

OSSERVAZIONE SULLA DEFINIZIONE DI UN PROCESSO SOLO ATTRAVERSO LE SUE DISTRIBUZIONI FINITO DIMENSIONALI.

Definizione 0.1. Dato un processo aleatorio $(X_t, t \in I)$ le distribuzioni finito-dimensionali sono la famiglia di misure

$$H \in \mathcal{B}(\mathbb{R}^k) \mapsto \mu_{t_1, t_2, \dots, t_k}(H) := \mathbb{P}\{(X_{t_1}, \dots, X_{t_k}) \in H\}, \text{ al variare di } k \geq 1 \text{ e } t_1, \dots, t_k \in I.$$

Definizione 0.2. Due processi aleatori $(X_t, t \in I)$ e $(Y_t, t \in I)$ sullo stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ si dicono stocasticamente equivalenti se

$$\mathbb{P}(X_t(\omega) = Y_t(\omega)) = 1 \text{ per ogni } t \in I.$$

(Si dice anche che $(Y_t, t \in I)$ è una versione di $(X_t, t \in I)$)

Lemma 0.1. Due processi stocasticamente equivalenti hanno le stesse distribuzioni finito-dimensionali.

Dimostrazione: Basta osservare che

$$\begin{aligned} & \mathbb{P}\{(X_{t_1}, \dots, X_{t_k}) \in H\} = \\ &= \mathbb{P}(\{(X_{t_1}, \dots, X_{t_k}) \in H\} \cap \bigcap_{i=1}^k \{X_{t_i} = Y_{t_i}\}) + \mathbb{P}(\{(X_{t_1}, \dots, X_{t_k}) \in H\} \cap \left(\bigcap_{i=1}^k \{X_{t_i} = Y_{t_i}\}\right)^c) = \\ &= \mathbb{P}(\{(Y_{t_1}, \dots, Y_{t_k}) \in H\} \cap \bigcap_{i=1}^k \{X_{t_i} = Y_{t_i}\}) = \mathbb{P}(\{(Y_{t_1}, \dots, Y_{t_k}) \in H\}) \end{aligned}$$

in quanto

$$\mathbb{P}\left(\left(\bigcap_{i=1}^k \{X_{t_i} = Y_{t_i}\}\right)^c\right) = \mathbb{P}\left(\bigcup_{i=1}^k \{X_{t_i} \neq Y_{t_i}\}\right) \leq \sum_{i=1}^k \mathbb{P}(\{X_{t_i} \neq Y_{t_i}\}) = 0.$$

□

Come abbiamo già notato, le proprietà 1) 2) 3) e 4) della definizione 2 del processo di Poisson permettono di trovare tutte le distribuzioni finito-dimensionali del processo. In alcuni testi viene detto processo di Poisson ogni processo con queste distribuzioni finito-dimensionali. Ciò però non permette di affermare che si tratti di un processo di conteggio. Daremo ora un esempio di processo che ha le stesse distribuzioni finito-dimensionali di N_t , ma le cui traiettorie non sono le tipiche traiettorie di un processo di conteggio, ovvero non sono costanti a tratti, non decrescenti, continue a destra e con salti di ampiezza unitaria.

Si definisca

$$M_t(\omega) := N_t(\omega) + f(t + U_1(\omega))$$

dove

$$\begin{cases} f(s) = 1 & \text{per } s \in \mathbb{Q} \\ f(s) = 0 & \text{per } s \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Essendo $M_t(\omega) \neq N_t(\omega)$ se e solo se $f(t + U_1(\omega)) = 1$ ovvero se e solo se $t + U_1(\omega) \in \mathbb{Q}$ ed inoltre

$$\mathbb{P}(t + U_1(\omega) \in \mathbb{Q}) = \sum_{r \in \mathbb{Q}} \mathbb{P}(U_1(\omega) = r - t) = 0$$

per ogni t si ha che

$$\mathbb{P}(M_t(\omega) = N_t(\omega)) = 1 \text{ per ogni } t \geq 0.$$

1 TEOREMA DI ESISTENZA DI KOLMOGOROV

Un problema noto nel caso unidimensionale è il seguente: data una distribuzione di probabilità μ su \mathbb{R} , esiste uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ ed una variabile aleatoria X che ammette come distribuzione μ ?

Come è noto ci sono due possibili risposte “classiche” a tale quesito: la prima consiste nel prendere lo spazio canonico $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$, X uguale all'identità, cioè $X(x) = x$, e infine $\mathbb{P} = \mu$; la seconda (Teorema di rappresentazione di Skorohod) consiste nel prendere $\Omega = (0, 1)$, $\mathcal{F} = \mathcal{B}(0, 1)$, \mathbb{P} la misura di Lebesgue ristretta a $(0, 1)$ ed $X(u) = F^{-1}(u) = \inf\{x \text{ tali che } u \leq F(x)\}$, cioè la **funzione inversa generalizzata**¹ della funzione di distribuzione $F(x) = \mu((-\infty, x])$.

Anche nel caso dei processi c'è qualcosa di simile. Il problema si esprime nel seguente modo: data una famiglia di distribuzioni finito-dimensionali μ_{t_1, \dots, t_k} , al variare di $k \geq 1$ e di t_1, \dots, t_k in I , esiste uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ ed un processo aleatorio $(X_t, t \geq 0)$ che ammette tali distribuzioni finito-dimensionali?

Lo spazio canonico, analogo ad \mathbb{R} , è $\mathbb{R}^I = \{x(\cdot) : I \rightarrow \mathbb{R}\}$, ed il processo canonico è il processo coordinata per il quale $X_t(x(\cdot)) := x(t)$. La scelta della σ -algebra viene fatta in modo che il processo canonico sia misurabile. Quindi necessariamente deve contenere gli insiemi del tipo

$$\{x(\cdot) \in \mathbb{R}^I \mid x(t) \in A\}, \text{ per } A \in \mathcal{B}(\mathbb{R}),$$

e delle intersezioni finite di insiemi di questo tipo, i rettangoli di base finita

$$\{x(\cdot) \in \mathbb{R}^I \mid x(t_1) \in A_1, \dots, x(t_k) \in A_k, \}, \text{ per } A_h \in \mathcal{B}(\mathbb{R}) \text{ e } h = 1, \dots, k.$$

Necessariamente, quindi, deve contenere anche i cilindri (o insiemi finito-dimensionali), ovvero degli insiemi del tipo

$$\mathcal{C} = \{x(\cdot) \in \mathbb{R}^I \mid (x(t_1), x(t_2), \dots, x(t_k)) \in H\}, \text{ dove } H \in \mathcal{B}(\mathbb{R}^k) \quad (0) \quad (1)$$

Si sceglie quindi la σ -algebra \mathcal{R}^I generata dall'algebra \mathcal{R}_0^I dei cilindri, al variare di $k \geq 1$, degli indici t_1, \dots, t_k in I e di H in $\mathcal{B}(\mathbb{R}^k)$ (il fatto che \mathcal{R}_0^I sia un'algebra sarà mostrato nel seguente teorema).

¹Se non si è familiari con il teorema di Skorohod basta considerare solo il caso in cui la funzione di distribuzione $F(\cdot)$ sia strettamente crescente e continua e quindi invertibile.

Infine come misura di probabilità \mathbb{P} , chiaramente, si vuole prendere una misura di probabilità su \mathcal{R}^I per la quale valga

$$\mathbb{P}(\mathcal{C}) = \mu_{t_1, \dots, t_k}(H),$$

per ogni cilindro \mathcal{C} di \mathcal{R}_0^I , definito come in (1).

A questo punto si apre il problema: esiste una tale probabilità \mathbb{P} ? e più precisamente

- (i) la definizione di \mathbb{P} è ben posta su \mathcal{R}_0^I ? (ii) \mathbb{P} si può estendere a tutto \mathcal{R}^I ?

La risposta è affermativa sotto alcune semplici condizioni di consistenza ed è il contenuto del Teorema di esistenza di Kolmogorov. Come si intuisce dal discorso precedente l'ingrediente essenziale della dimostrazione di tale teorema è il procedimento di estensione di una misura definita su un'algebra alla σ -algebra da essa generata (Teorema di Caratheodory).

Le **condizioni di consistenza** sono le seguenti:

Sia $k > 1$ e sia π una permutazione di $\{1, \dots, k\}$ e sia

$$\Phi_\pi : \mathbb{R}^k \rightarrow \mathbb{R}^k, (x_1, \dots, x_k) \rightarrow \Phi_\pi(x_1, \dots, x_k) := (x_{\pi_1}, \dots, x_{\pi_k}).$$

Chiaramente, essendo Φ_π biunivoca, $(x_1, \dots, x_k) \in H$ se e solo se $\Phi_\pi(x_1, \dots, x_k) \in \Phi_\pi(H)$, per $H \in \mathcal{B}(\mathbb{R}^k)$, e quindi

$$\mu_{t_1, \dots, t_k}(H) = \mu_{t_{\pi_1}, \dots, t_{\pi_k}}(\Phi_\pi(H)), \quad (1) \quad (2)$$

infatti

$$\begin{aligned} \mu_{t_1, \dots, t_k}(H) &= \mathbb{P}\{(X_{t_1}, \dots, X_{t_k}) \in H\} = \mathbb{P}\{\Phi_\pi(X_{t_1}, \dots, X_{t_k}) \in \Phi_\pi(H)\} = \\ &= \mathbb{P}\{(X_{t_{\pi_1}}, \dots, X_{t_{\pi_k}}) \in \Phi_\pi(H)\} = \mu_{t_{\pi_1}, \dots, t_{\pi_k}}(\Phi_\pi(H)), \end{aligned}$$

inoltre

$$\mu_{t_1, \dots, t_k, t_{k+1}}(H \times \mathbb{R}) = \mu_{t_1, \dots, t_k}(H). \quad (2) \quad (3)$$

È immediato constatare che le condizioni di consistenza (2) e (3) si possono riscrivere nel seguente modo:

