UNIVERSITÀ DEGLI STUDI DI TERAMO

ECONOMIA

METODI STATISTICI PER L'ANALISI ECONOMICA E AZIENDALE

LA PROBABILITÀ E LE PRINCIPALI VARIABILI CASUALI DISCRETE E CONTINUE

FABRIZIO ANTOLINI
fantolini@unite.it

PROBABILITÀ (CONCETTI BASE, PROBABILITA' CONGIUNTA E CONDIZIONATA)

In termini molto sintetici la probabilità può essere intesa come la misura del grado di incertezza connesso al risultato scaturito da una prova.

- La prova è un esperimento avente due o più risultati possibili
- L'evento rappresenta uno dei possibili risultati della prova

La probabilità viene espressa per mezzo di un numero compreso tra 0 ed 1 associato al grado di incertezza del verificarsi di un evento.

In una data prova, l'evento E si verifica con una probabilità ***P(E):*** Nel classico esempio del lancio del dato (ben bilanciato), infatti il verificarsi dell'evento "faccia contrassegnata dal numero 5) si presenta con probabilità 1/6= 0,16.

Definizione classica di probabilità

La probabilità è data dal rapporto tra il numero dei casi favorevoli all'evento e il numero dei casi possibili purché essi siano tutti ugualmente possibili.

$$\boldsymbol{P(E) = \frac{n.\, casi\ favorevoli}{n.\, casi\ possibili}}$$

PROBABILITÀ (CONCETTI BASE, PROBABILITA' CONGIUNTA E CONDIZIONATA)

L'insieme di tutti i possibili eventi elementari viene chiamato **spazio campionario** e viene indicato con il simbolo **Ω**.

- Un evento viene definito <u>impossibile</u> se non può mai verificarsi e può essere definito come ***P(E) = 0***
- Un evento viene definito <u>certo</u> se si verifica sempre in quanto comprende tutti i possibili risultati dell'esperimento. Può essere definito come ***P(E) = 1***
- Due eventi A e B si dicono incompatibili se la loro probabilità congiunta è uguale a zero. Una probabilità congiunta, nella teoria della probabilità, si riferisce alla probabilità che due eventi si verifichino entrambi. In altre parole, la probabilità congiunta è la probabilità che due eventi si verifichino insieme.

In simboli $A \cap B = 0$

Una prima definizione più appropriata della probabilità può quindi essere intesa come la funzione di insieme che associa ad ogni evento $E_i \in E$ un numero reale.

Si definisce, invece, **probabilità condizionata (o probabilità a posteriori)** la probabilità che si verifichi un evento A sapendo che si è già verificato un altro evento B. La probabilità condizionata di A rispetto a B si indica con P(A|B) e si può definire a patto che la probabilità B sia diversa da zero:

$$\boldsymbol{P(A|B) = \frac{n.\,casi\ favorevoli\ ad\ (A \cap B)}{n.\,casi\ favorevoli\ a\ B}}$$

Ossia:

$$\boldsymbol{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$

La probabilità condizionata entra in gioco tutte le volte che si vuole calcolare la probabilità di un evento A, detto <u>evento condizionato</u>, assumendo che si sia già verificato un altro evento B, chiamato <u>evento condizionante</u>.

Affinché i calcoli di probabilità congiunti funzionino, gli eventi devono essere indipendenti. In altre parole, gli eventi non devono potersi influenzare a vicenda. Per determinare se due eventi sono indipendenti o dipendenti, è importante chiedersi se il risultato di un evento avrebbe un impatto sul risultato dell'altro evento. Se il risultato di un evento non influisce sul risultato dell'altro evento, gli eventi sono indipendenti.

In simboli, dati 2 eventi A e B tali che $P(A)>0$ ***e*** $P(B)>0$***:***

$$P(A \cap B) = P(A)\,P(B|A) = P(B)P(A|B)$$

Due eventi si dicono indipendenti se il verificarsi di B non influenza la probabilità di A e il verificarsi di A non influenza la probabilità di B.

In simboli

$$P(A|B)=P(A) \text{ e } P(B|A)=P(B)$$

Ossia:

$$P(A \cap B) = P(A) * P(B)$$

TEOREMA DI BAYES

Il teorema di Bayes conosciuto anche come formula di Bayes o teorema della probabilità delle cause), enunciato da Thomas Bayes (1702-1761) viene impiegato per calcolare la probabilità di una causa che ha provocato l'evento verificato.

Data la probabilità a posteriori, il teorema è così formulato:

$$P(A_i|B) = \frac{P(A_i) * P(B|A_i)}{P(A_1) * P(B|A_1) + \cdots \ldots . + P(A_k) * P(B|A_k)}$$

- $P(A_i)$ è la probabilità a posteriori
- $P(B|A_i)$ è la probabilità condizionante o verosimiglianze
- $P(A_i|B)$ è la probabilità a posteriori in quanto si riferiscono agli eventi A_i, dopo aver osservato l’evento B.

