Mathematical Statistics

MAS 713

Chapter 3.1

Previous lectures

1. Descriptive statistics
2. Probability theory

Questions?

This lecture

3.1 Random variables

- 3.1.1 Introduction
- 3.1.2 Random variables
- 3.1.3 Discrete Random Variables
- 3.1.4 Continuous Random Variables
- 3.1.5 Expectation of a random variable
- 3.1.6 Variance of a random variable

Additional reading : Chapter 3 in the textbook

Introduction

Often, we are not interested in all of the details of the experimental result, but only in the value of some numerical quantity determined by the outcome

Example 1 : tossing two dice when playing a board game

$S = \{(1,1),(1,2),\dots,(6,5),(6,6)\}$

... but usually only the **sum** of the points matters

$\rightsquigarrow$ each possible outcome ω is characterized by a real number (when interested in the sum)

Introduction

Example 2 : buying 2 electronic items
each of which may be either defective or acceptable

$S = \{(d, d), (d, a), (a, d), (a, a)\}$

... but we might only be interested in the **number of acceptable items** obtained in the purchase

$\rightsquigarrow$ again, each possible outcome ω is characterised by a real number

It is often much more natural
to directly think in terms of the numerical quantity of interest
$\rightsquigarrow$ **random variables**

Random variables

Random variable : definition

The formal way of defining random variables involves measure theory.

Definition

A random variable is a real-valued function defined over the sample space :

$$X : S \to \mathbb{R}$$
$$\omega \to X(\omega)$$

Usually*, a random variable is denoted by an uppercase letter

Define S_X the domain of variation of X, that is the set of possible values taken by X

*except in your textbook

Random variable : definition

Example 1 : tossing two dice when playing a board game

$X =$ sum of the points, $\quad S_X = \{2, 3, 4, \dots, 12\}$

Example 2 : buying 2 electronic items

$X =$ number of acceptable items, $\quad S_X = \{0, 1, 2\}$

Events defined by random variables

For any fixed real value $x \in S_X$, assertions like "$X = x$" or "$X \leq x$" correspond to a set of possible outcomes

$$(X = x) = \{\omega \in S : X(\omega) = x\}$$
$$(X \leq x) = \{\omega \in S : X(\omega) \leq x\}$$

$\leadsto$ they are events !

$\leadsto$ meaningful to talk about their probability

Events defined by random variables

Example 1 : tossing two fair dice when playing a board game

- $X =$ sum of the points,
- $S_X = \{2, 3, 4, \ldots, 12\}$

for example:

$$(X = 2) = \{(1, 1)\} \quad \Longrightarrow \quad \mathbb{P}(X = 2) = 1/36$$

$$(X \geq 11) = \{(5, 6), (6, 5), (6, 6)\} \quad \Longrightarrow \quad \mathbb{P}(X \geq 11) = 3/36 = 1/12$$

Events defined by random variables

The usual properties of probabilities obviously apply to these ones, e.g.

- $\mathbb{P}(X \in S_X) = 1$
- $\mathbb{P}((X = x_1) \cup (X = x_2)) = \mathbb{P}(X = x_1) + \mathbb{P}(X = x_2)$ (if $x_1 \neq x_2$)
- $\mathbb{P}(X < x) = 1 - \mathbb{P}(X \geq x)$ ('$X < x$' is the complement of '$X \geq x$')

Notes

Note 1

It is important not to confuse:

- X, the name of the random variable
- $X(\omega)$, the numerical value taken by the random variable at some sample point ω
- x, a generic numerical value

Notes

Note 2

Most interesting problems can be stated, often naturally, in terms of random variables

- $\rightsquigarrow$ many inessential details about the sample space can be left unspecified, and one can still solve the problem
- $\rightsquigarrow$ often more helpful to think of random variables simply as variables whose values are likely to lie within certain ranges of the real number line

Cumulative distribution function

Cumulative distribution function

A random variable is essentially described by its cumulative distribution function (cdf) (or just distribution)

Definition

The cdf of the random variable X is defined for any real number x, by

$$F(x) = \mathbb{P}(X \leq x)$$

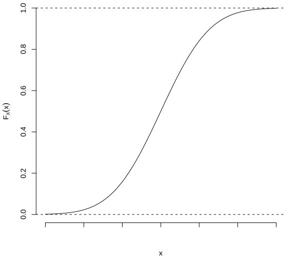
- All probability questions about X can be answered in terms of its distribution.
- We will denote $X \sim F$

