# STAT 850 Notes

BY LANGTIAN MA

## 1 Randomized Block Design

**Problem Setting:**

- One block factor with $b$ levels. and one treatment factor with $t$ levels.
- Treatments randomized with blocks.
- No replicate observations. Our observed data $y_{ij}$ represents the data in $i$th block with $j$th factor.

| | Treatment | | | | |
|---|---|---|---|---|---|
| | A | B | C | D | Mean |
| Block 1 | 89 | 88 | 97 | 94 | 92 |
| Block 2 | 84 | 77 | 92 | 79 | 83 |
| Block 3 | 81 | 87 | 87 | 85 | 85 |
| Block 4 | 87 | 92 | 89 | 84 | 88 |
| Block 5 | 79 | 81 | 80 | 88 | 82 |
| Mean | 84 | 85 | 89 | 86 | 86 |

**Table 1.**

**Model without block-treatment interactions:**

$$y_{ij} = \mu + b_i + t_j + \epsilon_{ij} \quad i \in 1\colon B, j = 1\colon T$$

with $\epsilon_{ij} \overset{\text{iid}}{\sim} N(0, \sigma^2)$ and $\sum_j t_j = 0$. $b_i$s and $t_j$s indicate block and factor effects.

There are two settings for the block effects in the model:

- **Fixed block effects:** $\sum_{i=1}^N b_i = 0$ (sum to zero constraint for block effects)
- **Random block effects:** $b_1, \ldots, b_B \overset{\text{iid}}{\sim} N(0, \sigma_B^2)$

**Data decomposition:** the observed data can be decomposed to a summation of the (estimated) effects:

$$\begin{aligned} y_{ij} &= y_{\cdot\cdot} + (y_{i\cdot} - y_{\cdot\cdot}) + (y_{\cdot j} - y_{\cdot\cdot}) + (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot}) \\ &= \hat{\mu} + \hat{b}_i + \hat{t}_j + \hat{\epsilon}_{ij} \end{aligned}$$

**ANOVA decomposition:** the total sum of squares can be decomposed to:

$$\underbrace{\sum_i \sum_j (y_{ij} - y_{\cdot\cdot})^2}_{\text{Total sum of squares}} = \underbrace{t \sum_i (y_i - y_{\cdot\cdot})^2}_{\text{Blocks}} + \underbrace{b \sum_j (y_{\cdot j} - y_{\cdot\cdot})^2}_{\text{Treatments}} + \underbrace{\sum_i \sum_j (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot})^2}_{\text{Erorrs}}$$

| Source | SS | df | $\mathbb{E}(\mathrm{MS} = \mathrm{SS}/\mathrm{df})$ |
|---|---|---|---|
| | Fixed block effects | | |
| Blocks | $t\sum_i (y_{i\cdot} - y_{\cdot\cdot})^2$ | $b-1$ | $\sigma^2 + t\,(b-1)^{-1}\sum_i b_i^2$ |
| Treatments | $b\sum_j (y_{\cdot j} - y_{\cdot\cdot})^2$ | $t-1$ | $\sigma^2 + b\,(t-1)^{-1}\sum_j t_j^2$ |
| Error | $\sum_{ij} (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot})^2$ | $(b-1)\,(t-1)$ | $\sigma^2$ |
| Total | $\sum_{ij} (y_{ij} - y_{\cdot\cdot})^2$ | $b\,t-1$ | |
| | Random block effects | | |
| Blocks | $t\sum_i (y_{i\cdot} - y_{\cdot\cdot})^2$ | $b-1$ | $\sigma^2 + t\,\sigma_b^2$ |
| Treatments | $b\sum_j (y_{\cdot j} - y_{\cdot\cdot})^2$ | $t-1$ | $\sigma^2 + b\,(t-1)^{-1}\sum_j t_j^2$ |
| Error | $\sum_{ij} (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot})^2$ | $(b-1)\,(t-1)$ | $\sigma^2$ |
| Total | $\sum_{ij} (y_{ij} - y_{\cdot\cdot})^2$ | $b\,t-1$ | |

