

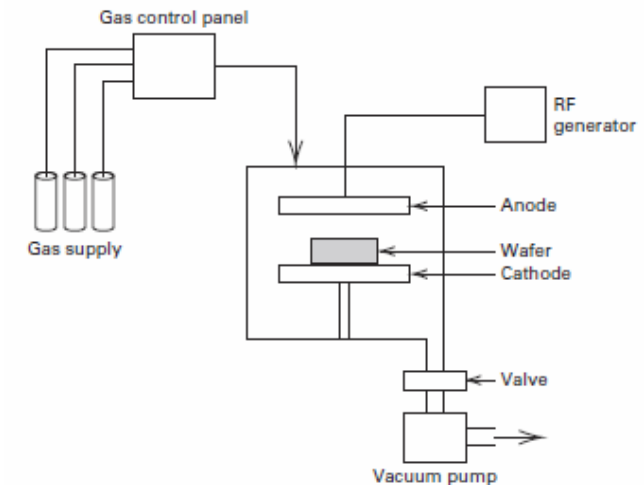
# Completely Randomized Experiments

Lin Wang

# An Example

An engineer is interested in investigating the relationship between the **RF power** setting and the etch rate for this tool. The objective of an experiment like this is to model the relationship between etch rate and RF power, and to specify the power setting that will give a desired target **etch rate**.

- X=RF power, Y=etch rate
- Other variables are fixed and not of our interest
- Design of the experiment:
  - **4 levels** of X: 160W, 180W, 200W, and 220W
  - **replicate 5 times** at each level of X
  - Runs made in random order



■ FIGURE 3.1 A single-wafer plasma etching tool

# Completely Randomized Design

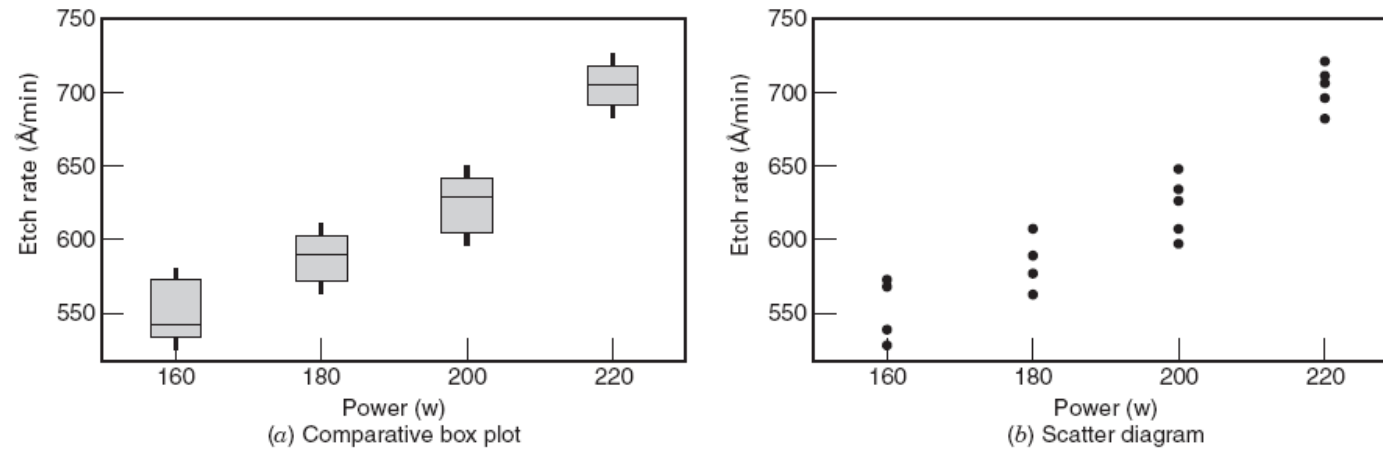
| Sequence | Power |
|----------|-------|
| 1        | 200   |
| 2        | 220   |
| 3        | 220   |
| 4        | 160   |
| 5        | 160   |
| 6        | 180   |
| 7        | 200   |
| 8        | 160   |
| 9        | 180   |
| 10       | 200   |
| 11       | 220   |
| 12       | 220   |
| 13       | 160   |
| 14       | 160   |
| 15       | 220   |
| 16       | 180   |
| 17       | 180   |
| 18       | 180   |
| 19       | 200   |
| 20       | 200   |

- This is an example of a single-factor experiment with  $a=4$  levels of the factor and  $n=5$  replicates. The 20 runs should be made in random order.
- In R, we can use the command “sample”.

```
> sample(20)
[1] 14 10  7  9 15  5 16 17 13 20
11  2 12  6  1  4  3 18  8 19
```

Let's arrange the data in a different way!

| <b><u>TABLE 3.1</u></b>   |                     |          |          |          |          |               |                 |
|---|---------------------|----------|----------|----------|----------|---------------|-----------------|
| <b>Etch Rate Data (in Å/min) from the Plasma Etching Experiment</b> |                     |          |          |          |          |               |                 |
|   | <b>Observations</b> |          |          |          |          |               |                 |
| <b>Power (W)</b>  | <b>1</b>            | <b>2</b> | <b>3</b> | <b>4</b> | <b>5</b> | <b>Totals</b> | <b>Averages</b> |
| 160   | 575                 | 542      | 530      | 539      | 570      | 2756          | 551.2           |
| 180   | 565                 | 593      | 590      | 579      | 610      | 2937          | 587.4           |
| 200   | 600                 | 651      | 610      | 637      | 629      | 3127          | 625.4           |
| 220   | 725                 | 700      | 715      | 685      | 710      | 3535          | 707.0           |



■ FIGURE 3.2 Box plots and scatter diagram of the etch rate data

- Does **changing** the power change the mean etch rate?
- Is there an **optimum** level for power?
- We would like to have an objective way to answer these questions
- The *t*-test really doesn't apply here – more than two factor levels

# The Problem

- We wish to test for differences between the mean etch rates at all  $a=4$  levels of RF power.
- Thus, we are interested in testing the equality of all four means.

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

$$H_a: \mu_i \neq \mu_j \text{ for at least one pair } (i, j)$$

- It might seem that this problem could be solved by performing a  $t$ -test for **all six possible pairs of means**. However, this is not the best solution to this problem.
  1. Performing all six pairwise  $t$ -tests is inefficient.
  2. Conducting all these pairwise comparisons inflates the type I error  $\alpha$

# Inflation of Type I error

## 1. What is Type I error?

A **Type I error** happens when we reject the null hypothesis  $H_0$  even though it is actually true.

- For a single test at significance level  $\alpha = 0.05$ , there's a 5% chance of making a Type I error.
- 

## 2. What happens if we run multiple tests?

If you test **six pairs of means** with separate  $t$ -tests, each test has its own 5% chance of error. These risks accumulate.

The probability of making **at least one Type I error** across all tests (the **familywise error rate**) is:

$$P(\text{at least one error}) = 1 - (1 - \alpha)^k$$

where  $k$  = number of tests.

For  $k = 6$ ,  $\alpha = 0.05$ :

$$1 - (1 - 0.05)^6 \approx 1 - 0.735 \approx 0.265$$

So instead of a 5% chance of false discovery, now there's about a **26.5% chance** that at least one test gives a false positive.

# A General Notation

**TABLE 3.2**

**Typical Data for a Single-Factor Experiment**

| Treatment (Level) |          | Observations |          |          | Totals   | Averages       |
|-------------------|----------|--------------|----------|----------|----------|----------------|
| 1                 | $y_{11}$ | $y_{12}$     | $\cdots$ | $y_{1n}$ | $y_{1.}$ | $\bar{y}_{1.}$ |
| 2                 | $y_{21}$ | $y_{22}$     | $\cdots$ | $y_{2n}$ | $y_{2.}$ | $\bar{y}_{2.}$ |
| $\vdots$          | $\vdots$ | $\vdots$     | $\cdots$ | $\vdots$ | $\vdots$ | $\vdots$       |
|                   |          |              | $\cdots$ |          |          |                |
| $a$               | $y_{a1}$ | $y_{a2}$     | $\cdots$ | $y_{an}$ | $y_{a.}$ | $\bar{y}_{a.}$ |

