Introdução à Análise de dados em FAE **(DATA)**

Exercício 2

Professores: NOME DOS PROFESSORES *Name:* Dalmo da Silva Dalto

TEXTO

EXERCICIO 1

1 Dedução da Propagação de Erros

Vamos deduzir a expressão para o erro propagado (Δu) em uma função $u = f(x, y)$ considerando que as incertezas em x e y estão correlacionadas.

Começamos com a expansão em série de Taylor até a primeira ordem em torno das médias $\langle x \rangle$ e $\langle y \rangle$:

$$f(x, y) \approx f(\langle x \rangle, \langle y \rangle) + \left(\frac{\partial f}{\partial x}\right)_{\langle x \rangle, \langle y \rangle} (x - \langle x \rangle) + \left(\frac{\partial f}{\partial y}\right)_{\langle x \rangle, \langle y \rangle} (y - \langle y \rangle)$$

O erro propagado (Δu) é dado por:

$$\Delta u = \sqrt{\left(\frac{\partial u}{\partial x}\sigma_x\right)^2 + \left(\frac{\partial u}{\partial y}\sigma_y\right)^2 + 2\text{cov}(\sigma_x, \sigma_y)\left(\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}\right)}$$

ou

$$\Delta u = \sqrt{\left(\frac{\partial u}{\partial x}\sigma_x\right)^2 + \left(\frac{\partial u}{\partial y}\sigma_y\right)^2 + \frac{2}{N}\left(\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}\right)\sigma_x\sigma_y}$$

onde $\text{cov}(\sigma_x, \sigma_y)$ é a covariância entre σ_x e σ_y, e as derivadas parciais são avaliadas nas médias $\langle x \rangle$ e $\langle y \rangle$.

EXERCICIO 2 a)

Vamos deduzir matematicamente que a média de u, denotada por $\overline{u}$, é igual à função da média de $\overline{x}$ e $\overline{y}$, denotada por $f(\overline{\overline{x}}, \overline{\overline{y}})$, onde $\overline{x}$ e $\overline{y}$ são as médias das variáveis x e y.

1. A média de u, denotada por $\overline{u}$, é dada por:

$$\overline{u} = \frac{1}{N}\sum_{i=1}^{N} u_i$$

onde N é o número de observações e u_i é o valor da função para a observação i.

2. Substituímos u pela função $f(x, y)$:

$$\overline{u} = \frac{1}{N}\sum_{i=1}^{N} f(x_i, y_i)$$

onde (x_i, y_i) são os pares de observações.

3. Agora, considerando as médias $\overline{x}$ e $\overline{y}$, a média de $f(x, y)$, denotada por $f(\overline{x}, \overline{y})$, é dada por:

$$f(\overline{x}, \overline{y}) = \frac{1}{N}\sum_{i=1}^{N} f(x_i, y_i)$$

4. Portanto, a média de u é igual à função da média de x e y:

$$\overline{u} = f(\overline{x}, \overline{y})$$

EXERCICIO 2 b)

Para $u = x + y$:

$$\sigma_u = \sqrt{(\frac{\partial u}{\partial x}\sigma_x)^2 + (\frac{\partial u}{\partial y}\sigma_y)^2 + 2r\sigma_x\sigma_y}$$
$$= \sqrt{\sigma_x^2 + \sigma_y^2 + 2r\sigma_x\sigma_y}$$

Para $u = x - y$:

$$\sigma_u = \sqrt{(\frac{\partial u}{\partial x}\sigma_x)^2 + (\frac{\partial u}{\partial y}\sigma_y)^2 + 2r\sigma_x\sigma_y}$$
$$= \sqrt{\sigma_x^2 + \sigma_y^2 - 2r\sigma_x\sigma_y}$$

