CONCURSO PÚBLICO DE PROVAS E TÍTULOS PARA O CARGO EFETIVO DE PROFESSOR DA CARREIRA DE MAGISTÉRIO SUPERIOR – EDITAL Nº 28/2023 – PROGRAD

FOLHA DE QUESTÕES

Área: **Número de C.P.F**___________________________

QUESTÃO 07: **(1 PONTO)** Seja X uma variável aleatória normalmente distribuída com média μ e variância, $\sigma^2 = 100$. Um pesquisador pretende testar as seguintes hipóteses:

$$H_0: \mu = 75 \quad \text{contra}$$
$$H_a: \mu < 75$$

Para isso, é retirada de X uma amostra aleatória de tamanho $n = 25$. Considerando que H_0 será rejeitada caso a média amostral seja menor ou igual a 70 (limite da Região Crítica, RC), determine:

a) **(0,4 PONTO)** o erro tipo I e sua probabilidade (α).
b) **(0,4 PONTO)** o erro tipo II e sua probabilidade (β), supondo o caso específico, em que $\mu = 68$, ou seja, $\beta(68)$.
c) **(0,2 PONTO)** Com base na amostra obtida da variável aleatória *X*, obteve-se $\bar{X} = 72$. Calcule o valor p do teste. Essa informação sugere que a média verdadeira seja menor que o valor-alvo estabelecido?

OBSERVAÇÃO: A Tabela da distribuição Normal padrão está anexada ao final da prova para consultas.

<u>CHAVE DE CORREÇÃO:</u>

a) A Região Crítica (RC) referente ao problema é dada por $RC = \{y \in \mathbb{R} |\ y \leq\ 70\}$. O erro de tipo I, pode ser descrito da seguinte forma:

Erro de tipo I: Dizer que a média é menor e igual a 70, enquanto na verdade, ela é de fato igual a 75.

Dessa forma, a probabilidade de se cometer o erro I pode ser determinada por:

$$P(\text{erro I}) \ = \ P(\ \bar{X} \ \in \ RC \mid \ H_0 \text{ é verdadeira}) = \alpha \ .$$

Sob H_0 verdadeira, tem-se que $\bar{X}$ possui distribuição normal com valor médio igual a 75 e desvio padrão dado por $\sigma_{\bar{X}} = \sigma/\sqrt{n} = \ 10/\sqrt{25} \ = \ 10/5 = 2$, isto é, $\bar{X} \sim N(\mu_{\bar{X}} = 75, \sigma^2_{\bar{X}} = 4)$. Assim, segue que:

$$\begin{aligned} \alpha = P(\text{erro I}) \ &= \ P(\bar{X} \ \leq \ 70 \ | \bar{X} \sim \ N(75,4)) \\ &= \ P\left(Z \leq \frac{70-75}{2}\right), \ \ Z \sim N(0,1) \\ &= \ P(Z \leq -2{,}5) = \Phi(-2{,}5) \ = \ 0{,}0062 \ = \ 0{,}62\ \% \end{aligned}$$

b) O erro tipo II pode ser descrito da seguinte forma:

Erro tipo II: Dizer que a média é maior que 70, sendo que de fato a média é igual a 68 ($\mu = 68$).

A probabilidade de se cometer o erro tipo II, é então calculada da seguinte forma:

$$P(\text{erro II}) = P(\bar{X} \notin RC \mid H_1 \text{ é verdadeira}) = \beta(68)$$

Sob H_1 verdadeira, tem-se que a v. a. $\bar{X}$ é tal que $\bar{X} \sim N(\mu_{\bar{X}} = 68, \sigma^2_{\bar{X}} = 4)$. Assim, segue que:

$$\begin{aligned} P(\text{erro II}) &= P(\bar{X} > 70 \mid \bar{X} \sim N(68, 4)) \\ &= P\left(Z > \frac{70-68}{2}\right) \\ &= 1 - P(Z \leq 1) \\ &= 1 - \Phi(1) \\ &= 1 - 0{,}8413 = 0{,}1586 = 15{,}86\% = \beta(68) \end{aligned}$$

c) Para o cálculo de valor p, segue pela definição que:

$$\text{valor } p = P(\bar{X} \leq 72 \mid \mu = 75) = P\left(Z \leq \frac{72-75}{2}\right) = P\left(Z \leq -\frac{3}{2}\right) = \Phi(-1{,}5) = 0{,}0668$$

$$\therefore \text{ valor } p = 0{,}0668.$$

E conclui-se que, estabelecido o valor do nível de significância α, o pesquisador não deve rejeitar a hipótese H_0 se o valor $p = 0{,}0668 > \alpha$. Caso contrário, se o valor $p = 0{,}0668 \leq \alpha$, rejeita -se a hipótese H_0 , em favor da Hipótese alternativa, H_a.

