Ch. 25 – Comparing Several Means

Use t-tools? NO!

→ Reason? *Compound uncertainty*

- In any test, there is uncertainty such that we reject H_0 when it's true, or Type I error. By comparing multiple means and using ONE t-test for each pair, the "overall" Type I error will compound.
- For example, consider 3 means that are equal and each t-test uses $\alpha = 0.05$. Thus, there's a 5% chance to show a difference when there isn't (recall H_0 assumes no diff.). The chance of detecting *at least* one difference among the three means is roughly $1 - 0.95^3 = 0.143$ or 14.3% when the means are EQUAL! (Note: 14.3% is the "overall" Type I error.)
- For 5 means, the "overall" Type I error becomes approximately 40%.

Def'n: ANalysis Of VAriance (ANOVA) is a procedure to test the equality of three or more population means. NOTE: the name of the test refers to comparing different sources of variability; it WILL test differences among means.

Test requires the following assumptions:

1. The samples from different populations are random and independent.
2. The populations are all normally distributed.
3. The populations all have the same standard deviation.

Checking Assumptions:

Assumption #1: Analyze the experimental design.

Assumption #2: Same method as Ch. 20.

Assumption #3: Same method as pooled t-test in Ch. 21-22.

Notation:

- y_{ij} = observation for i^{th} subject in j^{th} group
- $j = 1, \ldots, k$ indexes groups
- $i = 1, \ldots, n_j$ indexes subjects within groups
- n_j = # of observations in j^{th} group; $N = \sum n_j$ = total # of observations
- $\bar{y}_j$ and s_j^2 are sample mean and variance for the j^{th} group
- $\bar{\bar{y}}$ = grand mean = mean for combined sample:

$$\bar{\bar{y}} = \frac{1}{N}\sum_j\sum_i y_{ij} = \frac{1}{N}\sum_j n_j \bar{y}_j$$

Statistical model, parameters, hypotheses:

Each observation can be represented by

$$Y_{ij} = \mu + \tau_j + \varepsilon_{ij}$$

where Y_{ij} are independent random observations, μ is the *overall mean*, τ_j is a parameter associated with the j^{th} group called the j^{th} *treatment effect*, and ε_{ij} is a random error.

H_0: $\mu_1 = \ldots = \mu_k$

H_A: the μ_j are not all equal

(OR at least 1 μ_j is different from another)

ANOVA F test statistic:

For sources of variability in the model above, the *ANOVA identity* is

$$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(y_{ij}-\bar{\bar{y}})^2 = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(\bar{y}_j-\bar{\bar{y}})^2 + \sum_{j=1}^{k}\sum_{i=1}^{n_j}(y_{ij}-\bar{y}_j)^2$$

$$SS_Y = SS_T + SS_E$$

where SS_Y is the <u>total sum of squares</u>, SS_T is the <u>treatment sum of squares</u> and SS_E is the <u>error sum of squares</u>. Also, the presence of "squares" suggests a ratio test. Thus, for the above H_0, we have

$$SS_T = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(\bar{y}_j-\bar{\bar{y}})^2 = \sum_j n_j(\bar{y}_j-\bar{y})^2 \qquad \text{(variability between samples)}$$

$$SS_E = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(y_{ij}-\bar{y}_j)^2 = \sum_j (n_j-1)s_j^2 \qquad \text{(variability within samples)}$$

$$F_0 = \frac{MS_T}{MS_E} = \frac{SS_T/(k-1)}{SS_E/(N-k)}$$

Under H_0, the F_0 test statistic has an F-distribution with $df_1 = k - 1$ and $df_2 = N - k$.

Def'n: The <u>*F* distribution</u> has the following properties:

1. It is continuous and skewed to the right.
2. It has two parameters: *df* for the numerator and *df* for the denominator.
3. The units of an *F* distribution are nonnegative.

Reject H_0 when F_0 is large (greater than 10 works for most values of α)

- Rationale:
  - If H_0 is true, then both types of variability are identical. → $F_0 \approx 1$
  - If H_A is true, then MS_T should be larger. → $F_0 > 1$

ANOVA Table:

Source	**df**	**SS**	**MS**	***F***	***P*-value**
Between	$k-1$	SS_T	MS_T	MS_T/MS_E	?
Within	$N-k$	SS_E	MS_E		
Total	$\boldsymbol{N-1}$	$\mathbf{SS_Y}$			

Note that $SS_Y = SS_T + SS_E$; df(total) = df(Between) + df(Within); MS = SS/df

Calculating SS is tedious! More important to understand values and how they relate to other values in the ANOVA table. If you received an incomplete table, you should be able to fill it in.

Ex25.1) Consider a *k*-mean problem. Five observations are gathered for each group. Use the given information in the table below to answer the questions. Assume all populations are normal with some common variance.
(a) What is *k*?
(b) What is the test statistic for the test to determine if any of the *k* groups are different?

Source	**df**	**SS**	**MS**	***F***	***P*-value**
Between					
Within			6		
Total	**54**	**360**			

(filled out in class)

(Additional example shown in class with full hypothesis test.)

<u>Summary</u>
- check assumptions.
- rejecting H_0 does NOT mean all means are different, AT LEAST ONE is.
- not rejecting H_0 finishes the analysis, rejecting H_0 requires subsequent determination of which means are significantly different from the rest.