Approximation of the Standard Normal Distribution

Yang Jie

2020/3/26

Abstract

The Monte Carlo methods are used to calculate the approximate value of the standard normal distribution, and the box plots show the error of the method intuitively.

Math Equations

The equation for the standard normal distribution $N(0,1)$ is

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

The equation for the Monte Carlo methods is

$$\hat{\Phi}(t) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq t) \tag{2}$$

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

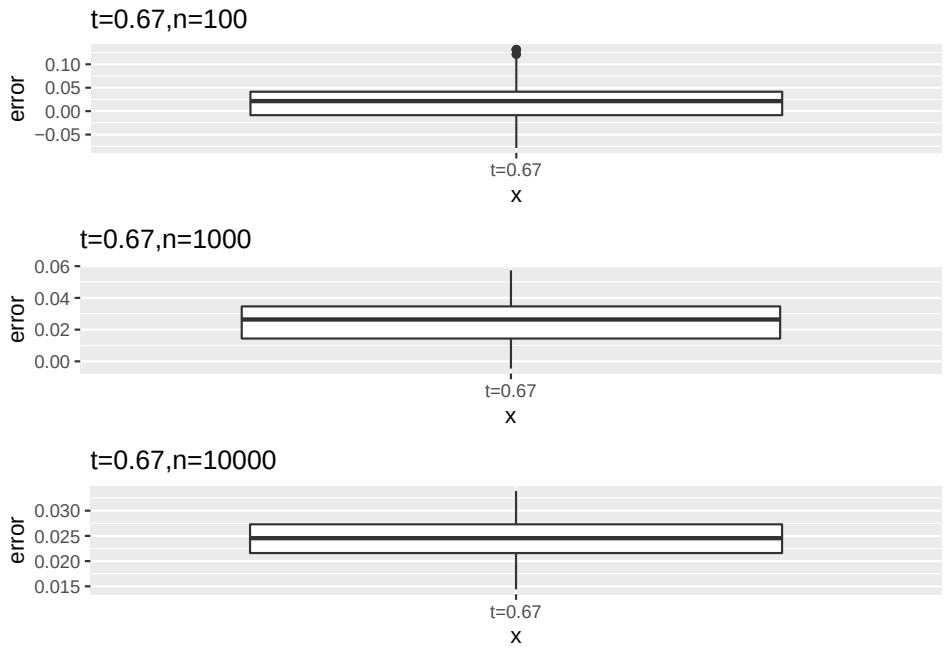
Table

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 approximation results are shown in the table.And it 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.Due to limited space, this article only lists the picture named "t=0.67".



Code Chunk

The following code is the main idea to draw box plots.

```r
library(ggplot2)
library(lattice)
library(plyr)
library(Rmisc)

t=c(0, 0.67, 0.84, 1.28, 1.65, 2.32, 2.58, 3.09, 3.72)
y=c(0.5000,0.7486,0.7995,0.8997,0.9505,0.9898,0.9951,0.9990,0.9999 )
n=c(100,1000,10000)

total=matrix(0,100,3)

for (j in 1:3) {
 err=rep(0,100)
 for (i in 1:100){
  a=rnorm(n[j],0,1)
  ypre=mean(a<y[2])
  err[i]=ypre-y[2]
  p=data.frame(err)
 }
 total[,j]=err
}
```


Summary

Get approximation of the distribution function of$N(0,1)$by the Monte Carlo methods.And there are still some difficulties in programming with R , especially on loop statements. Logic must be clear when programming.