



FUNDAÇÃO INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE

EDITAL DE PROCESSO SELETIVO SIMPLIFICADO Nº02/2019

NÍVEL SUPERIOR

ANALISTA CENSITÁRIO

MÉTODOS QUANTITATIVOS

Nome do Candidato _____

Inscrição _____



COMPOSIÇÃO DO CADERNO

Língua Portuguesa 01 a 20

Conhecimentos Específicos 21 a 60



INSTRUÇÕES

1. Confira seu nome, o número do seu documento e o número de sua inscrição na Folha de Respostas. Além disso, não se esqueça de conferir seu Caderno de Questões quanto a falhas de impressão e de numeração. Preencha os campos destinados à assinatura e ao número de inscrição. Qualquer divergência, comunique ao fiscal.
2. O único documento válido para avaliação é a Folha de Respostas. Só é permitido o uso de caneta esferográfica transparente de cor azul ou preta para o preenchimento da Folha de Respostas, que deve ser preenchida da seguinte maneira: ●
3. O prazo de realização da prova é de 4 (quatro) horas, incluindo a marcação da Folha de Respostas. Após 60 (sessenta) minutos do início da prova, o candidato estará liberado para utilizar o sanitário ou deixar definitivamente o local de aplicação, não podendo, no entanto, levar o Caderno de Questões e nenhum tipo de anotação de suas respostas.
4. Ao término de sua prova, comunique ao fiscal, devolvendo-lhe a Folha de Respostas devidamente preenchida e assinada. O candidato somente poderá levar consigo o Caderno de Questões, ao final da prova, se sua saída ocorrer nos 30 (trinta) minutos anteriores ao horário determinado para o término das provas.
5. Os 3 (três) últimos candidatos só poderão retirar-se da sala juntos, após assinatura do Termo de Fechamento do envelope de retorno.
6. As provas e os gabaritos preliminares estarão disponíveis no site do Instituto AOCP - www.institutoaocp.org.br, no dia posterior à aplicação da prova.
7. O NÃO cumprimento a qualquer uma das determinações constantes em Edital, no presente Caderno ou na Folha de Respostas incorrerá na eliminação do candidato.

PROVA
01

LEMBRE-SE DE MARCAR O NÚMERO
CORRESPONDENTE À SUA PROVA NA
FOLHA DE RESPOSTAS!

ATENÇÃO!

NÃO SE ESQUEÇA de marcar, na Folha de Respostas, o número de sua prova indicado na capa deste caderno.

Língua Portuguesa

TEXTO I

O último paradoxo da vida moderna: por que ficamos presos ao celular, mas odiamos falar por telefone?

Não deixe uma ligação rápida arruinar uma longa e confusa série de mensagens de WhatsApp

SILVIA LÓPEZ

Para iniciar um texto, Hemingway dizia a si mesmo: “Escreva a frase mais verdadeira que você conhece”. Neste caso, seria: a psicóloga Cristina Pérez, do Siquia, respondeu por meio de mensagens de áudio às perguntas que lhe enviamos por email. Essa curiosidade metajornalística não tem importância, não altera a qualidade de suas respostas, só ilustra a variedade e fluidez de opções com as quais podemos nos comunicar hoje. Recebemos um email? Respondemos com um áudio. Chegou um áudio de WhatsApp? Respondemos com um texto. Recebemos um telefonema? Não respondemos. Esperamos. Esperamos. E escrevemos: “Você me ligou? Não posso falar, é melhor me escrever”. O paradoxo do grande vício do século XXI é que estamos presos ao celular, mas temos fobia das ligações telefônicas.

A ligação telefônica – que, até não muito tempo atrás, esperávamos com alegria ou tolerávamos com resignação, mas nunca evitávamos com uma rejeição universal – se tornou uma presença intrusiva e incômoda, perturbadora e tirânica, mas por quê? “Uma das razões é que quando recebemos uma ligação, ela interrompe algo que estávamos fazendo, ou simplesmente não temos vontade de falar nesse momento”, explica a psicóloga Cristina Pérez. “Por outro lado, também exige de nós uma resposta imediata, ao contrário do que ocorre na comunicação escrita, que nos permite pensar bem no que queremos dizer. E a terceira razão seria o fato de não poder saber de antemão qual será a duração do telefonema”, acrescenta.

Adaptado de: <https://brasil.elpais.com/brasil/2019/06/01/tecnologia/1559392400_168692.html>. Acesso em: 25 jun. 2019

As questões de 1 a 10 se referem ao texto I.

- 1. Em relação ao excerto “Não deixe uma ligação rápida arruinar uma longa e confusa série de mensagens de WhatsApp”, é correto afirmar que**
 - (A) ele é um conselho para as pessoas não receberem ligações rápidas.
 - (B) a palavra “rápida” caracteriza “ligação” enquanto as palavras “longa”, “confusa” e “série” caracterizam “mensagens”.
 - (C) ele é irônico, pois o conteúdo que se pretende veicular possui o significado contrário daquilo que é posto.
 - (D) ele está no modo imperativo, indicando um pedido para que as pessoas não diminuam o uso de WhatsApp.
 - (E) a palavra “série” está sendo utilizada com o mesmo significado que na frase: “A antiga 5ª série equivale hoje ao 6º ano”.
- 2. De acordo com o texto, é correto afirmar que**
 - (A) a autora seguiu o conselho de Hemingway.
 - (B) a curiosidade metajornalística citada pela autora não tem importância para a temática do texto.
 - (C) não há resposta para a pergunta presente no título.
 - (D) a ligação telefônica está presente na vida das pessoas.
 - (E) respondemos e-mails com áudios porque nos comunicamos melhor oralmente do que pela escrita.
- 3. No segundo parágrafo do texto, a informação entre travessões**
 - (A) descreve três atitudes negativas em relação à ligação telefônica.
 - (B) é constituída por três ações que se excluem mutuamente.
 - (C) é constituída por verbos que não requerem um complemento.
 - (D) restringe o sentido do termo “ligação telefônica”.
 - (E) completa o sentido do termo “ligação telefônica”.