**Table 2.** ANOVA table for Randomized Block Design

# 2 Factorial treatment structure

## 2.1 Model Setting

Consider an experiment with two factors $P$ and $Q$ (***P* and *Q* may have interactions**) with levels $j = 1, 2, \ldots, p$, and $k = 1, 2, \ldots, q$, replicated $r$ times ($l = 1, 2, \ldots, r$), with model

$$y_{jkl} = \mu_{jk} + \epsilon_{jkl}$$

**Group effect parameters:**

- Grand mean: $\mu_{\cdot\cdot} = (pq)^{-1}\sum_j \sum_k \mu_{jk}$
- Group means for factor $P$: $\mu_{j\cdot} = q^{-1}\sum_k \mu_{jk}$
- Group means for factor $Q$: $\mu_{\cdot k} = p^{-1}\sum_j \mu_{jk}$
- Effect of factor $P$: $p_j = \mu_{j\cdot} - \mu_{\cdot\cdot}$
- Effect of factor $Q$: $q_k = \mu_{\cdot k} - \mu_{\cdot\cdot}$

We have sum to zero constrains under this setting:

$$\sum_j p_j = 0, \quad \sum_k q_k = 0$$

**Interaction effect parameters:**

$$(pq)_{jk} = \mu_{jk} - (\mu_{\cdot\cdot} + p_j + q_k) = (\mu_{jk} - \mu_{\cdot k}) - (\mu_{j\cdot} - \mu_{\cdot\cdot})$$

Also, we have

$$\sum_j (pq)_{jk} = 0 \quad \text{for all } k, \quad \sum_k (pq)_{jk} = 0 \quad \text{for all } j$$

Then the model can be expand as:

$$\mu_{jk} = \mu_{\cdot\cdot} + p_j + q_k + (pq)_{jk}. \tag{1}$$

**Remark 1.** $p$ and $q$ without subscripts denote the number of levels for fator $P$ and $Q$, while $p_i$ and $q_j$ denote the effect parameter for each level.

**Example 1.** (4 × 4 Design)

| | $\mu_{jk}$ | | | | $\mu_{j\cdot}$ | $p_j$ | $(pq)_{jk}$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 4 | 10 | 20 | 30 | 16 | -6 | -3 | -4 | 1 | 6 |
| | 14 | 20 | 40 | 50 | 31 | 9 | -8 | -9 | 6 | 11 |
| | 14 | 20 | 10 | 20 | 16 | -6 | 7 | 6 | -9 | -4 |
| | 20 | 30 | 30 | 20 | 25 | 3 | 4 | 7 | 2 | -13 |
| $\mu_{\cdot k}$ | 13 | 20 | 25 | 30 | 22 | | | | | |
| $q_k$ | -9 | -2 | 3 | 8 | | | | | | |

**Table 3.** 4 × 4 design data table

## 2.2 Compare differences between treatments

To compare differences between treatments, we define **contrast** and **interaction contrast**:

**Definition 1.** *A contrast for the main effects of factor $P$ is defined as*

$$C_P = \sum_{j=1}^{p} l_j \mu_{j\cdot},$$

*where $l_1, \ldots, l_p$ are coefficients with $\sum_{j=1}^{p} l_j = 0$.*

**Example 2.** (Simple Pairwise Comparison)

$$C_P = \mu_{1\cdot} - \mu_{2\cdot}$$

**Definition 2.** *An interaction contrast is defined as:*

$$C_{PQ} = \sum_{j=1}^{p} \sum_{k=1}^{q} l_j m_k \mu_{jk},$$

*where $m_1, \ldots, m_q$ are also coefficients with $\sum_{k=1}^{q} m_k = 0$.*

**Example 3.** Test whether the difference between levels of $P$ depends on trhe level of $Q$.

$$C_{\text{PQ}} = (\mu_{11} - \mu_{12}) - (\mu_{21} - \mu_{22})$$

**Interpretation of main and interaction effects**:

1. **Always start by checking main effects.** Interactions modify these effects and only make sense in that context.

2. **If interactions are negligible,** simplify the interpretation and focus on main effects**.**

3. **If 3 or higher order interactions are negligible**, but second-order interactions are significant, then we should focus on both main effects and second-order interactions.

4. If a two-factor interaction is **very important**, and its **mean square (MS) value is similar to the MS values for main effects**, then the best way to interpret results is by **looking at the mean values for two-factor combinations** rather than just reporting main effects.

5. If ar two-factor interaction is significant, but **one or both main effects are much larger than the interaction**, then the interpretation should consider main effects first, with adjustments for interaction effects.

