Legge binomiale e Teorema del limite centrale
(quesiti da 1 a 5)

Considero ora la variabile aleatoria binomiale che è la somma di n variabili aleatorie di Bernulli.
Si può simulare questa variabile aleatoria con il lancio di n monete tutte dello stesso tipo tali che sia p (½ nel caso di moneta non truccata) la probabilità di testa e (1-p) la probabilità di croce.
Anche nel caso della variabile binomiale vogliamo rappresentare la legge di distribuzione ad esempio nel caso in cui n=10.

Lanciamo quindi 10 volte una moneta; p è la probabilità di ottenere testa in ogni lancio.
Calcoliamo ora queste probabilità:

Probabilità che il numero di successi/testa sia uguale a $0=(1-p)^{10}$
Probabilità che il numero di successi/testa sia uguale a $1=10^{*}\ p^{*}(1-p)^{9}$
Probabilità che il numero di successi/testa sia uguale a $2=C(10,2)^{*}p^{2*}(1-p)^{8}$
Probabilità che il numero di successi/testa sia uguale a $3=C(10,3)^{*}p^{3*}(1-p)^{7}$
Probabilità che il numero di successi/testa sia uguale a $4=C(10,4)^{*}p^{4}(1-p)^{6}$
…...
Probabilità che il numero di successi/testa sia uguale a $k=C(10,k)^{*}p^{k}(1-p)^{10-k}$
con k<10
…….
Probabilità che il numero di successi/testa sia uguale a $10=p^{10}$

Con C(10,k) indico le combinazioni di 10 elementi a classe k.

Schematizzo ad esempio la
Probabilità che il numero di successi/testa sia uguale a $3=C(10,3)^{*}p^{3*}(1-p)^{7}$

p	p	p	1-p	1-p	1-p	1-p	1-p	1-p	1-p

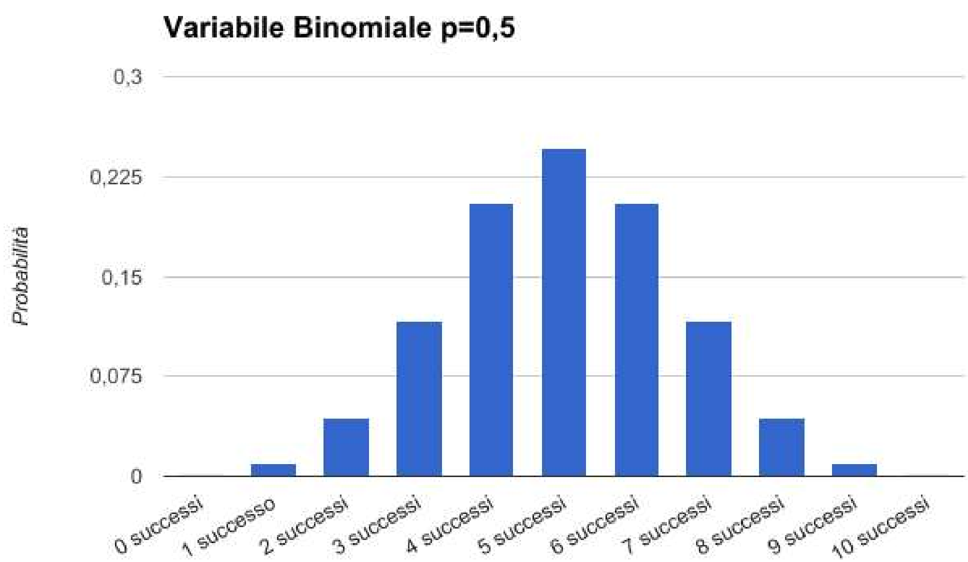
Una possibile configurazione favorevole prevede i successi sulle prime tre monete; sono configurazioni favorevoli anche quelle in cui i successi avvengono su tre monete qualsiasi in questa sequenza. I casi possibili sono quindi le combinazioni di 10 elementi a classe 3.

