Corso di Statistica per Ingegneria Fisica
Anno accademico 2020/2021

Esercitazione 8 - Test d'ipotesi per media, varianza e proporzione di una popolazione

Esercizio 1. Per decidere se sia il caso di revisionare i propri impianti, una nota marca di pasta decide di verificare l'attendibilità del peso dichiarato sulle proprie confezioni di spaghetti. Su un campione di 60 scatole da 500 g viene trovata una media campionaria di 496 g. Si supponga che la quantità di spaghetti contenuta in una confezione, espressa in grammi, possa essere descritta con una variabile normale di media μ e di varianza $200\,\mathrm{g}^2$, nota da precedenti indagini.

(a) Si introduca un opportuno test statistico e si tragga una conclusione al livello di significatività del 1%, 5% e 10%. *[Test per $H_0 : \mu = 500$ vs. $H_1 : \mu \neq 500$. Non si rifiuta H_0 al livello 1%, ma la si rifiuta ai livelli 5% e 10%.]*

(b) Si calcoli il minimo livello di significatività con cui si sarebbe giunti alla revisione degli impianti. *[2.8% .]*

(c) Per quale valore di $\mu \in \{490, 493.5, 496, 497.1\}$ è maggiore la probabilità di errore di seconda specie dei test costruiti al punto (a)? *[$\mu = 497.1$]*

Si supponga ora che gli stessi dati siano stati ottenuti in un'indagine condotta da un'associazione di consumatori.

(d) Possono i consumatori concludere con un opportuno test statistico al livello $\alpha = 1\%$ che le confezioni contengono meno di quanto dichiarato? *[Test per $H_0 : \mu = 500$ vs. $H_1 : \mu < 500$. Non si rifiuta H_0 al livello 1%.]*

Esercizio 2. Una certa ditta produce filo spinato in matasse nominalmente da 100 m l'una, ma la cui lunghezza X in realtà è una variabile aleatoria distribuita secondo la densità $N(\mu, 4)$.

(a) Scrivere la regione critica di un test di significatività α per sottoporre a verifica le ipotesi

$$H_0 : \ \mu = 100 \qquad \text{vs.} \qquad H_1 : \ \mu \neq 100$$

sulla base di un campione di numerosità n. *[$RC(\alpha) = \left\{ \frac{\sqrt{n}}{2} \left| \overline{X}_n - 100 \right| > z_{1-\frac{\alpha}{2}} \right\}$]*

Vengono esaminate 10 matasse, rilevando le lunghezze:

$$97.8 \quad 96.5 \quad 99.6 \quad 102.5 \quad 100.9 \quad 102.3 \quad 98.3 \quad 103.1 \quad 98.7 \quad 104.0\,.$$

(b) Al livello di significatività $\alpha = 5\%$, si può rifiutare l'ipotesi nulla?
[$\frac{\sqrt{n}}{2} |\overline{x}_n - 100| = 0.585 \not> z_{0.975} = 1.96 \Rightarrow$ Non si può rifiutare H_0.]

(c) Determinare, sulla base dei dati osservati, il p-value del test. *[55.856%]*

Esercizio 3. In un processo chimico viene utilizzato come reagente una soluzione il cui pH deve necessariamente avere valore $\mu > 8.20$; in caso contrario, infatti, il processo non può avvenire. Il metodo usato per determinare il pH della soluzione fornisce misurazioni che sono distribuite secondo una normale con media uguale a μ e deviazione standard uguale a 0.02. Il campione su cui ci baseremo sono 25 di tali misurazioni.

(a) Scrivere le ipotesi statistiche di un test in cui, coerentemente con quanto detto sopra, l'errore più grave consista nel ritenere $\mu > 8.20$ quando in realtà ciò non è vero. *[$H_0 : \mu \leq 8.20$ vs. $H_1 : \mu > 8.20$]*

(b) Se per testare le ipotesi statistiche del punto (a) usiamo la regione critica $RC = \{\bar{X}_{25} > 8.2093\}$, qual è il livello di significatività del nostro test? *[$\alpha = 1.0036\%$]*

(c) Calcolare la probabilità di errore di seconda specie del test del punto (b) quando:

(1) $\mu = 8.2093$; *[50%]*

(2) $\mu = 8.215$. *[7.7078%]*

(d) Calcolare la potenza del test del punto (b) quando:

(1) $\mu = 8.2093$; *[50%]*

(2) $\mu = 8.215$. *[92.2922%]*

(e) Eseguire il test del punto (b) sapendo che la media campionaria delle 25 misurazioni ha preso il valore $\bar{x}_{25} = 8.206$. *[Non si può rifiutare H_0]*

Esercizio 4. In uno studio clinico si vogliono osservare le concentrazioni di colesterolo totale, espresse in mg/dl, in n pazienti a cui viene somministrato il nuovo farmaco Colmeno. Supponiamo che i valori $X_1, \ldots, X_n$ di tali concentrazioni siano variabili aleatorie indipendenti e con densità gaussiana di media μ incognita e varianza $50\,(\text{mg/dl})^2$.

(a) Scrivere ipotesi nulla, ipotesi alternativa e regione critica al livello α di un test statistico volto a dimostrare che la concentrazione di colesterolo media, in un paziente che assuma il farmaco Colmeno, è inferiore a 200 mg/dl. *[$H_0 : \mu \geq 200$ vs. $H_1 : \mu < 200$; $RC(\alpha) = \left\{ \frac{\bar{x}-200}{\sigma}\sqrt{n} < -z_{1-\alpha} \right\}$]*

(b) Scegliere il livello α e la numerosità n del campione, affinché il test impostato al punto (a) abbia una probabilità di errore di primo tipo pari al 5% e una potenza, nel caso in cui $\mu = 190$ mg/dl, almeno del 90%. *[$\alpha = 0.05, n \geq 5$]*

(c) Se su 10 pazienti si è osservata una media campionaria pari a 198.5 mg/dl, stabilire quale conclusione si può trarre ad un livello del 5%. *[Non si può rifiutare H_0]*

Esercizio 5. In uno studio atmosferico si sono misurate, su 8 diversi campioni di aria di una certa città, le seguenti concentrazioni della sostanza tossica COUGH (espressi in microgrammi per metro cubo, $\mu\text{g/m}^3$):

$$2.3 \quad 1.7 \quad 3.2 \quad 2.1 \quad 2.3 \quad 2.0 \quad 2.2 \quad 1.2\,.$$

(a) Utilizzando stimatori non distorti, fornire una stima puntuale della media e della varianza della concentrazione di COUGH nell'atmosfera. *[$\widehat{\mu} = 2.125$; $\widehat{\sigma^2} = 0.325$]*

Si assuma ora che la concentrazione di COUGH abbia una distribuzione normale. Recenti studi clinici hanno dimostrato che il COUGH diventa pericoloso per la salute dell'uomo quando la sua concentrazione media nell'aria è superiore a $2.7\,\mu\text{g/m}^3$.

(b) Utilizzando un opportuno test statistico al 5% di significatività, stabilire se i dati raccolti consentono di affermare che l'aria da cui provengono i campioni non è dannosa per la salute dell'uomo. *[Test per le ipotesi $H_0 : \mu \geq 2.7$ vs. $H_1 : \mu < 2.7$. Si rifiuta H_0 al livello del 5%.]*

(c) Calcolare il p-value del test del punto precedente. Con un tale p-value, vi sentireste tranquilli a respirare l'aria da cui provengono i campioni? *[1% < p-value < 2.5%. Mi sentirei abbastanza tranquillo.]*

(d) Supponete invece che i recenti studi clinici dimostrino che il COUGH diventa pericoloso quando la sua concentrazione media nell'aria supera i $2.4\,\mu\text{g/m}^3$. Come cambia il p-value del test precedente? Vi sentireste sempre tranquilli a respirare l'aria da cui provengono i campioni? *[p-value $> 10\%$. Non mi sentirei affatto tranquillo.]*

Esercizio 6. Dall'esperienza è noto che il numero di guasti a un certo apparecchio è in media di 3 alla settimana. Nel corso dell'ultimo anno ci sono stati 234 guasti in tutto, con una deviazione standard campionaria del numero di guasti alla settimana pari a 1.8 .
Si può affermare che il numero di guasti è aumentato in modo significativo rispetto agli anni precedenti? *[Sì. Perché?]*

Esercizio 7. La concentrazione dello zinco nell'acqua di Lago Lungo è normalmente distribuita. Un campione di cinque misurazioni indipendenti fornisce i seguenti risultati (in mg/m^3):

$$10.20 \quad 9.85 \quad 10.15 \quad 9.97 \quad 10.40\,.$$

(a) Stimare media e varianza della concentrazione dello zinco nell'acqua di Lago Lungo. *[$\overline{x}_5 = 10.114$; $s^2 = 0.04523$]*

(b) I dati permettono di concludere che il valore della media è diverso da 10 mg/m^3? Rispondere effettuando un test a livello di significatività $\alpha = 1\%$, 5% e 10%. *[Test per le ipotesi $H_0 : \mu = 10$ vs. $H_1 : \mu \neq 10$. In nessun caso si rifiuta H_0.]*

(c) Valutare il p-value del test del punto (b). *[$0.2 <$ p-value < 0.3 (valore esatto: p-value $= 0.297$).]*

Esercizio 8. L'altezza media dei soldati di leva di un certo paese era di 170 cm nel 1950. Su 100 reclute nel 1960 la media campionaria era $\overline{x} = 171.2$ cm con una varianza campionaria $s^2 = 15.9$ cm^2.