Siano $k \geq h \geq 1$, sia π una permutazione di $\{1, \dots, k\}$ e sia $\Psi_{\pi, h} : \mathbb{R}^k \rightarrow \mathbb{R}^h, (x_1, \dots, x_k) \rightarrow \Psi_{\pi, h}(x_1, \dots, x_k) := (x_{\pi_1}, \dots, x_{\pi_h})$, allora, per ogni $(t_1, \dots, t_k) \in I^k$ e $H \in \mathcal{B}(\mathbb{R}^h)$

$$\mu_{t_1, \dots, t_k}(\Psi_{\pi, h}^{-1}(H)) = \mu_{t_{\pi_1}, \dots, t_{\pi_h}}(H), \quad (3) \quad (4)$$

(infatti, nel caso in cui H sia un rettangolo cioè $H = A_1 \times \dots \times A_h$, posto π^{-1} la permutazione inversa di π , pensata quindi come funzione biunivoca di $\{1, \dots, k\}$ in sé, e posto $A_m = \mathbb{R}$ per $m = h+1, \dots, k$, la relazione (3) diviene

$$\begin{aligned} \mathbb{P}((X_{t_{\pi_1}}, \dots, X_{t_{\pi_h}}) \in H) &= \mathbb{P}(X_{t_{\pi_1}} \in A_1, \dots, X_{t_{\pi_h}} \in A_h) = \\ &= \mathbb{P}(X_{t_{\pi_1}} \in A_1, \dots, X_{t_{\pi_h}} \in A_h, X_{t_{\pi_{h+1}}} \in A_{h+1}, \dots, X_{t_{\pi_k}} \in A_k) = \\ &= \mathbb{P}(X_{t_{\pi_1}} \in A_{\pi^{-1}(\pi_1)}, \dots, X_{t_{\pi_h}} \in A_{\pi^{-1}(\pi_h)}, X_{t_{\pi_{h+1}}} \in A_{\pi^{-1}(\pi_{h+1})}, \dots, X_{t_{\pi_k}} \in A_{\pi^{-1}(\pi_k)}) = \end{aligned}$$

(riordinando opportunamente le condizioni richieste)

$$\begin{aligned} &= \mathbb{P}(X_{t_1} \in A_{\pi^{-1}(1)}, \dots, X_{t_h} \in A_{\pi^{-1}(h)}, X_{t_{h+1}} \in A_{\pi^{-1}(h+1)}, \dots, X_{t_k} \in A_{\pi^{-1}(k)}) = \\ &= \mathbb{P}((X_{t_1}, \dots, X_{t_k}) \in \Psi_{\pi, h}^{-1}(H)), \end{aligned}$$

e se vale per i rettangoli, poi vale per ogni cilindro.

Prima di enunciare e dimostrare il teorema di Kolmogorov, menzioniamo il fatto che esiste un'altra possibilità, più simile a quanto fatto nel teorema di rappresentazione di Skorohod, nel caso in cui I è numerabile, cioè $I = \mathbb{N}$, ed illustreremo questo caso in seguito. In tale caso lo spazio canonico è di nuovo $(0,1)$ con la σ -algebra dei boreliani $\mathcal{B}(0,1)$ e la misura di Lebesgue ristretta a $(0,1)$, e quindi, a partire dalle distribuzioni finito-dimensionali non deve essere definita la misura di probabilità, bensì il processo stesso. Ingrediente essenziale per la dimostrazione è il teorema di esistenza delle versioni regolari delle distribuzioni condizionali di una variabile aleatoria, dato una variabile aleatoria multidimensionale.

Cominciamo con l'enunciare il Teorema di Kolmogorov

Teorema 1.1 (di Kolmogorov). *Sia data una famiglia μ_{t_1, \dots, t_k} di distribuzioni finito-dimensionali consistente, cioè che verifica le condizioni di consistenza (2) e (3), allora esiste uno spazio di probabilità ed un processo aleatorio che ammette μ_{t_1, \dots, t_k} come distribuzioni finito-dimensionali. Inoltre è sempre possibile prendere come spazio di probabilità lo spazio canonico \mathbb{R}^I e come processo il processo canonico $X_t(x(\cdot)) = x(t)$.*

Dimostrazione: Cominciamo con l'ultima parte del teorema. Sia $(\Omega, \mathcal{F}, \mathbb{Q})$ uno spazio di probabilità ed $(Y_t, t \in I)$ un processo su tale spazio che ammette come distribuzioni finito-dimensionali μ_{t_1, \dots, t_k} . Si definisca la funzione $\xi : \Omega \rightarrow \mathbb{R}^I, \omega \rightarrow \xi(\omega)(\cdot) = Y(\cdot, \omega)$, cioè $\xi(\omega)$ è la funzione $t \rightarrow \xi(\omega)(t) = Y_t(\omega)$.

Risulta che ξ è una funzione $\mathcal{F}/\mathcal{R}^I$ misurabile: infatti per ogni cilindro \mathcal{C} come in (1), l'insieme

$$\xi^{-1}(\mathcal{C}) = \{\omega \text{ tali che } (Y_{t_1}, \dots, Y_{t_k}) \in H\} \in \mathcal{F},$$

e ciò è sufficiente a dimostrare la misurabilità di ξ .

A questo punto ponendo

$$\mathbb{P}(\mathcal{A}) = \mathbb{Q}(\xi^{-1}(\mathcal{A})), \text{ per } \mathcal{A} \in \mathcal{R}^I,$$

si ottiene la rappresentazione canonica.

Continuiamo dando lo **schema della dimostrazione**

punto 1) \mathcal{R}_0^I è un'algebra.

punto 2) La definizione \mathbb{P} su \mathcal{R}_0^I come $\mathbb{P}(\mathcal{C}) := \mu_{t_1, \dots, t_k}(H)$, dove \mathcal{C} è il cilindro definito in (1), è ben posta e risulta \mathbb{P} una misura di probabilità finitamente additiva su \mathcal{R}_0^I .

punto 3) La misura \mathbb{P} è numerabilmente additiva su \mathcal{R}_0^I , e quindi si può estendere in modo univoco ad una misura alla σ -algebra \mathcal{R}^I generata da \mathcal{R}_0^I .

Sia il punto 1) che il punto 2) sono basati sulla osservazione che un cilindro \mathcal{C} può avere più di una rappresentazione:

$$\mathcal{C} = \{x(\cdot) \in \mathbb{R}^I \mid (x(t_1), x(t_2), \dots, x(t_k)) \in H\}, \text{ con } H \in \mathcal{B}(\mathbb{R}^k),$$

oppure

$$\mathcal{C} = \{x(\cdot) \in \mathbb{R}^I \mid (x(s_1), x(s_2), \dots, x(s_m)) \in J\}, \text{ con } J \in \mathcal{B}(\mathbb{R}^m).$$

Notazione: per k e t_1, \dots, t_k prefissati, indicheremo con $\mathcal{R}^{\{t_1, \dots, t_k\}}$ la famiglia dei cilindri del tipo precedente al variare di $H \in \mathcal{B}(\mathbb{R}^k)$.

Supponiamo che $k \leq m$, e che $\{t_1, \dots, t_k\} \subseteq \{s_1, \dots, s_m\}$, allora esiste una permutazione σ di $\{1, \dots, m\}$ per cui $t_i = s_{\sigma_i}$, $i = 1, \dots, k$, per cui $\Psi_{\sigma, k}(J) = H$ e per cui $J = \Psi_{\pi, h}^{-1}(H)$ o, equivalentemente $\Phi_{\sigma}(J) = H \times \mathbb{R}^{m-k}$.

Si noti allora che dalla (4) si ottiene che

$$\mu_{s_1, \dots, s_m}(J) = \mu_{s_1, \dots, s_m}(\Psi_{\sigma, k}^{-1}(H)) = \mu_{s_{\sigma_1}, \dots, s_{\sigma_k}}(H) = \mu_{t_1, \dots, t_k}(H). \quad (4) \quad (5)$$

Inoltre, due cilindri \mathcal{C} e \mathcal{C}' , con indici di tempo $\{t_1, \dots, t_k\}$ e $\{t'_1, \dots, t'_{k'}\}$ rispettivamente, si possono sempre rappresentare come cilindri con indici comuni $\{s_1, \dots, s_m\} = \{t_1, \dots, t_k\} \cup \{t'_1, \dots, t'_{k'}\}$.

Più precisamente se

$$\mathcal{C} = \{x(\cdot) \in \mathbb{R}^I \mid (x(t_1), x(t_2), \dots, x(t_k)) \in H\} \in \mathcal{R}^{\{t_1, \dots, t_k\}}$$

e

$$\mathcal{C}' = \{x(\cdot) \in \mathbb{R}^I \mid (x(t'_1), x(t'_2), \dots, x(t'_{k'})) \in H'\} \in \mathcal{R}^{\{t'_1, \dots, t'_{k'}\}}$$

allora, posto

$$\{s_1, \dots, s_m\} = \{t_1, \dots, t_k\} \cup \{t'_1, \dots, t'_{k'}\}$$

si può supporre, senza ledere in generalità (si tratta eventualmente di ricorrere a permutazioni opportune), che

$$\{s_1, \dots, s_h, s_{h+1}, \dots, s_k\} = \{t_1, \dots, t_k\}, \{s_{h+1}, \dots, s_m\} = \{t'_1, \dots, t'_{k'}\}$$

ed infine che

$$\{s_{h+1}, \dots, s_m\} = \{t'_1, \dots, t'_{k'}\}.$$

Infatti, per due opportune permutazioni π di $\{1, \dots, k\}$ e π' di $\{1, \dots, k'\}$, si può riscrivere $(t_{\pi_1}, \dots, t_{\pi_k}) = (s_1, \dots, s_k)$ e $(t'_{\pi'_1}, \dots, t'_{\pi'_k}) = (s_{h+1}, \dots, s_m)$ e quindi

$$\begin{aligned} \mathcal{C} &= \{x(\cdot) \in \mathbb{R}^I \mid (x(s_1), x(s_2), \dots, x(s_k)) \in \Phi_{\pi}(H)\} \\ &= \{x(\cdot) \in \mathbb{R}^I \mid (x(s_1), \dots, x(s_m)) \in \Phi_{\pi}(H) \times \mathbb{R}^{m-k}\} \in \mathcal{R}^{\{s_1, \dots, s_m\}} \end{aligned} \quad (5) \quad (6)$$

e

$$\begin{aligned} \mathcal{C}' &= \{x(\cdot) \in \mathbb{R}^I \mid (x(s_{h+1}), x(s_{h+2}), \dots, x(s_m)) \in \Phi_{\pi'}(H')\} \\ &= \{x(\cdot) \in \mathbb{R}^I \mid (x(s_1), \dots, x(s_m)) \in \mathbb{R}^h \times \Phi_{\pi'}(H')\} \in \mathcal{R}^{\{s_1, \dots, s_m\}} \end{aligned} \quad (6) \quad (7)$$

