Statistica II - Esercitazione 6

La verifica di ipotesi secondo Neyman e Pearson

Neyman e Pearson sono considerati in letteratura quali artefici della formalizzazione del problema inferenziale della verifica di ipotesi, avendo per primi affiancato all'ipotesi nulla, già prevista dai test di significatività, l'ipotesi alternativa, ossia un'ipotesi ad essa contrapposta.
Nel prosieguo, dopo aver introdotto alcuni concetti preliminari, la teoria dei test secondo Neyman e Pearson viene affrontata mediante la trattazione dell'omonimo Lemma.
Ciò premesso, si supponga che lo *spazio parametrico* Θ sia partizionato in due sottoinsiemi disgiunti Θ_0 e Θ_1, ovvero $\Theta = \Theta_0 \cup \Theta_1$ e $\Theta_0 \cap \Theta_1 = \emptyset$, ove Θ_0 e Θ_1 rappresentano gli insiemi dei valori di θ rispettivamente postulati dall'*ipotesi nulla*

$$\mathrm{H}_0 : \theta \in \Theta_0$$

e dall'*ipotesi alternativa*

$$\mathrm{H}_1 : \theta \in \Theta_1.$$

A tal proposito, si ricordi che un'ipotesi statistica è detta *semplice* se, supposta vera, è tale da identificare una e una sola legge di distribuzione della variabile casuale X di partenza rappresentante la popolazione statistica di riferimento. Un'ipotesi statistica è detta invece *composta* in ogni altro caso.
Si voglia fornire un criterio che, sulla base delle osservazioni campionarie $(x_1, \ldots, x_n)$, consenta di scegliere una fra le due suddette ipotesi, consentendo dunque di concludere a favore dell'appartenenza di θ al sottoinsieme Θ_0 o al sottoinsieme Θ_1 di Θ. Inoltre, sia Θ lo *spazio campionario*, ossia l'insieme di tutte le possibili n-uple di osservazioni campionarie ottenibili conducendo n prove indipendenti nella variabile casuale X di partenza. Si noti che un test determina una partizione di Ω in due sottoinsiemi C e $\overline{C}$, tali che $\Omega = C \cup \overline{C}$ e $C \cap \overline{C} = \emptyset$, dove C e $\overline{C}$ rappresentano rispettivamente le regioni di *rifiuto* e *non rifiuto* dell'ipotesi nulla (la regione di rifiuto è detta *anche regione critica*).
Ovviamente un test di ipotesi non può essere infallibile. La decisione nei confronti del sistema di ipotesi statistiche di interesse può infatti essere affetta da due differenti tipologie di errore. In primis, si supponga che il reale valore di θ appartenga a Θ_0. Può accadere che l'n-upla campionaria osservata appartenga a C e dunque si decida a favore di H_1. L'errore di rifiutare l'ipotesi nulla quando in realtà essa è vera è cosiddetto *errore di I tipo* (o di prima specie). Viceversa, si supponga che il reale valore di θ appartenga a Θ_1. Può accadere che l'n-upla campionaria osservata appartenga a $\overline{C}$ e dunque si decida a favore di H_0. L'errore di non rifiutare l'ipotesi nulla quando in realtà essa è falsa è cosiddetto *errore di II tipo* (o di II specie).
Le probabilità con cui vengono commessi gli errori di I e di II tipo sono rispettivamente indicate con α e β e, per quanto appena esplicitato, corrispondono a:

$$\alpha = \Pr\{(X_1, \ldots, X_n) \in C | \mathrm{H}_0\}, \qquad \beta = \Pr\{(X_1, \ldots, X_n) \in \overline{C} | \mathrm{H}_1\}$$

Particolare interesse riveste il complemento all'unità della probabilità di commettere l'errore di II tipo, ossia la quantità

$$1 - \beta = \Pr\{(X_1, \ldots, X_n) \in C | \mathrm{H}_1\}$$

che è detta *potenza del test* e rappresenta la probabilità di rifiutare l'ipotesi nulla quando effettivamente essa è falsa.

Esercizio 1

Sia:

$$(8.2,\ 5.3,\ 8.6,\ 10.9)$$

un campione indipendente e identicamente distribuito estratto da una variabile casuale normale $X \sim \mathcal{N}(\mu, \sigma^2)$, con varianza nota $\sigma^2 = 4$. Ai fini della verifica dell'ipotesi nulla $\mathrm{H}_0 : \mu = 10$ contro l'ipotesi alternativa $\mathrm{H}_1 : \mu = 5$, si supponga di considerare il test caratterizzato da regione critica $C = \{(x_1, x_2, x_3, x_4) : \overline{x} < 8\}$.

