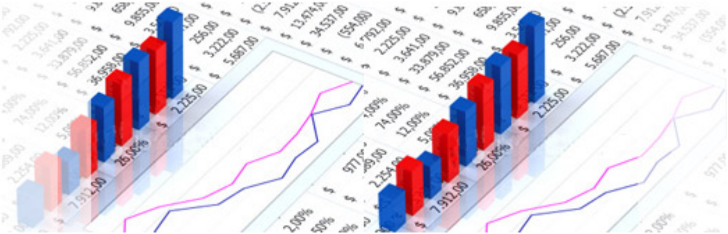
MTMS.01.099 Mathematical Statistics

Lecture 2. Discrete random variables

Tõnu Kollo



Fall 2017

Random variables (rv)

A **random variable** is a variable whose possible values are numerical outcomes of a random phenomenon. There are two main types of random variables, discrete and continuous.

Examples of random variables:

- number of cars at a parking place
- outside temperature
- number of children in a family
- height and weight of a person
- the daily change in a stock price
- ...

Random variables are denoted by capital letters such as X, Y, Z. The values (realisations) of random variables are denoted by small letters with indices $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots$. The set of possible values (realisations) of rv X is denoted by S_X.

Random variables

Random variable

A random variable is a numerical description of the outcome of an experiment (trial).

Discrete random variable

A discrete random variable may take either a finite number of values or an infinite countable number of values.

Continuous random variable

A continuous random variable may take any numerical value in an interval, in a collection of intervals or whole real line.

Discrete random variable. Definition

Discrete random variable

If a random variable can take only a finite or countable infinite number of different values, the variable is said to be *discrete*.

Values of a discrete random variables are usually (but not necessarily) integers.

Examples of discrete random variables:

- the number of children in a family
- the number of car accidents in a day in Estonia
- the number of students in a lecture
- the number of defective light bulbs in a box of ten
- ...

Discrete random variable

- Discrete random variable with a finite number of possible values is called also a simple random variable
  - Let X = number of laptops sold at the store in one day, where X can take on 5 values (0, 1, 2, 3, 4)
- Discrete random variable with an infinite countable set of possible values
  - Let X = number of daily customers in a Maxima foodstore, where X can take on the values 0, 1, 2, . . .

    (We can count the customers, but there is no finite upper limit on the number that might occur)

Discrete random variable and its probability distributions

Definition

The **probability distribution** of a random variable describes how probabilities are distributed over the values of the random variable.

- The probability model for a random variable is its probability distribution.
- The probability distribution of a random variable gives its possible values and their probabilities.

Discrete probability distribution **can be described with** a table, graph, or equation.

Discrete random variable and its probability distributions

Suppose a random variable X may take k different values, with the probability that $X = x_i$ is defined to be

$$P(X = x_i) = p(x_i) = p_i.$$

The probability distribution of a discrete random variable X lists the values x_i and their probabilities p_i:

Value	x_1	x_2	x_3	$\ldots$
Probability	p_1	p_2	p_3	$\ldots$

The probabilities p_i must satisfy two requirements:

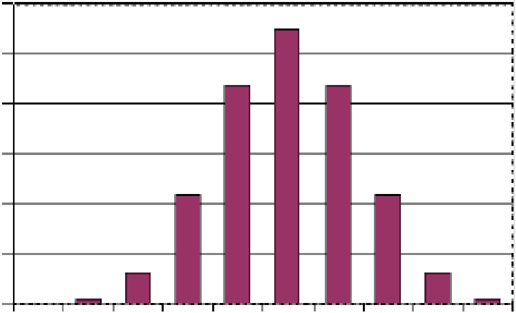
1. Every probability p_i is a number between 0 and 1.
2. The sum of the probabilities is 1.

Example 1.1

Suppose a variable X can take the values $\{0, 1, 2, 3, 4, 5, 6, 7, 8\}$. The probabilities associated with each outcome are described by the following table:

x_i	0	1	2	3	4	5	6	7	8
$p(x_i)$	0.004	0.031	0.109	0.219	0.274	0.219	0.109	0.031	0.004

This distribution may also be described by the graph shown below:



Distribution function of the random variable

One way to express the distribution of a rv is to use the distribution function of the variable. For a given real number x we compute the probability $P(X \leq x)$. We obtain a function $F_X(x) = P(X \leq x)$ defined for all x in the interval $-\infty < x < \infty$.

Definition

$F_X(x) = P(X \leq x)$ is called the **distribution function** of the random variable X.

