Parte I

Elementi di Teoria della Probabilità

Capitolo 1

Distribuzioni di probabilitá

Testi per la consultazione:
J. R. Taylor – Introduzione all'analisi degli errori
P. R. Bevington – Data reduction and error analysis for the physical sciences.

1.1 Premessa

Ricordiamo brevemente alcuni concetti di Probabilitá e Statistica giá introdotti nei Corsi precedenti.

Teoria della Probabilitá

Si chiama **grandezza casuale** o variabile stocastica una grandezza che, in un dato esperimento, puó assumere uno qualsiasi dei suoi valori possibili, in modo *imprevedibile a priori*.

La probabilitá di un evento casuale é il rapporto fra il numero dei casi favorevoli all'evento ed il numero totale dei casi possibili; la somma della probabilitá di un evento e dell'evento contrario é uguale all'unitá.

Teorema della probabilitá totale

Quando un evento puó avverarsi in piú modi diversi, mutuamente esclusivi, la sua probabilitá é la somma delle probabilitá che si verifichi in ciascuno di quei modi.

Teorema della probabilitá composta

Quando un evento dipende dal verificarsi di altri eventi indipendenti, la sua probabilitá é uguale al prodotto delle probabilitá degli eventi singoli; se gli eventi non sono indipendenti e si succedono in un dato ordine, le loro probabilitá sono quelle che essi hanno qualora gli eventi precedenti si siano giá verificati.

In un esperimento si effettui la misura di una grandezza x; quando il numero di misure effettuate diventa molto grande, la distribuzione delle frequenze sperimentali tende ad una qualche funzione (continua o discreta a seconda della natura della grandezza x) che prende il nome di *distribuzione limite o funzione di probabilitá, $f(x)$ o $f(x_i)$, tale per cui, nel caso continuo, la frazione di misure che cadono in un qualunque piccolo intervallo compreso tra un valore x ed un altro $x + dx$ é dato da: $f(x)dx$, cioé $f(x)dx$ é la probabilitá che una misura dia un risultato compreso tra x e $x + dx$; nel caso discreto, se x_i, con i=1, ... N, é l'insieme dei possibili valori della grandezza x, $f(x_i)$ é la probabilitá del valore x_i.* La frazione di misure comprese tra un valore a ed un valore b é data da:

$$F(b,\ a) \ = \ F(x,b) \ - \ F(x,a) \ = \ \int_{-\ \infty}^{b} f(x)\ dx \ - \ \int_{-\ \infty}^{a} f(x)\ dx$$

ovvero da:

$$F(b,\ a) \ = \ F(x,b) \ - \ F(x,a) \ = \ \sum_{x_i\ <b} f(x_i) \ - \ \sum_{x_i\ <a} f(x_i)$$

dove la funzione $F(x, a)$ si chiama *funzione di partizione di probabilitá* e rappresenta la probabilitá di ottenere un valore $\leq a$. La funzione di probabilitá gode della seguente proprietá:

$$\int_{-\infty}^{+\infty} f(x)\ dx \ = \ 1$$

ovvero

$$\sum_{i=1}^{N} f(x_i) \ = \ 1$$

e si dice che essa é normalizzata.

Di grande importanza per l'applicazione del Calcolo delle Probabilitá ai casi concreti é la **legge dei grandi numeri** per la quale *al crescere del numero di prove (effettuate tutte nelle stesse condizioni), la frequenza di un evento tende a diventare uguale alla sua probabilitá.*

Nota la distribuzione di probabilitá, $f(x)$ o $f(x_i)$, di una certa grandezza x, il valore medio di tale grandezza é definito come:

$$E(x) \ = \ \frac{\int_{-\infty}^{+\infty}\ x\ f(x)\ dx}{\int_{-\infty}^{+\infty}\ f(x)\ dx}$$

nel caso continuo, oppure come:

$$E(x) \ = \ \frac{\sum_{i=1}^{N}\ x_i\ f(x_i)}{\sum_{i=1}^{N}\ f(x_i)}$$

nel caso discreto. Esso rappresenta il valore che ci si aspetta di ottenere in media, effettuando un gran numero di misure.

La varianza della distribuzione dei valori della grandezza x é definita come:

$$Var(x) \ = \ E\left((x \ - \ E(x))^2\right)$$

cioé:

$$Var(x) \ = \ \int_{-\infty}^{+\infty} (x \ - \ E(x))^2 \ f(x) \ dx$$

ovvero

$$Var(x) \ = \ \sum_{i=1}^{N} (x_i \ - \ E(x))^2 \ f(x_i)$$

nei due casi. Si dimostra, inoltre che:

$$Var(x) \ = \ E(x^2) \ - \ (E(x))^2$$

La radice quadrata della varianza é la deviazione standard:

$$\sigma \ = \ \sqrt{Var(x)}$$

e fornisce una misura della dispersione dei valori attorno al valore medio. Di particolare importanza é la distribuzione gaussiana o normale di probabilità:

$$f_{\mu, \ \sigma}(x) \ = \ \frac{1}{\sigma \ \sqrt{2\pi}} \ e^{-(x \ - \ \mu)^2/2 \ \sigma^2}$$

che possiede valor medio $E(x) = \mu$ e varianza $Var(x) = \sigma^2$. Inoltre l'area al di sotto della curva, compresa tra $\mu - \sigma$ e $\mu + \sigma$ é all'incirca il 68% dell'area totale, quella compresa tra $\mu \ - \ 2\ \sigma$ e $\mu \ + \ 2\ \sigma$ il 95% e quella compresa tra $\mu \ - \ 3\ \sigma$ e $\mu \ + \ 3\ \sigma$ il 99%: questo significa che una singola misura della grandezza x ha una probabilitá del 68% circa di cadere tra $\mu \ - \ \sigma$ e $\mu \ + \ \sigma$ e così via. Nella figura 1.1 é riportato un sommario delle proprietá di tale distribuzione.

Viene spesso usata la distribuzione normale standard, con media $\mu = 0$ e varianza $\sigma^2 = 1$, che rappresenta la distribuzione della variabile standard $z \ = \ \frac{x \ - \ \mu}{\sigma}$:

$$f_{0, \ 1}(z) \ = \ \frac{1}{\sqrt{2\pi}} \ e^{-z^2/2}$$

L'importanza della distribuzione normale é dovuta alla sussistenza del **Teorema limite centrale**, per il quale *la distribuzione di una grandezza somma di un gran numero di grandezze indipendenti, distribuite ciascuna*

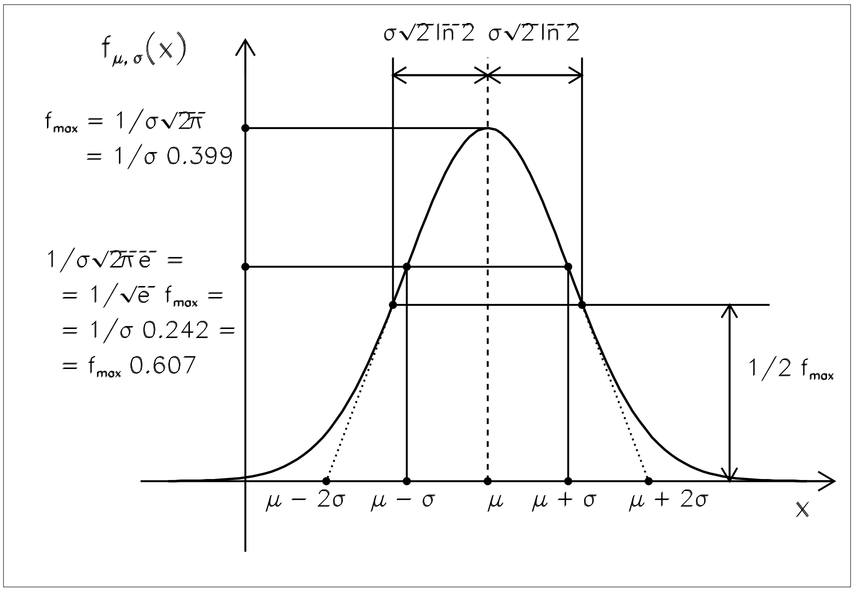

Figura 1.1:

secondo qualunque funzione di probabilitá, tende alla distribuzione normale, purché ciascuna delle grandezze abbia effetto trascurabile sul valore della somma. Esso giustifica l'assunzione che la distribuzione delle misure di una grandezza, a cui si somma inevitabilmente un gran numero di errori casuali infinitesimi, sia una gaussiana centrata sul valore vero della grandezza, essendo la distribuzione degli errori casuali una gaussiana centrata sullo zero.

Statistica Matematica

Quando si esegue una serie di n prove di saggio di una certa grandezza casuale, tutte nelle stesse condizioni, si effettua un **campionamento** della grandezza. La distribuzione dei risultati delle n misure prende il nome di **distribuzione campionaria** e da essa si possono dedurre informazioni sulla distribuzione di probabilitá della variabile in esame, usando i metodi della Statistica Matematica. La distribuzione di frequenze introdotta all'inizio é di fatto una quantitá statistica; altre quantitá statistiche importanti sono:

- la **media aritmetica** o **valor medio empirico**, definito come:

$$\bar{x} \;=\; \frac{\sum_{i=1}^{n}\; x_i}{n}$$

- la **varianza empirica**:

$$s_x^2 = \frac{1}{n - 1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

- la **varianza empirica della media**:

$$s_{\bar{x}}^2 = \frac{s_x^2}{n}$$

Teoria degli Errori

L' **errore** in una misura scientifica é l'inevitabile incertezza che é presente in tutte le misure. Gli errori non sono sbagli e non si possono evitare. Non esistono misure esatte, cioé prive di errore: si puó solo cercare di ridurre il piú possibile gli errori e di stimarne l'entitá.

Gli errori si classificano in **accidentali** e **sistematici**. Gli errori sistematici agiscono sempre nello stesso verso (algebricamente) e sono dovuti ad imperfezioni dell'apparato sperimentale, della teoria applicata o ad aberrazioni costanti dell'osservatore; possono essere compensati fino ad eliminarli. Gli errori accidentali sono dovuti al verificarsi di varie circostanze che, ripetendo le misure in condizioni supposte identiche, agiscono casualmente ora in un verso ora nell'altro. Questi errori NON POSSONO ESSERE ELIMINATI e sono l'oggetto della Teoria degli Errori.

Supponendo di ripetere la misura di una grandezza n volte, sempre con lo stesso apparato, il valore che verrá riportato come risultato del processo di misura é la media aritmetica $\bar{x}$ dei valori ottenuti nelle singole mi-sure. L'errore commesso nella singola misura i–esima é la differenza: $x_i - \bar{x}$. L'**errore (quadratico) medio** su una singola misura é la radice quadrata della varianza empirica:

$$s_x = \sqrt{\frac{1}{n - 1} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

L' **errore (quadratico) medio della media aritmetica** é dato da:

$$s_{\bar{x}} = \frac{s_x}{\sqrt{n}}$$

Il risultato delle n misurazioni viene indicato come: $x = \bar{x} \pm s_{\bar{x}}$. Non ha significato fisico il risultato di una misura riportato senza indicare l'errore da cui é affetto.

Se le misure vengono effettuate con apparati diversi, aventi sensibilitá diverse, allora il risultato dovrá essere dato dalla **media pesata** dei valori e dal suo errore:

$$\bar{x}_P = \frac{\sum w_i \; x_i}{\sum w_i} \quad w_i = \frac{1}{s_i^2}$$

dove s_i é l'errore sulla i–esima misura, mentre l'errore sulla media é:

$$s_{\bar{x}_P} = \sqrt{\frac{1}{\sum w_i}}$$

Se, invece, si effettua solamente una misura di una grandezza, si riporta il valore ottenuto indicando come errore la risoluzione della misura.

Il valor medio e l'errore vanno riportati con 2 cifre significative.

Se una grandezza y viene misurata in maniera indiretta, cioé se essa dipende da altre grandezze x_i che vengono misurate direttamente ed in modo indipendente per poi calcolare il valore di y, l'errore su y si ottiene con la regola di **propagazione degli errori**:

$$s_y = \sqrt{\sum \left(\frac{\partial y}{\partial x_i}\right)^2 s_i^2}$$

Ricordiamo infine alcune definizioni utili:

- Sensibilitá di uno strumento: é il minimo valore rilevabile con lo strumento
- Risoluzione di uno strumento: é il minimo intervallo di validitá di una misura rilevabile con lo strumento
- Accuratezza: é la capacitá di ottenere nella misura un valore prossimo al valore vero della grandezza
- Precisione: é il grado di esattezza con cui viene dato un risultato

1.2 Distribuzione binomiale

La distribuzione binomiale descrive la probabilitá del risultato di una successione di n **prove bernoulliane**, cioé una successione di n **esperimenti indipendenti con due sole possibilitá per ogni prova: successo e insuccesso e con probabilitá p di successo e $q = (1 - p)$ di insuccesso per ogni prova.**

Se lanciamo una moneta in aria e la lasciamo cadere, c'e' una probabilitá del 50% che essa cada con il lato "testa" verso l'alto e del 50% che essa cada con il lato "croce" verso l'alto (a meno che la moneta non sia truccata ed una delle due facce non sia piú 'pesante' dell'altra). Con questo si intende dire che, se il lancio della moneta viene ripetuto piú volte, la frazione di volte in cui essa cadrá a terra con il lato testa verso l'alto si avvicina asintoticamente a 1/2, indicando che c'e' una probabilitá di 1/2 per tale evento. Per ogni singolo lancio la probabilitá non puó determinare se la moneta cadrá con testa verso l'alto o meno; puó solo descrivere come ci aspettiamo che l'esito di un gran numero di lanci si divida tra le due possibilitá.

Supponiamo di lanciare due monete insieme. Ci sono ora quattro diverse possibili combinazioni del modo in cui esse possono cadere a terra: entrambe con testa verso l'alto, entrambe con croce verso l'alto e due combinazioni di testa e croce a seconda di quale delle due ha il lato testa verso l'alto. Poiché ciascuna di queste combinazioni é ugualmente probabile, la probabilitá per ciascuna di esse é 1/4 o 25% (probabilitá composta di due eventi indipendenti con probabilitá 50% ciascuno). Per trovare quale sia la probabilitá di ottenere una combinazione testa–croce senza distinguere tra le due, dobbiamo combinare le probabilitá delle due combinazioni (testa–croce) e (croce–testa), sommandole (probabilitá totale di eventi mutuamente esclusivi): così la probabilitá della combinazione testa–croce é 1/2. In altri termini, la probabilitá della combinazione testa–croce é data dal prodotto del numero di permutazioni che danno la combinazione (non ordinata) testa–croce, moltiplicato per la probabilitá della singola permutazione. Notiamo che la somma delle probabilitá di tutte le possibilitá $(1/4 + 1/4 + 1/4 + 1/4)$ é sempre 1 poiché un qualche risultato si deve comunque produrre.

Estrapoliamo queste idee al caso generale. Supponiamo di lanciare in aria n monete contemporaneamente, dove n é un intero qualunque. Qual'e' la probabilitá che esattamente x di queste monete cadano con il lato testa verso l'alto, senza distinguere quale moneta appartenga di fatto al gruppo "testa verso l'alto" e quale al gruppo "croce verso l'alto" ? Possiamo considerare che la probabilitá $P(x,\ n)$ sia una funzione del numero n di monete lanciate simultaneamente e del numero x di monete che cadono con il lato testa verso l'alto. Naturalmente in ogni esperimento reale x sará un intero, ma possiamo supporre che sia una variabile continua e che la probabilitá vari lentamente con esso per effettuare calcoli matematici.

Se vengono lanciate n monete ci sono 2^n differenti possibili modi o combinazioni in cui esse cadono a terra. Infatti la prima moneta puó cadere con due possibili orientazioni; per ognuna di queste la seconda moneta ha anch'essa due possibili orientazioni, per ogni orientazione della seconda anche la terza moneta ha due possibilitá e così via. Poiché i modi sono tutti ugualmente probabili, la probabilitá che ognuno di questi modi si produca ad ogni lancio simultaneo delle n monete é $(1/2)^n$ (probabilitá della singola combinazione).

