

Probabilità e statistica

November 14, 2010

Prof.ssa Paola Giacconi

1 Caso e probabilità

Per *probabilità* di un particolare risultato A nell'osservazione o misura di una grandezza, si intende **la nostra stima della frazione di osservazioni ripetute che darà quel particolare risultato**. Se immaginiamo di ripetere un'osservazione n volte e indichiamo con n_A la nostra **stima** della frequenza del risultato A , allora

$$p(A) \equiv \frac{n_A}{n} \quad \text{con} \quad 0 \leq p(A) \leq 1 \quad \text{e} \quad \sum_A p(A) = 1, \quad (1.1)$$

indica la probabilità $p(A)$ di ottenere un certo risultato A su n osservazioni totali. Questa definizione di probabilità poggia sull'assunzione, non banale, che *i nostri esperimenti e quindi le nostre osservazioni e misure siano ripetibili*. Noi supporremo che ciò sia vero fino a quando le nostre conoscenze sul fenomeno non divengano tali da imporre una modifica all'esperimento stesso e quindi alla nostra stima di n_A . Sottolineamo il fatto che i numeri n ed n_A non si riferiscono ad osservazioni o misure reali: n_A è la miglior *stima* di ciò che noi pensiamo possa accadere su n *esperimenti immaginati*. In questo senso la probabilità dipende dalla nostra conoscenza e/o dalla nostra ignoranza. Fu Bruno de Finetti, uno dei più grandi matematici del 900, ad approfondire questo aspetto ¹ che, tuttavia, in queste note non approfondiremo.

La definizione (1.1) rappresenta il rapporto tra quelli che noi stimiamo essere i casi favorevoli relativi al risultato in questione rispetto a quelli che noi supponiamo essere i casi possibili. Per esempio, nel lancio di una moneta

¹Bruno de Finetti: "Teoria della Probabilità" (Einaudi 1970)

non truccata la probabilità di avere testa $p(T)$ è uguale a quella di avere croce $p(C)$ in altre parole $p(T) = p(C)$ ed anche $p(T) + p(C) = 1$ da cui $p(T) = p(C) = 1/2$. Analogamente, nel lancio di un dado non truccato la probabilità attribuita ad una qualunque delle sei facce si ottiene, per quanto sopra esposto, osservando che ad ogni faccia va attribuita la stessa probabilità $p(F)$ e che la somma di tutti i risultati possibili deve dare uno. Quindi $6p(F) = 1$ da cui $p(F) = 1/6$.

Le probabilità associate all'uscita di testa nel lancio della moneta e/o all'uscita di un certo numero nel lancio del dado **non sono determinate sulla base di esperimenti realmente effettuati** ma rappresentano la nostra **stima** del valore più probabile determinata, a priori, a partire dalle nostre conoscenze del fenomeno in questione. In altre parole, se le nostre conoscenze mutassero anche la nostra stima cambierebbe e così anche la probabilità associata all'evento.

2 Distribuzione binomiale e variabili di Bernoulli

Cominceremo il nostro studio sulla probabilità, che applicheremo poi a delle situazioni sperimentali, dalla distribuzione discreta binomiale in quanto essa è abbastanza semplice e significativa. Inoltre, la distribuzione binomiale si generalizza facilmente alla distribuzione di Poisson e, come qualsiasi altra distribuzione di probabilità, alla distribuzione continua di Gauss che rappresenta il fondamento di tutta la statistica.

Supponiamo quindi di fare un **esperimento reale** con soli due risultati possibili. Esempi comuni di variabili binomiali possono essere:

1. passare/fallire un esame
2. vincere/perdere al gioco
3. osservare testa/croce lanciando una moneta
4. inserire una persona tra due categorie [fumatori — non fumatori]
5. prendere la sufficienza o l'insufficienza nel compito di matematica

L'evento è allora descritto da una variabile che prende il nome di *variabile dicotomica di Bernoulli*, che può assumere solo due valori: p e $\bar{p} = 1 - p$ dove p indica la probabilità di *successo* ed $1 - p$ la probabilità di *insuccesso*. Possiamo pensare di avere due contenitori da riempire, il contenitore p ed il contenitore $1 - p$. Il primo contenitore contiene il numero di successi il secondo il numero di insuccessi. La probabilità che su n eventi vi siano k

successi, cioè la probabilità che su n eventi k eventi siano nella scatola p , è data, come noto, da:

$$P_k(n, p) = C_{n,k} \cdot p^k \cdot (1 - p)^{n-k} \quad (2.1)$$

dove $C_{n,k}$ rappresenta la combinazione semplice di n oggetti presi k a k .

E' evidente che risulta:

$$\sum_{k=0}^n P_k(n, p) = \sum_{k=0}^n C_{n,k} \cdot p^k \cdot (1 - p)^{n-k} = (p + \bar{p})^n = 1 \quad (2.2)$$

in accordo con l'ultima dell' eq. (1.1). Infatti, come noto, ogni distribuzione di probabilità deve soddisfare l'ultima relazione della (1.1) che nel caso della distribuzione binomiale si riduce alla (2.2) e cioè alla formula di Leibnitz-Newton della potenza di un binomio la quale afferma, appunto, che la somma di tutti i casi possibili rappresenta la certezza che corrisponde alla probabilità uguale ad uno.

La (2.1) è la **distribuzione di probabilità binomiale**.

