Note sulla Distribuzione Binomiale

La distribuzione binomiale è relativa ad una variabile aleatoria discreta, che descrive i possibili risultati di un esperimento composto da *n* prove.
In particolare, definisce la probabilità di ottenere *k* successi sulle n prove, dove k è compreso tra 0 e *n*.
Le ipotesi per una corretta definizione della binomiale sono:

- ogni prova dell'esperimento può avere due esiti, detti "successo" e "insuccesso" rispettivamente
- il risultato di una prova non condiziona né è legato ai risultati delle altre prove
- la probabilità dell'evento successo in ogni prova è *p* e la probabilità di insuccesso è *q*=1-*p*

Come esempio si consideri un esperimento che consiste in 5 lanci di un dado e si consideri evento successo ottenere 5 o 6 su un singolo lancio.
La probabilità di ottenere 5 o 6 è p=2/6=1/3.
Un possibile risultato dell'esperimento è:

1 1 3 5 4 dove si mostrano i risultati dei 5 lanci. In questo caso si è avuto un successo. La probabilità di questa combinazione di lanci è q*q*q*p*q=p*q^4

Se volessimo stimare la probabilità di ottenere 1 successo su 5 lanci, dovremmo considerare le in quanti modi potrebbe verificarsi il successo. Potrebbe infatti risultare essere in corrispondenza della prima o della seconda o così via.
Questo numero, in questo caso banale, si può trovare con il coefficiente binomiale

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \frac{5!}{1!\,(4)!} = 5$$

Quindi la probabilità di ottenere 1 successo su 5 lanci sarà

$$5pq^4 = 5\,\frac{1}{3}\,\frac{2}{3}^4 = 0.3292$$

Nel caso di 2 successi in 5 lanci, sono possibili diverse combinazioni di risultati. Possibili risultati sono:
-successi nei lanci 1 e 2, oppure 1 e 3, oppure 1 e 4 etc. Analogamente, i successi potrebbero essere nei lanci 2 e 3 oppure 2 e 4

Il coefficiente binomiale fornisce il numero di combinazioni di n elementi presi a gruppi di *k*, senza ripetizione e quindi fornisce in questo caso il numero di tutte le configurazioni compatibili con 2 successi su 5 lanci

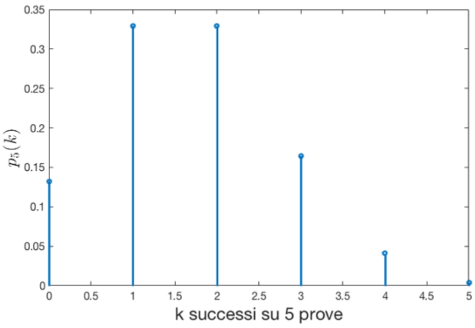
$$\binom{5}{2} = \frac{5!}{2!\,(3)!} = 10$$

In questo caso la probabilità di avere 2 successi su 5 lanci si indica con $p_n(k) = p_5(2)$ e vale $p_5(2) = 10p^2q^3 = 10\frac{1}{3}^2\frac{2}{3}^3 = 0.3292$

La densità di probabilità binomiale, quindi fornisce la probabilità di ottenere k successi su n prove e si può scrivere come

$$p_n(k) = \binom{n}{k} p^k q^{n-k}$$

In relazione al caso precedente, è possibile fare il grafico della densità di probabilità in funzione del numero di successi, quindi tra 0 e 5 successi su 5 lanci.



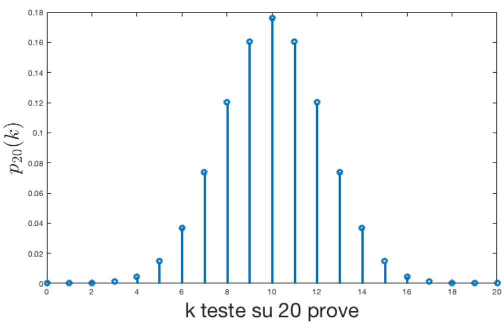
Nota matlab In Matlab la densità di proabilità binomiale si trova con il comando *binopdf* La figura precedente è stata trovata con facendo il grafico di y rispetto a x, dove - x=[0:5] sono i possibili successi su 5 prove - y=binopdf(x,5,1/3)

Esempio.

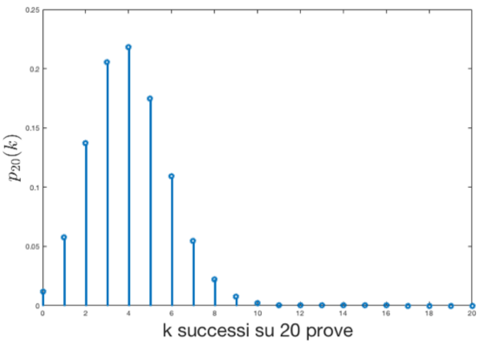
Dato un esperimento costituito da 20 prove, consistenti ognuna in un lancio di una moneta, la probabilità di ottenere 6 teste su 20 lanci è ottenibile come

$$p_{20}(6) = \frac{20!}{6!\,14!} 0.5^6 0.5^{14} = 0.0370$$

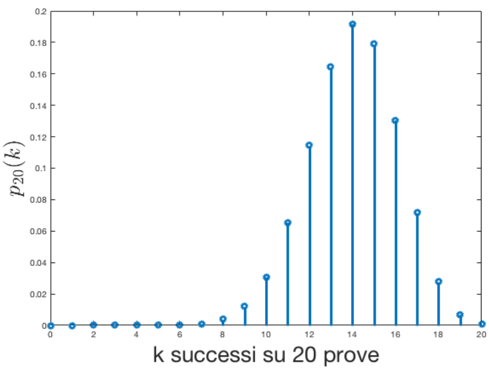
di seguito la densità di probabilità completa, ovvero calcolata per il ogni possibile numero di successi. Si può notare la simmetria della distribuzione. Questà è legata al fatto che *p=5=0.5*