Portanto para $u = x \pm y$

$$\sigma_u = \sqrt{\sigma_x^2 + \sigma_y^2 \pm 2r\sigma_x\sigma_y}$$

Para $u = xy$:
A propagação do erro (σ_u) para $u = xy$ é dada por:

$$\sigma_u = \sqrt{\left(\frac{\partial u}{\partial x}\sigma_x\right)^2 + \left(\frac{\partial u}{\partial y}\sigma_y\right)^2 + 2r\sigma_x\sigma_y}$$
$$= \sqrt{(y \cdot \sigma_x)^2 + (x \cdot \sigma_y)^2 + 2r\sigma_x\sigma_y}$$
$$= \sqrt{y^2\sigma_x^2 + x^2\sigma_y^2 + 2r\sigma_x\sigma_y}$$

Para $u = \frac{x}{y}$:
A propagação do erro (σ_u) para $u = \frac{x}{y}$ é dada por:

$$\sigma_u = \sqrt{\left(\frac{\partial u}{\partial x}\sigma_x\right)^2 + \left(\frac{\partial u}{\partial y}\sigma_y\right)^2 + 2r\sigma_x\sigma_y}$$
$$= \sqrt{\left(\frac{1}{y} \cdot \sigma_x\right)^2 + \left(\frac{-x}{y^2} \cdot \sigma_y\right)^2 + 2r\sigma_x\sigma_y}$$
$$= \sqrt{\frac{\sigma_x^2}{y^2} + \frac{x^2\sigma_y^2}{y^4} + 2r\sigma_x\sigma_y}$$

EXERCICIO 2 c)

Para $u = xy$:
A propagação do erro (σ_u) para $u = xy$ é dada por:

$$\sigma_u = \sqrt{\left(\frac{\partial u}{\partial x}\sigma_x\right)^2 + \left(\frac{\partial u}{\partial y}\sigma_y\right)^2 + 2r\sigma_x\sigma_y}$$
$$= \sqrt{(y \cdot \sigma_x)^2 + (x \cdot \sigma_y)^2 + 2r\sigma_x\sigma_y}$$
$$= \sqrt{y^2\sigma_x^2 + x^2\sigma_y^2 + 2r\sigma_x\sigma_y}$$

Para $u = \frac{x}{y}$:
A propagação do erro (σ_u) para $u = \frac{x}{y}$ é dada por:

$$\begin{aligned}
\sigma_u &= \sqrt{\left(\frac{\partial u}{\partial x}\sigma_x\right)^2 + \left(\frac{\partial u}{\partial y}\sigma_y\right)^2 + 2r\sigma_x\sigma_y} \\
&= \sqrt{\left(\frac{1}{y}\cdot\sigma_x\right)^2 + \left(\frac{-x}{y^2}\cdot\sigma_y\right)^2 + 2r\sigma_x\sigma_y} \\
&= \sqrt{\frac{\sigma_x^2}{y^2} + \frac{x^2\sigma_y^2}{y^4} + 2r\sigma_x\sigma_y}
\end{aligned}$$

EXERCICIO 3

A média ponderada $\overline{x}$ é definida como:

$$\overline{x} = \frac{\sum_{n=1}^{N} \frac{x_i}{\sigma_i^2}}{\sum_{n=1}^{N} \frac{1}{\sigma_i^2}}$$

Multiplicando e dividindo o numerador pelo denominador, obtemos:

$$\overline{x} = \frac{\frac{1}{\sum_{n=1}^{N} \frac{1}{\sigma_i^2}} \sum_{n=1}^{N} \frac{x_i}{\sigma_i^2}}{1}$$

O fator $\frac{1}{\sum_{n=1}^{N} \frac{1}{\sigma_i^2}}$ é apenas um fator de normalização que garante que os pesos somem a 1, tornando a expressão uma média ponderada. Assim, podemos reescrever a equação como:

$$\overline{x} = \sum_{n=1}^{N} w_i x_i$$

Onde $w_i = \frac{\frac{1}{\sigma_i^2}}{\sum_{n=1}^{N} \frac{1}{\sigma_i^2}}$ são os pesos associados a cada valor x_i. Essa é a definição padrão de média ponderada, onde os pesos são proporcionais ao inverso do quadrado dos desvios padrão.

