Homework 6 Solution

Chuan Tian

2017-11-18

```
library(tidyverse)
library(pander)
```

1. Properties of two sample t-tests

a. Show algebraically, that Welch's t-test and two sample equal variance t-test, have same test statistic when the sample sizes are the same (i.e. $n = m$).

   For both Welch's t-test and two sample equal variance t-test,

$$t = \frac{\overline{Y} - \overline{X}}{\widehat{Var}\left(\overline{Y} - \overline{X}\right)}$$

   The only difference between the two tests is the denominator, i.e. the estimate of the variance of $\overline{Y} - \overline{X}$. When $n = m$, we estimate the variance of $\overline{Y} - \overline{X}$ with

$$\frac{s_Y^2}{n} + \frac{s_X^2}{n}$$

   in Welch's t-test. And in two sample equal variance t-test, we estimate the variance of $\overline{Y} - \overline{X}$ with

$$s_p^2(\frac{1}{n} + \frac{1}{n}) = \frac{2 \cdot s_p^2}{n}$$

   , where

$$s_p^2 = \frac{(n-1)s_X^2 + (n-1)s_Y^2}{n+n-2} = \frac{s_X^2 + s_Y^2}{2}$$

   . Hence,

$$\frac{2 \cdot s_p^2}{n} = \frac{2}{n} \cdot \frac{s_X^2 + s_Y^2}{2} = \frac{s_Y^2}{n} + \frac{s_X^2}{n}$$

   , which is equal to the estimate of variance of $\overline{Y} - \overline{X}$ in Welch's t-test.

b. Recall that in the case where $n = m$, Welch's test estimates the variance of $\overline{Y} - \overline{X}$ with:

$$\frac{s_Y^2}{n} + \frac{s_X^2}{n}$$

   whereas the paired t-test estimates the variance of $\overline{Y} - \overline{X}$ with

$$\frac{s_Y^2}{n} + \frac{s_X^2}{n} - 2\frac{s_{YX}}{n}$$

c. If the data are truly paired, i.e $\sigma_{XY} \neq 0$, on average, what will the two tests estimate for the variance of $\overline{Y} - \overline{X}$? For paired t-test,

$$E(\widehat{Var}(\overline{Y} - \overline{X}))$$
$$= E(\frac{s_Y^2}{n} + \frac{s_X^2}{n} - 2\frac{s_{YX}}{n})$$

$$= E(\frac{s_Y^2}{n}) + E(\frac{s_X^2}{n}) - 2E(\frac{s_{YX}}{n})$$

$$= \frac{\sigma_Y^2}{n} + \frac{\sigma_X^2}{n} - 2\frac{\sigma_{YX}}{n}$$

Whereas for Welch's t-test,

$$E(\widehat{Var}(\overline{Y} - \overline{X}))$$

$$= E(\frac{s_Y^2}{n} + \frac{s_X^2}{n})$$

$$= E(\frac{s_Y^2}{n}) + E(\frac{s_X^2}{n})$$

$$= \frac{\sigma_Y^2}{n} + \frac{\sigma_X^2}{n}$$

ii. Use your answer to (i) to argue that when the data are truly paired, using the unpaired test (Welch's) may be misleading.

If the data is truely paired, i.e. $\sigma_{XY} \neq 0$, Welch's t-test will be misleading since it didn't take this into account. For instance, if $\sigma_{XY} > 0$, Welch's t-test will over estimate the variance of $\overline{Y} - \overline{X}$, and vice versa.

iii. If the data are not paired, i.e $\sigma_{XY} = 0$, on average, what will the two tests estimate for the variance of $\overline{Y} - \overline{X}$?

The estimate of variance will stay the same as in (i), except that they are equal now since $\sigma_{XY} = 0$.

iv. Use your answer to (iii) to argue that if the data are not paired, using the paired test is not misleading. The answer follows from (iii), since using which test doesn't make a difference on the estimate of variance of $\overline{Y} - \overline{X}$, on average.

v. In the the case where the data are unpaired, although the paired test is not misleading, it has less power than the two sample t-test. Design and conduct a simulation to demonstrate this fact.

Chuan's simulations:

Let $Y_i, i = 1, \ldots, n$ be i.i.d $N(0, 1)$ and $X_i, i = 1, \ldots, n$ be i.i.d $N(1, 1)$.

