Student t Tests

March 09, 2021

Review

Central Limit Theorem

$$\sqrt{n}\left(\frac{\overline{X} - \mu}{\sigma}\right) \sim N(0, 1)$$

- Approximation as n increases
- Larger n needed if X is skewed or has larger variance
- Assumes true value of σ^2 is known

Often, however, σ^2 is not known and must be approximated with

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n}(X_i - \overline{X})^2$$
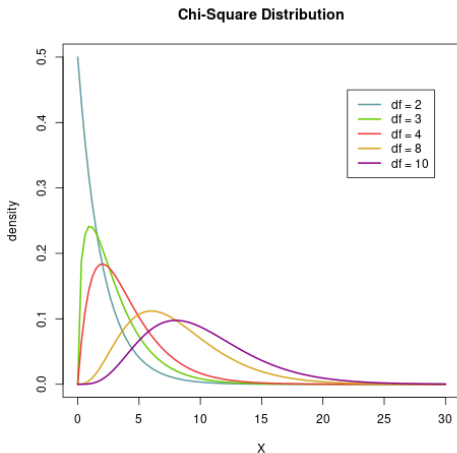
Degrees of Freedom

Note that unlike the calculation of $\overline{X}$, the sample variance is divided by the term $n-1$. This is known as the *degrees of freedom*.

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$$

The degrees of freedom indicate the number of values in our collection that are free to vary given a number of constraints (n - total constraints). As our computation of S^2 requires that we first compute $\overline{X}$, we have one constraint and thus the degrees of freedom are $n-1$.

χ^2 Distribution

The χ^2 distribution is commonly used for a number of statistical tests. It has a single parameter, ν, indicating the degrees of freedom. Although σ^2 is typically unknown, it can be shown that $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$.

Chi-Square Distribution

Legend:
- df = 2
- df = 3
- df = 4
- df = 8
- df = 10

Student's t Distribution

The ratio of a standard normal and the square root of a χ^2 with n degrees of freedom follows a Student's t distribution with $n-1$ degrees of freedom. With algebra, we can show that this is the same as the ratio used in CLT along with the distribution of $(n-1)S^2/\sigma^2$

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

Note that the accuracy of this statement depends upon the degree to which $\overline{X}$ is normally distributed

Student t vs Normal

As n increases, the Student t distribution approaches normality

Normal vs t Distribution
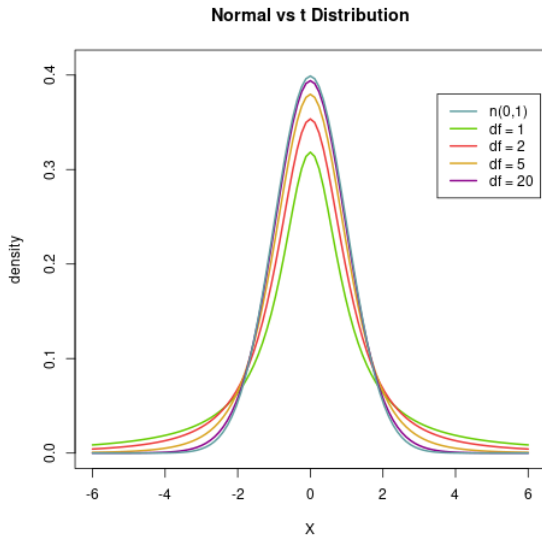
Types of t-tests

t-tests can be classified as either *unpaired* (*independent*) tests or *paired*, where our null will be stated in terms of $H_0 : \mu_1 = \mu_2$, or equivalently, $H_0 : \mu_1 - \mu_2 = 0$

The unpaired test simply involves comparing a value (typically a mean) between two groups that are otherwise unrelated, for example an experimental group and a control group.

