



EXÉRCITO BRASILEIRO
ESCOLA DE SAÚDE E FORMAÇÃO COMPLEMENTAR DO EXÉRCITO

CONCURSO DE ADMISSÃO/2024
PARA MATRÍCULA NO CURSO DE FORMAÇÃO DE OFICIAIS DO QUADRO COMPLEMENTAR/2025
E NO CURSO DE FORMAÇÃO DE OFICIAIS DO QUADRO DE CAPELÃES MILITARES/2025

006. PROVA OBJETIVA

CURSO DE FORMAÇÃO DE OFICIAIS DO QUADRO COMPLEMENTAR

ÁREA: ESTATÍSTICA

- Você recebeu sua folha de respostas e este caderno contendo 60 questões objetivas.
- Confira seus dados impressos na capa deste caderno e na folha de respostas.
- Certifique-se de que a letra referente ao modelo de sua prova é igual àquela constante em sua folha de respostas.
- Quando for permitido abrir o caderno, verifique se está completo ou se apresenta imperfeições.
- Caso haja alguma divergência de informação, comunique ao fiscal da sala para a devida substituição desse caderno.
- Leia cuidadosamente todas as questões e escolha a resposta que você considera correta.
- Marque, na folha de respostas, com caneta de tinta azul ou preta, a letra correspondente à alternativa que você escolheu.
- A duração da prova é de 4 horas, já incluído o tempo para o preenchimento da folha de respostas.
- Só será permitida a saída definitiva da sala e do prédio após transcorridas 3 horas do início da prova.
- Até que você saia do prédio, todas as proibições e orientações continuam válidas.

AGUARDE A ORDEM DO FISCAL PARA ABRIR ESTE CADERNO.

Nome do candidato _____

RG _____ Inscrição _____ Prédio _____ Sala _____ Carteira _____

CONHECIMENTOS GERAIS

LÍNGUA PORTUGUESA

01. Leia as manchetes:

- Pequim se nega a receber jogo da Argentina em _____ a Messi

(<https://www.uol.com.br/esporte,10.02.2024>)

- _____ de Direitos na Rede aprofunda diálogo com ANPD sobre regulamentação de inteligência artificial

(<https://www.gov.br/anpd/pt-br/assuntos/noticias,21.03.2024>)

- _____ ao mieloma múltiplo: novo tratamento aprovado no Brasil

(<https://saude.abril.com.br,29.03.2024>)

- Em sociedades _____, universidades devem ser os principais líderes sociais em justiça restaurativa

(<https://jornal.usp.br/,11.11.2022>)

De acordo com a ortografia oficial da língua portuguesa, as lacunas devem ser preenchidas, correta e respectivamente, com:

- (A) retalhação ... Coalisção ... Contraataque ... multirraciais
- (B) retalhação ... Coalisção ... Contra-ataque ... multi-raciais
- (C) retalição ... Co-alizção ... Contrataque ... multi raciais
- (D) retalição ... Coalizção ... Contra ataque ... multi-raciais
- (E) retalição ... Coalizção ... Contra-ataque ... multirraciais

Leia o texto para responder às questões de números 02 a 07.

O descobridor das coisas

A gente vinha de mãos dadas, sem pressa de nada pela rua. Totoca vinha me ensinando a vida. E eu estava muito contente porque meu irmão mais velho estava me dando a mão e ensinando as coisas. Mas ensinando as coisas fora de casa. Porque em casa eu aprendia descobrindo sozinho e fazendo sozinho, fazia errado e fazendo errado acabava sempre tomando umas palmadas. Até bem pouco tempo ninguém me batia. Mas depois descobriram as coisas e vivem dizendo que eu era o cão, que eu era capeta, gato ruço de mau pelo. Não queria saber disso. Se não estivesse na rua eu começava a cantar. Cantar era bonito. Totoca sabia fazer outra coisa além de cantar, assobiar. Mas eu por mais que imitasse, não saía nada. Ele me animou dizendo que era assim mesmo, que eu ainda não tinha boca de soprador. Mas como eu não podia cantar por fora, fui cantando por dentro. Aquilo era es-

quisito, mas se tornava muito gostoso. E eu estava me lembrando de uma música que Mamãe cantava quando eu era bem pequenininho. Ela ficava no tanque, com um pano amarrado na cabeça para tapar o sol. Tinha um avental amarrado na barriga e ficava horas e horas, metendo a mão na água, fazendo sabão virar muita espuma. Depois torcia a roupa e ia até a corda. Prendia tudo na corda e suspendia o bambu. Ela fazia igualzinho com todas as roupas. Estava lavando a roupa da casa do Dr. Faulhaber para ajudar nas despesas da casa. Mamãe era alta, magra, mas muito bonita. Tinha uma cor bem queimada e os cabelos pretos e lisos. Quando ela deixava os cabelos sem prender, davam até na cintura. Mas bonito era quando ela cantava e eu ficava junto aprendendo.

(José Mauro de Vasconcelos. *O meu pé de laranja lima*, 1975. Adaptado)

02. Na rua com Totoca, o narrador sente-se

- (A) ansioso por achar difícil aprender os ensinamentos de seu irmão.
- (B) apreensivo com a possibilidade de o irmão dar-lhe umas palmadas.
- (C) acolhido para viver uma realidade diversa daquela vivida em casa.
- (D) triste porque o irmão tinha o mesmo comportamento que em casa.
- (E) extasiado com a vida fora de casa, onde poderia fazer o que quisesse.

03. Na passagem – Mamãe era alta, magra, mas muito bonita. Tinha uma cor bem queimada e os cabelos pretos e lisos. Quando ela deixava os cabelos sem prender, davam até na cintura. – prevalece o tipo textual

- (A) dissertativo, com o qual o narrador faz uma análise subjetiva da beleza e da dedicação de sua mãe.
- (B) descritivo, com o qual o narrador faz uma caracterização predominantemente objetiva de sua mãe.
- (C) narrativo, com o qual o narrador relata momentos de interação com sua mãe, na lavagem das roupas.
- (D) expositivo, com o qual o narrador elenca as qualidades físicas e morais de sua mãe com emotividade.
- (E) injuntivo, com o qual o narrador estabelece um diálogo mais próximo com o leitor para falar de sua mãe.

04. Assinale a alternativa em que o termo destacado pertence à mesma classe gramatical e exerce a mesma função sintática que o destacado em: “Até bem pouco tempo **ninguém** me batia.”

- (A) **Aquilo** era esquisito...
- (B) **Totoca** vinha me ensinando a vida.
- (C) Mas depois descobriram as **coisas**...
- (D) Não queria saber **disso**.
- (E) ... mas se tornava muito **gostoso**.

05. Nas passagens – Aquilo era esquisito, **mas se tornava muito gostoso**. – e – Prendia tudo na corda **e suspendia o bambu**. –, as orações destacadas expressam, correta e respectivamente, sentidos de

- (A) conclusão e adição.
- (B) adição e conclusão.
- (C) oposição e alternância.
- (D) oposição e adição.
- (E) conclusão e alternância.

06. A concordância atende à norma-padrão em:

- (A) Haviam boas lembranças, como uma música que Mamãe cantava.
- (B) Na minha opinião, cantar e assobiar eram bonito e me distraiam.
- (C) Eu e Totoca vinha de mãos dadas, sem pressa de nada pela rua.
- (D) Mas depois foi descoberto as coisas e vivem dizendo que eu era o cão...
- (E) O sabão e a água usados pela minha mãe eram para lavar a roupa alheia.

07. Considere as frases:

- Meu irmão mais velho estava de mãos dadas comigo e ensinando as coisas para mim. _____, porém, as coisas fora de casa.
- Eu fazia errado e, fazendo errado, as pessoas sempre _____ umas palmadas.
- Meu irmão sabia assobiar. Mas eu, por mais que _____, não saía nada.

