CS 147:
Computer Systems Performance Analysis
One-Factor Experiments

2015-06-15

CS147

CS 147:
Computer Systems Performance Analysis
One-Factor Experiments

Overview

Overview
Introduction
The Model
Finding Effects
Calculating Errors
ANOVA
Allocation
Analysis
Verifying Assumptions
Unequal Sample Sizes

Characteristics of One-Factor Experiments

- Useful if there's only one important categorical factor with more than two interesting alternatives
  - Methods reduce to 2^1 factorial designs if only two choices
- If single variable isn't categorical, should use regression instead
- Method allows multiple replications

Characteristics of One-Factor Experiments

- Useful if there's only one important categorical factor with more than two interesting alternatives
  - Methods reduce to 2^1 factorial designs if only two choices
- If single variable isn't categorical, should use regression instead
- Method allows multiple replications

Comparing Truly Comparable Options

- Evaluating single workload on multiple machines
- Trying different options for single component
- Applying single suite of programs to different compilers

Comparing Truly Comparable Options
- Evaluating single workload on multiple machines
- Trying different options for single component
- Applying single suite of programs to different compilers

When to Avoid It

- Incomparable “factors”
  - E.g., measuring vastly different workloads on single system
- Numerical factors
  - Won’t predict any untested levels
  - Regression usually better choice
- Related entries across level
  - Use two-factor design instead

When to Avoid It

- Incomparable “factors”
  - E.g., measuring vastly different workloads on single system
- Numerical factors
  - Won’t predict any untested levels
  - Regression usually better choice
- Related entries across level
  - Use two-factor design instead

An Example One-Factor Experiment

- Choosing authentication server for single-sized messages
- Four different servers are available
- Performance measured by response time
  - Lower is better

An Example One-Factor Experiment

- Choosing authentication server for single-sized messages
- Four different servers are available
- Performance measured by response time
  - Lower is better

The One-Factor Model

- $y_{ij} = \mu + \alpha_j + e_{ij}$
- y_{ij} is i^{th} response with factor set at level j
- μ is mean response
- α_j is effect of alternative j

$$\sum \alpha_j = 0$$

- e_{ij} is error term

$$\sum e_{ij} = 0$$

The One-Factor Model
- $y_{ij} = \mu + \alpha_j + e_{ij}$
- y_{ij} is i^{th} response with factor set at level j
- μ is mean response
- α_j is effect of alternative j
$\sum \alpha_j = 0$
- e_{ij} is error term
$\sum e_{ij} = 0$

One-Factor Experiments With Replications

- Initially, assume r replications at each alternative of factor
- Assuming a alternatives, we have a total of ar observations
- Model is thus

$$\sum_{i=1}^{r}\sum_{j=1}^{a} y_{ij} = ar\mu + r\sum_{j=1}^{a}\alpha_j + \sum_{i=1}^{r}\sum_{j=1}^{a} e_{ij}$$

One-Factor Experiments With Replications

- Initially, assume r replications at each alternative of factor
- Assuming a alternatives, we have a total of ar observations
- Model is thus

$$\sum_{i=1}^{r}\sum_{j=1}^{a} y_{ij} = ar\mu + r\sum_{j=1}^{a}\alpha_j + \sum_{i=1}^{r}\sum_{j=1}^{a} e_{ij}$$

Sample Data for Our Example

- Four alternatives, with four replications each (measured in seconds)

A	B	C	D
0.96	0.75	1.01	0.93
1.05	1.22	0.89	1.02
0.82	1.13	0.94	1.06
0.94	0.98	1.38	1.21

Sample Data for Our Example

- Four alternatives, with four replications each (measured in seconds)

A	B	C	D
0.96	0.75	1.01	0.93
1.05	1.22	0.89	1.02
0.82	1.13	0.94	1.06
0.94	0.98	1.38	1.21

Computing Effects

- Need to figure out μ and α_j
- We have various y_{ij}'s
- Errors should add to zero:

$$\sum_{i=1}^{r}\sum_{j=1}^{a} e_{ij} = 0$$

- Similarly, effects should add to zero:

$$\sum_{j=1}^{a} \alpha_j = 0$$

Computing Effects

- Need to figure out μ and α_j
- We have various y_{ij}'s
- Errors should add to zero:

$$\sum_{i=1}^{r}\sum_{j=1}^{a} e_{ij} = 0$$

- Similarly, effects should add to zero:

$$\sum_{j=1}^{a} \alpha_j = 0$$

Calculating μ

- By definition, sum of errors and sum of effects are both zero:

$$\sum_{i=1}^{r}\sum_{j=1}^{a} y_{ij} = ar\mu + 0 + 0$$

- And thus, μ is equal to **grand mean** of all responses

$$\mu = \frac{1}{ar}\sum_{i=1}^{r}\sum_{j=1}^{a} y_{ij} = \bar{y}_{..}$$

Calculating μ
- By definition, sum of errors and sum of effects are both zero:
$$\sum_{i=1}^{r}\sum_{j=1}^{a} y_{ij} = ar\mu + 0 + 0$$
- And thus, μ is equal to **grand mean** of all responses
$$\mu = \frac{1}{ar}\sum_{i=1}^{r}\sum_{j=1}^{a} y_{ij} = \bar{y}_{..}$$

Calculating μ for Our Example

Thus,

$$\begin{aligned}\mu &= \frac{1}{4 \times 4} \sum_{i=1}^{4} \sum_{j=1}^{4} y_{ij} \\ &= \frac{1}{16} \times 16.29 \\ &= 1.018\end{aligned}$$

Calculating μ for Our Example

Thus,

$$\begin{aligned}\mu &= \frac{1}{4 \times 4} \sum_{i=1}^{4} \sum_{j=1}^{4} y_{ij} \\ &= \frac{1}{16} \times 16.29 \\ &= 1.018\end{aligned}$$

Calculating α_j

- α_j is vector of responses
  - One for each alternative of the factor
- To find vector, find **column means**

$$\bar{y}_{.j} = \frac{1}{r}\sum_{i=1}^{r} y_{ij}$$

- Separate mean for each j
- Can calculate directly from observations

Calculating α_j

- α_j is vector of responses
  - One for each alternative of the factor
- To find vector, find **column means**

$$\bar{y}_{.j} = \frac{1}{r}\sum_{i=1}^{r} y_{ij}$$

- Separate mean for each j
- Can calculate directly from observations

Calculating Column Mean

- We know that y_{ij} is defined to be

$$y_{ij} = \mu + \alpha_j + e_{ij}$$

- So,

$$\begin{aligned} \bar{y}_{\cdot j} &= \frac{1}{r}\sum_{i=1}^{r}(\mu + \alpha_j + e_{ij}) \\ &= \frac{1}{r}\left(r\mu + r\alpha_j + \sum_{i=1}^{r} e_{ij}\right) \end{aligned}$$

Calculating Column Mean
- We know that y_{ij} is defined to be
$$y_{ij} = \mu + \alpha_j + e_{ij}$$
- So,
$$\begin{aligned} \bar{y}_{\cdot j} &= \frac{1}{r}\sum_{i=1}^{r}(\mu + \alpha_j + e_{ij}) \\ &= \frac{1}{r}\left(r\mu + r\alpha_j + \sum_{i=1}^{r} e_{ij}\right) \end{aligned}$$

Calculating Parameters

- Sum of errors for any given row is zero, so

$$\begin{aligned} \overline{y}_{.j} &= \frac{1}{r}(r\mu + r\alpha_j + 0) \\ &= \mu + \alpha_j \end{aligned}$$

- So we can solve for α_j:

$$\alpha_j = \overline{y}_{.j} - \mu = \overline{y}_{.j} - \overline{y}_{..}$$

Calculating Parameters

- Sum of errors for any given row is zero, so

$$\begin{aligned} \overline{y}_{.j} &= \frac{1}{r}(r\mu + r\alpha_j + 0) \\ &= \mu + \alpha_j \end{aligned}$$

- So we can solve for α_j:

$$\alpha_j = \overline{y}_{.j} - \mu = \overline{y}_{.j} - \overline{y}_{..}$$

Parameters for Our Example

Server	A	B	C	D
Col. Mean	.9425	1.02	1.055	1.055

Subtract μ from column means to get parameters:

Parameters	-.076	.002	.037	.037

Parameters for Our Example

Server A B C D
Col. Mean .9425 1.02 1.055 1.055

Subtract μ from column means to get parameters:

Parameters -.076 .002 .037 .037

Estimating Experimental Errors

- Estimated response is $\hat{y}_{ij} = \mu + \alpha_{ij}$
- But we measured actual responses
  - Multiple responses per alternative
- So we can estimate amount of error in estimated response
- Use methods similar to those used in other types of experiment designs

