RELIABILITY THEORY

Alexander A. Yefremov
alexyefremov@tpu.ru

Tomsk Polytechnic University, 2025

Contact Details

Alexander A. YEFREMOV

Senior Lecturer,
Division for Automatics and Robotics,
School of Computer Science & Robotics

Office 112a, building 10
email: *alexyefremov@tpu.ru*

INTRODUCTION

As per IEEE standards, *reliability* is defined as the ability of a system or component to perform its required functions under stated conditions for a specified period of time.

The key elements of the definition are

- ability,
- required function,
- conditions,
- and specified period of time.

Ability is expressed quantitatively with probability.

Required function relates to expected performance.

Stated conditions usually refer to environmental conditions of operation.

Specified period of time is also referred as mission time which provides expected duration of operation.

We can distinguish between three main branches of reliability theory:

- Hardware reliability
- Software reliability
- Human reliability

This course is concerned with the first of these branches: the reliability of technical components and systems.

Introduction

Within hardware reliability we may use two different approaches:

- The physical approach
- The actuarial (probabilistic) approach

In the physical approach the strength of a technical item is modeled as a random variable S. The item is exposed to a load L that is also modeled as a random variable.

A failure will occur as soon as the load is higher than the strength.

The physical approach is mainly used for reliability analyses of structural elements, like beams and bridges. The approach is therefore often called *structural reliability analysis*.

In the actuarial approach, we describe all our information about the operating loads and the strength in the probability distribution function $F(t)$ of the time to failure T.

Introduction

Various approaches can be used to model the reliability of systems of several components and to include maintenance and replacement of components.

When several components are combined into a system, the analysis is called a *system reliability analysis*.

Introduction

The concept of reliability as a probability means that any attempt to quantify it must involve the use of probabilistic and statistical methods.

An understanding of probability theory and statistics as applicable to reliability engineering is therefore a necessary basis for progress.

The Institute of Electrical and Electronics Engineers (IEEE) – *Eye-triple-E*

stated (*adj*) – *заданный, известный*

specified (*adj*) – *заданный заранее, конкретный,*
[ˈspesəˌfaɪd] *точно определенный*

maintenance (*n*) – *техническое обслуживание и ремонт,*
[ˈmeɪnt(ə)nəns] *поддержание работоспособного состояния*

to maintain (*v*) – *поддерживать, содержать*
[meɪnˈteɪn]

Introduction

to define / definition / defined, predefined;
to perform / performance;
to rely / reliance, reliability;
to state / statement;
to specify / specification / specific;
to relate / relation / related;
to refer / reference;
to occur / occurrence;
to deduce / deduction;
to vary / variety, variance;
to note / notation;
to consider / consideration;

to provide;
to distinguish;
to concern;
to involve;
to interact;
to expose;
to deteriorate;
to obtain;
to approach;
to propose;
to prefer.

BASIC CONCEPTS
OF
PROBABILITY THEORY

Randomness is the lack of pattern or predictability in events.

Individual random events are by definition unpredictable, but in many cases the frequency of different outcomes over a large number of events (or "trials") is predictable.

Randomness is a measure of uncertainty of an outcome.

Basic Concepts of Probability Theory

Probability is the measure of the likelihood that an event will occur.

Probability quantifies as a number between 0 and 1, where, loosely speaking, 0 indicates impossibility and 1 indicates certainty. The higher the probability of an event, the more likely it is that the event will occur.

In any situation in which one of a number of possible outcomes may occur, the theory of probability provides methods for quantifying the chances, or likelihoods, associated with the various outcomes.

In probability, an experiment refers to any action or activity whose outcome is subject to uncertainty.

Although the word "experiment" generally suggests a planned or carefully controlled laboratory testing situation, we use it here in a much wider sense.

Thus experiments that may be of interest include tossing a coin once or several times, rolling a die or several dice, measuring time to failure of a particular device.

The sample space of an experiment, denoted by $\boldsymbol{\Omega}$, is the set of all possible outcomes of that experiment.

