Media e momenti di una variabile aleatoria
Comportamento globale di una variabile aleatoria

▶ Si definisce valor medio o media della variabile aleatoria X il numero reale

$$\mathbb{E}\left[X\right] = \begin{cases} \displaystyle\sum_{x \in \mathcal{A}_X} x P_X(x) & \text{se } \; X \; \text{ è una v.a. discreta} \\ \displaystyle\int_{-\infty}^{+\infty} x f_X(x)\, dx & \text{se } \; X \; \text{ è una v.a. continua} \end{cases}$$

e viene anche denotato con μ_X, omettendo eventualmente il pedice che indica la variabile aleatoria al quale il valor medio si riferisce.

Variabile aleatoria Bernoulliana, $X \sim \mathcal{B}(1, p)$

$$\mathbb{E}\left[X\right] = 1 \times p + 0 \times q = p$$

che coincide, quindi con la probabilità che la v.a. assuma il valore 1. medskip

Variabile aleatoria esponenziale, $X \sim \mathcal{E}x(\lambda)$

$$\mathbb{E}\left[X\right] = \int_{-\infty}^{+\infty} x\, \lambda e^{-\lambda x}\, u(x)\, dx = \frac{1}{\lambda}$$

che coincide col reciproco del parametro della v.a. esponenziale.

Variabile aleatoria gaussiana, $X \sim \mathcal{N}(\mu, \sigma^2)$

$$\mathbb{E}\left[X\right] = \mathbb{E}\left[\sigma X_0 + \mu\right] = \sigma \mathbb{E}\left[X_0\right] + \mu$$

Inoltre, risultando

$$\mathbb{E}\left[X_0\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-x^2/2}\, dx = 0$$

si ha

$$\mathbb{E}\left[X\right] = \mu \ .$$

Il parametro di locazione μ di una v.a. gaussiana coincide quindi con la sua media.

► La precedente operazione si basa sul fatto che gli operatori di somma e integrale sono lineari. Dimostrare che $\mathbb{E}[aX] = a\mathbb{E}[X]$ e $\mathbb{E}[a] = a$.

Variabile aleatoria Binomiale, $X \sim \mathcal{B}(n, p)$

$$\mathbb{E}\left[X\right] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}\left[X_i\right] = np \ .$$

Teorema fondamentale per il calcolo della media
La relazione seguente risulta particolarmente utile per il calcolo della media di una funzione di variabile aleatoria

Considerata una v.a. X, la media di una sua funzione è calcolabile come

$$\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x)\, dx$$

se la v.a. è continua, ovvero come

$$\mathbb{E}[g(X)] = \sum_{x \in \mathcal{A}} g(x) P_X(x)$$

se la v.a. è discreta, con alfabeto $\mathcal{A}$.

► Pertanto, per calcolare la media di una funzione di variabili aleatorie non è necessario il calcolo della PDF (o della PMF) della v.a. $g(X)$, ma è sufficiente la PDF (o la PMF) della v.a. X.

► Il momento del secondo ordine $\mathbb{E}\left[X^2\right]$ è anche denominato valore quadratico medio della v.a. X, mentre la sua radice quadrata

$$X_{rms} \triangleq \sqrt{\mathbb{E}\left[X^2\right]}$$

è detta valore *rms* (root mean square) o valore efficace.

► I momenti intorno alla media μ_X di una v.a. X, cioè le quantità:

$$\mathbb{E}\left[(X - \mu_X)^k\right]$$

sono detti momenti centrali. In particolare, il momento centrale

$$\sigma_X^2 \equiv \mathrm{Var}[X] \triangleq \mathbb{E}\left[(X - \mu_X)^2\right]$$

è la varianza della v.a. X e la sua radice quadrata

$$\sigma_X \triangleq \sqrt{\mathbb{E}\left[(X - \mu_X)^2\right]}$$

è la deviazione standard. Nel seguito, se dal contesto è chiara la v.a. a cui ci si riferisce, il pedice X sarà omesso.

▶ *Varianza di una trasformazione affine di una v.a.*

$$\begin{aligned}\mathrm{Var}[aX+b] &= \mathbb{E}\left[\Big((aX+b)-\mathbb{E}\left[aX+b\right]\Big)^2\right]\\ &= \mathbb{E}\left[\Big(aX-a\mathbb{E}\left[X\right]\Big)^2\right]\\ &= a^2\sigma_X^2\end{aligned}$$

Pertanto una traslazione non ha effetto sulla varianza, ma solo sulla media, mentre un cambiamento di scala con fattore a dà luogo ad una variazione secondo un fattore a^2 della varianza.

Momenti di una variabile aleatoria Bernoulliana.
Per una v.a. Bernoulliana, il momento di ordine k può essere calcolato come segue:

$$\mathbb{E}\left[X^k\right] = 1^k \times p + 0^k \times q = p$$

Dunque tutti i momenti di una v.a. Bernoulliana sono uguali a p. La varianza vale

$$\sigma_X^2 = p - p^2 = pq$$

come è facile verificare.

