

First-year Statistics for Psychology Students
Through Worked Examples

2. Probability and Bayes' Theorem

by

Charles McCreery, D.Phil

**Formerly Lecturer in Experimental Psychology
Magdalen College
Oxford**

Copyright © Charles McCreery, 2007

OXFORD FORUM

Psychological Paper No. 2007-2



Contents

1. Introduction.....	3
2. A worked examination question	
2.1 The question.....	4
2.2 The answer.....	5
2.2.1 The data, and probability tree.....	5
2.2.2 Probability of randomly choosing a green Smartie.....	5
2.2.3 Probability of randomly choosing green.....	6
2.2.4 Probability of Smartie given green.....	6
2.2.5 Penultimate part of question.....	7
2.2.6 Final part: probability of choosing three of one brand and one of the other.....	8
3. Recommended reading.....	10
 Appendix – Probability and Bayes’ Theorem: Summary of some key points.....	 11

Introduction

This tutorial is based on handouts I developed when teaching statistics to first-year psychology students at Magdalen College, Oxford. The rationale for basing such teaching on detailed worked examples is given in the Introduction to my tutorial on the Chi-square test¹.

A word of encouragement

Probability theory is a subject which is well-known for producing what seem at first sight to be counter-intuitive results. In addition, Bayes' theorem may seem difficult to grasp at first, because it seems to involve us in 'thinking backwards' in a way we are not used to. However, like most ideas, it is actually quite simple, and indeed obvious, once grasped. The problem is that having grasped it once may not guarantee that the understanding of it sticks – it may be necessary to think it through a second and even a third time. To this end some people may find the visual method of representing probabilities, via 'probability trees', helpful.

Concerning the layout of this tutorial

Although section 2.2 is called 'Answer', it is not intended to be a model answer, such as one might give in an examination. Even without the sections of commentary which I have hived off within square brackets for the sake of clarity, section 2.2 contains much more information than one would need to give in an examination, since I am attempting to explain what I am doing as I go along.

Comments

If anyone has any comment to make on this tutorial, or notices any typographical or other error which should be corrected, I should be pleased to hear from them, via:

reply@oxfordforumpublisher.com

¹ McCreery, C. (2007). 'First-year statistics for psychology students through worked examples: I – The Chi-square test, a test of association between categorical variables'. *Psychological paper No. 2007, 1*. Oxford: Oxford Forum. Online pdf at: <http://www.celiagreen.com/charlesmccreery.html>

2. A worked examination question

2.1 Question²

M&M's and Smarties³ are two different brands of small milk chocolates in a crisp coloured shell. Each item of confectionery is about the same size and each brand comes in a mixture of colours.

A large bowl contains a mixture of the two brands in the ratio of five M&M's to four Smarties in just four colours - red, yellow, orange and green.

The proportions of the M&M's which are red, yellow, orange and green are 0.3, 0.4, 0.1 and 0.2 respectively, while the equivalent proportions for Smarties are 0.25, 0.2, 0.3 and 0.25.

A sweet is chosen at random from the bowl. What is the probability that it is (i) a green Smartie; (ii) green; (iii) a Smartie if it is green?

The shades of green used by the two brands are very different and can be readily identified, but it is not possible to differentiate the two brands from the other colours. A single sweet is drawn at random from the container and a statistician is shown its colour. If it is green, the statistician will identify its brand correctly, but if it is not green, the statistician will toss a fair coin to decide which brand she thinks it is. What is the probability that she correctly identifies the brand?

If four sweets are chosen from this very large pool of sweets, what is the probability that three will be of one brand and the other sweet of the other brand?

(3,3,3;7;4)

[The numbers in the bottom right hand corner of the question are as they appear in the Oxford Prelims exam paper and indicate to the candidate how many marks each part of the question is potentially worth if answered correctly.]

² The question is taken from the Prelims Statistics paper for first-year psychology students at Oxford University, Hilary Term, 1999.

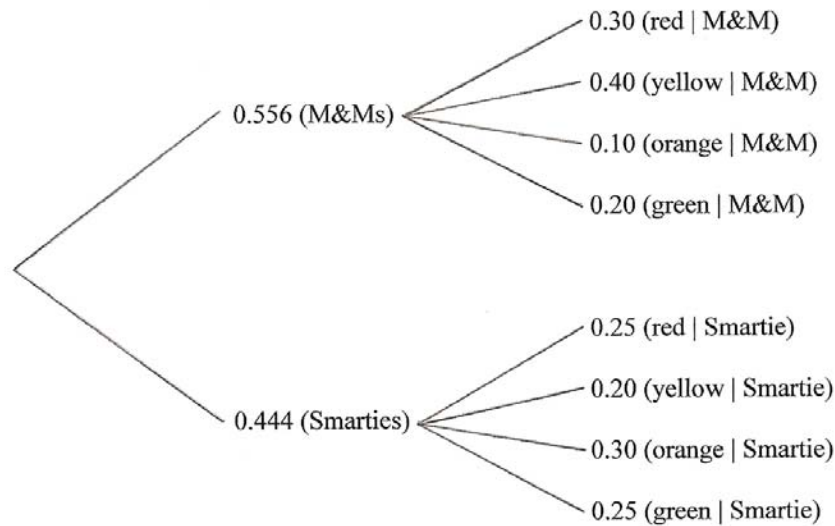
³ 'M&Ms' and 'Smarties' are registered tradenames.

