C2M2_peer_reviewed

June 27, 2023

1 C2M2: Peer Reviewed Assignment

1.0.1 Outline:

The objectives for this assignment:

1. Utilize contrasts to see how different pairwise comparison tests can be conducted.
2. Understand power and why it's important to statistical conclusions.
3. Understand the different kinds of post-hoc tests and when they should be used.

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.

[6]:
```
library(ggplot2)
```

2 Problem 1: Contrasts and Coupons

Consider a hardness testing machine that presses a rod with a pointed tip into a metal specimen with a known force. By measuring the depth of the depression caused by the tip, the hardness of the specimen is determined.

Suppose we wish to determine whether or not four different tips produce different readings on a hardness testing machine. The experimenter has decided to obtain four observations on Rockwell C-scale hardness for each tip. There is only one factor - tip type - and a completely randomized single-factor design would consist of randomly assigning each one of the 4×4=16 runs to an experimental unit, that is, a metal coupon, and observing the hardness reading that results. Thus, 16 different metal test coupons would be required in this experiment, one for each run in the design.

[15]:
```
tip    = factor(rep(1:4, each = 4))
coupon = factor(rep(1:4, times = 4))
y = c(9.3, 9.4, 9.6, 10,
       9.4, 9.3, 9.8, 9.9,
       9.2, 9.4, 9.5, 9.7,
       9.7, 9.6, 10, 10.2)
hardness = data.frame(y, tip, coupon)
```

```
hardness
```

A data.frame: 16 × 3

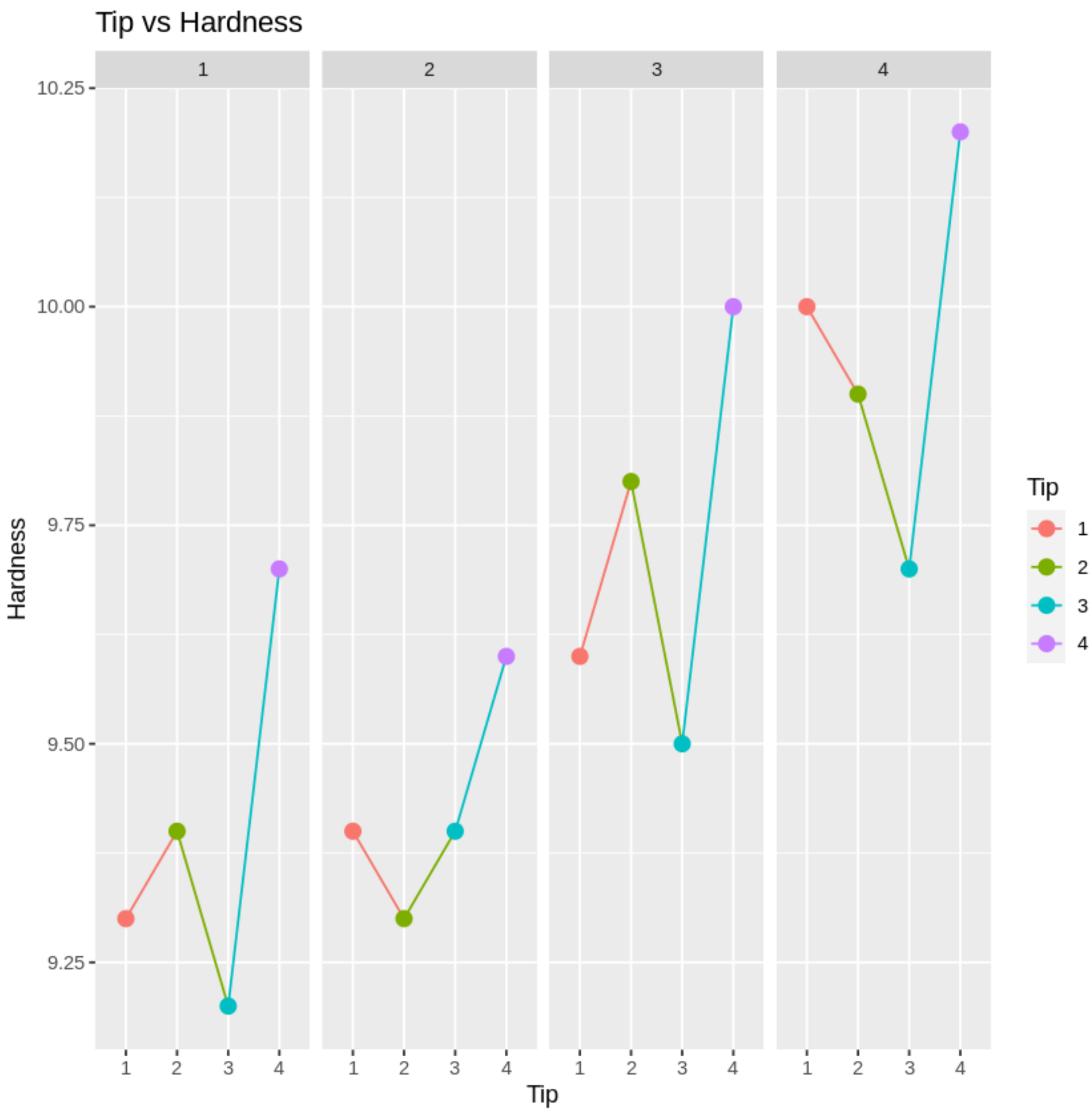
y	tip	coupon
<dbl>	<fct>	<fct>
9.3	1	1
9.4	1	2
9.6	1	3
10.0	1	4
9.4	2	1
9.3	2	2
9.8	2	3
9.9	2	4
9.2	3	1
9.4	3	2
9.5	3	3
9.7	3	4
9.7	4	1
9.6	4	2
10.0	4	3
10.2	4	4

