Randomness – náhodnosť, náhoda

Definition:
The term randomness suggests unpredictability. A simple example of randomness is the tossing of a coin. The outcome is uncertain; it can either be an observed head (H) or an observed tail (T). Because the outcome of the toss cannot be predicted for sure, we say it displays randomness.

Note:
This is an example of an easily describable random process. However, other random processes can be quite intricate; for example, the fluctuating prices of stocks are difficult to explain because there are so many variables and combinations of variables that are influencing the prices.

Uncertainty – neistota

Definition:
At some time or another, everyone will experience *uncertainty*. For example, you are approaching the traffic signals and the light changes from green to amber. You have to decide whether you can make it through the intersection or not. You may be uncertain as to what the correct decision should be.

Probability – pravdepodobnosť

Definition:
The concept of *probability* is used to quantify this measure of doubt. If you believe that you have a 0.99 probability of getting across the intersection, you have made a clear statement about your doubt. The probability statement provides a great deal of information, much more than statements such as "Maybe I can make it across," "I should make it across," etc.

Sample space – výberový priestor

Definition:
The sample space for an experiment is the list or set of all possible outcomes for the experiment.

Event – jav, udalosť

Definition:
An event is a subset of the sample space.

Classical probability – klasická definícia pravdepodobnosti

Definition:
If the outcomes in a sample space are equally likely to occur, then the classical probability of an event A is defined to be:

$$P(A) = \frac{number\ of\ events\ in\ A}{total\ number\ of\ events\ in\ the\ sample\ space}$$

Random variable – náhodná veličina

Definition:
A random variable assigns one and only one numerical value to each point in the sample space for a random experiment.

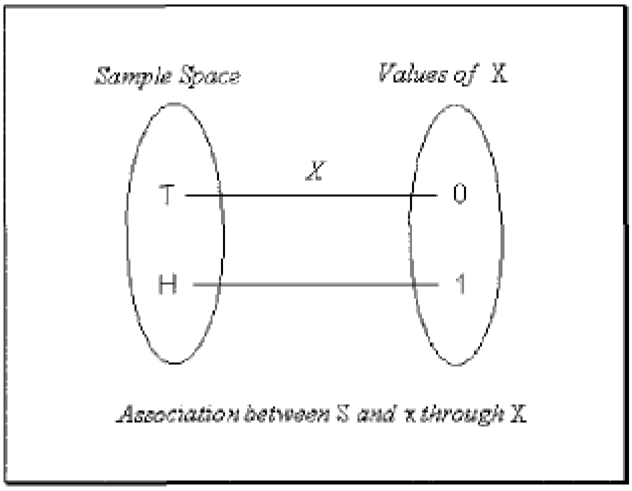


Fig. 8-1: Outcomes of tossing a single coin and values of X

Note:
Random variables are usually denoted by uppercase letters near the end of the alphabet, such as X, Y, and Z. We will use lowercase letters to represent the values of the random variables, such as x, y, and z.

We will encounter two types of random variables:
- discrete
- continuous.

Discrete random variable – diskrétna náhodná veličina

Definition:
A discrete random variable is one that can assume a countable number of possible values (f.e. the number of days it rained during the month of March).

Continuous random variable – spojitá náhodná veličina

Definition:
A continuous random variable is one that can assume any value in an interval on the real number line (f.e. the amount (in inches) of rainfall in your community during the month of March).

Normal distribution – normálne rozdelenie

Definition:
A normal distribution is a continuous, symmetrical, bell-shaped distribution of a normal random variable.

Summary of the Properties of the Normal Distribution:

- The curve is continuous.
- The curve is bell-shaped.
- The curve is symmetrical about the mean.
- The mean, median, and mode are located at the center of the distribution and are equal to each other.
- The curve is unimodal (single mode).
- The curve never touches the x axis.
- The total area under the normal curve is equal to 1.

The mathematical equation for the normal distribution is:

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

μ = population mean, and σ = population standard deviation.

