

Coefficiente di correlazione

- Con riferimento alla terza proprietà elencata, si può dimostrare che se viceversa il coefficiente di correlazione è 1, le v.a. sono *linearmente dipendenti*
- Per $|\rho_{XY}| = 1$ è possibile una previsione perfetta, ovvero se è noto X , ricavo esattamente $Y = aX + b$
- Il coefficiente di correlazione fornisce perciò una misura della dipendenza lineare tra due variabili aleatorie



Correlazione e covarianza di v.a. indipendenti

Nel caso di variabili aleatorie *indipendenti*:

$$\begin{aligned} \textcircled{r}_{XY} &= \underline{E\{XY\}} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \underline{xy} f_{XY}(x, y) dx dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_X(x) f_Y(y) dx dy = \\ &= \underbrace{\int_{-\infty}^{+\infty} x f_X(x) dx}_{\eta_X} \cdot \underbrace{\int_{-\infty}^{+\infty} y f_Y(y) dy}_{\eta_Y} \rightarrow \textcircled{r}_{XY} = \eta_X \eta_Y \end{aligned}$$

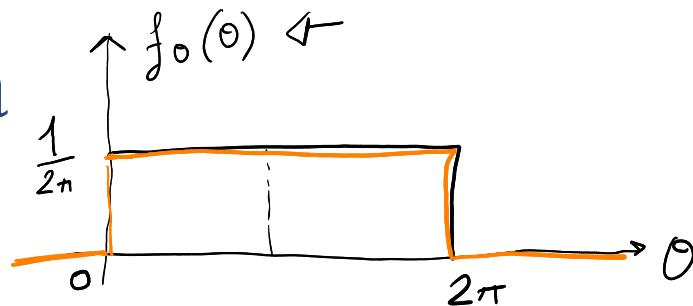
$$\textcircled{c}_{XY} = r_{XY} - \eta_X \eta_Y = 0$$

$$\textcircled{\rho}_{XY} = \frac{c_{XY}}{\sigma_X \sigma_Y} = 0$$

■ Se X e Y sono *indipendenti* sono anche *incorrelate* (covarianza nulla); viceversa, se la covarianza è nulla (variabili aleatorie incorrelate), non è detto che X e Y siano indipendenti

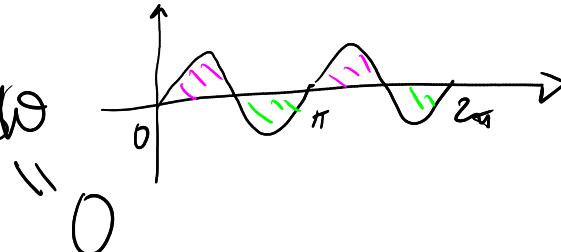


Incorrelazione ~~Indipendenza~~ Indipendenza



- Lo verifichiamo con l'**Esempio 8.8 – Libro LV**
- Sia Θ una variabile aleatoria uniformemente distribuita sull'intervallo $[0, 2\pi[$
- A partire da Θ si definiscono due ulteriori variabili aleatorie:
 $X = \sin(\Theta)$ e $Y = \cos(\Theta)$ di cui vogliamo valutare la correlazione statistica

$$\begin{aligned} r_{XY} &= \mathbb{E}\{XY\} = \mathbb{E}\{\sin(\theta) \cdot \cos(\theta)\} = \int_{-\infty}^{+\infty} \sin \theta \cdot \cos \theta \cdot f_{\theta}(\theta) d\theta \\ &= \int_0^{2\pi} \sin \theta \cdot \cos \theta \cdot \frac{1}{2\pi} d\theta = \frac{1}{4\pi} \int_0^{2\pi} \sin 2\theta d\theta \end{aligned}$$



$$\mu_X = \mathbb{E}\{\cos \theta\} = \int_0^{2\pi} \cos \theta \cdot \frac{1}{2\pi} d\theta = 0$$

$$\mu_Y = \mathbb{E}\{\sin \theta\} = \int_0^{2\pi} \sin \theta \cdot \frac{1}{2\pi} d\theta = 0$$

$$C_{XY} = r_{XY} - \mu_X \mu_Y = 0 - 0 \cdot 0 = 0$$

$$\sin^2 \theta + \cos^2 \theta = 1$$

$$X^2 + Y^2 = 1 \rightarrow X^2 = 1 - Y^2$$



Viti e dadi



$$X \in \mathcal{N}(5; 0,01)$$

Esercizio su sistema di v.a. indipendenti

- Una prima macchina produce viti il cui diametro X (in mm) può essere considerato una v.a. normale di parametri $(5; 0,01)$
- Una seconda macchina produce dadi il cui diametro Y (in mm) può considerarsi una v.a. normale di parametri $(5.2; 0,04)$
- Un dado si avvia correttamente su una vite se il suo diametro è maggiore di quello della vite di almeno 0.04 mm, ma non più di 0.36 mm
- Qual è la probabilità che una vite ed un dado scelti a caso si avvitino correttamente?

X, Y indipendenti

$$D = Y - X \rightarrow \mathcal{N}(\eta_D, \sigma_D^2)$$

$$\eta_D = \mathbb{E}\{Y - X\} = \mathbb{E}\{Y\} - \mathbb{E}\{X\} = 0,2$$



$$\begin{aligned}\sigma_D^2 &= \mathbb{E}\{D^2\} - \mathbb{E}^2\{D\} = \mathbb{E}\{(Y-X)^2\} - [\mathbb{E}\{Y\} - \mathbb{E}\{X\}]^2 \\ &= \mathbb{E}\{Y^2\} - 2\mathbb{E}\{XY\} + \mathbb{E}\{X^2\} - \mathbb{E}^2\{Y\} + 2\mathbb{E}\{X\} \cdot \mathbb{E}\{Y\} - \mathbb{E}^2\{X\} \\ &= \sigma_Y^2 + \sigma_X^2 - 2[\underbrace{\mathbb{E}\{XY\} - \mathbb{E}\{X\} \cdot \mathbb{E}\{Y\}}_{C_{XY}}]\end{aligned}$$

$$\sigma_D^2 = \sigma_Y^2 + \sigma_X^2 - 2C_{XY} \quad \text{VAR. IND. Sono inconosciute} \rightarrow C_{XY} = 0$$

$$\sigma_D^2 = 0,04 + 0,01 = 0,05$$

$$D \in \mathcal{N}(0,2; 0,05)$$

$$D = \sigma_D Z + \eta_D \rightsquigarrow Z = \frac{D - \eta_D}{\sigma_D}$$

$$\begin{aligned}P\{0,04 \leq D \leq 0,36\} &= F_D(0,36) - F_D(0,04) \\ &\Downarrow \\ &= 1 - Q\left(\frac{0,36 - \eta_D}{\sigma_D}\right)\end{aligned}$$

$$1 - Q\left(\frac{0,36 - 0,2}{\sqrt{0,05}}\right) = 1 + Q\left(\frac{0,04 - 0,2}{\sqrt{0,05}}\right) = -Q\left(\frac{0,16}{\sqrt{0,05}}\right) + Q\left(\frac{-0,16}{\sqrt{0,05}}\right)$$

$$\begin{aligned}Q(-x) &= 1 - Q(x) \\&= -Q\left(\frac{0,16}{\sqrt{0,05}}\right) + 1 - Q\left(\frac{+0,16}{\sqrt{0,05}}\right) \\&= 1 - 2 \cdot Q\left(\frac{0,16}{\sqrt{0,05}}\right) \\&= 1 - 2 \cdot Q(0,716) \\&\quad \underbrace{_{0,237}} \\&= 0,525\end{aligned}$$