Basic Concepts of Probability Theory

When dealing with experiments that are random, probabilities can be numerically described by the number of desired outcomes divided by the total number of all outcomes.

$$P = \frac{n}{N}$$

The probability of an event *A* is written as

$$P(A), p(A), Pr\{A\}$$

In probability and statistics, a random variable is a variable whose possible values are outcomes of a random phenomenon.

More specifically, a random variable X is defined as a function that maps the outcomes of an unpredictable process to numerical quantities, typically *real numbers*.

$$X:\ \Omega \rightarrow D$$

We can consider other sets D, such as Boolean values, complex numbers, matrices, categorical values, etc.

Random variables can be classified into two categories, namely, *discrete* and *continuous* random variables.

A random variable is said to be *discrete* if its *sample space is countable*.

If the elements of the sample space are *infinite in number* and *sample space is continuous*, the random variable defined over such a sample space is known as *continuous* random variable.

Basic Concepts of Probability Theory

event [ɪˈvɛnt] - *событие*

outcome [ˈaʊtkʌm] - *исход, результат, следствие*

trial [ˈtrʌɪəl] - *испытание*

to occur [əˈkəː] - *происходить, случаться*

occurrence [əˈkʌr(ə)ns] - *возникновение, появление*

uncertainty [ʌnˈsəːt(ə)nti] - не*определённость*

sample [ˈsɑːmp(ə)l] - *выборка, образец, экземпляр*

sample space - *пространство элементарных событий (исходов)*

to map - *отображать*

Try to explain the difference between these concepts:

- *possibility* (*possible, impossible*);
- *probability* (*probable, improbable*);
- *likelihood* (*likely, unlikely*).

DISTRIBUTION FUNCTIONS

Distribution Functions

In probability theory and statistics, a ***probability distribution*** is a mathematical function that provides the probabilities of occurrence of different possible outcomes in an experiment.

In more technical terms, the probability distribution is a description of a random phenomenon in terms of the probabilities of events.

To specify a random variable is to define a probability distribution function.

For a *discrete random variable*, the distribution is often specified by just a list of the possible values along with the probability of each.

In some cases, it is convenient to express the probability in terms of a formula.

A *probability mass function (pmf)* is a function that gives the probability that a discrete r.v. *X* is exactly equal to some value *x*.

$$f_X(x) = Pr\{X = x\}$$

In other words, for every possible value *x* of the random variable, the *pmf* specifies the probability of observing that value when the experiment is performed.

Each *pmf* must satisfy two conditions:

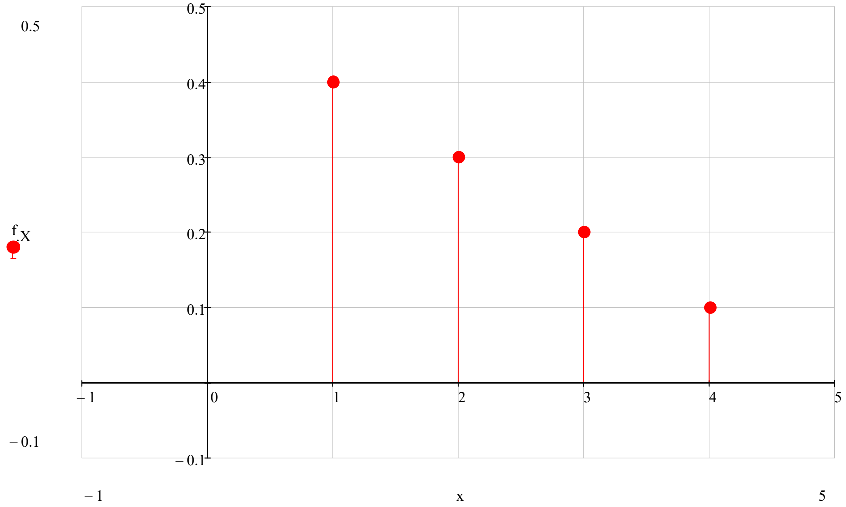
$$\forall x \in D \quad f_X(x) \geq 0$$

$$\sum_{x \in D} f_X(x) = 1$$

The *pmf* can be presented compactly in a tabular form:

x	1	2	3	4
$f_X(x)$	0.4	0.3	0.2	0.1

where any x value not listed receives zero probability.



