Ordinary Least squares vs gradient descent

Parameter estimation – put CI on the parameters, likelihood of having an effect on final model?

Statistical significance tests:

Make an assumption about the distribution:

i.e. normal - two parameters u and sigma these two plug into this PDF:

Which describes the Gaussian distribution

Parametric Test:

T-Test – accepts/rejects null hypothesis.

Null hypothesis = a statement we are trying to **disprove** by running our test

e.g. two samples from the same population (e.g. right and left batters have no difference)

a sample is drawn from a probability distribution (e.g. have 20 baseball players, test how likely it is that those 20 people are in the MLB population).

Specified in terms of a test statistic.

Test Statistic: One number that helps accept or reject the null hypothesis, for a t-test this test statistic is T.

One sample: u = u0 – population mean = sample mean

Two sample : u0 = u1 – two population means are equal

T-tests make assumptions.

Equal sample sizes

Equal variance

Welches T-test does not assume these.

Degrees of freedom:

Where:

Calculate t and calculate 𝜈 to estimate p.

p-value is the probability of obtaining a test statistic **at least** as extreme as the result observed given the null hypothesis is true.

e.g. to test is left or right handed batters are better.

Null hypothesis = u0 = u1 (i.e. no difference between the two groups)

Alternate hypothesis: u0 <> u1 (i.e. there is a difference between the two groups)

A p-value of 0.05 would mean that 5 % of the time we would see a difference between the two groups due to random sampling error.

Set a critical value of p (e.g. α) e.g 0.01, 0.05, 0.1:

If p < α then:

Reject null hypothesis, i.e. there is a difference

Else if p > α then:

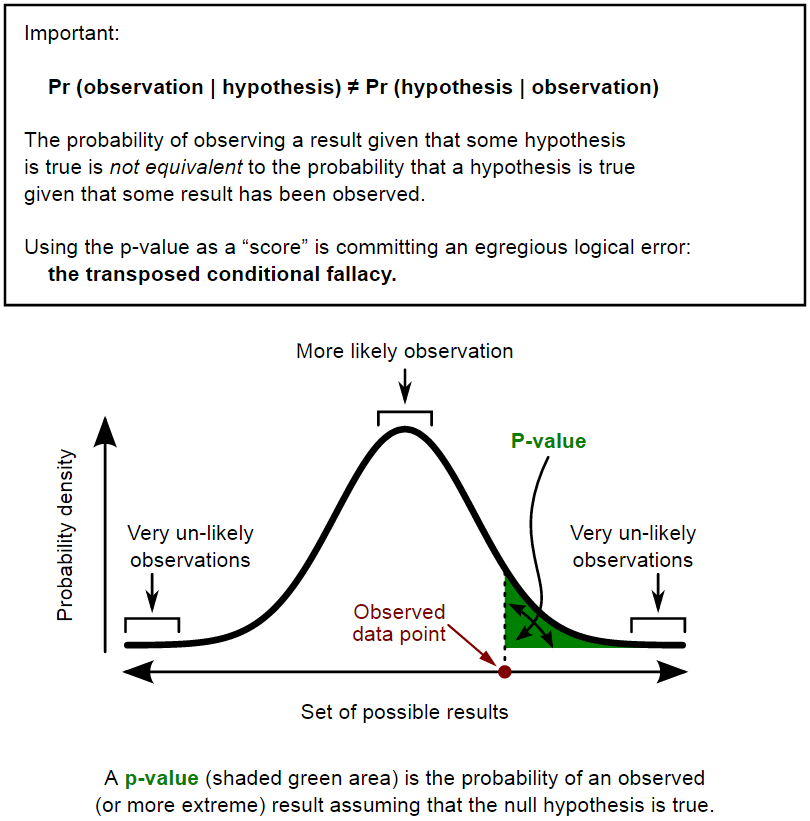
Accept null hypothesis i.e. that there is no difference

For α = 0.05 is less than a 5% chance of the test statistic being >= the observed result if the null hypothesis is true.

If you think the means are equal then the probability of the test statistic being further from the mean than the observed results is less than 5%.

One tailed - all 5% in one critical region, if know which way hypothesis is not correct,

Two tailed – split 2.5% into two critical regions if don’t know.



**A large p-value should not automatically be construed as evidence in support of the null hypothesis**. Perhaps the failure to reject the null hypothesis was caused by an inadequate sample size. When you see a large p-value in a research study, you should also look for one of two things:

1. a **power calculation** that confirms that the sample size in that study was adequate for detecting a clinically relevant difference; and/or
2. a **confidence interval** that lies entirely within the range of clinical indifference.

**The probability of observing the results from your sample with results more extreme, assuming the null hypothesis is true**. The smaller the p-value, the greater the inconsistency.