Handout 9: Single Factor Experiments with a Blocking Variable

Text: Read section 4.1

**Example**:A study was conducted to compare four different tips used to measure hardness on a testing machine. The metal sheets used to compare tips vary in density due to uncontrollable variables. It is thought that one particular sheet is fairly consistent, but consistency between sheets is a known concern. These metal sheets are referred to as coupons.

Suppose the experimenter has four different coupons available to test the tip.



Furthermore, suppose there was no restriction on treatment randomization. If this were the case, the following randomization of a completely randomized design is possible.



Questions:

1. Suppose that Coupon 1 is atypically soft. How might this affect the outcomes on that coupon?
2. Suppose that Coupon 2 is atypically hard. How might this affect the outcomes on that coupon?

Up to this point, we have considered only one type of experimental design, the completely randomized design. This design is appropriate when the experimental units are homogeneous. If this is not the case, then the experimental units should be grouped into blocks of homogeneous units to reduce the experimental error variance. This type of design is known as a **randomized complete block design (RCBD)**.

Key Elements of Randomized Complete Block Designs

* Block: This is a homogeneous group of experimental units. A RCBD consists of putting the experimental units into blocks.
* Complete: Each block consists of one complete replication of the set of treatments. Therefore, each treatment will show up once within each block.
* Randomized:The treatments are randomly assigned to the experimental units separately WITHIN each block.

When blocking the experimental units, keep the following objectives in mind:

1. Within blocks, make the experimental units as HOMOGENEOUS as possible with respect to the response variable.
2. Make the different blocks as HETEROGENEOUS as possible with respect to the response variable.

Comments

* When blocking is effective, there is a reduction in the unexplained error of the experiment.
* A reduced unexplained error will lead to more precision in the estimates of interest, which means more power to identify statistically significant effects, i.e. the factor that truly matter.
* When the homogeneous groups of experimental units, i.e. blocks, are selected at *random* from a larger possible set of units, then the blocking factor is said to be a **random effect**. For example, if coupon is known to effect the response, then coupon should be used as a blocking variable. Furthermore, if the goal is to extend the scope-of-inference beyond these four coupons, the coupon would be considered a *random* blocking variable. On the other hand, if the characteristics of these four coupons match exactly the characteristics of all future coupons to be used in this process, then coupon would be considered a fixed effect.

**Back to the Example**: Recall that each coupon in our study is fairly consistent, but consistency between coupons is not likely. Therefore, we will set up the experiment so that each coupon represents a block, and each tip will be tested on each block. One possible randomization scheme is given as follows.



***Comment****: Blocking is used to reduce the effect of one or more uncontrollable variables. If you can control the variable or you have a genuine interest in the effect of a variable on the response, then it is* ***not*** *a blocking variable.*

The data for this experiment are shown below (and can be found in the file **Harness.xlsx**).



Graphical Displays of the Data

The best graphical display in Minitab for these can be constructed by selecting **Graph > Dotplots,** using the **With Groups** type.



Place the response variable in the Graph Variable box, and then Coupon (blocking variable) followed by Type (the treatment variable) in the Categorical variables box.



Minitab returns the following set of stacked dotplots (with groups):



This plot gives us an indication of the effect of coupon on the response. From looking at this plot, it appears that Coupon 4 was the hardest because it consistently gave the highest measurements for all four Tip types. On the other hand, it appears that Coupon 1 or 2 was the softest.

|  |  |
| --- | --- |
| Coupon 1 | Coupon 2 |
| Coupon 3 | Coupon 4 |

The **best** graphical display is to literally adjust the responses according to the observed effect of each coupon. This is adjustment is called block centering.

|  |  |
| --- | --- |
|  |  |

Model Specification

The following general model is expressed in terms of treatment effects:

,

for i = 1, 2, … a; j = 1, 2,… b

where μ is the overall mean;

 τi is the effect of the ith treatment;

 βj is the effect of the jth block (this represents the average deviation of the units

in block j from the overall mean);

 eij is the random error term.

Model Assumptions:

1. The error terms are independently and normally distributed.
2. The error variance is constant.
3. The treatment and block effects are assumed to be additive (there is no interaction between treatments and blocks).

Testing Hypotheses

1. The usual F statistic is used to test the null hypothesis of no difference among the treatment means.
2. Little interest exists in formal inference about block effects, so we typically **ignore the F-statistic given for blocks** in the ANOVA table.
3. If a significant difference in treatments exists, you should proceed in the usual manner by figuring out where the differences exist.

Analysis in Minitab

Select **Stat > ANOVA > General Linear Model**, and enter the following.



**Specification of Fixed vs. Random Effect**



**The Overall F-test**



What are our conclusions?

Why ignore the Coupon effect?

How to measure the effectiveness of blocking variable?

|  |  |
| --- | --- |
| Total Variance: | Mean-Square Error in model without the blocking variable |
| Error Variance: | Mean-Square Error in the model with the blocking variable |
| Coupon Variance:  | The amount of variance that is taken out of “unexplained error” by using coupon as the blocking variable |
|  |

Comparing Tip Types (i.e. Multiple Comparisons)

Select **Stat > ANOVA > General Linear Model > Comparisons…** and enter the following.



Letterings…



95% Confidence Intervals for Pairwise Differences





Coefficients, Predicted Values, and Residuals





Factor Plots (Means on plot being averaged over blocks)





Checking Model Assumptions with Residual Plots



