One-Way ANOVA

PSY 5101: Advanced Statistics for Psychological and Behavioral Research 1

One-Way ANOVA: F-test

One-Way ANOVA F-test

1. Situation/hypotheses	One factor $J \geq 2$ independent samples $H_0: \mu_1 = \mu_2 = \ldots = \mu_J$
2. Test statistic	$F = \frac{ns^2_{\bar{X}}}{s^2_{pooled}} = \frac{MS_B}{MS_W} = \frac{\frac{SS_B}{df_B}}{\frac{SS_W}{df_W}}$
3 .Distribution	$F_{J-1, N-J}$
4. Assumptions	1. Populations are normal 2. $\sigma^2_1 = \sigma^2_2 = \ldots \sigma^2_J$ 3. Observations are independent

Basic Theory of ANOVA

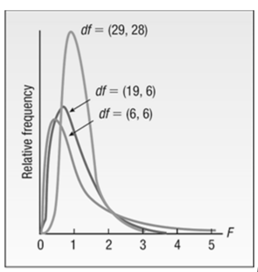
- We compare the amount of variability that is explained by the model (e.g., experimental manipulation) to the error in the model (e.g., individual differences in the outcome variable that are not attributed to the manipulation)
  - This ratio is called the *F*-ratio
- If the model explains a lot more variability than it is unable to explain, then the F-ratio is statistically significant (e.g., the experimental manipulation has had a significant effect on the outcome variable)

Basic Theory of ANOVA

- If the experiment is successful, then the model will explain more variance than it can't
  - SS_M will be greater than SS_R

Characteristics of *F*-Distribution

- There is a "family" of *F* Distributions
- Each member of the family is determined by two parameters: the <u>numerator degrees of freedom</u> and the <u>denominator degrees of freedom</u>
- *F* cannot be negative and it is a continuous distribution
- The *F* distribution is positively skewed
- Its values range from 0 *to* ∞
- As *F* → ∞ the curve approaches the *X*-axis



One-Way ANOVA: Introduction

- What happens when we want to compare two or more groups?
- Now we examine a test statistic that will let us test hypotheses about two or more means, so we can use two or more groups
- The two-sample t-tests could work with only two groups
- The one-way ANOVA uses two or more groups

Why Not Use Multiple t-Tests?

Problem 1: You would need 3 t-tests to test the equality of 3 groups (1 vs. 2, 1 vs. 3, & 2 vs. 3)

Problem 2: Those multiple t-tests are not independent

- The tests involving group 1 (1 vs. 2 & 1 vs. 3) have some overlap.

Problem 3: The probability of a Type I error increases as a function of the number of t-tests

- This idea will be discussed during our section on Multiple Comparison Procedures

Problem 4: You would have multiple tests for one hypothesis

- How many tests would need to be significant in order to reject the null?

Conceptual Similarity to Two-Independent-Samples *t*-Test

- $t = \dfrac{(\overline{X}_1 - \overline{X}_2)}{\sqrt{\left[\dfrac{s^2_1(n_1-1)+s^2_2(n_2-1)}{n_1+n_2-2}\right]\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}$
- What does the numerator of this *t* formula tell you?
  - It tells you about the difference between groups
- What does the denominator of this *t* formula tell you?
  - It tells you about the pooled variance within groups
- What about ANOVA?
  - It is a ratio of two variances
    - Between-groups variance in the numerator
      - It uses between-groups variance because you cannot simply look at the difference scores between three or more groups
    - Within-groups variance in the denominator

