2 Distribuzione di Poisson

La distribuzione di Poisson descrive gli eventi che entrano in un conteggio. Questi soddisfano generalmente i seguenti requisiti:

- ciascun evento è indipendente da quelli occorsi in precedenza e non influenza quelli successivi
- la distribuzione degli eventi lungo un *asse* che serva ad ordinarli, ad esempio una linea temporale o spaziale, non dipende in nessun modo da come si suddivide tale asse. Questo significa che immaginando di suddividere l'asse in intervalli uguali, la probabilità di occorrere in un dato intervallo è identica per tutti gli intervalli.
- gli eventi occorrono in modo sequenziale. In altre parole, esiste sempre un intervallo lungo l'asse di riferimento arbitrariamente piccolo che separa due eventi.

Da questi assunti possiamo dire che la probabilità di trovare un evento in un intervallo di lunghezza dl dipende dalla dimensione dell'intervallo, ovvero $p = k\,dl$, con k una costante, che interpreteremo come la frequenza media di occorrenza di un evento. Naturalmente, perché $k\,dl$ esprima il senso di una probabilità deve essere $k\,dl \leq 1$. Questo è garantito dall'arbitrarietà di dl, che possiamo considerare piccolo a piacere (e comunque sufficientemente piccolo perché sia trascurabile la probabilità di trovarci dentro due eventi).

Se $k\,dl$ è la probabilità di trovare un evento in dl, allora la probabilità di non trovarne nessuno è il suo complemento ad 1, ovvero $1 - k\,dl$.

Se stiamo considerando un intervallo L suddiviso da N intervalli dl, allora la probabilità che nessuno di questi contenga un evento vale

$$P_0(L) = (1 - k\,dl)^N.$$

D'altronde $dl = \frac{L}{N}$, quindi la precedente equazione diventa

$$P_0(L) = (1 - k\frac{L}{N})^N.$$

Immaginando di mandare $N \to \infty$, ovvero di riducendo sempre di più dl, possiamo scrivere

$$P_0(L) = \lim_{N\to\infty} (1 - k\frac{L}{N})^N = e^{-kL}.$$

A questo punto definiamo $I_1(l)$ come la densità di probabilità (p.d.f.) di intervalli con 1 solo evento. L'espressione seguente

$$I_1(l)\,dl \tag{2.1}$$

corrisponde alla probabilità di avere un evento nell'intervallo tra l e $l + dl$. In altre parole, la (2.1) indica la probabilità di trovare un impulso in un intervallo infinitesimale dl, dopo che non si è avuto un solo evento per tutto un intervallo di lunghezza l dall'ultimo evento. Insomma: stiamo parlando della sequenza di due eventi nell'intervallo compreso da 0 a $l+dl$.

Quindi

$$\begin{aligned} I_1(l)\,dl &= P_0(l)\times P_1(dl) = e^{-kl}\,k\,dl; \\ I_1(l) &= k\,e^{-kl}. \end{aligned}$$

A questo punto viene naturale chiederci prima quale sia la probabilità di avere una sequenza di 3 eventi, e poi generalizzare ad un numero arbitrario N di eventi (sempre inclusi in un intervallo compreso da 0 a $l+dl$).

Affrontiamo prima il caso di 3 eventi. Relativamente al caso della sequenza di due eventi, si tratta di inserirne uno tra 0 e $l+dl$. In pratica lungo l'asse di riferimento, l'evento da aggiungere si presenta entro un certo dl' dopo che non ci sono stati eventi tra l'inizio e $l'<l$. Pertanto la probabilità di tale sequenza a tre eventi (che occorrono rispettivamente in 0, $l'+dl'$ e $l+dl$) vale:

$$\begin{aligned} I_2(l)\,dl &= P_0(l')\times k\,dl'\times P_0(l-l')\times k\,dl \\ &= dl\int_0^l I_1(l')\times I_1(l-l')\,dl' \\ &= dl\int_0^l k\,e^{-kl'}\times k\,e^{-k(l-l')}\,dl' \\ &= k^2\,dl\,e^{-kl}\,l \\ &= k(kl)e^{-kl}\,dl. \end{aligned}$$

Pertanto

$$\begin{aligned} I_2(l) &= k(kl)e^{-kl} \\ &= (kl)\,I_1(l). \end{aligned} \tag{2.2}$$

Dalla (2.2) appare subito la relazione recursiva. Un Aumentando di un evento si dimostra che

$$\begin{aligned} I_3(l) &= \frac{(kl)^2\,I_1(l)}{2} \\ I_4(l) &= \frac{(kl)^3\,I_1(l)}{6} \\ I_N(l) &= \frac{(kl)^{N-1}\,I_1(l)}{(N-1)!} \end{aligned} \tag{2.3}$$

Possiamo quindi utilizzare la (2.3) per scrivere la probabilità di avere N eventi nell'intervallo $(0,L+dl)$ in modo che:

$$\begin{aligned} P_N(L+dl) &= I_N(L)\,dl \\ &= \frac{(kL)^{N-1}\,e^{-kL}}{(N-1)!}k\,dl \qquad (2.4) \\ &= P_{N-1}(L)P_1(dl) \qquad (2.5) \end{aligned}$$

Confrontando la (2.4) con la (2.5) si conclude che la probabilità di avere $M = N - 1$ eventi nell'intervallo di larghezza L vale

$$P_M(L) = \frac{(kL)^M e^{-kL}}{M!}. \tag{2.6}$$

La (2.6) è la formulazione matematica della distribuzione di Poisson.

