

Sam Tenney

Homework 2

1. Glyburide versus Insulin

- a) **Response variable:** percent fat mass of neonates of women with GDM, measured using total body electrical conductivity.

Factor: Medical treatment (insulin or glyburide)

Experimental Unit: Mothers to neonates who received either glyburide or insulin

- b) This study is an experiment because the women were randomly assigned different treatments of glyburide or insulin, they did not choose their medical treatment.
- c) The delimiter is a comma separating data in each column in the text file. There are three column headers in the file.
- d)
- e) The SAS file and the dataset are similar. There are three variables: id, treatment, and percentfatmass in both files. There are 82 observations in both files. The id and percentfatmass variables are both numeric, while treatment is a character type variable.
- f) It looks like the average percent fat mass for infants whose mothers were treated with glyburide were a little higher than those whose mothers were treated with insulin (12.8 versus 11.2). The spread for the glyburide treatment was a little larger as well (5.6997 versus 4.2008). See the table below for the means and standard deviation of the percent fat mass of the infants for the two treatments.

Summary of Neonate Percent Fat Mass

The MEANS Procedure

Analysis Variable : percentfatmass		
treatment	Mean	Std Dev
glyburid	12.8002439	5.6997370
insulin	11.2000000	4.2008089

g)

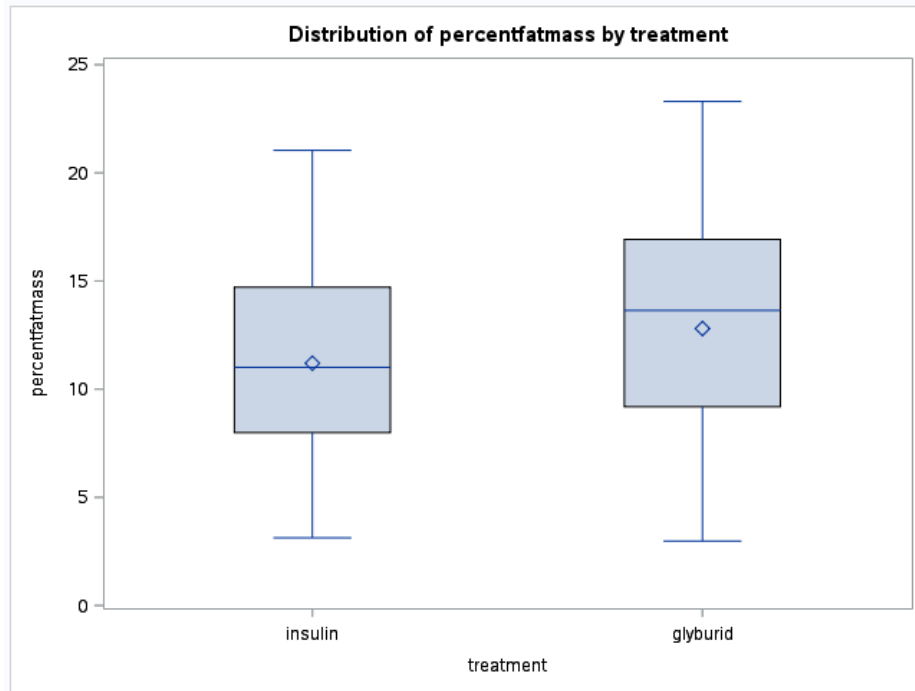


FIGURE 2.0.f. Boxplots of Percent Fat Mass in Neonates of Mothers with GDM. Boxplots for the percent fat mass (measured using total body electrical conductivity) in newborns of mothers with GDM. The 41 mothers were given glyburide and 41 other mothers were given insulin to test the difference in body composition of neonates based on the treatments expecting mothers received.

The distributions appear similar, although the glyburide treatment group has a larger spread and appears slightly left-skewed while the insulin appears symmetrical. The median and mean values of the glyburide group is slightly higher than that of the insulin group. There are no outliers.

h) I uploaded my glyburide.sas file to Learning Suite.

2. Finnish Fish

a) **Response Variable:** Weight in the different kinds of fishes.

Factor: Species of fish (Bream, Whitewash, Roach, Silver Bream, Smelt, Pike, Perch)

Experimental Unit: None. No treatment is given.

b) This is an observational study because no treatment is assigned to the subject. The scientists are merely collecting data on the different kinds of fish.

c) The data is tab-delimited and has no headers.

- d) The “\t” tab delimiter did not work because of the preceding tab before the data on each row. Instead, I used `sep=""` because it accounts for any amount of space larger than one whitespace.
- e) Calling the sex variable male allows us to use numbers instead of characters for the values in the column making it easier to aggregate the data. For example, instead of using “M” or “F”, we can use 1 or 0 to denote if the fish is male or not.
- f) The data is similar in the R file and in the dataset. There are nine variables all with the same names. There are 159 observations in the file and R dataset and all the variable types are what I expect (integers and numbers).

