STATISTICA

I modelli probabilistici

Statistico, detec anziere?

LINGUAGGIO della PROBABILITA'

« Per risolvere i casi comp ati bisogna essere capaci di **costruire una storia**, partendo dagli **indizi disponibili**, che contenga una **spiegazione plausibile** di t tti gli elementi che abbi vuole una ce i fantasia ed è u simile a quello di uno scrittore. Una volta costruita questa storia che è, in sostanza, ur come potrebbero essersi svolti i fatti bisogna andare alla ricerca delle **conferme**.»

DATI

MODELLI

INFERENZA

(G. Carofiglio)

Dal campione alla popolazione

Il Genetics and IVF Institute di Fairfax in Virginia ha sviluppato una tecnica in grado, sostengono, di aumentare la probabilità per una coppia di avere una figlia femmina. **Di 14 coppie** su cui è stata messa in atto, 13 hanno avuto una femmina.

Possiamo affermare che tale tecnica aumenta la probabilità di avere una femmina per *qualunque* coppia?

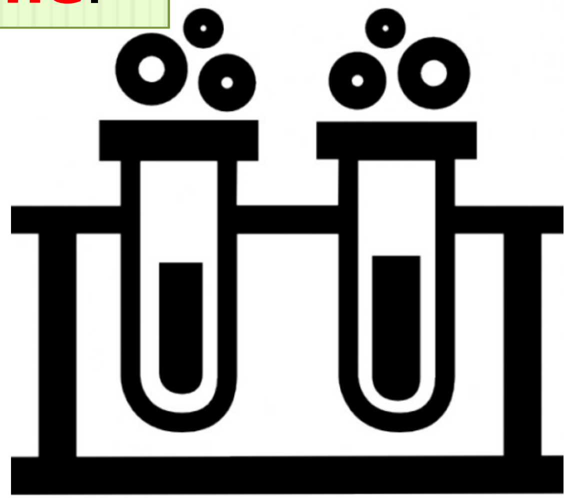
Dal campione alla popolazione

Durante le verifiche sulla qualità dell'acqua potabile di Milano si è misurata la quantità presente nell'acqua di arsenico, ottenendo i seguenti **valori campionari**: 15, 12.5, 7.0, 13.0, 7.5 $\mu\mathrm{g/L}$.

Sapendo che la legge impone una soglia massima per l'arsenico di 10 $\mu\mathrm{g/L}$, **si può affermare che *l'acqua potabile di Milano* non è a norma di legge**?

Dal campione alla popolazione

Ci serve un modo per capire se quello che vediamo sul **campione** vale anche sull'intera **popolazione**.



Dal campione alla popolazione

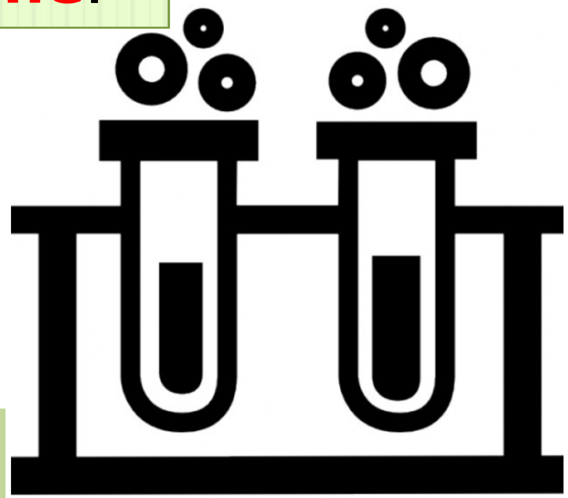
Tutti
i possibili
risultati

modello probabilistico

Ci serve un modo per capire se quello che vediamo sul **campione** vale anche sull'intera **popolazione**.

modello probabilistico

Uno dei tanti
possibili risultati

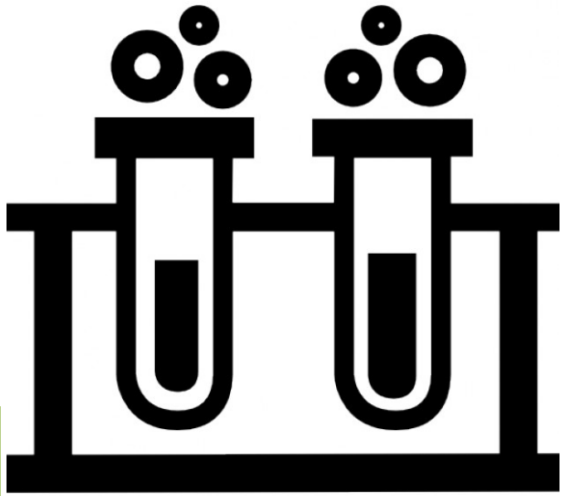


Dal campione alla popolazione

Tutti
i possibili
risultati

modello probabilistico

Il ***modello*** è una descrizione del meccanismo con cui si sviluppa il fenomeno (e, quindi, con cui sono prodotti i dati nel campione):



modello probabilistico

Uno dei tanti
possibili risultati

Dal campione alla popolazione

Tutti
i possibili
risultati

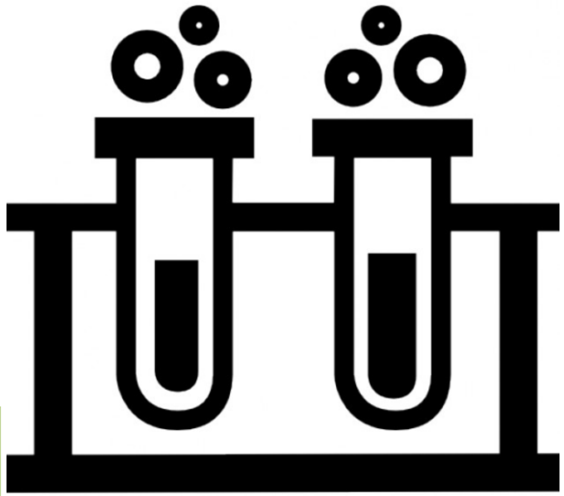
modello probabilistico

Il ***modello*** è una descrizione del meccanismo con cui si sviluppa il fenomeno (e, quindi, con cui sono prodotti i dati nel campione):

