MA4605 Lecture 11A : Experimental Design

From Last class

Suppose there are 24 observations divided evenly between 4 treatment groups.

Complete the ANOVA table and comment on the outcome of this test.

	DF	Sum Sq	Mean Sq	F value	
Treatment	**....**	**....**	**....**	**....**	3.59e-06 ***
Residuals	**....**	**....**	**6.51**		
Total	**....**	**513**			

1) There are 23 Degrees of freedom in total. (24-1) and 3 degrees of freedom for treatments. Necessarily that means that there are 20 degrees of freedom of residuals.

	DF	Sum Sq	Mean Sq	F value	
Treatment	**3**	**....**	**....**	**....**	3.59e-06 ***
Residuals	**20**	**130.2**	**6.51**		
Total	**23**	**513**			

2) The Mean Square for Residuals is the Sum of Square for Residuals divided by the associated degrees of freedom.
   Therefore the value for Sum of Squares is ***130.2***

	DF	Sum Sq	Mean Sq	F value	
Treatment	**3**	**382.8**	**….**	**….**	3.59e-06 ***
Residuals	**20**	**130.2**	**6.51**		
Total	**23**	**513**			

3) The sums of squares add up to 513. Necessarily the Sum of Squares for Treatment is (513-130.2 =) ***382.8***

	DF	Sum Sq	Mean Sq	F value	
Treatment	**3**	**382.8**	**127.60**	**19.61**	3.59e-06 ***
Residuals	**20**	**130.2**	**6.51**		
Total	**23**	**513**			

4) The mean square for treatment is the ratio of the sum of squares to the degrees of freedom (i.e. 382.8/3=) ***127.60***. The test statistic (i.e. the F-value) is the ratio of the two mean square values, i.e. (127.60/6.51 =) ***19.61***.

The p-value is very low. We reject the null hypothesis that all treatments have the same effect.

Two Way ANOVA – Randomized Block Design

In last week's class we considered the case of a completely randomized experimental design, which was analysed using the one-way ANOVA procedure.

Previously we have discussed the concept of blocking. Recall that blocking helps to mitigate or remove the effect of nuisance variables. However the blocking variable is not of interest itself.

The general procedure for a randomized block design consists of selecting **b** blocks and running a randomized design within each block. Recall that this can be described as applying ***a*** treatments

Following from the statistical model seen in last class, this can be described using the following model, with the addition of a term $\boldsymbol{\beta}$ to describe the effect for each block

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases}$$

The terms are the overall mean, the treatment effect term, the block effect term and the random error term (i.e. residual term).The random error terms are assumed to be normally distributed with mean of zero, and independently identical distributed.

The Treatment effect and Block effect are deviations from the overall mean, so the expected value of each is zero.

We also assume that treatments and blocks do not interact. (This is in contrast with forthcoming material, where interaction is normally assumed)

We are usually interested in testing the equality of treatment effects.

$$H_0: \tau_1 = \tau_2 = \cdots = \tau_a = 0$$
$$H_1: \tau_i \neq 0 \text{ at least one } i$$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Treatments	$SS_{\text{Treatments}}$	$a - 1$	$\dfrac{SS_{\text{Treatments}}}{a-1}$	$\dfrac{MS_{\text{Treatments}}}{MS_E}$
Blocks	SS_{Blocks}	$b - 1$	$\dfrac{SS_{\text{Blocks}}}{b-1}$	
Error	SS_E (by subtraction)	$(a-1)(b-1)$	$\dfrac{SS_E}{(a-1)(b-1)}$	
Total	SS_T	$ab - 1$		

Example 1

A standard solution was prepared, containing 16.00% (by weight) of chloride. Three titration methods, each with a different technique of end-point determination, were used to analyse the standard solution.

The procedure was carried out by different clinical analysts. The order of the experiments was randomized. The results for the chloride found (% w/w) are shown below:

Analyst	Method A	Method B	Method C
1	16.03	16.13	16.09
2	16.05	16.13	16.15
3	16.02	15.94	16.12
4	16.12	15.97	16.10

Here the treatment is the titration method and we are interested in determining if there is uniformity between each method. Four analysts performed an experiment using each of the titration methods. This allows the analysts to remove any effect due to the analysts.

To construct the model using ***R***, we use the `aov()` command, specifying the treatment factor and the blocking. Importantly we express the model additively (i.e”. `.Meth+Anlt..”`).

```
> Model=aov(Titr ~ Meth + Anlt)
>
> summary(Model)
            Df  Sum Sq  Mean Sq F value  Pr(>F)
Meth         2 0.01202 0.006008   1.279   0.345
Anlt         3 0.01109 0.003697   0.787   0.543
Residuals    6 0.02818 0.004697
```

This additive model is the appropriate specification if an interaction is not assumed. A variant of this ***R*** implementation, that does specify an interaction term, will be discussed in due course.

Our conclusion of this procedure is that there is no difference in the titration method.

Example 2

Assume that we have three fertilizers to be tested. We wish to determine if there is any difference is the mean yields for the three different types of fertilizer.

Fertilizer	Yields			
A	5.6	6.4	6.6	5.8
B	5.1	6.2	6.4	5.7
C	5.0	6.1	5.8	5.5

The following R output is a ***One-Way ANOVA*** procedure for testing multiple means.

```
> ModelA=aov(Yield~Fert,Study1)
>
> summary(ModelA)
             Df Sum Sq Mean Sq F value Pr(>F)
Fert          2   0.50  0.2500   0.957   0.42
Residuals     9   2.35  0.2611
```

Further to this test, we conclude that there are no differences between the three fertilizer types.

An agricultural analysts points out that there should a blocking to account for the different types of field. Field 1 is very boggy, Field 2 and Field 3 are reasonably good, and Field 4 is full of rocks.

The procedure was carried out again. Three plots of land in each field were given the fertilizer treatment.

	Blocks			
Fertilizers	1	2	3	4
A	5.6	6.4	6.6	5.8
B	5.1	6.2	6.4	5.7
C	5.0	6.1	5.8	5.5

The following R output is a ***Two-Way ANOVA*** procedure for analysing this data

```
> summary(Model2)
            Df Sum Sq Mean Sq F value   Pr(>F)
Fert         2 0.5000  0.2500   10.23 0.011667 *
Block        3 2.2033  0.7344   30.05 0.000519 ***
Residuals    6 0.1467  0.0244
```

This output informs us that there is a difference between fertilizer treatments, contrary to what was previously thought.