Estatística para Cursos de Engenharia e Informática, 4 ed.: Solução Analítica de Exercícios Selecionados

Os capítulos 4 a 8 envolvem mais fortemente o raciocínio lógico, sendo, portanto, importante que o estudante faça bastante exercícios. Neste contexto, apresentamos, a seguir, a solução analítica dos exercícios especificados a seguir:

Capítulo 4: exercícios 23 e 32;

Capítulo 5: exercícios 24 e 28;

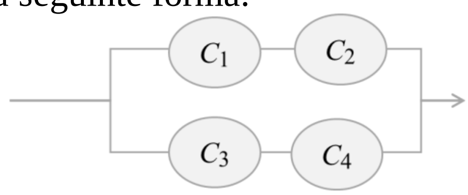
Capítulo 6: exercícios 7, 10 e 19;

Capítulo 7: exercícios 4 e 23;

Capítulo 8: exercício 27.

Cap. 4 – Probabilidade

23) Um sistema tem quatro componentes que operam independentemente, sendo que cada componente tem probabilidade 0,1 de não funcionar. O sistema é ligado da seguinte forma:



Determine a probabilidade de o sistema funcionar.

Solução.

Sejam os eventos: C_i: componente i funciona (i = 1, 2, 3, 4) e S = sistema funciona.

$$P(S) = P\{(\text{Caminho de baixo}) \cup (\text{caminho de cima})\} =$$
$$= P\{(C_1 \cap C_2) \cup (C_3 \cap C_4)\}$$

Usando a regra da soma de probabilidades:

$$P(S) = P(C_1 \cap C_2) + P(C_3 \cap C_4) - P\{(C_1 \cap C_2) \cap (C_3 \cap C_4)\}$$

Como as componentes operam independentemente, temos:

$$P(S) = P(C_1) \times P(C_2) + P(C_3) \times P(C_4) - P(C_1) \times P(C_2) \times P(C_3) \times P(C_4) =$$
$$= (0{,}9)^2 + (0{,}9)^2 - (0{,}9)^4 = 0{,}6939$$

32) A caixa I tem 8 peças boas e 2 defeituosas; a caixa II tem 6 peças boas e 4 defeituosas; a caixa III tem 15 peças boas e 5 defeituosa.

a) Tira-se, aleatoriamente, uma peça de cada caixa. Determinar a probabilidade de serem todas boas.

b) Escolhe-se uma caixa ao acaso e tira-se uma peça. Determinar a probabilidade de ser defeituosa.

c) Escolhe-se uma caixa ao acaso e tira-se uma peça. Calcular a probabilidade de ter sido escolhida a caixa I, sabendo-se que a peça é defeituosa.

Solução.

a) Sejam os eventos B_i: peça boa extraída da caixa C_i (i = 1, 2, 3).

$$P(\text{todas boas}) = P(B_1 \cap B_2 \cap B_3) = P(B_1) \times P(B_2) \times P(B_3) =$$
$$= \frac{8}{10} + \frac{6}{10} + \frac{15}{20} = \frac{9}{25}.$$

b) Sejam os eventos D: peça defeituosa e C_i: seleção da caixa i (i = 1, 2, 3). Usando o teorema da probabilidade total, temos:

$$P(D) = P(C_1) \times P(D \mid C_1) + P(C_2) \times P(D \mid C_2) + P(C_3) \times P(D \mid C_3) =$$
$$= \frac{1}{3} \times \frac{2}{10} + \frac{1}{3} \times \frac{4}{10} + \frac{1}{3} \times \frac{5}{20} = \frac{1}{3} \times \frac{17}{20} = \frac{17}{60}.$$

c) Usando o teorema de Bayes:

$$P(C_1 \mid D) = \frac{P(C_1) \times P(D \mid C_1)}{P(D)} = \frac{\frac{1}{3} \times \frac{2}{10}}{\frac{17}{60}} = \frac{4}{17}.$$

Cap. 5 – Variáveis Aleatórias Discretas

24) Suponha que o número de falhas em certo tipo de placa plástica tenha distribuição de Poisson, com taxa média de 0,05 defeito por m2. Na construção de um barco, é necessário cobrir uma superfície de 3m x 2m com essa placa.

a) Qual é a probabilidade de que não haja falhas nessa superfície?

b) Qual é a probabilidade de que haja mais que uma falha nessa superfície?

c) Na construção de 5 barcos, qual é a probabilidade de que pelo menos 4 não apresente defeito na superfície plástica.