The paired test on the other hand involves tests on the same group at different times i.e., before and after treatment. Because we are looking at the same subjects at multiple times, we have reason to suspect that their responses will be correlated

Note that the paired test carries an extra assumption: more assumptions equates to greater power

Unpaired Tests

The unpaired test carries with it a number of assumptions, some more important than others

- We have two groups of independent samples of size n_1 and n_2 respectively
- Each of these samples follows a normal distribution, $X_i \sim N(\mu_i, \sigma_i^2)$
- The variances in each of these groups is equal, i.e., $\sigma_1^2 = \sigma_2^2 = \sigma^2$. This assumption is known as *homoscedasticity*. (The case in which the variances *are not* considered equal is known as *heteroscedasticity*)
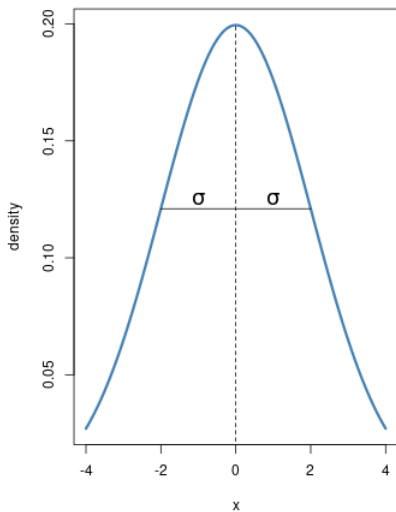
Unpaired Variance

Each of our tests must have some measure of variance for use in the t statistic. Here, as we assume the variances are equal between groups, we simply pool them together

$$S_p^2 = \frac{\sum_{i=1}^{n_1}(X_{i1} - \overline{X}_1)^2 + \sum_{j=1}^{n_2}(X_{j2} - \overline{X}_2)^2}{n_1 + n_2 - 2}$$
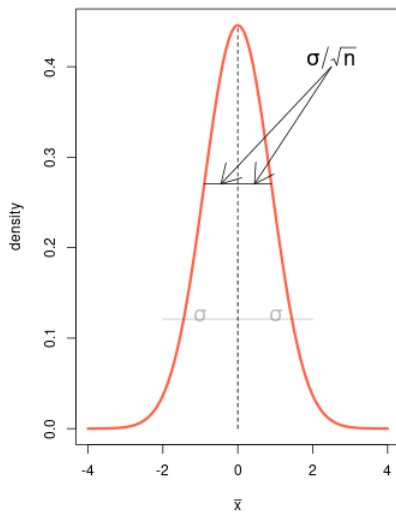
Note here that the degrees of freedom is $n_1 + n_2 - 2$. The -2 comes from the estimation of $\overline{X}_1$ and $\overline{X}_2$

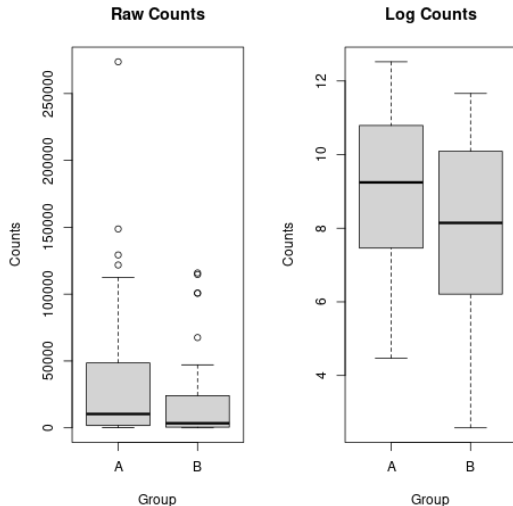# Standard deviation vs Standard error

Unpaired *t*-test

Using the same algebra trick as before, we end up with the following derivation, where the standard error or *precision* of our estimate of $\overline{X}_1 - \overline{X}_2$ is in the denominator

.

$$\frac{\overline{X}_1 - \overline{X}_2}{S_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n-2)$$

Unpaired Example

Here, we consider counts of *Step mutans* in children in two after school programs, with 46 in each program

