MATH 2B RECITATION 2/23/12

ALDEN WALKER

1. t-Tests

Again, the notes on the course website are highly recommended.

These tests arise when you have a data sample, and you assume it's normal, but you know neither the mean nor the variance. Suppose you want to estimate the mean. Well, if you did know the standard deviation, then you would use the test statistic $\frac{\bar{X}_i - \mu_0}{\sigma/\sqrt{n}}$, which would be normal, and everything would be good. However, if you don't know the standard deviation, then you use the obvious replacement—the sample standard deviation $s = \sqrt{\frac{\sum_i (X_i - \bar{X})^2}{n-1}}$. Because of the variability of the sample standard deviation, though, that statistic is no longer normal. It has what is called a t-distribution with $n-1$ degrees of freedom.

1.1. Basic.

You just use the statistic

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

Which, under the null hypothesis that $\mu = \mu_0$, will have a t distribution with $n-1$ degrees of freedom.

1.2. 2 Population t-Tests.

This is a similar situation, but here we have two populations which are normal with the same standard deviation and different means. Then to test hypotheses involving the difference of the means, we can use the statistic

$$T = \frac{\bar{X}_i - \bar{Y}_i - (\mu_X - \mu_Y)}{s\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

Where

$$s = \sqrt{\frac{\sum_i (X_i - \bar{X}_i)^2 + \sum_i (\bar{Y}_i - Y_i)^2}{m + n - 2}}$$

Where there are n X_i's and m Y_i's. The statistic T has a t-distribution with $n + m - 2$ degrees of freedom, so you can use a similar method to the above to set up tests. The setup and stuff is virtually identical once you have the distribution.

1.3. Paired t-Tests.

Here you have a set of data pairs, and you want to evaluate the average differences. This does not require the assumption that the values in the pairs are normal, but only that the differences are normal. Here you will use the statistic

$$T = \frac{(\bar{X}_i - \bar{Y}_i)}{\frac{1}{\sqrt{n}}\sqrt{\frac{\sum_i \left((X_i - Y_i) - (\bar{X}_i - \bar{Y}_i)\right)^2}{n-1}}}$$

Which has a t distribution with $n-1$ degrees of freedom. Note this is basically just a one population t-test, applied to the differences.

1.4. Example.

Suppose that I have 10 people, and I do the following experiment. During the first test, participants play a video game for a few minutes and are then tested for reaction time. During the second test, they don't play a video game but are still tested for reaction time. I get the following reaction times that I just made up (each person is a row).

game	no game
0.1	0.27
0.2	0.3
0.15	0.22
0.11	0.28
0.21	0.33
0.18	0.14
0.12	0.22
0.16	0.11
0.13	0.18
0.16	0.13

Does playing a video game improve your reaction time? Let's see.

1.4.1. *2-population test.* Let's do a two-population t-test to compare the means for the "game" population vs the "no game" population at the 0.01 significance level. Our hypotheses are:

$$H_0 : \mu_X = \mu_Y \text{ vs. } H_1 : \mu_X \neq \mu_Y$$

So we set $\mu_X = \mu_Y$ in the 2-population t statistic above, and we have

$$S_p^2 = \frac{\sum_{i=1}^{10}(X_i - \bar{X})^2 + \sum_{i=1}^{10}(Y_i - \bar{Y})^2}{10 + 10 - 2} = 0.00362$$

and

$$T = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{10} + \frac{1}{10}}} = \frac{0.152 - 0.21}{\sqrt{0.00362}\sqrt{1/5}} = 2.153$$

Under the null, T has a t distribution with $10 + 10 - 2 = 18$ degrees of freedom, and our rejection rule is "reject if $|T| \geq C$". We choose C such that $P(|T| \geq C \,|\, \mu_X = \mu_Y) = 0.01$), so $C = t_{18,0.005} = 2.88$. Notice that our statistic is 2.153, so we do *not* reject the null, and we conclude that playing a video game does not lower your reaction time

1.4.2. *Paired test.* Now let's do a paired t-test at the same significance level of 0.01. We will assume the difference between the times for each person are normal, and we want to know the mean (is the mean = 0?). Our hypotheses are $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$, where μ here is the true mean of the *difference* between each pair. Under the null hypothesis, our statistic

$$T = \frac{(\bar{X}_i - \bar{Y}_i)}{\frac{1}{\sqrt{n}}\sqrt{\frac{\sum_i\left((X_i - Y_i) - (\bar{X}_i - \bar{Y}_i)\right)^2}{n-1}}} = -5.4$$

Has a t distribution with 9 degrees of freedom. We reject if $|T| \geq C$, and we choose C so that $P(|T| \geq C \,|\, \mu = 0) = 0.01$. This means that $C = t_{9,0.005} = 3.25$.