Similarities Between Two-Independent-Samples t-Test and One-Way ANOVA

Type of Analysis	**Two-Independent-Samples t-Test**	**One-Way ANOVA**
Information about differences between group means	$\overline{X}_1$ - $\overline{X}_2$	$MS_{Between}$
Information about the variability of scores within groups	$\sqrt{\left[\frac{s^2_1(n_1-1)+s^2_2(n_2-1)}{n_1+n_2-2}\right]\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$	MS_{Within}
Degrees of freedom	df = n_1 + n_2 - 2	Two df terms: $df_{Between}$ = J – 1 df_{Within} = N - J
Test Statistic	$\frac{(\overline{X}_1 - \overline{X}_2)}{\sqrt{\left[\frac{s^2_1(n_1-1)+s^2_2(n_2-1)}{n_1+n_2-2}\right]\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$	$\frac{ns^2_{\overline{X}}}{s^2_{pooled}} = \frac{MS_B}{MS_W} = \frac{\frac{SS_B}{J-1}}{\frac{SS_W}{N-J}}$
Information represented by the test statistic	$\frac{\text{Between} - \text{Group Differences}}{\text{Within} - \text{Group Differences}}$	$\frac{\text{Between} - \text{Group Differences}}{\text{Within} - \text{Group Differences}}$

Why Is It Called "One-Way ANOVA"?

- "One-Way" refers to the number of factors (i.e., variables that classify the subjects into groups)
  - Note: ANOVA can be used for either experimental or non-experimental data
- "ANOVA" is an abbreviation for "Analysis of Variance"
- A One-Way ANOVA is a procedure that tests the effects of one factor (several independent groups) on the means of one continuous outcome variable

One-Way ANOVA: Example

Question: Does smoking impact your thinking?
Groups:
Non-Smokers (NS)
Active Smokers (AS, had just smoked)
Deprived Smokers (DS, not smoked for 3 hours)
The three groups performed several cognitive tasks that ranged from simple to complex
For the complex tasks, there were significant differences between the groups such that the AS group did the worst

One-Way ANOVA: Logic

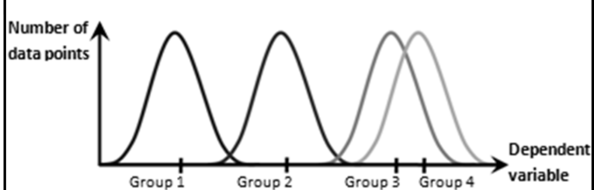
- You obtain two estimates of σ^2
  - One estimate is based on the variance of sample means and the other estimate is based on the variance of observations within the groups
- Both estimate the population variance (σ^2) if the null hypothesis of equal population means is true…but they estimate different quantities if H_0 is false
- With these two estimates of σ^2, you form the F ratio
  - You expect the *F* ratio to be approximately one if H_0 is true but larger than one if H_0 is false
- You reject H_0 if *F* is larger than or equal to a critical value

One-Way ANOVA: Logic

- The ANOVA F-test uses a different logic than $z_{\overline{X}}$, r, or any of the t-tests
  - They were all based on a logic that looked for how far the test statistic was from a middle value of zero
  - If the statistic was far enough away from zero – and in agreement with H_1 – then you rejected H_o
- The ANOVA's logic forms an F-ratio of two sample variances, one based on the group means (Between) and the other based on scores within groups (Within)
  - Total variance can be partitioned into "between-groups" and "within-groups"
  - If H_o of equal population means is true, then both variances should be equal and the average F will be about 1
  - If the population means are not equal, then we expect Between Variance > Within Variance which results in an average F > 1 (this leads us to reject H_o if $F \geq F_{crit}$)

One-Way ANOVA: Logic

- ANOVA will use two estimates of σ^2
- Estimate of σ^2 from the group means
  - $ns^2_{\overline{X}}$ is n (number of observations per group) times the unbiased sample variance of the $\overline{X}$ values (group means)
  - Remember $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$ …solve for σ^2 to get $\sigma^2 = n\sigma^2_{\overline{X}}$
  - $s^2_{\overline{X}}$ is our best estimate of $\sigma^2_{\overline{X}}$
- Estimate of σ^2 based on the variance of observations within groups
  - s^2_{pooled} is the estimate of σ^2 based on pooling the values of s^2_j
    - For equal sample sizes per group, s^2_{pooled} is the average of the s^2_j values



