Advances in Social Science, Education and Humanities Research, volume 385
3rd International Conference on Education, Economics and Management Research (ICEEMR 2019)

On the Teaching of Axiomatic Probability for Students of Science and Engineering

Kunping Zhu
Department of Mathematics
East China University of Science and Technology
Shanghai 200237, China
*Corresponding Author: kpzhu@ecust.edu.cn

***Abstract*—Axiomatic method is an important characteristic of Mathematics. The teaching of axiomatic method can cultivate students' rigorous logical deduction ability, and also enhance students' cognition of relevant mathematical knowledge. For most students majoring in science and engineering, the axiomatic method of probability is abstract and difficult. The teaching of axiomatic probability is also considered to be only a matter of strictness by some teachers. Thereby, the teaching of axiomatic method of probability is often neglected for the sake of easy understanding. However, the unfulfilled teaching in axiomatic probability will affect students' understanding of conditional probability. It may even lead to mistakes in probability calculation. Such an example with different solutions is discussed in this paper. Meanwhile, based on the axiomatic probability method, the reasons for several typical misreading of conditional probability are analyzed respectively. Furthermore, the key teaching points of axiomatic probability for students of science and engineering are discussed.**

Keywords*—*axiomatic method; probability space; conditional probability; active learning

I. INTRODUCTION

Axiomatic method is one of the prominent characteristics of mathematics. It emphasizes logical deduction on the basis of axioms. The cultivation of reasoning ability was also considered as the core issue of mathematics education in the late 1990s [1]. Under the background of the implementation of quality education, the concept of mathematics education has also undergone profound changes in china. Today's mathematics education emphasizes the cultivation of students' comprehensive qualities, such as mathematical awareness, problem solving, logical reasoning and information interchanging [2]. Although the significance of axiomatic method in mathematics education has declined [3], it is still indispensable. According to the theory of Active Learning, understanding information is the key to learning, whereas the process of understanding must be realized by students' intuition and experience. By reasoning, the bridge between perceptual knowledge and rational knowledge can be built. That is, the teaching of axiomatic method is propitious to the improvement of students' cognitive competence. The basic knowledge of probability, including conditional probability, has been integrated into mathematics teaching in middle school. But the concept of probability is explained by possibility based on examples, and the teaching of conditional probability is generally regarded as a difficult point in middle school mathematics [4]. The difference between probability teaching in universities and middle schools lies not only in more contents, but also in the introduction of axiomatic methods. Without the teaching of axiomatic probability, it is very difficult for students to achieve proper understanding of probability. Although the axiomatic probability established by Kolmogorov in 1933 is the cornerstone of the whole modern probability theory [5], the teaching of axiomatic method is generally not taken seriously in the college probability course for science and engineering. Most Chinese probability textbooks do not talk about probability space, but only the three probability axioms. The three axioms are used to prove the probability function of conditional probability. However, the proof is often overlooked or replaced by the words "it is provable". As a result, students can only understand probability based on possibility, and have no concept of probability functions. This is just the reason for many misinterpretations of conditional probability [6]. There are often debates about conditional probability on the Internet [7]. The misreading of conditional probability is also common among doctors and lawyers [8]. Moreover, there are imprecise expressions of conditional probability in some textbooks and published papers, where the most typical mistake is to consider $P(B|A)$ as the probability of event "$B|A$". All of the above shows that the teaching of axiomatic probability in college mathematics is relatively inadequate. In this paper, the problem of different solutions to a given probability question is discussed firstly. The theory of probability space is then introduced as a supplement to the teaching of probability and statistics. It is followed by the analysis of several typical misreading on conditional probability. Finally, the key points of probability teaching for students of science and engineering are discussed in the view of active learning.

II. ARGUMENTS ON A PROBABILITY PROBLEM

The following probability calculation problem comes from the Chinese graduate entrance examination for students of science and engineering. Different solutions to the problem are given in two tutorial textbooks. More arguments on similar issues can also be found on the Internet [7]. The question and two solutions are given as follows.

