Probability for Kids

Using Model-Eliciting Activities to Investigate Probability Concepts

Scott A. Chamberlin, Ph.D.

Hawker Brownlow
EDUCATION

Table of Contents

Introduction

This book is unlike any other on the market, for several reasons. Aspirant probability experts may commonly engage in the study of a book with the intent of significant learning. However, more often than not, afterward readers are more confused than they were prior to reading the book. The intent of this book is to provide activities that will help readers progress from being confused about probability to being informed about probability. As Winkler (1996) stated, probability is ubiquitous. Each day, decisions are made about unimportant to terribly important outcomes in life – and, generally, such decisions centre on probability. Some common decisions may be:

- Should I take an umbrella to work/school today?
- Should I go out to eat or eat at home?
- Should I go to university?
- Should I ask my partner to marry me?
- Should I buy the business from my business partner?
- Should I try out for the cricket team?

A second reason for another book on probability is that most of the other books on probability are outdated. Although the concepts in probability have not changed dramatically in the past few decades, the methods of introducing such concepts are age old (in some cases up to a century or more). The respective years of publication for competing books are 1994, 1969, 1966 and 2010. Minus the last book, the most recent book on probability was printed more than 20 years ago. Times have changed and some teaching and learning needs may be better met by up-to-date approaches. This is not to suggest that

Characteristics of MEAs

MEAs are powerful for various reasons. For example, MEAs have multiple entry points (Chamberlin, 2002) and are accessible by constituents at various levels of mathematical understanding. Thus, MEAs have great potential in magnet schools, pull-out programs and inclusive classrooms. The latter category may be the majority of environments in which gifted and talented mathematics students are served.

In addition, as Chamberlin's (2002) line of research suggests, MEAs have been used with gifted and talented students perhaps more than any other group of learners. Hence, many of the six principles have been implicitly adapted with gifted and talented students in mind.

Recent education standards increasingly demand that mathematical modelling be infused in F–12 classrooms. However, few teachers or learning facilitators have been provided with any training to meet such an expectation. This is where MEAs rescue many individuals looking to meet increased demands. They have been designed and piloted for approximately four decades and will help facilitators infuse mathematical modelling in a systematic, student-driven manner. To date, no printed materials exist in one central repository that expects learners to create mathematical models to make sense of mathematical phenomena or situations.

One of the most significant reasons that MEAs were chosen for this book is that they facilitate pre-university level thinking in mathematical situations (Lesh et al., 2000). Young problem solvers cannot be expected to use algorithms their entire school career and then be successful magically creating mathematical models when they get to the workforce and/or the tertiary level. The emphasis on mathematical modelling at the primary and secondary level carries over to higher level undergraduate and graduate courses at the university level. To expect promising mathematics students to know how to create mathematical models in college-level mathematics courses without adequate exposure to such demands prior to university is unreasonable.

MEAs are incredibly real world in nature. Consequently, they may be used simultaneously to engage various types of students in enrichment *and* acceleration. MEAs are in fact so realistic that on several occasions, problem solvers have asked about the status of the client or person for whom the problem is being solved. Such marketing or salesmanship may precipitate high-quality products on behalf of problem solvers because they feel they are truly solving an authentic problem-solving task.

Chapter 1

Likelihood

The Junior Secondary School Problem

For some students, the Junior Secondary School Problem is the introductory model-eliciting activity that they will do in probability. Consequently, the content and context are rather important. This problem, written specifically with advanced Years 4–6 students in mind, but aptly challenging for students up to Year 9, should sync well with their interests, as many of them will be making the transition from upper primary to junior secondary school in the coming months or years. Identifying a focus in probability for high-ability students was difficult in the respect that the problem needed to be sufficiently, but not overly, challenging. That is, on either end of the developmental spectrum, a problem may pose barriers for students. If the problem is too difficult for the students, then they may become frustrated and disengaged. If, on the other hand, the problem is too pedestrian for them, they may not take it seriously and produce less-than-satisfactory solutions. The hope with the Junior Secondary School Problem is that the context and content meet expectations and provide ideal challenge for upper primary high-ability students and middle year general population students.

The focus of the Junior Secondary School Problem is to get students to quantify the probability of an event transpiring. More specifically, this MEA is about compound probability because two events must occur, that is, both friends being placed in the same classroom. At its most basic definition, *probability is the likelihood of an event happening and it is represented by the number of favourable outcomes (numerator) relative to the number of potential outcomes (denominator)*. This is one of the reasons that simple events are used in introductory probability activities, such as flipping a coin (restricted to 2 outcomes),

Solution 1

The first solution is a strong explanation, but it is confined to only the first homeroom. Consequently, many teachers might desire a more comprehensive mathematical model than this one and encourage problem solvers to explain what could happen in homerooms after the first one. The first solution reflects the idea that there is a 1 in 100 chance of the two friends being placed in the same class. According to the problem solvers, this is the case because there are 10 homerooms and each boy has a 1 in 10 chance of being selected for the first homeroom.

The problem solvers therefore suggested that the probability of both boys being placed in the same homeroom is 1/10 for Maurice and 1/10 for Marvin or $(1/10)^2$. Other problem solvers may represent this as 1/10 multiplied by 1/10. Nevertheless, the probability of the event happening is 1 in 100, 1/100, or it can be expressed as 1%. These are not great odds and some problem solvers might suggest that with a much smaller school and fewer homerooms, the likelihood of being placed together would be increased.

Solution 2

The second solution is only somewhat more complex than the first solution. This solution represents the model expressed in the first solution – 1 in 100 – but with this model, the problem solvers added the caveat that if either of the boys were selected for one homeroom and the other was not, the conditions of the problem could not be met. The second solution provided an example to substantiate this claim. Solution 2 suggests that if Marvin was placed in the first homeroom and Maurice was not, then compound probability could not be met in this case because the problem statement prompted individuals for the likelihood of both boys being placed in the same homeroom. So, if Marvin were placed in the first homeroom and Maurice in the sixth homeroom, they would not be together. This solution does not specify the precise probability of this event transpiring, but does acknowledge that it was far more likely that the boys were separated than together in one homeroom.

Solution 3

With this solution, the problem solvers realise that the probability of being placed together could be quantified quite precisely if certain conditions were met. Each homeroom could be given an actual probability if the boys were not separated in the drawing. Figure 1.1 shows the probability for each homeroom, assuming no separation in the drawing.

Homeroom	Probability of Being Together	Probability of Being Separated
1	$(1/10)^2$ or 1.00%	99.00%
2	$(1/9)^2$ or 1.23%	98.77%
3	$(1/8)^2$ or 1.56%	98.44%
4	$(1/7)^2$ or 2.04%	97.96%
5	$(1/6)^2$ or 2.78%	97.22%
6	$(1/5)^2$ or 4.00%	96.00%
7	$(1/4)^2$ or 6.25%	93.75%
8	$(1/3)^2$ or 11.11%	88.89%
9	$(1/2)^2$ or 25.00%	75.00%
10	$(1/1)^2$ or 100%	0.0%

Figure 1.2. *Probability and inverse probability of each homeroom.*

Instructor Questions

Imperative to the success of MEAs is anticipating instances in which problem solvers may reach an impasse and not be able to successfully construct a suitable mathematical model. These questions might help problem solvers clarify thinking, change directions altogether and/or come to a refined solution. Appropriate use and timing of questions can help problem solvers attain considerable insight as to the solution of the problem and help them create strong mathematical models.

Three caveats are issued with the instructor questions in each chapter. First, not all questions will be used with the MEA. Second, the list is intended to be comprehensive, but it does not represent all possible questions. Third, the questions are not arranged in any particular order and sometimes an alteration of the wording in the questions can change them dramatically – for better or worse.

1. Can you provide an example of chance in your life?
2. Can you name a situation in which the chance of something happening was 100% or nearly so? Can you name a situation in which the chance of something happening was 0% or nearly so?
3. Does it have applications to this problem?
4. What have you done so far and what does not seem clear to you?