One-Way ANOVA: Logic

- Notation
  - n = # of observations per group
  - J = # of groups
  - N = nJ
- Two sample variances:
  - One variance is based on group means (Between)
    - Compute $s^2_{\bar{X}}$ (i.e., the unbiased variance of the J group means) and multiply it by n. This also is called MS_B, so $ns^2_{\bar{X}} = MS_B$
  - One variance is based on scores within groups (Within)
    - Compute s^2 of observations within each of the J groups, and the average of these J values of s^2 is s^2_{pooled}, also called MS_W
- Form the test statistic:

$$F = \frac{ns^2_{\bar{X}}}{s^2_{pooled}} = \frac{MS_B}{MS_W} = \frac{\frac{SS_B}{J-1}}{\frac{SS_W}{N-J}}$$

-Any unbiased sample variance consists of a sum of squares – such as $\Sigma(X-\bar{X})^2$ – divided by degrees of freedom

$$s^2 = \frac{\Sigma(X-\bar{X})^2}{N-1}$$

-Mean square (MS) is sample variance

-Sum of squares (SS) is a sum of squared deviations

- Hypotheses:
  - $H_o: \mu_1 = \mu_2 = \ldots = \mu_J$
  - H_1: any differences in μ_js

One-Way ANOVA: Logic

	H_o True	H_0 False
$ns^2_{\bar{X}} = MS_B = \frac{SS_B}{df_B}$	Estimates σ^2	Estimates σ^2 + positive quantity (i.e., treatment effect)
$s^2_{pooled} = MS_W = \frac{SS_W}{df_W}$	Estimates σ^2	Estimates σ^2
$F = \frac{ns^2_{\bar{X}}}{s^2_{pooled}} = \frac{MS_B}{MS_W} = \frac{\frac{SS_B}{df_B}}{\frac{SS_W}{df_W}}$	Expect $F \approx 1$	Expect $F > 1$

We reject H_o if $F \geq F_{crit}$ or p-value is $\leq .05$

One-Way ANOVA: Formulas

F (F ratio) $= \frac{ns^2_{\bar{X}}}{s^2_{pooled}} = \frac{MS_B}{MS_W} = \frac{\frac{SS_B}{df_B}}{\frac{SS_W}{df_W}}$

MS_B (Mean Squares Between) $= \frac{SS_B}{df_B}$

SS_B = Sum of Squares Between: Sum of squared deviations for each group mean about the grand mean (i.e., mean of the total sample)

df_B = J-1 *(J = # of groups)*

MS_W (Mean Squares Within) $= \frac{SS_W}{df_W}$

SS_W = Sum of Squares Within: Sum of squared deviations for all observations within each group from that group mean, summed across all groups

df_W = N-J *(n = # participants; N = n*J)*

One-Way ANOVA: Computation

- Compute $SS_{Between}$ and SS_{Within}
- Compute $MS_{Between}$ by dividing $SS_{Between}$ by its df
- Compute MS_{Within} by dividing SS_{Within} by its df
- Compute an F ratio by dividing $MS_{Between}$ by MS_{Within}
- Compare this value for F with the critical value for F based on $df_{Between}$ and df_{Within}

One-Way ANOVA: Computation

- Sum of Squares Between ($SS_{Between}$)

  $SS_{Between} \sum_{i=1}^{J} n_i(\overline{X}_i - \overline{X}_{Grand})^2$

- Sum of Squares Within (SS_{Within})

  $SS_{Within} \sum_{i=1}^{J} SS_i = SS_1 + SS_2 + _{\dots} + SSJ$

- Sum of Squares Total (SS_{Total})

  $SS_{Total} = SS_{Between} + SS_{Within}$

One-Way ANOVA: Factors vs. Levels

- The ANOVA is a general statistical tool, including the one-way ANOVA, the two-way ANOVA, and beyond
- The "one" in "one-way" refers to the number of factors (variables that classify the subjects into groups)
- A one-way layout looks like this:

1	2	3	4

- A two-way layout looks like this:
- Levels are the values of the factors
  - In these examples, the one-way above has four levels, and the two-way has 2 levels of one factor and 3 levels of the second and is called a 2X3 ANOVA

