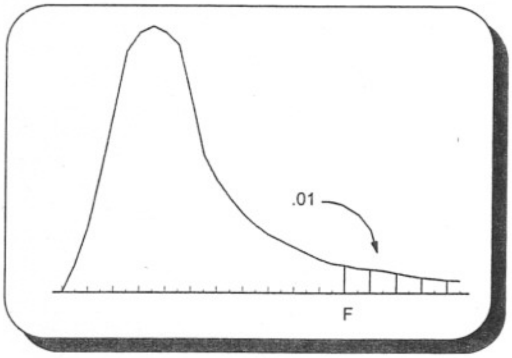
Chapter 18 Analysis of Variance

I. Introduction

A. Understanding variability is important.
   1. Quality manufacturing, successful marketing, and effective training require an understanding of variability.
   2. Using a t-distribution (page 99, section V,C) requires determining the equality of population variances.

B. Analysis of variance will require using the F distribution.

C. Characteristics of the F distribution (named after originator R. Fisher)
   1. It is positively skewed, unimodal, continuous, and asymptotic.
   2. Basic assumptions of the F distribution
      a. Experiments are of random design.
      b. Variables are independent.
      c. Populations are normally distributed with equal variances.
      d. Both interval and ratio levels of data may be analyzed.
   3. Degrees of freedom for the numerator and degrees of freedom for the denominator determine the shape of this family of curves.



II. Testing two sample variances from normal populations

A. Linda wants to compare sales variability of two stores. A sample of 5 from Store #1 measured mean daily sales at \$110. The standard deviation was \$16. A sample of 8 from Store #2 measured mean daily sales at \$125. The standard deviation was \$14. Test at the .02 level of significance whether these two stores have equal sales variances. This is a two-tail test. One-tail tests involve testing a difference in one direction.

B. The 5-step approach to hypothesis testing
   1. These are the null hypothesis and alternate hypothesis.
      $H_0 : \sigma_1^2 = \sigma_2^2$ and $H_1 : \sigma_1^2 \neq \sigma_2^2$
   2. The level of significance will be .02 and $\alpha/2 = .01$.
   3. The relevant statistic F, is the ratio of the sample variances.
      a. The larger variance is always put on top.
      b. This means F is a positive number greater than one.
      c. F's value increases as the difference in variability increases for this one-tail test.

$$F = \frac{s_1^2}{s_2^2}$$

   4. The decision rule will be, if F from the test statistic is large enough (beyond the critical value), the difference in variability is high and the null hypothesis is rejected.
      a. Degrees of freedom is df and df = n - 1.
      b. For the numerator, df is n - 1 = 5 - 1 = 4.
      c. For the denominator, df is n - 1 = 8 - 1 = 7.
      d. From the table, f = 7.85.
   5. Apply the decision rule.

$$F = \frac{s_1^2}{s_2^2} = \frac{16^2}{14^2} = 1.31$$

Accept H_0 because 1.31 < 7.85.
Variances are equal.

F Distribution, Critical Values for the Upper .01

Denominator degrees of freedom	Numerator degrees of freedom: 1	2	3	4
1	4052	4999	5403	5624
2	98.49	99.00	99.17	99.25
3	34.12	30.82	29.46	28.71
4	21.20	18.00	16.69	15.98
5	16.26	13.27	12.06	11.39
6	13.74	10.92	9.78	9.15
7	12.25	9.55	8.45	7.85
8	11.26	8.65	7.59	7.01
9	10.56	8.02	6.99	6.42

See pages ST 5A and 5B for more complete F tables.

III. Testing 3 or more sample means from normally-distributed populations

A. These **analysis of variance** tests are called **ANOVA**.

B. ANOVA measures whether a **treatment variable** has caused a change in a **response variable**.

C. ANOVA is used to measure training effectiveness and product quality when 3 or more samples are involved.

D. Basic procedures
   1. ANOVA uses two separate measures of population variance.
   2. Each is part of the f ratio.
   3. The numerator measures **between treatment variance**.
      This variance is due to differences among sample means.

$$F = \frac{\text{Estimated variance between the treatments}}{\text{Estimated variance within the treatments}}$$

   4. The denominator measures **within treatment variance**. This variance, which is only due to within group differences, is variation due to error.
   5. If the null hypothesis is true, the population means are equal, the expected value of the two measures of population variance will be equal, and F will be one. Otherwise, F will be larger than one.
   6. If, based upon some level of significance, the test statistic F is larger than the critical value of F, the means are not equal and the null hypothesis is rejected.