Portanto, demonstramos que a média ponderada $\overline{x}$ é igual à média ponderada padrão.

Além disso:

$$\sigma_x = \sigma = \frac{1}{\sum_{n=1}^{N} \frac{1}{\sigma_i^2}}$$

EXERCICIO 4

Partícula	Laboratório A	Laboratório B	Compatibilidade	Resultado Combinado
Píon	139.57 MeV ± 0.04 MeV	139.57 MeV ± 0.05 MeV	Compatível	139.57 MeV ± 0.045 MeV
Próton	938.27 MeV ± 0.05 MeV	938.27 MeV ± 0.06 MeV	Compatível	938.27 MeV ± 0.055 MeV
Nêutron	939.57 MeV ± 0.04 MeV	939.57 MeV ± 0.05 MeV	Compatível	939.57 MeV ± 0.045 MeV
Kaon	497.61 MeV ± 0.04 MeV	497.61 MeV ± 0.05 MeV	Compatível	497.61 MeV ± 0.045 MeV
Múon	105.66 MeV ± 0.15 MeV	105.66 MeV ± 0.18 MeV	Compatível	105.66 MeV ± 0.165 MeV

EXERCICIO 5

Queremos demonstrar a expressão:

$$S(a,b) = \sum_{n=1}^{N} \left(\frac{y_i - y(x_i)}{\sigma_i}\right)^2$$

Vamos começar expressando os resíduos ponderados:

$$\text{Resíduo ponderado} = \frac{y_i - y(x_i)}{\sigma_i}$$

Então, a soma dos quadrados dos resíduos ponderados é dada por:

$$S(a,b) = \sum_{n=1}^{N} (\text{Resíduo ponderado})^2 = \sum_{n=1}^{N} \left(\frac{y_i - y(x_i)}{\sigma_i} \right)^2$$

Devido a $y(x_i) = ax_i + b$:

$$S(a,b) = \sum_{n=1}^{N} \left(\frac{y_i - (ax_i + b)}{\sigma_i} \right)^2$$

EXERCICIO 6

Dedução de $a = r\frac{\sigma_y}{\sigma_x}$:
1. A equação da regressão linear simples é $y = b + ax$.
2. Temos a fórmula para o coeficiente de correlação linear r:

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

3. Substituímos a na equação da regressão linear: $b = \bar{y} - b\bar{x}$.
4. Utilizamos a fórmula de b em termos de r, σ_x e σ_y:

$$a = r\frac{\sigma_y}{\sigma_x} = \frac{\sigma_{xy}}{\sigma_x^2}$$

5. Substituímos b na equação de a:

$$b = \bar{y} - r\frac{\sigma_y}{\sigma_x}\bar{x}$$

EXERCICIO 6 b)

Prova de $\sigma_a = \frac{\sigma_y}{\sigma_x\sqrt{n}}$:
1. A variância do coeficiente linear a na regressão linear simples é dada por:

$$\text{Var}(a) = \frac{\sigma^2}{n\sum(x_i - \bar{x})^2}$$

2. O desvio padrão do coeficiente linear a é então:

$$\sigma_a = \sqrt{\frac{\sigma^2}{n\sum(x_i - \bar{x})^2}}$$

3. Na regressão linear, a variância dos resíduos σ^2 está relacionada ao desvio padrão dos erros de predição ϵ_y pela equação:

$$\sigma^2 = \frac{\sum(\epsilon_y)^2}{n-2}$$

4. Substituindo a variância dos resíduos na expressão do desvio padrão σ_a:

$$\sigma_a = \sqrt{\frac{\frac{\sum(\epsilon_y)^2}{n-2}}{n\sum(x_i - \bar{x})^2}}$$

5. Simplificando a expressão, chegamos em:

$$\sigma_a = \frac{\sigma_y}{\sigma_x\sqrt{n}}$$

Portanto, provamos que $\sigma_a = \frac{\sigma_y}{\sigma_x \sqrt{n}}$.