Quanti di questi modi contribuiscono alla nostra osservazione di x monete con il lato testa verso l'alto? Immaginamo due scatole, una con una etichetta "testa" e divisa in x scomparti, e la seconda con una etichetta "croce". Consideriamo dapprima il problema di quante "permutazioni" delle n monete danno come risultato la separazione voluta di x monete in una scatola e $n - x$ nell'altra e poi considereremo il problema di quante combinazioni di queste permutazioni possono essere considerate diverse una dall'altra.

Per contare il numero di *permutazioni* $Pm(n, x)$ prendiamo le monete una alla volta dall'insieme iniziale di n monete, identificandole con il loro numero d'ordine, e mettiamone x nella scatola "testa". Possiamo scegliere la prima moneta tra n, mentre la seconda puó essere scelta solo piú tra $n - 1$, la terza tra $n - 2$ e così via finché si arriva alla moneta x che puó essere scelta tra le $n - x + 1$ monete rimanenti. Il numero totale di scelte che possiamo operare per mettere x monete negli scomparti della scatola "testa" é il prodotto dei numeri delle scelte per le singole monete:

$$Pm(n, x) = n\,(n - 1)\,(n - 2)\,...(n - x + 2)\,(n - x + 1) \tag{1.1}$$

Questa espansione puó essere espressa piú facilmente in termini di fattoriali:

$$Pm(n, x) = \frac{n!}{(n - x)!} \tag{1.2}$$

dove la funzione fattoriale é definita (per valori n interi) come:

$$n! = n\,(n - 1)\,(n - 2)\,...(3)\,(2)\,(1)$$

$$1! = 1 \qquad 0! = 1$$

Finora abbiamo calcolato il numero di permutazioni $Pm(n, x)$ che danno x monete nella scatola "testa" e $n - x$ nella scatola "croce", avendo identificato in base al suo numero d'ordine quale moneta sia stata messa in ciascuna delle fessure della prima scatola: cioé abbiamo *ordinato* le x monete nella scatola “testa”. Invece, per determinare quale sia la probabilitá che x monete su n cadano con il lato testa verso l’alto, noi dobbiamo considerare solamente quali monete hanno dato testa e quali croce delle n lanciate, senza ordinarle. Pertanto dobbiamo considerare contributi differenti solo se ci sono monete differenti nelle due scatole e non se le x monete nella scatola “testa” sono state solo permutate in differenti sequenze temporali di scelta. (Cioé, se dovendo scegliere $x = 3$ monete da mettere nella scatola “testa” in una prima prova si scelgono la 1, la 3 e la 9 delle $n = 10$ disponibili, la prova che fornisce nell’ordine la 3, la 1 e la 9 non dà un contributo diverso al risultato di avere le monete 1, 3 e 9 nella scatola in qualunque ordine esse vengano estratte).

Il numero delle differenti *combinazioni* $C(n,\ x)$ delle permutazioni enumerate sopra é dato dalla combinazione degli $x!$ differenti modi in cui le x monete nella scatola “testa” possono essere permutate tra di loro. Per ogni $x!$ permutazioni ci sará una sola nuova combinazione. Così, il numero di diverse combinazioni $C(n,\ x)$ é il numero delle permutazioni $Pm(n,\ x)$ diviso per il fattore di degenerazione delle permutazioni:

$$C(n,\ x) \ = \ \frac{Pm(n,\ x)}{x!} \ = \ \frac{n!}{(n\ -\ x)!\ x!} \ \equiv \begin{pmatrix} n \\ x \end{pmatrix} \tag{1.3}$$

Questo é il numero di differenti possibili combinazioni di n oggetti, presi x alla volta. Il simbolo $\begin{pmatrix} n \\ x \end{pmatrix}$ prende il nome di *fattore binomiale*.

La probabilitá $P(x,\ n)$ di osservare x monete con testa e $n\ -\ x$ con croce é il prodotto del numero di diverse combinazioni $C(n,\ x)$, cioé il numero di permutazioni ’diverse’ che contribuiscono ad una osservazione di x teste su n monete, per la probabilitá della singola combinazione, che sappiamo essere $(1/2)^n$.

Di fatto, dobbiamo separare la probabilitá di ogni combinazione in due parti: una parte é la probabilitá di x teste, pari a $(1/2)^x$, e l’altra é la probabilitá di $n - x$ croci, pari a $(1/2)^{(n\ -\ x)}$. Infatti ogni combinazione rappresenta un evento composto dai due eventi indipendenti: x teste e $(n\ -\ x)$ croci. In generale la probabilitá p di successo (nel nostro caso “testa”) di un evento non é uguale alla probabilitá $q\ =\ 1\ -\ p$ di insuccesso (nel nostro caso “croce”). La probabilitá per ognuna delle combinazioni di x monete “testa” e $n - x$ monete “croce” é $p^x q^{n\ -\ x}$.

Con queste considerazioni su p e q, la probabilitá $P_B(x,\ n,\ p)$ di osservare x oggetti (successi) su n in uno stato con probabilitá p é data dalla **distribuzione binomiale**:

$$P_B(x,\ n,\ p) \ = \ \begin{pmatrix} n \\ x \end{pmatrix} p^x\ q^{(n\ -\ x)} \ = \ \frac{n!}{(n\ -\ x)!\ x!}\ p^x\ (1-p)^{(n\ -\ x)} \tag{1.4}$$

dove $q \; = \; 1 \; - \; p$. Il nome della distribuzione binomiale viene dal fatto che, come detto, la probabilità $P_B(x, \; n, \; p)$ sono esattamente i termini che rappresentano lo sviluppo della potenza di un binomio. Infatti, il teorema binomiale puó essere scritto nel seguente modo:

$$(p \; + \; q)^n \; = \sum_{x=0}^{n} \left[\left(\begin{array}{c} n \\ x \end{array} \right) \; p^x \; q^{(n \; - \; x)} \right] \tag{1.5}$$

Il termine *(j + 1)* della somma, che corrisponde a $x = j$, di questa espansione, pertanto, é uguale alla probabilitá $P_B(j, \; n, \; p)$.

Possiamo usare questo teorema per dimostrare che i coefficienti della distribuzione binomiale sono normalizzati a 1, cioé che la somma delle probabilitá per tutti i valori di x tra 0 e n é unitaria. Infatti il secondo termine dell'equazione (1.5 é la somma delle probabilitá per tutti i valori di x tra 0 e n e il primo termine é $(p \; + \; q)^n \; = \; (1)^n \; = \; 1$.

Valutiamo ora la **media** o **valore atteso** della distribuzione binomiale.

$$\mu \; = \; E(X) \; = \; \sum_{x=0}^{n} x \, P_B(x, \, n, \, p) \; = \; \sum_{x=0}^{n} x \, \frac{n!}{(n \; - \; x)! \; x!} \, p^x \, q^{(n \; - \; x)} \tag{1.6}$$

possiamo tralasciare il termine con $x \; = \; 0$ in quanto il suo valore é nullo ed estrarre np da ciascuno dei termini rimanenti:

$$E(X) \; = \; np \; \sum_{x=1}^{n} \; \frac{(n \; - \; 1)!}{(x \; - \; 1)! \; (n \; - \; x)!} \; p^{(x \; - \; 1)} \; q^{(n \; - \; x)} \tag{1.7}$$

Poniamo $s \; = \; x \; - \; 1$ nella sommatoria: mentre x assume i valori compresi tra 1 e n, s assume i valori compresi tra 0 e $n \; - \; 1$. Pertanto:

$$E(X) \; = \; np \; \sum_{s=0}^{n-1} \; \frac{(n \; - \; 1)!}{s! \; (n \; - \; 1 \; - \; s)!} \; p^s \; q^{(n \; - \; 1 \; - \; s)} \; = \tag{1.8}$$

$$np \; \sum_{s=0}^{n-1} \; P_B(s, \; n \; - \; 1, \; p) \; = \; np \tag{1.9}$$

nell'ultimo passaggio si é tenuto conto della normalizzazione della distribuzione binomiale (ovvero del teorema binomiale). Questo vuole dire che se facciamo un esperimento con n elementi ed osserviamo il numero x di successi, dopo un gran numero di esperimenti ripetuti la media $\bar{x}$ del numero di successi si avvicinerá ad un valore medio μ dato dal prodotto della probabilitá del singolo successo p per il numero degli elementi.