2.0.1 1. (a) Visualize the Groups

Before we start throwing math at anything, let's visualize our data to get an idea of what to expect from the eventual results.

Construct interaction plots for `tip` and `coupon` using ggplot(). Be sure to explain what you can from the plots.

[9]:
```
# Your Code Here
ggplot(hardness, aes(x = tip, y = y, group = coupon, colour = tip)) +
  geom_line() +
  geom_point(size = 3) +
  labs(title = "Tip vs Hardness",
       x = "Tip",
       y = "Hardness",
       colour = "Tip") +
  facet_grid(. ~ coupon)
```



This visualization approach provides a separate line plot for each coupon level. This makes it easier to examine the relationship between tip and hardness reading within each coupon level.

2.0.2 1. (b) Interactions

Should we test for interactions between `tip` and `coupon`? Maybe there is an interaction between the different metals that goes beyond our current scientific understanding!

Fit a linear model to the data with predictors `tip` and `coupon`, and an interaction between the two. Display the summary and explain why (or why not) an interaction term makes sense for this data.

[14]:
```
# Your Code Here
lm_fit = lm(y ~ tip * coupon, data = hardness)
```

```r
summary(lm_fit)
```

```
Call:
lm(formula = y ~ tip * coupon, data = hardness)

Residuals:
ALL 16 residuals are 0: no residual degrees of freedom!

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   9.300e+00         NA      NA       NA
tip2          1.000e-01         NA      NA       NA
tip3         -1.000e-01         NA      NA       NA
tip4          4.000e-01         NA      NA       NA
coupon2       1.000e-01         NA      NA       NA
coupon3       3.000e-01         NA      NA       NA
coupon4       7.000e-01         NA      NA       NA
tip2:coupon2 -2.000e-01         NA      NA       NA
tip3:coupon2  1.000e-01         NA      NA       NA
tip4:coupon2 -2.000e-01         NA      NA       NA
tip2:coupon3  1.000e-01         NA      NA       NA
tip3:coupon3 -3.758e-15         NA      NA       NA
tip4:coupon3 -3.869e-15         NA      NA       NA
tip2:coupon4 -2.000e-01         NA      NA       NA
tip3:coupon4 -2.000e-01         NA      NA       NA
tip4:coupon4 -2.000e-01         NA      NA       NA

Residual standard error: NaN on 0 degrees of freedom
Multiple R-squared:      1,Adjusted R-squared:    NaN
F-statistic:   NaN on 15 and 0 DF,  p-value: NA
```

By fitting a linear model with the interaction term between tip and coupon, we're examining if the effect of one predictor on the response variable depends on the level of the other. The significance of the interaction term in the model summary, denoted by a small p-value, suggests the presence of such an interaction. If not significant, the interaction term may not be needed, indicating that the influence of tip on hardness doesn't vary with different coupon levels.

