

2ª Freq. + Exame de Época Normal
2014/2015

1. X - n.º anual de instituições científicas q participam no OCJF
 Y - n.º " " alunos q participam no OCJF
 $n_1 = n_2 = 18$

a.i) $\bar{y} = \frac{\sum_{i=1}^{18} y_i}{n_2} = \frac{13214}{18} = 734,11$ - logo em média o n.º de alunos q participaram anualmente no OCJF é de cerca 734 alunos

$\hat{y} = P_{50} = 747,50$ - portanto, em 50% dos anos, o n.º anual de alunos q participaram no OCJF é inferior ou igual a 747,5 alunos e nos restantes anos é superior ou igual a esse valor.

ii) $s_x = ?$
 Como $s_x^2 = \frac{1}{n_1 - 1} \left(\sum_{i=1}^{18} x_i^2 - n_1 \bar{x}^2 \right) = \frac{1}{17} (75409 - 18 \bar{x}^2)$

e $\bar{x} = \frac{\sum_{i=1}^{18} x_i}{n_1} = \frac{1091}{18} = 60,6111$, então tem-se que

$s_x^2 = \frac{1}{17} (75409 - 18 \times 60,6111^2) = 546,0178 \Rightarrow s_x = 23,367$

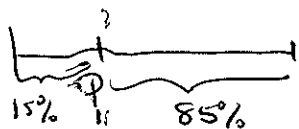
logo, o desvio típico em torno da média do n.º anual de instituições científicas q participam no OCJF é cerca de 23,367 instituições.

$\max_x = ?$

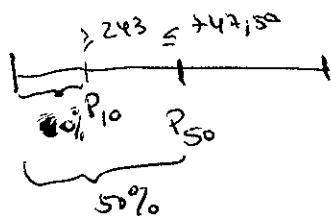
Como $a = \max - \min \Rightarrow 83 = \max - 11 \Rightarrow \max = 94$

portanto, o n.º máximo de instituições científicas q participaram anualmente no OCJF foi 94 instituições

b. c) Em 85% dos anos, o n.º anual de alunos... foi \geq a $318,90 \approx$ 319 alunos.



ii) Em cerca de 40 % dos anos...



c) X - n.º anual Inst. Científicas

Assimetria: Como $G_B = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1} = \frac{83,25 + 44 - 2 \times 62}{39,25} =$

$= \frac{3,25}{39,25} = 0,083 > 0$ logo a distribuição do n.º anual de Inst. Científicas é ligeira/assimétrica positiva.

Achatamento:

Como $k_p = \frac{\text{kurtosis}}{\text{std. Error kurtosis}} = \frac{-0,454}{1,038} = -0,4374$, então

$-1,96 < k_p = -0,4374 < 1,96$. Portanto, a distribuição do n.º anual de Inst. Científicas é mesocúrtica.

d) Como $CV_Y = 42,68\%$ e $CV_X = \frac{23,367}{60,611} \times 100\% = 38,55\%$

então, temos que $CV_X < CV_Y$, portanto podemos concluir que os dados relativos ao n.º anual de instituições q. partic. para no OCJF é mais homogêneo q. o conj. de dados relativos ao n.º anual de alunos. logo, a afirmação é verdadeira.

2.

X - v.a.

x	$1-k$	$k-1$	k	$2k$
$f(x)$	p	$3p$	p	p

a) $E[X] = 8$ $p = ?$ $k = ?$

Como $\sum_{i=1}^4 f(x_i) = 1$ então tem-se $p + 3p + p + p = 1 \Rightarrow 6p = 1$

$E(X) = 8 \Rightarrow \sum_{i=1}^4 x_i f(x_i) = 8 \Rightarrow (1-k)p + (k-1)3p + kp + 2kp = 8$

$\Rightarrow (1-k + 3k - 3 + k + 2k)p = 8 \Rightarrow (-2 + 5k)p = 8$

Logo vem $\begin{cases} 6p = 1 \\ (-2 + 5k)p = 8 \end{cases} \Rightarrow \begin{cases} p = 1/6 \\ -2 + 5k = 48 \end{cases} \Rightarrow \begin{cases} p = 1/6 \\ k = \frac{50}{5} = 10 \end{cases}$