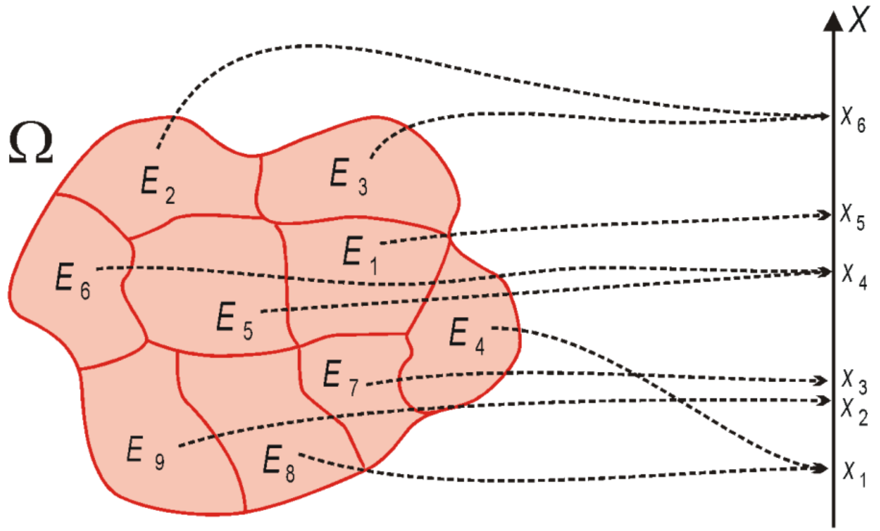
CONCEZIONI DELLA PROBABILITÀ (accenni)

- ***Frequentista***: basata sul postulato empirico del caso, ovvero dove in un gruppo di prove, ripetute più volte nelle stesse condizioni, ciascuno degli eventi possibili compare con una frequenza quasi eguale alla sua probabilità; generalmente l’approssimazione migliora quando il numero delle prove cresce

- ***Soggettivista:*** la probabilità di un evento è la misura del grado di fiducia che un individuo (il soggetto) coerente attribuisce al verificarsi dell’evento, in base alle informazioni in suo possesso

- ***Probabilità soggettiva:*** la probabilità di un evento E, secondo l’opinione di un individuo coerente, è il prezzo p che egli stima equo attribuire ad un importo unitario (ad esempio 1 euro) esigibile solo al verificarsi di E

Una variabile casuale ***(v.c) X*** è una funzione definita sullo spazio campionario ***Ω*** che associa ad ogni evento ***E ⊂ Ω*** un unico numero reale.



- Una <u>variabile casuale discreta</u> può assumere un insieme discreto (finito o numerabile) di numeri reali
- Una <u>variabile casuale continua</u> può assumere tutti i valori compresi in un intervallo reale.

VARIABILI CASUALI DISCRETE

La funzione di probabilità di una variabile casuale discreta ***X*** associa ad ognuno dei valori x_i la corrispondente probabilità $P(X=x_i)$

$$P(X=x_i) \rightarrow \text{Probabilità che la v.c. } X \text{ assuma il valore } x_i$$

- $\sum_i P(xi) = 1$
- $P(xi) \geq 0$

Funzione di Ripartizione

Data una ***v.c*** discreta ***X***, la funzione che fa corrispondere ai valori ***x*** le probabilità cumulate viene detta funzione di ripartizione ed indicata con:

$$F(x) = P(X \leq x) = \sum_{w \leq x} P(X = w)$$

Esempio

X	2	3	4	5	6	7	8	9	10	11	12
$P(x)$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$
$F(x)$	$1/36$	$3/36$	$6/36$	$10/36$	$15/36$	$21/36$	$26/36$	$30/36$	$33/36$	$35/36$	1

VARIABILI CASUALI DISCRETE

- La funzione di ripartizione è non decrescente:

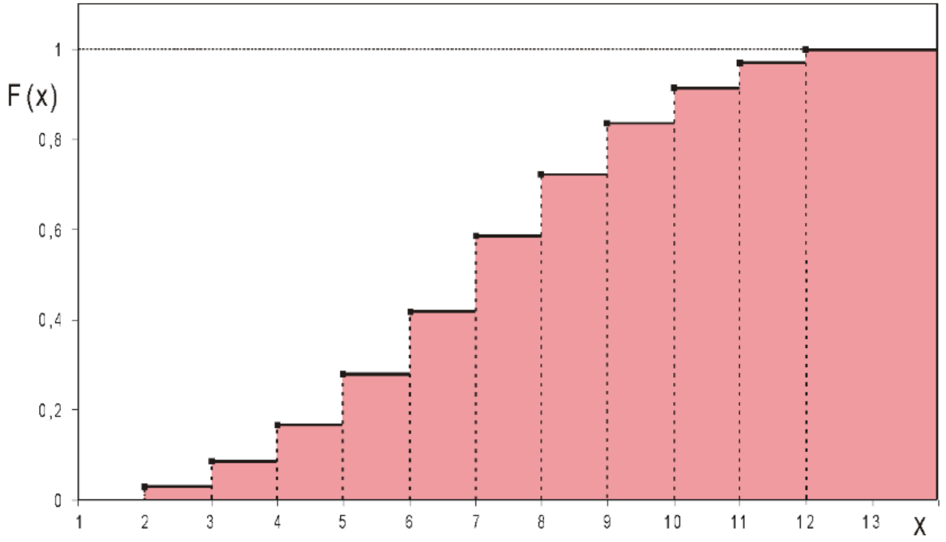
$$\boldsymbol{x_1 < x_2 \rightarrow F(x_1) \leq F(x_2)}$$

