



## **EXAME NACIONAL DE SELEÇÃO 2024**

### **PROVA DE ESTATÍSTICA**

**1º Dia: 27/09/2023 – QUARTA-FEIRA  
HORÁRIO: 11h00m às 12h30m (horário de Brasília)**

### Instruções

1. Este **CADERNO** é constituído de **dez** questões objetivas.
2. Nas questões do tipo A, recomenda-se 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 outros(as) candidatos(as).
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

8. 02/10/2023 – 14 horas – Divulgação dos gabaritos das provas objetivas, no endereço: <http://www.anpec.org.br>.
9. 02/10 a 03/10/2023 – Recursos identificados pelo autor serão aceitos até às 14h do dia 03/10 do corrente ano. Não serão aceitos recursos fora do padrão apresentado no Manual do Candidato.
10. 06/11/2023 – 14 horas – Divulgação do resultado na Internet, no *site* acima citado.

### OBSERVAÇÕES:

11. Em nenhuma hipótese a ANPEC informará resultado por telefone.
12. É **proibida** a reprodução total ou parcial deste material, por qualquer meio ou processo, sem autorização expressa da ANPEC.
13. 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**. Para **evitar a desclassificação** do candidato, **pelo menos um item de pelo menos uma questão** deve ser respondido na folha ótica de respostas.
14. 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**.
15. Atenção: o algarismo das **DEZENAS** deve ser obrigatoriamente marcado, mesmo que seja igual a **ZERO**.

## QUESTÃO 01

Em um concurso para professor do departamento de economia de determinada universidade, 12 candidatos foram aprovados e serão contratados ao longo de três anos. No primeiro ano, serão contratados 4 desses 12 candidatos. Os 12 candidatos estão divididos em 3 áreas: 6 em Microeconomia, 4 em Macroeconomia e 2 em Econometria. Os candidatos não estão classificados por nota, todos foram aprovados em igualdade de condições, e a discussão no departamento de economia é definir um critério para escolher os 4 primeiros candidatos a ingressar. Decide-se, então, fazer um sorteio aleatório para escolher esses 4 candidatos entre os 12 aprovados.

Defina as variáveis aleatórias  $X$ ,  $Y$  e  $Z$  da seguinte forma:

$X$  = número de candidatos na área de Microeconomia contratados no primeiro ano.

$Y$  = número de candidatos na área de Macroeconomia contratados no primeiro ano.

$Z$  = número de candidatos na área de Econometria contratados no primeiro ano.

São corretas as afirmativas:

- Ⓐ A probabilidade de ter exatamente 2 candidatos da área de Microeconomia e 1 de Macroeconomia entre os 4 contratados no primeiro ano é igual a  $\frac{8}{33}$ .
- Ⓑ ① A probabilidade de ter exatamente 2 candidatos da área de Microeconomia entre os 4 contratados no primeiro ano é igual a  $\frac{5}{11}$ .
- Ⓒ ② A probabilidade de ter pelo menos um candidato da área de Econometria entre os 4 contratados no primeiro ano é igual a  $\frac{12}{33}$ .
- Ⓓ ③  $E(X) = 1$
- Ⓔ ④ A probabilidade de ter exatamente 1 candidato da área de Microeconomia e 1 de Macroeconomia entre os 4 contratados no primeiro ano é igual a  $\frac{8}{165}$ .

## QUESTÃO 02

Suponha que  $X$  e  $Y$  sejam variáveis aleatórias independentes, com mesma média e variâncias  $V(X) = 0,5$  e  $V(Y) = 2$ . O valor da média populacional é desconhecido e se propõe estimá-lo a partir do estimador  $B$ :

$$B = aX + (1 - a)Y$$

Qual é o valor de  $a$  que produz o estimador mais eficiente? Multiplique o resultado por 5.

