International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2020; 5(1): 64-67

www.mathsjournal.com
Received: 25-11-2019
Accepted: 27-12-2019

Xintong Yang
School of Mathematics and Statistics, Northeastern University at Qinhuangdao, Qinhuangdao, China

Corresponding Author:
Xintong Yang
School of Mathematics and Statistics, Northeastern University at Qinhuangdao, Qinhuangdao, China

The research on the characteristics of approximation process of chi-square distribution function based on MATLAB

Xintong Yang

Abstract
Chi-square distribution is asymptotically normal, which has been proved by predecessors. However, there is no specific description of this approximation process. Based on the previous research, this paper analyzed the specific-characteristics of chi-square distribution further by using MATLAB software. The results indicated that the chi-square distribution moves from the left to the right to gradually approach the corresponding normal distribution as its maximum values, skewness values, and kurtosis values decreasing.

Keywords: Chi-square, normal, distribution, approximation process, MATLAB

1. Introduction

The chi-square distribution is very important and valuable in probability statistics. So far, many people have researched on it in order to understand and apply it better [1-8]. According to the previous researches, the chi-square distribution has asymptotic normality, that is, when its degree of freedom tends to be infinite, this distribution gradually approaches the normal distribution [9-12]. This is a beautiful result undoubtedly. However, there is a problem still unclear: How does it gradually approach the normal distribution? What are the characteristics of this process? The predecessors did not give relevant explanations yet, therefore, this paper intends to research the characteristics of the process in which the chi-square distribution approaches the normal distribution based on previous studies by using the MATLAB software.

2. The Characteristics of Chi-square Distribution Approaching to Normal Distribution

As for the chi-square distribution, the following conclusions have been obtained in previous researches: When the degree of freedom n tends to be infinite, the chi-square distribution gradually becomes a normal distribution $N(n,2n)$ [10]. However, predecessors have not proposed what the characteristics of this approximation process has. To figure this out, we applied MATLAB software to analyze it. The methods and processes we used are as follows:

We let the degrees of freedom of the chi-square distribution increasing from 1 to 100, then calculated the maximum values, skewness values, and kurtosis values of each chi-square distribution function and the corresponding normal distribution, and plotted their function images.

The codes entered in MATLAB software are as follows

```
Clear all; clc;
x=-1:0.1:130;
n=linspace (5,100,10);
For i=1:10
A (i,:)=chi2pdf(x,n(i));
B (i,:)=normpdf(x,n(i),sqrt(2*n(i)));
MAXA (1,i)=max(A(i,:));
MAXA (2,i)=((find(A(i,:)==MAXA(1,i))-10))*0.1;
MAXB (1,i)=max(B(i,:));
```

```
MAXB(2,i)=((find(B(i,:)==MAXB(1,i))-10))*0.1;
SKEWA (i)=skewness(A(i,:));
KURTA (i)=kurtosis(A(i,:));
SKEWB (i)=skewness(B(i,:));
KURTB (i)=kurtosis(B(i,:));
End
Plot(x,A,'color','r','linewidth',1.8);
Hold on;
Plot(x,B)
Axis ([0 130 0 0.16])
MAXA
MAXB
SKEWA
SKEWB
KURTA
KURTB
Grid on
Xlabel ('x');
Ylabel ('y');
Legend ('n from 5 to 100')
```

After calculation, the function image obtained is shown in Figure 1, and the maximum values, kurtosis values, and skewness values are shown in Tables 1, 2, and 3 respectively.

