Design of Experiments: One Factor and Randomized Block Experiments

Prof. Daniel A. Menasce

Dept. of Computer Science

George Mason University

Basic Notions in Design of Experiments

- Response: what you want to measure.
- Factor: what affects the response.
- Level: value of a factor.

	Factors			Response
	CPU Clock Frequency (MHz)	Number of CPUs	Main Memory (MB)	Benchmark Execution Time (sec)
Levels	550	1	128	25.0
	750	1	128	32.0
	1000	1	128	48.0
	550	2	128	19.0
	750	2	128	13.5
	1000	2	128	10.0
	550	1	256	23.0
	750	1	256	29.0
	1000	1	256	45.0
	550	2	256	16.5
	750	2	256	11.8
	1000	2	256	8.8

Comparing Means of Various Groups

- ANOVA: Analysis of Variance.
- Consider c groups (each group is a level of a factor).
- Subdivide total variation in the response into variations attributable to differences among the c groups and differences within the c groups (experimental error).

- A, B, C, and D are different page replacement algorithms.
- Factor: page replacement algorithm.
- Levels: A, B, C, and D.
- Number in each column: running times of programs under each replacement algorithm.

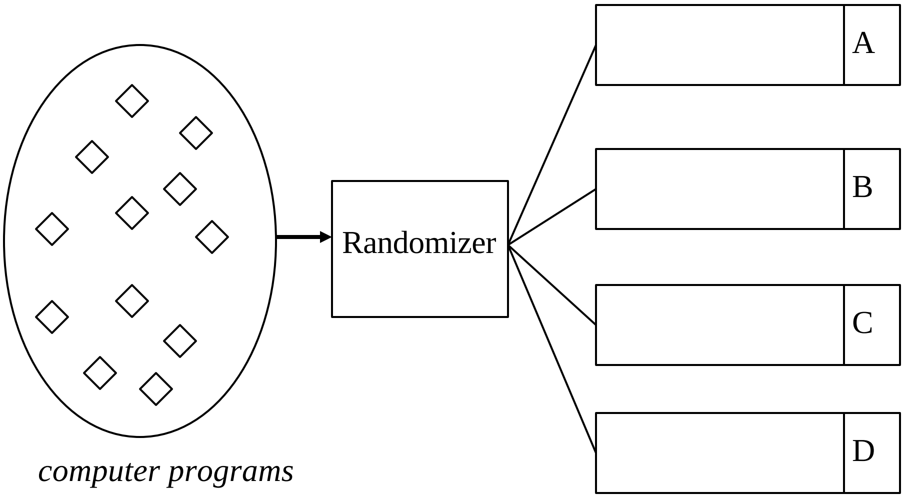
Page Replacement Algorithm

A	B	C	D
11	12	18	11
13	14	16	12
17	17	18	16
17	19	20	15
15	21	22	14
16	18	15	17
14	19	17	13
10	18	21	16
12	16	16	17
14	18	20	18

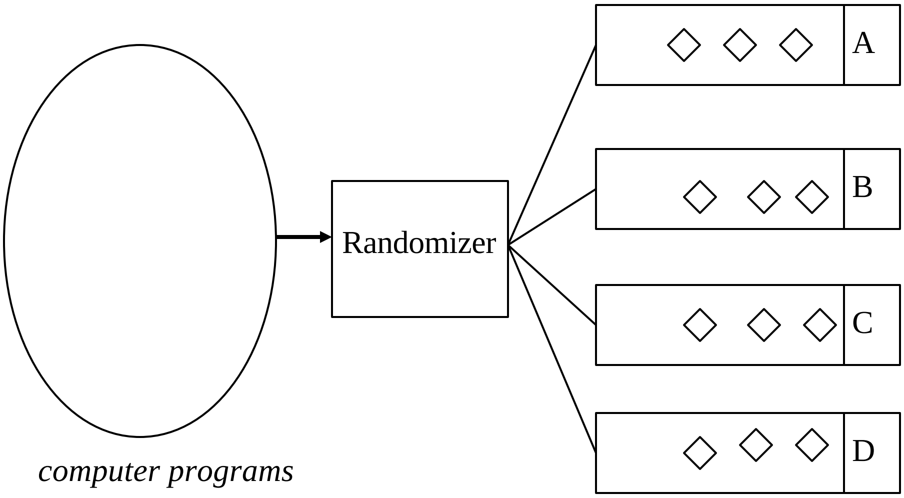
How about the Influence of Uncontrolled and Unforeseen Factors?

- The running time of a program depends on many other factors. Its locality of reference plays a role in the effectiveness of a page replacement algorithm.
- Randomization: consider a large set of programs and randomly assign programs to each group.

