Esercizio 13. Un comune conduce un'indagine sulla quota di cittadini soddisfatti del servizio di trasporto pubblico locale. Un cittadino è soddisfatto se esprime un giudizio pari ad almeno 4 su una scala da 1 a 5.

Livello di soddisfazione	1	2	3	4	5	Totale
Frequenza assoluta	30	45	60	180	185	500

(a) Stimare la proporzione di cittadini soddisfatti nella popolazione.

(b) Costruire un intervallo di confidenza al 95% per tale proporzione.

(c) Se si volesse un intervallo di confidenza con minore ampiezza, cosa si potrebbe provare a fare?

(d) Verificare $H_0 : p = 0.70$ contro $H_1 : p \neq 0.70$.

a) totale cittadini soddisfatti = 365

proporzione di cittadini soddisfatti = $\frac{365}{500} = 0.73$

$X_i = \begin{cases} 1 & \text{se cittadino}_i \text{ è soddisfatto} \\ 0 & \text{altrimenti} \end{cases}$

$X_i \sim \text{Bernoulli}(p)$

p = probabilità di essere soddisfatto

stima sul campione $\hat{p} = 0.73 = \bar{x}$

b) $\bar{X}_n \dot{\sim} N\left(p ; \frac{p(1-p)}{n}\right)$ del TLC

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

$\text{var}(\bar{X}_n) = \text{var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\text{var}\left(\sum_{i=1}^{n} X_i\right) \overset{ind}{=} \frac{1}{n^2}\sum_{i=1}^{n}\text{var}(X_i)$

$= \frac{1}{n^2}\sum_{i=1}^{n} p(1-p) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$

$\frac{\bar{X}_n - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \dot{\sim} N(0,1)$ quantità pivotale

$\mathbb{P}\left(z_{\frac{\alpha}{2}} \leq \frac{\bar{X}_n - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \leq z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$

simmetria: $z_{\frac{\alpha}{2}} = -z_{1-\frac{\alpha}{2}}$

se $1-\alpha = 0.95 \Rightarrow 1-\frac{\alpha}{2} = 0.975 \Rightarrow z_{0.975} = 1.96$

$\mathbb{P}\left(-1.96 \leq \frac{\bar{X}_n - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \leq 1.96\right) = 0.95$

$\mathbb{P}\left(-1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq \bar{X}_n - p \leq 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = 0.95$

$\mathbb{P}\left(\bar{X}_n - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \bar{X}_n + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = 0.95$

poiché $\hat{p} = \bar{x} = 0.73$ otteniamo

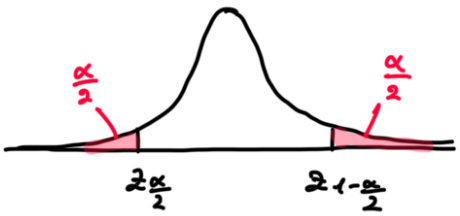
$$IC_{0.95}(p) = 0.73 \pm 1.96 \sqrt{\frac{0.73(1-0.73)}{500}}$$

c) aumentare la dimensione del campione n, oppure ridurre il livello di confidenza $1-\alpha$ (es. 0.90 invece di 0.95)

d) $\begin{cases} H_0: p = 0.70 \\ H_1: p \neq 0.70 \end{cases}$

statistica test $T = \dfrac{\bar{X}_n - 0.70}{\sqrt{\frac{0.70(1-0.70)}{n}}} \overset{\cdot}{\sim} N(0,1)$ sotto H_0

i) se fisso la significatività $\alpha = 0.05$



regione di rifiuto:

$$R = (-\infty ; z_{\frac{\alpha}{2}}) \cup (z_{1-\frac{\alpha}{2}} ; +\infty)$$

$$= (-\infty ; -1.96) \cup (1.96 ; +\infty)$$

valore osservato $t^{obs} = \dfrac{0.73 - 0.70}{0.0205} = 1.46$

$t^{obs} \notin R \Rightarrow$ non rifiuto H_0 al livello $\alpha = 0.05$

ii) pvalue: $\alpha^{obs} = \mathbb{P}_{H_0}(|T| > |t^{obs}|) =$

$$= 2\left(1 - \Phi(1.46)\right) = 0.14$$

la proporzione osservata non è sorprendente rispetto all'ipotesi che $p = 0.7 \Rightarrow$ non rifiuto H_0