Performance Evaluation:

One-factor Experiments

Hongwei Zhang
http://www.cs.wayne.edu/~hzhang

Experiments should be reproducible: they should all fail in the same way.

--- Finagle’s Rule

Acknowledgement: this lecture is partially based on the slides of Dr. Raj Jain.
Outline

- Model & Computation of Effects
- Estimating Experimental Errors
- Allocation of Variation
- ANOVA Table and F-Test
- Confidence Intervals for Effects
- Unequal Sample Sizes
- Visual Diagnostic Tests
Outline

- Model & Computation of Effects
- Estimating Experimental Errors
- Allocation of Variation
- ANOVA Table and F-Test
- Confidence Intervals for Effects
- Unequal Sample Sizes
- Visual Diagnostic Tests
One-factor experiments: Model

- Used to compare alternatives of a single categorical variable.
  
  \[ y_{ij} = \mu + \alpha_j + e_{ij} \]

  For example, several processors, several caching schemes

  \[
  \begin{align*}
  r & = \text{Number of replications} \\
  y_{ij} & = \text{ith response with jth alternative} \\
  \mu & = \text{mean response} \\
  \alpha_j & = \text{Effect of alternative j} \\
  e_{ij} & = \text{Error term}
  \end{align*}
  \]

  \[
  \sum \alpha_j = 0
  \]
Computation of effects

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

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

*Assuming error sum is 0*
Computation of effects (contd.)

\[
\bar{y}.j = \frac{1}{r} \sum_{i=1}^{r} y_{ij} \\
= \frac{1}{r} \sum_{i=1}^{r} (\mu + \alpha_j + e_{ij}) \\
= \frac{1}{r} \left( r\mu + r\alpha_j + \sum_{i=1}^{r} e_{ij} \right) \\
= \mu + \alpha_j + 0
\]

\[\alpha_j = \bar{y}.j - \mu = \bar{y}.j - \bar{y}..\]
Example:
code size on three diff. processors R, V, Z

<table>
<thead>
<tr>
<th></th>
<th>R</th>
<th>V</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>144</td>
<td>101</td>
<td>130</td>
</tr>
<tr>
<td>2</td>
<td>120</td>
<td>144</td>
<td>180</td>
</tr>
<tr>
<td>3</td>
<td>176</td>
<td>211</td>
<td>141</td>
</tr>
<tr>
<td>4</td>
<td>288</td>
<td>288</td>
<td>374</td>
</tr>
<tr>
<td>5</td>
<td>144</td>
<td>72</td>
<td>302</td>
</tr>
</tbody>
</table>

Entries in a row are unrelated.
(Otherwise, need a two factor analysis.)
Example (contd.)

<table>
<thead>
<tr>
<th>R</th>
<th>V</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>144</td>
<td>101</td>
<td>130</td>
</tr>
<tr>
<td>120</td>
<td>144</td>
<td>180</td>
</tr>
<tr>
<td>176</td>
<td>211</td>
<td>141</td>
</tr>
<tr>
<td>288</td>
<td>288</td>
<td>374</td>
</tr>
<tr>
<td>144</td>
<td>72</td>
<td>302</td>
</tr>
</tbody>
</table>

Col Sum  \[ \sum y_1 = 872 \quad \sum y_2 = 816 \quad \sum y_3 = 1127 \quad \sum y_{..} = 2815 \]

Col Mean  \[ \bar{y}_1 = 174.4 \quad \bar{y}_2 = 163.2 \quad \bar{y}_3 = 225.4 \quad \mu = \bar{y}_{..} = 187.7 \]

Col Effect  \[ \alpha_1 = \bar{y}_1 - \bar{y}_{..} = -13.3 \quad \alpha_2 = \bar{y}_2 - \bar{y}_{..} = -24.5 \quad \alpha_3 = \bar{y}_3 - \bar{y}_{..} = 37.7 \]

- Average processor requires 187.7 bytes of storage.
- The effects of the processors R, V, and Z are -13.3, -24.5, and 37.7, respectively. That is,
  - R requires 13.3 bytes less than an average processor
  - V requires 24.5 bytes less than an average processor, and
  - Z requires 37.7 bytes more than an average processor.
Outline