Estimating Experimental Errors
- Estimated response is $\hat{y}_{ij} = \mu + \alpha_{ij}$
- But we measured actual responses
  - Multiple responses per alternative
- So we can estimate amount of error in estimated response
- Use methods similar to those used in other types of experiment designs

Sum of Squared Errors

- SSE estimates variance of the errors:

$$\text{SSE} = \sum_{i=1}^{r} \sum_{j=1}^{a} e_{ij}^2$$

- We can calculate SSE directly from model and observations
- Also can find indirectly from its relationship to other error terms

Sum of Squared Errors
- SSE estimates variance of the errors:
$$\text{SSE} = \sum_{i=1}^{r} \sum_{j=1}^{a} e_{ij}^2$$
- We can calculate SSE directly from model and observations
- Also can find indirectly from its relationship to other error terms

SSE for Our Example

Calculated directly:

$$
\begin{aligned}
SSE &= (.96 - (1.018 - .076))^2 \\
&\quad + (1.05 - (1.018 - .076))^2 + \ldots \\
&\quad + (.75 - (1.018 + .002))^2 \\
&\quad + (1.22 - (1.018 + .002))^2 + \ldots \\
&\quad + (.93 - (1.018 + .037))^2 \\
&= .3425
\end{aligned}
$$

SSE for Our Example

Calculated directly:

$$
\begin{aligned}
SSE &= (.96 - (1.018 - .076))^2 \\
&\quad + (1.05 - (1.018 - .076))^2 + \ldots \\
&\quad + (.75 - (1.018 + .002))^2 \\
&\quad + (1.22 - (1.018 + .002))^2 + \ldots \\
&\quad + (.93 - (1.018 + .037))^2 \\
&= .3425
\end{aligned}
$$

Allocating Variation

- To allocate variation for model, start by squaring both sides of model equation

$$
\begin{aligned}
y_{ij}^2 &= \mu^2 + \alpha_j^2 + e_{ij}^2 + 2\mu\alpha_j + 2\mu e_{ij} + 2\alpha_j e_{ij} \\
\sum_{i,j} y_{ij}^2 &= \sum_{i,j} \mu^2 + \sum_{i,j} \alpha_j^2 + \sum_{i,j} e_{ij}^2 \\
&\quad + \text{cross-products}
\end{aligned}
$$

- Cross-product terms add up to zero

Allocating Variation
To allocate variation for model, start by squaring both sides of model equation
$y_{ij}^2 = \mu^2 + \alpha_j^2 + e_{ij}^2 + 2\mu\alpha_j + 2\mu e_{ij} + 2\alpha_j e_{ij}$
$\sum_{i,j} y_{ij}^2 = \sum_{i,j} \mu^2 + \sum_{i,j} \alpha_j^2 + \sum_{i,j} e_{ij}^2$ + cross-products
Cross-product terms add up to zero

Variation In Sum of Squares Terms

$$
\begin{aligned}
\mathrm{SSY} &= \mathrm{SS0} + \mathrm{SSA} + \mathrm{SSE} \\
\mathrm{SSY} &= \sum_{i,j} y_{ij}^2 \\
\mathrm{SS0} &= \sum_{i=1}^{r}\sum_{j=1}^{a} \mu^2 = ar\mu^2 \\
\mathrm{SSA} &= \sum_{i=1}^{r}\sum_{j=1}^{a} \alpha_j^2 = r\sum_{j=1}^{a} \alpha_j^2
\end{aligned}
$$

Gives another way to calculate SSE

2015-06-15

CS147
└ANOVA
 └Allocation
  └Variation In Sum of Squares Terms

Variation In Sum of Squares Terms

SSY = SS0 + SSA + SSE
SSY = $\sum_{i,j} y_{ij}^2$
SS0 = $\sum_{i=1}^{r}\sum_{j=1}^{a} \mu^2 = ar\mu^2$
SSA = $\sum_{i=1}^{r}\sum_{j=1}^{a} \alpha_j^2 = r\sum_{j=1}^{a} \alpha_j^2$

Gives another way to calculate SSE

Sum of Squares Terms for Our Example

- SSY = 16.9615
- SS0 = 16.58256
- SSA = .03377
- So SSE must equal 16.9615-16.58256-.03377
  - $= 0.3425$
  - Matches our earlier SSE calculation

Sum of Squares Terms for Our Example
- SSY = 16.9615
- SS0 = 16.58256
- SSA = .03377
- So SSE must equal 16.9615-16.58256-.03377
  - = 0.3425
  - Matches our earlier SSE calculation

Assigning Variation

- SST is total variation
- SST = SSY − SS0 = SSA + SSE
- Part of total variation comes from model
- Part comes from experimental errors
- A good model explains a lot of variation

Assigning Variation

- SST is total variation
- SST = SSY − SS0 = SSA + SSE
- Part of total variation comes from model
- Part comes from experimental errors
- A good model explains a lot of variation

Assigning Variation in Our Example

- SST $=$ SSY $-$ SS0 $=$ 0.376244
- SSA $=$ .03377
- SSE $=$ .3425
- Percentage of variation explained by server choice:

$$= 100 \times \frac{.03377}{.3762} = 8.97\%$$

2015-06-15

CS147
└ANOVA
  └Allocation
    └Assigning Variation in Our Example

Assigning Variation in Our Example

- SST $=$ SSY $-$ SS0 $=$ 0.376244
- SSA $=$ .03377
- SSE $=$ .3425
- Percentage of variation explained by server choice:

$$= 100 \times \frac{.03377}{.3762} = 8.97\%$$

Analysis of Variance

- Percentage of variation explained can be large or small
- Regardless of size, may or may not be statistically significant
- To determine significance, use ANOVA procedure
  - Assumes normally distributed errors

Analysis of Variance

- Percentage of variation explained can be large or small
- Regardless of size, may or may not be statistically significant
- To determine significance, use ANOVA procedure
  - Assumes normally distributed errors

Running ANOVA

- Easiest to set up tabular method
- Like method used in regression models
  - Only slight differences
- Basically, determine ratio of Mean Squared of A (parameters) to Mean Squared Errors
- Then check against F-table value for number of degrees of freedom

Running ANOVA

- Easiest to set up tabular method
- Like method used in regression models
  - Only slight differences
- Basically, determine ratio of Mean Squared of A (parameters) to Mean Squared Errors
- Then check against F-table value for number of degrees of freedom

ANOVA Table for One-Factor Experiments

Component	Sum of Squares	% of Variation	Degrees of Freedom	Mean Square	F-Computed	F-Table
y	$\mathrm{SSY} = \sum y_{ij}^2$		N			
$\bar{y}_{..}$	$\mathrm{SS0} = N\mu^2$		1			
$y - \bar{y}_{..}$	$\mathrm{SST} = \mathrm{SSY} - \mathrm{SS0}$	100	$N-1$			
A	$\mathrm{SSA} = r\sum \alpha_j^2$	$\frac{\mathrm{SSA}}{\mathrm{SST}}$	$a-1$	$\mathrm{MSA} = \frac{\mathrm{SSA}}{a-1}$	$\frac{\mathrm{MSA}}{\mathrm{MSE}}$	$F[1-\alpha; a-1, N-a]$
e	$\mathrm{SSE} = \mathrm{SST} - \mathrm{SSA}$	$\frac{\mathrm{SSE}}{\mathrm{SST}}$	$N-a$	$\mathrm{MSE} = \frac{\mathrm{SSE}}{N-a}$		

$N = ar \qquad s_e = \sqrt{\mathrm{MSE}}$

ANOVA Procedure for Our Example

Component	Sum of Squares	% of Variation	Degrees of Freedom	Mean Square	F-Computed	F-Table
y	16.96		16			
$\bar{y}_{..}$	16.58		1			
$y - \bar{y}_{..}$	0.376	100	15			
A	.034	9.0	3	.011	0.394	2.61
e	.342	91.0	12	.028		

ANOVA Procedure for Our Example

Interpretation of Sample ANOVA

- Done at 90% level
- F-computed is .394
- Table entry at 90% level with $n = 3$ and $m = 12$ is 2.61
- Thus, servers are not significantly different

Interpretation of Sample ANOVA

- Done at 90% level
- F-computed is .394
- Table entry at 90% level with $n = 3$ and $m = 12$ is 2.61
- Thus, servers are not significantly different

One-Factor Experiment Assumptions

- Analysis of one-factor experiments makes the usual assumptions:
  - Effects of factors are additive
  - Errors are additive
  - Errors are independent of factor alternatives
  - Errors are normally distributed
  - Errors have same variance at all alternatives
- How do we tell if these are correct?