Portanto, tem-se

x	-9	9	10	20
$f(x)$	$1/6$	$1/2$	$1/6$	$1/6$

b) $\text{var}(X) = ?$

Por def. tem-se $\text{var}(X) = E(X^2) - E^2(X) = ?$

Como $E(X^2) = \sum_{i=1}^4 x_i^2 f(x_i) = (-9)^2 \times \frac{1}{6} + 9^2 \times \frac{1}{2} + 10^2 \times \frac{1}{6} + 20^2 \times \frac{1}{6} =$

$= \frac{81}{2} + \frac{1}{6}(81 + 100 + 400) = \frac{81}{2} + \frac{581}{6} = \frac{824}{6} = \frac{412}{3}$, então

tem-se $\text{var}(X) = \frac{412}{3} - 8^2 = \frac{412}{3} - 64 = \frac{412 - 192}{3} = \frac{220}{3} \approx 73,3$

c) $P(X \leq x) \geq 0,75$

Como $P(X \leq -9) = P(X = -9) = 1/6$

$P(X \leq 9) = f(-9) + f(9) = \frac{1}{6} + \frac{1}{2} = \frac{4}{6} = \frac{2}{3} \approx 0,7$

$P(X \leq 10) = P(X \leq 9) + P(X = 10) = \frac{4}{6} + \frac{1}{6} = \frac{5}{6} \approx 0,8$

Portanto, tem-se que $x = 10$.

d) $y = x^3$

Como $x = -9 \Rightarrow y = (-9)^3 = -9^3 \Rightarrow f_{x,y}(-9^3) = 1/6$

Como $x = 9 \Rightarrow y = 9^3$ logo $f_{x,y}(9^3) = 1/2$

$x = 10 \Rightarrow y = 10^3$ logo $f_{x,y}(10^3) = 1/6$

$x = 20 \Rightarrow y = 20^3$ logo $f_{x,y}(20^3) = 1/6$

Dado $f_{x,y}(x,y) = P(X=x, Y=y)$ e como $y = x^3$ então x e y não são independentes, portanto tem-se

$f_{x,y}(-9, 9^3) = P(X=-9, Y=9^3) = 0$

$f_{x,y}(-9, -9^3) = P(X=-9, Y=-9^3) = 1/6 = f_{x,y}(-9, 10^3) = f_{x,y}(-9, 20^3)$

$f_{x,y}(9, -9^3) = P(X=9, Y=-9^3) = 0 = f_{x,y}(9, 10^3) = f_{x,y}(9, 20^3)$

$f_{x,y}(9, 9^3) = P(X=9, Y=9^3) = 1/2$

$f_{x,y}(10, -9^3) = P(X=10, Y=-9^3) = 0 = f_{x,y}(10, 9^3) = f_{x,y}(10, 20^3)$

$f_{x,y}(10, 10^3) = P(X=10, Y=10^3) = 1/6$

$f_{x,y}(20, -9^3) = 0 = f_{x,y}(20, 9^3) = f_{x,y}(20, 10^3)$ e $f_{x,y}(20, 20^3) = 1/6$

Portanto, a função de probabilidade conjunta é definida

por

$x \backslash y$	-9^3	9^3	10^3	20^3
-9	$1/6$	0	0	0
9	0	$1/2$	0	0
10	0	0	$1/6$	0
20	0	0	0	$1/6$

3. X - tempo de vida de um computador

$$X \sim N(\mu, \sigma^2 = 0,7225) \quad \sigma^2 = 0,7225 \Rightarrow \sigma = 0,85$$

a) $P(X > \frac{\mu}{2}) = 0,983 \Leftrightarrow 1 - P(X \leq \frac{\mu}{2}) = 0,983 \Leftrightarrow$

$$\Leftrightarrow P(X \leq \frac{\mu}{2}) = 1 - 0,983 \Leftrightarrow P(Z \leq \frac{\frac{\mu}{2} - \mu}{0,85}) = 0,017 \Leftrightarrow \Phi\left(\frac{\frac{\mu}{2}}{0,85}\right) = 0,983 =$$

$$= \Phi(2,12) \Leftrightarrow \frac{\frac{\mu}{2}}{0,85} = 2,12 \Leftrightarrow \frac{\mu}{2} = 1,802 \Leftrightarrow \mu = 3,604 \approx 3,6$$

logo, o tempo médio de vida dos computadores é 3,6 anos.

b. i) $P(1,5 \leq X \leq 4) = P\left(\frac{1,5 - 3,6}{0,85} \leq Z \leq \frac{4 - 3,6}{0,85}\right) =$

$$= P(-2,47 \leq Z \leq 0,47) = \Phi(0,47) - \Phi(-2,47) =$$

$$= \Phi(0,47) - 1 + \Phi(2,47) = 0,6808 - 1 + 0,9932 = 0,674$$

Não considere \rightarrow não!!

