The 2010 Crystal Ball Competition

The Royal Statistical Society and *Significance* magazine have launched our first Crystal Ball Competition. It is open to everyone – school and college team entries are especially welcome – and it is certainly not restricted to statisticians. Furthermore, it ought to be fun. All you have to do is try to forecast the results of 10 different events, some sporting, some cultural, some just odd, that will take place in the months of May to July 2010. Send in your forecasts before the last week of May and the events themselves will prove you right or wrong.

Cash prizes are on offer to the winners, as well as books generously provided by Wiley-Blackwell and, in addition to all that, a warm feeling of triumph inside. Even losers will gain: they can watch Wimbledon tennis, or the rise or fall of the Dow Jones index, or even the lengthening holiday queues at Heathrow, with the added interest of knowing that their own closeness to a prize depends on the outcome.

What to do

The things that are to be forecast are numbers. Question 1, for instance, is the number of goals that will be scored in the 2010 football World Cup. You have to predict – or estimate, or guess – that number; how you do it is entirely up to you. We give you data from the immediate past that might or might not help, and indicate where further data is to be found. You might want to perform intricate statistical analyses for some questions; in some cases – if you are a football expert? – other knowledge or judgements may come in useful; for some questions you might simply take a flying guess. If you have access to a crystal ball that foretells the future you are welcome to try even that. You will be judged strictly according to your results. Time will reveal the winners.

There is just one more thing. Many people are required to estimate things in their daily lives, but what sets statisticians apart is that they are asked to say how accurate their estimates are. This is a statistical competition. We want you to tell us how accurate you think your predictions are.

How to tell us your uncertainty

There are two common ways to describe uncertainty. One is to provide a range of values around an estimate, with, say, a 95% degree of assurance that the true value will lie in the range between them. Thus you might estimate that 100 goals will be scored in the World Cup; but you think that you might be out by 13 goals either way, and you are 95% certain that your spread of 13 goals will cover the actual result. Then your estimate is 100 goals; 87 goals is the lower limit and 113 is the upper limit for your 95% probability interval (or confidence interval, or credible interval).

The second way is to provide a standard deviation which simply says how close on average you expect the estimate to be to the true value. In this competition we are asking for a standard deviation.

You can use whatever method you like to decide on your standard deviation, but here is one simple way to get it. Fix upper and lower limits so that you think the true value is twice as likely to be inside your range as outside it. (Technically you are setting a 67% probability interval, but don't let that worry you.) Subtract the lower limit from the upper one to find the width of your range. Halve it. And that is your standard deviation.

So again, if you predict that the Wimbledon final will last 4 hours, and believe there is a two-thirds chance of its lasting between 3 hours 30 minutes and 4 hours 30 minutes, your 67% limits are 3 hr 30 min and 4 hr 30 min. The width of that range is 60 minutes; half of it is 30 minutes – and that is your standard deviation.

Another way of deciding on your 67% limits is that if you were offered a bet in which you would win £1 if the true value was between 3hr 30 min and 4 hr 30 min but would lose £2 otherwise, then this would feel like a fair bet.

How to enter

So all you have to do is give a prediction and a standard deviation for each of the ten questions. At the end of any analysis or research that you feel is useful, your prediction should be your best judgement of what the true value of the quantity will be.

Your standard deviation is also a judgement. Statistical analysis will often produce a standard deviation (or standard error) for the estimate, but you may wish to modify that if you have applied other knowledge or judgement to produce your estimate.

Scan or photocopy the form – or it is available to download at `www.interscience.wiley.com/journal/significance`. Post your form to: Significance Crystal Ball Competition, Royal Statistical Society, 12 Errol Street, London EC1Y 8LX, or e-mail them to: crystalball@rss.org.uk. The closing date is 28th May 2010.

Prizes

Anyone can enter. The competition is open to individuals and to school and college teams. The individual winner will receive a cash prize of £200, plus a year's free membership of the RSS (or a refund if you have already paid your subscription). The winner of the second prize will receive his or her selection of books published by Wiley-Blackwell to the value of £150. The third prize is £75 worth of Wiley-Blackwell books**.** The winning schools entry will receive £150 cash and associate schools membership of the RSS.

The scoring rules

Full technical details of the scoring rules can be found on the *Significance* website, `www.interscience.wiley.com/journal/significance`, but the main features are these. The scoring system aims to reward not just good estimation but also good assessment of the accuracy of the estimates.

For each question, the score looks at the difference between the true answer, as revealed by the event itself, and the contestant's estimate of it. It does this both in absolute terms, and relative to the contestant's standard deviation. An entry that has estimates which are in absolute terms close to the true values for all of the questions will receive a good (i.e. low) score, but the winner is likely also to have given estimates that are not generally better or worse than implied by their standard deviations.

To win, your estimates not only have to be accurate but they should not be a lot more accurate than you say they are!

Conditions

Entries must be received by 28th May 2010. Only complete entries, giving forecasts in the required format for all ten questions, will be accepted. In the event of one or more of the questions being rendered inoperative owing to the hazards of time – the Wimbledon final being rained off and settled by tossing a coin, for example, or a significant information website no longer operating – those questions will be omitted from the scoring, according to the Editor's best judgement. The results will be published in *Significance*. The Editor's decision will be final.