There are two boxes of the same kind of parts. The first box contains 50 parts, of which 10 are first-class. There are 30 pieces in the second box, of which 18 are first class. First pick out a box randomly from the two boxes, and then take out two parts from the box in turn without putting them back. The question is to find out the probability that the part taken out second is first-class, under the condition that the first taken out is first-class.

Let A_1, A_2 be the events that the first and second boxes are picked out respectively, and B_1, B_2 be the events that the parts taken out first and second are first-class respectively. For $P(A_1)=P(A_2)=0.5$, it holds immediately by the total probability formula that $P(B_1)=P(A_1)P(B_1\,|\,A_1)+P(A_2)P(B_1\,|\,A_2)=0.4.$

Solution 1: The probability to be determined can be expressed by $P(B_2|B_1)$. It follows

$$P(B_2\,|\,B_1)=\frac{P(B_1B_2)}{P(B_1)}=\frac{P(A_1)P(B_1B_2\,|\,A_1)+P(A_2)P(B_1B_2\,|\,A_2)}{P(B_1)}\approx 0.4856$$

Solution 2: Let C be the event that the part taken out second is first-class, under the condition that the first taken out is first-class. Thereby, the probability to be determined is $P(C)$. The conditional probability of C given A_1 is $P(C\,|\,A_1)=9/49$, and $P(C\,|\,A_2)=17/29$. By the total probability formula, it obtains:

$$P(C)=P(A_1)P(C\,|\,A_1)+P(A_2)P(C\,|\,A_2)=\frac{1}{2}\times\frac{9}{49}+\frac{1}{2}\times\frac{17}{29}\approx 0.3849.$$

There are two different solutions to the same problem, one of which must be wrong. Putting the argument on the question aside, we can use MATLAB to make a random simulation of the extracting process. That is, by generating uniformly distributed random numbers in interval [0,1] to simulate the process of box selection and parts extraction. The flow chart of simulation is shown in Figure 1, where the initial values of N1, N2, M1, M2, and i are all set to be zeros.

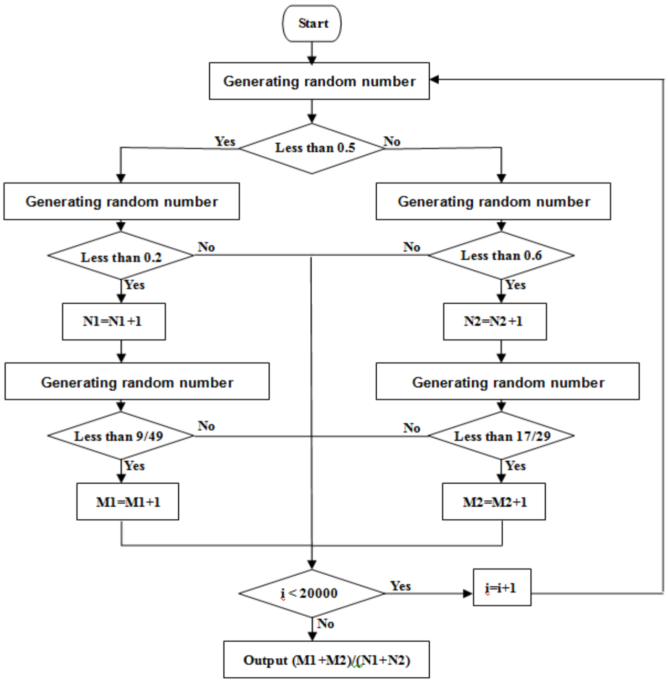

Fig.1. Flow chart of the simulation

The output of simulation is 0.48548, which is the frequency of 2000 tests that the part taken out second is first-class under the condition that the first taken out is first-class. The outputs by repeating the simulation are all in the neighbor of the result of Solution 1, which shows that the first solution is correct. So, what's wrong with Solution 2? In fact, the mistake of Solution 2 is that "*C*" is not an event at all. This kind of mistake is very common among students. To clarify the reasons, we must resort to the axiomatic method of probability.

