Esercizio 5. La seguente tabella riporta il numero di componenti di 100 nuclei familiari estratti casualmente in una certa città.

Componenti	1	2	3	4	5	6	7
Frequenza assoluta	11	19	28	26	11	4	1

(a) Stimare la dimensione media di tutti i nuclei familiari della città.

(b) Costruire un intervallo di confidenza al 95% per la dimensione media.

(c) Verificare, al livello $\alpha = 0.05$ e calcolando il p-value, l'ipotesi $H_0 : \mu = 3$ contro $H_1 : \mu \neq 3$.

a) $\bar{X}_n$ stimatore di μ : dimensione media in popolazione

$\mathbb{E}[\bar{X}_n] = \mu$

stima $\bar{x}_n = \frac{1}{100} \sum_{k=1}^{7} c_k \cdot F_k = \frac{1}{100} (1 \cdot 11 + 2 \cdot 19 + \dots + 7 \cdot 1) = 3.23$

b) dal TLC $\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} \dot{\sim} N(0,1)$ ma σ^2 non nota $\rightarrow$ la stimiamo

$\sum_{i=1}^{100} x_i^2 = \sum_{k=1}^{7} c_k^2 F_k = 1 \cdot 11 + 2^2 \cdot 19 + 3^2 \cdot 28 + \dots + 7^2 \cdot 1 = 1223$

$\Rightarrow \sum_{i=1}^{100} (x_i - \bar{x})^2 = \sum_{i=1}^{100} x_i^2 - \left(\sum_{i=1}^{100} x_i\right)^2 = 1223 - 100 (3.23)^2 = 179.71$

$\Rightarrow$ la stima corretta di σ^2 è $s_c^2 = \frac{179.71}{99} = 1.81$

lo stimatore corrispondente è $S_c^2 = \frac{1}{99} \sum_{i=1}^{100} (X_i - \bar{X}_n)^2$

QUANTITA' PIVOTALE : $Z_n = \frac{\bar{X}_n - \mu}{\sqrt{S_c^2/n}} \dot{\sim} N(0,1)$

$\mathbb{P}\left(z_{\frac{\alpha}{2}} \le Z_n \le z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$



$1 - \alpha = 0.95 \rightarrow \alpha = 0.05$

$1 - \frac{\alpha}{2} = 0.975$

$z_{0.975} = 1.96 \Rightarrow z_{\frac{\alpha}{2}} = -z_{1-\frac{\alpha}{2}}$

$\mathbb{P}\left(-1.96 \le \frac{\bar{X} - \mu}{\sqrt{S_c^2/n}} \le 1.96\right) = 0.95$

$\mathbb{P}\left(\bar{X} - 1.96\sqrt{\frac{S_c^2}{n}} \le \mu \le \bar{X} + 1.96\sqrt{\frac{S_c^2}{n}}\right) = 0.95$

sui dati : $\bar{x} = 3.23 \quad s_c^2 = 1.81$

$IC_{0.95}(\mu) = 3.23 \pm 1.96\sqrt{\frac{1.81}{100}} = [2.97, 3.49]$

se $1-\alpha = 0.99 \Rightarrow \alpha = 0.01 \qquad \frac{\alpha}{2} = 0.005$

$1-\frac{\alpha}{2} = 0.995 \Rightarrow z_{0.995} = 2.57$

$$IC_{0.99}(\mu) = 3.23 \pm 2.57\sqrt{\frac{1.81}{100}}$$

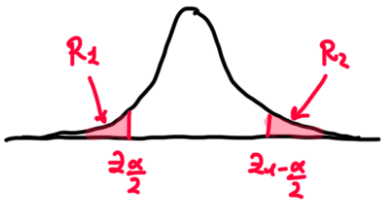
(c) $\begin{cases} H_0: \mu = 3 \\ H_1: \mu \neq 3 \end{cases}$

sotto H_0 $\quad \dfrac{\bar{X}_n - 3}{\sqrt{\sigma^2/n}} \dot{\sim} N(0,1)$

σ^2 non è noto: lo stimiamo $\Rightarrow$ $T = \dfrac{\bar{X}_n - 3}{\sqrt{S_c^2/n}} \dot{\sim} N(0,1)$ sotto H_0

1) fisso $\alpha = 0.05 \Rightarrow \alpha = P_{H_0}(T \in R)$

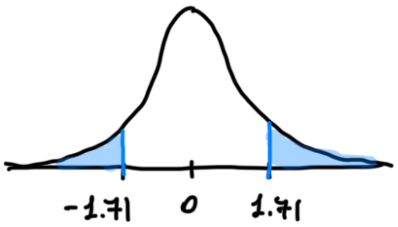
R = regione di rifiuto



$R = R_1 \cup R_2$
$= (-\infty ; a) \cup (b ; +\infty)$
$= (-\infty ; -z_{0.975}) \cup (z_{0.975} ; +\infty)$
$= (-\infty ; -1.96) \cup (1.96 ; +\infty)$

valore osservato della statistica test $\quad t^{obs} = \dfrac{3.23-3}{\sqrt{1.81/100}} = 1.71$

$t^{obs} \notin R \Rightarrow$ non rifiuto H_0



2) pvalue $\quad \alpha^{obs} = P_{H_0}(|T| > |t^{obs}|) =$

$= 2\, P_{H_0}(T > |t^{obs}|) =$

$= 2\,(1 - \Phi(1.71)) =$

$= 2\,(1 - 0.956) = 0.08$

se $\alpha = 0.05$ → non rifiuto H_0

se $\alpha = 0.1$ → rifiuto H_0