

Students' Understanding of Conditional Probability on Entering University

Robyn Reaburn
University of Tasmania
<Robyn.Reaburn@utas.edu.au>

An understanding of conditional probability is essential for students of inferential statistics as it is used in Null Hypothesis Tests. Conditional probability is also used in Bayes' theorem, in the interpretation of medical screening tests and in quality control procedures. This study examines the understanding of conditional probability of students entering an introductory applied statistics unit at an Australian university. These students answered questions that tested their ability to interpret conditional statements in two-way tables and their ability to discern the difference between a conditional statement and its inverse in written form. They also answered questions to determine if they held the time-axis fallacy.

Conditional probability is that which arises in a random experiment when the only elements of interest are those that are in some subset of the entire sample space (Tarr & Jones, 1997). Formally, the formula for the calculation of a conditional probability is written as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

This expression denotes that the only events of interest are those in the subset A that can be found in the subset B. That is, the sample space is reduced to only those events in B. Conditional probabilities also arise when there is a lack of independence, that is, when the probability of an event changes owing to a previous event. These types of conditional probabilities occur in "non-replacement" random experiments (Tarr & Jones, 1997). For example, if there are 5 white and 5 red marbles in a bag the original probability of getting a white marble is 1/2 (similarly for a red marble). If a white marble is taken out and not replaced then the probability of getting a white marble on the second draw is then 4/9 and of getting a red marble is 5/9.

Conditional probabilities are the basis of Bayes' theorem, which is important in the medical field when clinical diagnoses need to be made. Bayes' theorem is used to recalculate probabilities when new knowledge is acquired. Formally, the formula is written as below, where A' is the complement of A.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')P(A')}$$

For an example of its use, consider a person who is selected at random from the population. This person will have a chance of having HIV that is equal to the proportion of the disease in the population. If, however, this person has a positive screening test (routinely carried out on blood donors), what now is that person's chance of having the disease? Bayes' theorem is used to calculate this probability. The actual probability is quite low, and this is essential information for people who are called back for further tests. Conditional probabilities are also used in inferential statistics in hypothesis tests, where the probability of sample statistics and those more extreme are calculated given that a hypothesis about the entire population is true.

Previous research has examined students' understanding of conditional probability in primary and secondary school. Fischbein and Gazit (1984) studied the responses to

questions that asked 590 students in Grades 5, 6 and 7 to find the probability of taking black and red marbles from a box with and without replacement. In all three grades, the proportion of correct answers was lower for the without replacement questions than for the with replacement questions. It was only in Grade 7 that more than half of the students could answer the without replacement questions correctly (Grade 5: approximately 22%, Grade 6: approximately 43% and Grade 7: approximately 72%). The researchers found that where there was no replacement the students tended to forget that both the number of expected outcomes and the total number of outcomes needed to be reduced by one.

Tarr and Jones (1997) examined the response of students from Grades 4 to 8 (3 in each grade) to questions that required reasoning in a variety of contexts where the probabilities changed as items were removed from a sample space without replacement. They placed the reasoning used by the students into four levels ranging from Level 1 where students tended to “ignore relevant quantitative information and generally believe[d] that they could control the outcome of an event” (p. 51) to Level 4 where the students used numerical reasoning and could “assign numerical probabilities spontaneously and with explanation” (p. 54). Whereas the Level 1 students were found only in the lower two grades there was not a consistent increase in the level of reasoning from one grade to the other.

In a study by Watson and Kelly (2007) 69 students from Grades 3 to 13 were asked a series of questions involving conditional probabilities in different formats. One question required students to distinguish $P(S|W)$ and $P(W|S)$ when the question was in a written format. The proportion of students who answered this correctly with an appropriate justification increased with age. Another question required the calculation of conditional probabilities based on data in a frequency table; most of the students answered this question correctly. There was also a question (the “taxi problem”) where the calculation of a probability was required that involved taking into account both the base rate of the taxi colours and the rate of correct identification of the colours by an eye witness. There was a tendency to answer this question more thoroughly as age increased, but only one student was completely correct and only students in Grade 9 and above attempted to quantify their answers. The other question was the “urn problem” described in the next paragraph.