## 2.3 Least-squares estimation for an unreplicated 2×3 design

Consider a two-way factorial design with $p=2$ and $q=3$

$$\mu_{jk}=\mu_{..}+p_j+q_k+(p\,q)_{jk}$$

| | 1 | 2 | 3 |
|---|---|---|---|
| 1 | $y_{11}$ | $y_{12}$ | $y_{13}$ |
| 2 | $y_{21}$ | $y_{22}$ | $y_{23}$ |

**Table 4.**

$$\mathbf{y}=\begin{pmatrix} y_{11}\\ y_{12}\\ y_{13}\\ y_{21}\\ y_{22}\\ y_{23}\end{pmatrix}=\mathbf{X}\beta=\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0\\ 1 & 1 & 0 & 1 & 0 & 1\\ 1 & 1 & -1 & -1 & -1 & -1\\ 1 & -1 & 1 & 0 & -1 & 0\\ 1 & -1 & 0 & 1 & 0 & -1\\ 1 & -1 & -1 & -1 & 1 & 1\end{pmatrix}\begin{pmatrix} \mu_{..}\\ p_1\\ q_1\\ q_2\\ (p\,q)_{11}\\ (p\,q)_{12}\end{pmatrix}+\epsilon$$

The columns of $\mathbf{X}$ w.r.t. different parameter groups are orthognal. In this case, the columns with respect to $p$ and columns with respect to $q$ are orthogonal. Also, they are orthogonal to the columns w.r.t. $pq$. Therefore, $\mathbf{X}^T\mathbf{X}$ appears to be block diagonal:

$$\mathbf{X}'\mathbf{X}=\begin{pmatrix} 6 & 0 & 0 & 0 & 0 & 0\\ 0 & 6 & 0 & 0 & 0 & 0\\ 0 & 0 & 4 & 2 & 0 & 0\\ 0 & 0 & 2 & 4 & 0 & 0\\ 0 & 0 & 0 & 0 & 4 & 2\\ 0 & 0 & 0 & 0 & 2 & 4\end{pmatrix}$$

$$(\mathbf{X}'\mathbf{X})^{-1}=\begin{pmatrix} 1/6 & 0 & 0 & 0 & 0 & 0\\ 0 & 1/6 & 0 & 0 & 0 & 0\\ 0 & 0 & 1/3 & -1/6 & 0 & 0\\ 0 & 0 & -1/6 & 1/3 & 0 & 0\\ 0 & 0 & 0 & 0 & 1/3 & -1/6\\ 0 & 0 & 0 & 0 & -1/6 & 1/3\end{pmatrix}$$

The least square solution gives

$$\hat{\beta}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}=\begin{pmatrix} y_{..}\\ y_{1\cdot}-y_{..}\\ y_{\cdot 1}-y_{..}\\ y_{\cdot 2}-y_{..}\\ y_{11}-y_{1\cdot}-y_{\cdot 1}+y_{..}\\ y_{12}-y_{1.}-y_{\cdot 2}+y_{..}\end{pmatrix}.$$

Note that the number of parameter equals to the number of observations, $\mathbf{X}$ is invertable, $\hat{\beta}$ is the solutoin to $\mathbf{X}\beta=\mathbf{y}$.

Now we consider an additive model without interactions:

$$\mu_{jk} = \mu_{..} + p_j + q_k,$$

the least square solution gives:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}\mathbf{y} = (\ y_{..} \quad y_{1\cdot} - y_{..} \quad y_{\cdot 1} - y_{..} \quad y_{\cdot 2} - y_{..}\ )^T,$$

which align with the solution of the model with interactions. This is due to the orthogonality.

### 2.4 Experiment with Replication

The model for an experiment with replication can be written as:

$$y_{jkl} = \mu + p_j + q_k + (pq)_{jk} + \epsilon_{jkl}, j = 1, \ldots, p; k = 1, \ldots q, l = 1, \ldots r,$$

where $\epsilon_{jkl}$ are i.i.d. $\mathcal{N}(0, \sigma^2)$. We can decompose data as:

$$\begin{aligned} y_{jkl} &= \hat{\mu} + \hat{p}_j + \hat{q}_k + \widehat{(p\,q)}_{jk} + \hat{\epsilon}_{jkl} \\ &= y_{...} + (y_{j..} - y_{...}) + (y_{\cdot k \cdot} - y_{...}) + (y_{jk\cdot} - y_{j..} - y_{\cdot k \cdot} + y_{...}) + (y_{jkl} - y_{jk\cdot}). \end{aligned}$$

