

Corso di Statistica

Cenni di calcolo delle probabilità

Variabili aleatorie

Domenico De Stefano

a.a. 2023/2024

Indice

- 1 Concetti base
- 2 Variabili aleatorie discrete
- 3 Media e varianza
- 4 Variabili aleatorie continue
- 5 La distribuzione normale
- 6 Normale come approssimazione
- 7 Riepilogo

Variabile aleatoria

Definizione di variabile aleatoria

Una v.a. è un numero ben determinato ma non noto per mancanza d'informazioni

Esempi, già incontrati e non

- Il risultato del lancio di un dado.
- Il numero uscito a una mano di *roulette*.
- La somma dei punti di due dadi.
- Il numero di teste su 3 lanci di moneta.
- Il ritardo con cui arriva un treno.
- Il peso di un neonato.

Variabile aleatoria

Definizione di variabile aleatoria

Una v.a. è un numero ben determinato ma non noto per mancanza d'informazioni

Esempi, già incontrati e non

- Il risultato del lancio di un dado.
- Il numero uscito a una mano di *roulette*.
- La somma dei punti di due dadi.
- Il numero di teste su 3 lanci di moneta.
- Il ritardo con cui arriva un treno.
- Il peso di un neonato.

Un'importante distinzione

DISCRETO

DISCRETO

DISCRETO

DISCRETO

CONTINUO

CONTINUO

Variabile aleatoria

Definizione di variabile aleatoria

Una v.a. è un numero ben determinato ma non noto per mancanza d'informazioni

Tratteremo separatamente i due casi:

- v.a. discreta, cioè che assume valori interi.
- v.a. continua, cioè che assume valori reali.

Indice

- 1 Concetti base
- 2 Variabili aleatorie discrete**
- 3 Media e varianza
- 4 Variabili aleatorie continue
- 5 La distribuzione normale
- 6 Normale come approssimazione
- 7 Riepilogo

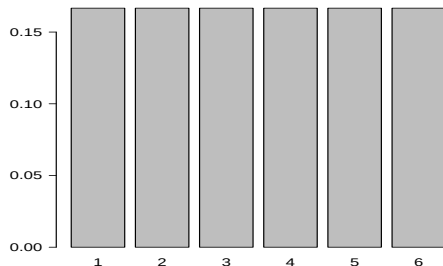
Distribuzione di probabilità per una v.a. discreta

Distribuzione di probabilità

La distribuzione di probabilità di una v.a. discreta non è altro che l'insieme dei possibili valori e delle probabilità di ciascuno.

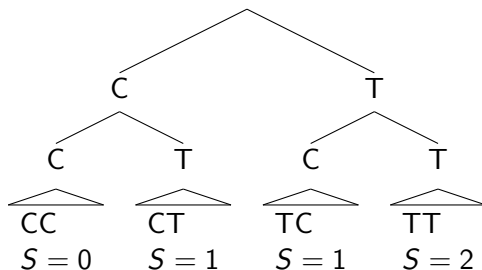
Ad esempio per il dado

Esito	Pr.
1	$1/6$
2	$1/6$
3	$1/6$
4	$1/6$
5	$1/6$
6	$1/6$



Numero di teste su più lanci

- Consideriamo due lanci di una moneta, i cui esiti possibili sono T (testa) o C (croce)
- Definiamo S = numero di teste su due lanci.
- S è una v.a. discreta con valori possibili 0, 1, 2.
- Otteniamo la distribuzione di probabilità di S considerando i possibili esiti

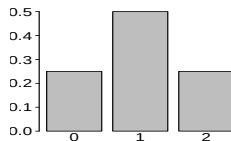


Numero di teste su più lanci

- Consideriamo due lanci di una moneta, i cui esiti possibili sono T (testa) o C (croce)
- Definiamo S = numero di teste su due lanci.
- S è una v.a. discreta con valori possibili 0, 1, 2.
- Otteniamo la distribuzione di probabilità di S considerando i possibili esiti

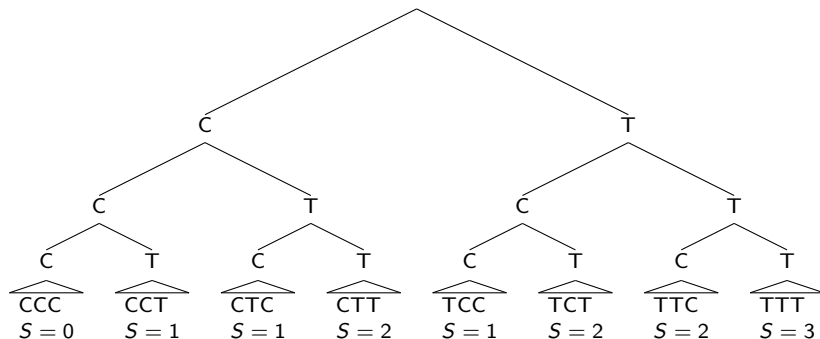
Esiti	Pr	S
CC	1/4	0
CT	1/4	1
TC	1/4	1
TT	1/4	2

S	Pr
0	$P(CC) = 1/4$
1	$P(CT \text{ o } TC) = 1/2$
2	$P(TT) = 1/4$



Numero di teste su tre lanci

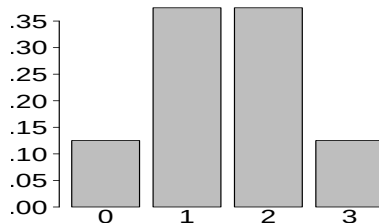
- Con tre lanci, S ha valori 0, 1, 2, 3.



Numero di teste su tre lanci

- Con tre lanci, S ha valori 0, 1, 2, 3.

Esiti	Pr	S	S	Pr
CCC	1/8	0	0	1/8
CCT	1/8	1	1	3/8
CTC	1/8	1		
TCC	1/8	1		
CTT	1/8	2	2	3/8
TTC	1/8	2		
TCT	1/8	2		
TTT	1/8	3	3	1/8



Numero di teste su n lanci

- S ha valori possibili $0, 1, 2, \dots, n$.
- Gli esiti possibili, del tipo

$TCTTCTT\dots TC$

sono 2^n .

- Ciascun esito ha probabilità $\left(\frac{1}{2}\right)^n$
- Per trovare $P(S = s)$ basta contare quanti ce ne sono con s teste, si trova che sono

$$\binom{n}{s} = \frac{n!}{s!(n-s)!}$$

- Si ha allora

$$P(S = s) = \binom{n}{s} \left(\frac{1}{2}\right)^n.$$

Parentesi: il coefficiente binomiale $\binom{n}{s}$

- Con questo simbolo si indica il numero di possibili sottoinsiemi di s elementi da un insieme di n .