Stanza con due lampade

Esercizio su sistema di v.a. indipendenti (MATLAB 8.6 – Libro LV)

- Una stanza priva di finestre è alimentata da un lampadario costituito da due lampade alimentate in parallelo
- I tempi di vita delle lampade sono descritti da due v.a. T_1 e T_2 **indipendenti** e caratterizzate da una ddp esponenziale monolatera con $\eta_1 = 2\eta$, $\eta_2 = \underline{\eta}$ ed $\eta = 6000$ ore

$$f_{T_i}(t_i) = \frac{1}{\eta_i} e^{-\frac{t_i}{\eta_i}} u(t_i), i = 1, 2$$

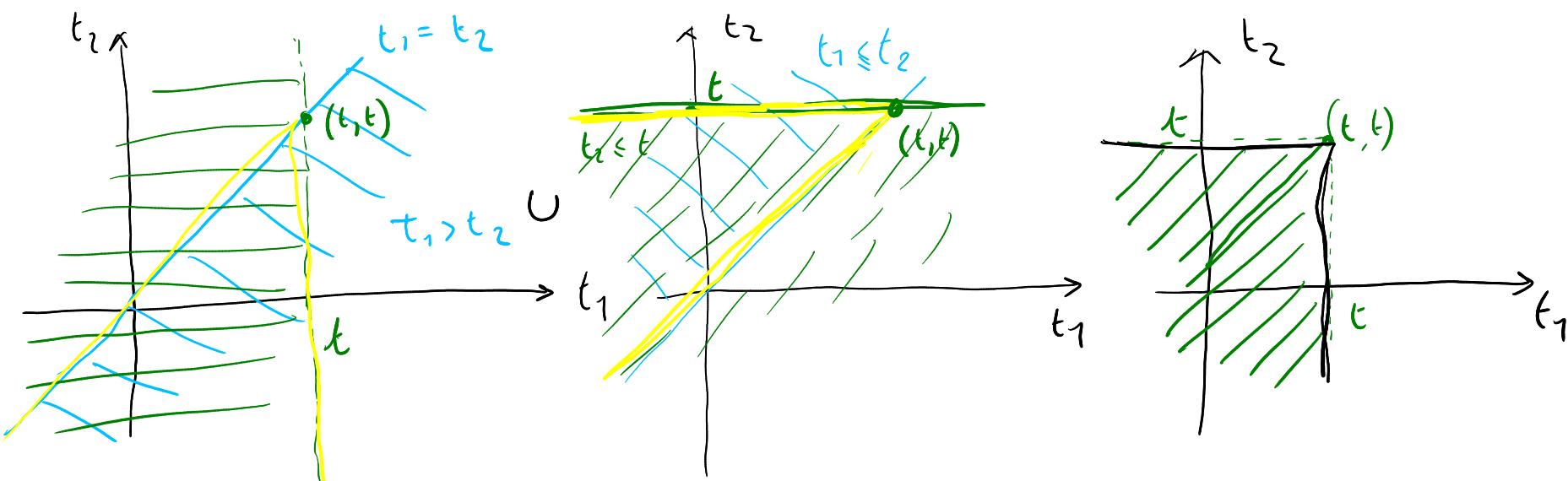
- Si calcoli il tempo medio di illuminazione della stanza

$$T = \max(T_1, T_2)$$

$$\begin{aligned} F_T(t) &= P\{T \leq t\} = P\{\max(T_1, T_2) \leq t\} \\ &= P\{(T_1 \leq t, T_1 > T_2) \cup (T_2 \leq t, T_1 \leq T_2)\} \end{aligned}$$

tempo di illuminazione delle lampade
 T





$$F_T(t) = P\{T_1 \leq t, T_2 \leq t\}$$

$$= \begin{cases} 0, & t \leq 0 \\ \int_0^t \int_0^t f_{T_1 T_2}(t_1, t_2) dt_1 dt_2, & t > 0 \end{cases}$$

$$f_{T_1 T_2}(t_1, t_2) = f_{T_1}(t_1) \cdot f_{T_2}(t_2) = \frac{1}{2\eta} \cdot e^{-\frac{|t_1|}{2\eta}} \cdot \frac{1}{\eta} e^{-\frac{|t_2|}{\eta}} u(t_1) u(t_2)$$

$$\begin{aligned}
 F_T(t) &= \int_0^t \int_0^t \frac{1}{2\eta} e^{-\frac{t_1}{2\eta}} \cdot \frac{1}{\eta} e^{-\frac{t_2}{\eta}} \cdot u(t) dt_1 dt_2 \\
 &= \underbrace{\int_0^t \frac{1}{2\eta} e^{-\frac{t_1}{2\eta}} dt_1}_{\left[-e^{-\frac{t_1}{2\eta}} \right]_0^t} \cdot \underbrace{\int_0^t \frac{1}{\eta} e^{-\frac{t_2}{\eta}} dt_2}_{\left[-e^{-\frac{t_2}{\eta}} \right]_0^t} u(t) \\
 &= \left[-e^{-\frac{t_1}{2\eta}} \right]_0^t \cdot \left[-e^{-\frac{t_2}{\eta}} \right]_{u(t)}^t = \left(1 - e^{-\frac{t}{2\eta}} \right) \cdot \left(1 - e^{-\frac{t}{\eta}} \right) u(t)
 \end{aligned}$$

$$\eta_T = \int_{-\infty}^{+\infty} t f_T(t) dt$$

$$\begin{aligned}
 f_T(t) &= \frac{dF_T(t)}{dt} = \left[\frac{1}{2\eta} e^{-\frac{t}{2\eta}} \left(1 - e^{-\frac{t}{\eta}} \right) + \left(1 - e^{-\frac{t}{2\eta}} \right) \cdot \frac{1}{\eta} e^{-\frac{t}{\eta}} \right] u(t) \\
 &= \left[\frac{1}{2\eta} e^{-\frac{t}{2\eta}} - \frac{1}{2\eta} e^{-\frac{3}{2}\frac{t}{\eta}} + \frac{1}{\eta} e^{-\frac{t}{\eta}} - \frac{1}{\eta} e^{-\frac{3}{2}\frac{t}{\eta}} \right] u(t)
 \end{aligned}$$

$$= \left[\frac{1}{\eta} e^{-\frac{t}{\eta}} + \frac{1}{2\eta} e^{-\frac{t}{2\eta}} - \frac{3}{2\eta} e^{-\frac{3}{2}\frac{t}{\eta}} \right] u(t)$$

$$\eta_T = \mathbb{E}\{T\} = \eta + 2\eta - \frac{2}{3}\eta = \frac{3+6-2}{3}\eta = \frac{7}{3}\eta = \frac{7}{3} \cdot 6000 = \underline{\underline{14000 \text{ ore}}}$$