(a) Con significatività $\alpha = 5\%$, si può dire che l'altezza media sia cambiata? *[Sì. Perché?]*

(b) Con significatività $\alpha = 1\%$ si può dire che l'altezza media sia aumentata? *[Sì. Perché?]*

(c) Calcolare il p-value del test del punto (b). C'è forte evidenza che l'altezza media sia aumentata? *[p-value $\simeq 0.131\%$. Sì, molto forte.]*.

Esercizio 9. L'etichetta delle bottiglie di champagne Veuve Coquelicot dichiara un contenuto di 730 ml. Un'associazione di consumatori decide di controllare questa affermazione e su 81 bottiglie esaminate riscontra una media campionaria $\bar{x} = 726$ ml ed una varianza campionaria $s^2 = 625\,\text{ml}^2$. Supponendo che la quantità di champagne contenuta in ogni bottiglia si possa modellizzare con una variabile aleatoria normale, si può concludere (al livello di significatività $\alpha = 5\%$) che in media le bottiglie contengono meno di quanto dichiarato? *[Test per le ipotesi $H_0 : \mu = 730$ vs. $H_1 : \mu < 730$. Non si può rifiutare H_0.]*

Esercizio 10. Per ottenere il marchio "DOC" dal consorzio dei produttori di vino della sua zona, un viticoltore deve dimostrare che la percentuale media di alcool contenuta nel suo vino è maggiore di 13 cl/l. Misurando il grado alcoolico in cinque bottiglie da lui prodotte, si rilevano i valori (in cl/l):

$$12.78 \qquad 15.83 \qquad 13.25 \qquad 13.81 \qquad 14.24\,.$$

(a) Supponendo che il grado alcolico X di ogni bottiglia sia normalmente distribuito, si effettui un test al livello di significatività del 5% per stabilire se in base a questi dati c'è evidenza che il produttore abbia diritto al marchio DOC. *[Non c'è evidenza.]*

(b) Supponendo in più di sapere che la deviazione standard di X è pari a 1, si ripeta il test precedente. *[C'è evidenza.]*

(c) Nelle situazioni dei punti (a) e (b), si calcoli il p-value. *[Nel punto (a),* $5\% < $ *p-value* $< 10\%$. *Nel punto (b), p-value* $= 1.405\%$.*]*

Esercizio 11. Una lamentela tipica degli utenti di file su server è la forte (alta) varianza σ^2 del tempo di risposta (che misuriamo in millisecondi quadrati, ms^2). La Yasdam sta valutando la possibilità di sostituire il vecchio server, il cui tempo di risposta ha varianza pari a $\sigma^2 = 20\,\mathrm{ms}^2$, con uno nuovo. Prima di procedere alla sostituzione, però, vuole essere abbastanza sicura che il tempo di risposta del nuovo server abbia effettivamente varianza minore di $20\,\mathrm{ms}^2$. Decide quindi di misurare un campione casuale di n nuovi tempi di risposta, in base al quale stabilire se vale la pena di fare l'acquisto oppure no. Si assuma che i tempi di risposta abbiano densità normale.

(a) Aiutate la Yasdam a decidere se acquistare il nuovo server o meno, impostando un opportuno test d'ipotesi al livello α basato sul campione di numerosità n. Si indichino ipotesi nulla, ipotesi alternativa, statistica test e regione critica. *[*$H_0 : \sigma^2 \geq 20 =: \sigma_0^2$ *vs.* $H_1 : \sigma^2 < \sigma_0^2$*; statistica test* $X_0^2 = \frac{(n-1)S_n^2}{\sigma_0^2}$*, dove* S_n^2 *è la varianza campionaria;* $\mathrm{RC}(\alpha) = \{(X_1, \ldots, X_n) : X_0^2 < \chi^2_\alpha(n-1)\}$*.]*

Vengono quindi misurati 51 tempi di risposta del server proposto alla Yasdam, trovando una varianza campionaria di $16.3\,\mathrm{ms}^2$.

(b) Si valuti il p-value dei dati raccolti. *[*$0.15 <$ *p-value* < 0.20*]*

(c) È il caso di procedere all'acquisto? *[No.]*

Esercizio 13. In un campione di 170 misure della temperatura di ebollizione di un certo liquido si è trovata una media campionaria $\bar{x} = 128.2°$ C, con una varianza campionaria $s^2 = 0.126°\,\mathrm{C}^2$. Supponendo che le osservazioni provengano da una popolazione normale con media μ e varianza σ^2 entrambe incognite:

(a) Qual è il livello di significatività minimo che porta a rifiutare l'ipotesi nulla che $\sigma^2 \geq 0.15°\,\mathrm{C}^2$? *[*$6.552\%$*.]*

(b) Al livello di significatività del 5%, i dati mostrano evidenza del fatto che $\sigma^2 < 0.15°\,\mathrm{C}^2$? *[No.]*

(c) Si definisca, si determini e si rappresenti graficamente la funzione di potenza del test al livello di significatività del 5% svolto nel punto precedente. *[*$\pi(\sigma^2) = F_{\chi^2(169)}(\frac{20.948}{\sigma^2})$*, definita quando* σ^2 *sta nell'intervallo* $(0\,,\,0.15)$*. È una funzione decrescente in* σ^2*, con* $\lim_{\sigma^2 \to 0^+} \pi(\sigma^2) = 1$ *e* $\pi(0.15) = 0.05$*.]*

(d) Quanto vale tale potenza in corrispondenza di $\sigma^2 = 0.11°\,\mathrm{C}^2$? *[*$87.698\%$*]*

Esercizio 14. La ditta Baltic Sea produce macchine per l'inscatolamento di caviale. A causa delle fluttuazioni casuali, la quantità di caviale dosata dalle macchine è una variabile aleatoria X gaussiana con media μ e varianza σ^2 entrambe incognite. Prima di procedere all'acquisto di una di queste macchine, controllo le quantità X_i (misurate in grammi) di caviale contenute in un campione casuale di 10 scatolette, ottenendo

$$\sum_{i=1}^{10} x_i = 300.6 \qquad \sum_{i=1}^{10} x_i^2 = 9071.1$$

Per me l'errore più grave consiste nell'acquistare una macchina che abbia un'imprecisione (= deviazione standard di X) effettiva maggiore o uguale di 2 g. Sono disposto a correre il rischio che questo succeda solo con una probabilità massima del 2.5%.

(a) Impostate un opportuno test sulla varianza specificando: ipotesi nulla, ipotesi alternativa e la regione critica del test. *[*$H_0 : \sigma^2 \geq 4$ *vs.* $H_1 : \sigma^2 < 4$*;* $\mathrm{RC}(2.5\%) = \{(X_1, \ldots, X_{10}) : \frac{9S^2}{4} < 2.7004\}$*, dove* S^2 *è la varianza campionaria.]*

(b) Fornite una stima puntuale della varianza σ^2. *[*$\widehat{\sigma^2} = 3.896$*.]*

(c) Quale decisione prendete nel test del punto (a)? *[Non si può rifiutare* H_0*.]*

(d) Fornite un intervallo entro cui cade la potenza del test del punto (a) quando $\sigma^2 = 1.5$. *[*$0.25 < \pi(1.5) < 0.50$*]*

Esercizio 15. I rappresentanti degli studenti del Politecnico propongono al Consiglio di Facoltà che l'1% delle tasse da loro pagate sia destinato a un'Università di un paese del terzo mondo. Per dimostrare che la proposta è condivisa da una gran parte dei loro rappresentati, i rappresentanti presentano al Consiglio i risultati di un'intervista fatta a 472 studenti, di cui ben 384 si sono dichiarati favorevoli. I rappresentanti sono convinti che ciò provi che almeno il 75% di tutti gli studenti è favorevole alla proposta; tuttavia, uno di loro che ha studiato Statistica preferisce effettuare un test d'ipotesi prima di trarre conclusioni affrettate.