- Model & Computation of Effects
- Estimating Experimental Errors
- Allocation of Variation
- ANOVA Table and F-Test
- Confidence Intervals for Effects
- Unequal Sample Sizes
- Visual Diagnostic Tests
Estimating experimental errors

- Estimated response for jth alternative:
  \[ \hat{y}_j = \mu + \alpha_j \]

- Error:
  \[ e_{ij} = y_{ij} - \hat{y}_j \]

- Sum of squared errors (SSE):
  \[ SSE = \sum_{i=1}^{r} \sum_{j=1}^{a} e_{ij}^2 \]
Example: for the processor-example

\[
\begin{bmatrix}
144 & 101 & 130 \\
120 & 144 & 180 \\
176 & 211 & 141 \\
288 & 288 & 374 \\
144 & 72 & 302
\end{bmatrix}
\begin{bmatrix}
187.7 & 187.7 & 187.7 \\
187.7 & 187.7 & 187.7 \\
187.7 & 187.7 & 187.7 \\
187.7 & 187.7 & 187.7 \\
187.7 & 187.7 & 187.7
\end{bmatrix}
\]

\[
\begin{bmatrix}
-13.3 & -24.5 & 37.7 \\
-13.3 & -24.5 & 37.7 \\
-13.3 & -24.5 & 37.7 \\
-13.3 & -24.5 & 37.7 \\
-13.3 & -24.5 & 37.7
\end{bmatrix}
\]

\[
\begin{bmatrix}
-30.4 & -62.2 & -95.4 \\
-54.4 & -19.2 & -45.4 \\
1.6 & 47.8 & -84.4 \\
113.6 & 124.8 & 148.6 \\
-30.4 & -91.2 & 76.6
\end{bmatrix}
\]

\[
SSE = (-30.4)^2 + (-54.4)^2 + \cdots + (76.6)^2 = 94365.20
\]
Outline

- Model & Computation of Effects
- Estimating Experimental Errors
- Allocation of Variation
- ANOVA Table and F-Test
- Confidence Intervals for Effects
- Unequal Sample Sizes
- Visual Diagnostic Tests
Allocation of variation

\[ y_{ij} = \mu + \alpha_j + e_{ij} \]

\[ y_{ij}^2 = \mu^2 + \alpha_j^2 + e_{ij}^2 + 2\mu\alpha_j + 2\mu e_{ij} + 2\alpha_j e_{ij} \]

\[ \sum_{i,j} y_{ij}^2 = \sum_{i,j} \mu^2 + \sum_{i,j} \alpha_j^2 + \sum_{i,j} e_{ij}^2 + \text{Cross product terms} \]

\[ SSY = SS0 + SSA + SSE \]

\[ SS0 = \sum_{i=1}^{r} \sum_{j=1}^{a} \mu^2 = ar\mu^2 \]
Allocation of variation (contd.)

\[
SSA = \sum_{i=1}^{r} \sum_{j=1}^{a} \alpha_{ij}^2 = r \sum_{j=1}^{a} \alpha_j^2
\]

- Total variation of \( y \) (SST):

\[
SST = \sum_{i,j} (y_{i,j} - \bar{y}_{..})^2
\]

\[
= \sum_{i,j} y_{i,j}^2 - ar\bar{y}_{..}^2
\]

\[
= SSY - SS0 = SSA + SSE
\]
Example: for the processor-example

\[
\begin{align*}
SSY &= 144^2 + 120^2 + \cdots + 302^2 = 633639 \\
SS0 &= ar\mu^2 \\
&= 3 \times 5 \times (187.7)^2 = 528281.7 \\
SSA &= r \sum_j \alpha_j^2 \\
&= 5[(-13.3)^2 + (-24.5)^2 + (37.6)^2] \\
&= 10992.1 \\
SST &= SSY - SS0 \\
&= 633639.0 - 528281.7 = 105357.3 \\
SSE &= SST - SSA \\
&= 105357.3 - 10992.1 = 94365.2
\end{align*}
\]
Example (contd.)