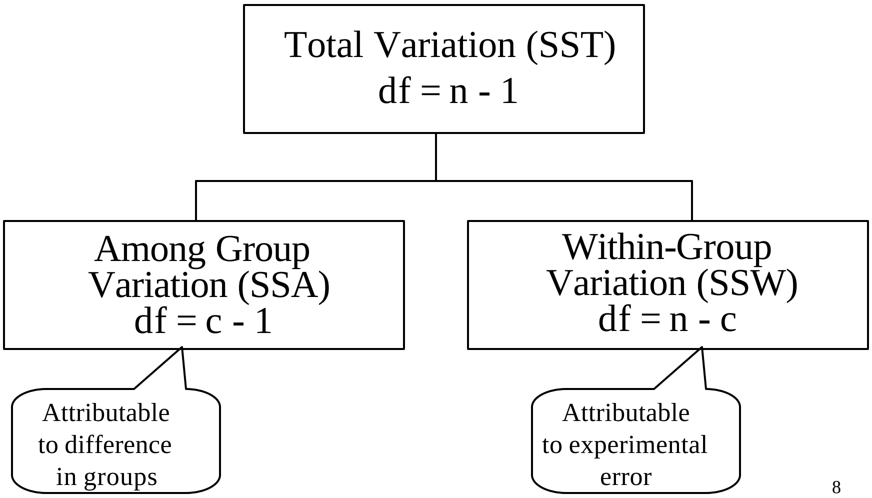
Randomization



Randomization



ANOVA Model



Total Variation (SST)
df = n - 1

Among Group
Variation (SSA)
df = c - 1

Attributable
to difference
in groups

Within-Group
Variation (SSW)
df = n - c

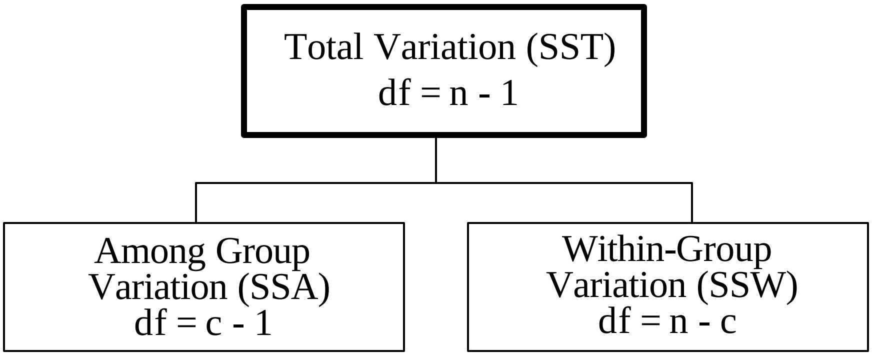
Attributable
to experimental
error

ANOVA

- Assumptions:
  - – c groups or levels of the factor being examined represent populations whose outcome measurements are randomly and independently drawn and follow a normal distribution and have equal variances.
- Hypotheses:

$$H_0 : m_1 = m_2 = ... = m_c$$
$$H_1 : \text{not all } m_j \text{ are equal (j} = 1,..., \text{c)}$$

ANOVA Model



Total Variation (SST)
df = n - 1

Among Group Variation (SSA)
df = c - 1

Within-Group Variation (SSW)
df = n - c

SST (Sum of Squares Total)

$$SST = \sum_{j=1}^{c} \sum_{i=1}^{n_j} \left(X_{ij} - \overline{\overline{X}} \right)^2$$

where

$$\overline{\overline{X}} = \frac{\sum_{j=1}^{c} \sum_{i=1}^{n_j} X_{ij}}{n}$$: overall or grand mean.

X_{ij} : i-th observation in group or level j.

n_j : number of observations in group or level j.

n: total number of observations: $\sum_{j=1}^{c} n_j$

SST Example

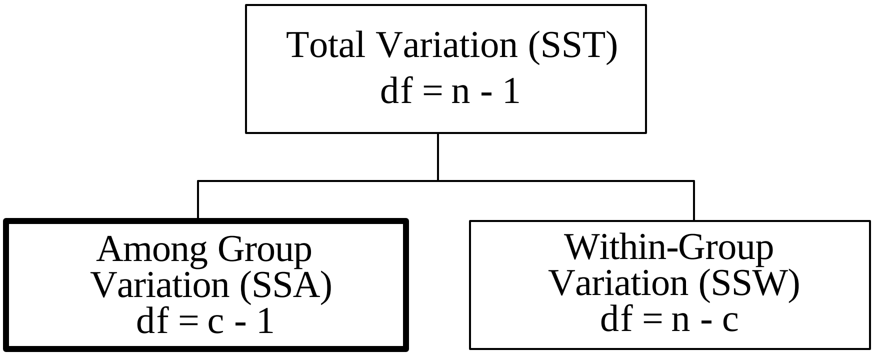
Page Replacement Algorithm

A	B	C	D
11	12	18	11
13	14	16	12
17	17	18	16
17	19	20	15
15	21	22	14
16	18	15	17
14	19	17	13
10	18	21	16
12	16	16	17
14	18	20	18

Grand Mean 16.075

SST = $(11-16.075)^2 + (13-16.075)^2 + \ldots + (18-16.075)^2 =$ 336.75

ANOVA Model



Total Variation (SST)
df = n - 1

Among Group Variation (SSA)
df = c - 1

Within-Group Variation (SSW)
df = n - c

SSA (Sum of Squares Among Groups)

$$SSA = \sum_{j=1}^{c} n_j \left(\overline{X}_j - \overline{\overline{X}} \right)^2$$

where

$$\overline{\overline{X}} = \frac{\sum_{j=1}^{c} \sum_{i=1}^{n_j} X_{ij}}{n}$$: overall or grand mean.

$\overline{X}_j$: sample mean corresponding to group or level j.