Table 1: Power for the two tests with $\alpha = 0.05$

Sample size, n	two sample t-test	paired t-test
10	0.567	0.525
20	0.863	0.849
50	0.999	0.998

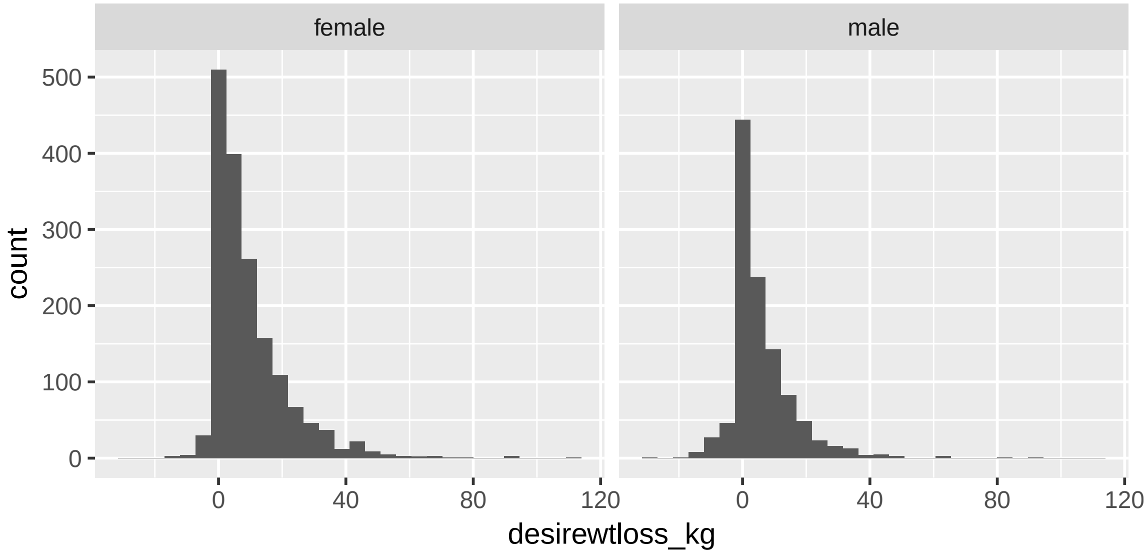
Looks like the power of paired t-test is smaller than the power of two sample t-test when data is actually unpaired. However, as sample size n increases, the difference of power between the two tests decreases.

2. Data Analysis

a. Using the BRFSS data from previous homeworks: Is the mean desired weight loss the same for male and female US residents?

```
download.file("http://st551.cwick.co.nz/data/brfss.rds",
  "brfss.rds", mode = "wb")
brfss <- read_rds("brfss.rds")
brfss <- brfss %>% mutate(desirewtloss_kg = weight_kg - wtdesire_kg)
```

```r
ggplot(data = brfss, aes(x = desirewtloss_kg)) +
  geom_histogram() +
  facet_wrap(~ sex)
```



The two samples (men and women) are independent from each other, so we use two sample test. Neither population seems normally distributed, but we have large enough sample size for C.L.T. to kick in. So two sample t-test can be used. Though it seems that the men's and women's desired weight loss has similar spread, we don't have any knowledge about the population variance for sure. It's safer to use Welch's t-test in this case.

```r
(t_test_dswt <- t.test(desirewtloss_kg ~ sex, data = brfss))
```

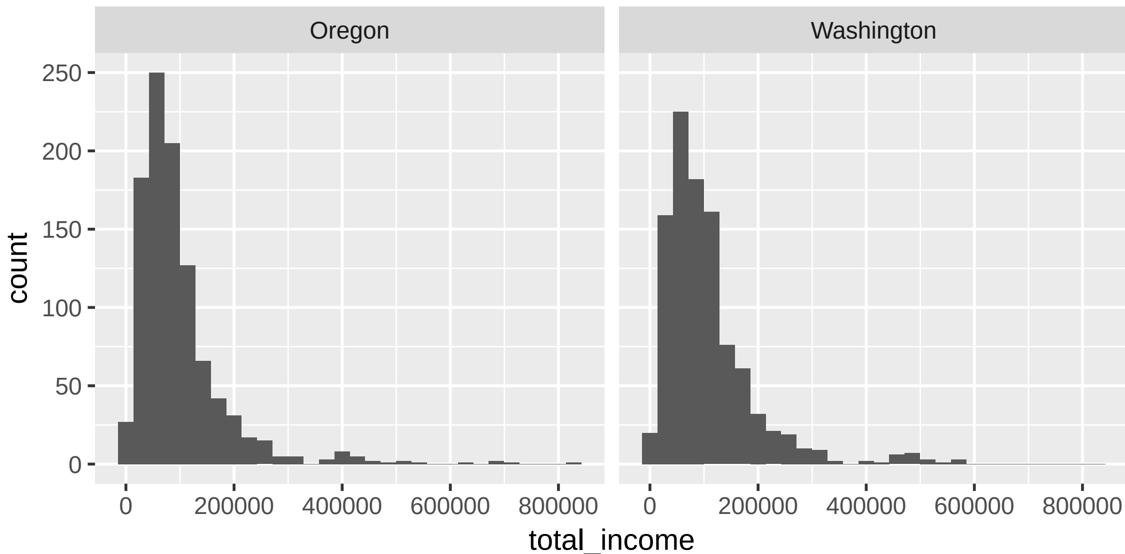
```
##
##  Welch Two Sample t-test
##
## data:  desirewtloss_kg by sex
## t = 8.8747, df = 2635.2, p-value < 0.00000000000000022
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.938029 4.604558
## sample estimates:
## mean in group female   mean in group male
##             9.621183             5.849890
```

