

Stat 324: Lecture 18

Two Sample Tests

Moo K. Chung
mchung@stat.wisc.edu

April 12, 2005

Concepts

1. *Example 1.* A study was conducted to compare the reaction times of men and women to a stimulus. 12 men and 9 women were employed in the experiment. Let X_i be the reaction time of the i -th man and Y_j be the reaction time of the j -th woman. There is a reason to believe that $X_i \sim N(\mu_X, 0.18)$ and $Y_j \sim N(\mu_Y, 0.14)$. The following measurements were taken: $\bar{x} = 3.6$ seconds, $\bar{y} = 3.8$ seconds. Is there any difference between the reaction times of men and women? Test this at $\alpha = 0.1$.

```
> z=(3.6-3.8)/sqrt(0.18/12+0.14/9)
> z
[1] -1.144155
```

2. Let $X_1, \dots, X_n$ and $Y_1, \dots, Y_m$ be two independent samples with mean μ_X, μ_Y and variance σ_X^2, σ_Y^2 respectively.

$$\mathbb{E}(\bar{X} - \bar{Y}) = \mu_X - \mu_Y.$$

$$\text{Var}(\bar{X} - \bar{Y}) = \sigma_X^2/n + \sigma_Y^2/m.$$

Let $X_1, \dots, X_n$ and $Y_1, \dots, Y_m$ be two independent samples with $X_i \sim N(\mu_X, \sigma_X^2)$ and $Y_j \sim N(\mu_Y, \sigma_Y^2)$. Then

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_X^2/n + \sigma_Y^2/m}} \sim N(0, 1).$$

3. *Example 2.* A study was conducted to compare the reaction times of men and women to a stimulus. 50 men and 65 women were employed in the experiment. Let X_i be the reaction time of the i -th man and Y_j be the reaction time of the j -th woman. The following measurements were taken: $\bar{x} = 3.6$ seconds, $\bar{y} = 3.8$ seconds, $s_X^2 = 0.18$, $s_Y^2 = 0.14$. Is there any difference between the reaction times of men and women?

```
> z=(3.6-3.8)/sqrt(0.18/50+0.14/65)
> z
[1] -2.63664
> qnorm(0.05)
[1] -1.64
```

What is the main reason you are getting the different results in Example 1 and 2? Increasing the sample size enable us to distinguish the small difference that was difficult to do so in the small sample.

4. Assume n and m to be large. The test statistic for testing

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

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{s_X^2/n + s_Y^2/m}} \sim N(0, 1).$$

5. *Example 3.* A study was conducted to compare the weights of cats and dogs.

Weights of cats: 20, 21, 35, 13, 21, 10.

Weights of dogs: 31, 10, 20, 40.

Assuming normality and equal variance for both cats and dogs, determine if there is any difference between the weights of cats and dogs at $\alpha = 0.1$.

```
> x<-c(20, 21, 35, 13, 21, 10)
> y<-c(31, 10, 20, 40)
> Sp<-sqrt((5*var(x)+3*var(y))/8)
[1] 10.52824
> t=(mean(x)-mean(y))/(Sp*sqrt(1/5+1/3))
> t
[1] -0.6827026
> qt(0.05, 8)
[1] -1.859548
```

6. Let $X_1, \dots, X_n$ and $Y_1, \dots, Y_m$ be two independent samples from normal distributions with the same population variance, i.e.

$$X_i \sim N(\mu_X, \sigma^2) \text{ and } Y_j \sim N(\mu_Y, \sigma^2).$$

The pooled estimator of σ^2 is

$$S_p^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}.$$

It can be shown that $\mathbb{E}S_p^2 = \sigma^2$. The proper test statistic for testing

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

is

$$T = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{S_p \sqrt{1/n + 1/m}} \sim t_{n+m-2}.$$

Reject H_0 if $|T| > t_{\alpha/2, n+m-2}$.

Note. R package has a built-in command `t.test` that can be used to perform two sample t tests.

Assigned problems. Exercise 9.7., Exercise 9.22 under the assumption of equal variance. HW 5 due April 14th. Solve the assigned problems in Lecture 15-18.