Introdução à Análise de dados em FAE **(29/03/2024)**

Lista de exercícios 2

Professores: Sandro Fonseca, Eliza Melo, Maurício Thiel e Diego Torres *Name:* Miguel Lopes

EXERCÍCIO 1

O primeiro exercíciopede para demonstrar a formula para o cálculo do erro de um função de duas variáveis que dependentes. Tendo a função $f(x, y)$ de duas variáveis x e y de incertezas associadas σ_x e σ_y.

Para isso, vamos fazer a expansão em serie de Taylor em torno dos valores médios de x e y. A primeira ordem é:

$$f(x, y) \approx f(\bar{x}, \bar{y}) + \left.\frac{\partial f}{\partial x}\right|_{\bar{x},\bar{y}} (x - \bar{x}) + \left.\frac{\partial f}{\partial y}\right|_{\bar{x},\bar{y}} (y - \bar{y}).$$

Sendo:$\Delta x = x - \bar{x}$ e $\Delta y = y - \bar{y}$. Representam desvios de dos valores médios de x e y.

$$f(x, y) \approx f(\bar{x}, \bar{y}) + \left.\frac{\partial f}{\partial x}\right|_{\bar{x},\bar{y}} \Delta x + \left.\frac{\partial f}{\partial y}\right|_{\bar{x},\bar{y}} \Delta y.$$

Para o erro de $f(x, y)$. Podemos dizer que $\Delta f = f(x, y) - f(\bar{x}, \bar{y})$ é o desvio de f do seu valor médio. O erro associado a f, σ_f, é a variação de Δf. Dessa forma:

$$\sigma_f^2 = \left(\left.\frac{\partial f}{\partial x}\right|_{\bar{x},\bar{y}}\right)^2 \sigma_x^2 + \left(\left.\frac{\partial f}{\partial y}\right|_{\bar{x},\bar{y}}\right)^2 \sigma_y^2 + 2\left(\left.\frac{\partial f}{\partial x}\right|_{\bar{x},\bar{y}}\right)\left(\left.\frac{\partial f}{\partial y}\right|_{\bar{x},\bar{y}}\right)\sigma_{xy}.$$

Sendo σ_{xy} a covariância entre x e y. Este último termo na equação é devido à covariância entre x e y. Se as variáveis são independentes, a covariância é zero.

EXERCÍCIO 2

Para obter o erro para o caso caso:

$$f = x + y.$$

$$\sigma_f = \sqrt{\sigma_x^2 + \sigma_y^2 + 2r\sigma_x\sigma_y}.$$

Basta aplicar a função f = x ± y na equação deduzida no exercício a cima.

$$\sigma_f^2 = \left(\left.\frac{\partial f}{\partial x}\right|_{\bar{x},\bar{y}}\right)^2 \sigma_x^2 + \left(\left.\frac{\partial f}{\partial y}\right|_{\bar{x},\bar{y}}\right)^2 \sigma_y^2 \pm 2\left(\left.\frac{\partial f}{\partial x}\right|_{\bar{x},\bar{y}}\right)\left(\left.\frac{\partial f}{\partial y}\right|_{\bar{x},\bar{y}}\right)\sigma_{xy}.$$

onde:

$$\frac{\partial f}{\partial x} = 1 \quad , \quad \frac{\partial f}{\partial y} = 1 \quad \text{e} \quad \sigma_{xy} = r\sigma_x\sigma_y.$$

Para o caso onde a equação é f = xy, podemos utilizar a equação a seguir, apenas realizando a derivada na nova função:

$$\sigma_f^2 = \left(\left.\frac{\partial f}{\partial x}\right|_{\bar{x},\bar{y}}\right)^2 \sigma_x^2 + \left(\left.\frac{\partial f}{\partial y}\right|_{\bar{x},\bar{y}}\right)^2 \sigma_y^2 \pm 2\left(\left.\frac{\partial f}{\partial x}\right|_{\bar{x},\bar{y}}\right)\left(\left.\frac{\partial f}{\partial y}\right|_{\bar{x},\bar{y}}\right)\sigma_{xy}.$$

E então dividir toda a equação por $(\bar{x}\bar{y})^2$

$$\sigma_f^2 = (\bar{y}\sigma_{\bar{x}})^2 + (\bar{x}\sigma_{\bar{y}})^2 + 2\bar{x}\bar{y}r\sigma_{\bar{x}}\sigma_{\bar{y}}.$$

$$\frac{\sigma_f}{\bar{f}} = \sqrt{(\frac{\sigma_{\bar{x}}}{\bar{x}})^2 + (\frac{\sigma_{\bar{y}}}{\bar{y}})^2 + 2\frac{r\sigma_{\bar{x}}\sigma_{\bar{y}}}{\bar{x}\bar{y}}}.$$

EXERCÍCIO 3

EXERCÍCIO 4

A média ponderada ($\bar{x}$) é:

$$\bar{x} = \frac{\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}$$

E o erro associado à média ponderada ($\sigma_{\bar{x}}$) é:

$$\sigma_{\bar{x}} = \sqrt{\frac{1}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}}$$

Para este cálculo vou utilizar os valores das massas e erros associados das seguintes partículas:

Partícula	Massa (GeV/c^2)	Erro associado (GeV/c^2)
W^+	80.377	0.012
W^-	80.377	0.012
Z^0	91.1876	0.0021
H	125.25	0.17
γ	≈ 0	≈ 0

Considerando a massa do $\gamma \approx 0$, temos que:

$$\bar{x} = \frac{\frac{80.377}{0.012^2} + \frac{80.377}{0.012^2} + \frac{91.1876}{0.0021^2} + \frac{125.25}{0.17^2}}{\frac{1}{0.012^2} + \frac{1}{0.012^2} + \frac{1}{0.0021^2} + \frac{1}{0.17^2}}$$
$$\bar{x} = 90,56865343$$

Para o erro, temos:

$$\begin{aligned}
\sigma_{\bar{x}} &= \sqrt{\frac{1}{\frac{1}{0.012^2} + \frac{1}{0.012^2} + \frac{1}{0.0021^2} + \frac{1}{0.17^2}}} \\
&= \sqrt{\frac{1}{\frac{1}{0.000144} + \frac{1}{0.000144} + \frac{1}{0.00000441} + \frac{1}{0.0289}}} \\
&= \sqrt{\frac{1}{6944.44 + 6944.44 + 226757.37 + 34.60}} \\
&= \sqrt{\frac{1}{240680.85}} \\
&\approx \sqrt{4.15487 \times 10^{-6}} \\
&\approx 0.002\,\mathrm{GeV/c^2}
\end{aligned}$$

Logo:

$$\bar{x} \approx 90,569 \pm 0.002$$

EXERCÍCIO 5

EXERCÍCIO 6

A começar pelos valores de a, b temos que expandir a função de ajuste dos mínimos quadrados:

$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

$$S(a,b) = \sum_{i=1}^{n} (y_i^2 - 2ax_iy_i - 2by_i + a^2x_i^2 + 2abx_i + b^2)$$

Fazendo o valor médio de x e y, a equação fica:

$$S(a,b) = (\bar{y}_i^2 + a^2\bar{x}_i^2 + b^2) - 2a\bar{x}_i\bar{y}_i - 2b\bar{y}_i + 2ab\bar{x}_i$$

Para encontrar os valores mínimos vamos derivar $S(a,b)$ em relação a a e b, e igualar os resultados a 0.

$$\frac{\partial S}{\partial a} = -2\bar{x}_i\bar{y}_i + 2a\bar{x}_i^2 + 2b\bar{x}_i = 0$$

$$\frac{\partial S}{\partial b} = -2\bar{y}_i + 2a\bar{x}_i + 2nb = 0$$

Dessa forma podemos resolvemos essas equações para obter os melhores valores para a e b no modelo dos mínimos quadrados. As soluções mais simplificadas para a e b são:

$$a = \frac{\sigma_{xy}}{\sigma_x^2} \qquad b = \bar{y} - a\bar{x}$$

EXERCÍCIO 7

No total temos 5419 estudante e queremos a amostragem de estudantes homens do curso da área de tecnologia. Podemos utilizar a formula a baixo e realizar a divisão entre o número de alunos homens que fazem a área de tecnologia. Onde A são os alunos homens de tecnologia e B é a quantidade total de alunos

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Onde:

$$P = \frac{1291}{5419} = 23,82\%$$

EXERCÍCIO 8

E essa probabilidade é definida usando o teorema de Bayes:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Onde:
- $P(A) = 0.15$ é a probabilidade de um táxi ser azul
- $P(B|A) = 0.80$ é a probabilidade de uma pessoa acertar o carro.
- Para $P(B)$ devemos usar a lei da probabilidade total:

$$P(B) = 0.80 \times 0.15 + 0.20 \times 0.85 = 0.29$$

Logo, a probabilidade do carro carro do acidente ser azul é de:

$$P(A|B) = \frac{0.80 \times 0.15}{0.27} \approx 0.444$$

EXERCÍCIO 9

Para determinar a probabilidade de que uma mulher esteja com câncer e receba o resultado de teste como positivo, vou considerar que todas as mulheres iram realizar o teste.

Sendo A o evento de uma mulher com câncer e B o evento de resultado positivo do teste.

- $P(B|A) = 1 - 0.10 = 0.90$, chance do teste ser positivo e a mulher ter câncer - $P(A) = 0.008$, chance de uma mulher ter câncer entre todas as mulheres. Logo:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Levando em consideração os possíveis testes falsos que dão positivo, $P(TB)$ será:

$$P(T) = 0,90 \times 0,008 + 0,07 \times 0,992 = 0,07664$$

Então:

$$P(A|B) = \frac{0,90 \times 0,008}{0,07664} \approx 0,094$$

Ou seja, a probabilidade de que uma mulher esteja com câncer ao receber um resultado de teste positivo é de 9,4%.

EXERCÍCIO 10

Novamente vamos utilizar o teorema de Bayes para calcular a probabilidade de uma bola preta ter sido retirada da terceira caixa.

Sendo

$$\begin{aligned} 1 &: \text{primeira urna.} \\ 2 &: \text{segunda urna.} \\ 3 &: \text{terceira urna.} \\ P &: \text{Retirar uma bola preta.} \end{aligned}$$

Pelo teorema de Bayes, temos:

$$P(3|P) = \frac{P(P|3) \cdot P(3)}{P(P)}$$

onde:

$$\begin{aligned} P(P|3) &: \text{Probabilidade de retirar uma bola preta da urna 3.} \\ P(3) &: \text{Probabilidade de selecionar a urna 3.} \\ P(P) &: \text{Probabilidade de escolher uma bola preta.} \end{aligned}$$

A probabilidade de uma bola preta ser escolhida vai ser:

$$P(B) = \frac{6}{11} \cdot \frac{1}{3} + \frac{5}{9} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3}$$

Logo, podemos dizer que:

$$P(3|P) = \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{6}{11} \cdot \frac{1}{3} + \frac{5}{9} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3}} \approx 0,312$$

Como resultado, a probabilidade de retirar uma bola preta da terceira urna é 31,2%.

EXERCÍCIO 11

EXERCÍCIO 12

EXERCÍCIO 13

Com o método binomial, para calcular a probabilidade de 2 pessoas aleatórias fazerem aniversário no mesmo dia, entre 720 pessoas, usamos:

$$P(X = k) = \binom{n}{k} \times p^k \times (1-p)^{n-k}$$

Onde:

P : (X = k) é a probabilidade.
n : é o número de pessoas total.
k : é o número de pessoas escolhidas.
p : é a probabilidade de duas pessoas terem o mesmo dia de aniversário.

$$P(X = 2) = \frac{720!}{2! \times (720-2)!} \times \left(\frac{1}{365}\right)^2 \times \left(\frac{364}{365}\right)^{718}$$

$$P(X = 2) = 25884 \times 1,04 \cdot 10^{-6} \times 0,1395$$

$$P(X = 2) = 0,271$$

Pelo método binomial, duas pessoas aleatórias teriam 27,1% de chance de fazerem aniversário no mesmo dia.

Agora calculando pelo método de Poisson temos que:

$$P(X = k) = \frac{e^{-\lambda} \times \lambda^k}{k!}$$

Onde:

P : é a probabilidade
k : é o número de pessoas escolhidasλ : é a média de eventos em um intervalo específico

$$\lambda = n \times p = 720 \times \frac{1}{365}$$

Então, aplicando os valores na expresão de Poisson temos:

$$P(X = 2) = \frac{e^{-\frac{720}{365}} \times \left(\frac{720}{365}\right)^2}{2!}$$

$$P(X = 2) = \frac{e^{-1,9726} \times 3,8911}{2!} = \frac{0,1390}{2!} = 0,069$$

Pelo método de Poisson a probabilidade sria de 6,9%.

EXERCÍCIO 14

(a) Para determinar a probabilidade de um aluno acertar 1, 2, 3, ..., 15 questões podemos usar o método binomial.

$$P(X = k) = \binom{n}{k} \times p^k \times (1-p)^{n-k}$$

Onde:

P : é a probabilidade de acerto.
n : é o número total de questões.
k : é o número de questões certas.
p : é a probabilidade de sucesso em uma tentativa.