1. Si verifichi se i dati osservati conducano al rifiuto dell'ipotesi nulla;
2. si calcoli la probabilità di commettere l'errore di prima specie α;
3. si calcoli la potenza $1 - \beta$ del test.

Soluzione

Sia $X \sim \mathcal{N}(\mu, \sigma^2)$ con $\sigma^2 = 4$ nota, $\mu \in \mathbb{R}$ ignota. Siano inoltre $X_1, \ldots, X_n$ i.i.d. a X, con $n = 4$. Trattasi di un problema di verifica di ipotesi per μ.

1. Essendo $(x_1, x_2, x_3, x_4) = (8.2, 5.3, 8.6, 10.9)$, risulta:

   $$\overline{x} = \frac{1}{4} \sum_{i=1}^{4} x_i = \frac{1}{4} \cdot 33 = 8.25 > 8,$$

   pertanto

   $$(x_1, x_2, x_3, x_4) \notin C,$$

   dunque non vi è sufficiente evidenza empirica a supporto del rifiuto di H_0.

2. Si noti che:

   $$X \overset{\mathrm{H}_0}{\sim} \mathcal{N}(10, 4) \quad \Longrightarrow \quad \overline{X} \overset{\mathrm{H}_0}{\sim} \mathcal{N}\left(10, \frac{4}{4} = 1\right).$$

   Pertanto:

   $$\begin{aligned} \alpha &= \Pr\{rifiutare\, \mathrm{H}_0 | \mathrm{H}_0\} = \Pr\{(X_1, \ldots, X_4) \in C | \mathrm{H}_0\} = \\ &= \Pr\{\overline{X} < 8 | \mu = 10\} = \Pr\{\overline{X} - 10 < -2 | \mu = 10\} = \\ &= \Pr\{Z < -2\} = 0.0228 \end{aligned}$$

3. Si noti ora che:

   $$X \overset{\mathrm{H}_1}{\sim} \mathcal{N}(5, 4) \quad \Longrightarrow \quad \overline{X} \overset{\mathrm{H}_1}{\sim} \mathcal{N}\left(5, \frac{4}{4} = 1\right).$$

   Pertanto:

   $$\begin{aligned} 1 - \beta &= \Pr\{rifiutare\, \mathrm{H}_0 | \mathrm{H}_1\} = \Pr\{(X_1, \ldots, X_4) \in C | \mathrm{H}_1\} = \\ &= \Pr\{\overline{X} < 8 | \mu = 5\} = \Pr\{\overline{X} - 5 < 3 | \mu = 5\} = \\ &= \Pr\{Z < 3\} = 0.9987 \end{aligned}$$

Esercizio 2

Un impianto di acquacoltura alleva gli storioni. Vengono spostati dal processo di nascita al processo di crescita intensiva quando la lunghezza degli individui supera i $10\,cm$. Al fine di valutare se siano pronti per essere spostati, un addetto estrae un campione casuale di pesci da una vasca, ottenendo:

$$(8.9,\ 10.1,\ 11.1,\ 12.3).$$

Assumendo che la varianza sia nota, pari a $\sigma^2 = 2$, e che il campione sia estratto da una variabile casuale Normale, si è interessati ad escludere che la lunghezza sia maggiore di $10\,cm$.

1. Specificare il parametro di interesse ed il sistema di ipotesi necessario;
2. Verificare l'ipotesi nulla specificata al punto precedente, ad un livello di significatività di 0.05, identificando la regione critica;
3. Calcolare il p-value del test;
4. Supponendo di avere un'ipotesi alternativa puntuale, $\mathrm{H}_1 : \mu = 12$, calcolare la probabilità di commettere un errore di II specie;
5. Supponendo di estrarre un secondo campione di numerosità pari a 100, e di ottenere la stessa media campionaria del primo campione, ripetere il test svolto al punto 2.
6. Supponendo di avere un'ipotesi alternativa puntuale, $\mathrm{H}_1 : \mu = 12$, calcolare la probabilità di commettere un errore di II specie per il nuovo campione di numerosità pari a 100;

Soluzione

Sia $X \sim \mathcal{N}(\mu, \sigma^2)$ con $\mu \in \mathbb{R}$ e $\sigma^2 = 2$. Siano inoltre $X_1, \ldots, X_n$ i.i.d. a X, con n = 4. Trattasi di un problema di verifica di ipotesi.

