Somma di numeri aleatori

Michele Impedovo

Riassunto L'operazione di somma (di numeri, di polinomi, di equazioni, di vettori, di matrici, di funzioni, …) è uno dei pilastri dell'intera matematica. Sommare due oggetti è spesso il modo più naturale di tener conto del loro effetto congiunto. Anche per i numeri aleatori (o *variabili aleatorie*) risulta naturale definire l'operazione di somma. Molti giochi, ad esempio, si basano sulla somma di due dadi, cioè di due particolari numeri aleatori con la stessa distribuzione di probabilità. Oppure: il guadagno totale di più scommesse è la somma dei guadagni aleatori in ciascuna di esse. La somma di numeri aleatori presenta forti analogie e significative differenze rispetto alla somma di numeri. Il problema più generale potrebbe essere così enunciato: note le distribuzioni di probabilità di due numeri aleatori X e Y, qual è la distribuzione di probabilità di $X+Y$?

Abstract The sum operation (sum of numbers, polynomials, equations, vectors, matrices, functions, …) is fundamental in mathematics. Adding two quantities together is often the most natural way to take into account their combined effects. Even for random numbers (i.e. *random variables*) it is natural to define the addition operation. Many games of chance, for example, are based on the sum of two dice, which is the sum of two particular random numbers; when gambling, the total payoff is the sum of random gains and losses from each gamble. The sum of random numbers has strong similarities and yet significant differences relative to the sum of numbers. The more general problem is: given the probability distributions of two random numbers X and Y, what is the probability distribution of $X+Y$?

Michele Impedovo Università Bocconi di Milano michele.impedovo@unibocconi.it

Chance, as we understand it, supposes the existence of things, and their general known properties: that a number of dice, for instance, being thrown, each of them shall settle upon one or other of its bases. After which, the probability of an assigned chance, that is of some particular disposition of the dice, becomes as proper a subject of investigation as any other quantity or ratio can be. But chance, in atheistical writings or discourse, is a sound utterly insignicant: it imports no determination to any mode of existence; nor indeed to existence itself, more than to non-existence; it can neither be defined nor understood: nor can any proposition concerning it be either affirmed or denied, excepting this one, "That it is a mere word."

Abraham de Moivre (1667-1754)

La mia tesi, paradossale e un po' provocatoria, ma genuina, è che semplicemente la probabilità non esiste.

Bruno de Finetti (1906-1985)

Un problema, per iniziare

Alla roulette (37 numeri da 0 a 36) gioco 1 € sul rosso.

Se esce un numero rosso ritiro dal banco 2 €, e quindi vinco 1 €; se esce un numero nero, oppure esce lo 0, perdo la mia puntata di 1 €.

La mia vincita ("vincita" in senso lato, potrebbe anche essere negativa), mentre il piatto della roulette sta ancora girando e non posso sapere che numero uscirà, è un *numero aleatorio* (o *variabile aleatoria*, *random variable* sui testi anglosassoni) che può assumere solo due valori: −1 o 1. Più precisamente, se ritengo che la roulette non sia truccata, perderò 1 con probabilità $q = 19/37$ e vincerò 1 con probabilità $p = 18/37$.

Fin qui nulla di particolarmente nuovo o interessante. Tranne la magia del *numero aleatorio*: alla domanda

"*Quanto vincerò?*"

possiamo rispondere, anziché con un ingenuo

"*Non posso saperlo, prima che la pallina della roulette si fermi*",

esibendo le informazioni che possediamo, che possono essere sintetizzate mediante la tabella seguente:

$$\begin{cases} -1 & 1 \\ q & p \end{cases} = \begin{cases} -1 & 1 \\ 19/37 & 18/37 \end{cases}$$

Questa tabella rappresenta la *distribuzione di probabilità* del numero aleatorio: sulla prima riga i valori che può assumere, sulla seconda riga le corrispondenti probabilità, di somma 1 (condizione di *normalizzazione*).

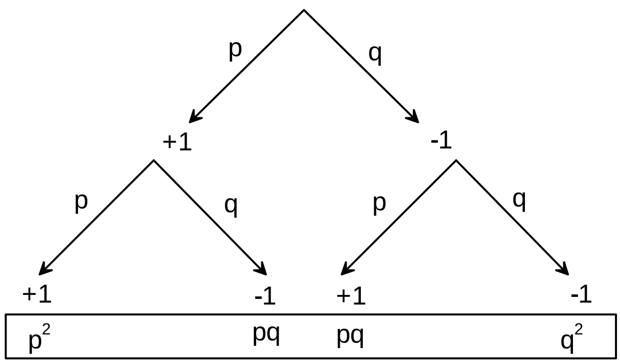
In un certo senso il numero aleatorio rappresenta il superamento della nostra ignoranza sugli eventi futuri: che faccia uscirà se lancio un dado? Non rispondiamo "*boh?*", bensì sfoderiamo valori possibili e corrispondenti probabilità; per esempio, se riteniamo che il dado sia regolare:

$$\begin{cases} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{cases}$$

La distribuzione di un numero aleatorio (coerentemente con l'impostazione *soggettivista* di Bruno de Finetti) rappresenta dunque lo stato corrente delle nostre conoscenze e delle nostre informazioni.

Torniamo alla roulette e alla mia puntata sul rosso. Poniamoci ora il seguente problema. Qual è la mia vincita se gioco due volte?

Poiché è naturale ritenere le due uscite tra loro indipendenti, possiamo considerare il grafo seguente:



e concludere che la vincita (aleatoria) totale è caratterizzata dalla seguente distribuzione:

$$\begin{cases} -2 & 0 & 2 \\ q^2 & 2pq & p^2 \end{cases}$$

Ecco una bella interpretazione del quadrato di un binomio: gli addendi di $(p+q)^2$ rappresentano le probabilità dei tre diversi esiti. Inoltre, poiché $p+q=1$, allora anche $(p+q)^2=1$.