The sum of squares are defined as:

$$\begin{aligned} S_P &= q\,r \sum_j (y_{j..} - y_{...})^2 \\ S_Q &= p\,r \sum_k (y_{\cdot k \cdot} - y_{...})^2 \\ S_{PQ} &= r \sum_j \sum_k (y_{jk\cdot} - y_{j..} - y_{\cdot k \cdot} + y_{...})^2 \\ S_R &= \sum_j \sum_k \sum_l (y_{jkl} - y_{jk\cdot})^2 \\ S_D &= \sum_j \sum_k \sum_l (y_{jkl} - y_{...})^2 \end{aligned}$$

| Source | SS | df | MS | Ratio |
|---|---|---|---|---|
| Factor P | $S_P = 1.03301$ | $p - 1 = 2$ | $s_P^2 = 0.51651$ | $s_P^2 / s_R^2 = 23.22$ |
| Factor Q | $S_B = 0.92121$ | $q - 1 = 3$ | $s_Q^2 = 0.30707$ | $s_Q^2 / s_R^2 = 13.81$ |
| Interaction | $S_{PQ} = 0.25014$ | $(p-1)\,(q-1) = 6$ | $s_{PQ}^2 = 0.04169$ | $s_{PQ}^2 / s_R^2 = 1.87$ |
| Residual | $S_R = 0.80073$ | $p\,q\,(r-1) = 36$ | $s_R^2 = 0.02224$ | |
| Total | $S_D = 3.00508$ | $p\,q\,r - 1 = 47$ | | |

**Table 5.** ANOVA table for two factors experiments with replications

### 2.5 Model Checking

1. Define the estimated value of $\mu_{\mathrm{jk}}$ under the full model as $\hat{\mu}_{\mathrm{jk}} = y_{\mathrm{jk}\cdot}$. Let $\tilde{\mu}_{jk} = y_{j..} + y_{\cdot k \cdot} - y_{...}$ denote the estimated value of $\mu_{jk}$ assuming no inetractions.

2. To assess the homogeneity of variance in interactions, create a plot of the residuals $y_{jkl} - \hat{\mu}_{\mathrm{jk}}$ against the fitted values $\hat{\mu}_{\mathrm{jk}}$. A consistent spread of residuals across different values of $\hat{\mu}_{\mathrm{jk}}$ suggests homogeneity, while a pattern or funnel shape may indicate variance issues.

3. To detect possible nonadditivity, plot $y_{jk} - \tilde{\mu}_{jk}$ against $\tilde{\mu}_{jk}$. If the plot exhibits a curvilinear pattern, this suggests the presence of transformable nonadditivity, meaning that a transformation of the response variable may be necessary for a better model fit.

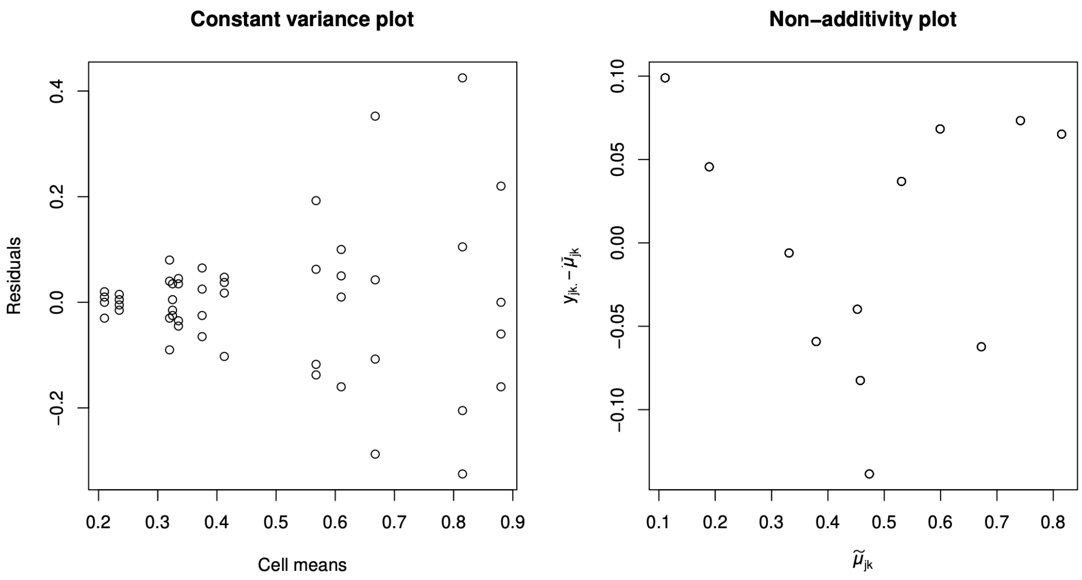


**Figure 1.** Model checking plots suggesting heteroscedasticity and non-additivity

## 2.6 Transformations

### 2.6.1 Taylor Power Transformation

Consider heterogeneous data $\mathrm{Var}(y_{jkl}) = \sigma_{jk}^2$ and assume that $\sigma_{jk}^2 \propto \mu_{jk}^\beta$ for some $\beta$. We can use **Taylor power transformation** to deal with the heteroscedasticity (Assume $y_{jk\cdot} > 0$).

1. Calculate the sample mean $y_{jk\cdot}$ and sample standard deviation $s_{jk}$ for each $(j, k)$ cell.

2. Fit a linear regression on $\log s_{jk} \sim \log y_{jk\cdot}$

3. The fitted slope $\hat{\beta}$ is an estimate of $\beta$.

4. Use the transformation $(y^\lambda - 1)/\lambda$ with $\lambda := 1 - \hat{\beta}$.

5. If $\lambda = 0$, apply the log transformation $y_{jkl}^* = \log(y_{jkl})$.

**Remark 2.** Simple power transformation uses $y^\lambda$, which doesn't smoothly transit to $\log y$ as $\lambda \to 0$.

**Justification for the method**:

Define

$$z = f_\lambda(y) = \begin{cases} (y^\lambda - 1)/\lambda, & \lambda \neq 0 \\ \log y, & \lambda = 0 \end{cases} \tag{2}$$

By Taylor expansion:

$$z_{jkl} = f_\lambda(y_{jkl}) \approx f_\lambda(\mu_{jk}) + f'_\lambda(\mu_{jk})\,(y_{jkl} - \mu_{jk}).$$

Since $f'(y) = y^{\lambda-1}$,

$$\begin{aligned}\mathrm{Var}(z_{jkl}) &\approx (f'_\lambda(\mu_{jk}))^2 \mathrm{Var}(y_{jkl}) \\ &= \mu_{jk}^{2(\lambda-1)}\,\sigma_{jk}^2 \\ &\propto \mu_{jk}^{2(\lambda-1)}\,\mu_{jk}^{2\beta} \\ &= \mu_{jk}^{2(\lambda-1+\beta)},\end{aligned}$$

and $\mathrm{Var}(z_{jkl})$ becomes a constant if $\lambda = 1 - \beta$.

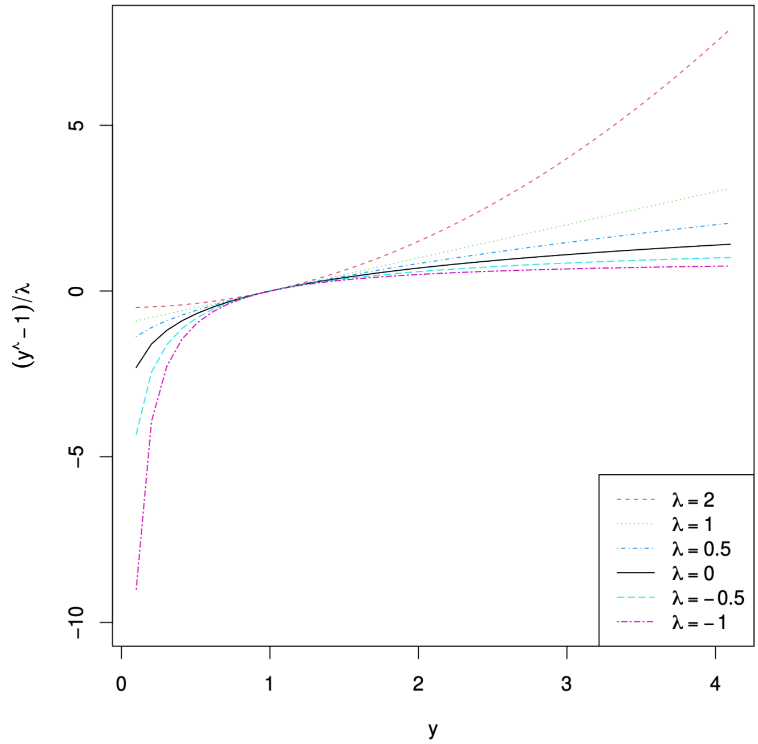


**Figure 2.** Power transformations

### 2.6.2 Box-Cox Transformation

**Assumptions:** There exists a $\lambda$ such that $\{f_\lambda(y_i)\}_{i=1}^n$:

- are mutually independent
- are normally distributed
- have constant variance
- satisfy a linear model $f_\lambda(\mathbf{y}) = \mathbf{X}\beta + \boldsymbol{\epsilon}$

