

## CONHECIMENTOS ESPECÍFICOS

X	0	1	2	3	4	5
frequência absoluta	80	47	30	20	6	1

Um levantamento foi realizado para se avaliar, por município, a quantidade X de obras que estão sob suspeita de irregularidade. Com base em uma amostra de municípios, foi obtida a distribuição de frequências mostrada na tabela acima. Com base nessas informações, julgue os itens a seguir.

- 51 O diagrama de setores (ou *pizza*) é um gráfico apropriado para se representar a distribuição de probabilidade da variável X.
- 52 Os extremos mínimo e máximo da variável X foram, respectivamente, iguais a 1 e 80.
- 53 O número de municípios que têm obras sob suspeita de irregularidades é superior a 120.
- 54 O total de municípios considerado no levantamento foi superior a 180.
- 55 A distribuição da quantidade X apresenta assimetria à direita (ou positiva).
- 56 Considerando que a média de X seja igual a 1,1 e que  $CV = 133\%$ , então a variância da distribuição de X é superior a 1,8.
- 57 O primeiro quartil e o quarto decil da distribuição de X são iguais a 0.

fonte de variação	graus de liberdade	soma dos quadrados	média dos quadrados	razão F
regressão	1	2.061,49	2.061,49	433,40
erro	78	371,01	4,75	
total	79	2.432,50		

Um analista deseja avaliar se o tempo — Y —, em dias, que um processo judicial leva para ser concluído está relacionado com a quantidade — X — de juizes disponíveis no tribunal em que tal processo foi julgado. O quadro acima apresenta a tabela de análise de variância (ANOVA) correspondente a essa avaliação por regressão linear simples, em que Y é a variável resposta e X é a variável regressora, com base no método de mínimos quadrados ordinários. Considerando essas informações e os conceitos de análise de regressão linear e inferência estatística, julgue os itens de 58 a 69.

- 58 Suponha que o modelo ajustado tenha aproximadamente a forma  $Y = -2X + B$ , em que B é uma estimativa do intercepto. Nessa situação, estima-se que cada juiz acrescentado ao tribunal representa uma redução de dois dias, em média, no tempo para a conclusão de um processo nesse tribunal.
- 59 Para a modelagem foram consideradas 79 observações.
- 60 O coeficiente de determinação  $R^2$  do modelo ajustado foi superior a 80%.
- 61 Se a estimativa do coeficiente angular do modelo for inferior a -2, com base nesse valor, é correto inferir que a variável X explica parte da variação de Y de modo significativo.
- 62 Considere que a distribuição dos tempos Y siga uma distribuição normal e que o analista efetue o ajuste do modelo pelo método da máxima verossimilhança. Nessa situação, o modelo ajustado será diferente do modelo ajustado pelo método de mínimos quadrados ordinários.
- 63 O desvio padrão amostral da variável resposta Y foi inferior a 6.

- 64 O método de mínimos quadrados ordinários produz estimadores BLUE (*best linear unbiased estimator*), pois as variâncias desses estimadores são iguais à variância do erro aleatório do modelo.
- 65 As variáveis X e Y possuem correlação linear superior a 0,90.
- 66 Suponha que a estatística t referente ao coeficiente angular do modelo tenha sido superior a 28. Nessa situação, é correto concluir que tal coeficiente é estatisticamente significativo a um nível de significância de 5%.
- 67 Se a distribuição dos resíduos não for normal e assimétrica, o teste não paramétrico dos sinais será apropriado para testar se a média dos resíduos é igual a 0.
- 68 Uma ferramenta descritiva para avaliação e diagnóstico do modelo é o gráfico de resíduos. Nesse gráfico, os resíduos devem apresentar-se dispostos aleatoriamente em torno do ponto zero.
- 69 Se o valor zero for contemplado pelo intervalo de confiança do coeficiente linear, é correto afirmar que o coeficiente ajustado não foi estatisticamente significativo.

RASCUNHO

Um fiscal deverá escolher aleatoriamente algumas áreas de certa floresta para serem visitadas. Com base em um mapa, o fiscal dividiu a floresta em regiões mutuamente exclusivas, formando uma partição. Essas regiões possuem áreas distintas.

Tendo como referência essa situação, julgue os itens de 70 a 81, com base nos conceitos de probabilidade e inferência estatística.