4. Assinale a alternativa que apresenta a reescrita adequada, sintática e semanticamente, para o seguinte excerto: “O paradoxo do grande vício do século XXI é que estamos presos ao celular, mas temos fobia das ligações telefônicas.”.
- (A) O paradoxo do grande vício do século consecutivo é estarmos presos ao celular, mas termos fobia das ligações telefônicas.
- (B) A incoerência do grande vício do século consecutivo é estarmos presos ao celular, mas ficarmos doentes sem as ligações telefônicas.
- (C) No século XXI, é contraditório estar preso ao celular por ter medo das ligações telefônicas.
- (D) Temos medo das ligações telefônicas porque estamos presos ao celular; esse é o paradoxo do grande vício do século XXI.
- (E) Estarmos presos ao celular, mas termos fobia das ligações telefônicas é a incoerência do grande vício do século vigente.
5. Assinale a alternativa em que a expressão em destaque NÃO indica circunstância de tempo, NÃO sendo, portanto, um adjunto adverbial de tempo ou uma oração subordinada adverbial temporal.
- (A) “O paradoxo do grande vício do século XXI é que estamos presos ao celular [...]”.
- (B) “[...] quando recebemos uma ligação, ela interrompe algo [...]”.
- (C) “[...] simplesmente não temos vontade de falar nesse momento [...]”.
- (D) “[...] saber de antemão qual será a duração do telefonema [...]”.
- (E) “[...] opções com as quais podemos nos comunicar hoje.”.
6. Qual é a relação de sentido estabelecida no excerto “Não posso falar, é melhor me escrever”?
- (A) Contraste.
- (B) Causalidade.
- (C) Adição.
- (D) Conformidade.
- (E) Finalidade.
7. Em “[...] esperávamos com alegria ou tolerávamos com resignação, mas nunca evitávamos com uma rejeição universal.”, há uma relação de oposição entre esperar com alegria e tolerar com resignação.
- (A) há uma relação de oposição entre esperar com alegria e tolerar com resignação.
- (B) a acentuação de “esperávamos”, “tolerávamos” e “evitávamos” se deve ao fato de serem verbos.
- (C) não é possível identificar os sujeitos dos verbos presentes no excerto.
- (D) todos os verbos apresentam o mesmo sujeito e estão no mesmo modo.
- (E) o uso da vírgula é facultativo.
8. É um sinônimo da palavra “ligação” o vocábulo
- (A) “email”.
- (B) “telefone”.
- (C) “telefonema”.
- (D) “celular”.
- (E) “áudio de WhatsApp”.
9. A repetição de “esperamos”, em “Recebemos um telefonema? Não respondemos. Esperamos. Esperamos. E escrevemos: [...]”,
- (A) indica que o retorno é dado depois de um tempo considerável após a ligação.
- (B) indica que o retorno é dado imediatamente após a ligação.
- (C) não dá nenhum indício sobre o tempo transcorrido até o retorno.
- (D) deveria ser eliminada, pois se trata de um problema de coesão.
- (E) enfatiza a pouca frequência com que se recebem telefonemas.
10. Assinale a alternativa que apresenta a reescrita adequada, sintática e semanticamente, para o seguinte excerto: “[...] ‘Você me ligou? Não posso falar, é melhor me escrever’.”.
- (A) Você ligou-me? Não posso falar, é melhor eu escrever.
- (B) Ligastes para mim? Não posso falar, escreve a mim que és melhor.
- (C) Me ligou? Não posso falar, escrever é melhor para mim.
- (D) Você ligou para mim? Não posso falar, é melhor que me escreva.
- (E) Me ligou? Não posso falar, me escreve que é melhor.

TEXTO II

Você é o que você se diz: a ciência do diálogo interno

PILAR JERICÓ

Se você quiser variar a percepção que tem sobre você, precisa alterar seu diálogo interior. A forma como você conversa consigo mesmo condiciona sua capacidade de enfrentar as dificuldades e determina a tomada de decisões. A autoafirmação, ou pensar coisas positivas sobre nós mesmos, é uma ferramenta muito útil para reforçar a autoestima. Entretanto, não vale qualquer comentário. Já ficou comprovado que frases como “aguento tudo” ou “sou uma pessoa superagradável” não ajudam muito. Quem as expressa não está realmente convencido disso, então essas expressões podem ter efeito contrário. A ciência do diálogo interior nos dá pistas sobre as técnicas que tornam nossas autoafirmações eficazes: devemos imaginar futuras situações agradáveis e nos tratar na segunda pessoa.

Adaptado de: <https://brasil.elpais.com/brasil/2019/05/05/ciencia/1557083642_455016.html>. Acesso em: 25 jun. 2019.

As questões de 11 a 20 se referem ao texto II.

11. Assinale a alternativa que apresenta um exemplo de tratamento na segunda pessoa do singular.

- (A) Eu sou capaz de vencer meus desafios.
- (B) Juntos, nós podemos transformar o mundo.
- (C) Você é forte para superar os obstáculos da vida.
- (D) As pessoas não vão abalar minha autoconfiança.
- (E) Sempre é possível evoluir como ser humano.

12. Em “Se você quiser variar a percepção que tem sobre você, precisa alterar seu diálogo interior.”, a relação de sentido que se estabelece é de

- (A) condição.
- (B) tempo.
- (C) conclusão.
- (D) concessão.
- (E) contraste.

13. Em “A forma como você conversa consigo mesmo condiciona sua capacidade [...]”, o termo em destaque poderia ser substituído adequadamente por

- (A) “que”.
- (B) “cuja”.
- (C) “onde”.
- (D) “para qual”.
- (E) “pela qual”.

14. Assinale a alternativa em que a palavra formada, assim como “autoafirmação” e “superagradável”, é grafada sem hífen.

- (A) auto + hipnose.
- (B) auto + observação.
- (C) super + herói.
- (D) super + requintado.
- (E) super + salário.

15. De acordo com o texto, é correto afirmar que

- (A) conversar consigo mesmo pode combater a depressão.
- (B) a imaginação é importante no processo de melhora da autoestima.
- (C) as pessoas que não conversam com elas mesmas não têm autoestima.
- (D) falar consigo mesmo na primeira pessoa não ajuda porque isso não reflete como as outras pessoas falam conosco.
- (E) o diálogo interior só é possível a partir de um diálogo exterior, no qual está presente uma segunda pessoa.

16. Em “A autoafirmação, ou pensar coisas positivas sobre nós mesmos, é uma ferramenta muito útil [...]”, a oração entre vírgulas tem qual função?

- (A) Apresentar uma alternativa para a autoafirmação.
- (B) Fornecer um exemplo de autoafirmação.
- (C) Reiterar a utilidade da autoafirmação.
- (D) Explicar o que é autoafirmação.
- (E) Caracterizar positivamente a autoafirmação.

17. Em “[...] frases como “aguento tudo” ou “sou uma pessoa superagradável” não ajudam muito.”, as aspas destacadas

- (A) reforçam que as expressões não pertencem à norma culta.
- (B) retratam ironia.
- (C) indicam citações.
- (D) apontam que se trata de neologismos.
- (E) poderiam ser retiradas sem que isso gerasse prejuízo sintático ou semântico.

18. Assinale a alternativa correta em relação ao excerto: “A ciência do diálogo interior nos dá pistas sobre as técnicas que tornam nossas autoafirmações eficazes [...]”.

- (A) A oração adjetiva restringe o sentido da palavra “técnicas”.
- (B) As palavras “ciência” e “interior” caracterizam o termo “diálogo”.
- (C) O verbo “dá” é transitivo direto, isto é, requer um complemento sem preposição.
- (D) Tanto “do” quanto “sobre” são preposições exigidas pelos substantivos que as precedem.
- (E) A substituição de “autoafirmações eficazes” por “eficazes autoafirmações” não prejudicaria a sintaxe ou a semântica do período.

19. Assinale a alternativa INCORRETA em relação à concordância verbal.

- (A) Pistas são nos dadas pela ciência do diálogo interior.
- (B) A ciência do diálogo interior dá-nos pistas.
- (C) A nós a ciência do diálogo interior dá pistas.
- (D) A ciência do diálogo interior dá pistas para nós.
- (E) Dá-se pistas para nós pela ciência do diálogo interior.