Parentesi: il coefficiente binomiale $\binom{n}{s}$

- Con questo simbolo si indica il numero di possibili sottoinsiemi di s elementi da un insieme di n .
- Ad esempio se l'insieme di $n = 5$ elementi è $\{a, b, c, d, e\}$ ha senso chiedersi quanti sono i sottoinsiemi di 0, 1, 2, 3, 4, 5 elementi, li si elenca nel seguito

0	1	2	3	4	5
$\emptyset$	a	ab	abc	$abcd$	$abcde$
	b	ac	abd	$abce$	
	c	ad	abe	$abde$	
	d	ae	acd	$acde$	
	e	bc	ace	$bcde$	
		bd	ade		
		be	bcd		
		cd	bce		
		ce	bde		
		de	cde		
$\binom{5}{0} = 1$	$\binom{5}{1} = 5$	$\binom{5}{2} = 10$	$\binom{5}{3} = 10$	$\binom{5}{4} = 5$	$\binom{5}{5} = 1$

Parentesi: il coefficiente binomiale $\binom{n}{s}$

- Con questo simbolo si indica il numero di possibili sottoinsiemi di s elementi da un insieme di n .
- In generale è

$$\binom{n}{s} = \frac{n!}{s!(n-s)!}$$

dove

$$n! = n(n-1)(n-2) \cdot \dots \cdot 3 \cdot 2 \cdot 1$$

- Alcuni esempi
 - Il numero di coppie che si possono formare in un gruppo di 15 individui è $\binom{15}{2} = 105$.
 - Il numero di cinquine che possono essere estratte al lotto (su 90 numeri) è $\binom{90}{5} = 43\,949\,268$

Cosa succede se testa e croce non hanno la stessa probabilità?

- Sia $P(T) = p$, quindi $P(C) = 1 - p$
- Gli esiti possibili

$TCTTCTT...TC$

hanno probabilità diverse, pari a

$$p^{\#T}(1-p)^{n-\#T}$$

dove $\#T$ è il numero di teste.

- Si ha allora la **distribuzione binomiale** con dimensione n e parametro p

$$P(S = s) = \binom{n}{s} p^s (1-p)^{n-s}.$$

Esempio binomiale I

Alle elezioni politiche del 24 e 25 febbraio 2013 sono andati a votare, per la Camera dei Deputati,

35 271 541

elettori, su un totale di

46 905 154

aventi diritto (da ambo i numeri sono esclusi i cittadini residenti all'estero). Se 'pesco' un residente maggiorenne a caso con diritto di voto, la probabilità che sia andato a votare è

$$p = 0.752$$

(valore pari all'affluenza alle urne).

Esempio binomiale II

Se ripeto la 'pesca' 10 volte, sia V il numero di persone, tra i 10, che sono andati a votare, qual è la probabilità che sia $V = 10$ (tutti e 10 sono andati a votare)?

$$P(V = 10) = \binom{10}{10} p^{10} (1 - p)^{10-10} = p^{10} = 0.0578$$

Se ripeto la 'pesca' 10 volte, sia V il numero di persone, tra i 10, che sono andati a votare, qual è la probabilità che sia $V = 5$ (5 dei 10 sono andati a votare)?

$$P(V = 5) = \binom{10}{5} p^5 (1 - p)^{10-5} = 252 p^5 (1 - p)^{10-5} = 0.0569$$

Esempio binomiale III

Qual è la probabilità che almeno 7 siano andati a votare?

$$\begin{aligned}P(V \geq 7) &= P(V = 7) + P(V = 8) + P(V = 9) + P(V = 10) \\&= \binom{10}{7} p^7 (1-p)^3 + \binom{10}{8} p^8 (1-p)^2 + \\&\quad + \binom{10}{9} p^9 (1-p) + p^{10} \\&= 0.7805\end{aligned}$$

Indice

- 1 Concetti base
- 2 Variabili aleatorie discrete
- 3 Media e varianza**
- 4 Variabili aleatorie continue
- 5 La distribuzione normale
- 6 Normale come approssimazione
- 7 Riepilogo

Media e varianza di una v.a.

Abbiamo già notato che la distribuzione di probabilità è analoga a una distribuzione di frequenze.

Sempre in analogia, si definiscono la media e la varianza.

Valori possibili di X	Probabilità
x_1	p_1
x_2	p_2
$\vdots$	$\vdots$
x_n	p_n

$$E(X) = \sum_{i=1}^n x_i p_i$$

$$E(X^2) = \sum_{i=1}^n x_i^2 p_i$$

$$V(X) = \sum_{i=1}^n (x_i - E(X))^2 p_i$$

Proprietà di media e varianza

Valgono regole analoghe a quelle viste per la media e la varianza campionarie

$$E(aX + b) = aE(X) + b$$

$$V(X) = E(X^2) - [E(X)]^2$$

$$V(aX + b) = a^2 V(X)$$

Media e varianza della binomiale

Nel caso della binomiale di dimensione n e probabilità p

$$P(S = s) = \binom{n}{s} p^s (1 - p)^{n-s}.$$

si ha

$$E(S) = np$$

$$V(S) = np(1 - p)$$

S è il numero di successi, calcoliamo la percentuale di successi, $X = S/n$, applicando le regole illustrate sopra troviamo che

$$E(X) = p$$

$$V(X) = p(1 - p)/n$$

Altre distribuzioni discrete

Naturalmente ci sono molte altre distribuzioni di probabilità discrete

- Poisson
- Geometrica
- Ipergeometrica
- Binomiale negativa

Non ci interessa entrare nei dettagli per il prosieguo.

Indice

- 1 Concetti base
- 2 Variabili aleatorie discrete
- 3 Media e varianza
- 4 Variabili aleatorie continue**
- 5 La distribuzione normale
- 6 Normale come approssimazione
- 7 Riepilogo

Esempi di v.a. continue

Una variabile aleatoria continua è una v.a. i cui valori possibili sono numeri reali (cioè, non solo numeri interi come negli esempi precedenti), ad esempio

- il ritardo di un treno;
- la temperatura di domani alle 12;
- il peso di un neonato;
- l'età di un individuo preso a caso nella popolazione italiana

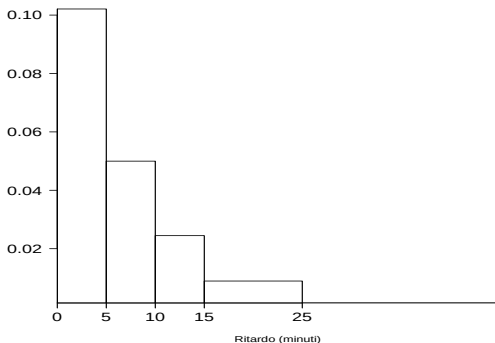
Perché le trattiamo a parte?

Principalmente perché è diverso il modo di esprimere la loro distribuzione di probabilità.