$$y_{i.} = \sum_{j=1}^n y_{ij} \quad \bar{y}_{i.} = y_{i.}/n \quad i = 1, 2, \dots, a$$

$$y_{..} = \sum_{i=1}^a \sum_{j=1}^n y_{ij} \quad \bar{y}_{..} = y_{..}/N$$

where  $N = an$  is the total number of observations. We see that the “dot” subscript notation implies summation over the subscript that it replaces.



# Models for the Data

- Means model

$$y_{ij} = \mu_i + \varepsilon_{ij}$$

- Effects model

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, n \end{cases}$$

$\mu$  = an overall mean,  $\tau_i$  =  $i$ th treatment effect,

Assume that  $\varepsilon_{ij}$  = experimental error, IID and follows  $N(0, \sigma^2)$

$$\sum_{i=1}^a \tau_i = 0$$

# The Test

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$$

$$H_1 : \mu_i \neq \mu_j \quad \text{for at least one pair } (i,j)$$

$$H_0 : \tau_1 = \tau_2 = \cdots \tau_a = 0$$

$$H_1 : \tau_i \neq 0 \quad \text{for at least one } i$$

# The Analysis of Variance

The basic single-factor ANOVA model is

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, n \end{cases} \quad \sum_{i=1}^a \tau_i = 0$$

$\mu$  = an overall mean,  $\tau_i$  =  $i$ th treatment effect,  
 $\varepsilon_{ij}$  = experimental error, IID (independent and identically distributed ) and follow  $N(0, \sigma^2)$

$y_{ij}$  are independent and follow  $N(\mu + \tau_i, \sigma^2)$

ANOVA is to **partition the total variability** in the response variable into components that are consistent with the model

# The Analysis of Variance

**Total variability** is measured by the total sum of squares:

$$\begin{aligned}SS_T &= \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 \\&= \sum_{i=1}^a \sum_{j=1}^n [(\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.})]^2 \\&= n \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 \\&\quad + 2 \sum_{i=1}^a \sum_{j=1}^n (\bar{y}_{i.} - \bar{y}_{..})(y_{ij} - \bar{y}_{i.}) \quad \text{Fortunately, the cross term is zero!} \\&= n \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2\end{aligned}$$

# The Analysis of Variance

$$SS_T = SS_{Treatments} + SS_E$$

- A large value of  $SS_{Treatments}$  reflects large differences in treatment means
- A small value of  $SS_{Treatments}$  likely indicates no differences in treatment means
- Degrees of freedom:

$$SS_T : N - 1$$

$$SS_{Treatments} : a - 1$$

$$SS_E : (N - 1) - (a - 1) = N - a$$

$$df_{Total} = df_{Treatments} + df_{Error}$$

# The Analysis of Variance

- A mean square is a sum of squares divided by its degrees of freedom:

$$MS_{Treatments} = \frac{SS_{Treatments}}{a - 1}, \quad MS_E = \frac{SS_E}{N - a}$$

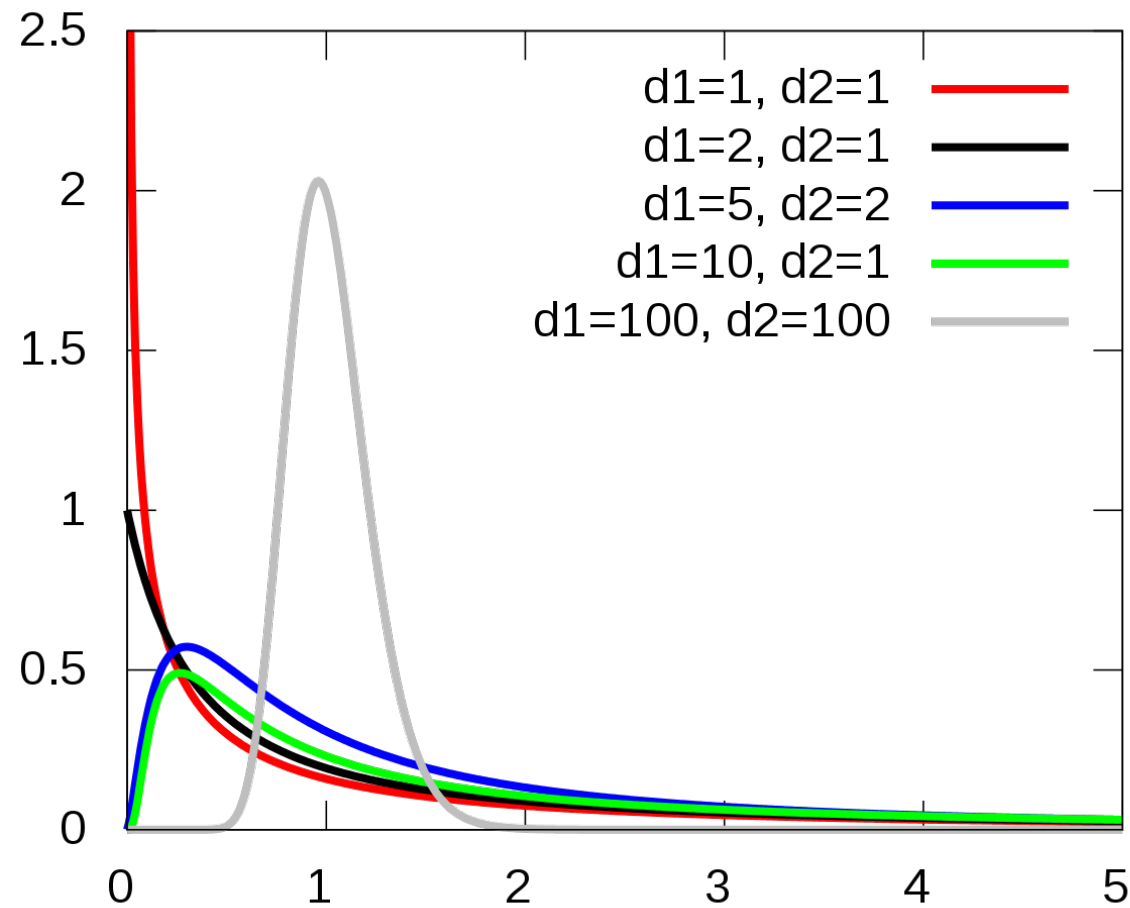
- While sums of squares cannot be directly compared to test the hypothesis of equal means, **mean squares** can be compared.

$SS_{Treatments} / \sigma^2$  and  $SS_E / \sigma^2$  are independently Chi-square distributed

$$F_0 = \frac{SS_{Treatments} / (a - 1)}{SS_E / (N - a)} = \frac{MS_{Treatments}}{MS_E}$$

follows  $F_{a-1, N-a}$  when null hypothesis holds

# F distribution



# The Analysis of Variance is Summarized in a Table

| <b>TABLE 3.3</b>   |   |                           |                          |   |
|--|---|---------------------------|--------------------------|---|
| <b>The Analysis of Variance Table for the Single-Factor, Fixed Effects Model</b> |   |                           |                          |   |
| <b>Source of Variation</b>   | <b>Sum of Squares</b>   | <b>Degrees of Freedom</b> | <b>Mean Square</b>       | <b><math>F_0</math></b>                     |
| Between treatments   | $SS_{\text{Treatments}} = n \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2$ | $a - 1$                   | $MS_{\text{Treatments}}$ | $F_0 = \frac{MS_{\text{Treatments}}}{MS_E}$ |
| Error (within treatments)  | $SS_E = SS_T - SS_{\text{Treatments}}$                                    | $N - a$                   | $MS_E$                   |   |
| Total  | $SS_T = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2$              | $N - 1$                   |                          |   |