Percent variation explained by processors $= 100 \times \frac{10992.13}{105357.3} = 10.4\%$

89.6\% of variation in code size is due to experimental errors (programmer differences).

Is 10.4\% statistically significant?
Outline

- Model & Computation of Effects
- Estimating Experimental Errors
- Allocation of Variation
- ANOVA Table and F-Test
- Confidence Intervals for Effects
- Unequal Sample Sizes
- Visual Diagnostic Tests
Analysis of variance (ANOVA)

- Importance ≠ Significance
- Important ⇒ Explains a high percent of variation
- Significance
  ⇒ High contribution to the variation compared to that by errors.
- Degree of freedom
  = Number of independent values required to compute

\[
SS_Y = SS_0 + SSA + SSE
\]
\[
\text{ar} = 1 + (a-1) + a(r-1)
\]

Note that the degrees of freedom also add up.
F-Test

- **Purpose:** To check if SSA is *significantly* greater than SSE.
- **Errors are normally distributed** ⇒ SSE and SSA have chi-square distributions.
  
The ratio \( \frac{SSA/v_A}{SSE/v_e} \) has an F distribution.
  
  where \( v_A = a-1 \) = degrees of freedom for SSA
  
  \( v_e = a(r-1) \) = degrees of freedom for SSE
  
  Computed ratio > \( F_{[1-\alpha; v_A, v_e]} \) ⇒
  
  SSA is significantly higher than SSE.

SSA/\( v_A \) is called mean square of A or (MSA).

Similary, MSE=SSE/\( v_e \)
# ANOVA table for one-factor experiments

<table>
<thead>
<tr>
<th>Component</th>
<th>Sum of Squares</th>
<th>%Variation</th>
<th>DF</th>
<th>Mean Square</th>
<th>F-Comp.</th>
<th>F-Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$</td>
<td>$SSY = \sum y_{ij}^2$</td>
<td></td>
<td>$ar$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{y}..$</td>
<td>$SS0 = ar \mu^2$</td>
<td></td>
<td>$1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y-\bar{y}..$</td>
<td>$SST = SSY - SS0$</td>
<td>100</td>
<td>$ar - 1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A$</td>
<td>$SSA = r \sum \alpha_i^2$</td>
<td>$100 \left( \frac{SSA}{SST} \right)$</td>
<td>$a - 1$</td>
<td>$MSA = \frac{SSA}{a - 1}$</td>
<td>$MSA \over MSE$</td>
<td>$F \left[ 1 - \alpha; a - 1, a(r - 1) \right]$</td>
</tr>
<tr>
<td>$e$</td>
<td>$SSE = SST - SSA$</td>
<td>$100 \left( \frac{SSE}{SST} \right)$</td>
<td>$a(r - 1)$</td>
<td>$MSE = \frac{SSE}{a(r - 1)}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example: for the processor-example

<table>
<thead>
<tr>
<th>Component</th>
<th>Sum of Squares</th>
<th>%Variation</th>
<th>DF</th>
<th>Mean Square</th>
<th>F-Comp.</th>
<th>F-Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>633639.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y..</td>
<td>528281.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y-y..</td>
<td>105357.31</td>
<td>100.0%</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>10992.13</td>
<td>10.4%</td>
<td>2</td>
<td>5496.1</td>
<td>0.7</td>
<td>2.8</td>
</tr>
<tr>
<td>Errors</td>
<td>94365.20</td>
<td>89.6%</td>
<td>12</td>
<td>7863.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ s_e = \sqrt{\text{MSE}} = \sqrt{7863.77} = 88.68 \]

- Computed F-value < F from Table
- The variation in the code sizes is mostly due to experimental errors and not because of any significant difference among the processors.

