CHAPTER 3

Distribuzioni di probabilità multidimensionali

3.1 Probabilità per variabili aleatorie in più dimensioni

Quando un evento è identificato da un vettore $\underline{x}$, si parla di di distribuzioni di probabilità multidimensionali o **joint pdf (probabiltà congiunta)**.

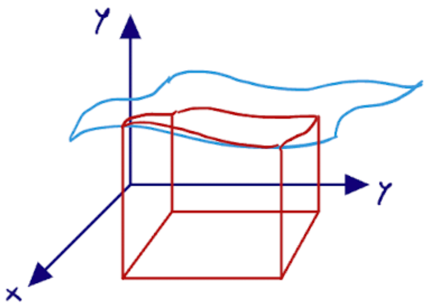
$$pdf(\underline{x}) : A \subseteq \mathbb{R}^n \to \mathbb{R}^+$$

e la cdf per estensione della definizione mono-dimensionale è:

$$cdf(\underline{x}) = \int_A pdf(\underline{x}) \cdot d\underline{x}$$

la probabilià di un evento per una variabile continua $\underline{x}$ è definita su una regione di spazio $A \subseteq \mathbb{R}^n$:

$$P(\underline{x} \in A) = \int_A pdf\, f(\underline{x}) \cdot d\underline{x}$$



Tale definizione di probabilità verifica gli assiomi di Kolgomorov.

3.2 Distribuzione di probabilità marginale

Data una distribuzione di probabilità multidimensionale pdf$(x_1, \cdots x_n)$ si definisce **distribuzione di probabilità marginale**:

$$f_{x_i}(x_1, \cdots, x_i, \cdots, x_n) = \int pdf(x_1, \cdots, x_i, \cdots, x_n) dx_1 \cdots dx_n \tag{3.1}$$

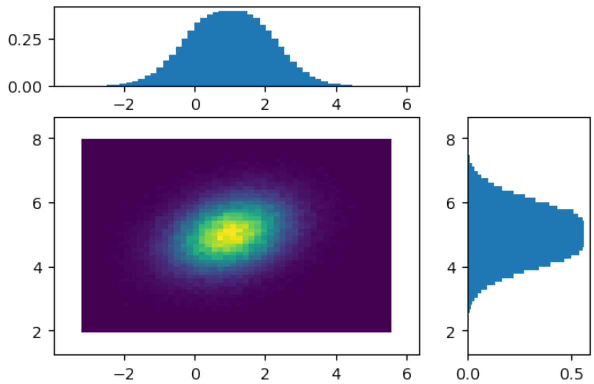


Figure 3.1: Distribuzione di probabilità marginale per una pdf(x,y)

Esempio

Nel caso bi-dimensionale si ha che le rispettive distribuzioni marginali di una joint pdf(x_1, x_2) sono date da:

$$f_x(y) = \int pdf(x, y) dy \qquad f_y(x) = \int pdf(x, y) dx \tag{3.2}$$

3.3 Distribuzione di probabilità condizionata

Per semplicità consideriamo una distribuzione di probabilità rispetto a due variabili aleatorie x ed y e le rispettive distribuzioni marginali f_x e f_y. Vogliamo determinare la probabilità che $P(x|y = y_0)$ o $P(y|x = x_0)$. Consideriamo due eventi $A, B \subset \Omega$ disgiunti. Possiamo identificare le pdf come:

- $P(A \cap B) \rightarrow$ joint pdf

- P(A) $\rightarrow$ pdf marginale
- $P(A|B)$ $\rightarrow$ pdf condizionata

Definiamo L'evento A = $\{y \mid x \in [x_0, x_0 + dx] \;,\; y \in \mathbb{R}\}$ e l'evento B = $\{x \mid y \in [y_0, y_0 + dx] \;,\; x \in \mathbb{R}\}$. Le probabilità associate ai singoli eventi sono P(A) = f_x e P(B) = f_y. Mentre la probabilità congiunta è $P(A \cap B) = \int pdf(x_0, y_0)dxdy$.

Di conseguenza possiamo scrivere la probabilità condizionata come:

$$pdf(x|y = y_0) = P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{pdf(x, y_0)}{fx(y_0)} \tag{3.3}$$

e analogamente:

$$pdf(y|x = x_0) = \frac{pdf(x_0, y)}{f_y(x_0)} \tag{3.4}$$

3.4 Valore di aspettazione di una joint pdf

Per una distribuzione di probabilità congiunta abbiamo che il valore di aspettazione è definito da un vettore:

$$E[\underline{x}] = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} = \begin{bmatrix} \int\int x \cdot pdf(x, y)dxdy \\ \int\int y \cdot pdf(x, y)dxdy \end{bmatrix} \tag{3.5}$$

3.5 Varianza di una pdf multidimensionale

Nel caso della Varianza, le componenti di una variabile aleatoria multidimensionale $\underline{x}$ possono essere legate tra loro, ovvero avere una relazione nel modo in cui variano, tale legame prende il nome di **covarianza**. Per due componenti x_i e x_j con $i \neq j$ si ha che:

$$\sigma_{i,j}^2 = E[(x_i - \mu_i)(x_j - \mu_j)] = E[x_i x_j] - E[x_i]E[x_j] \tag{3.6}$$

mentre

$$\sigma_{i,i}^2 = E[(x_i - \mu_i)^2] \tag{3.7}$$

di conseguenza la varianza di una pdf($\underline{x}$) multidimensionale è rappresentata da una matrice che prende il nome di **matrice di covarianza** ed ha dimensione $n \times n$ nel caso $\underline{x}$ sia un vettore dimensione n.

$$V[\underline{x}] = \begin{bmatrix} V[x_1] & \cdots\cdots & Cov[x_1, x_n] \\ \vdots & \ddots & \vdots \\ Cov[x_n, x_1] & \cdots\cdots & V[x_n] \end{bmatrix} \tag{3.8}$$

Nel caso in cui le componenti della variabile aleatoria $\underline{x}$ siano indipendenti tra loro si ha che la covarianza è nulla e dunque la (3.3) diventa una matrice diagonale.

$$V[\underline{x}] = \begin{bmatrix} V[x_1] & \cdots\cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots\cdots & V[x_n] \end{bmatrix} \tag{3.9}$$

La covarianza gode delle seguenti proprietà:

- Avere $Cov[x_i, x_j] = 0$ non implica necessariamente che le due variabili aleatorie siano statisticamente indipendenti.
- Se due variabili aleatorie sono statisticamente indipendenti $\Rightarrow Cov[x_i, x_j] = 0$
- Se due variabili aleatorie sono linearmente dipendenti $\Rightarrow Cov[x_i, x_j] = 0$
- La matrice di covarianza è simmetrica $\Rightarrow$ è diagonalizzabile.

3.6 Correlazione

Gli elementi V_{ij} della matrice di covarianza misurano il grado di correlazione tra le variabili x_i e x_j. Dato che ogni variabili mostra una varianza finita e

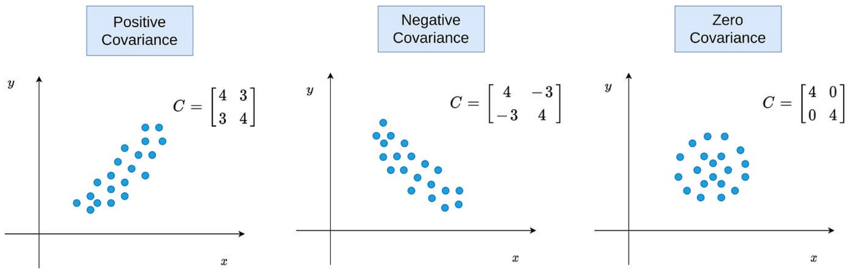


Figure 3.2: Esempi di come la covarianza si presenta in una distribuzione

positiva è utile confrontare la covarianza rispetto alle loro rispettive varianze, per farlo si introduce il coefficiente di correlazione:

$$\rho_{ij} = \frac{Cov[x_i, x_j]}{\sigma_i \sigma_j} \tag{3.10}$$

tale grandezza $\rho_{ij} \in [-1, 1]$.

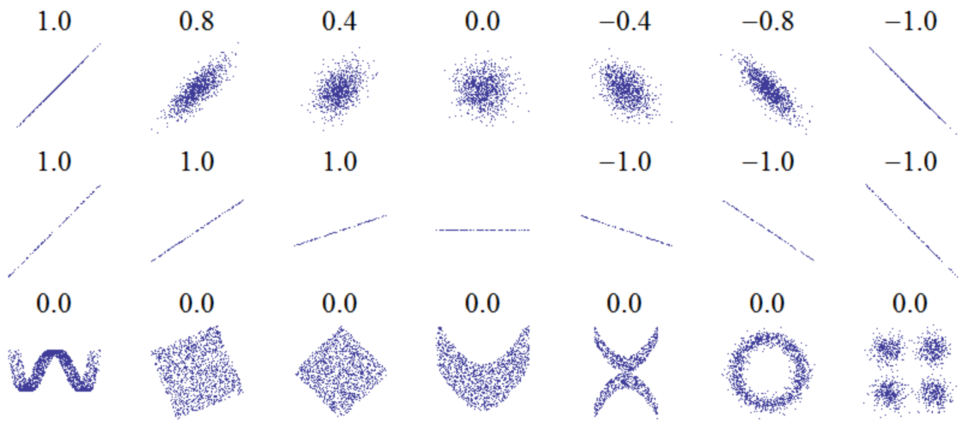


Figure 3.3: Indice di correlazione di Pearson per un campione di misure.

