

3. Distribuzioni

Corso di Simulazione

Anno accademico 2008/09

Ω *spazio campione*

$\mathcal{F} \subseteq 2^\Omega$ *spazio degli eventi:*

(i) $\Omega \in \mathcal{F}$

(ii) $A \in \mathcal{F} \implies \Omega \setminus A \in \mathcal{F}$

(iii) $A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}$

$P: \mathcal{F} \rightarrow [0, 1]$ *funzione di probabilità:*

(i) $P(A) \geq 0$

(ii) $P(\Omega) = 1$

(iii) $A \cap B = \emptyset \implies P(A \cup B) = P(A) + P(B)$

Eventi

$$A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F} \text{ e } B \setminus A \in \mathcal{F}$$

Probabilità

$$P(\Omega \setminus A) = 1 - P(A)$$

$$A \subseteq B \implies P(B \setminus A) = P(B) - P(A)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$X : \Omega \rightarrow \mathbb{R}$ **variabile casuale** se

$$\{\omega \in \Omega : X(\omega) \leq r\} \in \mathcal{F}$$

Funzione di distribuzione di X

$$F_X(r) := P(X \leq r) = P(\{\omega \in \Omega : X(\omega) \leq r\})$$

Arrivo di 100 pazienti urgenti, gravi o non gravi al pronto soccorso

X_u = numero pazienti urgenti (analogamente X_g , X_{ng})

$\Omega = \{u, g, ng\}^{100}$, $\mathcal{F} = 2^\Omega$, $P = ?$

$X_u(w) = \text{card}(\{i \in \{1, \dots, 100\} : w_i = u\})$

nessun paziente è urgente: $\{X_u = 0\}$

almeno tre pazienti sono gravi: $\{X_g \geq 3\}$

Tempo di interarrivo dei 100 pazienti

$\Omega = \mathbb{N}^{100}$ (oppure $\Omega = \mathbb{R}_+^{100}$), $\mathcal{F} = 2^\Omega$, $P = ?$

w_i = tempo tra l'arrivo del paziente $i - 1$ e l'arrivo del paziente i

X_k = tempo necessario per l'arrivo dei primi k pazienti

$X_k(\omega) = w_1 + \dots + w_k$

k pazienti arrivati in 60 minuti: $\{X_k \leq 60\}$

dopo 10 minuti non è arrivato nessun paziente: $\{X_1 > 10\}$

Ad uno sportello di banca si possono fare solamente tre operazioni:

a incasso di un assegno

b bonifico

c versamento

Assumiamo che un singolo cliente faccia una sola di esse.

Consideriamo come esperimento l'arrivo del prossimo cliente, come esito dell'esperimento la richiesta di una delle operazioni, *a*, *b* e *v*, e come evento il fatto che il cliente chieda una in un sottoinsieme delle operazioni (ad esempio la *a* o la *v*).

Poniamo allora $\Omega = \{a, b, v\}$, $\mathcal{F} = 2^\Omega$. Sia poi X la funzione così definita:

$$X(a) = 0,$$

$$X(b) = 1,$$

$$X(v) = 2.$$

La funzione X è una variabile casuale, infatti si ha:

$$\begin{aligned}r < 0 & \quad \{\omega : X(\omega) \leq r\} = \emptyset, \\0 \leq r < 1 & \quad \{\omega : X(\omega) \leq r\} = \{a\}, \\1 \leq r < 2 & \quad \{\omega : X(\omega) \leq r\} = \{a, b\}, \\2 \leq r & \quad \{\omega : X(\omega) \leq r\} = \Omega.\end{aligned}$$

X è una **variabile casuale discreta** se l'insieme dei valori che può assumere è numerabile: $X(\Omega) = \{x_1, x_2, \dots\}$

$$f_X(x) = \begin{cases} P(X = x), & \text{se } x = x_i, \text{ per qualche } i = 1, 2, \dots, \\ 0, & \text{altrimenti.} \end{cases}$$

$f_X(x)$ è una **Funzione di densità discreta**

$$F_X(x) = \sum_{x_i \leq x} f_X(x_i)$$

è la corrispondente **funzione di distribuzione**

$$f_X(x_i) = F_X(x_i) - \lim_{h \rightarrow 0^+} F_X(x_i - h)$$

Media di X

$$\mu_X = E[X] := \sum_i x_i f_X(x_i)$$

Varianza di X

$$\begin{aligned}\sigma_X^2 = \text{Var}[X] &:= E[(X - \mu_X)^2] \\ &= \sum_i (x_i - \mu_X)^2 f_X(x_i)\end{aligned}$$

