FACOLTA' DI INGEGNERIA

CORSO DI LAUREA IN INGEGNERIA CIVILE

CORSO DI IDROLOGIA

PROF. PASQUALE VERSACE

SCHEDA DIDATTICA N°1

ARGOMENTO:

CALCOLO DELLE PROBABILITA' E VARIABILI CASUALI

A.A. 2010-11

Calcolo delle probabilità e statistica

La descrizione dei fenomeni naturali è uno dei principali obiettivi della ricerca scientifica. In particolare risulta essenziale la realizzazione di modelli matematici che simulino in modo efficace il fenomeno analizzato. L'approccio deterministico prevede la costruzione di modelli matematici che descrivono in maniera precisa il mondo reale. Esistono le leggi fisiche e matematiche le quali, una volta definite (anche sperimentalmente) le grandezze in ingresso ed i parametri, consentono di descrivere, in maniera sufficientemente precisa, il comportamento di fenomeni più o meno complessi: ad esempio permettono di stabilire la traiettoria di un satellite, di determinare la legge di caduta di un grave, etc..

Altri eventi, per i quali le variabili in gioco sono molte e non tutte note, sono estremamente difficili da modellare matematicamente, e l'approccio deterministico non è sempre percorribile. L'esempio classico è quello dei giochi di azzardo che, come suggerisce il nome, consistono in azioni quali far girare la ruota di una roulette, gettare i dadi, lanciare una moneta, estrarre una carta, estrarre da un'urna una pallina colorata o numerata, etc., per le quali l'esito è incerto.

In tal caso, per sopperire alla nostra ignoranza sulle modalità di accadimento degli eventi reali, si ricorre alla formulazione di un modello probabilistico, attribuendo ad elementi casuali ciò che in realtà non è pienamente spiegato e compreso.

Un modello di probabilità non consente la previsione esatta del fenomeno, ma consente di valutare la probabilità con cui si può verificare un certo evento. Il risultato di un singolo tentativo è incerto, ma può essere previsto meglio il risultato complessivo di più prove.

Ad esempio, su numerosi lanci di una moneta regolare (bilanciata e simmetrica) è intuitivo stabilire che circa metà dei tentativi darà testa.

Se si lancia un dado, esistono 6 possibili risultati, che si escludono a vicenda, ugualmente probabili; ci si aspetta, cioè, che, effettuando un gran numero di lanci, ogni faccia apparirà all'incirca lo stesso numero di volte, pari a circa 1/6 dei lanci effettuati.

Negli esempi considerati, la stima della probabilità è stata effettuata attraverso il rapporto, che risulta minore o al più uguale ad 1, tra i casi favorevoli, n_A, ed il numero totale di casi possibili n. Al risultato si giunge per mezzo del seguente ragionamento **deduttivo** (non è stato necessario effettuare alcun esperimento): poiché ci sono due modi possibili con cui la moneta può cadere, se questa è regolare e ben bilanciata è ragionevole aspettarsi che in ciascun lancio la probabilità di osservare testa sia 1/2.

L'applicazione di quest'approccio è abbastanza immediata in casi semplici, ma non sempre è così ovvia: è di difficile attuazione quando il numero di possibili esiti è infinito, o ancora quando gli eventi possibili non sono ugualmente probabili (una moneta o un dado non perfettamente bilanciato) o non si escludono a vicenda.

In tal caso risulta indispensabile ricorrere ad un approccio alternativo nella definizione della probabilità che è quello **induttivo**. Questo secondo approccio (indicato anche come frequentista) si basa sull'analisi di uno o più campioni di dati: la probabilità di un evento è, infatti, approssimata con la frequenza osservata, ed è pertanto indispensabile effettuare numerosi esperimenti o disporre di molte osservazioni. Il risultato ottenuto sulla base del campione è poi, in genere, considerato rappresentativo del fenomeno considerato.

Alcuni esempi di applicazione dell'approccio deduttivo a casi più complessi sono di seguito illustrati.

1. Si consideri ora la probabilità di ottenere due volte testa se si lancia due volte una moneta. I risultati possibili sono quattro: {TT, CC, TC, CT}. La probabilità corretta è

perciò pari ad 1/4. Lanciando tre volte la moneta, la probabilità di ottenere tre volte testa è 1/8; infatti i casi possibili sono {TTT, TTC, TCT, TCC, CTT, CTC, CCT, CCC}. Ripetendo il lancio 4 volte la probabilità di avere sempre testa si riduce ad 1/16, essendo possibili i seguenti casi {TTTT, TTTC, TTCT, TTCC, TCTT, TCTC, TCCT, TCCC, CTTT, CTTC, CTCT, CTCC, CCTT, CCTC, CCCT, CCCC }

2. Utilizzando gli stessi dati di cui al punto 1 è possibile calcolare le probabilità di avere esattamente due volte croce in 2, 3 e 4 lanci di una moneta. Si ottengono, come è facile controllare, i seguenti valori:

n° lanci della moneta	Probabilità di avere due volte croce
2	1/4
3	3/8
4	6/16

3. La probabilità di estrarre a caso una determinata carta da un mazzo di carte preparate per giocare a poker (dal quale cioè sono state eliminate le carte con valore 2, 3, 4, 5 e 6) è pari a 1/32. Si supponga di voler stimare la probabilità di fare poker avendo in mano 3 assi, e scartando le altre due carte. La probabilità che la prima carta estratta sia asso è pari a $\frac{1}{(32-5)}=\frac{1}{27}$; infatti il giocatore conosce già 5 delle 32 carte del mazzo e sa che nelle altre 27 ci sarà un solo asso: il quarto. Ovviamente la probabilità che la prima carta estratta non sia asso sarà pari a $\frac{26}{27}$. Se la prima carta pescata non è asso, allora restano 26 carte e un solo asso, quindi la probabilità che l'asso sia la seconda carta estratta è $\frac{1}{26}$, mentre la probabilità che non lo sia risulta $\frac{25}{26}$. Se, invece, la prima carta è asso, allora la probabilità che lo sia la seconda sarà pari a 0 mentre che non lo sia sarà pari a $\frac{26}{26}=1$. La seconda estrazione è dipendente dalla prima, in quanto, non reinserendo la prima carta nel mazzo, viene modificato il numero di casi possibili (da 27 si passa a 26) e di conseguenza la probabilità di uscita dell'asso. Se invece la prima carta venisse reinserita, l'uscita dell'asso alla seconda estrazione avrebbe sempre probabilità pari a $\frac{1}{27}$. Alla luce di tali considerazioni, le due estrazioni si dicono *dipendenti* e, come meglio specificato nelle pagine successive, le probabilità associate ad ognuna si moltiplicano, ovvero **la probabilità che la seconda carta estratta sia asso, se la prima carta non lo è, risulta** $\frac{26}{27}\cdot\frac{1}{26}=\frac{1}{27}$. A ciò va sommata la probabilità che l'asso esca alla prima estrazione (e di conseguenza alla seconda estrazione non potrà più uscire), ovvero $\frac{1}{27}\cdot\frac{26}{26}=\frac{1}{27}$**. Complessivamente, la probabilità di uscita dell'asso da due estrazioni senza reimmissione della prima carta risulta:**

$$\frac{1}{27}+\frac{1}{27}=\frac{2}{27}=0.0741$$

4. Ancora, sempre facendo riferimento al gioco del poker, si supponga di avere un 8, un 9, un Jack ed una Regina, e si stimi la probabilità di fare una scala ad incastro, cioè di estrarre un 10 dal rimanente mazzo di carte. Tale probabilità vale:

$$\frac{4}{27} = 0.148 .$$

5. Si consideri, infine, un ultimo esempio relativo alla classica schedina. La probabilità di indovinare l'esito (1, X, 2) di una singola partita è, in questo caso, pari ad 1/3. La probabilità che il pronostico per due partite sia corretto, ha ovviamente una probabilità più bassa. Infatti, le possibili combinazioni dei risultati sono 9, cioè {11, 1X, 12, X1, XX, X2, 21, 2X, 22}, e, pertanto, la probabilità ricercata risulta $1/9=1/3^2$. Applicando lo stesso ragionamento è possibile stimare la probabilità di indovinare i risultati di tre partite come $1/27=1/3^3$. La probabilità di fare 13, quindi, è $1/3^{13}=1/1.594.323$.
   In realtà, questo è un esempio non del tutto corretto perché gli esiti di ciascuna possibile combinazione non sono realmente *equiprobabili* (in quasi tutte le partite ci sono alcuni risultati che per la forza delle squadre in campo sono più probabili di altri).

   Altri due esempi tra loro collegati possono essere utili a capire la differenza tra l'approccio deduttivo e quello induttivo.

A. Supponiamo di avere un'urna con 10000 palline colorate: 9000 sono bianche e 1000 sono nere. È facile capire che 90 volte su cento prenderò una pallina bianca e 10 volte su cento una pallina nera. Si può facilmente calcolare la probabilità di prendere in due estrazioni successive due palline nere, o due bianche, o prima una nera e poi una bianca, e così via. Il fatto di partenza è che conosco la composizione dell'urna.

B. Supponiamo di avere un'urna con miliardi di palline bianche o nere. Ma non sappiamo quante sono le bianche e quante sono le nere. Sono troppe per osservarle una ad una.
   Se estraiamo una sola pallina ed esce nera potremmo pensare che l'urna contiene solo palline nere. Se ne estraiamo due e sono una bianca e una nera potremmo ritenere che nell'urna ci sono metà palline bianche e metà palline nere. Ovviamente più palline estraiamo e più capiamo la reale composizione dell'urna. Per cui se estraiamo 10000 palline e vediamo che 6000 sono bianche e 4000 nere, possiamo ritenere con buona precisione che nell'urna c'è il 60% di palline bianche. E che quindi su 100 palline ce ne aspettiamo 60 bianche e 40 nere. Abbiamo raggiunto questa conclusione non in base ad una deduzione basata sull'esatta conoscenza del contenuto dell'urna, come nel caso A, ma con la analisi di alcuni *dati campionari* provenienti dall'urna. Si è cioè utilizzata una procedura induttiva che va dal particolare (l'osservazione) al generale (il contenuto dell'urna). Qualche cosa di simile avviene in occasione delle elezioni con le proiezioni degli istituti specializzati.

CONCETTI ELEMENTARI DEL CALCOLO DELLE PROBABILITÀ

L'impostazione seguita in questa breve introduzione alla teoria della probabilità si basa su un approccio 'assiomatico' che prevede concetti elementari, postulati e teoremi. I concetti elementari del calcolo delle probabilità sono:
Prova – è un esperimento soggetto ad incertezza.
Evento – è uno dei possibili risultati della prova.
Probabilità – è un numero associato al presentarsi di un evento.

I tre concetti primitivi e la loro reciproca relazione sono ben illustrati dalla frase:

"*La prova genera l'evento con una specifica probabilità*".

La totalità dei possibili risultati di una prova è indicato come **spazio campione**. L'evento pertanto è un sottoinsieme dello spazio campione.

Esempi
Prova = lancio di un dado
Eventi = $\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}$
Probabilità di ognuno dei sei possibili eventi = 1/6

Prova = lancio di due dadi
Eventi = $\{1,1\},\{1,2\} \ldots,\{6,6\}$
Probabilità di ognuno dei 36 possibili eventi = 1/36

ALGEBRA DEGLI EVENTI

Si indichino con *A,B,C,D,E* gli eventi risultati di una prova ben definita. Tra essi sono definite le seguenti tre operazioni principali:

UNIONE	$C = A \cup B$	L'evento *C* si verifica quando si verifica *A* oppure *B*
INTERSEZIONE	$D = A \cap B$	L'evento *D* si verifica quando si verificano sia *A* che *B*
NEGAZIONE	$E = \overline{A}$	L'evento *E* si verifica quando non si verifica *A*

L'**evento certo** = I è l'evento che si verifica sempre.
L'**evento impossibile** = $\varnothing$ è l'evento che non si verifica mai.

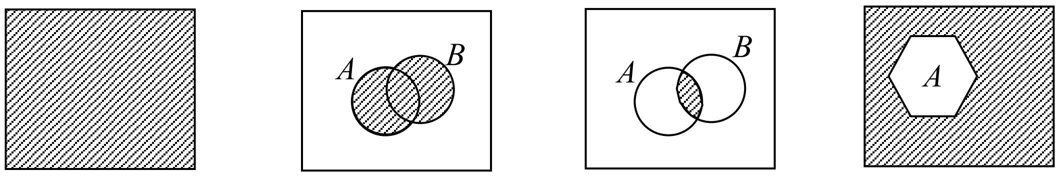
Due eventi A e B tali che non possono verificarsi contemporaneamente vengono definiti **incompatibili**, ovvero:

A e B sono eventi incompatibili se $A \cap B = \varnothing$

Mentre due eventi C e D tali che almeno uno di essi deve verificarsi, sono definiti **necessari**:

C e D sono eventi necessari se $C \cup D = I$

Le relazioni dell'algebra degli eventi, sono spesso illustrate su un piano mediante dei grafici, definiti **diagrammi di Venn**, nei quali lo spazio campione è rappresentato come un quadrangolo all'interno del quale gli eventi sono rappresentati come insiemi chiusi. Di seguito sono riportati alcuni esempi di diagrammi di Venn che descrivono alcune relazioni tra eventi.



Spazio campione I | Unione $A \cup B$ | Intersezione $A \cap B$ | Negazione di A = $\overline{A}$

Osservazioni

La superficie occupata dall'evento certo I coincide con quella dello spazio campione. In generale se una prova consiste di k eventi E_1, E_2, …, E_k allora, l'evento certo coincide con $I = E_1 \cup E_2 \cup \ldots \cup E_k$.

POSTULATI DEL CALCOLO DELLE PROBABILITÀ

I postulati del calcolo delle probabilità, proprio perché tali, non sono dimostrabili, ma rappresentano le basi sulle quali si sviluppa la teoria. Essi rappresentano la formalizzazione di convincimenti legati all'osservazione empirica.

Nel seguito si indica con $P(A)$ la probabilità che si verifichi l'evento A.

1. La probabilità $P(A)$ di un evento A è un numero non negativo: $P(A) \geq 0$;
2. La probabilità $P(I)$ di un evento certo I è sempre uno: $P(I) = 1$;
3. Se A e B sono incompatibili la probabilità della loro unione è la somma delle probabilità dei singoli eventi:
   se $A \cap B = \varnothing \Rightarrow P(A \cup B) = P(A)+P(B)$
4. La probabilità dell'evento B dato che si è verificato l'evento A, $P(B/A)$, è pari al rapporto tra le probabilità che si verifichino contemporaneamente A e B e la probabilità di A, se questa è diversa da zero:
   se $P(A) \neq 0 \Rightarrow P(B/A) = \dfrac{P(A \cap B)}{P(A)}$

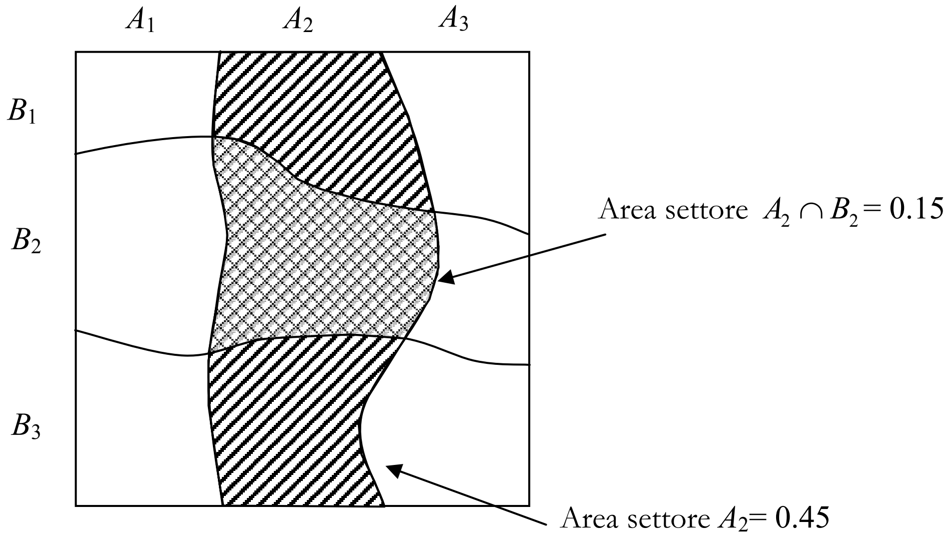
Osservazioni

- La probabilità che si verifichi un evento è un numero (compreso tra 0 e 1) e, quindi, nei calcoli si utilizzano le usuali operazioni di algebra (+, -, *, /). Gli eventi invece non sono identificati da un numero, ma possono essere caratterizzati da una descrizione (esempio: esce testa). Pertanto, come illustrato in precedenza, le operazioni tra gli eventi utilizzano un'algebra specifica.