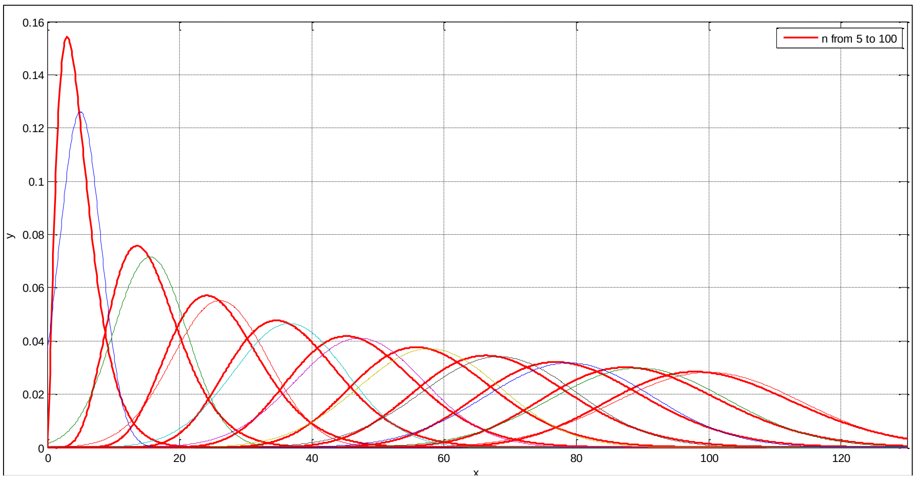

Fig 1: The function images of Chi-squared distribution and normal distribution with the degree of freedom increasing

It can be seen from Figure 1 that: (1) The images of probability density function of chi-square distribution and normal distribution are similar roughly; (2) When the degree of freedom *n* increases gradually, the probability density image of chi-square distribution gradually becomes low and wide; (3) When the degree of freedom *n* increases gradually, the function image moves to the right. It can be concluded from these changes that chi-square distribution approaches the corresponding normal distribution from the left.

Table 1: The coordinates of the maximum values point of the chi-square and normal distributions

MAXchi2_x	**3.1**	**13.7**	**24.2**	**34.8**	**45.3**	**55.9**	**66.4**	**77**	**87.5**	**98.1**
MAXchi2_y	0.1542	0.0757	0.0571	0.0477	0.0418	0.0377	0.0345	0.0321	0.0301	0.0284
MAXnorm_x	5.1	15.7	26.2	36.8	47.3	57.9	68.4	79	89.5	100.1
MAXnorm_y	0.1261	0.0715	0.0552	0.0466	0.0411	0.0371	0.0341	0.0318	0.0289	0.0282

As can be seen from the data in Table 1, when the degree of freedom *n* increases, (1) The maximum values of the probability density function of the chi-square distribution and its corresponding normal distribution gradually decrease, and the maximum point gradually moves to the right; (2) The maximum values point of the square distribution is always to the left of the maximum values point of the normal distribution, and the difference of their abscissasis a fixed value of 2; (3) The maximum values of the two distributions gradually decrease, and their differences also gradually decrease. Therefore, the maximum values point of the chi-square distribution and the corresponding normal distribution function approaches as the degree of freedom increases, which is mainly reflected in the vertical direction.

Table 2: The skewness values of the chi-square and normal distributions

SKEW_chi2	**4.4073**	**2.5659**	**2.078**	**1.7884**	**1.5837**	**1.4256**	**1.297**	**1.1884**	**1.0925**	**1.0041**
SKEW_norm	3.5864	2.4601	2.0213	1.75	1.5547	1.4023	1.2775	1.1718	1.0793	0.9922

Table 3: The kurtosis values of the chi-square and normal distributions

KURT_chi2	**18.8042**	**8.332**	**5.9305**	**4.7481**	**4.0216**	**3.5232**	**3.158**	**2.878**	**2.6533**	**2.4547**
KURT_norm	14.9209	7.7569	5.679	4.6014	3.9236	3.4525	3.1043	2.8362	2.6221	2.4329

Tables 2 and 3 are the skewness values and kurtosis values of the chi-square and normal distributions. It can be seen from the data of these two tables that as the degree of freedom *n* increases, the skewness values and kurtosis values of the chi-square distribution decrease gradually, so that they approach the corresponding skewness values and kurtosis values of the normal distribution gradually.

3. The Relationship between the Chi-square Distribution and the Standard Normal Distribution

When the above conclusion was proved by the predecessors, it was also proved that: Supposing the probability density function of the chi-square distribution with degree of freedom n is $f(x)$, $\zeta(n)$ distribution is recorded as the one of the probability density function $\sqrt{2n}\,f(n+\sqrt{2n}\,x)$, then when n tends to infinity, the $\zeta(n)$ distribution tends to the standard normal distribution [10]. So what are the characteristics of this approximation process? We also applied MATLAB software to figure it out. The methods and processes we used are as follows:

When the degree of freedom n increases from 1 to 300, we used MATLAB software to plot multiple probability density function images of $\zeta(n)$ distribution and standard normal distribution, and calculate their maximum values, skewness values, and kurtosis values. So as to analyze the characteristics of this process.