### 2.6.3 Maximum Likelihood Estimation of Box-Cox $\boldsymbol{\lambda}$

Assume $\mathbf{y}^{(\lambda)} = f_\lambda(\mathbf{y}) = \mathbf{X}\beta + \boldsymbol{\epsilon}$ where $f_\lambda$ is defined in (2) with $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2\mathbf{I})$ for some $\lambda$. Then the likelihood function for the untransformed data follows:

$$L(\lambda, \beta, \sigma) = \frac{1}{(2\,\pi\,\sigma^2)^{n/2}} \exp\left\{-\frac{(\mathbf{y}^{(\lambda)} - \mathbf{X}\boldsymbol{\beta})^T\,(\mathbf{y}^{(\lambda)} - \mathbf{X}\boldsymbol{\beta})}{2\,\sigma^2}\right\} J(\lambda, \mathbf{y})$$

with Jacobian

$$J(\lambda, y) = \prod_{i=1}^{n} y_i^{\lambda-1}$$

We first find the LSEs of $\beta$ and $\sigma_\lambda$ for fixed $\lambda$:

$$\hat{\beta}_\lambda = (X^T X)^{-1} X^T \mathbf{y}^{(\lambda)}, \quad \hat{\sigma}_\lambda = \frac{(\mathbf{y}^{(\lambda)} - \mathbf{X}\hat{\boldsymbol{\beta}}_\lambda)^T (\mathbf{y}^{(\lambda)} - \mathbf{X}\hat{\boldsymbol{\beta}}_\lambda)}{n-p-1},$$

then we have

$$\begin{aligned}\ell(\lambda, \hat{\boldsymbol{\beta}}_\lambda, \hat{\sigma}_\lambda) &= \frac{\exp(-(n-p-1)/2)}{(2\pi)^{n/2}\, \hat{\sigma}_\lambda^n} J(\lambda, \mathbf{y}) \\ &= \hat{\sigma}_\lambda^{-n} \prod_{i=1}^{n} y_i^{\lambda-1} \frac{\exp(-(n-p-1)/2)}{(2\pi)^{n/2}}\end{aligned}$$

Let $\ell(\lambda) = \log L(\lambda, \hat{\beta}_\lambda, \hat{\sigma}_\lambda)$, and find the MLE $\hat{\lambda}$ that maximizes $\ell(\lambda)$.

The **confidence interval** follows from the standard result that the log-likelihood ratio statistic follows a chi-square distribution with 1 degree of freedom

$$2(\ell(\hat{\lambda}) - \ell(\lambda)) \sim \chi_1^2.$$

Then the confidence interval is the root for $\ell(\lambda) = \ell(\hat{\lambda}) - 0.5\chi_{1,\alpha}^2$.

**Remark 3.** The `boxcox` function in `MASS` library gives the MLE of $\lambda$.

```
library(MASS)
bc <- boxcox(y ~ p+q)
title(paste("Without interaction, lambda =",round(bc$x[which.max(bc$y)],2)))
bc <- boxcox(y ~ p*q)
title(paste("With interaction, lambda =",round(bc$x[which.max(bc$y)],2)))
```

