STAT 8020 R Session 9: Completely Randomized Designs

Whitney Huang, Clemson University

Contents

Completely Randomized Design (CRD)

In a Completely Randomized Design (CRD), experimental units are randomly assigned to treatments. The primary goal is to determine whether the treatment means differ significantly.

In this example, we compare four treatments, each with five observations.

Create the Dataset

We first enter the observations for each treatment group.

```r
# Treatment 1 observations
r1 <- c(9.8, 8.8, 8.4, 9.5, 9.2)

# Treatment 2 observations
r2 <- c(8.2, 6.9, 7.5, 7.1, 6.5)

# Treatment 3 observations
r3 <- c(6.8, 6.6, 5.9, 7.3, 7.2)

# Treatment 4 observations
r4 <- c(4.8, 5.2, 5.4, 5.9, 4.6)

# Combine all observations into a single response vector
times <- c(r1, r2, r3, r4)

# Create treatment labels
# each = 5 means each treatment appears 5 times
trt <- rep(1:4, each = 5)

# Create a data frame for analysis
dat <- data.frame(y = times, trt = as.factor(trt))

# Display the dataset
dat
```

```
##      y trt
## 1  9.8   1
## 2  8.8   1
## 3  8.4   1
## 4  9.5   1
## 5  9.2   1
## 6  8.2   2
## 7  6.9   2
## 8  7.5   2
## 9  7.1   2
## 10 6.5   2
## 11 6.8   3
## 12 6.6   3
## 13 5.9   3
## 14 7.3   3
## 15 7.2   3
## 16 4.8   4
## 17 5.2   4
## 18 5.4   4
## 19 5.9   4
## 20 4.6   4
```

Summary Statistics by Treatment

We compute the sample mean and sample variance for each treatment group.

```r
# Compute treatment means
(means <- tapply(dat$y, dat$trt, mean))
```

```
##    1    2    3    4
## 9.14 7.24 6.76 5.18
```

```r
# Compute treatment variances
(vars <- tapply(dat$y, dat$trt, var))
```

```
##     1     2     3     4
## 0.308 0.418 0.313 0.262
```

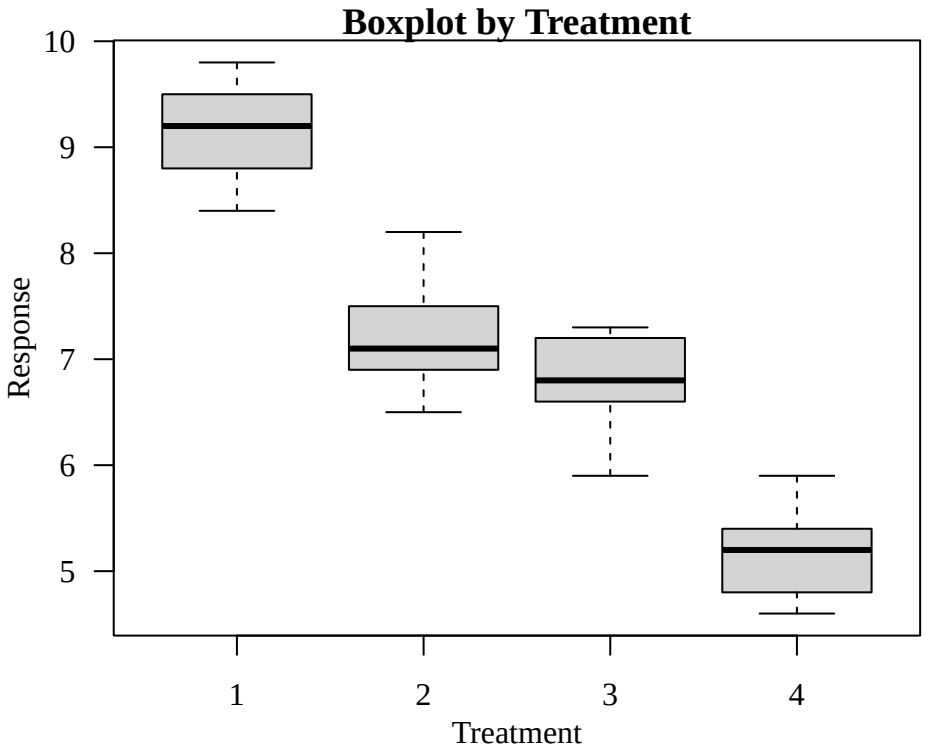
Interpretation

- The treatment means provide an initial comparison of average performance.
- The treatment variances describe variability within each treatment group.
- ANOVA will formally test whether the observed mean differences are statistically significant.