Significance and Royal Statistical Society Crystal Ball Competition

For each question give your prediction of the quantity involved, and an assessment – the standard deviation – of how accurate you think your prediction will be. You are welcome to use any method you like to arrive at your predictions and standard deviations.

1. How many goals in total will be scored in all 64 matches of the 2010 World Cup competition in South Africa, up to and including the final on 11th July?

Penalty shoot-outs are *not* included. Totals from the last three World Cups were: 2006, 147; 2002, 161; 1998, 171. See http://www.fifa.com/worldcup/ for earlier data.

Prediction:.....
Standard Deviation:.....

2. How long will the 2010 men's singles final match at Wimbledon, scheduled for 4th July, last?

We are looking for the playing time on court, as given by the scoreboard clock, and not including knock-ups and time out for rain, injury, etc. In 2009 the playing time was 4 hr 16 min; in 2008 it was 4 hr 48 min; in 2007 it was 3 hr 45 min.

Prediction:.....
Standard Deviation:.....

3. What will be the winning score in the Eurovision Song Contest on 29th May?

The last three winning scores were: 2009, 387; 2008, 272; 2007, 268. See http://www.eurovision.tv for earlier data.

Prediction:.....
Standard Deviation:.....

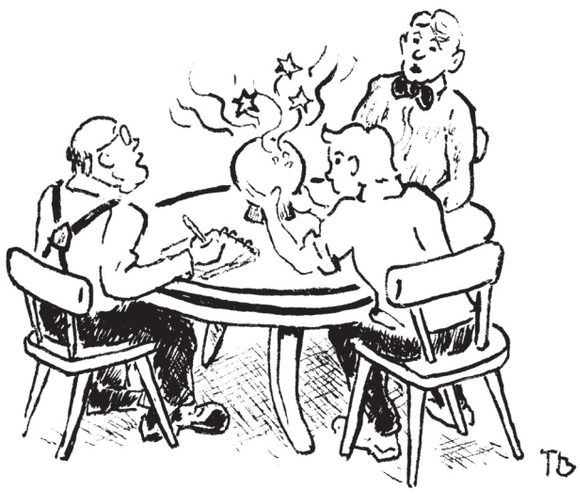
4. How many Twitter followers will Arnold Schwarzenegger have at midday (UK time) on 30th June?

It was 11,147 on 2nd February 2009; 1,437,199 on 24th November; and 1,503,892 on 14th December. See http://twitter.com/schwarzenegger for daily updates.

Prediction:.....
Standard Deviation:.....

5. At the US box office, on the weekend starting the 2nd July 2010, what will be the return in dollars of the most successful film?

On the first weekend of January 2009 it was $24,263,763; on the first weekend of June 2009 it was $44,979,314; on the first weekend of December 2009 it was $20,043,181. See http://www.imdb.com/boxoffice/ for further and archive data.



Prediction:.....
Standard Deviation:.....

6. According to Met Office figures, how many hours of sunshine will Scotland have in June 2010?

In 2009 it was 204.4 hours; in 2008 it was 146.0 hours; in 2007 it was 127.6 hours.

See http://www.metoffice.gov.uk/climate/uk/2009/june/averages.html for archive data, substituting month and year in the URL.

Prediction:.....
Standard Deviation:.....

7. At what value will the Dow Jones Index stand at close of business on 11th July 2010?

At the close on 11th November 2009 it was at 10,291.26. See, for example, http://www.google.co.uk/finance/historical?q=INDEXDJX:.DJI for historical data.

Prediction:.....
Standard Deviation:.....

8. How many passengers will pass through the terminals at Heathrow airport in June 2010?

We want the figure as given by the British Airports Authority (BAA). Recent data is: November 2009, 5,028,739; October 2009, 5,687,438; September 2009, 5,783,611; December 2009, 5,316,129.

Google "BAA monthly traffic statistics" to get to the BAA home page; then follow link "view BAA's monthly traffic statistics since 2003" for an Excel spreadsheet containing further data.

Prediction:.....
Standard Deviation:.....

9. How many viewers will watch the Friday 4th June 2010 episode of the popular BBC1 soap series *EastEnders*?

We want the number as calculated by the Broadcasters Audience Research Board (BARB). On 4th September 2009 it was 8.17 million; on 2nd October 2009 it was 8.18 million; on 6th November 2009 it was 9.23 million; see http://www.barb.co.uk/report/weeklyTopProgrammesOverview for other data.

Prediction:.....
Standard Deviation:.....

10. How many entries will there be to this competition?

The only useful information we can give is that a copy of Significance is sent to all members of the RSS and the magazine is also available via the libraries of subscribing institutions. However, statisticians frequently are asked to give estimates from inadequate information…

Prediction:.....
Standard Deviation:.....

Your details:

Name: .....

Address: .....

.....

.....

.....

Postcode: .....

Country: .....

E-mail address (optional): .....

Individual entry, or school/college team? Please tick box:

Individual: ☐ School/college team: ☐

Send your answers to: Significance Crystal Ball Competition, Royal Statistical Society, 12 Errol Street, London EC1Y 8LX, or e-mail them to: crystalball@rss.org.uk. Entries to be received by 28th May 2010.