One-Factor Experiment Assumptions
- Analysis of one-factor experiments makes the usual assumptions:
  - Effects of factors are additive
  - Errors are additive
  - Errors are independent of factor alternatives
  - Errors are normally distributed
  - Errors have same variance at all alternatives
- How do we tell if these are correct?

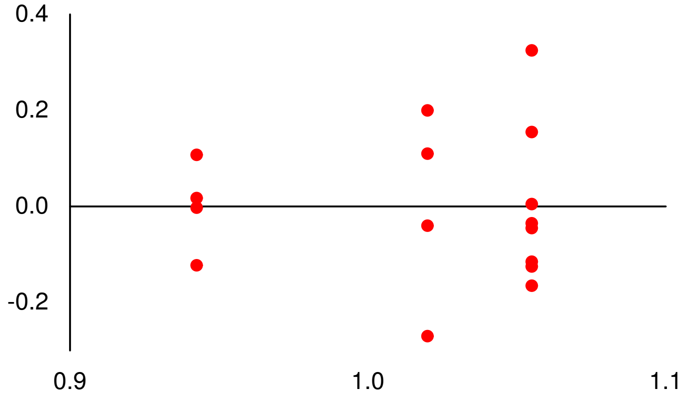
Visual Diagnostic Tests

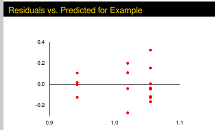
- Similar to those done before
  - Residuals vs. predicted response
  - Normal quantile-quantile plot
  - Residuals vs. experiment number

Visual Diagnostic Tests

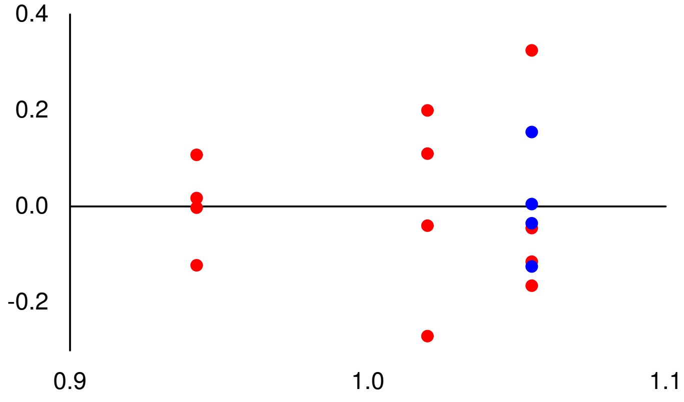
- Similar to those done before
  - Residuals vs. predicted response
  - Normal quantile-quantile plot
  - Residuals vs. experiment number

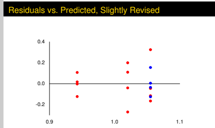
Residuals vs. Predicted for Example





Residuals vs. Predicted, Slightly Revised





In the alternate rendering, the predictions for server D are shown in blue so they can be distinguished from server C.

What Does The Plot Tell Us?

- Analysis assumed size of errors was unrelated to factor alternatives
- Plot tells us something entirely different
  - Very different spread of residuals for different factors
- Thus, one-factor analysis is not appropriate for this data
  - Compare individual alternatives instead
  - Use pairwise confidence intervals

What Does The Plot Tell Us?
- Analysis assumed size of errors was unrelated to factor alternatives
- Plot tells us something entirely different
  - Very different spread of residuals for different factors
- Thus, one-factor analysis is not appropriate for this data
  - Compare individual alternatives instead
  - Use pairwise confidence intervals

Could We Have Figured This Out Sooner?

- Yes!
- Look at original data
- Look at calculated parameters
- Model says C & D are identical
- Even cursory examination of data suggests otherwise

Could We Have Figured This Out Sooner?
- Yes!
- Look at original data
- Look at calculated parameters
- Model says C & D are identical
- Even cursory examination of data suggests otherwise