Momenti di una variabile aleatoria esponenziale.
Consideriamo una v.a. esponenziale di parametro unitario, $X_0 \sim \mathcal{E}x(1)$. I suoi momenti sono dati da

$$\mathbb{E}\left[X_0^k\right] = \int_0^{+\infty} x^k e^{-x} dx$$

e si possono valutare ricorsivamente. Infatti,

$$\mathbb{E}\left[X_0^k\right] = \int_0^{+\infty} x^k \mathrm{e}^{-x}\, dx = -\int_0^{+\infty} x^k \frac{d}{dx}\left[\mathrm{e}^{-x}\right]\, dx$$

da cui, integrando per parti si ricava il legame di tipo ricorsivo

$$\mathbb{E}\left[X_0^k\right] = k\mathbb{E}\left[X_0^{k-1}\right]$$

Iterando tale relazione, tenendo presente che $\mathbb{E}\left[X_0^0\right] = 1$, si ricava l'espressione esplicita

$$\mathbb{E}\left[X_0^k\right] = k!$$

CDF congiunta di due variabili aleatorie

Calcolo della probabilità congiunta degli eventi $\{X \le x\}$ *e* $\{Y \le y\}$.

▶ Notiamo che se due v.a sono indipendenti allora

$$P(\{X \le x\} \cap \{Y \le y\}) = P(\{X \le x\})\, P(\{Y \le y\}) = F_X(x)\, F_Y(y)$$

quindi la probabilità congiunta dei due eventi si ricava semplicemente come prodotto delle CDF delle due variabili aleatorie X e Y. Le CDF delle singole v.a. $F_X(x)$ ed $F_Y(y)$ sono dette CDF marginali.

▶ Se le due v.a non sono indipendenti, è necessaria una conoscenza più approfondita dell'esperimento, esprimibile mediante la funzione di distribuzione cumulativa congiunta

$$F_{XY} : (x, y) \in \mathbb{R}^2 \;\longrightarrow\; F_{XY}(x, y) = P(\{X \le x\} \cap \{Y \le y\}) \;.$$

Per la CDF congiunta di una coppia di variabili aleatorie valgono proprietà simili a quelle della CDF di una singola variabile aleatoria.

PDF e PMF congiunta di due variabili aleatorie

- Per v.a. discrete, si definisce la PMF congiunta

$$P_{XY} : (x, y) \in \mathcal{A}_X \times \mathcal{A}_Y \;\longrightarrow\; P_{XY}(x, y) = P(\{X = x\} \cap \{Y = y\})$$

dove $\mathcal{A}_X$ e $\mathcal{A}_Y$ denotano gli alfabeti delle v.a. X e Y, rispettivamente.

- Per v.a. continue, si definisce la PDF congiunta delle due v.a. X e Y definita come derivata seconda mista della CDF congiunta, cioè

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

Anche per la PMF e PDF congiunta di una coppia di variabili aleatorie valgono proprietà simili a quelle di una singola variabile aleatoria.

Distribuzioni e densità condizionate

▶ *CDF condizionata ad un evento*
La CDF di una v.a. (discreta o continua) X condizionata ad un evento B (avente probabilità non nulla), è la funzione

$$F_X(x|B) = P(\{X \leq x\}|B) = \frac{P(\{X \leq x\} \cap B)}{P(B)} \qquad P(B) \neq 0$$

▶ *PMF condizionata ad un evento*
Se X è una v.a. discreta, la PMF condizionata a B è la funzione

$$P_X(\cdot|B) : x \in \mathcal{A}_X \longrightarrow P_X(x|B) = \frac{P(\{X = x\} \cap B)}{P(B)}$$

▶ *PDF condizionata ad un evento*
Se X è una v.a. continua, la PDF condizionata è la derivata della CDF condizionata, cioè

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

▶ *Esempio: Radar ad impulsi*

In un sistema radar, gli impulsi riflessi hanno un'ampiezza R, tuttavia sullo schermo sono visualizzati solo gli impulsi la cui ampiezza supera una certa soglia x_0. Siamo interessati a determinare la CDF e la PDF degli impulsi visualizzati.

La CDF e la PDF degli impulsi visualizzati coincide con la CDF e la PDF dell'ampiezza R condizionate all'evento $\{R > x_0\}$. Si ha quindi

$$\begin{aligned} F_R(x|\{R > x_0\}) &= \frac{P(\{R \le x\} \cap \{R > x_0\})}{P(\{R > x_0\})} \\ &= \frac{F_R(x) - F_R(x_0)}{1 - F_R(x_0)} \mathrm{u}(x - x_0)\,. \end{aligned}$$

In particolare, se R è di tipo Rayleigh con parametro σ^2, la sua CDF è

$$F_R(x) = 1 - \exp\left[-\frac{x^2}{2\sigma^2}\right] \quad x \ge 0$$

Di conseguenza, si ottiene

$$F_R(x|R > x_0) = \left(1 - \exp\left[-\frac{x^2 - x_0^2}{2\sigma^2}\right]\right) \mathrm{u}(x - x_0)$$

La PDF degli impulsi visualizzati si ottiene effettuando la derivata della CDF condizionata. Si ha pertanto

$$f_R(x|\{R > x_0\}) = \frac{x}{\sigma^2} \exp\left[-\frac{x^2 - x_0^2}{2\sigma^2}\right] \mathrm{u}(x - x_0)$$

▶ *PMF condizionata*

Nel caso che X ed Y siano v.a. discrete, la PMF condizionata della v.a. X data la v.a. Y si ottiene particolarizzando la definizione generale al caso $\{Y = y\}$

$$\begin{aligned} P_X(x|\{Y=y\}) &= P(\{X=x\}|\{Y=y\}) \\ &= \frac{P(\{X=x\}\cap\{Y=y\})}{P(\{Y=y\})} = \frac{P_{XY}(x,y)}{P_Y(y)} \end{aligned}$$

La notazione più comune per la PMF condizionata è $P_{X|Y}(x|y)$ invece di $P_X(x|\{Y=y\})$.