Cumulative distribution function

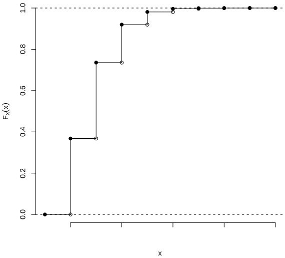
Some properties:

- for any $a \leq b$, $\mathbb{P}(a < X \leq b) = F(b) - F(a)$
- F is right-continuous, i.e. $F(x) = \lim_{h \downarrow 0} F(x+h)$
- F is a nondecreasing function
- $\lim_{x \to +\infty} F(x) = F(+\infty) = 1$
- $\lim_{x \to -\infty} F(x) = F(-\infty) = 0$

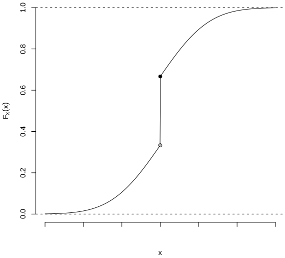
Cumulative distribution functions



Continuous distribution
$\leadsto$ continuous r.v.



Discrete distribution
$\leadsto$ discrete r.v.



Hybrid distribution
$\leadsto$ hybrid r.v.

Note : hybrid distributions will not be introduced in this course

Discrete random variables

Discrete random variables

Definition

A random variable is said to be discrete if it can only assume a finite (or at most countably infinite) number of values

Suppose that those values are $S_X = \{x_1, x_2, \ldots\}$

Discrete random variables

Definition

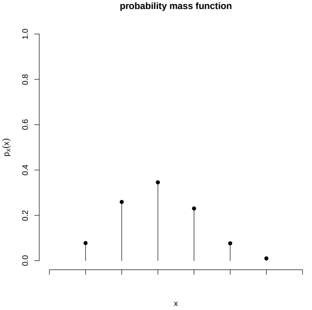
The probability mass function (pmf) of a discrete random variable X is defined for any real number x, by

$$\mathbb{P}(x) = \mathbb{P}(X = x)$$

$\rightsquigarrow p_X(x) > 0$ for $x = x_1, x_2, \ldots$, and $p_X(x) = 0$ for any other value of x

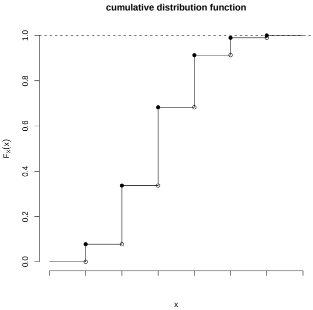
Obviously :

$$\mathbb{P}(X \in S_X) = \mathbb{P}((X = x_1) \cup (X = x_2) \cup \ldots) = \sum_{x \in S_X} p(x) = 1$$



Probability mass function:

- "spikes" at $x_1, x_2, \ldots$
- height of spike at $x_i = p(x_i)$



Cumulative distribution function:

- $F(x) = \sum_{i:x_i \le x} p(x_i)$
- step function
- jumps at $x_1, x_2, \ldots$
- magnitude of jump at $x_i = p(x_i)$

Discrete random variables : examples

Examples of discrete random variables include :

- number of scratches on a surface,
- number of defective parts among 1000 tested,
- number of transmitted bits received in error,
- the sum of the points when tossing 2 dice, ...

$\rightsquigarrow$ discrete random variables generally arise when we count things

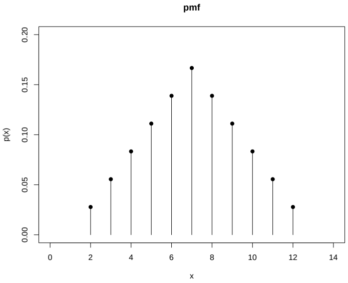
Discrete random variables : examples

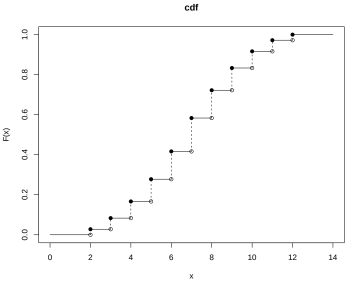
Example : tossing 2 dice

X = sum of the points **Goal:** represent $\mathbb{P}(x)$ and $F(x)$

x	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Check that $\mathbb{P}(x) = (6 - |7 - x|)/36$ for $x \in S_X = \{2, 3, 4, \ldots, 12\}$





Discrete random variables : examples

Example : production line

Suppose that a day's production of 850 manufactured parts contains 50 parts that do not conform to customer requirements.
Two parts are selected at random, without replacement, from the batch.

