Variabili aleatorie continue e simulazioni

Michele Impedovo

Riassunto Nell'introdurre le variabili aleatorie continue ci si scontra con la difficoltà di far comprendere che cosa sia la funzione di densità di probabilità. Mediante simulazioni numeriche e grafiche al calcolatore la "densità" assume un valore percettivo, e quindi semantico, di grande evidenza, che diventa punto di partenza per esplorazioni e scoperte matematiche.

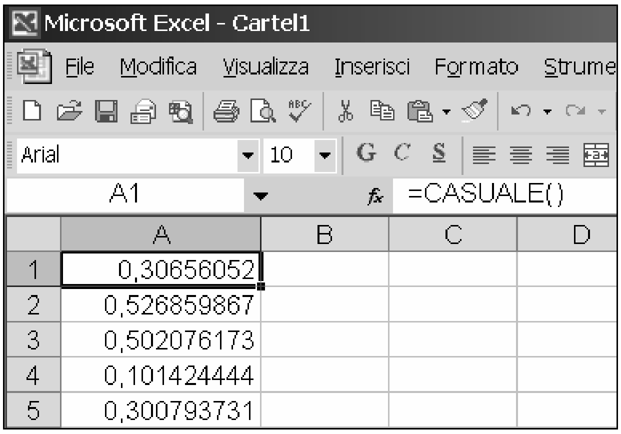
Abstract The outstanding learning difficulty, when introducing continuous random variables, is understanding what the probability density function is. The "density" can become a visual object by using numerical and graphical simulation on a computer and may be a starting point to mathematical explorations and discoveries.

Michele Impedovo
Università Bocconi di Milano
michele.impedovo@uni-bocconi.it

Prendi un numero casuale compreso tra 0 e 1 ...

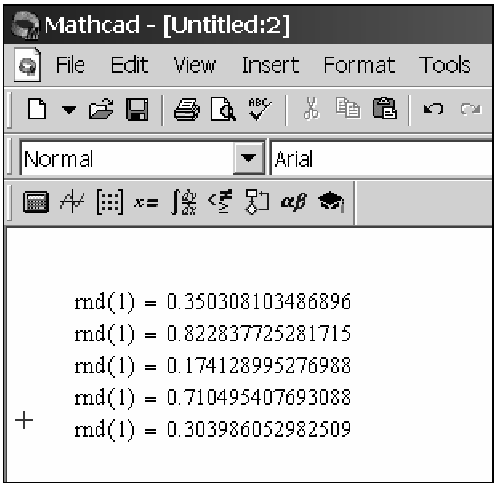
Siamo abituati ad usare il calcolatore o la calcolatrice per generare numeri casuali. Ogni software di matematica, sia esso di tipo numerico o di tipo simbolico, possiede un comando per generare un numero "casuale" compreso tra 0 e 1. Si tratta, naturalmente, di numeri generati da un algoritmo e per questo sarebbe meglio chiamarli "pseudo-casuali"; si può vedere, in proposito, il cap. 3, *Random numbers*, dell'intramontabile *The art of computer programming* di D. Knuth, Addison Wesley.

Per esempio con Excel il comando che genera un numero casuale compreso tra 0 e 1 è "`=casuale()`".



Per tutto ciò che segue si potrebbe dunque utilizzare Excel; è solo per motivi pratici e per mettere in campo una certa potenza di calcolo numerico che utilizzerò, per i calcoli e per i grafici, un software professionale di matematica, Mathcad (www.mathcad.com).

Con Mathcad il comando per generare un numero casuale tra 0 e 1 è "`rnd(1)`", dove "rnd" sta per "random".



Che cosa si intende per "numero casuale compreso tra 0 e 1"? Anche se dal punto di vista intuitivo l'espressione sembra chiara, non è facile precisarla. Siamo tentati di dire qualcosa del tipo "ogni numero ha la stessa probabilità di essere estratto". Le cose tuttavia sono un po' più complicate, perché la probabilità di estrarre un certo numero reale da un intervallo è sempre 0: non c'è modo di assegnare a ciascun numero reale dell'intervallo [0, 1] una probabilità positiva in modo che la somma delle probabilità sia 1.

Un'immagine adeguata potrebbe essere la seguente: vogliamo scegliere, sul segmento [0, 1], un numero reale (un punto) x in modo tale che, comunque si fissi un numero naturale M, x possa cadere con uguale probabilità in uno degli M sotto-intervalli individuati dai punti

$$0, \frac{1}{M}, \frac{2}{M}, \dots, \frac{M-1}{M}, 1.$$

In statistica si dice in questo caso che il numero casuale ha una **distribuzione uniforme**.

Un calcolatore non gestisce i numeri reali ma un opportuno sottoinsieme (finito) di numeri razionali: in linea di principio potremmo assegnare a ciascuno di essi una probabilità positiva, ma non è ciò che ci interessa. Noi vogliamo <u>simulare</u> la scelta di un punto a caso sul segmento [0, 1]; il generatore di numeri casuali può servire allo scopo; è questo il modo con cui lo utilizzeremo.

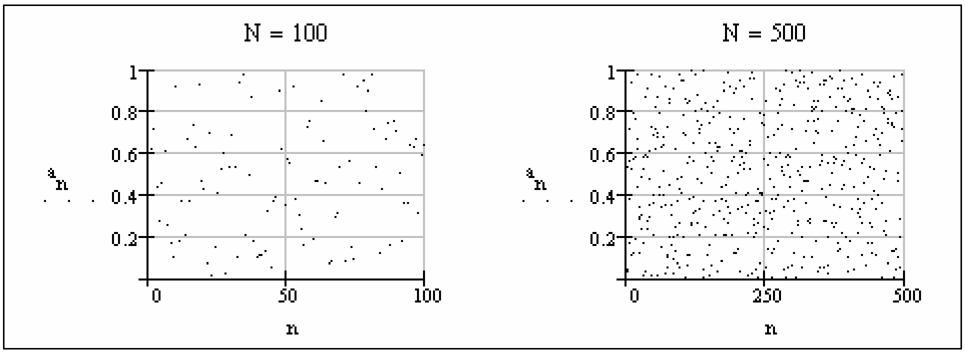
Un po' di esperimenti

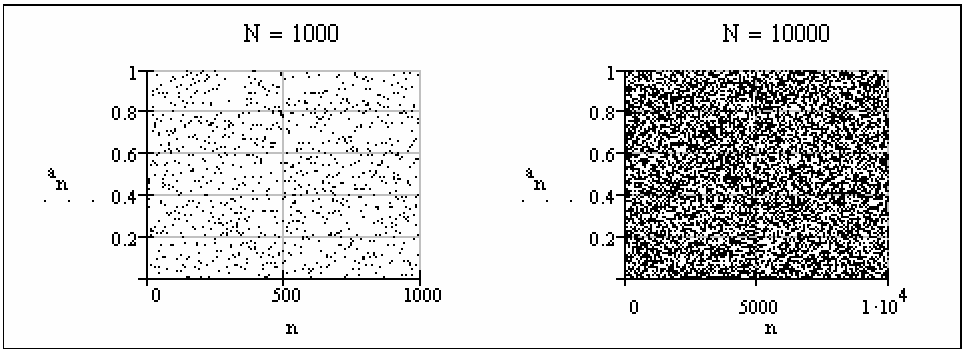
Generiamo un certo numero *N* di numeri casuali, raccolti nel vettore *a*.