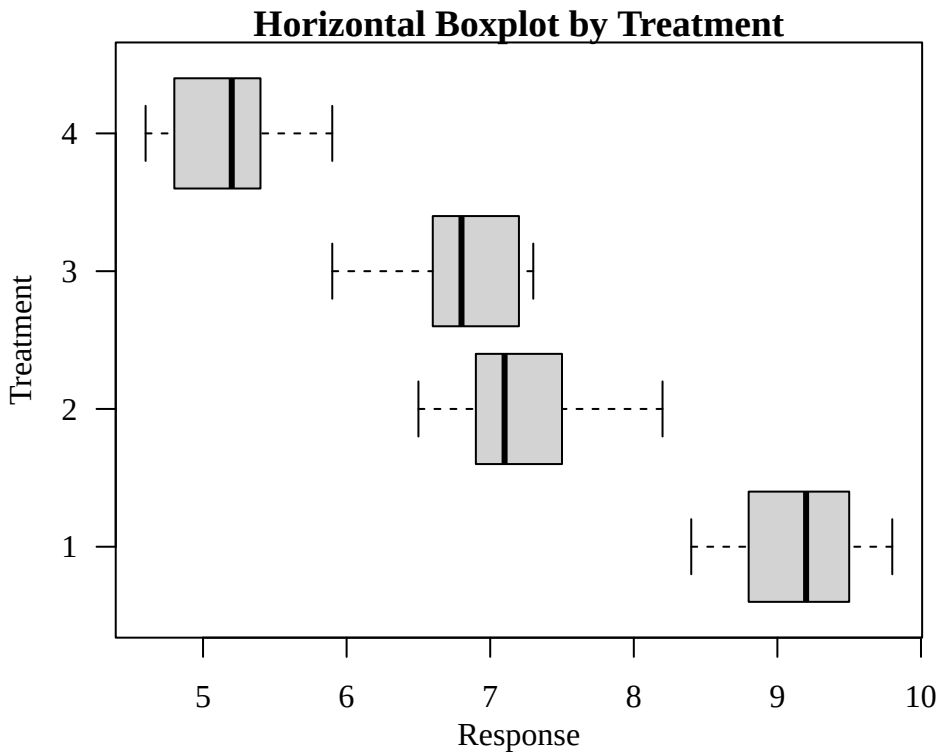
Visualize the Data

Boxplots provide a graphical comparison of treatment distributions.

```r
par(mar = c(3.5, 3.5, 1, 0.5), mgp = c(2, 1, 0), las = 1, family = "serif")
# Vertical boxplot
boxplot(y ~ trt, data = dat,
        main = "Boxplot by Treatment",
        xlab = "Treatment", ylab = "Response")
```



```r
# Horizontal boxplot
boxplot(y ~ trt, data = dat,
        las = 1, horizontal = TRUE,
        main = "Horizontal Boxplot by Treatment",
        xlab = "Response", ylab = "Treatment")
```



From the boxplots, we can visually assess:

- Differences in treatment centers (medians)
- Spread and variability within treatments
- Potential outliers
- Overall treatment patterns

The treatment groups appear to have different average responses, suggesting that treatment effects may exist.

Perform One-Way ANOVA

We now perform a one-way Analysis of Variance (ANOVA) to formally test whether the treatment means are equal.

