

CONHECIMENTOS ESPECÍFICOS

Nas chamadas de suporte de uma empresa de telecomunicações, o funcionário Pedro resolve o problema do cliente em duas de cada três vezes em que é solicitado, enquanto Marcos resolve em três de cada quatro chamadas.

A partir dessa situação hipotética, julgue os itens seguintes, considerando que os funcionários sejam suficientemente experientes para que a tentativa de resolução do problema de qualquer chamada não esteja subordinada a tentativas anteriores.

- 51 Considere que, para $i = 1, 2, 3, \dots$, X_i seja a variável aleatória que assume o valor 1 quando, na i -ésima chamada, Marcos resolve o problema, e o valor 0, em caso contrário.

Se $S_n = X_1 + X_2 + \dots + X_n$, então, quase certamente, $\frac{S_n}{n}$ converge para $\frac{3}{4}$.

- 52 Se Pedro não resolver o problema de um cliente, considerando-se que nenhuma informação a respeito da tentativa é repassada a Marcos, a probabilidade de que este também não resolva o referido problema será inferior a 20%.
- 53 Se Pedro e Marcos tentarem, independentemente, resolver o mesmo problema desse cliente, a chance de ambos resolverem o problema desse cliente será inferior a 45%.
- 54 Se ambos tentarem, independentemente, resolver o mesmo problema de um cliente, então a chance de o problema ser resolvido por Pedro ou por Marcos será superior a 90%.
- 55 A probabilidade de que Marcos consiga resolver o problema do cliente em exatamente três das quatro últimas chamadas em que foi solicitado é superior a 50%.

Um vendedor de certo tipo de equipamento de telecomunicações pode visitar, em um dia, um ou dois clientes, com probabilidades de $\frac{1}{3}$ e $\frac{2}{3}$, respectivamente. De cada contato pode resultar a venda

de um equipamento por R\$ 50.000, com probabilidade de $\frac{1}{10}$,

ou nenhuma venda, com probabilidade de $\frac{9}{10}$. Considerando que

V seja a variável aleatória que indica o valor total de vendas diárias desse vendedor, em milhares de reais, julgue os itens que se seguem.

- 56 O numeral 2 é um elemento do domínio da função de probabilidade de V , e indica o fechamento de duas vendas.
- 57 O valor esperado da variável aleatória V é maior que 5 e indica, em milhares, o valor total esperado de vendas em cada dia.
- 58 Se p representar a função de probabilidade de V , então $p(0) = 0,84$.
- 59 A probabilidade de esse vendedor fechar exatamente uma venda em um dado dia é superior a 0,09.
- 60 Supondo-se que X_i seja a variável aleatória que indica o número de visitas do vendedor a clientes no i -ésimo dia do mês de novembro, que $Y_i = X_i - 1$, e que $Z = Y_1 + Y_2 + \dots + Y_{30}$, é correto afirmar que Z será uma distribuição binomial de parâmetros $n = 30$ e $p = \frac{2}{3}$.

Espaço livre

Considerando que $Y_1, Y_2, \dots, Y_n, \dots$ sejam variáveis aleatórias independentes que satisfazem $P(Y_j = j) = P(Y_j = -j) = \frac{1}{2}$ para $j = 1, 2, \dots$, julgue os itens que se seguem.

- 61 O valor esperado para a variável aleatória Y_j é nulo para todo número natural positivo j .
- 62 As variáveis aleatórias Y_j , em que $j = 1, 2, \dots$, possuem variâncias iguais.
- 63 As variáveis aleatórias $Y_1, Y_2, \dots, Y_n, \dots$ são identicamente distribuídas.
- 64 O conjunto de valores que a variável aleatória $Y_1 + Y_2 + Y_3$ pode assumir é igual a $\{-6, -4, -2, 0, 2, 4, 6\}$.

Roberto comprou, por R\$ 2.800,00, rodas de liga leve para seu carro, e, ao estacionar no *shopping*, ficou indeciso sobre onde deixar o carro, pois, caso o coloque no estacionamento público, correrá o risco de lhe roubarem as rodas, ao passo que, caso o coloque no estacionamento privado, terá de pagar R\$ 70,00, com a garantia de que eventuais prejuízos serão ressarcidos pela empresa administradora.

Considerando que p seja a probabilidade de as rodas serem roubadas no estacionamento público, que X seja a variável aleatória que representa o prejuízo, em reais, ao deixar o carro no estacionamento público, e que Y seja a variável aleatória que representa o valor, em reais, desembolsado por Roberto ao deixar o carro no estacionamento pago, julgue os itens subsequentes.

