"It is easy to lie with statistics.
It is hard to tell the truth without statistics."

-Andrejs Dunkels

COHEN CHAP 12 ONE-WAY ANOVA

For EDUC/PSY 6600

MOTIVATING EXAMPLES

Dr. Vito randomly assigns 30 individuals to 1 of 3 study groups to evaluate whether one of **2 new approaches** to therapy for adjustment disorders with mixed anxiety and depressed mood are more effective than the **standard approach**. Participants are matched on current levels of anxiety and depressed mood at baseline. Scores from the BAI and BDI are collected after 2 months of therapy.

Dr. Creft wishes to assess differences in oral word fluency **among three groups of participants**: Right hemisphere stroke, left hemisphere stroke, and healthy controls. Scores on the COWAT are collected from 20 participants per group and the means of each group are compared.

DIFFERING RESEARCH DESIGNS

Fixed or random effects

- *Fixed effects design:* Levels of each factor systematically chosen by researcher
  - Individual therapy, group therapy, self-help/no-therapy
- *Random factors design*: Levels of each factor are chosen randomly from a larger subset (rarer)
  - Therapy groups coordinated by: Mrs. May, Mr. June, or Ms. July

Independent (Between-Subjects) or Repeated (Within-Subjects) factors

- *Independent*: Participants randomly allocated to each level of a factor
  - Chapter 12: one-way ANOVA
- *Repeated measures design*: Participants are paired or a dependency exists (multiple observations)
  - *Chapter 15: Repeated measures ANOVA*

Experimental design

- Participants are randomly **assigned** to levels and at least one factor is manipulated
  - Drug A, drug B, or placebo
- Participants are randomly selected from multiple **preexisting populations**
  - Single parent home, two parent home, does not live with parent

Note: If the levels of the Dependent Variable are **highly ordinal or continuous in nature, **regression** or a rank type test will be more powerful than ANOVA, which is appropriate in cases where the groups are more nominal in nature.*

Some variables can be construed as both!!! (e.g. Grade level)… probably want to analyze both ways

ANOVA = "ANALYSIS OF VARIANCE"

ANOVA designs can be used for…

- Experimental research
- Quasi-experimental studies
- Field/observational research

1-way ANOVA (or AOV)

- Also called:
  - Single factor ANOVA
  - *Univariate ANOVA*
  - Simple ANOVA
  - Independent-ANOVA
  - Between-subjects ANOVA

ONE Dependent Variable (DV)

Continuous (interval/ratio)
&
normally distributed

ONE Independent Variable (IV)

Categorical (nominal)
$\geq$ 3 independent samples or groups
Factor with k *levels*
or
groups levels can be chosen
experimentally or occur naturally

Omnibus test for group (mean) differences

- Overall pattern in the data

F-DISTRIBUTION

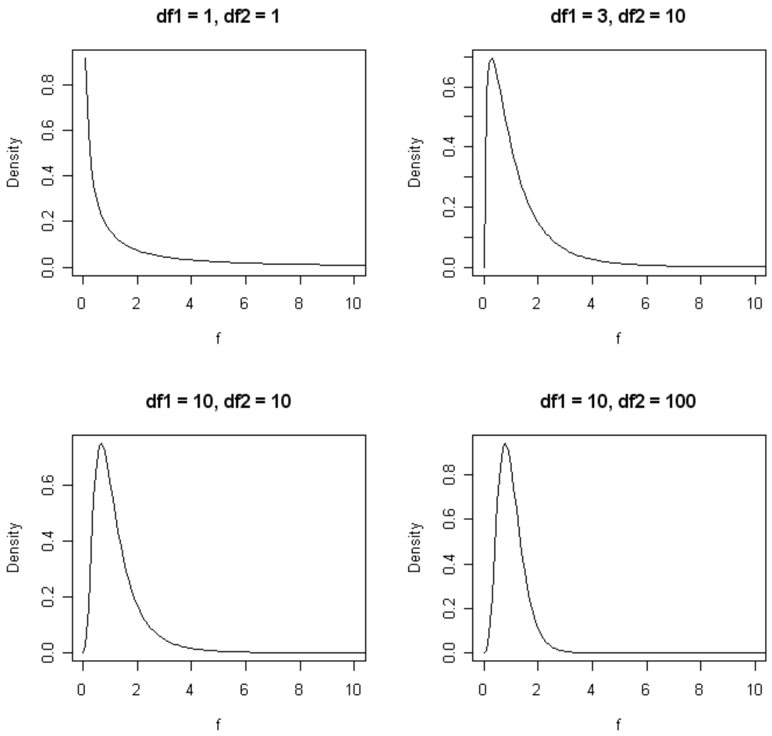
Sir Ronald A. Fisher (1920s) & agricultural experiments