Il teorema-limite centrale

Consideriamo la variabile aleatoria: $Y_n = \sum_{i=1}^n X_i$

Supponiamo che le n variabili aleatorie X_1, X_2, \dots, X_n siano *indipendenti* e con la stessa funzione densità di probabilità con valor medio η e varianza σ^2 , allora

$$\underbrace{E\{Y_n\} = \eta_n = n \cdot \eta}_{\text{e}} \quad \underbrace{E\{(Y_n - \eta_n)^2\} = \sigma_n^2 = n \cdot \sigma^2}_{\text{e}}$$

Se consideriamo un numero n di variabili man mano crescente, la densità di probabilità della variabile aleatoria *normalizzata*

$$S_n = \frac{Y_n - \eta_n}{\sigma_n} = \frac{Y_n - n \cdot \eta}{\sqrt{n} \cdot \sigma}$$

tende ad una densità normale standard, cioè con valor medio nullo e varianza unitaria

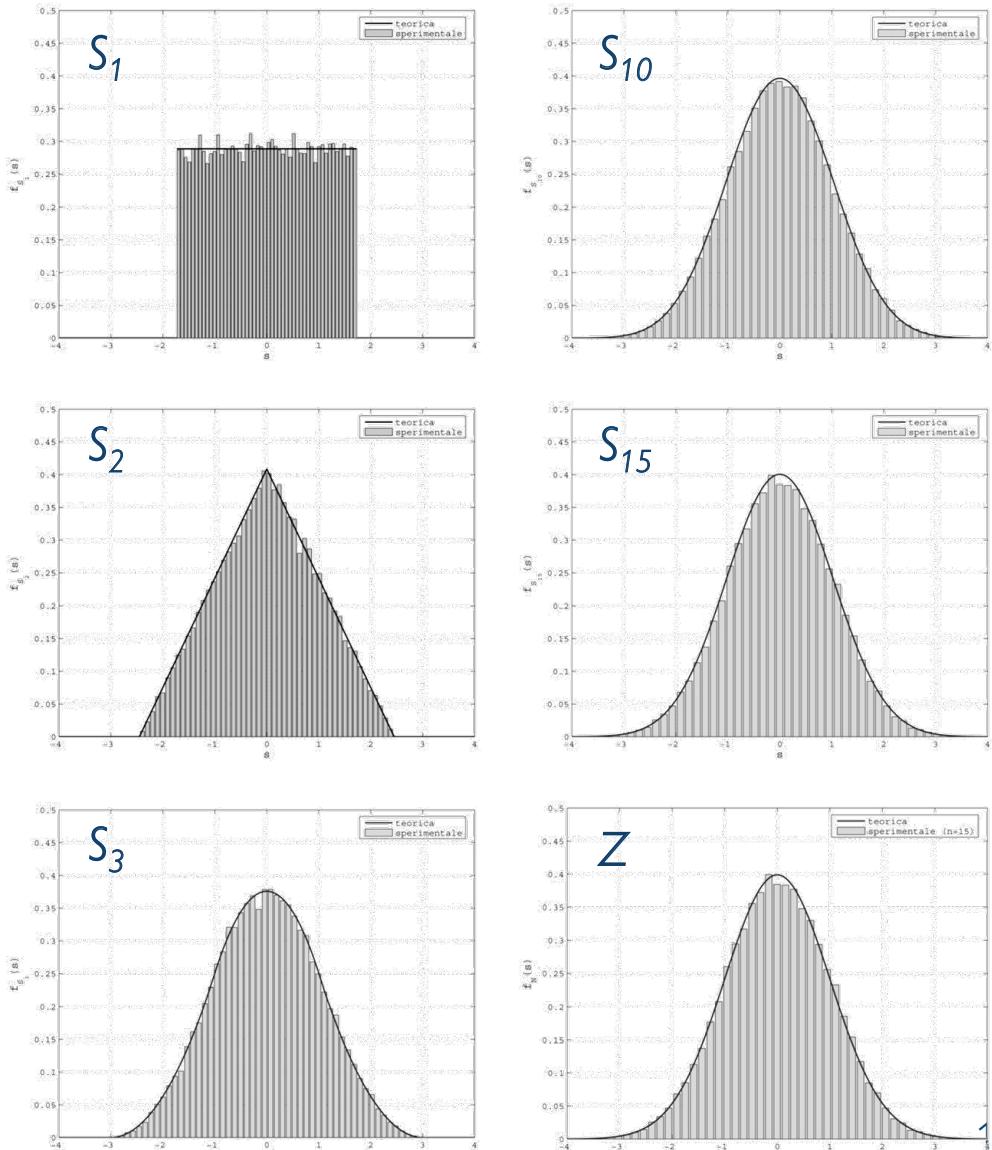
$$\lim_{n \rightarrow +\infty} f_{S_n}(s) = f_Z(s) = \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}}$$



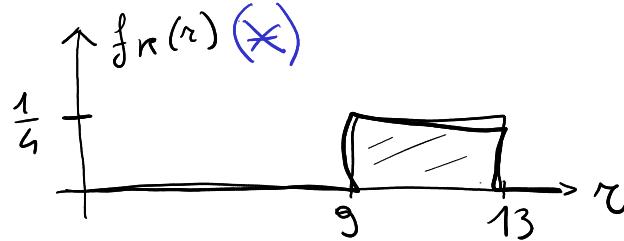
Il teorema-limite centrale

- Esempio: MATLAB 8.7 – Libro LV

Verifichiamo la tendenza della somma normalizzata S_n alla Gaussiana standard Z utilizzando $X_i \in U(0,1)$



Il teorema-limite centrale



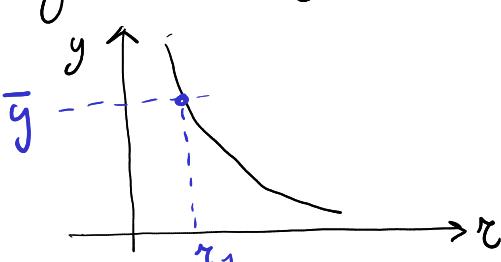
Esercizio: Si hanno a disposizione dei resistori la cui resistenza in Ω può essere descritta da una v.a. R avente distribuzione uniforme di parametri $a=9$ e $b=13$.

1. Se si collegano in parallelo 100 di tali resistori scelti a caso, qual è la probabilità che la resistenza complessiva superi 0,11 Ω ?
2. Se si collegano in serie 100 di tali resistori scelti a caso, qual è la probabilità che la resistenza complessiva superi 1120 Ω ?

$$Y_n = \frac{1}{R_n} \rightarrow f_Y(y) = ?$$

$$y = g(r) = \frac{1}{r}$$

$$g'(r) = -\frac{1}{r^2}$$



$$f_Y(y) = \sum_{i=1}^k \frac{f_R(r_i)}{|g'(r_i)|} \Big|_{r_i = g^{-1}(y)}$$

$$f_Y(y) = \begin{cases} 0, & y \leq 0 \\ \frac{f_R(r_1)}{|g'(r_1)|} & |r_1 = g^{-1}(y)| \end{cases} = \frac{f_R(r_1)}{\frac{1}{r_1^2}} \Big|_{r_1 = \frac{1}{y}}$$