- 65 O conjunto de valores possíveis para a variável aleatória X é $X(W) = \{2.800, 70, 0\}$.
- 66 Admitindo-se que Roberto tome sua decisão escolhendo aleatoriamente entre suas opções, dado que $p = 0,05$, então a probabilidade de ele ter prejuízo de R\$ 2.800,00 será superior a 0,01.
- 67 Supondo-se que Roberto tome sua decisão escolhendo aleatoriamente entre suas opções, se $p = 0,05$, então o valor esperado para o prejuízo/valor desembolsado por Roberto será inferior a R\$ 100,00.
- 68 Se Roberto tomar sua decisão escolhendo aquela cuja variável aleatória correspondente tenha menor valor esperado, então será mais vantajoso para ele deixar o carro no estacionamento pago apenas se $p > 0,025$.
- 69 A variável aleatória Y é contínua.

Em uma cidade, as companhias A, B e C detêm, respectivamente, 40%, 35% e 25% do mercado de telecomunicações. As probabilidades de um cliente de cada uma dessas empresas estar insatisfeito com os serviços prestados são, respectivamente, 0,1, 0,15 e 0,08.

Com base nessa situação, julgue os próximos itens.

- 70 Caso um cliente, escolhido ao acaso, esteja insatisfeito com os serviços prestados, a probabilidade de ele ser cliente da empresa C será inferior a 0,15.
- 71 Em termos relativos, no que se refere à prestação de serviços, há mais clientes insatisfeitos com a empresa B do que com a empresa A.
- 72 Selecionando-se um cliente ao acaso em todo o mercado dessa cidade, a probabilidade de ele estar satisfeito com os serviços prestados pelas companhias de telecomunicações é superior a 0,8.

A demanda diária de dados do pacote de Internet de um cliente de uma operadora, em MB, é uma variável aleatória X cuja função densidade de probabilidade é expressa por:

$$f(x) = \begin{cases} \frac{2}{3}x, & \text{se } 0 \leq x < 1; \\ -\frac{x}{3} + 1, & \text{se } 1 \leq x < 3; \\ 0, & \text{se } x < 0 \text{ ou } x > 3. \end{cases}$$

Com base nessas informações, julgue os itens seguintes.

- 73 Em 30 dias, a operadora espera transmitir 40 MB para esse cliente.
- 74 A probabilidade de, em um dia escolhido ao acaso, o cliente consumir mais de 1,5 MB em dados é superior a 0,3.

Considerando que os principais métodos para a estimação pontual são o método dos momentos e o da máxima verossimilhança, julgue os itens a seguir.

- 75 O estimador da máxima verossimilhança para a variância da distribuição normal é expresso por $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$, e este estimador é não viciado.
- 76 Para a distribuição normal, o método dos momentos e o da máxima verossimilhança fornecem os mesmos estimadores aos parâmetros μ e σ .

Para estimar a porcentagem de eleitores que votariam a favor de um candidato presidencial, foi escolhida uma amostra aleatória de 200 pessoas. Dessa amostra, uma avaliação indicou que 60 eleitores votariam no referido candidato. Considerando que $\Phi(1,645) = 0,95$ e que $\Phi(1,96) = 0,975$ em que a função Φ representa a função distribuição acumulada da distribuição normal padronizada, julgue os seguintes itens.

- 77 Um intervalo de confiança (IC) de 95% é dado por $IC = [0,3 - \varepsilon, 0,3 + \varepsilon]$ em que $\varepsilon = \sqrt{\frac{0,3 \times 0,7}{200}} \Phi^{-1}(0,95)$.
- 78 O erro máximo provável do intervalo de confiança é inferior a 0,07.
- 79 A estimativa pontual para o parâmetro p — proporção de eleitores na população favorável ao candidato — é superior a 25%.

A distribuição da resistência X de uma linha elétrica (em ohms) é $N(\mu, \sigma^2)$. Por meio de 25 mensurações, foram determinadas as estatísticas amostrais $\bar{x} = 50$ ohms (média amostral) e $S = 5$ ohms (desvio padrão amostral). Considerando que se T tiver distribuição t de Student com 24 graus de liberdade, então $P(T \leq 0,95) = 1,711$; e que se T tiver distribuição t de Student com 25 graus de liberdade, então $P(T \leq 0,95) = 1,708$, julgue os próximos itens.

