CONHECIMENTOS ESPECÍFICOS

RASCUNHO

quantidade	São Paulo ($j = 1$)	Rio de Janeiro ($j = 2$)	Minas Gerais ($j = 3$)	Rio Grande do Sul ($j = 4$)	total
casos novos (X, em milhões)	5	2	1	2	18
casos pendentes (Y, em milhões)	16	8	3	2	48
processos baixados (Z, em milhões)	5	2	2	2	18
sentenças e decisões (W, em milhões)	4	3	1	1	16

CNJ. **Justiça em números 2010**, Departamento de Pesquisas Judiciárias. Agosto/2011 (com adaptações).

O quadro acima mostra uma síntese da movimentação processual dos tribunais de justiça dos estados de São Paulo, Rio de Janeiro, Minas Gerais, Rio Grande do Sul e do total da justiça estadual no Brasil em 2010. Considere que o estoque de processos em andamento no estado j (E_j), no final de 2010, seja um indicador que se define como $E_j = X_j + Y_j - Z_j - W_j$, em que $j = 1, 2, ..., 27$; X_j representa o número de casos novos registrados em 2010 no estado j; Y_j seja a quantidade de casos pendentes no estado j (i.e., casos anteriores que não foram solucionados até o final de 2010); Z_j denota o total de processos baixados (arquivados) no estado j durante 2010 e W_j seja o número de sentenças e decisões proferidas no estado j até o final de 2010. Considere, por fim, que, para todos os efeitos, o Distrito Federal seja um estado. Com base nessas informações e no quadro acima, julgue os itens que se seguem.

51 Considerando-se apenas os dados relativos aos estados de São Paulo, Rio de Janeiro, Minas Gerais e Rio Grande do Sul quanto à dispersão entre duas variáveis, é correto afirmar que a covariância entre Z e W é superior a 1 e inferior a 2.

52 O quadro apresentado é uma tabela de contingência que mostra o cruzamento entre uma variável qualitativa nominal com 4 níveis de resposta (estados) e outra variável qualitativa com quatro níveis de resposta (casos novos, pendentes, baixados e resolvidos).

53 Considerando-se que $\bar{X}, \bar{Y}, \bar{Z}$ e $\bar{W}$ representem, respectivamente, as médias aritméticas das variáveis X, Y, Z e W, então $\bar{X} + \bar{Y} - \bar{Z} - \bar{W}$ representa a média aritmética da distribuição dos estoques de processos observados nos tribunais estaduais.

54 O estoque de processos em andamento no estado de São Paulo no final de 2010 representou 37,5% do total dos estoques de processos em andamento nos tribunais estaduais do país nesse mesmo período.

55 Considerando-se que Var(E) seja a variância da distribuição dos estoques de processos existentes nos tribunais estaduais, então Var(E) = Var(X) + Var(Y) − Var(Z) − Var(W).

Nas estatísticas do Poder Judiciário, a taxa de congestionamento (X), que consiste em um indicador que permite medir a efetividade da movimentação processual de um tribunal, é uma variável aleatória contínua com função de densidade $f(x)$ expressa por

$$f(x) = \begin{cases} \beta x^8 (1-x)^2, & \text{se } x \in [0,1] \\ 0, & \text{se } x \notin [0,1] \end{cases}, \text{ em que } \beta \text{ é uma constante real.}$$

Com base nessas informações, julgue os próximos itens.

56 A distribuição da taxa de congestionamento X é simétrica em torno de 0,5.

57 A variável aleatória X segue uma distribuição especial denominada de Beta, que é, *a priori*, conjugada das distribuições geométrica e de Bernoulli.

58 O valor de β é superior a 450 e inferior a 500.

59 A média da taxa de congestionamento é inferior a $0{,}10 \times \beta$.

Considerando A e B dois eventos aleatórios, com probabilidades $P(A) = 0{,}4$ e $P(B) = 0{,}1$, e o evento complementar B^c, julgue os itens seguintes, relativos a probabilidade condicional.