- Inoltre:

$$\lim_{x \to -\infty} \boldsymbol{F(x) = 0} \; e \; \lim_{x \to \infty} \boldsymbol{F(x) = 1}$$

- La funzione di ripartizione è continua a destra:

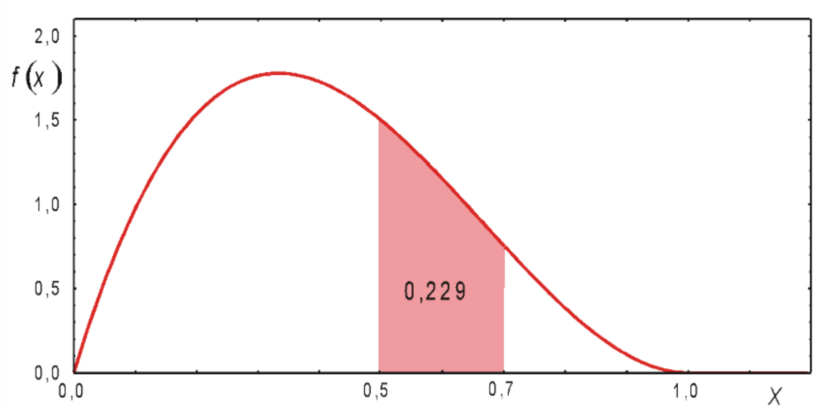
$$\lim_{x \to x_0^+} \boldsymbol{F(x) = F(x_0)}$$



VARIABILI CASUALI CONTINUE

Si definisce funzione di densità, la funzione matematica ***f(x)*** per cui l'area sottesa alla funzione, corrispondente ad un certo intervallo, è uguale alla probabilità che ***X*** assuma un valore a quell'intervallo.

Graficamente



- **F(x) ≥ 0** sempre
- L'area totale sottesa alla funzione è uguale ad 1:

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

- La probabilità che la v.c assuma un particolare valore dell'intervallo è zero.

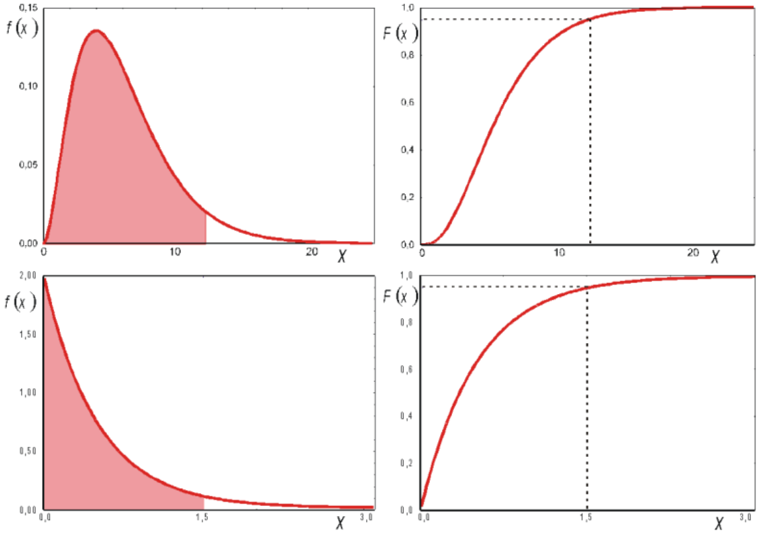
VARIABILI CASUALI CONTINUE

Funzione di Ripartizione

Data una ***v.c.*** continua ***X,*** la funzione che fa corrispondere ai valori ***x*** le probabilità cumulate ***P(X ≤ x)*** viene detta funzione di ripartizione:

$$\boldsymbol{F(x) = P(X \leq x) = \int_{-\propto}^{x} f(w)dw}$$

Esempio: funzione di densità di una *v.c* continua e corrispondente funzione di ripartizione



VALORE ATTESO E VARIANZA DI UNA V.C

Il valore medio di una v.c è definito come:

Se la v.c è discreta	Se la v.c è continua
$E(x) = \sum_i x_i P(x_i)$	$E(x) = \int_{-\infty}^{+\infty} x f(x) dx$

La varianza V(X) di una variabile casuale X è definita come:

Se la v.c è discreta	Se la v.c è continua
$V(x) = \sum_i (x_i - E(X))^2 P(x_i)$	$V(x) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x) dx$

Chiaramente, anche in questo caso è possibile ricavare la deviazione standard ricavabile semplicemente tramite:

$$SD(X) = \sqrt{V(X)}$$

V.C STANDARDIZZATA

I valori standardizzati esprimono la distanza tra I valori osservati e la media in termini di deviazione standard.

