Exercise Sheet 3

1. In the context of experimental design, what is:
   * a factor?

A factor of an experiment is a controlled explanatory variable; a variable whose levels are set by the experimenter.

* + a level?

A level represents different values of the controlled explanatory variable.

* + a treatment?

Different treatments constitute different levels of a factor, if an experiment has more than one factor then each combination of factors is a treatment.

1. A mechanical engineer is studying the thrust force developed by a drill press. He suspects that the drilling speed and the feed rate of the material affect the thrust force. He selects four feed rates and uses a high and low drill speed chosen to represent the extreme operating conditions.
   * How many factors does this experiment have?

Two factors, drilling speed and feed rate.

* + How many levels does each of the factors have?

Drilling speed has two levels, high and low. Feed rate has four levels.

* + How many treatments does this experiment have?

8 treatments

1. Give an example of an experiment that uses a randomized block design.
2. A manufacturer wants to test 4 different formulae for a solution, however, each formulation is mixed from a batch of raw material that is only large enough for four formulations and there might be variation from one batch of raw material to another. The following randomized block design can be used:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Formula** | **Batch** | | | | |
|  | **1** | **2** | **3** | **4** | **5** |
| **1** |  |  |  |  |  |
| **2** |  |  |  |  |  |
| **3** |  |  |  |  |  |
| **4** |  |  |  |  |  |

Note that the number of batches used to test each formula can vary, in the table above, 5 batches are used.

1. Why was a randomized block design appropriate?

The randomized block design will control for the effect of Batch.

1. If we test five null hypotheses (which are all true) using 0.01 as the critical significance level what is the probability that we will make at least one type I error?

Let number of type 1 errors = X

P(X >= 1) = 1 - P(X = 0) =

(1-0.01) x (1-0.01) x (1-0.01) x (1-0.01) x (1-0.01) = 1 - 0.95099 = 0.04901

1. A statistician carries out an ANOVA on a data set. Name two assumptions that should be satisfied before carrying out the test.

Random sampling (observations are independent).

Check that variances are equal

1. The Kenton Food Company wished to test four different package designs for a new breakfast cereal. The different designs were sold in twenty shops, with approximately equal sales figures. Each shop was randomly assigned one of the package designs so that each design was sold in five different shops. The number of cases sold were represented by the variable ‘Cases’ and the type of packaging was represented by the variable ‘Package’ which had four different levels, identified as 1, 2, 3 and 4. After the preliminary data analysis, a one way ANOVA was carried out on the data set, followed by Tukey’s HSD post hoc test. The R output is shown below:

**Figure 1**

Fligner-Killeen test of homogeneity of variances

data: kenton$Cases by kenton$Package

Fligner-Killeen:med chi-squared = 1.1321, df = 3, p-value = 0.7693

**Figure 2**

Df Sum Sq Mean Sq F value Pr(>F)

Package 3 586.8 195.6 19.56 1.33e-05 \*\*\*

Residuals 16 160.0 10.0

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**Figure 3**

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 14.600 1.414 10.32 1.76e-08 \*\*\*

kenton$Package2 -1.200 2.000 -0.60 0.5569

kenton$Package3 4.600 2.000 2.30 0.0352 \*

kenton$Package4 12.600 2.000 6.30 1.06e-05 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.162 on 16 degrees of freedom

Multiple R-squared: 0.7858, Adjusted R-squared: 0.7456

F-statistic: 19.56 on 3 and 16 DF, p-value: 1.333e-05

**Figure 4**

Tukey multiple comparisons of means

95% family-wise confidence level

Fit: aov(formula = kenton$Cases ~ Package)

$Package

diff lwr upr p adj

2-1 -1.2 -6.92203965 4.52204 0.9305700

3-1 4.6 -1.12203965 10.32204 0.1395792

4-1 12.6 6.87796035 18.32204 0.0000569

3-2 5.8 0.07796035 11.52204 0.0463912

4-2 13.8 8.07796035 19.52204 0.0000194

4-3 8.0 2.27796035 13.72204 0.0051230

**Figure 5**



**Figure 6**

# A tibble: 4 x 3

Package mean\_Cases sd\_Cases

<fct> <dbl> <dbl>

1 14.6 2.30

2 13.4 3.65

3 19.2 2.39

4 27.2 3.96

* + What does the result of the Fligner-Killeen test (see Figure 1) tell you?

The variance of sales is not significantly different for different packaging designs, p = 0.7693.

* + What are the null and alternative hypotheses associated with this analysis?

H₀ : the mean number of cases sold for each packaging type are equal i.e. μ₁ = μ₂ = μ₃ = μ4

HA : the mean number of cases sold is different for at least one packaging type.

* + Why does the variable ‘Package’ have 3 degrees of freedom?

