The general scheme of the Monte Carlo method

Central limit theorem

Karol Kołodziej

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

Normal random variables

A random variable ζ defined on the entire axis $(-\infty, +\infty)$ and characterized by the density

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}},$$

where a and $\sigma > 0$ are numerical parameters, is said to be *normal* or *Gaussian* random variable.

The parameter a does not affect the shape of the $p(x)$ curve. It just shifts the curve as a whole along the x-axis. On the contrary, variation of σ changes the shape of the curve. It is easily seen that

$$\max p(x) = p(a) = \frac{1}{\sqrt{2\pi}\sigma}.$$

Hence, a decrease of σ increases $\max p(x)$, but as $\int_{-\infty}^{+\infty} p(x)dx = 1$, the entire area below the $p(x)$ curve is equal to 1.

Normal random variables

A random variable ζ defined on the entire axis $(-\infty, +\infty)$ and characterized by the density

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}},$$

where a and $\sigma > 0$ are numerical parameters, is said to be *normal* or *Gaussian* random variable.

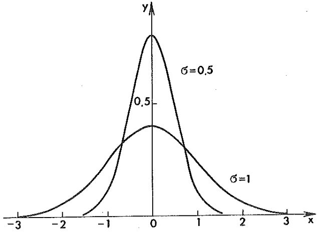
The parameter a does not affect the shape of the $p(x)$ curve. It just shifts the curve as a whole along the x-axis. On the contrary, variation of σ changes the shape of the curve. It is easily seen that

$$\max p(x) = p(a) = \frac{1}{\sqrt{2\pi}\sigma}.$$

Hence, a decrease of σ increases $\max p(x)$, but as $\int_{-\infty}^{+\infty} p(x)dx = 1$, the entire area below the $p(x)$ curve is equal to 1.

Normal random variables

Therefore the $p(x)$ curve will stretch upward in the vicinity of $x = a$, but it will decrease at sufficiently large of x. The normal densities for $a = 0$ and $\sigma = 1$, as well as for $a = 0$ and $\sigma = 0.5$ are plotted below.



Exercise. Prove that

$$\boldsymbol{M}\zeta = a, \qquad \boldsymbol{D}\zeta = \sigma^2.$$

Normal random variables are frequently encountered in investigation of many diverse problems. Why it is so, will be discussed later.
For example, the experimental error δ is usually a normal random variable. If there is no systematic error (bias) then

$$a = \mathbf{M}\delta = 0.$$

The value

$$\sigma = \sqrt{\mathbf{D}\zeta},$$

called *standard deviation*, characterizes the error of the measurement.

The “3 sigma” rule

It can be shown that regardless of the values of a and σ

$$\int_{a-3\sigma}^{a+3\sigma} p(x)dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{a-3\sigma}^{a+3\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = 0.997.$$

Hence the corresponding probability of finding the normal random variable ζ in the interval $(a - 3\sigma, a + 3\sigma)$ is equal to

$$\boldsymbol{P}\{a - 3\sigma < \zeta < a + 3\sigma\} = 0.997.$$

The probability 0.997 is so close to 1 that sometimes this formula is interpreted in the following way: *it is almost impossible to obtain in a single trial a value of ζ deviating from $\boldsymbol{M}\zeta$ by more than 3σ.*

Central limit theorem

This spectacular theorem was first formulated and proved by the French mathematician Pierre-Simon Laplace in 1810. A number of outstanding mathematicians, P.L. Chebyshev, A.A. Markov and A.M. Lyapunov among them, investigated the problem of generalizing this theorem. Its proof is rather complicated, therefore we will not present it here.

Let us consider N identical independent random variables $\xi_1, \xi_2, ..., \xi_N$ such that their probability distributions coincide. Consequently, both their expected values and variances coincide as well.