$(*) \neq 0$ se $9 \leq r_1 \leq 13$

$$\frac{1}{y} \geq y \geq \frac{1}{13}$$



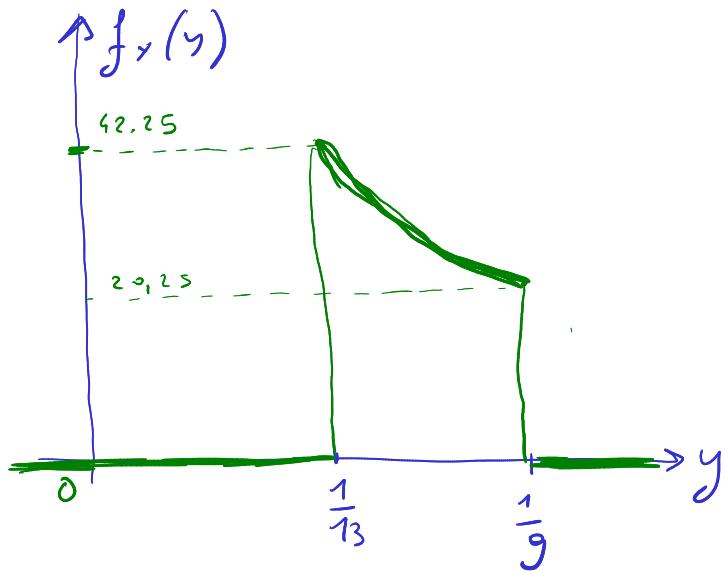
$$f_y(y) = \begin{cases} 0, & y \leq 0 \\ \frac{1/4}{y^2} = \frac{1}{4y^2}, & \frac{1}{13} \leq y \leq \frac{1}{9} \\ 0, & \text{altrimenti} \end{cases}$$

$$\mathbb{E}\{Y\} = \mathbb{E}\{g(R)\} = \int_{-\infty}^{+\infty} \frac{1}{r} \cdot f_R(r) dr$$

$$= \int_9^{13} \frac{1}{r} \cdot \frac{1}{4} dr = \frac{1}{4} (\ln r) \Big|_9^{13} = \frac{1}{4} \cdot \ln\left(\frac{13}{9}\right) \cong \underline{0,0919} = \eta_x$$

$$\mathbb{E}\{Y^2\} = \int_{-\infty}^{+\infty} y^2 \cdot f_y(y) dy = \int_{1/13}^{1/9} y^2 \cdot \frac{1}{4y^2} dy = \int_{1/13}^{1/9} \frac{1}{4} dy = \frac{1}{4} \left[\frac{1}{9} - \frac{1}{13} \right] = \frac{1}{4} \cdot \frac{\cancel{11}}{13 \cdot 9}$$

$$\sigma_y^2 = \mathbb{E}\{Y^2\} - \eta_x^2 = \frac{1}{13 \cdot 9} - (0,0919)^2 \cong \underline{9,55 \cdot 10^{-5}}$$



$$Y_{100} = \sum_{m=1}^{100} Y_m$$

$$\mathbb{E}\{Y_{100}\} = 100 \cdot \eta_y \approx 9,19 = \eta_{100}$$

$$\mathbb{E}\{(Y_{100} - \eta_{100})^2\} = \sigma_{100}^2 = 100 \cdot \sigma_y^2 \approx 9,55 \cdot 10^{-3}$$

$$R_{100} = \frac{1}{Y_{100}}$$

TÉO. CLT. CENTRALE

$$S_{100} = \frac{Y_{100} - \eta_{100}}{\sigma_{100}} \approx Z$$

$$\mathbb{P}\{R_{100} > 0,11\}$$

$$\mathbb{P}\left\{\frac{1}{Y_{100}} > 0,11\right\}$$

$$\mathbb{P}\left\{Y_{100} < \frac{1}{0,11}\right\} = \mathbb{P}\left\{Y_{100} < \underline{9,09}\right\} =$$

$$I' = 1 - \mathbb{P}\{Y_{100} > 9,09\} = 1 - Q\left(\frac{9,09 - \eta_{100}}{\sigma_{100}}\right)$$

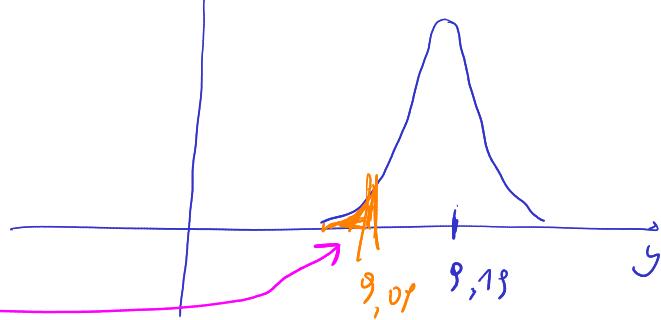
$$I' = 1 - Q\left(\frac{9,09 - 9,19}{0,0378}\right) = 1 - Q(-1,045)$$

$$I' = Q(1,045) \approx 0,148$$

$$Y_{100} \approx \sigma_{100} \cdot Z + \eta_{100}$$

$$Y \stackrel{\sim}{\in} \mathcal{N}(\eta_{100}, \sigma_{100}^2)$$

$$f_{Y_{100}}(y)$$



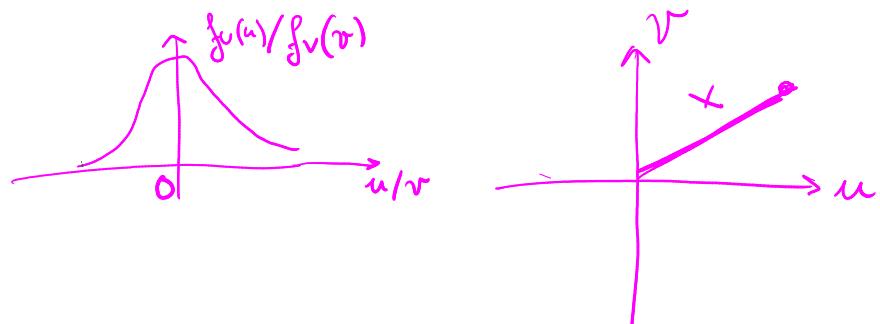
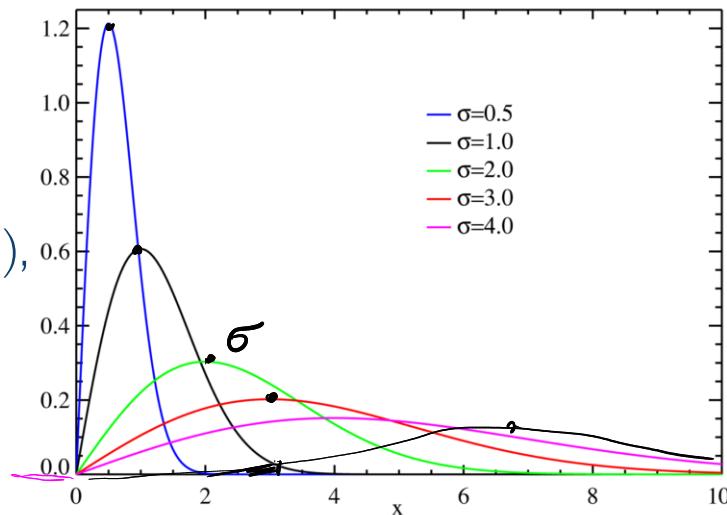
Esercizio: v. a. di Rayleigh

