

# An Exercise of R Markdown

Chen Xinglin

2019/6/13

ABSTRACT: This experiment simulates the probability values of a series of random numbers obeying the standard normal distribution under specific values by Monte Carlo simulation, and compares them with the true probability values. Subsequently, the experiment was repeated 100 times, and a series of errors were obtained. The boxes of error distribution under different sample sizes and different values were drawn respectively. The report is output by R Markdown.

## 1 Experimental Requirements

Consider a Standard Normal Distribution  $X \sim N(0, 1)$  Monte Carlo method is used to simulate the probability of the standard normal distribution at a given point (t) when the sample size is 100, 1000 and 10000, respectively. The approximate formula is frequency approximation.  $\hat{Phi}(t) = \frac{1}{n} \sum_{i=1}^n I(x_i \leq t)$

## 2 Experimental Process

### 2.1 Monte Carlo simulates probability value

The standard normal distribution with sample size of 100, 1000 and 10000 is generated, and the probability values under different T values are simulated. The codes and results are as follows:

```
n = c(100,1000,10000)
t = c(0.0,0.67,0.84,1.28,1.65,2.32,2.58,3.09,3.72)
data1 = data.frame()
for (i in n){
  p_e_list = c()
  s = 1
  for (j in t){
    x = rnorm(i)
    p_e = length(x[x<=j])/i
    p_e_list[s]=p_e
    s = s+1
  }
  data1<- rbind(data1, p_e_list)
}
p_t = pnorm(t)
data1 <-rbind(data1,p_t)
colnames(data1) <- t
rownames(data1) <- c("p_100","p_1000","p_10000","p_true")
data1
```

Table 1 The Results

tvalues	0.0	0.67	0.84	1.28	1.65	2.32	2.58	3.09	3.72
p_100	0.490	0.800	0.760	0.920	0.960	0.990	1.000	1.000	1.000

tvalues	0.0	0.67	0.84	1.28	1.65	2.32	2.58	3.09	3.72
p_1000	0.511	0.759	0.797	0.896	0.949	0.988	0.997	1.000	1.000
p_10000	0.491	0.749	0.800	0.896	0.950	0.988	0.995	0.999	1.000
p_true	0.500	0.749	0.800	0.900	0.951	0.990	0.995	0.999	1.000

## 2.2 Repeat the experiment 100 times and the error boxes

When the sample size are 100, 1000 and 10000, repeat the experiment 100 times, and get the simulation probability and the real probability corresponding to different t values, then calculate the error, which is expressed by the difference between the real probability and the simulation probability. This report lists the results when the sample size is 10000, the same is true when the sample size is 100 and 1000.

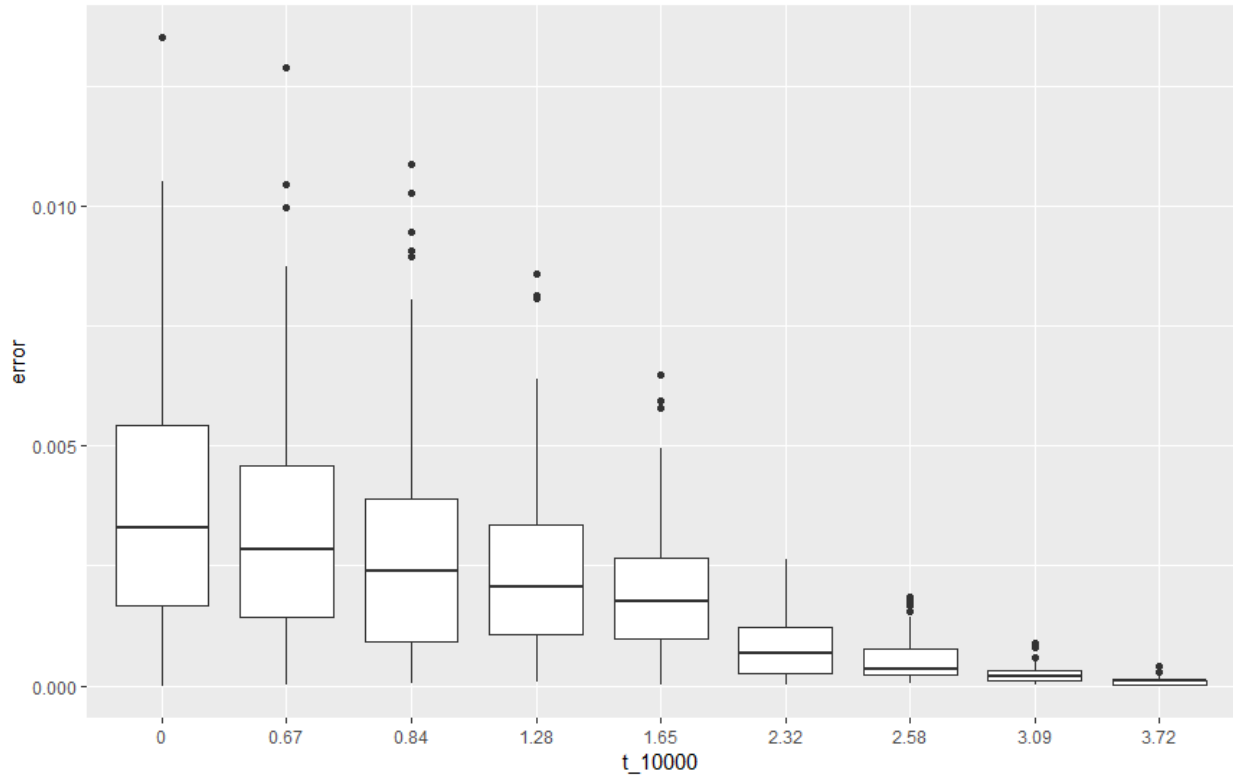


Figure 1 The box plots when n=10000

## 3 Conclusion

In this report, Monte Carlo method is used to simulate the fitting probability and real probability under different cost and different T values, and compare them. Then we repeat 100 experiments to get the error boxes. As the sample size increases, the error decreases. It is proved that the frequency converges in probability when the sample size is large enough.