- Due eventi A e B sono indipendenti quando il verificarsi di A non altera le probabilità di verificarsi di B. Si può, quindi, scrivere $P(B/A) = P(B)$
- Dall'osservazione precedente e dal quarto postulato si evince che la probabilità dell'intersezione di eventi indipendenti è uguale al prodotto delle probabilità dei singoli eventi:

$$P(B/A) = P(B) = \frac{P(A \cap B)}{P(A)} \Rightarrow P(A \cap B) = P(A) \cdot P(B).$$

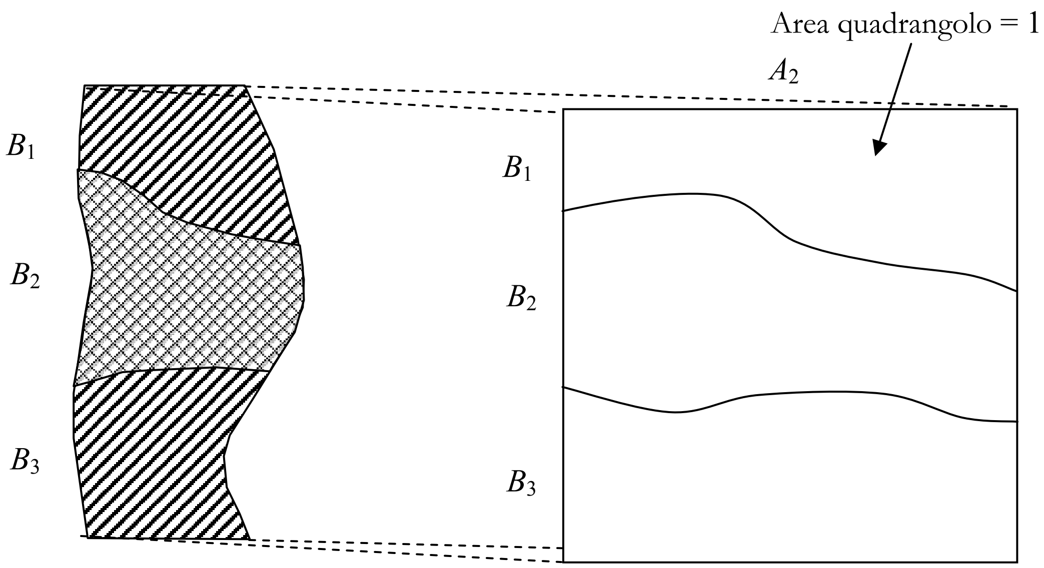
Per capire meglio il senso del quarto postulato può essere utile il seguente esempio.
Lo spazio campione della figura sottostante, complessivamente di area unitaria, è ripartito in settori, verticali rispetto agli eventi A_1, A_2, A_3 e orizzontali per B_1, B_2, B_3, con aree proporzionali alla probabilità di accadimento dell'evento. Gli eventi A_i e B_j non sono indipendenti. A seconda dell'area occupata dai diversi elementi si può calcolare la probabilità di un evento sia esso un evento singolo tipo A_2 (area tratteggiata) oppure un evento composto tipo $A_2 \cap B_2$ (area a doppio tratteggio).



Il quarto postulato in pratica ci fa capire che una volta che si sia verificato A_2 la probabilità che si verifichi anche B_2 è data dal rapporto tra $P(A_2 \cap B_2)$, che rappresenta i casi favorevoli, e $P(A_2)$ che rappresenta i casi possibili. Pertanto

$$P(B_2/A_2) = P(A_2 \cap B_2)/P(A_2) = 0.15/\ 0.45 = 0.333$$

In altri termini è come se al diagramma di Venn ne avessimo sostituito un altro che contiene non più tutti i nove casi all'inizio possibili, ma solo i tre che dopo il verificarsi di A_2 sono ancora possibili.



La quantità $P(A_2)$ al denominatore permette di ristabilire le proporzioni, assicurando la normalizzazione $[P(B_1/A_2) + P(B_2/A_2) + P(B_3/A_2) = 1]$.

Esempio

Si consideri un'urna con 1000 palline, così composta: 600 palline gialle, 200 palline rosse e 200 palline verdi. All'interno di ciascuna pallina si trova un bigliettino che può essere contrassegnato dalla lettera α, β o γ. Complessivamente ci sono 400 bigliettini α, 100 bigliettini β e 500 bigliettini γ. La tabella sottostante illustra le diverse possibili combinazioni tra i colori delle palline ed i bigliettini in esse contenute.

	Giallo G	Rosso R	Verde V	
α	300	50	50	400
β	50	10	40	100
γ	250	140	110	500
	600	200	200	1000

Da quanto detto, ad esempio, la probabilità di estrarre una pallina gialla contenente un bigliettino α è:

$$P(G \cap \alpha) = 300/1000 = 0.3 \quad (30\%)$$

Analogamente, la probabilità di estrarre una pallina verde contenente un bigliettino β è:

$$P(V \cap \beta) = 40/1000 = 0.04 \quad (4\%).$$

Prendiamo ora dall'urna una pallina gialla, la probabilità di estrarre tale pallina è:

$P(G) = 600/1000 = 0.6$

La probabilità che avendo già estratto una pallina gialla, al suo interno si trovi un bigliettino contrassegnato con la lettera α è, invece:

$$P(\alpha / G) = 300/600 = 0.5$$

Allo stesso risultato si giunge applicando il quarto postulato:

$$P(\alpha / G) = P(G \cap \alpha) / P(G) = 0.3/0.6 = 0.5.$$

TEOREMI DEL CALCOLO DELLE PROBABILITÀ

I teoremi del calcolo delle probabilità non saranno dimostrati, ma possono essere facilmente dedotti dai postulati.

1. $P(\emptyset) = 0$
2. $P(\overline{A}) = 1 - P(A)$
3. $P(\overline{A} \cap B) = P(B) - P(A \cap B)$
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Osservazioni

L'ultimo teorema generalizza al caso di eventi non incompatibili, la probabilità associata all'unione di due eventi definita nel postulato 3.

Gli effetti del considerare eventi incompatibili o indipendenti sulle probabilità delle unioni ed intersezioni di tali eventi, sono di seguito sintetizzati:

1. eventi incompatibili ($A \cap B = \emptyset$) $\Rightarrow$ $P(A \cup B) = P(A) + P(B)$
2. eventi non incompatibili ($A \cap B \neq \emptyset$) $\Rightarrow$ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
3. eventi dipendenti $\Rightarrow$ $P(A \cap B) = P(A) \cdot P(B / A)$
4. eventi indipendenti $\Rightarrow$ $P(A \cap B) = P(A) \cdot P(B)$

Un ulteriore esempio che può contribuire a chiarire la differenza nella trattazione di eventi dipendenti ed eventi indipendenti è riportato di seguito.

Si consideri l'estrazione di due palline da un'urna contenente dieci palline nere e cinque bianche.

Si vuole calcolare la probabilità che esca una pallina nera in entrambe le estrazioni. Indicati con A_1 e A_2 gli eventi "esce una pallina Nera", rispettivamente, alla prima e alla seconda estrazione, l'evento ricercato è, quindi, $A_1 \cap A_2$.

Se le singole estrazioni avvengono con conseguente reinserimento della pallina estratta nell'urna, gli eventi sono indipendenti, e si ha:

$$P(A_1) = 10/15$$

$$P(A_2/A_1) = 10/15 = P(A_2)$$

e, quindi,

$$P(A_1 \cap A_2) = P(A_1)\, P(A_2) = 4/9 = 0.444.$$

Se, invece, le estrazioni non prevedono il reinserimento della pallina nell'urna, gli eventi non sono più indipendenti. La seconda estrazione dipende da quanto si è verificato nella prima. Nello specifico la probabilità di avere una seconda pallina nera estratta risulta:

$$P(A_2/A_1) = 9/14,$$

e la probabilità cercata è

$$P(A_1 \cap A_2) = P(A_1)\, P(A_2/A_1) = 6/14 = 0.429.$$

Sempre dalla stessa urna, consideriamo ora la probabilità di estrarre una pallina nera alla prima prova ed una bianca alla seconda. Con la notazione introdotta si ricerca la probabilità di avere $A_1 \cap \overline{A_2}$.

Nel caso di prova con reinserimento si ha:

$$P(A_1) = 10/15$$

$$P(\overline{A_2}/A_1) = 5/15 = P(\overline{A_2})$$

e, quindi,

$$P(A_1 \cap \overline{A_2}) = P(A_1)\, P(\overline{A_2}) = 2/9 = 0.222.$$

Effettuando la prova prevedendo il reinserimento della pallina nell'urna, gli eventi sono dipendenti, e quindi si ha:

$$P(A_1) = 10/15$$

$$P(\overline{A_2}/A_1) = 5/14$$

$$P(A_1 \cap \overline{A_2}) = P(A_1)\, P(\overline{A_2}/A_1) = 5/21 = 0.238.$$

Esercizi svolti

1. Si effettua il lancio di un dado. Calcolare:

a) la probabilità che esca un 2 oppure un 5;

b) la probabilità che esca un numero pari;

c) la probabilità che esca un numero divisibile per 3;

d) dati gli eventi:

- A_1= "esce 1 oppure 2", che si può anche indicare come A_1= {l, 2}
- A_2= "esce 2 oppure 3", che si può anche indicare come A_2= {2, 3}

calcolare $P(A_1 \cup A_2)$

<u>Svolgimento</u>

a) l'evento che si verifica quando esce un 2 o un 5 si indica con $2 \cup 5$; sapendo che $P(1) = P(2) = ... = P(6) = \frac{1}{6}$ si ha:

$$P(2 \cup 5) = P(2) + P(5) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

b) $P(2 \cup 4 \cup 6) = P(2) + P(4) + P(6) = \frac{1}{2}$

c) $P(3 \cup 6) = P(3) + P(6) = \frac{1}{3}$

d) gli eventi A_1 e A_2 sono compatibili, poiché $A_1 \cap A_2$={2} $\neq \emptyset$; di conseguenza si ha:

$A_1 \cup A_2$={1, 2, 3}

$$P(A_1) = P(A_2) = \frac{1}{3}$$

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = \frac{1}{3} + \frac{1}{3} - \frac{1}{6} = \frac{1}{2}$$

2. Si estrae una carta a caso da un mazzo di 52 carte. Calcolare la probabilità che sia:

a) un asso;

b) un fante di cuori;

c) un 3 di picche o un 6 di fiori;

d) una carta di cuori;

e) un seme diverso da cuori;

f) un 10 oppure una carta di quadri;

g) né un 4 né una carta di picche.

<u>Svolgimento</u>

Si usano le seguenti notazioni:

1=asso, …, 11=fante, 12=regina, 13=re,

C=cuori, Q=quadri, P=picche, F=fiori.

a) $P(1)=\frac{4}{52}$

b) $P(11\cap C)=\frac{1}{52}$

c) $P((13\cap P)\cup(6\cap F))=P(13\cap P)+(6\cap F)=\frac{1}{52}+\frac{1}{52}=\frac{1}{26}$

d) $P(C)=\frac{13}{52}=\frac{1}{4}$

e) $P(\overline{F})=1-P(F)=1-\frac{1}{4}=\frac{3}{4}$

f) una carta 10 ed una carta di quadri sono eventi compatibili, quindi:

$$P(10\cup Q)=P(10)+P(Q)-P(10\cap Q)=\frac{1}{13}+\frac{1}{4}-\frac{1}{52}=\frac{4}{13}$$

g) l'evento è la negazione di " uscita di un 4 oppure di una carta di picche", quindi:

$$P\left(\overline{(4\cup P)}\right)=1-P(4\cup P)=1-[P(4)+P(P)-P(4\cap P)]=1-\left[\frac{1}{13}+\frac{1}{4}-\frac{1}{52}\right]=\frac{9}{13}$$

3. Si lancia due volte un dado. Calcolare la probabilità di ottenere 4, 5 o 6 al primo lancio e 1, 2, 3 o 4 al secondo.

Siano

$A=\{4,5,6\}$ e $B=\{1,2,3,4\}$

Si deve calcolare la probabilità $P(A\cap B)$.

<u>Svolgimento</u>

Il risultato del secondo lancio è indipendente dal primo, cioè i due eventi A e B sono indipendenti, perciò:

$$P(A \cap B) = P(A) \cdot P(B) = \frac{3}{6} \cdot \frac{4}{6} = \frac{1}{3}$$

4. Trovare la probabilità che in due lanci di un dado si presenti almeno una volta il 5.

Svolgimento

L'evento A="uscita del 5 almeno una volta su due lanci" costituisce la negazione dell'evento E="la faccia 5 non esce mai in due lanci", da cui:

$$P(A) = 1 - P(E) = 1 - \frac{5}{6} \cdot \frac{5}{6} = \frac{11}{36}$$

5. Si effettua un lancio di un dado; si considerino i seguenti eventi:

A="esce un numero dispari"={1, 3, 5}

B="esce un numero minore di 4"= {1, 2, 3}.

Si calcoli la probabilità di ottenere un numero minore di 4, sapendo che il risultato è un numero dispari

Svolgimento

La probabilità dell'evento A vale $P(A) = \frac{1}{2}$; inoltre $A \cap B = \{1,3\}$ con $P(A \cap B) = \frac{2}{6} = \frac{1}{3}$, da cui:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{2}{3}$$

Eventi e variabili casuali

Finora si è parlato di eventi come risultato di una prova, ed al verificarsi di ciascun evento si è associato un valore di probabilità dotato delle proprietà analizzate in precedenza. Spesso è utile poter 'quantificare' gli eventi, cioè creare una corrispondenza tra i risultati della prova ed un insieme (o un sottoinsieme) di numeri reali.

Una **variabile casuale** può essere definita come una regola che associa ad ogni evento un numero reale, ovvero, in maniera più rigorosa, una v.c. è una funzione reale e misurabile definita sullo spazio campione.
Considerato l'evento E che si verifica con probabilità $P(E)$, se ad esso associamo un determinato valore x_E della variabile casuale X, le seguenti due affermazioni risultano equivalenti:

L'evento E si verifica con probabilità P(E)

La v.c. X assume il valore x_E con probabilità P(E)

Ad esempio, considerando la prova lancio di un dado, la v.c. può essere definita facendo corrispondere ordinatamente ai sei possibili risultati dell'esperimento l'insieme dei numeri reali $\{1, 2, 3, 4, 5, 6\}$. Naturalmente la v.c. può assumere ciascuno di questi valori con probabilità pari ad 1/6.