2.2 Answer

2.2.1 The data, and probability tree

Out of every nine items in the bowl five are M&Ms and four are Smarties.
So the probability of picking an M&M at random is $5/9 = 0.556$.
The probability of picking a Smartie is $4/9 = 0.444$.

Probability Tree:



2.2.2 Probability of randomly choosing a green Smartie

[p = probability
| = given that
* = multiplied by]

$$\begin{aligned} & p(\text{Smartie and green}) \\ &= p(\text{Smartie}) * p(\text{green} | \text{Smartie}) \\ &= 0.444 * 0.25 \\ &= 0.111 \end{aligned}$$

[Illustrates the mnemonic: 'multiply for AND'. See Appendix.]

2.2.3 Probability of randomly choosing green

There are two ways of getting a green sweet; one can choose **EITHER** a green M&M **OR** a green Smartie, but not both at once - the possibilities are mutually exclusive, or 'disjunctive'. So this is a case of **addition** of probabilities. [Remember the mnemonic: 'Add for OR'.]

$$\begin{aligned} p(\text{Smartie and green}) &= 0.111 \text{ (from the first part of the question, above).} \\ p(\text{M\&M and green}) &= p(\text{M\&M}) * p(\text{green} | \text{M\&M}) \\ &= (0.556 * 0.20) \\ &= 0.111 \end{aligned}$$

$$\begin{aligned} p(\text{green}) &= p(\text{Smartie and green}) + p(\text{M\&M and green}) \\ &= 0.111 + 0.111 \\ &= 0.222 \end{aligned}$$

2.2.4 Probability of Smartie given green

[This is a Bayes-type question, since we are asked to find the probability of one of a number of possible antecedents to a given consequent.]

There are two ways of being green (as we saw in the previous part of the question): being an M&M and green, and being a Smartie and green. The probability that a given green sweet is a Smartie is therefore the probability of being a green Smartie expressed as a proportion of all the ways of being green; i.e. the probability of being a green Smartie divided by the probability of being green. More formally:

$$\begin{aligned} p(\text{Smartie} | \text{green}) &= p(\text{Smartie and green}) / (p(\text{Smartie and green}) + p(\text{M\&M and green})) \\ &= 0.111 \text{ [from first part of question]} / 0.222 \text{ [from second part]} \\ &= 0.5 \end{aligned}$$

[Notice that this result makes intuitive sense if you look at the probability tree above. The proportion of M&Ms to Smarties in the bowl is very roughly equal; the proportion of greens within the M&M group is very roughly equal to the proportion of greens within the Smarties group; and the disproportions in the two cases work in opposite directions to each other (i.e more M&Ms than Smarties overall, but more greens among

the Smarties than among the M&Ms). So it makes sense that the resulting probability (Smartie given green) is exactly half.

General point: the answer to a Bayes-type question may always be thought of as a fraction, in which the numerator (on the top) is the probability of the outcome arriving via the particular route you are interested in, while the denominator (on the bottom) consists of that same probability plus the probabilities of that same outcome arriving by all the other possible routes.

N.B1. The probability question in this paper is often progressive in the way this one is; i.e. successive parts of the question often make use of the results of previous parts, which can save you work if you notice it.

N.B2. I have presented the formula for Bayes' theorem in the form above for heuristic purposes: i.e. in order to give the reader an intuitive grasp of what the theorem is saying. The standard statement of the formula is given at the end of the Appendix.]

2.2.5 Penultimate part of question

A single sweet is taken at random.

Probability it is green = 0.222 (from section 2.2.3).

Probability it is not green = $1 - 0.222$ [because the probabilities must sum to 1]
= 0.778.

Probability of guessing whether a sweet is an M&M or a Smartie correctly by chance (e.g. by tossing a coin) = 0.5.

[Because every time she guesses M&M the probability she is correct is 0.556, i.e. rather more than 1/2, but every time she guesses Smartie, the probability she is correct is only 0.444, i.e. rather less than half, and these two deviations from 0.5 cancel each other out in the long run.

Note that a key premise of this part of the question is the idea that the statistician 'will toss a fair coin' to decide which sweet to guess when she cannot immediately tell from the colour. This ensures that in the long run she guesses M&M' and Smartie in equal proportions, i.e. 50% each. If she was truly 'guessing', i.e. making up her own mind on each trial which sweet to guess, there would most likely be some bias in the proportion of her guesses, so the proportion of correct guesses might well not be 0.5. For example, in the limiting case in which she guessed M&M on every trial, she would be right 56% of the time.]