Theorem 1

The distribution function $F_X(x)$ of a rv X has the properties:

$$F_X(x) \to \begin{cases} 0 \\ 1 \end{cases} \text{when } x \to \begin{cases} -\infty \\ \infty \end{cases}$$

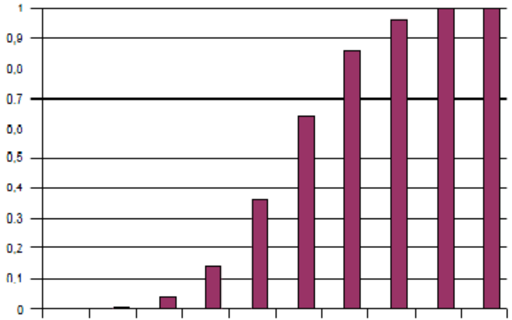
$F_X(x)$ is a nondecreasing function of x;

$F_X(x)$ is continuous from the right of any x.

Let's determine the distribution function based on the previous example. Taking into account that $F_X(x) = P(X \leq x)$, we get

	F(0)	F(1)	F(2)	F(3)	F(4)	F(5)	F(6)	F(7)	F(8)
F(x)	0.004	0.035	0.144	0.363	0.637	0.856	0.965	0.996	1.000

This table may also be described by the graph shown below:



The expected value of a discrete random variable

The **expected value** (often use the term **mean**) of any discrete random variable is an average of the possible outcomes, with each outcome weighted by its probability.

Definition

The *expectation* of the discrete rv X is defined by

$$E(X) = \sum_{x_i \in S_X} x_i p(x_i)$$

- Usually denoted by $E(X)$ or μ_X.
- The expected value of a function g of a rv X, $E[g(X)]$, is given by
$$E[g(X)] = \sum_{x_i \in S_X} g(x_i) p(x_i).$$
- The expectation $E(X)$ is a *measure of location* that shows where the values of the random variable are situated "on the average."

Example 1.3

Consider the previous distribution table

x_i	0	1	2	3	4	5	6	7	8
$p(x_i)$	0.004	0.031	0.109	0.219	0.274	0.219	0.109	0.031	0.004

Compute the mean of the random variable X.

$$\mu_X = E(X) = \sum_{x_i \in S_X} x_i p(x_i) = ?$$

Example 1.3

Consider the previous distribution table

x_i	0	1	2	3	4	5	6	7	8
$p(x_i)$	0.004	0.031	0.109	0.219	0.274	0.219	0.109	0.031	0.004

Compute the mean of the random variable X.

$$\mu_X = E(X) = \sum_{x_i \in S_X} x_i p(x_i) = ?$$

$$\mu_X = 0 \cdot 0.004 + 1 \cdot 0.031 + \ldots + 7 \cdot 0.031 + 8 \cdot 0.004 = 4$$

Some general properties of mean

- $Ec = c$ (mean of a constant is a constant itself)
- $E(cX) = cEX$
- $E(X + Y) = EX + EY$ (mean of the sum of rv's is the sum of means of rv's)
- $E(X - Y) = EX - EY$
- $E(X \cdot Y) = EX \cdot EY$ (holds only in case of independent random variables)

Independence:
Discrete random variables are independent when

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$$

The variance of a discrete random variable

Beside the mean as the measure of location for a discrete random variable, we need also a measure of dispersion. For that purpose notions of variance and standard deviation are in use.

Definition

The variance V(X) of the random variable X is defined by

$$V(X) = E[(X - E(X))^2].$$

Note that instead of $V(X)$ is common to write $Var(X)$ or $D(X)$ or σ_X^2.
The **variance of a discrete rv** X is defined by

$$Var(X) = \sigma_X^2 = \sum_{x_i} (x_i - E(X))^2 p(x_i).$$

To get the **standard deviation of a random variable**, take the square root of the variance: $\sigma_X = \sqrt{Var(X)}$.

Example 1.4

Consider the previous distribution table

x_i	0	1	2	3	4	5	6	7	8
$p(x_i)$	0.004	0.031	0.109	0.219	0.274	0.219	0.109	0.031	0.004

Compute the standard deviation of the random variable X.

$$\sigma_X^2 = \sum_{x_i} (x_i - E(X))^2 p(x_i) = ?$$

Example 1.4

Consider the previous distribution table

x_i	0	1	2	3	4	5	6	7	8
$p(x_i)$	0.004	0.031	0.109	0.219	0.274	0.219	0.109	0.031	0.004

Compute the standard deviation of the random variable X.

$$\sigma_X^2 = \sum_{x_i} (x_i - E(X))^2 p(x_i) = ?$$

$$\sigma_X^2 = (0-4)^2 \cdot 0.004 + (1-4)^2 \cdot 0.031 + \ldots + (8-4)^2 \cdot 0.004 = 1.996$$

Example 1.4

Consider the previous distribution table

x_i	0	1	2	3	4	5	6	7	8
$p(x_i)$	0.004	0.031	0.109	0.219	0.274	0.219	0.109	0.031	0.004

Compute the standard deviation of the random variable X.

