Variabili aleatorie discrete

Distribuzione binomiale: $X \sim B(n,p)$

$$p(x) = \begin{cases} \begin{pmatrix} n \\ x \end{pmatrix} p^x (1-p)^{n-x} & \text{se} \ \ x = 0,1,\dots,n \\ 0 & \text{altrimenti} \end{cases}$$

Distribuzione di Poisson: $X \sim \mathrm{Pois}(\lambda)$

$$p(x) = \begin{cases} e^{-\lambda} \frac{\lambda^x}{x!} & \text{se} \ \ x = 0,1,\dots,n \\ 0 & \text{altrimenti} \end{cases}$$

Distribuzione ipergeometrica:
$n = b + r$ oggetti, b di tipo 1, r di tipo 2

$$p(x) = \begin{cases} \dfrac{\begin{pmatrix} b \\ x \end{pmatrix}\begin{pmatrix} r \\ n-x \end{pmatrix}}{\begin{pmatrix} b+r \\ n \end{pmatrix}} & \text{se} \ \ x = 0,1,\dots,n \\ 0 & \text{altrimenti} \end{cases}$$

Distribuzione geometrica: $X \sim \mathrm{Geom}(p)$

$$p(x) = \begin{cases} p(1-p)^{x-1} & \text{se} \ \ x = 1,\dots,n \\ 0 & \text{altrimenti} \end{cases}$$

Distribuzione multinomiale: $X \sim B(n, q_1, q_2, \dots, q_n)$

$$p(x) = \begin{cases} \frac{n!}{\omega_1!\omega_2!\dots\omega_n!} \, q_1^{\omega_1} q_2^{\omega_2} \dots q_n^{\omega_n} & \text{se} \ \ \omega_k = 0,1,\dots,n \ \ \text{e} \ \ \sum_{k=1}^m \omega_k = n \\ \\ 0 & \text{altrimenti} \end{cases}$$

Variabili aleatorie continue

Distribuzione uniforme in $[a,b]$: $X \sim U([a,b])$

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{se} \ \ x \in [a,b] \\ \\ 0 & \text{altrimenti} \end{cases}$$

$$F_X(t) = \begin{cases} 0 & \text{se} \ \ t < a \\ \frac{t-a}{b-a} & \text{se} \ \ t \in [a,b] \\ 1 & \text{se} \ \ t > b \end{cases}$$

Legge normale standard: $X \sim N(0,1)$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Legge normale: $X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Distribuzione esponenziale: $X \sim \mathrm{Exp}(\lambda)$

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se} \ \ x > 0 \\ \\ 0 & \text{se} \ \ x \leq 0 \end{cases}$$

$$F_X(t) = \begin{cases} 0 & \text{se} \ \ t \leq 0 \\ 1 - e^{-\lambda t} & \text{se} \ \ t > 0 \end{cases}$$

Disuguaglianze

Disuguaglianza di Markov:

$$P(X \geq \alpha) \leq \frac{E[X]}{\alpha} \quad \forall \alpha \in \mathbb{R}$$

Disuguaglianza di Čebišëv:

$$P(|X - E[X]| \geq \eta) \leq \frac{\mathrm{Var}(X)}{\eta^2} \quad \forall \eta \in \mathbb{R}$$

Intervalli di confidenza

Media (varianza nota):

$$I = \left[\overline{X}_n - \frac{\sigma}{\sqrt{n}}\phi_{1-\frac{\alpha}{2}}, \overline{X}_n + \frac{\sigma}{\sqrt{n}}\phi_{1-\frac{\alpha}{2}}\right]$$

Media (varianza non nota):

$$I = \left[\overline{X}_n - \frac{\bar{S}_n}{\sqrt{n}}t_{1-\frac{\alpha}{2}}, \overline{X}_n + \frac{\bar{S}_n}{\sqrt{n}}t_{1-\frac{\alpha}{2}}\right]$$

