

## MAE 229 – Introdução à Probabilidade e à Estatística II

## 2ª Lista de Exercícios – 2º semestre/2017 - FEA - Diurno

- 1) Certo processo industrial produz um grande número de cilindros de aço cujos comprimentos tem distribuição aproximadamente normal, com média de 3,25 metros e variância de 0,0008 metros quadrados. Justapõem-se dois cilindros escolhidos ao acaso. Qual a probabilidade do comprimento ser inferior a 6,55 metros ?

**Solução:** Seja  $X$ : Comprimento de um cilindro.  $X \sim N(3,25; 0,0008)$ . Seja  $Y = X_1 + X_2$  o comprimento de dois cilindros justapostos e portanto,  $Y \sim N(6,5; 0,0016)$ .

$$\begin{aligned} P(Y < 6,55) &= P\left(\frac{Y - 6,5}{\sqrt{0,0016}} < \frac{6,55 - 6,5}{\sqrt{0,0016}}\right) \\ &= P\left(\frac{Y - 6,5}{\sqrt{0,0016}} < \frac{6,55 - 6,5}{\sqrt{0,0016}}\right) \\ &= P(Z < 1,25) = 0,8943. \end{aligned}$$

- 2) Uma companhia utiliza anualmente milhares de lâmpadas elétricas que permanecem acesas continuamente, dia e noite. Admite-se que em tais condições, a vida de uma lâmpada pode ser considerada como uma variável distribuída normalmente com média de 50 dias e desvio padrão de 19 dias. Ao primeiro dia de janeiro a Companhia colocou em serviço 5000 lâmpadas novas. Qual o número esperado das que deveriam ser substituídas no primeiro dia de fevereiro ?

**Solução:**  $Y$  : Tempo de vida de uma lâmpada.  $Y \sim N(50, 19^2)$ .

$$\begin{aligned} E &= 5000P(1 < Y < 31) \\ &= 5000P\left(\frac{(1 - 50)}{19} < \frac{(Y - 50)}{19} < \frac{(31 - 50)}{19}\right) \\ &= 5000P(-2,568 < Z < -1) \\ &= 5000 \times 0,15354 = 767,7 \approx 768 \text{ lâmpadas} \end{aligned}$$

- 3) Admita que a população dos pesos dos passageiros de uma companhia de aviação tenha média de 150 libras e desvio padrão de 25 libras. Certo avião tem capacidade de 7800 libras. Qual a probabilidade de uma lotação de 50 passageiros ultrapassar aquele limite? Se quisermos reduzir para 0,01 a chance de sobrecarga, de quanto devemos aumentar a capacidade ?

**Solução:**  $X$  é o peso dos passageiros de uma companhia de aviação, em que  $X \sim N(150, 25^2)$ . Sabemos que o avião tem capacidade máxima de 7800 libras. Definimos,  $Y = \sum_{i=1}^{50} X_i$  como a soma dos pesos.  $Y \sim N(7500, 50 * 25^2)$ .

$$\begin{aligned} P(Y > 7800) &= P\left(\frac{y - 7500}{\sqrt{50 * 25^2}} > \frac{7800 - 7500}{\sqrt{50 * 25^2}}\right) \\ &= P\left(Z > \frac{7800 - 7500}{\sqrt{50 * 25^2}}\right) \\ &= P\left(Z > 1,697\right) = 1 - P\left(Z \leq 1,697\right) = 0,0448 \end{aligned}$$

Temos que

$$\begin{aligned} P(Y > d) &= 0,01 \\ P\left(Z > \frac{d - 7500}{\sqrt{50 * 25^2}}\right) &= 0,01 \end{aligned}$$

$$\begin{aligned} \frac{d - 7500}{\sqrt{50 * 25^2}} &= 2,33 \\ d &= 7500 + 2,33\sqrt{50 * 25^2} = 7911,89 \end{aligned}$$