Hypotheses

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_a : \text{At least one treatment mean differs}$$

```r
# Fit one-way ANOVA model
AOV <- aov(y ~ trt, data = dat)

# Display ANOVA table
summary(AOV)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## trt          3  39.91  13.303    40.9 9.92e-08 ***
## Residuals   16   5.20   0.325
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Multiple Comparisons

After finding a significant ANOVA result, we often want to determine **which treatment means differ**.

These procedures are called **multiple comparison methods**.

In this example, we consider:

- Fisher's Least Significant Difference (LSD) procedure with Bonferroni adjustment
- Tukey's Honest Significant Difference (HSD) procedure

Fisher's LSD Test with Bonferroni Adjustment The Bonferroni adjustment controls the family-wise error rate when conducting multiple pairwise tests.

```r
# Load agricolae package
library(agricolae)

# Perform Fisher's LSD test with Bonferroni adjustment
LSD_bon <- LSD.test(AOV, "trt", p.adj = "bonferroni")

# Display grouping results
LSD_bon$groups
```

```
##      y groups
## 1 9.14      a
## 2 7.24      b
## 3 6.76      b
## 4 5.18      c
```

Interpretation

Treatments sharing the same letter are **not significantly different**.

Treatments with different letters indicate statistically significant mean differences.

Tukey's Honest Significant Difference (HSD)

Tukey's HSD provides simultaneous confidence intervals for all pairwise comparisons.

```
# Perform Tukey HSD procedure
HSD <- TukeyHSD(AOV, conf.level = 0.95)

# Display pairwise comparison results
HSD$trt
```

```
##      diff       lwr       upr        p adj
## 2-1 -1.90 -2.931952 -0.868048 4.024593e-04
## 3-1 -2.38 -3.411952 -1.348048 3.310735e-05
## 4-1 -3.96 -4.991952 -2.928048 4.112087e-08
## 3-2 -0.48 -1.511952  0.551952 5.577630e-01
## 4-2 -2.06 -3.091952 -1.028048 1.708962e-04
## 4-3 -1.58 -2.611952 -0.548048 2.363679e-03
```

Interpretation

For each pairwise comparison:

- A confidence interval that does not contain 0 suggests a significant difference.
- The adjusted p-value accounts for multiple testing.

Tukey's method is generally more conservative than Fisher's LSD.

Model Assumptions

ANOVA relies on several important assumptions:

1. Independence of observations
2. Equal variance across treatments
3. Normally distributed errors

We now investigate these assumptions using the following balloon experiment.

Balloon Experiment *(Taken from Dean and Voss, Exercise 3.12)*

The experimenter (Meily Lin) observed that some balloon colors seemed harder to inflate than others.

The purpose of the experiment was to determine whether balloon colors differ in terms of the time required to inflate balloons to a diameter of 7 inches.

Four balloon colors were selected from the same manufacturer.

- Treatment 1 = Pink
- Treatment 2 = Yellow
- Treatment 3 = Orange
- Treatment 4 = Blue

An assistant inflated the balloons while the experimenter recorded inflation times using a stopwatch.

Table 3.13 Times (in seconds) for the balloon experiment

Time order	1	2	3	4	5	6	7	8
Coded color	1	3	1	4	3	2	2	2
Inflation time	22.0	24.6	20.3	19.8	24.3	22.2	28.5	25.7
Time order	9	10	11	12	13	14	15	16
Coded color	3	1	2	4	4	4	3	1
Inflation time	20.2	19.6	28.8	24.0	17.1	19.3	24.2	15.8
Time order	17	18	19	20	21	22	23	24
Coded color	2	1	4	3	1	4	4	2
Inflation time	18.3	17.5	18.7	22.9	16.3	14.0	16.6	18.1
Time order	25	26	27	28	29	30	31	32
Coded color	2	4	2	3	3	1	1	3
Inflation time	18.9	16.0	20.1	22.5	16.0	19.3	15.9	20.3

Figure 1: Source: Table 3.13 of Dean and Voss Exercise 3.12

Read the Data into R

We first import the dataset and examine its structure.

