1. Pseudo-random numbers

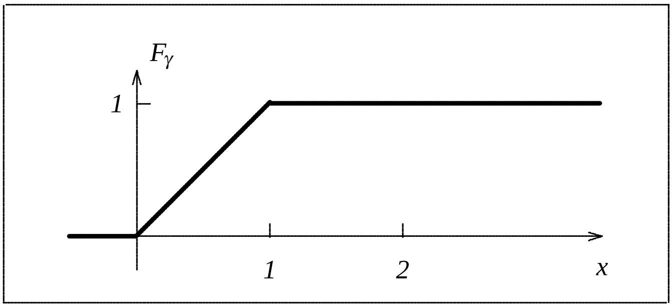
To simulate a random variable means to construct its numerical sample of an arbitrary large length. The basic idea is to represent the random variable in terms of one or more independent, uniformly distributed random variables. A random variable which is uniformly distributed in $[0,1]$ is called the random number. Therefore, the construction of a numerical sample of the original random variable is transformed to a generation of a sample of random numbers. In this way the random number has an exclusive importance in simulations. The first chapter is devoted to a study of theoretical background of simulations of random numbers and to testing quality of simulated sequences.

1.1. Random number and Monte Carlo simulation

For Monte Carlo simulations of random variables the central role is held by the random variable which is uniformly distributed in the interval $[0,1]$. Because of this significant role this random variable is called *random number* and it is denoted by γ. The distribution of random number is defined by the function:

$$F_\gamma(x) = \begin{cases} 0 & \text{for} \quad x \in (-\infty, 0), \\ x & \text{for} \quad x \in [0,1), \\ 1 & \text{for} \quad x \in [1,\infty), \end{cases}$$

and its graph is illustrated in Figure 1.1.



Distribution of random number

Figure 1.1.

The essence of Monte Carlo method can be explained by an example. Let us assume that a random variable is defined by $\xi = g(\gamma)$, where g is a function mapping the interval $[0,1]$ into real numbers (including $\pm\infty$). For instance, the function:

$$g(x) = \begin{cases} -\ln x & \text{for} \quad x \in (0,1], \\ \infty & \text{for} \quad x = 0, \end{cases}$$

has such properties. Our aim is to generate a sequence of numbers $y_1, y_2, \ldots$, which has the same properties as independent outcomes of the variable ξ. For this purpose a sequence of independent outcomes $\{\beta_k : k = 1,2,\ldots\}$ of the random number γ must be available. This is a sequence of numbers with values in the interval $[0,1]$. Then the sequence of numbers $\{y_k : k = 1,2,\ldots\}$, where $y_k = g(\beta_k)$, defines a sequence of independent outcomes of ξ. This simple numerical procedure is called a Monte Carlo simulation of random variable ξ. It can be split methodologically into three entities. These entities are called tasks.

S1) A random variable ξ is defined by its distribution $x \longrightarrow F_\xi(x)$. The first task of simulation is to find out a function g on $[0,1]$ such that

$$\xi = g(\gamma). \tag{1.1}$$

S2) The second task of simulation is the generation of a sequence $\{\beta_k : k = 1,2,\ldots\} \subset [0,1]$ which must have properties as a sequence of independent outcomes of the random number γ. Then the sequence of numbers $y_k = g(\beta_k)$ has properties as a sequence of independent outcomes of the random variable ξ.

S3) The third task is an analysis of the quality of simulation by using statistical tests.

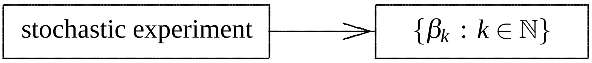
The procedure defined by Tasks S1) - S3) is called the Monte Carlo simulation of a random variable ξ. If the objects of simulations (such as random variables) are not specified in advance we prefer to use the terminology Monte Carlo method.

The described procedure can be easily generalized to the cases in which the random variable ξ is represented by a sequence of independent random numbers,

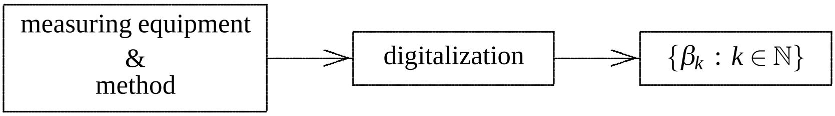
$$m = g_0(\gamma_0), \quad \xi = g_m(\gamma_1, \gamma_2, \ldots, \gamma_m), \tag{1.2}$$

where the function g_0 maps the interval $[0,1]$ into natural numbers $\mathbb{N}$, while the functions g_m map $[0,1]^m$ into $\mathbb{R}$. However, the idea of simulation is unchanged. In Task S1) one has to use (1.2) instead of (1.1) and in Task S2) the function $g(\cdot)$ must be replaced by pairs $g_0(\cdot), g_m(\cdots)$.