70 Considere que a distribuição de uma variável de interesse do fiscal siga uma distribuição assimétrica e bimodal. Nessa situação, se  $\bar{x}$ ,  $\mu$  e  $\sigma$  representam, respectivamente, a média amostral, a média populacional e o desvio padrão populacional dessa distribuição, então é correto afirmar que  $\frac{(\bar{x}-\mu)\sqrt{n}}{\sigma} \sim N(0,1)$ , se o tamanho da amostra  $n$  for suficientemente grande.

71 Suponha que o fiscal escolherá aleatoriamente duas regiões. Considere que a probabilidade de a região A ser escolhida seja igual a 0,3, enquanto a probabilidade de se escolher a região B seja igual a 0,1. Nesse caso, considerando que há independência entre os eventos considerados, se A foi a primeira região escolhida, então a probabilidade de a segunda região escolhida ser B é igual a 0,1.

72 Considere que uma variável a ser avaliada pelo fiscal seja a proporção da área desmatada em cada região visitada. Se essa proporção de área desmatada for uma variável aleatória, então um modelo de distribuição discreta permite adequadamente representar a distribuição dessas proporções.

73 Suponha que o fiscal deseje estudar a distribuição do número de irregularidades por região visitada. Nessa situação, o número médio de irregularidades por região é uma variável aleatória discreta, e a quantidade total esperada de irregularidades na floresta é igual à soma simples das quantidades médias de irregularidades por região.

74 Considere que a floresta tenha sido dividida em 10 regiões e que a soma das probabilidades de seleção das 8 maiores regiões seja igual a 0,8. Nessa situação, a probabilidade de seleção da menor região é superior a 0,2.

75 Suponha que a distância percorrida (em km) pelo fiscal seja uma variável aleatória  $X$ , contínua e positiva. Nesse caso, a probabilidade de o fiscal percorrer mais de 100 km é

dada por  $\int_0^{100} f(x)dx$ , em que  $f(x)$  é a função densidade de probabilidade da variável aleatória  $X$ .

76 Considere que uma das variáveis de interesse do fiscal seja a concentração (em ppm) de certo poluente em estações de observação e que tal concentração siga uma distribuição normal  $X$  cujo desvio padrão seja igual a 5 ppm. Nessa situação, a probabilidade de se registrar uma concentração superior à média de  $X$  é inferior a 0,49.

77 Suponha que um índice de qualidade do ar a ser monitorado pelo fiscal siga uma distribuição normal  $X$  com média 30 e variância 25. Sabendo-se que  $P(Z > 1,645) = 0,05$ , em que  $Z \sim N(0, 1)$ , é correto afirmar que  $P(X > 38,225) = 0,05$ .

78 Suponha que a mediana amostral  $Md$  seja uma estatística considerada pelo fiscal para efetuar inferências acerca da média  $\mu$  de uma distribuição simétrica cujo desvio padrão populacional é igual a  $\sigma$ . Nessa situação, para uma amostra de tamanho  $n$  suficientemente grande, pelo teorema central do limite, é correto afirmar que  $\frac{(Md - \mu)\sqrt{n}}{\sigma} \sim N(0,1)$ .

79 Considere que a floresta tenha sido dividida em quatro regiões, A, B, C e D, e que  $P_A$ ,  $P_B$ ,  $P_C$  e  $P_D$  sejam as correspondentes probabilidades de seleção dessas regiões. Se duas regiões forem escolhidas aleatoriamente pelo fiscal, então a probabilidade de a amostra contemplar a região A será igual a  $(P_A)^2$ .

80 Suponha que o número  $X$  de espécies de plantas nas áreas selecionadas pelo fiscal siga uma distribuição de Poisson com média  $\lambda$ . Nessa situação, dependendo da magnitude das contagens  $X$ , a distribuição de Poisson pode ser aproximada pela distribuição normal com média  $\lambda$  e desvio padrão  $\sqrt{\lambda}$ .

81 Se a probabilidade de se escolher a região G cai pela metade depois de se escolher a região F e se  $P(F) = 0,12$  e  $P(G) = 0,08$ , então é correto afirmar que a probabilidade de se escolherem os locais F ou G é superior a 0,2.

RASCUNHO

Com relação aos testes não paramétricos, julgue os próximos itens.

- 82 O teste F utilizado na análise de variância não é válido em situações em que os dados não seguem distribuição normal e(ou) a variância da variável resposta não é constante em função das condições de avaliação do experimento. Nesses casos, o teste de Kruskal-Wallis é mais apropriado.
- 83 Para a comparação de duas amostras independentes, são utilizados tanto o teste de Wilcoxon como o de Wilcoxon-Mann-Whitney.
- 84 Quando a hipótese de normalidade é satisfeita, em pequenas amostras, o teste dos sinais é mais poderoso que o teste  $t$  para amostras pareadas, já que a distribuição  $t$  possui caudas mais pesadas que a distribuição normal padrão.
- 85 O teste dos sinais de McNemar permite efetuar comparações entre dois tratamentos cuja variável resposta é binária.

