Matematica e BioStatistica con Applicazioni Informatiche
Esercitazione in aula del 10 gennaio 2018

Quesito 1. Della v.a. discreta X conosciamo la distribizione di probabilità

$$\Pr(X=4)=\frac{1}{2} \qquad \Pr(X=5)=\frac{1}{2}$$

Della v.a. discreta Y conosciamo la distribuzione condizionata a X

$$\Pr(Y=3 \mid X=4)=\frac{1}{2} \qquad \Pr(Y=3 \mid X=5)=\frac{1}{3}$$

$$\Pr(Y=2 \mid X=4)=\frac{1}{2} \qquad \Pr(Y=2 \mid X=5)=\frac{2}{3}$$

Calcolare la distribuzione di probablità di Y

Esprimere i numeri razionali come frazioni.

Risposta

$$\left.\begin{aligned}\Pr(Y=3) &= \Pr(Y=3 \mid X=4)\cdot\Pr(X=4) + \Pr(Y=3 \mid X=5)\cdot\Pr(X=5) = \frac{5}{12}\\ \Pr(Y=2) &= 1-\Pr(Y=3) = \frac{7}{12}\end{aligned}\right\} \text{Risposta}$$

Quesito 2. Assume the null hypothesis is true and denote by P the random variable that gives the p-value you would get if you run a test.

1. What is the probability that $\Pr(P<0.05)$?
2. If we run the tests 8 times (independently), what is the probability of incorrectly rejecting at least once the null hypotheses with a significance $\alpha = 5\%$?
3. If we run the tests 8 times (independently), how small do we have to make the cutoff (α above) to lower to 5% the probability of incorrectly rejecting at least once the null hypotheses?

Risposta

$\Pr(P<0.05) = 0.05$ Risposta 1

$1-\left(1-\frac{1}{20}\right)^8 = 0.3366$ Risposta 2

$$1-\left(1-\frac{x}{100}\right)^8 = \frac{1}{20}, \quad \text{risolvendo}$$

$$x = 100\left(1-\sqrt[8]{\frac{19}{20}}\right)$$

$= 0.6391\%$ Risposta 3

Quesito 3. A manufacturer claims that the mean lifetime of a lightbulb is on average at least 10 thousand hours with a standard deviation of 0.3. In a sample of 9 lightbulbs, it was found that they only last 9.5 thousand hours on average. The sample standard deviation is 0.6 thousand hours. Can we reject the manufacturer's claim? Answer the following questions:

1. H_0? H_1?
2. What test is required?
3. What is the value of the statistic?
4. What is the p-value?

Risposta

$\mu_0 = 10$

$\mathrm{H}_0: \ \mu = \mu_0, \qquad \mathrm{H}_1: \ \mu < \mu_0$ Risposta 1

We use a one tail t-test (lower tail) Risposta 2

$n = 9$ sample size

$s = 0.6$ sample standard deviation

$\bar{x} = 9.5$ sample mean

$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = -2.5$ value of the t-test statistic Risposta 3

$n - 1 = 8$ degrees of freedom

$P\big(T_{n-1} < t\big)$ = `t.cdf(-2.5, 8) = 0.0185` p-value Risposta 4

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

$$T = \frac{\bar{X} - \bar{Y}}{S \cdot \sqrt{1/n_x + 1/n_y}} \quad \text{dove } S^2 = \frac{n_x - 1}{n_x + n_y - 2} \cdot S_x^2 + \frac{n_y - 1}{n_x + n_y - 2} \cdot S_y^2 \quad \text{ha distribuzione } t(n_x + n_y - 2)$$

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)` $=$ `k` dove `k` è tale che $\Pr\big(X \leq \mathtt{k}\big) \cong \mathtt{q}$ per $X \sim B(\mathtt{n}, \mathtt{p})$
`nbinom.xxx(...)`, è l'analogo per $X \sim NB(\mathtt{n}, \mathtt{p})$.
`norm.xxx(...)`, è l'analogo per $Z \sim N(0, 1)$.
`t.xxx(..., ν)`, è l'analogo per $T \sim t(\nu)$.