

GENERALITÀ SUI PROCESSI STOCASTICI

I Processi Stocastici

Prendiamo $X(t)$ come un processo stocastico reale, ovvero una variabile casuale reale che evolvono con determinate leggi nel tempo. Da un punto di vista matematico esso è completamente definito se si conoscono tutte le funzioni densità di probabilità congiunta

$$p(x_1, t_1; x_2, t_2; \dots; x_n, t_n)$$

definite in modo che

$$p(x_1, t_1; x_2, t_2; \dots; x_n, t_n) dx_1 dx_2 \dots dx_n$$

rappresenti la probabilità che il processo assuma un valore tra x_k e $x_k + dx_k$ al tempo t_k per ogni k da 1 a n .

Utilizzando il concetto di *probabilità condizionata* possiamo scrivere la funzione di densità di probabilità come:

$$p(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = p(x_1, t_1 | x_2, t_2; \dots; x_n, t_n) \dots p(x_{n-1}, t_{n-1} | x_n, t_n) p(x_n, t_n)$$

Dove con $p(\dots | \dots)$ si è indicato la funzione di densità di probabilità condizionata subordinata all'avverarsi delle condizioni che si trovano a destra della sbarretta verticale.

Valgono inoltre le relazioni fondamentali derivanti dalla definizione di funzione densità di probabilità:

$$\int_{-\infty}^{+\infty} p(x, t) dx = 1$$

In generale sommando su tutti gli eventi mutuamente esclusivi di un tipo in una probabilità congiunta si elimina quella variabile:

$$p(x_1, t_1; x_2, t_2; \dots; x_{n-1}, t_{n-1}) = \int_{-\infty}^{+\infty} p(x_1, t_1; x_2, t_2; \dots; x_{n-1}, t_{n-1}; x_n, t_n) dx_n$$

Ricordiamo ora le definizioni di media, autocorrelazione (R) e autocovarianza (C) di una variabile stocastica.

Dato il valor medio di $x(t)$:

$$\eta(t) = E[x(t)] = \int_{-\infty}^{+\infty} x \cdot p(x, t) dx$$

$$R(t_1, t_2) = E[x(t_1)x(t_2)] = \int_{-\infty}^{+\infty} x_1 \cdot x_2 \cdot p(x_1, t_1; x_2, t_2) dx_1 dx_2$$

$$C(t_1, t_2) = E\{[x(t_1) - \eta(t_1)] \cdot [x(t_2) - \eta(t_2)]\} = \int_{-\infty}^{+\infty} [x(t_1) - \eta(t_1)] \cdot [x(t_2) - \eta(t_2)] \cdot p(x_1, t_1; x_2, t_2) \cdot dx_1 \cdot dx_2$$

Dove con $E[]$ si intende la **media di insieme ovvero la media fatta su un insieme di realizzazioni diverse dello stesso processo stocastico**

Si ha:

$$C(t_1, t_2) = R(t_1, t_2) - \eta(t_1)\eta(t_2)$$

Introduciamo ora il coefficiente di autocorrelazione:

$$\rho_t = \frac{C(t_1, t_2)}{\sqrt{\sigma(t_1)\sigma(t_2)}}$$

dove $\sigma^2(t)$ è la varianza ovvero il particolare valore della funzione di autocovarianza definito come:

$$\sigma^2(t) = C(t, t) = E[(x(t) - \eta(t))^2]$$

Tenendo conto della definizione di probabilità condizionata si ottiene:

$$E[x(t)|x(t_0)] = \int_{-\infty}^{+\infty} x \cdot p(x(t)|x(t_0)) dx$$

Stazionarietà

Un processo stocastico si dice stazionario in *sensu stretto* quando le sue caratteristiche statistiche non cambiano se si sposta l'origine dei tempi, vale cioè la seguente relazione per qualunque ε :

$$p(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = p(x_1, t_1 + \varepsilon; x_2, t_2 + \varepsilon; \dots; x_n, t_n + \varepsilon)$$

se questa relazione vale solo per ogni $i \leq k$ si dice che è stazionario al k -esimo ordine.

