# 1-Factor ANOVA

## Engineering Statistics II
## Section 10.1

Josh Engwer

TTU

2018

PART I:

Many-Sample Inference
Experimental Design Terminology

# Many-Sample Inference (Example)

Suppose we wish to determine whether three light bulb brands all have similar lifetimes or not. A sample of 5 bulbs from each brand has their lifetimes measured (in years) and recorded in the below table:

| **BULB BRAND:** | **SAMPLE SIZE:** | **LIFETIMES (in yrs):** |
|---|---|---|
| Brand 1 ($x_{1\bullet}$) | 5 | $9.22,\ 9.07,\ 8.95,\ 8.98,\ 9.54$ |
| Brand 2 ($x_{2\bullet}$) | 5 | $8.92,\ 8.88,\ 9.10,\ 8.71,\ 8.85$ |
| Brand 3 ($x_{3\bullet}$) | 5 | $9.08,\ 8.99,\ 9.06,\ 8.93,\ 9.02$ |

or expressed in terms of means and standard deviations:

| **BULB BRAND:** | **SAMPLE SIZE:** | **MEAN LIFETIMES (in yrs):** | **STD DEV:** |
|---|---|---|---|
| Brand 1 ($x_{1\bullet}$) | 5 | $\bar{x}_{1\bullet} = 9.152$ | $s_1 \approx 0.2410$ |
| Brand 2 ($x_{2\bullet}$) | 5 | $\bar{x}_{2\bullet} = 8.892$ | $s_2 \approx 0.1406$ |
| Brand 3 ($x_{3\bullet}$) | 5 | $\bar{x}_{3\bullet} = 9.016$ | $s_3 \approx 0.0594$ |

Now, the appropriate hypotheses are:

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

$H_A :$ At least two of the $\mu$'s differ

where $\mu_i \equiv$ ( Population Mean of all Brand $i$ light bulbs )

# Experimental Design Terminology

## Definition

The collection of $I$ samples to determine cause & effect is an **experiment**.
A **balanced experiment** has equal-sized samples/groups.
Each data point of a sample is called an **observation** or **measurement**.
The dependent variable to be measured is called the **response**.
The manner of sample collection & grouping is called **experimental design**.
The main characteristic distinguishing all the samples is called the **factor**.
The factor's particular values or settings are called its **levels**.
Each sample corresponding to a level is called a **group**.

| **FACTOR A:** | **GROUP SIZE:** | **GROUPS:** |
|---|---|---|
| Level 1 | $J$ | $x_{1\bullet}: \quad x_{11}, x_{12}, \cdots, x_{1J}$ |
| Level 2 | $J$ | $x_{2\bullet}: \quad x_{21}, x_{22}, \cdots, x_{2J}$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| Level $I$ | $J$ | $x_{I\bullet}: \quad x_{I1}, x_{I2}, \cdots, x_{IJ}$ |

This section (§10.1) & §10.2 involve only balanced experiments.
This chapter's last section (§10.3) considers <u>unbalanced</u> experiments.

# Experimental Design Terminology

## Definition

The collection of $I$ samples to determine cause & effect is an **experiment**.
A **balanced experiment** has equal-sized samples/groups.
Each data point of a sample is called an **observation** or **measurement**.
The dependent variable to be measured is called the **response**.
The manner of sample collection & grouping is called **experimental design**.
The main characteristic distinguishing all the samples is called the **factor**.
The factor's particular values or settings are called its **levels**.
Each sample corresponding to a level is called a **group**.

| **FACTOR A:** | **GROUP SIZE:** | **GROUP MEAN:** | **GROUP STD DEV:** |
|---|---|---|---|
| Level 1 $(x_{1\bullet})$ | $J$ | $\bar{x}_{1\bullet}$ | $s_1$ |
| Level 2 $(x_{2\bullet})$ | $J$ | $\bar{x}_{2\bullet}$ | $s_2$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| Level $I$ $(x_{I\bullet})$ | $J$ | $\bar{x}_{I\bullet}$ | $s_I$ |

This section (§10.1) & §10.2 involve only balanced experiments.
This chapter's last section (§10.3) considers <u>unbalanced</u> experiments.

# Experimental Design Terminology (Example)

| FACTOR A: (BULB BRAND) | GROUP SIZE: | GROUPS: (BULB LIFETIMES in yrs) |
|---|---|---|
| Level 1 $(x_{1\bullet})$ | 5 | $9.22,\ 9.07,\ 8.95,\ 8.98,\ 9.54$ |
| Level 2 $(x_{2\bullet})$ | 5 | $8.92,\ 8.88,\ 9.10,\ 8.71,\ 8.85$ |
| Level 3 $(x_{3\bullet})$ | 5 | $9.08,\ 8.99,\ 9.06,\ 8.93,\ 9.02$ |

or expressed in terms of means and standard deviations:

| FACTOR A: (BULB BRAND) | GROUP SIZE: | GROUP MEAN: | GROUP STD DEV: |
|---|---|---|---|
| Level 1 $(x_{1\bullet})$ | 5 | $\bar{x}_{1\bullet} = 9.152$ | $s_1 \approx 0.2410$ |
| Level 2 $(x_{2\bullet})$ | 5 | $\bar{x}_{2\bullet} = 8.892$ | $s_2 \approx 0.1406$ |
| Level 3 $(x_{3\bullet})$ | 5 | $\bar{x}_{3\bullet} = 9.016$ | $s_3 \approx 0.0594$ |

$H_0$ : $\mu_1 = \mu_2 = \mu_3$
$H_A$ : At least two of the $\mu$'s differ

where $\mu_i \equiv$ ( Population Mean of all Brand $i$ light bulbs )

REMARK: More about experimental design in later sections.

PART II:

The Problem with Many-Sample $t$-Tests

# The Problem with Many-Sample $t$-Tests

Suppose a designed experiment calls to test four independent samples:

$$\begin{aligned} H_0 &: \quad \mu_1 = \mu_2 = \mu_3 = \mu_4 \\ H_A &: \quad \text{At least two of the } \mu\text{'s differ} \end{aligned}$$

One way to do this is perform $\binom{4}{2}$ independent $t$-tests, each at signif. level $\alpha$:

$$\begin{array}{ccc} H_0^{(1)} : \ \mu_1 = \mu_2 & H_0^{(2)} : \ \mu_1 = \mu_3 & H_0^{(3)} : \ \mu_1 = \mu_4 \\ H_A^{(1)} : \ \mu_1 \neq \mu_2 & H_A^{(2)} : \ \mu_1 \neq \mu_3 & H_A^{(3)} : \ \mu_1 \neq \mu_4 \\ \hline H_0^{(4)} : \ \mu_2 = \mu_3 & H_0^{(5)} : \ \mu_2 = \mu_4 & H_0^{(6)} : \ \mu_3 = \mu_4 \\ H_A^{(4)} : \ \mu_2 \neq \mu_3 & H_A^{(5)} : \ \mu_2 \neq \mu_4 & H_A^{(6)} : \ \mu_3 \neq \mu_4 \end{array}$$

# The Problem with Many-Sample $t$-Tests

Suppose a designed experiment calls to test four independent samples:

$$\begin{array}{rl} H_0: & \mu_1 = \mu_2 = \mu_3 = \mu_4 \\ H_A: & \text{At least two of the } \mu\text{'s differ} \end{array}$$