▶ *PDF condizionata*

Nel caso di v.a. continue, in analogia a quanto esposto per le v.a. discrete, si definisce la PDF condizionata della v.a. X data la v.a. Y la funzione

$$f_{X|Y}(x|y): \quad x \in \mathbb{R} \longrightarrow \frac{f_{XY}(x,y)}{f_Y(y)} \qquad f_Y(y) \neq 0$$

Leggi della probabilità per la PDF

▶ Dalla definizione di PDF condizionata segue la *legge della probabilità composta* per le PDF, cioè

$$f_{XY}(x,y) = f_Y(y) f_{X|Y}(x|y) = f_X(x) f_{Y|X}(y|x)$$

conseguentemente risulta anche

$$f_{X|Y}(x|y) = f_{Y|X}(y|x) \frac{f_X(x)}{f_Y(x)}$$

che esprime la **legge di Bayes** per le pdf.

▶ Dalla relazione $f_{XY}(x,y) = f_X(x) f_{Y|X}(y|x)$, integrando entrambe i membri rispetto alla variabile x, si ottiene

$$\int_{-\infty}^{\infty} f_{XY}(x,y) dx = \int_{-\infty}^{\infty} f_{X|Y}(x|y) f_Y(y) dx = f_X(x)$$

da cui si evince che la pdf marginale si ottiene integrando la pdf congiunta rispetto all'altra variabile indipendente.

▶ Integrando entrambe i membri l'equazione per la probabilità composta rispetto alla variabile y, si ottiene invece

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x|y) f_Y(y) dy$$

che esprime la **legge della probabilità totale** per le PDF.

Definizioni e proprietà della PDF congiunta e della PDF condizionata di due v.a.

PDF congiunta	$f_{XY}(x,y) = \dfrac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$
PDF condizionata	$f_{X\|Y}(x\|y) = \dfrac{f_{XY}(x,y)}{f_Y(y)}$
PDF marginali	$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy; \quad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx$
Legge della probabilità composta	$f_{XY}(x,y) = f_Y(y) f_{X\|Y}(x\|y) = f_X(x) f_{Y\|X}(y\|x)$
Legge di Bayes	$f_{X\|Y}(x\|y) = f_{Y\|X}(y\|x) \dfrac{f_X(x)}{f_Y(x)}$
Legge della probabilità totale	$f_X(x) = \int_{-\infty}^{\infty} f_{X\|Y}(x\|y) f_Y(y) dy$

Variabili aleatorie indipendenti

▶ L'indipendenza tra n v.a. si ha quando la loro CDF congiunta è fattorizzabile nel prodotto delle singole CDF marginali, se cioè:

$$F_{X_1 \cdots X_n}(x_1, x_2, \ldots x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \cdots F_{X_n}(x_n)$$

Equivalentemente, se le v.a. sono continue si ricava che per v.a. indipendenti la PDF congiunta è il prodotto delle PDF marginali, cioè:

$$f_{X_1 \cdots X_n}(x_1, x_2, \ldots x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n)$$

Si noti che la relazione precedente è valida anche per v.a. discrete, a patto di sostituire le PMF alle PDF.

Momenti congiunti di due variabili aleatorie

▶ Nel caso in cui si considerano due o più variabili aleatorie è utile definire dei momenti che indichino il grado di influenza reciproca tra le variabili aleatorie. Considerate due v.a. X ed Y, si definisce momento congiunto di ordine $k = m + r$ la media

$$\mathbb{E}\left[X^m Y^r\right]$$

▶ Particolare importanza rivestono i momenti centrali e non centrali del primo e del secondo ordine, corrispondenti alle coppie di indici $(m, r) = (10, 01, 20, 02, 11)$.
I momenti del primo ordine: $\mu_X = \mathbb{E}\left[X\right]$ e $\mu_Y = \mathbb{E}\left[Y\right]$ sono le medie delle singole variabili aleatorie mentre i momenti del secondo ordine sono i valori quadratici medi delle singole variabili aleatorie e la correlazione tra le due variabili

$$r_{XY} = \text{corr}\left[X, Y\right] = \mathbb{E}\left[XY\right].$$

Analogamente, se si considerano i momenti centrali del secondo ordine, si riottengono le varianze della singole v.a. e la loro covarianza

$$c_{XY} \equiv \text{cov}[X, Y] \triangleq \mathbb{E}\left[(X - \mu_X)(Y - \mu_Y)\right]$$

▶ Proprietà

Sviluppando il prodotto che compare nella definizione di covarianza si ha

$$\begin{aligned}\mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right] &= \mathbb{E}\left[XY-\mu_Y X-\mu_X Y+\mu_X\mu_Y\right] \\ &= \mathbb{E}\left[XY\right]-\mu_Y\mathbb{E}\left[X\right]-\mu_X\mathbb{E}\left[Y\right]+\mu_X\mu_Y \\ &= \mathbb{E}\left[XY\right]-\mu_X\mu_Y\ ,\end{aligned}$$

pertanto correlazione e covarianza sono legate tra loro dalla relazione $r_{XY} = c_{XY} + \mu_X\mu_Y$

Incorrelazione e indipendenza

Due variabili aleatorie si dicono incorrelate se la loro covarianza è nulla o, equivalentemente, se la correlazione è pari al prodotto delle medie. Nel caso in cui la correlazione sia nulla, le variabili si dicono ortogonali.

N.B. *Se almeno una delle due v.a. incorrelate ha media nulla la correlazione è zero e le variabili, oltre che incorrelate, sono anche ortogonali,*

► L'incorrelazione è una condizione più debole dell'indipendenza statistica e non va confusa con quest'ultima, nel senso che sussiste la seguente implicazione.