The codes entered in MATLAB are as follows

```
Clear all;clc;
x=-350:0.1:350;
n=linspace (1,300,20);
For i=1:20
A (i,:)=(chi2pdf(x,n(i)))*sqrt(2*n(i));
B (i,:)=(x-n(i))/sqrt(2*n(i));
Plot (B(i,:),A(i,:));hold on;
Axis ([-4 4 0 0.5]);
MAX (1,i)=max(A(i,:));
MAX(2,i)=(((find(A(i,:)==MAX(1,i))-3500))*0.1-
n(i))/sqrt(2*n(i));
SKEW (i)=skewness(A(i,:));
KURT (i)=kurtosis(A(i,:));
End
MAX
SKEW
KURT
y=normpdf(x,0,1);
Plot (x,y,'color','r','linewidth',2.2)
Grid on
Xlabel ('x')
Ylabel ('y')
Legend ('n from 1 to 300')
```

After calculation, the images we obtained are shown in Figure 2, and the maximum values, kurtosis values and skewness values are shown in Tables 4, 5 and 6.

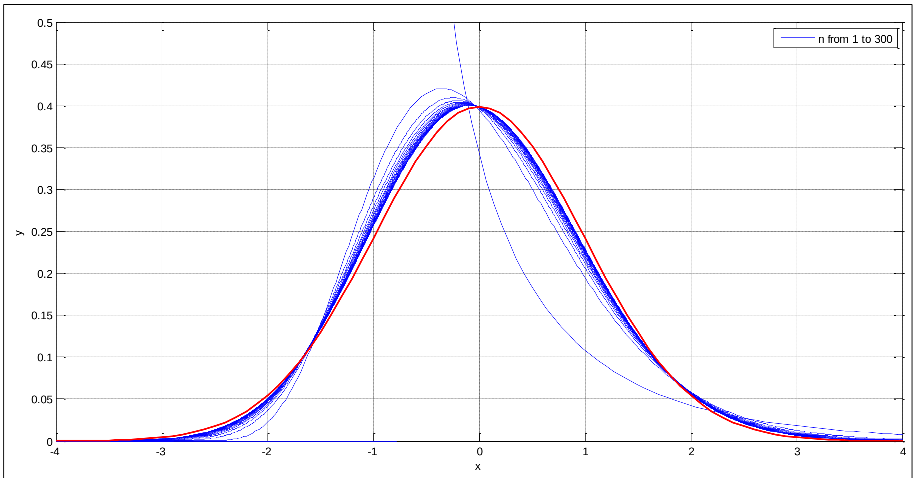

Fig 2: $\zeta(n)$ distribution and standard normal distribution function with degree of freedom increasing

As can be seen from Figure 2, the $\zeta(n)$ distributions are all positively skewed distributions. When the degree of freedom increases, the image of the probability density function of the $\zeta(n)$ distribution approaches the image of the standard normal distribution from the left side; when n is small, the change is large, and when n is larger, the change becomes much smaller.

Table 4: Coordinates of maximum values points

MAX_x	**-0.6364**	**-0.3348**	**-0.2325**	**-0.1946**	**-0.1722**	**-0.1493**	**-0.1391**	**-0.1246**	**-0.1189**	**-0.1144**
MAX_y	inf	0.4204	0.4096	0.406	0.4042	0.4032	0.4025	0.402	0.4016	0.4013
MAX_x	-0.105	-0.1021	-0.0997	-0.0927	-0.0911	-0.0851	-0.084	-0.0831	-0.0781	-0.0776
MAX_y	0.4011	0.4009	0.4007	0.4006	0.4005	0.4004	0.4003	0.4002	0.4001	0.4001

It can be seen from the data in Table 4 that as the degree of freedom increases, the maximum values point of the probability density function of the $\zeta(n)$ distribution gradually moves from the left to the right; Its Abscissa approaches 0, and the maximum values decreases approaching 0.4 gradually.

Table 5: Skewness values

SKEW	**NaN**	**6.7193**	**5.5394**	**4.9421**	**4.5526**	**4.268**	**4.0459**	**3.8649**	**3.373**	**3.5824**
	3.4683	3.3672	3.2766	3.1946	3.1198	3.0512	2.8977	2.9288	2.8736	2.8311

From the data in Table 5, it can be seen that the skewness values of the $\zeta(n)$ distribution is always greater than 0. As the degree of freedom increases, its skewness values gradually decreases, and the decreasing speed becomes slower and slower. It indicates that the $\zeta(n)$ distribution function is positively skewed, and as the degree of freedom increases, it gradually approaches the normal distribution, and the approaching speed becomes slower in the process.

