Random variables

2

Outline

1. Definition of random variable
2. Discrete and continuous random variables
   - Discrete random variables (probability mass function and cdf)
   - Continuos random variables (density mass function and cdf)
3. Characteristic features of a random variable
   - Location, scatter and shape parameters
4. Transformations of random variables
5. Random vectors
   - Joint distribution and independence
   - Characteristic features, covariance matrix

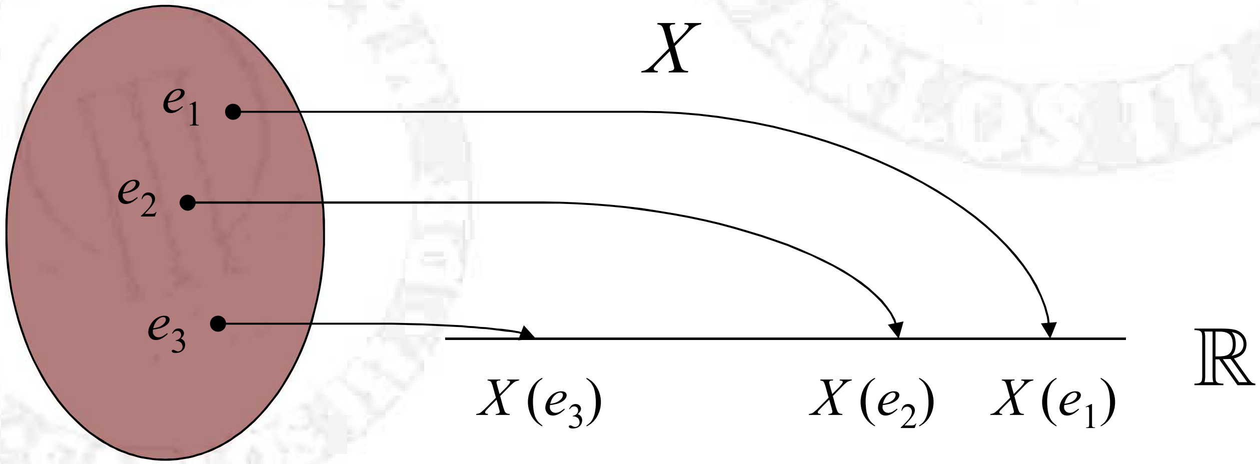
Definition of random variable

A random variable assesses a real number to each possible outcome of the random experiment.

It is random because we do not know its value before carrying out the random experiment.

Definition of random variable

- Definition. A random variable X is a mapping $X: E \to \mathbb{R}$, where E is the sample space associated with a random experiment.



Definition of random variable

- The events are now of the type $X \in A$, where A is a subset of $\mathbb{R}$.

Their probabilities are $P(X \in A) = P(\{e \in E\colon X(e) \in A\})$.

Properties:

1. $P(X \in A) \geq 0$;
2. $P(X \in \mathbb{R}) = 1$;
3. if $A_1, A_2, \ldots \subset \mathrm{IR}$ satisfy $A_i \cap A_j = \emptyset$ for $i \neq j$,

   then $P(X \in \cup_{i=1,\infty} A_i) = \Sigma_{i=1,\infty} P(X \in A_i)$.

Outline

1. Definition of random variable
2. Discrete and continuous random variables
   - Discrete random variables (probability mass function and cdf)
   - Continuos random variables (density mass function and cdf)
3. Characteristic features of a random variable
   - Location, scatter and shape parameters
4. Transformations of random variables
5. Random vectors
   - Joint distribution and independence
   - Characteristic features, covariance matrix

Discrete and continuous random variables

The support of a random variable is the set of values that it can assume.

- A random variable is discrete if its support is finite or denumerable.
  - Examples: number of defective items, number of rolls of a die until a 5 occurs.
- A random variable is continuous if its support contains an interval.
  - Example: battery life.