- 80 Nessa situação, o parâmetro μ satisfaz $P(-1,711 \leq 50 - \mu \leq 1,711) = 0,1$.
- 81 A estatística $T = \frac{(\bar{x} - \mu) \sqrt{n}}{S}$, em que n representa o tamanho da amostra, tem distribuição t de Student com n graus de liberdade.

	valor esperado, μ (em kg)	desvio padrão σ (em kg)
Japão	100	15
Taiwan	110	20

Um varejista de motocicletas e acessórios encontrou uma caixa de parafusos especiais de origem desconhecida para um modelo da marca Honda. Esses parafusos são produzidos apenas no Japão e Taiwan. As características da resistência à tração X dos parafusos são apresentadas na tabela. Uma amostra de 20 parafusos da caixa foi testada e encontrou-se a resistência à tração média \bar{x} .

Considere o teste a respeito da procedência dos parafusos constituído das seguintes hipóteses. H_0 : os parafusos procedem do Japão: $\mu = 100$; e H_1 : os parafusos procedem de Taiwan: $\mu = 110$. A regra da decisão do teste é não rejeitar H_0 se $\bar{x} < x_c$, em que x_c é um valor a ser encontrado; e rejeitar H_0 no caso contrário.

A respeito dessa situação, julgue os itens subsequentes.

- 82 O valor crítico x_c para o qual vale $P(\text{erro tipo I}) = P(\text{erro tipo II})$ é dado por $\frac{100 - x_c}{15} = \frac{x_c - 110}{20}$.
- 83 O teste descrito é um teste de hipóteses composto.
- 84 O erro do tipo I é cometido caso seja dito que o parafuso procede de Taiwan; quando na verdade, procede do Japão. O erro do tipo II é cometido caso seja dito que o parafuso procede do Japão; quando na verdade, procede de Taiwan.
- 85 A probabilidade do erro do tipo I em dependência do valor crítico x_c é dada por $1 - \Phi\left(\frac{x_c - 100}{\frac{15}{\sqrt{20}}}\right)$.

Um estudo realizado para avaliar a associação linear entre o índice de qualidade dos serviços prestados por empresas de TV por assinatura (X) e o índice de fidelização de seus usuários (Y) contou com informações fornecidas por 20 empresas e foram considerados três diferentes modelos de regressão linear simples:

$$\begin{aligned} Y_k &= a_0 + a_1 X_k + \varepsilon_k, \\ Y_k &= b_1 X_k + \varepsilon_k, \\ X_k &= c_0 + c_1 Y_k + \varepsilon_k, \end{aligned}$$

em que $k = 1, \dots, 20$; Y_k representa o índice de fidelização na empresa k ; e X_k denota a qualidade dos serviços prestados pela empresa k . Os termos $\varepsilon_1, \dots, \varepsilon_{20}$ representam erros aleatórios independentes e identicamente distribuídos, com média zero e desvio padrão σ .

A partir dessas informações, julgue os próximos itens, considerando que a estimativa do coeficiente a_1 obtida pelo método de mínimos quadrados ordinários seja $\hat{a}_1 = 0,5$ e que

$$\bar{X} = \sum_{k=1}^{20} \frac{X_k}{20} = 3 \quad ; \quad \bar{Y} = \sum_{k=1}^{20} \frac{Y_k}{20} = 4 ;$$

$$\sum_{k=1}^{20} (X_k - \bar{X})^2 = 20 \quad ; \quad \sum_{k=1}^{20} (Y_k - \bar{Y})^2 = 80.$$

- 86 Com base no método de mínimos quadrados ordinários, é correto afirmar que a estimativa do intercepto a_0 é maior que 2.

- 87 A correlação linear de Pearson entre as variáveis Y e X é menor que 0,30.
- 88 Na regressão linear que passa pela origem, a estimativa de mínimos quadrados ordinários do coeficiente b_1 é igual a $\frac{4}{3}$.
- 89 A estimativa de mínimos quadrados ordinários do coeficiente c_1 é maior ou igual a 2.
- 90 Considerando-se o modelo ajustado $\hat{Y}_k = \hat{a}_0 + \hat{a}_1 X_k$, em que \hat{a}_0 e \hat{a}_1 são as respectivas estimativas de mínimos quadrados ordinários dos coeficientes a_0 e a_1 , é correto afirmar que $\text{Var}[\hat{Y}_k] = \sigma^2$.