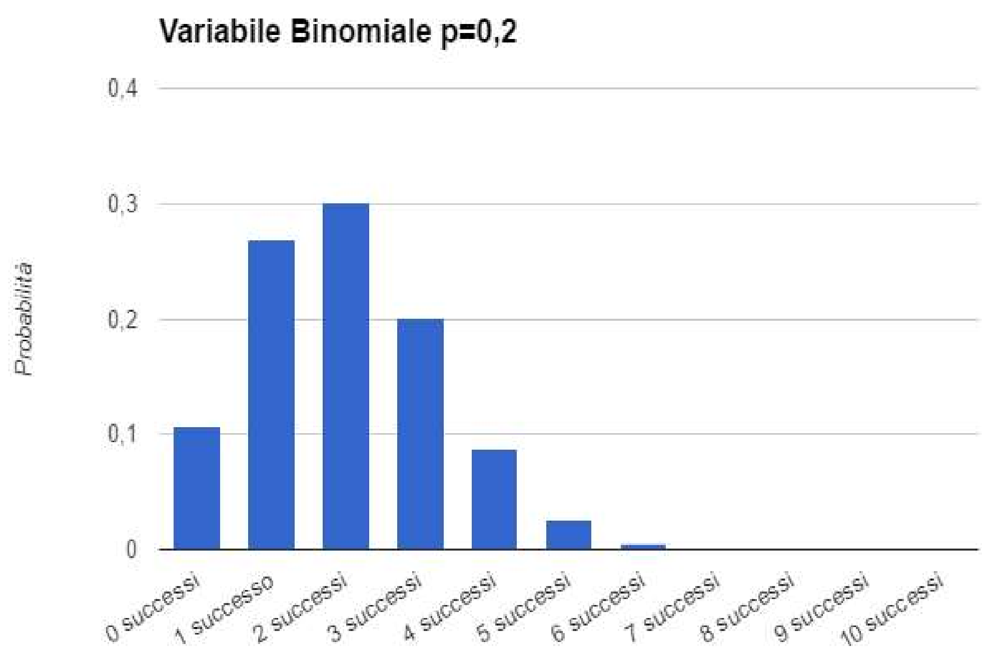
Un altro esempio di legge binomiale può essere il numero di guasti in un lotto di n elementi se si considera p la probabilità di guasto di ogni elemento o il numero di malati in una popolazione di n individui.

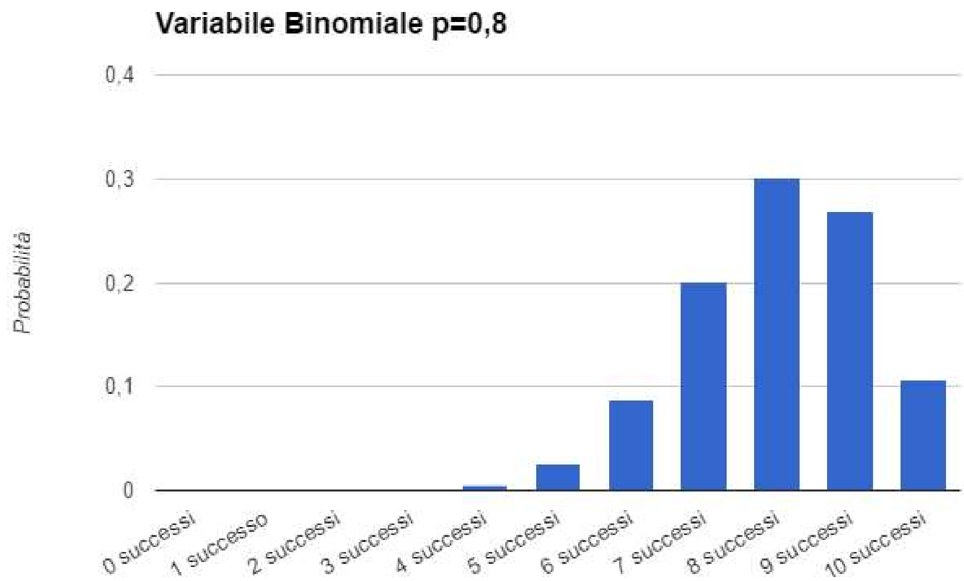
Qui sotto sono riportati tre grafici di probabilità di variabili binomiali, rispettivamente con questi parametri

n=10 e p=0,5
e

n=10 e p=0,2
e
n=10 e p=0,8







La somma di variabili aleatorie indipendenti è una variabile aleatoria con distribuzione “a campana” come nei grafici sopra illustrati.

Per ribadire quest’ultimo concetto e per comprendere il significato del teorema del limite centrale, che esporrò più avanti, propongo il seguente quesito.

