

## CONHECIMENTOS ESPECÍFICOS

ano de início dos processos	especificação		
	em trâmite	para parecer	juulgado
2010	200	30	600
2011	240	30	580
2012	260	50	700

Considerando os dados da tabela acima, que mostra a quantidade e situação de processos, nos anos 2010, 2011 e 2012, em um tribunal, julgue os itens subsequentes.

- 51 Se, em 2011, 5 juizes atuavam no referido tribunal, então a relação juiz/processo era de, aproximadamente, 1:170.
- 52 Menos de 5% dos processos aguardam parecer.
- 53 A probabilidade de um processo ser julgado é superior a 75%.
- 54 Se determinado processo está em trâmite, a probabilidade de ele ser do ano de 2012 é superior a 30%.
- 55 A variável ano é uma variável qualitativa ordinal, uma vez que é possível definir uma ordem entre os anos.
- 56 Com relação a 2010, em 2012 ocorreu um acréscimo de mais de 200 processos.

Um estagiário deve organizar uma pilha de  $n$  processos de acordo com o valor, em reais, das sentenças e por número, em três estantes: I, II e III. O desvio padrão do valor das sentenças é R\$ 50. A estante I é para processos referentes a sentenças com valores inferiores a R\$ 500; a II, para processos com sentenças de valores entre R\$ 500 e R\$ 2.000 e a III, para processos com sentenças de valores acima de R\$ 2.000.

A respeito dessa organização de processos, julgue os itens a seguir.

- 57 Para se conhecer a variação da quantidade média de páginas dos processos basta dividir o desvio padrão pela média.
- 58 Se o estagiário escolher aleatoriamente uma amostra com 20 desses processos, então o erro padrão da média amostral do valor das sentenças será igual ou inferior a R\$ 10.
- 59 Considerando que, em média, os processos na estante I tenham 100 páginas, os da estante II, 150 páginas e os da estante III, 300 páginas, e que as probabilidades de um processo pertencer às estantes I, II ou III sejam iguais a  $\frac{4}{5}$ ,  $\frac{3}{20}$  e  $\frac{5}{100}$ , respectivamente, então a quantidade média de páginas de um processo será superior a 130.

Uma máquina de café expresso precisa ser reiniciada algumas vezes durante o dia, devido ao uso excessivo. A tabela abaixo mostra a distribuição de probabilidade conjunta do número de vezes que ela é reiniciada na parte da manhã (M) e na parte da tarde (T).

M\T	0	1	2
0	0,10	0,10	0,30
1	0,04	0,06	0,12
2	0,06	0,10	$x$

Considerando essa tabela, julgue os próximos itens.

- 60 O número médio de vezes que a máquina é reiniciada na parte da manhã é inferior a 1.
- 61 A covariância entre M e T é positiva.
- 62 A probabilidade da máquina ser reiniciada duas vezes à tarde, considerando que ela tenha sido reiniciada uma vez pela manhã, é superior a 0,6.
- 63 A probabilidade da máquina ser reiniciada três vezes em um mesmo dia é inferior a 0,1.
- 64 O valor de  $x$  é inferior a 0,15.
- 65 O número de vezes que a máquina é reiniciada na parte da tarde depende do número de vezes que ela é reiniciada pela manhã.

RASCUNHO

$$f(t) = \begin{cases} \frac{2}{3}t & , \text{se } 0 \leq t < 1 \\ -\frac{t}{4} + \frac{5}{6} & , \text{se } 1 \leq t \leq 3 \\ 0 & , \text{caso contrário} \end{cases}$$

A função  $f(t)$  mostrada acima corresponde à função densidade de probabilidade do tempo gasto ( $t$ , em meses) para se analisar um processo em determinada vara civil.

Com relação essa função, julgue os itens seguintes.

- 66 Cada processo demora, em média, pelo menos 1,5 mês para ser analisado.
- 67 O tempo médio para se avaliar sequencialmente 10 processos é superior a 1 ano.
- 68 A curtose da distribuição é maior que 5.
- 69 A probabilidade de um processo, escolhido ao acaso, demorar mais de dois meses para ser analisado é superior a 0,4.
- 70 A probabilidade de um processo, escolhido ao acaso, demorar menos de três meses para ser analisado é superior a 0,99.