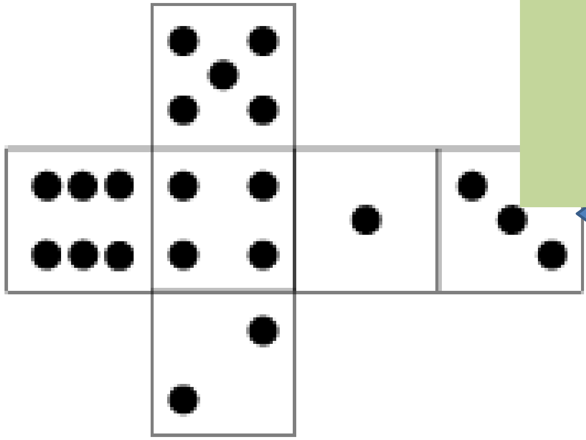
è in grado di darci informazioni sulle osservazioni *prima* di estrarre il campione (***osservazioni potenziali***)

modello probabilistico

Uno dei tanti
possibili risultati



Dal campione alla popolazione



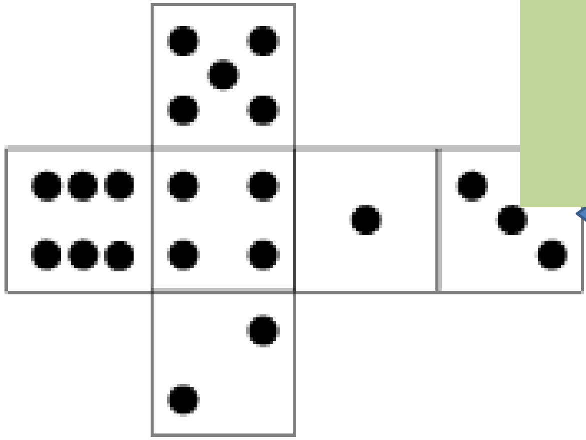
Tutti
i possibili
risultati

modello probabilistico

modello probabilistico

Uno dei tanti
possibili risultati: ***6***

Dal campione alla popolazione



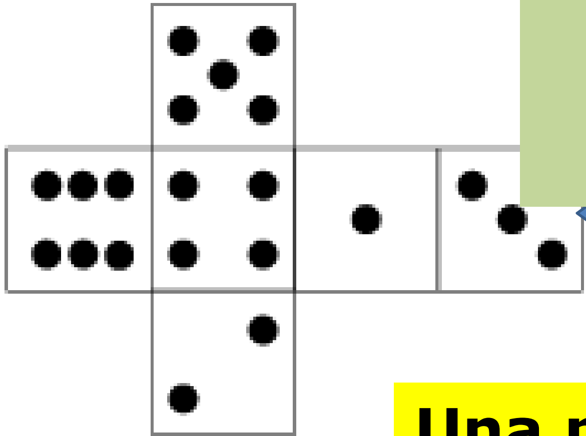
Tutti
i possibili
risultati

Modello: dado è equilibrato
(tutti i punteggi equiprobabili)

modello probabilistico

Uno dei tanti
possibili risultati: ***6***

Dal campione alla popolazione



Tutti
i possibili
risultati

Modello: dado è equilibrato
(tutti i punteggi equiprobabili)

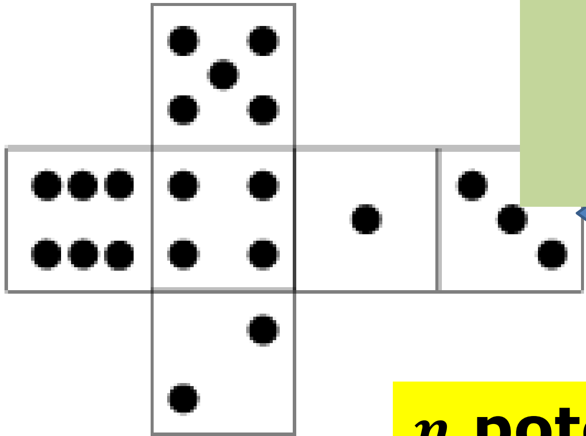
Una potenziale osservaz.: ***X***
variabile casuale che assume i valori
$1, 2, 3, 4, 5, 6$ e tale che

$$P(X = k) = \frac{1}{6}, \quad k = 1, \dots, 6$$

modello probabilistico

Uno dei tanti
possibili risultati: ***6***

Dal campione alla popolazione



Tutti
i possibili
risultati

Modello: dado è equilibrato
(tutti i punteggi equiprobabili)

$\boldsymbol{n}$ **potenziali osservaz.**: $X_1, X_2, \dots, X_n$
indipendenti (come i lanci) **e** tutte
identiche con distribuzione

$$P(X_i = k) = \frac{1}{6}, \qquad k = 1, \dots, 6$$

modello probabilistico

Uno dei tanti
possibili risultati: ***6***

Dal campione alla popolazione

Modello: dado è equilibrato (tutti i punteggi equiprobabili)