Se ***X*** è una ***v.c.*** con valore $E(X)$ e $SD(X)$ allora **Y** è una ***v.c.*** standardizzata con $E(Y)=0$ e $V(Y)=1$

$$Y = \frac{X - E(X)}{SD(X)}$$

VARIABILI CASUALI: LE DISTRIBUZIONI DI PROBABILITA’ (BERNOULLIANA-BINOMIALE, NORMALE, STUDENT, etc..)

La distribuzione di probabilità è un modello che associa una probabilità a ogni modalità osservabile di una <u>variabile aleatoria (o variabile casuale)</u>.

- <u>Distribuzione di probabilità discreta</u>: il fenomeno è osservabile con un numero intero di modalità, ad esempio il lancio del dado è un fenomeno statistico discreto perché il numero delle modalità osservabili è pari a 6. La <u>variabile aleatoria K</u> può assumere soltanto sei valori {1, 2, 3, 4, 5, 6}. Pertanto, la distribuzione di probabilità del fenomeno è discreta

- <u>Distribuzione di probabilità continua</u>: il fenomeno statistico è osservabile con un numero infinito o troppo elevato di modalità, ad esempio la distribuzione di probabilità della temperatura corporea è continua perché si tratta di un fenomeno statistico continuo, i valori della variabile aleatoria variano in modo continuo (es. 36.1°C, 36.2°C, ecc.).

DISTRIBUZIONE DI BERNOULLI/BINOMIALE

Eseguendo il lancio di una moneta sappiamo che la probabilità che possa uscire testa o croce è del 50%. Lanciando una seconda volta le probabilità rimangono invariate. Un tale evento, descritto da una variabile binaria (anche detta dicotomica) dove le probabilità di successo/insuccesso sono indipendenti dagli eventi temporalmente precedenti è detto evento di Bernoulli.

Una v.c. di Bernoulli può assumere il valore 1 con probabilità ***π*** e il valore 0 con probabilità ***1-π.*** La sua funzione di probabilità può essere espressa come:

$$\boldsymbol{P(X=x)=\pi^{x}\,(1-\pi)^{1-x};\,per\;x=0;1}$$

Tutte le prove che producono solo due possibili risultati generano ***v.c. di Bernoulli***: il lancio di una moneta, il sesso di un nascituro, il superamento o meno di un certo livello di inflazione. In alcuni casi però può essere utile chiedersi quale è la probabilità che su ***n*** di questi eventi, ***k*** abbiano un determinato esito. Per esempio, ci si può chiedere quale è la probabilità che lanciando 5 volte la moneta possa uscire testa in 2 occasioni. In questo caso, è necessario descrivere probabilisticamente non il singolo evento ma la serie dei 5 eventi. Una serie di eventi di Bernoulli di questo tipo è conosciuta come processo di Bernoulli ed è descritta da una distribuzione di probabilità per variabili discrete detta **binomiale.**

DISTRIBUZIONE DI BERNOULLI/BINOMIALE

Una ***v.c. Binomiale*** rappresenta il numero di successi che si presentano in una sequenza di n sotto prove indipendenti nelle quali è costante la probabilità di successo π, la sua funzione è definita come:

$$\boldsymbol{P(X) = \binom{n}{x} \pi^x (1-\pi)^{n-x}}$$

Considerando il coefficiente binomiale:

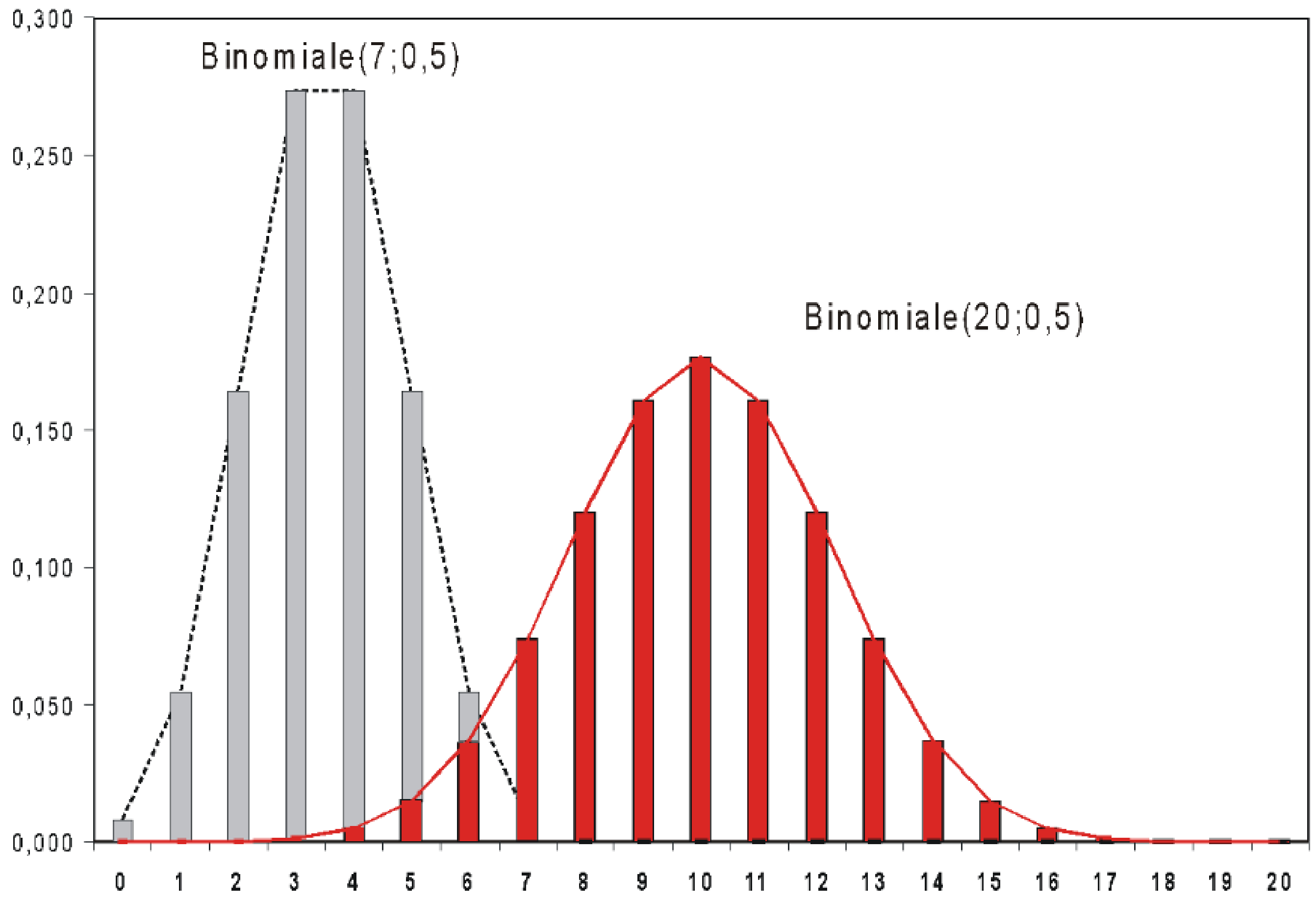
$$\boldsymbol{P(X) = \frac{n!}{x!\, n-x!} \pi^x (1-\pi)^{n-x}}$$

La distribuzione Binomiale può essere ottenuta considerando la somma di ***v.c. di Bernoulli***, indipendenti e identicamente distribuite:

- Media: $\boldsymbol{E(X) = E(X_1 + X_2 + \cdots + X_n) = \pi + \pi + \cdots + \pi = n\pi}$
- Varianza: $\boldsymbol{V(X) = V(X_1 + X_2 + \cdots + X_n) = \pi(1-\pi) + \pi(1-\pi) + \cdots + \pi(1-\pi) = n\pi(1-\pi)}$
- Il valore atteso e la varianza crescono al crescere di n
- Per ***π = 0.5*** la distribuzione è simmetrica rispetto al valor medio ***(n/2)***
- Per tendente a più infinito la distribuzione tende ad essere simmetrica rispetto al valor medio.

DISTRIBUZIONE DI BERNOULLI/BINOMIALE

Esempio di Binomiale



DISTRIBUZIONE DI POISSON

La distribuzione di Poisson descrive la probabilità che si verifichino ***n*** eventi di un tipo sapendo che in media si verificano con un ***rate λ*** in un intervallo (o area o lunghezza o altro). La sua formula è:

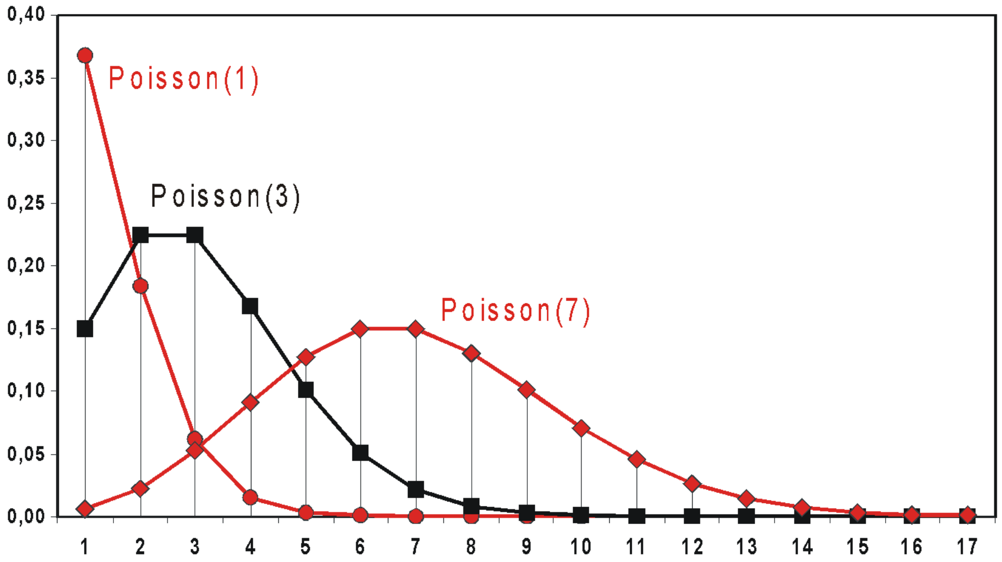
$$P(X) = \frac{\lambda^{x}}{x!} e^{-\lambda}; con\ x = 0; 1; 2; \dots e\ 0 < \lambda < \infty$$

- Media: $E(X) = \lambda$
- Varianza: $V(X) = \lambda$

La distribuzione di Poisson è conosciuta anche come ***distribuzione degli eventi rari*** in quanto approssima la più famosa distribuzione binomiale per probabilità molto basse di avvenimento.

