GeMan
Maintenance Management

Statistics and System Reliability

SAPIENZA
UNIVERSITÀ DI ROMA

Lorenzo Fedele
Department of Mechanical and Aerospace Engineering

From Statistics to Reliability
Trustworthiness

TRUSTWORTHINESS (*Dependability*)

- Reliability (*Reliability*)
- Availability (*Availability*)
- Maintainability (*Maintainability*)
- Safety (*Safety*)

- **RAMS**

From Statistics to Reliability

Trustworthiness

RAMS

Acronym for the quantities on which the theory of Reliability is based, which through the study of laws of occurrence of failures seeks to deepen the understanding of systems related to the issues of:

1. Survival
2. Average lifespan
3. Percentage of uptime.

From Statistics to Reliability
Trustworthiness

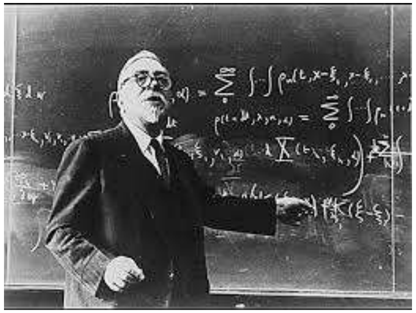
R'**Eliability** aims to investigate the trustworthiness characteristics of a system of components through:

- mathematical and statistical theories and methods
- organizational methods
- operational practices
- analysis techniques and methods
- numerical and computational analysis methods (algorithms)

From Statistics to Reliability
Systems Engineering

The systems

A **system** is a set of elements or components characterized by specific functional conditions and reliability, each of which contributes to the functioning of the entire system at a certain level of reliability.



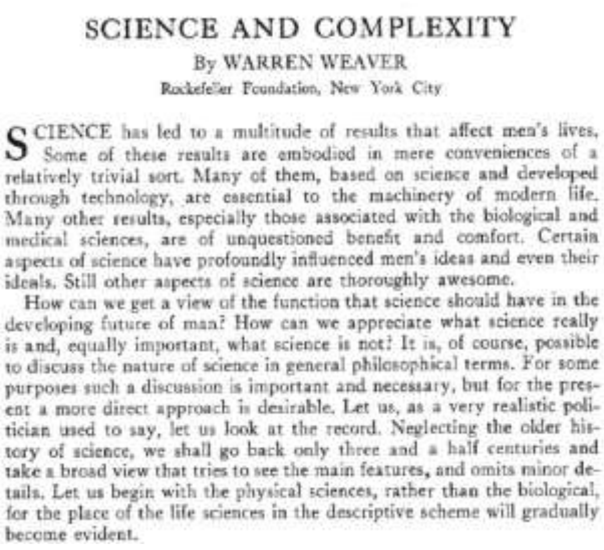
Warren Weaver
(Reedsburg, July 17, 1894 – New Milford, November 24, 1978)

Claude Shannon
(Petoskey, April 30, 1916 – Medford, February 24, 2001)

Herbert Simon
(Milwaukee, June 15, 1916 – Pittsburgh, February 9, 2001)

From Statistics to Reliability

Systems Engineering



SCIENCE AND COMPLEXITY

By WARREN WEAVER

Rockefeller Foundation, New York City

SCIENCE has led to a multitude of results that affect men's lives. Some of these results are embodied in mere conveniences of a relatively trivial sort. Many of them, based on science and developed through technology, are essential to the machinery of modern life. Many other results, especially those associated with the biological and medical sciences, are of unquestioned benefit and comfort. Certain aspects of science have profoundly influenced men's ideas and even their ideals. Still other aspects of science are thoroughly awesome.

How can we get a view of the function that science should have in the developing future of man? How can we appreciate what science really is and, equally important, what science is not? It is, of course, possible to discuss the nature of science in general philosophical terms. For some purposes such a discussion is important and necessary, but for the present a more direct approach is desirable. Let us, as a very realistic politician used to say, let us look at the record. Neglecting the older history of science, we shall go back only three and a half centuries and take a broad view that tries to see the main features, and omits minor details. Let us begin with the physical sciences, rather than the biological, for the place of the life sciences in the descriptive scheme will gradually become evident.

A Mathematical Theory of Communication

By C. E. SHANNON

INTRODUCTION

THE recent development of various methods of modulation such as PCM and PPM which exchange bandwidth for signal-to-noise ratio has intensified the interest in a general theory of communication. A basis for such a theory is contained in the important papers of Nyquist[1] and Hartley[2] on this subject. In the present paper we will extend the theory to include a number of new factors, in particular the effect of noise in the channel, and the savings possible due to the statistical structure of the original message and due to the nature of the final destination of the information.

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design.

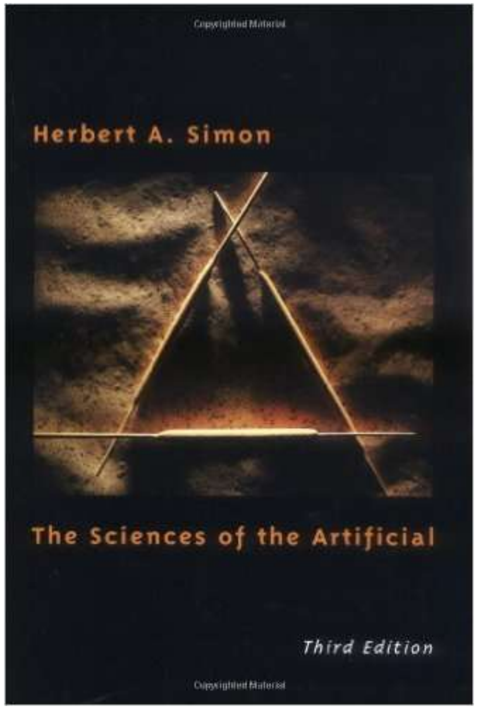
If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely. As was pointed out by Hartley the most natural choice is the logarithmic function. Although this definition must be generalized considerably when we consider the influence of the statistics of the message and when we have a continuous range of messages, we will in all cases use an essentially logarithmic measure.

The logarithmic measure is more convenient for various reasons:

1. It is practically more useful. Parameters of engineering importance

[1] Nyquist, H., "Certain Factors Affecting Telegraph Speed," *Bell System Technical Journal*, April 1924, p. 324; "Certain Topics in Telegraph Transmission Theory," *A. I. E. E. Trans.*, v. 47, April 1928, p. 617.

[2] Hartley, R. V. L., "Transmission of Information," *Bell System Technical Journal*, July 1928, p. 535.



✓ Systems and Complexity

The evolution of ***maintenance engineering*** *is closely linked to the development of* ***systems engineering****..*

The term **state of a system** at a certain moment denotes the set of relevant properties of all the elements of the system at that moment.

The complexity of a system generally increases with the number of component elements, with the multiplicity of existing relationships and with the multiplicity of states that the system can assume.

From Statistics to Reliability

Systems Engineering

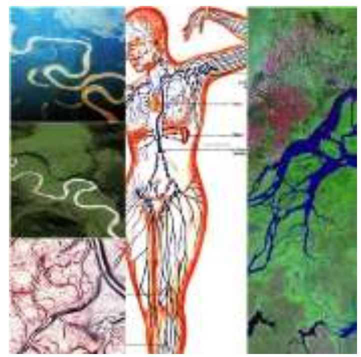
✓ Systems and Complexity

- The *complexity* has the characteristics of the multiple, the recursive, the uncertain, the unstable: it also pertains to the domain of information..
- Complexity is the property and characteristic size of any system, it increases with the rise in the **number of possible states**, so that the multiplicity (or the information required to identify each state) can be considered, in the first place, as a measure of the complexity of the system..

From Statistics to Reliability

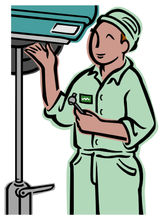
Systems Engineering

✓ Systems and Complexity



From Statistics to Reliability
Systems Engineering

✓ Systems and Complexity


1940 1970 2000



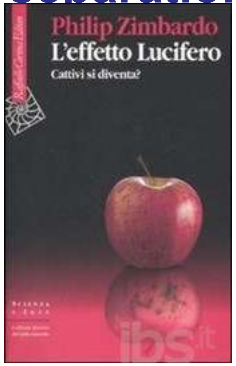


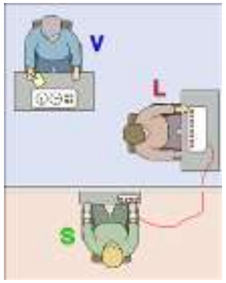

From Statistics to Reliability
Small worlds

The**small world theory** is a theory that claims that all complex networks present in nature are such that any two nodes can be connected by a path consisting ofa relatively small number of links..

Mathematically, the theory is studied as a branch ofgraph theory..

The small world theory originates from a series of experiments conducted by**StanleyMilgram** (psychologist, New York, 1933-1984) that examined the average path length forsocial networks among residents in the United States. (**"six degrees of separation"**).





In 1961 StanleyMilgramconducts a series of controversial experiments at Yale University.
The experiments involved ordinary people who were made to believe they were sending electric shocks to other people, in order totest the easy conditioning of humans towards authority..

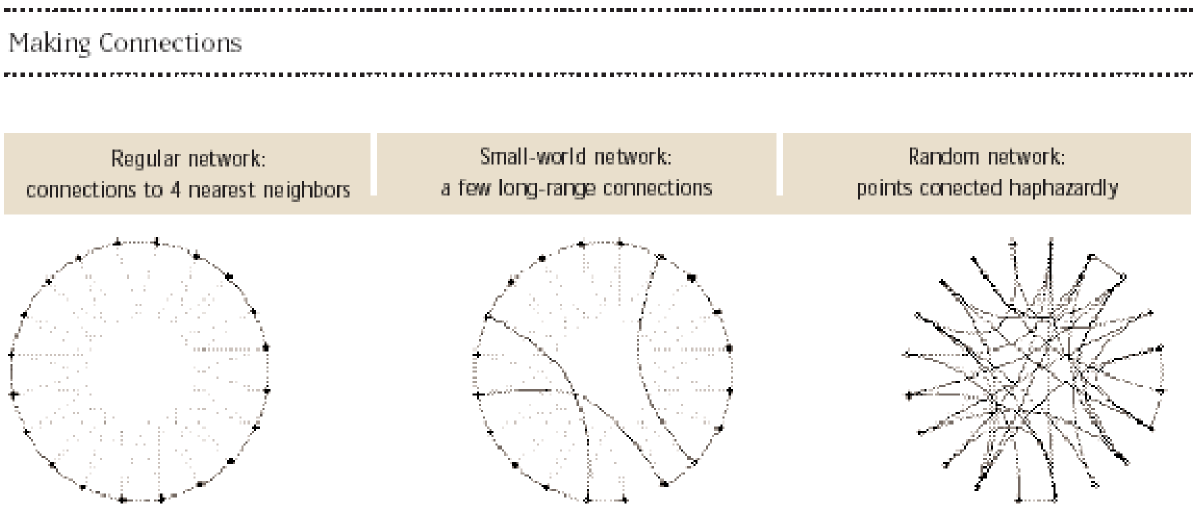
From Statistics to Reliability
Small worlds

The theory was developed in 1998 in the article published in the magazine*Nature.*, ”*Collective dynamics of «small-world» networks”*by mathematiciansDuncanWatts andStevenStrogatz.:

Networks of fireflies, routers, actors, sexual partners, etc. have at least two similar characteristics: thehigh level of clusteringand thelow degree of separation.:
despite the fact that each element tends to have relationships predominantly with a few others (high aggregation), this does not prevent obtaining its 'closeness' through a few intermediaries with any other element of the network (low degree of separation.)

The mathematician Paul Erdos studied random graphs, adding edges randomly between nodes in the given set.
Erdos proved that a small percentage of edges relative to the total is enough to have a connected graph.
the degree of separation of such graphs is extraordinarily small.