Let the random variable X equal the number of nonconforming parts in the sample. What is the cumulative distribution function of X?

Discrete random variables : examples

$$S_x = \{0, 1, 2\}$$

$$\mathbb{P}(x = 0) = \frac{800}{850} \times \frac{799}{849} = 0.886$$
$$\mathbb{P}(x = 1) = \frac{50}{850} \times \frac{800}{849} + \frac{800}{850} \times \frac{50}{849} = 0.111$$
$$\mathbb{P}(x = 2) = \frac{50}{850} \times \frac{49}{849} = 0.003$$

$$F(0) = Pr(x \leq 0) = 0.886$$
$$F(1) = Pr(x \leq 1) = 0.886 + 0.111 = 0.997$$
$$F(2) = Pr(x \leq 2) = 0.886 + 0.111 + 0.003 = 1$$

Bernoulli random variable

Bernoulli random variable

- It can only assume 2 values,
- $S_X = \{0, 1\}$
- Its pmf is given by

$$\mathbb{P}(1) = \pi = 1 - \mathbb{P}(0)$$

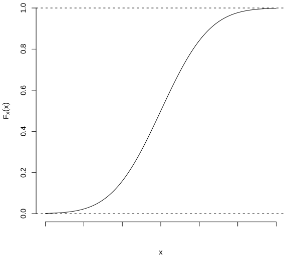
for some value $\pi \in (0, 1)$

- It is often used to characterise the occurrence/non-occurrence of a given event, or the presence/absence of a given feature

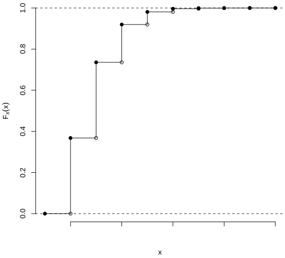
$$Bernoulli(y|\pi) = \pi^y (1 - \pi)^{1-y}$$

Continuous random variables

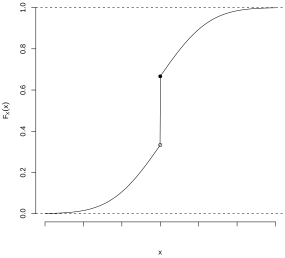
Cumulative distribution functions



Continuous distribution
$\rightsquigarrow$ continuous r.v.



Discrete distribution
$\rightsquigarrow$ discrete r.v.



Hybrid distribution
$\rightsquigarrow$ hybrid r.v.

Continuous random variables

- As opposed to discrete r.v., a continuous random variable X is expected to take on an uncountable number of values.
- S_X is therefore an uncountable set of real numbers (like an interval), and can even be $\mathbb{R}$ itself.

However, this is not enough as a definition

Continuous random variables

Definition

A random variable X is said to be continuous if there exists a nonnegative function $f(x)$ defined for all real $x \in \mathbb{R}$ such that for any set B of real numbers,

$$\mathbb{P}(X \in B) = \int_B f(x)dx$$

Consequence: $\mathbb{P}(X = x) = 0$ for any x !

$\rightsquigarrow$ the probability mass function is useless

$\rightsquigarrow$ the probability density function (pdf) $f(x)$ will play the central role

Continuous random variables : remark

Note 1 : the fact that $\mathbb{P}(X = x) = 0$ for any x should not be disturbing

$\rightsquigarrow$ coherent when dealing with measurements, e.g.

if we report a temperature of 74.8 degrees centigrade, owing to the limits of our ability to measure (accuracy of measuring devices), we really mean that the temperature lies "close to" 74.8, for instance between 74.75 and 74.85 degrees

Continuous random variables : remark

Note 2 : when there is a zero probability that a random variable X will take on any value x, this does not mean that it is impossible that X will take on the value x!

In the continuous case, zero probability does not imply logical impossibility

$\rightsquigarrow$ this should not be disturbing either, as we are always interested in probabilities connected with intervals and not with isolated points

Probability density function: properties

- $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(y)dy$, that is

$$\boxed{f(x) = \frac{dF(x)}{dx} = F'(x)}$$

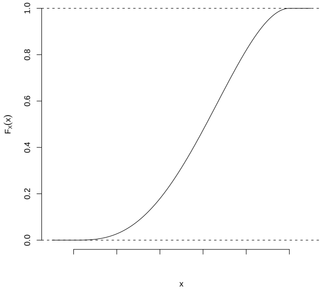
(wherever F is differentiable)

Probability density function: properties

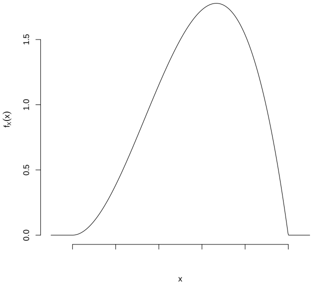
- $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(y)dy$, that is

$$\boxed{f(x) = \frac{dF(x)}{dx} = F'(x)}$$

(wherever F is differentiable)



cdf $F(x)$



pdf $f(x) = F'(x)$

Probability density function: properties

- $F(x) = \mathbb{P}(X \le x) = \int_{-\infty}^{x} f(y)dy$, that is

  $$\boxed{f(x) = \frac{dF(x)}{dx} = F'(x)}$$

  (wherever F is differentiable)

- $\boxed{f(x) \ge 0}$ $\quad \forall x \in \mathbb{R}$ $\qquad$ ($F(x)$ nondecreasing)

Probability density function: properties

- $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(y)dy$, that is

$$\boxed{f(x) = \frac{dF(x)}{dx} = F'(x)}$$

(wherever F is differentiable)

- $\boxed{f(x) \geq 0} \quad \forall x \in \mathbb{R} \qquad (F(x)$ nondecreasing)

- $\mathbb{P}(a \leq X \leq b) = \int_{a}^{b} f(x)dx = F(b) - F(a)$

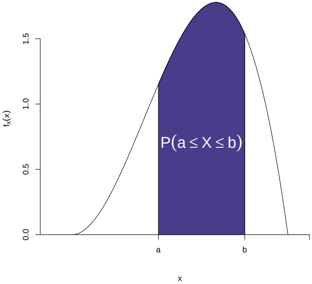
Probability density function: properties

- $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(y)dy$, that is

$$\boxed{f(x) = \frac{dF(x)}{dx} = F'(x)}$$

(wherever F is differentiable)

- $\boxed{f(x) \geq 0}$ $\quad \forall x \in \mathbb{R}$ $\quad$ ($F(x)$ nondecreasing)
- $\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx = F(b) - F(a)$



Probability density function: properties

- $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(y)dy$, that is

  $$\boxed{f(x) = \frac{dF(x)}{dx} = F'(x)}$$

  (wherever F is differentiable)

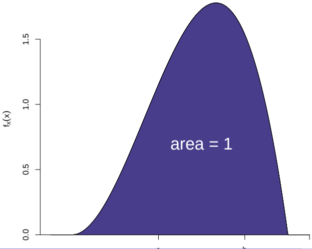
- $\boxed{f(x) \geq 0} \quad \forall x \in \mathbb{R} \qquad (F(x)$ nondecreasing)

- $\mathbb{P}(a \leq X \leq b) = \int_{a}^{b} f(x)dx = F(b) - F(a)$

- $$\boxed{\int_{-\infty}^{+\infty} f(x)dx = 1}$$

Probability density function: properties

- $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(y)dy$, that is
  $$\boxed{f(x) = \frac{dF(x)}{dx} = F'(x)}$$
  (wherever F is differentiable)
- $\boxed{f(x) \geq 0} \quad \forall x \in \mathbb{R} \qquad (F(x) \text{ nondecreasing})$
- $\mathbb{P}(a \leq X \leq b) = \int_{a}^{b} f(x)dx = F(b) - F(a)$
- $\boxed{\int_{-\infty}^{+\infty} f(x)dx = 1}$



Probability density function: properties

- $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(y)dy$, that is

$$\boxed{f(x) = \frac{dF(x)}{dx} = F'(x)}$$

(wherever F is differentiable)

- $\boxed{f(x) \geq 0} \quad \forall x \in \mathbb{R} \qquad (F(x) \text{ nondecreasing})$
- $\mathbb{P}(a \leq X \leq b) = \int_{a}^{b} f(x)dx = F(b) - F(a)$
- $\boxed{\int_{-\infty}^{+\infty} f(x)dx = 1}$
- for a small ε, $\mathbb{P}(x - \varepsilon/2 \leq X \leq x + \varepsilon/2) = \int_{x-\varepsilon/2}^{x+\varepsilon/2} f(y)dy \simeq \varepsilon f(x)$