In generale, si definisce *legge di distribuzione di probabilità* relativa ad una variabile aleatoria ogni relazione che stabilisce una corrispondenza tra i valori possibili di tale variabile e la loro probabilità di verificarsi. Il fatto che una variabile aleatoria si distribuisca secondo una data distribuzione di probabilità ci permette di trarre alcune conclusioni importanti sulla variabile stessa e quindi sul fenomeno che si sta studiando. Tra queste, la possibilità di definire quello che viene comunemente chiamato *livello di confidenza*, che altro non è se non la *la probabilità che l'affermazione a cui esso si riferisce sia vera*. In altre parole, conoscere la distribuzione di probabilità di una determinata variabile aleatoria associata ad un evento, ci permette di determinare la media con la quale l'evento in questione si dovrebbe verificare se facessimo un numero arbitrario n di prove, con la relativa incertezza. Vale a dire che conoscere la distribuzione di probabilità di una variabile aleatoria, ci permette di costruire un "ponte" tra la teoria e l'esperimento dandoci la possibilità di fare delle previsioni. Ad esempio se lanciando un dado 10 volte uscisse per per 10 volte il numero 6(sei), la matematica, servendosi della distribuzione di Benoulli, ci direbbe che molto probabilmente c'è qualcosa, nel dado che stiamo lanciando, che non va!

Infatti conoscendo la distribuzione di probabilità di una variabile aleatoria x possiamo rispondere alla seguente domanda: *ripetendo un esperimento n volte quale è la media con cui si presenta l'evento di probabilità p* ? In altre parole su dieci lanci del dado quante volte in media si presenterà il numero 6(sei)? Ricordando che in generale la media di una variabile aleatoria discreta

x_j alla quale è associata una probabilità P_j è data da:

$$\bar{x} = \langle x \rangle = \sum_{j=1}^n x_j P_j, \quad (2.3)$$

per rispondere alla nostra domanda dobbiamo eseguire la seguente somma che, nel caso di distribuzione di Bernoulli, dà come risultato:

$$\bar{k} = \langle k \rangle = \sum_{k=1}^n k P_k(n, p) = \sum_{k=1}^n k C_{n,k} \cdot p^k \cdot (1-p)^{n-k} = np. \quad (2.4)$$

Quindi, tornando al nostro esempio del lancio del dado, l'eq. (2.4) ci permette di prevedere che su 10 lanci il numero 6(sei) in media dovrebbe uscire circa 2 volte infatti

$$\bar{k} = np = 10 \frac{1}{6} = \frac{5}{3} \sim 2. \quad (2.5)$$

Questo significa che la distribuzione di Bernoulli ci indica che dobbiamo essere "sospettosi" anche nel caso in cui il numero 6(sei) esca una, tre, quattro, cinque,... volte su dieci? In altre parole quanto possiamo fidarci del risultato teorico ottenuto con la relazione eq. (2.5)? Come abbiamo già detto conoscere la distribuzione di probabilità associata ad un evento ci permette anche di determinare quale è l'incertezza associata alla nostra stima ottenuta, ad esempio, con l'equazione (2.5). *Ricordiamo*, a questo proposito, che la **fluttuazione quadratica media o scarto quadratico medio** di una variabile casuale x è definita come:

$$\begin{aligned} \sigma_x^2 &= \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 = \frac{1}{N} \left[\sum_{i=1}^N (x_i)^2 - 2\bar{x} \sum_{i=1}^N (x_i) + N\bar{x}^2 \right] \\ &= \frac{1}{N} \sum_{i=1}^N (x_i)^2 - \frac{2\bar{x}}{N} \sum_{i=1}^N x_i + \bar{x}^2 = \langle x^2 \rangle - \langle x \rangle^2, \end{aligned} \quad (2.6)$$

mentre la **varianza o deviazione standard**, è data da

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}, \quad (2.7)$$

σ_x rappresenta quindi una stima delle fluttuazioni casuali di cui è affetta ogni singola misura x_i . Osserviamo inoltre che la media di tutte le misure contiene più informazioni di ogni singola misura, possiamo in altre parole affermare che essa è un parametro più stabile e costituisce quindi una stima migliore

della grandezza fisica che si stà misurando. Per questi motivi possiamo comprendere come l'incertezza ad essa associata sia minore di quella associata ad ogni singola misura, risulta che la deviazione standard della media sia \sqrt{N} più piccola, dove N rappresenta il numero di misure effettuate. In altre parole la miglior stima per l'incertezza della media è

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}}, \quad (2.8)$$

quantità che tende rapidamente a zero per $N \rightarrow \infty$. $\sigma_{\bar{x}}$ prende il nome di **deviazione standard della media**.

Tornando alla distribuzione di Bernoulli si ha:

$$\begin{aligned} \sigma_k^2 &= \langle k^2 \rangle - \langle k \rangle^2 = \sum_{k=0}^n k^2 P_k(n, p) - (np)^2 \\ &= np[p(n-1) + 1] - (np)^2 = np(1-p). \end{aligned} \quad (2.9)$$

Questo significa che, per una variabile che segue la distribuzione di Bernoulli, su n eventi il numero medio di successi è dato da:

$$\langle k \rangle \pm \sigma = np \pm \sqrt{np(1-p)}. \quad (2.10)$$

Quindi, nel nostro esempio precedente, la nostra stima teorica può sbagliare di ± 1 , vale a dire che se il numero 6(sei) uscisse una sola volta o tre volte potremmo continuare a giocare senza il timore che ci stiano "fregando" !