Una potenziale osservaz.: X variabile casuale che assume i valori $1, 2, 3, 4, 5, 6$ e tale che

$$P(X = k) = \frac{1}{6}, \quad k = 1, \dots, 6$$

$\boldsymbol{n}$ **potenziali osservaz.**: $X_1, X_2, \dots, X_n$ indipendenti (come i lanci) e tutte identiche con distribuzione

$$P(X_i = k) = \frac{1}{6}, \qquad k = 1, \dots, 6$$

Campione casuale osservato : esito di n lanci: $\{x_1, x_2, \dots, x_n\} = \{1,3,2,1,1,6,2,3,4,4, \dots, 6\}$

Variabili casuali discrete

Modello probabilistico discreto:

variabile casuale X :

$$\boldsymbol{X}: \quad (\cdots) \rightarrow \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\}, \quad \boldsymbol{N} \leq +\infty$$

distribuzione di X :

$$\boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x}_i) \text{ per tutti gli } \boldsymbol{x}_i$$

$$\mathbf{0} \leq \boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x}_i) \leq \mathbf{1}$$

$$\sum_{i=1}^{N} \boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x}_i) = \mathbf{1}$$

Variabile casuale **uniforme**

Modello probabilistico discreto:

variabile casuale X :

$\boldsymbol{X} : \quad (\cdots) \rightarrow \{\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}, \mathbf{5}, \mathbf{6}\}$

distribuzione di X :

$\boldsymbol{P(X = k) = 1/6}, \quad \boldsymbol{k = 1, \ldots, 6}$

$$0 \leq P(X = x_i) \leq 1$$

$$\sum_{i=1}^{N} P(X = x_i) = 1$$

Variabile casuale **bernoulliana**

Modello probabilistico discreto:

variabile casuale X :

$\boldsymbol{X}: \quad (\cdots) \rightarrow \{\mathbf{0}, \mathbf{1}\}$

(numero di *successi* in una singola prova)

distribuzione di X :

$\boldsymbol{P(X=1)=p}, \qquad \boldsymbol{P(X=0)=1-p}$

$$0 \leq P(X=x_i) \leq 1$$

$$\sum_{i=1}^{N} P(X=x_i) = 1$$

Variabile casuale **bernoulliana**

Modello probabilistico discreto:

variabile casuale X :

$\boldsymbol{X}: \quad (\cdots) \to \{\mathbf{0}, \mathbf{1}\}$

(numero di *successi* in una singola prova)

distribuzione di X :

$\boldsymbol{P(X = 1) = 0.5}, \qquad \boldsymbol{P(X = 0) = 0.5}$

Variabile casuale **bernoulliana**

Modello probabilistico discreto:

variabile casuale X :

$\boldsymbol{X} : \quad (\cdots) \to \{\mathbf{0}, \mathbf{1}\}$

(numero di *successi* in una singola prova)

distribuzione di X :

$\boldsymbol{P(X = 1) = 1/6}, \qquad \boldsymbol{P(X = 0) = 5/6}$

Variabile casuale **bernoulliana**

Modello probabilistico discreto:

variabile casuale X :

$\boldsymbol{X}: \quad (\cdots) \rightarrow \{\mathbf{0}, \mathbf{1}\}$

(numero di *successi* in una singola prova)

distribuzione di X :

$$\boldsymbol{P(X=1)=p}, \qquad \boldsymbol{P(X=0)=1-p}$$

La variabile bernoulliana

Il Genetics and IVF Institute di Fairfax in Virginia ha sviluppato una tecnica in grado, sostengono, di aumentare la probabilità per una coppia di avere una figlia femmina. Di 14 coppie su cui è stata messa in atto, 13 hanno avuto una femmina.

X variabile casuale che dice il sesso del neonato: 1=F, 0=M

$$P(X=1)=P(X=0)=\frac{1}{2}, normalmente$$

$X_1, X_2, \dots, X_{14}$ variabili casuali indipendenti ed identiche a X che dicono il sesso dei neonati di 14 coppie

Campione casuale osservato: esito di 14 parti: $\{x_1, x_2, \dots, x_{14}\} = \{1,1,1,1,1,1,1,1,1,1,1,1,1,0\}$

La variabile bernoulliana

Il Genetics and IVF Institute di Fairfax in Virginia ha sviluppato una tecnica in grado, sostengono, di aumentare la probabilità per una coppia di avere una figlia femmina. Di 14 coppie su cui è stata messa in atto, 13 hanno avuto una femmina.

X variabile casuale che dice il sesso del neonato: 1=F, 0=M

$$P(X=1)=p, \qquad P(X=0)=q=1-p, \qquad con\ qualche\ tecnica$$

X è una variabile bernoulliana di *parametro* $p \in (0,1)$

$X_1 + X_2 + \ldots + X_{14}$ conta il numero di F (*successi*) nei 14 parti (*prove*)

$(1+1+\cdots+1+0=13)$

$S_n = X_1 + X_2 + \ldots + X_n$ è una variabile **Binomiale**: $Bin(n,p)$

se possiamo supporre che le variabili siano indipendenti ed identiche

Variabile casuale **Binomiale**

$S_n = X_1 + X_2 + \ldots + X_n$ è una variabile **Binomiale**: $Bin(n, p)$

Modello probabilistico discreto:

S_n: $(\cdots) \rightarrow \{\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots, \boldsymbol{n}\}$

(numero di *successi* in n prove indipendenti ed identiche)

distribuzione di S_n:

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(S_n = k) = \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}$$