The first of defined tasks, i.e. the search for a function g of (1.1) or (1.2), is a genuine mathematical problem. The greatest part of analysis in this book is devoted to this problem. The second task is related directly to the stochastic experiment or outcomes. Schematically, this problem is illustrated by the following chart:

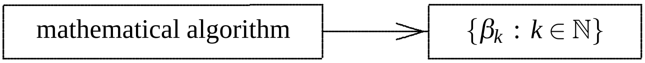


By the very nature of random variable the stochastic experiment is not numerically realizable by mathematical algorithms. Therefore the mathematical procedure defined by this chart is often replaced by another one,


measuring equipment & method
digitalization
$\{\beta_k : k \in \mathbb{N}\}$

which can be materialized. Such procedures are often used and finite sequences of numbers β_k are generated and saved in the form of tables or recorded on electronic carriers. Let us point out that this procedure is not mathematical because its essential part depends on objects of our material world.

We know that a sequence $\{\beta_k : k = 1,2,\ldots\}$ must have certain statistical properties implied by the characteristics of random number. Properties are verified by statistical tests. This fact gives us an opportunity to make an essential step towards generalizations. Our interest is not the origin of a sequence $\{\beta_k : k = 1,2,\ldots\}$. Our interest are statistical properties of this sequence which can be verified by using statistical tests. Hence, sequences generated by mathematical algorithms are also acceptable if they have the same properties as sequences of independent outcomes of random number. A tremendous advantage of such possibility is obvious since in a process of simulation we do not leave objects of mathematical nature:



A sequence $\{\beta_k : k = 1,2,\ldots\}$ generated by a mathematical algorithm and possessing properties of a sequence of independent outcomes of the random number, is called a sequence of *pseudo-random* numbers. Most of mathematical algorithms for generating pseudo-random numbers are based on the following one:

$$\beta_k = M\beta_{k-1} + \delta \ [\text{mod } 1], \quad k = 1,2,\ldots, \tag{1.3}$$

where β_0, δ and M are the parameters fulfilling the following constraints: $\beta_0 \in (0,1)$, $\delta \in \mathbb{R}$, while M is any natural number larger than 1. The algorithm (1.3) is called by various names, some of them are *multiplicative* and *congruental*.

A *generator of pseudo-random numbers* is any device which can generate a finite or infinite sequence of pseudo-random numbers. Thus tables and other electronic carriers of pseudo-random numbers are generators. A particular class of generators is made up by mathematical algorithms. They are briefly called algorithms as only mathematical algorithms are considered in this book. In order to avoid guessing what an algorithm might be we define it here with a sufficient generality to cover all the cases considered throughout this book. An algorithm is a function with a domain D and range R defined as follows. The set D consists of p ordered numbers (reals, integers or mixed), and R consists of infinite sequences of numbers (reals, integers).

1.1.1. Required properties of pseudo-random numbers

It is clear now that Monte Carlo methods deal with numbers, i.e. Monte Carlo methods make a part of numerical methods of Theory of probability. Passing from the world of random variables to the world of their outcomes, i.e. numbers, is always accompanied with certain undesirable features which must be excluded. For instance, in the introductory description of Monte Carlo simulation the notion of independent outcomes of the random number γ is used rather vaguely. An infinite sequence of independent outcomes of γ may have anomalous features failing to imitate statistical properties of the mathematical sample of γ. Such outcomes of γ are theoretically exceptional. Therefore, a correct formulation would be ”almost each sequence of independent outcomes of γ”. Unfortunately, such undesirable features occur also in the analysis of pseudo-random numbers. In the following it is assumed that repeated outcomes of a random variable, which are used in definitions, proofs, etc., have all the necessary properties possessed by the corresponding mathematical sample (see examples in Exercises 1.7).

A sequence of independent outcomes of a random variable ξ possesses statistical properties of the corresponding mathematical sample. These properties must be maintained by a sequence of simulated outcomes of ξ. Pseudo-random numbers at hand and which are used in an actual simulation of the given random variable may lack some of required properties of the corresponding mathematical sample. In a brief discussion that follows such deficiencies are analyzed.

Pseudo-random numbers $\{\beta_k : k = 1,2,\dots\}$ simulate a sequence of independent outcomes of the random number. Therefore, they must have the property which is described as follows. Let $I = [a,b] \subset [0,1]$ and let the amount of numbers $\beta_1,\beta_2,\dots,\beta_N$ falling into the interval I be denoted by $\nu_1(I,N)$. Apparently, $\nu_1(I,N)/N$ is the relative frequency of the considered finite sequence of numbers falling in I. Therefore, for any infinite sequence of pseudo-random numbers the following equality

$$\lim_{N\to\infty} \frac{\nu_1(I,N)}{N} = b - a \tag{1.4}$$

must be valid. This requirement follows from the basic characteristics of random number, $\mathbf{P}(\gamma \in I) = b - a$.

