Problems with Probability

Robert Peard
Queensland University of Technology

This research examines misconceptions in probability held by a sample of pre-service primary teacher education students. Questions were selected and modified from those reported in the research literature in order to examine; the misuse of heuristics, a false assumption of equal likelihood, misunderstanding of independence, awareness of counter-intuitive probabilities, and belief in other fallacies. Questions were accompanied with a request to explain the reasoning employed. Explanations were partitioned into disjoint categories according to the type of misconception and cognitive level of response. Results of selected questions and a summary of the response levels are presented here for discussion.

Rationale

Recent curriculum developments in primary school mathematics have seen a much greater emphasis on the role of probability in the classroom. The importance of understanding probabilistic concepts in modern technological societies is now well established but the inclusion of probability into the mathematics curriculum is a relatively recent development. Furthermore the teaching of the topic is "a very difficult task, fraught with ambiguity and illusion" (Garfield & Ahlgren, 1988, p. 57) and studies have shown that the content knowledge of topics in probability and statistics for both primary and secondary teachers is often deficient. (See for example, Peard, 1987; Shaughnessy, 1992).

Thus it is contended that pedagogical problems in teaching probability are compounded by lack of teacher knowledge of content and the presence of misconceptions. In order for the effective use the limited time available for the topic of probability in the content area of pre-service primary teacher education courses, research is required to determine what conceptions and misconceptions our pre-service primary teacher education students bring with them. Watson (1992, p. 560) has suggested that "it may be possible to use the results of research to make more precise and confident statements to help teachers and students understand and apply the concepts in the Chance and Data statement." She expressed concern that recent initiatives in curriculum development "have been taken without the benefit of previous educational research in Australia on the learning of probability" (p. 556).

Aims

This research aims to determine the nature and extent of misconceptions in probability held by second year pre-service primary teacher education students enrolled in the B.Ed course at the Queensland University of Technology, Semester 2, 1995. These misconceptions include:

- the use of the heuristics of representativeness and availability;
- the false assumption of equal likelihood;

- independence of events;
- awareness of counter-intuitive probabilities; and
- belief in other fallacies.

The study examines the types of reasonings associated with use of these misconceptions in order to "search for alternatives ... to explain peoples' flawed probability estimates" (Shaughnessy, 1992, p. 471).

Background to the Study

The difficulties pupils encounter with probabilistic concepts have been confirmed by the results of a number of recent studies. For example, in the Fourth National Assessment Education Program, (NAEP), Brown, Carpenter, Kouba, Lindquist, Silver and Swafford (1988) reported widespread "difficulty with items involving probability, except those involving very simple concepts" (p. 242) and that "only 5% of (Year 11) students correctly answered questions involving compound probabilities" (p. 243). In examining areas of difficulty in probabilistic reasoning, Garfield and Ahlgren (1988) reported that some concepts that are difficult because they are unlike anything the student has thought of before and other concepts cause difficulties because they run counter to intuitive ideas that the students already have. Fischbein, Nello, & Marino, (1991) commented, there appears to be "no natural intuition for evaluating the probability of a compound event" (p. 534). The present study examined some of the students' beliefs about probability prior to instruction in their second year mathematics subject.

Some Misconceptions in Stochastics

Tversky and Khaneman (1982) claimed that most misconceptions in probability concepts among adults could be attributed to one of two heuristics; *representativeness* and *availability*. Shaughnessy (1981) also reported that misconceptions used in estimating probabilities by college entry students could often be attributed to these heuristics. However, not all misconceptions can be attributed to either of these heuristics, nor can the use of heuristics explain some of the types of reasonings associated with misconceptions. The assumption of equal likelihood in situations where none exists is one example. Shaughnessy (1992) summarised research in stochastics. In addition to discussing the use of the heuristics of *representativeness* (p. 470) and *availability* (p. 472), he examined research into:

The base-rate fallacy. Shaughnessy (1992) reports that the phenomenon is often attributed to representativeness, but claims that other factors such as "fundamental attribution error .. or some sort of causal reasoning" (p. 472) may be the real cause of the errors of probability estimates.

The failure to recognise independence. This is sometimes attributable to representativeness in, for example, the gambler's fallacy. In other instances it may be attributable to availability such as, for example, the selection of Lotto numbers. However some occurrences may not be attributable to either of these heuristics and the present research explores this possibility.

The assumption of equal likelihood when none exists. Again, this is often attributed to some other factor or heuristic. However, the present research examines instances where this may not be possible and looks for alternative explanations.

The confusion between a conditional and its inverse. Shaughnessy (1992, p.474) cites the example "the difference between the probability that I have measles given that I have a rash, and the probability that I have a rash given that I have measles."

The conjunction fallacy. Shaughnessy (1992, p. 473) cited research studies in which students were reported to confuse situations of conditional probability P (A I B) with the conjunction P (A n B) and reported that "the problem seems to occur primarily with the students' translation of conditional probability tasks, which then affects their understanding of the problem." He reports that many of the studies rely on a "forced-choice task methodology ... and were unable to obtain detailed insight into why the students were confusing conditionals with conjunctions."

The awareness of counter-intuitive probabilities. Instances of this too are often attributed to some other factor or heuristic, such as representativeness in the case of the "birthday problem".

Methodology

A set of 12 data questions was selected and modified from those reported in the research literature in order to examine the above misconceptions. This set of data questions formed the *major research instrument* of the study. Many previous research studies into misconceptions in probability have used a multiple choice format. (See, for example, Brown et al., 1988; Fishbein et al., 1991). Shaughnessy (1992, p. 479) comments that information from multiple choice items is often sketchy, incomplete and "not very helpful in clarifying students' thinking." Thus in the present research, all questions, whether multiple choice or open response, are open-ended and are accompanied with a request to explain the reasoning employed.

The questions were given to a sample of 50 students enrolled in Year 2 Mathematics Curriculum. Most of the students had completed a Year 12 mathematics subject which included some studies in probability. The analysis of the responses in the present study identified four levels according to level of sophistication in a similar manner to that of Watson & Collis (1993).

Level 1. In interpreting probability situations no analysis or evidence of use of probability principles is demonstrated. Features may include: the use of irrelevent information, subjective judgements, disregarding quantitative information, guessing at random, belief in control of probability and absence of any reason. Responses that use recent experiences to predict or estimate probabilities, availabilty, are included in this level.

Level 2. Some evidence of use of the use of probability principles and appropriate quantitative information is evident, but they may be incomplete or are incorrectly used. Probabilistic reasoning based on the assumption of equal liklihood when none exists and The use of the representativeness hueristic is considered to be illustrative of this level.

Level 3. Probability principles are used correctly used and an awareness of the role of quantification is evident. However, such quantification is precise or numerical.

Level 4. Probability principles are used correctly and relationships are explained quantitatively.

[Note, using this classification it is possible that "correct" answers can be accompanied with low-level or incorrect reasoning while "incorrect" responses may still show some high level reasoning.]

Selected Questions, Results and Discussion

Question 1

An unbiased coin is tossed five times in a row and lands *heads* each time.
On the next toss:

(a) *tails* is the more likely outcome
(b) *heads* is the more likely outcome
(c) *heads* and *tails* are equally likely.

Explain you choice: ______________________________

Results for Question 1. Only two students demonstrated the use of representativeness. Thirty stated that the coin was unbiased or that the probability remained the same (*L3*). Eighteen quantified the probabilities as 1/2 of 50% (*L4*).

Question 2

On a roulette wheel there are 18 red numbers, 18 black numbers and one green (zero). On any throw, each number is equally likely to occur. Suppose that the ball lands on "red" six times in a row. On the next roll is it now more likely to land on the "red", the "black", or are they both still equally likely? Explain your response: ______________________________

Results for Question 2. Four students used the representativeness heuristic to conclude that black was now more likely.
Discussion of Question 2. The question explores what is referred to in the literature as "the gambler's fallacy." (del Mas & Bart, 1989, p. 45) This misconception is well documented as an example of "local representativeness" (Tversky & Kahneman, 1982, p. 5). Research by Shaughnessy (1981) of its use by college students in his study showed a much higher incidence than that of the present study. However, of the remainder who recognised equal probabilities only three quantified the probability as 18/37, (*L4*).

Question 3

In a six-from-40 LOTTO,
Jim chose: 1, 2, 3, 4, 5, 6
Mary chose: 32, 33, 34, 35, 36, 37
and Bill chose: 9, 12, 27, 31, 35, 38
Does any one have a better chance of winning than the others? If so who? Why?

