RANDOM VARIABLES

Definitions

A variable is defined as random, or stochastic, when it describes mathematically, in accordance with probability laws, the numerical outcomes of experiments related to random phenomena.
A continuous random variable can assume continuous values on the real axis.
A discrete random variable can assume values in a discrete set of numbers.
A mixed random variable can assume both continuous and discrete values.
The following notes focus on continuous random variables.

Distribution function

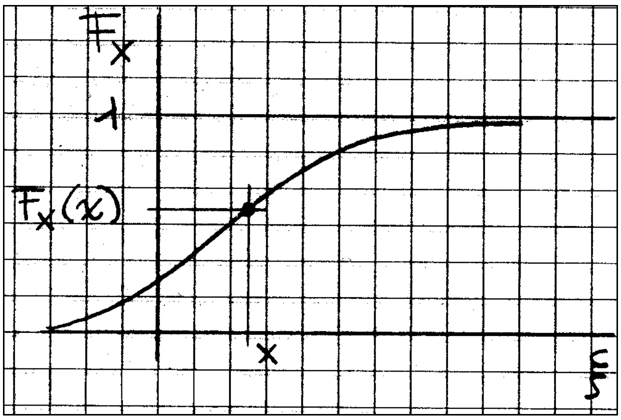
The distribution function, also called cumulative distribution, $F_X(x)$, is the probability that the random variable X assumes values less or equal to x:

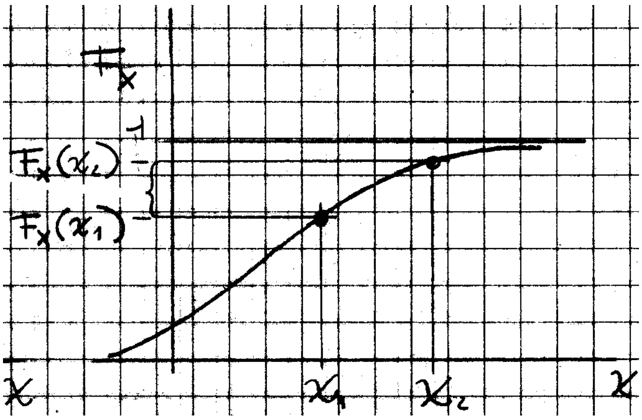
$$\boxed{F_X(x) = P(X \leq x)} \tag{1}$$

This function is always defined on the whole real axis $(-\infty < x < +\infty)$; x is called the state variable.
The distribution function has several noteworthy properties. In particular:

$$\boxed{F_X(-\infty) = P(X \leq -\infty) = P(0) = 0}$$
$$\boxed{F_X(+\infty) = P(X \leq +\infty) = 1}$$
$$\boxed{P\{x_1 < X \leq x_2\} = F_X(x_2) - F_X(x_1)}$$

$P\{x_1 < X \leq x_2\} \geq 0 \Rightarrow F_X(x_1) \leq F_X(x_2)$. So, $F_X(x)$ is a not decreasing function.





Density function

The density function $p_X(x)$, or simply the density of a random variable X, is the prime derivative of the distribution function with respect to x:

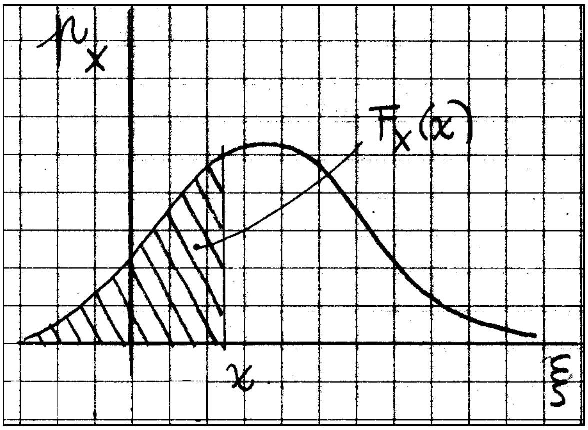
$$\boxed{p_X(x) = \frac{dF_X(x)}{dx}}$$

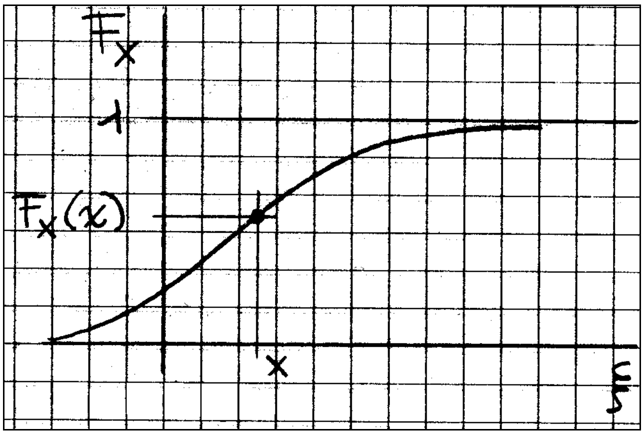
Since $F_X(x)$ is a not decreasing function, $p_X(x)$ is not negative:

$$p_X(x) \geq 0$$

Remembering that $F_X(-\infty)=0$, the application of the fundamental theorem of the integral calculus provides the relationship:

$$\boxed{F_X(x)=\int_{-\infty}^{x} p_x(\xi)\,d\xi}$$





It derives:

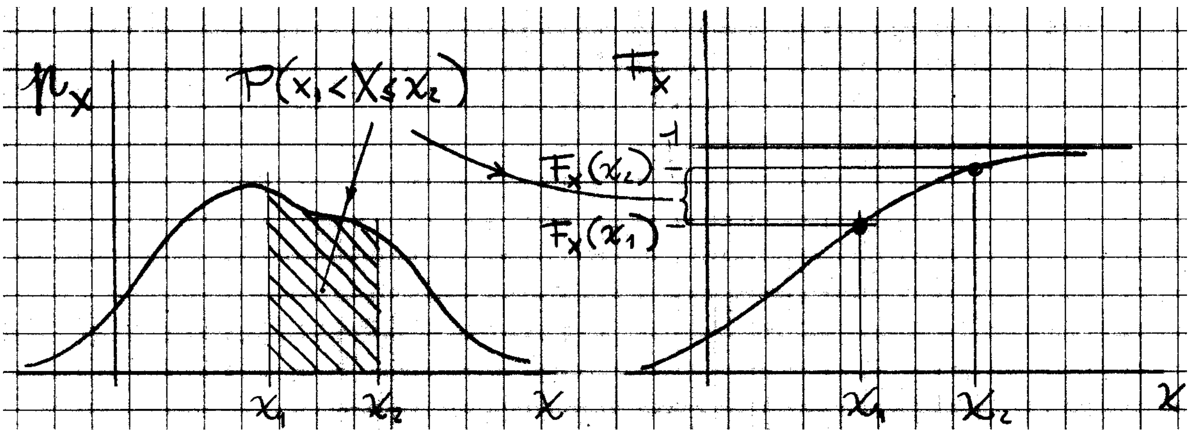
$$F_X(x_2)-F_X(x_1)=\int_{x_1}^{x2} p_x(x)\,dx$$

and then:

$$\boxed{P(x_1 < X \leq x_2)=\int_{x_1}^{x_2} p_X(x)\,dx}$$

$$\boxed{\int_{-\infty}^{\infty_2} p_X(x)\,dx=1}$$

which represent fundamental properties of the density function.



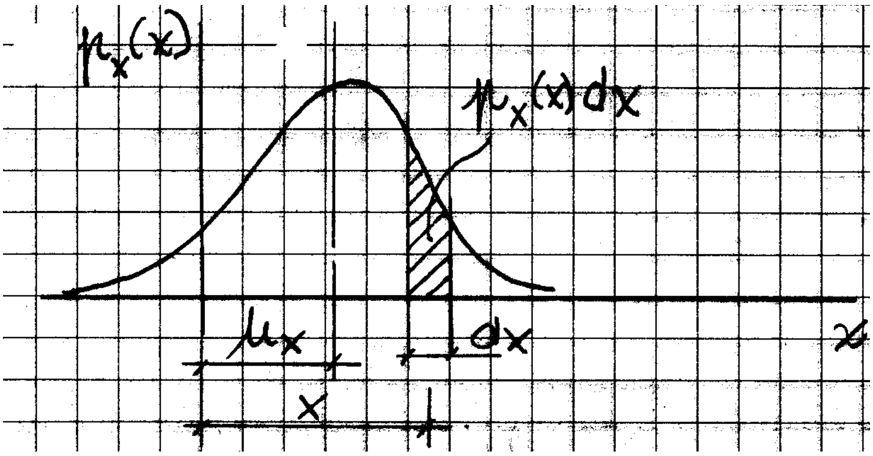
Indexes of a random variable

The distribution function $F_X(x)$ and the density function $p_X(x)$ give a complete description of X. The indexes of a random variable are quantities that provide a synthetic description of the random variable. Indexes are framed into 3 classes:

(1) Position indexes – provide the position of the distribution, i.e. the position of the values of x assumed by X; the most important position index is the mean value.
(2) Variability indexes – characterise the variability of the values assumed by X; they comprehend the mean square value, the variance, the standard deviation and the coefficient of variation.
(3) Shape indexes – provide an information on the shape of the distribution; they comprehend the skewness and the kurtosis.

Let us define as the mean value of the random variable X the quantity:

$$\boxed{\mu_X = \int_{-\infty}^{\infty} x p_X(x) dx}$$



Let us define as root mean square value of the random variable X the quantity:

$$\boxed{\varphi_X^2 = \int_{-\infty}^{\infty} x^2 p_X(x) dx}$$

Let us define as variance of the random variable X the quantity:

$$\boxed{\sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_x)^2 p_X(x) dx}$$

Developing the above equations:

$$\boxed{\sigma_X^2 = \varphi_X^2 - \mu_X^2}$$

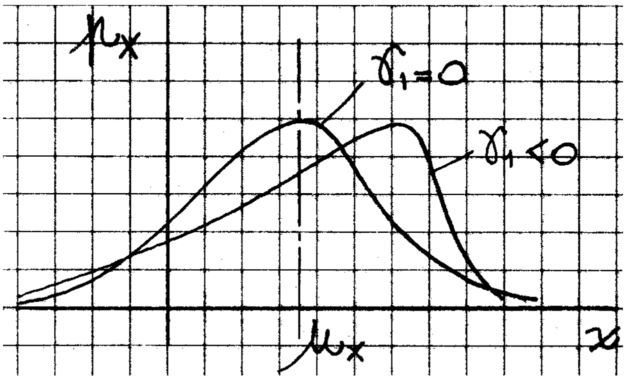
The standard deviation σ_X is the square root of the variance σ_X^2. It defines the dispersion of X around μ_X. Let us define as coefficient of variation the nondimensional ratio:

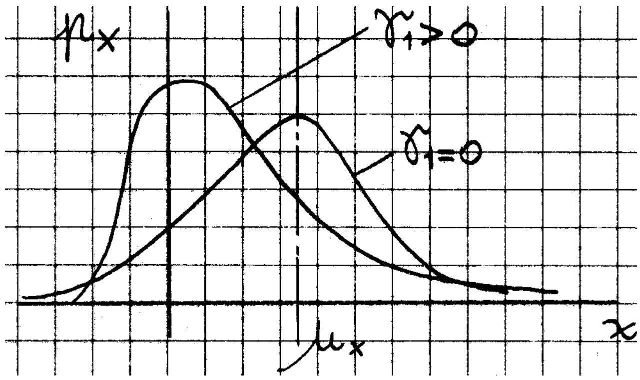
$$\boxed{V_X = \frac{\sigma_X}{\mu_X}}$$

It exists provided that the mean value is different from zero.
Let us define as <u>skewness</u> the non-dimensional quantity:

$$\boxed{\gamma_1 = \frac{1}{\sigma_X^3} \int_{-\infty}^{\infty} (x - \mu_X)^3 p_X(x)\, dx}$$

If $p_X(x)$ is symmetric, μ_X lies on its symmetry axis and $\gamma_1 = 0$.

