Azzolini Riccardo 2020-04-02

Momenti e funzioni generatrici delle variabili aleatorie

1 Momenti di una variabile aleatoria

Definizione: Sia X una variabile aleatoria. Si dice **momento di ordine n** di X, per $n = 1, 2, \dots$, il valore $E(X^n)$.

In particolare:

- il momento di ordine 1, $E(X^1) = E(X)$, è semplicemente il valore medio;
- il momento di ordine 2, $E(X^2)$, può essere usato (insieme al valore medio) per calcolare la varianza, sfruttando la formula

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

  già introdotta in precedenza.

Ci sono vari modi per calcolare i momenti di una variabile aleatoria. Il più immediato è quello che sfrutta la formula e le proprietà del valore medio:

$$E(X^n) = \sum_i x_i^n p(x_i)$$

(dove x_i sono i valori assunti da X, e p è la densità di X).

Nota: Il momento di ordine n di X esiste ed è finito solo se la serie $\sum_i x_i^n p(x_i)$ converge assolutamente.

2 Funzione generatrice dei momenti

Definizione: Si definisce **funzione generatrice dei momenti** $m_X(t)$ di una variabile aleatoria discreta X la funzione

$$m_X(t) = E(e^{tX}) = \sum_i e^{tx_i} p(x_i) \qquad t \in \mathbb{R}$$

X possiede la funzione generatrice dei momenti se $\exists \delta > 0$ tale che

$$\left|E(e^{tX})\right| < \infty \quad \forall t \in (-\delta, \delta)$$

cioè se esiste un intorno $(-\delta, \delta)$ di $t = 0$ nel quale il valore medio $E(e^{tX})$ esiste ed è finito. In tal caso, per ogni $r = 1, 2, \ldots$, esiste finito $E(X^r)$ (il momento di ordine r di X), e questa funzione generatrice permette di calcolarlo, determinando la derivata di ordine r della funzione e valutandola in $t = 0$:

$$E(X^r) = \frac{d^r}{dt^r} E(e^{tX}) \Big|_{t=0}$$

Infatti, facendo la derivata si "porta davanti" all'esponenziale il valore di X, ma l'esponenziale rimane, quindi si valuta in $t = 0$ per "farla sparire" (mettendo a 0 il suo esponente, essa assume valore 1), ottenendo così le formule dei momenti. Ad esempio:

- per $r = 1$:

$$\begin{aligned}
\frac{d}{dt} E(e^{tX}) &= \frac{d}{dt} \sum_i e^{tx_i} p(x_i) \\
&= \sum_i \frac{d}{dt} e^{tx_i} p(x_i) \\
&= \sum_i x_i e^{tx_i} p(x_i) \\
\frac{d}{dt} E(e^{tX}) \Big|_{t=0} &= \sum_i x_i e^{tx_i} p(x_i) \Big|_{t=0} \\
&= \sum_i x_i e^0 p(x_i) \\
&= \sum_i x_i p(x_i) = E(X)
\end{aligned}$$

- per $r = 2$:

$$\begin{aligned}
\frac{d^2}{dt^2} E(e^{tX}) &= \frac{d}{dt} \left(\frac{d}{dt} E(e^{tX}) \right) \\
&= \frac{d}{dt} \sum_i x_i e^{tx_i} p(x_i) \\
&= \sum_i \frac{d}{dt} x_i e^{tx_i} p(x_i) \\
&= \sum_i x_i^2 e^{tx_i} p(x_i) \\
\frac{d^2}{dt^2} E(e^{tX}) \Big|_{t=0} &= \sum_i x_i^2 e^{tx_i} p(x_i) \Big|_{t=0} \\
&= \sum_i x_i^2 e^0 p(x_i) \\
&= \sum_i x_i^2 p(x_i) = E(X^2)
\end{aligned}$$

2.1 Utilizzo

Le funzioni generatrici dei momenti hanno varie applicazioni, tra cui le seguenti:

- Quando sono calcolabili, permettono di ricavare i parametri di media e varianza.
- Permettono di identificare una distribuzione, poiché si può dimostrare che, se due variabili aleatorie hanno la stessa funzione generatrice dei momenti, allora hanno anche la stessa distribuzione.
- Permettono il calcolo di operazioni su più variabili aleatorie: si fanno i calcoli sulle funzioni generatrici, ottenendo una nuova funzione dalla quale può essere identificata la distribuzione della variabile aleatoria risultante da tali operazioni.

2.2 Esempio: distribuzione binomiale

La funzione generatrice dei momenti di una variabile aleatoria binomiale $X \sim B(n, p)$ è:

$$\begin{aligned}
E(e^{tX}) &= \sum_{x=0}^{n} e^{tx} p(x) \\
&= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} \binom{n}{x} p^x e^{tx} (1-p)^{n-x} \\
&= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x} \qquad \text{(sviluppo di un binomio elevato alla } n\text{)} \\
&= (pe^t + (1-p))^n
\end{aligned}$$

Essa può poi essere usata per determinare la media e la varianza di X:

1. si calcola il momento di ordine 1, cioè il valore medio:

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

$$\begin{aligned}
E(X) &= \left.\frac{d}{dt}(pe^t + 1 - p)^n\right|_{t=0} \\
&= npe^t(pe^t + 1 - p)^{n-1}\big|_{t=0} \\
&= npe^0(pe^0 + 1 - p)^{n-1} \\
&= np(p + 1 - p)^{n-1} \\
&= np \cdot 1^{n-1} \\
&= np
\end{aligned}$$

2. si calcola il momento di ordine 2:

$$
\begin{aligned}
\frac{d^2}{dt^2}(pe^t+1-p)^n &= \frac{d}{dt} npe^t(pe^t+1-p)^{n-1} \\
&= np\big[e^t(pe^t+1-p)^{n-1} + e^t(n-1)(pe^t+1-p)^{n-2}pe^t\big] \\
&= npe^t\big[(pe^t+1-p)^{n-1} + (n-1)(pe^t+1-p)^{n-2}pe^t\big] \\
&= npe^t(pe^t+1-p)^{n-2}\big[(pe^t+1-p) + pe^t(n-1)\big] \\
&= npe^t(pe^t+1-p)^{n-2}\big[pe^t+1-p+npe^t-pe^t\big] \\
&= npe^t(pe^t+1-p)^{n-2}(1-p+npe^t)
\end{aligned}
$$

$$
\begin{aligned}
E(X^2) &= \left.\frac{d^2}{dt^2}(pe^t+1-p)^n\right|_{t=0} \\
&= \left. npe^t(pe^t+1-p)^{n-2}(1-p+npe^t)\right|_{t=0} \\
&= npe^0(pe^0+1-p)^{n-2}(1-p+npe^0) \\
&= np(1-p+np)
\end{aligned}
$$

3. si calcola la varianza sfruttando i due momenti appena ricavati:

$$
\begin{aligned}
\mathrm{Var}(X) &= E(X^2) - (E(X))^2 \\
&= np(1-p+np) - (np)^2 \\
&= np(1-p+np-np) \\
&= np(1-p)
\end{aligned}
$$

2.3 Esempio: distribuzione di Poisson

Sia X una variabile aleatoria di Poisson di parametro λ. La sua funzione generatrice dei momenti è:

$$\begin{aligned}
E(e^{tX}) &= \sum_{x=0}^{\infty} e^{tx} p(x) \\
&= \sum_{x=0}^{\infty} e^{tx} e^{-\lambda} \frac{\lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x e^{tx}}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} \qquad \left(\text{serie di Taylor } \sum_{x=0}^{\infty} \frac{(f(t))^x}{x!} = e^{f(t)}, \text{ con } f(t) = \lambda e^t\right) \\
&= e^{-\lambda} e^{\lambda e^t} \\
&= e^{\lambda e^t - \lambda} \\
&= e^{\lambda(e^t - 1)}
\end{aligned}$$

Si possono allora determinare la media e la varianza della distribuzione di Poisson:

1. valore medio (momento di ordine 1):

$$\begin{aligned}
\frac{d}{dt} e^{\lambda(e^t-1)} &= e^{\lambda(e^t-1)} \lambda e^t \\
&= \lambda e^{\lambda(e^t-1)+t}
\end{aligned}$$

$$\begin{aligned}
E(X) &= \left. \frac{d}{dt} e^{\lambda(e^t-1)} \right|_{t=0} \\
&= \left. \lambda e^{\lambda(e^t-1)+t} \right|_{t=0} \\
&= \lambda e^{\lambda(e^0-1)+0} \\
&= \lambda e^{\lambda(1-1)} \\
&= \lambda e^0 \\
&= \lambda
\end{aligned}$$

2. momento di ordine 2:

$$\begin{aligned}
\frac{d^2}{dt^2}\, e^{\lambda(e^t-1)} &= \frac{d}{dt}\lambda e^{\lambda(e^t-1)+t} \\
&= \lambda e^{\lambda(e^t-1)+t}(\lambda e^t + 1) \\
E(X^2) &= \left.\frac{d^2}{dt^2}\, e^{\lambda(e^t-1)}\right|_{t=0} \\
&= \lambda e^{\lambda(e^t-1)+t}(\lambda e^t + 1)\big|_{t=0} \\
&= \lambda e^{\lambda(e^0-1)+0}(\lambda e^0 + 1) \\
&= \lambda e^0(\lambda + 1) \\
&= \lambda(\lambda + 1) \\
&= \lambda^2 + \lambda
\end{aligned}$$

3. varianza:

$$\begin{aligned}
\mathrm{Var}(X) &= E(X^2) - (E(X))^2 \\
&= \lambda^2 + \lambda - \lambda^2 \\
&= \lambda
\end{aligned}$$

Si osserva così che, per la distribuzione di Poisson, sia la media che la varianza sono uguali al parametro λ.

3 Funzione generatrice delle probabilità

Sia X una variabile aleatoria discreta a valori interi ≥ 0. Si chiama **funzione generatrice delle probabilità** di X la funzione

$$\psi_X(z) = E(z^X) \qquad z \in \mathbb{R}$$

ossia

$$\begin{aligned}
\psi_X(z) &= \sum_{n=0}^{\infty} z^n p(n) \\
&= p(0) + zp(1) + z^2 p(2) + z^3 p(3) + \cdots
\end{aligned}$$

Osservazione: Le funzioni generatrici delle probabilità e dei momenti sono di fatto equivalenti:

$$\begin{aligned}
m_X(z) &= E(e^{zX}) = \psi_X(e^z) \\
\psi_X(z) &= E(z^X) = E(e^{X \log z}) = m_X(\log z)
\end{aligned}$$

3.1 Calcolo delle densità

Dalla funzione generatrice delle probabilità, è possibile ritrovare la densità di X:

$$P\{X = n\} = p(n) = \frac{1}{n!}\frac{d^n}{dz^n}\psi_X(z)\bigg|_{z=0}$$

Ad esempio:

$$\begin{aligned}
\frac{1}{0!}\frac{d^0}{dz^0}\psi_X(z)\bigg|_{z=0} &= \psi_X(0) \\
&= p(0) + 0p(1) + 0p(2) + 0p(3) + 0\cdots \\
&= p(0) \\
\frac{1}{1!}\frac{d}{dz}\psi_X(z)\bigg|_{z=0} &= \frac{d}{dz}\big[p(0) + zp(1) + z^2p(2) + z^3p(3) + \cdots\big]\bigg|_{z=0} \\
&= p(1) + 2zp(2) + 3z^2p(3) + \cdots\big|_{z=0} \\
&= p(1) + 2\cdot 0p(2) + 3\cdot 0p(3) + 0\cdots \\
&= p(1) \\
\frac{1}{2!}\frac{d^2}{dz^2}\psi_X(z)\bigg|_{z=0} &= \frac{1}{2}\frac{d}{dz}\big[p(1) + 2zp(2) + 3z^2p(3) + \cdots\big]\bigg|_{z=0} \\
&= \frac{1}{2}\big[2p(2) + 3\cdot 2zp(3) + \cdots\big]\bigg|_{z=0} \\
&= \frac{1}{2}\big[2p(2) + 6\cdot 0p(3) + 0\cdots\big] \\
&= \frac{1}{2}\cdot 2p(2) \\
&= p(2) \\
\frac{1}{3!}\frac{d^3}{dz^3}\psi_X(z)\bigg|_{z=0} &= \frac{1}{6}\frac{d}{dz}\big[2p(2) + 6zp(3) + \cdots\big]\bigg|_{z=0} \\
&= \frac{1}{6}\big[6p(3) + \cdots\big]\bigg|_{z=0} \\
&= \frac{1}{6}\big[6p(3) + 0\cdots\big] \\
&= \frac{1}{6}\cdot 6p(3) \\
&= p(3)
\end{aligned}$$

3.2 Calcolo di media e varianza

La funzione generatrice delle probabilità può anche essere usata per determinare la media e la varianza di X, se la serie $\psi_X(z) = \sum_{n=0}^{\infty} z^n p(n)$ converge in $z = 1$ (per la precisione, se ha raggio di convergenza > 1).

1. La derivata prima di $\psi_X(z)$ è

$$\psi'_X(z) = \frac{d}{dz}\psi_X(z) = \frac{d}{dz}\sum_{n=0}^{\infty} z^n p(n) = \sum_{n=0}^{\infty} \frac{d}{dz} z^n p(n) = \sum_{n=1}^{\infty} n z^{n-1} p(n)$$

   (si osserva che la serie risultante inizia da $n = 1$, perché la derivata per $n = 0$ vale 0, quindi non serve includerla nella somma), che, valutata in $z = 1$, corrisponde al valore medio:

$$\psi'_X(1) = \left.\frac{d}{dz}\psi_X(z)\right|_{z=1} = \sum_{n=1}^{\infty} n 1^{n-1} p(n) = \sum_{n=1}^{\infty} n p(n) = E(X)$$

2. Calcolando la derivata seconda,

$$\psi''_X(z) = \frac{d^2}{dz^2}\psi_X(z) = \frac{d}{dz}\sum_{n=1}^{\infty} n z^{n-1} p(n) = \sum_{n=2}^{\infty} n(n-1) z^{n-2} p(n)$$

   e valutandola in $z = 1$, si ottiene

$$\psi''_X(1) = \left.\frac{d^2}{dz^2}\psi_X(z)\right|_{z=1} = \sum_{n=2}^{\infty} n(n-1) 1^{n-2} p(n) = \sum_{n=2}^{\infty} (n^2 - n) p(n) = E(X^2 - X)$$

   che, per la linearità del valore medio, corrisponde alla differenza di momenti $E(X^2) - E(X)$.

3. Si può infine calcolare la varianza:

$$\begin{aligned}
\mathrm{Var}(X) &= E(X^2) - (E(X))^2 \\
&= \underbrace{E(X^2) - E(X)}_{\psi''_X(1)} + \underbrace{E(X)}_{\psi'_X(1)} - (\underbrace{E(X)}_{\psi'_X(1)})^2 \\
&= \psi''_X(1) + \psi'_X(1) - (\psi'_X(1))^2 \\
&= \left.\frac{d^2}{dz^2}\psi_X(z)\right|_{z=1} + \left.\frac{d}{dz}\psi_X(z)\right|_{z=1} - \left(\left.\frac{d}{dz}\psi_X(z)\right|_{z=1}\right)^2
\end{aligned}$$

3.3 Esempio: distribuzione geometrica

Sia X una variabile aleatoria geometrica di parametro p. La sua funzione generatrice $\psi_X(z)$ viene calcolata nel modo seguente:

$$
\begin{aligned}
\psi_X(z) &= \sum_{n=0}^{\infty} z^n p(n) \\
&= \sum_{n=0}^{\infty} z^n p(1-p)^n \\
&= p\sum_{n=0}^{\infty} (z(1-p))^n \qquad\qquad \text{(serie geometrica di ragione } q = z(1-p)) \\
&= p\frac{1}{1-z(1-p)} \\
&= \frac{p}{1-z(1-p)}
\end{aligned}
$$

Nota: Questa serie geometrica converge se e solo se $|z(1-p)| < 1$, quindi $\psi_X(z)$ è definita solo per $|z| < \frac{1}{1-p}$.