Valutiamo adesso la **varianza** della distribuzione binomiale:

$$\sigma^2 = E(X^2) - \mu_X^2. \tag{1.10}$$

$$E(X^2) \;=\; \sum_{x=0}^{n} x^2\, P_B(x,\, n,\, p) \;=\; \sum_{x=0}^{n} x^2\, \frac{n!}{(n\; -\; x)!\; x!}\, p^x\, q^{(n\; -\; x)} \tag{1.11}$$

possiamo di nuovo tralasciare il termine con $x \;=\; 0$ in quanto il suo valore é nullo ed estrarre np da ciascuno dei termini rimanenti:

$$E(X^2) \;=\; np\, \sum_{x=1}^{n} x\, \frac{(n\; -\; 1)!}{(x\; -\; 1)!\; (n\; -\; x)!}\, p^{(x\; -\; 1)}\, q^{(n\; -\; x)} \tag{1.12}$$

Poniamo nuovamente $s \;=\; x \;-\; 1$ nella sommatoria ed otteniamo:

$$E(X^2) \;=\; np\, \sum_{s=0}^{n-1} (s\; +\; 1)\, \frac{(n\; -\; 1)!}{s!\; (n\; -\; 1\; -\; s)!}\, p^s\, q^{(n\; -\; 1\; -\; s)} \;= \tag{1.13}$$

$$np\, \sum_{s=0}^{n-1} (s\; +\; 1)\; P_B(s,\; n\; -\; 1,\; p) \tag{1.14}$$

Ma

$$\sum_{s=0}^{n-1} (s\; +\; 1)\; P_B(s,\; n\; -\; 1,\; p) \;= \tag{1.15}$$

$$\sum_{s=0}^{n-1} s\; P_B(s,\; n\; -\; 1,\; p) \;+\; \sum_{s=0}^{n-1} P_B(s,\; n\; -\; 1,\; p) \;= \tag{1.16}$$

$$(n\; -\; 1)\; p +\; 1 \;=\; np \;+\; 1 \;-\; p \;=\; np \;+\; q \tag{1.17}$$

avendo usato la definizione di valore medio trovata prima. Pertanto:

$$E(X^2) \;=\; np\; (np\; +\; q) \;=\; (np)^2 \;+\; npq \tag{1.18}$$

e

$$\sigma^2 \;=\; E(X^2) \;-\; \mu_X^2 \;=\; (np)^2 \;+\; npq \;-\; (np)^2 \;=\; npq \tag{1.19}$$

Se la probabilitá di un singolo successo p é uguale alla probabilitá dell'insuccesso $p \;=\; q \;=\; 1/2$, la distribuzione é simmetrica attorno alla media μ e la mediana ed il valore piú probabile coincidono con la media. In questo caso la varianza σ^2 é data da: $\sigma^2 \;=\; \mu/2$. Se p e q non sono uguali la distribuzione é asimmetrica ed ha una varianza piú piccola. Due esempi

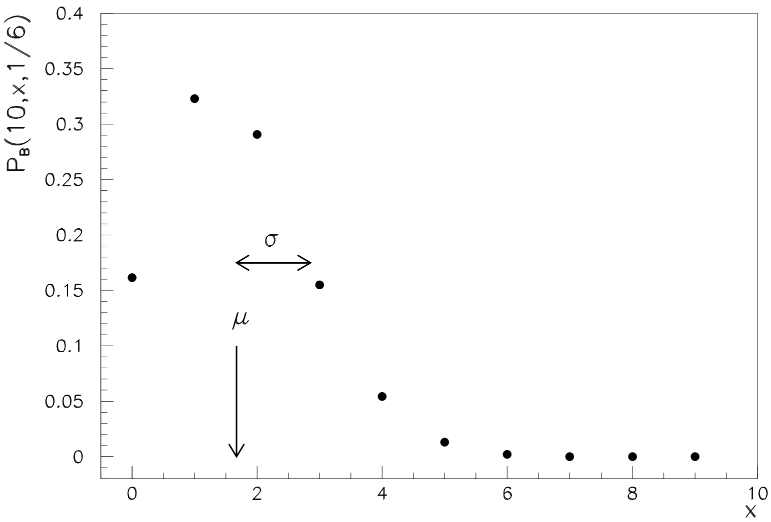

Figura 1.2:

di distribuzioni binomiale sono riportati nelle figure 1.2, nel caso $\mu = 10/6$, $p = 1/6$, e 1.3 nel caso $\mu = 5.0$ e $p = 1/2$.

Inoltre vale il seguente **limite**: **se n é grande e se né p né q sono troppo prossimi allo zero, la distribuzione binomiale puó essere molto bene approssimata da una distribuzione normale standardizzata, in cui la variabile z sia data da:**

$$z = \frac{x - np}{\sqrt{npq}}$$

L'approssimazione migliora al crescere di n e, al limite, le due distribuzioni coincidono.

Riassumendo:

la distribuzione binomiale di probabilitá $P_B(x,\ n,\ p)$ fornisce la probabilitá di osservare x successi in n prove bernoulliane se la probabilitá del singolo successo é p; essa puó essere applicata solo se tutti i suoi parametri, e, in particolare n e p, sono definiti con certezza. É una distribuzione di probabilitá discreta, con le seguenti proprietá:

Somma	1
Media	$\mu = np$
Varianza	$\sigma^2 = npq$
Scarto quadratico medio	$\sigma = \sqrt{npq}$

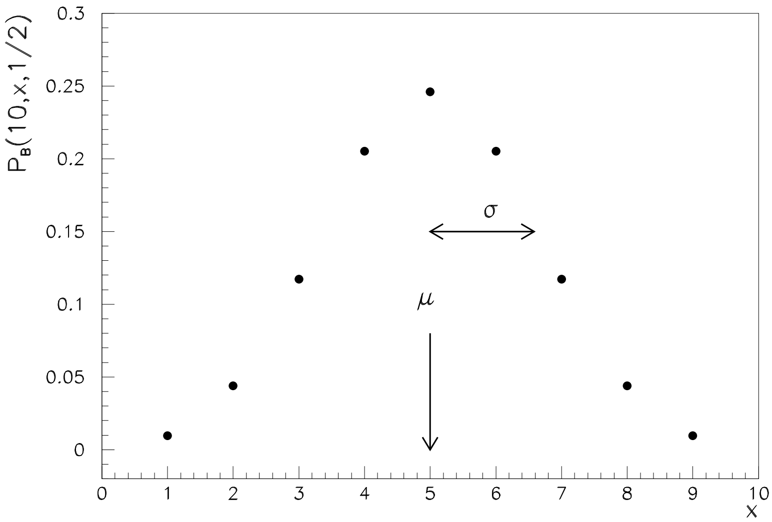

Figura 1.3:

Vediamo alcuni esempi:

1. Trovare la probabilitá che lanciando una moneta buona tre volte si presentino *(a)* tre teste, *(b)* due teste e una croce, *(c)* 2 croci e una testa e *(d)* tre teste.

   Intendendo come successo dell'evento l'uscita della testa, che ha probabilitá 1/2 per una moneta non truccata, si avrá:

   (a) $P_B(3,\ 3,\ 1/2) \ = \ \frac{3!}{3!\ 0!}\ (1/2)^3\ (1/2)^0 \ = \ 1/8$

   (b) $P_B(2,\ 3,\ 1/2) \ = \ \frac{3!}{2!\ 1!}\ (1/2)^2\ (1/2)^1 \ = \ 3/8$

   (c) $P_B(1,\ 3,\ 1/2) \ = \ \frac{3!}{1!\ 2!}\ (1/2)^1\ (1/2)^2 \ = \ 3/8$

   (d) $P_B(0,\ 3,\ 1/2) \ = \ \frac{3!}{0!\ 3!}\ (1/2)^0\ (1/2)^3 \ = \ 1/8$

   Si puó verificare che la somma delle probabilitá dei quattro casi, che rappresentano la totalitá delle combinazioni possibili con due monete, vale 1.