The concept similar to *pmf*, but applied for a continuous random variable, is called ***probability density function (pdf)***.

A *pdf*, or density of a continuous r.v., is a function whose value at any given sample (or point) in the sample space can be interpreted as providing a *relative likelihood* that the value of the random variable would equal that sample.

A *pmf* differs from a *pdf* in that the values of the *pdf* are not probabilities as such: a *pdf* must be integrated over an interval to yield a probability.

$$Pr\{a \leq x \leq b\} = \int_a^b f_X(x)dx$$

Similar to *pmf*, any *pdf* must satisfy two conditions:

$$\forall x \in D \; f_X(x) \geq 0 \qquad \int_{x \in D} f_X(x)dx = 1$$

For some fixed value *x*, we often wish to compute the probability that the observed value of *X* will be at most *x*.

For example, let *X* be the number of beds occupied in a hospital's emergency room at a certain time of day, and suppose the *pmf* of *X* is given by

x	0	1	2	3	4
$f_X(x)$	0.2	0.25	0.3	0.15	0.1

Then the probability that at most two beds are occupied is

$$Pr\{X \leq 2\} = f_X(0) + f_X(1) + f_X(2) = 0.75$$

Furthermore, we also have $Pr\{X \leq 2.7\} = 0.75$,
and similarly $Pr\{X \leq 2.999\} = 0.75$.

Since 0 is the smallest possible *X* value,

$$Pr\{X \leq -1.5\} = 0, \qquad Pr\{X \leq -10\} = 0$$

and in fact for any negative number *x*, $Pr\{X \leq x\} = 0$.

And because 4 is the largest possible value of *X*,

$$Pr\{X \leq 4\} = 1, \qquad Pr\{X \leq 9.8\} = 1,$$

and so on.

Distribution Functions

In probability theory and statistics, the ***cumulative distribution function (cdf)*** of a real-valued random variable *X*, or just distribution function of *X*, evaluated at *x*, is the probability that *X* will take a value less than or equal to *x*.

$$F_X(x) = Pr\{X \leq x\}$$

In this definition, the "less than or equal to" sign is a convention, not a universally used one, but is important for discrete distributions.