- 4) Suponha que um agrimensor está procurando determinar a área de um campo retangular, em que as medidas do comprimento,  $X$ , e da largura,  $Y$ , são variáveis aleatórias independentes que flutuam em redor dos verdadeiros valores, de acordo com as seguintes distribuições de probabilidade:

|            |      |      |     |
|------------|------|------|-----|
| $X$        | 8    | 10   | 11  |
| $P(X = x)$ | 0,25 | 0,25 | 0,5 |

|            |     |     |
|------------|-----|-----|
| $Y$        | 4   | 6   |
| $P(Y = y)$ | 0,5 | 0,5 |

A área calculada  $A = XY$  é, naturalmente, uma variável aleatória, e é usada para estimar a área verdadeira. Se o verdadeiro comprimento e a verdadeira largura são 10 e 5, respectivamente :

- a)  $X$  é um estimador não viesado do verdadeiro comprimento ? **Solução:**  $E(X) = 8 * 0,25 + 10 * 0,25 + 11 * 0,5 = 10$ . Logo,  $X$  é não viesado para o comprimento.

- b)  $Y$  é um estimador não viesado da verdadeira largura ?

**Solução:**  $E(Y) = 4 * 0,5 + 6 * 0,5 = 5$ . Logo,  $Y$  é não viesado para a largura.

- c)  $A$  é um estimador não viesado da verdadeira área ?

**Solução:**  $E(A) = 24 * 0,125 + 40 * 0,125 + 44 * 0,125 + 48 * 0,125 + 60 * 0,25 + 66 * 0,25 = 49$ .

|            |       |       |      |       |       |      |
|------------|-------|-------|------|-------|-------|------|
| $A$        | 24    | 40    | 44   | 48    | 60    | 66   |
| $P(A = a)$ | 0,125 | 0,125 | 0,25 | 0,125 | 0,125 | 0,25 |

$$66 * 0,25 = 49u.m.(unidades de medida)$$

- d) Qual é a variância de  $A$  ?

$$\begin{aligned} Var(A) &= E(A^2) - E(A)^2 \\ &= (24^2 + 40^2 + 44^2 + 48^2 + 60^2 + 66^2) * 0,125 + (44^2 + 66^2) * 0,25 - 49^2 \\ &= 182u.q.(unidades quadráticas) \end{aligned}$$

- 5) Suponha que dois economistas estimem (despesa familiar média com alimentação) com dois estimadores não viesados  $\hat{\mu}_1$  e  $\hat{\mu}_2$  (estatisticamente independentes). O segundo economista é menos cuidadoso que o primeiro, de maneira que o desvio padrão de  $\hat{\mu}_2$  é cinco vezes maior do que o de  $\hat{\mu}_1$ . Solicitado a informar sobre como combinar  $\hat{\mu}_1$  e  $\hat{\mu}_2$  para obter uma estimativa global publicável, um grupo de estatísticos apresentou as seguintes propostas :

$$1) \hat{W}_1 = \frac{1}{2}(\hat{\mu}_1 + \hat{\mu}_2), 2) \hat{W}_2 = \frac{4}{5}\hat{\mu}_1 + \frac{1}{5}\hat{\mu}_2, 3) \hat{W}_3 = \frac{5}{6}\hat{\mu}_1 + \frac{1}{6}\hat{\mu}_2, 4) \hat{W}_4 = \hat{\mu}_1$$

- a) Quais dos estimadores são não viesados ?

**Solução:** Todos os estimadores são não viesados.

- b) Dispor esses estimadores por ordem decrescente de eficiência.

**Solução:** Sabendo que  $\sqrt{\text{var}(\hat{\mu}_2)} = 5\sqrt{\text{var}(\hat{\mu}_1)} \rightarrow \text{var}(\hat{\mu}_2) = 25\text{var}(\hat{\mu}_1)$ .