1. Il parametro d'interesse è il valor medio nella popolazione, ovvero $\mathbb{E}[X] = \mu$. Volendo *"escludere che la lunghezza sia maggiore di* $10\,cm$*"*, il sistema di ipotesi da sottoporre a verifica, avente ipotesi alternativa unilaterale, è definito dall'ipotesi nulla $\mathrm{H}_0 : \mu = 10$ contro l'ipotesi alternativa $\mathrm{H}_1 : \mu > 10$.

2. Posto che $X \sim \mathcal{N}(\mu, \sigma^2)$, con σ^2 nota, è possibile ricondursi ad un test di verifica di ipotesi per il valore atteso di una variabile casuale Normale, con varianza nota. Assumendo vera l'ipotesi nulla si ha che:

$$X \overset{\mathrm{H}_0}{\sim} \mathcal{N}(10, 2) \qquad \Longrightarrow \qquad \overline{X} \overset{\mathrm{H}_0}{\sim} \mathcal{N}\left(10, \frac{2}{4}\right)$$

Quindi è possibile ricondursi alla statistica test:

$$Z = \frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \overset{\mathrm{H}_0}{\sim} \mathcal{N}(0, 1).$$

Sia $\underline{X}$ un vettore aleatorio, Z la variabile aleatoria descritta dalla statistica test e $z_{1-\alpha}$ il quantile di di ordine $1 - \alpha$ di tale distribuzione. La regione critica è l'insieme di tutti i valori che portano ad un rifiuto dell'ipotesi nulla H_0, ovvero, essendo l'ipotesi alternativa unilaterale:

$$C = \{\underline{X} : Z > z_{1-\alpha}\} = \{\underline{X} : Z > 1.64\}$$

essendo $z_{1-\alpha} = z_{0.95} = 1.64$.

La media campionaria, per il campione osservato, risulta essere pari a $\overline{x} = 10.6$, quindi la statistica test osservata:

$$z_{oss} = \frac{\overline{x} - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{10.6 - 10}{\sqrt{\frac{1}{2}}} = 0.85$$

ed essendo $0.85 < 1.64 \Rightarrow z_{oss} \in \overline{C}$, il valore osservato della statistica test non cade nella regione critica, quindi non viene rifiutata l'ipotesi nulla. Analogamente è possibile definire la regione critica basata sul supporto della variabile casuale X:

$$C' = \left\{ \underline{X} : \overline{X} > 1.64\sqrt{\frac{\sigma^2}{n}} + \mu \right\} = \left\{ \underline{X} : \overline{X} > 11.16 \right\}$$

Ed essendo che $\overline{x} = 10.6 < 11.16 \Rightarrow \overline{x} \in \overline{C}'$ ed analogamente a quanto concluso precedentemente, l'ipotesi nulla non viene rifiutata.
In conclusione, stando al test effettuato, gli storioni non vengono spostati dal processo di nascita.

3. Il p-value del test è pari a:

$$\begin{aligned}\alpha_{oss} &= \Pr(Z > z_{oss}|\mathrm{H}_0) = \Pr(Z > 0.85|\mathrm{H}_0) = \\ &= 1 - \Pr(Z \leq 0.85|\mathrm{H}_0) = 1 - \Phi(0.85) = \\ &= 1 - 0.8 = 0.2\end{aligned}$$

4. Supponendo di avere un'ipotesi alternativa puntuale, la probabilità di commettere un errore di II specie risulta essere:

$$\begin{aligned}\beta &= \Pr(\text{non rifiuto } \mathrm{H}_0|\mathrm{H}_1) = \Pr(\overline{X} \in \overline{C}|\mathrm{H}_1) = \\ &= \Pr(\overline{X} \leq 11.16|\mathrm{H}_1) = \Pr\left(Z \leq \frac{11.16 - \mu}{\sqrt{\frac{\sigma^2}{n}}} \middle| \mathrm{H}_1 \right) = \\ &= \Pr\left(Z \leq \frac{11.16 - 12}{\sqrt{\frac{2}{4}}} \middle| \mathrm{H}_1 \right) = \Phi(-1.19) = 0.12\end{aligned}$$

5. Avendo estratto un secondo campione di numerosità pari a 100, e con ugual media campionaria, la regione critica identificata nella prima parte del punto 2, corrispondente alla statistica test Z, non cambia, essendo costruita partendo da una variabile casuale Normale standard. Sia quindi $C = \{\underline{X} : Z > 1.64\}$ la regione critica. Per il nuovo campione la statistica test osservata è pari a:

$$z_{oss} = \frac{\overline{x} - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{10.6 - 10}{\frac{1}{\sqrt{50}}} = 4.24$$

ed essendo $4.24 > 1.64 \Rightarrow z_{oss} \in C$, l'ipotesi nulla viene rifiutata, esistono evidenze empiriche contro l'ipotesi che la media sia uguale a 10.

