

Stat 312: Lecture 16

Two sample test

Moo K. Chung
mchung@stat.wisc.edu

November 4, 2004

- 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.

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

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

- 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 we can show that

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

Proof. See lecture note.

- Based on the above statistical assumptions on X_i and Y_j , we are interested in testing

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

Example 1. 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?

```
> z=(3.6-3.8)/sqrt(0.18/12+0.14/9)
> z
[1] -1.144155
> qnorm(0.025)
[1] -1.959964
```

- Let $X_1, \dots, X_n$ and $Y_1, \dots, Y_m$ be two independent samples with mean μ_X , μ_Y respectively.

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

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

is

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

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
```

What is the main reason you are getting the different results in Example 1 and 2?

Review Problems. Determine the P -values in Example 1 and 2. Example 9.1., 9.4.