C2M2_peer_reviewed

April 30, 2024

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.

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.

[1]:
```
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

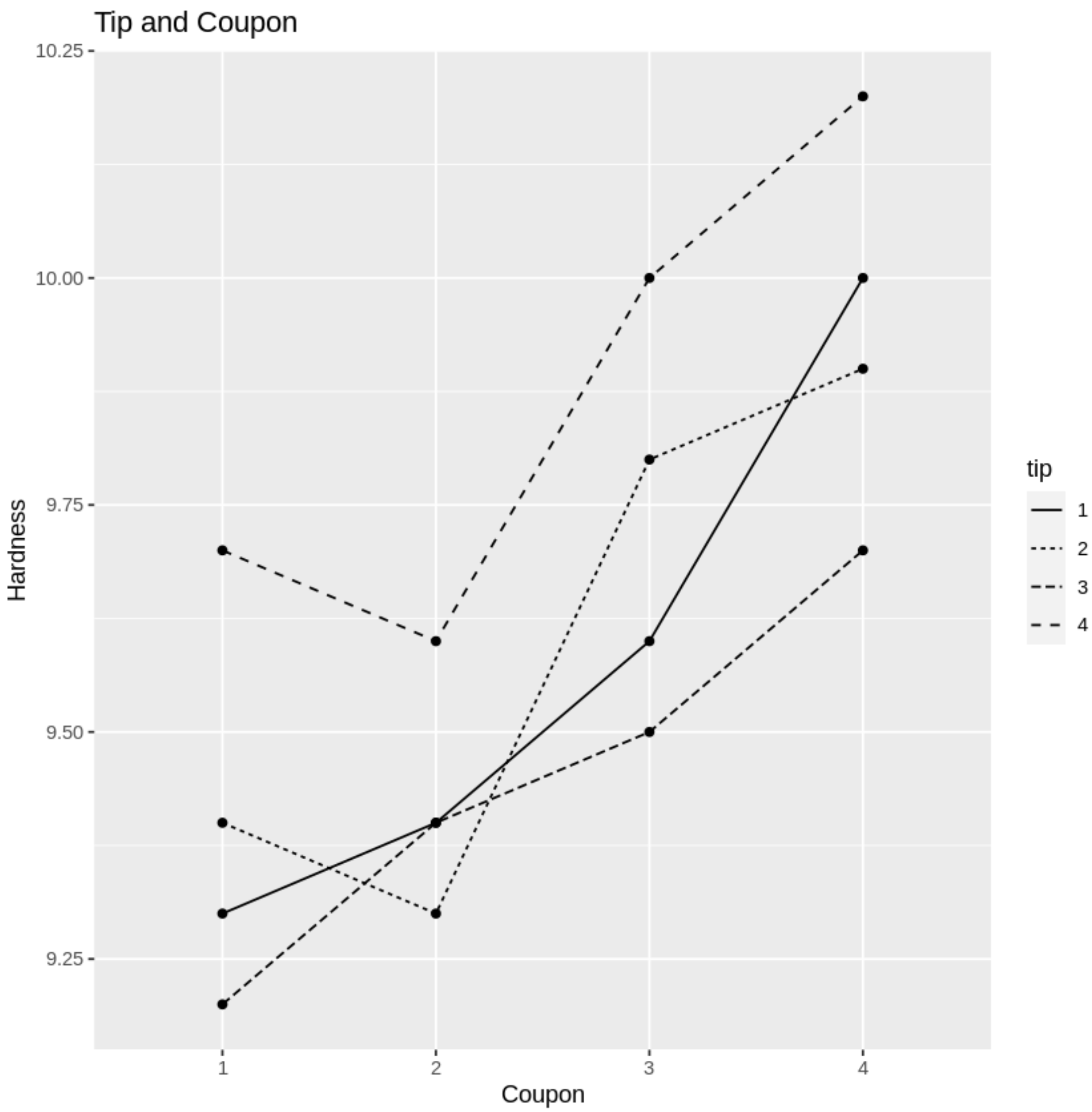
2.0.1 1. (a) Visualize the Groups

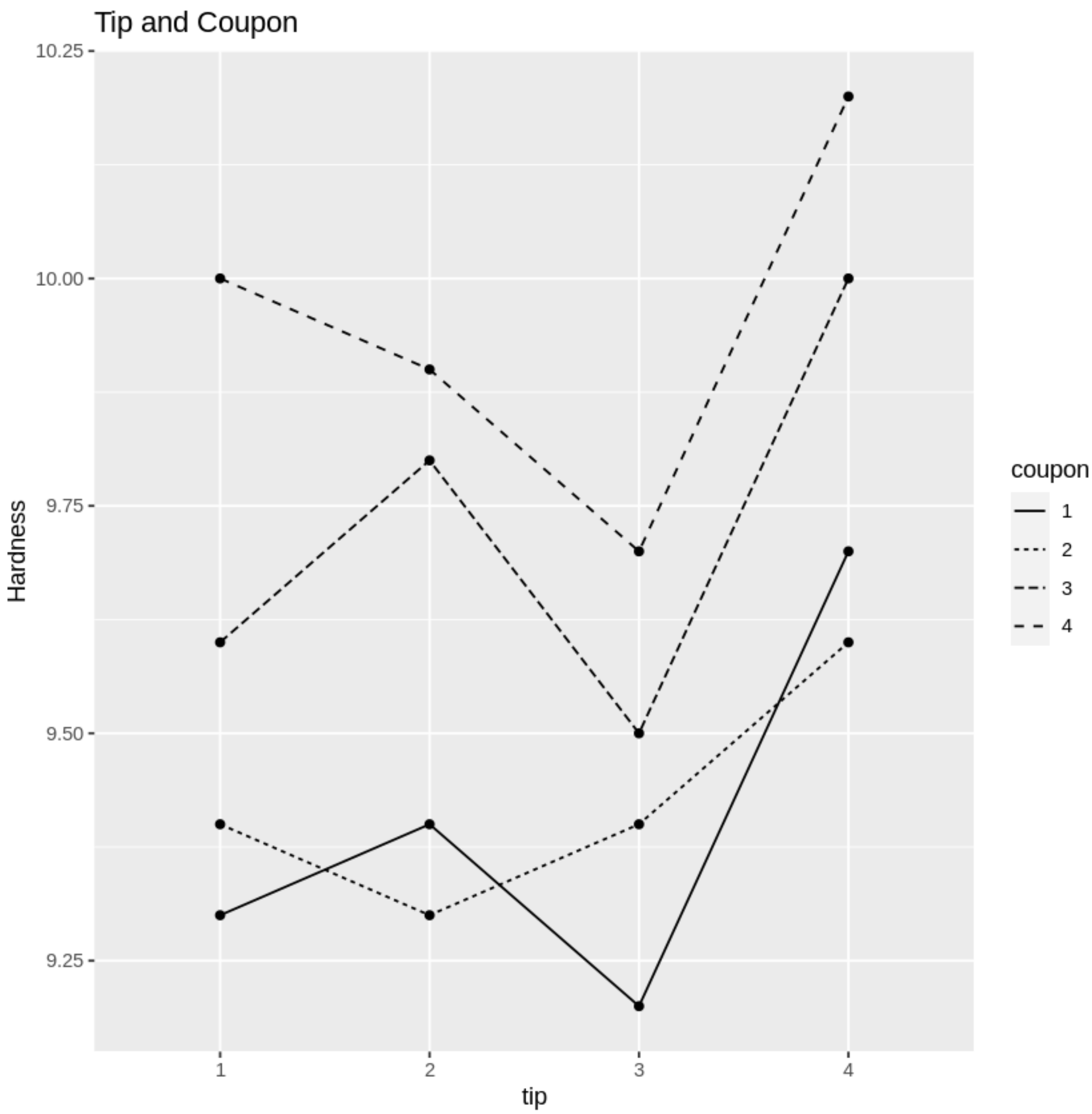
Before we start throwing math at anything, let's visualize our data to get an idea of what to expect from the eventual results.

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

[8]:
```
library(ggplot2)
# Your Code Here
ggplot(hardness, aes(x = coupon, y = y,group = tip, linetype =
tip)) +
  geom_line() +
  geom_point() +
  labs(x = "Coupon", y = "Hardness", color = "Tip") +
  ggtitle("Tip and Coupon")

ggplot(hardness, aes(x = tip, y = y, group = coupon, linetype =
coupon)) +
  geom_line() +
  geom_point() +
  labs(x = "tip", y = "Hardness", color = "Tip") +
  ggtitle("Tip and Coupon")
```


Tip and Coupon
10.25
10.00
9.75
9.50
9.25
Hardness
1
2
3
4
Coupon
tip
1
2
3
4



the plots do not show a clear evidence of interaction

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.

[9]:
```
lmod_int = lm(y ~ coupon + tip + tip:coupon, data = hardness)
summary(lmod_int)
```

