# Statistica II - Esercitazione 5

# Stima intervallare

L'affidabilità di un metodo di stima puntuale risiede esclusivamente nella garanzia probabilistica offerta dalle proprietà teoriche del corrispondente stimatore. Stante, però, la possibilità di disporre dei soli valori campionari effettivamente osservati, tali proprietà , essendo basate sull'ipotesi di campionamento ripetuto, non consentono di conoscere quanto bene la stima prodotta approssimi il reale e ignoto valore del parametro $\theta \in \Theta$ di interesse. Qualunque sia il metodo di stima impiegato, pertanto, le operazioni di individuazione di un singolo valore in $\Theta$ tramite sintesi dei risultati campionari e di sostituzione di esso al reale e ignoto valore di $\theta$ non possono essere accompagnate da alcuna quantificazione dell'errore inevitabilmente commesso. Quanto esplicitato induce a preferire, ai fini della stima di $\theta$, la determinazione, anziché di un solo valore, di un insieme di valori ragionevolmente sostituibili a $\theta$ e plausibili per esso; l'insieme in questione, spesso rappresentato da un intervallo, è detto identificare una stima intervallare di $\theta$.
Siano dunque $L(x_1, \dots, x_n)$ e $U(x_1, \dots, x_n)$ due funzioni delle $n$ osservazioni campionarie per le quali valga che $L(x_1, \dots, x_n) \leq U(x_1, \dots, x_n)$ qualunque sia l'$n-$upla campionaria osservata e tali che l'insieme

$$[L(x_1, \dots, x_n), U(x_1, \dots, x_n)]$$

possa contenere il reale e ignoto valore di $\theta$. Allora, l'intervallo casuale $[L, U] = [L(X_1, \dots, X_n), U(X_1, \dots, X_n)]$ è detto stimatore intervallare di $\theta$, mentre l'intervallo numerico $[l, u] = [L(x_1, \dots, x_n), U(x_1, \dots, x_n)]$ è detto stima intervallare di $\theta$.

**Intervallo di confidenza di livello** $1-\alpha$. *L'intervallo* $[l, u] = [L(x_1, \dots, x_n), U(x_1, \dots, x_n)]$ *è detto intervallo di confidenza per* $\theta$ *di livello* $1-\alpha$ *se*:

$$\Pr\left\{\theta \in [L(X_1, \dots, X_n), U(X_1, \dots, X_n)]\right\} = 1-\alpha, \quad \forall \theta \in \Theta.$$

Si noti che l'oggetto di natura aleatoria coinvolto nell'affermazione probabilistica appena proposta è l'intervallo casuale $[L(X_1, \dots, X_n), U(X_1, \dots, X_n)]$; pertanto, essa deve essere letta come segue:

"la probabilità che l'intervallo casuale $[L(X_1, \dots, X_n), U(X_1, \dots, X_n)]$ contenga il vero valore di $\theta$ è pari a $1-\alpha$".

Inoltre, l'affermazione probabilistica in questione si riferisce all'intervallo casuale $[L(X_1, \dots, X_n), U(X_1, \dots, X_n)]$, non all'intervallo osservato $[l, u]$; pertanto, una volta osservato il campione $n-$dimensionale $(x_1, \dots, x_n)$ e determinato l'intervallo $[l, u]$, non è corretto affermare che "l'intervallo $[l, u]$ contiene il vero valore di $\theta$ con probabilità pari a $1-\alpha$"; essendo un intervallo completamente determinato, o esso contiene il vero valore del parametro di interesse o non lo contiene. Nonostante ciò, è legittimo affermare che l'intervallo $[l, u]$ è stato costruito sulla base di una procedura che, applicata ad ipotetiche replicazioni dell'esperimento casuale di campionamento, $(x'_1, \dots, x'_n)$, fornisce intervalli contenenti il vero valore di $\theta$ nell'$(1-\alpha)\%$ dei casi. A differenza di una qualsivoglia stima puntuale di $\theta$, dunque, una stima intervallare è sempre accompagnabile da un numero che rappresenta la probabilità con cui il corrispondente stimatore intervallare contenga il reale e ignoto valore di $\theta$ e che ne identifica la garanzia di affidabilità .

Ciò premesso, il principale metodo di costruzione di intervalli di confidenza è basato sull'inversione analitica di particolari variabili casuali dette "quantità pivotali".

**Quantità pivotale**. *Una funzione* $Q(X_1, \dots, X_n, \theta)$ *del campione casuale* $(X_1, \dots, X_n)$ *e del parametro* $\theta$ *è detta quantità pivotale per* $\theta$ *se avente legge di distribuzione non dipendente da* $\theta$, *ossia identica* $\forall \theta \in \Theta$.

Ad esempio, si supponga che $X_1, \dots, X_n$ siano i.i.d. a $X \sim N(\theta, 1)$, con $\theta \in \mathrm{R}$ ignota; allora:

$$Q(X_1, \dots, X_n, \theta) = \frac{\bar{X} - \theta}{\frac{1}{\sqrt{n}}} = \sqrt{n}\left(\bar{X} - \theta\right) \sim N(0, 1)$$

è quantità pivotale per $\theta$. Pertanto, la costruzione di un intervallo di confidenza di livello $1-\alpha$ per $\theta$ può essere realizzata tramite la considerazione dell'affermazione probabilistica:

$$1-\alpha = \Pr\left(z_{\frac{\alpha}{2}} \leq \sqrt{n}\left(\bar{X} - \theta\right) \leq z_{1-\frac{\alpha}{2}}\right)$$

e la conseguente inversione analitica della quantità pivotale introdotta:

$$1-\alpha = \Pr\left(\underbrace{\bar{X} + z_{\frac{\alpha}{2}} \frac{1}{\sqrt{n}}}_{L(X_1, \dots, X_n)} \leq \theta \leq \underbrace{\bar{X} + z_{1-\frac{\alpha}{2}} \frac{1}{\sqrt{n}}}_{U(X_1, \dots, X_n)}\right);$$

così facendo, si ha che:

$$[l, u] = [L(x_1, \ldots, x_n), U(x_1, \ldots, x_n)] = \left[\bar{x} + z_{\frac{\alpha}{2}} \frac{1}{\sqrt{n}}, \bar{x} + z_{1-\frac{\alpha}{2}} \frac{1}{\sqrt{n}}\right]$$

risulta intervallo di confidenza di livello $1-\alpha$ per $\theta$. Si noti che, stante la natura continua della quantità pivotale per $\theta$ considerata, includere o meno gli estremi dell'intervallo è assolutamente indifferente: nel prosieguo, infatti, gli estremi verranno solitamente esclusi.