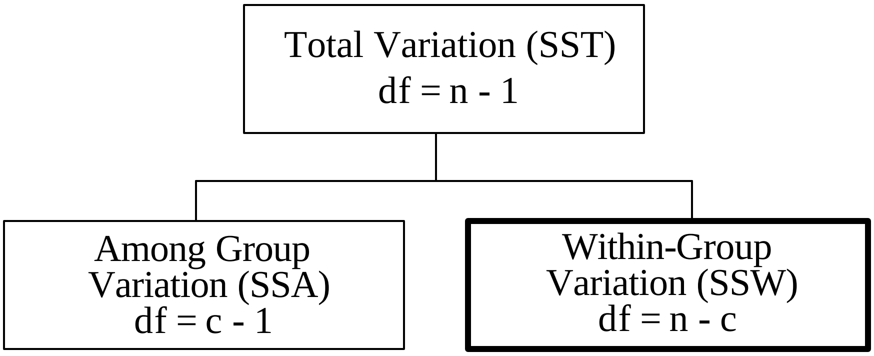
n_j : number of observations in group or level j.

SSA Example

	Page Replacement Algorithm			
	A	B	C	D
	11	12	18	11
	13	14	16	12
	17	17	18	16
	17	19	20	15
	15	21	22	14
	16	18	15	17
	14	19	17	13
	10	18	21	16
	12	16	16	17
	14	18	20	18
Mean	13.9	17.2	18.3	14.9
Grand Mean	16.075			

$$SSA = 10\ (13.9-16.075)^2 + 10\ (17.2-16.075)^2 + 10\ (18.3- 16.075)^2 + 10\ (14.9- 16.075)^2 = 123.275$$

ANOVA Model



Total Variation (SST)
df = n - 1

Among Group
Variation (SSA)
df = c - 1

Within-Group
Variation (SSW)
df = n - c

SSW (Sum of Squares Within Groups)

$$SSW = \sum_{j=1}^{c}\sum_{i=1}^{n_j}\left(X_{ij} - \overline{X}_j\right)^2$$

where

X_{ij} : i-th observation in group or level j.

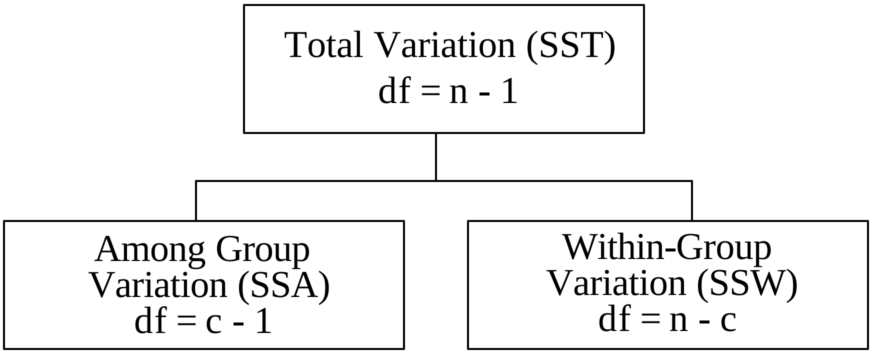
$\overline{X}_j$: sample mean corresponding to group or level j.

SSW Example

	Page Replacement Algorithm			
	A	B	C	D
	11	12	18	11
	13	14	16	12
	17	17	18	16
	17	19	20	15
	15	21	22	14
	16	18	15	17
	14	19	17	13
	10	18	21	16
	12	16	16	17
	14	18	20	18
Mean	13.9	17.2	18.3	14.9

SSW = $(11-13.9)^2$ + …+$(14-13.9)^2$ +
$(12-17.2)^2$ + … + $(18-17.2)^2$ +
$(18-18.3)^2$ + … + $(20-18.3)^2$ +
$(11-14.9)^2$ + … + $(18-14.9)^2$
= 213.5

ANOVA Model



Total Variation (SST)
df = n - 1

Among Group
Variation (SSA)
df = c - 1

Within-Group
Variation (SSW)
df = n - c

$$\boxed{SST = SSA + SSW}$$

ANOVA Model: Mean Squares

$$MSA = \frac{SSA}{c-1}$$

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

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

The mean squares are variances!
If there are no real differences among the c groups, MSA, MSW, and MST provide estimates for the variance inherent in the data.

The one-way ANOVA F Test Static

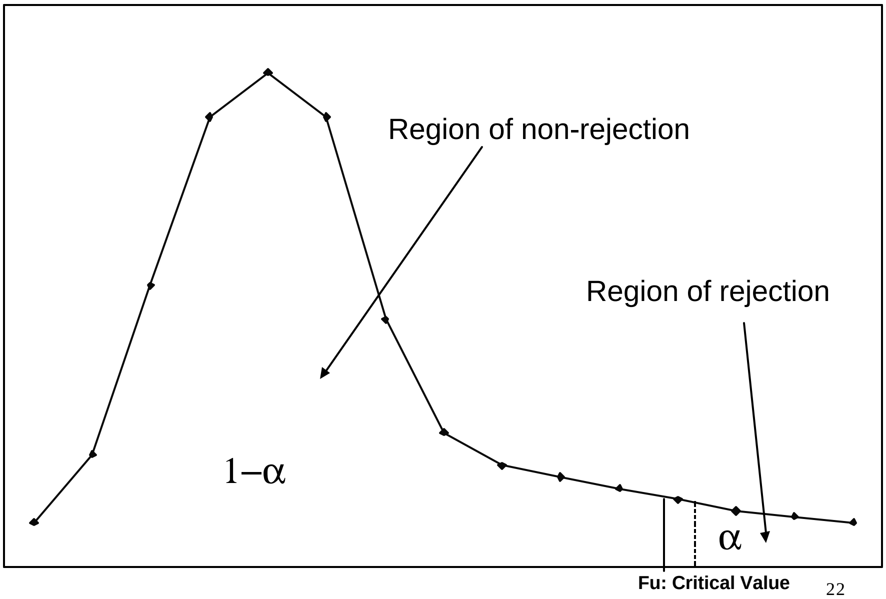
$$F = \frac{MSA}{MSW}$$

- The F-test statistic follows an F distribution with c-1 degrees of freedom in the numerator corresponding to MSA and n-c degrees of freedom in the denominator corresponding to MSW.

- Null hypothesis:
  H_0: μ1=μ2=…=μc

- Alternative hypothesis:
  H_1:Not all μj are equal (j=1,…c)

Reject Ho if F > Fu



ANOVA Summary Table

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square (Variance)	F
Among groups Within Groups Total	c-1 n-c n-1	SSA SSW SST	MSA= SSA/(c-1) MSW=SSW/(n-c)	F=MSA/MSW

ANOVA Example

Page Replacement Algorithm

	A	B	C	D
	11	12	18	11
	13	14	16	12
	17	17	18	16
	17	19	20	15
	15	21	22	14
	16	18	15	17
	14	19	17	13
	10	18	21	16
	12	16	16	17
	14	18	20	18
Mean	13.9	17.2	18.3	14.9

Grand Mean	16.075				
SSA	47.30625	12.65625	49.50625	13.80625	**123.275**
SSW	213.5				
SST	336.775				
MSA	41.091667				
MSW	5.9305556				
F	6.93				
df numer.	3				
df denom.	36				
Fu	2.87	(from table)			

F=MSA/MSW

c-1 = 4-1

n-c = 40-4

F > Fu ⇨ reject Ho. Algorithms A, B, C, and D have a significant difference at 0.05 level of significance.

α

ANOVA With Excel

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Column 1	10	139	13.9	5.877778
Column 2	10	172	17.2	6.844444
Column 3	10	183	18.3	5.566667
Column 4	10	149	14.9	5.433333

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	123.275	3	41.09167	6.928806	0.000844	2.866265
Within Groups	213.5	36	5.930556			
Total	336.775	39				

Since the p-value is less than $\alpha = 0.05$, reject Ho.

Multiple Comparisons: The Tukey-Kramer Procedure

- If Ho is rejected, then the question is “Which groups are different?”
- Use the Tukey-Kramer procedure to compare all pairs of groups simultaneously.
- Must compute the differences $\overline{X}_j - \overline{X}_{j'}$ for $j \neq j'$ among all c(c-1)/2 pairs of means.

Multiple Comparisons: The Tukey-Kramer Procedure

- Obtain the critical range:

$$\text{critical range} = q_u \sqrt{\frac{MSW}{2}\left(\frac{1}{n_j} + \frac{1}{n_{j'}}\right)}$$

where q_u is the upper-tail critical value from a *Studentized range** distribution with c degrees of freedom in the numerator and (n-c) degrees of freedom in the denominator.

* See statistical table.

Multiple Comparisons: The Tukey-Kramer Procedure

- A pair is considered significantly different if the absolute difference between the sample means exceeds the critical range.

Multiple Comparisons: The Tukey-Kramer Procedure

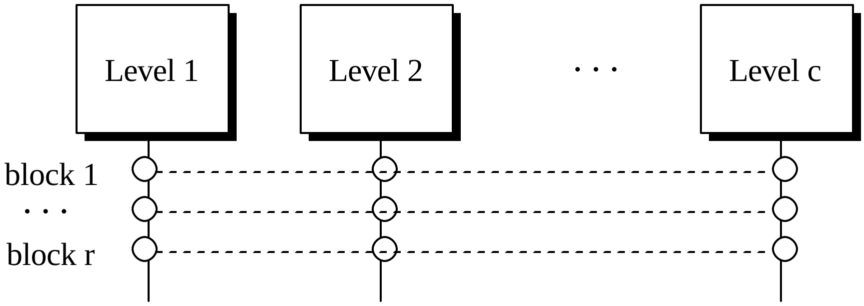
			critical range	
|XA-XB|	3.3	>	2.9341	A significantly different than B
|XA-XC|	4.4	>	2.9341	A significantly different than C
|XA-XD|	1	<	2.9341	A not significantly different than D
|XB-XC|	1.1	<	2.9341	B not significantly different than C
|XB-XD|	2.3	<	2.9341	B not significantly different than D
|XC-XD|	3.4	>	2.9341	C significantly different than D

qu	3.81	(from table)
MSW	5.930556	

Reviewing ANOVA Assumptions

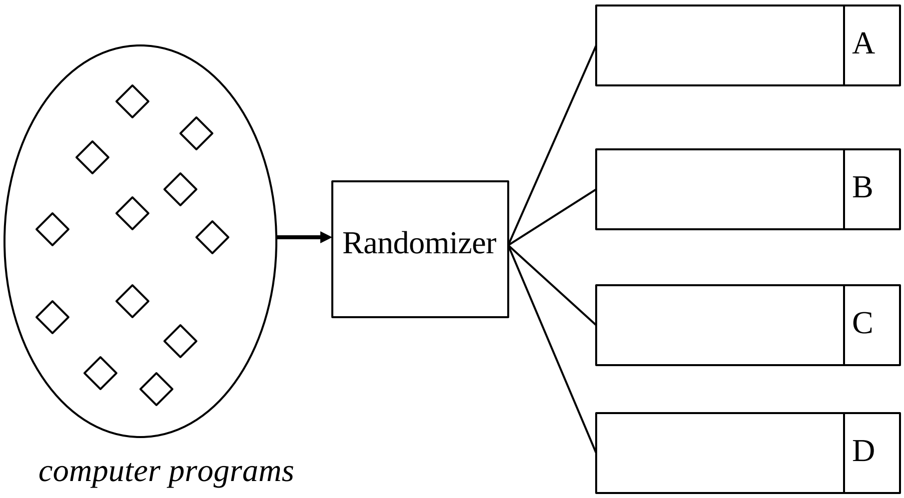
- Randomness and independence: must always be met.
- Normality: ANOVA F test is robust as long as distributions are not extremely different from a normal distribution particularly for large samples.
- Homogeneity of variance: $s_1^2 = s_2^2 = ... = s_c^2$
  - If unequal sample sizes between groups, different variances is a problem.
  - Should try to use same-size groups.