E.g., programmer difference
Outline

- Model & Computation of Effects
- Estimating Experimental Errors
- Allocation of Variation
- ANOVA Table and F-Test
- Confidence Intervals for Effects
- Unequal Sample Sizes
- Visual Diagnostic Tests
Confidence intervals for effects

- **Estimates are random variables**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>$\bar{y}.\quad$</td>
<td>$\frac{s_e^2}{ar}$</td>
</tr>
<tr>
<td>$\alpha_j$</td>
<td>$\bar{y}.j - \bar{y}.\quad$</td>
<td>$\frac{s_e^2(a - 1)}{ar}$</td>
</tr>
<tr>
<td>$\mu + \alpha_j$</td>
<td>$\bar{y}.j\quad$</td>
<td>$\frac{s_e^2}{r}$</td>
</tr>
<tr>
<td>$\sum_{j=1}^{a} h_j \alpha_j, \sum_{j=1}^{a} h_j = 0$</td>
<td>$\sum_{j=1}^{a} h_j \bar{y}.j\quad\sum_{j=1}^{a} \frac{s_e^2 h_j^2}{ar}$</td>
<td></td>
</tr>
<tr>
<td>$s_e^2\quad$</td>
<td>$\sum_{i,j} c_{i,j}^2\quad\frac{a(r-1)}{a(r-1)}$</td>
<td></td>
</tr>
</tbody>
</table>

Degrees of freedom for errors = $a(r-1)$

- **For the confidence intervals, use $t$ values at $r(a-1)$ degrees of freedom.**
- **Mean responses:** $\hat{y}_j = \mu + \alpha_j$
- **Contrasts $\sum h_j \alpha_j$: Use for $\alpha_1 - \alpha_2$**
Example: for the processor-example

Error variance \( s^2_e = \frac{94365.2}{12} = 7863.8 \)

\[
\text{Std Dev of errors} = \sqrt{\text{Var. of errors}} = 88.7
\]

\[
\text{Std Dev of } \mu = s_e/\sqrt{ar} = 88.7/\sqrt{15} = 22.9
\]

\[
\text{Std Dev of } \alpha_j = s_e/\sqrt{\{(a-1)/(ar)\}} = 88.7/\sqrt{(2/15)} = 32.4
\]
Example (contd.)

- For 90% confidence, $t_{[0.95; 12]} = 1.782$.
- 90% confidence intervals:

  $\mu = 197.7 \mp (1.782)(22.9) = (146.9, 228.5)$

  $\alpha_1 = -13.3 \mp (1.782)(32.4) = (-71.0, 44.4)$

  $\alpha_2 = -24.5 \mp (1.782)(32.4) = (-82.2, 33.2)$

  $\alpha_3 = 37.6 \mp (1.782)(32.4) = (-20.0, 95.4)$

The code size on an average processor is significantly different from zero.

Processor effects are not significant.
Example (contd.)

- Using $h_1=1$, $h_2=-1$, $h_3=0$, $(\sum h_j=0)$:
  
  Mean $\alpha_1 - \alpha_2 = \bar{y}_1 - \bar{y}_2 = 174.4 - 163.2 = 11.2$

  Std dev of $\alpha_1 - \alpha_2 = \frac{s_e}{\sqrt{(\sum h^2_j/\alpha r)}}$

  $= \frac{88.7}{\sqrt{(2/15)}} = 56.1$

  90% CI for $\alpha_1 - \alpha_2 = 11.2 \pm (1.782)(56.1)$

  $= (-88.7, 111.1)$

  CI includes zero $\Rightarrow$ one isn't superior to other.
Example (contd.)

- Similarly,
  
  90% CI for $\alpha_1 - \alpha_3$

  $$= (174.4 - 225.4) \mp (1.782)(56.1)$$

  $$= (-150.9, 48.9)$$

  90% CI for $\alpha_2 - \alpha_3$

  $$= (163.2 - 225.4) \mp (1.782)(56.1)$$

  $$= (-162.1, 37.7)$$

Any one processor is not superior to another.
Outline

- Model & Computation of Effects
- Estimating Experimental Errors
- Allocation of Variation
- ANOVA Table and F-Test
- Confidence Intervals for Effects
- Unequal Sample Sizes
- Visual Diagnostic Tests
Unequal sample sizes

\[ y_{ij} = \mu + \alpha_j + e_{ij} \]

By definition:

\[ \sum_{j=1}^{a} r_j \alpha_j = 0 \]

Here, \( r_j \) is the number of observations at \( j \)th level.

