C2M4_peer_reviewed

February 18, 2022

1 C2M4: Peer Reviewed Assignment

1.0.1 Outline:

The objectives for this assignment:

1. Get a better understanding of Experimental design patterns.
2. Prove some of the background intuition in blocking and interblock interactions.
3. Understand how and when to apply different model strucutres for different experimental designs.

General tips:

1. Read the questions carefully to understand what is being asked.
2. This work will be reviewed by another human, so make sure that you are clear and concise in what your explanations and answers.

2 Problem 1: Experimental Design

This problem is to get you thinking about how experiments are designed and how data is collected, because those influence what models we end up using.

2.0.1 1. (a)

In your own words, define experimental design. Describe some negative effects of making an incorrect experimental design decision.

2.0.2 Answer:

Experimental design is a research process in which one manipulates an enviroment, using a treatment variable in a controlled design, that results in maximum accuracy. If done correctly, a hypothesis statement can be drawn and causality may be defined in a relationship. An incorrect design can result in invalid conclusions; this can cause practical and ethical issues.

2.0.3 1. (b)

In your own words, describe the difference between an experimental unit and a treatment unit. Why does this distinction matter?

2.0.4 Answer:

Experimental units are the objects that have the possibility of being given the treatement. Treatment units are the subgroup of the experimental units that are chosen to have the treatment applied to them.

3 Problem 2: Proving the Intuition

Show that, for the randomized complete block design:

$$SS_{total} = SS_{treat} + SS_{block} + SS_R$$

$$TSS = \sum_{i=1}^{t}\sum_{j=1}^{r}(Y_{i,j} - \bar{Y})^2$$

$$= \sum_{i=1}^{t}\sum_{j=1}^{r}(Y_{i,j} + \bar{Y}_{i.} - \bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{.j} - \bar{Y}_{..} + \bar{Y}_{..} - \bar{Y})^2$$

$$= \sum_{i=1}^{t}\sum_{j=1}^{r}[(Y_{i,j} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}) + (\bar{Y}_{.j} - \bar{Y}_{..}) + (\bar{Y}_{i.} - \bar{Y}_{..})]^2$$

$$= \sum_{i=1}^{t}\sum_{j=1}^{r}(Y_{i,j} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 + \sum_{j=1}^{r} t(\bar{Y}_{.j} - \bar{Y}_{..})^2 + \sum_{i=1}^{t} r(\bar{Y}_{i.} - \bar{Y}_{..})^2$$

$$= SS_{treat} + SS_{block} + SS_R$$

4 Problem 3: Interblock Interactions

Describe why, in a randomized complete block design (RCBD), it is not possible to test whether interactions exist between the treatment and blocks.

Because each treatment occurs only one ine each block. We would use the data single data point to estimate our degree of freedoms for this factor/block interaction. Thus, each data point would be used as a parameter: this would result of our model having 0 degrees of freedom. Thus, we cannot calculate standard errors or variance (SSE / df). Therefore, we cannot carry out any hypothesis tests.

5 Problem 4: 99 Designs for 99 Problems

For each of the following design patterns, give an example (that wasn't given in the videos) for an experiment that would best lend itself to the specified design pattern. Make sure to explain why the specified design is more applicable for your experiment than the other design patterns.

1. Complete Randomized Design (CRD)
2. Complete Randomized Block Design (CRBD)
3. Factorial Design

5.0.1 Answer:

1. A doctor may want to see if a pill will decrease headaches for all of his adult female patients. Therefore, he tests different variety of pills, including a placebo. 50 if his patients agree to be a part of the study. He randomly assigns each person to a pill and evaluates their progress over the next few weeks. A completely Randomized Design is best because we are applying the experiment to adult females; we do not need to control for outside factors such as gender or age group. Therefore, no blocking is needed.
2. A football coach wants to develop a program that relies on speed. Therefore, he decides to conduct an experiment on his players. Currently, his players only develop speed by increasing their squat strength. To find the best method, the coach tests the following exercises along with continuing squats: power cleans, mobility/stretching, form running, and plyometrics. The control group will continue squating, but will do none of the other listed exercises. Also, the coach knows different position groups develop differently; therefore, he divides his players into the following groups: Lineman (Offensive and Defensive), Big Skill (Quarterback, Tight Ends, Linebackers, Specialists), and Small Skill (Wide Receivers and Defensive Backs). Each experimental unit within each group will be randomly assigned to an exercise. The Coach wil evaluate the results once their offseason is over. Complete Randomized Block Design is best because we want to evaluate the effects of the different exercises; however, different groups develop speed differently. Other factors, at first thought, won't have a large impact. Thus, we do not need more than 1 block.,.
3. An experiment where we evaluate a high school student's math test score by evaluating their gender, grade level, whether they participate in extracurricular activities, and socioeconomic status (Poor, Middle Class, Wealthy). Factorial design is best in this situation because all of the factos can make a large difference in a person's math score. Thus, we need to account for all the factors. Counting for less than these 4 factors would be inappropriate.

[ ]: