MATEMATICA SPERIMENTALE

Note per un Laboratorio del Progetto Lauree Scientifiche

Pierluigi Contucci

Anno Accademico 2006-2007, Università di Bologna

Introduzione

Queste brevi note (in forma ancora non definitiva) sono state scritte come guida schematica al laboratorio di matematica sperimentale che fa parte del Progetto Lauree Scientifiche dell'università di Bologna.

Lo scopo del nostro laboratorio è quello di avvicinare gli studenti alla matematica e in particolare a un suo settore: il calcolo della probabilità. Questa parte della matematica come forse poche altre unisce una grande versatilità nelle applicazioni e quindi nel mondo del lavoro a una struttura fortemente contro-intuitiva e difficile da conciliare col senso comune (si vedano gli esempi sul libro di M. Piatelli Palmarini [1]).

Usiamo il termine sperimentale *in riferimento all'approccio metodologico che seguiremo. In fisica la teoria viene affiancata dagli esperimenti in cui vengono verificate leggi esistenti o ne vengono cercate di nuove. Il nostro progetto è quello di utilizzare il calcolatore come un laboratorio dove testare teoremi e cercarne di nuovi. (si vedano gli interventi di W. Thurston [2, 3] e di Epstein-Levi [4]).*

L'uso del calcolatore per il calcolo scientifico nonostante concettualmente semplice richiede un cospicuo investimento di tempo. Per avvicinare gli studenti alla matematica e in particolare alla probabilità abbiamo scelto la strategia degli applets *che al contrario possono essere usati senza alcun prerequisito. Essi sono programmi preconfezionati che appaiono come pagine internet (scritte in html e contenenti Java). Dopo essere stato introdotto all'argomento in una lezione classica lo studente esegue in laboratorio esperimenti interagendo con la macchina attraverso un browser (explorer, netscape,* *etc) senza la necessità di alcuna preparazione informatica. L'esperimento si effettua letteralmente con un* click *che avvia l' applet. In pochi istanti per esempio si potrà simulare il lancio di milioni di freccette su uno schermo circolare e approssimare pi greco sfruttando la legge dei grandi numeri (metodo Monte Carlo) oppure lanciare centinaia di monete ripetendo il lancio un milione di volte e osservare la statistica delle teste e delle croci per approssimare la funzione Gaussiana (teorema del limite centrale).*

Desidero ringraziare la Dott. Nicoletta Bruno per la realizzazione degli applets utilizzati.

Desidero ringraziare la Prof. Emanuela Caliceti per il continuo supporto e incoraggiamento a questo laboratorio e per la generosa fiducia che mi ha accordato.

email: contucci@dm.unibo.it
homepage: http://www.dm.unibo.it/~contucci
Bologna, Ottobre 2006, (prima stesura maggio 2006).

Sperimentare

PROBABILITÀ E INTUITO

È opinione diffusa tra gli studiosi di probabilità che tale disciplina sia contro intuitiva. Si consideri per esempio il gioco delle tre scatole. Il gestore mette un premio dentro una di esse (il giocatore non sa in quale) e invita il giocatore a sceglierne una. La scatola non viene aperta. Dopo tale scelta il gestore elimina dal gioco una delle due scatole rimanenti, quella (o una di quelle) che non contiene il premio. Rimangono nel tavolo due scatole chiuse di cui una scelta dal giocatore. Il gestore offre ora al giocatore la possibilità di cambiare scelta (scegliere l'altra scatola) o di confermare la scelta iniziale. Qual'e' la strategia migliore per trovare il premio? Conviene confermare la prima scelta o cambiarla? O forse le due strategie si equivalgono? Quando il problema viene posto a un gruppo di persone con un test esse rispondono in larga maggioranza che le due strategie sono equivalenti. Restringendo il test a gruppi speciali, quali persone con una elevata educazione matematica, la risposta non cambia apprezzabilmente se esse vengono invitate a rispondere in tempi brevissimi e senza l'ausilio di calcoli. Un breve ed elementare calcolo mostra che la probabilità di vincita è di due terzi per la strategia di cambiare scelta mentre è di un terzo per la strategia di confermare la propria scelta. Si rappresenti infatti lo spazio delle delle possibili scelte iniziali con (P,P,V) cioè perde, perde, vince. Dopo che il giocatore ha fatto la sua scelta iniziale il gestore elimina una P. Il giocatore è invitato quindi a decidere se cambiare la sua scelta o confermarla e dovrà scegliere su (P,V). Le possibile giocate sono quindi:

1. (P,P)
2. (P,V)
3. (P,P)
4. (P,V)
5. (V,P)
6. (V,V)

Le giocate con conferma di scelta sono la 1 la 3 e la 6. Con questa strategia si vede che il giocatore vince solo una volta (giocata 6) su tre. Le giocate con cambio di scelta sono la 2 la 4 e la 5. Le giocate vincenti in questo caso sono due (la 2 e la 4) su tre.

Un altro esempio è quello del gioco del lotto in cui i numeri ritardatari divengono oggetto di un elevatissimo numero di giocate, molto superiore al numero che si avrebbe senza tali ritardi. In questo caso è una erronea e vaga idea sulla legge dei grandi numeri a causare l'errore. In particolare la convinzione popolare è che la conoscenza delle uscite passate possa suggerire qualcosa sulle uscite future mentre invece nel gioco del lotto ogni uscita è indipendente dalle precedenti.

Nel corso di questo laboratorio ci concentreremo su esperimenti al calcolatore che ci aiutino a una corretta comprensione di due leggi fondamentali della probabilità:

1. la legge (o teorema) dei grandi numeri

2. il teorema del limite centrale

Legge dei grandi numeri

Consideriamo il lancio di una moneta simmetrica e ripetiamo il lancio un numero di volte N. Chiamiamo N_T il numero di volte in cui esce testa e N_C quello in cui esce croce. Chiaramente $N_T + N_C = N$. Il quoziente $\frac{N_T}{N}$ rappresenta la frequenza relativa delle teste. Ci aspettiamo che per un gran numero di lanci (N grande) tale frequenza tenda alla probabilità teorica cioè a un mezzo. Ma in che senso va inteso questo limite? Per capirlo useremo l'applet 1. Esso simula il lancio di un certo numero di monete e calcola le frequenze (in percentuale) di teste e croci. Eseguiamo i seguenti esperimenti:

- Iniziamo col lanciare una sola moneta (N=1) e ripetiamo l'esperimento più volte. In tal caso la frequenza in ciascun lancio sarà di 0 o 1 e ripetendo l'esperimento diverse volte vedremo comparire 0 e 1 in modo casuale. Notiamo che l'esperimento singolo è il lancio di *una* moneta e noi lo ripetiamo diverse volte.

