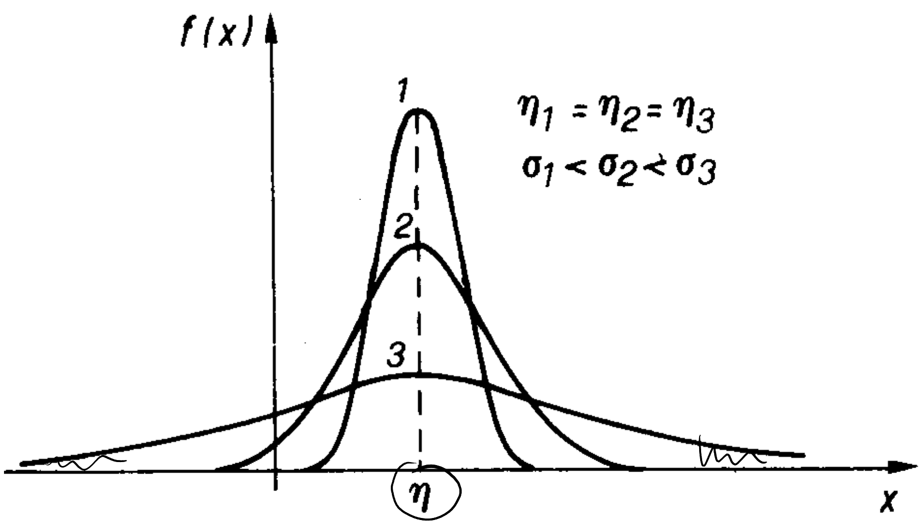
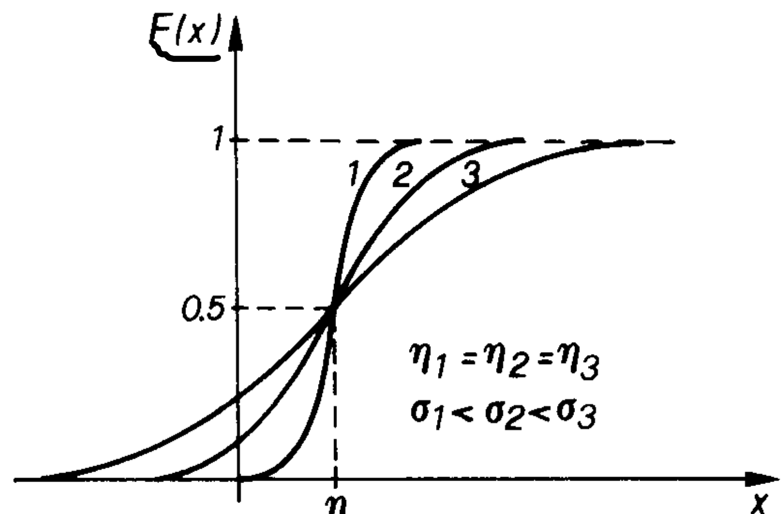
Variabile aleatoria Gaussiana (o normale)

Una v.a. X è detta **Gaussiana** o **normale** di parametri (η, σ^2), e si indica con $X \in \mathcal{N}(\eta, \sigma^2)$, se la sua ddp è:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\eta}{\sigma}\right)^2\right] = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\eta)^2}{2\sigma^2}\right]$$

con $\sigma > 0$

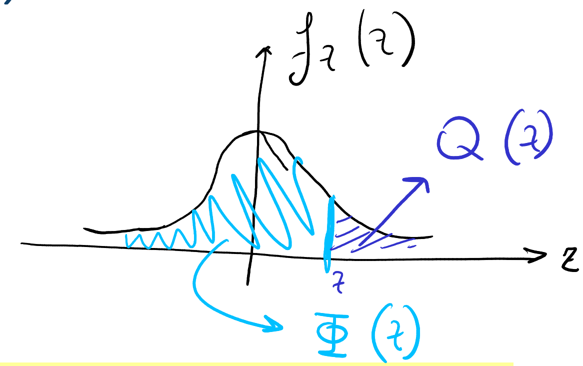




Variabile aleatoria Gaussiana (o normale)

$Z \in \mathcal{N}(0,1)$ → v.a. Gaussiana (o normale) standard

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$



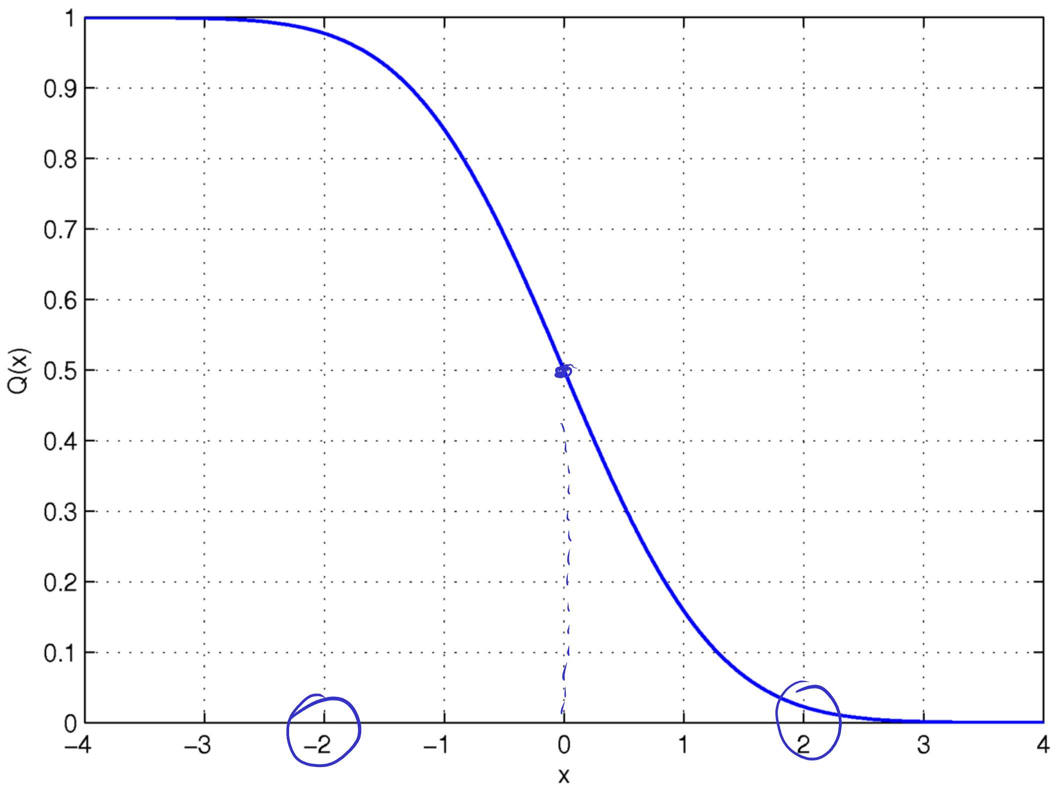
$$\Phi(z) \triangleq F_Z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{\alpha^2}{2}\right) \mathrm{d}\alpha$$

