CASO DI ALFABETO INFINITO

Variabili aleatorie geometriche

$X \sim Ge(p)$ con $p \in (0,1)$

se ha $\begin{cases} \text{alfabeto} & \mathcal{X} = \mathbb{N} = \{1, 2, \dots\} \\ \text{dens. discr.} & p_X(k) = (1-p)^{k-1} p, \quad k \in \mathcal{X} \end{cases}$

Inoltre, $E(X) = \frac{1}{p}$ e $Var(X) = \frac{1-p}{p^2}$

La v.al. geom. corrisponde al nr. di prove che devo effettuare per osservare il primo successo in uno schema di Bernoulli dove il <u>nr. di prove non è necessariamente predefito</u> e dove la prob. di successo è p.

<u>Contesto</u>: le v.al. $X_i \sim Be(p)$ $(i \geq 1)$ sono gli esiti delle prove ripetute (indip.) e la v. al.

$$X = \inf_{i \in \mathbb{N}} \{X_i = 1\} \sim Ge(p)$$

rappresenta la prova a cui osservo il primo successo.
Cruciale: indip. degli eventi $\{X_1 = 1\}, \{X_2 = 1\}, \dots$

<u>Interpretazione densità:</u>

$$p_X(k) = P(\text{ottenere il primo successo alla prova } k)$$
$$= P(\underbrace{X_1 = \dots = X_{k-1} = 0}_{\text{sequenza di insuccessi}}, \ X_k = 1)$$

$$\xrightarrow{\text{indip.}} = \prod_{j=1}^{k-1} P(X_j = 0) \cdot P(X_k = 1)$$

$$= (1-p)^{k-1} \, p$$

<u>**VARIANTE**</u>: se invece di guardare all'istante del primo successo, guardo al nr. di insuccessi che ho avuto prima di ottenere il primo successo, ottengo una v.al. $X' = X - 1$, con $X \sim Ge(p)$, detta v. al. geometrica traslata.

$$\underbrace{\overset{X_1}{0} \quad \overset{X_2}{0} \quad \overset{X_3}{0} \quad \overset{X_4}{0}}_{X'=4} \quad \overset{X_5}{1} \ \leftarrow x = 5$$

Densità discreta: $P(X' = k) = P(X = k+1) = (1-p)^k p$
per $k \in \mathbb{N}_0 = \{0, 1, \dots\}$

Media: $E(X') = E(X-1) = E(X) - 1 = \frac{1}{p} - 1 = \frac{1-p}{p}$

Varianza: $Var(X') = Var(X-1) = Var(X) = \frac{1-p}{p^2}$

Probabilità di lunga attesa. Nelle applicazioni è interessante conoscere la prob. di dover attendere il primo successo per più di k prove. Per ogni $k \geq 0$, si ha

$$P(X > k) = \sum_{i=k+1}^{+\infty} P(X=i) = \sum_{i=k+1}^{+\infty} (1-p)^{i-1} p$$

$$\xrightarrow[h=i-k-1]{\text{cambio indice}} = \sum_{h=0}^{+\infty} (1-p)^{h+k} p = p(1-p)^k \underbrace{\sum_{h=0}^{+\infty} (1-p)^h}_{= \frac{1}{1-(1-p)} \text{ serie geometrica}}$$

$$= (1-p)^k$$

Riassumendo, abbiamo ottenuto $P(X>k) = (1-p)^k$.

$\boxed{\text{ESERCIZIO}}$

Alla roulette (numeri da 0 a 36) si scommette ripetutamente su un numero tra 1 e 12 (compresi).

Si calcoli: (a) la probabilità di perdere nelle prime cinque giocate; (b) la probabilità di vincere alla sesta giocata per la prima volta.

Soluzione. Costruiamo uno schema di Bernoulli. Per ogni $i \geq 1$, definiamo le v.al.

$$X_i = \begin{cases} 1 & \text{se all'i-esima giocata esce un nr. tra } 1 \text{ e } 12 \quad (\text{vinco}) \\ 0 & \text{altrimenti} \end{cases}$$

Si ha $X_i \sim Be(12/37)$. Inoltre, assumiamo $\{X_1 = 1\}$, $\{X_2 = 1\}, \ldots$ sono indipendenti. La prima vincita è descritta dalla v.al. $X = \inf_{i \in \mathbb{N}} \{X_i = 1\} \sim Ge(12/37)$. Pertanto

(a) $P(X > 5) = (1 - 12/37)^5 \approx 0.14$

(b) $P(X = 6) = (1 - 12/37)^5 \cdot 12/37 \approx 0.046$

Variabili aleatorie di Poisson

$X \sim Po(\lambda)$ con $\lambda > 0$

se ha
$$\begin{cases} \text{alfabeto} & \mathcal{X} = \mathbb{N}_0 = \{0, 1, \dots\} \\ \text{dens. discr.} & p_X(k) = e^{-\lambda} \dfrac{\lambda^k}{k!}, \quad k \in \mathcal{X} \end{cases}$$

Inoltre, $E(X) = \lambda$ e $Var(X) = \lambda$.

ESERCIZIO

Il numero di meteoriti che colpisce un satellite durante la sua orbita si distribuisce come una variabile aleatoria di Poisson di parametro λ. Nel compiere la sua orbita il satellite impiega un giorno ed è mediamente colpito da 3 meteoriti. Si calcoli la probabilità che nel percorrere 5 orbite il numero di meteoriti che colpiscono il satellite sia minore o uguale a 3.

Soluzione. Il nr. medio di meteoriti che colpiscono il satellite in un periodo di 5 giorni è dato da $5 \cdot 3 = 15$. Quindi il nr. di meteoriti che colpiscono

il satellite nel percorrere 5 orbite è $X \sim Po(15)$.
Calcoliamo

$$P(X \leq 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$$
$$= e^{-15} + e^{-15} \cdot (15) + e^{-15} \cdot \frac{(15)^2}{2!} + e^{-15} \cdot \frac{(15)^3}{3!}$$
$$\approx 0.000\,2$$

TEOREMA LIMITE DI POISSON

Una v.al. binomiale con n molto grande e p molto piccolo può essere approssimata da una v.al. di Poisson di parametro $\lambda = np$.

<u>Teorema</u>. Siano $X_n \sim Bin(n, \lambda/n)$ e $Y \sim Po(\lambda)$. Allora, per ogni $k \in \mathbb{N}_0$ fissato, si ha

$$\lim_{n \to +\infty} p_{X_n}(k) = p_Y(k).$$

<u>Dimostrazione</u>. Si ha

$$p_{X_n}(k) = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{n!}{k!(n-k)!} \cdot \frac{\lambda^k}{n^k} \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{n^k} \cdot \frac{\lambda^k}{k!} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-k}$$

Prendiamo il limite. Analizziamo i termini uno a uno:

- $\lim_{n \to +\infty} \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{n^k} = 1$
- $\lim_{n \to +\infty} \left(1 - \frac{\lambda}{n}\right)^{-k} = 1$
- $\lim_{n \to +\infty} \left(1 - \frac{\lambda}{n}\right)^{n} = e^{-\lambda}$

Quindi, mettendo tutto insieme, otteniamo

$$\lim_{n \to +\infty} p_{X_n}(k) = e^{-\lambda} \frac{\lambda^k}{k!} = p_Y(k).$$

$\blacksquare$

Oss. Quando vale l'approssimazione?
Euristica: $n > 100$, $\boxed{p < 0.01}$ e $np \leq 20$.
EVENTI RARI

COME SI USA IL TEOREMA LIMITE DI POISSON NELLE APPLICAZIONI?