De acordo com a norma-padrão, as lacunas do texto devem ser preenchidas, respectivamente, com:

- (A) Me ensinando ... davam-me ... imitasse-o
- (B) Ensinando-me ... me davam ... o imitasse
- (C) Me ensinando ... me davam ... imitasse-o
- (D) Ensinando-me ... me davam ... imitasse-o
- (E) Ensinando-me ... davam-me ... o imitasse

08. Leia a charge.

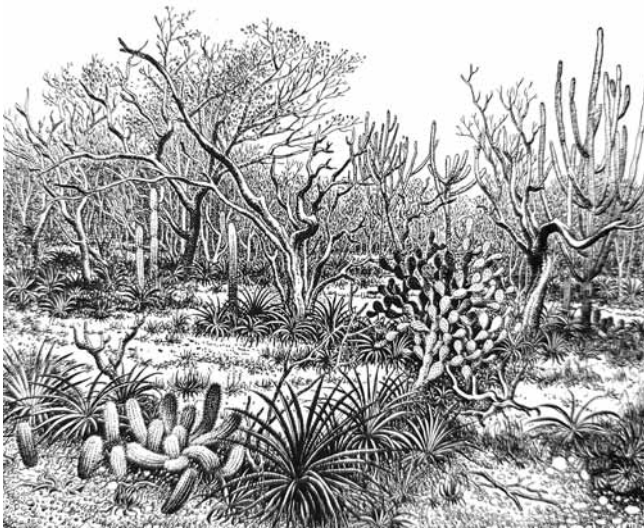


(Bob Thaves, “Frank & Ernest”. Em: <https://www.estadao.com.br/cultura/quadrinhos>. Acesso em 25.03.2024)

Do ponto de vista da coerência semântica, o termo empregado intencionalmente com duplo sentido, do qual decorre o efeito de humor na charge, é:

- (A) nosso.
- (B) clube.
- (C) folhinha.
- (D) cartão.
- (E) jardinagem.

09. Observe a imagem a seguir.



(AB'SABER, Aziz. Nacib. *Os domínios de natureza no Brasil: potencialidades paisagísticas*, 2003, pg. 33. Adaptado)

Trata-se de uma representação típica da vegetação presente no Domínio Morfoclimático

- (A) do Cerrado.
- (B) dos Mares de Morros.
- (C) das Pradarias.
- (D) das Caatingas.
- (E) das Araucárias.

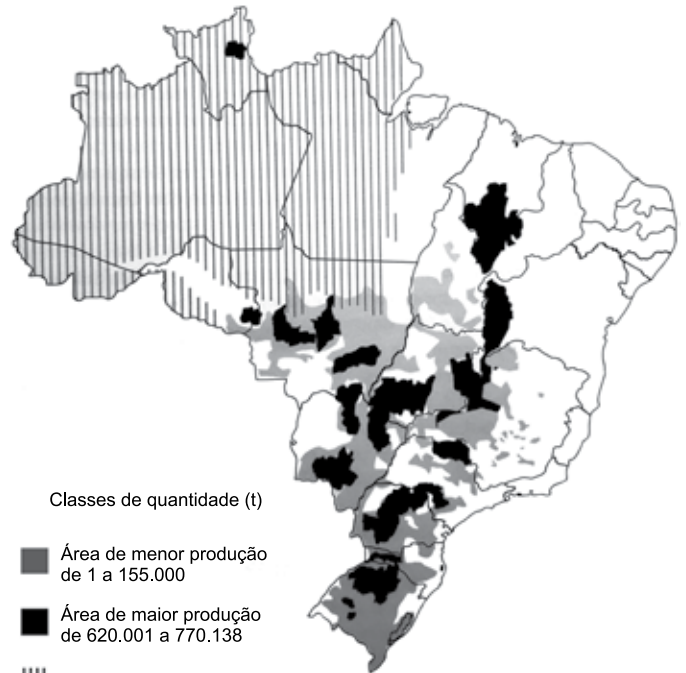
10. A escarpa da Serra do Mar que margeia o litoral do sudeste brasileiro apresenta altitudes médias de 800 m e alguns pontos superam os 2000 m. Essa elevação e a proximidade com o oceano, aliados aos ventos de quadrante sudeste na latitude do trópico, resultam em um importante processo para essa região.

(COUTINHO, Leopoldo Magno. *Biomias Brasileiros*, 2016, pg. 50-51. Adaptado)

Esse processo ocorre, marcadamente, em toda a mata Atlântica da costa leste brasileira e é definido como:

- (A) aquecimento adiabático.
- (B) efeito orográfico.
- (C) efeito Foehn.
- (D) inversão térmica.
- (E) efeito de Coriolis.

11. Observe o mapa a seguir.



(ROSS, J.L.S. *Ecogeografia do Brasil*, 2006, pg. 132. Adaptado)

O conteúdo do mapa apresenta a distribuição espacial de uma importante cultura agrícola brasileira, que também se destaca como produto de exportação. Trata-se

- (A) do arroz.
- (B) da mandioca.
- (C) do algodão.
- (D) da soja.
- (E) da cana-de-açúcar.

12. Este recurso mineral está associado a locais em que se encontram rochas sedimentares e possui grande importância econômica. O Brasil possui apenas 0,1% das reservas desse mineral. A produção brasileira desse recurso mineral é realizada em terreno sedimentar da Bacia do Paraná e se concentra nos estados de Santa Catarina e Paraná.

(TEIXEIRA, W. et. al. (org.). *Decifrando a Terra*, 2000, pg. 294-295. Adaptado)

O excerto trata do

- (A) carvão mineral.
- (B) urânio.
- (C) minério de bauxita.
- (D) potássio.
- (E) minério de ferro.

13. A compartimentação atual do relevo brasileiro tem fortes ligações genéticas com o soerguimento da plataforma sul-americana e com os processos erosivos que ocorreram principalmente no terciário e se estenderam até o quaternário, em concomitância com o soerguimento da plataforma sul-americana.

(ROSS, J. L. S. (org). *Geografia do Brasil*, 2019, pg. 52. Adaptado)

Nesse contexto, o relevo brasileiro apresenta três tipos de unidades geomorfológicas, que refletem suas gêneses e que são:

- (A) as montanhas, os vales e as planícies.
- (B) os planaltos, as depressões e a planície litorânea.
- (C) os planaltos, os vales e as serras.
- (D) as montanhas, as depressões e a planície litorânea.
- (E) os planaltos, as depressões e as planícies.

14. No contexto de culturas especializadas, estas apresentaram produção muito reduzida no território brasileiro, em função de suas exigências (temperatura, água, nutrientes) satisfeitas apenas em alguns ecossistemas.

(THÉRY, H.; MELLO, N. A. *Atlas do Brasil. Disparidades e dinâmicas do território*, 2018, pg. 164. Adaptado)

Como exemplo da especialização mencionada no excerto, pode-se citar:

- (A) o arroz e o feijão nas chapadas do centro-oeste.
- (B) a manga e a banana no sertão semiárido nordestino.
- (C) o cacau e a pimenta do reino na região amazônica.
- (D) o açaí e a borracha nas áreas litorâneas do sudeste.
- (E) a soja e a erva mate nos campos da região sul.

15. A primeira remessa de algodão brasileiro para o exterior, data, ao que parece, de 1760, e provém do Maranhão, que nesse ano exporta 651 arrobas. De Pernambuco exporta-se a partir de 1778, sendo em quantidade insignificante até 1781. A Bahia e o Rio de Janeiro seguirão o passo. Mas é no Maranhão que o progresso da cultura algodoeira é mais interessante, porque ela parte aí do nada, de uma região pobre e inexpressiva no conjunto da colônia. O algodão dar-lhe-á vida e transformá-la-á, em poucos decênios, numa das mais ricas e destacadas capitanias.

(Caio Prado Júnior, *Formação do Brasil Contemporâneo*, p. 143-144)

Contribuiu, decisivamente, para o desenvolvimento do algodão no Maranhão:

- (A) a busca por escravizados africanos que já haviam trabalhado com esse produto na África ou em outra capitania da América portuguesa.
- (B) a Companhia Geral do Comércio do Grão-Pará e do Maranhão, que forneceu créditos, escravos e ferramentas aos lavradores.
- (C) o exponencial aumento das exportações do produto para as principais cidades industriais dos Estados Unidos e do Caribe.
- (D) o apoio do Conselho Ultramarino pela oferta de conhecimento técnico, ferramentas sofisticadas e aportes de capital.
- (E) o desenvolvimento, em algumas comarcas mineiras, da produção artesanal de tecido grosso, voltada para a vestimenta dos escravizados.