Dimostrazione in dettaglio dei punti.

punto 1)

È immediato capire che, fissato $k \geq 1$ e $\{t_1, \dots, t_k\}$, la famiglia $\mathcal{R}^{\{t_1, \dots, t_k\}}$ dei cilindri del tipo

$$\mathcal{C} = \{x(\cdot) \in \mathbb{R}^I \mid (x(t_1), x(t_2), \dots, x(t_k)) \in H\},$$

con $H \in \mathcal{B}(\mathbb{R}^k)$, forma un'algebra (anche se in realtà si tratta di una σ -algebra):

Per ottenere \mathbb{R}^I , basta prendere $H = \mathbb{R}^k$.

Per ottenere \mathcal{C}^c , basta prendere H^c .

Per ottenere $\mathcal{C} \cup \mathcal{C}'$, con $\mathcal{C}' = \{x(\cdot) \in \mathbb{R}^I \mid (x(t_1), x(t_2), \dots, x(t_k)) \in H'\}$, basta prendere $H \cup H'$.

Dalle osservazioni precedenti sappiamo che comunque presi due cilindri essi si possono esprimere come cilindri con indici comuni. Si ottiene che quindi \mathcal{R}_0^I è un'algebra.

punto 2)

Le osservazioni iniziali e le proprietà di consistenza permettono di ottenere immediatamente che la definizione è ben posta. Infatti se i cilindri \mathcal{C} e \mathcal{C}' di (6) e (7) sono uguali allora necessariamente $\Phi_\pi(H) \times \mathbb{R}^{m-k} = \mathbb{R}^h \times \Phi_{\pi'}(H') = J$, e allora per la (4)

$$\mu_{s_1, \dots, s_m}(J) = \mu_{t_1, \dots, t_k}(H) = \mu_{t'_1, \dots, t'_k}(H').$$

La finita additività dipende dal fatto che ciascuna μ_{t_1, \dots, t_k} è una misura di probabilità su $\mathcal{R}^{\{t_1, \dots, t_k\}}$ e che, dato un numero finito di cilindri di \mathcal{R}_0^I , si può sempre pensare che tutti i cilindri siano appartenenti ad un'algebra del tipo $\mathcal{R}^{\{t_1, \dots, t_k\}}$.

punto 3)

Per provare la σ -additività è sufficiente mostrare la continuità, e per questo, a sua volta, è sufficiente mostrare che se $\mathcal{A}_n \in \mathcal{R}_0^I$ ed $\mathcal{A}_n \downarrow \emptyset$ allora $\mathbb{P}(\mathcal{A}_n) \downarrow 0$. La prova procede per assurdo. Esista, per assurdo, un $\epsilon > 0$ tale che $\mathbb{P}(\mathcal{A}_n) \geq \epsilon$ per ogni n . Se mostriamo che allora $\bigcap_n \mathcal{A}_n$ non può essere l'insieme vuoto la dimostrazione è completa.

Poiché per rappresentare un cilindro si può sempre aumentare il numero degli indici, cioè degli istanti di tempo coinvolti nella sua definizione, possiamo supporre che esista una successione $\{t_n, n \geq 1\}$ e una successione di boreliani $H_n \in \mathcal{B}(\mathbb{R}^n)$ per cui

$$\mathcal{A}_n = \{x \in \mathbb{R}^I, (x(t_1), \dots, x(t_n)) \in H_n\}$$

(eventualmente ripetendo qualche \mathcal{A}_n , infatti in generale ciascun \mathcal{A}_n coinvolgerà gli istanti t_1, \dots, t_{m_n} ; in tale caso si dovrà ripetere \mathbb{R}^I per $m_1 - 1$ volte, con $H_n = \mathbb{R}_n$ per $n = 1, \dots, m_1 - 1, \mathcal{A}_1$ per $m_2 - m_1$ volte e così via).

Ovviamente $\mathbb{P}(\mathcal{A}_n) = \mu_{t_1, \dots, t_n}(H_n)$, ed essendo μ_{t_1, \dots, t_n} una **misura internamente regolare**¹ è possibile trovare un compatto $K_n \subseteq H_n$ per cui

$$\mu_{t_1, \dots, t_n}(H_n \setminus K_n) \leq \frac{\epsilon}{2^{n+1}},$$

di modo che se $\mathcal{B}_n = \{x \in \mathbb{R}^I, (x(t_1), \dots, x(t_n)) \in K_n\}$ allora

$$\mathbb{P}(\mathcal{A}_n \setminus \mathcal{B}_n) \leq \frac{\epsilon}{2^{n+1}}.$$

Posto $\mathcal{C}_n = \bigcap_{m=1}^n \mathcal{B}_m$, di modo che

$$\mathcal{C}_n = \{x \in \mathbb{R}^I, (x(t_1), \dots, x(t_n)) \in \Gamma_n\}, \text{ con } \Gamma_n = \bigcap_{m=1}^n (K_m \times \mathbb{R}^{n-m}) \subseteq K_n \text{ compatto,}$$

allora

$$\mathcal{C}_n \subseteq \mathcal{B}_n \subseteq \mathcal{A}_n \text{ e } \mathbb{P}(\mathcal{A}_n \setminus \mathcal{C}_n) < \frac{\epsilon}{2}.$$

Infatti

$$\begin{aligned} \mathbb{P}(\mathcal{A}_n \setminus \mathcal{C}_n) &= \mathbb{P}(\mathcal{A}_n \cap \left(\bigcap_{m=1}^n \mathcal{B}_m\right)^c) = \mathbb{P}\left(\bigcup_{m=1}^n \mathcal{A}_n \cap \mathcal{B}_m^c\right) \leq \\ &\leq \sum_{m=1}^n \mathbb{P}(\mathcal{A}_m \cap \mathcal{B}_m^c) = \sum_{m=1}^n \mathbb{P}(\mathcal{A}_m \setminus \mathcal{B}_m) \leq \sum_{m=1}^n \frac{\epsilon}{2^{m+1}} < \frac{\epsilon}{2}. \end{aligned}$$

Di conseguenza $\mathbb{P}(\mathcal{C}_n) \geq \frac{\epsilon}{2} > 0$ (essendo $\mathbb{P}(\mathcal{A}_n) \geq \epsilon$ e $\mathbb{P}(\mathcal{A}_n \setminus \mathcal{C}_n) < \frac{\epsilon}{2}$). Quindi \mathcal{C}_n non è vuoto e lo stesso vale per $\Gamma_n \subseteq K_n$.

Si scelga per ogni n un punto appartenente al cilindro \mathcal{C}_n , ovvero una funzione, $x^{(n)}(\cdot) \in \mathcal{C}_n$. Se $n \geq k$ allora $x^{(n)}(\cdot) \in \mathcal{C}_n \subseteq \mathcal{C}_k \subseteq \mathcal{B}_k$ e quindi il punto $(x^{(n)}(t_1), \dots, x^{(n)}(t_k)) \in K_k \subseteq J_k \times J_k \times \dots \times J_k$ (k volte, con J_k un intervallo limitato). Di conseguenza, ciascuna successione del tipo $\{x^{(1)}(t_k), x^{(2)}(t_k), \dots, x^{(n)}(t_k), \dots\}$ è contenuta in J_k , e quindi è limitata.

In realtà la successione $\{x^{(1)}(t_k), x^{(2)}(t_k), \dots, x^{(n)}(t_k), \dots\}$ è contenuta in J_k solo definitivamente, anzi $x^{(n)}(t_k) \in J_k$ per $n \geq k$, comunque è una successione che ammette almeno una sottosuccessione convergente. Anche questo è un punto da tenere presente nel caso in cui si voglia fare una generalizzazione a processi a valori in spazi più generali di \mathbb{R} o \mathbb{R}^d .

Con il metodo diagonale si può scegliere una successione n_h crescente in modo che, qualunque sia $k \geq 1$, la successione $\{x^{(n_h)}(t_k)\}_{h \geq 1}$ sia convergente ad un punto y_k . Si noti che per ogni

¹Ricordiamo che, se S è uno spazio metrico, si dice che una misura μ sui boreliani di S è **internamente regolare** se per ogni insieme misurabile B si ha che $\mu(B) = \sup\{\mu(J), J \subset B, J \text{ compatto}\}$, e quindi esiste una successione di compatti J_h contenuti in B per cui la misura di B è il limite delle misure di J_h .

Comunque in \mathbb{R}^n ciò significa che per ogni boreliano H è possibile trovare una successione monotona di iper-rettangoli chiusi e limitati $\{K_m\}_{m \geq 1}$ e convergente ad H . Per la proprietà di continuità delle probabilità vale allora $\mu(K_m) \uparrow \mu(H)$.

Questo è un punto cruciale nella dimostrazione e che vale per tutte le misure di probabilità su uno spazio metrico localmente compatto. Anzi ciò vale addirittura per spazi di Hausdorff sempre localmente compatto (confrontare, ad esempio, il Teorema di rappresentazione di Riesz). Questa osservazione permette di estendere il teorema di Kolmogorov non solo al caso di processi a valori reali, ma anche al caso di processi a valori in spazi metrici localmente compatti, ed in particolare a \mathbb{R}^d .

