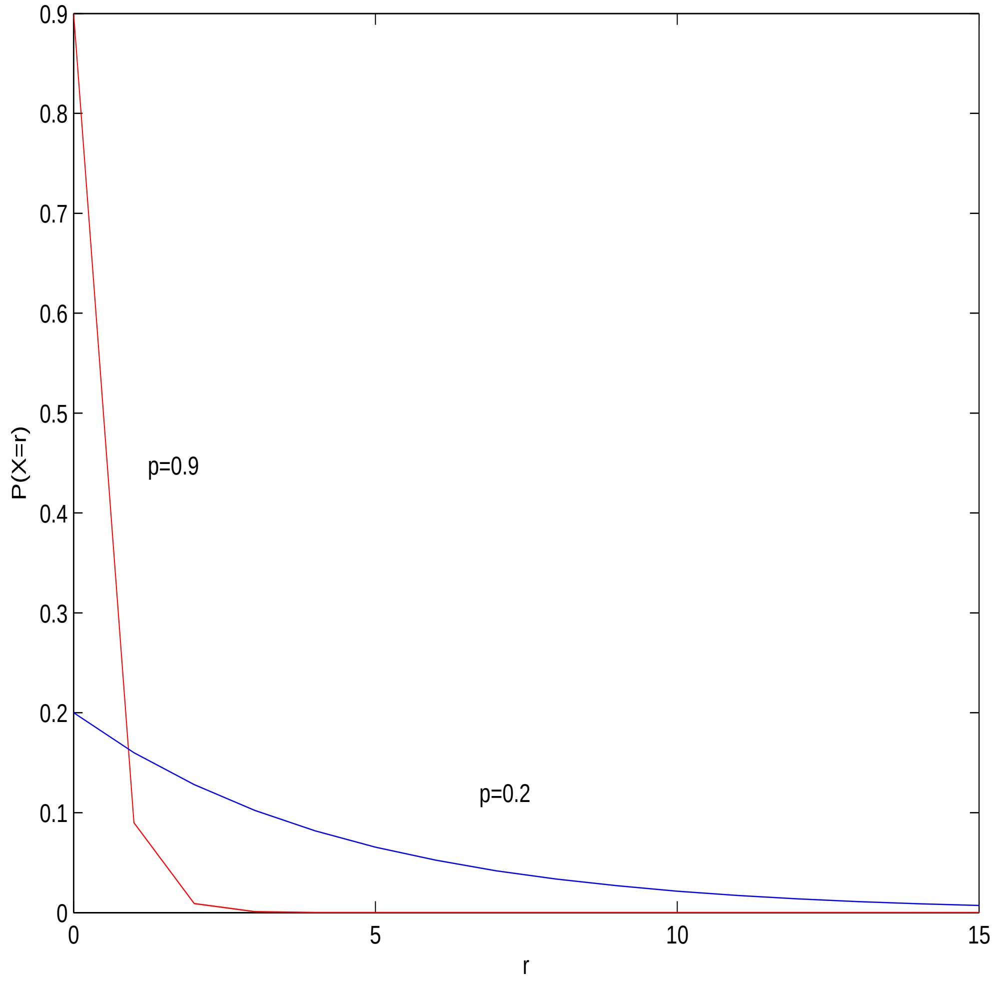
Distribuzione geometrica

Una variabile casuale discreta X segue la distribuzione geometrica se

$$P(X = r) = (1-p)^{r-1}p, \qquad r = 1, 2, \ldots$$

Condizioni:

1. C'è una successione di prove;
2. Due possibili risultati (successo/insuccesso);
3. Le prove sono indipendenti;
4. La probabilità ad ogni prova rimane costante;
5. La variabile è il numero di prove necessarie per avere il primo successo


0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
P(X=r)
p=0.9
p=0.2
0
5
10
15
r

Nota: Le probabilità associate ad r sono i termini di una serie geometrica:

$$p(1 + (1-p) + (1-p)^2 + \ldots) = p \sum_{r=0}^{\infty} (1-p)^r.$$

La serie è convergente con somma $p\frac{1}{1-(1-p)} = 1$ ($\Rightarrow$ la distribuzione geometrica è ben definita)

Esempio: In una produzione di chiodi con macchina automatica, in media un 5% della produzione viene scartata perchè inferiore al minimo permesso di 3 cm. Uno schema di controllo consiste nel prendere chiodi a caso dalla produzione e nel contare quanti ne vengono presi prima di prenderne uno imperfetto. Rappresentare i primi termini di questa distribuzione.

