

## INFERENCE FOR ONE-WAY ANOVA

To test equality of means for different treatments/levels, we can use the null hypothesis

$$H_0: \mu_1 = \mu_2 = \dots = \mu_v$$

Rephrase:

1. In terms of effects: \_\_\_\_\_
2. In terms of differences of effects: \_\_\_\_\_
3. In terms of contrasts  $\tau_i - \bar{\tau}$ , where  $\bar{\tau} = \frac{1}{v} \sum_{i=1}^v \tau_i$ : \_\_\_\_\_

The *treatment degrees of freedom* is the minimum number of equations needed to state the null hypothesis, in other words \_\_\_\_\_.

Alternate hypothesis:  $H_a$ : \_\_\_\_\_

Idea of the test: Compare ssE under the *full* model (with all parameters) with the error sum of squares  $ssE_0$  under the *reduced* model -- i.e., the one assuming  $H_0$  is true.

To calculate  $ssE_0$ : If  $H_0$  is true, let  $\tau$  be the common value of the  $\tau_i$ 's. Then the reduced model is

- $Y_{it} = \mu + \tau + \varepsilon_{it}^0$
- $\varepsilon_{it}^0 \sim N(0, \sigma^2)$
- the  $\varepsilon_{it}^0$ 's are independent,

where  $\varepsilon_{it}^0$  denotes the  $it^{\text{th}}$  error in the reduced model.

To find  $ssE_0$ , we use least squares to minimize  $g(m) = \sum_{i=1}^v \sum_{t=1}^{r_i} (y_{it} - m)^2$ :

$$g'(m) = \sum_{i=1}^v \sum_{t=1}^{r_i} 2(-1)(y_{it} - m) = 0,$$

which yields estimate  $\bar{y}_{..}$  for  $\mu + \tau$  -- that is, the least squares estimate of  $\mu + \tau$  is  $(\mu + \tau)^\wedge = \bar{y}_{..}$ . (By abuse of notation, we call this  $\hat{\mu} + \hat{\tau}$ ). So

$$ssE_0 = \sum_{i=1}^v \sum_{t=1}^{r_i} (y_{it} - \bar{y}_{..})^2,$$

which can be shown (proof might be homework) to equal  $\sum_{i=1}^v \sum_{t=1}^{r_i} y_{it}^2 - n(\bar{y}_{..})^2$

Note that  $ssE$  and  $ssE_0$  can be considered as minimizing the same expression, but over different sets:  $ssE$  minimizes  $\sum_{i=1}^v \sum_{t=1}^{r_i} (y_{it} - m - t_i)^2$  over the set of all  $v + 1$ -tuples

$(m, t_1, t_2, \dots, t_v)$ , whereas  $ssE_0$  can be considered as minimizing the same expression over the subset where all  $t_i$ 's are zero. Thus  $ssE_0$  must be at least as large as  $ssE$ :  $ssE_0 \geq ssE$ .

However, if  $H_0$  is true, then  $ssE$  and  $ssE_0$  should be about the same. This suggests the idea of using the ratio  $(ssE_0 - ssE)/ssE$  as a test statistic for the null hypothesis: If  $H_0$  is true, this ratio should be small; so an unusually large ratio would be reason to reject the null hypothesis.

The difference  $ssE_0 - ssE$  is called the *sum of squares for treatment*, or *treatment sum of squares*, denoted  $ssT$ . Using the alternate expressions for  $ssE_0$  and  $ssE$ , we have:

$$\begin{aligned} ssT = ssE_0 - ssE &= \sum_{i=1}^v \sum_{t=1}^{r_i} y_{it}^2 - n(\bar{y}_{..})^2 - \left( \sum_{i=1}^v \sum_{t=1}^{r_i} y_{it}^2 - \sum_{i=1}^v r_i (\bar{y}_{i\cdot})^2 \right) \\ &= \sum_{i=1}^v r_i (\bar{y}_{i\cdot})^2 - n(\bar{y}_{..})^2 \\ &= \sum_{i=1}^v \frac{(y_{i\cdot})^2}{r_i} - \frac{(y_{..})^2}{n} \quad (\text{using definitions}) \\ &= \sum_{i=1}^v r_i (\bar{y}_{i\cdot} - \bar{y}_{..})^2 \quad (\text{possible homework}) \end{aligned}$$

This last expression can be considered as a "between treatments" sum of squares --- we are comparing each treatment sample mean  $\bar{y}_{i\cdot}$  with the grand (overall) mean  $\bar{y}_{..}$ . By

contrast, our denominator,  $ssE = \sum_{i=1}^v \sum_{t=1}^{r_i} (y_{it} - \bar{y}_{i\cdot})^2$  is a "within treatments" sum of squares:

it compares each value with the mean for the treatment group from which the value was obtained.

Using the model assumptions, it can be proved that:

- $ssE/\sigma^2 \sim \chi^2(n - v)$
- If  $H_0$  is true,  $ssT/\sigma^2 \sim \chi^2(v - 1)$
- If  $H_0$  is true, then  $ssT$  and  $ssE$  are independent.

Thus, if  $H_0$  is true,

$$\frac{ssT/\sigma^2(v - 1)}{ssE/\sigma^2(n - v)} \sim F_{v-1, n-v}.$$

Now  $\frac{ssT/\sigma^2(v-1)}{ssE/\sigma^2(n-v)}$  simplifies to  $\frac{ssT/(v-1)}{ssE/(n-v)}$ , which we can calculate from our sample.

We originally wanted to test  $ssT/ssE$ , but  $\frac{ssT/(v-1)}{ssE/(n-v)}$  is just a constant multiple of  $ssT/ssE$ , so is good enough for our purposes:  $\frac{ssT/(v-1)}{ssE/(n-v)}$  will be unusually large exactly when  $ssT/ssE$  is unusually large. Thus, we can use an F test, with test statistic  $\frac{ssT/(v-1)}{ssE/(n-v)}$ , to test our hypothesis.

*Note:* We can look at  $ssT/(v-1)$  and  $ssE/(n-v)$  as we did in the equal-variance, two-sample t-test:  $ssE/(n-v)$  is a pooled estimate of the common variance  $\sigma^2$ , and if  $H_0$  is true, then  $ssT/(v-1)$  can be regarded as an estimate of  $\sigma^2$ .

*Notation:*  $ssT/(v-1)$  is called  $msT$  (*mean square for treatment* or *treatment mean square*) and  $ssE/(n-v)$  is called  $msE$  (*mean square for error* or *error mean square*). So the test statistic is  $F = msT/msE$ .