Therefore probability of guessing a not-green sweet correctly = $0.778 * 0.5$
= 0.389

[Because we have just established above that 77.8% of the sweets are not-green, and that every time she has to guess a not-green sweet she is right half the time.]

$$0.222 + 0.389 = 0.611$$

So overall the probability of correctly guessing the brand is 0.611.

[Here we have applied the additive rule, 'Add for OR'. The two possibilities on each trial - having to guess a green sweet, or having to guess a not-green sweet – are disjunctive, or mutually exclusive. So we have to add the probabilities of the two possible results. This again produces what may seem a rather paradoxical result: in neither situation – guessing a green sweet, and guessing a not-green sweet – can she hope to get more than half of the results right, and yet overall she will be right more than half of the time.

One way of thinking of the logic of this part of the question is as follows: imagine the person choosing a succession of 100 sweets. 22 will be green and will therefore be correctly identified. 78 will be other colours, so only half of them will be correctly identified. $78/2 = 39$. $22 + 39 = 61$. So 61 out of the hundred sweets will be correctly identified. $61/100 = 0.61$.

In general it is often helpful in probability questions to imagine the probabilities as percentages, i.e. as actual numbers of people/things out of a population of 100. This is equivalent to thinking of probabilities as areas, such as subsections of a pie-chart, which can also be helpful.]

2.2.6 Final part: probability of choosing three of one brand and one of the other

There are eight ways of choosing three sweets of one brand and one of the other, which may be represented as follows:

MMMS
MMSM
MSMM
SMMM

SSSM
SSMS
SMSS
MSSS

The first four possible arrangements each have the probability: $0.555^3 * 0.444 = 0.0759$.

The second four each have the probability: $0.444^3 * 0.555 = 0.0486$.

$$4 * 0.0759 = 0.3036.$$

$$4 * 0.0486 = 0.1944.$$

$$0.3036 + 0.1944 = 0.498.$$

[This result may seem counter-intuitively large at first sight. And in fact a similar calculation shows, somewhat surprisingly, that the probability of getting three sweets of one brand and one of the other is greater than the probability of getting two of each.

This arises because there are more possible ways of getting three and one than two and two. There are only six of the latter:

MMSS
MSMS
MSSM

SSMM
SMSM
SMMS

Note also that this last part of the question refers to a 'very large' pool of sweets. This is to allow you to assume that, although this is not a case of random selection 'with replacement', the pool is sufficiently large for the picking of four sweets from it not to affect the ratio of M&Ms to Smarties in what remains.]

3. Recommended reading

Hoel, Paul G. (1976). *Elementary Statistics* (4th edition). New York: Wiley. Chapter 3.

Covers the basic ideas of probability, the addition and multiplication rules, probability trees, and Bayes' theorem.

Hays, William L. (1994). *Statistics* (5th edition). Orlando, Florida: Harcourt Brace. Chapter 1.

Also covers the basic ideas of probability, the addition and multiplication rules, and Bayes' theorem, but does not introduce probability trees. Has a fuller discussion than Hoel of the concept of conditional probability.

Appendix 1

Probability and Bayes' Theorem Summary of some key points

1. Some key concepts:

Independence (of two or more events)

Criterion: 'If A1 occurs, does that change the chances that A2 will occur from what they would be if A1 were completely ignored?' (Hoel, p.55)

Mutual exclusion (of two or more events)

Criterion: 'If A1 occurs, does that make it impossible for A2 to occur?' (Hoel, p.55)

Conditional probability

The probability of B given A; represented as $p(B | A)$.⁴

2. The Additive Rule

When A1 and A2 are mutually exclusive events,

$$p\{A1 \text{ or } A2\} = p\{A1\} + p\{A2\}. \quad (\text{Hoel, p 47})$$

'Given a set of mutually exclusive events, the occurrence of one event or another is equal to the sum of their separate probabilities.'⁵

⁴ Most of the information we deal with in everyday life as the basis for the choices we must make has a conditional character [...] The very best information we have to go on is usually no more than a probability. These probabilities are conditional, because virtually all our information is of an "if-then" character.' Hays, *op. cit.*, p. 45.

⁵ Howell, David C. (1997). *Statistical Methods for Psychology* (4th edition). London: Duxbury Press, p.112.

3. The Multiplicative Rule

'The probability of the joint occurrence of two or more independent events is the product of their individual probabilities.'⁶

Mnemonic: Add for 'OR',
Multiply for 'AND'

4. Bayes' Theorem

Applies when we are 'working backwards' from a known outcome to the probability of one of several possible antecedent events.

General formulation:

$$p(A|B) = p(B|A) * p(A) / p(B)$$

Where p = probability

| = given that

* = multiplied by

and B is a consequent of some antecedent A

[E.g. A is one of the two possibilities on the left of the probability tree in section 2.2.1 above, and B is one of the four possibilities branching out from A, on the right of the tree.]

⁶ Howell, *op. cit.*, p.113.