Uniformly Distributed Random Numbers

Generation of Uniformly Distributed Random Numbers

In the processes of statistical simulation random numbers uniformly distributed over the interval $(0,1)$ play an important role.

It has been proved that they can be obtained from transcendental numbers such as, for example, $\pi = 3.141\,592\,653\,\ldots$, by "cutting" the mantissa of this number into parts containing the definite amount of digits and considering every part of the mantissa as a random number: 0.141, 0.592, $0.653, \ldots$.

In micro-calculators the sequence of uniformly distributed numbers γ is often obtained with the aid of the recurrence formula

$$\gamma_{i+1} = M(37\gamma_i),$$

where the symbol $M(x)$ means mantissa of the number x.

Software packages for computers uses the similar idea.

Uniformly distributed random numbers can be given in package Wolfram Mathematica by using function RandomReal[ ].

These numbers are not certainly random in the literal sense of this word as they have been obtained by a certain algorithm in practical calculations, however, they work like "true" random numbers.

It is why it called ***pseudo-random*** numbers.

Their statistical regularity manifests itself in the fact that every sufficiently large set of such numbers covers the interval $(0,1)$ with the "flat layer".

Uniformity of distribution means that probability density for them does not depend on x: $w(x) = C$.

The numerical value of constant C is found by integration of this equality taking into account the normalization condition

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

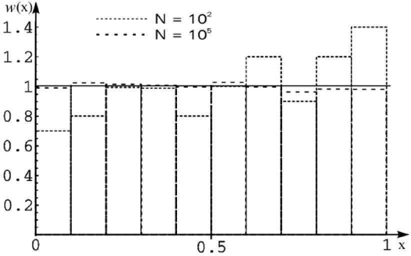
that gives $C = 1$.

The probability density was estimated by the formula

$$w(x) \approx \frac{N(x \in \Delta x)}{N\ \Delta x},$$

where $N(x \in \Delta x)$ - is the number of random points that got into Δx.

Next figure shows the distribution of random numbers γ on the interval $(0, 1)$.



From this figure one can see that random fluctuations for this case decreases with increase of N.

Simulation of Coin or Die Flipping and a Toss-up

.

The multiple toss of a "fair" coin generates the sequence of random numbers 0 (head) and 1 (tail) with equal probabilities.

One can obtain a statistically equivalent sequence of numbers of 0 and 1 on a computer if you generate random numbers γ uniformly distributed over the interval $(0,1)$, and write down 0 if $\gamma < 0.5$, and 1, if $\gamma > 0.5$.

The results of such simulation are shown on next slide.

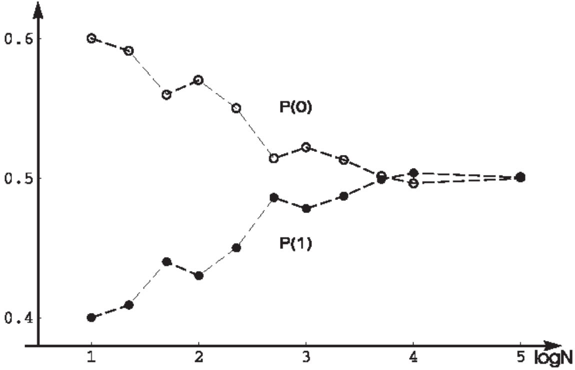


Figure: The results of the coin tossing simulation.

It is seen from the figure that with increasing of a number of trials N the ratios $N(''0'')/N$ and $N(''1'')/N$ lead to the same limit value equal to 1/2.

To simulate tossing the "fair" dice, the range $(0, 1)$ is divided into 6 equal parts and the interval index where the next random number γ occurs is taken as the number of sices on the die.

To simulate the toss up for n participants the range $(0, 1)$ is divided into n equal intervals and such a simulation resembles pictures of gambling at casinos using a roulette wheel.

The method described is easily generalized and can be used to solve problems where probabilities P_k of random outcomes A_k are not equal to each other.

In this case the range $(0, 1)$ is to be divided into parts proportional to the probabilities of corresponding outcomes P_k. The random interval number is determined by the random number γ.

In practical problems the following distributions are widely used.

The Binomial distribution with a finite number of intervals:

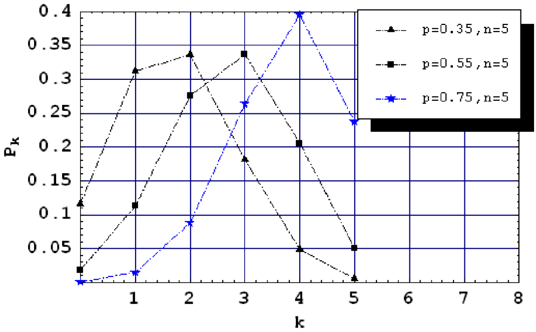
$$P_k = \frac{n!}{k!(n-k)!}\ p^k\ (1-p)^{n-k}, \quad 0 < p < 1,\ k = 0, 1, 2, ...n, \qquad (1)$$

n and p - are the distribution parameters.