  $\rightsquigarrow S_X = \{x \in \mathbb{R} : f(x) > 0\}$

Probability density function: properties

- $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(y)dy$, that is
$$\boxed{f(x) = \frac{dF(x)}{dx} = F'(x)}$$
(wherever F is differentiable)
- $\boxed{f(x) \geq 0}$ $\quad \forall x \in \mathbb{R}$ $\quad$ ($F(x)$ nondecreasing)
- $\mathbb{P}(a \leq X \leq b) = \int_{a}^{b} f(x)dx = F(b) - F(a)$
- $\boxed{\int_{-\infty}^{+\infty} f(x)dx = 1}$
- for a small ε, $\mathbb{P}(x - \varepsilon/2 \leq X \leq x + \varepsilon/2) = \int_{x-\varepsilon/2}^{x+\varepsilon/2} f(y)dy \simeq \varepsilon f(x)$

$\rightsquigarrow S_X = \{x \in \mathbb{R} : f(x) > 0\}$

Note: as $\mathbb{P}(X = x) = 0$, $\mathbb{P}(X < x) = \mathbb{P}(X \leq x)$ (for **contin.** r.v. only!)

Continuous random variables : examples

Examples of continuous random variables include :

- electrical current,
- length,
- pressure,
- temperature,
- time,
- voltage,
- weight,
- speed of a car,
- amount of alcohol in a person's blood,
- efficiency of solar collector,
- strength of a new alloy, . . .

$\rightsquigarrow$ continuous r.v. generally arise when we measure things

Continuous random variables : examples

Example

Let X denote the current measured in a thin copper wire (in mA). Assume that the pdf of X is

$$f(x) = \begin{cases} C(4x - 2x^2) & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

1. What is the value of C ?
2. Find $\mathbb{P}(X > 1.8)$

Continuous random variables : examples

Solution:

1) We must have $\int_{-\infty}^{+\infty} f(x)\,dx = 1$, so

$$C\int_0^2 (4x - 2x^2)\,dx = C \times \frac{8}{3} = 1,$$

that is $C = 3/8$

2) Then,

$$\mathbb{P}(X > 1.8) = \int_{1.8}^{+\infty} f(x)\,dx = 3/8 \times \int_{1.8}^{2} (4x - 2x^2)\,dx = 0.028$$

Discrete vs. Continuous random variables

Discrete vs. Continuous random variables

Discrete r.v.	Continuous r.v.

Domain of variation

Discrete r.v.	Continuous r.v.
$S_X = \{x_1, x_2, \ldots\}$	$S_X = [m, M]$

Probability mass function (pmf)

Discrete r.v.:

$\mathbb{P}(x) = \mathbb{P}(X = x) \geq 0$ for all $x \in \mathbb{R}$

- $\mathbb{P}(x) > 0$ if and only if $x \in S_X$
- $\sum_{x \in S_X} \mathbb{P}(x) = 1$

Continuous r.v.: useless

Probability density function (pdf)

Discrete r.v.: does not exist

Continuous r.v.:

$0 \leq f(x)\ (= F'(x))$ for all $x \in \mathbb{R}$

- $f(x) > 0$ if and only if $x \in S_X$
- $\int_{x \in S_X} f(x) = 1$

Note the similarity between the conditions for pmf and pdf

Parameters of a distribution

Parameters of a distribution

Fact

Some quantities characterise a random variable more usefully (although incompletely) than the whole cumulative distribution function

$\leadsto$ focus on certain general properties of the distribution of the r.v.

The two most important such quantities are:

- the expectation (or mean) and
- the variance

of the random variable

Often, we talk about the expectation or the variance of a distribution, understood as the expectation or the variance of a random variable having that distribution

Expectation of a random variable

Expectation

The **expectation** or the **mean** of a random variable X, denoted $\mathbb{E}(X)$ or μ, is defined by

Discrete r.v.

$$\mu = \mathbb{E}(X) = \sum_{x \in S_X} x\,\mathbb{P}(x)$$

Continuous r.v.

$$\mu = \mathbb{E}(X) = \int_{S_X} x\, f(x) dx$$

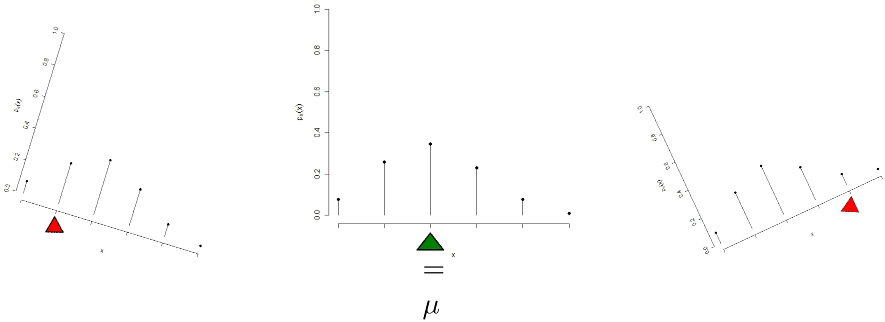
$\rightsquigarrow$ $\mathbb{E}(X)$ is a weighted average of the possible values of X, each value being weighted by the probability that X assumes it