Discrete random variables

Given a discrete random variable *X*, its probability mass function assigns to each possible value of the variable the probability that X assumes such a value.

$$p: \mathbb{R} \to [0,1]$$

$$x \to p(x) = P(X=x)$$

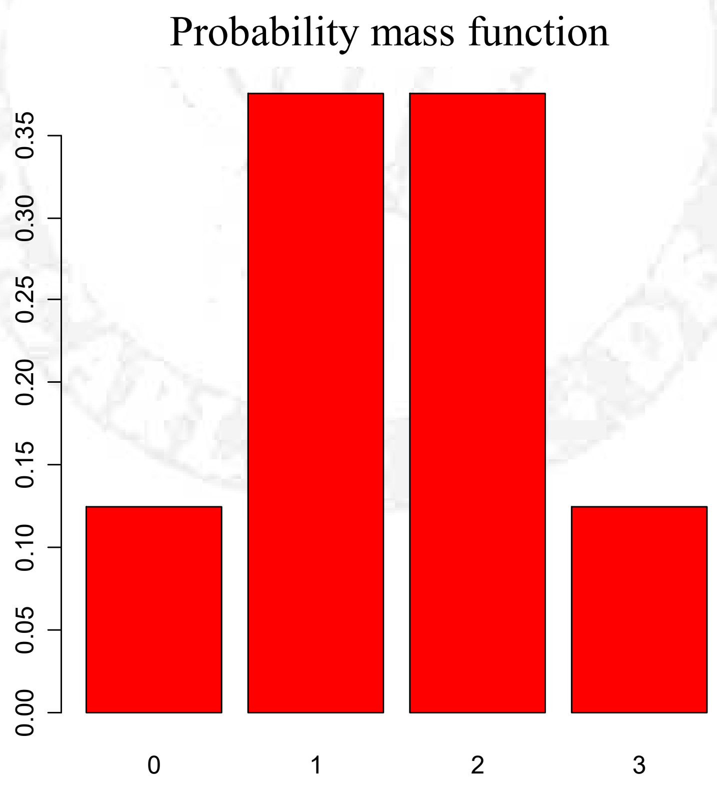
It satisfies that $0 \leq p(x) \leq 1$ for every x and if X assumes n different possible values $x_1,\ldots,x_n$, then $\Sigma_i\, p(x_i) = 1$.

We have $P(X \in A) = \Sigma_{x_i \in A}\, p(x_i)$.

Discrete random variables

Let X be the number of heads after tossing 3 times a fair coins. The probability mass function of X is:

x	$p(x) = P(X=x)$
0	0.125
1	0.375
2	0.375
3	0.125



Discrete random variables

The cumulative distribution function (cdf) of X at x is the probability that X is smaller or equal to x.

