Modello probabilistico

- S spazio campionario o spazio degli esiti = l’ insieme di tutti i possibili esiti di un esperimento

- Evento = ogni sottoinsieme di S

- Per ogni evento A è definita una probabilità P(A). Le probabilità soddisfano le tre proprietà
    1. $0 \leq P(A) \leq 1$
    2. $P(S)=1$
    3. $P(A \cup B)=P(A) + P(B)$ se A e B disgiunti

Esempio 1

Lancio di un dado. Determinare la probabilità che esca un valore ≤ 3.

$S=\{1, 2, 3, 4, 5, 6\}$

$A = \{1, 2, 3\}$

P(A)= # eventi favorevoli / # eventi possibili

$= 3/6 = 1/2 = 0{,}5$

$P(A) = P(E_1) + P(E_2) + P(E_3)$

Esempio 2

In una cassa di 12 bottiglie di vino, 3 sono di cattiva qualità. Scegliamo a caso 2 bottiglie. Determinare la probabilità che

(a) La prima bottiglia sia buona

(b)La seconda bottiglia sia buona

(c) Entrambe le bottiglie siano buone

(d)Una sia buona, una sia cattiva

Esempio 2

Numeriamo le bottiglie da 1 a 12. Le bottiglie 1, 2 e 3 sono di cattiva qualità.

- $S=\{(i, j),\ \ i, j=1,..,12,\ i \neq j\}$ $|S|= 12 \times 11=132$

(a) $A=\{ (i, j),\ \ i\neq 1, 2, 3,\ \ j$ qualsiasi$\}$ $|A|= 9 \times 11= 99$

(b) $B=\{ (i, j),\ \ i$ qualsiasi$,\ \ j \neq 1, 2, 3\}$ $|B| = 9 \times 11=99$

(c) $C=\{ (i,j)\ \ i, j = 4, \ldots, 12,\ i\neq j\ \}$ $|C|= 9 \times 8 = 72$

(d) $D = D_1 \cup D_2$

$D_1 = \{ (i, j)\ \ i= 4,\ldots,12,\ \ j = 1, 2, 3\}$

$D_1 = \{ (i, j)\ \ i= 1, 2, 3,\ \ j = 4,..,12\}$ $|D|= 2 \times 9 \times 3=54$

Esempio 3

Un’associazione sportiva ha 120 membri, di cui 44 giocano a tennis, 30 a calcio e 18 sia a tennis che a calcio. Sceglendo un membro a caso, detreminare la probabilità che

1) Non giochi a tennis
2) Non giochi a calcio
3) Non giochi né a tennis né a calcio

Probabilità condizionata e eventi indipendenti

Probabilità che si verifichi A sapendo che si è verificato B

$$P(A \mid B) = P(A \cap B) / P(B)$$

Due eventi A e B sono indipendenti se

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

Esempio 1

Si stima che il 30% degli adulti negli stati uniti siano obesi, che il 3% degli adulti siano diabetici e che il 2% siano sia obesi che diabetici. Determinare la probabilità condizionata che un individuo scelto casulamente

(a) Sia diabetico se è obeso

(b) Sia obeso se è diabetico

Esempio 2

La probabilità che un paziente muoia durante un'operazione a cuore aperto è 0,2. Un paziente che sopravvive in sala operatoria ha il 15% di probabilità di morire in ospedale per le consegunze dell'operazione.

Quale frazione di pazienti sopravvive sia all'operazione che alle conseguenze?

(Risultato: 68%)

Esempio 3

Determinare la probabilità di ciascuno dei seguenti esiti se una moneta onesta viene lanciata 6 volte

T T T T T T,   T C T C T C,   T T C T C T

Esempio 4

La moneta viene lanciata più volte fino a ottenere croce.

Qual è la probabilità che la moneta debba essere lanciata almeno 5 volte?

Esempio 5

Le analisi del sangue effettuate su un campione di 2000 individui hanno prodotto i seguenti risultati sui gruppi sanguigni:

0	48%	B	15%
A	31%	AB	6%

- Qual è la probabilità che il gruppo sanguigno di3 individui scelti casualmente 0?
- E che siano 2 del gruppo 0 e 1 del gruppo A?

Definizione frequentista di probabilità

- **Probabilità 0.5** significa che ci aspetta che l’evento di interesse “si verifichi metà delle volte su un gran numero di prove”.

- Si potrebbe sospettare che la probabilità che esca testa è 0.5 perché la moneta ha 2 facce, ma non basta (esistono le monete truccate).
- Analogamente anche i bambini possono essere di sesso M o F e le probabilità, calcolate statisticamente, non sono uguali: la probabilità di M è circa 0.51.

- Questa idea di probabilità è empirica.

