Analysis of Covariance

1. Introduction

The Analysis of Covariance (generally known as ANCOVA) is a technique that sits between analysis of variance and regression analysis. It has a number of purposes but the two that are, perhaps, of most importance are:

1. to increase the precision of comparisons between groups by accounting to variation on important prognostic variables;
2. to "adjust" comparisons between groups for imbalances in important prognostic variables between these groups.

In what follows it is the second of these that will occupy most of our attention, although the example will also illustrate the advantage of the first purpose. One reason for this is the second purpose is one that is new in ANCOVA, the first being a purpose shared by many techniques you have already met, specifically simple linear regression. Another reason to concentrate on point 2 is that "adjustments" for other variables, or "adjusted means" or "adjusted relative risks" are often encountered in the medical literature. Adjusted relative risks can be found in virtually any piece of modern epidemiological research, a good example on hospital readmission can be found in Fisher *et al.*, *New England Journal of Medicine*, 1994, **331**, 989-995.

Whether relative risks, hazard rates or simply means are adjusted depends on whether the outcome measurement is, respectively, binary/categorical, survival data or continuous. In the following the example is based on an outcome that is taken to be continuous, because it is easier to demonstrate the ideas behind "adjustment" with this type of outcome. Nevertheless the underlying ideas are the same for all types of outcome.

The example concerns thyroid-associated ophthalmopathy. The data are from the first visit of each of 101 patients made to a combined thyroid-eye clinic and have been made available by kind permission of Dr P. Perros, Freeman Hospital. For our purposes we will consider only three of several variables measured, namely the age of the patient, the sex of the patient and the ophthalmic index (OI) which is a composite score measuring several aspects of ophthalmic performance, the larger the value of OI, the poorer the performance. Preliminary analysis suggested the analysis be based on the log the OI. More details of the study can be found in Perros *et al. Clinical Endocrinology*, 1993, **38,** 367-372.

As with most statistical techniques, the availability of software means that it is not necessary to know the numerical procedures needed in order to use ANCOVA, so most emphasis will be placed on understanding the approach behind the method.

2. The Data and the Problem

Interest surrounds whether the OI is worse for men or women. Superficially this can be answered by looking at the means in the two groups. Some summary statistics are:

	Sample size	Mean $\log_{10}$ OI (SD)
Male	20	0.885 (0.218)
Female	81	0.699 (0.215)

Thus it appears that the ophthalmic performance of the men is worse. Indeed, this is confirmed by a *t*-test, which gives P=0.0008 and a 95% confidence interval for the difference in log OI (males - females) of 0.08, 0.29.

However, before concluding that this analysis is all that is needed it is, as ever, wise to plot the data.

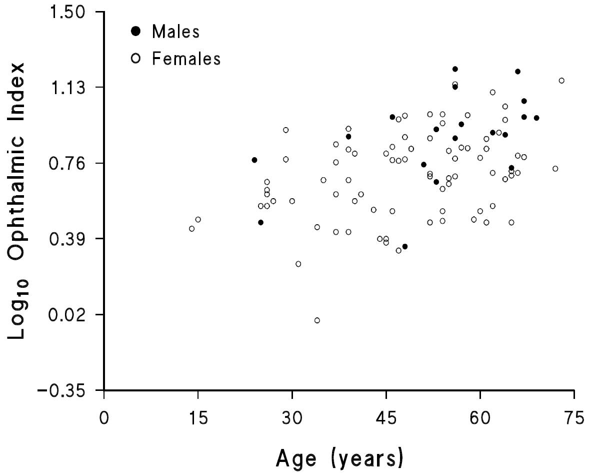


Figure 1: plot of log OI against age, with sex of patient indicated by type of symbol.

From figure 1 several things become apparent:

1 males clearly have larger values of the OI;
2 the ophthalmic index deteriorates with age;
3 there is a suggestion that the male patients may be older than the female patients.

If these observations are correct the preceding analysis is unreliable because the apparently larger OI in males may not reflect a difference between the sexes *per se* but simply that OI is larger in older patients and in this sample the males are older than the females. ANCOVA is a technique which attempts to make allowance for imbalances between groups and in this instance would try to determine whether there is a difference between the sexes in OI, independent of any age differences between the sexes that may exist.

The mean ages in the men and women are rather different, as the following table shows:

	Sample size	Mean age (yrs.) (SD)
Male	20	53.8 (12.8)
Female	81	48.2 (13.5)

A formal hypothesis test gives *t*=1.70, P=0.09 with 95% confidence interval for the difference in means, male -female, of -1.0, 12.3. Although there is no conclusive evidence of an age difference between the groups, this analysis certainly does not rule out this possibility. It may be, e.g., that there really is a difference in the ages of male and female patients with thyroid-associated eye disease but because there were only 20 males in the sample the test lacked the power to attain nominal levels of statistical significance. From this analysis it is important to realise that an age difference may exist and be sufficiently large to render any conclusion on the sex difference in OI unsafe unless some account is taken of age in the analysis.

3. Taking Account of Age

That age and log OI are related is confirmed by a simple regression analysis.

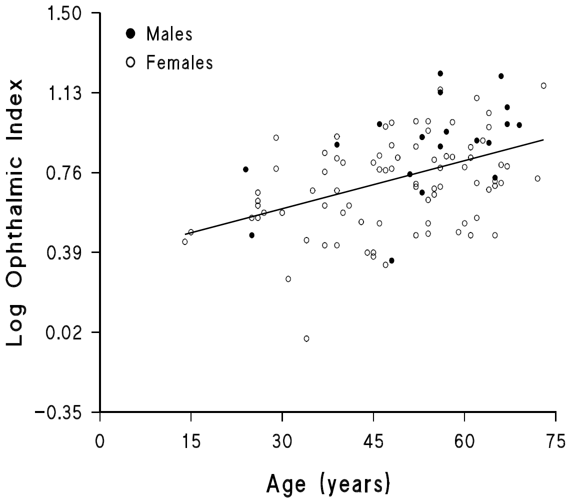


Figure 2: regression of $\log_{10}$ OI on age: log OI = 0.369 + 0.00743age, P<0.001

Figure 2 shows the relation between age and log OI and the accompanying regression line indicates quantitatively how the mean log OI changes with age. The idea behind ANCOVA is to extend this type of analysis: if differences in log OI due to age can be predicted then the differences in log OI between males and females that would be expected due to the age difference between these groups can also be predicted. Any difference between males and females beyond this prediction cannot be put down to differences in age.

This is achieved by fitting separate regression lines to males and females. However, for reasons that will become apparent this has to be done in a special way, as the fitted lines in the males and females must be parallel. If a line regression line is fitted to the females and the another one to the males, then the position in figure 3 would generally obtain, with the fitted lines not being parallel.

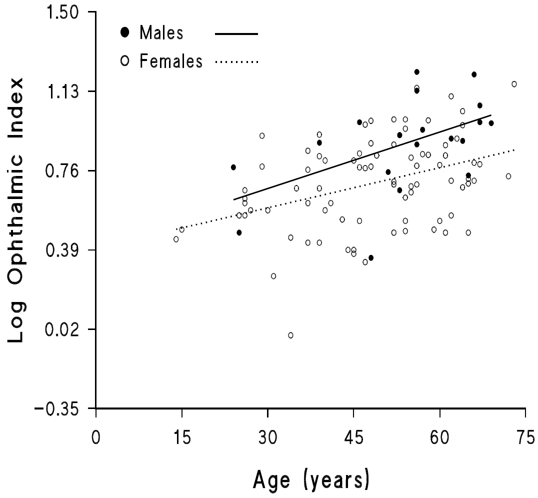


Figure 3: separate regression lines for males and females

Using a method, the details of which need not concern us, it is possible to fit a pair of regression lines, one for males and one for females, that are constrained to be parallel: (this method is essentially ANCOVA although in routine use it is not usual for the graph of the parallel lines to be displayed, indeed finding their equations from the output is not straightforward). When this is done the result is as in figure 4.

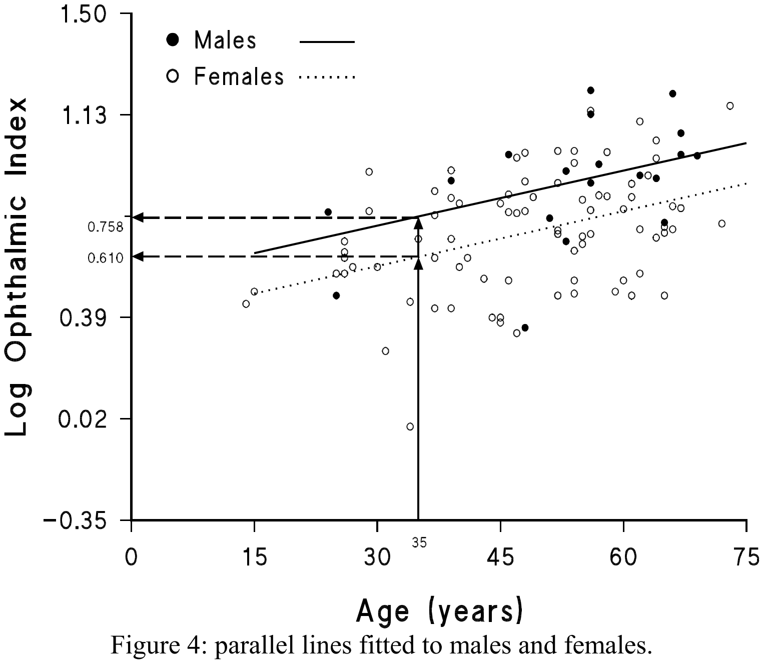


Figure 4: parallel lines fitted to males and females.

Two questions arise: first, how does this help us to "adjust" for sex differences and second, is it legitimate to fit parallel lines in this way?

The answer to the first point is embodied in the graph. As an example consider a 35-year-old patient, then the expected log OI is shown on the graph: if the patient is female it is 0.610, whereas for a male it is 0.758. The difference between these, 0.758-0.610 = 0.148, is the mean difference in log OI between males and females aged 35. However, because the regression lines are parallel, this difference is the same whatever the age of the patient. Consequently the mean sex difference, *adjusted for age* is simply the vertical difference between the lines. Any hypothesis test for a sex difference, adjusting for age, amounts to testing whether the sample provides evidence that the distance between the lines is zero or not. The adjustment for age arises because a method for modelling age has been incorporated into a comparison of males and females.

The adjustment is, strictly, only for the *difference* between the sexes and from figure 4 there is no obvious definition of the adjusted mean log OI for men and for women. A method, albeit rather arbitrary, that is often used is to find the mean log OI from figure 4 using not age 35, but the mean age in the combined sample, which is 49.3 years, giving:

adjusted mean log OI (males) = 0.854 adjusted mean log OI (females) = 0.707

These are the figures that the Minitab ANCOVA command will report.

The second problem mentioned above, namely whether the analyst is entitled to fit parallel lines, is important but often rather neglected in discussions of ANCOVA. Only a few general remarks will be made here.

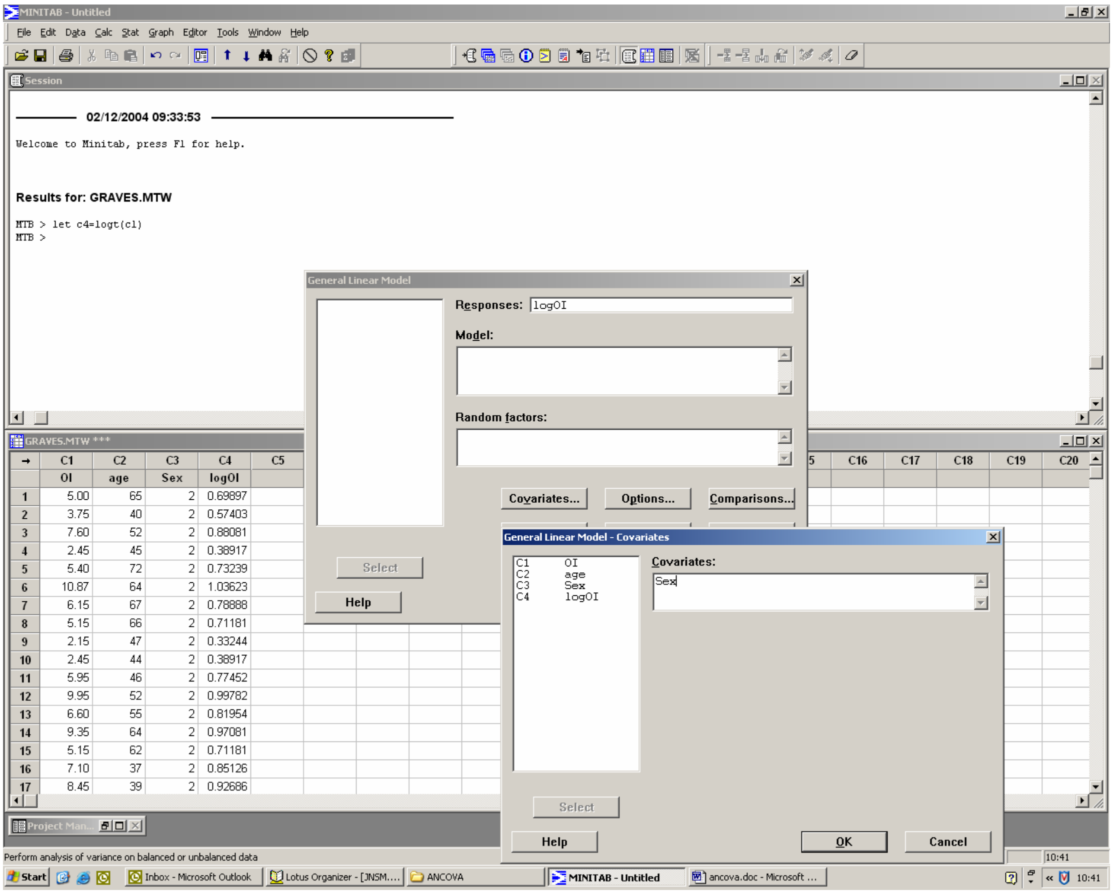
i) Clearly in most analyses the only justification for any statistical model is that it fits the data, and it may be that two non-parallel straight lines fit the data far better than two parallel lines. In the present example the separate lines in figure 3 are not far from parallel and the deviation from parallelism may be due to sampling error.

ii) If parallel lines cannot be fitted then it is impossible to report a single adjusted sex difference, as the vertical distance between the lines will change depending on the age of the patient. It must then be admitted that the method loses much of its appeal.

iii) In practice several options are available. A formal significance test of non-parallelism can be performed, although discussion of this is beyond the scope of the present lecture. If non-parallelism is a problem then some transformation of the response may be helpful.

4. Performing ANCOVA in Minitab

Using Minitab it is possible to perform an ANCOVA. Suppose three columns are available, 'logOI', 'age' and 'sex' (coded 1=male, 2=female), containing the data. The ANCOVA is performed by selecting General Linear Model… from the ANOVA part of the **Stat** menu. When confronted with the screen in figure 5, fill in the response as logOI. The model is the variable which, in terms of the present discussion, describes the groups, namely male and female, so Sex should be entered here. The covariate in this application is age and to enter this you need to click on the box marked Covariates… and enter age in the Covariates: box. The present example is the simplest possible: the 'model' could be complicated by having, e.g., a factorial structure and also several covariates can be handled at once.


