anpec
associação nacional de
centros de pós graduação
em economia

EXAME NACIONAL DE SELEÇÃO 2021

PROVA DE ESTATÍSTICA

1º Dia: 21/10/2020 – QUARTA-FEIRA
HORÁRIO: 10h30m às 12h30m (horário de Brasília)

associação nacional de centros de pós graduação em economia

EXAME NACIONAL DE SELEÇÃO 2021

PROVA DE ESTATÍSTICA

1º Dia: 21/10 - QUARTA-FEIRA (Manhã)

HORÁRIO: 10h30m às 12h30m

INSTRUÇÕES

1. Esta **PROVA** é constituída de **quinze** questões objetivas.
2. Nas questões do tipo A, recomenda-se não marcar ao acaso: cada item cuja resposta divirja 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 **duas horas**.
5. Durante a realização das provas **não** é permitida a utilização de calculadora, qualquer material de consulta ou equipamentos eletrônicos além do utilizado para realização das provas.
6. Durante a realização das provas somente será permitida a saída do candidato após a autorização, por meio do *chat online*, do fiscal de prova.
7. O candidato só poderá desconectar-se, após o término da prova de cada disciplina.
8. Se a conexão cair, o candidato deve reiniciar a máquina. Caso a conexão não volte após o reinício da máquina, o candidato deve rotear a internet/wi-Fi de alguma pessoa próxima ou entrar em contato com o suporte técnico, cujo contato está no Comprovante de Inscrição.
9. A desobediência a qualquer uma das recomendações constantes nas presentes Instruções poderá implicar a anulação das provas do(a) candidato(a). A desobediência ao fiscal de prova também poderá implicar a anulação da prova do(a) candidato(a).

AGENDA

- 26/10/2020 – 14 horas – Divulgação dos gabaritos das provas objetivas, no endereço: **http://www.anpec.org.br.**
- 26/10 a 27/10/2020 – Recursos identificados pelo autor serão aceitos até às 14h do dia 27/10 do corrente ano. Não serão aceitos recursos fora do padrão apresentado no Manual do Candidato.
- 16/11/2020 – 14 horas – Divulgação do resultado na Internet, no *site* acima citado.

OBSERVAÇÕES:

- Em nenhuma hipótese a ANPEC informará resultado por telefone.
-
- Nas questões de **1** a **15** (não numéricas), marque, de acordo com o comando de cada uma delas: itens **VERDADEIROS**, marque **V**; itens **FALSOS**, marque **F**; ou deixe a resposta **EM BRANCO (SEM MARCAR)**.

- Caso a **resposta seja numérica**, marque os dígitos da **DEZENA** e da **UNIDADE**, ou deixe a resposta **EM BRANCO (SEM MARCAR)**. **Atenção:** o algarismo das **DEZENAS** deve ser obrigatoriamente marcado, mesmo que seja igual a **ZERO** (para números entre zero e nove, marque: 01, 02, 03, ..., 09).

QUESTÃO 01

Seja p_t^i o preço do bem i no período t, e seja q_t^i a quantidade vendida do bem i no período t. Considerando dois bens (i = 1, 2) e dois períodos (t = 1, 2), verifique se as afirmativas abaixo são corretas, supondo que $p_1^1 < p_2^1,\ p_1^2 < p_2^2,\ q_1^1 > q_2^1,\ q_1^2 > q_2^2$:

⓪ O Índice de Preço de Laspeyres do período 2 com base no período 1 é maior que um.

① O Índice de Preço de Paasche do período 2 com base no período 1 é maior que um.

② O Índice de Preço de Laspeyres do período 2 com base no período 1 é dado por:

$$\frac{r^1 v_2^1 + r^2 v_2^2}{v_1^1 + v_1^2},$$

em que $v_t^i = p_t^i q_t^i$ e $r^i = p_2^i / p_1^i$.