Note: $\mathbb{E}(X)$ is in the same units as X

Expectation

Expectation = expected value, mean value, average value of X
= “central” value, around which X is distributed
= “centre of gravity” of the distribution

In the discrete case :



$\rightsquigarrow$ localisation parameter

Expectation : examples

Example 1

What is the expectation of the outcome when a fair die is rolled?

Solution:

- $X =$ outcome,
- $S_X = \{1, 2, 3, 4, 5, 6\}$
- $\mathbb{P}(x) = 1/6$ for any $x \in S_X$

$$\begin{aligned}\mu = \mathbb{E}(X) &= 1 \times 1/6 + 2 \times 1/6 + 3 \times 1/6 + 4 \times 1/6 + 5 \times 1/6 + 6 \times 1/6 \\ &= 3.5\end{aligned}$$

$\leadsto$ μ need not be a possible outcome !

$\leadsto$ μ is not the most likely outcome (this is called the mode)

Expectation : examples

Example 2

What is the expected sum when two fair dice are rolled?

Solution:

- $X = \text{sum}$,
- $S_X = \{2, 3, \ldots, 12\}$
- $\mathbb{P}(x) = (6 - |7 - x|)/36$ for any $x \in S_X$

$$\Longrightarrow \mu = \mathbb{E}(X) = 2 \times 1/36 + 3 \times 2/36 + \ldots + 12 \times 1/36 = 7$$

Expectation : examples

Example 3

What is the expectation of a Bernoulli r.v.?

Solution:

- $S_X = \{0, 1\}$
- $\mathbb{E}(X) = 0 \times (1 - \pi) + 1 \times \pi = \pi$

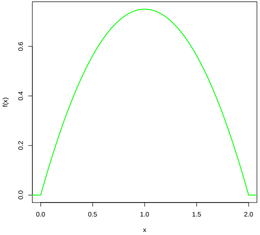
Expectation : examples

Example 4

Find the mean value of the copper current measurement X for Example on Slide 37, that is, with

$$f(x) = \begin{cases} \frac{3}{8}(4x - 2x^2) & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

The density is



By symmetry, it can be directly concluded that $\mu = 1$ mA

It can also easily be checked that

$$\mu = \mathbb{E}(X) = \int_{-\infty}^{+\infty} x\, f(x)\, dx$$
$$= \frac{3}{8} \int_0^2 x\,(4x - 2x^2)\, dx$$
$$= 1$$

Expectation of a function of a random variable

Sometimes we are interested, not in the expected value of X, but in the expected value of a function of X, say $g(X)$

There is actually no need for explicitly deriving the distribution of $g(X)$ Indeed, it can be shown

If X is a discrete r.v.

$$\mathbb{E}(g(X)) = \sum_{x \in S_X} g(x)\, \mathbb{P}(x)$$

If X is a continuous r.v.

$$\mathbb{E}(g(X)) = \int_{S_X} g(x)\, f(x)\, dx$$

Properties of the expectation $\mathbb{E}(\cdot)$

1) *Monotonicity:* $\mathbb{E}(X_1) \geq \mathbb{E}(X_2)$ if $X_1 \geq X_2$
2) *constant preserving:* $\mathbb{E}(c) = c \quad \forall c \in \mathbb{R}$
3) *homogeneity:* $\mathbb{E}(\lambda X) = \lambda\,\mathbb{E}(X) \quad \forall \lambda \in \mathbb{R}$
4) *additivity:* $\mathbb{E}(X_1 + X_2) = \mathbb{E}(X_1) + \mathbb{E}(X_2)$

In particular, for constants a and b :

Linear transformation

$$\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$$

Relation $\mathbb{P}$ and $\mathbb{E}$

Note that for any event E

$$\mathbb{P}(E) = \mathbb{E}[\mathbb{1}_E], \qquad \text{where } \mathbb{1}_E : S \to \mathbb{R} \text{ is defined by } \begin{cases} 1 & \text{if } \omega \in E \\ 0 & \text{if } \omega \notin E \end{cases}$$