Distribuzione di probabilità di una v.a. continua

- Ricordiamo le lezioni di statistica descrittiva.
- Per variabili continue la distribuzione di frequenza veniva rappresentata per intervalli (distribuzioni in classi).
- Ad esempio con riferimento ai ritardi di un treno potremmo avere

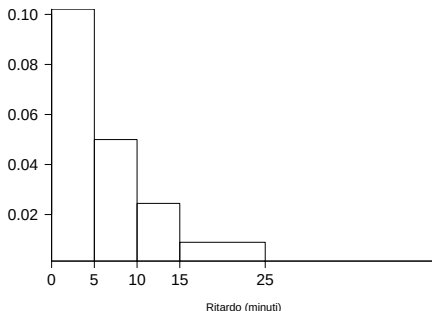
classe	prob
$(0,5]$	0.51
$(5,10]$	0.25
$(10,15]$	0.12
$(15,25]$	0.09
$(25,\text{Inf}]$	0.03



Da istogramma a densità

- Abbiamo infinite osservazioni, possiamo affinare l'istogramma

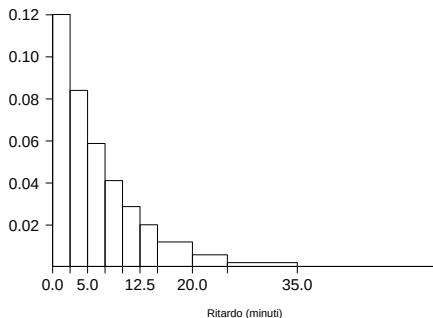
classe	prob
$(0,5]$	0.51
$(5,10]$	0.25
$(10,15]$	0.12
$(15,25]$	0.09
$(25,\text{Inf}]$	0.03



Da istogramma a densità

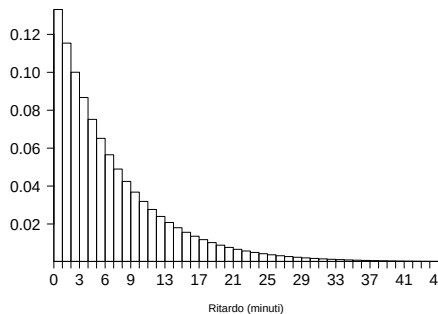
- Abbiamo infinite osservazioni, possiamo affinare l'istogramma

classe	prob
(0,2.5]	0.30
(2.5,5]	0.21
(5,7.5]	0.15
(7.5,10]	0.10
(10,12.5]	0.07
(12.5,15]	0.05
(15,20]	0.06
(20,25]	0.03
(25,35]	0.02
(35,Inf]	0.01



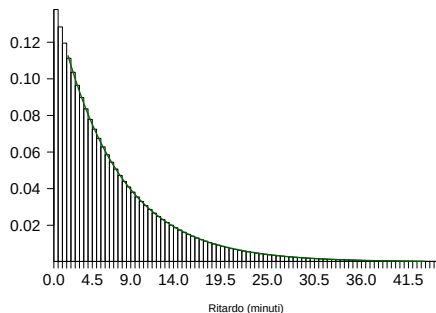
Da istogramma a densità

- Abbiamo infinite osservazioni, possiamo affinare l'istogramma



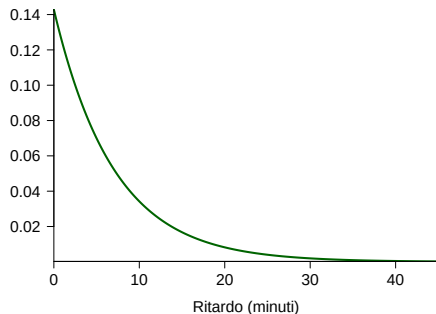
Da istogramma a densità

- Abbiamo infinite osservazioni, possiamo affinare l'istogramma
- fino a sostituirlo con una curva: la **funzione di densità**



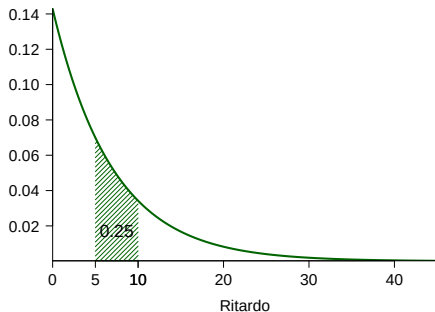
Da istogramma a densità

- Abbiamo infinite osservazioni, possiamo affinare l'istogramma
- fino a sostituirlo con una curva: la **funzione di densità**



Interpretazione della f.d.d.

L'interpretazione è analoga all'istogramma: le aree rappresentano probabilità anziché frequenze.



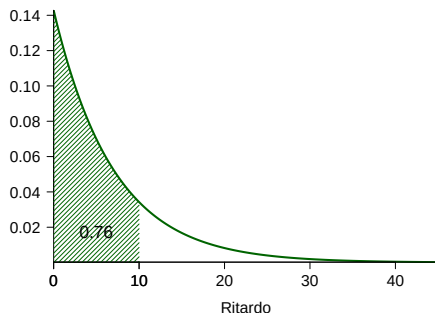
La probabilità di osservare un valore tra 5 e 10 è 0.25, in simboli

$$P(5 \leq X \leq 10) = 0.25$$

F.d.d. e probabilità

Di particolare importanza è

$$P(X \leq x)$$

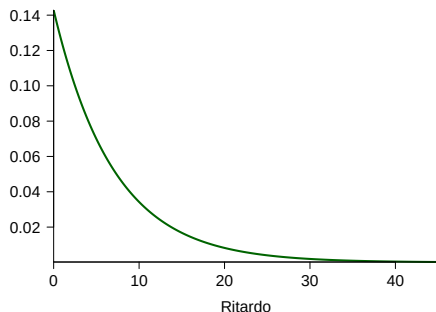


La probabilità di osservare X inferiore a 10 è 0.76, in simboli

$$P(X \leq 10) = 0.76$$

Rappresentazione della f.d.d.

La funzione di densità è descritta da una funzione reale a valori non negativi



ad esempio quella in figura è

$$f(x) = \frac{1}{7}e^{-x/7}$$

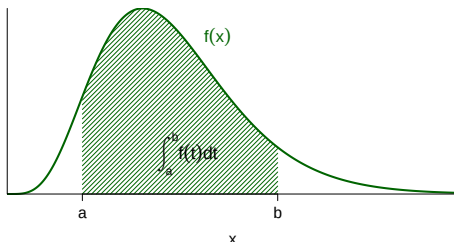
per $x \geq 0$.

Funzione di densità

Funzione di densità

La funzione di densità di una v.a. X è una funzione $f(x)$, non negativa, e la cui area sottesa (integrale) è la probabilità di X :

$$\int_a^b f(t)dt = P(a \leq X \leq b)$$

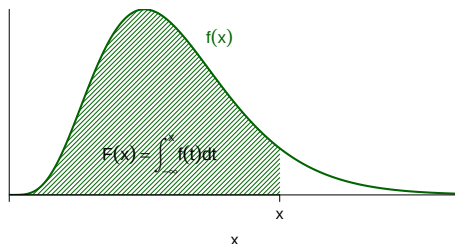


Funzione di ripartizione

Funzione di ripartizione

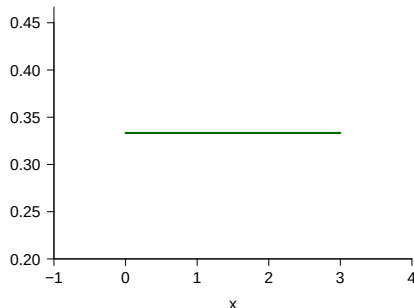
La funzione di ripartizione di una v.a. X è la funzione

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$$



F.d.d.: esempio

Ci sono numerose funzioni di densità, qualunque funzione a valori non negativi con integrale (area totale sottesa) pari a 1 è una funzione di densità.