\( N = \text{total number of observations:} \)

\[ N = \sum_{j=1}^{a} r_j \]
## Parameter estimation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>$\bar{y}..$</td>
<td>$s_e^2/N$</td>
</tr>
<tr>
<td>$\alpha_j$</td>
<td>$\bar{y}. j - \bar{y}..$</td>
<td>$s_e^2(N - r_j)/(N r_j)$</td>
</tr>
<tr>
<td>$\mu + \alpha_j$</td>
<td>$\bar{y}. j$</td>
<td>$s_e^2/r_j$</td>
</tr>
<tr>
<td>$\sum h_j\alpha_j, \sum h_j = 0$</td>
<td>$h_j \bar{y}. j$</td>
<td>$s_e^2 \sum_{j=1}^a (h_j^2/r_j)$</td>
</tr>
<tr>
<td>$s_e^2$</td>
<td>$\sum e_{ij}^2/(N - a)$</td>
<td></td>
</tr>
</tbody>
</table>

Degrees of freedom for errors = $N-a$
## Analysis of variance

<table>
<thead>
<tr>
<th>Component</th>
<th>Sum of Squares</th>
<th>%Variation</th>
<th>DF</th>
<th>Mean Square</th>
<th>F-Comp.</th>
<th>F-Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>SSY = $\sum y_{ij}^2$</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{y}_..$</td>
<td>SS0 = $N \mu^2$</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y-$\bar{y}_..$</td>
<td>SST = SSY - SS0</td>
<td>100</td>
<td>N-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>SSA = $\sum_{j=1}^{a} r_j \alpha_j^2$</td>
<td>a-1</td>
<td></td>
<td>MSA = $\frac{SSA}{a-1}$</td>
<td>$\frac{MSA}{MSE}$</td>
<td>$F_{[1-\alpha; a-1, N-a]}$</td>
</tr>
<tr>
<td>e</td>
<td>SSE = SST - SSA</td>
<td>N-a</td>
<td></td>
<td>MSE = $\frac{SSE}{N-a}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example: for the processor-example

<table>
<thead>
<tr>
<th></th>
<th>R</th>
<th>V</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>144</td>
<td>101</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>144</td>
<td>180</td>
<td></td>
</tr>
<tr>
<td>176</td>
<td>211</td>
<td>141</td>
<td></td>
</tr>
<tr>
<td>288</td>
<td>288</td>
<td>144</td>
<td></td>
</tr>
<tr>
<td>Column Sum</td>
<td>872</td>
<td>744</td>
<td>451</td>
</tr>
<tr>
<td>Column Mean</td>
<td>174.40</td>
<td>186.00</td>
<td>150.33</td>
</tr>
<tr>
<td>Column effect</td>
<td>2.15</td>
<td>13.75</td>
<td>-21.92</td>
</tr>
</tbody>
</table>

- All means are obtained by dividing by the number of observations added.
- The column effects are 2.15, 13.75, and -21.92.
Example (contd.)

\[
\begin{bmatrix}
144 & 101 & 130 \\
120 & 144 & 180 \\
176 & 211 & 141 \\
288 & 288 \\
144
\end{bmatrix}
= \begin{bmatrix}
172.25 & 172.25 & 172.25 \\
172.25 & 172.25 & 172.25 \\
172.25 & 172.25 & 172.25 \\
172.25 & 172.25 \\
172.25
\end{bmatrix}
+ \begin{bmatrix}
2.15 & 13.75 & -21.92 \\
2.15 & 13.75 & -21.92 \\
2.15 & 13.75 & -21.92 \\
2.15 & 13.75 \\
2.15
\end{bmatrix}
\]
\[
+ \begin{bmatrix}
-30.40 & -85.00 & -20.33 \\
-54.40 & -42.00 & 29.67 \\
1.60 & 25.00 & -9.33 \\
113.60 & 102.00 \\
-30.40
\end{bmatrix}
\]
Example (contd.)