Da $\psi_X(z)$, si possono dedurre media e varianza della distribuzione geometrica:

$$
\begin{aligned}
E(X) &= \frac{d}{dz}\left.\frac{p}{1-z(1-p)}\right|_{z=1} \\
&= p(-1)\left.\frac{1}{(1-z(1-p))^2}(-(1-p))\right|_{z=1} \\
&= p\frac{1}{(1-1(1-p))^2}(1-p) \\
&= \frac{p(1-p)}{(1-1+p)^2} \\
&= \frac{p(1-p)}{p^2} \\
&= \frac{1-p}{p}
\end{aligned}
$$

$$
\begin{aligned}
E(X^2) - E(X) &= \frac{d}{dz}\left.\frac{p(1-p)}{(1-z(1-p))^2}\right|_{z=1} \\
&= p(1-p)(-2)\left.\frac{1}{(1-z(1-p))^3}(-(1-p))\right|_{z=1} \\
&= 2p(1-p)\frac{1}{(1-1(1-p))^3}(1-p) \\
&= 2\frac{p(1-p)^2}{p^3} \\
&= 2\frac{(1-p)^2}{p^2}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(X) &= E(X^2) - E(X) + E(X) - (E(X))^2 \\
&= 2\frac{(1-p)^2}{p^2} + \frac{1-p}{p} - \left(\frac{1-p}{p}\right)^2 \\
&= 2\frac{(1-p)^2}{p^2} + \frac{1-p}{p} - \frac{(1-p)^2}{p^2} \\
&= \frac{(1-p)^2}{p^2} + \frac{1-p}{p} \\
&= \frac{1-p}{p}\left(\frac{1-p}{p} + 1\right) \\
&= \frac{1-p}{p} \cdot \frac{1-p+p}{p} \\
&= \frac{1-p}{p} \cdot \frac{1}{p} \\
&= \frac{1-p}{p^2}
\end{aligned}
$$

Quindi, ricapitolando, la variabile aleatoria geometrica X (di parametro p) ha:

$$
E(X) = \frac{1-p}{p} \qquad \mathrm{Var}(X) = \frac{1-p}{p^2}
$$

4 Media e varianza della distribuzione ipergeometrica

Sia X una variabile aleatoria ipergeometrica di parametri n_1, n_2, n, e sia $N = n_1 + n_2$. Tramite calcoli piuttosto complicati, si determina che essa ha valore medio

$$
E(X) = n\frac{n_1}{N}
$$

e varianza

$$\begin{aligned}
\mathrm{Var}(X) &= n\frac{n_1}{N}\left(1-\frac{n_1}{N}\right)\frac{N-n}{N-1} \\
&= n\frac{n_1}{N}\frac{N-n_1}{N}\frac{N-n}{N-1} \\
&= n\frac{n_1}{N}\frac{n_2}{N}\frac{N-n}{N-1}
\end{aligned}$$

Data la difficoltà dei calcoli relativi alla distribuzione ipergeometrica, è spesso utile approssimarla con la binomiale $B(n, \frac{n_1}{N})$:

$$\frac{\binom{n_1}{k}\binom{n_2}{n-k}}{\binom{N}{n}} \approx \binom{n}{k}\left(\frac{n_1}{N}\right)^k\left(1-\frac{n_1}{N}\right)^{n-k}$$

Questa è una buona approssimazione se $n < \frac{N}{10}$ (cioè se il numero di prove n è piccolo rispetto al numero di elementi disponibili N: in tal caso, non c'è molta differenza tra uno schemi successo-insuccesso con rimpiazzo e uno senza).