The urn problem, originally described by Falk (1986), was used in a study by Fischbein and Schnark (1997). In this problem an urn was described that contained two white and two black balls. The first question asked for the probability of taking a white ball from the urn given that one white ball has already been removed. The second question then asked for the probability that the first ball was white, given that it is known that the second ball was white. The most common reasoning used in the second question was that the second ball could not affect the outcome of the first ball; therefore the probability was one half. This is a result of misplaced causal reasoning (known as the time axis fallacy), in that whereas the probability of the colour of the second ball depends on that of the first ball, the reverse is not the case (Falk, 1986). In general, people do not see that later knowledge can be used to revise the calculation of the probability of the colour of the first ball. Fischbein and Schnark’s study involved 20 students from Grade 5, 20 students in Grade 7, 20 students in Grade 9, 20 students in Grade 11 and 18 college students who were training to be mathematics teachers. It was found that apart from the college students, the frequency of the time axis fallacy increased with age. Of all the students, 77% answered the first part correctly, and 37% of the students answered the second part correctly. In the study by Watson and Moritz, 72% of students answered the first part of the question correctly and 34% of the students answered the second part correctly.

The aim of this study was to examine tertiary students understanding of conditional probability in different contexts on entry to an Australian university.

Methodology

This study was part of a larger study on students' knowledge of probability and statistics at the beginning of an introductory one-semester applied statistics unit at an Australian university (Reaburn, 2011). This study took place over four teaching semesters, each with a new cohort of students. The students were studying computing, biomedical science, sports science and aquaculture. During the first week of the semester, a total of 75 volunteers over the four cohorts answered a series of questions some of which were required interpretation and calculation of questions that involved conditional probabilities (Figure 1). The students were also asked whether or not they had experience in statistics in their previous mathematical studies. All the students had previously either successfully completed a pre-tertiary mathematics unit or successfully completed a bridging calculus unit.

1. The table below shows the number of defective TV's produced every week at two factories by the day shifts and by the night shifts.

	Factory A	Factory B
Day	40	30
Night	40	60

a. How many defective TV's are produced at Factory B every week?
b. How many defective TV's are produced by a night shift every week?
c. If you were told that a defective TV was produced at Factory A, what is the probability it was produced by a day shift?
(From Watson & Kelly, 2007)

2. Which probability do you think is bigger?
a. The probability that a woman is a schoolteacher
OR
b. The probability that a schoolteacher is a woman.
c. Both (a) and (b) are equally likely.
Please explain your answer.
(From Watson & Kelly, 2007)

3. An urn has 2 white balls and 2 black balls in it. Two balls are drawn out without replacing the first ball.
a. What is the probability that the second ball is white, given that the first ball was white? Please explain your answer
b. What is the probability that the first ball was white, given that the second ball was white? Please explain your answer.
(From Falk, 1986)

Figure 1. Questions related to conditional probability.

The responses were coded according to the level of statistical reasoning shown by the students, where more sophisticated forms of reasoning received higher scores. These scores were then used to carry out a Rasch analysis on the entire questionnaire according to the Partial Credit Model (Masters, 1982). The Rasch model provides independent ratings of students' ability and difficulty of the items on unidimensional scales (in logits with a mean of zero) and in this study was used to compare the difficulty of each item for the

entire questionnaire. The median scores were compared for each question between students who indicated that they had previous statistical experience and those who did not via the Mann-Whitney U test. The Kruskal-Wallis H test was used to compare the median scores for each question between students of the four cohorts. Parametric tests were not used as the assumptions of normality and homogeneity of variance could not be met because the data were in the form of scores from zero to two. These scores were also used to determine if significant correlations existed between the students' performance on these questions.

Results

For all the questions, there was no significant difference between the median scores between the semesters, and no significant difference between the median scores of each question for those who did and did not claim previous statistical experience. In addition, there were no significant correlations between the questions.

Question 1

For the conditional probability part of this question, part (c), 87% of the answers were correct. With this number of correct responses, the question did not fit on the unidimensional scale of the Rasch analysis and an actual difficulty rating was not obtained.

Question 2

The scoring of this question was as follows. A score of '2' was given to those who chose (b) as the answer and also explained their reasoning appropriately. Those responses that chose (b) and used personal experience as their justification received a code of '1.' Overall the question rated at 0.61 logits below zero, and it was one of the easiest questions on the entire questionnaire. The types of reasoning used are described in Table 1.

Table 1
Forms of reasoning used to answer Question 2

Form of reasoning used (score)	Number of responses
There are two choices for a teacher, male or female, but there are many occupations for a woman to choose (Score 2).	32
The probability of a teacher being a woman is approximately 50%, the probability of a woman being a teacher is much less (Score 2).	13
There are more woman than schoolteachers (Score 1).	9
In my school the majority of the teachers were women (Score 1).	2
Both (a) and (b) are the same (Score 0).	10
No response (Score 0).	6
Idiosyncratic ^a (Score 0).	3

^a Idiosyncratic indicates that the answer had no relationship with the question or was unintelligible

Question 3

For both parts (a) and (b) a score of '2' was given to correct responses with an appropriate explanation. A score of '1' was given to correct responses with no explanation, or to those who calculated the joint probability of getting both balls to be white (1 in 6).

Part (a). In part (a) 84% of the responses received one of these two codes. This question, with a rating of 0.70 logits below zero, was one of the easiest questions on the entire questionnaire. The types of reasoning used are described in Table 2.

Table 2
Forms of reasoning used to answer Question 3(a)

Form of reasoning used	Number of responses
1 in 3 – There are three balls left, one of which is white (Score 2).	57
1 in 6 – Calculated a joint probability (Score 1).	8
Other numerical answer, excluding 1 in 6 (Score 0).	7
No response (Score 0).	3

Part (b). In contrast to part (a) only 19% of the responses received a score of ‘1’ or ‘2.’ This question was the hardest of all for the entire questionnaire, with a rating of 1.80 logits above zero, 2.5 logits above the rating of Question 3(a). The types of reasoning used are described in Table 3.

Table 3
Forms of reasoning used to answer Question 3(b)

Form of reasoning used	Number of responses
1 in 3 – There are 3 other balls, one of which is white (Score 2).	3
1 in 3 – Drew a tree diagram (Score 2).	1
1 in 3 – I am not sure why (Score 1).	5
1 in 6 – Calculated a joint probability (Score 1).	3
1 in 2 – The result of the second draw cannot affect the first draw (Score 0).	30
1 in 4 – There are four balls to start with (Score 0).	10
No response (Score 0)	23

Discussion

These questions were chosen as they do not require formal probabilistic calculations in their answers. Whereas the numerical information would be required to answer questions 1 and 3, no formal probability theory would be required. Even though 76% of the participants claimed previous knowledge of statistics, only one of them used a formula in any of the answers. This participant unsuccessfully attempted to use Bayes’ theorem for part (b) of Question 3.

These three questions required different forms of conditional probabilistic reasoning. The first question required participants to identify the condition that the sample space had been reduced to Factory A, and then to select those that had been produced in the day shift. Most of the participants (87%) were able to do this correctly. One student selected the items of interest but used the whole sample space as the denominator (40/170), and two students incorrectly reduced this fraction to 4/7. This is a higher correct response rate than with the study by Watson and Kelly (2007) which is not surprising as all these students had

either completed a Year 12 pre-tertiary mathematics subject or successfully completed a calculus bridging unit.