2. Trovare la probabilitá che in 5 lanci di un dado non truccato il 3 si presenti *(a)* mai, *(b)* una volta, *(c)* due volte, *(d)* tre volte, *(e)* quattro volte, *(f)* cinque volte.

   In questo caso il successo é l'uscita del numero 3, che ha probabilitá 1/6 per un dado non truccato. Allora:

   (a) $P_B(0,\ 5,\ 1/6) \ = \ \frac{5!}{0!\ 5!}\ (1/6)^0\ (5/6)^5 \ = \ 3125/7776$

(b) $P_B(1,\ 5,\ 1/6)\ =\ \frac{5!}{1!\ 4!}\ (1/6)^1\ (5/6)^4\ =\ 3125/7776$

(c) $P_B(2,\ 5,\ 1/6)\ =\ \frac{5!}{2!\ 3!}\ (1/6)^2\ (5/6)^3\ =\ 625/3888$

(d) $P_B(3,\ 5,\ 1/6)\ =\ \frac{5!}{3!\ 2!}\ (1/6)^3\ (5/6)^2\ =\ 125/3888$

(e) $P_B(4,\ 5,\ 1/6)\ =\ \frac{5!}{4!\ 1!}\ (1/6)^4\ (5/6)^1\ =\ 25/7776$

(f) $P_B(5,\ 5,\ 1/6)\ =\ \frac{5!}{5!\ 0!}\ (1/6)^5\ (5/6)^0\ =\ 1/7776$

Anche adesso la somma delle probabilitá di tuttl le combinazioni possibili vale 1.

3. Trovare la probabilitá che in una famiglia con quattro figli ci sia *(a)* almeno un maschio, *(b)* almeno un maschio e una femmina.

   (a) La probabilitá di avere almeno un maschio, che é da considerarsi come successo dell'evento, con probabilitá 1/2, puó essere trovata sommando le probabilitá di avere 1 maschio, 2 maschi, 3 maschi e 4 maschi: $\sum_{i=1}^{4} \frac{4!}{i!\ (4\ -\ i)!}\ 1/2^i\ 1/2^{(4\ -\ i)}\ =\ 1/4 + 3/8 + 1/4 + 1/16\ =\ 15/16$, oppure considerando che essa é data anche dall'unitá meno la probabilitá di non avere maschi: $1\ -\ \frac{4!}{0!\ 4!}\ 1/2^0 1/2^4\ =\ 1\ -\ 1/16\ =\ 15/16$

   (b) La probabilitá di avere almeno un maschio e almeno una femmina é data dall'unitá meno la probabilitá di non avere nessun maschio, meno la probabilitá di non avere nessuna femmina (cioé di avere tutti maschi): $1\ -\ \frac{4!}{0!\ 4!}\ 1/2^0 1/2^4\ -\ \frac{4!}{4!\ 0!}\ 1/2^4 1/2^0\ =\ 1\ -\ 1/16\ -\ 1/16\ =\ 7/8$

4. Trovare la probabilitá che, in 10 lanci di una moneta, testa si presenti da 3 a 6 volte comprese usando *(a)* la distribuzione binomiale, *(b)* l'approssimazione normale alla distribuzione binomiale.

   (a)

   $P_B(3,\ 10,\ 1/2)\ =\ \frac{10!}{3!\ 7!}\ (1/2)^3\ (1/2)^7\ =\ 15/128$

   $P_B(4,\ 10,\ 1/2)\ =\ \frac{10!}{4!\ 6!}\ (1/2)^4\ (1/2)^6\ =\ 105/512$

   $P_B(5,\ 10,\ 1/2)\ =\ \frac{10!}{5!\ 5!}\ (1/2)^5\ (1/2)^5\ =\ 63/256$

   $P_B(6,\ 10,\ 1/2)\ =\ \frac{10!}{6!\ 4!}\ (1/2)^6\ (1/2)^4\ =\ 105/512$

   la probabilitá voluta sará la somma: $P_B(3-6,\ 10,\ 1/2)\ =\ 99/128\ =\ 0.7735$.

   (b) La distribuzione di probabilitá di teste in 10 lanci di una moneta é visibile in grafico nella figura 1.4, dove la figura di destra tratta i dati come se fossero continui. La probabilitá richiesta é la somma delle aree dei rettangoli ombreggiati e puó essere approssimata dall'area sotto la curva normale (linea tratteggiata). Se si considerano i dati come continui, l'intervallo da 3 a 6 teste puó essere considerato, in

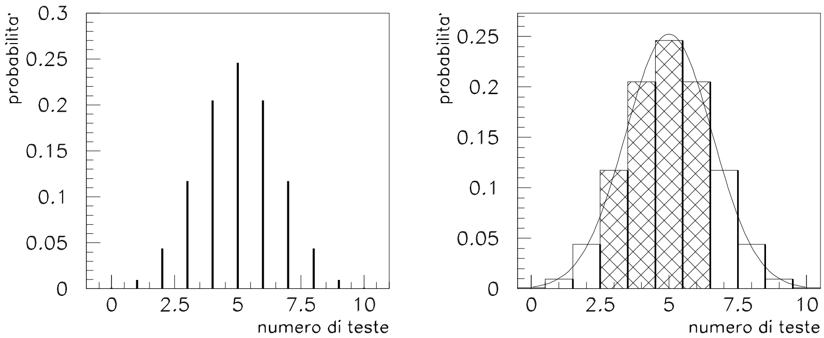

Figura 1.4:

realtá, come l'intervallo da 2.5 a 6.5 teste. La media e la varianza della distribuzione binomiale sono date da $\mu = np = 5$ e $\sigma = \sqrt{npq} = \sqrt{10\ 0.5\ 0.5} = 1.58$. 2.5 in unitá standard vale (2.5 - 5)/1.58 = -1.58 e 6.5 in unitá standard vale (6.5 - 5)/1.58 = 0.95. La probabilitá richiesta sará l'area sotto la curva della distribuzione normale standardizzata, compresa tra z = -1.58 e z = 0.95 (vedi figura 1.5) , ovvero (area tra z = -1.58 e z =0) + (area tra z = 0 e z = 0.95) = 0.4429 + 0.3289 = 0.7718, che si approssima molto al valore vero 0.7735 ottenuto al punto *(a)*: infatti lo scarto é $\sim$ - 0.2%. L'approssimazione é anche migliore per valori piú grandi di n.

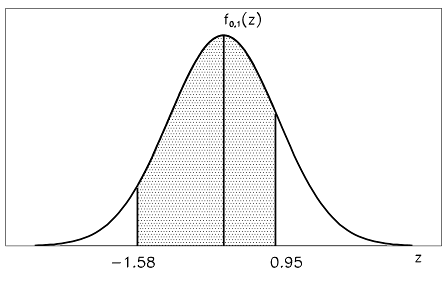

Figura 1.5:

5. Adattamento della distribuzione binomiale ai dati sperimentali.

   Un gruppo di 5 monete viene lanciato 1000 volte e vengono registrate le frequenze con cui si presentano 0, 1, 2, 3, 4, 5 teste.

numero di teste, x	frequenze osservate f(x)	frequenze teoriche	$P_B(x,\ 5,\ 1/2)$
0	38	33,2 o 33	0.0332
1	144	161,9 o 162	0.1619
2	342	316,2 o 316	0,3162
3	287	308,7 o 309	0,3087
4	164	150,7 o 151	0,1507
5	25	29,4 o 29	0,0294

Le frequenze sono riportate nella tabella; adattate una distribuzione binomiale ai dati.

La probabilitá di avere x teste nel lancio di 5 monete é: $P_B(x,\ 5,\ p) = \frac{5!}{x!\ (5\ -\ x)!}\ p^x\ q^{(5\ -\ x)}$; si vuole determinare il valore di p per il quale la distribuzione binomiale, con $n = 5$ e x che varia tra 0 e 5, riproduce la distribuzione delle frequenze osservate. Il numero medio di teste atteso per la distribuzione binomiale é: $\mu\ =\ np\ =\ 5p$; dalla distribuzione delle frequenze osservate, il numero medio di teste risulta: $\frac{\sum\ x\ f(x)}{\sum\ f(x)} = 2,47$. Eguagliando le medie teorica e reale si ottiene $5p\ =\ 2,47$ da cui p = 0,494. Così la distribuzione binomiale adattata é data da: $P_B(x,\ 5,\ 0.494)\ =\ \frac{5!}{x!\ (5\ -\ x)!}\ 0.494^x\ 0.506^{(5\ -\ x)}$. Nella tabella sono anche riportati i valori delle probabilitá, insieme alle frequenze teoriche, pari alle probabilitá moltiplicate per il numero totale di prove (1000).