Variabile casuale Binomiale

$S_n = X_1 + X_2 + \ldots + X_n$ è una variabile **Binomiale**: $Bin(n, p)$

Modello probabilistico discreto:

$S_n: \quad (\cdots) \rightarrow \{\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots, \boldsymbol{n}\}$

(numero di *successi* in n prove indipendenti ed identiche)

distribuzione di S_n:

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(S_n = k) = \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}$$

Se $p = 0.5, n = 14 \Rightarrow P(S_{14} = 13) = \frac{14!}{13!(14-13)!} 0.5^{13} (1-0.5)^{14-13} = 14 \times 0.5^{14} =$

$0.0008544922 \approx 0.001$

Variabile casuale Binomiale

$S_n = X_1 + X_2 + \ldots + X_n$ è una variabile **Binomiale**: $Bin(n,p)$

Modello probabilistico discreto:

$S_n: \quad (\cdots) \to \{\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots, \boldsymbol{n}\}$

(numero di *successi* in n prove indipendenti ed identiche)

distribuzione di S_n:

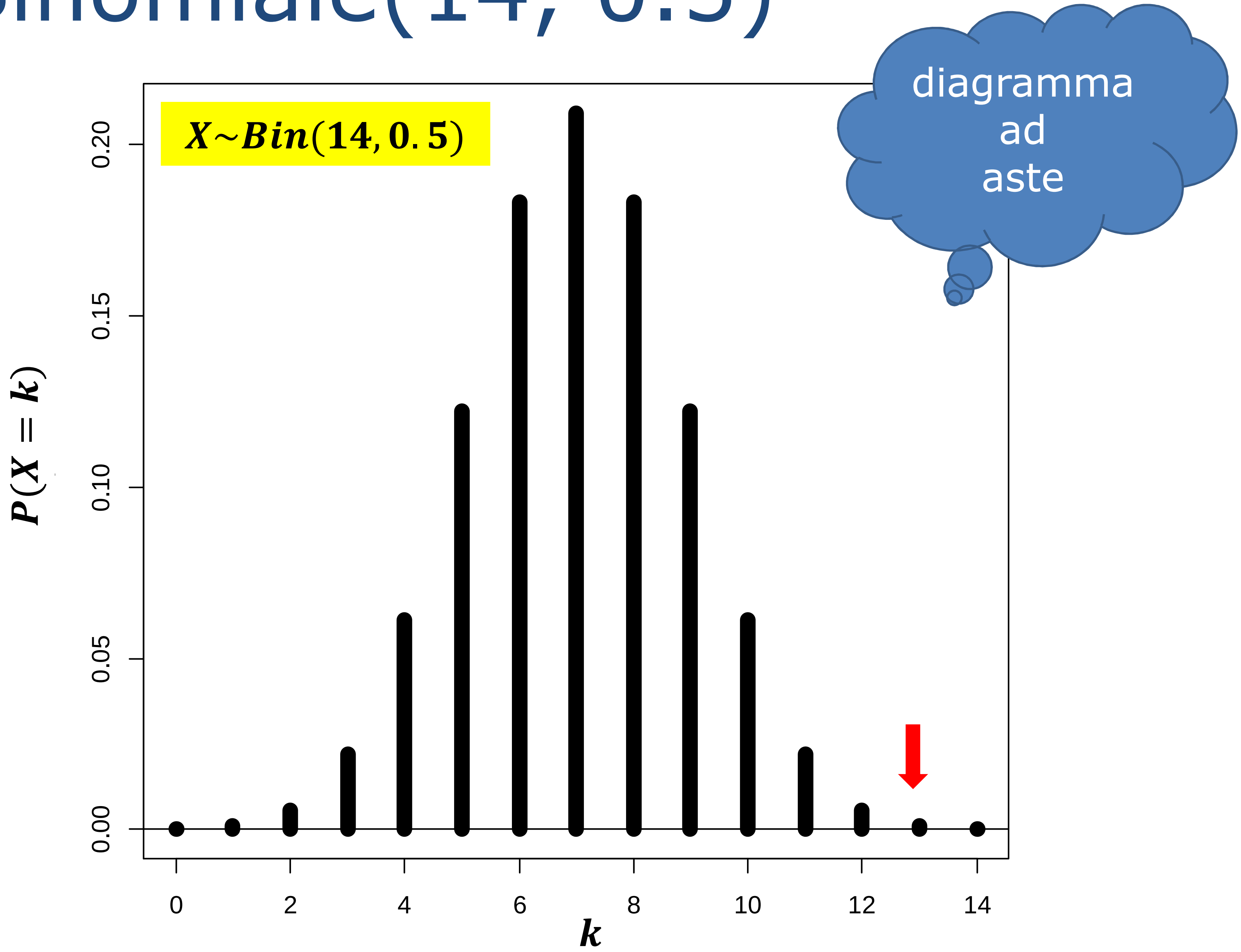
$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(S_n = k) = \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}$$