```
# Read the dataset
balloon <- read.csv("cr_assumptions.csv", header = TRUE)

# Display first few rows
head(balloon)
```

```
##   ORDER COLOR TIME
## 1     1     1 22.0
## 2     2     3 24.6
## 3     3     1 20.3
## 4     4     4 19.8
## 5     5     3 24.3
## 6     6     2 22.2
```

```
# Summary statistics
summary(balloon)
```

```
##      ORDER           COLOR           TIME
##  Min.   : 1.00   Min.   :1.00   Min.   :14.00
##  1st Qu.: 8.75   1st Qu.:1.75   1st Qu.:17.40
##  Median :16.50   Median :2.50   Median :19.70
##  Mean   :16.50   Mean   :2.50   Mean   :20.24
##  3rd Qu.:24.25   3rd Qu.:3.25   3rd Qu.:22.60
##  Max.   :32.00   Max.   :4.00   Max.   :28.80
```

```
# Display first 10 observations
head(balloon, 10)
```

```
##    ORDER COLOR TIME
```

```
## 1      1     1 22.0
## 2      2     3 24.6
## 3      3     1 20.3
## 4      4     4 19.8
## 5      5     3 24.3
## 6      6     2 22.2
## 7      7     2 28.5
## 8      8     2 25.7
## 9      9     3 20.2
## 10    10     1 19.6
```

Interpretation

The dataset contains:

- Inflation time (response variable)
- Balloon color (treatment variable)
- Observation order

Convert COLOR to a Factor Variable

Treatment variables should be stored as factors in ANOVA models.

```r
# Attach dataset
attach(balloon)

# Convert COLOR to factor
colorf <- as.factor(COLOR)

# Display factor levels
colorf
```

```
##  [1] 1 3 1 4 3 2 2 2 3 1 2 4 4 4 3 1 2 1 4 3 1 4 4 2 2 4 2 3 3 1 1 3
## Levels: 1 2 3 4
```

Interpretation

R now treats balloon color as a categorical treatment variable rather than a numeric variable.

Fit the ANOVA Model

We fit a one-way ANOVA model relating inflation time to balloon color.

```r
# Fit linear model
mod1 <- lm(TIME ~ colorf)

# Model summary
summary(mod1)
```

```
##
## Call:
## lm(formula = TIME ~ colorf)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.8750 -2.2500  0.0687  2.0531  6.2250
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   18.337      1.162  15.778 1.83e-15 ***
## colorf2        4.237      1.644   2.578   0.0155 *
## colorf3        3.538      1.644   2.152   0.0401 *
## colorf4       -0.150      1.644  -0.091   0.9279
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.287 on 28 degrees of freedom
## Multiple R-squared:  0.2967, Adjusted R-squared:  0.2214
## F-statistic: 3.938 on 3 and 28 DF,  p-value: 0.01836
```

```r
# ANOVA table
anova(mod1)
```

```
## Analysis of Variance Table
##
## Response: TIME
##           Df Sum Sq Mean Sq F value  Pr(>F)
## colorf     3 127.66  42.554  3.9379 0.01836 *
## Residuals 28 302.58  10.806
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interpretation

The ANOVA table tests whether the mean inflation times differ among balloon colors.

Obtain Residuals

Residuals are essential for diagnostic checking.

```r
# Raw residuals
r <- residuals(mod1)

# Standardized residuals
s <- rstandard(mod1)

# Variance of standardized residuals
var(s)
```

```
## [1] 1.032258
```

```r
# Studentized residuals
t <- rstudent(mod1)
```

Interpretation

- Raw residuals measure prediction errors
- Standardized residuals account for variability
- Studentized residuals are useful for detecting unusual observations

---

Assess Equal Variance

Equal variance (homoscedasticity) is an important ANOVA assumption.