Uniforme tra 0 e 3

$$f(x) = 1/3$$

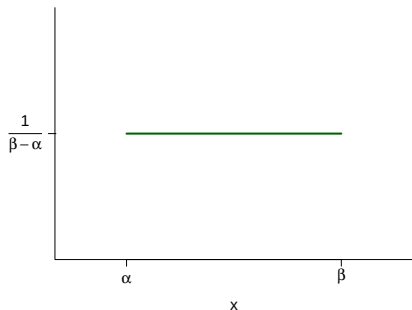
$$0 \leq x \leq 3$$

Qual è $P(1 \leq X \leq 2)$?

F.d.d. e parametri

- Molto spesso è utile definire una f.d.d. mediante parametri (come visto per la binomiale),
- che è lo stesso che dire che si definisce un insieme di f.d.d.

Ad esempio

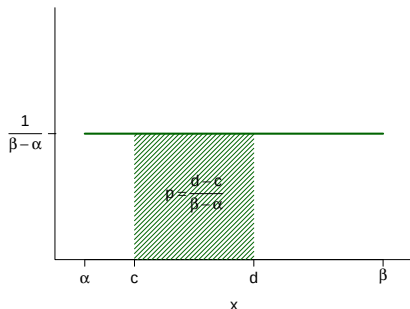


$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{se } \alpha \leq x \leq \beta \\ 0 & \text{altrimenti} \end{cases}$$

- è una f.d.d. qualunque siano α e β .

Distribuzione uniforme su $[a, b]$

La distribuzione uniforme è molto semplice, possiamo calcolare le probabilità ad essa riferite



$$P(c \leq X \leq d) = \frac{d - c}{\beta - \alpha}$$

Media e varianza per v.a. continue

Anche per le v.a. continue si definiscono media e varianza, se $f(x)$ rappresenta la f.d.d. si ha

$$E(X) = \int xf(x)dx$$

$$E(X^2) = \int x^2f(x)dx$$

$$V(X) = \int (x - E(X))^2 f(x)dx$$

Con le solite proprietà

$$E(aX + b) = aE(X) + b$$

$$V(X) = E(X^2) - [E(X)]^2$$

$$V(aX + b) = a^2 V(X)$$

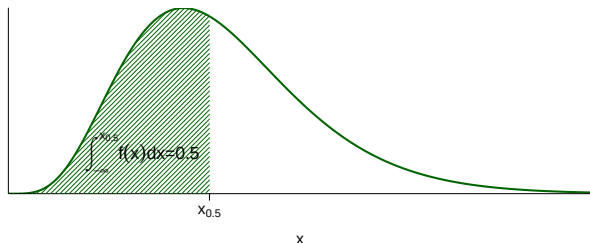
Mediana, quartili e quantili

Si definisce la mediana di una v.a. continua X con distribuzione $f(x)$ come il valore $Me(X)$ o $x_{0.5}$ tale che

$$F(Me(X)) = \int_{-\infty}^{Me(X)} f(x)dx = 0.5$$

Si definisce q -quantile di una v.a. continua X con distribuzione $f(x)$ come il valore x_q tale che

$$F(x_q) = \int_{-\infty}^{x_q} f(x)dx = q$$

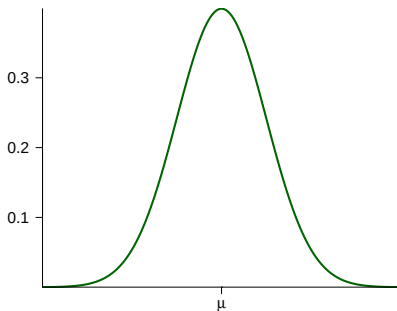


Indice

- 1 Concetti base
- 2 Variabili aleatorie discrete
- 3 Media e varianza
- 4 Variabili aleatorie continue
- 5 La distribuzione normale**
- 6 Normale come approssimazione
- 7 Riepilogo

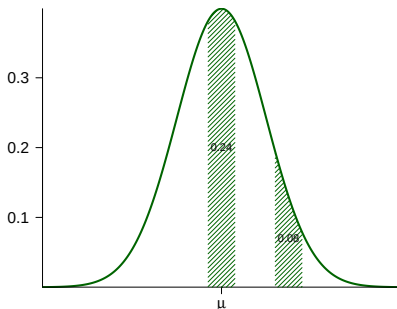
La distribuzione normale: principali caratteristiche

La distribuzione continua che si incontra più frequentemente è la normale o gaussiana, la cui densità ha la seguente forma



La distribuzione normale: principali caratteristiche

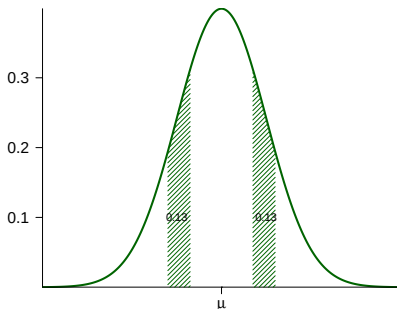
La distribuzione continua che si incontra più frequentemente è la normale o gaussiana, la cui densità ha la seguente forma



- Sono più probabili valori intorno a un valore centrale.

La distribuzione normale: principali caratteristiche

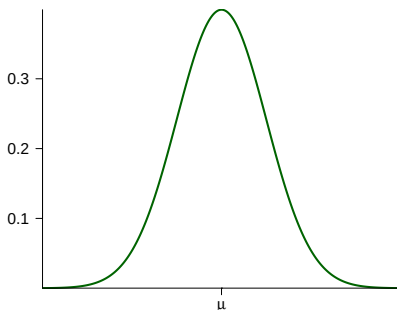
La distribuzione continua che si incontra più frequentemente è la normale o gaussiana, la cui densità ha la seguente forma



- Sono più probabili valori intorno a un valore centrale.
- Scostamenti a destra e sinistra hanno la stessa probabilità.

La distribuzione normale: principali caratteristiche

La distribuzione continua che si incontra più frequentemente è la normale o gaussiana, la cui densità ha la seguente forma



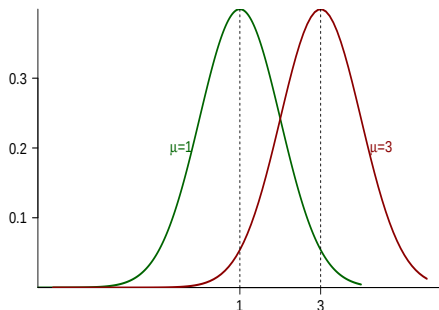
- Sono più probabili valori intorno a un valore centrale.
- Scostamenti a destra e sinistra hanno la stessa probabilità.
- Funzione: $f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$.

Parametri della normale

La formula della densità della normale:

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

non è importante, è importante il significato dei parametri μ e σ , tanto che la indicheremo semplicemente con $N(\mu, \sigma^2)$ e scriveremo $X \sim N(\mu, \sigma^2)$. (Attenzione: in alcuni testi si indica $N(\mu, \sigma)$!)



μ è la media della distribuzione.

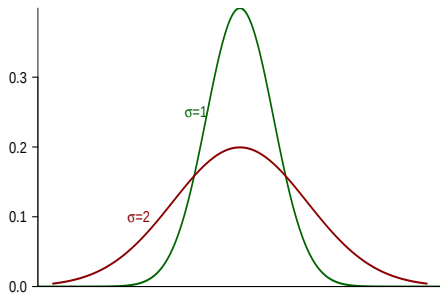
Al variare di μ la distribuzione si sposta lungo l'asse x .