In fact, our statistic is not in this interval, so we reject the null; playing a video game does alter your reaction time. Clearly, it reduces the time, and we would also reject the null hypothesis that the mean of the "game - no game" distribution is nonnegative.

Thus, with the two-population t-test, we failed to conclude that playing the game is helpful, while the paired t-test did conclude there was a difference. Of course, if we were less strict, we might have gotten the same conclusion.

There is also a difference in the assumptions you must make: the two-population t-test assumes the populations are normal, while the paired t-test assumes the differences are normal.

2. p-values

In the above example, we rejected using a paired t test but failed to reject using a 2-population t test. This indicates that the paired t test is more powerful, and if you think of some examples, it's clear why it might be more discerning. However, note that if we set, for example, $\alpha = 0.05$, we would have rejected using either test.

Somehow in real life you don't really want to know whether you accept or reject at a certain significance level — you want to know whether you accept or reject for any given significance level. This is where the p-value comes it.

The p-value of a statistic is the propability of observing something as or more extreme under the null hypothesis. For example, with our 2-population t test above, it's the probability of observing something with $|T| \geq 2.153$ under the null hypothesis that $\mu_X = \mu_Y$. Since T has a t distribution, we can just calculate this: it's $p = 0.023$.

Similarly, for the paired t test, we want the probability under the null that $|T| \geq 5.4$, which is $p = 0.00022$. Note that you will reject the null for any significance level α such that $\alpha \geq p$.

3. Confidence Intervals

The idea is that, given some statistic (observations), you want to give an interval such that with probability α, the true value of a parameter lies within the interval.

The way that you do all of these follows the same idea. You somehow write down what you want, or what you know, in terms of probabilities, and then fiddle with it until you get what you want.

3.1. **Basic Example.** Let's give a 90% confidence interval for the mean of a normal distribution, given 20 samples from that distribution with sample mean 10 and variance 5.

Ok: start with what we know. The statistic $\frac{\bar{X}_i - \mu}{s/\sqrt{n}}$, where μ is the true mean and s is the sample standard deviation, has a t distribution with $n-1$ degrees of freedom, so we know that

$$P\left(t_{n-1,0.05} \leq \frac{\bar{X}_i - \mu}{s/\sqrt{n}} \leq t_{n-1,0.95}\right) = 0.9$$

Rearranging, we know that

$$P\left(\bar{X}_i - |t_{n-1,0.05}|\frac{s}{\sqrt{n}} \leq \mu \leq \bar{X}_i + |t_{n-1,0.95}|\frac{s}{\sqrt{n}}\right) = 0.9$$

And thus with probability 0.9, the true mean lies within the interval with endpoints $\bar{X}_i \pm |t_{n-1,0.05}|\frac{s}{\sqrt{n}}$. Note that the t distribution is symmetric.

In our example, then, our 90% interval is $10 \pm |t_{19,0.05}|\frac{\sqrt{5}}{\sqrt{20}} = 10 \pm 0.865$

3.2. **Example.** Consider the distribution with density function $f_\Theta(x) = \frac{1}{2}e^{-|x-\Theta|}$. Note that the mle for Θ is actually the median of $\{X_i\}$, because the function to maximize (after taking logs) is $-\sum_i |X_i - \Theta|$.

Let's try to get a 0.95% confidence interval for Θ. Let $\hat{\Theta}$ be the mle (the median). One thing to note is that if X has the distribution f_Θ, then $X - \Theta$ has the distribution f_0.

We want to find C such that the probability that Θ is smaller than $\hat{\Theta} + C$ is $\sqrt{0.95}$. We want this because everything is symmetric and thus the probability that Θ is larger than $\hat{\Theta} - C$ will also be $\sqrt{0.95}$, so the probability of both happening will be 0.95.

Note that $P(-C \leq \hat{\Theta} - \Theta \leq C)$ does not depend on the value of Θ due to the fact above about translation. We therefore have that $P(\Theta \leq \hat{\Theta} + C)$ does not depend on Θ, and this is a probability that we can get a handle on, in the following way:

$$P(\Theta \leq \hat{\Theta} + C) = P(\leq (n/2) \text{ successes in } n \text{ trials})$$

Where a success is defined as having $\Theta \geq X_i + C$, and thus success on any trial has probability $P(X_i \leq \Theta - C) = \int_{-\infty}^{-C} (1/2)e^{-|x|}dx = (1/2)e^{-C}$. The probability of $\geq n/2$ successes in n trials is then Binomial, and has probability

$$\sum_{k=0}^{n/2} \binom{n}{k} \left(\frac{1}{2}e^{-C}\right)^k (1 - \frac{1}{2}e^{-C})^{n-k}$$

We then just find C which has this sum equal to $\sqrt{0.95}$