From Statistics to Reliability
Small worlds



Source: Watts & Strogatz, "Collective Dynamics of Small-World Networks," Nature, *June 1998.*

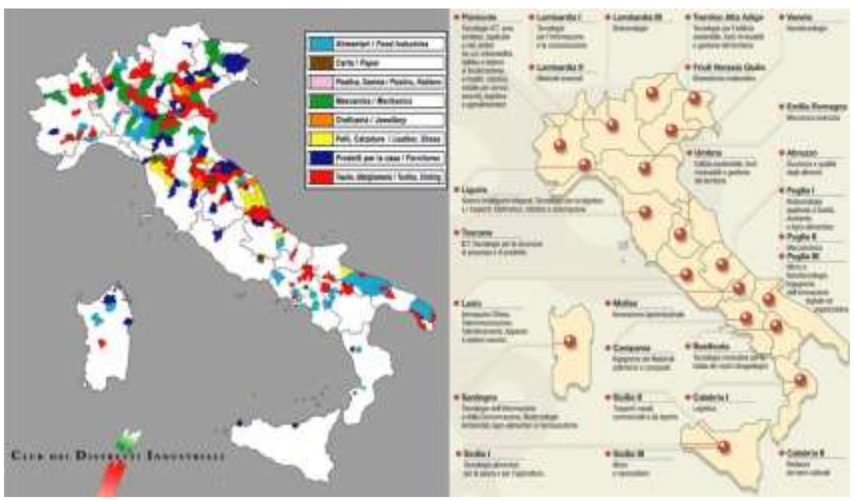
From Statistics to Reliability

Small worlds: the importance of weak ties

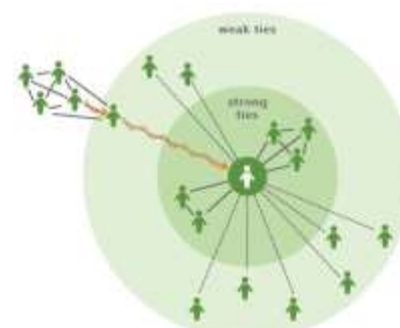
Mark Granovetter (Jersey City, October 20, 1943, American sociologist) demonstrated that:

individuals engaged in weak ties, that is, made of friendly acquaintances not too close, have more access to information<u>and thus to potential work positions of interest,</u> compared to those who socially invest only in strong ties, that is, family, relatives, and close friends

It is from this theory that in Italy derive the .



industrial districts**<u>, studies on social capital, informal economy, etc.</u>**Reliability analysis



Probability

One of the fundamental quantities on which Statistics is based is

Reliability analysis of Components and Systems

TOOL

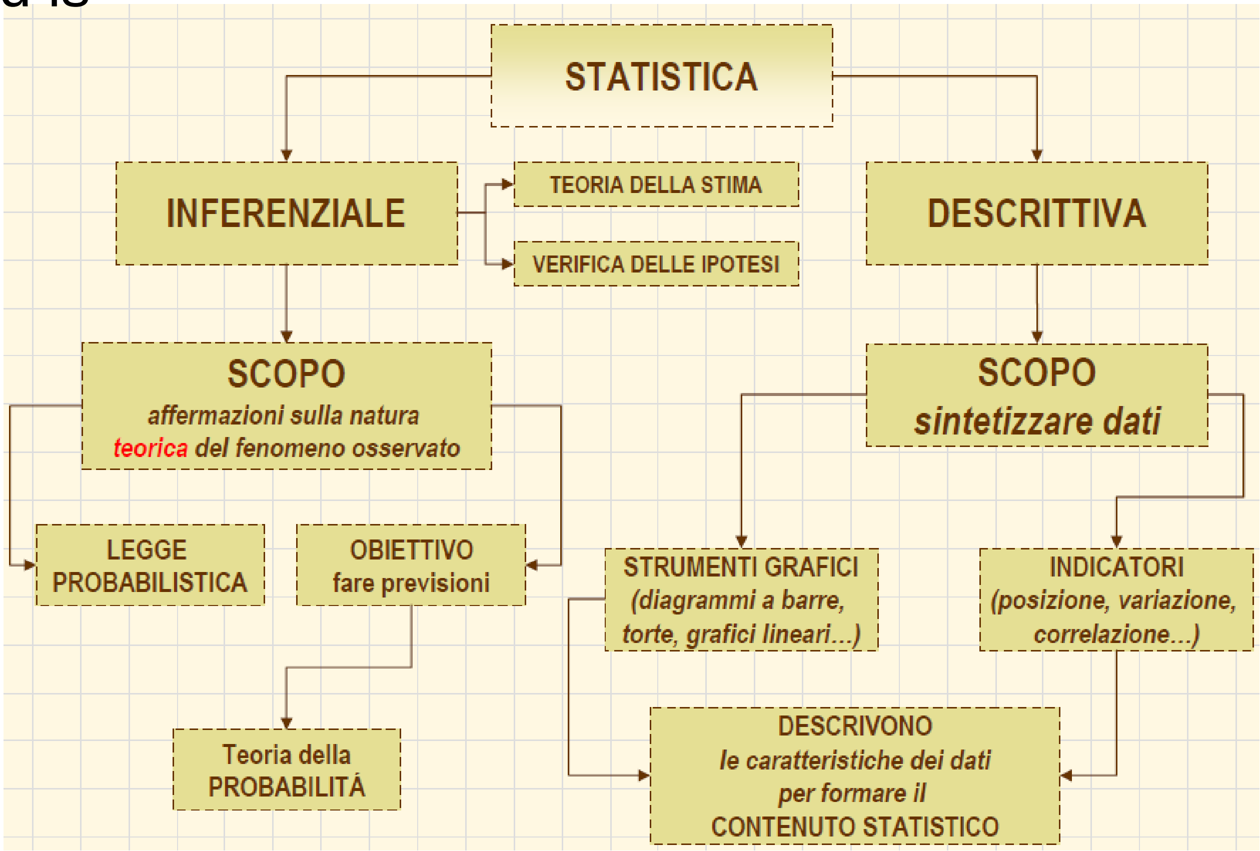
Statistics

Girolamo Cardano
(Pavia, 24 settembre 1501 – Roma, 21 settembre 1576 circa)

Fra Luca Bartolomeo de Pacioli o anche *Paciolo* (Borgo Sansepolcro, 1445 circa – Roma, 19 giugno 1517)

Probability

One of the fundamental quantities on which Statistics is based is



Probability
probability

- , introduced for the first time in the seventeenth century with gambling.**These games, therefore, involve actions with uncertain outcomes but with a** predictable

- In the case of tossing a sufficiently symmetrical and balanced coin, one can expect substantial equivalence between the times one face of the coin comes up and the other.***Probability has been the subject of*** .
- subsequent definitions over time

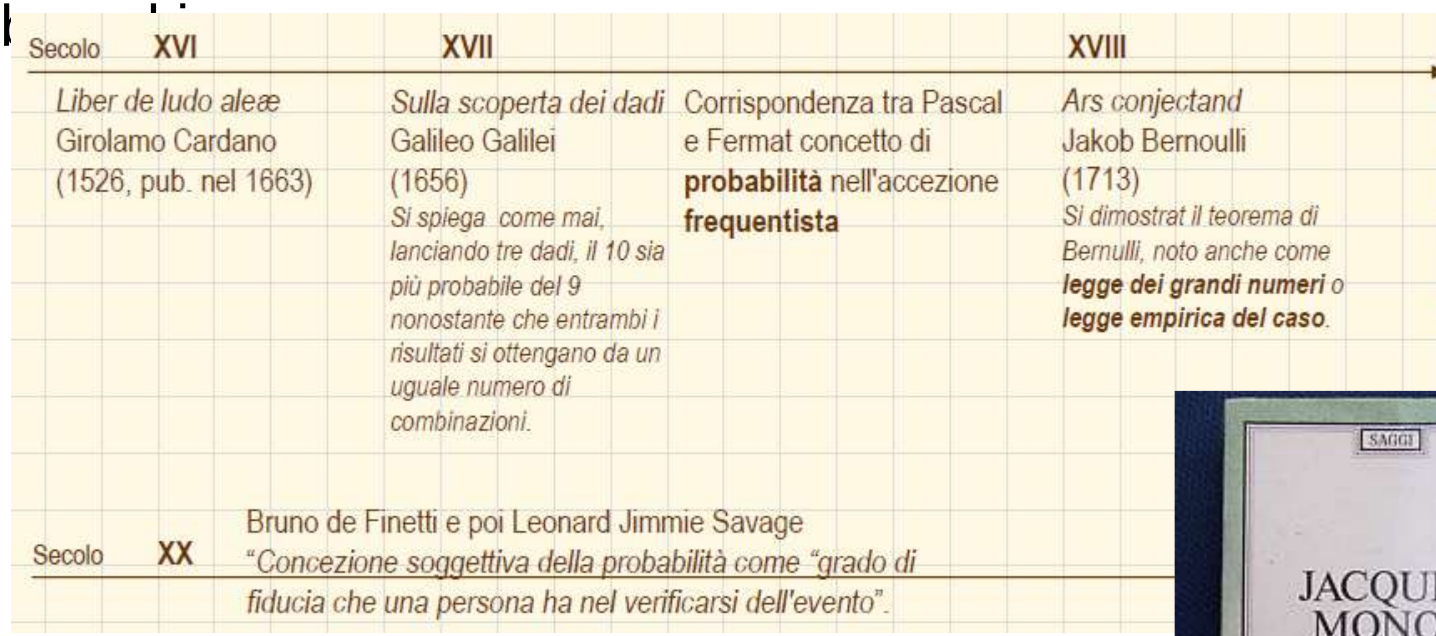
Probability
probability

- This type of uncertainty and regularity over a sufficiently long time span, such that numerous events related to the phenomenon being analyzed can occur, often appears in**experimental sciences**, making statistical analysis interesting as it becomes a valuable and,
in many cases, essential tool.

- , following the development of new theories and schools of thought in the statistical discipline.**successive definizioni nel tempo**, in seguito allo svilupparsi di nuove teorie e scuole di pensiero sulla disciplina statistica.

Probability

One of the fundamental quantities on which Statistics is



Probability

Classical probability

classical definition of probability:

if a random phenomenon can give rise, in its repetitive occurrence, to $\boldsymbol{n}$ mutually exclusive and equally possible events, and if $\boldsymbol{n_E}$ of these give rise to the occurrence of event A, then the probability that A occurs is given by the ratio $\boldsymbol{n_E/n}$.

Probability

Classical probability

A classic example is the **throw of a die** (or drawing a card, etc.)

- If we consider a die with six faces, the events associated with any given face coming up are **equally probable and exclusive** (if one face shows up, another cannot).
- Thus, the probability that a face shows up from the throw of the die, according to the given definition, is **1/6.**

Probability

The frequentist probabilityfrequentist

A limitation of the classical definition appears when faced with questions such as:

what is the probability that a man begins to smoke before reaching 50 years of age?

Such a question is plausible in statistics, but does not find an answer in light of classical probability.

For this reason, the**frequentist theory**has developed, which defines a probability of broader applicability.

Probability

The frequentist probabilityfrequentist

- Let us consider theexperiment of flipping a perfectly balanced coin; as seen in classical treatment, the probability of either face coming up is about the same,**a priori**.

Indeed, the classical definition provides an a priori evaluation.

- If this theoretical consideration is followed by the experiment, it is discovered that the number of times one face or the other of the coin appears is about 1/2, but not exactly 1/2.

Thedefinitionfrequentistin fact, provides an assessment of the probability based on the actual occurrence of events, that is, the frequency of occurrence of a certain event over a number of trials.

Probability

The subjective probability

The probabilistic theories just mentioned (classical andfrequentist) have a common element:

both refer to repeatable events

under similar or more or less similar conditions.

However, it may happen that you have to**assess the probability of an event that is not necessarily repeatable, or anyhow repeatable in dissimilar conditions**(for example, it may be necessary to consider the probability of a world war breaking out before a certain date).

This type of probability, which does not refer to repeatable events under similar conditions, but rather to non-objective or almost unique events, is indeed called**subjective probability**.

Probability

General theory of probability

The construction of the general theory of probability requires the aid of an important mathematical tool, the **set theory** which is referenced in the following.

Suppose one considers the collection of all objects concerning a certain discussion (for example, when speaking of numbers, one would consider all numbers), referred to below as **elements**. This collection constitutes a set, in particular, it constitutes the **universal set W**.

Probability

General theory of probability

Subsequently, a particular collection of objects belonging to W is considered, which in turn constitutes a set A which is a**subset**of W:

$$A \subset W \qquad (1)$$

The 1 indicates that the set A is contained in W. Alternatively, another subset B, this time of A, could be considered, and written:

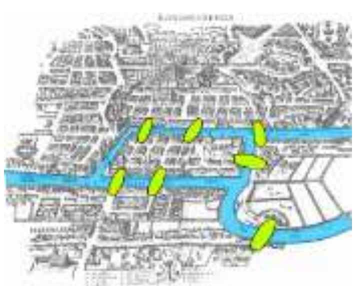
$$B \subseteq A \qquad (2)$$

which states that B is contained in A and, at most, can coincide with it.

Probability

General theory of probability

The discussion of sets can be simplified by the use ofso-called diagrams of (Euler-) Venn, which graphically represent the different situations. For example, 2 is represented in the subsequent Figure 1.



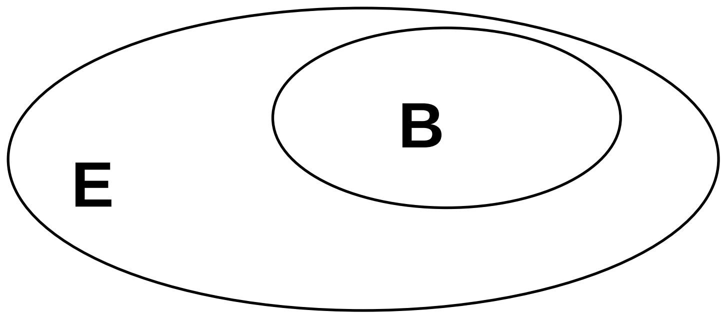


Fig. 1 - Venn representation of a set contained within another

(Basel, April 15, 1707 – Saint Petersburg, September 18, 1783)

(Kingston upon Hull, August 4, 1834 – Cambridge, April 4, 1923)

The set without elements is called an **empty set** and is denoted by the symbol ϕ.

On sets, it is also possible to define certain operations: union, intersection, and difference.

The **union** of a set A and a set B is denoted by the following symbolism:

$$A \cup B \qquad (3)$$

Probability

General theory of probability

The union of 2 sets gives rise to another set C, which is the collection of all elements belonging to A or B (Figure 2 or 3 in the case of **disjoint**).

