10, educ = 12)} = 1.101 + 0.054 \times 10 - 0.001 \times 10^2 + 0.080 \times 12 = 2.501$$

Defina $\theta_0 = \beta_0 + 10\beta_1 + 100\beta_2 + 12\beta_3$. Note que $\beta_0 = \theta_0 - 10\beta_1 - 100\beta_2 - 12\beta_3$, logo, a equação (1) pode ser reparametrizada desta forma:

$$salario = (\theta_0 - 10\beta_1 - 100\beta_2 - 12\beta_3) + \beta_1 exper + \beta_2 exper^2 + \beta_3 educ + u$$

$$salario = \theta_0 + \beta_1(exper - 10) + \beta_2(exper^2 - 100) + \beta_3(educ - 12) + u$$

Ao estimar esta última regressão obtém-se diretamente $\hat{\theta}_0$ e $\hat{\sigma}_{\hat{\theta}_0}$ e consegue-se construir o intervalo de confiança para θ_0 .

- (15) (d) Deduza o efeito parcial da experiência profissional sobre o salário mensal. A partir de que nível de escolaridade muda o sinal deste efeito?

Resposta:

O efeito parcial médio é dado por:

$$\frac{\partial \widehat{salario}}{\partial exper} = \hat{\beta}_1 + 2\hat{\beta}_2 = 0.054 + 2 \times (-0.001)exper$$

O efeito troca de sinal a partir do ponto: $0.054 - 0.002exper = 0$. Logo, $exper^* = 27$, ou seja, o efeito de *exper* será positivo no salário enquanto $exper < 27$. Para $exper > 27$, um ano adicional de experiência profissional tem um efeito negativo no salário.

- (20) (e) Decidiu-se retirar o termo quadrático, $exper^2$, ao modelo da equação (2). A nova equação estimada é a seguinte:

$$\widehat{salario} = \underset{(0.110)}{0.950} + \underset{(0.009)}{0.050} exper + \underset{(0.020)}{0.098} educ. \quad (3)$$

$$SSR = 220, R^2 = 0.41$$

Apresente a estatística de teste que permite testar se o efeito parcial de *educ* é duas vezes superior ao efeito parcial da *exper*, efetue o teste estatístico ($\alpha = 0.05$) e mostre as suas conclusões, assumindo que $\hat{\sigma}_{\hat{\beta}_{educ}, \hat{\beta}_{exper}} = -0.001$ e $\alpha = 0.05$.

Resposta: Quer-se testar a $H_0 : \beta_{educ} = 2\beta_{exper}$, que é equivalente a $H_0 : \beta_{educ} - 2\beta_{exper} = 0$. Defina $\delta = \beta_{educ} - 2\beta_{exper}$. Quer-se, portanto, testar $H_0 : \delta = 0$ vs $H_0 : \delta > 0$

$$T = \frac{\hat{\delta}}{\hat{\sigma}_{\hat{\delta}}} \sim t_{(200-3-1)} \xrightarrow{d} N(0, 1)$$

$$\begin{aligned} \hat{\sigma}_{\hat{\delta}} &= \sqrt{Var(\hat{\beta}_{educ}) + Var(2\hat{\beta}_{exper}) - 2 \times Cov(\hat{\beta}_{educ}, 2\hat{\beta}_{exper})} \\ &= \sqrt{0.020^2 + 2^2 \times 0.009^2 - 4 \times (-0.001)} = \sqrt{0.004724} = 0.069 \end{aligned}$$

$$t_{obs} = \frac{0.098 - 2 \times 0.050}{0.069} = -0.029$$

$$W_{0.05} = \{t : t > 1.645\}$$

Não se rejeita a H_0 ao nível de 5%. Ou seja, não encontramos evidência estatística contra a afirmação “o efeito parcial de *educ* é duas vezes superior ao efeito parcial da *exper*”.