Worksheet 7 - Randomized block, Split-plot and repeated measures

Multifactor ANOVA

- Quinn & Keough (2002) - Chpt 10-11

Question 1 - Randomized block

A plant pathologist wanted to examine the effects of two different strengths of tobacco virus on the number of lesions on tobacco leaves. She knew from pilot studies that leaves were inherently very variable in response to the virus. In an attempt to account for this leaf to leaf variability, both treatments were applied to each leaf. Eight individual leaves were divided in half, with half of each leaf inoculated with weak strength virus and the other half inoculated with strong virus. So the leaves were blocks and each treatment was represented once in each block. A completely randomised design would have had 16 leaves, with 8 whole leaves randomly allocated to each treatment.

Format of tobacco.csv data files

LEAF	TREAT	NUMBER
1	Strong	35.898
1	Week	25.02
2	Strong	34.118
2	Week	23.167
3	Strong	35.702
3	Week	24.122
...	...	...

LEAF The blocking factor - Factor B

TREAT Categorical representation of the strength of the tobacco virus - main factor of interest Factor A

NUMBER Number of lesions on that part of the tobacco leaf - response variable

Open the tobacco data file.

Since each level of treatment was applied to each leaf, the data represents a randomized block design with leaves as blocks.

The variable LEAF contains a list of Leaf idenfification numbers and is supposed to represent a factorial blocking variable. However, because the contents of this variable are numbers, R initially treats them as numbers, and therefore considers the variable to be numeric rather than categorical. In order to force R to treat this variable as a factor (categorical) it is necessary to first **convert this numeric variable into a factor** (HINT).

Q1-1. What are the main hypotheses being tested?

a. H_0 Factor A:

b. H_0 Factor B:

Q1-2. In the table below, list the assumptions of a randomized block design along with how violations of each assumption are diagnosed and/or the risks of violations are minimized.

a.

Assumption	Diagnostic/Risk Minimization
I.	
II.	
III.	
IV.	

b. Is the proposed model **balanced?** (Yor N)

Q1-3. Plot the number of lesions for each treatment and leaf combination (ie. an **interaction plot** (HINT). Any evidence of an interaction? Note that we cannot formally test for an interaction because we don't have replicates for each treatment-leaf combination!

Q1-4. Analyse these data with a **randomized block ANOVA** (HINT) to test the H_0 that there is no difference in the number of lesions produced by the different strength viruses. Complete the table below.

Source of variation	df	Mean Sq	F-ratio	P-value
LEAF (block)				
TREAT				
Residuals				

Q1-5. Interprete the results.

Question 2 - Simple Repeated measures

Driscoll and Roberts (1997) investigated the impact of fuel-reduction burning on the number of individual male frogs calling. Matched burnt and unburnt sites were blocked within six drainages, and the difference in number of calling male frogs between the sites was recorded for each drainage on three occasions (a 1992 pre-burn and two post burns in 1993 and 1994). They were primarily interested in investigating whether the mean difference in number of calling frogs between burn and control sites differed between years.

Format of driscoll.csv data file

BLOCK	YEAR	CALLS
logging	1	4
angove	1	-10
newpipe	1	-15
oldquinE	1	-14
newquinW	1	-4
newquinE	1	0
logging	2	17
...	...	...

BLOCK The name of the catchments - the blocking factor - (Factor B)

YEAR Categorical listing of the years in which the measurements were taken (1=1992 pre-burn, 2=1993 post-burn and 3=1994 post-burn) - main factor of interest (Factor A)

CALLS Difference in the number of calling male frogs between burnt and unburnt sites - response variable

Open the driscoll data file. HINT. Since years are listed as numbers, ensure that you define YEAR as a factorial variable (HINT).

Q2-1. What are the main hypotheses being tested?

a. H_0 Main Effect 1 (Factor A):

b. H_0 Blocking Effect (Factor B):

Q2-2. What are the assumptions of a randomized block design?

Q2-3 Test and comment on the above assumptions by;

a. Investigate normality and homogeneity of variance for the H_0 Main Effect 1 (Factor A): (HINT)

b. Investigate the likelihood of an interaction between the blocking factor and the main treatment factor (HINT). Note that if an interaction is likely, the test of the blocking factor many not be reliable.

c. Investigate the assumption of compound symmetry by calculating the **Greenhouse-Geisser and Huynh-Feldt epsilons**(HINT). Is there any evidence that the assumption has been violated?

d. Check whether or not the model is **balanced?**. (HINT)

Q2-4. Analyse these data with a **repeated measures (same as randomized block) ANOVA** to test the H_0 (HINT) that there is no difference in the mean difference in number of calling male frogs between the burnt and control sites between the different years. Check the diagnostics (primarily residual plot, HINT). If these summaries do not reveal any additional assumption violations, complete the table below (HINT).

Source of variation	df	Mean Sq	F-ratio	P-value	GG.P	HF.P
BLOCK (=block)						
YEAR						
Residuals						

Q2-5. What conclusions would you draw?

Q2-6. In addition to establishing a difference between years, the researchers might also have been interested in determining whether there was a linear trend in burnt-unburnt differences in frog numbers

through the years. **Re-fit the model with linear polynomial contrasts and since there was a very real possibility that sphericity was not met, the error term from the original model should not be used.**

Source of variation	df	Mean Sq	F-ratio	P-value	GG.P	HF.P
BLOCK (=block)						
YEAR: Linear						
Residuals						

Q2-7. What conclusions would you draw?

Question 3 - Split-plot

In an attempt to understand the effects on marine animals of short-term exposure to toxic substances, such as might occur following a spill, or a major increase in storm water flows, a it was decided to examine the toxicant in question, Copper, as part of a field experiment in Honk Kong. The experiment consisted of small sources of Cu (small, hemispherical plaster blocks, impregnated with copper), which released the metal into sea water over 4 or 5 days. The organism whose response to Cu was being measured was a small, polychaete worm, Hydroides, that attaches to hard surfaces in the sea, and is one of the first species to colonize any surface that is submerged. The biological questions focused on whether the timing of exposure to Cu affects the overall abundance of these worms. The time period of interest was the first or second week after a surface being available.

The experimental setup consisted of sheets of black perspex (settlement plates), which provided good surfaces for these worms. Each plate had a plaster block bolted to its centre, and the dissolving block would create a gradient of [Cu] across the plate. Over the two weeks of the experiment, a given plate would have pl ain plaster blocks (Control) or a block containing copper in the first week, followed by a plain block, or a plain block in the first week, followed by a dose of copper in the second week. After two weeks in the water, plates were removed and counted back in the laboratory. Without a clear idea of how sensitive these worms are to copper, an effect of the treatments might show up as an overall difference in the density of worms across a plate, or it could show up as a gradient in abundance across the plate, with a different gradient in different treatments. Therefore, on each plate, the density of worms (#/cm^2) was recorded at each of four distances from the center of the plate.

Format of copper.csv data file

COPPER	PLATE	DIST	WORMS
..	..	..	..

COPPER Categorical listing of the copper treatment (control = no copper applied, week 2 = copper treatment applied in second week and week 1= copper treatment applied in first week) applied to whole plates. Factor A (between plot factor).

PLATE Substrate provided for polychaete worm colonization on which copper treatment applied. These are the plots (Factor B). Numbers in this column represent numerical labels given to each plate.

DIST Categorical listing for the four concentric distances from the center of the plate (source of copper treatment) with 1 being the closest and 4 the furthest. Factor C (within plot factor)

WORMS Density (#/cm_2) of worms measured. Response variable.