(a) Si imposti un test opportuno (motivando la risposta) e lo si esegua al livello di significatività del 10%. *[Test per le ipotesi* $H_0 : p \leq 0.75$ *vs.* $H_1 : p > 0.75$. *Si rifiuta* H_0 *al* 10%*.]*

(b) Si determini il p-value del test e si commenti il risultato ottenuto. *[p-value* $= 0.00071$*.]*

(c) Calcolare la probabilità che il test al 10% del punto (a) induca erroneamente a credere che meno del 75% degli studenti è favorevole alla proposta, quando in realtà la percentuale di favorevoli è dell'80%. *[*0.09176*.]*

Esercizio 16. Su 2350 cittadini intervistati, 1908 sono favorevoli alla costruzione di un nuovo cinema multisala.

(a) Al livello di significatività del 5%, il sondaggio mostra evidenza del fatto che almeno l'80% di tutti i cittadini è favorevole al nuovo cinema? *[No.]*

(b) Calcolare il p-value del test eseguito al punto (a). *[p-value* $= 7.493\%$*.]*

(c) Calcolare la potenza del test al livello del 5% svolto al punto (a) quando la vera percentuale di cittadini favorevoli è l'83%. *[*$\pi(0.83) = 98.300\%$*.]*

Esercizio 17. Una certa pianta produce piselli che possono essere verdi o gialli. Secondo un classico modello genetico, la proporzione p dei piselli verdi deve essere pari a 0.75, mentre secondo un nuovo modello genetico dovrebbe essere pari a 0.8. Si vuole quindi verificare l'ipotesi H_0: $p = 0.75$ contro H_1: $p = 0.8$, esaminando un campione (casuale e numeroso) di n di questi piselli. Indicando con $\bar{X}_n$ la frazione di piselli verdi trovati nel campione, si intende usare il test di regione critica

$$\mathrm{RC} = \left\{ \bar{X}_n > 0.75 + \frac{0.71231}{\sqrt{n}} \right\}.$$

(a) Si determini la probabilità α di commettere un errore di *prima* specie esaminando un campione di $n = 100$ piselli. *[*$\alpha = 5\%$*]*

(b) Si determini la probabilità β di commettere un errore di *seconda* specie esaminando un campione di $n = 100$ piselli. *[*$\beta_{100} = 70.194\%$*]*

(c) Si calcoli quanto deve essere numeroso il campione se si vuole $\beta < 10\%$. *[*$n \geq 600$*]*

Viene esaminato un campione di 999 piselli, trovandone 777 verdi e 222 gialli.

(d) Si dica quale modello genetico risulta più corretto in base al test proposto. *[Il nuovo modello]*

Esercizio 18. Sia $X_1, \dots, X_n$ un campione casuale da una distribuzione di Bernoulli di parametro θ, e indichiamo con $\bar{X}_n$ la sua media campionaria. Si vogliono verificare le ipotesi

$$H_0 : \theta \leq 1/2 \qquad \text{vs.} \qquad H_1 : \theta > 1/2 \,.$$

(a) Sia $n = 2$. Per verificare H_0 contro H_1 viene proposto un test con regione critica

$$\mathrm{RC} = \{\bar{x}_2 \geq 1\} \,.$$

Al variare di θ, si determinino la probabilità di errore di primo tipo e quella di secondo tipo. Quanto vale la massima probabilità di commettere un errore di primo tipo? *[Se $\theta \leq 1/2$, si ha P_θ(Errore I tipo) $= \theta^2$. Se $\theta > 1/2$, si ha P_θ(Errore II tipo) $= 1 - \theta^2$. La massima probabilità di commettere l'errore di primo tipo è $\alpha = 1/4$.]*

(b) Sia ora $n = 1000$. Si determini k in modo tale che

$$\mathrm{RC} = \{\bar{x}_{1000} \geq k\}$$

sia la regione critica di un test di livello $\alpha = 5\%$ per verificare H_0 contro H_1. *[$k = 0.526$]*

Esercizio 19. Un lotto di molte migliaia di pezzi viene ritenuto inaccettabile dalla ditta acquirente se contiene più del 10% di pezzi difettosi. Pertanto, prima di procedere all'acquisto, viene sottoposto a controllo un campione di 100 pezzi. Sia m il numero massimo di pezzi difettosi tra i 100 collaudati per cui si decide comunque di acquistare l'intero lotto. Come si deve scegliere m affinché sia inferiore al 5% la probabilità di accettare un lotto con più del 10% di pezzi difettosi? *[$m = 5$]*

Soluzioni di alcuni esercizi

Soluzione 1.

(a) *Prima* di fare le 60 misure, abbiamo un campione aleatorio $X_1, \ldots, X_{60}$, con

$$X_i = \text{peso degli spaghetti che troveremo nell'}i\text{-esima confezione}\,.$$

Il testo dell'esercizio dice che $X_i \sim N(\mu, \sigma^2)$, dove $\sigma^2 = 200\,\mathrm{g}^2$ è nota, mentre μ è il peso medio incognito degli spaghetti in ciascuna confezione. Gli impianti sono da revisionare se $\mu \neq 500\,\mathrm{g}$; mettiamo pertanto questa affermazione come ipotesi alternativa. Ciò vuol dire che siamo disposti a revisionare gli impianti solo se c'è una forte evidenza del fatto che essi siano fuori norma; in altre parole, di default supponiamo che gli impianti funzionino regolarmente. In conclusione, dobbiamo fare uno Z-test per le ipotesi

$$H_0 : \mu = 500 =: \mu_0 \qquad \text{vs.} \qquad H_1 : \mu \neq \mu_0\,.$$

Dal formulario, la regola di un test di significatità α per tali ipotesi è

$$\text{“rifiuto } H_0 \text{ se } \left|\frac{\bar{X} - \mu_0}{\sigma}\sqrt{n}\right| > z_{1-\frac{\alpha}{2}}\text{”}\,,$$

dove $\bar{X}$ è la media campionaria delle 60 misure. Coi risultati trovati *dopo* le 60 misure,

$$\left|\frac{\bar{x} - \mu_0}{\sigma}\sqrt{n}\right| = \left|\frac{496 - 500}{\sqrt{200}}\sqrt{60}\right| = |-2.191| = 2.191\,.$$

Abbiamo:

- se $\alpha = 1\%$: $z_{1-\frac{\alpha}{2}} = z_{0.995} = 2.58 \Rightarrow$ non si può rifiutare H_0;
- se $\alpha = 5\%$: $z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96 \Rightarrow$ si rifiuta H_0;
- se $\alpha = 10\%$: $z_{1-\frac{\alpha}{2}} = z_{0.95} = 1.645 \Rightarrow$ si rifiuta H_0.

Da notare che in questo test non c'era il rischio di scambiare H_0 con H_1. Infatti, se fosse stato $H_1 : \mu = \mu_0$, non avremmo saputo fare alcun test con questa ipotesi alternativa (non conosciamo nessun test in grado di fornire evidenza forte del fatto che $H_1 : \mu = \mu_0$).

(b) Si richiede di calcolare il p-value del test precedente. Dalla risposta del punto (a), vediamo che $1\% < p\text{-value} < 5\%$. Più precisamente, il p-value è quel valore di α per cui

$$\left|\frac{\bar{x} - \mu_0}{\sigma}\sqrt{n}\right| \equiv z_{1-\frac{\alpha}{2}}\,,$$

da cui, applicando la f.d.r. della $N(0,1)$ a entrambi i membri e ricordandoci il valore di $\left|\frac{\bar{x}-\mu_0}{\sigma}\sqrt{n}\right|$ trovato al punto (a),

$$\Phi\left(\left|\frac{\bar{x} - \mu_0}{\sigma}\sqrt{n}\right|\right) = \Phi(z_{1-\frac{\alpha}{2}}) \quad \Leftrightarrow \quad \Phi(2.191) = 1 - \frac{\alpha}{2}$$
$$\Leftrightarrow \quad \alpha = 2\,[1 - \Phi(2.191)] = 2\,[1 - 0.98574] = 2.852\%\,.$$

Quest'ultimo è il valore cercato: p-value $= 2.852\%$.