Deviazione standard di X

$$\sigma_X := \sqrt{\text{Var}[X]}$$

X, Y variabili aleatorie, $\alpha \in \mathbb{R}$

$$E[\alpha] = \alpha$$

$$E[\alpha X] = \alpha E[X]$$

$$E[X + Y] = E[X] + E[Y]$$

Proprietà della varianza (1)

$$\begin{aligned} \text{Var}[X] &= E[(X - \mu_X)^2] \\ &= E[X^2] + \mu_X^2 - 2\mu_X E[X] \\ &= E[X^2] + \mu_X^2 - 2\mu_X^2 \\ &= E[X^2] - (E[X])^2 \end{aligned}$$

Proprietà della varianza (2)

$$\text{Var}[\alpha] = 0$$

$$\text{Var}[\alpha X] = \alpha^2 \text{Var}[X]$$

$$\begin{aligned}\text{Var}[X + Y] &= E[(X + Y - E[X + Y])^2] \\ &= E[(X + Y - \mu_X - \mu_Y)^2] \\ &= E[(X - \mu_X)^2 + (Y - \mu_Y)^2 + 2(X - \mu_X)(Y - \mu_Y)] \\ &= \sigma_X^2 + \sigma_Y^2 + 2\text{Cov}[X, Y],\end{aligned}$$

$$\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])] = \text{Covarianza di } X \text{ e } Y.$$

Proprietà della varianza (3)

Se X ed Y sono indipendenti, allora la loro covarianza è nulla e si ha:

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y].$$

$$P(|X - \mu_X| > r\sigma_X) \leq 1/r^2$$

Disuguaglianza di Chebyshev



$$P(|X - \mu_X| < r\sigma_X) \geq 1 - 1/r^2$$

$$\begin{aligned}m_X(t) &= E[e^{tX}] = E[1 + Xt + \frac{1}{2!}(Xt)^2 + \frac{1}{3!}(Xt)^3 + \dots] \\ &= 1 + \mu_{1X}t + \frac{1}{2!}\mu_{2X}t^2 + \dots = \sum_{i=0}^{\infty} \frac{1}{i!}\mu_{iX}t^i.\end{aligned}$$

μ_{rX} , il *momento* r^{esimo} di X , è la media della variabile casuale X^r .

È immediato verificare che risulta:

$$\frac{d^r}{dt^r} m_X(0) = \mu_{rX}.$$

$$X : \Omega \rightarrow \{1, \dots, n\}$$

Tutti i possibili valori sono equiprobabili

$$f_X(x) := \begin{cases} \frac{1}{n} & x \in \{1, \dots, n\} \\ 0 & x \notin \{1, \dots, n\} \end{cases} = \frac{1}{n} I_{\{1,2,\dots,n\}}(x)$$

$$E[X] = \frac{1}{n} \sum_{i=1}^n i = \frac{(n+1)n}{2n} = \frac{n+1}{2}$$

$$\begin{aligned} \text{Var}[X] &= E[X^2] - (E[X])^2 = \frac{1}{n} \sum_{i=1}^n i^2 - \frac{(n+1)^2}{4} \\ &= \frac{n(n+1)(2n+1)}{6n} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12} \end{aligned}$$

