

Monte Carlo simulation in Rstudio

Li Minghui

2020/3/25

Abstract

This report uses Monte Carlo method to make an approximate estimate of the standard normal function and compares it with the true value. And then, it repeats the experiment 100 times. Finally, drawing box plots of the 100 approximation errors at each t for each n . From this experiment, it's a conclusion that larger the number of computations, the greater the probability (according to the law of large numbers) of approximating the future events-provided the maximum-amount of known-data is incorporated into the model.

Introduction

Monte Carlo method is based on the principle of probability and statistics to simulate the formation process of things, so as to realize the characteristics of things and the law of change. The premise of this method is that the uncertain parameters can be described by probability distribution. When the solution of the problem is the probability of an event, or the mathematical expectation of a random variable, or the quantity related to the probability and mathematical expectation, the probability of the event, or the arithmetic mean value of several specific observation values of the random variable can be obtained by some experimental method, through which the solution of the problem can be obtained. This is the basic idea of Monte Carlo methods. Comparing the error between the Monte Carlo estimate and the true value, we can get a conclusion that there are errors between the results of Monte Carlo simulation and the true value, but as n increases, the errors decrease, and the simulation results are closer to the true value.

Math Equations

Consider approximation of the distribution function of $N(0, 1)$,

$$\Phi(t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$

by the Monte Carlo methods:

$$\hat{\Phi}(t) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq t)$$

where X_i 's are a random sample from $N(0, 1)$, and $I(\cdot)$ is the indicator function.

Table

Experiment with the approximation at $n \in \{10^2, 10^3, 10^4\}$ at $t \in \{0.0, 0.67, 0.84, 1.28, 1.65, 2.32, 2.58, 3.09, 3.72\}$ to form a table. The table includes the true value for comparison.

	t=0	t=0.67	t=0.84	t=1.28	t=1.65	t=2.32	t=2.58	t=3.09	t=3.72
n=100	0.5000	0.8000	0.8300	0.9100	0.9800	1.0000	1.0000	1.0000	1.0000
n=1000	0.5390	0.7460	0.8150	0.9040	0.9420	0.9910	0.9940	0.9990	1.0000
n=10000	0.4911	0.7518	0.7963	0.8986	0.9489	0.9903	0.9946	0.9992	1.0000
True Value	0.5000	0.7486	0.7995	0.8997	0.9505	0.9898	0.9951	0.9990	0.9999

Figures

Further, repeat the experiment 100 times. Draw box plots of the 100 approximation errors at each t using **ggplot2** for each n .

There are 9 pictures to show the result. For example, the figure of `boxploterror(error in t = 0, n = 100, 1000, 10000)` is that :

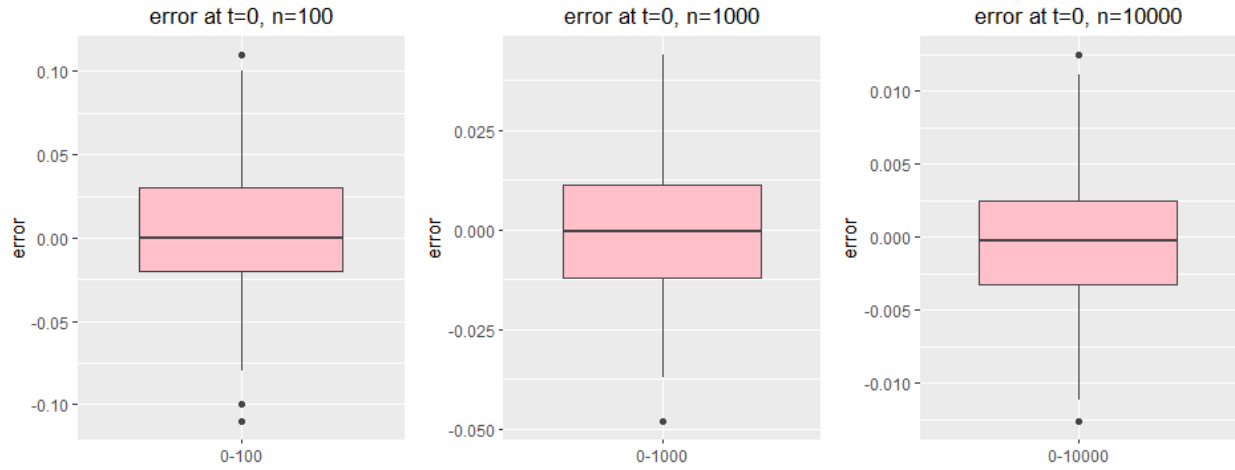


Figure 1: avatar

Conclusion

According to the above results, I come to the following conclusion: there are errors between the results of Monte Carlo simulation and the true value, but larger the number of computations, the greater the probability (according to the law of large numbers) of approximating the future events-provided the maximum-amount of known-data is incorporated into the model.