(c) Si commette l'errore di seconda specie quando si accetta H_0 quando in realtà questa è falsa. In un test bilatero come quello del punto (a), la probabilità di commettere tale errore è tanto più piccola quanto più il vero valore μ della media è lontano dal valore $\mu_0 = 500$ assegnato, o, più precisamente, quanto più $|\mu - \mu_0|$ è grande. Ciò è perfettamente ragionevole: significa che, a mano a mano che $|\mu - \mu_0|$ aumenta, diventa sempre più difficile che il test faccia (erroneamente!) accettare $H_0 : \mu = \mu_0$. Per dimostrare quanto detto, calcoliamo

$$
\begin{aligned}
\text{prob. errore II specie} &= \mathbb{P}_{H_0 \text{ falsa}}(\text{“accetto } H_0\text{”}) \\
&= \mathbb{P}_\mu\left(\left| \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} \right| \le z_{1-\frac{\alpha}{2}} \right) \\
&= \mathbb{P}_\mu\left(-z_{1-\frac{\alpha}{2}} \le \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} \le z_{1-\frac{\alpha}{2}} \right) \qquad \text{qui } \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \textit{ non} \text{ è ancora standardizzata} \\
&= \mathbb{P}_\mu\left(-z_{1-\frac{\alpha}{2}} \le \frac{\bar{X} - \mu}{\sigma}\sqrt{n} + \frac{\mu - \mu_0}{\sigma}\sqrt{n} \le z_{1-\frac{\alpha}{2}} \right) \\
&= \mathbb{P}_\mu\left(-z_{1-\frac{\alpha}{2}} - \frac{\mu - \mu_0}{\sigma}\sqrt{n} \le \underbrace{\frac{\bar{X} - \mu}{\sigma}\sqrt{n}}_{N(0,1)} \le z_{1-\frac{\alpha}{2}} - \frac{\mu - \mu_0}{\sigma}\sqrt{n} \right) \qquad \textit{ora } \bar{X} \text{ è standardizzata} \\
&= \Phi\left(z_{1-\frac{\alpha}{2}} - \frac{\mu - \mu_0}{\sigma}\sqrt{n} \right) - \Phi\left(-z_{1-\frac{\alpha}{2}} - \frac{\mu - \mu_0}{\sigma}\sqrt{n} \right) \\
&= \Phi\left(z_{1-\frac{\alpha}{2}} - \frac{\mu - \mu_0}{\sigma}\sqrt{n} \right) - \left[1 - \Phi\left(z_{1-\frac{\alpha}{2}} + \frac{\mu - \mu_0}{\sigma}\sqrt{n} \right) \right] \\
&= \Phi\left(z_{1-\frac{\alpha}{2}} - \delta \right) + \Phi\left(z_{1-\frac{\alpha}{2}} + \delta \right) - 1
\end{aligned}
$$

dove abbiamo posto $\delta = \frac{\mu - \mu_0}{\sigma}\sqrt{n}$. Da notare che l'espressione precendente dà lo stesso valore se sostituiamo δ con $-\delta$; questo vuol dire che essa dipende solo dal modulo $|\delta| = |\mu - \mu_0| \frac{\sqrt{n}}{\sigma}$, cioè da $|\mu - \mu_0|$. Più precisamente,

$$\text{prob. errore II specie} = \Phi\left(z_{1-\frac{\alpha}{2}} - |\delta| \right) + \Phi\left(z_{1-\frac{\alpha}{2}} + |\delta| \right) - 1$$

(per convincervene, verificate l'uguaglianza separatamente nei due casi con $\delta > 0$ e $\delta < 0$). Dobbiamo ancora verificare che questa espressione è crescente in $|\delta|$. Per farlo, prendendiamone la derivata

$$\frac{\mathrm{d}}{\mathrm{d}|\delta|}(\text{prob. errore II specie}) = -f\left(z_{1-\frac{\alpha}{2}} - |\delta| \right) + f\left(z_{1-\frac{\alpha}{2}} + |\delta| \right) ,$$

dove $f(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}$ è la densità normale standard. Per ogni $\alpha \in (0,1)$, abbiamo sempre $z_{1-\frac{\alpha}{2}} > z_{0.5} = 0$, e quindi (vedere sul grafico)

$$f\left(z_{1-\frac{\alpha}{2}} + |\delta| \right) < f\left(z_{1-\frac{\alpha}{2}} - |\delta| \right) .$$

Questo dimostra che $\frac{\mathrm{d}}{\mathrm{d}|\delta|}$(prob. errore II specie) è sempre negativa, e quindi la prob. errore II specie è decrescente in $|\delta|$. Ciò è esattamente quanto affermato all'inizio: la probabilità di errore di II specie decresce se aumenta $|\mu - \mu_0|$.

Per rispondere al punto (c), si tratta quindi di vedere quale dei $\mu \in \{490, 493.5, 496, 497.1\}$ rende minimo $|\mu - 500|$ (l'esercizio chiede il μ che rende *massima* la probabilità di errore di II specie $\Rightarrow$ $|\mu - \mu_0|$ deve essere *minimo*). Chiaramente, la risposta è $\mu = 497.1$.

(d) In questo caso, i consumatori possono accusare il produttore di pasta solo se c'è una forte evidenza del fatto che le sue confezioni hanno un contenuto medio di spaghetti minore di quanto da lui dichiarato, cioè solo se $\mu < 500\,\mathrm{g}$; questa dev'essere quindi la loro ipotesi alternativa. Invece, l'ipotesi di default è che il produttore sia onesto, e di conseguenza $\mu = 500\,\mathrm{g}$ dev'essere l'ipotesi nulla. Ricapitolando, i consumatori devono fare un test per le ipotesi

$$H_0 : \mu = 500\,\mathrm{g} =: \mu_0 \qquad \text{vs.} \qquad H_1 : \mu < \mu_0 \,.$$

In questo modo, l'errore più grave, cioè quello di I specie, è: *rifiutare H_0 (cioè citare il produttore di pasta in tribunale) quando in realtà H_0 è vera (cioè il produttore di pasta in realtà è onesto).*
Essendo sempre nota la varianza, si può usare uno Z-test, la cui regola di rifiuto diventa in questo caso:

$$\text{“rifiuto } H_0 \text{ se } \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} < -z_{1-\alpha}\text{”} \,.$$

Con gli stessi dati di prima, abbiamo ancora

$$\frac{\bar{x} - \mu_0}{\sigma}\sqrt{n} = -2.191$$

(però senza prendere il modulo, questa volta). La soglia della regione critica al livello $\alpha = 1\%$ risulta invece

$$-z_{1-\alpha} = -z_{1-0.01} = -z_{0.99} = -2.33 \,.$$

Poiché $-2.191 \not< -2.33$, non possiamo rifiutare H_0, e quindi non abbiamo nessun elemento per accusare il produttore di pasta.

Soluzione 4.