$N := 100$

$n := 1 .. N$ $\qquad$ $a_n := \mathrm{rnd}(1)$

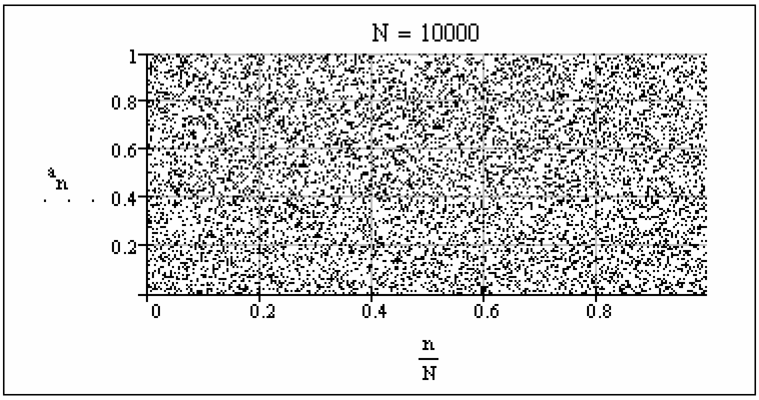
Tracciamo ora il grafico di questi *N* punti: in ascissa i numeri naturali da 1 a *N*, in ordinata i valori a_n, che variano in [0, 1]. Per analizzare il comportamento del grafico al crescere di *N*, tracciamo i grafici per $N := 100$, 500, 1000, 10000.



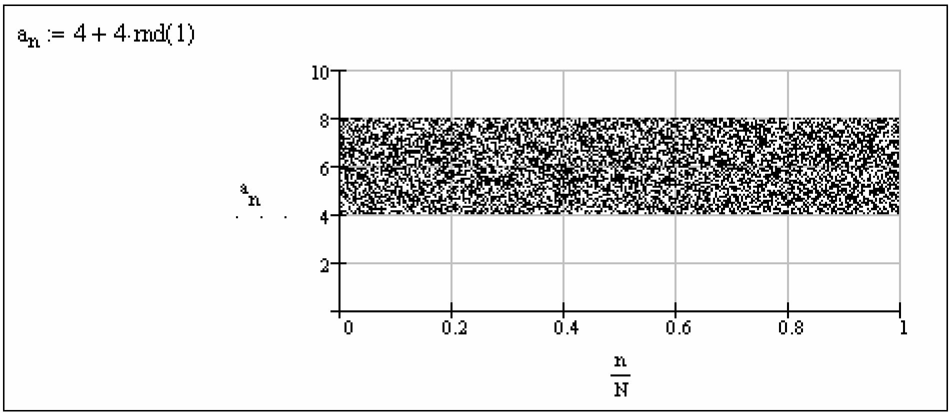


Il grafico è poco significativo: tutto ciò che possiamo osservare è che i punti sembrano distribuirsi in modo uniforme. Scorrendo il rettangolo $[0, N]\times[0, 1]$ sia orizzontalmente che verticalmente, al crescere di *N* non si notano (si notano sempre meno) regioni con maggiore densità di punti rispetto ad altre. L'uniformità orizzontale conferma empiricamente il fatto che ogni numero estratto è **indipendente** dai precedenti, nel senso che osservare il grafico fino ad un certo *n* non fornisce informazioni su quanto possa accadere da n a N; in altri termini l'algoritmo non degenera, continua a conservare "nel tempo" le proprie caratteristiche. L'uniformità verticale costituisce un test empirico del fatto che la distribuzione è **uniforme**: il generatore di numeri casuali "funziona bene".

Se vogliamo non tener conto del numero N di punti generati, possiamo "normalizzare" la grandezza in ascissa, dividendo per N; così in ascissa abbiamo ora la percentuale del numero totale di numeri generati, che varia da 0 a 1 qualunque sia N.



Costruiamo ora numeri casuali compresi (per esempio) tra 4 e 8, con il comando `4+4*rnd(1)`, e tracciamone il grafico.



Anche in questo caso non sembra di vedere, nel rettangolo $[0, 1]\times[4, 8]$, regioni con densità maggiore di punti; dal punto di vista empirico possiamo ipotizzare che se X è un numero casuale con distribuzione uniforme, anche $aX+b$ ha distribuzione uniforme.

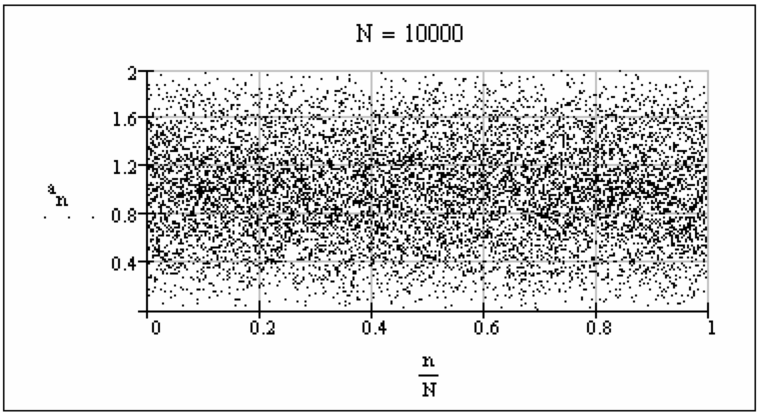
Somma di numeri casuali

Le cose cambiano se invece costruiamo la somma `rnd(1)+rnd(1)` di due numeri casuali compresi tra 0 e 1.

$N := 10000$

$n := 1..N \qquad a_n := \text{rnd}(1) + \text{rnd}(1)$

Se rnd(1) è compreso tra 0 e 1, allora rnd(1)+rnd(1) è compreso tra 0 e 2. Questa volta il grafico mostra una distribuzione di punti non uniforme: se scorriamo verticalmente il rettangolo $[0, 1]\times[0, 2]$ del grafico notiamo che la densità è maggiore al centro (vicino a 1) e minore agli estremi (vicino a 0 e a 2).

