Comparing Means with an ANOVA F-Test

H$_0$: $\mu_1 = \mu_2 = \ldots = \mu_k$

H$_a$: The population means are not all equal.

***F*-statistic**:

$$F = \frac{\text{Variation among sample means}}{\text{Natural variation within groups}}$$

1

$$F = \frac{\text{Variation among sample means}}{\text{Natural variation within groups}}$$

Variation among sample means is 0 if all k sample means are equal and gets larger the more spread out they are.

If large enough → evidence at least one population mean is different from others → **reject null hypothesis**.

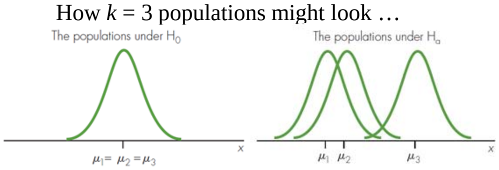
***p*-value** found using an ***F*-distribution**

2

Assumptions for the *F*-Test

- Samples are **independent random samples**.
- Distribution of response variable is a **normal** curve within each population.
- Different populations **may have different means**.
- All populations have **same standard deviation**, σ.

How k = 3 populations might look …



3

Details of One-Way Analysis of Variance

Fundamental concept: the variation among the data values in the overall sample can be separated into:

(1) differences **between group** means

(2) natural variation among observations **within a group**

Total variation = Variation **between** groups + Variation **within** groups

ANOVA Table displays this information.

4

Measuring Variation Between Groups

Sum of squares (between) groups = SS Groups

$$\text{SS Groups} = \sum_{groups} n_i (\bar{y}_i - \bar{y})^2$$

Numerator of F-statistic = **mean square for groups**

$$\text{MS Groups} = \frac{\text{SS Groups}}{k-1}$$

5

Measuring Variation within Groups

Sum of squared errors = SS Error

$$\text{SSE} = \sum_{groups} (n_i - 1) s_i^2$$

Denominator of F-statistic = **mean square error**

$$\text{MSE} = \frac{\text{SSE}}{N-k}$$

Pooled standard deviation: $s_p = \sqrt{\text{MSE}}$

6

Measuring Total Variation

Total sum of squares = SS Total = SSTO

$$\text{SS Total} = \sum_{values} \left(y_{ij} - \bar{y}\right)^2$$

SS Total = SS Groups + SSE

7

General Format of a One-Way ANOVA Table

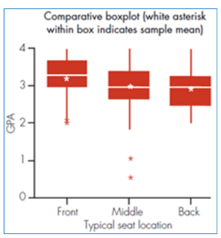
Source	Df	SS	MS	F
Groups/Factor (Between groups)	*K-1*	SSGroups	MSGroups	MSGroups/MSE
Error (Within groups)	*N – k*	SSE	MSE	
Total	*N-1*	SSTO		

8

Seat Location and GPA (R version on website)

Q: Do best students sit in the front of a classroom?

Data on seat location and GPA for *n* = 384 students; 88 sit in front, 218 in middle, 78 in back



Level	N	Mean	StDev
Front	88	3.2029	0.5491
Middle	218	2.9853	0.5577
Back	78	2.9194	0.5105

Students sitting in the front generally have slightly higher GPAs than others.

9

Seat Location and GPA

H_0: $\mu_1 = \mu_2 = \mu_3$

H_a: The three population means are not all equal.

Analysis of Variance for GPA

Source	DF	SS	MS	F	P
Location	2	3.994	1.997	6.69	0.001
Error	381	113.775	0.299		
Total	383	117.769			

The *F*-statistic is 6.69 and the *p*-value is 0.0001.

p-value so small → reject H_0 and conclude there are differences among the population means.

10

Multiple Comparisons

Multiple comparisons: two or more comparisons are made to examine specific pattern of differences among means.

Most common: ***all pairwise comparisons***.

Ways to make inferences about each pair of means:

- **Significance test** to assess if two means significantly differ.
- Find a **Confidence interval** for the difference and if 0 is *not* in the interval, there is a statistically significant difference.

11

Multiple Comparisons

Many statistical tests done → increased risk of making at least one type I error (erroneously rejecting a null hypothesis). Several procedures to control the overall **family type I error rate** or overall **family confidence level.**

- **Family error rate** for set of significance tests is probability of making one or more type I errors when more than one significance test is done.
- **Family confidence level** for procedure used to create a set of confidence intervals is the proportion of times all intervals in set capture their true parameter values.

