5 Valore atteso di una funzione di variabile casuale

Il concetto di valore atteso viene esteso in maniera naturale anche ad una funzione di variabile casuale. Se X è una variabile casuale e $g : \mathbb{R} \to \mathbb{R}$ una funzione reale di variabile reale, allora $Y = g(X)$ è una variabile casuale.

DEFINIZIONE: Chiamiamo VALORE ATTESO della variabile casuale $g(X)$ la quantità:

- $E[g(X)] = \sum_i g(x_i) f(x_i)$ se X è variabile casuale discreta.
- $E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x) dx$ se X è variabile casuale continua.

OSSERVAZIONI

- $g(X) = X$
  $\Rightarrow E[g(X)] = E[X] = \mu_X$
- $g(X) = (X - \mu_X)^2$
  $\Rightarrow E[g(X)] = E[(X - \mu_X)^2] = \text{var}[X] = \sigma_X^2$

Esempi

1) Sia X la variabile casuale che rappresenta il numero di automobili lavate in un autolavaggio tra le 16 e le 17 di un dato giorno. Supponiamo che X abbia la seguente distribuzione di probabilità:

x	4	5	6	7	8	9
$f(x) = P[X=x]$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Figura 18

Sia $g(X) = 2X - 1$ la variabile casuale che rappresenta la quantità, in euro, pagata all'addetto dal gestore. Calcolare il guadagno atteso dall'addetto, in euro, tra le 16 e le 17.

$$E[g(X)] = E[2X-1] = \sum_{i=4}^{9}(2x_i - 1)f(x_i) =$$

$$= 7 \cdot \frac{1}{12} + 9 \cdot \frac{1}{12} + ... + 15 \cdot \frac{1}{6} + 17 \cdot \frac{1}{6} = 12.67$$

2) Sia X la variabile casuale (continua) che rappresenta il tempo impiegato per trovare un guasto in un impianto elettrico (misurato in ore).

Supponiamo che X abbia la seguente funzione di densità:

$$f(x) = \begin{cases} 1 & \text{se } 0 < x < 1\,, \\ 0 & \text{altrove}\,. \end{cases}$$

Causa il guasto, l'interruzione di x ore dell'impianto provoca un danno economico $g(X) = X^3$.
Calcolare il valore atteso del danno.

$$E[g(X)] = \int_{-\infty}^{+\infty} x^3 f(x)dx = \int_0^1 x^3 dx = \frac{1}{4} = 0.25$$

PROPRIETÀ DEL VALORE ATTESO

Elenchiamo ora le proprietà del valore atteso di una funzione di variabile casuale, che possono quindi essere espresse sia per variabili casuali discrete che continue. Nel caso continuo la somma è sostituita da un integrale.

1. $E[c] = c$ con c costante $\in \mathbb{R}$. Infatti

$$E[c] = \sum_i cf(x_i) = c \underbrace{\sum_i f(x_i)}_{=1} = c$$

2. $E[cg(X)] = cE[g(X)]$ con c costante $\in \mathbb{R}$. Infatti

$$\begin{aligned} E[cg(X)] &= \sum_i cg(x_i)f(x_i) = \\ &= c\sum_i g(x_i)f(x_i) = cE[g(X)] \end{aligned}$$

3. $E[c_1g_1(X)+c_2g_2(X)] = c_1E[g_1(X)]+c_2E[g_2(X)]$ con c_1, c_2 costanti $\in \mathbb{R}$. Infatti

$$\begin{aligned} &E[c_1g_1(X) + c_2g_2(X)] = \\ &\sum_i (c_1g_1(x_i) + c_2g_2(x_i))f(x_i) = \\ &= c_1\sum_i g_1(x_i)f(x_i) + c_2\sum_i g_2(x_i)f(x_i) \\ &= c_1E[g_1(X)] + c_2E[g_2(X)] \end{aligned}$$

4. se $g_1(X) \leq g_2(X)$, $\forall X$ $E[g_1(X)] \leq E[g_2(X)]$. Infatti

$$\begin{aligned} &E[g_2(X)] - E[g_1(X)] = E[g_2(X) - g_1(X)] = \\ &= \sum_i \underbrace{(g_2(x_i) - g_1(x_i))}_{\geq 0}\underbrace{f(x_i)}_{\geq 0} \geq 0 \end{aligned}$$