Results for Question 3. Eleven students said Bill had the better chance. Of these six gave reasons that showed a clear use of availability while the other five gave no specific reason, although they may nevertheless have been using the heuristic.
Discussion of Question 3. There is a common misconception that apparently random combinations such as Bill's are more likely than sequences or patterns. This misconception has been widely reported. Its use by up to 22% of respondents in the present study is consistent with that reported in the literature.

Question 6

If I throw a pair of fair dice is it harder to get a pair of sixes than it is to get a five and a six? Why or why not?

Results for Question 6. Only 17 recognised that a pair was less likely. Of these only three quantified the probabilities (L4). Five cited experience with dice as the reason.
Discussion of Questions 6. Fischbein et al. (1991) and Peard (1994) used these questions in researching pupils understanding of compound events. Fischbein et al. concluded that "there is no natural intuition for evaluating the probability of a compound event" (p. 534). The results in the present study of 17/50 or 34% correct would at first appear to be much better than those reported in the literature. However only 3 or 6% demonstrated quantitative reasoning. Peard (1994) reported that 10/ 40 or 25% of Year 11 students in his study answered correctly, six of these citing experience and four using computation.

Question 7

The spinner shown (Figure 1) is spun once.

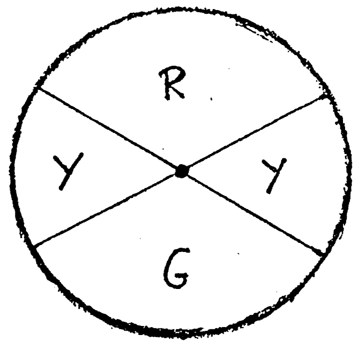

Figure 1

What is the probability that *yellow* will show? Explain your reasoning.

Results for Question 7. Thirty six responded 1/2 or 2/4. Thirteen answered 1/3, and one answered 2/3.
Discussion of Question 7. This question examines the misconceptions relating to the assumption of equal likelihood when none exists. Its use by such a high proportion of students is significant.

Question 9

A group of 30 people are selected randomly from the population.
The probability that at least two of them will have the same birthday is:

(a) zero
(b) very low
(c) more than 50%
(d) more than 90%

Explain your choice ______________________________

Results for Question 9. Forty five students selected (b). Of these, 23 gave an incorrect quantification (*L2*); 19 gave no real reason (*L1*); and three reasoned that the ratio of the number of people to the number of days in the year was low (*L3*).

Discussion of Question 10. Shaughnessy (1992) reported that most people are surprised to learn that the probability is high (70%) since this is counter-intuitive. Tversky and Khaneman (1982) cite it as an example of the use of the representativeness heuristic.

Question 11

In a pick-a-box quiz. there are three boxes. One box contains the prize and the other two are empty. You are a contestant who gets to select a box. After making your selection, the host, Monty, opens one of the two boxes to show that it is empty. (Monty does this with all contestants and Monty always knows which box holds the prize). Monty asks you whether you want to stick with your choice or change your selection.

Which of the following do you do now?

(a) Stick with your original choice
(b) Change you selection to the other box
(c) Toss a coin to decide
(d) Something else.

Explain your choice ______________________________

Results for Question 11. Forty students selected (a). Of these 23 reasoned new probabilities of 1/2 (*L2*), while 17 were unable to formulate any reason (*L1*).

Discussion of Question 12. Shaughnessy (1992, p. 475) has used this problem extensively with students, including pre-service teachers, to point out some of the difficulties with conditional information. He reported similar results. Various forms of this misconception have been recorded as a paradox in probabilistic reasoning. The probability of guessing the right box is clearly 1/3. Thus if you decide to stick with your original choice, this probability remains 1/3. However, if you use the strategy of changing your choice *after* Monty has opened one of the empty boxes (conditional probability), your probability of winning improves, not to 1/2 but to 2/3. This result is clearly counter-intuitive, but can readily be explained by the fact that using the strategy there is only one out of three ways the contestant can lose; that is by choosing the winning box initially.

Question 12

A deck of 21 green and yellow cards are shuffled and dealt twice.
The following sequences are observed:

A: Y Y Y G G Y Y G G Y Y Y G G G G G G Y Y Y
B: G Y G Y G Y G Y Y G Y G Y Y G Y G G Y G Y

Which sequence would represent the better shuffled (more random) deck?
A, B, or are both equally random? Why?