### QUESTÃO 03

Um pesquisador interessado em investigar o retorno educacional considera estimar o seguinte modelo:

$$\text{Salário}_i = \beta_0 + \beta_1 \text{Escolaridade}_i + u_i \quad (1)$$

Onde,  $\text{Salário}_i$  é dado em reais por hora e  $\text{Escolaridade}$  é dada em anos. O estimador de mínimos quadrados ordinários dos  $\beta$ 's são  $b$ 's. Supondo que o modelo verdadeiro (sob o qual vale a hipótese de exogeneidade das variáveis independentes) inclui uma variável de habilidade dos indivíduos ( $x_2$ ), julgue as afirmativas abaixo como verdadeiras ou falsas:

- Ⓐ O estimador de mínimos quadrados ordinários para  $\beta_1$  apresenta um viés de tamanho,  $\beta_2 \cdot d_{12}$ , onde  $d_{12}$  é o coeficiente de inclinação da regressão de habilidade em escolaridade.
- Ⓑ A variância do estimador da regressão simples para a equação 1 será sempre menor do que a variância da regressão múltipla para o modelo verdadeiro.
- Ⓒ A variância dos salários será dada por  $\beta_0^2 + \beta_1^2 V(\text{escolaridade}_i) + V(u_i)$ ; em que  $V$  = variância.
- Ⓓ A covariância amostral entre os valores preditos pela equação 1, estimada por mínimos quadrados ordinários, e os seus resíduos é sempre igual a zero.
- Ⓔ Se a correlação entre escolaridade e habilidade é igual a 0,95, a variância dos estimadores de mínimos quadrados ordinários para o modelo verdadeiro aumenta e este estimador será, portanto, não eficiente, mesmo se os erros forem homocedásticos e não correlacionados na cross-section.

## QUESTÃO 04

Considere que a distribuição da renda domiciliar do País A seja representada por uma variável aleatória  $X_A$  com distribuição lognormal, e que a distribuição da renda domiciliar do País B seja representada por uma variável aleatória  $X_B$ , também com distribuição lognormal. Suponha que  $\ln(X_A) = Y_A$ , onde  $Y_A$  tem média  $\theta_A$  e variância  $\gamma_A^2$ , e que  $\ln(X_B) = Y_B$ , onde  $Y_B$  tem média  $\theta_B$  e variância  $\gamma_B^2$ . Então, julgue as afirmativas abaixo como verdadeiras ou falsas. Para a resolução dessa questão, considere que  $\ln(10) = 2,30$ .

- Ⓐ Mesmo que tenhamos  $\theta_A > \theta_B$ , não podemos garantir que  $E(X_A) > E(X_B)$ .
- Ⓑ Se  $\theta_A = 2\theta_B$  e  $\gamma_A^2 = \gamma_B^2$ , então  $Var(X_A) = Var(X_B)$ .
- Ⓒ Suponha que  $\theta_A = 10$  e  $\gamma_A^2 = 1$ . A probabilidade de que um domicílio selecionado aleatoriamente no País A tenha renda domiciliar menor que 100.000 é superior a 0,90.
- Ⓓ Suponha ainda que  $\theta_A = 10$  e  $\gamma_A^2 = 1$ . Um domicílio selecionado aleatoriamente no País A estará entre os 1% mais ricos desse país se a sua renda  $x_A$  for tal que  $x_A \geq e^{9,80}$ .
- Ⓔ Em ambos os países, os domicílios com renda abaixo de 10.000 recebem um benefício do governo. Se  $\theta_A = 10$  e  $\gamma_A^2 = 1$ ; e,  $\theta_B = 12$  e  $\gamma_B^2 = 2$ , a probabilidade de um domicílio selecionado aleatoriamente no país B receber o benefício é maior que a probabilidade de um domicílio selecionado aleatoriamente no país A receber o benefício.

## QUESTÃO 05

Considere o modelo de regressão linear:  $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i$

Defina  $b_0$ ,  $b_1$  e  $b_2$  como os estimadores de Mínimos Quadrados Ordinários (MQO) para  $\beta_0$ ,  $\beta_1$  e  $\beta_2$ , respectivamente. Supondo que a equação acima tenha sido estimada pelo método de MQO usando uma amostra com  $n$  observações, julgue as afirmativas abaixo como verdadeiras ou falsas:

- Ⓒ  $\sum_{i=1}^N (y_i - b_0 - b_1 x_{i1} - b_2 x_{i2}) \neq 0$ .
- ① Seja  $z_i = a_0 + a_1 x_{i1} + a_2 x_{i2}$ , onde  $a_0$ ,  $a_1$  e  $a_2$  são constantes. Portanto, podemos afirmar que  $\sum_{i=1}^N [z_i(y_i - b_0 - b_1 x_{i1} - b_2 x_{i2})] = 0$ .
- ② Suponha que  $\sum_{i=1}^N x_{i2} > \sum_{i=1}^N x_{i1}$ . Então, é possível afirmar que  $\sum_{i=1}^N [x_{i2}(y_i - b_0 - b_1 x_{i1} - b_2 x_{i2})] > \sum_{i=1}^N [x_{i1}(y_i - b_0 - b_1 x_{i1} - b_2 x_{i2})]$ .
- ③  $\sum_{i=1}^N [y_i - b_1 x_{i1} - b_2 x_{i2}] > b_0$ .
- ④ Se  $x_{i2} = 5 + 4x_{i1}$ , então a hipótese de homocedasticidade fica comprometida.

