Module CH923
Statistics for Data Analysis
2009/10

Assignment 2 – Experimental Design

This coursework is concerned with experimental design, and you will be assessed by a group oral presentation on Thursday 5th November between 9.30 and 12.30 – note that you will need to be present throughout. Each of the six scenarios detailed below will be worked on in detail by a team of 4 or 5 students, but all students are expected to have considered all six scenarios, and therefore will be expected to contribute questions about the presentations made by the other teams.

Each team will be expected to give a presentation (20 minutes) discussing the design issues associated with their scenario, identifying some possible solutions, and briefly describing any implications of these solutions on the analysis. The presentation should be aimed at the experimenter (consider yourself to be acting as a statistical consultancy team!), who is someone doing a PhD in a biological science.

All PowerPoint/Impress files must be submitted electronically to the MOAC administrator (moac2) in advance of the assessment, and must be received by 1pm on Wednesday 4th November.

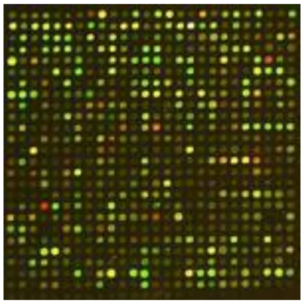
All members of the group should contribute to the presentation. Marks will be awarded based on understanding of the statistical design issues, identification of potential design solutions, identification of appropriate analytical approaches, and clarity of presentation (both visual and oral) –but please concentrate on the experimental design issues in your presentation. Marks will also be available for good questions asked about the presentations made by other groups.

Allocation of scenarios to students

Scenario 1: Damon Daniels (SB), Ingrid Tigges (SB), Jack Heal (MOAC), Mathew Wright (MOAC), Neil Pearson (DTG)
Scenario 2: Kashi Gorton (SB), Benjamin Wareham (SB), Ciara McCarthy (MOAC), Helen Armes (DTG), Georgia Roberton-Paterson (DTG)
Scenario 3: Jo Hulsmans (SB), Robert Deller (MOAC), Anthony Nash (MOAC), Charles Cameron (DTG)
Scenario 4: Philip Law (SB), Charlotte George (MOAC), Sang Young Noh (MOAC), Nathan Field (DTG)
Scenario 5: Seyed Rasooli-Hejad (SB), Vicent Hall (MOAC), Daniel Pearce (MOAC), Sarah Harvey (DTG)
Scenario 6: Ireve Sui (SB), Edward Harry (MOAC), Andrew Soulby (MOAC), Rachel Lipscombe (DTG)

Scenarios

1) A research student is planning a two-colour microarray experiment to identify genes associated with pathogen response in *Arabidopsis thaliana*. The plants will be grown in an enclosed controlled environment cabinet, with a fixed pattern of days and nights (16 hour day, 8 hour night).



To assess the impact of the pathogen on gene expression, he will have two sets of 4 plants (though there is both plenty of seed available, and plenty of space in the cabinet), one set inoculated with a single isolate of the pathogen, and the other uninoculated. The pathogen inoculation will take place simultaneously across all four plants, once all the plants have at least 10 leaves, and at the start of a 16 hour day. Immediately following inoculation, a random leaf will be excised from each of the 8 plants, and a further random leaf will be excised from each plant every 8 hours for the following 48 hours (this is the period over which it is expected to see most impact of the pathogen inoculation). The samples will be stored in liquid nitrogen prior to being prepared for microarray analysis (a standard technique for plant samples). Eight leaves will therefore be sampled on each of the seven occasions, each excised leaf being prepared separately and producing a separate mRNA sample to be compared in the microarray experiment. The 56 samples will be denoted by the treatment (Uninoculated (U), Pathogen (P)) and the sample time (0, 8, 16, 24, 32, 40 and 48 hours (0, 1, …6)), with the four samples for each combination of treatment and time being discriminated as sample a, b, c or d.

