Unit 10: Tests of Significance (Chapter 26)

Marius Ionescu

11/10/2011

What is due to chance?

Example (Questions)

- Suppose that you toss a coin 100 times and 60 times of the times you get a head? Is this due to chance error or is the coin biased?
- If your theory says 70% of students have blue hair and a simulation shows 80% have blue hair, is the simulation wrong?
- When is a difference **significant**?

Another example

Example

A die is rolled 100 times. The total number of spots is 367 instead of the expected 350. Is the die loaded?

Null hypothesis and alternate hypothesis

Fact

We test a hypothesis:

- *Null hypothesis: just chance variation*
- *Alternate hypothesis: there is a real difference between the two answers.*

Method

Fact

1. *Form a test statistics which measures the difference between the data and what is expected under the Null hypothesis*
2. *Find the chance that the test statistics takes on that value*
3. *Select a burden of proof to show that Null is wrong!*

The Z-statistics:

$$z = \frac{Actual - EV}{SE}.$$

Method

Fact

- *To use a test of significance, the null hypothesis should be stated in therms of a box model.*
- *The Z-test is used for large samples.*

Example

Example (Loaded die?)

Recall that we rolled a die 100 times and the number of spots was 367. The null hypothesis was that this is due to chance.

- Form a Z-statistics:

$$z = \frac{367 - 350}{17} = 1$$

- What is the chance that $z \geq 1$? (this is the z-test; it is used for large samples)
- We say that the **significance level** is 16%.
- Is 16% small enough to reject the null hypothesis?

Burden of proof

Fact (Levels of significance)

- *Usually* 5% *is the cut off used for significance*
- 5% *is statistically significant.*
- 1% *is highly significant.*

Example

Example

The registrar at a big University says that 67% of the 25,000 students are male. 100 students are chosen at random. 53 of them are men and 47 are women. Is this a simple random sample?

P-value

Fact

- *The P-value of a test is the chance of getting a big test statistic – assuming that the null hypothesis to be right.*
- *P is not the chance of the null hypothesis being right.*
- *The smaller the chance is, the stronger the evidence against the null.*

Example

Example

100 draws are made at random with replacement from a box. The average of the draws is 102.7 and their SD is 10. Someone claims that the average of the box is 100.

- Is this plausible?
- What if the average of the box was 101.1?

Example

Example (ESP (Extrasensory perception))

Suppose that you are tested for ESP. The computer picks at random one of the 10 options and you pick what you thing it will be. In 1000 trials you get 173 correct answers. Do you have ESP?

Example

Example

A coin is tossed 10,000 times and it lands 5,167 times. Is the chance of a head 50%?

- Formulate the Null and alternative hypothesis in terms of a box model.
- Compute z and P.
- What do you conclude?

True or False?

Example

If P-value is 1%, there is 1 chance in 100 that the Null hypothesis is correct?

t-test

Fact (t-test)

- *For small samples the SD of the sample is not a very good estimate of the SD_{box}.*
- *The statistics*

$$\frac{observed - expected}{SE}$$

is not normally distributed.

Procedure for the t-test

Fact (t-test, (small sample size))

1. *Calculate the sample SD and correct for the small sample size:*

$$SD^+ = \sqrt{\frac{\#\text{ of measurements}}{\#\text{ of measurements} - 1}} \cdot SD_{sample}.$$

2. *We approximate SD_{box} by SD^+.*
3. *Form the t-statistic*

$$t = \frac{\text{observed} - \text{expected}}{SE}.$$

4. *t is not normally distributed so we need to look at the t-statistics table in the back of the book*
    - *degree of freedom=number of measurements -1.*

Example

Example

You fly Useless Air six times one year and the arrival times seem to always be late. The times were 30, 10, 40, 10, 40, and 50 minutes late. Is this due to the chance?

Example

Example

Suppose that a thermometer us being checked for a calibration. The temperature in the room is held at 70°F. Six measurements are taken: 72, 79, 65, 84, 67, 77. Is the thermometer calibrated?