16. Em 28 de janeiro de 1808, D. João assinou a primeira medida régia na nova sede do Império lusitano: a carta de abertura dos portos brasileiros às nações amigas. A partir de então ficava permitida a importação “de todos e quaisquer gêneros, fazendas e mercadorias transportadas ou em navios estrangeiros das potências que se conservavam em paz e harmonia com a minha Real Coroa”, ou em navios da metrópole.

(Lília Moritz Schwartz e Heloisa Murgel Starling, *Brasil: Uma biografia*, p. 173. Adaptado)

O primeiro ato de D. João no Brasil resultou

- (A) no desmoronamento do exclusivo comercial da metrópole, estabelecido desde o princípio da colonização.
- (B) na forte diminuição das importações e exportações do Brasil, em razão do aumento das tarifas alfandegárias.
- (C) no importante desenvolvimento das manufaturas brasileiras, que serviam para a troca de escravizados na África.
- (D) no aumento do comércio exterior com as nações formadas a partir da fragmentação da América francesa.
- (E) na organização política da elite colonial, grupo bastante prejudicado com a nova ordem estabelecida.

17. Uma peculiaridade da Carta de 1824 foi incluir um artigo reproduzindo quase palavra por palavra a Declaração dos Direitos do Homem emitida na França em 1789. Comparado ao original havia, no entanto, algumas omissões bastante significativas e curiosas. Não foi incluído na Carta outorgada o artigo que, na versão original francesa, dizia: “O princípio de toda soberania reside essencialmente na nação. Nenhum corpo nem indivíduo podem exercer autoridade que não emane expressamente dela”. Também faltava o artigo VI: “A lei é expressão da vontade geral”. Finalmente, o artigo II: “O objetivo de toda associação política é a preservação dos direitos naturais e inalienáveis do homem. Estes direitos são a liberdade, a propriedade, a segurança e a resistência perante a opressão” foi reproduzido omitindo-se as seis últimas palavras.

(Emília Viotti da Costa, *Da monarquia à república: momentos decisivos*, p. 141-142. Adaptado)

Para Emília Viotti da Costa, tais omissões podem revelar

- (A) o descuido dos legisladores com a estruturação legal da nação.
- (B) a preponderância do Poder Legislativo sobre o Executivo.
- (C) a forte influência das estruturas políticas latino-americanas.
- (D) o desprestígio de Dom Pedro I junto ao povo brasileiro.
- (E) as tendências antidemocráticas e oligárquicas das elites brasileiras.

18. Uma comissão de cinco pessoas foi encarregada de redigir um projeto de Constituição, submetido, depois, à profunda revisão por parte de Rui Barbosa. A seguir, encaminhou-se o projeto à apreciação da Assembleia Constituinte, que, após muitas discussões e algumas emendas, promulgou o texto a 24 de fevereiro de 1891.

(Boris Fausto, *História do Brasil*, p. 249)

Segundo Boris Fausto, a Constituição de 1891,

- (A) baseada no direito consuetudinário, instituiu uma forte centralização político-administrativa, impedindo que os estados tivessem constituições, além disso os orçamentos das unidades federativas passavam pelo crivo federal.
- (B) fundamentada no direito natural, estabeleceu uma separação tênue entre público e privado e determinou que a educação pública fosse dirigida por ordens religiosas.
- (C) influenciada pelas principais constituições europeias, reforçou o Poder Judiciário em detrimento dos outros poderes, porque todos os níveis do processo eleitoral estavam ligados ao Tribunal Superior Eleitoral.
- (D) inspirada no modelo norte-americano, consagrou a República federativa liberal e os estados ficaram com a faculdade de organizar uma justiça própria, estabelecendo o sistema do voto direto e universal.
- (E) calcada no direito inglês, limitou as prerrogativas estaduais, na medida em que toda política tributária estaria centralizada no governo federal, e, com isso, nacionalizou todos os recursos minerais.

19. Um assunto dominava a atenção de Getúlio: a política trabalhista. Foi nessa área que ele mostrou quem era e a que viera. Dividiu sua política em duas metades. Numa, criou as leis de proteção ao trabalhador.

(Lilia M. Schwarcz e Heloisa M. Starling, *Brasil: uma biografia*, p. 362)

Na outra metade, Getúlio

- (A) reprimiu qualquer esforço de organização dos trabalhadores fora do controle do Estado e enquadrou os sindicatos como órgãos de colaboração com o Estado.
- (B) privilegiou os trabalhadores do campo, que foram contemplados com aposentadoria especial, além da obtenção de pequenas propriedades derivadas de reforma agrária.
- (C) estabeleceu acordos com os sindicatos mais progressistas e limitou a ação das entidades de trabalhadores que defendiam a presença do Estado na economia.
- (D) apoiou projetos que ampliavam a ação política dos sindicatos, ao mesmo tempo em que garantiu o amplo direito de greve de todos os trabalhadores.
- (E) reforçou a inaptidão dos trabalhadores no trato da questão político-partidária, ainda que tenha estendido o voto para todos os brasileiros, mesmo os analfabetos.

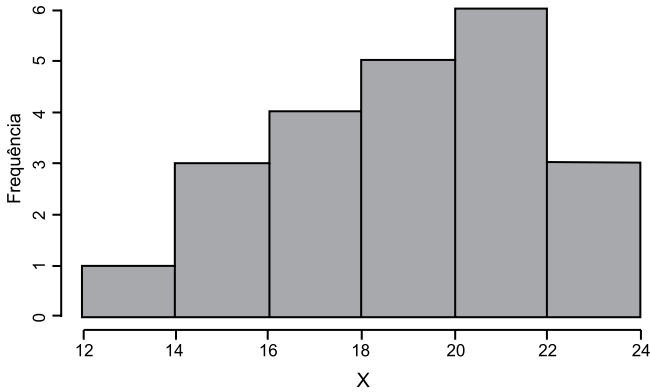
20. No início da década de 1950, o governo promoveu várias medidas destinadas a incentivar o desenvolvimento econômico, com ênfase na industrialização. Foram feitos investimentos públicos no sistema de transportes e de energia, com a abertura de um crédito externo de 500 milhões de dólares.

(Boris Fausto, *História do Brasil*, p. 409. Adaptado)

No contexto apresentado, também fez parte do esforço governamental

- (A) a criação do Banco Central do Brasil, instituição responsável pela taxa de câmbio e taxa de juros, que anteriormente eram determinadas pela dinâmica do mercado.
- (B) a autorização para que empresas estrangeiras, ligadas à educação e aos meios de comunicação, estabelecessem filiais nas capitais dos estados.
- (C) a implementação do Plano Salte, com maciços investimentos em saúde, alimentação, transporte e energia, a partir do capital nacional, público e privado.
- (D) a fundação do Banco Nacional de Desenvolvimento Econômico (BNDE), diretamente orientado para o propósito de acelerar o processo de diversificação industrial.
- (E) a permissão para a formação de empresas privadas de exploração de petróleo, desde que tais instituições contassem com a parcela mínima de 30% de capitais estatais.

21. Analise o histograma a seguir.



(Arquivo pessoal; imagem usada com autorização)

Com base nas informações obtidas pelo histograma, é correto afirmar que a

- (A) média de x é maior do que a mediana de x .
 - (B) moda de x é maior do que a mediana de x .
 - (C) média de x é maior do que a moda de x .
 - (D) média de x é igual à moda de x .
 - (E) moda de x é igual à mediana de x .
22. Considere uma amostra $x = \{x_1, x_2, x_3, x_4\}$, em que os três primeiros desvios de cada observação x_i para a média (\bar{x}) são dados por:

$$(x_1 - \bar{x}) = -13, (x_2 - \bar{x}) = 32 \text{ e } (x_3 - \bar{x}) = -4.$$

Sabendo que a média $\bar{x} = 34$, é correto afirmar que o valor de x_4 é igual a:

- (A) -15
 - (B) -7
 - (C) -49
 - (D) 15
 - (E) 19
23. Seja $X = \{-1, 0, 0, 0, 0, 0, 1, 1, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 6\}$. Considere o estimador da curtose (k) dado por:

$$k = \frac{Q_3 - Q_1}{2(D_9 - D_1)}$$

em que Q_1 e Q_3 são o primeiro e o terceiro quartil, respectivamente, e D_1 e D_9 são o primeiro e nono decil, respectivamente. Dessa forma, é correto afirmar que, para o conjunto de dados X , a curtose pode ser classificada, em relação à distribuição normal, como

- (A) microcúrtica.
- (B) leptocúrtica.
- (C) platicúrtica.
- (D) mesocúrtica.
- (E) megacúrtica.