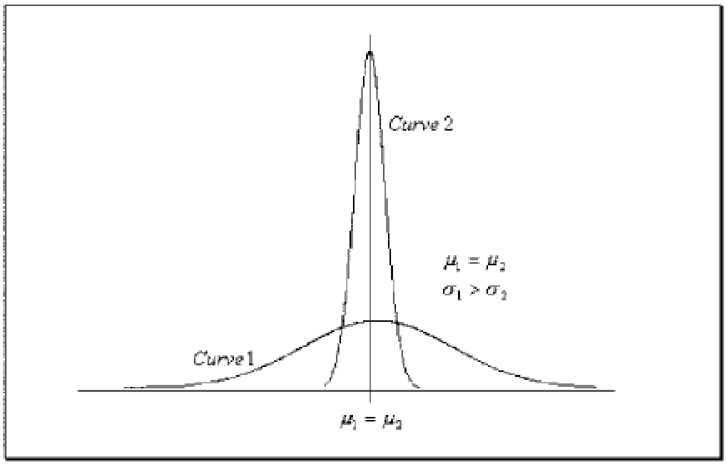


Fig. 9-8: Normal distributions with the same mean but with different standard deviations

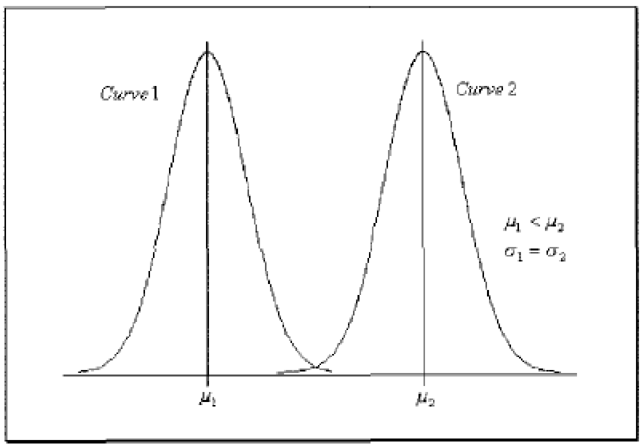


Fig. 9-9: Normal distributions with different means but with the same standard deviation

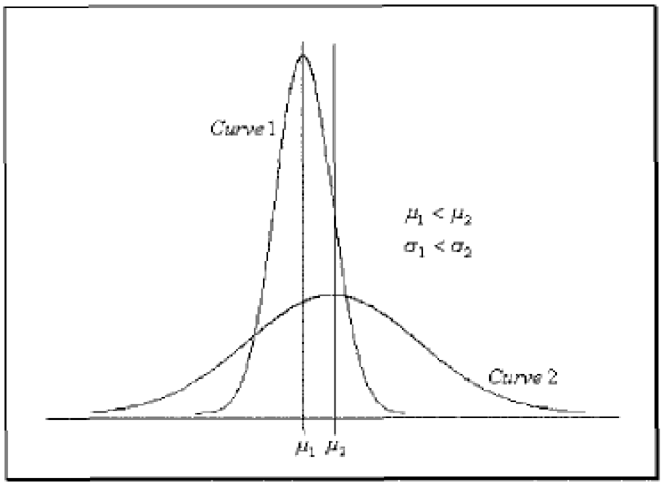


Fig. 9-10: Normal distributions with different means and different standard deviations

Empirical rule (rule of three sigma) – pravidlo troch sigma

Definition:
One-sigma rule: Approximately **68 percent** of the data values *should lie within one standard deviation of the mean*. That is, regardless of the shape of the normal distribution, the probability that a normal random variable will be within one standard deviation of the mean is approximately equal to 0.68. This is illustrated in Fig. 9-11.

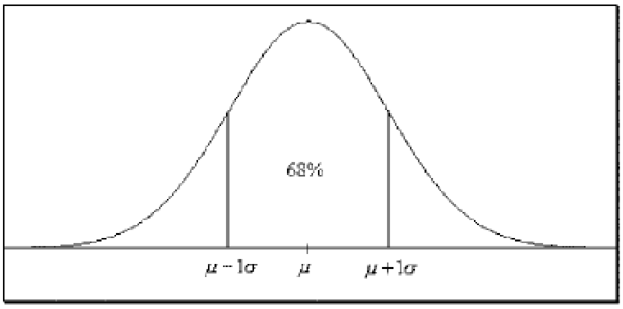


Fig. 9-11: One-sigma rule

Two-sigma rule: Approximately **95 percent** of the data values *should lie within two standard deviations of the mean.* That is, regardless of the shape of the normal distribution, the probability that a normal random variable will be within two standard deviations of the mean is approximately equal to 0.95. This is illustrated in Fig. 9-12.