Da questa definizione segue che deve essere

$$p(x, t) = p(x, t + \varepsilon) \text{ per ogni } \varepsilon$$

e quindi

$$p(x, t) = p(x).$$

Analogamente $p(x_1, t_1; x_2, t_2) = p(x_1, t_1 + \varepsilon; x_2, t_2 + \varepsilon)$ per qualunque ε implica

$$p(x_1, t_1; x_2, t_2) = p(x_1; x_2, \tau) \text{ con } \tau = t_2 - t_1 \quad \forall t$$

Un processo si dice stazionario in *senso ampio* se:

$$E[x(t)] = \eta = \text{costante}$$

$$R(t_1, t_2) = R(\tau) = E[x(t + \tau)x(t)]$$

$$\text{con } \tau = t_2 - t_1 \quad \forall t$$

cioè R dipende solo dalla differenza tra t_1 e t_2 e non dal tempo assoluto.

La funzione di autocorrelazione $R(\tau)$ quantifica il ricordo che il processo, all'istante $t + \tau$, ha di quello che è successo all'istante t .

Si può dimostrare che la stazionarietà in senso stretto implica la stazionarietà in senso ampio.

Ergodicità

Un processo stazionario si dice **ergodico** se tutte le proprietà statistiche possono essere determinate da un'unica realizzazione del processo.

In altre parole le medie di insieme possono essere sostituite da medie temporali:

$$E[x(t)] = \frac{1}{2T} \cdot \int_{-T}^T x(t) dt = \eta \text{ per } T \rightarrow \infty$$

Processo di Rumore Bianco

Si definiscono processi di rumore bianco i processi stocastici caratterizzati dalle seguenti proprietà:

$$E[x(t)] = 0$$

$$\sigma^2(t) = \text{costante}$$

$$C(\tau) = 0$$

$\forall \tau \neq 0$ essendo $\tau = t - t'$

La funzione di densità di probabilità è Gaussiana con media nulla e varianza costante. Il fatto che la covarianza sia nulla significa che non c'è correlazione per tempi differenti.

In sintesi in un processo di rumore bianco i valori di variabili casuali misurate ad un tempo t sono completamente scorrelati da quelli misurati ad un tempo t' con $t' \neq t$, cioè non vi è memoria del passato.

Processo di Markov

Consideriamo un generico processo stocastico ed esprimiamo la sua funzione di densità di probabilità in funzione della probabilità condizionata

$$p(x_1, t_1 | x_2, t_2; \dots; x_n, t_n)$$

$$t_1 > t_2 > t_3 \dots$$

Se la generica probabilità condizionata può essere espressa solo in funzione di x_2, t_2 allora il processo si dice di Markov.

$$p(x_1, t_1 | x_2, t_2; \dots; x_n, t_n) = p(x_1, t_1 | x_2, t_2)$$

Sostanzialmente in questo tipo di processi la statistica del passato non influenza quella del futuro, fissata la condizione del presente (x_2, t_2).

Applicando la definizione di probabilità condizionata si ha:

$$p(x_1, t_1) = \int_{-\infty}^{+\infty} p(x_1, t_1; x_2, t_2) \cdot dx_2 = \int_{-\infty}^{+\infty} p(x_1, t_1 | x_2, t_2) \cdot p(x_2, t_2) \cdot dx_2$$

che fornisce la probabilità al tempo $t_1 > t_2$, nota la probabilità condizionata $p(x_1, t_1 | x_2, t_2)$.

Per la probabilità condizionata si ha analogamente:

$$p(x_1, t_1 | x_3, t_3) = \int_{-\infty}^{+\infty} p(x_1, t_1; x_2, t_2 | x_3, t_3) \cdot dx_2 = \int_{-\infty}^{+\infty} p(x_1, t_1 | x_2, t_2; x_3, t_3) \cdot p(x_2, t_2 | x_3, t_3) \cdot dx_2$$

applicando infine la definizione di processo di Markov si ha:

$$\int_{-\infty}^{+\infty} p(x_1, t_1 | x_2, t_2; x_3, t_3) \cdot p(x_2, t_2 | x_3, t_3) \cdot dx_2 = \int_{-\infty}^{+\infty} p(x_1, t_1 | x_2, t_2) \cdot p(x_2, t_2 | x_3, t_3) \cdot dx_2$$

