DM DEPARTAMENTO DE MATEMÁTICA TÉCNICO LISBOA

Probabilidades e Estatística

TODOS OS CURSOS

2º Semestre – 2022/2023
28/07/2023
10:30–12:30

Duração: **120** minutos **Exame Época Especial**

- As questões de escolha múltipla têm uma e apenas uma resposta correta. Caso assinale uma resposta errada, serão descontados 0.5 valores.
- Se assinalar mais do que uma resposta numa questão de escolha múltipla, o resultado será classificado como errado e serão descontados 0.5 valores.
- As respostas às questões de escolha múltiplas têm de ser assinaladas na folha de instruções previamente distribuída.
- Para as restantes questões, terá de justificar convenientemente as suas respostas.

Pergunta 1 2 valores

Um sistema electrónico com três componentes funciona sempre que todas as componentes estão operacionais; funciona com probabilidade 0.7 quando exactamente duas componentes estão operacionais; e nunca funciona quando menos de duas componentes estão operacionais.

Sabendo que cada uma das componentes está operacional com probabilidade 0.9, independentemente das restantes, calcule a probabilidade de o sistema estar em funcionamento.

- **V.a. auxiliar**

  X = número de componentes que estão operacionais, em três existentes

  $X \sim \text{binomial}(n = 3, p = 0.9)$, em virtude de cada uma das componentes estar operacional com probabilidade 0.9, independentemente das restantes.

  $P(X = x) = \binom{3}{x} 0.9^x (1-0.9)^{3-x}, \quad x = 0,1,2,3$

- **Acontecimentos e probabilidades**

Acontecimento	**Probabilidade**
$A_0 = \{X = 0\}$	$P(A_0) = (1-0.9)^3 = 0.001$
$A_1 = \{X = 1\}$	$P(A_1) = 3 \times 0.9 \times (1-0.9)^2 = 0.027$
$A_2 = \{X = 2\}$	$P(A_2) = 3 \times 0.9^2 \times (1-0.9) = 0.243$
$A_3 = \{X = 3\}$	$P(A_3) = 0.9^3 = 0.729$
F = “o sistema está em funcionamento”	$P(F) = ?$
	$P(F \mid A_0) = P(F \mid A_1) = 0$
	$P(F \mid A_2) = 0.7$
	$P(F \mid A_3) = 1$

- **Cálculo da probabilidade pedida**

  Ao aplicar-se a lei da probabilidade total, temos

$$
\begin{aligned}
P(F) &= \sum_{i=0}^{3} P(F \mid A_i) \times P(A_i) \\
&= P(F \mid A_2) \times P(A_2) + P(F \mid A_3) \times P(A_3) \\
&= 0.7 \times 0.243 + 1 \times 0.729 \\
&= 0.1701 + 0.729 \\
&= 0.8991.
\end{aligned}
$$

Pergunta 2	2 valores

Seja X uma variável aleatória discreta que se distribui uniformemente no conjunto $\{1/8, 1/4, 3/8\}$.

Determine o valor esperado e a mediana de X.

- **V.a. e f.p.**

  $X \sim$ uniforme discreta$(\{1/8, 1/4, 3/8\})$

  $P(X = x) = \frac{1}{3}, \quad x = 1/8, 1/4, 3/8$

- **Valor esperado e mediana de** X

$$E(X) = \sum_{x \in \{1/8,1/4,3/8\}} x \times P(X = x) = (1/8 + 1/4 + 3/8) \times \frac{1}{3} = \frac{1}{4}$$

$$me = me(X) \in \{1/8, 1/4, 3/8\} \quad : \quad \frac{1}{2} \leq F_X(me) \leq \frac{1}{2} + P(X = me) \quad \Leftrightarrow \quad F_X(me^-) \leq \frac{1}{2} \leq F_X(me)$$

  Uma vez que a f.d. de X, $F_X(x) = \sum_{x_i \leq x} P(X = x_i)$ toma os valores 1/3, 2/3 e 1, para $x = 1/8, 1/4, 3/8$ (respectivamente), segue-se

$$1/3 = F_X(1/4^-) = F_X(1/8) \leq \frac{1}{2} \leq F_X(1/4) = 2/3,$$

  pelo que $me = me(X) = 1/4$.