Solução. Sejam as variáveis aleatórias:

X = Número de falhas por m^2, supostamente Poisson com λ = 0,05.

Y = Número de falhas na superfície de 3m x 2m = $6m^2$, então Y tem distribuição de Poisson com $\lambda = 6\times0{,}05 = 0{,}30$.

a) $P(Y=0)=p(0)=\dfrac{e^{-\lambda}\times\lambda^{y}}{y!}=\dfrac{e^{-0,3}\times(0,3)^{0}}{0!}=e^{-0,3}=0,7408$

b)

$P(Y>1)=1-[p(0)+p(1)].$

$p(1)=\dfrac{e^{-\lambda}\times\lambda^{y}}{y!}=\dfrac{e^{-0,3}\times(0,3)^{1}}{1!}=(0,3)e^{-0,3}=0,2222$. Então:

$P(Y>1)=1-[0,7408+0,2222]=0,037.$

c) Observar que, neste caso, temos outra variável aleatória envolvida na questão, que vamos definir como: W = *Número de barcos que não apresentam defeito, numa amostra de cinco barcos*; e queremos $P(W \geq 4)$. Podemos dizer que W é o resultado de cinco ensaios independentes com probabilidade p = 0,7408, a qual foi calculada no item (a). Então W tem distribuição binomial com parâmetros n = 5 e p = 0,7408.

$P(W\geq 4)=p(4)+p(5).$

$$p(4) = \binom{5}{4} \times (0,7408)^4 \times (1 - 0,7408)^1 =$$

$$= \frac{5!}{4!(5-4)!} \times (0,7408)^4 \times (0,2592) = 0,3903$$

$$p(5) = \binom{5}{5} \times (0,7408)^5 \times (1 - 0,7408)^0 = 0,2231.$$ Logo:

$$P(W \geq 4) = p(4) + p(5) = 0,3903 + 0,2231 = 0,6134.$$

28) Um armazém é abastecido mensalmente, sendo que a taxa média de abastecimento é 30 unidades/dia, com desvio padrão de 3 unidades/dia. A demanda média é de 25 unidades/dia, com desvio padrão de 4 unidades/dia. Suponha que o abastecimento e a demanda sejam independentes e, além disso, a demanda e o abastecimento em um dia não alteram o abastecimento e a demanda nos dias seguintes. Qual é o valor esperado e o desvio padrão do excedente de produtos, no período de um mês?

Solução. Sejam as variáveis aleatórias:

- A_i = Abastecimento no dia i; D_i = Demanda no dia i;
- $A = A_1 + A_2 + ... + A_{30}$ = Abastecimento no mês;
- $D = D_1 + D_2 + ... + D_{30}$ = Demanda no mês; e
- $Y = A - D$ = Excedente no mês.

Temos:

$E(A_i) = 30$; $E(D_i) = 25$; $V(A_i) = 3^2 = 9$; e $V(D_i) = 4^2 = 16$.

No mês: $E(A) = E(A_1) + E(A_2) + ... + E(A_{30}) = 900$

$$E(D) = E(D_1) + E(D_2) + ... + E(D_{30}) = 750.$$

Como é suposto que A_i e A_j são independentes, $\forall i, j = 1, 2, ..., 30$, então:

$V(A) = V(A_1) + V(A_2) + ... + V(A_{30}) = 270$ e

$V(D) = V(D_1) + V(D_2) + ... + V(D_{30}) = 480$. Logo:

$E(Y) = E(A - D) = E(A) + E(-D) = E(A) - E(D) = 900 - 750 = 150.$

$V(Y) = V(A - D) = V(A) + V(-D) = V(A) + (-1)^2\, V(D) =$

$= V(A) + V(D) = 270 + 480 = 750$. Portanto,

$DP(D) = \sqrt{750} = 27{,}39.$

Cap. 6 - Variáveis Aleatórias Contínuas

7) Seja T uma variável aleatória com distribuição exponencial. Usando a expressão de probabilidade condicional (Capítulo 4), mostrar que para qualquer $s, t > 0$, vale a seguinte relação, conhecida como propriedade de *falta de memória*:

$$P\left(T > s + t \mid T > s\right) = P(T > t)$$

Solução.

$$P\left(T > s + t \mid T > s\right) = \frac{P\{(T > s + t) \cap (T > s)\}}{P(T > s)}$$

Observar que:

$(T > s) \subseteq (T > s+t)$, já que $t > 0$ e $s > 0$. Logo:

$$P\left(T > s+t \,|\, T > s\right) = \frac{P(T > s+t)}{P(T > s)} =$$

$$= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = \frac{e^{-\lambda t} \times e^{-\lambda s}}{e^{-\lambda s}} = e^{-\lambda t} = P(T > t)$$

16) A vida útil de um certo componente eletrônico é, em média, 10.000 horas e apresenta distribuição exponencial. Após quantas horas se espera que 25% dos componentes tenham falhado?

