

Variabili casuali gaussiane

Ancora sulle v.c. gaussiane

$$X \sim N(\mu, \sigma^2)$$

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\int_{-\infty}^{\mu} f(x) dx = \int_{\mu}^{+\infty} f(x) dx = \frac{1}{2}$$

$$\phi(t) = e^{t\mu + \frac{t^2}{2}\sigma^2} \Rightarrow \begin{aligned} E[X] &= \mu \\ \text{Var}(X) &= \sigma^2 \end{aligned}$$

$$\mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+$$

$$F(a) = \int_{-\infty}^a f(x) dx = \int_{-\infty}^a \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Non c'è una primitiva analitica di questo integrale, sicuramente converge, ma non riusciamo ad esprimere con combinazioni semplici di funzioni ad esprimere il valore della funzione di ripartizione. Sigma e mu possono assumere infiniti valori diversi e questo crea dei problemi nella determinazione della funzione di ripartizione. Per questo motivo, attraverso una proprietà, si passa da una variabile casuale gaussiana con parametri mu e sigma quadro a una variabile casuale gaussiana con parametri 0 e 1, la normale standard.

PROPRIETÀ

$$1) \text{ Se } X \sim N(\mu, \sigma^2) \text{ e } Y = \alpha X + \beta \text{ con } \begin{aligned} &\beta \in \mathbb{R} \\ &\alpha \in \mathbb{R} \setminus \{0\} \end{aligned}$$

$$\Rightarrow Y \sim N(\alpha\mu + \beta, \alpha^2\sigma^2)$$

$$\begin{aligned}
 \boxed{\text{Dim}} \quad \phi_X(t) &= e^{t\mu + \frac{t^2}{2}\sigma^2} \\
 \phi_Y(t) &= E[e^{t(\alpha X + \beta)}] = E[e^{t\alpha X} e^{t\beta}] = \\
 &= e^{t\beta} E[e^{(t\alpha)X}] = e^{t\beta} \phi_X(t\alpha) = \\
 &= e^{t\beta} \left(e^{t\alpha\mu + \frac{t^2\alpha^2}{2}\sigma^2} \right) = e^{t(\alpha\mu + \beta) + \frac{t^2}{2}(\alpha^2\sigma^2)}
 \end{aligned}$$

$$Y \stackrel{||}{\sim} (\alpha\mu + \beta, \alpha^2\sigma^2)$$

Questa proprietà è valida sempre, si può impiegare in qualunque caso si voglia applicare una trasformazione. Ma tra tutte ci concentriamo su una trasformazione particolare.

APPLICAZIONE DELLA PROPRIETÀ

$$X \sim N(\mu, \sigma^2)$$

Considero Z.

$$Z = \frac{1}{\sigma} X - \frac{\mu}{\sigma}$$

È una trasformazione lineare.

$$\frac{1}{\sigma} = \alpha$$

$$-\frac{\mu}{\sigma} = \beta$$

\Downarrow

$$Z \sim N\left(\frac{1}{\sigma}\mu - \frac{\mu}{\sigma}, \frac{1}{\sigma^2}\sigma^2\right)$$

In poche parole, stiamo ottenendo una gaussiana di parametri 0 e 1.

$$= N(0, 1) \quad \text{NORMALE STANDARD}$$

Passo da X gaussiana di parametri qualunque a Z, che ha parametri molto particolari, specifici, per cui si chiama normale standard, si potrebbe chiamare normale di riferimento, gaussiana di riferimento presentata nelle tavole.

Viene presa una gaussiana qualunque, si applica la trasformazione:

$$\frac{1}{\sigma} = \alpha$$

$$-\frac{\mu}{\sigma} = \beta$$

E passo alla normale standard.

A cosa serve? Se voglio calcolare la funzione di ripartizione della X lo posso fare con l'integrale in linea di principio, ma in realtà con carta e penna non si può calcolare. Quindi si applica l'idea della trasformazione lineare.

$$F_X(b) = P(X \leq b) = P(X - \mu \leq b - \mu)$$

Divido per sigma.

$$\sigma = \sqrt{\sigma^2} > 0$$

$$P\left(\frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma}\right) =$$

$$= P(Z \leq \frac{b - \mu}{\sigma}) = F_Z\left(\frac{b - \mu}{\sigma}\right)$$

A questo punto intervengono le tavole.

