

# MTMS.01.099 Mathematical Statistics

## Lecture 4. Normal distribution. Standard normal distribution

Tõnu Kollo



Fall 2017

## Definition

A continuous rv  $X$  is said to have a **normal distribution** with parameters  $\mu$  and  $\sigma$ , where  $-\infty < \mu < \infty$  and  $\sigma > 0$ , if the density function of  $X$  is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

It is written as  $X \sim N(\mu, \sigma^2)$ .

The mean and variance of a random variable  $X$  having the normal distribution:  $E(X) = \mu$  and  $Var(X) = \sigma^2$ .

# Standard normal distribution

The normal distribution with parameter values  $\mu = 0$  and  $\sigma = 1$  is called a **standard normal distribution**. The random variable is usually denoted by  $Z$ . The p.d.f. of  $Z$  is

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty.$$

The distribution function of  $Z$  is denoted by  $\Phi$ . Thus

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^z \phi(x) dx.$$

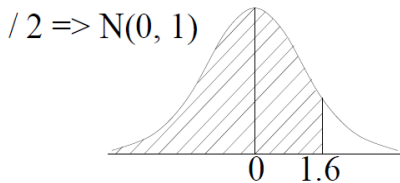
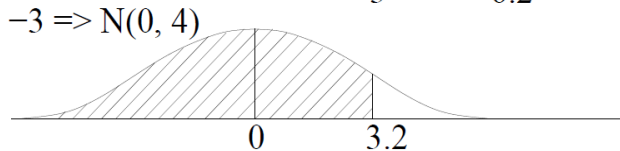
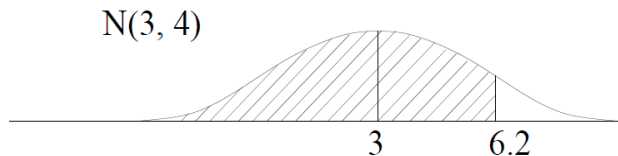
Recall that  $\Phi(-z) = 1 - \Phi(z)$ .

## Theorem

*If  $X \sim N(\mu, \sigma^2)$  and  $Y = \frac{X - \mu}{\sigma}$ , then  $Y \sim N(0, 1)$ .*

## Standard normal distribution (2)

Suppose  $X \sim N(3, 4)$ , what is  $P(X < 6.2)$ ?



## Standard normal distribution (3)

The probability that a random variable  $X$ ,  $X \sim N(\mu, \sigma^2)$ , is less or equal to number (value)  $x$  can be found using the following formula:

$$P(X \leq x) = F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right),$$

and it follows that

$$P(a \leq X \leq b) = F(b) - F(a) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right).$$

If  $Z \sim N(0, 1)$ , then the probability that  $Z$  lies between two values  $a$  and  $b$  is

$$P(a < Z \leq b) = \Phi(b) - \Phi(a).$$

# Example 1.1

## Example

If the random variable  $X$  is described by the distribution  $N(60, 0.0004)$  then what is the transformation to obtain the standardised normal variable?

## Example 1.1

### Example

If the random variable  $X$  is described by the distribution  $N(60, 0.0004)$  then what is the transformation to obtain the standardised normal variable?

Here,  $\mu = 60$  and  $\sigma^2 = 0.0004$  so that  $\sigma = 0.02$ . Hence  $Z = \frac{X-60}{0.02}$  is the required transformation.

\* When the random variable  $X$  takes values between 59.95 and 60.05, between which values does the random variable  $Z$  lie?

## Example 1.1

### Example

If the random variable  $X$  is described by the distribution  $N(60, 0.0004)$  then what is the transformation to obtain the standardised normal variable?

Here,  $\mu = 60$  and  $\sigma^2 = 0.0004$  so that  $\sigma = 0.02$ . Hence  $Z = \frac{X-60}{0.02}$  is the required transformation.

\* When the random variable  $X$  takes values between 59.95 and 60.05, between which values does the random variable  $Z$  lie?

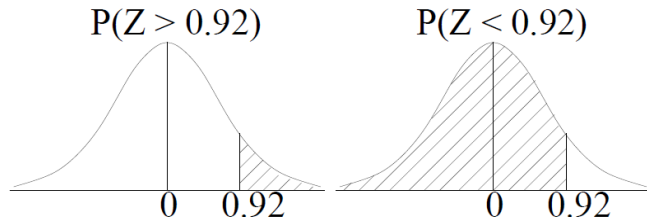
$$X = 60.05 \Rightarrow Z = \frac{60.05 - 60}{0.02} = 2.5,$$

$$X = 59.95 \Rightarrow Z = \frac{59.95 - 60}{0.02} = -2.5$$

Hence  $Z$  lies between -2.5 and 2.5.