We assess this assumption using:

- Levene's test
- Brown–Forsythe test

Levene's Test and Brown-Forsythe Test

```r
# Load package
library(lawstat)

# Levene's test using group means
levene.test(TIME, colorf, location = "mean")
```

```
##
##  Classical Levene's test based on the absolute deviations from the mean
##  ( none not applied because the location is not set to median )
##
## data:  TIME
## Test Statistic = 2.1682, p-value = 0.1141
```

```r
# Brown-Forsythe test using group medians
levene.test(TIME, colorf, location = "median")
```

```
##
##  Modified robust Brown-Forsythe Levene-type test based on the absolute
##  deviations from the median
##
## data:  TIME
## Test Statistic = 1.3975, p-value = 0.2642
```

Interpretation

Hypotheses:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$$

$$H_a : \text{At least one variance differs}$$

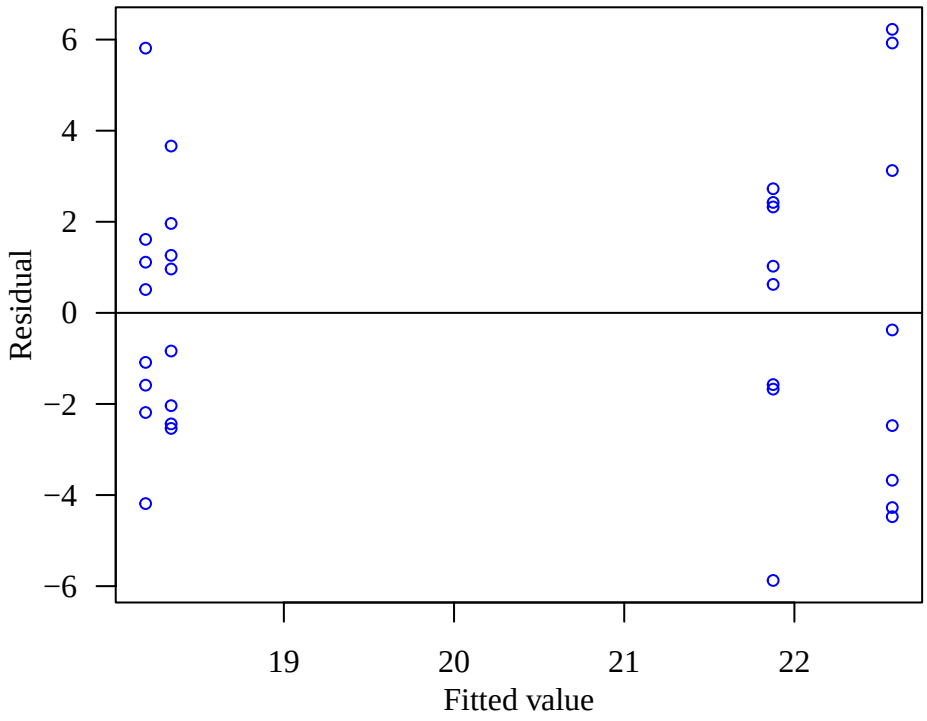
- A large p-value suggests equal variances are reasonable.
- A small p-value suggests unequal variances.