Proposizione

Date due v.a. X ed Y, statisticamente indipendenti esse sono anche incorrelate.

Prova: *supponiamo, per fissare le idee che le due v.a. siano di tipo continuo. In tal caso, essendo statisticamente indipendenti, la loro PDF congiunta è il prodotto delle due PDF marginali; conseguentemente risulta*

$$\begin{aligned}\mathbb{E}[XY] &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy f_{XY}(x,y)\,dx\,dy = \\ &= \int_{-\infty}^{+\infty} x f_X(x)\,dx \int_{-\infty}^{+\infty} y f_Y(y)\,dy = \mathbb{E}[X]\mathbb{E}[Y]\end{aligned}$$

Esempio: variabili aleatorie incorrelate, ma dipendenti
Si consideri la coppia di v.a. $X = \cos\Theta$ e $Y = \sin\Theta$, con $\Theta \sim U(-\pi, \pi)$. Avendosi:

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[\cos\Theta] = \int_{-\pi}^{+\pi} \frac{1}{2\pi}\cos\alpha\, d\alpha &= 0 \\ \mathbb{E}[Y] &= \mathbb{E}[\sin\Theta] = \int_{-\pi}^{+\pi} \frac{1}{2\pi}\sin\alpha\, d\alpha &= 0 \end{aligned}$$

le v.a. sono entrambe a media nulla; inoltre, poiché

$$\mathbb{E}[XY] = \mathbb{E}[\cos\Theta\sin\Theta] = \int_{-\pi}^{+\pi} \frac{1}{2\pi}\cos\alpha\sin\alpha\, d\alpha = 0$$

le variabili sono anche ortogonali e incorrelate. Tuttavia esse non sono statisticamente indipendenti in quanto i valori della X e della Y sono, a meno del segno, funzionalmente legate, con legame funzionale implicitamente definito dall'equazione:

$$X^2 + Y^2 = 1$$

In altri termini le due v.a. assumono, con probabilità 1, valori sulla circonferenza del cerchio di raggio unitario e centro l'origine.

Coppie di v.a. congiuntamente Gaussiane

- Una coppia di v.a. gaussiane si ottiene da una trasformazione affine non singolare di v.a. gaussiane standard indipendenti, cioè

$$\begin{aligned} X_1 &= a_{11}X_{01} + a_{12}X_{02} + \mu_1 \\ X_2 &= a_{21}X_{01} + a_{22}X_{02} + \mu_2 \end{aligned}$$

con $X_{01},\ X_{02}$ v.a. gaussiane standard indipendenti. Indicando con $\mathbf{X} = (X_1, X2)$ il vettore delle v.a. e con $\mathbf{x} = (x_1, x_2)$ il vettore delle variabili indipendenti, la PDF congiunta ha la forma generale

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\mathbf{C}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$$

dove la matrice $\mathbf{C}$ è detta matrice di covarianza e $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$ è il vettore delle medie. Il generico termine a_{ij} della matrice di covarianza è la covarianza

$$a_{ij} = \mathbb{E}[(X_1 - \mu_1)(X_2 - \mu_2)]$$

per cui si ha

$$\mathbf{C} = \begin{bmatrix} \mathbb{E}[(X_1-\mu_1)^2] & \mathbb{E}[(X_1-\mu_1)(X_2-\mu_2)] \\ \mathbb{E}[(X_1-\mu_1)(X_2-\mu_2)] & \mathbb{E}[(X_2-\mu_2)^2] \end{bmatrix}$$

I termini sulla diagonale principale coincidono con le varianze delle singole v.a. mentre i termini sulla diagonale secondaria possono essere espressi attraverso il coefficiente di correlazione, definito come

$$\rho = \frac{\mathbb{E}[(X_1 - \mu_1)(X_2 - \mu_2)]}{\sigma_1 \sigma_2} \ .$$

Calcolando il determinante e l'inversa della matrice $\mathbf{C}$ che, come è facile verificare, valgono

$$\det(\mathbf{C}) = \sigma_1^2 \sigma_2^2 (1 - \rho^2) \qquad\qquad \mathbf{C}^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} \dfrac{1}{\sigma_1^2} & -\dfrac{\rho}{\sigma_1 \sigma_2} \\ -\dfrac{\rho}{\sigma_1 \sigma_2} & \dfrac{1}{\sigma_2^2} \end{bmatrix}$$

è possibile esplicitare la forma generale della PDF congiunta in termini di medie, varianze e coefficiente di correlazione. Infatti si ha

$$f_{X_1 X_2}(x_1, x_2) \ = \ \frac{1}{2\pi \sigma_1 \sigma_2 (1-\rho^2)^{1/2}} \exp\left[-\frac{(x_1 - \mu_1)^2}{2\sigma_1^2 (1-\rho^2)} + \frac{\rho (x_1 - \mu_1)(x_2 - \mu_2)}{(1-\rho^2)\sigma_1 \sigma_2} - \frac{(x_2 - \mu_2)^2}{2\sigma_2^2 (1-\rho^2)} \right]$$

Si noti infine che, ponendo $\rho = 0$ nell'espressione della PDF congiunta di due v.a. gaussiane, si ottiene