EXERCICIO 6 c)

edução de $\sigma_b = \sigma_a \sqrt{x^2}$:
1. A variância do coeficiente angular b na regressão linear simples é dada por:

$$\text{Var}(b) = \frac{\sigma^2}{\sum(x_i - \bar{x})^2}$$

2. O desvio padrão do coeficiente angular b é então:

$$\sigma_b = \sqrt{\frac{\sigma^2}{\sum(x_i - \bar{x})^2}}$$

3. Substituímos a variância dos resíduos σ^2 pela relação $\sigma_a = \frac{\sigma_y}{\sigma_x \sqrt{n}}$:

$$\sigma_b = \sqrt{\frac{\sigma_a^2 \sum(x_i - \bar{x})^2}{\sum(x_i - \bar{x})^2}}$$

4. Simplificando a expressão, temos:

$$\sigma_b = \sigma_a \sqrt{\frac{\sum(x_i - \bar{x})^2}{\sum(x_i - \bar{x})^2}}$$

5. Como $\frac{\sum(x_i - \bar{x})^2}{\sum(x_i - \bar{x})^2} = 1$, concluímos que:

$$\sigma_b = \sigma_a \sqrt{1} = \sigma_a$$

Portanto, $\sigma_b = \sigma_a \sqrt{x^2}$.

EXERCICIO 6 d)

Prova de $\sigma_y = \sqrt{\frac{N}{N-2}(1 - r^2)\sigma^2}$:
1. Relação entre Variância dos Resíduos e Coeficiente de Determinação:

$$r^2 = 1 - \frac{\sigma^2}{\sigma_y^2}$$

2. Expressão para σ_y:

$$\sigma_y^2 = \frac{\sigma^2}{1 - r^2}$$

$$\epsilon_y = \sigma_y \sqrt{\left(\frac{N}{N-2}\right)(1 - r^2)}$$

Portanto, provamos que $\epsilon_y = \sqrt{\frac{N}{N-2}(1 - r^2)\sigma^2}$.

EXERCICIO 7

Detonando H o conjunto de alunos e M o conjunto das alunas. Portanto a probabilidade de sortear um aluno é igual ou uma aluno são, respectivamente,

$$P(H) = 2587/(2587 + 2832) \approx 0,48eP(M) = 0,52$$

.

Para um aluno ou uma aluna na área tecnológica são, respectivamente,

$$P(T|H) = 129/2587 \approx 0,5ep(T|M) = 547/2832 \approx 0.19.$$

Logo a probabildade de uma pessoa sorteada estudar em algum curso da área tecnológic é expressa como:

$$P(T) = P(H) * P(T|H) + P(T|M) * P(T|H) = 0,339$$

De acordo com a fórmula de Bayes, a probabilidade $P(H|T)$ é dada por:

$$P(H|T) = \frac{P(T|H) * P(H)}{P(T)} \approx 0,71$$

EXERCICIO 8

Sabendo que $P(A)$ é a probabilidade de ser azul e $P(V)$ a probabilidade de ser verde, a probabilidade de cada um, respectivamente, são dadas por:

$$P(A) = 0,15 => P(V) = 1 - P(A) = 0,85$$

As probabilidade condicionais de identificação são definadas como:

$$P(I_A|A) = 0,8eP(I_A|V) = 0,2$$

De acordo com a probabilidade de Bayes de ter sido azul o carro identificado:

$$P(A|I_A) = \frac{P(I_A|A) * P(A)}{P(I_A|A) * P(A) + P(V) * P(I_A|V)} \approx 0,41$$

EXERCICIO 8

Sabendo que a taxa de incidência ($P(C)$) igual a $0,0008$, a probabilidade $P(\overline{C})$ de que uma mulher qualquer não tenha a doença é igual a $0,992$.

Como a taxa de falso-positivos ($P(P|\overline{C})$) é igual a 0,07 e a de ocorrência de um resultado positivo quando a doença não existe:

$$P(P|C) = 1 - P(P|\overline{C}) = 1 - 0,1 = 0,9$$

Pela formula de Bayes, a probabilidade de ocorrência da doença quando o teste for positivo, $P(C,P)$, é determinada por:

$$P(C|P) = \frac{P(P|C) * P(C)}{P(P|C) * P(C) + P(P|\overline{C}) * P(\overline{C})} \approx 0,094$$

EXERCICIO 9

Como a escolha de qualquer uma das urnas é igualmente propavel, ou seja:

$$P(U_i) = \frac{1}{3} \quad (i = 1,2 \quad \text{ou} \quad 3)$$

$$U_1 = \{5B, 6P\} \quad U_2 = \{4B, 5P\} \quad U_3 = \{4B, 4P\}$$

Onde B é bola branca e P é bola preta. Portanto a probabilidade condicionais $P(P|U_i)$ de que a bola sorteada seja preta são dadas por:

$$P(P|U_1) = 6/11 \quad P(P|U_2) = 5/9 \quad P(P|U_3) = 1/2$$

Pela fórmula de bayes a probabilidade de $P(U_3|P)$ é dada por:

$$P(U_3|P) = \frac{P(P|U_3) * P(U_3)}{P(P|U_3) * P(U_3) + P(P|U_2) * P(U_2)P(P|U_1) * P(U_1} \approx 0,71$$

EXERCICIO 10

A densidade de probilidade é expressa por:

$$\rho(x) = Asin^2\frac{\pi}{a}x$$

Em que a constante de normalização é dada por:

$$< x >= A\int_0^a \sin^2\left(\frac{\pi}{a}x\right)dx = \frac{A}{2}\int_0^a (1-\cos(2\pi x/a))\,dx = 1 => A = \frac{2}{a}$$

O valor de $< x >$ é dada por:

$$\frac{2}{a}\int_0^a x\sin^2\left(\frac{\pi}{a}x\right) = \frac{1}{a}\{\int_0^a xdx - \int_0^a xcos(\frac{2\pi x}{a})\}dx$$

Como

$$\int_0^a xcos(\frac{2\pi x}{a})dx = 0$$

$$< x >= \frac{a}{2}$$

O valor médio quadrático é dado por:

$$< x^2 >= \frac{1}{a}\{\int_0^a x^2dx - \int_0^a x^2cos(\frac{2\pi x}{a})\}dx$$

$$= \frac{1}{a}\left\{\frac{a^3}{3} - \frac{a}{\pi}\int_0^a x\sin\left(\frac{2\pi x}{a}\right)dx\right\}$$

$$\left(\frac{a^2}{3} + \frac{a}{2\pi^2}x\cos\left(\frac{2x}{\pi}\right)\right)\Bigg|_a^b = \frac{a^2}{3} + \frac{a^2}{2\pi^2}$$

Sendo a variância $\sigma_x^2 =< x^2 > - < x >^2= \frac{a^2}{3} + \frac{a^2}{2\pi^2} - \frac{a^2}{4}$

$$\sigma_x = a\sqrt{1/12 + 1/\pi^2} = \frac{a}{2\sqrt{3}}\sqrt{1+\frac{6}{\pi^2}}$$

EXERCICIO 11

A eficiência para 4 dectetores é dada por:

$$B(3|4;0,6) = \binom{4}{3}\cdot 0,6^3 \cdot 0,4^1 = 4\cdot 0,6^3\cdot 0,4^1 = 0,3456$$

$$B(4|4;0,6) = \binom{4}{4}\cdot 0,6^4 \cdot 0,4^0 = 1\cdot 0,6^4\cdot 0,4^0 = 0,1296$$

Eficiência do sistema $= B(3|4;0,6) + B(4|4;0,6) = 0,3456 + 0,1296 = 0,4752 = 47,52\%$
A eficiência para 5 dectetores é dada por:

$$B(3|5;0,6) = \binom{5}{3}\cdot 0,6^3 \cdot 0,4^2 = 10\cdot 0,6^3\cdot 0,4^2 = 0,3456$$

$$B(4|5;0,6) = \binom{5}{4}\cdot 0,6^4 \cdot 0,4^1 = 5\cdot 0,6^4\cdot 0,4^1 = 0,2304$$

$$B(5|5;0,6) = \binom{5}{5}\cdot 0,6^5 \cdot 0,4^0 = 1\cdot 0,6^5\cdot 0,4^0 = 0,07776$$

A eficiência do sistema para pelo menos três pontos em câmaras distintas é a soma dessas probabilidades:
Eficiência do sistema $= B(3|5;0,6) + B(4|5;0,6) + B(5|5;0,6) = 0,3456 + 0,2304 + 0,07776 = 0,65376$

EXERCICIO 12

Não levando em conta os anos bissextos, o número total de possibilidades de que a data de anivesário de duas pessoas quaisquer possam coincidir durante um ano é dado por

$$365 * 365 = 365^2$$

A probabilidade de que um par de pessoas anivesariem em um dia qualquer do ano é dado por:

$$P(i,j) = \frac{365}{365^2} = \frac{1}{365}$$

Generalizando para M pessoas:

$$P(M) = \frac{1}{365^{M-1}}$$

Para anivesariar em um mesmo dia:

$$C_N^2 = \frac{N!}{2!(N-2)!} = \frac{N(N-1)}{2}$$

Segundo a distribuição binomial a probabilidade de que não ocorra coincidências de aniversários (m=0 sucessos), é dada por:

$$P(m=0) = (1 - \frac{1}{365})^{C_N^2}$$

A probabilidade para duas pessoas anivesariem em um mesmo dia pode ser expressa como:

$$P_{m=2} = 1 - (1 - \frac{1}{365})^{\frac{N(N-1)}{2}}$$

Se a probabilidade de que um par comemore no mesmo dia é muito pequena podemos aproximar para uma distribuição de poisson e expressar para um $m = 0$:

$$P(m=0) = e^{-\frac{C_N^2}{365}}$$

Portanto a probabilidade para que menos duas pessoas aniversariem em um mesmo dia é dado por:

$$P_{m=2} = 1 - e^{-\frac{N(N-1)}{730}}$$

Nas aproximações binomial e de Poisson é praticamente igual à unidade para mais de 60 pessoas, sendo igual a 0,5 para 23 pessoas.

N	**Binomial**	**Poisson**
10	$P_{m=2} \approx 0.1173$	$P_{m=2} \approx 0.1169$
20	$P_{m=2} \approx 0.2314$	$P_{m=2} \approx 0.2310$
30	$P_{m=2} \approx 0.3392$	$P_{m=2} \approx 0.3388$
40	$P_{m=2} \approx 0.4427$	$P_{m=2} \approx 0.4422$
50	$P_{m=2} \approx 0.5402$	$P_{m=2} \approx 0.5397$
60	$P_{m=2} \approx 0.6321$	$P_{m=2} \approx 0.6317$

EXERCICIO 13

Como a probabilidade de acerto (p) de uma questão é $\frac{1}{4}$ a distribuição de probabilidade de acertos ao acaso de m das 15 questões é dada por:

$$P_m = \frac{15!}{m!(15-m)!}(\frac{1}{4})^m(\frac{3}{4})^{15-m}$$

m	P_m
0	0,0134
1	0.0668
2	0.1559
3	0.2252
4	0.2252
5	0.1651
6	0.0917
7	0.0393
8	0.0131
9	0.0039
10	0.0068
11	1.029e-4
12	1.144e-5
13	8.8e-7
14	4.191e-10
11	9.313e-10