$$\sigma_X^2 = \sum_{x_i} (x_i - E(X))^2 p(x_i) = ?$$

$$\sigma_X^2 = (0-4)^2 \cdot 0.004 + (1-4)^2 \cdot 0.031 + \ldots + (8-4)^2 \cdot 0.004 = 1.996$$

Now we can easily obtain standard deviation

$$\sigma_X = \sqrt{1.996} = 1.412$$

Some general properties of variance

Let us assume that X and Y are random variables. Then

- Variance of a constant: $Var(c) = 0$
- $Var(cX) = c^2 Var(X)$
- $Var(X) = E(X - EX)^2 = E(X^2 - 2XEX + (EX)^2) = EX^2 - 2EXEX + E(EX)^2 = EX^2 - (EX)^2$
- When X and Y are independent $Var(X + Y) = Var(X) + Var(Y)$
- When X and Y are independent $Var(X - Y) = Var(X) + Var(Y)$

Note that $Var(X) \geq 0$!

Some general properties of variance

Assume now that X and Y are dependent random variables and let us have $a, b \in R$. Then

$$Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2ab \cdot cov(X, Y),$$

where $cov(X, Y) = E((X - EX)(Y - EY))$ is the **covariance of X and Y.**

Remarks:

- If $cov(X, Y) > 0$, then X and Y are said to be positively correlated.
- If $cov(X, Y) < 0$, then X and Y are said to be negatively correlated.
- If X and Y are independent, then $cov(X, Y) = 0$, e.g. they are uncorrelated.

If $cov(X, Y) = 0$, does it mean that X and Y are independent?

Some general properties of variance

... if $cov(X, Y) = 0$, does it mean that X and Y are independent?

Example.

Consider rv X which has two possible values: -1 and 1, and both values have probability 0.5:

x_i	-1	1
p_i	0.5	0.5

It's easy to see that $EX = 0$ and $EX^3 = 0$. Now let's take $Y = X^2$.

Some general properties of variance

… if $cov(X, Y) = 0$, does it mean that X and Y are independent?

Example.

Consider rv X which has two possible values: -1 and 1, and both values have probability 0.5:

x_i	-1	1
p_i	0.5	0.5

It's easy to see that $EX = 0$ and $EX^3 = 0$. Now let's take $Y = X^2$.

Then:

$$cov(X, Y) = E(XY) - EX \cdot EY = E(X \cdot X^2) - 0 \cdot EY = EX^3 - 0 = 0.$$

Some general properties of variance

... if $cov(X, Y) = 0$, does it mean that X and Y are independent?

Example.

Consider rv X which has two possible values: -1 and 1, and both values have probability 0.5:

x_i	-1	1
p_i	0.5	0.5

It's easy to see that $EX = 0$ and $EX^3 = 0$. Now let's take $Y = X^2$.

Then:

$$cov(X, Y) = E(XY) - EX \cdot EY = E(X \cdot X^2) - 0 \cdot EY = EX^3 - 0 = 0.$$

So, the answer is NO.

Some discrete distributions. Bernoulli distribution

Definition

The random variable X that counts the success in a random trial is said to have a **Bernoulli distribution** with parameter p, written $X \sim Be(p)$.

The probability function of a Bernoulli random variable X with parameter p is

$$p(k) = P(X = k) = p^k (1 - p)^{1-k}$$

for $k = 0, 1$.

Some discrete distributions. Binomial distribution

Definition

The random variable X that counts the number of successes, k, in the n trials is said to have a **binomial distribution** with parameters n and p, written $X \sim Bin(n, p)$.

The probability mass function of a binomial random variable X with parameters n and p is

$$p(k) = P(X = k) = C_n^k p^k (1 - p)^{n-k}$$

for $k = 0, 1, \ldots, n$.

Some discrete distributions. Binomial distribution

Definition

The random variable X that counts the number of successes, k, in the n trials is said to have a **binomial distribution** with parameters n and p, written $X \sim Bin(n, p)$.

The probability mass function of a binomial random variable X with parameters n and p is

$$p(k) = P(X = k) = C_n^k p^k (1 - p)^{n-k}$$

for $k = 0, 1, \ldots, n$.
$C_n^k = \frac{n!}{k!(n-k)!}$ counts the number of outcomes that include exactly k successes and $n - k$ failures.

Some discrete distributions. Discrete uniform distribution

Definition

Let the discrete random variable X have k different values $x_1, \ldots, x_k$. X has **discrete uniform distribution**, if it's probability function is

$$p(j) = P(X = x_j) = \frac{1}{k}, \quad j = 1, \ldots, k$$

Some discrete distributions. Poisson distribution

A Poisson distributed random variable is often useful in estimating the number of occurrences over a specified interval of time or space.

Definition

If the random variable X has probability function

$$p(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

where $k = 0, 1, \ldots$ and distribution parameter $\lambda > 0$ is a constant, then X is said to have a **Poisson distribution**, written $X \sim Po(\lambda)$.