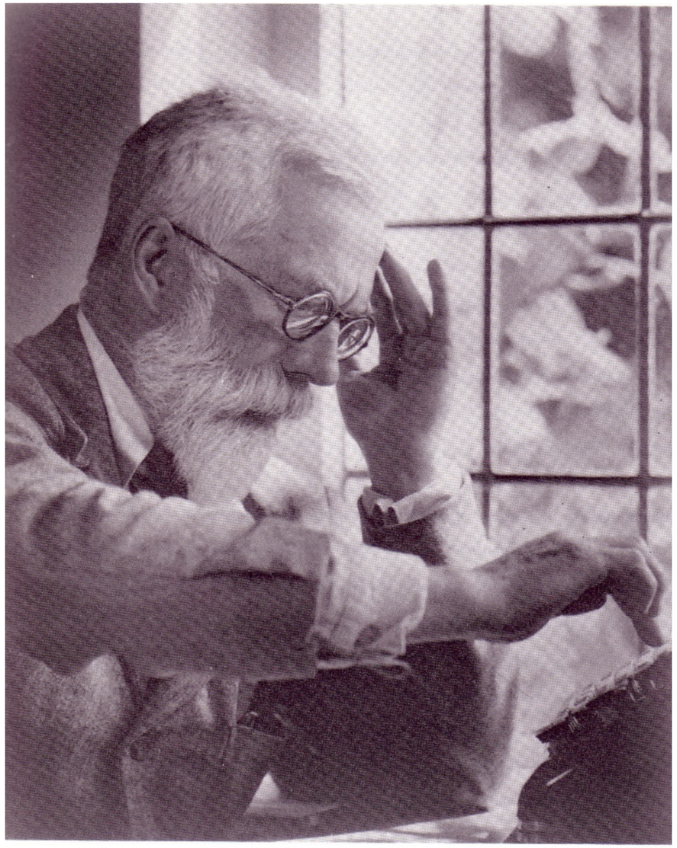
F-distribution
- Continuous theoretical probability distribution
- Probability of ratios of variance between groups to variance within groups

$$F = t^2$$

Positively skewed
- Range: 0 to ∞
- one-tailed
- More "normal" as $N\uparrow$
- Mean ≈1… $M = \frac{df_W}{df_W - 2}$

Family of distributions
- Need 2 df and α to determine F_{crit}
  - df_{Within} and $df_{Between}$ (more later…)





Fisher at his desk calculator at Whittingehame Lodge, 1952

EXAMPLE: NOISE & WORDS MEMORIZED

Group		
A	B	C
8	7	4
10	8	8
9	5	7
10	8	5
9	5	7

Study to determine if noise inhibits learning

$N = 15$

Students randomized to 1 of 3 groups ($n = 5$)

- Group A: No noise (no music, quiet room)
- Group B: Moderate noise (classical music)
- Group C: Extreme noise (rock music)

Participants are given 1 minutes to memorize list of 15 nonsense words

DV = # of correct nonsense words recalled

HYPOTHESES OF ANOVA

<u>Label sets of means and variances as…</u>

- $\mu_1, \mu_2, \mu_3, \ldots, \mu_k$
- $\sigma_1^2, \sigma_2^2, \sigma_3^2, \ldots, \sigma_k^2$

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

H_1: Not H_0

- Many ways to reject H_0
- **NOT H_1: $\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_k$**

Reject H_0: <u>low</u> p(samples drawn from <u>same</u> population)

Retain H_0: <u>high</u> p(samples drawn from <u>same</u> population)

EXAMPLE: NOISE & WORDS MEMORIZED

Group A	B	C
8	7	4
10	8	8
9	5	7
10	8	5
9	5	7

Null Hypothesis:

Alternative Hypothesis:

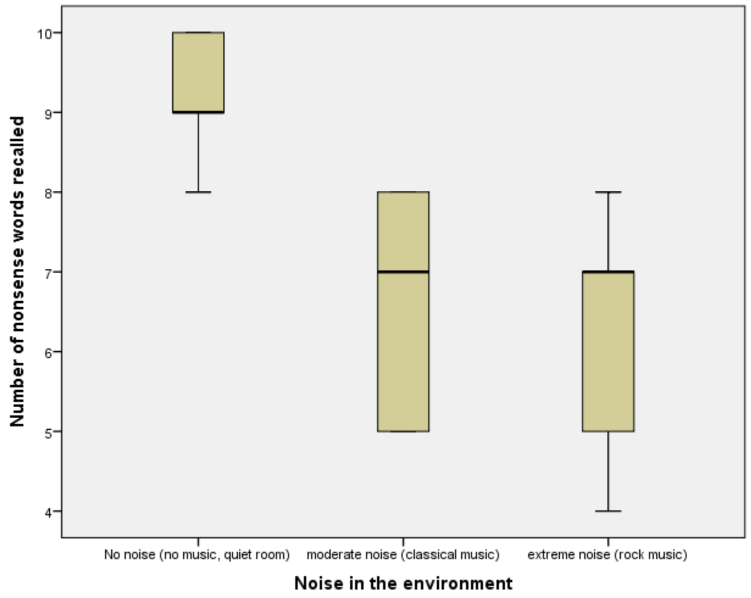
```
SORT CASES by noise.
TEMPORARY.
SPLIT FILE by noise.
FREQUENCIES words
   /FORMAT NOTABLE
   /STATISTICS MEAN STDDEV.
SPLIT FILE off.
```

```
EXAMINE Words BY Noise
  /PLOT=BOXPLOT
  /STATISTICS=NONE.
```

Statistics

Words Number of nonsense words recalled

1 No noise (no music, quiet room)	N	Valid	5
		Missing	0
	Mean		9.20
	Std. Deviation		.837
2 moderate noise (classical music)	N	Valid	5
		Missing	0
	Mean		6.60
	Std. Deviation		1.517
3 extreme noise (rock music)	N	Valid	5
		Missing	0
	Mean		6.20
	Std. Deviation		1.643



LINK: INDEPENDENT SAMPLE "T-TEST" & ANOVA

Same question as before…

- Do group means significantly differ?
- Or…Do mean differences on DV '<u>between</u>' groups EXCEED differences '<u>within</u>' groups?
  - Between-groups differences
    - Differences in DV due to IV (group)
  - Within-groups differences
    - Differences in DV due to pooled random error or variation

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{s_{\bar{X}_1 - \bar{X}_2}} = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Could rewrite as… $t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{s_p^2\left(\frac{2}{n_j}\right)}}$,

Where n_j = sample size for any group j. Then....

$$t^2 = \frac{(\bar{X}_1 - \bar{X}_2)^2}{\frac{2s_p^2}{n_j}} = \frac{n_j(\bar{X}_1 - \bar{X}_2)^2}{2s_p^2} = \frac{n_j\frac{(\bar{X}_1 - \bar{X}_2)^2}{2}}{s_p^2} = F$$

Same analysis approach as before…

$$\boxed{F = t^2}$$

LINK: INDEPENDENT SAMPLE "T-TEST" & ANOVA

2 samples, when $n_1 = n_2$

Numerator: Variation between (among) group means

'Variance' of 2 means multiplied by n_i

Mean Square Between (MS_B) or Mean Square Treatment (MS_T)

$$t^2 = \frac{n_j \dfrac{\left(\bar{X}_1 - \bar{X}_2\right)^2}{2}}{s_p^2} = F$$

$$\boxed{t^2 = F}$$

Denominator: Pooled variation within groups

Pooled variance (s_p^2) = average of 2 variances when ns are equal

Mean Square Within (MS_W) or Mean Square Error (MS_E)

'Mean Square' or ***MS***

is a fancy statistical term for the **variance**

'Square': Refers to the sum of SQUARED (SS) deviations from the mean

Mean: AVERAGE of the SS deviations

SS is divided by N or $N - 1$ to yield variance

So, Mean of the sum of SQUARED deviations = Variance

All we want to know is whether variation among group means exceeds that variation within groups

