Agricultural Econometrics

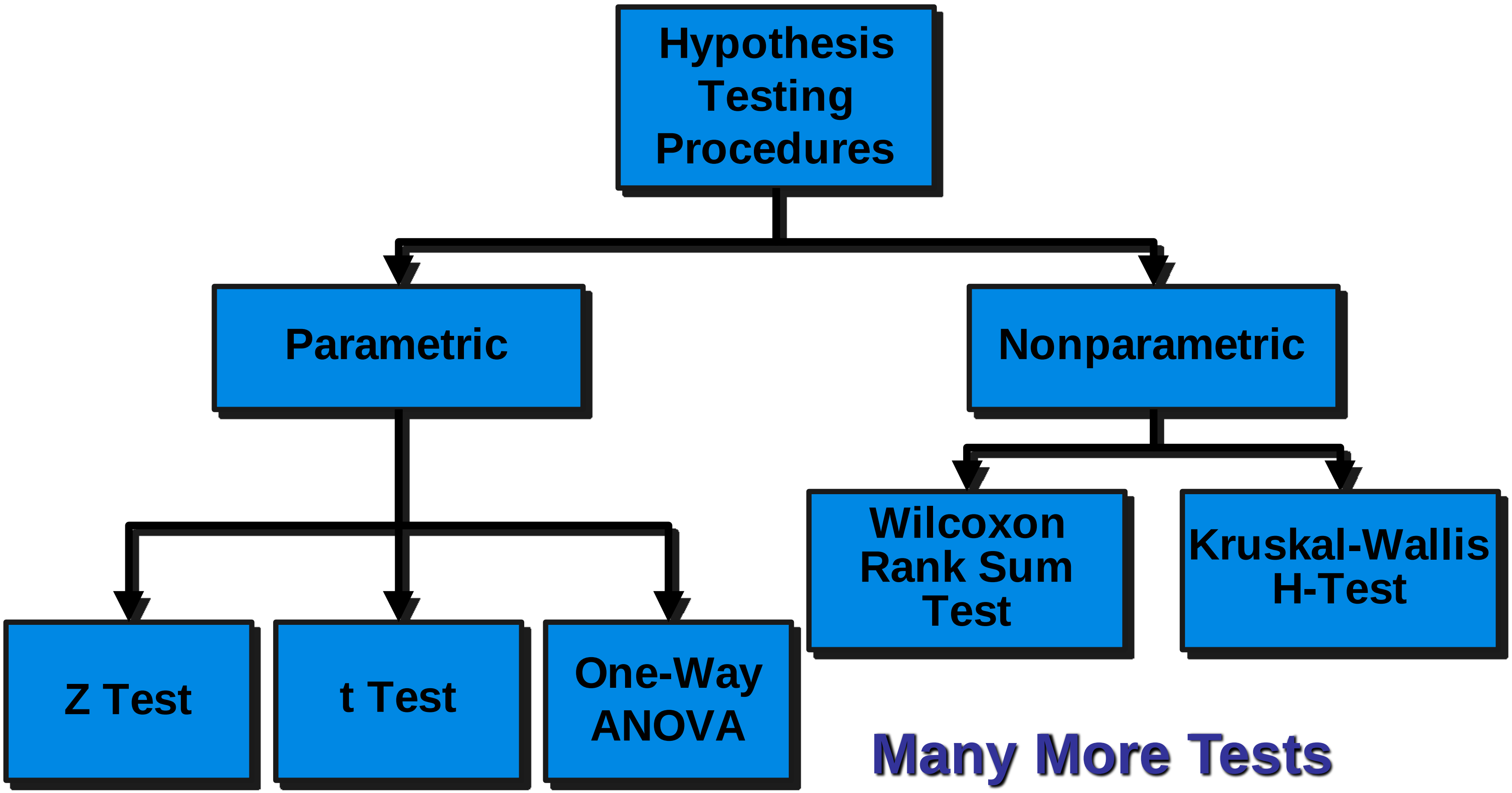
Chapter 1

1.1 Parametric Statistics : Z-test, T test
Analysis of Variance: One way , Two way

Learning Objectives

1. Distinguish Parametric & Nonparametric Test Procedures
2. Explain commonly used Nonparametric Test Procedures
3. Perform Hypothesis Tests Using Nonparametric Procedures

Hypothesis Testing Procedures



Many More Tests Exist!

Parametric Test Procedures

1. Involve Population Parameters (Mean)

2. Have Stringent Assumptions (Normality)

3. Examples: Z Test, t Test, χ^2 Test, F test

Nonparametric Test Procedures

1. Do Not Involve Population Parameters

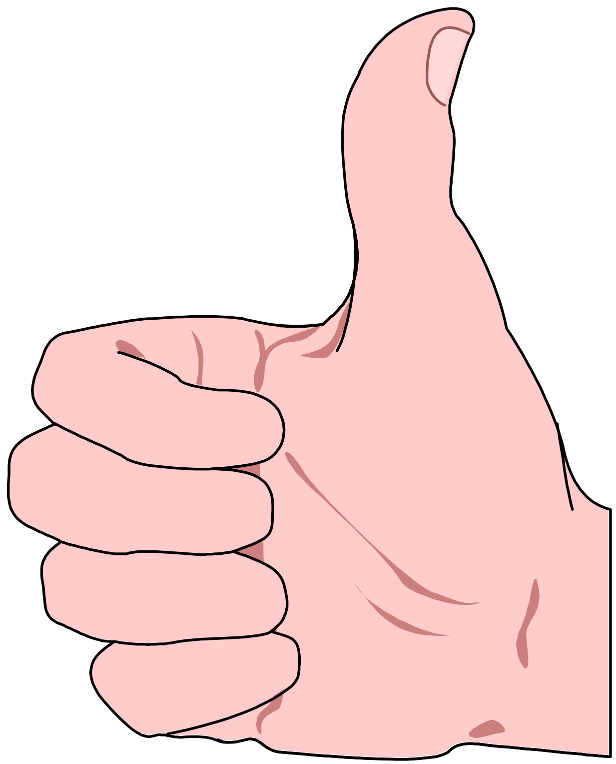
   Example: Probability Distributions, Independence

2. Data Measured on Any Scale (Ratio or Interval, Ordinal or Nominal)

3. Example: Wilcoxon Rank Sum Test

Advantages of Nonparametric Tests

1. Used With All Scales
2. Easier to Compute
3. Make Fewer Assumptions
4. Need Not Involve Population Parameters
5. Results May Be as Exact as Parametric Procedures



Disadvantages of Nonparametric Tests

1. May Waste Information

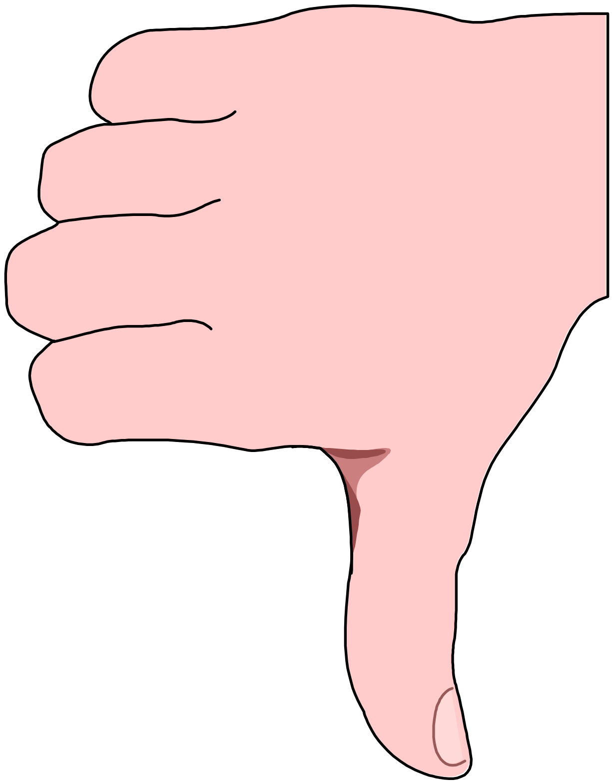
   Parametric model more efficient

   if data Permit

2. Difficult to Compute by

   hand for Large Samples

3. Tables Not Widely Available



Popular Nonparametric Tests

1. Sign Test

2. Wilcoxon Signed Rank Test

3. Man Whitney test

Agricultural Econometrics

Chapter 1

1.2 Analysis of Variance: One way Two way

Continuous outcome (means)

Outcome Variable	Are the observations independent or correlated?		Alternatives if the normality assumption is violated (and small sample size):
	independent	correlated	
Continuous (e.g. pain scale, cognitive function)	**Ttest:** compares means between two independent groups **ANOVA:** compares means between more than two independent groups **Pearson's correlation coefficient** (linear correlation): shows linear correlation between two continuous variables **Linear regression:** multivariate regression technique used when the outcome is continuous; gives slopes	**Paired ttest:** compares means between two related groups (e.g., the same subjects before and after) **Repeated-measures ANOVA:** compares changes over time in the means of two or more groups (repeated measurements) **Mixed models/GEE modeling**: multivariate regression techniques to compare changes over time between two or more groups; gives rate of change over time	Non-parametric statistics **Wilcoxon sign-rank test**: non-parametric alternative to the paired ttest **Wilcoxon sum-rank test** (=Mann-Whitney U test): non-parametric alternative to the ttest **Kruskal-Wallis test:** non-parametric alternative to ANOVA **Spearman rank correlation coefficient:** non-parametric alternative to Pearson's correlation coefficient

15.2

One-Way Analysis of Variance

- Evaluate the difference among the means of three or more groups

Examples:	Average production for 1^{st}, 2^{nd}, and 3^{rd} shifts
	Expected mileage for five brands of tires

- Assumptions
  - Populations are normally distributed
  - Populations have equal variances
  - Samples are randomly and independently drawn

Hypotheses of One-Way ANOVA

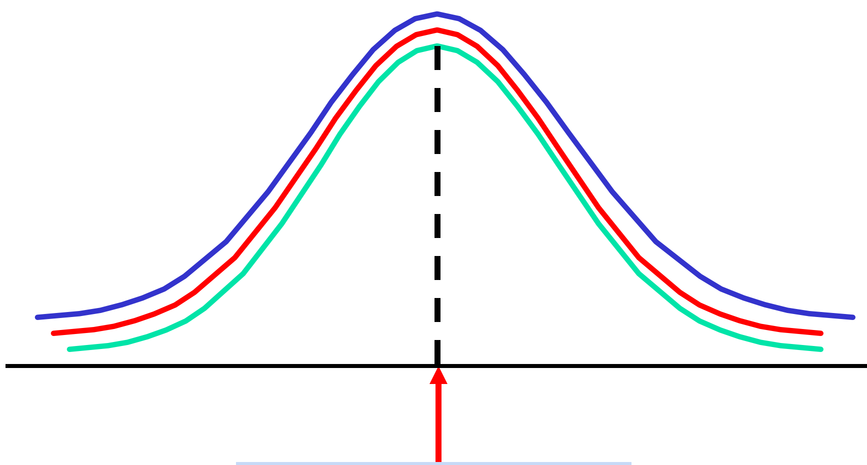
- $H_0 : \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_K$
  - All population means are equal
  - i.e., no variation in means between groups
- $H_1 : \mu_i \neq \mu_j \quad \text{for at least one } i, j \text{ pair}$
  - At least one population mean is different
  - i.e., there is variation between groups
  - Does not mean that all population means are different (some pairs may be the same)

