11.5 Analysis of Covariance (ANCOVA)

In an ideal experimental design we would choose all of our subjects to be similar (i.e. replicate measurements), except for the specific treatments that are applied as part of the experiment. However, in practice, the available subjects may differ in many different ways.

In some circumstances we can allow for some differences between subjects if we know how their differences may affect the experimental response variable.

Consider a simple example where a factory is developing a new production process, and wishes to see if method A is faster than method B.
Two groups of workers are to be allocated to carry out the **two different methods, A and B**, and each worker records the time taken, *T*, as being the measured **response variable**.

The problem in the experiment design is that there are only 14 workers available to carry out the test, and their individual performance is also affected by the number of **hours experience, *n*,** that each of them already has on this type of process.
Their individual times, *T*, would be **covariant** with the number of hours experience, *n*, they have had.
The workers are allocated to groups A and B alternately in pairs with increasing experience (or randomly for those with the same experience).

The results are given in the table below

Method A

Employee number	8	6	5	12	2	1	13
Hours experience, *n*	13	16	17	19	19	21	25
Time, *T*	189	179	156	152	172	160	135

Method B

Employee number	10	14	9	3	4	7	11
Hours experience, *n*	13	15	17	18	19	20	27
Time, *T*	180	159	150	154	157	135	120

Scatter plots (Figure 11.7) of time, *t*, against experience hours, *n*, for the two methods shows:

- The time taken, *t*, decreases with increased experience, *n* - a best-fit straight line (bold) is drawn using all data.

- 'Best-fit' straight lines drawn separately for each method, appear to show that data for method B is to the bottom-left and data for method A tends to the top-right.

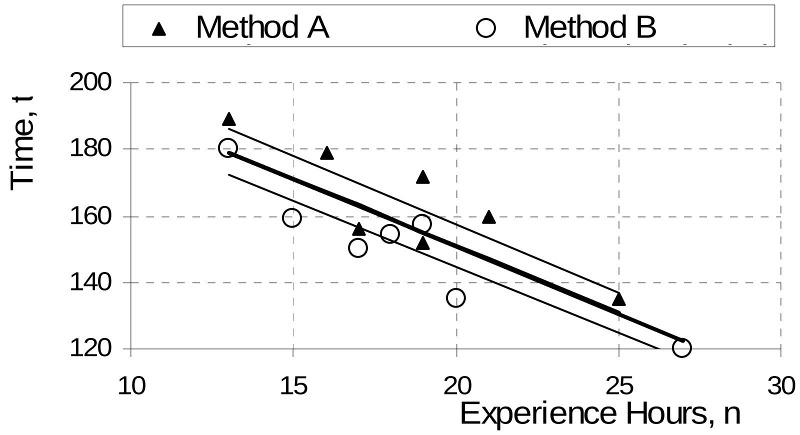

Figure 11.7: Time against Experience

If we **ignore the effect of experience** on the employees' times, a two-sample *t*-test between the times, *T*, for the two methods gives

$p = 0.230$,

which suggests initially that there is NO significant difference between the two methods.

The simple two-sample *t*-test on time only looks for difference in **vertical** values in the Figure 11.7, and ignores the apparent displacement in horizontal values.

However we can apply an **'adjustment'** to the time (the response variable), by taking into account the hours practiced for each particular subject.
The process of 'adjustment' is as follows:

- Calculate the coefficients ('slope', *m*, and 'intercept', *c*) for the average **'best-fit' straight line** through the data.
- Calculate the **residuals** (p320) between each data point and the 'average' best-fit straight line.
- Perform a *t*-test for the values of residuals in each sample.

The calculations are given in the Excel file: ANCOVA.xls.

The plot of **residual times v hours of experience** is given in Figure 11.8.

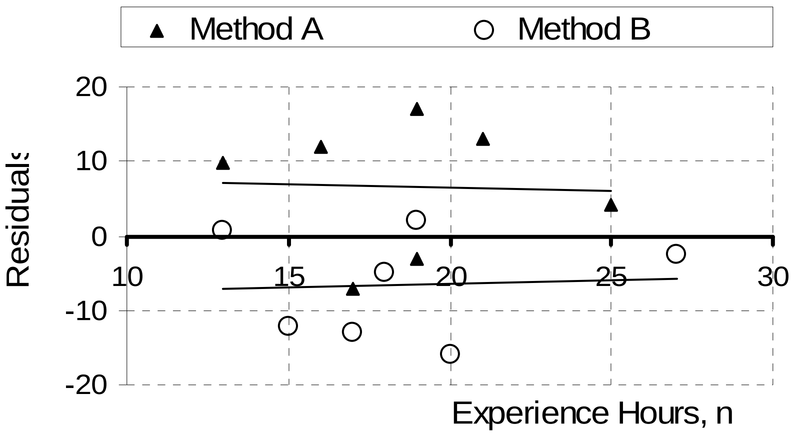

Figure 11.8: Residual Times against Experience

The differences between methods A and B now appear just as 'vertical' differences in the 'residuals' graph.

A two-sample *t*-test between the times, *T*, for the two methods gives

$p = 0.010$,

which suggests that there IS a significant difference between the two methods.

In Minitab, a univariate ANCOVA calculation can be performed by using the **General Linear Model**, and specifying the variables that are acting as the covariates.

Minitab calculations give:

- *p*-value *without* including covariance: $p = 0.230$
- *p*-value *including* covariance: $p = 0.014$

The covariance calculation used by Minitab is slightly different from the process given above, but the principle is the same and does allow for uncertainties in the calculation of the coefficients of the best-fit straight line. The calculation in specialist software should be taken as more reliable than the two-step process used above as a demonstration.

Two-Sample T-Test and CI: Time, Group

```
Two-sample T for Time

Group  N   Mean  StDev  SE Mean
A      7  163.3   18.1      6.9
B      7  150.7   19.0      7.2


Difference = mu (A) - mu (B)
Estimate for difference:  12.5714
95% CI for difference:  (-9.0680, 34.2108)
T-Test of difference = 0 (vs not =): T-Value = 1.27  P-Value = 0.230  DF =
12
Both use Pooled StDev = 18.5806
```

Correlations: Time, Practice

```
Pearson correlation of Time and Practice = -0.840
P-Value = 0.000
```

General Linear Model: Time versus Group

```
Factor     Type Levels Values
Group     fixed      2 A B

Analysis of Variance for Time, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Group       1      553.1      553.1      553.1    1.60  0.230
Error      12     4142.9     4142.9      345.2
Total      13     4696.0
```

General Linear Model: Time versus Group with covariate Practice

```
Factor     Type Levels Values
Group     fixed      2 A B

Analysis of Variance for Time, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Practice    1     3312.0     3363.6     3363.6   47.48  0.000
Group       1      604.7      604.7      604.7    8.54  0.014
Error      11      779.3      779.3       70.8
Total      13     4696.0
```