Soluzione all'esercizio del 2022-05-27 creato per luigi.miazzo

Sia $(X_1, \dots, X_n)$ un campione casuale (i.e. le X_i sono indipendenti ed identicamente distribuite) da una distribuzione di media μ e varianza σ^2.

Sia $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ la media campionaria, stimatore non distorto del valore atteso, μ, della distribuzione. Ricordiamo che $\mathbb{V}\mathrm{ar}(\bar{X}_n) = \frac{\sigma^2}{n}$.

Quesiti e soluzioni

Quesito 1

Si consideri lo stimatore di σ^2: $S^2 = \frac{1}{n}\sum_{i=1}^n \left(X_i - \bar{X}_n\right)^2$.

Lo stimatore è non distorto, cioè $\mathbb{E}(S^2) = \sigma^2$? Rispondere `TRUE` o `FALSE`.

Per verificare se lo stimatore S^2 è non distorto σ^2, calcoliamone il valore atteso:

$$\begin{aligned}
\mathbb{E}(S^2) &= \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left((X_i - \bar{X}_n)^2\right) \\
&= \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^n (X_i^2 + \bar{X}_n^2 - 2X_i\bar{X}_n)\right) \\
&= \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^n X_i^2 + \sum_{i=1}^n \bar{X}_n^2 - 2\bar{X}_n\sum_{i=1}^n X_i\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^n \mathbb{E}(X_i^2) + n\mathbb{E}(\bar{X}_n^2) - 2n\mathbb{E}(\bar{X}_n^2)\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^n \mathbb{E}(X_i^2) - n\mathbb{E}(\bar{X}_n^2)\right) \\
&= \frac{1}{n}\left(n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2\right) \\
&= \frac{n-1}{n}\sigma^2.
\end{aligned}$$

Dove abbiamo utilizzato:

- la linearità del valore atteso (valore atteso della somma è la somma dei valori attesi),
- il fatto che $\sum_{i=1}^n X_i = n\bar{X}_n$,
- la scomposizione della varianza $\mathbb{V}\mathrm{ar}(T) = \mathbb{E}(T^2) - \mathbb{E}^2(T)$,
- $\sigma^2 = \mathbb{V}\mathrm{ar}(X_i) = \mathbb{E}(X_i^2) - \mu^2$ e
- $\mathbb{V}\mathrm{ar}(\bar{X}_n) = \frac{\sigma^2}{n}$.

In particolare combinando gli ultimi quattro punti abbiamo che

$$\frac{\sigma^2}{n} = \mathbb{V}\mathrm{ar}(\bar{X}_n) = \mathbb{E}(\bar{X}_n^2) - \mathbb{E}^2(\bar{X}_n)$$

da cui $\mathbb{E}(\bar{X}_n^2) = \frac{\sigma^2}{n} + \mu^2$.

Lo stimatore S^2 è distorto, $\mathbb{E}(S^2) \neq \sigma^2$.

- La risposta corretta è: FALSE
- La risposta inserita è: FALSE

Quesito 2.

Dato il seguente campione estratto

$$(3, 6, 3, 7, 0, 4, 9, 6, 5, 6, 8, 4, 2, 3, 11)$$

stimarne la varianza, usando S^2, se non distorto, o la sua versione opportunamente "corretta", se S^2 è distorto.

S^2 è distorto. Rendiamolo non distorto:

$$\mathbb{E}(S^2) = \frac{n-1}{n}\sigma^2 \quad \Leftrightarrow \quad \frac{n}{n-1}\mathbb{E}(S^2) = \sigma^2$$

Da cui $\frac{n}{n-1}S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ è uno stimatore non distorto (chiamato varianza campionaria e spesso indicato con S_n^2).

Non ci resta che calcolare la stima per il campione dato. Otteniamo $\bar{x}_n$ con la funzione `mean(x)` ove `x` è l'array in cui abbiamo salvato i dati; e poi `1 / (length(x) - 1) * sum((x - mean(x))^ 2)` o, semplicemente `var(x)`.

- La risposta corretta è: 8.2666667
- La risposta inserita è: 8.266667

Quesito 3.

Ci vengono ora date la media esatta della distribuzione $\mu = 5$ e $n = 15$, la dimensione del campione casuale. Modifichiamo quindi il nostro stimatore S^2 usando la media nota, i.e.

$$S^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2.$$

Calcolare la distorsione del nuovo stimatore.

Come per il quesito 1:

$$\begin{aligned}
\mathbb{E}(S^2) &= \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left((X_i - \mu)^2\right) \\
&= \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^n (X_i^2 + \mu^2 - 2X_i\mu)\right) \\
&= \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^n X_i^2 + \sum_{i=1}^n \mu^2 - 2\mu\sum_{i=1}^n X_i\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^n \mathbb{E}(X_i^2) + n\mu^2 - 2n\mu^2\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^n \mathbb{E}(X_i^2) - n\mu^2\right) \\
&= \frac{1}{n}\left(n\sigma^2 + n\mu^2 - n\mu^2\right) \\
&= \sigma^2.
\end{aligned}$$

Per cui la distorsione di S^2 con media nota è 0.

- La risposta corretta è: 0
- La risposta inserita è: 0