Variance of a random variable

Variance of a random variable

Definition

The **variance** of a random variable X, usually denoted by $\mathbb{V}\mathrm{ar}(X)$ or σ^2, is defined by

$$\mathbb{V}\mathrm{ar}(X) = \mathbb{E}\left((X-\mu)^2\right)$$

Clearly, $\boxed{\mathbb{V}\mathrm{ar}(X) \geq 0}$

Note

An alternative formula for $\mathbb{V}\mathrm{ar}(X)$ is the following :

$$\sigma^2 = \mathbb{V}\mathrm{ar}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \mathbb{E}(X^2) - \mu^2$$

$\rightsquigarrow$ in practice, this is often the easiest way to compute $\mathbb{V}\mathrm{ar}(X)$

Variance of a random variable

If X is a discrete r.v.

$$\sigma^2 = \mathbb{V}\mathrm{ar}(X) = \sum_{x \in S_X} (x-\mu)^2 \mathbb{P}(x)$$

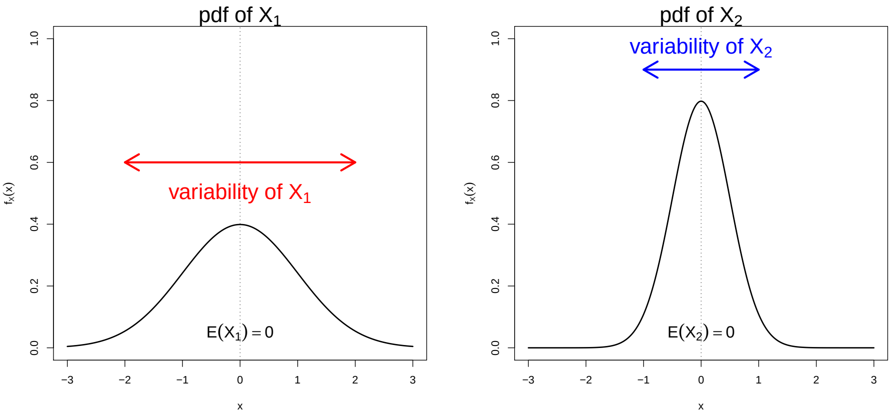
If X is a continuous r.v.

$$\sigma^2 = \mathbb{V}\mathrm{ar}(X) = \int_{S_X} (x-\mu)^2 f(x)dx$$

- $\leadsto$ expected square of the deviation of X from its expected value
- $\leadsto$ the variance quantifies the dispersion of the possible values of X around the "central" value μ, that is, the variability of X

Variance : illustration

Two random variables X_1 and X_2, with $\mathbb{E}(X_1) = \mathbb{E}(X_2)$



$$\rightsquigarrow \mathbb{V}\text{ar}(X_1) > \mathbb{V}\text{ar}(X_2)$$

Standard deviation

Note 2

The variance σ^2 is not in the same units as X, which may make interpretation difficult

$\leadsto$ often, we adjust for this by taking the square root of σ^2

This is called the **standard deviation** σ of X :

$$\sigma = \sqrt{\sigma^2} = \sqrt{\mathbb{V}\mathrm{ar}(X)}$$

Variance : linear transformation

A useful identity is that, for any constants a and b, we have

Linear transformation

$$\mathbb{V}\text{ar}(aX + b) = a^2\, \mathbb{V}\text{ar}(X)$$

- Take $a = 1$, it follows that for any b, $\boxed{\mathbb{V}\text{ar}(X + b) = \mathbb{V}\text{ar}(X)}$

$\rightsquigarrow$ variance not affected by translation

- Take $a = 0$, if follows that for any b, $\boxed{\mathbb{V}\text{ar}(b) = 0}$

("degenerate" random variable)

Variance : examples

Example 1

What is the variance of the number of points shown when a fair die is rolled?

Solution:

- $X =$ outcome,
- $S_X = \{1, 2, 3, 4, 5, 6\}$
- $\mathbb{P}(x) = 1/6$ for any $x \in S_X$

$$\begin{aligned}\mathbb{E}(X^2) &= 1^2 \times 1/6 + 2^2 \times 1/6 + 3^2 \times 1/6 + 4^2 \times 1/6 + 5^2 \times 1/6 + 6^2 \times 1\\ &= 91/6\end{aligned}$$

- We know that $\mu = 3.5$ (Slide 45), so that

$$\sigma^2 = \mathbb{E}(X^2) - \mu^2 = 91/6 - 3.5^2 \simeq 2.92$$

- The standard deviation is $\sigma = \sqrt{2.92} \simeq 1.71$

Variance : examples

Example 2

What is the variance of the sum of the points when 2 fair dice are rolled ?