L'esempio svolto è uno dei più semplici esempi di *somma di numeri aleatori*: dati i numeri aleatori (che riteniamo indipendenti) X_1 e X_2, che rappresentano rispettivamente la vincita nella prima e nella seconda puntata sul rosso, si tratta di comprendere quale sia la distribuzione del numero aleatorio X_1+X_2, che rappresenta la vincita totale. La vincita totale ha la seguente distribuzione:

$$X_1+X_2 \sim \begin{cases} -2 & 0 & 2 \\ q^2 & 2pq & p^2 \end{cases} = \begin{cases} -2 & 0 & 2 \\ 26.4\% & 50\% & 23.7\% \end{cases}$$

Si osservi che nonostante X_1 (vincita alla prima giocata) e X_2 (vincita alla seconda giocata) abbiano la stessa distribuzione, dobbiamo indicarli con nomi differenti; se chiamassimo X entrambi potremmo concludere erroneamente:

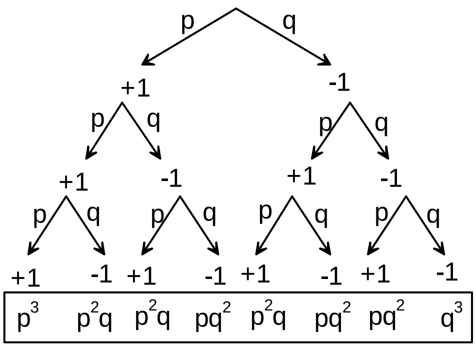
$$X+X=2X \sim \begin{cases} -2 & 2 \\ q & p \end{cases}$$

Invece con $2X$ è sensato indicare la vincita aleatoria di un'unica giocata in cui punto 2 €.

Ai numeri aleatori possiamo associare dunque le operazioni tipiche di uno spazio vettoriale su **R**: somma e multiplo reale.

Ancora un passo avanti: qual è la mia vincita se gioco tre volte?

Un rapido sguardo al grafo seguente



(che mostra ancora una bella interpretazione del cubo del binomio $p+q$) ci permette di concludere che la vincita totale è

$$X_1+X_2+X_3 \sim \begin{cases} -3 & -1 & 1 & 3 \\ q^3 & 3pq^2 & 3p^2q & p^3 \end{cases} = \begin{cases} -3 & -1 & 1 & 3 \\ 13.5\% & 38.5\% & 36.5\% & 11.5\% \end{cases}$$

La generalizzazione è immediata: qual è la vincita totale se gioco n volte 1 € sul rosso? La risposta è la seguente:

$$X_1+\ldots+X_n \sim \begin{cases} -n & -n+2 & \ldots & n-2 & n \\ q^n & \binom{n}{1} pq^{n-1} & \ldots & \binom{n}{n-1} p^{n-1}q & p^n \end{cases}$$

Per esempio, con $n = 5$:

$$X_1+\ldots+X_5 \sim \begin{cases} -5 & -3 & -1 & 1 & 3 & 5 \\ 3.6\% & 16.9\% & 32.0\% & 30.4\% & 14.4\% & 2.7\% \end{cases}$$

La difficoltà di una generalizazzione

Il problema più generale potrebbe essere così formulato: dati due numeri aleatori X eY, che supponiamo indipendenti sulla base delle nostre informazioni, di distribuzioni rispettive

$$X \sim \begin{cases} x_1 & \ldots & x_n \\ a_1 & \ldots & a_n \end{cases} \quad Y \sim \begin{cases} y_1 & \ldots & y_m \\ b_1 & \ldots & b_m \end{cases},$$

determinare la distribuzione del numero aleatorio

$$S = X+Y.$$

Ci interessa soprattutto una soluzione algoritmica di questo problema: dati in *input* i valori

$$x_1, \ldots, x_n, y_1, \ldots, y_m, a_1, \ldots a_n, b_1, \ldots, b_m$$

vogliamo fornire in uscita la matrice che ha sulla prima riga i valori possibili di $X+Y$ e sulla seconda riga le corrispondenti probabilità.

I valori che S può assumere sono ovviamente tutte le possibili somme

$$s_{ij} = x_i+y_j$$

con $i = 1, \ldots, n$ e $j = 1, \ldots, m$.
e, poiché riteniamo X e Y indipendenti, ciascuno di questi valori si presenta con probabilità a_ib_j:

$$\Pr(S=s_{ij}) = a_ib_j.$$

Potremmo concludere allora che la distribuzione di $X+Y$ è

$$X+Y \sim \begin{cases} x_1+y_1 & \ldots & x_n+y_m \\ a_1b_1 & \ldots & a_nb_m \end{cases}$$

Però c'è un problema: può essere che sulla prima riga si ottengano valori che si presentano più di una volta; se il numero s_{ij} si ottiene con diverse combinazioni di addendi x_i e y_j, occorrerà sommare le corrispondenti probabilità.

Per esempio, consideriamo

$$X \sim \begin{cases} 1 & 2 & 3 \\ 0.2 & 0.5 & 0.3 \end{cases} \qquad Y \sim \begin{cases} 4 & 5 \\ 0.4 & 0.6 \end{cases}$$

Le possibili somme tra i valori di X e i valori di Y sono $3 \cdot 2 = 6$:

$X \backslash Y$	**4**	**5**
1	5	6
2	6	7
3	7	8

Allora $S = X+Y$ può assumere i valori 5, 6, 7, 8. Ma mentre 5 e 8 si ottengono in un unico modo (rispettivamente 1+4 e 3+5) e quindi le corrispondenti probabilità sono

$$\Pr(S = 5) = 0.08$$
$$\Pr(S = 8) = 0.18,$$

i valori 6 e 7 si ottengono entrambi con due combinazioni:

6 = 1+5 (con probabilità $0.1\cdot 0.6=0.12$) = 2+4 (con probabilità $0.5\cdot 0.4=0.2$)

7 = 2+5 (con probabilità $0.5\cdot 0.6=0.3$) = 3+4 (con probabilità $0.3\cdot 0.4=0.12$)

Dunque

$$\Pr(S=6)=0.32$$
$$\Pr(S=7)=0.42$$

Ecco infine la distribuzione di $X+Y$:

$$S = X+Y \sim \begin{cases} 5 & 6 & 7 & 8 \\ 0.08 & 0.32 & 0.42 & 0.18 \end{cases}$$

È proprio la generalizzazione di questi passaggi, e cioè riconoscere quante e quali volte si ottiene la stessa somma da diverse combinazioni di addendi, a costituire la difficoltà maggiore per la generalizzazione che cerchiamo.

Un problema di dadi

Per comprendere meglio tale difficoltà consideriamo il problema seguente, apparentemente innocuo:

Qual è la probabilità che lanciando 5 dadi la somma delle facce sia 18?