Residual Plots

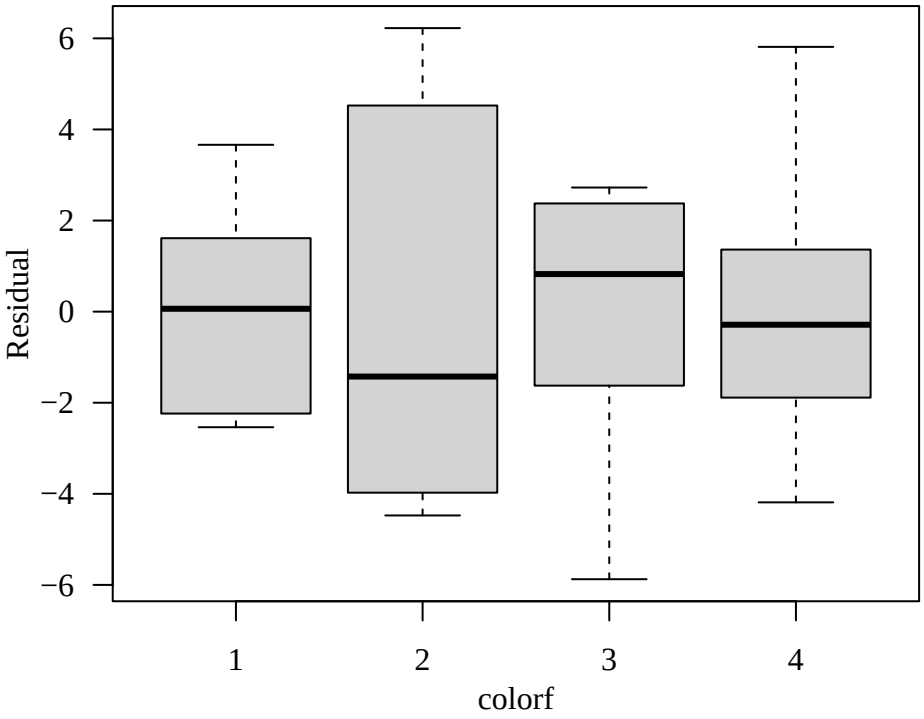
Residual plots help visually assess model assumptions.

```
# Residuals versus fitted values
par(mar = c(3.5, 3.5, 1, 0.5), mgp = c(2, 1, 0), las = 1, family = "serif")
plot(mod1$fitted, mod1$resid,
     xlab = "Fitted value", ylab = "Residual", cex = 0.75, col = "blue")

abline(h = 0)
```




```
# Residuals by treatment
plot(mod1$resid ~ colorf, ylab = "Residual", las = 1)
```



Interpretation

We look for:

- Constant spread across fitted values
- No obvious patterns
- Similar variability across treatments

A random scatter around zero supports model assumptions.

Assess Normality

ANOVA assumes residuals are approximately normally distributed.

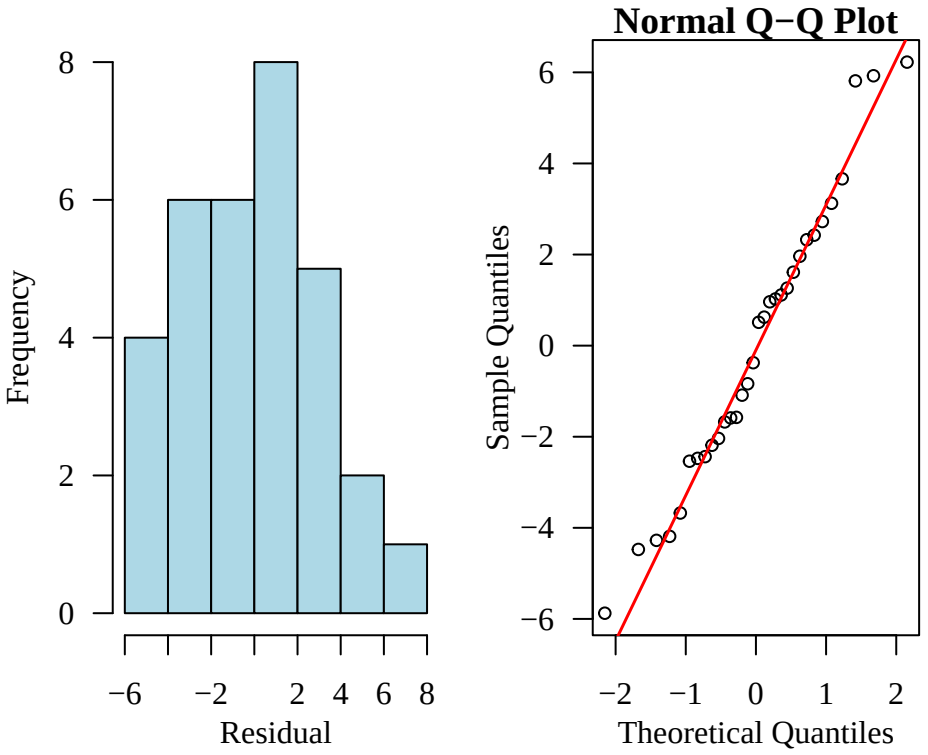
Histogram and QQ-Plot

```
# Set plotting layout
par(mfrow = c(1, 2), mar = c(3.5, 3.5, 1, 0.5), mgp = c(2, 1, 0),
    las = 1, family = "serif")

# Histogram of residuals
hist(mod1$resid, 8,
     main = "", xlab = "Residual", col = "lightblue")

# Normal QQ-plot
qqnorm(mod1$resid, cex = 0.8)

# Add reference line
qqline(mod1$resid, col = "red", lwd = 1.5)
```