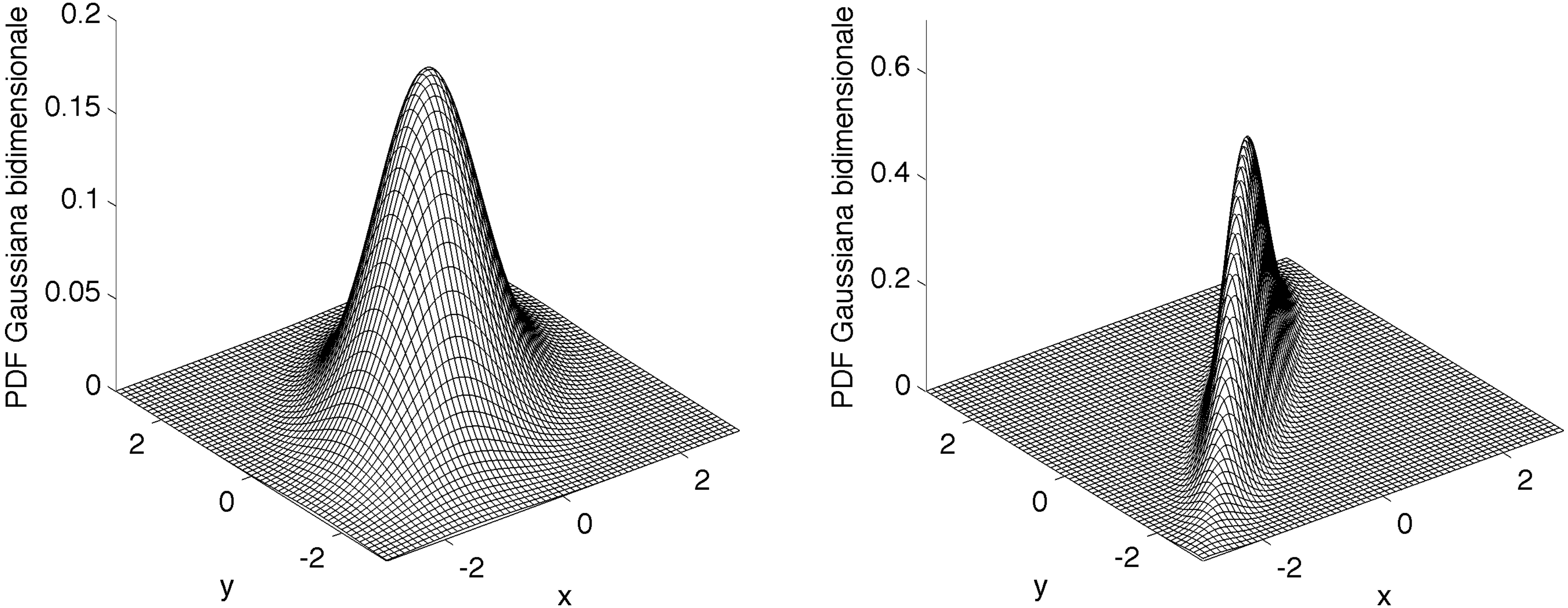
$$f_{X_1 X_2}(x_1, x_2) \ = \ \frac{1}{\sqrt{2\pi \sigma_1^2}} \exp\left[-\frac{1}{2} \frac{(x_1 - \mu_1)^2}{\sigma_1^2} \right] \times \frac{1}{\sqrt{2\pi \sigma_2^2}} \exp\left[-\frac{1}{2} \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right]$$

Pertanto, se due v.a. congiuntamente gaussiane sono incorrelate sono anche indipendenti.

In figura è mostrata la PDF congiunta di due v.a. gaussiane. È immediato notare che quando ρ si tende ad 1 la superficie si concentra intorno alla retta $y = x$; in tal caso, infatti, deve risultare $\mathbb{E}[(X_1 - \mu_1)(X_2 - \mu_2)] = \sigma_1\sigma_2$ che si ottiene se le variabili $(X_1 - \mu_1)$ e $(X_2 - \mu_2)$ sono proporzionali. Infatti, ponendo $(X_1 - \mu_1) = \alpha(X_2 - \mu_2)$ si ricava