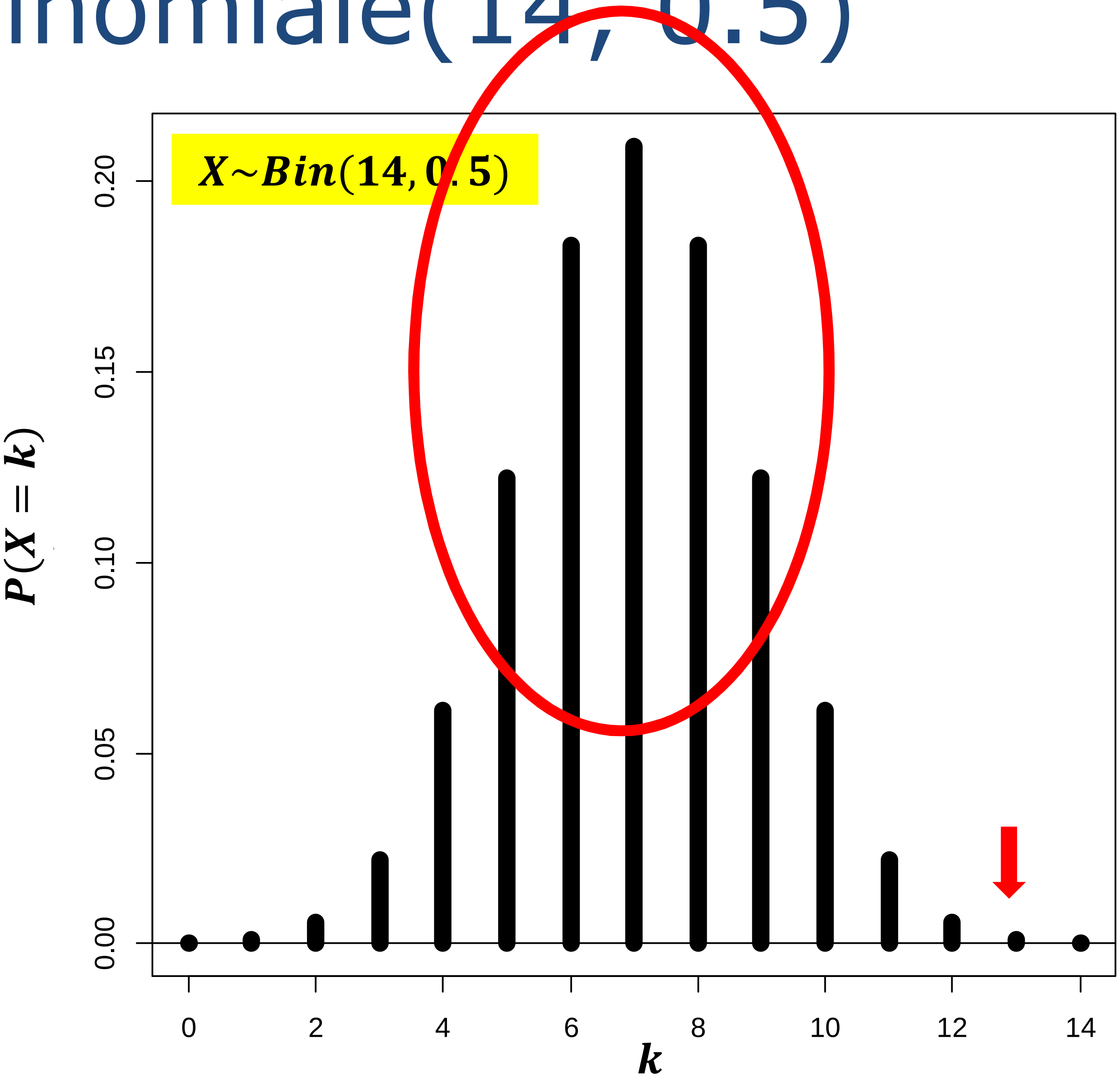
Se $p = 0.5, n = 14 \Rightarrow P(S_{14} = 13) = \frac{14!}{13!(14-13)!} 0.5^{13}(1-0.5)^{14-13} = 14 \times 0.5^{14} = 0.0008544922 \approx 0.001$

Se la tecnica fosse inefficace, sarebbe altamente improbabile avere 13 F su 14 parti!

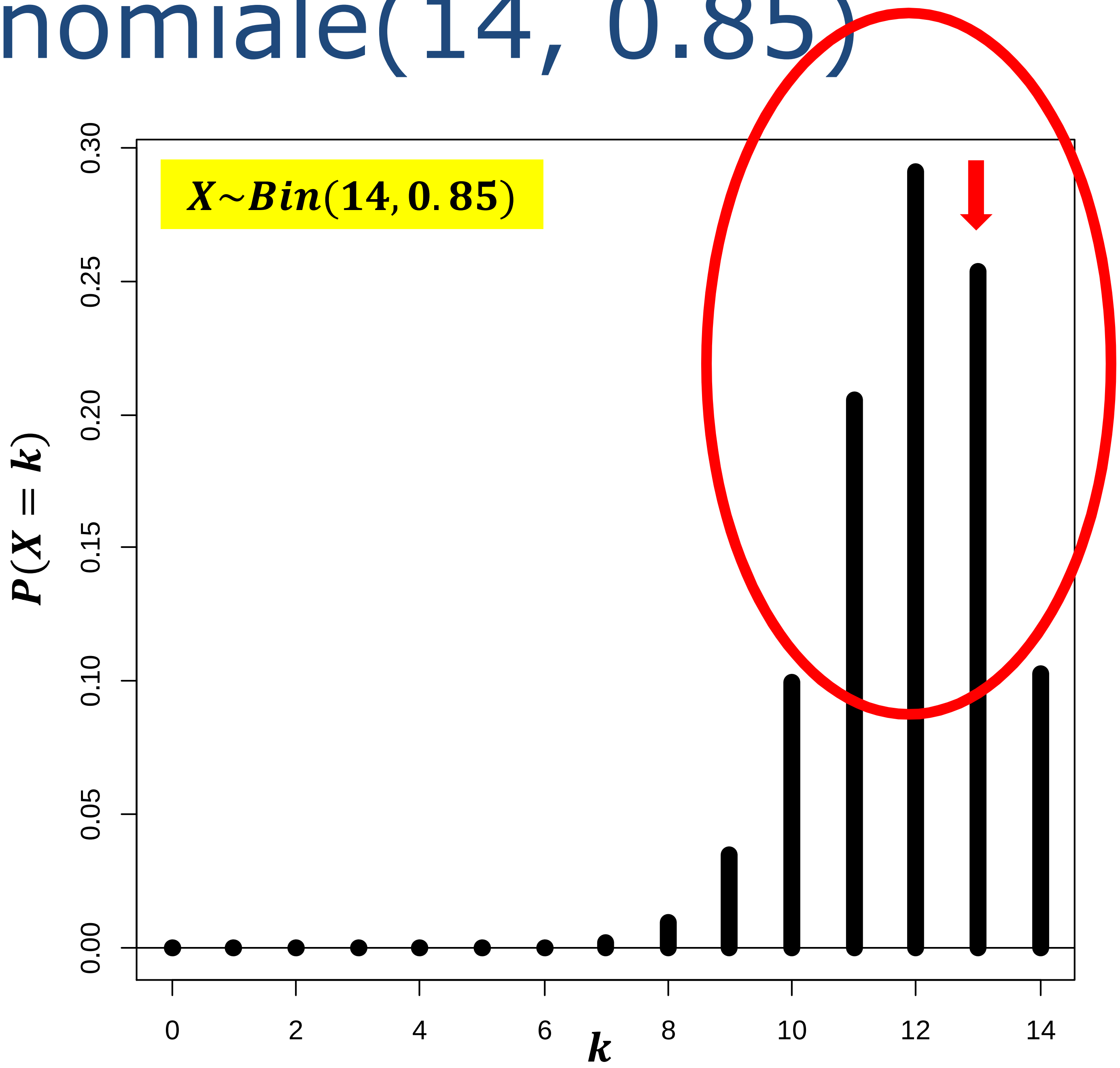
Binomiale(14, 0.5)