Normal distribution

z	$f(z)$	$F(z)$	$1 - F(z)$	z	$f(z)$	$F(z)$	$1 - F(z)$
-4.00	0.0001	0.0000	1.0000	-3.50	0.0009	0.0002	0.9998
-3.99	0.0001	0.0000	1.0000	-3.49	0.0009	0.0002	0.9998
-3.98	0.0001	0.0000	1.0000	-3.48	0.0009	0.0003	0.9998
-3.97	0.0001	0.0000	1.0000	-3.47	0.0010	0.0003	0.9997
-3.96	0.0002	0.0000	1.0000	-3.46	0.0010	0.0003	0.9997
-3.95	0.0002	0.0000	1.0000	-3.45	0.0010	0.0003	0.9997
-3.94	0.0002	0.0000	1.0000	-3.44	0.0011	0.0003	0.9997
-3.93	0.0002	0.0000	1.0000	-3.43	0.0011	0.0003	0.9997
-3.92	0.0002	0.0000	1.0000	-3.42	0.0011	0.0003	0.9997
-3.91	0.0002	0.0001	1.0000	-3.41	0.0012	0.0003	0.9997
-3.90	0.0002	0.0001	1.0000	-3.40	0.0012	0.0003	0.9997
-3.89	0.0002	0.0001	1.0000	-3.39	0.0013	0.0003	0.9997
-3.88	0.0002	0.0001	1.0000	-3.38	0.0013	0.0004	0.9996
-3.87	0.0002	0.0001	1.0000	-3.37	0.0014	0.0004	0.9996
-3.86	0.0002	0.0001	0.9999	-3.36	0.0014	0.0004	0.9996

La prima colonna z è la colonna dell'argomento della funzione di ripartizione, la terza è la funzione di ripartizione. La seconda è la funzione di ripartizione. La quarta colonna è il complementare (z sia maggiore di quel valore). Stesso ragionamento nelle seconde 4 colonne.

z è approssimata di 2 cifre e va da -4 a 4 nella tavola.

$$P\left(Z \leq \frac{b-\mu}{\sigma}\right) = F_Z\left(\frac{b-\mu}{\sigma}\right)$$

⇓

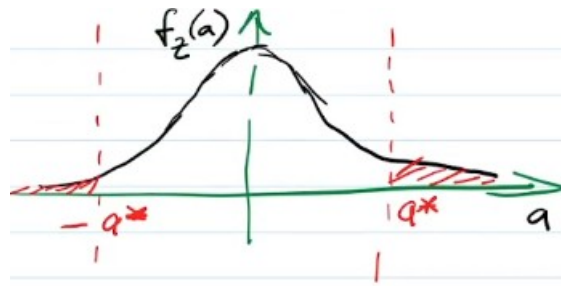
TAVOLE PER LA NORMALE STANDARD

ALCUNE PROPRIETÀ USATE A VOLTE NELLE TAVOLE DELLA NORMALE STANDARD

Simmetrica rispetto all'asse delle y .



Essendo simmetrica vuol dire che se io considero a^* e $-a^*$:



$$P(Z > a^*) = P(Z < -a^*)$$

$$\Downarrow$$

$$1 - P(Z \leq a^*) = F_Z(-a^*)$$

$$\Downarrow$$

$$1 - F(a^*) = F_Z(-a^*)$$

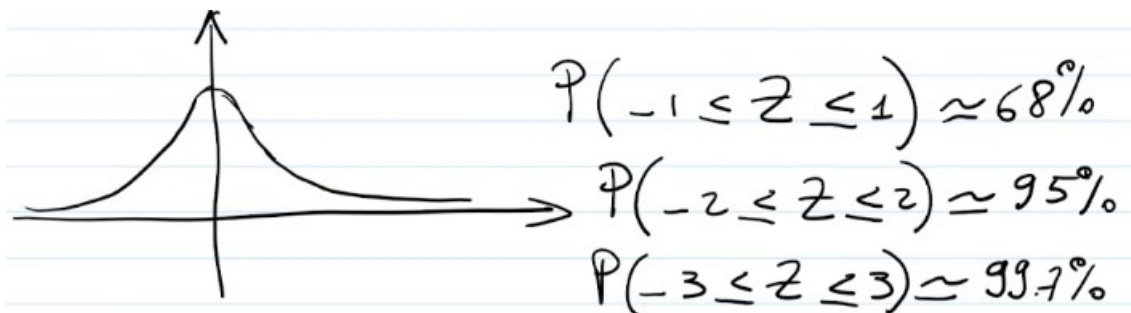
Per risparmiare carta, alcune tavole mettono a disposizione solo gli argomenti positivi e tramite questa formula si possono calcolare le funzioni di ripartizione per argomenti negativi.