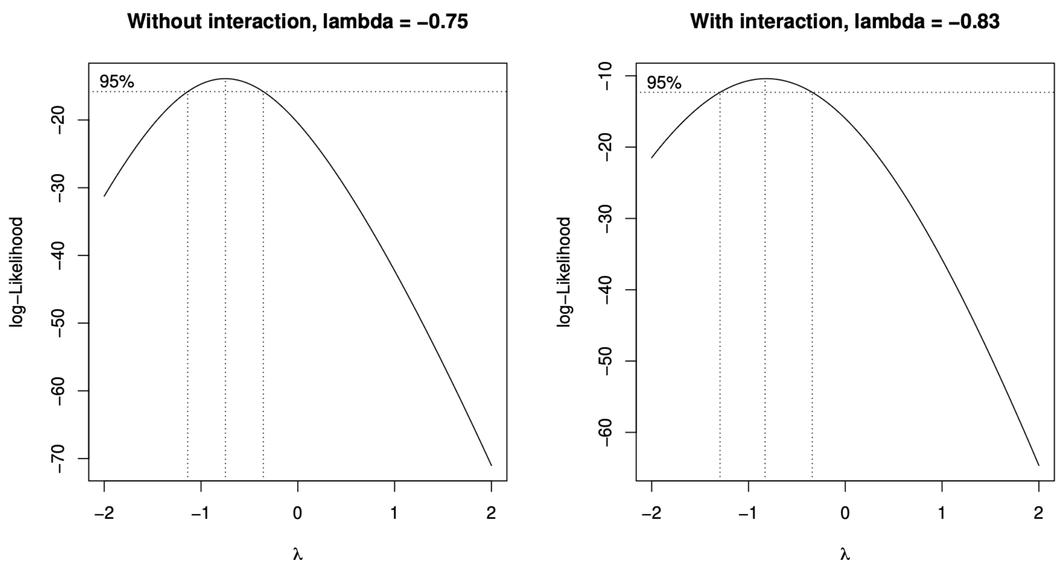


**Figure 3.** Likelihood of Box-Cox Transformation

**Example 4.** For the Poison data, we present the ANOVA table after the two transformations.

| | Df | Sum Sq | Mean Sq | F value | P-value |
|---|---|---|---|---|---|
| Poison | 2 | 34.877 | 17.4386 | 70.6302 | 5.17e−13 |
| Treatment | 3 | 20.414 | 6.8048 | 27.5610 | 2.48e−09 |
| Interaction | 6 | 1.571 | 0.2618 | 1.0603 | 0.4046 |
| Residuals | 35 | 8.643 | 0.2469 | | |
| Poison | 2 | 11.926 | 5.9631 | 66.5525 | 1.18e−12 |
| Treatment | 3 | 7.158 | 2.3860 | 26.6295 | 3.76e−09 |
| Interaction | 6 | 0.486 | 0.0810 | 0.9040 | 0.5032 |
| Residuals | 35 | 3.136 | 0.0896 | | |

**Table 6.** ANOVA table after simple Taylor power transformation ($y^{(\lambda)} = y^{\lambda}$ with $\lambda = -1$) (above) and Box-Cox transformation (below).

**Remark 4.** The residual Df is reduced by 1 to compensate for the estimation of $\lambda$.

## 2.7 Confidence Intervals

### 2.7.1 When interactions are not significant

Let $u_j = \mu + p_j$ and $v_k = \mu + q_k$. Let the unbiased estimators be $\hat{u}_j = y_{j\cdot\cdot}$ and $\hat{v}_k = y_{\cdot k\cdot}$ with $\mathrm{Var}(\hat{u}_j) = \sigma^2/(pr)$, $\mathrm{Var}(\hat{v}_j) = \sigma^2/(qr)$. The sample standard deviation is

$$s(\hat{u}_j) = S_R/\sqrt{pr}, s(\hat{v}_k) = S_R/\sqrt{qr}.$$

Then $100(1-\alpha)\%$ confidence interval for $u_j$ is $\hat{u}_j \pm t_{\nu_R;\alpha/2} \times \mathrm{se}(\hat{u}_j)$, where $\nu_R = pq(r-1)$ is the degree of freedom for $S_R$.

For simultaneous confidence intervals, we define the contrast of interest be

$$L = \sum_{j=1}^{p} c_j u_j, \quad \text{where} \sum_{j=1}^{p} c_j = 0$$

with its estinator

$$\hat{L} = \sum_{j=1}^{p} c_j y_{j\cdot\cdot} \quad \text{with} \quad s(\hat{L}) = S_R \sqrt{(qr)^{-1} \sum_{j=1}^{p} c_j^2}$$

Then a $100(1-\alpha)\%$ simultaneous confidence interval for $L$ takes the form:

$$\hat{L} \pm T s(\hat{L}),$$

where $T$ is a multiplier that depends on the type of the inference method used.

1. **Tukey's Method (Pairwise Comparisons):** Tukey's method is designed for simultaneous confidence intervals when comparing all possible pairwise differences between group means. The multiplier is:

   $$T = \frac{q(p, \nu_R; \alpha)}{\sqrt{2}}$$

   where $q(p, \nu_R; \alpha)$ is the studentized range statistic for $p$ groups and residual degrees of freedom $\nu_R$.