La probabilità di un evento

In una serie di prove, ripetute un gran numero di volte ed eseguite tutte nelle stesse condizioni, la **frequenza relativa** tende ad assumere valori prossimi alla probabilità dell'evento stesso e l'approssimazione è tanto maggiore quanto più numerose sono le prove eseguite.

Esempio: Se negli ultimi 30 anni nella nostra città ha nevicato 18 volte, la probabilità che nevichi quest' anno è 18/30=3/5.

Esempio: Si vuole calcolare la probabilità che un neonato sia femmina. Su 100.000 nascite si sono avute 48.500 femmine. Essendo il numero di prove sufficientemente elevato ed ogni prova indipendente dall'altra, utilizziamo la <u>definizione frequentista:</u>
P(F) = 48500 / 100000 = 0,485 P(M) = 51500 / 100000 = 0,515

Esempio. Si è osservato **un campione** di 730 nidi di una particolare specie di uccello in una determinata foresta e si è costruita la distribuzione di frequenze del numero di uova per nido

N° uova	Frequenze	Frequenze relative
0	**90 nidi**	**0.12**
1	**165**	**0.23**
2	**209**	**0.29**
3	**187**	**0.26**
4	**67**	**0.09**
5	**12**	**0.01**
Totale	**730**	**1.00**

Distribuzioni di probabilità – Modelli probabilistici

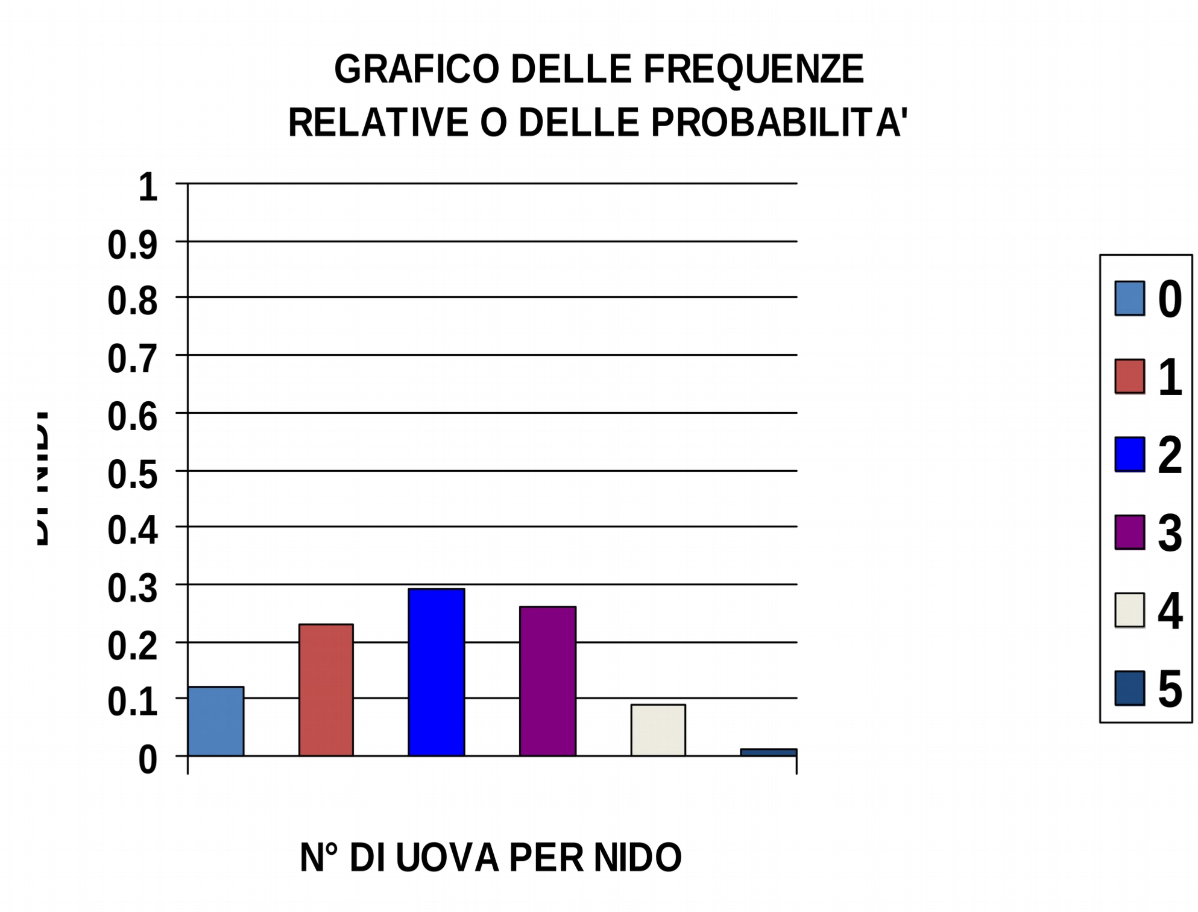
Distribuzioni di frequenze relative (**campione**)

Distribuzioni di probabilità (**popolazione**)

Esempio. Distribuzione del numero di uova per nido di una particolare specie di uccello in una foresta.