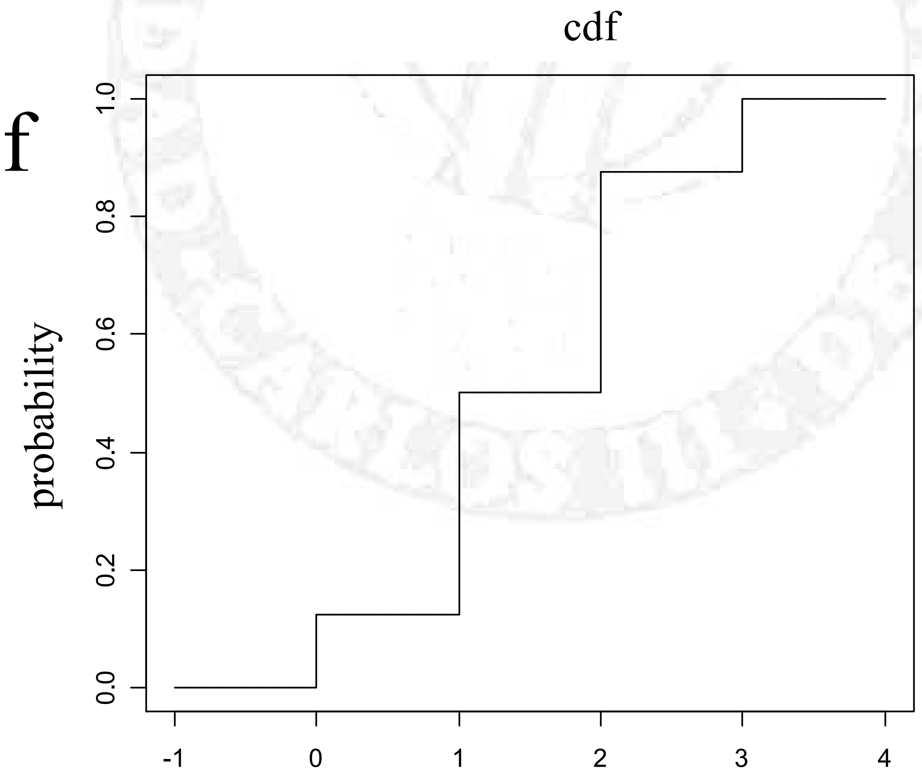
$$F(x) = P(X \leq x)$$

1. $\lim_{x \to -\infty} F(x) = 0$;
2. $\lim_{x \to \infty} F(x) = 1$;
3. F is nondecreasing ;
4. F is right-continuous.

Discrete random variables

The cumulative distribution function of a discrete random variable is stepwise,

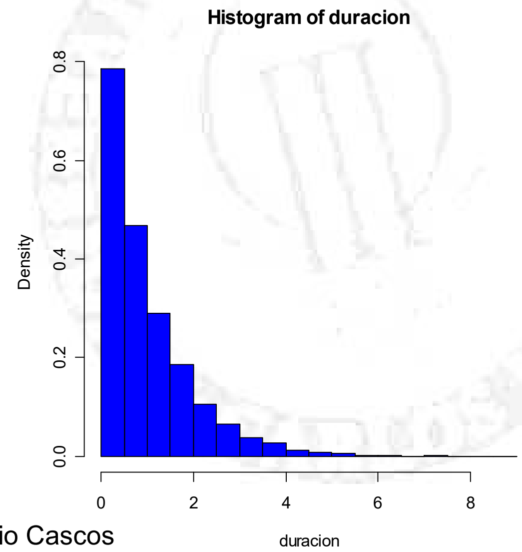
$$F(x) = P(X \leq x)$$
$$= \Sigma_{x_i \leq x} p(x_i)$$

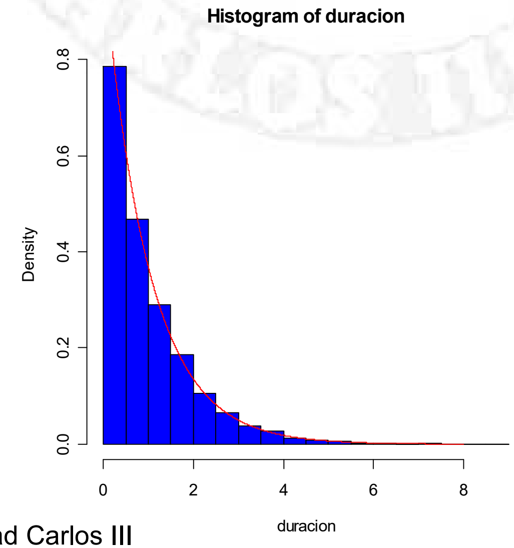


Continuous random variables

Since the range of a continuous random variable is not denumerable, an expression like $\Sigma_i\, p(x_i) = 1$ makes no sense.

Histogram for the lifetime of 10000 batteries.