24. Considere as idades em anos completos dos grupos A e B, apresentadas a seguir.

Grupo A	20	24	24	25	32
Grupo B	20	22	22	32	

Com base nesses dados, é correto afirmar que

- (A) a amplitude total da idade do grupo A é menor do que a amplitude total da idade do grupo B.
- (B) a moda da idade do grupo A é menor do que a moda da idade do grupo B.
- (C) a mediana da idade do grupo A é menor do que a mediana da idade do grupo B.
- (D) o desvio-padrão da idade do grupo A é menor do que o desvio-padrão da idade do grupo B.
- (E) a média da idade do grupo A é menor do que a média da idade do grupo B.

25. Um experimento foi conduzido em delineamento inteiramente casualizado com 3 tratamentos e 4 repetições, totalizando 12 unidades amostrais. Os resultados da Análise de Variância são observados na tabela a seguir.

Causas de Variação	Graus de Liberdade	Soma de Quadrados	Quadrado Médio	F calculado
Tratamento	2	1 000	500	B
Resíduo	9	180	A	–

Os valores de A e B são, respectivamente,

- (A) 20 e 0,04.
- (B) 20 e 25.
- (C) 0,18 e 0,22.
- (D) 90 e 5,5.
- (E) 90 e 0,18.

R A S C U N H O

26. Um experimento foi conduzido em esquema fatorial duplo, sendo que o primeiro fator tinha 3 níveis e o segundo fator tinha 2 níveis. Na Análise de Variância, os valores da estatística F calculada obtidos foram: 0,21 para o primeiro fator, 0,92 para o segundo fator e 2,80 para o efeito da interação entre os dois fatores. Sabendo que o experimento tinha 18 graus de liberdade para o resíduo, considere os valores de F tabelados para o nível de significância de 5% a seguir, em que V1 são os graus de liberdade para o numerador e V2 são os graus de liberdade do denominador.

V2	V1					
	1	2	3	4	5	6
18	4,414	3,555	3,160	2,928	2,773	2,661

Considerando o nível de significância de 5%, assinale a alternativa correta.

- (A) Não se rejeita a hipótese nula para o primeiro fator nem para o segundo fator e a interação.
- (B) Rejeita-se a hipótese nula para o primeiro e o segundo fator, mas não se rejeita para a interação.
- (C) Rejeita-se a hipótese nula para o segundo fator, mas não se rejeita para o primeiro fator e a interação.
- (D) Rejeita-se a hipótese nula para a interação, mas não se rejeita para o primeiro e o segundo fator.
- (E) Rejeita-se a hipótese nula para o primeiro fator, mas não se rejeita para o segundo fator e a interação.
27. Um experimento agrícola foi conduzido no esquema fatorial duplo em delineamento em blocos casualizados, sendo os níveis do primeiro fator as doses 0, 10, 20 e 30 kg/ha e do segundo fator os adubos X, Y e Z. Considere que foram utilizados 5 blocos e uma repetição por bloco. O grau de liberdade para o resíduo é igual a:
- (A) 40
- (B) 48
- (C) 44
- (D) 50
- (E) 59
28. Seja um experimento unifatorial considerando um modelo de efeito aleatório com 3 tratamentos e 4 repetições. Sabendo que o quadrado médio dos tratamentos é igual a 360 e que o quadrado médio dos resíduos é igual a 60, então a estimativa do componente de variância do tratamento é igual a:
- (A) 90
- (B) 100
- (C) 60
- (D) 360
- (E) 75

29. Sejam X_1, \dots, X_n uma amostra aleatória da variável aleatória $X \sim N(\mu, 1)$. Sabendo que $\hat{\mu} = \bar{X} = \left(\sum_{i=1}^n x_i \right) / n$ é o estimador de máxima verossimilhança de μ , então, pelo princípio da invariância, o estimador de máxima verossimilhança de $g(\mu) = e^{-\mu}$ será:

- (A) $g(\hat{\mu}) = \bar{X}$
- (B) $g(\hat{\mu}) = e^{-\bar{X}}$
- (C) $g(\hat{\mu}) = -\bar{X}$
- (D) $g(\hat{\mu}) = -\bar{X}e^{-\bar{X}}$
- (E) $g(\hat{\mu}) = -e^{-\bar{X}}$

30. Considere uma amostra aleatória X_1, \dots, X_n de uma variável aleatória X com função densidade de probabilidade dada por:

$$f(x; \theta) = \frac{1}{2} (1 + \theta x),$$

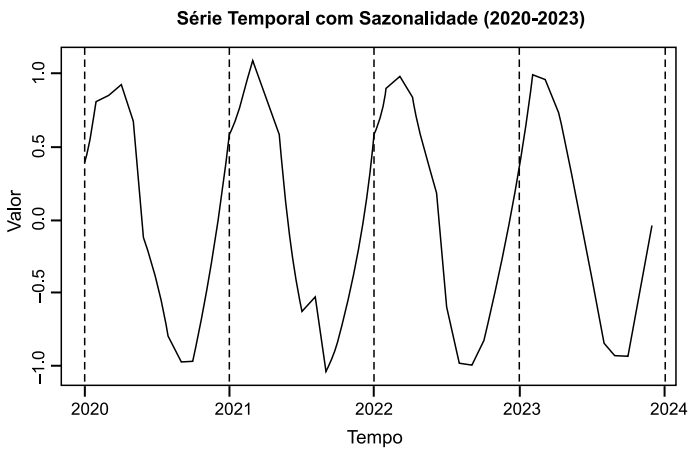
com $-1 < x < 1$ e $-1 < \theta < 1$. Sendo θ o parâmetro da função, nos procedimentos para a obtenção do estimador de máxima verossimilhança de θ , considerando $\ln()$ a função logaritmo natural, a função log-verossimilhança é dada por:

- (A) $l(\theta; \mathbf{x}) = \sum_{i=1}^n \frac{x_i}{1 + \theta x_i}$
- (B) $l(\theta; \mathbf{x}) = \ln(1) - \ln(2) + \ln(1 + (\theta x))$
- (C) $l(\theta; \mathbf{x}) = \sum_{i=1}^n \ln(1 + (\theta x_i))$
- (D) $l(\theta; \mathbf{x}) = n \ln\left(\frac{1}{2}\right) + \sum_{i=1}^n \ln(1 + (\theta x_i))$
- (E) $l(\theta; \mathbf{x}) = \ln\left(\frac{n}{2}\right) + \sum_{i=1}^n \ln(n + (\theta x_i))$

31. Sejam X_1, \dots, X_n uma amostra aleatória da variável aleatória X com distribuição geométrica de parâmetro θ e função de probabilidade $f(x|\theta) = \theta(1 - \theta)^{x-1}$, $x = 1, 2, 3, \dots$, e $0 < \theta < 1$. Seja $\hat{\theta}_{EMV}$ o estimador de máxima verossimilhança de θ e seja $\hat{\theta}_{MM}$ o estimador pelo método dos momentos de θ , é correto afirmar que

- (A) $\hat{\theta}_{EMV} > \hat{\theta}_{MM}$
- (B) não é possível encontrar um estimador para θ pelo método da máxima verossimilhança.
- (C) $\hat{\theta}_{EMV} < \hat{\theta}_{MM}$
- (D) não é possível encontrar um estimador para θ pelo método dos momentos.
- (E) $\hat{\theta}_{EMV} = \hat{\theta}_{MM}$

32. Considere a componente de sazonalidade de uma série temporal sem tendência, apresentada na imagem a seguir.

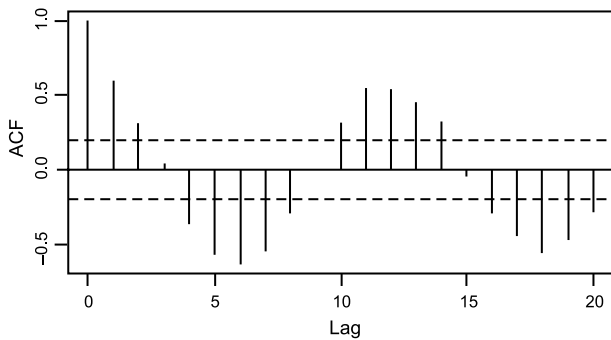


(Arquivo pessoal; imagem usada com autorização)

Seja a constante π (pi) e seja $x = \text{Tempo}$ (em meses), a melhor função para modelar a componente de sazonalidade, entre as opções a seguir, é:

- (A) $f(x) = \text{sen}(\pi x/6)$
- (B) $f(x) = \text{cos}(\pi x/4)$
- (C) $f(x) = \text{cos}(\pi x/6)$
- (D) $f(x) = \text{sen}(\pi x/12)$
- (E) $f(x) = \text{sen}(\pi x/4)$

33. Analise o gráfico da função de autocorrelação (ACF) a seguir.



(Arquivo pessoal; imagem usada com autorização)

É correto afirmar que a série temporal é uma série

- (A) com tendência e sem sazonalidade.
- (B) estacionária.
- (C) aleatória (ruído branco).
- (D) sem tendência e com sazonalidade.
- (E) sem tendência e sem sazonalidade.