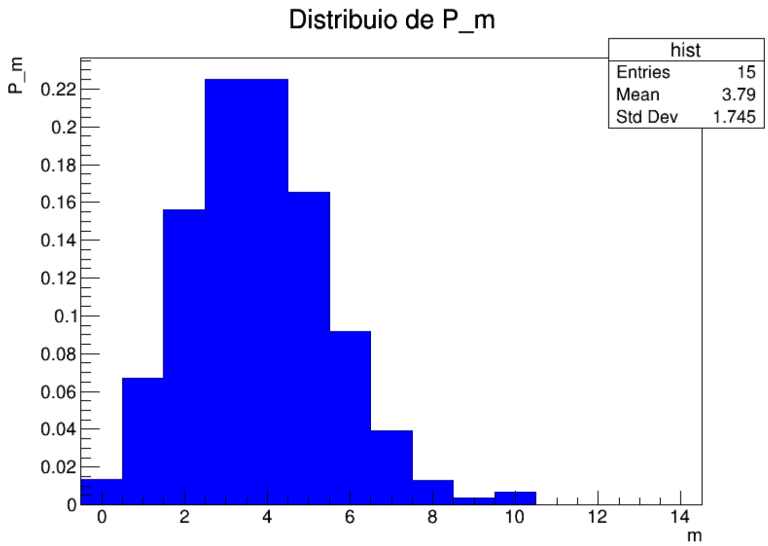

Figura 1: Distribuição de P_m.

A expectativa de erro para 3 questões é dada por:

$$P(m = 3) = 1 - (P_0 + P_1 + P_2) = 0,0764$$

Ou seja, em média, cerca de 764 deveriam acertar pelo 3 questões de 1000 alunos.

$$P(m = 3) = 1 - (P_0 + P_1 + P_2) = 0,0764$$

EXERCICIO 13

m	f_m
0	57
1	203
2	383
3	525
4	532
5	408
6	273
7	139
8	45
9	27
10	10
11	4
12	2
13	0
14	0

O número médio de contagens em cada intervalo de 7,5s é dado por:

$$\overline{m} = \frac{\sum_{m=0}^{14} m f_m}{2608} = 3,87$$

A probabilidade pode ser expressa por:

$$p = \frac{3,87}{N}$$

Assim, as contagens em cada intervalo obedecem a distribuição binomial:

$$2608 * B(m|N, p = 3,87/N) = 2608 * \frac{N!}{m!(N-m)!}(\frac{3,87}{N})^m(1 - \frac{3,87}{N})^{N-m}$$

Uma vez que $N >> 1$ e $p << 1$ e $m << N$, a distribuição tende a distribuição de Poisson:

$$2608\frac{3,87^m}{m!}e^{-3,87}$$

EXERCICIO 14

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-0.5(\frac{x-\mu}{\sigma})^2}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-z^2/2}dz = 0,9$$

implicando que:

$$z = \frac{x-\mu}{\sigma} = 1,29 => x = 176,4cm$$

EXERCICIO 15

$$z = \frac{0,491 - 0,482}{0,004} = \left|\frac{0,491 - 0,482}{0,004}\right| = 2,25$$

$$P(|z| < 2,25) = 2(1 - \phi(2,25)) => \phi(2,25) = 0,9878$$

Portanto

$$P(|z| < 2,25) = 2,44\%$$

EXERCICIO 16 a)

Prova da relação $\chi^2 = \left(\frac{\sigma_y}{\sigma}\right)^2(1 - r^2)$ para uma amostra heterocedástica:

$$
\begin{aligned}
\chi^2 &= \sum_{i=1}^{n} \left(\frac{y_i - \hat{y}_i}{\sigma_i} \right)^2 \\
&= \sum_{i=1}^{n} \frac{(y_i - (a + bx_i))^2}{\sigma_i^2} \\
&= \sum_{i=1}^{n} \frac{(y_i - a - bx_i)^2}{\sigma_i^2} \\
&= \frac{1}{\sigma^2} \sum_{i=1}^{n} (y_i - a - bx_i)^2 \\
&= \frac{1}{\sigma^2} \sum_{i=1}^{n} \left(y_i - \bar{y} + \bar{y} - a - bx_i + a + bx_i\right)^2 \\
&= \frac{1}{\sigma^2} \sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2 + 2b \sum_{i=1}^{n} x_i \left(\bar{y} - a\right) + b^2 \sum_{i=1}^{n} x_i^2 \\
&= \frac{1}{\sigma^2} \sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2 + b^2 \sum_{i=1}^{n} x_i^2 \\
&= \left(\frac{\sigma_y}{\sigma} \right)^2 (1 - r^2)
\end{aligned}
$$