20. Assinale a alternativa em que o verbo auxiliar indica possibilidade.

- (A) “[...] devemos imaginar futuras situações [...]”.
- (B) “[...] essas expressões podem ter efeito contrário.”.
- (C) “Já ficou comprovado [...]”.
- (D) “Se você quiser variar a percepção [...]”.
- (E) “[...] precisa alterar seu diálogo interior.”.

21. Uma pesquisa sobre gastos semanais com alimentação, em reais, foi realizada em uma cidade de 400000 residências. Para compor a amostra, foram selecionadas, aleatoriamente, 100 residências e a variável gastos semanais, representada por Y, resultou nos seguintes valores:

$$\sum_{i=1}^{100} y_i = 11000,00 \quad \text{e} \quad \sum_{i=1}^{100} y_i^2 = 1900000,00.$$

Assuma que a variável gastos semanais com alimentação na cidade pesquisada tem distribuição de frequências em forma de sino (simétrica em torno da média) e assinale a alternativa que apresenta o desvio-padrão e a porcentagem das residências que se espera que seus gastos estejam nos intervalos $\bar{y} \pm s$, $\bar{y} \pm 2s$ e $\bar{y} \pm 3s$ (aproximadamente).

- (A) 83,06, 68%, 95%, 99,7%, respectivamente.
- (B) 83,4847, 65%, 90%, 99%, respectivamente.
- (C) 6969,7, 65%, 90%, 98,7%, respectivamente.
- (D) 83,4847, 68%, 95%, 99,7%, respectivamente.
- (E) 83,06, 65%, 90%, 99%, respectivamente.

22. A seguir tem-se o tempo gasto, em minutos, por 9 atletas para a realização de uma corrida de 25 km:

55; 45; 60; 58; 52; 56; 43; 59; 47.

Diante do exposto, assinale a alternativa que apresenta as principais medidas descritivas de dispersão.

- (A) $S^2 = 36$; Desvio médio = 5,111; $S = 6,000$.
(B) $\bar{X} = 53$; Mediana = 55; Moda = 45.
(C) $\bar{X} = 53$; Mediana = 51; Moda = 45.
(D) $S^2 = 36$; $\bar{X} = 53$; Mediana = 55.
(E) $\bar{X} = 53$; $S^2 = 36$; Moda = 45.
23. Em uma escola que atende alunos de primeira a quinta série, existem duas turmas de cada série. A diretora dessa escola pretende fazer um estudo quanto à aprendizagem comparando as turmas dentro de cada série. Quais medidas estatísticas descritivas ela pode utilizar para avaliar a homogeneidade das turmas e a diferença de aprendizagem entre as turmas?
- (A) Mediana e moda.
(B) Média e variância.
(C) Variância e desvio-padrão.
(D) Variância e coeficiente de variação.
(E) Média e coeficiente de variação.
24. O gerente de um supermercado da periferia de uma cidade fez um levantamento das vendas de pés de alface durante 30 dias com o objetivo de prever a compra diária do produto. O resultado se encontra na seguinte tabela.

Pés de alface	Frequência de vendas
0	9
1	10
2	8
3	2
4	1

Nesse caso, quais são o número médio, a mediana e a moda de pés de alface vendidos por dia?

- (A) 9; 1; 1, respectivamente.
(B) 4; 3; 1, respectivamente.
(C) 5; 3; 3, respectivamente.
(D) 6; 10; 10, respectivamente.
(E) 1,2; 1; 1, respectivamente.

25. A palavra momento é usada com frequência em contextos estatísticos e refere-se à soma dos desvios de uma variável X em relação à média relacionada ao tamanho da amostra. A fórmula geral para o cálculo dos momentos amostrais em torno da média é dada por:

$$M_k = \frac{\sum_{i=1}^n (x_i - \bar{x})^k}{n} \quad \left(M_k = \frac{\sum_{i=1}^c f_i (x_i - \bar{x})^k}{\sum_{i=1}^c f_i}, \text{ para dados agrupado onde "c" é o número de classes} \right)$$

Sendo M_k representante do k -ésimo momento em torno da média amostral, considerando uma variável X , é correto afirmar que

- (A) o primeiro momento em torno da média é a variância.
 (B) o primeiro momento em torno da média é a média e o segundo a variância.
 (C) o primeiro momento em torno da média é a moda.
 (D) o primeiro momento em torno da média é sempre igual a zero.
 (E) o primeiro momento em torno da média é o desvio-padrão e o segundo a variância.
26. Os valores a seguir representam os salários, em números de salários mínimos (s.m.), de um grupo de 6 operários de uma empresa:

3 s.m.; 2 s.m.; 3,7 s.m.; 5 s.m.; 2,7 s.m.; 3 s.m.

Diante do exposto, assinale a alternativa que representa a seguinte medida relativa de assimetria: $a_3 = \frac{m_3}{s^3}$ (aproximada na segunda casa decimal) para essas medidas.

- (A) 0,874
 (B) 0,817
 (C) 0,602
 (D) 0,935
 (E) 0,737
27. Para um conjunto de dados, uma medida de dispersão alternativa ao desvio-padrão é o intervalo interquartil, ou amplitude interquartilica, definida por: $d_q = q_3 - q_1$ em que q_1 e q_3 denotam o primeiro e o terceiro quartil, respectivamente. Um possível *outlier* será qualquer ponto do conjunto localizado
- (A) abaixo de $q_1 - q_3$ e acima de $q_3 - q_1$
 (B) abaixo de $q_1 - 1,5 \times d_q$ e acima de $q_3 + 1,5 \times d_q$
 (C) abaixo de $q_1 - 1,5 \times M_d$ e acima de $q_3 + 1,5 \times M_d$
 (D) abaixo de $q_1 - 1,5 \times M_d$ e acima de $q_3 - 1,5 \times M_d$
 (E) abaixo de $q_1 - 1,5 \times M_o$ e acima de $q_3 + 1,5 \times M_o$
28. Considere o seguinte espaço amostral $S = \{1, 2, 3, 4, 5, 6, 7, 8\}$. Sendo os eventos de S : $A = \{1, 2, 3, 5, 7\}$ e $B = \{2, 5, 7\}$, assinale a alternativa correta.
- (A) $B \subset A$ e $P(B) > P(A)$
 (B) $B \supset A$ e $P(B) > P(A)$
 (C) $B \subset A$ e $P(B) < P(A)$
 (D) $B \subset A$ e $P(B) = P(A)$
 (E) $B \supset A$ e $P(B) = P(A)$

29. Em uma pesquisa cujo objetivo foi investigar as condições do saneamento básico do País junto às prefeituras municipais e empresas contratadas para a prestação desses serviços, foram incluídas questões como captação, adequação da qualidade e fornecimento de água através de rede geral de distribuição, entre outras. Os resultados obtidos para um pequeno município de 10000 habitantes, quanto à satisfação com as medidas de prevenção de alagamentos e erosões nos municípios, estão sumarizados na Tabela a seguir, sendo que os sujeitos participantes da pesquisa foram classificados em: morador da cidade (C), subúrbio (S) e morador da área rural (R).