Table 6: Kurtosis values

KURT	**NaN**	**49.2567**	**33.8838**	**27.237**	**23.3127**	**20.6491**	**18.6902**	**17.1719**	**15.9507**	**14.941**
	14.0882	13.3556	12.7175	12.1552	11.6548	11.2058	10.8001	10.4305	10.0904	9.8084

From the data in Table 6, it can be seen that as the degree of freedom increases, the kurtosis values of the $\zeta(n)$ distribution gradually decreases, and the decreasing speed becomes slower and slower, which also indicates that the distribution gradually approaches the normal distribution, and the approaching speed becomes slower and slower.

4. Conclusion

It can be concluded from the analysis above that the chi-square distribution has the following characteristics in the process of the degree of freedom *n* increasing and approaching the normal distribution:

1. The chi-square distribution moves from the left to the right to gradually approach the corresponding normal distribution $N(n,2n)$.
2. The maximum values of the probability density function of the chi-square distribution keeps decreasing, so the maximum points of the chi-square distribution and its corresponding normal distribution function are approaching gradually.
3. The the skewness values and kurtosis values of the chi-square distribution keep decreasing, approaching values of the corresponding normal distribution.
4. Setting $f(x)$ to be the probability density function of the chi-square distribution with degree of freedom *n*, so when the degree of freedom *n* approaches infinity, the distribution of the probability density function $\sqrt{2n}\,f(n+\sqrt{2n}x)$ approaches the standard normal distribution $N(0,1)$. In the approximation process, the maximum value of the chi-square distribution function decreases, and its function image is closer to the image of the standard normal function, but the approaching speed is getting slower.

This research mainly applied MATLAB software to analyze the characteristics in the approximation process. In order to conduct the research feasibly and conveniently, the degree of freedom *n* has only been set in a range that is not too large. Therefore, it requires further study on whether there are any other characteristics in the process of chi-square distribution tending to normal distribution.

Additionally, in this study, we found that the kurtosis values and skewness values calculated by MATLAB software are inaccurate, and even the kurtosis and skewness values of the standard normal distribution are inaccurate, so this problem also needs study further.

5. Acknowledgment

I sincerely thank Mr. Shimei Ma in the School of Mathematics and Statistics and Mr. Tieyu Zhao in Information Science Teaching and Research Section at Northeastern University at Qinhuangdao for his guiding a part of this paper.

6. References

1. Cai ZY. Understanding and specific properties of three sampling distributions. Mathematics learning and research. 2019; 12:143.
2. Zhang ZX, Lin X, Wang HY. Transformation of three distribution of statistics. Journal of Jiamusi University. 2017; 35(04):643-644.
3. Zhu ZR, Chen SQ. Convergence rate of maximum value of chi square distribution sequence. 2016; 38(09):137-142.
4. Zeng Z, Zhang Y. The relationship between χ^2 distribution, t distribution, F distribution and normal distribution. Journal of Hubei Normal University. 2015; 35(03):62-66.
5. Zhang ZX, Liu RY. Hypothesis test of chi square distribution and F distribution discrimination. Journal of science. 2008; 03:33-35.
6. Zhou SG, Qi XL, Yang SH. Study on the relationship between χ^2 distribution and F distribution. Journal of Nanyang Normal University. 2007; 03:21-23, 27.
7. Li KC, Meng ZL. Uniformly asymptotical normality of χ^2 distribution, t distribution and F distribution. Journal of Beijing Printing Institute. 2004; 03:30-33.
8. Liang Q. Normal approximation of χ^2·t· F distribution. Journal of Hubei Institute of Nationalities. 1995; (02):33-38.
9. Liu H. Limit distribution of independent chi square distribution sequence. Journal of Jingchu Institute of technology. 2019; 34(03):45-46, 90.
10. Liang XJ. Further study on three important distributions and their properties. Xinjiang Normal University, 2014.
11. Wang ZX. Asymptotic distribution of chi square random variables with random degrees of freedom. Journal of Jiamusi University. 2009; 27(05):783-785.
12. Wang FC, Cao HR. Approximate calculation of χ^2 distribution, t distribution and F distribution. Journal of Institute of disaster prevention technology. 2008; (01):89-94.