MINITAB - Untitled
File Edit Data Calc Stat Graph Editor Tools Window Help
Session
General Linear Model
Model:
Random factors:
Covariates...
Options...
Comparisons...
Select
Help
General Linear Model - Covariates
Select
Help
OK
Cancel
Project Man...
Perform analysis of variance on balanced or unbalanced data
Start
Inbox - Microsoft Outlook
Lotus Organizer - [JNSM....
ANCOVA
MINITAB - Untitled
ancova.doc - Microsoft ...
10:41

```
——————   02/12/2004 09:33:53   ————————————————————

Welcome to Minitab, press F1 for help.


Results for: GRAVES.MTW

MTB > let c4=logt(c1)
MTB >
```

Responses: logOI

Covariates: Sex

```
C1   OI
C2   age
C3   Sex
C4   logOI
```

GRAVES.MTW ***

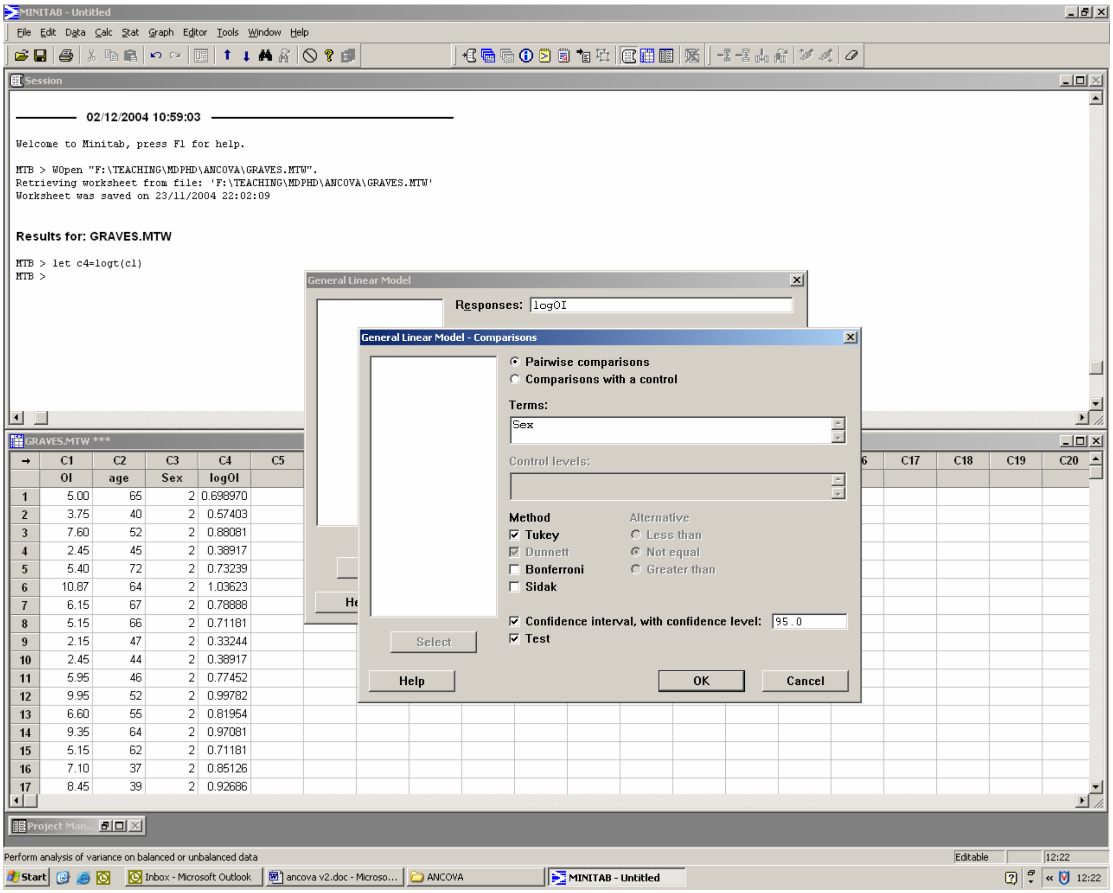
→	C1	C2	C3	C4	C5
	OI	age	Sex	logOI	
1	5.00	65	2	0.69897	
2	3.75	40	2	0.57403	
3	7.60	52	2	0.88081	
4	2.45	45	2	0.38917	
5	5.40	72	2	0.73239	
6	10.87	64	2	1.03623	
7	6.15	67	2	0.78888	
8	5.15	66	2	0.71181	
9	2.15	47	2	0.33244	
10	2.45	44	2	0.38917	
11	5.95	46	2	0.77452	
12	9.95	52	2	0.99782	
13	6.60	55	2	0.81954	
14	9.35	64	2	0.97081	
15	5.15	62	2	0.71181	
16	7.10	37	2	0.85126	
17	8.45	39	2	0.92686	

Figure 5: first screen encountered when performing ANCOVA

After clicking on OK on the General Linear Models - Covariates box it is worth exploring a few of the other buttons. The Results… button can be clicked and, among other items, a white box headed Display least squares means corresponding to the terms: is presented. By enetering Sex in this box you get the least squares means of log OI for males and females. This is just another name for the adjusted means discussed above.

A slight sang with ANCOVA in Minitab is that it does not routinely produce confidence intervals for differences between adjusted means. In this case, where we are comparing two groups, there is a device which allows the intervals to be computed. You need to click on the Comparisons… box and enter the grouping variable, Sex in this example, in the Terms: box: see below. When comparing just two groups, but not for more than two, the other selections in this dialogue box do not affect the output. The techniques in the comparisons box are largely aimed at allowing for multiple comparison when you compare more than two groups. Such techniques are rather overused and often used inappropriately, so the range of techniques (Tukey, Bonferroni etc.) available here, will apply adjustments to the widths of the confidence interval which you may well not want.

{n.b. For comparisons between just two groups such methods make no adjustments to the width of the interval, so the use of this option is perhaps the easiest way to obtain confidence intervals for the adjusted difference between males and females.}



MINITAB - Untitled
File Edit Data Calc Stat Graph Editor Tools Window Help
Session
GRAVES.MTW ***
Project Man...
Editable
12:22
Start
Inbox - Microsoft Outlook
ancova v2.doc - Microso...
ANCOVA
MINITAB - Untitled

```
02/12/2004 10:59:03

Welcome to Minitab, press F1 for help.

MTB > WOpen "F:\TEACHING\MDPHD\ANCOVA\GRAVES.MTW".
Retrieving worksheet from file: 'F:\TEACHING\MDPHD\ANCOVA\GRAVES.MTW'
Worksheet was saved on 23/11/2004 22:02:09

Results for: GRAVES.MTW

MTB > let c4=logt(cl)
MTB >
```

General Linear Model

Responses: logOI

General Linear Model - Comparisons

● Pairwise comparisons
○ Comparisons with a control

Terms: Sex

Control levels:

Method
☑ Tukey
☑ Dunnett
☐ Bonferroni
☐ Sidak

Alternative
○ Less than
● Not equal
○ Greater than

☑ Confidence interval, with confidence level: 95.0
☑ Test

Select | Help | OK | Cancel

→	C1	C2	C3	C4	C5
	OI	age	Sex	logOI	
1	5.00	65	2	0.698970	
2	3.75	40	2	0.57403	
3	7.60	52	2	0.88081	
4	2.45	45	2	0.38917	
5	5.40	72	2	0.73239	
6	10.87	64	2	1.03623	
7	6.15	67	2	0.78888	
8	5.15	66	2	0.71181	
9	2.15	47	2	0.33244	
10	2.45	44	2	0.38917	
11	5.95	46	2	0.77452	
12	9.95	52	2	0.99782	
13	6.60	55	2	0.81954	
14	9.35	64	2	0.97081	
15	5.15	62	2	0.71181	
16	7.10	37	2	0.85126	
17	8.45	39	2	0.92686	

Perform analysis of variance on balanced or unbalanced data

The output from this set of selections, plus annotations, is below.

General Linear Model: logOI versus Sex

```
Factor  Type   Levels  Values
Sex     fixed       2  1, 2


Analysis of Variance for logOI, using Adjusted SS for Tests

Source   DF   Seq SS   Adj SS   Adj MS      F      P
age       1  1.00327  0.79104  0.79104  20.33  0.000
Sex       1  0.34051  0.34051  0.34051   8.75  0.004
Error    98  3.81409  3.81409  0.03892
Total   100  5.15786
```

[the above shows that both the age and the sex of the patient affects the log OI. The only items you really need from this table are the P-values. In fact it is largely the P-value associated with Sex that you need: the value of 0.004 indicates that there is strong evidence of a difference between males and females, even after adjusting for age. The P-value associated with age merely shows that age is related to log OI and confirms we were right to take account of it in the analysis. However, the P-value for age is of limited interest in itself]

```
S = 0.197280   R-Sq = 26.05%   R-Sq(adj) = 24.54%
```

[S is the residual standard deviation about the fitted lines, rather as in a regression analysis]

```
Term           Coef    SE Coef     T      P
Constant   0.45046   0.07965  5.66  0.000
age       0.006693  0.001485  4.51  0.000
```

[This part of the output is only of indirect interest: the coefficient of age is the slope of the parallel lines]

```
Unusual Observations for logOI
```

Output on unusual observations suppressed (beyond scope of present discussion)

```
Means for Covariates

Covariate   Mean  StDev
age        49.32  13.48


Least Squares Means for logOI

Sex    Mean  SE Mean
1    0.8544  0.04462
2    0.7066  0.02198
```

[These are the adjusted means and are amongst the most important part of the output]

```
Tukey 95.0% Simultaneous Confidence Intervals
Response Variable logOI
All Pairwise Comparisons among Levels of Sex
Sex = 1  subtracted from:

Sex    Lower   Center     Upper  -----+---------+---------+---------+-
2    -0.2470  -0.1478  -0.04864  (-------------*-------------)
                                 -----+---------+---------+---------+-
                                   -0.210    -0.140    -0.070     0.000
```

[The difference in adjusted means (female – male) together with 95% confidence interval; another important part of the output]

```
Tukey Simultaneous Tests
Response Variable logOI
All Pairwise Comparisons among Levels of Sex
Sex = 1  subtracted from:

     Difference       SE of           Adjusted
Sex    of Means  Difference  T-Value   P-Value
2       -0.1478     0.04997   -2.958    0.0039
```

[This gives the difference in the adjusted means of log OI: females – males. The Adjusted P-value (i.e. after adjustment for age) is,as given in the Analysis of variance table, albeit to more decimal places]

5. Summary and Concluding Remarks

The method of ANCOVA allows the analyst to make comparisons between groups that are not comparable with respect to some important variable, often referred to as a covariate. This is done by making an adjustment based on fitting a particular kind of regression line. When the imbalance between the groups is not large this method may be very helpful, however it is worth bearing in mind that the "adjustment" is done through a particular statistical model and it may be unwise to rely on such a device to bring into balance two highly divergent groups.

In addition to allowing for imbalances, the method removes variation due to the covariate and therefore provides a more precise analysis. A geometrical interpretation is that the 'unexplained variation' with respect to which the significances of group differences are ultimately assessed is the variation about the lines in figure 4, whereas an analysis ignoring the covariate would use the variation about the group means, which will clearly be greater.