60 Em face dos dados apresentados, é correto afirmar que $P(A|B) < P(A \cap B)$.

61 Considerando-se que A e B sejam eventos mutuamente excludentes, é correto afirmar que $P(A|B^c) = 0$.

62 Se A e B forem eventos independentes, então $P(A|B^c) = P(A|B) = 0{,}4$.

O núcleo de assistência jurídica de um fórum que presta assistência jurídica gratuita a pessoas carentes recebe diariamente X casos novos, conforme uma distribuição condicional na forma $P(X = k|Y = b) = \dfrac{e^{-b} b^k}{k!}$, em que $k = 0, 1, 2, ..., b > 0$ e Y segue uma distribuição exponencial com função de densidade $f(y) = 2e^{-2y}$, em que $y > 0$.

Considerando essa situação hipotética, julgue os itens subsequentes.

63 O coeficiente de correlação linear entre as variáveis X e Y é negativo.

64 O valor b representa a taxa diária de chegada de casos novos, sendo essa taxa o valor esperado da quantidade diária X condicionada ao evento $Y = b$, ou seja, $E(X|Y = b) = b$.

65 Se o valor b for desconhecido, a quantidade média diária de casos novos — $E(X)$ — também será desconhecida.

66 Espera-se que a probabilidade de não chegar casos novos em determinado dia seja superior a 0,6.

67 O desvio padrão da variável aleatória Y é igual a 2.

68 Considerando-se que, em certo dia, o núcleo tenha recebido 5 casos novos, é correto afirmar que a distribuição condicional $Y|X = 5$ segue a distribuição *gamma* na forma $f(y|5) = \dfrac{3^6 y^5 e^{-3y}}{5!}$.

RASCUNHO

A produtividade do magistrado (Z) é um indicador que permite medir a celeridade dos processos judiciais. Ela é definida como uma razão na forma $Z = \frac{X}{N}$, em que X representa o total anual de processos julgados pelos magistrados de certo tribunal e N, uma constante, representa o total de magistrados existentes nesse tribunal. Embora X seja uma variável aleatória discreta, ela pode ser aproximada por uma distribuição normal com média μ e desvio padrão σ.

Com base nessa situação hipotética, julgue os itens a seguir.

69 O indicador Z representa o número médio anual de processos julgados por um magistrado no referido tribunal.

70 A produtividade do magistrado é uma variável aleatória que segue, aproximadamente, uma distribuição normal com média $\frac{\mu}{N}$ e desvio padrão $\frac{\sigma}{\sqrt{N}}$.

Considerando que X seja uma variável aleatória contínua, tal que $E(X) = 1$ e $E(X^2) = 4$, julgue os itens seguintes.

71 $\mathrm{Var}(X) = 2$.

72 O coeficiente X de variação é igual ou superior a 2.

73 $P(X > 4) \leq \frac{1}{4}$.

Considerando $X_1, X_2, ..., X_n$ uma amostra aleatória simples retirada de uma distribuição com média μ e desvio padrão σ, e a estatística $T_n = \sum_{k=1}^{n} \frac{X_k}{k}$, julgue os itens que se seguem.

74 De acordo com o teorema central do limite, a sequência de estatísticas T_n converge em distribuição para uma distribuição normal.

75 Se $A = \sum_{k=1}^{n} \frac{1}{k}$, então $\frac{T_n}{A}$ é um estimador não tendencioso (ou não viciado) da média amostral.

Com o propósito de produzir inferências acerca da proporção populacional (p) de pessoas satisfeitas com determinado serviço oferecido pelo judiciário brasileiro, foi considerada uma pequena amostra de 30 pessoas, tendo cada uma de responder 1, para o caso de estar satisfeita, ou 0, para o caso de não estar satisfeita. Os dados da amostra estão registrados a seguir.

0 0 0 1 1 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 1 1 0 0 0 1 0 0 0 1

Com base nessas informações, julgue os itens seguintes.

