DOI: 10.53469/jerp.2024.06(07).02

A Case Study of MATLAB-based Teaching of Skewness of Two Types of Discrete Random Variable

Kuaile Luo, Ruopeng Wang*

Department of Mathematics, Beijing Institute of Petrochemical Technology, Beijing 102617, China
wangruopeng@bipt.edu.cn

Abstract: ***A case study of teaching skewness of discrete random variables based on MATLAB software is presented. By calculating the skewness of binomial and Poisson distributions and realizing the independent learning process with the help of visual teaching of MATLAB software, it can not only simplify the teaching process of teachers, but also stimulate the learning motivation of students, which only indirectly improves the classroom participation of students in the degree of learning and also lays the foundation for students to engage in the subsequent scientific research work.***

Keywords: MATLAB, Skewness, Discrete random variable, Case study.

1. Introduction

Probability theory and mathematical statistics, as a professional foundation course, provides a powerful mathematical tool for subsequent study and research in the professional field. And skewness is an important numerical characteristic of random variables in the mathematical statistics part, reflecting the degree and direction of skewness of the distribution of random variables. In the existing literature, there are more discussions on the application of skewness, but the calculation of skewness is not very perfect, especially the formula for calculating the skewness of descriptive statistics is even less. And with the development of the information age, various educational methods based on information technology have appeared, such as distance education and webcasting. MATLAB software developed by MathWorks in the United States has also been developed in the field of mathematics learning, and this software plays a prominent role in the realization of numerical analysis, computation, and visualization research. Based on the rich advantages of this software, in order to better enhance the learning efficiency of students, cultivate students' mathematical literacy, and enhance the professionalism of teachers, this software is introduced into the beginning of mathematical statistics students, for the calculation of skewness formula is relatively complex, most students only need to understand the concept of skewness, and the calculation of skewness can be realized with the help of MATLAB software.

Based on the above analysis, this paper proposes the teaching practice of skewness in probability theory and mathematical statistics based on Maltab to help students better understand and master the basic concepts by materializing the difficult calculations, enhancing students' self-confidence in learning, and combining image thinking and abstract thinking by the software. Combining image thinking and abstract thinking, not only simplifies the teaching process of teachers, but also helps students to build the process of thinking and improve students' motivation to learn, hope that this aspect of the practice of the front-line teachers can provide some ideas and help for the integration of the software into the teaching to throw a brick to attract a jade.

2. Background

As the number of university enrollment continues to increase, teaching methods, student differences and many other factors will affect the effectiveness of teaching. Our school belongs to the Beijing municipal colleges and universities, enrollment of students in general mathematical knowledge is relatively weak, and probability theory and mathematical statistics as a professional required courses, each professional students of all involved. This paper takes the course of the institution as an example, mainly from the content of the extreme value, teaching methods and student differences in three aspects to explain the current situation and shortcomings of this content.

From the curriculum of probability theory and mathematical statistics, probability theory and mathematical statistics are usually offered in the third semester, with three credit hours per week, for a total of 48 hours. And the skewness content syllabus will be designed for the content of less than one classroom hour, the actual classroom hours and the content of the required academic fashion gap. In the actual teaching can only catch up the class time to reduce the difficulty. Although this way to build a knowledge system is not much, but the subsequent study of the course and the application of the students can not consolidate the foundation, so that in the actual application, students always feel at a loss.

From the perspective of teaching methods, the main alternative to PowerPoint teaching, still lack of intuitive analysis and practical operation, teacher-student interaction and student participation is low, students can not hand can be feedback and application, encountered in the perspective of the actual problem is at a loss. The training goal of engineering students is to emphasize the application, this approach has a certain gap with the actual goal.

From the students' side, with the massification and popularization of higher education, the students who enter our school generally do not have solid knowledge, for the complicated deduction, often insufficient information, for the theory put into practice link is relatively weak.

Based on the above analysis, combining the content of "skewness" in probability theory and mathematical statistics as an example, through the teaching design, MATLAB is introduced into the teaching and practice, avoiding complicated calculations, increasing students' participation in the classroom, improving the efficiency of the classroom, and enhancing students' motivation to learn.

3. A Case for Teaching Skewness

The distribution function of a random variable can completely describe the statistical characteristics of a random variable, but in some practical problems, there is no need to go to a comprehensive examination of the variation of the random variable, but only need to know some characteristics of the random variable, and therefore do not need to find out its distribution function. Skewness is an important numerical characteristic of random variables, which is of great significance in theory and practice.

3.1 Definition of Skewness

Skewness (denoted by S_k) refers to an asymmetric distribution in statistics. A distribution with asymmetric tails extending in the positive direction is said to be "positively skewed" ($S_k > 0$), while a distribution with asymmetric tails extending in the negative direction is said to be "negatively skewed" ($S_k < 0$), and a distribution is said to be symmetric if $S_k = 0$ when the statistical distribution is symmetric. The definition is as follows:

Skewness, also known as the skewness coefficient, measures the degree of skewness by dividing the third-order central moment by the third power of the standard deviation, i.e. $S_k = \frac{\nu_3}{\sigma^3}$, where ν_3 is the third-order central moment and σ is the standard deviation which ranges between -3~3 .