Tabela – Resultados da pesquisa quanto à prevenção de alagamentos e erosões

	Satisfeito (com as medidas)	Insatisfeito (com as medidas)
C	1500	2500
S	2500	1500
R	500	1500

Definindo os eventos:

A : a pessoa escolhida vive na cidade

B: a pessoa não aprova as medidas de prevenção de alagamentos e erosões do município.

Diante do exposto, a probabilidade do evento $D = \{A^c \cap B\}$ é

- (A) $P(D) = 0,55$.
- (B) $P(D) = 0,7$.
- (C) $P(D) = 0,33$.
- (D) $P(D) = 0,30$.
- (E) $P(D) = 0,82$.

30. Uma empresa financiadora de veículos constatou que 30% dos seus clientes não pagam as prestações mensais. Consultando os arquivos da empresa, foi observado que 85% dos não pagadores eram aqueles cujo valor da prestação representava 25% da renda familiar ou mais e que, para 75% dos bons pagadores, a mensalidade representava menos que 25% da renda familiar. Qual é a probabilidade de um cliente não pagar

a prestação, dado que sua prestação é elevada em relação a sua renda familiar? (Definindo: P= “Pagar”; NP= “Não pagar”; E= “prestação em relação à renda é maior que 25%” e E^c= “prestação em relação à renda menor ou igual a 25%”).

- (A) 0,59
- (B) 0,85
- (C) 0,43
- (D) 0,25
- (E) 0,20

31. Seja X_1, X_2, \dots, X_n uma amostra aleatória de uma variável aleatória com distribuição normal de média θ e variância 1 ($X_1, X_2, \dots, X_n \sim N(\theta, 1)$), considerando a variável aleatória máxima da amostra $Y = \max(X_1, X_2, \dots, X_n)$, a função de densidade de Y é

- (A) $f(y) = n\{\Phi(y - \theta)\}^{n-1}\phi(y - \theta)$, y pertencente aos reais.
- (B) $f(y) = \{\Phi(y - \theta)\}^n\phi(y - \theta)$, y pertencente aos reais.
- (C) $f(y) = \{\Phi(y - \theta)\}^n$, y pertencente aos reais.
- (D) $f(y) = \{\Phi(y - \theta)\}^{n-1}\phi(y - \theta)$, y pertencente aos reais.
- (E) $f(y) = \{\Phi(y - \theta)\}^{n-1}$, y pertencente aos reais.

32. Suponha que, em determinada esquina de uma rua, ocorram acidentes de carro aleatoriamente e independente um do outro, em uma taxa média de três por semana. Então, a probabilidade de que ocorra exatamente um acidente na primeira semana de agosto é

- (A) 0,5976.
- (B) 0,1494.
- (C) 0,4024.
- (D) 0,8506.
- (E) 0,3528.

33. Sendo X e Y duas variáveis aleatórias contínuas com função de densidade conjunta dada por:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{4}(2x+y), & 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0, & \text{caso contrário} \end{cases}$$

e considerando as variáveis aleatórias X e Y e a função de densidade conjunta, é correto afirmar que

- (A) $f_X(x) = \frac{1}{2}(x+1), 0 \leq x \leq 1.$
- (B) $f_Y(y) = \frac{1}{4}(2+y), 0 \leq y \leq 2.$
- (C) o coeficiente de correlação entre X e Y é igual a 0 (zero).
- (D) $E[X|Y=y] = \frac{4+3y}{6(1+y)}.$
- (E) as variáveis aleatórias X e Y são independentes.

34. Seja X uma variável aleatória com distribuição normal de média μ e variância σ^2 , assim,

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)/2\sigma^2}, \quad -\infty \leq x \leq \infty, \quad -\infty \leq \mu \leq \infty \text{ e } \sigma^2 > 0. \text{ Seja } Y = \int_0^X \frac{1}{\sqrt{2\pi\sigma}} e^{-(u-\mu)/2\sigma^2} du,$$

então, $Y=F(X)$, em que F é a função de distribuição acumulada de uma variável aleatória com distribuição normal padrão. Com essas informações, é correto afirmar que

- (A) Y tem distribuição normal padrão.
- (B) A média de Y é igual a μ e a variância é igual a σ^2 .
- (C) Y tem distribuição normal com a μ e variância igual a σ^2 .
- (D) Y tem distribuição qui-quadrado, $0 \leq Y \leq \infty$.
- (E) Y tem distribuição uniforme em (0, 1).

35. Considere as variáveis aleatórias X_1, X_2, \dots, X_n , independentes e identicamente distribuídas conforme a distribuição de Bernoulli, com parâmetro p . Seja $S_n = \sum_{i=1}^n X_i$, em relação à probabilidade da média das variáveis aleatórias X_1, X_2, \dots, X_n pertencer ao intervalo $(p-c; p+c)$, é correto afirmar que

- (A) $P(p-c \leq \frac{S_n}{n} \leq p+c) = P\left(\left|\frac{S_n}{n} - p\right| \geq c\right) \geq 1 - \frac{p(1-p)}{nc^2} \rightarrow 1$, quando $n \rightarrow \infty$
- (B) $P(p-c \leq \frac{S_n}{n} \leq p+c) = P\left(\left|\frac{S_n}{n} - p\right| \leq c\right) \geq \frac{p(1-p)}{nc^2} \rightarrow 0$, quando $n \rightarrow \infty$
- (C) $P(p-c \leq \frac{S_n}{n} \leq p+c) = P\left(\left|\frac{S_n}{n} - p\right| \geq c\right) \geq \frac{p(1-p)}{nc^2} \rightarrow 0$, quando $n \rightarrow \infty$
- (D) $P(p+c \leq \frac{S_n}{n} \leq p-c) = P\left(\left|\frac{S_n}{n} - p\right| \leq c\right) \geq 1 - \frac{p(1-p)}{nc^2} \rightarrow 1$, quando $n \rightarrow \infty$
- (E) $P(p-c \leq \frac{S_n}{n} \leq p+c) = P\left(\left|\frac{S_n}{n} - p\right| \leq c\right) \geq 1 - \frac{p(1-p)}{nc^2} \rightarrow 1$, quando $n \rightarrow \infty$

36. Sendo X_1, X_2, \dots, X_k variáveis aleatórias independentes e distribuídas conforme a distribuição Normal de média μ_i e variância σ_i^2 , $i=1, 2, \dots, k$, considere as seguintes igualdades:

$$U = \sum_{i=1}^k \left(\frac{X_i - \mu_i}{\sigma}\right)^2 \quad V = \sum_{i=1}^k \left(\frac{X_i - \mu_i}{\sigma_i^2}\right)^2 \quad W = \sum_{i=1}^k \left(\frac{X_i - \mu_i}{\sigma_i}\right) \quad L = \sum_{i=1}^k \left(\frac{X_i - \mu_i}{\sigma_i^2}\right)^2$$

Informe se é verdadeiro (V) ou falso (F) o que se afirma a seguir e assinale a alternativa com a sequência correta.