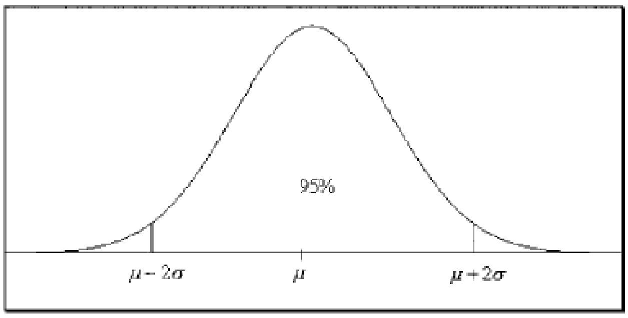


Fig. 9-12: Two-sigma rule

Three-sigma rule: Approximately **99.7 percent** of the data values *should lie within three standard deviations of the mean.* That is, regardless of the shape of the normal distribution, the probability that a normal random variable will be within three standard deviations of the mean is approximately equal to 0.997. This is illustrated in Fig. 9-13.

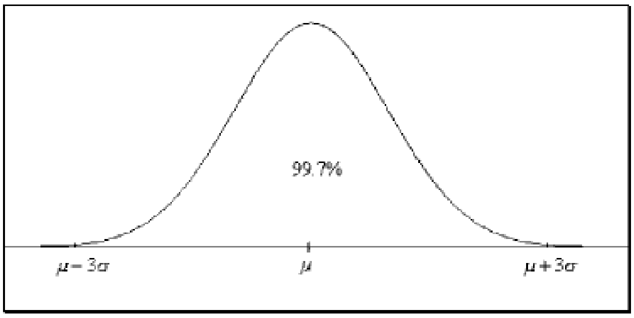


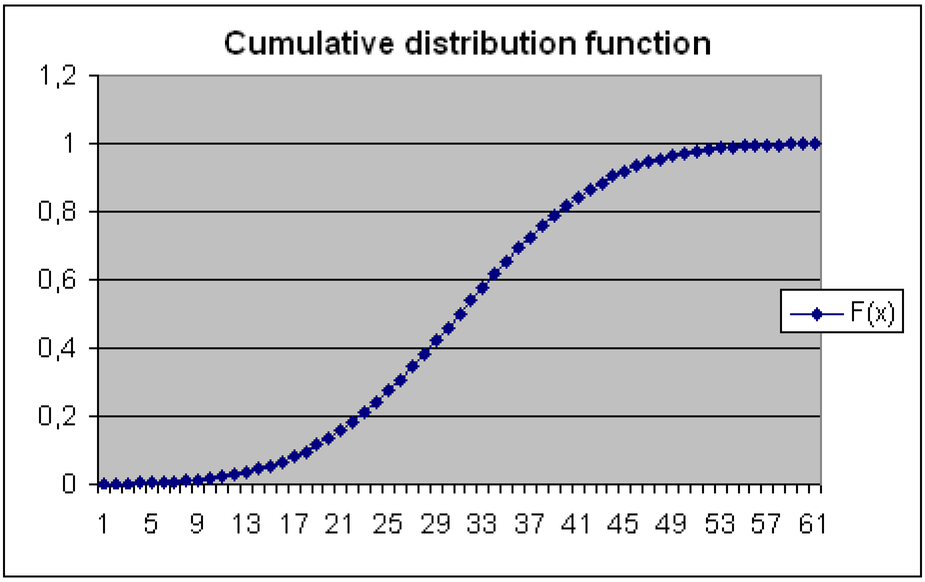
Fig. 9-13: Three-sigma rule

Cumulative distribution function – distribučná funkcia

Definition:
In probability theory, the **cumulative distribution function** (CDF), also called **probability distribution function** or just **distribution function**, completely describes the probability distribution of a real-valued random variable *X*. For every real number *x*, the CDF of X is given by

$$x \to F_X(x) = \mathrm{P}(X \leq x),$$

where the right-hand side represents the probability that the random variable *X* takes on a value less than or equal to *x*. It is conventional to use a capital *F* for a cumulative distribution function, in contrast to the lower-case *f* used for probability density functions and probability mass functions.



Probability density function – funkcia hustoty pravdepodobnosti

Definition:
A **probability density function (PDF)** is a function that represents a probability distribution in terms of integrals, such that the probability of the interval [*a*, *b*] is given by:

$$\int_a^b f(x)\,dx$$

for any two numbers *a* and *b*.

A **probability density function** is any function *f*(*x*) that describes the probability density in terms of the input variable *x* in a manner described below.

- *f*(*x*) is greater than or equal to zero for all values of *x*
- The total area under the graph is 1:
$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

The actual probability can then be calculated by taking the integral of the function *f*(*x*) by the integration interval of the input variable *x*.