Com relação a inferência estatística, julgue os itens a seguir.

- 71 Considerando-se que, para comparar a proporção de sucessos em uma variável aleatória de Bernoulli com determinada constante, tenha sido empregado um teste de hipóteses em uma amostra de tamanho 25, é correto afirmar que o limite superior da variância da estatística do teste é 0,01.
- 72 A região crítica definida pela razão entre as verossimilhanças sob a hipótese nula e a hipótese alternativa é aquela que minimiza a soma das probabilidades dos erros do tipo I e do tipo II.
- 73 Sob a perspectiva bayesiana, para se estimar a média de uma distribuição normal com variância conhecida, com base na distribuição *a priori* não informativa de Jeffrey, a moda da distribuição *a posteriori* deve coincidir com o estimador clássico de máxima verossimilhança.
- 74 Considere que  $X$  seja uma variável aleatória com média  $\mu$  e variância  $\sigma^2$ , que  $\{X_1, X_2, \dots, X_n\}$  represente uma amostra aleatória simples de  $X$  de tamanho  $n$ , e que  $\bar{X}$  represente o estimador de momentos da média  $\mu$ . Nesse caso, o estimador  $\hat{\sigma}^2$ , para a variância de  $X$ , obtido pelo método dos momentos para a referida amostra, é corretamente representado por  $\hat{\sigma}^2 = \frac{1}{n} \left( \sum_{i=1}^n X_i^2 - n\bar{X}^2 \right)$ .
- 75 O estimador de máxima verossimilhança para um parâmetro  $\theta$  é obtido, independentemente da distribuição dos dados, igualando-se a zero a derivada do logaritmo da função de verossimilhança.
- 76 O nível de significância de determinado intervalo de confiança corresponde à probabilidade de o verdadeiro valor do parâmetro estar contido nesse intervalo.
- 77 Entre todos os intervalos que possuem o mesmo nível de credibilidade, o intervalo HPD (*highest probability density*) é o que proporciona a maior amplitude possível.
- 78 Se para um intervalo de confiança conservativo, cuja amplitude é inferior a 0,196, a proporção de sucessos em lançamentos de Bernoulli é de 95%, é correto afirmar que o tamanho da amostra é superior a 25.
- 79 Em um teste de hipóteses para a média  $\mu$ , cuja hipótese nula é  $\mu = 0$ , o poder do teste depende apenas do nível de significância.

RASCUNHO

Acerca dos métodos não paramétricos, julgue o item abaixo.

RASCUNHO

- 80 Considerando-se a tabela abaixo, é correto afirmar que, para a homogeneidade marginal aplicada nessa tabela, sem a correção de Yates, a estatística do teste de McNemar é inferior a 18,00 e tem distribuição qui-quadrado com 1 grau de liberdade.

	após (+)	após (-)	total
antes (+)	46	19	65
depois (+)	1	34	35
total	47	53	100

No que se refere aos estimadores dos parâmetros dos modelos de regressão, julgue os itens seguintes.

- 81 Um elevado coeficiente de determinação ( $R^2 \geq 0,70$ ) referente a um modelo de regressão linear para uma amostra não muito grande, não implica, necessariamente, que a reta de regressão passe próxima a todos os pontos amostrados e que o modelo esteja bem ajustado.
- 82 Se a inclinação da reta de regressão com relação ao eixo das abscissas for igual a  $45^\circ$ , então, para cada unidade acrescentada na variável independente ( $X$ ), ocorre um acréscimo de duas unidades na variável dependente ( $Y$ ).
- 83 Se  $\hat{\beta}_0$  e  $\hat{\beta}_1$  forem, respectivamente, estimadores do intercepto e da inclinação de uma reta de regressão linear simples, então a distribuição condicional  $Y|\hat{\beta}_0, \hat{\beta}_1$  não depende de  $\beta_0$  e  $\beta_1$ .
- 84 Se  $\hat{\beta}_0$  e  $\hat{\beta}_1$  forem, respectivamente, estimadores do intercepto e da inclinação da reta de regressão, é correto afirmar que  $\lim_{n \rightarrow \infty} P(|\hat{\beta}_1 - \beta_1| > \epsilon) > \lim_{n \rightarrow \infty} P(|\hat{\beta}_0 - \beta_0| > \epsilon)$ .

Julgue o item abaixo acerca dos parâmetros dos modelos de regressão.

- 85 Considere que, para um modelo de regressão linear, um vetor  $x^*$  de variáveis independentes tenha sido observado com o objetivo de se prever a resposta. Considere, ainda, que um *leverage* seja igual a 0,45, que o quadrado médio do resíduo seja 120 e que o percentil de ordem 97,5% da distribuição t-Student correspondente seja 2,04. Nesse caso, a amplitude do intervalo de predição é superior a 360.

A respeito dos modelos de análise de variância, julgue os itens subsequentes.

RASCUNHO

**86** Em um modelo de análise de variância (ANOVA) para um fator fixo com três níveis, o valor absoluto do efeito de um dos níveis é igual ao valor absoluto da soma dos efeitos dos demais níveis.

**87** Considere que, em um modelo de análise de variância (ANOVA) com um fator fixo em quatro níveis, ajustado em uma amostra de tamanho 25, tenha sido observada uma estatística F igual a 3,5. Nesse caso, se a soma de quadrados totais (SQT) for igual a 31,5, então a soma de quadrados associada ao fator é igual a quatro vezes o valor da soma de quadrados do resíduo.

Julgue os itens a seguir, referentes à análise dos resíduos e da qualidade de ajuste dos modelos de regressão.

**88** Considere que um pesquisador, ao ajustar um modelo de regressão de  $Y$  explicado pelas variáveis  $X_1$  e  $X_2$ , tenha observado que, no gráfico dos resíduos do modelo de  $Y$  explicado por  $X_2$  contra os resíduos de  $X_1$  explicado por  $X_2$ , havia uma relação quadrática. Nesse caso, o diagnóstico indica que a relação entre a variável resposta  $Y$  e  $X_2$  é quadrática.

**89** Considere que um modelo de regressão com intercepto e três variáveis regressoras tenha sido ajustado com base em uma amostra de tamanho 32 e que a análise da qualidade do ajuste não tenha detectado valor *outlier* nas variáveis independentes. Nessa situação, o menor valor do resíduo padronizado deletado que define observações influentes na amostra é  $\sqrt{3}$ .

Com relação às técnicas de amostragem, julgue os itens seguintes.

**90** Para se aplicar a amostragem sistemática faz-se necessário o conhecimento do tamanho populacional.

**91** Considere que, em uma amostragem aleatória simples sem reposição, a variância da média amostral observada seja  $V(\bar{X}) = \frac{\sigma^2}{N}$ , em que  $N$  é o tamanho da população e  $\sigma^2$  é a variância populacional da variável observada. Nesse caso, o tamanho amostral é igual à metade do tamanho populacional.

**92** A ordenação da variância  $V(\cdot)$  da média amostral, realizando-se o cálculo dessa média da forma usual  $(\bar{Y})$ , mediante o estimador de regressão  $(\bar{Y}_{\text{Reg}})$  e o estimador de razão  $(\bar{Y}_{\text{R}})$ , é dada por  $V(\bar{Y}_{\text{Reg}}) \leq V(\bar{Y}) \leq V(\bar{Y}_{\text{R}})$ , em que o coeficiente  $b_0$  do estimador de regressão é aquele que minimiza a variância no estimador tipo regressão.

**93** Os estratos, diferentemente dos conglomerados, possuem grande variabilidade interna.

RASCUNHO

Um método popular para a obtenção de números pseudoaleatórios (NPAs) é o gerador de congruência linear, no qual os NPAs são construídos da seguinte maneira:

- 1) escolha um número natural  $X_0$ ;
- 2) para  $i = 1, 2, \dots$ , faça  $X_i = (aX_{i-1} + c) \bmod (m)$ , em que  $a$ ,  $c$ , e  $m$  representam constantes inteiras adequadas.
- 3) Para  $i = 1, 2, \dots$ , faça  $U_i = X_i / m$ .

Com relação a esse método, julgue os itens a seguir.

- 94** Se  $X_0 = 5$  e  $X_i = (3X_{i-1} + 2) \bmod (7)$ , então, para  $i = 91$  e  $i = 360$ , os NPAs correspondentes são  $U_{91} = 3/7$  e  $U_{360} = 1/7$ .
- 95** O gerador em que  $X_0 = 4$ ,  $X_i = (aX_{i-1} + 3) \bmod (10)$  e  $a$  é um número par é um gerador de período inteiro.
- 96** O gerador em que  $X_0 = 1$  e  $X_i = (5X_{i-1} + 3) \bmod (8)$  é um gerador de período inteiro.

O Teorema de Transformação Inversa afirma que se  $U$  for uma variável aleatória uniformemente distribuída no intervalo  $[0, 1]$  e se  $F$  for uma função de distribuição (acumulada) de uma variável aleatória contínua, então  $X = F^{-1}(U)$  tem função de distribuição  $F$ . Considerando essa informação e a função acumulada da distribuição logística  $F(x) = \frac{1}{1 + e^{-(x-a)/\beta}}$ , em que  $\beta > 0$ , julgue os

itens seguintes nos quais os números  $u_i$  representam realizações da variável  $U$  acima.

- 97** Os valores da forma  $-3\ln(1 - \sqrt{u_i})$  representam realizações de uma variável aleatória com função de distribuição dada por  $F(x) = (1 - e^{-3x})^2$ , para  $x \geq 0$ .
- 98** Os valores  $x_i = -\frac{1}{5}\ln\left[\frac{1}{u_i} - 1\right]$  representam realizações de uma variável aleatória logística.
- 99** Os valores  $x_i = \frac{1}{u_i^{1/3}} - 1$  representam realizações da variável aleatória em que  $f(x) = \frac{3}{(x+1)^4}$ , para  $x \geq 0$ , é função densidade de probabilidade.

A quantidade de chamadas que uma central telefônica recebe por hora é modelado por uma distribuição de Poisson, com parâmetro  $\lambda = 12$  chamadas por hora. Nesse caso, a probabilidade de a central telefônica receber

- 100** menos de três chamadas em uma hora é igual a  $85e^{-12}$ .
- 101** exatamente duas chamadas em 20 minutos é igual a  $\frac{12^6}{6!}e^{-12}$ .

RASCUNHO

A população de um país é dividida em classes alta (A), média (M) e baixa (B). Um estudo estatístico mostra que, atualmente, 10% da população pertence à classe A, 40% à classe M e 50% à classe B. Considera-se um modelo simplificado para as mudanças de classes, na forma de uma cadeia de Markov, em que as mudanças de uma geração para a próxima acontecem de acordo com a seguinte matriz de transição:

$$M = \begin{matrix} & \begin{matrix} A & M & B \end{matrix} \\ \begin{matrix} A \\ M \\ B \end{matrix} & \begin{pmatrix} 0,9 & 0,1 & 0 \\ 0,1 & 0,6 & 0,3 \\ 0 & 0,2 & 0,8 \end{pmatrix} \end{matrix}$$

Assim, por exemplo, as probabilidades dos filhos de uma família da classe M pertencerem às classes A, M ou B são iguais a 10%, 60% e 30%, respectivamente.

Com base nessa situação hipotética, julgue os itens subsequentes.

- 102** De acordo com o modelo apresentado, os netos de uma família da classe B têm 70% de probabilidade de pertencer a essa mesma classe e 3% de probabilidade de pertencer à classe A.
- 103** Na próxima geração, 13% da população pertencerá à classe A, 35% à classe M e 52% à classe B.
- 104** Na hipótese de que o modelo tenha sido válido para a formação da geração atual, então as classes A, M e B na geração anterior eram formadas por 5%, 30% e 65% da população, respectivamente.
- 105** Se o modelo descrito valer por tempo indeterminado, então as proporções das classes A, M e B tenderão para as probabilidades estacionárias  $2/7$ ,  $2/7$  e  $3/7$ , respectivamente.

A densidade da distribuição normal bivariada pode ser escrita na forma  $f(x_1, x_2) =$

$$\frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{1/2}} \times \exp\left(-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right),$$

em que  $\mu_i$  é a expectativa de  $X_i$ ,  $\sigma_i^2 = \sigma_{i,i}$  é a variância de  $X_i$  para  $i = 1$  e  $2$  e  $\rho = \frac{\sigma_{1,2}}{\sigma_1 \times \sigma_2}$  é o coeficiente de correlação linear entre

$X_1$  e  $X_2$ .

Considerando essas informações e a função de densidade bivariada

$$f(x, y) = \frac{3}{4\pi\sqrt{2}} \exp\left[-\frac{9}{16}\left(x^2 - \frac{2}{3}xy + y^2\right)\right],$$

para  $x$  e  $y$  reais, julgue os

próximos itens.

- 106** A densidade é simétrica em relação aos eixos  $x$  e  $y$ .
- 107** As densidades marginais seguem uma distribuição  $N(0, 1)$ .
- 108** A função densidade de probabilidade de  $X$ , dado que  $Y = y$  é expressa por  $\frac{3}{4\sqrt{\pi}} \exp\left[-\frac{1}{16}(3x - y)^2\right]$ .
- 109** O coeficiente de correlação linear entre  $X$  e  $Y$  é igual a  $\frac{1}{4}$ .

Um procedimento básico da análise fatorial é o modelo ortogonal, que visa representar um vetor aleatório  $X = (X_1, \dots, X_p)^T$  na forma  $X - \mu = LF + \varepsilon$ , em que  $\mu = E(X)$ ,  $F = (F_1, \dots, F_m)^T$ , com  $m < p$ , é o vetor de fatores comuns,  $L$  é a  $p \times m$ -matriz de cargas fatoriais e  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_p)^T$  é o vetor de erros (o T sobrescrito ao vetor indica o vetor transposto).

Considerando essa informação, julgue os itens a seguir.

- 110** No modelo ortogonal, os seguintes pressupostos adicionais sobre os vetores aleatórios  $F$  e  $\varepsilon$  são impostos:  $F$  e  $\varepsilon$  são independentes; ambos têm expectância zero, e as covariâncias de  $F$  e  $\varepsilon$  são a matriz unidade e uma matriz positiva definida qualquer, respectivamente.
- 111** Considere que, para um vetor aleatório  $X$  de 5 componentes, o método de componente principal tenha sido aplicado, com dois fatores comuns,  $F_1$  e  $F_2$ , com as seguintes cargas estimadas:

variável	$F_1$	$F_2$
$X_1$	0,5	0,8
$X_2$	0,7	-0,5
$X_3$	0,7	0,7
$X_4$	0,9	-0,1
$X_5$	0,8	-0,5

De acordo com a tabela, é correto afirmar que as communalidades estimadas são 0,89; 0,74; 0,98; 0,82 e 0,89.

Na literatura de séries temporais, para se detectar uma tendência são conhecidos, entre outros, o teste de sinais de Cox-Stuart, o teste com base no coeficiente de correlação de Spearman e o *run test* de Wald-Wolfowitz.

Acerca desse assunto e considerando que  $Z_1, \dots, Z_N$  seja uma série temporal, julgue os itens seguintes.

- 112** Em um teste bilateral de sinais, se o nível de significância for de 5%, a hipótese de ausência de tendência é aceita para a série temporal 2, 4, 3, 5, 6, 3, 5, 4, 6, 5.

- 113** O teste com base no coeficiente de correlação de Spearman emprega a estatística  $T_3 = \sum_{i=1}^N (R_i - i)^2$ , em que  $R_i$  é o posto de  $Z_i$ .

No caso de observações empatadas, são usados os postos médios. A hipótese de que não exista tendência será rejeitada apenas se  $T_3$  for grande.

- 114** Para a série temporal 2, 8, 7, 9, 12, 6, 11, 10, o valor da estatística  $T_3$  do teste com base no coeficiente de correlação de Spearman é 34.

- 115** Se a serie temporal for 7, 8, 10, 12, 11, 9, 15, 16, 18, 20, então a quantidade de *runs* obtidos será igual a 4.

Acerca de planejamento e gestão estratégica no âmbito do Poder Judiciário, regidos pela Resolução n.º 70/2009, julgue o próximo item.

- 116** É assegurado ao serventuário que tenha sido indicado pela respectiva entidade de classe participar da elaboração e da execução da proposta orçamentária e do planejamento estratégico do tribunal correspondente.

Com relação aos princípios do Sistema de Estatística do Poder Judiciário (SIESPJ), regidos pela Resolução n.º 76/2009, julgue os itens subsecutivos.

- 117** É atribuição indelegável da presidência de cada tribunal a função de gerar e transmitir os dados estatísticos.
- 118** Cabe ao CNJ verificar o conteúdo dos dados estatísticos transmitidos eletronicamente pelos tribunais.
- 119** Os tribunais regionais eleitorais e os tribunais de justiça militar integram o SIESPJ.
- 120** Cabe à Comissão de Estatística e Gestão Estratégica, composta por três conselheiros eleitos pelo plenário do CNJ, requisitar o auxílio temporário de magistrados, que continuarão a exercer suas funções regulares nos órgãos de origem.

RASCUNHO

**PROVA DISCURSIVA**

- Nesta prova, faça o que se pede, usando, caso deseje, os espaços para rascunho indicados no presente caderno. Em seguida, transcreva os textos para o **CADERNO DE TEXTOS DEFINITIVOS DA PROVA DISCURSIVA**, nos locais apropriados, pois **não será avaliado fragmento de texto escrito em local indevido**.
- Tanto na questão quanto na dissertação, qualquer fragmento de texto além da extensão máxima de linhas disponibilizadas será desconsiderado. Será também desconsiderado o texto que não for escrito na folha de texto definitivo correspondente.
- No caderno de textos definitivos, identifique-se apenas no cabeçalho da primeira página, pois não será avaliado texto que tenha qualquer assinatura ou marca identificadora fora do local apropriado.
- Ao domínio do conteúdo serão atribuídos, **para a questão**, até **30,00 pontos**, dos quais até **1,50 ponto** será atribuído ao quesito apresentação e estrutura textual (legibilidade, respeito às margens e indicação de parágrafos) — e, **para a dissertação**, até **40,00 pontos**, dos quais até **2,00 pontos** serão atribuídos ao referido quesito.

**QUESTÃO**

Considere que  $X$  seja uma variável aleatória com distribuição exponencial com parâmetro  $\lambda = \frac{1}{4}$ . Explique, sucintamente, a utilidade dessa distribuição, calcule sua média e mediana com base no valor do parâmetro fornecido e discuta, se for o caso, por que há diferença entre essas duas medidas de posição. Considere:  $\ln 1 = 0$ ;  $\ln 0,5 = -0,69$ ; e  $\ln 0 = -\infty$ .

**RASCUNHO – QUESTÃO**

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	



**DISSERTAÇÃO**

A análise de variância é uma técnica estatística que permite tanto verificar a qualidade de ajuste em modelos de regressão linear quanto comparar as médias de uma variável quantitativa obtidas em diferentes tratamentos. Considerando a importância dessa técnica, redija um texto dissertativo acerca das relações entre os modelos de análise de variância e os modelos de regressão linear. Ao elaborar seu texto, explique, necessariamente:

- ▶ como a análise de variância pode ser usada para avaliar a qualidade de ajuste de um modelo de regressão; **[valor: 13,00 pontos]**
- ▶ como um modelo de análise de variância pode ser escrito como uma regressão linear; **[valor: 12,50 pontos]**
- ▶ como os resultados do modelo de regressão podem ser usados para construir uma tabela de análise de variância. **[valor: 12,50 pontos]**

**RASCUNHO – DISSERTAÇÃO**

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	