Alas, since each successive $t$-test is performed with the same dataset, the **experiment-wise significance level**, $\alpha_{exp}$, grows with each $t$-test:

$$\begin{aligned} \alpha_{exp} &:= \mathbb{P}(\text{Committing a Type I Error in at least one } t\text{-test}) \\ &= 1 - \mathbb{P}(\text{Never Committing a Type I Error in any of the } t\text{-tests}) \\ &= 1 - \mathbb{P}\left(\bigcap_{i=1}^{6}(\text{Not Committing a Type I Error in } i^{th}\ t\text{-test})\right) \\ &\overset{IND}{=} 1 - \prod_{i=1}^{6} \mathbb{P}(\text{Not Committing a Type I Error in } i^{th}\ t\text{-test}) \\ &\overset{\alpha}{=} 1 - \prod_{i=1}^{6}(1-\alpha) \\ &= 1 - (1-\alpha)^6 \end{aligned}$$

$$\left[\alpha := \mathbb{P}\left(\text{Rejecting } H_0^{(k)} \mid H_0^{(k)} \text{ is True}\right)\right]$$

# The Problem with Many-Sample $t$-Tests

$H_0$ : $\mu_1 = \mu_2 = \mu_3 = \mu_4$
$H_A$ : At least two of the $\mu$'s differ

$$N_{t\text{-}tests} \equiv (\#\ t\text{-tests}) = \binom{4}{2} = 6$$

Alas, with successive $t$-tests, $\alpha_{exp}$ grows (AKA $\alpha$**-inflation**):

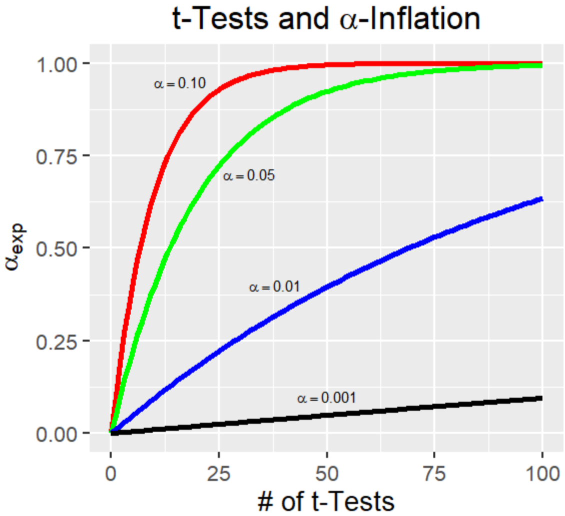
| Chosen $\alpha$ | Resulting $\alpha_{exp} = 1 - (1-\alpha)^6$ |
|---|---|
| 0.10 | **0.4686** |
| 0.05 | **0.2649** |
| 0.01 | **0.0585** |
| 0.001 | **0.0060** |

One can determine which $\alpha$ achieves a desired $\alpha_{exp}$, but often it's not feasible:

| Required $\alpha = 1 - (1-\alpha_{exp})^{1/6}$ | Desired $\alpha_{exp}$ |
|---|---|
| **0.0174** | 0.10 |
| **0.0085** | 0.05 |
| **0.0017** | 0.01 |
| **0.0002** | 0.001 |

A loose (rough) upper bound for $\alpha_{exp}$ is $\alpha \times$ (# $t$-tests): $\alpha_{exp} \leq \alpha N_{t\text{-}tests}$

# The Problem with Many-Sample $t$-Tests




To prevent $\alpha$-inflation, all means should be simultaneously tested.

# PART III

PART III:

1-Factor Fixed Effects Linear (Statistical) Models:

Definitions, Examples

Least Squares Estimators (LSE's)

Best Linear Unbiased Estimators (BLUE's)

Gauss-Markov Theorem

# 1-Factor Fixed Effects Linear (Statistical) Models

With many-sample inference, it's convenient to use a **linear model**:

## Definition

(1-Factor Fixed Effects Linear Model)

Given a 1-factor balanced experiment with $I > 2$ groups, each of size $J$.

Let $X_{ij} \equiv$ random variable for $j^{th}$ measurement in the $i^{th}$ group.

Then, the **fixed effects linear model** for the experiment is defined as:

$$X_{ij} = \mu + \alpha_i^A + E_{ij} \quad \text{where} \quad E_{ij} \overset{iid}{\sim} \text{Normal}(0, \sigma^2)$$

where:

$$\begin{aligned} \mu &\equiv \text{population grand mean of all } I \text{ population means} \\ \alpha_i^A &\equiv \text{deviation of } i^{th} \text{ population mean } \mu_i \text{ from } \mu \text{ due to Factor A} \\ E_{ij} &\equiv \text{rv for error/noise applied to } j^{th} \text{ measurement in } i^{th} \text{ group} \end{aligned}$$

**Fixed effects** means <u>all relevant levels</u> of factor A are considered in model.

$$X_{ij} = \mu$$

$$\mu := 3.2$$

$$\mu_1 = 3.2,\ \mu_2 = 3.2,\ \mu_3 = 3.2$$

| FACTOR A: | MEASUREMENTS: | | | |
|---|---|---|---|---|
| Level 1 ($x_{1\bullet}$) | $x_{11} = 3.2$, | $x_{12} = 3.2$, | $x_{13} = 3.2$, | $x_{14} = 3.2$ |
| Level 2 ($x_{2\bullet}$) | $x_{21} = 3.2$, | $x_{22} = 3.2$, | $x_{23} = 3.2$, | $x_{24} = 3.2$ |
| Level 3 ($x_{3\bullet}$) | $x_{31} = 3.2$, | $x_{32} = 3.2$, | $x_{33} = 3.2$, | $x_{34} = 3.2$ |

# 1-Factor Linear Models (Motivating Example)

$$X_{ij} = \mu + \alpha_i^A$$

$$\mu := 3.2$$

$$\alpha_1^A := -5.5,\ \alpha_2^A := -2.0,\ \alpha_3^A := 7.5$$

$$\mu_1 = -2.3,\ \mu_2 = 1.2,\ \mu_3 = 10.7$$

| **FACTOR A:** | **MEASUREMENTS:** | | | |
|---|---|---|---|---|
| Level 1 $(x_{1\bullet})$ | $x_{11} = -2.3,$ | $x_{12} = -2.3,$ | $x_{13} = -2.3,$ | $x_{14} = -2.3$ |
| Level 2 $(x_{2\bullet})$ | $x_{21} = 1.2,$ | $x_{22} = 1.2,$ | $x_{23} = 1.2,$ | $x_{24} = 1.2$ |
| Level 3 $(x_{3\bullet})$ | $x_{31} = 10.7,$ | $x_{32} = 10.7,$ | $x_{33} = 10.7,$ | $x_{34} = 10.7$ |