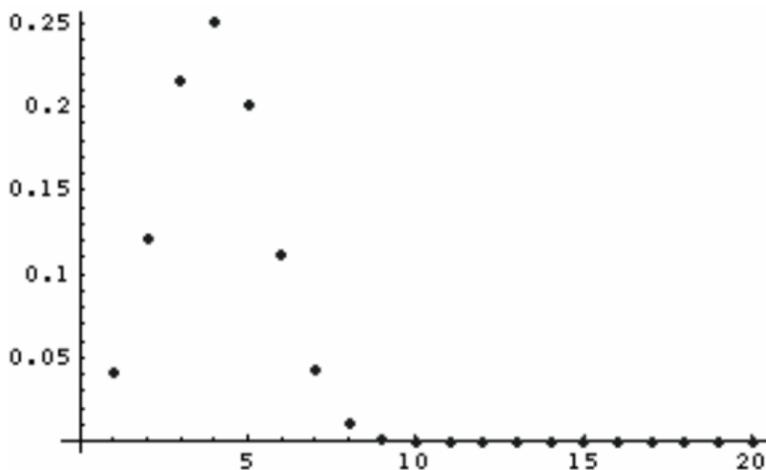
Distribuzione binomiale

$$X : \Omega \rightarrow \{0, 1, \dots, n\}$$

Numero di successi in un esperimento ripetuto n volte

$p \in [0, 1]$ probabilità di successo in un singolo esperimento

$$f_X(k) := \binom{n}{k} p^k (1-p)^{n-k} I_{\{0,1,2,\dots,n\}}(k)$$



Distribuzione binomiale (2)

$$m_X(t) = \sum_{i=0}^n e^{ti} \binom{n}{i} p^i q^{n-i} = \sum_{i=0}^n \binom{n}{i} (pe^t)^i q^{n-i} = (q + pe^t)^n$$

$$\frac{dm_X(t)}{dt} = pe^t n (pe^t + q)^{n-1}$$

$$\begin{aligned} \frac{d^2 m_X(t)}{dt^2} &= pe^t n (pe^t + q)^{n-1} + pe^t n (n-1) pe^t (pe^t + q)^{n-2} \\ &= pe^t n (pe^t + q)^{n-2} (npe^t + q) \end{aligned}$$

Distribuzione binomiale (3)

$$E[X] = \frac{dm_X(0)}{dt} = pn,$$

$$E[X^2] = \frac{d^2m_X(0)}{dt^2} = np(np + q)$$

$$\begin{aligned} \text{Var}[X] &= E[X^2] - (E[X])^2 \\ &= np(np + q) - (np)^2 = npq \end{aligned}$$

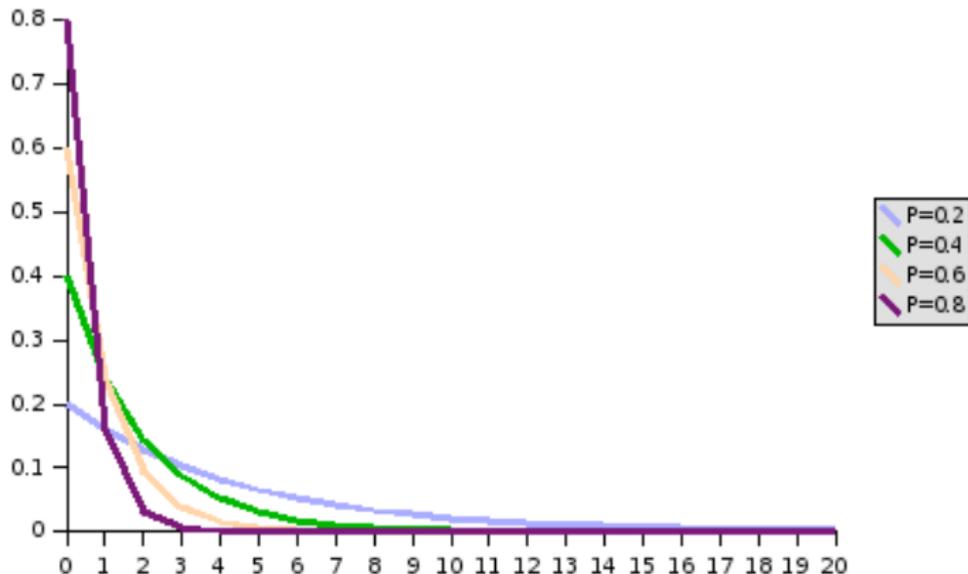
- Numero di viaggiatori prenotati che non si presentano alla partenza di un volo.
- Numero di pezzi difettosi in un lotto di n pezzi prodotti.
- Numero di clienti che richiedono un particolare servizio.