- () U tem distribuição qui-quadrado com k graus de liberdade.
- () V tem distribuição t de Student com k graus de liberdade.
- () W é a soma de k variáveis aleatórias normal padrão.
- () U tem distribuição qui-quadrado com $(k-1)$ graus de liberdade.
- () L tem distribuição F de Snedecor.

- (A) V – F – V – V – F.
- (B) F – F – V – F – V.
- (C) V – F – V – F – F.
- (D) V – F – F – V – F.
- (E) F – V – F – V – F.

37. Sendo (X_1, X_2) uma amostra aleatória independente de uma variável aleatória normal padrão, informe se é verdadeiro (V) ou falso (F) o que se afirma a seguir e assinale a alternativa com a sequência correta.

- () $\frac{X_2 - X_1}{\sqrt{2}}$ tem distribuição normal padrão.

() $\frac{X_2 - X_1}{\sqrt{2}}$ não tem distribuição normal padrão.

() $\frac{(X_1 + X_2)^2}{(X_1 - X_2)^2}$ tem distribuição F(2, 2).

() $\frac{(X_1 + X_2)^2}{(X_1 - X_2)^2}$ tem distribuição F(1, 1).

() $\frac{(X_1 + X_2)}{\sqrt{(X_1 - X_2)^2}}$ tem distribuição t com 1 grau de liberdade.

(A) V - F - F - V - V.

(B) V - F - V - F - V.

(C) F - F - V - F - F.

(D) V - F - F - V - F.

(E) F - V - F - V - F.

38. O tempo de atendimento em um ambulatório público é de grande interesse para os gestores de um município. Uma amostra com 25 pessoas indicou tempo médio de atendimento $\bar{X} = 5,5$ minutos e desvio-padrão $S = 0,36$. Supondo que a variável tempo de atendimento seja distribuída conforme uma distribuição normal de média μ e variância igual a σ^2 , uma estatística L tal que $P(L \leq \mu) = 1 - \alpha$ representa o limite inferior do intervalo de confiança unilateral à esquerda para a média populacional μ . Com essas informações, assinale a alternativa que representa o intervalo de confiança unilateral para a média populacional do tempo de atendimento ambulatorial se

$$P(L \leq \mu) = 0,95 \text{ . (} t_{24; 0,95} = 1,711 \quad z_{0,95} = 1,65 \text{)}$$

(A) $IC_{uni;inf}(\mu; 95\%) = (-5,377; 5,377)$

(B) $IC_{uni;inf}(\mu; 95\%) = (5,377; \infty)$

(C) $IC_{uni;inf}(\mu; 95\%) = (-\infty; 5,377)$

(D) $IC_{uni;inf}(\mu; 95\%) = (5,381; \infty)$

(E) $IC_{uni;inf}(\mu; 95\%) = (-\infty; 5,381)$

39. O Instituto Brasileiro de Geografia e Estatística (IBGE) realiza diversas pesquisas que abrangem temas como Educação, Trabalho, Economia, População, Saúde e Território, a nível nacional, estadual e municipal. Suponha que seja necessário realizar um levantamento sobre o número de pessoas infectadas por certa doença em um município. Dessa forma, a variável de interesse tem distribuição Binomial. Com essas informações, assinale a alternativa que representa o menor tamanho de amostra (utilizando o teorema Chebyshev) para o qual pode-se afirmar que

$$P\left(\left|\frac{X_n}{n} - p\right| < 0,1\right) \geq 0,90.$$

(X_n : soma dos valores observados para a variável aleatória de interesse; suponha $p(1-p) \leq \frac{1}{4}$).

(A) $n=150$.

(B) $n=320$.

(C) $n=250$.

(D) $n=350$.

(E) $n=280$.

40. Um intervalo de confiança $(1 - \alpha)$ para um parâmetro $\theta \in \Theta$ é um intervalo $C_n = (a, b)$ em que $a = a(X_1, X_2, \dots, X_n)$ e $b = b(X_1, X_2, \dots, X_n)$ são funções dos dados de tal modo que $P_\theta(\theta \in C_n) \geq 1 - \alpha$ para todo $\theta \in \Theta$. Denomina-se $(1 - \alpha)$ de cobertura do intervalo de confiança. Considerando essas informações, assinale a alternativa correta a respeito da construção de intervalo de confiança.

(A) Os limites do intervalo não dependem dos valores amostrais.

(B) $C_n = (a, b)$ é aleatório e θ é fixado.

(C) θ é aleatório e $C_n = (a, b)$ é fixado.

(D) α é o nível de significância do intervalo.

(E) O intervalo definido acima é um intervalo unilateral à esquerda.

41. Sendo X uma variável aleatória com distribuição Binomial, suponha que seja de interesse testar a hipótese $H_0 : p = 0,8$ contra a hipótese $H_1 : p < 0,6$, sendo $\alpha = 0,03$, fixado. Assinale a alternativa que apresenta o erro tipo II, β , para uma amostra de 10 sujeitos e $H_1 : p = 0,6$.

(Pela Tabela $P(X \leq 5 | p = 0,8) \cong 0,03$ e $P(X \leq 5 | p = 0,6) \cong 0,367$).

- (A) 0,050
- (B) 0,970
- (C) 0,367
- (D) 0,733
- (E) 0,950

42. Seja (X_1, X_2, \dots, X_k) uma amostra aleatória de uma distribuição exponencial com parâmetro θ , o teste da razão de verossimilhança para a hipótese nula $H_0 : \theta = \theta_0$ contra a hipótese alternativa $H_1 : \theta \neq \theta_0$ rejeita a hipótese nula se $\sum_{i=1}^n X_i \geq m_1$ ou $\sum_{i=1}^n X_i \leq m_2$. Com essas informações, assinale a alternativa correta.

(A)
$$m_1 = \frac{1}{\ln k - n \ln \left(\frac{\theta_1}{\theta_0} \right)} \left(\frac{\theta_0 - \theta_1}{\theta_1 \theta_0} \right) \text{ e } m_2 = \left\{ \ln k - n \ln \left(\frac{\theta_1}{\theta_0} \right) \right\} \left(\frac{\theta_1 \theta_0}{\theta_0 - \theta_1} \right)$$

(B)
$$m_1 = \ln k - n \ln \left(\frac{\theta_1}{\theta_0} \right) \text{ e } m_2 = \left\{ \ln k + n \ln \left(\frac{\theta_1}{\theta_0} \right) \right\}$$

(C)
$$m_1 \leq \sum_{i=1}^n X_i \text{ e } m_2 \leq \sum_{i=1}^n X_i$$

(D)
$$m_1 = \frac{1}{\ln k - n \ln \left(\frac{\theta_1}{\theta_0} \right)} \left(\frac{\theta_0 + \theta_1}{\theta_1 \theta_0} \right) \text{ e } m_2 = \left\{ \ln k - n \ln \left(\frac{\theta_1}{\theta_0} \right) \right\} \left(\frac{\theta_1 \theta_0}{\theta_0 + \theta_1} \right)$$