Will create a **ratio of the *MS*s**, the F-statistic, to see if this ratio is significantly **different from 1**

PRIOR EXAMPLE

Applying data from independent-samples *t*-test example (drug v. placebo and depression)

- Recall, t = 1.96, p = .085

Group 1 - Drug	Group 2 - Placebo
11	11
1	11
0	5
2	8
0	4

depression

1 Drug	N	Valid	5
		Missing	0
	Mean		2.80
	Std. Deviation		4.658
2 Placebo	N	Valid	5
		Missing	0
	Mean		7.80
	Std. Deviation		3.271

$$t^2 = \frac{n_j \frac{\left(\bar{X}_1 - \bar{X}_2\right)^2}{2}}{s_p^2} = \frac{5 * \frac{(7.8-2.8)^2}{2}}{\left(\frac{3.27^2 + 4.66^2}{2}\right)} = \frac{5*12.5}{\left(\frac{32.41}{2}\right)} = \frac{62.5}{16.21} = 3.84 = F$$

1.96^2 = 3.84

$$\boxed{t^2 = F}$$

LINK: INDEPENDENT SAMPLE "T-TEST" & ANOVA

Same principle underlies many statistical tests

$$F = \frac{MS_B}{MS_W} = \frac{\text{Measure of effect (or treatment) assessed by examining variance (or differences) between groups}}{\text{Measure of random variation (or error) assessed by examining variance (or differences) within groups}}$$

How do we rearrange t-statistic equation to accommodate >2 samples with <u>equal</u> n_j?

<u>Numerator</u>

MS_B : Compute variance <u>between</u> (among) sample means, multiply by n_j

<u>Denominator</u>

MS_W : Compute average of sample variances

ASSUMPTIONS

Large or multiple violations will GREATLY increase risk of inaccurate p-values
Increased probability of Type I or II error

Independent, Random Sampling (for the IV)

- For preexisting populations: randomly select a sample from each population
- For experimental conditions: randomly divide your sample *(of convenience)* for assignment to groups
- Ensure no connection between subjects in the different groups (no matching!) ← MUST!!!

Normally distributed (DV)

- Robust requirement…if samples are large, this isn't as important
- If not normal (or small samples)…alternatives : use the Krukal-Wallis H test

HOV: homogeneity of Variance (DV)

- Since an average variance is computed for denominator of F-statistic, variance should be similar for all groups: $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_k^2$
- σ_e^2, pooled or averaged variance, must be representative of each group so that MS_W is accurate
- Testing: F-test is rarely use, Levene's Test is more common
- All test for HOV are underpowered if samples are small, so you have to use judgement ;)
- If NOT HOV…alternatives: Welch, Brown-Forsythe, ect.

LOGIC OF "ANOVA"

$$F_{Ratio} \text{ or } F_{Statistic} = \frac{MS_B}{MS_W}$$

In ANOVA, 2 independent estimates of same population (error) variance are computed: σ^2, now called σ_e^2

- MS_B : Variance between group means corrected by sample sizes (n_j)
- MS_W : Average variance within groups

Ratio of 2 estimates of population variance

- Hence the term *Analysis of Variance,* instead of something related to means comparisons (even though that is what we are interested in doing)

F-ratio increases as variance among means increases

- Given within-group variance is constant

Increased **variance among means** indicates means are **spread out**

&likely differ from one another or come from different populations

Large F-ratio indicates differences among means is NOT likely due to chance

ANOVA is simply….

$$\frac{\text{Between-Group Measure of Variation Due to Estimate of Random Variation (Error)} + \text{Effect of IV (Group)}}{\text{Within-Group Estimate of Random Variation (Error)}}$$

F-STATISTIC: NUMERATOR = MS_B

Recall from CLT, relationship between

variance of population (σ^2) &

variance of SDM ($SE^2 = \sigma^2_{\bar{X}}$)

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n_j}} \rightarrow \sigma^2_{\bar{X}} = \frac{\sigma^2}{n_j} \rightarrow \sigma^2_{\bar{X}} \cdot n_j = \sigma_e^{\ 2} = MS_B$$

One estimate of population variance: $\sigma_e^{\ 2}$

Cannot compute population variance of all possible means as we only have a sample

- Estimate population **variance** with sample means and **multiply** by sample size:

Equal

Sample Sizes

$$MS_B = n \cdot s^2_{\bar{X}}$$

UN-equal

Sample Sizes

$$MS_B = \frac{\sum n_j \left(\bar{X}_j - \bar{X}_G\right)^2}{k \ - 1}$$

If H_0 true, $MS_B = \sigma_e^{\ 2}$

Have drawn k independent samples

From the SAME population

(i.e. group differences = 0)

If H_0 false, $MS_B \neq \sigma_e^{\ 2}$

MS_B reflects BOTH

population variance

AND

group differences

EXAMPLE: NOISE & WORDS MEMORIZED

1. *Find grand mean:*

2. *Find the SD of the means:*

3. *Multiply by n*

Equal

Sample Sizes

$$MS_B = n \cdot s_{\bar{X}}^2$$

Statistics

Words Number of nonsense words recalled

1 No noise (no music, quiet room)	N	Valid	5
		Missing	0
	Mean		9.20
	Std. Deviation		.837
2 moderate noise (classical music)	N	Valid	5
		Missing	0
	Mean		6.60
	Std. Deviation		1.517
3 extreme noise (rock music)	N	Valid	5
		Missing	0
	Mean		6.20
	Std. Deviation		1.643

F-STATISTIC: DENOMINATOR $= MS_W$

Goal should be to obtain equal *ns*
BUT…
1 group > 50% larger other group: too much

Second estimate of population variance: σ_e^2

Pooling sample variances yields best estimate

- $\sigma_1^2 = s_1^2$; $\sigma_2^2 = s_2^2$; … ; $\sigma_j^2 = s_j^2$

k = # subgroups
"j" denotes the j-th subgroup

Average subgroup (*j*) variance: $\sigma_e^2 = s_e^2$

Equal
Sample Sizes

$$MS_W = \sigma_e^2 = \frac{\sum s_j^2}{k}$$

UN-equal
Sample Sizes

$$MS_W = \sigma_e^2 = \frac{\sum (n_j - 1) s_j^2}{n_T - k}$$

Regardless of whether H_0 true:

$$MS_W = \sigma_e^2$$

Not affected by group MEANS

EXAMPLE: NOISE & WORDS MEMORIZED

Statistics

Words Number of nonsense words recalled

1 No noise (no music, quiet room)	N	Valid	5
		Missing	0
	Mean		9.20
	Std. Deviation		.837
2 moderate noise (classical music)	N	Valid	5
		Missing	0
	Mean		6.60
	Std. Deviation		1.517
3 extreme noise (rock music)	N	Valid	5
		Missing	0
	Mean		6.20
	Std. Deviation		1.643

Equal

Sample Sizes

$$MS_W = \sigma_e{}^2 = \frac{\sum s_j{}^2}{k}$$

1. *Average the **VARIANCES's**:*

LOGIC OF "ANOVA"

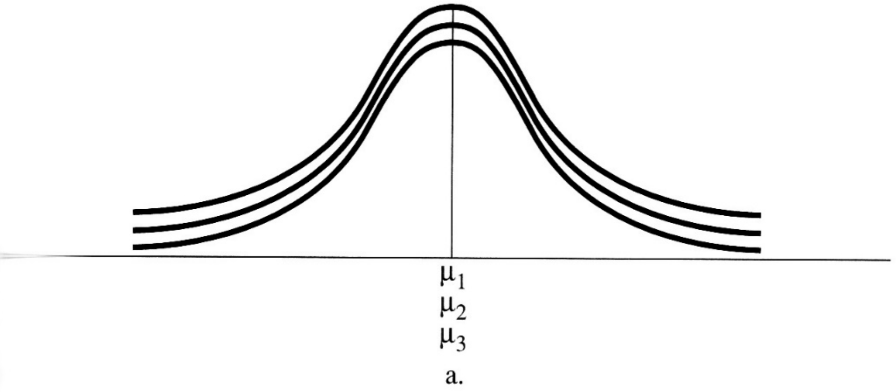
Groups all same sample sizes

$$\left.\begin{aligned} MS_B &= \sigma_e{}^2 = n_j \cdot s^2{}_{\bar{X}} \\ MS_W &= \sigma_e{}^2 = \frac{\sum s_j{}^2}{k} \end{aligned}\right\} \xrightarrow{ratio} \boldsymbol{F} = \frac{MS_B}{MS_W}$$

When estimates of $\sigma_e{}^2$ are…

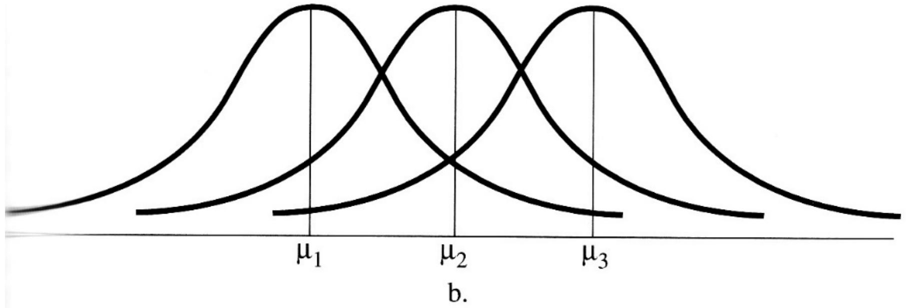
Equal: Fail to reject H_0

- All means come from same population
- Both are estimates of the same population variance $\sigma_e{}^2$
- F-ratio ≈ 1



Unequal: Reject H_0

- **Unlikely** that all means come from same population
- **Effect of IV surpasses random error/variation within groups**
- F-ratio significantly > 1, $\boldsymbol{MS_B > MS_W}$



CALCULATIONS:

SUMMARY STATS KNOWN ← shown on previous few slides
SUM OF SQUARES (SS) APPROACH ← alternate formulas here

$$SS = \sum_{i=1}^{n}(X_i - \bar{X})^2$$

Can '**partition**' total variation in DV due to group effects (IV) and error

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

Total

How different are ALL individuals from the "GRAND MEAN"

Inner Sum: individuals in each subgroup

Outter Sum: subgroups in the whole

$$SS_{Total} = \sum_{j=1}^{k}\sum_{i=1}^{n}(X_{ij} - \bar{X}_{GM})^2$$

$$\boxed{df_T = n_T - 1}$$

Between

How different are "GROUP MEANS" from the "GRAND MEAN"

$$SS_{Between} = n_j \sum_{j=1}^{k}(\bar{X}_j - \bar{X}_{GM})^2$$

$$\boxed{df_B = k - 1}$$

$$MS_{Between} = \frac{SS_B}{df_B} = \frac{n_j \sum_{j=1}^{k}(\bar{X}_j - \bar{X}_{GM})^2}{k-1}$$

Within

How different are individuals from their "GROUP's MEAN"

Inner Sum: individuals in each subgroup

Outter Sum: subgroups in the whole

$$SS_{Within} = \sum_{j=1}^{k}\sum_{i=1}^{n}(X_{ij} - \bar{X}_j)^2$$

$$\boxed{df_W = n_T - k}$$

$$MS_{Within} = \frac{SS_W}{df_W} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n}(X_{ij} - \bar{X}_j)^2}{N-k}$$

$$\boxed{F_{Ratio} \text{ or } F_{Statistic} = \frac{MS_B}{MS_W}}$$

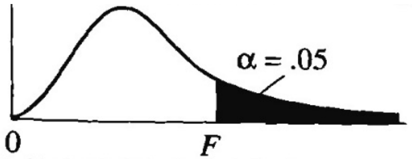
F-STATISTIC

$$F_{Ratio} \text{ or } F_{Statistic} = \frac{MS_B}{MS_W}$$

$F_{\text{crit}} \rightarrow F$-distribution table

(different table per α)

- *Across the top: find df_B*
- *Down the side: find df_W*



If H_0 is true, $MS_B = MS_W$

F-statistic ≈ 1

Both are estimates of variance of **same** population

If H_0 is false, $MS_B > MS_W$

F-statistic exceeds F_{crit} by some amount

At least one mean significantly differs from another

	df Numerator											
df Denominator	1	2	3	4	5	6	7	8	9	10	12	15
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07
27	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

EXAMPLE: NOISE & WORDS MEMORIZED

$MS_B = \mathbf{13.267}$ $MS_W = \mathbf{1.90}$

Statistics

Words Number of nonsense words recalled

1 No noise (no music, quiet room)	N	Valid	5
		Missing	0
	Mean		9.20
	Std. Deviation		.837
2 moderate noise (classical music)	N	Valid	5
		Missing	0
	Mean		6.60
	Std. Deviation		1.517
3 extreme noise (rock music)	N	Valid	5
		Missing	0
	Mean		6.20
	Std. Deviation		1.643

Test statistic: F-score observed

Critical Value: F-crit for α=.05

Conclusion:

SPSS

```
ONEWAY Words BY Noise
  /STATISTICS DESCRIPTIVES HOMOGENEITY
  /PLOT MEANS
  /MISSING ANALYSIS.
```

	Noise	Words	var
1	1	8	
2	1	10	
3	1	9	
4	1	10	
5	1	9	
6	2	7	
7	2	8	
8	2	5	
9	2	8	
10	2	5	
11	3	4	
12	3	8	
13	3	7	
14	3	5	
15	3	7	

Descriptives

Words Number of nonsense words recalled

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1 No noise (no music, quiet room)	5	9.20	.837	.374	8.16	10.24	8	10
2 moderate noise (classical music)	5	6.60	1.517	.678	4.72	8.48	5	8
3 extreme noise (rock music)	5	6.20	1.643	.735	4.16	8.24	4	8
Total	15	7.33	1.877	.485	6.29	8.37	4	10

Test of Homogeneity of Variances

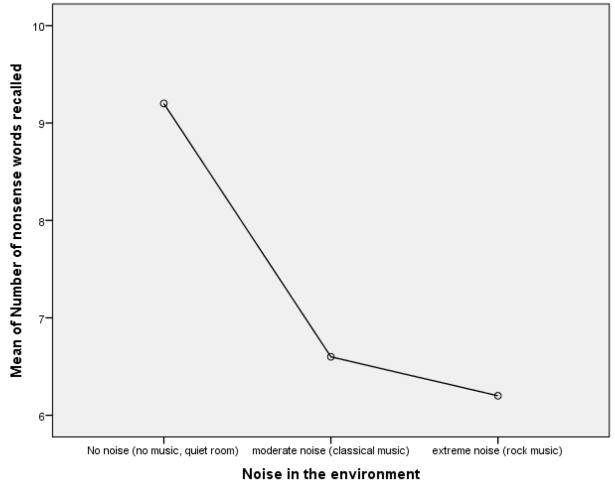
Words Number of nonsense words recalled

Levene Statistic	df1	df2	Sig.
2.821	2	12	.099

ANOVA

Words Number of nonsense words recalled

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	26.533	2	13.267	6.982	.010
Within Groups	22.800	12	1.900		
Total	49.333	14			



MEASURES OF ASSOCIATION

<u>Term preferred over "Effect size" for ANOVA</u>

- Amount or % of variation in DV explained/accounted for by knowledge of group membership (IV)
- Correlation between grouping variable (IV) and outcome variable (DV)

<u>4 measures:</u>

- Eta-squared (η^2)
- Omega-squared (ω^2)
- Cohen's f
- Intra-class Correlation Coefficients (ρ)

ω^2 is least biased, but unfamiliarity and 'difficulty' of computation have limited use

η^2 probably sufficient in many cases

MEASURES OF ASSOCIATION: ETA-SQUARED

Range: 0 to 1

Small: .01 to .06

Medium: .06 to .14

Large: > .14

η^2 : Measure of % reduction in error IN THIS DATA (SAMPLES)

- SS_{Total} = Error in DV around grand mean
- SS_{Within} = Error around group means
- By knowing group membership we reduce error by $SS_{Between} = SS_{Total} - SS_{Within}$

% reduction in error expressed as…

$$\eta^2 = \frac{SS_B}{SS_T} = \frac{df_B \cdot F}{df_B \cdot F + df_W}$$

η^2 can be biased with sample data

- Adjusted $\eta^2 = 1 - \frac{MS_W}{MS_T}$

Compute using information from ANOVA summary table

- $\eta^2 = SS_B / SS_T$
- $\eta^2_{adj} = 1 - (MS_W / MS_T)$

EXAMPLE: NOISE & WORDS MEMORIZED

$df = (3-1, 15-3) = (\mathbf{2}, \mathbf{12})$

$F(2,12) = \dfrac{13.267}{1.90} = \mathbf{6.98}$

$MS_B = \mathbf{13.267} \xRightarrow{\text{"}SS = MS/df} SS_B = \mathbf{13.267} * (\mathbf{2}) = \mathbf{26.534}$