Distribuzione geometrica

$$X : \Omega \rightarrow \mathbb{N}$$

Numero di fallimenti prima di un successo

$$f_X(k) = P(X = k) := (1 - p)^k p$$



Distribuzione geometrica (2)

$$m_X(t) = \sum_{i=0}^{\infty} e^{it} p(1-p)^i = p \sum_{i=0}^{\infty} [e^t(1-p)]^i.$$

e se assumiamo che essa sia definita in un intorno sufficientemente piccolo dello 0 per cui risulti $e^t(1-p) < 1$, allora è

$$m_X(t) = \frac{p}{1 - e^t(1-p)}$$

$$m'_X(t) = \frac{pe^t(1-p)}{(1 - e^t(1-p))^2}$$

$$m''_X(t) = \frac{pe^t(1-p)(1 - e^t(1-p))^2 + 2(1 - e^t(1-p))e^t(1-p)pe^t(1-p)}{(1 - e^t(1-p))^4}$$

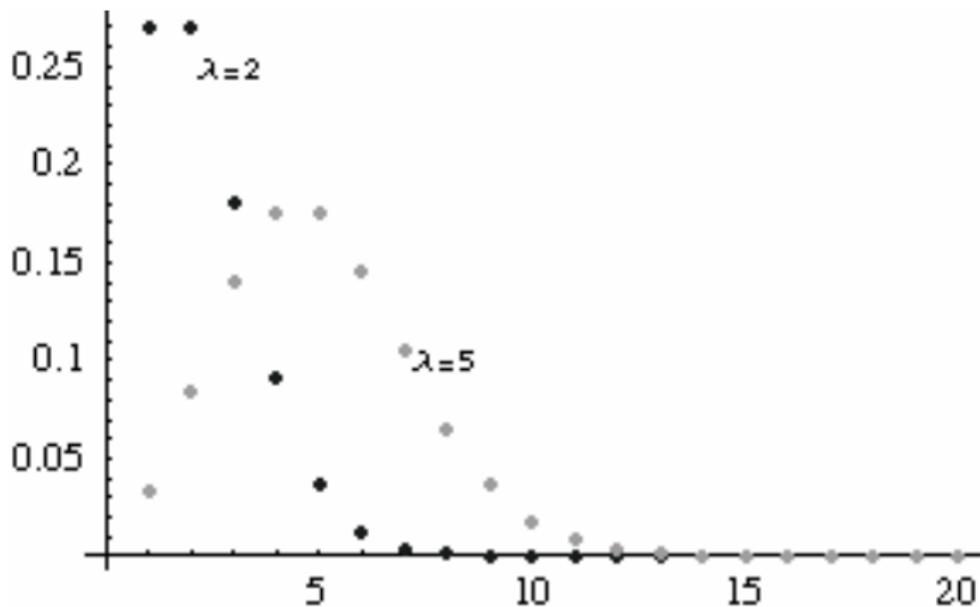
$$E[X] = m'_X(0) = \frac{1-p}{p};$$

$$\begin{aligned} \text{Var}[X] &= E[X^2] - E[X]^2 \\ &= m''_X(0) - E[X]^2 \\ &= \frac{(1-p)(2-p)}{p^2} - \frac{(1-p)^2}{p^2} = \frac{1-p}{p^2}. \end{aligned}$$

Distribuzione di Poisson (1)

Numero di occorrenze di un evento in $[0, t]$

$$f_X(k) := \frac{\lambda^k e^{-\lambda}}{k!}$$



$\lambda := \nu t$ (ν numero medio di occorrenze nell'unità di tempo)

Distribuzione di Poisson (2)

$$\frac{\lambda^k e^{-\lambda}}{k!}$$

$$m_X(t) = e^{-\lambda} \sum_{i=0}^{\infty} \frac{(e^t \lambda)^i}{i!} = e^{\lambda(e^t - 1)}$$

$$\frac{dm_X(t)}{dt} = \lambda e^{\lambda(e^t - 1) + t}, \quad \frac{d^2 m_X(t)}{dt^2} = \lambda e^{\lambda(e^t - 1) + t} (\lambda e^t + 1)$$

$$E[X] = \frac{dm_X(0)}{dt} = \lambda,$$

$$E[X^2] = \frac{d^2 m_X(0)}{dt^2} = \lambda(\lambda + 1)$$

$$\text{Var}[X] = E[X^2] - E[X]^2 = \lambda(\lambda + 1) - \lambda^2 = \lambda$$

Evento ripetuto nel tempo: ipotesi

- numeri di occorrenze in intervalli disgiunti sono indipendenti
- Se $h \ll 1$, allora

$$P(X_{[\tau, \tau+h]} = 1) = \nu h + o(h)$$

$$P(X_{[\tau, \tau+h]} > 1) = o(h)$$

dove $o(h)$ indica un infinitesimo di ordine superiore al primo:

$$\lim_{h \rightarrow 0} o(h)/h = 0$$

Evento ripetuto nel tempo: discretizzazione

Dividere $[0, t]$ in n intervalli di ampiezza t/n

Al più una singola occorrenza in ciascun intervallo $[t_k, t_{k+1}]$
successo = occorrenza nell'intervallo $[t_k, t_{k+1}]$ considerato

