The Monte Carlo method

Introduction

Karol Kołodziej

Institute of Physics
University of Silesia, Katowice
http://kk.us.edu.pl

The origin of the Monte Carlo method

The Monte Carlo method is a numerical method of solving mathematical, physical, etc., problems by means of random sampling.

It is conventionally assumed to be born in 1949 when the paper *The Monte Carlo method* was published:

N. Metropolis and S. Ulam, J. Amer. statistical assoc., **44** No. 247, 335-341.

The origin of the Monte Carlo method

The Monte Carlo method is a numerical method of solving mathematical, physical, etc., problems by means of random sampling.

It is conventionally assumed to be born in 1949 when the paper *The Monte Carlo method* was published:

N. Metropolis and S. Ulam, J. Amer. statistical assoc., **44** No. 247, 335-341.

John von Neumann and Stanisław Ulam - a member of the Lwów Mathematical School, who also became the US citizen in 1943 - are thought to have developed the method.

The origin of the Monte Carlo method

The Monte Carlo method is a numerical method of solving mathematical, physical, etc., problems by means of random sampling.

It is conventionally assumed to be born in 1949 when the paper *The Monte Carlo method* was published:

N. Metropolis and S. Ulam, J. Amer. statistical assoc., **44** No. 247, 335-341.

John von Neumann and Stanisław Ulam - a member of the Lwów Mathematical School, who also became the US citizen in 1943 - are thought to have developed the method.

Actually, theoretical basics of the methods were known much earlier, as some statistical problems were solved by means of random sampling, however, it could not be used on a significant scale because manual simulation of random variables is a very time consuming procedure.

The origin of the Monte Carlo method

The Monte Carlo method is a numerical method of solving mathematical, physical, etc., problems by means of random sampling.

It is conventionally assumed to be born in 1949 when the paper *The Monte Carlo method* was published:

N. Metropolis and S. Ulam, J. Amer. statistical assoc., **44** No. 247, 335-341.

John von Neumann and Stanisław Ulam - a member of the Lwów Mathematical School, who also became the US citizen in 1943 - are thought to have developed the method.

Actually, theoretical basics of the methods were known much earlier, as some statistical problems were solved by means of random sampling, however, it could not be used on a significant scale because manual simulation of random variables is a very time consuming procedure.

Thus, the Monte Carlo method became feasible only in the era of computers.

The origin of the Monte Carlo method

The Monte Carlo method is a numerical method of solving mathematical, physical, etc., problems by means of random sampling.

It is conventionally assumed to be born in 1949 when the paper *The Monte Carlo method* was published:

N. Metropolis and S. Ulam, J. Amer. statistical assoc., **44** No. 247, 335-341.

John von Neumann and Stanisław Ulam - a member of the Lwów Mathematical School, who also became the US citizen in 1943 - are thought to have developed the method.

Actually, theoretical basics of the methods were known much earlier, as some statistical problems were solved by means of random sampling, however, it could not be used on a significant scale because manual simulation of random variables is a very time consuming procedure.

Thus, the Monte Carlo method became feasible only in the era of computers.

The origin of the Monte Carlo method

The term *Monte Carlo* is due to the name of town Monte Carlo in the Principality of Monaco which is famous for its casino. The point is that the roulette is one of the simplest mechanical devices for generation of random numbers.

However, the Monte Carlo method does not obviously help to increase the chance of win in the game.

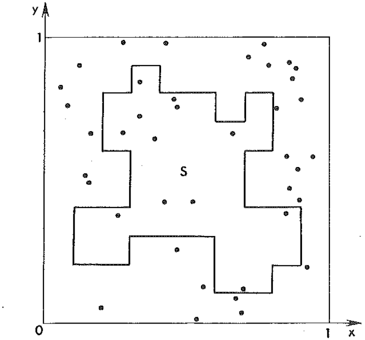
The origin of the Monte Carlo method

The term *Monte Carlo* is due to the name of town Monte Carlo in the Principality of Monaco which is famous for its casino. The point is that the roulette is one of the simplest mechanical devices for generation of random numbers.
However, the Monte Carlo method does not obviously help to increase the chance of win in the game.

Example

To clarify the idea behind the Monte Carlo method, let us consider a simply example.

Example Assume that we want to calculate the area of a flat figure S depicted in the figure below.

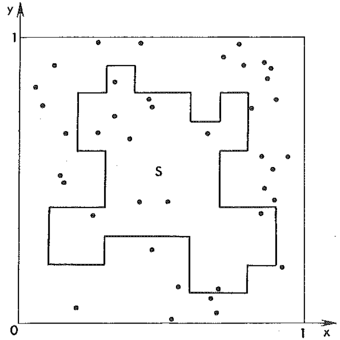


Assume that S is enclosed within a unit square. Now, let us select N random points, N' of which fall within S.

Example

To clarify the idea behind the Monte Carlo method, let us consider a simply example.

Example Assume that we want to calculate the area of a flat figure S depicted in the figure below.



Assume that S is enclosed within a unit square. Now, let us select N random points, N' of which fall within S.

Example

It is obvious that the area of S should be equal approximately to the ratio N'/N.

In our example $N = 40$ of which $N' = 12$ is inside S, thus

$$\frac{N'}{N} = \frac{12}{40} = 0.30,$$

while the true area of S is 0.35. Obviously, our approximation would become better, if we used more random points.

Example

It is obvious that the area of S should be equal approximately to the ratio N'/N.
In our example $N = 40$ of which $N' = 12$ is inside S, thus

$$\frac{N'}{N} = \frac{12}{40} = 0.30,$$

while the true area of S is 0.35. Obviously, our approximation would become better, if we used more random points.
In practise the Monte Carlo (MC) method is not used to calculate areas of flat figures, as there are much better methods, though a bit more complicated, which are much better in this respect.
However, the MC method becomes indispensable, if we want to calculate *volume* of a body in a multidimensional space.

Example

It is obvious that the area of S should be equal approximately to the ratio N'/N.
In our example $N = 40$ of which $N' = 12$ is inside S, thus

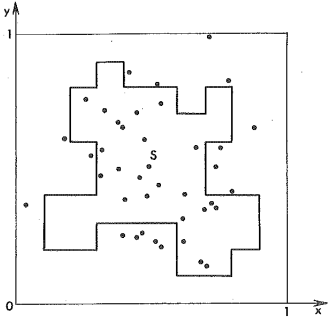
$$\frac{N'}{N} = \frac{12}{40} = 0.30,$$

while the true area of S is 0.35. Obviously, our approximation would become better, if we used more random points.
In practise the Monte Carlo (MC) method is not used to calculate areas of flat figures, as there are much better methods, though a bit more complicated, which are much better in this respect.
However, the MC method becomes indispensable, if we want to calculate *volume* of a body in a multidimensional space.

More about the example

How the random points should actually be selected?

Consider our example once more, but this time imagine that we fix our figure on the wall and let some dart player to hit it with darts from some distance.



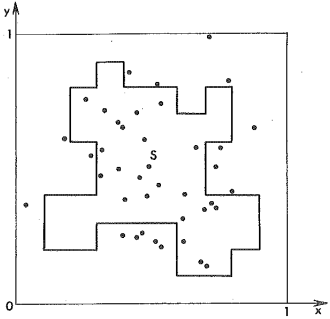
We see, that now $N = 40$ of which $N' = 24$ is inside S, thus

$$\frac{N'}{N} = \frac{24}{40} = 0.60,$$

More about the example

How the random points should actually be selected?

Consider our example once more, but this time imagine that we fix our figure on the wall and let some dart player to hit it with darts from some distance.



We see, that now $N = 40$ of which $N' = 24$ is inside S, thus

$$\frac{N'}{N} = \frac{24}{40} = 0.60,$$

More about the example

which is almost twice the true value of the area that is equal to 0.35.

Obviously, if the dart player was better our result would become even worse, as almost all hits would be inside S.

More about the example

which is almost twice the true value of the area that is equal to 0.35.

Obviously, if the dart player was better our result would become even worse, as almost all hits would be inside S.

We see, that our approximation is better if the random points are *uniformly scattered* over the whole square.

More about the example

which is almost twice the true value of the area that is equal to 0.35.

Obviously, if the dart player was better our result would become even worse, as almost all hits would be inside S.

We see, that our approximation is better if the random points are *uniformly scattered* over the whole square.

Random variables

The term *random variable*, in its commonplace sense is used to emphasize that we do not know what specific value this variable will assume.

Mathematicians, however, use the term *random variable* in quite a definite sens. Although we indeed do not know what value will this variable take on in a given case, we do know what values it can assume and what are the respective probabilities.