He is intending to divide the microarray experiment into two parts, the first to consider the time course responses for each treatment (Uninoculated, Pathogen), and the second to compare the Uninoculated and Pathogen treatments. For the first part, he will label the eight "0 hour" samples using Cy3 and all the other samples using Cy5. He will then run 48 microarrays, split into 8 groups of 6, with each group including comparisons of a "0 hours" sample with samples collected at the other 6 time points, as indicated below:

Microarray	Cy3	Cy5
1	U0a	U1a
2	U0a	U2a
3	U0a	U3a
4	U0a	U4a
5	U0a	U5a
6	U0a	U6a
7	U0b	U1b
…		

13 …	U0c	U1c
25 …	P0a	P1a
31 …	P0b	P1b

For the second part, he will label all the Uninoculated samples with Cy3 and all the Pathogen samples with Cy5. He will then run a further 28 microarrays comparing Uninoculated and Pathogen samples as indicated below:

Microarray	Cy3	Cy5
49	U0a	P0a
50	U1a	P1a
51	U2a	P2a
52	U3a	P3a
53	U4a	P4a
54	U5a	P5a
55	U6a	P6a
56 …	U0b	P0b
63 …	U0c	P0c
70 …	U0d	P0d

Your presentation should focus on the controlled environment experiment to collect the plant samples, though you should also consider how the design of this part of the experiment interacts with the design of the microarray phase.

2) A field trial is concerned with assessing the efficacy of five post-emergence herbicides in controlling weeds in carrots. The experimenter has been allocated an area of land large enough to accommodate 30 experimental plots, each of which will be three beds wide (a bed is 1.8m wide, the distance between tractor wheelings) by 10m long, with a 2m guard area between plots along beds. Each bed contains three rows of carrots (as seen in the photo). The plot dimensions have been set to allow for the width of the spray boom, and to allow sufficient length of plot for the tractor to be operating at the correct speed for the central 8m of each plot length. The experimenter expects there to be some spatial variability in the density of weeds within the area, and is therefore proposing to arrange six replicates of five treatments following an extended Latin Square design as shown below.

A	C	D	B	E	D
B	D	E	A	C	B
C	A	B	E	D	A
D	E	A	C	B	C
E	B	C	D	A	E

The herbicides will all be applied when the carrots have reached the 4 true-leaf stage (about 2 weeks after sowing), and the experimenter has just assessed the weed density on the middle bed of each plot a few days before the expected application date. The observed weed densities (recorded as numbers per square meter) have been classified into four grades (Low, Average, High, Very High) as shown below:

L	A	VH	VH	H	H
A	A	L	L	H	A
H	H	H	L	A	A
VH	H	H	A	A	L
VH	VH	H	A	L	L

As well as assessing the efficacy of each herbicide by assessing the weed density at different time points after herbicide application, the experimenter is also interested in assessing any phytotoxic (i.e. detrimental) effects that each herbicide might have on the growth of the carrots by measuring the harvest yield at maturity

3) This experiment is concerned with the survival and germination of sclerotia, the resting bodies that cause onion white rot, a major soil borne disease of onions. A research student believes that exposing these sclerotia to periods at different temperatures will affect both the conditioning of the sclerotia, altering the germination response when placed in optimal conditions near to onion roots, and the premature mortality of the sclerotia due to germination in the absence of onion roots.

The student intends to study the effect of 4 constant temperature treatments (10°C, 20°C, 30°C or 40°C) each over 4 different durations (1, 2, 3 or 4 weeks), and has 4 incubators available. She has devised a plan in which each temperature and duration will appear once in each of the incubators, so that the experiment can be completed in 10 weeks:

Incubator	A	B	C	D
1st treatment	10°C, 1 wk	20°C, 2 wks	30°C, 3 wks	40°C, 4 wks
2nd treatment	30°C, 2 wks	40°C, 1 wk	10°C, 4 wks	20°C, 3 wks
3rd treatment	40°C, 3 wks	30°C, 4 wks	20°C, 1 wk	10°C, 2 wks
4th treatment	20°C, 4 wks	10°C, 3 wks	40°C, 2 wks	30°C, 1 wk