Interpretation

- Histogram should appear approximately symmetric
- QQ-plot points should follow the reference line reasonably well

QQ-Plot by Treatment

We also examine normality separately within each treatment group.

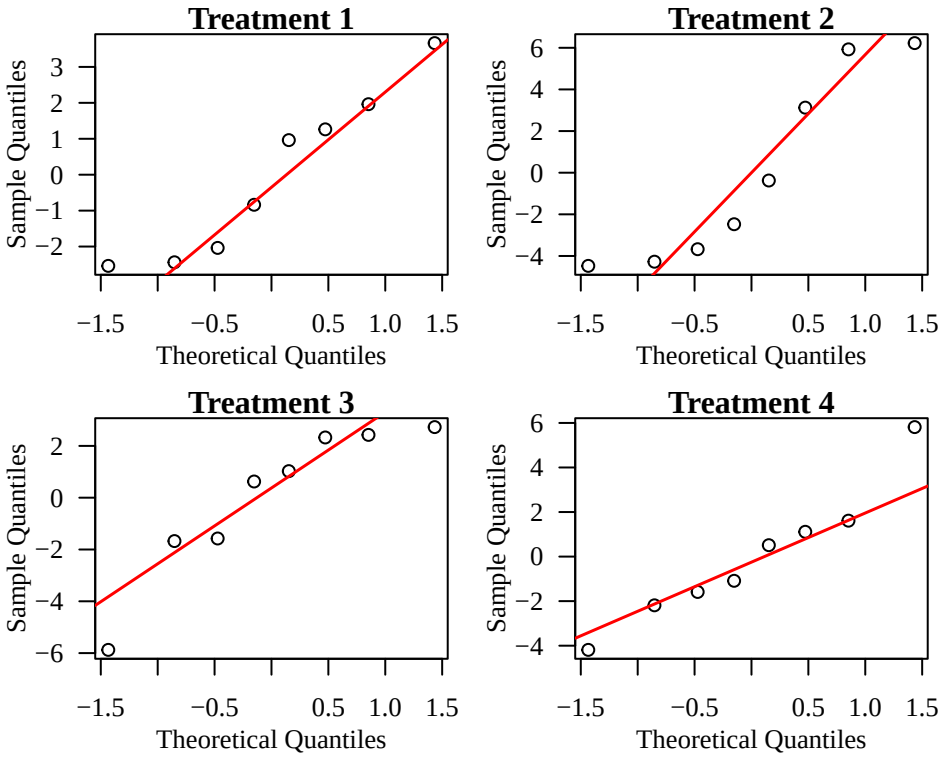
```
# Set plotting layout
par(mfrow = c(2, 2), mar = c(3.5, 3.5, 1, 0.5), mgp = c(2, 1, 0),
    las = 1, family = "serif")

# Create data frame of residuals
new <- data.frame(colorf, mod1$resid)

# Titles for plots
trt <- paste("Treatment", 1:4)

# Loop through treatments
for (i in 1:4){
  # Subset data for treatment i
  newc1 <- new[colorf == i,]
  # QQ-plot
  qqnorm(newc1$mod1.resid, main = trt[i])

  # Add reference line
  qqline(newc1$mod1.resid, col = "red", lwd = 1.5)
}
```