Non valutabile in forma chiusa; si trova tabulata o si calcola numericamente

Funzione Q :
$$Q(z) = \frac{1}{\sqrt{2\pi}} \int_{z}^{+\infty} \exp\left(-\frac{\alpha^2}{2}\right) \mathrm{d}\alpha = 1 - \Phi(z)$$

Funzione Q

$$Q(z)=\frac{1}{\sqrt{2\pi}}\int_{z}^{+\infty}\exp\left(-\frac{\alpha^2}{2}\right)\mathrm{d}\alpha=1-\Phi(z)$$



x	Q(x)
.0	.5000
.1	.4602
.2	.4207
.3	.3821
.4	.3446
.5	.3085
.6	.2743
.7	.2420
.8	.2119
.9	.1841
1.0	.1587
1.1	.1357
1.2	.1151
1.3	.0968
1.4	.0808
1.5	.0668
1.6	.0548
1.7	.0446
1.8	.0359
1.9	.0287
2.0	.0228
2.1	.0179
2.2	.0139
2.3	.0107
2.4	.0082
2.5	.0062
2.6	.0047
2.7	.0035
2.8	.0026
2.9	.0019
3.0	.0013

$$Q(-z)=1-Q(z)$$

Variabile aleatoria Gaussiana (o normale)

Se $X \in \mathcal{N}(\eta, \sigma^2)$, la funzione di distribuzione si ricava da quella della normale standard mediante la seguente relazione:

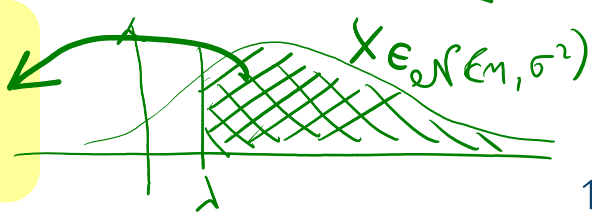
$$F_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{(\alpha-\eta)^2}{2\sigma^2}\right] \mathrm{d}\alpha$$

$$\left[\text{ponendo}: \; y = \frac{\alpha-\eta}{\sigma} \Rightarrow \mathrm{d}y = \frac{\mathrm{d}\alpha}{\sigma}\right]$$

$d\alpha = \sigma \, dy$

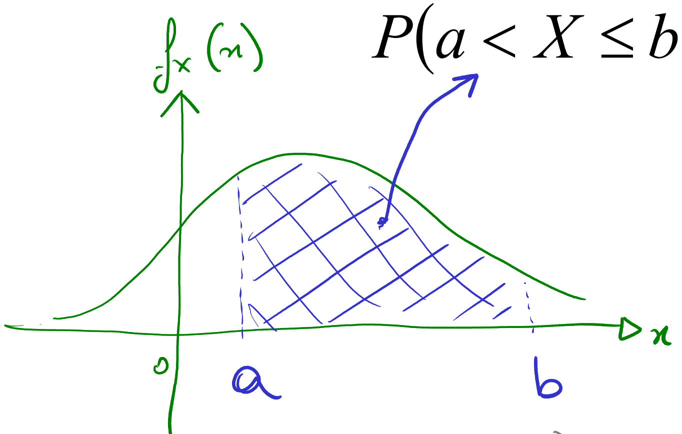
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-\eta)/\sigma} \exp\left(-\frac{y^2}{2}\right) \mathrm{d}y = \Phi\left(\frac{x-\eta}{\sigma}\right) = 1 - Q\left(\frac{x-\eta}{\sigma}\right)$$

$$\text{quindi}: \; \mathrm{P}(\mathrm{X} > \lambda) = 1 - F_X(\lambda) = Q\left(\frac{\lambda-\eta}{\sigma}\right)$$



Variabile aleatoria Gaussiana (o normale)

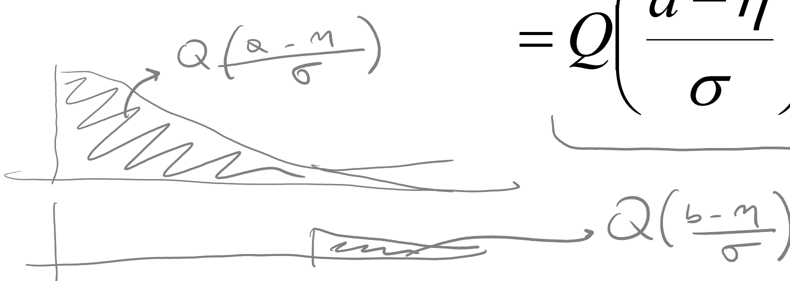
In generale, data una v.a. Gaussiana di parametri (η, σ^2), la probabilità associata ad un qualunque evento di interesse può essere calcolata mediante la funzione $\Phi(\cdot)$ oppure equivalentemente mediante la funzione $Q(\cdot)$:



$$P(a < X \le b) = \int_a^b f_X(\alpha)\,\mathrm{d}\alpha = F_X(b) - F_X(a)$$

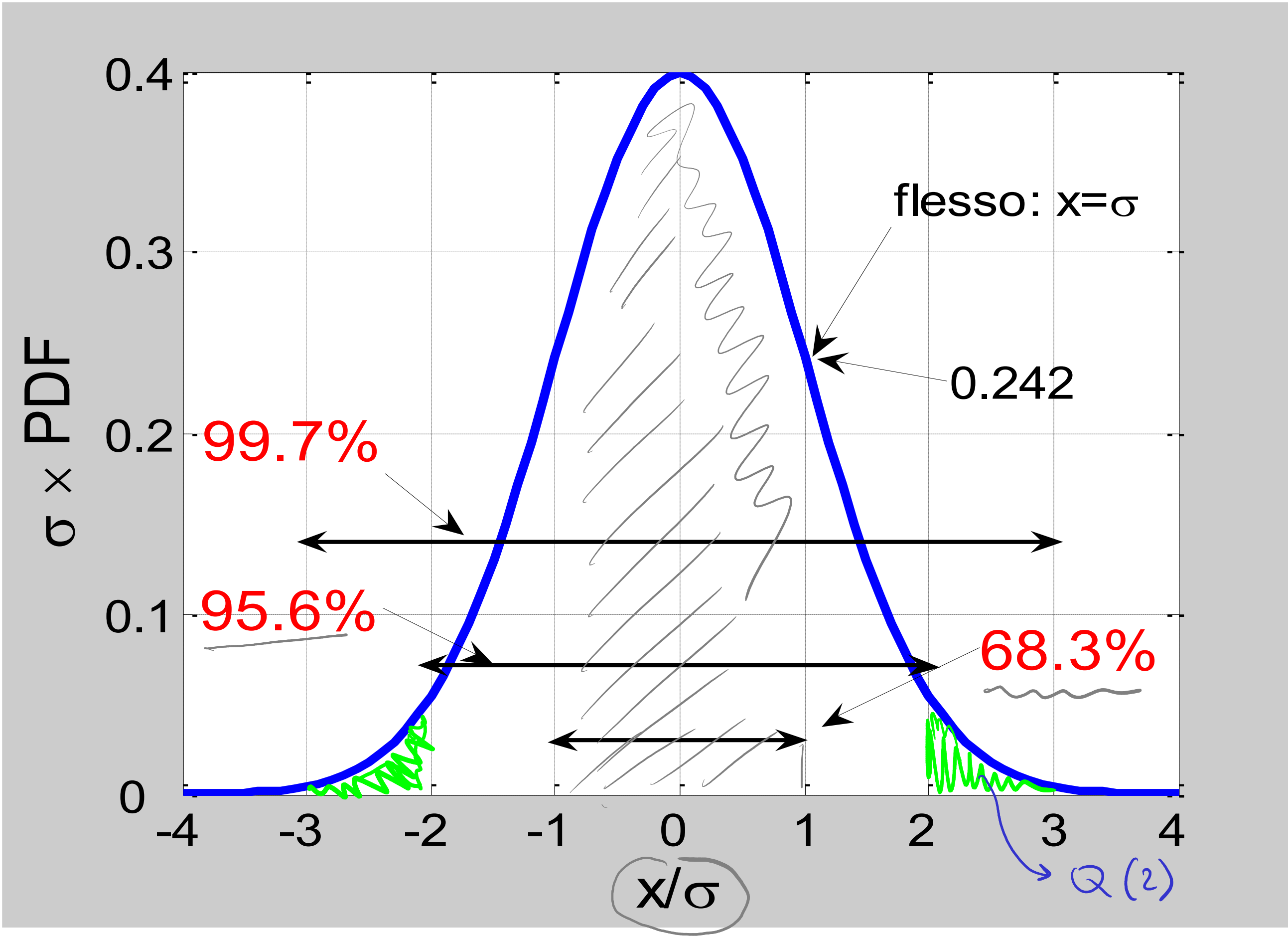
$$= \Phi\left(\frac{b-\eta}{\sigma}\right) - \Phi\left(\frac{a-\eta}{\sigma}\right) =$$