(E)
$$m_1 = \left\{ \ln k - n \ln \left(\frac{\theta_1}{\theta_0} \right) \right\} \left(\frac{\theta_1 \theta_0}{\theta_0 - \theta_1} \right) \text{ e } m_2 = \frac{1}{\ln k - n \ln \left(\frac{\theta_1}{\theta_0} \right)} \left(\frac{\theta_0 - \theta_1}{\theta_1 \theta_0} \right)$$

43. Seja (X_1, X_2, X_3) uma amostra aleatória de tamanho $n=3$ de uma distribuição de média desconhecida μ , $-\infty < \mu < \infty$ e variância σ^2 é um número positivo. Considere os estimadores $\hat{\theta}_1 = \bar{X}$ e $\hat{\theta}_2 = \frac{2X_1 + X_2 + 5X_3}{8}$ para a média μ . Então, considerando-se as variâncias de $\hat{\theta}_1$ e de $\hat{\theta}_2$, é correto afirmar que

- (A) $\hat{\theta}_1 = \bar{X}$ é um estimador viesado melhor, no sentido que sua variância é menor.
- (B) $\hat{\theta}_2 = \frac{2X_1 + X_2 + 5X_3}{8}$ é um estimador viesado melhor, no sentido que sua variância é maior.
- (C) $\hat{\theta}_1 = \bar{X}$ é um estimador não viesado melhor, pois sua variância é menor.
- (D) $\hat{\theta}_2 = \frac{2X_1 + X_2 + 5X_3}{8}$ é um estimador não viesado melhor, no sentido que sua variância é menor.
- (E) $\hat{\theta}_2 = \frac{2X_1 + X_2 + 5X_3}{8}$ é um estimador viesado melhor, no sentido que sua variância é menor.

44. É conhecido que o erro quadrático médio (EQM($\hat{\theta}$)) mede, em média, quão perto um estimador $\hat{\theta}$ chega ao valor real do parâmetro θ . Diante do exposto, é correto afirmar que

- (A) EQM($\hat{\theta}$) não pode ser usado como critério para comparar estimadores.
- (B) EQM($\hat{\theta}$) é uma medida que combina viés e variância de um parâmetro.
- (C) para estimadores viesados, EQM($\hat{\theta}$) é igual à variância.

- (D) $\text{EQM}(\hat{\theta}) = \hat{\theta} - \text{Vies}(\hat{\theta})$.
- (E) em geral, é fácil encontrar um estimador que minimize EQM($\hat{\theta}$).

45. Seja (X_1, X_2, \dots, X_k) uma amostra aleatória de uma distribuição $N(\mu, \sigma^2)$ com variância conhecida. A estatística da razão de verossimilhança para testar valores específicos de μ é dada por $W(\mu; x) = 2\{l(\hat{\mu}_{MV}; x) - l(\mu; x)\} = \left(\frac{\bar{x} - \mu}{\sigma / \sqrt{n}}\right)^2$. Nesse caso, é correto afirmar que

- (A) a estatística de razão de verossimilhança não é um pivô exato e tem exatamente uma distribuição normal de média zero e variância um, $N(0, 1)$.
- (B) a estatística de razão de verossimilhança não é um pivô exato e tem uma distribuição t de Student com $(n-1)$ graus de liberdade.
- (C) a estatística de razão de verossimilhança é um pivô exato e tem exatamente uma distribuição de qui-quadrado com $(n-1)$ grau de liberdade, $\chi^2(n-1)$.
- (D) a estatística de razão de verossimilhança é um pivô exato e tem exatamente uma distribuição de qui-quadrado com 1 grau de liberdade, $\chi^2(1)$.
- (E) a estatística de razão de verossimilhança é um pivô exato e tem exatamente uma distribuição normal de média zero e variância um, $N(0, 1)$, pois a variância é conhecida.

46. O Ministério da Saúde vem desenvolvendo ações de combate ao mosquito *Aedes Aegypti* e eliminação das arboviroses, como zika, dengue e chikungunya. Uma dessas ações é a conscientização da população para a importância do permanente combate ao mosquito transmissor. Supondo que uma pesquisa para inferir sobre o conhecimento que a população de um município tem sobre o assunto precise ser planejada e que uma pesquisa piloto mostrou que 60% da população do município têm conhecimento sobre o assunto, determine o tamanho da amostra (número de sujeitos na amostra) de forma que a proporção de sujeitos na amostra difira de menos de 0,02 da proporção verdadeira da população, com 99% de confiança com base no resultado da pesquisa piloto. ($Z_{0,025} = 1,96$, $Z_{0,01} = 2,33$, $Z_{0,005} = 2,58$).

- (A) $n=2305$.
- (B) $n=3258$.
- (C) $n=3994$.
- (D) $n=3394$.
- (E) $n=4161$.

47. Considere uma população de empresas de prestação de serviços que pode ser dividida em 3 estratos quanto ao número de trabalhadores que emprega: pequenas – 50 ou menos trabalhadores; médias – entre 51 e 100; grandes – mais de 101 trabalhadores. Supondo que a população de interesse é constituída por 1800 empresas, destas 45% são pequenas, 35% são médias e 20% grandes. Se o tamanho da amostra for fixado em 120 empresas, selecionadas com igualdade de proporção (estratificada proporcional), quantas empresas devem ser selecionadas de cada estrato?

- (A) 40 empresas de cada estrato.
- (B) 54, 42, 24, respectivamente.
- (C) 45, 35, 20, respectivamente.
- (D) 50, 40, 30, respectivamente.
- (E) 55, 43, 20, respectivamente.

48. Relacione as colunas e assinale a alternativa com a sequência correta.

1. Amostra Intencional.
 2. Amostra Snowball ou Bola de Neve.
 3. Amostra por Conveniência.
- () O investigador localiza, de início, um grupo de indivíduos que tenha as características desejadas, ou que consiga indicar indivíduos que as tenham.
 - () Neste tipo de amostragem, a escolha dos indivíduos é feita não tanto pela “representatividade”, mas porque eles podem prestar a colaboração de que se necessita.
 - () Neste tipo de amostra, os elementos são escolhidos porque se encontram onde os dados estão a ser recolhidos – a sua participação, no estudo, é como que “acidental”.
 - () A amostra, assim escolhida, pode não ser representativa da População, mas é de interesse para o investigador.

- (A) 2 – 1 – 3 – 1.
- (B) 3 – 1 – 1 – 2.
- (C) 1 – 2 – 2 – 3.
- (D) 3 – 1 – 2 – 3.
- (E) 1 – 2 – 3 – 3.

49. O número médio de pacientes que um médico atende durante um dia de trabalho foi tema de uma pesquisa. Iniciou-se o plano amostral com a ideia, a priori, de que quanto mais experiência tem um médico mais clientes tem. Isso levou a população de médicos a ser classificada em 3 grupos: os “iniciantes” (classe 1), os “intermediários” (classe 2) e os “experientes” (classe 3). Além disso, sabe-se, a partir do quadro de amostragem dos médicos, a classe de cada um (1 ou 2 ou 3). Assim, foram listados 500 médicos na classe 1, 1000 na classe 2 e 2500 na classe 3. Por meio de amostragem aleatória simples, foram selecionados 200 médicos em cada classe. Observou-se, em cada classe, o número médio de pacientes por dia para o médico amostrado: 10 na classe 1, depois 15 na classe 2 e 20 na classe 3. Diante do exposto, qual é a estimativa da média de pacientes atendidos por dia?

- (A) $\hat{Y} = 17,5$
- (B) $\hat{Y} = 15,0$
- (C) $\hat{Y} = 0,11$
- (D) $\hat{Y} = 8,6$
- (E) $\hat{Y} = 0,43$

50. A junta comercial de um estado tem os registros das empresas em atividades, as quais estão enumeradas de 22335 a 50780. Deve-se selecionar uma amostra aleatória de 350 empresas. Qual é o número de elementos dessa população e qual é o melhor método de amostragem nesse caso?

- (A) 28447; amostragem aleatória simples.
- (B) 28447; amostragem sistemática.
- (C) 28446; amostragem sistemática.
- (D) 28446; amostragem por conglomerado.
- (E) 28447; amostragem estratificada.

51. Pretende-se realizar uma amostragem sobre um conjunto de 9 regiões de uma cidade, que estão dispostas aproximadamente conforme uma matriz de 3 colunas e 3 linhas. Para facilitar o estudo, decidiu-se construir amostras de dimensão 3. A região 5 é considerada central e todas as outras são consideradas periféricas. Acredita-se que essas regiões periféricas apresentam comportamentos semelhantes quanto à variável de interesse, quando são vizinhas próximas. Foi decidido construir um plano de amostragem que permita a escolha de qualquer amostra com três unidades, mas dê menos peso à possibilidade de escolha das amostras contendo as regiões de um dos seguintes conjuntos: {1; 2; 4}, {2; 3; 6}, {4; 7; 8} ou {6; 8; 9}. A essas amostras atribui-se a probabilidade p , enquanto a todas as outras amostras possíveis se atribui a probabilidade $2p$. Com essas informações, é correto afirmar que o valor de p é

- (A) $p = \frac{1}{4}$
- (B) $p = \frac{1}{84}$
- (C) $p = \frac{1}{252}$
- (D) $p = \frac{1}{164}$
- (E) $p = \frac{1}{3}$

52. Em uma maternidade, foram observadas 20 repetições das variáveis Y: peso do bebê (kg) e X: comprimento do bebê (cm), cujos resultados foram: $\sum_{i=1}^{20} y_i = 75$
- $$\sum_{i=1}^{20} x_i = 938, \quad S_{xx} = \sum_{i=1}^{20} x_i^2 - \frac{1}{20} \left(\sum_{i=1}^{20} x_i \right)^2 = 377,8 \quad \text{e} \quad SQ \text{ Resíduo} = 8,2 \text{ (Soma de quadrados dos resíduos).}$$

Assinale a alternativa que apresenta as estimativas da variância dos estimadores $\hat{\beta}_0$ e $\hat{\beta}_1$ dos parâmetros da reta de regressão.

- (A) 592,50 e 0,2670, respectivamente.
 (B) 2,09 e 0,0008, respectivamente.
 (C) 0,49 e 0,0148, respectivamente.
 (D) 8,80 e 0,2671, respectivamente.
 (E) 32,92 e 0,0148, respectivamente.
53. Na estimação do vetor de parâmetros β , em modelos lineares, podem ser utilizados vários métodos. Os mais comuns são o Método dos Mínimos Quadrados e o Método da Máxima Verossimilhança. O propósito desses métodos é encontrar um estimador para o vetor de parâmetros β tal que o somatório dos quadrados das distâncias entre cada ponto observado e seu correspondente estimado pelo modelo seja mínimo. Para esses dois métodos citados, os estimadores obtidos são iguais, no entanto, pode ocorrer, que para métodos diferentes, resultam-se estimadores diferentes. Nesses casos, é necessário escolher o melhor estimador. Assinale a alternativa que apresenta o(s) principal(is) critério(s) utilizado(s) para avaliar um estimador $\hat{\beta}_i$ de um parâmetro β_i para um modelo linear.

Considere o seguinte:

1. Não viesado - $E(\hat{\beta}_i) = \beta_i$;
 2. Consistência - $\lim_{n \rightarrow \infty} P(|\hat{\beta}_i - \beta_i| \geq \varepsilon) = 0$ para qualquer $\varepsilon > 0$;
 3. Suficiência - quando a função de densidade de probabilidade conjunta condicional das observações amostrais, dado $\hat{\beta}_i$, não depende do parâmetro β_i ;
 4. Variância Mínima - um estimador $\hat{\beta}_i$ é de variância mínima de β_i se para qualquer outro estimador $\hat{\beta}_i^*$ $\text{var}(\hat{\beta}_i) \leq \text{var}(\hat{\beta}_i^*)$ para todo $\hat{\beta}_i^*$.
 5. Normalidade – os parâmetros $\hat{\beta}_i$ devem distribuir-se conforme a distribuição normal padrão.
- (A) Não viesado.
 (B) Não viesado, Consistência, Suficiência e Variância Mínima.
 (C) Não viesado, Suficiência, Variância Mínima e normalidade.
 (D) Não viesado, Variância Mínima e normalidade.
 (E) Viesado, Consistência, Suficiência.

54. Na aplicação de modelos lineares com mais de uma variável explicativa, a Multicolinearidade é um problema que geralmente ocorre. Para medir o grau de multicolinearidade, pode-se utilizar a matriz de correlações que mede a dependência linear de primeira ordem entre as variáveis explicativas. Utilizando essas informações, relacione as colunas e assinale a alternativa com a sequência correta.

1. Multicolinearidade Perfeita.
2. Ausência de Multicolinearidade.
3. Alto Grau de Multicolinearidade.
4. Baixo Grau de Multicolinearidade.

- () Se isso ocorrer, o coeficiente de uma variável vai depender da outra, não refletindo, assim, o efeito individual da variável à qual está associado, mas somente um efeito parcial ou marginal, condicionado a outra variável.
- () Se isso ocorrer, o coeficiente de uma variável não irá depender da outra, situação perfeita para a aplicação da análise de regressão múltipla.
- () Nesse caso, é impossível estimar o vetor de parâmetros porque a matriz $X'X$ tem o determinante igual a zero; logo, não possui inversa.
- () Quando as variáveis não estão correlacionadas entre si, a matriz $X'X$ é diagonal.

- (A) 1 – 2 – 3 – 4.
 (B) 2 – 3 – 4 – 1.
 (C) 3 – 4 – 1 – 2.
 (D) 4 – 2 – 3 – 1.
 (E) 2 – 3 – 1 – 4.

55. Um investidor solicitou a uma imobiliária os preços (Y , em mil reais) de casas aleatoriamente selecionadas de certa vizinhança suspeita, as correspondentes idades das casas (x_1 , em anos) e o tamanho (x_2 , em metros quadrados). Uma tabela de análise de variância (ANOVA) foi construída para os dados observados, conforme segue:

Fonte	Graus de liberdade	Soma de quadrados	Quadrado médio	F-Snedecor	Valor p
Regressão	2	B	478,2	D	0,286
Resíduos	2	382,7	C		
Total	A	1339,2			

Considerando as informações contidas na Tabela ANOVA, é correto afirmar que

- (A) $A=4$, $B=1721,9$, $C=191,35$ e $D=2,5$.
- (B) $R^2 = 0,99$.
- (C) não é possível estimar a variância dos erros.
- (D) o número de casas selecionadas foi $n=5$ e a regressão é significativa ao nível de 5% no teste F de Snedecor.
- (E) pelo teste F para as hipóteses " $H_0 : \beta_1 = \beta_2 = 0$ contra $H_1 : \text{pelo menos um dos } \beta_i \neq 0$ " com base no valor p, não se pode rejeitar a hipótese nula para nível de significância menor do que 30%.