Então, todos os cálculos ficam desta maneira:

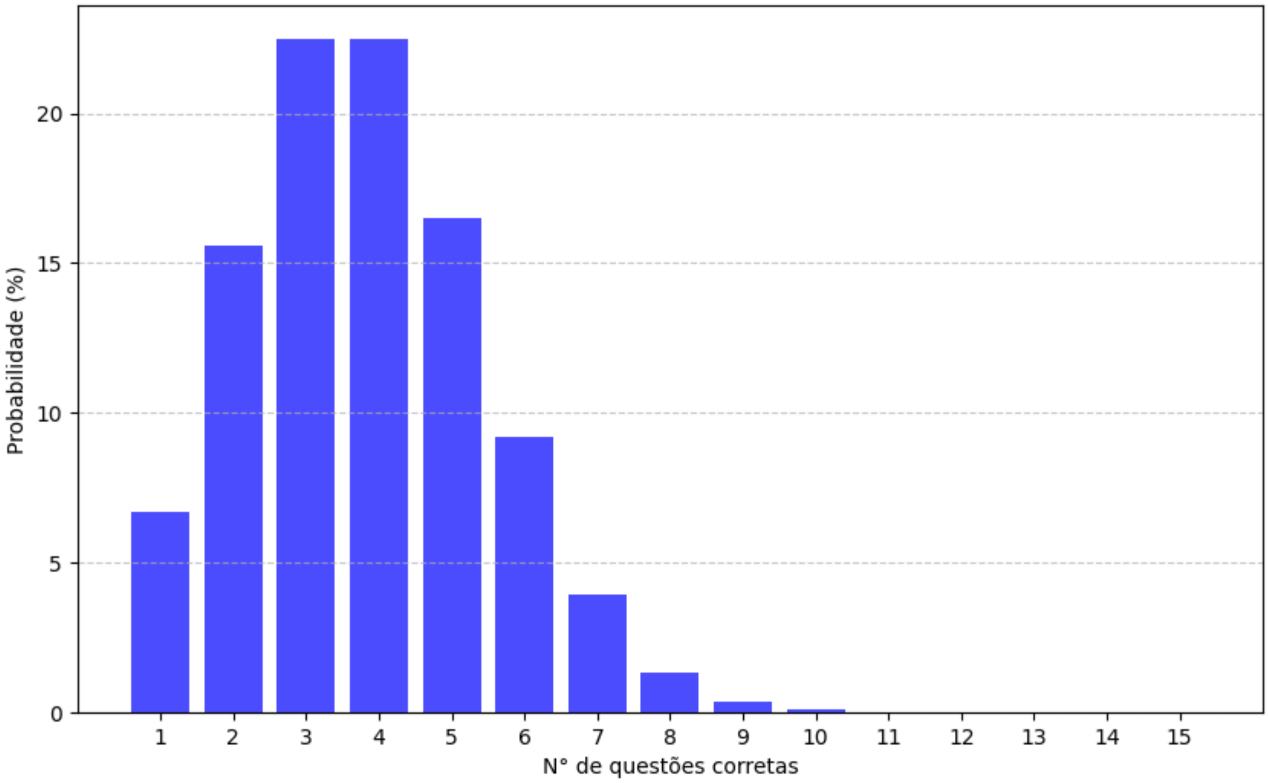
$$P(X=1) = \binom{15}{1} \times \left(\frac{1}{4}\right)^{1} \times \left(\frac{3}{4}\right)^{14}$$

$$P(X=2) = \binom{15}{2} \times \left(\frac{1}{4}\right)^{2} \times \left(\frac{3}{4}\right)^{13}$$

$$\vdots$$

Probabilidades		
$P_1 = 6,7\%$	$P_2 = 15,6\%$	$P_3 = 22,5\%$
$P_4 = 22,5\%$	$P_5 = 16,5\%$	$P_6 = 9,2\%$
$P_7 = 3,9$	$P_8 = 1,3\%$	$P_9 = 0,34$
$P_{10} = 0,07$	$P_{11} = 0,01$	$P_{12} = 0,001$
$P_{13} = 0,00009$	$P_{14} = 0,000004$	$P_{15} = 0,0000001$

(b)



(c)

Para calcular quantos alunos, de 1000 acertariam pelo menos 3 questões podemos multiplicar o número de alunos pela probabilidade de 1 aluno acertar as 3 questões.

$$1000 \times P_3 = 1000 \times 0,225199 = 225,2$$

Logo, 225 alunos acertariam pelo menos 3 questões.

EXERCÍCIO 15

EXERCÍCIO 16

Para calcular a porcentagem de peças com defeito podemos usar a distribuição normal padrão:

$$Z = \frac{X - \mu}{\sigma}$$

Onde:

X : é o diâmetro (0,491 cm ou 0,473 cm)
μ : é a média dos diametros
σ : desvio padrão

Usando a tabela de probabilidade para a distribuição normal padrão vamos conseguir calcular a porcentagem de peças com defeito.

FAzendo:

$$Z_1 = \frac{0,491 - 0,482}{0,004} = 2,25$$
$$Z_2 = \frac{0,473 - 0,482}{0,004} = -2,25$$

Usando a tabela, temos que a probabilidade do diâmetro da peça estar de dentro do intervalo é de 98,78%, o que nos da uma probabilidade de 1,22% para peças defeituosas.