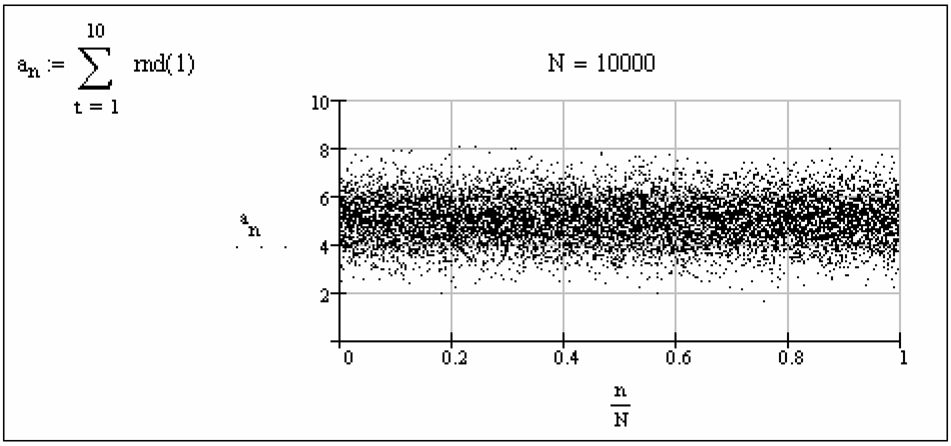


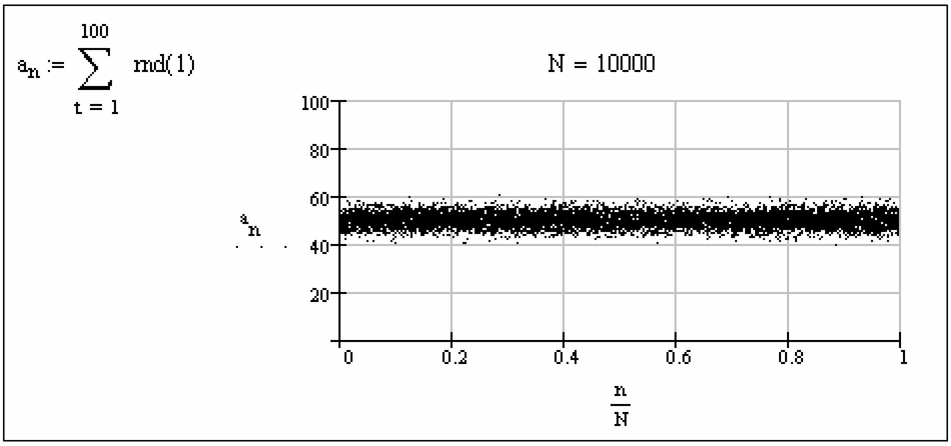
Dunque

$$\text{rnd}(1) + \text{rnd}(1) \neq 2\text{rnd}(1).$$

In altri termini, se X e Y hanno entrambe distribuzione uniforme tra 0 e 1 allora $X+Y$ non ha distribuzione uniforme tra 0 e 2. È ragionevole: per ottenere un numero vicino a 0 (oppure vicino a 2) occorre che entrambi gli addendi siano vicini a 0 (oppure entrambi vicini a 1). È esattamente ciò che accade, nel discreto, con il lancio di due dadi regolari: ciascuno di essi ha una distribuzione uniforme, ma non è uniforme la probabilità della loro somma: come è noto i valori estremi 2 e 12 hanno la probabilità minima di uscita (1/36) mentre il valore centrale 7 ha probabilità massima (6/36).

Questo fatto appare con maggiore evidenza se sommiamo, ad esempio, 10 numeri casuali, oppure addirittura 100. Si ottengono, rispettivamente, numeri compresi tra 0 e 10 e tra 0 e 100.



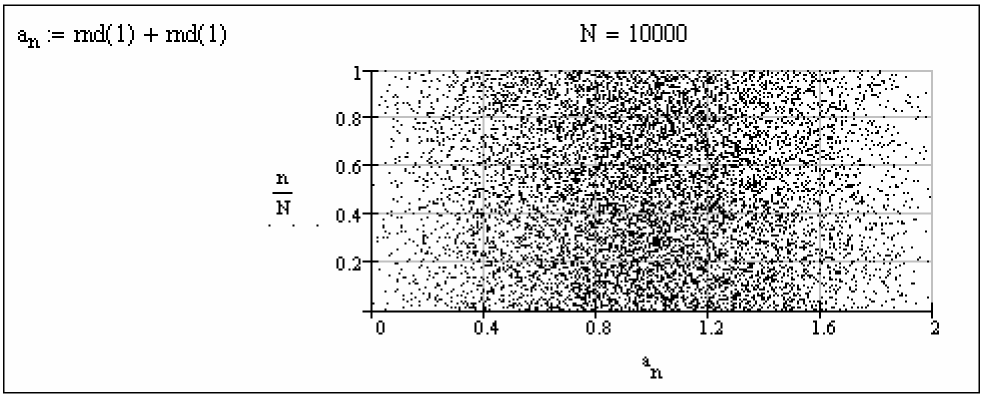


I punti si addensano, rispettivamente, intorno a 5 e intorno a 50, e la loro dispersione relativa sembra diminuire all'aumentare del numero di addendi.

Un colpo di bacchetta magica: la funzione di ripartizione

Fino ad ora abbiamo visto grafici di punti generati in modo casuale e perciò disordinato; l'unica caratteristica che siamo riusciti a cogliere in modo empirico è stata la densità dei punti nell'intervallo dei valori che i numeri casuali potevano assumere.

In effetti è proprio questa densità che vorremmo descrivere, in funzione dei valori assunti dal numero casuale; per esempio nel caso di $X+Y$, che può variare tra 0 e 2, vorremmo una funzione che descrivesse la densità dei punti al variare di x in $[0, 2]$. Un primo passo, puramente qualitativo, consiste semplicemente nell'invertire, sul grafico, l'ascissa n/N con l'ordinata a_n. Il grafico seguente mostra, ad esempio, la densità di $X+Y$ sull'intervallo $[0, 2]$, dove X e Y sono, come al solito, due numeri casuali distribuiti uniformemente tra 0 e 1.



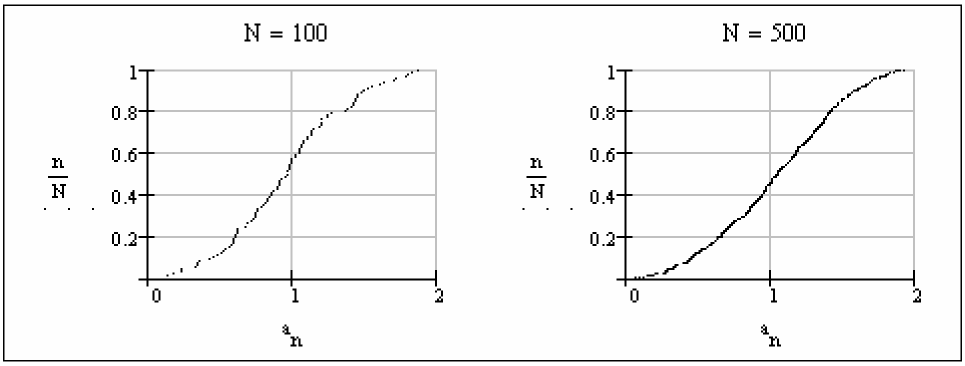
Al variare di x tra 0 e 2, sul grafico precedente possiamo leggere, almeno dal punto di vista qualitativo, come varia la densità dei punti.