Julgue os itens seguintes considerando o cenário de amostragem aleatória simples, processo de amostragem probabilística no qual todas as  $n$  unidades amostrais, das  $N$  unidades que compõem a população ( $N > n$ ), possuem as mesmas probabilidades de seleção.

- 86 Na amostragem aleatória simples, a ordem dos elementos na amostra não é relevante: as amostras  $\{X, Y, Z\}$  e  $\{Z, X, Y\}$ , por exemplo, são consideradas idênticas.
- 87 Suponha que de uma população com  $N = 12$  elementos seja retirada a amostra  $\{10; 7; 13\}$  de tamanho  $n = 3$ . Nessa situação, estima-se que a variância da média amostral seja igual a 3.
- 88 Ao se realizar o cálculo do tamanho da amostra para um levantamento sobre uma população finita de tamanho  $N$ , o fator de correção para populações finitas corresponde a uma grandeza sempre inferior a 1.
- 89 Em uma população de  $N = 15$  elementos, o número de possíveis amostras de tamanho  $n = 2$  é igual a 210.

Considere que uma população composta por 100 indivíduos seja estratificada da seguinte maneira: o estrato 1 contempla 50 indivíduos, o estrato 2 é composto por 30 indivíduos e o estrato 3 é formado por 20 indivíduos. Com base nessas informações, julgue os itens subsecutivos.

- 90 Suponha que sejam retiradas amostras de tamanhos iguais nos três estratos e que as estimativas da média por estrato sejam  $\bar{X}_{\text{estrato 1}} = 10$ ,  $\bar{X}_{\text{estrato 2}} = 6$  e  $\bar{X}_{\text{estrato 3}} = 8$ . Nessa situação, a estimativa global da média populacional é igual a 8,4.
- 91 Suponha que se deseje tomar uma amostra dessa população optando-se pela alocação proporcional ao tamanho dos estratos. Nessa situação, sendo amostrados 9 indivíduos do estrato 2, é correto inferir que a quantidade de indivíduos na amostra era igual a 27.

A respeito de amostragem por conglomerados, julgue o próximo item.

- 92 Na amostragem por conglomerados em estágio único, é possível fixar, *a priori*, o tamanho exato da amostra total ainda que os tamanhos dos conglomerados sejam diferentes.

Considerando  $X_1, X_2$  e  $X_3$  três valores gerados independentemente a partir de uma distribuição de Bernoulli com probabilidade de sucesso igual a 0,3, julgue os itens seguintes.

- 93 A soma  $Y = 2X_1 + 2X_2 + 2X_3$  pode ser vista como uma realização de uma distribuição binomial com parâmetros  $n = 6$  e  $p = 0,3$ .
- 94 O valor  $W$ , em  $W = 3 - X_1 - X_2 - X_3$ , pode ser visto como uma realização de uma distribuição binomial com média igual a 2,1.

Com relação à estatística computacional, julgue os itens que se seguem.

- 95 O algoritmo da aceitação-rejeição permite simular valores de variáveis aleatórias contínuas, mas não realizações de variáveis discretas.
- 96 Considere que  $X$  seja uma variável aleatória com distribuição exponencial com média igual a  $\lambda$ . Nessa situação, de acordo com o método da transformação inversa, uma realização  $x$  dessa distribuição exponencial pode ser gerada por  $x = (1/\lambda)\ln(1 - u)$ , em que  $u$  é uma realização da distribuição uniforme no intervalo  $[0, 1]$ , e  $\ln$  representa o logaritmo natural.
- 97 Para se gerar uma realização  $x$  de uma distribuição normal com média 5 e variância 4 a partir de uma distribuição normal padrão, basta gerar uma realização  $z$  de uma distribuição  $N(0, 1)$  e, em seguida, aplicar a transformação  $x = 5 + 4z$ .

RASCUNHO

$$P = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 0,3 & 0,7 & 0 & 0 & 0 \\ 0,5 & 0,5 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0,2 & 0,8 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{matrix}$$

A matriz acima corresponde à matriz de transição de estados de uma cadeia de Markov em tempo discreto. Julgue os itens a seguir com base na matriz.