## Example 2.1

If  $Z \sim N(0, 1)$ , what is  $P(Z > 0.92)$ ?



We know that  $P(Z > 0.92) = 1 - P(Z < 0.92)$  and we can calculate  $P(Z < 0.92)$  from table. Thus,  
 $P(Z > 0.92) = 1 - 0.8212 = 0.1788$ .

## Example 3.1

### Example

Let us assume that a person's intelligence index IQ is a random variable  $X$  with the distribution  $N(125, 100)$ . Find the probability that randomly chosen person has IQ **a)** lower than 110, **b)** greater or equal to 130.

## Example 3.1

### Example

Let us assume that a person's intelligence index IQ is a random variable  $X$  with the distribution  $N(125, 100)$ . Find the probability that randomly chosen person has IQ **a)** lower than 110, **b)** greater or equal to 130.

**a)**

$$P(X < 110) = F(110) = \Phi\left(\frac{110 - 125}{10}\right) = \Phi(-1.5) = 0.067.$$

Hence, the probability that a randomly chosen person's IQ is under 110, is 0.067.

## Example 3.1

### Example

Let us assume that a person's intelligence index IQ is a random variable  $X$  with the distribution  $N(125, 100)$ . Find the probability that randomly chosen person has IQ **a)** lower than 110, **b)** greater or equal to 130.

**a)**

$$P(X < 110) = F(110) = \Phi\left(\frac{110 - 125}{10}\right) = \Phi(-1.5) = 0.067.$$

Hence, the probability that a randomly chosen person's IQ is under 110, is 0.067.

**b)**

$$\begin{aligned} P(X \geq 130) &= 1 - P(X < 130) = 1 - \Phi\left(\frac{130 - 125}{10}\right) \\ &= 1 - \Phi(0.5) = 1 - 0.69 = 0.31. \end{aligned}$$

If  $X$  and  $Y$  are two *independent* normal variable such that  $X \sim N(\mu_1, \sigma_1^2)$  and  $Y \sim N(\mu_2, \sigma_2^2)$ , then

- $X - Y \sim N(\mu_1 - \mu_2, \sigma_2^2 + \sigma_1^2)$
- $X + Y \sim N(\mu_1 + \mu_2, \sigma_2^2 + \sigma_1^2)$
- $aX \sim N(a\mu_1, a^2\sigma_1^2)$
- $aX + bY \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$

# Why is the normal distribution useful?

- Many things actually are normally distributed, or very close to it. For example, height and intelligence are approximately normally distributed; measurement errors also often have a normal distribution.
- The normal distribution is easy to work with mathematically. In many practical cases, the methods developed using normal theory work quite well even when the distribution is not normal.
- There is a very strong connection between the size of a sample  $n$  and the extent to which a sampling distribution approaches the normal form. Many sampling distributions based on large  $n$  can be approximated by the normal distribution even though the population distribution itself is definitely not normal.

# Distribution of sums of squared random variables

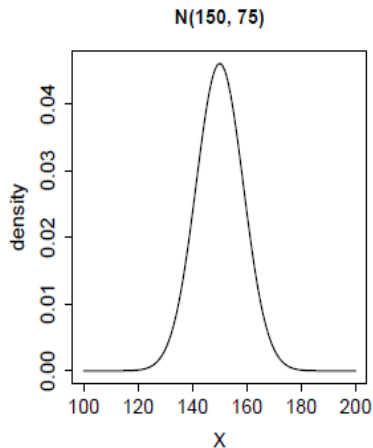
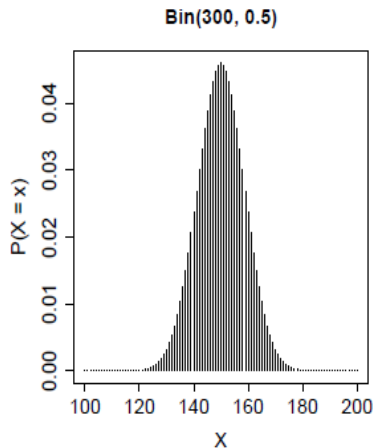
If  $X_1, X_2, \dots, X_m$  are  $m$  independent random variables having the standard normal distribution, then

$$Y = X_1^2 + X_2^2 + \dots + X_m^2$$

follows a Chi-Squared distribution with  $m$  degrees of freedom:  
 $Y \sim \chi^2(m)$ .