3.7 Variabili Statisticamente indipendenti

Applichiamo la definizione di probabilità congiunta data dall'equazione (3.3), poichè per ipotizziamo che $P(x_2|x_1)$ e x_2 non dipenda da x_1 avremo che:

$$pdf(x_2|x_1) = \frac{pdf(x_1, x_2)}{f_{x_1}} \Rightarrow pdf((x_1, x_2) = f_{x_1} \cdot pdf(x_2|x_1)$$

si ha che

$$f_{x_2}(x_1) = \int pdf(x_1, x_2)dx_1 = pdf(x_2|x_1) \int f_{x_1} dx_1 = pdf(x_2|x_1)$$

dunque si ha che:

$$pdf(x_1, x_2) = f_{x_1} \cdot f_{x_2}$$

Definizione d'indipendenza

Due variabili aleatorie x_1 ed x_2 descritte da una joint pdf(x_1, x_2) si definiscono **indipendenti** $\iff$ la joint pdf(x_1, x_2) $= pdf(x_1) \cdot pdf(x_2)$ dove $pdf(x_1)$ e $pdf(x_2)$ coincidono con le marginali.

Teorema

Due variabili aleatorie x_1 ed x_2 descritte da una joint pdf(x_1, x_2) sono indipendenti $\Rightarrow$ la covarianza può essere scritta come:

$$E[(x_1 - \mu_{x_1})(x_2 - \mu_{x_2})] = E[(x_1 - \mu_{x_1})] \cdot E[(x_2 - \mu_{x_2})] \tag{3.11}$$

3.8 Cambiamento di variabili per una pdf multidimensionale

La matrice della varianza (3.6) è diagonalizzabile, questo vuol dire che esiste un cambio di base che la diagonalizza, in fisica è equivalente ad avere un cambio di variabile. Sperimentalmente raccolto un campione di misure esiste sempre un cambio di variabile tale per cui la matrice di covarianza è diagonalizzabile. Anche se con un cambio di variabili le grandezze diventano decorrelate $Cov[x_i, x_j] = 0$ questo non vuol dire che siano statisticamente indipendenti.

Analogamente al caso mono-dimensionale per il cambio di variabili si ha che date delle funzioni:

$$\begin{cases} x = u(\alpha, \beta) \\ y = w(\alpha, \beta) \end{cases}$$

la join pdf nelle nuove coordinate sarà data da:

$$pdf(\alpha, \beta) = pdf(x, y) \cdot |detJ| \tag{3.12}$$

dove J è la matrice Jacobiana associata alla trasformazione.

3.8.1 Propagazione degli errori

Ipotizziamo di avere un insieme di N misure $\{x_i\}_i^N$ e usiamo tali valori come componenti di un vettore $\underline{x} \in \mathbb{R}^N$, descritto da una pdf($\underline{x}$) di cui conosciamo $\underline{\mu}$ e matrice di covarianza $V[\underline{x}]$, vogliamo calcolare $y = u(\underline{x})$. Per Farlo approssimiamo u($\underline{\text{x}}$) con un sviluppo di Taylor al primo ordine in un intorno di $\underline{\mu}$:

$$u(\underline{x}) \approx u(\underline{\mu}) + \nabla u\big|_{x=\mu} \cdot (\underline{x} - \underline{\mu})$$

Dunque i momenti della $pdf(\underline{y})$ sono:

- E[y] = E[u($\underline{\mu}$)] + $\sum_{i=1}^n \frac{\partial}{\partial x_i} u(\underline{\mu}) E[(x_i - \mu_i)] = u(\underline{\mu})$
- $\sigma_y^2 = E[y^2] - E[y]^2 = E[(\underline{x} - \underline{\mu})^T H(u(\underline{x}))\big|_{x=\mu} (\underline{x} - \underline{\mu})] =$