N° uova	Frequenze	Frequenze relative	Probabilità
0	**90**	**0.12**	**0.12**
1	**165**	**0.23**	**0.23**
2	**209**	**0.29**	**0.29**
3	**187**	**0.26**	**0.26**
4	**67**	**0.09**	**0.09**
5	**12**	**0.01**	**0.01**
Totale	**730**	**1.00**	**1.00**

Grafico a segmenti della distribuzione di probabilità (frequenze relative) della v.a. n° di uova per nido



Distribuzioni di probabilità discrete

- Grafico a segmenti delle frequenze relative del campione

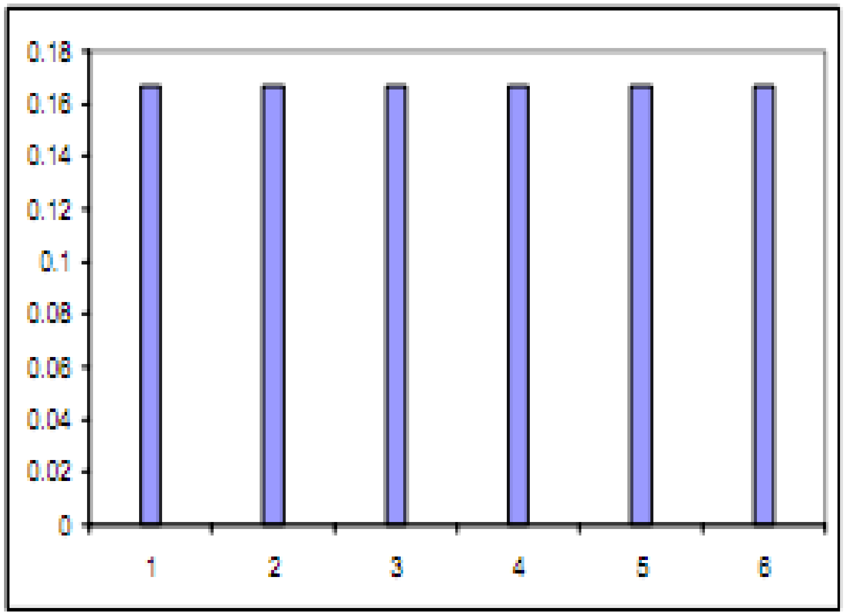
  Grafico della distribuzione di probabilità della variabile aleatoria discreta “n° di uova per nido”.

La distribuzione di probabilità serve a fare previsioni sull’intera popolazione.

Variabili aleatorie discrete

- Una variabile aleatoria o casuale X è una quantità di interesse determinata dal risultato di un esperimento casuale.
- Una v. a. discreta assume un numero finito o un'infinità numerabile di valori.
- E' completamente descritta quando sia nota la probabilità con cui si può verificare ciascun valore: $P(X = x_i) = p(x_i)$ con $\Sigma\, p(x_i) = 1$
- Media e varianza sono indici riassuntivi delle proprietà di tali variabili

Distribuzione uniforme



Distribuzione teorica di probabilità dei valori possibili che si possono ottenere nel lancio di un dado equilibrato: è discreta e uniforme

Distribuzione binomiale

- Siamo interessati alla v. a. discreta X che conta il numero di femmine in una famiglia di 3 figli.
- X=" il numero di successi su 3 prove"
- Il modello ideale per rappresentare tale situazione è dato dalla: distribuzione binomiale di parametri n, p, dove n=3 e p=0.5

Valori di X:	0	1	2	3
Probabilità di tali valori:	0.125	0.375	0.375	0.125

Distribuzione binomiale

Stiamo assumendo che:

- le 3 prove, ossia le 3 nascite sono indipendenti tra loro
- in ciascuna prova sono possibili 2 esiti (“successo”, “insuccesso”)
- In ciascuna prova è nota e costante la probabilità di successo p (la prob di insuccesso sarà $1-p$)