Interpretation

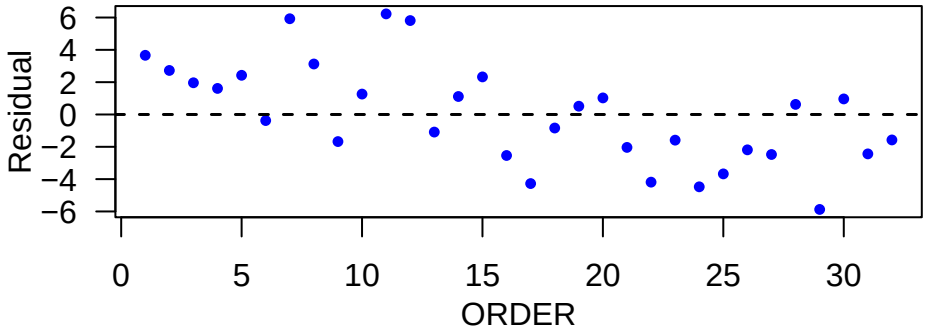
We assess whether residuals within each treatment approximately follow a normal distribution.

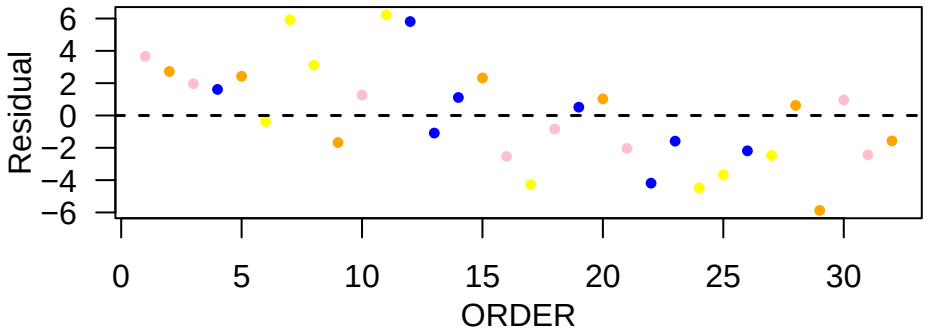
Assess Independence

Independence is often evaluated using residual plots against observation order.

Patterns over time may indicate serial correlation.

Residuals versus Observation Order





Interpretation

We look for:

- Random scatter around zero
- No trends or cycles
- No clustering patterns

Systematic patterns may suggest correlated errors.

Durbin-Watson Test

The Durbin-Watson test formally assesses autocorrelation.

```
# Load package
library(lmtest)

# Durbin-Watson test
dwtest(TIME ~ colorf, data = balloon)
```

```
##
##  Durbin-Watson test
##
## data:  TIME ~ colorf
## DW = 1.1617, p-value = 0.006005
## alternative hypothesis: true autocorrelation is greater than 0
```

Interpretation

Hypotheses:

$$H_0 : \text{No autocorrelation}$$

$$H_a : \text{Autocorrelation exists}$$

- A small p-value suggests correlated residuals.
- Values of the Durbin-Watson statistic near 2 indicate weak autocorrelation.

Fit a Model with Correlated Errors

If residuals are correlated, we may fit a model with AR(1) errors.

Generalized Least Squares Model

```
# Load package
library(nlme)

# Fit GLS model with AR(1) correlation structure
mod2 <- gls(TIME ~ colorf, correlation = corARMA(p = 1, q = 0))

# Display model summary
mod2
```

```
## Generalized least squares fit by REML
##   Model: TIME ~ colorf
##   Data: NULL
##   Log-restricted-likelihood: -74.42885
##
## Coefficients:
## (Intercept)     colorf2     colorf3     colorf4
##  18.5860865   3.7248742   3.4233901  -0.3578644
##
## Correlation Structure: AR(1)
##  Formula: ~1
##  Parameter estimate(s):
##       Phi
## 0.4285025
## Degrees of freedom: 32 total; 28 residual
## Residual standard error: 3.321057
```

Interpretation

This model allows neighboring observations to be correlated.

The AR(1) structure assumes:

$$\text{Corr}(\varepsilon_i, \varepsilon_{i+k}) = \rho^k$$

where ρ is the autocorrelation parameter.