(a) Voler dimostrare che la concentrazione di colesterolo media in un paziente che assume il farmaco Colmeno è inferiore a $200\,\mathrm{mg/dl}$, significa cercare evidenza forte del fatto che $\mu < 200\,\mathrm{mg/dl}$. Di conseguenza, dovremo mettere questa affermazione come ipotesi alternativa. Dopodiché, scegliamo come ipotesi nulla $\mu \geq 200$ (o $\mu = 200$, che dà comunque lo stesso test) *come conseguenza della scelta di H_1*. Ricapitolando, le ipotesi corrette sono

$$H_0 : \mu \geq 200 =: \mu_0 \;\; (\text{o } \mu = \mu_0) \qquad \text{vs.} \qquad H_1 : \mu < \mu_0 \,.$$

Fissate H_0 e H_1, si tratta di fare uno Z-test per un campione normale a varianza nota. Dal formulario, la statistica test è

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$$

e la regola di rifiuto al livello di significatività α è

$$\text{“rifiuto } H_0 \text{ se } Z_0 < -z_{1-\alpha}\text{”} \,.$$

(b) La probabilità di errore di primo tipo è il livello di significatività del test; con la regione di rifiuto precedente, abbiamo appena visto che questa è α. Di conseguenza, deve essere $\alpha = 5\% = 0.05$. La potenza invece è

$$
\begin{aligned}
\pi &= \mathbb{P}_{H_0 \text{ falsa}}(\text{“rifiuto } H_0\text{”}) \\
&= \mathbb{P}_\mu(Z_0 < -z_{1-\alpha}) \qquad \text{con } \mu < \mu_0 \text{ come in } H_1 \\
&= \mathbb{P}_\mu\left(\frac{\bar{X}-\mu_0}{\sigma}\sqrt{n} < -z_{1-\alpha}\right) \\
&= \mathbb{P}_\mu\left(\frac{\bar{X}-\mu}{\sigma}\sqrt{n} < -z_{1-\alpha} + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) \\
&= \Phi\left(-z_{1-\alpha} + \frac{\mu_0-\mu}{\sigma}\sqrt{n}\right) \\
&= \Phi\left(-z_{0.95} + \frac{200-190}{\sqrt{50}}\sqrt{n}\right) \\
&= \Phi\left(-1.645 + 1.4142\sqrt{n}\right) .
\end{aligned}
$$

L'esercizio vuole $\pi \geq 90\% = 0.90$. Di conseguenza, deve essere

$$
-1.645 + 1.4142\sqrt{n} > z_{0.90} = 1.28 \quad \Leftrightarrow \quad n > \left(\frac{1.28+1.645}{1.4142}\right)^2 = 4.278 .
$$

Poiché n deve essere intero, servono almeno $n = 5$ misure.

(c) La regola del test al livello $\alpha = 5\%$ è

$$
\text{“rifiuto } H_0 \text{ se } Z_0 < -z_{1-0.05} = -z_{0.95} = -1.645\text{”} .
$$

Dopo aver fatto le misure, abbiamo trovato la realizzazione $\bar{x} = 198.5\,\text{mg/dl}$. Di conseguenza,

$$
z_0 = \frac{\bar{x}-\mu_0}{\sigma}\sqrt{n} = \frac{198.5-200}{\sqrt{50}}\sqrt{10} = -0.671 .
$$

Poiché $-0.671 \not< -1.645$, non possiamo rifiutare H_0. I dati non forniscono quindi alcuna evidenza che il Colmeno sia efficace.

Soluzione 5.

(a) Abbiamo un campione aleatorio $X_1, \ldots, X_8$. Indichiamo con

$$
\mu = \mathbb{E}[X_i] \qquad \qquad \sigma^2 = \operatorname{Var}(X_i)
$$

i parametri media e varianza del campione. Gli *stimatori* naturali (non distorti e consistenti in media quadratica) di tali parametri sono rispettivamente la media campionaria

$$
\bar{X} = \frac{X_1 + \ldots + X_8}{8} ,
$$

e la varianza campionaria

$$
S^2 = \frac{1}{8-1}\left[(X_1-\bar{X})^2 + \ldots + (X_8-\bar{X})^2\right] ,
$$

Le *stime* corrispondenti (= realizzazioni sui risultati *dopo* la misura) sono

$$
\begin{aligned}
\bar{x} &= \frac{2.3 + \ldots + 1.2}{8} = 2.125 \\
s^2 &= \frac{1}{7}\left[(2.3-2.125)^2 + \ldots + (1.2-2.125)^2\right] = 0.325 ,
\end{aligned}
$$

(b) Dobbiamo dimostrare che l'aria da cui provengono i campioni non è dannosa, cioè provare che $\mu < 2.7\,\mu\mathrm{g/m}^3$. Vogliamo evidenza forte di questo fatto; di conseguenza lo mettiamo come ipotesi alternativa. Le ipotesi statistiche corrette sono dunque

$$H_0 : \mu \geq 2.7 =: \mu_0 \quad (\text{o } \mu = \mu_0) \qquad \text{vs.} \qquad H_1 : \mu < \mu_0$$

Con questa scelta, l'errore più grave è l'errore di I specie

rifiutare H_0 (= concludere che l'aria è sana) quando in realtà H_0 è vera (= l'aria è dannosa) ,

che è ragionevole. Poiché l'esercizio ci dice che abbiamo a che fare con un campione normale, ma non ci dà la sua varianza, l'unico test possibile è il T-test. La sua regola di rifiuto al livello α è

$$\text{“rifiuto } H_0 \text{ se } T_0 := \frac{\bar{X} - \mu_0}{S}\sqrt{n} < -t_{1-\alpha}(n-1)\text{”} . \tag{*}$$

Con $\alpha = 5\%$ e $n = 8$, il quantile è

$$-t_{1-\alpha}(n-1) = -t_{0.95}(7) = -1.8946\,,$$

mentre coi risultati delle misure

$$t_0 = \frac{2.125 - 2.7}{\sqrt{0.325}}\sqrt{8} = -2.8528\,.$$

Poiché $-2.853 < -1.8946$, i nostri dati ci costringono a rifiutare H_0 al 5% di significatività.

(c) Con la regola di rifiuto (*), il p-value è il valore di α che risolve l'equazione

$$t_0 \equiv -t_{1-\alpha}(7)\,. \tag{o}$$

Dalle tabelle abbiamo

$$-t_{0.975}(7) = -2.3646 < t_0 = -2.8528 < -2.9979 = -t_{0.99}(7)\,. \tag{oo}$$

Confrontando (o) e (oo), abbiamo

$$0.975 < 1 - \alpha < 0.99 \quad \Rightarrow \quad 0.01 < \alpha < 0.025\,.$$

Questo significa che il p-value è compreso tra l'1% e il 2.5%. Con un p-value così basso, abbiamo forte evidenza del fatto che bisogna rigettare H_0 e accettare H_1; dunque possiamo sentirci tranquilli a respirare l'aria da cui provengono i campioni.

(d) In questo punto, cambia solo il valore di μ_0, che ora è $\mu_0 = 2.4\,\mu g/m^3$, ma il test e i risultati delle misure restano gli stessi del punto precedente. Ricalcolando la statistica test troviamo

$$t_0 = \frac{2.125 - 2.4}{\sqrt{0.325}}\sqrt{8} = -1.3644\,.$$

Ora abbiamo dalle tabelle

$$-t_{0.85}(7) = -1.1192 < t_0 = -1.3644 < -1.4149 = -t_{0.90}(7)\,.$$

e di conseguenza

$$0.85 < 1 - \alpha < 0.90 \quad \Rightarrow \quad 10\% < \alpha = p\text{-value} < 15\%\,.$$

Questo p-value è alto (ben al di sopra dell'usuale soglia del 5%). Non possiamo quindi rigettare H_0; in altre parole, non possiamo affatto escludere che l'aria da cui provengono i campioni sia dannosa.

Soluzione 6.

L'ipotesi di default è l'esperienza degli anni passati, cioè che il numero di guasti medi alla settimana sia rimasto $\mu = 3$. Di conseguenza, questa sarà l'ipotesi nulla. Se vogliamo dimostrare che nell'ultimo anno il parametro μ è aumentato in modo significativo (cioè cerchiamo evidenza forte di questo fatto), dobbiamo scegliere $\mu > 3$ come ipotesi alternativa. Perció, dobbiamo fare un test per le ipotesi

$$H_0 : \mu = 3 =: \mu_0 \qquad \text{vs.} \qquad H_1 : \mu > \mu_0 \,.$$

Ora, il numero totale di guasti nell'ultimo anno è la v.a.

$$Y = X_1 + X_2 + \ldots + X_{52} \,,$$

dove

$$X_i = \text{numero di guasti nell'}i\text{-esima settimana dell'ultimo anno}$$

(in un anno ci sono 52 settimane). Possiamo supporre che le v.a. $X_1, X_2, \ldots, X_{52}$ siano i.i.d., con valore atteso $\mathbb{E}[X_i] = \mu$. Si tratta di un campione numeroso ($n = 52 > 30$), in cui la densità delle X_i senz'altro NON potrà essere gaussiana; infatti, le X_i sono discrete (in un giorno può capitare solo un numero intero di guasti), e tipicamente prenderanno pochi valori (p.es. $\{0, 1, 2, 3, 4, 5, 6\}$). Faremo quindi un test per la media di un campione numeroso qualsiasi a varianza incognita. La regola del test al livello α sarà pertanto

$$\text{“rifiuto } H_0 \text{ se } T_0 := \frac{\bar{X} - \mu_0}{S}\sqrt{n} > z_{1-\alpha}\text{”} \,. \qquad (*)$$

Dal momento che l'esercizio non fissa α, troveremo direttamente il p-value del test e in base a esso decideremo noi se accettare H_0 o meno.