EXERCICIO 16 b)

Prova da relação $\chi^2 = \left(\frac{\sigma_y^2}{\epsilon_y^2/N} \right)(1 - r^2)$ para uma amostra homocedástica:

$$
\begin{aligned}
\chi^2 &= \sum_{i=1}^{n} \left(\frac{y_i - \hat{y}_i}{\sigma_y} \right)^2 \\
&= \frac{1}{\sigma_y^2} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \\
&= \frac{1}{\sigma_y^2} \sum_{i=1}^{n} (y_i - (a + bx_i))^2 \\
&= \frac{1}{\sigma_y^2} \sum_{i=1}^{n} (y_i - a - bx_i)^2 \\
&= \frac{1}{\sigma_y^2} \sum_{i=1}^{n} (y_i - a - bx_i + a + bx_i - \hat{y}_i + \hat{y}_i)^2 \\
&= \frac{1}{\sigma_y^2} \sum_{i=1}^{n} (y_i - \bar{y} + \bar{y} - a - bx_i + a + bx_i - \hat{y}_i + \hat{y}_i)^2 \\
&= \frac{1}{\sigma_y^2} \sum_{i=1}^{n} [(y_i - \bar{y}) - b(x_i - \bar{x})]^2 \\
&= \frac{1}{\sigma_y^2} \sum_{i=1}^{n} (\epsilon_{yi})^2 \quad \text{[onde } \epsilon_{yi} = y_i - \bar{y} \text{ é o resíduo]} \\
&= \frac{n}{\epsilon_y^2} \sigma_y^2 \\
&= \left(\frac{\sigma_y^2}{\epsilon_y^2/N} \right) (1 - r^2)
\end{aligned}
$$

EXERCICIO 17

E(J)	$D(cm)$	$\epsilon_D(cm)$
0.07	4.9	0.3
0.18	6.7	0.3
0.3	7.3	0.4
0.45	8.1	0.4
0.69	9.2	0.4

Expressando a equação da reta como:

$$\ln D = n \ln E + \ln k$$

Sendo 0,25 o valor esperado de n, os dados podem ser tabelados e representados graficamente como:

x	y	ϵ_y
-2.659	1.589	0.061
-1.715	1.902	0.045
-1.204	1.988	0.055
-0.799	2.092	0.049
-0.371	2.219	0.043

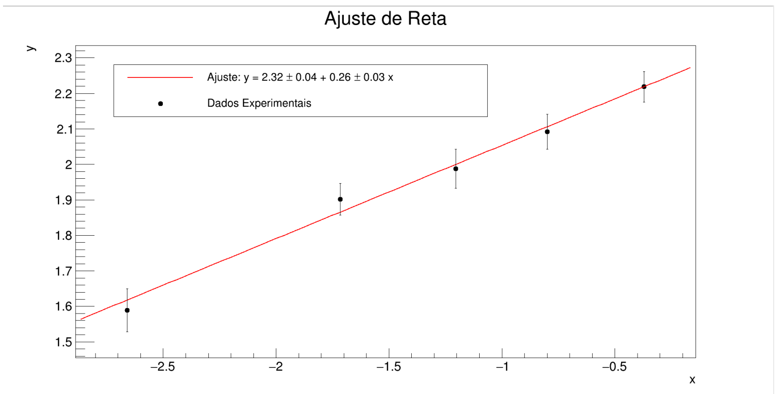

Figura 2: Ajuste de reta do exercício 17.

Com isso obtemos que a estimativa do expoente da relação esperada é dado por:

$$n \pm \sigma_n = (0,26 \pm 0,03)$$