12

Pairwise Comparisons

- With k means, there are $k(k-1)/2$ comparisons.
- For instance $k = 3$; there are 3 comparisons.
- For $k = 4$, there are 6 comparisons, and so on.
- We need a way to control the overall probability of making a Type 1 error.

13

Two Methods

Fisher's LSD (least significant difference)

- First carry out ANOVA F-test. If *not* significant, stop.
- If significant, compute CI for $(\mu_j - \mu_k)$ for all pairs j, k (next slide). Do *not* adjust confidence level.

Tukey simultaneous confidence intervals (HSD)

- Find multiplier that will give 95% (or other) confidence that the interval with the biggest difference in sample means covers the truth; use that same multiplier for all pairwise intervals. (Need table or computer.)

14

Confidence Intervals for difference in two Population Means in ANOVA

In one-way analysis of variance, a **confidence interval for a difference in population means** is

$$(\bar{y}_j - \bar{y}_k) \pm multiplier\left(s_p \sqrt{\left(\frac{1}{n_j} + \frac{1}{n_k}\right)} \right)$$

where $s_p = \sqrt{\text{MSE}}$ and the *multiplier* is:

Fisher: From a t-distribution with df = $N - k$.
Tukey: From a "studentized range" distribution. (But we let the computer do the work.)

15

Seat Location and GPA

Pairwise Comparison Output:
Tukey: Family confidence level of 0.95
Fisher: 0.95 level for each individual interval

Here, both give same conclusions:
Only 1 interval covers 0, $\mu_{\text{Middle}} - \mu_{\text{Back}}$

Appears population mean GPAs differ for front and middle students and for front and back students.

Tukey 95% Simultaneous Confidence Intervals

Seat = Back subtracted from:

Seat	Lower	Center	Upper
Middle	−0.1028	0.0659	0.2347
Front	0.0846	0.2835	0.4824

Seat = Middle subtracted from:

Seat	Lower	Center	Upper
Front	0.0561	0.2176	0.3791

Fisher 95% Individual Confidence Intervals

Seat = Back subtracted from:

Seat	Lower	Center	Upper
Middle	−0.0759	0.0659	0.2077
Front	0.1164	0.2835	0.4506

Seat = Middle subtracted from:

Seat	Lower	Center	Upper
Front	0.0819	0.2176	0.3533

16

Go over the two examples posted on website

Example 1:

Response Y = GPA

Factor = preferred seat location

Example 2:

Response Y = Days student attends parties per month

Factor = preferred seat location

17

Two-Way ANOVA

Two-way analysis of variance: to examine how two categorical explanatory variables affect the mean of a quantitative response variable.

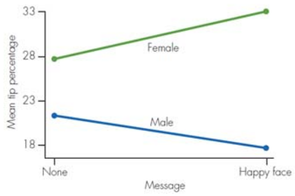
Main effect: overall effect of a single explanatory variable.

Interaction: effect on response variable of one explanatory variable depends upon the specific value or level for the other explanatory variable.

18

Happy Faces and Tips

Q: Does drawing a happy face on the restaurant bill increase average tip to server?

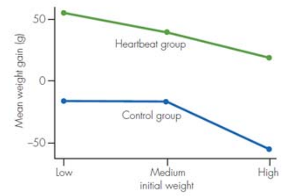


Effect of drawing happy face **depended** on gender. Tips went up for female, down for male. Speculated customers felt happy face not gender appropriate for males.

19

You've Got to Have Heart

Response: Weight gain in Infants
Explanatory: Heartbeat Status (Yes or No)
Initial weight (low, med, high)



Weight gain generally greater for heartbeat group.

There is a main effect for the *heartbeat status*.

Approximately parallel lines => little/no interaction

20

Happy Faces and Tips

Two-way ANOVA:

Three *F*-statistics are computed – one for each main effect and one for interaction.

Source	DF	Adj SS	Adj MS	F	P
Message	1	14.7	14.7	0.13	0.715
Sex	1	2602.0	2602.0	23.69	0.000
Interaction	1	438.7	438.7	3.99	0.049
Error	85	9335.5	109.8		
Total	88	12407.9			

Since interaction effect is significant
→ difficult to interpret the main effect.

21

Two Factor Example on Website

Response Y = GPA

Factor A: Preferred seat location

Factor B: Alcohol consumption in drinks/week, categorized as none (0), moderate (1 to 7), heavy (more than 7)

22