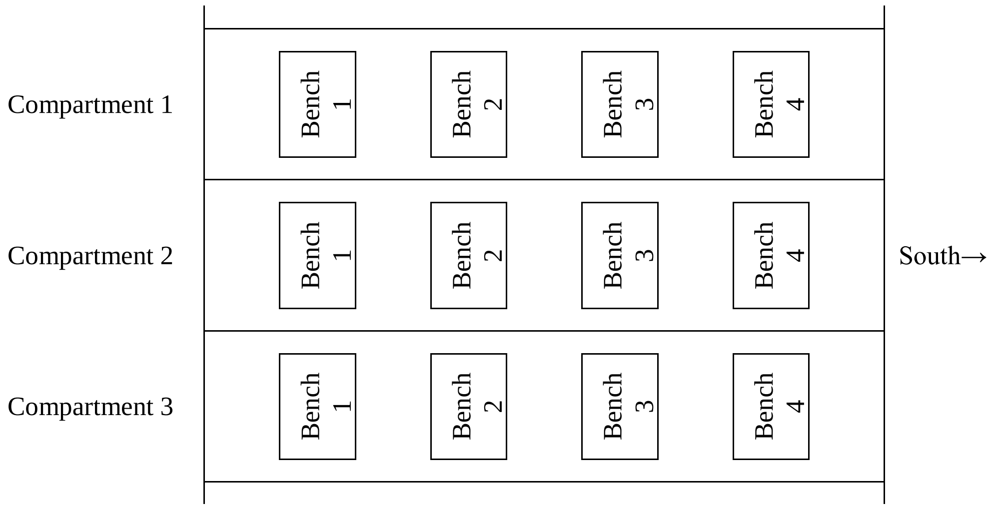
Each incubator contains 3 shelves, with the possibility of some variation in temperature between shelves. Each shelf can accommodate 4 pots of soil, so that each incubator can accommodate 12 pots in total at any one time. For each combination of temperature and duration, the student is intending to place 6 batches of 50 sclerotia in separate pots of soil, 2 pots per incubator shelf.

The sclerotia for use in this experiment will be produced on infected onions, and although each onion will produce a large number of sclerotia, it is anticipated that more than one production run will be required to produce sufficient sclerotia for the complete experiment. The student is intending to combine sclerotia from two or three runs, as required, to obtain a sufficient quantity for the experiment.

On removal from the incubator, each batch of 50 sclerotia will be assessed for survival (and the proportion surviving recorded). The surviving sclerotia will then be placed in optimal germination conditions, and the time to germination recorded for each sclerotia, based on daily assessments.

4) A researcher is studying the biocontrol of western flower thrips on poinsettias using two different, non-flying, predator species, *Orius laevigatus* (X) and *Neoseilus cucumeris* (Y). He is planning a glasshouse experiment where he will assess thrip presence and damage on poinsettias after introducing either predator species X, or predator species Y, or both. It is anticipated that the level of biocontrol provided by both species together will be better than the sum of the effects of the two species acting alone, due to their different modes of action.

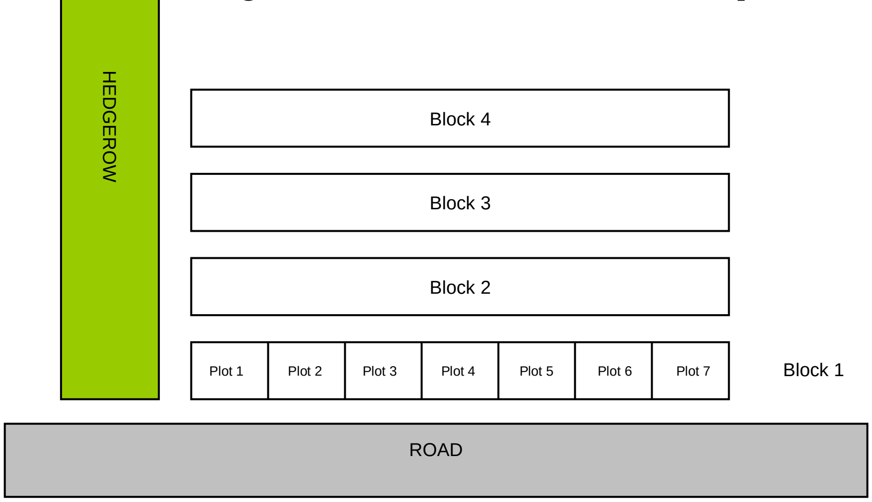
With three glasshouse compartments available, he is planning to use each compartment for a different predator treatment (predator species X only, predator species Y only, both predator species together), with four separate plots of poinsettias within each compartment to provide replication. Plots will be on separate benches, with netting around each plot to stop movement of either thrips or predators between plots. Each bench is oriented in an East-West direction, parallel to the South-facing outer-wall of the glasshouse compartment (see below). The researcher has sufficient predators to perform the experiment on two separate occasions, and access to the glasshouse compartments for two separate experimental runs, and so intends to have two replicates in time.