Randomized Block Model

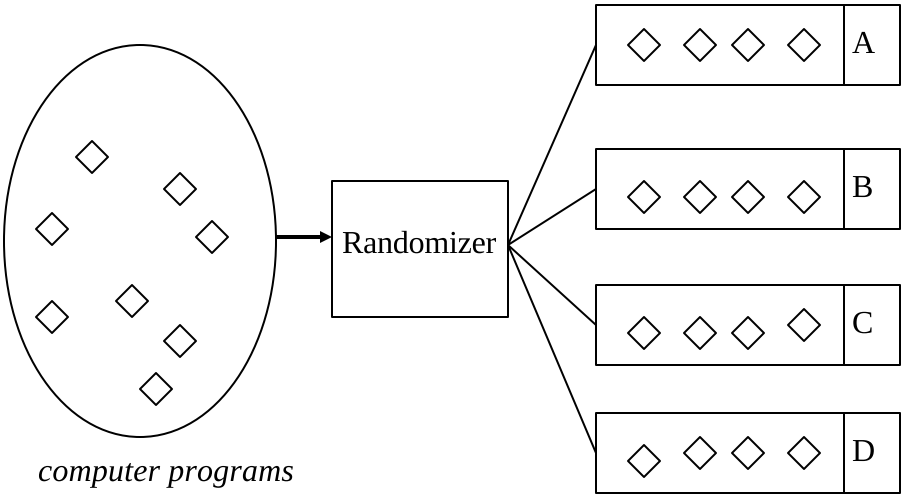


- each block contains the response of the same item to the c levels of the factor being analyzed.
- Purpose: remove as much block or subject variability as possible by reducing experimental error.

Randomized Block Model



Randomized Block Model

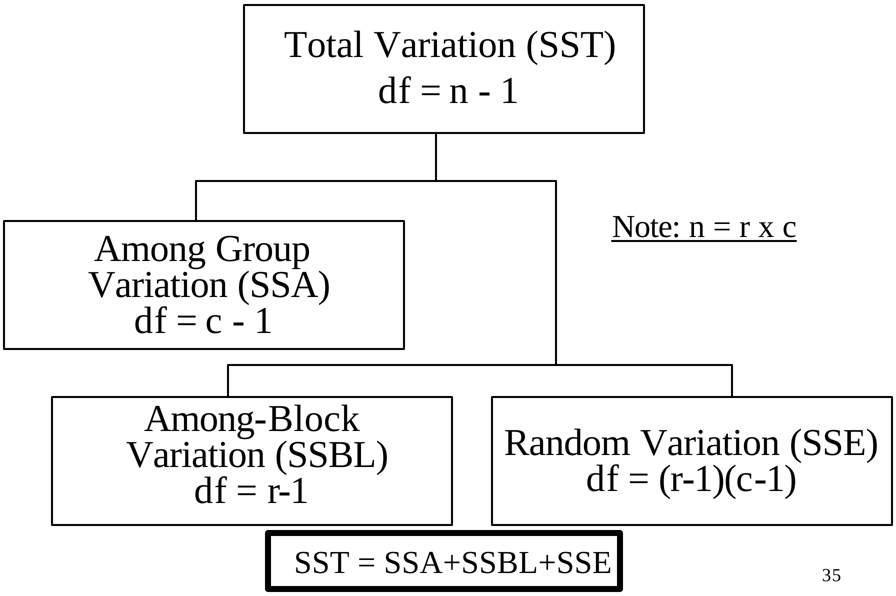


Page Replacement Algorithm

	A	B	C	D	
block 1	11.0	12.0	18.0	11.0	*program 1*
block 2	13.0	14.0	19.0	12.0	*program 2*
block 3	17.0	18.4	23.4	16.5	*program 3*
block 4	14.0	14.9	20.0	12.5	*program 4*
block 5	15.0	16.0	21.0	13.5	*program 5*