One-Way ANOVA

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

$$H_1: \text{Not all } \mu_i \text{ are the same}$$

All Means are the same:
The Null Hypothesis is True
(No variation between groups)



$$\mu_1 = \mu_2 = \mu_3$$

One-Way ANOVA

(continued)

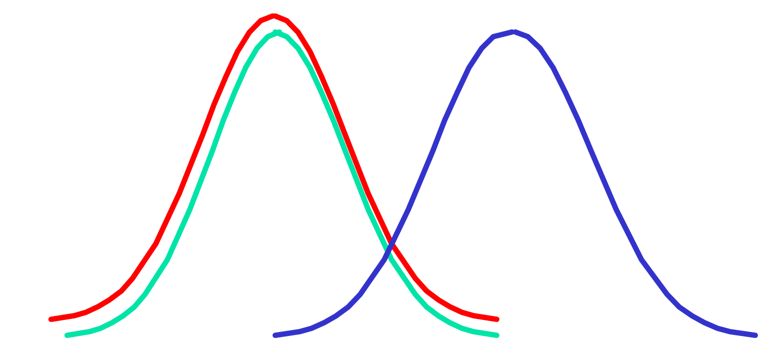
$$H_0 : \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_K$$

$$H_1 : \text{Not all } \mu_i \text{ are the same}$$

At least one mean is different:

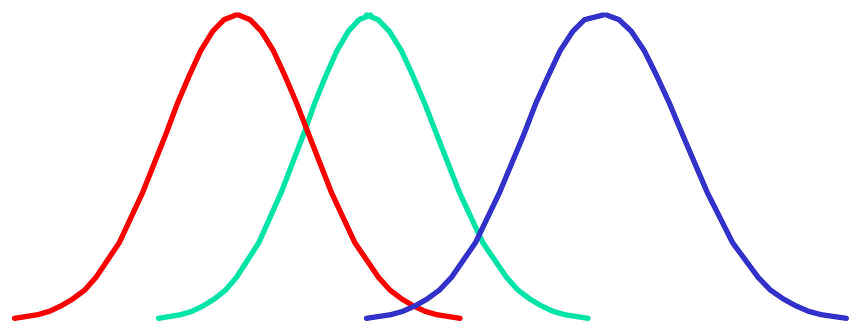
The Null Hypothesis is NOT true

(Variation is present between groups)



$$\mu_1 = \mu_2 \neq \mu_3$$

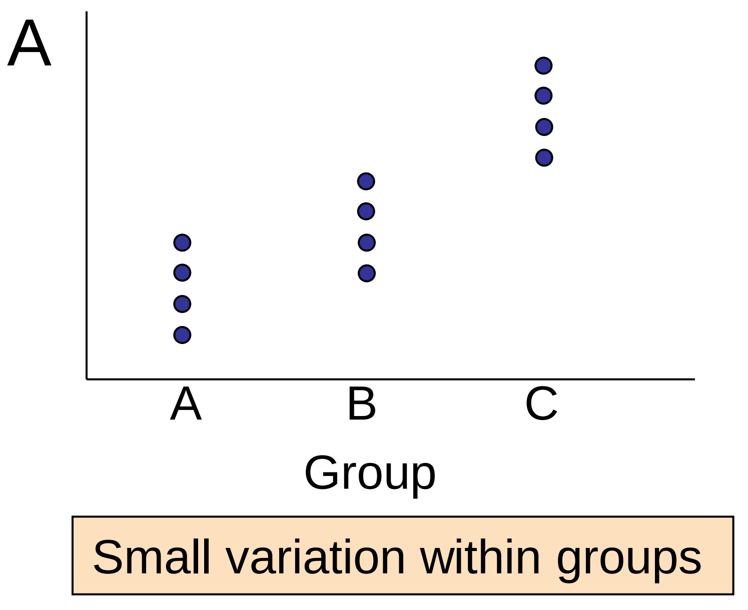
or

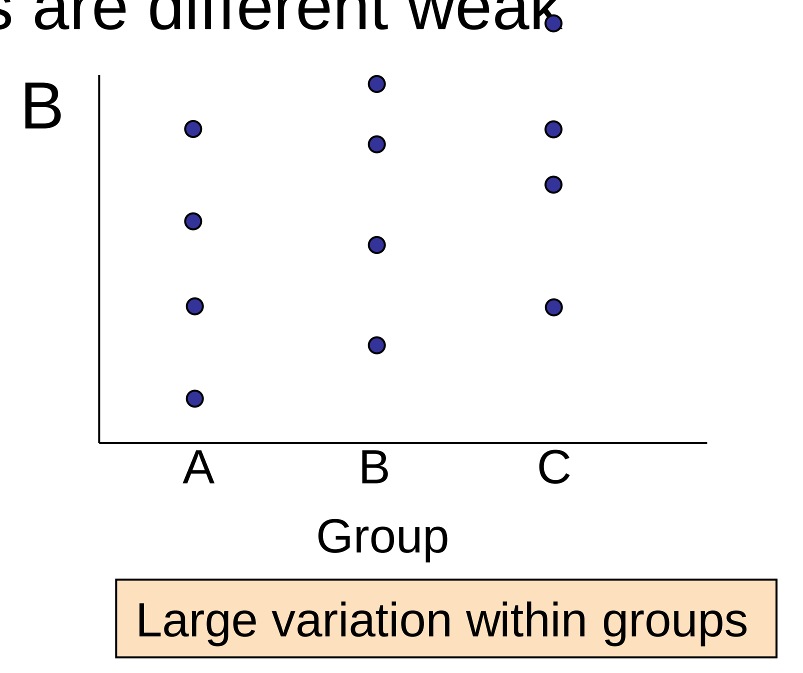


$$\mu_1 \neq \mu_2 \neq \mu_3$$

Variability

- The variability of the data is key factor to test the equality of means
- In each case below, the means may look different, but a large variation within groups in B makes the evidence that the means are different weak





Partitioning the Variation

- Total variation can be split into two parts:

SST = SSW + SSG

SST = Total Sum of Squares

Total Variation = the aggregate dispersion of the individual data values across the various groups

SSW = Sum of Squares Within Groups

Within-Group Variation = dispersion that exists among the data values within a particular group

SSG = Sum of Squares Between Groups

Between-Group Variation = dispersion between the group sample means

Partition of Total Variation

Total Sum of Squares (SST)

= **Variation due to random sampling (SSW)** + **Variation due to differences between groups (SSG)**

Total Sum of Squares

SST = SSW + SSG

$$SST = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$$

Where:

SST = Total sum of squares

K = number of groups (levels or treatments)

n_i = number of observations in group i

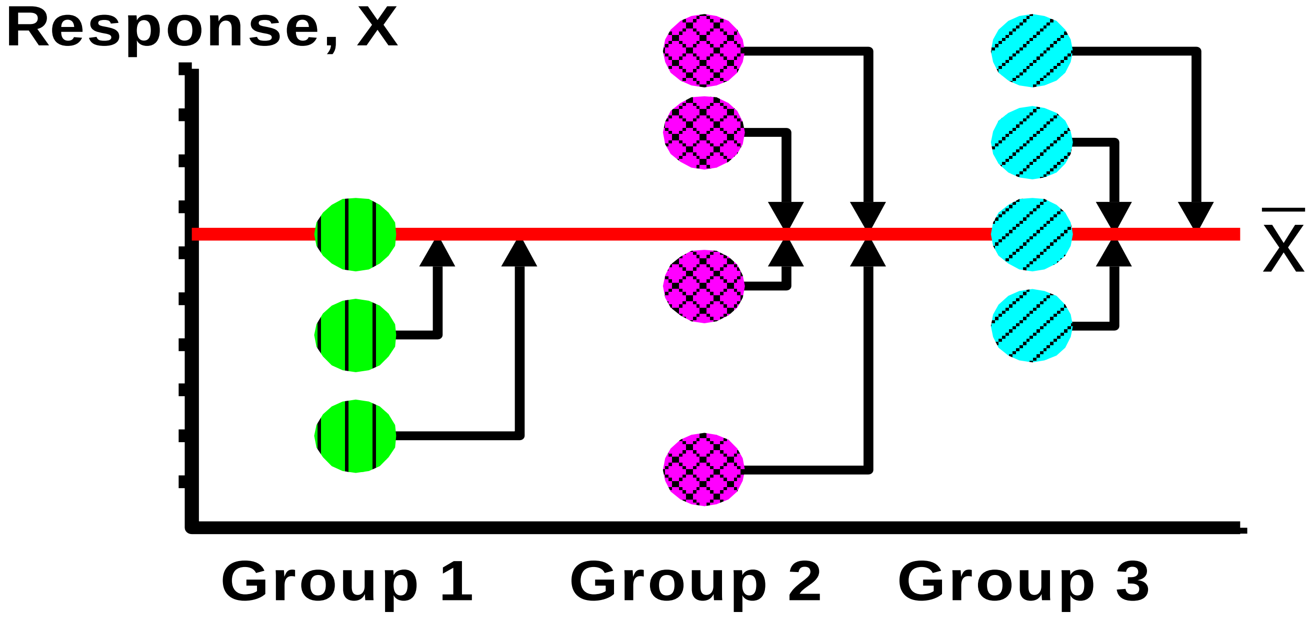
x_{ij} = j^{th} observation from group i

$\bar{x}$ = overall sample mean

Total Variation

(continued)

$$SST = (x_{11} - \bar{x})^2 + (X_{12} - \bar{x})^2 + ... + (x_{Kn_K} - \bar{x})^2$$



Within-Group Variation

SST = SSW + SSG

$$SSW = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

Where:

SSW = Sum of squares within groups

K = number of groups

n_i = sample size from group i

$\bar{x}_i$ = sample mean from group i

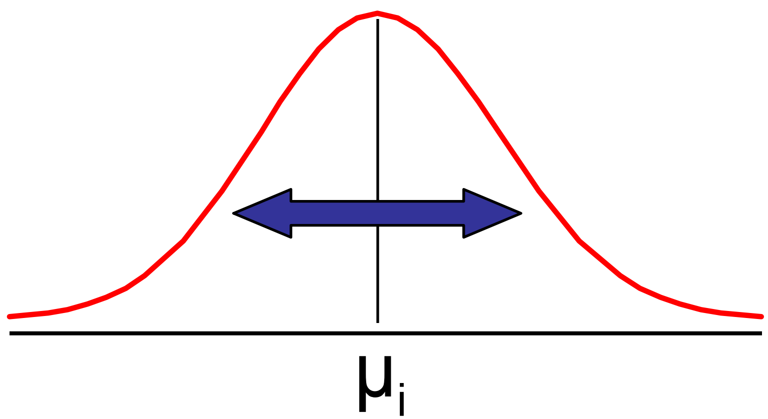
x_{ij} = j^{th} observation in group i

Within-Group Variation

(continued)

$$SSW = \sum_{i=1}^{K} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

Summing the variation within each group and then adding over all groups



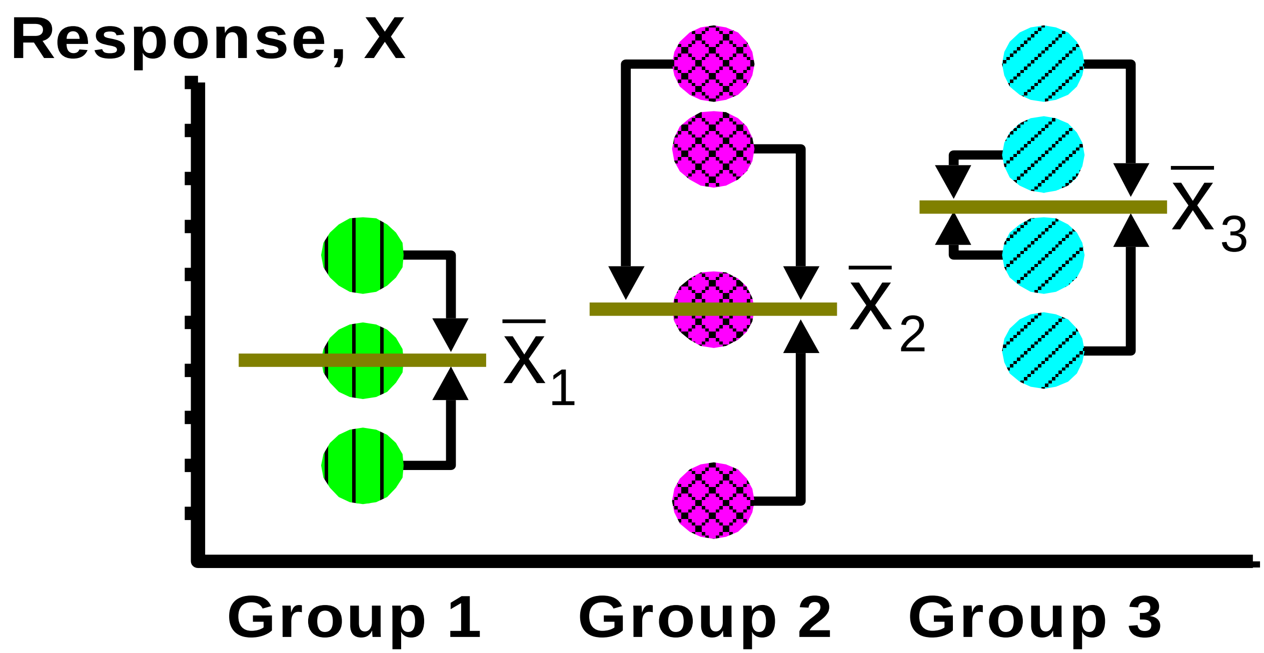
$$MSW = \frac{SSW}{n - K}$$

Mean Square Within = SSW/degrees of freedom

Within-Group Variation

(continued)

$$SSW = (x_{11} - \bar{x}_1)^2 + (x_{12} - \bar{x}_1)^2 + \ldots + (x_{Kn_K} - \bar{x}_K)^2$$



Between-Group Variation

SST = SSW + SSG

$$SSG = \sum_{i=1}^{K} n_i(\bar{x}_i - \bar{x})^2$$

Where:

SSG = Sum of squares between groups

K = number of groups

n_i = sample size from group i

$\bar{x}_i$ = sample mean from group i

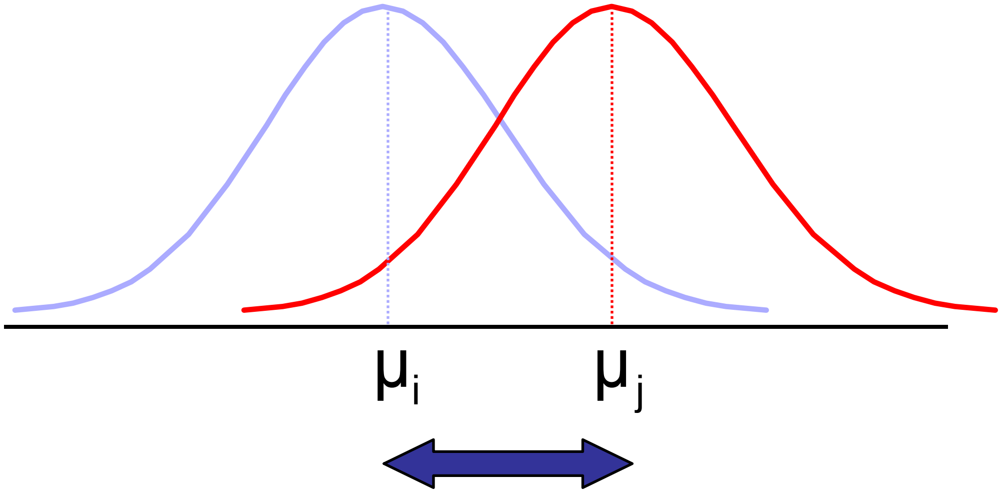
$\bar{x}$ = grand mean (mean of all data values)

Between-Group Variation

(continued)

$$SSG = \sum_{i=1}^{K} n_i(\bar{x}_i - \bar{x})^2$$

Variation Due to
Differences Between Groups



$$MSG = \frac{SSG}{K-1}$$

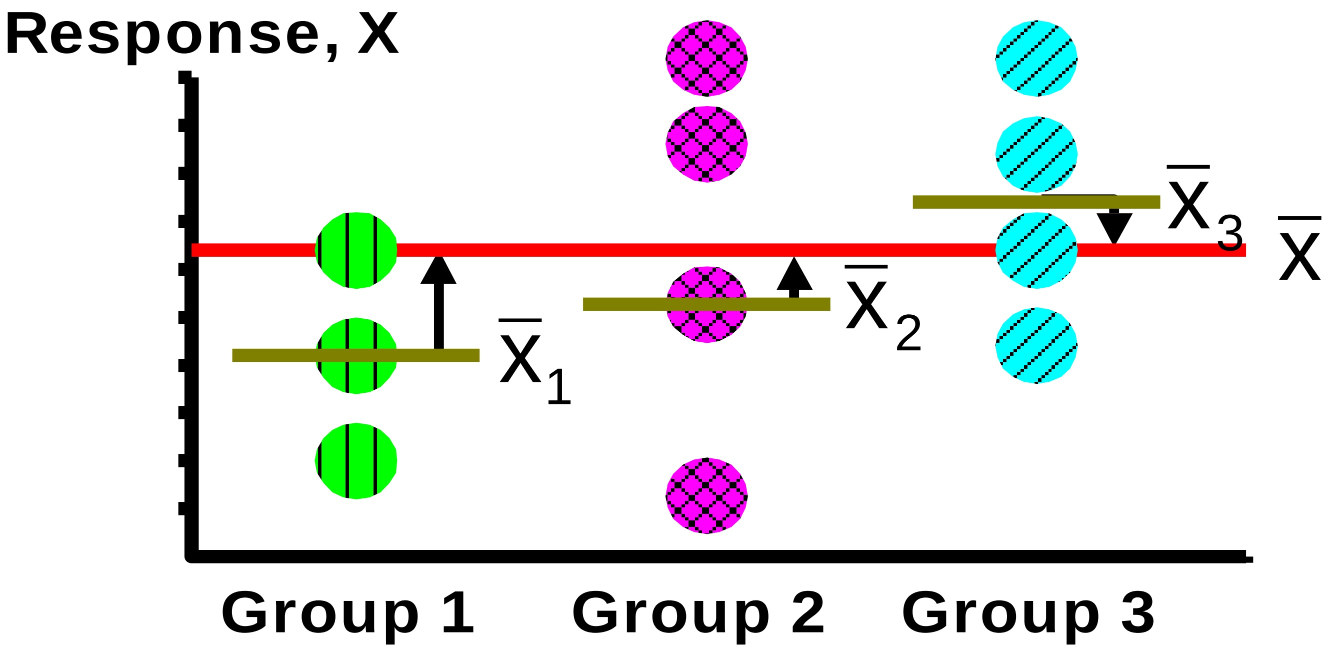
Mean Square Between Groups =

SSG/degrees of freedom

Between-Group Variation

(continued)

$$SSG = n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + \ldots + n_K(\bar{x}_K - \bar{x})^2$$



Obtaining the Mean Squares

$$MST = \frac{SST}{n-1}$$

$$MSW = \frac{SSW}{n-K}$$

$$MSG = \frac{SSG}{K-1}$$

One-Way ANOVA Table

Source of Variation	SS	df	MS (Variance)	F ratio
Between Groups	SSG	K - 1	$MSG = \frac{SSG}{K - 1}$	$F = \frac{MSG}{MSW}$
Within Groups	SSW	n - K	$MSW = \frac{SSW}{n - K}$	
Total	SST = SSG+SSW	n - 1		

K = number of groups

n = sum of the sample sizes from all groups

df = degrees of freedom

One-Factor ANOVA
F Test Statistic

H_0: $\mu_1 = \mu_2 = \ldots = \mu_K$

H_1: At least two population means are different

- Test statistic

$$F = \frac{MSG}{MSW}$$

MSG is mean squares between variances

MSW is mean squares within variances

- Degrees of freedom
  - $df_1 = K - 1$ (K = number of groups)
  - $df_2 = n - K$ (n = sum of sample sizes from all groups)

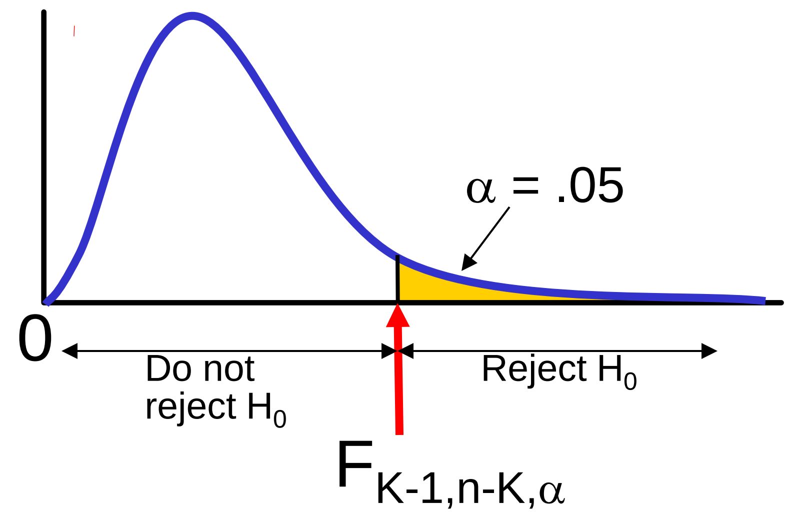
Interpreting the F Statistic

- The F statistic is the ratio of the between estimate of variance and the within estimate of variance
  - The ratio must always be positive
  - $df_1 = K - 1$ will typically be small
  - $df_2 = n - K$ will typically be large

Decision Rule:

- Reject H_0 if

  $F > F_{K-1,n-K,\alpha}$

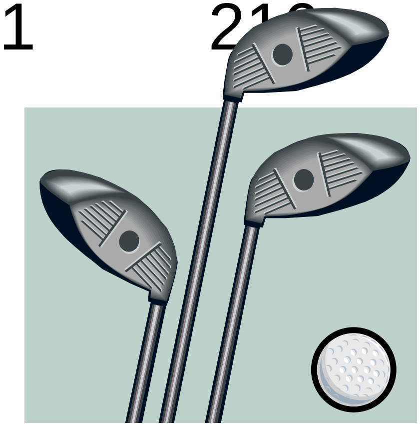


One-Factor ANOVA F Test Example

(compare the results with Excel)

You want to see if three different golf clubs yield different distances. You randomly select five measurements from trials on an automated driving machine for each club. At the .05 significance level, is there a difference in mean distance?

Club 1	**Club 2**	**Club 3**
254	234	200
263	218	222
241	235	197
237	227	206
251	21[illegible]	204

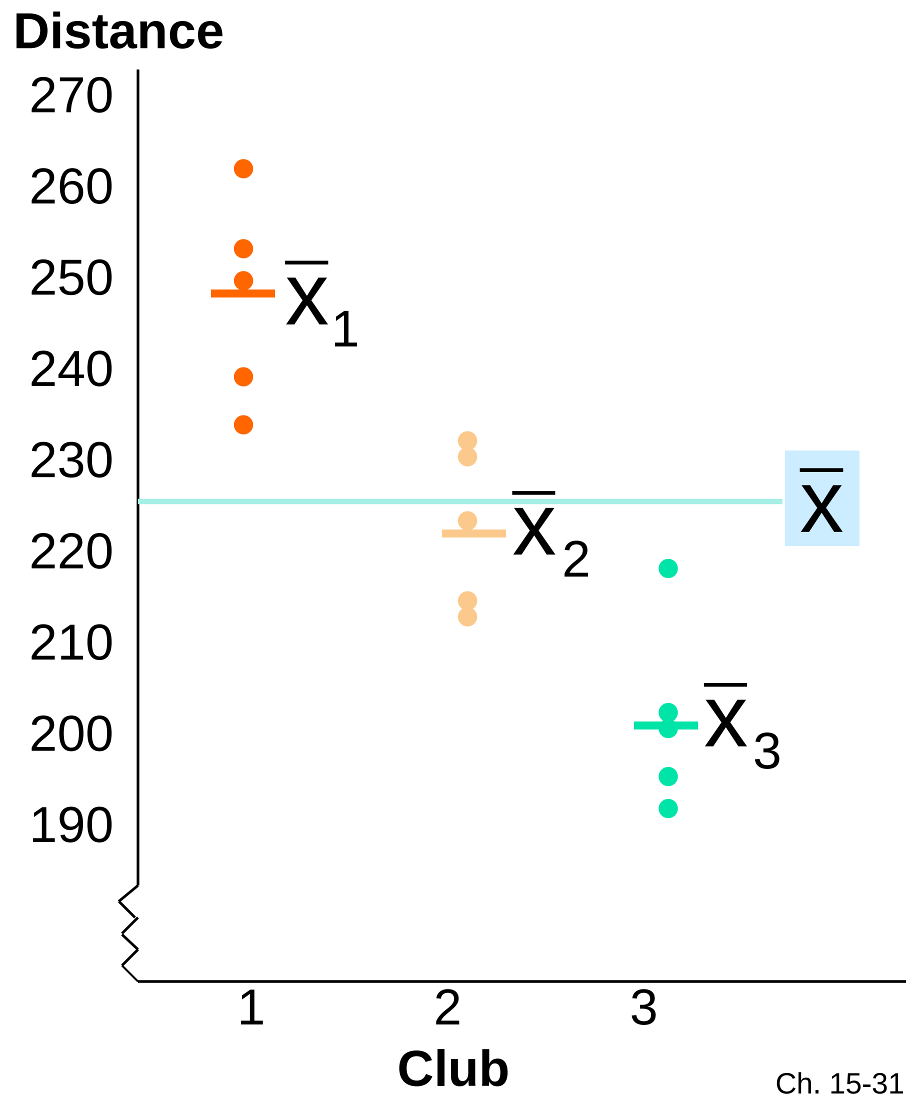


One-Factor ANOVA Example: Scatter Diagram

Club 1	Club 2	Club 3
254	234	200
263	218	222
241	235	197
237	227	206
251	216	204

$\bar{x}_1 = 249.2$ | $\bar{x}_2 = 226.0$ | $\bar{x}_3 = 205.8$

$\bar{\bar{x}} = 227.0$



One-Factor ANOVA Example Computations

Club 1	Club 2	Club 3
254	234	200
263	218	222
241	235	197
237	227	206
251	216	204

$\bar{x}_1 = 249.2$ $n_1 = 5$

$\bar{x}_2 = 226.0$ $n_2 = 5$

$\bar{x}_3 = 205.8$ $n_3 = 5$

$n = 15$

$\bar{\bar{x}} = 227.0$

$K = 3$

$SSG = 5\,(249.2 - 227)^2 + 5\,(226 - 227)^2 + 5\,(205.8 - 227)^2 = 4716.4$

$SSW = (254 - 249.2)^2 + (263 - 249.2)^2 + \ldots + (204 - 205.8)^2 = 1119.6$

$MSG = 4716.4 / (3-1) = 2358.2$

$MSW = 1119.6 / (15-3) = 93.3$

$$F = \frac{2358.2}{93.3} = 25.275$$

One-Factor ANOVA Example Solution

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

H_1: μ_i not all equal

$\alpha = .05$

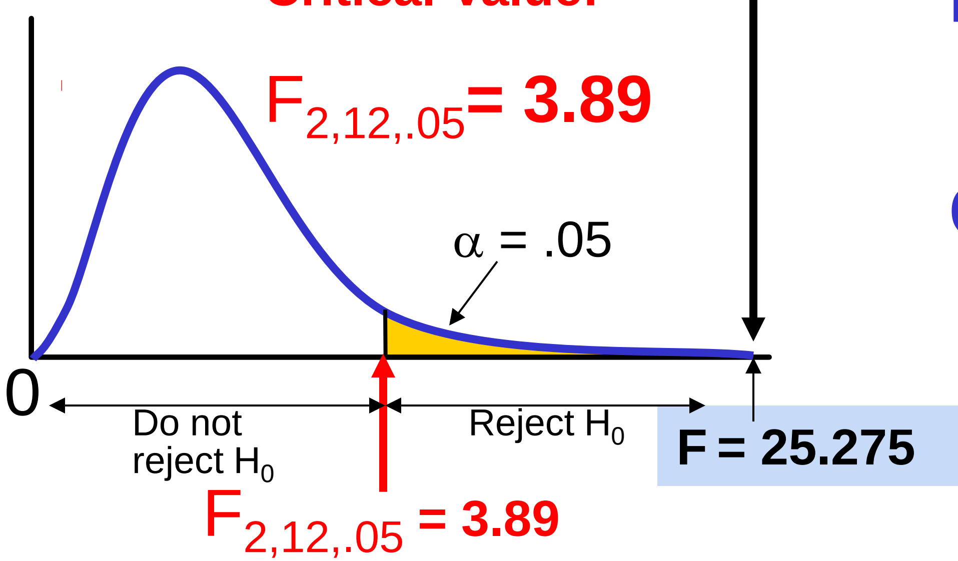
$df_1 = 2$ $df_2 = 12$

Test Statistic:

$$F = \frac{MSA}{MSW} = \frac{2358.2}{93.3} = 25.275$$

Critical Value:

$F_{2,12,.05}$ **= 3.89**



Decision:

Reject H_0 at $\alpha = 0.05$

Conclusion:

There is evidence that at least one μ_i differs from the rest

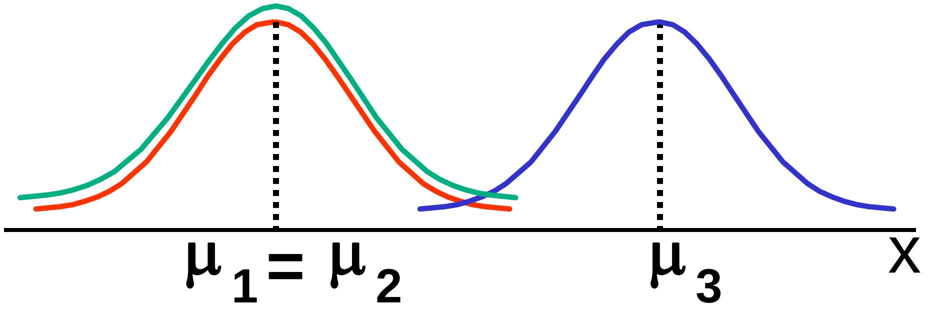
ANOVA -- Single Factor: Excel Output

EXCEL: data | data analysis | ANOVA: single factor

SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
Club 1	5	1246	249.2	108.2		
Club 2	5	1130	226	77.5		
Club 3	5	1029	205.8	94.2		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*

Multiple Comparisons Between Subgroup Means

- To test which population means are significantly different
  - e.g.: $\mu_1 = \mu_2 \neq \mu_3$
  - Done after rejection of equal means in single factor ANOVA design
- Allows pair-wise comparisons
  - Compare absolute mean differences with critical range



Continuous outcome (means)

Outcome Variable	Are the observations independent or correlated?		Alternatives if the normality assumption is violated (and small sample size):
	independent	correlated	
Continuous (e.g. pain scale, cognitive function)	**Ttest:** compares means between two independent groups **ANOVA:** compares means between more than two independent groups **Pearson's correlation coefficient** (linear correlation): shows linear correlation between two continuous variables **Linear regression:** multivariate regression technique	**Paired ttest:** compares means between two related groups (e.g., the same subjects before and after) **Repeated-measures ANOVA:** compares changes over time in the means of two or more groups (repeated measurements) **Mixed models/GEE modeling**: multivariate regression techniques to compare changes over time between two or more groups; gives rate of change over time	Non-parametric statistics **Wilcoxon sign-rank test**: non-parametric alternative to the paired ttest **Wilcoxon sum-rank test** (=Mann-Whitney U test): non-parametric alternative to the ttest **Kruskal-Wallis test:** non-parametric alternative to ANOVA **Spearman rank correlation coefficient:** non-parametric alternative to Pearson's correlation coefficient

ANOVA example

Mean micronutrient intake from the school lunch by school

		S1[a], *n*=28	S2[b], *n*=25	S3[c], *n*=21	*P*-value[d]
Calcium (mg)	Mean	117.8	158.7	206.5	0.000
	SD[e]	62.4	70.5	86.2	
Iron (mg)	Mean	2.0	2.0	2.0	0.854
	SD	0.6	0.6	0.6	
Folate (μg)	Mean	26.6	38.7	42.6	0.000
	SD	13.1	14.5	15.1	
Zinc (mg)	Mean	1.9	1.5	1.3	0.055
	SD	1.0	1.2	0.4	

[a] School 1 (most deprived; 40% subsidized lunches).
[b] School 2 (medium deprived; <10% subsidized).
[c] School 3 (least deprived; no subsidization, private school).
[d] ANOVA; significant differences are highlighted in bold (P<0.05).