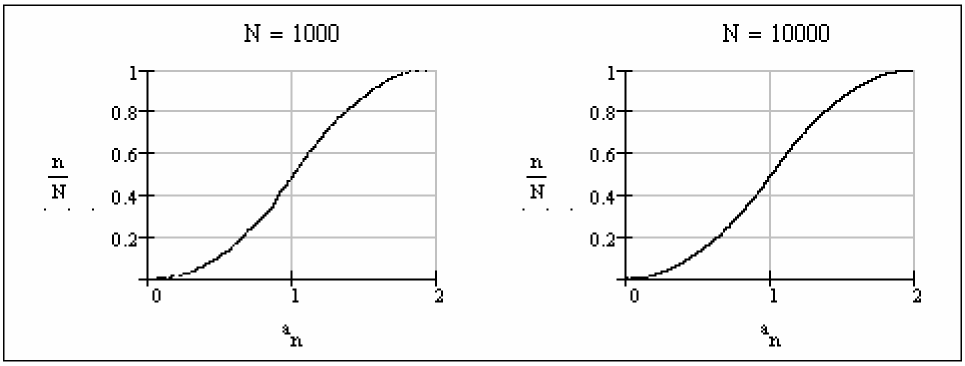
Ci piacerebbe ora quantificare in qualche modo ciò che fino ad ora abbiamo chiamato "densità", per passare dalla fase di esplorazione a quella di scoperta matematica. Vediamo come fare.

Ogni software di matematica possiede un comando per l'ordinamento degli elementi di un vettore; con Mathcad il comando "`sort(a)`" costruisce un vettore con gli stessi elementi del vettore a ordinati in senso crescente. Proviamo allora a prendere la nuvola di punti del grafico precedente e ordiniamone in senso crescente gli elementi. I grafici seguenti mostrano il risultato dei comandi

$n := 1..N \qquad a_n := \mathrm{rnd}(1) + \mathrm{rnd}(1) \qquad a := \mathrm{sort}(a)$

con valori di N (cioè il numero di punti generati) uguali rispettivamente a 100, 500, 1000, 10000.





Magia!

Dalla nuvola informe di punti emerge netto il profilo di una curva: di che curva si tratta? Che cosa descrive? É possibile esprimerla come funzione elementare?

Cerchiamo di capire innanzitutto di che curva si tratti. Fissiamo sulle ascisse un dato valore x compreso tra 0 e 2, cioè uno dei valori che il numero casuale $X+Y$ può assumere. Saliamo in verticale fino ad incontrare la curva in un certo valore $F(x)$ compreso tra 0 e 1: poiché il vettore a è ordinato, $F(x)$ è la percentuale di numeri estratti minori o uguali a x.

Per ogni $x \in [0, 2]$ il numero $F(x)$ è la percentuale di numeri generati che risultano minori o uguali a x.

Per esempio sul grafico leggiamo $F(1.5) \approx 0.87$: circa l'87% degli N numeri calcolati con il comando `rnd(1)+rnd(1)` sono minori o uguali a 1.5.

Se N è sufficientemente grande, possiamo interpretare $F(x)$ come la **probabilità** che $X+Y$ sia minore o uguale a x.

$F(x)$ è una funzione crescente; risulta ovviamente $F(x) = 0$ per ogni $x < 0$

(nessun numero estratto è minore di 0) e $F(x) = 1$ per ogni $x > 2$ (tutti i numeri estratti sono minori o uguali a 2).

La funzione $F(x)$ è chiamata in statistica **funzione di ripartizione**, o **funzione di distribuzione** del numero casuale $X+Y$. Per approssimare la probabilità che $X+Y$ sia compreso tra a e b è sufficiente calcolare $F(b) - F(a)$, cioè la percentuale di numeri generati compresa tra a e b.

Per esempio nella nostra simulazione risulta che la probabilità che $X+Y$ sia compreso tra 0.4 e 0.6 (intervallo di ampiezza 0.2) è

$$F(0.4) - F(0.2) \approx 0.177 - 0.081 = 0.096 = 9.6\%.$$

mentre la probabilità che sia compreso tra 0.9 e 1.1 (intervallo di uguale ampiezza) è

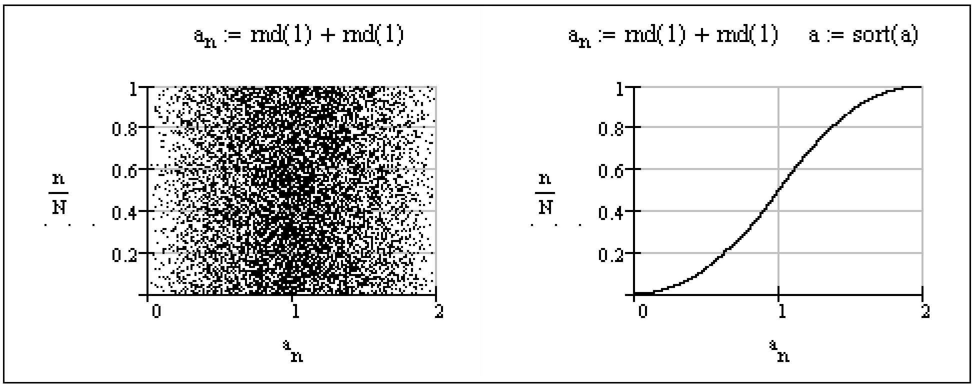
$$F(1.1) - F(0.9) \approx 0.589 - 0.403 = 0.186 = 18.6\%,$$

una probabilità circa doppia. Otteniamo dunque un primo dato quantitativo che conferma ciò che abbiamo osservato solo qualitativamente sul grafico della nuvola di punti di $X+Y$.

Ma c'è di più.

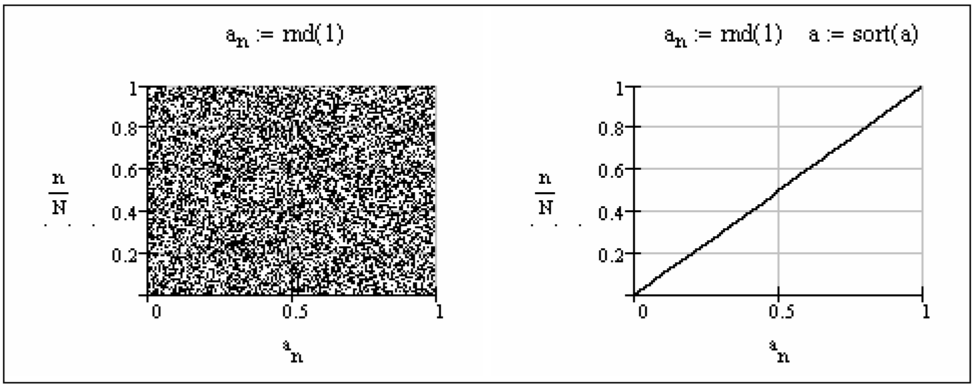
La funzione di densità di probabilità

Avrete osservato che il grafico di $F(x)$ per il numero casuale $X+Y$ non è lineare: tra 0 e 1 cresce in modo convesso (cioè la pendenza aumenta) mentre cresce in modo concavo tra 1 e 2 (la pendenza diminuisce). Confrontiamo i grafici del numero casuale $X+Y$:= rnd(1)+rnd(1) e della sua funzione di ripartizione.



