ERRATA DO LIVRO *PROBABILIDADE E ESTATÍSTICA: QUANTIFICANDO A INCERTEZA*

Agradecimentos

- ... Bianca Moreira, aluna ~~de~~ dos cursos ...

Cap1

- Página 12
  Penúltima linha:

  ... = $\dfrac{12!}{3!(9)!}$ Excluir os parênteses

- Página 34
  Primeira linha

  Da relação (IV), vemos que cada coordenada do vetor $\mathbf{x}$ é igual ao produto interno de uma linha da matriz Σ^{-1} pelo vetor $\mathbf{y}$, que, por sua vez, é um múltiplo do vetor $\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$. Então todas as 4 coordenadas do vetor $\mathbf{x}$ são...

- Página 36
  Ao final do Exercício R1.9, acrescentar o seguinte:

  (Ver subseção 6.1.1)

Cap2

- Página 59

  Notas:
  1. ...
  2. ...
  3. Os valores das probabilidades relativas a uma Binomial podem ser obtidos diretamente a partir de um software adequado.

- Página 64
  Logo após as Notas 1 e 2:

  Em outras palavras, quando N é grande, a distribuição de probabilidade hipergeométrica pode ser aproximada por uma binomial de parâmetros n e p = K/N. (Ver Exercício Resolvido 2.3 ~~2.2 item(b)~~)

- Páginas 67 e 68
  No Resumo do Cap2, a itemização das Propriedades da Função de Distribuição Acumulada é de (a) até (d), e não de (c) até (f).

- Página 69

  Na solução do Exercício R2.1 Turbulência no avião

  (b) P(X $\geq$ 3) = 1 – P(X $\leq$ 2) = …

  No enunciado do Exercício R2.2 Examinando alunos, somente o item (b) deve ficar marcado com um asterisco, não o item (a).

- Página 71

  Na solução do Exercício R2.3 Aproximação da Hipergeométrica pela Binomial

  a) Nos 3 primeiros casos X é Hipergeométrica(5,N,K), isto é,

  $$P(X=k) = \frac{\binom{K}{k}\binom{N-K}{5-k}}{\binom{N}{5}} \quad \text{para } k = 0,1,2,3,4,5$$

- Página 73

  Na solução do Exercício R2.4 Avaliação da qualidade de um processo produtivo, acrescentar o que está em vermelho abaixo:

  b) Usando o valor de p = 0,8 calculado no item (a), temos:

  P(no mínimo bom) = P(X $\geq$ 3) = $10(1-p)^2p^3 + 5(1-p)p^4 + p^5$ = 0,942

  e P(ótimo) = P(X $\geq$ 4) = $5(1-p)p^4 + p^5$ = 0,737

  Então, P(X≥4|X≥3) = $\frac{P(X\geq 4)}{P(X\geq 3)} = 0{,}7825$.

Cap3

- Página 82

  Na definição de variável aleatória contínua:

  3. Para quaisquer a, b reais (a<b), P(a≤X≤b) = $\int_a^b f(x)dx$ (Excluir os parênteses)

- Página 94

  A convenção adotada no livro é que a palavra Exponencial deve ser usada sempre com “E” maiúsculo em lugar de “e” minúsculo quando se faz referência à distribuição ou modelo Exponencial.

  Isso precisa ser corrigido uma vez na página 92 (na definição), cinco vezes na página 93, nove vezes na página 94, três vezes na página 95 e uma vez na página 96.

  Quando a referência é à função exponencial exp(.) tal que exp(x) = e^x, aí sim usa-se “e” minúsculo

- Página 103

  Na Figura 3.17 – Trabalhando com a Normal Padrão, corrigir:

  Função densidade φ() FDA Φ(.)

- Página 105

  No item 6:

  • Se X ~ (μ; 25) e P(X < 32) = 0,35...

- Página 109

  No enunciado do Exercício R3.1:

  ... sendo f(x) = 0 quando x está fora do intervalo (a-k; a+k), ~~sendo~~ h e k constantes positivas.