# 1-Factor Linear Models (Motivating Example)

$$X_{ij} = \mu + \alpha_i^A + E_{ij}$$

$$\mu := 3.2$$

$$\alpha_1^A := -5.5,\ \alpha_2^A := -2.0,\ \alpha_3^A := 7.5$$

$$\mu_1 = -2.3,\ \mu_2 = 1.2,\ \mu_3 = 10.7$$

$$E_{ij} \overset{iid}{\sim} \text{Normal}(0, \sigma^2 := 3.24)$$

| **FACTOR A:** | **MEASUREMENTS:** | | | |
|---|---|---|---|---|
| Level 1 $(x_{1\bullet})$ | $x_{11} = -1.23$, | $x_{12} = -1.17$, | $x_{13} = 0.05$, | $x_{14} = -3.08$ |
| Level 2 $(x_{2\bullet})$ | $x_{21} = 0.54$, | $x_{22} = 1.03$, | $x_{23} = 0.62$, | $x_{24} = 1.63$ |
| Level 3 $(x_{3\bullet})$ | $x_{31} = 13.64$, | $x_{32} = 12.30$, | $x_{33} = 11.74$, | $x_{34} = 10.60$ |

# 1-Factor Linear Models (Least-Squares Estimators)

Like all population parameters, linear model parameters can be estimated:

## Proposition

*Given a 1-factor linear model:*

$$X_{ij} = \mu + \alpha_i^A + E_{ij} \quad \textit{where} \quad E_{ij} \overset{iid}{\sim} \textit{Normal}(0, \sigma^2)$$

*Then:*

*(a) The **least-squares♠♣ estimators (LSE's)** for the model parameters are:*

$$\begin{aligned} \hat{\mu} &= \bar{x}_{\bullet\bullet} \\ \hat{\alpha}_i^A &= \bar{x}_{i\bullet} - \bar{x}_{\bullet\bullet} \end{aligned} \quad \textit{where} \quad \begin{aligned} \bar{x}_{\bullet\bullet} &\equiv \textit{Grand sample mean} \\ \bar{x}_{i\bullet} &\equiv \textit{Sample mean of } i^{th} \textit{ group} \end{aligned}$$

*(b) For these least-squares estimators, it's required that* $\sum_i \hat{\alpha}_i^A = 0$.

*(c) These least-squares estimators are all unbiased.*

PROOF: The general case is left as an ungraded exercise for the reader.

♠A.M. Legendre, *Nouvelles Méthodes pour la Détermination des Orbites des Comètes*, 1806.

♣Gauss, *Theoria Motus Corporum Coelestrium in Sectionibus Conicis Solem Ambientium*, 1809.

# 1-Factor Linear Models (Predicted Responses)

With the model parameter estimators in hand, responses can be predicted:

**Definition**

(Predicted Responses)

Given a 1-factor linear model:

$$X_{ij} = \mu + \alpha_i^A + E_{ij} \quad \text{where} \quad E_{ij} \overset{iid}{\sim} \text{Normal}(0, \sigma^2)$$

Then the corresponding **predicted responses**, denoted $\hat{x}_{ij}$, are:

$$\hat{x}_{ij} := \hat{\mu} + \hat{\alpha}_i^A = \bar{x}_{\bullet\bullet} + (\bar{x}_{i\bullet} - \bar{x}_{\bullet\bullet}) = \bar{x}_{i\bullet}$$

<u>SYNONYMS:</u> Predicted values, fitted values

# 1-Factor Linear Models (Residuals)

With the predicted responses in hand, residuals can be computed:

## Definition

(Residuals)

Given a 1-factor linear model:

$$X_{ij} = \mu + \alpha_i^A + E_{ij} \quad \text{where} \quad E_{ij} \overset{iid}{\sim} \text{Normal}(0, \sigma^2)$$

Then the corresponding predicted responses, denoted $\hat{x}_{ij}$, are:

$$\hat{x}_{ij} := \hat{\mu} + \hat{\alpha}_i^A = \bar{x}_{\bullet\bullet} + (\bar{x}_{i\bullet} - \bar{x}_{\bullet\bullet}) = \bar{x}_{i\bullet}$$

Moreover, the corresponding **residuals**, denoted $x_{ij}^{res}$, are:

$$x_{ij}^{res} := x_{ij} - \hat{x}_{ij} = x_{ij} - \bar{x}_{i\bullet}$$

Point estimators for a linear model should be ideal ones:

**Definition**

(Best Linear Unbiased Estimators – BLUE's)

A point estimator $\hat{\theta}$ is called a **best linear unbiased estimator (BLUE)** if:

- It estimates a parameter $\theta$ of a linear model.
- It is a linear combination of the data points: $\hat{\theta} := \sum_{k=1}^{n} c_k x_k$
- It is an unbiased estimator: $\mathbb{E}[\hat{\theta}] = \theta$
- It has minimum variance of all such unbiased estimators.

<u>REMARK:</u> BLUE's are generally easier to construct & prove than UMVUE's.

For a 1-factor linear model: $X_{ij} = \mu + \alpha_i^A + E_{ij}$

$\hat{\mu}, \hat{\alpha}_i^A$ are each linear combinations of data points in the linear model.
A particular example of demonstrating this is done in EX 10.1.1.

Ideally, point estimators for linear model parameters should be BLUE's:

## Theorem

*(Gauss[1]-Markov[2] Theorem)*

*Given a 1-factor linear model:* $X_{ij} = \mu + \alpha_i^A + E_{ij}$
*Moreover, suppose the following conditions are all satisfied:*

$$\begin{array}{rcll} \mathbb{E}[E_{ij}] & = & 0 & (\textit{errors are all centered at zero}) \\ \mathbb{V}[E_{ij}] & = & \sigma^2 & (\textit{errors all have the same finite variance}) \\ \mathbb{C}[E_{ij}, E_{i'j'}] & = & 0 & (\textit{errors are uncorrelated when } i \neq i' \textit{ or } j \neq j') \end{array}$$

*Then, the least-squares estimators (LSE's)* $\hat{\mu}, \hat{\alpha}_i^A$ *are all BLUE's.*

PROOF: Omitted due to time.

[1] C.F. Gauss, "Theoria Combinationis Observationum Erroribus Minimis Obnoxiae", (1823), 1-58.

[2] A.A. Markov, *Calculus of Probabilities*, $1^{st}$ Edition, 1900.

PART IV:

1-Factor Analysis of Variance (ANOVA):