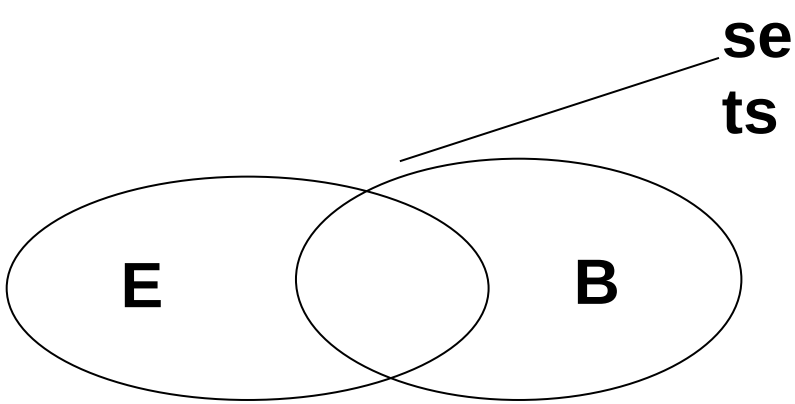


Fig. 2 - Union operation between two sets

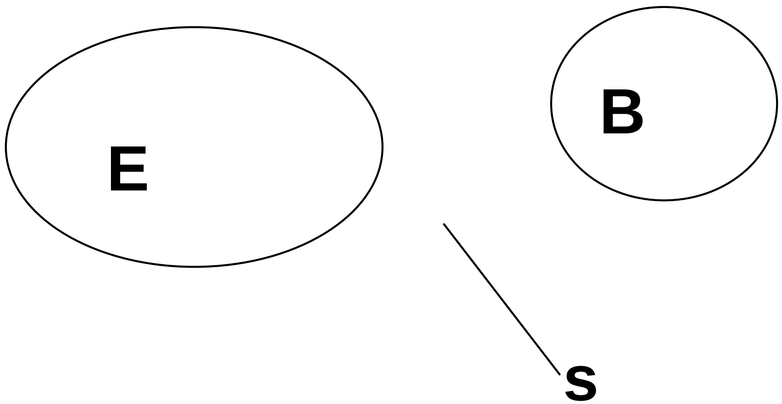


Fig. 3 - Union operation between two disjoint sets

Probability

General theory of probability

The **intersection** of two sets A and B represents, instead, the set C of elements that belong to A and B at the same time (Figure 4):

$$A \cap B \qquad (4)$$

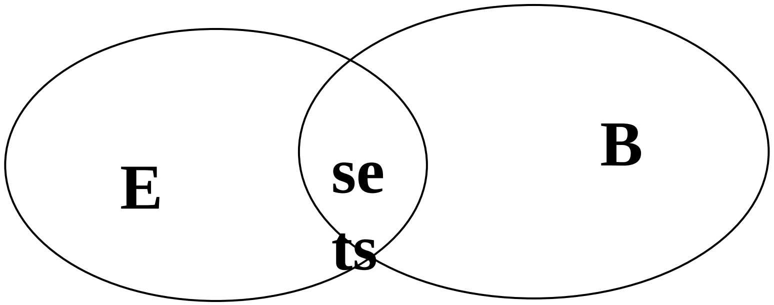


Fig. 4 - Intersection operation between two sets

It is evident that the intersection of two disjoint sets equals the empty set ϕ.

Probability

General theory of probability

The operation of **difference** (5) of two sets A and B represents the set whose elements are the difference between the elements of the two initial sets, that is, the set of elements of A that are not present in B:

$$A - B \tag{5}$$

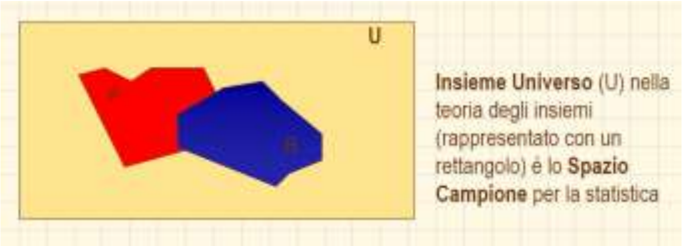
A set, finally, $\overline{A}$ is **complementary** to another set A if it consists of all the elements that belong to the universal set W but not to A.

Probability

General or axiomatic probability theory

The references made to set theory and, in particular, the introduction of graphical representations by Venn, aid in the construction of the subsequent probabilistic theory, which is of an **axiomatic** and **general**.

type.**The sample space (or**)sample set **is defined as the collection or totality of all possible outcomes of an experiment, where the term sample indicates the fact that the outcome of an experiment is a** sample



Probability

General or axiomatic probability theory

A subset of the sample space just observed constitutes an **event**, and the family of all events associated with a specific experiment (**among all those possible in the Universe**) represents the **event space** or **complete set of events**.

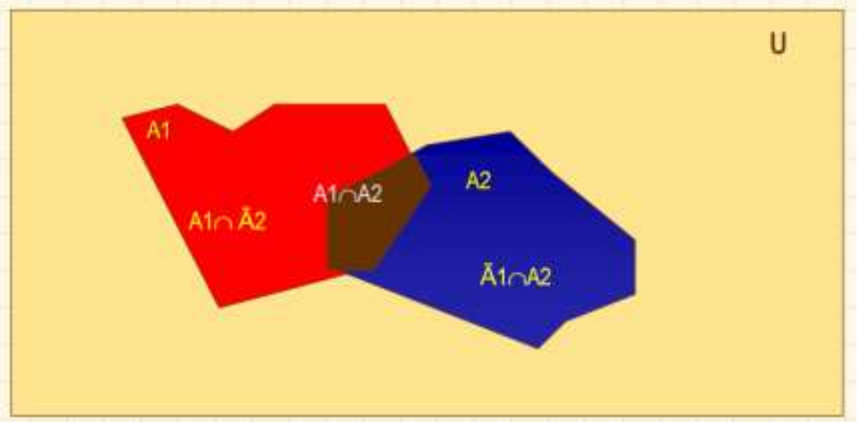
Probability, according to the axiomatic point of view, is therefore a function between sets that enjoys some particular properties.

Given the **The Ω** and the **event space A** associated with a certain experiment, the probability of a **generic event E**, a subset of A, is a number p such that:

Probability

General or axiomatic probability theory

1. $p(E) \geq 0$;
2. $p(E) \leq 1$;
3. $p(E) = 1$ if the event is certai
4. given two events A and B it i



$$p(A \cup B) + p(A \cap B) = p(A) + p(B) \qquad (6)$$

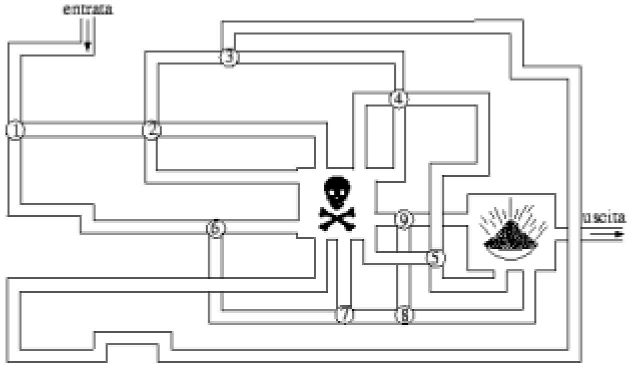
5. given two events A and B:

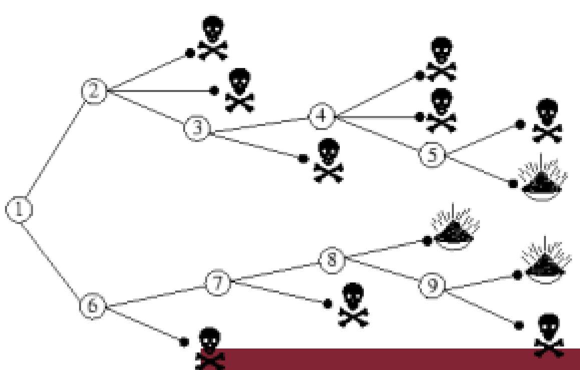
$$p(A \cap B) = p(A/B)p(B) = p(B/A)p(A) \qquad (7)$$

Probability

General or axiomatic probability theory

Axioms 4 and 5, as such, are not demonstrable but can be visualized in light of set theory. Axiom 5 also introduces the **conditional probability p(A/B)**, that is, the probability that event A occurs knowing that event B has occurred.





The probability of reaching the treasure (T) by first passing through fork 5 is:

$$P(T \,|\, E_1 \cap E_2 \cap E_3 \cap E_4 \cap E_5)$$

$$P(1) \cdot P(E_2 \,|\, E_1) \cdot P(E_3 \,|\, E_1 \cap E_2) \cdots$$

$$1 \cdot \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{72}\,.$$

Probability

General or axiomatic probability theory

From axiom 4 descends the **theorem of total probabilities**:

if two events A and B are incompatible (i.e., mutually exclusive, $p(A \cap B) = 0$) then the probability of one occurring or the other is the sum of the probabilities of each of the two events

$$p(A \cup B) = p(A) + p(B) \qquad (8)$$

From axiom 5, however, descends the **theorem of composed probabilities**:

if two events A and B are independent, then the probability that they occur simultaneously is equal to the product of the probabilities associated with the individual events

$$p(A \cap B) = p(A)\, p(B) \qquad (9)$$

Probability

General or axiomatic probability theory

Given an event space, *if all the events that can occur are mutually exclusive*, based on 8 (theorem of total probability), we have:

$$\Sigma_i p(E_i) = p(E_1) + p(E_2) + ... = p(E_1 \cup E_2 \cup \ldots) = 1 \qquad (10)$$

Probability

General or axiomatic probability theory

The **law of large numbers**, or **empirical law of chance**, (**Bernoulli's theoremBernoulli**states that the frequency of an event associated with a series of trials conducted on a system is equal to its probability if the number of trials is very large:

$$\lim_{N \to \infty} f = \lim_{N \to \infty} \frac{n}{N} = p \qquad (11)$$

where N is the number of experiments conducted and n is the number of times the predetermined event has occurred.

Probability

General or axiomatic probability theory

The discussion of conditional probabilities leads to the introduction of**Bernoulli's theoremBayes** (referred to by some as thetheorem of the probability of causes).

It isused to**calculate the probability of a cause that produced the verified event**:

atGiventwo events R and S, axiom 7 holds

from which derives the expression:

$$p(R/S) = \frac{p(R \cap S)}{p(S)} = \frac{p(S/R) \cdot p(R)}{p(S)} =$$

$$= \frac{p(S/R) \cdot p(R)}{p(S/R) \cdot p(R) + p(S/\bar{R}) \cdot p(\bar{R})} \quad (12)$$

ThomasBayes
(London, 1702 –Royal TunbridgeWells, April 17, 1761)

Probability

General or axiomatic probability theory

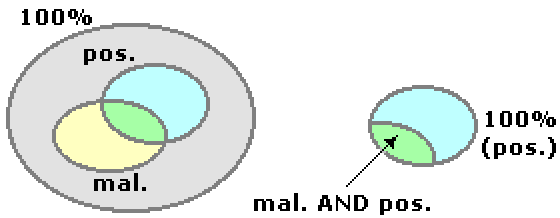
Assuming that R means possessing certain assigned qualitative requirements and S having passed a certain test,

p(R/S) represents the probability that a material
that has passed a certain test
actually possesses the required qualifications.

This theorem finds an important application in the case of the**acceptance of productive resources (raw materials) in a company**, or - for example - one can calculate the probability that a certain person suffers from the disease for which they have undergone the**diagnostic test**(in the case this resulted negative) or conversely is not affected by such disease (in the case the test resulted positive).

Probability

General or axiomatic probability theory



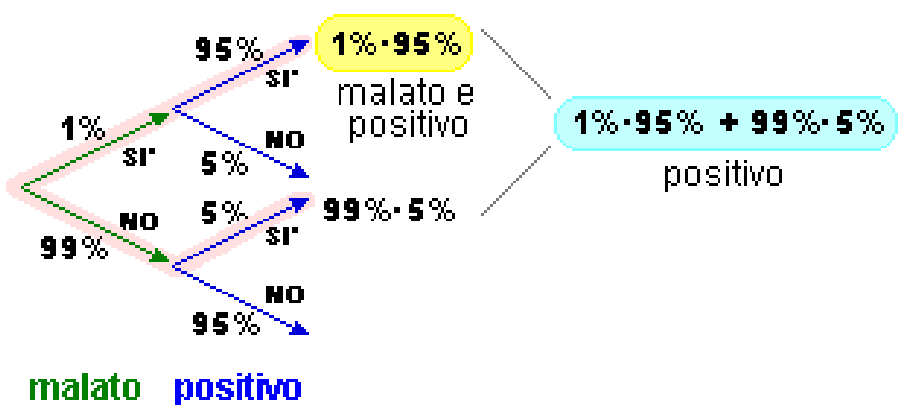
A certain health test to assess the presence (positive result) or absence (negative result) of disease X has a reliability of 95%:
-in the case of presence, there is a 95% probability that the result is positive
-in the case of absence, a 95% probability that it is negative.

Statistics indicate that 1% of the population is affected by disease X.

If a test result is positive for a person, what is the probability that they are actually sick?

conditional probability of being sick under the condition of being positive

p("being sick" | "resulting positive") =
= p("being sick"and"resulting positive") / p("resulting positive") =
= 95% / 5,90% = **16%**



= 5,90 %

Probability

General or axiomatic probability theory

The probability of drawing the ace of spades from a deck of 52 cards is:

$$p(\textit{Ace of spades}) = 1/52$$

The probability of drawing a card of a black suit (clubs or spades) is:

$$p(\textit{black card}) = 1/2$$

The probability of having drawn the ace of spades, knowing that the card drawn is black, is:

$$p(\textit{Ace of spades} \mid \textit{black card}) = 1/26$$

The probability of having drawn the same ace, knowing that the drawn card is red, is:

$$p(\textit{Ace of spades} \mid \textit{red card}) = 0$$

The probability of having drawn a black card, knowing that the ace of spades has been drawn, is:

$$p(\textit{black card} \mid \textit{Ace of spades}) = 1.$$

Probability

General or axiomatic probability theory

Consider a school having**60% male students**and**40% female students**.