# Esercizio 1

*Si supponga di effettuare $n$ prove indipendenti nella variabile casuale $X \sim N(\mu, \sigma^2)$.*

1. *Si costruisca un intervallo di confidenza per $\mu$ di livello $1-\alpha$ e si studi il comportamento della relativa ampiezza al divergere di $n$.*

*Si ponga ora $n = 13$, $1-\alpha = 0.9$ e si supponga di aver ottenuto $\sum_{i=1}^{13} x_i = 17$ e $\sum_{i=1}^{n} x_i^2 = 23.48$:*

2. *si determini numericamente l'intervallo di confidenza costruito in termini generali al punto 1;*
3. *si individui l'intervallo di confidenza per $\mu$ basato, a differenza del precedente, sui quantili di ordine $\frac{\alpha}{4}$ e $1-\frac{3\alpha}{4}$ dell'opportuna v.c. $\mathcal{T}$ di Student e si decida, motivandone la risposta, quale fra i due intervalli individuati risulti preferibile (suggerimento: può essere utile sapere che $\Pr(\mathcal{T}_{12} < 1.53796) = 0.925$);*
4. *supponendo $\sigma^2 = 1$, si dica quante prove indipendenti nella v.c. $X$ si dovrebbero effettuare per far sì che l'intervallo di confidenza per $\mu$ di livello $1-\alpha = 0.9$ abbia lunghezza inferiore a $\frac{2}{10}$.*

# Soluzione

Sia $X \sim N(\mu, \sigma^2)$, $\mu \in \mathrm{R}$ ignota ed oggetto dell'interesse inferenziale, $\sigma^2 \in \mathrm{R}^+$ supposta ignota nei punti 1,2,3 e nota nel punto 4 dell'esercizio; siano inoltre $X_1, \ldots, X_n$ i.i.d. a $X$. Trattasi di un problema di stima intervallare per $\mu$.

1. Poiché $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, risulta $Z = \frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$. Inoltre $U = \frac{n-1}{\sigma^2}S^2 \sim \chi^2_{n-1}$ ed è indipendente da $Z$; pertanto:

$$\frac{Z}{\sqrt{\frac{U}{n-1}}} = \frac{\bar{X}-\mu}{\sqrt{\frac{S^2}{n}}}$$

è una variabile casuale $\mathcal{T}$ di Student con $n-1$ gradi di libertà . In definitiva:

$$Q(X_1, \ldots, X_n, \mu) = \frac{\bar{X}-\mu}{\sqrt{\frac{S^2}{n}}} \sim \mathcal{T}_{n-1},$$

essendo funzione di $(X_1, \ldots, X_n)$ e di $\mu$ ed avente legge di distribuzione non dipendente da $\mu$, è una quantità pivotale per $\mu$. Pertanto, ai fini della costruzione di un intervallo di confidenza di livello $1-\alpha$ per $\mu$, si consideri l'affermazione probabilistica:

$$1-\alpha = \Pr\left(t_{n-1,\frac{\alpha}{2}} < \frac{\bar{X}-\mu}{\sqrt{\frac{S^2}{n}}} < t_{n-1,1-\frac{\alpha}{2}}\right),$$

da cui, essendo

$$t_{n-1,\frac{\alpha}{2}} = -t_{n-1,1-\frac{\alpha}{2}},$$

in virtù della simmetria della densità di probabilità di $\mathcal{T}_{n-1}$ rispetto alla propria media (pari a zero), discende tramite inversione analitica della quantità pivotale introdotta:

$$1-\alpha = \Pr\left(\bar{X} - t_{n-1,1-\frac{\alpha}{2}}\sqrt{\frac{S^2}{n}} < \mu < \bar{X} + t_{n-1,1-\frac{\alpha}{2}}\sqrt{\frac{S^2}{n}}\right).$$

In definitiva:

$$\left(\bar{x} - t_{n-1,1-\frac{\alpha}{2}}\sqrt{\frac{s^2}{n}}, \bar{x} + t_{n-1,1-\frac{\alpha}{2}}\sqrt{\frac{s^2}{n}}\right)$$

è intervallo di confidenza di livello $1-\alpha$ per $\mu$.
La relativa ampiezza risulta:

$$2t_{n-1,1-\frac{\alpha}{2}}\sqrt{\frac{s^2}{n}}$$

ed è tale che:

$$\lim_{n\to\infty} 2t_{n-1,1-\frac{\alpha}{2}}\sqrt{\frac{s^2}{n}} = 0.$$

2. Essendo $\alpha = 0.1$, dalle tavole della v.c. $\mathcal{T}$ di Student, risulta:

$$t_{n-1,1-\frac{\alpha}{2}} = t_{12,0.95} = 1.7823;$$

inoltre, si ha che:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{17}{13} = 1.3077,$$

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}{n-1} = 0.1041.$$

In definitiva, l'intervallo di confidenza per $\mu$ di livello $1 - \alpha = 0.9$ è:

$$(1.1482, 1.4672)\,.$$

3. L'intervallo di confidenza per $\mu$ richiesto ha espressione:

$$\left(\bar{x} + t_{n-1,\frac{\alpha}{4}}\sqrt{\frac{s^2}{n}}, \bar{x} + t_{n-1,1-\frac{3\alpha}{4}}\sqrt{\frac{s^2}{n}}\right)$$

e, al pari del precedente, risulta caratterizzato da livello:

$$1 - \frac{3\alpha}{4} - \frac{\alpha}{4} = 1 - \alpha.$$

Poiché $\alpha = 0.1$, si ha che:

$$1 - \frac{3\alpha}{4} = 1 - 0.075 = 0.925 :$$

essendo noto dal testo dell'esercizio che $\Pr(\mathcal{T}_{12} < 1.53796) = 0.925$, risulta dunque:

$$t_{n-1,1-\frac{3\alpha}{4}} = t_{12,0.925} = 1.53796;$$

d'altra parte, dalle tavole della v.c. $\mathcal{T}$ di Student, risulta:

$$t_{n-1,1-\frac{\alpha}{4}} = t_{12,0.975} = 2.1788,$$

dunque, in virtù della simmetria della densità di probabilità della v.c. $\mathcal{T}$ di Student rispetto alla propria media (pari a zero), si ha che:

$$t_{n-1,\frac{\alpha}{4}} = -t_{n-1,1-\frac{\alpha}{4}} = -2.1788.$$

Pertanto, l'intervallo di confidenza per $\mu$ di livello $1 - \alpha = 0.9$ richiesto risulta:

$$(1.1127, 1.4453)\,.$$

4. In generale, l'intervallo di confidenza per $\mu$ di livello $1 - \alpha$ con $\sigma^2$ nota ha espressione:

$$\left(\bar{x} + z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}, \bar{x} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}\right),$$

ove, essendo la densità di probabilità della v.c. Normale Standard simmetrica rispetto alla propria media (pari a zero), si ha che: $z_{\frac{\alpha}{2}} = -z_{1-\frac{\alpha}{2}}$; pertanto, esso risulta:

$$\left(\bar{x} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}, \bar{x} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}\right),$$

con lunghezza pari a:

$$2z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}.$$

Ai fini della determinazione dell'ampiezza campionaria $n$ che garantisca che la lunghezza dell'intervallo di confidenza in questione sia inferiore ad una quantità $k > 0$ prefissata, si deve risolvere rispetto $n$ la seguente disequazione:

$$2z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}} < k,$$

da cui discende che:

$$2z_{1-\frac{\alpha}{2}} \cdot \sigma < k\sqrt{n} \Leftrightarrow \sqrt{n} > \frac{2z_{1-\frac{\alpha}{2}} \cdot \sigma}{k} \Leftrightarrow n > \left(\frac{2z_{1-\frac{\alpha}{2}} \cdot \sigma}{k}\right)^2.$$

Nel caso in esame, si ha che $\alpha = 0.1$, quindi dalle tavole della v.c. Normale Standard risulta: $z_{1-\frac{\alpha}{2}} = z_{0.95} = 1.65$ (o, equivalentemente, 1.64); inoltre: $\sigma = 1$ e $k = \frac{2}{10} = 0.2$. In definitiva, deve essere:

$$n > \left(\frac{2 \cdot 1.65 \cdot 1}{0.2}\right)^2 = 272.25,$$

dunque $n \geq 273$, ossia per garantire che l'intervallo di confidenza per $\mu$ di livello $1 - \alpha = 0.9$ con $\sigma^2 = 1$ nota abbia lunghezza inferiore a $\frac{2}{10}$ si devono realizzare almeno 273 prove nella v.c. $X$ in questione.

## Esercizio 2

*Siano $X_1, \ldots, X_n$ i.i.d. a $X \sim N(\mu, \sigma^2)$ e si supponga che $n = 25$, $\sum_{i=1}^n x_i = 258.5$ e $\sum_{i=1}^n x_i^2 = 2759.23$.*

1. *Si costruisca una quantità pivotale per $\sigma^2$ nel caso di $\mu$ nota e sulla base di essa si determini un intervallo di confidenza per $\sigma^2$ di livello 90% supponendo $\mu = 10$;*

2. *si costruisca una quantità pivotale per $\sigma^2$ nel caso di $\mu$ ignota e sulla base di essa si determini un intervallo di confidenza per $\sigma^2$ di livello 90%. Si confronti l'intervallo di confidenza così ottenuto con quello determinato al punto precedente fornendo opportuno commento.*

## Soluzione

Sia $X \sim N(\mu, \sigma^2)$, $\sigma^2 \in \mathrm{R}^+$ ignota ed oggetto dell'interesse inferenziale, $\mu \in \mathrm{R}$ supposta nota nel punto 1 e ignota nel punto 2 dell'esercizio; siano inoltre $X_1, \ldots, X_n$ i.i.d. a $X$. Trattasi di un problema di stima intervallare per $\sigma^2$.

1. Poiché $X_i \sim N\left(\mu, \sigma^2\right)$, si ha che: $Z_i = \frac{X_i - \mu}{\sigma} \sim N(0,1)$ e $Z_i^2 = \left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi_1^2$, $\forall i = 1, \ldots, n$. Inoltre, essendo $X_1, \ldots, X_n$ i.i.d., altresì $Z_1^2, \ldots, Z_n^2$ sono i.i.d. e, in virtù della proprietà riproduttiva della variabile casuale Chi-Quadrato rispetto al parametro numero di gradi di libertà , si ha che:

$$\sum_{i=1}^n Z_i^2 = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi_n^2.$$

In definitiva:

$$Q\left(X_1, \ldots, X_n, \sigma^2\right) = \frac{\sum_{i=1}^n \left(X_i - \mu\right)^2}{\sigma^2} \sim \chi_n^2,$$

essendo funzione di $(X_1, \ldots, X_n)$ e di $\sigma^2$ ed avente legge di distribuzione non dipendente da $\sigma^2$, è una quantità pivotale per $\sigma^2$. Pertanto, ai fini della costruzione di un intervallo di confidenza di livello $1-\alpha$ per $\sigma^2$, si consideri l'affermazione probabilistica:

$$1 - \alpha = \Pr\left(c_{n,\frac{\alpha}{2}} < \frac{\sum_{i=1}^n \left(X_i - \mu\right)^2}{\sigma^2} < c_{n,1-\frac{\alpha}{2}}\right),$$

ove $c_{n,\frac{\alpha}{2}}$ e $c_{n,1-\frac{\alpha}{2}}$ identifichino, rispettivamente, i quantili di ordine $\frac{\alpha}{2}$ e $1-\frac{\alpha}{2}$ della v.c. Chi-Quadrato con $n$ gradi di libertà ; tramite inversione analitica della quantità pivotale introdotta, discende dunque:

$$\begin{aligned} 1 - \alpha &= \Pr\left(\frac{c_{n,\frac{\alpha}{2}}}{\sum_{i=1}^n \left(X_i - \mu\right)^2} < \frac{1}{\sigma^2} < \frac{c_{n,1-\frac{\alpha}{2}}}{\sum_{i=1}^n \left(X_i - \mu\right)^2}\right) = \\ &= \Pr\left(\frac{\sum_{i=1}^n \left(X_i - \mu\right)^2}{c_{n,1-\frac{\alpha}{2}}} < \sigma^2 < \frac{\sum_{i=1}^n \left(X_i - \mu\right)^2}{c_{n,\frac{\alpha}{2}}}\right). \end{aligned}$$

In definitiva:

$$\left(\frac{\sum_{i=1}^n \left(x_i - \mu\right)^2}{c_{n,1-\frac{\alpha}{2}}}, \frac{\sum_{i=1}^n \left(x_i - \mu\right)^2}{c_{n,\frac{\alpha}{2}}}\right)$$

è intervallo di confidenza per $\sigma^2$ di livello $1-\alpha$.
Essendo $1 - \alpha = 0.9$, dunque $\alpha = 0.1$ e $n = 25$, dalle tavole della v.c. $\chi^2$ risultano:

$$c_{n,\frac{\alpha}{2}} = c_{25,0.05} = 14.6114, \quad c_{n,1-\frac{\alpha}{2}} = c_{25,0.95} = 37.6525;$$

inoltre, si ha che:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{258.5}{25} = 10.34,$$

$$\sum_{i=1}^n (x_i - \mu)^2 = \sum_{i=1}^n x_i^2 + n\mu^2 - 2n\bar{x}\mu = 89.23.$$

In conclusione, l'intervallo di confidenza per $\sigma^2$ di livello $1 - \alpha = 0.9$ con $\mu = 10$ nota risulta:

$$(2.3698, 6.1069).$$

2. Essendo $X_i, \ldots, X_n$ i.i.d. a $X \sim N\left(\mu, \sigma^2\right)$, con riferimento allo stimatore Varianza Campionaria Corretta vale che:
$$S^2 = \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}{n-1} = \frac{\sigma^2}{n-1}\chi^2_{n-1};$$
pertanto:
$$Q\left(X_1, \ldots, X_n, \sigma^2\right) = \frac{n-1}{\sigma^2} S^2 \sim \chi^2_{n-1},$$
essendo funzione di $(X_1, \ldots, X_n)$ e di $\sigma^2$ ed avente legge di distribuzione non dipendente da $\sigma^2$, è una quantità pivotale per $\sigma^2$. Ai fini della costruzione di un intervallo di confidenza di livello $1-\alpha$ per $\sigma^2$, si consideri dunque l'affermazione probabilistica:
$$1-\alpha = \Pr\left(c_{n-1,\frac{\alpha}{2}} < \frac{n-1}{\sigma^2} S^2 < c_{n-1,1-\frac{\alpha}{2}}\right),$$
ove $c_{n-1,\frac{\alpha}{2}}$ e $c_{n-1,1-\frac{\alpha}{2}}$ identifichino, rispettivamente, i quantili di ordine $\frac{\alpha}{2}$ e $1-\frac{\alpha}{2}$ della v.c. Chi-Quadrato con $n-1$ gradi di libertà ; tramite inversione analitica della quantità pivotale introdotta, discende:
$$\begin{aligned} 1-\alpha &= \Pr\left(\frac{(n-1)S^2}{c_{n-1,1-\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)S^2}{c_{n-1,\frac{\alpha}{2}}}\right) = \\ &= \Pr\left(\frac{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}{c_{n-1,1-\frac{\alpha}{2}}} < \sigma^2 < \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}{c_{n-1,\frac{\alpha}{2}}}\right). \end{aligned}$$
In definitiva:
$$\left(\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{c_{n-1,1-\frac{\alpha}{2}}}, \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{c_{n-1,\frac{\alpha}{2}}}\right)$$
è intervallo di confidenza per $\sigma^2$ di livello $1-\alpha$.
Essendo $1-\alpha = 0.9$, dunque $\alpha = 0.1$ e $n = 25$, dalle tavole della v.c. $\chi^2$ risultano:
$$c_{n-1,\frac{\alpha}{2}} = c_{24,0.05} = 13.8484, \quad c_{n-1,1-\frac{\alpha}{2}} = c_{24,0.95} = 36.4150;$$
inoltre, si ha che:
$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2 = 86.34.$$
In conclusione, l'intervallo di confidenza per $\sigma^2$ di livello $1-\alpha = 0.9$ con $\mu$ ignota risulta:
$$(2.3710, 6.2346).$$
Infine, si noti che l'intervallo di confidenza determinato al punto precedente, supponendo $\mu$ nota, ha lunghezza $6.1069 - 2.3698 = 3.7371$, mentre quello appena determinato, supponendo $\mu$ ignota, ha lunghezza pari a $6.2346 - 2.3710 = 3.8636$, dunque maggiore di quella del primo; infatti, nel caso in cui la media della popolazione sia ignota, dunque ne sia necessaria la stima, l'intervallo di confidenza in questione è meno preciso (ossia più lungo) rispetto al caso in cui la media sia nota a causa dell'errore commesso con tale ulteriore stima.

# Esercizio 3

*è luogo comune considerare il caffè come responsabile di un innalzamento della frequenza cardiaca (numero di battiti del cuore al minuto). Al fine di trarre evidenze a supporto o meno di tale considerazione, viene condotta una sperimentazione su un campione di dieci soggetti, i quali vengono sottoposti alla misurazione della frequenza cardiaca a riposo e a seguito della somministrazione di una tazzina di caffè. Le frequenze cardiache osservate a riposo sono risultate*:

$$\underline{x} = (74, 68, 67, 71, 69, 65, 70, 70, 66, 67)\,;$$

*a seguito dell'assunzione di caffè esse sono risultate invece:*

$$\underline{y} = (71, 72, 69, 66, 73, 77, 69, 68, 71, 78)\,.$$

*Si assuma che le variabili casuali* $D_i = X_i - Y_i$, $i = 1, \ldots, 10$, *che descrivano la differenza tra la frequenza cardiaca a riposo e a seguito dell'assunzione di caffè, siano i.i.d. con legge di distribuzione di tipo Normale. Si determini un intervallo di confidenza per* $\mathrm{E}(X) - \mathrm{E}(Y)$ *di livello* $1 - \alpha = 0.95$ *e si commenti il risultato ottenuto con riferimento all'attendibilità del luogo comune proposto.*

# Soluzione

Sia $D = X - Y \sim N(\mu, \sigma^2)$, $\mu = \mathrm{E}(X - Y) \stackrel{\text{lin.}}{=} \mathrm{E}(X) - \mathrm{E}(Y) \in \mathrm{R}$ ignota ed oggetto dell'interesse inferenziale, $\sigma^2 \in \mathrm{R}^+$ ignota; siano inoltre $D_1, \ldots, D_n$ i.i.d. a $D$. Trattasi di un problema di stima intervallare per $\mu$ sulla base di dati cosiddetti "appaiati".