Nell’esempio assumiamo i due esiti equiprobabili

$$p = 0.5 \qquad\qquad 1-p = 0.5$$

Distribuzione binomiale

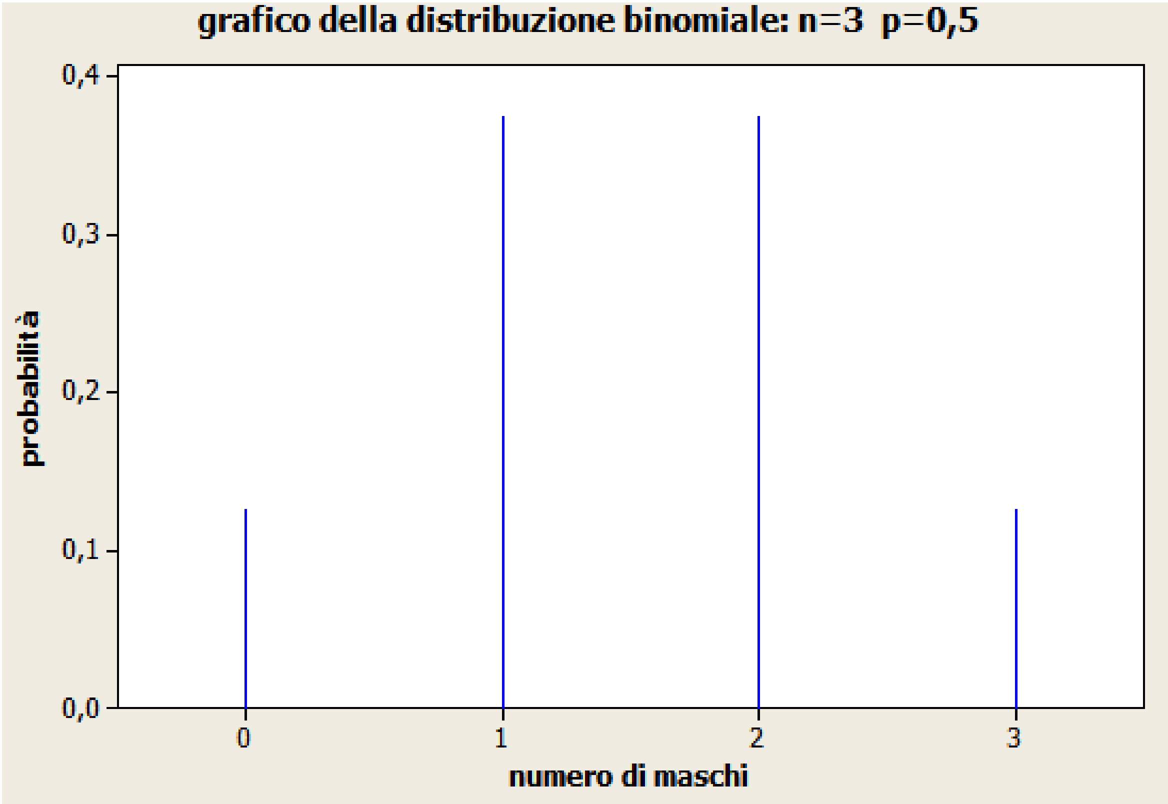
- X := numero di successi su n prove
- *P(X=k)* := probabilità che su n prove il n° di successi sia uguale a k

k= 0,1,…,n

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

Media = np   Varianza = np(1-p)

Grafico della distribuzione binomiale
di parametri $n = 3$ e $p = 0.5$



I parametri individuano in modo univoco la distribuzione binomiale

Esempio

Si suppone che la probabilità di sopravvivenza dei piccoli di tordo sia indipendente dal sesso e dal numero di uccellini presenti nel nido. Calcolare la probabilità che in un nido in cui sono sopravvissuti 5 uccellini il numero di femmine sia