Fazendo os cálculos, teremos por ordem de eficiência,  $V(\hat{W}_4) = \text{var}(\hat{\mu}_1)$ ,

$V(\hat{W}_3) = \frac{50}{36}\text{var}(\hat{\mu}_1)$ ,  $V(\hat{W}_2) = \frac{41}{25}\text{var}(\hat{\mu}_1)$ , e  $V(\hat{W}_1) = \frac{26}{4}\text{var}(\hat{\mu}_1)$ .

- c) Determinar a combinação linear mais eficiente possível.

**Solução:** Seja  $W_5 = \theta\hat{\mu}_1 + (1 - \theta)\hat{\mu}_2$  o estimador mais eficiente possível. Temos que

$$\begin{aligned} \text{var}(W_5) &= \theta^2\text{var}(\hat{\mu}_1) + (1 - \theta)^2\text{var}(\hat{\mu}_2) \\ &= \theta^2\text{var}(\hat{\mu}_1) + 25(1 - \theta)^2\text{var}(\hat{\mu}_1) \end{aligned}$$

Derivando com relação a  $\theta$ ,  $\frac{d}{d\theta}\text{var}(W_5) = 2\theta\text{var}(\hat{\mu}_1) - 50(1 - \theta)\text{var}(\hat{\mu}_1)$ , igualando a zero e isolando  $\theta$  obteremos  $\theta = \frac{50}{52}$ . Portanto,  $W_5 = \frac{50}{52}\hat{\mu}_1 + \frac{2}{52}\hat{\mu}_2$ .

- 6) Um fazendeiro tem um campo quadrado cuja área deseja estimar. Quando ele mede o comprimento do campo, comete um erro aleatório, de modo que o comprimento observado  $X$  é uma variável aleatória normal centrada em  $\mu$  (valor verdadeiro porém desconhecido), com desvio padrão  $\sigma$ . Cômescio deste erro possível, ele decide fazer duas observações independentes de  $X$ ,  $X_1$  e  $X_2$  e tirar a média. Mas ainda tem dúvidas como proceder.

- (1) Tirar a média de  $X_1$  e  $X_2$  e elevar ao quadrado? Ou
- (2) Calcular primeiro o quadrado e, então, tirar a média dos resultados ?
- a) Qual método acusa o menor viés ?

**Solução:**  $Z_{(1)} = \left(\frac{X_1+X_2}{2}\right)^2 = \bar{X}^2$  e  $Z_{(2)} = \frac{X_1^2+X_2^2}{2}$ . Deste modo,

$$\begin{aligned} E(Z_{(1)}) &= E(\bar{X}^2) = E(\bar{X})^2 + Var(\bar{X}) = \mu^2 + \frac{\sigma^2}{2} \\ E(Z_{(2)}) &= E\left(\frac{X_1^2 + X_2^2}{2}\right) = \frac{E(X_1)^2 + Var(X_1) + E(X_2)^2 + Var(X_2)}{2} \\ &= \frac{2\mu^2 + 2\sigma^2}{2} = \mu^2 + \sigma^2 \end{aligned}$$

Portanto, o método 1 fornece o menor viés.

- b) Generalizar a resposta em a) para uma amostra de tamanho  $n$ .

**Solução:** Para uma amostra de tamanho  $n$  temos que os estimadores serão:

$$E(Z_{(1)}) = \mu^2 + \frac{\sigma^2}{n} \text{ e } E(Z_{(2)}) = \mu^2 + \sigma^2.$$

- c) Como estimador alternativo para a área, qual o viés de  $X_1.X_2$  ?

**Solução:**

$$E(X_1.X_2) \stackrel{ind}{=} E(X_1)E(X_2) = \mu^2$$

Portanto,  $X_1.X_2$  é um estimador não viesado, ou seja, seu viés é zero.

- 7) Faz-se uma pesquisa sobre o consumo (em milhares de reais) de uma amostra aleatória de 25 famílias com mesmo padrão de renda. Seja 8,2 a média amostral. Se  $\sigma = 0,72$ , construir um intervalo de confiança de 95% para a média de consumo de todas as famílias com este padrão de renda.