Vediamo ancora due esempi variando il numero n di prove.

1. **Lancio del dado:** se lanciamo un dado $n = 60$ volte, il numero medio di volte in cui può uscire il numero 4 è dato da:

$$\langle k \rangle \pm \sigma = np \pm \sqrt{np(1-p)} = 60 \cdot \frac{1}{6} \pm \sqrt{60 \cdot \frac{1}{6} \cdot \frac{5}{6}} = 10 \pm \sqrt{\frac{25}{3}},$$

mentre il numero medio di volte in cui **non** uscirà il numero 4 sarà dato da:

$$\langle k \rangle \pm \sigma = np \pm \sqrt{np(1-p)} = 60 \cdot \frac{5}{6} \pm \sqrt{60 \cdot \frac{1}{6} \cdot \frac{5}{6}} = 50 \pm \sqrt{\frac{25}{3}}.$$

2. **Lancio della moneta** Se lanciamo in aria una moneta non truccata $n = 60$ volte, il numero medio di volte in cui può uscire testa è

$$\begin{aligned} \langle k \rangle \pm \sigma &= np \pm \sqrt{np(1-p)} = \frac{n}{2} \pm \frac{\sqrt{n}}{2} \\ &= 60 \cdot \frac{1}{2} \pm \sqrt{60 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 30 \pm \sqrt{15} \end{aligned} \quad (2.11)$$

Supponiamo ora di tracciare l'istogramma dell'esperimento lancio del dado. L'istogramma rappresenta il grafico della distribuzione di probabilità relativa all'evento che stiamo studiando e ci fornisce tutte le informazioni sulla variabile in questione. L'area racchiusa dall'istogramma e dall'asse delle ascisse è sempre uguale ad 1 in quanto essa rappresenta tutti i casi possibili e cioè la certezza.

Facciamo quindi l'esperimento del lancio del dado che consiste, ad esempio, nel ripetere per 100 volte 60 lanci e valutare quale è la media con cui esce un determinato numero, per fissare le idee il numero quattro. Tracciamo quindi l'istogramma per valutare se il nostro dado è o meno truccato. Quale tipo di istogramma ci aspettiamo?

Poniamo in ascissa tante caselle quanti sono i lanci di ogni esperimento, nel nostro caso 60, ed annotiamo quante volte in ogni esperimento di 60 lanci è uscito il numero quattro, sarà un numero compreso tra 0 e 60. Ripetiamo tutto ciò per 100 volte facciamo cioè 100 esperimenti così da costruire su ogni casella una colonna la cui altezza è data dal numero di esperimenti che hanno dato come esito il corrispondente numero di lanci. Una attenta riflessione fa capire che non a tutte 60 caselle corrisponderà una colonna di altezza non nulla, inoltre sommando il contenuto di tutte le colonne si otterrà 100. Per ottenere un **istogramma normalizzato** dobbiamo dividere l'altezza di ogni colonna per il numero di esperimenti effettuati, nell'esempio 100, riportare cioè in ordinata la frequenza dei successi. Se il nostro dado non è truccato dovremmo ottenere una distribuzione con il massimo sul valore 10. Se inoltre avessimo diviso i valori delle ascisse per il numero di lanci il nostro istogramma avrebbe avuto un massimo sul valore $1/6$ e l'istogramma sarebbe stato tanto più piccato attorno a questo valore quanto più grande è il numero di lanci di ogni esperimento. Quindi nel nostro caso abbiamo ottenuto un istogramma non simmetrico in quanto la probabilità di ottenere il numero quattro è proprio $p = \frac{1}{6}$ mentre quella di non ottenere il numero quattro è $(1 - p) = \frac{5}{6}$. La **frequenza**, che come già detto, è il rapporto tra la molteplicità ed il numero totale di prove, cioè il numero di volte in cui l'esperimento si ripete, nell'esempio 100, rappresenta una estensione della definizione di probabilità alle situazioni sperimentali. Tale definizione prende il nome di definizione frequentista di probabilità e tende a quella assiomatica all'aumentare del numero di prove, nel limite di infinite prove la definizione frequentista si riduce a quella assiomatica.

Ripetendo lo stesso esperimento con una moneta avremmo ottenuto un istogramma simmetrico centrato su 30 in quanto, se la moneta non è truccata, la probabilità di ottenere testa è uguale a quella di ottenere croce $p = (1 - p) = \frac{1}{2}$. Se la moneta fosse truccata avremmo $p \neq (1 - p)$ e quindi di nuovo un istogramma non simmetrico.

Supponiamo ora di voler risolvere il seguente problema: data una moneta vogliamo capire se abbiamo a che fare con una moneta "onesta" o "truccata". Possiamo lanciarla in aria tante volte, n volte, e contare quante volte esce testa, tanto maggiore è il numero di lanci tanto più attendibile sarà il risultato. Se indichiamo con n_T il numero di volte in cui esce testa, da una moneta "onesta" ci aspettiamo che il rapporto n_T/n , si avvicini sempre più ad $1/2$ all'aumentare di n , in altre parole ci si aspetta che $\lim_{n \rightarrow \infty} (n_T/n) = 1/2$. Infatti la **fluttuazione relativa** definita dalla relazione:

$$\frac{\sigma_x}{\langle x \rangle} = \frac{\sqrt{\sigma_x^2}}{\langle x \rangle} = \frac{\sqrt{\langle x^2 \rangle - \langle x \rangle^2}}{\langle x \rangle}, \quad (2.12)$$

che nel caso di variabile discreta, ricordando le eq.ni (2.4) e (2.9), assume la forma

$$\frac{\sigma_k}{\langle k \rangle} = \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}}, \quad (2.13)$$

per grandi n tende a zero come $1/\sqrt{n}$, **una variabile aleatoria che manifesta tale comportamento si dice normale**.

