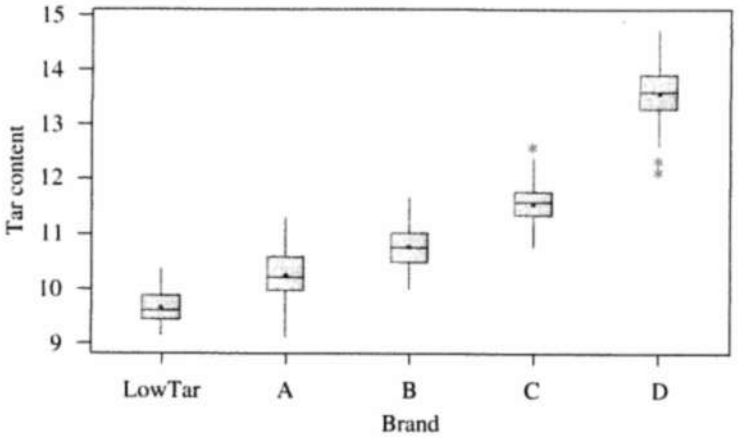
Bus. **8.7** A cigarette manufacturer has advertised that it has developed a new brand of cigarette, LowTar, that has a lower average tar content than the major brands. To evaluate this claim, a consumer testing agency randomly selected 100 cigarettes from each of the four leading brands of cigarettes and 100 from the new brand. The tar content (milligrams) of the cigarettes gave the following results:

| Brand | $\bar{y}_i$ | $s_i$ | $n_i$ |
|---|---|---|---|
| LowTar | 9.64 | .291 | 100 |
| A | 10.22 | .478 | 100 |
| B | 10.77 | .372 | 100 |
| C | 11.57 | .352 | 100 |
| D | 13.59 | .469 | 100 |

A boxplot of the data used to produce the table are given here.

Boxplots of tar content by brand for Exercise 8.7 (means are indicated by solid circles)




**a.** Based on the information contained in the boxplot, does the LowTar brand appear to have a lower average tar content than the other brands?

**b.** Using the computer output shown here, is there a significant ($\alpha = .01$) difference in the average tar content of the five brands of cigarettes?

**c.** What is the $p$-value of the test statistic in (b)?

**d.** What are the practical consequences of making a Type I error with respect to your test in (b)?

```
One-Way Analysis of Variance for Exercise 8.7

Analysis of Variance for Tar Cont
Source      DF        SS        MS         F      P
Brand        4   941.193   235.298  1478.39  0.000
Error      495    78.784     0.159
Total      499  1019.976

                                    Individual 95% CIs for Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  -+---------+---------+---------+-----
1         100     9.644     0.291   (*)
2         100    10.221     0.478        (*)
3         100    10.775     0.372             (*)
4         100    11.570     0.352                  (*)
5         100    13.592     0.469                                    (*)
                                    -+---------+---------+---------+-----
Pooled StDev =    0.399           9.6       10.8       12.0       13.2
```

### 8.3 The Model for Observations in a Completely Randomized Design

**Theory** **8.8** Four populations are to be compared based on differences in their means. Suppose the population means are given as follows:

$$\mu_1 = 20 \qquad \mu_2 = 25 \qquad \mu_3 = 15 \qquad \mu_4 = 35$$

Using the relationship $\mu_i = \mu + \tau_i$, compute the values of $\mu$ and $\tau_i$.

**Consum.** **8.9** Refer to Example 8.1. Apply the model $y_{ij} = \mu + \tau_i + e_{ij}$ to the data in this example by identifying the values of $t$, $n_1$, $n_2$, and $n_3$. Also, estimate the values of $\mu$, $\tau_i$, $\sigma$ from the observed data.

**Med.** **8.10** Refer to Example 8.2. Apply the model $y_{ij} = \mu + e_{ij}$ to the data in this example by identifying the values of $t$, $n_1$, $n_2$, and $n_3$. Also, estimate the values of $\mu$, $\tau_i$, $\sigma$ from the observed data.

### 8.4 Checking on the AOV Conditions

**Theory** **8.11** Suppose that in a study for comparing five population means $n_i = 10$ for $i = 1, \ldots, 5$. The data yield $s_W^2 = 0$. What can we conclude about the 50 residuals: $e_{ij} = y_{ij} - \bar{y}_i$?