1) 0

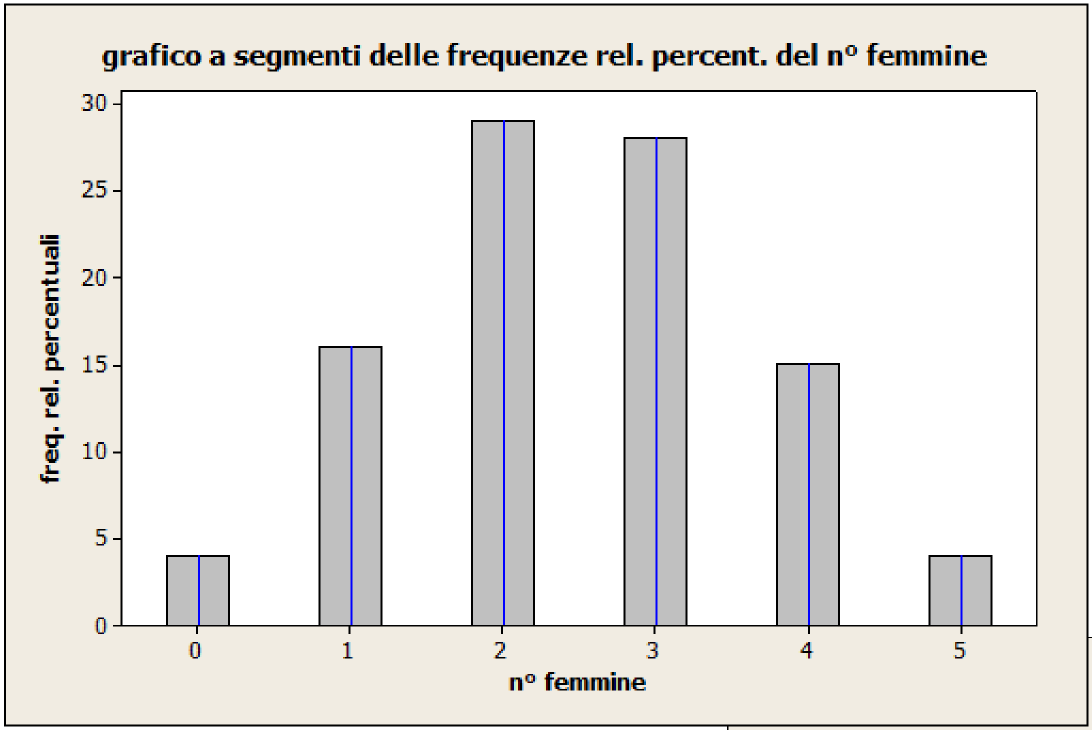
2) 5

3) 2

Distribuzione binomiale

Sono stati esaminati 480 nidi di tordo dove erano sopravissuti 5 uccellini. Si osserva il n° di femmine sopravissute in ogni nido.

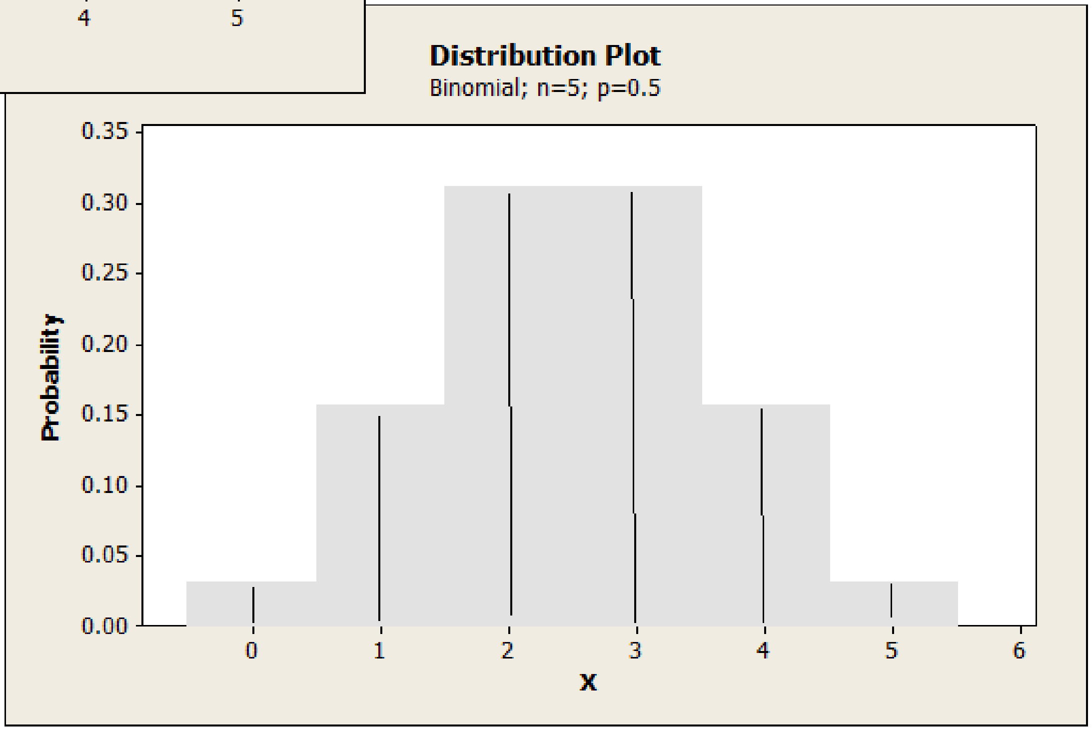
Femmine	maschi	Numero di nidi	frequenze relative
5	0	21	0.043
4	1	76	0.158
3	2	138	0.287
2	3	142	0.295
1	4	80	0.166
0	5	23	0.047



Esempio di modellizzazione

A confronto il grafico freq. rel. perc. **osservate** per il fenomeno e il grafico della Distrib.binom. **n=5; p= 0.5**

Ottimo modello



ESEMPIO

Una certa malattia ha un'evoluzione per cui non si conoscono terapie, tuttavia tra le persone colpite il 40% guarisce spontaneamente nell'arco di due mesi.

Non conoscendo particolarità della malattia, la possibilità di guarigione nell'arco di due mesi viene vista come puramente casuale.

- Con quale probabilità tra 6 persone colpite dalla malattia 2 guariranno spontaneamente nell'arco di due mesi?
- Qual è il numero medio di guarigioni spontanee? Quanto vale la varianza?
- Con quale probabilità nessuno guarirà spontaneamente?

ESEMPIO

Quattro bambini vengono vaccinati contro il morbillo. Il vaccino attecchisce con probabilità 0.8, garantendo l'immunità del bambino alla malattia.