1.2.1 Test di ipotesi

Una applicazione della distribuzione binomiale di particolare interesse é la verifica di un'ipotesi: una volta che si sappia, o che si assuma di sapere, come dovrebbero essere distribuiti i risultati di un esperimento, si puó verificare se effettivamente lo sono. Consideriamo un insieme di n prove similari ma indipendenti (o "corse"), ciascuna delle quali ha le stesse due possibili uscite, "successo" o "insuccesso". Dapprima si formula una ipotesi, in questo caso semplicemente si assume un valore per la probabilità p di successo in una prova. Questo valore assunto di p determina il numero medio atteso di successi, $\bar{\nu} = np$, in n prove. Se il numero effettivo di successi, ν, nelle nostre n prove è vicino a np, allora non vi è evidenza contro l'ipotesi. Se ν è apprezzabilmente più grande di np, allora calcoliamo la probabilità (data la nostra ipotesi) di ottenere ν o più successi. Se questa probabilità è minore del nostro "livello di significatività" scelto, (cioè 5% o 10%), allora arguiamo che il nostro numero osservato è improbabile in modo inaccettabile (se la nostra ipotesi è corretta) e quindi che la nostra ipotesi dovrebbe essere rigettata. Nello stesso modo, se il nostro numero di successi è apprezzabilmente minore di np, allora possiamo ragionare analogamente, eccetto che dovremmo calcolare la probabilità di ottenere ν o meno successi.

Come avremmo dovuto aspettarci, questo procedimento non ci fornisce una risposta semplice che la nostra ipotesi è certamente vera o certamente falsa. Quello che ci da è una valutazione quantitativa della ragionevolezza dei nostri risultati alla luce dell'ipotesi, così possiamo scegliere un criterio oggettivo, anche se arbitrario, per il rigetto della nostra ipotesi. Quando uno sperimentatore trae una conclusione basandosi su questo tipo di ragionamento, è importante che dica chiaramente quale criterio ha usato e qual era la probabilità calcolata, in modo che il lettore possa giudicare la ragionevolezza delle conclusioni.

Un sondaggio di opinione Come esempio, consideriamo un'elezione tra due candidati, A e B. Suponiamo che il candidato A rivendichi che una ricerca estensiva ha stabilito che egli è favorito dal 60% dell'elettorato, e supponiamo che il candidato B ci chieda di verificare questa affermazione (nella speranza, naturalmente, di mostrare che il numero che è in favore di A è significativamente minore del 60 %).

In questo caso la nostra ipotesi statistica dovrebbe essere che il 60% di elettori è in favore di A; così la probabilità che un campione di elettori selezionato casualmente voterà per A dovrebbe essere $p = 0.6$. Poichè non possiamo chiedere il parere di ogni singolo elettore, con cura ne selezioniamo un campione casuale di 600 e chiediamo le loro preferenze. Se il 60 % è realmente in favore di A, allora il numero atteso nel nostro campione che favorisce A è $np = 600 \times 0.6 = 360$. Se di fatto 330 dicono di preferire A, possiamo dire di aver messo un dubbio significativo sull'ipotesi che il 60% è in favore di A?

Per rispondere a questa domanda, notiamo che secondo l'ipotesi la probabilità che ν elettori voterà A è la probabilità binomiale:

$$\mathrm{P}(\nu \text{ elettori in favore di A}) = \mathrm{P}_B(n,\, p,\, \nu)$$

con n=600 e p=0.6. Poichè n è così grande, è una approssimazione eccellente rimpiazzare la funzione binomiale con l'approppriata funzione di Gauss, con centro in $np = 360$ e deviazione standard $\sigma_\nu = \sqrt{np(1-p)} = 12$. Il numero medio atteso in favore di A è 360. Così il numero che di fatto è stato in favore di A nel nostro campione (330) è minore di quello atteso di 30. Poichè la deviazione standard è 12, il nostro risultato è 2.5 deviazioni standard al di sotto della media supposta. La probabilità di un risultato così basso o più basso (si veda la tabella dell'integrale normale degli errori) è 0.6%. Così il nostro risultato è "altamente significativo" e al livello dell' 1% possiamo confidentemente rigettare l'ipotesi che A è favorito dal 60% degli elettori.

Questo esempio illustra due aspetti generali di questo tipo di test. In primo luogo, avendo trovato che 330 elettori sono in favore di A (cioè 30 in meno del numero atteso) abbiamo calcolato la probabilità che il numero in favore di A sia 330 "o meno". Dapprima si potrebbe aver considerato la probabilità che il numero in favore di A sia esattamente $\nu = 330$. Comunque, questa probabilità è estremamante piccola (0.15%, in realtà), ed anche il risultato più probabile ($\nu = 360$) ha una probabilità bassa (3.3%). Per ottenere una valutazione più approppriata di quanto sia inatteso il risultato $\nu = 330$, dobbiamo includere $\nu = 330$ e ogni risultato che sia ancor più al di sotto della media.

Il nostro risultato $\nu = 330$ è 30 al di sotto di quello atteso. La probabilità di un risultato che è 30 o più al di sotto della media è detta "probabilità ad una coda", poichè è l'area sotto una coda della curva di distribuzione. In alcuni test la probabilità rilevante è la "probabilità a due code" di ottenere un risultato che differisca dalla media attesa di 30 o più in entrambe le direzioni, cioè la probabilità di ottenere $\nu \leq 330$ oppure $\nu \geq 390$. Il fatto di utilizzare la probabilità a una coda o a due code in un test statistico dipende da quale si considera l'alternativa interessante all'ipotesi originale. In questo esempio volevamo mostrare che il candidato A era favorito da "meno" del 60% che egli rivendicava; così la probabilità ad una coda era approppriata. Se avessimo voluto mostrare che il numero in favore di A era diverso dal 60% (in entrambe le direzioni), allora la probabilità a due code sarebbe stata approppriata. In pratica, di solito è chiaro quale probabilità debba essere usata. In ogni caso lo sperimentatore deve sempre stabilire chiaramente quale probabilità e quale livello di significatività siano stati scelti, e qual'era la probabilità calcolata. Con questa informazione il lettore può giudicare il significato dei risultati per se stesso.

1.3 Distribuzione di Poisson

La distribuzione di Poisson rappresenta una approssimazione alla distribuzione binomiale per il caso speciale in cui il numero medio di successi é molto piccolo, cioé quando $\mu << n$ perché $p << 1$. Tale approssimazione é generalmente approppriata per descrivere i risultati di esperimenti di conteggio, in cui i dati rappresentano il numero di conteggi effettuati nell'intervallo di tempo unitario (per esempio lo studio del decadimento radioattivo). Per tali esperimenti la distribuzione binomiale descrive correttamente la probabilitá $P_B(x, n, p)$ di osservare x eventi nell'intervallo di tempo unitario (i nuclei decaduti in un certo intervallo di tempo Δt) su n possibili eventi (il numero totale dei nuclei presenti nel campione), ciascuno dei quali ha una probabilitá p di avvenire (cioé di decadere), peró il fatto che il numero di possibili eventi n sia molto grande rende impossibile un calcolo esatto. Inoltre, in questo tipo di esperimenti, in genere, non sono noti con esattezza né il numero n di possibili eventi, né la probabilitá p del singolo evento. Si conosce, invece, il numero medio di eventi per unitá di tempo μ o la sua stima $\bar{x}$.

Consideriamo, allora, la distribuzione binomiale nel caso limite $p << 1$ ed, in particolare, il suo andamento quando n diventa infinitamente grande, mentre il valore medio $\mu = np$ rimane costante. L'equazione (1.4) per la funzione di probabilitá della distribuzione binomiale puó essere scritta come:

$$P_B(x, n, p) = \frac{1}{x!} \frac{n!}{(n - x)!} p^x (1 - p)^{-x} (1 - p)^n \tag{1.20}$$

Se espandiamo il secondo termine nel seguente modo:

$$\frac{n!}{(n - x)!} = n (n - 1) (n - 2) ... (n - x + 2) (n - x + 1) \tag{1.21}$$

possiamo considerarlo come il prodotto di x termini, ciascuno dei quali é molto prossimo ad n, dato che $x << n$ nella regione di interesse. Questo termine, pertanto, si avvicina asintoticamente a n^x. Allora il prodotto del secondo e del terzo termine diventa: $(np)^x = \mu^x$. Il quarto termine é approssimativamente uguale a $1 + px$ (sviluppo in serie di potenze di $(1 - p)^{-x}$), che tende a 1 quando p diventa molto piccolo. L'ultimo termine puó essere riscritto sostituendo μ/p al posto di n:

$$\lim_{p\to 0}(1 - p)^n = \lim_{p\to 0}\left[(1 - p)^{1/p}\right]^{\mu} = \left(\frac{1}{e}\right)^{\mu} = e^{-\mu} \tag{1.22}$$

Combinando queste approssimazioni, si trova che la distribuzione binomiale di probabilitá $P_B(x, n, p)$ approssima asintoticamente la **distribuzione di Poisson** $P_P(x, \mu)$.

$$\lim_{p\to 0} P_B(x, n, p) = P_P(x, \mu) = \frac{\mu^x}{x!} e^{-\mu} \tag{1.23}$$

La distribuzione di Poisson, come la distribuzione binomiale per $p \neq 1/2$, é asimmetrica attorno al suo valore medio μ, come mostrato nella figura 1.6, in cui $\mu = 1.69$. La distribuzione di Poisson é comunque una distribuzione discreta. Osserviamo che $P_P(x, \mu)$ non va a 0 per $x = 0$, ma non é definita

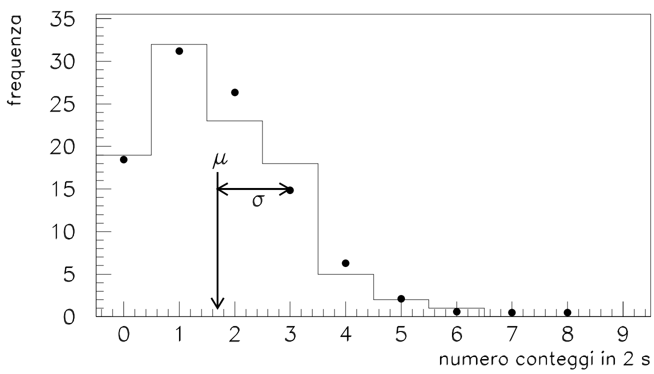

Figura 1.6:

per valori negativi si x. Questa restrizione non é penalizzante per esperimenti di conteggio, dato che il numero di conteggi per unitá di tempo non puó mai essere negativo.

Verifichiamo che la normalizzazione della distribuzione sia corretta:

$$\sum_{x=0}^{\infty} P_P(x, \mu) = \sum_{x=0}^{\infty} \frac{\mu^x}{x!} e^{-\mu} \tag{1.24}$$

Poiché é noto dall'Analisi che:

$$e^{\mu} = \sum_{x=0}^{\infty} \frac{\mu^x}{x!} \tag{1.25}$$

si avrá:

$$\sum_{x=0}^{\infty} P_P(x, \mu) = e^{\mu} e^{-\mu} = 1 \tag{1.26}$$

Valutiamo ora la **media** o **valore atteso** della distribuzione poissoniana.

$$E(X) = \sum_{x=0}^{\infty} x P_P(x, \mu) = \sum_{x=0}^{\infty} x \frac{\mu^x}{x!} e^{-\mu} = \mu \sum_{x=1}^{\infty} \frac{\mu^{x-1}}{(x-1)!} e^{-\mu} \tag{1.27}$$

dove si é trascurato il primo termine della sommatoria poiché il suo valore é nullo. Poniamo poi $s = x - 1$: mentre x assume i valori compresi tra 1 e ∞, s assume i valori compresi tra 0 e ∞. Pertanto:

$$E(X) = \mu \sum_{s=0}^{\infty} \frac{\mu^s}{s!} e^{-\mu} = \mu \tag{1.28}$$

avendo usato la normalizzazione della distribuzione di Poisson dimostrata sopra.

Valutiamo adesso la **varianza** della distribuzione poissoniana:

$$\sigma^2 = E(X^2) - (E(X))^2. \tag{1.29}$$

$$\begin{aligned} E(X^2) &= \sum_{x=0}^{\infty} x^2 P_P(x, \mu) = \sum_{x=0}^{\infty} x^2 \frac{\mu^x}{x!} e^{-\mu} = \\ &= \mu \sum_{x=1}^{\infty} x \frac{\mu^{x-1}}{(x-1)!} e^{-\mu} \end{aligned} \tag{1.30}$$

dove si é nuovamente trascurato il primo termine della sommatoria, che é nullo; poniamo nuovamente $s = x - 1$:

$$E(X^2) = \mu \sum_{s=0}^{\infty} (s + 1) \frac{\mu^s}{s!} e^{-\mu} = \mu \sum_{s=0}^{\infty} (s + 1) P_P(s, \mu) \tag{1.31}$$

Ma:

$$\sum_{s=0}^{\infty} (s+1) P_P(s, \mu) = \sum_{s=0}^{\infty} s P_P(s, \mu) + \sum_{s=0}^{\infty} P_P(s, \mu) = \mu + 1 \tag{1.32}$$

per la definizione di media e la normalizzazione della distribuzione di Poisson, sopra ottenute. Di conseguenza:

$$E(X^2) = \mu (\mu + 1) \tag{1.33}$$

e perció

$$\sigma^2 = E(X^2) - (E(X))^2 = \mu (\mu + 1) - \mu^2 = \mu \tag{1.34}$$

Inoltre vale il seguente **limite: se μ é grande, ed esattamente se tende ad infinito, la distribuzione di Poisson si approssima alla distribuzione normale con variabile standardizzata:**

$$\frac{x - \mu}{\sqrt{\mu}}$$

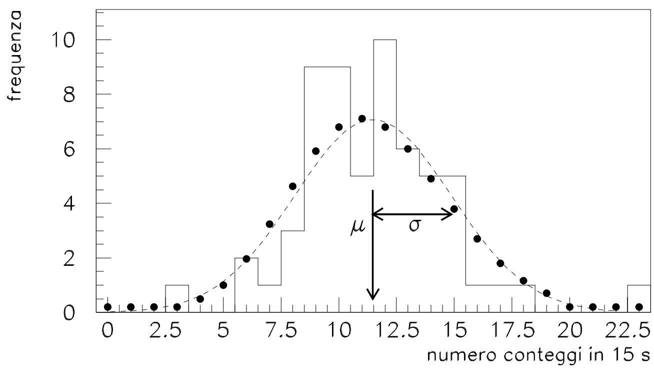

Figura 1.7:

L'approssimazione migliora al crescere di μ e, al limite, le due distribuzioni coincidono. Questa approssimazione, come pure quella valida per la distribuzione binomiale, sono una diretta conseguenza del **teorema limite centrale**. Si veda, a proposito, la figura 1.7, nella quale $\mu = 11.48$

Riassumendo:

la distribuzione poissoniana di probabilitá $P_p(x,\ \mu)$ fornisce la probabilitá di osservare x successi per unitá di tempo quando il numero medio di successi per unitá di tempo é μ. Si applica quando non sono noti con esattezza né il numero di possibili successi (n), che é supposto essere molto grande, né la probabilitá del singolo successo (p), che é supposta essere molto piccola, ma il loro prodotto é costante. É una distribuzione di probabilitá discreta, con le seguenti proprietá:

Somma	1
Media	E(x) $= \mu$
Varianza	$\sigma^2 = \mu$
Scarto quadratico medio	$\sigma = \sqrt{\mu}$

Riassumiamo anche le approssimazioni enunciate:

- **la distribuzione binomiale é ben approssimata dalla distribuzione di Poisson quando $n \rightarrow \infty$ e $p \rightarrow 0$ di modo che il prodotto np risulti costante (evento raro)**;
- **la distribuzione binomiale é ben approssimata dalla distribu-**

zione normale con variabile standardizzata $z = \frac{x - np}{\sqrt{npq}}$ **quando** n **cresce e se né** p **né** q **sono troppo prossimi allo zero;**

- **la distribuzione di Poisson é ben approssimata dalla distribuzione normale con variabile standardizzata** $z = \frac{x - \mu}{\sqrt{\mu}}$ **al crescere di** μ**.**

Vediamo alcuni esempi di applicazione della distribuzione di Poisson a casi diversi da esperimenti di conteggio:

- Il 10% degli utensili prodotti in un certo processo produttivo é stato trovato difettoso. Trovare la probabilitá che in un campione di 10 utensili scelti a caso, esattamente 2 siano difettosi ricorrendo *(a)* alla distribuzione binomiale, *(b)* all'approssimazione di Poisson alla distribuzione binomiale.