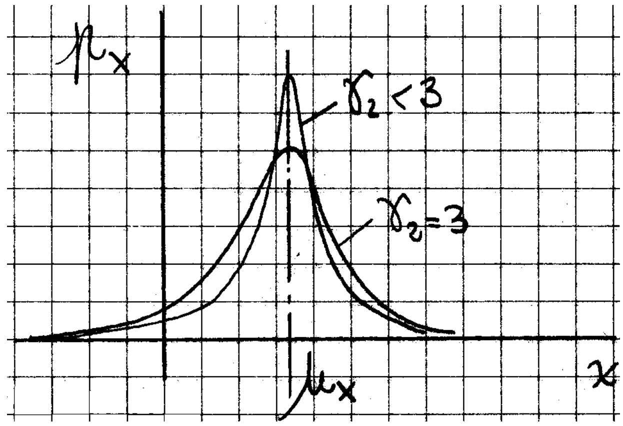


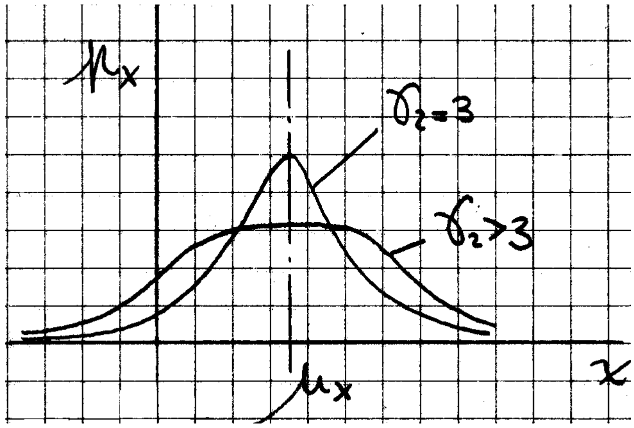


Let us define as <u>kurtosis</u> the non-dimensional coefficient:

$$\boxed{\gamma_2 = \frac{1}{\sigma_X^4} \int_{-\infty}^{\infty} (x - \mu_X)^4 p_X(x)\, dx}$$

If X is a normal random variable, $\gamma_2 = 3$.





Normal distribution

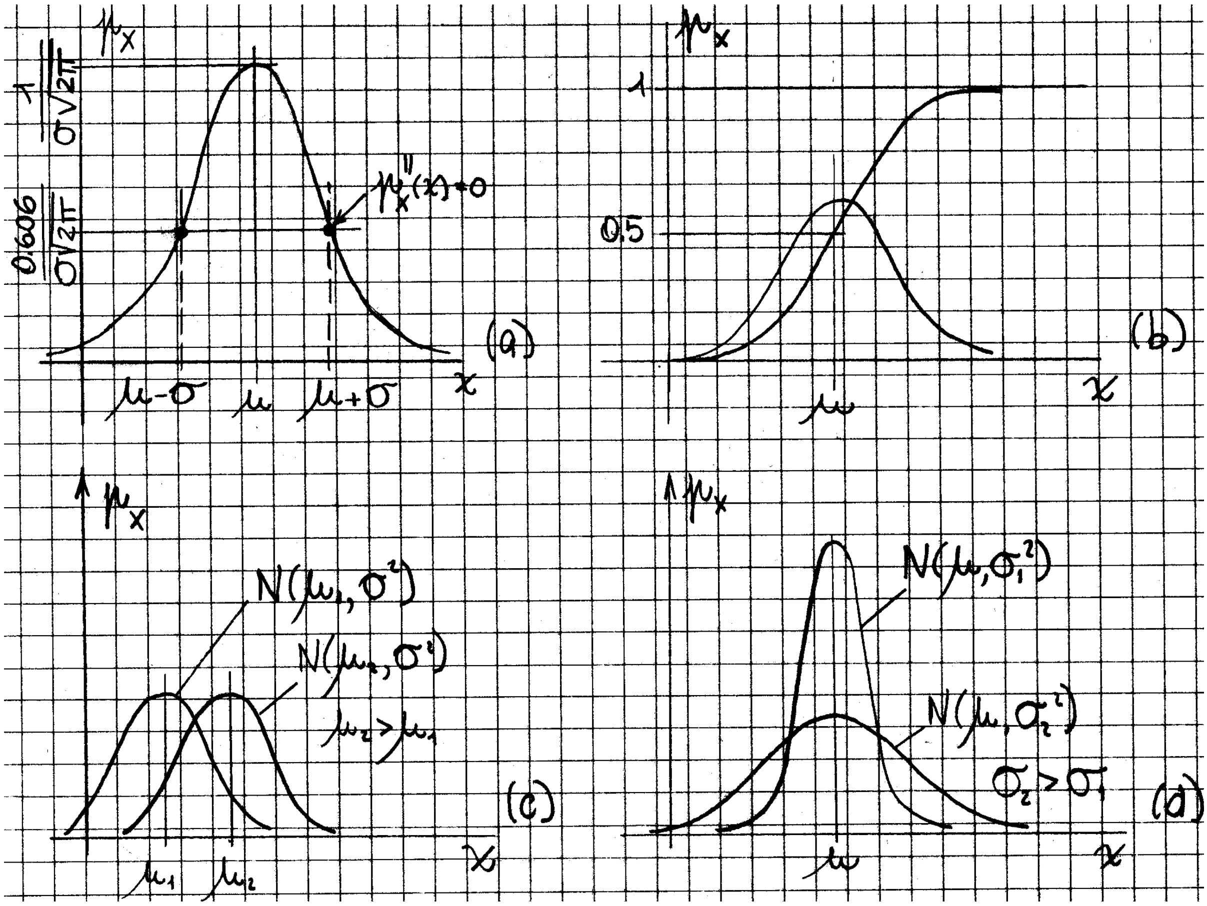
A continuous random variable X has normal distribution if its density function has the form:

$$\boxed{p_X(x) = \frac{1}{\sigma_X \sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left(\frac{x - \mu_X}{\sigma_X} \right)^2 \right\}}$$

The distribution function, $F_X(x) = P[X \le x]$, is given by the expression:

$$\boxed{F_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2\right\} du}$$

The following figure shows some qualitative diagrams of the normal distribution and its most important properties.



The reduced or standard normal random variable is defined as:

$$\boxed{Z = \frac{X-\mu_X}{\sigma_X}}$$

Thus the mean value and the standard deviation of Z are 0 and 1, respectively.

COUPLE OF RANDOM VARIABLES

Definitions

A couple of random variables X, Y is called a bi-variate random variable or a 2-component vector.

Joint distribution function

Let us consider a couple of random variables X, Y. The joint distribution function $F_{XY}(x, y)$ is the probability that X ≤ x, Y ≤ y:

$$\boxed{F_{XY}(x, y) = P(X \le x, Y \le y)}$$

It is always defined on the bi-dimensional space $(-\infty < x < +\infty)$, $(-\infty < y < +\infty)$.

The <u>marginal distribution functions</u> of X and Y, $F_X(x)$ e $F_Y(y)$, are the distribution functions of each variable X and Y. In general, the knowledge of F_X and F_Y does not allow to determine F_{XY}. Instead, given F_{XY}, it is possible to derive F_X e F_Y:

$$\boxed{F_X(x) = P(X \le x, Y < +\infty) = F_{XY}(x, +\infty)}$$

$$\boxed{F_Y(y) = P(X < +\infty, Y \le y) = F_{XY}(+\infty, y)}$$

The joint distribution function has some relevant properties:

$$F_{XY}(-\infty, y) = P(X < -\infty, Y \le y) = P(0) = 0$$
$$F_{XY}(x, -\infty) = P(X \le x, Y < -\infty) = P(0) = 0$$
$$F_{XY}(-\infty, -\infty) = P(X < -\infty, Y < -\infty) = P(0) = 0$$
$$F_{XY}(+\infty, +\infty) = P(X < +\infty, Y < +\infty) = 1$$

$F_{XY}(x, y)$ is a not decreasing function of x, y.

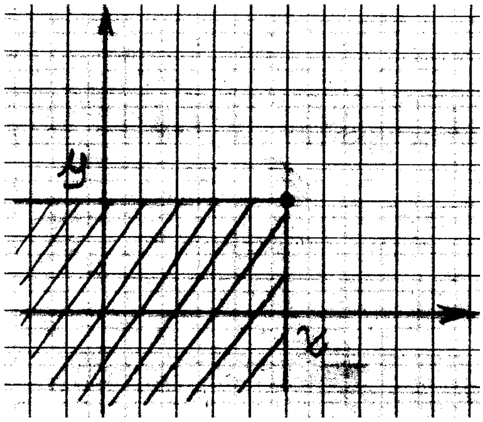
Joint density function

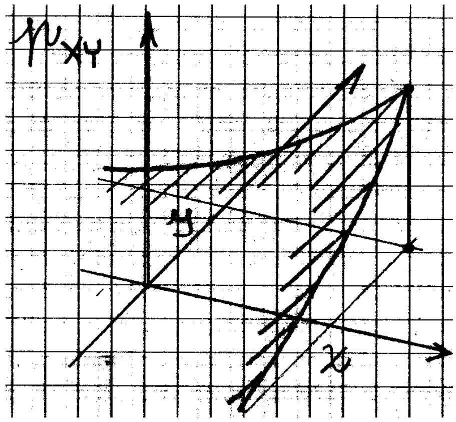
The joint density function $p_{XY}(x, y)$, or simply the joint density of the random variables X and Y, is given by:

$$\boxed{p_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y}}$$

The application of the bi-dimensional form of the fundamental theorem of the integral calculus gives rise to the expression:

$$\boxed{F_{XY}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} p_{XY}(\xi, \eta)\, d\xi\, d\eta}$$