$$= Q\left(\frac{a-\eta}{\sigma}\right) - Q\left(\frac{b-\eta}{\sigma}\right)$$



Nota: $Q(z) = 1 - Q(-z)$ e $\Phi(z) = 1 - \Phi(-z)$

Variabile aleatoria Gaussiana (o normale)



> 2σ

$$P\left(|X-\eta|>n\sigma\right)=\int_{-\infty}^{\eta-n\sigma} f_X(\alpha)\,\mathrm{d}\alpha+\int_{\eta+n\sigma}^{+\infty} f_X(\alpha)\,\mathrm{d}\alpha=2Q(n),\quad n=1,2,3,\cdots$$

Variabile aleatoria Gaussiana (o normale)

$$P(|X-\eta| \le n\sigma) = 1 - 2Q(n)$$

$$P(|X-\eta| \le \sigma)$$

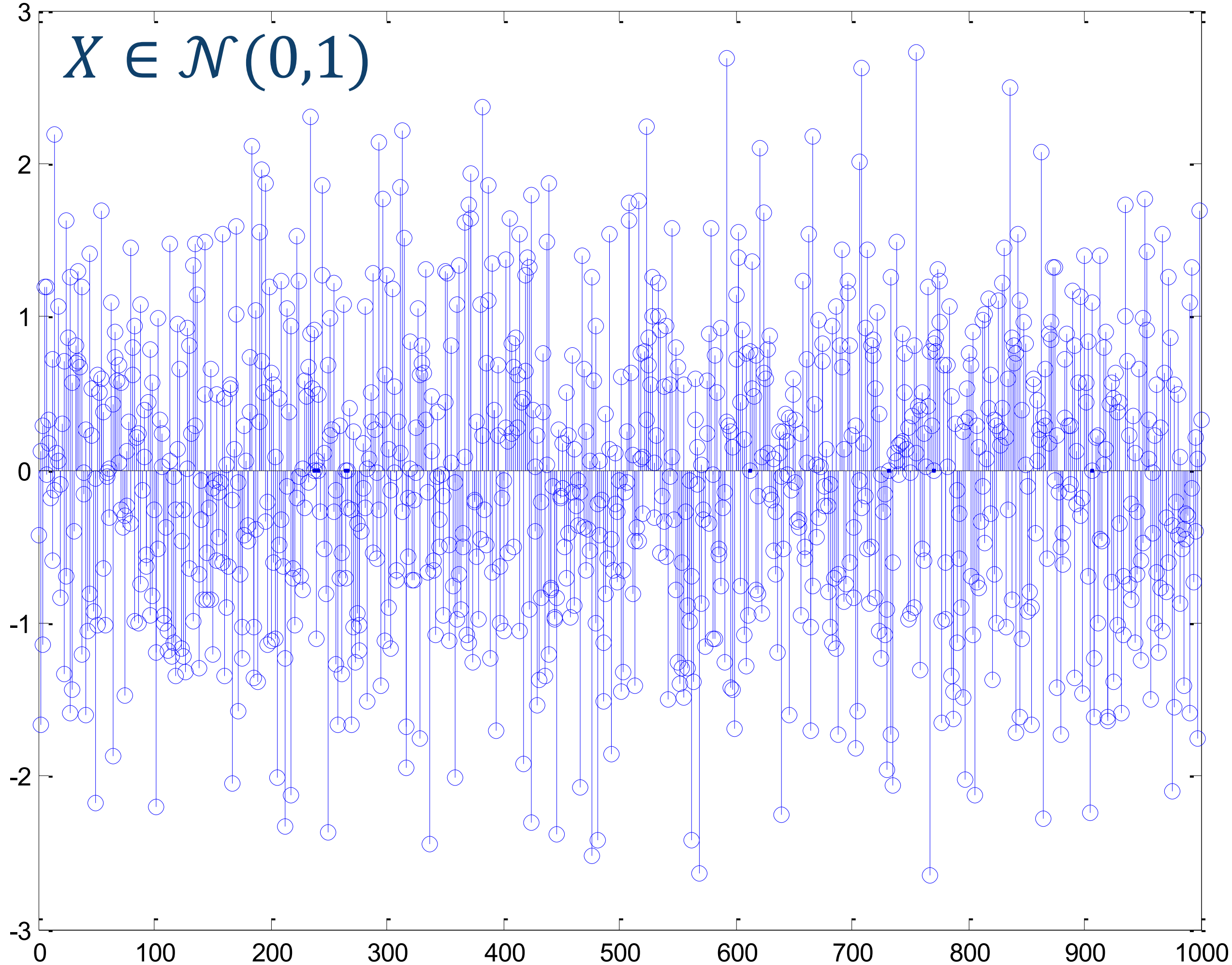
$$\cong 0.683$$

$$P(|X-\eta| \le 2\sigma)$$

$$\cong 0.956$$

$$P(|X-\eta| \le 3\sigma)$$

$$\cong 0.997$$



Valor medio e varianza di una v.a. Gaussiana

Esercizio:
Si dimostri che i due parametri della Gaussiana coincidono rispettivamente con il valore medio e con la varianza

DISP PARI

- $Z \in \mathcal{N}(0;1)$ $\quad E\{Z\} = \int_{-\infty}^{+\infty} z \cdot f_z(z)\, dz = \int_{-\infty}^{+\infty} z \cdot \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = 0 = \eta_z$

