Matematica e BioStatistica con Applicazioni Informatiche
Esercitazione in aula del 6 dicembre 2018

Quesito 1.

Una macchina è calibrata in modo da fare un taglio in un punto di altezza $\mu_0 = 20$. Se calibrata bene, l'altezza del taglio è distribuita normalmente con media μ_0 deviazione standard $\sigma = 2$.

Ogni tanto (per effetto delle vibrazioni) la macchina si sposta, va quindi fermata e ricalibrata. Idealmente vorremmo fermare la macchina quando la nuova media μ differisce più di 1.3 da μ_0.

1. Misuriamo quindi la posizione del taglio. Chiamiamo $\bar{x}$ la media fatta su un campione di $n = 25$. Calibreremo la macchina se $|\mu_0 - \bar{x}|$ è maggiore di un valore critico c. Quale dev'essere questo valore per non fermare inutilmente ma macchina più del 10% delle volte?
2. Dato il valore critico al punto 1, qual'è (nel peggiore dei casi) la probabilità di non ricalibrare la macchina quando invece necessita di essere ricalibrata?

3 alcune delle quantita calcolate nelle domande precedenti vengono generalmente denominate α, β, δ, e p-valore. Specificare quali.

Risposta 1. Sia $x_{10\%}$ il valore critico. Dev'essere che

$$\Pr\left(|\bar{X} - \mu_0| \geq c_{10\%}\right) = 0.1 = \alpha \qquad \text{dove } \bar{X} \sim N(\mu_0, \sigma^2/\sqrt{n}). \qquad \text{Risposta 3}$$

Standardizzando

$$\Pr\left(|Z| \geq \left|\frac{c_{10\%}}{\sigma/\sqrt{n}}\right|\right) = 2\cdot\Pr\left(Z \leq -\left|\frac{c_{10\%}}{\sigma/\sqrt{n}}\right|\right).$$

$$= 2\cdot\Pr\left(Z \leq -c_{10\%}/0.4\right) = 0.1$$

Quindi $c_{10\%}$ = `- 0.4 * norm.ppf(0.05)` $\cong$ `0.66` Risposta 1

2. I casi più sfavorevoli sono quando $\mu = \mu_0 \pm \delta$. Dove $\delta = 1.3$ Risposta 3

Per simmetria possiamo considerare solo $\mu = \mu_0 + \delta$. Dobbiamo calcolare

$$\Pr\left(|\bar{X} - \mu_0| \leq c_{10\%}\right) = \Pr\left(-c_{10\%} \leq \bar{X} - \mu_0 \leq c_{10\%}\right) \qquad \text{dove } \bar{X} \sim N(\mu,\ \sigma^2/\sqrt{n}).$$

$$= \Pr\left(-c_{10\%} - \delta \leq \bar{X} - \mu \leq c_{10\%} - \delta\right)$$

$$= \Pr\left(-\frac{c_{10\%} + \delta}{\sigma/\sqrt{n}} \leq Z \leq \frac{c_{10\%} - \delta}{\sigma/\sqrt{n}}\right)$$

Standardizzando

$$= \Pr\left(Z \leq \frac{c_{10\%} - \delta}{\sigma/\sqrt{n}}\right) - \Pr\left(Z \leq -\frac{c_{10\%} + \delta}{\sigma/\sqrt{n}}\right).$$

La coda sinistra è trascurabile

$$\cong \Pr\left(Z \leq \frac{c_{10\%} - \delta}{\sigma/\sqrt{n}}\right).$$

= `norm.cdf( (` $c_{10\%}$ `- 1.3 ) / 0.4 ) )` Risposta 2

$$\cong 0.0548 \cong \beta \qquad \text{Risposta 3}$$

Formulario: se $X \sim B(\mathtt{n},\mathtt{p})$ allora $E(X) = np$
se $X \sim NB(\mathtt{n},\mathtt{p})$ allora $E(X) = n(1-p)/p$

Si assuma noto il valore delle seguenti funzioni della libreria `scipy.stats` di `Python`
`binom.pmf(k, n, p)` $= \Pr\big(X = \mathtt{k}\big)$ dove $X \sim B(\mathtt{n},\mathtt{p})$
`binom.cdf(k, n, p)` $= \Pr\big(X \leq \mathtt{k}\big)$ dove $X \sim B(\mathtt{n},\mathtt{p})$
`bimom.ppf(q, n, p)` $= \mathtt{k}$ dove $\mathtt{k}$ è tale che $\Pr\big(X \leq \mathtt{k}\big) \cong \mathtt{q}$ per $X \sim B(\mathtt{n},\mathtt{p})$

`nbinom.xxx(k, n, p)`, è l'analogo per $X \sim NB(\mathtt{n},\mathtt{p})$.

`norm.xxx(z)`, è l'analogo per $Z \sim N(0,1)$.