e quindi

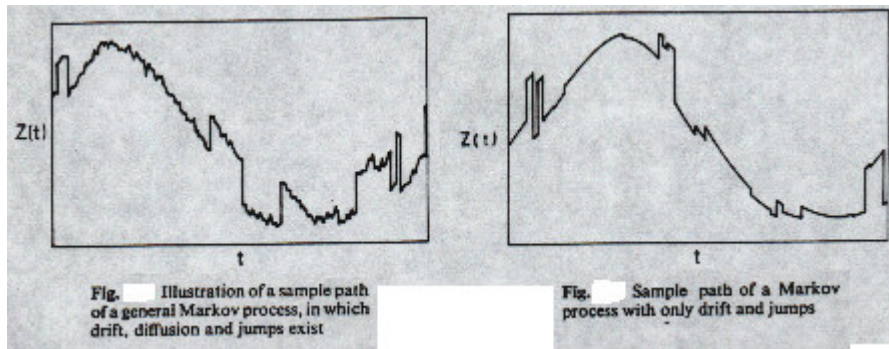
$$p(x_1, t_1 | x_3, t_3) = \int_{-\infty}^{+\infty} p(x_1, t_1 | x_2, t_2) \cdot p(x_2, t_2 | x_3, t_3) \cdot dx_2$$

chiamata equazione di *Chapman-Kolmogorov* che lega le probabilità condizionate l'una all'altra.

Si può derivare la forma differenziale dell'equazione di *Chapman-Kolmogorov*:

$$\begin{aligned} \frac{\partial p(x, t | y, t')}{\partial t} = & -\frac{\partial}{\partial x} \cdot [A(x, t) \cdot p(x, t | y, t')] \\ & + \frac{1}{2} \cdot \frac{\partial^2}{\partial x^2} \cdot [B(x, t) \cdot p(x, t | y, t')] \\ & + \int dz [W(x | z, t) p(z, t | y, t') - W(z | x, t) p(x, t | y, t')] \end{aligned}$$

dove $A(x, t)$ è chiamato termine di drift (deriva), $B(x, t)$ termine di diffusione e $W(x | y, t)$ definisce i processi di salto. A , B e W sono funzioni caratteristiche del particolare problema che si vuole studiare.



Un processo di Markov si dice continuo se per qualunque $\epsilon > 0$ si ha:

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|x-z| > \epsilon} dx p(x, t + \Delta t | z, t) = 0$$

uniformemente in z , t e Δt .

In particolare si può dimostrare che l'equazione di *Chapman-Kolmogorov* definisce un processo continuo se è nullo il termine W . (figura)

Un caso particolare dell'equazione di *Chapman-Kolmogorov*, che definisce un processo markoviano continuo, è l'equazione di *Fokker-Planck*, ottenuta dalla precedente ponendo a zero il termine discontinuo W :

$$\frac{\partial p(x, t | y, t')}{\partial t} = -\frac{\partial}{\partial x} [A(x, t) \cdot p(x, t | y, t')] + \frac{1}{2} \cdot \frac{\partial^2}{\partial x^2} [B(x, t) \cdot p(x, t | y, t')]$$

Usando la relazione:

$$p(x, t) = \int p(x, t; y, t') \cdot dy = \int p(x, t | y, t') \cdot p(y, t') \cdot dy$$

si può mostrare che l'equazione di *Fokker Planck* è valida anche per la generica $p(x, t)$ con la condizione iniziale

$$p(x, t) \Big|_{t=t'} = p(x, t')$$

e quindi possiamo riscrivere questa equazione come

$$\frac{\partial p(x, t)}{\partial t} = -\frac{\partial}{\partial x} [A(x, t) \cdot p(x, t)] + \frac{1}{2} \cdot \frac{\partial^2}{\partial x^2} [B(x, t) \cdot p(x, t)]$$

L'equazione di *Fokker-Planck* può essere utilizzata per la soluzione ad esempio del problema del moto Browniano. Infatti, si può dimostrare che esiste una relazione tra quest'ultima e l'equazione di *Langevin*, equazione stocastica differenziale lagrangiana che descrive questo tipo di moto.

L'equazione di *Langevin* può essere scritta come:

$$dx = a(x, t) \cdot dt + b(x, t) \cdot \epsilon(t) \cdot dt$$

dove $a(x, t)$ e $b(x, t)$ sono funzioni tipiche del particolare problema in esame e $\epsilon(t)$ è un termine casuale rapidamente variabile tale che:

$$\langle \epsilon(t) \cdot \epsilon(t') \rangle = \delta(t - t')$$

$$\langle \mathcal{E}(t) \rangle = 0$$

cioè un processo di rumore bianco.