- $X \in \mathcal{N}\left(0, \frac{1}{\alpha}\right) \rightarrow f_x(x) = \sqrt{\frac{\alpha}{2\pi}} \cdot e^{-\alpha \frac{x^2}{2}}$

$$\int_{-\infty}^{+\infty} f_x(x)\, dx = 1 = \underbrace{\sqrt{\frac{\alpha}{2\pi}} \int_{-\infty}^{+\infty} e^{-\alpha\frac{x^2}{2}}\, dx}_{} \rightarrow \frac{1}{\sqrt{\alpha}} \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\alpha \frac{x^2}{2}}\, dx = \alpha^{-\frac{1}{2}}$$

DERIVO AMBO I MEMBRI
RISPETTO AD α

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left(-\frac{x^2}{2}\right) \cdot e^{-\frac{x^2}{2}\cdot\alpha}\, dx = \left(-\frac{1}{2}\right) \alpha^{-\frac{3}{2}} \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}\alpha} = \alpha^{-\frac{3}{2}} \qquad \alpha = 1$$

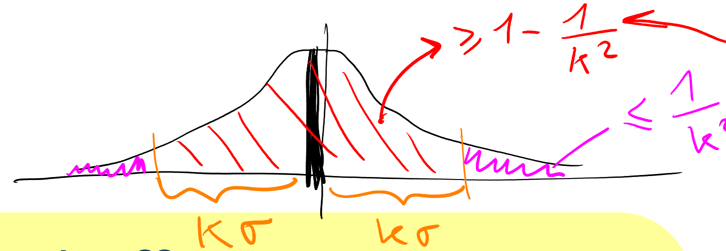
$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}} dx = E\{Z^2\} = 1 \qquad \sigma_z^2 = E\{z^2\} - \eta_z^2 = 1$$

$$X \in \mathcal{N}(\eta_x, \sigma_x^2) \longrightarrow X = \sigma_x Z + \eta_x$$

$$\mathbb{E}\{X\} = \sigma_x \cdot \underbrace{\mathbb{E}\{Z\}}_{=0} + \eta_x = \eta_x$$

$$\mathbb{E}\{(X - \eta_x)^2\} = \sigma_x^2 \cdot \underbrace{\mathbb{E}\{Z^2\}}_{=1} = \sigma_x^2$$

Proprietà della varianza



Teorema di Tchebycheff

Data una qualsiasi variabile aleatoria X con varianza finita, per ogni k>1 si ha: $\boxed{P\left(\left|X-\eta_X\right|\geq k\sigma_X\right)\leq \frac{1}{k^2}}$

$$\Rightarrow \quad P\left(\eta_X - k\sigma_X \leq X \leq \eta_X + k\sigma_X\right) \geq 1-\frac{1}{k^2}$$

Il teorema stabilisce un *limite inferiore* alla probabilità che la v.a. X stia entro un intervallo centrato attorno al valore medio e di ampiezza pari a $2k$ volte la deviazione standard, ***indipendentemente dalla densità di probabilità***

Proprietà della varianza

Esempio:
Scelgo k=5 → La prob. che X stia entro ±5 volte la deviazione standard rispetto al valor medio è *almeno* 1-1/25=0.96

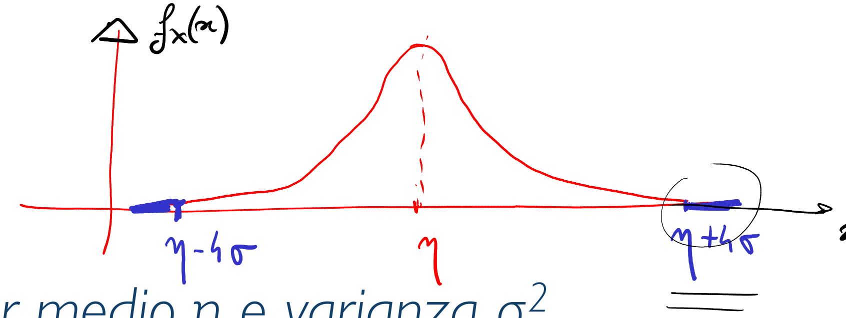
Nota la deviazione standard, possiamo determinare l'intervallo centrato sul valor medio di ampiezza 2ε in cui, con probabilità molto elevata, cadrà il valore osservato della v.a. X:

$$\boxed{\varepsilon = k\sigma_X} \Rightarrow \underline{\frac{1}{k^2} = \frac{\sigma_X^2}{\varepsilon^2}}$$

$$\underline{P\left(\left|X-\eta_X\right| \le \varepsilon\right)} = P\left(\eta_X - \varepsilon \le X \le \eta_X + \varepsilon\right) \ge 1 - \frac{\sigma_X^2}{\varepsilon^2}$$

mette in evidenza che la prob. che X sia compresa in $(\eta_X - \varepsilon, \eta_X + \varepsilon)$ è prossima all'unità per $\sigma_X \ll \varepsilon$

Proprietà della varianza



Esempio:

Si consideri una v.a. Gaussiana con valor medio η e varianza σ^2.
Si determini attraverso il teorema di Tchebycheff il limite superiore per la probabilità che X non appartenga all'intervallo (η - 4σ, η + 4σ). $k=4$
Si confronti quindi il risultato con il valore vero di tale probabilità.

$$P\left(\left|X-\eta_X\right|\geq k\sigma\right)\leq \overset{1/16}{\frac{1}{k^2}}=6.25\cdot 10^{-2}$$

Tchebycheff

$$P\left(\left|X-\eta_X\right|\geq k\sigma\right)=2\int_{\eta+4\sigma}^{\infty}\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\eta)^2}{2\sigma^2}\right]\mathrm{d}x$$

$$=2\int_{4}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\alpha^2}{2}\right)\mathrm{d}\alpha=2\mathrm{Q}(4)\cong 6.3\cdot 10^{-5}$$

Calcolo esatto

Variabili aleatorie condizionate

$X|C$

- La **funzione di distribuzione** della v.a. X condizionata all'evento C è definita come:

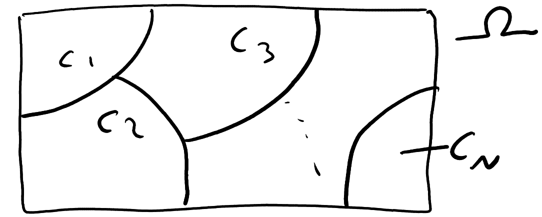
$$F_{X|C}(x|C) = P(X \leq x|C) = \frac{P(X \leq x, C)}{P(C)}$$

dove l'evento $\{X \leq x, C\}$ rappresenta tutti i punti dello spazio campione per cui $\{X(\omega) \leq x\}$ e $\{\omega \in C\}$, cioè l'evento $\{X \leq x\} \cap C$