Results for Question 12. Twenty three students assumed equal likelihood. Of these 14 gave a Level 2 reason, while nine gave either no reason or used irrelevant information such as "the number of cards is the same" (L1). Six correctly recognised that A is "too regular".

Discussion of Question 12. It can be shown that if a coin is tossed *n* times, then there occurs a run of heads of length log2 *n* with a probability converging to 1 as *n* approaches infinity. Applying this theorem to the deal of 21 cards, a run of 4 or 5 yellow or green cards is to be expected. (log2 21 = 4.4). Thus sequence A is more likely than B. Falk (1981) asked a similar question to secondary students and reported that students tended to pick the sequences where more "switches" occurred. The responses of the tertiary students in the present study are similar to those reported by Falk (1981).

Summary

It would appear that the level of misconception in the use of the representativeness and availability heuristics, independence of events, compound probability and the lack of awareness of counter-intuitive probabilities are widespread and exhibited by the students in much the same proportion as is reported in the literature. The assumption of equal likelihood when none exists was much higher than expected. This misconception was evident in Question 6, 36%; Question 8, 72%; Question 12, 46%; and Question 13, 46%. Furthermore an analysis of the level of responses shows that on relatively few occasions did the students use quantitative or relational reasoning with only 17% of all responses categorised as Level 4. Probability principles were correctly used in another 23% of all responses categorised as Level 3. A full 60% of responses were made without evidence of analysis or with an incorrect use of probability principles.

Bearing in mind that the majority of students have had some formal education in probability further education as a part of their pre- service teacher education is clearly needed. It cannot be assumed that they are necessarily free of the confusion and misconceptions that permeate this difficult topic.

References

Brown, C., Carpenter, T., Kouba, V., Lindquist, M., Silver, E., & Swafford, J. (1988). Secondary school results for the fourth National Asessment Education Program, maths assessment. *Mathematics Teacher, 81* (4), 241-248.

del Mas, R. C., & Bart, W. M. (1989). The role of an evaluation exercise in the resolution of misconceptions of probability. *Focus on Learning Problems in Mathematics, 11*(3), 43-53.

Falk, R. (1981). The perception of randomness. In Proceedings of the *Third International Conference for the Psychology of Mathematics Education* (pp. 64-66). Warwick, U.K.

Fischbein, E., Nello, M., & Marino, M. (1991). Factors affecting probabilistic judgements in children and adolescents. *Educational Studies in Mathematics, 22*(6), 523- 549.

Garfield, J. B., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education, 19*(1), 44-59.

Peard, R. (1987). Teaching statistics in Queensland: Qualifications and attitudes of teachers. *Teaching Mathematics, 12*(4), 13-18.

Peard, R. (1994). *The effect of social background on probabilistic reasoning*. Unpublished doctoral thesis. Deakin University. Geelong.

Shaughnessy, J. M. (1981). Misconceptions of probability: From systematic errors to systematic experiments and decisions. In A. P. Shulte (Ed.), *Teaching statistics and probability* (pp. 90-100). Reston, VA: National Council of Teachers of Mathematics.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). New York: MacMillan.

Tversky, A., & Kahneman, D. (1982). Judgement under uncertainty: Heuristics and biases. In Kahneman, D., Slovic, P. & Tversky, A. (Eds.), *Judgement under uncertainty: Heuristics and biases* (pp. 3-22). Cambridge: Cambridge University Press.

Watson, J. M. (1992). What research is needed in probability and statistics education in Australia in the 1990s? In B. Southwell, B. Perry & K. Owens (Eds.), *Space- the first and final frontier. Proceedings of the Fifteenth Annual Conference of the Mathematics Education Research Group of Australiasia* (pp. 556-567). Kingswood: Mathematics Education Research Group of Australasia.

Watson, J. M., & Collis, K. (1993). Initial considerations concerning the understanding of probabilistic and statistical concepts in Australian students. In W. Atweh, C. Kanes, M. Carss & G. Booker (Eds.), *Contexts in mathematics education. Proceedings of the Sixteenth Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 575-580). Brisbane: Mathematics Education Research Group of Australasia.