EVENTO	VARIABILE CASUALE
Esce 1	1
Esce 2	2
Esce 3	3
Esce 4	4
Esce 5	5
Esce 6	6

Ancora, considerando l'esperimento lancio di una sola moneta, la v.c. X associata all'esito della prova può essere così definita:

$$\begin{cases} X = 1 & se\ E = testa \\ X = 0 & se\ E = croce \end{cases}$$

In questo caso ad ognuno dei due possibili valori della v.c è associata una probabilità pari a 1/2.

Variabili casuali discrete

Una variabile casuale è detta discreta se l'insieme dei valori che può assumere è numerabile. Per una variabile X discreta, che può assumere un insieme numerabile di valori $x_1, x_2, \ldots, x_k$, si definisce la **funzione di densità discreta** (o di probabilità di massa), $P(X = x)$, come la funzione che associa a ciascun valore x della v.c. la corrispondente probabilità.

L'esempio precedente della v.c. associata alla prova lancio di un dado rappresenta una chiara dimostrazione di v.c. discreta, la cui funzione di densità vale:

$$\begin{cases} P(X=x)=1/6 & se\ x=1,2,3,4,5,6 \\ P(X=x)=0 & altrimenti \end{cases}$$

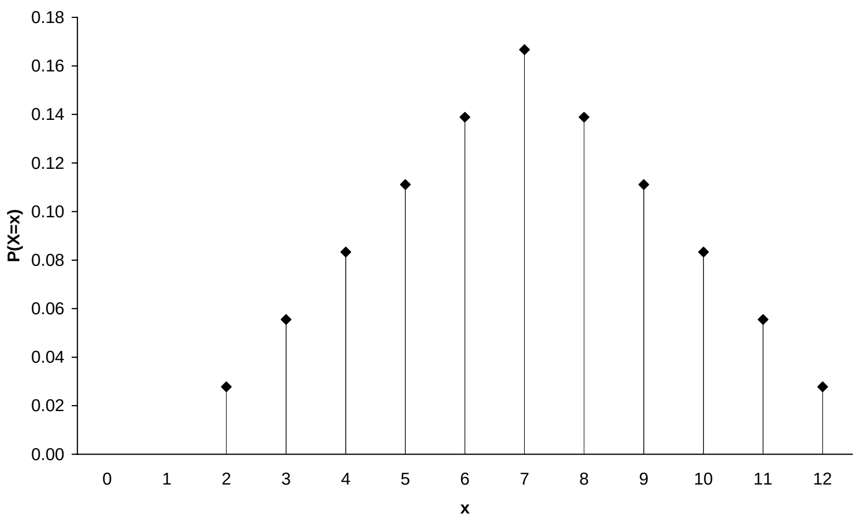
In accordo con i postulati, condizioni necessarie e sufficienti affinchè una funzione possa rappresentare una funzione di densità discreta sono:

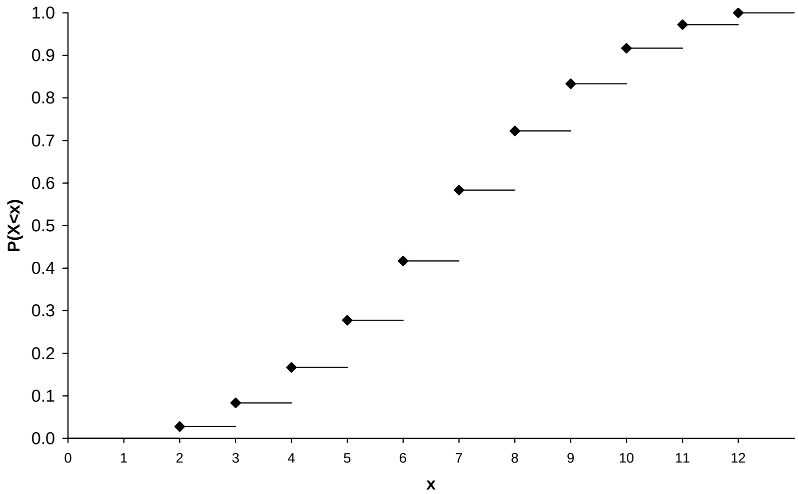
1. $P(X=x)\geq 0$ se $x = x_1, x_2, \ldots, x_k$
2. $\sum_{i=1}^{k} P(X=x_i)=1$

Si definisce, inoltre, **funzione di ripartizione** o funzione cumulativa di frequenze di una v.c. discreta, la funzione definita sullo spazio campione come $P(X \leq x)$.
La funzione di ripartizione gode delle seguenti proprietà:

1. $\lim_{x\to-\infty} P(X\leq x)=0$; $\lim_{x\to+\infty} P(X\leq x)=1$;
2. è una funzione monotona non decrescente $\Rightarrow$ se $a < b$, $P(X \leq a) \leq P(X \leq b)$;

Esempio
Consideriamo l'esempio del lancio di due dadi. Sia la v.c. X il totale delle facce verso l'alto. I possibili valori che X può assumere sono 2,3,4,5,6,7,8,9,10,11,12. Nelle figure che seguono sono rappresentate rispettivamente la funzione di densità discreta e la funzione di ripartizione. La prima indica la probabilità di occorrenza di ciascun valore della v.c. (es. P(X=2)=1/36; P(X =3)=2/36,…), la seconda la probabilità cumulata, ossia $P(X \leq x)$ (es. $P(X \leq 3) = 3/36$).





Variabili casuali continue

Una variabile casuale è continua se può assumere valori in un insieme continuo che appartiene ad un intervallo reale (anche l'intera retta reale tra -∞ e +∞). A differenza di una v.c. discreta, una v.c. continua X può assumere un'infinità non numerabile di valori e, pertanto, non è possibile elencare tali valori ed attribuire a ciascuno di essi la probabilità corrispondente.

In maniera analoga a quanto visto per le v.c. discrete, a ciascun punto dell'intervallo (a,b) in cui è definita la v.c. X si associa una funzione matematica, $f_X(x)$, denominata **funzione di densità di probabilità** (spesso indicata in forma abbreviata come **fd** o **pdf**), così definita:

$$f_X(x)dx = P[x \leq X \leq x + dx] \qquad x \in (a,b)$$

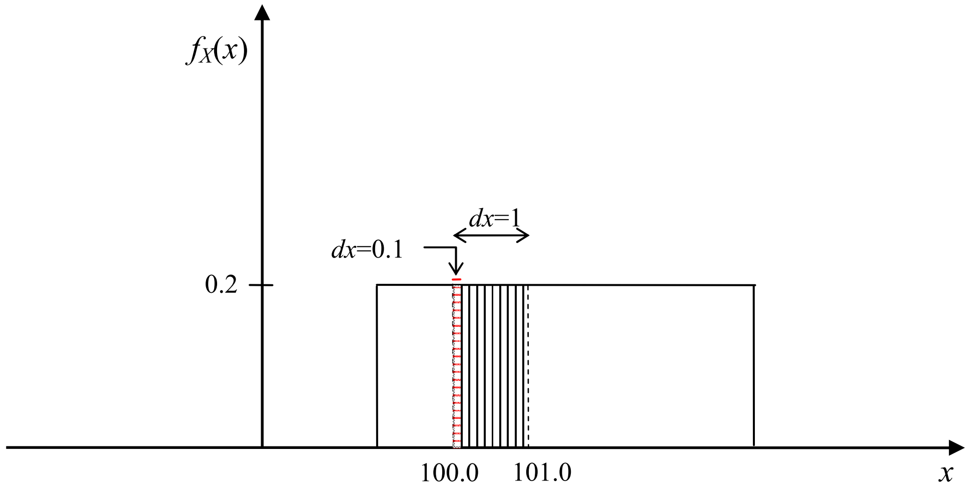
che è proporzionale alla probabilità di un intervallino 'sufficientemente piccolo'.
Si noti che la funzione di densità di probabilità $f_X(x)$ di una variabile casuale X non definisce la probabilità associata al singolo valore puntuale, ad esempio all'evento $X = x$. Si è, piuttosto, interessati alla probabilità che X assuma valori in un qualche intervallo non degenere del suo dominio di definizione.
Quindi, punti del dominio di definizione a cui corrisponde una elevata (bassa) funzione di densità indicano una alta (bassa) probabilità che X assuma valori in un loro intorno.

Esempio
Si consideri una v.c. X, definita su un intervallo $[a, b]$, la cui funzione di densità di probabilità, $f_X(x)$, è pari ovunque a 0.2.
Per uno specifico valore della v.c. in esame, ad esempio $x = 100$, ed un intervallo $dx = 1$, la probabilità che X sia compresa tra 100 e 101 è, per la definizione fornita, pari a 0.2.

$$P[100 \leq X \leq 101] = f_X(100)dx = 0.2$$

Se adesso si prende un intervallo di ampiezza dx =0.1, la probabilità che X sia compresa tra 100.0 e 100.1 risulta inferiore ed in particolare pari a 0.2·0.1=0.02.

$$P[100.0 \leq X \leq 100.1] = f_X(100)dx = 0.2 \cdot 0.1 = 0.02$$



Al fine di uniformare la simbologia, è usuale ipotizzare che la v.c. sia definita sull'intervallo (-∞ , +∞), con la convenzione che la funzione di densità sia nulla per quei valori esterni al campo in cui la v.c. può assumere valori non nulli. Così se la v.c. X può assumere valori compresi tra a e b, la funzione densità di probabilità è così definita:

$$\begin{cases} f_X(x) = 0 & \text{per } -\infty < x < a \\ f_X(x) = g(x) & \text{per } a \leq x \leq b \\ f_X(x) = 0 & \text{per } b < x < \infty \end{cases}$$

dove con $g(x)$ si è indicata l'espressione matematica della specifica funzione di densità di probabilità nell'intervallo (a,b).

Le condizioni necessarie e sufficienti perché la funzione $f_X(x)$ individui una funzione di densità di probabilità di una v.c. X sono:

1. $f_X(x) \geq 0 \quad -\infty < x < +\infty$
2. $\int_{-\infty}^{+\infty} f_X(x)dx = 1$ l'area totale sottesa dalla $f_X(x)$ sul campo di definizione è uguale ad 1.

Nota la funzione di densità di una v.c. X continua si può calcolare la probabilità di un qualsiasi evento utilizzando le nozioni elementari del calcolo integrale.
Volendo calcolare la probabilità che la v.c. assuma un valore compreso nell'intervallo (c,d) incluso nel campo di definizione, si tratta di integrare la $f_X(x)$ per x che va da c fino a d:

$$P(c \leq X \leq d) = \int_c^d f_X(x)dx$$

Importanti conseguenze che discendono dalla definizione di v.c. continua sono le seguenti:

1. $P(X = x_0) = \int_{x_0}^{x_0} f_X(x)dx = 0$
2. $P(X \leq x_0) = P(X < x_0)$

Indicando con X una v.c. continua, è possibile individuare una funzione $F_X(x)$ (indicata come **funzione di ripartizione** o **funzione di probabilità cumulata** o **di non superamento**) la quale, per ciascun valore reale x che la X può assumere, indica la probabilità che si verifichino valori di X minori o al più uguali a x:

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

Considerando la definizione della pdf si ha, quindi:

$$F_X(x) = \int_{-\infty}^{x} f_X(x)dx$$

La probabilità di qualsiasi evento è deducibile dalla $F_X(x)$, ad esempio:

$$P(a < X < b) = \int_{a}^{b} f_X(x)dx = F_X(b) - F_X(a)$$

Per i punti x in cui la $F_x(x)$ è differenziabile vale, inoltre:

$$f_X(x) = \frac{dF_X(x)}{dx}$$

Esercizi svolti

1. Sia data la funzione

$$f(x)=\begin{cases}\dfrac{x}{8} & 0\le x\le 4\\ 0 & altrimenti\end{cases}$$

Verificare che $f(x)$ è una densità di probabilità di una variabile aleatoria continua X e calcolare la probabilità che:

a) $X\le 2$

b) $1\le X\le 3$

Svolgimento

Affinché $f(x)$ sia una densità di probabilità, quindi riscrivibile come $f_X(x)$ deve essere:

1) $f_X(x)\ge 0 \qquad \forall x\in R$

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

La prima condizione è verificata $\forall x\in R$. Inoltre si ha:

$$\int_{-\infty}^{+\infty} f_X(x)dx=\int_0^4 \frac{x}{8}dx=\left.\frac{x^2}{16}\right|_0^4=1$$

a) $P(X\le 2)=\int_{-\infty}^{2} f_X(x)dx=\int_0^2 \frac{x}{8}=\left.\frac{x^2}{16}\right|_0^2=\frac{1}{4}$

b) $P(1\le X\le 3)=\int_1^3 f_X(x)dx=\int_1^3 \frac{x}{8}=\left.\frac{x^2}{16}\right|_1^3=\frac{9}{16}-\frac{1}{16}=\frac{1}{2}$

2. Sapendo che la funzione di ripartizione di una variabile aleatoria X è:

$$F_X(x)=\begin{cases}0 & x<0\\ \left(\dfrac{x}{3}\right)^3 & 0\le x\le 3\\ 1 & x>3\end{cases}$$

Calcolare la probabilità $P\left(1\le X\le \frac{3}{2}\right)$ e $P(2\le X\le 4)$

Svolgimento

$$P\left(1 \leq X \leq \frac{3}{2}\right) = F_X\left(\frac{3}{2}\right) - F_X(1) = \left(\frac{3}{6}\right)^3 - \left(\frac{1}{3}\right)^3 = \frac{1}{8} - \frac{1}{27} = \frac{19}{216} = 0.088 = 8.8\ \%$$

$$P(2 \leq X \leq 4) = F_X(4) - F_X(2) = 1 - \left(\frac{2}{3}\right)^3 = \frac{19}{27} = 0.704 = 70.4\ \%$$

3. Trovare il valore della costante $c \in R$ in modo che la funzione:

$$f(x) = \begin{cases} cx^2 & 0 \leq x \leq 3 \\ 0 & altrimenti \end{cases}$$

sia una densità di probabilità, ovvero riscrivibile come $f_X(x)$.

Trovare, inoltre, la funzione di ripartizione $F_X(x)$ e calcolare la probabilità $P(1 \leq X \leq 2)$.