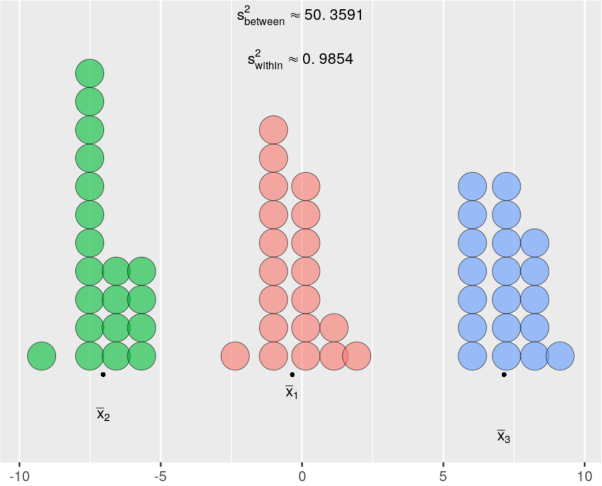
Motivation

Basic Model Assumptions

$F$-Test Statistic Value

# 1-Factor Analysis of Variance (Motivation)

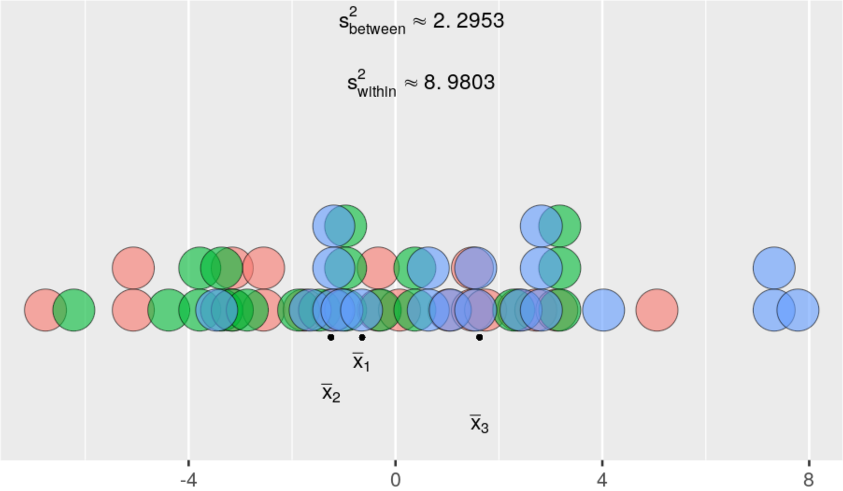
High variance between groups
Low variance within groups




$s^2_{between} / s^2_{within} \gg 1 \implies$ Factor A clearly has a significant effect!!

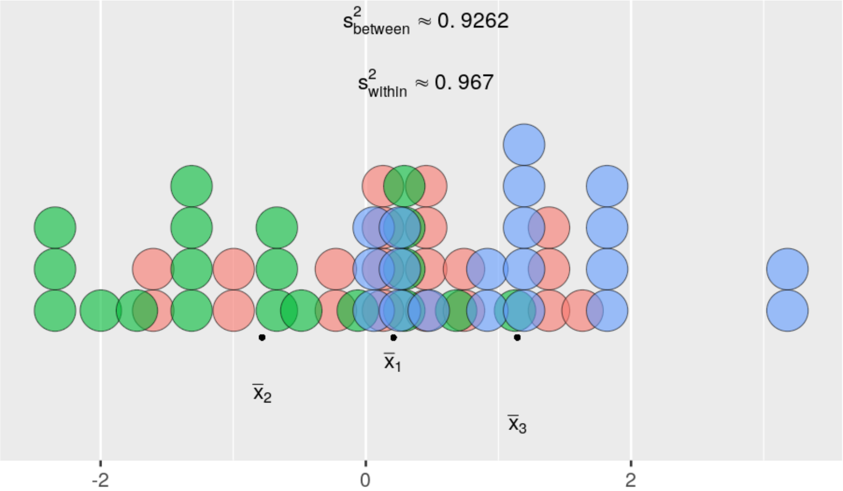
# 1-Factor Analysis of Variance (Motivation)

Low variance between groups
High variance within groups




$s^2_{between}/s^2_{within} \ll 1 \implies$ Factor A clearly has no significant effect!

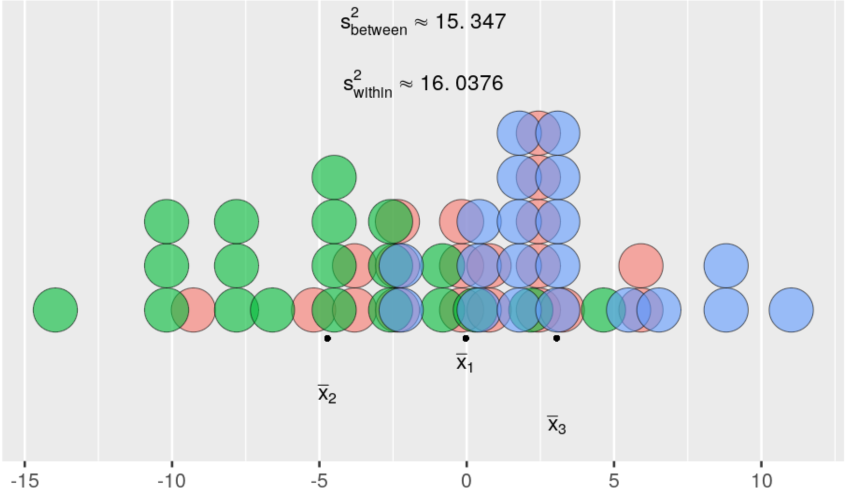
Low variance between groups
Low variance within groups




$s^2_{between}/s^2_{within} \approx 1 \implies$ Hard to tell if factor A has a significant effect...

# 1-Factor Analysis of Variance (Motivation)

High variance between groups
High variance within groups




$s^2_{between}/s^2_{within} \approx 1 \implies$ Hard to tell if factor A has a significant effect...

# 1-Factor ANOVA Basic Model Assumptions

In order for the forthcoming ANOVA test to bear good statistical properties and to utilize the Gauss-Markov Theorem, certain assumptions regarding the samples & populations must be imposed (similarly to $t$-tests & $F$-tests):

**Proposition**

*(1-Factor ANOVA Basic Model Assumptions)*

- *All measurements on units are independent.*
- *All groups are approximately normally distributed.*
- *All groups have approximately same variance.*

# 1-Factor ANOVA Test Statistic

The preceding four slides suggest the natural statistic is the $F$-Test Statistic:

## Proposition

*(Best Test Statistic Value for 1-Factor ANOVA♠♣)*

*Given an experiment with one factor and $I > 2$ groups.*
*Moreover, suppose the 1-factor basic ANOVA assumptions are all satisfied.*
*Then, the $F$-test using the following test statistic value:*

$$f = \frac{s^2_{between}}{s^2_{within}}$$

*is the most-powerful test that prevents $\alpha$-inflation for hypotheses:*

$$\begin{aligned} H_0 &: \quad \mu_1 = \mu_2 = \cdots = \mu_I \\ H_A &: \quad \textit{At least two of the } \mu\textit{'s differ} \end{aligned}$$

♠ R.A. Fisher, "The Correlation between Relatives on the Supposition of Mendelian Inheritance", *Transactions of the Royal Society of Edinburgh*, **52** (1918), 399-433.
♣ R.A. Fisher, *Statistical Methods for Research Workers*, 1925. (Ch VII)

# $s^2_{between}$ in terms of a Mean Square & Sum of Squares

## Proposition

*Given a 1-factor experiment involving $I$ groups, each of size $J$.*

*Then the variance between groups, $s^2_{between}$, is the variance of sample consisting of the $I$ group means $\overline{x}_{i\bullet}$ scaled by common treatment size $J$:*

$$s^2_{between} := \frac{J \cdot \sum_i (\overline{x}_{i\bullet} - \overline{x}_{\bullet\bullet})^2}{I-1} = \frac{\sum_i \sum_j (\hat{\alpha}_i^A)^2}{I-1} := \frac{SS_A}{\nu_A} := MS_A$$

*where the **grand mean**, $\overline{x}_{\bullet\bullet}$, is the mean of the $I$ group means, $\overline{x}_{i\bullet}$:*

$$\overline{x}_{\bullet\bullet} := \tfrac{1}{I} \sum_i \overline{x}_{i\bullet} = \tfrac{1}{IJ} \sum_i \sum_j x_{ij}$$

*In essence, a large variance between groups indicates much of the observed variation is explained by the chosen Factor A – hence, subscript $A$.*

$SS_A$ and $\nu_A$ are used later for computing $F$-cutoffs/$P$-values and interpretation.