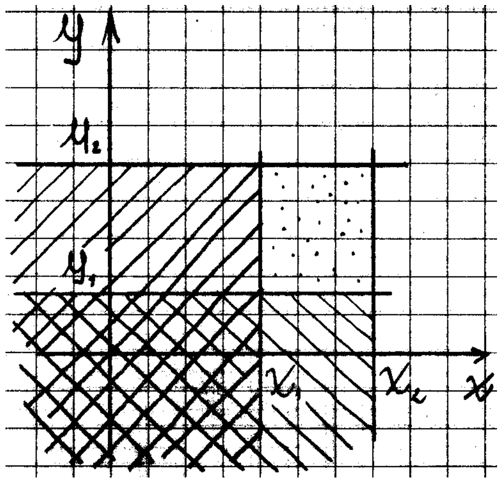


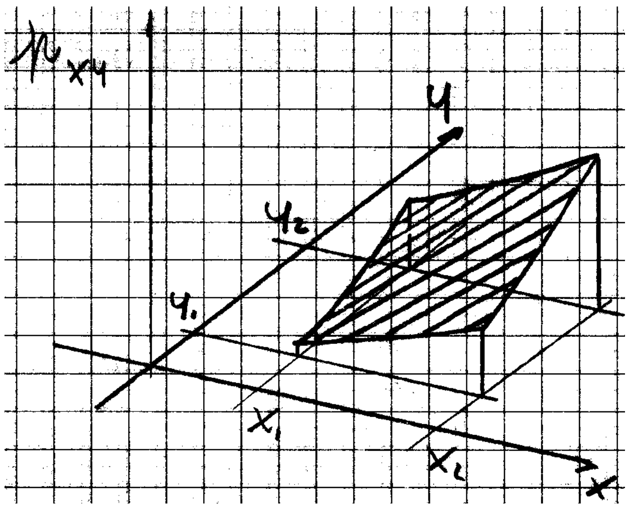


By virtue of the above equation it results:

$$\boxed{P\left(x_1 \le X \le x_2, y_1 \le Y \le y_2\right) = \int_{x_1}^{x_2}\int_{y_1}^{y2} p_{XY}\left(x,y\right)dxdy}$$

$$\boxed{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p_{XY}\left(x,y\right)dxdy = 1}$$





Moreover:

$$\boxed{p_X(x) = \int_{-\infty}^{+\infty} p_{XY}(x,y)\,dy}$$

$$\boxed{p_Y(y) = \int_{-\infty}^{+\infty} p_{XY}(x,y)\,dx}$$

The density functions $p_X(x)$ and $p_Y(y)$ of X and Y are called herein <u>marginal density functions</u>.

Independent random variables

Two random variables X and Y are defined as independent if the events $\{X \le x\}$ and $\{Y \le y\}$ are independent for any x and y, i.e. if $P(X \le x, Y \le y) = P(X \le x)P(Y \le y)$. It follows:

$$\boxed{F_{XY}(x,y) = F_X(x)F_Y(y)}$$

$$\boxed{p_{XY}(x,y) = p_X(x)p_Y(y)}$$

These equations are necessary and sufficient conditions of independence.

Indexes of a couple of random variables

The joint distribution function $F_{XY}(x,y)$ and the joint density function $p_{XY}(x,y)$ give a complete probabilistic description of X and Y. The indexes are quantities that provide a synthetic information on the couple of random variables. They can be framed into two classes:

(1) Indexes of X, Y dealt with separately – They are the same indexes of each random variable.
(2) Indexes of X, Y dealt with together – They express the probabilistic link between X and Y; they comprehend the correlation, the covariance and the coefficient of correlation.

The correlation of X, Y is defined as:

$$\boxed{R_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{+\infty} xy p_{XY}(x,y)\,dxdy}$$

The covariance of X, Y is defined as:

$$\boxed{C_{XY} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y) p_{XY}(x,y)\,dxdy}$$

Expanding the above equations it follows:

$$\boxed{C_{XY} = R_{XY} - \mu_X\mu_Y}$$

If $\mu_X = 0$ or $\mu_Y = 0$, then $C_{XY} = R_{XY}$. Moreover, if X = Y, then $R_{XX} = \varphi_X^2$, $C_{XX} = \sigma_X^2$.
The normalised covariance of X, Y, called also the coefficient of correlation, is defined as:

$$\boxed{\rho_{XY} = \frac{C_{XY}}{\sigma_X\sigma_Y}}$$

It should be remembered that X, Y are independent if:

$$p_{XY}(x,y) = p_X(x)\, p_Y(y)$$

These variables are not correlated if:

$$R_{XY} = \mu_X\ \mu_Y$$

$$C_{XY} = \rho_{XY} = 0$$

If X, Y are independent, they are also not correlated. It follows:

$$R_{XY} = \int_{-\infty}^{\infty} x p_X(x)\, dx \int_{-\infty}^{\infty} y p_Y(y)\, dy$$

which implies the independence of X and Y. The inverse statement is generally not true: if X, Y are not correlated, not necessarily they are also independent (this occurs only if X, Y are normal random variables). Thus, the condition of independence is stronger that the condition of not correlation.

The normalised covariance ρ_{XY} expresses the degree of correlation between X and Y. We already noted that $\rho_{XY} = 0$ if X and Y are not correlated. It is easy to demonstrate that:

$$-1 \le \rho_{XY} \le 1$$

It is possible to demonstrate that:

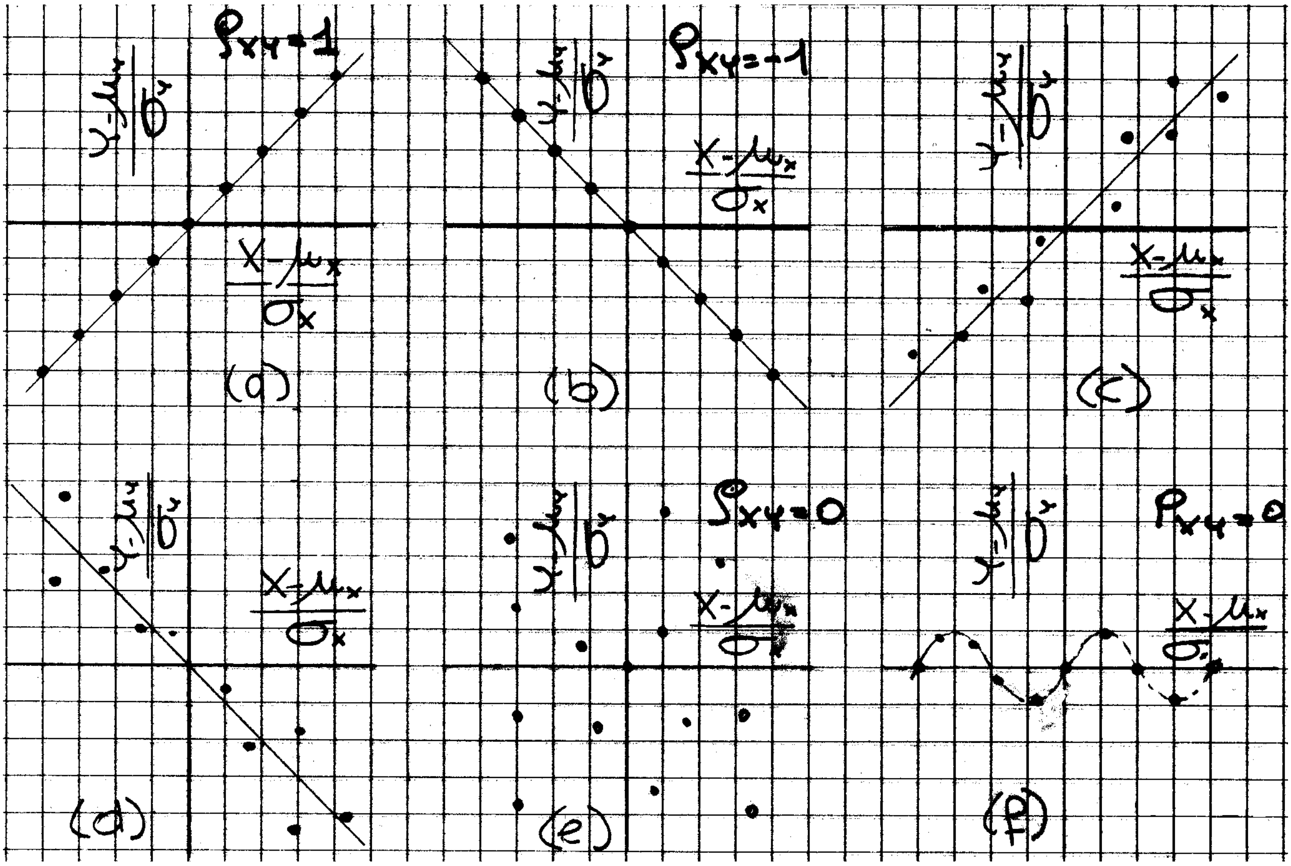
$$\rho_{XY} = +1 \Leftrightarrow Y = aX + b,\ \ a > 0$$
$$\rho_{XY} = -1 \Leftrightarrow Y = aX + b,\ \ a < 0$$

Tis equations is equivalent to the position:

$$\rho_{XY} = \pm 1 \Leftrightarrow \frac{X - \mu_X}{\sigma_X} = \pm \frac{Y - \mu_Y}{\sigma_Y} \Rightarrow \tilde{X} = \tilde{Y}$$

The above equations explain the difference between independence and not correlation. The statistical dependence involves any functional link between X and Y. The correlation involves a functional link of the linear type. Thus, the correlation is a particular case of the independence.

The following figure shows some typical examples of the statistical link between X, Y. If $\rho_{XY} = +1$ (a) X and Y are proportional; if $\rho_{XY} = -1$ (b) X and Y are inversely proportional. For intermediate values of $|\rho|$ between 0 and 1 (c), (d) X and Y tend to be roughly proportional or inversely proportional. The scattering is complete for $\rho_{XY} = 0$; it is worth noting that ρ_{XY} may be equal to zero also in the presence of a strong functional link of not linear type (f).

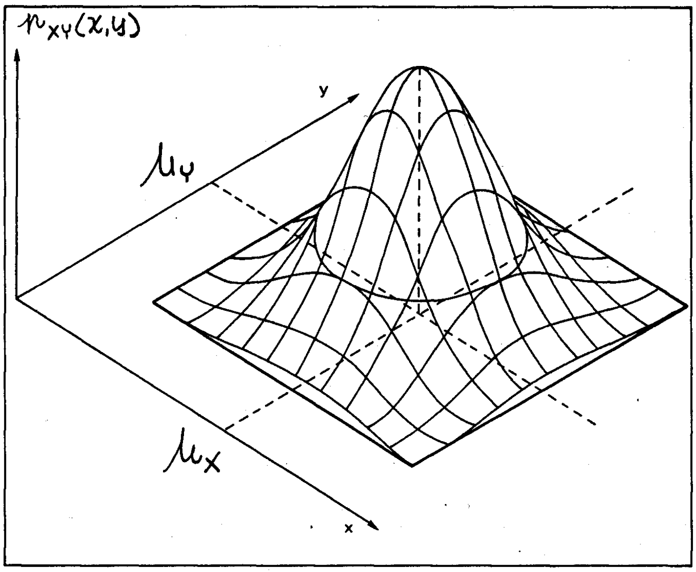


Bi-variate normal distribution

Two random variables X, Y have a bi-variate normal distribution if $p_{XY}(x,y)$ has the form:

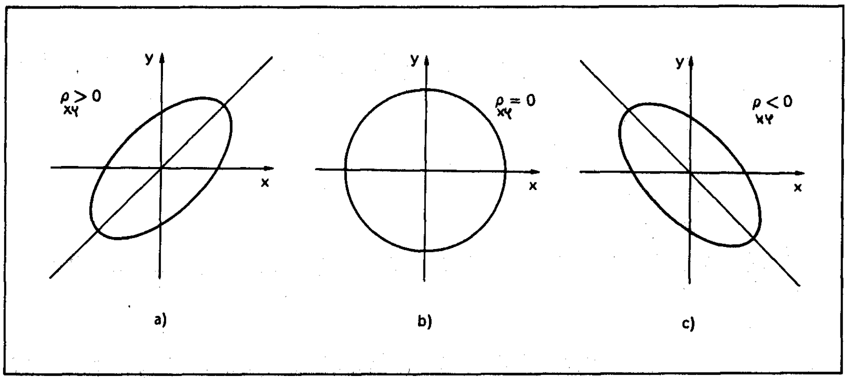
$$p_{XY}(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}}\cdot$$
$$\cdot\exp\left\{-\frac{\sigma_Y^2(x-\mu_X)^2-2\sigma_X\sigma_Y\rho_{XY}(x-\mu_X)(y-\mu_y)+\sigma_X^2(y-\mu_y)^2}{2\sigma_X^2\sigma_Y^2(1-\rho_{XY}^2)}\right\}$$

where μ_X and μ_Y are the mean values of X and Y, σ_X^2 and σ_Y^2 are the variances of X and Y, ρ_{XY} is the coefficient of correlation.