- Sums of Squares:
  \[
  \begin{align*}
  \text{SSY} &= \sum y_{ij}^2 = 397375 \\
  \text{SS0} &= N\mu^2 = 356040.75 \\
  \text{SSA} &= 5\alpha_1^2 + 4\alpha_2^2 \\
  &\quad + 3\alpha_3^2 = 2220.38 \\
  \text{SSE} &= (-30.40)^2 + (-54.40)^2 + \cdots \\
  &\quad + (-9.33)^2 = 39113.87 \\
  \text{SST} &= \text{SSY} - \text{SS0} = 41334.25
  \end{align*}
  \]

- Degrees of Freedom:
  \[
  \begin{align*}
  \text{SSY} &= \text{SS0} + \text{SSA} + \text{SSE} \\
  N &= 1 + (a-1) + N-a \\
  12 &= 1 + 2 + 9
  \end{align*}
  \]
Example (contd.)

<table>
<thead>
<tr>
<th>Component</th>
<th>Sum of Squares</th>
<th>%Variation</th>
<th>DF</th>
<th>Mean Square</th>
<th>F-Comp.</th>
<th>F-Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>397375.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y-</td>
<td>356040.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y-y-</td>
<td>41334.25</td>
<td>100.00%</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>2220.38</td>
<td>5.37%</td>
<td>2</td>
<td>1110.19</td>
<td>0.26</td>
<td>3.01</td>
</tr>
<tr>
<td>Errors</td>
<td>39113.87</td>
<td>94.63%</td>
<td>9</td>
<td>4345.99</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ s_e = \sqrt{\text{MSE}} = \sqrt{4345.99} = 65.92 \]

**Conclusion:** Variation due to processors is insignificant as compared to that due to modeling errors.
Outline

- Model & Computation of Effects
- Estimating Experimental Errors
- Allocation of Variation
- ANOVA Table and F-Test
- Confidence Intervals for Effects
- Unequal Sample Sizes
- Visual Diagnostic Tests
Visual diagnostic tests for the one-factor experimental analysis

Assumptions:
1. Factor effects are additive.
2. Errors are additive.
3. Errors are independent of factor levels.
4. Errors are normally distributed.
5. Errors have the same variance for all factor levels.

Tests:
- Residuals versus predicted response:
  No trend $\Rightarrow$ Independence
  Scale of errors $\ll$ Scale of response $\Rightarrow$ Ignore visible trends.
- Normal quantile-quantile plot linear $\Rightarrow$ Normality
Example: for the processor-example

- Horizontal and vertical scales similar
  ⇒ Residuals are not small ⇒ Variation due to factors is small compared to the unexplained variation

- No visible trend in the spread
- Q-Q plot is S-shaped ⇒ shorter tails than normal.
Summary

- Model & Computation of Effects
- *Estimating Experimental Errors*
- *Allocation of Variation*
- *ANOVA Table and F-Test*
- Confidence Intervals for Effects
- *Unequal Sample Sizes*
- *Visual Diagnostic Tests*
Further reading

- Chapters 21: two-factor full factorial design without replications
  - Model: \( y_{ij} = \mu + \alpha_j + \beta_i + e_{ij} \)
    - \( y_{ij} \) = Observation with A at level j and B at level i
    - \( \mu \) = mean response
    - \( \alpha_j \) = effect of factor A at level j
    - \( \beta_i \) = effect of factor B at level i
    - \( e_{ij} \) = error term

- Chapters 22: two-factor full factorial design with replications
  - Model: \( y_{ijk} = \mu + \alpha_j + \beta_i + \gamma_{ij} + e_{ijk} \)
    - \( y_{ijk} \) = Response in the kth replication with factor A at level j and factor B at level i
    - \( \mu \) = mean response
    - \( \alpha_j \) = Effect of factor A at level j
    - \( \beta_i \) = Effect of Factor B at level i
    - \( \gamma_{ij} \) = Effect of interaction between factors A and B
    - \( e_{ijk} \) = Experimental error