Corso di Statistica per Ingegneria Fisica
Anno accademico 2020/2021

Esercitazione 9 - Test d'ipotesi per due popolazioni indipendenti

Esercizio 1. Si vuole confrontare, mediante un esperimento clinico, l'efficacia di due farmaci antiallergici, che indichiamo con A e B. Su $n = 50$ pazienti che acconsentono a sottoporsi a tale esperimento, 27 selezionati a caso vengono curati con il farmaco A, mentre ai restanti 23 viene assegnato il farmaco B. Le risposte a tali farmaci da parte dei pazienti, valutate sulla base di un opportuno indice clinico adimensionale, sono indipendenti e distribuite secondo una $N(\mu_A, 0.9)$ per il farmaco A e secondo una $N(\mu_B, 0.9)$ per il farmaco B.

(a) Per mezzo di un opportuno test al livello di significatività del 5%, si decida se è evidente dai dati che $\mu_A > \mu_B$, sapendo che l'esperimento clinico ha fornito media campionaria delle risposte pari a 11.51 per il farmaco A e pari a 10.69 per il farmaco B. *[Test per le ipotesi $H_0 : \mu_A \leq \mu_B$ (o $H_0 : \mu_A = \mu_B$) vs. $H_1 : \mu_A > \mu_B$. Si rifiuta H_0 al 5% di significatività.]*

(b) Si determini il p-value del test del punto (a) e se ne tragga una conclusione. È una conclusione debole o forte? *[p-value = 0.16% $\Rightarrow$ Fortissima evidenza a favore dell'ipotesi alternativa $H_1 : \mu_A > \mu_B$. Conclusione forte.]*

(c) Calcolare la potenza del test costruito al punto (a) quando $\mu_A = \mu_B + 0.39$.
[42.2695%]

(d) Supponendo ora di suddividere un gruppo di n volontari in due campioni ugualmente numerosi, uno per ciascun farmaco, quanto deve valere n affinché la potenza del test del punto (a) sia maggiore o uguale all'80% quando $\mu_A = \mu_B + 0.39$? *[$n \geq 147$.]*

Esercizio 2. Si considerino due campioni aleatori con la medesima numerosità n: il primo è $X_1, \ldots, X_n$, in cui $X_i \sim N(\mu, 3)$ per ogni i, mentre il secondo è $Y_1, \ldots, Y_n$, con $Y_j \sim N(\nu, 3)$ per ogni j. Si intende sottoporre a test l'ipotesi nulla $H_0 : \nu = \mu$ contro l'ipotesi alternativa $H_1 : \nu < \mu$.

(a) Si indichi la regione critica per un test di livello 5%.
[$\mathrm{RC}(0.05) = \left\{(x_1, \ldots, x_n, y_1, \ldots, y_n) : \frac{\bar{x}_n - \bar{y}_n}{\sqrt{6}}\sqrt{n} > 1.645\right\}$]

(b) In funzione di n, si trovi la probabilità di accettare erroneamente H_0 nel caso $\mu = \nu + 1$.
[$\Phi\left(1.645 - \sqrt{\frac{n}{6}}\right)$]

(c) Si ricavi la minima numerosità n_0 per cui la probabilità calcolata al punto (b) non supera il 50%.
[$n_0 = 17$]

Esercizio 3. Le ditte H305 ed Antares producono sfere luminose. Una volta attivata la reazione chimica che rende luminosa una di queste sfere, la sua durata, misurata in ore, è descritta da una variabile aleatoria normale la cui varianza è nota e dipende dal processo produttivo impiegato da ciascuna ditta: nel caso delle sfere prodotte dalla ditta H305 la varianza è $\sigma_H^2 = 9\,\mathrm{h}^2$, nel caso delle sfere prodotte dalla ditta Antares si ha $\sigma_A^2 = 25\,\mathrm{h}^2$. La ditta H305 sta progettando una campagna pubblicitaria comparativa nella quale intende affermare che la durata media μ_A delle proprie sfere luminose è superiore a quella μ_B della ditta rivale Antares; tuttavia, ritenendo molto sconveniente fare erroneamente un'affermazione di questo tipo, prima di intraprendere la campagna vuole essere ragionevolmente sicura di quello che dice. Vengono pertanto esaminati un campione casuale di $n_H = 35$ sfere H305, trovando la media campionaria $\bar{x}_H = 51.15\,\mathrm{h}$, e un campione casuale di $n_A = 25$ sfere Antares, ottenendo la media campionaria $\bar{x}_A = 49.68\,\mathrm{h}$.

(a) Si fissino le ipotesi nulla e alternativa di un test volto a dimostrare che la durata media delle sfere H305 è maggiore di quella delle sfere Antares. Si tragga la conclusione del test al livello di significatività del 5%. *[Test per le ipotesi $H_0 : \mu_H = \mu_A$ (o $H_0 : \mu_H \leq \mu_A$) vs. $H_1 : \mu_H > \mu_A$. Non si può rifiutare H_0.]*

(b) Usando uno stimatore non distorto, si fornisca una stima della differenza fra le durate medie dei due tipi di sfera. *[$\widehat{\mu_H - \mu_A} = 1.47\,\mathrm{h} = 1\,\mathrm{h}$ e $28.2\,\mathrm{min}$.]*

(c) Si calcoli la probabilità che il test al livello 5% del punto (a) induca la H305 a rinunciare erroneamente alla sua campagna pubblicitaria comparativa, quando in realtà $\mu_H = \mu_A + 1$. *[77.337%]*

(d) Si calcoli quanto dovrebbe valere n_A per ridurre la probabilità calcolata in (c) sotto il 50%. *[$n_A \geq 223$.]*

(e) Si calcoli quanto dovrebbe valere n_A per ridurre la probabilità calcolata in (c) sotto il 10%. *[Impossibile! Va aumentato anche n_H.]*

Esercizio 4. Una grossa industria vuole confrontare le caratteristiche tra due resine plastiche per costruire tubature. La resina A è meno costosa della resina B, ma si sospetta che la resina B abbia un maggior allungamento medio percentuale (capacità del materiale di allungarsi senza superare la rottura). Un campione sperimentale di numerosità 25 della resina A ha fornito un allungamento medio percentuale pari a 2.1 e una deviazione standard campionaria pari a 0.51, mentre un campione di numerosità 16 della resina B ha fornito un allungamento medio percentuale pari a 2.5 e una deviazione standard campionaria pari a 0.58. Si supponga che i campioni delle due resine siano normali e indipendenti.

(a) Stabilire con un opportuno test se si può accettare al livello di significatività del 10% l'ipotesi nulla che i due campioni abbiano varianze uguali contro l'alternativa che le varianze siano diverse. *[Non possiamo rifiutare H_0]*

(b) Alla luce della risposta precedente, si fornisca una stima della varianza di ciascuna delle due resine. *[$\widehat{\sigma^2_A} = \widehat{\sigma^2_B} = 0.289446$.]*

(c) Al livello di significatività dell'2.5%, i dati a disposizione evidenziano che l'allungamento medio della resina B è maggiore di quello della resina A? Si costruisca un opportuno test statistico. *[Sì.]*

(d) Si determini il p-value del test del punto precedente. *[$1\% < p\text{-value} < 2.5\%$.]*

Ora si supponga invece che $\sigma_A = 0.51$ e $\sigma_B = 0.58$ siano le deviazioni standard delle due serie di misurazioni, *già note* dall'esperienza passata per questo genere di campioni.

