



EXAME NACIONAL DE SELEÇÃO 2025

PROVA DE ESTATÍSTICA

1º Dia: 18/09/2024 - QUARTA-FEIRA
HORÁRIO: 11h00m às 12h30m (horário de Brasília)

Instruções

1. Este **CADERNO** é constituído de **dez** questões **objetivas**.
2. Recomenda-se, nas questões apresentadas a seguir, não marcar ao acaso: cada item cuja resposta diverja do gabarito oficial acarretará a perda de $\frac{1}{n}$ ponto, em que n é o número de itens da questão a que pertença o item, conforme consta no Manual do Candidato.
3. Durante as provas, o(a) candidato(a) não deverá levantar-se ou comunicar-se com outras pessoas.
4. A duração da prova é de **uma hora e trinta minutos**, já incluído o tempo destinado à identificação do(a) candidato(a) – que será feita no decorrer da prova – e ao preenchimento da **FOLHA DE RESPOSTAS**.
5. Durante a realização das provas **não** é permitida a utilização de calculadora, equipamentos eletrônicos ou qualquer material de consulta.
6. A desobediência ao fiscal de prova ou a qualquer uma das recomendações constantes nas presentes Instruções e na **FOLHA DE RESPOSTAS** poderá implicar a anulação das provas do(a) candidato(a).
7. Só será permitida a saída de candidatos, levando o Caderno de Provas, **somente a partir de 1 hora após o início da prova** e nenhuma folha pode ser destacada.

AGENDA

- 23/09/2024 – 14 horas – Divulgação dos gabaritos das provas objetivas, no endereço: <http://www.anpec.org.br>.
- 23/09 a 24/09/2024 – Recursos identificados pelo autor serão aceitos até às 14h do dia 24/09 do corrente ano. Não serão aceitos recursos fora do padrão apresentado no Manual do Candidato.
- 28/10/2024 – 14 horas – Divulgação do resultado na Internet, no *site* acima citado.

OBSERVAÇÕES:

- Em nenhuma hipótese a ANPEC informará resultado por telefone.
- É **proibida** a reprodução total ou parcial deste material, por qualquer meio ou processo, sem autorização expressa da ANPEC.
- Nas questões de **1 a 10 (não numéricas)**, marque de acordo com a instrução de cada uma delas: itens **VERDADEIROS** na coluna **V** itens **FALSOS** na coluna **F** ou deixe a resposta **EM BRANCO**.
- Caso a **resposta seja numérica**, marque o dígito da **DEZENA** na coluna D e o dígito da **UNIDADE** na coluna U, ou deixe a resposta **EM BRANCO**.
- Atenção: o algarismo das **DEZENAS** deve ser obrigatoriamente marcado, mesmo que seja igual a **ZERO**.

QUESTÃO 01

Para essa questão, considere a seguinte notação: se X e Y são eventos de um espaço amostral Ω , $P(X)$ representa a probabilidade de ocorrência do evento X , $P(X|Y)$ representa a probabilidade de ocorrência do evento X condicionada à ocorrência do evento Y , e \bar{X} é o complemento de X . Julgue as alternativas abaixo como verdadeiras ou falsas.

- Ⓐ Sejam A e B eventos do espaço amostral de um experimento aleatório S . Se A e B são independentes, A e \bar{B} também são independentes.
- Ⓑ Sejam A e B eventos independentes do espaço amostral de um experimento aleatório S , onde $P(A) = \frac{1}{4}$ e $P(B) = \frac{1}{3}$. A probabilidade de que pelo menos um desses dois eventos (A e B) ocorra é $\frac{7}{12}$.
- Ⓒ Sejam A e B dois eventos do espaço amostral de um experimento aleatório S , onde $P(A) = \frac{2}{3}$, $P(B) = \frac{1}{3}$, e $P(\bar{A}|B) = \frac{1}{4}$. Então: $P(A|B) = P(B|A)$.
- Ⓓ Sejam A , B e C três eventos do mesmo espaço amostral de um experimento aleatório T , onde $P(A) = \frac{2}{5}$, $P(C) = \frac{1}{2}$, $P(A \cup B) = \frac{3}{4}$, $P(B|A) = \frac{3}{10}$ e $P(C|A) = \frac{1}{4}$. Então, $P(\bar{A}|C) = \frac{4}{5}$.
- Ⓔ Sejam A , B e C três eventos do mesmo espaço amostral de um experimento aleatório T , onde $P(A) = \frac{2}{5}$, $P(C) = \frac{1}{2}$, $P(A \cup B) = \frac{3}{4}$, $P(B|A) = \frac{3}{10}$ e $P(C|A) = \frac{1}{4}$. Então, $P(\bar{A}|\bar{C}) = \frac{1}{5}$.

