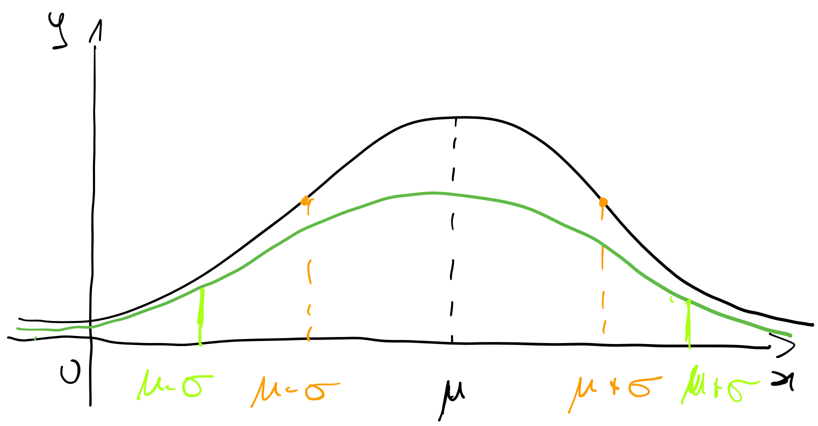
Distribuzione di Gauss o Normale, $N(\mu, \sigma)$

Definizione

Una variabile aleatoria X è detta distribuita secondo una normale (o gaussiana) di parametri $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$ ($X \sim N(\mu, \sigma)$) se ha densità di probabilità

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



Proprietà:

1) $f_X(x)$ è simmetrica rispetto alla retta $x=\mu$
2) ha un massimo assoluto in $\left(\mu, \frac{1}{\sigma\sqrt{2\pi}}\right)$
3) ha asintoto orizzontale $y=0$
4) ha due flessi in

$$\left(\mu-\sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right), \left(\mu+\sigma, \frac{1}{\sigma\sqrt{2\pi e}}\right)$$

5) $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$

6) all'aumentare di σ il max diminuisce

7) se $\mu = 0$ è simmetrica rispetto all'asse y

8) $E[X] = \mu$, $V[X] = \sigma^2$

La funzione di ripartizione associata è

$$F_X(t) = P[X \le t] = \int_{-\infty}^{t} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}\,dx$$

Definizione

Distribuzione normale standardizzata

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

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

$$z = \frac{x-\mu}{\sigma}$$

$$x = \sigma z + \mu$$

$$dx = \sigma\,dz$$

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,dx =$$

$$= \frac{1}{\cancel{\sigma}\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}z^2} \cancel{\sigma}\, dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}z^2} dz$$

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

Vettori aleatori

Definizione

È detto vettore aleatorio, o variabile aleatoria multidimensionale, un vettore le cui componenti sono variabili aleatorie

$$X \equiv (X_1, X_2, \dots, X_n)$$

$$\boxed{P[X_1 \cap X_2 \cap X_3 \dots] \equiv P[X_1, X_2, \dots, X_n]}$$

Se abbiamo le variabili aleatorie X, Y discrete

- $\sum_i \sum_j P[X=i, Y=j] = 1$, $\sum_i P[X=i] = 1$

  $\sum_j P[Y=j] = 1$
- I valori $P[X=i, Y=j]$ formano la densità

discreta congiunta del vettore (X, Y)

- Le densità $P[X=i]$, $P[Y=j]$ sono dette densità marginali

Oss: è sempre possibile ricavare le prob. marginali dalle prob. congiunte, ma non è vero il viceversa.

Definizione (Funzione di ripartizione congiunta)

Dato un vettore aleatorio (X, Y), è detta funzione di ripartizione congiunta la funzione

$$F_{X,Y}(x, y) = P[X \leq x, Y \leq y]$$

$$\Big(P\big[(X \leq x) \cap (Y \leq y) \big] \Big)$$

Definizione (Funzione di ripartizione marginale)

Data una funzione di ripartizione congiunta $F_{X,Y}$, è definita la funzione di ripartizione marginale come

$$F_X(x) = P[X \leq x] = P[X \leq x, Y < +\infty] =$$

$$= \lim_{y \to +\infty} F_{X,Y}(x, y)$$

Proprietà

$$\forall y \in \mathbb{R} \quad \lim_{x \to -\infty} F_{XY}(x,y) = 0, \quad \lim_{x \to +\infty} F_{XY}(x,y) = F_Y(y)$$

$$\forall x \in \mathbb{R} \quad \lim_{y \to -\infty} F_{XY}(x,y) = 0, \quad \lim_{y \to +\infty} F_{XY}(x,y) = F_X(x)$$

- Densità congiunta (caso discreto)