Nel grafico di sinistra osserviamo che la densità cresce da 0 a 1 e decresce da 1 a 0. Nel grafico di destra leggiamo che la pendenza di $F(x)$ cresce da 0 a 1 e decresce da 1 a 2. Forse abbiamo trovato come descrivere in termini quantitativi la "densità" della nuvola di punti al variare dei valori assunti da $X+Y$: la densità in un certo x è la pendenza della funzione di ripartizione $F(x)$.

Chiediamoci allora: per quale numero casuale la funzione di ripartizione è lineare, cioè con pendenza costante? L'intuizione ci suggerisce che se la distribuzione di un numero casuale fosse uniforme allora la funzione di ripartizione dovrebbe crescere in modo lineare. Proviamo.



Magnifico!

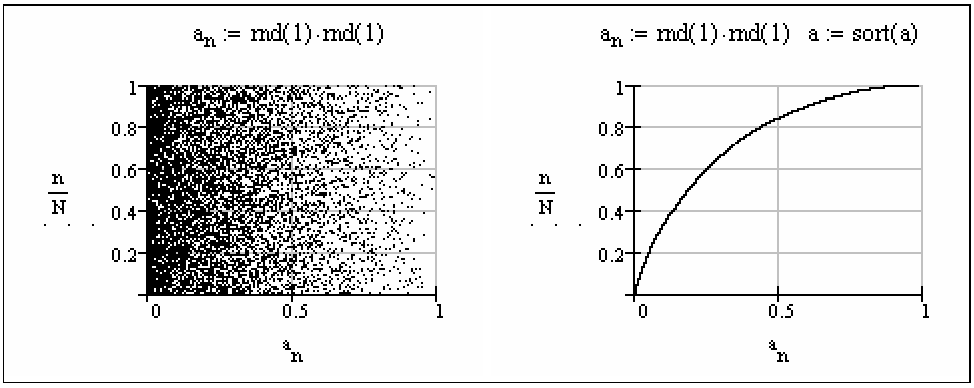
Dunque abbiamo trovato ciò che è costante nella distribuzione uniforme: la pendenza della funzione di ripartizione. Se riprendiamo il concetto intuitivo di "densità" con il quale siamo partiti, possiamo ora darne una definizione quantitativa: la densità di un numero casuale in corrispondenza di un certo valore x è la pendenza della sua funzione di ripartizione $F(x)$ in x.

Se la densità è costante allora $F(x)$ ha pendenza costante (è lineare).

Se la densità non è costante allora la pendenza di $F(x)$ varierà, in generale, da punto a punto: la densità che vogliamo definire non è altro che la derivata $F'(x)$ di $F(x)$. In statistica essa è chiamata **funzione di densità di probabilità**. La indicheremo con $f(x)$.

$$f(x) := F'(x).$$

Per fissare le idee osserviamo la seguente coppia di grafici, che mostra la nuvola di punti del numero casuale XY ottenuto mediante il comando `rnd(1)·rnd(1)`, e la simulazione della sua funzione di ripartizione.



Sul grafico di sinistra si legge che la densità decresce tra 0 e 1. Sul grafico di destra si legge che la pendenza di $F(x)$ decresce tra 0 e 1.

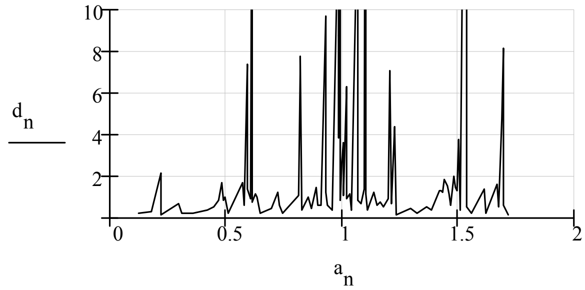
Se fosse possibile esprimere $F(x)$ in forma simbolica come funzione elementare, la sua derivata $f(x)$ descriverebbe punto per punto la densità della distribuzione. Purtroppo la formulazione esplicita di $F(x)$ richiede in generale un bagaglio teorico che non è alla portata di uno studente secondario; inoltre non è raro il caso in cui $F(x)$ non è esprimibile come funzione elementare. Per questo motivo vogliamo continuare ad utilizzare il calcolatore come strumento di simulazione, di esplorazione e di scoperta.

Una prima idea per simulare la funzione $f(x)$ potrebbe essere quella di calcolare la pendenza di $F(x)$ (che noi conosciamo solo in forma numerica, come coppia di vettori), per ogni coppia di punti consecutivi; costruiamo perciò il vettore formato dagli elementi

$$\frac{\frac{n+1}{N}-\frac{n}{N}}{a_{n+1}-a_n}=\frac{1}{N\left(a_{n+1}-a_n\right)}.$$

Ci aspettiamo, per $X+Y$, una curva che cresce da 0 a 1 e decresce da 1 a 2. Il risultato però è inaspettatamente (e notevolmente) irregolare, come mostra la seguente figura, che in un certo senso ci riporta alla realtà: non dimentichiamo che i punti sono stati scelti in modo (pseudo) casuale.

$$d_n := \frac{1}{N \cdot (a_{n+1} - a_n)}$$



Mentre $F(x)$, al crescere di N, appare via via più regolare, ciò non accade per $F'(x)$: poiché analizziamo il comportamento "microscopico" tra a_n e a_{n+1}, l'irregolarità non decresce al crescere di N.

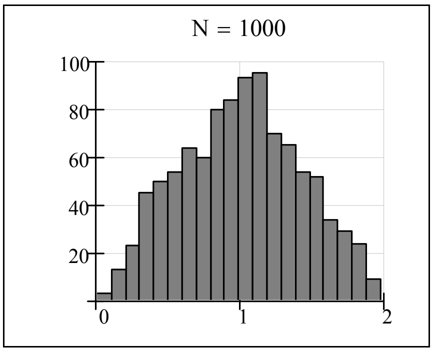
Seguiamo allora un'altra strada, più diretta, che riesca a sfruttare il crescere di N. La derivata in x di $F(x)$ può essere approssimata, se h è sufficientemente piccolo, dal rapporto

$$\frac{F(x+h)-F(x)}{h}$$

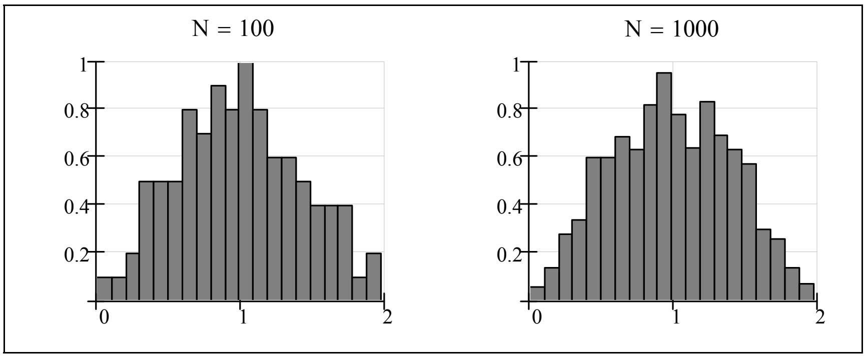
che, a meno del fattore $1/h$, è la percentuale di punti che cadono tra x e $x+h$. Costruiamo allora l'istogramma delle frequenze relative del vettore a, ponendo (per esempio) $h := 0.1$ e dividendo perciò l'intervallo $[0, 2]$ nei 20 sotto-intervalli $[0, 0.1)$, $[0.1, 0.2)$, …, $[1.9, 2]$.

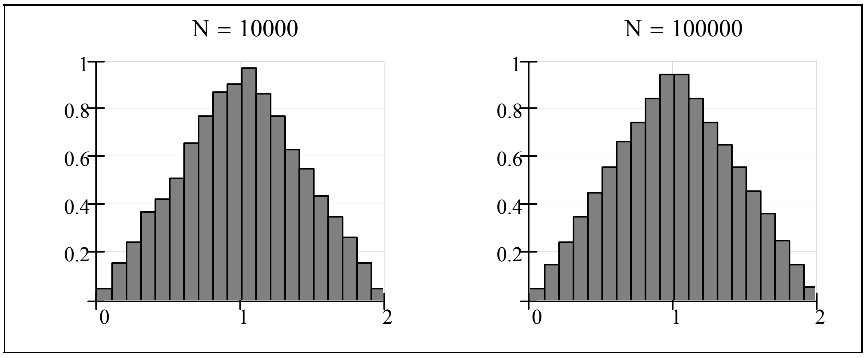
$N := 1000$	$n := 1 .. N$
$a_n := \mathrm{rnd}(1) + \mathrm{rnd}(1)$	$H := \mathrm{histogram}(20, a)$

Il comando "`H:=histogram(20,a)`" prende in ingresso il numero di intervalli in cui suddividere l'intervallo $[0, 2]$ e il vettore a; fornisce in uscita il vettore v delle frequenze assolute.

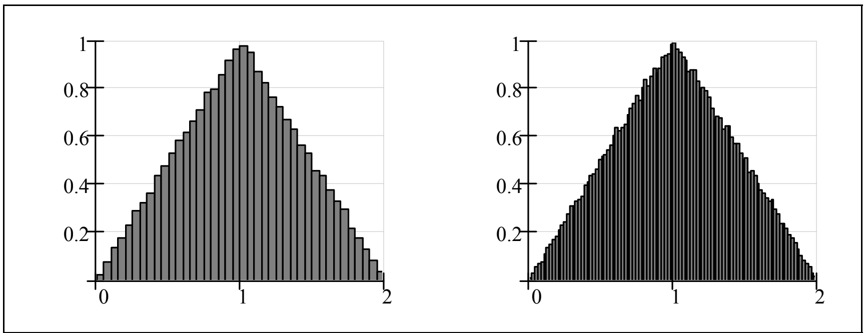


Per approssimare $F'(x)$ dividiamo le frequenze assolute (che variano da 0 a N) per N, ottenendo così le frequenze relative (che variano tra 0 e 1) e dividiamo ancora per h. I grafici seguenti si riferiscono a $N := 100$, 1000, 10000, 100000.





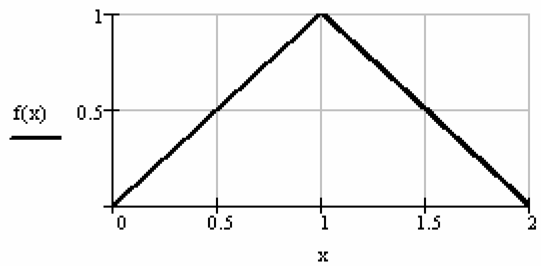
Al crescere di N i grafici mostrano senza ambiguità che se X e Y hanno distribuzione uniforme tra 0 e 1 allora la densità di $X+Y$ cresce linearmente da 0 a 1 e decresce linearmente da 1 a 2, con un punto angoloso in 1. Anche diminuendo l'ampiezza h degli intervalli il risultato è confermato; i grafici seguenti si riferiscono, con $N := 10^5$, a $h := 0.05$, 0.02.



Possiamo allora ipotizzare che l'espressione simbolica della funzione di densità di probabilità di $X+Y$, in $[0, 2]$, sia la seguente.

$$f(x) := \begin{cases} x & 0 \le x \le 1 \\ 2-x & 1 < x \le 2 \end{cases}.$$

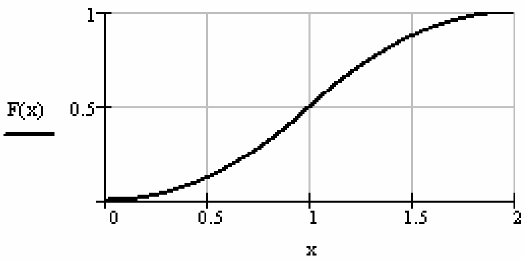
$$f(x) := \begin{array}{|l} x \text{ if } 0 \le x \le 1 \\ 2-x \text{ if } 1 < x \le 2 \\ 0 \text{ otherwise} \end{array}$$



Se tale congettura è corretta, poiché $F'(x) = f(x)$, allora in $[0, 1]$ $F(x)$ è la antiderivata di $f(x) = x$ tale che $F(0) = 0$, cioè $x^2/2$, mentre in $[1, 2]$ $F(x)$ è l'antiderivata di $f(x) = 2-x$ tale che $F(2) = 1$, cioè $2x - x^2/2 - 1$. In definitiva dovrebbe risultare:

$$F(x) := \begin{cases} 0 & x < 0 \\ \frac{1}{2}x^2 & 0 \le x \le 1 \\ 2x - \frac{1}{2}x^2 - 1 & 1 < x \le 2 \\ 1 & 2 < x \end{cases}.$$

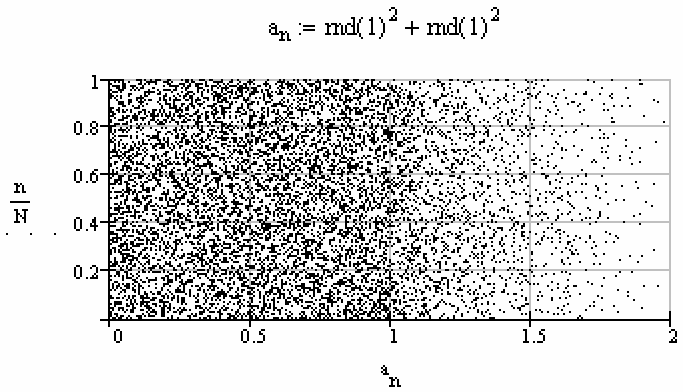
$$F(x) := \begin{cases} 0 & \text{if } x < 0 \\ \text{otherwise} & \\ \quad \begin{cases} \frac{1}{2}\cdot x^2 & \text{if } x \le 1 \\ \text{otherwise} & \\ \quad \begin{cases} 2\cdot x - \frac{1}{2}\cdot x^2 - 1 & \text{if } x \le 2 \\ 1 & \text{otherwise} \end{cases} \end{cases} \end{cases}$$