**Consum.** **8.12** Refer to Example 8.1.

**a.** From the data set compute the 15 residuals.

**b.** Are the conditions for conducting the AOV $F$ test satisfied by this data set?

**Med.** **8.13** Refer to Example 8.2. Are the conditions for conducting the AOV $F$ test satisfied by this data set?

**Med.** **8.14** Refer to Exercise 8.6. Are the conditions for conducting the AOV $F$ test satisfied by this data set?

**Med.** **8.15** Refer to Exercise 8.7. Are the conditions for conducting the AOV $F$ test satisfied by this data set?

### 8.5 An Alternative Analysis: Transformations of the Data

**Envir.** **8.16** Refer to Example 8.4.

**a.** Apply the AOV $F$ test to the original measurements using $\alpha = .05$.

**b.** Apply the AOV $F$ test to the transformed data using $\alpha = .05$

**c.** Did transforming the data alter your conclusion whether the oxygen content is related to the distance to the mouth of the Mississippi River?

**Pol.** **8.17** Refer to Example 8.6.

**a.** Apply the AOV $F$ test to the original measurements using $\alpha = .05$.

**b.** Apply the AOV $F$ test to the transformed data using $\alpha = .05$

**c.** Did transforming the data alter your conclusion whether there is a difference in the four geographical regions with respect to their opinion of the EPA regulations on air pollution?

**Engin.** **8.18** Refer to Example 7.9. The consumer testing agency was interested in evaluating whether there was a difference in the mean percentage increase in mpg of the three additives. In Example 7.9, we showed that the data did not appear to have a normal distribution.

**a.** Apply the natural logarithm transformation to the data. Do the conditions for applying the AOV $F$ test appear to hold for the transformed data?

**b.** Test for a difference in the means of the three additives using $\alpha = .05$.

**Biol.** **8.19** Refer to Exercise 7.20.

**a.** The biologist hypothesized that the mean weight of deer raised in a zoo would differ from the mean weight of deer raised either in the wild or on a ranch. Do the conditions necessary for applying the AOV $F$ test appear to be valid?

**b.** If the conditions for AOV $F$ test are satisfied, then conduct the test to evaluate the biologist's claim. If not, then suggest a transformation, and conduct the test on the transformed data.

Hort. **8.29** Researchers from the Department of Fruit Crops at a university compared four different preservatives to be used in freezing strawberries. The researchers prepared the yield from a strawberry patch for freezing and randomly divided it into four equal groups. Within each group they treated the strawberries with the appropriate preservative and packaged them into eight small plastic bags for freezing at 0°C. The bags in group I served as a control group, while those in groups II, III, and IV were assigned one of three newly developed preservatives. After all 32 bags of strawberries were prepared, they were stored at 0°C for a period of 6 months. At the end of this time, the contents of each bag were allowed to thaw and then rated on a scale of 1 to 10 points for discoloration. (Note that a low score indicates little discoloration.) The ratings are given here:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Group I** | 10 | 8 | 7.5 | 8 | 9.5 | 9 | 7.5 | 7 |
| **Group II** | 6 | 7.5 | 8 | 7 | 6.5 | 6 | 5 | 5.5 |
| **Group III** | 3 | 5.5 | 4 | 4.5 | 3 | 3.5 | 4 | 4.5 |
| **Group IV** | 2 | 1 | 2.5 | 3 | 4 | 3.5 | 2 | 2 |

**a.** Use the following plots of the residuals and a test of the homogeneity of variances to assess whether the conditions needed to use AOV techniques are satisfied with this data set.

**b.** Test whether there is a difference in the mean ratings using $\alpha = .05$.

**c.** Place 95% confidence intervals on the mean ratings for each of the groups.

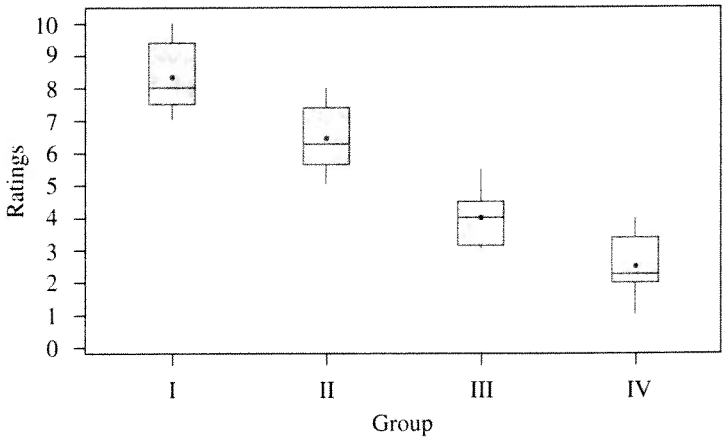
**d.** Confirm your results with the computer output given here.

