

# Appunti di Statistica

Appunti dalle lezioni

Nicola Vanello

27 dicembre 2018



# Capitolo 1

## Variabili Aleatorie Discrete

### 1.1 Variabile aleatoria di Bernoulli

Una variabile aleatoria di Bernoulli, può assumere due valori, denominati nel seguito come successo e insuccesso, di valori rispettivamente, 1 e 0. La probabilità di ottenere un successo è pari a  $p$ , mentre la possibilità di ottenere un insuccesso è pari a  $q=1-p$ . Quindi

$$\begin{aligned}P\{X = 1\} &= p \\P\{X = 0\} &= 1 - p = q\end{aligned}$$

La variabile aleatoria di Bernoulli può essere utilizzata per descrivere un processo di Bernoulli, ovvero una serie di  $n$  prove indipendenti, ciascuna delle quali può assumere valore 0 o 1, con la stessa probabilità.

In particolare, il numero di successi in  $n$  prove può essere descritto dalla variabile aleatoria

$$X = X_1 + X_2 + \dots + X_n$$

dove la variabile  $X_i$  descrive il risultato della prova  $i$ -esima.

Si può dimostrare che tale variabile ha una distribuzione di tipo binomiale. Infatti se consideriamo la probabilità di un evento composto nel quale si hanno  $k$  successi e  $n-k$  insuccessi, tale probabilità vale  $p^k q^{n-k}$ . Visto che il numero di eventi di questo tipo è pari al numero di combinazioni di  $k$  oggetti in  $n$  posizioni (o equivalentemente al numero di modi differenti nei quali i  $k$  successi si possono presentare nelle  $n$  prove), la probabilità di avere  $k$  successi in  $n$  prove è pari alla massa di probabilità binomiale

$$p_n(k) = \binom{n}{k} p^k q^{n-k}$$

Se la probabilità di successo è molto piccola, allora l'evento "successo" è un evento raro e la variabile aleatoria è descritta da una distribuzione di Poisson. Ora introdurremo la variabile di Poisson.

### 1.1.1 Tempo del $k$ -esimo successo: variabile binomiale negativa

Consideriamo i risultati di un processo di Bernoulli e consideriamo delle variabili aleatorie che descrivano i tempi intercorrenti tra due eventi caratterizzati da un successo. Ad esempio  $T_1$  indica il numero dell'evento che è stato caratterizzato dal primo successo. Quindi  $T_1=5$  indica che il primo successo coincide con l'evento 5. Se poi indichiamo  $T_2=3$ , significa che il secondo successo si è verificato al tempo  $T_1+T_2=8$ .

Consideriamo adesso il tempo del  $k$ -esimo successo

$$\tau_k = T_1 + T_2 + \dots + T_k$$

Questa è un variabile aleatoria che può assumere i valori interi da  $k$  a  $+\infty$ . La massa di probabilità può essere definita come, la probabilità che il  $k$ -esimo successo avvenga al tempo  $k+h$  quindi

$$P\{\tau_k = k + h\}$$

Il successo  $k$ -esimo in corrispondenza del tempo  $k+h$ , avviene quando  $X_k + h=1$  e ci sono stati altri  $k-1$  successi nelle precedenti  $k+h-1$  variabili. I modi nei quali i  $k-1$  successi possono verificarsi nei  $k+h-1$  eventi sono  $\binom{k+h-1}{k-1}$ . Ciascuno di questi eventi ha una probabilità pari a  $p^{k-1}q^h$  e considerando anche l'evento successo al tempo  $k+h$  abbiamo  $p^kq^h$ . In definitiva, siamo in grado di calcolare la massa di probabilità cercata, che prenderà il nome di binomiale negativa di parametri  $k$  e  $p$  ed è pari a

$$P\{\tau_k = k + h\} = \binom{k+h-1}{k-1} p^k q^h$$

Tale variabile ha valore medio e varianza pari a

$$\tau_k = \frac{k}{p}, \quad \sigma^2 = k \frac{q}{p^2}$$

Chi fosse interessato alla dimostrazione può fare riferimento alle dispense del Prof. Flandoli.

## 1.2 Variabile aleatoria di Poisson

La distribuzione di Poisson serve per contare il numero di volte in cui un evento ha luogo in un determinato intervallo di tempo. La stessa distribuzione può essere estesa anche per contare gli eventi che hanno luogo in una determinata porzione di spazio. Tale distribuzione trova un'applicazione biomedica, ad esempio, nella descrizione delle immagini ottenute tramite la Tomografia ad emissione di positroni o PET, nella quale la formazione delle

immagini è legata alla rivelazione di eventi di decadimento radioattivo. La distribuzione di Poisson si può vedere come limite della distribuzione binomiale dove la probabilità di successo  $p$  è pari a  $\frac{\Lambda}{n}$  con  $n$  che tende a  $\infty$ .

$$\lim_{n \rightarrow \infty} \binom{n}{k} p^k q^{n-k} = e^{-\Lambda} \frac{\Lambda^k}{k!}$$

dove  $\Lambda$  è il numero medio di eventi e può essere visto come  $\Lambda = \lambda T$  dove  $T$  è l'intervallo di tempo e  $\lambda$  il numero medio di eventi nell'unità di tempo. Si fa notare che il prodotto  $np$  è pari al valore finito  $\Lambda$ .

