Math Revision Session

Statistics (4): Representative Continuous Random Variables

Jukina HATAKEYAMA

The University of Osaka, Department of Economics

April 8, 2026

Normal Distribution

The normal distribution, also called the **Gaussian distribution**, is one of the most important continuous probability distributions.

It is characterised by two parameters:

- the mean μ,
- the variance σ^2 (or standard deviation σ).

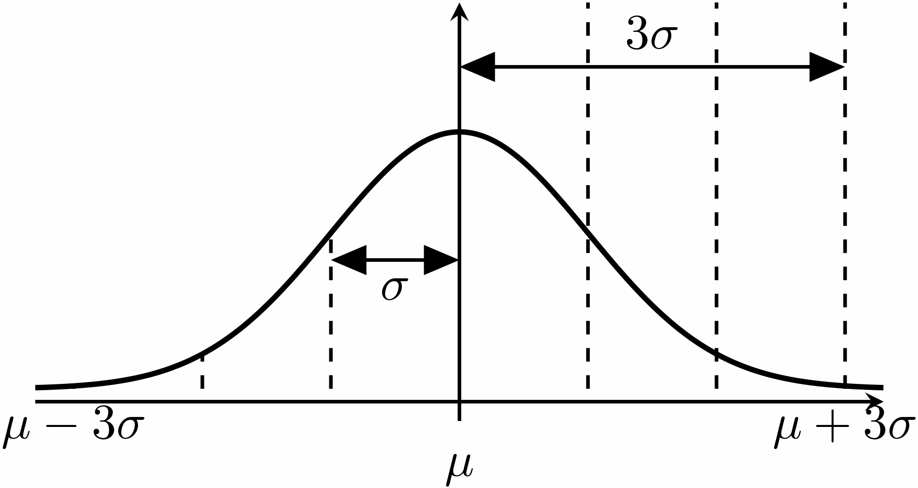
Its probability density function (PDF) is

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \qquad x \in \mathbb{R}.$$

Shape of the Normal Distribution

- The graph of the normal distribution is bell-shaped.
- It is symmetric around the mean μ.
- The total area under the curve is 1.
- The normal distribution is important partly because of the Central Limit Theorem.

Graph of a Normal Distribution



The 3σ Rule

For a normal distribution,

- about 68% of the probability mass lies within $\mu \pm \sigma$,
- about 95% lies within $\mu \pm 2\sigma$,
- about 99.7% lies within $\mu \pm 3\sigma$.

So the interval

$$[\mu - 3\sigma,\ \mu + 3\sigma]$$

covers almost all of the distribution.

Notation

If a random variable X follows a normal distribution with mean μ and variance σ^2, we write

$$X \sim N(\mu, \sigma^2).$$

Then

$$E[X] = \mu, \qquad \mathrm{Var}(X) = \sigma^2.$$

Standard Normal Distribution

The **standard normal distribution** is the special case

$$Z \sim N(0, 1).$$

Its PDF is

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right),$$

and its CDF is

$$\Phi(z) = \int_{-\infty}^{z} \phi(t)\, dt.$$

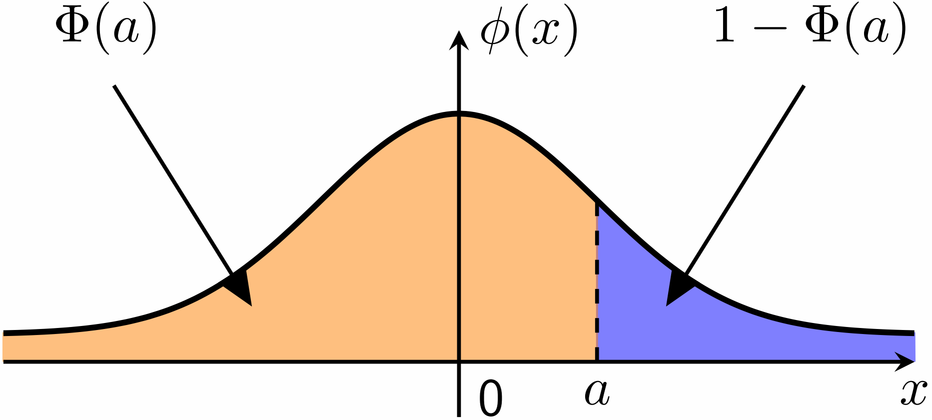
Symmetry of the Standard Normal Distribution

Because the standard normal distribution is symmetric about 0,

$$\phi(-z) = \phi(z), \qquad \Phi(-z) = 1 - \Phi(z).$$

This symmetry is very useful in probability calculations.