$k \geq 1$ il punto $(y_1, \dots, y_k) \in K_k$, in quanto definitivamente $(x^{(n)}(t_1), \dots, x^{(n)}(t_k)) \in K_k$.

Richiamo sul metodo diagonale:

Sia data una successione a due indici

$$\begin{array}{cccc} x_{1,1}, & x_{1,2}, & \cdots & \cdots \\ x_{2,1}, & x_{2,2}, & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ x_{n,1}, & x_{n,2}, & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{array}$$

per cui, qualunque sia n , la successione $\{x_{n,h}\}_{h \geq 1}$ è limitata (o relativamente compatta). Allora esiste una successione $\{h_m\}_{m \geq 1}$ di interi per cui, qualunque sia n , la successione $\{x_{n,h_m}\}_{m \geq 1}$ è convergente.

Si considera la sottosuccessione convergente $\{x_{1,h_{1,m}}\}_{m \geq 1}$ ad y_1 . Poi si considera la successione $\{x_{2,h_{1,m}}\}_{m \geq 1}$, che essendo limitata ammette una sottosuccessione convergente $\{x_{2,h_{2,m}}\}_{m \geq 1}$ ad y_2 . Si noti che quindi anche $\{x_{1,h_{2,m}}\}_{m \geq 1}$ è ancora una successione convergente ad y_1 e che anche ogni sua sottosuccessione è convergente ad y_1 . Si continua in questo modo fino ad ottenere che $\{x_{n,h_{n,m}}\}_{m \geq 1}$ è una successione convergente ad y_n e $\{h_{n,m}\}_{m \geq 1}$ è una sottosuccessione di $\{h_{n-1,m}\}_{m \geq 1}$, ed anche ogni sottosuccessione di $\{x_{n,h_{n,m}}\}_{m \geq 1}$ converge ad y_n . A questo punto abbiamo una nuova configurazione

$$\begin{array}{cccc} x_{1,h_{1,1}}, & x_{1,h_{1,2}}, & \cdots & \rightarrow y_1 \\ x_{2,h_{2,1}}, & x_{2,h_{2,2}}, & \cdots & \rightarrow y_2 \\ \cdots & \cdots & \cdots & \cdots \\ x_{n,h_{n,1}}, & x_{n,h_{n,n}}, & \cdots & \rightarrow y_n \end{array}$$

Basta prendere $h_m = h_{m,m}$, infatti, qualunque sia n , la successione $\{x_{n,h_{m,m}}\}_{m \geq 1}$ è (definitivamente) una sottosuccessione di $\{x_{n,h_{n,m}}\}_{m \geq 1}$ e quindi converge ad y_n , per m che tende ad infinito.

Perciò, ogni funzione $x(\cdot)$ di \mathbb{R}^I tale che $x(t_h) = y_h$, per ogni $h \geq 1$, appartiene ad ogni \mathcal{B}_k , $k \geq 1$, e quindi ad ogni \mathcal{A}_k , $k \geq 1$, e quindi $\bigcap_n \mathcal{A}_n$ risulta un insieme non vuoto.

Si noti che l'argomento della prova è lo stesso della dimostrazione del teorema di Tikhonov, cioè che il prodotto infinito di compatti è un compatto. Una dimostrazione simile appare anche nella dimostrazione che l'intersezione di una successione di compatti non vuoti è non vuota.

1.1 ESEMPI

1) PROCESSI A COORDINATE INDIPENDENTI

Data una famiglia di distribuzioni $\{\mu_t, t \in I\}$ ad un tempo, si consideri la famiglia di distribuzioni finito-dimensionali $\mu_{t_1, \dots, t_k} := \mu_{t_1} \times \cdots \times \mu_{t_k}$. Tale famiglia è chiaramente una famiglia consistente e quindi esiste un processo con tali distribuzioni finito-dimensionali. Si noti che non si richiede che I sia numerabile.

2) PROCESSI GAUSSIANI

Siano date una funzione $m : I \rightarrow \mathbb{R}$, $t \rightarrow m(t)$ e una funzione $K : I \times I \rightarrow \mathbb{R}$, $(t, s) \rightarrow K(t, s)$ definita non negativa, cioè tale che comunque scelti $n \geq 1$, $t_1, \dots, t_n \in I$, $\eta_1, \dots, \eta_n \in \mathbb{C}$ valga

$$\sum_{h,k}^{1,n} \eta_h \bar{\eta}_k K(t_h, t_k) \geq 0.$$

Si noti che quindi necessariamente $K(t, s) = K(s, t)$ e $K(t, t) \geq 0$.

Sia μ_{t_1, \dots, t_k} la distribuzione Gaussiana di media $(m(t_1), \dots, m(t_k))$ e matrice di covarianza Γ definita da $\Gamma_{i,j} = K(t_i, t_j)$ per $i, j = 1, \dots, k$.

Per l'esistenza di tali distribuzioni finito-dimensionali si consideri il caso in cui $K(t, s)$ sia strettamente definita positiva e si noti che se (Y_1, \dots, Y_k) è un vettore aleatorio con componenti indipendenti e ciascuna con distribuzione normale $N(0, 1)$, allora il vettore definito da $Z = (Z_1, \dots, Z_k) = A(Y_1, \dots, Y_k) + (m(t_1), \dots, m(t_k)) = AY + m$, dove $A = \Gamma^{1/2}$ (cioè $\Gamma = A^t A = AA^t$), è un vettore con la distribuzione cercata.

Per la consistenza della famiglia μ_{t_1, \dots, t_k} così definita, si noti che l'operatore Φ_π di (2) è una trasformazione lineare e che trasformazioni lineari di vettori gaussiani sono ancora gaussiani. Inoltre, indicando ancora con Φ_π la matrice associata all'operatore di (2), allora $\Phi_\pi Z = \Phi_\pi AY + \Phi_\pi m$, segue una legge gaussiana di media $\Phi_\pi m = (m(t_{\pi_1}), \dots, m(t_{\pi_k}))$ e con matrice di covarianza $(\Phi_\pi A)(\Phi_\pi A)^t = \Phi_\pi AA^t \Phi_\pi^t = \Phi_\pi \Gamma \Phi_\pi^t = (\Gamma_{\pi_i, \pi_j})_{i,j=1, \dots, k}$.

Da queste osservazioni si deduce immediatamente che $\mu_{t_1, \dots, t_k}(H) = \mu_{t_{\pi_1}, \dots, t_{\pi_k}}(\Phi_\pi(H))$, ovvero la (2).

Per quanto riguarda la (32), ogni sottovettore di un vettore gaussiano è ancora un vettore gaussiano.

In particolare possiamo ad esempio stabilire l'esistenza del processo di Wiener standard, detto anche Moto Browniano, cioè del processo gaussiano con $m(t) = 0$ e $K(t, s) = t \wedge s$.

Bisognerebbe ovviamente controllare che la funzione $K(\cdot, \cdot)$ sia definita non negativa. Ciò può essere fatto direttamente, ma con una certa fatica.

Un sistema decisamente più semplice è il seguente: consiste nell'osservare che la funzione di correlazione

$$K(t, s) := Cov(N_t, N_s)$$

di un processo N_t di Poisson standard (cioè con $\lambda = 1$) è proprio $t \wedge s$, ciò è sufficiente a garantire la sua non negatività. Infatti

$$\begin{aligned} Cov(N_t, N_s) &= Cov(N_{t \wedge s}, N_{t \vee s}) = \\ &= \mathbb{E}[N_{t \wedge s}(N_{t \wedge s} + N_{t \vee s} - N_{t \wedge s})] - \mathbb{E}[N_{t \wedge s}]\mathbb{E}[N_{t \wedge s} + N_{t \vee s} - N_{t \wedge s}] = \\ &= \mathbb{E}[N_{t \wedge s} N_{t \wedge s}] + \mathbb{E}[N_{t \wedge s}(N_{t \vee s} - N_{t \wedge s})] - \mathbb{E}[N_{t \wedge s}](\mathbb{E}[N_{t \wedge s}] + \mathbb{E}[N_{t \vee s} - N_{t \wedge s}]) = \\ &= Var(N_{t \wedge s}) + \mathbb{E}[N_{t \wedge s}(N_{t \vee s} - N_{t \wedge s})] - \mathbb{E}[N_{t \wedge s}]\mathbb{E}[N_{t \vee s} - N_{t \wedge s}] = Var(N_{t \wedge s}) = t \wedge s \end{aligned}$$

Inoltre la funzione di correlazione di un processo X_t qualsiasi è sempre definita non negativa:

$$\begin{aligned}
\sum_{h,k}^{1,n} \eta_h \bar{\eta}_k K(t_h, t_k) &= \sum_{h,k}^{1,n} \eta_h \bar{\eta}_k \mathbb{E} \left[(X_{t_h} - \mathbb{E}[X_{t_h}])(X_{t_k} - \mathbb{E}[X_{t_k}]) \right] = \\
&= \mathbb{E} \left[\sum_{h,k}^{1,n} \eta_h \bar{\eta}_k (X_{t_h} - \mathbb{E}[X_{t_h}])(X_{t_k} - \mathbb{E}[X_{t_k}]) \right] = \\
&= \mathbb{E} \left[\sum_{h=1}^n \eta_h (X_{t_h} - \mathbb{E}[X_{t_h}]) \sum_{k=1}^n \bar{\eta}_k (X_{t_k} - \mathbb{E}[X_{t_k}]) \right] = \\
&= \mathbb{E} \left[\left| \sum_{h=1}^n \eta_h (X_{t_h} - \mathbb{E}[X_{t_h}]) \right|^2 \right] \geq 0.
\end{aligned}$$

Si noti che la proprietà del processo di Wiener di avere la stessa funzione di correlazione del processo di Poisson standard, implica che gli incrementi sono non correlati. Poiché gli incrementi del processo di Wiener W_t hanno distribuzione gaussiana, sono quindi anche indipendenti. Inoltre $\mathbb{P}(W_0 = 0) = 1$.