5. DISUGUAGLIANZA DI MARKOV
   Data una funzione $g : \mathbb{R} \to \mathbb{R}$ tale che $g(X) \geq 0$ $\forall x \in \mathbb{R}$ per cui esista $E[g(X)]$, allora:

$$P[g(X) \geq r] \leq \frac{1}{r} E[g(X)], \quad \forall r > 0$$

6. DISUGUAGLIANZA DI CHEBYSHEV
   Data una variabile casuale X con media μ_X e varianza σ_X^2 si ha:

$$P[|X - \mu_X| \geq k\sigma_X] \leq \frac{1}{k^2}, \ \forall k > 0$$

   o equivalentemente, $\forall k > 0$

$$P[|X - \mu_X| < k\sigma_X] =$$

$$= P[\mu_X - k\sigma_X < X < \mu_X + k\sigma_X] \geq 1 - \frac{1}{k^2}$$

La disuguaglianza di Chebyshev si ottiene dalla disuguaglianza di Markov tramite un'opportuna scelta di $g(X)$ e di r. Più precisamente siano $g(X) = (X - \mu_X)^2$ ed $r = k^2\sigma_X^2$ (> 0). Dalla disuguaglianza di Markov si ottiene:

$$P[(X - \mu_X)^2 \geq k^2\sigma_X^2] \leq \frac{1}{k^2\,\sigma_X^2} \underbrace{E[(X - \mu_X)^2]}_{=\mathrm{var}[X]=\sigma_X^2} = \frac{1}{k^2}$$

e

$$P[(X-\mu_X)^2 \geq k^2\sigma_X^2] = P[|X-\mu_X| \geq k\sigma_X]$$

OSSERVAZIONE

Se $g(X) = X$ con $X > 0$ per Markov

$$P[X \geq a] \leq \frac{1}{a}E[X], \ \forall a > 0.$$

COMMENTI

1) Verifichiamo la disuguaglianza di Markov nel caso continuo, prima con la variabile casuale X $(X > 0)$ e poi con una variabile casuale arbitraria $g(X)$ (con $g(X) > 0$):

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx \underset{\substack{\uparrow \\ x>0}}{=} \int_0^{\infty} xf(x)dx =$$

$$= \underbrace{\int_0^a xf(x)dx}_{>0, f(x)>0} + \underbrace{\int_a^{\infty} xf(x)dx}_{>0, f(x)>0} \geq$$

$$\geq \int_a^{\infty} xf(x)dx \underset{\substack{\uparrow \\ x>a}}{\geq} \int_a^{\infty} af(x)dx = aP[X \geq a]$$

Nell'ultimo passaggio abbiamo utilizzato

$$P[X \geq a] = \int_a^\infty f(x)dx.$$

Analogamente, per la variabile casuale $g(X)$:

$$\begin{aligned}
E[g(X)] &= \int_{-\infty}^{\infty} g(x)f(x)dx = \\
&= \overbrace{\underbrace{\int g(x)f(x)dx}_{x:g(x)\geq r}}^{\geq 0} + \overbrace{\underbrace{\int g(x)f(x)dx}_{x:g(x)<r}}^{\geq 0} \geq \\
&\geq \int_{x:g(x)\geq r} g(x)f(x)dx \geq \int_{x:g(x)\geq r} rf(x)dx = \\
&= r\int_{x:g(x)\geq r} f(x)dx = rP[g(X) \geq r]
\end{aligned}$$

2) La disuguaglianza di Chebyshev afferma che più la varianza della variabile casuale X è piccola, più è piccola la probabilità che X assuma valori lontani dalla sua media.

3) La disuguaglianza di Chebyshev afferma che la probabilità che una qualsiasi variabile casuale X assuma un valore all'interno di k deviazioni standard dalla media è almeno $1 - \frac{1}{k^2}$, ad esempio:

- per $k = 2$, X ha una probabilità di almeno $3/4$ di cadere all'interno di 2 deviazioni standard dalla media μ $(\mu \pm 2\sigma)$;
- per $k = 3$, X ha una probabilità di almeno $8/9$ di cadere all'interno di 3 deviazioni standard dalla media μ $(\mu \pm 3\sigma)$.

4) Il valore determinato con la disuguaglianza di Chebyshev è riferito SOLO AL LIMITE INFERIORE, cioè si sa che la probabilità di una variabile casuale X di cadere all'interno di 2 deviazioni standard è almeno $3/4$, ma non si sa di quanto potrebbe essere maggiore di $3/4$.
Questo accade perchè NON È NOTA la funzione di densità $f(x)$ della variabile casuale X. Perciò l'utilizzo della disuguaglianza di Chebyshev è limitato ai casi in cui non è nota la forma della distribuzione della variabile casuale.

Esempio. Sia X la variabile casuale corrispondente al numero di pezzi prodotti da una fabbrica in una settimana. Si sa che $\mu_X = 50$.

- Calcolare la $P[X \geq 75]$.
- Sapendo che $\sigma_X^2 = 25$, dare un limite inferiore alla $P[40 < X < 60]$.

La funzione di densità $f(x)$ non è nota; posso però applicare Chebyshev e Markov.

- $P[X \geq 75] \leq \frac{1}{75}E[X] = \frac{1}{75} \cdot 50 = \frac{2}{3}$
- $P[40 < X < 60] =$
  $= P[40 - 50 < X - 50 < 60 - 50] =$
  $= P[|X - 50| < 10] = 1 - P[|X - 50| \geq 10]$
  ma
  $P[|X - 50| \geq 10] = P[(X - 50)^2 \geq 100]$,
  quindi, applicando Chebyshev
  $1 - P[40 < X < 60] \leq \frac{1}{100} \underbrace{E[(X - 50)^2]}_{=\mathrm{var}[X]} =$
  $= \frac{\sigma_X^2}{100} = \frac{25}{100} = \frac{1}{4}$
  da cui $P[40 < X < 60] \geq \frac{3}{4}$

Nota bene: se fosse nota la funzione di densità $f(x)$ associata ad X, potremmo calcolare ESATTAMENTE la $P[X \geq 75]$, poichè
$P[X \geq a] = 1 - P[X < a]$ e

- se X è continua $P[X < a] = \int_{-\infty}^{a} f(x)dx$
- se X è discreta $P[X < a] = P[X \leq a] - P[X = a] = F(a) - f(a)$

PROPRIETÀ DELLA VARIANZA

1. $\text{var}[X] = E[(X - E[X])^2] = E[(X - \mu_X)^2]$
   Segue direttamente dalla definizione di varianza e di valore atteso di una funzione di variabile casuale $(g(X) = (X - \mu_X)^2)$
2. $\text{var}[X] = E[X^2] - E[X]^2 = E[X^2] - \mu_X^2$

$$E[(X - E[X])^2] = E[X^2 - 2XE[X] + E[X]^2] =$$
$$= E[X^2] - 2E[XE[X]] + E[E[X]^2]$$

Ma $E[X]$ è un valore numerico, quindi si comporta come una costante, e dalla relativa proprietà del valore atteso $(E[cg(X)] = cE[g(X)])$,

$$\text{var}[X] = E[X^2] - 2E[X] \cdot E[X] + E[X]^2 =$$
$$= E[X^2] - E[X]^2$$

3. $\mathrm{var}[aX] = a^2\,\mathrm{var}[X], \ \forall a \in \mathbb{R}$

$$\begin{aligned}
&\mathrm{var}[aX] = E[(aX - E[aX])^2] = \\
&= E[(aX - aE[X])^2] = a^2 E[(X - E[X])^2] = \\
&= a^2\,\mathrm{var}[X]
\end{aligned}$$

4. $\mathrm{var}[X + a] = \mathrm{var}[X], \ \forall a \in \mathbb{R}$

$$\begin{aligned}
&\mathrm{var}[X + a] = E[((X + a) - E[X + a])^2] = \\
&= E[(X + a - E[X] - E[a])^2] = \\
&= E[(X + \cancel{a} - E[X] - \cancel{a})^2] = \mathrm{var}[X]
\end{aligned}$$

Nota bene: $\mathrm{var}[X + X] = 4\,\mathrm{var}[X]$ (per la proprietà 3) con $a = 2$)

Esempio

Sia X una variabile casuale continua e si consideri la funzione

$$f(x) = \begin{cases} \dfrac{c}{x^2} & 1 \le x \le 2\,, \\ 0 & \text{altrove}\,. \end{cases}$$

- Determinare il valore di c affinchè f sia una densità e disegnare il grafico di $f(x)$.
- Determinare la funzione di ripartizione $F(x)$ e disegnarne il grafico.

- Calcolare $P[1.5 < X < 2]$.
- Calcolare la media e la varianza di X.

1) $\forall x \in \mathbb{R}\,,\ f(x) \geq 0 \Rightarrow c \geq 0$
Inoltre vale:
$\int_{-\infty}^{\infty} f(x)dx = 1$
cioè:

$$1 = \int_1^2 \frac{c}{x^2}dx = c\int_1^2 \frac{1}{x^2}dx = c\,\frac{1}{-2+1}x^{-2+1}\Big|_1^2 =$$

$$= -c\,\frac{1}{x}\Big|_1^2 = \frac{c}{2} \Rightarrow c = 2$$

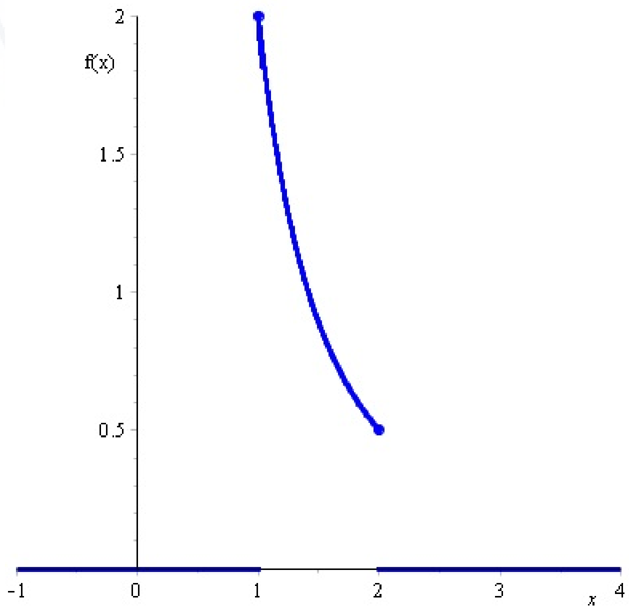

Figura 19

2)

$$\begin{cases} \text{se } x < 1 & F(x) = \int_{-\infty}^{x} f(t)dt = 0, \\ \\ \text{se } 1 \leq x < 2 & F(x) = \int_{-\infty}^{1} f(t)dt + \int_{1}^{x} f(t)dt, \\ & = \int_{1}^{x} \frac{2}{t^2}dt = 2(-\dfrac{1}{x} + 1), \\ \text{se } x \geq 2 & F(x) = \int_{-\infty}^{1} f(t)dt + \int_{1}^{2} f(t)dt + \\ & = + \int_{2}^{x} f(t)dt = 1, \end{cases}$$

quindi

$$F(x) = \begin{cases} 0 & \text{se } x < 1\,, \\ 2 - \dfrac{2}{x} & \text{se } 1 \leq x < 2\,, \\ 1 & \text{se } x \geq 2\,. \end{cases}$$

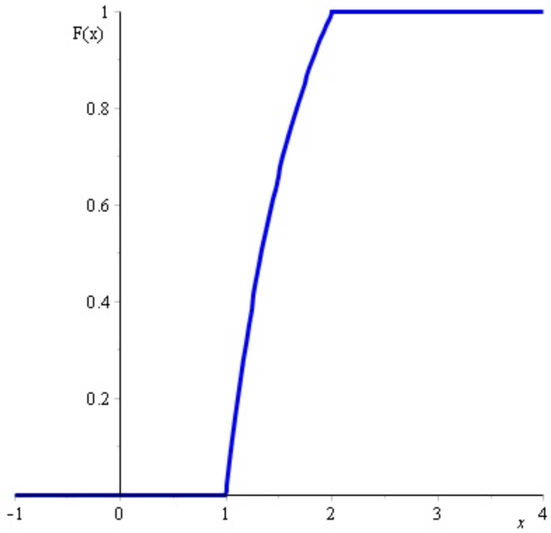

Figura 20

3) $P[1.5 < X < 2] = F(2) - F(1.5) =$
$= 1 - \left(2 - \frac{2}{x}\right)\Big|_{x=3/2} = 1 - \frac{2}{3} = \frac{1}{3}$
oppure, tramite la $f(x)$
$P[1.5 < X < 2] = \int_{1.5}^{2} \frac{2}{x^2} dx = 2\left(-\frac{1}{x}\right)\Big|_{1.5}^{2} = \frac{1}{3}$