# The Normal approximation to the Binomial

Under certain conditions we can use the Normal distribution to approximate the Binomial distribution.



# The normal approximation to the Binomial

When you have a binomial distribution where  $n$  is large and  $p$  is middle-of-the road (not too small, not too big, close to 0.5), then the graph of binomial probabilities starts to look like the density of a normal distribution.

So, for large  $n$  the binomial distribution can be approximated by a normal distribution.  $X$  is then approximately normally distributed with mean  $E(X) = np$  and variance  $Var(X) = npq$ , so we have approximately  $X \sim N(np, npq)$ . It follows

$$P(a \leq X \leq b) \approx \Phi\left(\frac{b - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - np}{\sqrt{npq}}\right).$$

What to do when  $X \sim B(n, p)$  and  $p$  is small,  $n$  is relatively big?  
Then  $X$  can be approximated by  $Y \sim Po(np)$ .

EXAMPLE. Let  $X$  be a random variable which counts, how many of you have been born today, September 20.

$$X \sim B(38, \frac{1}{365}).$$

What is the probability that at least two of you have birthday today?

$$P(X \geq 2) = 1 - (P(X = 0) + P(X = 1)).$$

$P(X = k) = C_{38}^k (\frac{1}{365})^k (\frac{364}{365})^{38-k}$ . Lot of work!

But if  $Y \sim Po(38 \times \frac{1}{365})$  then  $38 \times \frac{1}{365} \approx 0.1$ , and  
 $P(Y = 0) + P(Y = 1) \approx e^{-0.1} + e^{-0.1} \times 0.1 \approx 0.995$ .

$$P(X \geq 2) \approx 0.005$$

Much simpler!

## Central Limit Theorem

If  $X_1, X_2, \dots$  is an infinite sequence of independent identically distributed random variables, each with expectation  $\mu$  and standard deviation  $\sigma > 0$ , and if we set

$$Y_n = X_1 + \dots + X_n.$$

then we have

$$P\left(a < \frac{Y_n - n\mu}{\sigma\sqrt{n}} < b\right) \rightarrow \Phi(b) - \Phi(a), \quad n \rightarrow \infty.$$

In other words, the central limit theorem states that the distribution of the sum (or average) of a large number of independent, identically distributed variables will be approximately normal, regardless of the underlying distribution.

## Definition

If  $Z_n, n = 1, 2, \dots$ , is an infinite sequence of random variables and it is possible to find numbers  $A_n$  and  $B_n, n = 1, 2, \dots$ , such that

$$P(a < \frac{Z_n - A_n}{\sqrt{B_n}} < b) \rightarrow \Phi(b) - \Phi(a), \quad n \rightarrow \infty.$$

then  $Z_n$  is said to be **asymptotically normally distributed** with parameters  $A_n$  and  $B_n$ , or more briefly

$$Z_n \sim \text{AsN}(A_n, B_n).$$

Using this definition, it follows from Central Limit Theorem that  $Y_n \sim \text{AsN}(n\mu, n\sigma^2)$ . Thus, for large  $n$  we may treat the rv  $Y_n$  as approximately normally distributed with the given parameters.

# Law of large numbers

Sum of i.i.d. random variables has another important property.

Let  $Y_n = X_1 + \dots + X_n$ ,  $X_i$  be i.i.d. random variables and

$E(X_i) = \mu$ .

Then  $\frac{Y_n}{n} \rightarrow \mu$ , when  $n \rightarrow \infty$ .

So, if  $E(X_k) = 0$  and  $\text{Var}(X_k) = 1$ , then

$$\frac{\sum_{k=1}^n X_k}{n} \rightarrow 0, \quad \frac{\sum_{k=1}^n X_k}{\sqrt{n}} \rightarrow Z \sim N(0, 1), \quad n \rightarrow \infty.$$

# Are my data normally distributed?

- Look at the histogram! Does it appear bell shaped?
- Compute descriptive characteristics – are mean, median, and mode close to each other?
- Is skewness close to zero? Is kurtosis close to 3?
- Run tests of normality (such as Kolmogorov-Smirnov). But, be cautious, highly influenced by sample size!
- ...

## First homework, maximum 3 points:

Using your mathematical skills show that

$$E(X) = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \dots = \mu$$

$$\text{Var}(X) = E(X - EX)^2 = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \dots = \sigma^2$$

Prove that  $E(X) = \lambda$  when  $X \sim \text{Po}(\lambda)$ .

**Deadline: 28.09.2017 at 8.15am** (in the beginning of the Lab)