Graph of the Standard Normal Distribution



Linear Transformation of a Normal Variable

If

$$X \sim N(\mu, \sigma^2),$$

then for constants a and b,

$$Y = aX + b$$

also follows a normal distribution.
In fact,

$$E[Y] = a\mu + b, \qquad \mathrm{Var}(Y) = a^2\sigma^2,$$

so

$$Y \sim N(a\mu + b,\ a^2\sigma^2).$$

Standardisation

Given

$$X \sim N(\mu, \sigma^2),$$

define

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

Then

$$Z \sim N(0, 1).$$

This is called **standardisation** or the **Z-score transformation**.

Chi-Squared Distribution

Let

$$Z_1, Z_2, \ldots, Z_n$$

be independent standard normal random variables:

$$Z_i \sim N(0, 1).$$

Define

$$W = \sum_{i=1}^{n} Z_i^2.$$

Then W follows a chi-squared distribution with n degrees of freedom:

$$W \sim \chi^2(n).$$

Interpretation of the Chi-Squared Distribution

The chi-squared distribution is:

- always nonnegative,
- right-skewed for small degrees of freedom,
- closer to a normal shape when the degrees of freedom are large.

It is widely used in statistical inference, especially in hypothesis testing.

Mean and Variance of the Chi-Squared Distribution

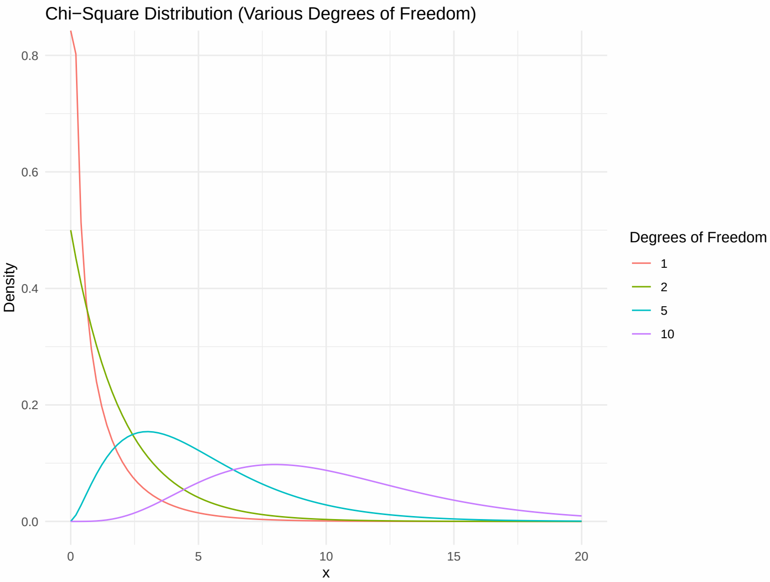
If

$$W \sim \chi^2(n),$$

then

$$E[W] = n, \qquad \mathrm{Var}(W) = 2n.$$

Graph of the Chi-Squared Distribution



1 Normal Distribution

2 Standard Normal Distribution

3 Chi-Squared Distribution

4 t-Distribution

5 Exponential Distribution

6 Degrees of Freedom

7 Summary

t-Distribution

Let

$$Z \sim N(0,1), \qquad W \sim \chi^2(n),$$

and assume that Z and W are independent.
Define

$$T = \frac{Z}{\sqrt{W/n}}.$$

Then T follows a t-distribution with n degrees of freedom:

$$T \sim t(n).$$

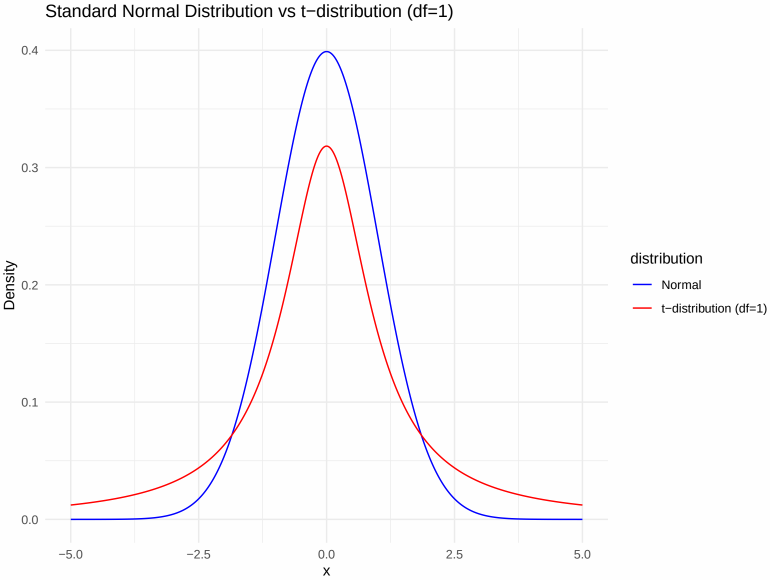
Characteristics of the t-Distribution

The t-distribution:

- is symmetric around 0,
- has heavier tails than the standard normal distribution,
- approaches the standard normal distribution as the degrees of freedom increase.