3.2 Skewness Computation for Two Types of Discrete Random Variables

Below we calculate the skewness of the Binomial and Poisson distributions separately.

3.2.1 Skewness of the Binomial distribution

The probability of an event succeeding is p and the number of independent repetitions of times, the number of successful occurrences of the event obeys the Binomial distribution. And the probability that the event succeeds k times is $P(X=k)=C_n^k p^k(1-p)^{n-k}$.

The mathematical expectation of the binomial distribution is $E(X) = np$ and the moment is $D(X) = np(1-p)$. Next, we compute the third-order central moment.

We have

$$E(X-EX)^3 = E\left(X^3 - 3X^2E(X) + 3X\left(E(X)\right)^2 - (EX)^3\right)$$
$$= E\left(X^3\right) - 3E\left(X^2\right)E(X) + 3E(X)\left(E(X)\right)^2 - \left(E(X)\right)^3$$
$$= \sum_{k=0}^{\infty} k^3 C_n^k p^k (1-p)^{n-k} - 3np\left(\left(E(X)\right)^2 + D(X)\right) + 2n^3p^3$$

where

$$\sum_{k=0}^{\infty} k^3 C_n^k p^k (1-p)^{n-k}$$
$$= \sum_{k=0}^{\infty} k(k-1)(k-2) + k(k-1) C_n^k p^k (1-p)^{n-k}$$
$$= \sum_{k=0}^{\infty}\left(\frac{n!}{(k-3)!(n-k)!} + \frac{n!}{(k-2)!(n-k)!} + \frac{n!}{(k-3)!(n-k)!}\right)p^k(1-p)^{n-k}$$
$$= \sum_{k=0}^{\infty} n(n-1)(n-2)\, p^3 C_{n-3}^{k-3} p^{k-3}(1-p)^{n-k}$$
$$+ \sum_{k=0}^{\infty} n(n-1)\, p^2 C_{n-2}^{k-2} p^{k-2}(1-p)^{n-k} + np C_{n-1}^{k-1} p^{k-1}(1-p)^{n-k}$$
$$= n(n-1)(n-2)\, p^3 + n(n-1)\, p^2$$

That is, $E(X-EX)^3 = np(1-p)(1-2p)$. Therefore, we get that the skewness of the Binomial distribution is $S_K = \frac{1-2p}{\sqrt{npq}}$.

3.2.2 Skewness of the Poisson distribution

Since the distribution law of the Poisson distribution is $P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda}$, where λ >0 and k=0,1,⋯. The mathematical expectation of the binomial distribution is, and the moment is. Next, we compute the third-order central moment.

We have

$$E(X-EX)^3 = E\left(X^3 - 3X^2E(X) + 3X\left(E(X)\right)^2 - \left(E(X)\right)^3\right)$$
$$= E\left(X^3\right) - 3E\left(X^2\right)E(X) + 3E(X)\left(E(X)\right)^2 - \left(E(X)\right)^3$$
$$= \sum_{k=0}^{\infty} k^3 \frac{\lambda^k}{k!}e^{-\lambda} - 3\lambda((E(X))^2 + D(X)) + 3\lambda^3 - \lambda^3$$

where

$$\sum_{k=0}^{\infty} k^3 \frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=0}^{\infty}\left(\left(k(k-1)\right)(k-2) + k(k-1) + k\right)\frac{\lambda^k}{k!}e^{-\lambda}$$
$$= \sum_{k=0}^{\infty}\left(\frac{\lambda^k}{(k-3)!}e^{-\lambda} + \frac{\lambda^k}{(k-2)!}e^{-\lambda} + \frac{\lambda^k}{(k-1)!}e^{-\lambda}\right)$$
$$= \lambda^3 \sum_{k=0}^{\infty}\frac{\lambda^{k-3}}{(k-3)!}e^{-\lambda} + \lambda^2\sum_{k=0}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}e^{-\lambda} + \lambda\sum_{k=0}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}$$
$$= \lambda^3 + \lambda^2 + \lambda$$

Thus, we have $E(X-EX)^3 = \lambda$. Therefore, we get that the skewness of the Poisson distribution is $S_k = \frac{1}{\sqrt{\lambda}}$.

4. MATLAB Visualization

To extend the application of skewness, the results of the obtained skewness of each distribution are written as a MATLAB function, respectively, binomial distribution (program see Appendix 2), Poisson distribution (program see Appendix 3), and in order to be intuitive, a MATLAB visualization is done (program see Appendix 1).