III. Axiomatic Probability

The axiomatic system of probability contains six axioms in all. Most textbooks of probability and statistics for students of science and engineering introduce only the three axioms for probability, without involving the three axioms for events. The comprehensive understanding of the axiomatic of probability for a teacher will help him to improve his teaching level of probability [9]. For references, a supplementary introduction to probability space is given as follows, which is also the basis for subsequent analysis.

A. Probability Space

Probability space (Ω, F, P) is a whole involving three elements [10], in which sample space Ω is a set of all the basic results of a random experiment, event field F is a set of some subsets of the sample space (probably not all subsets of Ω), and probability P is a set function (also known as probability measure) from event field F to interval [0,1]. The event field is also called sigma algebra or Borel field. There are three axioms for event field F in probability space (Ω, F, P).

(i) $\Omega\in F;$

(ii) If $A\in F$, then $\overline{A}\in F;$

(iii) If $A_i\in F\ (i=1,2,\cdots)$, then $\bigcup_{i=1}^{\infty}A_i\in F.$

In addition, there are three axioms for probability measure P, namely, the non-negative axiom, the normative axiom and the countable additive axiom respectively [11].

(I) $P(A)\geq 0$ for all $A\in F;$

(II) $P(\Omega)=1;$

(III) If $A_1, A_2,\cdots$ are pair-wise disjoint elements of F, then

$$P(\bigcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}P(A_i).$$

The above six axioms of probability space give the most fundamental rules for events and probabilities. All the functions from F to [0,1] satisfying the axioms are called probabilities. In essence, probability is a measure, and cannot be defined for the un-measurable subsets of Ω . All such subsets are excluded from event field F by the first three axioms. In the teaching of probability and statistics, the question what subset is not an event is often asked by students. For ease understanding, a concrete subset based on Vitali subset [12] is given below.

B. Subset without Probability

Suppose that a number is randomly selected from interval [0,1], then the sample space is $\Omega=[0,1]$. Let A be a subset of [0,1], and the length of A be denoted by $|A|$ for convenience. Obviously, the probability that the number comes from set A is the length $|A|$. If $|A|$ does not exist, then there is no probability for A. That is, the subset A of Ω is not an event. The construction of such a subset is given as follows.

For any $x\in[0,1]$, put together all the numbers in the interval [0,1] whose difference from x is a rational number to generate a class. That is

$$[x]=\{y\,|\,(y-x)\in Q, y\in[0,1]\}.$$

It's easy to prove that all the different classes defined above form a division of the interval [0,1]. Take a number form each of the different classes, and put all the numbers together to make up a set denoted by A. We will prove that set A is the required set.

From $A\subseteq[0,1]$ it follows that $0\le|A|\le1$, namely, $|A|=0$ or $|A|>0$. Define set B by

$$B=\bigcup_{q\in Q\bigcap[-1,1]}(A+q)=\bigcup_{q\in Q\bigcap[-1,1]}\{a+q\,|\,a\in A\} \tag{1}$$

Then, we have

$$|B|=\sum_{q\in Q\bigcap[-1,1]}|(A+q)|=\sum_{q\in Q\bigcap[-1,1]}|A| \tag{2}$$

If $|A|=0$, then $|B|=0$ by (2). On the other hand, for any $x\in[0,1]$, there is an element $a\in A$ such that $a\in[x]$. For $x-a$ is a rational number in the interval [-1,1], there exists $q\in Q\bigcap[-1,1]$ such that $x=a+q$. It follows that $x\in B$ by (1). Thereby, it holds $[0,1]\subseteq B$. That is $|B|\ge1$, which is contradictory to $|B|=0$.

According to definition of B in (1), it follows that $B\subseteq[-1,2]$. That is $|B|\le3$. By (2), $|B|$ is the sum of infinitely countable $|A|$. If $|A|>0$, it follows that $|B|$ is infinite, which is contradictory to $|B|\le3$.