**Solução:** Sabendo que  $\bar{X} \sim N(\mu, \sigma^2/25)$

$$P(-1,96 \leq \frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} \leq 1,96) = 0,95$$

Assim, para  $\bar{x} = 8,2$  e  $\sigma = 0,72$  e  $n = 25$ ,

$$IC(\mu, 95\%) = \left( \bar{x} - 1,96 \frac{\sigma}{\sqrt{n}}; \bar{x} + 1,96 \frac{\sigma}{\sqrt{n}} \right) = (7,92; 8,48)$$

- 8) Para os tempos de reação de 30 motoristas selecionados aleatoriamente, encontrou-se uma média de 0,83 Segundos e um desvio padrão de 0,2 segundos. Determinar um intervalo de 95% de confiança para o tempo médio de reação de toda a população de motoristas utilizando a aproximação normal.

**Solução:** Seguindo os passos da questão 7 e usando  $\bar{x} = 0,83$  e  $\sigma = 0,2$  e  $n = 30$ , obtemos  $IC(\mu, 95\%) = (0,76; 0,90)$ .

- 9) De uma classe muito grande de Estatística, extraíram-se aleatoriamente as 40 notas :

|    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|
| 71 | 74 | 65 | 72 | 64 | 42 | 62 | 62 | 58 | 82 |
| 49 | 83 | 58 | 65 | 68 | 60 | 76 | 86 | 74 | 53 |
| 78 | 64 | 55 | 87 | 56 | 50 | 71 | 58 | 57 | 75 |
| 58 | 86 | 64 | 56 | 45 | 73 | 54 | 86 | 70 | 73 |

Construir um intervalo de 90% de confiança para a nota média de toda a classe. (Reduzir o seu trabalho a proporções convenientes agrupando os dados em classes de amplitude 5).

**Solução**

| Classe de Salários | Proporção | Freq. abs. |
|--------------------|-----------|------------|
| 42 ┤ 47            | 0,050     | 2          |
| 47 ┤ 52            | 0,050     | 2          |
| 52 ┤ 57            | 0,125     | 5          |
| 57 ┤ 62            | 0,150     | 6          |
| 62 ┤ 67            | 0,175     | 7          |
| 67 ┤ 72            | 0,1       | 4          |
| 72 ┤ 77            | 0,175     | 7          |
| 77 ┤ 82            | 0,025     | 1          |
| 82 ┤ 87            | 0,150     | 6          |

Temos que  $\bar{x} = 44,5 \times 0,05 + 49,5 \times 0,05 + \dots + 84,5 \times 0,175 = 66,375$  e  $S^2 = \frac{40}{39}\sigma^2 = \frac{40}{39}\left(44,5^2 \times 0,05 + 49,5^2 \times 0,05 + \dots + 84,5^2 \times 0,175 - 66,375^2\right) = 132,93$ . Assim,  $T = \frac{(\bar{X}-\mu)\sqrt{n}}{S} \sim t - Student(39)$ . Em que,

$$P(-t_\gamma < \frac{(\bar{X} - \mu)\sqrt{n}}{S} < t_\gamma) = \gamma = 0,9.$$

O intervalo de confiança é  $IC(\mu, \gamma) = \left(\bar{x} - t_\gamma \frac{S}{\sqrt{n}}; \bar{x} + t_\gamma \frac{S}{\sqrt{n}}\right)$ . Como número de graus de liberdade é grande, podemos aproximar para a distribuição normal, ou seja,  $IC(\mu, \gamma) = \left(\bar{x} - z_\gamma \frac{S}{\sqrt{n}}; \bar{x} + z_\gamma \frac{S}{\sqrt{n}}\right)$ . Fazendo isto, teremos  $IC(\mu, 90\%) = \left(66,375 - 1,65 \frac{\sqrt{132,93}}{\sqrt{40}}; 66,375 + 1,65 \frac{\sqrt{132,93}}{\sqrt{40}}\right) = (63,367; 69,382)$

- 10) Uma amostra aleatória de 120 trabalhadores de uma grande fábrica leva em média 22 minutos para executar determinado serviço, com variância de 4. Em uma segunda fábrica, para executar a mesma tarefa, uma amostra aleatória de 120 operários gasta em média 19 minutos, com variância 10. Construir um intervalo de 95% de confiança para a diferença entre as médias das duas populações.