The second question required participants to be able to distinguish between $P(W|S)$ and $P(S|W)$ in written form. Of the participants, 13% could not determine the difference. This compares with 51% of the participants who could not determine the difference in Watson and Kelly's study. This reduction in proportions is again not surprising considering the type of participants involved.

The responses to the third question indicate the presence of the time-axis fallacy in the participants. Most of the students answered part (a) correctly, whereas few could answer part (b) correctly. Most of the participants used arguments based on the situation at the beginning, for example, "In the beginning, there are 2 white and 2 black." A small number showed signs of causal reasoning as described by Falk (1986), for example, "The first ball doesn't affect the drawing of the second ball." These students did not see that knowledge of later events can affect a person's knowledge of earlier events. The difficulty of part (b) not only is shown by the low number of correct responses and the larger number of non-responses but also by the reasoning used in the answers. Of those who did answer correctly, most indicated that they were unsure of the answer or did not give an explanation. Some of the participants who answered 1 in 2 were also unsure of their answers. For example, "2/4 - not sure, the fact the second was white may decrease the chance of the first being white but I don't know by how much." Falk indicates that this question "usually triggers a lively discussion in class" (p. 292) with some students refusing even to think about the problem; this is the experience of this researcher as well.

Confusion between a probability and its inverse is common (Falk, 1986). The inability to discern the difference may lead to confusion in the interpretation of Null Hypothesis Tests (Falk), in understanding the true situation for positive results in screening tests for diseases such as HIV (Shaughnessy, 1992), and in understanding Bayes' theorem. Whereas most students will not study inferential statistics or be in a profession where knowledge of conditional probabilities is important, this lack of understanding may become important as they become older and undergo screening tests themselves. It may also be important when these students go on to view or read media reports that are deliberately alarming or are alarming as a result of a journalist's poor understanding of statistics (Blastland & Dilnot, 2008).

Implications for Teaching

Rittle-Johnson, Siegal and Alibali (2002) define procedural knowledge as the "ability to execute action sequences to solve problems" (p. 346) in contrast to conceptual knowledge defined "as implicit or explicit understanding of the principles that govern a domain and the interrelation between units of knowledge in a domain" (pp. 346-347). This conceptual knowledge is "flexible and not tied to specific problem types and is therefore generalizable" (p. 347). These two ideas are related to the ideas of surface learning, where the content of a text maybe just reproduced, and deep learning, where the student can relate parts of the material to the whole, integrate it with existing knowledge and apply it in real world situations (Boulton-Lewis, 1995). If students are encouraged to solve problems in a variety of ways their progress to conceptual understanding can be enhanced (Moreno & Duran, 2004). Question 1 was solved by the participants in this study without the formal formula; encouraging students to solve a problem such as this with the formula as well can improve their ability to generalise. Similarly, once students can solve part (b) of Question 3 by informal means they may be able to see how the problem can be solved with Bayes'

theorem. Questions without a specific numerical value such as the schoolteacher problem above can also assist students in developing a conceptual understanding of the nature of conditional probability.

Using various ways of solving problems not only can encourage students to make connections and enhance conceptual understanding, but also can assist students' learning of the fundamental principles. What students find difficult with one method may be found easier using another method. Gigerenzer and Hoffrage (1995) indicate that "statistically naïve participants" had much more success with Bayesian problems when the information was presented in frequency formats. An example is found in Table 4.

Table 4

A conditional probability question posed in probability and frequency formats

Question in probability format	Question in frequency format.
1% of women at age forty who have a mammogram have breast cancer.	10 out of every 1,000 women at age forty who have a mammogram will have breast cancer,
80% of these women who have cancer will get a positive mammogram	8 out of every 10 women with breast cancer will have a positive mammogram,
9.6% of these women without cancer will get a positive, mammogram	95 out of every 990 without breast cancer will get a positive mammogram.
What is the probability that a woman who had a positive mammogram actually has cancer?	Here is a representative sample of women at age 40 who had a positive mammogram. How many do you expect will actually have breast cancer?

Another method is to convert the question into a two-way table. Pfannkuch, Seber and Wild (2002) have found students find it easier to understand such questions when presented this way. Table 5 shows the same question presented in this way.

Table 5

Mammogram question in a two-way table where the data are presented as probabilities

	Positive test	Negative test	Total
Has breast cancer	0.008	0.002	0.01
Does not have breast cancer	0.0950	0.895	0.99
Total	0.103	0.897	1.000

At the Year 10 level The Australian Curriculum Mathematics states that a student should "Use the language of 'if...then', 'given', 'of', 'knowing that' to investigate conditional statements and identify common mistakes in interpreting such language (ACMSP247)" (Australian Curriculum Assessment and Reporting Authority (ACARA), 2012a). The elaboration goes on to mention two-way tables, Venn diagrams and tree diagrams. The proposed curriculum for Mathematical Methods states that students should "Understand the notion of a conditional probability and recognise and use language that indicates conditionality" and "Use the notation $P(A|B)$ and the formula $P(A \cap B) = P(A|B)P(B)$ " (ACARA, 2012b). This formula, as well as two-way tables, Venn diagrams and tree diagrams are all useful tools in calculating conditional probabilities. The challenge for busy teachers is to encourage students to develop deep learning.

Acknowledgements

I would like to thank Emerita Professor Jane Watson for her valuable advice in reviewing this paper.

References

- Australian Curriculum Assessment and Reporting Authority (2012a). *The Australian Curriculum, Mathematics*, Version 3.0. Australia: ACARA.
- Australian Curriculum Assessment and Reporting Authority (2012b). *The Australian Curriculum, Mathematical Methods*, retrieved from:
<http://www.australiancurriculum.edu.au/Static/docs/senior%20secondary/Senior%20Secondary%20Curriculum%20-%20Mathematical%20Methods%20November%202012.pdf>
- Blastland, M., & Dilnot, A. (2008). *The tiger that isn't: Seeing through a world of numbers*. Great Britain: Profile Books.
- Boulton-Lewis, G. (1995). The SOLO taxonomy as a means of shaping and assessing learning in higher education. *Higher Education Research and Development*, 14(2), 143-154.
- Falk, R. (1986). Conditional probabilities: Insights and difficulties. In R. Davidson & J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 714-716). Victoria, BC: University of Victoria.
- Fischbein, E., & Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? An exploratory research study. *Educational Studies in Mathematics*, 15(1), 1-24.
- Fischbein, E., & Schnark, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education*, 28(1), 96-105.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning with instruction: Frequency formats. *Psychological Review*, 102(4), 684-704.
- Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149-174.
- Moreno, R., & Duran, R. (2004). Do multiple representations need explanations? The role of verbal guidance and individual differences in multimedia mathematics learning. *Journal of Educational Psychology*, 96(3), 492-503.
- Pfannkuch, M., Seber, G., & Wild, C. (2002). Probability with less pain. *Teaching Statistics*, 24(1), 24-30.
- Reaburn, R. (2011). *Students' understanding of statistical inference: Implications for teaching*. (Doctoral thesis). University of Tasmania, Australia.
- Rittle-Johnson, B., Siegler, R., & Alibali, M. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93(2), 346-362.
- Shaughnessy, J. (1992). Research in probability and statistics: Reflections and directions. In D.A. Grouws (Ed.), *Handbook of research in mathematics teaching and learning* (pp. 465-494). New York: Macmillan.
- Tarr, J., & Jones, G. (1997). A framework for assessing middle school students' thinking in conditional probability and independence. *Mathematics Education Research Journal*, 9(1), 39-59.
- Watson, J., & Kelly, B. (2007). The development of conditional probability reasoning. *International Journal of Mathematics Education in Science and Technology*, 38(2), 213-235.