Sol. X: numero di chiodi presi fino al (ed incluso il) primo imperfetto.

$$P(X = r) = (0.95)^{r-1}(0.05)$$

r	1	2	3	4	5
P(X=r)	0.05	0.0475	0.0451	0.0429	0.0407

Esempio: In un'analisi di laboratorio, un esperimento ha il 30% di probabilità di dare una risposta positiva. Quante prove occorre fare per avere una probabilità del 90% di avere una risposta positiva?

Sol. La variabile X è il numero di prove. Si ha $p = 0.3$.

$P(X = 1) = 0.3$

$P(X = 2) = 0.3 \cdot 0.7 = 0.21$

$\vdots$

$P(X = 6) = 0.3(1 - 0.3)^5 = 0.88$

$P(X = 7) = 0.3(1 - 0.3)^6 = 0.91$

Alcune proprietà

Sia $p_r = p(1-p)^{r-1}$.

Media:

$$\begin{aligned} E[X] &= p_1 + 2p_2 + 3p_3 + \cdots = \sum_{r=1}^{\infty} r p_r \\ &= \sum_{r=1}^{\infty} r p (1-p)^{r-1} = p \sum_{r=1}^{\infty} r (1-p)^{r-1} \\ &= p \frac{1}{(1-(1-p))^2} = p \frac{1}{p^2} = \frac{1}{p}. \end{aligned}$$

Varianza:

$$Var[X] = (1-p)/p^2$$

Distribuzione di Poisson

Si dice che una variabile aleatoria discreta X segue la distribuzione di Poisson se, fissato $\lambda > 0$, vale

$$P(X = r) = \frac{\lambda^r}{r!} e^{-\lambda} \qquad r = 0, 1, \ldots, \lambda.$$

Condizioni:

1. Gli eventi sono *casuali* nello spazio (tempo) continuo $\star$;
2. Gli eventi hanno luogo singolarmente e sono esclusivi;
3. Il numero di eventi che ha luogo in un dato intervallo è proporzionale alla lunghezza dell'intervallo $\star$;
4. Gli eventi sono indipendenti $\star$;
5. La variabile è *il numero di eventi* aventi luogo nell'intervallo considerato.

Eventi rari

La distribuzione di Poisson si può derivare come limite della distribuzione binomiale per

$N \to \infty$ e $p \to 0$, ponendo $\lambda = N \cdot p$.

Esercizio: Analizzare graficamente la distribuzione binomiale e Poissoniana al variare del parametro λ.