Pergunta 3	2 valores

Seja X a variável aleatória relativa ao tempo de espera de um passageiro (arbitrário) pelo comboio. Admita que X possui distribuição exponencial com valor esperado igual a 1 minuto.

Calcule a probabilidade do passageiro, que está à espera há 2 minutos pelo comboio, ainda ter de esperar 4 minutos adicionais pela chegada do mesmo. **A**: 0.9817 **B**: 0.1353 **C**: 0.8647 **D**: 0.0183

- **V.a. de interesse, f.d.p. e f.d.**

  $X =$ tempo de espera (em minutos) de um passageiro pelo comboio

  $X \sim$ exponencial(λ)

  $f_X(x) \stackrel{form.}{=} \begin{cases} \lambda\, e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$

  $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\, dt = \begin{cases} 0, & x < 0 \\ \int_0^x \lambda\, e^{-\lambda t}\, dt = 1 - e^{-\lambda x}, & x \geq 0 \end{cases}$

  $\lambda > 0 : E(X) = 1 \Leftrightarrow \frac{1}{\lambda} = 1 \Leftrightarrow \lambda = 1.$

- **Probabilidade pedida**

  Ao invocarmos a propriedade da falta de memória da distribuição exponencial, obtemos

$$P(X > 2 + 4 \mid X > 2) \quad = \quad P(X > 4) = 1 - F_X(4) = e^{-4} \simeq 0.0183.$$

  [Em alternativa, $P(X > 2 + 4 \mid X > 2) = \frac{P(X>2+4, X>4)}{P(X>2)} = \frac{P(X>2+4)}{P(X>2)} = \frac{e^{-(2+4)}}{e^{-2}} = e^{-4} \equiv 1 - F_X(4) \simeq 0.018316.$]

Pergunta 4	2 valores

Admita que os números diários de defeitos em peças produzidas por uma pequena unidade fabril são variáveis aleatórias independentes com distribuição comum de Poisson com valor esperado igual a 1.8.

Obtenha um valor aproximado da probabilidade de o número total de defeitos em peças produzidas em 50 dias exceder 100 defeitos.

- **V.a.; valor esperado e variância comuns**

  X_i = número de defeitos detetados em peças produzidas no dia i, $\quad i = 1, \ldots, n$

  $n = 50 > 30$

  $X_i \overset{i.i.d.}{\sim} X \sim \text{Poisson}(\lambda = 1.8), \quad i = 1, \ldots, n$

  $E(X_i) = E(X) = \mu = \lambda = 1.8$

  $V(X_i) = V(X) = \sigma^2 = \lambda = 1.8$

- **V.a. de interesse**

  $S_n = \sum_{i=1}^n X_i$ = número total de defeitos detetados em peças produzidas em n dias

- **Valor esperado e variância de S_n**

  $E(S_n) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) \overset{X_i \sim X}{=} n\,E(X) = n\,\mu$

  $V(S_n) = V\left(\sum_{i=1}^n X_i\right) \overset{X_i\ indep.}{=} \sum_{i=1}^n V(X_i) \overset{X_i \sim X}{=} n\,V(X) = n\,\sigma^2$

- **Distribuição aproximada de S_n**

  Segundo o teorema do limite central (TLC),

$$\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \frac{S_n - n\,E(X)}{\sqrt{n\,V(X)}} = \frac{S_n - n\,\mu}{\sqrt{n\sigma^2}} \overset{a}{\sim} \text{normal}(0,1).$$