L'esercizio ci dice che, *dopo* aver misurato le X_i, abbiamo trovato

$$y = x_1 + x_2 + \ldots + x_{52} = 234 \quad \Rightarrow \quad \bar{x} = \frac{x_1 + x_2 + \ldots + x_{52}}{52} = \frac{234}{52} = 4.5 \,.$$

e la deviazione standard campionaria

$$s = \sqrt{\frac{1}{52-1}\sum_{i=1}^{52}(x_i - \bar{x})^2} = 1.8 \,.$$

Inserendo questi valori in T_0 troviamo

$$t_0 = \frac{4.5 - 3}{1.8}\sqrt{52} = 6.009 \,.$$

Il p-value dato dalla regola (*) è il valore di α per cui

$$t_0 \equiv z_{1-\alpha} \quad \Rightarrow \quad 6.009 \equiv z_{1-\alpha} \quad \Rightarrow \quad 1 - \alpha \simeq 1 \quad \Rightarrow \quad \alpha \simeq 0 \,.$$

In altre parole, il p-value è del tutto trascurabile. H_0 è pertanto da rigettare a tutti i livelli di significativià sensati e c'è fortissima evidenza che nell'ultimo anno μ sia aumentata.

Soluzione 13.

(a) Test per le ipotesi

$$H_0 : \sigma^2 \geq 0.15 =: \sigma_0^2 \qquad \text{vs.} \qquad H_1 : \sigma^2 < \sigma_0^2 \,.$$

Rifiuto H_0 al livello α se $\frac{(n-1)S_n^2}{\sigma_0^2} < \chi^2_\alpha(n-1)$. Coi dati a disposizione

$$\frac{(n-1)s_n^2}{\sigma_0^2} = 141.96\,,$$

che va confrontata con

$$\chi^2_\alpha(n-1) = \chi^2_\alpha(169) \simeq \frac{(z_\alpha + \sqrt{2\cdot 169 - 1})^2}{2} \qquad \text{perché 169 è grande}\,.$$

Il p-value è il valore di α che soddisfa

$$\frac{(n-1)s_n^2}{\sigma_0^2} \equiv \chi^2_\alpha(n-1) \quad \Leftrightarrow \quad 141.96 = \frac{(z_\alpha + \sqrt{2\cdot 169 - 1})^2}{2}$$
$$\Leftrightarrow \quad z_\alpha = -1.508 \quad \Leftrightarrow \quad \alpha = 1 - 0.93448 = 6.552\%\,.$$

Cioè, p-value $= 6.552\%$. Quindi, il livello di significatività minimo richiesto dalla domanda è 6.552%.

(b) Poiché $5\% < 6.552\% = p$-value, al livello del 5% dobbiamo accettare H_0, cioè non abbiamo evidenza in favore del fatto che $\sigma^2 < 0.15°\,\mathrm{C}^2$.

(c) La potenza è una funzione del vero valore del parametro σ^2, cioè

$$\begin{aligned}\pi(\sigma^2) &= \mathbb{P}_{\sigma^2}(\text{“rifiuto } H_0\text{”}) \\ &= \mathbb{P}_{\sigma^2}\left(\frac{(n-1)S_n^2}{\sigma_0^2} < \chi^2_\alpha(n-1)\right) = \mathbb{P}_{\sigma^2}\left(\frac{(n-1)S_n^2}{\sigma^2} < \frac{\sigma_0^2}{\sigma^2}\,\chi^2_\alpha(n-1)\right) \\ &= F_{\chi^2(n-1)}\left(\frac{\sigma_0^2}{\sigma^2}\,\chi^2_\alpha(n-1)\right) = F_{\chi^2(169)}\left(\frac{0.15}{\sigma^2}\,\chi^2_{0.05}(169)\right)\,.\end{aligned}$$

Nell'espressione precedente, dobbiamo ancora calcolare

$$\chi^2_{0.05}(169) \simeq \frac{(z_{0.05} + \sqrt{2\cdot 169 - 1})^2}{2} = \frac{(-1.645 + \sqrt{2\cdot 169 - 1})^2}{2} = 139.655\,.$$

Quindi, in conclusione,

$$\pi(\sigma^2) = F_{\chi^2(169)}\left(\frac{20.948}{\sigma^2}\right)\,.$$

La funzione potenza è definita per $\sigma^2 \in (0, \sigma_0^2)$, cioè per tutti e soli i valori di σ^2 che soddisfano l'ipotesi alternativa H_1. Per studiare il suo andamento (crescente o decrescente), osserviamo che la funzione di ripartizione $F_{\chi^2(169)}$ è decrescente e continua, con $F_{\chi^2(169)}(0) = 0$ (perché la densità χ^2 è diversa da 0 solo in $[0, +\infty)$) e $\lim_{x\to+\infty} F_{\chi^2(169)}(x) = 1$. Di conseguenza $\pi(\sigma^2)$ è *de*crescente e continua, con

$$\lim_{\sigma^2\to 0^+} \pi(\sigma^2) = \lim_{x\to+\infty} F(x) = 1 \qquad\qquad \pi(0.15) = F_{\chi^2(169)}\left(\chi^2_{0.05}(169)\right) = 0.05 = 5\%\,.$$

Questo è proprio quello che ci aspettavamo: la potenza del test è tanto maggiore quanto più σ^2 è spostata a sinistra di σ_0^2, cioè 'quanto più soddisfa' l'ipotesi alternativa $H_1 : \sigma^2 < \sigma_0^2$. Al contrario, quando σ^2 raggiunge il limite destro $\sigma^2 \equiv \sigma_0^2$, allora la potenza diventa il livello di significatività

$$\pi(\sigma_0^2) = \mathbb{P}_{\sigma^2=\sigma_0^2}(\text{“rifiuto } H_0\text{”}) = \mathbb{P}_{H_0\ \text{vera}}(\text{“rifiuto } H_0\text{”}) = \text{prob. errore I tipo} = 5\%\,.$$

(d) Abbiamo

$$\pi(0.11) = F_{\chi^2(169)}\left(\frac{20.948}{0.11}\right) = F_{\chi^2(169)}(190.436)\,.$$

Per calcolare la funzione di ripartizione $F_{\chi^2(169)}(190.436)$, cerchiamo l'ordine γ del quantile $\chi^2_\gamma(169) = 190.436$. Infatti, una volta trovato γ, avremo

$$F_{\chi^2(169)}(190.436) = F_{\chi^2(169)}(\chi^2_\gamma(169)) = \gamma\,.$$

A questo scopo, usiamo ancora una volta il fatto che, quando i gradi di libertà k sono tanti,

$$\chi^2_\gamma(k) \simeq \frac{(z_\gamma + \sqrt{2k-1})^2}{2} \qquad \Rightarrow \qquad z_\gamma \simeq \pm\sqrt{2\,\chi^2_\gamma(k)} - \sqrt{2k-1}\,.$$

L'ambiguità del segno $\pm$ è dovuta al fatto che non sappiamo se prendere la radice positiva o negativa del quadrato. Tuttavia, osserviamo che le due funzioni $\gamma \mapsto z_\gamma$ e $\gamma \mapsto \chi^2_\gamma(k)$ devono essere entrambe crescenti, perché i quantili aumentano sempre al crescere del loro ordine. Di conseguenza, la soluzione col segno '$-$' va necessariamente scartata, e rimane

$$z_\gamma \simeq \sqrt{2\,\chi^2_\gamma(k)} - \sqrt{2k-1}\,.$$

Nel nostro caso, con $k = 169$ e $\chi^2_\gamma(169) = 190.436$,

$$z_\gamma \simeq \sqrt{2 \cdot 190.436} - \sqrt{2 \cdot 169 - 1} = 1.158 \qquad \Rightarrow \qquad \gamma \simeq 0.87698\,.$$

In conclusione,

$$\pi(0.11) = \gamma \simeq 0.87698 = 87.698\%\,.$$

Soluzione 14.

(a) In un test d'ipotesi, l'errore più grave è sempre quello di prima specie, cioè

rifiutare H_0 quando in realtà H_0 è vera.