③ O Índice de Preço de Laspeyres do período 2 com base no período 1 é menor que o Índice de Preço de Paasche do período 2 com base no período 1.

④ O Índice de Preço de Paasche do período 2 com base no período 1 pode ser representado por:

$$\frac{v_2^1 + v_2^2}{\frac{v_2^1}{r^1} + \frac{v_2^2}{r^2}},$$

em que $v_t^i = p_t^i q_t^i$ e $r^i = p_2^i / p_1^i$.

QUESTÃO 02

A tabela de contingência a seguir apresenta os dados de uma amostra de 104 trabalhadores de uma empresa durante a pandemia, classificados segundo a decisão de trabalho e se estes trabalhadores moram ou não com pessoas classificadas no grupo de risco.

Decisão	Moram com pessoas do grupo de risco	Não moram com pessoas do grupo de risco	Total
Trabalham na empresa durante a pandemia	14	18	32
Trabalham em casa durante a pandemia	10	12	22
Não trabalham durante a pandemia	32	18	50
Total	56	48	104

Com base nestas informações, verifique as seguintes afirmações:

⓪ Se selecionarmos um trabalhador ao acaso, a probabilidade deste trabalhador não trabalhar durante a pandemia e não morar com pessoas do grupo de risco é igual a 36%.

① Se selecionarmos um trabalhador ao acaso, a probabilidade deste trabalhador morar com uma pessoa classificada no grupo de risco é de, aproximadamente, 46%.

② Entre os trabalhadores que trabalham durante a pandemia, 22% trabalham em casa e moram com pessoas classificadas no grupo de risco.

③ A probabilidade de que trabalhadores que moram com pessoas que não estão classificadas no grupo de risco continuar trabalhando é de, aproximadamente, 19,6 p.p. maior do que trabalhadores que moram com pessoas classificadas no grupo de risco e continuam trabalhando.

④ Sabendo que um trabalhador optou por continuar trabalhando, a probabilidade deste trabalhador não morar com pessoas classificadas no grupo de risco é de, aproximadamente, 55,5%.

QUESTÃO 03

Seja X uma variável aleatória com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \dfrac{x^2}{3}, & -1 \leq x \leq 2, \\ 0, & caso\ contrário \end{cases}$$

Encontre o valor esperado de $h(X) = 4X + 3$.

QUESTÃO 04

Sejam X, Y e Z três variáveis aleatórias. Definindo $cov(A, B)$ como a covariância entre as variáveis A e B, julgue as proposições:

⓪ Sendo V = Y + Z, então $cov(X, V) = cov(X, Y) + cov(X, Z)$.

① $cov(X, 2Y) = 4cov(X, Y)$.

② $cov(X, 4) = 0$.

③ Sendo W = 2Y + 3Z, então $cov(X, W) = 2cov(X, Y) + 3cov(X, Z)$.

④ Sendo T = 4 + 2Z, então $cov(X, T) = 4cov(X, Z)$.

QUESTÃO 05

Após análise dos últimos 2 anos da demanda por seu produto, uma empresa concluiu o seguinte: se um consumidor adquiriu seu produto em 2018, a probabilidade deste consumidor voltar a consumir em 2019 é de 40%. Por outro lado, a probabilidade deste consumidor adquirir seu produto pela primeira vez em 2019 é de 70%. Imaginando que as preferências dos consumidores por seu produto sejam estáveis, a empresa gostaria de saber a probabilidade de um consumidor que não adquiriu seu produto em 2018 passe a consumi-lo em 2020. Multiplique o resultado por 100.

QUESTÃO 06

Considere o processo ARMA(3,1):

$$\boldsymbol{x_t = 2 + 0,2x_{t-1} + 0{,}5x_{t-2} - 0,25x_{t-3} + 0,75\eta_{t-1} + \eta_t}.$$

Sabendo que $\boldsymbol{x_0 = 10}$, $\boldsymbol{x_1 = 8}$, $\boldsymbol{x_2 = 12}$ e $\boldsymbol{\eta_2 = 2}$. Qual o valor esperado de $\boldsymbol{x_5}$? Multiplique o resultado por 10 e marque a parte inteira.