Let us denote

$$\boldsymbol{M}\xi_1 = \boldsymbol{M}\xi_2 = ... = \boldsymbol{M}\xi_N = m,$$
$$\boldsymbol{D}\xi_1 = \boldsymbol{D}\xi_2 = ... = \boldsymbol{D}\xi_N = b^2$$

and denote the sum of all these values as ρ_N

$$\rho_N = \xi_1 + \xi_2 + ... + \xi_N.$$

Central limit theorem

This spectacular theorem was first formulated and proved by the French mathematician Pierre-Simon Laplace in 1810. A number of outstanding mathematicians, P.L. Chebyshev, A.A. Markov and A.M. Lyapunov among them, investigated the problem of generalizing this theorem. Its proof is rather complicated, therefore we will not present it here.

Let us consider N identical independent random variables $\xi_1, \xi_2, ..., \xi_N$ such that their probability distributions coincide. Consequently, both their expected values and variances coincide as well.

Let us denote

$$\boldsymbol{M}\xi_1 = \boldsymbol{M}\xi_2 = ... = \boldsymbol{M}\xi_N = m,$$
$$\boldsymbol{D}\xi_1 = \boldsymbol{D}\xi_2 = ... = \boldsymbol{D}\xi_N = b^2$$

and denote the sum of all these values as ρ_N

$$\rho_N = \xi_1 + \xi_2 + ... + \xi_N.$$

Central limit theorem

Now we know that the expected value and variance of the sum of independent random variables are additive, hence we get

$$\boldsymbol{M}\rho_N = \boldsymbol{M}(\xi_1 + \xi_2 + ... + \xi_N) = \boldsymbol{M}\xi_1 + \boldsymbol{M}\xi_2 + ... + \boldsymbol{M}\xi_N = Nm,$$
$$\boldsymbol{D}\rho_N = \boldsymbol{D}(\xi_1 + \xi_2 + ... + \xi_N) = \boldsymbol{D}\xi_1 + \boldsymbol{D}\xi_2 + ... + \boldsymbol{D}\xi_N = Nb^2.$$

Let us consider now a normal random variable ζ_N with $a = Nm$ and $\sigma^2 = Nb^2$. The central limit theorem states that for any interval (a', b') for large N

$$\boldsymbol{P}\{a' < \rho_N < b'\} \approx \int_{a'}^{b'} p_{\zeta_N}(x)dx.$$

Central limit theorem

Now we know that the expected value and variance of the sum of independent random variables are additive, hence we get

$$\boldsymbol{M}\rho_N = \boldsymbol{M}(\xi_1 + \xi_2 + ... + \xi_N) = \boldsymbol{M}\xi_1 + \boldsymbol{M}\xi_2 + ... + \boldsymbol{M}\xi_N = Nm,$$
$$\boldsymbol{D}\rho_N = \boldsymbol{D}(\xi_1 + \xi_2 + ... + \xi_N) = \boldsymbol{D}\xi_1 + \boldsymbol{D}\xi_2 + ... + \boldsymbol{D}\xi_N = Nb^2.$$

Let us consider now a normal random variable ζ_N with $a = Nm$ and $\sigma^2 = Nb^2$. The central limit theorem states that for any interval (a', b') for large N

$$\boldsymbol{P}\{a' < \rho_N < b'\} \approx \int_{a'}^{b'} p_{\zeta_N}(x)dx.$$

The physical meaning of this theorem is obvious: the sum ρ_N of a large number of identical random variables is approximately normal, i.e. $p_{\rho_N}(x) \approx p_{\zeta_N}(x)$.

Now we know that the expected value and variance of the sum of independent random variables are additive, hence we get

$$\boldsymbol{M}\rho_N = \boldsymbol{M}(\xi_1 + \xi_2 + ... + \xi_N) = \boldsymbol{M}\xi_1 + \boldsymbol{M}\xi_2 + ... + \boldsymbol{M}\xi_N = Nm,$$
$$\boldsymbol{D}\rho_N = \boldsymbol{D}(\xi_1 + \xi_2 + ... + \xi_N) = \boldsymbol{D}\xi_1 + \boldsymbol{D}\xi_2 + ... + \boldsymbol{D}\xi_N = Nb^2.$$

Let us consider now a normal random variable ζ_N with $a = Nm$ and $\sigma^2 = Nb^2$. The central limit theorem states that for any interval (a', b') for large N

$$\boldsymbol{P}\{a' < \rho_N < b'\} \approx \int_{a'}^{b'} p_{\zeta_N}(x)dx.$$

The physical meaning of this theorem is obvious: the sum ρ_N of a large number of identical random variables is approximately normal, i.e. $p_{\rho_N}(x) \approx p_{\zeta_N}(x)$.