- Una v.a. X è detta **Rayleigh** di parametro σ ($\sigma > 0$), e si indica con $X \in \mathcal{R}(\sigma^2)$, se la sua ddp è:

$$f_X(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} u(x)$$

- Se $U, V \in \mathcal{N}(0, \sigma^2) \rightarrow X = \sqrt{U^2 + V^2}$: X rappresenta la distanza dall'origine di un punto (U, V) nel piano euclideo le cui coordinate siano v. a. indipendenti aventi entrambe distribuzione normale
- Importante per modellare l'ampiezza di un segnale ricevuto quando vi sono numerosi cammini multipli tra trasmettitore e ricevitore

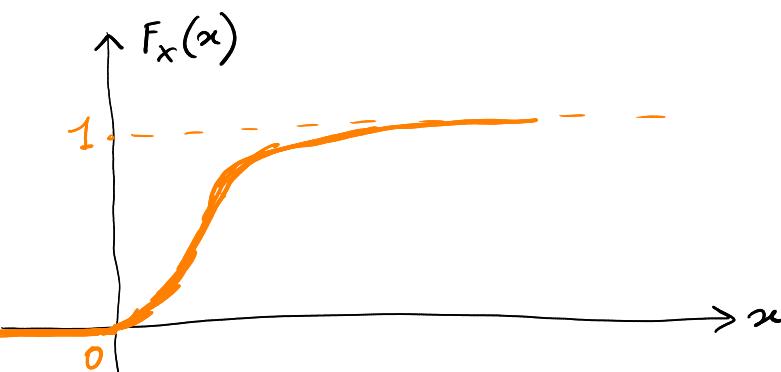
- $F_X(x) = ?$
- Se $\sigma = 7$, $P(X \leq 3) = ?$
- Se $\sigma = 7$, $P(X \leq 3 | X \leq 10) = ?$



$$1) F_x(x) = \int_{-\infty}^x f_x(\alpha) d\alpha = \int_0^x \frac{\alpha}{\sigma^2} \cdot e^{-\frac{\alpha^2}{2\sigma^2}} d\alpha \stackrel{x > 0}{=} \left[-e^{-\frac{\alpha^2}{2\sigma^2}} \right]_0^x = -e^{-\frac{x^2}{2\sigma^2}} + 1$$

/1 \$0, x < 0\$

$$F_x(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\frac{x^2}{2\sigma^2}}, & x \geq 0 \end{cases}$$



$$2) P\{X \leq 3\} = F_x(3) = 1 - e^{-\frac{9}{2 \cdot 49}} = 1 - e^{-0,082} \approx 0,0877$$

$$3) P\{X \leq 3 | X \leq 10\} = \frac{P\{X \leq 3, X \leq 10\}}{P\{X \leq 10\}}$$