In summary, whether $|A|=0$ or $|A|>0$ will lead to contradictions. That is, the length of A doesn't exist. Such un-measurable set is not an event. This is why the event field may not include all subsets of the sample space.

IV. Problems and Analysis

In the course of probability and statistics, the six axioms above are used not only to the proofs of properties of probability, but also to the proofs of conditional probability functions. In teaching practice, the proofs of the latter are easily ignored for good understanding, and the students also prefer to understand the definition of conditional probability through examples rather than axioms. However, aside from the axiomatic method, it may lead to misreading problems on conditional probability.

A. Problem of Misreading

If A and B are two events and $P(A)>0$, the conditional probability of B for given A is defined by:

$$P(B\,|\,A)=\frac{P(AB)}{P(A)} \tag{3}$$

The definition can be explained intuitively with Figure 2, where the given event is regarded as the new reduced sample space. Thereby, each event B in Ω is reduced to event AB in Ω_A. Taking geometric probability as an example, the possibility of B for given A is the ratio of the area of region AB to the area of region A, which is equal to the probability ratio by (3) obviously.

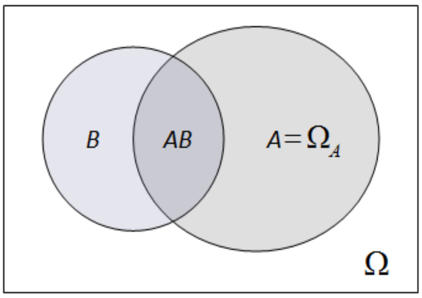

Fig.2. Intuitive Explanation of Conditional Probability

However, it is not enough to understand the definition of conditional probability based on experience and examples. As the argument discussed in section 2, there are lots of misuse and misreading on conditional probability, such as:

1) Consider $P(B|A)$ as the probability of "event $B|A$";
2) Consider definition (3) as a provable theorem;
3) Confuse the probability functions on both sides of equation (3);
4) Consider $P(A)$=1 in equation (3), for A is given;
5) Take the relationship between A and B in (3) as the occurrence of event A before the occurrence of B;
6) Misreading of the interaction between events.

The above problems are quit common among students of science and technology. To clarify the misunderstandings, the axiomatic method is indispensable.

B. Axiomatic Analysis

Based on the axiomatic probability discussed in section 3, the definition of conditional probability involves two different probability spaces. The symbol P on the right of (3) stands for the probability function of probability space (Ω,F,P) with $F\xrightarrow{P(*)}[0,1]$. The left side P is a different probability function from the right. For the given A, the conditional probability space is defined as follows [13]:

$$\begin{cases}(\Omega_A,F_A,P_A)\\ F_A\xrightarrow{P_A(AB)=\frac{P(AB)}{P(A)}=:P(B|A)}[0,1]\end{cases} \tag{4}$$

where $\Omega_A=A$ and $F_A=\{BA\,|\,B\in F\}$. It is easy to verify that the six axioms hold in (Ω_A,F_A,P_A). That is, the conditional probability $P(B|A)$ stands for the probability of event AB in (Ω_A, F_A, P_A). In particular, there is equation $P_A(A) = P(A)/P(A) = 1$ in the conditional probability space, whereas $P(A)$ may not equal 1 in the original probability space without condition. The probability space formulated by (4) is consistent with the explanation in Figure 2. In many cases, the calculation of conditional probability may be simplified by directly taking the given event as a new sample space. The conditional probability space is not unique. It can also be formulated as [10]:

$$\begin{cases} (\Omega, F, P_A) \\ F \xrightarrow{P_A(B)=\frac{P(AB)}{P(A)}=:P(B|A)} [0,1] \end{cases} \tag{5}$$

That is, both sample space and event field remain unchanged. In this case, the conditional probability $P(B|A)$ stands for in fact the probability of event B in (Ω, F, P_A). Whether the conditional probability function is defined by (4) or (5), it corresponds to the same value. Anyway, $P(B|A)$ is not the probability of "$B|A$", for "$B|A$" is not an event at all. If the probability functions are confused, it may lead to mistakes in probability calculation as shown in section 2.