```
Call:
lm(formula = y ~ coupon + tip + 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
coupon2       1.000e-01         NA      NA       NA
coupon3       3.000e-01         NA      NA       NA
coupon4       7.000e-01         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:tip2 -2.000e-01         NA      NA       NA
coupon3:tip2  1.000e-01         NA      NA       NA
coupon4:tip2 -2.000e-01         NA      NA       NA
coupon2:tip3  1.000e-01         NA      NA       NA
coupon3:tip3 -3.689e-15         NA      NA       NA
coupon4:tip3 -2.000e-01         NA      NA       NA
coupon2:tip4 -2.000e-01         NA      NA       NA
coupon3:tip4 -3.784e-15         NA      NA       NA
coupon4:tip4 -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
```

it does not make sense to include an interaction term

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]))

fit.gh2 = glht(lmod, linfct = mcp(tip = c(0,0,-1,1)))
summary(fit.gh2)
#estimate of mu_4 - mu_3
with(hardness, sum(y[tip == 4])/length(y[tip == 4]) -
     sum(y[tip == 3])/length(y[tip == 3]))
```

```
Loading required package: mvtnorm

Loading required package: survival

Loading required package: TH.data

Loading required package: MASS


Attaching package: 'TH.data'


The following object is masked from 'package:MASS':

    geyser
```

0.0250000000000021

```
 Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: User-defined Contrasts


Fit: lm(formula = y ~ tip + coupon, data = hardness)

Linear Hypotheses:
       Estimate Std. Error t value Pr(>|t|)
1 == 0  0.42500    0.06667   6.375 0.000129 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Adjusted p values reported -- single-step method)
```

0.425000000000001

since estimate = .425 and t test is significant, so there is a difference between 4 and 3

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?

[11]:
```
fit.gh = glht(lmod, linfct = mcp(tip = c(0,-1,1,0)))
summary(fit.gh)



fit.gh = glht(lmod, linfct = mcp(tip = "Tukey"))
summary(fit.gh)
```

```
 Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: User-defined Contrasts


Fit: lm(formula = y ~ tip + coupon, data = hardness)

Linear Hypotheses:
       Estimate Std. Error t value Pr(>|t|)
1 == 0 -0.15000    0.06667   -2.25    0.051 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)



 Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts


Fit: lm(formula = y ~ tip + coupon, data = hardness)

Linear Hypotheses:
           Estimate Std. Error t value Pr(>|t|)
2 - 1 == 0  0.02500    0.06667   0.375  0.98090
```

```
3 - 1 == 0 -0.12500    0.06667  -1.875  0.30292
4 - 1 == 0  0.30000    0.06667   4.500  0.00648 **
3 - 2 == 0 -0.15000    0.06667  -2.250  0.18165
4 - 2 == 0  0.27500    0.06667   4.125  0.01116 *
4 - 3 == 0  0.42500    0.06667   6.375  < 0.001 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)
```

there is evidence that there is difference between 4&1, 4&2, and 4&3

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. how likely a test is to find a real effect. It tells us if our test can spot something meaningful when it's actually there.
2. shows if our test is good at finding real effects. If power is high, the test is good at catching important things and avoiding mistakes where it misses them.
3. we might miss real effects in our data, or we might think we found something important when we didn't.

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. Different post-hoc tests exist because they tackle different problems in data analysis.

2. Tukey's Method is good when you want to compare pairs of groups. It helps prevent making too many Type I errors. Tukey's Method works well there are lots of comparisons, but it works best if the groups have similar variances and sizes.

3. Scheffé's method adjusts significance levels in linear regression to handle multiple comparisons. Unlike Tukey's method, Scheffé's method covers all possible contrasts among factor level means.