QUESTÃO 02

Julgue as afirmativas a seguir como verdadeiras ou falsas:

- Ⓒ Uma estatística é uma função de valores de uma amostra enquanto um parâmetro descreve uma característica da população.
- ① Quanto maior o tamanho da amostra, menor a variância da média amostral.
- ② A média amostral é uma variável aleatória.
- ③ A média amostral é uma estatística, enquanto a média populacional é um estimador.
- ④ Para amostras de tamanho 100, de uma variável aleatória X , o desvio padrão da média amostral é igual a um centésimo do desvio padrão de X .

QUESTÃO 03

Sejam X e Y variáveis aleatórias com a seguinte função densidade de probabilidade conjunta:

$$f(x,y) = \begin{cases} 2(x+y-2xy) & \text{para } 0 \leq x \leq 1, \quad 0 \leq y \leq 1, \\ 0 & \text{caso contrário} \end{cases}$$

Julgue as afirmativas abaixo como verdadeiras ou falsas:

Ⓒ A variável aleatória X tem distribuição uniforme.

① $Prob\left(\frac{1}{4} \leq X \leq \frac{3}{4}\right) = \frac{1}{3}$

② $Prob\left(0 \leq Y \leq \frac{3}{4}\right) = \frac{1}{2}$

③ $E(X) = \frac{1}{2}$

④ $Var(X) = 2$

QUESTÃO 04

Suponha que em um grupo de 10 pessoas, 4 estejam desempregadas. Obtenha a probabilidade de que em uma amostra de 4 pessoas desse grupo, 2 pessoas estejam desempregadas. Multiplique o resultado por 100 e considere a parte inteira.

QUESTÃO 05

Sejam X_1, X_2, \dots, X_n variáveis aleatórias independentes e identicamente distribuídas com média μ_X e variância σ_X^2 , onde $\sigma_X^2 < \infty$. Além disso, as variáveis X_1, X_2, \dots, X_n têm distribuição normal. Considere que $plim$ representa o limite em probabilidade, e defina $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$.

Pela Lei dos Grandes Números, é correto afirmar:

- Ⓐ Mesmo se as variáveis aleatórias X_1, X_2, \dots, X_n não fossem normalmente distribuídas, teríamos $plim(\bar{X}) = \mu_X$.
- Ⓑ Defina $\omega = h(\mu_X)$, onde $h(\mu_X) = a + b\mu_X$, sendo a e b constantes positivas. Definindo $H = a + b\bar{X}$ como estimador para ω , temos $plim(H) = a + b\mu_X$.
- Ⓒ Sejam T_1, T_2, \dots, T_n variáveis aleatórias independentes e identicamente distribuídas com média μ_T e variância σ_T^2 , onde $\mu_T > 0$ e $\sigma_T^2 < \infty$. Se $\mu_T > \mu_X$, então:
 $plim\left(\frac{\bar{X}}{\bar{T}}\right) = 0$, onde $\bar{T} = \frac{\sum_{i=1}^n T_i}{n}$.
- Ⓓ Sejam Y_1, Y_2, \dots, Y_n variáveis aleatórias independentes e identicamente distribuídas com média μ_Y e variância σ_Y^2 , onde $\sigma_Y^2 < \infty$. Então, $plim(\bar{X} + \bar{Y}) = \mu_X + \mu_Y$, onde $\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}$.
- Ⓔ Sejam Z_1, Z_2, \dots, Z_n variáveis aleatórias independentes com distribuição de Bernoulli com parâmetro p , onde $0 < p < 1$. Definindo $\bar{Z} = \sum_{i=1}^n \frac{Z_i}{n}$, podemos dizer que a variância de \bar{Z} se aproxima de zero quando $n \rightarrow \infty$, e que $plim(\bar{Z}) = p$.

QUESTÃO 06

Suponha que uma amostra aleatória de n observações independentes X_1, X_2, \dots, X_n seja retirada de uma população com função densidade de probabilidade dada por:

$$f(x) = \begin{cases} \frac{x e^{-\frac{x}{\lambda}}}{\lambda^2}, & x > 0, \\ 0 & \text{caso contrário} \end{cases}$$

Onde λ é um parâmetro desconhecido, tal que $\lambda > 0$. Definindo \bar{X} como a média amostral, ou seja, $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$, é proposto o seguinte estimador para λ : $\hat{\lambda} = \frac{\bar{X}}{2}$. Usando essas informações, são corretas as afirmativas abaixo:

- Ⓒ Podemos dizer que $\hat{\lambda}$ é um estimador não tendencioso para λ .
- ① $Var(\hat{\lambda}) = \frac{\lambda}{4n}$
- ② $\hat{\lambda}$ é um estimador consistente para λ .
- ③ Considere o seguinte estimador para λ : $\tilde{\lambda} = \frac{\bar{X}}{3}$. Para $n = 4$, o Erro Quadrático Médio (EQM) de $\tilde{\lambda}$ é menor que o EQM de $\hat{\lambda}$.
- ④ Suponha que $n = 3$, e que sejam propostos os seguintes estimadores para λ :

$$\dot{\lambda} = \left(\frac{X_1}{4}\right) + \left(\frac{X_2}{8}\right) + \left(\frac{X_3}{8}\right)$$

$$\ddot{\lambda} = \left(\frac{X_1}{3}\right) + \left(\frac{X_2}{12}\right) + \left(\frac{X_3}{12}\right)$$

Podemos dizer que $\ddot{\lambda}$ é eficiente em relação a $\dot{\lambda}$ como estimador para o parâmetro λ .

QUESTÃO 07

Suponha que se deseja estimar os parâmetros β_0 , β_1 e β_2 na equação abaixo:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i.$$

Usando uma amostra aleatória da população com 23 observações, e estimando essa equação pelo método de Mínimos Quadrados Ordinários (MQO), são encontrados os seguintes resultados: $\hat{\beta}_1 = -4$ e $\hat{\beta}_2 = 2$, em que $\hat{\beta}_1$ é o estimador de MQO para o parâmetro β_1 , e $\hat{\beta}_2$ é o estimador de MQO para o parâmetro β_2 . Para essa mesma amostra, sabe-se também que: $\bar{Y} = 25$, $\bar{X}_1 = 4$, e $\bar{X}_2 = 2$, onde $\bar{Y} = \frac{\sum_{i=1}^{23} Y_i}{23}$, $\bar{X}_1 = \frac{\sum_{i=1}^{23} X_{1i}}{23}$, e $\bar{X}_2 = \frac{\sum_{i=1}^{23} X_{2i}}{23}$. Obtenha o resultado encontrado para $\hat{\beta}_0$ nessa mesma regressão por MQO, em que $\hat{\beta}_0$ é o estimador de MQO para β_0 .

QUESTÃO 08

Um pesquisador deseja estimar o seguinte modelo:

$$(1) Y = \beta_0 + \beta_1 Z + u,$$

Esse modelo satisfaz as seguintes condições: $E[u|Z] = 0$ e $Var[u|Z] = \sigma^2$. No entanto, a variável Z não é observada. O pesquisador decide, então, estimar o modelo de regressão linear representado pela equação (2) usando o método de Mínimos Quadrados Ordinários (MQO).

$$(2) Y = \alpha_0 + \alpha_1 X + \varepsilon,$$

A variável X está relacionada da seguinte maneira com a variável não observada Z :

$$(3) X = Z + w,$$

onde w tem média zero e variância σ_w^2 . Além disso, w é distribuído de maneira independente de u e de Z . Considere também que a variância populacional da variável não observada Z é igual a σ_Z^2 . Para estimar os parâmetros do modelo na equação (2), o pesquisador tem uma amostra aleatória da população com n observações $\{(X_i, Y_i, S_i) : i = 1, 2, \dots, n\}$, onde S é uma variável correlacionada com X . Julgue as afirmativas abaixo.

- Ⓒ Sendo $\hat{\alpha}_1$ o estimador de MQO para α_1 na equação (2), então o limite em probabilidade de $\hat{\alpha}_1$ é dado por:

$$\text{plim } \hat{\alpha}_1 = \beta_1 - \frac{\beta_1 \sigma_w^2}{\sigma_Z^2 + \sigma_w^2}$$

- Ⓐ O estimador de MQO para α_1 na equação (2) apresenta um viés de atenuação em relação ao parâmetro β_1 . Para um valor fixo de σ_w^2 , esse viés de atenuação aumenta à medida que σ_Z^2 aumenta.

- Ⓑ Defina $\hat{\gamma}_1 = \frac{\sum_{i=1}^n (S_i - \bar{S}) Y_i}{\sum_{i=1}^n (S_i - \bar{S}) X_i}$, onde $\bar{S} = \frac{\sum_{i=1}^n S_i}{n}$. Então, $\hat{\gamma}_1$ é um estimador consistente para o parâmetro β_1 .