Proporzione:

$$I = \left[\overline{X}_n - \frac{\sqrt{\overline{X}_n(1-\overline{X}_n)}}{\sqrt{n}}\phi_{1-\frac{\alpha}{2}}, \overline{X}_n + \frac{\sqrt{\overline{X}_n(1-\overline{X}_n)}}{\sqrt{n}}\phi_{\frac{\alpha}{2}}\right]$$

Varianza:

$$I = \left[\frac{\bar{S}_n^2(n-1)}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}, \frac{\bar{S}_n^2(n-1)}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right]$$

Correlazione e regressione lineare

Correlazione lineare:

$$\rho = \frac{\mathrm{Cov}(x,y)}{\sqrt{\mathrm{Var}(x)\mathrm{Var}(y)}} = \frac{\bar{\sigma}_{xy}}{\bar{\sigma}_x\bar{\sigma}_y}$$

Coefficiente di determinazione:

$$R = \frac{\mathrm{Cov}^2(x,y)}{\mathrm{Var}(x)\mathrm{Var}(y)} = \frac{\bar{\sigma}_{xy}^2}{\bar{\sigma}_x^2\bar{\sigma}_y^2}$$

Retta dei minimi quadrati:

$$y = b_0 + b_1 x$$

$$b_0 = \bar{y} - \frac{\bar{\sigma}_{xy}}{\bar{\sigma}_x^2}, \quad b_1 = \frac{\bar{\sigma}_{xy}}{\bar{\sigma}_x^2}$$

Test d'ipotesi

Z-test- Z_n

Zona rigetto:

$$C_R : \begin{cases} |Z_n| > \phi_{1-\frac{\alpha}{2}} \text{ bilatero} \\ Z_n < \phi_\alpha \text{ unilatero a sx} \\ Z_n > \phi_{1-\alpha} \text{ unilatero a dx} \end{cases}$$

p-value:

$$\text{p-value} = \begin{cases} 2(1-\Phi(|Z_n|) \text{ bilatero} \\ \Phi(Z_n) \text{ unilatero a sx} \\ 1-\Phi(Z_n) \text{ unilatero a dx} \end{cases}$$

- Media con varianza nota: $Z_n = \frac{\overline{X}_n - \mu_0}{\sigma}\sqrt{n}$
- Proporzione: $Z_n = \frac{\overline{X}_n - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}$
- Media di coppie di poplazioni X, Y con varianza nota: $Z_n = \frac{\overline{X}_n - \mu_X - (\overline{Y}_m - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$

t-test- T_n

Zona rigetto:

$$C_R : \begin{cases} |T_n| > t_{1-\frac{\alpha}{2}} \text{ bilatero} \\ T_n < t_\alpha \text{ unilatero a sx} \\ T_n > t_{1-\alpha} \text{ unilatero a dx} \end{cases}$$

p-value:

$$\text{p-value} = \begin{cases} 2(1-F_T(|T_n|) \text{ bilatero} \\ F_T(T_n) \text{ unilatero a sx} \\ 1-F_T(T_n) \text{ unilatero a dx} \end{cases}$$

- Media con varianza non nota: $T_n = \frac{\overline{X}_n - \mu_0}{\bar{S}_n}\sqrt{n}$
- Media di popolazioni accoppiate $X, Y : T_n = \frac{\overline{D}_n}{\bar{S}_n}$, con $D_i = X_i - Y_i$.

test del χ^2- W_n

Zona rigetto:

$$C_R : \begin{cases} W_n > \chi^2_{1-\frac{\alpha}{2}}(n-1) \text{ o } W_n < \chi^2_{\frac{\alpha}{2}}(n-1) \text{ bilatero} \\ W_n < \chi^2_\alpha(n-1) \text{ unilatero a sx} \\ W_n > \chi^2_{1-\alpha}(n-1) \text{ unilatero a dx} \end{cases}$$

- Varianza: $W_n = \frac{\bar{S}_n(n-1)}{\sigma_0^2}$
- Multinomiale m-dimensionale – test di Pearson: $W_n = n\sum_{k=1}^m \frac{(\bar{p}_k - p_k^0)^2}{p_k^0} \sim \chi^2(m-1)$.