- 98 A probabilidade de o processo sair do estado 1 e passar para o estado 2, em dois passos, é igual a 0,49.
- 99 Os estados 1 e 2 dessa cadeia de Markov são recorrentes.
- 100 O estado 4 da matriz  $P$  é transiente.
- 101 Os estados 3 e 5 são absorventes.

Considerando que o número de consumidores que chegam a determinada loja seja descrito por um processo de Poisson homogêneo com taxa de 30 consumidores por hora, julgue os itens subsequentes.

- 102 Sendo o percentual de consumidores do sexo feminino que frequentam essa loja igual 70%, a probabilidade de nenhum consumidor homem entrar nessa loja durante um intervalo de 15 minutos corresponde a  $\exp\{-9/4\}$ .
- 103 A média do intervalo de tempo entre as chegadas de dois consumidores consecutivos é de 30 segundos.

Considerando o modelo de fila M/M/1, em que os usuários entram na fila conforme um processo de Poisson com taxa igual a 5 usuários por hora e o fato de que, nesse sistema de servidor único, a taxa de atendimento é de 6 usuários por hora, julgue os itens seguintes.

- 104 Com o sistema em equilíbrio, o número esperado de usuários no sistema (em atendimento ou esperando na fila) é igual a 5; se a capacidade de atendimento elevar-se para 7 indivíduos por hora, o número médio de indivíduos no sistema será reduzido em 50%.
- 105 Se um novo usuário entra na fila quando 5 usuários já se encontram no sistema, então o seu tempo de espera até ser atendido segue uma distribuição gama cuja média é igual a 1 hora.

Considerando que, em uma fila única do caixa-rápido em um supermercado, os clientes cheguem, de acordo com um processo de Poisson, com taxa igual a 20 pessoas por hora e que o número de atendimentos de cada caixa seja representado por um processo de Poisson com taxa de 15 atendimentos por hora, julgue o próximo item.

- 106 É necessário haver, pelo menos, 5 caixas em operação para que o tamanho da fila não cresça indefinidamente.

Considerando que as variáveis aleatórias  $X_1$ ,  $X_2$  e  $X_3$  sejam conjuntamente distribuídas segundo uma distribuição normal multivariada cujo vetor de médias e matriz de covariâncias sejam, respectivamente,

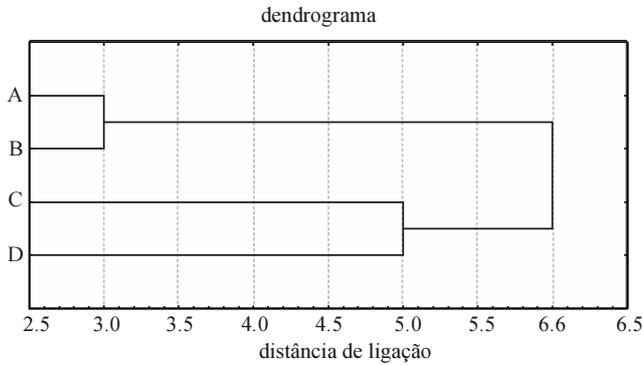
$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix} \text{ e } \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 9 & 3 \\ 0 & 3 & 4 \end{bmatrix},$$

em que  $\sigma_{ij} = \text{Cov}(X_i, X_j)$ ,  $i, j = 1, 2, 3$ , julgue os itens seguintes.

- 107 A variável  $Y = 2X_1 + X_2$  tem distribuição normal com média 4 e variância 17.
- 108 As variáveis  $X_1$  e  $X_3$  são independentes.
- 109 A distribuição condicional  $X_3|X_2 = 4$  (distribuição condicional de  $X_3$ , dado  $X_2 = 4$ ) segue uma distribuição normal com média igual a 2.

RASCUNHO

$$\text{distâncias} = \begin{matrix} & \begin{matrix} A & B & C & D \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \end{matrix} & \begin{bmatrix} 0 & & & \\ 3 & 0 & & \\ 7 & 9 & 0 & \\ 8 & 6 & 5 & 0 \end{bmatrix} \end{matrix}$$



Considerando a matriz de distâncias e o dendrograma acima, que se referem aos objetos *A*, *B*, *C* e *D* de um problema de análise de conglomerados, julgue os itens que se seguem.

- 110 Com base no método hierárquico aglomerativo, é correto afirmar que o agrupamento representado no dendrograma foi realizado conforme os seguintes passos: fusão dos objetos *A* e *B*, formando-se o grupo (*AB*); fusão dos objetos *C* e *D*, formando-se o grupo (*CD*); e agrupamento do grupo (*AB*) com o (*CD*).
- 111 Caso se faça a opção por dois conglomerados, um deles será composto pelos objetos *A* e *B* e o outro, pelos objetos *C* e *D*, composições essas que serão mantidas mesmo se a distância entre os objetos *A* e *B* for menor.

Com relação à análise discriminante e à rotação *varimax*, julgue os próximos itens.

- 112 Muito utilizada em análise fatorial, a rotação *varimax* é ortogonal, ou seja, os fatores rotacionados não são correlacionados.
- 113 A análise discriminante constitui técnica multivariada que permite classificar os objetos que formam um conjunto de dados e, com base nela, é possível segmentar um conjunto de objetos em *k* grupos homogêneos sem se dispor de uma informação *a priori* acerca da alocação desses objetos nesses grupos.

Em uma análise multivariada, as variáveis  $X_1$ ,  $X_2$  e  $X_3$  possuem matriz de covariâncias dada por

$$\Sigma = \begin{bmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \text{Cov}(X_1, X_3) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \text{Cov}(X_2, X_3) \\ \text{Cov}(X_3, X_1) & \text{Cov}(X_3, X_2) & \text{Var}(X_3) \end{bmatrix} = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

e as seguintes componentes principais:

$$C_1 = 0,7071X_1 + 0,7071X_2;$$

$$C_2 = 0,7071X_1 - 0,7071X_2;$$

$$C_3 = X_3.$$

Com base nessas informações, julgue os itens a seguir.

- 114 A terceira componente principal explica 20% da variação total.
- 115 A correlação entre as componentes  $C_1$  e  $C_2$  é maior que 0,1 e menor que 0,5.
- 116 A soma das variâncias das três componentes é igual a 10.

Considerando uma série temporal  $\{Z_t\}$ ,  $t = 1, 2, \dots, n$ , em que

$$Z_t = a_t - \frac{a_{t-1}}{2} \text{ e } a_t \text{ é o ruído branco com média } 0 \text{ e variância } 4,$$

julgue os itens que se seguem.

- 117 A autocorrelação entre  $Z_t$  e  $Z_{t-2}$  é nula.
- 118 A variância de  $Z_t$  é igual a 3.

Julgue o item abaixo, a respeito da estacionariedade e invertibilidade dos processos que descrevem as séries temporais.

- 119 Para o modelo de médias-móveis (MA) de ordem finita, não há restrições sobre os parâmetros para que o processo seja estacionário.

Considerando que uma série temporal  $\{Y_t\}_{t=1, 2, \dots, n}$  segue o processo ARIMA(0, 2, 1) e que  $W_t = Y_t - 2Y_{t-1} + Y_{t-2}$ , julgue o item a seguir.

- 120 A série temporal  $W_t$  é estacionária e segue um processo de médias móveis de ordem 1.

RASCUNHO

## PROVA DISCURSIVA

- Faça o que se pede a seguir, usando o espaço para rascunho indicado no presente caderno. Em seguida, transcreva o texto para a **FOLHA DE TEXTO DEFINITIVO DA PROVA DISCURSIVA**, no local apropriado, pois **não serão avaliados fragmentos de texto escritos em locais indevidos**.
- Qualquer fragmento de texto além da extensão máxima de **trinta** linhas será desconsiderado.
- Na **folha de texto definitivo**, identifique-se apenas no cabeçalho da primeira página, pois **não será avaliado** texto que tenha qualquer assinatura ou marca identificadora fora do local apropriado.

### Dilma sanciona comissão da verdade e lei de acesso à informação

A presidenta Dilma Rousseff sancionou, em 18 de novembro, a lei que permite aos cidadãos ter acesso a informações públicas e a lei que cria a Comissão da Verdade. Em cerimônia no Palácio do Planalto, Dilma destacou que essas duas leis "representam um grande avanço nacional e um passo decisivo na consolidação da democracia brasileira". "A informação torna-se aberta em todas as suas instâncias. O poder público torna-se mais transparente", completou a presidenta.

Internet: <www.estadao.com.br> (com adaptações).

---

Considerando que o fragmento de texto acima tem caráter unicamente motivador, redija um texto dissertativo a respeito do seguinte tema.

### ACESSO À INFORMAÇÃO NO SERVIÇO PÚBLICO

Ao elaborar seu texto, comente, necessariamente, sobre os seguintes aspectos:

- ▶ origens e objetivo da Comissão da Verdade;
- ▶ objetivo da Lei de Acesso à Informação.

**RASCUNHO**

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	