Factor B

Factor A	B_1	B_2	B_3
A_1			
A_2			

One-Way ANOVA: Test Statistic

- Hypotheses: if J=4
  - H_o: $\mu_1 = \mu_2 = \mu_3 = \mu_4$
  - H_1: any differences in μs
- The test statistic is the F-ratio,

$$F = \frac{ns^2_{\bar{X}}}{s^2_{pooled}} = \frac{MS_B}{MS_W} = \frac{\frac{SS_B}{df_B}}{\frac{SS_W}{df_W}}, \text{ where } df_B = J\text{-}1 \text{ and } df_W = N\text{-}J$$

- Example: if SS_B=410, SS_W=630, n=30, and J=4 then df_B=J-1=4-1=3, and df_W=N-J =J(n-1)=4(29)=116, so

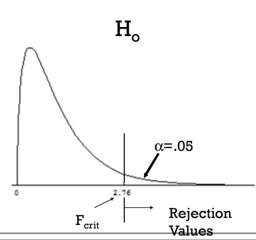
$$F = \frac{\frac{410}{3}}{\frac{630}{116}} = \frac{136.67}{5.43} = 25.16$$

One-Way ANOVA: F Distribution

- The F distribution is a positively skewed distribution with a minimum of zero
- It has two parameters, the df for the numerator variance (df_B) and the df for the denominator variance (df_W)
- The F table of critical values is organized by df_B, df_W, and α (.05 and .01)
- Only upper-tail critical values are given because we expect the F only to get large if H_1 is true

One-Way ANOVA: F Distribution

- Here is a picture of the F distribution with df_B=3 and df_W=60, with the critical value that cuts off α=.05 in the upper tail



One-Way ANOVA: Assumptions

- The one-way ANOVA F statistic is distributed as $F_{J-1,N-J}$ only if all of the assumptions are met
- If any of the assumptions are not met, then F only approximately has this distribution and we need to ask questions about robustness for each assumption
  - Normality: like the two-independent-samples *t*-test, *F* is reasonably robust to non-normality, except for mixed distributions
  - Equal variances: unlike the *t*, *F* is not robust to very unequal variances, even with large and equal sample sizes
  - Independence: like the *t*, *F* is not robust to dependence in the data but we typically meet this assumption

One-Way ANOVA: Unequal Variances

- Unlike the *t*, *F* is not robust to very unequal variances, even with large and equal sample sizes (if $J > 2$)
- For example, for J=4, n=50, if the population variances are in the ratio of 16:1:1:1, then the true α is .088 when α is set at .05
  - Note that .088 is larger than the .06 that we set as an upper boundary on α=.05
  - Also note that the n=50 per group is considerably larger than the n=15 that it took to make the t robust to any ratio in variances
- The F is robust to slightly unequal variances but you do not know the population variances
- This problem of the F's lack of robustness to very unequal variances will be resolved when we get to the next statistical procedures: multiple-comparison procedures

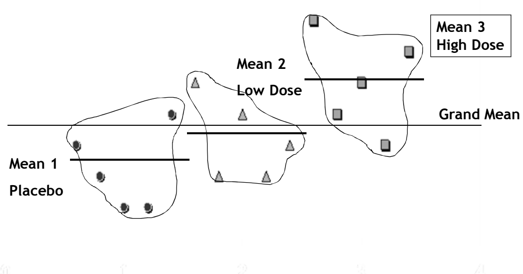
ANOVA Example

- Testing the effects of Viagra on Libido using three groups:
  - Placebo (Sugar Pill)
  - Low Dose Viagra
  - High Dose Viagra
- The Outcome/Dependent Variable (DV) was an objective measure of Libido

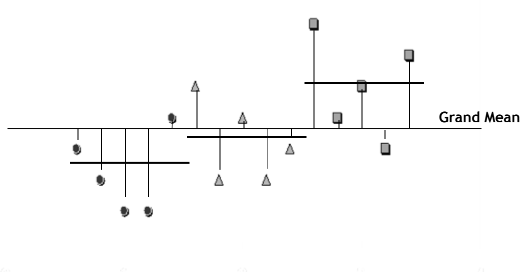
The Data

	Placebo	Low Dose	High Dose
	3	5	7
	2	2	4
	1	4	5
	1	2	3
	4	3	6
$\bar{X}$	2.20	3.20	5.00
s	1.30	1.30	1.58
s^2	1.70	1.70	2.50
	Grand Mean = 3.467	Grand SD = 1.767	Grand Variance = 3.124

The data



Total Sum of Squares (SS_T)



Step 1: Calculate SS_T

$$SS_T = \sum (x_i - \bar{x}_{grand})^2$$

$$s^2 = \frac{SS}{(N-1)}$$

$$SS = s^2(N-1)$$

$$SS_T = s^2_{grand}(N-1)$$

$$SS_T = 3.124(15-1) = 43.74$$

Grand variance

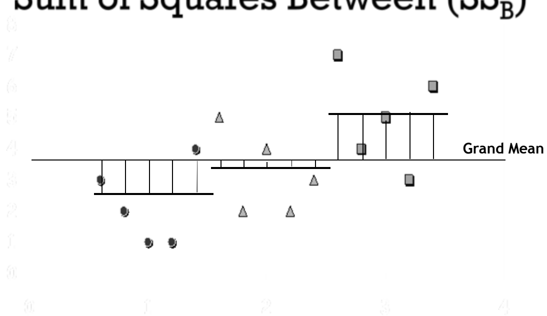
# of participants

Degrees of Freedom (df)

- Degrees of Freedom (*df*) are the number of values that are free to vary
- In general, the *df* are one less than the number of values used to calculate the SS

$$df_T = (N-1) = 15 - 1 = 14$$

Sum of Squares Between (SS_B)



Step 2: Calculate SS_B

$$SS_B = \sum n_i (\bar{x}_i - \bar{x}_{grand})^2$$

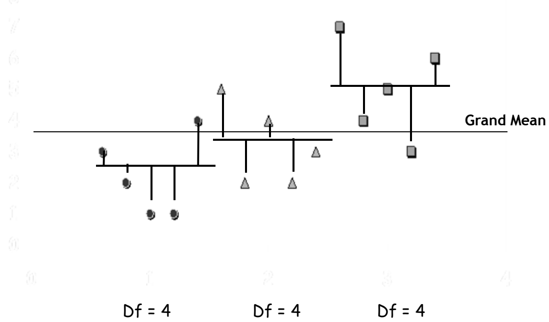
$$\begin{aligned} SS_B &= 5(2.2-3.467)^2 + 5(3.2-3.467)^2 + 5(5.0-3.467)^2 \\ &= 5(-1.267)^2 + 5(-0.267)^2 + 5(1.533)^2 \\ &= 8.025 + 0.355 + 11.755 \\ &= 20.135 \end{aligned}$$

Model Degrees of Freedom

- How many values did we use to calculate SS_B?
  - We used the 3 means

$$df_B = (J-1) = 3-1 = 2$$

Sum of Squares Within (SS_W)



Step 3: Calculate SS_W

$$SS_W = \sum (x_i - \bar{x}_i)^2$$

$$s^2 = \frac{SS}{(N-1)}$$

$$SS = s^2(N-1)$$

$$SS_W = \sum s_i^2 (n_i - 1)$$

$$SS_W = s_{group1}^2 (n_1 - 1) + s_{group2}^2 (n_2 - 1) + s_{group3}^2 (n_3 - 1)$$

Step 3: Calculate SS_W

$$\begin{aligned} SS_W &= s_{group1}^2 (n_1 - 1) + s_{group2}^2 (n_2 - 1) + s_{group3}^2 (n_3 - 1) \\ &= (1.70)(5-1) + (1.70)(5-1) + (2.50)(5-1) \\ &= (1.70 \times 4) + (1.70 \times 4) + (2.50 \times 4) \\ &= 6.8 + 6.8 + 10 \\ &= 23.60 \end{aligned}$$

Residual Degrees of Freedom

- How many values did we use to calculate SS_W?
  - We used the 5 scores for each of the SS for each group

$$\begin{aligned} df_W &= df_{group1} + df_{group2} + df_{group3} \\ &= (n_1 - 1) + (n_2 - 1) + (n_3 - 1) \\ &= (5-1) + (5-1) + (5-1) \\ &= 12 \end{aligned}$$