Dal punto di vista operativo non sembra di scorgere intoppi insormontabili. Possiamo partire dal caso semplice di 2 dadi: si tratta di sommare i numeri aleatori X_1 e X_2, entrambi con distribuzione

$$\begin{cases} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{cases}$$

Le somme tra tutti i valori di X_1 e X_2 sono $6\cdot 6=36$, e ciascuna di esse si ottiene con probabilità

$$\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36}.$$

$X\backslash Y$	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

I valori possibili per X_1+X_2 sono dunque

$$2, 3, \ldots, 11, 12.$$

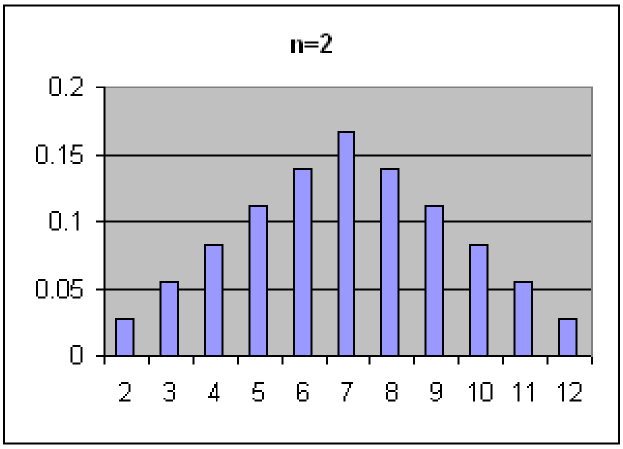
Ma solo il 2 e il 12 si ottengono con un'unica combinazione di addendi (rispettivamente 1+1 e 6+6), dunque

$$\Pr(2) = \Pr(12) = 1/36.$$

Per tutti gli altri valori possiamo osservare la tabella e contare il numero di occorrenze. La distribuzione di X_1+X_2 è allora la seguente:

$$X_1+X_2 \sim \begin{cases} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 1/36 & 2/36 & 3/36 & 4/36 & 5/36 & 6/36 & 5/36 & 4/36 & 3/36 & 2/36 & 1/36 \end{cases}$$

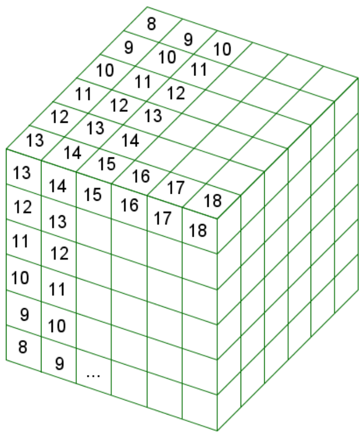
Il grafico mostra la tipica distribuzione "a tetto" della somma di due numeri aleatori *uniformi* (che assumono cioè n valori distinti ciascuno con probabilità $1/n$).



Ma come fare a rendere algoritmico quell'*osservare la tabella e contare il numero di occorrenze*?

Se saliamo a tre dadi il conteggio si presenta subito proibitivo: si tratta di sommare in tutti i modi possibili i numeri 2, 3, …, 12 (ma attenzione, non più con probabilità tutte uguali) con i numeri 1, 2, …, 6.

Oppure si può pensare a una tabella tridimensionale, cubica, su ogni spigolo della quale sono disposti gli addendi 1, 2, …, 6. Ecco una parziale rappresentazione di questa tabella vista da un osservatore dalla parte opposta rispetto all'origine.



Per un'attività didattica su questo tema si veda il bell'articolo di Mario Barra *Argomenti e termini nell'insegnamento del Calcolo delle probabilità*, Progetto Alice 2011 vol. XII n° 34.

Comunque sia, la generalizzazione a n dadi ci sembra per ora irraggiungibile.

Torniamo all'esempio della somma di numeri aleatori $X+Y$, con

$$X \sim \begin{cases} 1 & 2 & 3 \\ 0.2 & 0.5 & 0.3 \end{cases} \qquad Y \sim \begin{cases} 4 & 5 \\ 0.4 & 0.6 \end{cases}$$

Abbiamo visto che la difficoltà di generalizzare il risultato consiste nel riconoscere tutte le possibili combinazioni x_i+y_j che danno lo stesso risultato: una volta determinate queste combinazioni, dobbiamo sommarne le corrispondenti probabilità.

A pensarci bene, è esattamente a quello che accade quando moltiplichiamo due polinomi: eseguiamo il prodotto di ogni termine del primo per ogni termine del secondo, e poi sommiamo i monomi simili, cioè quelli per cui la variabile ha lo stesso esponente. Associamo allora a un numero aleatorio il polinomio (per esempio in t) che ha

- come esponenti di t: i valori $x_1, \ldots, x_n$ che il numero aleatorio può assumere
- come coefficienti: le corrispondenti probabilità $p_1, \ldots, p_n$.

Nel nostro esempio:

$$X \sim \begin{cases} 1 & 2 & 3 \\ 0.2 & 0.5 & 0.3 \end{cases} \qquad \Rightarrow \qquad X(t) = 0.2t + 0.5t^2 + 0.3t^3$$

$$Y \sim \begin{cases} 4 & 5 \\ 0.4 & 0.6 \end{cases} \qquad \Rightarrow \qquad Y(t) = 0.4t^4 + 0.6t^5$$

Le funzioni $X(t) = 0.2t + 0.5t^2 + 0.3t^3$ e $Y(t) = 0.4t^4 + 0.6t^5$ prendono il nome di *funzioni generatrici* rispettivamente di X e Y.

Moltiplichiamo ora i due polinomi; per farlo dobbiamo moltiplicare tra loro i coefficienti (cioè le probabilità: ecco l'indipendenza) e sommare tra loro gli esponenti (ecco la somma di numeri aleatori):

$$\left(0.2t + 0.5t^2 + 0.3t^3\right)\left(0.4t^4 + 0.6t^5\right) =$$
$$0.08t^5 + 0.2t^6 + 0.12t^7 + 0.12t^6 + 0.3t^7 + 0.18t^8$$

Ora sommare i monomi simili significa esattamente *contare il numero di combinazioni* con cui si ottiene una stessa somma e sommarne i coefficienti, cioè le probabilità:

$$\left(0.2t + 0.5t^2 + 0.3t^3\right)\left(0.4t^4 + 0.6t^5\right) = 0.08t^5 + 0.32t^6 + 0.42t^7 + 0.18t^8$$

Interpretiamo il risultato: la somma $X+Y$ ha come valori possibili gli esponenti di t e come probabilità i corrispondenti coefficienti:

$$S(t) = 0.08t^5 + 0.32t^6 + 0.42t^7 + 0.18t^8$$

cioè, come già sappiamo:

$$S = X+Y \sim \begin{cases} 5 & 6 & 7 & 8 \\ 0.08 & 0.32 & 0.42 & 0.18 \end{cases}$$

Magia!

Abbiamo trovato dunque una generalizzazione.