Si è nuovamente nel caso di una popolazione Normale con media ignota oggetto dell'interesse inferenziale e varianza ignota; pertanto:

$$\left(\bar{d} - t_{n-1,1-\frac{\alpha}{2}}\sqrt{\frac{s^2}{n}}, \bar{d} + t_{n-1,1-\frac{\alpha}{2}}\sqrt{\frac{s^2}{n}}\right)$$

è l'espressione generale dell'intervallo di confidenza richiesto, ove $\bar{d} = \frac{\sum_{i=1}^n d_i}{n}$ e $s^2 = \frac{\sum_{i=1}^n \left(d_i - \bar{d}\right)^2}{n-1}$.
Essendo $1 - \alpha = 0.95$, dunque $\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$, $1 - \frac{\alpha}{2} = 0.975$ e $n = 10$, dalle tavole della v.c. $\mathcal{T}$ di Student, risulta:

$$t_{n-1,1-\frac{\alpha}{2}} = t_{9,0.975} = 2.2622;$$

inoltre, le differenze tra le frequenze cardiache a riposo e a seguito dell'assunzione di caffè dei dieci soggetti coinvolti nella sperimentazione risultano:

$$\underline{d} = (3, -4, -2, 5, -4, -12, 1, 2, -5, -11)\,,$$

da cui:

$$\bar{d} = \frac{\sum_{i=1}^n d_i}{n} = -2.7,$$

$$s^2 = \frac{\sum_{i=1}^n \left(d_i - \bar{d}\right)^2}{n-1} = \frac{\sum_{i=1}^n d_i^2 - n\bar{d}^2}{n-1} = \frac{365 - 10 \cdot 7.29}{9} = 32.4556.$$

In definitiva, l'intervallo di confidenza per $\mu$ di livello $1 - \alpha = 0.95$ risulta:

$$(-6.7754, 1.3754)\,,$$

ossia è possibile concludere con una confidenza pari al 95% che il reale e ignoto valore di $\mu$ sia compreso in tale intervallo. A questo punto, ai fini del commento del risultato ottenuto con riferimento all'attendibilità del luogo comune proposto, si noti che il valore $\mu = 0$ corrisponde al caso $\mathrm{E}(X) = \mathrm{E}(Y)$ in cui le frequenze cardiache medie a riposo e a seguito della somministrazione del caffè coincidono, emblematico del fatto che l'assunzione del caffè non ha alcun effetto sulla frequenza cardiaca. Poichè il valore 0 è incluso nell'intervallo, non è possibile escludere con una confidenza del 95% che il caffè non sia responsabile dell'innalzamento della frequenza cardiaca.

## Esercizio 4

*Si supponga che la durata di una tipologia di lampade alogene possa essere interpretata mediante una variabile casuale Esponenziale di parametro $\theta$. Avendo registrato le durate $(x_1, \ldots, x_n)$ di un campione $n-$dimensionale di lampade in questione, si risponda ai seguenti quesiti:*

1. *si individui una quantità pivotale per $\theta$ e sulla base di essa si determini un intervallo di confidenza per $\theta$ di livello $1-\alpha$;*
2. *si individui una quantità pivotale per $\theta$ supponendo che la durata delle lampade in questione sia interpretabile da una variabile casuale* $\text{Gamma}(3, \theta)$;
3. *si determini numericamente l'intervallo costruito in termini generali al punto 1 supponendo $n = 40$, $\alpha = 0.05$ e $\bar{x} = 2.46$.*

## Soluzione

Sia $X \sim \text{Exp}(\theta)$, $\theta \in \mathrm{R}^+$ ignoto; siano inoltre $X_1, \ldots, X_n$ i.i.d. a $X$. Trattasi di un problema di stima intervallare per $\theta$.

1. Poichè $X \sim \text{Exp}(\theta) \equiv \text{Gamma}(1, \theta)$ ed in virtù della proprietà riproduttiva della variabile casuale Gamma rispetto al parametro di forma, risulta:
$$\sum_{i=1}^{n} X_i \sim \text{Gamma}(n, \theta);$$
inoltre, in quanto trasformazione di scala di $\sum_{i=1}^{n} X_i$, si ha che:
$$2\theta \sum_{i=1}^{n} X_i \sim \text{Gamma}\left(n, \frac{\theta}{2\theta}\right) \equiv \text{Gamma}\left(\frac{2n}{2}, \frac{1}{2}\right) \equiv \chi^2_{2n}.$$
In definitiva:
$$Q\left(X_1, \ldots, X_n, \theta\right) = 2\theta \sum_{i=1}^{n} X_i \sim \chi^2_{2n},$$
essendo funzione di $(X_1, \ldots, X_n)$ e $\theta$ ed avente legge di distribuzione non dipendente da $\theta$, è una quantità pivotale per $\theta$.
Pertanto, ai fini della costruzione di un intervallo di confidenza di livello $1-\alpha$ per $\theta$, si consideri l'affermazione probabilistica:
$$1 - \alpha = \Pr\left(c_{2n,\frac{\alpha}{2}} < 2\theta \sum_{i=1}^{n} X_i < c_{2n,1-\frac{\alpha}{2}}\right),$$
ove $c_{2n,\frac{\alpha}{2}}$ e $c_{2n,1-\frac{\alpha}{2}}$ identifichino, rispettivamente, i quantili di ordine $\frac{\alpha}{2}$ e $1-\frac{\alpha}{2}$ della v.c. Chi-Quadrato con $2n$ gradi di libertà ; tramite inversione analitica della quantità pivotale introdotta, discende dunque:
$$1 - \alpha = \Pr\left(\frac{c_{2n,\frac{\alpha}{2}}}{2\sum_{i=1}^{n} X_i} < \theta < \frac{c_{2n,1-\frac{\alpha}{2}}}{2\sum_{i=1}^{n} X_i}\right).$$
In definitiva:
$$\left(\frac{c_{2n,\frac{\alpha}{2}}}{2\sum_{i=1}^{n} x_i}, \frac{c_{2n,1-\frac{\alpha}{2}}}{2\sum_{i=1}^{n} x_i}\right)$$
è intervallo di confidenza per $\theta$ di livello $1-\alpha$.

2. Supponendo che $X \sim \text{Gamma}(3, \theta)$, in virtù della proprietà riproduttiva della variabile casuale Gamma rispetto al parametro di forma, risulta:
$$\sum_{i=1}^{n} X_i \sim \text{Gamma}(3n, \theta);$$

inoltre, in quanto trasformazione di scala di $\sum_{i=1}^{n} X_i$, si ha che:

$$2\theta \sum_{i=1}^{n} X_i \sim \text{Gamma}\left(3n, \frac{\theta}{2\theta}\right) \equiv \text{Gamma}\left(\frac{6n}{2}, \frac{1}{2}\right) \equiv \chi^2_{6n}.$$

In definitiva:

$$Q\left(X_1, \ldots, X_n, \theta\right) = 2\theta \sum_{i=1}^{n} X_i \sim \chi^2_{6n},$$

essendo funzione di $(X_1, \ldots, X_n)$ e $\theta$ ed avente legge di distribuzione non dipendente da $\theta$, è una quantità pivotale per $\theta$.