- Con quale probabilità tutti i bambini risultano immunizzati?
- Se 100 bambini vengono vaccinati, qual è il numero medio di bambini immunizzati?
- Quanto vale la varianza?

Variabile aleatoria di Poisson

- Assume valori interi non negativi 0,1,2,…
- Approssima una binomiale di parametri (n,p) quando n è molto grande e p molto piccolo

Esempi

- Il numero di refusi in un libro
- Il numero di individui in una regione che raggiungono i cento anni d'età
- Il numero di transistor che si guastano nel primo giorno di utilizzo
- La quantità di particelle alfa emesse in un periodo di tempo fissato da un campione di materiale radioattivo

Variabile aleatoria di Poisson

$$P(X = k \quad successi) = e^{-\lambda}\frac{\lambda^k}{k!}$$

X v.a. di Poisson di parametro λ

- X può assumere tutti I valori interi non negativi
- λ è il numero medio di successi

E(X)= λ; Var(X)=λ

Esempio

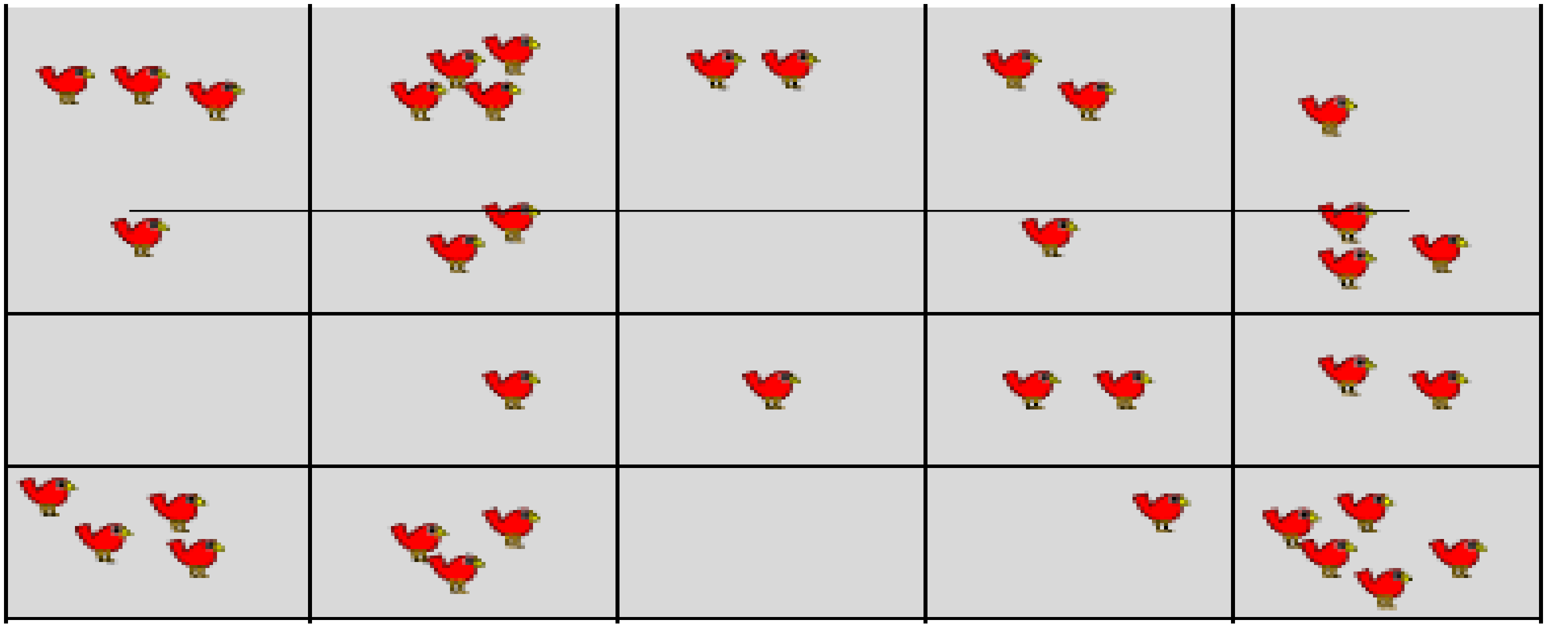
Campione di 20 nidi di una specie di uccelli. Si conta il numero di piccoli sopravvissuti per nido.

Il numero totale di piccoli sopravvissuti risulta pari a 38.

Il numero medio di piccoli sopravvissuti per nido è 38/20.

