

Prova Escrita Objetiva – Nível Superior

# ANALISTA CENSITÁRIO - MÉTODOS QUANTITATIVOS

Tipo 1 – BRANCA



## SUA PROVA

Além deste caderno de prova, contendo cinquenta questões objetivas, você receberá do fiscal de sala:

- uma folha destinada às respostas das questões objetivas



## TEMPO

- **4 horas** é o período disponível para a realização da prova, já incluído o tempo para a marcação da folha de respostas da prova objetiva
- **2 horas e 30 minutos** após o início da prova é possível retirar-se da sala, sem levar o caderno de prova
- **30 minutos** antes do término do período de prova é possível retirar-se da sala levando o caderno de prova



## NÃO SERÁ PERMITIDO

- Qualquer tipo de comunicação entre os candidatos durante a aplicação da prova
- Levantar da cadeira sem autorização do fiscal de sala
- Usar o sanitário ao término da prova, após deixar a sala



## INFORMAÇÕES GERAIS

- As questões objetivas têm cinco alternativas de resposta (A, B, C, D, E) e somente uma delas está correta
- Verifique se seu caderno está completo, sem repetição de questões ou falhas. Caso contrário, notifique imediatamente o fiscal da sala, para que sejam tomadas as devidas providências
- Confira seus dados pessoais, especialmente nome, número de inscrição e documento de identidade e leia atentamente as instruções para preencher a folha de respostas
- Use somente caneta esferográfica, fabricada em material transparente, com tinta preta ou azul
- Assine seu nome apenas nos espaços reservados
- Marque na folha de respostas o campo relativo à confirmação do tipo/cor de prova, conforme o caderno recebido
- O preenchimento das respostas da prova objetiva é de sua responsabilidade e não será permitida a troca da folha de respostas em caso de erro
- Reserve tempo suficiente para o preenchimento de suas respostas. Para fins de avaliação, serão levadas em consideração apenas as marcações realizadas na folha de respostas da prova objetiva, não sendo permitido anotar informações relativas às respostas em qualquer outro meio que não seja o caderno de prova
- A FGV coletará as impressões digitais dos candidatos
- Os candidatos serão submetidos ao sistema de detecção de metais quando do ingresso e da saída de sanitários durante a realização das provas
- Boa Sorte!



## Língua Portuguesa

Texto 1 – ENTREVISTA COM O FÍSICO HOWARD GELLER

**O Brasil passou por um período de racionamento de energia em 2001. Isso pode se repetir? O que pode ser feito para evitar um novo racionamento?**

O racionamento foi resultado da política de privatização e desregulamentação que não incentivou suficientemente a construção de novas usinas. O governo também não permitiu que o setor público investisse nessa área. Não planejou nem implementou uma política para o setor. O problema principal foi esse e não tinha uma carência de energia ou da capacidade de fornecê-la, embora o volume de chuvas tenha sido pequeno nos anos anteriores.

No futuro, o desafio será adotar uma política energética que estimule o fornecimento de energia, através de eletricidade ou de combustíveis, a um custo acessível para os consumidores e as empresas, protegendo inclusive o meio ambiente. É preciso levar em conta questões econômicas e sociais. No Brasil, há pelo menos 20 milhões de pessoas que vivem em áreas rurais das regiões Norte e Nordeste, sem acesso à eletricidade. Uma boa política expandiria o fornecimento para essa população. (*Ciência Hoje*, maio de 2004 - adaptado)

**1**

Considerando as duas perguntas formuladas pelo entrevistador, pode-se ver que o entrevistado:

- (A) não respondeu a nenhuma delas;
- (B) respondeu claramente às duas questões;
- (C) respondeu somente à segunda das perguntas;
- (D) respondeu somente à primeira das perguntas;
- (E) respondeu integralmente à primeira das perguntas e parcialmente à segunda.

**2**

“O racionamento foi resultado da política de privatização e desregulamentação que não incentivou suficientemente a construção de novas usinas”.

O início da resposta do entrevistado corresponde a uma pergunta que não foi formulada diretamente; essa pergunta, se formulada, seria:

- (A) O que é racionamento de energia?
- (B) A política interferiu com o racionamento de energia ocorrido?
- (C) A privatização é uma boa política do Estado?
- (D) A solução para a crise é a construção de novas usinas?
- (E) Como se pode chegar à regulamentação do setor energético?

**3**

**“O que pode ser feito para evitar um novo racionamento?”**

A oração “para evitar um novo racionamento” pode ser desenvolvida em forma de uma nova oração do seguinte modo:

- (A) Para evitar-se um novo racionamento?
- (B) Para que se evitasse um novo racionamento?
- (C) Para que um novo racionamento fosse evitado?
- (D) Para que se evite um novo racionamento?
- (E) Para ser evitado um novo racionamento?

**4**

No texto 1 há um conjunto de termos precedidos da preposição DE; o termo abaixo em que essa preposição tem emprego não exigido por um termo anterior é:

- (A) “racionamento de energia”;
- (B) “construção de novas usinas”;
- (C) “capacidade de fornecê-la”;
- (D) “volume de chuvas”;
- (E) “fornecimento de energia”.

**5**

“O governo também não permitiu que o setor público investisse nessa área”.