Binomiale(14, 0.5)



Binomiale(14, 0.85)



L'avevamo già incontrata!

Un ricercatore di genetica vi ha insospettito: ogni volta che sceglie a caso 20 bambini, il campione da lui scelto ha 10 F e 10M. Secondo lui è normale.

c) Qual è la probabilità di ottenere 10M e 10F su 20 neonati?

$$P(M) = P(F) = 0.5$$

$$\Downarrow$$

$$\frac{\binom{20}{10}}{2^{20}}$$

tutti i modi in cui posso mettere esattamente 10 caselle rosa, come da b)

tutti i modi con cui posso riempire la riga sopra di azzurro e rosa, come da punto a).

$$= \binom{20}{10} \times \left(\frac{1}{2}\right)^{20} = 0.176$$

L’avevamo già incontrata!

Un ricercatore di genetica vi ha insospettito: ogni volta che sceglie a caso 20 bambini, il campione da lui scelto ha 10 F e 10M. Secondo lui è normale.

c) Qual è la probabilità di ottenere 10M e 10F su 20 neonati?

$P(M) = P(F) = 0.5$

$$P(10F\ tra\ 20\ bimbi) = \binom{20}{10} \times \left(\frac{1}{2}\right)^{20} = \binom{20}{10} \times \left(\frac{1}{2}\right)^{10} \times \left(\frac{1}{2}\right)^{10} = 0.176$$

X variabile casuale che conta il numero di F in 20 bambini scelti a caso: $X \sim Bin(20, 0.5)$ e $= P(X = 10)$

L’avevamo già incontrata!

Un ricercatore di genetica vi ha insospettito: ogni volta che sceglie a caso 20 bambini, il campione da lui scelto ha 10 F e 10M. Secondo lui è normale.

c) Qual è la probabilità di ottenere 10M e 10F su 20 neonati?

$\boldsymbol{P(M) = 0.52, P(F) = 0.48}$

$\Downarrow$

$$\frac{\binom{20}{10}}{2^{20}}$$

tutti i modi in cui posso mettere esattamente 10 caselle rosa, come da b)

tutti i modi con cui posso riempire la riga sopra di azzurro e rosa, come da punto a).

$$= \binom{20}{10} \times \left(\frac{1}{2}\right)^{20} = 0.176$$

L’avevamo già incontrata!

Un ricercatore di genetica vi ha insospettito: ogni volta che sceglie a caso 20 bambini, il campione da lui scelto ha 10 F e 10M. Secondo lui è normale.

d) Sulla base di questi risultati, siamo d’accordo col ricercatore che 10M e 10F è normale? No!

$\boldsymbol{P(M) = 0.52,} \qquad \boldsymbol{P(F) = 0.48}$

$$P(10F\ tra\ 20\ bimbi) = \binom{20}{10} \times (0.48)^{10} \times (0.52)^{10} = 0.173\ (cfr.\, 0.176)$$

X variabile casuale che conta il numero di F in 20 bambini scelti a caso: $X \sim Bin(20, 0.48)$ e $= P(X = 10)$

L’avevamo già incontrata!

Un ricercatore di genetica vi ha insospettito: ogni volta che sceglie a caso 20 bambini, il campione da lui scelto ha 10 F e 10M. Secondo lui è normale.

d) Sulla base di questi risultati, siamo d’accordo col ricercatore che 10M e 10F è normale? No!

$\boldsymbol{P(M) = 0.20}$, $\boldsymbol{P(F) = 0.80}$

$$P(10F\ tra\ 20\ bimbi) = \binom{20}{10} \times (0.80)^{10} \times (0.20)^{10} = 0.002\ (cfr.\, 0.176)$$

X variabile casuale che conta il numero di F in 20 bambini scelti a caso: $X \sim Bin(20, 0.80)$ e $= P(X = 10)$

L'avevamo già incontrata!