Parametri della normale

La formula della densità della normale:

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

non è importante, è importante il significato dei parametri μ e σ , tanto che la indicheremo semplicemente con $N(\mu, \sigma^2)$ e scriveremo $X \sim N(\mu, \sigma^2)$. (Attenzione: in alcuni testi si indica $N(\mu, \sigma)$!)



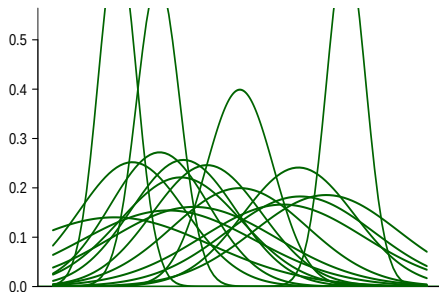
σ^2 è la varianza della distribuzione.
Al variare di σ la distribuzione mantiene lo stesso centro, ma si 'allarga': sono più probabili valori grandi (in valore assoluto)

Parametri della normale

La formula della densità della normale:

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

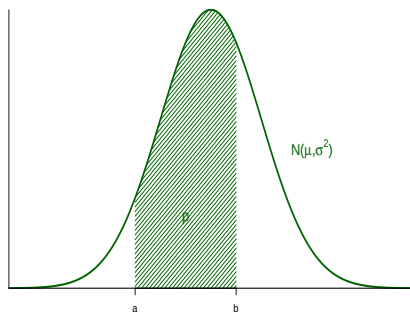
non è importante, è importante il significato dei parametri μ e σ , tanto che la indicheremo semplicemente con $N(\mu, \sigma^2)$ e scriveremo $X \sim N(\mu, \sigma^2)$. (Attenzione: in alcuni testi si indica $N(\mu, \sigma)$!)



Insomma, abbiamo un'infinità di distribuzioni al variare di μ e σ .

Parametri della normale

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



Ci servirà, sapendo che $X \sim N(\mu, \sigma^2)$, poter calcolare probabilità del tipo

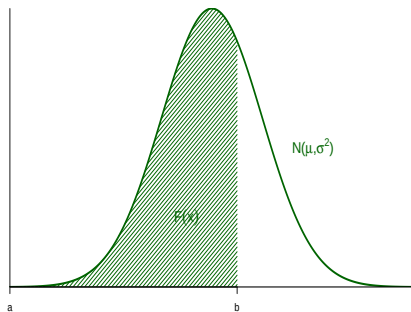
$$p = P(a \leq X \leq b)$$

dove

$$p = \int_a^b f(t) dt$$

Parametri della normale

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



Ci basta, in realtà, poter valutare

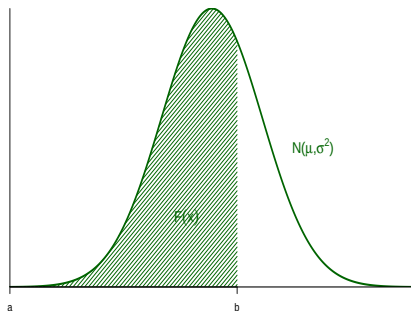
$$F(x) = P(X \leq x)$$

dove

$$F(x) = \int_{-\infty}^x f(t) dt$$

Parametri della normale

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



Sfortunatamente, non c'è un'espressione esplicita per calcolare queste probabilità, però potremo calcolarle riferendoci a una particolare normale.

Normale Standard

Un ruolo particolare lo ha la normale standard, con $\mu = 0$ e $\sigma = 1$

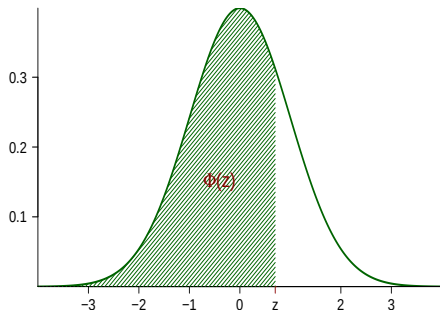
$$Z \sim N(0, 1) \quad f(z; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

Definiamo

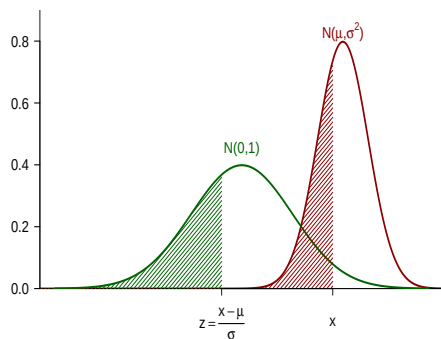
$$\Phi(z) = P(Z \leq z)$$

area sottesa dalla curva tra $-\infty$ e z , cioè l'area scurita in figura.

(Si dà un nome particolare alla FdR della normale standard.)



Normale Standard e normale (μ, σ)



Si verifica che se X è distribuito secondo una $N(\mu, \sigma^2)$, allora

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

cioè l'area in rosso (pari a $P(X \leq x)$) è uguale all'area verde.

Questo permette di ricavare qualunque probabilità associata ad una generica $N(\mu, \sigma^2)$ conoscendo $\Phi(z)$.

Tavole della normale

- La funzione $\Phi(z)$ non è esprimibile in forma analitica, ma è tabulata.
- Cioè i valori di $\Phi(z)$ corrispondenti a diversi valori di z sono riportati in una tabella la cui forma è la seguente

	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
					.					
					.					
0.2					.					
0.3	.	.	.	.	$\Phi(0.34)$					
0.4										

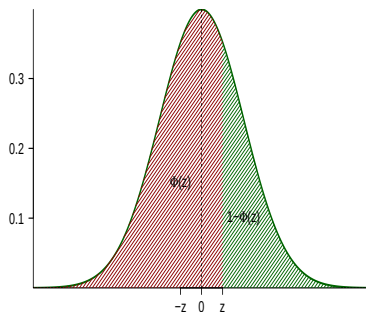
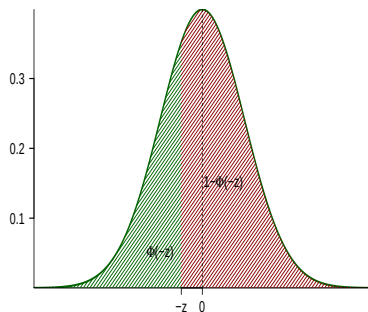
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

E i valori negativi?

