T-tests: one-sample, paired, and two-sample tests

Review

- Random samples
- Standard normal distribution
    - What is a z-score?
    - What are the quantiles?
    - How do you capture height vs. area?
- Sampling distribution of the sample mean
- Standard error
- Confidence intervals
- What is a z-score?
- t-statistics vs. z-statistics : when is each appropriate?
- P values : one- vs. two-sided
- What does qt(0.975) represent?

One-sample, paired-sample, and two-sample t-tests

This worksheet illustrates basic concepts and practical application of significance tests for different experimental scenarios.

We will use a simple case study where a drug was administered to 10 random patients (or test subjects) and its effect was compared to that of a placebo pill. Three different experimental designs are illustrated:

- Treatment sample vs. control population (one-sample test)
- Two random samples: treatment vs. control (two-sample test)
- One sample measured before and after treatment (paired test)

In all three scenarios, measurements were collected to answer the question:

Is there a significant difference between the control subjects and the those who were given the drug?

```
# the population
Placebo_pop = read.csv("placebo_pop", header = FALSE)
Placebo_pop = Placebo_pop[,1]

# sample data
Placebo = c(54,51,58,44,55,52,42,47,58,46)
Drug = c(54,73,53,70,73,68,52,65,65,60)
```

Confidence Interval

The standard error is very helpful because it gives us an idea of how close our data are to the actual mean. We can use the SE to help define **Confidence Intervals (CI)** of what the the actual population mean is.

Here, we assume that we have random samples and that the measurements are normally distributed in the population. The estimated distribution of the sample means will then follow a t-distribution, which is similar to a standard normal distribution but with heavier tails. As the sample size increases, the tails of the t-distribution become smaller and it resembles closely the Z-distribution.

Let's take our Drug sample and try to calculate our 95% CI for the mean of the population, assuming that the drug has no effect.

- Estimate the mean and SE for the Drug sample.
- Use the `qt()` *quantile function* to identify the critical t-score for the the 95% CI with the appropriate df.
- Use the estimated mean and t_{crit} to get our 95% CI.

```
Dse=sd(Drug)/sqrt(length(Drug))
Dse
## [1] 2.564934

t_0.975 = qt(p = .975, df = length(Drug)-1)
t_0.975   # 2.26 = z-score for 97.5% area
## [1] 2.262157

Dse * t_0.975
## [1] 5.802285
# now can calculate 95% CI
Dmean = mean(Drug)
c(Dmean - Dse * t_0.975, Dmean + Dse * t_0.975)
## [1] 57.49772 69.10228

# use sample SD to approximate population SD
# now can get an estimate how how likely it is that we have captured the true pop mean
```

One-sample t-test

We can check our answer by using the `t.test()` function. Do this for both the Drug and the Placebo samples. What do you conclude?

```
t.test(Placebo, mu=mean(Placebo_pop))
##
##  One Sample t-test
##
## data:  Placebo
## t = 0.0014947, df = 9, p-value = 0.9988
## alternative hypothesis: true mean is not equal to 50.6973
## 95 percent confidence interval:
##   46.6107 54.7893
## sample estimates:
## mean of x
##      50.7
t.test(Drug, mu=mean(Placebo_pop))
##
##  One Sample t-test
##
## data:  Drug
## t = 4.9135, df = 9, p-value = 0.0008321
## alternative hypothesis: true mean is not equal to 50.6973
```

```
## 95 percent confidence interval:
##  57.49772 69.10228
## sample estimates:
## mean of x
##      63.3
```

One-sample t-test using an estimate for the population mean

Now let's pretend we don't know the population, so that Placebo is just a sample from some unknown population. Since we don't know the true σ, we will estimate it by calculating the *sd* of the sample. Note that this only works if the population is truly normally distributed.

Calculate the 95% CI for the true Placebo mean using the Placebo sample.

```
Pse=sd(Placebo)/sqrt(length(Placebo))
Pse  # 1.8 -- this is one SD from the sample mean
## [1] 1.8077

t_0.975 = qt(p = .975, df = length(Placebo)-1)
t_0.975  # 2.26 = z-score for 97.5% area
## [1] 2.262157

# now we can compute
# point on x-axis in original distribution that corresponds to z = 2.26
Pse * t_0.975
## [1] 4.089301
# now can calculate 95% CI
Pmean = mean(Placebo)
c(Pmean - Pse * t_0.975, Pmean + Pse * t_0.975)
## [1] 46.6107 54.7893

# use sample SD to approximate population SD
# now can get an estimate how how likely it is that we have captured the true pop mean
```

Let's test our estimates for the Placebo population mean using a t-test. WE have to provide $\bar{x}$ as the estimate for the true population mean since we are pretending that we don't know μ.

```
t.test(Placebo, mu=mean(Placebo))
##
##  One Sample t-test
##
## data:  Placebo
## t = 0, df = 9, p-value = 1
## alternative hypothesis: true mean is not equal to 50.7
## 95 percent confidence interval:
##  46.6107 54.7893
## sample estimates:
## mean of x
##      50.7
t.test(Drug, mu=mean(Placebo))
##
##  One Sample t-test
##
## data:  Drug
```

```
## t = 4.9124, df = 9, p-value = 0.0008333
## alternative hypothesis: true mean is not equal to 50.7
## 95 percent confidence interval:
##  57.49772 69.10228
## sample estimates:
## mean of x
##     63.3
```

How do these results differ from the previous test using the true population mean?

Difference between sample means - two-sample t-test

Now let's start to consider that Placebo and Drug groups really are two separate populations. Now we want to estimate (with 95% confidence) the true difference in the means of the two populations.

