Stats 135, Fall 2019
Lecture 21, Monday, 10/21/2019

CLASS ANNOUNCEMENTS:
There is no lecture 20.

The remaining 2 quizzes will be pushed back by 1 week. Quiz 3 will be on November 8, and Quiz 4 will be November 22.

1 Review

Last time, we compared the means of independent normal populations that have the same variance. That is,
We introduced "pooled sample variance" in the case that we need to approximate σ^2 via s_p^2.

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

Then we have a 2-sample t-test where

$$t = \frac{(\overline{X} - \overline{Y}) - (\mu_X - \mu_Y)}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

follows a t distribution with $m+n-2$ degrees of freedom.
Then under these conditions, a $100(1-\alpha)\%$ confidence interval for $\mu_X - \mu_Y$ is:

$$(\overline{X} - \overline{Y}) \pm t_{m+n-2}(\alpha/2)s_{\overline{X}-\overline{Y}}.$$

2 Upcoming Topics

We'll be considering "normal theory", which is looking at the **parametric** t-test, comparing the means of two distributions.

We'll also consider nonparametric tests, comparing the means of 2 continuous populations (§11.2, 11.3).

Then we'll look at Chi square tests, testing whether 2 or more categorical variables have the same distribution.

Also, we'll look at ANOVA (Normal theory F-test), comparing the means of 2 or more normal populations. §12.2

Then we'll look at Regression (Ch 14), and then finally...

Bayesian statistics!

3 Example: Hypothesis Testing and the Confidence Interval

Let's verify that the two sample t-test at the significance level α of

$$H_0 : \mu_X - \mu_Y = 0$$
$$H_1 : \mu_X - \mu_Y \neq 0$$

(This is in the written notes; Lucas goes over this very briefly in the lecture).

4 §11.2

We have assumed before that X, Y both have the same variance. Now if we remove (relax) this assumption, then we'll have a new unbiased estimator that we're going to use. Take the sample variance, divided by the sample size:

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

is a natural unbiased estimator of $\mathrm{Var}(\overline{X} - \overline{Y})$. We check that the expectation of the sample variance over X and that over Y is:

$$\begin{aligned}
\mathbb{E}\left(\frac{s_X^2}{n} + \frac{s_Y^2}{m}\right)^2 &= \frac{\mathbb{E}(S_X^2)}{n} + \frac{\mathbb{E}(S_Y^2)}{m} \\
&= \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m} \\
&= \mathrm{Var}(\overline{X}) + \mathrm{Var}(\overline{Y}) \\
&= \mathrm{Var}(\overline{X} - \overline{Y}) \quad \text{(assuming independence between) } X, Y
\end{aligned}$$

So this implies

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

is an unbiased estimator of $\mathrm{Var}(\overline{X} - \overline{Y})$. But we may ask, what is the value of

$$\frac{\overline{X} - \overline{Y} - 0}{\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}}?$$

Lucas skips to note that we have the following formula for an APPROXIMATION (approximately a t distribution) with degrees of freedom (round this to the nearest integer):

$$\mathrm{df} = \frac{[s_X^2/n) + (s_Y^2/m)]^2}{\frac{(s_X^2/n)^2}{n-1} + \frac{(s_Y^2/m)^2}{m-1}}$$

Lucas notes that we'll essentially program this into R, and nothing that we need to memorize (and can put on our cheat-sheet). What we'll get is a t distribution with this degrees of freedom.

5 Example

50 rats were randomly divided into 2 groups of 25 each. The rats in one group were given steroids. They were timed running a maze. The rats in the steroid group had an average time of 10 seconds, with a SD of 2 seconds. The rats in the steroid group had an average time of 9 seconds and SD = 5.

(a) What is the SE of differences of the 2 averages?

We have:

$$s_{\overline{X}_T - \overline{X}_C} = \sqrt{\frac{s_T^2}{n_T} + \frac{s_C^2}{n_C}} = 1.08.$$

(b) Calculate a test statistic for testing the null hypothesis that steroids make no difference.