La funzione generatrice di due numeri aleatori indipendenti

$$X \sim \begin{cases} x_1 & \dots & x_n \\ a_1 & \dots & a_n \end{cases} \quad Y \sim \begin{cases} y_1 & \dots & y_m \\ b_1 & \dots & b_m \end{cases}.$$

è il prodotto delle funzioni generatrici di X e Y.

Dunque

$$X(t) = a_1 t^{x_1} + \dots + a_n t^{x_n},\ Y(t) = b_1 t^{y_1} + \dots + b_m t^{y_m}$$
$$S(t) = X(t)Y(t) = \left(a_1 t^{x_1} + \dots + a_n t^{x_n}\right)\left(b_1 t^{y_1} + \dots + b_m t^{y_m}\right) =$$
$$a_1 b_1 t^{x_1+y_1} + \dots + a_n b_m t^{x_n+y_m}$$

da cui si ricava la distribuzione di $X+Y$.

Questa magia funziona anche se $X(t)$ e $Y(t)$ non sono polinomi, cioè se i valori assunti da X e Y non sono numeri naturali. Per esempio, se torniamo all'esempio iniziale della roulette, possiamo ritrovare in altro modo i risultati già ottenuti; la vincita a ogni giocata ha distribuzione

$$X_i \sim \begin{cases} -1 & 1 \\ q & p \end{cases}$$

dunque la funzione generatrice è

$$X(t) = qx^{-1} + px = \frac{q}{x} + px$$

Se gioco due volte la vincita aleatoria ha la distribuzione corrispondente al quadrato di $X(t)$:

$$\left(\frac{q}{x} + px\right)^2 = \frac{q^2}{x^2} + 2pqx^0 + p^2x^2$$

dunque

$$X_1 + X_2 = \begin{cases} -2 & 0 & 2 \\ q^2 & 2pq & p^2 \end{cases}$$

La potenza n-esima di $(q/x+px)$ fornisce la funzione generatrice, e quindi la distribuzione, della vincita aleatoria dopo n giocate.

La somma di *n* dadi con la funzione generatrice

Con la funzione generatrice possiamo esprimere la soluzione del precedente problema dei dadi: Qual è la probabilità che lanciando 5 dadi la somma delle facce sia 18?

Se $X_1, \ldots, X_5$ sono le uscite (aleatorie) di ciascun dado, poiché hanno tutti la stessa distribuzione, e quindi la stessa funzione generatrice

$$X(t) = \frac{1}{6}x + \frac{1}{6}x^2 + \frac{1}{6}x^3 + \frac{1}{6}x^4 + \frac{1}{6}x^5 + \frac{1}{6}x^6$$

la funzione generatrice di $X_1+\ldots+X_5$ è $X(t)^5$. Con un software di calcolo simbolico (in questo caso Ti-Nspire) otteniamo:

$$f(t):=\sum_{k=1}^{6}\left(\frac{1}{6}\cdot t^{k}\right)$$ *Fatto*

$\text{expand}\left((f(t))^5\right)$

$$\frac{t^{30}}{7776}+\frac{5\cdot t^{29}}{7776}+\frac{5\cdot t^{28}}{2592}+\frac{35\cdot t^{27}}{7776}+\frac{35\cdot t^{26}}{3888}+\frac{7\cdot t^{25}}{432}+\frac{205\cdot t^{24}}{7776}+\frac{305\cdot t^{23}}{7776}+\frac{35\cdot t^{22}}{648}+\frac{5\cdot t^{21}}{72}+\frac{217\cdot t^{20}}{2592}+\frac{245\cdot t^{19}}{2592}+\frac{65\cdot t^{18}}{648}+\dots$$

Scorrendo il polinomio otteniamo che il termine t^{18} ha coefficiente 65/648; dunque la probabilità che la somma delle facce di 5 dadi sia 18 è $65/648 \approx 10\%$.

In assenza di un software di calcolo simbolico possiamo ricorrere al sito di Mathworld (www.wolframalpha.com).

Per esempio, per catturare il coefficiente di t^{18} nello sviluppo di $X(t)^5$ possiamo calcolarne la derivata 18-esima, dividere per 18! e calcolare il polinomio così ottenuto in 0.

WolframAlpha computational knowledge engine

```
18-th derivative of ((x+x^2+x^3+x^4+x^5+x^6)/6)^5/18! where x=0
```

Examples Random

Input interpretation:

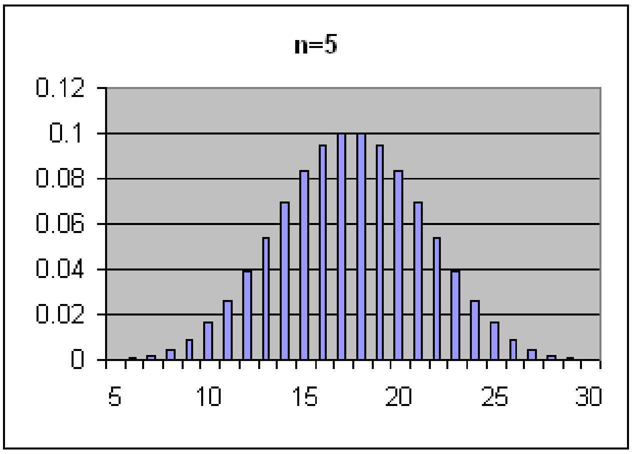
$$\frac{\partial^{18}}{\partial x^{18}}\frac{\left(\frac{1}{6}\left(x+x^2+x^3+x^4+x^5+x^6\right)\right)^5}{18!} \quad \text{where } x=0$$

$n!$ is the factorial function »

Result: More digits

$$\frac{65}{648}\approx 0.100309$$

Ecco il grafico della distribuzione della somma di 5 dadi, che già assomiglia a una curva gaussiana.



In modo analogo possiamo calcolare la distribuzione di probabilità della somma di *n* dadi.
Anzi, possiamo generalizzare il risultato alla somma di *n* "dadi" con un numero qualsiasi *f* di facce, cioè a numeri aleatori con distribuzione

$$X \sim \begin{cases} 1 & 2 & \dots & f \\ 1/f & 1/f & \dots & 1/f \end{cases}$$

per i quali la funzione generatrice è

$$X(t) = \sum_{k=1}^{f} \frac{1}{f} t^k$$

Ma … il lettore attento si sarà chiesto: già, ma come si fa a calcolare, "by fair means", il coefficiente di t^{18} nello sviluppo del polinomio

$$\left(\frac{1}{6}t + \frac{1}{6}t^2 + \frac{1}{6}t^3 + \frac{1}{6}t^4 + \frac{1}{6}t^5 + \frac{1}{6}t^6\right)^5 ?$$

D'accordo, abbiamo scoperto un prezioso "è come se": la probabilità che la somma di 5 dadi sia 18 è il coefficiente di t^{18} nello sviluppo di quel polinomio, ma dal punto di vista computazionale come si calcola tale coefficiente?