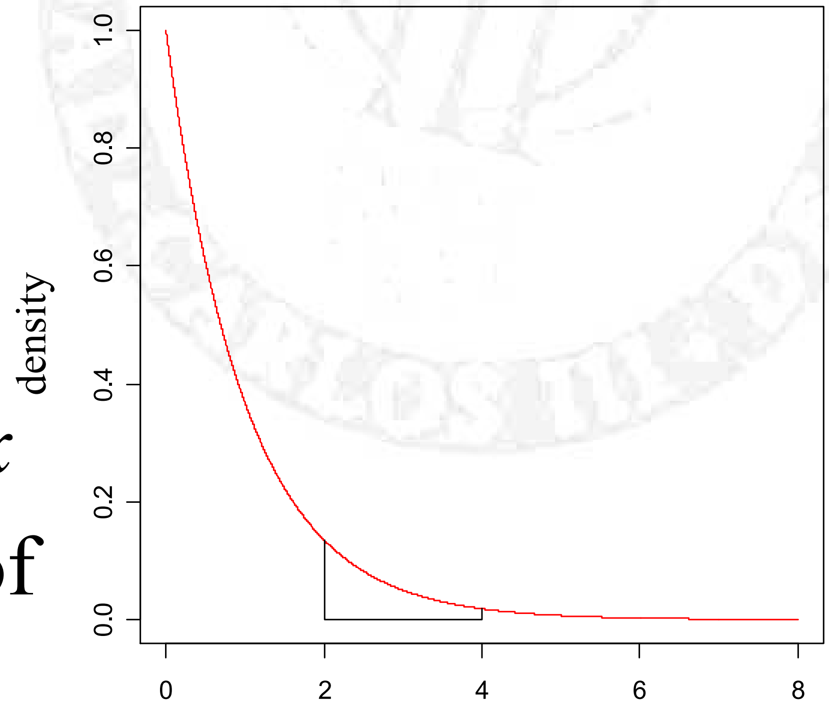


Continuous random variables

Function f describes the curve drawn together with the histogram on the right. We have

$$P(2 \leq X \leq 4) \approx \int_2^4 f(x)\mathrm{d}x$$

where X is the lifetime of a battery.



Continuous random variables

The density mass function f describes the probability distribution of a continuous random variable. It satisfies:

1. $f(x) \geq 0$;
2. $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$.
3. We have $P(a \leq X \leq b) = \int_a^b f(x)\mathrm{d}x$.

Given X continuous r.v., it satisfies

- $P(X = a) = 0$;
- $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$

Continuous random variables

In order to compute the cumulative distribution Function (cdf) of a continuous random variable, we must integrate its density mass function,

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$$

1. $\lim_{x\to -\infty} F(x) = 0$;
2. $\lim_{x\to \infty} F(x) = 1$;
3. F is nondecreasing ;
4. F is continuous .

Continuous random variables

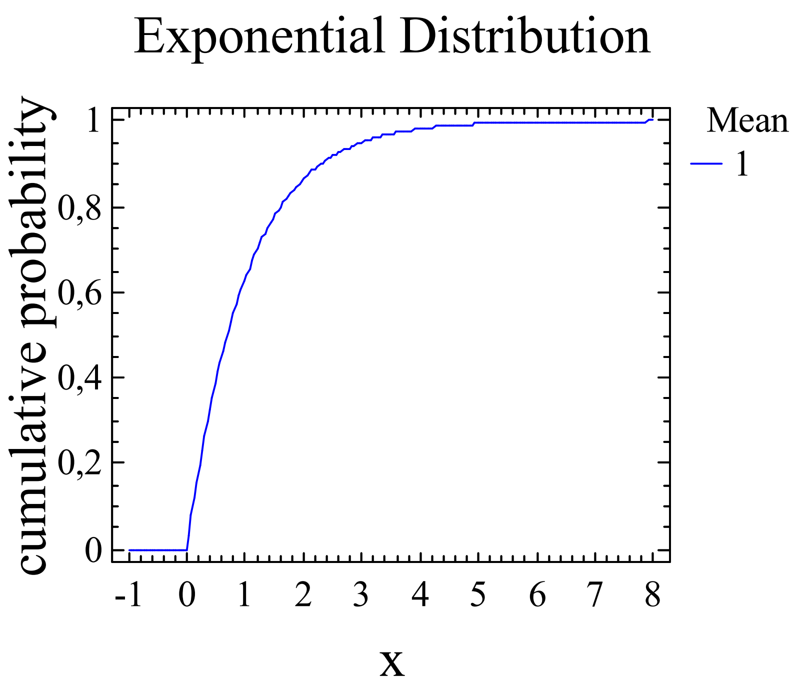
Since the cumulative distribution function is a primitive of the density mass function, deriving the cdf, we obtain the density mass function,

$$f(x) = \mathrm{d}F(x)/\mathrm{d}x\ .$$

We are working with

$f(x) = e^{-x}$ if $x > 0$

$F(x) = 1 - e^{-x}$ if $x > 0$



Outline

1. Definition of random variable
2. Discrete and continuous random variables
   - Discrete random variables (probability mass function and cdf)
   - Continuos random variables (density mass function and cdf)
3. Characteristic features of a random variable
   - Location, scatter and shape parameters
4. Transformations of random variables
5. Random vectors
   - Joint distribution and independence
   - Characteristic features, covariance matrix