(Exercise): Check that $\sigma^2 \simeq 5.83$, $\sigma \simeq 2.41$

Example 3

What is the variance of a Bernoulli r.v.?

Solution:

- $\mathbb{E}(X^2) = 0^2 \times (1 - \pi) + 1^2 \times \pi = \pi$
- $\mathbb{E}(X) = \pi$

$$\Longrightarrow \mathbb{V}\text{ar}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \pi - \pi^2 = \pi(1 - \pi)$$

Variance : examples

Example 4

What is the variance of the copper current measurement X for Example on Slide 36, that is, with

$$f(x) = \begin{cases} \frac{3}{8}(4x - 2x^2) & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

Solution:

- We have $\mathbb{E}(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\, dx = \frac{3}{8} \int_0^2 x^2(4x - 2x^2)\, dx = 1.2$
- We know that $\mu = 1$ (Slide 23), so that $\sigma^2 = 1.2 - 1^2 = 0.2$ mA^2

$\rightsquigarrow \sigma \simeq 0.45$ mA

Standardisation

Standardisation

Standardisation is a very useful linear transformation

Suppose you have a random variable X with mean μ and variance σ^2. Then, the associated standardised random variable, often denoted Z, is given by

$$Z = \frac{X - \mu}{\sigma},$$

that is $Z = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$. Hence, using the linear transformations properties,

$$\mathbb{E}(Z) = \frac{1}{\sigma}\mathbb{E}(X) - \frac{\mu}{\sigma} = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0$$

$$\mathbb{V}\text{ar}(Z) = \frac{1}{\sigma^2}\,\mathbb{V}\text{ar}(X) = \frac{\sigma^2}{\sigma^2} = 1$$

$\leadsto$ a standardised random variable has always mean 0 and variance 1

Note 1 : Z is a dimensionless variable (no unit)
Note 2 : a standardised value of X is sometimes called z-score

Markov's Inequality

Proposition

Let X be a r.v. that takes only nonnegative values, and $a > 0$. Then

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

Proof:

By monotonicity and homogeneity of the expectation

$$\mathbb{P}(X \geq a) = \mathbb{E}(\mathbb{1}_{\{X \geq a\}}) \leq \mathbb{E}(\mathbb{1}_{\{X \geq a\}} \tfrac{X}{a}) \leq \mathbb{E}\left(\tfrac{X}{a}\right) = \tfrac{\mathbb{E}(X)}{a}.$$

Chebyshev's inequality

Proposition

Let X be a random variable with mean μ and variance σ^2. Then, for any value $k > 0$,

$$\mathbb{P}(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

Proof :

Apply Markov's inequality to the nonnegative r.v. $(X - \mu)^2$:

$$\mathbb{P}(|X - \mu| \geq k) = \mathbb{P}((X - \mu)^2 \geq k^2) \leq \frac{\mathbb{E}((X - \mu)^2)}{k^2} = \frac{\sigma^2}{k^2}$$

Chebyshev's inequality

Note

Markov's and Chebyshev's inequalities provide bounds on some probabilities of interest which are valid for any random variable $\rightsquigarrow$ herein lies their power

Of course, if the actual distribution of X was known, then the probability could be exactly computed and we would not need to resort to bounds

Chebyshev's inequality : example

Example

Suppose it is known that the number of items produced in a factory during a week is a random variable with mean 50.

a) What can be said about the probability that this week's production will exceed 75?

b) If the variance of a week's production is known to equal 25, then what can be said about the probability that this week's production will not be between 25 and 75?

a) By Markov's inequality,

$$\mathbb{P}(X > 75) \leq \frac{\mathbb{E}(X)}{75} = \frac{50}{75} \simeq 0.67$$

b) By Chebyshev's inequality,

$$\mathbb{P}(|X - 50| \geq 25) \leq \frac{\sigma^2}{25^2} = \frac{1}{25} = 0.04$$

Objectives

Now you should be able to :

- understand the differences between discrete and continuous random variables □
- for discrete r.v., determine probabilities from pmf and the reverse □
- for continuous r.v., determine probabilities from pdf and the reverse □
- for discrete r.v., determine probabilities from cdf and cdf from pmf and the reverse □
- for continuous r.v., determine probabilities from cdf and cdf from pdf and the reverse □
- calculate means and variances for both discrete and continuous random variables □

Put yourself to the test ! $\rightsquigarrow$ Q1 p.127, Q2 p.127 Q3 p.127, Q4 p.128