Svolgimento

Affinché la funzione assegnata risulti una densità di probabilità devono essere rispettate le condizioni:

1) $f_X(x) \geq 0 \qquad \forall x \in R$

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

Da cui $cx^2 \geq 0 \Rightarrow c \geq 0$; inoltre:

$$\int_{-\infty}^{+\infty} f_X(x)dx = \int_0^3 cx^2 dx = c\frac{x^3}{3}\Bigg|_0^3 = 9c$$

$$9c = 1 \Rightarrow c = \frac{1}{9}$$

la densità di probabilità è pertanto:

$$f_X(x) = \begin{cases} \frac{1}{9}x^2 & 0 \leq x \leq 3 \\ 0 & altrimenti \end{cases}$$

Si calcola ora la funzione di ripartizione:

per $x < 0 \qquad F_X(x) = 0$

per $0 \le x \le 3$ $$F_X(x) = \int_{-\infty}^{x} f_X(x)dx = \int_{0}^{x} f_X(x)dx = \int_{0}^{x} \frac{1}{9}x^2 dx = \frac{1}{27}x^3$$

per $x > 3$ $$F_X(x) = \int_{-\infty}^{x} f_X(x)dx = \int_{0}^{3} f_X(x)dx + \int_{3}^{x} f_X(x)dx = \int_{0}^{3} f_X(x)dx = \frac{1}{27}x^3\Big|_0^3 = 1$$

riassumendo:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{27}x^3 & 0 \le x \le 3 \\ 1 & x > 3 \end{cases}$$

Calcolando $P(1 \le X \le 2)$, si ha:

$$P(1 \le X \le 2) = F_X(2) - F_X(1) = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

Distribuzioni di probabilità

Di seguito sono illustrate le caratteristiche principali delle distribuzioni di probabilità introdotte durante il corso.
Per ogni distribuzione esaminata è riportata una scheda contenente un breve cenno sul significato della distribuzione, il significato della variabile casuale in esame, l'espressione della funzione di probabilità di massa (per le distribuzioni discrete), della funzione di densità di probabilità (per le distribuzioni continue), dei momenti (media, varianza, coefficiente di variazione, coefficiente di asimmetria, coefficiente di Kurtosi), le equazioni per la stima dei parametri con il metodo dei momenti e della massima verosimiglianza.
Le distribuzioni esaminate sono le seguenti:

<u>Distribuzioni discrete</u>

1) Distribuzione di Bernoulli
2) Distribuzione Binomiale
3) Distribuzione Geometrica
4) Distribuzione Binomiale Negativa
5) Distribuzione di Poisson
6) Distribuzione Uniforme discreta

<u>Distribuzioni continue</u>

1) Distribuzione Uniforme
2) Distribuzione Normale
3) Distribuzione LogNormale
4) Distribuzione Esponenziale
5) Distribuzione Gamma
6) Distribuzione Gumbel

La distribuzione di Bernoulli

Si dice che una variabile casuale X ha una distribuzione di Bernoulli se la funzione di densità discreta di X è data da:

Funzione di densità discreta: $P(X=x)=\begin{cases} p & se\ x=1 \\ 1-p & se\ x=0 \\ 0 & altrimenti \end{cases}$

Momenti

Media	p
Varianza	(1-p)p
Coefficiente di variazione	$\dfrac{\sqrt{p(1-p)}}{p}$
Coefficiente di asimmetria	$\dfrac{1-2p}{\sqrt{(1-p)p}}$
Kurtosi	$3+\dfrac{1-6p(1-p)}{\sqrt{(1-p)p}}$

Si definisce *prova bernoulliana* un esperimento casuale il cui esito può essere soltanto di due tipi: ‘successo’ o ‘insuccesso’. Posto che il valore della variabile casuale X sia 1 quando la prova dà luogo ad un successo, e 0 quando dà luogo ad un insuccesso, X ha una distribuzione di Bernoulli con parametro p che rappresenta la probabilità di avere un successo nella singola prova.

La distribuzione Binomiale

La distribuzione Binomiale è una distribuzione discreta a due parametri molto utilizzata in statistica. Essa, infatti, si basa su un processo molto semplice che può essere utilizzato per schematizzare i più svariati fenomeni. Si consideri una serie di n prove Bernoulliane indipendenti a due a due, ciascuna con probabilità di successo p. Il numero di successi X in n prove è una variabile casuale distribuita secondo la legge Binomiale.

Funzione di densità discreta: $P_n(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ $\qquad x = 0, 1, 2, \ldots, n$

Il simbolo $\binom{n}{x}$ è il coefficiente binomiale (si legge 'n su x') ed è definito come segue:

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

e rappresenta il numero di combinazioni di n oggetti presi x alla volta.

Variabile casuale in esame:	numero di successi X in n prove bernoulliane indipendenti, ciascuna con probabilità di successo p. $0 \le x \le n$
Parametri:	p,n $\qquad 0 \le p \le 1;\ n \ge 0$
Interpretazione fisica dei parametri:	p – probabilità di successo ad ogni prova n – numero di prove

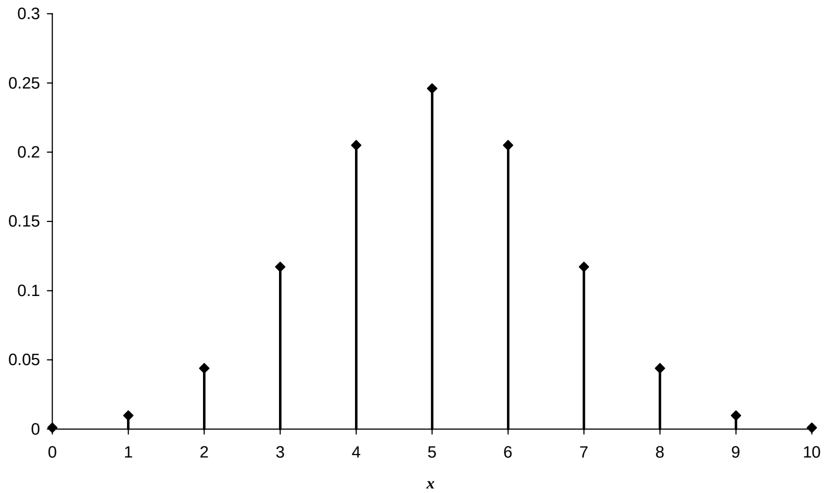
Momenti

Media	np
Varianza	$np(1-p)$
Coefficiente di variazione	$\sqrt{\frac{1-p}{np}}$
Coefficiente di asimmetria	$\frac{1-2p}{\sqrt{np(1-p)}}$
Kurtosi	$3 + \frac{1-6p+6p^2}{\sqrt{np(1-p)}}$

Osservazioni: per n=1 la distribuzione Binomiale coincide con quella di Bernoulli; la distribuzione Binomiale è simmetrica se e solo se p = 0.5.

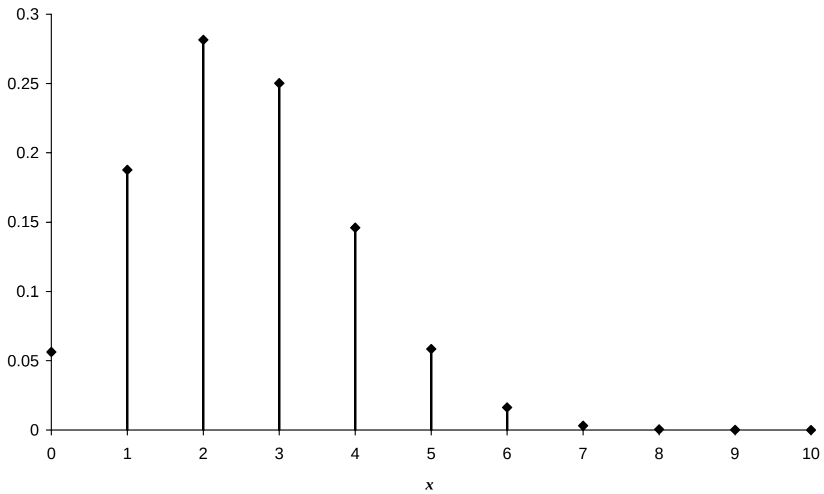
Funzione di densità discreta distribuzione Binomiale

$n = 10 \quad p = 0.5$

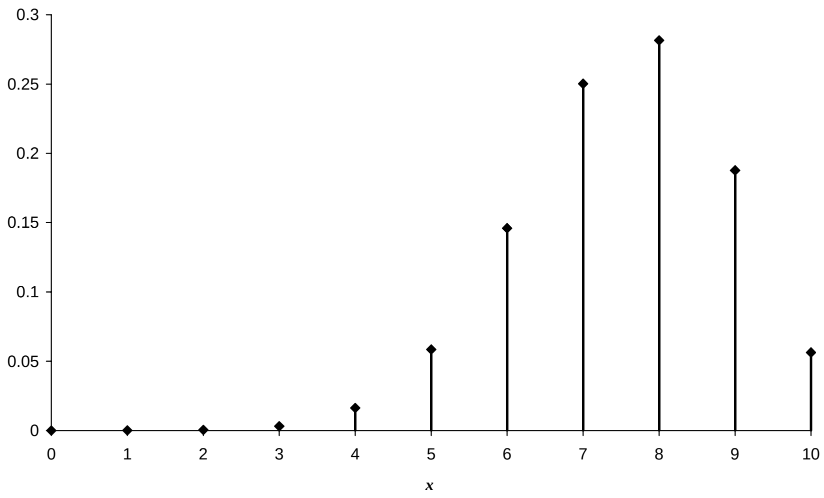


Dal diagramma a sinistra si evince, ad esempio, che la probabilità di avere 0 successi in 10 prove, ciascuna con probabilità di successo pari a 0.5, risulta 0.0009. La probabilità di avere 5 successi è, invece, pari a 0.246.

$n = 10 \quad p = 0.25$

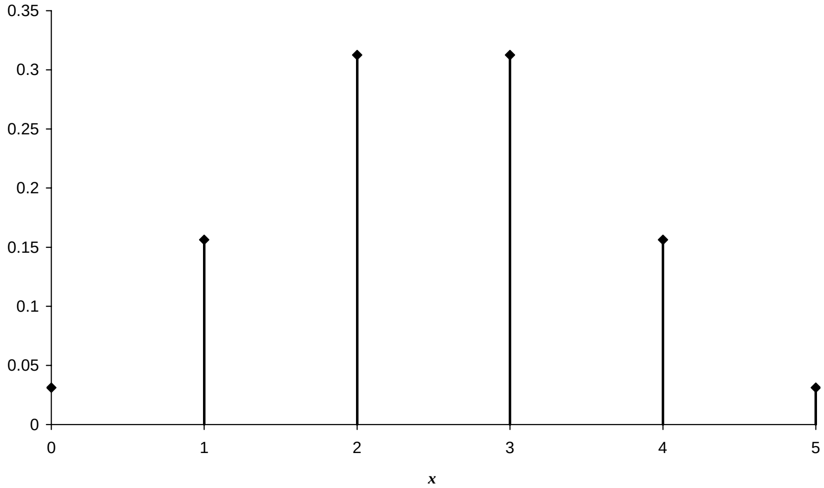


$n = 10 \quad p = 0.75$

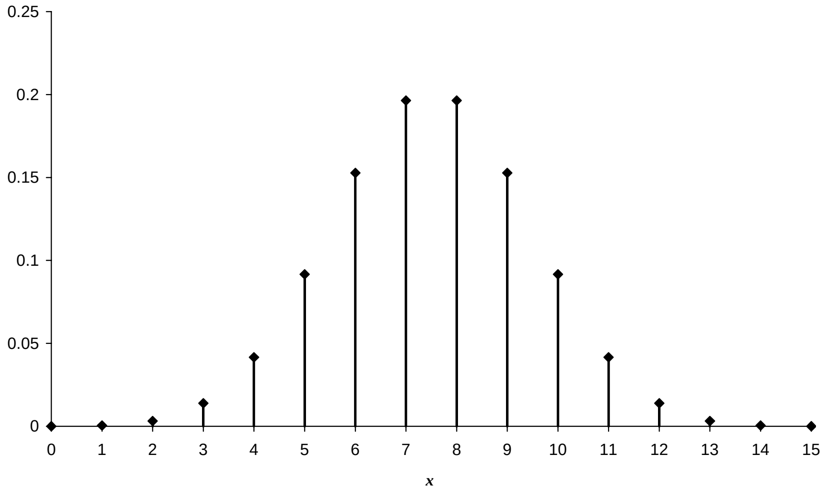


Funzione di densità discreta distribuzione Binomiale

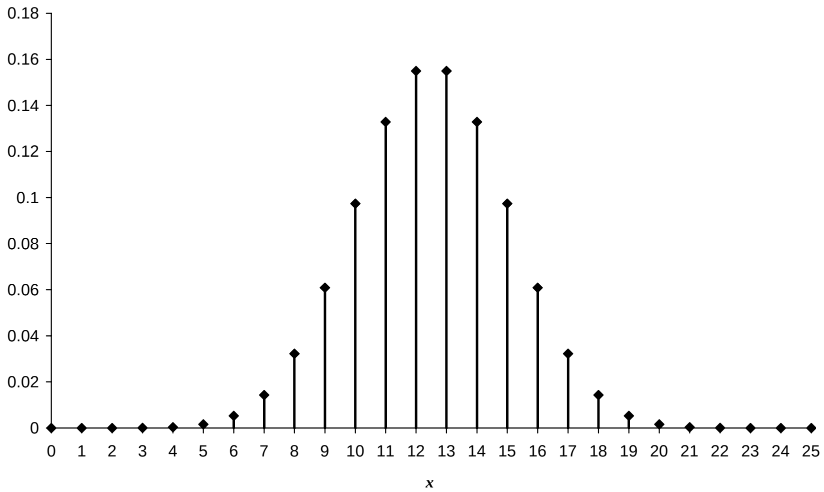
$n = 5 \quad p = 0.5$



$n = 15 \quad p = 0.5$



$n = 25 \quad p = 0.5$



La distribuzione Geometrica

Si consideri una serie di n prove Bernoulliane indipendenti a due a due, ciascuna con probabilità di successo p.
La distribuzione Geometrica è una distribuzione di probabilità discreta a un parametro e rappresenta la probabilità di ottenere un successo dopo un numero n di prove bernoulliane.
Un successo di verifica all'ennesimo tentativo se, evidentemente accade che:

- i primi n-1 tentativi sono insuccessi
- l'ennesimo tentativo è un successo

Funzione di densità discreta: $P_N(n) = (1-p)^{n-1}p$

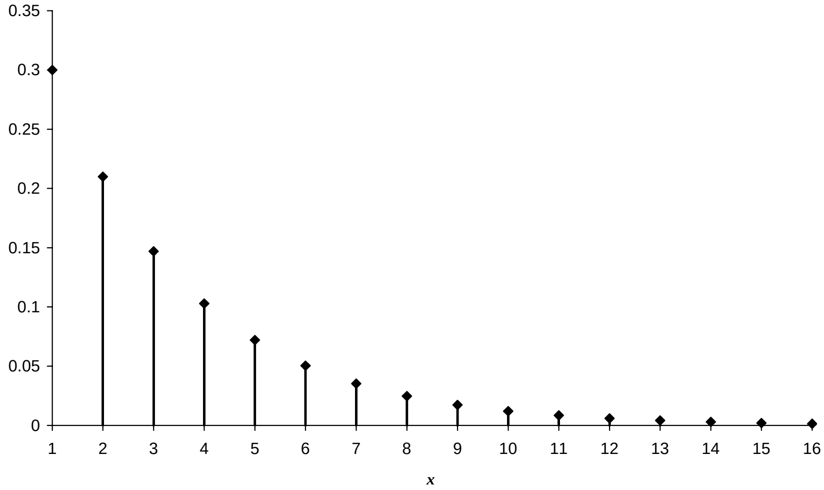
Variabile casuale in esame:	numero N di prove necessarie per ottenere il primo successo in una serie di prove bernoulliane indipendenti, ciascuna con probabilità di successo p. n >0
Parametri:	p $\quad 0 \le p \le 1$
Interpretazione fisica dei parametri:	p – probabilità di successo ad ogni prova

Momenti

Media	$\frac{1}{p}$
Varianza	$\frac{(1-p)}{p^2}$
Coefficiente di variazione	$\sqrt{1-p}$

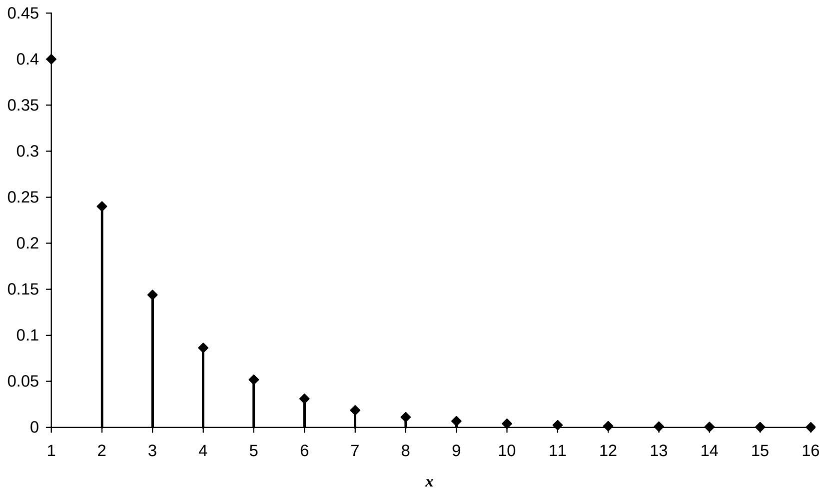
Funzione di densità discreta distribuzione Geometrica

$p = 0.3$

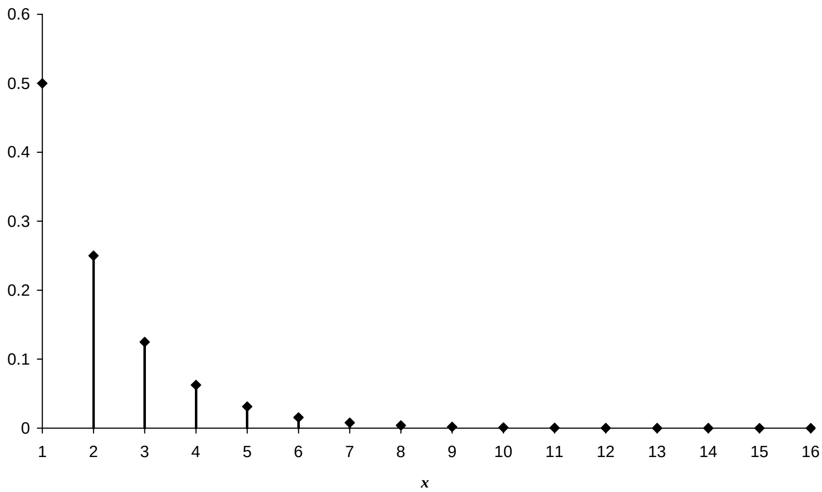


Per un esperimento, per il quale la probabilità di successo in ciascuna prova è pari a 0.3, la probabilità di avere il 1° successo alla 1° prova, risulta 0.3.
La probabilità di avere il 1° successo alla 3° prova, è, invece, pari a 0.147.

$p = 0.4$



$p = 0.5$



La distribuzione Binomiale Negativa

La distribuzione Binomiale Negativa è una distribuzione discreta a due parametri. Essa può essere vista come una generalizzazione della distribuzione Geometrica. Se, infatti, la distribuzione geometrica descrive la distribuzione di probabilità del numero di prove necessarie per ottenere il primo successo, la distribuzione Binomiale Negativa descrive la distribuzione di probabilità del numero di prove necessarie per ottenere il *k*-esimo successo in una serie di prove Bernoulliane indipendenti, ciascuna con probabilità di successo *p*.
La distribuzione Geometrica può essere, quindi, considerata come un caso particolare della Binomiale Negativa per *k*=1.

Funzione di densità discreta: $P_k(X = x) = \binom{x-1}{k-1}(1-p)^{x-k} p^k \qquad x \geq k$

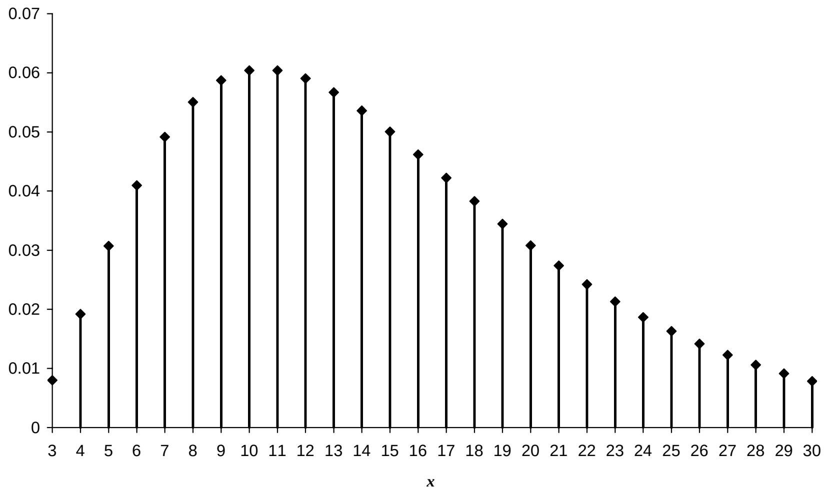
Variabile casuale in esame:	numero X di prove necessarie per ottenere il k-esimo successo in una serie di prove bernoulliane indipendenti, ciascuna con probabilità di successo p. $x \geq k$
Parametri:	$p,k \qquad 0 \leq p \leq 1;\ k \geq 1$
Interpretazione fisica dei parametri:	p – probabilità di successo ad ogni prova k – numero di successi

Momenti

Media	$\frac{k}{p}$
Varianza	$\frac{k(1-p)}{p^2}$
Coefficiente di variazione	$\sqrt{\frac{1-p}{k}}$

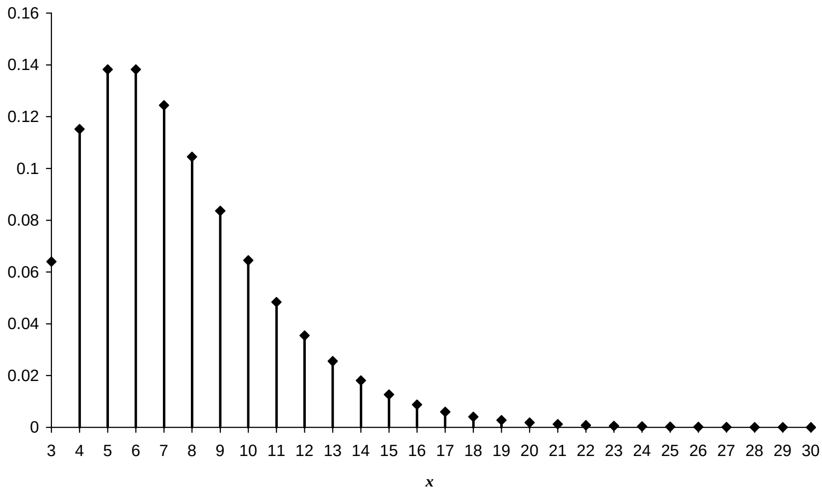
Funzione di densità discreta distribuzione Binomiale Negativa

$k = 3$ $p = 0.2$

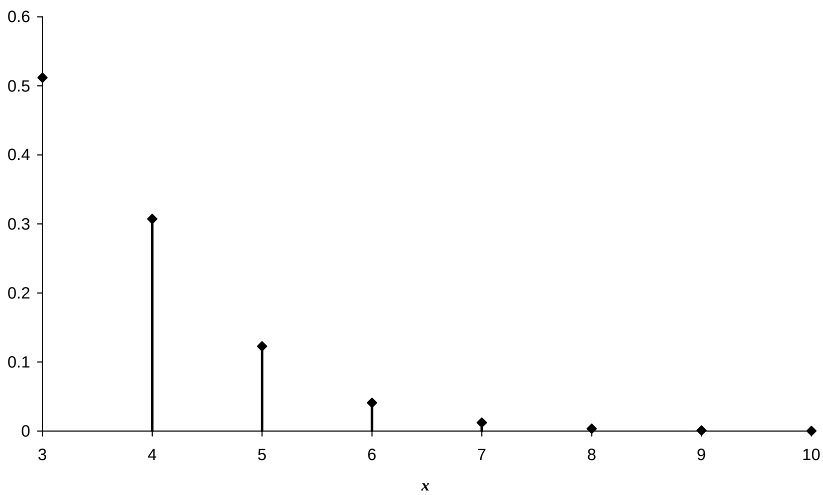


Per un esperimento in cui ciascuna prova ha probabilità di successo pari a 0.2, la probabilità di avere il 3° successo 3° prova è pari a 0.08.
La probabilità di avere il 3° successo alla 10° prova, è, invece, pari circa a 0.06.

$k = 3$ $p = 0.4$

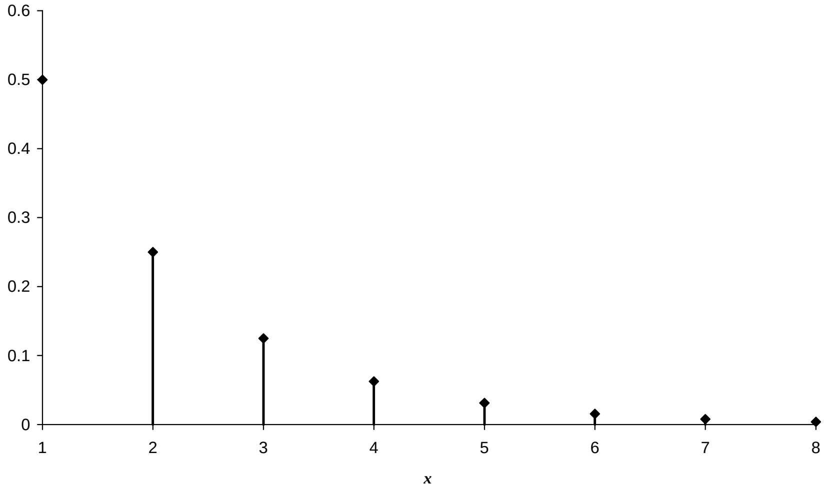


$k = 3$ $p = 0.8$

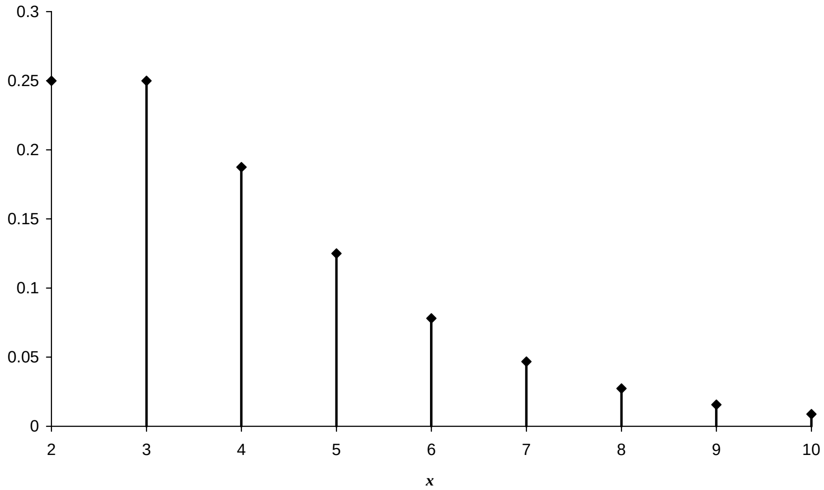


Funzione di densità discreta distribuzione Binomiale Negativa

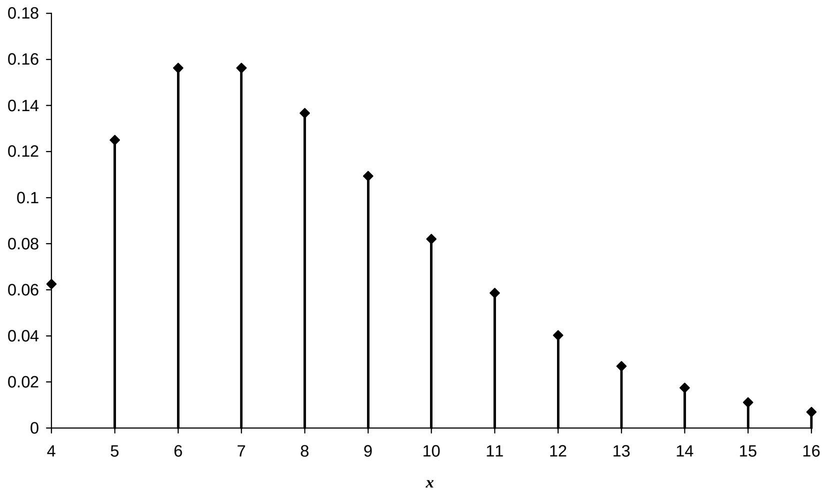
$k = 1$ $p = 0.5$



$k = 2$ $p = 0.5$



$k = 4$ $p = 0.5$



La distribuzione di Poisson

La distribuzione di Poisson è una distribuzione di probabilità discreta ad un parametro, utilizzata per rappresentare una vasta gamma di fenomeni casuali, in particolare quelli connessi ad un qualche "conteggio" di eventi (ad esempio il numero di errori, di incidenti o nell'ambito idrologico il numero di perturbazioni) in un prefissato periodo di tempo.
Si ipotizzi di essere interessati a conoscere il numero totale di veicoli che arrivano in una certa località in un fissato periodo di tempo tra 0 e t. Si ipotizzi quindi di suddividere il periodo temporale t in n intervalli (ognuno dei quali rappresenta una prova) e di conoscere, inoltre, p la probabilità che arrivi un veicolo in ciascuno di questi intervalli. La probabilità ricercata può essere determinata con l'ausilio della distribuzione Binomiale. Si può dimostrare che la distribuzione di Poisson, con parametro $\nu = np$, è il limite a cui tende una distribuzione Binomiale per $p \to 0$ e per $n \to \infty$. Poiché p rappresenta la probabilità di avere un successo, la variabile casuale di Poisson può essere assimilata ad una Binomiale nella quale la probabilità che si verifichi un successo è prossima allo zero: per tale motivo la variabile casuale di Poisson si presta bene ad interpretare i cosiddetti *eventi rari*.

Funzione di densità discreta: $P(X = x) = \dfrac{\nu^x e^{-\nu}}{x!}$ $\qquad x = 0, 1, 2, \ldots$

Variabile casuale in esame:	numero X di eventi che si osservano in un determinato periodo di tempo. $x \geq 0$
Parametri:	ν
Interpretazione fisica dei parametri:	ν – numero medio di eventi nel prefissato periodo di tempo

Momenti

Media	ν
Varianza	ν
Coefficiente di variazione	$\dfrac{1}{\sqrt{\nu}}$
Coefficiente di asimmetria	$\dfrac{1}{\sqrt{\nu}}$
Kurtosi	$3 + \dfrac{1}{\nu}$

La distribuzione di Poisson è spesso utilizzata nell'analisi dei processi stocastici, ossia dei processi per i quali la distribuzione di probabilità delle variabili casuali presenta la caratteristica di essere dipendente dal tempo.
L'esempio sopra descritto può essere posto come un processo stocastico *X(t)*, il cui valore (casuale) ad ogni istante *t* è il numero di arrivi occorsi da *t*=0.
In tal caso è possibile riscrivere la distribuzione di Poisson considerando $\nu=\lambda t$, in cui λ è il numero medio di arrivi nell'intervallo unitario di tempo:

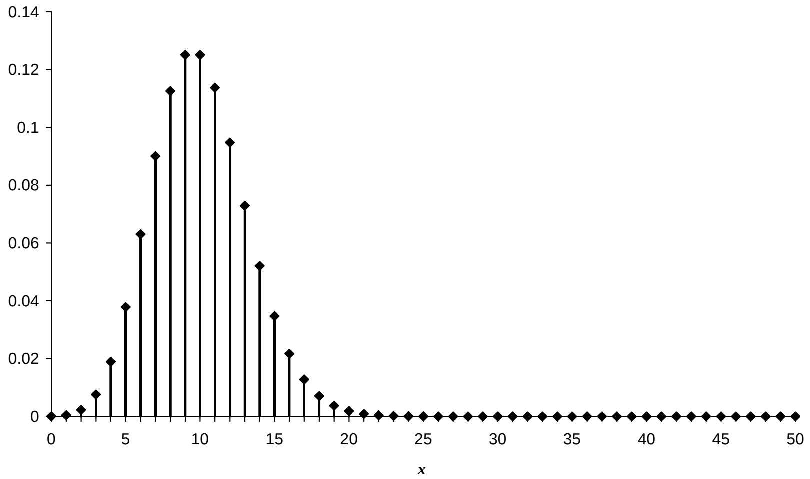
$$P(X = x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

La relazione precedente vale però solo nel caso di processo stocastico Poissoniano e cioè di un processo che ha le seguenti caratteristiche:

1. **proprietà di stazionarietà:** la probabilità che in un intervallo (t, $t+h$), di ampiezza h opportunamente piccola, si verifichi un solo evento è approssimativamente uguale a λh e non dipende dall'istante t;
2. **proprietà di non molteplicità:** la probabilità che in un intervallo di tempo piccolo si verifichi più di un evento è nulla;
3. **proprietà di indipendenza:** i numeri di eventi in intervalli di tempo disgiunti (non sovrapposti) sono indipendenti.