Let ξ be a random variable and $F(\cdot)$ be its distribution. Firstly we consider the case where ξ is represented by (1.1) and $F(g(z)) = z$ for each $z \in (0,1)$. Simulated outcomes are defined by $y_k = g(\beta_k)$. Analogously to $\nu_1(I,N)$ of (1.4) we also need the quantity $\rho((-\infty,x],N)$ defined as the amount of numbers $y_1,y_2,\ldots,y_N$ falling into the interval $(-\infty,x]$. Apparently, $\rho((-\infty,x],N) = \nu_1((0,F(x)],N)$. A sequence of numbers $y_k = g(\beta_k)$, simulating independent outcomes of ξ, must have the property:

$$\lim_{N\to\infty} \frac{\rho((-\infty,x],N)}{N} \;=\; F(x).$$

This property is always valid because of (1.4). Hence, in the case of representation (1.1) and $F \circ g = 1$, the basic statistical property (1.4) of pseudo-random numbers ensures the required statistical properties of the simulated sequence $\{y_k : k = 1,2,\ldots\}$.

However, Condition (1.4) is not the only one which must be fulfilled by a sequence of pseudo-random numbers. This assertion will become clear after an example of simulation.

Very often Monte Carlo methods are used in applications as an alternative approach to estimate probabilities of events. Monte Carlo methods can be advantageous whenever numerical evaluations of probabilities by deterministic methods represent a complex numerical procedure. The following simple example can illustrate this alternative approach. Let us consider the 5-dimensional normal random variable $(\xi_1,\xi_2,\ldots,\xi_5)$ with zero expectations $\mathbf{E}[\xi_k] = 0$, and the covariance matrix $c_{ij} = \mathbf{E}[\xi_i\xi_j]$ defined by:

$$c_{ij} \;=\; \frac{2}{\pi}\exp\left(-\frac{1}{5}|i-j|\right)\frac{\sin\left(\frac{\pi}{2}(i-j)\right)}{i\,-\,j}, \quad i,j = 1,2,\ldots,5. \tag{1.5}$$

To estimate the probability of event $\mathcal{K}_0 = \{\xi_1 < \xi_2+\xi_3+\xi_4 < \xi_5\}$, i.e. the number $\mathbf{P}(\mathcal{K}_0)$, we can use the definition of this probability in terms of 5-dimensional density. The corresponding expression is a 5-dimensional integral over $\mathbb{R}^5$. This expression is complex from the numerical point of view so that Monte Carlo simulations of the event and statistical estimates of its probability yield a more efficient method. Of course, any other event $\mathcal{K}$ defined by a relation among the random variables $\xi_1,\xi_2,\ldots,\xi_5$, can be simulated by Monte Carlo methods as well.

Suppose that the random variables ξ_i of the example are represented as $\xi_i = g_i(\gamma_1,\gamma_2,\ldots,\gamma_5)$, $i = 1,2,\ldots,5$, where g_i are functions from $(0,1)^5$ to $\mathbb{R}$ and $\gamma_1,\gamma_2,\ldots,\gamma_5$ are independent random numbers. Then the simulated sequence has the form

$$y_{5(k-1)+1}, y_{5(k-1)+2}, \ldots, y_{5k}, \quad k \;=\; 1,2,\ldots,$$

where

$$y_{5(k-1)+r} \;=\; g_r(\beta_{5(k-1)+1},\, \beta_{5(k-1)+2},\, \ldots,\, \beta_{5k}).$$

A sequence of numbers $y_{5(k-1)+r}, r = 1,2,\ldots,5$, simulating an event $\mathcal{K}$, can be used for an unbiased estimation of $\mathbf{P}(\mathcal{K})$ if the sequence of points

$$\mathbf{z}_k \;=\; \left(\beta_{5(k-1)+1},\, \beta_{5(k-1)+2},\, \ldots,\, \beta_{5k}\right) \;\in\; [0,1]^5, \quad k \;=\; 1,2,\ldots,$$

has a property which generalizes (1.4). Instead of interval I in (1.4) we consider a rectangle $I = [a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_5,b_5] \subset [0,1]^5$ with the volume $\mathrm{vol}\,(I) = \prod_1^5 (b_i - a_i)$. Among

N simulated points $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_N \in [0,1]^5$ there are $\nu_5(I,N)$ points in the set I. Then there must hold

$$\lim_{N\to\infty} \frac{\nu_5(I,N)}{N} = \mathrm{vol}\,(I). \tag{1.6}$$

A sequence of pseudo-random numbers and the corresponding sequence of derived points $\boldsymbol{z}_k \subset [0,1]^5$ may lack the property (1.6). The validity of (1.6) is to be considered as an additional property of the sequence $\{\beta_k : k = 1,2,\ldots\}$.

The equality (1.6) is defined for $\nu_d(I,N)$ with $d = 5$. No doubt it can be defined for any $d \in \mathbb{N}$. Let the function g_0 in (1.2) have its range equal to $\mathbb{N}$. Then the corresponding random variable ξ is simulated by a sequence of numbers y_k which are defined by $m+1$ subsequent numbers of $\{\beta_k : k = 1,2,\ldots\}$. In addition, the number m varies with k. The resulting sequence $\{y_k : k = 1,2,\ldots\}$ simulates independent outcomes of ξ iff

$$\lim_{N\to\infty} \frac{\nu_d(I,N)}{N} = \mathrm{vol}\,(I), \tag{1.7}$$

is valid for any $d \in \mathbb{N}$. In a simulation of sample paths of stochastic processes we are faced with the requirements (1.7) which must be valid for all d. In Chapters 3. – 7. certain stochastic processes are studied from the point of view of simulations. It follows that the sample paths of these processes are defined as functions of a countable number of independent random numbers. Simulations of such sample paths will be acceptable if the utilized sequence of pseudo-random numbers has the property (1.7) for any $d = 1,2,\ldots$.