$$\left[\begin{aligned} P(X \leq 3,5) &= P\left(Z \leq \frac{3,5 - 3,6}{0,85}\right) = P(Z \leq -0,18) = \Phi(-0,18) = \\ &= 1 - \Phi(0,18) = 1 - 0,5714 = 0,4286 \end{aligned} \right]$$

ii) $P(X < 1) = P\left(Z < \frac{1 - 3,6}{0,85}\right) = P(Z < -3,06) = \Phi(-3,06) =$

$$= 1 - \Phi(3,06) = 1 - 0,9989 = 0,0011$$

c) $n = 25$

Y - v.a. q representa o n.º de computadores substituídos

$Y \sim B(n=25, p=0,0011)$, pois dado 1 computador só

há duas possibilidades, ou é substituído ou não é substituído. logo, $f(y) = P(Y=y) = {}^{25}C_y (0,0011)^y (1-0,0011)^{25-y}$, $y = 0, \dots, 25$

$$P(Y \leq 3) = P(Y \leq 2) = P(Y=0) + P(Y=1) + P(Y=2) =$$

$$= {}^{25}C_0 (0,0011)^0 (1-0,0011)^{25} + {}^{25}C_1 (0,0011)^1 (1-0,0011)^{24} + {}^{25}C_2 (0,0011)^2 \times$$

$$(1 - 0,0011)^{23} = 1 - 0,9989^{25} + 25 \times 0,0011 \times 0,9989^{24} + 300 \times 0,0011^2 \times 0,9989^{23} \approx 0,97286 + 0,02678 + 0,00035 \approx 0,99999$$

4.

$n_1 = 12$ X - % de indivíduos, com idades entre 16 e 74 anos, com acesso à internet em casa, em Itália
 $n_2 = 12$ Y - % " " , com idades entre 16 e 74 anos, " " à internet em casa, em Portugal

a) Teste da normalidade dos dados *

Teste: $H_0: \mu = 50\%$ vs $H_1: \mu \neq 50\%$

Estatística de Teste: $T = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} \sim t_{12-1} = t_{11}$

C.A.:
 $\alpha = 10\% = 0,10$
 \Downarrow
 $1 - \frac{\alpha}{2} = 1 - 0,05 = 0,95$

* $H_0: Y \sim N(\mu_2, \sigma_2^2)$ vs $H_1: Y \not\sim N(\mu_2, \sigma_2^2)$

Regra de decisão: Rejeitar H_0 se p-value $\leq \alpha$.

Como $n = 12$, então aplica-se o Teste de Shapiro-Wilk, portanto

p-value $_{sw} = 0,312 > \begin{cases} \alpha = 0,01 \\ \alpha = 0,05 \\ \alpha = 0,10 \end{cases}$. Donde, aos níveis de significância usuais, não se rejeita H_0 ,

ou seja, assume-se que Y segue uma distribuição Normal.

Regra de decisão: Rejeitar H_0 se $|T_{obs}| > t_{11; 0,95}$.

Como $t_{11; 0,95} = 1,796$ então R.A. = $[-1,796; 1,796]$ e

R.R.: $]-\infty; -1,796[\cup]1,796; +\infty[$.

Devido que $T_{obs} = \frac{32,83 - 50}{4,651} = -3,692 \in R.R.$

logo rejeita-se a hipótese H_0 , ao nível de significância de 10%.

24) Como I.C._{95%}(μ_2) = $]22,60; 43,07[\supset I.C._{90%}(μ_2) \not\supset 50$

então, rejeita-se a hipótese H_0 , ao nível de significância de 5%.

Portanto, existe evidência estatística para afirmar que a percentagem média de indivíduos com internet em casa em Portugal é diferente de 50%.

20 min

b) $\alpha = 5\% = 0,05 \Rightarrow 1 - \alpha = 0,95$

Teste de normalidade para X :

$$H_0: X \sim N(\mu_1; \sigma_1) \quad \text{vs} \quad H_a: X \not\sim N(\mu_1; \sigma_1)$$

Regra de decisão: Rejeitar H_0 se $p\text{-value} \leq \alpha$.

Como $n_1 = 12 < 30$, então aplica-se o teste de Shapiro-Wilk.

Portanto, como $p\text{-value}_{S-W} = 0,206 > \begin{cases} \alpha = 0,01 \\ \alpha = 0,05 \\ \alpha = 0,10 \end{cases}$, então aos níveis usuais de significância, não se rejeita H_0 , ou seja, assume-se que $X \sim N(\mu_1; \sigma_1)$.

Teste: $H_0: \sigma_1^2 \leq 144 \quad \text{vs} \quad H_1: \sigma_1^2 > 144$

Est. Teste: $\chi^2 = \frac{(n-1) \cdot S^2}{\sigma_0^2} = \frac{11 \times S^2}{144} \sim \chi^2_{12-1} = \chi^2_{11}$

Regra de decisão: Rejeitar H_0 se $\chi^2_{obs} > \chi^2_{11; 0,95}$

Como $\chi^2_{11; 0,95} = 19,68$, então R.A. = $[0; 19,68]$ e

R.R. = $]19,68; +\infty[$. Dado que $\chi^2_{obs} = \frac{11 \times 12,281^2}{144} = 11,52$

é R.A., então não se rejeita a hipótese H_0 , ou seja, ao nível de significância de 5%, não existe evidência estatística para afirmar que a variância da percentagem de indivíduos de Itália com acesso à internet excede 144.

c) $p\text{-value} = P(\chi^2 > \chi^2_{obs}) = P(\chi^2 > 11,52) = 1 - P(\chi^2 \leq 11,52) = 1 - 0,6 = 0,4$

d) $\alpha = 10\% \Rightarrow 1 - \alpha = 0,90$

Teste: $H_0: \mu_1 - \mu_2 = 0 \quad \text{vs} \quad H_1: \mu_1 - \mu_2 \neq 0$

como X e Y seguem distribuições normais tem-se que se

verificam se podemos assumir a igualdade das variâncias. ⑧

Teste de Levene:

$$H_0: \sigma_1^2 = \sigma_2^2 \quad \text{vs} \quad H_1: \sigma_1^2 \neq \sigma_2^2$$

Regra de decisão: Rejeitar H_0 se $p\text{-value}_{T. \text{ Levene}} \leq \alpha$.

Como $p\text{-value}_{T. L.} = 0,220 > \begin{cases} \alpha = 0,01 \\ \alpha = 0,05 \\ \alpha = 0,10 \end{cases}$, então não se rejeita H_0 , ou seja, aos níveis de significância usuais, assume-se a igualdade das variâncias. Assim, tem-se que

Est. Teste:

$$T = \frac{\bar{X} - \bar{Y} - \mu_0}{\sqrt{\frac{11 \cdot S_1^2 + 11 \cdot S_2^2}{22} \left(\frac{1}{12} + \frac{1}{12} \right)}} \sim t_{22}$$

Regra de decisão: Rejeitar H_0 se $p\text{-value} \leq \alpha$.

Como $p\text{-value} = 0,778 > 0,10 = \alpha$, então não se rejeita H_0 , ou seja, ao nível de significância de 10%, não existe evidência estatística para afirmar que existem diferenças significativas entre Itália e Portugal quanto à percentagem média de indivíduos com acesso à internet.

10u) Regra de decisão: Rejeitar H_0 se $|T_{\text{obs}}| > t_{22; 0,90}$

$$\text{Como } I.C._{95\%}(\mu_1 - \mu_2) =]-1,106 - \epsilon; 13,795[= (*)$$

$$13,795 = \bar{X} - \bar{Y} + \epsilon \quad \Rightarrow \quad \epsilon = 13,795 - (\bar{X} - \bar{Y}) = 13,795 - (3,545 - 4,651) \\ = 13,795 - (-1,106) = 14,901$$

$(*) =]-16,007; 13,795[\supset I.C._{90\%}(\mu_1 - \mu_2) \ni 0$, logo \bar{n} se rejeita H_0 .

5.

9

X-estágios por área científica

 $k = 5$ - classes

a) $A = O_1 = ? = 33 //$

$O_1 - E_1 = -4,4 \Rightarrow O_1 = E_1 + 4,4 = 37,4 - 4,4 = 33 //$

$B = E_2 = ? = O_2 - (O_2 - E_2) = 17 - (-4,8) = 21,8 //$

$C = E_3 = ? = n - (E_1 + E_2 + E_4 + E_5) = 104 - (37,4 + 21,8 + 22,9 + 16,6) = 5,3 //$

$\Rightarrow D = O_3 - E_3 = 8 - 5,3 = 2,7 //$

$E = O_4 = n - O_1 - O_2 - O_3 - O_5 = 104 - 33 - 17 - 8 - 16 = 30 //$

$F = O_4 - E_4 = 30 - 22,9 = 7,1 //$

$G = df = k - 1 = 5 - 1 = 4 //$

 $\alpha = 5\%$

b) Teste de ajustamento:

$H_0: p_1 = 0,36 \wedge p_2 = 0,21 \wedge p_3 = 0,05 \wedge p_4 = 0,22 \wedge p_5 = 0,16$

vs

$H_1: p_1 \neq 0,36 \text{ ou } p_2 \neq 0,21 \text{ ou } p_3 \neq 0,05 \text{ ou } p_4 \neq 0,22 \text{ ou } p_5 \neq 0,16$

Estatística teste:

$$\chi^2 = \sum_{i=1}^5 \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_4$$

Regra de decisão: Rejeitar H_0 se $\chi^2_{obs} > \chi^2_{4; 0,95}$ Como $\chi^2_{obs} = 5,347$ e $\chi^2_{4; 0,95} = 11,14$, então

$\chi^2_{obs} < \chi^2_{4; 0,95}$, logo não se rejeita H_0 , ou seja, ao nível de significância de 5%, não existe evidência estatística para afirmar que a distribuição do número de estágios por área científica sofreu alteração significativa.

ou

Regra de decisão: Rejeitar H_0 se $p\text{-value} \leq \alpha$.

Como $p\text{-value} = 0,253 > 0,05 = \alpha$, então não se rejeita H_0 , ao nível de significância de 5%.

6.

 X - faixa etária X_1 - até aos 25 anos X_2 - mais de 25 anos $L=2$ Y - classificação dos pacientes Y_1 - Recuperados Y_2 - \bar{N} na ep. mas em trata / Y_3 - \bar{N} " e desistiu do trata / $C=3$

a) o teste de hipóteses q permite concluir se o estado dos pacientes ao fim de 3 anos de trata / é independente da idade e o teste de independência

 H_0 : X e Y são independentes

vs

 H_1 : X e Y não são independentes

b) Est. Teste:

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^3 \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2_{1 \times 2} = \chi^2_2$$

Regra de decisão: Rejeitar H_0 se $(\chi^2_{obs} > \chi^2_2)$
p-value $\leq \alpha$.

Como

		Y			
		Y_1	Y_2	Y_3	Total
X	X_2	$O_{11}=56$ $E_{11}=58,3$	$O_{12}=72$ $E_{12}=60,2$	$O_{13}=52$ $E_{13}=61,5$	$180=O_{1.}$
	X_1	$O_{21}=34$ $E_{21}=31,7$	$O_{22}=21$ $E_{22}=32,8$	$O_{23}=43$ $E_{23}=33,5$	$98=O_{2.}$
Total		$O_{.1}=90$	$O_{.2}=93$	$O_{.3}=95$	$n=278$

$$E_{12} = \frac{O_{1.} \times O_{.2}}{n} = \frac{180 \times 93}{278} = 60,2$$

$$E_{13} = \frac{O_{1.} \times O_{.3}}{n} = \frac{180 \times 95}{278} = 61,5$$

$$E_{22} = \frac{98 \times 93}{278} = 32,8$$

$$E_{23} = \frac{98 \times 95}{278} = 33,5$$

então tem-se que

$$\begin{aligned} \chi^2_{obs} &= \frac{(56-58,3)^2}{58,3} + \frac{(72-60,2)^2}{60,2} + \frac{(52-61,5)^2}{61,5} + \frac{(34-31,7)^2}{31,7} \\ &+ \frac{(21-32,8)^2}{32,8} + \frac{(43-33,5)^2}{33,5} = 0,0907 + 2,3130 + 1,4675 + 0,1669 \\ &+ 4,2451 + 2,6940 = 10,977 \end{aligned}$$

Portanto, como $P(\chi^2 > \chi^2_{obs}) = p\text{-value} = P(\chi^2 > 10,977) =$
 $= 1 - P(\chi^2 \leq 10,977) = 1 - 0,995 = 0,005 < \alpha = 0,05$

então, rejeita-se a hipótese H_0 , ou seja, ao nível de significância de 5%, existe evidência estatística para afirmar que o estado dos pacientes ao fim de 3 anos de tratamento, não é independente da idade.