Looking Back at the Data

A	B	C	D
0.96	0.75	1.01	0.93
1.05	1.22	0.89	1.02
0.82	1.13	0.94	1.06
0.94	0.98	1.38	1.21

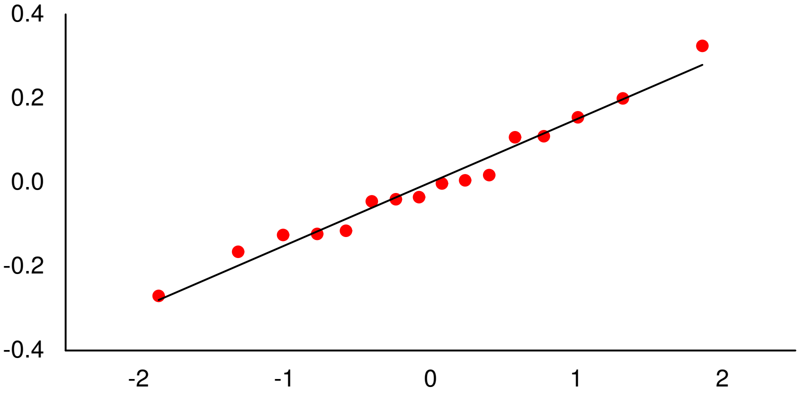
Parameters:

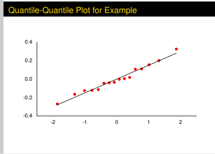
-.076	.002	.037	.037

Looking Back at the Data

A B C D
0.96 0.75 1.01 0.93
1.05 1.22 0.89 1.02
0.82 1.13 0.94 1.06
0.94 0.98 1.38 1.21
Parameters:
-.076 .002 .037 .037

Quantile-Quantile Plot for Example





What Does This Plot Tell Us?

- Overall, errors are normally distributed
- If we only did quantile-quantile plot, we'd think everything was fine
- The lesson: **test ALL assumptions, not just one or two**

What Does This Plot Tell Us?
- Overall, errors are normally distributed
- If we only did quantile-quantile plot, we'd think everything was fine
- The lesson: **test ALL assumptions, not just one or two**

One-Factor Confidence Intervals

- Estimated parameters are random variables
  - Thus, can compute confidence intervals
- Basic method is same as for confidence intervals on $2^k r$ design effects
- Find standard deviation of parameters
  - Use that to calculate confidence intervals
  - Possible typo in book, p. 336, example 20.6, in formula for calculating α_j
  - Also might be typo on p. 335: degrees of freedom is $a(r-1)$, not $r(a-1)$

One-Factor Confidence Intervals
- Estimated parameters are random variables
  - Thus, can compute confidence intervals
- Basic method is same as for confidence intervals on $2^k r$ design effects
- Find standard deviation of parameters
  - Use that to calculate confidence intervals
  - Possible typo in book, p. 336, example 20.6, in formula for calculating α_j
  - Also might be typo on p. 335: degrees of freedom is $a(r-1)$, not $r(a-1)$

Confidence Intervals For Example Parameters

- $s_e = .158$
- Standard deviation of $\mu = .040$
- Standard deviation of $\alpha_j = .069$
- 95% confidence interval for $\mu = (.932, 1.10)$
- 95% CI for $\alpha_1 = (-.225, .074)$
- 95% CI for $\alpha_2 = (-.148, .151)$
- 95% CI for $\alpha_3 = (-.113, .186)$
- 95% CI for $\alpha_4 = (-.113, .186)$

Confidence Intervals For Example Parameters
- $s_e = .158$
- Standard deviation of $\mu = .040$
- Standard deviation of $\alpha_j = .069$
- 95% confidence interval for $\mu = (.932, 1.10)$
- 95% CI for $\alpha_1 = (-.225, .074)$
- 95% CI for $\alpha_2 = (-.148, .151)$
- 95% CI for $\alpha_3 = (-.113, .186)$
- 95% CI for $\alpha_4 = (-.113, .186)$

Unequal Sample Sizes in One-Factor Experiments

- Don't really need identical replications for all alternatives
- Only slight extra difficulty
- See book example for full details

Unequal Sample Sizes in One-Factor Experiments
- Don't really need identical replications for all alternatives
- Only slight extra difficulty
- See book example for full details

Changes To Handle Unequal Sample Sizes

- Model is the same
- Effects are weighted by number of replications for that alternative:

$$\sum_{j=1}^{a} r_j a_j = 0$$

- Slightly different formulas for degrees of freedom

Changes To Handle Unequal Sample Sizes
- Model is the same
- Effects are weighted by number of replications for that alternative:
$\sum_{j=1}^{a} r_j a_j = 0$
- Slightly different formulas for degrees of freedom