(e) Con questa nuova informazione, i dati a disposizione permettono di concludere a livello del 2.5% che l'allungamento medio della resina B è maggiore di quello della resina A? *[Sì.]*

(f) Calcolare la potenza del test del punto (e) sapendo che il vero allungamento medio della resina A è pari a 1.9 e quello della resina B è 2.4. *[80.511%]*

Esercizio 5. Si vuole valutare il tempo di clock dei microprocessori prodotti da due linee di produzione diverse. Vengono quindi scelti due campioni di numerosità $n = 10$ e $m = 12$ rispettivamente e ad ognuno di essi viene fatto eseguire uno stesso programma, misurando il tempo impiegato. Si ottengono i risultati seguenti per la media e la varianza empirica dei due campioni

$$\bar{x} = 8.5, \qquad s^2_X = 3.6, \qquad\qquad \bar{y} = 5.2, \qquad s^2_Y = 1.1.$$

Si può affermare che in media i due tempi di clock siano diversi? Si supponga che i dati siano normali e che le varianze reali dei due microprocessori siano uguali *[Sì , ad ogni usuale livello di significatività si può respingere l'ipotesi nulla che i microprocessori prodotti dalle due linee abbiano lo stesso tempo medio di clock. Infatti per il test bilatero si ottiene p-value* $< 0.05\%$*.]*

Esercizio 6. Vengono sottoposti a confronto i consumi delle autovetture CS 1.1 e VP 1.0 alla velocità costante di 120 Km/h. La CS in 20 prove consuma mediamente 6.51 lt/100 Km, la VP in 23 prove consuma mediamente 6.84 lt/100 Km. Le relative varianze campionarie sono rispettivamente di $0.64\,(\mathrm{lt}/100\,\mathrm{Km})^2$ e $0.48\,(\mathrm{lt}/100\,\mathrm{Km})^2$. Si supponga che i dati provengano da una distribuzione gaussiana.

(a) Possiamo ritenere che i consumi delle due autovetture abbiano la stessa varianza al livello di significatività del 5%? *[Sì.]*

(b) Possiamo ritenere che le due autovetture abbiano lo stesso consumo medio al livello di significatività del 5%? *[Sì.]*

Esercizio 7. La Molino Blanco, una famosa ditta spagnola di biscotti, ha prodotto una nuova versione dei suoi biscotti più famosi, i Tarallucciος. La particolarità di questi nuovi biscotti dovrebbe essere il diametro più grande. Prima di lanciare il nuovo prodotto sul mercato, però, il responsabile alle vendite vuole essere sicuro di quanto si intende affermare nella campagna pubblicitaria. Pertanto, fa misurare i diametri di 13 biscotti del vecchio tipo (variabili $X_1, \dots, X_{13}$, assunte gaussiane) e 16 del nuovo (variabili $Y_1, \dots, Y_{16}$, assunte gaussiane). I dati ottenuti sono riassunti dai seguenti valori delle media e varianze campionarie:

$$\begin{aligned} \bar{x} &= 1.92\,\mathrm{cm} & \qquad \bar{y} &= 2.23\,\mathrm{cm} \\ s_X^2 &= 0.094\,\mathrm{cm}^2 & \qquad s_Y^2 &= 0.132\,\mathrm{cm}^2\,. \end{aligned}$$

(a) Aiutate il responsabile alle vendite a stabilire se i nuovi biscotti abbiano un diametro medio superiore a quelli vecchi, effettuando un opportuno test al 5%. *[Test per le ipotesi* $H_0 : \mu_X = \mu_Y$ *(o* $\mu_X \geq \mu_Y$*) vs.* $H_1 : \mu_X < \mu_Y$*. Si rifiuta* H_0*.]*

(b) Si calcoli il p-value del test implementato al punto precedente. L'azienda secondo voi può lanciare la campagna pubblicitaria? *[*$1\% <$ *p-value* $< 2.5\% \Rightarrow$ *Sì.]*

Esercizio 8. Un laboratorio informatico ha elaborato un nuovo protocollo P_{new} per la trasmissione di dati. Per confrontare P_{new} con il vecchio protocollo P_{old}, si procede a inviare un certo file per 7 volte da un server a un altro usando il protocollo P_{new}, e altre 6 volte usando il protocollo P_{old}. Si misurano quindi i tempi (in secondi) intercorrenti tra ciascun invio e la corrispondente ricezione. I risultati ottenuti per P_{new} sono:

$$x_i : \quad 1.49 \quad 1.50 \quad 1.96 \quad 2.33 \quad 1.45 \quad 1.71 \quad 2.83$$

e quelli per P_{old} sono

$$y_i : \quad 1.85 \quad 3.47 \quad 4.44 \quad 1.75 \quad 2.16 \quad 3.93$$

Supponendo i dati gaussiani, utilizzarli per stabilire se il nuovo protocollo P_{new} sia migliore del vecchio P_{old}, dove, in modo naturale, un protocollo è migliore di un altro se trasferisce i dati in meno tempo. *[Dai dati abbiamo evidenza – sebbene non fortissima – che* P_{new} *sia migliore di* P_{old}*: p-value* $\in (2.5\%\,, 5\%)$ *per un test delle ipotesi* $H_0 : \mu_X \geq \mu_Y$ *vs.* $H_1 : \mu_X < \mu_Y$*.]*

Esercizio 9. Due macchine A e B producono filo di rame, di cui si è già stabilito che il diametro ha il valore medio $\mu_0 = 2\,\mathrm{mm}$. Per controllare la qualità del processo, vengono ispezionati 10 fili prodotti dalla macchina A e 16 prodotti dalla macchina B, e per ogni filo viene registrato l'errore E_i nella lunghezza del diametro (dove $E_i = i$-esimo valore misurato $- \mu_0$). Dalle misurazioni effettuate si ottiene

$$\sum_{i=1}^{10} e_{i,A}^2 = 0.017\,\text{mm}^2 \qquad \sum_{i=1}^{10} e_{i,A} = 0.1304\,\text{mm}$$
$$\sum_{i=1}^{16} e_{i,B}^2 = 0.095\,\text{mm}^2 \qquad \sum_{i=1}^{16} e_{i,B} = 0.3194\,\text{mm}\,,$$

in cui gli indici A e B si riferiscono alla macchina che ha prodotto l'errore. Si può assumere che gli errori nella lunghezza del diametro siano variabili aleatorie indipendenti gaussiane.

(a) Impostando un opportuno test di verifica di ipotesi di livello $\alpha = 10\%$, i dati mostrano evidenza che la macchina A è più precisa della macchina B? *[Test per le ipotesi* $H_0 : \sigma_A^2 = \sigma_B^2$ *(o* $H_0 : \sigma_A^2 \geq \sigma_B^2$*) vs.* $H_1 : \sigma_A^2 < \sigma_B^2$*. Si rifiuta* H_0*.]*

(b) Calcolare il p-value del test del punto (b). *[*$2.5\% <$ *p-value* $< 5\%$ *.]*

(c) Fornite una stima puntuale della varianza σ_A^2 dell'errore nella lunghezza del diametro dei fili prodotti dalla macchina A e della varianza σ_B^2 dell'errore nella lunghezza del diametro dei fili prodotti da B. *[*$\widehat{\sigma_A^2} = 0.001700$*,* $\widehat{\sigma_B^2} = 0.005908$*]*

Esercizio 10. La ditta A produce funi per arrampicata sportiva. L'azienda chimica B sostiene di aver sviluppato un nuovo materiale che, se utilizzato per produrre funi da arrampicata, permetterebbe di ottenere funi con resistenza media maggiore di quelle attualmente vendute dalla A. La B ha proposto alla A di comprare il brevetto del nuovo materiale, e, per convincerla a fare l'acquisto, le ha inviato un documento coi risultati delle misure di resistenza di un certo numero di funi prodotte attualmente dalla A, e di altre prodotte invece con il nuovo materiale. Nel documento sono riportati i seguenti dati:

- $n_A = 40$: il numero di funi, prodotte da A, che sono state testate;
- $n_B = 60$: il numero di funi, prodotte con il nuovo materiale, che sono state testate;
- $\bar{x}_A = 5953\,\text{N}$ e $s_A^2 = 25140\,\text{N}^2$: la media campionaria e la varianza campionaria delle resistenze misurate per il campione di funi prodotte da A;
- $\bar{x}_B = 6088\,\text{N}$ e $s_B^2 = 32367\,\text{N}^2$: la media campionaria e la varianza campionaria delle resistenze misurate per il campione di funi prodotte con il nuovo materiale.

