# Probability & Statistics

## *for Engineers & Scientists*

NINTH EDITION

WALPOLE | MYERS | MYERS | YE

# Probability & Statistics for Engineers & Scientists

*This page intentionally left blank*

# Probability & Statistics for Engineers & Scientists

NINTH EDITION

Ronald E. Walpole
*Roanoke College*

Raymond H. Myers
*Virginia Tech*

Sharon L. Myers
*Radford University*

Keying Ye
*University of Texas at San Antonio*

**Prentice Hall**

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montréal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Editor in Chief: *Deirdre Lynch*
Acquisitions Editor: *Christopher Cummings*
Executive Content Editor: *Christine O'Brien*
Associate Editor: *Christina Lepre*
Senior Managing Editor: *Karen Wernholm*
Senior Production Project Manager: *Tracy Patruno*
Design Manager: *Andrea Nix*
Cover Designer: *Heather Scott*
Digital Assets Manager: *Marianne Groth*
Associate Media Producer: *Vicki Dreyfus*
Marketing Manager: *Alex Gay*
Marketing Assistant: *Kathleen DeChavez*
Senior Author Support/Technology Specialist: *Joe Vetere*
Rights and Permissions Advisor: *Michael Joyce*
Senior Manufacturing Buyer: *Carol Melville*
Production Coordination: *Lifland et al. Bookmakers*
Composition: *Keying Ye*
Cover photo: *Marjory Dressler/Dressler Photo-Graphics*



**Library of Congress Cataloging-in-Publication Data**

Probability & statistics for engineers & scientists/Ronald E. Walpole ... [et al.] — 9th ed.
p. cm.
ISBN 978-0-321-62911-1
1. Engineering—Statistical methods. 2. Probabilities. I. Walpole, Ronald E.
TA340.P738 2011
519.02'462–dc22

2010004857



1 2 3 4 5 6 7 8 9 10—EB—14 13 12 11 10

**Prentice Hall**
is an imprint of

PEARSON

**www.pearsonhighered.com**

ISBN 10: 0-321-62911-6
ISBN 13: 978-0-321-62911-1

# Chapter 13

# One-Factor Experiments: General

## 13.1 Analysis-of-Variance Technique

In the estimation and hypothesis testing material covered in Chapters 9 and 10, we were restricted in each case to considering no more than two population parameters. Such was the case, for example, in testing for the equality of two population means using independent samples from normal populations with common but unknown variance, where it was necessary to obtain a pooled estimate of $\sigma^2$.

This material dealing in two-sample inference represents a special case of what we call the *one-factor problem*. For example, in Exercise 10.35 on page 357, the survival time was measured for two samples of mice, where one sample received a new serum for leukemia treatment and the other sample received no treatment. In this case, we say that there is *one factor*, namely *treatment*, and the factor is at *two levels*. If several competing treatments were being used in the sampling process, more samples of mice would be necessary. In this case, the problem would involve one factor with more than two levels and thus more than two samples.

In the $k > 2$ sample problem, it will be assumed that there are $k$ samples from $k$ populations. One very common procedure used to deal with testing population means is called the **analysis of variance**, or **ANOVA**.

The analysis of variance is certainly not a new technique to the reader who has followed the material on regression theory. We used the analysis-of-variance approach to partition the total sum of squares into a portion due to regression and a portion due to error.

Suppose in an industrial experiment that an engineer is interested in how the mean absorption of moisture in concrete varies among 5 different concrete aggregates. The samples are exposed to moisture for 48 hours. It is decided that 6 samples are to be tested for each aggregate, requiring a total of 30 samples to be tested. The data are recorded in Table 13.1.

The model for this situation may be set up as follows. There are 6 observations taken from each of 5 populations with means $\mu_1, \mu_2, \ldots, \mu_5$, respectively. We may wish to test

$$H_0: \ \mu_1 = \mu_2 = \cdots = \mu_5,$$
$$H_1: \ \text{At least two of the means are not equal.}$$

Table 13.1: Absorption of Moisture in Concrete Aggregates

| **Aggregate:** | **1** | **2** | **3** | **4** | **5** | |
|---|---|---|---|---|---|---|
| | 551 | 595 | 639 | 417 | 563 | |
| | 457 | 580 | 615 | 449 | 631 | |
| | 450 | 508 | 511 | 517 | 522 | |
| | 731 | 583 | 573 | 438 | 613 | |
| | 499 | 633 | 648 | 415 | 656 | |
| | 632 | 517 | 677 | 555 | 679 | |
| Total | 3320 | 3416 | 3663 | 2791 | 3664 | 16,854 |
| Mean | 553.33 | 569.33 | 610.50 | 465.17 | 610.67 | 561.80 |

In addition, we may be interested in making individual comparisons among these 5 population means.

### Two Sources of Variability in the Data

In the analysis-of-variance procedure, it is assumed that whatever variation exists among the aggregate averages is attributed to (1) variation in absorption among observations *within* aggregate types and (2) variation *among* aggregate types, that is, due to differences in the chemical composition of the aggregates. The **within-aggregate variation** is, of course, brought about by various causes. Perhaps humidity and temperature conditions were not kept entirely constant throughout the experiment. It is possible that there was a certain amount of heterogeneity in the batches of raw materials that were used. At any rate, we shall consider the within-sample variation to be **chance or random variation**. Part of the goal of the analysis of variance is to determine if the differences among the 5 sample means are what we would expect due to random variation alone or, rather, due to variation beyond merely random effects, i.e., differences in the chemical composition of the aggregates.