Perché le tavole riportate vanno da -4 a 4? Perché sono così vicine a 0 le code che l'area sottesa tra la curva e l'asse è 0.

$$P(Z > 4) \simeq 0$$

per via delle "code" sempre più vicine a 0

A questo proposito, si osserva anche questa proprietà particolare:



Se ci muoviamo oltre la probabilità -3, 3 o addirittura -4, 4 parliamo di una probabilità veramente trascurabile.

Da qui deriva la regola del 3 sigma, oltre il 3 sigma i dati vengono scartati. Questa regola deriva dalla proprietà della normale standard, oltre 3 o -3 ha una probabilità di osservazione molto piccola.

Se chiediamo che Z sia compreso tra -3 e 3, stiamo dicendo che $X - \mu$ deve essere compreso tra -3 sigma e 3 sigma. Per cui quando si fanno degli esperimenti, i dati distanti dalla media aritmetica più di 3 sigma, vengono scartati perché è più probabile che sia stato fatto qualche errore nella sperimentazione.

Questa appendice nasce dalla proprietà generale che la trasformazione lineare di una gaussiana è sempre una gaussiana.

2) PROPRIETÀ DELLA GAUSSIANA: RIPRODUCIBILITÀ

$$X_1 \sim N(\mu_1, \sigma_1^2) \quad X_2 \sim N(\mu_2, \sigma_2^2)$$

INDEPENDENT

$$Y = X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

Dimostriamola con la funzione generatrice dei momenti.

DIM

$$\phi_{X_1}(t) = e^{t\mu_1 + \frac{t^2}{2}\sigma_1^2}$$

$$\phi_{X_2}(t) = e^{t\mu_2 + \frac{t^2}{2}\sigma_2^2}$$

$$\phi_Y(t) = \phi_{X_1+X_2}(t) \stackrel{\text{INDIPENDENZA}}{=} \phi_{X_1}(t) \phi_{X_2}(t) =$$

$$= e^{t\mu_1 + \frac{t^2}{2}\sigma_1^2} e^{t\mu_2 + \frac{t^2}{2}\sigma_2^2} = e^{t(\mu_1 + \mu_2) + \frac{t^2}{2}(\sigma_1^2 + \sigma_2^2)}$$

Abbiamo imparato che questa ha sempre la struttura di una funzione generatrice di una gaussiana.

$$\Downarrow$$

$$Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

N.B. Questa proprietà si può estendere
alle somme di n variabili casuali
gaussiane indipendenti

es. 1

Precipitazioni annuali a LA hanno distribuzione
gaussiana con $\mu = 12.08$ pollici, $\sigma = 3.1$ pollici
Si suppone l'indipendenza delle precipitazioni
tra anni diversi

$$\begin{aligned} X_1 &= \text{'precipitazioni anno 1'} \sim N(12.08, (3.1)^2) \\ X_2 &= \text{'precipitazioni anno 2'} \sim N(-, -) \end{aligned} \left. \vphantom{\begin{aligned} X_1 &= \text{'precipitazioni anno 1'} \sim N(12.08, (3.1)^2) \\ X_2 &= \text{'precipitazioni anno 2'} \sim N(-, -) \end{aligned}} \right\} \text{INDIP.}$$

$$P(X_1 + X_2 > 25)$$

$$\begin{aligned} Y = X_1 + X_2 &\sim N(12.08 + 12.08, (3.1)^2 + (3.1)^2) = \\ &\quad \text{RIPRODUCIBILITÀ} \\ &= N(24.16, 2(3.1)^2) \end{aligned}$$

$$P(Y > 25) = 1 - P(Y \leq 25)$$

Facciamo la standardizzazione.

$$= 1 - P\left(\frac{Y - 24.16}{\sqrt{2 \cdot (3.1)^2}} \leq \frac{25 - 24.16}{\sqrt{2 \cdot (3.1)^2}}\right)$$

$$= 1 - P\left(Z \leq \frac{25 - 24.16}{\sqrt{2 \cdot (3.1)^2}}\right)$$

$$= 1 - F_Z(0.1916)$$

Prendo la tavola.