3. Essendo $\alpha = 0.05$, dunque $\frac{\alpha}{2} = 0.025$ e $n = 40$, dalle tavole della v.c. $\chi^2$ risultano:

$$c_{2n,\frac{\alpha}{2}} = c_{80,0.025} = 57.1532, \quad c_{2n,1-\frac{\alpha}{2}} = c_{80,0.975} = 106.6286;$$

inoltre, si ha che:

$$\sum_{i=1}^{n} x_i = n\bar{x} = 98.4.$$

In conclusione, l'intervallo di confidenza per $\theta$ di livello $1 - \alpha = 0.95$ risulta:

$$(0.2904, 0.5418)\,.$$

# Esercizio 5

*Si conduca una prova nella variabile casuale* $X \sim U\left[\theta - \frac{1}{2}, \theta + \frac{1}{2}\right]$, $\theta \in \mathrm{R}$.

1. *Si verifichi che* $Y = X - \theta$ *è quantità pivotale per* $\theta$;

2. *supponendo di aver ottenuto il valore* $x = 3.7$, *si individui, sulla base di quanto determinato al punto 1, l'intervallo di confidenza centrale per* $\theta$ *di livello* $1 - \alpha = 0.9$ *(suggerimento: si ricorda che ai fini della costruzione di un intervallo di confidenza centrale, nell'inversione analitica della quantità pivotale* $Y$, *si deve assumere come punto di partenza la considerazione dell'intervallo* $(a, b)$ *sottoinsieme del supporto di* $Y$ *tale che* $\Pr(Y < a) = \Pr(Y > b) = \frac{\alpha}{2}$*).*

# Soluzione

Sia $X \sim U\left[\theta - \frac{1}{2}, \theta + \frac{1}{2}\right]$, $\theta \in \mathrm{R}$ ignoto. Trattasi di un problema di stima intervallare per $\theta$.

1. Essendo $Y = X - \theta$ funzione di $X$ e $\theta$, ai fini della verifica della relativa natura di quantità pivotale per $\theta$, si rende necessario mostrare che la relativa legge di distribuzione non dipende da $\theta$. Poichè $Y = f(X) = X - \theta$ è trasformazione biunivoca di $X$ con inversa $X = f^{-1}(Y) = Y + \theta$ ed essendo:

$$\varphi_X(x;\theta) = I_{\left[\theta-\frac{1}{2},\theta+\frac{1}{2}\right]}(x),$$

in virtù del metodo basato sulla densità , la legge di distribuzione di $Y$ risulta:

$$\varphi_Y(y) = \varphi_X\left(f^{-1}(y);\theta\right)\left|\frac{dx}{dy}\right| = I_{\left[\theta-\frac{1}{2},\theta+\frac{1}{2}\right]}(y+\theta),$$

ove:

$$I_{\left[\theta-\frac{1}{2},\theta+\frac{1}{2}\right]}(y+\theta) = 1 \Leftrightarrow \theta - \frac{1}{2} \leq y + \theta \leq \theta + \frac{1}{2} \Leftrightarrow -\frac{1}{2} \leq y \leq \frac{1}{2};$$

pertanto:

$$\begin{aligned} \varphi_Y(y) &= I_{\left[-\frac{1}{2},+\frac{1}{2}\right]}(y) = \\ &= 1, \quad y \in \left[-\frac{1}{2}, +\frac{1}{2}\right], \end{aligned}$$

dunque:

$$Y \sim U\left[-\frac{1}{2}, \frac{1}{2}\right].$$

In conclusione, avendo legge di distribuzione non dipendente da $\theta$, $Y = X - \theta$ è quantità pivotale per $\theta$.

2. Ai fini dell'identificazione dell'intervallo di confidenza centrale per $\theta$ di livello $1 - \alpha = 0.9$, si consideri l'affermazione probabilistica:

$$0.9 = \Pr(a < Y < b),$$

ove, come suggerito dal testo dell'esercizio, $a$ e $b$ devono essere tali che:

$$\Pr(Y < a) = \Pr(Y > b) = 0.05,$$

ossia:

$$\int_{-\frac{1}{2}}^{a} dy = 0.05 \Leftrightarrow y|_{-\frac{1}{2}}^{a} = 0.05 \Leftrightarrow a + \frac{1}{2} = 0.05 \Leftrightarrow a = -0.45,$$

$$\int_{b}^{\frac{1}{2}} dy = 0.05 \Leftrightarrow y|_{b}^{\frac{1}{2}} = 0.05 \Leftrightarrow \frac{1}{2} - b = 0.05 \Leftrightarrow b = 0.45.$$

A questo punto, tramite inversione analitica della quantità pivotale introdotta, discende:

$$\begin{aligned} 0.9 &= \Pr(-0.45 < X - \theta < 0.45) = \\ &= \Pr(X - 0.45 < \theta < X + 0.45), \end{aligned}$$

da cui, avendo osservato $x = 3.7$, l'intervallo di confidenza centrale per $\theta$ di livello 0.9 risulta:

$$(x - 0.45, x + 0.45) = (3.7 - 0.45, 3.7 + 0.45) = (3.25, 4.15).$$

# Esercizio 6

*Siano $X_1, \dots, X_n$ i.i.d. a $X \sim N(\mu, \sigma^2)$.*

1. *Sia $\sigma^2$ noto. Trovare il valore di n più piccolo affinché l'intervallo di confidenza per $\mu$ di livello $1-\alpha = 0.95$ abbia lunghezza non superiore a $\sigma/4$;*

2. *Sia $\sigma^2$ incognito. Trovare il valore di n più piccolo affinché l'intervallo di confidenza per $\mu$ di livello $1-\alpha = 0.95$ abbia lunghezza non superiore a $\sigma/4$ con probabilità 0.9.;*

# Soluzione

1. L'intervallo di confidenza per $\mu$ di livello $1 - \alpha = 0.95$ è pari a

$$\left[\bar{x} - z_{0.975}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{0.975}\frac{\sigma}{\sqrt{n}}\right]$$

dove $z_{0.975} \approx 1.96$ quindi la lunghezza dell'intervallo è pari a:

$$2(1.96)\frac{\sigma}{\sqrt{n}}.$$

Pertanto

$$2(1.96)\frac{\sigma}{\sqrt{n}} \leq \frac{\sigma}{4} \Leftrightarrow \sqrt{n} \geq 4 \cdot 2 \cdot 1.96 \Leftrightarrow n \geq 64 \cdot (1.96)^2 \approx 245.9$$

quindi $n = 246$.

2. La lunghezza dell'intervallo è pari a:

$$2t_{0.975,n-1}\frac{s}{\sqrt{n}}$$

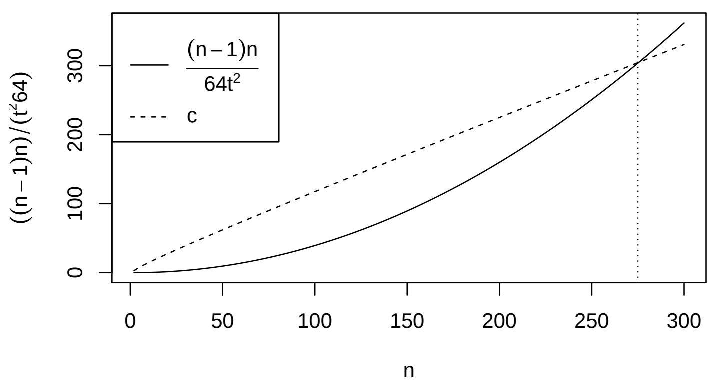
quindi

$$\mathrm{P}\left(2t_{0.975,n-1}\frac{S}{\sqrt{n}} \leq \frac{\sigma}{4}\right) = 0.9$$

ovvero

$$\mathrm{P}\left(4t^2_{0.975,n-1}\frac{S^2}{n} \leq \frac{\sigma^2}{16}\right) = 0.9 \Leftrightarrow \mathrm{P}\left(\underbrace{(n-1)\frac{S^2}{\sigma^2}}_{\chi^2_{n-1}} \leq \frac{(n-1)n}{64 \cdot t^2_{0.975,n-1}}\right) = 0.9$$

quindi $\frac{(n-1)n}{64 \cdot t^2_{0.975,n-1}} = \chi^2_{0.9,n-1}$. Numericamente si ottiene $n = 276$:



# Esercizio 7

*Siano $X_1, \ldots, X_n$ i.i.d. a $X \sim N(\mu, \sigma^2)$. Per $\alpha = 0.1$ e $n = 5$, si calcoli la probabilità che la lunghezza dell'intervallo di confidenza per $\mu$ con $\sigma^2$ incognito sia inferiore della lunghezza dell'intervallo di confidenza per $\mu$ che assume $\sigma^2$ noto. .*

# Soluzione

La lunghezza dell'intervallo di confidenza per $\mu$ con $\sigma^2$ incognito è

$$2t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}}$$

mentre la lunghezza dell'intervallo di confidenza per $\mu$ con $\sigma^2$ noto è

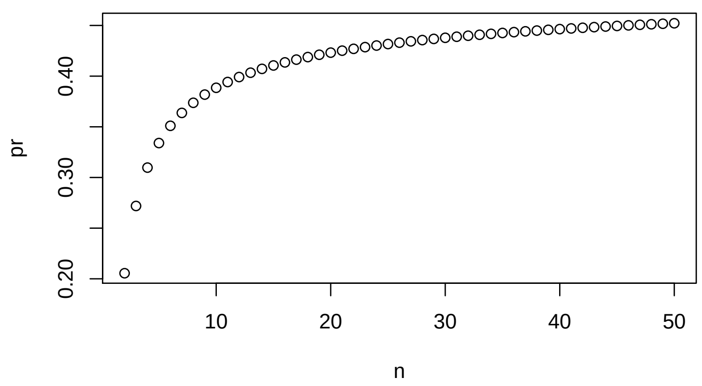
$$2z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

quindi

$$\mathrm{P}\left(2t_{1-\alpha/2,n-1}\frac{S}{\sqrt{n}} \leq 2z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \Leftrightarrow \mathrm{P}\left(\underbrace{\frac{(n-1)S^2}{\sigma^2}}_{\chi^2_{n-1}} \leq \left(\frac{z_{1-\alpha/2}}{t_{1-\alpha/2,n-1}}\right)^2 (n-1)\right)$$

Per $\alpha = 0.1$ e $n = 5$ abbiamo

$$\mathrm{P}\left(\chi^2_4 \leq \left(\frac{1.64}{2.13}\right)^2 \cdot 4\right) \approx 0.334$$



## Esercizio 8

*Siano* $X_1, \ldots, X_n$ *i.i.d. a* $X \sim N(\mu, \sigma^2)$. Si calcoli la lunghezza attesa dell'intervallo di confidenza per $\mu$ di livello $1 - \alpha$ per

1. $\sigma^2$ *noto*;
2. $\sigma^2$ *incognito (si ricordi che* $\mathrm{E}(S) = k\sigma$ *con* $k = \left[\frac{\Gamma(\frac{n-1}{2})\sqrt{n-1}}{\sqrt{2}\Gamma(n/2)}\right]^{-1}$ *).*

## Soluzione

1. La lunghezza attesa è pari a
$$\mathrm{E}\left(2z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 2z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$
2. La lunghezza attesa è pari a
$$\mathrm{E}\left(2t_{1-\alpha/2,n-1}\frac{S}{\sqrt{n}}\right) = 2t_{1-\alpha/2,n-1}\frac{\mathrm{E}(S)}{\sqrt{n}}$$
dove $\mathrm{E}(S) = k\sigma$ con $k^{-1} = \frac{\Gamma(\frac{n-1}{2})\sqrt{n-1}}{\sqrt{2}\Gamma(n/2)}$.

La differenza tra le due lunghezze attese è $\frac{2\sigma}{\sqrt{n}}(z_{1-\alpha/2} - kt_{1-\alpha/2,n-1})$ che tende a 0 per $n \to \infty$.