$X_{[0,t]}^{(n)}$ ha distribuzione binomiale di parametro $p = \nu t/n$

$$\begin{aligned} P[X = k] &= \binom{n}{k} \left(\frac{\nu t}{n}\right)^k \left(1 - \frac{\nu t}{n}\right)^{n-k} \\ &= \frac{n(n-1)\dots(n-k+1)}{k!n^k} (\nu t)^k \left(1 - \frac{\nu t}{n}\right)^n \left(1 - \frac{\nu t}{n}\right)^{-k} \\ &\xrightarrow{n \rightarrow \infty} \frac{(\nu t)^k e^{-\nu t}}{k!}. \end{aligned}$$

$$\lim_{n \rightarrow +\infty} \left(1 + \frac{1}{n}\right)^n = e$$

$$\begin{aligned} \lim_{n \rightarrow +\infty} \left(1 - \frac{\nu t}{n}\right)^n &= \lim_{m \rightarrow +\infty} \left(1 - \frac{1}{m+1}\right)^{(m+1)\nu t} \\ &= \lim_{m \rightarrow +\infty} \left(\frac{m}{m+1}\right)^{m\nu t + \nu t} \\ &= \lim_{m \rightarrow +\infty} \left(1 + \frac{1}{m}\right)^{-m\nu t} \left(\frac{m}{m+1}\right)^{\nu t} \\ &= \lim_{m \rightarrow +\infty} \left(\left(1 + \frac{1}{m}\right)^m\right)^{-\nu t} \\ &= e^{-\nu t} \end{aligned}$$

$$m = \frac{n}{\nu t} - 1$$

$$P[X = k] = \frac{(\nu t)^k e^{-\nu t}}{k!}$$

$E[X] = \nu t \Rightarrow \nu$ è il numero medio di arrivi per unità di tempo

X è una variabile casuale continua se $X(\Omega) \in \mathbf{R}$, ed esiste una funzione reale f_X per cui :

$$F_X(x) = \int_{-\infty}^x f_X(u) du, \quad \forall x \in \mathbf{R}$$

(f_X è la *funzione di densità*)

$$F_X(x) \text{ differenziabile} \Rightarrow f_X(x) = \frac{dF_X(x)}{dx}.$$

$$E[X] = \mu_X = \int_{-\infty}^{\infty} xf_X(x)dx.$$

$$\text{Var}[X] = E[(X - \mu_X)^2] = E[X^2] - (E[X])^2$$

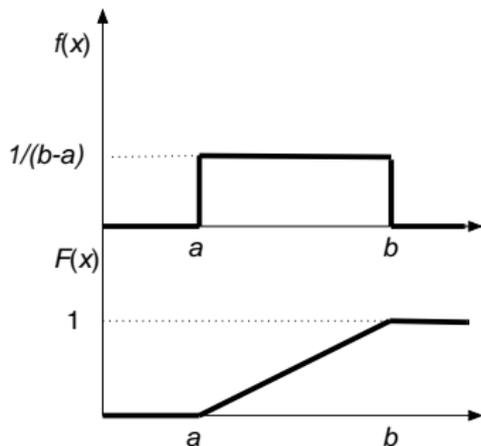
$$m_X(t) = E[e^{tX}]$$

Distribuzione uniforme

X è una variabile casuale uniformemente distribuita nell'intervallo reale $[a, b]$ se