Un ricercatore di genetica vi ha insospettito: ogni volta che sceglie a caso 20 bambini, il campione da lui scelto ha 10 F e 10M. Secondo lui è normale.

d) Sulla base di questi risultati, siamo d'accordo col ricercatore che 10M e 10F è normale? No!

$\boldsymbol{P(M) = 0.40}$, $\boldsymbol{P(F) = 0.60}$

$$P(10F\ tra\ 20\ bimbi) = \binom{20}{10} \times (0.60)^{10} \times (0.40)^{10} = 0.117\ (cfr.\ 0.176)$$

X variabile casuale che conta il numero di F in 20 bambini scelti a caso: $X \sim Bin(20, 0.60)$ e $= P(X = 10)$

Binomiale

Un test di ammissione ad un corso molto selettivo consiste di 15 domande a risposta multipla, con 3 risposte ciascuna di cui 1 sola corretta. Si è ammessi al corso solo rispondendo correttamente ad almeno 13 domande. Calcolare la probabilità di essere ammessi al corso rispondendo a caso a ciascuna domanda.

Binomiale

Un test di ammissione ad un corso molto selettivo consiste di 15 domande a risposta multipla, con 3 risposte ciascuna di cui 1 sola corretta. Si è ammessi al corso solo rispondendo correttamente ad almeno 13 domande. Calcolare la probabilità di essere ammessi al corso rispondendo a caso a ciascuna domanda.

X = numero di risposte esatte

$P(X \geq 13) = ?$

risposte **indipendenti**,
la prob. di indovinare è la stessa ad ogni domanda

Binomiale

Un test di ammissione ad un corso molto selettivo consiste di 15 domande a risposta multipla, con 3 risposte ciascuna di cui 1 sola corretta. Si è ammessi al corso solo rispondendo correttamente ad almeno 13 domande. Calcolare la probabilità di essere ammessi al corso rispondendo a caso a ciascuna domanda.

X = numero di risposte esatte

$P(X \geq 13) = ?$

$$X \sim Binomiale\left(15, \frac{1}{3}\right)$$

risposte **indipendenti**,
la prob. di indovinare è la stessa ad ogni domanda

$$P(X \geq 13) = P(X = 13) + P(X = 14) + P(X = 15) = \sum_{k=13}^{15} \binom{15}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{15-k}$$

$$\frac{15 \times 14}{2}\left(\frac{1}{3}\right)^{13}\left(\frac{2}{3}\right)^{2} + 15\left(\frac{1}{3}\right)^{14}\left(\frac{2}{3}\right)^{1} + \left(\frac{1}{3}\right)^{15}\left(\frac{2}{3}\right)^{0} = 0.0000314$$

Esercizio 5 p. 128

(testo adattato nella traduzione)

Si supponga che in un test per valutare l'efficacia di una tecnica di selezione del sesso sia stato scelto un campione casuale di 14 nascite, di cui 9 sono risultate essere femmine.

a) determinare la probabilità di avere 9 femmine in 14 parti se F e M sono ugualmente probabili;
b) determinare la probabilità di avere almeno 9 femmine in 14 parti se F e M sono ugualmente probabili;
c) Quale tra a) e b) è rilevante per determinare se 9 F è un esito insolitamente alto?
d) Se le 9 femmine sono state ottenute dopo l'applicazione di una tecnica di selezione del sesso, direste che tale tecnica è efficace?

Esercizio 5 p. 128

9F su 14 parti.

$k = num.\,di\ F$	$P(S_{14} = k)$
0	0.000
1	0.001
2	0.006
3	0.022
4	0.061
5	0.122
6	0.183
7	0.209
8	0.183
9	0.122
10	0.061
11	0.022
12	0.006
13	0.001
14	0

Esercizio 5 p. 128

9F su 14 parti.

a) $P(S_{14} = 9) = 0.122$

$k = num.\, di\, F$	$P(S_{14} = k)$
0	0.000
1	0.001
2	0.006
3	0.022
4	0.061
5	0.122
6	0.183
7	0.209
8	0.183
9	0.122
10	0.061
11	0.022
12	0.006
13	0.001
14	0

Esercizio 5 p. 128

9F su 14 parti.

a) $P(S_{14} = 9) = 0.122$

b) "almeno 9": $P(S_{14} \geq 9) = 0.122 + 0.061$
$+0.022 + 0.006 + 0.001 + 0 = 0.212$

$k = num.\,di\ F$	$P(S_{14} = k)$
0	0.000
1	0.001
2	0.006
3	0.022
4	0.061
5	0.122
6	0.183
7	0.209
8	0.183
9	0.122
10	0.061
11	0.022
12	0.006
13	0.001
14	0

Esercizio 5 p. 128

9F su 14 parti.

a) $P(S_{14} = 9) = 0.122$

b) "almeno 9": $P(S_{14} \geq 9) = 0.122 + 0.061$
$+ 0.022 + 0.006 + 0.001 + 0 = 0.212$

c) => b)

$k = num.\, di\ F$	$P(S_{14} = k)$
0	0.000
1	0.001
2	0.006
3	0.022
4	0.061
5	0.122
6	0.183
7	0.209
8	0.183
9	0.122
10	0.061
11	0.022
12	0.006
13	0.001
14	0

Esercizio 5 p. 128

9F su 14 parti.

a) $P(S_{14} = 9) = 0.122$

b) "almeno 9": $P(S_{14} \geq 9) = 0.122 + 0.061$
$+0.022 + 0.006 + 0.001 + 0 = 0.212$

c) => b)

d) Non sembra che la tecnica sia particolarmente efficace: anche *naturalmente* potevamo avere 9 o più F con buona probabilità.