Open the copper data file. HINT. Notice that both the PLATE variable and the DIST variable contain only numbers. Make sure that you define both of these as factors (HINT)

Q3-1. What are the main hypotheses being tested?

a. H_0 Main Effect 1 (Factor A):

b. H_0 Main Effect 2 (Factor C):

c. H_0 Main Effect 3 (A*C):

Q3-2. The usual ANOVA assumptions apply to split-plot designs, and these can be tested by constructing boxplots for each of the main effects. However, it is important to consider what data the boxplots should be based upon. For each of the main hypothesis tests, describe what data should be used to construct boxplots (remember that the assumptions of normality and homogeneity of variance apply to the residuals of each hypothesis test, and therefore the data used in the construction of boxplots for each hypothesis test should represent the respective residuals, or sources of unexplained variation).

a. H_0 Main Effect 1 (Factor A):

b. H_0 Main Effect 2 (Factor C):

c. H_0 Main Effect 3 (A*C):

Q3-3. For each of the hypothesis tests, indicate which Mean Square term should be used as the residual (denominator) in the F-ratio calculation. Note, COPPER and DIST are fixed factors and PLATE is a random factor.

a. H_0 Main Effect 1 (Factor A): F-ratio = MS_{COPPER}/MS

(choose correct option)

b. H_0 Main Effect 2 (Factor C): F-ratio = MS_{DIST}/MS (choose correct option)

c. H_0 Main Effect 3 (A*C): F-ratio = $MS_{COPPER:DIST}/MS$ (choose correct option)

Q3-4. Construct a boxplot to investigate the assumptions as they apply to the test of H_0 Main Effect 1 (Factor A): This is done in two steps

a. **Aggregate the data set by the mean number of WORMS within each plate**(HINT)

b. **Construct a boxplot of aggregated mean number of WORMS against COPPER treatment**(HINT)

c. Any evidence of violations of the assumptions (y or n)?

Q3-5. Construct a boxplot to investigate the assumptions as they apply to the test of H_0 Main Effect 2 (Factor C): Since Factor C is tested against the overal residual in this case, this is a relatively straight forward procedure.(HINT)

a. Any evidence of violations of the assumptions (y or n)?

Q3-6. Construct a boxplot to investigate the assumptions as they apply to the test of H_0 the main interaction effect (A:C): Since A:C is tested against the overal residual, this is a relatively straight forward procedure.(HINT)

a. Any evidence of violations of the assumptions (y or n)?

Q3-7. In addition to the above assumptions, the test of PLATE assumes that there is no PLATE by DIST interaction as this is the overal residual (the replicates). That is, the test assumes that the effect of DIST is consistent in all PLATES. Construct an interaction plot to examine whether there is any evidence of an interation between PLATE and DISTANCE (HINT)

a. Any evidence of an interaction (y or n)?

b. Is the design (model) **unbalanced** ?(HINT) (Yor N)

Q3-8. Write out the linear model

Q3-8. Perform a **split-plot ANOVA** (HINT), and complete the following table (HINT). To obtain the hypothesis test for the random factor (Factor B: PLATE), run the model as if all factors were fixed and thus all terms are tested against the overall residuals, HINT)

Source of variation	df	Mean Sq	F-ratio	P-value
COPPER				
PLATE				
DIST				
COPPER:DIST				
Residuals				

Q3-10. Construct an **interaction plot** showing the density of worms against distance from Cu source, with each treatment as different lines (or different bars). HINT

Q3-11. What conclusions would you draw from the analysis (and graph)?

Question 4 - Repeated Measures

In an honours thesis from (1992), Mullens was investigating the ways that cane toads (Bufo marinus) respond to conditions of hypoxia. Toads show two different kinds of breathing patterns, lung or buccal, requiring them to be treated separately in the experiment. Her aim was to expose toads to a range of O_2 concentrations, and record their breathing patterns, including parameters such as the expired volume for individual breaths. It was desirable to have around 8 replicates to compare the responses of the two breathing types, and the complication is that animals are expensive, and different individuals are likely to have different O_2 profiles (leading to possibly reduced power). There are two main design options for this experiment;

- One animal per O_2 treatment, 8 concentrations, 2 breathing types. With 8 replicates the experiment would require 128 animals, but that this could be analysed as a completely randomized design
- One O_2 profile per animal, so that each animal would be used 8 times and only 16 animals are required (8 lung and 8 buccal breathers)

Mullens decided to use the second option so as to reduce the number of animals required (on financial and ethical grounds). By selecting this option, she did not have a set of independent measurements for each oxygen concentration, by repeated measurements on each animal across the 8 oxygen concentrations.