FROM: Gould R, Russell J, Barker ME. School lunch menus and 11 to 12 year old children's food choice in three secondary schools in England-are the nutritional standards being met? *Appetite*. 2006

Example

Treatment 1	Treatment 2	Treatment 3	Treatment 4
60 inches	50	48	47
67	52	49	67
42	43	50	54
67	67	55	67
56	67	56	68
62	59	61	65
64	67	61	65
59	64	60	56
72	63	59	60
71	65	64	65

Example

Step 1) calculate the sum of squares between groups:

Mean for group 1 = 62.0

Mean for group 2 = 59.7

Mean for group 3 = 56.3

Mean for group 4 = 61.4

Treatment 1	Treatment 2	Treatment 3	Treatment 4
60 inches	50	48	47
67	52	49	67
42	43	50	54
67	67	55	67
56	67	56	68
62	59	61	65
64	67	61	65
59	64	60	56
72	63	59	60
71	65	64	65

Grand mean= 59.85

SSB = $[(62-59.85)^2 + (59.7-59.85)^2 + (56.3-59.85)^2 + (61.4-59.85)^2]$ *xn per group*= 19.65*x10* = 196.5

Example

Step 2) calculate the sum of squares within groups:

$(60-62)^2+(67-62)^2+(42-62)^2+(67-62)^2+(56-62)^2+(62-62)^2+(64-62)^2+(59-62)^2+(72-62)^2+(71-62)^2+(50-59.7)^2+(52-59.7)^2+(43-59.7)^2+67-59.7)^2+(67-59.7)^2+(69-59.7)^2$...+....(sum of 40 squared deviations) = **2060.6**

Treatment 1	Treatment 2	Treatment 3	Treatment 4
60 inches	50	48	47
67	52	49	67
42	43	50	54
67	67	55	67
56	67	56	68
62	59	61	65
64	67	61	65
59	64	60	56
72	63	59	60
71	65	64	65

Step 3) Fill in the ANOVA table

Source of variation	d.f.	Sum of squares	Mean Sum of Squares	F-statistic	p-value
Between	3	196.5	65.5	1.14	.344
Within	36	2060.6	57.2	-	-
Total	39	2257.1	-	-	-

Step 3) Fill in the ANOVA table

Source of variation	d.f.	Sum of squares	Mean Sum of Squares	F-statistic	p-value
Between	3	196.5	65.5	1.14	.344
Within	36	2060.6	57.2	-	-
Total	39	2257.1	-	-	-

INTERPRETATION of ANOVA:

How much of the variance in height is explained by treatment group?

R^2=“Coefficient of Determination” = SSB/TSS = 196.5/2275.1=9%

Coefficient of Determination

$$R^2 = \frac{SSB}{SSB + SSE} = \frac{SSB}{SST}$$

The amount of variation in the outcome variable (dependent variable) that is explained by the predictor (independent variable).

15.3

Kruskal-Wallis Test (No example is provided)

- Use when the normality assumption for one-way ANOVA is violated
- Assumptions:
  - The samples are random and independent
  - variables have a continuous distribution
  - the data can be ranked
  - populations have the same variability
  - populations have the same shape

Kruskal-Wallis Test Procedure

- Obtain relative rankings for each value
  - In event of tie, each of the tied values gets the average rank
- Sum the rankings for data from each of the K groups
  - Compute the Kruskal-Wallis test statistic
  - Evaluate using the chi-square distribution with K – 1 degrees of freedom

Kruskal-Wallis Test Procedure

(continued)

- The Kruskal-Wallis test statistic:

  (chi-square with K – 1 degrees of freedom)

$$W = \left[\frac{12}{n(n+1)} \sum_{i=1}^{K} \frac{R_i^2}{n_i} \right] - 3(n+1)$$

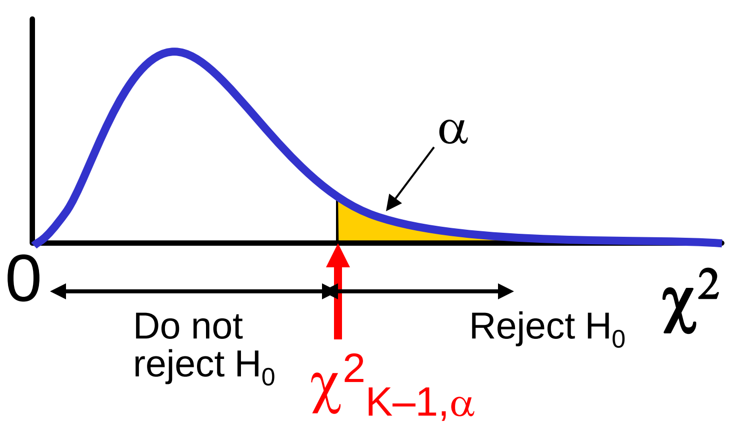
where:

n = sum of sample sizes in all groups
K = Number of samples
R_i = Sum of ranks in the i^{th} group
n_i = Size of the i^{th} group

Kruskal-Wallis Test Procedure

(continued)

- Complete the test by comparing the calculated H value to a critical χ^2 value from the chi-square distribution with K – 1 degrees of freedom



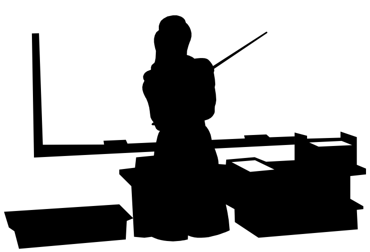
Decision rule

- Reject H_0 if $W > \chi^2_{K-1,\alpha}$
- Otherwise do not reject H_0

Kruskal-Wallis Example

- Do different departments have different class sizes?

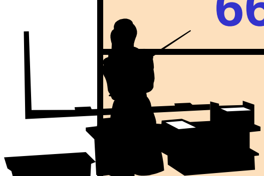
Class size (Math, M)	Class size (English, E)	Class size (Biology, B)
23	55	30
45	60	40
54	72	18
78	45	34
66	70	44



Kruskal-Wallis Example

- Do different departments have different class sizes?

Class size (Math, M)	Ranking	Class size (English, E)	Ranking	Class size (Biology, B)	Ranking
23	**2**	**55**	**10**	**30**	**3**
41	**6**	**60**	**11**	**40**	**5**
54	**9**	**72**	**14**	**18**	**1**
78	**15**	**45**	**8**	**34**	**4**
66	**12**	**70**	**13**	**44**	**7**
	Σ = 44		**Σ = 56**		**Σ = 20**



Kruskal-Wallis Example

(continued)

$$H_0: \text{Mean}_M = \text{Mean}_E = \text{Mean}_B$$

$$H_1: \text{Not all population means are equal}$$

- The W statistic is

$$W = \left[\frac{12}{n(n+1)} \sum_{i=1}^{K} \frac{R_i^2}{n_i} \right] - 3(n+1)$$

$$= \left[\frac{12}{15(15+1)} \left(\frac{44^2}{5} + \frac{56^2}{5} + \frac{20^2}{5} \right) \right] - 3(15+1) = 6.72$$

Kruskal-Wallis Example

(continued)

- Compare W = 6.72 to the critical value from the chi-square distribution for 3 – 1 = 2 degrees of freedom and α = .05:

$$\chi^2_{2,0.05} = 5.991$$

Since H = 6.72 **>** $\chi^2_{2,0.05} = 5.991$,

reject H_0

There is sufficient evidence to reject that the population means are all equal

15.4

Two-Way Analysis of Variance

- Examines the effect of
  - Two factors of interest on the dependent variable
    - e.g., Percent carbonation and line speed on soft drink bottling process
  - Interaction between the different levels of these two factors
    - e.g., Does the effect of one particular carbonation level depend on which level the line speed is set?

Two-Way ANOVA

(continued)

- Assumptions
  - Populations are normally distributed
  - Populations have equal variances
  - Independent random samples are drawn

Randomized Block Design

Two Factors of interest: A and B

K = number of groups of factor A

H = number of levels of factor B

(sometimes called a blocking variable)

Block	Group			
	1	2	…	K
1	x_{11}	x_{21}	…	x_{K1}
2	x_{12}	x_{22}	…	x_{K2}
.	.	.	.	.
.	.	.	.	.
H	x_{1H}	x_{2H}	…	x_{KH}

Two-Way Notation

- Let x_{ji} denote the observation in the j^{th} group and i^{th} block
- Suppose that there are K groups and H blocks, for a total of n = KH observations
- Let the overall mean be $\bar{x}$
- Denote the group sample means by

$$\bar{x}_{j\bullet} \quad (j = 1,2,\ldots,K)$$

- Denote the block sample means by

$$\bar{x}_{\bullet i} \quad (i = 1,2,\ldots,H)$$

Partition of Total Variation

- SST = SSG + SSB + SSE

Total Sum of Squares (SST)	=	Variation due to differences between groups (SSG)
		+
		Variation due to differences between blocks (SSB)
		+
The error terms are assumed to be independent, normally distributed, and have the same variance	→	Variation due to random sampling (unexplained error) (SSE)

Two-Way Sums of Squares

- The sums of squares are

Degrees of Freedom:

Total: $$SST = \sum_{j=1}^{K}\sum_{i=1}^{H}(x_{ji} - \bar{x})^2$$ $n - 1$

Between-Groups: $$SSG = H\sum_{j=1}^{K}(\bar{x}_{j\bullet} - \bar{x})^2$$ $K - 1$

Between-Blocks: $$SSB = K\sum_{i=1}^{H}(\bar{x}_{\bullet i} - \bar{x})^2$$ $H - 1$

Error: $$SSE = \sum_{j=1}^{K}\sum_{i=1}^{H}(x_{ji} - \bar{x}_{j\bullet} - \bar{x}_{\bullet i} + \bar{x})^2$$ $(K - 1)(K - 1)$

Two-Way Mean Squares

- The mean squares are

$$MST = \frac{SST}{n-1}$$

$$MSG = \frac{SST}{K-1}$$

$$MSB = \frac{SST}{H-1}$$

$$MSE = \frac{SSE}{(K-1)(H-1)}$$

Two-Way ANOVA: The F Test Statistic

F Test for Groups

H_0: The K population group means are all the same

$$F = \frac{MSG}{MSE}$$

Reject H_0 if

$F > F_{K-1,(K-1)(H-1),\alpha}$

F Test for Blocks

H_0: The H population block means are the same

$$F = \frac{MSB}{MSE}$$

Reject H_0 if

$F > F_{H-1,(K-1)(H-1),\alpha}$

General Two-Way Table Format

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F Ratio
Between groups	SSG	K – 1	$MSG = \frac{SSG}{K-1}$	$\frac{MSG}{MSE}$
Between blocks	SSB	H – 1	$MSB = \frac{SSB}{H-1}$	$\frac{MSB}{MSE}$
Error	SSE	(K – 1)(H – 1)	$MSE = \frac{SSE}{(K-1)(H-1)}$	
Total	SST	n - 1		