Interpretiamo il numero di sopravvissuti in un nido come una variabile di Poisson di parametro $\lambda=38/20=1,9$

Variabile aleatoria di Poisson



x	n_i	n_i x
0	3	0
1	6	6
2	5	10
3	3	9
4	2	8
5	1	5
TOTALI	**20**	**38**

Variabile aleatoria di Poisson

La distribuzione di Poisson descrive il numero di successi in intervalli (o blocchi) spaziali, volumetrici, o temporali quando

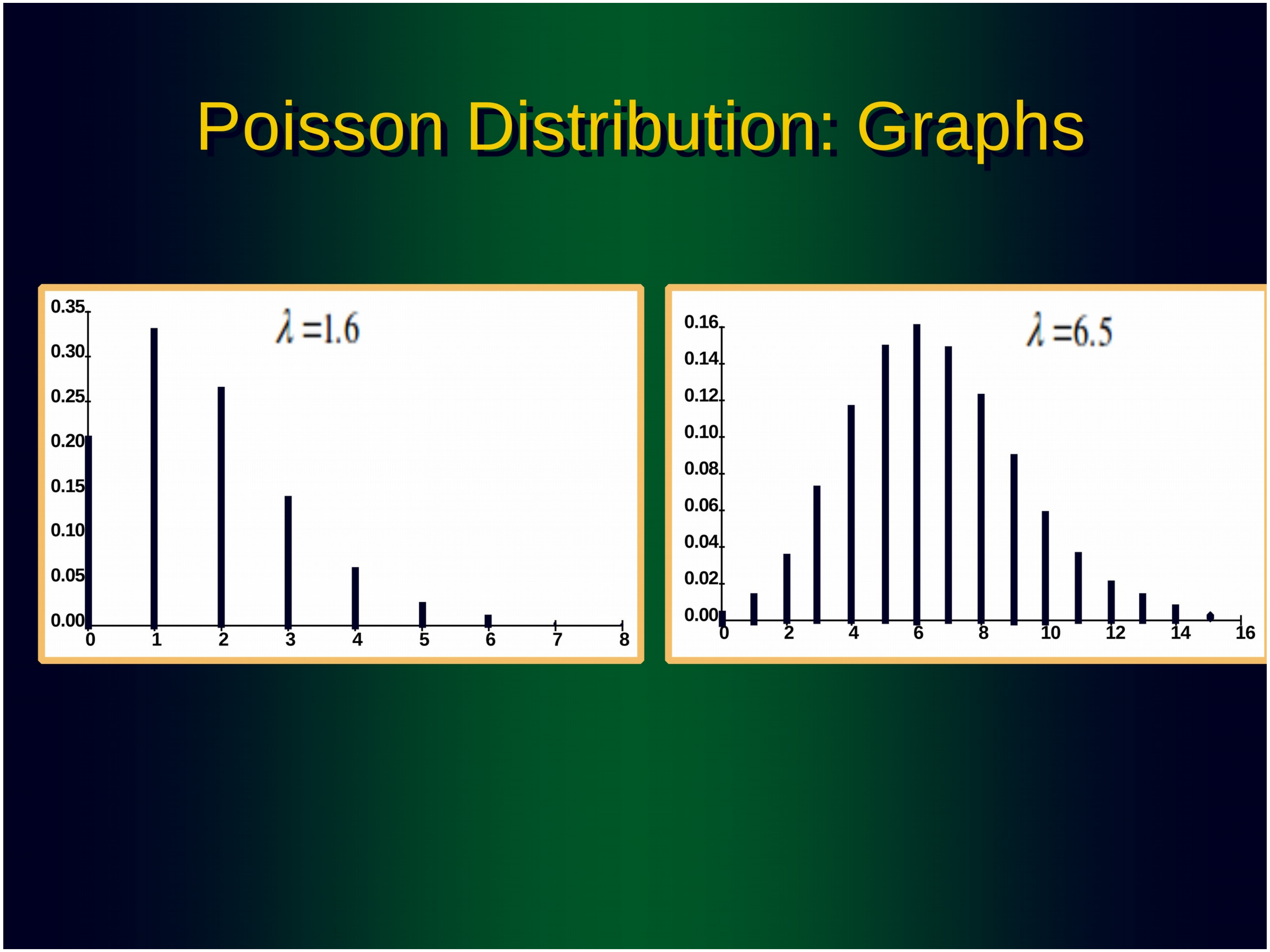
- i successi avvengono indipendentemente l'uno dall'altro
- i successi hanno la stessa probabilità di verificarsi in ogni punto dello spazio, di volume, o di tempo.

Variabile aleatoria di Poisson

<u>Esempi</u>

- numero di semi di una pianta infestante in un certo volume di terriccio
- numero di mutazioni in un certo intervallo di tempo
- numero di casi di influenza in un paese in una settimana
- numero di incidenti stradali mortali in un mese in una città
- numero di pezzi difettosi in una giornata di produzione
- numero di cetacei presenti in un tratto di mare

In tutti questi casi si può immaginare che <u>nell'area, volume</u>, <u>tempo</u> analizzati ci sia la possibilità teorica di osservare un numero elevatissimo di eventi tipo "presenza", ma che quelli realmente osservati siano invece "rari".