- **Reject** the null hypothesis (equal treatment means) if

$$F_0 > F_{\alpha, a-1, N-a}$$



# ANOVA Table

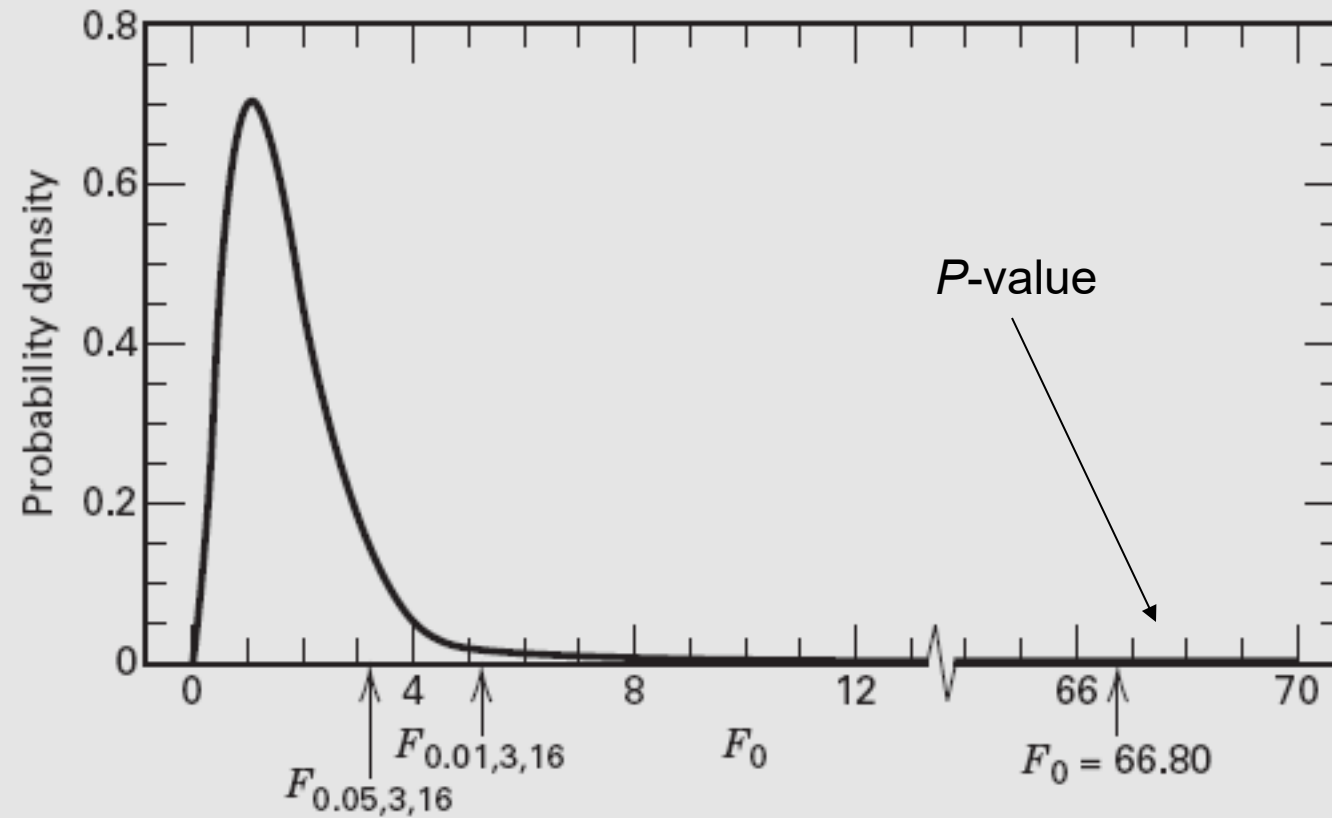
## Example 3-1

**TABLE 3.4**

**ANOVA for the Plasma Etching Experiment**

| Source of Variation | Sum of Square | Degrees of Freedom | Mean Squares | $F_o$         | $P$ -Value |
|---------------------|---------------|--------------------|--------------|---------------|------------|
| RF Power            | 66,870.55     | 3                  | 22,290.18    | $F_o = 66.80$ | < 0.01     |
| Error               | 5339.20       | 16                 | 333.70       |               |            |
| Total               | 72,209.75     | 19                 |              |               |            |