Let $\mathcal{F}(d)$ be the algebra of events generated by d independent random numbers $\gamma_1, \ldots, \gamma_d$. For instance, the event $\{\gamma_1 < \gamma_2 < \cdots < \gamma_d\}$ is an element of that algebra. By using a sequence of pseudo-random numbers an event in $\mathcal{F}(d)$ can be simulated and its probability estimated by the corresponding statistics. This numerical process gives an acceptable result if the utilized sequence of pseudo-random numbers has the property defined by (1.7). Therefore, sequences of pseudo-random numbers are classified additionally with respect to the largest d for which (1.7) is valid.

1.1.2. Required properties of functions $g(\cdot)$ and $g_m(\cdots)$

Conditions (1.7) for all values of d are not the only ones to be satisfied in a successful simulation of random variables. Here we begin with an example demonstrating a possibility of an unwanted result of simulation of a random variable. Let $\{\beta_k, k \in \mathbb{N}\}$ be a sequence of pseudo-random numbers which is used in simulations of $\xi = g(\gamma)$. The function g in this example is defined as follows. Let $R = \{x_1, x_2, \ldots\} \subset [0,1]$ be an infinite sequence and

$$g(z) = \begin{cases} \frac{1}{2}z^2 & \text{for} \quad z \in R, \\ z^2 & \text{for} \quad z \in [0,1] \setminus R. \end{cases}$$

By using the sets $R(x) = \{x_i \in R : x_i \leq x\}$ we can write

$$F_\xi(x) = \mathbf{P}(\xi \leq x) = \int_{R(x)} dz + \int_{[0,\sqrt{x}]\setminus R(x)} dz = \sqrt{x}.$$

Hence, the random variable $\xi = g(\gamma)$ is equivalent to the random variable $\hat{\xi} = \gamma^2$ because both of them have the same distribution, $f(x) = \sqrt{x}$. If by some chance we have $R = \{\beta_k, k \in \mathbb{N}\}$ then the simulated sequence $y_k = g(\beta_k)$ will have statistical properties of the random variable $\frac{1}{2}\gamma^2$, contrary to our expectations. Similar non-regularities can be expected if R and $\{\beta_k, k \in \mathbb{N}\}$ have a sufficiently large intersection. In order to avoid such deviations in a process of simulations of random variables, the functions $g(\cdot)$, as well as $g_m(\cdots)$, must have certain properties. It is clear that any sequence of pseudo-random numbers is acceptable for simulation of $\xi = g(\gamma)$ if the function g has a finite number of discontinuities. Even if the number of discontinuities is not finite, the number of its accumulation points must be finite. In all such cases the integral of $g(\cdot)$ over $[0,1]$ can be defined as the Riemann integral. In our study of Monte Carlo simulations, the functions $g(\cdot), g_m(\cdots)$, etc., will belong to a particular class of functions to be defined at the beginning of next section. These functions produce no anomalous behavior of simulated sequences $\{y_k = g(\beta_k) : l \in \mathbb{N}\}$.

1.2. Equidistributed numbers in the Weyl sense

How to recognize that an infinite sequence of numbers with values in $[0,1]$ is a sequence of independent outcomes of the random number? An answer to this question comes from results of statistical tests. The number of statistics which is necessary for testing the mentioned sequence $\{\beta_k, k \in \mathbb{N}\}$ cannot be finite. Objectives of the present section are constrictions of statistics which are sufficient to answer the question whether or not a sequence of numbers $\{\beta_k, k \in \mathbb{N}\}$ has properties as a sequence of independent outcomes of the random number. A sequence of such statistics is called a complete sequence (or set) of statistics. To construct statistics we need the notion of piece-wise continuous function.

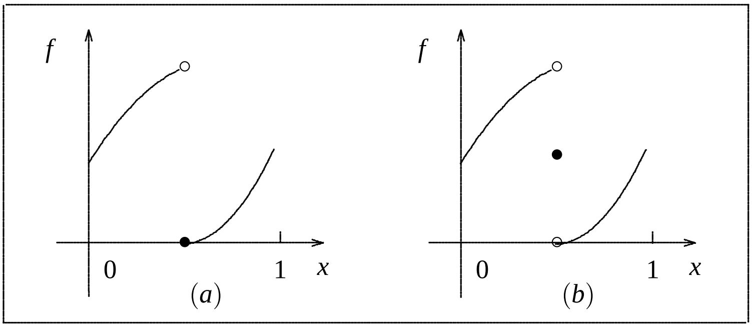
A real-valued function f from $D \subset \mathbb{R}^d$ into $\mathbb{R}$ is denoted by various symbols such as $f : D \to \mathbb{R}$, $D \ni x \to f(x) \in \mathbb{R}$, $x \to f(x)$, or simply $f(\cdot)$. The indicator of an interval $[a,b] \subset [0,1]$ is denoted as $\mathbb{1}_{[a,b]}$ and defined by:

$$\mathbb{1}_{[a,b]}(x) = \begin{cases} 1 & \text{for} \quad x \in [a,b], \\ 0 & \text{for} \quad x \notin [a,b]. \end{cases}$$

Analogously is defined the indicator of a set $D \subset \mathbb{R}^d$. This is a function with values $\mathbb{1}_D(\boldsymbol{x}) = 1$ for $\boldsymbol{x} \in D$ and $\mathbb{1}_D(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \notin D$.