$r = 5$
$c = 4$

ANOVA Model for Randomized Block Design



Total Variation (SST)
df = n - 1

Among Group
Variation (SSA)
df = c - 1

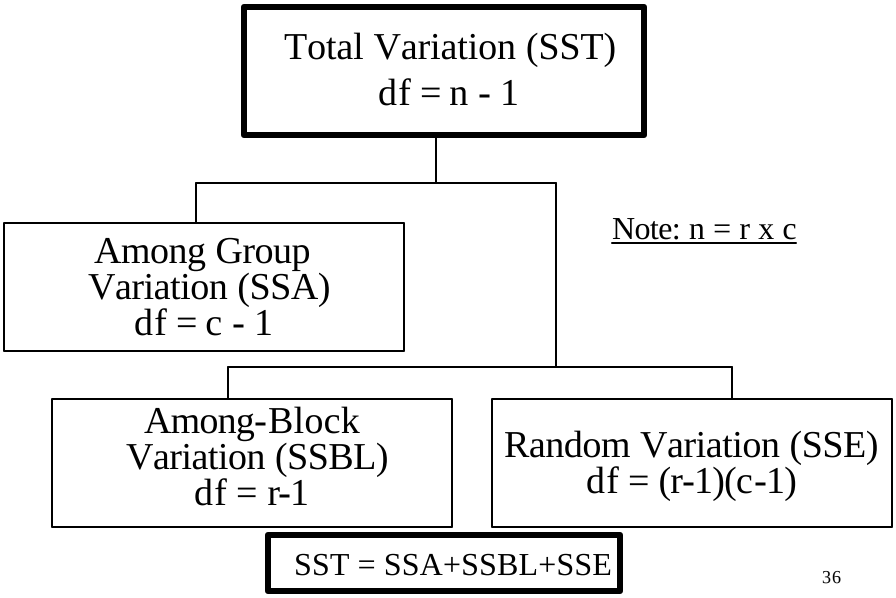
Among-Block
Variation (SSBL)
df = r-1

Random Variation (SSE)
df = (r-1)(c-1)

Note: n = r x c

SST = SSA+SSBL+SSE

ANOVA Model for Randomized Block Design



Total Variation (SST)
df = n - 1

Among Group
Variation (SSA)
df = c - 1

Among-Block
Variation (SSBL)
df = r-1

Random Variation (SSE)
df = (r-1)(c-1)

Note: n = r x c

SST = SSA+SSBL+SSE

SST (Sum of Squares Total)

$$SST = \sum_{j=1}^{c}\sum_{i=1}^{r}\left(X_{ij} - \overline{\overline{X}}\right)^2$$

where

$$\overline{\overline{X}} = \frac{\sum_{j=1}^{c}\sum_{i=1}^{r} X_{ij}}{rc}$$: overall or grand mean.

X_{ij} : observation in i-th block and level j.

r : number of blocks.

Page Replacement Algorithm

	A	B	C	D	
block 1	11.0	12.0	18.0	11.0	*program 1*
block 2	13.0	14.0	19.0	12.0	*program 2*
block 3	17.0	18.4	23.4	16.5	*program 3*
block 4	14.0	14.9	20.0	12.5	*program 4*
block 5	15.0	16.0	21.0	13.5	*program 5*

Grand Mean = 15.61

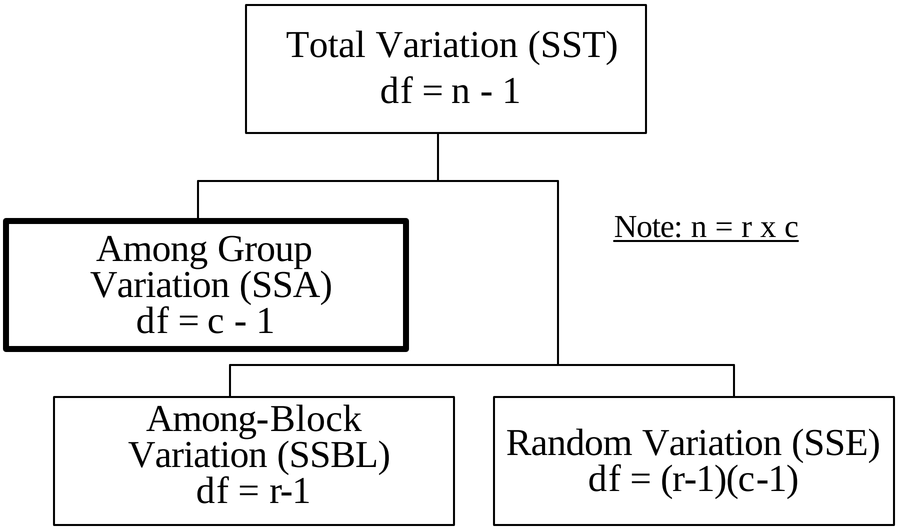
$SST = (11.0-15.61)^2+(13.0-15.61)^2+\ldots+(15.0-15.61)^2+$

$\ldots$

$(11.0-15.61)^2+(12.0-15.61)^2+\ldots+(13.5-15.61)^2$

$= 232.44$

ANOVA Model for Randomized Block Design



Total Variation (SST)
df = n - 1

Note: n = r x c

Among Group
Variation (SSA)
df = c - 1

Among-Block
Variation (SSBL)
df = r-1

Random Variation (SSE)
df = (r-1)(c-1)

SST = SSA+SSBL+SSE

SSA (Sum of Squares Among Group)

$$SSA = r\sum_{j=1}^{c}\left(\overline{X}_{.j} - \overline{\overline{X}}\right)^{2}$$

where

$$\overline{X}_{.j} = \frac{\sum_{i=1}^{r} X_{ij}}{r}: \text{ group mean.}$$

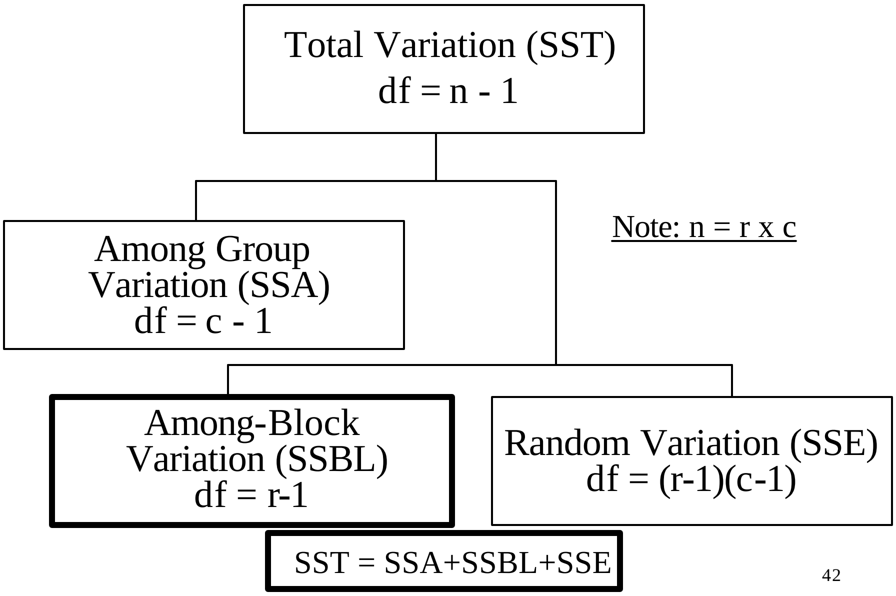
Page Replacement Algorithm

	A	B	C	D	
block 1	11.0	12.0	18.0	11.0	*program 1*
block 2	13.0	14.0	19.0	12.0	*program 2*
block 3	17.0	18.4	23.4	16.5	*program 3*
block 4	14.0	14.9	20.0	12.5	*program 4*
block 5	15.0	16.0	21.0	13.5	*program 5*
Mean	14.0	15.1	20.3	13.1	

Grand Mean 15.61

$$SSA = 5 * [\ (14.0-15.61)^2+(15.1-15.61)^2+\ldots+(13.1-15.61)^2]$$
$$= 155.018$$

ANOVA Model for Randomized Block Design



Total Variation (SST)
df = n - 1

Note: n = r x c

Among Group
Variation (SSA)
df = c - 1

Among-Block
Variation (SSBL)
df = r-1

Random Variation (SSE)
df = (r-1)(c-1)

SST = SSA+SSBL+SSE

SSBL (Sum of Squares Among Blocks)

$$SSBL = c\sum_{i=1}^{r}\left(\overline{X}_{i.} - \overline{\overline{X}}\right)^2$$

where

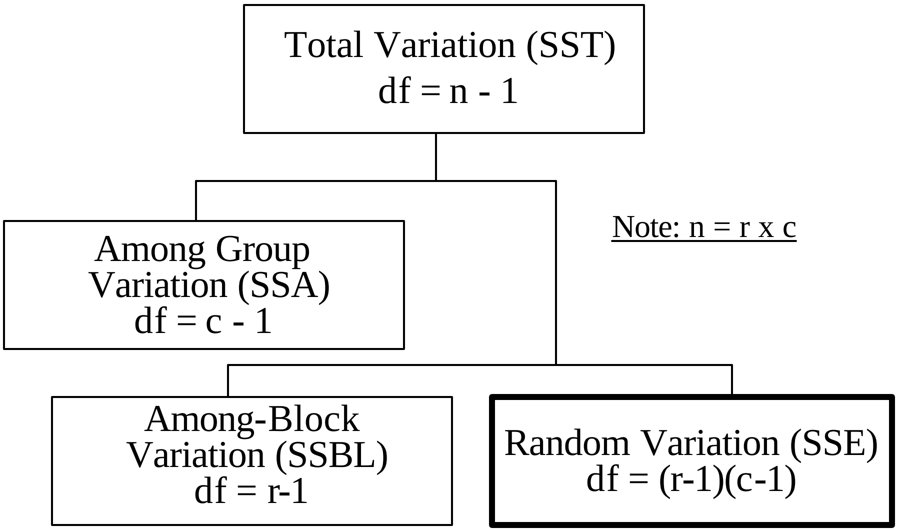
$$\overline{X}_{i.} = \frac{\sum_{j=1}^{c} X_{ij}}{c} : \text{ block mean.}$$

	Page Replacement Algorithm					
	A	**B**	**C**	**D**		Mean
block 1	11.0	12.0	18.0	11.0	*program 1*	13.0
block 2	13.0	14.0	19.0	12.0	*program 2*	14.5
block 3	17.0	18.4	23.4	16.5	*program 3*	18.8
block 4	14.0	14.9	20.0	12.5	*program 4*	15.4
block 5	15.0	16.0	21.0	13.5	*program 5*	16.4
Mean	14.0	15.1	20.3	13.1		

Grand Mean 15.61

SSBL = 4 * $[(13.0-15.61)^2+(14.5-15.61)^2+\ldots+(16.4-15.61)^2]$
= 76.133

ANOVA Model for Randomized Block Design



Total Variation (SST)
df = n - 1

Among Group
Variation (SSA)
df = c - 1

Among-Block
Variation (SSBL)
df = r-1

Random Variation (SSE)
df = (r-1)(c-1)

Note: n = r x c

SST = SSA+SSBL+SSE

SSE (Random Error)

$$SSE = \sum_{j=1}^{c}\sum_{i=1}^{r}\left(X_{ij} - X_{.j} - X_{i.} + \overline{\overline{X}}\right)^2$$

Page Replacement Algorithm

	A	B	C	D		Mean
block 1	11.0	12.0	18.0	11.0	*program 1*	13.0
block 2	13.0	14.0	19.0	12.0	*program 2*	14.5
block 3	17.0	18.4	23.4	16.5	*program 3*	18.8
block 4	14.0	14.9	20.0	12.5	*program 4*	15.4
block 5	15.0	16.0	21.0	13.5	*program 5*	16.4
Mean	14.0	15.1	20.3	13.1		

Grand Mean 15.61

$$SSE = (11.0-14.0-13.0-15.61)^2+(13.0-14.0-14.5+15.61)^2+$$

$$\ldots.$$

$$(21.0-20.3-16.4+15.61)^2+(13.5-13.1-16.4+15.61)^2$$

$$= 1.287$$

ANOVA Model: Mean Squares

$$MSA = \frac{SSA}{c-1}$$

$$MSBL = \frac{SSBL}{r-1}$$

$$MSE = \frac{SSE}{(r-1)(c-1)}$$

The mean squares are variances!