Every *cdf* F_X is non-decreasing and right-continuous function. Furthermore,

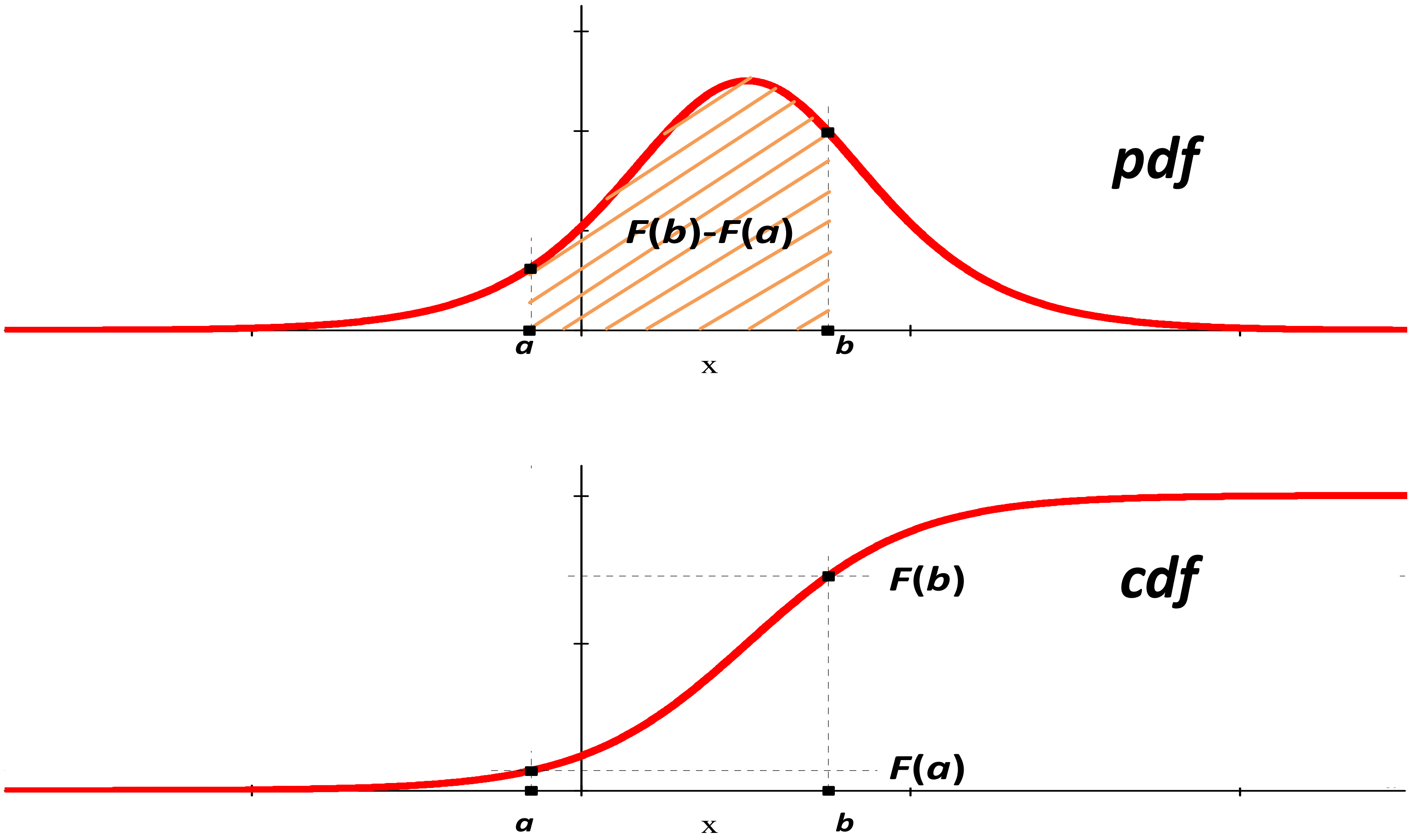
$$\lim_{x \to -\infty} F_X(x) = 0, \qquad \lim_{x \to +\infty} F_X(x) = 1$$

$$F_X(x) = \sum_{x_i \le x} f_X(x_i) = \int_{-\infty}^{x} f_X(\theta)\, d\theta$$

$$Pr\{a \le x \le b\} = F_X(b) - F_X(a) = \int_{a}^{b} f_X(x) dx$$

$$f_X(x) = \frac{dF_X(x)}{dx}$$

Distribution Functions



MOMENTS
OF RANDOM VARIABLES

Moments of Random Variables

In order for us to give a brief description of the distribution of a random variable, it is obviously not very convenient to present a table of the distribution function.

It would be better to present some suitable characteristics.

Two important classes of such characteristics are *measures of location* and *measures of dispersion*.

Moments of Random Variables

Let X be a r.v. with *cdf* F_X and *pdf* f_X. The most common measure of location is the mean or expected value, which is defined as

$$E[X] = M[X] = \mu_X = \begin{cases} \sum_{x_i \in D} x_i f_X(x_i) & \text{if } X \text{ is discrete} \\ \int_{x \in D} x f_X(x)\, dx & \text{if } X \text{ is continuous} \end{cases}$$

If we think of the distribution as the mass of some body, the mean corresponds to the center of gravity.

NB! The expected value may not exist!