34. Sobre os métodos de estimação e propriedades dos estimadores, é correto afirmar:

- (A) Quando uma amostra de tamanho n for grande e $\hat{\theta}$ for o estimador de máxima verossimilhança para o parâmetro θ , $\hat{\theta}$ tem uma distribuição normal aproximada, porém não se pode garantir que é um estimador não tendencioso.
- (B) Quando uma amostra de tamanho n for grande e $\hat{\theta}$ for o estimador de máxima verossimilhança para o parâmetro θ , pode-se garantir que $\hat{\theta}$ é um estimador não tendencioso e a variância de $\hat{\theta}$ é aproximadamente tão pequena quanto a variância que poderia ser obtida para qualquer outro estimador.
- (C) Considerando uma amostra aleatória proveniente de uma distribuição normal, com parâmetros μ e σ o estimador obtido pelo método dos momentos para σ é não tendencioso.
- (D) Quando uma amostra de tamanho n for grande e $\hat{\theta}$ for o estimador obtido pelo método dos momentos para o parâmetro θ , pode-se garantir que $\hat{\theta}$ é um estimador não tendencioso de menor variância entre os estimadores.
- (E) Considerando uma amostra aleatória proveniente de uma distribuição normal, com parâmetros μ e σ o estimador de máxima verossimilhança para σ é não tendencioso.

35. Considere uma pesquisa realizada em um restaurante para avaliar a proporção de clientes satisfeitos com o atendimento. Foram avaliados $n = 200$ clientes dos quais 130 afirmaram que estão satisfeitos com o restaurante. Dado: $\phi(1,645) = 0,95$ e $\phi(1,96) = 0,975$, sendo ϕ a função de distribuição acumulada normal padrão e os valores aproximados $\sqrt{10} = 3,16$; $\sqrt{11} = 3,32$; $\sqrt{12} = 3,46$ e $\sqrt{13} = 3,61$.

O intervalo de confiança 95% para a proporção de clientes satisfeitos é dado por:

- (A) [0,53; 0,77]
- (B) [0,51; 0,79]
- (C) [0,55; 0,75]
- (D) [0,61; 0,69]
- (E) [0,58; 0,72]

36. Um pesquisador interessado em estimar a idade média dos estudantes que frequentam um curso gratuito de inglês em uma pequena cidade coletou informações de 9 alunos, obtendo as estimativas para a média $\bar{x} = 55$ e para variância $s^2 = 9$. Com base nessas informações, ele obteve o intervalo com 95% de confiança para a idade média dos estudantes. $F(1,860) = 0,95$; $F(2,306) = 0,975$; $\phi(1,645) = 0,95$ e $\phi(1,96) = 0,975$; sendo F a função de distribuição acumulada t de Student com 8 graus de liberdade e ϕ a função de distribuição acumulada normal padrão.

O intervalo de confiança para a média das idades é:

- (A) [53,35; 56,65]
 - (B) [53,04; 56,96]
 - (C) [52,69; 57,31]
 - (D) [52,39; 57,61]
 - (E) [53,14; 56,86]
37. Com o objetivo de estimar a idade média das crianças de um bairro, foram coletadas as idades de 81 crianças, obtendo-se uma média de 6 anos e desvio-padrão de 3 anos. Sejam os valores da função acumulada da distribuição normal padrão $\phi(1,645) = 0,95$ e $\phi(1,96) = 0,975$, o intervalo de confiança de 95% obtido para a idade média das crianças será:

- (A) [5,10; 6,90]
- (B) [5,45; 6,55]
- (C) [4,80; 7,20]
- (D) [5,35; 6,65]
- (E) [5,00; 7,00]

38. Um empreendedor que recentemente investiu em uma franquia de alimentação gostaria de saber qual a proporção de clientes que está satisfeita com o atendimento, para decidir sobre a manutenção do negócio. Considere o caso de uma amostragem aleatória simples e o nível de confiança de 95%. Dado: $\phi(1,645) = 0,95$ e $\phi(1,96) = 0,975$, sendo ϕ a função de distribuição acumulada normal padrão.

O tamanho da amostra que o empreendedor deve utilizar na pesquisa para um erro máximo de 2% é igual a:

- (A) 2401
- (B) 1028
- (C) 271
- (D) 1692
- (E) 385

39. Considere um estudo para investigar o número médio de crianças (menores de 10 anos) por residência em uma cidade com $N = 385$ residências, em que se sabe, de estudos anteriores, que a variância do número de crianças por residência é 1. Considere o caso de uma amostragem aleatória simples. Dado: $\Phi(1,645) = 0,95$ e $\Phi(1,96) = 0,975$, sendo Φ a função de distribuição acumulada normal padrão.

Considerando o nível de confiança de 95% e a margem de erro máxima de 0,2, o tamanho da amostra necessário é igual a:

- (A) 77
- (B) 96
- (C) 68
- (D) 58
- (E) 87

40. Uma decisão importante na utilização da amostragem estratificada é a forma pela qual o tamanho total da amostra será alocado ou distribuído nos estratos. Uma das formas de realizar essa alocação é utilizando a alocação ótima.

Sobre a alocação ótima, pode-se afirmar que

- (A) considera na alocação somente o tamanho, a média e a variabilidade de cada estrato.
- (B) considera na alocação o tamanho, a variabilidade e o custo de coleta de cada estrato.
- (C) sua utilização é recomendada apenas nos casos em que o número de elementos que compõem os estratos (tamanho dos estratos) difere muito entre si.
- (D) considera na alocação apenas o tamanho e a média de cada estrato.
- (E) a amostra é distribuída proporcionalmente ao tamanho dos estratos, sem considerar outras variáveis.

41. Considere uma população de 10 elementos, da qual se deseja obter uma amostra com 4 elementos. Considere uma amostragem aleatória simples, sem reposição.
- A probabilidade de se obter uma amostra particular é:
- (A) $1/210$
 - (B) $1/10$
 - (C) $1/5040$
 - (D) $1/120$
 - (E) $1/37800$
42. Considere a teoria de decisão Bayesiana. Sobre uma priori não-informativa, é possível afirmar:
- (A) Uma priori não-informativa deve ser utilizada nas situações em que há conhecimentos que permitam especificar uma distribuição priori informativa.
 - (B) Utilizando uma priori não informativa, sempre obtemos uma família conjugada.
 - (C) Sempre que empregada uma priori não-informativa, as interpretações utilizando análise Bayesiana serão equivalentes à análise clássica.
 - (D) A única forma de utilizar uma priori não-informativa é assumindo para priori uma distribuição uniforme.
 - (E) Mesmo utilizando uma priori não informativa, a interpretação obtida para o intervalo de credibilidade é distinta da obtida utilizando o intervalo de confiança convencional.
43. Considere a teoria Bayesiana e as famílias conjugadas de distribuição. Seja F uma família de distribuições para a verossimilhança $p(x|\theta)$ e P uma família de distribuição para a priori $p(\theta)$. Dizemos que F e P são famílias conjugadas de distribuições se:
- (A) a distribuição posterior $p(\theta|X)$ também for um membro de F .
 - (B) as distribuições a priori, a verossimilhança e a distribuição posterior forem membros de F .
 - (C) a distribuição posterior $p(\theta|X)$ também for um membro de P .
 - (D) a família de distribuições para a verossimilhança for um membro de P .
 - (E) a distribuição posterior for membro de P ou de F .

