Statistica I - A.A. 2017-2018

Prova scritta - 31 gennaio 2018

Exercise 1 *(pt 20, $\sim$ 4+4+4+4+4) Un produttore di pile sa, sulla base di numerosissime osservazioni sperimentali del passato, che le sue pile utilizzate in un certo tipo di torcia hanno una durata aleatoria T di tipo normale, con parametri $\mu = 10$ ore e $\sigma = 2$ ore.*

1. Calcolare la probabilità che una pila, in quel tipo di torcia, duri meno di 7 ore. Trovare inoltre la durata minima, a meno del 5% dei casi. Raffigurare ciascuna risoluzione con un disegno.

2. In una spedizione archeologica che può durare settimane prima di rientrare in luoghi in cui si possono ricomprare le pile, un esploratore si porta 10 pile. Si supponga che la torcia utilizzi una sola pila per volta. In media, per quante ore sarà utilizzabile la torcia? Con che probabilità non potrà più utilizzare la torcia dopo 95 ore? Spiegare bene il perché della validità del calcolo effettuato.

3. Viene successivamente utilizzato lo stesso tipo di pila su una torcia diversa, il cui consumo non è del tutto chiaro a priori. Per semplicità, si immagini che la deviazione standard sia sempre di due ore. Invece, da un campione sperimentale $t_1, ..., t_{80}$ di numerosità 80, si misura che $\frac{1}{80}\sum_{i=1}^{80} t_i = 8.75$. Al 95%, in che intervallo cadrà la durata media vera per questo tipo di torcia? Trovare ora la durata minima, secondo la stima più pessimistica possibile, spiegando bene il ragionamento.

4. Torniamo alla torcia delle prime due domande. Viene ideata una pila che dovrebbe avere maggiore durata; se ne predispongono 20 esemplari come prototipo e si provano su quella torcia, ottenendo i valori $t_1, ..., t_{20}$ con $\frac{1}{20}\sum_{i=1}^{20} t_i = 10.5$, $\frac{1}{19}\sum_{i=1}^{20} (t_i - 10.5)^2 = 4.5$. Si esegue un test; scrivere le ipotesi e la regione di rifiuto al 97,5% e dire se si può ritenere che la nuova pila duri di più.

5. Non convinta del risultato, l'azienda è disposta ad eseguire un campionamento più massiccio al fine di confermare la validità della nuova pila. Se si ipotizza che la vera durata media della nuova pila sia di 10.5 ore, si assume che la deviazione standard sia 2 e si vuole che il test rifiuti l'ipotesi nulla con probabilità 0.9, quanti esemplari vanno esaminati? Si imposti il problema formulando letteralmente, non con formule prefatte, poi si svolgano i conti. La precisione della formulazione iniziale è considerata la parte principale della risoluzione.

Exercise 2 *(pt 10, $\sim$ 3+4+3) Una v.a. X a valori positivi si dice lognormale se $\log X$ (usiamo qui il logaritmo in base e) è una $N(\mu, \sigma^2)$ ed i numeri μ, σ^2 verranno detti "parametri" della lognormale.*

1. Supponendo $\mu = 3$ e $\sigma = 1$, calcolare il numero λ tale che $P(X < \lambda) = 0.1$.

2. Esprimere la funzione di ripartizione $F_X(x)$ di una v.a. lognormale X di parametri μ e σ tramite la funzione Φ e trovare la densità di una v.a. lognormale X di parametri μ, σ^2.

3. Impostare la ricerca dello stimatore di massima verosimiglianza per il parametro μ (con σ noto), arrivando dove si riesce nel suo calcolo.

1 SOLUZIONI

Esercizio 1.

1)

$$\begin{aligned} P(T<7) &= ... = 0.06 \\ 0.05 &= P(T<\lambda) = ...; \lambda = 6.7. \end{aligned}$$

2) Dette $T_1, ..., T_{10}$ le durate delle 10 pile, che supponiamo indipendenti $N(10, 2^2)$, la durata totale è $S = T_1 + ... + T_{10}$ e vale

$$E[S] = 10 \cdot 10 = 100$$

$$P(S \leq 95) = P\left(\frac{S - 10 \cdot 10}{\sqrt{10} \cdot 2} \leq \frac{95 - 10 \cdot 10}{\sqrt{10} \cdot 2}\right) = P(Z \leq -0.79) = ... = 0.21.$$

Abbiamo usato il fatto che la somma di gaussiane indipendenti è gaussiana; non avremmo potuto applicare in modo affidabile l'approssimazione del TLC, se ce ne fosse stato bisogno, perché la numerosità è troppo bassa.

3) Vista la numerosità elevata e l'ipotesi di varianza nota (che basterebbe da sola), usiamo l'intervallo gaussiano; esso è

$$8.75 \pm \frac{2 \cdot q_{0.975}}{\sqrt{80}} = 8.75 \pm 0.438.$$

In relazione al fatto che temiamo una breve durata, il valore più pessimistico è $8.75 - 0.438$. Rispetto a questo, preso come media, calcoliamo λ tale che

$$0.05 = P(T < \lambda) = ...; \lambda = 5.022.$$

4) Siccome viene specificato all'inizio che dovrebbe durare di più, usiamo un test unilaterale: $\mathcal{H}_0$) nuova media $\leq$10, $\mathcal{H}_1$) nuova media $>$10. Regione di rifiuto:

$$\left\{(t_1, ..., t_{20}) : \frac{\bar{t} - 10}{\sqrt{4.5}} \sqrt{20} > t_{19,0.025}\right\}$$

dove $t_{19,0.05} = 2.093$. Il valore empirico è $\bar{t} = 10.5$. Quindi

$$\frac{\bar{t} - 10}{\sqrt{4.5}} \sqrt{20} = 1.054$$

che non supera 2.093. Quindi il test non è significativo, non possiamo affermare che la nuova pila duri di più.

5) Intanto, riscriviamo la regione di rifiuto con le gaussiane, per poter eseguire il conto richiesto:

$$\left\{(t_1, ..., t_n) : \frac{\bar{t} - 10}{2} \sqrt{n} > q_{0.975}\right\}.$$

Vogliamo che

$$0.9 = P^{10.5,2}\left(\frac{\overline{T}-10}{2}\sqrt{n} > q_{0.975}\right) = ...$$

da cui alla fine si trova $n = 8$.

Esercizio 2.

1)

$$0.1 = P\left(X < \lambda\right) = P\left(\log X < \log \lambda\right) = ...$$

(vedi punto 2).

2)

$$F_X\left(x\right) = P\left(X < x\right) = P\left(\log X < \log x\right) = P\left(\frac{\log X - \mu}{\sigma} < \frac{\log x - \mu}{\sigma}\right) = \Phi\left(\frac{\log x - \mu}{\sigma}\right)$$

$$f_X\left(x\right) = F_X'\left(x\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(\frac{\log x - \mu}{\sigma}\right)^2}{2}} \frac{1}{\sigma x} = \frac{1}{\sqrt{2\pi\sigma^2 x^2}} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}}$$

3)

$$\begin{aligned} \log L\left(\mu, x_1, ..., x_n\right) &= \log \frac{1}{\sqrt{2\pi\sigma^2 x_1^2}} + ... + \log \frac{1}{\sqrt{2\pi\sigma^2 x_1^2}} \\ &\quad - \frac{\left(\log x_1 - \mu\right)^2}{2\sigma^2} - ... - \frac{\left(\log x_n - \mu\right)^2}{2\sigma^2} \end{aligned}$$

ecc.