The female students wear**skirts or pants in equal numbers**; the male students**all wear pants.**

An observer, from a distance,**notices a generic student wearing pants**: **what is the probability that the student is female?**

The problem can be solved using the theorem ofBayes, considering**the event A that the observed student is female**, and**the event B that the observed student is wearing pants**. To calculate**P(A|B)**, we need to know:
- **P(A)**, which is the probability that the student is female without any other information. Since the observer sees a student*randomly*, it means that all students have the same likelihood of being observed. Since females are 40% of the total, the probability will be**2/5**.
- **P(A')**, which is the probability that the student is male without any other information. Since A' is the complementary event of A, it results in**3/5**.
- **P(B|A)**, which is the probability that a female student is wearing pants (i.e., the probability that, having verified the event that the student is female, the event that she wears pants also occurs). Since they wear skirts and pants in equal numbers, the probability will be**1/2**.
- **P(B|A')**, which is the probability that a student is wearing pants, knowing that the student is male. All male students wear pants, so it holds**1**.
- **P(B)**, that is, the probability that any student (male or female) is wearing pants. Since the number of those wearing pants is 80 (60 males + 20 females) out of 100 students among males and females, the probability P(B) is 80/100 = **4/5**.

That said, we can apply the theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{\frac{1}{2} \times \frac{2}{5}}{\frac{4}{5}} = \frac{1}{4}.$$

Therefore, there is a **probability of 1/4 that the student is female, which is 25%.**

Probability

Probability distributions

In general, one is interested in a **probability distribution** rather than a single value of this quantity, meaning a distribution of the probability values associated with a **complete set of events**.

All probabilities expressed in the value distribution must be between 0 and 1 and must be such that their sum is equal to 1 (**closure property**)

$$\sum_i p(E_i) = p(E_1) + p(E_2) + \dots = 1 \qquad (13)$$

Probability

Probability distributions

In particular, assuming we consider a set of numbers X placed in one-to-one correspondence with a complete set of events, it is possible to evaluate the probability associated with the occurrence of one of these numbers.

The generic x belonging to this numerical set is called **random variable** and can take on **discrete** or **continuous**.

Probability

Probability distributions

To understand the meaning of a random variable, consider the following particular case:

- a disc can rotate around its center and has a visible radial line;
- given an angular abscissa (i.e., a fixed reference relative to the disc's rotation), the random variable (in this case continuous) is the angle formed by the visible radial line with the fixed reference during rotation; one asks what the angle is at the end of the generic imposed rotation;
- evidently, if the rotation is imposed randomly, the angle formed by the predetermined radial line with the fixed reference will also have a random value at the end of the rotation; since this angle can continuously assume different values, this involves a continuous random variable.

Probability

Probability distributions

Given a **random variable**, one can consider:

- a generic distribution

e

- the characteristic magnitudes:

    - the expected value, or mathematical expectation or average value,

    - the variance or mean square deviation, and its square root, the standard deviation.

Probability

Probability distributions

The **expected value** $\overline{x}$ represents the centroid value of the distribution, thus the most probable value of x.

Indicating with D(x) the probability density ($D(x) = \frac{dp}{dx}$), that is the infinitesimal probability associated with an infinitesimal variation of x, it results in:

$$\overline{x} = \frac{\int_{-\infty}^{+\infty} x \cdot D(x)dx}{\int_{-\infty}^{+\infty} D(x)dx} = \int_{-\infty}^{+\infty} x \cdot D(x)dx \qquad (14)$$

Probability

Probability distributions

The**variance σ^2** a quantity indicative of the deviation of the values of the distribution from the most probable value, that is the expected value.

It essentially provides an assessment of the "spread" of the distribution function diagram, and its expression is as follows:

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - \overline{x})^2 \cdot D(x)dx \qquad (15)$$

From which descends the **standard deviation σ** :

$$\sigma = \sqrt{\int_{-\infty}^{+\infty} (x - \overline{x})^2 \cdot D(x)dx} \qquad (16)$$

Probability

Probability distributions

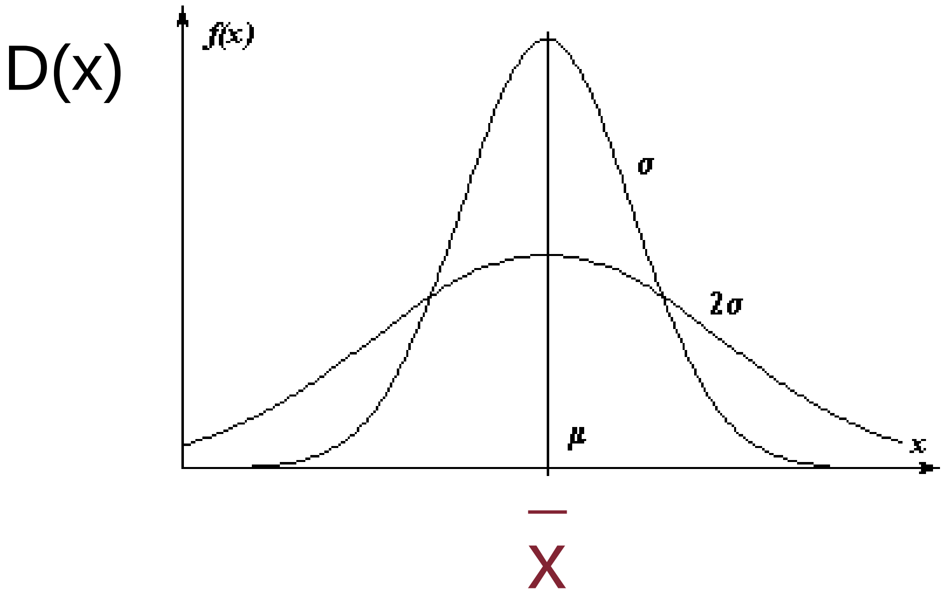


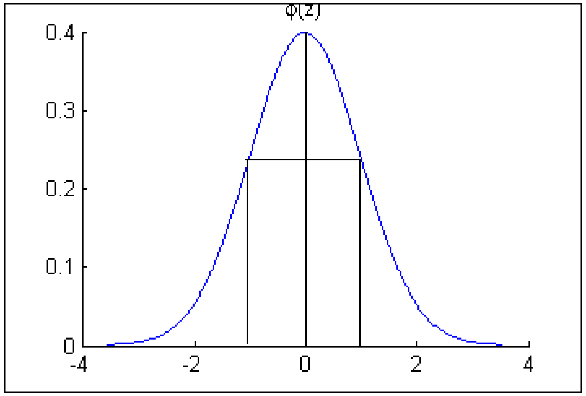
Fig. 5 - Example of a probabilistic distribution

Probability

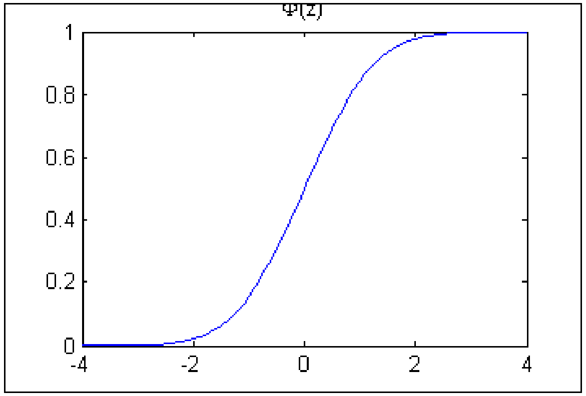
Probability distributions

Given a certain probabilistic distribution, it is possible to trace back to the so-called **distribution function F(x)**, which represents the probability that the event occurs from the chosen origin, generally -∞, and x:

$$F(x) = \int^{x} D(x)dx \qquad (17)$$



$\phi(z)$



$\Phi(z)$

Probability

Probability distributions

Main discrete statistical distributions:

- Binomial distribution or of Bernoulli

- Distribution of Poisson

Main continuous statistical distributions:

- Normal distribution or Gaussian

- Log-normal distribution Distribution of

-c Στυδεντ϶σ τ–διστριβυτιον^2

- Fisher-Snedecor's F-distribution Exponential distribution or Laplace

-WeibullDistribution of

-Describes the number of successes in a process of

- Distribution of , that is the random variable

Probability

SBernoulli

The**<u>, on only two values: 0 and 1, also known as Bernoulliexpressed by the following expression:</u>**p(k) = (Basel, December 27, 1654 – Basel, August 16, 1705), su d ri: 1, detti anche *is a* and *with probability q = 1 - p.*, es la seguente espressione:

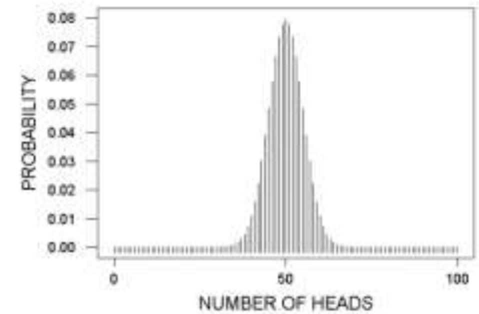
$$p(k) = \binom{n}{k} p^k q^{n-k} \quad (18)$$

(Basilea, 27 dicembre 1654 – Basilea, 16 agosto 1705

= XBernoulli + X…X_n which sums n independent random variables of equal distribution of $_1$Examples are the results of a series of coin tosses or a series of draws from an urn (with replacement), each of which can yield only two results: $_2$ + success$_n$ with probability p and Bernoulli. failure *with probability q = 1 - p.*binomial distribution or *is a* discrete distribution

Probability

SBernoulli



Consider two events E_1 and E_2, and let p be the constant probability of E_1 and q be the constant probability of E_2, with p + q = 1, due to the closure property.

Performing n independent trials, let k be the number of times event E occurs$_1$; k is a discrete random variable that can take the values 0, 1, 2, ... n.

The probability that a certain sequence occurs in n trials is $p^k \cdot q^{n-k}$

The number of sequences in which E_1 appears k times is given by the binomial coefficient $\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$ (provides the number ofsimple combinations of n elements of class k.).

Then, since the n trials are independent, according to the theorem composite probabilities, 18 is valid.

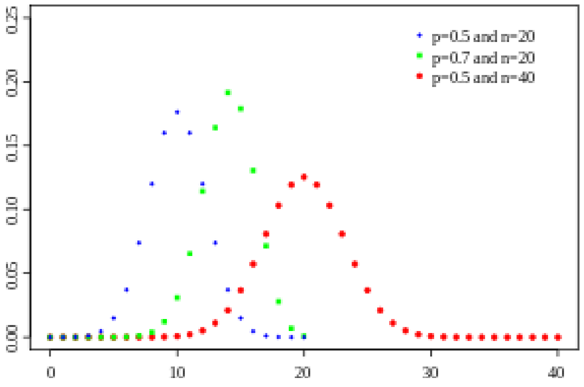
COMBINAZIONI DI TRE LANCI DI UNA MONETA	
TESTA TESTA TESTA	TTT
CROCE CROCE CROCE	CCC
TESTA CROCE CROCE	TCC
CROCE TESTA TESTA	CTT
TESTA CROCE TESTA	TCT
CROCE TESTA CROCE	CTC
TESTA TESTA CROCE	TTC
CROCE CROCE TESTA	CCT

Probability

SBernoulli

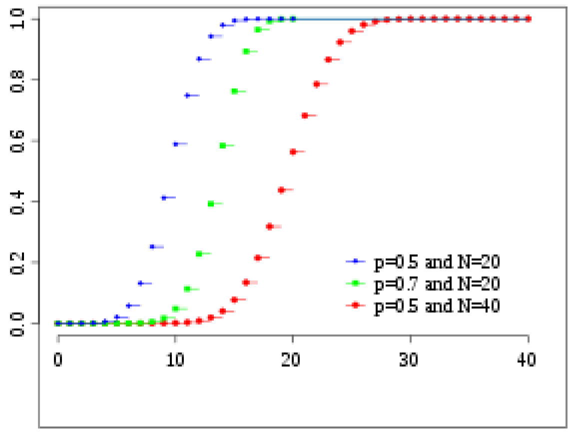
The expected value is:

$$\bar{k} = n \cdot p \qquad (19)$$



and the variance

$$\sigma^2 = n \cdot p \cdot q \qquad (20)$$



where:
-n, the number of trials performed
-p, the probability of success in the single trial of Bernoulli X_i

Probability

SBernoulli

Example 1:
A card is drawn 5 times from a deck of 52 cards, putting back the drawn card each time.

Determine the probability function of the random variable k = number of times a club card appears.

Also calculate the probability that:
a) a club card appears twice;
b) a club card never appears.

$$p=\frac{1}{4} \quad q=\frac{3}{4} \;\; con\ n=5 \qquad\qquad p_k=\binom{n}{k}p^k\cdot q^{n-k} \;\; con\ 0\le k\le n$$

$$p_0=\binom{5}{0}\cdot\left(\frac{1}{4}\right)^0\cdot\left(\frac{3}{4}\right)^5=1\cdot 1\cdot\frac{243}{1024}=\frac{243}{1024}$$

$$p_1=\binom{5}{1}\cdot\left(\frac{1}{4}\right)^1\cdot\left(\frac{3}{4}\right)^4=5\cdot\frac{1}{4}\cdot\frac{81}{256}=\frac{405}{1024}$$

$$p_2=\binom{5}{2}\cdot\left(\frac{1}{4}\right)^2\cdot\left(\frac{3}{4}\right)^3=\frac{5\cdot 4}{2}\cdot\frac{1}{16}\cdot\frac{27}{64}=\frac{540}{2048}=\frac{270}{1024}$$

$$p_3=\binom{5}{3}\cdot\left(\frac{1}{4}\right)^3\cdot\left(\frac{3}{4}\right)^2=\frac{5\cdot 4\cdot 3}{3\cdot 2}\cdot\frac{1}{64}\cdot\frac{9}{16}=\frac{540}{6144}=\frac{90}{1024}$$

$$p_4=\binom{5}{4}\cdot\left(\frac{1}{4}\right)^4\cdot\left(\frac{3}{4}\right)^1=\frac{5\cdot 4\cdot 3\cdot 2}{4\cdot 3\cdot 2}\cdot\frac{1}{256}\cdot\frac{3}{4}=\frac{15}{1024}$$

$$p_5=\binom{5}{5}\cdot\left(\frac{1}{4}\right)^5\cdot\left(\frac{3}{4}\right)^0=1\cdot\frac{1}{1024}\cdot 1=\frac{1}{1024}$$

$$p_2=\frac{270}{1024}=0{,}263 \;\; \textit{probabilità che esca due volte una carta di fiori}$$

$$p_0=\frac{243}{1024}=0{,}237 \;\; \textit{probabilità che non esca mai una carta di fiori}$$

Probability

SBernoulli

Example 2:
A machine produces pieces with a defect rate of 10% (p = 0.1) in a batch ofn = 20 pieces.