### Solução

Sabendo que  $X \sim N(\mu_X, 4)$ ,  $Y \sim N(\mu_Y, 10)$  em que,  $\bar{X} \sim N(\mu_X, 4/120)$ ,  $\bar{Y} \sim N(\mu_Y, 10/120)$  e  $\bar{X} - \bar{Y} \sim N(\mu_X - \mu_Y, 14/120)$ ,

$$P(-1,96 \leq \frac{(\bar{X} - \bar{Y} - (\mu_X - \mu_Y))\sqrt{n}}{\sigma_X + \sigma_Y} \leq 1,96) = 0,95.$$

Assim, para  $\bar{x} = 22$ ,  $\bar{y} = 19$ ,  $\sigma_X^2 = 4$ ,  $\sigma_Y^2 = 10$  e  $n_X = n_Y = 120$ ,

$$IC(\mu_X - \mu_Y, 95\%) = \left(\bar{x} - \bar{y} - 1,96 \frac{(\sigma_X + \sigma_Y)}{\sqrt{n}}; \bar{x} - \bar{y} + 1,96 \frac{(\sigma_X + \sigma_Y)}{\sqrt{n}}\right) = (2,33; 3,67)$$

- 11) Considere o modelo  $Y_t = \mu + \epsilon_i$ , onde  $\mu$  é um parâmetro desconhecido.

- a) Com base numa amostra  $Y_1, Y_1, Y_2, \dots, Y_n$  obtenha o estimador de mínimos quadrados de  $\mu$ .

**Solução:** Para estimar  $\mu$  pelo método de mínimos quadrados devemos minimizar  $S(\mu) = \sum_{i=1}^n \epsilon^2 = \sum_{i=1}^n (y_i - \mu)^2$ . Derivando  $S(\mu)$  com relação a  $\mu$  e igualando a zero, teremos  $\hat{\mu} = \bar{y}$ .

- b) Determine o valor desse estimador se, para uma amostra de 5 elementos, os valores de  $\mathbf{Y}_t$  foram: 3, 5, 6, 8, 16. **Solução:**  $\hat{\mu} = 7,6$

- 12) Antes de uma eleição, um determinado partido está interessado em estimar a proporção  $p$  de eleitores favoráveis ao seu candidato. Uma amostra piloto de tamanho 100 revelou que 60% dos eleitores eram favoráveis ao candidato em questão.

- a) Determine o tamanho da amostra necessário para que o erro cometido na estimação seja de, no máximo, 0,01 com probabilidade de 80%.

**Solução:** Sabemos que  $n \approx \frac{z_\gamma^2 p(1-p)}{\epsilon^2} = \frac{1,28^2 \times 0,6 \times 0,4}{0,01^2} \approx 3932$ , em que  $z_\gamma = z_{80\%} = 1,28$ .

- b) Se na amostra final, com tamanho igual ao obtido em (a), observou-se que 55% dos eleitores eram favoráveis ao candidato em questão, construa um intervalo de confiança para a proporção  $p$  (Utilize  $\gamma = 0,95$ .)

**Solução:** Temos que  $\hat{p} = 55\%$ ,  $n = 3932$  e sabemos que  $IC(p, 95\%) = \left( \hat{p} - z_\gamma \sqrt{\hat{p}(1-\hat{p})/n}; \hat{p} + z_\gamma \sqrt{\hat{p}(1-\hat{p})/n} \right)$ . Assim,  $IC(p, 95\%) = (0,534; 0,565)$ .