0.19	0.3918	0.5754	0.4247
------	--------	--------	--------

Possiamo già usare il 4° argomento.

$$= 1 - F_2(0.1916) = 0.4247$$

$$2) P(X_1 > X_2 + 3)$$

Sapevamo che sommando due gaussiane si ottiene una gaussiana. Qui possiamo immaginare di portare X_2 dall'altra parte del maggiore e dobbiamo ricordarci della proprietà della trasformazione lineare della gaussiana.

$$= P(X_1 - X_2 > 3)$$

$$X_2, X_1 \sim N(12.08, (3.1)^2)$$

$$-X_2 \text{ è ancora gaussiana } Y = \alpha X_2 + \beta \\ \text{con } \beta = 0, \alpha = -1$$

$$-X_2 \sim N(-12.08, (-1)^2 (3.1)^2)$$

$$= N(-12.08, (3.1)^2)$$

\Downarrow

$$S = X_1 - X_2 \sim N(12.08 - 12.08, (3.1)^2 + (3.1)^2)$$

Fare attenzione alle varianze, sono sempre positive (mai nulle).

$$= N(0, 2 \cdot (3.1)^2)$$

Rispetto alla domanda iniziale:

$$\begin{aligned} P(S > 3) &= P\left(\frac{S - 0}{\sqrt{2 \cdot (3.1)^2}} > \frac{3 - 0}{\sqrt{2 \cdot (3.1)^2}}\right) = \\ &= P(Z > 0.68) = 1 - P(Z \leq 0.68) = \\ &= 1 - F_Z(0.68) \end{aligned}$$

+47

Di nuovo, vado di tavole.

	0.68		0.3166		0.7518		0.2482	
--	------	--	--------	--	--------	--	--------	--

Di nuovo, il 4° argomento.

$$= 1 - F_Z(0.68) \simeq 0.2482 \simeq 25\%$$

Processi stocastici di Poisson

Una famiglia di variabili casuali che cambiano al variare di un parametro, spesso è il tempo perché frequentemente si usano delle variabili casuali per descrivere dei fenomeni al variare del tempo.

PROCESSO STOCASTICO = FAMIGLIA DI VARIABILI CASUALI AL VARIARE DI)
DIPENDENTI DA
UN PARAMETRO (REALE)

N.B. SPESSO IL PARAMETRO GIOCA IL RUOLO DEL TEMPO

Tra tutte le famiglie di variabili casuali, c'è un tipo particolare di famiglie, chiamato "processo stocastico di Poisson" che soddisfa alcune condizioni particolari.

PROCESSO STOCASTICO DI POISSON DI COSTANTE λ ($\lambda \in \mathbb{R}^+$) = processo stocastico tale che

$$N(t) = \text{n° eventi in }]0, t] \quad (t = \text{tempo})$$

con le seguenti 5 proprietà

1) $N(0) = 0$

2) Il n° di eventi di due intervalli di tempo disgiunti è indipendente

INDIPENDENZA
DEGLI INCREMENTI



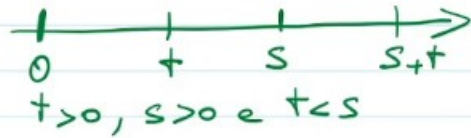
n° eventi in $]0, t]$ è indipendente dal n° di eventi in $]s, w]$

Rispetto alla definizione, il numero di eventi che avvengono tra s e w come si calcolano?

$$N(t) = \text{n° eventi in }]0, t] \\ N(w) - N(s) = \text{n° eventi in }]s, w]$$

3) Il n° di eventi dipende dall'ampiezza dell'intervallo, ma non dalla sua posizione sull'asse reale

STAZIONARIETÀ
DEGLI INCREMENTI



n° eventi in $]0, t] = N(t)$
n° eventi in $]s, s+t] = N(t)$

Perché l'intervallo è di lunghezza t.

