

Indice

1	Il Moto Browniano	2
1.1	La distribuzione gaussiana multivariata	2
1.2	Una prima interpretazione	5
1.3	Definizioni	6
1.4	Caratteristiche del moto Browniano	9
1.5	Il ponte Browniano	11
1.6	Proprietà delle traiettorie	13
1.6.1	Funzioni a variazioni finite	14
1.6.2	Funzioni Hölderiane	15
2	Simulazione di traiettorie di moti Browniani	17
2.1	La simulazione	17
2.1.1	Il metodo della trasformazione inversa	18
2.1.2	L'algoritmo di Box-Muller	19
2.1.3	La simulazione multivariata	21
2.2	Un algoritmo di simulazione per moti Browniani	21
2.3	Proprietà del moto Browniano simulato	23
A	Qualche richiamo di probabilità	26
B	Variabili di Rademacker	28

Capitolo 1

Il Moto Browniano

Il moto Browniano è il più importante processo stocastico continuo e trova numerose applicazioni in quasi ogni scienza. In particolare, in fisica rappresenta il moto disordinato delle particelle presenti in fluidi, come fu osservato per la prima volta dal botanista inglese Robert Brown nel 1828. I primi studi su questo fenomeno si devono però ad Albert Einstein (1905), grazie ai quali confermò indirettamente l'esistenza di atomi e molecole. Tuttavia la prima succinta definizione matematica di questo processo fu data da Norbert Wiener nel 1918: infatti il moto Browniano si può anche chiamare processo di Wiener, in onore del matematico statunitense.¹ Oggi il moto Browniano è utilizzato in diverse materie: in economia (fluttuazioni di borsa), meccanica quantistica, cosmologia fisica e molte altre.

Fissiamo fin da subito uno spazio di probabilità (Ω, \mathcal{F}, P) .

1.1 La distribuzione gaussiana multivariata

Ricordiamo le caratteristiche e le proprietà di questa distribuzione che è fondamentale per la comprensione del processo stocastico che si andrà a studiare.

Definizione 1.1. *Siano $X = (X_1, \dots, X_n)$ un vettore aleatorio a valori in \mathbb{R}^n , C una matrice quadrata di dimensione n definita positiva e simmetrica e $m = (m_1, \dots, m_n) \in \mathbb{R}^n$. Se la densità di X calcolata nel punto $x = (x_1, \dots, x_n)$ è*

$$f_X(x) = \frac{1}{(2\pi)^{n/2} \sqrt{\det(C)}} \exp\left(-\frac{(x-m)^t C^{-1} (x-m)}{2}\right)$$

allora X è un vettore gaussiano e si indica $X \sim N^n(m, C)$.

¹Nel proseguo del lavoro useremo indistintamente i due termini, riferendoci alla stessa entità matematica.

Chiamiamo m vettore delle medie e C matrice di covarianza. La funzione caratteristica di un vettore così definito è, per $t \in \mathbb{R}$

$$\varphi_X(t) = \exp(i \langle t, m \rangle) \exp\left(-\frac{1}{2} \langle Ct, t \rangle\right)$$

Se le componenti del vettore X sono delle variabili normali standard indipendenti ed identicamente distribuite, allora la densità della legge congiunta sarà

$$f_X(x) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} x^t x\right)$$

In questo caso $X \sim N^n(\underline{0}, I_n)$ e la funzione caratteristica è

$$\varphi_X(t) = \exp\left(-\frac{1}{2} \langle t, t \rangle\right)$$

Ricordiamo ora due teoremi fondamentali.

Teorema 1.2. *Sia X un vettore gaussiano. Allora le densità marginali sono indipendenti se e solo se la matrice di covarianza è diagonale.*

Questa è una proprietà peculiare delle variabili normali: infatti per un qualsiasi vettore aleatorio è sempre vero che l'indipendenza delle densità marginali implica la diagonalità della matrice di covarianza, ma non è vero il viceversa.

Teorema 1.3. *Trasformazioni affini² di vettori gaussiani sono ancora vettori gaussiani.*

Grazie a questo teorema possiamo effettuare le seguenti osservazioni.

Osservazione 1.4. *Consideriamo $Z \sim N^k(\underline{0}, I_n)$, una matrice reale A di dimensione $n \times k$, con $k \geq n$, di rango pieno e un vettore $m \in \mathbb{R}^n$. Il vettore aleatorio X ottenuto con la trasformazione affine $X = m + AZ$ sarà ancora gaussiano. In più dato $C = AA^t$ ³, allora $X \sim N^n(m, C)$.*

Effettuiamo un'ulteriore trasformazione affine $Y = p + BX$, dove $p \in \mathbb{R}^r$ e B è una matrice reale di dimensione $r \times n$, con $n \geq r$, di rango pieno. Allora $Y \sim N^r(p + Bm, BA(BA)^t)$.

Osservazione 1.5. *Le densità marginali di vettori gaussiani sono gaussiane.*

²Una trasformazione affine è una trasformazione del tipo $x \mapsto Ax + b$, ovvero la composizione di una trasformazione lineare determinata da una matrice A e di una traslazione determinata da un vettore b .

³Sappiamo che una matrice C definita in questo modo è di rango massimo e quindi invertibile, oltre ad essere simmetrica.

Ciò è vero perchè le densità marginali si ottengono attraverso trasformazioni affini. Il viceversa invece non è vero, ossia che variabili marginalmente normali lo siano anche congiuntamente, come mostrerà il seguente esempio. Tuttavia se le variabili marginali sono normali e indipendenti, allora congiuntamente formano un vettore gaussiano per il Teorema 1.2

Esempio 1.6. *Siano X e Z due variabili aleatorie indipendenti, la prima gaussiana standard e la seconda con valori $+1$ o -1 con probabilità $1/2$. Consideriamo la variabile $Y = ZX$ e calcoliamone la densità.*

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(ZX \leq y) \\ &= P(X \leq y, Z = 1) + P(-X \leq y, Z = -1) \\ &= \frac{1}{2}F_X(y) + \frac{1}{2}P(X \geq -y) = F_X(y) \end{aligned}$$

Allora anche Y segue una distribuzione normale standard.

Consideriamo ora il vettore (X, Y) e verifichiamo che non si tratta di un vettore gaussiano, nonostante le componenti siano delle variabili normali standard.

Calcoliamo la covarianza tra queste due variabili.

$$\text{Cov}(X, Y) = \text{Cov}(X, ZX) = E(XZX) - E(X)E(ZX) = E(ZX^2) = E(Z)E(X^2) = 0$$

avendo sfruttato il fatto che anche X^2 e Z sono indipendenti tra loro.

Affinchè il vettore (X, Y) sia gaussiano allora, per il teorema (1.2), deve essere che X sia indipendente da Y . Ossia deve succedere che $P(X < x, Y < y) = P(X < x)P(Y < y)$, per ogni $x, y \in \mathbb{R}$. Scegliamo arbitrariamente $x = 1$ e $y = 2$.

La parte destra dell'equazione precedente si calcola andando a controllare direttamente i valori delle tavole per le normali standardizzate. Per la parte sinistra osserviamo che:

$$\begin{aligned} P(X < 1, Y < 2) &= P(X < 1, ZX < 2) \\ &= P(X < 1, X < 2, Z = 1) + P(X < 1, -X < 2, Z = -1) \\ &= \frac{1}{2}P(X < 1) + \frac{1}{2}P(X < 1, X > -2) \\ &= \frac{1}{2}F_X(1) + \frac{1}{2}P(-2 < X < 1) \\ &= F_X(1) - \frac{1}{2}F_X(-2) \end{aligned}$$

A questo punto anche per la parte sinistra dell'uguaglianza è possibile consultare le tavole della normale. Si verifica che i due valori ottenuti sono diversi e che quindi le variabili X e Y non sono indipendenti. Quindi il vettore (X, Y) non può essere gaussiano.

Prima di concludere con un ultimo teorema, dobbiamo sistemare alcune notazioni.

Sia $X \sim N^n(m, C)$, con le caratteristiche della Definizione 1.1, partizionato in due blocchi, $X = (X_1^t, X_2^t)^t$, di dimensioni rispettivamente n_1 e n_2 , con $n = n_1 + n_2$. Allora la media e la matrice di covarianza risultano anch'esse ripartite in blocchi di opportune dimensioni:

$$m = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

Teorema 1.7. *Sia X un vettore gaussiano partizionato nei vettori X_1 e X_2 e C_{11} invertibile. Allora la distribuzione condizionata di X_2 dato $X_1 = x_1$ è ancora gaussiana.*

Dimostrazione. Consideriamo la trasformazione affine

$$X' = \begin{bmatrix} X_1 \\ X_2 - C_{21}C_{11}^{-1}X_1 \end{bmatrix} = \begin{bmatrix} I_{n_1} & \mathbf{0} \\ -C_{21}C_{11}^{-1} & I_{n_2} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

Per il Teorema 1.3 anche X' è gaussiano. Calcoliamo quindi la matrice di covarianza di X' che per l'Osservazione 1.4 sappiamo essere

$$\begin{aligned} \text{Cov}(X') &= \begin{bmatrix} I_{n_1} & \mathbf{0} \\ -C_{21}C_{11}^{-1} & I_{n_2} \end{bmatrix} \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} I_{n_1} & -(C_{11}^{-1})^t C_{12} \\ \mathbf{0} & I_{n_2} \end{bmatrix} \\ &= \begin{bmatrix} C_{11} & C_{12} \\ \mathbf{0} & -C_{21}C_{11}^{-1}C_{12} + C_{22} \end{bmatrix} \begin{bmatrix} I_{n_1} & -(C_{11}^{-1})^t C_{12} \\ \mathbf{0} & I_{n_2} \end{bmatrix} \\ &= \begin{bmatrix} C_{11} & \mathbf{0} \\ \mathbf{0} & C_{22} - C_{21}C_{11}^{-1}C_{12} \end{bmatrix} \end{aligned} \quad (1.1)$$