E' possibile dimostrare che un fenomeno di questo tipo è rappresentabile come un processo *Markoviano* e continuo, quindi descrivibile dal punto di vista statistico da una funzione di densità di probabilità $p(x,t|x_1,t_1)$.

Esistono delle relazioni che legano le funzioni $a(x,t)$, $b(x,t)$ e $A(x,t)$, $B(x,t)$ ottenibili a partire dalla formula di *Ito* riguardante la definizione del differenziale di una funzione $f(x(t))$ dove, ovviamente, $x(t)$ rappresenta un processo stocastico *Markoviano* continuo (Gardiner, 1990).

Queste relazioni sono:

$$a(x,t) = A(x,t)$$

$$b^2(x,t) = B(x,t)$$

Quindi possiamo scrivere l'equazione di *Fokker-Planck* come:

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x} \cdot [a(x,t) \cdot p(x,t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [b^2(x,t) \cdot p(x,t)]$$

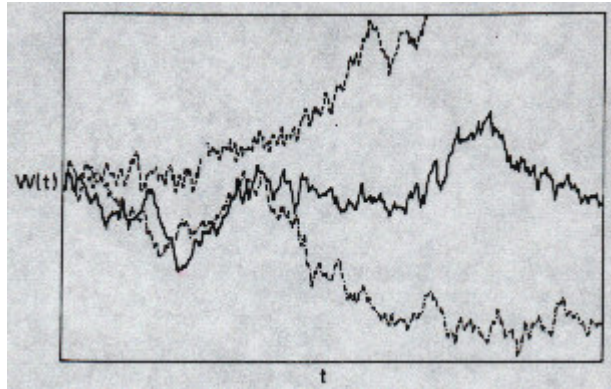
Data la funzione densità di probabilità $p(x,t)$ possiamo usare l'equazione di *Fokker-Planck* per ricavare le quantità $a(x,t)$ e $b(x,t)$ che utilizzate nell'equazione di *Langevin* permettono di descrivere, ad esempio il moto *Browniano* o il moto di una particella in un campo di velocità turbolento.

Processo di Wiener

Una particolare soluzione dell'equazione di Fokker-Planck, nel caso in cui si abbia $A(x,t)=0$ e $B(x,t)=1$ è il processo di Wiener, un processo stocastico con media nulla e varianza proporzionale a t per cui:

$$E[x(t)] = 0$$

$$E[(x(t) - \eta(t))^2] = E[x^2(t)] = \alpha t$$



La distribuzione di probabilità di questo processo è Gaussiana:

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

Per tempi ordinati $t_1 > t_2 > t_3$ vale anche la seguente condizione:

$$E[(x(t_1) - x(t_2)) \cdot (x(t_2) - x(t_3))] = 0$$

che significa che gli incrementi in un processo di Wiener sono scorrelati (indipendenti).

Un processo di Wiener può essere scritto come:

$$x(t) = \int_0^t n(\tau) \cdot d\tau$$

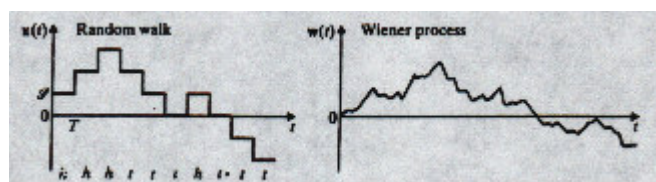
dove $n(\tau)$ è un processo di rumore bianco.

Da ciò segue che:

$$dx(t) = n(t)dt$$

quindi gli incrementi infinitesimi di questo tipo di processo costituiscono un processo di rumore bianco.

Il processo di Wiener può essere considerato come caso limite continuo di un processo discreto detto di **random walk**.



Momenti di una distribuzione di probabilità

La descrizione statistica di una generica variabile s può essere ottenuta attraverso la conoscenza dei suoi momenti o equivalentemente della sua funzione di densità di probabilità $B(s)$. Si ha per il momento di ordine n :

$$\langle s^n \rangle = \overline{s^n} = \int_{-\infty}^{+\infty} s^n B(s) \cdot ds$$

e per il momento centrato di ordine n :

$$\overline{s^n} = \int_{-\infty}^{+\infty} (s - \bar{s})^n B(s) \cdot ds$$

Teoricamente n può andare fino all'infinito, cioè teoricamente possono esistere infiniti momenti della grandezza statistica s . Ai fini pratici ci si ferma ad utilizzare i primi tre o quattro momenti.