Quesito/Esercitazione 1

Utilizzando il foglio elettronico costruire almeno 35 colonne di numeri casuali compresi tra 0 e 10. Ogni colonna deve essere costituita da almeno 100 righe di numeri casuali. Ad esempio la prima colonna conterrà numeri casuali tra le celle A2 e A101 la seconda colonna conterrà numeri casuali tra B2 e B101 e così via. La prima riga deve essere riservata ai titoli: var1, var2, var3…… Ogni colonna simula quindi una variabile uniforme. Sommando le 35 variabili uniformi e facendo la media sulle 35 variabili uniformi si ottengono variabili aleatorie con distribuzione “a campana”. Verifica che la variabile aleatoria somma e media hanno una distribuzione a campana (Per visualizzare la distribuzione di probabilità della variabile variabile media ad esempio si devono calcolare le frequenze relative delle volte in cui si è ottenuto 1,2,........10).
NB quando fai la media dei numeri casuali arrotonda il risultato ottenuto alle unità.

Teorema del limite centrale

Se X1, X2, X3,.......Xn sono n (n>>30) variabili aleatorie con media μ e varianza σ^2

La media di queste variabili aleatorie presenta una distribuzione Gaussiana (“a campana”), simmetrica, con media μ, anche se le variabili X1,X2,X3,......Xn hanno una distribuzione di probabilità di altro tipo (uniforme/asimmetrica…….)

La variabile aleatoria media presenta una variabilità molto più bassa rispetto alla variabilità delle singole variabili:

La Varianza della media è σ^2 /n.

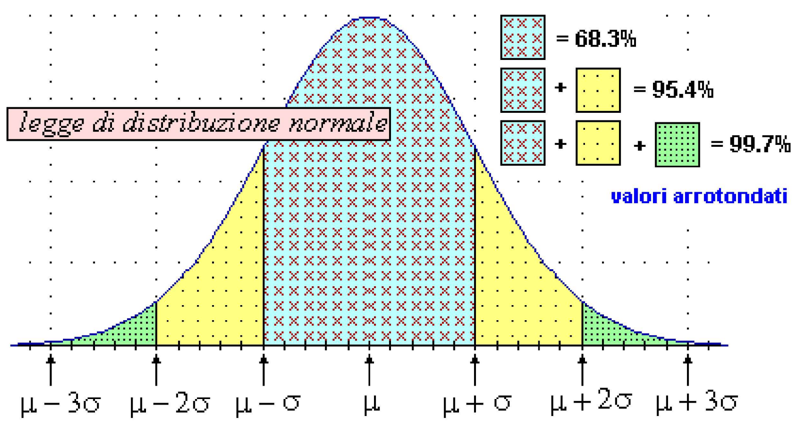
Questa distribuzione a campana garantisce che la maggior parte dei dati (delle medie) è molto vicina alla Media μ.

e precisamente

il 68,3% dei dati sono compresi tra (μ-σ ; μ+σ)

il 95,4% dei dati sono compresi tra (μ-2σ ; μ+2σ)

il 99,7% dei dati sono compresi tra (μ-3 ; μ+3σ)



Quesito 2

La tabella qui sotto riportata rappresenta il numero di errori di lettura registrati su un campione di prove di lettura riguardanti alunni esenti da problemi specifici di apprendimento.

4	3	5	2	0	5	3	2
4	5	0	0	0	5	3	2
4	5	0	3	2	5	2	3
5	8	2	3	4	2	3	2
0	3	4	3	2	4	1	1
2	5	6	2	3	1	2	0
5	3	1	0	2	4	7	3
4	1	0	2	3	0	1	1
3	3	5	2	1	3	0	1

Utilizzando il teorema del limite centrale costruire un intervallo di fiducia al 95% della media campionaria

Quesito 3

Un'azienda artigiana produce listelli di legno aventi lunghezza media 150 cm con scarto quadratico medio di 4 cm. Per verificare se la produzione è sotto controllo si estrae un campione casuale di 55 listelli da cui si rileva una lunghezza media di 149 cm.
Verificare se il campione cade nell'intervallo della media con probabilità rispettivamente del 70% e del 95%.

Quesito 4

Per controllare la produzione di tubi metallici un'impresa estrae dalla produzione un campione casuale di 60 tubi da cui rileva una lunghezza media di 55 cm. Dare una stima per intervallo della lunghezza media a livello di fiducia del 95% sapendo che lo scarto quadratico medio (o deviazione standard $\boldsymbol{\sigma}$) della produzione è 4 cm.

Quesito 5

Una macchina riempie automaticamente delle lattine di birra. Si estrae un campione di 60 lattine per un controllo e si rileva un contenuto medio di 33,2 cl.
Nella produzione lo scarto quadratico medio è $\boldsymbol{\sigma}$=2,5 (cl).
Determinare l'intervallo di fiducia del contenuto medio al livello del 95% e del 99%.