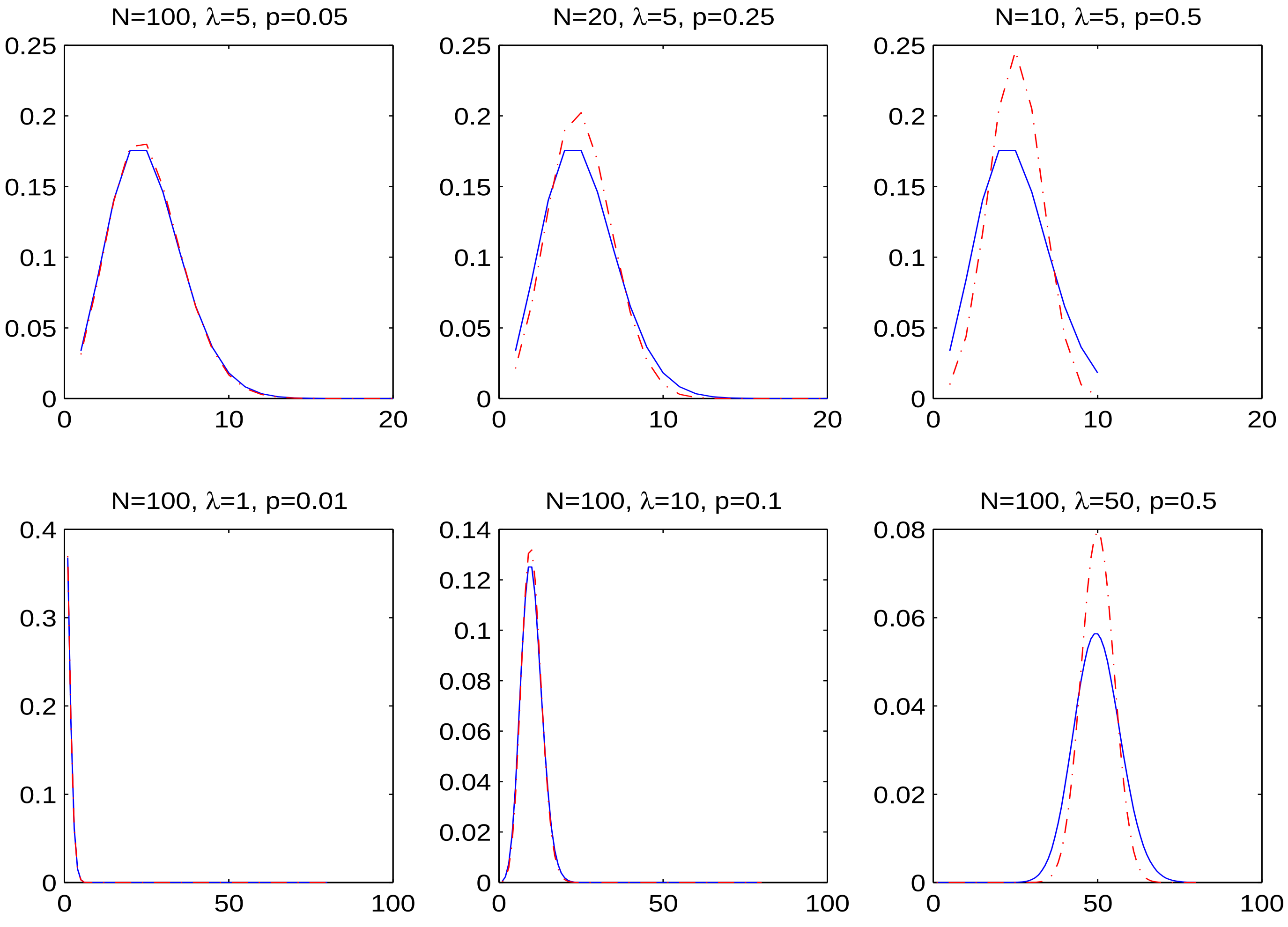


Figura 1: Poligoni per le distribuzioni binomiale e di Poisson per vari valori dei parametri di riferimento, con $\lambda = N \cdot p$. Poisson. Binomiale

Esempio: Il numero medio di errori di battitura per pagina in un libro è 1.2. Qual'é la probabilità di trovare in una particolare pagina (di 2000 lettere): (a) nessun errore; (b) tre o più errori?

Sol. La variabile aleatoria X è il numero di errori.

La probabilità di trovare un errore in una pagina di $N = 2000$ lettere è $p = 1.2/2000 = 0.0006$. Con la distribuzione di Poisson per $\lambda = Np = 1.2$:

(a) $P(X = 0) = 1e^{-1.2} \approx 0.30$

(b) $P(X \geq 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2) =$
$1 - 0.30 - 0.36 - 0.22 = 0.12$

Esempio: Una malattia rara viene presa in media dallo 0.5% dei neonati. Cento bambini nascono in un ospedale in una settimana. Qual'è la probabilità che esattamente tre di loro abbiano la malattia?

Sol. La probabilità è bassa. Ponendo $\lambda = 100 \cdot 0.005 = 0.5$, si ottiene usando la distribuzione di Poisson,

$$P(X = 3) = \frac{\lambda^3}{3!} e^{-5} = 0.1263$$

Alcune proprietà

Media:

$$\begin{aligned} E[X] &= \sum_{r=0}^{\infty} r P(X=r) = \sum_{r=1}^{\infty} \frac{r}{r!} \lambda^r e^{-\lambda} \\ &= \lambda e^{-\lambda} \sum_{r=1}^{\infty} \frac{1}{(r-1)!} \lambda^{r-1} = \lambda e^{-\lambda} e^{\lambda} = \lambda \end{aligned}$$

Varianza:

$$Var[X] = \lambda$$

Distribuzioni continue di variabili casuali

X può assumere tutti i valori in un intervallo $[a, b]$

Definizione. Densità di probabilità di X in $[a, b]$:

la funzione f integrabile a cui è associata la probabilità

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Nota: $P(X = a) = 0$

Probabilità di ottenere un valore $c \in [a, b]$:

$$P(X = c) = \int_{c-\delta}^{c+\delta} f(x)dx \approx f(c) \cdot (2\delta)$$

Condizioni per f:

- f continua
- $f \geq 0, \; \forall x$
- $\int_{-\infty}^{\infty} f(x)dx = 1$

La terza condizione: Per X definita in $[a, b]$:

$$\int_a^b f(x)dx = 1$$

Funzione di distribuzione cumulativa:

$$F(b) = P(X \leq b) = \int_{-\infty}^{b} f(x)dx$$

Mediana:

$$M \text{ tale che } \int_{-\infty}^{M} f(x)dx = \frac{1}{2}.$$

Moda:

$$M \text{ tale che } \max_{x \in [a,b]} f(x) = f(M).$$

Valore atteso (**media**) nel continuo:

$$\mu = E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

(momento di ordine 1)

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

Varianza nel continuo:

$$Var[X] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx = E[X^2] - \mu^2$$

(momento di ordine 2 dalla media)