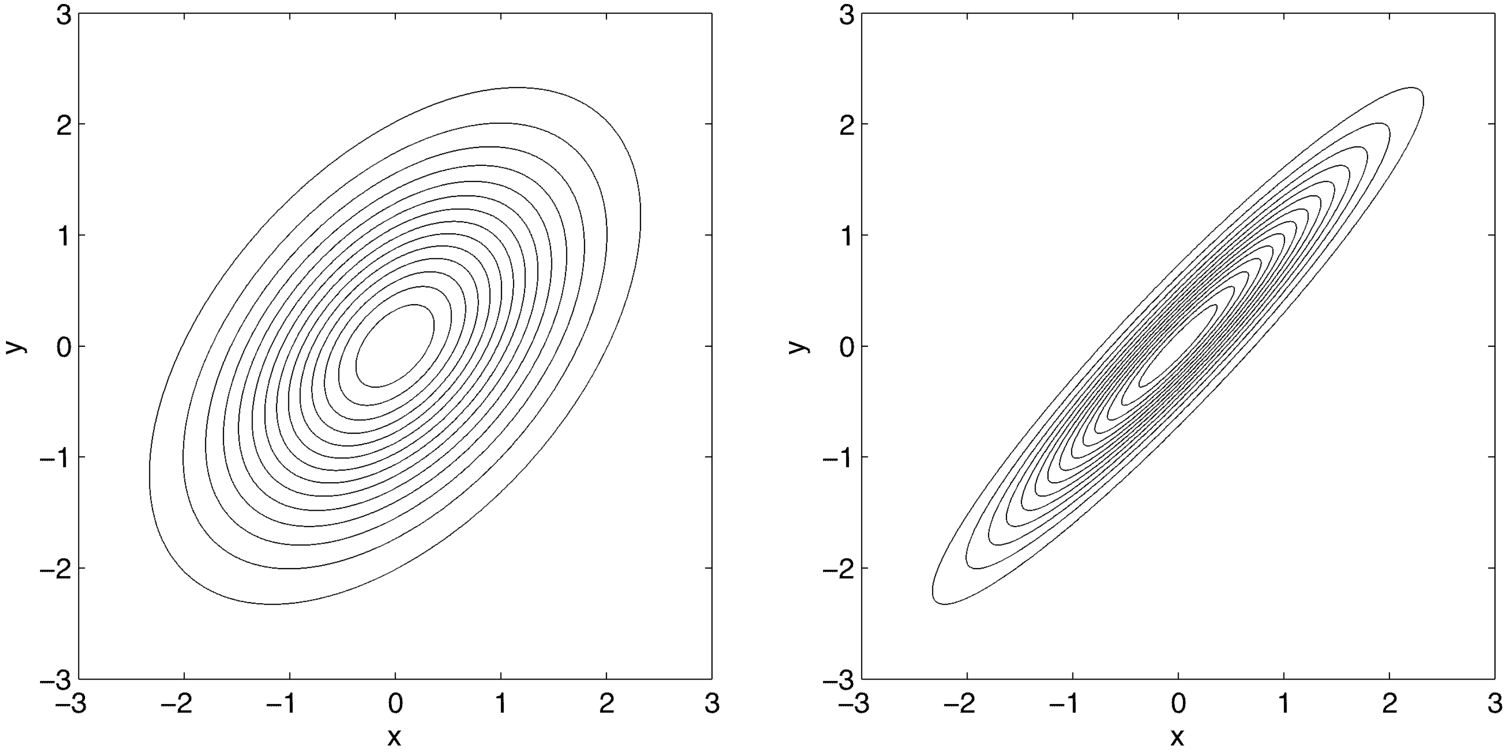
$$\rho = \frac{\mathbb{E}[\alpha(X_2 - \mu_2)(X_2 - \mu_2)]}{\alpha\sigma_2^2} = 1$$

Funzione di densità di probabilità gaussiana bidimensionale
($\sigma_X = \sigma_Y = 1$, $\rho = 0.5$, $\rho = 0.95$.)



▶ Le curve di livello sono delle ellissi che si ottengono ponendo $(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu}) = cost.$

Curve di livello della funzione di densità di probabilità gaussiana bidimensionale
$\sigma_X = \sigma_Y = 1,\ \rho = 0.5,\ \rho = 0.95.$



Stime basate su sequenze di osservazioni

Siamo abituati ad osservare che l'accuratezza con cui si misurano grandezze non deterministiche dipende dal numero di osservazioni disponibili. La teoria della stima, di cui si presenteranno nel seguito alcuni elementi, si occupa di studiare quantitativamente tale affermazione.

▶ *Media campionaria*
Iniziamo con lo studiare il comportamento asintotico della media aritmetica di n v.a., detta anche media campionaria, definita come

$$S_n \triangleq \frac{1}{n}\sum_{i=1}^{n} X_i \ .$$

La media campionaria è a sua volta una variabile aleatoria. La media statistica si ricava semplicemente utilizzando la proprietà di linearità; si ha infatti

$$\mathbb{E}\left[S_n\right] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[X_i\right] = \mu$$

dove μ è la media statistica delle variabili aletorie X_i. Si dice che S_n è uno stimatore della media statistica μ. Risultando $\mathbb{E}\left[S_n\right] = \mu$ la media dello stimatore coincide con la grandezza da stimare, per cui lo stimatore è *non polarizzato*.

▶ *Varianza della media campionaria*
La varianza dello stimatore si calcola come

$$\begin{aligned}
\mathrm{Var}[S_n] &= E\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[X_i\right]\right)^2\right] = \\
&= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}\left[X_i\right])\right)^2\right] = \\
&= \frac{1}{n^2}\mathbb{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - \mathbb{E}\left[X_i\right])(X_j - \mathbb{E}\left[X_j\right])\right] = \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}\left[(X_i - \mathbb{E}\left[X_i\right])(X_j - \mathbb{E}\left[X_j\right])\right]
\end{aligned}$$

dove si può riconoscere che il generico addendo dell'ultima sommatoria è la covarianza delle variabili aleatorie X_i ed X_j. In definitiva si ottiene

$$\mathrm{Var}[S_n] = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Cov}[X_i, X_j]\ .$$

Nel caso in cui le variabili aleatorie siano indipendenti, risulta

$$\mathrm{Cov}\,[X_i, X_j] = \begin{cases} \mathrm{Var}[X_i] = \sigma^2_{X_i} & \text{se} \quad i = j \\ 0 & \text{se} \quad i \neq j \end{cases}$$

per cui la varianza della somma di variabili aleatorie incorrelate è data dalla somma delle varianze delle singole variabili aleatorie,

$$\mathrm{Var}[S_n] = \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2_{X_i} \ .$$

Se infine le variabile, oltre che incorrelate, hanno tutte la stessa varianza, se cioè $\sigma^2_{X_i} = \sigma^2_X, \quad \forall i \in \{1, 2, \dots n\}$ si ottiene

$$\mathrm{Var}[S_n] = \frac{\sigma^2}{n}$$

Quanto stabilito per la media campionaria è riassunto dal seguente asserto.