# $s^2_{within}$ in terms of a Mean Square & Sum of Squares

## Proposition

*Given a 1-factor experiment involving $I$ groups, each of size $J$.*

*Then the variance within groups, $s^2_{within}$, is the mean of the group variances:*

$$s^2_{within} := \frac{(J-1)\cdot\sum_i s_i^2}{I(J-1)} = \frac{\sum_i\sum_j (x_{ij}-\bar{x}_{i\bullet})^2}{I(J-1)} = \frac{\sum_i\sum_j (x_{ij}^{res})^2}{I(J-1)} := \frac{SS_{res}}{\nu_{res}} := MS_{res}$$

*Effectively, a large variance within the groups indicates that much of the observed variation is not explained by the chosen Factor A. Therefore, the within variance is considered unexplained error in the experimental design.*

$SS_{res}$ and $\nu_{res}$ are used later for finding $F$-cutoffs/$P$-values and interpretation.

# $F$-Test Statistic Value in terms of Mean Squares

We can now express the $F$-Test statistic value in terms of mean squares:

## Proposition

*(Test Statistic Value for 1-Factor ANOVA in terms of Mean Squares)*

$$f_A = \frac{s^2_{between}}{s^2_{within}} = \frac{MS_A}{MS_{res}}$$

*The test statistic value for 1-Factor ANOVA will be denoted $f_A$ instead of $f$.*

In terms of the $F$-test notation in section 9.5, $f_A$ is always $f_+$.
The following slides explain why this is always the case for ANOVA.

PART V:

1-Factor Balanced Completely Randomized ANOVA
(1F bcrANOVA)

1-Factor Balanced Completely Randomized Design

Fixed Effects Model Assumptions

Fixed Effects Linear Model

Sums of Squares

$F$-Test Procedure

Expected Mean Squares

Point Estimators of $\sigma^2$

# 1-Factor Balanced Completely Randomized Design

An example **balanced completely randomized design** entails:

- Collect $12$ relevant experimental units (EU's): $\text{EU}_1, \text{EU}_2, \cdots, \text{EU}_{12}$
- Produce a random shuffle sequence using software: $(4, 12, 5, 10;\ 7, 2, 1, 11;\ 3, 6, 8, 9)$
- Use random shuffle sequence to assign the EU's into the $I$ levels:

| FACTOR A: | MEASUREMENTS: | | | |
|---|---|---|---|---|
| Level $1$ | $\text{EU}_4,$ | $\text{EU}_{12},$ | $\text{EU}_5,$ | $\text{EU}_{10}$ |
| Level $2$ | $\text{EU}_7,$ | $\text{EU}_2,$ | $\text{EU}_1,$ | $\text{EU}_{11}$ |
| Level $3$ | $\text{EU}_3,$ | $\text{EU}_6,$ | $\text{EU}_8,$ | $\text{EU}_9$ |

- Measure each EU appropriately (note the change in notation):

| FACTOR A: | MEASUREMENTS: | | | |
|---|---|---|---|---|
| Level $1$ $(x_{1\bullet})$ | $x_{11},$ | $x_{12},$ | $x_{13},$ | $x_{14}$ |
| Level $2$ $(x_{2\bullet})$ | $x_{21},$ | $x_{22},$ | $x_{23},$ | $x_{24}$ |
| Level $3$ $(x_{3\bullet})$ | $x_{31},$ | $x_{32},$ | $x_{33},$ | $x_{34}$ |

$$
\begin{aligned}
\text{EU}_k &\equiv \left(k^{th} \text{ experimental unit collected}\right) \\
x_{ij} &\equiv \left(\text{Measurement of } j^{th} \text{ experimental unit in } i^{th} \text{ level}\right) \\
x_{i\bullet} &\equiv \left(\text{Group of all measurements in } i^{th} \text{ level}\right)
\end{aligned}
$$

# How to Produce Random Shuffle Sequence

How to produce random shuffle sequence of numbers 1 through $N$:

| LANGUAGE: | MINIMUM CODE: |
|---|---|
| Matlab | `s=1:`$N$`;`<br>`s(randperm(length(s)))` |
| Python | `import random`<br>`random.sample(range(1,`$N$`+1),`$N$`)` |
| R | `sample(`$N$`)` |

# 1-Factor ANOVA Fixed Effects Model Assumptions

**Fixed effects** means all relevant levels of factor A are considered in model.

## Proposition

*(1F bcrANOVA Fixed Effects Model Assumptions)*

- *(**1 Desired Factor**) Factor A has $I$ levels.*
- *(**All Factor Levels are Considered**) AKA Fixed Effects.*
- *(**Balanced Replication in Groups**) Each group has $J > 1$ units.*
- *(**Distinct Exp. Units** ) All $IJ$ units are distinct from each other.*

---

- *(**Random Assignment across Groups**)*

---

- *(**Independence**) All measurements on units are independent.*
- *(**Normality**) All groups are approximately normally distributed.*
- *(**Equal Variances**) All groups have approximately same variance.*

*Mnemonic:* ***1DF AFLaC BRiG DEU*** | ***RAaG*** | ***I.N.EV***

**Fixed effects** means all relevant levels of factor A are considered in model.

**1F bcrANOVA Fixed Effects Linear Model**

| | | |
|---|---|---|
| $I$ | $\equiv$ | # groups to compare |
| $J$ | $\equiv$ | # measurements in each group |
| $X_{ij}$ | $\equiv$ | rv for $j^{th}$ measurement taken from $i^{th}$ group |
| $\mu_i$ | $\equiv$ | Mean of $i^{th}$ population or true average response from $i^{th}$ group |
| $\mu$ | $\equiv$ | Common population mean or true average overall response |
| $\alpha_i^A$ | $\equiv$ | Deviation from $\mu$ due to $i^{th}$ group |
| $E_{ij}$ | $\equiv$ | Deviation from $\mu$ due to random error |

ASSUMPTIONS: $E_{ij} \overset{iid}{\sim}$ Normal $\left(0, \sigma^2\right)$

$$X_{ij} = \mu + \alpha_i^A + E_{ij} \quad \text{where} \quad \textstyle\sum_i \alpha_i^A = 0$$

$H_0^A :$ All $\alpha_i^A = 0$

$H_A^A :$ Some $\alpha_i^A \neq 0$

$X_{ij} \overset{IND}{\sim} \ldots \equiv$ rv's $X_{ij}$ are independently distributed as ...

$E_{ij} \overset{iid}{\sim} \ldots \equiv$ rv's $E_{ij}$ are independently and identically distributed as ...

# 1F bcrANOVA (Sums of Squares "Partition" Variation)

$$\underbrace{\mathsf{SS}_{total}}_{Total\ Variation\ in\ Experiment} = \underbrace{\mathsf{SS}_{A}}_{Variation\ due\ to\ Factor\ A} + \underbrace{\mathsf{SS}_{res}}_{Unexplained\ Variation}$$

$$\sum_{ij}(x_{ij} - \hat{\mu})^2 = \sum_{ij}(\hat{\alpha}_i^A)^2 + \sum_{ij}(x_{ij}^{res})^2$$

$$\sum_i \sum_j (x_{ij} - \bar{x}_{\bullet\bullet})^2 = \sum_i \sum_j (\bar{x}_{i\bullet} - \bar{x}_{\bullet\bullet})^2 + \sum_i \sum_j (x_{ij} - \bar{x}_{i\bullet})^2$$

---

$$\underbrace{\nu}_{Total\ dof's\ in\ Experiment} = \underbrace{\nu_A}_{'Between\ Groups'\ dof's} + \underbrace{\nu_{res}}_{'Within\ Groups'\ dof's}$$

$$\nu = IJ - 1 \qquad \nu_A = I - 1 \qquad \nu_{res} = I(J-1)$$

# 1F bcrANOVA $F$-Test (Given Means $\overline{x}_{i\bullet}$ & SD's $s_i$)

1. Determine df's: $n = IJ, \quad \nu_A = I - 1, \quad \nu_{res} = I(J-1)$
2. Compute Grand Mean: $\overline{x}_{\bullet\bullet} = \frac{1}{I}\sum_i \overline{x}_{i\bullet}$
3. Compute $\text{SS}_{res} := \sum_{ij}(x_{ij}^{res})^2 = (J-1)\cdot\sum_i s_i^2$
4. Compute $\text{SS}_A \quad := \sum_{ij}(\hat{\alpha}_i^A)^2 = J\cdot\sum_i(\overline{x}_{i\bullet} - \overline{x}_{\bullet\bullet})^2$
5. Compute Mean Squares: $\text{MS}_{res} := \frac{\text{SS}_{res}}{\nu_{res}}, \quad \text{MS}_A := \frac{\text{SS}_A}{\nu_A}$
6. Compute Test Statistic Value: $f_A = \frac{\text{MS}_A}{\text{MS}_{res}}$
7. Compute $F$-cutoff/P-value:

   | | |
   |---|---|
   | By hand, lookup | $f^*_{\nu_A,\nu_{res};\alpha}$ |
   | By SW, compute | $p_A = 1 - \Phi_F(f_A; \nu_A, \nu_{res})$ |

8. Render Decision:

   | | | |
   |---|---|---|
   | If | $f_A \geq f^*_{\nu_A,\nu_{res};\alpha}$ | , then reject $H_0^A$; else accept $H_0^A$. |
   | If | $p_A \leq \alpha$ | , then reject $H_0^A$; else accept $H_0^A$. |

1. Determine df's: $n = IJ, \quad \nu_A = I - 1, \quad \nu_{res} = I(J-1)$
2. Compute Grand Mean: $\bar{x}_{\bullet\bullet} = \frac{1}{I}\sum_i \bar{x}_{i\bullet}$
3. Compute Group Std. Dev's: $s_i = \sqrt{J} \cdot \widehat{\sigma}_{\bar{x}_{i\bullet}}$
4. Compute $\text{SS}_{res} := \sum_{ij}(x_{ij}^{res})^2 = (J-1) \cdot \sum_i s_i^2$
5. Compute $\text{SS}_A := \sum_{ij}(\hat{\alpha}_i^A)^2 = J \cdot \sum_i (\bar{x}_{i\bullet} - \bar{x}_{\bullet\bullet})^2$
6. Compute Mean Squares: $\text{MS}_{res} := \frac{\text{SS}_{res}}{\nu_{res}}, \quad \text{MS}_A := \frac{\text{SS}_A}{\nu_A}$
7. Compute Test Statistic Value: $f_A = \frac{\text{MS}_A}{\text{MS}_{res}}$
8. Compute $F$-cutoff/P-value:
   - By hand, lookup $f^*_{\nu_A, \nu_{res}; \alpha}$
   - By SW, compute $p_A = 1 - \Phi_F(f_A; \nu_A, \nu_{res})$
9. Render Decision:
   - If $f_A \geq f^*_{\nu_A, \nu_{res}; \alpha}$, then reject $H_0^A$; else accept $H_0^A$.
   - If $p_A \leq \alpha$, then reject $H_0^A$; else accept $H_0^A$.

# 1F bcrANOVA $F$-Test (Given Observations $x_{ij}$)

1. Determine df's: $n = IJ, \quad \nu_A = I - 1, \quad \nu_{res} = I(J-1)$
2. Compute Group Means: $\bar{x}_{i\bullet} := \frac{1}{J}\sum_j x_{ij}$
3. Compute Group Variances: $s_i^2 := \frac{1}{J-1}\sum_j (x_{ij} - \bar{x}_{i\bullet})^2$
4. Compute Grand Mean: $\bar{x}_{\bullet\bullet} = \frac{1}{I}\sum_i \bar{x}_{i\bullet}$
5. Compute $\mathrm{SS}_{res} := \sum_{ij}(x_{ij}^{res})^2 = (J-1)\cdot\sum_i s_i^2$
6. Compute $\mathrm{SS}_A := \sum_{ij}(\hat{\alpha}_i^A)^2 = J\cdot\sum_i(\bar{x}_{i\bullet} - \bar{x}_{\bullet\bullet})^2$
7. Compute Mean Squares: $\mathrm{MS}_{res} := \frac{\mathrm{SS}_{res}}{\nu_{res}}, \quad \mathrm{MS}_A := \frac{\mathrm{SS}_A}{\nu_A}$
8. Compute Test Statistic Value: $f_A = \frac{\mathrm{MS}_A}{\mathrm{MS}_{res}}$
9. Compute $F$-cutoff/P-value:
   By hand, lookup $f^*_{\nu_A,\nu_{res};\alpha}$
   By SW, compute $p_A = 1 - \Phi_F(f_A; \nu_A, \nu_{res})$
10. Render Decision:
    If $f_A \geq f^*_{\nu_A,\nu_{res};\alpha}$, then reject $H_0^A$; else accept $H_0^A$.
    If $p_A \leq \alpha$, then reject $H_0^A$; else accept $H_0^A$.

# 1F bcrANOVA $F$-Test (Summary Table)

| 1-Factor ANOVA Table (Significance Level $\alpha$) | | | | | | |
|---|---|---|---|---|---|---|
| **Variation Source** | **df** | **Sum of Squares** | **Mean Square** | **$F$ Stat Value** | **P-value** | **Decision** |
| Factor A | $\nu_A$ | $SS_A$ | $MS_A$ | $f_A$ | $p_A$ | Acc/Rej $H_0^A$ |
| Unknown | $\nu_{res}$ | $SS_{res}$ | $MS_{res}$ | | | |
| Total | $\nu$ | $SS_{total}$ | | | | |

# 1F bcrANOVA (Expected Mean Squares)

## Proposition

*Given 1-factor experiment satisfying the 1F bcrANOVA assumptions. Then:*

$$(i) \quad \mathbb{E}[MS_{res}] = \sigma^2$$

$$(ii) \quad \mathbb{E}[MS_A] = \sigma^2 + \frac{J}{I-1}\sum_i (\alpha_i^A)^2$$

For the proof of part (i): There's nothing too terribly tricky involved.
For the proof of part (ii):

We will proceed by producing a simplified expression for $\mathbb{E}[\mathrm{MS}_A]$ in terms of $\sigma^2$ and $\alpha_i^A$ using the error group means $\overline{E}_{i\bullet}$ and the grand error mean $\overline{E}_{\bullet\bullet}$ as was done in Hays' statistics textbook†.

†W.L. Hays, *Statistics*, $5^{th}$ Edition, 1994.

Alternatives involve tricky uses of covariance and/or tedious determinations of the distributions of the squares of the means (which are Gamma distributions.)

# 1F bcrANOVA Expected Mean Squares: Proof of (i)

$$
\begin{aligned}
\mathbb{E}[\mathsf{SS}_{res}] &:= \mathbb{E}\left[\sum_{ij}(X_{ij}^{res})^2\right] \\
&= \mathbb{E}\left[\sum_i \sum_j (X_{ij} - \hat{X}_{ij})^2\right] \\
&= \mathbb{E}\left[\sum_i \sum_j \left(X_{ij} - (\hat{\mu} + \hat{\alpha}_i^A)\right)^2\right] \\
&\overset{BLUE}{=} \mathbb{E}\left[\sum_i \sum_j (X_{ij} - \overline{X}_{i\bullet})^2\right] \\
&\overset{CIO}{=} \tfrac{J-1}{J-1} \cdot \mathbb{E}\left[\sum_i \sum_j (X_{ij} - \overline{X}_{i\bullet})^2\right] \\
&= (J-1) \cdot \sum_i \mathbb{E}\left[\tfrac{1}{J-1} \sum_j (X_{ij} - \overline{X}_{i\bullet})^2\right] \\
&= (J-1) \cdot \sum_i \mathbb{E}\left[S_i^2\right] = (J-1) \cdot \sum_i \sigma^2 \\
&= I(J-1)\sigma^2
\end{aligned}
$$

$$
\Longrightarrow \mathbb{E}\left[\mathsf{MS}_{res}\right] := \mathbb{E}\left[\frac{\mathsf{SS}_{res}}{\nu_{res}}\right] = \frac{\mathbb{E}\left[\mathsf{SS}_{res}\right]}{I(J-1)} = \frac{I(J-1)\sigma^2}{I(J-1)} = \sigma^2 \qquad \square
$$

---

CIO $\equiv$ "**C**lever **I**nsertion of **O**ne"

# 1F bcrANOVA Expected Mean Squares: Proof of (ii)

$$\begin{array}{lcll}
\text{Given} & X_{ij} = \mu + \alpha_i^A + E_{ij} & \text{s.t.} & E_{ij} \overset{IND}{\sim} \text{Normal}(0, \sigma^2) \;\; \& \; \sum_i \alpha_i^A = 0 \\
\implies & \overline{X}_{i\bullet} = \mu + \alpha_i^A + \overline{E}_{i\bullet} & \overset{CLT}{\implies} & \overline{E}_{i\bullet} \overset{IND}{\sim} \text{Normal}(0, \frac{\sigma^2}{J}) \\
\implies & \overline{X}_{\bullet\bullet} = \mu + \overline{E}_{\bullet\bullet} & \overset{CLT}{\implies} & \overline{E}_{\bullet\bullet} \sim \text{Normal}(0, \frac{\sigma^2}{IJ})
\end{array}$$

$$\begin{aligned}
\mathbb{E}[\text{SS}_A] &:= \mathbb{E}\left[\textstyle\sum_{ij} (\hat{\alpha}_i^A)^2\right] \overset{BLUE}{=} \textstyle\sum_i \sum_j \mathbb{E}\left[(\overline{X}_{i\bullet} - \overline{X}_{\bullet\bullet})^2\right] \\
&= \textstyle\sum_i \sum_j \mathbb{E}\left[(\alpha_i^A + \overline{E}_{i\bullet} - \overline{E}_{\bullet\bullet})^2\right] \\
&\overset{(1)}{=} J \cdot \textstyle\sum_i \mathbb{E}\left[(\alpha_i^A)^2\right] + J \cdot \sum_i \mathbb{E}\left[(\overline{E}_{i\bullet})^2 - 2(\overline{E}_{i\bullet}\overline{E}_{\bullet\bullet}) + (\overline{E}_{\bullet\bullet})^2\right] \\
&\overset{(2)}{=} J \cdot \textstyle\sum_i (\alpha_i^A)^2 + J \cdot \sum_i \mathbb{E}\left[(\overline{E}_{i\bullet})^2\right] + \mathbb{E}\left[-IJ(\overline{E}_{\bullet\bullet})^2\right] \\
&\overset{(3)}{=} J \cdot \textstyle\sum_i (\alpha_i^A)^2 + J \cdot \sum_i \left[\left(\mathbb{E}\left[\overline{E}_{i\bullet}\right]\right)^2 + \mathbb{V}\left[\overline{E}_{i\bullet}\right]\right] + \mathbb{E}\left[-IJ(\overline{E}_{\bullet\bullet})^2\right] \\
&= J \cdot \textstyle\sum_i (\alpha_i^A)^2 + J \cdot \sum_i \left[(0)^2 + \frac{\sigma^2}{J}\right] - IJ \cdot \mathbb{E}\left[(\overline{E}_{\bullet\bullet})^2\right] \\
&\overset{(3)}{=} J \cdot \textstyle\sum_i (\alpha_i^A)^2 + I\sigma^2 - IJ \cdot \left(\left(\mathbb{E}\left[\overline{E}_{\bullet\bullet}\right]\right)^2 + \mathbb{V}\left[\overline{E}_{\bullet\bullet}\right]\right) \\
&= J \cdot \textstyle\sum_i (\alpha_i^A)^2 + I\sigma^2 - IJ \cdot \left((0)^2 + \frac{\sigma^2}{IJ}\right) \\
&= J \cdot \textstyle\sum_i (\alpha_i^A)^2 + (I-1)\sigma^2
\end{aligned}$$

---

(1) $\sum_i (\overline{E}_{i\bullet} - \overline{E}_{\bullet\bullet}) = 0$ (2) $\sum_i \overline{E}_{i\bullet} = I \cdot \overline{E}_{\bullet\bullet}$ (3) $\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

# 1F bcrANOVA Expected Mean Squares: Proof of (ii)

$\mathbb{E}[\mathsf{SS}_A] = (I-1)\sigma^2 + J \cdot \sum_i (\alpha_i^A)^2$

$$\implies \mathbb{E}[\mathsf{MS}_A] := \mathbb{E}\left[\frac{\mathsf{SS}_A}{\nu_A}\right] = \frac{\mathbb{E}[\mathsf{SS}_A]}{I-1} = \frac{(I-1)\sigma^2 + J \cdot \sum_i (\alpha_i^A)^2}{I-1}$$

$$\therefore \ \mathbb{E}[\mathsf{MS}_A] = \sigma^2 + \frac{J}{I-1} \cdot \sum_i (\alpha_i^A)^2 \qquad \square$$

## Proposition

*(Point Estimation of Mean Squares)*
*Given 1-factor balanced experiment satisfying ANOVA assumptions. Then:*

*(i) $MS_{res}$ is always an unbiased point estimator of the population variance:*

$$H_0 \text{ is indeed true OR } H_0 \text{ is indeed false} \implies \mathbb{E}[MS_{res}] = \sigma^2$$

*(ii) If the status quo prevails, $MS_A$ is an unbiased estimator of pop. variance:*

$$H_0 \text{ is indeed true} \implies \mathbb{E}[MS_A] = \sigma^2$$

*(iii) If the status quo fails, $MS_A$ tends to overestimate population variance:*

$$H_0 \text{ is indeed false} \implies \mathbb{E}[MS_A] > \sigma^2$$

PROOF OF PART (i):

Follows from part (i) of Excepted Mean Squares proposition.