  $= \sum_{i,j=1}^n \frac{\partial u(\underline{x})}{\partial x_i} \cdot \frac{\partial u(\underline{x})}{\partial x_j}\Big|_{x=\mu} \cdot V_{ij}$

Per un caso bidimensionale l'incertezza su z = f(x,y) è data da:

$$\sigma_z^2 = \Big(\frac{\partial f(x, y)}{\partial x}\Big)^2\Big|_{x=\mu} \sigma_x^2 + \Big(\frac{\partial f(x, y)}{\partial y}\Big)^2\Big|_{x=\mu} \sigma_y^2 + 2\frac{\partial f(x, y)}{\partial x} \cdot \frac{\partial f(x, y)}{\partial y}\Big|_{x=\mu} Cov[x, y] \tag{3.13}$$

3.8.2 Decorrelazione delle variabili

Consideriamo un vettore $\underline{x}$ le cui componenti sono variabili aleatorie dipendenti tra loro. Poichè la matrice di covarianza è simmetrica possiamo definire una matrice U che la diagonalizza $D[\underline{y}] = U^T V[\underline{x}]U$, tale matrice definisce un cambio di coordinate $\underline{y} = U\underline{x}$ dove le componenti di $\underline{y}$ risultano essere decorrelate tra loro.

3.9 Distribuzione di Gauss Multidimensionale

L'espressione di una Gaussiana in più dimensioni è data da:

$$f(\underline{x} \mid \underline{\mu}, V) = \frac{1}{(2\pi)^{\frac{n}{2}} detV^{\frac{1}{2}}} \cdot \exp\left[-\frac{1}{2} < (\underline{x} - \underline{\mu}), V^{-1}(\underline{x} - \underline{\mu}) > \right] \quad (3.14)$$

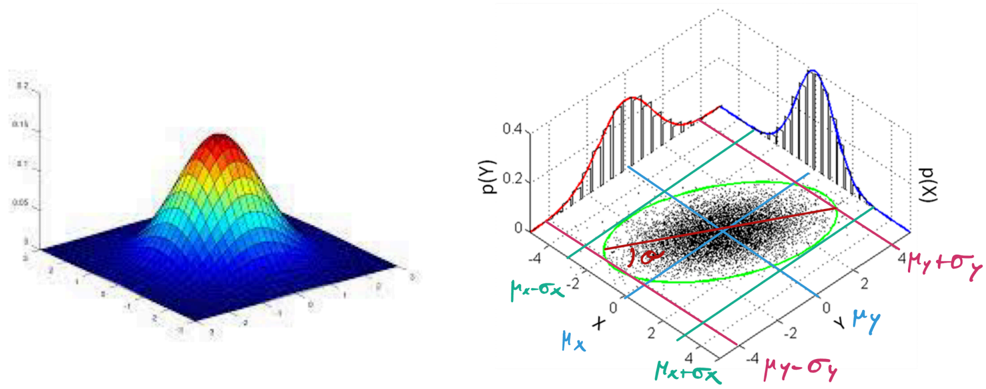


Figure 3.4: Gaussiana in due variabili (x,y) e le sue distribuzioni marginali

La gaussiana in due dimensioni in generale ha un profilo ellittico per 1σ. L'angolo d'inclinazione del semiasse maggiore è legato al coefficiente di correlazione:

$$\theta = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2} \quad (3.15)$$

Se esplicitiamo l'equazione (3.11) per due variabili si ha:

$$f(x, y \mid \mu_x, \mu_y, V) =$$

$$= \frac{1}{(2\pi)\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\Big\{ -\frac{1}{2(1-\rho^2)}\Big[\Big(\frac{x-\mu_x}{\sigma_x}\Big)^2 - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \Big(\frac{y-\mu_y}{\sigma_y}\Big)^2\Big]\Big\}$$

Se il coefficiente di correlazione di Pearson $\rho = 0$ possiamo riscrivere l'equazione precedente come:

$$f(x, y \mid \mu_x, \mu_y, V) = \frac{1}{(2\pi)\sigma_x\sigma_y} \exp\Big\{ -\frac{1}{2}\Big[\Big(\frac{x-\mu_x}{\sigma_x}\Big)^2 + \Big(\frac{y-\mu_y}{\sigma_y}\Big)^2\Big]\Big\} =$$

$$= G(x, \mu_x, \sigma_x) \cdot G(y, \mu_y, \sigma_y)$$

In generale è solo vero che due variabili statisticamente indipendenti sono anche decorrelate, nel caso Gaussiano vale anche il viceversa, ovvero se due variabili sono decorrelate allora sono anche statisticamente indipendenti.