Notiamo che la matrice della formula (1.1) è diagonale a blocchi, per cui X_1 è indipendente da $X_2 - C_{21}C_{11}^{-1}X_1$. Consideriamo quindi $X_2 - C_{21}C_{11}^{-1}X_1 | X_1 = x_1$, che per l'indipendenza appena mostrata e per l'Osservazione 1.5 sappiamo essere un vettore gaussiano. In più, siccome condizionatamente a $X_1 = x_1$, il termine $-C_{21}C_{11}^{-1}X_1$ si comporta come una costante, possiamo scrivere X_2 come somma dei termini $X_2 - C_{21}C_{11}^{-1}X_1$ e della costante $-C_{21}C_{11}^{-1}X_1$, che sappiamo essere una trasformazione affine di un vettore gaussiano. \square

1.2 Una prima interpretazione

Prima di definire in maniera formale il processo di Wiener, diamone un'interpretazione abbastanza intuitiva, partendo da una passeggiata aleatoria in \mathbb{Z} , nella quale le probabilità di transizione sono

$$P_{i,i+1} = \frac{1}{2} = P_{i,i-1} \quad \text{per } i = 0, \pm 1, \dots$$

Immaginiamo di velocizzare il processo, effettuando passi sempre più piccoli in intervalli di tempo sempre minori. Quindi per ogni unità di tempo Δt si effettua un passo di lunghezza Δx a destra o a sinistra con uguale probabilità. Il processo che si otterrà facendo tendere a zero la lunghezza dei passi e degli intervalli di tempo, sarà effettivamente un moto Browniano. Definiamo ora:

$$X(t) = \Delta x (X_1 + \dots + X_{\lfloor t/\Delta t \rfloor})$$

dove

$$X_i = \begin{cases} +1, & \text{se l'i-esimo passo di lunghezza } \Delta x \text{ è a destra} \\ -1, & \text{se è a sinistra} \end{cases}$$

$[t/\Delta t]$ è il più grande intero minore o uguale a $t/\Delta t$ e le X_i sono supposte indipendenti con

$$P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$$

Siccome $E(X_i) = 0$ e $Var(X_i) = E(X_i^2) = 1$, ricordando la proprietà di linearità del valore atteso e il comportamento della varianza per somme di variabili indipendenti e per il prodotto con una costante, si ottiene che

$$E(X(t)) = 0,$$

$$Var(X(t)) = (\Delta x)^2 \left[\frac{t}{\Delta t} \right]$$

Passiamo quindi al limite, ponendo $\Delta x = \sigma\sqrt{\Delta t}$ ⁴ per una qualche costante positiva σ e, facendo tendere Δt a 0, si ottiene che

$$E(X(t)) = 0,$$

$$Var(X(t)) \rightarrow \sigma^2 t$$

Ricordiamo due delle proprietà di una passeggiata aleatoria in \mathbb{Z} :

- I valori assunti dal processo in intervalli di tempo disgiunti sono indipendenti.
- La distribuzione di probabilità degli spostamenti della passeggiata aleatoria in un qualsiasi intervallo di tempo dipende solo dalla lunghezza dell'intervallo stesso.

Queste due proprietà seguono direttamente dal fatto che una passeggiata aleatoria in \mathbb{Z} è una catena di Markov omogena.

Risulta quindi plausibile considerare il processo, ottenuto passando al limite, stazionario e ad incrementi indipendenti. Inoltre, ricordando il teorema del limite centrale, $X(t)$ segue una distribuzione normale di media 0 e varianza $\sigma^2 t$. Infine, questa interpretazione del moto Browniano come limite di una passeggiata aleatoria, suggerisce che $X(t)$ debba essere una funzione continua in t . Queste quattro proprietà alle quali si è giunti per via intuitiva risultano essere caratterizzanti di un processo di Wiener, come si vedrà a breve.

1.3 Definizioni

Prima di dare una definizione formale di moto Browniano bisogna però ancora definire alcune concetti, in particolare quelli di filtrazione e di traiettoria.

⁴Questa scelta permette di ottenere un processo non banale. Se infatti avessimo posto $\Delta x = \Delta t$, allora $X(t)$ sarebbe stato $= 0$ con probabilità 1. Scegliendo invece $\Delta x = 1/\Delta t$ il risultato del limite sarebbe stato $+\infty$.

Definizione 1.8. Dato un processo $\{X_t\}_{t \in T}$ definito su uno spazio di probabilità (Ω, \mathcal{F}, P) , la mappa $t \rightarrow X_t(\omega)$, a valori in \mathbb{R} , si dice *traiettoria relativa a $\omega \in \Omega$* . Le traiettorie di un processo $\{X_t\}_{t \in T}$ sono tutte le mappe $t \rightarrow X_t(\omega)$ al variare di $\omega \in \Omega$.

Definiamo quindi il concetto di filtrazione.

Definizione 1.9. Dato uno spazio misurabile (Ω, \mathcal{F}) e un insieme $T = [0, \infty)$, si dice *filtrazione una famiglia crescente $\{\mathcal{F}_t\}_{t \in T}$ di sotto σ -algebre di \mathcal{F} , cioè tale che $\mathcal{F}_s \subseteq \mathcal{F}_t$ per ogni $s, t \in T$ con $s \leq t$* .

La σ -algebra \mathcal{F}_t rappresenta l'informazione disponibile fino all'istante t . Più precisamente, \mathcal{F}_t contiene tutti gli eventi conoscibili entro l'istante t , ossia quelli per cui è possibile dire se si siano verificati o no.

Facciamo due esempi di filtrazione: la filtrazione banale e la filtrazione naturale.

La filtrazione banale è quella filtrazione $\{\mathcal{F}_t\}_{t \in T}$ formata da una sola sotto σ -algebra, la σ -algebra banale $\{\Omega, \emptyset\}$, per ogni $t \in T$.

La filtrazione naturale, invece, estende il concetto di σ -algebra generata da una variabile aleatoria⁵, prendendo in considerazione più variabili. Definiamo $\mathcal{F}_s^X := \sigma(\{X_u\}_{0 \leq u \leq s})$ come la σ -algebra generata dalle variabili $\{X_u\}_{0 \leq u \leq s}$, cioè la più piccola σ -algebra che renda misurabili tutte le applicazioni X_u , per $0 \leq u \leq s$. La famiglia $\{\mathcal{F}_s^X\}_{s \in T}$ è detta *filtrazione naturale del processo X*.

Esiste anche una classe molto importante di filtrazioni, ossia le filtrazioni complete. Una filtrazione si dice *completa* se ogni σ -algebra \mathcal{F}_t contiene tutti gli insiemi P -trascurabili, cioè tutti gli eventi di probabilità nulla. Data una generica filtrazione $\{\mathcal{F}_t\}_{t \in T}$ è sempre possibile estenderla a una filtrazione completa: basta sostituire, per ogni $t \in T$, \mathcal{F}_t con $\bar{\mathcal{F}}_t = \sigma(\mathcal{F}_t, \mathcal{N})$, dove \mathcal{N} rappresenta tutti gli insiemi P -trascurabili.⁶

Definizione 1.10. Uno spazio di probabilità (Ω, \mathcal{F}, P) munito di una filtrazione $\{\mathcal{F}_t\}_{t \in T}$ è detto *spazio di probabilità filtrato* e si indica $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in T}, P)$. Se la filtrazione è completa lo spazio si chiama anche *base stocastica*.

Definizione 1.11. Sia assegnato un processo $\{X_t\}_{t \in T}$ su una base stocastica $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in T}, P)$ a valori nello spazio misurabile (\mathbb{R}, τ) , dove τ è la σ -algebra boreliana. Si dice che $\{X_t\}_{t \in T}$ è *adattato alla filtrazione (o adattato tout court)* se, per ogni $t \in T$, la variabile X_t è \mathcal{F}_t -misurabile, cioè se X_t è misurabile come applicazione da (Ω, \mathcal{F}_t) in (\mathbb{R}, τ) .

Per come è stata definita la filtrazione naturale, risulta evidente che il processo X è adattato alla sua filtrazione naturale $\{\mathcal{F}_t^X\}_{t \in T}$. Inoltre si tratta

⁵Si veda l'appendice A

⁶Perchè vale anche la relazione $\{\mathcal{F}_s \cup \mathcal{N}\} \subseteq \{\mathcal{F}_t \cup \mathcal{N}\}$ per $s < t$

della più piccola filtrazione a cui X sia adattato: ciò significa che, se X è adattato a $\{\mathcal{F}_t\}_{t \in T}$, si deve avere che $\mathcal{F}_t^X \subseteq \mathcal{F}_t$ per ogni $t \in T$.

Siamo ora pronti per dare la definizione di moto Browniano (uni-dimensionale).