- la media o momento del primo ordine, $\langle s \rangle = \bar{s}$

$$\langle s \rangle = \bar{s} = \int_{-\infty}^{+\infty} s B(s) \cdot ds$$

- la varianza o momento centrato del secondo ordine, $\langle s^2 \rangle = \overline{s^2}$ (o σ^2)

$$\langle s^2 \rangle = \overline{s^2} = \int_{-\infty}^{+\infty} (s - \bar{s})^2 B(s) \cdot ds$$

- la skewness (non normalizzata), momento centrato del terzo ordine, $\langle s^3 \rangle = \overline{s^3}$

$$\langle s^3 \rangle = \overline{s^3} = \int_{-\infty}^{+\infty} (s - \bar{s})^3 B(s) \cdot ds$$

- la kurtosis (non normalizzata), momento centrato del quarto ordine, $\langle s^4 \rangle = \overline{s^4}$

$$\langle s^4 \rangle = \overline{s^4} = \int_{-\infty}^{+\infty} (s - \bar{s})^4 B(s) \cdot ds$$

(Normalmente con *skewness* e *kurtosis* si intendono i momenti normalizzati con opportune potenze di σ)

Distribuzione di probabilità (PDF) bi-Gaussiana

La distribuzione e' ottenuta a partire da una combinazione lineare, nella variabile w , di due distribuzioni Gaussiane:

$$P(w, z) = A \cdot P_A(w_A, \sigma_A) + B \cdot P_B(w_B, \sigma_B)$$

Queste due distribuzioni Gaussiane sono caratterizzate dall'aver valore medio non nullo (w_A e w_B rispettivamente), e, in generale, diversi valori di varianza (σ_A , σ_B). I coefficienti A e B tengono conto del 'peso' delle due gaussiane.

Si noti che la distribuzione bi-Gaussiana può essere pensata semplicemente come un'artificio matematico per poter introdurre valori dei momenti superiori al secondo. La corrispondenza fisica con la distribuzione reale e' assicurata unicamente dalla coincidenza dei valori assunti da tali momenti.

Possiamo scrivere questa PDF quindi come:

$$P(w, z) = A \cdot \frac{\exp\left(-\frac{1}{2} \cdot \left(\frac{w - w_A}{\sigma_A}\right)^2\right)}{\sqrt{2 \cdot \pi} \cdot \sigma_A} + B \cdot \frac{\exp\left(-\frac{1}{2} \cdot \left(\frac{w - w_B}{\sigma_B}\right)^2\right)}{\sqrt{2 \cdot \pi} \cdot \sigma_B}$$

Si devono quindi determinare w_A , w_B , A , B , σ_A e σ_B in modo tale che la $P(w, z)$ soddisfi le n equazioni per la definizione degli n momenti della distribuzione:

$$\overline{w^n} = \int_{-\infty}^{+\infty} w^n \cdot P(w, z) \cdot dw = A \int_{-\infty}^{+\infty} w^n \cdot P_A(w, z) \cdot dw + B \int_{-\infty}^{+\infty} w^n \cdot P_B(w, z) \cdot dw$$

dove $\overline{w^n}$ sono i momenti misurati delle PDF reali.

In particolare, assumendo il valor medio nullo, per i primi quattro momenti si ottengono le seguenti relazioni:

$$A + B = 1$$

$$Aw_A + Bw_B = 0$$

$$A(w_A^2 + \sigma_A^2) + B(w_B^2 + \sigma_B^2) = \overline{w^2}$$

$$A(w_A^3 + 3w_A\sigma_A^2) + B(w_B^3 + 3w_B\sigma_B^2) = \overline{w^3}$$

$$A(w_A^4 + 6w_A^2\sigma_A^2 + 3\sigma_A^4) + B(w_B^4 + 6w_B^2\sigma_B^2 + 3\sigma_B^4) = \overline{w^4}$$

Risolvendo questo sistema si determinano i valori di A, B, w_A , w_B , σ_A , σ_B , noti ovviamente i 4 momenti della PDF.

Come si può vedere il sistema è formato da 5 equazioni in cui compaiono 6 incognite. Per essere risolto univocamente occorre fissare una condizione supplementare che ci permetta di legare una delle incognite alle altre o ad una parte delle altre. E' da notare che la conoscenza del quinto momento della distribuzione permetterebbe di ottenere una soluzione esatta senza dover ricorrere a nessuna condizione supplementare. Nella realtà, però, può essere difficile, in molti casi avere stime attendibili del quinto momento.