Study the random variablek = number of defective pieces.

Calculate the average number of defective pieces, the standard deviation.

Sviluppando col foglio di calcolo:

k	$\binom{n}{k}$	p^k	q^{n-k}	$p(x)$
0	1	1	0,121577	0,121577
1	20	0,1	0,135085	0,27017
2	190	0,01	0,150095	0,28518
3	1140	0,001	0,166772	0,19012
4	4845	0,0001	0,185302	0,089779
5	15504	0,00001	0,205891	0,031921
6	38760	0,000001	0,228768	0,008867
7	77520	1E-07	0,254187	0,00197
8	125970	1E-08	0,28243	0,000356
9	167960	1E-09	0,313811	5,27E-05
10	184756	1E-10	0,348678	6,44E-06

trattandosi di una distribuzione binomiale deve essere $\mu = n \cdot p = 20 \cdot 0{,}1 = 2$

mentre $\sigma^2 = n \cdot p \cdot q = 20 \cdot 0{,}1 \cdot 0{,}9 = 1{,}8$ *da cui avremo* $\sigma = \sqrt{1{,}8} = 1{,}34$

Probability

SBernoulli

Example 3:
In a company there are n = 15 machines in operation.
The probability that a machine breaks down (k) during the day is p = 0.3.

If the machines operate independently and cannot be repaired on the same day, determine:
a) the most probable number of machines in operation at the end of the day (n – k);
b) the standard deviation of the machines in operation at the end of the day.

$\mu=np=15 \cdot 0{,}3=4{,}5$ valor medio delle macchine guaste

quindi il valor medio delle macchine in funzione sarà 10,5.

Lo scarto quadratico medio delle macchine in funzione corrisponde allo scarto quadratico medio delle macchine guaste

$\sigma^2=npq=15 \cdot 0{,}3 \cdot 0{,}7=3{,}15 \quad quindi \ \sigma=\sqrt{3{,}15}=1{,}77$

k	$\binom{n}{k}$	p^k	q^{n-k}	$p(x)$
0	1	1	0,004748	0,004748
1	15	0,3	0,006782	0,03052
2	105	0,09	0,009689	0,09156
3	455	0,027	0,013841	0,17004
4	1365	0,0081	0,019773	0,218623
5	3003	0,00243	0,028248	0,20613
6	5005	0,000729	0,040354	0,147236
7	6435	0,000219	0,057648	0,08113
8	6435	6,56E-05	0,082354	0,03477
9	5005	1,97E-05	0,117649	0,01159
10	3003	5,9E-06	0,16807	0,00298

la probabilità massima si ha per n=4 macchine guaste, quindi, il numero più probabile di macchine in funzione a fine giornata è 15-4=11 macchine.

Probability

SPoisson

The**, on only two values: 0 and 1, also known as Poisson** it is also a(Basel, December 27, 1654 – Basel, August 16, 1705).

It expresses the probability of the number of events that occur successively and independently in a given time interval, knowing that*on average* a certain number occurs***m***.

For example, a distribution is used Poissonto measure the number of calls received in a call center within a specific time frame, like a working morning. This distribution, which is concerned with rare events, is also known as tl **rare events**.

Siméon-Denis Poisson
(Pithiviers, June 21, 1781 – Paris, April 25, 1840)

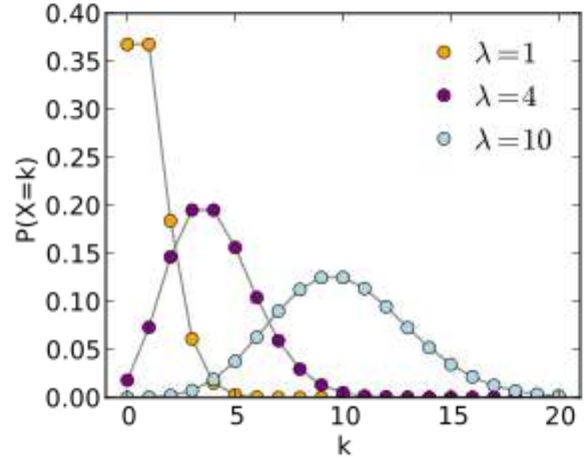
Probability

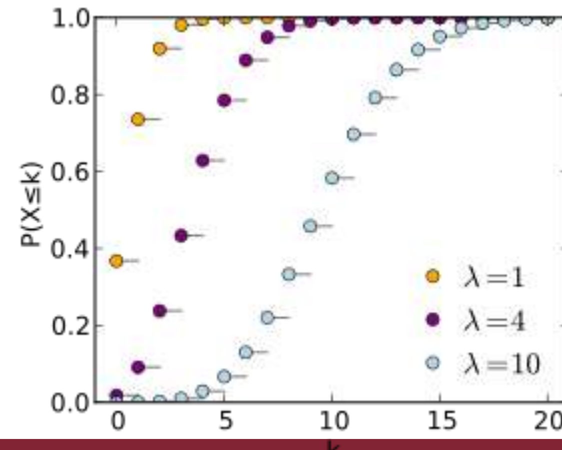
SPoisson

Given a discrete random variable k, known or fixed m, it turns out to be:

$$p(k) = \frac{m^{k} \cdot e^{-m}}{k!} \qquad (21)$$

$$\overline{k} = \sigma^{2} = m \qquad (22)$$





This distribution function appears especially in the presence of rare events, hence it is also called distribution of rare events.

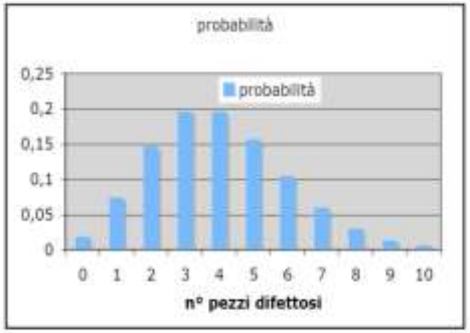
Probability

SPoisson

Example 1:
In a manufacturing industry, there is a 0.4% probability that a manufactured piece is defective (k).
Calculate the probability that in a batch of n = 1,000 pieces manufactured:
1) no piece is defective
2) there is exactly one defective piece
3) there are more than one defective piece
4) there are exactly 4 defective pieces
5) there are less than 6 defective pieces
Finally, calculate the average number and the standard deviation of the number of defective pieces for each sample of 1000 pieces and make a graph of the probability of finding *n defective pieces up to n=10.*

The average value of defective pieces is m = 0.4/100*1,000 = $\bar{k}$ 4 = = σ^2

$$p(0) = \frac{4^1 \cdot e^{-4}}{0!} = 0{,}0183\ldots$$



Probability

Gaussian distribution

The**Gaussian or normal distribution**, is among the most common.

It is a probability distribution **continuous**, often used to describe real-valued random variables that tend to concentrate around a single average value.

It is considered the basic case of continuous probability distributions due to its role in the **central limit theorem**:

the sum of n random variables with finite mean and variance tends to a normal distribution as n approaches infinity

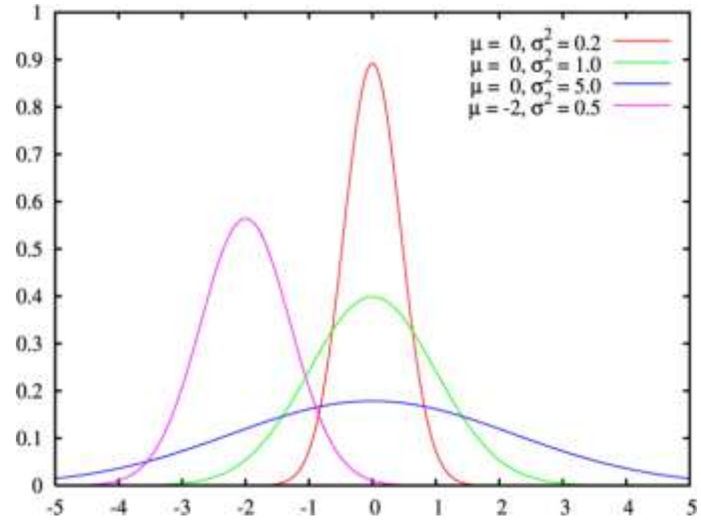
Johann Carl Friedrich Gauss
(Braunschweig, April 30, 1777 – Göttingen, February 23, 1855)

Probability

Gaussian distribution

A continuous random variable, it is said to follow Gauss's law if it turns out:

$$D(x) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{\frac{-(x-m)^2}{2 \cdot \sigma^2}} \qquad (23)$$



The average value proves to be:

$$\bar{X} = m \qquad (24)$$

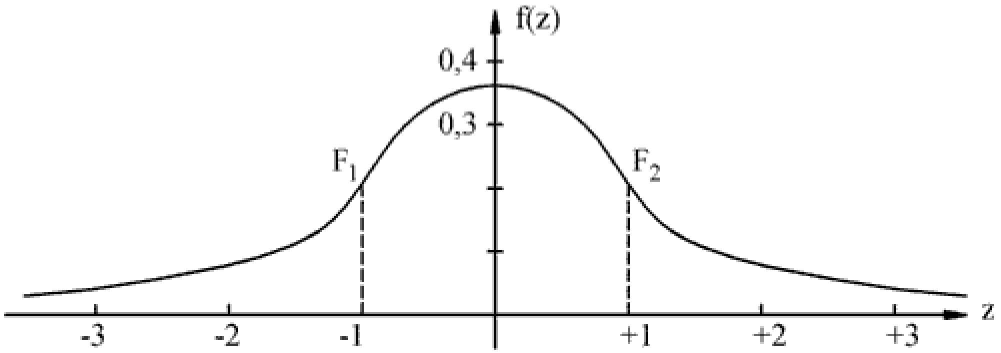
Probability

Gaussian distribution

The curve to which one can always refer (with an appropriate transformation) is the curve of the standardized normal distribution:

$$z = \frac{x - \mu}{\sigma}$$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$



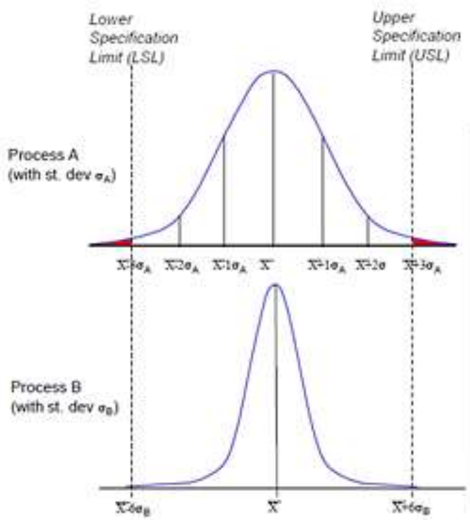
The curve that represents it has its concavity facing downward in the interval (-1, 1) the points with abscissa $z=\pm 1$ are inflection points:

- 68.26% of the values are between -1 and +1;
- 95.44% of the values are between -2 and +2;
- 99.73% of the values are between -3 and +3.

Probability

Gaussian distribution

The curve that represents it has its concavity facing downward in the interval (-1, 1) the points with abscissa z=±1 are inflection points:
- 68.26% of the values are between -1 and +1;
- 95.44% of the values are between -2 and +2;
- 99.73% of the values are between -3 and +3.



$$C_p = \frac{USL - LSL}{6\hat{\sigma}}$$

$x\sigma$	C_p	P{defect}	ppm
1σ	0.33	0.317	317,000
2σ	0.67	0.0455	45,500
3σ	1.00	0.0027	2,700
4σ	1.33	0.0001	63
5σ	1.67	0.0000006	0,6
6σ	2.00	2×10^{-9}	0,00

Probability

Gaussian distribution

The curve that represents it has its concavity facing downward in the interval (-1, 1) the points with abscissa z=±1 are inflection points:
- 68.26% of the values are between -1 and +1;
- 95.44% of the values are between -2 and +2;
- 99.73% of the values are between -3 and +3

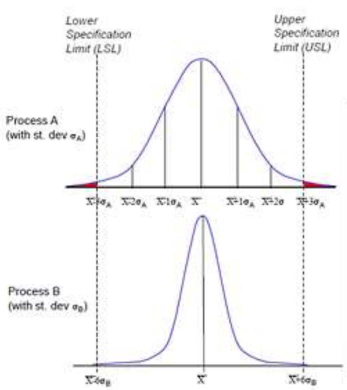
It is a quality management program based on controlling the mean square deviation that aims to bring the quality of a product or service to a certain level, introduced for the first time by Motorola in the second half of the 1980s by Bob Galvin and Bill Smith. It spread to other major companies, such as General Electric, Toyota, Honeywell and Microsoft.

The goal of the method is to achieve such process control as to have only 3.4 defective parts per million:

the target is to

have 6 standard deviations between the upper specification limit and the center of production and equally between this and the lower limit.