Notes on Implementation

As mentioned previously, teachers implementing the problem must invest substantive time solving the problem in as many ways as possible. Moreover, the timing of questions can help elucidate problem solvers' thinking when they are on the precipice of creating a mathematical model. To that end, some of the questions are designed to help problem solvers create a more efficient model by helping them realise prospective faults in their current mathematical model. In this section, some commentary is provided on each question to help teachers identify situations in which the questions may be used.

With Question 1, the intent is to (hopefully) alleviate some anxiety on behalf of problem solvers by getting them to discuss a real-life situation in which chance was evident. The response by problem solvers may be superficial or somewhat in depth. Nevertheless, simply discussing probability and helping them clarify the relationship between terms such as *probability*, *chance* and *likelihood* might serve to illustrate the notion that creating a mathematical model is accessible by this age group of students.

Question 2 is a follow-up to Question 1. The objective of Question 2 is to push problem solvers to realise that a likelihood of 0% (0 as a decimal) or 100% (1 as a decimal) is in fact quite low. Problem solvers should come to the realisation that all probabilities must be represented as 0 to 1 (or 0 to 100%), as reiterated by Solution 4. By providing a real-life example of a situation in which probability is applicable, problem solvers are also coming to the realisation that probability is a content area in which they may not be formally trained, but that does not suggest that they have no knowledge of the concept. Assuming a very sharp group of problem solvers, they may come to the mathematical realisation that in situations in which the probability is zero, assuming two events, as is the case with this problem, the events must be mutually exclusive ones. In addition, if two events are discussed and a 100% probability exists, the two events must overlap each other perfectly. Although the question does not expect problem solvers to name two events, it only directs their attention to one event. The following Venn diagrams may help students gain understanding about probability. In Figure 1.3, there is no union or intersection of events. In Figure 1.4, there is perfect (100%) union or intersection of the two events.

come out to be the same. If/when they make that realisation, they should be able to quickly come up with the solution that a 12-sided die with 1 000 rolls should result in 83.33 rolls per number (note the .33 is the repeating decimal .33, not .330).

Questions 5 and 6 are intended to help problem solvers operationally define what constitutes the terms similar and quite different. As expressed earlier in the chapter, a main objective of this MEA is to get problem solvers to determine what constitutes a significant difference in two data sets. In this MEA, the concepts of theoretical and empirical probability were used for groups to compare. Groups are asked to quantify a somewhat nebulous concept in identifying whether a significant difference existed or not. This is a prime example of not using formal statistical jargon (statistical significance) because it would likely not resonate with upper primary students.

Question 7 investigates what procedure problem solvers used to identify the theoretical probability in case Question 4 seems too broad. Problem solvers may not provide specifics regarding the procedure on identifying the actual mechanics on theoretical probability for this data set. With Questions 4 and 7, problem solvers should be capable of creating the theoretical probability.

Question 8 pulls problem solvers away from simply the theoretical probability in an attempt to investigate the procedure used to compare the two data sets. This is a rather open-ended question and facilitators may be faced with the prospect of posing another, more direct, question to individual groups. At the heart of Question 8 is: What are you going to do (what process are you going to use) when you compare the two groups?

Question 9 should be answered with a "no" by problem solvers. There is not a pattern in the provided data set, other than that the final seven numerals (6–12) alternate perfectly from below, above, below, above, below, above and below the expected values. This was simply circumstantial in the way the data occurred.

Question 10, on the other hand, has no real pattern, per se, as the data should all be the same for each roll (83.33). Some may consider that a consistent pattern, while others will not consider it a pattern because there is no variation in the numbers. The response to question 10 is contingent upon how one defines a pattern. This is another question that can be used (i.e. what constitutes a pattern?) in the data. It may be best not to dwell on Questions 9 and 10, as patterns are not relevant to the problem. Simply asking such questions may help them realise that patterns will not work as a heuristic to solve this MEA.