- Lanciamo ora due monete. In ciascun lancio la frequenza prenderà i valori possibili di 0, 0.5 e 1. Ripetiamo l'esperimento diverse volte. I valori che otterremo oscilleranno tra le tre possibilità date ma ora notiamo che il valore 0.5 compare più spesso di zero e di uno (ripetendo il lancio 10 volte o 20 volte si conti la frequenza di ciascun esito possibile).

- Passiamo ora al lancio di 10 monete. Le possibili frequenze in ciascun lancio (un esperimento) sono i decimali a una cifra: 0, 0.1, 0.2... fino a 1. Ripetendo l'esperimento vedremo che la frequenza è spesso vicino al valore 0.5 e raramente vicino a 0 o 1 (ripetendo il lancio 10 volte o 20 volte si conti la frequenza di ciascun esito possibile).

Notiamo che stiamo parlando di frequenza a due diversi livelli: la frequenza $f_T(N)$ *dell'evento testa in un gruppo di* N *monete lanciate in un tavolo e la frequenza con cui, ripetendo* M *volte l'esperimento del lancio del gruppo di monete, il valore di* $f_T(N)$ *cade in un certo intervallo. Questi due livelli compaiono come livelli fondamentali da tenere ben distinti anche nella prova rigorosa della legge dei grandi numeri che riportiamo in appendice.*

- Procediamo ora al lancio di 100 monete e ripetiamo l'esperimento come fatto in precedenza 10 o 20 volte. Vedremo che i valori oscillano spesso tra 0.45 e 0.55 e difficilmente si esce da tale intervallo.

- Se poi si lanciano 1000 monete e si osservano gli esiti dei lanci ripetuti (sempre 10 o 20 volte) si osserva che tutti gli esperimenti cadono dentro all'intervallo a meno di fenomeni rarissimi.

Si usi l'applet 2 per fare i test con un numero maggiore al migliaio. Il funzionamento di questo applet è identico a quello precedente con la sola differenza che non scrive esplicitamente la stringa di teste e croci ma calcola solo la frequenza.

Dal punto di vista sperimentale quindi un modo per riassumere le osservazioni fatte (legge dei grandi numeri) è quello di dire che se fissiamo un intervallo intorno al valore 0.5 sarà possibile, aumentando il numero di monete lanciate in ciascun lancio (N) essere sicuri che prove successive di lancio producano, con grande sicurezza, una frequenza dentro l'intervallo fissato. Le nostre osservazioni sperimentali invece non ci dicono che è *impossibile* ottenere un esito al di fuori dell'intervallo dato per grandi numeri di monete, ma che la frequenza con cui tale violazione si presenta è molto bassa. In particolare i fenomeni osservati non permettono di fare nessun tipo di affermazione che leghi le proprietà del lancio successivo conosciuti gli esisti di un gran numero di lanci precedenti. Nelle appendici viene data una dimostrazione della legge dei grandi numeri utilizzando la teoria assiomatica della probabilità e la distribuzione binomiale.

Il metodo Monte Carlo

Come interessante applicazione della legge dei grandi numeri si può illustrare il metodo Monte Carlo e il suo uso per la determinazione di π. Si lanciano freccette contro uno schermo quadrato di lato 2 in cui è iscritto un cerchio di raggio uno. Se ciascuna freccetta è lanciata casualmente nel quadrato ci aspettiamo che il numero di freccette che cadono dentro il cerchio rispetto al numero di freccette totali tenda, nel senso della legge dei grandi numeri, a $\pi/4$. Utilizzando l'applet 3 si facciano esperimenti con numeri crescenti di lanci e si ottengano *stime* delle cifre successive del Pi Greco. Applet 3. Per coloro che hanno qualche nozione di calcolo integrale è utile calcolare l'area sottesa dalla parabola $y = x^2$ tra zero e uno. Il risultato esatto è $1/3$ Eseguirne stime con l'ausilio dell'applet 4.

Il teorema limite centrale

La serie di esperimenti che ci proponiamo di fare in questa sede vuole affrontare una questione più sottile della precedente che era sui valori medi di certe variabili aleatorie. Abbiamo visto che la variabile aleatoria frequenza $f_T(N)$ diventa *certa* quando N cresce. Ora invece ci chiediamo come sono distribuiti i diversi valori che essa può assumere e soprattutto come possiamo individuare la scala in cui confrontarli. Per fare ciò faremo uso di un applet che realizza l'istogramma sugli esiti di lancio di N monete ripetuti M volte. Per semplicità associeremo alla testa il numero 1 e alla croce il numero -1 e faremo esperimenti che riguardano la somma degli esiti su ciascun lancio ripetuta tante volte. Applet 5.

- Si scelga inizialmente $N = 1$ e $M = 1$, una moneta lanciata una volta. L'istogramma che osserviamo ci da la stessa informazione che avevamo nell'applet precedente. Ripetiamo l'esperimento alcune volte per rendercene conto.

- Passiamo ora al caso $N = 2$ e $M = 2$ e ripetiamo l'esperimento più volte. Vedremo che il suo aspetto cambia casualmente tra quello a due barre a quello a una barra sola.

- Sempre con $N = 2$ proviamo i valori $M = 10, 100$ e 1000. Vedremo che l'aspetto dell'istogramma tende a stabilizzarsi con una barra alta al centro e due piccole e simili ai lati. Per esempio il caso $N = 2$ e $M = 1000$ ha la barra centrale che oscilla di poco intorno il 500 e due barre laterali che oscillano intorno al 250.

- Procediamo ora al caso $N = 10$ e $M = 10$. L'istogramma al ripetersi dei lanci si presenta piuttosto erratico e non mostra una precisa regolarità.

- Tenendo fisso $N = 10$ aumentiamo M al valore 100. Notiamo che le barre centrali sono tendenzialmente più alte e le laterali descrescono ma il profilo al ripetersi dei lanci ha ancora una forte instabilità.

- Per vedere stabilizzarsi la figura facciamo il test (sempre con $N = 10$) coi valori successivi $M = 1000$, $M = 10000$, $M = 100000$ e persino $M = 1000000$. Con un milione di lanci di 10 monete l'istogramma diventa praticamente stabile e osserviamo la barra centrale dello zero vicina ai 250000 le due barre successive, quelle del 2, vicine ai 205.000, quelle del 4 intorno al 120000, quelle del 6 intorno al 45.000 quelle del 8 intorno al 10000 e infine quelle del 10 pressochè invisibili dell'ordine dei 1000.