### Esempio 1

Supponiamo che un evento accada 300 volte all'ora e noi siamo interessati alla probabilità che in un minuto accadano esattamente 2 eventi. In questo caso quindi la nostra unità di riferimento è il minuto e in un minuto il numero medio di eventi è pari a  $\Lambda = \frac{300}{60} = 5$ . Quindi la probabilità che in un minuto si abbiano 2 eventi è pari a

$$e^{-5} \frac{5^2}{2!} = 0.0842$$

La probabilità che un evento accada almeno 8 volte in un minuto è data da

$$P\{X \geq 8\} = 1 - P\{X \leq 7\} = 1 - \sum_{i=0}^7 e^{-5} \frac{5^i}{i!} = 1 - 0.8666$$

Tale risultato può essere ottenuto anche tramite tabelle che mostrano i valori di

$$\sum_{i=0}^k e^{-\Lambda} \frac{\Lambda^i}{i!}$$

al variare di  $\Lambda$  e  $k$ .

### Esempio 2

Supponiamo che un evento in media avvenga 6 volte al minuto, si calcoli la probabilità che avvenga un numero di eventi uguale o inferiore a 3 in un minuto.

In questo caso  $\Lambda = 6$ , quindi si ha che la probabilità richiesta è pari a

$$P\{X \leq 3\} = \sum_{i=0}^3 e^{-6} \frac{6^i}{i!} = 0.0025 + 0.0149 + 0.045 + 0.0892 = 0.1512$$

Si calcoli la probabilità che avvenga un numero di eventi pari a 2 in 30 secondi. In questo caso  $\lambda = \frac{6}{\text{minuto}}$  mentre  $\Lambda = 6T = 6 * 0.5 = 3$ . Quindi si ha che la probabilità richiesta è pari a

$$P\{X = 2\} = \sum_{i=0}^3 e^{-6} \frac{6^i}{i!} e^{-3} \frac{3^2}{2!} = 0.0025 + 0.0149 + 0.045 + 0.0892 = 0.2240$$

### 1.2.1 Alcune proprietà della variabile di Poisson

*Valore medio.* Il valore medio della variabile di Poisson è pari a

$$E\{X\} = \eta_X = \sum_{k=0}^{\infty} k p_k = \sum_{k=1}^{\infty} k e^{-\Lambda} \frac{\Lambda^k}{k!} = e^{-\Lambda} \Lambda \sum_{k=1}^{\infty} \frac{\Lambda^{k-1}}{k-1!}$$

considerato che

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$$

tramite un semplice cambio di variabile si può dire che

$$E\{X\} = e^{-\Lambda} \Lambda \sum_{k=1}^{\infty} \frac{\Lambda^{k-1}}{k-1!} = \Lambda$$

*Varianza.* Calcoliamo la varianza della variabile di Poisson. Per fare questo è utile calcolare la grandezza  $E\{X(X-1)\}$ . Tale grandezza risulta di semplice calcolo e permette di ottenere il valore quadratico medio come  $E\{X^2\} = E\{X(X-1)\} + E\{X\}$

$$\begin{aligned} E\{X(X-1)\} &= \sum_{k=0}^{\infty} k(k-1)p_k = \sum_{k=2}^{\infty} k(k-1)p_k = \sum_{k=2}^{\infty} k(k-1) \frac{\Lambda^k e^{-\Lambda}}{k!} = \\ &= e^{-\Lambda} \Lambda^2 \sum_{k=2}^{\infty} \frac{\Lambda^{k-2}}{(k-2)!} = \Lambda^2 \end{aligned}$$

In questo modo  $E\{X^2\} = E\{X(X-1)\} + E\{X\} = \Lambda^2 + \Lambda$

e

$$\sigma_X^2 = E\{X^2\} - E^2\{X\} = \Lambda^2 + \Lambda - \Lambda^2 = \Lambda$$

*Somma di variabili aleatorie di Poisson.* La variabile  $Z$ , somma di due variabili Poisson  $X$  e  $Y$  di valore atteso  $\Lambda_X$  e  $\Lambda_Y$ , rispettivamente, è una variabile di Poisson con valore atteso  $\Lambda_Z = \Lambda_X + \Lambda_Y$ .

La differenza di due variabili di Poisson non è una variabile di Poisson. Basti pensare che può assumere valori negativi.

*Esempio 3 (dal libro di Ciampi, Del Corso, Verrazzani)*

In campo biomedico, le immagini la tomografia ad emissioni di positroni possono essere studiate tramite un modello basato sulla distribuzione di Poisson. Il processo di decadimento radioattivo può essere schematizzato come una variabile di Poisson. Consideriamo ad esempio una sostanza radioattiva il cui numero medio di decadimenti al secondo è pari a  $\lambda = 0.5$ . Si calcoli qual è la probabilità di avere almeno una disintegrazione al minuto. In questo caso possiamo dire che il problema si può studiare utilizzando  $\Lambda = 0.5T = 0.5 * 60 = 30$ , quindi la probabilità richiesta è

$$P\{X \geq 1\} = 1 - P\{X = 0\} = 1 - e^{-30} \frac{30^0}{0!} = 1 - e^{-30} \simeq 1$$

In modo analogo si può vedere il numero di decadimenti in un minuto come la somma di 60 variabili, ognuna di Poisson con  $E\{X\} = \lambda$ . Quindi tale variabile sarà anch'essa una variabile di Poisson con valore atteso pari a  $60\lambda$ .





# Bibliografia

- [1] Luigi Landini, Nicola Vanello (2016) *Analisi e Modelli di Segnali Biomedici*, Plus Pisa University Press ed.
- [2] Luigi Landini (2005) *Fondamenti di Analisi di Segnali Biomedici* CON ESERCITAZIONI IN MATLAB, Plus Pisa University Press ed.
- [3] F. Flandoli (2001) *Elementi di Probabilità, Statistica e Processi Stocastici*
- [4] M. Ciampi, G. Del Corso, L. Verrazzani (1994) *Teoria dei Segnali. Segnali Aleatori* ETS ed.