(a) Impostate un opportuno test d'ipotesi per aiutare la ditta A a decidere se vale la pena di acquistare il brevetto della B. Sulla base dei dati ricevuti, voi lo acquistereste? *[Test per le ipotesi* $H_0 : \mu_A = \mu_B$ *(o* $H_0 : \mu_A \geq \mu_B$*) vs.* $H_1 : \mu_A < \mu_B$*. p-value* $= 0.004\% \Rightarrow$ *Fortissima evidenza che vale la pena di acquistare il brevetto.]*

(b) Nel test del punto precedente, bisogna fare delle ipotesi particolari sulla densità dei due campioni? *[No.]*

(c) Fornite una stimatore non distorto della differenza delle medie dei due campioni, e trovate la stima corrispondente. *[Stimatore:* $\hat{\Theta} = \bar{X}_A - \bar{X}_B$*. Stima:* $\hat{\theta} = 135\,\text{N}$*.]*

Esercizio 11. Un campione di 300 dei votanti della regione A e di 200 della regione B ha mostrato che rispettivamente il 56% e il 48% è favorevole a un certo candidato. Si può affermare che a un livello di significatività dello 0.05 vi sia differenza tra le due regioni? *[No.]*

Esercizio 12. La POSL Corsico, una gloriosa società di calcio del sud-ovest milanese, ha nella Nuova Amatese uno dei suoi più acerrimi nemici. Il presidente della POSL è convinto della superiorità della sua squadra. Si consideri che la POSL ha vinto 85 delle 126 gare disputate nella sua storia, e che la Nuova Amatese ne ha vinte 51 su 99. Si faccia inoltre l'assunzione che la probabilità di vincere non dipenda dall'avversario. È possibile affermare che la probabilità di vittoria in una partita della POSL è maggiore di quella della Nuova Amatese a un livello $\alpha = 0.05$? *[Sì.]*

Esercizio 13. Un ufficio studi di una certa assicurazione ha constatato che nella località A, dove conta 25 automobili assicurate, vi sono stati 5 furti d'auto; nella località B, a fronte di 45 auto assicurate, vi sono stati 8 furti d'auto. L'ufficio studi può concludere che le due località siano ugualmente pericolose? In caso contrario, qual è la più pericolosa? *[Non si può rifiutare l'ipotesi nulla che una macchina abbia la stessa probabilità di essere rubata in ciascuna delle due località, quindi abbiamo una conclusione (debole) a favore del fatto che le due località siano ugualmente pericolose.]*

Esercizio 14. Si teme che nel prossimo referendum non venga raggiunto il quorum (percentuale dei votanti superiore al 50%).

(a) Per prevedere se il quorum verrà raggiunto, si vuole costruire un intervallo di confidenza bilatero di livello 0.99 per la proporzione p dei votanti. A tale scopo, si intervistano n persone scelte a caso nella popolazione degli elettori e si chiede loro se si recheranno alle urne. Quanto grande deve essere n affinchè l'intervallo abbia ampiezza al più pari a 0.02? *[$n \geq 16577$]*

(b) Si supponga che in Lombardia, su 1000 persone intervistate, ve ne siano solo 480 che dichiarano che si recheranno alle urne. I dati a disposizione consentono di affermare che in Lombardia il quorum non verrà raggiunto, con livello di significatività del 5%? *[No.]*

(c) Supponiamo inoltre che in Emilia Romagna su 1000 persone intervistate solo 470 si dichiarino intenzionate a votare. Tenendo conto dei dati al punto (b), possiamo concludere che in Lombardia la percentuale di votanti sarà maggiore che in Emilia Romagna, con livello di significatività ancora del 5%? *[No.]*

Soluzione 1.

(a) Se vogliamo verificare se i dati mostrano evidenza che $\mu_A > \mu_B$, dobbiamo fare un test che abbia questa affermazione come ipotesi alternativa:

$$H_0 : \mu_A \leq \mu_B \;\; (\text{o } H_0 : \mu_A = \mu_B) \qquad \text{contro} \qquad H_1 : \mu_A > \mu_B \,.$$

La scelta di H_0 tra le due possibilità date sopta è del tutto ininfluente; è H_1 difatti che fissa la forma della regione di rifiuto (intervallo unilatero destro, sinistro o bilatero). Siccome il testo ci dice che i due campioni sono gaussiani con varianza nota ($\sigma_A^2 = \sigma_B^2 = 0.9$ per entrambi i campioni), possiamo fare uno Z-test. La regola di rifiuto al livello α di un tale test per la differenza delle medie è

$$\text{“rifiuto } H_0 \text{ se } Z_0 = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} > z_{1-\alpha}\text{”} \,. \qquad (*)$$

Dai dati ricaviamo

$$z_0 = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} = \frac{11.51 - 10.69}{\sqrt{\frac{0.9}{27} + \frac{0.9}{23}}} = 3.046$$

Per fare il test al livello $\alpha = 5\%$, dobbiamo confrontare questo valore con il quantile

$$z_{1-\alpha} = z_{0.95} = 1.645 \,.$$

Dal momento che $z_0 = 3.046 > 1.645 = z_{0.95}$, dobbiamo rifiutare H_0 e concludiamo che $\mu_A > \mu_B$ al 5% di significatività.

(b) Il p-value si ricava imponendo nella regola (*) l'uguaglianza

$$\begin{aligned} z_0 \equiv z_{1-\alpha} \quad &\Leftrightarrow \quad \Phi(z_0) = \Phi(z_{1-\alpha}) = 1 - \alpha \\ &\Leftrightarrow \quad \alpha = 1 - \Phi(z_0) = 1 - \Phi(3.046) = 1 - \frac{0.9982 + 0.9986}{2} = 0.0016 = 0.16\% \,. \end{aligned}$$

Quindi, p-value $= 0.16\%$, e perciò si rifiuta H_0 a tutti i livelli di significatività ragionevoli. Quando un test si conclude rifiutando H_0, la conclusione è sempre forte (si ha evidenza forte a favore di H_1).

(c) Ricordando che la differenza delle medie campionarie di due campioni gaussiani indipendenti ha *sempre* densità

$$\bar{X}_A - \bar{X}_B \sim N\left(\mu_A - \mu_B, \, \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}\right)$$

(indipendentemente dal fatto che H_0 sia vera oppure no), abbiamo

$$\begin{aligned}
\text{potenza} &= \mathbb{P}_{H_0 \text{ falsa}}(\text{“rifiuto } H_0\text{”}) \\
&= \mathbb{P}_{\mu_A=\mu_B+0.39}\left(\frac{\bar{X}_A-\bar{X}_B}{\sqrt{\frac{\sigma_A^2}{n_A}+\frac{\sigma_B^2}{n_B}}} > z_{1-\alpha}\right) \\
&= \mathbb{P}_{\mu_A=\mu_B+0.39}\Big(\underbrace{\frac{\bar{X}_A-\bar{X}_B-(\mu_A-\mu_B)}{\sqrt{\frac{\sigma_A^2}{n_A}+\frac{\sigma_B^2}{n_B}}}}_{\sim N(0,1)} > z_{1-\alpha}-\frac{\mu_A-\mu_B}{\sqrt{\frac{\sigma_A^2}{n_A}+\frac{\sigma_B^2}{n_B}}}\Big) \\
&= 1-\Phi\left(z_{1-\alpha}-\frac{\mu_A-\mu_B}{\sqrt{\frac{\sigma_A^2}{n_A}+\frac{\sigma_B^2}{n_B}}}\right) \qquad \text{con } \mu_A=\mu_B+0.39 \\
&= 1-\Phi\left(1.645-\frac{0.39}{\sqrt{\frac{0.9}{27}+\frac{0.9}{23}}}\right) = 1-\Phi(0.1962) = 1-\frac{0.57535+0.57926}{2} \\
&= 0.422695 = 42.2695\% \,.
\end{aligned}$$

(d) Recuperando l'espressione della potenza dal punto precedente, abbiamo

$$\begin{aligned}
\text{potenza} = 1-\Phi\left(z_{1-\alpha}-\frac{\mu_A-\mu_B}{\sqrt{\frac{\sigma_A^2}{n_A}+\frac{\sigma_B^2}{n_B}}}\right) &\qquad \text{con } \mu_A=\mu_B+0.39 \\
= 1-\Phi\left(1.645-\frac{0.39}{\sqrt{\frac{0.9}{n/2}+\frac{0.9}{n/2}}}\right) &\qquad \text{se ora } n_A=n_B=\frac{n}{2} \\
= 1-\Phi\left(1.645-\frac{0.39}{\sqrt{\frac{3.6}{n}}}\right) &\,.
\end{aligned}$$

Se vogliamo che la potenza del test sia maggiore o uguale all'80%, dobbiamo imporre

$$\begin{aligned}
&1-\Phi\left(1.645-\frac{0.39}{\sqrt{\frac{3.6}{n}}}\right) \geq 0.80 \\
&\Rightarrow \quad \Phi\left(1.645-\frac{0.39}{\sqrt{\frac{3.6}{n}}}\right) \leq 0.20 \\
&\Rightarrow \quad 1.645-\frac{0.39}{\sqrt{\frac{3.6}{n}}} \leq z_{0.20} = -z_{1-0.20} = -z_{0.80} = -0.84 \\
&\Rightarrow \quad \frac{3.6}{n} \leq \left(\frac{0.39}{0.84+1.645}\right)^2 \\
&\Rightarrow \quad n \geq 3.6\cdot\left(\frac{0.84+1.645}{0.39}\right)^2 = 146.159\,.
\end{aligned}$$

Poiché n deve essere un numero intero, la soluzione è $n \geq 147$.

Soluzione 4.

(a) Dobbiamo fare un test per le ipotesi

$$H_0 : \sigma_A^2 = \sigma_B^2 \qquad \text{contro} \qquad H_1 : \sigma_A^2 \neq \sigma_B^2 \,.$$

Poiché abbiamo a che fare con due popolazioni gaussiane indipendenti, una possibile regola di rifiuto al livello α è

$$\text{“rifiuto } H_0 \text{ se } F_0 < f_{\frac{\alpha}{2}}(m-1, n-1) \text{ oppure } F_0 > f_{1-\frac{\alpha}{2}}(m-1, n-1)\text{”}$$

dove

$$F_0 := \frac{S_X^2}{S_Y^2} \,.$$

Tuttavia, è spesso più facile lavorare con la regola equivalente che si ottiene negando la precedente affermazione, cioè

$$\text{“accetto } H_0 \text{ se } f_{\frac{\alpha}{2}}(m-1, n-1) \leq F_0 \leq f_{1-\frac{\alpha}{2}}(m-1, n-1)\text{”} \,.$$

Dai risultati della misura,

$$f_0 = \frac{s_A^2}{s_B^2} = 0.77319 \,,$$

mentre dalle tabelle, al livello $\alpha = 10\%$, si ha

$$\begin{aligned} f_{\frac{\alpha}{2}}(m-1, n-1) &= f_{0.05}(24, 15) = \frac{1}{f_{1-0.05}}(15, 24) = \frac{1}{f_{0.95}(15, 24)} = \frac{1}{2.108} = 0.4744 \\ f_{1-\frac{\alpha}{2}}(m-1, n-1) &= f_{0.95}(24, 15) = 2.288 \end{aligned}$$

Dal momento che $f_0 = 0.77319 \in (0.4744\,,\, 2.288)$, non possiamo rifiutare H_0 al 10% di significatività. Equivalentemente, p-value $> 10\%$, e con un p-value così elevato non possiamo azzardarci a rifiutare H_0 in nessun modo (conclusione debole). Dunque d'ora in poi potremo assumere $\sigma_A^2 = \sigma_B^2 =: \sigma^2$

(b) Uno stimatore non distorto della varianza comune σ^2 è la varianza pooled

$$S_P^2 = \frac{(n_A - 1)S_A^2 + (n_B - 1)S_B^2}{n_A + n_B - 2} \,.$$

Infatti, per verificarne la non distorsione dobbiamo dimostrare che $\mathbb{E}\left[S_P^2\right] = \sigma^2$; questa è in effetti una conseguenza immediata della linearità di $\mathbb{E}$ e del fatto che S_A^2 e S_B^2 sono stimatori non distorti di $\sigma_A^2 = \sigma_B^2 = \sigma^2$:

$$\begin{aligned} \mathbb{E}\left[S_P^2\right] = \mathbb{E}\left[\frac{(n_A - 1)S_A^2 + (n_B - 1)S_B^2}{n_A + n_B - 2}\right] &= \frac{(n_A - 1)\mathbb{E}\left[S_A^2\right] + (n_B - 1)\mathbb{E}\left[S_B^2\right]}{n_A + n_B - 2} \\ \underset{\substack{b(S_A^2;\sigma^2)=0 \\ b(S_B^2;\sigma^2)=0}}{=} \frac{(n_A - 1)\sigma^2 + (n_B - 1)\sigma^2}{n_A + n_B - 2} &= \sigma^2 \,. \end{aligned}$$

Se usiamo S_P^2 come stimatore, la stima (= realizzazione sui dati) corrispondente è

$$s_P^2 = \frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2} = \frac{(25-1) \cdot 0.51^2 + (16-1) \cdot 0.58^2}{25 + 16 - 2} = 0.289446 \,.$$

(c) Cerchiamo evidenza del fatto che $\mu_B > \mu_A \Rightarrow$ dobbiamo mettere questa affermazione come ipotesi alternativa. Impostiamo dunque il test per le ipotesi

$$H_0 : \mu_A \geq \mu_B \quad (\text{o } H_0 : \mu_A = \mu_B) \qquad \text{contro} \qquad H_1 : \mu_A < \mu_B$$

Avendo a che fare con due campioni gaussiani indipendenti a varianze incognite ma uguali (per il risultato del test al punto (a)), possiamo fare un T-test. Con le H_0, H_1 appena scelte, la regola di rifiuto è

$$\text{“rifiuto } H_0 \text{ se } T_0 := \frac{\bar{X} - \bar{Y}}{S_P\sqrt{\frac{1}{m} + \frac{1}{n}}} < -t_{1-\alpha}(m + n - 2)\text{”}$$

Dai dati otteniamo

$$t_0 = \frac{\bar{x}_A - \bar{x}_B}{s_P\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} = \frac{2.1 - 2.5}{\sqrt{0.289446}\sqrt{\frac{1}{25} + \frac{1}{16}}} = -2.3223$$

che al livello $\alpha = 2.5\%$ dobbiamo confrontare con

$$-t_{1-\alpha}(n_A + n_B - 2) = -t_{0.975}(39) \simeq -t_{0.975}(40) = -2.0211\,.$$

Dato che $-2.3223 < -2.0211$, rifiutiamo H_0 al 2.5% (conclusione forte).