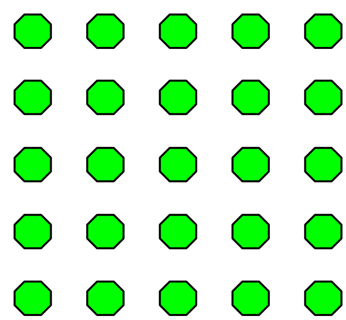


Each plot will contain 30 plants in a 10-by-3 array, but the researcher does not have sufficient time to assess all plants, and so is intending to randomly select 8 plants to be assessed in each plot. Numbers of thrips on each of these plants will be counted, together with an assessment of the damage on each of the leaves of each of these plants.

5) This trial is concerned with the damage to flowers caused by a flying insect pest. The research student has identified 7 treatments, comprising an untreated (unsprayed) control, 3 rates (Low, Medium, High) of a chemical which it is believed will act as a repellent to the pest, and 3 rates (Low, Medium, High) of a pesticide which it is believed will kill the pest.

She is intending to arrange the trial as a randomised complete block design with the blocks arranged as shown below. The hedgerow contains species thought to act as an alternative host for the pest.





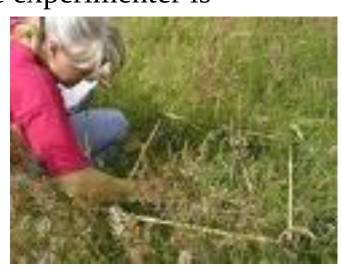
Plot Detail

The plant species

chosen for the trial produces multiple (between 10 and 20) sprays of flowers from each plant, and it is proposed that each plot will include 25 plants, arranged in a 5-by-5 array as shown above. The gap between adjacent plots will be twice the spacing between plants within a plot. Each spray has between 40 and 60 individual flowers, and so it takes a long time to accurately count the number of damaged flowers. Consequently in order to assess the experiment in a single day, only 10 sprays per plot can be assessed in this way. The experimenter intends to select a random spray from 10 random plants per plot, and count the number of damaged flowers in each spray.

6) A field trial is concerned with the impact of three different mechanical weeders (flexible tine weeder, finger weeder, steerage hoe) on the control of weeds in 7 different brassica crops (Cauliflower, Cabbage, Calabrese, Brussels sprouts, Kale, Borecole, Kohl Rabi). The experimenter also proposes to include an irrigation treatment (some plots irrigated, some not), supplied by spraying from pipes above each plot, to see how this interacts with the weeding control provided. There are a total of 42 treatments (7 crops x 3 weeders x 2 irrigation levels)

The experimenter is proposing to arrange the trial as a randomised complete block design, with each plot occupying a 10m length of bed (between tractor wheelings – approximately 1.8m wide), with a 2m guard area between plots along a bed. It is estimated that it will take 30 seconds to weed each plot, and 5 minutes to change over the mechanical weeder between plots. As it will therefore take just under four hours to mechanically weed a single replicate block (42 plots x 5m 30s = 3h 51m), the experimenter is proposing to only include 2 replicates so that the trial can be weeded in a single day, though there is sufficient land and resources available for more.



Within a plot, plants are at a 50cm spacing in both directions (between row, within row), with 3 rows of plants per bed. The experimenter wants to assess weed cover at various points during crop growth, within 4 randomly thrown 1m-by-1m quadrats per plot (these quadrats are rigid, square metal frames, which are thrown into a plot to designate a random area of ground to be assessed). He also wants to assess the impact of the different treatments on brassica yield at harvest.

	Plot 1	Plot 2	Plot 3	Plot 4	Plot 5	Plot 6	Plot 7
Rep 1	Plot 8	Plot 9	…				
	…						
					…	Plot 41	Plot 42

	Plot 43	Plot 44	Plot 45	Plot 46	Plot 47	Plot 48	Plot 49
Rep 2	Plot 50	Plot 51	…				
	…						
					…	Plot 83	Plot 84

Plot Detail: