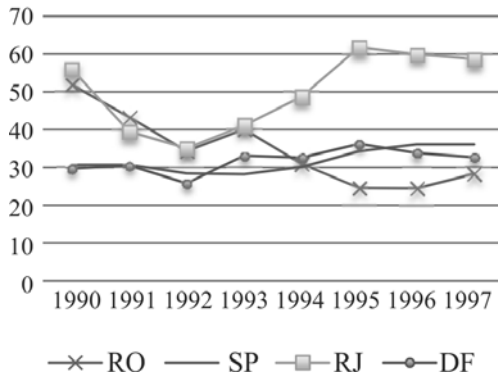


## CONHECIMENTOS ESPECÍFICOS

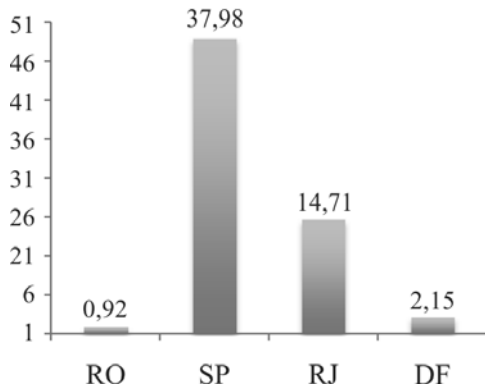
**QUESTÃO 21**

**RASCUNHO**

**Taxas de homicídio/100 mil hab.**



**população (milhões de hab.) em 2011**



IBGE: Séries estatísticas e históricas.

Os gráficos acima mostram a evolução da taxa relativa de homicídios (por 100 mil habitantes) nos estados de Rondônia, São Paulo, Rio de Janeiro e no Distrito Federal; e as populações residentes nesses estados, em 2011 (em milhões de habitantes). Com base nesses gráficos, assinale a opção correta.

- A** Suponha que, em 1997, as quantidades de homicídios ocorridos no estado de Rondônia e no do Rio de Janeiro tenham sido iguais. Nesse caso, nesse ano, as taxas de homicídios nesses estados foram iguais.
- B** Suponha que as populações nessas unidades da Federação cresceram, de 2001 até 2011, na mesma proporção, e que em 2011 a população de Rondônia atingiu 1 milhão de habitantes. Nesse cenário, é correto afirmar que, em 2011, o estado de São Paulo tinha menos de 40 milhões de habitantes.
- C** Se, em 1997, o Distrito Federal tivesse 10% mais habitantes do que o registrado, e se a taxa de homicídios fosse a mesma, então teriam ocorrido, nesse ano, aproximadamente mais 650 casos de homicídio.
- D** Em 1997, ocorreram mais homicídios no estado de São Paulo que no estado do Rio de Janeiro.
- E** A cada ano, entre 1990 e 1997, a quantidade de homicídios no Distrito Federal foi aproximadamente igual à do estado de São Paulo.

## RASCUNHO

## QUESTÃO 22

Considerando que  $x_1, x_2, \dots, x_n$  seja uma amostra aleatória e que  $\bar{x}$  seja a média aritmética dessa amostra, é correto afirmar que

$$\phi = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})$$

- A** é uma medida de dispersão, pois os termos da soma dependem apenas da distância entre os valores amostrais e a média correspondente.
- B** é uma medida de dispersão que se denomina amplitude média.
- C** é uma medida de dispersão chamada erro padrão.
- D** não é uma medida de dispersão, a menos que a amostra tenha uma distribuição assimétrica.
- E** não é uma medida de dispersão, pois, independentemente da distribuição dos dados, o valor de  $\phi$  será sempre nulo.

## QUESTÃO 23

Considerando a sequência de dados ordenados: 1,  $a$ ,  $b$ ,  $c$ , 8, em que  $1 \leq a \leq b \leq c \leq 8$ , assinale a opção correta.

- A** As sequências  $\{1, 1, 1, 8, 8\}$  e  $\{1, 1, 8, 8, 8\}$  correspondem a caso em que a variância amostral é máxima.
- B** Se  $a = b = c$  então o intervalo interquartilício é igual a 7.
- C** Se  $2b = c + a$ , é correto concluir que há simetria entre os dados.
- D** A menor variância amostral possível para esses dados é zero.
- E** Independentemente dos dados, a variância amostral será sempre inferior a 10.

## QUESTÃO 24

Um perito investiga se determinada empresa manipulou os resultados fiscais com o objetivo de sonegar impostos. Considere que  $M$  e  $L$  representam, respectivamente, os eventos “manipula os resultados fiscais” e “obteve lucro no último trimestre”, e que  $M^c$  e  $L^c$  denotam os eventos complementares correspondentes. Com base em investigações anteriores, é conhecida a proporção de empresas que manipulam os resultados fiscais, ou seja,  $P(M)$ . Conhece-se também a probabilidade condicional  $P(M|L)$ . O perito determinou a proporção de empresas que tiveram lucro no último trimestre,  $P(L)$ , com base em dados fornecidos pela junta comercial. Com base nessas informações, assinale a opção correta.

- A**  $P(M) = P(M|L^c) \times P(L^c) + P(M|L) \times P(L)$ .
- B**  $P(M|L^c) = 1 - P(M|L)$ .
- C** Não há qualquer relação entre as probabilidades  $P(L^c)$  e  $P(L)$  e a probabilidade  $P(M)$ .
- D** Suponha que  $P(L) = 0,8$ ,  $P(M|L^c) = 0,4$  e  $P(M) = 0,1$ . Nesse caso, é correto afirmar que  $P(M|L) > 0,9$ .
- E** A probabilidade condicional  $P(M|L^c)$  pode ser determinada pela fórmula de Bayes.

## RASCUNHO

## QUESTÃO 25

$X$  é uma variável aleatória cuja função geradora de momentos é dada por  $\psi(t) = \exp\left[\mu t + \frac{(\sigma t)^2}{2}\right]$ , em que  $\mu$  e  $\sigma$  são, respectivamente, a média e o desvio padrão de  $X$ . Considerando que  $\{X_1, X_2, \dots, X_{10}\}$  é uma sequência de cópias independentes de  $X$ , assinale a opção correta.

- A** A função geradora de momentos da soma de quadrados

$$\sum_{i=1}^{10} X_i^2 \text{ é dada pela expressão } \psi_Z(t) = \left(\frac{1}{1-2t}\right)^5.$$

- B** A distribuição de  $X$  é assimétrica.  
**C** Considerando as estatísticas de ordem  $X_{(1)} < X_{(2)} < \dots < X_{(10)}$ , é correto afirmar que  $P(X_{(3)} < \mu) > 15/130$ .  
**D** A covariância entre duas variáveis distintas  $X_i$  e  $X_k$ , em que  $i \neq k$ , é dada por

$$\text{cov}(X_i, X_k) = \frac{d^2}{dt^2} \psi(t) \Big|_{t=0} - \left[ \frac{d}{dt} \psi(t) \Big|_{t=0} \right]^2.$$

- E** Considerando que  $X_{(10)} = \max \{X_1, X_2, \dots, X_{10}\}$ , é correto afirmar que a função de probabilidade acumulada de  $X_{(10)}$  é

$$\text{dada por } f_{X_{(n)}}(z) = \left[ \exp\left(z - \mu + \frac{\sigma^2}{2}\right) \right]^{10}.$$

## QUESTÃO 26

Na sequência  $\{X_1, X_2, \dots, X_n\}$ , de variáveis aleatórias independentes e identicamente distribuídas, cada variável possui a função geradora de momentos na forma  $\psi(t) = pe^t + 1 - p$ , para  $0 < p < 1$ . Com base nessas informações, assinale a opção correta.

- A** Independentemente do valor de  $p$ , a distribuição da soma de  $\{X_1, X_2, \dots, X_n\}$  será sempre simétrica.  
**B** A variância dessa variável aleatória é igual à sua média.  
**C** A variável aleatória é contínua.  
**D** A variância dessa variável aleatória é maior que a sua média.  
**E** A função geradora de momentos da soma dessas variáveis aleatórias é dada por  $\psi_s(t) = (pe^t + 1 - p)^n, 0 < p < 1$ .

## RASCUNHO

## QUESTÃO 27

Assinale a opção que resulta diretamente da definição axiomática de probabilidade estabelecida por Kolmogorov.

- Ⓐ Para eventos  $A$  e  $B$  é válida a seguinte relação:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}.$$

- Ⓑ Se  $X_1, X_2, \dots, X_n$  é uma sequência de eventos quaisquer, então

$$P\left(\bigcup_{i=1}^n X_i\right) = \sum_{i=1}^n P(X_i).$$

- Ⓒ Se  $X_1, X_2, \dots, X_n$  é uma sequência de eventos e  $X_i^c$ ,  $i = 1, 2, \dots, n$

é o evento complementar de  $X_i$ , então  $P\left(\left[\bigcup_{i=1}^n X_i\right]^c\right) = P\left(\bigcap_{i=1}^n X_i^c\right)$ .

- Ⓓ Se  $A^c$  for o evento complementar do evento  $A$ , então

$$P(A^c) = 1 - P(A).$$

- Ⓔ Se o evento  $B$  e o seu complemento  $B^c$  forem eventos condicionantes do evento  $A$ , então  $P(A|B^c) \neq 1 - P(A|B)$ .

## QUESTÃO 28

Considere que  $\bar{X}$  representa a média amostral de uma amostra aleatória simples de tamanho  $n$  retirada de uma distribuição  $X$ , e  $\mu$  e  $\sigma$  denotam, respectivamente, a média e o desvio padrão dessa distribuição  $X$ . Com relação a inferência estatística, é correto afirmar que a relação  $P\left(\left|\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right| \leq 1,96\right) \xrightarrow{n \rightarrow \infty} 0,95$  é consequência do importante resultado que se intitula

- Ⓐ princípio da invariância.  
 Ⓑ lei fraca dos grandes números.  
 Ⓒ desigualdade de Chebyshev.  
 Ⓓ lei forte dos grandes números.  
 Ⓔ teorema limite central.

## QUESTÃO 29

Os estimadores  $\hat{\theta}_n$  e  $\hat{\theta}_n^*$  são estimadores pontuais do parâmetro  $\theta$  de certa distribuição, em que  $n$  representa o tamanho da amostra. Nesse caso, o estimador  $\hat{\theta}$  é dito ser consistente se

- Ⓐ  $\lim_{n \rightarrow \infty} P(|\hat{\theta}_n - \theta| \geq \epsilon) = 0$ ; para todo  $\epsilon > 0$ .  
 Ⓑ  $P(x_1, x_2, \dots, x_n | \hat{\theta}_n)$ , em que  $x_1, x_2, \dots, x_n$  são as observações amostrais, não depende de  $\theta$ .  
 Ⓒ  $\sigma^2(\hat{\theta}_n) \leq \sigma^2(\hat{\theta}_n^*)$ , em que  $\sigma^2(\cdot)$  é a variância do estimador.  
 Ⓓ  $L(\hat{\theta}_n | x_1, x_2, \dots, x_n) \geq L(\hat{\theta}_n^* | x_1, x_2, \dots, x_n)$ , em que  $L(\cdot)$  é a função de verossimilhança associada ao modelo e  $x_1, x_2, \dots, x_n$  são as observações amostrais.  
 Ⓔ  $E(\hat{\theta}_n) = \theta$ .

RASCUNHO

**QUESTÃO 30**

Com relação a acurácia e precisão, é correto afirmar que o estimador ENVUMV (estimador não viciado uniformemente de mínima variância)

- A é um estimador acurado e é o mais preciso.
- B não é um estimador acurado e nem é o mais preciso.
- C é um estimador acurado mas não é o mais preciso.
- D é um estimador não acurado mas é o mais preciso.
- E é um estimador acurado e assintoticamente o mais preciso.

**QUESTÃO 31**

Com relação a intervalo de confiança e intervalo de credibilidade, é correto afirmar que, se  $I = [a; b]$  for um intervalo de

- A confiança para o parâmetro  $\theta$ , então esse parâmetro é uma variável aleatória.
- B confiança para o parâmetro  $\theta$  de nível  $1 - \alpha$ , então a probabilidade do verdadeiro valor do parâmetro estar fora do intervalo é igual a  $\alpha$ .
- C credibilidade para o parâmetro  $\theta$ , então esse parâmetro é uma variável aleatória.
- D confiança para o parâmetro  $\theta$ , então o valor esperado para esse parâmetro é  $\hat{\theta} = \frac{b + a}{2}$ .
- E credibilidade para o parâmetro  $\theta$ , então a mediana deste parâmetro é  $\hat{\theta} = \frac{b + a}{2}$ .

**QUESTÃO 32**

Nas aplicações de regressão não-paramétrica, a função de densidade de um conjunto de dados pode ser estimada pelo método do *kernel*, que consiste em uma suavização na forma

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right),$$

em que  $h$  é a largura de banda e  $K(\cdot)$  é a função *kernel*. Com relação a seus aspectos característicos, é correto afirmar que a função *kernel* é

- A simplesmente simétrica.
- B simétrica, positiva e de área 1.
- C qualquer função de área é 1.
- D simétrica e de área 1.
- E positiva e de área 1.

## RASCUNHO

## QUESTÃO 33

É correto afirmar que o teste de Kolmogorov-Smirnov é um teste

- A que permite avaliar a simetria dos dados.
- B de aderência que permite avaliar a hipótese de assimetria dos dados.
- C de aderência que somente permite verificar se uma amostra segue a distribuição Normal ou t-Student.
- D de aderência usado para verificar se uma amostra segue determinada distribuição de probabilidade.
- E específico para avaliação da normalidade dos dados.

## QUESTÃO 34

	teste 2 positivo	teste 2 negativo
teste 1 positivo	$a$	$b$
teste 1 negativo	$c$	$d$

Com relação ao teste de McNemar aplicado à tabela de contingência acima, assinale a opção correta.

- A O teste pode ser usado para testar a independência entre as variáveis “teste 1” e “teste 2”.
- B A hipótese do teste é  $H_0: p_b = p_c$  versus  $H_1: p_b \neq p_c$ , em que  $p_b$  e  $p_c$  são as respectivas proporções em  $b$  e  $c$ .
- C A estatística do teste é dada por  $\xi = \frac{(b - c)^2}{a + d}$ .
- D A substituição de  $(b - c)^2$  por  $(b - c)^2 - 0,5$  na estatística do teste foi proposta por Yates como uma correção de continuidade.
- E A estatística do teste tem distribuição qui-quadrado com 4 graus de liberdade.

## QUESTÃO 35

Com respeito ao modelo de regressão linear simples, assinale a opção correta.

- A O parâmetro de inclinação da reta é igual à tangente do ângulo formado entre a reta e o eixo  $Oy$ .
- B A inclinação da reta é proporcional à correlação entre a variável resposta e a variável preditora.
- C Se o modelo linear estiver bem ajustado, a correlação entre o resíduo do modelo e a variável resposta deve estar próxima de  $-1$ .
- D Se o intercepto do modelo for nulo, a variável resposta assume o valor zero quando a variável preditora for igual ao inverso da inclinação da reta.
- E O parâmetro de inclinação da reta é igual ao cosseno do ângulo formado entre a reta e o eixo  $Ox$ .

RASCUNHO

QUESTÃO 36

	gl	SQ	QM	F
regressão	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
erro	<i>e</i>	<i>f</i>	<i>g</i>	

Com relação à tabela de análise de variância mostrada acima, em que  $gl =$  graus de liberdade;  $SQ =$  soma de quadrados;  $QM =$  quadrados médios;  $F =$  estatística  $F$ , assinale a opção correta.

- A O valor  $d$  é igual a  $\frac{b}{c}$ .
- B O valor  $d$  é igual a  $\frac{c}{g}$ .
- C A variância da variável resposta do modelo de regressão é estimada pelo valor  $g$ .
- D O valor  $a + e$  é igual ao tamanho amostral.
- E O valor  $a$  é igual ao valor  $e$ .

QUESTÃO 37

No modelo de regressão múltipla com variável resposta  $y_i$  e regressoras  $x_{i1}, x_{i2}, \dots, x_{ip}$ , a estimativa da variância dos erros aleatórios segue a forma  $s^2 = \frac{\sum \phi_i^2}{\psi}$ . A respeito desse modelo,

assinale a opção correta.

- A  $\phi = x_i - \hat{y}_i$  e  $\psi = n - 2$ .
- B  $\phi = y_i - \hat{y}_i$  e  $\psi = n - p$ .
- C  $\phi = x_i - \hat{x}_i$  e  $\psi = n - 2$ .
- D  $\phi = x_i - \hat{x}_i$  e  $\psi = n - p$ .
- E  $\phi = x_i - \hat{y}_i$  e  $\psi = n - p$ .

QUESTÃO 38

	gl	SQ	QM	F
A	$a - 1$	SQA	QMA	F1
B	$b - 1$	SQB	QMB	F2
AB	X	SQAB	QMAB	F3
Erro	$ab(n - 1)$			

A tabela acima mostra os resultados de uma análise de variância (ANOVA) de um estudo com múltiplos fatores. Nesse caso, o valor dos graus de liberdade da interação (AB) é igual a

- A  $(a - 1)b$ .
- B  $nab - 1$ .
- C  $a(b - 1)$ .
- D  $(a - 1) \times (b - 1)$ .
- E  $a - b$ .

## RASCUNHO

## QUESTÃO 39

O teste de Breusch-Pagan permite avaliar a homocedasticidade em um conjunto de dados e define-se a estatística desse teste como

$$X_{BP}^2 = \frac{SQR^*}{2} \times \left( \frac{SQE}{n} \right)^{-2},$$
 em que  $SQE$  é a soma de quadrados do

resíduo do modelo; e  $SQR^*$  é a soma de quadrados da regressão de  $\varepsilon_i^2$

explicado por  $x_i$ , considerando-se o modelo de regressão linear simples  $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ ,  $\varepsilon_i \sim N(0, \sigma^2)$ . Com relação a esse teste, assinale a opção correta.

- A** Se o modelo for heterocedástico, então  $SQR^*$  será pequeno, fazendo que a estatística do teste tenha um valor baixo.
- B** Se o modelo for heterocedástico, então  $SQR^*$  será muito maior que  $SQE$ , fazendo que a estatística do teste tenha um valor alto.
- C** A sensibilidade do teste para detectar heterocedasticidade não depende do tamanho amostral.
- D** Se o modelo for homocedástico, então  $SQR^*$  será muito maior que  $SQE$ , fazendo que a estatística do teste tenha um valor alto.
- E** A heterocedasticidade é a propriedade dos modelos de regressão cujas observações amostrais são independentes.

## QUESTÃO 40

Assinale a opção correspondente à estatística que pode ser obtida diretamente da estatística do teste t-Student para a comparação de médias entre dois grupos não pareados em pequenas amostras.

- A** estatística  $F$  de uma análise de variância para um fator em dois níveis
- B** estatística  $t$  para a comparação de média entre dois grupos pareados
- C** estatística  $F$  de uma análise de variância para dois fatores
- D** estatística do teste de Wilcoxon
- E** estatística  $z$  para a comparação de média entre dois grupos não pareados



**QUESTÃO 41****RASCUNHO**

estabelecimento	população carcerária
penitenciárias	2.000
presídios e cadeias públicas	1.000
<b>total</b>	<b>3.000</b>

A tabela acima mostra a distribuição da população carcerária em determinada região brasileira segundo o tipo do estabelecimento prisional. Um levantamento estatístico foi planejado para se inferir sobre as condições psicofisiológicas desses presos. Esse estudo contempla uma amostra de 600 presos, dos quais 300 são selecionados aleatoriamente entre os que se encontram nas penitenciárias, e os restantes são selecionados ao acaso dos presídios e das cadeias públicas.

Com base nessas informações, assinale a opção correta a respeito de técnicas de amostragem.

- A** A técnica de amostragem empregada nesse estudo é a por conglomerados.
- B** Em cada estrato retira-se uma amostra aleatória simples de tamanho igual a 300.
- C** A fração amostral do levantamento é igual a 50%.
- D** Cada estabelecimento representa uma unidade amostral.
- E** A alocação da amostra é proporcional ao tamanho dos estratos.

**QUESTÃO 42**

Um estatístico utilizou uma amostragem aleatória estratificada sobre uma população que se divide nos estratos A e B, de tamanhos  $N_A = 20$  mil e  $N_B = 30$  mil, respectivamente. Sabe-se que as variâncias da variável de interesse dentro desses estratos são, respectivamente,  $S_A = 9$  e  $S_B = 4$ . O estatístico retirou uma amostra aleatória de tamanho  $n = 500$ , de acordo com a alocação ótima de Neyman. Com base nessas informações, assinale a opção correspondente às quantidades observadas pelo estatístico nos estratos A e B, respectivamente.

- A** 400 e 100
- B** 200 e 300
- C** 250 e 250
- D** 300 e 200
- E** 350 e 150

## RASCUNHO

## QUESTÃO 43

Um analista estudou o pagamento dos valores  $Y$  (em R\$ mil) das custas processuais em ações trabalhistas. Com base em uma amostra aleatória simples de processos judiciais, ele concluiu que a variável  $Y$  se relaciona linearmente com o valor da causa  $X$  (em R\$ mil), conforme uma reta ajustada pelo método de mínimos quadrados ordinários na forma  $Y = 0,1 \times X + 200$ . A média populacional e a amostral da variável  $X$  foram, respectivamente, iguais a R\$ 100 mil e R\$ 90 mil.

Nesse caso, é correto afirmar que a estimativa de regressão para a média populacional de  $Y$  foi igual a

- Ⓐ R\$ 210 mil.
- Ⓑ R\$ 211 mil.
- Ⓒ R\$ 212 mil.
- Ⓓ R\$ 208 mil.
- Ⓔ R\$ 209 mil.

## QUESTÃO 44

Para um estudo estatístico por simulações computacionais, o analista utilizará um algoritmo para a geração de números pseudo-aleatórios. Ele optou pelo gerador linear congruencial — uma vez que esse é o mais comumente encontrado nos *softwares* estatísticos — que se define pela relação recursiva  $X_{n+1} = (a X_n + b) \bmod M$ , em que  $\{X_n\}$ ,  $n \geq 1$ , representa uma sequência pseudo-aleatória e  $a$ ,  $b$  e  $M$  são constantes. Com relação a esse método de geração de números aleatórios escolhido pelo analista, assinale a opção correta.

- Ⓐ O período da sequência pseudo-aleatória  $\{X_n\}$  depende de uma escolha apropriada do valor inicial  $X_0$ . Por isso, recomenda-se que se utilize sempre o mesmo valor inicial  $X_0$  nas simulações, como, por exemplo, a constante  $\pi = 3,14159$ .
- Ⓑ Para inicializar o gerador de forma apropriada, deve-se escolher um número inteiro — denominado semente — para a constante  $M$ , de tal maneira que se tenha  $M < a$ .
- Ⓒ O valor de  $M$  se relaciona com o período da sequência pseudo-aleatória  $\{X_n\}$ . Para um estudo estatístico por simulações computacionais, o analista deve utilizar valores grandes, como, por exemplo,  $M = 2^{24}$ ,  $2^{32}$  ou  $2^{64}$ .
- Ⓓ Para gerar uma sequência de números pseudo-aleatórios, permite-se ao analista atribuir à constante  $a$  — denominada incremento — qualquer valor real não nulo.
- Ⓔ Se  $b = 0$ , o método chama-se gerador linear congruencial inversível (*invertible congruential generator*); e se  $b < 0$ , o gerador intitula-se gerador de Park-Miller.

## RASCUNHO

## QUESTÃO 45

Em um estudo acerca de incêndios criminosos no cerrado, o perito concluiu que a área queimada  $X$  é uma variável aleatória que segue uma distribuição na forma  $P(X \leq x) = \frac{\exp(ax + b)}{1 + \exp(ax + b)}$ , em que  $a > 0$

e  $-\infty < b < +\infty$  são os seus parâmetros. Esse mesmo estudo considerou simulações de Monte Carlo da variável aleatória  $X$  pelo método da transformação integral (ou método da transformação inversa). Por esse método, a partir de uma realização  $u$  de uma distribuição uniforme no intervalo  $(0, 1)$ , uma realização  $x$  pode ser obtida com base na expressão

A  $x = \frac{1}{a} \ln \frac{u}{1-u} - \frac{b}{a}$ .

B  $x = a \ln \frac{u}{1+u} - b$ .

C  $x = -\frac{\ln u - b}{a}$ .

D  $x = \frac{\exp(au + b)}{1 + \exp(au + b)}$ .

E  $x = \frac{1}{a} \exp \frac{u}{1-u} - \frac{b}{a}$ .

## QUESTÃO 46

O método computacional que se destina à estimação do vício e do desvio padrão de uma estatística é conhecido como

- A *jackknife*.
- B *bootstrap*.
- C *cross-validation*.
- D algoritmo de Metropolis-Hastings.
- E algoritmo EM (*expectation-maximization*).

## QUESTÃO 47

Um estudo sobre reincidência criminal considera uma cadeia de Markov que assume dois estados:  $X(t) = 1$  significa que um indivíduo comete crime no instante  $t$  e  $X(t) = 0$  indica que ele não comete crime no instante  $t$ , em que  $t = 0, 1, 2, 3, \dots$ . Esse estudo apresenta as probabilidades de transição  $P[X(t+1) = 0 | X(t) = 0] = 0,9$  e  $P[X(t+1) = 1 | X(t) = 1] = 0,8$ , de modo que a matriz de transição de estados é dada por  $P = \begin{bmatrix} 0,9 & 0,1 \\ 0,2 & 0,8 \end{bmatrix}$ .

Com base nessas informações, assinale a opção correta acerca do processo estocástico  $\{X(t)\}$ .

- A A cadeia de Markov  $\{X(t)\}$  é um processo estocástico periódico.
- B Os estados dessa cadeia de Markov são transientes (ou transitórios).
- C A cadeia de Markov em questão é duplamente estocástica.
- D No regime estacionário tem-se  $\lim_{n \rightarrow \infty} P^n = \begin{bmatrix} 0,5 & 0,5 \\ 0,5 & 0,5 \end{bmatrix}$ .
- E O processo  $\{X(t)\}$  é uma cadeia de Markov irreduzível.

## RASCUNHO

## QUESTÃO 48

Um cartório distribui a um oficial de justiça, em média, 15 mandados por dia, segundo um processo de Poisson. Em suas diligências diárias, o oficial cumpre, em média, 60% dos mandados que lhe foram distribuídos. Os mandados não cumpridos são devolvidos no mesmo dia para o cartório, e este os redistribui a um oficial plantonista. Nesse caso, a quantidade diária de mandados redistribuídos ao oficial plantonista segue um processo Poisson com taxa igual a

- A 9 mandados por dia.
- B 6 mandados por dia.
- C 3 mandados por dia.
- D 15 mandados por dia.
- E 12 mandados por dia.

## QUESTÃO 49

Considere que  $N(t)$  representa a quantidade de veículos que chegam ao estacionamento de um tribunal durante o intervalo de tempo  $t$  (em minutos). Suponha que  $N(t)$  segue um processo de renovação, de maneira que  $P[N(t) = k] = \frac{A(t) \cdot t^k}{k!}$ , em que  $A(t) > 0$  é uma constante de normalização. Nesse caso, é correto afirmar que o intervalo de tempo entre chegadas de dois veículos consecutivos segue uma distribuição

- A de Weibull, com média igual a  $A(2)$ .
- B uniforme, com média igual a  $A(1)$ .
- C exponencial, com média igual a  $A(0)$ .
- D gama, com média igual a  $\frac{1}{A(t)}$ .
- E geométrica, com variância igual a  $A(t)$ .

## QUESTÃO 50

Neste estudo, a teoria de filas foi empregada para descrever as filas que se formam nos caixas de certo supermercado. O interesse particular desse estudo é a determinação do tempo médio que cada cliente gasta em espera na fila. Dois modelos foram considerados: (i) sistema M/M/m em fila única e (ii) sistema M/M/1 em  $m$  filas paralelas e independentes.

R. Morabito e F. C. R. de Lima. Um modelo para analisar o problema de filas em caixas de supermercados: um estudo de caso. Pesquisa Operacional, vol. 20, n.º 1, jun./2000, p. 59-71 (com adaptações).

Com relação ao assunto abordado no texto acima, assinale a opção correta.

- A O modelo G/M/m descreve uma fila cuja distribuição dos intervalos de tempos entre as chegadas de dois clientes consecutivos segue uma distribuição arbitrária, não necessariamente exponencial.
- B Uma fila do tipo M/M/m é composta por  $M$  servidores e a taxa de entrada na fila é igual a  $m$ .
- C O tempo médio de espera por cliente na fila tende a ser menor no sistema M/M/1 em  $m$  filas paralelas e independentes do que no sistema M/M/m em fila única.
- D No sistema M/M/1, considera-se que há correlação linear entre os tempos de atendimentos de dois clientes consecutivos.
- E O sistema M/G/1 representa uma fila em que a taxa de chegada de clientes na fila segue uma distribuição geométrica, e cada cliente é atendido por um único servidor.

## RASCUNHO

## QUESTÃO 51

Em determinado órgão, a dinâmica da entrada e saída de documentos administrativos foi descrito conforme um processo de nascimento e morte (*birth and death process*). Quando  $N$  documentos administrativos se encontram nesse órgão, a taxa de entrada de novos documentos (nascimento) e a de saída dos documentos ora existentes no local (morte) são, respectivamente, iguais a  $10 \times N$  unidades por dia e  $15 \times N$  unidades por dia. Nesse caso, a probabilidade de haver transição de  $N$  para  $N + 1$  documentos administrativos nesse órgão é igual a

- A 0,3.
- B 0,4.
- C 0,5.
- D 0,6.
- E 0,2.

## QUESTÃO 52

Um indicador da qualidade —  $X(t)$  — da gestão do Poder Judiciário segue um processo gaussiano com média nula e variância igual a  $t$ . Em determinado instante  $t$ , um analista observou que  $X(t) = 2$ . Com base nesse resultado, esse analista deseja calcular a média e a variância no instante  $\frac{t}{2}$ . Nesse caso, é correto afirmar que os momentos condicionais  $E[X(\frac{t}{2}) | X(t) = 2]$  e  $\text{Var}[X(\frac{t}{2}) | X(t) = 2]$  são, respectivamente, iguais a

- A  $2$  e  $\frac{t}{2}$ .
- B  $2$  e  $t$ .
- C  $0$  e  $t$ .
- D  $1$  e  $\frac{t}{4}$ .
- E  $1$  e  $\frac{t}{2}$ .

## QUESTÃO 53

O vetor  $(X, Y, Z)^T$  é um vetor aleatório que segue uma distribuição normal multivariada em que o vetor  $(1, 2, 3)^T$  é o vetor de médias e

a matriz de covariância é  $\begin{bmatrix} 4 & 0 & 1 \\ 0 & 9 & 1 \\ 1 & 1 & 4 \end{bmatrix}$ . Considerando essas informações,

assinale a opção correta.

- A A variável  $X$  se correlaciona com a variável  $Z$ , e esta se correlaciona com a variável  $Y$ . Logo, há dependência entre as variáveis  $X$  e  $Y$ .
- B A variância da soma  $X + Y + Z$  é igual a 21.
- C As distribuições marginais não são necessariamente gaussianas.
- D A correlação linear entre  $X$  e  $Z$  é igual a 0,5.
- E O vetor aleatório  $(X, Y)^T$  segue uma distribuição normal bivariada em que a matriz  $\begin{bmatrix} 6 & 1 \\ 1 & 11 \end{bmatrix}$  é a matriz de covariância.

## RASCUNHO

## QUESTÃO 54

Uma análise de componentes principais considerou 20 variáveis. Com base na matriz de covariância entre essas variáveis, observou-se que os cinco maiores autovalores foram iguais a 6, 4, 3, 2 e 1. Considerando esses resultados, assinale a opção correspondente ao percentual de variação explicada por esses cinco maiores autovalores.

- A 80%
- B 40%
- C 50%
- D 60%
- E 70%

## QUESTÃO 55

Um analista efetuou uma pesquisa sobre o perfil do menor infrator. Para cada menor observado na amostra, foram observadas 15 medidas supostamente gaussianas. O analista deseja classificar as unidades amostrais em grupos, de modo que as pessoas que pertencem a um mesmo grupo tenham, estatisticamente, um tipo de similaridade com base nas 15 medidas consideradas. Com base nessas informações, é correto afirmar que a técnica multivariada apropriada para a finalidade desejada pelo analista é a análise

- A de conglomerados.
- B fatorial.
- C de correspondência.
- D de componentes principais.
- E discriminante.

## QUESTÃO 56

$lag\ h$	FAC( $h$ )	FACP( $h$ )	FACOV( $h$ )
0	---	---	5
1	0,5	---	---
2	0,2	-0,1	---
3	0,1	0,05	---
4	0,05	-0,05	---
5	-0,05	0,01	---

A tabela acima mostra alguns valores da função de autocorrelação amostral (FAC), da função de autocorrelação parcial amostral (FACP) e da função de autocovariância amostral (FACOV) referentes a um estudo da evolução temporal da quantidade mensal de mercadorias em determinado terminal de cargas. Com base nessas informações, assinale a opção correta.

- A O valor do semi-variograma empírico no  $lag\ 0$  é igual a 1.
- B A variância do processo é superior a 7.
- C A estimativa da função de autocovariância no  $lag\ 1$  se encontra no intervalo  $[0,2; 0,4]$ .
- D A função de autocovariância amostral no  $lag\ 4$  é superior a 0,1 e inferior a 0,3.
- E O valor da função de autocorrelação no  $lag\ 0$  se encontra no intervalo  $[-1,0; -0,8]$ .

**QUESTÃO 57**

Na série temporal  $\{X(t)\}$ , deseja-se efetuar o teste de hipóteses:  $H_0 : X(t) = X(t-1) + a(t)$  versus  $H_A : X(t) = \phi X(t-1) + a(t)$ , em que  $\phi$  é tal que  $|\phi| < 1$  e  $a(t)$  representa o choque aleatório. Nesse caso, o teste apropriado para essa situação é o teste

- A de McLeod-Li.
- B de Dickey-Fuller.
- C de Ljung-Box.
- D  $t$  de Student.
- E dos sinais.

**QUESTÃO 58**

número de coeficientes de médias móveis	número de coeficientes autorregressivos			
	0	1	2	3
0	6,0	5,5	5,1	5,6
1	4,9	4,4	4,0	5,5
2	4,5	4,3	4,2	5,5
3	4,7	4,9	5,0	6,5

A tabela acima mostra os valores do AIC (Akaike *Information Criterion*) correspondentes aos modelos ARMA( $p, q$ ), em que  $0 \leq p \leq 3$  e  $0 \leq q \leq 3$ . Nesse caso, o modelo sugerido pelo critério de informação de Akaike é o

- A ARMA(2, 1).
- B ARMA(2, 2).
- C ARMA(3, 3).
- D ARMA(0, 0).
- E ARMA(1, 2).

**QUESTÃO 59**

Um modelo de séries temporais tem a forma  $(1 - 0,5B)Y_t = (1 - 0,5B)a_t$ , em que  $a_t$  representa um choque aleatório no instante  $t$ ,  $B$  é o operador de atraso (*backward shift operator*) e  $Y_t = (1 - B^6)X_t$ . Nesse caso, é correto afirmar que o processo  $Y_t$

- A não é estacionário.
- B é ruído branco.
- C segue uma tendência linear na forma  $\alpha t + \beta$ .
- D é estacionário.
- E segue um processo SARIMA (1, 0, 1)  $\times$  (0, 6, 0).

**QUESTÃO 60**

Para  $-\pi \leq \omega \leq \pi$ , a densidade espectral  $f(\omega)$  de um processo AR(1), na forma  $X_t = 0,5X_{t-1} + a_t$ , em que  $a_t$  representa um choque aleatório com desvio padrão igual a 1, é

- A  $f(\omega) = \left( \frac{1}{\pi} \sum_{h=0}^{+\infty} 0,5^h \cos^h(\omega) \right) - \frac{1}{2\pi}$ .
- B  $f(\omega) = \left( \frac{1}{\pi} \sum_{h=0}^{+\infty} 0,5^h \right) - \frac{1}{2\pi}$ .
- C  $f(\omega) = \left( \frac{1}{\pi} \sum_{h=0}^{+\infty} \cos(\omega h) \right) - \frac{1}{2\pi}$ .
- D  $f(\omega) = \left( \frac{1}{\pi} \sum_{h=0}^{+\infty} 0,5^h \cos^2(\omega h) \right) - \frac{1}{2\pi}$ .
- E  $f(\omega) = \left( \frac{1}{\pi} \sum_{h=0}^{+\infty} 0,5^h \cos(\omega h) \right) - \frac{1}{2\pi}$ .

**RASCUNHO**



**cespeUnB**

Centro de Seleção e de Promoção de Eventos