Definition 1.1. (P.W.C. FUNCTION) *A function f on $\mathbb{R}^d$ is* piece-wise continuous with respect to the decomposition $\mathbb{R}^d = \cup_k D_k$ *if there exists a finite collection of L measurable and disjoint subsets $D_k \subset \mathbb{R}^d$, and bounded, uniformly continuous functions on $\mathbb{R}^d$, $f_1, f_2, \ldots, f_L$, such that $\mathbb{R}^d = \cup_{k=1}^L D_k$ and $f = \sum_{j=1}^L f_j \mathbb{1}_{D_j}$.*

If no misunderstanding can happen, the terminology ‘piece-wise continuous with respect. . . ’ is replaced by a simpler one ’piece-wise continuous’. A function f_D on $D \subset \mathbb{R}^d$ is *piece-wise continuous* if there exists a piece-wise continuous f on $\mathbb{R}^d$ such that $f_D = f|D$. Piece-wise constant functions are special cases of piece-wise continuous ones. For instance, indicators and their linear combinations are piece-wise constant functions. A piece-wise continuous function is often written as p.w.c. function.

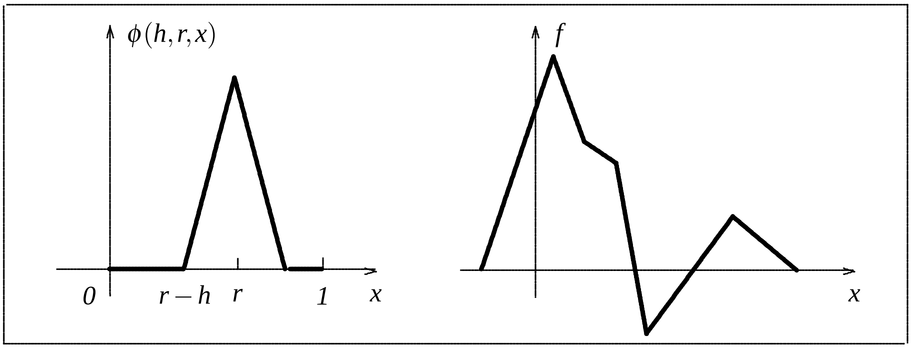


Piece-wise continuous functions on $[0,1]$.
$(a): \ f(0.5) = 0, \quad (b): \ f(0.5) = 0.5.$

Figure 1.2.

In the case of a function from an interval $[a,b]$ into $\mathbb{R}$, the following possibility is obtained. A function $f : [a,b] \to \mathbb{R}$ is piece-wise continuous on $[a,b]$, if there exists a finite number of points $a = x_0 < x_1 < x_2 < \cdots < x_m = b$, such that f is continuous on the intervals $(x_j, x_{j+1}), j = 0,1,\ldots,m-1$, and has finite values at the boundary points of these intervals. Two examples of piece-wise continuous functions are illustrated in Figure 1.2. It is interesting to notice that in case (b) the boundary values of f at $x = 1/2$ differ mutually and also from the value of f at $x = 1/2$.

A continuous function on the interval $[a,b] \subset \mathbb{R}$ is called piece-wise linear if there exists a finite number of points $x_0 = a, x_1, \ldots, x_N = b$ and a finite number of real numbers $f_0, f_1, \ldots, f_N$ such that $f(x_i) = f_i, i = 0,1,\ldots,N$, and f is linear on subintervals $[x_{i-1}, x_i]$. A continuous function $f : \mathbb{R} \mapsto \mathbb{R}$ is called piece-wise linear if there exists a bounded interval $[a,b] \subset \mathbb{R}$ such that $f|[a,b]$ is piece-wise linear and $f(x) = 0$ for $x \notin [a,b]$. An illustration of a piece-wise linear function is given in Figure 1.3.



A hat-function A piece-wise linear function

Figure 1.3.

A hat-function is an example of a piece-wise linear function. It is defined by

$$\phi(h,r,x) \;=\; \begin{cases} 1-\dfrac{|x-r|}{h} & \text{for} \quad |x-r| \le h, \\ 0 & \text{otherwise,} \end{cases}$$

and illustrated in Figure 1.3. We say that the hat-function $\phi(h,r,\cdot)$ has its center in r and support $[r-h,r+h]$.

A piece-wise linear function is a linear combination of hat-functions. In other words, the hat-functions span the linear space of piece-wise linear functions.

