Hypothesis Tests - General (3 pages; 22/5/20)

See also "Hypothesis Tests - Types"

(1) When an extreme value occurs, there are two possible explanations: (a) H_0 is correct, and the extreme value occurred by chance, and (b) H_0 is not correct.

The hypothesis test establishes whether explanation (a) is too implausible to be the correct one.

If H_0 really is true, there will still be occasions when implausibly extreme values occur, but we are prepared to be wrong 5% of the time (if the significance level is 5%). For a one-tailed test, this will happen when H_0 is true and the value is greater than or equal to some critical value (CV), such that $P(X \geq CV) = 0.05$

(2) In order for a test to be properly valid, a new sample needs to be taken when investigating a hypothesis (rather than basing the test on the data that gave rise to the hypothesis in the first place).

(3) One-tailed or two-tailed?

Situation A: A casino is being investigated to see whether a particular die being used is biased. (Suppose that a gambler wins money if a 6 is thrown.)

Situation B: A dice-making factory is suspected of producing biased dice because of faulty equipment.

Binomial trials can be carried out in each case. Let X be the number of sixes obtained in 120 trials, so that $X \sim B(120, p)$, where p is the probability of obtaining a six when a die is thrown.

For situation A, we have $H_0: p = \frac{1}{6}$, $H_1: p < \frac{1}{6}$ (one-tailed test),

and for situation B, we have $H_0: p = \frac{1}{6}$, $H_1: p \neq \frac{1}{6}$ (two-tailed test).

Let the significance level be 5%.

Suppose that a value of 28 is obtained for X.

In the case of situation A, H_0 could be defended if it could be shown that results as suspicious as this - or more so - are expected to occur at least 5% of the time.

But it can be found that $P(X \geq 28|H_0$ is true$) = 0.0373$ (3sf), and so H_0 is rejected, as $0.0373 < 0.05$.

In the case of situation B, a suspicious event would be a large value for X or a small one. H_0 could again be defended if it could be shown that results as suspicious as this - or more so - are expected to occur at least 5% of the time, but as these are spread over the two tails, for H_0 to be accepted we want $\text{P}(X \geq 28|H_0$ is true) to be no less than 2.5%. So, as $0.0373 > 0.025$, we can accept H_0.

(4) In general, a hypothesis test can be carried out in either of the following ways:

(a) Reject H_0 if $P(X \geq TS) < 0.05$ (for a one-tailed test at the 5% significance level), where TS (the test statistic) is the value of X obtained from the sample.

(b) Establish the critical region, and reject H_0 if TS lies within it.

This may be the best approach if multiple tests are to be carried out.

(5) It is recommended to express the conclusion of a hypothesis test both (a) formally and (b) 'in context', in layman's terms.

For example:

"As 13.2 [the test statistic] $>$ 12.7 [the critical value], reject H_0 at the 5% significance level. There is enough evidence to conclude that average rainfall has increased."

(6) It is also possible to use a confidence interval to test a hypothesis (ie reject H_0 if the value proposed by H_0 lies outside the appropriate interval).