- 13) Suponha que estejamos interessados em estimar a porcentagem de consumidores de um certo produto. Se uma amostra de tamanho 300 forneceu 100 indivíduos que consomem o dado produto, determine:

- a) O intervalo de confiança de  $p$ , com coeficiente de confiança de 95% (interprete o resultado).

**Solução:** Usando o mesmo procedimento do item b da questão 12, teremos que  $IC(p, 95\%) = (0,28; 0,39)$ . Isto quer dizer que, se a pesquisa fosse realizada 100 vezes, esperasse que em 95 delas o valor de  $\hat{p}$  pertença ao intervalo  $(0,28; 0,39)$ .

- b) O tamanho da amostra para que o erro da estimativa não exceda a 0,02 unidades com probabilidade de 95% (interprete o resultado).

**Solução:** Temos que  $n \approx \frac{1,96^2}{4 \times 0,02^2} = 2401$ , em que  $z_\gamma = z_{95\%} = 1,96$ . Assim, para uma amostra de 2401, o erro da estimativa será de 0,02 com confiança de 95%.

- 14) No modelo  $y_t = \alpha + \beta x_t^2 + \epsilon_t$  com base numa amostra de  $n$  pares de observações  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  determine os estimadores de mínimos quadrados de  $\alpha$  e  $\beta$ .

**Solução** Para estimar pelo método de mínimos quadrados devemos minimizar com relação aos parâmetros a seguinte função  $S(\alpha, \beta) = \sum_{i=1}^n (y_i - (\alpha + \beta x_i^2))^2$ . Derivando parcialmente com relação aos parâmetros e igualando a zero, obteremos as seguintes equações para formar o sistema,  $\hat{\alpha} = \frac{\sum y_i - \hat{\beta} \sum x_i^2}{n}$  e  $\hat{\beta} = \frac{\sum y_i x_i^2 - \hat{\alpha} \sum x_i^2}{\sum x_i^4 - (\sum x_i^2)^2}$ . Resolvendo o sistema, encontramos os seguintes resultados:

$$\hat{\alpha} = \bar{y} - \hat{\beta} \frac{\sum_{i=1}^n x_i^2}{n} \text{ e } \hat{\beta} = \frac{n \sum_{i=1}^n y_i x_i^2 - \sum_{i=1}^n y_i \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n x_i^4 - (\sum_{i=1}^n x_i^2)^2}.$$

- 15) Numa pesquisa de mercado para estudar a preferência da população de uma cidade em relação a um determinado produto, colheu-se uma amostra aleatória de 300 indivíduos, dos quais 180 preferiam esse produto.

- a) Determine um intervalo de confiança para a proporção da população que prefere o produto em estudo.

**Solução:** Assumindo um nível de confiança de 95%, o intervalo será  $IC(p, 95\%) = (0,54; 0,65)$

- b) Determine a probabilidade de que a estimativa pontual dessa proporção não difira do verdadeiro valor em mais de 0,001.

**Solução:**

$$\begin{aligned} P(|p - \hat{p}| < 0,001) &= P(-0,001 < \hat{p} - p < 0,001) \\ &= P\left(\frac{-0,001\sqrt{300}}{\sqrt{18/30(12/30)}} < \frac{(\hat{p} - p)\sqrt{n}}{\sqrt{\hat{p}(1 - \hat{p})}} < \frac{0,001\sqrt{300}}{\sqrt{18/30(12/30)}}\right) \\ &= P(-0,035 < Z < 0,035) = 2 \times P(0 < Z < 0,035) \\ &= 0,027 \end{aligned}$$

- c) É possível obter uma estimativa pontual dessa proporção que não difira do valor verdadeiro em mais de 0,0005 com probabilidade 0,95? Caso contrário, determine o que deve ser feito.

**Solução:** Para obter esta estimativa é necessário uma amostra de tamanho  $n = 38.416$ . Neste caso, o tamanho da amostra pode representar alto custo. Uma alternativa seria aumentar o erro entre o verdadeiro valor e a estimativa. Um possível valor seria 0,01.

