The Guardian

When the moon is full, sink your teeth into a p-value

Are crime rates really higher when there's a full moon? That's a question for the statistical werewolves among you

Nathan Green

Fri 23 Dec 2011 12.23 GMT

I was a particularly cowardly child. I'm not such a brave adult either, but the subjects of my cowardice have changed somewhat. As an adult I'm more scared of losing my job in a recession or having my identity stolen on the internet. As a child, I was terrified of werewolves. Every full moon I would worry about being on the wrong end of gnashing, razor-sharp teeth.

Of course I shouldn't have been any more worried when there was a full moon than on any other night. Or should I? A classic article in the British Medical Journal sought to answer a similar question: are crime rates higher when there's a full moon?

Science is all about formulating and testing hypotheses. In this case the hypothesis would be: "Crime rates are higher when there's a full moon." Often, scientists set out to test the "null hypothesis": the default statement that, if true, would indicate that their experiment had not detected any real effect. In the case in question, the null hypothesis could be expressed as: "There is no difference in crime rates when there is a full moon compared with other nights."

The problem with data like crime rates is that it contains random noise - patterns can appear and disappear by chance alone. So we first need to ask ourselves how sure we want to be that there is a real difference between crime rates on a full moon and those on any other night. Quite sure? Fairly sure? Almost certain?

In other words, how high are we going to set the bar? Statisticians quantify the degree of certainty they want to achieve as the significance level: a threshold between 0 and 1 where the closer you get to 0, the more certain you can be that there's a genuine difference between two sets of numbers. This is what scientists mean when they talk about how "significant" their results are.

The significance level can also be expressed as a percentage between 0 and 100. For example, if we want to be at least 95% sure there's a real difference between sets of figures, this is called a 5% significance level (0.05). In other words, if we compared the number of crimes on lots of full moon and non-full moon nights, the numbers would be different at least 95% of the time.

To recap, we've formally described the hypothesis we want to test - that more crimes occur on nights when the moon is full - and we've said how certain we want to be that there's a genuine difference in these crime rates - the significance level.

Now we just need to find out how likely it is that the null hypothesis is true. This number is called a p-value and is a form of probability. If this probability is smaller than the significance level we've chosen then we can reject the null hypothesis (at least for now).

And that's it. Now you have the tools to answer the original question, because when the p-value is smaller than the significance level, we can say with some statistical confidence that crime rates go up when there's a full moon.

Applying these methods to their data, the researchers found there was a significant difference between the number of crimes committed when there was a full moon and the number of crimes on other nights: their p-value was much smaller than 0.001 (usually written as $p<0.001$). This means there was less than a 0.1% chance of the full moon and non-full moon crime rates actually being the same, despite the numbers that were observed in this particular analysis.

In other words, the researchers showed that if they were to repeat this analysis over and over again, say every year, they'd expect to get a difference in the number of crimes between nights with a full moon and non-full moon most of the time.

The strict statistical interpretation of what a significance test tells us is actually a little more subtle and is often misunderstood, but for now this explanation is just fine. (I'll cover "confidence intervals" in a later post.)

The scientists in this particular study did not claim to be able to explain their result. However, they speculated that it could be something to do with biochemical changes in the body under the influence of the moon's gravity. I'm sure there will be a number of more convincing explanations, on which I am not qualified to comment. After all, I'm a statistician not a criminologist.

(The number of crimes on equinox and solstice days did not differ significantly from those on other days, the authors reported, "suggesting that the sun probably does not influence the incidence of crime".)

This story just goes to show it's often harder to explain a statistical finding than to calculate it. Somehow I doubt that much crime is carried out by werewolves, but I'm going to keep my door firmly locked just the same on the night of Monday 9 January - the next full moon.