- Ⓓ Definindo $\hat{\alpha}_0$ como o estimador de MQO para α_0 na equação (2), podemos dizer que o limite em probabilidade de $\hat{\alpha}_0$ é dado por:

$$\text{plim } \hat{\alpha}_0 = \beta_0 + \beta_1 \left(\frac{\sigma_w^2}{\sigma_Z^2 + \sigma_w^2} \right) \bar{X}, \text{ onde } \bar{X} = \frac{\sum_{i=1}^n X_i}{n}.$$

- Ⓔ Suponha que o pesquisador tenha acesso a uma amostra aleatória da população com n observações $\{(T_i, Y_i) : i = 1, 2, \dots, n\}$, e que a variável T seja tal que: $T = Z + v$. Suponha que v tenha média $\mu_v > 0$ e variância σ_v^2 , e que v seja distribuído de maneira independente de u e de Z . Sendo $\hat{\delta}_1$ o estimador de MQO para δ_1 na equação

$$Y = \delta_0 + \delta_1 T + \varphi, \text{ temos: } \text{plim } \hat{\delta}_1 = \beta_1 - \frac{\beta_1 \sigma_v^2}{\sigma_Z^2 + \sigma_v^2}$$

QUESTÃO 09

Julgue as afirmativas abaixo, a respeito dos estimadores de Mínimos Quadrados Ordinários (MQO), em um modelo de regressão linear múltipla, $y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + u$, com observações obtidas de uma amostra aleatória da população e que não apresentam multicolinearidade perfeita.

- Ⓐ A hipótese de que $V(u|X) = \sigma^2$ não é necessária para que os estimadores de mínimos quadrados ordinários sejam consistentes.
- Ⓑ A omissão da variável explicativa relevante para explicar a variável dependente torna a estimativa dos parâmetros $(\beta_1, \dots, \beta_k)$ tendenciosa e inconsistente, se e somente se, a variável omitida for correlacionada com a respectiva variável explicativa (x_1, \dots, x_k) incluída no modelo.
- Ⓒ Se o termo de erro for perfeitamente correlacionado com a variável explicativa é impossível obter o estimador de MQO.
- Ⓓ A presença de colinearidade imperfeita entre as variáveis explicativas gera estimadores ineficientes.
- Ⓔ Os estimadores de mínimos quadrados ordinários e de máxima verossimilhança coincidem quando os erros são independentes e identicamente distribuídos.

QUESTÃO 10

Considere os dois modelos de séries de tempo abaixo.

$$(I) Y_t = \alpha Y_{t-1} + u_t + \beta u_{t-1},$$

onde $0 < \alpha < 1$, Y_0 é um valor inicial não-aleatório para Y , e u_t é um ruído branco, que tem distribuição normal e satisfaz as seguintes condições: $E(u_t) = 0$ e $E(u_t^2) = \sigma^2 > 0$ para todo t , e $E(u_t u_s) = 0$ para $t \neq s$.

$$(II) Z_t = c + Z_{t-1} + \theta t + \varepsilon_t,$$

onde c é uma constante, Z_0 é um valor inicial não-aleatório para Z , $\theta > 0$, e ε_t é um ruído branco, que tem distribuição normal e satisfaz as seguintes condições: $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = \sigma^2 > 0$ para todo t , e $E(\varepsilon_t \varepsilon_s) = 0$ para $t \neq s$.

Julgue como verdadeiras ou falsas as afirmativas abaixo referentes a esses dois modelos:

Ⓐ Em relação ao modelo (I), podemos escrever:

$$\gamma_0 = \alpha \gamma_1 + \sigma^2 + \beta(\alpha + \beta)\sigma^2, \text{ onde } \gamma_0 = E(Y_t^2) \text{ e } \gamma_1 = E(Y_t Y_{t+1}).$$

Ⓑ Em relação ao modelo (I), podemos escrever:

$$\gamma_1 = \alpha(\gamma_0 + \sigma^2) + \beta\sigma^2, \text{ onde } \gamma_0 = E(Y_t^2) \text{ e } \gamma_1 = E(Y_t Y_{t+1}).$$

Ⓒ Sendo $\rho_h = \frac{\gamma_h}{\gamma_0}$, onde $\gamma_h = E(Y_t Y_{t+h})$, temos o seguinte resultado para o modelo (I):

$$\rho_h = \frac{(\beta + \alpha)(1 + \alpha\beta)}{(1 + 2\alpha\beta + \beta^2)}.$$

Ⓓ O modelo (II) pode ser representado por: $Z_t = ct + \left(\frac{\theta}{2}\right)t^2 + Z_0 + \sum_{j=1}^t \varepsilon_j$

Ⓔ Em relação ao modelo (II), a variância de Z_t é igual a: $Var(Z_t) = (c + \theta t)\sigma^2$.