There is convincing evidence that the mean desired weight loss is different for male and female US residents (Welch's t-test, two-sided p-value < 0.001). It is estimated that US females have a desired weight loss 3.8 kg higher than US males', on average. With 95% confidence the desired weight loss of US females is between 2.9 and 4.6 kg higher than the desired weight loss of US males, on average.

b. Using the ACS data described below: Do Washington households have the same mean total household income as Oregon households?

```r
library(tidyverse)
download.file("http://st551.cwick.co.nz/data/couples_wide.csv",
  "couples_wide.csv")
couples_wide <- read_csv("couples_wide.csv")
couples_wide <- couples_wide %>%
  mutate(total_income = husband_total_income + wife_total_income, age_difference = husband_age - wi
```

```r
ggplot(data = couples_wide, aes(x = total_income)) +
  geom_histogram() +
  facet_wrap(~ state)
```



The two samples (OR and WA) are independent from each other, so we use two sample test. Neither population seems normally distributed, but we have large enough sample size for C.L.T. to kick in. Though it seems that the OR's and WA's total household income has similar spread, we don't have any knowledge about the population variance for sure. It's safer to use Welch's t-test in this case.

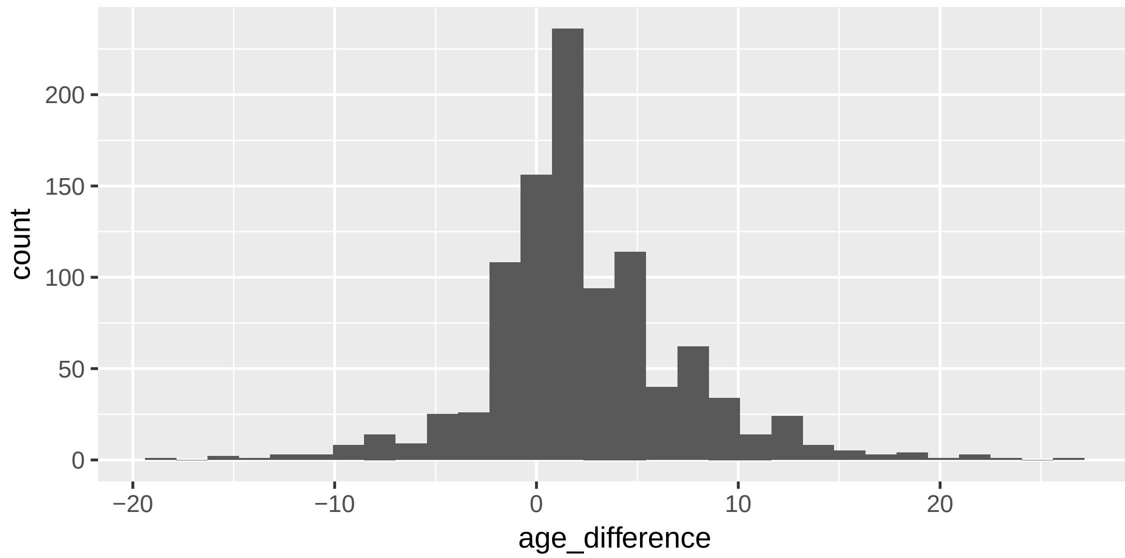
```r
(t_test_income <- t.test(total_income ~ state, data = couples_wide))
```

```
##
##  Welch Two Sample t-test
##
## data:  total_income by state
## t = -2.2539, df = 1996.5, p-value = 0.02431
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -16275.44  -1130.22
## sample estimates:
##     mean in group Oregon mean in group Washington
##                 97718.11                106420.93
```

There is moderate to strong evidence that Washington households have different mean total household income from Oregon households (Welch's t-test, p-value = 0.024). It is estimated the Washington households have total household income 8702.8 USD higher than Oregon households, on average. With 95% confidence the Washington households have total household income between 1130.2 and 16275.4 USD higher than Oregon house households, on average.

c. Using the ACS data described below: How much older are husbands than their wives in Oregon opposite sex married couples?

```r
couples_wide_OR = filter(couples_wide, state == "Oregon")
ggplot(data = couples_wide_OR, aes(x = age_difference)) +
  geom_histogram()
```



Since husbands and wives in the same household are paired, we use one sample test (or paired-test). The age difference in Oregon opposite sex married couples seems somewhat bell-shaped although has long tails compared to a Normal. And again, we have sufficient sample size for C.L.T. to kick in. So one sample t-test (paired t-test) can be used.

```
(t_test_age <- t.test(couples_wide_OR$age_difference))
```

```
##
##  One Sample t-test
##
## data:  couples_wide_OR$age_difference
## t = 15.095, df = 999, p-value < 0.00000000000000022
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  2.04016 2.64984
## sample estimates:
## mean of x
##     2.345
```

There is convincing evidence that husbands are older than their wives on average in Oregon households (one sample t-test, two-sided p-value < 0.001). It is estimated that husbands are 2.3 years older than their wivs in Oregon households, on average. With 95% confidence the husbands in Oregon households are between 2 and 2.6 years older than their wives, on average.