44. Em uma pesquisa, realizada em uma determinada cidade, foram coletadas informações das residências por bairro, sendo obtido um conjunto de dados, nomeado como `dado`, no seguinte formato:

<code>resid</code>	<code>bairro</code>	<code>npf</code>	<code>rf</code>	<code>nc</code>
1	A	3	7	1
2	A	4	3	1
3	A	7	2	3
1	B	5	7	3
2	B	4	2	2
...				

Em que `resid` representa o número da residência por bairro, `bairro` representa o bairro da residência, `npf` representa o número de pessoas que residem na residência, `rf` representa a renda familiar (em quantidade de salários mínimos) e `nc` representa o número de pessoas com menos de 10 anos que residem na residência.

Considere que se deseja obter os valores médios das variáveis por bairro. Utilizando o pacote `dplyr`, os comandos que retornam um novo conjunto com os valores médios por bairro para as variáveis número de pessoas que residem na residência, renda familiar e número de pessoas com menos de 10 anos que residem na residência é apresentado em:

- (A) `summarise(group_by(dado, bairro.agrup=bairro), npf.medio= (npf), rf.medio= (rf), nc.medio= (nc))`
 - (B) `tapply(dado, bairro,mean)`
 - (C) `apply(dado,1,mean)`
 - (D) `summarise(group_by(dado, bairro.agrup=bairro), npfm=mean(npf), rfm=mean(rf), ncm=mean(nc))`
 - (E) `ddply(dado, dado.agrup=bairro, mean)`
45. Em um levantamento para analisar as escolas de um determinado município, foram obtidos dados do número de alunos matriculados por turma em cada escola, sendo criado um arquivo `.RData`, nomeado como `dado`, com as colunas `Escola`, `Turma` e `Matriculados`. As três primeiras linhas do objeto `dado` são apresentadas a seguir.

<code>Escola</code>	<code>Turma</code>	<code>Matriculados</code>
1	1A	30
1	1B	26
2	2A	20
...		

O comando que pode ser empregado para obter a variância por escola para o número de alunos matriculados por turma é:

- (A) `apply(Matriculados,Escola, var)`
 - (B) `var(Matriculados, by=Escola)`
 - (C) `var(Matriculados)`
 - (D) `with(dado, tapply(Matriculados,Escola,var))`
 - (E) `tapply(dado,Escola,var)`
46. Considere que um professor de estatística deseja avaliar se a nota obtida pelos alunos pode ser descrita em função do tempo de estudo deles. Para isso, decidiu realizar o ajuste de um modelo de regressão linear e organizou os dados das notas dos alunos e do tempo de estudo em dois objetos no ambiente R, nomeados como “nota” e “tempo”, ambos na mesma ordem de entrada. A sequência de comandos que realiza o ajuste de um modelo de regressão linear e apresenta o intervalo de confiança (95%) para os coeficientes de regressão é:
- (A) `regressaolinear=lm(nota~tempo) ; intervalo(regressaolinear)`
 - (B) `regressaolinear=lm(nota~tempo) ; confint(regressaolinear)`
 - (C) `regressaolinear= reg(nota~tempo) ; summary(regressaolinear)`
 - (D) `reg(nota~tempo) ; summary(reg)`
 - (E) `lm(nota~tempo) ; intervalo(regressaolinear)`

47. Considere X uma variável aleatória discreta, em que $X \sim \text{Binomial}(n, p)$.

Sobre essa distribuição, é correto afirmar que

- (A) a função geradora de momentos é dada por:
 $M_X(t) = (p e^t + 1 - p)^n$
- (B) o terceiro momento dessa variável aleatória não existe, então não é possível determinar o coeficiente de curtose.
- (C) a função geradora de momentos é dada por:
 $M_X(t) = (p e^t + 1 - n)^p$
- (D) o segundo momento dessa variável aleatória é dado por: $p(n p - p + 1)$
- (E) o segundo momento dessa variável aleatória é dado por: $n p(n p + p - 1)$

Notação: $f(y) = e^y$: é a função exponencial cuja base é o número de Euler.

48. Arrecadações de objetos novos e usados com a população são ações que o exército costuma realizar para ajudar os necessitados depois de alguma tragédia. Contudo, objetos usados nem sempre chegam em bom estado e, por isso, necessitam de algum conserto antes de serem distribuídos. Um levantamento feito em uma dessas ações, que arrecadou 900 itens, gerou a seguinte relação de materiais e a situação desses itens:

Objeto	Situação	
	Bom	Conserto
Eletrodoméstico	112	65
Roupa	340	209
Calçado	149	25

A probabilidade de pegar, ao acaso, uma roupa ou um item que precise de conserto é igual a:

- (A) 0,71
- (B) 0,94
- (C) 0,23
- (D) 0,69
- (E) 0,90
49. Considere que a função de densidade da variável aleatória contínua uniforme, X , no intervalo $[13, 25]$ modela razoavelmente um fenômeno de interesse. Dessa forma, o valor esperado e a variância dessa variável aleatória serão respectivamente:
- (A) 354; 38
- (B) 38; 354
- (C) 38; 19
- (D) 19; 14
- (E) 19; 12

50. Considere a realização dos seguintes experimentos:

- Experimento I: anota-se a face superior do lançamento de três moedas.
- Experimento II: anota-se a face superior do lançamento de dois dados.
- Experimento III: anota-se a face superior do lançamento de duas moedas e três dados.

Considere que todos os dados utilizados nesses experimentos têm seis faces. O número de elementos do espaço amostral de cada experimento é respectivamente:

- (A) 6, 12 e 864
- (B) 4, 18 e 432
- (C) 8, 36 e 72
- (D) 8, 36 e 864
- (E) 6, 12 e 72

51. Um fabricante de pilhas AAA afirma que a vida útil delas tem distribuição aproximadamente normal com média de 0,17 ano e desvio-padrão de 0,3 ano. Uma amostra aleatória de 37 dessas pilhas apresentou um desvio-padrão de 0,4 ano. Considerando a hipótese alternativa de o desvio-padrão ser maior que 0,3 ano, o resultado do valor da estatística calculada e a conclusão desse teste de hipótese ao nível de significância de 0,05 serão respectivamente:

- (A) 69,4; o desvio-padrão é maior que 0,3 ano.
- (B) 64,0; o desvio-padrão é maior que 0,3 ano.
- (C) 64,0; o desvio-padrão é igual a 0,3 ano.
- (D) 20,8; o desvio-padrão é diferente de 0,3 ano.
- (E) 69,4; o desvio-padrão é igual a 0,3 ano.

Dados: $P(Y > 52,1923) = 0,05$; $P(Q < 0,4935) = 0,95$; $P(D < 0,6635) = 0,95$; $P(X < 50,9985) = 0,95$ e $P(Z < 1,96) = 0,975$, em que Y é uma variável aleatória que segue uma distribuição qui-quadrado com 37 graus de liberdade; Q é uma variável aleatória que segue uma distribuição normal com média 0 e variância 0,09; D é uma variável aleatória que segue uma distribuição normal com média 0,17 e variância 0,09; X é uma variável aleatória que segue uma distribuição qui-quadrado com 36 graus de liberdade; e Z é uma variável aleatória que segue uma distribuição normal padrão.

52. Em uma fábrica de ar-condicionado, nove máquinas do mesmo modelo foram selecionadas aleatoriamente a fim de determinar o efeito da limpeza do filtro de ar no gasto de energia elétrica. Todas as máquinas novas foram instaladas em um mesmo lado de um prédio, e durante dois meses (numa mesma estação do ano) foram ligadas durante o mesmo período por dia, numa mesma temperatura. O gasto médio diário em kW da última semana apresentou um valor de 156. Terminado esse mês, foi realizada a limpeza do filtro de ar de todas as máquinas e, durante mais uma semana, elas foram ligadas nas mesmas condições. No final do último dia, calculou-se o consumo médio, resultando no valor de 140 kW. O desvio-padrão da diferença entre o consumo antes da limpeza menos o consumo depois da limpeza foi de 15 kW. Ao nível de 5%, de significância, foram testadas as hipóteses: de o consumo médio antes ser igual ao consumo médio depois da limpeza das máquinas contra o consumo médio antes ser maior que o consumo médio depois da limpeza. O valor calculado da estatística de teste e sua conclusão para esse teste de hipóteses são, respectivamente:

- (A) 8,4. Então, aceita-se H_0 , conclui-se que o consumo médio de energia antes da limpeza do filtro de ar das máquinas é diferente do consumo médio de energia depois da limpeza.
- (B) 2,8. Então, rejeita-se H_0 , conclui-se que o consumo médio de energia antes da limpeza do filtro de ar das máquinas é menor que o consumo médio de energia depois da limpeza.
- (C) 3,2. Então, rejeita-se H_0 , conclui-se que o consumo médio de energia antes da limpeza do filtro de ar das máquinas é maior que o consumo médio de energia depois da limpeza.
- (D) -3,2. Então, rejeita-se H_0 , conclui-se que o consumo médio de energia antes da limpeza do filtro de ar das máquinas é diferente do consumo médio de energia depois da limpeza.
- (E) 2,8. Então, aceita-se H_0 , conclui-se que o consumo médio de energia não mudou depois da limpeza do filtro de ar das máquinas.