## QUESTÃO 06

Em uma empresa que fabrica determinado utensílio doméstico, metade das peças produzidas apresentam algum defeito, gerando muitas reclamações dos clientes. Um novo processo de produção foi adotado para reduzir a taxa de defeitos. Para testar esse novo processo, é retirada uma amostra aleatória de 5 itens. Definindo  $Y$  como a soma do número de itens produzidos sem defeito nessa amostra, o teste proposto consiste em rejeitar  $H_0$  caso  $Y$  seja maior ou igual a 4, e não rejeitar  $H_0$  caso contrário. Supondo que  $Y$  tenha distribuição binomial com parâmetro  $p = \frac{1}{2}$ , obtenha o nível de significância desse teste. Multiplique o nível de significância por 100 e assinale a parte inteira.



## QUESTÃO 07

Um pesquisador deseja analisar a relação entre renda per capita ( $X$ ) e expectativa de vida ( $Y$ ) nas cidades brasileiras. Nessa análise, cada cidade é classificada de acordo com a população como pequena, média ou grande. O pesquisador pretende estimar dois modelos usando o método de Mínimos Quadrados Ordinários (MQO):

$$(1) Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 X_i + u_i.$$

$$(2) Y_i = \gamma_1 D_{1i} + \gamma_2 D_{2i} + \gamma_3 D_{3i} + \gamma_4 X_i + e_i.$$

Onde  $D_{1i} = 1$  se a cidade  $i$  é pequena, e igual a zero caso contrário;  $D_{2i} = 1$  se a cidade  $i$  é média, e igual a zero caso contrário; e  $D_{3i} = 1$  se a cidade  $i$  é grande, e igual a zero caso contrário;  $u_i$  e  $e_i$  são termos de erro das equações (1) e (2), tais que  $E[u_i | D_{2i}, D_{3i}, X_i] = 0$  e  $E[e_i | D_{1i}, D_{2i}, D_{3i}, X_i] = 0$ . Julgue as afirmativas abaixo como verdadeiras ou falsas:

- Ⓒ Se  $\beta_2 > 0$  na equação (1), podemos dizer que a expectativa de vida é maior, em média, nas cidades médias do que nas cidades pequenas.
- ① O parâmetro  $\beta_2$  na equação (1) é idêntico ao parâmetro  $\gamma_2$  na equação (2).
- ② O parâmetro  $\beta_4$  na equação (1) é idêntico ao parâmetro  $\gamma_4$  na equação (2).
- ③ Se  $\beta_3 > 0$  na equação (1), então  $\gamma_1 < \gamma_3$  na equação (2).
- ④ A equação (2) tem a vantagem em relação a equação (1) de permitir que seja obtido  $E[Y_i | D_1 = 1, X_i = x]$ .

## QUESTÃO 08

Considere o modelo de regressão linear abaixo:

$$(1) \quad y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u.$$

Onde,  $u$  é um termo de erro tal que  $E(u|x_1, x_2) = 0$  e  $Var(u|x_1, x_2) = \sigma^2$ .

Suponha que esteja disponível uma amostra aleatória da população com  $n$  observações,  $\{(x_{1i}, x_{2i}, y_i) : i = 1, 2, \dots, n\}$ , e que nenhuma das variáveis independentes seja constante. Suponha também que a correlação amostral entre  $x_1$  e  $x_2$  seja igual a zero. Decide-se, então, estimar também as duas equações abaixo pelo método de Mínimos Quadrados Ordinários (MQO):

$$(2) \quad y = \alpha_0 + \alpha_1 x_1 + e.$$

$$(3) \quad y = \delta_0 + \delta_2 x_2 + \epsilon.$$

Julgue as afirmativas abaixo como verdadeiras ou falsas:

⊙ Definindo  $\hat{\beta}_1$  como o estimador de MQO para  $\beta_1$  na equação (1) e  $\hat{\alpha}_1$  como o estimador de MQO para  $\alpha_1$  na equação (2),  $\hat{\beta}_1 = \hat{\alpha}_1$ .