- Infine, testiamo il lancio di gruppi di 100 monete e aumentiamo progressivamente il numero di esperimenti M. Il nostro applet arriva persino a fare un milione di lanci (in 40 secondi di tempo). Quello che ci troviamo di fronte è lo spettacolare profilo della funzione Gaussiana. Con il nostro set di parametri la funzione raggiunge il suo massimo in corrispondenza della barra dello zero all'incirca a una altezza di 80000 e si schiaccia velocemente verso il valore ± 40 in orizzontale.

Possiamo ora fare una stima sull'ordine di grandezza della larghezza dei nostri istogrammi. Le curve a campana che abbiamo visto comparire, con approssimazione crescente col crescere del numero di monete lanciate, mostrano un cambio di convessità per certi valori σ_N *dell'asse orizzontale. Analizzando* *l'istogramma (fissiamo ora* $M = 1000000$*) per* $N = 10$ *vediamo che il cambio di convessità avviene intorno al valore* 3*. Per* $N = 20$ *attorno al valore* 4.5*. Se riportiamo i valori della funzione* σ_N *in un grafico in termini di* N *vediamo che essa si fa ben approssimare da* $\sqrt{N}$*. Questa preziosa indicazione ci suggerisce che sarà la variabile aleatoria*

$$\frac{\sum_{i=1}^{N} x_i}{\sqrt{N}} \tag{1}$$

ad avere una ben definita distribuzione di probabilità (in particolare Gaussiana) al tendere di N *all'infinito. Questa conclusione infatti può essere rigorosamente dimostrata. Nell'appendice riportiamo una prova per il caso di variabili di Bernoulli.*

Appendici

Riportiamo qui come ausilio al lettore alcune semplici dimostrazioni di teoria della probabilità invitandolo anche a consultare testi introduttivi per le scuole superiori come il [5] o di livello di un corso di ingresso universitario come il [6].

CONTARE

Consideriamo una macchina da scrivere che stampa 3 lettere, a, b e c. Quante parole diverse di 2 lettere si possono scrivere? Elencondole avremo aa, ab, ac, ba, bb, bc, ca, cb e cc. Cioè 9 parole. Invece che elencarle e successivamente contarle possiamo contarle direttamente osservando che la prima lettera può essere scelta tra 3 possibilità, la seconda tra 3 possibilità quindi il numero totale sarà di $3 \cdot 3 = 9$. In generale con una macchina da scrivere che stampa n lettere diverse possiamo scrivere

$$n \cdot n \cdots n = n^k \tag{2}$$

parole distinte lunghe k lettere.

Consideriamo ora invece un problema diverso. Invece di una macchina da scrivere abbiamo un sacchetto che contiene 3 lettere a, b e c. Quante parole di 2 lettere si possiano scrivere? In questo caso ciascuna lettera potrà essere utilizzata una sola volta e l'elenco fornisce: ab, ac, ba, bc, ca e cb. Cioè 6 possibilità. In generale con un sacchetto che contiene n lettere distinte possiamo scrivere

$$n \cdot (n-1) \cdot (n-2) \cdots (n-k+1) \tag{3}$$

parole diverse lunghe k dato che la prima lettera potrà essere scelta tra n possibilità, la seconda tra $n-1$ etc. Introducendo il simbolo di *fattoriale*

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1 \tag{4}$$

che conta il numero di parole lunghe n che si possono scrivere con un sacchetto con n lettere distinte il numero precedente si può esprimere come:

$$n \cdot (n-1) \cdot (n-2) \cdots (n-k+1) = \frac{n!}{(n-k)!} \tag{5}$$

Consideriamo infine un sacchetto che contiene due a e una b. Quante parole diverse possiamo formare usando tutte le tre lettere? Avremo *aab*, *aba*, *baa*. In generale se il sacchetto contiene n lettere di cui k a e $(n-k)$ b possiamo procedere al conteggio in questo modo: dapprima fingiamo che le lettere siano tutte distinte, in tal caso si avrebbero $n!$ parole diverse. Cosi facendo tuttavia abbiamo contato parole uguali come se fossero distinte. Dobbiamo quindi dividere il risultato per i modi in cui si permutano tra loro le lettere coincidenti di tipo a cioè $k!$ e di tipo b cioè $(n-k)!$. Avremo quindi che il numero di parole distinte è il coefficiente binomiale:

$$\frac{n!}{k!(n-k)!} \tag{6}$$

Il nome di tale coefficiente deriva dal fatto che esso compare nello sviluppo del binomio di Newton:

$$(a+b)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} a^k b^{n-k} \tag{7}$$

Lanciando una moneta simmetrica molte volte ci aspettiamo di ottenere mediamente un numero di teste o di croci non molto diverso dalla metà dei lanci. Questa affermazione che il senso comune ci suggerisce innocuamente abbisogna tuttavia di un certo lavoro per trasformarsi in un teorema, lavoro che storicamente ha coperto un lasso di tempo di oltre due secoli.

Consideriamo un insieme di indici $i \in I$ e ad essi associamo la collezione di numeri $p_i \in P$ che verifica le proprietà

$$0 \leq p_i \leq 1 \tag{8}$$

$$\sum_i p_i = 1 \tag{9}$$

Ciascun indice i identifica un *evento elementare* e p_i è la sua *probabilità*. La coppia (I, P) è uno spazio di *probabilità discreto* e le relazioni (8) e (9) sono il primo e secondo assioma.

Esempio 1: la moneta simmetrica ha spazio di probabilità identificato da

$$T \equiv testa\ , \tag{10}$$

$$C \equiv croce\ , \tag{11}$$

quindi

$$I = (T, C)\ ; \tag{12}$$

e dalle relative probabilit'a

$$p_T = \frac{1}{2}\ , \tag{13}$$

$$p_C = \frac{1}{2}\ . \tag{14}$$

$$P \equiv \left(\frac{1}{2}, \frac{1}{2}\right) . \tag{15}$$

Per esercizio si verfichi che i due assiomi sono soddisfatti.