3. The Art of Oreo Dunking

- a) **Response variable:** The amount of milk absorbed after dunking an Oreo for 10 seconds in 2% milk, measured in grams (absorption is the difference in grams before and after dunking).

Factor: The type of Oreo cookie (Regular, Trader Joe’s, Double Stuffed).

Experimental Unit: Oreo cookies.

- b) This is an experiment because the cookie types were randomly assigned when they’d be dunked and there were various controls in place (how long to dunk the cookies, amount of milk, type of milk, etc.)
- c) See `myoreos.R`
- d) The R data object looks like the data from the project. There are 5 variables (`run`, `experiment`, `treatment`, `weightBefore`, and `weightAfter`) and 12 observations in each. The variable types are what I expected. `run` and `experiment` are integers since they are basic numbers with no decimals. `treatment` is a factor with the 3 levels of DS (Double Stuffed), Reg (Regular), and TJ (Trader Joe’s). `weightBefore` and `weightAfter` are both numeric types since they include decimals.
- e)
- f) The mean difference in weight (grams) after dunking the Regular Oreos for 10 seconds in milk is 6.18 grams, while for Trader Joe’s cookies, it was 4.845 grams. On average, the Regular Oreos absorbed more milk than the Trader Joe’s cookies. The spread for the Regular Oreos was a little bit larger, with a standard deviation of 0.8658, while Trader Joe’s had a standard deviation of 0.4535.

- g) The Trader Joe's Joe-Joe's absorption is mostly symmetrical while the Regular Oreos absorption is right skewed. There are no outliers. It appears all the Regular Oreos tested absorbed more milk than the Trader Joe's Joe-Joe's, as the left tail on the Regular Oreos is higher than the top of the right tail of the Trader Joe's Joe-Joe's.

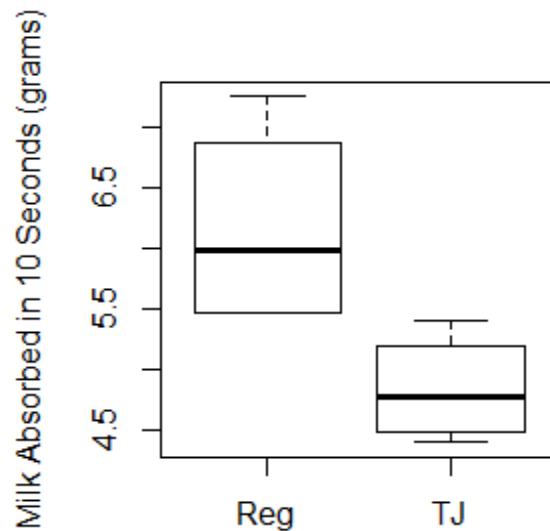


Figure 2.3.g. Boxplot of Milk Absorbed in 10 Seconds for Regular Oreos and Trader Joe's Joe-Joe's. This shows the amount of milk absorbed in 10 seconds in grams for the different cookies, Regular Oreos and Trader Joe's Joe-Joe's.

- h) The Regular Oreos distribution is higher than the distribution of the Trader Joe's Joe-Joe's cookies. Regular Oreos appear to be slightly right skewed, while Trader Joe's Joe-Joe's are nearly symmetrical. There don't appear to be any outliers. The spread for Regular Oreos is larger than the spread of Trader Joe's Joe-Joe's.

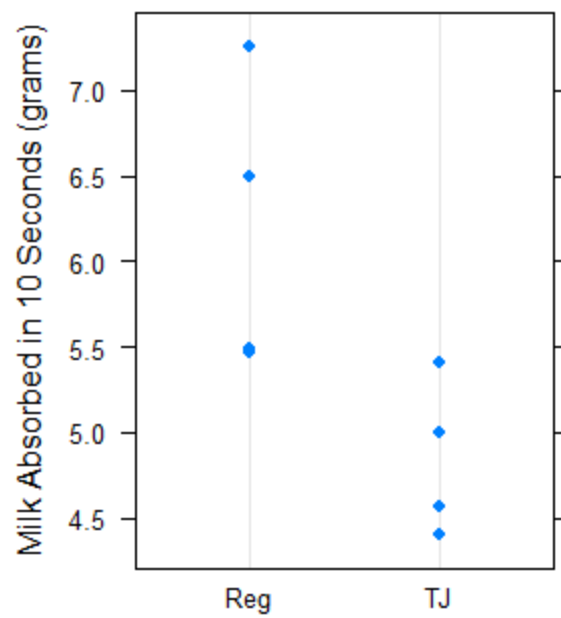


Figure 2.3.h. Dotplot of Milk Absorbed in 10 Seconds for Regular Oreos and Trader Joe's Joe-Joe's. This shows the amount of milk absorbed in 10 seconds in grams for the different cookies, Regular Oreos and Trader Joe's Joe-Joe's.