Mostriamo invece la densità di probabilità per un esperimento composto da 20 prove, ma nel quale la probabilità di successo è pari a p=0.2



Nel caso invece di probabilità di successo sulla singola prova p=0.7, si ha



Si nota come in questi due ultimi casi la simmetria della distribuzione sia modificata.

Nel caso di *npq>>1* il teorema di Moivre-Laplace mostra che la densità di probabilità binomiale tende ad una distribuzione gaussiana con valore medio $\mu = np$ e varianza pari a npq.

$$p_n(k) = \binom{n}{k} p^k q^{n-k} \cong \frac{1}{\sqrt{2\pi npq}} e^{\frac{-(k-np)^2}{2npq}}$$

Esercizio.

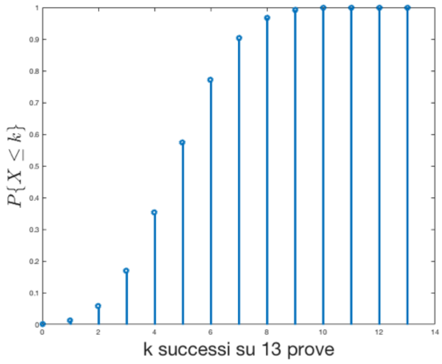
Si consideri un esperimento composto da 13 prove, caratterizzato da una probabilità di successo sulla singola prova pari a 0.4
Si calcolino a) la probabilità di ottenere 5 successi, e b) la probabilità di ottenere tra 0 e 2 successi.

a) $p_{13}(5) = \binom{13}{5} 0.4^5 0.6^8$ =0.2214
b) la probabilità richiesta è
c) $p_{0\_2} = \sum_{k=0}^{2} \binom{n}{k} p^k q^{n-k} = \binom{13}{0} 0.4^0 0.6^{13} + \binom{13}{1} 0.4^1 0.6^{12} + \binom{13}{2} 0.4^2 0.6^{11} = 0.0579$

Nota matlab

In Matlab la funzione di distribuzione binomiale si trova con il comando *binocdf*
Questa fornisce la somma cumulativa tra 0 e un valore indicato del numero di successi.

Nella figura seguente si mostra la funzione di distribuzione relativa all'esempio precedente.



Esercizio (Tratto, con modifiche, da *Teoria dei Segnali, Segnali aleatori*, M. Ciampi, G. del Corso, L. Verrazzani, ETS 1994)

Una fabbrica produce 10000 viti al giorno. Il diametro delle viti D è schematizzabile come una variabile normale di valore medio pari a 100 e varianza pari a 1.
Si chiede quale sia la probabilità che in un giorno vengano prodotte almeno 200 viti con diametro maggiore di 102.

La probabilità che una vite abbia un diametro maggiore di 102 è pari a

$$P\{D > 102\} = 1 - P\{D \leq 102\}$$

Opero una standardizzazione della variabile

$$P\{D \leq 102\} = P\left\{\frac{D-100}{1} \leq \frac{102-100}{1} = 2\right\} = \\ = P\{z \leq 2\} = \Phi(2)$$

dove $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{\alpha^2}{2}} d\alpha$ è la funzione di distribuzione della variabile normale standard.
Dalle tabelle della distribuzione standard si vede che

$P\{D > 102\} = 1 - \Phi(2)$=0.02275

Il numero di viti con diametro maggiore di 102 su 10000, si può vedere come una binomiale con n=1000 e p=0.02275.

Per trovare la probabilità richiesta dovremmo sommare i valori della densità di probabilità binomiale per k maggiore di 199.
Ovviamente con il calcolatore questo sarebbe facilmente ottenibile tramite il comando

1-*binocdf(199,10000,0.02275)*=0.9717

Se dovessimo utilizzare le tabelle, potremmo sfruttare l'approssimazione della binomiale con la gaussiana di valore medio $\mu = 227.5$ e varianza 222.3244. La variabile aleatoria X descrive il numero di viti prodotte giornalmente con diametro superiore a 102. Quindi la probabilità richiesta è

$$P\{200 \leq X \leq 10000\} = \Phi\left\{\frac{10000 - 227.5}{14.9105}\right\} - \Phi\left\{\frac{200 - 227.5}{14.9105}\right\} \cong 1 - \Phi\{-1.8142\} =$$

= $\Phi\{1.8142\}$=0.9649

Usando le tabelle il valore più prossimo sarebbe 1-0.03515=0.96485.