The number of intervals, into which the range $(0, 1)$ is split here, equals $n+1$.

The maximum position of distribution is defined by the parameter value p. With increasing p maximum position is shifted to the right.

The view of the distribution is shown on next slide.



The Binomial distribution (1) defines the probability of the thing that in a set of n independent trials the event A occurring with the probability p and not occurring with the probability $1 - p$ will show up k times. Possible meanings of a random value k are equal to $0,\ 1,\ \ldots\ n$.

The geometrical distribution:

$$P_k = p\ (1-p)^{k-1}, \quad 0 < p < 1,\ k = 1, 2, \ldots \infty, \tag{2}$$

p is the distribution parameter. Here the number of intervals is infinite. The view of the distribution is given in Fig.2. With increasing the parameter p the function P_k decreases faster.

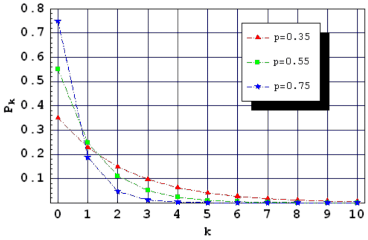


Figure: The geometrical distribution

The geometrical distribution (2) defines the probability of the thing that random event A, which occurs with the probability p and does not occur with the probability $1-p$, will show up at k trials. Possible quantities of the random value k are $1, 2, \ldots \infty$.

The Poisson distribution where the number of intervals is also infinite:

$$P_k = \frac{m^k}{k!}\ e^{-m}, \quad k = 0, 1, 2, ...\infty, \tag{3}$$

m is the distribution parameter. If $m \leq 1$, the function P_k monotonously decreases when k is growth. If $m > 1$ the distribution has the maximum that is shifted to the right with increasing m (next slide).

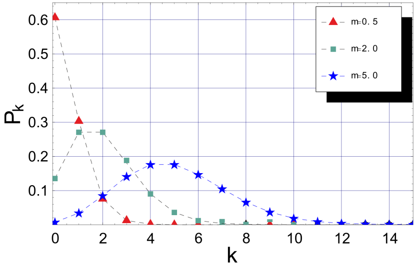


Figure: The Poisson distribution.

The examples of problems where a random value has a Poisson distribution is given on next lectures.

Dealing with statistical simulation of distributions (1)- (3) it is convenient to write them in the form of recurrence ratios

$$P_{k+1} = r_k P_k \ , \tag{4}$$

giving the expressions for P_0 or P_1 (depends on distribution) and r_k by independent formulas.

Simulation of Breeding Processes.

Let us consider a population made up of m living organisms. Each of them will produce a random number of descendants k $(k = 0, 1, 2, \dots)$ during its life. If probability to pruduce the k descendants are known, then the population development can be simulated.

Simulation is marking these probabilities P_k on the interval $(0, 1)$ and then a random number γ is generated for each organism and according to it an interval index is found which is taken as a random number of descendants k_i $(i = 1, 2, \dots m)$ of the given organism.

The sum

$$N = \sum_{i=1}^{m} k_i$$

gives the number of organisms in the next generation.

Repeating this procedure we get data for the further population development.

In this case a generation number plays the role of time.

The examples of the random function $N(n)$ showing the number of organisms in n generation are given on next fig.

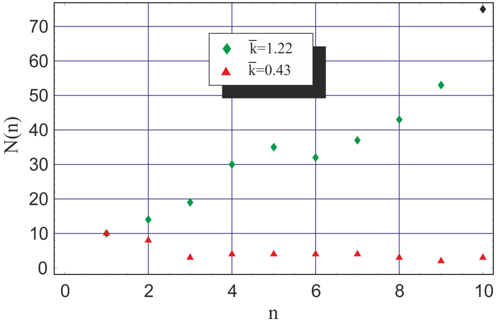


Figure: The dependence of the quantity of organisms in the population on the generation number.

According to calculations we suppose that the number of descendants of one organism is random and has the geometrical distribution (3).

The two curves in the figure correspond to two values of the parameter m determining a mean number of descendants from a single organism:

$$\bar{k} = \sum_{k=0}^{\infty} k \; P_k = m.$$

If $\bar{k} > 1$, the number of organisms of the population increases on the average and, if $\bar{k} < 1$, it decreases.

Analogous problems occur in other fields, for instance, in physics studying the reproduction of neutrons in a nuclear reactor.