Se per me l'errore più grave consiste nello

acquistare una macchina quando in realtà la sua imprecisione è maggiore o uguale di 2 g,

confrontando le due frasi dalla parola 'quando' in poi segue che

H_0 : l'imprecisione della macchina è maggiore o uguale di 2 g, cioè $H_0 : \sigma \geq 2$.

Di conseguenza,

H_1 : l'imprecisione della macchina è minore di 2 g, cioè $H_1 : \sigma < 2$.

Si vede che questa è la scelta corretta dalla serie di equivalenze

acquistare una macchina quando in realtà la sua imprecisione è maggiore o uguale di 2 g
$\Leftrightarrow$ ritenere che l'imprecisione della macchina sia minore di 2 g (e quindi comprarla) quando in realtà l'imprecisione è maggiore o uguale di 2 g
$\Leftrightarrow$ accettare H_1 quando in realtà è vera H_0
$\Leftrightarrow$ rifiutare H_0 quando in realtà H_0 è vera.

Tutto questo è ragionevole: per decidermi ad acquistare una macchina della Baltic Sea, voglio avere evidenza che la sua imprecisione sia piccola (cioè che $\sigma < 2$); pertanto, devo mettere $\sigma < 2$ nell'ipotesi alternativa. Ricapitolando, le ipotesi statistiche corrette sono

$$H_0 : \sigma \geq 2 =: \sigma_0 \qquad \text{vs.} \qquad H_1 : \sigma < \sigma_0\,.$$

Il testo dice che "Sono disposto a correre il rischio [di acquistare una macchina quando in realtà la sua imprecisione è maggiore o uguale di 2 g] solo con una probabilità pari al più al 2.5%". In altre parole, ammetto per l'errore di I specie una probabilità massima del 2.5%. La massima probabilità di errore di I specie è – per definizione! – il livello di sigificatività di un test con ipotesi nulla composta (e tale è la nostro H_0). Quindi, voglio fare un test per le ipotesi precedenti al livello di significatività del 2.5%.
A questo punto, una volta fissate H_0, H_1 e il livello di significatività $\alpha = 2.5\%$, si tratta solo di copiare la regola di rifiuto dal formulario:

$$\text{“rifiuto } H_0 \text{ se } \frac{(n-1)S_n^2}{\sigma_0^2} < \chi^2_\alpha(n-1)\text{ ”} \quad \Leftrightarrow \quad \frac{(10-1)S_n^2}{2^2} < \chi^2_{0.025}(10-1) \quad \Leftrightarrow \quad \frac{9\,S_n^2}{4} < 2.7004\,.$$

(b) Uno stimatore della varianza σ^2 non distorto e consistente in media quadratica è la varianza campionaria

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 = \frac{1}{n-1}\left[\left(\sum_{i=1}^{n} X_i^2\right) - n\,\bar{X}_n^2\right] \qquad \text{(formula alternativa)}\,.$$

Per ottenere dai dati la corrispondente stima di σ^2, è più comodo usare la formula alternativa, calcolando

$$s_n^2 = \frac{1}{10-1}\left[9071.1 - 10\cdot\left(\frac{300.6}{10}\right)^2\right] = 3.896\,.$$

(c) Calcoliamo la realizzazione della statistica test sui dati:

$$\frac{9\,s_n^2}{4} = \frac{9\cdot 3.896}{4} = 8.766\,.$$

Questo valore NON è minore di 2.7004. Di conseguenza, NON posso rifiutare H_0, e concludo che non vale la pena di acquistare la macchina della Baltic Sea.

(d) La potenza del test è la seguente funzione di σ^2, definita solo quando è soddisfatta l'ipotesi alternativa $H_1 : \sigma^2 < 4$:

$$\begin{aligned}
\pi(\sigma^2) &= \mathbb{P}_{\sigma^2}\,(\text{“rifiuto } H_0\text{”}) = \mathbb{P}_{\sigma^2}\left(\frac{(n-1)S_n^2}{\sigma_0^2} < \chi^2_{0.025}(n-1)\right) \\
&= \mathbb{P}_{\sigma^2}\left(\frac{(n-1)S_n^2}{\sigma^2} < \frac{\sigma_0^2}{\sigma^2}\,\chi^2_{0.025}(n-1)\right) = F_{\chi^2(n-1)}\left(\frac{\sigma_0^2}{\sigma^2}\,\chi^2_{0.025}(n-1)\right) \\
&= F_{\chi^2(9)}\left(\frac{4}{1.5}\cdot 2.7004\right) = F_{\chi^2(9)}(7.2011)\,.
\end{aligned}$$

Sulle tavole, vediamo che

$$\chi^2_{0.25}(9) = 5.8988 < 7.2011 < 8.3428 = \chi^2_{0.50}(9)\,,$$

e quindi

$$0.25 < F_{\chi^2(9)}(7.2011) < 0.50\,.$$

Ne concludiamo che

$$25\% < \pi(1.5) < 50\%\,.$$

Soluzione 15.

(a) Per dimostrare che almeno il 75% degli studenti è favorevole alla proposta, bisogna fare un test che abbia questa affermazione come ipotesi alternativa (= cercare evidenza forte in favore di questa affermazione). Di conseguenza, se p è la vera probabilità che uno studente a caso del Politecnico sia favorevole, l'ipotesi alternativa dovrà essere $p > 0.75$. Pertanto, scegliamo

$$H_0 : p \leq 0.75 =: p_0 \quad (\text{o } H_0 : p = p_0) \qquad \text{vs.} \qquad H_1 : p > p_0 \, .$$

Introdotto il campione aleatorio $X_1, \ldots, X_{472}$, con

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esimo dei 472 intervistati è favorevole} \\ 0 & \text{in caso contrario} \end{cases}$$

abbiamo

$$X_i \sim B(1, p) \, ,$$

e quindi dobbiamo fare un test per la frequenza di un campione bernoulliano numeroso. Con le ipotesi statistiche precedenti, la regola di rifiuto al livello $\alpha = 10\%$ è

$$\text{“rifiuto } H_0 \text{ se } Z_0 := \frac{\bar{X} - p_0}{\sqrt{p_0(1-p_0)}} \sqrt{n} > z_{1-\alpha} = z_{0.90} = 1.28\text{”} \, . \qquad (*)$$

La frequenza empirica trovata (= realizzazione di $\bar{X}$ dopo le misure) è

$$\bar{x} = \frac{384}{472} = 0.813559 \, ,$$

a cui corrisponde

$$z_0 := \frac{0.813559 - 0.75}{\sqrt{0.75 \cdot (1 - 0.75)}} \sqrt{472} = 3.189 \, .$$

Essendo $3.189 > 1.28$, dobbiamo rifiutare H_0 al 10% e concludere che effettivamente almeno il 75% degli studenti è favorevole alla proposta.

(b) Con la regola (*), il p-value è l'α che risolve l'equazione

$$\begin{aligned} z_0 \equiv z_{1-\alpha} \quad &\Rightarrow \quad 3.189 = z_{1-\alpha} \quad \Rightarrow \quad \Phi(3.189) = \Phi(z_{1-\alpha}) = 1 - \alpha \quad \Rightarrow \quad 0.99929 = 1 - \alpha \\ &\Rightarrow \quad \alpha = 0.00071 = 0.071\% = p\text{-value} \, . \end{aligned}$$

Con un p-value così piccolo, possiamo tranquillamente rifiutare H_0 a tutti i livelli di significatività ragionevoli, e concludere che almeno il 75% degli studenti è favorevole alla proposta con fortissima evidenza.