Esempio 2: la moneta asimmetrica (o truccata) ha spazio di probabilità identificato dallo stesso spazio di indici

$$I \equiv (T, C) , \tag{16}$$

ma da un diverso valore delle probabilità degli eventi

$$p_T = p , \tag{17}$$

con

$$0 \leq p \leq 1 , \tag{18}$$

e da

$$p_C = 1 - p , \tag{19}$$

quindi

$$P \equiv (p, 1 - p) . \tag{20}$$

Per esercizio si verfichi che anche in questo caso che generalizza quello simmetrico il primo e secondo assioma sono soddisfatti. Si diano inoltre alcuni esempi numerici di p.

Esempio 3: il dado regolare a sei facce ha spazio di probabilità identificato da

$$I \equiv (1, 2, 3, 4, 5, 6) , \tag{21}$$

e da

$$P \equiv (\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}) . \tag{22}$$

Un possibile esempio di dado truccato è quello in cui l'evento 6 da solo ha probabilità un mezzo e gli altri eventi $1, 2, 3, 4, 5$ hanno la probabilità p: in tal caso l'assioma (9) impone

$$5p + \frac{1}{2} = 1 \,, \tag{23}$$

e quindi $p = \frac{1}{10}$. Si diano altri esempi di spazi di proabilità di un dado truccato.

È di grande importanza considerare il lancio della moneta (o del dado) ripetuto più volte in modo indipendente. Consideriamo dapprima il lancio della moneta simmetrica ripetuto due volte e costruiamo il suo spazio delle probabilità. Avremo che gli eventi elementari saranno

$$I \quad \equiv \quad (TT, TC, CT, CC) \,, \tag{24}$$

e le probabilità

$$P \quad \equiv \quad (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}) \,. \tag{25}$$

È utile nell' esempio appena dato analizzare alcuni eventi che non siano eventi elementari: l' evento A definito dall' uscita di croce al primo lancio potrà realizzarsi o come CT o come CC e utilizzando il simbolo di unione insiemistica (elementi che stanno o nel primo o nel secondo insieme) scriveremo

$$A \,=\, CT \,\cup\, CC \;; \tag{26}$$

talvolta si utilizza per esso anche il simbolo

$$CX = CT \,\cup\, CC \,, \tag{27}$$

e la notazione significa che siamo nello spazio di probabilità di due lanci in cui il primo è croce e il secondo non è specificato. Similmente l' evento TXC

indica l'uscita, nello spazio dei tre lanci di testa al primo e croce al terzo e vale:

$$TXC = TTC \cup TCC \ . \tag{28}$$

Gli esempi precedenti illustrano che l'unione di eventi è essa stessa un evento. La stessa proprietà vale per l'intersezione (elementi che stanno sia nel primo che nel secondo insieme). Per esempio se consideriamo

$$TX \cap XC \ , \tag{29}$$

ci riferiamo all'evento nello spazio dei 2 lanci che ha testa al primo e croce al secondo quindi

$$TX \cap XC = TC \ , \tag{30}$$

In generale non c'è nessuna relazione tra tra le probabilità di certi eventi e quelle della loro intersezione. Tuttavia quando per due eventi A e B vale la

$$P(A \cap B) \ = \ P(A)P(B) \ , \tag{31}$$

i due eventi si dicono *indipendenti*. Due eventi A e B si dicono mutuamente esclusivi (o disgiunti) quando la loro intersezione è l' insieme vuoto:

$$A \cap B = \emptyset, \tag{32}$$

per essi vale

$$P(A \cup B) = P(A) + P(B) \ . \tag{33}$$

Più in generale quando la loro intersezione è non vuota l'assioma si generalizza in

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \, . \tag{34}$$

Verifichiamo che i due eventi TX e XC sono indipendenti.

$$P(TX) = P(TT \cup TC) = P(TT) + P(TC) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}\,, \tag{35}$$

dove la prima uguaglianza viene dall'assioma 2 dato che $TT \cap TC = \emptyset$. Analogamente si ha $P(XC) = 1/2$ e quindi poichè $TX \cap XC = TC$ e $P(TC) = 1/4$ la relazione di indipendenza è verificata.

Problema. Dati due eventi indipendenti A e B con $P(A) = 1/2$ e $P(B) = 1/3$ calcoliamo la probabilità della loro unione. Osserviamo anzitutto che senza l'informazione di indipendenza il problema sarebbe indeterminato dato che manca il valore di $P(A \cap B)$. Grazie all'indipendenza tuttavia sappiamo che $P(A \cap B) = P(A)P(B)$ quindi per la (34) abbiamo:

$$P(A \cup B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{2} \cdot \frac{1}{3} = \frac{2}{3} \tag{36}$$

Esercizio. Si dica quali delle seguenti coppie di eventi sono indipendenti: (TX, XT), (CX, CX), (XT, CX), (CT, CC), (TC, XT).

Problema. Consideriamo il lancio di un dado ripetuto 3 volte, l?evento A in cui esce il 2 al secondo lancio e l?evento B in cui esce il 6 al primo lancio e il 4 al terzo. Ci proponiamo di vedere quali sono gli eventi elementari che compongono A e quelli che compongono B, di calcolare le relative probabilità e infine di determinare se A e B sono eventi indipendenti. Dalla definizione di A si ha:

$$A = X2X\,, \tag{37}$$

quindi A à formato dall? unione dei $6 \cdot 6$ eventi ottenibili inserendo una delle cifre del dado al posto delle X, come per esempio $A_1 = 121$ (da leggere uno-due-uno e non centoventuno!) oppure $A_2 = 122$, etc. Analogamente avremo

$$B = 6X4\,, \tag{38}$$

quindi B sarà formato dai 6 eventi elementari che si ottengono inserendo una delle sei cifre del dado al posto della X, quali $B_1 = 614$, $B_2 = 624$, etc. Per il calcolo di P(A) la regola di somma fornisce:

$$P(A) = P(\cup_{i=1}^{36} A_i) = \sum_{i=1}^{36} P(A_i) = 36P(A_1) = 36\frac{1}{216} = \frac{1}{6} \tag{39}$$

e analogamente per il calcolo di P(B)

$$P(B) = P(\cup_{i=1}^{6} B_i) = \sum_{i=1}^{6} P(B_i) = 6P(B_1) = 6\frac{1}{216} = \frac{1}{36}\,. \tag{40}$$

Infine per stabilire se i due eventi A e B sono indipendenti notiamo anzitutto che l'evento intersezione $A \cap B = 624$ quindi $P(A \cap B) = \frac{1}{6^3} = \frac{1}{216}$ e $P(A)P(B) = \frac{1}{6}\frac{1}{36} = \frac{1}{216}$, quindi $P(A \cap B) = P(A)P(B)$ cioè gli eventi A e B sono indipendenti.

