Lab 8: Two-sample inference

STAT218

This lab focuses on two-sample inference for differences in population means using paired and independent data. The main objectives are:

1. Learn to implement two-sample tests in R
2. Practice distinguishing between paired and independent data settings

Our goal is to learn implementations, so less focus will be placed on interpretation in these instructions; however, we will discuss interpretations as we go.

## Datasets

The lab uses several datasets for which we will consider two-sample comparisons:

* swim: compare swimming velocities in a bodysuit and in a swimsuit
* famuss: compare changes in dominant and nondominant arm strength after resistance training
* finch: compare finch beak depths the generation before a drought and the generation after
* diets: compare blood pressure on a fish oil diet and a regular oil diet
* sleep: compare extra sleep gained on two different soporific drugs

You should run the lines below before beginning the lab to load all five datasets into your environment.

library(oibiostat)  
library(Sleuth3)  
data(swim)  
data(famuss)  
data(sleep)  
finch <- case0201  
diets <- ex0112  
cloud <- case0301

## Performing two sample tests

First we’ll cover examples of tests for paired and independent tests, and you’ll practice implementing each on a new dataset.

The goal of two-sample inference is to make a comparison of means between two groups based on two sets of observations (one per group). This is usually structured as follows:

* one variable of interest
* a grouping of observations into two groups
* possibly, a pairing of observations between groups

### Paired data

Recall that **pairing** refers to matching the observations between the two groups of interest. This is most often done when the study units are the same in each group. In the swim data, for example, each swimmer’s velocity is measured once in a bodysuit and once in a swimsuit. If denotes mean velocity, we want to test:

Because of the pairing, we can compute the difference in velocity for each swimmer and use that to test the hypotheses. There are two steps in this case:

1. Calculate pairwise differences
2. Perform a one-sample test with the appropriate direction

This is shown below:

# step 1: compute difference wetsuit - swimsuit  
velocity.diff <- swim$wet.suit.velocity - swim$swim.suit.velocity  
  
# step 2: perform one-sample t test  
t.test(velocity.diff, mu = 0, alternative = 'greater')

One Sample t-test  
  
data: velocity.diff  
t = 12.318, df = 11, p-value = 4.443e-08  
alternative hypothesis: true mean is greater than 0  
95 percent confidence interval:  
 0.06620114 Inf  
sample estimates:  
mean of x   
 0.0775

|  |
| --- |
| Your turn |
| In the FAMuSS data, changes in arm strength by dominant and nondominant arms are paired by participant. Use this pairing to test the hypotheses  where denotes mean change in arm strength. Be careful about the direction of the test; make sure this matches the sign of the pairwise differences appropriately.  # inspect data head(famuss)  # compute difference ndrm.ch - drm.ch strength.diff <- ...  # perform t test t.test(...) |

### Independent data

For independent data, the problem is the same but without the pairing. If the sample sizes differ between groups, that’s a dead giveaway that data are very likely not paired (unless observations are missing). Otherwise, you’ll need to reflect on whether it is possible to match the observations in each group.

In the finch data, the observations in each group come from different individual finches of generations two years apart, so there is no sensible way to pair the data. The goal is to determine whether beak depth increased as a result of selection pressure following a drought in 1977. So, with denoting mean beak depth, let’s test the hypotheses:

The implementation is shown below. Notice that *unlike* in the one-sample case, we provide the data as a data frame with a grouping column and a column for the variable of interest, and specify the grouping using the formula variable ~ group. Otherwise the input format is analogous to the one-sample case.

# examine data; note that grouping is shown as a variable  
head(finch)

Year Depth  
1 1976 6.2  
2 1976 6.8  
3 1976 7.1  
4 1976 7.1  
5 1976 7.4  
6 1976 7.8

# perform t test  
t.test(Depth ~ Year, data = finch, mu = 0, alternative = 'less')

Welch Two Sample t-test  
  
data: Depth by Year  
t = -4.5833, df = 172.98, p-value = 4.37e-06  
alternative hypothesis: true difference in means between group 1976 and group 1978 is less than 0  
95 percent confidence interval:  
 -Inf -0.427321  
sample estimates:  
mean in group 1976 mean in group 1978   
 9.469663 10.138202

|  |
| --- |
| Your turn |
| Try implementing the two-sample test for independent data using a new dataset. The diets data show reductions in diastolic blood pressures for 14 participants who were randomly assigned one of two diets: fish oil and regular oil. Because the participants in each group are unrelated, there is no pairing of observations.  Use these data to test whether fish oil leads to a greater reduction in blood pressure:  Again, careful about the direction of the alternative. You might need to try once and inspect output to make sure you performed the test indicated above.  # inspect data head(diets)  # perform test t.test(..., data = ..., mu = ..., alternative = ...) |

## Distinguishing paired from independent data

Here you’ll practice distinguishing the paired and independent settings, as well as get further practice implementing the tests above. You’re given two datasets, and must:

1. Determine whether pairing is present
2. Conduct the appropriate test

You should discuss the pairing question with your group after inspecting the datasets but before carrying out any inference.

### Cloud seeding data

The cloud seeding data are shown below, and comprise observations of rainfall in a target area on 26 days when clouds were seeded with silver iodide and on 26 different days when clouds were not seeded.

# inspect data  
head(cloud)

Rainfall Treatment  
1 1202.6 Unseeded  
2 830.1 Unseeded  
3 372.4 Unseeded  
4 345.5 Unseeded  
5 321.2 Unseeded  
6 244.3 Unseeded

With denoting mean rainfall (in acre-feet), the goal is to test:

### Sleep data

The sleep data are shown below and comprise observations of extra hours of sleep of 10 study participants on one of two drugs. The variable extra records extra hours of sleep, group records which drug was taken, and ID records the study participant.

# inspect data after ordering by participant ID  
dplyr::arrange(sleep, ID) |> head()

extra group ID  
1 0.7 1 1  
2 1.9 2 1  
3 -1.6 1 2  
4 0.8 2 2  
5 -0.2 1 3  
6 1.1 2 3

With denoting mean extra hours of sleep, the goal is to test with this data whether one drug is better than the other:

The hypotheses are non-directional because there is not a prior hypothesis about which drug might be better.

|  |
| --- |
| Your turn |
| # test whether cloud seeding increases rainfall  # test whether sleep drugs produce different effects |