MT1010
Topics in Mathematics:
Problem-solving techniques
2014/15

Dealing with Uncertainty

David Borchers

University of St. Andrews

School of Mathematics and Statistics

Contents

Part I
Lecture Notes

1 Introduction

This part of the course aims to introduce some ideas that will be built on more formally in later parts of MT2504. We will do this by considering a series of data sets that were generated by some processes that include randomness. The focus will be on constructing reasonable probability models for the observed data and getting useful and appropriate interpretations of the data from these models.

2 Can an octopus predict World Cup winners?

Well no, obviously not - but then what is the appropriate explanation for the fact that in 2010 Paul the German octopus (originally from Weymouth, England) did this? Paul came to the attention of the media after he predicted the outcomes of the first four of Germany's Wold Cup football matches with 100% accuracy. He did this by choosing one of two boxes containing clams that were lowered into his tank before each match - one box with the German flag on it, the other with their opponent's flag. (See http://www.youtube.com/watch?v=HEnhiJRep3g.) He then went on to predict the outcomes of Germany's remaining matches with 100% accuracy too. Was this just luck, or was Paul psychic, as the media said, or maybe he just liked the German flag (although he was technically English). If it was just luck, how lucky was he?

Figure 1: Paul the octopus' 2010 World Cup football predictions. (*From Wikipedida.*)

Australia	World Cup 2010	group stage	13 June 2010	Germany[31]	4–0	Correct
Serbia	World Cup 2010	group stage	18 June 2010	Serbia[31]	0–1	Correct
Ghana	World Cup 2010	group stage	23 June 2010	Germany[31]	1–0	Correct
England	World Cup 2010	round of 16	27 June 2010	Germany[32]	4–1	Correct
Argentina	World Cup 2010	quarter-finals	3 July 2010	Germany[23]	4–0	Correct
Spain	World Cup 2010	semi-finals	7 July 2010	Spain[33]	0–1	Correct
Uruguay	World Cup 2010	3rd place play-off	10 July 2010	Germany	3–2	Correct

If Paul has no ability to predict football results and was just choosing boxes at random on each occasion (this is our "null hypothesis"), what is the probability of getting the results shown in Figure 1? Well if he is choosing randomly with equal probability between two alternatives, then the probability of him making the correct choice for any one game (we will call the probability p) is 0.5. And if he is choosing independently on each occasion, then the probability of making the correct choice on 7 occasions is $p^7 \approx 0.009$. That is, if we let Paul make these 7 choices 128 times we'd expect him not to get it right about 127 times and to get it right only once. So either Paul was just very lucky, or there is something else going on.

Questions:

Q1 What would you conclude: that Paul was lucky, or that something else was going on?

Q2 In general, how unlikely[1] would an event have to be for you to conclude that there was something else going on (1 in 20, 1 in 100, 1 in 250, 1 in 1,000, ...)?

Paul came to the attention of the media only after correctly predicting the results of Germany's first four World Cup matches. Presumably we would never have heard about him if he had not got all of the first four right. Should we take this into account when calculating the probability of seeing what we saw with Paul?

Consider this scenario: Suppose that 2,000 octopuses worldwide are "asked" to predict the results in the same way that Paul was "asked". It is certain (well actually 99.99998% certain - see later) that at least one of them makes 7 correct predictions. One that makes correct predictions makes it into the media and is the one we hear about. Should we regard the event of 7 correct predictions as an event that has a probability of only 1/128 of happening by chance, - and hence is in fact very unlikely to have occurred by chance (which would incline us to conclude that the octopus has some intelligent way of choosing winners)? Or should we regard it as an event that was virtually certain to happen and so provides no evidence about the said octopus' predictive abilities? (The latter is the correct interpretation.)

There are really two different kinds of random events at work in Paul's case, and to build an appropriate probability model for what we observed, we need to take both into account. The first (event A, say) is the event "A World Cup-predicting animal makes it into the media", the second (event B, say) is "A World Cup-predicting animal predicts the last 3 of 7 results correctly".

To work out how unlikely Paul's predictions really are, we need to calculate $\mathbb{P}(A \cap B)$, the probability of A *and* B happening: of some animal making it into the media *and* this animal predicting the last four results correctly. We can do this by calculating $\mathbb{P}(A)$ and $\mathbb{P}(B|A)$ and then multiply them together to get $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B|A)$. And if A and B are independent, then $\mathbb{P}(B|A) = \mathbb{P}(B)$ and $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

2.1 Event A: Making it into the media

In predicting World Cup results, Paul had competitors in the animal world. "Paul's main rival was Mani the Parrokeet in Singapore, who was roundly beaten by Paul in the psychic showdown after picking Netherlands to win the final. In fact many animals around the world were making predictions – porcupines, guinea pigs and so on." (Spiegelhalter, http://understandinguncertainty.org/).