Mathematical expectation or mean

The expecation or mean (μ) of a random variable is a central point with respect to its distribution

- X discrete, $\mu = \mathrm{E}[X] = \sum x_i p(x_i)$
- X continuous, $\mu = \mathrm{E}[X] = \int x f(x) \mathrm{d}x$

Properties: Given X,Y and two real numbers, a,b

1. $\mathrm{E}[a+bX] = a+b\mathrm{E}[X]$;
2. $\mathrm{E}[X+Y] = \mathrm{E}[X]+\mathrm{E}[Y]$.

Mathematical expecation or mean

Given a function $g: \mathbb{R} \to \mathbb{R}$, the expectation of random variable $g(X)$ can be computed as

- X discrete, $E[g(X)] = \sum g(x_i)p(x_i)$
- X continuous, $E[g(X)] = \int g(x)f(x)\mathrm{d}x$

Median

The median of a random variable X is a value Me such that

$$F(\text{Me}) \geq 1/2 \;\; ; \;\; P(X \geq \text{Me}) \geq 1/2$$

If X is a continuous r.v., then $F(\text{Me}) = 1/2$.

Quantiles

For $0 < \alpha < 1$, the α-quantile of random variable X is a value x_α such that the probability of X being not greater than x_α is, at least, α and the probability of X being not smaller than x_α is, at least, $1-\alpha$.

$$F(x_\alpha) = P(X \le x_\alpha) \ge \alpha \;;\; P(X \ge x_\alpha) \ge 1-\alpha$$

Percentiles and quartiles are defined in a similar way

$$\mathrm{P}_a = x_{a/100} \;;\; \mathrm{Q}_i = \mathrm{P}_{25i}$$

where $1 \le a \le 99$ and $1 \le i \le 3$.

Scatter parameters

The variance of a random variable X is given by

$$\sigma^2 = \mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2]$$

- X discrete, $\sigma^2 = \mathrm{Var}[X] = \sum (x_i - \mu)^2 p(x_i)$
- X continuous, $\sigma^2 = \mathrm{Var}[X] = \int (x - \mu)^2 f(x)\mathrm{d}x$

The standard deviation is the (positive) square root of the variance, $\sigma = (\mathrm{Var}[X])^{1/2}$.

Scatter parameters

Property. $\mathrm{Var}[X] = \mathrm{E}[X^2]-\mathrm{E}[X]^2 = \mathrm{E}[X^2]-\mu^2$

Given $a,b\in\mathrm{IR}$ and a random variable X, we have the following properties of the variance

1. $\mathrm{Var}[b] = 0$;
2. $\mathrm{Var}[aX] = a^2\mathrm{Var}[X]$;
3. $\mathrm{Var}[aX+b] = a^2\mathrm{Var}[X]$.

Shape parameters

Describe features from the distribution of a r.v. different from location and scatter

k-th moment about the origin, $m_k = E[X^k]$

k-th moment about the mean, $\mu_k = E[(X-\mu)^k]$

- Skewness. Skew = μ_3/σ^3
- Kurtosis. Kurt = μ_4/σ^4-3

Example (Flange thickness)

The thickness of a flange on an aircraft component, X in mm, has the following density mass function

$$f_X(x) = \begin{cases} 200(x - 0.95) & \text{if } 0.95 < x < 1.05 \\ 0 & \text{otherwise} \end{cases}$$

- Determine the cumulative distribution function of flange thickness.
- Determine the proportion of flanges that exceeds 1.02 mm.
- What thickness is exceeded by 90% of the flanges?
- Determine the mean and variance of flange thickness.