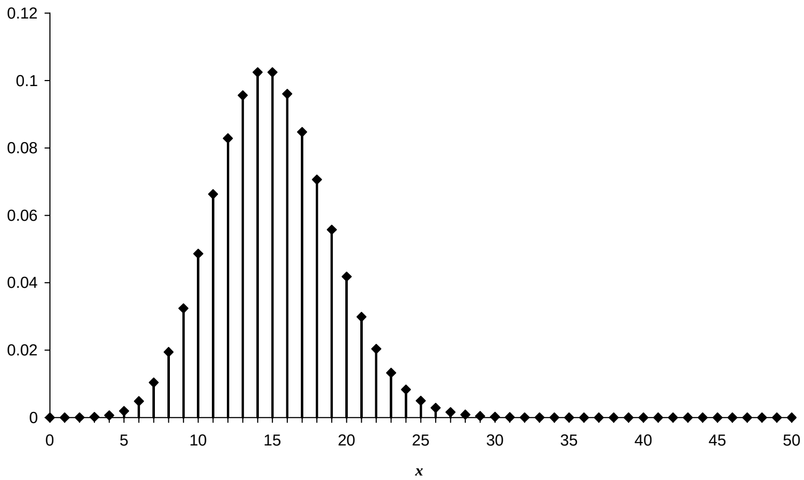
Funzione di densità discreta distribuzione Poisson

$\nu = 10$

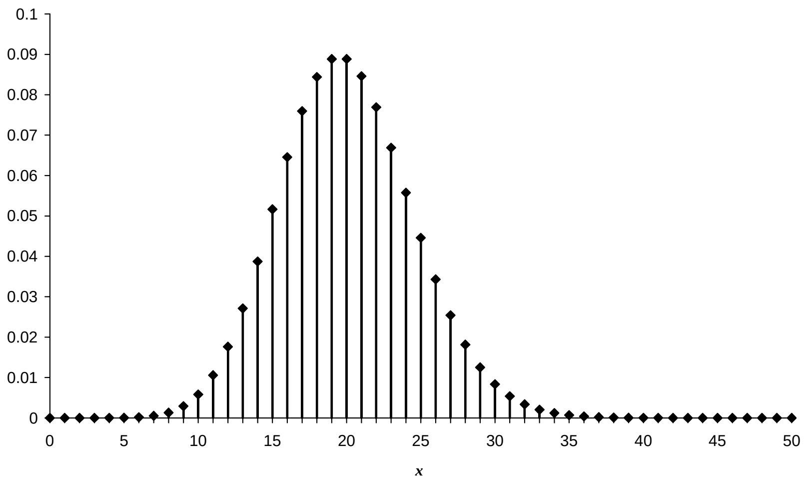


Noto il numero medio di eventi nel prefissato periodo di tempo, $\nu = 10$, la probabilità di avere 5 eventi nello stesso periodo di riferimento risulta 0.037.
La probabilità di avere 10 successi, è, invece, pari circa a 0.125.

$\nu = 15$



$\nu = 20$



La distribuzione Uniforme discreta

Sia X una variabile casuale che assume i valori 1,2, …, N; essa ha una distribuzione uniforme discreta se risulta:

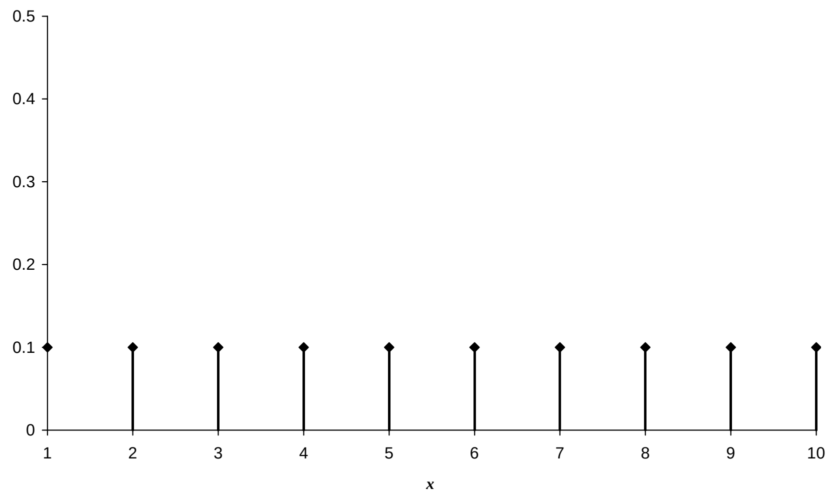
Funzione di densità discreta: $P(X=x)=\begin{cases}1/N & x=1,2,\ldots,N\\ 0 & altrove\end{cases}$ $\qquad N \geq 0$

Momenti

Media	$\frac{N+1}{2}$
Varianza	$\frac{(N^2-1)}{12}$

Funzione di densità discreta distribuzione Uniforme discreta

$N = 10$



Esercizi svolti sulle distribuzioni discrete

1. In un lago le trote formano il 33% della popolazione ittica presente, mentre lucci e carpe formano il restano 67%. Estraendo a caso 15 pesci, calcolare la probabilità che:

a)nessun pesce sia una trota

b)siano tutte trote

Svolgimento

Dai dati forniti, la probabilità di pescare una trota è 0.33 mentre la probabilità di non pescarla è 0.67

a)la probabilità che tra 15 pesci non ci sia nemmeno una trota si calcola applicando la distribuzione binomiale:

$$P_{15}(X=0)=\binom{15}{0}\cdot(0.33)^{0}\cdot(0.67)^{15}=0.00246 \qquad (0.24\%)$$

b)la probabilità che siano tutte trote è invece:

$$P_{15}(X=15)=\binom{15}{15}\cdot(0.33)^{15}\cdot(0.67)^{0}=0.00000006 \qquad (0.000006\%)$$

2. È statisticamente dimostrato che in tutte la popolazioni umane nascono più maschi che femmine, per cui si può assumere che la probabilità della nascita di un maschio è approssimativamente 0.52 e quella di una femmina 0.48. usando la distribuzione binomiale, calcolare la probabilità di avere 3 figli maschi nelle famiglie con 4 figli.

Svolgimento

Applicando la distribuzione binomiale si ha:

$$P_{4}(X=3)=\binom{4}{3}\cdot(0.52)^{3}\cdot(0.48)^{1}=0.28 \qquad (28\ \%)$$

3. L'inquinamento atmosferico di una grande città provoca reazioni allergiche tanto gravi che in un mese di 30 giorni si registrano 27 ricoveri urgenti. Se la v.c. reazione allergica è distribuita in accordo con la legge di Poisson, trovare la probabilità di avere al massimo 1 caso di allergia al giorno.

Svolgimento

Dopo aver calcolato la media giornaliera delle allergie $\nu = 0.9(27/30)$, la probabilità di avere al massimo 1 caso di allergia al giorno è data dalla somma:

$$P(X=0)+P(X=1)=\frac{e^{-0.9}\cdot 0.9^0}{0!}+\frac{e^{-0.9}\cdot 0.9^1}{1!}=0.4066+0.3659=0.7725\ (77.3\ \%)$$

4. In uno scaffale di cioccolata ci sono 1000 tavolette, di cui 20 contenenti un tagliando con cui si vince un premio. Comprando 20 tavolette a caso, calcolare la probabilità di trovare almeno 3 tagliandi vincenti.

Svolgimento

La probabilità di estrarre un tagliando vincente risulta essere $p=20/1000=0.02$; applicando la legge di distribuzione binomiale, la probabilità di avere almeno tre tagliandi vincenti su 20 tavolette di cioccolata risulta:

$$P(almeno\,3\,tagliandi\,vincenti\,su\,20\,tavolette)=1-[P_{20}(0)+P_{20}(1)+P_{20}(2)]=$$

$$=1-\binom{20}{0}\cdot(0.02)^0\cdot(0.98)^{20}-\binom{20}{1}\cdot(0.02)^1\cdot(0.98)^{19}-\binom{20}{2}\cdot(0.02)^2\cdot(0.98)^{18}=$$

$$=1-0.6676-0.2725-0.0528=0.00710\ (0.7\%)$$

5. La puntura di un insetto porta conseguenze tossiche gravi con lunghe degenze in 1 caso su 5000 persone. Risolvendo l'esercizio sia con la legge binomiale che con la legge di Piosson, si valutino le probabilità che in un campione casuale di 800 persone punte dall'insetto si possano avere:

a)esattamente due casi di tossicità grave;

b)non ci siano casi di tossicità grave.

almeno 3 tagliandi vincenti.

Svolgimento

Poiché la probabilità di avere tossicità risulta essere pari a $p=1/5000=0.0002$, applicando la legge binomiale si ottiene:

a) $P_{800}(X=2)=\binom{800}{2}(0.0002)^2(0.9998)^{798}=0.01090=1.09\%$

b) $P_{800}(X=0)=\binom{800}{0}(0.0002)^0(0.9998)^{800}=0.8521=85.21\%$

È più semplice risolvere l'esercizio con la legge di Poisson, visto che *n* è molto elevato e *p* molto piccolo, per cui si può scrivere:

$\nu=n\cdot p=800\cdot 0.0002=0.16$

a) $P(X=2)=\dfrac{e^{-0.16}0.16^2}{2!}=0.01091=1.09\%$

b) $P(X=0)=\dfrac{e^{-0.16}0.16^0}{0!}=0.8521=85.21\%$

6. Le trote salmonate rappresentano il 20% di tutte le trote presenti in un allevamento artificiale. Pescando 12 trote, nell'ipotesi che le probabilità di estrazione delle trote non siano influenzate dalla pesca, calcolare la probabilità che:

a)almeno 1 trota pescata sia salmonata

b)al massimo 2 trote pescate siano salmonate.

Svolgimento

Dai dati forniti, la probabilità di pescare una trota salmonata è 0.20 mentre la probabilità di non pescarla è 0.80.

a)la probabilità che tra le 12 trote non ce ne sia nessuna salmonata è:

$$P_{12}(X=0)=\binom{12}{0}\cdot(0.20)^0\cdot(0.80)^{12}=0.06872\quad(6.9\ \%)$$

b)la probabilità che al massimo 2 trote pescate siano salmonate è:

$$P_{12}(X=0)+P_{12}(X=1)+P_{12}(X=2)=$$
$$=\binom{12}{0}\cdot(0.20)^0\cdot(0.80)^{12}+\binom{12}{1}\cdot(0.20)^1\cdot(0.80)^{11}+\binom{12}{2}\cdot(0.20)^2\cdot(0.80)^{10}=$$
$$=0.0682+0.2062+0.2835=0.5579\quad(55.8\ \%)$$

7. In una zona priva di strutture fognarie il rischio di infezione da colera è di 2 casi su 1000 persone. Valutare la probabilità che in un campione casuale di 300 persone vi siano:

a)esattamente 2 casi di colera

b)non più di 1 caso di colera.

Svolgimento

La probabilità di infezione da colera è $p=2/1000=0.002$. Utilizzando la legge binomiale si ottiene:

a) $P_{300}(X=2)=\binom{300}{2}(0.002)^2(0.998)^{298}=0.0988=9.88\%$

b)

$$P_{300}(X=0)+P_{300}(X=1)=\binom{300}{0}(0.002)^{0}(0.998)^{300}+\binom{300}{1}(0.002)^{1}(0.998)^{299}=$$

$$=0.5485+0.3287=0.8782=87.82\ \%$$

è praticamente la stessa cosa, oltre che numericamente più semplice, risolvere l'esercizio con la legge di Poisson, essendo *n* elevato e *p* piccolo:

$$\nu=n\cdot p=300\cdot 0.002=0.60$$

a) $P(X=2)=\dfrac{e^{-0.60}\cdot 0.60^{2}}{2!}=0.0988=9.88\%$

b) $P(X=0)+P(X=1)=\dfrac{e^{-0.60}\cdot 0.60^{0}}{0!}+\dfrac{e^{-0.60}\cdot 0.60^{1}}{1!}=0.8781=87.81\%$

8. La probabilità che un apparecchio subisca guasti è $p=0.05$. Calcolare la probabilità che su 16 di tali apparecchi:

a)al più 2 si guastino;

b)almeno 2 si guastino.

Svolgimento

La variabile aleatoria X indica il numero dei guasti. I valori dei parametri in esame sono: $n=16\quad p=0.05\quad 1-p=0.95$, per cui si ha:

a) $P_{16}(X\le 2)=P_{16}(X=0)+P_{16}(X=1)+P_{16}(X=2)=$

$$=\binom{16}{0}\cdot(0.05)^{0}\cdot(0.95)^{16}+\binom{16}{1}\cdot(0.05)^{1}\cdot(0.95)^{15}+\binom{16}{2}\cdot(0.05)^{2}\cdot(0.95)^{14}=$$

$$=0.4401+0.3706+0.1463=0.9570=95.7\ \%$$

b) $P_{16}(X\ge 2)=1-P_{16}(X<2)=1-[P_{16}(X=0)+P_{16}(X=1)]=$

$$=1-\left[\binom{16}{0}\cdot(0.05)^{0}\cdot(0.95)^{16}+\binom{16}{1}\cdot(0.05)^{1}\cdot(0.95)^{15}\right]=$$

$$=1-(0.4401+0.3706)=0.1893=18.93\ \%$$

9. Se il 5% dei microchip venduti da una ditta sono difettosi, determinare la probabilità che su 4 microchip scelti a caso:

a)1 microchip sia difettoso

b)meno di 2 siano difettosi

c)calcolare la media del numero di microchip difettosi su un totale di 300 pezzi.

Svolgimento

La variabile aleatoria X indica il numero dei microchip difettosi. I valori dei parametri in esame sono: $n = 4 \quad p = 0.05 \quad 1 - p = 0.95$, per cui:

a) $P(X = 1) = \binom{4}{1} \cdot (0.05)^1 \cdot (0.95)^3 = 0.1715 = 17.15\ \%$

b)

$$P(X < 2) = P(X = 0) + P(X = 1) = \binom{4}{0} \cdot (0.05)^0 \cdot (0.95)^4 + \binom{4}{1} \cdot (0.05)^1 \cdot (0.95)^3 =$$
$$= 0.8145 + 0.1715 = 0.9860 = 98.6\ \%$$

c)n° di pezzi=300, p=0.05, da cui, $\mu = n \cdot p = 300 \cdot 0.05 = 15$

10. Un libro di 500 pagine contiene 50 errori di stampa. Calcolare la probabilità di trovare almeno 3 errori di stampa su una pagina aperta a caso.

