



EXAME NACIONAL DE SELEÇÃO 2027

PROVA DE ESTATÍSTICA

1º Dia: 22/09/2026 - TERÇA-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 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 outros(as) candidatos(as).
4. A duração da prova é de **uma hora e trinta minutos**, já incluído o tempo destinado à identificação – 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 ou qualquer material de consulta.
6. A desobediência a qualquer uma das recomendações constantes nas presentes Instruções e na **FOLHA DE RESPOSTAS** 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).
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

- 25/09/2026 – 14 horas – Divulgação dos gabaritos das provas objetivas, no endereço: <http://www.anpec.org.br>.
- 25/09 a 28/09/2026 – Recursos identificados pelo autor serão aceitos até às 14h do dia 28/09 do corrente ano. Não serão aceitos recursos fora do padrão apresentado no Manual do Candidato.
- 21/10/2026 – 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**.

Para evitar a desclassificação do candidato, pelo menos um item de pelo menos uma questão deve ser respondido na folha ótica de respostas

QUESTÃO 01

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

$$f(x, y) = \begin{cases} 3(x + y) & \text{para } x > 0, y > 0, x + y < 1 \\ 0 & \text{caso contrário} \end{cases}$$

Obtenha o valor esperado de X e multiplique por 80.

QUESTÃO 02

Sejam X_1 , X_2 e X_3 variáveis aleatórias independentes, cada uma com distribuição normal padrão, ou seja, $X_i \sim N(0,1)$ para $i = 1, 2, 3$. Definindo as variáveis $Z_1 = \frac{X_1 + X_2}{\sqrt{2}}$, $Z_2 = X_1^2 + X_2^2$, e $Z_3 = X_1^2 + X_2^2 + X_3^2$, indique se as afirmativas abaixo são verdadeiras ou falsas.

- Ⓐ A variável aleatória Z_1 tem distribuição normal com média igual a zero e variância igual a dois.
- Ⓑ A variável aleatória Z_3 tem distribuição qui-quadrado com média igual a 3 e variância igual a 6.
- Ⓒ Definindo $U = \frac{Z_1}{\sqrt{\frac{Z_3}{3}}}$, podemos dizer que U tem distribuição t de Student com 3 graus de liberdade, e que U tem média igual a 0 e variância igual a 1.
- Ⓓ Definindo $S = \frac{Z_2/2}{Z_3/3}$, podemos dizer que $S \sim F_{2,3}$.
- Ⓔ Para $U = \frac{Z_1}{\sqrt{\frac{Z_3}{3}}}$, é definida a variável $V = U^2$. A variável V é tal que $V \sim F_{1,3}$.

QUESTÃO 03

Verifique se as afirmativas abaixo são verdadeiras ou falsas.

- ⓐ Suponha que 10 pontos tenham sido escolhidos de maneira aleatória e independente no intervalo $[0,1]$. Sendo b um número tal que $0 < b < 1$, a probabilidade de apenas um dos 10 pontos pertencer ao intervalo $[0, b]$ é dada por: $b(1 - b)^9$.

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

$$f(x) = \begin{cases} 1 & \text{se } 0 \leq x \leq 1 \\ 0, & \text{caso contrário} \end{cases}$$

Se $Y = X^2$, a função densidade de probabilidade de Y é dada por:

$$g(y) = \begin{cases} y^{-\left(\frac{1}{2}\right)} & \text{se } 0 \leq y \leq 1 \\ 0, & \text{caso contrário} \end{cases}$$

- ⓒ Sejam X e Y duas variáveis aleatórias independentes, cada uma delas com distribuição uniforme no intervalo entre 0 e 1. Se $Z = \min(X, Y)$, então $E(Z) = \frac{1}{3}$.

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

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

Definindo $Y = X^2$, podemos dizer que a função densidade de probabilidade de Y é dada por:

$$g(y) = \begin{cases} \frac{1}{2} e^{-\left(y^{\frac{1}{2}}\right)} & \text{se } y > 0 \\ 0, & \text{caso contrário} \end{cases}$$

- ⓔ Sejam X , Y e Z variáveis aleatórias independentes, cada uma delas com distribuição uniforme no intervalo entre 0 e 1. Podemos dizer que $\text{Prob}(X < Y < Z) = \frac{1}{8}$.

QUESTÃO 04

Julgue as afirmativas abaixo como verdadeiras ou falsas:

- Ⓒ Seja X uma variável aleatória com distribuição geométrica com parâmetro $p = \frac{1}{4}$. Então, pelo Teorema de Tchebycheff, podemos dizer que:
- $$Prob(|X - E(X)| \geq 6) \leq \frac{1}{3}.$$
- ① Considere o lançamento de uma moeda equilibrada (com probabilidades de 50% de sair “cara” e de 50% de sair “coroa”). Aplicando o Teorema Central do Limite, a probabilidade de saírem mais de 55 caras em 100 lançamentos é maior que a probabilidade de saírem mais de 220 caras em 400 lançamentos.
- ② Considere um determinado dispositivo eletrônico desenvolvido pela firma A. A duração desse dispositivo segue uma distribuição normal com média desconhecida igual a m horas e variância conhecida igual a 10.000 horas. Suponha que as durações de diferentes aparelhos desse tipo sejam independentes. Aplicando o Teorema Central do Limite, precisamos de no mínimo 50 aparelhos para que a duração média estimada tenha um nível de confiança de 95% tal que a margem de erro não exceda 20 horas.
- ③ Suponha que o número de maçãs produzidas por uma macieira no ano seja representado por uma variável aleatória com valor esperado igual a 50. Se essa variável aleatória tem variância igual a 25, a probabilidade dessa macieira produzir 75 ou mais maçãs no ano é inferior a 2%, de acordo com o Teorema de Tchebycheff.
- ④ Suponha que 20% dos consumidores se declarem influenciados por propagandas na internet. Se 100 consumidores são selecionados aleatoriamente, a probabilidade de que pelo menos 28 se declarem influenciados por propagandas na internet é menor que 0,05, de acordo com o Teorema Central do Limite.

QUESTÃO 05

Seja Y uma variável aleatória com média $\frac{1-p}{p}$ e variância $\frac{1-p}{p^2}$, onde $0 < p < 1$. Suponha que $Y_1, Y_2, Y_3, \dots, Y_n$ seja uma amostra aleatória de tamanho n retirada dessa população.

Definindo $\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}$, indique se as alternativas a seguir são verdadeiras ou falsas:

Ⓒ $\bar{Y}$ é um estimador não tendencioso para $\frac{1}{p}$.

① $\hat{p} = \frac{1}{\bar{Y}+1}$ é um estimador não tendencioso para p .

② $\bar{Y} + 1$ é um estimador tendencioso para $\frac{1}{p}$.

③ A variância de $\bar{Y} + 1$ é igual a $\frac{(1-p)}{np^2}$.

④ $\bar{Y} + 1$ é um estimador consistente para $\frac{1}{p}$.

QUESTÃO 06

Considere o seguinte modelo linear para uma relação entre variáveis Y , X_1 e X_2 numa população de interesse:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + U, \quad E[U|X_1, X_2] = 0, \quad V[U|X_1, X_2] = \sigma_U^2,$$

onde, ademais, sabemos que X_2 é *independente* de X_1 . Suponha que o pesquisador obteve acesso a uma amostra aleatória $\{Y_i, X_{i,1}, X_{i,2}\}_{i=1}^n$ da população em que vale o modelo de interesse, e que na amostra os regressores $X_{i,1}$, $X_{i,2}$ e o intercepto não exibem colinearidade perfeita. Suponha, ademais, que todas as variáveis aleatórias têm variância positiva e finita.

No que segue, indique se as alternativas são verdadeiras ou falsas:

- Ⓐ O estimador de mínimos quadrados ordinários de β_1 computado a partir da especificação que omite X_2 da regressão não apresentará viés somente se $\beta_2 = 0$.
- Ⓑ Se $\beta_2 \neq 0$, a variância do estimador de mínimos quadrados de β_1 da especificação que omite X_2 da regressão não pode, no geral, ser ranqueada com respeito à variância do estimador de β_1 da especificação que inclui X_2 .
- Ⓒ Se o tamanho amostral n é suficientemente grande e $\beta_2 \neq 0$, esperamos que a variância do estimador de mínimos quadrados ordinários de β_1 da especificação que omite X_2 da regressão seja inferior à do estimador de β_1 da especificação que inclui X_2 .
- Ⓓ Se $\beta_2 = 0$, a variância do estimador de β_1 da especificação que omite X_2 da regressão é inferior à variância do estimador de β_1 da especificação que inclui X_2 .
- Ⓔ A hipótese de independência entre X_1 e X_2 pode ser falseada a partir da estimação e teste de hipótese numa especificação linear que regride X_2 em um intercepto e em X_1 , mas não em uma especificação que regride X_1 em um intercepto, em X_2 e em X_2^2 .

QUESTÃO 07

Considere o seguinte modelo linear para uma relação entre duas variáveis, Y e X , numa população de interesse:

$$Y = \beta_0 + \beta_1 X + U, \quad E[U|X] = 0.$$

Denote, ademais, a função variância condicional por $g(x) = V[U|X = x]$.

Suponha que se tem acesso a uma amostra aleatória $\{Y_i, X_i\}_{i=1}^n$ da população na qual vale o modelo de interesse e que na amostra o regressor X_i varie entre observações. No que segue, julgue como verdadeiras ou falsas as asserções sobre o estimador de mínimos quadrados ponderados que usa como pesos individuais valores w_i , $i = 1, \dots, n$:

Ⓐ O estimador de β_1 obtido por mínimos quadrados ponderados é dado por:

$$\frac{\sum_{i=1}^n w_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n w_i (X_i - \bar{X})^2}, \text{ onde } \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \text{ e } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

- Ⓑ Se $g(x) = 10$, o estimador de mínimos quadrados ponderados que toma como pesos $w_i = \exp(-X_i)$ apresentará viés.
- Ⓒ Se $g(x) = \exp(x)$, o estimador de mínimos quadrados ponderados que toma como pesos $w_i = \exp(-X_i)$ é um estimador linear não viciado de variância mínima.
- Ⓓ Suponha que $g(x) = \gamma_0 + \gamma_1 X_i$. Considere a escolha de pesos $w_i = \hat{\gamma}_0 + \hat{\gamma}_1 X_i$, onde $(\hat{\gamma}_0, \hat{\gamma}_1)$ é obtido em duas etapas: primeiro, calculam-se os resíduos $\hat{U}_i$ obtidos através da estimação por mínimos quadrados ordinários de (β_0, β_1) ; em seguida, estima-se $(\hat{\gamma}_0, \hat{\gamma}_1)$ regredindo-se $\hat{U}_i^2$ num intercepto e em X_i . O estimador de mínimos quadrados ponderados que usa esses pesos é um estimador linear não viciado de variância mínima.
- Ⓔ Seja $g(x) = \tau^2$ para algum $\tau^2 > 0$. Considere a escolha de pesos $w_i = \hat{\tau}^2$, onde $\hat{\tau}^2$ é calculado a partir da variância amostral dos resíduos obtidos da estimação preliminar de (β_0, β_1) por mínimos quadrados ordinários. O estimador de mínimos quadrados ponderados que usa esses pesos é um estimador linear não viciado de variância mínima.

QUESTÃO 08

Um pesquisador estima a seguinte especificação linear, com base em uma amostra de $n = 155$ observações do modelo de interesse:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + V.$$

A estimação da seguinte especificação resulta em um R^2 de 0,6.

Em seguida, o pesquisador define uma variável auxiliar $Z = Y - 2X_4$ e regride-a contra um intercepto, X_1 e X_2 . Nesse caso, o R^2 reportado pelo pacote econométrico utilizado na análise é de 0,4. Por fim, o pesquisador calcula as seguintes somas de quadrados:

$$S_1 = \sum_{i=1}^n (Y_i - \bar{Y})^2 = 500,$$
$$S_2 = \sum_{i=1}^n (Z_i - \bar{Z})^2 = 400,$$

onde $\bar{Y}$ e $\bar{Z}$ denotam, respectivamente, as médias amostrais das variáveis Y e Z .

Com base nas informações acima descritas, calcule o valor da estatística F do teste da hipótese nula conjunta $H_0: \beta_3 = 0$ e $\beta_4 = 2$ contra a alternativa $H_1: \beta_3 \neq 0$ ou $\beta_4 \neq 2$.

QUESTÃO 09

Considere um problema em que um pesquisador está interessado em aferir o efeito de uma causa X sobre um resultado Y em uma população de interesse. O pesquisador postula o seguinte modelo para a relação que almeja estudar:

$$Y = \beta_0 + \beta_1 X + U,$$

onde β_1 é definido como o efeito causal de interesse, e U são outras causas não observáveis.