Actually this theorem remains valid for much more general conditions: all the addends $\xi_1, \xi_2, ..., \xi_N$ need not be independent, it is only essential that none of them plays too great a role in the sum.

It is this theorem that explains why normal random variables are met so often in nature. Indeed, whenever we come across an aggregate effect of a large number of negligible random factors we find that the resulting random variable is normal.

Actually this theorem remains valid for much more general conditions: all the addends $\xi_1, \xi_2, ..., \xi_N$ need not be independent, it is only essential that none of them plays too great a role in the sum.

It is this theorem that explains why normal random variables are met so often in nature. Indeed, whenever we come across an aggregate effect of a large number of negligible random factors we find that the resulting random variable is normal.

For example, deviation of an artillery shell off the target is almost always a normal random variable because it depends on meteorological conditions in various sections of the trajectory as well as on many other factors.

Central limit theorem

Actually this theorem remains valid for much more general conditions: all the addends $\xi_1, \xi_2, ..., \xi_N$ need not be independent, it is only essential that none of them plays too great a role in the sum.

It is this theorem that explains why normal random variables are met so often in nature. Indeed, whenever we come across an aggregate effect of a large number of negligible random factors we find that the resulting random variable is normal.

For example, deviation of an artillery shell off the target is almost always a normal random variable because it depends on meteorological conditions in various sections of the trajectory as well as on many other factors.

Let us assume that we want to calculate some unknown value m. We shall try to find a random variable ξ such that $\boldsymbol{M}\xi = m$. Let us assume also that $\boldsymbol{D}\xi = b^2$.

Let us consider N random independent variables $\xi_1, \xi_2, ..., \xi_N$ with distributions identical to that of ξ. If N sufficiently large than the distribution of the sum $\rho_N = \xi_1 + \xi_2 + ... + \xi_N$ will be approximately normal with the parameters $a = Nm$, $\sigma^2 = Nb^2$ and following the 3 sigma rule

$$\boldsymbol{P}\{a - 3\sigma < \rho_N < a + 3\sigma\} \approx 0.997,$$

we will have

$$\boldsymbol{P}\{Nm - 3b\sqrt{N} < \rho_N < Nm + 3b\sqrt{N}\} \approx 0.997.$$

If we divide the inequalities in braces by N we will get an equivalent inequality and the probability will not change. Hence

$$\boldsymbol{P}\left\{m - \frac{3b}{\sqrt{N}} < \frac{\rho_N}{N} < m + \frac{3b}{\sqrt{N}}\right\} \approx 0.997.$$

The general scheme of the MC method

Let us assume that we want to calculate some unknown value m. We shall try to find a random variable ξ such that $\boldsymbol{M}\xi = m$. Let us assume also that $\boldsymbol{D}\xi = b^2$.
Let us consider N random independent variables $\xi_1, \xi_2, ..., \xi_N$ with distributions identical to that of ξ. If N sufficiently large than the distribution of the sum $\rho_N = \xi_1 + \xi_2 + ... + \xi_N$ will be approximately normal with the parameters $a = Nm$, $\sigma^2 = Nb^2$ and following the 3 sigma rule

$$\boldsymbol{P}\{a - 3\sigma < \rho_N < a + 3\sigma\} \approx 0.997,$$

we will have

$$\boldsymbol{P}\{Nm - 3b\sqrt{N} < \rho_N < Nm + 3b\sqrt{N}\} \approx 0.997.$$

If we divide the inequalities in braces by N we will get an equivalent inequality and the probability will not change. Hence

$$\boldsymbol{P}\left\{m - \frac{3b}{\sqrt{N}} < \frac{\rho_N}{N} < m + \frac{3b}{\sqrt{N}}\right\} \approx 0.997.$$