DISTRIBUZIONE DI POISSON

Esempio



Come si può notare dal grafico, all'aumentare del valore λ la distribuzione di Poisson si abbassa, si allarga e si sposta verso valori più alti. Se quindi abbiamo a che fare con eventi con ***rate di accadimento*** molto alte, la distribuzione di Poisson aumenta la sua dispersione.

N.B: Gli eventi descritti dalla distribuzione di probabilità di Poisson, non necessariamente devono prevedere 2 possibili risultati (come per la binomiale). Il rate di accadimento deve invece essere costante nel l'intervallo di tempo (lunghezza, area, etc..) preso in considerazione e deve essere tale che due eventi non possono accadere allo stesso istante (o punto nello spazio).

DISTRIBUZIONE NORMALE (O GAUSSIANA)

La distribuzione normale o di Gauss è la più comune tra le distribuzione di probabilità per variabili continue. La sua popolarità è dovuta all’enorme quantità di fenomeni fisici, e non, descritti mediante l’utilizzo di tale distribuzione. Inoltre, secondo il teorema del limite centrale, la distribuzione di alcuni valori statistici come la media o la somma di campioni di valori non distribuiti normalmente è normale se il numero di campioni è molto grande.

La ***v.c. Normale X***, è dunque una v.c. continua che può assumere valori su tutto l’asse reale. La funzione di densità è data da:

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

- Con σ deviazione standard e μ contemporaneamente moda, media e mediana della distribuzione.
- Media: $E(X) = \mu$
- Varianza: $V(X) = \sigma^2$

DISTRIBUZIONE NORMALE (O GAUSSIANA)

La distribuzione normale ha le seguenti proprietà:

- Ogni trasformazione lineare di una ***v.c. Normale*** è ancora una **v.c. Normale**
- La somma di due ***v.c. Normali*** indipendenti è ancora una ***v.c. Normale*** con media e varianza pari, rispettivamente, alla somma delle medie e delle varianze delle due ***v.c. Normali***
- La variabile causale normale è simmetrica rispetto alla media $\boldsymbol{\mu}$
- Si tratta di una distribuzione unimodale in quanto la sua derivata prima (punto di massimo) è zero in corrispondenza di $\boldsymbol{\mu}$
- L'area della curva è pari a 1
- Presenta due punti di flesso in corrispondenza di $\boldsymbol{\mu} + \boldsymbol{\sigma}\ e\ \boldsymbol{\mu} - \boldsymbol{\sigma}$
- L'area della curva compresa tra determinati valori di sigma è:
  0,68 (68%) tra $\boldsymbol{\mu} - \boldsymbol{\sigma}\ e\ \boldsymbol{\mu} + \boldsymbol{\sigma}$
  0,95 (95%) tra $\boldsymbol{\mu} - \mathbf{2}\boldsymbol{\sigma}\ e\ \boldsymbol{\mu} + \mathbf{2}\boldsymbol{\sigma}$
  0,99 (99%) tra $\boldsymbol{\mu} - \mathbf{3}\boldsymbol{\sigma}\ e\ \boldsymbol{\mu} + \mathbf{3}\boldsymbol{\sigma}$

DISTRIBUZIONE NORMALE (O GAUSSIANA)

Distribuzione normale standardizzata

Si tratta di una particolare distribuzione normale caratterizzata da $\boldsymbol{\mu = 0}$ e $\boldsymbol{\sigma = 1}$.