Dados: $P(X < 16,92) = 0,95$; $P(Y < 2,73) = 0,05$; $P(T < -1,86) = 0,05$; $P(W < -1,83) = 0,05$; $P(Q < -1,74) = 0,05$; $P(Z < 1,64) = 0,95$; e $P(Z < 1,96) = 0,975$; em que X é uma variável aleatória que segue uma distribuição qui-quadrado com 9 graus de liberdade; Y é uma variável aleatória que segue uma distribuição qui-quadrado com 8 graus de liberdade; T é uma variável aleatória que segue uma distribuição t -Student com 8 graus de liberdade; W é uma variável aleatória que segue uma distribuição t -Student com 9 graus de liberdade; Q é uma variável aleatória que segue uma distribuição t -Student com 17 graus de liberdade; e Z é uma variável aleatória que segue uma distribuição normal padrão.

53. A abordagem do teste de hipóteses para a inferência estatística é muito próxima à abordagem do intervalo de confiança. Essa equivalência se estende às diferenças entre duas médias, variâncias, razão de variâncias e assim por diante. Para o caso de uma única média populacional μ com variância σ^2 conhecida, considerando um nível de significância α e uma amostra aleatória de tamanho n dessa população, é correto afirmar:

- (A) Para um teste de hipóteses $H_0: \mu = \mu_0$ contra $H_1: \mu < \mu_0$ é equivalente a calcular o intervalo de confiança $100(1 - \alpha)\%$ para μ : $[\bar{x} - z_{(\alpha/2)} \frac{\sigma}{\sqrt{n}}; +\infty]$
- (B) Para um teste de $H_0: \mu = \mu_0$ contra $H_1: \mu < \mu_0$ é equivalente a calcular o intervalo de confiança $100(1 - \alpha)\%$ para μ : $[-\infty; \bar{x} + z_{(\alpha/2)} \frac{\sigma}{\sqrt{n}}]$
- (C) Para um teste de $H_0: \mu = \mu_0$ contra $H_1: \mu \neq \mu_0$ é equivalente a calcular o intervalo de confiança $100(1 - \alpha)\%$ para μ : $[-\infty; \bar{x} + z_{(\alpha/2)} \frac{\sigma}{\sqrt{n}}]$
- (D) Para um teste de $H_0: \mu = \mu_0$ contra $H_1: \mu < \mu_0$ é equivalente a calcular o intervalo de confiança $100(1 - \alpha)\%$ para μ : $[\bar{x} \mp z_{(\alpha/2)} \frac{\sigma}{\sqrt{n}}]$
- (E) Para um teste de $H_0: \mu = \mu_0$ contra $H_1: \mu > \mu_0$ é equivalente a calcular o intervalo de confiança $100(1 - \alpha)\%$ para μ : $[\bar{x} \mp z_{(\alpha/2)} \frac{\sigma}{\sqrt{n}}]$

54. Considere a população de crianças, do sexo masculino, de faixa etária de 6 a 7 anos de uma determinada região. É de desejo realizar o seguinte teste de hipóteses para a proporção (p) de crianças com o índice de massa corpórea (IMC) maior que 30 dessa população (que é normalmente distribuída para essa variável): $p = 0,6$ contra $p > 0,6$. Fixando um nível de significância de 5%; considerando p_c o um ponto crítico para a tomada de decisão e o estimador \hat{p} da verdadeira proporção de crianças com o índice de massa corpórea (IMC) maior que 30, p , é correto afirmar que:

- (A) o nível de significância é obtido calculando a probabilidade de $\hat{p} < p_c$ dado que $p > 0,6$, ou seja, é a probabilidade de ocorrer o erro tipo I.
- (B) o poder do teste é obtido calculando a probabilidade de $\hat{p} < p_c$ dado que $p = 0,6$ menos 100%.
- (C) o nível de confiança é obtido calculando a probabilidade de $\hat{p} < p_c$ dado que $p = 0,6$.
- (D) a probabilidade de $\hat{p} < p_c$ dado que $p > 0,6$ é igual à probabilidade de ocorrer o erro tipo II.
- (E) a probabilidade de $\hat{p} < p_c$ dado que $p = 0,6$ é igual à probabilidade de ocorrer o erro tipo I.

55. São vários os procedimentos para a busca do “subconjunto ótimo” de variáveis, na ausência da ortogonalidade, para obter uma equação de estimação adequada que relaciona uma variável Y a todas ou a um subconjunto de variáveis independentes. Considere o seguinte procedimento:

PASSO 1: Escolha a variável que fornece a maior soma de quadrados da regressão em regressão linear simples com Y ou, de maneira equivalente, que forneça o maior valor de R^2 . Chamaremos essa variável inicial de X_1 .

PASSO 2: Escolha a variável que, quando inserida no modelo, fornece o maior aumento em R^2 , na presença de X_1 , sobre o valor de R^2 encontrado no passo 1, isto é, a variável X_j para a qual:

$$R(\beta_j|\beta_1) = R(\beta_1, \beta_j) - R(\beta_1)$$

é maior. Vamos chamá-la de variável X_2 . O modelo de regressão com X_1 e X_2 é, então, ajustado e R^2 é observado.

PASSO 3: Escolha a variável X_j que fornece o maior valor de:

$$R(\beta_j|\beta_1, \beta_2) = R(\beta_1, \beta_2, \beta_j) - R(\beta_1, \beta_2),$$

resultando novamente em um aumento em R^2 sobre aquele dado no PASSO 2. Ao chamar essa variável de X_3 , agora temos um modelo de regressão que envolve X_1 , X_2 e X_3 . Esse processo é continuado até que a variável inserida mais recentemente falhe ao produzir um aumento significativo na regressão explicada. Tal aumento pode ser determinado em cada passo, devendo-se usar o teste F (ou t) apropriado.

Por exemplo, no PASSO 2, o valor: $f = R(\beta_2|\beta_1)/s_{1,2}^2$ pode ser determinado para testar a adequação de X_2 no modelo. De maneira similar, no PASSO 3 a razão: $f = R(\beta_3|\beta_1, \beta_2)/s_{1,2,3}^2$ testa a adequação de X_3 no modelo.

Se $f < f_{(1, n-3; \alpha)}$ no PASSO 2, para um nível de significância preestabelecido, X_2 não é incluído e o processo é encerrado, resultando em uma equação linear simples que relaciona Y e X_1 .

Contudo, se $f > f_{(1, n-3; \alpha)}$ deve-se seguir para o PASSO 3. Novamente, se $f < f_{(1, n-4; \alpha)}$ no PASSO 3, X_3 não é incluído e o processo é encerrado com a equação de regressão apropriada que contém as variáveis X_1 e X_2 .

Notações utilizadas:

R^2 é o coeficiente de determinação do modelo de regressão;

$R(\cdot)$ é a soma dos quadrados do modelo de regressão em questão;

β_j é o coeficiente do modelo de regressão que acompanha a variável X_j ;

A notação ‘|’ indica a probabilidade condicional;

$s_{1,2}^2$ é o quadrado do erro médio para o modelo que contém as variáveis X_1 e X_2 ;

$s_{1,2,3}^2$ é o quadrado do erro médio para o modelo que contém as variáveis X_1 , X_2 e X_3 .