~~20~~ 3 10

$$= \frac{P\{X \leq 3\}}{P\{X \leq 10\}} = \frac{0,0877}{0,6386} = 0,137$$

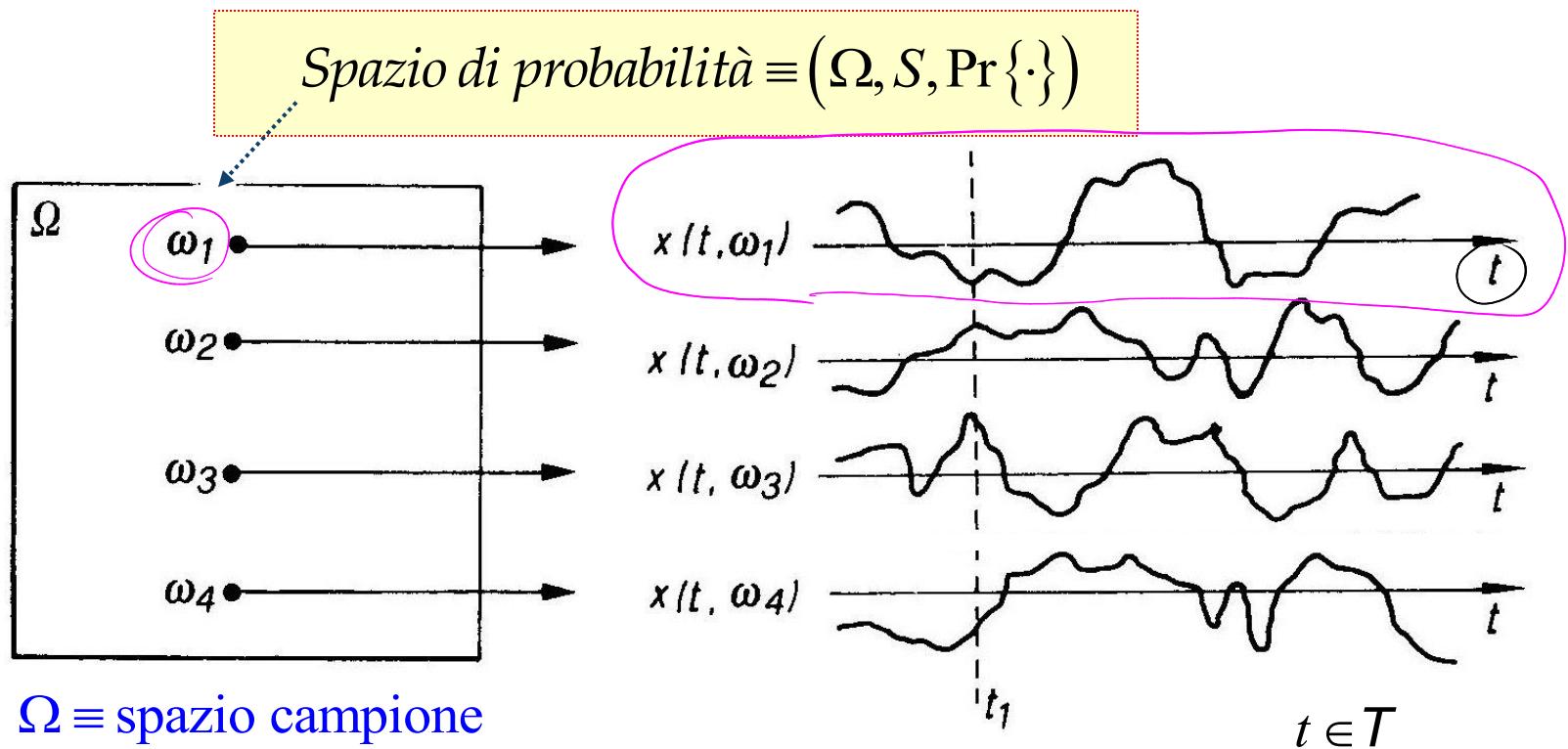
↓

$$F_x(10)$$

Processi aleatori



Definizione di processo aleatorio

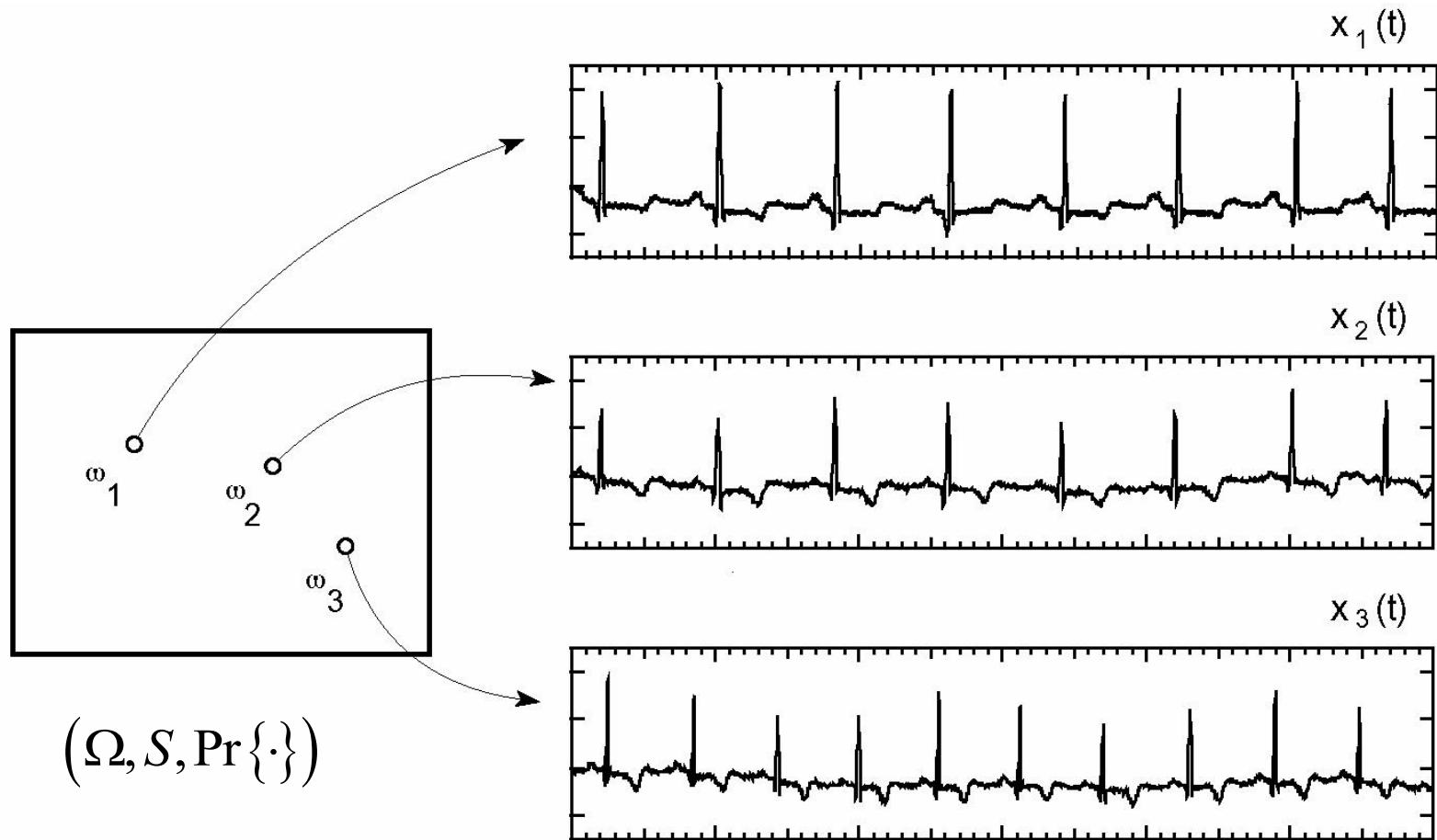


Dato un **esperimento casuale** di modello di probabilità assegnato, ad ogni suo risultato ω_i si associa una funzione reale $x(t, \omega)$ della variabile $t \rightarrow$ Risulta così definito un insieme di funzioni $X(t, \omega)$, detto **processo aleatorio** (o casuale o stocastico), che verrà indicato in breve con $X(t)$, omettendo così la dipendenza da ω



Rappresentazione grafica della definizione di p.a.

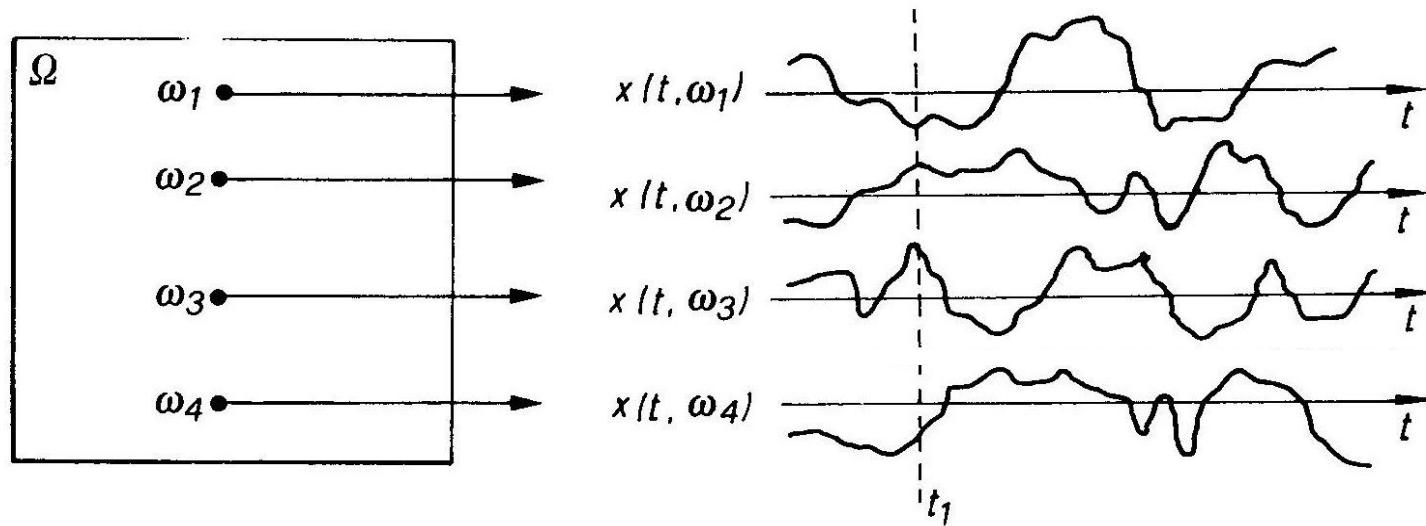
Esempi di elettrocardiogramma in pazienti affetti da aritmia