Suppose there were N animals around the world making predictions. What is the probability that at least one of these makes it into the media? If we assume that predicting the outcome of your team's first four games correctly gets you into the media, this question is "What is the probability that at least one of the N animals predicts the first four results correctly?". The probability of making 4 correct predictions is $p^4 = 1/16$. Using this, we can calculate the probability that at least one of the N animals predicts the first four results correctly as follows:

[1] This level of "unlikeliness" would be called the "significance level" by a statistician.

$$
\begin{aligned}
&\mathbb{P}(\text{At least 1 of } N \text{ animals make 4 correct predictions})\\
&\quad = 1 - \mathbb{P}(\text{No animals make 4 correct predictions})\\
&\quad = 1 - \mathbb{P}(\text{All } N \text{ animals make at least one error in 4 predictions})\\
&\quad = 1 - \mathbb{P}(\text{An animal make at least one error in 4 predictions})^N\\
&\quad = 1 - \{1 - \mathbb{P}(\text{An animal makes no errors in 4 predictions})\}^N\\
&\quad = 1 - \left(1 - p^4\right)^N = 1 - \left(1 - \frac{1}{16}\right)^N = 1 - \left(\frac{15}{16}\right)^N
\end{aligned}
$$

Spiegelhalter (http://understandinguncertainty.org/) reckons that there were about $N = 20$ animals around the world making predictions, so the probability of at least one of them making it into the media (event A) is about $1 - (15/16)^{20} \approx 0.725$.

2.2 Event B: Predicting the last 4 results correctly

Under our (null) hypothesis that animals are choosing winners randomly and independently with probability $p = 0.5$, the probability of any animal predicting the last four results correctly (event B) is $\mathbb{P}(B) = p^3 = 1/8$. This probability is unaffected by whether or not this animal made it into the media (event A) and therefore, $\mathbb{P}(B|A)$=$\mathbb{P}(B) = 1/8$ in Paul's case.

2.3 Event $A \cap B$ and a more correct probability

The probability of what we observed is the probability of "an animal making it into the media and then predicting the last 4 results correctly" ($\mathbb{P}(A \cap B)$). (We assume that the media focusses on only one animal if more than one predict the first 4 results correctly.) By the definition of conditional probability, $\mathbb{P}(A \cap B) = \mathbb{P}(B|A)\mathbb{P}(A) \approx (1/8) \times 0.725 \approx 0.09$. That is, the chance of a randomly-choosing animal doing as well as Paul did is about 1 in 11 (rather than 1 in 128, as we previously calculated).

Questions:

Q3 What would you conclude: that Paul was lucky, or that something else was going on[2]?

Q4 Is your answer to Q3 consistent with your answer to Q2?

Q5 Suppose your decision was to conclude that something else was going on. What is the probability that you made the wrong decision if Paul really was choosing randomly?

[2]Concluding that something else was going on is what statisticians would call "rejecting the null hypothesis" (where in this case our null hypothesis was that Paul selected winners independently with probability 0.5 for each game).

3 The binomial distribution

In the course of working out probabilities related to Paul's predictive abilities above, we derived expressions for the following when the probability of succeeding on any one trial is p and trials are independent:

(a) the probability of getting N successes out of N trials , which is just p^N, and

(b) the probability of getting at least one success in N trials. This is $1-(1-p)^N$. (Note that when we did it for Paul's case, the "success" was "4 correct predictions" and the associated probability was p^4: here we are calling a generic success probability p, whatever the success itself is.)

We can generalise these results by deriving expressions for the probability of obtaining *exactly* n successes in N trials (we have done this only for $n = N = 4$ so far) and for the probability of obtaining *at least* n successes in N trials (we have done this only for $n = 1$ so far).

3.1 The binomial probability mass function

For brevity, we will uses 1s and 0s to indicate successful and unsuccessful predictions, respectively. So the event that Paul predicts correctly 4 times in 4 trials is encoded as 1111, while the event that he predicts all but the last trial correctly is encoded as 1110, and the event the he predicts only the second trial correctly is encoded as 0100, and so on. You should be able to see that:

$$\mathbb{P}(1100) = p^2(1-p)^2$$

$$\mathbb{P}(1000) = p^1(1-p)^{4-1}$$

$$\mathbb{P}(0100) = p^1(1-p)^{4-1}$$

$$\mathbb{P}(0010) = p^1(1-p)^{4-1}$$

$$\mathbb{P}(0001) = p^1(1-p)^{4-1}$$

and that in general the probability of any particular sequence of successes and failures involving n successes and $(N-n)$ failures is $p^n(1-p)^{N-n}$.

Knowing this, what is the probability of Paul correctly predicting exactly 1 out of 4 trials? Well there are four possible ways that he could do this (they are the last four listed above), so $\mathbb{P}(n = 1)$ when there are four trials is $4 \times p(1-p)^{N-1}$.

Now, what is the probability that Paul correctly predicts exactly n out of N trials? Well there are $\binom{N}{n}$ possible ways that he could do this (see the "Bionomial coefficients" section of Chapter 2 of MT2504 notes), so writing the probability of Paul correctly predicting exactly n out of N trials as $\mathbb{P}(n)$, we see that $\mathbb{P}(n) = \binom{N}{n}p^n(1-p)^{N-n}$.

For generality we are going to introduce a variable X whose value encodes random events. We will say "$X = 0$" if the event is that there are zero successes in our N trials, "$X = 1$" if the event is that there is exactly one successes in our N trials, "$X = N$" if the event is that there are exactly N successes in our N trials, and in general "$X = n$" if the event is that there are exactly n success in our N trials. Obviously $X \in \{0, 1, 2, \ldots, N\}$.