76 A estatística do teste para verificar se p é igual a 0,5 possui 29 graus de liberdade.

77 Caso o p-valor do teste H_0: $p = 0{,}5$ *versus* H_1: $p \neq 0{,}5$ seja igual a 0,0295, então, se a hipótese alternativa fosse alterada para H_1: $p < 0{,}5$, o teste seria significativo ao nível de significância de 2%.

78 A estimativa pontual para o parâmetro p é inferior a 0,20.

79 A variância amostral para a proporção de pessoas satisfeitas e não satisfeitas é a mesma.

80 Considerando que Z represente a distribuição normal padrão, que $P(Z > 2) \approx 0{,}975$ e $P(Z > 1{,}645) = 0{,}95$ e que 2,51 é valor aproximado para $\sqrt{6{,}3}$, é correto afirmar que o intervalo $[a; b]$ que representa um intervalo de 95% de confiança para a proporção de pessoas não satisfeitas está contido no intervalo [0,4; 0,9].

81 Considerando-se 0,145 valor aproximado para $\sqrt{0{,}021}$, é correto afirmar que o coeficiente de variação da distribuição de zeros e uns é superior a 50%.

RASCUNHO

Segundo notícia veiculada recentemente, em rede nacional, os processos do judiciário estão demorando mais que o razoável porque os juízes têm de analisar, em média, 3 mil processos por ano. Para verificar o fato, um analista coletou a quantidade de processos de uma amostra de 10 juízes, estando os resultados dispostos a seguir (em mil processos por ano).

2 5 4 3 2 2 3 3,5 2,5 5

Com base nessas informações e considerando que μ representa a média populacional por juiz, julgue os itens subsequentes.

82 A estatística do teste para se testar se H_0: $\mu = 3$ mil possui 8 graus de liberdade, dada a necessidade de se estimarem a média e o desvio-padrão e sabendo que os dados seguem uma distribuição normal.

83 Sabendo-se que $\sum_{i=1}^{10} (x_i - \bar{x})^2 = 12{,}1$, em que x_i representa a quantidade anual de processos com o juiz i (i = 1, ..., 10) e $\bar{x}$ é a média amostral dessas quantidades, conclui-se que o erro padrão da média utilizado para o cálculo do intervalo de confiança para a média é superior a 100.

84 A estimativa pontual da média μ é superior a 3 mil.

85 A mediana dos processos é igual a 2 mil.

86 Caso fosse utilizado o estimador da média dado pela soma dos 5 primeiros registros, então esse estimador seria não viciado e consistente.

87 Se o objetivo fosse apenas verificar a veracidade da afirmação de que os juízes analisam em média 3 mil processos por ano, então a hipótese nula seria do tipo bilateral, isto é, H_0: $\mu \neq 3$ mil.

88 Para se calcular o poder do teste para a média populacional, bastaria alterar a hipótese nula para 2,5 mil, por exemplo, e depois calcular $1 - \beta$, em que β é o erro do tipo II.

RASCUNHO

Para verificar se a escolaridade dos servidores de determinado tribunal estaria relacionada à eficiência no atendimento ao público, um analista pesquisou alguns servidores, dispondo as informações obtidas na tabela a seguir.

escolaridade	eficiência		
	baixa	média	alta
fundamental	20	10	8
médio	10	30	25
superior	10	40	47

Com base nessas informações e considerando que a escolaridade de cada servidor entrevistado, apresentada na tabela, corresponda à maior escolaridade que possui, julgue os itens seguintes.

89 Para verificar se as variáveis estão associadas, pode-se utilizar o teste qui-quadrado com 4 graus de liberdade.

90 Os valores esperados na hipótese de independência da diagonal principal da tabela de dados são, respectivamente, 7, 6, 26 e 38,8.

91 Caso se pretenda fazer um teste qui-quadrado de homogeneidade no que se refere à eficiência entre os níveis de escolaridade, então a estatística do teste teria apenas 2 graus de liberdade.

92 Considere que os níveis críticos da distribuição qui-quadrado com 1 a 4 graus de liberdade sejam, respectivamente,

$P\left(\chi_1^2 > 3{,}84\right) = 0{,}05$, $P\left(\chi_2^2 > 5{,}991\right) = 0{,}05$,

$P\left(\chi_3^2 > 7{,}815\right) = 0{,}05$ e $P\left(\chi_4^2 > 9{,}488\right) = 0{,}05$,

e que $\sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} = 36{,}15$, em que O e E correspondam

às contagens observadas e esperadas. Nesse caso, é correto afirmar, com 5% de significância, que não há evidências estatísticas que permitam rejeitar a hipótese de independência.

93 Para calcular o *p*-valor da estatística qui-quadrado do respectivo problema, utilizando-se uma tabela da distribuição qui-quadrada, basta encontrar o valor mais próximo da estatística dentro da tabela, independentemente dos graus de liberdade.

94 O estudo em questão insere-se entre as restrições para o uso do teste qui-quadrado, visto que todos os valores esperados são maiores que 5.

95 Foram pesquisados mais de 200 servidores.

RASCUNHO

O administrador de uma organização, antes de promover um processo de treinamento de pessoal, fez um treinamento piloto com 10 empregados para verificar a eficácia da metodologia aplicada no treinamento. A tabela a seguir mostra a quantidade de processos resolvidos por cada um desses 10 empregados, numerados de 1 a 10, no mês anterior ao treinamento piloto e no mês seguinte.

empregado	1	2	3	4	5	6	7	8	9	10
n.° processos (antes)	5	8	4	3	2	5	7	7	6	5
n.° processos (depois)	6	8	5	6	4	6	8	9	6	6

Considerando as informações acima e que os dados da tabela seguem uma distribuição normal, julgue os itens subsequentes.

96 O teste dos sinais tende a ser mais liberal do que o teste a ser aplicado aos dados, uma vez que a magnitude do ganho da metodologia empregada não é considerada.

97 A hipótese nula é a mesma, tanto para um teste com dados pareados quanto para um teste com dados independentes.

98 Para verificar se a metodologia tem efeito, deve-se aplicar um teste para dados pareados.

99 O teste a ser aplicado possui 18 graus de liberdade.

100 Sendo o valor crítico, a 95% de confiança, igual a 2,26, os dados apresentam indícios de que a metodologia produzirá ganhos de produtividade.

Com relação aos modelos de regressão, julgue os itens subsecutivos.

101 O estimador de mínimos quadrados para um modelo de regressão linear simples para uma variável resposta IID, é não viciado e possui mínima variância.

102 Em um modelo de regressão linear simples, o coeficiente de determinação cresce à medida que a correlação entre a variável resposta e a variável regressora aumenta.

103 Suponha que um advogado pretenda estimar o valor concedido para processos de danos morais com relação à idade do proponente. Para isso, ele observou que a relação entre essas variáveis é descrita por $Y = -3.500 + 100 \cdot X$. Suponha, ainda, que com o objetivo de simplificar a interpretação do modelo, o advogado decida considerar uma nova variável, $Z = X - 35$, como regressora, criando um modelo com intercepto igual a zero. Nessa situação, é correto afirmar que a variância dos estimadores permanece inalterada.

104 Em um modelo de regressão linear, a variância associada às estimativas obtidas pelo método da máxima verossimilhança é menor que as variâncias associadas às estimativas obtidas por mínimos quadrados.

RASCUNHO

Com relação à inferência para os parâmetros de modelos de regressão linear, julgue os seguintes itens.

105 Em um modelo de regressão linear simples, com resposta Y e variável explicativa X, o valor esperado para a variável resposta no ponto $X = \bar{X}$ será igual a $\bar{Y}$.

106 Em um modelo de regressão linear simples, a média dos valores observados na variável resposta é maior que a média dos valores preditos.