- Página 113

  No enunciado do Exercício R3.1:

  a) ... = 1, ou seja, φ é uma densidade... (Usar o mesmo símbolo que foi usado dentro da integral)

  b) ...

  c) $\int_{-\infty}^{+\infty} z^2 \varphi(z)dz$ = ... (Novamente usar o símbolo que a ser corrigido no item (a),

Cap4

- Página 134

  Na solução do Exemplo 4.9:

  Para a variância temos:

  $\text{Var}(Y) = (200^2 \times 0{,}8849 + 120^2 \times 0{,}1144 + 0^2 \times 0{,}0007) - (190{,}7)^2 = \mathbf{673{,}82}$

  Daí , DP(Y) = **25,96**

  Portanto, o preço médio dos rolos é de R$ 190,70, com desvio padrão igual a R$ **25.96**.

Cap5

- Página 161

Na última linha:

Note também que $p_{X|Y}(x_i|y_j) = p_X(x_i)$ e que $p_{Y|X}(y_j|x_i) = p_Y(y_j)$.

- Página 163

Logo após a definição de covariância:

Obs.:

1. Note que esta definição é válida tanto para o caso discreto quanto para o caso contínuo.
2. Mais adiante, na sub-seção 5.8.2, o leitor encontrará uma metodologia geral para o cálculo da esperança de uma função de duas variáveis aleatórias. Em particular, os valores esperados que constam nas expressões acima podem ser obtidos através dessa metodologia.
3. A veracidade da expressão alternativa na definição anterior deriva do seguinte:

   $E[(X - \mu_X)(Y - \mu_Y)] = E(XY - X\mu_Y - \mu_X Y + \mu_X \mu_Y) =$

   $= E(XY) - \mu_Y E(X) - \mu_X E(Y) + \mu_X \mu_Y = E(XY) - \mu_X \mu_Y$

- Página 167

  Na solução do Exemplo 5.15:

  Dado que X e Y assumem valores inteiros não negativos, Z assumirá também valores inteiros não negativos. Temos, para um dado valor k de Z, sendo k = 0,1,2,... :

  P(Z=k) = P(X+Y=k) = $\cancel{\sum_{i=0}^{k} P(X = k - i,\ Y = i)}$

- Página 168

  Corrigir a Figura 5.7

- Página 190

  P5.7) **Retomando o tema da** Pesquisa de Mercado

- Página 193

  No enunciado do Exercício P5.13:

  Alinhar horizontalmente a 2ª linha, onde a palavra “e” está abaixo das demais.

Cap6

- Página 207

  Ao final do Exemplo 6.6:

  $$P[30W \geq 30000] = P[W \geq 1000] = P\left[Z \geq \frac{1000 - 1100}{\sqrt{4250}}\right] = P[Z \geq -1{,}534] = 0{,}9375.$$

  (Ver também o Exercício P6.15)

- Página 207

  Logo após o enunciado do Teorema Central do Limite:

  Este é um dos resultados mais importantes da Teoria das Probabilidades e nos mostra claramente a relevância da distribuição Normal. Com efeito, conforme este Teorema, a distribuição da soma de ~~quaisquer~~ n variáveis aleatórias independentes e identicamente distribuídas ...

- Página 212

  Na última linha:

  ... a distribuição Normal de média r/p e variância r**(1-p)**/p^2.

- Página 213

Na 2ª e 3ª linhas:

$$P(a \le X \le b) \cong P\left(a - \tfrac{1}{2} \le W \le b + \tfrac{1}{2}\right) = \Phi\left(\frac{b + \frac{1}{2} - \frac{r}{p}}{\sqrt{r(1-p)/p^2}}\right) - \Phi\left(\frac{a - \frac{1}{2} - \frac{r}{p}}{\sqrt{r(1-p)/p^2}}\right),$$

onde W ~ Normal(r/p, $r(1-p)/p^2$).