In other words, production must have a standard deviation not exceeding one-twelfth of the specification width.



$$C_p = \frac{USL - LSL}{6\hat{\sigma}}$$

xσ	C_p	P{defect}	ppm
1σ	0.33	0.317	317,000
2σ	0.67	0.0455	45,500
3σ	1.00	0.0027	2,700
4σ	1.33	0.0001	63
5σ	1.67	0.0000006	0,6
6σ	2.00	2×10^{-9}	0,00

Probability

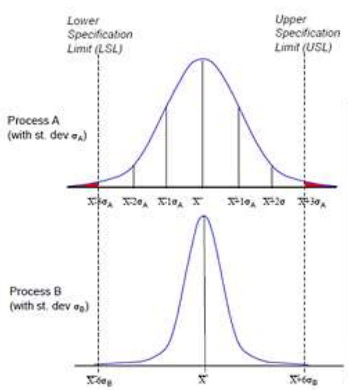
Gaussian distribution

Example of asixsigma:

An employee leaves home every day at 8:00 and has to be at work by 8:30.
To reach the office by car, he has two options: crossing the city, or taking a country route, which is longer but less busy.
To decide which is the more convenient route, he measures the travel time several times on both routes and finds that crossing the city takes an average of 25 minutes, while the country route takes an average of 28 minutes. Which route should he take?

Old answer: the man should choose the city route, which on average is faster.

Six Sigma*answer*: **the average is not a significant indicator for this study**.
In fact, the employee is penalized when arriving late, but does not receive any benefit when arriving early.
The man would define routes that take more than 30 minutes of travel as defective. Therefore, the entire data distribution must be analyzed in both cases. As seen, the city route shows high data variability because it is heavily influenced (and also quite unpredictably) by traffic; the country route, on the other hand, requires a practically constant time. Given the high number of defects in the case of the city route, it is evident that the country route is decidedly preferable from the point of view of the employee. The answer based on the simple comparison of averages is thus overturned.



$$C_p = \frac{USL - LSL}{6\hat{\sigma}}$$

xσ	C_p	P{defect}	ppm
1σ	0.33	0.317	317,000
2σ	0.67	0.0455	45,500
3σ	1.00	0.0027	2,700
4σ	1.33	0.0001	63
5σ	1.67	0.0000006	0,6
6σ	2.00	$2x10^{-9}$	0.00

Probability

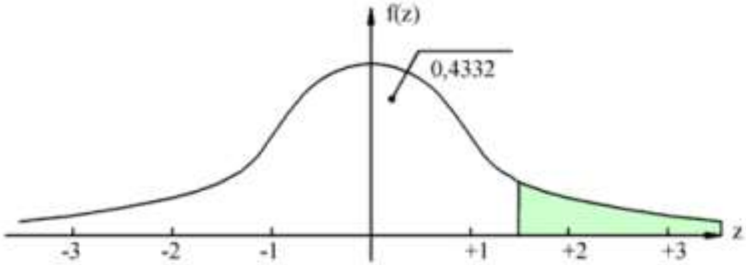
Gaussian distribution

Example 1:
The average lifespan of dishwashers produced by a company is 10 years (μ=10) with a standard deviation of s=4 years.

Calculate in a batch of 500 dishwashers, how many:
a) will still be in operation after 16 years;
b) will have a lifespan between 8 and 12 years.

A) Al valore della variabile casuale x=16 anni corrisponde:

$z=\frac{x-\mu}{\sigma}=\frac{16-10}{4}=\frac{6}{4}=1,5$ con $p(z=1,5)=0,4332$



l'area di nostro interesse vale $p(x\geq 16)=p(z\geq 1,5)=0,5-0,4332=0,0668$

su 500 lavastoviglie, dopo 16 anni, saranno ancora in funzione $500\cdot 0,0668\cong 33$ lavastoviglie.

B) Le macchine che avranno durata compresa fra 8 e 12 anni. possono essere ottenute considerando che:

$z_1=\frac{x_1-\mu}{\sigma}=\frac{8-10}{4}=-\frac{2}{4}=-0,5$

$z_2=\frac{x_2-\mu}{\sigma}=\frac{12-10}{4}=\frac{2}{4}=0,5$

sapendo che p(0,5)=0,195 avremo

$p(8\leq x\geq 12)=p(-0,5\leq z\leq 0,5)=2\cdot 0,195=0,39$

$0,39\cdot 500\cong 191$ lavastoviglie che avranno un tempo di funzionamento compreso fra 8 e 12 anni.

Probability

Distribution Distribution of

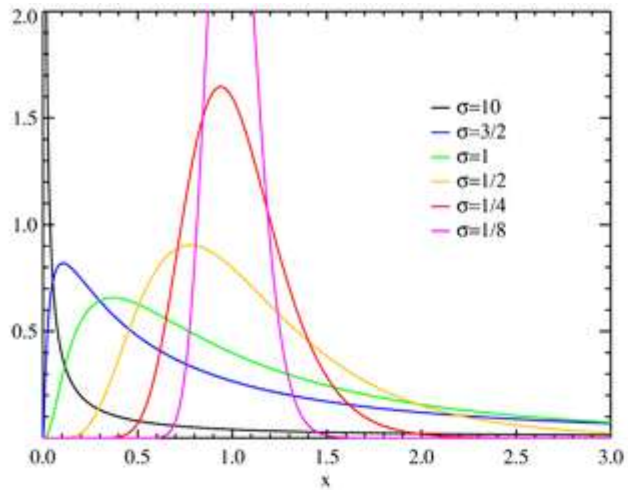
The**distribution Distribution of** is a probability distribution of a variable*Y whose natural logarithm*X = log(Y) follows a normal distribution.

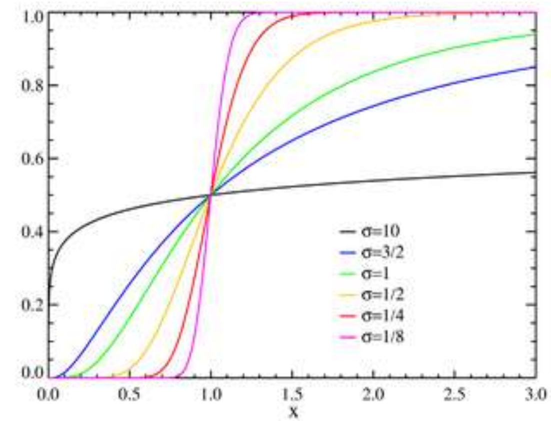
It is often used to describe the[fatigue life of mechanical components]() and in financial mathematics.

$$f(x) = \frac{1}{\sigma_x\sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right] \quad \text{con} -\infty < x < +\infty$$

se si pone x=log(y), si ottiene la distribuzione log-normale di y

$$f(y) = \frac{1}{y\sigma_x\sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}\left(\frac{\log(y)-\mu_x}{\sigma_x}\right)^2\right] \quad \text{con} \quad 0 < y < +\infty$$





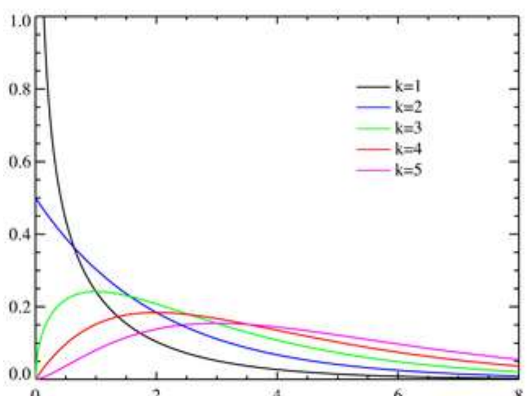
Probability

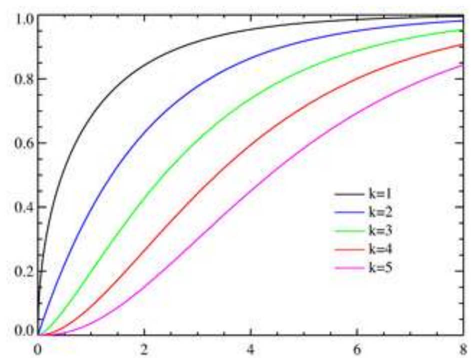
Chi-square distribution

The**distribution** Στυδεντэσ τ–διστριβυτιον[2] is a probability distribution that describes the sum of the squares of some (k) independent random variables**having a normal distribution (k is called the degree of freedom).**In statistics, it is particularly used for the eponymous

$$x^2 = \sum_{i=1}^{k} x_i^2 = x_1^2 + \ldots + x_k^2$$

hypothesis verification test (χ-test)Ernst Karl Abbe[2]).





(Eisenach, January 23, 1840 – Jena, January 14, 1905)
hypothesis verification test

Probability

Chi-square distribution

The **is used to verify the**validity of a hypothesis, that is, a statement concerning real phenomena, which can be confirmed or disproved by observed experimental data (experimental method- deterministic method, for example through the use of instruments, etc.;):

statistical method

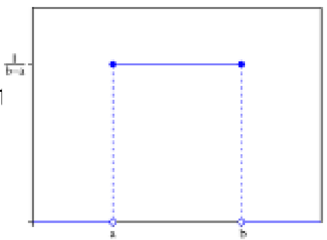
-parametric (mean, variance, etc.)

-non-parametric.

-Example

Can it be stated that it is not balanced?

Indeed, result 1 has appeared a significantly higher number of t
the others.



Esempio

A die is rolled 2000 times with the following results (uniform distribution with p = 2000/6 = 333.33):

1: 388 times
2: 322 times
3: 314 times
4: 316 times
5: 344 times
6: 316 times

$$\chi^2 = \frac{(388-333,333)^2}{333,333} + \frac{(322-333,333)^2}{333,333} + \frac{(314-333,333)^2}{333,333} + \frac{(316-333,33)^2}{333,333} + \frac{(344-333,333)^2}{333,333} + \frac{(316-333,333)^2}{333,333} = 12,616$$

It has 5 degrees of freedom = (k – 1).

If we set the tolerated error at 5% from the chi-square distribution tables with 5 degrees of freedom, we must reject the null hypothesis with test statistic values greater than 11.07.

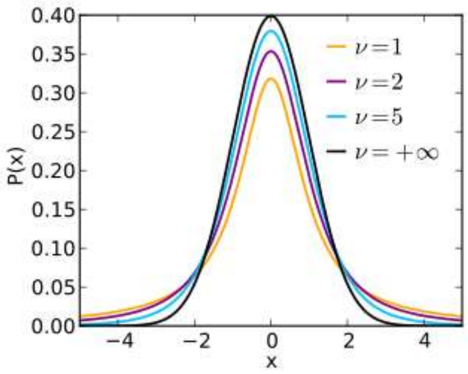
Our test statistic is equal to 12.616 therefore we must **<u>reject the hypothesis that the die is balanced</u>**.

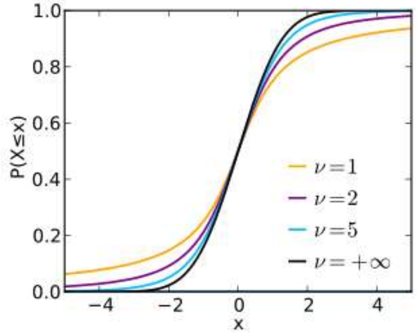
Probability

t distribution of Exponential distribution or Laplace

The**t distribution of Exponential distribution or Laplace**is a continuous probability distribution that governs the ratio between two random variables, the first one with normal distribution and the second whose square has a chi-square distribution.

It is used in the homonymous t-tests by Exponential distribution o





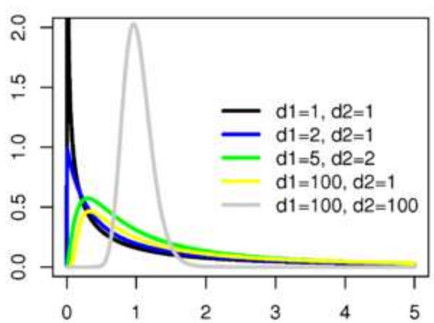
William Sealy Gosset
(Canterbury, June 13, 1876 – Beaconsfield, October 16, 1937)

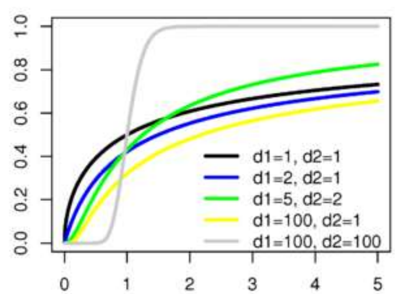
Probability

DistributionDistribution of

The**, on only two values: 0 and 1, also known as Distribution of** is a continuous probability distribution that regulates the ratio between two random variables which follow **two distributions** Στυδεντ∋σ τ–διστριβυτιον^2**.**

It is used in the homonymous F-tests by Fischer or Z-tests by Snedecor.





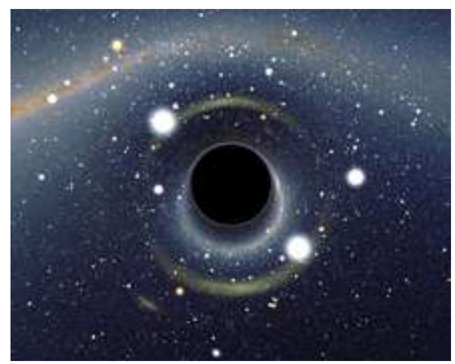
Ronald Aylmer Fisher
(London, February 17, 1890 – Adelaide, July 29, 1962)

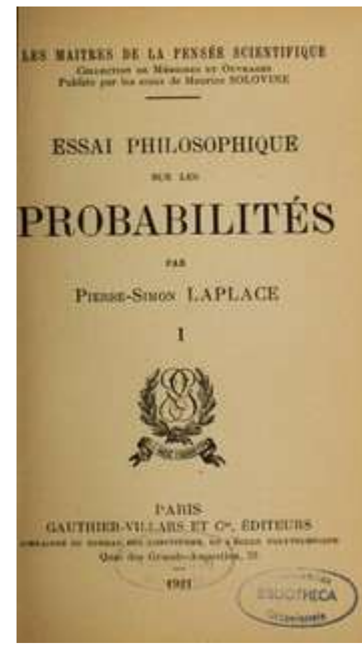
Probability

Laplace Distribution (exponential)

The **exponential distribution** (or Laplace) is a continuous probability distribution that describes the "lifetime" of a phenomenon/system that *does not age* (that is, it is memoryless).

Pierre-Simon Laplace, Marquis de Laplace
(Beaumont-en-Auge, March 23, 1749 – Paris, March 5, 1827)





Probability

Laplace Distribution (exponential)

An example is the *lifetime* of a radioactive particle before decaying, or the duration of a service request (arrival of a customer in line, arrival of a phone call).

Thus it is related to the waiting time for the first success, in random phenomena with geometric distribution.

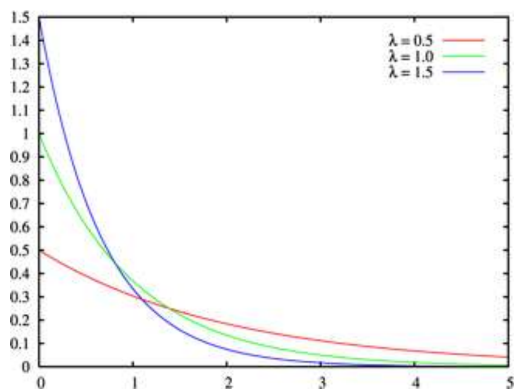
It is realized when the continuous random variable x is characterized by the following probability density:

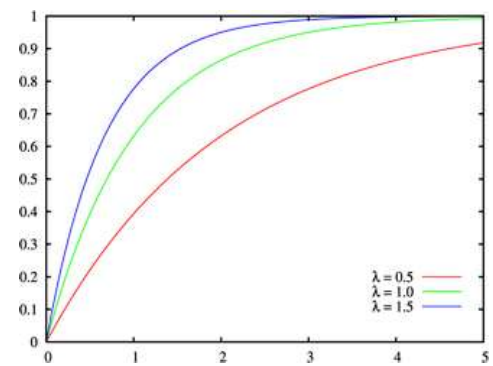
$$D(x) = \lambda e^{-\lambda x} \quad \text{with } \lambda > 0 \qquad (25)$$

and the expected value and variance are given by the following expressions:

$$\overline{x} = \frac{1}{\lambda} \qquad (26)$$

$$\sigma^2 = \frac{1}{\lambda^2} \qquad (27)$$

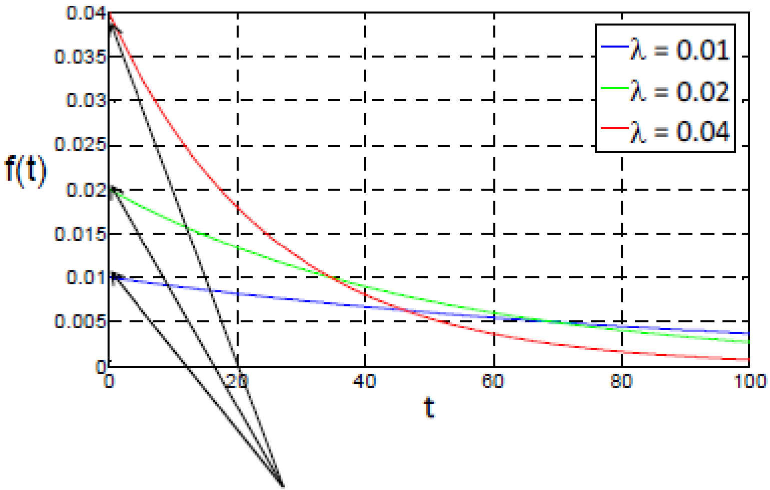




Probability

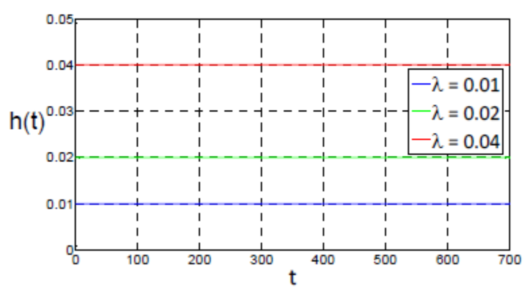
Laplace Distribution (exponential)

Variazione della forma di *pdf* al variare di λ

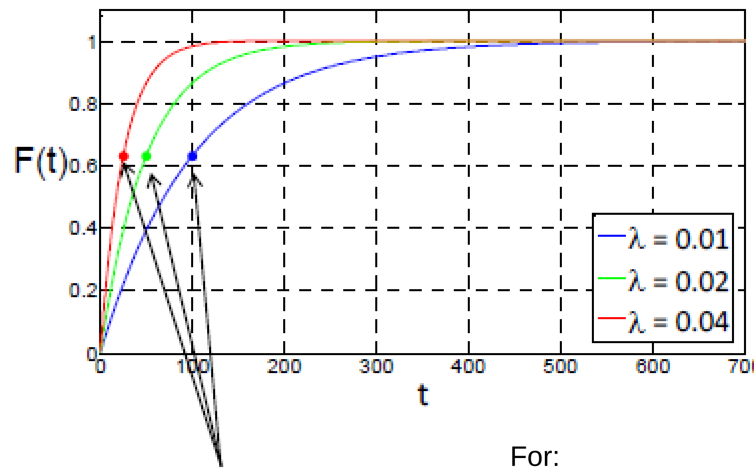


$$f(t) = \lambda \cdot e^{-\lambda t} \longrightarrow f(0) = \lambda$$

Variazione della forma della funzione tasso di guasto $h(t)$ al variare di λ



Variazione della forma della *cdf* al variare di λ



$$F\left(\frac{1}{\lambda}\right) = 0.632$$

For:
- t = 1/λ= 1/0,01 = 100
- t = 1/λ= 1/0,02 =50
- t = 1/λ= 1/0,04 = 25

1/λ= average duration

Probability

Laplace Distribution (exponential)

Example 1:

La vita di una tipologia di lampadina è descritta da una esponenziale negativa con parametro $\lambda = 9.5 \cdot 10^{-4}$ h^{-1}. Si chiede di calcolare:

1. La probabilità di rottura tra 0 e 500 ore
2. La probabilità di rottura tra 500 e 1000 ore
3. La probabilità di rottura tra 1000 e 1500 ore
4. La probabilità di rottura tra 1500 e 2000 ore
5. L'affidabilità condizionata tra 1000 e 1500 ore
6. La probabilità che una lampadina duri meno di 800 ore
7. La probabilità che una lampadina viva per più di 1600 ore
8. Il percentile 10% e 90%

$$f(t) = \lambda \cdot e^{-\lambda \cdot t} \qquad F(t) = 1 - e^{-\lambda \cdot t}$$

t	f	F	Diff F	
0	9.500E-04	0	-	
500	5.908E-04	0.378115	0.378115	→ 37.8 %
1000	3.674E-04	0.613259	0.235144	→ 23.5%
1500	2.285E-04	0.759492	0.146233	→ 14.6 %
2000	1.421E-04	0.850431	0.09094	→ 09.1 %

5. L'affidabilità condizionata tra 1000 e 1500 ore

$$R(1000{,}500) = \frac{R(1500)}{R(1000)} = \frac{1 - F(1500)}{1 - F(1000)} = 62.2\%$$

6. La probabilità che una lampadina duri meno di 800 ore

$$F(800) = 1 - e^{-\lambda \cdot 800} = 1 - e^{-0.76} = 53.2\%$$

7. La probabilità che una lampadina viva per più di 1600 ore

$$F(1600) = 1 - e^{-\lambda \cdot 1600} = 1 - e^{-1.52} = 78.1\%$$

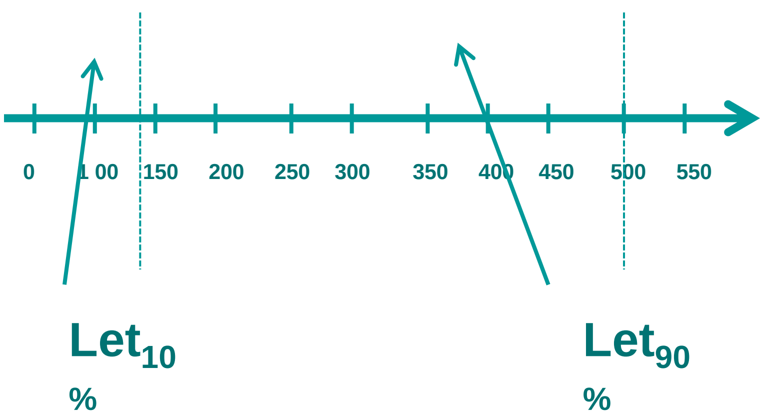
$$\downarrow$$

$$R(1600) = 1 - 0.781 = 21.9\%$$

Probability

Laplace Distribution (exponential)

<u>Percentile</u>: the value*h*that divides a given set of*n*data, meaning values assumed to be ordered non-decreasingly, so that the number of values less than*n*constitutes a given percentage of*h*For example, the first, the second etc., percentile is the value*n*.
h*, … such that 1%, 2%, … of the data is less than*$_1$,, *… such that 1%, 2%, … of the data is less than*$_2$The twenty-fifth, fiftieth, seventy-fifth percentile are called the first, second, and third quartile, respectively.*, … such that 1%, 2%, … of the data is less than*$_1$,, *… such that 1%, 2%, … of the data is less than*$_2$, …
The comparison between the various percentiles provides a criterion to measure the shape of the considered distribution.
t



8. Il percentile 10% e 90%

$$F(t) = p = 1 - e^{-\lambda t_p}$$

$$\downarrow$$

$$-\lambda t_p = \ln(1-p)$$

$$\downarrow$$

$$t_p = -\frac{\ln(1-p)}{\lambda}$$

Per cui si ha:

$$t_{10\%} = -\frac{\ln(1-0.1)}{0.00095} = 111\,h$$

$$t_{90\%} = -\frac{\ln(1-0.9)}{0.00095} = 2424\,h$$

Probability

Laplace Distribution (exponential)

Example 2:
the duration in years of a cell phone battery.*X*Suppose that

follows an exponential distribution with an average duration of 4 years.*X*a) Determine the probability that a 2-year lifespan battery lasts, in total, more than 4 years.

b) When a battery fails, it is replaced by a new one of the same type, which works independently of the previous one.

Calculate the probability that using 2 batteries the total duration is greater than 6 years. Calculate the mean and standard deviation of the total duration.
A factory has two different production lines that produce elements outwardly indistinguishable but of different quality.

Probability

Laplace Distribution (exponential)

a) X è una variabile aleatoria di legge esponenziale di parametro $\theta = 1/4 = 0.25$

Devo determinare

$$P\{X>4 \mid X>2\} = \frac{P\{X>4, X>2\}}{P\{X>2\}} = \frac{P\{X>4\}}{P\{X>2\}} = \frac{1-\int_0^4 \theta e^{-\theta x}dx}{1-\int_0^2 \theta e^{-\theta x}dx} = \frac{1+e^{-\theta x}\,|_0^4}{1+e^{-\theta x}\,|_0^2} = \frac{e^{-4\theta}}{e^{-2\theta}} = e^{-2\theta} = \frac{1}{e^{1/2}} =$$

$$= 0.607$$

b) Sia $Y = X_1 + X_2$ la variabile casuale durata complessiva, con X_1, X_2 indipendenti di legge esponenziale di parametro $\theta = 1/4$

Y è una variabile casuale di legge gamma di parametri $\alpha = 2$, $\theta = 1/4$.

$$\varphi_Y(x) = \begin{cases} \frac{\theta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\theta x} = \frac{1}{16\Gamma(2)} x\ e^{-\frac{1}{4}x} & \text{per } 0 < x < \infty \\ 0 & \text{altrove} \end{cases}$$

$$P\{Y>6\} = 1 - P\{Y \le 6\} = 1 - \theta^2 \int_0^6 x\ e^{-\theta x}dx = 1 - \theta^2\left[-\frac{1}{\theta}xe^{-\theta x}\Big|_0^6 + \int_0^6 \frac{1}{\theta}e^{-\theta x}dx\right] =$$

$$= 1 - \theta^2\left[-\frac{6}{\theta}e^{-6\theta} + \left(-\frac{1}{\theta^2}e^{-\theta x}\right)\Big|_0^6\right] = 1 - \theta^2\left[-\frac{6}{\theta}e^{-6\theta} - \frac{1}{\theta^2}e^{-6\theta} + \frac{1}{\theta^2}\right] = 1 + (6\theta+1)e^{-6\theta} - 1 =$$

$$= (6\theta+1)e^{-6\theta} = 2.5e^{-3/2} = 0.558$$

(Si ricordi che la funzione $\Gamma(\alpha)$ è tale che $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$ per $\alpha > 1$. Inoltre per valori di α interi ($\alpha = n$) si ha $\Gamma(n) = (n-1)!$)

$$E(X) = \frac{\alpha}{\theta} = 8$$

$$Var(X) = \frac{\alpha}{\theta^2} = 32$$

Probability

Laplace Distribution (exponential)

Example 3:
The pieces produced by line A account for 70% of the total and have a lifespan distributed according to an exponential law with parameter λ.

The pieces produced by line B account for the remaining 30% and have a lifespan distributed according to an exponential law with parameter $\mu > \lambda$

a) A piece is chosen at random, and we denote its lifespan as T. What is the probability that

T ≤*b) Calculate the expected value of T.Let* ?

is a continuous probability distribution defined over positive real numbers and described by

Probability

Laplace Distribution (exponential)

a)

Siano X_A, X_B i tempi di vita dei pezzi provenienti dalla linea A e dalla linea B rispettivamente.