$MS_W = \mathbf{1.90} \xRightarrow{\text{"}SS = MS/df} SS_B = \mathbf{1.9} * (\mathbf{12}) = \mathbf{22.8}$

Statistics

Words Number of nonsense words recalled

1 No noise (no music, quiet room)	N	Valid	5
		Missing	0
	Mean		9.20
	Std. Deviation		.837
2 moderate noise (classical music)	N	Valid	5
		Missing	0
	Mean		6.60
	Std. Deviation		1.517
3 extreme noise (rock music)	N	Valid	5
		Missing	0
	Mean		6.20
	Std. Deviation		1.643

Range: 0 to 1

Small: .01 to .06

Medium: .06 to .14

Large: > .14

$$\eta^2 = \frac{SS_B}{SS_T} = \frac{df_B \cdot F}{df_B \cdot F + df_W}$$

Using SS

Using F & df's

Conclusion:

MEASURES OF ASSOCIATION: OMEGA-SQUARED

Range: 0 to 1
Small: .01 to .06
Medium: .06 to .14
Large: > .14

ω^2 : Measure of % reduction in error IN THIS POPULATION (ESTIMATE TRUTH)

Alternative for "fixed-effects" ANOVA

- More conservative than η^2 (and less biased)
- Range: 0 to 1 (can be negative when $F < 1$)
  - Same interpretation as η^2
- Compute using information from ANOVA summary table
  - Equation for fixed effects ANOVA only

$$\omega^2 = \frac{SS_B - (k-1)MS_W}{SS_T + MS_W} = \frac{(k-1)(F-1)}{(k-1)(F-1) + n_j \cdot k}$$

MEASURES OF ASSOCIATION: COHEN'S F

Range: 0 to infinity

Small: .10 to .25

Medium: .25 to .40

Large: > .40

Traditional effect size index

- Not a measure of association
- Generalization of Cohen's d to ANOVA
- Compute using ANOVA summary information

$$f = \sqrt{\frac{\omega^2}{1-\omega^2}} = \sqrt{\frac{\frac{k-1}{n_j \cdot k}\left(MS_B - MS_W\right)}{MS_W}}$$

Converting from f to ω^2 →

$$\omega^2 = \frac{f^2}{1+f^2}$$

MEASURES OF ASSOCIATION: INTRA-CLASS CORRELATION COEFFICIENT (ICC)

Measure of association for <u>random-effects</u> ANOVA

At least 6 ICCs available

- Type selected depends on data structure
- Range: 0 to 1
- Commonly used measure of agreement for continuous data

Basic form:

$$\rho_{\text{intraclass}} = \frac{MS_B - MS_W}{MS_B + (n_j - 1)MS_W}$$

- Measures extent to which observations within a treatment are similar to one another relative to observations in different treatments

APA RESULTS

Methods

- Describe statistical and sample size analyses
- Describe factor and its levels
- Results of data screening

Results

- Reporting F-test:
  - $F(df_B, df_W)$ = F-statistic, p = / <, measure of association and effect/effect size, power (optional)
- Don't need to include MSE (or MS_W) as Cohen suggests
- Discuss any follow-up tests, if any (next lecture)

EXAMPLE:

Method

"A 1-way ANOVA was used to test the hypothesis that the means of the three groups (Control, Moderate Noise, and Extreme Noise) were different following the experiment. A sample size analysis conducted prior to beginning the study indicated that five participants per group would be sufficient to reject the null hypothesis with at least 80% power if the effect size were moderate (Cohen's f = .95)."

Results

"Results indicated a significant difference among the group means, $F(2, 12) = 6.98$, $p < .01$, $\eta^2 = .54$"

ANOVA VS. MULTIPLE T-TESTS

Why not run series of independent-samples t-tests?

Could, and will usually get same results, but this approach becomes more difficult under 2 conditions:

- Large k
  - $k(k-1) / 2$ different t-tests!
- Factorial designs

Danger of increased risk of Type I error when conducting multiple t-tests on same data set

- In next lecture we explain ways to potentially limit this risk