3) PROCESSI DI MARKOV REGOLARI

I processi di Markov regolari sono processi costruiti attraverso una **famiglia di probabilità di transizione regolari**, cioè una famiglia $p(s, t, x, A)$

$$p : (\mathbb{R}^+ \times \mathbb{R}^+)_+ \times \mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1], (s, t, x, A) \rightarrow p(s, t, x, A)$$

dove

$$(\mathbb{R}^+ \times \mathbb{R}^+)_+ = \{(s, t) \in \mathbb{R}^+ \times \mathbb{R}^+, \text{ tali che } s \leq t\},$$

che soddisfa le seguenti proprietà:

(i) per ogni $0 \leq s \leq t$ ed $x \in \mathbb{R}$,

$$p(s, t, x, \cdot) : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1], A \rightarrow p(s, t, x, A)$$

è una misura di probabilità,

(ii) per ogni $0 \leq s \leq t$ ed $A \in \mathcal{B}(\mathbb{R})$,

$$p(s, t, \cdot, A) : \mathbb{R} \rightarrow [0, 1], x \rightarrow p(s, t, x, A)$$

è una funzione misurabile,

(iii) la famiglia $p(\cdot, \cdot, \cdot, \cdot)$ soddisfa l'**equazione di Chapman-Kolmogorov**, cioè per ogni $0 \leq r \leq s \leq t$, per ogni $x \in \mathbb{R}$ e per ogni $A \in \mathcal{B}(\mathbb{R})$ vale

$$p(r, t, x, A) = \int_{\mathbb{R}} p(r, s, x, dy) p(s, t, y, A)$$

Si noti che quindi necessariamente $p(t, t, x, A) = \delta_x(A)$, cioè vale 1 se $x \in A$ e vale 0 altrimenti e che le proprietà (i) e (ii) permettono di dare senso all'integrale in (iii).

Sia inoltre data una misura di probabilità μ_0 , detta misura delle **probabilità iniziali**.

Si definiscano, per $0 \leq t_1 \leq \dots \leq t_k$, e per $H \in \mathcal{B}(\mathbb{R}^{k+1})$,

$$\mu_{0,t_1,\dots,t_k}(H) := \int_H \mu_0(dx_0) p(0, t_1, x_0, dy_1) p(t_1, t_2, y_1, dy_2) \cdots \\ \cdots p(t_{k-2}, t_{k-1}, y_{k-2}, dy_{k-1}) p(t_{k-1}, t_k, y_{k-1}, dy_k),$$

e

$$\mu_{t_1,\dots,t_k}(H) := \mu_{0,t_1,\dots,t_k}(\mathbb{R} \times H), \quad \text{per } H \in \mathcal{B}(\mathbb{R}^k).$$

Se invece non vale $0 \leq t_1 \leq \dots \leq t_k$, si definiscono attraverso un'opportuna permutazione π per la quale $0 \leq t_{\pi_1} \leq \dots \leq t_{\pi_k}$ ed in modo che valga la (2).

La proprietà (iii) permette di verificare immediatamente la (3). La famiglia di distribuzioni finito-dimensionali così ottenute soddisfa quindi le proprietà di consistenza del teorema di Kolmogorov, e pertanto esiste un processo con queste distribuzioni finito-dimensionali. Un processo le cui distribuzioni finito-dimensionali si possono ottenere attraverso una famiglia di probabilità di transizione come sopra, viene detto Processo di Markov.

Il processo di Wiener ed il processo di Poisson sono processi Markoviani in questo senso con $\mu_0 = \delta_{\{0\}}$ per entrambi i processi e

$$p(s, t, x, dy) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left\{-\frac{(y-x)^2}{2(t-s)}\right\} dy$$

per il processo di Wiener,

$$p(s, t, n, m) = \frac{(\lambda(t-s))^{n-m}}{(n-m)!} \exp\{-\lambda(t-s)\}, \quad 0 \leq m \leq n$$

per il processo di Poisson.

In entrambi questi casi $p(s, t, x, A)$ dipende solo dalla differenza $t-s$. Più in generale ogni volta che $p(s, t, x, A) = p(s+h, t+h, x, A) = q(t-s, x, A)$ per ogni s, t, h, x, A , si parla di processo di Markov omogeneo. In realtà si tratta di esempi di processi a incrementi indipendenti ed omogenei, che sono l'argomento del prossimo esempio.

4) PROCESSI AD INCREMENTI INDIPENDENTI ED OMOGENEI

Si tratta di una generalizzazione del processo di Wiener, del processo di Poisson e dei processi di Poisson composti, e anche dei processi di Markov. Infatti le loro probabilità di transizione godono anche della proprietà che $p(s, t, x, A) = q(t-s, A-x)$, dove $A-x = \{y = z-x \text{ per } z \in A\}$, di modo che le distribuzioni finito-dimensionali vengono caratterizzate dalla famiglia di distribuzioni $q(t, \cdot)$, che rappresenta la distribuzione degli incrementi $X_{t+h} - X_h$, oltre ovviamente alla

distribuzione iniziale μ_0 .

La relazione (iii) per $q(r, B)$ diviene

$$(iii)' \quad q(u + v, B) = \int_{\mathbb{R}} q(u, dz)q(v, B - z).$$

Infatti, essendo $p(r, t, x, A) = q(t - r, A - x)$ ed anche

$$\begin{aligned} \int_{\mathbb{R}} p(r, s, x, dy)p(s, t, y, A) &= \int_{\mathbb{R}} q(s - r, d(y - x))q(t - s, A - y) = \\ & \text{(cambiando variabile } z = y - x) \\ &= \int_{\mathbb{R}} q(s - r, dz)q(t - s, A - x - z), \end{aligned}$$

la (iii)' si ottiene dalla (iii) ponendo, per semplicità di notazione, $u = s - r$, $v = t - s$ e $B = A - x$.

Si noti infine che posto $F_t(x) = q(t, (-\infty, x])$ la (iii)' equivale a chiedere che

$$F_{u+v}(x) = F_u * F_v(x),$$

dove $*$ indica la convoluzione. È possibile caratterizzare di conseguenza le distribuzioni $q(t, \cdot)$ per le quali esiste un processo ad incrementi indipendenti ed omogenei. Chiaramente deve valere che $F_t = (F_{t/n})^{*n}$, qualunque sia $n \geq 1$, ovvero si tratta delle distribuzioni infinitamente divisibili.

Qui si potrebbero citare le varie caratterizzazioni delle funzioni caratteristiche delle distribuzioni infinitamente divisibili.

1.2 TEMPO DISCRETO: METODO DIRETTO

Passiamo ora ad illustrare il preannunciato metodo costruttivo di dimostrazione del teorema di Kolmogorov, ma prima di tutto, proviamo ad ottenere una variabile aleatoria bidimensionale con funzione di distribuzione $F(x, y)$ data. Notiamo che si può riscrivere

$$F(x, y) = \int_{-\infty}^x F_{Y|X}(y | z) dF_X(z).$$

Daremo per il momento per scontato che tale scomposizione (o meglio disintegrazione) sia sempre possibile.

Si possono inoltre definire le inverse generalizzate $\Gamma_X(\cdot)$ di $F_X(\cdot)$ e $\Gamma_{Y|X}(\cdot | x)$ di $F_{Y|X}(\cdot | x)$, qualunque sia x .

Siano ora U e V due v.a. uniformi in $(0, 1)$ ed indipendenti, si definiscano

$$\tilde{X}(\omega) = \Gamma_X(U(\omega))$$

ed

$$\tilde{Y}(\omega) = \Gamma_{Y|X}(V(\omega) | \tilde{X}(\omega)) = \Gamma_{Y|X}(V(\omega) | \Gamma_X(U(\omega))).$$

È facile verificare che

$$\mathbb{P}(\tilde{X}(\omega) \leq x, \tilde{Y}(\omega) \leq y) = F(x, y).$$

Infatti

$$\begin{aligned} \mathbb{P}(\tilde{X}(\omega) \leq x, \tilde{Y}(\omega) \leq y) &= \mathbb{P}(\Gamma_X(U(\omega)) \leq x, \Gamma_{Y|X}(V(\omega) | \Gamma_X(U(\omega))) \leq y) = \\ &= \mathbb{P}\{U(\omega) \leq F_X(x), V(\omega) \leq F_{Y|X}(y | \Gamma_X(U(\omega)))\} = \\ &= \int_{(0, F_X(x)]} F_{Y|X}(y | \Gamma_X(u)) du = \\ &\hspace{15em} \text{(applicando il cambio di variabile } z = \Gamma_X(u)) \\ &\hspace{15em} \text{(considerando che } u \leq F_X(x) \Leftrightarrow \Gamma_X(u) = z \leq x) \\ &= \int_{-\infty}^x F_{Y|X}(y | z) dF_X(z) = F(x, y). \end{aligned}$$

È facile generalizzare al caso in dimensione n e costruire una variabile aleatoria n -dimensionale (X_1, \dots, X_n) con funzione di distribuzione data, a partire da n variabili aleatorie uniformi in $(0, 1)$ ed indipendenti.

A questo punto è chiaro come estendere il procedimento al caso di una successione di v.a. $\{X_n\}_{n \geq 1}$ con distribuzioni finito-dimensionali $\mu_{1, \dots, n}$ assegnate (e relativa funzione di ripartizione $F(x_1, x_2, \dots, x_n)$) in modo che

$$\mu_{1, \dots, n+1}(H \times \mathbb{R}) = \mu_{1, \dots, n}(H),$$

pur di avere a disposizione una successione di v.a. uniformi in $(0, 1)$ ed indipendenti.