Question 11 focuses problem solvers' attention on somewhat abnormal data (either very high or very low). Presumably, data point 4 is most noticeable to problem solvers because it is the only triple digit result. Even though the number of rolls for number 4 (102) is close to data for numbers 7 and 9 (93 and 94 respectively), the triple digit is very noticeable to most problem solvers. In fact, 102 is only 21 rolls away from the expected value. Roll number 12 is almost as far away from the expected value, 16 with 67 total rolls, as is roll number 4 (102).

Questions 12 and 13 may seem identical on the surface, but are slightly different. Question 12 focuses on a broader topic than just mathematical or statistical procedures. Tools may be defined in various manners such as physical or mental. Mathematical or statistical procedures are related, but are a much more specific category of an approach. Procedures may imply something like using measure of spread, central tendency, variability, or a line of best fit.

In considering the aforementioned statistical procedures, Question 14 leads problem solvers to quantify what constitutes "far enough" away from the expected value to earn consideration as a prospective data point that is significantly different than the expected value. Again, data points (rolls) 4 and 12 are perhaps the most noticeable in the data set. The challenge in identifying statistical significance is to look at the entire data set and see if it varies enough as a whole, rather than just looking at certain data points.

Questions 15 and 16 pertain to why the data set may appear as it does. In the respective questions, problem solvers are asked if there is a reason why some numbers were rolled more or less than others. It is hoped that problem solvers will realise that the rolls occurred as they did simply due to circumstance. That is to say, data points that align well with expected values or deviate from them considerably were just due to chance. If the same trial was performed again, there is no reason to believe that the number 4 would have more than 100 rolls.

Question 17 focuses problem solvers' attention on the fact that there is some variation in the numbers rolled, but again this is a follow-up to Questions 15 and 16. Problem solvers should realise that the expected values, or theoretical probability, is what *should* happen, but those events transpiring perfectly is a statistical impossibility if calculated due to the decimal in expected values. Very wise problem solvers may realise that sometimes expected values cannot actually transpire.

Question 18 asks what problem solvers would do with the number 83.33 because they cannot actually get a fraction of a roll. Problem solvers may need to be reminded that the expected values are just that, values that one can expect to get if the problem is solved theoretically. However, with actual rolls, problem solvers will need to realise that there will be a discrepancy in the two data sets given whole numbers in one set and decimals in another.

Question 19 is somewhat generic and intended to get problem solvers to test their mathematical data with the sets and to inquire as to whether their mathematical model would work with another data set that may or may not be similar to those presented here (i.e. theoretical and empirical probability).

In this chapter, problem solvers are asked to create a mathematical model to explain how to identify the probability of several (compound) events happening with dependent probability. The objective is to get problem solvers to realise that early events affect subsequent ones. Moreover, an ancillary concept is to help problem solvers realise how probability changes and needs to be adjusted with each event. To accomplish these objectives, the Volleyball Program Problem is provided.

Potential Student Responses

Facilitators should first solve the problem on their own to see what solution path is chosen. However, having a series of solutions can prove fruitful in anticipating student responses and questions. In this section, the most common solutions or mathematical models are presented.

Solution 1

With Solution 1, problem solvers determine that the probability of winning a prize is 5 out of 100 or 1 in 20. This is not altogether correct or incorrect. Stating that the probability of winning at any given time during the distribution of programs is approximately 5% is a fair generalisation. However, groups are unable to state that the probability changed depending on the previous outcome. This solution does not take into account that every time one person won the probability of winning was decreased by at least one in the numerator (number of favourable outcomes) and one in the denominator (number of overall outcomes). For instance, just before the first person wins, the probability of winning is 5/94 (better than 5/100 or 5 out 100). However, after the first person wins, the new probability of winning is 4/93, which is considerably worse than 5/94. With this solution, problem solvers do not take into account the fact that the most recent event changes the probability of the next outcome. This may suggest that the individuals who created the solution do not have a complete understanding of dependent events.

Solution 2

In Solution 2, the problem solvers appear to grasp the concept of dependent events in probability, although their solution lacks insight as to how this model can be applied to future situations (potential lack of generalisability). Problem solvers get the probability of each event correct and their solution looks something like Figure 5.1.