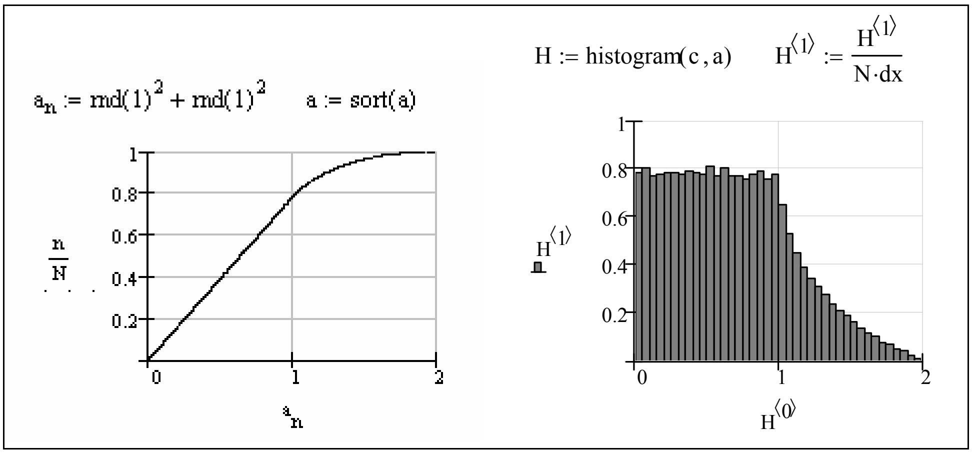
Otteniamo proprio lo stesso grafico di $F(x)$ che avevamo simulato con il comando `sort(a)`.

Abbiamo ora uno strumento molto potente: possiamo simulare la funzione di ripartizione e la funzione di densità di un numero casuale, comunque generato. Per curiosità, proviamo con

$$X^2 + Y^2 := \texttt{rnd(1)}^2 + \texttt{rnd(1)}^2.$$



Strano: la densità sembra costante da 0 a 1 e decrescente da 1 a 2. I grafici seguenti mostrano la distribuzione $F(x)$ e la densità $f(x)$, e confermano la prima impressione.



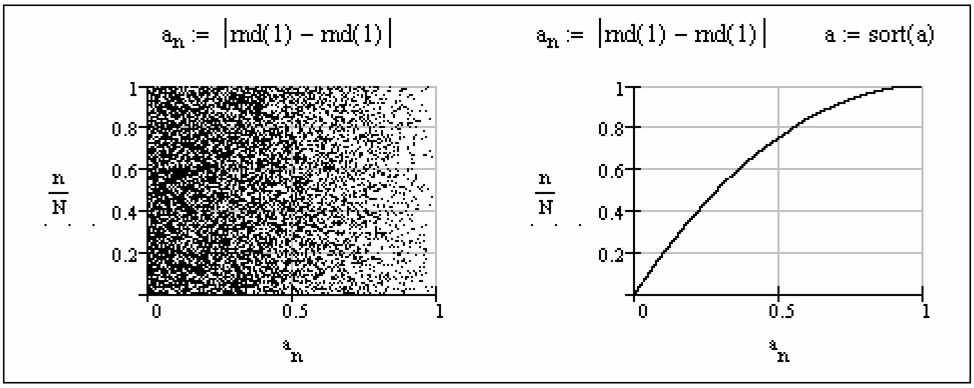
Le sorprese non mancano.

Due problemi

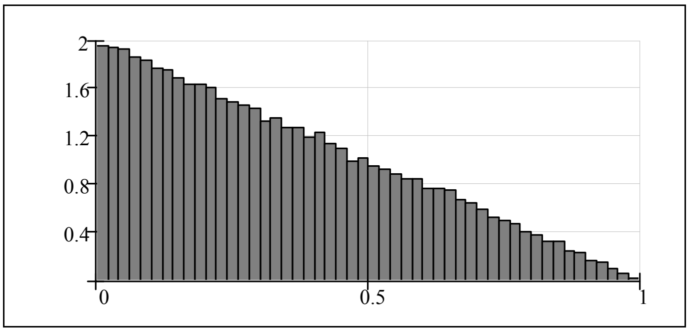
PROBLEMA 1. *Si scelgono due punti a caso sul segmento* [0, 1]. *Come si distribuisce la loro distanza?*

Possiamo simulare la distanza tra due punti scelti a caso (cioè con distribuzione uniforme) in [0, 1] con il comando `|rnd(1)-rnd(1)|`.

Tale numero varia tra 0 e 1; ci aspettiamo che la distribuzione non sia uniforme. Infatti è più probabile che la distanza sia vicino a 0 (è sufficiente che i due punti siano vicini, non importa dove) piuttosto che vicino a 1 (è necessario che un punto sia vicino a 0 e l'altro vicino a 1): l'intuito ci suggerisce che la densità di probabilità sia decrescente da 0 a 1. Vediamo i grafici della nuvola di punti e della funzione di ripartizione.



In effetti la pendenza di $F(x)$ va decrescendo da 0 a 1; di che curva si tratta? Proviamo a simulare la funzione di densità $f(x)$.



La densità sembra decrescere linearmente nell'intervallo [0, 1]; risulta in effetti

$$f(x) := 2 - 2x, \qquad F(x) := 2x - x^2$$

Così, ad esempio, la probabilità che il numero casuale $|X - Y|$ sia compreso tra 0 e 0.1 è

$$\int_0^{0.1} (2-2x)\,dx = F(0.1) - F(0) = 0.19$$

mentre la probabilità che sia compreso tra 0.9 e 1 è

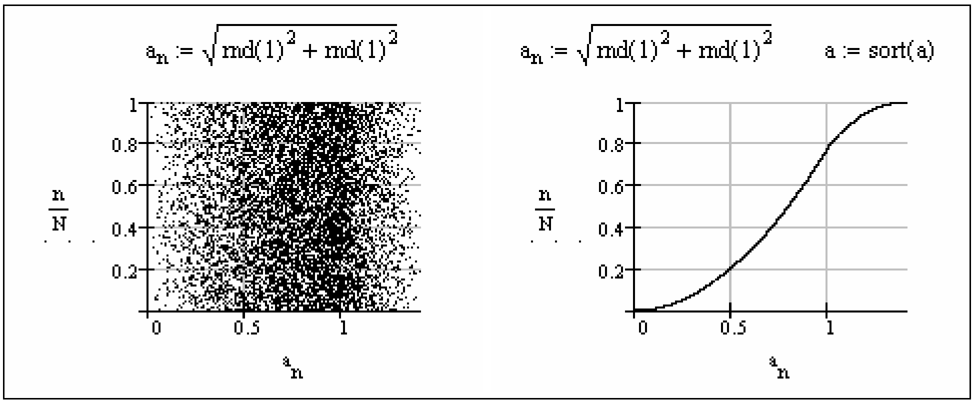
$$\int_{0.9}^{1} (2-2x)\,dx = F(1) - F(0.9) = 0.01.$$