(d) Il p-value si ricava imponendo l'uguaglianza nella regola di rifiuto del test, cioè risolvendo l'equazione

$$t_0 \equiv -t_{1-\alpha}(39)\,.$$

in α. Sulle tabelle troviamo

$$-t_{0.99}(39) \simeq -t_{0.99}(40) = -2.4233 < t_0 = -2.3223 < -2.0211 = -t_{0.975}(40) \simeq -t_{0.975}(39)$$

e quindi

$$0.975 < 1 - \alpha < 0.99 \quad \Rightarrow \quad 0.01 < \alpha = p\text{-value} < 0.025\,.$$

Un p-value $\in (1\%\,,\, 2.5\%)$, è molto basso, e quindi abbiamo una forte evidenza in favore dell'ipotesi alternativa H_1.

(e) Le ipotesi da testare sono le stesse di prima:

$$H_0 : \mu_A \geq \mu_B \quad (\text{o } H_0 : \mu_A = \mu_B) \qquad \text{contro} \qquad H_1 : \mu_A < \mu_B$$

Ora tuttavia, poiché le varianze sono note a priori, possiamo fare uno Z-test al posto del T-test. Da notare che in questo caso non dobbiamo richiedere che $\sigma_A^2 = \sigma_B^2$ (e difatti $0.51^2 \neq 0.58^2$). La regola di rifiuto diventa

$$\text{“rifiuto } H_0 \text{ se } Z_0 := \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} < -z_{1-\alpha}\text{”}\,.$$

Poiché

$$z_0 = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} = \frac{2.1 - 2.5}{\sqrt{\frac{0.51^2}{25} + \frac{0.58^2}{16}}} = -2.2563$$

$$-z_{1-\alpha} = -z_{1-0.025} = -z_{0.975} = -1.96\,.$$

abbiamo $z_0 < -z_{1-\alpha}$ con $\alpha = 2.5\%$, e quindi rifiutiamo H_0 anche in questo caso.
Notare che il p-value ora si può calcolare esattamente, ed è

$$p\text{-value} = \Phi(z_0) = \Phi(-2.2563) = 1 - \frac{0.98778 + 0.98809}{2} = 1.2065\%\,,$$

confrontabile con quello del punto (d).

(f) La potenza è

$$
\begin{aligned}
\text{potenza} &= \mathbb{P}_{H_0 \text{ falsa}}(\text{“rifiuto } H_0\text{”}) \\
&= \mathbb{P}_{\mu_A=1.9,\, \mu_B=2.4}\left(\frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} < -z_{1-\alpha} \right) \\
&= \mathbb{P}_{\mu_A=1.9,\, \mu_B=2.4}\Bigg(\underbrace{\frac{\bar{X}_A - \bar{X}_B - (\mu_A - \mu_B)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}}}_{\sim N(0,1)} < -z_{1-\alpha} - \frac{\mu_A - \mu_B}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \Bigg) \\
&= \Phi\left(-z_{1-\alpha} - \frac{\mu_A - \mu_B}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \right) \qquad \text{con } \mu_A = 1.9,\ \mu_B = 2.4 \\
&= \Phi\left(-1.96 - \frac{1.9 - 2.4}{\sqrt{\frac{0.51^2}{25} + \frac{0.58^2}{16}}} \right) = \Phi(0.860) \\
&= 0.80511 = 80.511\% \,,
\end{aligned}
$$

dove abbiamo usato il fatto che per le medie campionarie di due campioni gaussiani indipendenti vale *sempre*

$$
\bar{X}_A - \bar{X}_B \sim N\left(\mu_A - \mu_B,\, \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B} \right)
$$

Soluzione 7.

(a) Abbiamo due campioni gaussiani indipendenti a medie e varianze incognite, e dobbiamo fare un test per la differenza delle loro medie. I due campioni sono poco numerosi, dunque l'unico test che possiamo fare è un T-test. Tale test richiede però che le varianze dei due campioni, benché incognite, si possano assumere uguali. Di conseguenza, dobbiamo fare preliminarmente un test per le ipotesi

$$
H_0 : \sigma_X^2 = \sigma_Y^2 \qquad \text{contro} \qquad H_1 : \sigma_X^2 \neq \sigma_Y^2 \,. \tag{*}
$$

Dopodiché, se il test si concluderà accettando H_0, potremo procedere al T-test. In caso contrario, non saremo in grado di rispondere alla domanda.
La regola di rifiuto (o, meglio, quella di *accettazione*) di un test al livello α per le ipotesi (*) è

$$
\text{“accetto } H_0 \text{ se } f_{\frac{\alpha}{2}}(m-1, n-1) \leq F_0 := \frac{S_X^2}{S_Y^2} \leq f_{1-\frac{\alpha}{2}}(m-1, n-1)\text{”}
$$

Scegliamo $\alpha = 5\%$, il livello di significatività più usuale. In questo modo,

$$
\begin{aligned}
f_{\frac{\alpha}{2}}(m-1, n-1) &= f_{0.025}(12, 15) = \frac{1}{f_{1-0.025}(15, 12)} = \frac{1}{f_{0.975}(15, 12)} = \frac{1}{3.177} = 0.3148 \\
f_{1-\frac{\alpha}{2}}(m-1, n-1) &= f_{0.975}(12, 15) = 2.963 \,,
\end{aligned}
$$

mentre dai dati otteniamo

$$
f_0 := \frac{s_X^2}{s_Y^2} = \frac{0.094}{0.132} = 0.7121
$$

Poiché $0.3148 < 0.7121 < 2.963$, non possiamo rifiutare H_0. Dunque possiamo assumere che le due varianze siano uguali e procedere al T-test.
Se vogliamo essere sicuri che il diametro medio dei nuovi Taralluccios sia più grande di quello dei vecchi, dobbiamo mettere quest'ultima affermazione nell'ipotesi alternativa, cioè

$$H_0 : \mu_X = \mu_Y \quad (\text{o } \mu_X \geq \mu_Y) \qquad \text{contro} \qquad H_1 : \mu_X < \mu_Y\,.$$

La regola di rifiuto per il T-test con queste ipotesi è

$$\text{“rifiuto } H_0 \text{ se } T_0 := \frac{\bar{X} - \bar{Y}}{S_P\sqrt{\frac{1}{m} + \frac{1}{n}}} < -t_{1-\alpha}(m+n-2)\text{”}$$

Dopo aver fatto le misure, abbiamo trovato

$$s_P^2 = \frac{(m-1)s_X^2 + (n-1)s_Y^2}{m+n-2} = \frac{(13-1)\cdot 0.094 + (16-1)\cdot 0.132}{13+16-2} = 0.115111$$

e quindi

$$t_0 = \frac{1.92 - 2.23}{\sqrt{0.115111}\cdot\sqrt{\frac{1}{13} + \frac{1}{16}}} = -2.4470\,.$$

Con $\alpha = 5\%$,

$$-t_{1-\alpha}(m+n-2) = -t_{0.95}(27) = -1.7033\,,$$

e quindi, poiché $-2.4470 < -1.7033$, rifiutiamo H_0 e accettiamo l'ipotesi alternativa che i nuovi Taralluccios abbiano diametro medio maggiore di quelli vecchi (conclusione *forte*).