Svolgimento

Il numero medio di errori su una pagina è $\nu = 50/500 = 0.10$. Con la distribuzione di Poisson si ha:

$$P(X \geq 3) = 1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1) + P(X = 2)] =$$
$$= 1 - \left[\frac{e^{-0.1} \cdot 0.1^0}{0!} + \frac{e^{-0.1} \cdot 0.1^1}{1!} + \frac{e^{-0.1} \cdot 0.1^2}{2!}\right] = 1 - 0.99985 = 0.00015$$

La distribuzione Uniforme Continua o Rettangolare

Una variabile casuale X è uniformemente distribuita nell'intervallo reale $[a,b]$ se è caratterizzata dalle seguenti funzioni di densità e di probabilità cumulata:

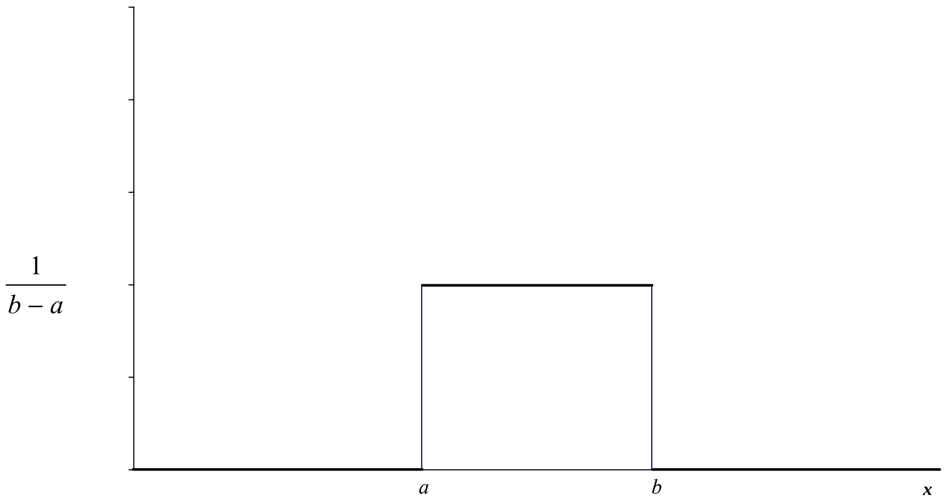
Funzione di densità di probabilità: $f_X(x) = \begin{cases} \dfrac{1}{b-a} & se\ a \le x \le b \\ 0 & altrimenti \end{cases}$ $\qquad a, b \in (-\infty, +\infty)$

Funzione di probabilità cumulata: $F_X(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & se\ a \le x \le b \\ 1 & x > b \end{cases}$

Momenti

Media	$\dfrac{a+b}{2}$
Varianza	$\dfrac{(b-a)^2}{12}$

Funzione di densità di probabilità distribuzione Uniforme continua



La distribuzione Normale

Storicamente tale distribuzione è stata ricavata dagli errori di misurazione accidentali di un fenomeno. E' funzione di due parametri μ e σ che variano rispettivamente tra (-∞, +∞) e (0,+∞).

Funzione di densità di probabilità: $f_X(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right]$ $\qquad x \in (-\infty, +\infty)$

Caratteristiche della funzione di densità di probabilità Normale:

1. $f_X(x)$ è tanto più grande quanto più l'argomento dell'esponenziale è piccolo e raggiunge il massimo (moda) per $x = \mu$;
2. per x che tende a (-∞, +∞), $f_X(x)$ tende a 0;
3. $f_X(x)$ ha due flessi in $\mu-\sigma$ ed in $\mu+\sigma$;
4. $f_X(x)$ è simmetrica intorno ad $x = \mu$ ossia $f_X(x-\mu) = f_X(x+\mu)$;

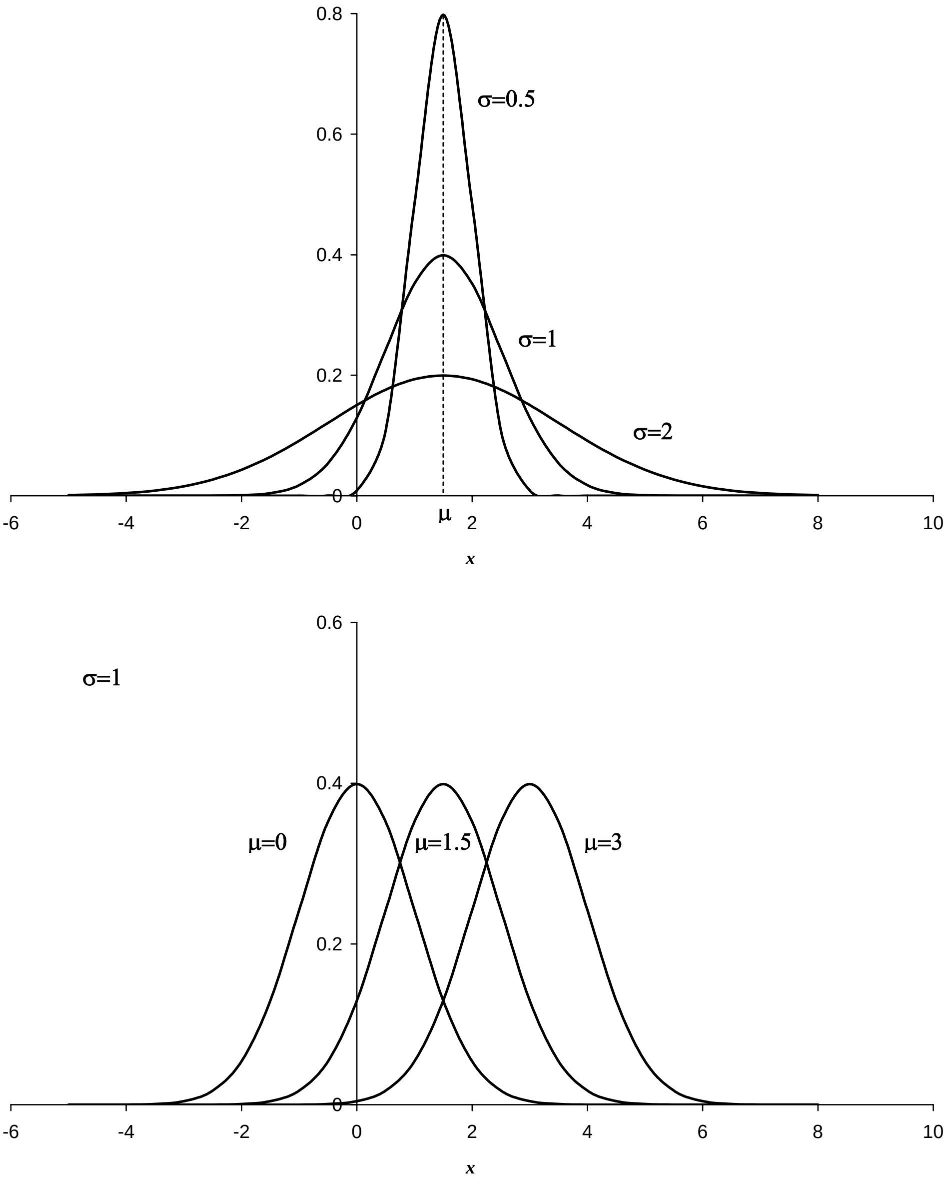
Si può dimostrare che una combinazione di v.c. Normali indipendenti è ancora una Normale. La distribuzione normale, inoltre, è la forma limite di molte altre distribuzioni di probabilità.

Variabile casuale in esame:	X assume valori tra (-∞, +∞)
Parametri:	μ, σ
Interpretazione fisica dei parametri:	μ – media σ – deviazione standard

Momenti

Media	μ
Varianza	σ^2
Coefficiente di variazione	$\dfrac{\sqrt{\sigma^2}}{\mu}$
Coefficiente di asimmetria	0
Kurtosi	3

Funzione di densità di probabilità distribuzione Normale



Una variabile casuale con distribuzione normale è detta *standard* se ha media 0 e varianza 1, e viene denotata con $N(0,1)$.

La distribuzione LogNormale

Si dice che la variabile casuale X è distribuita secondo la legge lognormale con parametri μ_y e σ_y se la variabile Y=log X è distribuita secondo la legge Normale.

Funzione di densità di probabilità: $f_X(x) = \frac{1}{\sigma_y \sqrt{2\pi x}} \exp\left[-\frac{1}{2\sigma_y^2}\left(\log x - \mu_y\right)^2\right]$ $\qquad x>0$

Variabile casuale in esame:	$X>0$
Parametri:	μ_y, σ_y
Interpretazione fisica dei parametri:	μ_y – media di Y=log X σ_y – deviazione standard di Y=log X

Momenti

Media	$e^{\mu_y + \frac{1}{2}\sigma_y^2}$
Varianza	$e^{2\mu_y + 2\sigma_y^2} - e^{\mu_y + \sigma_y^2}$

La distribuzione Esponenziale

La distribuzione esponenziale è stata introdotta ed usata come modello per la descrizione probabilistica degli intervalli di tempo intercorrenti tra successive manifestazioni di diversi fenomeni sia naturali che artificiali. Infatti, si può mostrare che, se il numero delle manifestazioni in un dato intervallo di tempo ha una distribuzione di Poisson, la lunghezza dell'intervallo di tempo fra manifestazioni successive del fenomeno ha una distribuzione esponenziale .

Funzione di densità di probabilità: $f_X(x) = \lambda e^{-\lambda x}$ $\quad x > 0; \lambda > 0$

Funzione di probabilità cumulata: $F_X(x) = (1 - e^{-\lambda x})$

Variabile casuale in esame:	$X > 0$ Lunghezza intervallo di tempo fra successive manifestazioni di un fenomeno temporale ad occorrenza poissoniana.
Parametri:	$\lambda > 0$
Interpretazione fisica dei parametri:	λ – inverso del valor medio dell'intervallo di tempo tra manifestazioni successive del fenomeno in esame

Momenti

Media	$\frac{1}{\lambda}$
Varianza	$\frac{1}{\lambda^2}$
Coefficiente di asimmetria	2
Kurtosi	9

Se si considera un processo Poissoniano, di parametro λ, il periodo di tempo T intercorrente tra due successi consecutivi (ovvero il tempo a cui si osserva il primo successo) è una variabile casuale che si distribuisce con legge esponenziale.
L'epressione della funzione di probabilità cumulata può essere facilmente ricavata a partire dall'espressione della distribuzione di Poisson. Considerando che la probabilità che la v.c. T superi un certo valore t è pari alla probabilità (calcolata con Poisson) che nessun evento si verifichi nell'intervallo di tempo di lunghezza t, si ha:

$$P(T > t) = 1 - F_T(t) = P(X = 0) = \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t}$$

da cui si ottiene $F_T(t) = (1 - e^{-\lambda t})$

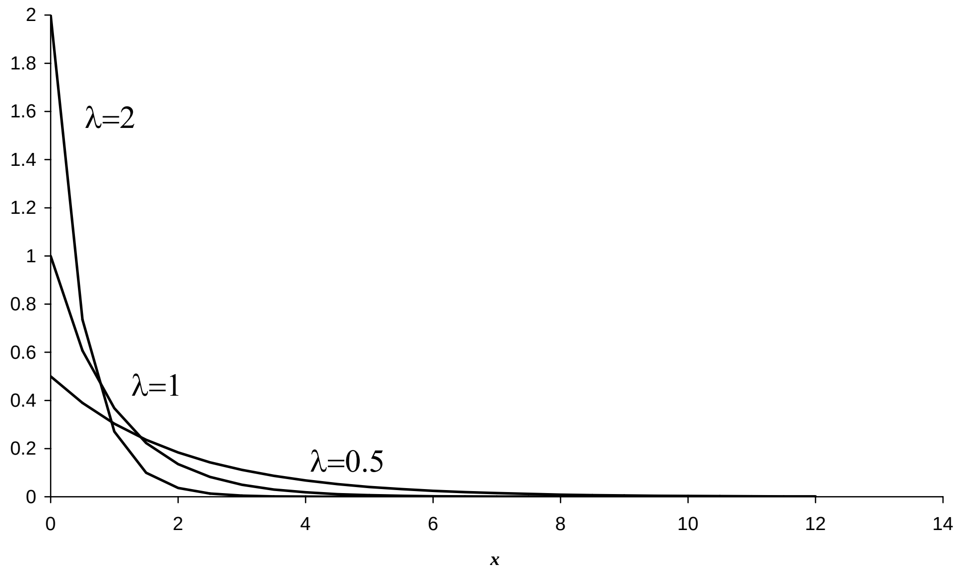
Per la distribuzione esponenziale si parla, inoltre, di proprietà di assenza di memoria. Per capirne il significato assumiamo di avere un processo di Poisson e che sia passato un tempo *s* dall'ultimo evento verificatosi. Ci si chiede quale sia la probabilità che passi ancora almeno un tempo *t* prima che si verifichi il prossimo evento. La proprietà di assenza di memoria per la distribuzione esponenziale ci dice che la probabilità cercata è indipendente dal valore di *s*: si può cioè fare come se il processo iniziasse nel momento in cui ci troviamo.

Stima dei parametri della distribuzione

La stima del parametro λ della distribuzione esponenziale, sia operando con il metodo dei momenti sia operando con il metodo della massima verosimiglianza, si ottiene mediante la relazione:

$$\hat{\lambda} = \frac{1}{\bar{x}}$$

Funzione di densità di probabilità distribuzione Esponenziale



La distribuzione Gamma

Si consideri un fenomeno fisico, naturale o artificiale, che si verifichi ad intervalli di tempo casuali. Assumendo che il numero di occorrenze del fenomeno in un dato intervallo di tempo sia distribuito con legge di Poisson, si può mostrare che la lunghezza dell'intervallo di tempo in cui si verificano α manifestazioni ha una distribuzione di tipo Gamma. Così una variabile casuale gamma può essere pensata come una variabile 'tempo di attesa' continua, indicante il tempo che si deve aspettare per avere a manifestazioni del fenomeno che si sta esaminando. Si può, inoltre, dimostrare che la somma di variabili casuali esponenziali indipendenti con uguale distribuzione ha una distribuzione di tipo Gamma.

Si ricordi che le variabili casuali geometriche e binomiali negative erano variabili casuali tempo di attesa discrete e rappresentano in un certo senso i corrispondenti discreti rispettivamente dell'esponenziale e della gamma.

Funzione di densità di probabilità: $f_X(x) = \dfrac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^\alpha \Gamma(\alpha)}$ $x>0$; $\alpha>0$; $\beta>0$

Variabile casuale in esame:	X >0 Lunghezza intervallo di tempo in cui si verificano α manifestazioni di un fenomeno temporale ad occorrenza poissoniana.
Parametri:	α, β α>0; β>0
Interpretazione fisica dei parametri:	α - è un parametro di forma e rappresenta il numero medio di manifestazioni che hanno luogo nell'intervallo di tempo di lunghezza *x*. (spesso si indica α= *k*) β - è un parametro di scala e rappresenta il valor medio dell'intervallo di tempo fra due manifestazioni successive del fenomeno. (β= 1/λ) Il prodotto dei due parametri rappresenta il valor medio dell'intervallo di tempo che trascorre per osservare α manifestazioni del fenomeno in esame.