  (a) $p = 0.1$; $P_B(2,\ 10,\ 0.1) = \frac{10!}{2!\ 8!}\ 0.1^2\ 0.9^8 = 0.1937$

  (b) $\mu = np = 1$; $P_P(2,\ 1) = 1^2\ e^{-1}\ /\ 2! = e^{-1}\ /\ 2 = 0.1839$. In generale **l'approssimazione é buona se** $p \leq 0.1$ **e** $\mu = np \leq 5$.

- Se la probabilitá che un individuo sia allergico ad un certo vaccino é 0.001, determinare la probabilitá che, su 2000 individui, *(a)* esattamente 3, *(b)* piú di 2 siano allergici al vaccino stesso.

  Se la probabilitá che un individuo sia allergico al vaccino é p=0.001, il numero atteso di persone allergiche su un campione di n=2000 individui sará: $\mu = np = 2$; la probabilitá che x individui siano allergici é: $P_P(x,\ 2) = \frac{2^x\ e^{-2}}{x!}$.

  (a) $P_P(3,\ 2) = \frac{2^3\ e^{-2}}{3!} = 0.180$

  (b) $P_P(0,\ 2) = \frac{2^0\ e^{-2}}{0!} = \frac{1}{e^2}$, $P_P(1,\ 2) = \frac{2^1\ e^{-2}}{1!} = \frac{2}{e^2}$, $P_P(2,\ 2) = \frac{2^2\ e^{-2}}{2!} = \frac{2}{e^2}$; $P_P(> 2,\ 2) = 1 - P_P(0,\ 2) - P_P(1,\ 2) - P_P(2,\ 2) = 1 - (\frac{1}{e^2} + \frac{2}{e^2} + \frac{2}{e^2}) = 1 - \frac{5}{e^2} = 0.323$

  Si noti che, secondo la distribuzione binomiale le probabilitá richieste sono:

  (a) $\frac{2000!}{3!\ 1997!}\ 0.001^3\ 0.999^{1997}$

  (b) $1 - \frac{2000!}{0!\ 2000!}\ 0.001^0\ 0.999^{2000} - \frac{2000!}{1!\ 1999!}\ 0.001^1\ 0.999^{1999} -$

  $$\frac{2000!}{2!\ 1998!}\ 0.001^2\ 0.999^{1998}$$

  che sarebbero molto difficili da calcolare esattamente.

1.3.1 La legge del decadimento radioattivo

Si puó dimostrare che la distribuzione di probabilitá di eventi rari, tali cioé da avere probabilitá piccola di avverarsi, é data dalla distribuzione di Poisson, senza ricavare tale distribuzione come limite della distribuzione binomiale, bensì a partire dalle seguenti semplici condizioni:

1. gli eventi siano indipendenti: il verificarsi di uno o piú eventi in un certo intervallo di tempo Δt_i non dipende dal verificarsi di uno o piú eventi nell'intervallo di tempo Δt_j, per qualsiasi valore di i e j;

2. la probabilitá di avere un evento nell'intervallo di tempo Δt sia proporzionale alla durata dell'intervallo stesso: $p = \mu \, \Delta t$;

3. la probabilitá di avere piú di un evento dell'intervallo Δt sia un infinitesimo di ordine superiore rispetto a Δt.

Considerata la generalità delle tre posizioni sopra enunciate, è evidente come la distribuzione che si ricava a partire da esse sia particolarmente importante per la vastità delle sue applicazioni. Come esempio consideriamo il decadimento radioattivo di una sostanza che emetta particelle α; il numero di particelle che raggiunge un dato contatore nel tempo t è l'esempio più noto di evento casuale che soddisfa alla legge di Poisson. Naturalmente, nel tempo, la sostanza continua a decadere, fino a far diminuire la densità di particelle α emesse. Tuttavia, se la vita media della sostanza presa in esame è sufficientemente lunga le condizioni di decadimento possono ritenersi costanti, almeno per intervalli temporali non troppo lunghi.

In una famosa esperienza del 1920 si è osservato il decadimento radioattivo di una sostanza per 2608 intervalli temporali di 7.5 s ciascuno e, per ogni periodo, sono state contate le particelle che raggiungevano un certo contatore. In totale sono state rivelate 10094 particelle α, per cui il numero medio di particelle rivelate in ogni intervallino di 7.5 s è stato 3.870.

Possiamo dire di stare esaminando un esperimento casuale ripetuto 2608 volte, nel quale definiamo una variabile aleatoria (che assume i valori k=0, 1, 2...) che rappresenta il numero di decadimenti osservati in un intervallo temporale unitario (7.5 s); vogliamo verificare se tale variabile ha distribuzione di probabilità poissoniana.

In altre parole, ci chiediamo come valutare il parametro $\lambda = \mu t$ affinchè il confronto tra i valori sperimentali e quelli calcolati in base alla distribuzione di Poisson sia soddisfacente. A questo scopo, notiamo che se n_k è il numero di intervalli temporali in cui si sono verificati k decadimenti, si ha:

$$n = n_0 + n_1 + n_2 + ... = 2608$$

inoltre, il numero totale di decadimenti osservati è:

$$r = n_1 + 2n_2 + 3n_3 + ... + kn_k + ... = 10094$$

Il rapporto n_k/n non è altro che la frequenza relativa dell'evento k, intendendo per evento k quello in cui ci sono k decadimenti in un intervallo temporale unitario.

Per la legge dei grandi numeri la frequenza relativa tende alla probabilità quando il numero di prove è molto grande; possiamo quindi porre: $\frac{n_k}{n} = P_P(k, \mu t)$, da cui otteniamo $n_k = nP_P(k, \mu t)$ ed anche:

$$r = n[P_P(1, \mu t) + 2P_P(2, \mu t) + 3P_P(3, \mu t) + ...] =$$

$$= n\mu t e^{-\mu t}\left[1 + \mu t + \frac{(\mu t)^2}{2} + ...\right] = n\mu t$$

In conclusione si ha $\lambda = \mu t = r/n$ e la distribuzione di Poisson è, in questo caso:

$$P_P(k, \lambda) = P_P(k, r/n) = P_P(k, 3.870)$$

e rappresenta la probabilità di avere k decadimenti radioattivi in un intervallo temporale unitario. Questa probabilità, moltiplicata per il numero totale di intervalli temporali n, rappresenta la distribuzione di intervalli temporali in funzione di k in quanto si passa dalla normalizzazione ad 1 (distribuzione di probabilità) alla normalizzazione ad n (distribuzione di intervalli temporali).

Possiamo perciò confrontare il numero di intervalli temporali che presentano k decadimenti, quale misurato nel corso dell'esperienza, con l'analoga quantità calcolata secondo la distribuzione di Poisson. Tale confronto è riportato in tabella.

k	n_k	$nP_P(k, 3.870)$
0	57	54.399
1	203	210.523
2	383	407.361
3	525	525.496
4	532	508.418
5	408	393.515
6	273	253.817
7	139	140.325
8	45	67.882
9	27	29.189
$\geq$10	16	17.075
Totale	2608	2608.00