6. Supponendo di avere un'ipotesi alternativa puntuale, ed una numerosità campionaria pari a $n = 100$, per calcolare la probabilità di commettere un errore di II specie è necessaria la soglia della regione critica, definita sul supporto della variabile casuale X, ovvero:

$$x_{1-\alpha} = z_{1-\alpha}\sqrt{\frac{2}{100}} + \mu|\mathrm{H}_0 = 1.64\sqrt{\frac{2}{100}} + 10 = 10.23,$$

ed usando questo valore è possibile calcolare:

$$\begin{aligned}\beta &= \Pr(\text{non rifiuto } \mathrm{H}_0|\mathrm{H}_1) = \Pr(\overline{X} \in \overline{C}|\mathrm{H}_1) = \\ &= \Pr(\overline{X} \leq 10.23|\mathrm{H}_1) = \Pr\left(Z \leq \frac{10.23 - \mu}{\sqrt{\frac{\sigma^2}{n}}} \middle| \mathrm{H}_1 \right) = \\ &= \Pr\left(Z \leq \frac{10.23 - 12}{\sqrt{\frac{2}{100}}} \middle| \mathrm{H}_1 \right) = \Phi(-12.51) \simeq 0\end{aligned}$$

Dato un sistema di ipotesi, al crescere della numerosità campionaria, diminuisce la probabilità di commettere un errore di II specie e, quindi, aumenta la potenza del test.

Esercizio 3

L'azienda BirraMi produce birra artigianale meneghina. Durante la produzione di birra, un aspetto rilevante al fine di ottenere un buon prodotto, è la densità del mosto. Sia 1000 la densità dell'acqua, espressa in g/l. Il mosto prodotto dall'azienda BirraMi ha una densità iniziale di 1040 g/l. Il prodotto, per rispettare gli standard imposti dalla produzione, deve avere una densità finale pari a 1010 g/l. Un addetto estrae un campione indipendente dal fermentatore, composto da 4 osservazioni, prelevate a diverse profondità:

$$(1008,\ 1011,\ 1014,\ 1015).$$

Assumendo che il campione sia estratto da una variabile casuale Normale, con varianza nota pari a $\sigma^2 = 10$, si è interessati a testare se il mosto è pronto per essere imbottigliato, rispettando gli standard di produzione.

1. Specificare il parametro di interesse ed il sistema di ipotesi necessario;
2. Calcolare il p-value del test e prendere una decisione in merito all'ipotesi specificata nel punto 1, ad un livello di significatività $\alpha = 0.05$;
3. Assumendo un'ipotesi alternativa puntuale $\mathrm{H}_1 : \mu = 1005$, calcolare la potenza del test.

Soluzione

Sia $X \sim \mathcal{N}(\mu, \sigma^2)$ con $\mu \in \mathbb{R}$ e $\sigma^2 = 10$. Siano inoltre $X_1, \ldots, X_n$ i.i.d. a X, con $n = 4$. Trattasi di un problema di verifica di ipotesi.

1. Sia X la variabile casuale che descrive la densità del mosto. Il parametro di interesse, che si vuole testare, è il valore atteso di tale distribuzione, $\mathbb{E}[X] = \mu$. Essendo interessati a *"testare se il mosto è pronto per essere imbottigliato, rispettando gli standard di produzione"*, il sistema di ipotesi da sottoporre a verifica, avente ipotesi alternativa bilaterale, è definito dall'ipotesi nulla $\mathrm{H}_0 : \mu = 1010$ contro l'ipotesi alternativa $\mathrm{H}_1 : \mu \neq 1010$.

