Modelação de sistemas complexos

Department of Physics, University of Aveiro, Aveiro, Portugal

Alexander V. Goltsev

Project 1.

Probability. Binomial and Poisson distributions. The central limiting theorem.

1. Probability.

Task 1.1 A random number x can take the values $1,2,3, \dots M$ uniformly at random with the same probability 1/M. Prove **analytically** that the averaged value <x> of the random number *x* is

$$< x >= \frac{M+1}{2} \quad , \quad (1)$$

and the variance is

$$< (x - < x >)^2 >= \frac{M^2-1}{12}. \quad (2)$$

Prove **analytically** that the probability that *x* is smaller or equal *q*=60 is 0.6 at *M*=100.

Task 1.2 Prove Eqs. (1) and (2) **numerically**.

Algorithm.
Generate *N* times a random integer numbers x_i, (where i =1, 2, ...*N*) where $x_i = 1, or\ 2, or\ 3, \dots or\ M$ uniformly at random. Find the average value and the variance:

$$< x >= \frac{1}{N} \sum_{i=1}^{N} x_i ,$$

$$< (x - < x >)^2 >= \frac{1}{N} \sum_{i=1}^{N} (x_i - < x >)^2 .$$

The number of attempts N = 1000, 10^4, 10^6. Compare numerical results with the theoretical predictions Eqs. (1) and (2). Show that the mean value <x> , the variance $< (x - < x >)^2 >$, and the probability *p* tend to the theoretical predictions.

Task 1.3 Prove that if two random numbers *x*= rand(1) and *y*= rand(1) are uncorrelated random variables then the mean value of a random variable *z=xy* is <z>=<x><y>.

Algorithm. Generate *N* pairs of random numbers $x = \text{rand}(1)$ and $y = \text{rand}(1)$. Find their product, $z_i = x_i y_i$, i =1, 2, ...*N* . Then find the mean values <x>, <y>, and <z>:

$$< x >= \frac{1}{N} \sum_{i=1}^{N} x_i , \ \langle y \rangle = \frac{1}{N} \sum_{i=0}^{N} y_i, \ < z >= \frac{1}{N} \sum_{i=1}^{N} z_i$$

and compare with <z>=<x><y>. The number of attempts N = 100, 10^4, 10^6

2. Probability density.

1. Find numerically the probability density *p(x)* for a random variable x=rand(1), i.e. x is a random number in the interval $0 \le x < 1$. Plot *p(x).*

2. Find the mean value <x>, and the variance $\sigma^2 = \int_0^1 (x - < x >)^2 p(x) dx$. Compare with the theoretical values <x>=1/2, $\sigma^2 = 1/12$.
3. Find the probability density *g*(x) of a random variable *x* defined as $= \sqrt{rand(1)}$. Plot *g*(x) versus *x*. Compare with the theoretical result
$$g(x) = 2x.$$

Use the algorithm in the task 3 below.

Parameters: the number of iterations N =100, 10^3, 10^6. The width of the bin is $\Delta x = 0.005$.

3. The central limiting theorem.

Generate n random numbers z_i with the mean value <z> and the variance σ^2. For example, you can use the random number generator z = rand(1). A random number Y is defined as a sum of random numbers z_i,

$$Y = \frac{1}{n} \sum_{i=1}^{n} z_i \; .$$

Find the distribution function $P(Y)$ of the random numbers Y. Plot $P(Y)$. Shows that the mean value of Y is equal to <z>. Calculate the variance of Y,

$$\Lambda^2 = <(Y - <Y>)^2>,$$

and show that Λ^2 tends to σ^2 / n when the number of trials tends to infinity.

Algorithm.

1. Using z = rand(1), generate n random numbers z_i and calculate Y.
2. Repeat step 1 N times. You will get N random numbers Y_m, m =1,2, ….N.
4. Divide the interval [0,1] into bins of width Δy, i.e., $(k\Delta y, (k+1)\Delta y)$ where

$k = 0,1,..., k_{\max} = \dfrac{1}{\Delta y} - 1$. Calculate the numbers of Y_m in each bin ($k \le Y_m < k + 1$). Let this number in the bin k is $M(k)$. Calculate the distribution function $P(y_k)$ (the probability density),

$$P(y_k) = \frac{M(k)}{N \Delta y}.$$

Here we define a parameter $y_k = (k + 0.5)\Delta y$ (the coordinate of the bin k). Check the normalization condition $\sum_{k=0}^{k_{max}} P(y_k) = 1$.

5. Calculate the mean value and the variance of Y_m,

$$< Y >= \sum_{k=0}^{k_{\max}} y_k P(y_k) \Delta y,$$

$$\Lambda^2 = \sum_{k=0}^{k_{\max}} (y_k - < Y >)^2 P(y_k) \Delta y.$$

Parameters: n=10, 100, 1000. $\Delta y = 0.005$, the number of trials N=10^6.

4. Throwing balls.

There are M boxes and N balls. The balls are distributed among the boxes uniformly at random. Find the probability to find n balls in a given box. M=9, N=21, the number of trials K= 10^3, 10^4, 10^6.

Algorithm.

1. Let us measure the number of balls in the box with the number 3. Generate at random N integer numbers x_i, (i =1, 2, ...N). Each $x_i = 1, or\ 2, or\ 3, ... or\ M$ uniformly at random. from 1 to M.

2. Calculate how many times the number 3 appear among these N integer numbers x_i. Let it be n times.
3. Repeat steps 1 and 2 K times (K trials). You will get K random numbers $n_j, j = 1,2, \ldots K$.
4. Calculate the number of trials $N_{tr}(n)$ which give n darts in the box 3. Calculate the probability to find n balls in the box of number 3 as follows:

$$P(n) = \frac{N_{tr}(n)}{K}.$$

5. Plot $P(n)$ versus n, $n = 0,1,2,..N$. Compare between $P(n)$ and the binomial, Poisson and Gaussian distributions:

$$B(n,N) = C_n^N p^n (1-p)^{N-n},$$

$$P_n(Np) = \frac{(Np)^n}{n!} e^{-Np},$$

$$G(n) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(n-<n>)^2}{2\sigma^2}\right).$$

Here $p = 1/M$, $<n> = \sigma^2 = Np$. What distribution, the Poisson or Gaussian, is a better approximation to the exact binomial distribution?