# $\mathrm{MS}_A$ as Point Estimator of $\sigma^2$: Proof of (ii) & (iii)

## Proposition

*(Point Estimation of Mean Squares)*

*Given $I$ size-$J$ random samples satisfying ANOVA assumptions. Then:*

*(ii)* $H_0$ *is indeed true* $\implies \mathbb{E}[MS_A] = \sigma^2$

*(iii)* $H_0$ *is indeed false* $\implies \mathbb{E}[MS_A] > \sigma^2$

From the Expected Mean Squares proposition, $\mathbb{E}[\mathrm{MS}_A] = \sigma^2 + \frac{J}{I-1} \cdot \sum_i (\alpha_i^A)^2$

(ii) $H_0$ is true

$$\begin{aligned}
&\implies \mu_1 = \mu_2 = \cdots = \mu_I \\
&\implies \mu = \mu_1 = \cdots = \mu_I && \text{(Since } \mu := \tfrac{1}{I}\textstyle\sum_i \mu_i) \\
&\implies \alpha_1^A = \alpha_2^A = \cdots = \alpha_I^A = 0 && \text{(Since } \alpha_i^A := \mu_i - \mu) \\
&\implies \mathbb{E}[\mathrm{MS}_A] = \sigma^2 \qquad \square
\end{aligned}$$

(iii) $H_0$ is false

$$\begin{aligned}
&\implies \text{At least two of the } \mu\text{'s differ} \\
&\implies \text{At least two of the } \alpha^A\text{'s } \neq 0 \\
&\implies \textstyle\sum_i (\alpha_i^A)^2 > 0 \\
&\implies \mathbb{E}[\mathrm{MS}_A] > \sigma^2 \qquad \square
\end{aligned}$$

PART VI:

Effect Size Measures for 1-Factor ANOVA:

Fisher $(\hat{\eta}_A^2)$, Kelley $(\hat{\epsilon}_A^2)$, Hays $(\hat{\omega}_A^2)$

# 1-Factor ANOVA (Effect Size Measures)

Recall that when performing a hypothesis test of any kind, statistical significance does not necessarily imply practical significance.

As Gravetter & Wallnau put it in §13.5 of their statistics textbook[GW]:

> "the term *significant* does not necessarily mean *large*,
> it simply means larger than expected by chance."

Q: How does one determine whether a statistically significant effect due to factor A in 1F ANOVA is a practical (i.e. large enough) effect??

A: Effect size measures! What follows are 3 such popular measures.

# 1-Factor ANOVA (Effect Size Measures)

| YEAR | NAME | MEASURE | HOW IT COMPARES* |
|---|---|---|---|
| 1925† | Fisher<br>$[GW],[H],[LH],[S]$ | $\hat{\eta}_A^2 := \frac{\mathrm{SS}_A}{\mathrm{SS}_{total}}$ | Most biased (positively)♠<br>Least SD, Most RMSE♠ |
| 1935‡ | Kelley | $\hat{\epsilon}_A^2 := \frac{\mathrm{SS}_A - \nu_A \mathrm{MS}_{res}}{\mathrm{SS}_{total}}$ | Least biased (negatively)♠<br>Most SD, Nearly Least RMSE♠ |
| 1963♣ | Hays<br>$[H],[LH],[S]$ | $\hat{\omega}_A^2 := \frac{\mathrm{SS}_A - \nu_A \mathrm{MS}_{res}}{\mathrm{SS}_{total} + \mathrm{MS}_{res}}$ | Moderately biased (negatively)♠<br>Moderate SD, Least RMSE♠ |

*Requires all 1F ANOVA assumptions (LADR'S RAIN EV) to be satisfied.
SD ≡ **S**tandard **D**eviation, RMSE ≡ **R**oot **M**ean **S**quared **E**rror

---

†R.A. Fisher, *Statistical Methods for Research Workers*, 1925. (Ch VIII, §45)

‡T.L. Kelley, "An Unbiased Correlation Ratio Measure", *Proceedings of the National Academy of Sciences*, **21** (1935), 554-559.

♣W.L. Hays, *Statistics for Psychologists*, 1963.

♠K. Okada, "Is Omega Squared Less Biased? A Comparison of Three Major Effect Size Indices in 1-Way ANOVA", *Behaviormetrika*, **40** (2013), 129-147.

# Effect Size Measures (General Remarks)

There are about 75 different effect size measures† that have been discovered!!

†R.E. Kirk, “The Importance of Effect Magnitude”, In S.F. Davis (Ed.), *Handbook of Research Methods in Experimental Psychology*, 2003.

Moreover, realize that many of these measures are ’measures of association’ and, hence, are tailored for either numerical-numerical (num-num) inference (Ch 12 & 13) or categorical-categorical (cat-cat) inference (Ch 14).

---

1. Cutoff values for “small”/“medium”/“large” effects vary by field[LH]:
   J. Cohen, *Statistical Power Analysis for Behavioral Sciences*, 1969. (§8.2)
2. Be very careful when interpreting values of effect size measures[S], especially for 2-Factor ANOVA or higher:
   - K.E. O’Grady, “Measures of Explained Variance: Cautions and Limitations”, *Psychological Bulletin*, **92** (1982), 766-777.
   - C.A. Pierce, R.A. Block, H. Aguinis, “Cautionary Note on Reporting Eta-Squared Values from Multifactor ANOVA Designs”, *Educational & Psychological Measurement*, **64** (2004), 916-924.

# References



| | | | | |
|---|---|---|---|---|
| $[GW]$ | F.J. Gravetter<br>L.B. Wallnau | *Statistics for the Behavioral Sciences* | $7^{th}$ Ed | 2007 |
| $[H]$ | D.C. Howell | *Statistical Methods for Psychology* | $7^{th}$ Ed | 2010 |
| $[LH]$ | R.G. Lomax<br>D.L. Hahs-Vaughn | *Statistical Concepts : A Second Course* | $4^{th}$ Ed | 2012 |
| $[S]$ | J.P. Stevens | *Intermediate Statistics A Modern Approach* | $3^{rd}$ Ed | 2007 |

# Textbook Logistics for Section 10.1

- Difference(s) in Terminology:

| TEXTBOOK TERMINOLOGY: | SLIDES/OUTLINE TERMINOLOGY: |
|---|---|
| Treatment/Cell | Group |

- Difference(s) in Notation:

| CONCEPT | TEXTBOOK NOTATION | SLIDES/OUTLINE NOTATION |
|---|---|---|
| Probability of Event | $P(E)$ | $\mathbb{P}(E)$ |
| Expected Value | $E(X)$ | $\mathbb{E}[X]$ |
| Variance | $V(X)$ | $\mathbb{V}[X]$ |
| Sum of Squares of Factor A | SSTr | $\text{SS}_A$ |
| Mean Square of Factor A | MSTr | $\text{MS}_A$ |
| Sum of Squares of Residuals | SSE | $\text{SS}_{res}$ |
| Mean Square of Residuals | MSE | $\text{MS}_{res}$ |
| Null Hypothesis for Factor A | $H_0$ | $H_0^A$ |
| Alt. Hypothesis for Factor A | $H_A$ | $H_A^A$ |

Fin.