È sufficiente conoscere $\Phi(z)$ per i valori positivi di z , se z è minore di 0 si usa la relazione

$$\Phi(-z) = 1 - \Phi(z)$$

Graficamente: a colori uguali corrispondono aree uguali



Alcuni esempi I

Sia X distribuito secondo una normale standard

$$P(X \leq 1.34) = ?$$

La parte rilevante della tavola è

	0.01	0.02	0.03	0.04	0.05	0.06	0.07
1							
1.1	0.8686	0.8708	0.8729	0.8749	0.8770		
1.2	0.8888	0.8907	0.8925	0.8944	0.8962		
1.3	0.9066	0.9082	0.9099	0.9115	0.9131		
1.4	0.9222	0.9236	0.9251	0.9265	0.9279		
1.5	0.9357	0.9370	0.9382	0.9394	0.9406		
1.6							

Alcuni esempi II

Sia X distribuito secondo una normale standard

$$P(X > -1.23) = ?$$

La parte rilevante della tavola è

	0	0.01	0.02	0.03	0.04	0.05	0.06
0.9							
1	0.8438	0.8461	0.8485	0.8508	0.8531		
1.1	0.8665	0.8686	0.8708	0.8729	0.8749		
1.2	0.8869	0.8888	0.8907	0.8925	0.8944		
1.3	0.9049	0.9066	0.9082	0.9099	0.9115		
1.4	0.9207	0.9222	0.9236	0.9251	0.9265		
1.5							

Alcuni esempi III

Sia X distribuito secondo una normale con media $\mu = 4$ e varianza $\sigma^2 = 16$

$$P(X \leq 13) = ?$$

Calcoliamo $z = (13 - 4)/4 = 2.25$, si ha allora

$$P(X \leq 13) = \Phi(2.25) = ?$$

La parte rilevante della tavola è

	0.01	0.02	0.03	0.04	0.05	0.06	0.07
1.9							
2		0.9783	0.9788	0.9793	0.9798	0.9803	
2.1		0.9830	0.9834	0.9838	0.9842	0.9846	
2.2		0.9868	0.9871	0.9875	0.9878	0.9881	
2.3		0.9898	0.9901	0.9904	0.9906	0.9909	
2.4		0.9922	0.9925	0.9927	0.9929	0.9931	
2.5							

Quantili della normale: z_q t.c. $\Phi(z_q) = q$!

Poniamoci il problema inverso, vogliamo trovare $z_{0.9}$, tale che

$$\Phi(z_{0.9}) = 0.9$$

Faremo il lavoro inverso, cerchiamo sulla tavola un valore vicino a 0.9, e leggiamo il corrispondente valore di z

	0.06	0.07	0.08	0.09
0.9				
1	0.8577	0.8599	0.8621	
1.1	0.8790	0.8810	0.8830	
1.2	0.8980	0.8997	0.9015	
1.3	0.9147	0.9162	0.9177	
1.4	0.9292	0.9306	0.9319	
1.5				

Quantili della normale: z_q t.c. $\Phi(z_q) = q$ ||

Sapendo che

$$\Phi(1.28) = 0.9$$

Possiamo trovare facilmente $z_{0.1}$ cioè il valore tale che

$$\Phi(z_{0.1}) = 0.1$$

Come?

Quantili della normale: z_q t.c. $\Phi(z_q) = q$ III

Troviamo anche $z_{0.95}$, il valore tale che

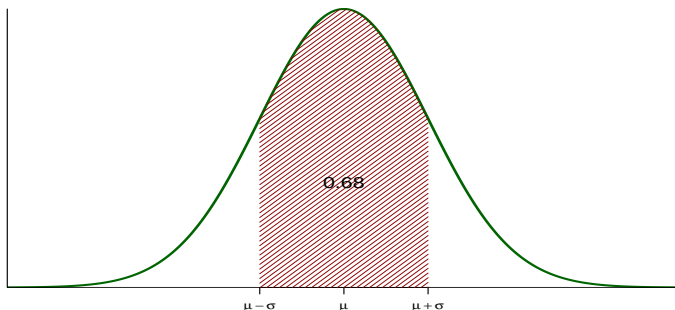
$$\Phi(z_{0.95}) = 0.95$$

Cerchiamo sulla tavola un valore vicino a 0.9, e leggiamo il corrispondente valore di z

	0.01	0.02	0.03	0.04	0.05	0.06	0.07
1.3							
1.4	0.9222	0.9236	0.9251	0.9265	0.9279		
1.5	0.9357	0.9370	0.9382	0.9394	0.9406		
1.6	0.9474	0.9484	0.9495	0.9505	0.9515		
1.7	0.9573	0.9582	0.9591	0.9599	0.9608		
1.8	0.9656	0.9664	0.9671	0.9678	0.9686		
1.9							

Intervalli di riferimento I

L'intervallo da $\mu - \sigma$ a $\mu + \sigma$ comprende una probabilità pari a 0.68.



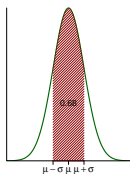
Come lo verifichiamo?

Intervalli di riferimento II

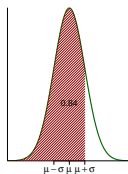
Un aiuto, la parte rilevante della tavola è

	0	0.01	0.02	0.03
0.7				
0.8	0.7881	0.7910	0.7939	
0.9	0.8159	0.8186	0.8212	
1	0.8413	0.8438	0.8461	
1.1	0.8643	0.8665	0.8686	
1.2	0.8849	0.8869	0.8888	
1.3				

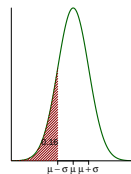
Intervalli di riferimento III



=



-



0.68

=

$$\Phi\left(\frac{(\mu + \sigma) - \mu}{\sigma}\right)$$

-

$$\Phi\left(\frac{(\mu - \sigma) - \mu}{\sigma}\right)$$

=

$$\Phi(1)$$

-

$$\Phi(-1)$$

=

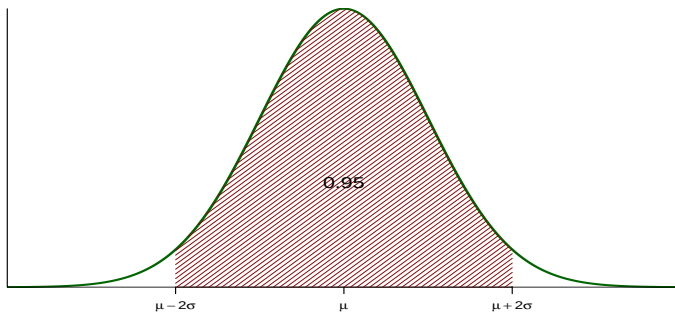
0.84

-

0.16

Intervalli di riferimento IV

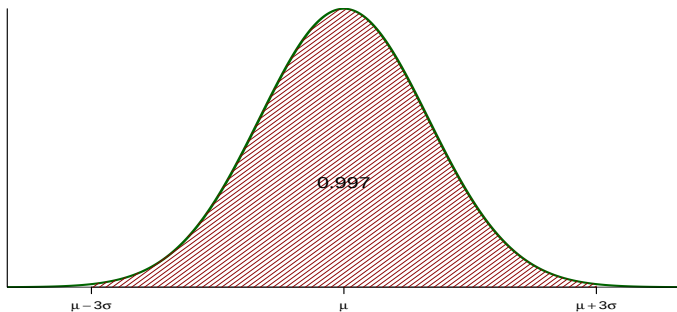
L'intervallo da $\mu - 2\sigma$ a $\mu + 2\sigma$ comprende una probabilità pari a 0.95.



Si verifica in maniera analoga. Provare!

Intervalli di riferimento V

L'intervallo da $\mu - 3\sigma$ a $\mu + 3\sigma$ comprende una probabilità pari a 0.997.



