Student t-test

Collin Nolte

February 22, 2022

Review and Preview

What we did last time:

- Type I and Type II Errors
- Sensitivity and Specificity
- Power and effect size

What we do today:

- Finish up power
- t-tests
- paired/unpaired tests

Test Result	True State of Nature	
	H_0 True	H_0 False
Fail to reject H_0	Correct $(1 - \alpha)$	Incorrect Type II Error (β)
Reject H_0	Incorrect Type I Error (α)	Correct $(1 - \beta)$

- Type I error $= P(\text{Reject } H_0 | H_0 \text{ true}) = $ false alarm
- Type II error $= P(\text{Fail to reject } H_0 | H_A \text{ true}) = $ missed opportunity

t-distribution

Until now, we have used the CLT to approximate a normal distribution with

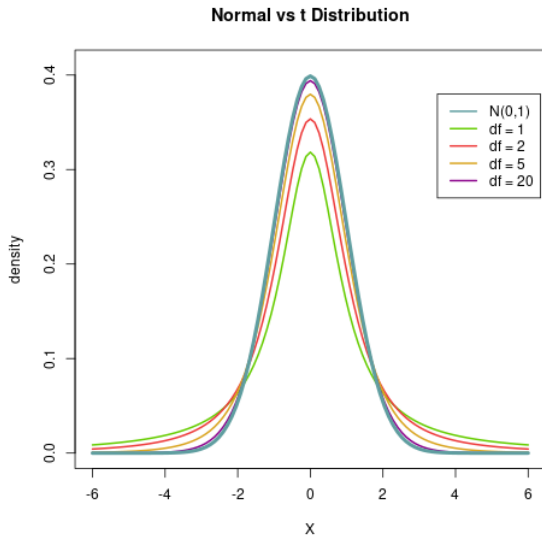$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

However, this assumes that we know the value of σ. More commonly, this value is estimated from the sample by $\hat{\sigma}$. Using this estimated value instead, we end up with a t distribution,

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

where n is the size of the sample

Student t vs Normal

As n increases, the Student t distribution approaches normality

Normal vs t Distribution

t-statistic

Functionally, the t-statistic works just as our z-statistic did, giving us a sampling distribution with which to assess our observation against a null hypothesis.

In particular, we are going to consider the following cases in which a t might be used

1. One sample test
2. Two sample test
   - paired t-test
   - unpaired t-test
     + equal size and variance
     + not equal size or variance

The one sample t-test works virtually identically to each of the examples we have covered thus far

As was the case when considering the CLT, our random variable X doesn't have to be normally distributed itself for the t-statistic to be valid, though more extreme or skewed distributions will typically require a larger sample for the approximation to be accurate

The mean value of a t-distribution is 0 (why?), and the variance is equal to $\frac{\nu}{\nu-2}$, where $\nu = n - 1$

Two sample t-test

Two sample t-tests start to get a bit more interesting. In these situations, we are typically interested in determining if there is a difference in mean between the two groups. That is, we typically have a null of the form $H_0 : \mu_1 = \mu_2$ or $H_0 : \mu_1 - \mu_2 = 0$

Unpaired t-tests involve comparing two groups that are otherwise unrelated. Each group may have a different underlying variance, and even the number of observations in each group may be different

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

Unpaired t-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, with $X_i \sim N(\mu_i, \sigma_i^2)$

- The variances between these groups are equal, i.e., $\sigma_1 = \sigma_2$. 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. When we assume that the population (not necessarily sample) variance are equal between groups, we simply pool them together:
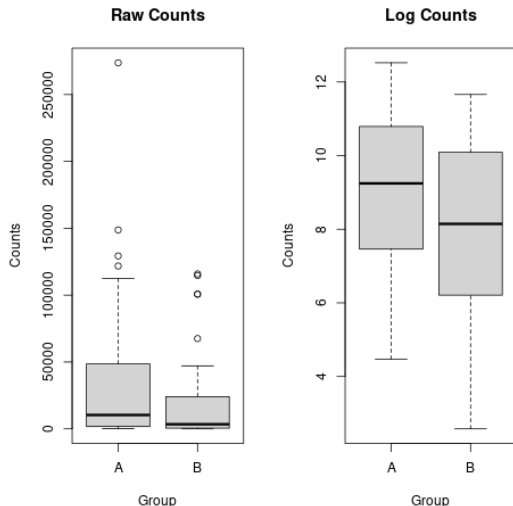
$$S_p^2 = \frac{\sum_{i=1}^{n_1}(X_{i1} - \overline{X}_1)^2 + \sum_{j=1}^{n_2}(X_{j2} - \overline{X}_2)^2}{n_1 + n_2 - 2}$$

This gives us the following construction for the t-statistic

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

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
```
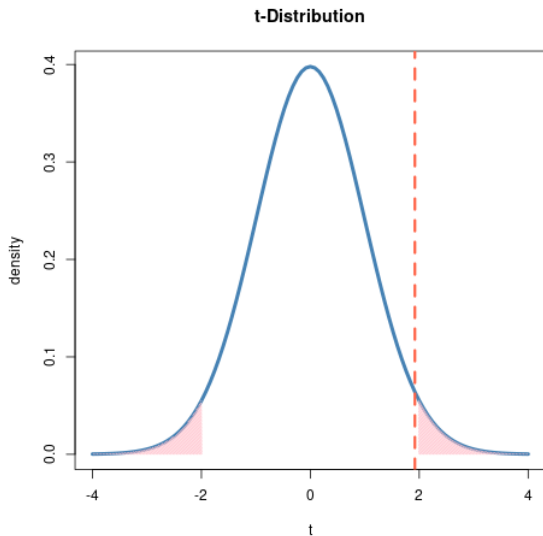
Visualize *t*-statistic

Paired *t*-tests

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

The primary consequence of this is ultimately a reduction in variance which, in turn, increases the precision of our estimate. A second consequence is the impact on sample size – n subjects are able to give us a total of $2n$ observations

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)

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

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

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

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
```

Review

- *t*-statistic uses estimate of variance rather than (unknown) true variance
- Results in normal distribution with fatter tails
- Can be paired, unpaired, or generated from a single collection of observations