Take $H_0 : \mu_T = \mu_C$ and $H_1 : \mu_T \neq \mu_C$. Then
(Lucas skips ahead)

(c) What is the p-value of the test? (Using the formula for df of t-test, we find that we have approximately 31 df (degrees of freedom). This is large so we can just do a normal z-test.)

$$\begin{aligned} p &= \mathbb{P}_0(|T| > |-.93|) \\ &= \mathbb{P}_0(T < .93, \text{ or } T > .93) \\ &= 2\Phi(-.93) = .35, \end{aligned}$$

which we find via `pnorm` in R.

6 Remarks

The calculation of power is an important part of planning an experiment in order to determine how large of a sample size we should use.

For simplicity, let's assume that σ is known, and that σ is the same in both populations with $n = m$.

To draw a picture, consider the hypothesis:

$$\begin{aligned} H_0 &: \mu_X - \mu_Y = 0 \\ H_1 &: \mu_X - \mu_Y = \Delta > 0, \end{aligned}$$

with a one-sided alternative. Then fix α as the level of significance.

We have a normal distribution (centered at 0) under the null distribution, and a shifted version for the alternative (the same normal distribution centered at Δ). Then α is the area under the H_0 distribution curve. Notice that we haven't normalized $\overline{X} - \overline{Y}$, and because $m = n$, we have:

$$z(\alpha) \cdot \sqrt{\frac{2}{n}}.$$

as our cutoff.

Then the standard deviation is:

$$SD = \sigma\sqrt{\frac{1}{n} + \frac{1}{n}} = \sigma\sqrt{\frac{2}{n}}.$$

Our rejection region is to the right of our point, and is the area under the H_1 alternative distribution curve. That is, we have:

$$\begin{aligned}
\text{Power} &= \mathbb{P}_1(\overline{X} - \overline{Y} > z(\alpha)\sigma\sqrt{\frac{2}{n}}) \\
&= \mathbb{P}_1\left(\frac{\overline{X} - \overline{Y} - \Delta}{\sigma\sqrt{\frac{2}{n}}} > \frac{z(\alpha)\sigma\sqrt{\frac{2}{n}} - \Delta}{\sigma\sqrt{\frac{2}{n}}}\right) \\
&= \boxed{1 - \Phi\left(z(\alpha) - \frac{\Delta}{\sigma\sqrt{\frac{2}{n}}}\right)}
\end{aligned}$$

Now we may ask, as we individually increase each of $\Delta, \alpha, \sigma, n$, what happens with power?

- $\delta \uparrow \implies$ Power $\uparrow$
- $\alpha \uparrow \implies$ Power $\uparrow$
- $\sigma \uparrow \implies$ Power $\downarrow$
- $n \uparrow \implies$ Power $\uparrow$

For intuition, we recall that Power is defined as $1 - \Phi(x)$, where Φ is the normal CDF.

7 Example of Power Calculation

Suppose n measurements are taken from a treatment group and independently, n are taken from a control group. An SD of a single measurement is $\sigma = 10$ for both groups. How large should n be so that the test

$$H_0 : \mu_X - \mu_Y = 0$$
$$H_1 : \mu_X - \mu_Y = \Delta > 0$$

has power .7 if $\Delta = \mu_X - \mu_Y = 1$ and $\alpha = 0.05$.

Solution. We set our significance level and want to get the desired power level of 0.7. From our formula before, we have that

$$\text{Power} = 1 - \Phi\left(z(\alpha) - \frac{\Delta}{\sigma}\sqrt{\frac{n}{2}}\right)$$

$$\implies \Phi\left(z(\alpha) - \frac{\Delta}{\sigma}\sqrt{\frac{n}{2}}\right) = 1 - \text{Power} = 1 - 0.7 = 0.3,$$

and now $\texttt{qnorm}(.3) = -.524$. This implies

$$z(\alpha) - \frac{\Delta}{\sigma}\sqrt{\frac{n}{2}} = -.524.$$

Then $\alpha := 0.05 \implies z(\alpha) = 1.645$ and $\Delta := 1$ and $\sigma := 10$ give:

$$\sqrt{\frac{n}{2}} = (.524 + 1.645)\frac{10}{1} \implies \boxed{n = 941}$$

□

Lecture ends here.