The above equations is often represented by the curves provided by the intersection of the joint density function with the planes $p_{XY}(x, y) = K$ parallel to the plane x, y:

$$\frac{\left(x^2 + y^2 - 2\rho_{XY}\, xy\right)}{\left(1-\rho_{XY}^2\right)} = K$$



For $\rho_{XY} = \pm 1$ these curves degenerate into the linear relationships $y = \pm x$.

It is relevant to remember that, if X, Y are statistically independent, then they are not correlated; the inverse statement is in general not true. In the particular case that X, Y have a bi-variate normal distribution, this statement is correct. Thus, if X, Y are not correlated, they are also independent. Setting $\rho_{XY} = 0$, the bi-variate normal density function becomes:

$$p_{XY}(x, y) = \frac{1}{\sigma_X\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_X}\right)^2\right\} \frac{1}{\sigma_Y\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\} \Rightarrow$$

$p_{XY}(x, y) = p_X(x) \cdot p_Y(y)$, $p_X(x)$, $p_Y(y)$ being marginal normal distributions. In other words, in this case, $\rho_{XY} = 0$ is a necessary and sufficient condition of statistical independence.

RANDOM VECTORS

Definitions

A vector that lists n random variables is a n-variate random vector. The probabilistic representation of the n-variate random vector $\mathbf{X} = \{x_1 x_2 \,..\, x_n\}^T$ implies the knowledge of the joint distribution of all the random variables $X_j \;(j = 1,...n)$ (n-th order distribution). If **X** is normal, its complete probabilistic representation involves the knowledge of the joint distribution of all the possible couples of random variables $X_i, X_j \;(i, j = 1,...n)$ composing the vector (2nd order distributions).

Joint distribution function

The joint distribution function $F_{\mathbf{X}}(\mathbf{x}) = F_{X_1 X_2 \ldots X_n}(x_1, x_2, ...x_n)$ of the vector $\mathbf{X} = \{x_1 x_2 \,..\, x_n\}^T$ is the probability that $X_1 \le x_1,\; X_2 \le x_2, ...X_n \le x_n$:

$$\boxed{F_{\mathbf{X}}(\mathbf{x}) = F_{X_1 X_2 \ldots X_n}(x_1, x_2, ...x_n) = P(X_1 \le x_1, X_2 \le x_2, ...X_n \le x_n)}$$

The joint distribution function of some of the random variables that compose the vector can be derived from the above equation setting as equal to infinite all the other variables. For instance:

$$F_{X_1 X_2}(x_1, x_2) = F_{X_1 X_2 \ldots X_n}(x_1, x_2, +\infty, ... +\infty)$$

Moreover, the following properties apply:

$$F_{X_1 X_2 \ldots X_n}(x_1, x_2, .., x_j = -\infty, .., x_n) = 0 \qquad \text{for } \forall j = 1, ...n$$

$$F_{X_1 X_2 \ldots X_n}(+\infty, +\infty, \;...\; +\infty) = 1$$

$F_{X_1 X_2 \ldots X_n}(x_1, x_2, ..x_n)$ is a not decreasing function of $x_1, x_2, .. x_n$.

Joint density function

Generalising to the vector **X** the considerations already developed with reference to the couple of random variables X,Y, the joint density function of **X** is given by the relationship:

$$\boxed{p_{X_1 X_2 \ldots X_n}(x_1, x_2, ...x_n) = \frac{\partial^n F_{X_1 X_2 \ldots X_n}(x_1, x_2, ...x_n)}{\partial x_1 \partial x_2 ... \partial x_n}}$$

$$\boxed{F_{X_1 X_2 \ldots X_n}(x_1, x_2, ...x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} ... \int_{-\infty}^{x_n} p_{X_1 X_2 \ldots X_n}(\xi_1, \xi_2, ...\xi_n)\, d\xi_1 d\xi_2 ... d\xi_n}$$

$$P(x_{1a} < X_1 < x_{1b},\, x_{2a} < X_2 < x_{2b},\, ...\, x_{na} < X_n < x_{nb}) = \\ = \int_{x_{1a}}^{x_{1b}} \int_{x_{2a}}^{x_{2b}} ... \int_{x_{na}}^{x_{nb}} p_{X_1 X_2 \ldots X_n}(x_1, x_2, \;...\; x_n)\, dx_1 dx_2 ... dx_n$$

It follows that:

$$\boxed{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}...\int_{-\infty}^{+\infty} p_{X_1X_2...X_n}(x_1,x_2,...x_n)\,dx_1dx_2...dx_n = 1}$$

The joint density function of some of the random variables composing the random vector can be derived from $p_{X_1X_2...X_n}(x_1,x_2,...x_n)$ integrating between -∞ and +∞ with respect to all the other variables. For instance:

$$p_{X_1X_2}(x_1,x_2) = \int_{-\infty}^{+\infty}...\int_{-\infty}^{+\infty}...p_{X_1X_2X_3...X_n}(x_1,x_2,x_3...x_n)\,dx_3...dx_n$$