4) $E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_{1}^{2} x \frac{2}{x^2} dx = 2 \int_{1}^{2} \frac{1}{x} dx =$
$= 2 \ln\left(|x|\right)\Big|_{1}^{2} = 2\ln(2) = \ln(4).$

$\mathrm{var}[X] = \sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x) dx,$
ma, anzichè applicare la definizione di varianza, utilizziamo la proprietà n. 2 della varianza, cioè:

$$\mathrm{var}[X] = E[X^2] - E[X]^2$$

visto che la media $\mu_X = E[X]$ è già stata calcolata.
Abbiamo
$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_1^2 x^2 \frac{2}{x^2} dx = 2x \mid_1^2 = 2$
quindi
$\text{var}[X] = 2 - (\ln(4))^2$

Quesito teorico

Sia X una variabile casuale. Se esiste $E[X]$ e la funzione di densità è simmetrica rispetto all'asse $x = a$, allora

$$E[X] = a$$

Una funzione $f(x)$ si dice simmetrica rispetto all'asse $x = a$ se

$$f(x) = f(2a - x)$$

Dimostrazione (caso continuo)

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{a} xf(x)dx + \int_{a}^{\infty} xf(x)dx$$

Poniamo $y = 2a - x \rightarrow dy = -dx$

$E[X] = \int_{-\infty}^{a} xf(x)dx + \int_{a}^{-\infty} (2a-y)f(2a-y)(-dy) =$

$$= \int_{-\infty}^{a} xf(x)dx - \int_{-\infty}^{a} (y-2a) \underbrace{f(2a-y)}_{=f(y)} dy =$$

$$= \cancel{\int_{-\infty}^{a} xf(x)dx} - \cancel{\int_{-\infty}^{a} yf(y)dy} + 2a\int_{-\infty}^{a} f(y)dy$$

ma f è simmetrica rispetto ad a, cioè:

$$\int_{-\infty}^{a} f(x)dx = \int_{a}^{\infty} f(x)dx = \frac{1}{2},$$

quindi

$$E[X] = 2a\int_{-\infty}^{a} f(y)dy = 2a \cdot \frac{1}{2} = a$$

I MOMENTI

Oltre al valore atteso e alla varianza esistono altre quantità che possono misurare le caratteristiche di una variabile casuale: sono i valori attesi delle potenze della variabile casuale, detti momenti.

DEFINIZIONE. Sia X una variabile casuale con media μ_X. Chiamiamo MOMENTO DI ORDINE r ($r \in \mathbb{N}$) la quantità:

$$\mu'_r = E[X^r]$$

dove

- $E[X^r] = \sum_i x_i^r f(x_i)$ se X è discreta.
- $E[X^r] = \int_{-\infty}^{\infty} x^r f(x) dx$ se X è continua.

Si noti che $\mu'_1 = E[X] = \mu_X$.

DEFINIZIONE: Sia X una variabile casuale con media μ_X. Chiamiamo MOMENTO CENTRALE DI ORDINE r ($r \in \mathbb{N}$) la quantità:

$$\mu_r = E[(X - \mu_X)^r]$$

dove

- $\mu_r = \sum_i (x_i - \mu_X)^r f(x_i)$ se X è discreta.
- $\mu_r = \int_{-\infty}^{\infty} (x - \mu_X)^r f(x) dx$ se X è continua.

Si noti che
$\mu_1 = E[X - \mu_X] = E[X] - \mu_X = 0$
$\mu_2 = E[(X - \mu_X)^2] = \mathrm{var}[X] = \sigma_X^2$

OSSERVAZIONE
Il momento centrale di ordine 3, μ_3 è un indice di asimmetria della curva $f(x)$; il momento centrale di ordine 4, μ_4 è un indice di appiattimento della curva $f(x)$ attorno al suo punto di massimo.