Definizione 1.12. *Un processo stocastico $B = \{B_t\}_{t \geq 0}$ a valori in \mathbb{R} definito su una base stocastica $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ è detto moto Browniano rispetto alla filtrazione $\{\mathcal{F}_t\}_{t \geq 0}$, se è adattato a $\{\mathcal{F}_t\}_{t \geq 0}$ e se soddisfa le seguenti proprietà:*

1. $B_0 = 0$ q.c.;
2. $B_t - B_s$ è indipendente da \mathcal{F}_s per ogni $0 \leq s \leq t$;
3. $B_t - B_s \sim N(0, \sigma^2(t - s))$ per ogni $0 \leq s \leq t$;
4. Esiste $A \in \mathcal{F}$ con $P(A) = 1$ t.c. per ogni $\omega \in A$ la funzione $t \mapsto B_t(\omega)$ è continua, cioè q.c. B ha traiettorie continue.

In questa definizione è facile ritrovare le proprietà dedotte intuitivamente in precedenza. In particolare la proprietà (2) della definizione corrisponde al richiedere che gli incrementi siano indipendenti, mentre in (2) e (3) è rintracciabile la stazionarietà del processo.

Quando $\sigma = 1$, il processo sopra definito è detto *moto Browniano standard*⁷. Dove non diversamente espresso, considereremo sempre moti Browniani standard.

È anche abbastanza intuitivo estendere la definizione appena introdotta al caso n -dimensionale: infatti è sufficiente modificare che il processo stocastico B sia a valori in \mathbb{R}^n e che $B_t - B_s \sim N^n(0, \sigma^2(t - s)I_n)$ per ogni $0 \leq s \leq t$.

Definiamo ora alcune variazioni sul moto Browniano.

Definizione 1.13. *Dato un moto Browniano standard $\{B_t\}_{t \geq 0}$, allora $X(t) = \sigma B(t) + \mu t$ si dice moto Browniano con coefficiente di deriva μ e parametro di varianza σ^2 .*⁸

Definizione 1.14. *Se $\{Y_t\}_{t \geq 0}$ è un moto Browniano con coefficiente di deriva μ e parametro di varianza σ^2 , allora il processo $\{X_t\}_{t \geq 0}$ definito come*

$$X(t) = \exp(Y(t))$$

è chiamato moto Browniano geometrico.

Il moto Browniano geometrico non è un moto Browniano come definito in (1.12). Questa variazione del processo di Wiener risulta particolarmente utile nel modellare i prezzi delle azioni in periodi in cui si pensa che le variazioni percentuali possano essere indipendenti ed identicamente distribuite.

⁷Ogni moto Browniano X può essere comunque trasformato in un moto Browniano standard ponendo $B(t) = X(t)/\sigma$

⁸ $X(t)$ è una normale di media μt e varianza $\sigma^2 t$

1.4 Caratteristiche del moto Browniano

Partiamo subito con un'osservazione.

Osservazione 1.15. *Il processo di Wiener è un processo Markoviano omogeneo.*

Dimostrazione. Abbiamo già osservato che un moto Browniano è stazionario. Ricordiamo ora la definizione di processo Markoviano continuo a stati continui. Siano $T = [0, +\infty)$, l'insieme degli stati $S = [0, +\infty)$ e $I \subseteq S$, allora il processo stocastico $\{X_t\}_{t \in T}$ è Markoviano se per ogni $s \in S$ e per ogni $t, s \in T$

$$\begin{aligned} P_{I,j}(t, s) &:= P(X_{t+s} \in I | X_s = j) \\ &= P(X_{t+s} \in I | X_s = j, X_u = x_u \text{ per ogni } u < s) \end{aligned}$$

Ma in uno spazio di probabilità filtrato, come è stato definito in precedenza per il moto Browniano, questo corrisponde a dire che $B_t - B_s$ è indipendente da F_s per ogni $0 \leq s < t$. \square

Introduciamo ora un'altra classe di processi di fondamentale importanza.

Definizione 1.16. *Se un processo stocastico $\{X_t\}_{t \geq 0}$ ha la proprietà che il vettore $(X_{t_1}, \dots, X_{t_n})$ segue una distribuzione multivariata Gaussiana per ogni scelta di $0 < t_1 \dots < t_n$ allora $\{X_t\}_{t \geq 0}$ è chiamato processo Gaussiano.*

Il moto Browniano è il processo Gaussiano per eccellenza e molti dei più importanti modelli in probabilità conducono a processi Gaussiani.

Proposizione 1.17. *Un moto Browniano $\{B_t\}_{t \geq 0}$ è un processo Gaussiano.*

Dimostrazione. Fissiamo dei tempi $0 < t_1 \dots < t_n$, chiamiamo gli incrementi $Y := B_{t_i} - B_{t_{i-1}}$, per $i = 1, \dots, n$ e $y = (y_1, \dots, y_n)$ (con $t_0 = 0$). Il vettore $Y = (Y_1, \dots, Y_n)$ ha componenti normali e indipendenti per le proprietà (2) e (3) della Definizione 1.12: quindi Y è un vettore normale. Sempre per la proprietà (3) della Definizione 1.12 e per la linearità del valore atteso sappiamo che la sua media è nulla. Verifichiamo ora che la sua matrice di covarianza è diagonale. Per $i < j$

$$\begin{aligned} Cov(Y_i, Y_j) &= Cov(B_{t_i} - B_{t_{i-1}}, B_{t_j} - B_{t_{j-1}}) \\ &= E((B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})) \\ &= E(B_{t_i} B_{t_j}) - E(B_{t_i} B_{t_{j-1}}) - E(B_{t_{i-1}} B_{t_j}) + E(B_{t_{i-1}} B_{t_{j-1}}) \\ &= t_i - t_i - t_{i-1} + t_{i-1} = 0 \end{aligned}$$

Quindi $\Gamma_{ij} := Cov(Y_i, Y_j) = (t_i - t_{i-1})\delta_{ij}$. Ora osserviamo che

$$f_{Y_i}(y_i) = \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left(\frac{-\frac{1}{2}y_i^2}{t_i - t_{i-1}}\right)$$

Ma siccome le variabili Y_1, \dots, Y_n sono indipendenti, allora

$$\begin{aligned} f_Y(y) &= \prod_{i=1}^n f_{Y_i}(y_i) \\ &= \frac{1}{(2\pi)^{n/2} \sqrt{t_1(t_2 - t_1) \dots (t_n - t_{n-1})}} \exp\left(-\frac{1}{2} \sum_{i=1}^n \frac{y_i^2}{t_i - t_{i-1}}\right) \end{aligned} \quad (1.2)$$

che è la distribuzione di una normale multivariata.

Dato che $(B_{t_1}, \dots, B_{t_n})$ è una trasformazione lineare (invertibile) di Y , più precisamente $B_{t_i} = Y_1 + \dots + Y_i$, segue che anche il vettore aleatorio $(B_{t_1}, \dots, B_{t_n})$ è normale. \square

È facile calcolare la funzione di covarianza per un moto Browniano (standard) semplicemente notando che per qualsiasi $s \leq t$ si ha che $Cov(B_s, B_t) = E((B_t - B_s + B_s)B_s) = E(B_s^2) = s$.⁹, per cui in generale succede che

$$Cov(B_s, B_t) = \min(s, t) \quad 0 < s, t < \infty \quad (1.3)$$

Il prossimo lemma fornisce una nuova visione del moto Browniano come caso particolare di un processo Gaussiano.

Lemma 1.18. *Se un processo Gaussiano $\{X_t\}_{t \geq 0}$ ha le seguenti caratteristiche:*

- $E(X_t) = 0$;
- $Cov(X_s, X_t) = \min(s, t) \quad 0 < s, t < \infty$;
- *Traiettorie q.c. continue;*
- $X_0 = 0$.

Allora $\{X_t\}_{t \geq 0}$ è un moto Browniano in $[0, \infty)$.

Questo lemma sfrutta i calcoli effettuati nella dimostrazione della Proposizione 1.17 e nel calcolo della covarianza di un moto Browniano: infatti se la covarianza ha la stessa forma dell'equazione (1.3), allora $\{X_t\}_{t \geq 0}$ deve avere incrementi indipendenti. A questo punto è facile ritrovare nel lemma tutte le proprietà presenti nella Definizione 1.12 del moto Browniano.

Vediamo ora un semplice esempio che sfrutta questo lemma.

Esempio 1.19. *Dato un processo di Wiener $\{B_t\}_{t \in [0, \infty)}$, verificare se i processi $\{X_t\} = -B_t$ e $\{Y_t\} = B_{2t} - B_t$ sono ancora processi di Wiener.*

Per controllare se i processi $\{X_t\}$ e $\{Y_t\}$ sono ancora processi di Wiener conviene utilizzare la caratterizzazione data dal Lemma 1.18. Infatti in entrambi i casi è immediato verificare

⁹Si utilizza la proprietà degli incrementi indipendenti.

che i processi sono gaussiani, in quanto le loro componenti sono funzioni lineari delle componenti del processo gaussiano $\{B_t\}_{t \in [0, \infty)}$. Anche la relazione $E(X_t) = E(Y_t) = 0$ si verifica facilmente, sfruttando la linearità del valore atteso. È verificata anche la continuità delle traiettorie dei nuovi processi perchè ottenute componendo le traiettorie di $\{B_t\}_{t \in [0, \infty)}$, che sappiamo essere continue, con funzioni continue. Il punto (d) del lemma è sempre verificato. I problemi possono nascere con il calcolo della covarianza. Per $s \leq t$

$$\text{Cov}(X_t, X_s) = \text{Cov}(-B_t, -B_s) = \text{Cov}(B_t, B_s) = s$$

Quindi X_t è ancora un moto Browniano. Proviamo con Y_t

$$\begin{aligned} \text{Cov}(Y_t, Y_s) &= \text{Cov}(B_{2t} - B_t, B_{2s} - B_s) \\ &= \text{Cov}(B_{2t}, B_{2s}) - \text{Cov}(B_{2t}, B_s) - \text{Cov}(B_t, B_{2s}) + \text{Cov}(B_t, B_s) \\ &= 2s - s + s - \text{Cov}(B_t, B_{2s}) \\ &= 2s - \text{Cov}(B_t, B_{2s}) \end{aligned}$$

che in ogni caso non è uguale al $\min(s, t)$. Quindi Y_t non è un moto Browniano.

Concludiamo con un'ultima osservazione. Le proprietà (1), (2) e (3) della Definizione 1.12 sono proprietà che la legge del generico vettore aleatorio $(B_{t_1}, \dots, B_{t_n})$ deve avere per $t_1 < \dots < t_n$. Di conseguenza, dato un processo stocastico $X = \{X_t\}_{t \geq 0}$, se il vettore aleatorio $(X_{t_1}, \dots, X_{t_n})$ segue una distribuzione normale multivariata di media nulla e varianza $t_1(t_2 - t_1) \dots (t_n - t_{n-1})$ e tale che $X_0 = 0$, è sufficiente mostrare che X ha traiettore q.c. continue e si ottiene che X è un moto Browniano.

1.5 Il ponte Browniano

Nella dimostrazione della Proposizione 1.17, abbiamo calcolato la distribuzione congiunta del vettore degli incrementi di un processo di Wiener, col fine di dimostrare che un moto Browniano è un processo Gaussiano. Nella sezione precedente non eravamo interessati a quale fosse la formula della densità di un moto Browniano; ci bastava sapere che fosse normale. Ora però per studiare il ponte Browniano ne abbiamo bisogno. Notiamo che per l'indipendenza degli incrementi

$$\prod_{i=1}^n f_{B_{t_i} - B_{t_{i-1}}}(b_{t_i}) = f_B(b)$$

dove $B = (B_{t_1}, \dots, B_{t_n})$ per $0 < t_1 < \dots < t_n$ e $b = (b_{t_1}, \dots, b_{t_n})$. Ora basandoci sulla formula (1.1) si vede subito che

$$\begin{aligned} f_B(b) &= f_{B_{t_1}}(b_{t_1}) f_{B_{t_2} - B_{t_1}}(b_{t_2} - b_{t_1}) \dots f_{B_{t_n} - B_{t_{n-1}}}(b_{t_n} - b_{t_{n-1}}) \\ &= \frac{\exp \left\{ -\frac{1}{2} \left[\frac{b_{t_1}^2}{t_1} + \frac{(b_{t_2} - b_{t_1})^2}{b_{t_2} - t_1} \dots \frac{(b_{t_n} - b_{t_{n-1}})^2}{t_n - t_{n-1}} \right] \right\}}{(2\pi)^{n/2} \sqrt{t_1(t_2 - t_1) \dots (t_n - t_{n-1})}} \end{aligned}$$

Calcoliamo anche la distribuzione di B_s condizionato a $B_t = R$ dove $s < t$

$$\begin{aligned}
 f_{B_s|B_t=R}(b_s|R) &= \frac{f_{B_s, B_t}(b_s, R)}{f_{B_t}(R)} \\
 &= \frac{f_{B_s}(b_s)f_{B_t-B_s}(R-b_s)}{f_{B_t}(R)} \\
 &\propto \exp\left(-\frac{b_s^2}{2s} - \frac{(R-b_s)^2}{2(t-s)}\right) \\
 &\propto \exp\left(-b_s^2\left(\frac{1}{2s} + \frac{1}{2(t-s)}\right) + \frac{Rb_s}{t-s}\right) \\
 &\propto \exp\left(-\frac{t}{2s(t-s)}\left(b_s^2 - 2\frac{sR}{t}b_s\right)\right) \\
 &\propto \exp\left(-\frac{(b_s - \frac{Rs}{t})^2}{\frac{2s(t-s)}{t}}\right)
 \end{aligned}$$

Quindi per il Teorema 1.7 sappiamo che la distribuzione condizionata di B_s dato $B_t = R$ è, per $s < t$, normale di media $\frac{s}{t}R$ e varianza $\frac{s}{t}(t-s)$.

Definizione 1.20. Dato un moto Browniano standard $\{B_t\}_{t \geq 0}$, il processo stocastico $\{B_t, 0 \leq t \leq 1 | B(1) = 0\}$ si chiama ponte Browniano.

Il nome "ponte" deriva dal fatto che il processo è obbligato ad essere uguale a 0 agli estremi del dominio.

Per il Teorema 1.7, sappiamo che la distribuzione di un generico vettore aleatorio per un ponte Browniano è una normale multivariata. Quindi anche questo processo è un processo gaussiano.

Calcoliamo ora media e covarianza per un ponte Browniano, fissati $s < t < 1$. Nei prossimi calcoli sfrutteremo i risultati appena ottenuti per il valore medio e la varianza di B_s dato $B_t = R$ e le proprietà del valore atteso condizionato.

$$E(B_s | B(1) = 0) = 0,$$

$$\begin{aligned}
 Cov(B_s | B(1) = 0, B_t | B(1) = 0) &= Cov((B_s, B_t) | B(1) = 0) \\
 &= E(B_s B_t | B(1) = 0) \\
 &= E(E(B_s B_t | B_t, B(1) = 0) | B(1) = 0) \\
 &= E(E(B_s B_t | B_t) | B(1) = 0) \\
 &= E(B_t E(B_s | B_t) | B(1) = 0) \\
 &= E\left(B_t \frac{s}{t} B_t | B(1) = 0\right) \\
 &= \frac{s}{t} E(B_t^2 | B(1) = 0) \\
 &= \frac{s}{t} t(1-t) \\
 &= s(1-t)
 \end{aligned}$$

Quindi il ponte Browniano può anche essere definito come un processo Gaussiano di media 0 e funzione di covarianza $s(1-t)$ per $s < t < 1$. Vediamo ora quali sono le trasformazioni possibili per passare da un moto Browniano ad un ponte Browniano e viceversa.

Proposizione 1.21. *Se $\{B_t\}_{0 \leq t \leq 1}$ è un ponte Browniano e Z una variabile gaussiana standard indipendente, allora $W_t = B_t + tZ$ è un processo di Wiener per $t \in [0, 1]$.*

Dimostrazione. Per la dimostrazione ci basiamo sul lemma (1.18). Bisogna prima di tutto dimostrare che W_t è un processo Gaussiano: ma questo è immediato in quanto W_t è una combinazione lineare di variabili normali. Inoltre, per costruzione, il ponte Browniano è q.c. continuo e uguale a 0 in 0 e anche dopo la trasformazione queste due proprietà rimangono valide. Controlliamo quindi il valore atteso e la covarianza, per $s < t < 1$.

$$E(W_t) = E(B_t + tZ) = E(B_t) + tE(Z) = 0$$

$$\begin{aligned} Cov(W_t, W_s) &= Cov(B_t + tZ, B_s + sZ) \\ &= Cov(B_t, B_s) + Cov(B_t, sZ) + \\ &\quad + Cov(tZ, B_s) + Cov(tZ, sZ) \\ &= s(1-t) + st = s - st + st = s \end{aligned}$$

□

Proposizione 1.22. *Se $\{X_t, \}_{t \geq 0}$ è un moto Browniano standard, allora $\{Z_t\}_{0 \leq t \leq 1}$ è un ponte Browniano dove $Z_t = X_t + tX(1)$.*

Si tralascia la dimostrazione di quest'ultima proposizione, in quanto segue lo stesso ragionamento della dimostrazione della Proposizione 1.21.

1.6 Proprietà delle traiettorie

Studieremo due caratteristiche delle traiettorie dei moti Browniani, partendo da due concetti di analisi: le funzioni a variazione finita e le funzioni Hölderiane. In particolare per le funzioni a variazione finita forniremo le dimostrazioni dei risultati che si otterranno, mentre per le funzioni Hölderiane si verificheranno solamente in seguito attraverso la simulazione. Tuttavia alcune considerazioni proprio su queste ultime funzioni ci permetteranno di dare un risultato molto forte sulla differenziabilità di un moto Browniano.

1.6.1 Funzioni a variazioni finita

Definizione 1.23. Data una funzione $f : \mathbb{R}^+ \rightarrow \mathbb{R}$, si dice variazione di f in $[a, b]$ la quantità

$$V_b^a f = \sup_{\pi} \sum_{i=0}^n |f(x_{i+1}) - f(x_i)|$$

dove π varia fra tutte le partizioni $a = x_0 < x_1 < \dots < x_{n+1} = b$ dell'intervallo $[a, b]$.

f si dice a variazione finita se $V_b^a f < +\infty$.

f si dice a variazione limitata se $V_{+\infty}^0 f < +\infty$.

Introdotta questa classe di funzioni enunciamo e dimostriamo la seguente proposizione.

Proposizione 1.24. Le traiettorie di un moto Browniano $\{B_t\}$ non sono q.c. a variazione finita su alcun intervallo.

Dimostrazione. Fissiamo $s = t_0 < t_1 \dots < t_m = t$ e siano

$$\pi = \{t_0, \dots, t_m\} \text{ una partizione dell'intervallo } [s, t],$$

$$|\pi| = \max_k \{t_{k+1} - t_k \text{ per } k = 0, \dots, m-1\},$$

$$\tau = t - s,$$

$$\text{e } S_{\pi} = \sum_{k=0}^{m-1} |B_{t_{k+1}} - B_{t_k}|^2.$$

Allora

$$\tau = \sum_{k=0}^{m-1} (t_{k+1} - t_k)$$

$$S_{\pi} - \tau = \sum_{k=0}^{m-1} [(B_{t_{k+1}} - B_{t_k})^2 - (t_{k+1} - t_k)]$$

$$E[(S_{\pi} - \tau)^2] = E \left[\left(\sum_{k=0}^{m-1} (B_{t_{k+1}} - B_{t_k})^2 - (t_{k+1} - t_k) \right)^2 \right]$$

Ora siccome le v.a. $(B_{t_{k+1}} - B_{t_k})^2 - (t_{k+1} - t_k)$ sono indipendenti e centrate e sfruttando la proprietà del valore atteso per il prodotto di v.a. indipendenti, succede che

$$\begin{aligned} E[(S_{\pi} - \tau)^2] &= \sum_{k=0}^{m-1} E[(B_{t_{k+1}} - B_{t_k})^2 - (t_{k+1} - t_k)]^2 \\ &= \sum_{k=0}^{m-1} (t_{k+1} - t_k)^2 E \left[\left(\frac{(B_{t_{k+1}} - B_{t_k})^2}{t_{k+1} - t_k} - 1 \right)^2 \right] \end{aligned}$$

ma per ogni k la v.a. $\frac{B_{t_{k+1}} - B_{t_k}}{\sqrt{t_{k+1} - t_k}}$ è $N \sim (0, 1)$, quindi

$$c = E \left[\left(\frac{(B_{t_{k+1}} - B_{t_k})^2}{t_{k+1} - t_k} - 1 \right)^2 \right]$$

non dipende da k , allora

$$E(S_\pi - \tau)^2 = c \sum_{k=0}^{m-1} (t_{k+1} - t_k)^2 \leq |\pi|(t - s) \xrightarrow{|\pi| \rightarrow 0} 0.$$

Per cui si ha che

$$\lim_{|\pi| \rightarrow 0} E((S_\pi - (t - s))^2) = 0.$$

Si dice che

$$\lim_{|\pi| \rightarrow 0} S_\pi = t - s \text{ in } L^2.^{10}$$

Riconsideriamo S_π e scriviamolo in un altro modo.

$$S_\pi = \sum_{k=0}^{m-1} |B_{t_{k+1}} - B_{t_k}|^2 \leq \sup_i |B_{t_{i+1}} - B_{t_i}| \sum_{k=0}^{m-1} |B_{t_{k+1}} - B_{t_k}|$$

Se le traiettorie fossero a variazione finita su $[s, t]$ si avrebbe

$$\lim_{|\pi| \rightarrow 0} S_\pi(\omega) = 0$$

ma questo è in contraddizione con quanto calcolato in precedenza. \square

Questo risultato indica che le traiettorie di un processo di Wiener, pur essendo continue per definizione, oscillano molto velocemente tanto che la somma dei moduli degli spostamenti al limite è uguale a $+\infty$, dandoci un primo segnale che possono avere dei comportamenti "strani".

1.6.2 Funzioni Hölderiane

In questa sezione daremo la definizione di funzione γ -Hölderiana in un intervallo e in un punto e vedremo un particolare tipo di funzioni Hölderiane, ossia le funzioni Lipschitziane. Enunceremo quindi alcuni risultati legati alle funzioni che stiamo per introdurre per i moti Browniani.

Definizione 1.25. Una funzione $f : [a, b] \rightarrow \mathbb{R}$ si dice γ -Hölderiana (su $[a, b]$), con $\gamma \in (0, 1]$, se

$$[f]_\gamma := \sup_{t, s \in [a, b], t \neq s} \frac{|f(t) - f(s)|}{|t - s|^\gamma} < \infty$$

¹⁰ L^2 è lo spazio delle funzioni quadrato integrabili.

Invece una funzione $g : [0, \infty) \rightarrow \mathbb{R}$ si dice γ -Hölderiana nel punto $s \geq 0$, con $\gamma \in (0, 1]$ se esiste $\eta > 0$, $b > 0$, per cui $|t - s| < \eta$, $t \geq 0$, implica

$$|g(t) - g(s)| \leq b|t - s|^\gamma$$

Le funzioni Lipschitziane sono funzioni 1-Hölderiane. Ovvero una funzione $f : [0, +\infty) \rightarrow \mathbb{R}$ è Lipschitziana in \mathbb{R} se $\exists K \geq 0$ tale che

$$|f(t) - f(s)| \leq K|t - s|$$

per ogni $t, s \in [0, +\infty)$.

Se una variabile reale è lipschitziana vuol dire che ha "crescita limitata", nel senso che il rapporto tra variazione di ordinata e variazione di ascissa non può mai superare un valore fissato K . La lipschizianità ha due implicazioni:

- la limitatezza del rapporto incrementale;
- la continuità della funzione in \mathbb{R} .

Enunciamo quindi i risultati che riguardano i moti Browniani

Proposizione 1.26. *Le traiettorie di un processo di Wiener $\{B_t\}$, q.c., su ogni $[0, T]$, $T > 0$, sono γ -Hölderiane per qualunque $\gamma \in (0, 1/2)$.*

Teorema 1.27. *Sia $\{B_t\}$ un processo di Wiener. Per ogni $\gamma > 1/2$, q.c. le traiettorie di $\{B_t\}$ non sono γ -Hölderiane in alcun punto $s \geq 0$.*

Corollario 1.28. *Le traiettorie di un processo di Wiener $\{B_t\}$, q.c., non sono derivabili in nessun $t \geq 0$.*

Dimostrazione. Il teorema (1.27) ci dice che le traiettorie di un moto Browniano non sono lipschitziane in alcun punto. Ma allora per ogni $\eta > 0$, $b > 0$ $\exists t \geq 0$ t.c. $|t - s| < \eta$ per cui

$$\frac{|B(t) - B(s)|}{|t - s|} \geq b.$$

Ma allora

$$\lim_{t \rightarrow s} \left| \frac{B(t) - B(s)}{t - s} \right| \quad \text{non esiste finito.}$$

□

Capitolo 2

Simulazione di traiettorie di moti Browniani

In questo secondo capitolo, dopo una breve panoramica su alcuni metodi di simulazione, introdurremo un algoritmo per simulare le traiettorie dei moti Browniani e rintracceremo le caratteristiche di questo processo stocastico, studiate nel capitolo precedente, nei grafici prodotti e all'interno dell'algoritmo.

2.1 La simulazione

Ci concentriamo su quell'aspetto della simulazione che riguarda la generazione di numeri casuali provenienti da una certa distribuzione di probabilità. Ossia si vorrebbe generare delle realizzazioni x_1, \dots, x_n provenienti da n variabili aleatorie reali indipendenti e identicamente distribuite X_1, \dots, X_n con una certa funzione di ripartizione F .

Definizione 2.1. *Si definisce flusso di simulazione da una distribuzione di probabilità, la cui funzione di ripartizione è F , una sequenza numerica x_1, \dots, x_n , generata artificialmente, di realizzazioni di variabili aleatorie reali X_1, \dots, X_n indipendenti ed identicamente distribuite con funzione di distribuzione cumulativa F .*

Il problema di generare un flusso di simulazione x_1, \dots, x_n da F si scompone in due sottoproblemi.

1. come generare un flusso di simulazione $u_1 \dots u_n$ dalla distribuzione Uniforme(0,1);
2. come ottenere $x_1 \dots x_n$ a partire da $u_1 \dots u_n$.

Il primo punto è necessario in quanto ogni algoritmo per la generazione di numeri provenienti da una generica distribuzione di probabilità utilizza

numeri provenienti da una Uniforme in $(0,1)$. Bisogna quindi spiegare come sia possibile reperire dei numeri casuali provenienti da una distribuzione uniforme. In genere però si richiede che la successione u_1, \dots, u_n sia in realtà il prodotto di un algoritmo matematico e assuma così un carattere deterministico, in contrasto con la richiesta di casualità. La richiesta è motivata dalla necessità di essere in grado di riprodurre l'esperimento di simulazione, in genere tramite l'ausilio di un calcolatore. Si parlerà allora tecnicamente di numeri *pseudocasuali*. Questo chiarisce il senso dell'espressione "generata artificialmente" nella definizione di flusso.

Un'ampia classe di soluzioni consiste nel ricorrere a generatori congruenziali lineari. Per questa classe di generatori l' i -esimo numero casuale (intero) si ottiene dal precedente in base alla relazione ricorrente

$$x_i \equiv (ax_{i-1} + c) \pmod{m}$$

con $m > 0$ e $a, c, x_0 < m$. Quindi x_i è il resto della divisione di $ax_{i-1} + c$ per m . Se $c = 0$ si dice che si è utilizzato un metodo moltiplicativo, mentre se $c > 0$ il metodo è additivo. Inoltre x_0 si chiama seme della simulazione. A questo punto la sequenza u_1, \dots, u_n si ricava ponendo $u_i = x_i/m$.

La scelta di a, c e m è cruciale, in quanto si richiede che la sequenza dei numeri generati abbia un periodo ¹ lungo: tuttavia non si approfondisce questo aspetto in questa sede.

A questo punto rimane da approfondire il secondo sottoproblema della generazione di un flusso di simulazione, ossia come ottenere dei numeri provenienti da una certa distribuzione di probabilità. Esistono numerosi metodi per generare una qualsiasi variabile aleatoria. Nelle pagine seguenti ne svilupperemo due: il primo, detto metodo della trasformazione inversa, è valido per la maggior parte delle variabili con distribuzione continua, mentre il secondo, l'algoritmo di Box-Muller, è specifico per la generazione di variabili normali.

2.1.1 Il metodo della trasformazione inversa

Il metodo della trasformazione inversa si basa sulla seguente proposizione.

Proposizione 2.2. *Sia U una v.a. uniforme in $(0,1)$. Per ogni funzione di ripartizione continua e strettamente crescente F , se definiamo la v.a. X come*

$$X = F^{-1}(U)$$

allora la v.a. X ha funzione di ripartizione F .

¹Con periodo si intende la quantità di numeri generati fino ad ottenerne uno già presente nella sequenza.

Dimostrazione.

$$\begin{aligned} F_X(a) &= P(X \leq a) \\ &= P(F^{-1}(U) \leq a) \end{aligned}$$

Siccome $F(x)$ è una funzione monotona, allora $F^{-1}(U) \leq a$ se e solo se $U \leq F(a)$. Allora notiamo che

$$\begin{aligned} F_X(a) &= P(U \leq F(a)) \\ &= F(a) \end{aligned}$$

□

Possiamo quindi simulare una variabile aleatoria X partendo dalla sua funzione di ripartizione F : basta infatti simulare dei numeri provenienti da una v.a. uniforme U e porre $X = F^{-1}(U)$.

Esempio 2.3. *Generiamo delle variabili esponenziali di parametro λ . Detta $X \sim \text{Exp}(\lambda)$, ricordiamo che la funzione di ripartizione è*

$$F_X(x) = 1 - e^{-\lambda x}$$

per le $x > 0$. Fissato $F_X(x) = u$, invertiamo questa equazione e otteniamo

$$x = -\frac{1}{\lambda} \log(1 - u)$$

Allora, per ottenere un flusso di simulazione x_1, x_2, \dots , per X , a partire da un flusso uniforme u_1, u_2, \dots , basta prendere

$$x_1 = -\frac{1}{\lambda} \log(1 - u_1), \quad x_2 = -\frac{1}{\lambda} \log(1 - u_2), \dots$$

2.1.2 L'algoritmo di Box-Muller

L'algoritmo di Box-Muller è un algoritmo *ad hoc* per la generazione di numeri casuali normali e si basa sulla trasformazione in coordinate polari. Consideriamo due variabili normali standard indipendenti X e Y , per cui la loro distribuzione congiunta risulta essere

$$f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}, \quad -\infty < x, y < +\infty$$

Consideriamo quindi la trasformazione in coordinate polari di queste due variabili aleatorie ²

$$\begin{cases} R^2 = X^2 + Y^2 \\ \Theta = \arctan(Y/X) \end{cases}$$

²La trasformazione è sempre invertibile se si esclude il punto $(0, 0)$, che ha comunque probabilità nulla.

Fissati $d = x^2 + y^2$ e $\theta = \arctan(y/x)$, calcoliamo la matrice Jacobiana della trasformazione.

$$\begin{aligned} J &= \begin{bmatrix} \frac{\partial d}{\partial x} & \frac{\partial d}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{bmatrix} \\ &= \begin{bmatrix} 2x & 2y \\ \frac{1}{1+y^2/x^2} \left(\frac{-y}{x^2}\right) & \frac{1}{1+y^2/x^2} \left(\frac{1}{x}\right) \end{bmatrix} \\ &= \begin{bmatrix} 2x & 2y \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{bmatrix} \end{aligned}$$

Calcoliamo ora il modulo del determinante della matrice J.

$$\begin{aligned} |\det(J)| &= \left| 2x \frac{x}{x^2 + y^2} + 2y \frac{y}{x^2 + y^2} \right| \\ &= \left| \frac{2x^2}{x^2 + y^2} + \frac{2y^2}{x^2 + y^2} \right| \\ &= \left| \frac{2(x^2 + y^2)}{x^2 + y^2} \right| \\ &= 2 \end{aligned}$$

Allora la densità congiunta di R^2 e Θ è data da

$$\begin{aligned} f_{R^2, \Theta}(d, \theta) &= \frac{1}{2\pi} e^{-d/2} \frac{1}{2} \\ &= \frac{1}{2} e^{-d/2} \frac{1}{2\pi}, \quad 0 < d < +\infty, \quad 0 < \theta < 2\pi. \end{aligned}$$

Ma allora $R^2 \sim \text{Exp}(1/2)$ e $\Theta \sim U(0, 2\pi)$ indipendenti.³

Effettuiamo ora il procedimento inverso, ossia dalle coordinate polari torniamo a quelle cartesiane: allora $X = \sqrt{R^2} \cos(\Theta)$ e $Y = \sqrt{R^2} \sin(\Theta)$ sono delle gaussiane standard indipendenti. Ricordando l'Esempio 2.3⁴, si nota che, se u_1 e u_2 sono due realizzazioni indipendenti da una distribuzione uniforme in $(0, 1)$, allora

$$\begin{aligned} x &= \sqrt{-2 \log u_1} \cos(2\pi u_2) \\ y &= \sqrt{-2 \log u_1} \sin(2\pi u_2) \end{aligned}$$

sono due realizzazioni indipendenti di variabili normali standard.

³Si noti che siamo riusciti a dimostrare che una distribuzione esponenziale di parametro $1/2$ è uguale a una chi-quadro a 2 gradi di libertà.

⁴Bisogna anche calcolare la funzione inversa della funzione di ripartizione di una v.a. uniforme in $(0, 2\pi)$. Ricordando che se $X \sim U(0, 2\pi)$, allora $F_X(x) = x/2\pi$. Posto $u = F_X(x)$, si vede che $x = 2\pi u$.

2.1.3 La simulazione multivariata

Fino a questo punto ci siamo concentrati sulla simulazione univariata, senza preoccuparci di come poter simulare dei vettori aleatori. Risulta però semplice la generazione multivariata in quanto si ricorre a una sequenza di simulazioni univariate. Infatti il generico flusso di simulazione x_1, \dots, x_n , che contiene delle realizzazioni provenienti dalle variabili aleatorie X_1, \dots, X_n (indipendenti ed identicamente distribuite), può anche essere considerato l'unica realizzazione del vettore aleatorio $X = (X_1, \dots, X_n)$. Si sceglierà caso per caso la sequenza più adeguata sulla base delle caratteristiche della distribuzione in esame.

Esempio 2.4. *Simuliamo una singola realizzazione da una normale multivariata. Nel caso standard, dove $Z \sim N^n(0, I_n)$, basta prendere un flusso di simulazione z_1, \dots, z_k normale standard. Nel caso generale, in virtù dell'Osservazione 1.4 e dato $z = (z_1, \dots, z_n)^t$, occorrerà invece prendere*

$$x = m + Az.$$

*Poichè la scelta di A non è univoca, risulta essere più efficace scegliere come A la fattorizzazione di Choleski di C , perchè si velocizzano i calcoli.*⁵

2.2 Un algoritmo di simulazione per moti Browniani

Nella sezione precedente abbiamo visto come generare variabili aleatorie uni/multi-variate; dobbiamo fare, però, ancora un altro passo e capire come poter simulare un processo stocastico. La simulazione di processi stocastici avviene semplicemente attraverso la generazione casuale di una sequenza di variabili aleatorie. A seconda del processo oggetto di studio si possono utilizzare algoritmi specifici, in quanto non esiste un algoritmo valido per ogni processo stocastico.

Per simulare le traiettorie di moti Browniani si è quindi utilizzato un algoritmo *ad hoc*. In particolare la simulazione riguarda un processo di Wiener W_t per $t \in [0, 1]$: questo però non restringe la simulazione ad una classe specifica di moti Browniani in quanto è sempre possibile estendere l'intervallo in cui il processo è definito. La motivazione di questo fatto sarà più chiara dopo aver analizzato il semplice algoritmo di simulazione:⁶

```
N=10000;
T=(1:N)/N;
X=randn(1,N)/sqrt(N);
W=cumsum(X,2);
plot(T,W);
```

⁵Si mantiene la nomenclatura dell'Osservazione 1.4.

⁶Il software utilizzato è Matlab, per cui riportiamo la sintassi specifica di questo programma.

Il numero di osservazioni del processo che vogliamo simulare è rappresentato da N .

T è l'insieme dei tempi. Con questa linea di codice si genera un vettore riga di dimensione N , nel quale la generica i -esima entrata assume il valore i/N . Ossia il vettore contiene una sequenza di numeri da $1/N$ a 1, in cui la relazione tra un numero e il suo successivo è: $T_{i+1} = T_i + 1/N$.

Il comando `X=randn(1,N)` della terza linea di codice genera un vettore riga⁷ a N entrate contenenti delle realizzazioni provenienti da N normali standard indipendenti. Si dividono le entrate per la radice di N in modo da sistemare i valori assunti dalla varianza dalle realizzazioni del processo di Wiener: mostreremo questo fatto conclusa l'analisi del codice di simulazione.

`W=cumsum(X,2)` genera il vettore delle somme cumulate a partire dal vettore X , ossia $W_i = X_i + X_{i-1} + \dots + X_1$ per ogni $i = 1, \dots, N$.⁸ W contiene le realizzazioni di un generico moto Browniano standard, come mostreremo a breve.