- Si definisce **densità di probabilità** della v.a. X condizionata all'evento C:

$$f_{X|C}(x|C) = \frac{\mathrm{d}F_{X|C}(x|C)}{\mathrm{d}x}$$

Teorema della probabilità totale



Se l'insieme dei C_i (con i=1,2,…,N) rappresenta una partizione dello spazio campione Ω, dal **teorema della probabilità totale** si ha:

$$P(X \le x) = \sum_{i=1}^{N} P(X \le x \mid C_i) P(C_i)$$

ovvero:

$$F_X(x) = \sum_{i=1}^{N} F_{X|C}(x \mid C_i) P(C_i)$$

dove: $$\sum_{i=1}^{N} P(C_i) = 1$$

Variabili aleatorie condizionate

Esempio 8.6 Luise-Vitetta: *tempo di guasto dopo rodaggio*

- Supponiamo di effettuare il collaudo di una partita di lampadine, lasciandole accese fino allo spegnimento per guasto (bruciatura del filamento)
- In prima approssimazione, il tempo di guasto è una variabile aleatoria esponenziale unilatera, avente densità di probabilità

$$f_X(x) = \frac{1}{\eta}\exp\left(-\frac{x}{\eta}\right)\mathrm{u}(x)$$

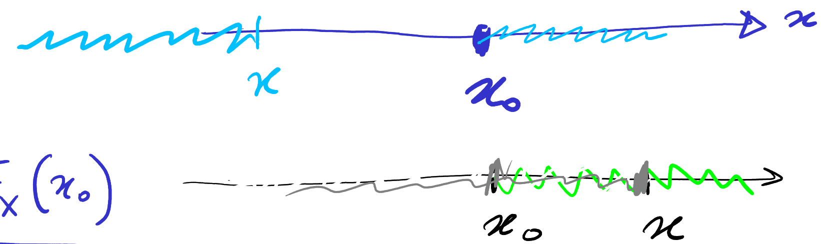
dove η è il tempo medio di guasto (MTBF, Mean Time Before Failure)

- Alteriamo adesso la modalità di collaudo: attendiamo un tempo fisso x_0 dall'accensione, scartiamo le lampadine che a tale istante risultano già guaste, e ripetiamo il collaudo come in precedenza sulle sole lampadine che a x_0 risultano funzionanti. Questa operazione di rodaggio avrà influenza sulla densità di probabilità del tempo di guasto delle lampadine rimaste?

Variabili aleatorie condizionate

- Identifichiamo l'evento *B* condizionante la variabile aleatoria X: *B* = $\{X > x_0\}$
- Vogliamo calcolare … $f_{X|B}(x|B)$

$$F_{X|B}(x|B) \triangleq \frac{P(\{X \le x\}, B)}{P\{B\}}$$



$$P(B) = P\{X > x_0\} = 1 - P\{X \le x_0\} = 1 - F_X(x_0)$$

$$P(\{X \le x\}, B) = P(\{X \le x, X > x_0\}) = \begin{cases} 0, & x < x_0 \\ F_X(x) - F_X(x_0), & x \ge x_0 \end{cases}$$

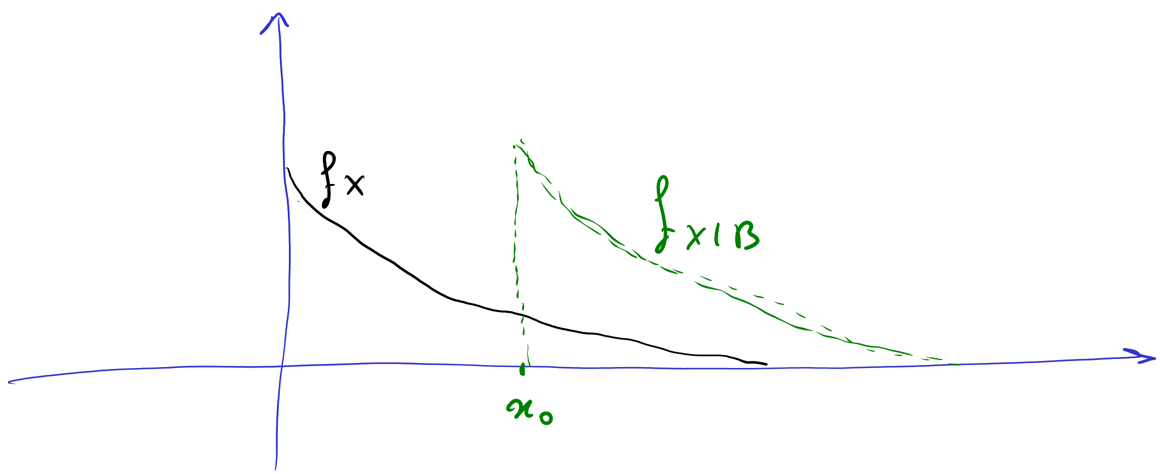
$$= [F_X(x) - F_X(x_0)] \cdot u(x - x_0)$$

$$F_{X|B}(x|B) = \frac{F_X(x) - F_X(x_0)}{1 - F_X(x_0)} \cdot u(x - x_0) \rightarrow f_{X|B}(x|B) = \frac{d F_{X|B}(x|B)}{dx} = \frac{f_X(x)}{1 - F_X(x_0)} u(x - x_0)$$

Variabili aleatorie condizionate

- Perché dal punto di vista del fabbricante non c'è alcuna convenienza a effettuare il rodaggio in fabbrica per il caso particolare delle lampadine con tempo di guasto esponenziale?

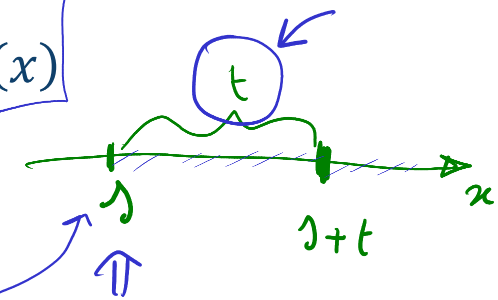
$$f_{X|B}(x|B) = \frac{1}{\eta} e^{-\frac{x-x_0}{\eta}} u(x-x_0)$$



$$f_X(x) = \frac{1}{\eta} e^{-\frac{x}{\eta}} u(x)$$

Assenza di memoria delle v.a. esponenziali

$$X \in Exp(\lambda) \rightarrow f_X(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \mathrm{u}(x); \quad \boxed{F_X(x) = \left[1 - e^{-\frac{x}{\lambda}}\right] \mathrm{u}(x)}$$



- Siano $s, t \geq 0$. Consideriamo gli eventi {X > t+s} e {X > s}

- $$P\{X > t + s | X > s\} = \frac{P\{(X>t+s)\cap(X>s)\}}{P\{X>s\}} = \frac{P\{X>t+s)\}}{P\{X>s\}} = \frac{1-F_X(t+s)}{1-F_X(s)} =$$

$$= \frac{\overbrace{\cancel{1-}\left[\cancel{1}-e^{-\frac{t+s}{\lambda}}\right]}^{F_X(t+s)}}{\underbrace{\cancel{1-}\left[\cancel{1}-e^{-\frac{s}{\lambda}}\right]}_{F_X(s)}} = \frac{e^{-\frac{t+s}{\lambda}}}{e^{-\frac{s}{\lambda}}} = \frac{e^{-\frac{t}{\lambda}} \cdot \cancel{e^{-\frac{s}{\lambda}}}}{\cancel{e^{-\frac{s}{\lambda}}}} = e^{-\frac{t}{\lambda}} = P\{X > t\}$$

Nota:

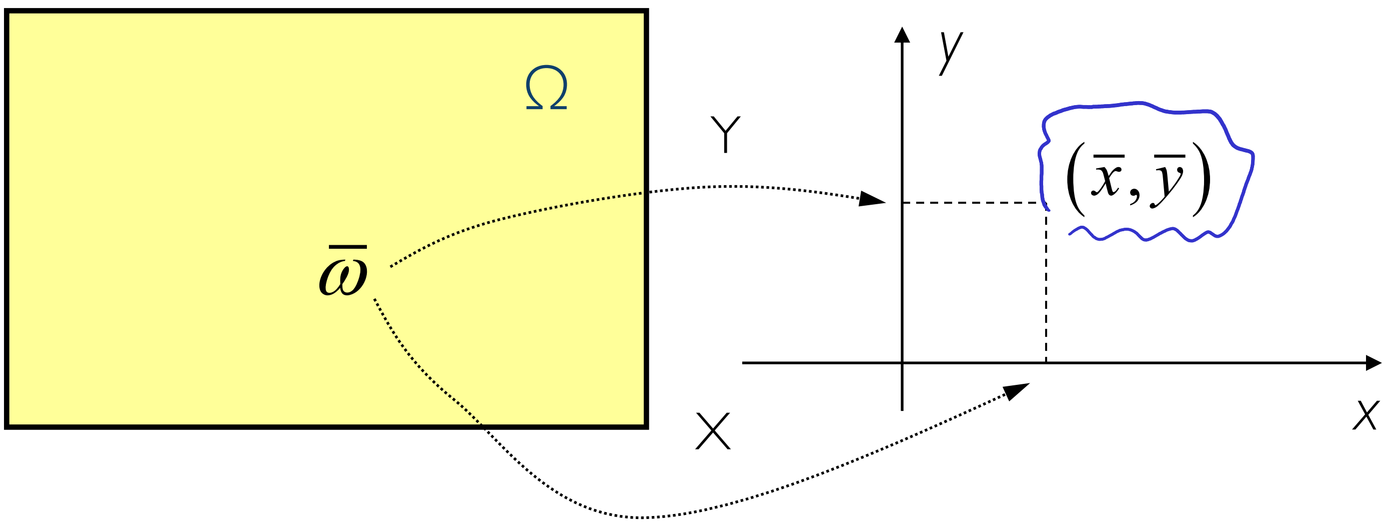
$X \in Exp(\lambda)$ rappresenta il tempo di guasto di un dispositivo: se ha funzionato per un tempo *s*, la probabilità che sopravviva un tempo addizionale *t* dipende *solo* da *t* (e *non* da *s*) ed è identica alla probabilità di funzionamento per un tempo *t* di un dispositivo nuovo → Il dispositivo «*non ricorda*» che è stato in esercizio per un tempo *s* (non accade per altre v.a. continue non negative!)

Sistemi di variabili aleatorie

Sistema di due variabili aleatorie

- Siano X e Y due v.a. definite **sullo stesso sistema** di probabilità $S=(\Omega, F, P)$

- Esse costituiscono un *sistema di 2 v.a.* (*vettore aleatorio bidimensionale*) che trasforma gli elementi dell'insieme Ω in punti nel piano (x, y)



Sistema di due variabili aleatorie

- Poiché X e Y sono v.a., gli insiemi $\{X \le x\}$ e $\{Y \le y\}$ costituiscono due eventi; la probabilità ad essi associata rappresenta la funzione di distribuzione di X e Y, rispettivamente

- La loro intersezione:

$$\{X \le x\} \cap \{Y \le y\} = \{X \le x, Y \le y\}$$

è un evento e la probabilità ad esso associata è nota come **Funzione distribuzione di probabilità congiunta** delle due v.a.:

$$F_{XY}(x, y) = P(X \le x, Y \le y)$$

Funzione distribuzione di probabilità congiunta

- $$F_{XY}(-\infty,-\infty)=F_{XY}(-\infty,y)=F_{XY}(x,-\infty)=0$$

$$\{X\le-\infty,Y\le-\infty\}=\varnothing$$

$\forall y$ $$\{X\le-\infty,Y\le y\}\subset\{X\le-\infty\}=\varnothing$$

$\forall x$ $$\{X\le x,Y\le-\infty\}\subset\{Y\le-\infty\}=\varnothing$$

- $$F_{XY}(\infty,\infty)=1$$

$$\{X\le\infty,Y\le\infty\}=\Omega$$

Funzioni distribuzione marginali

- Le funzioni di distribuzione di ciascuna delle due v.a. X e Y si ottengono dalle **regole marginali**:

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

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

Infatti:

$$\{X \leq x, Y \leq \infty\} = \{X \leq x\}$$

$$\{X \leq \infty, Y \leq y\} = \{Y \leq y\}$$

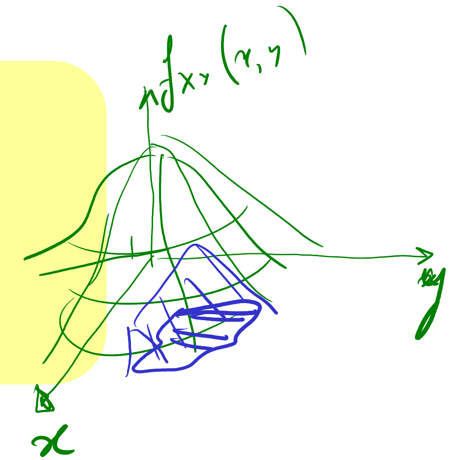
quindi:

$$F_{XY}(x, \infty) = P(X \leq x, Y \leq \infty) = P(X \leq x) = F_X(x)$$

Funzione densità di probabilità congiunta

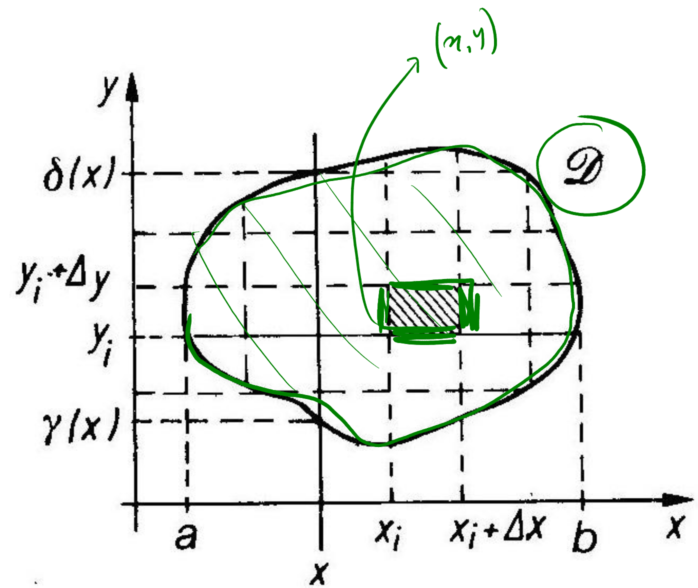
$$f_{XY}(x,y)=\frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$$