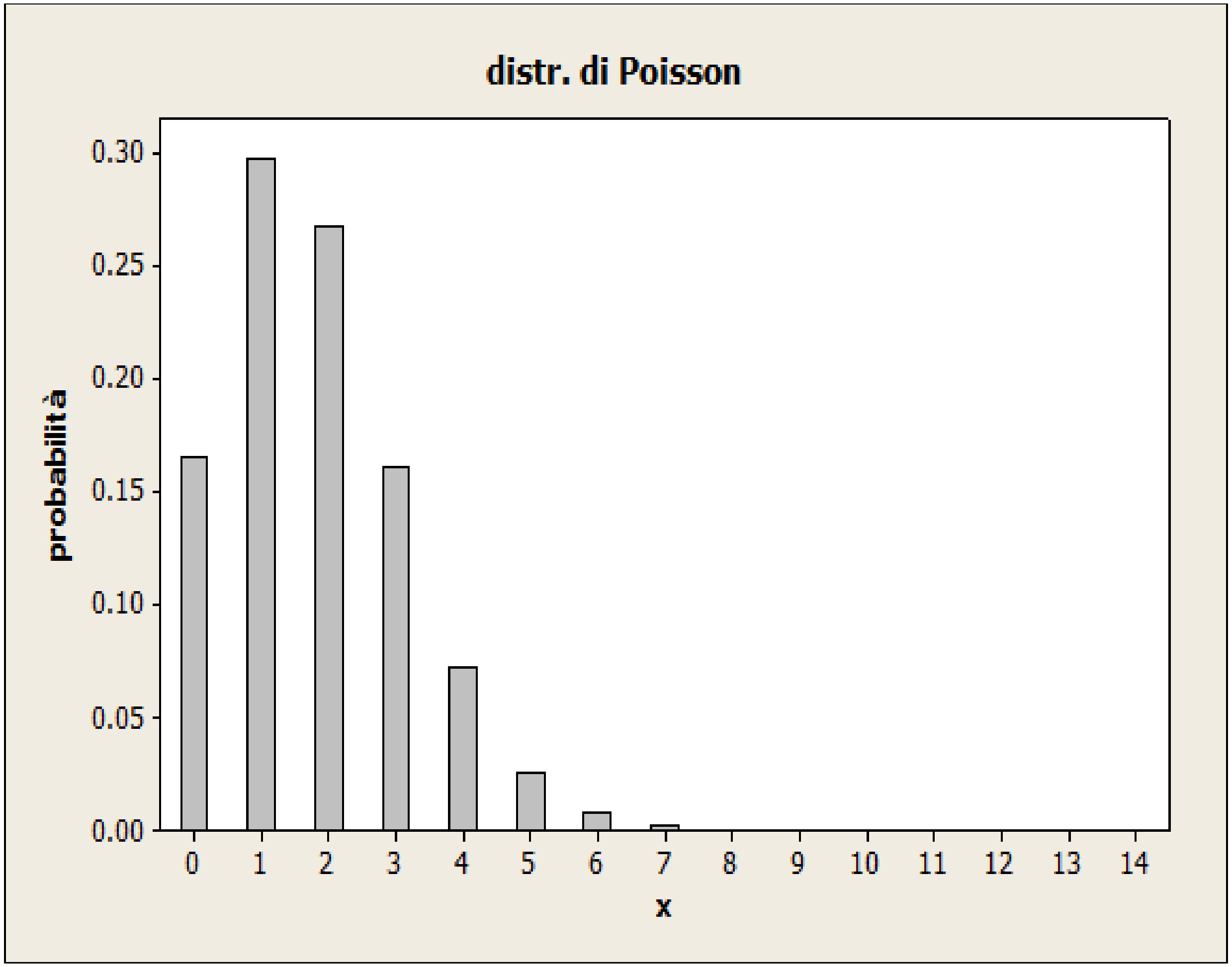


Al crescere del parametro λ la distribuzione diventa sempre più simmetrica e ha il massimo nel punto λ.

Esempio

In un ospedale le nascite avvengono casualmente e ci sono mediamente 1.8 nascite all'ora.
Qual è la prob di osservare 4 nascite fra le 21 e le 22 di un qualsiasi giorno?

Grafico relativo all'esempio precedente



parametro λ =1.8

x	p(x)
0	0.165299
1	0.297538
2	0.267784
3	0.160671
4	0.072302
5	0.026029
6	0.007809
7	0.002008
8	0.000452
9	0.000090

Esempi

1) Per analizzare le tracce delle bombe V-I della seconda guerra mondiale, la zona meridionale di Londra è stata suddivisa in 576 regioni , ognuna delle quali di area 0.25 Km^2. Un totale di 535 bombe ha colpito l'area delle 576 regioni.

   Qual è la prob che, una regione scelta a caso, sia stata colpita 2 volte?