In generale per una variabile aleatoria A che si verifica in media $\langle n_A \rangle$ volte su n prove l'intervallo:

$$I_A \equiv \left[\frac{\langle n_A \rangle}{n} - \frac{\sigma}{\langle n_A \rangle}, \frac{\langle n_A \rangle}{n} + \frac{\sigma}{\langle n_A \rangle} \right] \quad (2.14)$$

$$= \left[\frac{\langle n_A \rangle}{n} - \frac{1}{\sqrt{n}}, \frac{\langle n_A \rangle}{n} + \frac{1}{\sqrt{n}} \right] \quad (2.15)$$

prende il nome di intervallo di confidenza della variabile aleatoria A .

È allora importante sottolineare il fatto che l'esempio da noi considerato della moneta è particolarmente istruttivo in quanto, come vedremo, per $n \rightarrow \infty$ ed $np \gg 1$, l'intervallo di confidenza, **per qualsiasi variabile aleatoria**, ha l'espressione data dalla (2.15).

3 Distribuzione di Poisson

La distribuzione di Poisson è il limite di una successione di distribuzioni binomiali. Essa diviene molto utile quando **non si conosce a priori il valore di n** , ma si sa solamente che è altissimo e che la probabilità p di ottenere un certo risultato in una singola prova, anche essa ignota a priori, è piccolissima in modo tale che $np \sim \text{costante}$. La sola quantità che

può essere misurata è il numero medio in cui si verifica l'evento. Poichè non si conoscono a priori nè n nè p , l'unico parametro della distribuzione è il numero di prove in cui ci si aspetta di ottenere quel certo risultato, in altre parole il numero medio di "successi" in quanto, come appena detto, il numero medio è la sola quantità misurabile.

Una tipica situazione che gode delle caratteristiche appena elencate è lo studio di un processo di decadimento radioattivo o di **conteggio di raggi cosmici**. Nel primo caso il numero di prove è costituito dal numero di nuclei che potenzialmente possono decadere (per una mole di materiale radioattivo il numero di nuclei è dell'ordine di 10^{23}), mentre la probabilità di "successo" (decadimento) per ogni nucleo è decisamente piccola.

Nel secondo caso di conteggio di **raggi cosmici** il numero di interazioni elettromagnetiche dovute alla presenza di particelle cariche nel rivelatore ($n \gg 1$) è molto elevato e fra tutte queste le interazioni dovute alle particelle cariche prodotte dagli sciami di raggi cosmici, nel nostro caso muoni μ^\pm , sono relativamente poche ($p \ll 1$). Quindi ancora $n \rightarrow \infty$ e $p \rightarrow 0$ con $np \sim \text{cost}$.

Per una variabile che segue la distribuzione di Poisson si suppone inoltre che la probabilità p sia costante e che tale probabilità di successo in un intervallo di tempo $[t, \Delta t + t]$ sia proporzionale, in prima approssimazione, a Δt . Il numero di decadimenti radioattivi o il numero di interazioni rivelate dal rivelatore dovute ai raggi cosmici soddisfano tali proprietà in quanto il numero di decadimenti e/o di interazioni dovute ai raggi cosmici rivelate cresce, in prima approssimazione, linearmente con il tempo.

La distribuzione di Poisson fu introdotta dal matematico, fisico, astronomo e statistico francese Simon-Denis Poisson nato a Pithiviers il 21 giugno 1781 e morto a Parigi il 25 aprile 1840, il quale formulò il seguente teorema:

Teorema

Data la distribuzione binomiale

$$P_k(n, p) = C_{n,k} \cdot p^k \cdot (1 - p)^{n-k} \quad (3.1)$$

e la funzione

$$r_k(\lambda) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (3.2)$$

nel limite $n \rightarrow \infty$ e $p \rightarrow 0$ in modo tale che

$$\lim_{n \rightarrow \infty} np = \lambda \quad (3.3)$$

si ha che $\forall k \in \mathbb{N}_0$

$$\lim_{n \rightarrow \infty} P_k(n, p) = r_k(\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}. \quad (3.4)$$

La distribuzione di probabilità (3.2) prende il nome di distribuzione di probabilità di Poisson, dalla definizione di esponenziale è immediato dimostrare che vale la relazione

$$\sum_{k=0}^{\infty} r_k(\lambda) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = 1 \quad (3.5)$$

in accordo con l'ultima di (1.1).

La distribuzione di Poisson è quindi il limite di una successione di distribuzioni binomiali. La (3.2) rappresenta quindi la probabilità di ottenere un determinato risultato un numero k volte durante il nostro esperimento. Ricordiamo che k è una quantità piccola ma misurabile mentre λ rappresenta, secondo la (3.3) il valore medio del risultato atteso.

Dimostrazione

La dimostrazione è per induzione rispetto a k , poniamo

$$p_n = \lambda_n/n \quad \text{con} \quad \lim_{n \rightarrow \infty} \lambda_n = \lambda \quad (3.6)$$

1. Verifichiamo che la (3.4) è vera per $k = 0$

$$\begin{aligned} \lim_{n \rightarrow \infty} P_0 \left(n, \frac{\lambda_n}{n} \right) &= \lim_{n \rightarrow \infty} C_{n,0} \cdot \left(\frac{\lambda_n}{n} \right)^0 \cdot \left(1 - \frac{\lambda_n}{n} \right)^{n-0} \\ &= \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda_n}{n} \right)^n = e^{-\lambda} = r_0(\lambda). \end{aligned} \quad (3.7)$$

2. Assumiamo che sia vera per k

$$\begin{aligned} \lim_{n \rightarrow \infty} P_k \left(n, \frac{\lambda_n}{n} \right) &= \lim_{n \rightarrow \infty} C_{n,k} \cdot \left(\frac{\lambda_n}{n} \right)^k \cdot \left(1 - \frac{\lambda_n}{n} \right)^{n-k} \\ &= \frac{\lambda^k}{k!} e^{-\lambda} = r_k(\lambda). \end{aligned} \quad (3.8)$$

3. Dimostriamo che vale per $k + 1$, a tale scopo consideriamo

$$\begin{aligned} \lim_{n \rightarrow \infty} P_{k+1} \left(n, \frac{\lambda_n}{n} \right) &= \lim_{n \rightarrow \infty} \frac{P_{k+1} \left(n, \frac{\lambda_n}{n} \right)}{P_k \left(n, \frac{\lambda_n}{n} \right)} P_k \left(n, \frac{\lambda_n}{n} \right) \\ &= \lim_{n \rightarrow \infty} \frac{\lambda_n}{k+1} \cdot \frac{n-k}{n-\lambda_n} P_k \left(n, \frac{\lambda_n}{n} \right) \\ &= \frac{\lambda}{k+1} r_k(\lambda) = r_{k+1}(\lambda) \quad \text{c.v.d} \end{aligned} \quad (3.9)$$

In sostanza il teorema afferma che per *eventi rari* per i quali p diventa sempre più piccolo al crescere di n , per $n \gg \lambda$, la distribuzione binomiale può essere bene approssimata dalla distribuzione di Poisson. Come già osservato questo fatto è di grande utilità quando non si conoscono a priori nè n nè p .

Determiniamo ora il **valore medio** dei *successi* e lo **scarto quadratico medio** di una variabile che segue la distribuzione di Poisson. Come risulta chiaro dalla (3.3) λ rappresenta la media dei *successi*, infatti

$$\bar{k} = \langle k \rangle = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} = \sum_{k=0}^{\infty} \lambda \frac{\lambda^k}{k!} e^{-\lambda} = \lambda, \quad (3.10)$$

mentre lo scarto quadratico medio, cioè l'errore da associare alla media, sarà

$$\sigma_k^2 = \langle k^2 \rangle - \langle k \rangle^2 = \sum_{k=0}^n k^2 r_k(\lambda) - \lambda^2 = \lambda.$$

Quindi il valore medio di una grandezza che segue la distribuzione di probabilità di Poisson è dato :

$$\langle k \rangle \pm \sigma = \lambda \pm \sqrt{\lambda}. \quad (3.11)$$

Inoltre, se $\lambda \gg 1$ ad esempio $\lambda \sim 10 - 20$ la distribuzione di Poisson così come la distribuzione binomiale, nel limite continuo, può essere approssimata da una distribuzione gaussiana.

È importante osservare come nella distribuzione di Poisson la media e la fluttuazione quadratica media coincidono. Questo ci permette di decidere se una variabile aleatoria può essere o meno considerata una variabile aleatoria di Poisson. Infatti a partire dai dati sperimentali si possono determinare λ e σ^2 e se questi differiscono di poco allora possiamo supporre che i dati siano distribuiti secondo una *puassoniana*.

Esempi

1. *La roulette* La vincita secca alla roulette è un evento raro in quanto la probabilità di vincere è $p = 1/37 \sim 0,027$ se si considerano tutti i numeri, compreso lo zero, equiprobabili. Quindi giocando $n = 37$ volte in media si vince $\lambda = np = 1$ volta. Valutiamo allora quale è la probabilità di vincere 0, 1, 2 volte e valutiamo tale probabilità usando la distribuzione binomiale e la distribuzione di Poisson rispettivamente. Distribuzione Binomiale

$$P_0 \left(37, \frac{1}{37} \right) = C_{37,0} \cdot \left(\frac{1}{37} \right)^0 \cdot \left(1 - \frac{1}{37} \right)^{37} \sim 0,363$$

$$\begin{aligned}
P_1\left(37, \frac{1}{36}\right) &= C_{37,1} \cdot \left(\frac{1}{37}\right)^1 \cdot \left(1 - \frac{1}{37}\right)^{37} \sim 0,373 \\
P_2\left(37, \frac{1}{37}\right) &= C_{37,2} \cdot \left(\frac{1}{37}\right)^2 \cdot \left(1 - \frac{1}{37}\right)^{35} \sim 0,186
\end{aligned} \tag{3.12}$$

Distribuzione di Poisson

$$r_0(1) = \frac{1^0}{0!} e^{-1} \sim 0,368 \tag{3.13}$$

$$r_1(1) = \frac{1^1}{1!} e^{-1} \sim 0,368 \tag{3.14}$$

$$r_2(1) = \frac{(1)^2}{2!} e^{-1} \sim 0,184 \tag{3.15}$$

2. *Guasti* Supponiamo che i microscopi forniti da una certa ditta abbiano avuto in media 6 rotture in dieci anni con una deviazione standard di 2,5.

Si può supporre che la distribuzione sia di tipo puassoniano in quanto $\lambda = 6 \simeq 2,5^2 = 6,25$. Possiamo allora determinare, ad esempio

$$r_0(6) = e^{-6} \sim 0,0025 \tag{3.16}$$

$$r_1(6) = \frac{6}{1!} e^{-6} \sim 0,015 \tag{3.17}$$

$$r_2(6) = \frac{6^2}{2!} e^{-6} \sim 0,045 \tag{3.18}$$

$$r_3(6) = \frac{6^3}{3!} e^{-6} \sim 0,089 \tag{3.19}$$

$$r_4(6) = \frac{6^4}{4!} e^{-6} \sim 0,133 \tag{3.20}$$

$$r_5(6) = \frac{6^5}{5!} e^{-6} \sim 0,160 \tag{3.21}$$

$$r_6(6) = \frac{6^6}{6!} e^{-6} \sim 0,161 \tag{3.22}$$

$$r_7(6) = \frac{6^7}{7!} e^{-6} \sim 0,14 \tag{3.23}$$

La probabilità che vi siano più di sette guasti è:

$$P(k > 7) = 1 - \sum_{k=0}^7 r_k(6) = 1 - 0,743 = 0,257 \tag{3.24}$$

I guasti di una rete telefonica seguono la distribuzione di Poisson, nella speranza che questi non avvengano di frequente!

4 Distribuzione Gaussiana

La distribuzione di Gauss riveste un ruolo particolarmente importante in quanto **qualsiasi distribuzione di probabilità, sia essa binomiale o di Poisson, per numero di prove sufficientemente grande e nel limite del continuo segue la distribuzione di probabilità di Gauss.**

Nella teoria delle probabilità la legge di distribuzione di Gauss riveste quindi un ruolo di notevole importanza in quanto essa costituisce una legge limite, cui tendono tutte le distribuzioni sotto condizioni che si verificano abbastanza di frequente. Karl Friedrich Gauss introdusse la curva che porta il suo nome studiando il moto dei corpi celesti. Altri scienziati dell'epoca la usarono per descrivere fenomeni molto diversi tra loro come i colpi di sfortuna nel gioco d'azzardo o la distribuzione dei tiri attorno a bersagli. Da qui i nomi "curva di Gauss" o "curva degli errori".

Fu Francis Galton ad intuire che la curva in questione, da lui detta anche *ogiva*, poteva essere applicata a moltissimi fenomeni molto diversi tra loro e non solo allo studio degli "errori". Questa idea dell'esistenza di una curva capace di descrivere "dati" in generale, lo portò ad usare il termine *Normale*. Tale curva rappresentava cioè uno substrato *normale*, ovvero la norma, per un grandissimo numero di distribuzioni presenti in natura.

La distribuzione di probabilità di Gauss è una funzione a due parametri λ e σ . Il primo rappresenta il valore medio della variabile aleatoria che coincide anche con il valore più probabile, *moda*, e con il valore centrale, *mediana*, in quanto la curva di Gauss è una curva simmetrica attorno al suo valore massimo, il secondo la *deviazione standard* che dà una stima della larghezza della curva stessa. La distribuzione di probabilità *normale* è descritta dalla funzione

$$f(t)_{\lambda,\sigma} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(t-\lambda)^2}{2\sigma^2}\right\} \quad (4.1)$$

con $t \in \mathbb{R}$ e

$$\int_{-\infty}^{\infty} f(t)_{\lambda,\sigma} dt = 1. \quad (4.2)$$

Ricordando le definizioni già viste per il caso discreto, eq.ni (2.3) e (2.6), la loro generalizzazione al caso continuo sarà:

$$\lambda = \langle t \rangle = \int_{-\infty}^{\infty} t f(t)_{\lambda,\sigma} dt \quad (4.3)$$

$$\sigma_t^2 = \langle t^2 \rangle - \langle t \rangle^2 = \int_{-\infty}^{\infty} t^2 f(t)_{\lambda,\sigma} dt - \left(\int_{-\infty}^{\infty} t f(t)_{\lambda,\sigma} dt \right)^2. \quad (4.4)$$

È inoltre utile sapere che

$$\int_{\lambda-\sigma}^{\lambda+\sigma} f(t)_{\lambda,\sigma} dt = 0,68; \quad (4.5)$$

$$\int_{\lambda-2\sigma}^{\lambda+2\sigma} f(t)_{\lambda,\sigma} dt = 0,95; \quad (4.6)$$

$$\int_{\lambda-3\sigma}^{\lambda+3\sigma} f(t)_{\lambda,\sigma} dt = 0,997. \quad (4.7)$$

Le equazioni (4.5), (4.6) e (4.7) ci dicono che effettuando una ulteriore misura della variabile t con il 68% della probabilità questa cadrà nell'intervallo $I_1 = [\bar{t} \pm \sigma]$, con il 95% della probabilità nell'intervallo $I_2 = [\bar{t} \pm 2\sigma]$, e quasi certamente, 99.7%, nell'intervallo $I_3 = [\bar{t} \pm 3\sigma]$. In altre parole l'intervallo I_3 corrisponde alla certezza.