(è richiesta la derivabilità fino al secondo ordine)



$$f_{XY}(x,y)\,\mathrm{d}x\mathrm{d}y = P(x < X \le x+\mathrm{d}x, y < Y \le y+\mathrm{d}y)$$

$$P[(X,Y)\in D]=\iint_D f_{XY}(x,y)\,\mathrm{d}x\mathrm{d}y$$

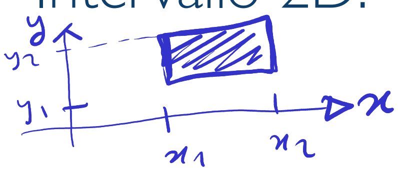


Sistema di due v.a. continue

Dalla relazione: $$P[(X,Y)\in D]=\iint_D f_{XY}(x,y)\,\mathrm{d}x\mathrm{d}y$$ si ricava:

Condizione di normalizzazione: $$P(X\le\infty, Y\le\infty)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,\mathrm{d}x\mathrm{d}y=1$$

Intervallo 2D: $$P(x_1<X\le x_2, y_1<Y\le y_2)=\int_{y_1}^{y_2}\int_{x_1}^{x_2} f_{XY}(x,y)\,\mathrm{d}x\mathrm{d}y$$



FD marginale di X: $$F_X(x)=P(X\le x, Y\le\infty)=\int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{XY}(\zeta,y)\,\mathrm{d}\zeta\,\mathrm{d}y$$

FD marginale di Y: $$F_Y(y)=P(X\le\infty, Y\le y)=\int_{-\infty}^{\infty}\int_{-\infty}^{y} f_{XY}(x,\eta)\,\mathrm{d}\eta\,\mathrm{d}x$$

Sistema di due v.a. continue

$$\underline{F_X(x)} = F_{XY}(x,\infty) \longrightarrow F_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{XY}(\zeta,y)\,\mathrm{d}\zeta\,\mathrm{d}y$$

Derivando rispetto ad *x* si ottiene:

$$f_X(x) = \boxed{\int_{-\infty}^{\infty} f_{XY}(x,y)\,\mathrm{d}y}$$

ddp marginale della v.a. *x*

Analogamente si ricava:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,\mathrm{d}x$$

ddp marginale della v.a. *y*

Sistema di due v.a. indipendenti

- Le variabili X e Y si dicono statisticamente indipendenti se gli eventi $\{X \leq x\}$ e $\{Y \leq y\}$ sono indipendenti, ovvero:

$$F_{XY}(x,y) = \underline{P(X \leq x, Y \leq y)} = P(X \leq x)P(Y \leq y)$$
$$= F_X(x)F_Y(y)$$

- Da cui segue che anche la ddp congiunta si fattorizza nel prodotto delle due ddp marginali:

$$f_{XY}(x,y) \triangleq \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y} = \underbrace{\frac{\partial F_X(x)}{\partial x}}_{f_X(x)} \cdot \underbrace{\frac{\partial F_Y(y)}{\partial y}}_{f_Y(y)} = f_X(x)f_Y(y)$$

Osservazione: nel caso di v.a. indipendenti le ddp marginali sono **sufficienti** per descrivere statisticamente il sistema

Trasformazione di un sistema di due v.a.

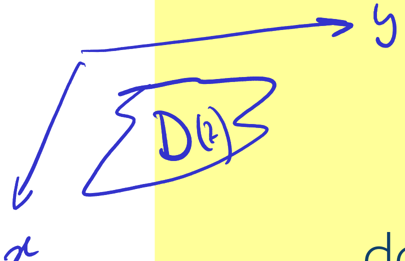
- Sia (X,Y) un sistema di v.a. definite per un esperimento di modello di probabilità assegnato; associando ad ogni coppia di valori (x,y) delle due v.a. il valore della funzione $z=g(x,y)$, risulta definita la nuova v.a. Z funzione del sistema (X,Y) e che indicheremo con:

$$Z = g(X,Y)$$

Problema: determinare la legge di distribuzione di Z nota la legge di distribuzione congiunta delle v.a. X e Y

Trasformazione di un sistema di 2 v.a. continue

Se X e Y sono v.a. continue, la **funzione di distribuzione** di Z è data dalla seguente espressione:



$$F_Z(z) = P(Z \le z) = P\left[g(X,Y) \le z\right]$$

$$= \iint_{D(z)} f_{XY}(x,y)\,\mathrm{d}x\mathrm{d}y$$

dove: $D(z) = \{(x,y) : g(x,y) \le z\}$

Trasformazione di un sistema di 2 v.a. continue

Per il calcolo del valore atteso η_Z della v.a. Z = g(X,Y), si può utilizzare il teorema dell'aspettazione:

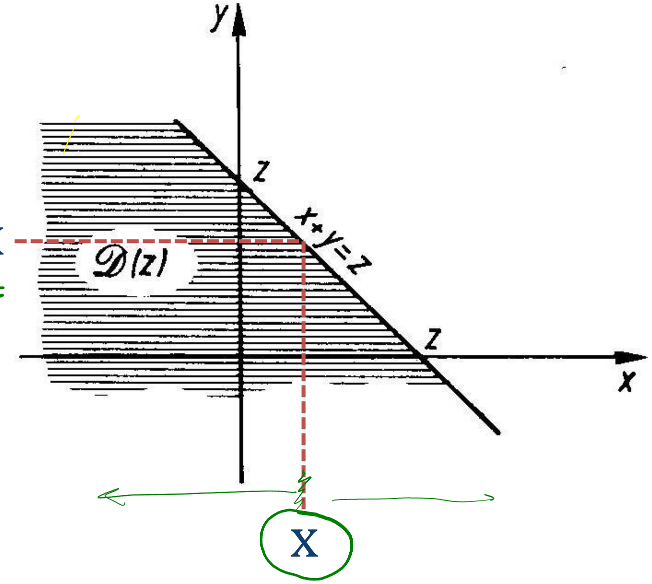
$$\eta_Z = E(Z) = E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{XY}(x,y)\mathrm{d}x\mathrm{d}y$$

Somma di due variabili aleatorie

Esempio: $Z = X + Y \qquad F_Z(z), f_Z(z), E(z), \sigma_Z^2 = ?$

$$D(z) = \{(x,y) : x + y \le z\}$$
$$= \{(x,y) : y \le z - x\}$$

y ≤ z − x



$$F_Z(z) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{z-x} f_{XY}(x,y)\,\mathrm{d}y \right] \mathrm{d}x$$

Derivando rispetto a z si ottiene la ddp della v.a. Z:

$$f_Z(z) = \int_{-\infty}^{\infty} f_{XY}(x, z - x)\,\mathrm{d}x$$

Somma di due v.a. indipendenti: ddp

- Se X e Y sono v.a. **indipendenti**:

$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(z-x)\,\mathrm{d}x = f_X(z) \otimes f_Y(z)$$