Sommario:

- Spazio di probabilità:

$$I \equiv (1, 2, 3 \cdots, n)$$

$$P \equiv (p_1, p_2, p_3 \cdots p_n)$$

$$\begin{aligned}
\mathcal{A}_0 \quad &: \qquad 0 \le p_i \le 1 \\
\mathcal{A}_1 \quad &: \qquad \sum_i p_i = 1 \\
\mathcal{A}_2 \quad &: P(A \cup B) = P(A) + P(B)\,, \quad A \cap B = \emptyset
\end{aligned}$$

- Probabilità condizionata:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{41}$$

- Eventi indipendenti:

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \quad \text{oppure} \\ P(A|B) &= P(A) \end{aligned}$$

VARIABILI ALEATORIE E LORO MEDIE

Una funzione che associa un numero a ciascun evento elementare di uno spazio di probabilità è chiamata *variabile aleatoria.* Per esempio la funzione che vale 1 in corrispondenza della testa della moneta e -1 in corrispondenza della croce:

$$\begin{aligned} I &= (T, C) \\ P &= (p, 1-p) \\ f &= (1, -1) \ . \end{aligned} \tag{42}$$

In generale avremo

$$\begin{aligned} I &= (1, 2, \cdots n) \\ P &= (p_1, p_2, \cdots p_n) \\ f &= (f_1, f_2, \cdots f_n) \ . \end{aligned}$$

Si definisce *valor medio* della funzione f la quantità

$$E(f) = \sum_{i=1}^{n} f_i p_i \ . \tag{43}$$

Notiamo che il valore medio di una combinazione lineare di due funzioni è uguale alla combinazione lineare dei valori medi:

$$E(\alpha f+\beta g) \;=\; \sum_{i=1}^{n}(\alpha f_i+\beta g_i)p_i \;=\; \alpha\sum_{i=1}^{n} f_i p_i+\beta\sum_{i=1}^{n} g_i p_i \;=\; \alpha E(f)+\beta E(g)\ . \tag{44}$$

Inoltre se una variabile aleatoria è non negativa anche la sua media è non negativa:

$$f_i \;\geq\; 0 \qquad \textit{implica} \qquad E(f)\geq 0\ .$$

Infine si osserva che la variabile aleatoria costante unitaria ha valor medio unitario:

$$f_i \;=\; 1 \qquad \textit{implica} \qquad E(f)=1\ .$$

Si definisce varianza di una variabile aleatoria la quantità

$$\begin{aligned} V(f) &= E([f-E(f)]^2) \;=\; E\left(f^2-2fE(f)+E(f)^2\right) \\ &= E(f^2)-E(f)^2 \;=\; \sum_{i=1}^{n} f_i^2 p_i-(\sum_{i=1}^{n} f_i p_i)^2\,, \end{aligned} \tag{45}$$

dove abbiamo usato le proprietà della media. Notiamo che per definizione si ha sempre $V(f)\geq 0$ e quindi $E(f^2)\;\geq\;E(f)^2$. Vale inoltre $V(\alpha f)\;=\;\alpha^2 V(f)$.

Problema. calcolare media e varianza della variabile aleatoria definita dalle (42) nello spazio della moneta asimmetrica:

$$E(f)=f_T p_T+f_C p_C=p-(1-p)=2p-1\,, \tag{46}$$

$$E(f^2)=f_T^2 p_T+f_C^2 p_C=p+(1-p)=1\,, \tag{47}$$

$$V(f) = E(f^2) - E(f)^2 = 1 - (2p-1)^2 = 1 - 4p^2 + 4p - 1 = 4p(1-p) \; . \quad (48)$$

Problema. Consideriamo un dado asimmetrico in cui la probabilità del 6 è di un mezzo mentre tutte le altre facce hanno uguale probabilità 1/10 (vedi la 23). Vogliamo calcolare nel lancio del dado ripetuto tre volte la probabilità degli eventi $X2X$ e $6X4$ e della loro intersezione 624; infine vogliamo calcolare media della variabile aleatoria $f_i = i$ e confrontarla col caso del dado simmetrico. Dal calcolo diretto per somma degli eventi elementari si ha

$$P(X2X) = \frac{1}{10} \, , \quad (49)$$

$$P(6X4) = \frac{1}{20} \, , \quad (50)$$

$$P(624) = \frac{1}{200} \, . \quad (51)$$

Notiamo quindi che l?indipendenza tra gli eventi A e B vale anche nel caso del dado asimmetrico. Infine abbiamo:

$$E(f) = \sum_{i=1}^{6} ip_i \; = \; \frac{1}{10}(1+2+3+4+5) + \frac{1}{2}6 = 4 \, , \quad (52)$$

mentre per il dado simmetrico vale

$$E(f) = \sum_{i=1}^{6} ip_i \; = \; \frac{1}{6}(1+2+3+4+5+6) = 3.5 \quad (53)$$

DISTRIBUZIONE BINOMIALE

Vogliamo ora concentrarci sul caso della moneta asimmetrica e considerarne il lancio un numero arbitrario di volte n. In particolare vogliamo stabilire qual? è la probabilità che in n lanci si abbiano k teste. Il problema si riconduce ad uno di tipo combinatorio del contare quante parole diverse possiamo formare con k lettere di tipo T e $n-k$ lettere di tipo C. Per cominciare facciamo il caso con $k=2$ e $n=3$. Le possibilità sono TTC, TCT, CTT ed ognuna di esse ha (per l'indipendenza dei lanci) probabilità:

$$p^2(1-p)\,. \tag{54}$$

In generale il numero di parole sarà dato dal fattore binomiale

$$\frac{n!}{k!(n-k)!}\,, \tag{55}$$

e la relativa probabilità di ciascuno degli eventi

$$p^k(1-p)^{(n-k)}\,. \tag{56}$$

Pertanto utilizzando la proprietà (33) otteniamo

$$P_n(k) = \frac{n!}{k!(n-k)!}p^k(1-p)^{(n-k)}\,. \tag{57}$$

La formula sopra scritta definisce la distribuzione binomiale della variabile aleatoria f = numero di teste in n lanci. Notiamo infatti che la somma delle probabilità totale risulta

$$\sum_{k=0}^{n} \frac{n!}{k!(n-k)!}p^k(1-p)^{(n-k)} = (p+1-p)^n = 1, \tag{58}$$

dove si è utilizzata la formula del binomio di Newton.