Se la funzione di densità $f(x)$ è simmetrica rispetto al valore atteso μ_X allora tutti i momenti centrali dispari $(r > 0)$ della variabile casuale X sono nulli, cioè:

$$\mu_{2r+1} = 0, \quad \forall r \in \mathbb{N}$$

Verifichiamolo (nel caso continuo) per $r = 1$, cioè per μ_3:

$$\mu_3 = \int_{-\infty}^{\infty} (x - \mu_X)^3 f(x) dx =$$

$$= \int_{-\infty}^{\mu_X} (x - \mu_X)^3 f(x) dx + \int_{\mu_X}^{\infty} (x - \mu_X)^3 f(x) dx =$$

$$x = 2a - y \;\; \text{con} \;\; a = \mu_X \quad f(y) = f(2a - y)$$

$$= \int_{-\infty}^{\mu_X} (x - \mu_X)^3 f(x) dx + \underbrace{\int_{-\infty}^{\mu_X} (\mu_X - y)^3 f(y) dy}_{-\int_{-\infty}^{\mu_X} (y - \mu_X)^3 f(y) dy} = 0$$

OSSERVAZIONE

Poichè $\text{var}[X] = E[X^2] - E[X]^2$ possiamo anche scrivere

$$\sigma_X^2 = \mu_2' - (\mu_1')^2 \quad \text{poichè} \quad \mu_1' = \mu_X$$

Sebbene i momenti di una variabile casuale possano essere determinati attraverso la definizione, è possibile utilizzare una procedura alternativa che si basa

sulla funzione generatrice dei momenti.
DEFINIZIONE: Data una variabile casuale X, $\forall t \in I$ con $I = [-h, h], h > 0$, definiamo FUNZIONE GENERATRICE DEI MOMENTI di X la quantità:

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

se tale valore è finito, dove

- $m(t) = \sum_i e^{tx_i} f(x_i)$ se X è discreta.
- $m(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$ se X è continua.

Se si può definire la funzione generatrice dei momenti di una variabile casuale X, allora tale funzione può essere utilizzata per generare tutti i momenti di X perchè

$$m^{(r)}(0) = \frac{d^r}{dt^r} m(t)\Big|_{t=0} = \mu'_r$$

Verifichiamolo nel caso continuo:

$$\frac{d}{dt} m(t) = \int_{-\infty}^{\infty} x e^{tx} f(x) dx$$
$$\frac{d^2}{dt^2} m(t) = \int_{-\infty}^{\infty} x^2 e^{tx} f(x) dx$$
$$\vdots$$

$$\vdots$$

$$\frac{d^r}{dt^r}m(t) = \int_{-\infty}^{\infty} x^r e^{tx} f(x)dx$$

Calcoliamo ora

$$\underbrace{\lim_{t\to 0} \frac{d^r}{dt^r}m(t)}_{=\frac{d^r}{dt^r}m(t)\big|_{t=0}} = \int_{-\infty}^{\infty} x^r f(x)dx = \mu'_r$$

Esempio. Durata di una conversazione telefonica

La variabile casuale continua X abbiamo visto essere rappresentata dalla funzione di densità:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x \geq 0, \quad \lambda > 0\,, \\ 0 & \text{se } x < 0\,. \end{cases}$$

Quindi

$$m(t) = E[e^{tX}] = \int_{-\infty}^{\infty} f(x)e^{tx}dx = \int_0^{\infty} \lambda e^{-\lambda x}e^{tx}dx$$

$$= \frac{\lambda}{-(\lambda - t)} \int_0^{\infty} \left(-(\lambda - t)e^{-(\lambda - t)x}\right) dx$$

$$= -\frac{\lambda}{\lambda - t} e^{-(\lambda - t)x}\Big|_0^{\infty} = \frac{\lambda}{\lambda - t}, \quad t < \lambda$$

Conoscendo $m(t)$, possiamo calcolare $E[X^r]$ per ogni r tramite derivazione. Abbiamo:
$m'(t) = \frac{\lambda}{(\lambda-t)^2}$, $m'(0) = \frac{1}{\lambda} = E[X] = \mu_X = \mu'_1$
$m''(t) = \frac{2\lambda}{(\lambda-t)^3}$, $m''(0) = \frac{2}{\lambda^2} = E[X^2] = \mu'_2 \Rightarrow$
$\Rightarrow \sigma_X^2 = m''(0) - m'(0)^2 = \frac{1}{\lambda^2}$

Esercizio

Mostrare che per una variabile casuale esponenziale di parametro λ il momento di ordine r è dato da $\mu'_r = \dfrac{r!}{\lambda^r}$.