- 16) Se  $X$  tem distribuição de Poisson com parâmetro  $\lambda$ , determine o estimador de máxima verossimilhança de  $\lambda$ .

**Solução**

$$f(x_1, x_2, \dots, x_n, \lambda) = \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}$$

Aplicando o logaritmo natural,

$$\log(f(x_1, x_2, \dots, x_n, \lambda)) = \sum_{i=1}^n x_i \log(\lambda) - n\lambda - \log\left(\prod_{i=1}^n x_i!\right)$$

Derivando com relação a  $\lambda$  e igualando a zero,

$$\frac{1}{\hat{\lambda}} \sum_{i=1}^n x_i - n = 0$$

$$\hat{\lambda} = \bar{x}$$

- 17) Seja  $X$  a Quantidade de um certo produto (em milhares de unidades) e  $Y$  o respectivo custo total de produção, avaliados para uma amostra de 10 observações:

|                     |   |    |    |    |    |    |    |    |    |    |
|---------------------|---|----|----|----|----|----|----|----|----|----|
| $X$ (1000 unidades) | 1 | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
| $Y$ (R\$1.000,00)   | 7 | 11 | 15 | 14 | 18 | 21 | 23 | 30 | 32 | 34 |

Obtenha os valores dos estimadores de mínimos quadrados para o custo fixo e para o custo marginal. Ajustado o modelo  $\hat{y} = \hat{\alpha} + \hat{\beta}x_i$ ,  $\hat{\alpha}$  é o estimador do custo fixo e  $\hat{\beta}$  o estimador do custo marginal).

**Solução:** Para estimar pelo método de mínimos quadrados devemos minimizar com relação aos parâmetros a seguinte função  $S(\alpha, \beta) = \sum_{i=1}^n (y_i - (\alpha + \beta x_i))^2$ . Derivando

parcialmente com relação aos parâmetros e igualando a zero, obteremos as seguintes equações para formar o sistema,  $\hat{\alpha} + \hat{\beta}\bar{x} = \bar{y}$  e  $\hat{\alpha} \sum_{i=1}^n x_i + \hat{\beta} \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i$ . Resolvendo o sistema, encontramos os seguintes resultados:

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \text{ e } \hat{\beta} = \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}.$$

Fazendo os cálculos com os dados da tabela, obteremos:  $\hat{\alpha} = 15$  e  $\hat{\beta} = 1$ .

- 18) Encontre o coeficiente de confiança de um intervalo de confiança para  $p$ , se  $n = 100$  e  $\hat{p} = 0,6$  a amplitude do intervalo deve ser igual a 0,090.

**Solução:**

$$\begin{aligned} P(|p - \hat{p}| < 0,09) &= P(-0,09 < \hat{p} - p < 0,09) \\ &= P\left(\frac{-0,09\sqrt{100}}{\sqrt{0,6 \times 0,4}} < \frac{(\hat{p} - p)\sqrt{n}}{\sqrt{\hat{p}(1 - \hat{p})}} < \frac{0,09\sqrt{100}}{\sqrt{0,6 \times 0,4}}\right) \\ &= P(-1,84 < Z < 1,84) = 2 \times P(0 < Z < 1,84) \\ &= 2 \times 0,467 = 0,934 \end{aligned}$$

- 19) Suponha que uma amostra de  $n = 100$  de uma distribuição normal  $N(\mu, \sigma^2)$  forneceu  $\bar{X} = 510,6$ . Supondo  $\sigma^2$  conhecido e igual a 16, obtenha um intervalo de confiança para  $\mu$ , com coeficiente de confiança  $\gamma = 0,95$ .

**Solução:** Seguindo os passos da questão 7 e usando  $\bar{x} = 510,6$ ,  $\sigma^2 = 16$  e  $n = 100$ , obtemos  $IC(\mu, 95\%) = (509,81; 511,38)$ .