- **Probabilidade pedida (valor aproximado)**

$$\begin{aligned} P(S_n > 100) &= 1 - P(S_n \le 100) = 1 - P\left(\frac{S_n - n\,\mu}{\sqrt{n\sigma^2}} \le \frac{100 - n\,\mu}{\sqrt{n\sigma^2}}\right) \overset{TLC}{\simeq} 1 - \Phi\left(\frac{100 - n\,\mu}{\sqrt{n\sigma^2}}\right) \\ &= 1 - \Phi\left(\frac{100 - 50 \times 1.8}{\sqrt{50 \times 1.8}}\right) \simeq 1 - \Phi(1.05) \overset{tabelas/calc}{=} 1 - 0.8531 = 0.1469. \end{aligned}$$

Pergunta 5	2 valores

Se X e Y forem duas variáveis aleatórias independentes e com função de probabilidade comum dada por $P(X = x) = P(Y = x) = (1-p)^x\,p$, para $x = 0, 1, 2, \ldots$, então a variável aleatória $X + Y$ possui função de probabilidade dada por $P(X + Y = z) = (1+z)\,(1-p)^z\,p^2$, para $z = 0, 1, 2, \ldots$.

Deduza a função de probabilidade de X condicional a $X + Y = z$, onde z está fixo e é um inteiro não negativo.

- **Par aleatório; f.p. marginais; f.p. conjunta**

  (X, Y)

  $P(X = x) = P(Y = x) = (1-p)^x\,p, \quad x = 0, 1, 2, \ldots$

  Uma vez que X e Y são v.a. independentes, temos

$$P(X = x, Y = y) \quad = \quad P(X = x) \times P(Y = y) = (1-p)^x\,p \times (1-p)^y\,p = (1-p)^{x+y}\,p^2, \quad x, y \in \mathbb{N}_0.$$

- **F.p. condicional pedida**

$$P(X = x \mid X + Y = z) \quad = \quad \frac{P(X = x, X + Y = z)}{P(X + Y = z)} = \frac{P(X = x, Y = z - x)}{P(X + Y = z)} = \frac{(1-p)^{x+z-x}\,p^2}{(1+z)\,(1-p)^z\,p^2}$$

$$P(X = x \mid X+Y=z) \quad = \quad \frac{1}{z+1}, \quad x = 0,1,\ldots,z \quad (z \in \mathbb{N}_0).$$

[I.e., $(X \mid X+Y=z) \sim$ uniforme discreta$(\{0,1,\ldots,z\})$, para z fixo em $\mathbb{N}_0$.]

Pergunta 6	2 valores

Admita que a variável aleatória X possui função de probabilidade

$$P(X=x) = \begin{cases} \binom{x+r-1}{x}(1-p)^x\, p^r, & x = 0,1,2,\ldots \\ 0, & \text{caso contrário,} \end{cases}$$

onde r é um inteiro positivo conhecido e p é um parâmetro desconhecido no intervalo $(0,1)$.

Deduza o estimador de máxima verosimilhança de p, baseado numa amostra aleatória de dimensão n proveniente de X.

- **V.a. de interesse; f.p.**

  X

  $$P(X=x) = \begin{cases} \binom{x+r-1}{x}(1-p)^x\, p^r, & x = 0,1,2,\ldots \\ 0, & \text{caso contrário,} \end{cases}$$

- **Parâmetro desconhecido**

  $p \quad (p \in [0,1])$

- **Amostra; amostra aleatória**

  $\underline{x} = (x_1,\ldots,x_n)$, amostra de dimensão n proveniente de X.

  $\underline{X} = (X_1,\ldots,X_n)$, amostra aleatória de dimensão n proveniente de X.