Conditional probability indicates the effect of one event on another. It is provable that such influence is mutual, and the degree of influence is the same. In fact, if $P(A)P(B)>0$, then

$$\frac{P(B|A)-P(B)}{P(B)} = \frac{P(AB)/P(A)-P(B)}{P(B)} = \frac{P(AB)-P(A)P(B)}{P(A)P(B)}$$
$$= \frac{P(B)[P(A|B)-P(A)]}{P(A)P(B)} = \frac{P(A|B)-P(A)}{P(A)}.$$

That is, the percentage of probability change of event B for given A, is the same as the percentage of probability change of event A for given B. In particular, when the percentage of change is zero, the two events are independent. It is to note that the independence of two events is not equivalent to the non-influence of two events. An event with zero probability is independent of any events, but it cannot be a given event in conditional probability. Since the impact of two events is mutual, it has nothing to do with the timing of their occurrence. That is, the event as condition does not necessarily occur first in conditional probability. Take the topic of section 2 for example, where event B_2 occurs after event B_1. However, the probability of B_1 for given B_2 can also be calculated:

$$P(B_1|B_2) = \frac{P(B_1B_2)}{P(B_2)} = \frac{P(A_1)P(B_1B_2|A_1)+P(A_2)P(B_1B_2|A_2)}{P(B_2)} \approx 0.48557.$$

That is, the later event B_2 also has an effect on the earlier event B_1. Moreover, the occurrence of either event B_1 or B_2 will increase the probability of the other by 21.39% according to the above discussion.

V. Teaching of Axiomatic Probability

The course of probability and statistics for science and engineering is generally offered for sophomores with three credits in most Chinese universities. The students have learned advanced mathematics and laid a solid mathematical foundation in the early stage. There is no problem for students to understand axiomatic probability in the case of appropriate teaching. However, the course of probability and statistics mainly focuses on cultivating students' application ability, and often pays insufficient attention to the method of axiomatic probability in teaching. The concept of probability space is not introduced in most textbooks for students of science and engineering. The latest edition of the syllabus by the Chinese Ministry of education also requires only the three axioms about probability and does not involve the three axioms about events [14]. In practice, the course requirements may be further reduced, therein the axiomatic method is very easily neglected in teaching. As mentioned in section 2, there are lots of arguments about the calculation of conditional probability, which shows that the problem of teaching for axiomatic probability is quite common.

By the above analysis in this paper, probability is a set function from event field to interval [0,1] and involves no concept of possibility. Apart from the axiomatic definition, it is inadequate to understand the concept of probability as the degree of possibility directly. The confusion of different probability functions may lead to computational errors as shown in the argument. The significance of the axiomatic method lies in the following three aspects:

- The axiomatic system is the theoretical basis of the whole course;
- The axiomatic method is also a theoretical tool for deducing and discovering new laws;
- In terms of axiomatic method, it is helpful to improve the understandings, so far as to clear up the misreading of probability knowledge.

In short, the teaching of axiomatic method is very meaningful and indispensable for the students of science and engineering.

The emphasis of axiomatic probability in teaching is the three axioms for probability. It is easy for students to understand these axioms by comparing them with the similar properties of frequency. These three axioms are used to prove the properties of probability. The proof can well demonstrate the characteristics of axiomatic methods. But the proof involves only one probability function in the same probability space. The more important application of these three axioms is the proof of conditional probability, which involves different probability spaces. By verifying that the conditional probability satisfies the three probability axioms, it helps students to distinguish the different probability functions in different probability spaces, which is critical for students to correctly understand the definition of conditional probability.

The difficulty of axiomatic probability in teaching lies in the three axioms about events. That is, the event field may be not the power set of the sample space. The question of what subset of sample space is not an event is also often asked by some students. For the students of science and engineering, the question is difficult to explicate in class because it is related to measure theory. It is enough to know that not all subsets of sample space have probabilities, when the sample space is an uncountable set. As extra-curricular materials, a specific example is given in section “B. Subset without Probability” for the reference of interested students.