2.1 Valore di aspettazione e varianza per la distribuzione di Poisson

Utilizzando la relazione (1.2), identificando $i = x_i = M$, e con $w_i = P_M(L)$si trova che il valore di aspettazione di una variabile che segue la distribuzione di Poisson vale proprio $\mu = kL$.

Infatti si ha che

$$\begin{aligned}
\mu &= E[M] \\
&= \sum_{M=0}^{\infty} M \frac{(kL)^M e^{-kL}}{M!} \\
&= e^{-kL} \sum_{M=0}^{\infty} M \frac{(kL)^M}{M!}
\end{aligned}$$

Poiché per $M = 0$ il relativo termine della somma è nullo, possiamo scrivere

$$\begin{aligned}
\mu &= e^{-kL} \sum_{M=1}^{\infty} M \frac{(kL)^M}{M!} \\
&= e^{-kL} \sum_{M=1}^{\infty} \frac{(kL)^M}{(M-1)!} \\
&= e^{-kL} kL \sum_{M=1}^{\infty} \frac{(kL)^{M-1}}{(M-1)!}
\end{aligned}$$

Se definisco $J = M - 1$

$$\mu = e^{-kL} kL \sum_{J=0}^{\infty} \frac{(kL)^J}{(J)!}$$

ma valendo $\sum_{J=0}^{\infty} \frac{(kL)^J}{(J)!} = e^{kL}$, si trova infine

$$\begin{aligned}
\mu &= e^{-kL} kL e^{kL} \\
&= kL
\end{aligned} \tag{2.7}$$

In sostanza μ è il valore di aspettazione del numero di eventi che possono occorrere in un intervallo di grandezza L. Per questo motivo, una formulazione equivalente alla (2.6) è la seguente:

$$P(r,\mu) = \frac{\mu^r \, e^{-\mu}}{r!}. \tag{2.8}$$

ovvero la probabilità di trovare r eventi quando me ne aspetto μ.

Per la varianza si applica la relazione (1.4):

$$\begin{aligned}
V[r] &= E[r^2] - \mu^2 \\
E[r^2] &= \sum_{r=0}^{\infty} r^2 \frac{e^{-\mu}\mu^r}{r!} \\
&= e^{-\mu} \sum_{r=1}^{\infty} r \frac{\mu^r}{(r-1)!} \\
&= e^{-\mu}\mu \sum_{r=1}^{\infty} r \frac{\mu^{(r-1)}}{(r-1)!} \\
&= e^{-\mu}\mu \sum_{i=0}^{\infty} (i+1) \frac{\mu^i}{i!} \\
&= e^{-\mu}\mu \left(\sum_{i=0}^{\infty} i \frac{\mu^i}{i!} + \sum_{i=0}^{\infty} \frac{\mu^i}{i!} \right) \\
&= e^{-\mu}\mu \left(\mu \sum_{i=1}^{\infty} \frac{\mu^{(i-1)}}{(i-1)!} + e^{\mu} \right) \\
&= e^{-\mu}\mu \left(\mu \sum_{j=0}^{\infty} \frac{\mu^j}{j!} + e^{\mu} \right) \\
&= e^{-\mu}\mu \left(\mu e^{\mu} + e^{\mu} \right) \\
&= \mu (\mu + 1)
\end{aligned}$$

Quindi si trova che

$$\begin{aligned}
V[r] &= \mu^2 + \mu - \mu^2 \\
&= \mu.
\end{aligned} \tag{2.9}$$

Questo implica che nel caso della distribuzione di Poisson, la varianza della variabile aleatoria coincide con il valore di aspettazione.

2.2 Ancora il caso delle entries di un istogramma

Finalmente possiamo ritornare ad occuparci delle entries di un bin di un istogramma. Per quanto affermato all'inizio della Sezione, la (1.13) può essere riscritta tenendo conto della (2.9), e quindi:

$$V[z] = \sum_j w_j^2 \mu_j \tag{2.10}$$

Ora, in generale, quando si costruisce un istogramma, il valore delle μ_i non è conosciuto. Possiamo però stimarlo nel modo più verosimile, ovvero trovare il suo miglior *estimatore* $\hat{\mu}$. Scopriremo come fare più avanti, quando parleremo del metodo della *Massima Verosimiglianza*, (*Maximum Likelihood* - ML, in inglese). Qui basta anticipare che data una misura di conteggi r, quindi essendo r una variabile aleatoria distribuita secondo una distribuzione di Poisson, la miglior stima del valore di aspettazione $\hat{\mu} = r$, cioè è proprio la misura effettuata. Di conseguenza possiamo riscrivere la (2.10) nel seguente modo:

$$\begin{aligned} V[z] &= \sum_j w_j^2 \hat{\mu}_i \\ &= \sum_j w_j^2 x_j \end{aligned} \tag{2.11}$$

con x_j i conteggi che hanno il peso w_j (l'avevamo già detto in precedenza).

Adesso se ogni entry ha un suo peso individuale, $x_j = 1$, possiamo scrivere finalmente che

$$V[z] = \sum_j w_j^2. \tag{2.12}$$

E' interessante notare che se ogni evento ha peso unitario ($w_j \equiv 1$), allora

$$V[z] = n_i.$$

Questo risultato è perfettamente in linea con il fatto che n_i è la misura di un conteggio, ovvero il valore misurato di una variabile aleatoria distribuita secondo una Poissoniana, e quindi avente la miglior stima della varianza coincidente con l'estimatore del suo valore di aspettazione.