Essa descrição se refere ao método de seleção de variáveis:

- (A) Forward (‘para frente’).
- (B) Lasso bayesiano (BLASSO).
- (C) Multicolinearidade.
- (D) Bootstrap.
- (E) Backward (‘para trás’).

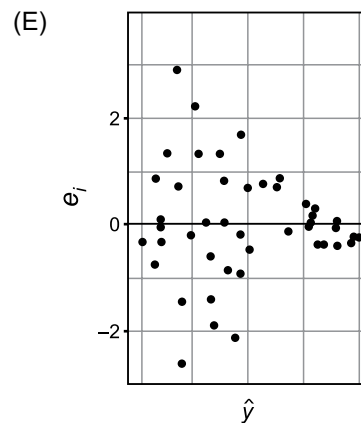
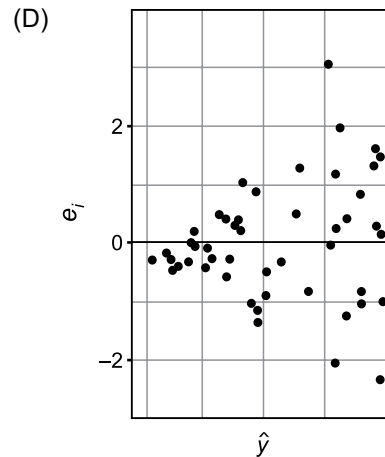
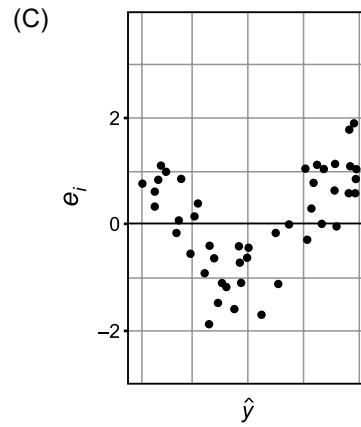
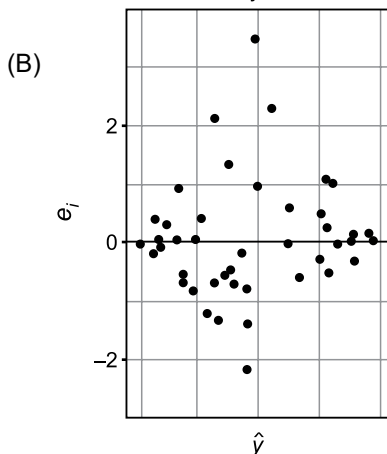
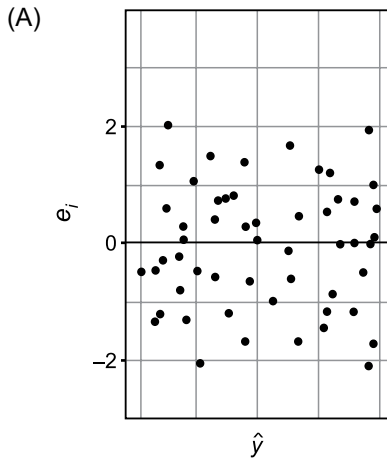
R A S C U N H O

56. Em relação ao coeficiente de correlação de Pearson populacional e amostral (ρ e r , respectivamente), é correto afirmar:

- (A) O teste de hipóteses para correlação de Pearson tem como suposição única para sua aplicação que as variáveis envolvidas sejam ambas contínuas.
- (B) A obtenção do coeficiente de correlação amostral é equivalente ao cálculo da variância amostral quando o tamanho da amostra é pequeno ($n < 30$).
- (C) Em um modelo de regressão linear múltipla ajustado aos dados, tirando a raiz quadrada do coeficiente de determinação é possível obter o coeficiente de correlação amostral (r) entre as variáveis desse modelo.
- (D) Não é possível estabelecer um modelo de regressão linear para duas variáveis quando o coeficiente de correlação de Pearson amostral é igual a zero.
- (E) O coeficiente de correlação entre duas variáveis é uma medida de sua relação linear e que um valor de $r = 0$ implica falta de linearidade e não falta de associação.

57. No ajuste de qualquer modelo estatístico, uma das principais tarefas do analista é avaliar se as pressuposições assumidas por determinada metodologia são satisfeitas, bem como a qualidade desse ajuste aos dados. Para isso, fazer uma análise de resíduos se torna imprescindível. Há três tipos de violações das suposições que são prontamente detectados por meio do uso de gráficos residuais; são elas: (i) presença de valores discrepantes; (ii) variância do erro heterogênea; e (iii) especificação do modelo inadequada.

Diante disso, o gráfico de valores preditos (\hat{y}) versus resíduos padronizados (e_i), que indica uma especificação do modelo inadequada para a situação em estudo, é:



R A S C U N H O

58. Um modelo de regressão linear simples ($Y_i = a + b X_i + e_i$, sendo $i = 1, 2, \dots, 33$) foi ajustado a uma amostra aleatória de uma determinada população, onde se obteve as seguintes informações referentes à análise de variância desse modelo: (i) a soma de quadrados referente a regressão foi igual a 3390; e (ii) a soma de quadrados totais foi igual a 3713. A estimativa não viciada para a variância populacional e a interpretação do coeficiente de determinação desse modelo são, respectivamente:
- (A) 116,03. O modelo ajustado aos dados explica 91,3% da variabilidade total da variável Y .
- (B) 10,42. O modelo ajustado aos dados explica 91,3% da variabilidade total da variável Y .
- (C) 10,77. O modelo ajustado aos dados explica 8,7% da variabilidade total da variável Y .
- (D) 323. O modelo ajustado aos dados explica 9,5% da variabilidade total da variável X .
- (E) 109,35. O modelo ajustado aos dados explica 8,7% da variabilidade total da variável X .

59. Um pesquisador em educação infantil deseja verificar se existe (ou não) associação entre desenvolvimento da fala e o engatinhar (se engatinhou ou não engatinhou) de bebês de uma determinada região. Para isso, ele observou uma amostra de 70 bebês de uma mesma faixa etária e sexo. Na tabela a seguir é apresentado um resumo (frequências absolutas) obtido desse estudo.

Fala	Engatinhou	
	Não	Sim
Não	23	14
Sim	15	18

Considerando um nível de significância de 10%, o valor aproximado calculado do teste (utilize para os cálculos os valores inteiros aproximados das frequências esperadas), juntamente com a sua conclusão, são, respectivamente:

- (A) 2,08. Então, rejeita-se H_0 , e conclui-se que o processo de fala depende do engatinhar (ou não) dos bebês dessa região.
- (B) 2,08. Então, aceita-se H_0 , e conclui-se que o processo de fala independe do engatinhar (ou não) dos bebês dessa região.
- (C) 1,94. Então, aceita-se H_0 , e conclui-se que o processo de fala depende do engatinhar (ou não) dos bebês dessa região.
- (D) 36,00. Então, aceita-se H_0 , e conclui-se que o processo de fala depende do engatinhar (ou não) dos bebês dessa região.
- (E) 1,94. Então, aceita-se H_0 , e conclui-se que o processo de fala independe do engatinhar (ou não) dos bebês dessa região.

Dados: $P(Z < 1,28) = 0,90$; $P(Z > 1,64) = 0,05$; $P(X > 85,53) = 0,10$; $P(Y < 84,42) = 0,90$ e $P(W < 2,7055) = 0,90$; em que Z é uma variável aleatória que segue uma distribuição normal padrão; X , Y e W são variáveis aleatórias que seguem uma distribuição qui-quadrado com respectivamente 70, 69 e 1 graus de liberdade.

60. Sobre o teste qui-quadrado, é correto afirmar:

- (A) O teste qui-quadrado para ajustamento de dados (aderência) só pode ser aplicado a dados contínuos.
- (B) O teste qui-quadrado é equivalente ao teste de Kolmogorov-Smirnov quando se tem amostras pequenas ($n \leq 30$).
- (C) O teste qui-quadrado é um teste versátil, pois com ele se pode testar a independência de duas variáveis quaisquer, a homogeneidade, a multicolinearidade e o ajustamento de dados contínuos (aderência).
- (D) O teste qui-quadrado testa a autocorrelação entre duas variáveis contínuas.
- (E) O teste qui-quadrado de aderência no ajustamento de uma distribuição contínua a uma amostra, deve-se proceder à agregação dos dados em classes.

R A S C U N H O