$$f_X(x) = f_X(x; a, b) = \frac{1}{b-a} I_{[a,b]}(x)$$

$$F_X(x) = \left(\frac{x-a}{b-a} \right) I_{[a,b]}(x) + I_{(b,\infty)}(x)$$



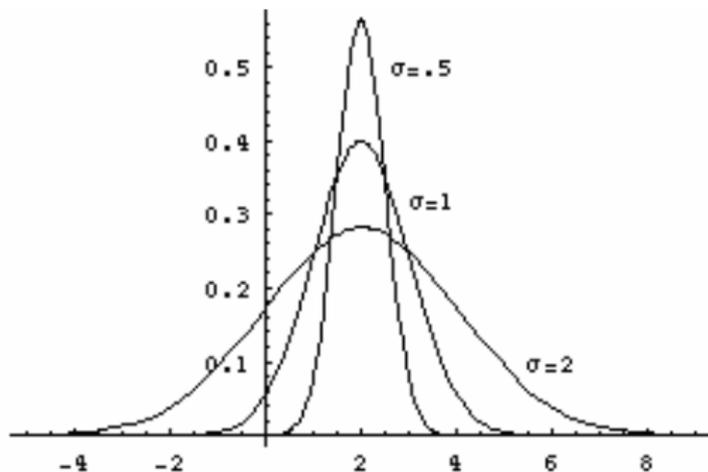
$$E[X] = \int_a^b \frac{x}{b-a} dx = \frac{b+a}{2}$$

$$\begin{aligned} \text{Var}[X] &= E[X^2] - (E[X])^2 = \int_a^b \frac{x^2}{b-a} dx - \left(\frac{a+b}{2}\right)^2 \\ &= \frac{b^3 - a^3}{3(b-a)} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12} \end{aligned}$$

$$m_X(t) = \int_a^b e^{tx} \frac{1}{b-a} dx = \frac{e^{bt} - e^{at}}{(b-a)t}$$

$$f_X(x) = f_X(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

μ e σ sono rispettivamente la media e la deviazione standard



$$\begin{aligned}m_X(t) &= E[e^{tX}] = e^{t\mu} E[e^{t(X-\mu)}] \\&= e^{\mu t} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{t(x-\mu)} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\&= e^{\mu t} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2 - 2\sigma^2 t(x-\mu)}{2\sigma^2}} dx \\&= e^{\mu t + \frac{\sigma^2 t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu-\sigma^2 t)^2}{2\sigma^2}} dx = e^{\mu t + \frac{\sigma^2 t^2}{2}}\end{aligned}$$

$$E[X] = \frac{d}{dt} m_X(0) = \mu$$

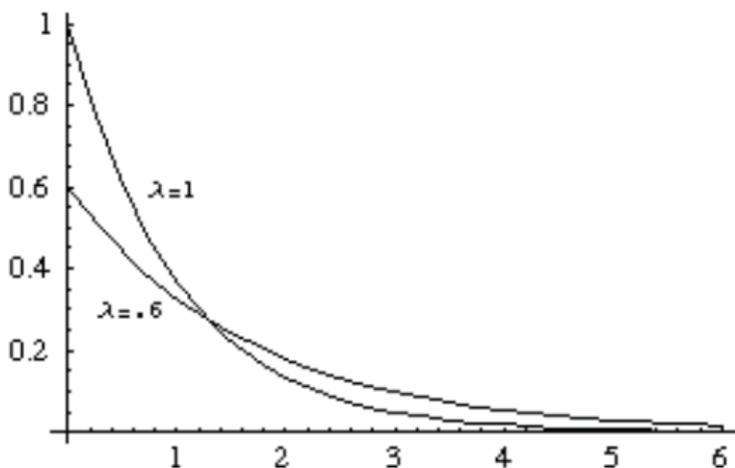
$$\text{Var}[X] = E[X^2] - (E[X])^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2$$

Distribuzione esponenziale

Variabile casuale definita nello spazio dei reali non negativi, con $\lambda > 0$

$$f_X(x; \lambda) = \lambda e^{-\lambda x},$$

$$F_X(x) = 1 - e^{-\lambda x},$$



$$\begin{aligned}m_X(t) &= E[e^{tX}] = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx \\ &= \lambda \int_0^{\infty} e^{-(\lambda-t)x} dx = \frac{\lambda}{\lambda - t}.\end{aligned}$$

$$E[X] = \frac{1}{\lambda}, \text{Var}[X] = \frac{1}{\lambda^2}.$$

Consideriamo un evento le cui occorrenze nel tempo hanno una distribuzione di Poisson. Supponendo che si sia appena verificata un'occorrenza, chiamiamo con X la variabile casuale "tempo da attendere prima della occorrenza successiva". Ricordando che, se Y è una v.c. Poisson, è $P[Y = k] = \frac{(\nu t)^k e^{-\nu t}}{k!}$, si ha allora

$$P[X > t] = P[\text{nessuna occorrenza fino al tempo } t] = e^{-\nu t}$$

e di conseguenza

$$F_X(t) = P[X \leq t] = 1 - e^{-\nu t}, t \geq 0$$

$$f_X(t) = \frac{dF_X(t)}{dt} = \nu e^{-\nu t}$$

Per ogni coppia (s, t) di reali positivi, vale la

$$P[X > s + t | X > s] = P[X > t].$$

Infatti è

$$P[X > s+t | X > s] = \frac{P[X > s+t]}{P[X > s]} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P[X > t].$$

Si parla in questo caso di proprietà di *assenza di memoria*.