- **Dedução do estimador de MV de p**

  Passo 1 — Função de verosimilhança

  $$\begin{aligned} L(p \mid \underline{x}) &= P(\underline{X} = \underline{x}) \overset{X_i\, indep}{=} \prod_{i=1}^n P(X_i = x_i) \overset{X_i \sim X}{=} \prod_{i=1}^n P(X = x_i) \\ &= \prod_{i=1}^n \left[\binom{x_i + r - 1}{x_i} (1-p)^{x_i}\, p^r \right] \\ &= \prod_{i=1}^n \left[\binom{x_i + r - 1}{x_i} \right] \times p^{r\,n} (1-p)^{\sum_{i=1}^n x_i}, \quad p \in [0,1] \end{aligned}$$

  Passo 2 — Função de log-verosimilhança

  $$\ln L(p|\underline{x}) \quad = \quad \sum_{i=1}^n \ln \left[\binom{x_i + r - 1}{x_i} \right] + r\,n\ln(p) + \ln(1-p) \times \sum_{i=1}^n x_i, \quad p \in [0,1]$$

  Passo 3 — Maximização

  A estimativa de MV de p é representada por $\hat{p}$ e

  $$\hat{p} \quad : \quad \begin{cases} \left.\frac{d\ln L(p|\underline{x})}{dp}\right|_{p=\hat{p}} = 0 & \text{(ponto de estacionaridade)} \\ \left.\frac{d^2 \ln L(p|\underline{x})}{dp^2}\right|_{p=\hat{p}} < 0 & \text{(ponto de máximo)} \end{cases}$$

  $$\begin{cases} \frac{r\,n}{\hat{p}} - \frac{\sum_{i=1}^n x_i}{1-\hat{p}} = 0 \\ -\frac{r\,n}{\hat{p}^2} - \frac{\sum_{i=1}^n x_i}{(1-\hat{p})^2} < 0 \quad \text{(prop. verdadeira)} \end{cases} \overset{\hat{p} \neq 0,1}{\Leftrightarrow} \begin{cases} r\,n - \hat{p}\left(r\,n + \sum_{i=1}^n x_i\right) = 0 \Leftrightarrow \hat{p} = \frac{r\,n}{r\,n + \sum_{i=1}^n x_i} \\ - \end{cases}$$

Passo 4 — Estimador de MV de p

$$EMV(p) \quad = \quad \frac{r\,n}{r\,n + \sum_{i=1}^{n} X_i} = \frac{r}{r + \bar{X}}$$

Pergunta 7 2 valores

A massa (em kg) dos aparelhos de ondas de choque radiais portáteis é uma variável aleatória X com distribuição normal com valor esperado desconhecido μ e variância igual a $\sigma^2 = 2.1$.

Determine a dimensão mínima da amostra que garante que o intervalo de confiança a 96% para μ possua amplitude inferior ou igual a 0.7. **A**: 93 **B**: 10 **C**: 73 **D**: 9

- **V.a. de interesse**

  X = massa (em kg) de aparelho de ondas de choque radiais portátil

- **Situação**

  $X \sim \text{normal}(\mu, \sigma^2)$

  $\mu = E(X)$ desconhecido

  $\sigma^2 = V(X) = 2.1$

- **Obtenção do IC para μ; dimensão mínima pedida**

  Passo 1 — Variável aleatória fulcral para μ

  $$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \text{normal}(0, 1)$$

  Passo 2 — Quantis de probabilidade

  Dado que $(1-\alpha) \times 100\% = 96\%$, usaremos os quantis

  $$(a_\alpha, b_\alpha) \qquad \begin{cases} a_\alpha = \Phi^{-1}(\alpha/2) = \Phi^{-1}(0.02) = -\Phi^{-1}(1-0.02) \stackrel{tabela/calc.}{=} -2.0537 \\ b_\alpha = \Phi^{-1}(1-\alpha/2) = \Phi^{-1}(0.98) \stackrel{tabela/calc.}{=} 2.0537. \end{cases}$$