Let $f:[0,1]\to\mathbb{R}$ be a p.w.c. function defining the random variable $\eta = f(\gamma)$ and let $\{\gamma_1,\gamma_2,\dots\}$ be an infinite sample of independent random numbers. Statistics of the random variable η are defined by:

$$s(f,N) \;=\; \frac{1}{N}\sum_{k=1}^{N} f(\gamma_k), \quad N \in \mathbb{N}.$$

The expectation and variance of $s(f,N)$ have the well known expressions:

$$\begin{aligned} \mathbf{E}[s(f,N)] \;&=\; \mathbf{E}[\eta] \;=\; \int_0^1 f(x)dx, \\ \mathbf{Var}[s(f,N)] \;&= \frac{1}{N}\mathbf{Var}[\eta] \;=\; \frac{1}{N}\left[\int_0^1 f(x)^2dx \;-\; \mathbf{E}[\eta]^2\right]. \end{aligned}$$

One of possibilities of a complete sequence of statistics is defined by using all the indicators $f(x) = \mathbb{1}_{[a,b]}(x)$, for which $[a,b]$ have rational endpoints: $a_k = (k-1)2^{-n}, b_k = k2^{-n}, k = 1,2,\dots,2^n, n = 1,2,\dots$. The set of such intervals is countable and

$$s(\mathbb{1}_{[a,b]},N) \;=\; \frac{1}{N}\sum_{k=1}^{N} \mathbb{1}_{[a,b]}(\gamma_k), \quad \mathbf{E}\left[s(\mathbb{1}_{[a,b]},N)\right] \;=\; b-a. \tag{1.8}$$

Another sequence of a complete set of statistics is defined by statistical moments of all orders, $s(x^m,N) = N^{-1}\sum_k(\gamma_k)^m$, $\mathbf{E}[s(x^m,N)] = (m+1)^{-1}$. There are other important sequences of a complete set of statistics which will be described later. They are defined by using various bases of the linear space of piece-wise continuous functions $f:[0,1]\to\mathbb{R}$. After Herman Weyl, we define equidistributed numbers or pseudo-random numbers by using (1.8):

Definition 1.2. (OF EQUIDISTRIBUTED NUMBERS (H. WEYL)) *A sequence of numbers* $\{\beta_k : k \in \mathbb{N}\} \subset [0,1]$ *is called equidistributed in* $[0,1]$ *if for any subinterval* $[a,b] \subset [0,1]$ *the following equality is valid:*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} \mathbb{1}_{[a,b]}(\beta_k) \;=\; b-a. \tag{1.9}$$

A sequence of numbers $\{\beta_k : k \in \mathbb{N}\}$ fulfilling (1.9) for all subintervals $[a,b]$ is also called a sequence of pseudo-random numbers. Both names are used, so that in Theory of numbers the name equidistributed is used almost regularly while the name pseudo-random numbers is a preferable name in the applications of Monte Carlo methods. In the next section we define and study multiply equidistributed numbers. Therefore, by Definition 1.2 only simply equidistributed numbers are defined.

It is necessary to answer several important questions about equidistributed numbers or pseudo-random numbers, such as the existence of such numbers, how to test their properties, about convenient algorithms to generate such numbers, etc.

Let $\{\beta_k : k \in \mathbb{N}\}$ be a sequence of pseudo-random numbers. For any p.w.c. function f we define

$$\langle f \rangle \;=\; \lim_N \frac{1}{N} \sum_{k=1}^{N} f(\beta_k), \tag{1.10}$$

whenever the right hand side exists. The set of all such piece-wise continuous functions is denoted by $\mathcal{L}$. Apparently, the set $\mathcal{L}$ is a linear space, i.e. $\langle \alpha_1 f_1 + \alpha_2 f_2 \rangle = \alpha_1 \langle f_1 \rangle + \alpha_2 \langle f_2 \rangle$ for any pair of p.w.c. functions $f_1, f_2 \in \mathcal{L}$ and a pair of real numbers α_1, α_2. The linear space $\mathcal{L}$ contains indicators of all the intervals $I \subset [0,1]$.

Lemma 1.3. *Let f be a p.w.c. function on $[0,1]$ and let for each $\varepsilon > 0$ there exists two functions $f_-, f_+ \in \mathcal{L}$ such that*

$$\begin{gathered} f_-(x) \le f(x) \le f_+(x), \\ \sup_{x\in[0,1]} \,|\, f(x) - f_\pm(x) \,| < \varepsilon. \end{gathered} \tag{1.11}$$

Then $f \in \mathcal{L}$.

A proof follows directly from the following sequence of inequalities:

$$\langle f_- \rangle \le \liminf_N \frac{1}{N} \sum_{k=1}^{N} f(\beta_k) \le \limsup_N \frac{1}{N} \sum_{k=1}^{N} f(\beta_k) \le \langle f_+ \rangle.$$

Let $\{\beta_k : k \in \mathbb{N}\}$ be a sequence of numbers with values in $[0,1]$. Let us consider a p.w.c. function on $[0,1]$ for which the following equality

$$\langle f \rangle \;=\; \lim_N \frac{1}{N} \sum_{k=1}^{N} f(\beta_k) \;=\; \int_0^1 f(x)\,dx \tag{1.12}$$

is valid. All such functions span a linear space denoted by $\mathcal{K}$. The space $\mathcal{L}$ is defined by a sequence of pseudo-random numbers, while the space $\mathcal{K}$ is defined by a sequence of numbers with values in $[0,1]$ which are not necessarily pseudo-random numbers.

Lemma 1.4. *Let $\{\beta_k : k \in \mathbb{N}\}$ be a sequence of numbers in $[0,1]$ and let $\mathcal{K}$ be the linear space of all p.w.c. functions f on $[0,1]$ for which (1.12) is valid. If for any interval $I \subset [0,1]$, and any positive number $\varepsilon > 0$ there exists a pair $f_-, f_+ \in \mathcal{K}$ fulfilling the inequalities*

$$\begin{gathered} f_-(x) \le \mathbb{1}_I(x) \le f_+(x), \\ \langle f_+ \rangle - \langle f_- \rangle < \varepsilon, \end{gathered}$$

then $\{\beta_k : k \in \mathbb{N}\}$ is a sequence of pseudo-random numbers.

A proof follows from the following relations:

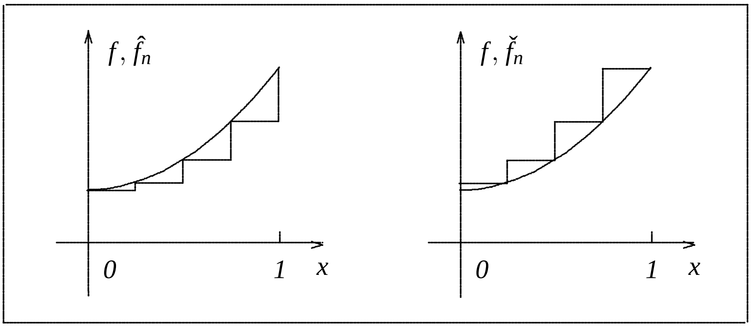
$$\begin{gathered}\tfrac{1}{N}\textstyle\sum_{k=1}^{N}\ f_{-}(\beta_k)\ \le\ \tfrac{1}{N}\textstyle\sum_{k=1}^{N}\ \mathbb{1}_I(\beta_k)\ \le\ \tfrac{1}{N}\textstyle\sum_{k=1}^{N}\ f_{+}(\beta_k),\\ b-a-\varepsilon\ \le\ \liminf_N\ \tfrac{1}{N}\textstyle\sum_{k=1}^{N}\ \mathbb{1}_I(\beta_k)\\ \le \limsup_N \tfrac{1}{N}\textstyle\sum_{k=1}^{N}\mathbb{1}_I(\beta_k)\ \le\ b-a+\varepsilon.\end{gathered}$$

Now we come to the first important result:

Theorem 1.5. (NECESSARY AND SUFFICIENT CONDITION ON EQUIDISTRIBUTION) *Let* $\{\beta_k : k\in\mathbb{N}\}\subset[0,1]$. *This sequence is equidistributed in* $[0,1]$ *iff the following equality*

$$\lim_N \frac{1}{N}\sum_{k=1}^{N} f(\beta_k)\ =\ \int_0^1 f(x)\,dx \tag{1.13}$$

is valid for any function f *which is continuous on* $[0,1]$ *and* $f(0)=f(1)$.



A continuous function f is approximated
by step-functions from below (left) and above (right)

Figure 1.4.

PROOF: *Necessity.* Let $\{\beta_k : k\in\mathbb{N}\}$ be eqidistributed in $[0,1]$ and generally $f(0)\neq f(1)$. The interval $[0,1]$ is divided into 2^n subintervals of equal length:

$$J(k)\ =\ \left[\frac{k-1}{2^n},\frac{k}{2^n}\right),\quad k=1,2,\dots,2^n-1,\quad J(2^n)\ =\ \left[1-\frac{1}{2^n},1\right].$$

Let $\underline{f_k}$ and $\overline{f_k}$ be the infimum and supremum of f on $J(k)$, respectively. The functions,

$$x\mapsto\ \hat{f}_n(x)\ =\ \sum_{k=1}^{2^n}\underline{f_k}\,\mathbb{1}_{J(k)}(x),\quad x\mapsto\ \check{f}_n(x)\ =\ \sum_{k=1}^{2^n}\overline{f_k}\,\mathbb{1}_{J(k)}(x),$$

are piece-wise constant functions approximating the function f from below and above as illustrated in Figure 1.4. These functions belong to $\mathcal{L}$ and:

$$\hat{f}_n(x) \ \leq f(x) \leq \ \check{f}_n(x),$$
$$\sup_{x\in[0,1]} \left\{ |f(x) - \check{f}_n(x)|, \quad |f(x) - \hat{f}_n(x)| \right\} \ < \ \varepsilon(n),$$

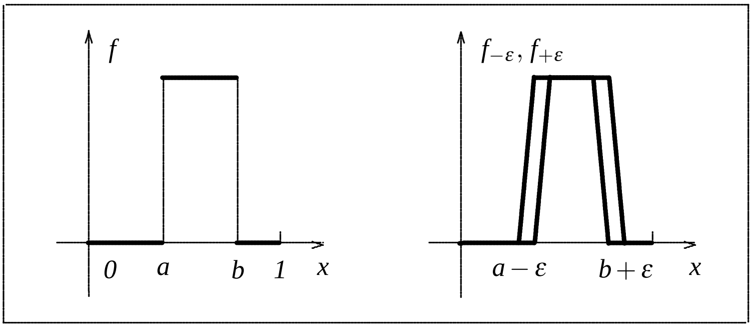
where

$$\lim_n \ \varepsilon(n) \ = \ 0.$$

Thus, by Lemma 1.3 $f \in \mathcal{L}$ and (1.13) is valid.