Prendendo $\Omega = (0, 1)$ è certamente possibile avere una successione di v.a. uniformi in $(0, 1)$ ed indipendenti. Per costruire tale successione si ricordi che scrivendo $\omega \in (0, 1)$ in forma diadica $\omega = \sum_{i=1}^{\infty} W_i(\omega) \frac{1}{2^i}$, le v.a. W_i risultano indipendenti e $\mathbb{P}(W_i = 0) = \mathbb{P}(W_i = 1) = \frac{1}{2}$. La successione U_n di v.a. uniformi ed indipendenti si può costruire, a partire dalle v.a. $\{W_i\}$, riordinandole in modo che formino una sequenza a doppio indice $\{\tilde{W}_{i,n}\}$ così da poter definire

$$U_n(\omega) = \sum_1^{\infty} \tilde{W}_{i,n}(\omega) \frac{1}{2^i},$$

ad esempio si può prendere $W_{i,n} = W_{2^{i-1}(2n+1)}$.

Questa costruzione è riportata perché sia chiaro che l'affermazione che esiste una successione di v.a. indipendenti ed uniformi in $(0, 1)$, non dipende dal Teorema di Kolmogorov dimostrato precedentemente.

Infine, va notato che sostanzialmente le condizioni di compatibilità servono solo a garantire che le distribuzioni finito-dimensionali della successione aleatoria così ottenuta, e relative a tempi non consecutivi siano quelle volute.

2 OSSERVAZIONE SU \mathcal{R}^I

La σ -algebra \mathcal{R}^I coincide con la famiglia degli insiemi del tipo

$$\mathcal{A} = \{x(\cdot) \mid \{x(t_k)\}_k \in \mathcal{N}\}, \quad (5)$$

al variare delle successioni $S = \{t_k\}_k \subseteq I$ ed $\mathcal{N} \in \mathcal{R}^{\mathbb{N}}$.

Anzi, più in generale, dato un processo, cioè una famiglia di variabili aleatorie $(Y_t, t \in I)$ in uno spazio $(\Omega, \mathcal{F}, \mathbb{P})$, e posto $\mathcal{F}_S = \sigma\{Y_s, \text{ per } s \in S\}$, valgono le seguenti affermazioni:

- (i) Se $A \in \mathcal{F}_I, \omega \in A$ e $Y_t(\omega) = Y_t(\omega')$ per ogni $t \in I$, allora $\omega' \in A$.
- (ii) Se $A \in \mathcal{F}_I$, allora esiste un $S \subseteq I$ e numerabile per cui $A \in \mathcal{F}_S$, ovvero $\mathcal{F}_I = \bigcup_S \mathcal{F}_S$, dove l'unione è fatta su tutti i sottoinsiemi S numerabili di I .

Entrambe le due proprietà si basano sul fatto che, posto $\xi : \Omega \rightarrow \mathbb{R}^I, \omega \rightarrow \xi(\omega)(\cdot)$, dove $\xi(\omega)(t) = Y_t(\omega)$ per ogni $t \in I$, allora

$$\mathcal{F}_I = \{\xi^{-1}(M), \quad M \in \mathcal{R}^I\}.$$

Infatti allora $Y_t(\omega) = Y_t(\omega')$ per ogni $t \in I$ equivale a $\xi(\omega) = \xi(\omega')$ e quindi $\omega \in A = \xi^{-1}(M)$, cioè $\xi(\omega) \in M$ implica $\xi(\omega') \in M$ e quindi $\omega' \in A$. Infatti ogni σ -algebra rispetto alla quale ξ è misurabile deve contenere le controimmagini dei cilindri e quindi deve contenere anche $\{\xi^{-1}(M), M \in \mathcal{R}^S\}$ per ogni S numerabile, ed inoltre la famiglia $\bigcup_S \mathcal{F}_S$ è una σ -algebra, come è facile verificare.

L'unico punto in cui è necessaria un pò di attenzione è il caso di unioni numerabili di eventi $E_n \in \bigcup_S \mathcal{F}_S$. Sia S_n tale che $E_n \in \mathcal{F}_{S_n}$. Poiché $\mathcal{F}_{S_n} \subseteq \mathcal{F}_{\cup_m S_m}$ ed $\mathcal{F}_{\cup_m S_m}$ è una σ -algebra, ovviamente $\bigcup_n E_n \in \mathcal{F}_{\cup_m S_m} \subseteq \bigcup_S \mathcal{F}_S$.

In particolare $\mathcal{R}^I = \bigcup_S \mathcal{R}^S$, dove l'unione è fatta su tutti i sottoinsiemi S numerabili di I .

2.1 CONSEGUENZE: PROBLEMI CON LO SPAZIO CANONICO

Consideriamo il caso in cui I è un insieme continuo e per fissare le idee $I = [0, T]$ oppure $[0, \infty)$. Allora non ha senso considerare la probabilità che le traiettorie abbiano particolari proprietà tipo siano crescenti, o continue, in quanto tali insiemi non sono misurabili nello spazio canonico non essendo esprimibili come in (8): non è possibile trovare una successione di tempi (ovvero un sottoinsieme numerabile di tempi) e condizioni relative solo a tali istanti per determinare se una funzione è continua oppure no. In altre parole: comunque sia data una successione di tempi esistono funzioni crescenti (o continue) su tale successione di tempi, ma non su tutto I .

Tipicamente, al massimo si può sperare di ottenere un processo $(X'_t, t \in I)$, che abbia le proprietà richieste, e che sia **stocasticamente equivalente a** $(X_t, t \in I)$, ovvero una **versione del processo** $(X_t, t \in I)$.

Definizione 2.1. *Sia dato un processo $(X_t, t \in I)$. Un processo $(X'_t, t \in I)$ si dice **stocasticamente equivalente a** $(X_t, t \in I)$ se*

$$\mathbb{P}(X'_t \neq X_t) = 0 \text{ per ogni } t \geq 0.$$

In tale caso si dice anche X'_t è una **versione** di X_t ,

ed ha quindi le stesse distribuzioni finito-dimensionali di X_t

Per proprietà tipo la continuità si può al massimo sperare di procedere nel seguente modo: si dimostra che, se ristretto ad un opportuno insieme numerabile di tempi S , il processo X_t risulta a traiettorie continue, tranne a parte un eventuale insieme di probabilità nulla. Si definisce poi un nuovo processo X'_t definito come limite delle traiettorie ristrette ad S . Si spera poi di dimostrare che il processo così ottenuto abbia le stesse distribuzioni finito-dimensionali del processo X_t .

Problemi di questo tipo sono connessi al problema della separabilità di un processo di cui non parleremo in questo corso. Rimandiamo gli interessati al libro di Billingsley, Probability and measures.

A titolo di esempio vediamo come si può procedere con il processo di Poisson. Il Teorema di Kolmogorov assicura che esiste una misura di probabilità \mathbb{P} su $(\mathbb{R}^I, \mathcal{R}^I)$, in modo che il processo $X_t(x(\cdot)) = x(t)$ abbia le distribuzioni finito-dimensionali uguali a quelle del processo di Poisson. Ovviamente non possiamo dire nulla sulle traiettorie, ad esempio non possiamo affermare che siano non decrescenti, costanti a tratti e con salti unitari. Proviamo a mostrare come procedere per ottenere una versione X'_t con traiettorie non decrescenti. Consideriamo l'insieme D dei diadici. La condizione che $X_t|_D$ sia a traiettorie non decrescenti e a valori in \mathbb{N} diviene $X_t|_D \in \mathcal{I}$, dove

$$\mathcal{I} = \bigcap_n \bigcap_{0 \leq i < 2^n} \{x(\cdot) \text{ tali che } x(\frac{i}{2^n}) \in \mathbb{N} \text{ e } x(\frac{i}{2^n}) \leq x(\frac{i+1}{2^n})\}.$$

Inoltre

$$\mathbb{P}\left(\bigcap_{0 \leq i < 2^n} \{x(\cdot) \text{ tali che } x(\frac{i}{2^n}) \in \mathbb{N} \text{ e } x(\frac{i}{2^n}) \leq x(\frac{i+1}{2^n})\}\right) = 1 \quad \text{per ogni } n,$$

e quindi vale anche che $(X_t, t \in D)$ è a traiettorie non decrescenti con probabilità 1. Si noti inoltre che

$$\lim_{s \downarrow t, s \in D} x(s)$$

esiste per ogni $x(\cdot)$ appartenente a \mathcal{I} .

Si definisca ora

$$\begin{cases} X'_t(x(\cdot)) := X_t(x(\cdot)) & \text{se } x(\cdot) \in \mathcal{I} \text{ e } t \in D \\ X'_t(x(\cdot)) := \lim_{s \downarrow t, s \in D} X_s(x(\cdot)) = \lim_{s \downarrow t, s \in D} x(s), & \text{se } x(\cdot) \in \mathcal{I} \text{ e } t \notin D \\ X'_t(x(\cdot)) := 0 & \text{se } x(\cdot) \notin \mathcal{I} \text{ e per ogni } t \in (0, 1) \end{cases}$$

Inoltre ovviamente $\{X'_t = X_t\}$ è tutto \mathbb{R}^I , per $t \in D$, ed è, se $t \notin D$,

$$\bigcap_{\epsilon \in \mathbb{Q}^+} \bigcup_{\delta \in \mathbb{Q}^+} \bigcap_{\substack{t \leq s \leq t+\delta \\ s \in D}} \{x(\cdot) \text{ tali che } |x(s) - x(t)| < \epsilon\}$$

e, se intersecato con \mathcal{I} , coincide con

$$\bigcup_{\delta \in \mathbb{Q}^+} \bigcap_{\substack{t \leq s \leq t+\delta \\ s \in D}} \{x(\cdot) \text{ tali che } |x(s) - x(t)| = 0\}.$$

Infine ¹

$$\mathbb{P}\left(\bigcap_{\substack{t \leq s \leq t+\delta \\ s \in D}} \{x(\cdot) \text{ tali che } |x(s) - x(t)| = 0\}\right) \geq \exp\{-\lambda\delta\} \rightarrow 1$$

Di conseguenza $\mathbb{P}(X'_t \neq X_t) = 0$ per ogni $t \geq 0$, ovvero X'_t è una versione di X_t , ed ha quindi le stesse distribuzioni finito-dimensionali di X_t e gode della proprietà di monotonia, continuità a destra delle traiettorie e di essere a valori in \mathbb{N} .