Para prever a demanda por determinado tipo de serviço de comunicação de dados, um especialista em gestão de telecomunicações considerou um modelo de regressão linear múltipla na forma $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, em que \mathbf{y} é o vetor de respostas, \mathbf{X} é a matriz de delineamento, $\boldsymbol{\beta}$ é o vetor de parâmetros, e $\boldsymbol{\varepsilon}$ denota o vetor de erros aleatórios independentes e identicamente distribuídos. Cada componente do vetor $\boldsymbol{\varepsilon}$ segue uma distribuição normal com média zero e variância v . O modelo ajustado é expresso por $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$, em que $\hat{\boldsymbol{\beta}}$ representa a estimativa de máxima verossimilhança do vetor $\boldsymbol{\beta}$.

Considerando que

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 0,6 & -0,3 & -0,2 \\ -0,3 & 0,2 & 0 \\ -0,2 & 0 & -0,2 \end{bmatrix},$$

em que \mathbf{X}' denota a transposta da matriz de delineamento, e que

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \text{ julgue os itens que se seguem.}$$

- 91 É correto afirmar que $\hat{\boldsymbol{\beta}} = \begin{bmatrix} -0,6 \\ 0,1 \\ 0,4 \end{bmatrix}$.
- 92 O vetor de resíduos é dado por $(\mathbf{I} - \mathbf{H})\mathbf{y}$, em que $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ é a matriz de projeção (*hat matrix*) e \mathbf{I} é a matriz identidade.
- 93 Conclui-se que $\text{Var}[\hat{\mathbf{y}}] = v\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.
- 94 O fator de inflação da variância (VIF) é obtido com base nos elementos da diagonal principal da matriz $(\mathbf{X}'\mathbf{X})^{-1}$, sendo sua principal função detectar possíveis pontos influentes ou valores atípicos (*outliers*) no vetor de resposta \mathbf{y} .
- 95 Se $v = 20$, então $\begin{bmatrix} 12 & -6 & -4 \\ -6 & 4 & 0 \\ -4 & 0 & 4 \end{bmatrix}$ será a matriz de covariância de $\hat{\boldsymbol{\beta}}$.

Um estudo a respeito do índice de cancelamento de assinaturas (Y) de uma operadora de telefonia celular no período de 2010 a 2014 produziu um ajuste na forma $\hat{Y}_t = \hat{a} + \hat{b}(t - 2012)$, em que $t = 2010, 2011, 2012, 2013, 2014$; \hat{Y}_t é a estimativa desse índice no ano t correspondente; e \hat{a} e \hat{b} representam as estimativas de mínimos quadrados ordinários dos coeficientes da reta ajustada. A tabela a seguir apresenta a análise de variância (ANOVA) do ajuste.

fonte de variação	graus de liberdade	soma de quadrados
modelo	1	324
erro	3	76
total	4	400

Considerando que $\hat{a} = 30$, julgue os itens subsequentes relativos ao referido ajuste.

- 96 Se, em um teste de significância para o coeficiente angular em questão, as hipóteses nula e alternativa forem, respectivamente, $H_0: b = 0$ e $H_1: b \neq 0$, então o valor da razão t desse teste será inferior a -4 ou superior a 4 .
- 97 A estatística F da tabela ANOVA, que permite testar conjuntamente a significância estatística dos coeficientes a e b , corresponde à razão entre a soma de quadrados do modelo e a soma de quadrados dos erros. As hipóteses nula e alternativa dessa estatística são, respectivamente, $H_0: a = 0$ e $b = 0$ e $H_1: a \neq 0$ ou $b \neq 0$.
- 98 A correlação linear de Pearson entre a variável resposta e a variável regressora foi superior a $-0,75$ e inferior a $0,75$.
- 99 Na análise de resíduos, recomenda-se realizar uma avaliação para se detectar possível correlação linear entre os erros consecutivos. Essa avaliação pode ser realizada, por exemplo, por meio das estatísticas C_p de Mallow.
- 100 No período de 2010 a 2014, a média aritmética do índice Y foi igual a 30.
- 101 O coeficiente de determinação do modelo (R^2) é superior a 0,90.
- 102 A estimativa da variância do erro aleatório em torno da tendência ajustada é superior a 27.
- 103 O quadrado do coeficiente angular \hat{b} , é inferior a 30.
- 104 A estimativa da variância de \hat{b} é inferior a 3.
- 105 No período de 2010 a 2014, o desvio padrão amostral do índice de cancelamento de assinaturas (Y) foi igual a 10.