QUESTÃO 08: (1,5 PONTO) Um dos modelos mais empregados em análise de dados é o modelo de regressão linear. Quando temos somente uma variável auxiliar ou regressora x, temos o modelo de regressão linear simples, dado por: $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, em que Y_i é o valor da variável resposta Y, β_0 e β_1 são os parâmetros do modelo, x_i é o valor pré-fixado da variável explicativa x (x_i é constante conhecida), ε_i é um termo de erro aleatório, com média 0 e variância constante σ^2. Além disso, os erros aleatórios ε_i são não correlacionados, ou seja, $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$, para todo $i \neq j$, com $i, j = 1, \ldots, n$.

Frequentemente, assumimos que os erros possuem distribuição normal, ou seja, $\varepsilon_i \sim N(0, \sigma^2)$. Assim, pode-se mostrar que a distribuição de probabilidade de Y_i é dada por $Y_i | x_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$.

CONCURSO PÚBLICO DE PROVAS E TÍTULOS PARA O CARGO EFETIVO DE PROFESSOR DA CARREIRA DE MAGISTÉRIO SUPERIOR – EDITAL Nº 28/2023 – PROGRAD

FOLHA DE QUESTÕES

Área: **Número de C.P.F**______________________________

a) **(0,7 PONTO)** Seja $(x_1, Y_1), (x_2, Y_2), \ldots, (x_n, Y_n)$ uma amostra aleatória observada segundo o modelo de regressão linear simples apresentado, obtenha os estimadores de mínimos quadrados para os parâmetros β_0 e β_1 do modelo.

b) **(0,8 PONTO)** Mostre que os estimadores de mínimos quadrados $\hat{\beta}_0$ e $\hat{\beta}_1$ são não tendenciosos, ou seja, $E(\hat{\beta}_1) = \beta_1$ e $E(\hat{\beta}_0) = \beta_0$. Para tal, reescreva $\hat{\beta}_1$ como combinação linear de Y_i, ou seja, $\hat{\beta}_1 = \sum_{i=1}^{n} k_i Y_i$, com $k_i = \frac{(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$.

CHAVE DE CORREÇÃO:

a) Do modelo de regressão linear simples, temos que

$$\varepsilon_i = Y_i - \beta_0 - \beta_1 x_i.$$

O método de mínimos quadrados consiste em encontrar os valores de β_0 e β_1 que minimizam a soma dos quadrados dos erros, definida por

$$S = \sum_{i=1}^{n} {\varepsilon_i}^2 = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 x_i)^2.$$

Para encontrar esse mínimo, precisamos derivar a função *S* em relação a β_0 e β_1 e igualar a zero, ou seja, os estimadores de mínimos quadrados $\hat{\beta}_0$ e $\hat{\beta}_1$ devem satisfazer:

$$\left.\frac{\partial S}{\partial \beta_0}\right|_{\hat{\beta}_0, \hat{\beta}_1} = -2 \sum_{i=1}^{n} \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right) = 0$$

e

$$\left.\frac{\partial S}{\partial \beta_1}\right|_{\hat{\beta}_0, \hat{\beta}_1} = -2 \sum_{i=1}^{n} \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right) x_i = 0$$

Simplificando, temos:

$$\sum_{i=1}^{n} Y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{n} x_i = 0 \tag{1}$$

$$\sum_{i=1}^{n} x_i Y_i - \hat{\beta}_0 \sum_{i=1}^{n} x_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0, \tag{2}$$

o qual é denominador sistema de equações normais.