Solução.

Seja T a variável aleatória que representa o tempo de vida do componente. É suposto que T tem distribuição exponencial com $E\left(T\right) = 10.000$. Considerando que $E(T) = 1/\lambda$, então temos que: $\lambda = 1/10.000$.

Para um dado *t* > 0, podemos ter $P\left(T \le t\right) = 0,25$. Então, para esse *t*, temos $P\left(T > t\right) = 0,75 \Rightarrow e^{-\lambda t} = 0,75 \Rightarrow -\lambda t = \ln(0,75)$.

$\Rightarrow t = -\dfrac{\ln(0,75)}{\lambda} = -10.000 \times \ln(0,75) \cong 2.877.$ Ou seja, 2.877 horas.

19) O tempo para que um sistema computacional execute uma determinada tarefa é uma variável aleatória com distribuição normal de média 320 segundos e desvio padrão de 7 segundos.

a) Qual é a probabilidade de a tarefa ser executada entre 310 e 330 segundos?

b) Se a tarefa é colocada para execução 200 vezes. Qual é a probabilidade dessa tarefa demorar mais que 325 segundos em pelo menos 50 vezes?

Solução.

a) Queremos $P(310 \leq X \leq 330)$, sendo X normal com média de 320 e desvio padrão de 7 segundos. Então, temos os escores padronizados:

$x_1 = 310 \Rightarrow z_1 = \dfrac{310 - 320}{7} \cong -1,43$ e

$x_2 = 330 \Rightarrow z_2 = \dfrac{330 - 320}{7} \cong 1,43$

Logo, $P(310 \leq X \leq 330) \cong P(-1,43 \leq Z \leq 1,43)$, sendo Z normal padrão. Assim:

$$P(310 \leq X \leq 330) \cong 1 - 2 \times P(Z > 1,43) = 1 - 2(0,0764) = 0,8472.$$

b) Vamos calcular, primeiro, a probabilidade p de a tarefa demorar mais que 325 segundos numa particular vez:

$$p = P(X > 325) = P\left(Z > \frac{325 - 320}{7} \right) = P(Z > 0,71) = 0,2389.$$

Seja Y a variável aleatória que representa o número de ocorrências do evento *demorar mais que 325 segundos*, nas 200 replicações do processo. Observe que podemos assumir que Y tem distribuição binomial com n = 200 e p = 0,2389. Queremos calcular: $P(Y \geq 50)$.

Como n é grande, podemos supor que Y tem distribuição aproximadamente normal com:

$\mu = n \times p = 200 \times 0,2389 = 47,78$ e

$\sigma = \sqrt{n \times p \times q} = \sqrt{200 \times 0,2389 \times (1-0,2389)} = 6,03$

Assim:

$$P(Y \geq 50) \cong P\left(Z \geq \frac{49,5-47,78}{6,03}\right) = P\left(Z \geq 0,29\right) = 0,3859.$$

Obs. Os resultados desse exercício são aproximados, porque resolvemos usando a tabela da distribuição normal padrão, arredondado o valor de *z* para duas decimais.

Cap. 7 - Distribuições Amostrais e Estimação de Parâmetros

4) Existem vários algoritmos computacionais que permitem gerar números aleatórios (ou, mais apropriadamente, números *pseudoaleatórios*) no intervalo [0, 1], com distribuição uniforme. Seja uma amostra aleatória simples formada pela geração independente de cem números pseudoaleatórios em [0, 1], representada por (X_1, X_2, ..., X_{100}). Seja $\bar{X}$ a média aritmética simples desses 100 números.

a) Qual é o valor esperado e a variância de X_1?

b) Qual é a probabilidade de X_1 assumir um valor no intervalo [0,47; 0,53]?

c) Qual é o valor esperado e a variância de $\bar{X}$?

d) Qual é a distribuição de probabilidade de $\bar{X}$?

e) Qual é a probabilidade de $\bar{X}$ assumir um valor no intervalo [0,47; 0,53]?

Solução.

a) $X_1 \sim U[0,1]$

$$E(X_1) = \int_0^1 x1dx = \frac{x^2}{2}\Bigg|_0^1 = \frac{1}{2} - \frac{0}{2} = \frac{1}{2}$$

$$E(X_1^2) = \int_0^1 x^2 1dx = \frac{x^3}{3}\Bigg|_0^1 = \frac{1}{3} - \frac{0}{3} = \frac{1}{3}$$

$$V(X_1) = E(X_1^2) - [E(X_1)]^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}.$$

b) $P(0,47 \le X_1 \le 0,53) = \int_{0,47}^{0,53} 1dx = 0,06$

c) $E\left(\bar{X}\right) = E(X_1) = \dfrac{1}{2}$

$$V\left(\bar{X}\right) = \frac{V(X_1)}{n} = \frac{1/12}{100} = \frac{1}{1.200}.$$

d) Como são n observações independentes com a mesma distribuição, $U[0, 1]$, e n é razoavelmente grande, então $\bar{X}$ tem distribuição aproximadamente normal com média e variância calculados no item anterior.

e) $P(0,47 \le \bar{X} \le 0,53) = ?$ Escores padronizados:

$$z_1 = \frac{\bar{x}_1 - E\left(\bar{X}\right)}{DP\left(\bar{X}\right)} = \frac{0,47 - \frac{1}{2}}{\sqrt{\frac{1}{1.200}}} = -1,04$$

$$z_2 = \frac{\bar{x}_2 - E\left(\bar{X}\right)}{DP\left(\bar{X}\right)} = \frac{0,53 - \frac{1}{2}}{\sqrt{\frac{1}{1.200}}} = 1,04.$$ Logo:

$$P\left(0,47 \le \bar{X} \le 0,53\right) = P\left(-1,04 \le Z \le 1,04\right) =$$
$$= 1 - 2(0,1492) = 0,7016$$

Obs. O resultado do último item é aproximado, porque resolvemos usando a tabela da distribuição normal padrão, arredondado o valor de z para duas decimais.

22) Uma empresa tem 2.400 funcionários que usam o refeitório. Deseja-se extrair uma amostra para verificar o grau de satisfação dos funcionários em relação à qualidade da comida no refeitório. Em uma amostra piloto, extraída de forma aleatória, com n_0 = 31 funcionários, o grau de satisfação teve nota média de 6,5 e desvio padrão de 2,0, numa escala de 0 a 10.

a) Determine o tamanho mínimo da amostra, supondo amostragem aleatória simples, com erro máximo de 0,3 unidades e nível de confiança de 95%.

b) Considere que a amostra planejada no item anterior tenha sido realizada, resultando numa média de 5,20 e desvio padrão de 1,80 pontos. Construa o intervalo de 95% de confiança para o parâmetro μ.

c) Considerando o resultado do item anterior, você diria, com nível de confiança de 95%, que a nota média seria superior a cinco se a pesquisa fosse aplicada a todos os 2.400 funcionários? Justifique.

d) Realizada a amostra planejada no item (a), suponha que 70 funcionários atribuíram notas iguais ou superiores a cinco. Apresente um intervalo de 95% de confiança para a porcentagem de funcionários de toda a empresa que atribuiriam notas iguais ou superiores a cinco.

Solução.

Tamanho da população: N = 2.400

Amostra piloto: n_0 = 31, $\bar{x}_0$ = 6,5 e S_0 = 2,0.

a) $n_0 = \dfrac{t_{0,95}^2 \times S_0^2}{E_0^2} = \left(\dfrac{2,42 \times 2}{0,3}\right)^2 = 185,32$.

Corrigindo pelo tamanho da população:

$$n \cong \frac{N \times n_0}{N \times + n_0 - 1} = \frac{2400 \times 185,32}{2400 + 185,32 - 1} = 172,11.$$

Levando para o menor inteiro superior a esse valor mínimo, temos: n = 173.

b) Sendo: n = 173, $\bar{x}$ = 5,20 e S = 1,80;

Erro padrão: $S_{\bar{X}} = \dfrac{S}{\sqrt{n}} \times \sqrt{\dfrac{N-n}{N-1}} = \dfrac{1,80}{\sqrt{173}} \times \sqrt{\dfrac{2.400-173}{2.400-1}} = 0,1319$

Intervalo de confiança:

$$IC(\mu, 95\%) = \bar{x} \pm t_{0,95} S_{\bar{X}} = 5,20 \pm 1,96 \times 0,1319 = 5,20 \pm 0,26.$$

c) Não, pois o intervalo onde deve estar a verdadeira média abrange também valores menores ou iguais a cinco.

d) $\hat{p} = \dfrac{70}{173} = 0,4046$

e) $S_{\hat{P}} = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} \times \sqrt{\dfrac{N-n}{N-1}} = \sqrt{\dfrac{0,4046(1-0,4046)}{173}} \times \sqrt{\dfrac{2.400-173}{2.400-1}} = 0,3595$

$$E = 1,96 \times 0,3595 = 0,0704$$

$$IC(p, 95\%) = \hat{p} \pm z_{0,95} S_{\hat{P}} = 0,4045 \pm 1,96 \times 0,0704 \cong 0,404 \pm 0,070.$$

Cap. 8 – Testes de Hipóteses

27) O controle estatístico de um processo estabeleceu que pelo menos 80% dos produtos têm que estar com certo padrão de qualidade. Para verificar a validade desta afirmação, coleta-se uma amostra de 400 produtos, obtendo-se uma proporção de 76% em conformidade com o padrão de qualidade.

a) Com nível de significância de 1%, há evidência que o processo esteja em desacordo com o esperado?

b) Se o percentual real de itens satisfazendo o padrão de qualidade for 74%, qual é a probabilidade de se tomar uma decisão errada no item (a)?

c) Suponha que se queira identificar, com 95% de probabilidade, a falsidade de H_0, quando a proporção de itens no padrão for, na realidade, igual a 74%. Considerando que o teste será realizado com 1% de significância, qual é o tamanho de amostra necessário?

Solução.

a) H_0: p = 0,80 vs. H_1:p < 0,80

Observar que na hipótese nula colocamos uma igualdade no ponto limite de conformidade do produto.

Dado: n = 400; $\hat{p}$ = 0,80.

Sob H_0, o valor esperado de itens em conformidade com o padrão é:

$$E(Y) = np_0 = 400(0,80) = 320$$

Na amostra foi observado: y = 400(0,76) = 304. Fazendo a correção de continuidade, temos: y' = 304 + 0,5 = 304,5.

Estatística do teste:

$$z = \frac{y' - np_0}{\sqrt{np_0(1-p_0)}} = \frac{304,5-320}{\sqrt{400(0,8)(0,2)}} \cong -1,94$$

Pela tabela da distribuição normal padrão, temos:

$$Valor\text{-}p = P(Z < -1,94) = 0,262$$

Como o *valor-p* é maior que o nível de significância adotado para esse teste (α = 0,01), então o teste aceita H_0, ou seja, <u>não</u> há evidência, ao nível de significância de 1%, de que o processo esteja em desacordo com o padrão de qualidade.

b) A decisão errada, no caso, seria o teste aceitar H_0, quando o verdadeiro valor de p é 0,74 (erro tipo II). Logo, o cálculo que se precisa fazer é:

$$\beta(0,74) = P(\textit{decisão errada} \mid p = 0,74)$$

Considerando a regra de decisão:

$$\text{Aceita } H_0 \Leftrightarrow z > z_c = -2,326$$

Então:

$$P(\textit{decisão errada} \mid p = 0,74) = P(Z > -2,326 \mid p = 0,74) =$$

$$= P\left(\frac{Y' - np}{\sqrt{np_0(1-p_0)}} > -2,326 \mid p = 0,74\right) =$$

$$= P\left(\frac{Y' - 320}{8} > -2,326 \mid p = 0,74\right) = P(Y' > 301,392 \mid p = 0,74) =$$

$$= P\left(Z > \frac{301,392 - 400(0,74)}{\sqrt{400(0,74)(0,26)}}\right) = P(Z > 0,6146) \cong$$

$$\cong P(Z > 0,61) = 0,27.$$

c) Usando:

$$n = \left(\frac{z_\alpha\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p(1-p)}}{p - p_0}\right)^2$$

Temos:

$$n = \left(\frac{2,326\sqrt{0,80(0,20)} + 1,645\sqrt{(0,74)(0,26)}}{0,74 - 0,80}\right)^2 = 758,04$$

Aproximando para o inteiro imediatamente superior, temos o tamanho da amostra igual a n = 759.