① Definindo  $\hat{\beta}_0$  como o estimador de MQO para  $\beta_0$  na equação (1),  $\hat{\alpha}_0$  como o estimador de MQO para  $\alpha_0$  na equação (2), e  $\hat{\delta}_0$  como o estimador de MQO para  $\delta_0$  na equação (3),  $\hat{\beta}_0 = \hat{\alpha}_0 + \hat{\delta}_0$ .

② Definindo  $Var(\hat{\beta}_1|x_1, x_2)$  como a variância do estimador de MQO para  $\beta_1$  na equação (1) e  $Var(\hat{\alpha}_1|x_1)$  como a variância do estimador de MQO para  $\alpha_1$  na equação (2),  $Var(\hat{\beta}_1|x_1, x_2) = Var(\hat{\alpha}_1|x_1)$ .

③ Defina  $\hat{r}_{1i}$  como os resíduos de uma regressão simples de  $x_1$  em  $x_2$  (incluindo uma constante), usando essa mesma amostra. Então, podemos representar o estimador de MQO para  $\beta_1$  na equação (1) pela seguinte equação:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n \hat{r}_{1i} y_i}{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2}.$$

④ Defina  $R_1^2$  como o coeficiente de determinação da regressão correspondente a equação (1) e  $R_2^2$  como o coeficiente de determinação da regressão correspondente a equação (2). Então, escolhendo um nível de significância, podemos testar a hipótese nula:  $H_0: \beta_2 = 0$  contra a hipótese alternativa  $H_1: \beta_2 \neq 0$ , usando o fato de que  $\frac{(R_1^2 - R_2^2)}{(1 - R_1^2)/(n - 2)} \sim F_{1, n - 2}$ .

## QUESTÃO 09

Considere o seguinte modelo:

$$q = \theta_1 z + u \quad (1)$$

$$w = \beta_1 q + \beta_2 z + e \quad (2)$$

Em que,

$$E[u] = E[e] = 0$$

$$E[u^2] = \sigma_u^2, E[e^2] = \sigma_e^2, \text{Cov}[u, e] = \mu \neq 0$$

$$\text{Cov}[u, z] = \text{Cov}[e, z] = 0$$

Julgue as afirmativas abaixo como verdadeiras ou falsas:

- Ⓐ O estimador de mínimos quadrados ordinários de  $\theta_1$  é consistente.
- Ⓑ Os estimadores de mínimos quadrados ordinários dos  $\beta$ s são não viesados.
- Ⓒ A equação 1 é exatamente identificada e a equação 2 é sobreidentificada
- Ⓓ Se  $\mu = 0$ , tanto a equação 1 quanto a equação 2 são exatamente identificadas.
- Ⓔ Se  $\mu = 0$ , os estimadores de mínimos quadrados ordinários dos  $\beta$ s são não viesados.

## QUESTÃO 10

Julgue as afirmativas abaixo como verdadeiras ou falsas:

- Ⓒ Considere o modelo  $Y_t = \beta_0 + \beta_1 Y_{t-1} + e_t$ , onde  $\{e_t\}$  é um processo ruído branco. Se  $\beta_1 = 1$ , então  $Y_t$  é um processo estacionário somente no caso em que  $\beta_0 = 0$ .
- ① Considere o seguinte modelo AR(2)  $Y_t = \frac{3}{4}Y_{t-1} - \frac{1}{8}Y_{t-2} + u_t$ , onde  $\{u_t\}$  é um processo ruído branco. As raízes da equação característica são 2 e 4.
- ② O modelo  $Y_t = \frac{3}{4}Y_{t-1} - \frac{1}{8}Y_{t-2} + u_t$ , onde  $\{u_t\}$  é um processo ruído branco, é estacionário.
- ③ Suponha que  $Y_t$  seja um processo definido como:

$$\begin{cases} Y_t = X_t, & \text{se } t \text{ é par} \\ Y_t = X_t + 1, & \text{se } t \text{ é ímpar} \end{cases}$$

Onde,  $X_t$  é um processo estacionário. Então,  $Y_t$  também é um processo estacionário.

- ④ Considere o processo:  $Y_t = u_1 + u_2 + \dots + u_t$ , em que  $t = 1, 2, 3, \dots, T$ . Onde,  $\{u_t\}$  é um processo ruído branco. Então,  $Y_t$  é um processo estacionário.