The variable package has 4 different levels. The degrees of freedom are given by a − 1 = 4 − 1 = 3

* + Report the results of the ANOVA including the pairwise comparisons, a table indicating which (if any) treatments are significantly different at the 5% level and the effect size η².

A one way analysis of variance indicated that the package design has a significant effect on the number of cases of cereal sold, F(3, 16) = 19.56, MSE = 10, p < 0.001, η2 = 0.786. As shown in Figure 6, Tukey’s HSD test indicated that the mean number of cases sold was significantly different between all designs except between designs 1 and 2 and 1 and 3.

* + Write down an ANOVA model for this data set, indicating clearly what each term in the model represents.

Using the output from the treatment effects we can write:

Where for the packaging type 1 treatment, , and .

For the packaging type 2 treatment , and .

For the packaging type 3 treatment , and .

For the packaging type 4 treatment , and .

1. A soft drink distributer knows that end-aisle displays are an effective way to increase sales of the product. However there are several ways to design these displays. The marketing group has designed three new end aisle displays and wants to test their effectiveness. They have identified 15 stores of similar size and type to participate in each study. Each store will test one of the displays for a period of one month. The displays are assigned at random around the stores with each display tested in 5 stores. The response variable is the percentage increase in sales activity over typical sales in that store when the display is not is use. The data from the experiment is available on Blackboard in the soft\_drink.txt file. Can you analyse the data to determine whether the end-aisle displays have a significant impact on the soft drink sales? Write up a brief report including the following outputs and any conclusions that you draw.
   * Exploratory data analysis
   * ANOVA table
   * Pairwise comparisons including a table summarizing which (if any) treatments are significantly different.
   * The effect size η²
   * An ANOVA model indicating clearly what each term in the model represents.
   * Model diagnostics

The summary sales statistics for each end-aisle design are shown in Table 1 below.

**Table 1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Design | Mean Increase (%) | SD Increase | Max Increase | Min Increase | No. Obs. (n) |
| 1 | 5.73 | 0.388 | 6.22 | 5.29 | 5 |
| 2 | 6.24 | 0.434 | 6.71 | 5.66 | 5 |
| 3 | 8.32 | 0.670 | 9.2 | 7.55 | 5 |

The greatest mean percentage increase in sales was for display 3, with a mean percentage increase of 8.32. The lowest percentage increase in sales was for design 1 with a mean percentage increase of 5.73.

The boxplots (see Fig. 1) and histograms (see Fig. 2) indicate that there are no potential outliers and there is no evidence of strong skew (Design 3 does shows some skew but we will check that the assumptions of normality are met for the error terms). Equality of variance was assessed using the Fligner – Killeen test which found variance of the percentage increase did not differ significantly across different displays, p-value = 0.3601 (2 d.f.).

**Figure 1**



**Figure 2**



A one way analysis of variance indicated that the end-aisle design has a significant effect on the percentage increase in sales, F(3, 16) = 19.56, MSE = 10, p < 0.001, η2 = 0.786 (see Table 2).

**Table 2**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | d.f. | Sum Sq | Mean Sq | F- Value | p-value |
| Display | 2 | 18.78 | 9.392 | 35.77 | < 0.001 |
| Residuals | 12 | 3.15 | 0.263 |  |  |

As shown in Figure 3 and Table 3, Tukey’s HSD test indicated that the mean percentage increase in sales was significantly different between designs 1 and 3 and 2 and 3.

**Figure 3**



**Table 3**

|  |  |  |  |
| --- | --- | --- | --- |
| **Design** | **Mean % Increase** | **Standard deviation** | **n** |
| 1 | 5.73A | 0.388 | 5 |
| 2 | 6.24A | 0.434 | 5 |
| 3 | 8.32 | 0.670 | 5 |

Note. Means sharing a letter in their superscript are not significantly different at the .05 level according to a Tukey’s HSD test.

Using the output from the treatment effects we can write:

Where for the design 1 treatment, , and .

For the design 2 treatment and .

For the design 3 treatment and .

Model diagnostic plots shown in Figure 4 indicate that the residual terms have a normal distribution and that homogeneity of variance across treatment groups is satisfied. No influential points were identified.

**Figure 4**



1. An experiment was run to determine whether four specific firing temperatures affect the density of a certain type of brick. A completely randomized experiment was run and the results are available on Blackboard in the brick.txt file. Write up a brief report including the following outputs and any conclusions that you draw.
   * Exploratory data analysis
   * ANOVA table
   * Pairwise comparisons including a table summarizing which (if any) treatments are significantly different.
   * The effect size η²
   * An ANOVA model indicating clearly what each term in the model represents.
   * Model diagnostics

See notes in R script.