(b) Per trovare il p-value del T-test precedente, troviamo la soglia α per cui si ha esattamente

$$t_0 \equiv -t_{1-\alpha}(m+n-2) \quad \Leftrightarrow \quad -2.4470 = -t_{1-\alpha}(27)\,.$$

Dal momento che

$$-t_{0.99}(27) = -2.4727 < t_0 = -2.4470 < -t_{0.975}(27) = -2.0518\,,$$

abbiamo

$$0.975 < 1-\alpha < 0.99 \quad \Rightarrow \quad 0.01 < \alpha = p\text{-value} < 0.025\,.$$

Con un p-value così basso, abbiamo evidenza molto forte in favore di H_1; quindi, l'azienda può sentirsi molto tranquilla nel lanciare la campagna pubblicitaria.

Soluzione 8.

Vogliamo stabilire se (= vedere se c'è evidenza del fatto che) P_{new} è migliore di P_{old}; cioè, vogliamo dimostrare che il parametro $\mu_X =$ tempo di trasmissione medio del nuovo protocollo è minore del tempo medio μ_Y di quello vecchio; di conseguenza, mettiamo quest'ultima affermazione come ipotesi alternativa: $H_1 : \mu_X < \mu_Y$. Impostiamo pertanto il test d'ipotesi

$$H_0 : \mu_X \geq \mu_Y \quad (\text{o } \mu_X = \mu_Y) \qquad \text{contro} \qquad H_1 : \mu_X < \mu_Y\,. \tag{*}$$

Dal momento che le varianze σ_X^2, σ_Y^2 dei due campioni sono entrambe incognite (l'esercizio non ci dice nulla a riguardo) e i due campioni sono poco numerosi, dobbiamo fare un T-test per la differenza delle medie di due campioni gaussiani a varianza incognita. Tale test richiede però che sia $\sigma_X^2 \equiv \sigma_Y^2$. Dobbiamo quindi controllare preliminarmente se si può assumere questa ipotesi oppure no (e sperare di sì, perché

altrimenti non avremmo nessun altro test a disposizione).
Testiamo quindi le ipotesi

$$H_0 : \sigma_X^2 = \sigma_Y^2 \qquad \text{contro} \qquad H_1 : \sigma_X^2 \neq \sigma_Y^2 \,. \tag{**}$$

Invece di fissare noi arbitrariamente il livello di significatività di questo test, calcoliamone il p-value e traiamone delle conclusioni. La regola di un test al livello α per le ipotesi (**) è

$$\text{“accetto } H_0 \text{ se } f_{\frac{\alpha}{2}}(m-1, n-1) \leq F_0 \leq f_{1-\frac{\alpha}{2}}(m-1, n-1)\text{”} \tag{o}$$

(come al solito, è più semplice scriversi la regola di accettazione anziché di rifiuto per H_0). Dai dati raccolti si ottiene

$$\begin{aligned} \bar{x} &= 1.89571 & \bar{y} &= 2.93333 \\ s_X^2 &= 0.27000 & s_Y^2 &= 1.34467 \end{aligned}$$

e quindi

$$f_0 = \frac{s_X^2}{s_2^Y} = \frac{0.27000}{1.34467} = 0.2008 \,.$$

L'intervallo (o) diventa tanto più piccolo quanto più α diventa grande; è tutto $\mathbb{R}$ al limite $\alpha \to 0^+$, mentre tende all'insieme di un solo elemento $\{1\}$ quando $\alpha \to 100\%^-$. Facendo variare α tra 0 e il 100%, il primo valore per cui f_0 (fissato) tocca uno dei due estremi è il p-value. Ora, il primo estremo $f_{\frac{\alpha}{2}}(m-1, n-1)$ è senz'altro < 1, mentre il secondo estremo $f_{1-\frac{\alpha}{2}}(m-1, n-1)$ è sempre > 1; infatti, la mediana della densità F di Fisher è $f_{0.50}(h, k) = 1$, e per ogni α compreso tra lo 0 e il 100% abbiamo sempre $\frac{\alpha}{2} < 50\%$ e $1 - \frac{\alpha}{2} > 50\%$. Poiché nel nostro caso $f_0 = 0.2008 < 1$, l'estremo toccato da f_0 dovrà essere per forza il primo. Di conseguenza, il p-value si ricava risolvendo in α l'equazione

$$f_0 \equiv f_{\frac{\alpha}{2}}(m-1, n-1) \quad \Leftrightarrow \quad 0.2008 \equiv f_{\frac{\alpha}{2}}(6, 5) \,.$$

Dobbiamo cercare sulle tavole quali sono i quantili che si avvicinano di più a 0.2008. Dal momento però che le tavole dànno solo i quantili > 1, bisogna trasformare l'equazione precedente prendendone il reciproco:

$$\frac{1}{0.2008} = 4.9801 \equiv \frac{1}{f_{\frac{\alpha}{2}}(6, 5)} = f_{1-\frac{\alpha}{2}}(5, 6) \,.$$

Ora possiamo cercare sulle tavole i quantili che si avvicinano di più a 4.9801 (ma coi gradi di libertà invertiti):

$$f_{0.90}(5, 6) = 3.108 \qquad f_{0.95}(5, 6) = 4.387 \qquad f_{0.975}(5, 6) = 5.988 \qquad \dots$$

Vediamo che

$$\begin{aligned} & f_{0.95}(5, 6) = 4.387 < 4.9801 \equiv f_{1-\frac{\alpha}{2}}(5, 6) < 5.988 = f_{0.975}(5, 6) \\ & \quad \Rightarrow \quad 0.95 < 1 - \frac{\alpha}{2} < 0.975 \\ & \quad \Rightarrow \quad 0.05 < \alpha < 0.10 \,. \end{aligned}$$

Abbiamo quindi trovato p-value $\in (5\%, 10\%)$. Non è un valore altissimo, ma comunque ci fa accettare $H_0 : \sigma_X^2 = \sigma_Y^2$ almeno al 5% di significatività. È lecito quindi supporre che le due varianze siano uguali (conclusione debole), e procedere quindi finalmente al T-test originario.
Le ipotesi da mettere alla prova col T-test erano (*). L'ipotesi alternativa $H_1 : \mu_X < \mu_Y$ impone che la regola del test sia

$$\text{“rifiuto } H_0 \text{ se } T_0 = \frac{\bar{X} - \bar{Y}}{S_P \sqrt{\frac{1}{m} + \frac{1}{n}}} < -t_{1-\alpha}(m+n-2)\text{”}$$