3 ESISTENZA DI UNA VERSIONE CONTINUA: CRITERIO DI CHENSOV.

Il seguente criterio sufficiente fornisce condizioni che assicurano in realtà l'esistenza di una versione hölderiana, non solo continua. Ricordiamo che una funzione $f(x)$ è **hölderiana** di esponente γ se per ogni x esistono un $\delta(x) > 0$ e un $L_\gamma(x)$ tali che, per ogni y per il quale $|y - x| \leq \delta(x)$, si abbia

$$|f(x) - f(y)| \leq L_\gamma(x)|x - y|^\gamma.$$

Nel caso in cui $\delta(x)$ e $L_\gamma(x)$ possano essere presi in modo indipendente da x , per $x \in I$, si dice che f è **uniformemente hölderiana** nell'insieme I .

CRITERIO DI CHENSOV

Sia X_t un processo per cui esistono α, β e C positivi, per cui

$$\mathbb{E}[|X_t - X_s|^\beta] \leq C|t - s|^{1+\alpha}.$$

Allora esiste una versione \tilde{X}_t di X_t , a traiettorie continue. Inoltre le traiettorie sono uniformemente hölderiane di esponente γ , per ogni $\gamma < \frac{\alpha}{\beta}$, in ogni intervallo limitato.

Dimostrazione: Per semplicità consideriamo il caso dell'intervallo $[0, 1]$. Sia come al solito D l'insieme dei diadici e $D_n = \{s = \frac{k}{2^n}, \text{ per } 0 \leq k \leq 2^n\}$. Si definisca Γ_n

$$\Gamma_n = \{\omega \text{ t.c. } \max_{k \leq 2^n - 1} |X_{k/2^n} - X_{(k+1)/2^n}| > 2^{-n\gamma}\}$$

Se mostriamo che

$$\sum_{n=1}^{\infty} \mathbb{P}(\Gamma_n) < +\infty,$$

allora, per il Lemma di Borel-Cantelli,

$$\mathbb{P}(\bigcap_{m \geq 1} \bigcup_{n \geq m} \Gamma_n) = 0$$

ovvero

$$\mathbb{P}(\bigcup_{m \geq 1} \bigcap_{n \geq m} \Gamma_n^c) = 1.$$

Ciò significa che esiste un evento \mathcal{N} , con $\mathbb{P}(\mathcal{N}) = 0$, e che per ogni $\omega \in \mathcal{N}^c$ esiste un $m = m(\omega)$ tale che per ogni $n \geq m = m(\omega)$

$$|X_{k/2^n} - X_{(k+1)/2^n}| \leq 2^{-n\gamma}.$$

Mostriamo ora che per $\omega \in \mathcal{N}^c$ le traiettorie sono uniformemente hölderiane su D con

$$\delta = \delta(\omega) = 1/2^{m(\omega)} \text{ ed } L_\gamma = \left(\frac{2}{2^\gamma - 1} + 1\right)2^\gamma,$$

e quindi sono hölderiane su D per quasi ogni ω .

Sia quindi $\omega \in \mathcal{N}^c$.

Cominciamo con l'osservare che se

$$r \in [j/2^m, (j+1)/2^m), \text{ con } m \geq m(\omega), \text{ ed } r \in D,$$

allora

$$r = j/2^m + \sum_{h=m+1}^{m+i} \alpha_h \frac{1}{2^h},$$

per un $i \geq 0$, con $\alpha_h \in \{0, 1\}$. Posto

$$r_0 = j/2^m \text{ ed } r_s = j/2^m + \sum_{h=m+1}^{m+s} \alpha_h \frac{1}{2^h},$$

si ha che $r = r_i$ e che

$$|X_{j/2^m} - X_r| \leq \sum_{s=1}^i |X_{r_s} - X_{r_{s-1}}| \leq \sum_{h=m+1}^{m+i} 2^{-h\gamma} \leq 2^{-m\gamma} \sum_{l=1}^i 2^{-l\gamma} \leq 2^{-m\gamma} \frac{1}{2^\gamma - 1}.$$

Se invece $|r - u| \leq 1/2^m$ con $r, u \in D$, allora sono possibili due casi:

(i) **esiste j per cui $r, u \in [j/2^m, (j+1)/2^m)$** , e allora

$$|X_u - X_r| \leq |X_{j/2^m} - X_u| + |X_{j/2^m} - X_r| \leq \frac{2}{2^\gamma - 1} 2^{-m\gamma}$$

(ii) **esiste j per cui $r \in [j/2^m, (j+1)/2^m)$ ed $u \in [(j+1)/2^m, (j+2)/2^m)$** , e allora

$$|X_u - X_r| \leq |X_{(j+1)/2^m} - X_u| + |X_{j/2^m} - X_{(j+1)/2^m}| + |X_{j/2^m} - X_r| \leq \left(\frac{2}{2^\gamma - 1} + 1 \right) 2^{-m\gamma}$$

Quindi, per quasi ogni ω , si ha definitivamente, per ogni $r, u \in D$, con $|r - u| \leq 1/2^m$,

$$|X_u - X_r| \leq M_\gamma 2^{-m\gamma},$$

con $M_\gamma = \frac{2}{2^\gamma - 1} + 1$. In particolare se $|r - u| \leq 1/2^m(\omega)$, ed $m \geq m(\omega)$ è tale che $1/2^{m+1} \leq |r - u| \leq 1/2^m$, allora

$$|X_u - X_r| \leq L_\gamma 2^{-(m+1)\gamma} \leq L_\gamma |r - u|^\gamma.$$

Avremo quindi, come preannunciato, che le traiettorie sono uniformemente hölderiane su D con $\delta = \delta(\omega) = 1/2^{m(\omega)}$ ed $L_\gamma = \left(\frac{2}{2^\gamma - 1} + 1 \right) 2^\gamma$, per quasi ogni ω .¹

Si definisce, sempre per $\omega \in \mathcal{N}^c$, $\tilde{X}_t(\omega)$ come il limite di $X_{r_n}(\omega)$ per una successione r_n di diadici convergente a t , e per $\omega \in \mathcal{N}$ si può definire, ad esempio, $\tilde{X}_t \equiv 0$. È facile vedere che la proprietà di

¹Si noti che questa parte della dimostrazione è puramente deterministica: nel senso che se una funzione $x(\cdot)$ definita sui diadici di $[0, 1]$, gode della proprietà che esiste un m tale che per ogni $n \geq m$ e per $k < 2^n$ si ha $|x(k/2^n) - x((k+1)/2^n)| \leq 2^{-n\gamma}$, allora $x(\cdot)$ è hölderiana sui diadici di $[0, 1]$

hölderianità si conserva. Il fatto che si tratta di una versione di X_t , deriva dal fatto che per l'ipotesi fatta X_t è continuo in L^β e quindi è continuo in probabilità e quindi contemporaneamente si ha

$$X_{r_n} \xrightarrow{Pr} X_t \text{ e } X_{r_n} \xrightarrow{q.c.} \tilde{X}_t,$$

da cui subito $\mathbb{P}(X_t = \tilde{X}_t) = 1$.

A questo punto rimane solo da controllare che $\sum_{n=1}^{\infty} \mathbb{P}(\Gamma_n) < +\infty$, e infatti:

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbb{P}(\Gamma_n) &\leq \sum_{n=1}^{\infty} \sum_{k=0}^{2^n-1} \mathbb{P}(|X_{k/2^n} - X_{(k+1)/2^n}| > 2^{-n\gamma}) \leq \sum_{n=1}^{\infty} \sum_{k=0}^{2^n-1} \mathbb{E}[|X_{k/2^n} - X_{(k+1)/2^n}|^\beta] / 2^{-n\gamma\beta} \leq \\ &\leq \sum_{n=1}^{\infty} \sum_{k=0}^{2^n-1} (1/2^n)^{1+\alpha} 2^{n\gamma\beta} = \sum_{n=1}^{\infty} 2^n (1/2^n)^{1+\alpha} 2^{n\gamma\beta} = \sum_{n=1}^{\infty} 2^{-n(\alpha-\gamma\beta)} < +\infty \end{aligned}$$

purché $\alpha - \gamma\beta > 0$, e ciò è vero per come abbiamo preso γ . (Si noti che è fondamentale nella prova che $1 + \alpha > 1$, in quanto altrimenti non ci sarebbe speranza di ottenere nessun tipo di convergenza ovvero la maggiorazione di $\mathbb{P}(\Gamma_n)$ non avrebbe nessun significato in quanto risulterebbe un numero maggiore di 1)

APPLICAZIONE

Il processo W_t di Wiener standard, o moto Browniano, ammette sempre una versione continua, anzi hölderiana per ogni $\gamma < \frac{1}{2}$. Infatti, per $s \leq t$,

$$\mathbb{E}[|W_t - W_s|^k] = \mathbb{E}[|W_{t-s}|^k] = C(k)|t - s|^{k/2}.$$

Il criterio di Chensov si può quindi applicare per $k \geq 3$ (così $\alpha = (k/2) - 1 > 0$) e ci garantisce che esiste una versione di W_t le cui traiettorie sono hölderiane di esponente γ per $\gamma < \frac{(k/2)-1}{k} = \frac{1}{2} - \frac{1}{k}$, e quindi per ogni $\gamma < \frac{1}{2}$. È possibile dimostrare che, a parte un eventuale insieme di misura nulla, le traiettorie non sono hölderiane di esponente $\frac{1}{2}$, e tantomeno di esponente maggiore di $\frac{1}{2}$.