Another measure of location is the median, which is a number *m* (not necessarily unique) such that

$$Pr\{X \leq m\} = F_X(m) = Pr\{X \geq m\} = 0.5$$

If the distribution is *symmetric*, then, clearly, the median and the mean coincide (provided that the latter exists).

If the distribution is skewed, the median might be a better measure of the "average" than the mean.

However, this also depends on the problem at hand.

It is clear that two distributions may well have the same mean and yet be very different. One way to distinguish them is via a measure of dispersion—by indicating how spread out the mass is.

The most commonly used such measure is the ***variance*** *Var(X)*, which is defined as

$$Var(X) = E[(x - \mu_X)^2]$$

Moments of Random Variables

The variance can be computed as

$$Var(X) = \begin{cases} \sum_{x_i \in D} (x_i - \mu_X)^2 f_X(x_i) & if\ X\ is\ discrete \\ \int_{x \in D} (x - \mu_X)^2 f_X(x)\, dx & if\ X\ is\ continuous \end{cases}$$

An alternative and, in general, more convenient way to compute the variance is via the relation

$$Var(X) = E[X^2] - E[X]^2$$

We also define the ***standard deviation*** of a random variable X by

$$\sigma_X = \sqrt{Var(X)}$$

It is often preferable to work with the standard deviation rather than with the variance of a random variable, because it is easier to interpret.

Indeed, the standard deviation is expressed in the same units of measure as X, whereas the units of the variance are the squared units of X.

Now we can generalize to introduce the concept of *raw moments* and *central moments* of a r.v.

Raw moment of the *k*-th order (*k*-th raw moment) is defined as

$$\mu_k = E[X^k] = \begin{cases} \sum_{x_i \in D} x_i^k f_X(x_i) & \text{if } X \text{ is discrete} \\ \int_{x \in D} x^k f_X(x)\, dx & \text{if } X \text{ is continuous} \end{cases}$$

Central moment of the k-th order (k-th central moment) is defined as

$$\nu_k = E[(X-\mu)^k] = \begin{cases} \sum_{x_i \in D} (x_i - \mu)^k f_X(x_i) & \text{if } X \text{ is discrete} \\ \int_{x \in D} (x-\mu)^k f_X(x)\, dx & \text{if } X \text{ is continuous} \end{cases}$$

Thus, the mean is the 1st raw moment, and the variance is the 2nd central moment.

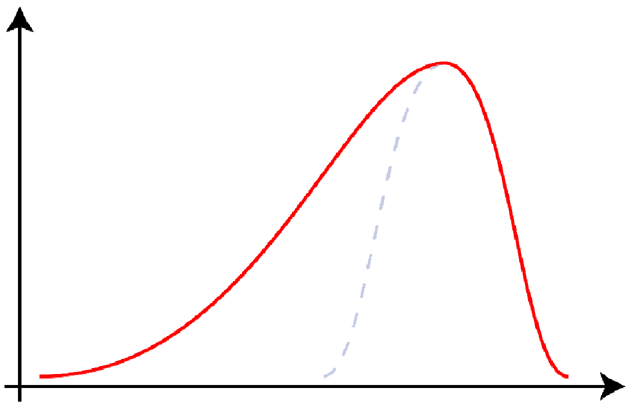
Beside the variance, 3rd and 4th central moments are often made use of in order to measure *skewness* and *kurtosis* of a random value.

Skewness is a measure of the *asymmetry* of the probability distribution of a real-valued r.v. about its mean. The skewness value can be positive or negative, or undefined.