Come sopra detto qualsiasi distribuzione di probabilità per numero di prove sufficientemente grande e nel limite del continuo segue la distribuzione di probabilità di Gauss. Infatti nel caso in cui n sia molto grande e la variabile aleatoria assuma più di due valori, il teorema di Laplace ci permette di generalizzare la distribuzione binomiale nel seguente modo. Riprendiamo allora il modello delle scatole e supponiamo ora che le scatole da riempire non siano più solamente due, come nel caso della distribuzione binomiale, bensì un numero qualsiasi che noi indicheremo con M . Ad ogni scatola sarà associata una probabilità ω_i con $i = 1, 2, \dots, M$, ed i numeri k_i ($i = 1, 2, \dots, M$) rappresenteranno il numero di eventi nella i -esima scatola, cioè il numero di successi dell' i -esimo evento². Per avere la probabilità totale si dovranno evidentemente moltiplicare le probabilità di tutte le possibili M configurazioni indipendenti e quindi la distribuzione di probabilità assumerà la forma

$$P(\{k_i\}, \omega, n) = n! \prod_{i=1}^M \frac{\omega_i^{k_i}}{k_i!} \quad (4.8)$$

con

$$\sum_{\{k_i\}} P(\{k_i\}, \omega, n) = 1, \quad \sum_{i=1}^M \omega_i = 1 \quad \text{e} \quad \sum_{i=1}^M k_i = n. \quad (4.9)$$

³ Il teorema di Laplace allora afferma che:

²L'analogia con la distribuzione binomiale si costruisce osservando che per $M = 2$ si ha $\omega_1 = p$ e $\omega_2 = 1 - p$ ed anche $k_1 = k$ e $k_2 = n - k$.

³Nel caso $M = 2$ si hanno i vincoli della distribuzione binomiale $\omega_1 + \omega_2 = p + (1 - p) = 1$, $k_1 + k_2 = k + n - k = n$ e $\sum_k P_k = 1$ che sono generalizzati dalle relazioni (4.9).

1. la distribuzione di probabilità (4.8) ammette un unico massimo per $k_i = \bar{k}_i$ ($i = 1, 2, \dots, M$);
2. nell'intorno di tale massimo e nel limite $n \gg 1$, $k_i \gg 1$ ($i = 1, 2, \dots, M$), la distribuzione di probabilità (4.8) diviene

$$P(\{k_i\}, \omega, n) \sim P(\{\bar{k}_i\}, \omega, n) \prod_{i=1}^M \exp\left\{-\frac{(k_i - \bar{k}_i)^2}{2\bar{k}_i}\right\} \quad (4.10)$$

da cui risulta evidente, confrontando con la (4.1), che la distribuzione di probabilità (4.8) si comporta come un prodotto di distribuzioni normali gaussiane con valori medi di k_i dati da \bar{k}_i e con varianze date da $\sigma_i^2 = \bar{k}_i$ con $i = 1, 2, \dots, M$.

Vale la pena rimarcare che la larghezza relativa della i -esima gaussiana data da

$$\frac{\sigma_i}{\bar{k}_i} = (\bar{k}_i)^{-1/2} \quad (i = 1, 2, \dots, M)$$

tende a zero per \bar{k}_i molto grandi. In altre parole le gaussiane risultano essere estremamente piccate attorno ai loro massimi, che coincidono con il valore medio \bar{k}_i , all'aumentare del valore medio stesso.

5 Conteggio di particelle

Applichiamo ora quanto sopra esposto al caso di conteggio di particelle o meglio ancora al caso di conteggio di muoni μ^\pm prodotti negli sciami di raggi cosmici.

Da quanto sopra esposto risulta chiaro che nel caso di conteggi, siano essi dovuti ad eventi rari per i quali vale la distribuzione di Poisson (3.2) o ad eventi con numero di prove $n \gg 1$ e valore medio \bar{n} anch'esso molto maggiore di uno, $\bar{n} \gg 1$, per i quali vale (4.10), il numero di conteggi è dato da

$$N = \bar{n} \pm (\bar{n})^{1/2}.$$

Supponiamo ora di voler determinare l'efficienza ε_2 della camera 2 posta tra la camera 1 e la camera 3. L'efficienza è definita come il rapporto tra il numero di conteggi effettuati da tutte e tre le camere ed il numero di conteggi effettuati dalla sola camera 2. Quindi si avrà

$$\varepsilon_2 = \frac{\bar{n}_{123}}{\bar{n}_{13}} \quad (5.1)$$

dove \bar{n}_{123} e \bar{n}_{13} sono rispettivamente il valor medio dei conteggi delle camere 1-2-3 in coincidenza ed il valor medio dei conteggi delle camere 1-3, le cui incertezze associate sono rispettivamente $(\bar{n}_{123})^{1/2}$ e $(\bar{n}_{13})^{1/2}$. Dobbiamo allora rispondere alla domanda: quale è l'errore associato a ε_2 ? A tale scopo determiniamo la legge di propagazione degli errori nelle misure indirette.