It is especially important when the sample size is small and the population variance is unknown.

Graph: Normal vs t-Distribution



Expectation and Variance of the t-Distribution

If

$$T \sim t(n),$$

then

$$E[T] = 0 \qquad \text{for } n > 1,$$

and

$$\mathrm{Var}(T) = \frac{n}{n-2} \qquad \text{for } n > 2.$$

For $n \leq 1$, the mean does not exist, and for $n \leq 2$, the variance does not exist.

Exponential Distribution

A random variable X follows an **exponential distribution** with rate parameter $\lambda > 0$ if its PDF is

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

We write

$$X \sim \mathrm{Exp}(\lambda).$$

Mean and Variance of the Exponential Distribution

If

$$X \sim \mathrm{Exp}(\lambda),$$

then

$$E[X] = \frac{1}{\lambda}, \qquad \mathrm{Var}(X) = \frac{1}{\lambda^2}.$$

Memoryless Property

The exponential distribution is **memoryless**:

$$\Pr(X > x + y \mid X > x) = \Pr(X > y).$$

This means that the remaining waiting time does not depend on how much time has already passed.

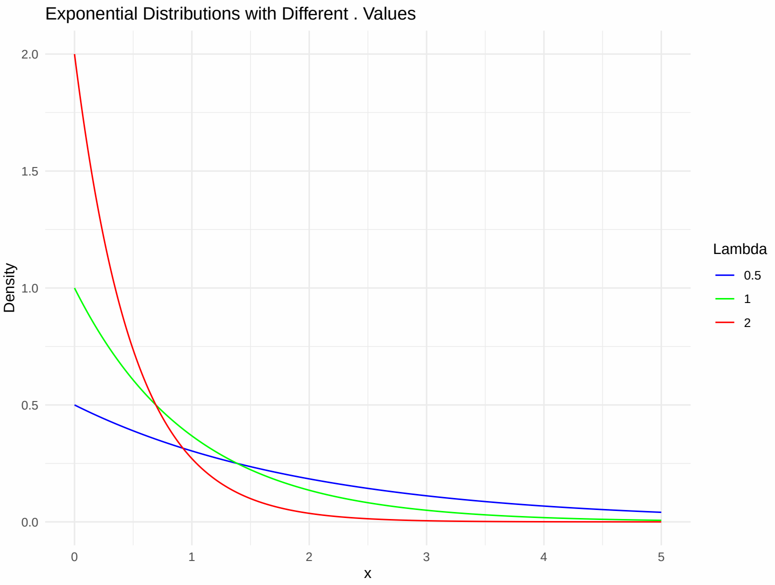
Interpretation as Waiting Time

The exponential distribution is often used to model waiting times between events in a Poisson process.

Examples:

- time between customer arrivals,
- time until failure of a machine,
- waiting time until the next event.

Graph of Exponential Distributions



Degrees of Freedom

Degrees of freedom refer to the number of independent pieces of information available after accounting for constraints.

In general:

$$\text{degrees of freedom} = \text{number of values} - \text{number of constraints.}$$

Example: Sample Mean

Suppose we have n observations

$$X_1, X_2, \ldots, X_n,$$

and their sample mean is fixed.
Once we know the first $n-1$ values, the last value is determined by the sample mean constraint.
So there are only

$$n-1$$

independent pieces of information.

Why Does This Matter?

Degrees of freedom are important because they affect:

- the sample variance,
- the t-distribution,
- the chi-squared distribution,
- statistical tests more generally.

As the degrees of freedom increase, some distributions become closer to the normal distribution.

A Simple Constraint Example

Suppose

$$x + y = 1.$$

Then once x is chosen, y is determined:

$$y = 1 - x.$$

So although there are two variables, there is only one degree of freedom.

Summary

In this lecture, we studied:

- the normal and standard normal distributions,
- standardisation,
- the chi-squared distribution,
- the t-distribution,
- the exponential distribution,
- degrees of freedom.