Random variables

The term *random variable*, in its commonplace sense is used to emphasize that we do not know what specific value this variable will assume.

Mathematicians, however, use the term *random variable* in quite a definite sens. Although we indeed do not know what value will this variable take on in a given case, we do know what values it can assume and what are the respective probabilities.

The result of a single trial representing the random variable cannot be predicted, but we can predict very reliably the result of a large number of trials.

The bigger the number of trials $\Rightarrow$ the more accurate prediction.

Random variables

The term *random variable*, in its commonplace sense is used to emphasize that we do not know what specific value this variable will assume.

Mathematicians, however, use the term *random variable* in quite a definite sens. Although we indeed do not know what value will this variable take on in a given case, we do know what values it can assume and what are the respective probabilities.

The result of a single trial representing the random variable cannot be predicted, but we can predict very reliably the result of a large number of trials.

The bigger the number of trials $\Rightarrow$ the more accurate prediction.

Discrete random variables

A random variable ξ which takes on values from a discrete set $\{x_1, x_2, ..., x_n\}$ is specified by the table

$$\xi = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix},$$

where $p_1, p_2, ..., p_n$ are the corresponding probabilities.
Denote the probability that the random variable ξ will be equal to x_i by

$$\boldsymbol{P}\{\xi = x_i\} = p_i.$$

Discrete random variables

A random variable ξ which takes on values from a discrete set $\{x_1, x_2, ..., x_n\}$ is specified by the table

$$\xi = \begin{pmatrix} x_1 & x_2 & ... & x_n \\ p_1 & p_2 & ... & p_n \end{pmatrix},$$

where $p_1, p_2, ..., p_n$ are the corresponding probabilities.
Denote the probability that the random variable ξ will be equal to x_i by

$$\boldsymbol{P}\{\xi = x_i\} = p_i.$$

While the values $x_1, x_2, ..., x_n$ can be arbitrary, the probabilities $p_1, p_2, ..., p_n$ must satisfy simultaneously two conditions:

1. all p_i's are non negative: $p_i > 0$;
2. the sum of all p_i's is equal to 1: $p_1 + p_2 + ... + p_n = 1$.

Discrete random variables

A random variable ξ which takes on values from a discrete set $\{x_1, x_2, ..., x_n\}$ is specified by the table

$$\xi = \begin{pmatrix} x_1 & x_2 & ... & x_n \\ p_1 & p_2 & ... & p_n \end{pmatrix},$$

where $p_1, p_2, ..., p_n$ are the corresponding probabilities.
Denote the probability that the random variable ξ will be equal to x_i by

$$\boldsymbol{P}\{\xi = x_i\} = p_i.$$

While the values $x_1, x_2, ..., x_n$ can be arbitrary, the probabilities $p_1, p_2, ..., p_n$ must satisfy simultaneously two conditions:

1. all p_i's are non negative: $p_i > 0$;
2. the sum of all p_i's is equal to 1: $p_1 + p_2 + ... + p_n = 1$.

Condition 2 means that in each trial ξ must assume one of the values $x_1, x_2, ..., x_n$.

The quantity

$$\boldsymbol{M}\xi = \sum_{i=1}^{n} x_i p_i$$

is called the *expected value* or *mean value* of the random variable ξ.

Discrete random variables

Condition 2 means that in each trial ξ must assume one of the values $x_1, x_2, ..., x_n$.
The quantity

$$\boldsymbol{M}\xi = \sum_{i=1}^{n} x_i p_i$$

is called the *expected value* or *mean value* of the random variable ξ. We see that values x_i with greater probabilities contribute more to the mean than those associated with smaller probabilities.

Discrete random variables

Condition 2 means that in each trial ξ must assume one of the values $x_1, x_2, ..., x_n$.
The quantity

$$\boldsymbol{M}\xi = \sum_{i=1}^{n} x_i p_i$$

is called the *expected value* or *mean value* of the random variable ξ. We see that values x_i with greater probabilities contribute more to the mean than those associated with smaller probabilities.

Basic properties of the expected value

If c is an arbitrary non-random number, then

$$\begin{aligned} \boldsymbol{M}(\xi + c) &= \boldsymbol{M}\xi + c, \\ \boldsymbol{M}(c\xi) &= c\boldsymbol{M}\xi. \end{aligned}$$

If ξ and η are two arbitrary random variables, then

$$\boldsymbol{M}(\xi + \eta) = \boldsymbol{M}\xi + \boldsymbol{M}\eta.$$

The quantity

$$\boldsymbol{D}\xi = \boldsymbol{M}[(\xi - \boldsymbol{M}\xi)^2]$$

is called the *variance* of the random variable ξ.