Segnali che portano informazione sono per loro natura aleatori



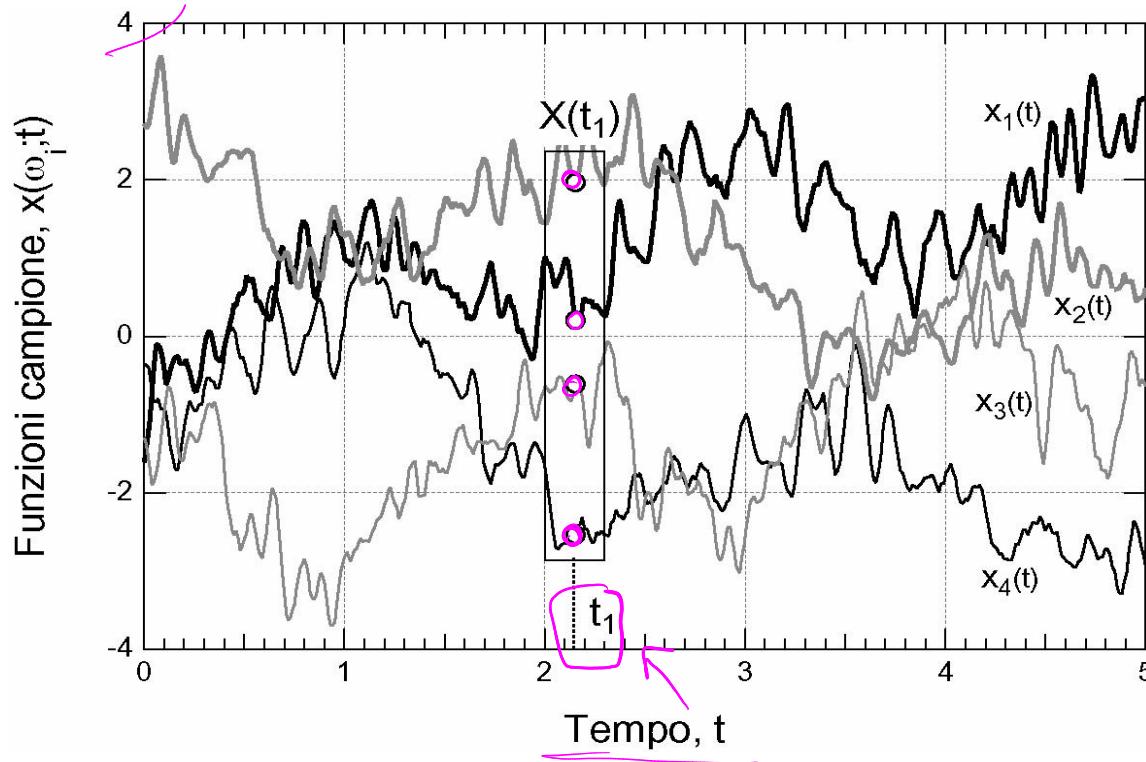
Definizione di processo aleatorio



- Nella maggior parte delle applicazioni t rappresenta il *tempo*
- Le funzioni $x(t, \omega)$ sono funzioni deterministiche, la casualità risiede solo nella presentazione di un particolare risultato dell'esperimento
- Fissato il valore di ω , $X(t, \omega)$ è una funzione deterministica detta **funzione campione** del processo
- La particolare $x(t, \omega)$ che si osserva in una data prova dell'esperimento aleatorio prende il nome di **realizzazione** del processo



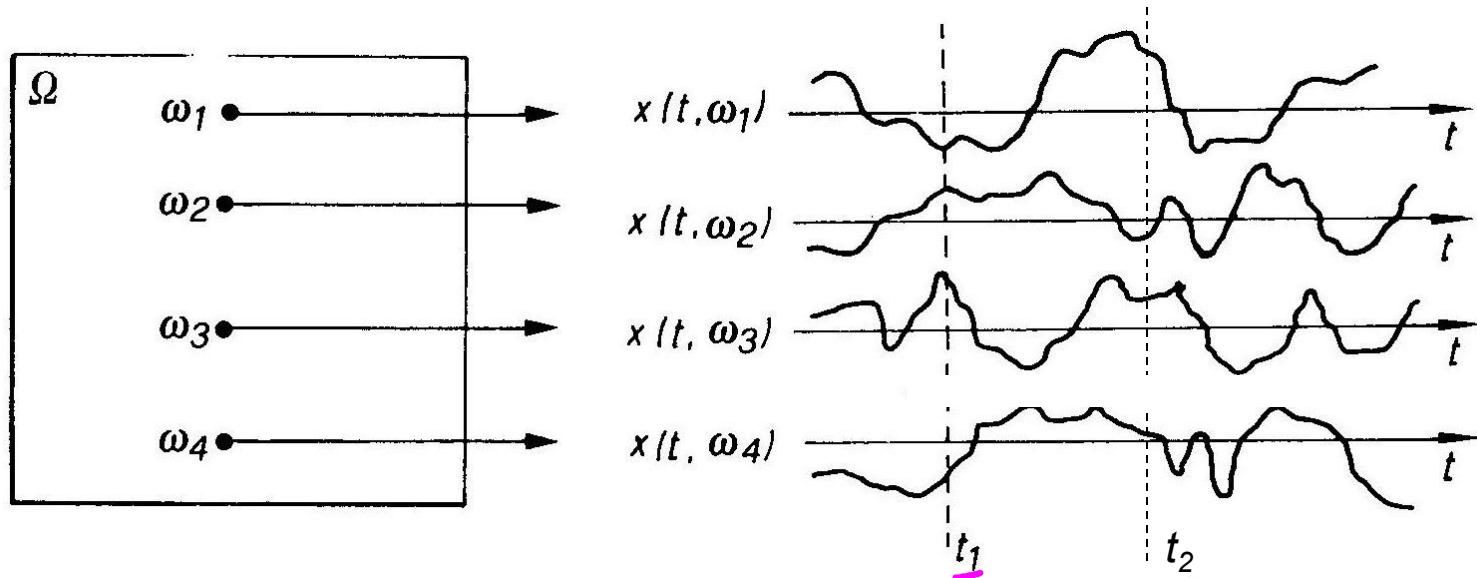
Variabile aleatoria estratta da un p.a.



- Qualora si fissi un determinato istante di tempo t_1 , ad ogni risultato ω dell'esperimento viene associato il valore numerico $x(t_1, \omega)$ della corrispondente realizzazione in quell'istante
- Si ottiene così una quantità dipendente da ω , cioè una v.a. indicata con $X(t_1)$
→ In altre parole, fissato il valore t_1 , il processo casuale $X(t)$ è una variabile aleatoria che indicheremo, per semplicità con $X(t)$



N v.a. estratte da un processo aleatorio



- Se si fissano due istanti distinti t_1 e t_2 si ottengono due distinte v.a. $X(t_1)$ e $X(t_2)$, che costituiscono un sistema di due variabili aleatorie, ovvero il vettore aleatorio $X = [X(t_1) \ X(t_2)]^T$
- Analogamente, fissando N istanti t_1, t_2, \dots, t_N , si ottiene un vettore di N variabili aleatorie $X = [X(t_1) \ X(t_2) \dots X(t_N)]^T$
- La **descrizione statistica** del processo implica perciò la conoscenza della legge di distribuzione di **tutti** i possibili sistemi così formati



Definizione di processo aleatorio

Riassumendo $\underline{X}(t,\omega)$, semplificato in $\underline{X}(t)$, può rappresentare:

- un insieme di funzioni delle variabili t ed ω (**processo aleatorio**)
 - una funzione deterministica della variabile t detta **funzione campione** del processo (ω fissato, t variabile)
 - una **variabile casuale** indicata con $X(t)$ (t fissato, ω variabile)
 - un **numero reale** (t e ω fissati)
-

- In molte applicazioni i risultati dell'esperimento sono già delle forme d'onda; in tal caso non vi è più distinzione tra risultato e funzione campione assegnatagli
- Esempi: misura della tensione di rumore, segnale musicale/video, segnale dati all'uscita di un PC, segnale raccolto da un idrofono



Descrizione statistica di un processo aleatorio

A. Specificazione diretta

Un processo $X(t)$ si dice **statisticamente determinato** se sono note le sue funzioni di distribuzione (CDF):

$$F_X(x_1, x_2, \dots, x_N; t_1, t_2, \dots, t_N) = \Pr\{X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_N) \leq x_N\}$$

per ogni N e per ogni N -upla di istanti t_1, t_2, \dots, t_N

Nota la CDF di ordine N è possibile ricavare tutte le CDF di ordine inferiore mediante le regole marginali (non vale il viceversa).