Inoltre:

$$P[cX \leq x] = P[X \leq \frac{x}{c}] = 1 - e^{-\frac{\lambda}{c}x}$$

X esponenziale con parametro $\lambda \Rightarrow cX$ esponenziale con parametro $\frac{\lambda}{c}$

Distribuzione Gamma con parametri n e λ

X_1, X_2, \dots, X_n variabili casuali indipendenti con distribuzione esponenziale e parametro λ .

$$Y_n = \sum_{i=1}^n X_i.$$

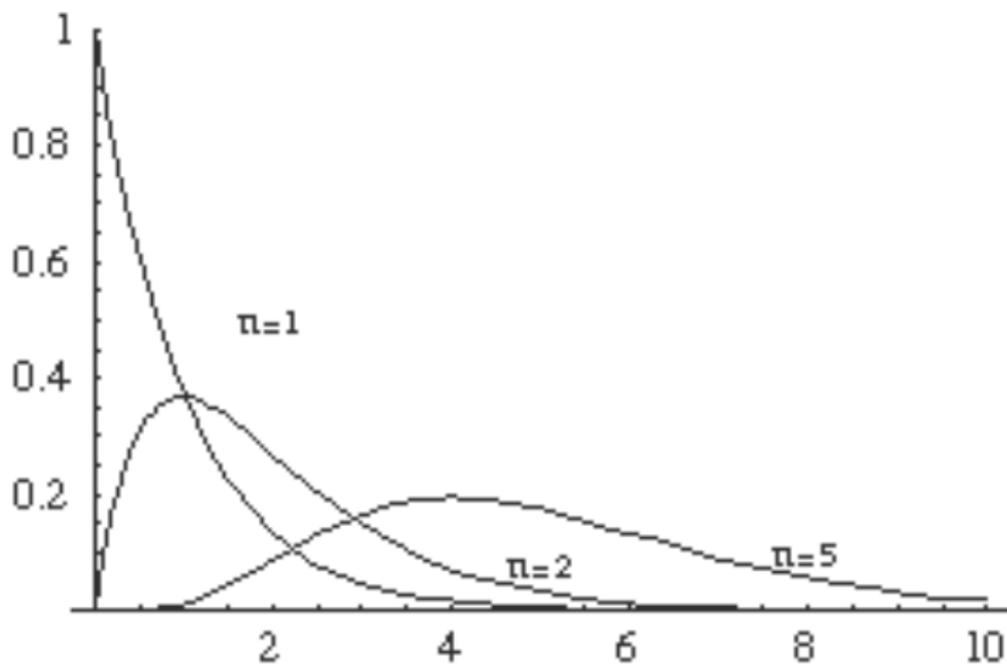
La probabilità che Y_n sia minore o uguale a t è pari alla probabilità che nel tempo t si verifichino almeno n eventi.

$$F_{Y_n}(t) = P[Y_n \leq t] = \sum_{j=n}^{\infty} \frac{(\lambda t)^j e^{-\lambda t}}{j!}.$$

$$\begin{aligned} f_{Y_n}(t) &= F'_{Y_n}(t) = \sum_{j=n}^{\infty} \frac{\lambda j (\lambda t)^{j-1} e^{-\lambda t} - \lambda (\lambda t)^j e^{-\lambda t}}{j!} \\ &= \lambda e^{-\lambda t} \left(\sum_{j=n}^{\infty} \frac{(\lambda t)^{j-1}}{(j-1)!} - \sum_{j=n}^{\infty} \frac{(\lambda t)^j}{j!} \right) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}. \end{aligned}$$

Gamma: media e varianza

$$E[Y_n] = \frac{n}{\lambda}, \quad \text{Var}[Y_n] = \frac{n}{\lambda^2}.$$



$\lambda = 1$