La quasi totalità delle osservazioni dovrebbe trovarsi tra $\mu - 3\sigma$ e $\mu + 3\sigma$

Esercizio: temperature a TS

- La temperatura media a Trieste in un giorno di aprile è una variabile aleatoria che è ben descritta da una normale di media 13.5.
- Sapendo che è 0.71 la probabilità che la temperatura sia inferiore a 15, possiamo ricavare la deviazione standard?

Esercizio: temperature a TS

- La temperatura media a Trieste in un giorno di aprile è una variabile aleatoria che è ben descritta da una normale di media 13.5.
- Sapendo che è 0.71 la probabilità che la temperatura sia inferiore a 15, possiamo ricavare la deviazione standard?
- La risposta è sì.

Esercizio: temperature a TS

- Ragioniamo così, sia σ la deviazione standard, non nota.
- Conoscendo σ , calcoleremmo la probabilità che la temperatura sia al di sotto di 15 come

$$\Phi\left(\frac{15 - 13.5}{\sigma}\right).$$

- Noi però sappiamo che questa è pari a 0.71.
- Guardiamo al solito la tavola:

Esercizio: temperature a TS

- Ragioniamo così, sia σ la deviazione standard, non nota.
- Conoscendo σ , calcoleremmo la probabilità che la temperatura sia al di sotto di 15 come

$$\Phi\left(\frac{15 - 13.5}{\sigma}\right).$$

- Noi però sappiamo che questa è pari a 0.71.
- Guardiamo al solito la tavola:

	0.02	0.03	0.04	0.05	0.06	0.07	0.08
0.2							
0.3		0.6293	0.6331	0.6368	0.6406	0.6443	
0.4		0.6664	0.6700	0.6736	0.6772	0.6808	
0.5		0.7019	0.7054	0.7088	0.7123	0.7157	
0.6		0.7357	0.7389	0.7422	0.7454	0.7486	
0.7		0.7673	0.7704	0.7734	0.7764	0.7794	
0.8							

Esercizio: temperature a TS

- Ragioniamo così, sia σ la deviazione standard, non nota.
- Conoscendo σ , calcoleremmo la probabilità che la temperatura sia al di sotto di 15 come

$$\Phi\left(\frac{15 - 13.5}{\sigma}\right).$$

- Noi però sappiamo che questa è pari a 0.71.
- Guardiamo al solito la tavola: $\Phi(0.55) = 0.71$ (approssimativamente).

Esercizio: temperature a TS

- Ragioniamo così, sia σ la deviazione standard, non nota.
- Conoscendo σ , calcoleremmo la probabilità che la temperatura sia al di sotto di 15 come

$$\Phi\left(\frac{15 - 13.5}{\sigma}\right).$$

- Noi però sappiamo che questa è pari a 0.71.
- Guardiamo al solito la tavola: $\Phi(0.55) = 0.71$ (approssimativamente).
- Allora abbiamo

$$\frac{15 - 13.5}{\sigma} = 0.55$$

Esercizio: temperature a TS

- Ragioniamo così, sia σ la deviazione standard, non nota.
- Conoscendo σ , calcoleremmo la probabilità che la temperatura sia al di sotto di 15 come

$$\Phi\left(\frac{15 - 13.5}{\sigma}\right).$$

- Noi però sappiamo che questa è pari a 0.71.
- Guardiamo al solito la tavola: $\Phi(0.55) = 0.71$ (approssimativamente).
- Allora abbiamo

$$\frac{15 - 13.5}{\sigma} = 0.55$$

- da cui

$$\sigma = \frac{15 - 13.5}{0.55} = 2.7$$

- La temperatura media giornaliera in aprile a Trieste è quindi distribuita secondo una normale con media 13.5 e deviazione standard 2.7.

Temperature a TS

Sapendo ora che la temperatura media giornaliera in aprile a Trieste è quindi distribuita secondo una normale con media 13.5 e deviazione standard 2.7.

- 1 S'individuì un intervallo di temperature che, al 99% contiene la temperatura media.
- 2 Qual è la probabilità che la temperatura media in un dato giorno sia pari a 3.0 gradi?
- 3 Qual è la probabilità che la temperatura media in un dato giorno sia inferiore a 3 gradi?
- 4 Qual è la probabilità che la temperatura media in un dato giorno sia inferiore o uguale a 3 gradi?

Trasformazioni e normalità

Trasformato di una normale

Se X è distribuito secondo una $N(\mu, \sigma^2)$, e a, b sono due numeri reali, allora

$$Y = aX + b$$

è distribuito secondo una $N(a\mu + b, a^2\sigma^2)$.

- Se il valore di un'azione in borsa in un dato giorno espresso in euro è $N(14, \sigma^2 = 2)$ allora il valore espresso in dollari ($1\$ = 0.78\text{€}$) è normale con media 0.78×14 e varianza $2 \times (0.78)^2$.

Esempio: temperature e trasformazioni

La temperatura media giornaliera in una località, espressa in gradi Fahrenheit è distribuita secondo una $N(68, \sigma^2 = 9)$, qual è la probabilità che si superino i 20 gradi Celsius?

- Dovremo anzitutto tenere conto che la temperatura in gradi Celsius si ottiene, a partire da quella in gradi Fahrenheit come

$$C = (5/9) \times (F - 32)$$

si ha dunque una trasformazione del tipo $aX + b$ con $a = 5/9$ e $b = -(5/9) \times 32$.

- Applicando la regola sopra si ottiene quindi che la temperatura in gradi Celsius è distribuita secondo una $N(20, \sigma^2 = 2.78)$.
- (Infatti $(5/9)(68 - 32) = 20$ e $(5/9)^2 \times 9 = 2.78$.)
- Qual è dunque la probabilità richiesta?

Indipendenza tra v.a.

- Indipendenza tra eventi A e B significa che $P(A|B) = P(A)$: sapere che B si è verificato non è influente per la probabilità di A .
- Indipendenza tra v.a.: è lo stesso per gli eventi riferiti alle v.a.

$$P(c \leq X \leq d | Y = y) = P(c \leq X \leq d)$$

qualunque siano c , d , y .

- Conoscere Y non dice nulla su X .
- I risultati dei lanci di due dadi sono indipendenti.
- Temperatura e pressione in una località non sono indipendenti.
- Le temperature di due località vicine non sono indipendenti.

Normalità della somma

Somma di normali

Se X_1 è distribuito secondo una $N(\mu_1, \sigma_1^2)$, e X_2 secondo una $N(\mu_2, \sigma_2^2)$ e X_1, X_2 sono indipendenti, allora

$$Y = a_1 X_1 + a_2 X_2$$

è distribuito secondo una

$$N(a_1 \mu_1 + a_2 \mu_2, a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2).$$

Esempio: somme

Supponiamo che il peso (espresso in kg) in una popolazione sia distribuito secondo una $N(70, \sigma^2 = 50)$.

Due individui, presi a caso nella popolazione, salgono su un montacarichi che porta al più 160kg, qual è la probabilità che la coppia non superi tale peso?

