Variabili aleatorie

Se il risultato di una misura è *x* (*x* individua un elemento nello spazio campione e come tale prende il nome di **variabile aleatoria**), la **funzione densità di probabilità *f(x)*** è definita dalla relazione seguente:

$$P\left(x \in \left[x, x + dx\right]\right) = f(x)dx$$

La condizione di normalizzazione è: $\int_{-\infty}^{+\infty} f(x)dx = 1$

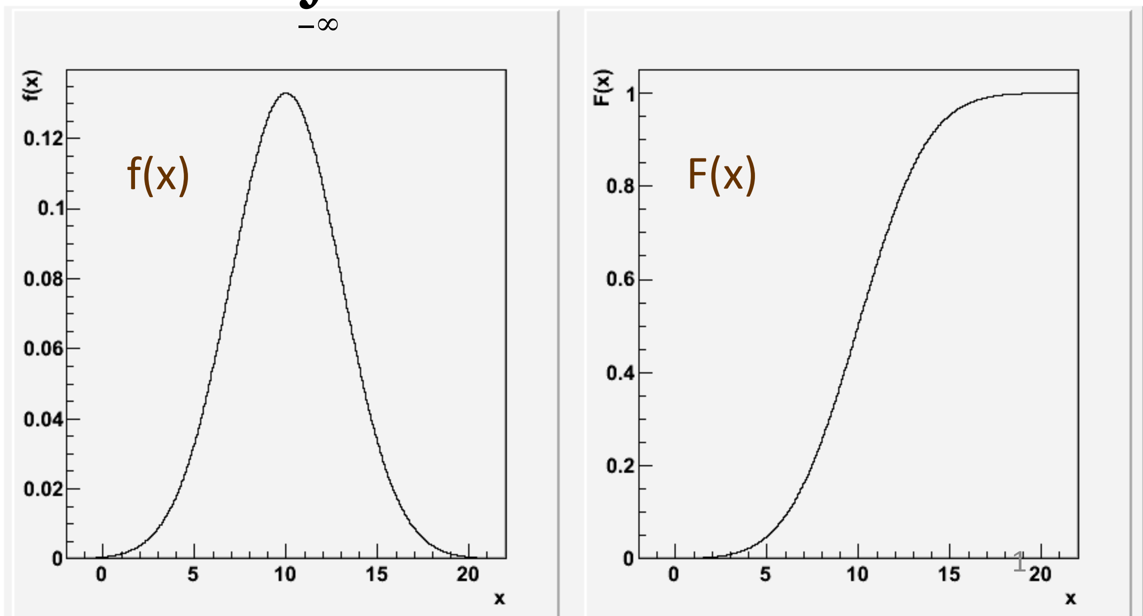
La funzione cumulativa di probabilità è $F(x) = \int_{-\infty}^{x} f(u)du$

Nel caso discreto:

$$f_i = P(x_i)$$

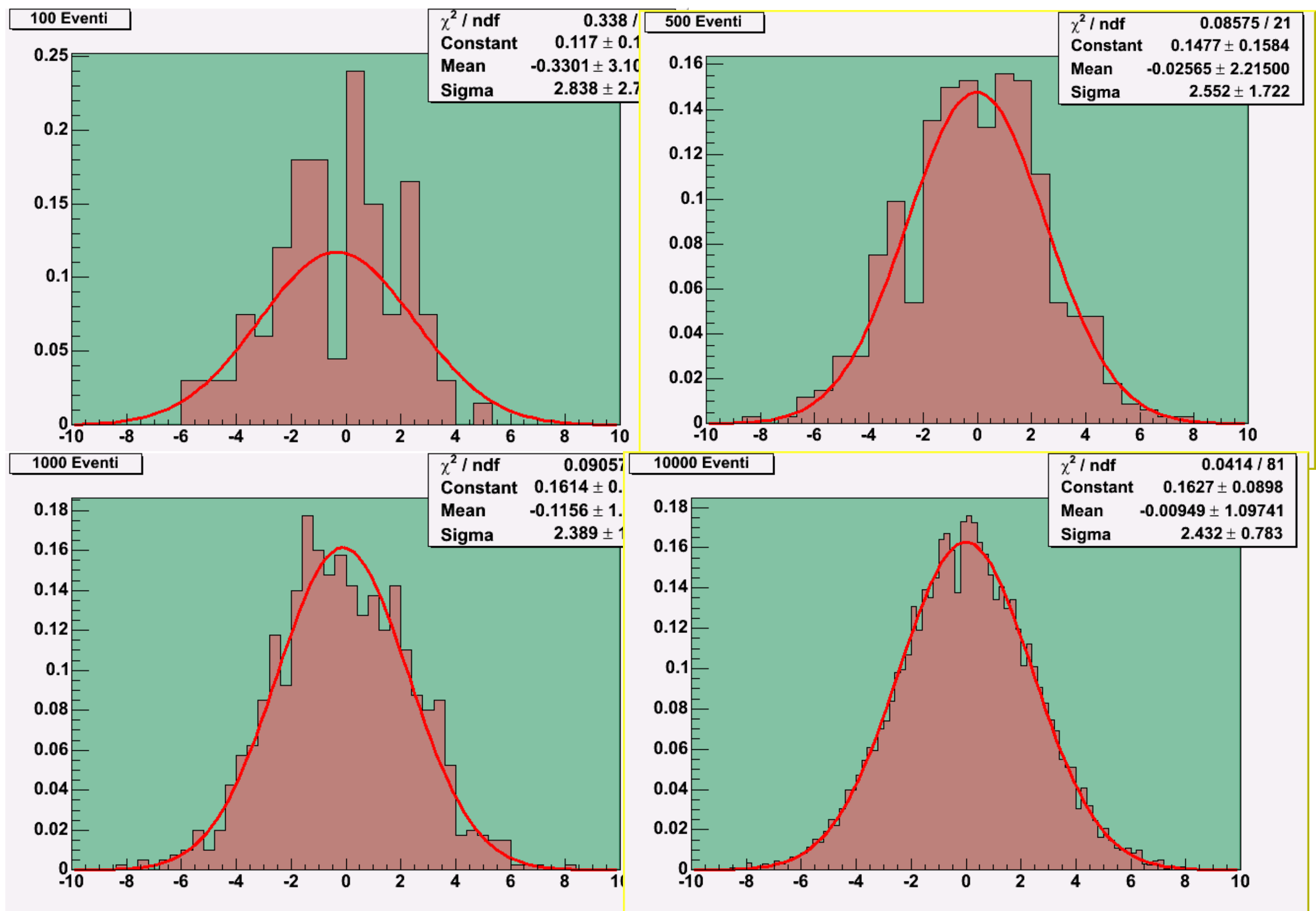
$$\sum_i f_i = 1$$

$$F(x) = \sum_{x_i < x} f_i$$



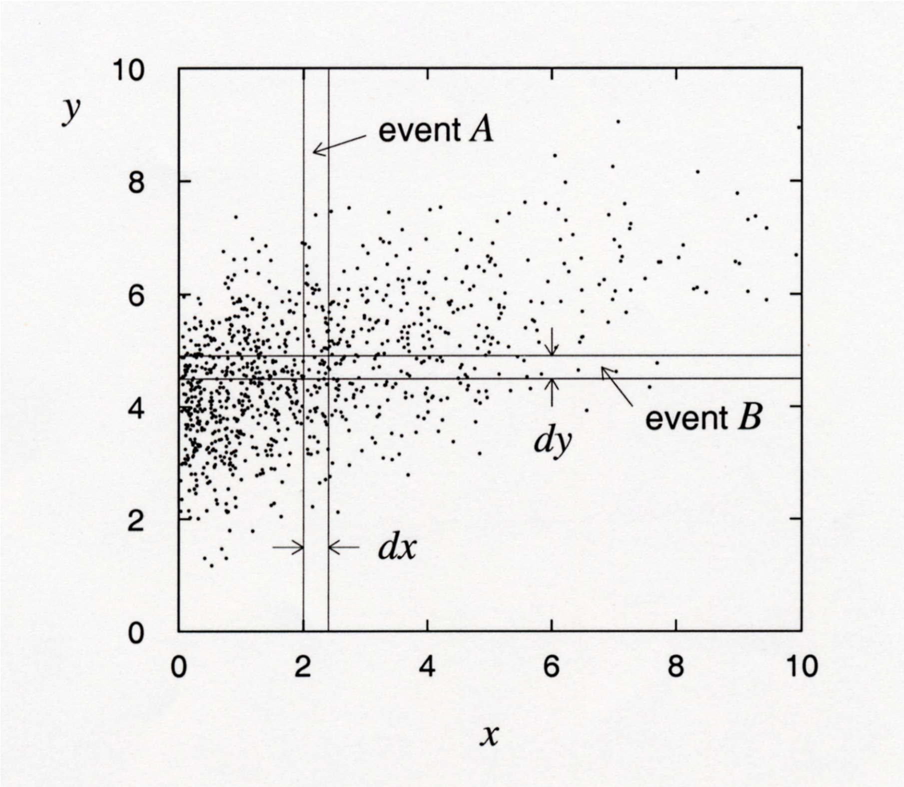
I risultati delle misure vengono spesso rappresentati mediante istogrammi, che, se normalizzati ad area unitaria, tendono a rappresentare la densità di probabilità nel limite di infinito numero di misure e di larghezza di ciascun elemento (cella o “bin”) dell’istogramma nulla

Nell’esempio riportato in figura, al crescere del campione, la distribuzione tende ad assumere la forma di una gaussiana



Una misura può condurre alla determinazione di un set di parametri: la definizione di densità di probabilità può essere facilmente estesa al caso **multivariato**.