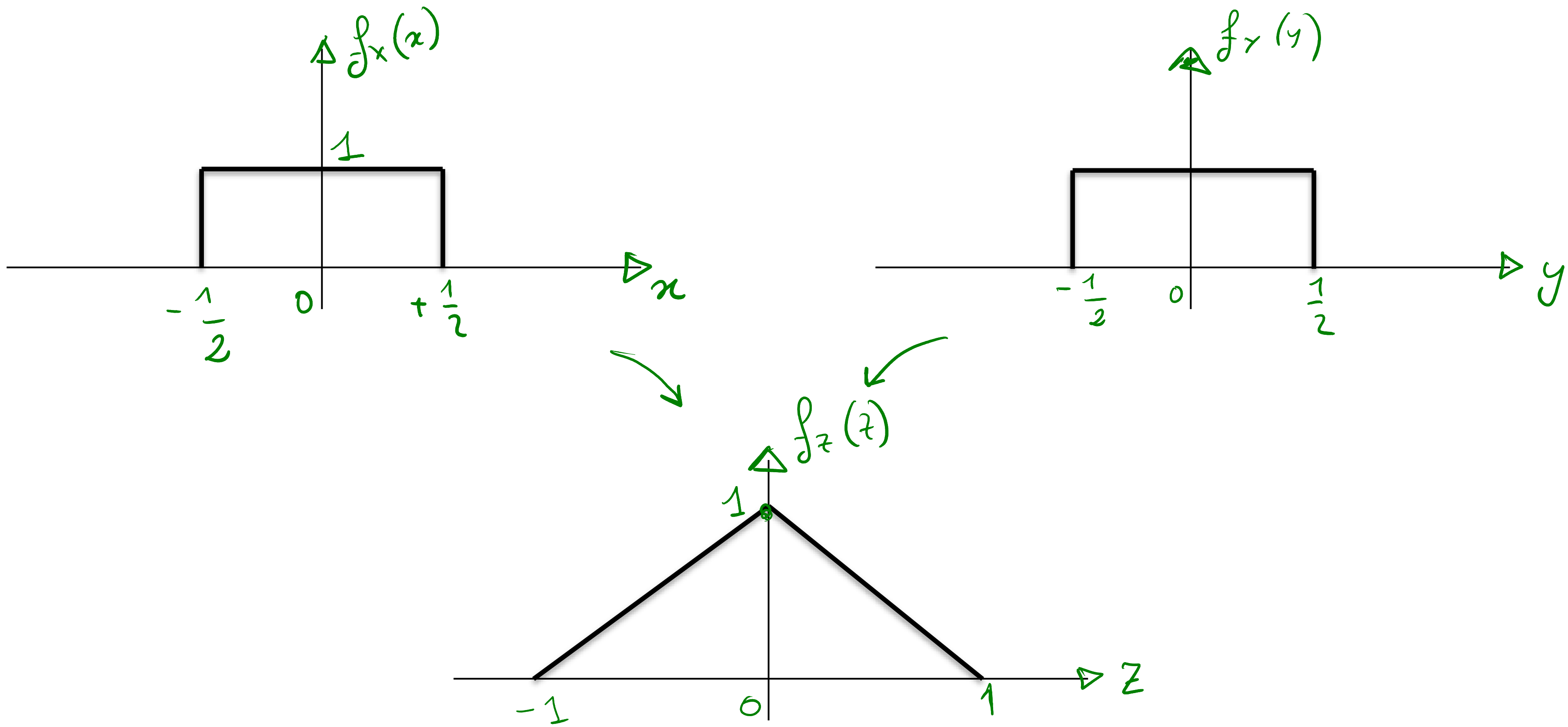
operatore "integrale di convoluzione"

Somma di due v.a. indipendenti: ddp

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,\mathrm{d}x$$

X, Y indip

$Z = X + Y$



Somma di due variabili aleatorie: valor medio

$Z = X + Y$

$g(x,y) \cdot f_{XY}(x,y)$

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

Risultato previsto, vista la linearità dell'operatore aspettazione

Correlazione e covarianza

- Il comportamento statistico di una variabile aleatoria *X* può essere caratterizzato in maniera incompleta ma talvolta sufficiente da alcuni parametri caratteristici, come il valor medio e la varianza
- Per una coppia di variabili aleatorie *(X,Y)* possiamo determinare alcuni parametri statistici semplificati che forniscono utili indicazioni per la comprensione del loro comportamento statistico *congiunto*

Correlazione:

$$r_{XY} = E\{XY\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\, y\, f_{XY}(x,y)\, dx\, dy$$

$$\boxed{c_{XY} = r_{XY} - \eta_X \eta_Y}$$

Covarianza:

$$c_{XY} = E\{(X-\eta_X)(Y-\eta_Y)\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x-\eta_X)(y-\eta_Y)\, f_{XY}(x,y)\, dx\, dy = r_{XY} - \eta_X\eta_Y$$

Somma di due variabili aleatorie: varianza

$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)+2c_{XY}$$

$Z = X + Y$

Infatti:

$[E(X)+E(Y)]^2$

$$\sigma_Z^2 = \underbrace{\mathrm{Var}(X+Y)} = E\left[(X+Y)^2\right]-E^2(X+Y)$$

$$= \underbrace{E(X^2)+E(Y^2)+2E(XY)}-\underbrace{E^2(X)-E^2(Y)-2E(X)E(Y)}$$

$$= \boxed{E(X^2)-E^2(X)}+\boxed{E(Y^2)-E^2(Y)}+2\big(\underbrace{E(XY)-E(X)E(Y)}_{r_{XY}-\eta_X\eta_Y \,=\, c_{XY}}\big)$$

$$\sigma_Z^2 = \mathrm{Var}(X)+\mathrm{Var}(Y)+2c_{XY}$$

Nota: se due v.a. sono tali che $c_{XY}=0$, cioè sono incorrelate, la varianza della somma è la somma delle varianze

Correlazione e covarianza

- La covarianza c_{XY} è un parametro statistico molto importante → Accerta se tra le due variabili X e Y esiste una relazione di dipendenza di tipo *lineare*, e che comunque misura la tendenza di variazione congiunta (covarianza) delle due
- Se la covarianza è grande e positiva, le due variabili aleatorie X e Y tendono a discostarsi dal rispettivo valor medio *nella stessa direzione*, cioè le due quantità $X-\eta_X$ e $Y-\eta_Y$ tendono ad avere lo stesso segno
- Se la covarianza tra due v.a. è nulla ($c_{XY}=0$), le variabili si dicono **incorrelate**
- Se la correlazione tra due v.a. è nulla ($r_{XY}=0$), le variabili si dicono **ortogonali**
- Il medesimo significato della covarianza ha il *coefficiente di correlazione* (o covarianza normalizzata) fra le variabili aleatorie X e Y:

$$\rho_{XY} = E\left\{\frac{X-\eta_X}{\sigma_X}\cdot\frac{Y-\eta_Y}{\sigma_Y}\right\} = \frac{c_{XY}}{\sigma_X\sigma_Y} = \frac{r_{XY}-\eta_X\eta_Y}{\sigma_X\sigma_Y}$$

Coefficiente di correlazione

$$\rho_{XY} = \frac{c_{XY}}{\sigma_X \sigma_Y}$$

Coefficiente di correlazione (*covarianza normalizzata*)

Proprietà:

- Poiché $c_{XY}^2 \leq \sigma_X^2 \sigma_Y^2$ si ha: $|\rho_{XY}| \leq 1$
- Se le v.a sono incorrelate: $\rho_{XY} = 0$
- Se $Y=a\,X+b$ (con $a \neq 0$): $|\rho_{XY}| = 1$