The general scheme of the MC method

Now, let us rewrite this expression in a somewhat different form:

$$\boldsymbol{P}\left\{\left|\frac{1}{N}\sum_{j=1}^{N}\xi_j - m\right| < \frac{3b}{\sqrt{N}}\right\} \approx 0.997.$$

This formula is very important to the Monte Carlo method. It gives us both the method of calculating m and the estimate of the error.

Let us indeed sample N values of the random variable ξ. Actually, determining a single value of the variables $\xi_1, \xi_2, ..., \xi_N$ is equivalent to determining N values of a single variable ξ, because all these random variables are identical, i.e. have identical distributions. Our formula shows that the arithmetic mean of these values will be approximately equal to m. The error of the approximation will most probably not exceed the value $3b/\sqrt{N}$ and it will approach zero if N increases.

The general scheme of the MC method

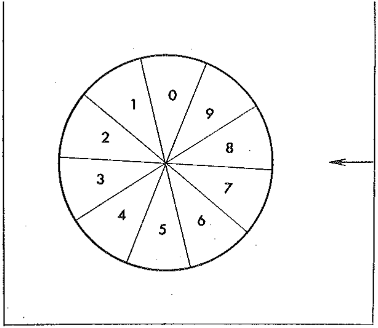
Now, let us rewrite this expression in a somewhat different form:

$$\boldsymbol{P}\left\{\left|\frac{1}{N}\sum_{j=1}^{N}\xi_j - m\right| < \frac{3b}{\sqrt{N}}\right\} \approx 0.997.$$

This formula is very important to the Monte Carlo method. It gives us both the method of calculating m and the estimate of the error.

Let us indeed sample N values of the random variable ξ. Actually, determining a single value of the variables $\xi_1, \xi_2, ..., \xi_N$ is equivalent to determining N values of a single variable ξ, because all these random variables are identical, i.e. have identical distributions. Our formula shows that the arithmetic mean of these values will be approximately equal to m. The error of the approximation will most probably not exceed the value $3b/\sqrt{N}$ and it will approach zero if N increases.

As we already mentioned, the roulette is a device that could in principle be used to generate random numbers. See the figure below.



However, this way of generating random numbers would be much too slow, as for practical calculations we need millions of them.

Various algorithms implemented in computers allow to generate pseudo-random numbers of quite good *quality*. The quality can be checked by means of dedicated tests.

If those tests are satisfied, then the question about the difference between the pseudo-random numbers and the true random numbers, which is an ideal mathematical concept, is rather of a philosophical nature and we will not bother about it here.

Pseudo-random numbers

Various algorithms implemented in computers allow to generate pseudo-random numbers of quite good *quality*. The quality can be checked by means of dedicated tests.

If those tests are satisfied, then the question about the difference between the pseudo-random numbers and the true random numbers, which is an ideal mathematical concept, is rather of a philosophical nature and we will not bother about it here.

Exercise. Plot large a number of points with coordinates randomly generated with different random number generators, i.e. the intrinsic generator of the Fortran or C compiler and RANLUX, on the plane and see if no visible patterns occur.

Pseudo-random numbers

Various algorithms implemented in computers allow to generate pseudo-random numbers of quite good *quality*. The quality can be checked by means of dedicated tests.

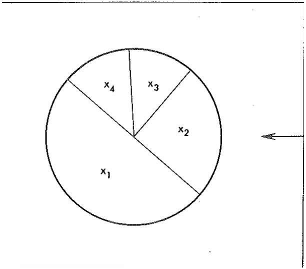
If those tests are satisfied, then the question about the difference between the pseudo-random numbers and the true random numbers, which is an ideal mathematical concept, is rather of a philosophical nature and we will not bother about it here.