QUESTÃO 07

Considere as principais distribuições de probabilidade e julgue as afirmativas:

⓪ A distribuição de probabilidade hipergeométrica é um caso particular da distribuição binomial.

① A distribuição t-Student é um caso particular da distribuição Normal.

② Seja X uma variável aleatória com distribuição Qui-quadrado com n graus de liberdade, então $Y = X^2$ segue uma distribuição $F_{(1,n)}$.

③ Seja X uma variável aleatória com distribuição log-Normal, então $Y = \ln(X)$ segue uma distribuição Normal.

④ Se $W_1, \cdots, W_n$ são variáveis aleatórias independentes e identicamente distribuídas, com distribuição Normal, então $Y = \sum_{i=1}^{n} W_1$ tem distribuição Qui-quadrado com n graus de liberdade.

QUESTÃO 08

Considere uma variável aleatória Y com média igual a 12 e variância igual a 4. Considere também que, usando o Teorema de Tchebycheff, temos:

$Prob(|Y - 12| < 10) \geq c$.

Calcule o valor de c e multiplique o resultado por 100.

QUESTÃO 09

Seja Y uma variável aleatória com distribuição χ^2 com k graus de liberdade. Defina μ como a média de Y. Para estimar 2μ, é proposto o seguinte estimador baseado em uma amostra aleatória da população $Y = (Y_1, Y_2, \dots, Y_n)$:

$\phi(Y) = \phi(Y_1, Y_2, \dots, Y_n) = (2\bar{Y}) - 1$, em que $\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$.

Considerando, portanto, que Y_i é independente de Y_j para $i \neq j$, julgue as afirmativas:

⓪ $E[\phi(\boldsymbol{Y})] = 2k - 1$.

① $\phi(\boldsymbol{Y})$ é um estimador viesado de 2μ.

② O estimador $\phi(\boldsymbol{Y})$ tem variância igual a $\frac{2k}{n}$.

③ Quando $n \to \infty$, o erro quadrático médio de $\phi(\boldsymbol{Y})$ tende para zero.

④ Considere outro estimador para 2μ, também baseado em uma amostra aleatória da população

$\boldsymbol{Y} = (Y_1, Y_2, \dots., Y_n)$: $\psi(\boldsymbol{Y}) = \psi(Y_1, Y_2, \dots., Y_n) = 2(\bar{Y} - 1)$, em que $\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$. Quando $n \to \infty$, $\{E[\psi(\boldsymbol{Y})] - 2\mu\}$ tende para zero.

QUESTÃO 10

Considere o modelo de regressão linear múltipla:

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$.

Suponha que está disponível uma amostra aleatória da população com n observações, $\{(x_{1i}, x_{2i}, x_{3i}, y_i): i = 1,2, \dots, n\}$, que nenhuma das variáveis independentes seja constante, e que não existam relações lineares entre as variáveis independentes. Defina $\hat{\beta}_j$ como o estimador de Mínimos Quadrados Ordinários (MQO) de β_j para $j = 1,2,3,$ em uma regressão de y em x_1, x_2, x_3 e uma constante. Considerando também que $E(u|x_1, x_2, x_3) = 0$, julgue as afirmativas:

⓪ Se $Var(u|x_1, x_2, x_3) = \sigma^2 x_{1i}$, então $\hat{\beta}_1$ é um estimador tendencioso de β_1.

① Se $Var(u|x_1, x_2, x_3) = \sigma^2 x_{1i}$, então $\hat{\beta}_1$ não é estimador consistente de β_1.

② Se $Var(u|x_1, x_2, x_3) = \sigma^2$, então $\hat{\beta}_1$ tem distribuição normal.

③ $\hat{\beta}_1$ é um estimador não-tendencioso de β_1, mesmo que β_2 seja igual a zero.

④ Se $Var(u|x_1, x_2, x_3) = \sigma^2$ e $u \sim Normal(0, \sigma^2)$, então $\hat{\beta}_1 \sim t_{n-4}$.

QUESTÃO 11

Considere o seguinte modelo de regressão linear simples:

$y_i = \beta_0 + \beta_1 x_i + u_i, \; i = 1,2, \dots\dots n$.

Para uma amostra com 11 observações, são obtidos os seguintes resultados:

$\sum_{i=1}^{11} x_i = 0$, $\sum_{i=1}^{11} y_i = 0$, $\sum_{i=1}^{11} x_i^2 = A$, $\sum_{i=1}^{11} y_i^2 = B$, $\sum_{i=1}^{11} x_i y_i = C$.

Suponha que esse modelo tenha sido estimado pelo método de Mínimos Quadrados Ordinários (MQO) usando essa amostra com 11 observações. Sendo $\hat{\beta}_0$ e $\hat{\beta}_1$ os estimadores de MQO para β_0 e β_1, respectivamente, e $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$, são corretas as afirmativas:

⓪ $\hat{\beta}_0 = 0$.

① $\hat{\beta}_1 = \frac{C}{A}$.

② $\sum_{i=1}^{11} (\hat{y}_i - \bar{y})^2 = \frac{C}{AB}$.

③ $\sum_{i=1}^{11} (y_i - \hat{y}_i)^2 = \frac{B-C}{A}$.

④ O coeficiente de determinação dessa regressão é: $R^2 = \frac{(A - C^2)}{AB}$.

QUESTÃO 12

Considere verdadeiro o modelo de regressão populacional $y_i = 5 + 10x_{1i} + 1{,}5x_{2i} + \varepsilon_i$ e considere que as suposições clássicas de Gauss-Markov sejam satisfeitas. No entanto, o modelo $y_i = \theta_0 + \theta_1 x_{1i} + u_i$ foi estimado por Mínimos Quadrados Ordinários. A covariância entre x_1 e x_2 é igual a 50, a variância de x_1 é igual a 30 e a variância de x_2 é igual a 15. Qual é o viés do estimador $\hat{\theta}_1$? Multiplique o resultado por 10 e marque a parte inteira.

QUESTÃO 13

Considere o sistema de duas equações simultâneas, em que a variável y_1 aparece no lado esquerdo das equações de oferta e demanda:

(I) $y_1 = \alpha_1 y_2 + \beta_1 z_1 + u_1$;

(II) $y_1 = \alpha_2 y_2 + \beta_2 z_2 + u_2$.

Suponha que z_1 seja não correlacionada com os termos aleatórios u_1e u_2, e que a variável z_2 também seja não correlacionada com u_1e u_2. Portanto, y_1 e y_2 são variáveis endógenas do sistema e z_1 e z_2 são variáveis exógenas do sistema. Julgue as afirmativas:

⓪ Se $\alpha_1 = 0$ e $\alpha_2 \neq 0$, a forma reduzida para y_1 é: $y_1 = \beta_1 z_1 + u_1$.

① Se $\alpha_1 \neq 0$ e $\alpha_2 = 0$, a forma reduzida para y_2 é: $y_2 = \frac{\beta_2}{\alpha_1} z_2 - \frac{\beta_1}{\alpha_1} z_1 + \frac{(u_2 - u_1)}{\alpha_1}$.

② Se $\alpha_1 \neq 0$, $\alpha_2 \neq 0$ e $\alpha_1 \neq \alpha_2$, a forma reduzida para y_1 é:

$$y_1 = \left(\frac{\alpha_1}{\alpha_1 - \alpha_2}\right) \beta_2 z_2 - \left(\frac{\beta_1}{\alpha_1 - \alpha_2}\right) z_1 - \left(\frac{1}{\alpha_1 - \alpha_2}\right) u_1 + \left(\frac{\alpha_1}{\alpha_1 - \alpha_2}\right) u_2.$$

③ Se $\alpha_1 \neq 0$, $\alpha_2 \neq 0$ e $\alpha_1 \neq \alpha_2$, a forma reduzida para y_2 é:

$$y_2 = \frac{\beta_1}{(\alpha_2 - \alpha_1)} z_1 - \frac{\beta_2}{(\alpha_2 - \alpha_1)} z_2 + \frac{1}{(\alpha_2 - \alpha_1)} (u_1 - u_2).$$

④ Se $\alpha_1 > 0$ e $\alpha_2 > 0$, não existem formas reduzidas para y_1e y_2.

QUESTÃO 14

Considere o seguinte processo:

$X_t = Y_t + 0{,}5Y_{t-1} - 0{,}2Y_{t-2}$, em que Y_t é um ruído branco, com distribuição normal, e satisfazendo as condições: E(Y_t) = 0, Var(Y_t) = σ^2 e E($Y_t Y_s$) = 0 para t ≠ s.

São corretas as seguintes afirmativas:

⓪ E(X_t) = 0.

① Var(X_t) = 1,21σ^2.

② Cov(X_t, X_{t-1}) = 0,4σ^2.

③ Cov(X_t, X_{t-2}) = 0.

④ Cov(X_t, X_{t-3}) = 0.

QUESTÃO 15

O seguinte modelo de regressão múltipla foi estimado por Mínimos Quadrados Ordinários com o objetivo de fazer previsões para o preço de 88 imóveis de uma amostra aleatória:

$$\boldsymbol{ln(preço)} = \underset{(\mathbf{0,51})}{\mathbf{0,26}} + \underset{(\mathbf{0,11})}{\mathbf{1,05}}\boldsymbol{ln(preço_{aval})} - \underset{(\mathbf{0,12})}{\mathbf{0,10}}\boldsymbol{ln(area)} + \underset{(\mathbf{0,02})}{\mathbf{0,02}}\boldsymbol{dorm} + \underset{(\mathbf{0,03})}{\mathbf{0,04}}\boldsymbol{col} + \boldsymbol{\hat{u}}$$

$$\boldsymbol{R^2 = 0,77} \quad \boldsymbol{SST = 8,01} \quad \boldsymbol{\hat{\sigma} = 0,14},$$

em que $\boldsymbol{preço}$ e $\boldsymbol{preço\_aval}$ são, respectivamente, preço de venda e preço de avaliação do imóvel, em milhares de dólares; $\boldsymbol{area}$ é o tamanho da área construída do imóvel, em pés quadrados; $\boldsymbol{dorm}$ é a quantidade de dormitórios do imóvel; e $\boldsymbol{col}$ é uma variável que indica se o imóvel possui estilo colonial. Os erros-padrão estão em parênteses.

Para a resolução desta questão talvez seja útil saber que se Z tem distribuição normal padrão, então P(|Z| > 1,645) = 0,10 e P(|Z| > 1,96) = 0,05. Julgue as afirmativas:

⓪ Ao nível de significância de 10%, é possível afirmar que a variável preço de avaliação (em log) não é relevante para explicar o preço de venda do imóvel (em log).

① O modelo $\boldsymbol{ln(\widehat{preço}) = \hat{\alpha}_0 + \hat{\alpha}_1 ln(preço\_aval) + \hat{\alpha}_2 dorm + \hat{\alpha}_3 col}$ possui maior R^2 ajustado que o modelo inicialmente estimado.

② A soma de quadrados dos resíduos (SSR) é igual a 1,84.

③ A estatística de significância da regressão é igual a 69,47.

④ Para dois imóveis com as mesmas características e o mesmo preço de avaliação, é esperada uma diferença de, aproximadamente, 4% dos preços de venda devido ao estilo do imóvel.