Outline

1. Definition of random variable
2. Discrete and continuous random variables
   - Discrete random variables (probability mass function and cdf)
   - Continuos random variables (density mass function and cdf)
3. Characteristic features of a random variable
   - Location, scatter and shape parameters
4. Transformations of random variables
5. Random vectors
   - Joint distribution and independence
   - Characteristic features, covariance matrix

Transformations of random variables

Given a random variable X and a function $g\colon \mathbb{R} \to \mathbb{R}$, we want to describe the distribution of the random varaible $Y=g(X)$.

$$F_Y(y) = P(Y \le y) = P(g(X) \le y) = P(X \in A_y)\ ,$$

where $A_y = \{x\colon g(x) \le y\}$.

It often happens that the set A_y can be fully described in a simple way.

Transformations of random variables

If X is a discrete random variable,

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$$

$$= \sum_{g(x_i) \leq y} p_X(x_i) \; .$$

Further, the probability mass function of Y is

$$p_Y(y) = P(Y = y) = P(g(X) = y)$$

$$= \sum_{g(x_i) = y} p_X(x_i) \; .$$

Transformations of random variables

If g is continuous and increasing

$$F_Y(y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$$

In general, given X continuous r.v. and $Y=g(X)$ with g injective and derivable, the density mass function of Y satisfies

$$f_Y(y) = f_X(x)\ |\mathrm{d}x/\mathrm{d}y|$$

Outline

1. Definition of random variable
2. Discrete and continuous random variables
   - Discrete random variables (probability mass function and cdf)
   - Continuos random variables (density mass function and cdf)
3. Characteristic features of a random variable
   - Location, scatter and shape parameters
4. Transformations of random variables
5. Random vectors
   - Joint distribution and independence
   - Characteristic features, covariance matrix

Joint distribution of a random vector

Given X and Y two random variables in the same sample space E, the mapping

$$(X,Y) : E \to \mathbb{R}^2$$

is a bivariate random vector.

The joint cumulative distribution function of (X,Y) at (x,y) is given by

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

Discrete random vectors

Given two discrete random variables X and Y, the random vector (X,Y) is discrete and its joint probability mass function is given by

$$p(x,y) = P(X = x\ ,\ Y = y).$$

It satisfies:

1. $p(x,y) \geq 0$
2. $\Sigma_x \Sigma_y p(x,y) = 1$

Joint cumulative distribution function

$$F(x_0,y_0) = P(X \leq x_0\ ,\ Y \leq y_0) = \Sigma_{x\leq x_0} \Sigma_{y\leq y_0} p(x,y)$$

Continuous random vectors

For X and Y continuous rv's, the random vector (X,Y) is continuous and its joint density function $f(x,y)$ satisfies

1. $f(x,y) \geq 0$;
2. $\int\int f(x,y)\mathrm{d}x\mathrm{d}y = 1$.

Joint cumulative distribution function

$$F(x_0,y_0) = P(X \leq x_0 \,,\, Y \leq y_0) = \int_{-\infty}^{x_0}\int_{-\infty}^{y_0} f(x,y)\mathrm{d}y\mathrm{d}x$$

That satisifes

$$P(a \leq X \leq b \,,\, c \leq Y \leq d) = \int_a^b\int_c^d f(x,y)\mathrm{d}y\mathrm{d}x$$

$$f(x,y) = \partial^2 F(x,y)/\partial x \partial y$$

Marginal distributions

- The distribution of each of the components of a random vector alone is referred to as marginal distribution.

Marginal distributions

- Discrete variables. Given X and Y whose joint probability mass function is denoted as $p(x,y)$,

$$p_X(x) = P(X{=}x) = \Sigma_y P(X{=}x, Y{=}y) = \Sigma_y p(x,y)$$

$$p_Y(y) = P(Y{=}y) = \Sigma_x P(X{=}x, Y{=}y) = \Sigma_x p(x,y)$$

- Continuous variables. Given X and Y whose joint density mass function is denoted as $f(x,y)$,

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x$$

Independent random variables

Two random variables X and Y are independent if for any $A,B\subset\mathrm{IR}$,

$$P\big((X\in A)\cap(Y\in B)\big) = P(X\in A)P(Y\in B)$$

Equivalently, for any $x,y\in\mathrm{IR}$

$$P\big((X\le x)\cap(Y\le y)\big) = P(X\le x)P(Y\le y)$$

Property. If X and Y are independent,

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X]+\mathrm{Var}[Y]$$

Characteristic features of a rand. vector

- Expectation. The mean vector of a random vector X is a column vector each of whose entries is the expectation of an entry of X

$$\mathrm{E}[X] = \begin{pmatrix} \mathrm{E}[X_1] \\ \mathrm{E}[X_2] \\ \vdots \\ \mathrm{E}[X_n] \end{pmatrix}$$

Characteristic features of a rand. vector

Given a function $g\colon \mathbb{R}^2 \to \mathbb{R}$, the expectation of random vector $g(X,Y)$ can be computed as

- (X,Y) discrete, $\mathrm{E}[g(X,Y)] = \sum g(x_i,y_j)p(x_i,y_j)$
- (X,Y) cont., $\mathrm{E}[g(X,Y)] = \int\int g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y$

Characteristic features of a rand. vector

- Covariance. The covariance is a measure of the linear relation between two variables,

$$\mathrm{Cov}(X,Y) = \mathrm{E}[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])]$$

- Properties:
  1. $\mathrm{Cov}(X,Y) = \mathrm{E}[XY]-\mathrm{E}[X]\mathrm{E}[Y]$;
  2. if X and Y are independent, $\mathrm{Cov}(X,Y) = 0$;
  3. $\mathrm{Cov}(X,Y) = 0$ does not guarantee that X and Y are independet ;
  4. $\mathrm{Cov}(aX+b,cY+d)=ac\mathrm{Cov}(X,Y)$.

Characteristic features of a rand. vector

- Correlation. The correlation is also a measure of the linear relation between two variables

$$\rho(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]}}$$

- Properties:
  1. if X and Y are independent, $\rho(X,Y) = 0$;
  2. if $\rho(X,Y) = 0$ they are said to be uncorrelated ;
  3. $|\rho(X,Y)| \leq 1$;
  4. $|\rho(X,aX+b)| = 1$.

Characteristic features of a rand. vector

- Covariance matrix.

$$M_X = \mathrm{E}\big[(X-\mu)(X-\mu)^{\mathrm{t}}\big]$$
$$= \begin{pmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}(X_1, X_2) & \cdots & \mathrm{Cov}(X_1, X_n) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Var}[X_2] & \cdots & \mathrm{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_n, X_1) & \mathrm{Cov}(X_n, X_2) & \cdots & \mathrm{Var}[X_n] \end{pmatrix}$$

Characteristic features of a rand. vector

- Variance of a linear combination.

$$\mathrm{Var}[aX+bY] = a^2\mathrm{Var}[X]+b^2\mathrm{Var}[Y]+2ab\,\mathrm{Cov}(X,Y)$$

$$\mathrm{Var}[aX+bY] = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} \mathrm{Var}[X] & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(Y,X) & \mathrm{Var}[Y] \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

Example (Dartboard)

A dartboard is a wooden square of 40cmx40cm. Inside the board there are several concentric circles, the greater among them having a diameter of 40cm. The darts of an unskilled player hit the board at a point that can be described by a random vector (X,Y) with joint density function

$$f_{X,Y}(x,y)=\begin{cases}1/1600 & \text{if } 0\le x,y\le 40\\ 0 & \text{otherwise}\end{cases}$$

where (X,Y) represent the coordinates (in centimeters) of the hitting point. The lower right corner of the board is represented by the origin, that is, the point (0,0).

Example (Dartboard)

- Whenever we hit one of the circles, we get some point. What is the probability of getting some point from a throw?
- What is the mean vector of (X,Y)?, and the probability that (X,Y) assumes that value?
- The inner circle has a diameter of 4cm. What is the probability of hitting it with a dart?