## The Reference Distribution:



■ **FIGURE 3.3** The reference distribution ( $F_{3,16}$ ) for the test statistic  $F_0$  in Example 3.1

# Model Adequacy Checking in the ANOVA

Text reference, Section 3.4, page 68

- **Checking assumptions** is important
- Normality
- Constant variance
- Independence
- Have we fit the right model?
- Later we will talk about what to do if some of these assumptions are **violated**

# Model Adequacy Checking in the ANOVA

- Examination of **residuals**

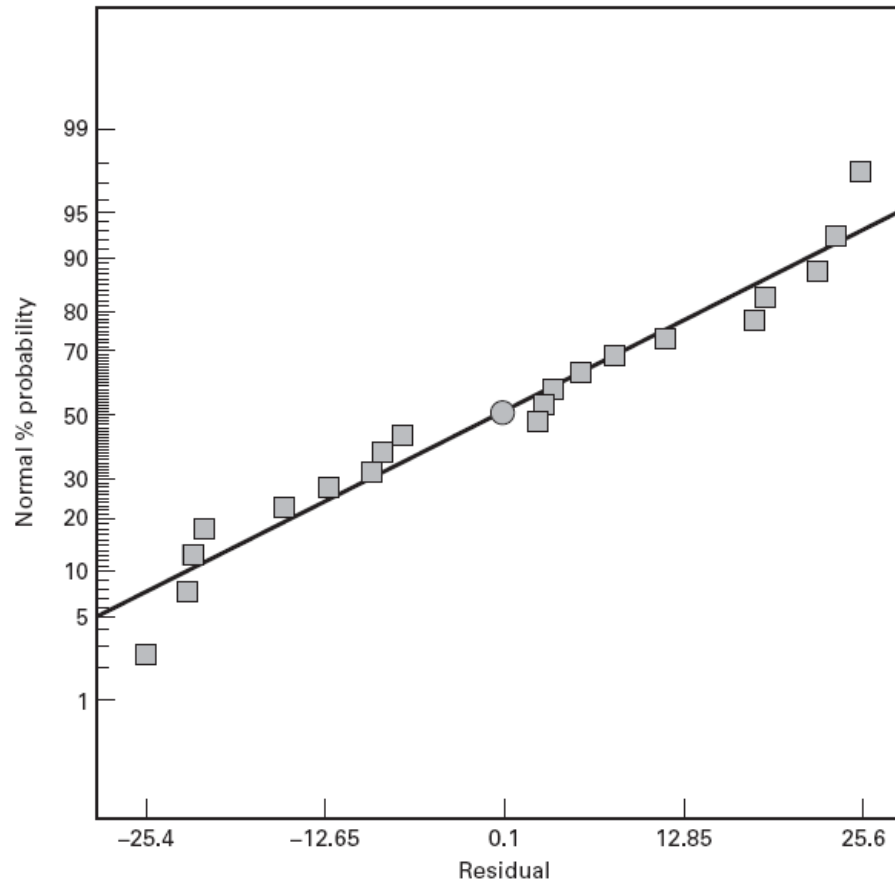
$$\begin{aligned}\hat{y}_{ij} &= \hat{\mu} + \hat{\tau}_i \\ &= \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) \\ &= \bar{y}_{i.}\end{aligned}$$

$$\begin{aligned}e_{ij} &= y_{ij} - \hat{y}_{ij} \\ &= y_{ij} - \bar{y}_{i.}\end{aligned}$$

- **Residual plots** on the following pages

# Residual Plot for Checking Normality

■ **FIGURE 3.4**  
Normal probability  
plot of residuals for  
Example 3.1



```
fit <- lm(y ~ x1 + x2, data = dat)
```

```
r <- rstandard(fit)
```

```
f <- fitted(fit)
```

```
ord # run order
```