Proposizione
Date n variabili aleatorie i.i.d., la media campionaria è uno stimatore non polarizzato della media statistica delle variabili e la sua varianza è uguale alla varianza delle variabili scalata per un fattore $1/n$, cioè

$$\mathbb{E}\,[S_n] = \mu \qquad\qquad Var[S_n] = \frac{\sigma^2}{n}$$

▶ La *Legge dei grandi numeri* stabilisce che la media campionaria di una successione di variabili aleatorie i.i.d. converge asintoticamente alla loro media statistica. In particolare, poichè la varianza dello stimatore tende asintoticamente a zero, la convergenza va intesa in *media quadratica*, cioè nel senso che

$$\lim_{n} \ \mathbb{E}\left[(S_n-\mu)^2\right]=0$$

N.B. La convergenza in media quadratica non comporta che vadano a zero le possibili determinazioni dello scarto assoluto $|S_n-\mu|$. In altre parole, è possibile avere un numero finito di determinazioni non nulle su un numero infinito di prove. Alla luce di tale definizione, e sulla scorta di quanto stabilito, resta dimostrato la seguente legge.

Legge dei grandi numeri

Data una successione di variabili aleatorie iid X_n, $n \in \mathbb{N}$ la sua media campionaria S_n converge, in media quadratica, al valore comune μ della media statistica delle singole variabili aleatorie

$$\lim_{n} \frac{1}{n}\sum_{i=1}^{n} X(i)=\mu \qquad \text{m.s.}$$

▶ La legge dei grandi numeri sostanzia, da un punto di vista analitico, l'affermazione che la media statistica coincide con la media aritmetica di un numero molto grande di osservazioni. Dunque, pur essendo la media campionaria una variabile aleatoria, al crescere del numero delle osservazioni essa tende a diventare una quantità deterministica. Inoltre, l'errore quadratico medio che si commette nell'approssimare la media statistica μ con la media campionaria è inversamente proporzionale al numero n di osservazioni.

▶ La legge dei grandi numeri è lo strumento base per la misura di grandezze probabilistiche; va però individuata una successione di variabili aleatorie la cui media statistica coincida con la grandezza che si intende stimare.

Esempio: stima della probabilità di un evento A
Definendo la v.a. indicatore dell'evento

$$I_A = \begin{cases} 1 & \text{se l'evento } A \text{ risulta verificato} \\ 0 & \text{altrimenti} \end{cases}$$

si ottiene $\mathbb{E}[I_A] = 1 \times P(A) + 0 \times (1 - P(A) = P(A))$ per cui, dalla legge dei grandi numeri, risulta che la media campionaria dell'indicatore tende alla probabilità dell'evento al divergere del numero di osservazioni. D'altra parte, la media campionaria dell'indicatore coincide con il numero di volte N_A in cui l'evento A si verifica; da cui si ottiene il seguente asserto.

Convergenza della frequenza relativa
La frequenza relativa di un evento A in n prove indipendenti, effettuate tutte nelle medesime condizioni, converge, in media quadratica, alla probabilità $P(A)$ dell'evento A

$$\lim_n \frac{N_A(n)}{n} = P(A) \qquad \text{m.s.}$$

In pratica non si opera al limite, ma con un numero finito, ancorchè grande di osservazioni. È possibile valutare il numero di osservazioni necessarie per misurare la $P(A)$ con un prefissato errore *rms* ovvero di calcolare l'errore *rms* che si commette effettuando la misura con un prefissato numero di osservazioni, come illustrato dal seguente esempio.

Esempio: misura della probabilità di un evento
L'utilizzo della variabile aleatoria I_A – indicatrice dell'evento A – ha consentito di tradurre il problema della stima di $P(A)$ nel calcolo di una media campionaria ed, in ultima analisi, in un problema di conteggio. La varianza dello stimatore è semplicemente data da $\mathrm{Var}(I_A) = \mathrm{Var}(X_A)/n$ dove la variabile aleatoria X_A è bernoulliana, con probabilità di successo uguale a $P(A)$. La varianza di X_A è data da $P(A)(1 - P(A))$ da cui si ottiene che la deviazione standard della stima è data da

$$\sigma = \sqrt{\frac{P(A)[1 - P(A)]}{n}} \leq \frac{1}{2\sqrt{n}}$$

ove la maggiorazione si giustifica immediatamente se si tiene presente che

$$\max_{0 \leq P(A) \leq 1} P(A)[1 - P(A)] = \frac{1}{4}$$

Si noti che, a differenza di σ, la sua maggiorazione è indipendente dal valore della $P(A)$ e coincide con σ per $P(A) = 0.5$. Valutiamo ora l'errore medio che si commette nella misura della $P(A)$ a partire da $n = 1000$ osservazioni i.i.d. ed il numero di osservazioni necessarie per avere un errore medio inferiore a 10^{-2}. Effettuando la stima su 1000 dati, si ha

$$\sigma \leq \frac{1}{2\sqrt{n}} = \frac{1}{2\sqrt{1000}} = 0.016$$

Volendo calcolare il numero di dati necessari al fine di ottendere un errore medio minore di di 10^{-2}, si ha

$$\frac{1}{2\sqrt{n}} \approx 10^{-2} \qquad \Longrightarrow n \approx 2500$$

In realta è sufficiente un numero inferiore di dati in quanto si è utilizzata una maggiorazione dell'errore medio, ma il calcolo preciso richiede la conoscenza della probabilità che si vuole stimare.