15.5

More than One Observation per Cell

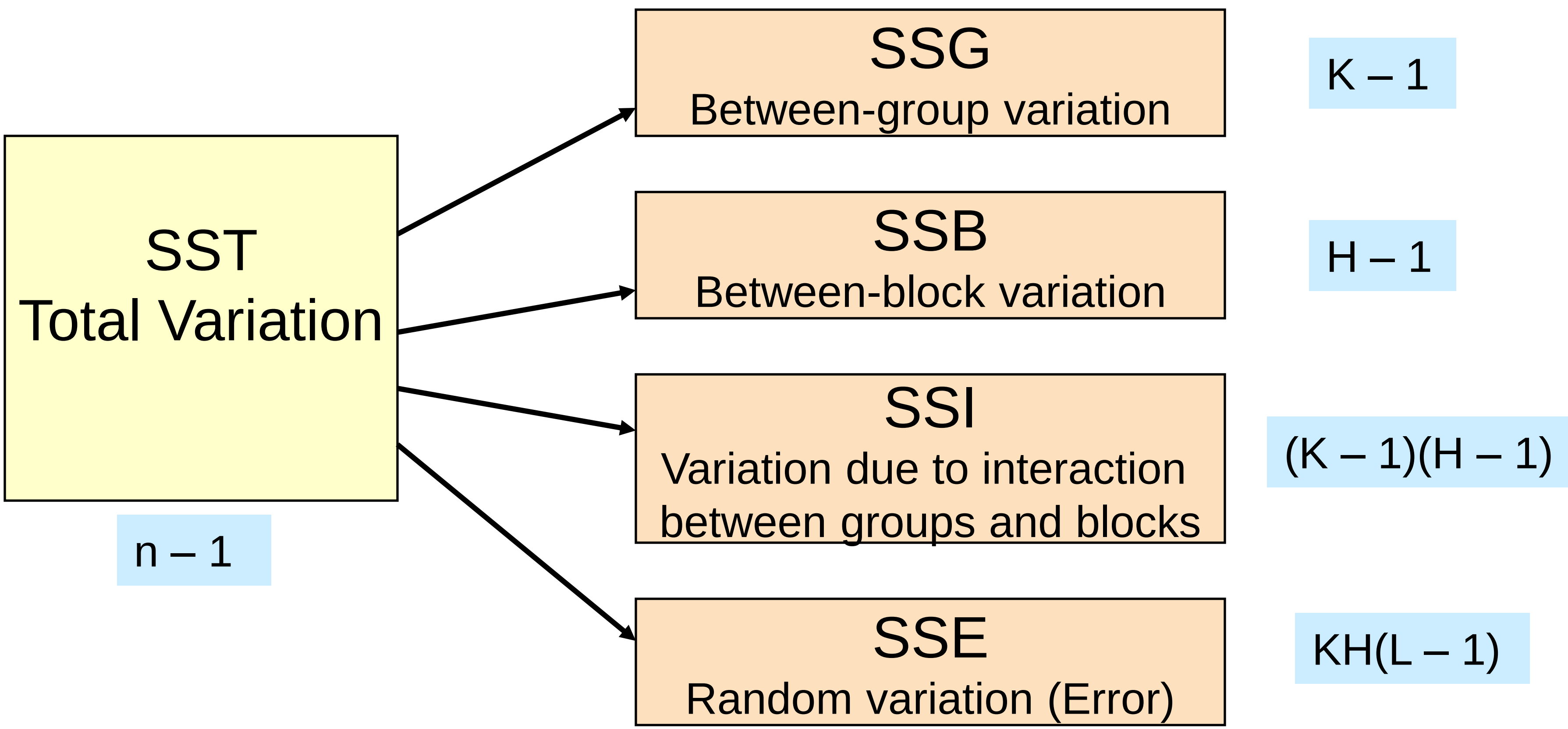
- A two-way design with more than one observation per cell allows one further source of variation
- The interaction between groups and blocks can also be identified
- Let
  - K = number of groups
  - H = number of blocks
  - L = number of observations per cell
  - n = KHL = total number of observations

More than One Observation per Cell

(continued)

SST = SSG + SSB + SSI + SSE

Degrees of Freedom:



Sums of Squares with Interaction

Degrees of Freedom:

Total: $SST = \sum_{j}\sum_{i}\sum_{l}(x_{jil} - \bar{x})^2$ — $n - 1$

Between-groups: $SSG = HL\sum_{j=1}^{K}(\bar{x}_{j\bullet\bullet} - \bar{x})^2$ — $K - 1$

Between-blocks: $SSB = KL\sum_{i=1}^{H}(\bar{x}_{\bullet i\bullet} - \bar{x})^2$ — $H - 1$

Interaction: $SSI = L\sum_{j=1}^{K}\sum_{i=1}^{H}(\bar{x}_{ji\bullet} - \bar{x}_{j\bullet\bullet} - \bar{x}_{\bullet i\bullet} + \bar{x})^2$ — $(K - 1)(H - 1)$

Error: $SSE = \sum_{i}\sum_{j}\sum_{l}(x_{jil} - \bar{x}_{ji\bullet})^2$ — $KH(L - 1)$

Two-Way Mean Squares with Interaction

- The mean squares are

$$MST = \frac{SST}{n-1}$$

$$MSG = \frac{SST}{K-1}$$

$$MSB = \frac{SST}{H-1}$$

$$MSI = \frac{SSI}{(K-1)(H-1)}$$

$$MSE = \frac{SSE}{KH(L-1)}$$

Two-Way ANOVA: The F Test Statistic

H_0: The K population group means are all the same

F Test for group effect

$$F = \frac{MSG}{MSE}$$

Reject H_0 if

$F > F_{K-1,KH(L-1),\alpha}$

H_0: The H population block means are the same

F Test for block effect

$$F = \frac{MSB}{MSE}$$

Reject H_0 if

$F > F_{H-1,KH(L-1),\alpha}$

H_0: the interaction of groups and blocks is equal to zero

F Test for interaction effect

$$F = \frac{MSI}{MSE}$$

Reject H_0 if

$F > F_{(K-1)(H-1),KH(L-1),\alpha}$

Two-Way ANOVA Summary Table

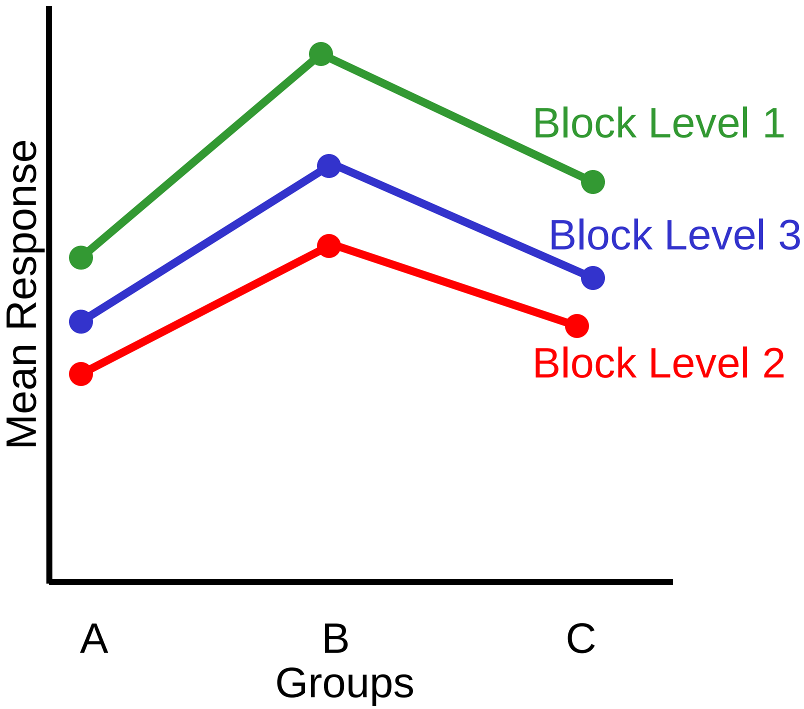
Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F Statistic
Between groups	**SSG**	**K – 1**	**MSG** = SSG / (K – 1)	$\frac{MSG}{MSE}$
Between blocks	**SSB**	**H – 1**	**MSB** = SSB / (H – 1)	$\frac{MSB}{MSE}$
Interaction	**SSI**	**(K – 1)(H – 1)**	**MSI** = SSI / (K – 1)(H – 1)	$\frac{MSI}{MSE}$
Error	**SSE**	**KH(L – 1)**	**MSE** = SSE / KH(L – 1)	

Features of Two-Way ANOVA F Test

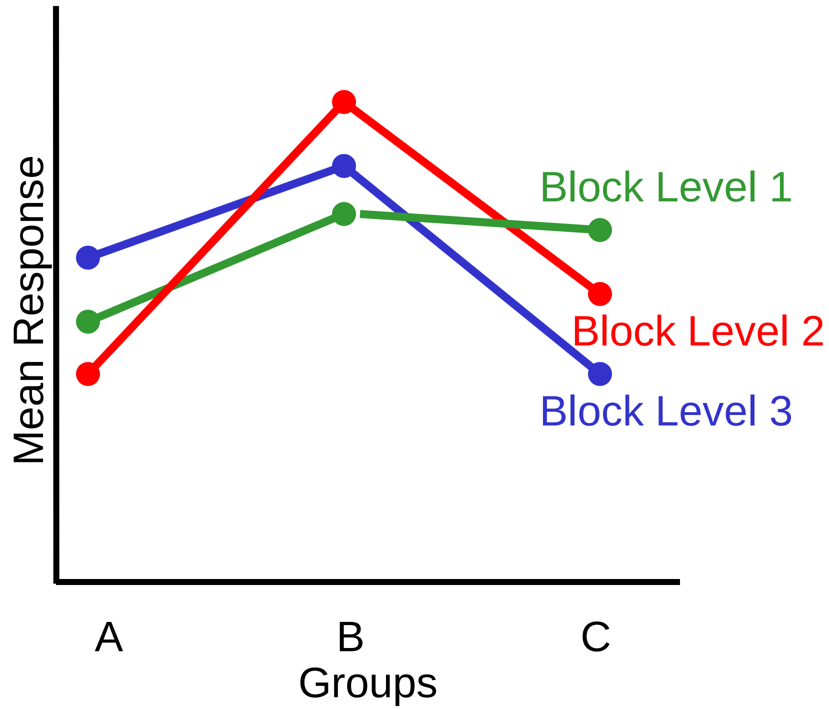
- Degrees of freedom always add up
  - n-1 = KHL-1 = (K-1) + (H-1) + (K-1)(H-1) + KH(L-1)
  - Total = groups + blocks + interaction + error
- The denominator of the F Test is always the same but the numerator is different
- The sums of squares always add up
  - SST = SSG + SSB + SSI + SSE
  - Total = groups + blocks + interaction + error