Vale la pena, ancora una volta, sottolineare come in matematica (e ovunque) il valore semantico del termine "risolvere" dipenda strettamente dagli strumenti che si ipotizza di poter utilizzare.

La somma di *n* dadi con un linguaggio di programmazione

Oppure, se abbiamo a disposizione un linguaggio di programmazione, possiamo risolvere il problema in forma iterativa, con 5 cicli "for" indentati. Ecco un esempio di programma di questo tipo.

```
Define dadi5(s)=
Func
Local c,d1,d2,d3,d4,d5
c:=0
For d1,1,6
  For d2,1,6
    For d3,1,6
      For d4,1,6
        For d5,1,6
          If d1+d2+d3+d4+d5=k Then
            c:=c+1
          EndIf
        EndFor
      EndFor
    EndFor
  EndFor
EndFor
c/6^5
EndFunc
```

Se calcoliamo la funzione **dadi5** con argomento 18 otteniamo il risultato che già conosciamo.

$$dadi5(18) \qquad \frac{65}{648}$$

Oppure possiamo risolvere il problema in forma ricorsiva. Perché la somma di 5 dadi (chiamiamola S_5) sia 18 è necessario che la somma dei primi 4 dadi sia compresa tra 12 e 17. Quindi

$$\begin{aligned}\Pr(S_5 = 18) = {} & \Pr(d_5 = 6|d_1+d_2+d_3+d_4 = 12)\cdot\Pr(d_1+d_2+d_3+d_4 = 12)+\\ & \Pr(d_5 = 5|d_1+d_2+d_3+d_4 = 13)\cdot\Pr(d_1+d_2+d_3+d_4 = 13)+\\ & \ldots+\Pr(d_5 = 1|d_1+d_2+d_3+d_4 = 17)\cdot\Pr(d_1+d_2+d_3+d_4 = 17) =\end{aligned}$$

$$\frac{1}{6}\Pr(d_1+d_2+d_3+d_4 = 12)+\frac{1}{6}\Pr(d_1+d_2+d_3+d_4 = 13)+\ldots+\frac{1}{6}\Pr(d_1+d_2+d_3+d_4 = 17)$$

La somma di 5 dadi è così ricondotta alla somma di 4 dadi. Una funzione ricorsiva che esegua questo calcolo per un numero qualsiasi *n* di dadi è la seguente.

```
Define dadin(s,n)=
Func
If s<n or s>6·n Then
  0
ElseIf n=1 Then
 1/6
Else
 1/6·∑(d=1..6) (dadin(s−d,n−1))
EndIf
EndFunc
```

Se calcoliamo la funzione **dadin** con argomenti 18 e 5 otteniamo il solito risultato.

$$dadin(18,5) \qquad \frac{65}{648}$$

Questa funzione è elegante, come tutte le funzioni ricorsive consente di "liberare la mente" dalla complessità del problema (cioè ogniqualvolta lo si riesca a ricondurre al caso precedente). Inoltre essa risolve il problema per un numero *n* qualsiasi di dadi, e si generalizza facilmente a "dadi" con *f* facce. Ma come sappiamo non è efficiente: se *n* è elevato la ricorsione è talmente profonda che i tempi di attesa diventano proibitivi.

La somma di *n* dadi secondo de Moivre

Insomma, dal punto di vista algoritmico non siamo ancora soddisfatti: abbiamo trovato problemi equivalenti, abbiamo riconosciuto una catena ricorsiva, ma non siamo riusciti a esprimere la probabilità della somma di *n* dadi in forma (relativamente) semplice, ipotizzando per esempio di avere a disposizione solo carta e penna.

Si attribuisce a de Moivre (A. de Moivre, *The doctrine of Chance*, 1718) la seguente soluzione diretta, che richiede solo la somma di qualche addendo.

$$\Pr(S_5 = 18) = \frac{1}{6^5}\left(\binom{5}{0}\binom{17}{4} - \binom{5}{1}\binom{11}{4} + \binom{5}{2}\binom{5}{4}\right) = \frac{1}{6^5}(2380 - 1650 + 50) =$$

$$\frac{780}{7776} = \frac{65}{648}$$

La funzione di de Moivre, per un numero qualsiasi n di dadi, è la seguente:

$$\Pr(S_n = s) = \frac{1}{6^n}\sum_{k=0}^{\text{floor}\left(\frac{s-n}{6}\right)} (-1)^k \binom{n}{k}\binom{s-6k-1}{n-1}$$

dove "floor" è la parte intera di $(s–n)/6$.

La dimostrazione di questa formula non è banale, ma vale la pena descrivere la sua genesi. Mostriamola nel caso $n = 5$. L'idea di partenza è ancora quella della funzione generatrice: per calcolare $\Pr(S_5 = 18)$ occorre calcolare il coefficiente di t^{18} nello sviluppo di

$$X(t)^5 = \left(\frac{1}{6}t + \frac{1}{6}t^2 + \frac{1}{6}t^3 + \frac{1}{6}t^4 + \frac{1}{6}t^5 + \frac{1}{6}t^6\right)^5$$

Raccogliamo $t/6$:

$$X(t)^5 = \left(\frac{1}{6}t\left(1 + t + t^2 + t^3 + t^4 + t^5\right)\right)^5 = \frac{1}{6^5}t^5\left(\frac{1-t^6}{1-t}\right)^5 =$$

$$\frac{1}{6^5}t^5\left(1-t^6\right)^5(1-t)^{-5}$$

Se si espande $\left(1-t^6\right)^5$ mediante il binomio di Newton si ottiene

$$\left(1-t^6\right)^5 = \sum_{h=0}^{5}\binom{5}{h}(-1)^h t^{6h}$$

Se si sviluppa $(1-t)^{-5}$ in serie di Taylor con centro in 0 si ottiene

$$(1-t)^{-5} = \sum_{k=0}^{+\infty}\binom{5+k-1}{k}t^k$$

Dunque

$$X(t)^5 = \frac{1}{6^5}t^5\left(\sum_{h=0}^{5}\binom{5}{h}(-1)^h t^{6h}\right)\left(\sum_{k=0}^{+\infty}\binom{5+k-1}{k}t^k\right)$$

