Overview of lecture

- What is ANCOVA?
- Partitioning Variability
- Assumptions
- Examples
- Limitations

Analysis of covariance

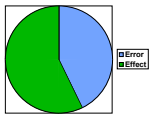
- Analysis of Covariance is used to achieve statistical control of error when experimental control of error is not possible.
- The Ancova adjusts the analysis in two ways:-
  - reducing the estimates of experimental error
  - adjusting treatment effects with respect to the covariate

Analysis of covariance

- In most experiments the scores on the covariate are collected before the experimental treatment
  - eg. pretest scores, exam scores, IQ etc
- In some experiments the scores on the covariate are collected after the experimental treatment
  - e.g.anxiety, motivation, depression etc.
- It is important to be able to justify the decision to collect the covariate after the experimental treatment since it is assumed that the treatment and covariate are independent.

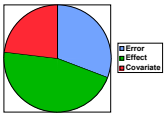
Partitioning variability in ANOVA

- In analysis of variance the variability is divided into two components
  - Experimental effect
  - Error - experimental and individual differences



Partitioning variability in ANCOVA

- In ancova we partition variance into three basic components:
  - Effect
  - Error
  - Covariate

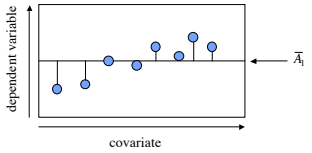


Estimating treatment effects

- When covariate scores are available we have information about differences between treatment groups that existed before the experiment was performed
- Ancova uses linear regression to estimate the size of treatment effects given the covariate information
- The adjustment for group differences can either increase of decrease depending on the dependent variables relationship with the covariate.

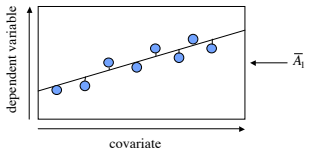
Error variability in ANOVA

- In between groups analysis of variance the error variability comes from the subject within group deviation from the mean of the group.
- It is calculated on the basis of the S/A sum of squares



Error variability in ANCOVA

- In regression the residual sum of squares is based on the deviation of the score from the regression line.
  - The residual sum of squares will be smaller than the S/A sum of squares
- This is how ANCOVA works



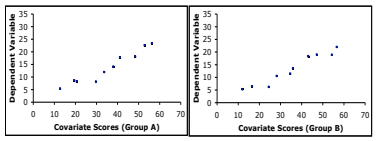
Assumptions of ANCOVA

- There are a number of assumptions that underlie the analysis of covariance
  - All the assumptions that apply to between groups ANOVA
    - normality of treatment levels
    - independence of variance estimates
    - homogeneity of variance
    - random sampling
  - Two assumptions specific to ANCOVA
    - The assumption of linear regression
    - The assumption of homogeneity of regression coefficients

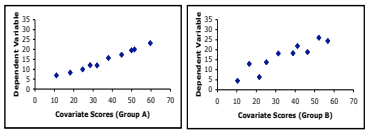
The assumption of linear regression

- This states that the deviations from the regression equation across the different levels of the independent variable have
  - normal distributions with means of zero
  - homoscedasticity.
- If linear regression is used when the true regression is curvilinear then
  - the ANCOVA will be of little use.
  - adjusting the means with respect to the linear equation will be pointless

Homoscedasticity - Equal Scatter



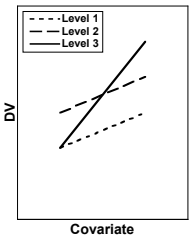
homoscedasticity



heteroscedasticity

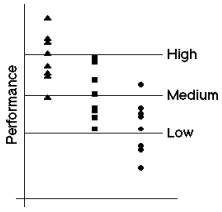
The homogeneity of regression coefficients

- Homogeneity of Regression Coefficients
  - The regression coefficients for each of the groups in the independent variable(s) should be the same.
  - Glass et al (1972) have argued that this assumption is only important if the regression coefficients are significantly different
  - We can test this assumption by looking at the interaction between the independent variable and the covariate



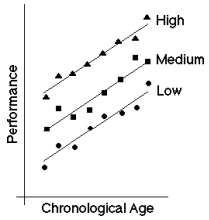
An Example Ancova

- A researcher is looking at performance on crossword clues.
- Subjects have been grouped into three vocabulary levels.
- An anova & tukeys on this data finds that the high group and low groups are different



An Example Ancova

- However amount of experience solving crosswords might make a difference.
- Plotting the scores against the age we obtain this graph.
- Ancova produces a significant effect of age and vocabulary. This time all the groups are significantly different



Example Results - ANOVA

	Mean	Std. Error
low	5.3750	.32390
medium	6.7500	.52610
high	7.6250	.32390
Total	6.5833	.29437

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	20.583	2	10.292	7.931	.003
Within Groups	27.250	21	1.298		
Total	47.833	23			

Example Results - ANCOVA

	Mean	Std. Error
low	5.679	.153
medium	6.699	.151
high	7.372	.152

Source	Sum of Squares	df	Mean Square	F	Sig.
AGE	23.623	1	23.623	130.268	.000
GROUP	11.042	2	5.521	30.445	.000
Error	3.627	20	.181		
Total	1088.000	24			

Example Results - Post hoc Tukey tests

		Mean Difference	Sig.
low	medium	-1.3750	.062
low	high	-2.2500	.002
medium	high	-.8750	.295

		Mean Difference	Sig.
low	medium	-1.021	.000
low	high	-1.694	.000
medium	high	-.673	.005

A Teaching Intervention Example

- Two groups of children either use maths training software or they do not.
- After using (or not using) the software the participants maths abilities are measured using a standardised maths test

Post-test results - Using t-test

GROUP	Mean	Std. Deviation
software	14.0000	4.18015
no software	14.1500	5.54669

t	df	Sig. (2-tailed)	Mean Difference
-.097	38	.924	-.1500

Problems with the design?

- It is quite possible that prior mathematical ability varies between the two groups of children
  - This needs to be taken into account
  - Prior to using the software the participants' maths abilities are measured using a standardised maths test

Post-test results - Using ANCOVA

	Mean	Std. Error
software	16.236	.573
no software	11.914	.573

Source	SS	df	MS	F	Sig.
PRETEST	703.676	1	703.676	122.307	.000
GROUP	145.472	1	145.472	25.285	.000
Error	212.874	37	5.753		
Total	8841.000	40			

- NB the means have been adjusted for the pre-test covariate

Limitations to analysis of covariance

- As a general rule a very small number of covariates is best
  - Correlated with the dv
  - Not correlated with each other (multi-collinearity)
- Covariates must be independent of treatment
  - Data on covariates be gathered before treatment is administered
  - Failure to do this often means that some portion of the effect of the IV is removed from the DV when the covariate adjustment is calculated.

Next Week

- MANOVA - Multivariate Analysis of Variance