Infine con l'ultima linea creiamo il grafico delle traiettorie di un moto Browniano. In ascissa abbiamo l'insieme dei tempi, mentre in ordinata i valori del vettore W : stiamo quindi rappresentando una traiettoria di un moto Browniano tra 0 e 1. Infatti ad ogni T_i , che sappiamo essere uguale a i/N , associamo il valore dell' i -esima entrata del vettore W , che abbiamo detto essere un moto Browniano.

Nella pagina successiva mostriamo l'output di questo algoritmo⁹, ossia il grafico delle traiettorie di moti Browniani.

Rimane da chiarire perchè questo algoritmo genera un moto Browniano. Prima di tutto notiamo che, detta $Z_i \sim N(0, 1)$, allora $X_i = Z_i/\sqrt{N}$ per ogni $i = 1, \dots, N$ e

$$E(X_i) = E(Z_i/\sqrt{N}) = 0$$

$$\text{Var}(X_i) = \text{Var}(Z_i/\sqrt{N}) = \frac{1}{N} \text{Var}(Z_i) = \frac{1}{N}$$

Quindi $X_i \sim N(0, 1/N)$ per ogni $i = 1, \dots, N$.

Invece la generica entrata del vettore W_i , che abbiamo definito come $W_i = X_i + X_{i-1} + \dots + X_1$, ha media e varianza pari a

$$E(W_i) = E(X_i + \dots + X_1) = E(X_i) + \dots + E(X_1) = 0$$

⁷Cambiando il valore 1 di questa linea di codice con un un'altro numero naturale > 1 è possibile simulare più moti Browniani contemporaneamente. Ad esempio il comando `X=randn(4,N)` genera una matrice $4 \times N$, nella quale ogni entrata è una realizzazione di una normale standard indipendente dalle altre.

⁸Il valore 2 all'interno della parentesi indica che le somme cumulate devono essere effettuate per riga. Se infatti X fosse una matrice, questo ci consente di simulare più processi di Wiener contemporaneamente, ognuno con N realizzazioni.

⁹Si sono simulate più traiettorie contemporaneamente, nel modo spiegato nelle note precedenti.

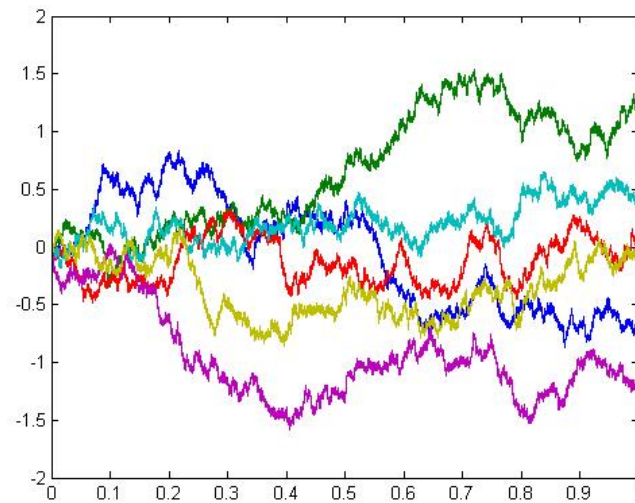


Figura 2.1: Traiettorie di un moto Browniano

$$\text{Var}(W_i) = \text{Var}(X_i + \dots + X_1) = \text{Var}(X_i) + \text{Var}(X_1) = \frac{i}{N}$$

e quindi $W_i \sim N(0, i/N)$.

Ricordando la corrispondenza che si crea con il comando plot tra i vettori T e W e che, dato un moto Browniano B , $B_t \sim N(0, t)$ (per la Definizione 1.12), allora W_t rappresenta una sequenza di realizzazioni di un moto Browniano tra 0 e 1. Si vorrebbe infatti che al tempo T_i , che abbiamo calcolato essere i/N , la realizzazione fosse estratta da una normale $N \sim (0, i/N)$, cosa che abbiamo ottenuto attraverso la divisione di ogni entrata del vettore X per \sqrt{N} . Bisogna ancora verificare che tutte le altre proprietà della Definizione 1.12 siano valide, cosa che ci apprestiamo a svolgere nella prossima sezione.

Prima però un'ultima osservazione. Più si aumenta il numero di osservazioni N , più queste saranno vicine tra di loro, in quanto è fissata l'ampiezza dell'intervallo in cui vengono plottate: per questo motivo si è deciso di fissare come dominio l'intervallo $(0,1]$. Se si vuole estendere questo intervallo è sufficiente generare altri N numeri provenienti da una normale standard e continuare con le somme cumulate dall'entrata $N + 1$ del vettore W .

2.3 Proprietà del moto Browniano simulato

Nella sezione precedente abbiamo dimostrato che $W_i \sim N(0, i/N)$, calcoliamo ora che distribuzione segue $W_t - W_s$. Notiamo che $W_t - W_s = X_t + \dots + X_1 -$

$X_s - \dots X_1 = X_t + \dots X_{s+1}$, allora

$$E(W_t - W_s) = E(X_t + \dots X_{s+1}) = 0$$

$$\text{Var}(W_t - W_s) = \text{Var}(X_t + \dots X_{s+1}) = t - s$$

Allora il punto (c) della Definizione 1.12 è verificato e ciò significa che la proprietà degli incrementi indipendenti è rispettata. Verifichiamo anche che il processo è stazionario. Infatti $W_{t+h} - W_t = X_{t+h} + \dots X_{t+1}$ e la media e varianza sono

$$E(W_{t+h} - W_t) = E(W_{t+h}) - E(W_t) = 0$$

$$\text{Var}(W_{t+h} - W_t) = \text{Var}(X_{t+h} \dots X_{t+1}) = h$$

Quindi i valori assunti dal processo non dipendono dalla posizione in cui ci si trova, ma solo dall'ampiezza dell'intervallo considerato: la stazionarietà è valida.

Nella definizione di moto Browniano si era richiesto che $X_0 = 0$, ma il moto Browniano simulato non è definito in 0: la prima realizzazione avviene al tempo $1/N$. Tuttavia, avendo calcolato il moto Browniano come somma cumulata di realizzazioni normali standard indipendenti, possiamo immaginare che la prima realizzazione del processo sia 0. Inoltre all'aumentare di N , la prima realizzazione, che avviene al tempo $1/N$, si avvicina sempre di più a zero.¹⁰

Nell'ultima sezione del capitolo precedente avevamo mostrato che le traiettorie del moto Browniano non sono a variazione finita su alcun intervallo. Verifichiamo se anche per il processo di Wiener simulato vale questo risultato. Per fare ciò consideriamo più simulazioni, aumentando il numero di realizzazioni simulate e diminuendo l'ampiezza dei sottointervalli della partizione. L'algoritmo usato è il seguente.

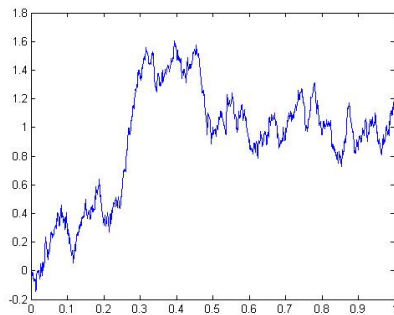
```
somma=0;
M=50;
for i=M:M:N
    somma=somma + abs(W(i)-W(i-(M-1)));
end;
```

La costante M rappresenta il numero di realizzazioni simulate all'interno di ogni sottointervallo della partizione del dominio, mentre *somma* contiene la variazione delle traiettorie. Lanciamo diverse simulazioni e calcoliamo la variazione per valori diversi di N e M . I risultati ottenuti sono riassunti nella seguente tabella:

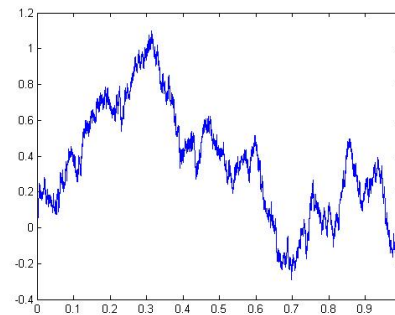
¹⁰I vettori in Matlab non hanno l'entrata indicizzata in 0, per cui risulta poi molto più complicato gestire le simulazioni con un vettore in cui nell'entrata indicizzata a 1, troviamo la realizzazione al tempo 0.

M / N	1000	10000	1 milione	10 milioni
50	3,87	10,16	111,03	353,34
20	7,04	16,47	173,39	549,20
5	10,67	31,93	319,08	1008,70
2	12,21	39,12	399,20	12615,00

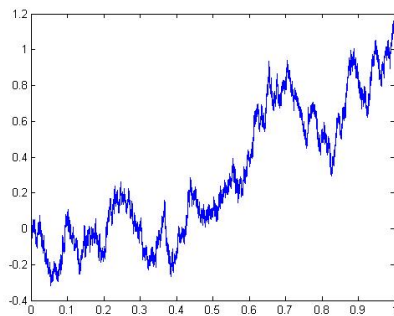
All'aumentare sia di N, sia di M, il valore di *somma* aumenta. Possiamo quindi pensare che, continuando ad aumentare il numero di realizzazioni della simulazione, il valore della variazione tenda ad infinito. Questo conferma il risultato ottenuto analiticamente nel capitolo 2. Mostriamo ora anche gli output delle simulazioni effettuate durante la verifica della non limitatezza della variazione delle traiettorie dei moti Browniani



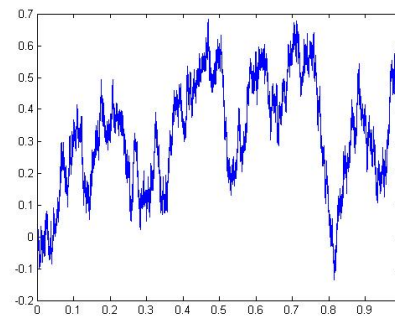
(a) 1000 real.