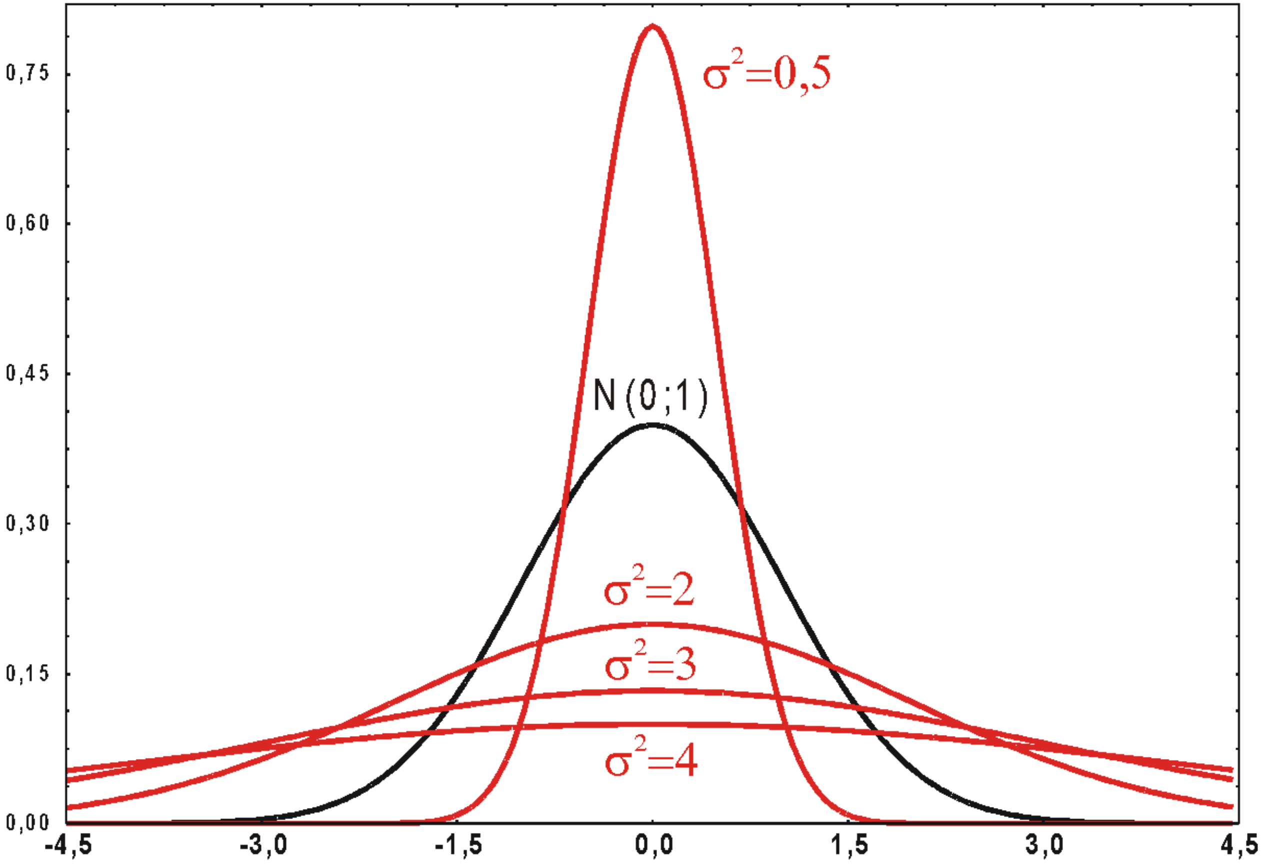
La distribuzione normale standardizzata risulta importante in quanto qualsiasi distribuzione Gaussiana può essere trasformata in essa. Se la v.c. X ha una distribuzione normale con parametri $\boldsymbol{\mu}$ ***e*** $\boldsymbol{\sigma^2}$, allora $\boldsymbol{Z = \frac{(x-\mu)}{\sigma}}$ è ancora una v.c. Normale con media nulla e varianza unitaria.

Essa è definita dalla funzione:

$$\boldsymbol{f(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}}$$

DISTRIBUZIONE NORMALE (O GAUSSIANA)

Esempio di distribuzione normale con media 0 e varianza 1



DISTRIBUZIONE NORMALE (O GAUSSIANA)

È infine possibile determinare la forma della distribuzione (lo spessore delle code ovvero il grado di appiattimento) per mezzo degli indici di Curtosi.

Indice di curtosi di Peason; $\boldsymbol{\gamma = E(\frac{X - E(X)}{SD(X)})^4}$

Indice di curtosi di Fisher = $\boldsymbol{\gamma - 3}$

In questo senso la ***funzione di Gauss*** può assumere tre formazioni:

- ***Leptocurtica*** per $\boldsymbol{\gamma > 0}$; distribuzione molto concentrata intorno alla sua media
- ***Normo-curtica* per** $\boldsymbol{\gamma = 0}$; distribuzione normalmente concentrata intorno alla sua media
- ***Platicurtica*** per $\boldsymbol{\gamma < 0}$; distribuzione poco concentrata intorno alla media

DISTRIBUZIONE CHI QUADRO

Tale distribuzione descrive l'andamento delle quadrati di v variabili aleatorie caratterizzate da una distribuzione di densità di probabilità normale standard con media 0 e varianza pari a 1.

$$\boldsymbol{X^2(v) = \sum_{i=1}^{v} x_i^2 = x_1^2 + x_2^2 + .. + x_v^2}$$

Tale distribuzione varia al variare del valore v detto anche numero gradi di libertà della funzione.