Calcoliamo media e varianza di f.

$$\begin{aligned} E(f) &= \sum_{k=0}^{n} k\frac{n!}{k!(n-k)!}p^k(1-p)^{(n-k)} \qquad (59)\\ &= \sum_{k=1}^{n} k\frac{n!}{k!(n-k)!}p^k(1-p)^{(n-k)} \\ &= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!}p^k(1-p)^{(n-k)} \end{aligned}$$

dove si è osservato che il primo termine della somma non contribuisce essendo zero e si è semplificato un fattore k dai rimanenti. Ora poniamo $l = k - 1$ e otteniamo

$$\begin{aligned} E(f) &= \sum_{l=0}^{n-1} \frac{n!}{l!(n-l-1)!} p^{l+1} (1-p)^{(n-l-1)} \qquad (60) \\ &= np \sum_{l=0}^{n-1} \frac{(n-1)!}{l!(n-1-l)!} p^{l} (1-p)^{(n-1-l)} \\ &= np \sum_{l=0}^{n-1} P_{n-1}(l) = np\,. \end{aligned}$$

Per il calcolo della varianza si procede in modo analogo

$$\begin{aligned} E(f^2) &= \sum_{k=0}^{n} k^2 \frac{n!}{k!(n-k)!} p^{k} (1-p)^{(n-k)} \qquad (61) \\ &= \sum_{k=1}^{n} k^2 \frac{n!}{k!(n-k)!} p^{k} (1-p)^{(n-k)} \\ &= \sum_{k=1}^{n} k \frac{n!}{(k-1)!(n-k)!} p^{k} (1-p)^{(n-k)} \end{aligned}$$

Ora poniamo $l = k - 1$ e otteniamo

$$\begin{aligned} E(f^2) &= \sum_{l=0}^{n-1} (l+1) \frac{n!}{l!(n-1-l)!} p^{l+1} (1-p)^{(n-1-l)} \qquad (62) \\ &= np \sum_{l=0}^{n-1} l P_{n-1}(l) + np \sum_{l=0}^{n-1} P_{n-1}(l) \\ &= np(n-1)p + np = n^2p^2 - np^2 + np\,. \end{aligned}$$

da cui

$$V(f) = np(1-p)\,. \qquad (63)$$

DISUGUAGLIANZA DI CHEBYSHEV E LEGGE DEI GRANDI NUMERI

In questa paragrafo ci proponiamo di dare un significato preciso all' affermazione che la frequenza empirica tende alla probabilità teorica in un grande numero di lanci della moneta. Partiamo da due semplici disuguaglianze. La prima dice che per due variabili aleatorie tali che $f_i \leq g_i$ si ha

$$\sum_i f_i p_i \leq \sum_i g_i p_i \ . \tag{64}$$

La seconda che per una coppia di eventi in cui uno contiene l'altro $A \subset B$ si ha

$$P(A) \leq P(B) \qquad \textit{equivalentemente} \qquad \sum_{i \in A} p_i \leq \sum_{i \in B} p_i \ , \tag{65}$$

che deriva dall'assioma di positività delle p_i (la seconda somma è più grande perchè si sommano più termini positivi). Ci proponiamo di stimare la quantità

$$P(|f - E(f)| \geq \delta) \ , \tag{66}$$

cioè la probabilità che una funzione scarti dalla propria media per più di una quantità assegnata δ. Per definizione avremo

$$P(|f - E(f)| \geq \delta) = \sum_{i, |f_i - E(f)| \geq \delta} p_i \leq \sum_{i, |f_i - E(f)| \geq \delta} \frac{(f_i - E(f))^2}{\delta^2} p_i \tag{67}$$

dove abbiamo utilizzato la (64). Inoltre per la (65) abbiamo

$$\sum_{i, |f_i - E(f)| \geq \delta} \frac{(f_i - E(f))^2}{\delta^2} p_i \leq \sum_i \frac{(f_i - E(f))^2}{\delta^2} p_i = \frac{1}{\delta^2} V(f) \ . \tag{68}$$

Definendo la *deviazione standard* come $\sigma(f) = \sqrt{V(f)}$ la precedente ha una interpretazione particolarmente efficace quando si sceglie $\delta = l\sigma(f)$. Si ha infatti (disuguaglianza di Chebyshev)

$$P(|f - E(f)| \geq l\sigma(f)) \leq \frac{1}{l^2} \tag{69}$$

cioè la probabilità che una variabile aleatoria scarti dal suo valor medio per pi di 2 sigma è minore di 1/4 , per più di 3 sigma è minore di 1/9 e cosi via.

Una importante applicazione della precedente si ha per la variabile aleatoria binomiale $f_k = k/n$. Anzitutto osserviamo da un calcolo diretto che

$$E(\frac{k}{n}) = p \qquad V(\frac{k}{n}) = \frac{p(1-p)}{n} \ , \tag{70}$$

da cui la (68) fornisce una stima della probabilità che la frequenza $\frac{k}{n}$ delle teste si discosti dalla probabilità p per una quantità piccola

$$P(|\frac{k}{n} - p| \geq \varepsilon) \leq \frac{p(1-p)}{\varepsilon^2 n} \ , \tag{71}$$

quantità questa che tende a zero all' aumentare del numero di lanci. La formula sopra scritta (legge dei grandi numeri) merita alcuni commenti. Essa fa un' affermazione che vale nello spazio di probabilità degli n-lanci e non del singolo lancio! I singoli lanci sono indipendenti e i lanci passati non possono in alcun modo influenzare quelli futuri.