(b) 10mila real.



(c) 1milione di real.



(d) 10milioni di real.

Figura 2.2: Traiettorie al variare di N

Appendice A

Qualche richiamo di probabilità

Utilizziamo questo primo appendice per richiamare o introdurre qualche concetto della teoria della probabilità e della teoria della misura che risultano utili per comprendere in maniera più profonda i temi sviluppati nella tesi.

Definizione A.1. Una σ -algebra ε su un insieme non vuoto E è una famiglia di parti (sottoinsiemi) di E che contiene E e che sia chiusa per unioni numerabili e passaggio al complementare. La coppia (E, ε) è detta spazio misurabile.

I due esempi "estremi" di σ -algre sono quella banale $\varepsilon = \{\emptyset, E\}$ e quella discreta $\varepsilon = \mathcal{P}(E)$, ossia l'insieme di tutti i possibili sottoinsiemi di E . Un altro esempio è dato dalla σ -algebra boreliana su \mathbb{R} generata dagli intervalli aperti (equivalentemente, dagli intervalli chiusi).¹ Per cui in generale, dato uno spazio misurabile (\mathbb{R}, τ) , dove τ rappresenta la sigma algebra boreliana, si avrà che $\emptyset \in \tau$, $\mathbb{R} \in \tau$ e per $a, b \in \mathbb{R}$, con $a < b$, $]a, b[\in \tau$.

Definizione A.2. Sia $X : E \rightarrow F$ un'applicazione tra due spazi misurabili (E, ε) e (F, φ) . L'applicazione X è detta misurabile se $X^{-1}(\varphi) \subseteq \varepsilon$, cioè se $X^{-1}(B) \in \varepsilon$ per ogni $B \in \varphi$.

Qualunque sia la σ -algebra φ su F , $X^{-1}(\varphi)$ è sempre una σ -algebra su E , detta la σ -algebra generata da X e indicata con $\sigma(X)$: si tratta della più piccola σ -algebra su E che renda X misurabile.

Definizione A.3. Dato uno spazio misurabile (E, ε) , una misura μ è una funzione $\mu : \varepsilon \rightarrow [0, +\infty]$ tale che $\mu(\emptyset) = 0$ e con la proprietà di σ -additività, cioè $\mu(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} \mu(A_n)$ per ogni successione $\{A_n\}_{n \in \mathbb{N}}$ di elementi di ε a due a due disgiunti ($A_n \cap A_m = \emptyset$ per $m \neq n$).

¹È equivalente per la richiesta di chiusura per passaggio al complementare.

Definizione A.4. Una misura P su uno spazio misurabile (Ω, F) tale che $P(\Omega) = 1$ è detta probabilità. La terna (Ω, F, P) è detta spazio di probabilità.

Ricordiamo ora cosa si intende per variabile aleatoria.

Definizione A.5. Sia (Ω, F, P) uno spazio di probabilità e (E, ε) uno spazio misurabile, se la mappa $X : \Omega \rightarrow E$ è misurabile allora è detta variabile aleatoria. Ossia X è una variabile aleatoria se per ogni $A \in \varepsilon$, $\{\omega \in \Omega : X(\omega) \in A\} = X^{-1}(A) = B \in F$. Se E coincide con \mathbb{R} allora è detta variabile aleatoria reale.

Quindi inserendo la misura P di uno spazio probabilizzato (mantenendo le stesse notazioni), succede che

$$P(\{\omega \in \Omega : X(\omega) \in A\}) = P(X^{-1}(A)) = P(B)$$

Appendice B

Variabili di Rademacker

Inseriamo questa appendice per dare un esempio concreto dei concetti di σ -algebra e di filtrazione, così da capirne meglio il significato. Fissiamo uno spazio di probabilità (Ω, F, P) , dove $\Omega = [0, 1)$ e F è la σ -algebra boreliana, ossia la σ -algebra generata da tutti gli intervalli aperti (o chiusi) di $[0, 1)$.

Definiamo:

$$X_1 = \begin{cases} -1, & \text{se } x_1 \in [0, \frac{1}{2}) \\ +1, & \text{se } x_1 \in [\frac{1}{2}, 1) \end{cases}$$
$$X_2 = \begin{cases} -1, & \text{se } x_2 \in [0, \frac{1}{4}) \cup [\frac{1}{2}, \frac{3}{4}) \\ +1, & \text{se } x_2 \in [\frac{1}{4}, \frac{1}{2}) \cup [\frac{3}{4}, 1) \end{cases}$$

con $P(X_1 = 1) = P(X_1 = -1) = P(X_2 = 1) = P(X_2 = -1) = \frac{1}{2}$.

Si vede che X_2 "deriva" da X_1 nel senso che le due v.a. sono uguali a meno dei domini, che passando da X_1 a X_2 in qualche modo si "dimezzano". Calcoliamo ora speranza, varianza e covarianza per queste due funzioni.

$$E(X_1) = P(X_1 = 1) - P(X_1 = -1) = 0 = E(X_2)$$

$$Var(X_1) = V(X_2) = E(X_1^2) = E(X_2^2) = 1$$

$$Cov(X_1, X_2) = E(X_1 X_2)$$

Dobbiamo quindi calcolare la distribuzione di probabilità per la v.a. $X_1 X_2$. Calcoliamo quali valori assume questa nuova variabile e con quale probabilità.

$X_1 X_2$	$X_2 = -1$	$X_2 = +1$	$P(X_1 X_2)$	$X_2 = -1$	$X_2 = +1$
$X_1 = -1$	+1	-1	$X_1 = -1$	1/4	1/4
$X_1 = +1$	-1	+1	$X_1 = +1$	1/4	1/4

Quindi $P(X_1 X_2 = 1) = P(X_1 X_2 = -1) = \frac{1}{2}$.

Allora $E(X_1 X_2) = 0 = Cov(X_1, X_2)$. In generale, si dice che X_1 è ortogonale a X_2 .

Scriviamo ora la generica variabile X_n .

$$X_n = \begin{cases} 1, & \text{se } x_n \in \bigcup_{k=0}^{2^n-1} [\frac{2k}{2^n}, \frac{2k+1}{2^n}] \\ -1, & \text{altrimenti.} \end{cases}$$

Queste variabili X_1, X_2 e la generica X_n sono chiamate variabili di Rademacker.

Consideriamo ora $S_n = X_1 + \dots + X_n = \sum_{i=1}^n X_i$ e calcoliamo la distribuzione di S_2 .

$X_1 + X_2$	$X_2 = -1$	$X_2 = +1$	$P(X_1 + X_2)$	$X_2 = -1$	$X_2 = +1$
$X_1 = -1$	-2	0	$X_1 = -1$	1/4	1/4
$X_1 = +1$	0	+2	$X_1 = +1$	1/4	1/4

Quindi $P(S_2 = 0) = \frac{1}{2}$ e $P(S_2 = -2) = P(S_2 = 2) = \frac{1}{4}$. Possiamo quindi definire S_2 , stando attenti agli intervalli del dominio.

$$S_2 = \begin{cases} +2, & \text{se } s_2 \in [\frac{3}{4}, 1) \\ -2, & \text{se } s_2 \in [0, \frac{1}{4}) \\ 0, & \text{altrimenti} \end{cases}$$

Siamo ora pronti ad elencare le σ -algebre di $S_1 = X - 1$ e S_2 e a trovare dei generici generatori delle σ -algebre.

$$\sigma(S_1) = \left\{ \emptyset, [0, 1), \left[0, \frac{1}{2}\right), \left[\frac{1}{2}, 1\right) \right\}$$

$$\sigma(S_2) = \left\{ \emptyset, [0, 1), \left[0, \frac{1}{4}\right), \left[\frac{1}{4}, 1\right), \left[\frac{3}{4}, 1\right), \left[0, \frac{3}{4}\right), \left[\frac{1}{4}, \frac{3}{4}\right), \left[0, \frac{1}{4}\right) \cup \left[\frac{3}{4}, 1\right) \right\}$$

Due dei generatori delle σ -algebre potrebbero essere:

$$\text{per } S_1 \quad \left\{ \emptyset, \left[0, \frac{1}{2}\right) \right\}$$

$$\text{per } S_2 \quad \left\{ \emptyset, \left[0, \frac{1}{4}\right), \left[\frac{3}{4}, 1\right), \left[\frac{1}{4}, \frac{3}{4}\right) \right\}$$

Proviamo a trovare una filtrazione per il processo S_n e scegliamo $F_1 = \sigma(S_1)$ e $F_2 = \sigma(S_2)$, ossia $F_n = \sigma(S_n)$. Questa però non è una filtrazione, in quanto non è una famiglia crescente di σ -algebre. Ad esempio, l'intervallo $[0, 1/2)$, che è presente in $\sigma(S_1)$, non fa parte di $\sigma(S_2)$.

Proviamo invece a definire $F_1 = \sigma(S_1)$ e $F_2 = \sigma(S_1, S_2)$, ossia $F_n = \sigma(S_1, \dots, S_n)$. F_2 in questo caso corrisponde a $\sigma(S_1) \cup \sigma(S_2)$. Risulta quindi semplice verificare che F_n è in questo caso una filtrazione.