$k = num.\, di\ F$	$P(S_{14} = k)$
0	0.000
1	0.001
2	0.006
3	0.022
4	0.061
5	0.122
6	0.183
7	0.209
8	0.183
9	0.122
10	0.061
11	0.022
12	0.006
13	0.001
14	0

Media e varianza

Modello probabilistico discreto:

variabile casuale X :

$\boldsymbol{X} : \quad (\cdots) \rightarrow \{\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_N}\}, \quad \boldsymbol{N} \leq +\infty$

distribuzione di X :

$\boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x_i})$ per tutti gli $\boldsymbol{x_i}$

ricordatevi la media pesata

$$\mu = E(X) = \sum_{i=1,\ldots,N} x_i P(X = x_i)$$

$$\sigma^2 = Var(X) = \sum_{i=1,\ldots,N} (x_i - \mu)^2 P(X = x_i)$$

Media e varianza

Modello probabilistico discreto:

variabile casuale X :

$\boldsymbol{X}: \quad (\cdots) \rightarrow \{\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_N}\}, \quad \boldsymbol{N} \leq +\infty$

distribuzione di X :

$\boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x_i})$ per tutti gli $\boldsymbol{x_i}$

$$\mu = E(X) = \sum_{i=1,\dots,N} x_i P(X = x_i)$$

$$\sigma^2 = Var(X) = \sum_{i=1,\dots,N} (x_i - \mu)^2 P(X = x_i)$$

$$\sigma^2 = \sum_{i=1,\dots,N} {x_i}^2 P(X = x_i) - \mu^2$$

Esercizio "di compito"

La variabile discreta X ha la distribuzione di frequenza data nelle seguente tabella:

x	**0**	**1**	**2**	**10**	**20**
$P(X = x)$	0.55	0.05	0.05	0.05	0.30

a) Rappresentare la distribuzione con un opportuno grafico
b) Calcolare $P(X \geq 10)$ e $P(X > 5)$
c) Calcolare $E(X)$ e $Var(X)$

Soluzioni parziali: b) 0.35, 0.35; c) 6.65, 81.0275

Media e varianza

Modello probabilistico binomiale:

variabile casuale $\boldsymbol{X \sim Bin(n,p)}$:

$\boldsymbol{X}$: $\quad (\cdots) \rightarrow \{\mathbf{0}, \mathbf{1}, \ldots, \boldsymbol{n}\}\,, \quad \boldsymbol{N} \leq +\infty$

distribuzione di X :

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\mu = \boldsymbol{E(X) = np}$$

$$\sigma^2 = \boldsymbol{Var(X) = np(1-p)} = npq$$

Media e varianza

$\boldsymbol{X \sim Bin(n, p)}$

$$\mu = E(X) = np$$

$$\sigma^2 = Var(X) = np(1-p) = npq$$

$Se\ p = 0.5$, il numero medio di F che mi aspetto in 14 nascite è $14 \times 0.5 = 7$

$Se\ p = \frac{1}{3}$, il numero medio di risposte esatte che mi aspetto in 15 domande a risposta multipla, con 3 risposte possibili, è $15 \times \frac{1}{3} = 5$

Media e varianza

$\boldsymbol{X \sim Bin(n, p)}$

$$\mu = E(X) = np$$

$$\sigma^2 = Var(X) = np(1-p) = npq$$

$Se\ p = 0.5$, il numero medio di F che mi aspetto in 14 nascite è $14 \times 0.5 = 7$

$Se\ p = \frac{1}{3}$, il numero medio di risposte esatte che mi aspetto in 15 domande a risposta multipla, con 3 risposte possibili, è $15 \times \dfrac{1}{3} = 5$

$$\boxed{\mu \mp \sigma} \quad \left.\begin{array}{l} 7 \mp \sqrt{14 \times 0.5 \times 0.5} = 7 \mp 1.87 \\ \\ 5 \mp \ \ldots \end{array}\right\}$$

Media e varianza

$\boldsymbol{X \sim Bin(n, p)}$

$$\mu = E(X) = np$$

$$\sigma^2 = Var(X) = np(1-p) = npq$$

$Se\ p = 0.5$, il numero medio di F che mi aspetto in 14 nascite è $14 \times 0.5 = 7$

***range* di normalità**:

$$(\mu - 2\sigma, \mu + 2\sigma) \Rightarrow \left(7 - 2\sqrt{14 \times 0.5 \times 0.5}, 7 + 2\sqrt{14 \times 0.5 \times 0.5}\right) = (3.26, 10.74)$$

$\Rightarrow$ da 4 a 10 F rientrano nel range di normalità, senza trattamenti specifici.

Media e varianza

$\boldsymbol{X \sim Bin(n, p)}$

$$\mu = E(X) = np$$

$$\sigma^2 = Var(X) = np(1-p) = npq$$

$Se\ p = 0.5$, il numero medio di F che mi aspetto in 14 nascite è $14 \times 0.5 = 7$

range* di normalità**: ***impareremo a fare di meglio!

$$(\mu - 2\sigma, \mu + 2\sigma) \Rightarrow \left(7 - 2\sqrt{14 \times 0.5 \times 0.5}, 7 + 2\sqrt{14 \times 0.5 \times 0.5}\right) = (3.26, 10.74)$$

$\Rightarrow$ da 4 a 10 F rientrano nel range di normalità, senza trattamenti specifici.