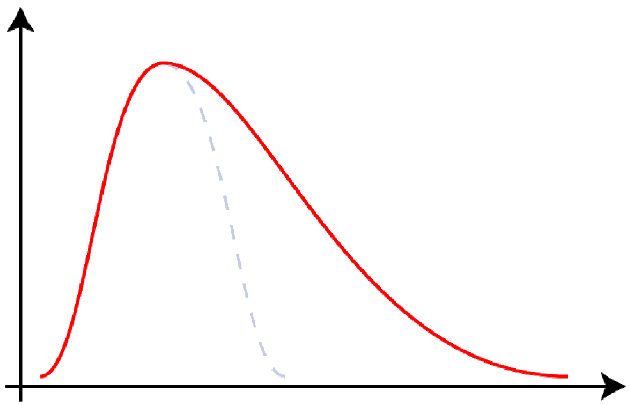
$$Sk[X] = \frac{\nu_3}{\sigma_X^3}$$

Within each graph, the values on the right side of the distribution taper differently from the values on the left side.

These tapering sides are called ***tails***, and they provide visual means to determine which of the two kinds of skewness a distribution has.

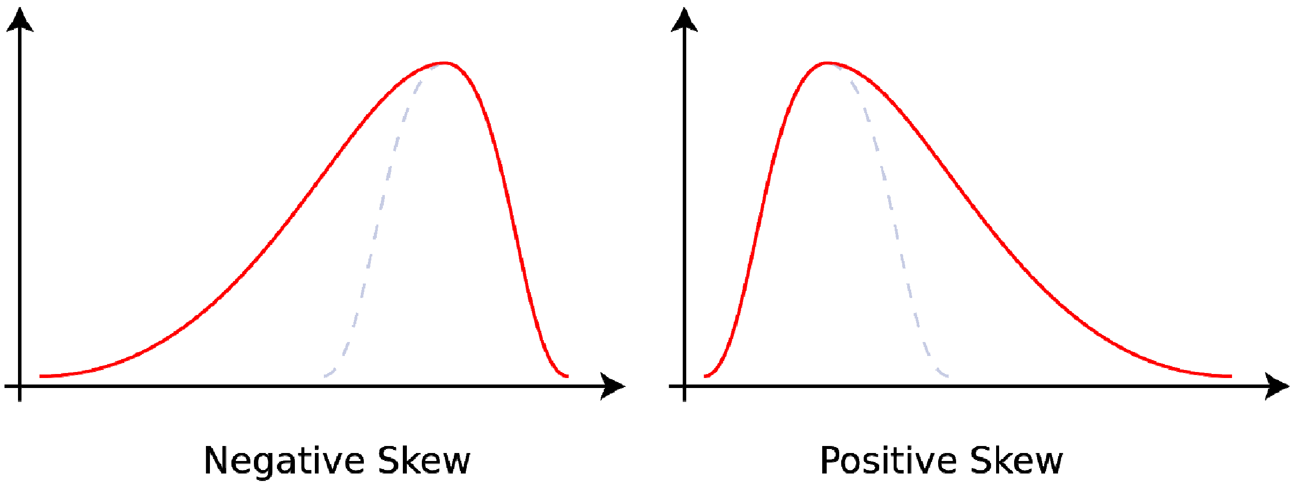


Negative Skew



Positive Skew

Moments of Random Variables



<u>negative skew</u>: The left tail is longer; the distribution is said to be *left-skewed*, *left-tailed*, or *skewed to the left*.
A left-skewed distribution usually appears as a *right-leaning curve*.

<u>positive skew</u>: The right tail is longer; the distribution is said to be *right-skewed*, *right-tailed*, or *skewed to the right*.
A right-skewed distribution usually appears as a *left-leaning curve*.

In probability theory and statistics, ***kurtosis*** is a measure of the "*tailedness*" of the probability distribution of a real-valued r.v.

$$Kurt[X] = \frac{\nu_4}{\sigma_X^4}$$

The kurtosis of any normal distribution is 3. It is common to compare the kurtosis of a distribution to this value.

If a distribution has kurtosis less than 3, it means the distribution produces *fewer* and *less extreme* outliers than does the normal distribution.

If a distribution has kurtosis greater than 3, it means the distribution produces more outliers than the normal distribution.

It is also common practice to use an adjusted version of kurtosis, the *excess kurtosis*, to provide the comparison to the normal distribution: $Ex(X) = Kurt(X) - 3$