Le ultime 2 proprietà sono più matematiche, c'è un limite.

$$4) \lim_{h \rightarrow 0} \frac{P(N(h)=1)}{h} = \lambda \quad \Rightarrow \quad P(N(h)=1) \approx \lambda h \text{ se } h \ll 1$$

La probabilità che il numero di eventi in un intervallino molto piccolo sia 1, diviso h è lambda.

Lambda caratterizza un processo stocastico rispetto ad un altro, perché cambiando lambda cambia la probabilità di osservare un evento in un intervallo molto piccolo. Questo evento avrà una probabilità di verificarsi piccola proporzionale all'ampiezza dell'intervallo (più è piccolo, più è piccola la possibilità di verificarsi).

$$5) \lim_{h \rightarrow 0} \frac{P(N(h) \geq 2)}{h} = 0 \quad \Rightarrow \quad P(N(h) \geq 2) \approx 0 \text{ se } h \ll 1$$

In un piccolo intervallo ci sarà una piccola probabilità di osservare 1 evento, ma 0 probabilità di osservare più di un evento.

Queste proprietà a cosa portano?

$N(t)$ come si distribuisce?

La risposta è nel nome, si distribuirà secondo una legge poissoniana.



$N(t) = \text{n° di eventi in }]0, t]$

Immaginiamo di dividere $]0, t]$ in tanti piccoli sottointervalli tutti uguali.



Divido $[0, t]$ in m intervallini di lunghezza $\frac{t}{m}$

$$\text{Se } m \gg 1 \quad P(N(\frac{t}{m})=1) \underset{\downarrow}{\approx} \lambda \frac{t}{m}$$

$$P(N(\frac{t}{m}) \geq 2) \underset{\downarrow}{\approx} 0$$

$$P(N(\frac{t}{m})=0) = 1 - \lambda \frac{t}{m}$$

Se io so che al variare dell'intervallo nell'asse reale si notano differenze, allora non rientra in Poisson, poi posso decidere che se queste differenze non sono eclatanti, non me ne occupo e continuo a usare Poisson.

Nei processi stocastici di Poisson rientrano per esempio il numero di studenti che accedono all'aula virtuale per il ricevimento. Però se si fa un ricevimento prima dell'esame, si ha sicuramente una variazione rispetto al resto dell'anno. Il numero fluttua il resto dell'anno rimanendo però nello stesso ordine di grandezza. Invece nel caso particolare il numero è molto più alto, quindi non rientra nella famiglia dei ricevimenti, volendo ha una costante lambda diversa.

Oppure, alla cassa del supermercato alla Vigilia di Natale c'è un numero molto più alto di clienti che in un qualunque giorno dell'anno. Queste fluttuazioni fanno sì che almeno queste situazioni eccezionali vadano escluse dal modello che prevede la stazionarietà degli incrementi.

Anche l'indipendenza degli incrementi non è scontata, è molto forte, perché può capitare che se un evento accade una volta non accada più.

Sotto le ipotesi prima fatte:

$$P(N(t)=K) \quad \text{se } K \in \mathbb{N}?$$

La risposta sembra non esserci, ma facendo riferimento a tutti i sottointervalli, trovare K eventi, ad esempio 3 eventi, significa che negli M sottointervalli, in 3 è capitato 1 evento, negli altri $M - 3$ 0 eventi.

Ogni sottointervallo è come un piccolo esperimento (c'è o non c'è l'evento?). Usando i sottointervalli come descriviamo $N(t)$? Ci devono essere K esperimenti che danno 1 evento e $M - K$ esperimenti che non danno alcun evento. È una somma di bernoulliane: in ogni sottointervallo c'è una bernoulliana che indica se l'evento si è verificato. Quindi, la somma di tante bernoulliane danno una binomiale.

$$N(t) \sim B(m, \lambda \frac{t}{m})$$

Sto sfruttando implicitamente tutte le ipotesi, sfruttando l'indipendenza.

perché ogni intervallo di lunghezza Δt presenta o no un evento INDIPENDENTEMENTE DAGLI ALTRI INTERVALLI (2) e GRAZIE A 3) OGNI INTERVALLINO HA PROB. $\lambda \Delta t$ di presentare un evento.

È una binomiale perché gli eventi si verificano in maniera identica e indipendente.