Basic properties of the expected value

If c is an arbitrary non-random number, then

$$\begin{aligned} \boldsymbol{M}(\xi + c) &= \boldsymbol{M}\xi + c, \\ \boldsymbol{M}(c\xi) &= c\boldsymbol{M}\xi. \end{aligned}$$

If ξ and η are two arbitrary random variables, then

$$\boldsymbol{M}(\xi + \eta) = \boldsymbol{M}\xi + \boldsymbol{M}\eta.$$

The quantity

$$\boldsymbol{D}\xi = \boldsymbol{M}[(\xi - \boldsymbol{M}\xi)^2]$$

is called the *variance* of the random variable ξ. Thus variance $\boldsymbol{D}\xi$ is the expected value of the square of the deviation of the random variable ξ from the mean value $\boldsymbol{M}\xi$. It is obvious that in all cases $\boldsymbol{D}\xi > 0$.

Basic properties of the expected value

If c is an arbitrary non-random number, then

$$\begin{aligned} \boldsymbol{M}(\xi + c) &= \boldsymbol{M}\xi + c, \\ \boldsymbol{M}(c\xi) &= c\boldsymbol{M}\xi. \end{aligned}$$

If ξ and η are two arbitrary random variables, then

$$\boldsymbol{M}(\xi + \eta) = \boldsymbol{M}\xi + \boldsymbol{M}\eta.$$

The quantity

$$\boldsymbol{D}\xi = \boldsymbol{M}[(\xi - \boldsymbol{M}\xi)^2]$$

is called the *variance* of the random variable ξ. Thus variance $\boldsymbol{D}\xi$ is the expected value of the square of the deviation of the random variable ξ from the mean value $\boldsymbol{M}\xi$. It is obvious that in all cases $\boldsymbol{D}\xi > 0$.

Expected value and variance are the two most important numerical characteristics of the random variable ξ. What is their practical importance?

Basic properties of the expected value

If c is an arbitrary non-random number, then

$$\begin{aligned} \boldsymbol{M}(\xi + c) &= \boldsymbol{M}\xi + c, \\ \boldsymbol{M}(c\xi) &= c\boldsymbol{M}\xi. \end{aligned}$$

If ξ and η are two arbitrary random variables, then

$$\boldsymbol{M}(\xi + \eta) = \boldsymbol{M}\xi + \boldsymbol{M}\eta.$$

The quantity

$$\boldsymbol{D}\xi = \boldsymbol{M}[(\xi - \boldsymbol{M}\xi)^2]$$

is called the *variance* of the random variable ξ. Thus variance $\boldsymbol{D}\xi$ is the expected value of the square of the deviation of the random variable ξ from the mean value $\boldsymbol{M}\xi$. It is obvious that in all cases $\boldsymbol{D}\xi > 0$.

Expected value and variance are the two most important numerical characteristics of the random variable ξ. What is their practical importance?

Practical importance of $\boldsymbol{M}\xi$ and $\boldsymbol{D}\xi$

If in the course of many observations of the random variable ξ we obtain a set of values $\xi_1, \xi_2, ..., \xi_N$, each of them being equal to one of the numbers $x_1, x_2, ..., x_n$, then the arithmetic mean of these numbers will be close to $\boldsymbol{M}\xi$, i.e.

$$\frac{1}{N}(\xi_1 + \xi_2 + ... + \xi_N) \approx \boldsymbol{M}\xi.$$

The variance $\boldsymbol{D}\xi > 0$ characterizes the spreading of $\xi_1, \xi_2, ..., \xi_N$ around the mean value $\boldsymbol{M}\xi$.

Practical importance of $\boldsymbol{M}\xi$ and $\boldsymbol{D}\xi$

If in the course of many observations of the random variable ξ we obtain a set of values $\xi_1, \xi_2, ..., \xi_N$, each of them being equal to one of the numbers $x_1, x_2, ..., x_n$, then the arithmetic mean of these numbers will be close to $\boldsymbol{M}\xi$, i.e.

$$\frac{1}{N}(\xi_1 + \xi_2 + ... + \xi_N) \approx \boldsymbol{M}\xi.$$

The variance $\boldsymbol{D}\xi > 0$ characterizes the spreading of $\xi_1, \xi_2, ..., \xi_N$ around the mean value $\boldsymbol{M}\xi$.

Using the definition and the above properties of the expected value variance $\boldsymbol{D}\xi$ can be written in the following way

$$\begin{aligned} \boldsymbol{D}\xi &= \boldsymbol{M}[(\xi - \boldsymbol{M}\xi)^2] = \boldsymbol{M}[\xi^2 - 2\boldsymbol{M}\xi \cdot \xi + (\boldsymbol{M}\xi)^2] \\ &= \boldsymbol{M}(\xi^2) - 2\boldsymbol{M}\xi \cdot \boldsymbol{M}\xi + (\boldsymbol{M}\xi)^2], \end{aligned}$$

whence

$$\boldsymbol{D}\xi = \boldsymbol{M}(\xi^2) - (\boldsymbol{M}\xi)^2.$$

Practical importance of $\boldsymbol{M}\xi$ and $\boldsymbol{D}\xi$

If in the course of many observations of the random variable ξ we obtain a set of values $\xi_1, \xi_2, ..., \xi_N$, each of them being equal to one of the numbers $x_1, x_2, ..., x_n$, then the arithmetic mean of these numbers will be close to $\boldsymbol{M}\xi$, i.e.

$$\frac{1}{N}(\xi_1 + \xi_2 + ... + \xi_N) \approx \boldsymbol{M}\xi.$$

The variance $\boldsymbol{D}\xi > 0$ characterizes the spreading of $\xi_1, \xi_2, ..., \xi_N$ around the mean value $\boldsymbol{M}\xi$.

Using the definition and the above properties of the expected value variance $\boldsymbol{D}\xi$ can be written in the following way

$$\begin{aligned} \boldsymbol{D}\xi &= \boldsymbol{M}[(\xi - \boldsymbol{M}\xi)^2] = \boldsymbol{M}[\xi^2 - 2\boldsymbol{M}\xi \cdot \xi + (\boldsymbol{M}\xi)^2] \\ &= \boldsymbol{M}(\xi^2) - 2\boldsymbol{M}\xi \cdot \boldsymbol{M}\xi + (\boldsymbol{M}\xi)^2], \end{aligned}$$

whence

$$\boldsymbol{D}\xi = \boldsymbol{M}(\xi^2) - (\boldsymbol{M}\xi)^2.$$

Basic properties of variance

If c is an arbitrary non-random number, then

$$\begin{aligned} \boldsymbol{D}(\xi + c) &= \boldsymbol{D}\xi + c, \\ \boldsymbol{D}(c\xi) &= c^2\boldsymbol{D}\xi. \end{aligned}$$

The concept of *independence* of random variables plays an important role in the probability theory. Although being quite complicated, it may be quite clear in the simplest cases.

If c is an arbitrary non-random number, then

$$\begin{aligned} \boldsymbol{D}(\xi + c) &= \boldsymbol{D}\xi + c, \\ \boldsymbol{D}(c\xi) &= c^2 \boldsymbol{D}\xi. \end{aligned}$$

The concept of *independence* of random variables plays an important role in the probability theory. Although being quite complicated, it may be quite clear in the simplest cases.

Let us assume a simultaneous observation of random variables ξ and η. If the distribution of ξ does not change after we have found the value of η, then it natural to consider ξ as being independent of η.

Basic properties of variance

If c is an arbitrary non-random number, then

$$\begin{aligned} \boldsymbol{D}(\xi + c) &= \boldsymbol{D}\xi + c, \\ \boldsymbol{D}(c\xi) &= c^2 \boldsymbol{D}\xi. \end{aligned}$$

The concept of *independence* of random variables plays an important role in the probability theory. Although being quite complicated, it may be quite clear in the simplest cases.

Let us assume a simultaneous observation of random variables ξ and η. If the distribution of ξ does not change after we have found the value of η, then it natural to consider ξ as being independent of η.

The following relations are valid for independent random variables ξ and η

$$\begin{aligned} \boldsymbol{M}(\xi\eta) &= \boldsymbol{M}\xi \cdot \boldsymbol{M}\eta, \\ \boldsymbol{D}(\xi + \eta) &= \boldsymbol{D}\xi + \boldsymbol{D}\eta. \end{aligned}$$

Basic properties of variance

If c is an arbitrary non-random number, then

$$\begin{aligned} \boldsymbol{D}(\xi + c) &= \boldsymbol{D}\xi + c, \\ \boldsymbol{D}(c\xi) &= c^2 \boldsymbol{D}\xi. \end{aligned}$$

The concept of *independence* of random variables plays an important role in the probability theory. Although being quite complicated, it may be quite clear in the simplest cases.

Let us assume a simultaneous observation of random variables ξ and η. If the distribution of ξ does not change after we have found the value of η, then it natural to consider ξ as being independent of η.

The following relations are valid for independent random variables ξ and η

$$\begin{aligned} \boldsymbol{M}(\xi\eta) &= \boldsymbol{M}\xi \cdot \boldsymbol{M}\eta, \\ \boldsymbol{D}(\xi + \eta) &= \boldsymbol{D}\xi + \boldsymbol{D}\eta. \end{aligned}$$

Example. Let us consider a random variable κ which is the number of points obtained in a single throw of a die. It can be represented by the following table

$$\kappa = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}.$$

Let us find the expected value and variance of κ.

$$\boldsymbol{M}\kappa = \sum_{i=1}^{6} x_i p_i = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + ... + 6 \cdot \frac{1}{6} = 3.5,$$

$$\boldsymbol{D}\kappa = \boldsymbol{M}(\kappa^2) - (\boldsymbol{M}\kappa)^2 = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + ... + 6^2 \cdot \frac{1}{6} - 3.5^2 = 2.917.$$

Example. Let us consider a random variable θ which is a single toss of a coin, assuming that a toss of heads brings 3 points an that of tails brings 4 points. It can be represented by the following table

$$\theta = \begin{pmatrix} 3 & 4 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

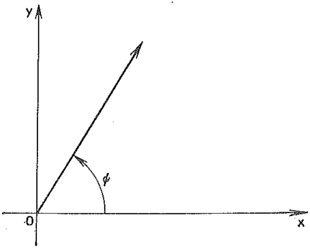
The expected value and variance of θ are

$$\boldsymbol{M}\theta = \sum_{i=1}^{2} x_i p_i = 3 \cdot \frac{1}{2} + 4 \cdot \frac{1}{2} = 3.5,$$

$$\boldsymbol{D}\theta = \boldsymbol{M}(\theta^2) - (\boldsymbol{M}\theta)^2 = 3^2 \cdot \frac{1}{2} + 4^2 \cdot \frac{1}{2} - 3.5^2 = 0.25.$$

We see that $\boldsymbol{M}\theta = \boldsymbol{M}\kappa$, but $\boldsymbol{D}\theta < \boldsymbol{D}\kappa$. The inequality can be easily understood, as the maximum deviation of θ from the mean 3.5 is ± 0.5, while the maximum deviation of κ from the mean is ± 2.5.

Let us assume that a certain amount of radium is placed on a plane at the origin of the coordinates. An α-particle is emitted in each decay of a radium atom. The direction of motion of this particle will be characterized by an angle ψ, see the figure below.



Since any direction of emission is possible, the random variable ψ can assume any value fro 0 to 2π.

We shall call a random variable ξ *continuous* if it takes on any value out of a certain interval (a, b).

A continuous random variable ξ is defined by specifying the interval (a, b) of its variation and the function $p(x)$ called the *probability density* of the random variable ξ.

The physical meaning of $p(x)$ is the following. Let (a', b') be an arbitrary interval contained within (a, b), i.e. $a \leq a'$ and $b' \leq b$. Then the probability that ξ falls inside (a', b') is given by the integral

$$\boldsymbol{P}\{a' < \xi < b'\} = \int_{a'}^{b'} p(x)dx.$$

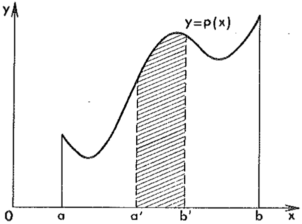
We shall call a random variable ξ *continuous* if it takes on any value out of a certain interval (a, b).

A continuous random variable ξ is defined by specifying the interval (a, b) of its variation and the function $p(x)$ called the *probability density* of the random variable ξ.

The physical meaning of $p(x)$ is the following. Let (a', b') be an arbitrary interval contained within (a, b), i.e. $a \leq a'$ and $b' \leq b$. Then the probability that ξ falls inside (a', b') is given by the integral

$$\boldsymbol{P}\{a' < \xi < b'\} = \int_{a'}^{b'} p(x)dx.$$

The integral $\int_{a'}^{b'} p(x)dx$ is equal to the hatched area in the following figure.



The set of values of ξ can be any interval. The cases when $a = -\infty$ as well as $b = +\infty$ are also possible.

Continuous random variables

However, the density $p(x)$ must satisfy two conditions, similar to those of probabilities p_i in the case of a discrete random variable:

1. the density $p(x)$ is positive: $p(x) > 0$,
2. the integral of the density $p(x)$ over the whole interval (a, b) is equal to 1:

$$\int_a^b p(x)dx = 1.$$

The value

$$\boldsymbol{M}\xi = \int_a^b xp(x)dx$$

is called *expected* or *mean value* of the continuous random variable.

Continuous random variables

However, the density $p(x)$ must satisfy two conditions, similar to those of probabilities p_i in the case of a discrete random variable:

1. the density $p(x)$ is positive: $p(x) > 0$,
2. the integral of the density $p(x)$ over the whole interval (a, b) is equal to 1:

$$\int_a^b p(x)dx = 1.$$

The value

$$\boldsymbol{M}\xi = \int_a^b xp(x)dx$$

is called *expected* or *mean value* of the continuous random variable.

Continuous random variables

All properties of the expected value and variance of the continuous random variable are the same as those of the discrete random variable.

We will mention here only one more expression for the expected value of a function of ξ. Let us assume that the random variable ξ is characterized by the probability density $p(x)$ and consider an arbitrary continuous function $f(x)$. We introduce a random variable $\eta = f(\xi)$. It can be shown that

$$\boldsymbol{M} f(\xi) = \int_a^b f(x) p(x) dx.$$

We must emphasize that in general $\boldsymbol{M} f(\xi) \neq f(\boldsymbol{M}\xi)$.

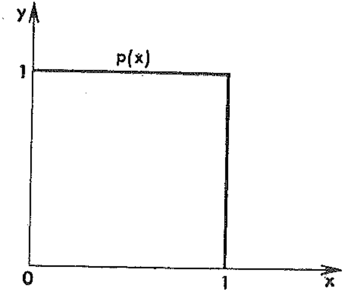
All properties of the expected value and variance of the continuous random variable are the same as those of the discrete random variable.

We will mention here only one more expression for the expected value of a function of ξ. Let us assume that the random variable ξ is characterized by the probability density $p(x)$ and consider an arbitrary continuous function $f(x)$. We introduce a random variable $\eta = f(\xi)$. It can be shown that

$$\boldsymbol{M} f(\xi) = \int_a^b f(x) p(x) dx.$$

We must emphasize that in general $\boldsymbol{M} f(\xi) \neq f(\boldsymbol{M}\xi)$.

A random variable γ, defined on the interval $(0, 1)$ with the density $p(x) = 1$ is said to be *uniformly distributed* in $(0, 1)$, see the figure below.



Continuous random variables

Indeed, for any interval (a', b') within $(0, 1)$ the probability that γ will take on the value within (a', b') is equal to

$$\int_{a'}^{b'} p(x)dx = b' - a',$$

i.e. to the length of this interval. If, in particular, we divide the interval $(0, 1)$ into an arbitrary number of intervals of equal length, the probability of γ falling within one of these intervals will be the same. Note that

$$\boldsymbol{M}\gamma = \int_0^1 xp(x)dx = \int_0^1 xdx = \frac{1}{2},$$

$$\boldsymbol{D}\gamma = \int_0^1 x^2 p(x)dx - (\boldsymbol{M}\gamma)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

Indeed, for any interval (a', b') within $(0, 1)$ the probability that γ will take on the value within (a', b') is equal to

$$\int_{a'}^{b'} p(x)dx = b' - a',$$

i.e. to the length of this interval. If, in particular, we divide the interval $(0, 1)$ into an arbitrary number of intervals of equal length, the probability of γ falling within one of these intervals will be the same. Note that

$$\boldsymbol{M}\gamma = \int_0^1 xp(x)dx = \int_0^1 xdx = \frac{1}{2},$$

$$\boldsymbol{D}\gamma = \int_0^1 x^2 p(x)dx - (\boldsymbol{M}\gamma)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$