Student-centered teaching encourages student’s active learning to meet their different needs. Some students who are interested in the course may want to learn more about axiomatic probability. It is feasible to provide extended content such as probability space to interested students for online learning. The misunderstandings of conditional probability discussed in this paper are particularly appropriate for students to discuss in flipped classroom. The problems from students themselves such as the wrong solutions can arouse their interesting and thinking. Through the discussion of the reasons for the mistakes, the students can deepen their understanding of the axiomatic probability. In addition, there are many online courses of probability and statistics for students of science and engineering. By encouraging students to use appropriate resources for active learning, it is a powerful complement to the classroom teaching.

VI. CONCLUSION

Some problems in the teaching of the axiomatic method of probability are discussed. The insufficient teaching of axiomatic method may lead to student’s misreading and misuse of conditional probability. A controversial example is given to illustrate that the wrong result by misuse can be identified by random simulation. The reasons for several typical misreading of conditional probability are analyzed by means of axiomatic method. It shows that the teaching of the three probability axioms is very important for students’ understanding of conditional probability. The key teaching points of axiomatic probability are pointed out. In addition, the student-centered teaching for the students of science and technology is discussed.

REFERENCES

[1] Wu Yaping. The Core Problem of American Mathematics Education - the cultivation of reasoning ability. Foreign educational materials [J]. 1999, (5): 59-55. (In Chinese)

[2] Zhu Kunping. Discussion on the Student-centered Teaching of College Mathematics [J]. American Journal of Educational Research, 2018, 4(4):144-148.

[3] Sun Mingfu, Gao Weizong, Zheng Qingquan. Conception of the Reform of Mathematical Axiomatic Method in Senior High School [J]. Journal of Mathematical Education, 2003, 12(2): 88-90. (In Chinese)

[4] He Shengqing. Sixth Grade Students' Cognition of Higher Probability Content: Potential and Limitation [J]. Journal of Mathematical Education, 2018, 27 (3): 57-61. (In Chinese)

[5] Xu Chuansheng, Yang Jun. Several Evolutions of Probability Philosophy [J]. Research on Dialectics of Nature, 2008, 24 (5): 78-82. (In Chinese)

[6] Xie Meihua. Misunderstanding in the Application of Conditional Probability [J]. Mathematical Bulletin, 2017, 56 (7): 33-38. (In Chinese)

[7] Online available from https://bbs.hupu.com/17449693.html

[8] John Allen Paulos. Innumeracy [M]. America: Hill and Wang Pub. 1988:28-29.

[9] Zeng Xiaoping, Liu Xiaoli, Tu Rongbao, Ning Lianhua. Research on the Influence of Teachers' Mathematics Accomplishment on Mathematics Teaching [J]. Journal of Mathematical Education, 2010, 19 (2): 42-46. (In Chinese)

[10] Li Xianping. Basis of Probability Theory [M]. Higher Education Press, 2010:38-49. (In Chinese)

[11] Xia Ningmao, Qin Yan, Ni Zhongxin. New edition of probability theory and mathematical statistics [M]. East China University of Technology Press, 2011:12-18. (In Chinese)

[12] Cheng Qixiang et al. Basis of Real Function and Functional Analysis [M]. Higher Education Press, 1983:74-78. (In Chinese)

[13] A.N.Kolmogorov. Foundations of the Theory of Probability [M]. Chelsea Publishing Co. New York, 1956:47.

[14] College Mathematics Teaching and Instruction Committee of Ministry of Education. Basic Requirements for College Mathematics Teaching [M]. Higher Education Press, 2014:8. (In Chinese)

[15] Chen Gang, Sun Ting, Ding Huiling. The Discussion of Random Variables Correlation and Random Variables Causal Relationship, Mathematics in Practice and Theory [J], 2018, 46 (5), 194-200. (In Chinese)