Nel caso bidimensionale *f(x,y)dxdy* è la probabilità che *x* sia compreso tra *x* e *x+dx* e che *y* sia compreso tra *y* e *y+dy*



$$P(A \cap B) = f(x,y)dxdy$$

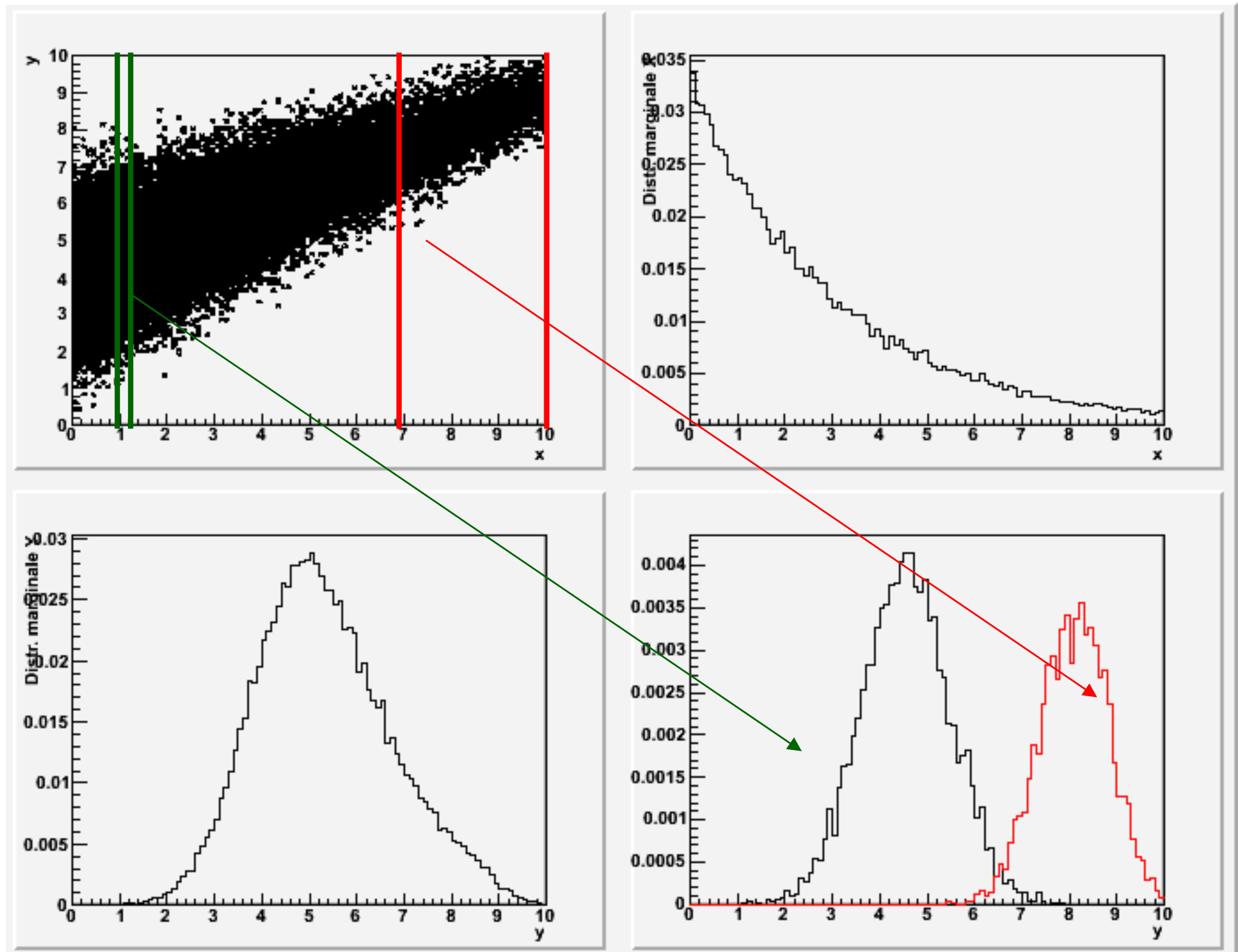
$$\iint f(x,y)dxdy = 1$$

$f(x_1,x_2,x_3,\dots x_n)$ è detta distribuzione congiunta di probabilità

La proiezione della distribuzione congiunta su uno degli assi x_i è detta **distribuzione marginale** di probabilità