Esercizi

Ex. 1
Si consideri un triangolo isoscele in cui i lati uguali hanno misura unitaria e l'angolo al vertice è una variabile aleatoria uniforme in $(0, \pi)$. Calcolare la media e la varianza dell'area del triangolo.

Ex. 2
Una variabile aleatoria X ha la seguente PDF

$$f_X(x) = \begin{cases} \dfrac{2x}{9} & 0 < x < 3 \\ 0 & \text{altrimenti} \end{cases}$$

Verificare che la funzione assegnata soddisfi le proprietà caratterizzanti la PDF. Calcolare inoltre

1. $P(\{X \leq 2\})$
2. $P(\{X < 2\})$
3. $P(\{-1 < X < 1.5\})$.

Ex. 3
Siano X e Y due *v.a.* a media nulla, identicamente distribuite, ma correlate con coefficiente di correlazione ρ. Calcolare valor medio, valore quadratico medio e varianza delle due v.a.

$$Z = X$$
$$V = Y - \rho X$$

Dimostrare inoltre che tali v.a. sono incorrelate.

Ex. 4
Si consideri un canale di trasmissione con ingresso X ed uscita $Y = X + D_1 + D_2$. Nell'ipotesi che le v.a. X, D_1, D_2 siano gaussiane indipendenti e precisamente si abbia:

$$X \sim \mathcal{N}(\mu_X, \sigma_X^2), \qquad D_1, \sim \mathcal{N}(0, \sigma_D^2) \qquad D_2 \sim \mathcal{N}(0, \sigma_D^2)$$

calcolare media, valore ms e varianza dell'ingresso e dell'uscita del canale; determinare inoltre il coefficiente di correlazione tra l'ingresso e l'uscita. Stabilire infine che cambia se, a parità di medie e varianze, le v.a. non son gaussiane, ma hanno PDF arbitraria.

Ex. 5
Si consideri un canale di trasmissione con ingresso $X \sim \mathcal{B}(1, 1/2)$ ed uscita Y, variabile aleatoria discreta, con alfabeto $\mathcal{A}_Y = \{-1, 0, 1\}$ e PMF condizionali riportate in tabella.

y	$f_{Y\|X}(y\|0)$	$f_{Y\|X}(y\|1)$
1	0	$1-\epsilon$
0	ϵ	ϵ
-1	$1-\epsilon$	0

Calcolare $E[Y]$, $E[Y^2]$, e σ_Y^2; determinare inoltre il coefficiente di correlazione tra l'ingresso e l'uscita del canale.

Ex. 6
Siano X_1 e X_2 due variabili aleatorie gaussiane indipendenti, $X_1 \sim \mathcal{N}(1, 4)$ e $X_2 \sim \mathcal{N}(0.5, 0.36)$. Date le due

variabili aleatorie $Y_1 = X_1$ e $Y_2 = X_1 + X_2$ si determinino medie e varianze di Y_1 e Y_2 ed il loro coefficiente di correlazione .

Ex. 7
Date le variabili aleatorie $X \sim \mathcal{N}(1, 0.16)$ e $Y \sim \mathcal{N}(2, 0.25)$, congiuntamente gaussiane con coefficiente di correlazione $\rho_{XY} = 0.8$, calcolare $P(Z > 4)$, dove $Z = X + Y$.

N.B. Se X e Y sono congiuntamente Gaussiane la loro somma è gaussiana.

Ex. 8
In un negozio di materiale elettrico è possibile trovare cavi la cui lunghezza può essere schematizzata come una variabile aleatoria uniforme nell'intervallo [0.9,1.1] metri, [0.95,1.05] metri e [0.98,1.02] metri, a seconda che essi siano stati prodotti negli stabilimenti A,B e C rispettivamente. Sapendo che un cavo proviene da A con probabilità 0.5, da B con probabilità 0.2 e da C con probabilità 0.3, calcolare:

1. la lunghezza media di un cavo;
2. la probabilità che la lunghezza di un cavo prodotto nello stabilimento A si discosti dalla media per più di 1 cm in valore assoluto;
3. la probabilità che un cavo sia stato prodotto nello stabilimento A, sapendo che la sua lunghezza si discosta dalla media per meno di 1 cm in valore assoluto.

Determinare inoltre come varia la lunghezza media se, a causa di un malfunzionamento, lo stabilimento C comincia a produrre cavi con lunghezza uniforme nell'intervallo [1.48,1.52].

Ex. 9
Si lancia una coppia di dadi e siano X_1 e X_2 i risultati del lancio. Considerando il valore assoluto della differenza $Y = |X_1 - X_2|$ si calcoli

a. la PMF di Y;
b. la probabilità che il risultato X_1 sia uguale a 4, conoscendo che $Y = 2$.

Ex. 10
Si considerino due variabili aleatorie $Y_1 = 3X_1 + 2$ e $Y_2 = 2X_2$ costruite a partire dalle variabili aleatorie gaussiane standard, indipendenti X_1 e X_2. Calcolare:
1. la matrice di covarianza C tra le due variabili Y_1 e Y_2;
2. la media e la varianza di $Z = Y_1 + Y_2$;
3. La probabilità $\mathrm{P}(Z \geq 3)$.

Ex. 11
Si consideri la coppia di variabili aleatorie discrete X e Y con PMF congiunta

$Y \backslash X$	0	1	2
0	0.10	0.20	0.15
1	0.05	0.10	0.20
2	0.05	0.05	0.10

Calcolare:
a. la PMF della v.a. X;
b. la PMF della v.a. $Z = X + Y$;
c. la probabilità che la $Y - X$ sia minore di 0.