Program ideas: known function expression, function input parameters, will not meet the conditions of the parameters re-entered until the input is correct, in accordance with the expression of the different functions, calculate the expression of the distribution of random variables. MATLAB visualization interface runs as follows:

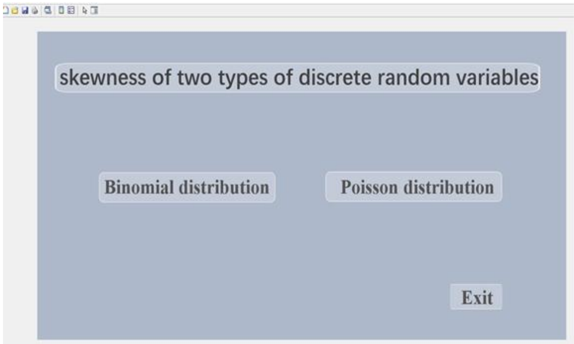

Figure 1: Visualization of the Binomial distribution and Poisson distribution

5. Conclusions

In the teaching of probability theory and mathematical statistics courses, traditional teaching is more theoretical, students are not easy to understand the abstract knowledge, and the calculation ability is weak. Based on the visualization function of MATLAB software, it helps the teaching of probability courses, improves classroom efficiency, cultivates students' divergent thinking, mathematical observation ability and analytical ability, and deepens their understanding and mastery of knowledge. In this paper, MATLAB visualization teaching, convenient and intuitive, easy to operate the conclusion, help students understand the limit process, improve the knowledge of the importance of mathematical knowledge, to cultivate students' interest in mathematics, but also for the students to contact this kind of computer mathematical programming software to lay a good foundation for the students to stimulate the motivation to learn, only indirectly enhance the students' participation in the classroom to learn the degree of learning, and also for the students to subsequently It also lays the foundation for students to engage in scientific research in the future.

Acknowledgement

This work was supported by Beijing College Students Innovation and Entrepreneurship Training Program Fund (No.2023J00067) and Project of Educational and Teaching Reform of Beijing Institute of Petrochemical Technology (No.2023J00055).

References

[1] Zheng Hongqing, “Assistance of Matlab Software to Higher Mathematics Teaching-Taking the Computer Major as an Example,” Science and Technology Information, 2024, (1), pp. 182-185.

[2] Chen Zhenmin, Wang Chuhan, “MATLAB-Based Teaching Strategies for Higher Mathematics and Their Practice,” Vocational training, 2023,12(4), pp. 479-491.

[3] Xue Dingyu, Chen Yangquan, Simulation and Applications Based on MATLAB/Simulink, Tsinghua-University-Press, Beijing, 2011.

[4] Zhang Yuejiao, “Application Research of MATLAB in Three Operations of Advanced Mathematics,” Heilongjiang Science, 2022,13(19), pp. 158-160.

[5] Qu Wei, “Reflections on Course Teaching of Advanced Mathematics Assisted by MATLAB,” Education and Teaching Forum, 2022 (47), pp. 173-176.

Author Profile

Ruopeng Wang is a associated professor in Department of Mathematics at Beijing Institute of Petrochemical Technology. He received thePHD degrees in Department of Mathematics from Beijing Jiaotong University in 1999.

Appendix

Appendix 1

```
function varargout = pd(varargin)
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                   'gui_Singleton',  gui_Singleton, ...
                   'gui_OpeningFcn', @pd_OpeningFcn, ...
                   'gui_OutputFcn',  @pd_OutputFcn, ...
                   'gui_LayoutFcn',  [] , ...
                   'gui_Callback',   []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end
if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end

function pd_OpeningFcn(hObject, eventdata, handles, varargin)
handles.output = hObject;
guidata(hObject, handles);

function varargout = pd_OutputFcn(hObject, eventdata, handles)
varargout{1} = handles.output;

function pushbutton1_Callback(hObject, eventdata, handles)
n=input('input a number:');
p=input('input a probability:');
skb=B(n,p)

function pushbutton4_Callback(hObject, eventdata, handles)
lam=input('input a number:');
skp=P(lam)
```

Appendix 2

```
function pushbutton5_Callback(hObject, eventdata, handles)
ss=questdlg('are you exit？','Exit','No！','Yes！');
switch ss
    case 'Yes！'
        delete(handles.figure1);
end
```

```
function skb=B(n,p)
a=1;
while a
    if n~=fix(n)
         n=input('input a number again:');
         a=1;
    else if p>1||p<0
        p=input('input a probability again:');
        a=1;
        else
            a=0;
            fprintf('Binomial distribution:');
            skb=(1-2*p)/sqrt(n*p*(1-p));
        end
     end
     end
end
```

Appendix 2function skp=P(lam)

```
a=1;
while a
    if lam<=0
         lam=input('input a number again:');
         a=1;
    else
         a=0;
         fprintf('Poisson distrubution:');
         skp=1.0/sqrt(lam);
    end
end
```