In R

```
> t.test(Count ~ Group, data = strep_log,
         var.equal = TRUE)

        Two Sample t-test

data:  Count by Group
t = 1.92, df = 90, p-value = 0.058
alternative hypothesis: true difference in means is
not equal to 0
95 percent confidence interval:
 -0.033096  1.895705
sample estimates:
mean in group A mean in group B
        8.9728          8.0415
```
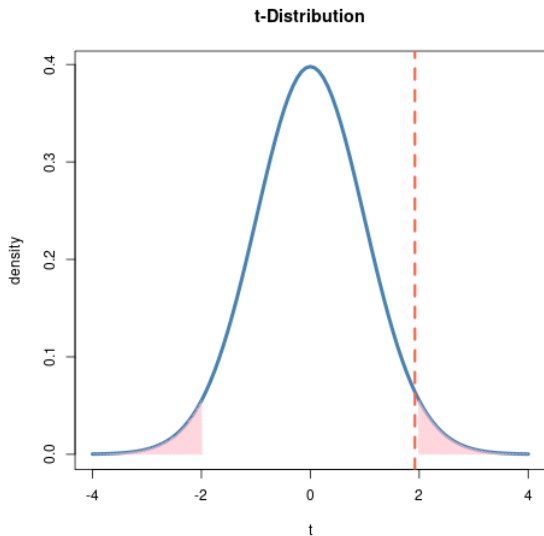
Visualize t-statistic

t-Distribution

Depending on the situation, there are a number of subtle differences in computing the unpaired t-test. The most common (and most widely used) is the *Welch's t-test*, which does not require the assumption that the variances are equal.

Fortunately, we can rely on R to compute the formulas for us

```
t.test(x, y = NULL,
       alternative = c("two.sided", "less", "greater"),
       mu = 0, paired = FALSE, var.equal = FALSE,
       conf.level = 0.95, ...)
```

Paired *t*-test

Often, we find ourselves working with the same observational group under separate conditions, that is, a case in which each subject is able to serve as their own control

The primary consequence of this is the reduction of total variance which (because of correlation), in turn, increases the precision of our estimate, where *n* is the total number of pairs (half the number of observations)

$$\frac{\overline{X}_1 - \overline{X}_2}{S_{diff}/\sqrt{n}} \sim t(n-1)$$

S_{diff}

This is more clear when we consider that, instead of looking at $n_1 + n_2$ observations, we are instead considering the *total number of differences*. As we are assuming this is paired, $n_1 = n_2 = n$. Hence, the total number of differences is n

The standard deviation is just the standard deviation of the differences. Let $d_i = X_{1i} - X_{2i}$, with $\overline{d}$ being the mean value of the differences. Then

$$S_{diff}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (d_i - \overline{d})^2$$

Example

Consider a case studying the use of an SSRI, with "well-being" measured on a Likert-type scale with scores ranging from 0-20 (low to high).

Subject	preMood	postMood	diff
1	3	5	2
2	0	1	1
3	6	5	-1
4	7	7	0
5	4	10	6
6	3	9	6
7	2	7	5
8	1	11	10
9	4	8	4

Example cont.

```
> t.test(post, pre, paired = TRUE)

        Paired t-test

data:  post and pre
t = 3.14, df = 8, p-value = 0.014
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.97633 6.35700
sample estimates:
mean of the differences
                 3.6667


> t.test(diff)

        One Sample t-test

data:  diff
t = 3.14, df = 8, p-value = 0.014
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.97633 6.35700
sample estimates:
mean of x
   3.6667
```

Example cont.

```
> t.test(post, pre, paired = TRUE)

        Paired t-test

data:  post and pre
t = 3.14, df = 8, p-value = 0.014
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.97633 6.35700
sample estimates:
mean of the differences
                  3.6667


> t.test(post, pre, var.equal = TRUE)

        Two Sample t-test

data:  post and pre
t = 2.91, df = 16, p-value = 0.01
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.99918 6.33416
sample estimates:
mean of x mean of y
   7.0000    3.3333
```