Coi dati raccolti,

$$
\begin{aligned}
s_P &= \sqrt{\frac{(m-1)s_X^2+(n-1)s_Y^2}{m+n-2}} = \sqrt{\frac{(7-1)\cdot 0.27000+(6-1)\cdot 1.34467}{7+6-2}} = 0.87091 \\
t_0 &= \frac{1.89571-2.93333}{0.87091\cdot\sqrt{\frac{1}{7}+\frac{1}{6}}} = -2.1415\,.
\end{aligned}
$$

Dal momento che l'esercizio non fissa il livello di significatività del test, troviamone il p-value e cerchiamo di trarre una conclusione dal suo valore. Dobbiamo risolvere l'equazione

$$t_0 \equiv -t_{1-p\text{-value}}(m+n-2) = -t_{1-p\text{-value}}(11)\,.$$

Dalle tavole ricaviamo

$$
\begin{aligned}
&-t_{0.95}(11) = -1.7959 < t_0 = -2.1415 < -t_{0.975}(11) = -2.2010 \\
&\quad\Rightarrow\quad 0.95 < 1-p\text{-value} < 0.975 \\
&\quad\Rightarrow\quad 2.5\% < p\text{-value} < 5\%\,.
\end{aligned}
$$

Con un p-value così piccolo – ma non piccolissimo – abbiamo evidenza forte – sebbene non fortissima – per rifiutare H_0 in favore di H_1 (rifiuteremmo H_0 al 5%, ma la accetteremmo al 2.5%).
Nota: Se i due campioni fossero stati entrambi numerosi (diciamo $m > 30$, $n > 30$), avremmo potuto fare lo Z-test che usa la statistica test $Z_0 = \frac{\bar{X}-\bar{Y}}{\sqrt{\frac{S_X^2}{m}+\frac{S_Y^2}{n}}}$. Quando è vera $H_0 : \mu_X = \mu_Y$, tale statistica è normale standard indipendentemente dalla densità delle X_i, Y_j ($\Rightarrow$ non sarebbe servita l'ipotesi di gaussianità) e indipendentemente dal fatto che $\sigma_X^2 = \sigma_Y^2$ ($\Rightarrow$ non sarebbe servita l'ipotesi di uguaglianza delle varianze).

Soluzione 9.

(a) La precisione di una macchina è la deviazione standard degli errori $E_i \sim N(0,\sigma^2)$ dei fili da essa prodotti: tanto più questa è piccola, tanto più la macchina è precisa. L'esercizio chiede se i dati *mostrano evidenza* del fatto che la macchina A è più precisa della macchina B (cioè che $\sigma_A < \sigma_B$); in altre parole, cerca una conclusione forte in questo senso. Di conseguenza, dobbiamo mettere $\sigma_A < \sigma_B$ come ipotesi alternativa. Le ipotesi corrette sono pertanto

$$H_0 : \sigma_A = \sigma_B \quad (\text{o } H_0 : \sigma_A \geq \sigma_B) \qquad \text{contro} \qquad H_1 : \sigma_A < \sigma_B\,. \tag{*}$$

La regola di un test di livello α per le ipotesi precedenti è

$$\text{“rifiuto } H_0 \text{ se } F_0 := \frac{S_A^2}{S_B^2} < f_\alpha(9,15)\text{”}\,, \tag{o}$$

dove S_A^2 è la varianza campionaria delle prime $m = 10$ misure, e S_B^2 quella delle ultime $n = 16$. Per calcolare il valore misurato per la statistica test F_0, bisogna innanzitutto determinare il valore delle realizzazioni s_A^2 e s_B^2 coi dati ottenuti. Siccome l'esercizio non dà s_A^2 e s_B^2 direttamente, ma fornisce invece le realizzazioni del momento primo e del momento secondo campionari, usiamo la formula alternativa per la varianza

$$
\begin{aligned}
s_A^2 &= \frac{1}{m-1}\left[\left(\sum_{i=1}^{m} e_{i,A}^2\right) - m\cdot \bar{e}_A^2\right] = \frac{1}{m-1}\left[\left(\sum_{i=1}^{m} e_{i,A}^2\right) - \frac{1}{m}\left(\sum_{i=1}^{m} e_{i,A}\right)^2\right] \\
&= \frac{1}{10-1}\left[0.017 - \frac{1}{10}\cdot(0.1304)^2\right] = 0.001700 \\
s_B^2 &= \frac{1}{n-1}\left[\left(\sum_{i=1}^{n} e_{i,B}^2\right) - \frac{1}{n}\left(\sum_{i=1}^{n} e_{i,B}\right)^2\right] = \frac{1}{16-1}\left[0.095 - \frac{1}{16}\cdot(0.3194)^2\right] = 0.005908\,.
\end{aligned}
$$

Quindi,

$$f_0 = \frac{0.001700}{0.005908} = 0.2877\,,$$

mentre con $\alpha = 10\%$

$$f_\alpha(9,15) = f_{0.10}(9,15) = \frac{1}{f_{0.90}(15,9)} = \frac{1}{2.340} = 0.4274\,.$$

Poiché $0.2877 < 0.4274$, seguendo la regola (o) siamo costretti a rifiutare H_0 al 10% e concludiamo che $\sigma_A < \sigma_B$. Questa tuttavia non è una conclusione fortissima: il 10% è comunque un livello di significatività abbastanza alto, mentre la conclusione a favore di H_1 è tanto più forte quanto più H_0 si rigetta a livelli di significatività bassi (5% o anche meno; si veda il calcolo del p-value al punto seguente).

(b) Dalla regola di rifiuto (o), il p-value si ricava risolvendo l'equazione

$$f_0 \equiv f_{p\text{-value}}(9,15) \quad \Leftrightarrow \quad \frac{1}{f_0} \equiv \frac{1}{f_{p\text{-value}}(9,15)} = f_{1-p\text{-value}}(15,9)$$

(siamo passati al reciproco perché $f_0 < 1$, mentre sulle tavole si trovano solo i quantili $f_\alpha(h,k) > 1$). Abbiamo

$$\frac{1}{f_0} = \frac{1}{0.2877} = 3.4758\,,$$

mentre guardando le tavole troviamo

$$f_{0.95}(15,9) = 3.006 \qquad\qquad f_{0.975}(15,9) = 3.769\,.$$

Di conseguenza,

$$
\begin{aligned}
& f_{0.95}(15,9) < \frac{1}{f_0} < f_{0.975}(15,9) \\
&\Rightarrow \quad 0.95 < 1 - p\text{-value} < 0.975 \\
&\Rightarrow \quad 2.5\% < p\text{-value} < 5\%\,.
\end{aligned}
$$

Tale p-value è abbastanza piccolo da poter rifiutare H_0 al livello tipico del 5%, e mostra pertanto che H_1 è molto verosimile (conclusione forte).