**One-Way Analysis of Variance for Exercise 8.29**

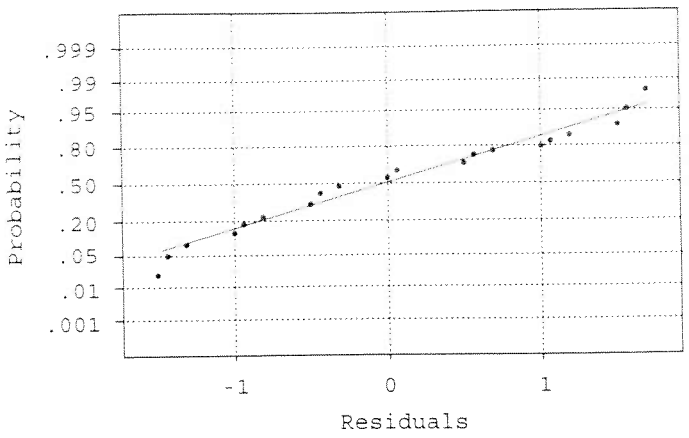
```
Analysis of Variance for Ratings
Source      DF        SS        MS        F        P
Group        3   159.187    53.062    55.67    0.000
Error       28    26.687     0.953
Total       31   185.875

                                      Individual 95% CIs for Mean
                                      Based on Pooled StDev
Group       N      Mean     StDev  --+---------+---------+---------+----
I           8    8.3125    1.0670                                   (---*--)
II          8    6.4375    1.0155                         (--*---)
III         8    4.0000    0.8452              (---*---)
IV          8    2.5000    0.9636    (---*--)
                                      --+---------+---------+---------+----
Pooled StDev =   0.9763               2.0       4.0       6.0       8.0
```

Boxplots of ratings by group for Exercise 8.29 (means are indicated by solid circles)



Normal probability plot of residuals for Exercise 8.29




Average: 0
StDev: 0.927840
N: 32

Anderson-Darling Normality Test
A-Squared: 0.503
P-Value: 0.191

**8.30** Refer to Exercise 8.29. In many situations in which the response is a rating rather than an actual measurement, it is recommended that the Kruskal–Wallis test be used.

**a.** Apply the Kruskal–Wallis test to determine whether there is a shift in the distribution of ratings for the four groups.

**b.** Is the conclusion reached using the Kruskal–Wallis test consistent with the conclusion reached in Exercise 8.29 using AOV?

H.R. **8.31** Salary disputes and their eventual resolutions often leave both employers and employees embittered by the entire ordeal. To assess employee reactions to a recently devised salary and fringe benefits plan, the personnel department obtained random samples of 15 employees from each of three divisions in the company: manufacturing, marketing, and research. The personnel staff asked each employee sampled to respond (in confidence) to a series of questions. Several employees refused to cooperate, as reflected in the unequal sample sizes. The data are given here:

| | Manufacturing | Marketing | Research |
|---|---|---|---|
| Sample size | 12 | 14 | 11 |
| Sample mean | 25.2 | 32.6 | 28.1 |
| Sample variance | 3.6 | 4.8 | 5.3 |

**a.** Write a model for this experimental situation.

**b.** Use the summary of the scored responses to compare the means for the three divisions (the higher a score, the higher the employee acceptance). Use $\alpha = .01$.

Ag. **8.32** Researchers record the yields of corn, in bushels per plot, for four different varieties of corn, A, B, C, and D. In a controlled greenhouse experiment, the researchers randomly assign each variety to eight of 32 plots available for the study. The yields are listed here:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **A** | 2.5 | 3.6 | 2.8 | 2.7 | 3.1 | 3.4 | 2.9 | 3.5 |
| **B** | 3.6 | 3.9 | 4.1 | 4.3 | 2.9 | 3.5 | 3.8 | 3.7 |
| **C** | 4.3 | 4.4 | 4.5 | 4.1 | 3.5 | 3.4 | 3.2 | 4.6 |
| **D** | 2.8 | 2.9 | 3.1 | 2.4 | 3.2 | 2.5 | 3.6 | 2.7 |