Uma pesquisa de satisfação dos usuários de um serviço de comunicação de dados prestado pela operadora X foi realizada em determinado estado. Na ocasião, a população de interesse dessa pesquisa era constituída por 2.000 empresas. Ela foi dividida em três estratos A, B e C, tendo sido retirada de cada um deles uma amostra aleatória simples de empresas, na qual se registrou o total de empresas satisfeitas com esse serviço, conforme mostra a tabela a seguir.

estrato	tamanho populacional	tamanho da amostra	total de empresas satisfeitas na amostra
A	1.250	125	40
B	500	50	40
C	250	25	20
total	2.000	200	100

Com base nas informações e na tabela apresentadas, julgue os itens seguintes.

- 106 A estimativa da variância da fração (proporção) de empresas satisfeitas no estrato B foi inferior a 0,0032.
- 107 Na ocasião da pesquisa, 50% das empresas da referida população estavam satisfeitas com o serviço de comunicação de dados em questão.
- 108 A fração amostral da pesquisa em questão foi superior a 50%.
- 109 Suponha que os estratos tenham sido estabelecidos com base no porte das empresas. Nesse caso, a unidade primária de amostragem corresponde ao porte, e cada empresa representa uma unidade secundária.
- 110 O plano amostral descrito no texto corresponde a uma amostragem aleatória estratificada com alocação proporcional ao tamanho dos estratos.

Um analista da área de estatística da TELEBRAS tem a tarefa de verificar se a atuação dos órgãos de defesa do consumidor em um processo referente a cobranças abusivas feitas por empresas operadoras de telefonia móvel resultou em efetiva alteração no valor das contas apresentadas aos clientes. Para isso, o analista dispõe de dados de dois grupos distintos, um com 300 clientes cujos dados foram coletados antes da atuação dos órgãos de defesa do consumidor, e um segundo com outros 350 clientes e dados coletados após essa atuação.

Considerando essa situação hipotética e com base nos conceitos de inferência estatística, julgue os itens a seguir.

- 111 Sabendo-se que as variâncias amostrais, s_1^2 e s_2^2 correspondentes, respectivamente, aos dados do grupo com 300 clientes e com 350 clientes, são $s_1^2 = s_2^2 = 310$, é correto afirmar que qualquer teste que o analista opte por realizar será constituído por mais de 300 graus de liberdade.
- 112 Considerando-se que o analista deseje fazer um teste bilateral, é correto afirmar que o valor crítico do teste para 95% de confiança será dado por 1,96, uma vez que $P(Z < 1,645) = 0,95$ e $P(Z < 1,96) = 0,975$.
- 113 O analista deverá analisar os dados utilizando um teste para dados pareados, uma vez que os dados são de clientes diferentes e foram mensurados antes e depois da atuação dos órgãos de defesa do consumidor.
- 114 Se o analista optar por utilizar um teste para diferenças de médias com dados independentes, ele deverá considerar que a estatística desse tipo de teste segue uma distribuição t de Student.

Um analista da TELEBRAS, a fim de verificar o tempo durante o qual um grupo de consumidores ficou sem o serviço de Internet do qual eram usuários, selecionou uma amostra de 10 consumidores críticos. Os dados coletados, em minutos, referentes a esses consumidores foram listados na tabela seguinte.

consumidor	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
tempo (em minutos)	8	2	3	5	7	7	10	9	4	5

Com base nessa situação hipotética, julgue os itens subsequentes.

- 115 O estimador da média é preferível ao da mediana, pois o primeiro apresenta mais de 30% de eficiência sobre o segundo.
- 116 O segundo momento dos dados é menor do que o primeiro momento ao quadrado.
- 117 Para verificar se o tempo médio sem Internet é igual a 5 minutos, o analista deverá realizar um teste com 8 graus de liberdade.
- 118 Se os dados seguissem uma distribuição normal, a expressão matemática que permite calcular a variância estimada pelo método de máxima verossimilhança teria denominador igual a 9.
- 119 Na situação hipotética em apreço, as estimativas pontuais da média e da mediana são as mesmas e ambos estimadores são consistentes, suficientes e não viesados.
- 120 Sabendo-se que $\sum_{i=1}^{10} x_i^3 = 3.276$, em que x_i corresponde ao tempo que o consumidor c_i , para $i = 1, 2, \dots, 10$ ficou sem internet, é correto afirmar que, no cálculo do terceiro momento central, $\frac{1}{10} \sum x_i^3$ pode ser maior ou menor que o terceiro momento.

Espaço livre