Lecture 38

Paired data

There are some situations where you have two observations (measurements) of the same variable for each individual.

eg individual	1	2	3	----	n
measurement 1	x_1	x_2	x_3	----	x_n
measurement 2	y_1	y_2	y_3	----	y_n

How do we handle these?

- Not two sample situation
- Work with differences

eg

$$D_1 = x_1 - y_1$$

$$D_2 = x_2 - y_2$$

$$\vdots$$

$$D_n = x_n - y_n$$

Some example situations

1. Before and after measurements
2. Dividing a sample unit into two halfs and then applying two different treatments. one to each half.
3. measurements in different locations
   eg top or bottom
   left or right
   inside or outside

Assuming that the X_i, Y_i's are normally distributed means the D_i's are also normally distributed (with mean μ_D and variance σ_D^2)

This knowledge leads us, in a similar manner as previous results, to the conclusion that

$$T = \frac{\bar{D} - \mu_D}{S_D/\sqrt{n}}$$ has t distribution with n-1 df

A CI for paired data

A $100(1-\alpha)\%$ CI for μ_D is

$$\bar{d} \pm t_{\alpha/2, n-1} \frac{s_D}{\sqrt{n}}$$

Hypothesis testing

Typically we will be testing for some specified

difference Δ_0. Often this will be 0.

To test the hypotheses

Null $H_0: \mu_D = \Delta_D$

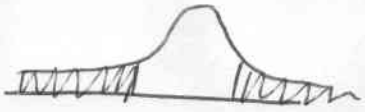
against any of the alternatives use

$$t = \frac{\bar{d} - \Delta_0}{s_0/\sqrt{n}}$$

with the region used to determine P-value (or rejection region) dependent on alternative hypothesis

$H_A: \mu_D > \Delta_0$ $\quad P(T > t)$ 

$H_A: \mu_D < \Delta_0$ $\quad P(T < t)$ 

$H_A: \mu_D \neq \Delta_0$ $\quad 2P(T > |t|)$ 

Matched Pairs

often in real life situations (particularly medical situations) it is not single individuals who get two measurements, but instead two individuals

are matched together on the basis of other variables (eg age, gender, weight, etc) that might also be responsible for changes in the variable of interest. Then one member of the pair gets assigned, randomly, one treatment and the other gets assigned a different treatment. The differences are then analysed.