Per ovviare a questi problemi sono stati proposti diversi schemi di chiusura del sistema tra i quali riportiamo il seguente, valido nel caso in cui si conoscano o si utilizzino solo i primi 3 momenti.

Luhar and Britter (1985), utilizzando solo le prime 4 equazioni del sistema, proposero come condizioni di chiusura le seguenti relazioni:

$$\sigma_A = |w_A| \quad \sigma_B = |w_B|$$

ottenendo:

$$w_B = \frac{\overline{w^3} - \sqrt{(\overline{w^3})^2 + 8 \cdot (\overline{w^2})^3}}{4 \cdot \overline{w^2}}$$

$$w_A = \frac{1}{2} \cdot \frac{\overline{w^3}}{\overline{w^2}} - w_B$$

$$A = \frac{w_B}{w_B - w_A}$$

$$B = 1 - A$$

PDF di Gram-Charlier

Il secondo tipo di PDF si basa su uno sviluppo in serie di derivate della funzione

Gaussiana standardizzata (dove si è effettuata la sostituzione $\frac{x - \mu_1}{\sigma} \rightarrow x$):

$$\alpha(x) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot \exp\left(-\frac{1}{2} \cdot x^2\right)$$

se si indica con $D^r = \frac{d^r}{dx^r}$ abbiamo che

$$(-D)^r \alpha(x) = H_r(x) \cdot \alpha(x)$$

con $H_r(x)$ polinomio di Hermite di grado r .

Scriviamo quindi la PDF come:

$$P(x) = \sum_{j=0}^{\infty} C_j \cdot H_j(x) \cdot \alpha(x)$$

dove moltiplicando per $H_r(x)$ e integrando da $+\infty$ a $-\infty$ si ottiene, tenendo conto delle condizioni di ortogonalità tra polinomi di Hermite:

$$C_r = \frac{1}{r!} \cdot \int_{-\infty}^{+\infty} P(x) \cdot H_r(x) \cdot dx$$

Sostituendo gli espliciti valori di $H_r(x)$ e considerando i momenti della distribuzione riferiti ad un valor medio nullo, i coefficienti C_r sono definiti come:

$$C_0 = 1$$

$$C_1 = 0$$

$$C_2 = \frac{1}{2} \cdot (\mu_2 - 1)$$

$$C_3 = \frac{1}{6} \cdot \mu_3$$

$$C_4 = \frac{1}{24} \cdot (\mu_4 - 6 \cdot \mu_2 + 3)$$

dove con μ_r si sono indicati i vari momenti della distribuzione. Otteniamo infine che la PDF nella forma standardizzata è uguale a:

$$P(x) = \alpha(x) \cdot [C_0 \cdot H_0 + C_1 \cdot H_1 + C_2 \cdot H_2 + C_3 \cdot H_3 + C_4 \cdot H_4 + \dots] =$$

$$= \alpha(x) \cdot [C_0 \cdot H_0 + C_3 \cdot H_3 + C_4 \cdot H_4 + \dots] = \alpha(x) \cdot \left[1 + \frac{1}{6} \cdot \mu_3 \cdot H_3 + \frac{1}{24} \cdot (\mu_4 - 3) \cdot H_4 + \dots \right]$$

dove con la standardizzazione si è sostituito a x il valore $\frac{x}{\sqrt{\mu_2}}$ da cui deriva $C_0=1$,

$C_1=0$, $C_2=0$ (in quanto nella forma standardizzata $\mu_2=\mu_2/\sigma^2=1$). I particolari valori di μ_3 , μ_4 così standardizzati sono chiamati rispettivamente Skewness e Kurtosis.

Questa particolare espressione della PDF è chiamata espansione in serie di Gram Charlier tipo A.

In particolare, possiamo riscriverla, considerando solo momenti fino al terzo ordine, come:

$$P(x, z) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} (1 + C_3 H_3)$$

dove $x = \frac{w}{\sqrt{w^2}}$, e:

$$H_3 = x^3 - 3x$$

$$C_3 = \frac{\mu_3}{6} = \frac{1}{6} \cdot \frac{\overline{w^3}}{(\overline{w^2})^{\frac{3}{2}}}$$