Exercise. Plot large a number of points with coordinates randomly generated with different random number generators, i.e. the intrinsic generator of the Fortran or C compiler and RANLUX, on the plane and see if no visible patterns occur.

Solution of various problems requires simulation of on various random variables. We could design a dedicated roulette for each random variable. For example, the random variable with the distribution

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ 0.5 & 0.25 & 0.125 & 0.125 \end{pmatrix}$$

could be generated with the roulette depicted below.



However, this proved to be absolutely unnecessary, as the values of

Transformations of random numbers

any random variable can be obtained by transforming a continuous random variable γ uniformly distributed in the interval $(0, 1)$.

How to draw a discrete random variable with a given distribution? Suppose that we want to calculate the random variable

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

Let us divide the interval $(0, 1)$ into n intervals with lengths equal to $p_1, p_2, ..., p_n$ by choosing points with the coordinates $p_1, p_1 + p_2, p_1 + p_2 + p_3, ...$

Transformations of random numbers

any random variable can be obtained by transforming a continuous random variable γ uniformly distributed in the interval $(0, 1)$.

How to draw a discrete random variable with a given distribution?

Suppose that we want to calculate the random variable

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

Let us divide the interval $(0, 1)$ into n intervals with lengths equal to $p_1, p_2, \dots, p_n$ by choosing points with the coordinates $p_1, p_1 + p_2, p_1 + p_2 + p_3, \dots$.

Now, let us select γ. Since it is uniformly distributed within $(0, 1)$ the probability of γ lying within one of the intervals is given by

$$\begin{aligned} &\boldsymbol{P}\{0 < \gamma < p_1\} = p_1, \\ &\boldsymbol{P}\{p_1 < \gamma < p_1 + p_2\} = p_2, \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &\boldsymbol{P}\{p_1 + p_2 + \dots + p_{n-1} < \gamma < 1\} = p_n. \end{aligned}$$

Transformations of random numbers

any random variable can be obtained by transforming a continuous random variable γ uniformly distributed in the interval $(0, 1)$.

How to draw a discrete random variable with a given distribution?

Suppose that we want to calculate the random variable

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

Let us divide the interval $(0, 1)$ into n intervals with lengths equal to $p_1, p_2, ..., p_n$ by choosing points with the coordinates $p_1, p_1 + p_2, p_1 + p_2 + p_3, ...$

Now, let us select γ. Since it is uniformly distributed within $(0, 1)$ the probability of γ lying within one of the intervals is given by

$$\boldsymbol{P}\{0 < \gamma < p_1\} = p_1,$$

$$\boldsymbol{P}\{p_1 < \gamma < p_1 + p_2\} = p_2,$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\boldsymbol{P}\{p_1 + p_2 + ... + p_{n-1} < \gamma < 1\} = p_n.$$

Transformations of random numbers

According to our procedure we may assign $\xi = x_i$ if

$$p_1 + p_2 + ... + p_{i-1} < \gamma < p_1 + p_2 + ... + p_i$$

and the probability of this event is equal to p_i.

Exercise. Write a computer program which realizes this transformation.

Transformations of random numbers

According to our procedure we may assign $\xi = x_i$ if

$$p_1 + p_2 + ... + p_{i-1} < \gamma < p_1 + p_2 + ... + p_i$$

and the probability of this event is equal to p_i.
Exercise. Write a computer program which realizes this transformation.

Drawing a continuous random variable

Now, let us assume that we want to generate the values of the random variable ξ distributed in the interval (a, b) with the density $p(x)$.

We will prove that ξ can be found from the equation

$$\int_a^{\xi} p(x)dx = \gamma,$$

where γ is the random variable uniformly distributed in $(0, 1)$.

Let us define the function

$$y(x) = \int_a^{x} p(x)dx.$$

It is obvious that $y(a) = 0$ and $y(b) = 1$.