Nota: Se nella scelta delle ipotesi (*), avessimo (erroneamente!) scambiato H_0 con H_1, cioè

$$H_0 : \sigma_A = \sigma_B \;\; (\text{o } H_0 : \sigma_A \leq \sigma_B) \qquad \text{contro} \qquad H_1 : \sigma_A > \sigma_B\,, \tag{**}$$

la regola di rifiuto al livello α sarebbe stata

$$\text{“rifiuto } H_0 \text{ se } \frac{S_A^2}{S_B^2} > f_{1-\alpha}(9,15)\text{”}\,.$$

Dal momento che senz'altro $f_{1-\alpha}(9,15) > 1$ per ogni α tra lo 0 e il 50%, mentre dai dati abbiamo visto che $s_A^2 = 0.001700 < s_B^2 = 0.005908 \Rightarrow f_0 < 1$, il test al 10% si sarebbe concluso accettando H_0 senza neanche guardare le tavole delle distribuzioni. In altre parole, se vediamo già dai dati che $s_A^2 < s_B^2$, non ha senso fare un test per le ipotesi (**); tale test, infatti, darebbe senz'altro un p-value altissimo (maggiore del 50%), e non permetterebbe di trarre nessuna conclusione forte. Ciò è una manifestazione particolare del seguente fatto generale: nei test unilateri, spesso le ipotesi corrette sono “suggerite” dai dati. Per esempio, in un test sulla differenza delle medie, se troviamo $\bar{x} < \bar{y}$, ha molto più senso fare un test con $H_1 : \mu_X < \mu_Y$ piuttosto che con $H_1 : \mu_X > \mu_Y$. Nel primo caso, infatti *forse potremmo trovare una conclusione forte* a favore di $H_1 : \mu_X < \mu_Y$; nel secondo, invece, *senza dubbio troveremmo una conclusione debole* a favore di $H_0 : \mu_X \leq \mu_Y$ (sapremmo che il p-value è maggiore del 50% senza nemmeno fare i calcoli del test).

(c) Uno stimatore non distorto e consistente in media quadratica della varianza incognita σ_A^2 è la varianza campionaria S_A^2 delle prime $m = 10$ misure. La stima corrispondente è il valore $s_A^2 = 0.001700$ trovato al punto (a). Similmente, una stima per σ_B^2 è la realizzazione della varianza campionaria S_B^2 sulla seconda serie di $n = 16$ misure, cioè il valore $s_B^2 = 0.005908$ già calcolato.

Nota: Se avessimo potuto assumere $\sigma_A^2 \equiv \sigma_B^2 =: \sigma^2$, avremmo potuto usare la varianza pooled $S_P^2 = \frac{(m-1)S_A^2+(n-1)S_B^2}{m+n-2}$ come unico stimatore non distorto della varianza comune σ^2. Ciò sarebbe stato preferibile, perchè si può verificare facilmente che

$$\text{mse}(S_P^2;\sigma^2) < \text{mse}(S_A^2;\sigma_A^2) \qquad \text{e} \qquad \text{mse}(S_P^2;\sigma^2) < \text{mse}(S_B^2;\sigma_B^2)\,.$$

Tuttavia, il risultato del test ai punti (a)-(b) ci impedisce di assumere $\sigma_A^2 = \sigma_B^2$, e quindi di usare S_P^2 come stimatore non distorto delle due varianze.

Soluzione 14.

(a) Abbiamo un campione bernoulliano di numerosità n, con $X_i \sim B(1,p)$, dove p è la probabilità incognita che un elettore a caso vada a votare. Un $IC_p(\gamma)$ bilatero è

$$p \in \left(\bar{x} - z_{\frac{1+\gamma}{2}} \sqrt{\frac{\bar{x}(1-\bar{x})}{n}}\,,\, \bar{x} + z_{\frac{1+\gamma}{2}} \sqrt{\frac{\bar{x}(1-\bar{x})}{n}} \right)\,,$$

la cui ampiezza è

$$\delta = 2\,z_{\frac{1+\gamma}{2}} \sqrt{\frac{\bar{x}(1-\bar{x})}{n}} \leq 2\,z_{\frac{1+\gamma}{2}} \sqrt{\frac{\frac{1}{4}}{n}} = \frac{1}{\sqrt{n}}\,z_{\frac{1+\gamma}{2}}\,.$$

in quanto $\bar{x}(1-\bar{x}) \in [0\,,\,1/4]$ *sempre*. Se vogliamo essere sicuri che $\delta \leq 0.02$, basta richiedere

$$\frac{1}{\sqrt{n}}\,z_{\frac{1+\gamma}{2}} \leq 0.02 \quad \Rightarrow \quad n \geq \left(\frac{z_{\frac{1+\gamma}{2}}}{0.02}\right)^2 = \left(\frac{z_{0.995}}{0.02}\right)^2 \qquad \text{al livello } \gamma = 0.99$$
$$= \left(\frac{2.575}{0.02}\right)^2 = 16576.56\,.$$

Bisogna quindi intervistare almeno $n = 16577$ persone.

(b) Chiedersi se i dati a disposizione consentono di affermare che in Lombardia il quorum *non* verrà raggiunto significa cercare evidenza forte del fatto che la proporzione p di votanti è minore del 50%. Questa affermazione va quindi messa come ipotesi alternativa:

$$H_0 : p \geq 0.5 \quad (\text{o } H_0 : p = 0.5) \qquad \text{contro} \qquad H_1 : p < 0.5\,.$$

La regola di un test al livello α per le ipotesi precedenti è

$$\text{“rifiuto } H_0 \text{ se } Z_0 := \frac{\bar{X} - 0.5}{\sqrt{0.5 \cdot (1 - 0.5)}}\sqrt{n} < -z_{1-\alpha}\text{”}\,,$$

dove $\bar{x}$ è la frequenza empirica di votanti. Coi dati dell'esercizio,

$$n = 1000 \qquad \bar{x} = \frac{480}{1000} = 0.48 \quad \Rightarrow \quad z_0 = \frac{0.48 - 0.5}{\sqrt{0.5 \cdot (1 - 0.5)}}\sqrt{1000} = -1.265\,,$$

che al 5% va confrontata col quantile

$$-z_{1-\alpha} = -z_{1-0.05} = -z_{0.95} = -1.645\,.$$

Dal momento che $-1.265 \not< -1.645$, H_0 non si rifiuta e quindi *non* possiamo affermare che in Lombardia il quorum *non* verrà raggiunto (conclusione debole a favore di $H_0 : p \geq 0.5$).

(c) Ora abbiamo un nuovo campione di altri 1000 intervistati in Emilia Romagna, che chiameremo $Y_1, \ldots, Y_{1000}$ ($X_1, \ldots, X_{1000}$ era il campione della Lombardia). Anche le Y_i sono bernoulliane, ma, per distinguere il loro parametro incognito da quello degli intervistati in Lombardia, chiamiamo p_{ER} e p_L le rispettive probabilità (p_L è la vecchia p dei punti (a) e (b)). Ora si richiede di fare un test per le ipotesi

$$H_0 : p_L \leq p_{ER} \qquad \text{contro} \qquad H_1 : p_L > p_{ER}\,.$$

Si tratta di un test per la differenza delle frequenze. Una regola di rifiuto al livello α è

$$\text{“rifiuto } H_0 \text{ se } Z_0 := \frac{\bar{X} - \bar{Y}}{\sqrt{\hat{P}(1 - \hat{P})\left(\frac{1}{m} + \frac{1}{n}\right)}} > z_{1-\alpha}\text{”}\,.$$

Abbiamo già trovato $\bar{x}$ al punto precedente, mentre ora in più ci viene detto che $\bar{y} = \frac{470}{1000} = 0.47$. Dobbiamo ancora calcolare

$$\hat{p} = \frac{m\bar{x} + n\bar{y}}{m + n} = \frac{480 + 470}{1000 + 1000} = 0.475$$

e quindi

$$z_0 = \frac{0.48 - 0.47}{\sqrt{0.475 \cdot (1 - 0.475) \cdot \left(\frac{1}{1000} + \frac{1}{1000}\right)}} = 0.448\,.$$

Al livello $\alpha = 5\%$, questo valore NON è maggiore di $z_{1-\alpha} = z_{0.95} = 1.645$, e di conseguenza NON possiamo rifiutare H_0.