A meno del fattore $1/6^5$ il coefficiente c_{18} di t^{18} si ottiene dalla somma di tutti gli addendi con

$$18 = 5+6h+k$$

da cui

$$k = 13-6h$$

quindi

$$c_{18} = \sum_{h=0}^{5}\binom{5}{h}(-1)^h\binom{5+13-6h-1}{13-6h} = \sum_{h=0}^{5}(-1)^h\binom{5}{h}\binom{17-6h}{13-6h}$$

ma $13-6h$ è positivo solo per $h \leq 13/6$ (in generale $(s-n)/6$); per $13-6h < 0$ il coefficiente binomiale vale 0. Dunque

$$c_{18} = \sum_{h=0}^{2}(-1)^h\binom{5}{h}\binom{17-6h}{13-6h}$$

Inoltre, poiché

$$\binom{17-6h}{13-6h} = \binom{17-6h}{17-6h-(13-6h)} = \binom{17-6h}{4}$$

si ottiene

$$c_{18} = \sum_{h=0}^{2}(-1)^h\binom{5}{h}\binom{17-6h}{4} = \binom{5}{0}\binom{17}{4}-\binom{5}{1}\binom{11}{4}+\binom{5}{2}\binom{5}{4} =$$

$$2380-1650+50 = 780$$

Dividendo c_{18} per 6^5 si ottiene la probabilità cercata:

$$\Pr(S_5 = 18) = \frac{780}{7776} = \frac{65}{648}$$

Questo procedimento si generalizza facilmente alla probabilità di ottenere somma s lanciando un numero n di dadi a f facce:

$$\frac{1}{f^n}\sum_{k=0}^{\mathrm{floor}\left(\frac{s-n}{f}\right)}(-1)^k\binom{n}{k}\binom{s-kf-1}{n-1}$$

Il caso continuo

Per andare al lavoro devo prendere un autobus della linea 94, che passa ogni 5 minuti; arrivo alla fermata in un istante casuale. Quanto tempo devo aspettare? Il numero aleatorio

X = tempo d'attesa (in minuti) alla fermata della 94 ha una distribuzione di probabilità che non è più *discreta*, cioè del tipo

$$\begin{cases} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{cases}$$

bensì il tempo di attesa si distribuisce sull'intervallo continuo [0,5] di **R**. Non ha più senso chiedersi

Qual è la probabilità di aspettare x minuti?

(tale probabilità è sempre nulla); la domanda corretta è

Qual è la probabilità di aspettare tra x_1 e x_2 *minuti?*

Come è noto, per descrivere la distribuzione di probabilità di un numero aleatorio continuo X occorre una funzione f, chiamata *funzione di densità di probabilità di X*, tale che per ogni intervallo $[a, b]$ la probabilità che X assuma valori tra a e b sia l'integrale definito di f in $[a, b]$, cioè per ogni intervallo $[a, b]$ risulti

$$\Pr(a \le X \le b) = \int_a^b f(x)\,\mathrm{d}x$$

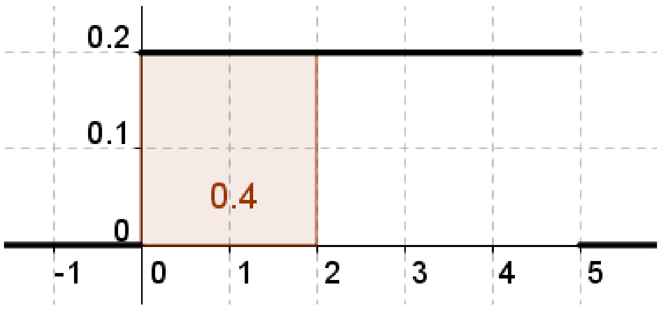
Di conseguenza f è una funzione non negativa e inoltre

$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$ (condizione di *normalizzazione*: la probabilità totale è 1).

Nel caso del tempo di attesa X è ragionevole ipotizzare che la distribuzione sia *uniforme*, cioè

$$f(x) = \begin{cases} 0.2 & 0 \le x \le 5 \\ 0 & \text{altrimenti} \end{cases}$$

In questo modo la probabilità $\Pr(a \le X \le b)$ è, per un intervallo $[a, b] \subseteq [0,5]$, proporzionale all'ampiezza stessa dell'intervallo, cioè a $b-a$. Per esempio, la probabilità di aspettare meno di 2 minuti è del 40%.



Fin qui nulla di particolarmente nuovo o interessante.

Ma supponiamo ora che per andare al lavoro io debba prendere due mezzi: la 94 (a Milano è una *filovia*, femminile), che passa ogni 5 minuti, e il 15 (è un tram, maschile), che anch'esso passa ogni 5 minuti. Supponiamo che i tempi di percorrenza delle due linee siano indipendenti (ipotesi del tutto ragionevole, almeno a Milano). Qual è il tempo di attesa totale?

Indichiamo con X_1 e X_2 i tempi aleatori di attesa rispettivamente della 94 e del 15, ciascuno con la propria funzione di densità (e supponiamo entrambe uniformi):

$$f_1(x) = f_2(x) = \begin{cases} 0.2 & 0 \le x \le 5 \\ 0 & \text{altrimenti} \end{cases}$$

Anche in questo caso si tratta di sommare due numeri aleatori: il tempo (aleatorio) di attesa totale è

X_1+X_2. Ebbene, come si distribuisce il numero aleatorio X_1+X_2? Più precisamente, qual è la sua funzione di densità f? È immediato constatare che non è f_1+f_2, sia perché non soddisfa la condizione di normalizzazione (la probabilità totale sarebbe 2):

$$f_1(x)+f_2(x) = \begin{cases} 0.4 & 0 \le x \le 5 \\ 0 & \text{altrimenti} \end{cases}$$

sia perché se X_1 e X_2 assumono valori tra 0 e 5, X_1+X_2 assume valori tra 0 e 10.

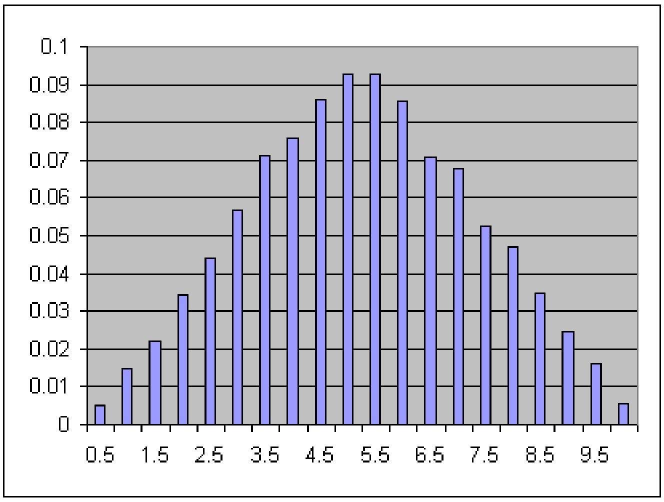
Simulazione e congettura

Una prima idea potrebbe essere quella di simulare il problema. Con Excel generiamo 10000 volte due numeri reali casuali uniformemente distribuiti tra 0 e 5 con il comando

=5*CASUALE()

e sommiamoli. Calcoliamo poi le frequenze relative con cui queste 10000 somme cadono negli intervalli [0,0.5), [0.5,1), …, [9.5,10] e ipotizziamo che tali frequenze siano una stima ragionevole delle corrispondenti probabilità.