Drawing a continuous random variable

Now, let us assume that we want to generate the values of the random variable ξ distributed in the interval (a, b) with the density $p(x)$.
We will prove that ξ can be found from the equation

$$\int_a^{\xi} p(x)dx = \gamma,$$

where γ is the random variable uniformly distributed in $(0, 1)$.
Let us define the function

$$y(x) = \int_a^{x} p(x)dx.$$

It is obvious that $y(a) = 0$ and $y(b) = 1$.

Indeed

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

due to the normalization condition of the probability density $p(x)$.
The derivative

$$y'(x) = p(x) > 0,$$

which means that the function monotonically increases from 0 to 1, as in the following figure.

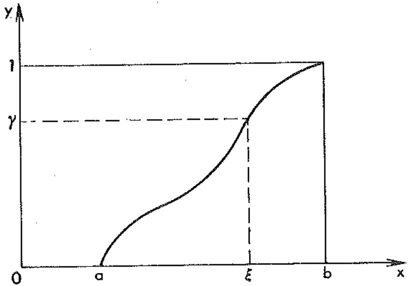
Indeed

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

due to the normalization condition of the probability density $p(x)$. The derivative

$$y'(x) = p(x) > 0,$$

which means that the function monotonically increases from 0 to 1, as in the following figure.



Any straight line $y = \gamma$, where $0 < \gamma < 1$, intersects the curve $y(x)$ at one and only one point whose abscissa is taken for the value of ξ. Thus, the equation

$$\int_a^{\xi} p(x)dx = \gamma$$

has always a unique solution.

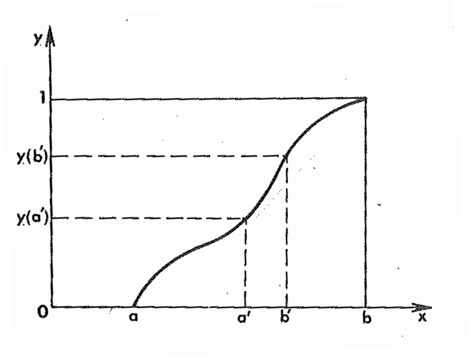
Drawing a continuous random variable

Now we select an arbitrary interval (a', b') within (a, b). The ordinates of the curve $y = y(x)$ satisfying the the inequalities

$$y(a') < y < y(b')$$

correspond to the points of the interval

$$a' < x < b'.$$



Consequently, if ξ belongs to the interval (a', b'), then γ belongs to the interval $(y(a'), y(b'))$, and vice versa. Hence

$$\boldsymbol{P}\{a' < \xi < b'\} = \boldsymbol{P}\{y(a') < \gamma < y(b')\}.$$

Since γ is uniformly distributed in $(0, 1)$, then

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

Thus

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

and this precisely means that the random variable ξ which is the root of the equation $\int_a^\xi p(x)dx = \gamma$ has the probability density $p(x)$.

Consequently, if ξ belongs to the interval (a', b'), then γ belongs to the interval $(y(a'), y(b'))$, and vice versa. Hence

$$\boldsymbol{P}\{a' < \xi < b'\} = \boldsymbol{P}\{y(a') < \gamma < y(b')\}.$$

Since γ is uniformly distributed in $(0, 1)$, then

$$\boldsymbol{P}\{y(a') < \gamma < y(b')\} = y(b') - y(a') = \int\limits_{a'}^{b'} p(x)dx.$$

Thus

$$\boldsymbol{P}\{a' < \xi < b'\} = \int\limits_{a'}^{b'} p(x)dx,$$

and this precisely means that the random variable ξ which is the root of the equation $\int_a^\xi p(x)dx = \gamma$ has the probability density $p(x)$.

Example. Find the random variable η uniformly distributed in the interval (a, b).