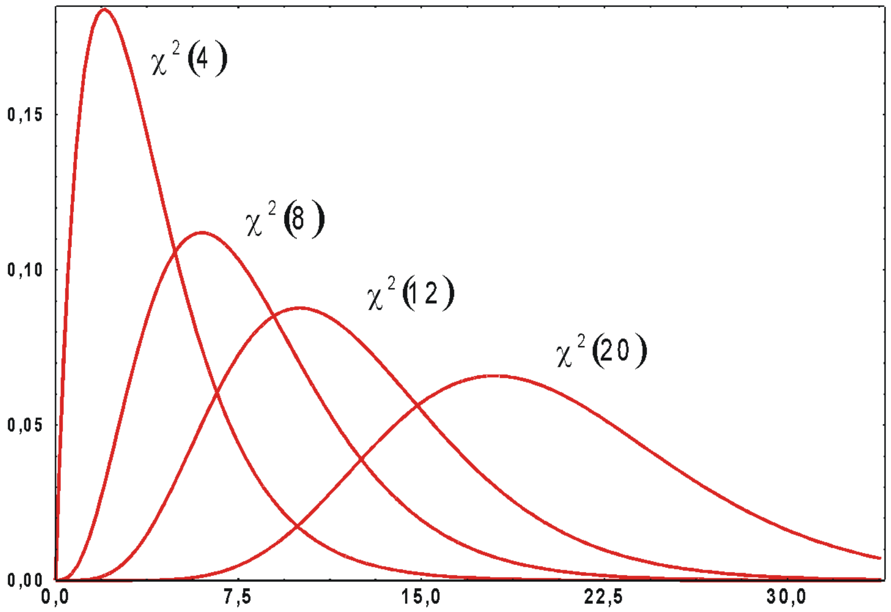
La v.c. Chi-quadrato è una distribuzione asimmetrica, continua e definita per valori reali non negativi. La funzione di densità dipende da un unico parametro, chiamato gradi di libertà, che è un intero positivo. Una generalizzazione della distribuzione $\boldsymbol{X^2}$ è la distribuzione Gamma $\boldsymbol{\tau(\frac{x}{2}, \frac{1}{2}) = X^2(x)}$

$$\boldsymbol{f(x) = \frac{1}{2^{\frac{v}{2}} \tau(\frac{v}{2})} \, x^{\frac{v}{2}-1} e^{\frac{-x}{2}}}$$

- Media: $\boldsymbol{E(X) = v}$
- Varianza: $\boldsymbol{V(X) = 2v}$

DISTRIBUZIONE CHI QUADRO

Esempio



- All'aumentare di $\boldsymbol{v}$ la distribuzione tende ad una normale (la distribuzione di densità di probabilità dipende dai gradi di libertà)
- Il massimo della funzione diminuisce e si sposta verso destra all'aumentare dei gradi di libertà

La distribuzione è asimmetrica

DISTRIBUZIONE CHI QUADRO

Utilizzo statistico del chi quadro:

Il test del chi quadro è un ***test di significatività statistica*** utilizzato per rifiutare o accettare un'ipotesi sulla distribuzione di probabilità di una o più variabili aleatorie (ipotesi nulla H_0).

La logica del test consiste nel confrontare le frequenze osservate di una variabile (o più variabili) aleatoria e le frequenze teoriche previste dalla distribuzione di probabilità ipotizzata.

Definita con H_0 l'ipotesi nulla e con H_1 la sua alternativa, il test del chi quadro consente di definire se le variazioni tra i dati osservati e quelli teorici siano dovute a oscillazioni casuali (ipotesi nulla accettata) oppure da altri fattori (ipotesi alternativa).

Approfondiremo meglio questi concetti nelle prossime unità didattiche

DISTRIBUZIONE T DI STUDENT

La distribuzione t di Student è una distribuzione di densità di probabilità di una variabile aleatoria continua definita come segue:

$$\boldsymbol{t} = \frac{\boldsymbol{Z}}{\sqrt{\frac{\boldsymbol{V}}{\boldsymbol{v}}}}$$

Con $\boldsymbol{Z}$ e $\boldsymbol{V}$ due variabili aventi distribuzioni di densità di probabilità indipendenti tra loro rispettivamente normale standard (media 0 e varianza 1) e chi quadro con $\boldsymbol{v}$ gradi di libertà.

- Come la distribuzione normale, la distribuzione t di Student ha una forma a campana ed è simmetrica rispetto allo 0.
- Diversamente dalla distribuzione normale, essa dipende dal numero di gradi di libertà (grandezza del campione – 1).
- In particolare, per un valore basso di gradi di libertà (campioni di dimensione molto piccola), la distribuzione risulta più bassa e più larga alle code di una distribuzione normale.
- All’aumentare dei gradi di libertà, la distribuzione diventa simile ad una distribuzione normale fino a coincidere da campione di 30/50 dati.

DISTRIBUZIONE T DI STUDENT

La formula della distribuzione t di Student è la seguente:

$$\boldsymbol{f(t) = \frac{1}{\sqrt{n}B(\frac{1}{2},\frac{n}{2})}(1+\frac{t^2}{n})\frac{-(n+1)}{2}}$$

Esempi di applicazione sono test delle ipotesi su:

- Media di una popolazione di cui non è conosciuta la deviazione standard della popolazione e calcolo dell’intervallo di confidenza.
- Differenza tra due medie
- Differenza tra due medie di 2 campioni dipendenti
- Coefficiente di correlazione.