Nonostante questo però le traiettorie delle versione continua non possono essere molto regolari, in quanto ad esempio sono anche a variazione non limitata con probabilità 1 su ogni intervallo $[s, t]$. Infatti se esistesse un intervallo $[s, t]$ ed un insieme A con $\mathbb{P}(A) > 0$, su cui la variazione di W_u , cioè

$$V(\omega) := \sup \left\{ \sum_{k=0}^n |W_{t_{k+1}} - W_{t_k}|, \text{ al variare delle partizioni } \pi : t_0 = s \leq t_1 \leq \dots \leq t_n = t \right\}$$

fosse finita per ogni $\omega \in A$, allora si avrebbe che

$$\begin{aligned} S_\pi &:= \sum_{k=0}^n |W_{t_{k+1}} - W_{t_k}|^2 \leq \sum_{k=0}^n \left(\sup_h |W_{t_{h+1}} - W_{t_h}| \right) |W_{t_{k+1}} - W_{t_k}| = \\ &= \left(\sup_h |W_{t_{h+1}} - W_{t_h}| \right) \sum_{k=0}^n |W_{t_{k+1}} - W_{t_k}| \leq \left(\sup_h |W_{t_{h+1}} - W_{t_h}| \right) V(\omega) \rightarrow 0 \end{aligned}$$

se si tratta di una versione continua di W_u . Questo fatto è in contraddizione con la seguente proprietà del moto browniano (valida anche per versioni non continue):

4 PROPRIETÀ: LE TRAIETTORIE DEL PROCESSO DI WIENER NON SONO A VARIAZIONE LIMITATA

Si definisca

$$S_\pi := \sum_{k=0}^n |W_{t_{k+1}} - W_{t_k}|^2,$$

allora $S_\pi \xrightarrow{L^2} (t-s)$, al tendere a zero di $|\pi| := \max_k |t_{k+1} - t_k|$.

DIMOSTRAZIONE: Si tratta di mostrare che $\mathbb{E}[(S_\pi - (t-s))^2] \rightarrow 0$ al tendere a zero di $|\pi| := \max_k |t_{k+1} - t_k|$, e infatti

$$\begin{aligned} \mathbb{E}[(S_\pi - (t-s))^2] &= \mathbb{E}[S_\pi^2 + (t-s)^2 - 2S_\pi(t-s)] = \\ &= \mathbb{E}\left[\left(\sum_{k=0}^n |W_{t_{k+1}} - W_{t_k}|^2\right)^2 + (t-s)^2 - 2\left(\sum_{k=0}^n |W_{t_{k+1}} - W_{t_k}|^2\right)(t-s)\right] = \\ &= \mathbb{E}\left[\sum_{k=0}^n |W_{t_{k+1}} - W_{t_k}|^2 \sum_{h=0}^n |W_{t_{h+1}} - W_{t_h}|^2 + (t-s)^2 - 2\left(\sum_{k=0}^n |W_{t_{k+1}} - W_{t_k}|^2\right)(t-s)\right] = \\ &= \sum_{k=0}^n \mathbb{E}[|W_{t_{k+1}} - W_{t_k}|^4] + \sum_{k=0}^n \sum_{\substack{h=0 \\ h \neq k}}^n \mathbb{E}[|W_{t_{k+1}} - W_{t_k}|^2 |W_{t_{h+1}} - W_{t_h}|^2] + \\ &\quad + (t-s)^2 - 2(t-s) \sum_{k=0}^n \mathbb{E}[|W_{t_{k+1}} - W_{t_k}|^2] = \\ &= \sum_{k=0}^n 3|t_{k+1} - t_k|^2 + \sum_{k=0}^n \sum_{\substack{h=0 \\ h \neq k}}^n |t_{k+1} - t_k| |t_{h+1} - t_h| + (t-s)^2 - 2(t-s) \sum_{k=0}^n |t_{k+1} - t_k| = \\ &= \sum_{k=0}^n 2|t_{k+1} - t_k|^2 + \sum_{k=0}^n \sum_{h=0}^n |t_{k+1} - t_k| |t_{h+1} - t_h| + (t-s)^2 - 2(t-s)(t-s) \leq \\ &\leq 2 \sum_{k=0}^n |t_{k+1} - t_k| (\max_h |t_{h+1} - t_h|) + (t-s)^2 + (t-s)^2 - 2(t-s)(t-s) = \\ &= 2(\max_h |t_{h+1} - t_h|) \sum_{k=0}^n |t_{k+1} - t_k| = 2(\max_h |t_{h+1} - t_h|)(t-s) \rightarrow 0. \end{aligned}$$

Questo fatto in particolare pone dei problemi nel tentare di definire $\int_0^t f(s, \omega) dW_s(\omega)$, in quanto non si può adottare la classica definizione di integrale di Lebesgue-Stieltjes, che avrebbe senso solo se W_s avesse traiettorie a variazione finita. La definizione che si dà è quindi diversa e si parlerà di integrale stocastico. Per ottenerla avremo però bisogno della nozione di Martingala.

5 APPENDICE: variabili Gaussianne

Cominciamo con il definire una variabile aleatoria gaussiana standard unidimensionale:

Definizione 5.1. Si dice che una variabile aleatoria reale Z è **gaussiana** di valore atteso μ e varianza σ^2 , se ammette densità

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

In questo caso si usa la notazione $Z \sim N(\mu, \sigma^2)$. Se $\mu = 0$ e $\sigma^2 = 1$ allora si dice che Z segue una **legge normale o gaussiana standard**.

Caso n -dimensionale: iniziamo con il caso di un vettore (colonna) aleatorio

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_k \\ \dots \\ Y_n \end{pmatrix}$$

(Y_1, Y_2, \dots, Y_n)

a componenti indipendenti e tutte gaussiane standard, ovvero il caso in cui

$$f_Y(\mathbf{y}) = \prod_{i=1}^n f_{Y_i}(y_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} y_i^2 \right\} = \frac{1}{(\sqrt{2\pi})^n} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n y_i^2 \right\} = \frac{1}{(2\pi)^{n/2}} \exp \left\{ -\frac{1}{2} \mathbf{y}' \mathbf{y} \right\}.$$

dove l'apice indica l'operazione di trasposizione, ovvero \mathbf{y}' è il vettore riga (y_1, y_2, \dots, y_n) .

È immediato verificare che $\mathbb{E}(Y_i) = 0$, $Var(Y_i) = 1$ e che $Cov(Y_i, Y_j) = 0$, per $i \neq j$.

Sia ora A una matrice non singolare e sia \mathbf{m} un vettore (colonna). Definiamo ora $Z = AY + \mathbf{m}$ e cerchiamo la sua densità. Sappiamo dai risultati generali che se Y ammette densità e $Z = \varphi(Y)$ con φ invertibile e con derivate continue, allora anche Z ammette densità:

$$f_Z(z) = f_Y(\varphi^{-1}(z)) \left| \det \left(\frac{\partial \varphi^{-1}(z)}{\partial z} \right) \right| = f_Y(\varphi^{-1}(z)) \frac{1}{\left| \det \left(\frac{\partial \varphi(\mathbf{y})}{\partial \mathbf{y}} \right) \right|_{\mathbf{y}=\varphi^{-1}(z)}}$$

di conseguenza, poiché nel nostro caso $\varphi(\mathbf{y}) = A\mathbf{y} + \mathbf{m}$ e $\varphi^{-1}(z) = A^{-1}(z - \mathbf{m})$

$$f_Z(z) = \frac{1}{\sqrt{2\pi}^n} \exp \left\{ -\frac{1}{2} (A^{-1}(z - \mathbf{m}))' A^{-1}(z - \mathbf{m}) \right\} \frac{1}{|\det(A)|}.$$

Essendo

$$(A^{-1}(z - \mathbf{m}))' A^{-1}(z - \mathbf{m}) = (z - \mathbf{m})' (A^{-1})' A^{-1}(z - \mathbf{m}) = (z - \mathbf{m})' (A')^{-1} A^{-1}(z - \mathbf{m}) = (z - \mathbf{m})' (AA')^{-1}(z - \mathbf{m})$$

si ottiene

$$f_Z(z) = \frac{1}{(2\pi)^{n/2}} \frac{1}{|\det(A)|} \exp \left\{ -\frac{1}{2} (z - \mathbf{m})' (AA')^{-1} (z - \mathbf{m}) \right\}.$$

La precedente espressione si basa sulle seguenti proprietà:

(i) $\boxed{(A')^{-1} = (A^{-1})'}$
in quanto

$$A'z = w \Leftrightarrow z = (A')^{-1}w$$

e inoltre

$$\begin{aligned} A'z = w &\Leftrightarrow (z'A)' = w \Leftrightarrow z'A = w' \Leftrightarrow z' = w'A^{-1} \\ &\Leftrightarrow z = (w'A^{-1})' \Leftrightarrow z = (A^{-1})'w. \end{aligned}$$

(ii) $\boxed{(AA')^{-1} = (A')^{-1}A^{-1}}$
in quanto

$$(AA')^{-1}z = w \Leftrightarrow z = AA'w \Leftrightarrow A^{-1}z = A'w \Leftrightarrow (A')^{-1}A^{-1}z = w.$$

È interessante notare che sia il vettore \mathbf{m} che la matrice $AA' = A'A$ hanno una interpretazione probabilistica:

$$\mathbb{E}(Z_i) = \mathbb{E}\left(\sum_{k=1}^n a_{k,i}Y_k\right) + m_i = \sum_{k=1}^n a_{k,i}\mathbb{E}(Y_k) + m_i = m_i$$

$$Cov(Z_i, Z_j) = \mathbb{E}[(Z_i - m_i)(Z_j - m_j)] = \mathbb{E}\left[\sum_{k=1}^n a_{i,k}Y_k \sum_{h=1}^n a_{j,h}Y_h\right] = \sum_{k=1}^n \sum_{h=1}^n a_{i,k}a_{j,h}\mathbb{E}[Y_k Y_h]$$

e quindi

$$Cov(Z_i, Z_j) = \sum_{k=1}^n a_{i,k}a_{j,k}\mathbb{E}[Y_k Y_k] + \sum_{k=1}^n \sum_{h \neq k}^n a_{i,k}a_{j,h}\mathbb{E}[Y_k Y_h] = \sum_{k=1}^n a_{i,k}a_{j,k} = (AA')_{i,j}$$