Examples: Interaction vs. No Interaction

- No interaction:



- Interaction is present:



Extra examples on ANOVA:

- **EX. : the following table gives the output of 3 years of an experimental farm that use 3 types of fertilizer . a sum that the output under each fertilizer are normally distributed with equal variance . test that the population mean are the same at 5% level of significant**

	F1	F2	F3
1	50	30	30
2	90	40	70
3	40	50	50

- EX.: **the following table gives the output of 3 years of an experimental farm that use 3 types of fertilizer and pesticides . a sum that the output under each fertilizer per pesticide are normally distributed with equal variance . test that the population mean are the same at 5% level of significant** $\bar{\bar{X}}$

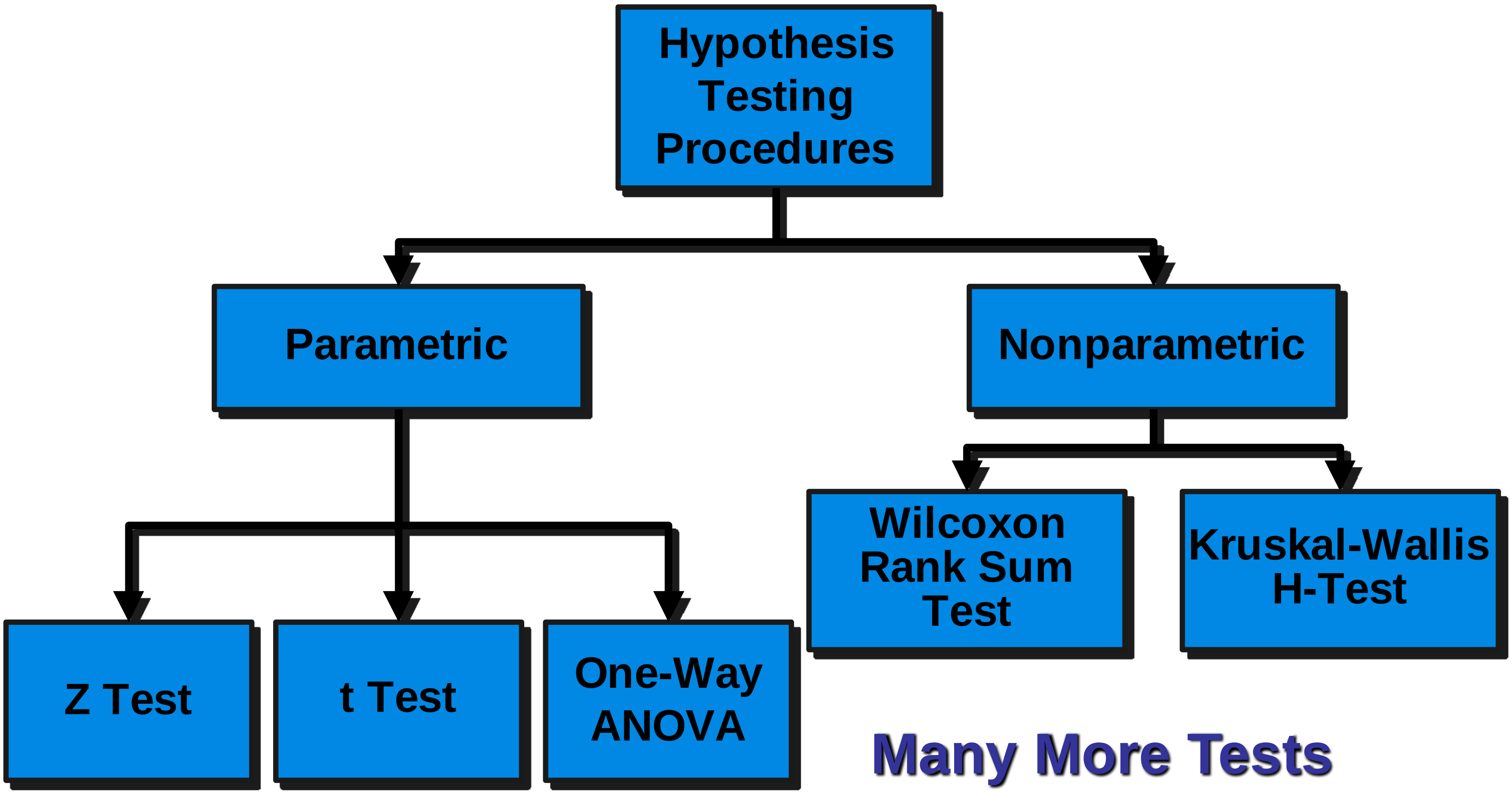
	Fert1	Fert2	Fert3	Fert4	Sample mean
Pest1	21	12	9	6	$_1 = 12$
Pest2	13	10	8	5	$_2 = 9$
Pest3	8	8	7	1	$_3 = 6$
Sample mean	$_1 = 14$	$_2 = 10$	$_3 = 8$	$_4 = 4$	

Chapter 1

1.3 Nonparametric Statistics

- Nonparametric Statistics
  - Fewer restrictive assumptions about data levels and underlying probability distributions
    - Population distributions may be skewed
    - The level of data measurement may only be ordinal or nominal

Hypothesis Testing Procedures



Many More Tests Exist!

Sign Test and Confidence Interval

- A sign test for paired or matched samples:
  - Calculate the differences of the paired observations
  - Discard the differences equal to 0, leaving n observations
  - Record the sign of the difference as + or –
- For a symmetric distribution, the signs are random and + and – are equally likely

Sign Test

(continued)

- Define + to be a "success" and let P = the true proportion of +'s in the population
- The sign test is used for the hypothesis test

$$H_0 : P = 0.5$$

- The test-statistic S for the sign test is

S = the number of pairs with a positive difference

- S has a binomial distribution with P = 0.5 and

  n = the number of nonzero differences

Determining the p-value

- The p-value for a Sign Test is found using the binomial distribution with n = number of nonzero differences, S = number of positive differences, and $P = 0.5$

- For an upper-tail test, $H_1: P > 0.5$, p-value = $P(x \geq S)$

- For a lower-tail test, $H_1: P < 0.5$, p-value = $P(x \leq S)$

- For a two-tail test, $H_1: P \neq 0.5$, 2(p-value)

Sign Test Example

- Ten consumers in a focus group have rated the attractiveness of two package designs for a new

Consumer	Rating		Difference	Sign of Difference
	Package 1	Package 2	Rating 1 – 2	
1	5	8	-3	**–**
2	4	8	-4	**–**
3	4	4	0	**0**
4	6	5	+1	**+**
5	3	9	-6	**–**
6	5	9	-4	**–**
7	7	6	-1	**–**
8	5	9	-4	**–**
9	6	3	+3	**+**
10	7	9	-2	**–**

Sign Test Example

(continued)

- Test the hypothesis that there is no overall package preference using $\alpha = 0.10$

$$H_0 : P = 0.5$$

The proportion of consumers who prefer package 1 is the same as the proportion preferring package 2

$$H_1 : P < 0.5$$

A majority prefer package 2

- The test-statistic S for the sign test is

S = the number of pairs with a positive difference
= 2

- S has a binomial distribution with P = 0.5 and n = 9 (there was one zero difference)

Sign Test Example

(continued)

- The p-value for this sign test is found using the binomial distribution with n = 9, S = 2, and P = 0.5:

- For a lower-tail test,

$$\text{p-value} = P(x \leq 2 | n=9, P=0.5) = 0.090$$

Since 0.090 < α = 0.10 we reject the null hypothesis and conclude that consumers prefer package 2

Sign Test: Normal Approximation

- If the number n of nonzero sample observations is large, then the sign test is based on the normal approximation to the binomial with mean and standard deviation

$$\mu = nP = 0.5n$$

$$\sigma = \sqrt{nP(1-P)} = \sqrt{0.25n} = 0.5\sqrt{n}$$

- The test statistic is

$$Z = \frac{S^* - \mu}{\sigma} = \frac{S^* - 0.5n}{0.5\sqrt{n}}$$

- Where S* is the test-statistic corrected for continuity:
  - For a two-tail test, $S^* = S + 0.5$, if $S < \mu$ or $S^* = S - 0.5$, if $S > \mu$
  - For upper-tail test, $S^* = S - 0.5$
  - For lower-tail test, $S^* = S + 0.5$

Sign Test for Single Population Median

- The sign test can be used to test that a single population median is equal to a specified value
  - For small samples, use the binomial distribution
  - For large samples, use the normal approximation

Wilcoxon Signed Rank Test for Paired Samples

- Uses matched pairs of random observations
- Still based on ranks
- Incorporates information about the magnitude of the differences
- Tests the hypothesis that the distribution of differences is centered at zero
- The population of paired differences is assumed to be symmetric

Wilcoxon Signed Rank Test for Paired Samples

(continued)

Conducting the test:

- Discard pairs for which the difference is 0
- Rank the remaining n absolute differences in ascending order (ties are assigned the average of their ranks)
- Find the sums of the positive ranks and the negative ranks
- The smaller of these sums is the Wilcoxon Signed Rank Statistic T:

$$T = \min(T_{+}, T_{-})$$

Where T_{+} = the sum of the positive ranks

T_{-} = the sum of the negative ranks

n = the number of nonzero differences

- The null hypothesis is rejected if T is less than or equal to the value in Appendix Table 10

Signed Rank Test Example

Consumer	Rating		Difference		
	Package 1	Package 2	Diff (rank)	Rank (+)	Rank (–)
1	5	8	-3 (5)		**5**
2	4	8	-4 (7 tie)		**7**
3	4	4	0 (-)		
4	6	5	+1 (2)	**2**	
5	3	9	-6 (9)		**9**
6	5	9	-4 (7 tie)		**7**
7	7	6	-1 (3)		**3**
8	5	9	-4 (7 tie)		**7**
9	6	3	+3 (1)	**1**	
10	7	9	-2 (4)	↓	**4**

- Ten consumers in a focus group have rated the attractiveness of two package designs for a new product

Signed Rank Test Example

(continued)

Test the hypothesis that the distribution of paired differences is centered at zero, using $\alpha = 0.10$

Conducting the test:

- The smaller of T_+ and T_- is the Wilcoxon Signed Rank Statistic T:

$$\mathbf{T = \min(T_+, T_-) = 3}$$

- Use Appendix Table 10 with $n = 9$ to find the critical value:

The null hypothesis is rejected if $T \le 34$

- Since $T = 3 < 34$, we Accept t the null hypothesis

Wilcoxon Signed Rank Test Normal Approximation

A normal approximation can be used when

- Paired samples are observed
- The sample size is large
- The hypothesis test is that the population distribution of differences is centered at zero

Wilcoxon Signed Rank Test Normal Approximation

(continued)

- The T statistic approaches a normal distribution as sample size increases
- If the number of paired values is larger than 20, a normal approximation can be used

Wilcoxon Matched Pairs Test for Large Samples

- The mean and standard deviation for Wilcoxon T :

$$E(T) = \mu_T = \frac{n(n+1)}{4}$$

$$Var(T) = \sigma_T^2 = \frac{(n)(n+1)(2n+1)}{24}$$

where n is the number of paired values

Wilcoxon Matched Pairs Test for Large Samples

(continued)