O segundo período da resposta do entrevistado desempenha o seguinte papel textual:

- (A) contraria o que é anteriormente dito;
- (B) retifica um erro cometido pelo entrevistado;
- (C) indica uma outra possibilidade de se verem os fatos;
- (D) confirma o que é informado antes;
- (E) acrescenta mais um argumento à opinião expressa antes.

**6**

“Não planejou nem implementou uma política para o setor”.

Sobre as duas ações citadas nesse segmento do texto 1, pode-se corretamente afirmar que:

- (A) a primeira ação depende da segunda;
- (B) a segunda ação depende da primeira;
- (C) as duas ações são independentes;
- (D) as ações são interdependentes;
- (E) a primeira ação implica obrigatoriamente a segunda.

**7**

No primeiro parágrafo do texto 1 há um conjunto de termos que recuperam elementos anteriores, o que dá coesão ao texto.

O termo cujo antecedente é uma oração é:

- (A) “...que não incentivou...”;
- (B) “...o setor público investisse nessa área”;
- (C) “...nem implementou uma política para o setor”;
- (D) “O problema principal foi esse...”;
- (E) “...ou da capacidade de fornecê-la”.

**8**

“...embora o volume de chuvas tenha sido pequeno nos anos anteriores”.

Esse segmento do texto 1 mostra o seguinte valor:

- (A) reduz a intensidade da crítica feita ao governo da época;
- (B) compara a situação citada com a de anos anteriores;
- (C) elogia a atuação estatal no setor energético;
- (D) atribui o racionamento a problemas naturais;
- (E) aumenta as críticas ao desempenho governamental.

**9**

“No futuro, o desafio será adotar uma política energética que estimule o fornecimento de energia, através de eletricidade ou de combustíveis, a um custo acessível para os consumidores e as empresas, protegendo inclusive o meio ambiente”.

Deduz-se desse segmento opinativo do entrevistado que:

- (A) a energia fornecida não deve obrigatoriamente gerar lucro para as empresas;
- (B) a sugestão do entrevistado é de fácil execução e só depende de vontade política;
- (C) a energia pode ser produzida, de forma acessível, por várias fontes;
- (D) a energia elétrica é de menor custo de produção que a de combustíveis;
- (E) a energia elétrica é a única que protege o meio ambiente.

**10**

“É preciso levar em conta questões econômicas e sociais”; se juntássemos os adjetivos sublinhados em forma de adjetivo composto, a forma correta, no contexto, seria:

- (A) econômicas-sociais;
- (B) econômico-social;
- (C) econômica-social;
- (D) econômico-sociais;
- (E) econômicas-social.

Texto 2 – ANTES DO PRIMEIRO CRIME

Um fenômeno bastante conhecido nos estudos criminológicos é o fato de que muitas coisas envolvem poucas pessoas e lugares. Estudos usando técnicas de análise para a detecção de “áreas quentes” de criminalidade mostram que crimes são fenômenos bastante concentrados no tempo e no espaço. Alguns exemplos: 1) as regiões metropolitanas de São Paulo e do Rio de Janeiro concentram 40% dos homicídios no Brasil, embora tenham 18% da população; 2) cerca de 20% desse tipo de crime acontecem em menos de 2% da área geográfica de um centro urbano; 3) a maioria dos assaltos ocorre em poucos locais. (*Ciência Hoje*, maio de 2004)

**11**

“Um fenômeno bastante conhecido nos estudos criminológicos é o fato de que muitas coisas envolvem poucas pessoas e lugares”.

O fato citado nesse segmento inicial do texto 2 se apoia:

- (A) na opinião do autor;
- (B) na credibilidade da revista que publica o texto;
- (C) no senso comum;
- (D) na autoridade de conhecimentos científicos;
- (E) no valor de estudos históricos.

**12**

A característica que mostra mais claramente ser o texto 2 de cunho jornalístico tradicional é:

- (A) uma seleção de informações de destaque social;
- (B) uma utilização de linguagem popular, para fácil acesso;
- (C) a indicação de pessoas de valor reconhecido;
- (D) a não participação direta do autor do texto;
- (E) a atualidade flagrante do tema selecionado.

**13**

Os exemplos citados no texto 2 têm a função de:

- (A) mostrar que as afirmações anteriormente feitas só se aplicam a Rio e São Paulo;
- (B) indicar que os estudos citados tiveram por base as regiões metropolitanas carioca e paulista;
- (C) comprovar a tese indicada anteriormente;
- (D) demonstrar o relativismo das informações jornalísticas;
- (E) contrariar as afirmações anteriores do texto.

**14**

Indique a frase abaixo em que uma outra concordância verbal é possível:

- (A) “...muitas coisas envolvem poucas pessoas”;
- (B) “...cerca de 20% desses tipos de crime acontecem...”;
- (C) “...a maioria dos assaltos ocorre em poucos locais”;
- (D) “...crimes são fenômenos bastante concentrados”;
- (E) “...as regiões metropolitanas de São Paulo e do Rio de Janeiro concentram 40% dos homicídios”.

**15**

“as regiões metropolitanas de São Paulo e do Rio de Janeiro concentram 40% dos homicídios no Brasil, embora tenham 18% da população”.

Para que a última oração desse segmento do texto 2 tivesse mais clareza, a forma adequada seria:

- (A) embora tenham apenas 18% da população;
- (B) embora tenham mais de 18% da população;
- (C) embora tenham talvez 18% da população;
- (D) embora tenham menos de 18% da população;
- (E) embora tenham cerca de 18% da população.

## Conhecimentos Específicos

16

A população de um estudo é dividida em quatro estratos, sendo o menor com 10% dos indivíduos e os demais com tamanhos acrescidos de dez pontos percentuais, progressivamente. Os estratos se distinguem por classes de renda com amplitude constante, sendo maiores quanto menor a renda.

Sobre os estratos sabe-se que:

$$Rd_{Estrato\ 1} = 65 \quad Rd_{Estrato\ 2} = 45 \quad e \quad Rd_{Estrato\ 4} = 5$$

onde os valores acima representam os limites inferiores da renda dos extratos, inclusive.

Portanto, é correto afirmar que:

- (A) tomando os pontos médios das classes como representativos, a renda média é igual a  $Md(Rd) = 38$ ;
- (B) a mediana da distribuição de renda,  $Me(Rd)$ , é menor que 45 e maior do que ou igual a 25;
- (C) tomando os pontos médios das classes como representativos, a moda da renda é igual a  $Mo(Rd) = 35$ ;
- (D) o valor máximo atingido pela renda nessa distribuição é igual a  $Mx(Rd) = 85$ ;
- (E) o valor do desvio-interquartilício da distribuição de renda deverá ser superior a 50.

17

Para a realização de uma pesquisa de campo, uma amostra por conglomerados deverá ser extraída. São considerados como unidades da população os 27 estados brasileiros que, para essa finalidade estão classificados como mais (11 UFs) ou menos (16 UFs) desenvolvidos.

Supondo que a amostra será de  $n = 5$ , a probabilidade de que sejam sorteados dois mais desenvolvidos e três menos desenvolvidos é de:

- (A)  $8 \cdot \left(\frac{11}{27}\right) \cdot \left(\frac{10}{26}\right) \cdot \left(\frac{11}{25}\right) \cdot \left(\frac{16}{24}\right) \cdot \left(\frac{15}{23}\right)$
- (B)  $\left(\frac{2}{11}\right)^2 \cdot \left(\frac{3}{16}\right)^3$
- (C)  $\left(\frac{11}{27}\right) \cdot \left(\frac{10}{26}\right) \cdot \left(\frac{16}{25}\right) \cdot \left(\frac{15}{24}\right) \cdot \left(\frac{14}{23}\right)$
- (D)  $10 \cdot \left(\frac{11}{27}\right) \cdot \left(\frac{10}{26}\right) \cdot \left(\frac{16}{25}\right) \cdot \left(\frac{15}{24}\right) \cdot \left(\frac{14}{23}\right)$
- (E)  $\frac{(11!) \cdot (16!)}{27!}$

18

Uma vez concluída a etapa de críticas de dados, relativa a um conjunto de registros obtido através de uma pesquisa de campo, inicia-se o trabalho de tabulação e elaboração de gráficos. Durante as análises, algumas variáveis surgem com destaque.

(i) número de indivíduos por faixa etária; (ii) percentuais do nível de escolaridade; e (iii) pares de valores de consumo e renda.

Portanto, os tipos de gráficos considerados adequados a serem empregados em cada caso são, respectivamente:

- (A) colunas, linha e setores;
- (B) setores, cartograma e barras;
- (C) barras, setores e linha;
- (D) linha, colunas e cartograma;
- (E) colunas, setores e linha.

19

A ideia de agrupar as observações de uma população ou amostra constitui uma técnica bem antiga de condensar as informações e assim facilitar o seu tratamento. No passado essa técnica era empregada com sucesso, mas com a ressalva de que os resultados não eram tão precisos quanto aqueles obtidos com dados não agrupados.

Considere a distribuição expressa em classes de frequências:

Classes	Frequências
10  -- 20	50
20  -- 30	28
30  -- 40	24
<b>Total</b>	<b>102</b>

Mesmo sem dispor dos dados de forma desagregada, sobre as estatísticas exatas, é correto afirmar que:

- (A) a moda não pertence à última classe;
- (B) a média é superior a 28;
- (C) a mediana é menor do que 23;
- (D) a média é superior a 16;
- (E) a moda é inferior a 20.

20

A figura abaixo apresenta um diagrama do tipo Box-Plot de uma amostra que, além de estatísticas de referência próprias da sua construção, assinala quatro elementos  $X_1, X_2, X_3$  e  $X_4$ ,



onde os  $Q_{i,s}$  são os quartis e LI e LS são os limites inferior e superior do Box-Plot.

Sobre esses valores e/ou a sua distribuição, é correto afirmar que:

- (A) os valores  $X_1$  e  $X_4$  podem ser considerados outliers;
- (B) o valor  $X_2$  é superior à média das observações;
- (C) a distribuição é assimétrica e o valor  $X_4$  é um outlier;
- (D) menos de 30% dos valores amostrais estão acima de  $X_3$ ;
- (E) o valor mínimo da amostra está situado abaixo de  $X_1$ .

21

Alguns economistas estão discutindo sobre a volatilidade dos preços em duas economias, relativamente parecidas, tendo como moedas peras (A) e maçãs (B). Sabe-se que as médias dos preços são 100 peras e 120 maçãs, respectivamente. É fornecido, ainda, o desvio-padrão dos preços em A, igual a 25 peras, e a variância em B, igual a 400 maçãs ao quadrado.

Considerando as principais medidas estatísticas de dispersão como medidas de volatilidade, é correto afirmar que:

- (A) o desvio padrão dos preços em A é inferior ao de B;
- (B) a taxa de conversão da moeda A para B é de 1,2;
- (C) a taxa de inflação em A deve ser menor do que em B;
- (D) os preços em B são, em média, mais caros do que em A;
- (E) a medida adimensional de dispersão de A é superior à de B.

**22**

Para uma distribuição de frequências apenas parcialmente conhecida são fornecidas as seguintes estatísticas,

$$Mo(X) = 19, E(X^2) = 625 \text{ e } Me(X) = 22$$

sendo  $Mo$ , a moda e  $Me$ , a mediana dos dados. Sabe-se ainda que a distribuição é unimodal.

Esse conjunto bem restrito de informações seria compatível apenas com:

- (A) a distribuição é assimétrica à esquerda;
- (B) a média da distribuição,  $Md(X)$ , é igual a 27;
- (C) a variância da distribuição é igual a 500;
- (D) o coeficiente de variação,  $CV(X)$ , é igual a 7,5;
- (E) a média da distribuição,  $Md(X)$ , é igual a 23.

**23**

Sejam  $X$  e  $Y$  duas variáveis aleatórias com variâncias iguais a 21 e 17, respectivamente. Além disso, sabe-se que a variável  $Z$  representada pela diferença entre as duas tem variância igual a 44.

Com base em tais informações, é correto deduzir que:

- (A) as variáveis  $Z$  e  $X$  são positivamente correlacionadas;
- (B) o momento de segunda ordem de  $Y$  é maior do que o de  $Z$ ;
- (C) a média de  $Z$  é menor do que ambas as médias, de  $X$  e de  $Y$ ;
- (D) a covariância entre  $X$  e  $Y$  é positiva;
- (E) as variáveis  $X$  e  $Y$  são negativamente correlacionadas.

**24**

Sejam  $X$  e  $Y$  variáveis aleatórias do tipo Bernoulli, assumindo valores  $x_1, x_2, y_1$  e  $y_2$ , respectivamente. Também é sabido que  $P(X = x_1 / Y = y_2) = 0,60$  e  $P(Y = y_1) = 0,75$ .

Então:

- (A)  $P(X = x_2 / Y = y_2) = 0,35$ ;
- (B)  $P(X = x_2; Y = y_2) = 0,10$ ;
- (C)  $P(X = x_1; Y = y_1) = 0,25$ ;
- (D)  $P(Y = y_2 / X = x_1) = 0,40$ ;
- (E)  $P(X = x_2) = 0,17$ .

**25**

Suponha que o número de pessoas aguardando em uma fila segue, por unidade de tempo, uma distribuição de Poisson, com parâmetro que depende do atendente. O funcionário de 2ª, 4ª e 6ª produz  $\lambda = 20$ , enquanto o de 3ª e 5ª  $\lambda = 15$ .

Assim, sobre a variável "número de pessoas esperando em um dia aleatório", é correto afirmar que:

- (A) é um processo de Poisson com parâmetro  $\lambda = 12$ ;
- (B) tem média igual a 17,5;
- (C) tem variância igual a 35;
- (D) tem esperança matemática igual a 18;
- (E) tem coeficiente de variação igual a 2.

**26**

Sabe-se que o tempo de aplicação de um questionário em uma pesquisa de campo é uma variável com distribuição uniforme entre 8 e 20 minutos. Um entrevistador pretende aplicar três questionários.

Logo, é correto afirmar que:

- (A) a probabilidade de que todas as entrevistas durem mais do que 15 minutos é de 27/64;
- (B) a probabilidade de que duas das entrevistas durem mais do que a média é igual a 5/8;
- (C) o desvio padrão do tempo de duração de cada entrevista é igual a 2 minutos;
- (D) a probabilidade de que apenas uma das entrevistas leve menos da metade do tempo máximo é igual a 25/72;
- (E) a probabilidade do tempo total de entrevista exceder 40 minutos é igual a 0,5.

**27**

Suponha que certa característica de uma dada população tem suas medidas distribuídas normalmente com média 40 e variância igual a 25. Um indivíduo deverá ser extraído ao acaso e sua característica observada. Considere também as seguintes informações:

$\phi(1,9) = 0,971$ ,  $\phi(1,6) = 0,945$ ,  $\phi(1,25) = 0,895$  e  $\phi(2,1) = 0,982$  onde  $\phi(\cdot)$  = função distribuição acumulada da normal-padrão.

A probabilidade de o valor sorteado diferir por mais de vinte por cento da média verdadeira é:

- (A) 3,6%;
- (B) 5,5%;
- (C) 10,5%;
- (D) 11%;
- (E) 21%.

**28**

Suponha que o tempo de vida útil da lâmpada de um Scanner seja distribuído exponencialmente com parâmetro  $\beta = 600$  horas.

Se  $T$  representa a durabilidade da lâmpada, é correto afirmar que:

- (A)  $P(T > 600) = 0,50$ ;
- (B)  $P(200 < T < 600) = 0,25$ ;
- (C)  $P(T > 1500) = 1 - e^{-2}$ ;
- (D)  $P(T > 1200 | T > 300) = P(Y > 900)$ ;
- (E)  $P(T < 450) = 1 - e^{-2/5}$ .

**29**

Suponha que de uma população muito numerosa de pessoas, formada por brancos, negros e mestiços, nas proporções 3:2:1 será extraída uma amostra de tamanho  $n = 12$ . O que se deseja é selecionar uma amostra que reflita perfeitamente as proporções de cores da população.

Então a probabilidade de que a amostra tenha a característica desejada é igual a:

- (A)  $30.240. \left(\frac{3}{6}\right)^6 \left(\frac{2}{6}\right)^4 \left(\frac{1}{6}\right)^2$ ;
- (B)  $13.860. \left(\frac{3}{6}\right)^6 \left(\frac{2}{6}\right)^4 \left(\frac{1}{6}\right)^2$ ;
- (C)  $\left(\frac{3}{6}\right)^6 \left(\frac{2}{6}\right)^4 \left(\frac{1}{6}\right)^2$ ;
- (D)  $\left(\frac{3}{12}\right)^6 \left(\frac{2}{12}\right)^4 \left(\frac{1}{12}\right)^2$ ;
- (E)  $60. \left(\frac{3}{6}\right)^6 \left(\frac{2}{6}\right)^4 \left(\frac{1}{6}\right)^2$ .

**30**

Seja  $(X, Y)$  uma variável aleatória bidimensional que em dada amostra assumiu o seguinte conjunto de valores:

$(1,16), (5,8)$  e  $(9, 3)$

PS: Use, nos cálculos,  $\sqrt{43} \cong 6,5$ .

Logo, a estimativa para o coeficiente de correlação de Pearson para o par  $(X, Y)$  obtido pelo método dos momentos será aproximadamente:

- (A) -1,00;
- (B) -0,75;
- (C) +0,45;
- (D) -0,50;
- (E) +0,80.

**31**

Considere a distribuição conjunta  $(X,Y)$  abaixo especificada.

		X		
		1	3	4
Y	-1	0.10	0.10	0.25
	0	0.15	0.05	0.05
	1	0.05	0.15	0.10

Com base nessa função de probabilidade temos que:

- (A)  $E(X | Y = 1) = 3$ ;
- (B)  $E(X | Y = 0) = 2$ ;
- (C)  $E(Y | X = 3) = 0,25$ ;
- (D)  $E(X | Y = -1) = 3,5$ ;
- (E)  $E(Y | X = 1) = -0,5$ .

**32**

O responsável pelo planejamento de uma pesquisa acredita que, *a priori*, a probabilidade de que um indivíduo tenha uma determinada opinião, positiva, é de 80%. Para avaliar melhor essa crença, o responsável realiza um experimento no qual a opinião é positiva em 40% dos casos, quando o responsável julga *a priori* que não será assim; sendo positiva em 70% dos casos, quando ele prevê uma opinião positiva. No experimento, a opinião se mostrou positiva (ExpPos).

Portanto, a distribuição *a posteriori*, ou seja, após a realização do experimento, para a crença do responsável depois do experimento é:

- (A)  $P(\text{Positiva} | \text{ExpPos}) = 0,56$  e  $P(\text{NÃO Positiva} | \text{ExpPos}) = 0,44$ ;
- (B)  $P(\text{Positiva} | \text{ExpPos}) = 0,68$  e  $P(\text{NÃO Positiva} | \text{ExpPos}) = 0,32$ ;
- (C)  $P(\text{Positiva} | \text{ExpPos}) = 0,875$  e  $P(\text{NÃO Positiva} | \text{ExpPos}) = 0,125$ ;
- (D)  $P(\text{Positiva} | \text{ExpPos}) = 0,68$  e  $P(\text{NÃO Positiva} | \text{ExpPos}) = 0,32$ ;
- (E)  $P(\text{Positiva} | \text{ExpPos}) = 2/3$  e  $P(\text{NÃO Positiva} | \text{ExpPos}) = 1/3$ .

**33**

Sejam  $X_1, X_2, X_3, \dots, X_{64}$  variáveis aleatórias discretas, com distribuição Binomial, todas com  $p = 0,25$  e  $n = 12$ . Também são conhecidos valores da função distribuição acumulada da normal-padrão, mais especificamente:

$\Phi(2) = 0,977, \Phi(1,5) = 0,933, \Phi(1,25) = 0,894$

No caso da extração de uma amostra ( $n = 64$ ), a probabilidade (desprezando o ajuste de continuidade) de que a soma dos valores seja superior a 207 é igual a:

- (A) 0,023;
- (B) 0,046;
- (C) 0,067;
- (D) 0,106;
- (E) 0,134.

**34**

Com o objetivo de verificar qual seria a forma funcional mais adequada a um modelo é feita uma transformação Box-Cox, estimando-se repetidas vezes o seguinte modelo:

$$Y^* = \alpha + \beta \cdot X^* + \varepsilon$$

onde  $Y^* = \frac{Y^\lambda - 1}{\lambda}$  e  $X^* = \frac{X^\delta - 1}{\delta}$ , sendo  $\lambda$  e  $\delta$  os parâmetros que mudam a cada nova rodada de estimações. As distribuições de  $\lambda$  e  $\delta$  foram identificadas para os testes de hipóteses:

$$H_0: \lambda = 0 \text{ vs } H_a: \lambda = 1 \text{ e } H_0: \delta = 1 \text{ vs } H_a: \delta = 0$$

Em ambos os testes  $H_0$  foi rejeitada.

Então a forma funcional mais adequada ao modelo inicial é:

- (A)  $Y = \alpha^* + \beta \cdot X + \varepsilon$ ;
- (B)  $\ln Y = \alpha^* + \beta \cdot X + \varepsilon$ ;
- (C)  $Y = \alpha^* + \beta \cdot \ln X + \varepsilon$ ;
- (D)  $\ln Y = \alpha^* + \beta \cdot \ln X + \varepsilon$ ;
- (E)  $Y = \alpha^* + \beta \cdot \left(\frac{1}{X}\right) + \varepsilon$ .

**35**

Para estimar a média de certa população  $\mu$ , desconhecida, partindo apenas de duas observações amostrais, cogita-se o emprego de um dos seguintes estimadores, onde  $X_1$  e  $X_2$  representam os indivíduos da amostra *ex ante*.

$$\hat{\theta} = \frac{2}{3}X_1 + \frac{1}{3}X_2 \text{ e } \tilde{\theta} = \frac{2}{7}X_1 + \frac{4}{7}X_2$$

Sobre os estimadores, é correto afirmar que:

- (A) o estimador  $\tilde{\theta}$  é mais eficiente do que  $\hat{\theta}$ ;
- (B) o estimador  $\tilde{\theta}$  subestima, em média, o valor verdadeiro de  $\mu$ ;
- (C) a tendenciosidade de  $\hat{\theta}$ ,  $T(\hat{\theta})$ , é igual a  $\mu/4$ ;
- (D) o Erro Quadrático Médio de  $\tilde{\theta}$ ,  $EQM(\tilde{\theta})$ ,  $\frac{20}{49} \cdot \sigma^2$ ;
- (E) a tendenciosidade de  $\tilde{\theta}$ ,  $T(\tilde{\theta})$ , é igual a  $\frac{-\mu}{7}$ .

**36**

Dois estimadores,  $\hat{\theta}_1$  e  $\hat{\theta}_2$ , para um parâmetro populacional  $\theta$ , têm seus Erros Quadráticos Médios (EQM) dados por:

$$EQM(\hat{\theta}_1) = \frac{3\sigma^2}{n} + \left(\frac{\theta-1}{n}\right)^2 \text{ e } EQM(\hat{\theta}_2) = \frac{\sqrt{n}\sigma^2}{5} + \left(1 + \frac{\theta}{n}\right)^2$$

Com base apenas nas expressões, onde as primeiras parcelas são as variâncias, é correto concluir que:

- (A)  $\hat{\theta}_1$  é não tendencioso;
- (B)  $\hat{\theta}_1$  é um estimador consistente;
- (C)  $\hat{\theta}_2$  é assintoticamente não tendencioso;
- (D)  $\hat{\theta}_2$  não é um estimador consistente;
- (E) ambos são assintoticamente eficientes.

**37**

Seja X uma variável aleatória com função de probabilidade dada por  $P(X = x) = p \cdot (1 - p)^{x-1}$  para  $x = 1, 2, 3, \dots$ , onde p é um parâmetro desconhecido. Dispondo de uma amostra de tamanho n,  $x_1, x_2, x_3, \dots, x_n$ , o estimador de Máxima Verossimilhança de p é:

- (A)  $\hat{p} = \sum x_i$ ;
- (B)  $\hat{p} = \frac{1}{\sum x_i}$ ;
- (C)  $\hat{p} = \frac{\sum x_i}{n}$ ;
- (D)  $\hat{p} = \frac{n}{\sum x_i}$ ;
- (E)  $\hat{p} = \frac{n}{\sqrt{\sum x_i}}$ .

**38**

Uma amostra de cinco indivíduos é extraída aleatoriamente de uma dada população, obtendo-se os seguintes valores:

$$X_1 = 3, \quad X_2 = 5, \quad X_3 = 4, \quad X_4 = 7 \text{ e } X_5 = 11$$

Então a variância amostral e a estimativa não tendenciosa da variância populacional seriam iguais a, respectivamente:

- (A) 8 e 10;
- (B) 9 e 8;
- (C) 8 e 12;
- (D) 8 e 6,67;
- (E) 6,67 e 10.

**39**

Para estimar por intervalo da proporção de indivíduos que, em certa população, são portadores de diabetes, é extraída uma amostra aleatória simples (AAS) com tamanho  $n = 2500$ . Do total, 375 indivíduos foram classificados como portadores da doença. Adicionalmente,  $\phi(\cdot)$ , a distribuição acumulada da normal-padrão assume os valores:

$$\phi(1,96) = 0,975, \phi(1,64) = 0,95, \phi(1,28) = 0,90$$

Fazendo uso do limite superior da variância de proporções e com nível de significância de 10%, o intervalo de confiança procurado é:

- (A) (0,1336, 0,1664);
- (B) (0,1372, 0,1628);
- (C) (0,1244, 0,1756);
- (D) (0,1168, 0,1832);
- (E) (0,1304, 0,1696).

**40**

O faturamento médio das empresas de determinado setor é desconhecido para os empresários de fora do mercado. Um deles, interessado em investir, já sabe que só vale a pena entrar no negócio caso o faturamento médio seja maior do que 80 unidades monetárias. Para avaliar esse mercado, um teste de hipóteses é realizado. Uma AAS (Amostra Aleatória Simples) de tamanho  $n = 100$  é extraída, obtendo-se  $\bar{X} = c$ . Sabe-se que o desvio padrão verdadeiro do faturamento é igual a 30 e a função distribuição acumulada de normal,  $\phi(\cdot)$ , toma valores  $\phi(1,96) = 0,975$ ,  $\phi(1,64) = 0,95$ ,  $\phi(1,28) = 0,90$ .

Sendo  $\alpha$  o nível de significância, a decisão do teste deve ser:

- (A) rejeitar a hipótese nula se  $\alpha = 10\%$  e  $c = 83$ ;
- (B) não rejeitar a hipótese nula se  $\alpha = 5\%$  e  $c = 85$ ;
- (C) rejeitar a hipótese nula se  $\alpha = 2,5\%$  e  $c = 86$ ;
- (D) rejeitar a hipótese nula se  $\alpha = 5\%$  e  $c = 84$ ;
- (E) aceitar a hipótese alternativa se  $\alpha = 2,5\%$  e  $c = 87$ .

**41**

Um teste de hipóteses será realizado para verificar se uma moeda é, de fato, honesta. Suspeita-se que, ao invés de um equilíbrio,  $P(\text{Cara}) = P(\text{Coroa}) = 0,5$ , há uma tendência para que as chances sejam de 3:2 favorável a Cara. Assim sendo, as hipóteses formuladas são:

Ho: Moeda equilibrada (1:1)

Ha: Moeda desequilibrada (3:2)

A decisão deverá seguir um critério bem simples. A tal moeda será lançada quatro vezes, rejeitando-se a hipótese nula caso aconteçam mais do que três Caras.

Com tal critério, é correto afirmar que:

- (A)  $P(\text{Erro Tipo I}) = 1/8$  e  $P(\text{Erro Tipo II}) = 1 - (0,125)$ ;
- (B)  $P(\text{Erro Tipo I}) = 0,25$  e  $P(\text{Erro Tipo II}) = 1 - (0,25)$ ;
- (C)  $P(\text{Erro Tipo I}) = 3/16$  e  $P(\text{Erro Tipo II}) = 1 - (0,4)^4$ ;
- (D)  $P(\text{Erro Tipo I}) = 1/16$  e  $P(\text{Erro Tipo II}) = 1 - (0,6)^4$ ;
- (E)  $P(\text{Erro Tipo I}) = 3/8$  e  $P(\text{Erro Tipo II}) = 1 - (0,75)^4$ .



**42**

Para o caso de variáveis aleatórias quaisquer, existem diversas propriedades que se aplicam diretamente à esperança matemática e ao momento central de segunda ordem.

Dentre essas propriedades está:

- (A)  $Var(X) > E(X^2)$ ;
- (B)  $Var(X \pm Y) = Var(X) \pm Var(Y)$ ;
- (C)  $E(X) \pm E(Y) = E(X) \pm E(Y)$ ;
- (D)  $Var(aX) = aVar(X)$ , sendo  $a$  uma constante positiva;
- (E)  $E(aX) = E(X)$ , sendo  $a$  uma constante qualquer.

**43**

Um modelo de regressão linear múltiplo será adotado para realizar imputações, em razão de algumas lacunas de dados em um levantamento de campo. As variáveis consideradas são o consumo de feijão (anual), o número de adultos equivalente e a renda (mensal do domicílio). O modelo formulado foi:

$$(Feijão) = \alpha + \beta \cdot (Renda_i) + \delta \cdot (Adulto_i) + \varepsilon_i$$

Para  $i = 1, 2, 3, \dots, 28$  (domicílios)

Os resultados da estimação estão na tabela abaixo:

Parâmetros	Estimativa	Erro-Padrão	t-Student	p-valor
$\alpha$ (constante)	4,25	1,85	2,30	0,0303
$\beta$ (renda)	0,047	0,027	1,74	0,0940
$\delta$ (adulto)	5,17	2,41	2,15	0,0418

A partir dos números acima, é correto concluir que:

- (A) se a renda de um domicílio for acrescida de 100 unidades monetárias, o consumo de feijão deverá sofrer um acréscimo médio de 4,7 Kg do produto;
- (B) ao nível de significância de 10% a renda é não significativa para explicar o consumo de feijão;
- (C) para cada 1% de variação no número de adultos equivalentes no domicílio, o consumo de feijão será acrescido de 5,17%;
- (D) se uma família com quatro adultos equivalentes ganha 1.000 unidades monetárias, o consumo médio imputado seria de 71,83 quilos de feijão;
- (E) o consumo médio de feijão de uma pessoa desempregada (sem renda) é de 4,25 quilos de feijão anuais.