$$f_x(x) = \int f(x,y)\,dy$$

$$f_y(y) = \int f(x,y)\,dx$$



bivar.C – emacs@ali26xl.to.infn.it

File Edit Options Buffers Tools C++ Help

```
void bivar(){
  TH2F *bv = new TH2F("bv","bivariato",100,0.,10.,100,0.,10.);
  for(Int_t i =1; i<50000; i++){
    Float_t x=gRandom->Exp(3.);
    Float_t mea=x*0.5+4.;
    Float_t sig=1.-0.05*x;
    if(sig<=0.1)sig=0.1;
    Float_t y=gRandom->Gaus(mea,sig);
    bv->Fill(x,y,1./50000);
  }
  gStyle->SetOptStat(kFALSE);
  gStyle->SetOptTitle(kFALSE);
  bv->GetXaxis()->SetTitle("x");
  bv->GetYaxis()->SetTitle("y");
  TCanvas *cv =  new TCanvas();
  cv->Divide(2,2);
  cv->cd(1);
  bv->Draw();
  cv->cd(2);
  TH1D *px=bv->ProjectionX();
  px->GetYaxis()->SetTitle("Distr. marginale X");
  px->Draw();
  cv->cd(3);
  TH1D *py=bv->ProjectionY();
  py->GetYaxis()->SetTitle("Distr. marginale Y");
  py->Draw();
  cv->cd(4);
  TH1D *uno = bv->ProjectionY("uno",10,13);
  TH1D *due = bv->ProjectionY("due",70,100);
  uno->Draw();
  due->SetLineColor(2);
  due->Draw("same");

}
```

--:-- **bivar.C** (C++ Abbrev)--L16--All------------------------------------

Beginning of buffer

Dalla definizione di probabilità subordinata $\quad P(B \mid A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{f(x,y)dxdy}{f_x(x)dx}$

Segue la definizione di **densità condizionale di probabilità** (e.g. di y ad x fisso o viceversa):

$$h(y \mid x) = \frac{f(x,y)}{f_x(x)} \qquad e \qquad g(x \mid y) = \frac{f(x,y)}{f_y(y)}$$

Se le variabili aleatorie x e y sono **indipendenti**, si ha

$$\boxed{h(y \mid x) = f_y(y) \quad e \quad g(x \mid y) = f_x(x) \quad \Rightarrow \quad f(x,y) = f_x(x) f_y(y)}$$

Il teorema di Bayes si può riscrivere come:

$$\boxed{g(x \mid y) = \frac{h(y \mid x) f_x(x)}{f_y(y)}}$$

Funzioni di variabili aleatorie

Una funzione *a(x)* di una variabile aleatoria *x* (con distribuzione di probabilità *f(x)*) è essa stessa una variabile aleatoria.

- Qual è la distribuzione di probabilità di *a*? [indicata con *g(a)*]

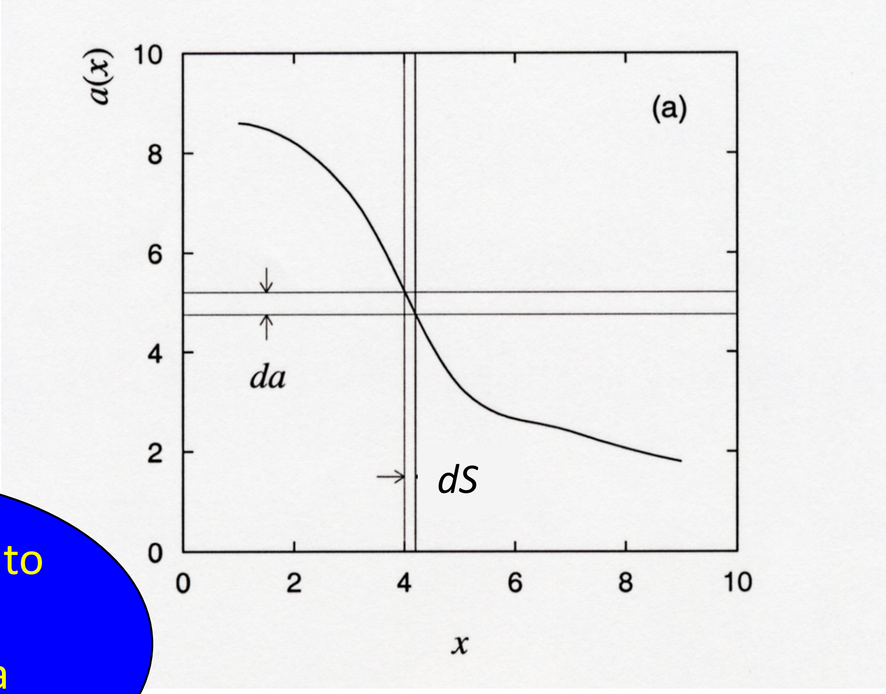
Consideriamo un dominio *dS* di *x* per cui $a \in [a, a+da]$ e supponiamo *a* monotona:

$$g(a)da = \int_{dS} f(x)dx \cong$$

$$\cong f[x(a)]\left[x(a) + \left|\frac{dx}{da}\right|da - x(a)\right] \Rightarrow$$

$$\boxed{g(a) = f[x(a)]\left|\frac{dx}{da}\right|}$$

Il valore assoluto assicura che l'integrale sia positivo



Caso multivariato

Nel caso di una funzione di n variabili aleatorie $a(x_1, x_2, \ldots, x_n)$, la densità congiunta di probabilità è

$$g(a)\,da = \int \ldots \int_{dS} f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n$$

Come esempio si considerino 2 v.a. indipendenti distribuite secondo due f.d.p. $g(x)$ e $h(y)$. La densità congiunta di probabilità è: $g(x)h(y)$

Si vuole trovare la densità di probabilità del prodotto $z=xy$

$$f(z)dz = \int \int_{dS} g(x)h(y)dxdy$$

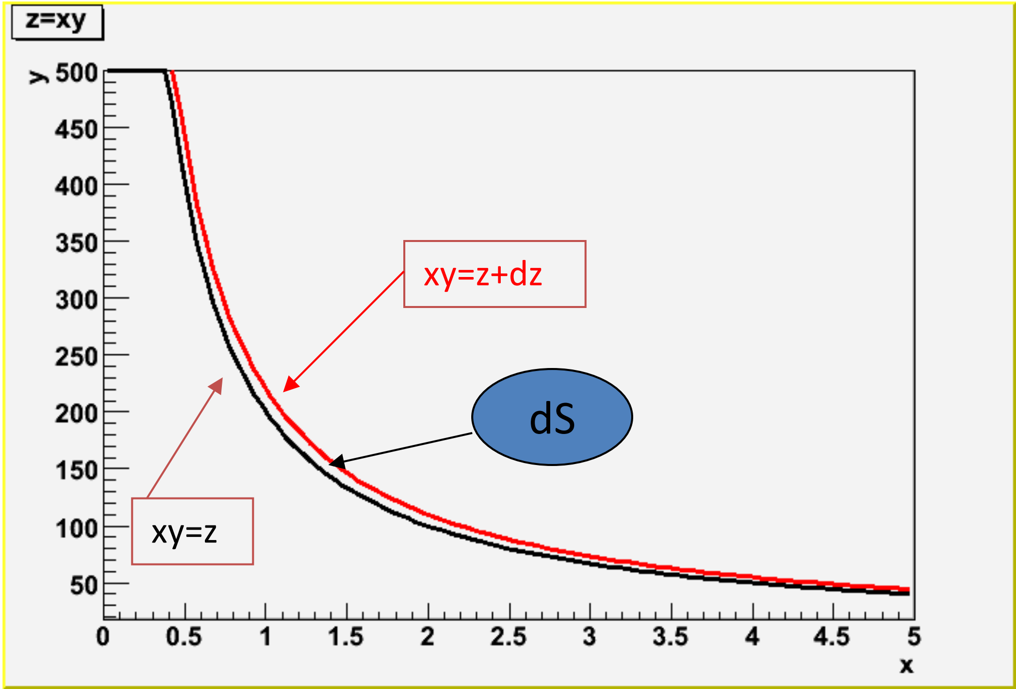
$$= \int_{-\infty}^{\infty} g(x)\,dx \int_{z/x}^{z/x+dz/|x|} h(y)dy$$

da cui segue

$$\boxed{f(z) = \int_{-\infty}^{+\infty} g(x)h(z/x)\frac{dx}{|x|}}$$

oppure

$$\boxed{f(z) = \int_{-\infty}^{+\infty} g(z/y)h(y)\frac{dy}{|y|}}$$



Convoluzione

Il risultato raggiunto nell'ultima slide per la densità di probabilità del prodotto di due variabili aleatorie indipendenti prende il nome di **convoluzione di Mellin** e viene indicata spesso con $f = g \otimes h$.