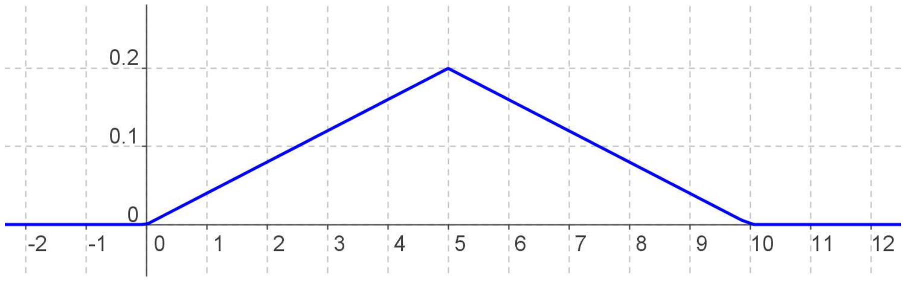
X1	X2	X1+X2	intervalli	fr. relative
4.690145	2.96614	7.656284	0.5	0.0049
1.513745	2.542711	4.056456	1	0.0147
3.023404	3.903103	6.926507	1.5	0.0221
3.433556	2.013488	5.447044	2	0.0342
0.517606	4.46099	4.978596	2.5	0.044
2.438785	2.365084	4.803869	3	0.0566
3.543309	4.895968	8.439277	3.5	0.071
4.354247	4.111744	8.46599	4	0.0758
2.491587	0.927665	3.419252	4.5	0.086
1.655013	4.936169	6.591182	5	0.0926
4.171175	1.16552	5.336695	5.5	0.0926
2.628865	4.140227	6.769092	6	0.0858
2.229705	3.986025	6.21573	6.5	0.0707
2.812528	2.898765	5.711294	7	0.068
4.831918	0.32092	5.152838	7.5	0.0527
2.811659	4.354764	7.166423	8	0.047
4.195706	2.086453	6.282159	8.5	0.0349
2.107753	1.300487	3.40824	9	0.0246
3.142032	3.951279	7.09331	9.5	0.0163
1.563062	1.893145	3.456208	10	0.0055
3.635029	0.671682	4.306711		
2.648179	3.820703	6.468882		
4.184881	0.768633	4.953513		



Otteniamo ancora una distribuzione "a tetto" nell'intervallo [0,10]: possiamo supporre che (in analogia con quanto accade per la somma di due dadi) la densità di probabilità di X_1+X_2 sia una funzione che parte da 0, cresce linearmente da 0 a 5 e decresce simmetricamente da 5 a 10.

L'unica funzione di questo tipo che soddisfa la condizione di normalizzazione è

$$f_{X_1+X_2}(x)=\begin{cases} 0.04x & 0\le x\le 5 \\ 0.4-0.04x & 5<x\le 10 \\ 0 & \text{altrimenti} \end{cases}$$



Se la congettura è corretta, allora la densità di probabilità non è uniforme, ma cresce (linearmente) avvicinandosi a 5. Per esempio, la probabilità che il tempo totale di attesa sia compreso tra 1 e 2 minuti è

$$\int_1^2 0.04x\,\mathrm{d}x = 0.07 = 6\%$$

mentre la probabilità che il tempo totale di attesa sia compreso tra 4 e 5 minuti è il triplo:

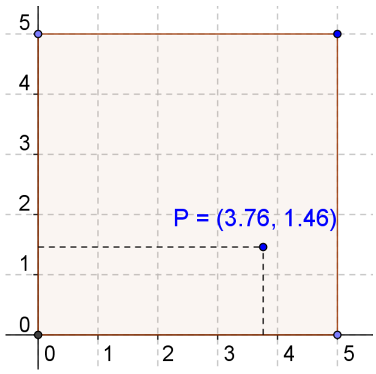
$$\int_4^5 0.04x\,\mathrm{d}x = 0.18 = 18\%$$

Validazione geometrica della congettura

X_1 e X_2 sono due numeri aleatori continui ritenuti indipendenti, con distribuzione uniforme tra 0 e 5. Si vuole sapere come è distribuito il numero aleatorio

$$Z = X_1+X_2.$$

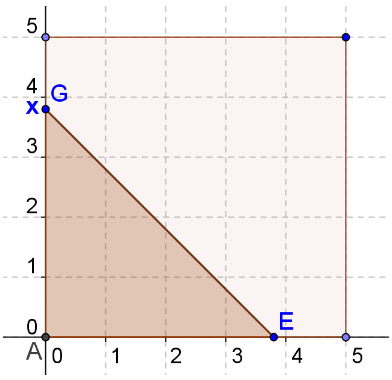
Poiché X_1 e X_2 sono indipendenti, scegliere un valore per X_1 e per X_2 significa scegliere a caso un punto nel quadrato unitario di vertici opposti (0,0) e (5,5).

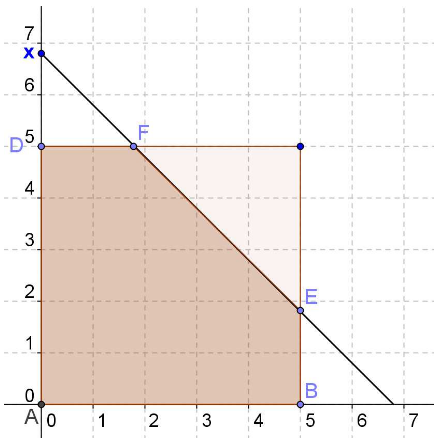


Fissato un x qualsiasi nell'intervallo [0,10], risulta

$$\Pr(Z \le x) = \Pr(X_1+X_2 \le x) = \Pr(X_2 \le -X_1+x).$$

dunque le coppie di valori (X_1, X_2) tali che $X_1+X_2 \le x$ sono rappresentate dai punti del quadrato "al di sotto" della retta $X_2 \le -X_1+x$. Possiamo allora interpretare la probabilità che X_1+X_2 sia minore di x come il rapporto tra l'area del triangolo AEG (o l'area del poligono ABEFD, a seconda del valore di x) e quella del quadrato nelle figure seguenti.