Molte prove ripetute	$n \gg 1$	entrambi NON CONOSCIUTI
Prob. di successo molto piccola	$p \ll 1$	

Si conosce invece, in quanto misurabile, $np = \lambda$.
Quindi il nr. successi nella seq. delle prove, che si distr. come $Bin(n, \lambda/n)$, con n incognito, si può approssimare con $Po(\lambda)$.
(approssimazione di Poisson).

CALCOLABILITÀ

<u>Esempio</u>

Modelliamo nr. accessi al minuto al sito unipol.it. Ci aspettiamo: nr. utenti collegati ad internet sia molto grande ($n \gg 1$) e che ciascuno di essi abbia una prob. bassa ($p \ll 1$) di collegarsi al sito. Sia n che p sono non noti. Possiamo guardare le statistiche degli accessi dei giorni precedenti e stimare il nr. medio di accessi al minuto, cioè np.

Se, ad esempio, volessimo calcolare la prob. che nel prossimo minuto ci fossero almeno due accessi, supponendo di aver stimato $np = 4$, otterremmo

X = # accessi per minuto

$P(X \geq 2)$ con $X \sim Bin(n, 4/n)$, n incognito

??

$P(Y \geq 2)$ con $Y \sim Po(4)$

$\parallel$

$1 - e^{-4} - 4e^{-4}$.

ESERCIZI

1. La probabilità che in una partita di poker venga servito un full è 0.0014. Vengono giocate 1000 partite. Calcolare la probabilità che vengano serviti al più tre full.

<u>Soluzione</u>. Costruiamo uno schema di Bernoulli. Per $i = 1, \dots, 1000$ definiamo le v. al.

$$X_i = \begin{cases} 1 & \text{se all'i-esima partita viene servito un full} \\ 0 & \text{altrimenti} \end{cases}$$

Si ha $X_i \sim Be(0.0014)$. Inoltre, assumiamo che $\{X_1 = 1\}, \dots, \{X_{1000} = 1\}$ siano indipendenti. Allora

$$X = \sum_{i=1}^{1000} X_i \sim Bin(1000,\ 0.0014)$$

è il nr. totale di full serviti. Calcoliamo

$$P(X \leq 3) \approx P(Y \leq 3) \qquad \text{dove} \quad Y \sim Po(1.4)$$

$$= e^{-1.4}\left[1 + 1.4 + \frac{(1.4)^2}{2!} + \frac{(1.4)^3}{3!}\right]$$

$$\approx 0.95$$

2. Il numero di volte che una persona contrae l'influenza in un anno si distribuisce come una variabile aleatoria di Poisson di parametro $\lambda = 5$.
Viene proposto un vaccino che fa diminuire a 3 il numero medio di influenze contratte annualmente. Il vaccino funziona però solo nel 75% dei casi.
Se un individuo si vaccina e quell'anno prende l'influenza solo 2 volte, qual è la probabilità che il vaccino sia stato efficace?

<u>Soluzione</u>. Definiamo gli eventi E = "il vaccino

è stato efficace", F = "l'individuo vaccinato contrae l'influenza 2 volte" e le v.a. $X_i \sim Po(i)$ con $i = 3, 5$

Dal testo ricaviamo che $P(E) = 0.75$.

Dobbiamo calcolare $P(E|F)$. Usiamo la formula di Bayes

$$P(E|F) = \frac{P(F|E)P(E)}{P(F)} = \frac{P(F|E)P(E)}{\underbrace{P(F|E)P(E) + P(F|E^c)P(E^c)}_{\text{prob. tot.}}}$$

$$= \frac{P(X_3 = 2)\, P(E)}{P(X_3 = 2)\, P(E) + P(X_5 = 2)\, P(E^c)}$$

$$= \frac{e^{-3} \cdot \frac{3^2}{2!} \cdot (0.75)}{e^{-3} \cdot \frac{3^2}{2!} \cdot (0.75) + e^{-5} \cdot \frac{5^2}{2!} \cdot (0.25)}$$

$$\approx 0.89.$$