2.0.3 1. (c) Contrasts

Let's take a look at the use of contrasts. Recall that a contrast takes the form

$$\sum_{i=1}^{t} c_i \mu_i = 0,$$

where $\mathbf{c} = (c_1, ..., c_t)$ is a constant vector and $\mu = (\mu_1, ..., \mu_t)$ is a parameter vector (e.g., μ_1 is the mean of the i^{th} group).

We can note that $\mathbf{c} = (1, -1, 0, 0)$ corresponds to the null hypothesis $H_0 : \mu_2 - \mu_1 = 0$, where μ_1 is the mean associated with tip1 and μ_2 is the mean associated with tip2. The code below tests this hypothesis.

Repeat this test for the hypothesis $H_0 : \mu_4 - \mu_3 = 0$. Interpret the results. What are your conclusions?

[10]:
```
library(multcomp)
lmod = lm(y~tip+coupon, data=hardness)
fit.gh2 = glht(lmod, linfct = mcp(tip = c(1,-1,0,0)))

#estimate of mu_2 - mu_1
with(hardness, sum(y[tip == 2])/length(y[tip == 2]) -
     sum(y[tip == 1])/length(y[tip == 1]))
```

0.0250000000000021

Examining contrast $\mathbf{c} = (0, 0, -1, 1)$ tests the null hypothesis $H_0 : \mu_4 - \mu_3 = 0$, meaning it checks if there's a significant difference in the hardness means for tips 3 and 4. Rejecting the null hypothesis implies a significant mean difference, while failing to reject it suggests insufficient evidence for a difference. These tests can help us understand the significance of differences between the tips' hardness readings.

2.0.4 1. (d) All Pairwise Comparisons

What if we want to test all possible pairwise comparisons between treatments. This can be done by setting the treatment factor (`tip`) to "Tukey". Notice that the p-values are adjusted (because we are conducting multiple hypotheses!).

Perform all possible Tukey Pairwise tests. What are your conclusions?

[13]:
```
# Your Code Here
pairwise.t.test(hardness$y, hardness$tip, p.adjust.method = "BH")
```

```
Pairwise comparisons using t tests with pooled SD

data:  hardness$y and hardness$tip

  1    2    3
2 0.90 -    -
3 0.64 0.64 -
4 0.36 0.36 0.29

P value adjustment method: BH
```

We use pairwise t-tests for each pair of tip levels, followed by adjusting p-values with the Benjamini-Hochberg procedure. This approach helps control the false discovery rate, providing a balance between Type I and Type II errors. The results indicate which pairs of tip levels differ significantly in hardness readings.

3 Problem 2: Ethics in my Math Class!

In your own words, answer the following questions:

- What is power, in the statistical context?
- Why is power important?
- What are potential consequences of ignoring/not including power calculations in statistical analyses?

1. Power as used in this context relates to the rejection of a false null hypothesis. It has proved vital in detecting the probabilities true effects.
2. A type II errors is an error in which the null hypothesis is falsely accepted. Power is very important in statistics, in that it has the ability to detects these type II errors. It plays a key role in the accuracy of results.
3. As stated, it plays a crucial role in assuring accuracy of results, this is very important in data based statistical studies.

4 Problem 3: Post-Hoc Tests

There's so many different post-hoc tests! Let's try to understand them better. Answer the following questsions in the markdown cell:

- Why are there multiple post-hoc tests?
- When would we choose to use Tukey's Method over the Bonferroni correction, and vice versa?
- Do some outside research on other post-hoc tests. Explain what the method is and when it would be used.

1. There are a variety of different statistical reserch studies that are conducted, and different studies use different post-hoc tests. Each post-hoc tests included its own set of assumptions, which is why it is important that multiple exist in order to aid reserchers and statisticians
2. Unlike other methods, both of these are type I error methods. Turkeys method is a more rigerous and resource taxing approuch, whereas Bonferronis' method is generally used when only a small number of comparisons are being made.
3. Student-Newman-Keuls (SNK): This test is similar to the Turkeys method in that they both perform pairwise comparisons of group means, but unlike Turkeys HSD which uses range distribution, SNK uses range statistic.