M3S12 BIOSTATISTICS: ASSESSED COURSEWORK 1

Deadline: Friday 12th March

1. The percentage protein serum binding capability of a number of different antibiotics was measured

	Penicillin G	Tetra-cyclin	Strepto-mycin	Erythro-mycin	Chloram-phenicol
	29.6	27.3	5.8	21.6	29.2
	24.3	32.6	6.2	17.4	32.8
	28.5	30.8	11.0	18.3	25.0
	32.0	34.8	8.3	19.0	24.2
Mean	28.6	31.4	7.8	19.1	27.8

Carry out a one-way ANOVA to assess whether there is any evidence of a difference between the antibiotic binding capabilities. You may use the following SPLUS code, or implement the analysis by hand using a calculator.

```
serum.type<-rep(c(1:5),4)
serum.binding<-c(29.6,27.3,5.8,21.6,29.2,24.3,32.6,6.2,17.4,32.8,28.5,30.8,11.0,18.3,25.0,32.0,34.8,8.3,19.0,24.2)
serum.data<-data.frame(serum.type,serum.binding)
names(serum.data)<-c("TYPE","BINDING")
summary(aov(BINDING~factor(TYPE),data=serum.data))
```

[8 MARKS]

2. The following study investigates the effect of a new drug combined with therapy on the **behaviour score** (a positive, continuous measure) of a cohort of schizophrenic patients. The **drug** has three dose levels (**Absent**, **Low**, **High**) and the **therapy** has four types (**BMod -** behaviour modification; **Psycho -** psychodynamic; **Group -** group counselling; **Nondir -** nondirective counselling) in a 3×4 **factorial design**. An elementary numerical summary of the cross-classified data is presented below (values are mean (standard deviation) within each group; a common sample size of 5 is used for each cross-classification).

Behaviour score		Therapy			
		BMod	Psycho	Group	Nondir
Drug	Absent	23.8 (1.30)	20.6 (2.70)	20.0 (2.12)	25.4 (1.14)
	Low	25.2 (4.15)	26.6 (2.07)	23.0 (2.45)	25.0 (2.74)
	High	21.8 (1.92)	23.2 (2.17)	25.4 (2.30)	22.2 (2.59)

Conduct a two-way ANOVA **with interaction** using the partially filled ANOVA-table below.

Source	D.F.	Sum of squares	Mean square	F
DRUG	*	66.533	*	*
THERAPY	*	*	4.950	*
INTERACTION	6	156.800	*	*
Residual	*	518.983		
Total	59	*		

Discuss any features of the data that might invalidate the ANOVA results.

[12 MARKS]

NOTE: The usual sum of squares decomposition is valid even in the **interaction** model case; however, there is an additional sum of squares term,

$$TSS = FSS_1 + FSS_2 + ISS + RSS$$

where ISS is the interaction sum of squares, which measures the amount of variability explained by the interaction between the two factors. The degrees of freedom for an interaction term for a two-way cross classification with K and L levels for the two factors respectively is

$$(K-1) \times (L-1)$$

and we have that, under the assumption that the interaction terms, γ_{kl}, in the model

$$y_{klj} = \beta_k + \delta_l + \gamma_{kl} + \varepsilon_{klj}$$

are **zero,** then

$$\frac{ISS}{\sigma^2} \sim \chi^2_{(L-1)(K-1)}$$