Statistical Models in R: Day 1

Challenge Questions for Experimental Design & ANOVA

Tiffany Timbers
Applied Statistics and Data Science Group
UBC Statistics

February 6, 2017

Experimental design questions

Balanced design - discuss

- What are the advantages of using a balanced experimental design?
- What are the disadvantages?
- Give an example of when a balanced design might not be possible.

Blocking factors and random effects - discuss

There are 3 undergraduates assisting you with your experiment that assess the addiction potential of Saturday morning cartoons in rats. You need to run the experiments every Saturday, but one of your undergraduate assistants can only help out 2 Saturdays a month, while the other two undergraduate assistants can be there every Saturday. Rat behaviour is sensitive to handler. What should you do?

Randomization - TRUE or FALSE?

A completely randomized design offers no control for lurking variables (a variable that is not included as an explanatory or response variable in the analysis)?

Answer:

FALSE! Although it does not control perfectly for lurking variables, a randomized design offers some control for lurking variables.

ANOVA questions

TRUE or FALSE?

ANOVA tests the null hypothesis that the sample means are all equal?

Answer:

FALSE! ANOVA tests the equality of the popula1on means.

TRUE or FALSE?

We use ANOVA to compare the variances of the population?

Answer:

FALSE! We use ANOVA to compare the popula1on means.

TRUE or FALSE?

A one-way ANOVA is equivalent to a t-test when there are 2 groups to be compared.

Answer:

TRUE! Two groups can be represented as a factor with 2 levels.

TRUE or FALSE?

In rejecting the null hypothesis, one can conclude that all the population means are different from one another?

Answer:

FALSE! We can only conclude that there are at least 2 different popula1on means. We cannot conclude that they are not all equal.

Perform a 1-way ANOVA

- Use built-in `iris` dataset in R and perform a 1-way ANOVA to assess differences in petal length and interpret results.
- `Species` is your explanatory variable
- `Petal.Length` is your response variable

```
head(iris)
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1         3.5          1.4         0.2  setosa
## 2          4.9         3.0          1.4         0.2  setosa
## 3          4.7         3.2          1.3         0.2  setosa
## 4          4.6         3.1          1.5         0.2  setosa
## 5          5.0         3.6          1.4         0.2  setosa
## 6          5.4         3.9          1.7         0.4  setosa
```

Perform a 2-way ANOVA with an interaction term

- Use built-in mtcars dataset in R to do a 2-way ANOVA (with interaction term) to assess the effect of number of forward gears and automatic vs. manual transmission on miles per gallon:
- `mpg` == miles per gallon (response variable)
- `gear` == number of forward gears
- `am` == automatic vs. manual transmission

```
head(mtcars)
```

```
##                   mpg cyl disp  hp drat   wt  qsec vs am gear carb
## Mazda RX4        21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag    21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710       22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive   21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7  8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant          18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

Sources/references

- Randomization question: http://stattrek.com/experiments/experimental-design.aspx?Tutorial=AP
- ANOVA true/false questions: Dr. Gabriela Cohen Freue's DSCI 562 course (UBC)