Format of mullens.csv data file					

BREATH	TOAD	O2LEVEL	FREQBUC	SFREQBUC
lung	a	0	10.6	3.256
lung	a	5	18.8	4.336
lung	a	10	17.4	4.171
lung	a	15	16.6	4.074
...	...	...	...	...

BREATH	Categorical listing of the breathing type treatment (buccal = buccal breathing toads, lung = lung breathing toads). This is the between subjects (plots) effect and applies to the whole toads (since a single toad can only be one breathing type - either lung or buccal). Equivalent to Factor A (between plots effect) in a split-plot design
TOAD	These are the subjects (equivalent to the plots in a split-plot design: Factor B). The letters in this variable represent the labels given to each individual toad.
O2LEVEL	0 through to 50 represent the the different oxygen concentrations (0% to 50%). The different oxygen concentrations are equivalent to the within plot effects in a split-plot (Factor C).
FREQBUC	The frequency of buccal breathing - the response variable
SFREQBUC	Square root transformed frequency of buccal breathing - the response variable

Open the mullens data file. HINT. Notice that both the O2LEVEL variable contains only numbers. Make sure that you define both of this as a factors (HINT)

Q4-1. What are the main hypotheses being tested?

a. H_0 Main Effect 1 (Factor A):

b. H_0 Main Effect 2 (Factor C):

c. H_0 Main Effect 3 (A*C):

Q4-2. We will now address all the assumptions Construct a boxplot to investigate the assumptions as they apply to the test of H_0 Main Effect 1 (Factor A): This is done in two steps

a. Construct a boxplot to investigate the assumptions as they apply to the test of H_0 Main Effect 1 (Factor A): Start by **aggregating the data set by TOAD** (since the toads are the replicates for the between subject effect - BREATH) (HINT). Then **Construct a boxplot of aggregated mean FREQBUC against BREATH treatment**(HINT). Any evidence of violations of the assumptions (y or n)? _____ ☐
Try a square-root transformation!

b. Construct a boxplot to investigate the assumptions as they apply to the test of H_0 Main Effect 2 (Factor C): Since Factor C is tested against the overal residual in this case, this is a relatively straight forward procedure.(HINT). Any evidence of violations of the assumptions (y or n)? _____ ☐

c. Construct a boxplot to investigate the assumptions as they apply to the test of H_0 the main interaction effect (A:C): Since A:C is tested against the overal residual, this is a relatively straight forward procedure.(HINT). Any evidence of violations of the assumptions (y or n)? _____ ☐

d. In addition to the above assumptions, the test of TOAD assumes that there is no TOAD by O2LEVEL interaction as this is the overal residual (the replicates). That is, the test assumes that the effect of O2LEVEL is consistent in all TOADS. Construct an interaction plot to examine whether there is any evidence of an interation between TOAD and O2LEVEL (HINT). Any evidence of an interaction (y or n)? _____ ☐

e. Finally, you must also check to see whether the proposed model **balanced. Is it?** (Yor N) _____ ☐

Q4-3. Assume that the assumption of compound symmetry/sphericity will be violated and perform a **split-plot ANOVA (repeated measures)** (HINT), and complete the following table with corrected p-values (HINT).

Source of variation	df	Mean Sq	F-ratio	P-value	GG.P	HF.P
BREATH	☐	☐	☐	☐		
TOAD	☐	☐				
O2LEVEL	☐	☐	☐	☐	☐	☐
BREATH:O2LEVEL	☐	☐	☐	☐	☐	☐
Residuals	☐	☐				

Q4-4. Construct an **interaction plot** showing the frequency of buccal breathing against oxygen level, with each breathting type as different lines (or different bars). HINT

Q4-5. What conclusions would you draw from the analysis (and graph)?

Welcome to the end of Worksheet 7!