PROBLEMA 2. *Si sceglie a caso un punto nel quadrato* $[0, 1]\times[0, 1]$. *Come si distribuisce la sua distanza dall'origine?*

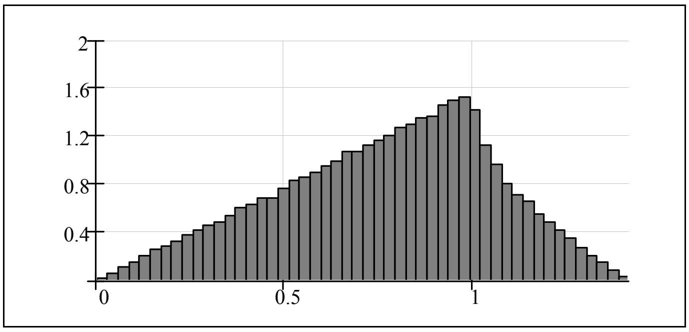
Scegliere un punto caso nel quadrato unitario è equivalente a scegliere l'ascissa X e l'ordinata Y con distribuzione uniforme in [0, 1]; possiamo simulare la distanza del punto dall'origine con il comando

$$\left|\sqrt{\mathtt{rnd(1)}^2 + \mathtt{rnd(1)}^2}\right|.$$

Tale numero casuale varia tra 0 e $\sqrt{2}$; la distribuzione certamente non è uniforme: è difficile che la distanza sia vicino a 0 (occorre che entrambe le coordinate siano vicine a 0), oppure che sia vicino a $\sqrt{2}$ (occorre che entrambe le coordinate siano vicine a 1). Proviamo a tracciare la nuvola di punti e a simulare la funzione di ripartizione.



Nel grafico di sinistra si osserva che la densità massima è in 1; questo è confermato dal grafico di $F(x)$, la cui pendenza passa da crescente a decrescente in 1. Simuliamo la funzione di densità.

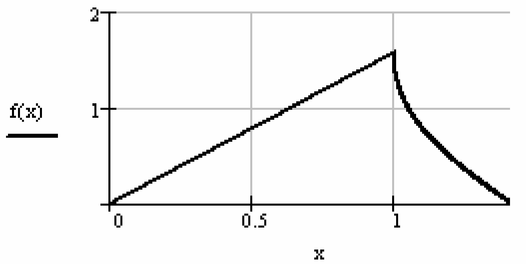


La figura precedente mostra che $f(x)$ cresce linearmente da 0 a 1, con pendenza $\pi/2$, e decresce da 1 a $\sqrt{2}$. In effetti il grafico di $F(x)$, in $[0, 1]$, è quello di $\pi x^2/4$, l'antiderivata di $\pi x/2$ che vale 0 in 0. Assai arduo è invece calcolare l'espressione simbolica di $f(x)$ nell'intervallo $[1, \sqrt{2}\,]$ e ci accontentiamo di un'approssimazione numerica. Per i più curiosi si tratta della funzione

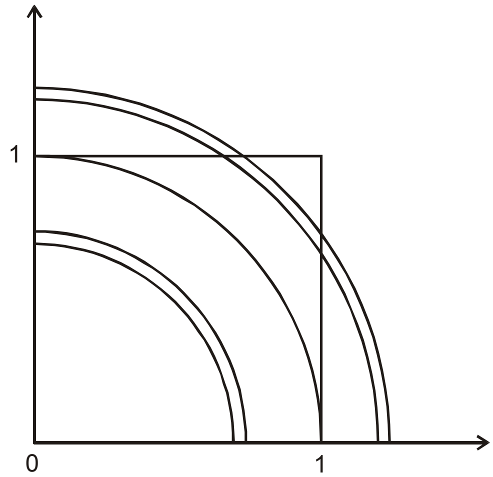
$$x \arcsin\left(\frac{2-x^2}{x^2}\right),$$

come mostra la seguente figura.

$$f(x) := \begin{cases} \frac{\pi}{2}\cdot x & \text{if } 0 \le x \le 1 \\ x\cdot \text{asin}\left(\frac{2-x^2}{x^2}\right) & \text{if } 1 < x \le \sqrt{2} \\ 0 & \text{otherwise} \end{cases}$$



Il significato di questo risultato apparentemente strano è illustrato dalla seguente figura: la probabilità di trovare un punto a distanza dall'origine compresa tra d e $d+h$ aumenta linearmente all'aumentare di d fino a 1, insieme all'area della striscia circolare compresa tra d e $d+h$.



Invece se $d > 1$, allora l'area della striscia circolare contenuta nel quadrato decresce rapidamente fino a 0.

Conclusioni

Nel percorso illustrato non si è fatto cenno alla definizione formale di "variabile aleatoria continua" (che non è stata nemmeno nominata, si è preferito utilizzare l'espressione "numero casuale"), lasciando che il concetto emergesse lentamente dal contesto e dai problemi proposti.

Gli unici strumenti indispensabili per questo tipo di approccio sono un generatore di numeri casuali, un comando per l'ordinamento di un vettore e un comando per il calcolo delle frequenze, tutti strumenti che si trovano anche in Excel o in qualunque software di matematica o di statistica.

Nel lavoro proposto la simulazione diventa strumento di approccio a nuovi concetti e al tempo stesso strumento di scoperta.

Si osservi che il problema del calcolo simbolico della funzione di ripartizione, o della funzione di densità, è intrinsecamente molto complesso, anche per variabili casuali semplici; la soluzione, come mostrato nei problemi proposti, è spesso assai laboriosa da calcolare e non è raro il caso in cui essa non sia esprimibile simbolicamente come funzione elementare. Di conseguenza la simulazione diventa anche uno strumento di approssimazione, secondo il più classico approccio che va sotto il nome di "metodo Montecarlo".

Una volta impostato il foglio di lavoro è sufficiente modificare la definizione del numero casuale per ottenere, a cascata, tutti i risultati, almeno dal punto di vista numerico.

Non resta che sottolineare il fatto che, almeno a livello secondario, un approccio mediante simulazione al concetto di variabile aleatoria continua permette di superare le note difficoltà didattiche legate alla definizione di "funzione di densità di probabilità" (i cui valori non sono delle probabilità) e di mostrare un contesto significativo in cui utilizzare l'integrazione definita e indefinita.

I grafici che si ottengono offrono spesso sorprese e costringono ad una interpretazione semantica non banale. La ricchezza di registri linguistici, numerici e grafici che occorre controllare e confrontare fa sì che il percorso mostrato possa essere remunerativo dal punto di vista didattico.