Sommario:

- Media

$$E(f) \ = \sum_{i=1}^{n} f_i p_i \ ,$$

- Varianza

$$V(f) \;=\; E([f-E(f)]^2) \;=\; E(f^2)-E(f)^2 \;=\; \sum_{i=1}^{n} f_i^2 p_i - (\sum_{i=1}^{n} f_i p_i)^2 \,,$$

- Distribuzione Binomiale

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{(n-k)}$$

- Distribuzione di Poisson

$$P_\lambda(k) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{72}$$

- Disuguaglianza di Chebyshev

$$P(|f-E(f)| \geq l\sigma(f)) \leq \frac{1}{l^2} \tag{73}$$

- Legge dei grandi numeri

$$P(|\frac{k}{n} - p| \geq \varepsilon) \leq \frac{p(1-p)}{\varepsilon^2 n} \tag{74}$$

DISTRIBUZIONI CONTINUE

In una distribuzione di probabilità continua gli eventi non sono numerabili (non possono essere messi in corrispondenza biunivoca con gli interi). Un esempio è la distribuzione di probabilità uniforme nell?intervallo $[0,1]$. Si dice che una variabile aleatoria ξ ha distribuzione uniforme in tale intervallo se per ogni $0 \leq a < b \leq 1$ si ha

$$P(a \leq \xi \leq b) = \int_a^b dx = b - a \,. \tag{75}$$

Più in generale si dice che una variabile aleatoria ζ ha *densità di probabilità* $p(x)$ nell?intervallo $[0,1]$ se vale la condizione di positività

$$p(x) \geq 0 \ , \tag{76}$$

e la condizione di normalizzazione

$$\int_0^1 p(x)dx = 1 \ . \tag{77}$$

Esercizi:

- Si dica quali tra le seguenti funzioni sono distribuzioni di probabilità nell'intervallo $[0,1]$: $2x$, $sen(x)$, $1-x$.
- Si diano esempi che violano la positività, che violino la normalizzazione o che violino entrambe.
- Si calcoli il valore della costante c che rende le seguenti funzioni delle distribuzioni di probabilità in $[0,1]$: cx^3, $c(x^2-1)$

Notiamo che la distribuzione uniforme è un caso particolare del precedente definito da $p(x) = 1.$.

Per le distribuzioni continue valgono le stesse definizioni di media e varianza date per quelle discrete in cui si utitlizza l'integrale al posto delle somme. Data la funzione f dell'intervallo $[0,1]$ (variabile aleatoria) si definisce

$$E(f) = \int_0^1 f(x)p(x)dx \tag{78}$$

$$E(f^2) = \int_0^1 f(x)^2 p(x)dx \ , \tag{79}$$

$$V(f) = E([\zeta^2 - E(\zeta)]^2) \tag{80}$$

Ovviamente le definizioni date per variabili nell' intervallo unitario si estendono a qualsiasi intervallo finito, e vedremo anche alla retta reale.

Esercizi:

- La variabile aleatoria $f(x) = x$ ha densità di probabilità $p(x) = 1 - x/2$ nell?intervallo $[0, 2]$. Si verifichi che la $p(x)$ data è una funzione densità di probabilità e si calcoli la media e la varianza di f.
- La variabile aleatoria $f(x) = x$ ha densità di probabilità cx^4 nell?intervallo [0,1] Si calcoli il valore di c e la media e la varianza di f.
- La variabile aleatoria $f(x) = 3x$ ha distribuzione di probabilità costante nell?intervallo $[0, 3]$. Si calcoli il valore della costante, la media e la varianza di f.

LA GAUSSIANA

La distribuzione di Gauss è definita (nell?intervallo reale) dalla densità di probabilità

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \,. \tag{81}$$

Per tale distribuzione valgono le seguenti formule

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = 1 \tag{82}$$

$$\int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = 0 \tag{83}$$

$$\int_{-\infty}^{\infty} \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = 1 \,, \tag{84}$$

da cui si può dedurre che per la Gaussiana di media μ e varianza σ^2 di equazione

$$p_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{85}$$

valgono le

$$\int_{-\infty}^{\infty} \quad p_{\mu,\sigma}(x) = 1 \tag{86}$$

$$\int_{-\infty}^{\infty} \quad x p_{\mu,\sigma}(x) = \mu \tag{87}$$

$$\int_{-\infty}^{\infty} \quad (x-\mu)^2 p_{\mu,\sigma}(x) = \sigma^2 \ . \tag{88}$$

È interessante vedere come si trasforma la curva $p_{\mu,\sigma}(x)$ al varioare di μ e σ. Quando μ aumenta la curva si sposta in modo rigido verso destra, quando diminuisce verso sinistra. Al variare di σ invece cambia la *larghezza* della curva: per σ piccoli la curva è molto alta e scende a zero in modo molto rapido. Per σ grandi la curva invece è sparpagliata.

TEOREMA LIMITE CENTRALE

Nell'espressione (57) per la distribuzione $P_n(k)$ compaiono i coefficienti binomiali $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ e la loro valutazione numerica la quando n e k sono molto grandi risulta grandemente facilitata dalla *formula asintotica di Stirling*:

$$n! = e^{-n}\, n^n \sqrt{2\pi n}\ \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right) . \tag{89}$$

Usando la (89) nella formula della distribuzione binomiale con $q = 1 - p$ otteniamo

$$\begin{aligned} P_n(k) &= \frac{n^n\, e^{-n} \sqrt{2\pi n}}{k^k\, e^{-k} \sqrt{2\pi k}\, (n-k)^{n-k}\, e^{-n+k} \sqrt{2\pi(n-k)}}\, p^k\, q^{n-k}\, (\, 1 + \epsilon(n,k)\,) \\ &= \frac{1}{\sqrt{2\pi n \frac{k}{n} \left(\frac{n-k}{n}\right)}} \left(\frac{k}{n}\right)^{-k} \left(\frac{n-k}{n}\right)^{-n+k} p^k\, q^{n-k}\, (\, 1 + \epsilon(n,k)\,) \\ &= \frac{1}{\sqrt{2\pi n t\, (1-t)}}\ e^{-nH(t)}\ (\, 1 + \epsilon(n,k)\,) \end{aligned} \tag{90}$$

con $\epsilon(n,k) = \mathcal{O}(1/n) + \mathcal{O}(1/k) + \mathcal{O}(1/(n-k))$ e dove si è posto $t = \frac{k}{n}$ e

$$H(t) = t\,\log\frac{t}{p} + (1-t)\,\log\frac{1-t}{q}. \tag{91}$$

Osserviamo che $H(p) = 0$. In particolare prendendo $k = np$ (supponendo per semplicità np intero) si trova

$$P_n(np) = \frac{1}{\sqrt{2\pi npq}}\,\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right) \tag{92}$$

Questa espressione asintotica si può interpretare con il fatto che in n prove indipendenti di un esperimento dicotomico la probabilità di avere esattamente np successi (cioè a dire un numero di successi pari al valor medio del numero di successi) tende a zero al crescere del numero di prove. Ad esempio in una sequenza di lanci della moneta, la probabilità di osservare *testa esattamente* la metà delle volte (o comunque un numero di volte che si mantiene vicino alla metà entro un margine fissato) tende a zero al crescere del numero dei lanci. Ciò non deve sorprendere: quale che sia il contenuto di una eventuale 'legge dei grandi numeri', non sarà certamente sensato aspettarsi esattamente 1000 volte testa in 2000 lanci! Vedremo che la domanda corretta da porre riguarda la probabilità che la *proporzione* delle occorrenze *testa* sia vicina a 1/2.

