

AVVISO: LA LEZIONE DI DOMANI È CANCELLATA

Studiamo il comportamento di $X \sim B(n, p)$ nel caso in cui $n \gg 1$ e $p \ll 1$.

Poniamo $\lambda = np$. Teniamo fisso λ , mantenendo $n \rightarrow +\infty$, dunque $p = \frac{\lambda}{n}$

$$\begin{aligned} P(X=K) &= \binom{n}{K} p^K (1-p)^{n-K} = \frac{n(n-1)\dots(n-K+1)}{K!} \frac{\lambda^K}{n^K} \left(1 - \frac{\lambda}{n}\right)^{n-K} \\ &= \frac{\lambda^K}{K!} \underbrace{\frac{n(n-1)\dots(n-K+1)}{n^K}}_{\xrightarrow{n \rightarrow +\infty} 1} \frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^K} \xrightarrow{n \rightarrow +\infty} e^{-\lambda} \end{aligned}$$

Risumando: se $X \sim B(n, \frac{\lambda}{n})$ allora

$$\lim_{n \rightarrow +\infty} P(X=K) = e^{-\lambda} \frac{\lambda^K}{K!}$$

Questo suggerisce, più in generale che se $X \sim B(n, p)$ $n \gg 1$, $p \ll 1$ allora

$$P(X=K) \simeq e^{-np} \frac{(np)^K}{K!}$$

Variabili aleatorie di Poisson

Diciamo che una v.e. X che assume i valori $0, 1, \dots, n$ (tutti gli interi ≥ 0) è di Poisson di parametro $\lambda > 0$ se $\forall k \geq 0$

$$P(X=k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Questa è una buona definizione in quanto

$$\sum_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \underbrace{\sum_{k=0}^{+\infty} \frac{\lambda^k}{k!}}_{\text{serie esponenziale}} = e^{-\lambda} \cdot e^{\lambda} = 1$$

Scriviamo $X \sim \text{Po}(\lambda)$

Teorema (Approssimazione di Poisson per una binomiale)

Sia $X \sim B(n, p)$ e sia $Y \sim \text{Po}(np)$

Allora per ogni insieme A di numeri interi non negativi

$$|P(X \in A) - P(Y \in A)| \leq np^2$$

Quindi tale approssimazione è "buona" se

$m p$ è molto piccolo (generalmente si richiede $m p \leq 0.01$)

Esempio Nella scrittura di un testo ogni carattere viene sbagliato con probabilità $\frac{1}{1000}$, indipendentemente dagli altri caratteri.

Calcolare la probabilità che in un testo di 2000 caratteri ci siano 3 o più errori.

$$X = \text{no di errori nel testo} \sim B(2000, \frac{1}{1000})$$

$$\begin{aligned} P(X \geq 3) &= 1 - P(X \leq 2) = 1 - P(X=0) - P(X=1) - P(X=2) \\ &= 1 - \left(1 - \frac{1}{1000}\right)^{2000} - 2000 \frac{1}{1000} \cdot \left(1 - \frac{1}{1000}\right)^{1999} - \\ &\quad - \frac{2000 \cdot 1999}{2} \left(\frac{1}{1000}\right)^2 \left(1 - \frac{1}{1000}\right)^{1998} \end{aligned}$$

Calcoliamo tale probab. con l'approx di Poisson

$$Y \sim \text{Po}(m p) \quad m p = 2$$

$$\begin{aligned} P(Y \geq 3) &= 1 - P(Y=0) - P(Y=1) - P(Y=2) \\ &= 1 - e^{-2} - e^{-2} \cdot 2 - e^{-2} \frac{2^2}{2!} = 1 - 5e^{-2} \end{aligned} \quad \boxed{P(Y=k) = e^{-2} \frac{2^k}{k!}}$$

La differenza fra $P(X \geq 3)$ e $P(Y \geq 3)$ è

al 'minimo' $np' = 2 \cdot 10^5 \cdot 10^{-6} = \frac{2}{1000}$

Dunque $P(X \geq 3) \approx 1 - 5e^{-2} \approx 0.002$

Sia $X \sim Po(1)$ e sia $X_n \sim B(n, \frac{1}{n})$

Sappiamo che $E(X_n) = n \cdot \frac{1}{n} = 1$

$$Var(X_n) = n \frac{1}{n} \left(1 - \frac{1}{n}\right) \xrightarrow{n \rightarrow \infty} 1$$

Questo significa che $E(X) = 1$, $Var(X) = 1$

Verifichiamo che è proprio così

$$E(X) = \sum_{k=0}^{+\infty} k P(X=k) = \sum_{k=0}^{+\infty} k e^{-1} \frac{1^k}{k!} =$$

$$= \sum_{k=1}^{+\infty} e^{-1} \frac{1^k}{(k-1)!} = \sum_{h=0}^{+\infty} e^{-1} \frac{1^{h+1}}{h!} =$$

$$= 1 \sum_{h=0}^{+\infty} e^{-1} \frac{1^h}{h!} = 1$$

Somme delle
densità

$$E(X^2) = \sum_{k=0}^{+\infty} k^2 e^{-1} \frac{1^k}{k!} = \sum_{k=1}^{+\infty} k e^{-1} \frac{1^k}{(k-1)!}$$

$$\begin{aligned}
 & \frac{\lambda^k}{k!} = \frac{\lambda^k}{k!} \cdot \frac{(k-1)!}{(k-1)!} \\
 & = \sum_{h=0}^{k-1} (h+1) e^{-\lambda} \frac{\lambda^{h+1}}{(h+1)!} = \\
 & = \underbrace{\lambda \sum_{h=0}^{+\infty} h e^{-\lambda} \frac{\lambda^h}{h!}}_{= E(X) = \lambda} + \underbrace{\lambda \sum_{h=0}^{+\infty} e^{-\lambda} \frac{\lambda^h}{h!}}_{= 1} \\
 & = \lambda^2 + \lambda \Rightarrow \text{Var}(X) = E(X^2) - (E(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda
 \end{aligned}$$

PROCESSO DI POISSON

Supponiamo di osservare le occorrenze di un fenomeno che avviene a tempi aleatori (terremoto, guasto, arrivo di un cliente ...)

Denotiamo con N_t il numero di occorrenze fra il tempo 0 e il tempo t

Il numero di occorrenze fra il tempo t e il tempo $t+h$ è $N_{t+h} - N_t$

Ipotesi 1 $P(N_{t+h} - N_t = 1) = \lambda h + o(h)$

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

Ipotesi 2 $P(N_{t+h} - N_t \geq 2) = o(h)$

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

Ipotesi 3 Siano $I_1, \dots, I_n$ intervalli di tempo disgiunti, e sia A_i l'evento $A_i =$ "nell'intervallo I_i si è verificata almeno un'occorrenza"

Allora gli eventi $A_1, A_2, \dots, A_n$ sono indipendenti.