The concepts, structures and thinking of probability and distributions underpin all statistics, but the development of these in statistics education has received far less attention over the past ten to fifteen years than the topics of what has become known as the “statistics education reform movement”. This reform, built on recommendations [1] for “more data and concepts, less theory, fewer recipes”, and much of the associated research effort, has been oriented to the understanding, skills and thinking of statistical data investigations, both at school and tertiary levels, and the use of statistical data analysis methods and technology in real contexts at the tertiary level. Many statisticians think it is time to bring the principles of data-driven, holistic, authentic and active learning into the development of probabilistic and distributional understanding, reasoning and modelling.

As is apparent in extensive research such as reported in [7], contexts such as coins, dice, spinners and balls in urns, play a valuable role in developing early concepts of chance, particularly for children. However such contexts become increasingly abstract and remote from reality as students progress, as well as prolonged use having the deleterious effects of over-focus on narrow contexts with negligible connection with the wide range of real processes of increasing importance in many disciplines.

The development of probability at school level beyond the realms of coins and dice needs urgent attention, and solutions are most likely to be found through connecting with data in real contexts. At the tertiary level, the persistent presence in data analysis texts and courses of tangential chapters on probability through Venn diagrams and the abstraction of coins, dice and balls, is a puzzling distraction in the “story” of statistical literacy and introductory data investigations. Probability is not just a sub-section of mathematics with limited applications, and distributions are not just about “lumps” of data. Just as the importance of data, statistical literacy and statistical thinking has been increasing across disciplines and society, so too has the importance of probabilistic and distributional thinking and modelling in themselves, as well as underlying all sound statistical data analysis techniques and the statistical thinking and practice of professional statisticians. Computing power and areas such as risk analysis are increasing the demand for stochastic and distributional thinking in applications and problem-solving in statistics, mathematical modelling, business, information technology, engineering and science, to name just some.

At the Queensland University of Technology, all students undertaking a mathematics degree, whether single or double degree program, are required to take two first
Examples from an introductory course in developing probabilistic statistical thinking: Part 1 – Helen MacGillivray

year statistics courses. The data analysis course is taken by all Science students and is similar in objectives and pedagogy to statistics courses taken by all students in Engineering, Surveying, Pharmacy, Medical Sciences, Medical Biosciences and Optometry. For information about these courses, oriented to, and built around data investigations, see [5, 6]. The other first year statistics course, compulsory for all mathematics and statistics majors, including mathematics education students, is in introductory probability and distributional modelling. This course builds skills and foundations in basic concepts, thinking and methods in probability, conditional arguments, distributional and stochastic modelling for applications in a wide range of areas, from communication systems and networks to traffic to law to biology to financial analysis. The course helps students to unpack, analyse and extend their prior knowledge, understanding and misunderstandings. The learning and assessment package of the course is built around problem-solving, with activities that link data, everyday processes, student experiences, modelling and simulation, and consolidate core mathematical skills.

Various aspects of this course are discussed in [3, 4, 5], including the group project in collecting and testing data in two everyday stochastic processes identified in free choice by the group. The content and the collaborative constructivist approach are discussed in [4] with examples of the preliminary activities and exercises for each topic that assist in the unpacking and extending approach.


Almost all the examples, exercises and problems in this course have been developed specifically for it over a decade as there are very few resources available for its approach. Although some examples are included in the above papers, space restrictions limit such inclusions and requests for more examples are frequently received. Hence two articles give a number of the problems developed for this course, with comments on their aim. The problems given in these articles tend to be the most challenging, with prior examples and exercises focussing on developing components of the reasoning in a wide variety of contexts. The problems selected for these articles are also in the sections that benefit most from the tutorial group exercise strategy outlined in [3] that creates the problem-solving learning environment envisioned by [2]. The ideas for contexts arise from many sources – consulting, research in other areas, everyday occurrences and conversations. My students often joke that they can tell where I was when I thought up certain questions.

Below are selected problems from the first few sections of the course on logical representations of events, basic probability, independence and conditional probability.

The second article will include selected problems from introductory Markov chains, special distributions, introductory queues, finding means through conditioning, and linear combinations of normal random variables.

Examples on events, basic probabilities and using independence

The examples in these sections focus on being able to move between word descriptions and their logical (Boolean) representations, and the use of basic probabilities and the assumption of independence within contexts. The network design and reliability exercises, of which the first question below is representative, are of direct importance to students combining IT or engineering with maths, and of indirect value to all students. The types of questions which question 2 represents are of great value in learning to logically dissect and represent wordy descriptions in other disciplines. Question 3 combines this with the use of independence and comfort with working with probabilities of “at least” and “at most”, which diagnostic tests have shown to be an unexpectedly difficult stumbling block for a considerable number of students.

1. Within a network design, eight identical components can be arranged according to either (A) or (B) below. Let \( W_i \) be the event component \( i \) functions.

   ![Fig 1 – Network design (A)](image1)

   ![Fig 2 – Network design (B)](image2)

   (i) Give an expression for the events that:

   (1) (A) functions, and

   (2) (B) functions.

   (ii) Each component functions with probability \( p \) independently of other components.

   System (A) is never more reliable than system (B). This means \( \Pr[(B) \text{ functions}] \geq \Pr[(A) \text{ functions}] \). Find the value of \( p \) for which the difference in the reliability of systems
(A) and (B) (that is, the difference \( \Pr(B) \) functions - \( \Pr(A) \) functions) is a maximum.