- Página 213

Na solução do Exemplo 6.10:

... distribuição Normal(40/0,6, 40×**(1 – 0,6)**/$0{,}6^2$), ou seja:

$$P(X \ge 71) \cong \mathbf{1} - \Phi\left(\frac{71 - \frac{1}{2} - \frac{40}{0{,}6}}{\sqrt{40 \times (\mathbf{1 - 0{,}6}) / 0{,}6^2}}\right) = \mathbf{0{,}283 \text{ ou } 28{,}3\%.}$$

- Página 217

No item **Aproximação Normal para a distribuição de Pascal**

$$P(a \le X \le b) = \Phi\left(\frac{b + \frac{1}{2} - \frac{r}{p}}{\sqrt{r(1-p)/p^2}}\right) - \Phi\left(\frac{a - \frac{1}{2} - \frac{r}{p}}{\sqrt{r(1-p)/p^2}}\right)$$

- Página 219

Ao final da solução do Exercício R6.2:

Isso significa que a chance da obra ser realizada com base nessas 30 quotas seria de aproximadamente 64%.

- Página 223

No enunciado do Exercício R6.6, inserir após o item (c):

d) Determine o vetor de médias, a matriz de covariâncias e a matriz de correlações do vetor aleatório (X_1,X_2,X_3), novamente no caso em que
n = 1000 p_1 = 0,50 p_2 = 0,40 p_3 = 0,10.

- Página 225

No enunciado do Exercício R6.8, fazer os seguintes consertos:

Voltemos ao Exemplo 5.1. Suponha que o custo de conserto de cada defeito grave do carro, em reais, é uma variável aleatória U ~ Normal (1250,**00**; 200,**00**^{2}) e para cada defeito menor, esse custo é uma variável aleatória V ~ Normal (250,**00** ; 50,00^2). ...

- Página 228

Na sugestão do Exercício P6.3:

... variável aleatória U_i tal que $U_i = \begin{cases} 1, \text{se } i \in \Gamma \\ 0, \text{se } i \notin \Gamma \end{cases}$.

- Página 229

P6.8) **Avaliando o interesse do público por um novo produto** ~~Pesquisa de Mercado~~

Cap7

- Página 249
  No último parágrafo antes do Exemplo 7.11:

  Um segundo critério (também muito usado, que se baseia em medidas mais resistentes**)** para a identificação de observações discrepantes ~~)~~ é apontar ...

- Página 250
  Na 8ª linha de cima para baixo:
  As cercas são: $Q1 - \frac{3}{2}DIQ = 45{,}75 \text{ kg}$ e $Q3 + \frac{3}{2}DIQ = 138{,}5 \text{ kg}$

- Página 250
  Na descrição do processo de construção do Box plot:

  b) Depois, desenha-se um retângulo cuja **ordenada** da base inferior corresponde ao valor do primeiro quartil Q1 e cuja **ordenada** da base superior corresponde ao valor do terceiro quartil Q3. A **ordenada** da mediana é indicada por um traço horizontal no interior desse retângulo.

  c) Em seguida, são traçados dois segmentos de reta verticais que vão, um deles desde o ponto médio da base inferior do retângulo até a **ordenada** da menor observação não discrepante, e o outro desde o ponto médio da base superior do retângulo até a **ordenada** da maior observação não discrepante.

- Página 254
  No terceiro parágrafo da subseção 7.9.2:

  Consideremos, então, duas variáveis quantitativas contínuas, X e Y. Cada **unidade amostral** ~~-dado~~ a partir de uma amostra de tamanho n será representad**a** por um par ordenado (x_i , y_i) para i = 1,2,...,n , onde x_i e y_i são, respectivamente, a i-ésima observação de X e de Y.

- Página 258
  Logo após a figura, eliminar as frases a seguir:

  ~~- Sua teoria está errada!~~
  ~~- Fora, mentiroso!~~

- Página 259
  No último parágrafo antes do Exemplo 7.17:

  Observe que este tema está sendo apresentado aqui apenas a nível de análise exploratória. Considerações que envolvem Inferência Estatística p**o**dem ser encontradas na literatura sobre **A**nálise de **R**egressão.

- Página 263
  Nas Propriedades da correlação amostral (3º bullet):

  • O coeficiente de correlação mede **o grau de** ~~inter~~dependência linear entre as variáveis.

- Página 280
  No enunciado do Exercício P7.8:

  Finalmente, sejam $a = q_1(X) - 1{,}5\big(q_3(X) - q_1(X)\big)$ e $b = q_3(X) + 1{,}5\big(q_3(X) - 1(X)\big)$.

Cap8

- Página 301
  Na definição de viés, ao final da página:

  O **viés** de $\hat{\theta}$ é igual à diferença entre o seu valor esperado e o valor do parâmetro. Simbolicamente, $B(\hat{\theta}) = E(\hat{\theta}) - \theta\,\cancel{q}$, onde $B(\hat{\theta})$ é o viés do estimador $\hat{\theta}$.

- Página 306
  No item b):

  $$E(\hat{\theta}) = E\left(\overline{X} - \overline{Y}\right) = E(\overline{X}) - E(\overline{Y}) = \mu_1 - \mu_2 = \cancel{\hat{\theta}}\ \boldsymbol{\theta} \text{ mg/g.}$$

- Página 307
  No penúltimo parágrafo:

  ... Portanto, ~~nota-se~~ de novo ~~que~~, quando trabalhamos por amostragem, não é possível garantir que o erro absoluto será obrigatoriamente menor do que uma constante pré-fixada (d)**. P**ode-se, todavia, calcular a probabilidade de que isso ocorra.

- Página 308
  No terceiro parágrafo da seção 8.8:

  Na discussão a seguir, estamos admitindo que será usada a Amostragem Aleatória Simples, (sabidamente um bom procedimento amostral, que pode ser visto com detalhes, por exemplo, na referência **4**). Portanto, é suficiente que nos preocupemos apenas em dimensionar corretamente o tamanho da amostra.

- Página 318
  Nas Propriedades dos estimadores pontuais:

  ...
  • S^2 é um estimador não tendencioso para $Var(X) = \sigma^2$.

- Página 319
  No que se refere a Precisão dos estimadores:

  ...
  • $P[|\hat{p} - p| \leq d] \geq 2\Phi(2d\sqrt{n}) - 1$.
- Página 336
  No Exercício P8.12:

  Mostre que se $(X_1, Y_1), \ldots, (X_n, Y_n)$ são n vetores aleatórios bivariados **iid** tais que

$$Cov(X_i, Y_j) = \begin{cases} C, & \text{se } i=j \\ 0, & \text{caso contrário} \end{cases} \quad \ldots$$

Cap9

- Página 347
  Na primeira fórmula:

$$1 - \alpha = P\left[\hat{p} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} < p < \hat{p} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{\mathbf{n}}}\right]$$

- Página 349
  Na **Nota** ao final da página:

  No caso acima foram supostas distribuições Normais para X e para Y. Contudo, **mesmo que essa suposição não seja válida,** se n e m forem suficientemente grandes, podemos usar o Teorema Central do Limite para encontrar um intervalo de confiança aproximado.

- Página 350
  No segundo parágrafo da Situação 2:

  Assim, ~~o intervalo de 100(1 − α)% para $\mu_X - \mu_Y$ é agora~~ nessas condições

- Página 351
  No último parágrafo antes da Situação 3:

  Como os dois limites são negativos concluímos que existem evidências, ao nível de confiança de 0,95, para **se** afirmar que o verdadeiro conteúdo médio de enxofre é maior no carvão proveniente da mineradora B do que no carvão fornecido pela mineradora A. Por quê?

- Página 354
  No segundo parágrafo da página:

  Desta maneira, a amostra aleatória:

  $[(X_1 , Y_1), (X_2 , Y_2),\ldots,(X_n ,Y_n)]$

  reduz-se a

  $(\Delta_1 , \Delta_2 ,\ldots, \Delta_n)$, onde $\Delta_i = X_i - Y_i$ **, para i = 1,2,...,n,**

  ou seja , voltamos ao caso de uma única variável aleatória.

- Página 356
  Para construirmos o intervalo de confiança, **que só pode depender dos dados,** p_1 e p_2 são substituídos por $\hat{p}_1$ e $\hat{p}_2$, respectivamente. ...

- Página 361
  No último parágrafo da Solução do exercício R9.3:

  Notemos que o intervalo inclui apenas valores positivos, indicando uma tendência no sentido de que a média do primeiro grupo é maior do que a média do segundo grupo para um nível de confiança de 95%. Então ~~podemos afirmar~~ **a análise sugere** que o método que usa multimídia tende a produzir notas superiores às do método tradicional, ao nível de confiança dado.

- Página 362
  Em um parágrafo $\pm$ no meio da Solução do exercício R9.4:

  Para verificar se a premissa de Normalidade **da variável Δ** está sendo atendida aqui, você pode construir, por exemplo, um gráfico de probabilidade Normal (ver Capítulo 12).

Cap10

- Página 371
  Logo após **O que podemos dizer sobre as probabilidades de erro α e β?:**

  Vimos acima que há dois tipos possíveis de erro de decisão:

  - o Erro I, rejeitar H_0 quando **essa** hipótese é verdadeira, tem probabilidade α de ocorrer
  - o Erro II, aceitar H_0 quando **essa** hipótese é falsa, tem probabilidade β de ocorrer

- Página 371
  Imediatamente antes de **O Erro I no centro das atenções:**

  ... Assim, a única forma de fazer com que ambos, α e β, diminuíssem, seria aumentar o tamanho n da amostra. (Ver Seção 10.6 **e também o Exercício R10.2**)

- Página 373
  Na Figura 10.1, as marcações para $\mu_1 = 150$ e $\mu_0 = 180$ devem ficar alinhadas verticalmente com os pontos de máximo das duas curvas.

- Página 381
  Na definição de poder:

  O **poder de um teste estatístico** é a capacidade do teste rejeitar corretamente uma hipótese nula ~~e~~ falsa ou seja, a probabilidade de rejeitar a hipótese nula quando a hipótese alternativa for verdadeira.

- Página 384
  Na Solução do Exemplo 10.10:

  H_0: $p \leq$ **0,05** vs. H_1:$p > 0{,}05$, $p_0 =$ 0,05 e $\sigma_0 = \sqrt{\dfrac{0{,}05 \times 0{,}95}{100}} = 0{,}022$

- Página 385
  Na Solução do Exemplo 10.10 item (c):

  c) Se p = 0,12, então $\sqrt{\dfrac{p(1-p)}{n}} = \sqrt{\dfrac{0{,}12 \times 0{,}88}{100}} = 0{,}0325$ e $\hat{p}$ ~ N(0,12; 0,0325²). Portanto,...

- Página 386
  Na terceira curva da figura, a Região de Rejeição é:

  $$Z < \cancel{z_\alpha}\ -z_{1-\alpha}$$

- Página 389
  Na Solução do Exercício R10.2 item (b)

  Assim, teremos, para cada valor possível de $\bar{x}_c$, $\alpha = P\left[Z > \dfrac{\bar{x}_c - 50}{4/\sqrt{30}}\right]$ e $\beta = P\left[Z \leq \dfrac{\bar{x}_c - 52}{4/\sqrt{\mathbf{30}}}\right]$.

- Página 390
  Na Solução do Exercício R10.2 item (c)

  $$0{,}01 = P\left[\mathbf{Z >} \dfrac{\bar{x}_c - 50}{4/\sqrt{\mathbf{n}}}\right] \Rightarrow \dfrac{\bar{x}_c - 50}{4/\sqrt{n}} = 2{,}33. \qquad \text{(I)}$$

- Página 393
  No enunciado do Exercício P10.2:

  (a) Para um nível de significância de 5%, construa o teste de hipóteses adequado para dirimir essa controvérsia, especificando a estatística de teste e as regiões de aceitação e de rejeição.
  (b) Se os valores de teor de nicotina dos 16 cigarros da amostra coletada são:

0,718	0,703	0,692	0,792	0,657	0,679	0,706	0,719
0,673	0,682	0,665	0,684	0,770	0,761	0,699	0,728

  qual a decisão a ser tomada?
  (c) Qual é o p-valor neste caso?