2. Posto che $X \sim \mathcal{N}(\mu, \sigma^2)$, con σ^2 nota, è possibile ricondursi ad un test di verifica di ipotesi per il valore atteso di una variabile casuale Normale, con varianza nota. Assumendo vera l'ipotesi nulla si ha che:

$$X \overset{\mathrm{H}_0}{\sim} \mathcal{N}(1010, 10) \qquad \Longrightarrow \qquad \overline{X} \overset{\mathrm{H}_0}{\sim} \mathcal{N}\left(1010, \frac{10}{4}\right).$$

Quindi è possibile ricondursi alla statistica test:

$$Z = \frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \overset{\mathrm{H}_0}{\sim} \mathcal{N}(0,1),$$

dove la statistica test osservata, assunta H_0 vera, è pari a:

$$z_{oss} = \frac{\overline{x} - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{1012 - 1010}{\sqrt{\frac{10}{4}}} = 1.27.$$

Dato il valore calcolato per z_{oss}, è possibile calcolare il p-value:

$$\begin{aligned}
\alpha_{oss} &= \Pr(|Z| > |z_{oss}|\,|\mathrm{H}_0) = \Pr(\{Z > |z_{oss}|\} \cup \{Z < -|z_{oss}|\}\,|\mathrm{H}_0) = \\
&= \Pr(Z > |z_{oss}|\,|\mathrm{H}_0) + \Pr(Z < -|z_{oss}|\,|\mathrm{H}_0) = \Pr(Z > 1.27\,|\mathrm{H}_0) + \Pr(Z < -1.27\,|\mathrm{H}_0) \\
&= 2\Pr(Z > 1.27\,|\mathrm{H}_0) = 2\,[1 - \Pr(Z \leq 1.27\,|\mathrm{H}_0)] = \\
&= 2\,[1 - \Phi(1.27)] = 0.20.
\end{aligned}$$

Essendo il p-value maggiore del livello di significatività, $\alpha_{oss} = 0.20 > 0.05 = \alpha$, allora l'ipotesi nulla non viene rifiutata, non sono state riscontrate sufficienti evidenze empiriche contro il fatto che la media della densità sia pari a 1010.

3. Il complemento della regione critica, identificata dal test effettuato al punto precedente, in accordo con il supporto della variabile aleatoria X, è l'insieme:

$$\begin{aligned}
\overline{C} &= \left\{ \underline{X} : \left| \frac{\overline{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}} \right| \leq z_{1-\frac{\alpha}{2}} \middle| \mathrm{H}_0 \right\} = \left\{ \underline{X} : z_{\frac{\alpha}{2}} \leq \frac{\overline{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}} \leq z_{1-\frac{\alpha}{2}} \middle| \mathrm{H}_0 \right\} = \\
&= \left\{ \underline{X} : z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma^2}{n}} + \mu \leq \overline{X} \leq z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma^2}{n}} + \mu \middle| \mathrm{H}_0 \right\} = \\
&= \left\{ \underline{X} : -1.96\sqrt{\frac{10}{4}} + 1010 \leq \overline{X} \leq 1.96\sqrt{\frac{10}{4}} + 1010 \middle| \mathrm{H}_0 \right\} = \\
&= \left\{ \underline{X} : 1006.9 \leq \overline{X} \leq 1013.1 \right\}
\end{aligned}$$

Supponendo di avere un'ipotesi alternativa puntuale $\mathrm{H}_1 : \mu = 1005$, la probabilità di commettere un errore di II specie risulta essere:

$$\begin{aligned}
\beta &= \Pr(\text{non rifiuto } \mathrm{H}_0 | \mathrm{H}_1) = \Pr(\overline{X} \in \overline{C} | \mathrm{H}_1) = \\
&= \Pr(1006.9 \leq \overline{X} \leq 1013.1 | \mathrm{H}_1) = \Pr\left(\frac{1006.9-\mu}{\sqrt{\frac{\sigma^2}{n}}} \leq Z \leq \frac{1013.1-\mu}{\sqrt{\frac{\sigma^2}{n}}} \middle| \mathrm{H}_1 \right) = \\
&= \Pr\left(\frac{1006.9-1005}{\sqrt{\frac{10}{2}}} \leq Z \leq \frac{1013.1-1005}{\sqrt{\frac{10}{4}}} \middle| \mathrm{H}_1 \right) = \Pr(1.20 \leq Z \leq 5.12) = \\
&= \Pr(Z \leq 5.12) - \Pr(Z \leq 1.20) = \Phi(5.12) - \Phi(1.20) = 0.12
\end{aligned}$$

La potenza del test è definita come $1-\beta$, dove β è l'errore di II specie, e risulta essere pari a $1-\beta = 1-0.12 = 0.88$.