Ci proponiamo ora di stimare le probabilità $P_n(k)$ quando k varia entro un insieme di valori 'ragionevolmente frequenti'. A questo scopo osserviamo che per la (63) la larghezza media della distribuzione $P_n(k)$ è data da $\sqrt{V(f)} = \mathcal{O}(\sqrt{n})$ e ci aspetteremo che i valori suddetti si trovino entro un intervallo di tale ampiezza attorno al valor medio np. Poniamo dunque

$$x = \frac{k - np}{\sqrt{n\,p\,q}}$$

e, dato $a > 0$, consideriamo i valori di t corrispondenti a

$$|x| \leq \alpha \quad \text{ossia} \quad |t-p| \leq \alpha\sqrt{p\,q/n}.$$

Essendo $t = p + x\sqrt{p\,q/n}$ e $1-t = q - x\sqrt{p\,q/n}$ si ha

$$t\,(1-t) = pq\left(1 + \frac{(q-p)x}{\sqrt{npq}} - \frac{x^2}{n}\right)$$

e dunque per il primo fattore della (90) otteniamo:

$$\frac{1}{\sqrt{2\pi nt\,(1-t)}} = \frac{1}{\sqrt{2\pi npq}}\,\left(1 - \frac{(q-p)x}{2\sqrt{npq}} + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

Inoltre si ha

$$H'(t) = \log\frac{t}{p} - \log\frac{1-t}{q}, \quad H''(t) = \frac{1}{t} + \frac{1}{1-t}, \quad H'''(t) = -\frac{1}{t^2} + \frac{1}{(1-t)^2},$$

da cui otteniamo

$$H'(p) = 0, \quad H''(p) = \frac{1}{pq}, \quad H'''(p) = \frac{p-q}{p^2q^2}.$$

Dunque, ricordando che $t = p + x\sqrt{p\,q/n}$ e usando la formula di Taylor in $t = p$ (cioè $x = 0$) troviamo

$$H(t) = \frac{x^2}{2n} - \left(\frac{q-p}{\sqrt{npq}}\right)\frac{x^3}{6n} + \mathcal{O}(|t-p|^4)$$

Da cui,

$$e^{-nH(t)} = e^{-\frac{x^2}{2}}\left(1 + \frac{(q-p)x^3}{6\sqrt{npq}} + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

Inserendo queste espressioni nella (90) ricaviamo il *teorema limite locale* di De-Moivre-Laplace: se $|x| = \left|(k-np)/\sqrt{npq}\right| \leq \alpha$ allora

$$P_n(k) = \frac{e^{-x^2/2}}{\sqrt{2\pi npq}}\,\left(1 + \frac{(q-p)(x^3-3x)}{6\sqrt{npq}} + \mathcal{O}\left(\frac{1}{n}\right)\right). \tag{93}$$

In altre parole, le probabilità $P_n(k)$, quando n e k tendono entrambi all'infinito in modo tale che $|(k-np)/\sqrt{npq}|$ si mantenga limitato, sono approssimate dai valori della distribuzione gaussiana $p(k) = e^{-(k-m)^2/2\sigma^2}/\sqrt{2\pi}\sigma$ con $m = np$, $\sigma = \sqrt{npq}$. Osserviamo inoltre che il primo termine correttivo nella (93) si annulla quando $p = q$ e dunque rende conto dell'asimmetria della distribuzione $P_n(k)$ quando $p \neq \frac{1}{2}$ (vedi sopra). D'altra parte esso tende ad annullarsi al crescere di n e pertanto, entro l'intervallo considerato di valori di k, la distribuzione $P_n(k)$ tende a divenire simmetrica.

Dalla (93) si ha inoltre

$$\sum_{k \in R_{np}(\alpha)} P_n(k) = \sum_{|x| \leq \alpha} \frac{e^{-x^2/2}}{\sqrt{2\pi}} \, \Delta x \, \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right)$$

dove $\Delta x = x_{k+1} - x_k = 1/\sqrt{npq}$ e $x_k = (k - np)/\sqrt{npq}$. Ragionando come nella sezione dedicata alle distribuzioni continue otteniamo il *teorema limite integrale* di De-Moivre-Laplace:

$$\lim_{n\to\infty} \sum_{k:|k-np| \leq \alpha\sqrt{npq}} P_n(k) = \frac{1}{\sqrt{2\pi}} \int_{-\alpha}^{\alpha} e^{-x^2/2} \, dx \, . \tag{94}$$

Una suggestiva 'verifica sperimentale' di questo risultato si può ottenere per mezzo del cosiddetto *tavolo di Galton*, ossia un piano inclinato di forma triangolare su cui sono piantati dei chiodi in modo regolare lungo n linee orizzontali, dove la k-esima linea (dall'alto) contiene k chiodi. Una pallina, lanciate dal vertice del triangolo, verrà deviata in corrispondenza di ogni linea di chiodi verso destra o verso sinistra, con la stessa probabilità $\frac{1}{2}$. Sotto l'ultima linea di chiodi si trova una serie di $n + 1$ scatoline entro le qualli le palline si accumulano. Affinchè una pallina finisca nella scatola k-esima il suo tragitto attraverso i chiodi dovrà consistere di k deviazioni verso destra e $n - k$ verso sinistra. La probabilità di un tale tragitto è evidentemente $\binom{n}{k} \cdot 2^{-n}$. Lanciando un gran numero di palline (molto piccole) la loro distribuzione nella scatole approssima in modo sorprendentemente preciso la curva gaussiana.

Bibliography

[1] M. Piattelli Palmarini, *L'illusione di Sapere*, Mondadori (Oscar Saggi), (1995)

[2] William P. Thurston: "On Proof and Progress in Mathematics" Bulletin of the American Mathematical Society 30 (2), 161-177, (1994).

[3] William P. Thurston: "Mathematical Education.", Notices of the American Mathematical Society 37, no. 7, September (1990): 844-850.

[4] D.Epstein and S. Levy: "Experimentation and Proof in Mathematics" Notices of the American Mathematical Society 42, no. 6, June (1995): 670-674

[5] G.Prodi, *Matematica Come Scoperta*, D'Anna, (1975)

[6] Y.G. Sinai, *Probability theory : an introductory course*, Springer, (1992)