Cap11

- Página 402
  Na Nota imediatamente antes do Exemplo 11.1, usar letra minúscula em z_p :

  Nota: Na exposição acima, **z_p** representa o quantil p da distribuição Normal Padrão, qualquer que seja p.

- Página 402
  No enunciado do Exemplo 11.1:

  Lembremos o Exemplo 9.4, no qual o engenheiro químico responsável por um processo produtivo quer verificar se o fato de se empregar um catalisador recentemente lançado no mercado em vez de empregar o catalisador **que está em uso** atualmente ...

- Página 408
  Imediatamente antes do Exercitando :

  Para verificar se a premissa de Normalidade está sendo atendida aqui, você pode construir, por exemplo, um gráfico de probabilidade Normal (Ver Capítulo 12) **para a variável Δ**.

- Página 409
  No último parágrafo da página:

  Quando fizemos inferências sobre a diferença de médias de duas variáveis aleatórias Normais e independentes foi utilizada, no caso de amostras pequenas, a distribuição t de *Student*, o que exigia a suposição de igualdade das variâncias das duas populações. Isto se aplica tanto **às** estimativas por intervalos quanto **aos** testes de hipóteses. Uma possível verificação do cumprimento dessa exigência é através da utilização de um gráfico de caixas (box-plot). Uma outra maneira de se proceder é por meio **da aplicação** de um teste de hipótese como o que veremos mais adiante. Ele faz uso da chamada distribuição F de *Fisher-Snedecor* que apresentamos a seguir.

- Página 411
  Na 2ª linha da página:

  $$\nu_1 = 10\ ,\ \ \nu_2 = 20\ \ \text{e}\ \ F_{p;\nu_1,\nu_2} = 4{,}92 \ \Rightarrow\ p = P(F \le F_p) > 0{,}995$$

- Página 417
  No 2º parágrafo da página:

  Quando o p-valor é menor que 0,05 porém maior do que 0,01, dizemos que a diferença observada é significativa. Se o **p-valor** é menor do que 0,01 – como ocorre no presente exemplo – costuma-se dizer que a diferença é altamente significativa.

- Página 423
  No 1º parágrafo da subseção 11.5.3:

  ... Esse procedimento é chamado de Teste de Aderência ou Teste da Qualidade do Ajustamento.

- Página 433
  Na Solução do Exercício R11.1 item (a):

  A estatística de teste é **$F = \frac{S_M^2}{S_T^2}$** , e sua distribuição de probabilidade...

Cap12

- Página 464
Imediatamente antes das Propriedades do R2:
  - S_e^2 é a variância amostral do erro **e**
  - S_y^2 é a variância amostral de **Y**
  - $r_{y\hat{y}}$ é o coeficiente de correlação amostral entre **Y** e $\hat{\mathbf{Y}}$, sendo $\hat{y}_i$ o chamado valor ajustado de y_i, que é calculado pela expressão: ...

- Página 486
Logo após a figura, eliminar as frases a seguir:

  "~~Pecados mortais: Não aleatorizou. Extrapolou. Aceitou a alternativa.~~"

- Página 491
Quase ao final da página:

  Usam-se transformações de variáveis para:
  - **simetrização** do perfil de freqüências de uma certa variável
  - linearizaçao da relação entre duas variáveis
  - homogeneização das variâncias de uma variável em diferentes grupos

- Página 514
Nos histogramas do Exercício P12.9:

  Inserir um espaço no início da última linha, para acertar o alinhamento vertical das 3 figuras.

- Página 533
No texto que está escrito no topo da tabela:

  Fornece os quantis $\mathbf{F_{0,975}}$ (em cima) e $\mathbf{F_{0,995}}$ (em baixo) em função do n^o de g.l. numerador ν_1 (coluna)

  e do n^o de g.l. denominador ν_2 (linha)

  F tem distribuição F com ν_1 g.l. no numerador e ν_2 g.l. no denominador

  $P(F \leq \mathbf{F_{0,975}}) = 0,975$ e $P(F \leq \mathbf{F_{0,995}}) = 0,995$

Contracapa:

Consertar a última frase:

No site do livro... as respostas de muitos dos exercícios propostos.