Approximation of the Distribution

2020/3/25

Abstract

This report fits the standard normal distribution by the Monte Carlo methods and compare the approximation of the distribution under different ‘t’ and ‘n’ values.

Introduction

This report fits approximation of the distribution fuction of $N(0,1)$ by the Monte Carlo methods,including three parts.

The first part, the truth value of each ‘t’ value is calculated according to the cumulative density function of the standard normal distribution.

Secondly, the fitting degree of monte carlo method is compared under different ‘t’ and ‘n’ values. the definite integral in the experiment with the approval 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.

Finally, repeat the experiment 100 times and draw box plots of the 100 approximation errors at each t for each n.

Math Equations

To solve this problem, I consider approximation the distribution function of $N(0,1)$,

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

by the Monte Carlo methods:

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

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

Table

I 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 resulting table is shown in table @ref(tab:norm).

(ref:norm) The table of approximation results.

```
n=c(100,1000,10000)
t=c(0,0.67,0.84,1.28,1.65,2.32,2.58,3.09,3.72)
p=matrix(0,nrow=3,ncol=9)
#a=0
for (i in 1:3)
  for(j in 1:9){
    num=rnorm(n[i],0,1)
    p[i,j]=mean(num<t[j])
  }
rownames(p)<-n
colnames(p)<-t
real_p<-c(pnorm(0),pnorm(0.67),pnorm(0.84),pnorm(1.28),pnorm(1.65),
          pnorm(2.32),pnorm(2.58),pnorm(3.09),pnorm(3.72))
p<-rbind(p,real_p)
p<-round(p,digits=3)

library(knitr)
```

```
## Warning: package 'knitr' was built under R version 3.6.3
```

```
library(magrittr)
```

```
## Warning: package 'magrittr' was built under R version 3.6.3
```

```
library(kableExtra)
```

```
## Warning: package 'kableExtra' was built under R version 3.6.3
```

```
kable(p, booktabs=TRUE, caption='(ref:norm)') %>%
  kable_styling(bootstrap_options = "striped",full_width = F) %>%
 column_spec(1,bold=T)
```

Table 1: (ref:norm)

	0	0.67	0.84	1.28	1.65	2.32	2.58	3.09	3.72
100	0.550	0.760	0.880	0.870	0.930	1.000	1.000	1.000	1.000
1000	0.495	0.770	0.794	0.906	0.952	0.991	0.995	0.999	0.999
10000	0.499	0.750	0.799	0.898	0.955	0.990	0.996	0.999	1.000
real_p	0.500	0.749	0.800	0.900	0.951	0.990	0.995	0.999	1.000