Analogamente, nel caso della somma di due v.a. indipendenti: $z=x+y$ si ottiene:

$$f(z) = \int_{-\infty}^{+\infty} g(x)h(z-x)dx = \int_{-\infty}^{+\infty} g(z-y)h(y)dy$$

Questa rappresenta la **convoluzione di Fourier** delle due funzione h e g e viene ancora indicata con la notazione $f = g \otimes h$

Se *a(x)* non ammette un unico inverso (e.g. $a=x^2$), *dS* va esteso a tutti gli intervalli di *x* che corrispondono a *da*

Caso multivariato-bis: $f(\vec{x}),\quad dove\quad \vec{x}=(x_1,x_2,\cdots,x_n)$

Consideriamo n funzioni linearmente indipendenti:

$$\vec{y}(\vec{x})=\left(y_1(\vec{x}),y_2(\vec{x}),\cdots,y_n(\vec{x})\right)$$

Che siano invertibili: $x_1(\vec{y}),x_2(\vec{y}),\cdots,x_n(\vec{y})$

La distr. congiunta di probabilità di ***y*** è:

$$\boxed{g(\vec{y})=|J|f(\vec{x})}$$

Dove *J* è il determinante jacobiano della trasformazione

$$J=\begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} & \cdots & \dfrac{\partial x_1}{\partial y_n} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} & \cdots & \dfrac{\partial x_2}{\partial y_n} \\ \vdots & & & \vdots \\ & & \cdots & \dfrac{\partial x_n}{\partial y_n} \end{vmatrix}$$

Momenti di una distribuzione

Il momento di ordine *n* di una variabile aleatoria *x* è:

$$\alpha_n \equiv E\left[x^n\right] = \int_{-\infty}^{+\infty} x^n f(x)dx$$

Il momento centrale di ordine *n* (o momento rispetto alla media) è:

$$m_n \equiv E\left[\left(x-\alpha_1\right)^n\right] = \int_{-\infty}^{+\infty} \left(x-\alpha_1\right)^n f(x)dx$$

Caveat: i momenti non sono una funzione di *x*, ma *parametri* della densità di probabilità *f(x)*

Media e varianza

I momenti più utilizzati sono la media e la varianza, definiti rispettivamente come:

$\mu \equiv \alpha_1$ e

$$\sigma^2 \equiv V[x] = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx = \int_{-\infty}^{+\infty} x^2 f(x)dx + \mu^2 \int_{-\infty}^{+\infty} f(x)dx - 2\mu \int_{-\infty}^{+\infty} x f(x)dx$$

$$= \alpha_2 - \mu^2$$

La media è un'espressione del "*centro di massa*" della distribuzione di probabilità mentre la varianza è il quadrato della sua larghezza. Come stima di dispersione si usa in generale la deviazione standard $\sigma \equiv \sqrt{\sigma^2}$.

Covarianza e correlazione

Nel caso di una distribuzione bivariata, il grado di correlazione tra le variabili aleatorie *x* e *y* è espresso dalla covarianza:

$$\mathrm{cov}[x,y] \equiv V_{xy} \equiv E\left[(x-\mu_x)(y-\mu_y)\right] = E[xy] - \mu_x\mu_y$$

Nel caso multivariato, V_{xy} prende il nome di matrice di covarianza o matrice degli errori. Il coefficiente adimensionale di correlazione è definito come:

$$\rho_{xy} = \frac{V_{xy}}{\sigma_x \sigma_y} \qquad -1 \le \rho_{xy} \le 1$$

Si veda appendice C del testo di M.Loreti per la dimostrazione

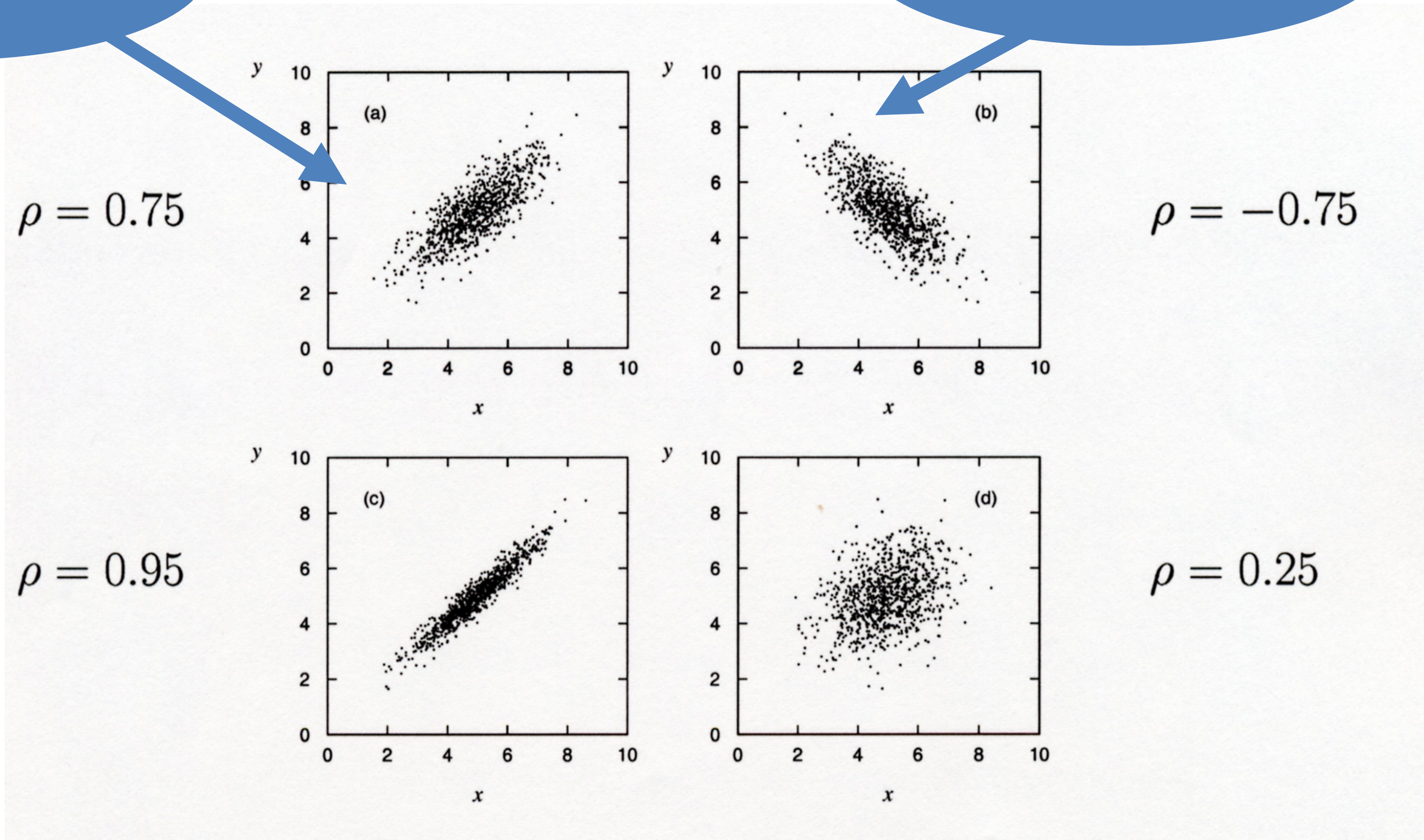
Se x e y sono indipendenti: $E[xy] = \mu_x\mu_y \quad \Rightarrow \quad \mathrm{cov}[x,y] = 0$

Caveat: Non è detto che 2 variabili scorrelate siano indipendenti !

Esempi di correlazione

Correlazione positiva

Correlazione negativa



Propagazione degli errori (1)

Supponiamo che un set di variabili aleatorie $\vec{x} = (x_1, \cdots, x_n)$ segua una distribuzione congiunta di probabilità $f(\vec{x})$ eventualmente non nota, ma che si abbiano stime delle covarianze $V_{ij} = \mathrm{cov}[x_i, x_j]$ e delle medie $\vec{\mu} = E[\vec{x}]$.

Consideriamo una funzione $\boxed{y(\vec{x})}$

- Qual è la media di *y* ?
- Qual è la varianza di *y* ? ($V[y] = E[y^2] - E[y]^2$)

Il momento di una funzione *y* della variabile aleatoria ***x***, vale:

$$E[y(\vec{x})] = \int_{-\infty}^{\infty} y\, g(y)\, dy = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} y(\vec{x}) f(\vec{x})\, dx_1 \cdots dx_n = \mu_y$$

Propagazione degli errori (2)

Se y non è nota, si può usare l'espansione al prim'ordine di *y* attorno a $\boldsymbol{\mu}$**:** (vedremo dopo i limiti di questo assunto)

$$\boxed{y(\vec{x}) \cong y(\vec{\mu}) + \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} (x_i - \mu_i)}$$

In altri termini:

$$E[x_i - \mu_i] = 0$$

$$E[y(\vec{x})] \cong E\left[y(\vec{\mu}) + \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} (x_i - \mu_i) \right] =$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left\{ y(\vec{\mu}) + \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} (x_i - \mu_i) \right\} f(\vec{x})\, dx_1 \cdots dx_n$$

$$= y(\vec{\mu}) + \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i f(\vec{x})\, dx_1 \cdots dx_n - \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} \mu_i = y(\vec{\mu})$$