We already know that the distribution of the sample mean is normally distributed, and that the sum or difference of two normally distributed variables is also normally distributed. Therefore, the difference in sample means must be normally distributed.

When we consider the means of two samples, the SE of the difference is the square root of the sum of the SE of each sample:

$$\sigma_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Compute the SE of the difference below.

```
se_diff = ( sqrt (
  ((sd(Placebo)**2)/length(Placebo)) +
    ((sd(Drug)**2)/length(Drug))))

se_diff
## [1] 3.13794
```

95% CI

We can calculate the 95% confidence interval by using the t-distribution. In this case, the degrees of freedom is the sum of the total values minus the number of samples ($df = 20 - 2$).

```
Diff_mean = mean(Placebo) - mean(Drug)   # -12.6
Diff_mean
## [1] -12.6

qt(p=0.025, df=18)*se_diff
## [1] -6.592567
qt(p=0.975, df=18)*se_diff
## [1] 6.592567

# these are all equivalent
Diff_mean + qt(p=0.025, df=18)*se_diff
## [1] -19.19257
Diff_mean - qt(p=0.025, df=18)*se_diff
## [1] -6.007433
```

```
Diff_mean - qt(p=0.975, df=18)*se_diff
## [1] -19.19257
Diff_mean + qt(p=0.975, df=18)*se_diff
## [1] -6.007433

Diff_mean + qt(p=0.025, df=18)*se_diff
## [1] -19.19257
Diff_mean + qt(p=0.975, df=18)*se_diff
## [1] -6.007433
```

Notice that the confidence interval does not span 0, which means it is extremely unlikely that the true difference between the Drug and Placebo populations is 0.

Let's check is this is true.

We can look at the t-statistic (which is the equivalent to the Z-score for a normal distribution). Compute the t-score for the difference in the means.

```
tstat = Diff_mean / se_diff
tstat
## [1] -4.015373
```

Two-sample t-test

Now we can calculate the p-value as to whether our null hypothesis is true. The **null hypothesis** is that the difference between the two populations is 0. We know this because if we were to plot a distribution from random variables with the same mean, we would expect to get a difference of 0.

$$H_o : \mu_{Drug} = \mu_{Placebo}$$

$$H_A : \mu_{Drug} \neq mu_{Placebo}$$

First, calculate the test statistic by hand, then confirm it using the `t.test()` function.

```
#using our calculations done by hand
2*pt(q = tstat,
     df = (length(Placebo)+length(Drug)-2),
     lower.tail = T)
## [1] 0.0008115625

#confirmation using the t.test() function.
t.test(Placebo, Drug, var.equal = T)
##
##  Two Sample t-test
##
## data:  Placebo and Drug
## t = -4.0154, df = 18, p-value = 0.0008116
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -19.192567  -6.007433
## sample estimates:
## mean of x mean of y
##      50.7      63.3
```

Paired *t*-test

What if the data are taken from the **same individuals**? For example, we could test the same patients before and after treatment. In this kind of experimental design, we say that the data are **paired**. If there is no difference between the two measurements for each individual – for example, a new drug for blood pressure has no measurable benefit – then we would expect that our before and after values would be about the same on average.

The **paired *t*-test** is performed in the same way as the one-sample *t*-test, except we use the **mean difference between paired measurements** from the two samples, $\bar{X}_D$, to compute a test statistic. Our **null hypothesis** is usually that the mean difference, D_o, is zero (we could set it to something else if our null hypothesis is that the difference between them is something else ...).

For paired data, we assume that the two sets of measurements are arranged in the same order as the corresponding individuals (because we have good record-keeping practices!) The *t*-statistic is:

$$t^* = \frac{\bar{X}_D - D_0}{\frac{s_D}{\sqrt{n}}} = \frac{\bar{X}_D - D_0}{SE_D} = \frac{\sqrt{n}}{s_D}(\bar{X}_D - D_0)$$

where $\bar{X}_D$ is the mean of the pairwise differences, s_D is the standard deviation of the pairwise differences, and D_0 is what we are testing (which in this case is 0).

To compute the test statistic, we can simply subtract one vector from the other to obtain pair-wise differences for each individual, take the mean, and divide by the standard error. We then find the *p* value in the usual way.

Study the above equation to convince yourself that the form of the *t*-statistic above is the same as that for a one-sample *t*-test, substituting $\bar{X}_D$ for $\bar{X}$, D_o for μ_o, and s_D for s.

Do this by hand first, and then using the `t.test()` function.

```
DrugDiff = Drug-Placebo

t_stat = mean(DrugDiff)/(sd(DrugDiff)/sqrt(length(DrugDiff)))
2*pt(t_stat, df=9, lower.tail = T)
## [1] 1.997358

t.test(DrugDiff, mu=0)
##
##  One Sample t-test
##
## data:  DrugDiff
## t = 4.1087, df = 9, p-value = 0.002642
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##   5.662718 19.537282
## sample estimates:
## mean of x
##      12.6
```

Importantly, the result from the *t*-test with the paired option gives the same result as the one-sample *t*-test using the Placebo sample to estimate the true population parameter $\mu_{Placebo}$!

Note: when performing a paired *t*-test, you must specify the **test set first**, and the **control set second** – otherwise you get a *t*-score that is on the opposite side of the distribution. This will not change your *p*-value, however your **confidence interval** will have the wrong sign!!!

```
# t.test(treatment, control, ...)
t.test(Drug, Placebo, paired = T, var.equal = T)
##
##  Paired t-test
##
## data:  Drug and Placebo
## t = 4.1087, df = 9, p-value = 0.002642
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##    5.662718 19.537282
## sample estimates:
## mean of the differences
##                    12.6
```