Sufficiency. Let us consider firstly the case $I = [a,b] \subset (0,1)$. For the chosen $\mathbb{1}_{[a,b]}$ and sufficiently small $\varepsilon > 0$ there are two trapezoidal functions $f_{-\varepsilon}$, $f_{+\varepsilon}$ as illustrated in Figure 1.5. They are continuous and have properties assumed in the theorem. Therefore, they belong to the space $\mathcal{K}$. The inequalities $\langle f_{-\varepsilon}\rangle < (b-a) < \langle f_{+\varepsilon}\rangle$ and $\langle f_{+\varepsilon}\rangle - \langle f_{-\varepsilon}\rangle < \varepsilon$ together with Lemma 1.4 imply the assertion. In case of $[0,b] \subset [0,1)$ or $[a,1] \subset (0,1]$ the same argument can be used with the function $f_{-\varepsilon}$ as before and the restriction $\tilde{f}_{+\varepsilon}|[0,1]$, where $\tilde{f}_{+\varepsilon}$ is the continuous periodic function with the period 1 extending the trapezoidal function $f_{+\varepsilon}$. □

A continuous function on $[0,1]$ can be represented by its Fourier series and the condition (1.13) can be derived from the corresponding conditions of trigonometric functions. In this way we come to the next important result:



An approximation of indicator by continuous
trapezoidal functions

Figure 1.5.

Theorem 1.6. (THE WEYL CRITERION) *Let* $\{\beta_k : k \in \mathbb{N}\} \subset [0,1]$ *be given. Then the equalities*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \exp\left(2\pi i m \beta_k\right) \ = \ 0, \tag{1.14}$$

for each $m \in \mathbb{N}$ *are valid iff* $\{\beta_k : k \in \mathbb{N}\}$ *is equidistributed.*

PROOF: *Necessity* follows from the previous theorem.

Sufficiency. A continuous function on $[0,1]$ with equal values at the end-points can be approximated arbitrarily well by a piece-wise linear function having also equal values at the end-points of $[0,1]$. Any piece-wise linear function f on $[0,1], f(0) = f(1)$, can be extended to a continuous and periodic function on $\mathbb{R}$ with the period equal to 1. This periodic function can be represented by its Fourier series converging absolutely to f. Piece-wise linear functions are finite combinations of hat-functions. Hence, in this proof, it is sufficient to consider a hat-function and its periodic extension for which the support has length less than 1.

Let $x \mapsto p(h,r,x)$ be a continuous function on $\mathbb{R}$ extending a hat-function $\phi(h,r,\cdot)$ (see Figure 1.3) by periodicity with the period 1. Its Fourier series has the form:

$$p(h,r,x) \;=\; h \,+\, \sum_{k=1}^{\infty} a_k \Big\{ \cos(2\pi kr)\cos(2\pi kx) \,+\, \sin(2\pi kr)\sin(2\pi kx) \Big\},$$

where

$$a_k \;=\; \frac{2}{h(\pi k)^2}\sin^2(\pi kh).$$

Because of the absolute convergence of the series the equalities (1.14) imply

$$\lim_N \frac{1}{N}\sum_{k=1}^{N} p(h,r,\beta_k) \;=\; h.$$

i.e. the equality (1.13) of Theorem 1.5. □

Let us define the quantities

$$W(\exp,m,N) \;=\; \frac{1}{N}\sum_{k=1}^{N} \exp\Big(2\pi i m\beta_k\Big) \tag{1.15}$$

in accordance with the previous theorem. Then the countable sequence of equalities

$$\lim_{N\to\infty} W(\exp,m,N) = 0, \quad \text{for} \quad m = 1,2,\ldots,$$

is called *the Weyl criterion.*

Corollary 1.7. *Let there be defined two sequences of numbers, x_k and $y_k = x_k + \alpha$, where $\alpha \in \mathbb{R}$. If the numbers $\beta_k = x_k$ $[\text{mod } 1]$ are equidistributed in $[0,1]$ then $\sigma_k = y_k$ $[\text{mod } 1]$ are also equidistributed.*

A proof of this result follows directly from the Weyl criterion.

Corollary 1.8. *Let t be irrational and let us consider the sequence of numbers*

$$\beta_k \;=\; tk \;[\text{mod } 1], \quad k = 1,2,\ldots. \tag{1.16}$$

Then β_k are equidistributed in $[0,1]$.