Def.
Si dice densità congiunta del vettore aleatorio discreto (X,Y), e si indica con p_{XY}, l'insieme dei

valori $$p_{XY}(x_i, y_j) = P[X = x_i, Y = y_j]$$

Definizione (densità marginali)

$$p_X(x_i) = \sum_{j\,(y_j)} p_{XY}(x_i, y_j)$$

$$p_Y(y_j) = \sum_{i\,(x_i)} p_{XY}(x_i, y_j)$$

- Densità congiunta per il caso continuo

Def.
Dato un vettore aleatorio continuo (X,Y), è

detta funzione di densità di probabilità la funzione $f_{XY}(x, y)$ per la quale, per ogni scelta di $a, b, c, d \in \mathbb{R}$,

$$P[a < X < b,\ c < Y < d] =$$

$$= \int_a^b \left(\int_c^d f_{XY}(x, y)\, dy \right) dx =$$

$$= \int_c^d \left(\int_a^b f_{XY}(x, y)\, dx \right) dy$$

$$\int_a^b dx \int_c^d dy\ f_{XY}(x, y)$$

Si dimostra che

$$F_{XY}(s, t) = P[X < s,\ Y < t] = \int_{-\infty}^{s} dx \int_{-\infty}^{t} dy\ f_{XY}(x, y)$$

$$\boxed{f_X(x) = \frac{dF_X(x)}{dx}}$$

$$f_{XY}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y)$$

Definizione (Funzioni densità marginali)

Sono dette funzioni densità marginali

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\,dy \quad , \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\,dx$$

$$\left[\int_{-\infty}^{+\infty} f_X(x)\,dx = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy\, f_{XY}(x,y) = 1 \right]$$

Definizione

Date n variabili aleatorie X_i, con $i=1,\dots,n$, esse si dicono indipendenti se, per ogni scelta degli intervalli A_i, si ha che

$$P[X_1 \in A_1, X_2 \in A_2 \dots, X_n \in A_n] =$$

$$= P[X_1 \in A_1] \cdot P[X_2 \in A_2] \cdot \dots \cdot P[X_n \in A_n]$$

Si dim. che (nel caso di variabili indipendenti)

$$P_{XY}(x_i, y_j) = P_X(x_i)\, P_Y(y_j)$$

$$f_{XY}(x_i, y_i) = f_X(x_i)\, f_Y(y_i)$$

$$F_{XY}(x, y) = F_X(x)\, F_Y(y)$$

Teorema

Dato un vettore aleatorio (X, Y) (discreto o continuo)

la somma $U = X + Y$ delle sue componenti

ha densità

$$p_U(u) = \sum_k P_{XY}(k, u-k) \quad \text{caso discreto}$$

$$f_U(u) = \int_{-\infty}^{+\infty} f_{XY}(t, u-t)\, dt \quad \text{caso continuo}$$

Definizione

Sia (X, Y) un vettore aleatorio e $g(X, Y)$

una funzione, $E[g(X, Y)]$, se esiste, è dato

da

$$E[g(X, Y)] = \sum_i \sum_j g(x_i, y_j)\, P_{XY}(x_i, y_j) \quad \text{c.d.}$$

$$E[g(X, Y)] = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy\; g(x, y)\, f_{XY}(x, y) \quad \text{c.c.}$$

$-\infty \quad -\infty$

Proprietà

Per ogni coppia di variabili aleatorie X, Y, il valore atteso della somma è uguale alla somma dei valori attesi

$$E[X+Y] = E[X] + E[Y]$$

Dim. (esercizio)

Definizione (covarianza)

$$Cov[X,Y] = E\left[(X - E[X])(Y - E[Y])\right]$$

Proprietà (esercizio)

$$Cov[X,Y] = E[XY] - E[X]E[Y]$$

Proprietà:

X, Y, V, W variabili aleatorie

$\alpha, \beta \in \mathbb{R}$

1) $Cov[X, \alpha] = 0$

2) $Cov[X,Y] = Cov[Y,X]$

3) $\mathrm{Cov}[X, X] = \mathrm{Var}[X]$

4) $\mathrm{Cov}[\alpha X, \beta Y] = \alpha\beta\,\mathrm{Cov}[X, Y]$

5) $\mathrm{Cov}[X + Y, V + W] =$

$= \mathrm{Cov}[X, V] + \mathrm{Cov}[X, W] + \mathrm{Cov}[Y, V] + \mathrm{Cov}[Y, W]$

Definizione (matrice di covarianza)

Dato $X = (X_1, X_2, \dots, X_n)$, la matrice di covarianza (C_X) è la matrice di elementi

$$(C_X)_{ij} = \mathrm{Cov}[X_i, X_j]$$

$$C_X = \begin{bmatrix} \mathrm{Cov}[X_1, X_1] & \mathrm{Cov}[X_1, X_2] & \dots & \mathrm{Cov}[X_1, X_n] \\ \mathrm{Cov}[X_2, X_1] & \mathrm{Cov}[X_2, X_2] & \dots & \\ & & & \end{bmatrix}$$

Definizione (coefficiente di correlazione)

Date due variabili aleatorie X, Y, è detto coefficiente di correlazione (ρ_{XY})

$$\rho_{XY} = \frac{Cov[X,Y]}{\sqrt{Var[X]\, Var[Y]}}$$

Proprietà

1) $-1 \leq \rho_{XY} \leq 1$

2) $|\rho_{XY}| = 1$ se unicamente tra X e Y esiste un valore deterministico lineare $Y = aX + b$ con a e $b \in \mathbb{R}$ che hanno lo stesso segno

Definizione

Si dicono incorrelate due variabili aleatorie X, Y per cui $Cov[X,Y] = 0$