(c) Poiché

$$\text{“il test mi fa credere che meno del 75\% degli studenti è favorevole”} = \text{“accetto } H_0\text{”} \, ,$$

la probabilità richiesta è la probabilità di errore di II specie

$$\mathbb{P}_{p=0.80}(\text{“accetto } H_0\text{”}) = \mathbb{P}_{p=0.80}\left(\frac{\bar{X}-p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n} \le z_{0.90}\right)$$

se $p = 0.80$, questa non è la standardizzazione corretta di $\bar{X} \approx N\left(p, \frac{p(1-p)}{n}\right)$

$$= \mathbb{P}_{p=0.80}\left(\bar{X} \le z_{0.90}\,\frac{\sqrt{p_0(1-p_0)}}{\sqrt{n}} + p_0\right)$$

$$= \mathbb{P}_{p=0.80}\left(\frac{\bar{X}-p}{\frac{\sqrt{p(1-p)}}{n}} \le \frac{z_{0.90}\,\frac{\sqrt{p_0(1-p_0)}}{\sqrt{n}} + p_0 - p}{\frac{\sqrt{p(1-p)}}{n}}\right)$$

standardizzazione corretta

$$= \Phi\left(\frac{z_{0.90}\,\frac{\sqrt{p_0(1-p_0)}}{\sqrt{n}} + p_0 - p}{\frac{\sqrt{p(1-p)}}{n}}\right)$$

$$= \Phi\left(z_{0.90}\sqrt{\frac{p_0(1-p_0)}{p(1-p)}} + \frac{p_0-p}{\sqrt{p(1-p)}}\sqrt{n}\right)$$

$$= \Phi\left(1.28\cdot\sqrt{\frac{0.75\cdot(1-0.75)}{0.80\cdot(1-0.80)}} + \frac{0.75-0.80}{\sqrt{0.80\cdot(1-0.80)}}\sqrt{472}\right)$$

$$= \Phi(-1.330) = 1-\Phi(1.330) = 1-0.90824$$

$$= 0.09176 = 9.176\%\,.$$

Soluzione 18.

(a) La probabilità di errore di I tipo è

$$\mathbb{P}_{H_0\text{ vera}}(\text{“rifiuto } H_0\text{”}) = \mathbb{P}_{\theta\le 1/2}\left(\bar{X}\ge 1\right) = \mathbb{P}_{\theta\le 1/2}\left(\frac{X_1+X_2}{2}\ge 1\right) = \mathbb{P}_{\theta\le 1/2}\left(X_1+X_2\ge 2\right)$$

$$= \sum_{k=2}^{2} p(k) = \theta^2\,,$$

dove $p : \{0,1,2\} \to [0,1]$ è densità binomiale $B(2,\theta)$. Tale probabilità dipende dal vero valore di θ, che, essendo H_0 un'ipotesi composta, non risulta univocamente determinato quando H_0 è vera. Si può comunque calcolare la probabilità massima di errore di I specie, e cioè

$$\max_{H_0\text{ vera}} \mathbb{P}(\text{“rifiuto } H_0\text{”}) = \max_{\theta\le 1/2}\theta^2 = 1/4\,.$$

Tale probabilità massima è – per definizione – la *significatività* del test.
La probabilità di errore di II tipo si calcola in modo simile:

$$\mathbb{P}_{H_1\text{ vera}}(\text{“accetto } H_0\text{”}) = \mathbb{P}_{\theta>1/2}\left(\bar{X}<1\right) = \mathbb{P}_{\theta>1/2}\left(\frac{X_1+X_2}{2}<1\right) = \mathbb{P}_{\theta>1/2}\left(X_1+X_2<2\right)$$

$$= \sum_{k=0}^{1} p(k) = 1-p(2) = 1-\theta^2\,.$$

(b) Ora vogliamo determinare k in modo che

$$\text{significatività} = \max_{H_0 \text{ vera}} \mathbb{P}_\theta\big(\text{“rifiuto } H_0\text{”}\big) = \max_{\theta \leq 1/2} \mathbb{P}_\theta\big(\overline{X}_{1000} > k\big) \equiv 5\% \,.$$

Come al solito (e come in particolare è già successo al punto precedente), il massimo si raggiunge quando θ è al limite dei valori ammessi in H_0, e cioè quando $\theta = 1/2$. Di conseguenza,

$$\begin{aligned}\text{significatività} &= \mathbb{P}_{\theta=1/2}\big(\overline{X}_{1000} > k\big) = \mathbb{P}_{\theta=1/2}\left(\frac{\overline{X}_{1000} - \mathbb{E}\left[\overline{X}_{1000}\right]}{\sqrt{\operatorname{Var}\left(\overline{X}_{1000}\right)}} > \frac{k - \frac{1}{2}}{\sqrt{\frac{1}{4000}}}\right) \\ &\simeq 1 - \Phi\left((2k-1)\sqrt{1000}\right)\end{aligned}$$

in quanto per il TLC abbiamo

$$\overline{X}_{1000} \approx N\left(\mathbb{E}\left[X_i\right], \frac{\operatorname{Var}\left(X_i\right)}{n}\right) = N\left(\theta, \frac{\theta(1-\theta)}{1000}\right) = N\left(\frac{1}{2}, \frac{1}{4000}\right) \qquad \text{se} \quad \theta = \frac{1}{2} \,.$$

Ne ricaviamo

$$\begin{aligned}5\% \equiv 1 - \Phi\left((2k-1)\sqrt{1000}\right) \quad &\Rightarrow \quad \Phi\left((2k-1)\sqrt{1000}\right) = 0.95 \\ &\Rightarrow \quad (2k-1)\sqrt{1000} = z_{0.95} = 1.645 \quad \Rightarrow \quad k = 0.526 \,.\end{aligned}$$

Soluzione 19.
Sia

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esimo lotto è difettoso} \\ 0 & \text{altrimenti} \,. \end{cases}$$

Abbiamo un campione di 100 misure $X_1, \ldots, X_{100}$, con $X_i \sim B(1,p)$, in cui p è la probabilità che un lotto a caso sia difettoso. Naturalmente, le ipotesi da mettere a confronto sono

$$p > 10\% \quad \Leftrightarrow \quad \text{l'intero lotto (e non solo i 100 pezzi collaudati) è INaccettabile}$$

contro

$$p < 10\% \quad \Leftrightarrow \quad \text{l'intero lotto è accettabile} \,.$$

Il punto è che non sappiamo quale delle due scegliere come H_0 e quale come H_1. L'esercizio inoltre ci dice che, se il numero di pezzi difettosi tra i 100 esaminati – cioè la v.a.

$$S_{100} = X_1 + X_2 + \ldots + X_{100}$$

– NON supera la soglia di m pezzi, allora decidiamo di acquistare l'intero lotto, cioè lo riteniamo accettabile. In altre parole,

$$\text{“}S_{100} \leq m\text{”} = \text{“riteniamo il lotto accettabile”} \,.$$

Ora l'esercizio vuole che la *massima probabilità* di accettare un lotto con più del 10% di pezzi difettosi sia inferiore al 5%, cioè vuole che

$$5\% \geq \max_{p > 10\%} \mathbb{P}(\text{“riteniamo il lotto accettabile”}) = \max_{p > 10\%} \mathbb{P}(S_{100} \leq m) \,.$$

L'ultima espressione è a tutti gli effetti la massima probabilità di errore di I specie (= la *significatività*) di un test avente come ipotesi nulla $H_0 : p > 10\%$ e come regola di rifiuto

$$\text{“rifiuto } H_0 \text{ se } S_{100} \leq m\text{”} .$$

Infatti, con tale scelta di H_0 e della regola di rifiuto, si ha

$$\text{massima probabilità di errore di I specie} = \max_{H_0 \text{ vera}} \mathbb{P}(\text{“rifiuto } H_0\text{”}) = \max_{p>10\%} \mathbb{P}(S_{100} \leq m) .$$

Da qui in poi, si tratta dunque solo di trovare m che renda tale probabilità massima pari al 5%. Abbiamo

$$\mathbb{P}(S_{100} \leq m) = \mathbb{P}\left(\underbrace{\frac{S_{100} - \mathbb{E}\left[S_{100}\right]}{\sqrt{\mathrm{Var}\left(S_{100}\right)}}}_{\approx N(0,1)} \leq \frac{m - 100 \cdot p}{\sqrt{100 \cdot p \cdot (1-p)}} \right) = \Phi\left(\frac{m - 100 \cdot p}{\sqrt{100 \cdot p \cdot (1-p)}} \right)$$

e, come sempre succede quando si calcola il livello di significatività di un test con ipotesi nulla composta, il massimo si ottiene per $p = p_0 = 10\%$. Dunque vogliamo che

$$\Phi\left(\frac{m - 100 \cdot 0.10}{\sqrt{100 \cdot 0.10 \cdot (1 - 0.10)}} \right) \equiv 5\% \quad \Leftrightarrow \quad \frac{m - 100 \cdot 0.10}{\sqrt{100 \cdot 0.10 \cdot (1 - 0.10)}} = z_{0.05} = -z_{0.95} = -1.645 .$$

La soluzione è $m = 5.065$. Siccome la regola del test fa rifiutare H_0 quando $S_{100} \leq m$, se come valore soglia vogliamo dare un numero intero, significa che dobbiamo scegliere $m = 5$ ($m = 6$ già non andrebbe bene).