Momenti

Media	$\alpha\beta$
Varianza	$\alpha\beta^2$
Coefficiente di asimmetria	$\dfrac{2}{\sqrt{\alpha}}$
Kurtosi	$3+6/\alpha$

Funzione matematica Gamma

Definizione: $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$

Proprietà: $\Gamma(x) = (x-1)\Gamma(x-1)$ ricorsività (vale anche per un x reale)

$\Gamma(n) = (n-1)!$ n=1,2,… (vale per n interi)

Relazione polinomiale approssimata per la funzione gamma (da Abramovitz-Stegun):

$$\Gamma(x+1) = 1 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + a_5 x^5 + \varepsilon(x) \qquad (*)$$

con $0 \le x \le 1$ $|\varepsilon(x)| \le 5 \cdot 10^{-5}$

Parametri

a_1=-0.5748646

a_2=0.9512363

a_3=-0.6998588

a_4=0.4245549

a_5=-0.10108678

Esempio: Calcolo di $\Gamma(3.1)$

$\Gamma(3.1)$= 2.1· $\Gamma(2.1)$= 2.1 · 1.1 ·$\Gamma(1.1)$

$\Gamma(1.1)$= $\Gamma(x+1)$ ponendo x=0.1

dalla (*) $\rightarrow \Gamma(0.1+1)$=0.951367489

in definitiva: $\Gamma(3.1)$= 2.1· $\Gamma(2.1)$= 2.1 · 1.1 · 0.951367489=2.1976589

Stima dei parametri della distribuzione

1) Metodo dei Momenti

$$\hat{\beta} = \frac{s^2}{\bar{x}} \qquad \hat{\alpha} = \frac{\bar{x}^2}{s^2}$$

2) Metodo della Massima Verosimiglianza (da verificare)
Il sistema di equazioni che permette di determinare i parametri della distribuzione Gamma è:

$$\begin{cases} \hat{\beta} = \bar{x} / \hat{\alpha} \\ \log(\bar{x}) - \frac{1}{n}\sum_{j=1}^{n} \log(x_j) = \log(\hat{\alpha}) - \psi(\hat{\alpha}) \end{cases}$$

dove $\psi(\hat{\alpha})$ = funzione digamma, $\psi(\alpha) = \frac{d \log[\Gamma(\alpha)]}{d\alpha}$

Questa funzione può essere calcolata in via approssimata sulla base del calcolo della funzione gamma come indicato di seguito. La seconda delle due equazioni di stima della massima verosimiglianza per funzione di distribuzione Gamma, nell'incognita α, deve essere risolta per tentativi. Come valore di primo tentativo per α si consideri la stima fornita dal metodo dei momenti.

Funzione digamma

Calcolo approssimato della funzione digamma.

Sulla base dei valori tabulati della funzione gamma si può calcolare $\psi(\alpha)$ come:

$$\psi(\alpha) \cong \frac{1}{2\Delta^2}\left\{(\alpha_c - \alpha)\ln\left[\frac{\Gamma(\alpha_c)}{\Gamma(\alpha_a)}\right] + (\alpha - \alpha_b)\ln\left[\frac{\Gamma(\alpha_d)}{\Gamma(\alpha_b)}\right]\right\}$$

dove $\alpha_b < \alpha < \alpha_c$ e $\Delta = \alpha_b - \alpha_a = \alpha_c - \alpha_b = \alpha_d - \alpha_c = 0.01$

se $\alpha \geq 2$ riduzione a $1 \leq \alpha < 2$ $\qquad \psi(\alpha) = \psi(\alpha - 1) + \dfrac{1}{\alpha - 1}$

se $\alpha < 1$ riduzione a $1 \leq \alpha < 2$ $\qquad \psi(\alpha) = \psi(\alpha + 1) + \dfrac{1}{\alpha}$

Esempio: Calcolo di $\psi(2.743)$

1. $\psi(2.743) = \psi(1.743) + \dfrac{1}{1.743}$
2. $\alpha = 1.743$

$\alpha_a = 1.73 \quad \Gamma(\alpha_a) = 0.91467$

$\alpha_b = 1.74 \quad \Gamma(\alpha_b) = 0.91683$

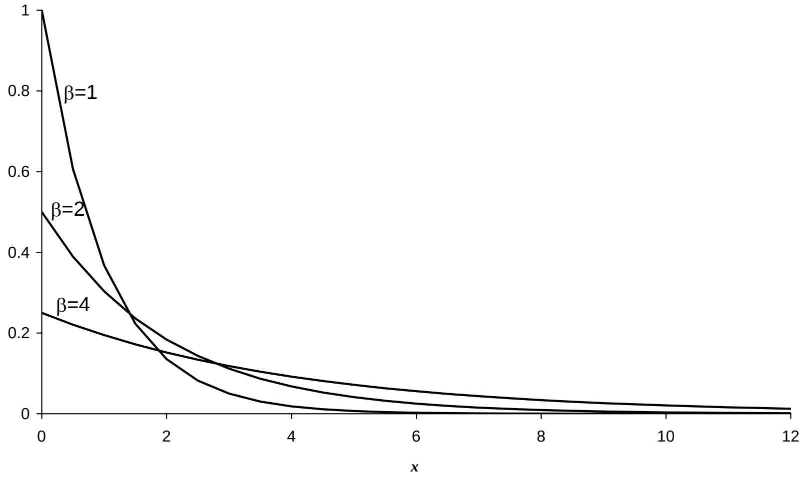
$\alpha_c = 1.75 \quad \Gamma(\alpha_c) = 0.91906$

$\alpha_d = 1.76 \quad \Gamma(\alpha_d) = 0.92137$

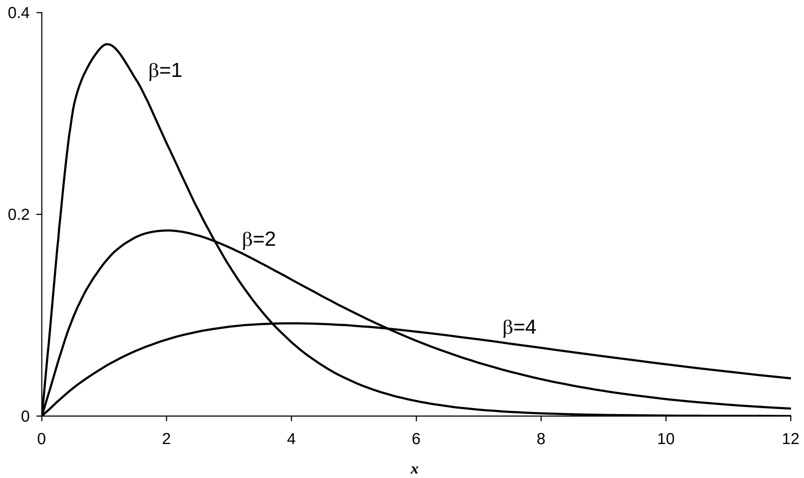
3. $\psi(1.743) \cong \dfrac{1}{2(0.01)^2}\left\{0.007\ln\left[\dfrac{0.91906}{0.91467}\right] + 0.003\ln\left[\dfrac{0.92137}{0.91683}\right]\right\} = 0.24$
4. $\psi(2.743) = 0.2417 + \dfrac{1}{1.743} = 0.8154$

Funzione di densità di probabilità distribuzione Gamma

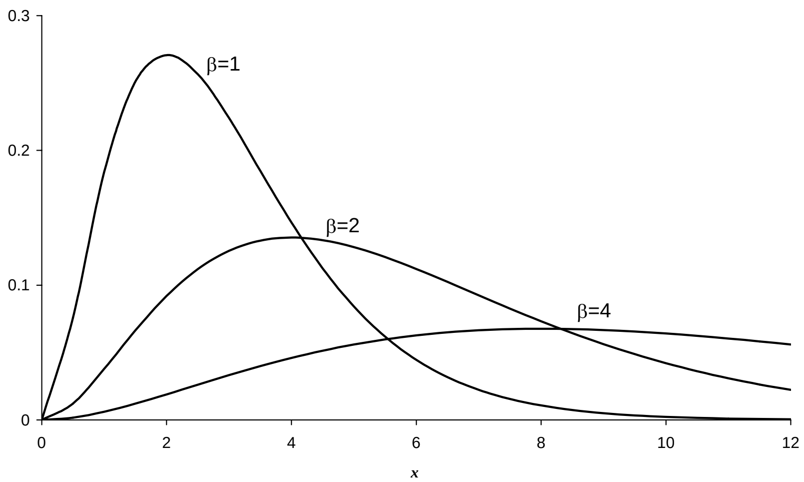
$\alpha = 1$



$\alpha = 2$



$\alpha = 3$



La distribuzione di Gumbel

La distribuzione di Gumbel a due parametri è molto utilizzata in Idrologia. La distribuzione nasce dalla teoria del valore estremo: si può dimostrare che se un fenomeno si presenta nel tempo con un processo di tipo Poissoniano, e se il fenomeno è associato ad una grandezza distribuita esponenzialmente, il massimo di suddetta grandezza in un prefissato intervallo temporale segue una distribuzione di Gumbel.

Funzione di densità di probabilità: $f_X(x) = \alpha e^{-\alpha(x-\varepsilon)} e^{-e^{-\alpha(x-\varepsilon)}}$ $x \geq 0$;

Funzione di probabilità cumulata: $F_X(x) = e^{-e^{-\alpha(x-\varepsilon)}}$ $x \geq 0$;

Variabile casuale in esame:	$X \geq 0$ Valore massimo di una grandezza in un determinato periodo di tempo.
Parametri:	α, ε α>0; ε>0
Interpretazione fisica dei parametri:	α - è un parametro di scala. ε - è un parametro di posizione che coincide con il valore modale della distribuzione.

Momenti

Media	$\varepsilon + \dfrac{0.5772}{\alpha}$
Varianza	$\dfrac{\pi^2}{6\alpha^2}$
Coefficiente di variazione	$\dfrac{\pi}{\sqrt{6}(\varepsilon\alpha + 0.5772)}$
Coefficiente di asimmetria	1.14

Stima dei parametri della distribuzione

1) Metodo dei Momenti

$$\hat{\alpha} = \frac{\pi}{\sqrt{6}s}; \quad \hat{\varepsilon} = \bar{x} - \frac{0.5772}{\hat{\alpha}}$$

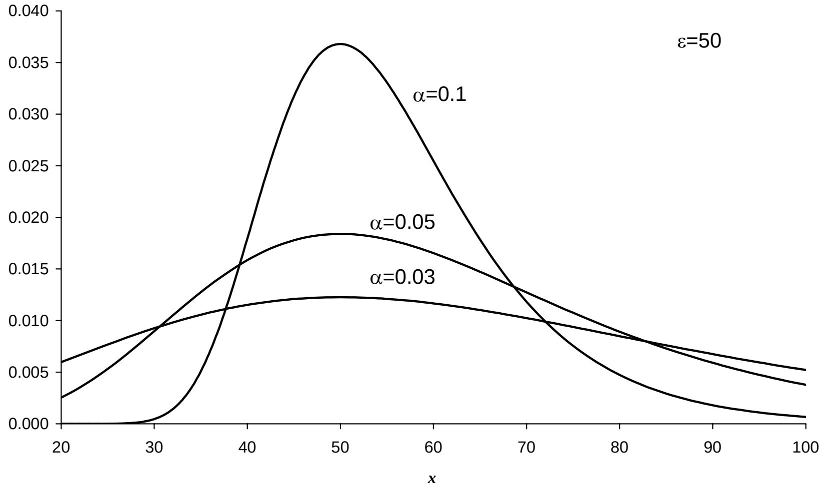
dove s e $\bar{x}$ sono rispettivamente la media campionaria e lo scarto quadratico medio campionario.

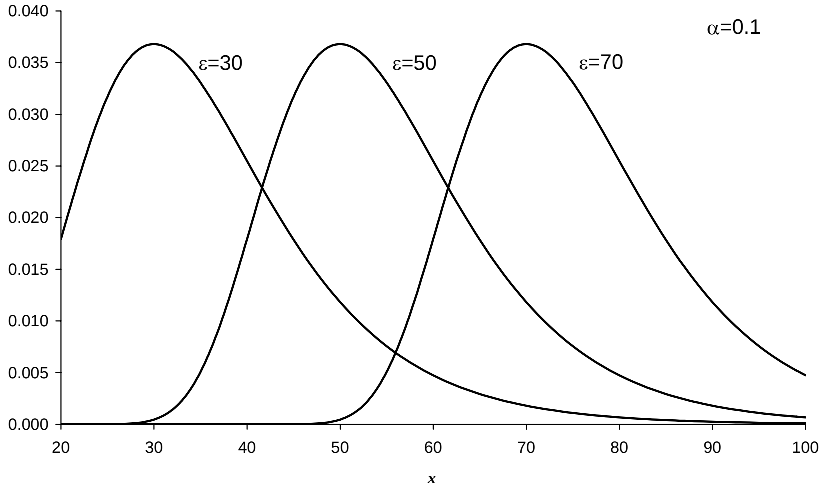
2) Metodo della Massima Verosimiglianza
La stima dei parametri con il metodo della Massima Verosimiglianza può essere ottenuta risolvendo per via numerica il seguente schema di due equazioni in due incognite:

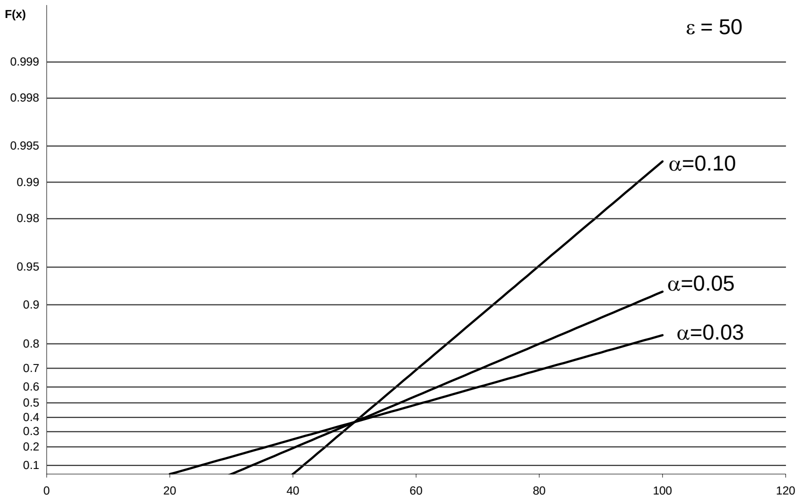
$$
\begin{cases}
\dfrac{1}{\hat{\alpha}} = \bar{x} - \dfrac{\sum_{i=1}^{n} x_i e^{-\hat{\alpha} x_i}}{\sum_{i=1}^{n} e^{-\hat{\alpha} x_i}} \\
e^{-\hat{\alpha}\hat{\varepsilon}} = \dfrac{1}{n}\sum_{i=1}^{n} e^{-\hat{\alpha} x_i}
\end{cases}
$$

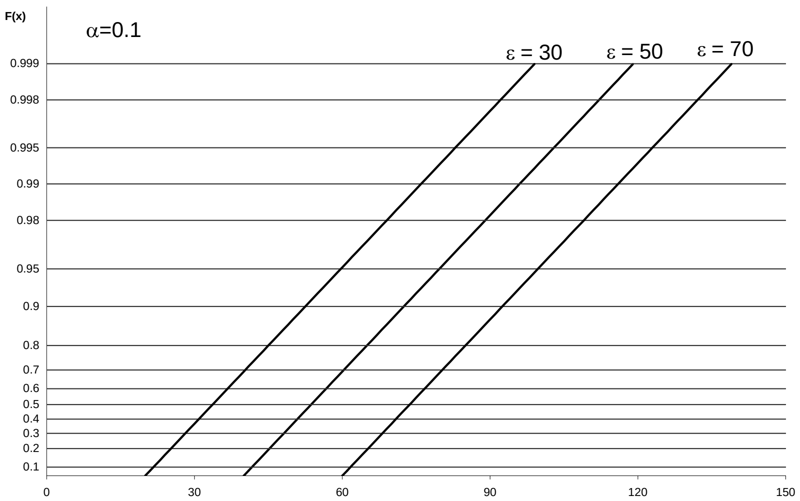
dove *n* è la numerosità del campione.

Funzione di densità di probabilità distribuzione Gumbel






F(x)
ε = 50
α=0.10
α=0.05
α=0.03
0.999
0.998
0.995
0.99
0.98
0.95
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
20
40
60
80
100
120


F(x)
α=0.1
ε = 30
ε = 50
ε = 70
0.999
0.998
0.995
0.99
0.98
0.95
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
30
60
90
120
150