Nota: la funzione di distribuzione di ordine N del processo è la funzione di distribuzione del vettore di v.a. $\mathbf{X} = [X(t_1) \ X(t_2) \ \dots \ X(t_N)]^T$ ottenuto fissando N istanti t_1, t_2, \dots, t_N

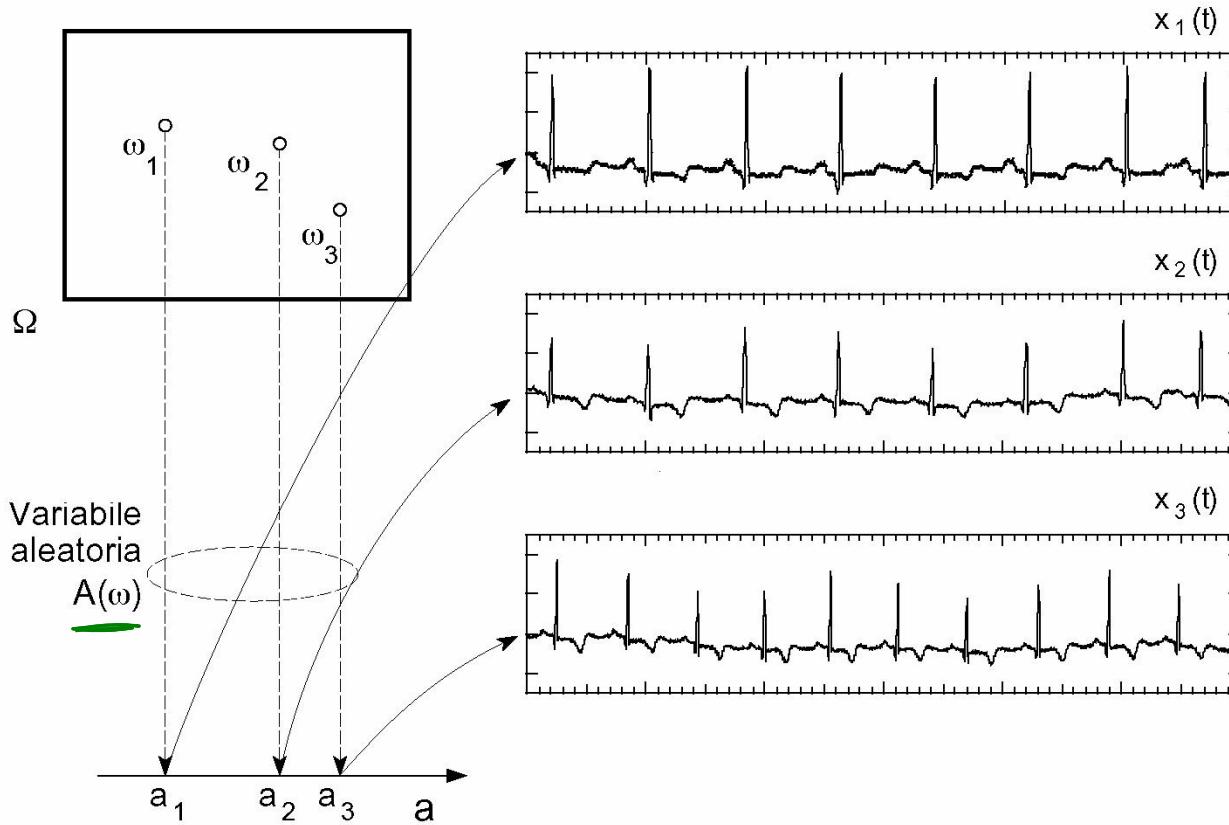
Nota: anche se il comportamento statistico di un processo stocastico è completamente determinato quando sono note le distribuzioni di tutti i possibili ordini, in alcune applicazioni è sufficiente conoscere alcune statistiche dei primi due ordini (**descrizione in potenza del processo**)



Descrizione statistica di un processo aleatorio

B. Specificazione in forma parametrica

Un processo $X(t)$ si dice **parametrico** quando può essere specificato attraverso la forma delle sue funzioni campione, che dipende parametricamente da un certo numero di variabili aleatorie:



$$X(t) = s(t; \Theta_1, \Theta_2, \dots, \Theta_K)$$

La caratterizzazione statistica completa del processo richiede la conoscenza della ddp congiunta dei parametri aleatori:

$$f_{\Theta}(\theta_1, \theta_2, \dots, \theta_K)$$

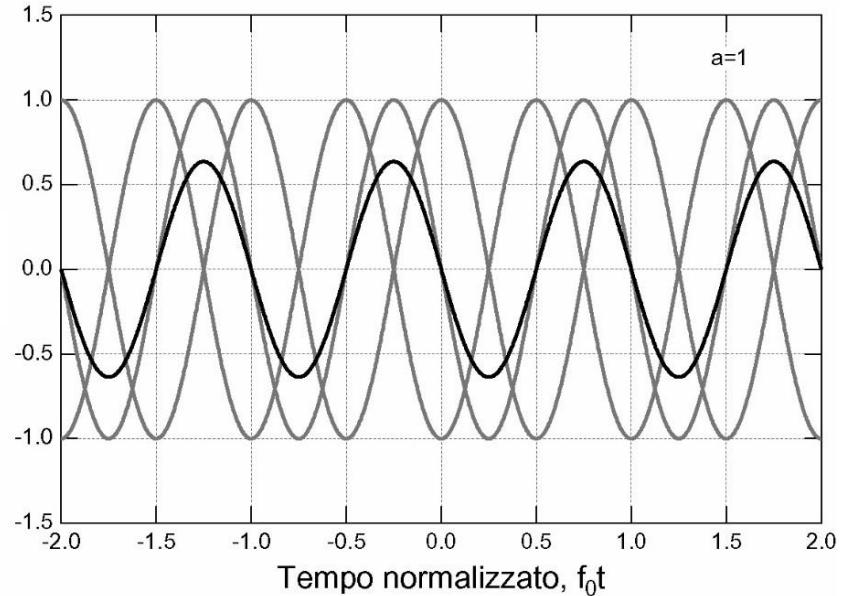


Esempi di p.a. parametrici

- Oscillazione cosinusoidale con fase iniziale incognita

$$X(t) = a \cos(2\pi f_0 t + \Theta)$$

con $\Theta \in U(-\pi, \pi)$



- Oscillazione cosinusoidale con ampiezza e fase iniziale aleatorie:

$$X(t) = A \cos(2\pi f_0 t + \Theta)$$

con $A \in R(\sigma^2)$ e $\Theta \in U(-\pi, \pi)$