CONCURSO PÚBLICO DE PROVAS E TÍTULOS PARA O CARGO EFETIVO DE PROFESSOR DA CARREIRA DE MAGISTÉRIO SUPERIOR – EDITAL Nº 28/2023 – PROGRAD

FOLHA DE QUESTÕES

Área: **Número de C.P.F**___________________________

Pela equação (1), obtemos

$$n\hat{\beta}_0 = \sum_{i=1}^{n} Y_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i \Rightarrow \hat{\beta}_0 = \frac{\sum_{i=1}^{n} Y_i}{n} - \hat{\beta}_1 \frac{\sum_{i=1}^{n} x_i}{n}.$$

E substituindo na equação (2), temos

$$\sum_{i=1}^{n} x_i Y_i - \left(\frac{\sum_{i=1}^{n} Y_i}{n} - \hat{\beta}_1 \frac{\sum_{i=1}^{n} x_i}{n}\right) \sum_{i=1}^{n} x_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i Y_i - \frac{\sum_{i=1}^{n} Y_i \sum_{i=1}^{n} x_i}{n} + \hat{\beta}_1 \frac{(\sum_{i=1}^{n} x_i)^2}{n} - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i Y_i - \frac{\sum_{i=1}^{n} Y_i \sum_{i=1}^{n} x_i}{n} = \hat{\beta}_1 \left[\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}\right]$$

Portanto,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i Y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} Y_i}{n}}{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}}$$

e

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x},$$

onde $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ e $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

b)

$$E(\hat{\beta}_1) = E\left(\sum_{i=1}^{n} k_i Y_i\right) = E(k_1 Y_1 + k_2 Y_2 + \cdots + k_n Y_n) \Rightarrow$$

$$E(\hat{\beta}_1) = k_1 E(Y_1) + k_2 E(Y_2) + \cdots + k_n E(Y_n) = \sum_{i=1}^{n} k_i E(Y_i).$$

Conforme apresentado no enunciado, e pode-se mostrar também, $E(Y_i) = \beta_0 + \beta_1 x_i$. Logo,

$$E(\hat{\beta}_1) = \sum_{i=1}^{n} k_i E(Y_i) = \sum_{i=1}^{n} k_i(\beta_0 + \beta_1 x_i) = \beta_0 \sum_{i=1}^{n} k_i + \beta_1 \sum_{i=1}^{n} k_i x_i.$$

Temos que

$$\sum_{i=1}^{n} k_i = \sum_{i=1}^{n} \frac{(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sum_{i=1}^{n} x_i - n\bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = 0.$$

CONCURSO PÚBLICO DE PROVAS E TÍTULOS PARA O CARGO EFETIVO DE PROFESSOR DA CARREIRA DE MAGISTÉRIO SUPERIOR – EDITAL Nº 28/2023 – PROGRAD

FOLHA DE QUESTÕES

Área: **Número de C.P.F**____________________________

e

$$\sum_{i=1}^{n} k_i\, x_i = \sum_{i=1}^{n} \frac{(x_i - \bar{x})x_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})x_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sum_{i=1}^{n}(x_i^2 - \bar{x}x_i)}{\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

Desenvolvendo o numerador, temos:

$$\sum_{i=1}^{n}(x_i^2 - \bar{x}x_i) = \sum_{i=1}^{n} x_i^2 - \bar{x}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i^2 - \bar{x}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2.$$

e desenvolvendo o denominador, temos:

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2.$$

Então,

$$\sum_{i=1}^{n} k_i\, x_i = \frac{\sum_{i=1}^{n}(x_i^2 - \bar{x}x_i)}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = 1.$$

Assim, concluímos que

$$E(\hat{\beta}_1) = \beta_0 \sum_{i=1}^{n} k_i + \beta_1 \sum_{i=1}^{n} k_i\, x_i = \beta_1.$$

Agora, queremos mostrar que $E(\hat{\beta}_0) = \beta_0$.

$$E(\hat{\beta}_0) = E(\bar{Y} - \hat{\beta}_1\bar{x}) = E\left(\frac{\sum_{i=1}^{n} Y_i}{n}\right) - \bar{x}E(\hat{\beta}_1) =$$

$$= E\left(\frac{Y_1 + Y_2 + \cdots + Y_n}{n}\right) - \bar{x}E(\hat{\beta}_1) = \frac{1}{n}E(Y_1 + Y_2 + \cdots + Y_n) - \bar{x}E(\hat{\beta}_1) =$$

$$= \frac{1}{n}\sum_{i=1}^{n} E(Y_i) - \bar{x}E(\hat{\beta}_1) = \frac{1}{n}\sum_{i=1}^{n}(\beta_0 + \beta_1 x_i) - \bar{x}\beta_1 =$$

$$= \frac{1}{n}\left[n\beta_0 + \beta_1\sum_{i=1}^{n} x_i\right] - \beta_1\bar{x} = \beta_0 + \beta_1\bar{x} - \beta_1\bar{x} = \beta_0.$$