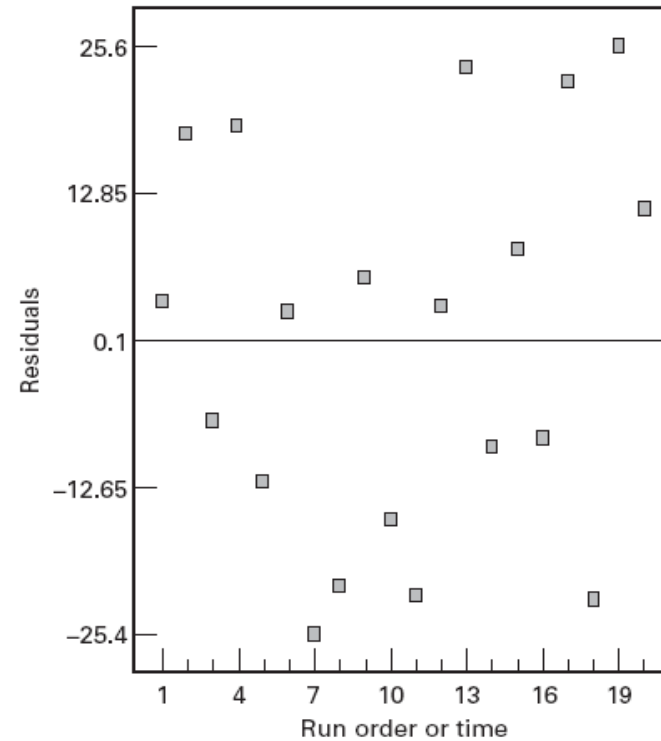
```
qqnorm(r, pch = 22, bg = "grey70",  
       main = "Normal probability plot of residuals",  
       xlab = "Theoretical quantiles", ylab =  
       "Standardized residuals")  
qqline(r, lwd = 2)
```

```
plot(ord, r, pch = 22, bg = "grey70",  
     xlab = "Run order or time", ylab = "Standardized  
residuals")  
abline(h = 0, lwd = 2)
```

```
plot(f, r, pch = 22, bg = "grey70",  
     xlab = "Predicted", ylab = "Standardized residuals")  
abline(h = 0, lwd = 2)
```

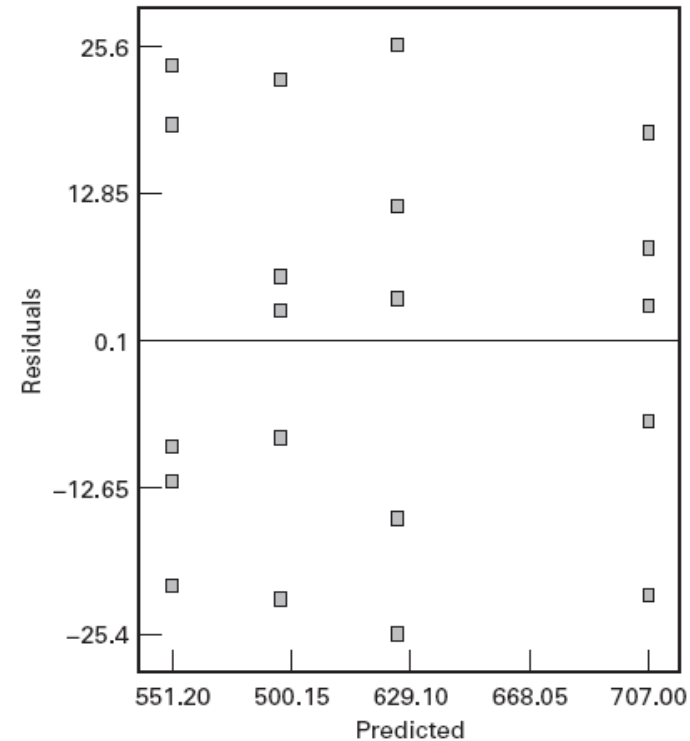
# Residual Plots for Checking Independent and Constant Variance

Test variance independence



■ **FIGURE 3.5** Plot of residuals versus run order or time

Test constant variance



■ **FIGURE 3.6** Plot of residuals versus fitted values