Preocupado com um possível problema de determinação simultânea entre Y e X , o pesquisador considera uma terceira variável, Z , como potencial instrumento. Entretanto, também preocupado com potenciais violações da hipótese de exclusão, o pesquisador pretende analisar, antes da implementação empírica, as propriedades de consistência dos estimadores de mínimos quadrados ordinários e de variável instrumental de β_1 baseando-se num experimento hipotético em que o tamanho amostral diverge para infinito.

No que segue, suponha que Y , X e Z têm variância estritamente positiva e finita.

Com base no acima exposto, identifique como verdadeiras ou falsas as afirmações abaixo.

- Ⓒ A diferença entre o limite em probabilidade do estimador de mínimos quadrados ordinários de β_1 e o parâmetro β_1 é dada por $\frac{\text{cov}(X,U)}{V[X]}$.
- Ⓐ A diferença entre o limite em probabilidade do estimador de variáveis instrumentais de β_1 e o parâmetro β_1 é dada por $\frac{\text{cov}(Z,U)}{V[Z]}$.
- Ⓑ Em termos de “viés assintótico”, definido como a diferença entre o limite em probabilidade do estimador e o parâmetro de interesse, o estimador de variável instrumental é sempre preferível ao estimador de MQO.
- Ⓓ Se tanto $\text{cov}(X,U) = 0$, como $\text{cov}(Z,U) = 0$, então, tanto o estimador de MQO como o estimador de variável instrumental são consistentes para β_1 .
- Ⓔ Se $\text{cov}(X,U) = 0$ e $\text{cov}(Z,U) = 0$ são satisfeitas e valem as hipóteses de homocedasticidade $E[U^2|X] = E[U^2|Z] = \sigma^2$, o estimador de mínimos quadrados ordinários é preferível ao estimador de variável instrumental sob o ponto de vista de eficiência em amostras grandes.

QUESTÃO 10

Considere o seguinte processo $\{X_t\}_{t \in N \cup \{0\}}$,

$$\begin{aligned} X_0 &= 0 \\ X_t &= X_{t-1} + \epsilon_t, \quad t > 0, \end{aligned}$$

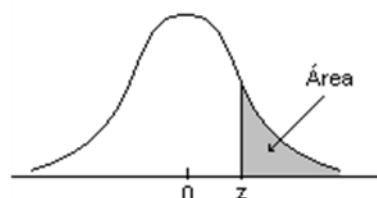
onde $\{\epsilon_t\}_{t \in N}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas, com a seguinte distribuição:

$$\begin{aligned} P[\epsilon_t = 1] &= \theta, \\ P[\epsilon_t = -1] &= 1 - \theta, \end{aligned}$$

para algum $\theta \in (0,1)$. Com base no acima exposto, julgue como verdadeiras ou falsas as seguintes afirmações:

- Ⓒ O valor esperado de X_t é dado por $E[X_t] = \theta t$.
- ① A variância de X_t é dada por $V[X_t] = 2\theta(1 - \theta)t$.
- ② A covariância entre X_t e X_s é dada por $4\theta(1 - \theta)\min\{t, s\}$.
- ③ O processo $\{X_t\}$ é estacionário para $\theta = 1/2$ e não estacionário para $\theta \neq 1/2$.
- ④ A probabilidade de que o processo eventualmente alcance $X_t = 1$ para algum $t > 0$ é igual a 1 se $\theta > 1/2$.

Distribuição Normal Padrão



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135									
3.5	.000233									
4.0	.0000317									
4.5	.00000340									
5.0	.000000287									

Valores críticos da distribuição t de Student

Graus de liberdade	Nível de significância (bicaudal)			
	0,20	0,10	0,05	0,02
1	3,078	6,314	12,706	31,821
2	1,886	2,920	4,303	6,965
3	1,638	2,353	3,182	4,541
4	1,533	2,132	2,776	3,747
5	1,476	2,015	2,571	3,365
6	1,440	1,943	2,447	3,143
7	1,415	1,895	2,365	2,998
8	1,397	1,860	2,306	2,896
9	1,383	1,833	2,262	2,821
10	1,372	1,812	2,228	2,764
20	1,325	1,725	2,086	2,528
30	1,310	1,697	2,042	2,457
40	1,303	1,684	2,021	2,423
60	1,296	1,671	2,000	2,390
120	1,289	1,650	1,980	2,358
∞	1,282	1,645	1,960	2,326