Independent random vectors

A random vector is composed by independent random variables if the following properties apply:

$$\boxed{F_{X_1X_2...X_n}(x_1,x_2,...x_n) = F_{X_1}(x_1)F_{X_2}(x_2)...F_{X_n}(x_n) = \prod_{i=1}^{n} F_{X_i}(x_i)}$$

$$\boxed{p_{X_1X_2...X_n}(x_1,x_2,...x_n) = p_{X_1}(x_1)p_{X_2}(x_2)...p_{X_n}(x_n) = \prod_{i=1}^{n} p_{X_i}(x_i)}$$

Indexes of a random vector

Using the indexes introduced for a couple of random variables, a random vector $\mathbf{X} = \{x_1x_2\ ..\ x_n\}^T$ is characterised, in synthetic form, by the vector of the mean values $\boldsymbol{\mu}_\mathbf{X}$ and by the matrix of the correlations $\mathbf{R}_\mathbf{X}$ or of the covariances $\mathbf{C}_\mathbf{X}$.

The mean vector, called also the vector of the mean values, is defined as:

$$\boldsymbol{\mu}_\mathbf{X} = E[\mathbf{X}] = \{\mu_{X_1}\ \mu_{X_2}\ \ ...\ \ \mu_{X_n}\}^T$$

Its components are the mean values of the random variables that compose the random vector **X**, i.e. $\mu_{Xi} = E[X_i]\ \ (i = 1,...n)$.
The correlation matrix is defined as:

$$\mathbf{R}_\mathbf{X} = E[\mathbf{X}\ \mathbf{X}^\mathbf{T}] = \begin{bmatrix} R_{X_1X_1} & R_{X_1X_2} & .. & R_{X_1X_n} \\ R_{X_2X_1} & R_{X_2X_2} & .. & R_{X_2X_n} \\ \vdots & \vdots & \vdots & \vdots \\ R_{X_nX_1} & R_{X_nX_2} & .. & R_{X_nX_n} \end{bmatrix}$$

The on-diagonal terms are the mean square values of each random variable $\left(R_{X_iX_i} = \varphi^2_{X_i}\right)$; the off-diagonal terms are the correlations of all the possible couples of the random variables.

The covariance matrix is defined as:

$$\mathbf{C_X} = E\left[(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})^T\right] = \begin{bmatrix} C_{X_1X_1} & C_{X_1X_2} & .. & C_{X_1X_n} \\ C_{X_2X_1} & C_{X_2X_2} & .. & C_{X_2X_n} \\ \vdots & \vdots & \vdots & \vdots \\ C_{X_nX_1} & C_{X_nX_2} & .. & C_{X_nX_n} \end{bmatrix}$$

The on-diagonal terms are the variances of each random variable $\left(C_{X_iX_i} = \sigma^2_{X_i}\right)$; the off-diagonal terms are the covariances of all the couples of random variables. Expanding the above equations it follows that:

$$\mathbf{R_X} = \mathbf{C_X} + \boldsymbol{\mu}_\mathbf{X}\,\boldsymbol{\mu}_\mathbf{X}^T$$

$\mathbf{R_X}$ and $\mathbf{C_X}$ are symmetric matrices. It is possible to show that they are also semi-positive defined. $\mathbf{R_X}$ is diagonal if all the couples of different random variables $X_i, X_j\,(i \neq j)$ are orthogonal, i.e. $R_{X_iX_j} = 0$ ($\forall$i,j, i$\neq$j). $\mathbf{C_X}$ is diagonal of all the couples of different random variables $X_i, X_j(i \neq j)$ are not correlated, i.e. $C_{X_iX_j} = 0$ ($\forall$i,j, i$\neq$j).

It is demonstrated later that, if **X** is a normal vector, then the knowledge of $\boldsymbol{\mu}_\mathbf{X}$ and $\mathbf{R_X}$ or $\mathbf{C_X}$ is enough to determine the joint distributions of any order.

n-variate normal distribution

Let us consider a n-variate random vector $\mathbf{X} = \{X_1\ X_2\ ...\ X_n\}^T$. Let $\boldsymbol{\mu}_\mathbf{X} = \{\mu_{X_1}\ \mu_{X_2}\ ...\ \mu_{X_n}\}^T$ be the mean vector and $\mathbf{C_X} = E\left[(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})^T\right]$ be the covariance matrix. **X** has a n-variate normal distribution if its joint density function $p_{X_1 X_2 ... X_n}(x_1\ x_2\ ...\ x_n)$ has the form:

$$p_{X_1 X_2 ... X_n}(x_1, x_2, ...x_n) = \frac{1}{(2\pi)^{n/2}\left|\mathbf{C_X}\right|^{1/2}}\ \exp\left\{-\frac{1}{2\left|\mathbf{C_X}\right|}\sum_{1}^{n}{}_j\sum_{1}^{n}{}_k \left|\mathbf{C_X}\right|_{jk}\left(x_j - \mu_{X_j}\right)\left(x_k - \mu_{X_k}\right)\right\}$$

where $\left|\mathbf{C_X}\right|$ is the modulus of the determinant of $\mathbf{C_X}$ e $\left|\mathbf{C_X}\right|_{jk}$ is the j,k-th co-factor of $\mathbf{C_X}$ (i.e. the determinant of the matrix obtained by cancelling the j-th row and the k-th column of $\mathbf{C_X}$, multiplied by $(-1)^{j+k}$). In matrix form:

$$p_\mathbf{X}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}\left|\mathbf{C_X}\right|^{1/2}}\ \exp\left\{-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})^T\,\mathbf{C}_x^{-1}(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})\right\}$$

Thus, the knowledge of the first order statistical mean $(\boldsymbol{\mu}_\mathbf{X})$ and of the second order statistical mean $(\mathbf{C_X})$ of **X** is enough to derive the joint density function of order n.