2. In an assessment moderation scheme used across schools in a state, two moderators independently consider schools' ratings of selected student portfolios. For each student portfolio, if both moderators agree with the school's rating, it is ratified. If at least one moderator disagrees with the rating, two other independent moderators also consider the rating. If at least two of the total of four moderators agree with the rating, it is ratified, otherwise it is returned to the school for further consideration and discussion. Let \( A_i \) be the event that the \( i \)th moderator agrees with the school's rating. Use set notation to denote the event that the school's rating is ratified.

3. Two species of trees are growing in an area. The tables below show the probabilities of AT LEAST certain numbers of young trees of each species appearing in a 400 sq m section

<table>
<thead>
<tr>
<th>Number of young trees in section, ( x )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Pr(\text{at least } x) )</td>
<td>0.9</td>
<td>0.7</td>
<td>0.2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 1 – Species A data

<table>
<thead>
<tr>
<th>Number of young trees in section, ( y )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Pr(\text{at least } y) )</td>
<td>0.9</td>
<td>0.8</td>
<td>0.75</td>
<td>0.6</td>
<td>0.4</td>
<td>0.2</td>
<td>0.1</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Table 2 – Species B data

The number of young trees of species A appearing in a section is independent of the number of young trees of species B appearing in the section. It is judged by forestry management that viable sections will have either at most 2 trees of species A and at most 7 trees of species B, or at most 3 trees of species A and at most 4 trees of species B. If a section is not naturally viable, intervention is necessary. What is the probability that a section does not need intervention?

Examples using conditional probability

The section on conditional probability starts with exercises that the students can do just with knowledge in handling %’s [4], and develops gradually through simple and classic Bayes exercises to multi-step and more complex situations. Although these involve important language understanding and interpretation similar to those described in [8], they are quintessential problem-solving learning activities, involving logical dissection, identification of knowns and unknowns, and careful handling of reference frameworks. Note that both problems below involve a key learning aspect in understanding and using conditional probabilities: all probabilities are conditional on their reference framework, but the condition needs specification only when different reference frameworks are combined.

1. Fred is a beagle “sniffer” dog at a cargo handling depot. Fred is 95% reliable in detecting contraband substances when they are present, and also has a probability of only 0.005 of indicating the presence of contraband substances when they are not present.

(i) If Fred indicates that contraband substances are present in 1% of cargoes he inspects, show that the probability, \( p \), that a cargo contains contraband substances is 0.0053.

Using this value of \( p \), obtain the probability of contraband substances in a cargo if Fred indicates their presence.

(iii) Fred’s younger brother Pete is still being trained. Pete is currently 90% reliable in detecting contraband substances when they are present, and has a probability of 0.01 of indicating the presence of contraband substances when they are not present. When a dog indicates the presence of contraband substances in a cargo, it is opened and inspected in detail. If Fred and Pete share the work, what proportion could Pete be allowed to do if it is desired to keep the percentage of cargoes opened and inspected in detail because a dog indicates the presence of contraband substances, to 1.2%. (Again use the value of \( p \) of 0.0053)

2. (i) A new marketing campaign for an existing product that currently has 5% of the market runs advertisements in newspapers. Based on circulation and previous survey information, the marketing company states that the probability an arbitrary person sees the ads is 0.6. During the newspaper ads campaign, the buyers of the product were asked if they’d seen the ads – 75% replied that they had. Assume that for those who did not see the ads, the probability of buying the product is still 0.05.

“...understanding of language in real contexts is an important aspect of statistical literacy. Interpreting descriptions of situations and extracting the essential information is also the first stage of mathematical and statistical/stochastic modelling.”
Show that the probability that a person who saw the ads buys the product is 0.1.

(ii) The campaign now extends to include ads on TV. Based on (i) above, we assume that a person who sees the newspaper ads, but not the TV ads, has a probability of 0.1 of buying the product. It is considered that a person who sees the TV ads, but not the newspaper ads, has a probability of 0.3 of buying the product, and that those who see both newspaper and TV ads have a probability of 0.35 of buying the product. As in (i), the probability that a person who sees no ads buys the product is 0.05, and the probability that a person sees the newspaper ads is still 0.6. It is also assumed that a person has a probability of 0.6 of seeing the TV ads, independently of whether they see the newspaper ads.

Find the probability that a buyer saw the TV ads but not the newspaper ads.

Conclusion

As Schield comments [8], understanding of language in real contexts is an important aspect of statistical literacy. Interpreting descriptions of situations and extracting the essential information is also the first stage of mathematical and statistical/stochastic modelling. The next stage is expressing the situation and information in the logical representative way that we call mathematical modelling. It is statistical/stochastic modelling if it involves probability and/or random variables and hence distributions. In addition, conditional probability is the implicit or explicit basis for an enormous amount of statistical/stochastic modelling both theoretical and applied, and building familiarity and comfort with conditional probability beyond definitions and standard simple exercises can make a significant difference to students as they proceed beyond first year courses.

The examples, exercises and problems from which the above are selections, aim to start the development of capabilities in these stages of problem-solving and in conditional probability. It is hoped that the above small selection of problems will whet statisticians' appetites for enriching and revolutionising the teaching of introductory probability and conditional probability.

References