Double Check

$$SS_T = SS_B + SS_W$$
$$43.74 = 20.14 + 23.60$$
$$43.74 = 43.74$$

Step 4: Calculate the Mean Squared Error

$$MS_B = \frac{SS_B}{df_B} = \frac{20.135}{2} = 10.067$$

$$MS_W = \frac{SS_W}{df_W} = \frac{23.60}{12} = 1.967$$

Step 5: Calculate the *F*-Ratio

$$F = \frac{MS_B}{MS_W}$$

$$F = \frac{MS_B}{MS_W} = \frac{10.067}{1.967} = 5.12$$

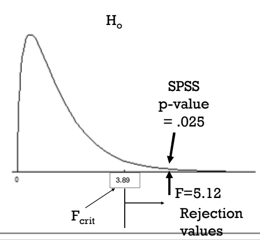
Step 6: Construct a Summary Table

Source	SS	df	MS	*F*
Between	20.14	2	10.067	5.12*
Within	23.60	12	1.967	
Total	43.74	14		

One-Way ANOVA

- Next, get an F_{crit} for df_B=2 and df_W=12. With α=.05 we have F_{crit}=3.89.

Now, reject H_o because $F=5.12>F_{crit}=3.89$.



Or, reject H_o because $p=.025< \alpha=.05$.

One-Way ANOVA

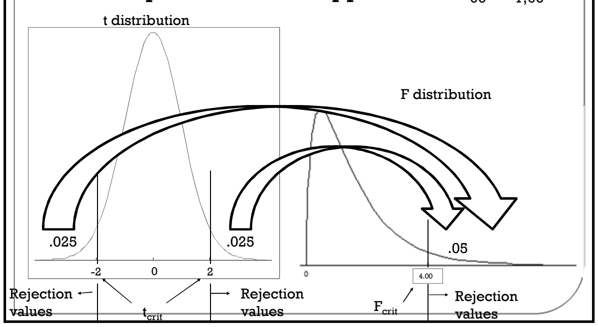
- The ANOVA does not tell us which groups are different
  - Do people in the High Dose group have higher libidos than those in the Low Dose group?
  - Do people in the Low Dose group have higher libidos than those in the Placebo group?
  - Do people in the High Dose group have higher libidos than both those in the Low Dose group and the Placebo group?
- ANOVA tells us that there is "some difference in the means" but not how many differences there are or which means are different
- How do we know which groups are different? We will have to use Multiple Comparison Procedures (we cannot rely on looking at means)

One-Way ANOVA: $t^2=F$

- The final topic for the ANOVA is to show the connection between the two-independent-sample t and the one-way ANOVA F when J=2
- When comparing two groups, either test is fine because they will lead to the same conclusion
- The relationship is: $t^2=F$

One-Way ANOVA: $t^2=F$

- Here is a picture of what happens with $t^2_{60}=F_{1,60}$



One-Way ANOVA: $t^2=F$

Two-Independent-Samples *t*-Test

Group Statistics

	group	N	Mean	Std. Deviation	Std. Error Mean
outcome	control	15	3.4667	1.45733	.37628
	experimental	15	8.9333	1.57963	.40786

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
outcome	Equal variances assumed	.004	.947	-9.851	28	.000	-5.46667	.55492	-6.60337	-4.32996
	Equal variances not assumed			-9.851	27.820	.000	-5.46667	.55492	-6.60370	-4.32963

One-Way ANOVA

ANOVA

outcome

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	224.133	1	224.133	97.047	.000
Within Groups	64.667	28	2.310		
Total	288.800	29			

Effect Size for One-Way ANOVA

- How strong is the actual effect? That is, what proportion of variability in the outcome variable is accounted for by the factor?
- What is needed is an estimate of the magnitude that is relatively independent of sample size
  - Estimates of magnitude or effect size tell us how strongly two or more variables are related or how large the difference is between groups
- Eta squared: $\eta^2 = \frac{df_{between} * F}{(df_{between} * F) + dfw_{ithin}}$