Propagazione degli errori (3)

$$y(\vec{x}) \cong y(\vec{\mu}) + \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} (x_i - \mu_i) \Rightarrow \begin{cases} E\left[y^2(\vec{x})\right] \cong y^2(\vec{\mu}) + \cancel{2y(\vec{\mu})\sum_{i=1}^{n}\left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} E\left[x_i - \mu_i\right]} + \\ + E\left[\left(\sum_{i=1}^{n}\left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} (x_i - \mu_i)\right)\left(\sum_{j=1}^{n}\left[\frac{\partial y}{\partial x_j}\right]_{\vec{x}=\vec{\mu}} (x_j - \mu_j)\right)\right] \Rightarrow \\ \boxed{E\left[y^2(\vec{x})\right] = y^2(\vec{\mu}) + \sum_{i,j=1}^{n}\left[\frac{\partial y}{\partial x_i}\frac{\partial y}{\partial x_j}\right]_{\vec{x}=\vec{\mu}} V_{ij}} \end{cases}$$

Espansione al primo ordine

Propagazione degli errori (4)

La varianza di ***y*** è quindi:

$$\boxed{\sigma_y^2 = E\left[y^2\right] - E\left[y\right]^2 \cong \sum_{i,j=1}^{n} \left[\frac{\partial y}{\partial x_i}\frac{\partial y}{\partial x_j}\right]_{\vec{x}=\vec{\mu}} V_{ij}}$$

Se le x_i sono tra loro scorrelate, si ha:

$$\sigma_y^2 \cong \sum_{i,j=1}^{n} \left[\frac{\partial y}{\partial x_i}\frac{\partial y}{\partial x_j}\right]_{\vec{x}=\vec{\mu}} V_{ij} = \sum_{i,j=1}^{n} \left[\frac{\partial y}{\partial x_i}\frac{\partial y}{\partial x_j}\right]_{\vec{x}=\vec{\mu}} \sigma_i^2 \delta_{ij} = \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}}^2 \sigma_i^2$$

Propagazione degli errori (5)

Nel caso in cui si abbia, in luogo di una funzione reale *y*, un set di funzioni $\vec{y}(\vec{x}) = \left(y_1(\vec{x}), y_2(\vec{x}), \cdots, y_n(\vec{x})\right)$ le formule precedenti possono essere estese come:

$$U_{kl} = \mathrm{cov}\left[y_k, y_l\right] \cong \sum_{i,j=1}^{n} \left[\frac{\partial y_k}{\partial x_i}\frac{\partial y_l}{\partial x_j}\right]_{\vec{x}=\vec{\mu}} V_{ij}$$

O, in notazione matriciale:

$$\mathbf{U} = \mathbf{A}\mathbf{V}\mathbf{A}^{\mathbf{T}} \quad \text{dove} \quad A_{ij} = \left[\frac{\partial y_i}{\partial x_j}\right]_{\vec{x}=\vec{\mu}}$$

Propagazione degli errori (6)

- Queste relazioni, che propagano le covarianze di un set di variabili aleatorie a variabili aleatorie derivate, sono le formule di propagazione degli errori, in quanto varianze e covarianze sono usate generalmente per esprimere gli errori delle grandezze oggetto di misura.
- Sono valide solo se $\vec{y}(\vec{x})$ è lineare. L'approssimazione non è valida se la funzione non è lineare in un dominio di estensione confrontabile con σ_i

Esempi (1)

$$y = x_1 + x_2 \quad \Rightarrow \quad \boxed{\sigma_y^2 = \sigma_1^2 + \sigma_2^2 + 2 \cdot \operatorname{cov}[x_1, x_2]}$$

$$y = x_1 x_2 \quad \Rightarrow \quad \sigma_y^2 = x_2^2 \sigma_1^2 + x_1^2 \sigma_2^2 + 2 \cdot x_1 x_2 \operatorname{cov}[x_1, x_2]$$

$$\boxed{\frac{\sigma_y^2}{y^2} = \frac{\sigma_1^2}{x_1^2} + \frac{\sigma_2^2}{x_2^2} + 2\frac{\operatorname{cov}[x_1, x_2]}{x_1 x_2}}$$

$$y = \frac{x_1}{x_2} \quad \Rightarrow \quad \sigma_y^2 = \frac{1}{x_2^2}\sigma_1^2 + \frac{x_1^2}{x_2^4}\sigma_2^2 - 2 \cdot \frac{x_1}{x_2^3}\operatorname{cov}[x_1, x_2]$$

$$\boxed{\frac{\sigma_y^2}{y^2} = \frac{\sigma_1^2}{x_1^2} + \frac{\sigma_2^2}{x_2^2} - 2\frac{\operatorname{cov}[x_1, x_2]}{x_1 x_2}}$$

Esempi (2)

- Nel caso di x_i scorrelate, gli errori si sommano in quadratura per la somma o la differenza
- Nel caso di x_i scorrelate, gli errori relativi si sommano in quadratura per il prodotto o il rapporto.