Solution. Let γ be the random variable uniformly distributed in the interval $(0, 1)$ and N be the normalization constant of the probability density of η and solve the equation

$$N \int_a^{\eta} dx = \gamma, \quad \text{where} \quad N \int_a^{b} dx = 1 \quad \Rightarrow \quad N = \frac{1}{b-a}.$$

The integration gives

$$N \, x|_a^{\eta} = \frac{\eta - a}{b - a} = \gamma \quad \Rightarrow \quad \eta = (b-a)\gamma + a.$$

We see, that if $\gamma = 0 \quad \Rightarrow \quad \eta = a$ and if $\gamma = 1 \quad \Rightarrow \quad \eta = b$, as it should be.

Example. Find the random variable η uniformly distributed in the interval (a, b).
Solution. Let γ be the random variable uniformly distributed in the interval $(0, 1)$ and N be the normalization constant of the probability density of η and solve the equation

$$N \int_a^\eta dx = \gamma, \quad \text{where} \quad N \int_a^b dx = 1 \quad \Rightarrow \quad N = \frac{1}{b-a}.$$

The integration gives

$$N\, x|_a^\eta = \frac{\eta - a}{b-a} = \gamma \quad \Rightarrow \quad \eta = (b-a)\gamma + a.$$

We see, that if $\gamma = 0 \quad \Rightarrow \quad \eta = a$ and if $\gamma = 1 \quad \Rightarrow \quad \eta = b$, as it should be.

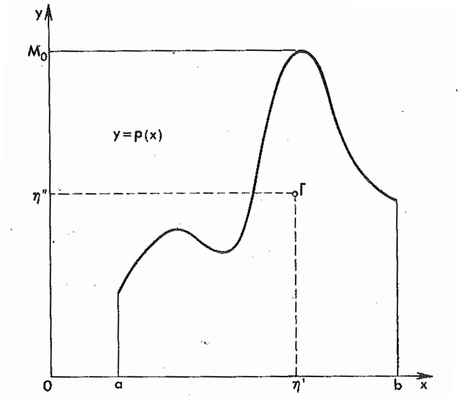
Exercise. Find the random variable ξ with the following densities $p(x)$

1. $x \in (x_0, x_{\max})$, with $x_0, x_{\max} > 0$, and
$$p(x) = \frac{N}{x},$$
2. $x \in (0, x_{\max})$, with $x_{\max} < 1$, and
$$p(x) = \frac{N}{1 - x},$$
3. $x \in (-c_0, c_0)$, with $0 < c_0 < 1$, and
$$p(x) = \frac{N}{1 - \beta^2 x^2}, \qquad \text{where} \quad \beta = 1 - \varepsilon, \quad \varepsilon > 0,$$
4. $x \in (-c_0, c_0)$, with $0 < c_0 < 1$, and
$$p(x) = \frac{N}{a + x}, \qquad \text{where} \quad a > 0, \quad a > c_0$$

Drawing a continuous random variable

It may happen that the equation $\int_a^\xi p(x)dx = \gamma$ cannot be solved analytically, as it is the case, e.g., for the normal distribution. Then the random variable can be drawn with von Neumann, method.

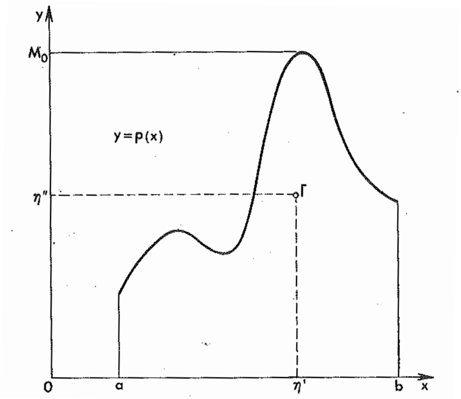
Let us assume that the random variable ξ is defined on a finite interval (a, b) and its density is limited $p(x) \leq M_0$, as in the figure below.



Drawing a continuous random variable

It may happen that the equation $\int_a^{\xi} p(x)dx = \gamma$ cannot be solved analytically, as it is the case, e.g., for the normal distribution. Then the random variable can be drawn with von Neumann, method.

