Statistica I - A.A. 2016-2017

Prova scritta - 31 gennaio 2017

Problema 1. (pt 15=3+4+4+4) Un'azienda produce componenti per elettrodomestici. Un suo reparto produce un numero di pezzi giornalieri che varia casualmente di giorno in giorno; con un certo grado di estrapolazione, si supponga che questo numero X aleatorio sia distribuito in modo gaussiano. Nel 2016, il numero medio di pezzi prodotti giornalmente è stato 23.3, con una deviazione standard pari a 6.5.

1. 90 giorni su 100 (mediamente parlando), su quale quantità minima di pezzi prodotti si può contare, in questo regime di lavoro? Spiegare bene come avete traddotto matematicamente questa richiesta.

2. Un dirigente ha studiato un nuovo metodo di produzione che si spera aumenti il numero medio di pezzi prodotti (si può escludere che li diminuisca); non ci sono invece ragioni di ritenere che modifichi la variabilità. Si vuole innanzi tutto stimare in modo bilaterale tale numero medio, nel regime del nuovo metodo, con una precisione assoluta di due pezzi ed un livello di fiducia del 90%. Quanti giorni di lavoro col nuovo metodo si devono aspettare?

3. Il nuovo metodo di produzione viene implementato e vengono registrati i valori giornalieri di produzione per 20 giorni, trovando una media dei valori pari a 25.8, con deviazione standard pari a 6. Possiamo affermare, al 95% che il nuovo metodo funziona? Formalizzare la risoluzione dichiarando le ipotesi e la regione di rifiuto. Dopo aver eseguito il test, calcolare anche il valore p utilizzando i quantili gaussiani e la nuova deviazione come vera ed eseguire nuovamente il test a posteriori.

4. Se la nuova media è in realtà solo pari a 25, che probabilità ha il test precedente di accorgersene (usare la nuova deviazione come vera)? Oltre a calcolare la quantità richiesta, come potreste inquadrare il quesito in uno schema più generale?

Problema 2. (pt 9) Al variare del parametro $\beta > 0$, si consideri la funzione

$$f_\beta(x) = \begin{cases} C_\beta - \beta\,|x| & \text{per } -\frac{C_\beta}{\beta} \le x < \frac{C_\beta}{\beta} \\ 0 & \text{altrove} \end{cases}$$

dove C_β è una costante positiva, dipendente da β.

1. Calcolare C_β in modo che f_β sia una densità di probabilità. Senza calcolare la deviazione standard, sapreste intuire se essa cresce o diminuisce al crescere di β, spiegando perché?

2. Se X è una v.a. X con densità f_β, trovare λ tale che $P(X > \lambda) = 0.1$.
3. Col metodo dei momenti, trovare uno stimatore di β.

Problema 3. (pt 6) Si consideri una v.a. $X \sim N(0, 4)$. Si ponga $Y = e^X$.

1. Calcolare $P(Y \leq 1)$ e trovare λ tale che $P(Y \leq \lambda) = 0.9$.
2. Trovare la densità di probabilità di Y.

1 Soluzioni

Esercizio 1.

1. Detto $X \sim N(23.3, 6.5^2)$, cerchiamo n (meglio se intero) tale che
$$P(X \geq n) \geq 0.9.$$
Cerchiamo x tale che $P(X \geq x) = 0.9$, cioè $P(X \leq x) = 0.1$. Vale allora
$$x = 23.3 + 6.5 \cdot q_{0.1} = 23.3 - 6.5 \cdot 1.28 = 14.98.$$

2. L'uso delle t di Student rende implicita la risoluzione e quindi ci affidiamo alle gaussiane. L'intervallo di fiducia per la media è
$$\overline{X} \pm \frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}}$$
e vogliamo $\frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}} \leq 2$. Sostituendo $\sigma = 6.5$, $\alpha = 0.1$, troviamo
$$n \geq \left(\frac{6.5 \cdot q_{0.95}}{2}\right)^2 = \left(\frac{6.5 \cdot 1.645}{2}\right)^2 = 28.582$$
quindi $n = 29$.

3. E' più naturale, viste le premesse, svolgere un test unilaterale. L'ipotesi nulla è $\mathcal{H}_0$) media $\leq$23.3, l'alternativa $\mathcal{H}_0$) media $>$23.3, la regione di rifiuto, usando le t di Student perché il campione è basso, osservando che $t_{19,0.05} = 1.729$ e che $23.3 + \frac{6 \cdot 1.729}{\sqrt{20}} = 25.62$
$$\left\{\overline{X} > 25.62\right\}$$
e pertanto il test risulta significativo. Il valore p è dato da
$$\begin{aligned} P^{(23.3,6^2)}\left(\overline{X} > 25.8\right) &= P^{(23.3,6^2)}\left(\frac{\overline{X} - 23.3}{6}\sqrt{20} > \frac{25.8 - 23.3}{6}\sqrt{20}\right) \\ &= 1 - \Phi\left(\frac{25.8 - 23.3}{6}\sqrt{20}\right) = 1 - \Phi(1.86) \\ &= 1 - 0.96856 = 0.03144. \end{aligned}$$
E' inferiore a 0.05 e quindi anche così il test risulterebbe significativo.

4.

$$P^{(25,6)}\left(\overline{X}>25.62\right)=P^{(25,6)}\left(\frac{\overline{X}-25}{6}\sqrt{20}>\frac{25.62-25}{6}\sqrt{20}\right)$$
$$=1-\Phi\left(\frac{25.62-25}{6}\sqrt{20}\right)=1-\Phi\left(0.46\right)$$
$$=1-0.67724=0.32276.$$

Molto bassa. E' il calcolo della potenza del test, che si può impostare più in generale calcolando la funzione

$$\mu\mapsto P^{(\mu,6)}\left(\overline{X}>25.62\right).$$

Esercizio 2.

1. Area del triangolo $= 2\cdot\frac{C_\beta}{\beta}C_\beta/2=\frac{C_\beta^2}{\beta}$, quindi $C_\beta=\sqrt{\beta}$ (radice positiva). Guardando il grafico, si capisce che al crescere di β il triangolo si schiaccia e quindi la varianza diminuisce.

2. Il punto λ dev'essere in $\left(0,\frac{C_\beta}{\beta}\right)$. Vale

$$P\left(X>\lambda\right)=\frac{1}{2}\left(\frac{\sqrt{\beta}}{\beta}-\lambda\right)\left(\sqrt{\beta}-\beta\lambda\right)=\frac{1}{2}\left(1-\lambda\sqrt{\beta}\right)^2$$

quindi dobbiamo risolvere $\frac{1}{2}\left(1-\lambda\sqrt{\beta}\right)^2=0.1$, $1-\lambda\sqrt{\beta}=\sqrt{0.2}$, $\lambda\sqrt{\beta}=1-\sqrt{0.2}$,

$$\lambda=\frac{1-\sqrt{0.2}}{\sqrt{\beta}}.$$

3. Media zero, quindi non basta.

$$Var\left[X\right]=E\left[X^2\right]=\int x^2 f_\beta\left(x\right)dx=2\int_0^{\frac{C_\beta}{\beta}}x^2\left(C_\beta-\beta x\right)dx=2\left[C_\beta\frac{x^3}{3}-\beta\frac{x^4}{4}\right]_0^{\frac{C_\beta}{\beta}}$$
$$=\frac{2C_\beta^4}{3\beta^3}-\frac{C_\beta^4}{2\beta^3}=\left(\frac{2}{3}-\frac{1}{2}\right)\frac{1}{\beta}=\frac{1}{6\beta}.$$

Ovvero

$$\beta=\frac{1}{6Var\left[X\right]}$$

e quindi

$$T=\frac{1}{6S^2}$$

è uno stimatore.

Esercizio 3.

1. $$P(Y \leq 1) = P\left(e^X \leq 1\right) = P(X \leq 0) = 0.5.$$

$$0.9 = P\left(e^X \leq \lambda\right) = P(X \leq \log \lambda) = P\left(\frac{X}{2} \leq \frac{\log \lambda}{2}\right) = \Phi\left(\frac{\log \lambda}{2}\right)$$

$\log \lambda = 2q_{0.9} = 2 \cdot 1.29 = 2.58.$

2. Se $y \leq 0$, $F_Y(y) = P(Y \leq y)$ è nulla, quindi $f_Y(y) = 0$ per $y < 0$. Se $y > 0$:

$$F_Y(y) = P(Y \leq y) = P(X \leq \log y) = P\left(\frac{X}{2} \leq \frac{\log y}{2}\right) = \Phi\left(\frac{\log y}{2}\right)$$

$$f_Y(y) = f\left(\frac{\log y}{2}\right)\frac{1}{2y} = \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(\frac{\log y}{2}\right)^2}{2}} \frac{1}{2y} = \frac{1}{2y\sqrt{2\pi}} e^{-\frac{\log^2 y}{8}}.$$