## Esercizio 9

*Siano* $X_1$ *e* $X_2$ *i.i.d. a* $X \sim \text{Uniforme}(\theta - 0.5, \theta + 0.5)$ *Si calcoli il livello di confidenza relativo al seguente intervallo di confidenza per* $\theta$*:*

$$[\min\{x_1, x_2\}, \max\{x_1, x_2\}].$$

*Si scriva infine la funzione di verosimiglianza di* $\theta$*.*

## Soluzione

Abbiamo

$$\mathrm{P}\left(\min\{X_1, X_2\} < \theta < \max\{X_1, X_2\}\right) = \mathrm{P}\left(\{X_1 < \theta < X_2\} \cup \{X_2 < \theta < X_1\}\right)$$

e poichè abbiamo due eventi disgiunti

$$= \mathrm{P}(X_1 < \theta < X_2) + \mathrm{P}(X_2 < \theta < X_1)$$

quindi per l'indipendenza di $X_1$ e $X_2$

$$= \mathrm{P}(X_1 < \theta)\mathrm{P}(\theta < X_2) + \mathrm{P}(X_2 < \theta)\mathrm{P}(\theta < X_1)$$

e per l'identica distribuzione (dove $\mathrm{P}(X < x) = \int_{\theta-0.5}^{x} dx = x - \theta + 0.5$)

$$= \mathrm{P}(X < \theta)\mathrm{P}(\theta < X) + \mathrm{P}(X < \theta)\mathrm{P}(\theta < X) = 0.5 \cdot 0.5 + 0.5 \cdot 0.5 = 0.5$$

La funzione di verosimiglianza è

$$L(\theta) = 1_{[\theta-0.5,\theta+0.5](x_1)} 1_{[\theta-0.5,\theta+0.5](x_2)}$$

che vale 1 se $\{x_i \geq \theta - 0.5\} \cap \{x_i \leq \theta + 0.5\}$, $i = 1, 2$ ovvero $\{\min\{x_1, x_2\} \geq \theta - 0.5\} \cap \{\max\{x_1, x_2\} \leq \theta + 0.5\}$; quindi

$$L(\theta) = 1_{[\max\{x_1,x_2\}-0.5, \min\{x_1,x_2\}+0.5]}(\theta).$$

## Esercizio 10

*Siano* $X_1, \ldots, X_n$ *i.i.d. a* $X \sim \text{Uniforme}(0, \theta)$, $\theta > 0$. *Sia* $Y = \max(X_1, \ldots, X_n)$ *(si ricordi che* $Y$ *ha funzione di densità* $\varphi_Y(y) = \frac{n}{\theta^n} y^{n-1}$ *per* $0 < y < \theta$*). Si mostri che l'intervallo di confidenza*

$$[y, (1-\alpha)^{-1/n} y]$$

*è di livello di confidenza* $\alpha$*.*

## Soluzione

Abbiamo

$$P(Y \leq \theta \leq (1-\alpha)^{-1/n} Y) = P(\theta(1-\alpha)^{1/n} \leq Y \leq \theta) = \int_{\theta(1-\alpha)^{1/n}}^{\theta} \varphi_Y(y) dy$$

ovvero

$$= \frac{n}{\theta^n} \int_{\theta(1-\alpha)^{1/n}}^{\theta} y^{n-1} dy = \frac{n}{\theta^n} \left[ \frac{\theta^n}{n} - \frac{(\theta(1-\alpha)^{1/n})^n}{n} \right] = 1 - (1-\alpha) = \alpha$$

# Esercizio 11

*Si supponga che ogni mese per dieci anni si siano contati i fiori di una pianta, nota per essere estremamente fiorifera, riscontrando un numero medio campionario di fiori pari a* 123. *Si supponga altresì che il numero di fiori presenti in un mese sulla pianta in questione sia descrivibile tramite una v.c. di Poisson di parametro* $\lambda$.

1. *Si determini una quantità asintoticamente pivotale per il numero medio mensile di fiori* $\lambda$ *della pianta in esame*;
2. *sulla base di quanto determinato al punto 1, si costruisca un intervallo di confidenza approssimato per* $\lambda$ *di livello* $1-\alpha=0.9$.

# Soluzione

Sia $X \sim \text{Poisson}(\lambda)$, $\lambda \in \mathrm{R}^+$ ignoto; siano inoltre $X_1, \ldots, X_n$ i.i.d. a $X$. Trattasi di un problema di stima intervallare per $\lambda = \mathrm{E}(X)$.

1. Sia $\hat{\lambda}_n$ stimatore di massima verosimiglianza per $\lambda$. Stando alle proprietà dello stimatore di massima verosimiglianza, è noto che asintoticamente si distribuisce come una variabile casuale Normale con media $E(\hat{\lambda}_n) = \lambda$ e varianza pari all'inversa dell'informazione attesa di Fisher $[\mathcal{I}(\lambda)]^{-1} = \frac{\lambda}{n}$. Conseguentemente si ha che:

$$\frac{\bar{X}-\lambda}{\sqrt{\frac{\lambda}{n}}} \xrightarrow{d} N(0,1).$$

   Essendo $\lambda$ parametro ignoto, al fine di ricavare una quantità pivotale può essere utile sostituire l'informazione attesa di Fisher con l'informazione osservata, con inversa pari a $[\mathcal{J}(\lambda)]^{-1} = \frac{\bar{X}}{n}$. In definitiva:

$$Q(X_1, \ldots, X_n, \lambda) = \frac{\bar{X}-\lambda}{\sqrt{\frac{\bar{X}}{n}}} \overset{n\to\infty}{\sim} N(0,1).$$

   essendo funzione di $(X_1, \ldots, X_n)$ e $\lambda$ ed avente legge di distribuzione indipendente da $\lambda$ per $n \to \infty$, è una quantità asintoticamente pivotale per $\lambda$.

2. Essendo l'ampiezza campionaria $n = 12 \cdot 10 = 120$ (mesi) sufficientemente grande, ai fini della costruzione di un intervallo di confidenza approssimato di livello $1-\alpha = 0.9$ per $\lambda = \mathrm{E}(X)$, è possibile sfruttare la quantità asintoticamente pivotale determinata al punto precedente. Si consideri dunque l'affermazione probabilistica:

$$1-\alpha \cong \Pr\left(-z_{1-\frac{\alpha}{2}} < \frac{\bar{X}-\lambda}{\sqrt{\frac{\bar{X}}{n}}} < z_{1-\frac{\alpha}{2}}\right),$$

   da cui, tramite inversione analitica della quantità pivotale introdotta, discende:

$$1-\alpha \cong \Pr\left(\bar{X} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}}{n}} < \lambda < \bar{X} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}}{n}}\right).$$

   Pertanto:

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

   è l'espressione generale dell'intervallo di confidenza richiesto, per la cui particolarizzazione al presente caso si noti che, essendo $1-\alpha = 0.9$, dunque $\alpha = 0.1$ e $\frac{\alpha}{2} = 0.05$, dalle tavole della v.c. Normale Standard risulta:

$$z_{1-\frac{\alpha}{2}} = z_{0.95} = 1.64$$

   (o, equivalentemente, 1.65). In definitiva, l'intervallo di confidenza approssimato di livello $1-\alpha = 0.9$ per $\lambda = \mathrm{E}(X)$ risulta:

$$(121.3397, 124.6603).$$