The Boxplots of Errors

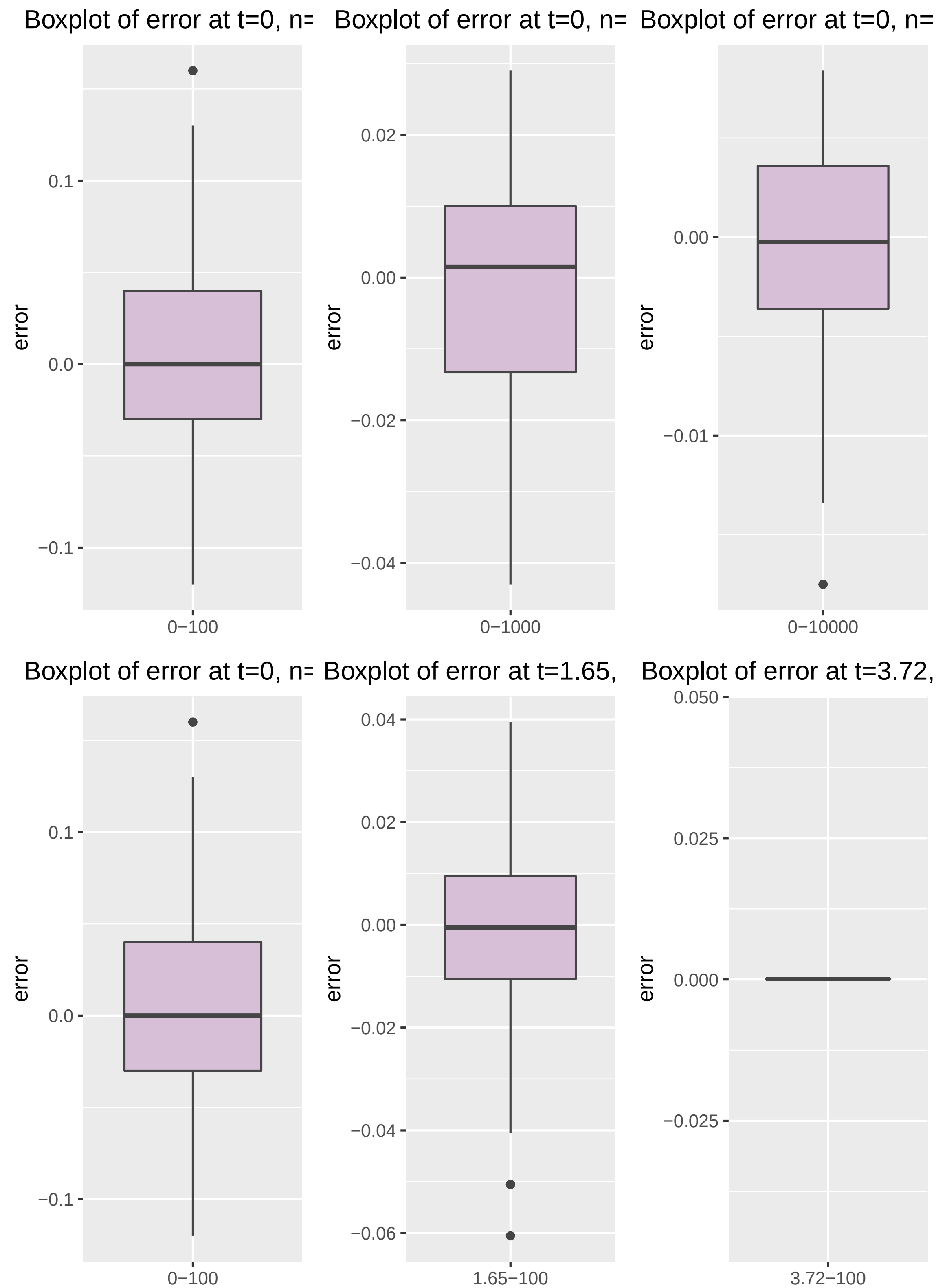
Further, repeat the experiment 100 times. Draw box plots of the 100 approximation errors at each t using **ggplot2** for each n.

(ref:cap-error) The boxplots of errors

```
## [1] 1
## [1] 2
## [1] 3
## [1] 4
## [1] 5
## [1] 6
## [1] 7
## [1] 8
## [1] 9
## [1] 10
## [1] 11
## [1] 12
## [1] 13
## [1] 14
## [1] 15
## [1] 16
## [1] 17
## [1] 18
## [1] 19
## [1] 20
## [1] 21
## [1] 22
## [1] 23
## [1] 24
## [1] 25
## [1] 26
## [1] 27
```

```
## Warning: package 'lattice' was built under R version 3.6.3
```

```
## Warning: package 'Rmisc' was built under R version 3.6.3
```


Boxplot of error at t=0, n=
Boxplot of error at t=0, n=
Boxplot of error at t=0, n=
error
0.1
0.0
−0.1
0−100
0.02
0.00
−0.02
−0.04
0−1000
0.00
−0.01
0−10000
Boxplot of error at t=0, n=
Boxplot of error at t=1.65,
Boxplot of error at t=3.72,
0.1
0.0
−0.1
0−100
0.04
0.02
0.00
−0.02
−0.04
−0.06
1.65−100
0.050
0.025
0.000
−0.025
3.72−100

Conclusion

From the table and boxplots of errors,we can conclude that as n gets bigger and bigger,approximation of the distribution function is getting closer and closer to the truth value.

R Code Reference

Refer to the code of other classmates in the process of compiling the code. [Reference website https://github.com/data-science-in-action/03-practicing-r-markdown-fanngguofang]

Finally,I would like to thank teachers and classmates.Through this course, I learned lots of skills and knowledge.