Let us assume that the random variable ξ is defined on a finite interval (a, b) and its density is limited $p(x) \leq M_0$, as in the figure below.



von Neumann's method

The variable ξ may be drawn as follows:

1. Select two values γ' and γ'' of the random variable γ and generate a random point $\Gamma(\eta', \eta'')$ with coordinates

$$\eta' = (b - a)\gamma' + a, \qquad \eta'' = M_0 \gamma''.$$

2. If the point Γ lies below the curve $y = p(x)$ then assume $\xi = \eta'$.
3. If the point Γ lies above the curve $y = p(x)$ then reject the pair (γ', γ'').

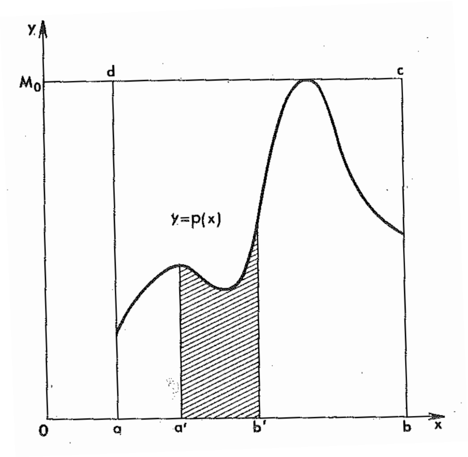
The variable ξ may be drawn as follows:

1. Select two values γ' and γ'' of the random variable γ and generate a random point $\Gamma(\eta', \eta'')$ with coordinates

$$\eta' = (b - a)\gamma' + a, \qquad \eta'' = M_0\gamma''.$$

2. If the point Γ lies below the curve $y = p(x)$ then assume $\xi = \eta'$.
3. If the point Γ lies above the curve $y = p(x)$ then reject the pair (γ', γ'').

Proof. The random point Γ is uniformly distributed in the rectangle abcd whose area is $M_0(b-a)$.



von Neumann's method

The probability for the point Γ to fall below the curve $y = p(x)$ and thus not to be rejected is equal to the ratio of the areas

$$\frac{\int_a^b p(x)dx}{M_0(b-a)} = \frac{1}{M_0(b-a)}.$$

The probability for the point Γ to fall below the curve $y = p(x)$ in the interval (a', b') is also equal to the ratio of areas

$$\frac{\int_{a'}^{b'} p(x)dx}{M_0(b-a)}.$$

von Neumann's method

The probability for the point Γ to fall below the curve $y = p(x)$ and thus not to be rejected is equal to the ratio of the areas

$$\frac{\int_a^b p(x)dx}{M_0(b-a)} = \frac{1}{M_0(b-a)}.$$

The probability for the point Γ to fall below the curve $y = p(x)$ in the interval (a', b') is also equal to the ratio of areas

$$\frac{\int_{a'}^{b'} p(x)dx}{M_0(b-a)}.$$

Hence, the fraction of ξ values comprised within the interval (a', b') among all selected values of ξ is equal to the ratio

$$\frac{\int_{a'}^{b'} p(x)dx}{M_0(b-a)} : \frac{1}{M_0(b-a)} = \int_{a'}^{b'} p(x)dx,$$

that is what was required to prove.

von Neumann's method

The probability for the point Γ to fall below the curve $y = p(x)$ and thus not to be rejected is equal to the ratio of the areas

$$\frac{\int_a^b p(x)dx}{M_0(b-a)} = \frac{1}{M_0(b-a)}.$$

The probability for the point Γ to fall below the curve $y = p(x)$ in the interval (a', b') is also equal to the ratio of areas

$$\frac{\int_{a'}^{b'} p(x)dx}{M_0(b-a)}.$$

Hence, the fraction of ξ values comprised within the interval (a', b') among all selected values of ξ is equal to the ratio

$$\frac{\int_{a'}^{b'} p(x)dx}{M_0(b-a)} : \frac{1}{M_0(b-a)} = \int_{a'}^{b'} p(x)dx,$$

that is what was required to prove.