

Introduction to Biostatistics (171:161)
Breheny

Lab #12

In lab #12, we are going to analyze two-group studies in which the outcome is continuous. Today's lab will focus on the use of the two-sample t -test, confidence intervals for the difference between two means, and transformations and rank-based methods.

A note about permutation tests: It is possible to perform permutation tests in SAS and R, but as these tests are computer intensive and not particularly common, we won't be going over them.

1 Infant diarrhea study

Diarrhea is a major public health problem in underdeveloped countries, especially for babies. Diarrhea leads to dehydration, which results in millions of deaths each year worldwide. Bismuth salicylate (the active ingredient in Pepto Bismol) has been shown to reduce diarrhea in adults. Researchers in Peru conducted a double-blind randomized controlled trial, published in *The New England Journal of Medicine*, to determine whether it would do so in infants suffering from diarrhea as well.

In their study, all infants received the standard therapy for diarrhea: oral rehydration. In addition to the rehydration, 85 babies received bismuth salicylate, while 84 babies received a placebo. The total stool volumes for all infants over the course of their illness was measured. To adjust for body size, the researchers divided by body weight to obtain their outcome of interest: stool output per kilogram of body weight. The results of their study are available on the course website.

2 Exploration

Download and import the data set `diarrhea.txt`. With continuous data, it is always a good idea to explore and plot your data before diving into the analysis. Box plots are useful:

SAS:

```
PROC SGPLOT DATA=diarrhea;  
  VBOX Stool / CATEGORY=Group;  
RUN;
```

R:

```
boxplot(Stool~Group,col="gray")
```

So are histograms:

SAS:

```
PROC SGPNEL DATA=diarrhea;
  PANELBY Group;
  HISTOGRAM Stool;
RUN;
```

R:

```
require(lattice)
histogram(~Stool|Group)
```

Clearly, this data is very right-skewed and not at all normal-looking. At this point, you face a choice about how to analyze the data: using t -tests which are best for normally distributed data, or using nonparametric methods which are better for non-normally distributed data. The nonparametric methods will probably be better for this study. However, with large sample sizes of 84 and 85 in each group, conducting a t -test would not be unreasonable.

Let's look at one more type of plot before we move on to testing, as it is a very common type of plot in certain fields: a bar chart of means in the two groups. In SAS:

```
PROC SGPLOT DATA=diarrhea;
  VBAR Group / RESPONSE=Stool;
RUN;
```

However, as you'll notice, this is not what we want. By default, SAS chooses to report the totals in the two groups. To get the means instead, type:

```
PROC SGPLOT DATA=diarrhea;
  VBAR Group / RESPONSE=Stool STAT=MEAN;
RUN;
```

When such graphs are included in journal articles, they are often accompanied by error bars. There are three common choices of error bars: the standard deviation, the standard error, and 95% confidence limits. SAS can give you any of the three, depending on the `LIMITSTAT` option. For example,

```
PROC SGPLOT DATA=diarrhea;
  VBAR Group / RESPONSE=Stool STAT=MEAN LIMITSTAT=CLM;
RUN;
```

Try replacing `CLM` with `STDDEV` and `STDERR` and observe how the three choices tell you quite different things.

R does not provide a similar function by default, but it is not hard to write one. I wrote a function called `error.bar` that provides the same output (the first line below links to the function I wrote):

```
source("http://bit.ly/HzbVt3")
error.bar(Stool,Group)
error.bar(Stool,Group,conf.level=.8)
error.bar(Stool,Group,bar="SE")
error.bar(Stool,Group,bar="SD")
```

3 Two-sample t test

Suppose that we decide to analyze our data using a two-sample t -test. This is quite straightforward: you just need to specify which variable separates the two groups (SAS calls this `CLASS`) and which

variable is the outcome (SAS calls this VAR):

SAS:

```
PROC TTEST DATA=diarrhea;
  CLASS Group;
  VAR Stool;
RUN;
```

R:

```
t.test(Stool~Group,
       var.equal=TRUE) ## Student's
t.test(Stool~Group) ## Welch's
```

In SAS, we get a lot of useful output, including some of the plots we drew in the previous section. SAS tells you the means, standard deviations, and ranges in the two groups, as well as the mean of the difference and the pooled standard deviation. In addition, it calculates p -values and confidence intervals based on Student's t -test and Welch's t -test.

Note: Welch and Satterthwaite independently came up with the same idea for a t -test that doesn't assume that the standard deviations (or variances – the variance is just the SD^2) of the two groups are equal. Some give credit to Welch, some give credit to Satterthwaite, some call the approach the Welch-Satterthwaite t -test. But they're all doing the same thing.

Note: SAS analyzes the data in two ways: assuming the variability is the same in each group (Student's) and allowing them to be different (Welch's). In R, you must specify whether or not you would like to assume equal variability or not through the `var.equal` option. By default, R takes a conservative approach and refrains from assuming equal variability.

The code above also provides confidence intervals. Note that we can obtain two different confidence intervals – one based on pooling the standard deviations and one which keeps them separate. The same issues apply to confidence intervals that applied to t -tests when it comes to choosing between these two intervals. In this study, we have pretty large sample sizes, so we might as well use the non-pooled (Satterthwaite) confidence interval, although the standard deviations and sample sizes are similar, so it should come as no surprise that the two methods give virtually identical confidence intervals.

4 Practice Problem

Some infants are born with congenital heart defects and require surgery very early in life. One approach to performing this surgery is known as “circulatory arrest.” A downside of this procedure, however, is that it cuts off the flow of blood to the brain, possibly resulting in brain damage. An alternative procedure, “low-flow bypass” maintains circulation to the brain, but does so with an external pump that potentially causes other sorts of injuries to the brain.

To investigate the treatments, surgeons at Harvard Medical School conducted a randomized controlled trial. In the trial, 70 infants received low-flow bypass surgery and 73 received the circulatory arrest approach. The researchers looked at two outcomes: the Psychomotor Development Index (PDI), which measures physiological development, and the Mental Development Index (MDI), which measures mental development. For both indices, higher scores indicate greater levels of development. The results of their study are on the course website.

- (a) Calculate the standard deviations of PDI and MDI for each group. Do you think that pooling the standard deviations will be a problem?
- (b) Conduct a t -test (whichever one you think is most appropriate) to determine whether the difference in physiological development between the infants in the circulatory arrest group and the low-flow bypass group is statistically significant.
- (c) Conduct a t -test (whichever one you think is most appropriate) to determine whether the difference in mental development between the infants in the circulatory arrest group and the low-flow bypass group is statistically significant.
- (d) Construct a 95% confidence interval for the difference in mental development between the two surgical groups. Describe the confidence interval in terms of the context of the study (what does the confidence interval mean? which group did better?).
- (e) If your child had to have open-heart surgery as an infant, which treatment option would you prefer? Why?