107 Se o coeficiente de inclinação da reta de um modelo de regressão linear simples pode ser escrito como $\beta_1 = \sum_{i=1}^{n} k_i X_i$, em que $k_i = \dfrac{X_i - \bar{X}}{\sum \left(X_i - \bar{X}\right)^2}$, para $i = 1, \ldots, n$, então $0 < \sum k_i X_i \leq 1$.

108 Considere que um analista judiciário cometeu um equívoco na especificação de um modelo de regressão linear simples, de modo que a variável explicativa, que era categorizada, foi codificada com os valores 1 e 2 e tratada como uma variável discreta. Nesse caso, se, para corrigir o erro, o analista transformou a variável em uma *dummy*, então essa transformação alterou o coeficiente de determinação do modelo.

Julgue os próximos itens, referentes à qualidade de ajuste de um modelo de regressão.

109 Considere que em uma tabela de ANOVA para ajuste de um modelo de regressão a esperança da soma de quadrados do resíduo é igual a 15 vezes a variância da variável resposta. Nesse caso, o tamanho amostral é inferior a 20 unidades.

110 Considere que, em uma tabela de ANOVA para ajuste de um modelo de regressão, a soma de quadrados totais não corrigida pela média $(SQTN = \sum y_i^2)$ tem associado n graus de liberdade. Nesse caso, o quadrado da média da variável resposta tem associado 2 graus de liberdade.

111 Considere que, em um modelo de regressão linear simples, o valor esperado do quadrado médio do resíduo seja dado por $E[QMR] = \sigma^2 + \beta_1^2 \cdot \sum \left(X_i - \bar{X}\right)^2$, em que β_1 é a inclinação da reta e σ^2 é a variância da resposta. Nesse caso, se a estatística F da ANOVA para qualidade de ajuste fosse unitária para um modelo, a inclinação da reta de regressão seria nula.

112 Se um modelo de regressão linear simples tivesse coeficiente de determinação igual a 0,75, então, nesse modelo, a soma de quadrados do resíduo seria menor que a metade da soma de quadrados totais.

113 Considere que um conjunto de dados usado para o ajuste de um modelo de regressão linear simples tenha a variância da resposta igual a 4 vezes a variância da variável explicativa. Nesse caso, se o coeficiente de determinação for igual a 0,95, então o coeficiente de inclinação da reta de regressão será menor que 3,5.

RASCUNHO

Com relação à análise de regressão linear, julgue os itens que se seguem.

114 Em um modelo de regressão linear, se a variável explicativa e a variável resposta não se correlacionam, o coeficiente de determinação seria próximo de 0. Além disso, se o coeficiente de determinação fosse próximo de 0, as variáveis explicativa e resposta seriam independentes.

115 A homocedasticidade é a propriedade conforme a qual o resíduo de um modelo de regressão tem média 0.

116 Em um modelo linear simples, se a correlação entre os quantis do resíduo padronizado e uma amostra aleatória da normal padrão for alta, o modelo não terá intercepto.

117 Um modelo de regressão linear múltipla com duas variáveis explicativas será inequivocamente ajustado se essas variáveis forem proporcionais.

118 Suponha que um modelo de regressão linear simples seja ajustado de modo que se obtenha um coeficiente de determinação próximo de 1. Nessa situação, o modelo não pode ser utilizado para previsão da variável resposta referente a valores da variável explicativa além do intervalo observado na amostra.

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

119 Considere que determinado tribunal pretenda avaliar a proporção de habitantes de um município que foram vítimas de algum tipo de violência e que não exista um banco de dados com a identificação dos habitantes desse município. Nesse caso, a aplicação da amostragem aleatória simples não será adequada para selecionar os habitantes do município.

120 Na amostragem aleatória simples sem reposição (AASs), o tamanho amostral n é calculado por

$$n = \frac{1}{\Delta / S^2 + 1/N},$$

em que N é o tamanho da população, S^2 é a variância amostral e $\Delta = (B/z)^2$, sendo B o erro máximo de estimação e z o quantil da distribuição normal. Dessa forma, é correto afirmar que o maior tamanho amostral na AASs será menor que N.

RASCUNHO