Noi sappiamo che data una funzione di tre o più variabili indipendenti, che indicheremo con $g = g(x, y, z)$, ognuna delle quali affetta da una incertezza pari rispettivamente a dx, dy, dz , si ha, secondo la propagazione lineare degli errori, che:

$$\sigma_g \equiv dg = \left| \frac{\partial g}{\partial x} \right| dx + \left| \frac{\partial g}{\partial y} \right| dy + \left| \frac{\partial g}{\partial z} \right| dz. \quad (5.2)$$

Esiste tuttavia un'altra relazione di propagazione degli errori nelle misure indirette dovuta a Gauss della quale la (5.2) ne è il limite superiore. Essa risulta particolarmente utile quando le variabili aleatorie che si stanno trattando seguono una distribuzione di probabilità simmetrica rispetto al valore medio, come è la distribuzione gaussiana ed inoltre tali variabili siano tutte indipendenti tra loro o al più debolmente interagenti in modo tale che la distribuzione di probabilità fattorizzi nelle variabili in questione. In questo caso si ha:

$$\sigma_g = \sqrt{\left(\frac{\partial g}{\partial x} \right)_{x=\bar{x}}^2 \sigma_x^2 + \left(\frac{\partial g}{\partial y} \right)_{y=\bar{y}}^2 \sigma_y^2 + \left(\frac{\partial g}{\partial z} \right)_{z=\bar{z}}^2 \sigma_z^2}, \quad (5.3)$$

dove con σ_x, σ_y e σ_z abbiamo indicato la deviazione standard associata ad ogni singola variabile. La relazione (5.3) può essere dimostrata come segue. Consideriamo per semplicità una funzione g di sole due variabili x, y e sviluppiamola in serie di Taylor attorno ai valori medi di x e di y che indicheremo rispettivamente con \bar{x} e \bar{y} . Al primo ordine si ha:

$$\begin{aligned} g(x, y) &\sim g(\bar{x}, \bar{y}) + (x - \bar{x}) \left(\frac{\partial g}{\partial x} \right)_{x=\bar{x}} + (y - \bar{y}) \left(\frac{\partial g}{\partial y} \right)_{y=\bar{y}} + \dots \\ &\sim \bar{g}(x, y) + (x - \bar{x}) \left(\frac{\partial g}{\partial x} \right)_{x=\bar{x}} + (y - \bar{y}) \left(\frac{\partial g}{\partial y} \right)_{y=\bar{y}} + \dots \end{aligned} \quad (5.4)$$

Nella (5.4) abbiamo giustamente supposto che $g(\bar{x}, \bar{y}) = \bar{g}(x, y)$ cioè che il valore medio della funzione $g(x, y)$ sia proprio il valore che la funzione assume quando $x = \bar{x}$ e $y = \bar{y}$, in altre parole la $g(x, y)$ è piccata attorno al valore $x = \bar{x}, y = \bar{y}$.

Calcoliamo ora la fluttuazione quadratica media della variabile $g(x, y)$ sec-

ondo la definizione (2.6)

$$\begin{aligned}
\sigma_g^2 \equiv \langle g^2 \rangle - \langle g \rangle^2 &= \left\langle \left[(x - \bar{x}) \left(\frac{\partial g}{\partial x} \right)_{x=\bar{x}} \right]^2 \right\rangle + \left\langle \left[(y - \bar{y}) \left(\frac{\partial g}{\partial y} \right)_{y=\bar{y}} \right]^2 \right\rangle \\
&+ \left\langle 2(x - \bar{x})(y - \bar{y}) \left(\frac{\partial g}{\partial y} \right)_{y=\bar{y}} \left(\frac{\partial g}{\partial x} \right)_{x=\bar{x}} \right\rangle \\
&+ \left\langle 2\bar{g}(x, y)(x - \bar{x}) \left(\frac{\partial g}{\partial x} \right)_{x=\bar{x}} \right\rangle \\
&+ \left\langle 2\bar{g}(x, y)(y - \bar{y}) \left(\frac{\partial g}{\partial y} \right)_{y=\bar{y}} \right\rangle \tag{5.5}
\end{aligned}$$

È allora immediato osservare che gli ultimi tre termini della (5.5), per variabili indipendenti, per le quali cioè la distribuzione di probabilità in x, y è il prodotto della distribuzione di probabilità in x e della distribuzione di probabilità in y , è zero se la distribuzione di probabilità è una funzione simmetrica rispetto al valore medio. Ciò posto è immediato osservare che dalla eq. (5.5) segue la (5.3) essendo $\sigma_x = (x - \bar{x})$ e $\sigma_y = (y - \bar{y})$.

Applichiamo ora quanto sopra esposto al calcolo dell'errore assoluto sull'efficienza ε_2 della camera 2 data dalla relazione (5.1), determiniamo cioè lo scarto quadratico medio dell'efficienza usando l'eq. (5.3) che in questo caso specifico assume la forma:

$$\begin{aligned}
\sigma_\varepsilon^2 &= \left(\frac{\partial \varepsilon}{\partial n_{123}} \right)_{n_{123}=\bar{n}_{123}}^2 \bar{n}_{123} + \left(\frac{\partial \varepsilon}{\partial n_{13}} \right)_{n_{13}=\bar{n}_{13}}^2 \bar{n}_{13} \\
&= \frac{\bar{n}_{123}}{\bar{n}_{13}^2} + \frac{\bar{n}_{123}^2}{\bar{n}_{13}^3} = \frac{\bar{n}_{123}^2}{\bar{n}_{13}^2} \left(\frac{1}{\bar{n}_{123}} + \frac{1}{\bar{n}_{13}} \right) \tag{5.6}
\end{aligned}$$

da cui si ottiene che l'efficienza della camera 2 è data da:

$$\varepsilon_2 = \frac{\bar{n}_{123}}{\bar{n}_{13}} \left(1 \pm \sqrt{\frac{1}{\bar{n}_{123}} + \frac{1}{\bar{n}_{13}}} \right). \tag{5.7}$$