When comparing all pairs, each pairwise difference is a contrast (with coefficients $c_j = 1$ for one group and $c_j = -1$ for the other, and 0 elsewhere). Each interval is given by:

$$\text{CI for } (u_j - u_k): \quad (y_{j\cdot\cdot} - y_k) \pm \frac{q\,(p, \nu_R; \alpha)}{\sqrt{2}}\, s_R \sqrt{\frac{1^2 + (-1)^2}{q\,r}}$$

Here, the multiplier $\frac{q\,(p, \nu_R;\, \alpha)}{\sqrt{2}}$ is applied to each pairwise contrast.

2. **Scheffé's Method (All Contrasts):** Scheffés method is more conservative and applies to all possible contrasts, not just pairwise comparisons. The multiplier is:

$$T = \sqrt{(p-1)\, F_{(p-1), \nu_R; \alpha}}$$

where $F_{(p-1), \nu_R; \alpha}$ is the critical value from the F-distribution with $p - 1$ and $\nu_R$ degrees of freedom.

For any contrast $L = \sum_j c_j\, u_j$, the simultaneous confidence interval is:

$$\hat{L} \pm \sqrt{(p-1)\, F_{(p-1), \nu_R; \alpha}}\, S_R \sqrt{\frac{1}{q\,r} \sum_j c_j^2}$$

This interval applies to every possible contrast you might form.

3. **Bonferroni's Method (For $g$ Comparisons):** The Bonferroni method controls the familywise error rate by adjusting the significance level for multiple comparisons. For any set of $g$ comparisons, the multiplier is:

$$T = t_{\nu_R; \alpha/(2\,g)}$$

where $t_{\nu_R; \alpha/(2\,g)}$ is the t-distribution critical value with residual degrees of freedom $\nu_R$ and a Bonferroni-adjusted significance level of $\alpha / (2\,g)$.

If you have a specific set of $g$ comparisons (contrasts) you plan to test, each interval is:

$$\hat{L} \pm t_{\nu_R; \alpha/(2\,g)}\, S_R \sqrt{\frac{1}{q\,r} \sum_j c_j^2}.$$

Each of the $g$ contrasts gets its own interval, with the critical value adjusted by dividing $\alpha$ by $2\,g$.

### 2.7.2 When interatcions are significant

In this case, each combination of factor levels has its own mean:

$$\mu_{jk} = \mu + p_j + q_k + (p\,q)_{jk}$$

If we are interested in comparing the means of two specific treatment combinations, say $\mu_{j_1 k_1}$ and $\mu_{j_2 k_2}$, we are comparing two of the $p\,q$ treatments.

1. **Tukey's Method:** for all possible pairs of treatment means,

$$(\hat{\mu}_{j_1 k_1} - \hat{\mu}_{j_2 k_2}) \pm \frac{q(p\,q, \nu_R; \alpha)}{\sqrt{2}}\, S_R \sqrt{\frac{2}{r}}$$

where $q(pq, \nu_R; \alpha)$ is the quantile of the studentized range statistic for $pq$ treatments and $\nu_R$ is the residual degrees of freedom.

2. **Bonferroni's Method:** for $g$ pairs of comprisons,

$$(\hat{\mu}_{j_1 k_1} - \hat{\mu}_{j_2 k_2}) \pm t_{\nu_R; \alpha/(2g)} \, s_R \sqrt{\frac{2}{r}},$$

where $t_{\nu_R; \alpha/(2g)}$ is the quantile from $t$-distirbution.

3. **Scheffé's Method for General Contrasts**: the contrast takes the form:

$$L = \sum_{j=1}^{p} \sum_{k=1}^{q} c_{jk} \, \mu_{jk}, \quad \text{with} \quad \sum_{j,k} c_{jk} = 0$$

The confidence interval for the contrast $L$ is given by:

$$\hat{L} \pm \sqrt{(p\,q - 1)\, F_{(pq-1), \nu_R; \alpha}} \, s(\hat{L})$$

where:

- $\hat{L} = \sum_{j,k} c_{jk} \hat{\mu}_{jk}$,
- $s(\hat{L}) = s_R \sqrt{\sum_{j,k} \frac{c_{jk}^2}{r}}$ (assuming balanced replication),
- $F_{(pq-1), \nu_R; \alpha}$ ris the critical value from the $F$-distribution with $pq - 1$ and $\nu_R$ degrees of freedom.

## 2.8 Two-way Factorial with Blocks