- Per un individuo a caso, dunque il peso è una $N(70, \sigma^2 = 100)$.
- Il peso totale dei due, $X_1 + X_2$ è distribuito secondo una

$$N(70 + 70 = 140, 50 + 50)$$

- la probabilità cercata è allora

$$\Phi\left(\frac{160 - 140}{10}\right) = \Phi(2) = 0.98$$

Esempio: somme

Supponiamo che il peso (espresso in kg) in una popolazione sia distribuito secondo una $N(70, \sigma^2 = 50)$.

Come cambia la risposta se gli individui sono tre?

- Il peso totale dei tre, $X_1 + X_2 + X_3$ è distribuito secondo una

$$N(70 + 70 + 70 = 210, 50 + 50 + 50 = 150)$$

- la probabilità cercata è allora

$$\Phi\left(\frac{160 - 210}{12.25}\right) = \Phi(-4.1) \approx 0$$

Esempio: somme

Supponiamo che il peso (espresso in kg) in una popolazione sia distribuito secondo una $N(70, \sigma^2 = 50)$.

Come cambia la risposta se gli individui sono n ?

- Il peso totale dei tre, $X_1 + X_2 + X_3 + \dots + X_n$ è distribuito secondo una

$$N(70n, 50n)$$

Esempio: somme

Supponiamo che il peso (espresso in kg) in una popolazione sia distribuito secondo una $N(70, \sigma^2 = 50)$.

Come cambia la risposta se gli individui sono n ?

- Il peso totale dei tre, $X_1 + X_2 + X_3 + \dots + X_n$ è distribuito secondo una

$$N(70n, 50n)$$

- Cosa possiamo dire della media?
- La media è la somma moltiplicata $1/n$: $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, pertanto è distribuita secondo una

$$N\left(70, \frac{50}{n}\right)$$

Somma e media di normali

Somma e media di n normali indipendenti

Siano $X_1, \dots, X_n$ normali con media μ e varianza σ^2 indipendenti, allora

$$\sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2)$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Indice

- 1 Concetti base
- 2 Variabili aleatorie discrete
- 3 Media e varianza
- 4 Variabili aleatorie continue
- 5 La distribuzione normale
- 6 Normale come approssimazione**
- 7 Riepilogo

Ruolo della normale

- Una ragione per cui la normale è usata molto di frequente, è che è una valida approssimazione per molte situazioni.
- In particolare, quando si sommano molte variabili, qualunque distribuzione essi abbiano (quasi), la loro somma è approssimativamente normale.
- Il risultato detto a parole sopra è il Teorema del limite centrale.

Normale come approssimazione della binomiale

- Consideriamo una binomiale di dimensione n e probabilità p

$$P(S = s) = \binom{n}{s} p^s (1 - p)^{n-s}.$$

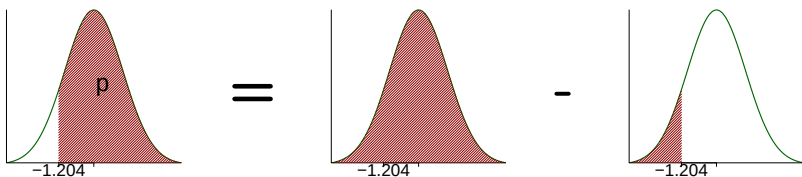
- Questa è approssimabile con una normale di media np e varianza $np(1 - p)$.

Elezioni: elettori M5S e approssimazione binomiale I

- Abbiamo visto che, preso un votante, la probabilità che abbia votato M5S il 24/25 febbraio 2013 è 0.2464.
- Supponiamo di prendere 1000 elettori, il numero di persone che hanno votato M5S è una binomiale con media 246.4 e varianza 185.687.
- (In realtà questa è un'approssimazione che presuppone di estrarre con rimpiazzo, cioè di poter estrarre due volte la stessa persona, l'effetto di questa approssimazione è però trascurabile.)
- Qual è la probabilità che, tra i 1000 più di 230 abbiano votato M5S
- Usando l'approssimazione normale, è

$$p = 1 - \Phi \left(\frac{230 - 246.4}{\sqrt{185.687}} \right) = 1 - \Phi(-1.2035) = 0.8856$$

Elezioni: elettori M5S e approssimazione binomiale II



$$\begin{aligned}
 p &= 1 - \Phi\left(\frac{230 - 246.4}{\sqrt{185.687}}\right) \\
 &= 1 - \Phi(-1.2035) \\
 &= 1 - 0.1144
 \end{aligned}$$

Elezioni: elettori M5S e approssimazione binomiale III

- Qual è la probabilità che la percentuale di votanti il M5S sia compresa tra 23 e 27 %?
- Possiamo ragionare osservando che, se il numero di votanti è

$$\text{normale}(np, np(1 - p))$$

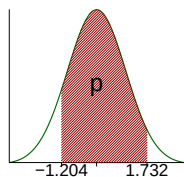
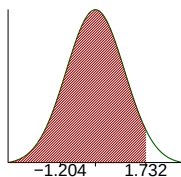
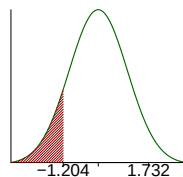
allora la percentuale è

$$\text{normale}(p, p(1 - p)/n)$$

- la probabilità richiesta è allora

$$\begin{aligned} p &= \Phi\left(\frac{0.27 - 0.2464}{\sqrt{0.0002}}\right) - \Phi\left(\frac{0.23 - 0.2464}{\sqrt{0.0002}}\right) = \\ &= \Phi(1.7319) - \Phi(-1.2035) = \\ &= 0.9584 - 0.1144 = 0.844 \end{aligned}$$

Elezioni: elettori M5S e approssimazione binomiale IV


 $=$

 $-$


$$\begin{aligned}
 p &= \Phi\left(\frac{0.27-0.2464}{\sqrt{\quad}}\right) - \Phi\left(\frac{0.23-0.2464}{\sqrt{\quad}}\right) \\
 &= \Phi(1.7319) - \Phi(-1.2035) \\
 &= 0.9584 - 0.1144
 \end{aligned}$$

Indice

- 1 Concetti base
- 2 Variabili aleatorie discrete
- 3 Media e varianza
- 4 Variabili aleatorie continue
- 5 La distribuzione normale
- 6 Normale come approssimazione
- 7 Riepilogo

Variabili aleatorie: fatti importanti

Si definiscono mediante parametri  Si dividono in

- Discrete

- si definisce la funzione di probabilità, che fornisce la probabilità per ciascun possibile valore. 
- Si definiscono media e varianza. 
- Esempio notevole: binomiale. 

- Continue

- si definisce la funzione di densità, il cui integrale è la probabilità. 
- Si definiscono media e varianza. 

Distribuzione normale: fatti importanti

Si è definita la distribuzione normale, con media μ e varianza σ^2 , .

Si definisce inoltre

$$\Phi(z) = P(Z \leq z), \quad Z \text{ normale standard} $$

per la quale e si hanno le tavole . Inoltre

- Si definisce z_q il valore tale che $\Phi(z_q) = q$. 
- $\Phi(z) = 1 - \Phi(-z)$. 
- Se X è distribuita secondo una $N(\mu, \sigma)$ allora

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

- Somma e media di normali 
- La binomiale (n, p) è approssimata dalla $N(np, np(1 - p))$. 