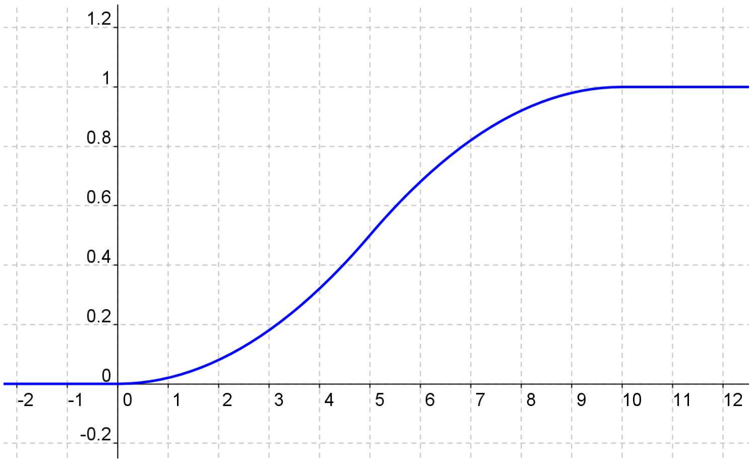
Distinguiamo due casi.

- $0 \le x \le 5$: allora $\Pr(X_1+X_2 \le x) = \dfrac{x^2}{50} = 0.02x^2$
- $5 \le z < 10$: allora $\Pr(X+Y) \le z = \dfrac{25-\frac{1}{2}(10-x)^2}{25} = -\dfrac{1}{50}x^2+\dfrac{2}{5}x-1 =$ $-0.02x^2+0.4x-1$

Dunque la funzione

$$F_Z(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & x < 0 \\ 0.02x^2 & 0 \le x \le 5 \\ -0.02x^2 + 0.4x - 1 & 5 < x \le 10 \\ 1 & 10 < x \end{cases}$$

rappresenta una funzione integrale (è la *funzione di ripartizione*) della funzione di densità f_Z di X_1+X_2.



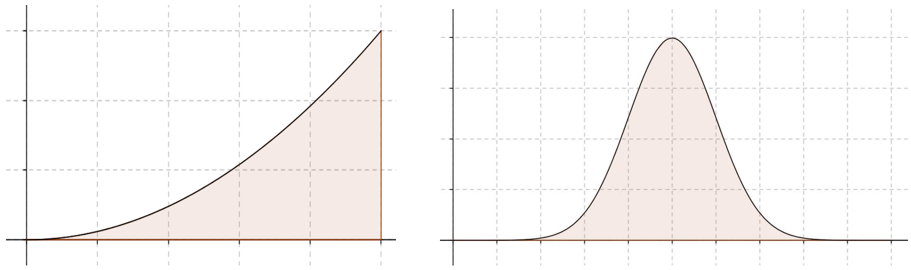
Derivando F_Z si ottiene f, che è proprio la funzione che avevamo congetturato:

$$f_Z(x) = \begin{cases} 0.04x & 0 \le x \le 5 \\ 0.4 - 0.04x & 5 < x \le 10 \\ 0 & \text{altrimenti} \end{cases}$$

Conclusioni

In un certo senso l'importanza dei numeri aleatori risiede nel magico concetto di *distribuzione di probabilità*, e cioè come si spalma la massa totale 1 di probabilità sui valori che il numero aleatorio può assumere. La somma di numeri aleatori mostra che anche se si parte da numeri aleatori con distribuzione uniforme, si arriva subito a distribuzioni che uniformi non sono.

Le fluttuazioni che questa funzione esibisce costituiscono una forma di conoscenza del mondo; associare distribuzioni di probabilità in modo immediato, visivo, a funzioni come le seguenti



aiuta a descrivere con un tratto di penna un intero sistema di informazioni. Tra i tanti paradigmi che si possono utilizzare a scuola per il concetto di funzione, quello di distribuzione di probabilità è forse uno dei più ricchi ed evocativi.

I problemi successivi che potremmo pensare di risolvere sono i seguenti.

- Se X_1 e X_2 sono ritenuti indipendenti, se X_1 è uniformemente distribuito su $[a, b]$ e X_2 è uniformemente distribuito su $[c, d]$, qual è la distribuzione di X_1+X_2?
- Se $X_1, X_2, \dots, X_n$ sono (a due a due, a tre a tre, …) ritenuti indipendenti e tutti uniformemente distribuiti in $[a, b]$, qual la distribuzione di $X_1+X_2+\dots+X_n$?
- Se X_1 e X_2 sono indipendenti e hanno densità di probabilità rispettivamente f_1 e f_2, qual è la densità di probabilità di X_1+X_2?

Bibliografia

Barra M., 2003, *Difficoltà nascoste nella didattica dei primi elementi di calcolo delle probabilità*, Progetto Alice n° 10

Barra M., 2003, *Aspetti storici e pedagogici relativi al calcolo combinatorio*, Progetto Alice n° 11

Barra M., 2005, *Alcuni problemi dell'insegnamento del calcolo delle probabilità*, Progetto Alice n° 18

Barra M., 2006, *La probabilità è nata in Italia*, Progetto Alice n° 19

Barra M., 2011, Argomenti e termini nell'insegnamento del Calcolo delle *probabilità*, Progetto Alice n° 34

Bucciarelli A., 2010, *L'insegnamento della matematica al liceo: de Finetti vs Laplace*, Bollettino dei Docenti di Matematica n° 61, CH
de Finetti B., 1970, *Teoria delle probabilità*, Einaudi
Eisenring M., 2011, Un'introduzione al calcolo delle probabilità, Bollettino dei Docenti di Matematica n° 62, CH
Feller W., 1950, *An introduction to probability theory*, John Wiley & Sons
Grimmett G., Stirzaker D., 1982, *Probability and random processes*, Oxford Science Publications
Invernizzi S., Rinaldi M., Sgarro A., 2000, *Moduli di matematica e statistica*, Zanichelli
Negrini P., Ragagni M., 2005, *La probabilità*, Carocci Faber
Prodi G., 1992, *Metodi matematici e statistici*, McGraw-Hill
Scozzafava R., 2001, *Incertezza e probabilità*, Zanichelli

Michele Impedovo


MOULIN ROUGE
MOULIN ROUGE
MOULIN ROUGE
CONCERT
BAL
TOUS Les SOIRS
LA GOULUE