**44**

Após formular e estimar um modelo de regressão simples, o estatístico responsável pela análise trabalha nos resultados, defrontando-se com a tabela a seguir:

Fonte	S. Quadrados	G.L.	Q. Médio	F-Snedecor	p-valor
Euação	450	1	450	12,00	0,21%
Resíduos	300	8	37,50		
Total	750	9	83,33		

A partir desses números, é correto concluir que:

- (A) com uma amostra de tamanho 10, o modelo é capaz de explicar 60% da variação total;
- (B) a variância estimada do erro aleatório é inferior a 6;
- (C) apesar de um poder de explicação de 60%, o modelo não passa no teste de significância da estatística F;
- (D) a variância da variável explicativa do modelo é igual a 75;
- (E) a estimativa do coeficiente angular da equação é igual a 2.

**45**

Os processos de seleção amostral podem ser de dois tipos, probabilísticos e não probabilísticos.

Sobre cada um desses tipos de obtenção das amostras, pode-se destacar que:

- (A) nos probabilísticos, todos os indivíduos da população têm igual probabilidade de ser selecionado;
- (B) nos probabilísticos, a tarefa de seleção é simplificada pela dispensa de um cadastro exaustivo das unidades amostrais;
- (C) nos não probabilísticos, a aplicação das técnicas de inferência estatística são facilitadas pelo uso da distribuição uniforme;
- (D) nos probabilísticos, a distribuição de probabilidades de seleção dos indivíduos da população é conhecida;
- (E) nos não probabilísticos, a probabilidade de seleção de cada indivíduo amostrado só é conhecida *a posteriori*.

**46**

Dentre os métodos de amostragem probabilística, podem ser destacados os de amostras aleatórias simples, sistemáticas, por estratos, por cluster e múltipla.

Sobre cada um desses métodos, e nessa mesma ordem, poderiam ser associadas, relativamente, as seguintes palavras-chave ou expressões:

- (A) equiprováveis, periodicidade, grupos homogêneos, grupos heterogêneos e sequencial;
- (B) uniforme, sem cadastro, maior variância por grupo, menor variância por grupo e parada endógena;
- (C) uniforme, periodicidade, menor custo, maior custo e geométrica;
- (D) hipergeométrica, cadastral, grupos homogêneos, grupos heterogêneos e geométrica;
- (E) equiprováveis, periodicidade, grupos heterogêneos, grupos homogêneos e "por etapas".

**47**

Dentre os métodos de amostragem não probabilística, podem ser destacados os realizados por conveniência, por cotas, por julgamento, por tipicidade e as bolas de neve.

Sobre cada um dos métodos, e nessa exata ordem, poderiam ser associadas às seguintes palavras-chave ou expressões:

- (A) praticidade, efeito de estratificação, arbitragem, para um subgrupo e indicações técnicas;
- (B) proximidade, juízo de valor, para um subgrupo, avaliações em sequência e baixíssimo custo;
- (C) baixo custo, arbitragem, seleção endógena, efeito cluster e para um subgrupo;
- (D) seleção endógena, para um subgrupo, indicações técnicas, efeito de estratificação e efeito cluster;
- (E) proximidade, avaliações em sequência, baixíssimo custo, efeito de estratificação e longa duração.

**48**

Uma das mais importantes características que uma amostra pode ter em relação à população é a representatividade.

Tal característica notabiliza-se especialmente pelo seguinte:

- (A) todos os indivíduos da população estarem representados, de alguma forma, na amostra;
- (B) todos os grupos de indivíduos estarem representados, de alguma forma, na amostra;
- (C) a probabilidade de indivíduos de grupos populacionais distintos terem a mesma probabilidade de serem sorteados;
- (D) as proporções populacionais, para os fins de interesse do levantamento, se reproduzirem no nível da amostra;
- (E) o conjunto de unidades da amostra, gerada a partir da população, seja o mais homogêneo possível.

**49**

Sejam dois eventos quaisquer de um mesmo espaço amostral  $S$ , independentes, tais que  $P(A \cup B) = 0,76$ .

Então é correto afirmar que:

- (A)  $A$  e  $B$  são mutuamente exclusivos;
- (B) se  $P(A) + P(B) = 0,26$ , então  $P(A \cap B) = 0,5$ ;
- (C) se  $P(A \cap B) > 0,24$ , então  $A \cup B = S$ ;
- (D)  $P(B) = 0,40$ , então  $P(A) = 0,60$ ;
- (E) se  $P(A) + P(B) = 1$ ,  $A \cup B = S$ .

**50**

Para estimar uma proporção populacional, será extraída uma amostra de tamanho  $n$ , usando a proporção amostral como estimador. Em pesquisas anteriores o valor da proporção foi avaliado em  $\left(\frac{9}{25}\right)$ . O erro máximo tolerado é de  $0,02$  ( $=E$ ) e o grau de confiança de  $95\%$  ( $z_{0,05} = 2$ ). Então, para os casos de variâncias estimadas através do uso da evidência empírica do passado e do valor máximo para proporções, os respectivos tamanhos ótimos de amostras são:

- (A) 1.600 e 2.560;
- (B) 8.600 e 1.280;
- (C) 2.048 e 5.000;
- (D) 4.096 e 10.000;
- (E) 2.304 e 2.500.



Realização