  Passo 3 — Inversão da desigualdade $a_\alpha \le Z \le b_\alpha$

  $$P(a_\alpha \le Z \le b_\alpha) = 1 - \alpha$$

  $$P\left[a_\alpha \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le b_\alpha\right] = 1 - \alpha$$

  $$P\left[\bar{X} - b_\alpha \times \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} - a_\alpha \times \frac{\sigma}{\sqrt{n}}\right] = 1 - \alpha$$

  Passo 4 — Concretização

  Atendendo à expressão geral do IC para μ, $IC_{(1-\alpha)\times 100\%}(\mu) = [\bar{x} \pm \Phi^{-1}(1-\alpha/2) \times \sigma/\sqrt{n}]$, temos

  $$IC_{96\%}(\mu) \quad = \quad \left[\bar{x} - 2.0537 \times \frac{\sqrt{2.1}}{\sqrt{n}},\ \bar{x} + 2.0537 \times \frac{\sqrt{2.1}}{\sqrt{n}}\right].$$

- **Dimensão mínima pedida**

  Este intervalo de confiança possuirá amplitude não superior a 0.7 desde que

  $$n \quad : \quad \left[\bar{x} + 2.0537 \times \frac{\sqrt{2.1}}{\sqrt{n}}\right] - \left[\bar{x} - 2.0537 \times \frac{\sqrt{2.1}}{\sqrt{n}}\right] \le 0.7 \quad \Leftrightarrow \quad 2 \times 2.0537 \times \frac{\sqrt{2.1}}{\sqrt{n}} \le 0.7$$

$$n \quad : \quad \sqrt{n} \geq 2 \times 2.0537 \times \frac{\sqrt{2.1}}{0.7} \quad \Leftrightarrow \quad n \geq 2^2 \times 2.0537^2 \times \frac{2.1}{0.7^2} \simeq 72.3031.$$

Assim sendo, a dimensão mínima da amostra que garante que a amplitude do $IC_{96\%}(\mu)$ seja inferior ou igual a 0.7 é $n_{\min} = 73$.

Pergunta 8 2 valores

Suponha que o comprimento (em mm) de uma peça é uma variável aleatória com distribuição normal com valor esperado μ e desvio padrão σ desconhecidos.

Teste a hipótese $H_0 : \sigma^2 = 4$ contra $H_1 : \sigma^2 > 4$, num dia em que foram recolhidas casualmente 21 peças e a média e a variância corrigida amostrais foram de $\bar{x} = 99.89$ e $s^2 = 6.7374$. Decida com base no valor-p.

- **V.a. de interesse**

  X = comprimento de uma peça metálica

- **Situação**

  $X \sim \text{normal}(\mu, \sigma^2)$

  μ desconhecido

  σ DESCONHECIDO

- **Hipóteses**

  $H_0 : \sigma^2 = \sigma_0^2 = 4 \quad \text{vs.} \quad H_1 : \sigma^2 > \sigma_0^2$

- **Estatística de teste**

  $\frac{(n-1)\,S^2}{\sigma_0^2} \sim_{H_0} \chi^2_{(n-1)}$

- **Região de rejeição de H_0 (para valores da estatística de teste)**

  O teste é unilateral superior ($H_1 : \sigma^2 > \sigma_0^2$), logo a região de rejeição de H_0 é do tipo $W = (c, +\infty)$.

- **Decisão (com base no valor-p)**

  Atendendo a que o valor observado da estatística de teste e o valor-p são iguais a

$$\begin{aligned} t &= \frac{(n-1)\,s^2}{\sigma_0^2} = \frac{(21-1) \times 6.7374}{4} = 33.687 \\ valor-p &= P(T > t \mid H_0) = 1 - F_{\chi^2_{(n-1)}}(t) = 1 - F_{\chi^2_{(20)}}(33.687) \overset{calc.}{=} 0.0283, \end{aligned}$$

  é suposto:

  - não rejeitar H_0 a qualquer n.s. $\alpha_0 \leq valor-p = 2.83\%$, designadamente ao nível usual de significância de 1%;
  - rejeitar H_0 a qualquer n.s. $\alpha_0 > valor-p = 2.83\%$, nomeadamente aos níveis usuais de significância de 5% e 10%.

  [Alternativamente, poderíamos recorrer às tabelas de quantis da distribuição do qui-quadrado e adiantar um intervalo para o valor-p:

$$\begin{aligned} F^{-1}_{\chi^2_{(20)}}(0.95) = 31.41 &< t = 33.687 < 34.17 = F^{-1}_{\chi^2_{(20)}}(0.975) \\ 0.95 &< F_{\chi^2_{(20)}}(33.687) < 0.975 \\ 1 - 0.975 &< 1 - F_{\chi^2_{(20)}}(33.687) < 1 - 0.95 \\ 0.025 &< valor-p < 0.05. \end{aligned}$$

Assim, é suposto:

- não rejeitarmos H_0 a qualquer n.s. $\alpha_0 \leq 2.5\%$, designadamente ao n.u.s. de 1%;
- rejeitarmos H_0 a qualquer n.s. $\alpha_0 \geq 5\%$, por exemplo, aos n.u.s. de 5% e 10%.]

Pergunta 9 2 valores

Conjectura-se que o peso (em kg) de uma saca de laranjas segue uma distribuição normal de valor esperado 6 e desvio padrão unitário (H_0).

Por forma a testar esta hipótese de ajustamento, recolheu-se uma amostra casual de dimensão 50, tendo-se obtido a seguinte tabela de frequências:

Massa da saca	≤ 5.8	$]5.8, 6.2]$	> 6.2
Frequência absoluta observada	20	26	4
Frequência absoluta esperada sob H_0	21.04	7.92	21.04

Teste H_0 ao nível de significância de 5%.

- **V.a. de interesse**

  X = massa da saca

- **Hipóteses**

  $H_0 : X \sim \text{normal}(6,1)$ vs. $H_1 : X \not\sim \text{normal}(6,1)$

- **Nível de significância**

  $\alpha_0 = 5\%$

- **Estatística de teste**

$$T = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \stackrel{a}{\sim}_{H_0} \chi^2_{(k-1)},$$

  k = no. de classes = 3; O_i = frequência absoluta observável da classe i; E_i = frequência absoluta esperada, sob H_0, da classe i.

- **Região de rejeição de H_0** (para valores de T)

  Tratando-se de um teste de ajustamento, a região de rejeição de H_0 é o intervalo à direita $W = (c, +\infty)$, onde

$$c = F^{-1}_{\chi^2_{(k-\beta-1)}}(1-\alpha_0) = F^{-1}_{\chi^2_{(3-1)}}(1-0.05) \stackrel{tabela/calc.}{=} 5.991.$$

- **Decisão**

i	Classe i	Freq. abs. obs. o_i	Freq. abs. esp. sob H_0 E_i	Parcelas valor obs. estat. teste $\frac{(o_i-E_i)^2}{E_i}$
1	≤ 5.8	20	21.04	$\frac{(20-21.04)^2}{21.04} \simeq 0.0511$
2	$]5.8, 6.2]$	26	7.92	41.274
3	> 6.2	4	21.04	13.800
		$\sum_{i=1}^k o_i = n = 50$	$\sum_{i=1}^k E_i = n = 50$	$t = \sum_{i=1}^k \frac{(o_i-E_i)^2}{E_i} \simeq 55.125$

Uma vez que $t \simeq 55.125 \in W = (5.991, +\infty)$, devemos rejeitar H_0 ao n.s. de $\alpha_0 = 5\%$ [bem como a qualquer outro n.s. superior a α_0].

Pergunta 10 2 valores

Pretende-se estudar a relação entre a percentagem de hidrocarbonetos presentes no condensador principal de uma unidade de destilação (x) e a pureza do oxigénio produzido (Y, em percentagem). Para o efeito recolhidas 20 observações que conduziram aos seguintes resultados:

$$\sum_{i=1}^{20} x_i = 23.92, \quad \sum_{i=1}^{20} x_i^2 = 29.2892, \quad \sum_{i=1}^{20} y_i = 1843.21, \quad \sum_{i=1}^{20} y_i^2 = 170044.5321, \quad \sum_{i=1}^{20} x_i y_i = 2214.6566,$$

onde $[\min_{i=1,\dots,20} x_i, \max_{i=1,\dots,20} x_i] = [0.87, 1.55]$.

Admita que as variáveis x e Y estão relacionadas de acordo com o modelo de regressão linear simples, $Y = \beta_0 + \beta_1 x + \epsilon$.

Estime $E(Y \mid x = 1.20)$. Calcule o coeficiente de determinação e interprete o valor obtido.

- **[Modelo de RLS e hipóteses de trabalho**

  Y = pureza do oxigénio produzido (em percentagem) (v.a. resposta)

  x = percentagem de hidrocarbonetos... (variável explicativa)

  $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \dots, n$

  $E(\epsilon_i) = 0, \ V(\epsilon_i) = \sigma^2, \quad i = 1, \dots, n$

  $cov(\epsilon_i, \epsilon_j) = 0, \ i \neq j, \ i, j = 1, \dots, n]$

- **Estimativas de MQ dos parâmetros desconhecidos β_0 e β_1**

  Temos

  - $n = 20$
  - $\sum_{i=1}^n x_i = 23.92 \qquad \bar{x} = 1.196 \qquad \sum_{i=1}^n x_i^2 = 29.2892 \qquad \sum_{i=1}^n x_i^2 - n\,\bar{x}^2 = 0.68088$
  - $\sum_{i=1}^n y_i = 1843.21 \qquad \bar{y} = 92.1605 \qquad \sum_{i=1}^n y_i^2 = 170044.5321 \qquad \sum_{i=1}^n y_i^2 - n\,\bar{y}^2 = 173.376895$
  - $\sum_{i=1}^n x_i y_i = 2214.6566 \qquad \sum_{i=1}^n x_i y_i - n\,\bar{x}\,\bar{y} = 10.17744.$

  Logo, as estimativas de MQ dos parâmetros desconhecidos β_1 e β_0 são dadas por:

$$\begin{aligned} \hat{\beta}_1 &= \frac{\sum_{i=1}^n x_i y_i - n\,\bar{x}\,\bar{y}}{\sum_{i=1}^n x_i^2 - n\,\bar{x}^2} = \frac{10.17744}{0.68088} = 14.947480; \\ \hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \times \bar{x} \simeq \bar{y} - \hat{\beta}_1 \bar{x} = 92.1605 - 14.947480 \times 1.196 = 74.283314; \end{aligned}$$

- **Estimativa pedida**

$$\hat{E}(Y \mid x = 1.20) = \hat{\beta}_0 + \hat{\beta}_1 \times 1.20 \simeq 74.283314 + 14.947480 \times 1.20 = 92.220290.$$

- **Cálculo do coeficiente de determinação**

$$r^2 = \frac{\left(\sum_{i=1}^n x_i y_i - n\,\bar{x}\,\bar{y}\right)^2}{\left(\sum_{i=1}^n x_i^2 - n\,\bar{x}^2\right) \times \left(\sum_{i=1}^n y_i^2 - n\,\bar{y}^2\right)} = \frac{(10.17744)^2}{0.68088 \times 173.376895} \simeq 0.877436$$

- **Interpretação coeficiente de determinação**

  Cerca de 88% da variação total da variável resposta Y é explicada pela variável x, através do modelo de regressão linear simples ajustado, donde possamos afirmar que a reta estimada parece ajustar-se bem ao conjunto de dados.