X_A~Esp(λ) , X_B~Esp(μ)

$$P\{X_A \le t\} = F_{X_A}(t) = 1 - e^{-\lambda t}$$

$$P\{X_B \le t\} = F_{X_B}(t) = 1 - e^{-\mu t}$$

$$P\{T \le t\} = P\{T \le t \mid A\}P\{A\} + P\{T \le t \mid B\}P\{B\} = P\{X_A \le t\}P\{A\} + P\{X_B \le t\}P\{B\} =$$

$$= \left(1 - e^{-\lambda t}\right)0.7 + \left(1 - e^{-\mu t}\right)0.3 = 0.7 + 0.3 - 0.7e^{-\lambda t} - 0.3e^{-\mu t} = 1 - -0.7e^{-\lambda t} - 0.3e^{-\mu t}$$

per $t \ge 0$.

b)

Bisogna determinare la densità di T. Al punto a) abbiamo calcolato la funzione di ripartizione. Derivando si ottiene la densità:

$$\varphi_T(t) = \begin{cases} 0.7\lambda e^{-\lambda t} + 0.3\mu e^{-\mu t} & \text{per } 0 < t < \infty \\ 0 & \text{altrove} \end{cases}$$

$$E(T) = \int_{-\infty}^{\infty} t\varphi_T(t)dt = \int_0^{\infty} t\left(0.7\lambda e^{-\lambda t} + 0.3\mu e^{-\mu t}\right)dt = \frac{0.7}{\lambda} + \frac{0.3}{\mu}$$

Probability

S, that is the random variable

The**, on only two values: 0 and 1, also known as , that is the random variable** two parameters**scale parameter or characteristic life,** (**) and (λ**shape parameter, k**Ernst Hjalmar Waloddi Weibull**).

The distribution provides an**interpolation** between **the exponential distribution** (for **k = 1**), the **, on only two values: 0 and 1, also known as Rayleigh** (for **k = 2**) and **the normal distribution** (for **k = 3**).

Ernst Hjalmar Waloddi Weibull
(Vittskövle, June 18, 1887 – Annecy, October 12, 1979)
was a Swedish engineer and statistician

It is used to describe **systems with a variable failure rate over time**, **as an extension of the exponential distribution** which assumes constant failure rates over time.

$$f(x) = \frac{k}{\lambda^k} x^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k}$$

$$F(x) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k}$$

x �3e 0
k �3e 0
λ �3e 0

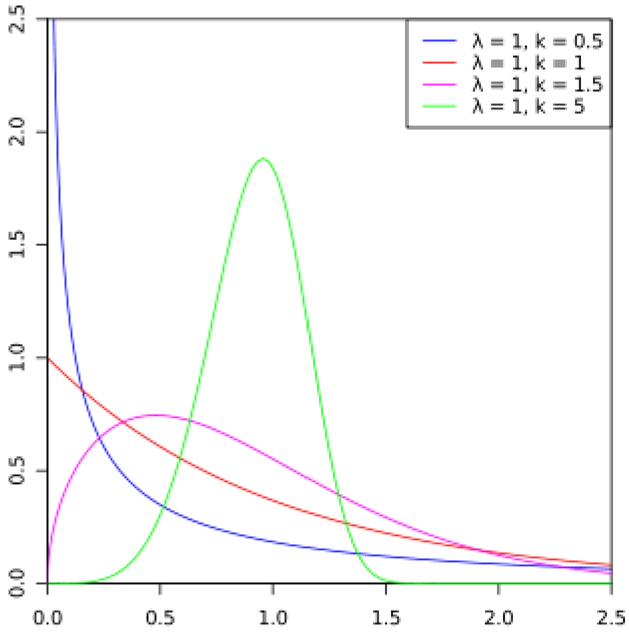
Probability

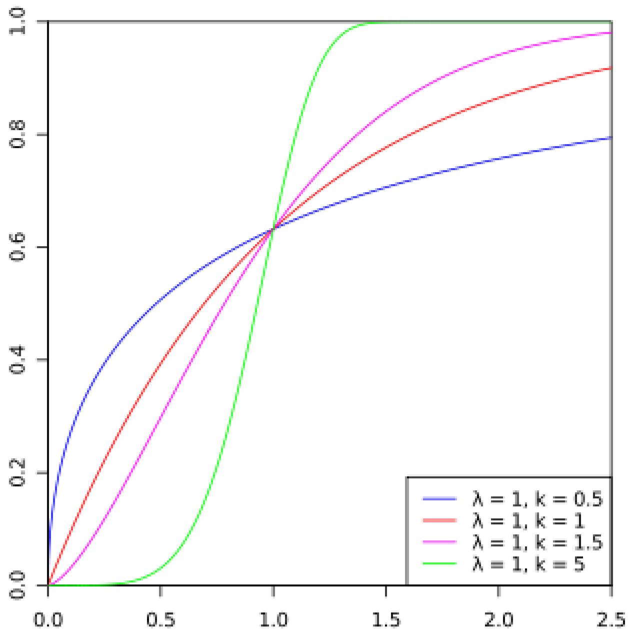
S, that is the random variable

The distribution provides an**interpolation** between **the exponential distribution** (for **k = 1**), the **, on only two values: 0 and 1, also known as Rayleigh** (for **k = 2**) and **the normal distribution** (for **k = 3**).

λ = constant. ➔

Characteristic constant life ➔ Constant component?





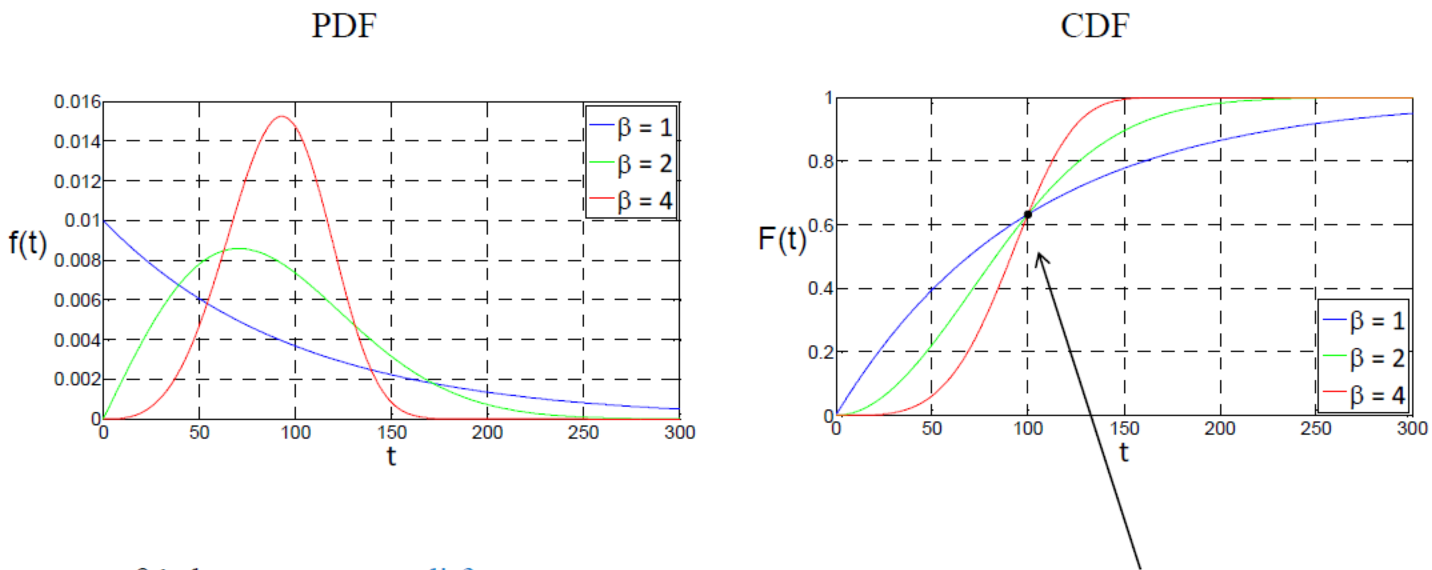
$$f(x) = \frac{k}{\lambda^k} x^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k}$$

$$F(x) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k}$$

Probability

S, that is the random variable

Variazione della forma di *pdf* e *cdf* al variare di β (α = 100 s cost)



β è detto parametro di forma

Da notare che $F(\alpha) = 63.2\%$

The distribution of , that is the random variable is also characterized by the so-called scale effect:

If there is a system composed of *n identical components in series, whose reliability is* described by a , that is the random variable with parameters α and β

then

the reliability of the entire system is a , that is the random variable , still described by a , that is the random variable , with parameters *β and* $\alpha_{total} = \alpha/n^{1/\beta}$

Probability

S, that is the random variable

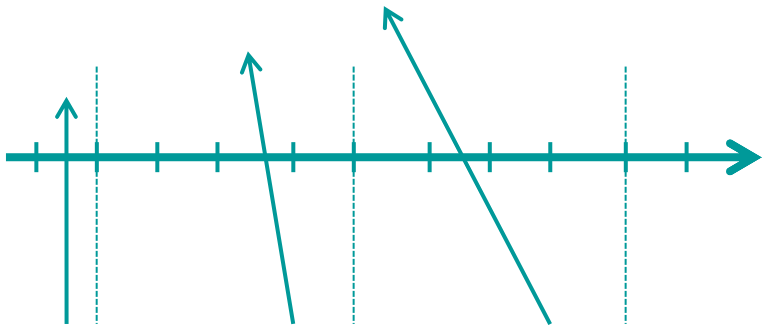
Example 1:

Consider car suspension springs, whose fatigue life is described by a , that is the random variable law with α = 300,000 cycles and β = 2.

Calculate:

1. **The 10%, 50%, and 90% percentiles**;
2. The percentage of pieces that break before 100,000 cycles;
3. The failure rate at 200,000 and 300,000 cycles;
4. If a car had these same springs mounted on both axles, at what number of cycles would the failure probability of 10% for the overall system correspond?

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right]$$



Let$_{10}$ % **Let$_{50}$ %** **Let$_{90}$ %**

$$t_p = \alpha \cdot [-\ln(1-p)]^{1/\beta}$$

$$t_{10\%} = 300000 \cdot [-\ln(1-0.1)]^{1/2} = 97378$$

$$t_{50\%} = 300000 \cdot [-\ln(1-0.5)]^{1/2} = 249766$$

$$t_{90\%} = 300000 \cdot [-\ln(1-0.9)]^{1/2} = 455228$$

Probability

S, that is the random variable

Example 1:

Consider car suspension springs, whose fatigue life is described by a , that is the random variable law with α = 300,000 cycles and β = 2.

Calculate:
1. The 10%, 50%, and 90% percentiles;
2. **The percentage of pieces that break before 100,000 cycles**;
3. The failure rate at 200,000 and 300,000 cycles;
4. If a car had these same springs mounted on both axles, at what number of cycles would the failure probability of 10% for the overall system correspond?

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right]$$

$$F(100000) = 1 - \exp\left[-\left(\frac{100000}{300000}\right)^{2}\right] = 0.1052 = 10.52\%$$

Probability

S, that is the random variable

Example 1:

Consider car suspension springs, whose fatigue life is described by a Weibull law with α = 300,000 cycles and β = 2.

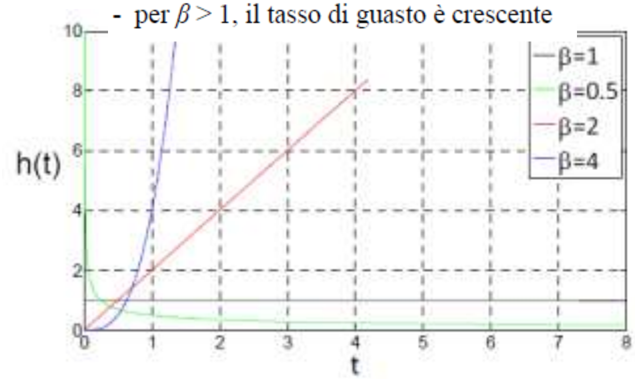
Calculate:
1. The 10%, 50%, and 90% percentiles;
2. The percentage of pieces that break before 100,000 cycles;
3. **The failure rate at 200,000 and 300,000 cycles**;
4. If a car had these same springs mounted on both axles, at what number of cycles would the failure probability of 10% for the overall system correspond?

$$h(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1-F(t)} = \frac{\beta}{\alpha^{\beta}} t^{\beta-1}$$

$$h(200000) = \frac{2}{300000^2} 200000^{2-1} = 4.44E-06 \quad \text{guasti/ciclo}$$

$$h(300000) = \frac{2}{300000}\left(\frac{300000}{300000}\right)^2 = \frac{2}{300000} = 6.66E-0.6 \quad \text{guasti/ciclo}$$

- per $\beta < 1$, il tasso di guasto è decrescente
- per $\beta = 1$, il tasso di guasto è costante
- per $\beta > 1$, il tasso di guasto è crescente



Probability

S, that is the random variable

Consider car suspension springs, whose fatigue life is described by a Weibull law with α = 300,000 cycles and β = 2.

Calculate:

1. The 10%, 50%, and 90% percentiles;
2. The percentage of pieces that break before 100,000 cycles;
3. The failure rate at 200,000 and 300,000 cycles;
4. **If a car had these same springs (4) mounted on both axles, at what number of cycles would the failure probability of 10% for the overall system correspond?**

Considerando inservibile il sistema quando uno solo dei suoi componenti va fuori uso → I vari componenti possono essere ricondotti ad un unico sistema in serie

$$\alpha_{tot} = \frac{\alpha}{n^{\frac{1}{\beta}}} = \frac{300000}{4^{\frac{1}{2}}} = \frac{300000}{\sqrt{4}} = 150000 \quad \text{cicli}$$

$$t_{10\%} = 150000 \cdot \left[-\ln(1-0.1)\right]^{1/2} = 48689 \quad \text{cicli}$$