If there are no real differences among the c groups, MSA, MSBL, and MSE provide estimates for the variance inherent in the data.

ANOVA Hypothesis Testing

$$H_0 : m_{.1} = m_{.2} = ... = m_{.c}$$

$$H_1 : \text{Not all } m_{.j} \ (j = 1,...,c) \text{ are equal.}$$

F-Test statistic: $F = \frac{MSA}{MSE}$

The F-test statistic follows an F distribution with (c-1) degrees
Of freedom in the numerator and (r-1)(c-1) in the denominator.

Reject H_0 if $F > Fu$

Page Replacement Algorithm

	A	B	C	D		Mean
block 1	11.0	12.0	18.0	11.0	*program 1*	13.0
block 2	13.0	14.0	19.0	12.0	*program 2*	14.5
block 3	17.0	18.4	23.4	16.5	*program 3*	18.8
block 4	14.0	14.9	20.0	12.5	*program 4*	15.4
block 5	15.0	16.0	21.0	13.5	*program 5*	16.4
Mean	14.0	15.1	20.3	13.1		

Grand Mean	15.61	
SSA	155.018	
SSE	1.287	
SSBL	76.133	
SST	232.438	
MSA	51.67267	
MSBL	19.03325	
MSE	0.10725	
F	481.80	F=MSA/MSE
df numer.	3	
df denom.	12	
Fu	7.23	(from table)

F > Fu ⇒ reject Ho

Anova: Two-Factor Without Replication

SUMMARY	*Count*	*Sum*	*Average*	*Variance*
Row 1	4	52	13	11.333
Row 2	4	58	14.5	9.667
Row 3	4	75.3	18.825	9.949
Row 4	4	61.4	15.35	10.590
Row 5	4	65.5	16.375	10.563
Column 1	5	70	14	5
Column 2	5	75.3	15.06	5.638
Column 3	5	101.4	20.28	4.292
Column 4	5	65.5	13.1	4.425

MSA

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows (blocks)	76.133	4	19.03325	177.4662	1.46E-10	3.25916
Columns (groups)	155.018	3	51.67267	481.79643	9.11E-13	3.4903
Error	1.287	12	0.10725			
Total	232.438	19				

MSE

=MSA/MSE

Since the p-value is less than α = 0.05, reject Ho.
Since F > F critical, reject Ho.

Estimated Relative Efficiency (RE)

$$RE = \frac{\overbrace{(r-1)MSBL}^{\text{SSBL}} + r(c-1)MSE}{\underbrace{(rc-1)}_{\text{n-1}}MSE}$$

- Used to assess if blocking results in an increase in precision in comparing the different groups.

MSA	51.67267
MSBL	19.03325
MSE	0.10725

RE = 38.2

- If blocking is not used, we would need 38.2 times as many observations to obtain the same precision in comparing the groups.

Multiple Comparisons: The Tukey-Kramer Procedure

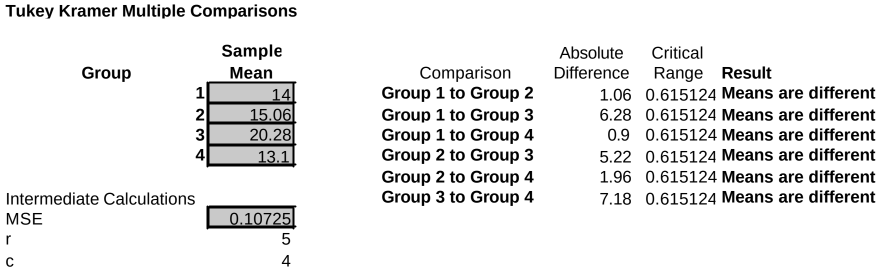
- Obtain the critical range:

$$\text{critical range} = q_u \sqrt{\frac{MSE}{r}}$$

where q_u is the upper-tail critical value from a *Studentized range** distribution with c degrees of freedom in the numerator and (r-1)(c-1) degrees of freedom in the denominator.

(See Statistical Tables).

Multiple Comparisons: The Tukey-Kramer Procedure



Tukey Kramer Multiple Comparisons

Group	**Sample Mean**
1	14
2	15.06
3	20.28
4	13.1

Intermediate Calculations	
MSE	0.10725
r	5
c	4

Comparison	Absolute Difference	Critical Range	**Result**
Group 1 to Group 2	1.06	0.615124	**Means are different**
Group 1 to Group 3	6.28	0.615124	**Means are different**
Group 1 to Group 4	0.9	0.615124	**Means are different**
Group 2 to Group 3	5.22	0.615124	**Means are different**
Group 2 to Group 4	1.96	0.615124	**Means are different**
Group 3 to Group 4	7.18	0.615124	**Means are different**