- Normal approximation for the Wilcoxon T Statistic:

$$z = \frac{T - \mu_T}{\sigma_T} = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

  - If the alternative hypothesis is one-sided, reject the null hypothesis if

$$\frac{T - \mu_T}{\sigma_T} < -z_{\alpha}$$

  - If the alternative hypothesis is two-sided, reject the null hypothesis if

$$\frac{T - \mu_T}{\sigma_T} < -z_{\alpha/2}$$

Mann-Whitney U-Test

Used to compare two samples from two populations

Assumptions:

- The two samples are independent and random
- The value measured is a continuous variable
- The two distributions are identical except for a possible difference in the central location
- The sample size from each population is at least 10

Mann-Whitney U-Test

(continued)

- Consider two samples
  - Pool the two samples (combine into a singe list) but keep track of which sample each value came from
  - rank the values in the combined list in ascending order
    - For ties, assign each the average rank of the tied values
  - sum the resulting rankings separately for each sample
- If the sum of rankings from one sample differs enough from the sum of rankings from the other sample, we conclude there is a difference in the population medians

Mann-Whitney U Statistic

- Consider n_1 observations from the first population and n_2 observations from the second
- Let R_1 denote the sum of the ranks of the observations from the first population
- The Mann-Whitney U statistic is

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

Mann-Whitney U Statistic

(continued)

- The null hypothesis is that the central locations of the two population distributions are the same
- The Mann-Whitney U statistic has mean and variance

$$E(U) = \mu_U = \frac{n_1 n_2}{2}$$

$$Var(U) = \sigma_U^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

- Then for large sample sizes (both at least 10), the distribution of the random variable

$$z = \frac{U - \mu_U}{\sigma_U}$$

is approximated by the normal distribution

Decision Rules for Mann-Whitney Test

The decision rule for the null hypothesis that the two populations have the same central location:

- For a one-sided upper-tailed alternative hypothesis:

$$\text{Reject } H_0 \text{ if } \; z = \frac{U - \mu_U}{\sigma_U} < -z_\alpha$$

- For a one-sided lower-tailed hypothesis:

$$\text{Reject } H_0 \text{ if } \; z = \frac{U - \mu_U}{\sigma_U} > z_\alpha$$

- For a two-sided alternative hypothesis:

$$\text{Reject } H_0 \text{ if } z = \frac{U - \mu_U}{\sigma_U} < -z_\alpha \quad \text{or} \quad \text{Reject } H_0 \text{ if } z = \frac{U - \mu_U}{\sigma_U} > z_\alpha$$

Mann-Whitney U-Test Example

Claim: Median class size for Math is larger than the median class size for English

A random sample of 10 Math and 10 English classes is selected (samples do not have to be of equal size)

Rank the combined values and then determine rankings by original sample

Mann-Whitney U-Test Example

(continued)

- Suppose the results are:

Class size (Math, M)	Class size (English, E)
23	30
45	47
34	18
78	34
34	44
66	61
62	54
95	28
81	40
99	96

Mann-Whitney U-Test Example

(continued)

Ranking for combined samples

Size	Rank	Size	Rank
18	1	47	11
23	2	54	12
28	3	61	13
30	4	62	14
34	6	66	15
34	6	78	16
34	6	81	17
40	8	95	18
44	9	96	19
45	10	99	20

tied

Mann-Whitney U-Test Example

(continued)

- Rank by original sample:

Class size (Math, M)	Rank	Class size (English, E)	Rank
23	2	**30**	4
45	10	**47**	11
34	6	**18**	1
78	16	**34**	6
34	6	**44**	9
66	15	**61**	13
62	14	**54**	12
95	18	**28**	3
81	17	**40**	8
99	20	**96**	19
	Σ = 124		Σ = 86

Mann-Whitney U-Test Example

(continued)

Claim: Median class size for Math is larger than the median class size for English

→

H_0: $Median_M \leq Median_E$

(Math median is not greater than English median)

H_A: $Median_M > Median_E$

(Math median is larger)

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - \sum R_1 = (10)(10) + \frac{(10)(11)}{2} - 124 = 31$$

Mann-Whitney U-Test Example

(continued)

H_0: $Median_M \leq Median_E$

H_A: $Median_M > Median_E$

$$z = \frac{U - \mu_U}{\sigma_U} = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} = \frac{31 - \frac{(10)(10)}{2}}{\sqrt{\frac{(10)(10)(10 + 10 + 1)}{12}}} = -1.436$$

- The decision rule for this one-sided upper-tailed alternative hypothesis:

$$\text{Reject } H_0 \text{ if } \; z = \frac{U - \mu_U}{\sigma_U} < -z_\alpha$$

- For $\alpha = 0.05$, $-z_\alpha = -1.645$
- The calculated z value is not in the rejection region, so we conclude that there is not sufficient evidence of difference in class size medians

Wilcoxon Rank Sum Test

- Similar to Mann-Whitney U test
- Results will be the same for both tests

Wilcoxon Rank Sum Test

(continued)

- n_1 observations from the first population
- n_2 observations from the second population
- Pool the samples and rank the observations in ascending order
- Let T denote the sum of the ranks of the observations from the first population
  - (T in the Wilcoxon Rank Sum Test is the same as R_1 in the Mann-Whitney U Test)

Wilcoxon Rank Sum Test

(continued)

- The Wilcoxon Rank Sum Statistic, T, has mean

$$E(T) = \mu_T = \frac{n_1(n_1 + n_2 + 1)}{2}$$

- And variance

$$Var(T) = \sigma_T^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

- Then, for large samples ($n_1 \geq 10$ and $n_2 \geq 10$) the distribution of the random variable

$$Z = \frac{T - \mu_T}{\sigma_T}$$

is approximated by the normal distribution

Wilcoxon Rank Sum Example

- We wish to test

H_0: $Median_1 \geq Median_2$

H_1: $Median_1 < Median_2$

- Use $\alpha = 0.05$
- Suppose two samples are obtained:
- $n_1 = 40$, $n_2 = 50$
- When rankings are completed, the sum of ranks for sample 1 is $\Sigma R_1 = 1475$ = T
- When rankings are completed, the sum of ranks for sample 2 is $\Sigma R_2 = 2620$

Wilcoxon Rank Sum Example

(continued)

- Using the normal approximation:

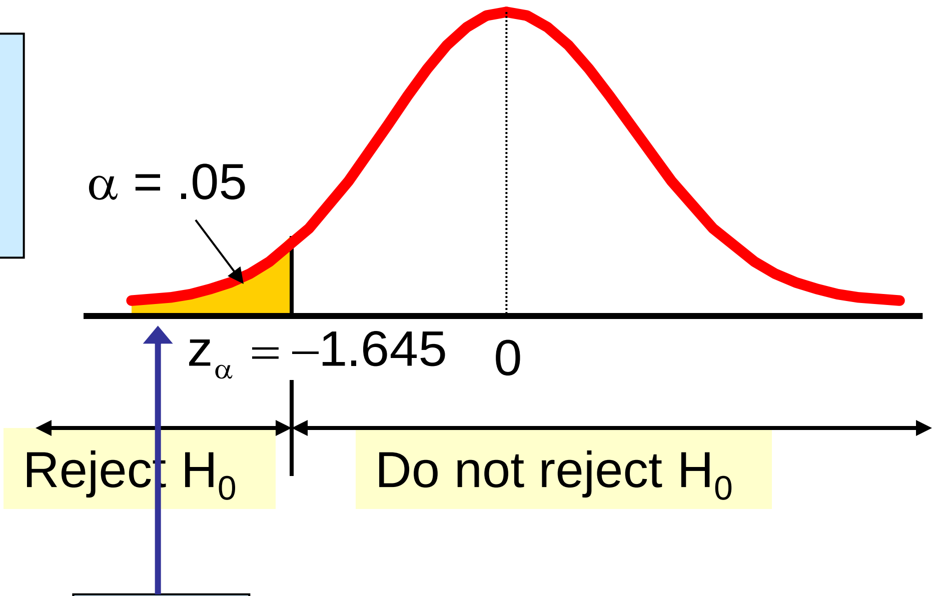
$$z = \frac{T - \mu_T}{\sigma_T} = \frac{T - \frac{n_1(n_1 + n_2 + 1)}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} = \frac{1475 - \frac{(40)(40 + 50 + 1)}{2}}{\sqrt{\frac{(40)(50)(40 + 50 + 1)}{12}}} = -2.80$$

Wilcoxon Rank Sum Example

(continued)

H_0: $\text{Median}_1 \geq \text{Median}_2$

H_1: $\text{Median}_1 < \text{Median}_2$



$$z = \frac{T - \mu_T}{\sigma_T} = -2.80$$

Since z = -2.80 < -1.645, we reject H_0 and conclude that median 1 is less than median 2 at the 0.05 level of significance

Spearman Rank Correlation

(Example is not provided)

- Consider a random sample $(x_1, y_1), \ldots, (x_n, y_n)$ of n pairs of observations
- Rank x_i and y_i each in ascending order
- Calculate the sample correlation of these ranks
- The resulting coefficient is called Spearman's Rank Correlation Coefficient.
- If there are no tied ranks, an equivalent formula for computing this coefficient is

$$r_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$$

where the d_i are the differences of the ranked pairs

Spearman Rank Correlation

(continued)

- Consider the null hypothesis

H_0: no association in the population

- To test against the alternative of positive association, the decision rule is

$$\text{Reject } H_0 \text{ if } r_S > r_{S,\alpha}$$

- To test against the alternative of negative association, the decision rule is

$$\text{Reject } H_0 \text{ if } r_S < -r_{S,\alpha}$$

- To test against the two-sided alternative of some association, the decision rule is

$$\text{Reject } H_0 \text{ if } r_S < -r_{S,\alpha/2} \quad \text{or} \quad r_S > r_{S,\alpha/2}$$

Extra Examples: Wilcoxon Signed rank test

- **The dos department publishes information about food cost document according to that document , a typical Jordanian family of 4 spends about 157JD per week on food , 10 randomly selected families have the weekly costs shown in table one . do the data provide sufficient evidence to conclude that the mean weekly food cost for the Amman families is less than the national mean (157JD):**

Sample mean weakly food costs:

143 , 169 , 149 , 135 , 161 , 138 , 152 , 150 , 141 , 159 .

Extra Examples: Man-Whitney

a nationwide shipping firm purchased a new computer system to track its shipment , pickups …… employees were expected to need about 2 hours to learn how to use the system . in fact , some employees could use the system in very little time, where as others took considerably longer.

Someone suggested that the reason for this difference might be that only some employees had experience with this kind of computer system . To test experience were randomly selected .

The times , in minutes , required for there employed to learn how to use the system are given in table 1 at the 5%significance level , do the data provide sufficient evidence to conclude that the mean learning time for all employed without experience exceeds the mean learning time for all employees with experience .

Without experience	With experience
139	142
118	109
164	130
151	107
182	155
140	88
134	95
	104