56. Os modelos lineares generalizados constituem uma extensão dos modelos lineares de regressão, permitindo alargar as hipóteses admitidas. A variável resposta do modelo passa a poder provir de um universo que segue uma lei de distribuição, chamada família exponencial, deixando de ter obrigatoriamente uma distribuição Normal. Para variáveis aleatórias Y_i ($i=1, 2, \dots, N$) independentes, de média μ_i , uma das notações de função de densidade de probabilidade pertencente à família exponencial é dada por: $f(y_i | \theta_i, \phi) = \exp\left\{\frac{y_i \theta_i - b(\theta_i)}{a_i(\phi)} + c(y_i, \phi)\right\}$. Existem diversas distribuições de probabilidades que podem ser escritas na forma da distribuição da família exponencial. Diante do exposto, considere a distribuição normal de média μ e variância σ^2 , escrita na forma exponencial, e assinale a alternativa correta para $\theta = \mu$ e $\phi = \sigma^2$.

(A) $a(\phi) = 1$, $b(\theta) = \mu$, $c(y, \phi) = -\frac{1}{2}[\ln(2\pi\phi)]$

(B) $E[Y] = b'(\theta)$ e $\text{Var}[Y] = a(\phi)b'(\theta)$

(C) $f(y | \theta, \phi) = \exp\left\{\frac{y\theta - \theta^2/2}{\phi} - \frac{1}{2}\left[\frac{y^2}{\phi} + \ln(2\pi\phi)\right]\right\}$

(D) $E[Y] = b'(\theta)$ e $\text{Var}[Y] = b''(\theta)$

(E) $f(y | \theta, \phi) = \exp\left\{\frac{y\theta - \theta^2/2}{\phi^2} - \frac{1}{2}\left[\frac{y^2}{\phi^2} + \ln(2\pi\phi)\right]\right\}$

57. Na aplicação da metodologia estatística, a função geradora de momentos é de grande utilidade, pois, a partir dela, é possível determinar os momentos de uma variável aleatória. Para a família exponencial na notação $f(y_i | \theta_i, \phi) = \exp\left\{\frac{y_i \theta_i - b(\theta_i)}{a_i(\phi)} + c(y_i, \phi)\right\}$ é dada por: $M_Y(t; \theta, \phi) = E[e^{tY}] = \exp\left\{\frac{1}{a(\phi)}\{b[a(\phi)t + \theta] - b(\theta)\}\right\}$. Considere uma variável aleatória Y com distribuição normal de média μ e variância σ^2 , escrita na forma exponencial, e assinale a alternativa que apresenta sua função geradora de momentos. (Seja: $a(\phi) = \sigma^2$, $\theta = \mu$ e $b(\theta) = \theta^2/2$).

(A) $M_Y(t) = \exp\left(t\mu + \frac{\sigma^2 t^2}{2}\right)$

(B) $M_Y(t) = \exp(\mu t + \sigma^2 t^2)$

(C) $M_Y(t) = \exp\left(\mu + \frac{\sigma^2 t^2}{2}\right)$

(D) $M_Y(t) = \exp(t\mu + \sigma^2 t^2)$

(E) $M_Y(t) = \exp\left(t\mu - \frac{\sigma^2 t^2}{2}\right)$

58. Segundo o Fundo de Defesa da Citricultura (Fundecitrus), associação privada, sem fins econômicos e em benefício público da citricultura, o greening (huanglongbing/HLB) é o principal desafio fitossanitário da citricultura mundial. Em um estudo sobre 50 talhões de plantação de laranjeiras, foram registrados os dados sobre idade da planta, número de plantas no talhão e número de casos de contaminação. Utilizando um modelo linear generalizado da família exponencial com função de densidade Binomial, obteve-se a equação de valores preditos: $\hat{\eta}(x_i) = -5,395 + 0,125x_i$. A razão de chances representa o aumento na probabilidade de sucesso associado à mudança de uma unidade no valor do preditor linear. Assinale a alternativa que representa a razão de chances de árvores serem contaminadas em 1 e 5 anos de exposição.

- (A) 30% e 90% aproximadamente.
- (B) 10% e 50% aproximadamente.
- (C) 3% e 12% aproximadamente.
- (D) 37% e 86% aproximadamente.
- (E) 13% e 87% aproximadamente.

59. Suponha um fenômeno que pode ser descrito por uma série temporal cujos valores flutuam aleatoriamente em torno de um valor fixo, sem apresentar qualquer tendência. Para uma série desse tipo, um modelo razoável é dado por: $Y_t = \mu + e_t$, em que Y_t representa os valores da série, μ é o valor em torno do qual os valores flutuam e e_t são os erros aleatórios. Para a estimação do parâmetro μ , um dos métodos sugeridos é a suavização exponencial, que consiste em supor que μ é uma média ponderada dos valores passados da série. Com essas informações, assinale a alternativa que contém o modelo matemático que representa esse método (α constante de suavização $0 < \alpha < 1$).

(A) $\bar{Y}_t = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \dots$

(B) $\bar{Y}_t = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha^2(1-\alpha)Y_{t-2} + \dots$

(C) $\bar{Y}_t = \alpha Y_{t-1} + \alpha(1-\alpha)Y_{t-2} - \alpha(1-\alpha)^2 Y_{t-3} + \dots$

(D) $\bar{Y}_t = \alpha Y_t - \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)Y_{t-2} - \dots$

(E) $\bar{Y}_t = \alpha Y_t - \alpha(1-\alpha)Y_{t-1}$

60. Um modelo autorregressivo de ordem p tem a forma

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t.$$

Considere as vendas anuais das fábricas (em milhões de unidades), em todo o mundo, de carros, caminhões e ônibus fabricados pela Marca XX. Suponha que o modelo que representa essas vendas seja dado pela série $Z_t = 1,8Z_{t-1} - 0,8Z_{t-2} + a_t$. A respeito desse modelo, é correto concluir que

- (A) $\phi(L) = 1 + 1,8L + 0,8L^2$ é o seu polinômio autorregressivo.
- (B) $\phi(L) = 1 - 1,8L + 0,8L^2$ é o seu polinômio autorregressivo.
- (C) a série temporal Z_t foi gerada por um processo AR(2) estacionário.
- (D) essa série não tem raiz unitária.
- (E) $\phi(L) = 1,8L - 0,8L^2$ é o seu polinômio autorregressivo.

.....
ATENÇÃO!

NÃO SE ESQUEÇA de marcar, na Folha de Respostas, o número de sua prova indicado na capa deste caderno.
.....