Many pointed questions appear at this stage concerning the preceding problem. For example, how many samples must be tested for each aggregate? This is a question that continually haunts the practitioner. In addition, what if the within-sample variation is so large that it is difficult for a statistical procedure to detect the systematic differences? Can we systematically control extraneous sources of variation and thus remove them from the portion we call random variation? We shall attempt to answer these and other questions in the following sections.

## 13.2 The Strategy of Experimental Design

In Chapters 9 and 10, the notions of estimation and testing for the two-sample case were covered under the important backdrop of the way the experiment is conducted. This falls into the broad category of design of experiments. For example, for the **pooled *t*-test** discussed in Chapter 10, it is assumed that the factor levels (treatments in the mice example) are assigned randomly to the experimental units (mice). The notion of experimental units was discussed in Chapters 9 and 10 and

illustrated through examples. Simply put, experimental units are the units (mice, patients, concrete specimens, time) that **provide the heterogeneity that leads to experimental error** in a scientific investigation. The random assignment eliminates bias that could result with systematic assignment. The goal is to distribute uniformly among the factor levels the risks brought about by the heterogeneity of the experimental units. Random assignment best simulates the conditions that are assumed by the model. In Section 13.7, we discuss **blocking** in experiments. The notion of blocking was presented in Chapters 9 and 10, when comparisons between means were accomplished with **pairing**, that is, the division of the experimental units into homogeneous pairs called **blocks**. The factor levels or treatments are then assigned randomly within blocks. The purpose of blocking is to reduce the effective experimental error. In this chapter, we naturally extend the pairing to larger block sizes, with analysis of variance being the primary analytical tool.

## 13.3 One-Way Analysis of Variance: Completely Randomized Design (One-Way ANOVA)

Random samples of size $n$ are selected from each of $k$ populations. The $k$ different populations are classified on the basis of a single criterion such as different treatments or groups. Today the term **treatment** is used generally to refer to the various classifications, whether they be different aggregates, different analysts, different fertilizers, or different regions of the country.

### Assumptions and Hypotheses in One-Way ANOVA

It is assumed that the $k$ populations are independent and normally distributed with means $\mu_1, \mu_2, \ldots, \mu_k$ and common variance $\sigma^2$. As indicated in Section 13.2, these assumptions are made more palatable by randomization. We wish to derive appropriate methods for testing the hypothesis

$$H_0: \ \mu_1 = \mu_2 = \cdots = \mu_k,$$
$$H_1: \ \text{At least two of the means are not equal.}$$

Let $y_{ij}$ denote the $j$th observation from the $i$th treatment and arrange the data as in Table 13.2. Here, $Y_{i.}$ is the total of all observations in the sample from the $i$th treatment, $\bar{y}_{i.}$ is the mean of all observations in the sample from the $i$th treatment, $Y_{..}$ is the total of all $nk$ observations, and $\bar{y}_{..}$ is the mean of all $nk$ observations.

### Model for One-Way ANOVA

Each observation may be written in the form

$$Y_{ij} = \mu_i + \epsilon_{ij},$$

where $\epsilon_{ij}$ measures the deviation of the $j$th observation of the $i$th sample from the corresponding treatment mean. The $\epsilon_{ij}$-term represents random error and plays the same role as the error terms in the regression models. An alternative and

Table 13.2: $k$ Random Samples

| **Treatment:** | **1** | **2** | $\cdots$ | ***i*** | $\cdots$ | ***k*** | |
|---|---|---|---|---|---|---|---|
| | $y_{11}$ | $y_{21}$ | $\cdots$ | $y_{i1}$ | $\cdots$ | $y_{k1}$ | |
| | $y_{12}$ | $y_{22}$ | $\cdots$ | $y_{i2}$ | $\cdots$ | $y_{k2}$ | |
| | $\vdots$ | $\vdots$ | | $\vdots$ | | $\vdots$ | |
| | $y_{1n}$ | $y_{2n}$ | $\cdots$ | $y_{in}$ | $\cdots$ | $y_{kn}$ | |
| Total | $Y_{1.}$ | $Y_{2.}$ | $\cdots$ | $Y_{i.}$ | $\cdots$ | $Y_{k.}$ | $Y_{..}$ |
| Mean | $\bar{y}_{1.}$ | $\bar{y}_{2.}$ | $\cdots$ | $\bar{y}_{i.}$ | $\cdots$ | $\bar{y}_{k.}$ | $\bar{y}_{..}$ |

preferred form of this equation is obtained by substituting $\mu_i = \mu + \alpha_i$, subject to the constraint $\sum_{i=1}^{k} \alpha_i = 0$. Hence, we may write

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij},$$

where $\mu$ is just the **grand mean** of all the $\mu_i$, that is,

$$\mu = \frac{1}{k}\sum_{i=1}^{k} \mu_i,$$

and $\alpha_i$ is called the **effect** of the $i$th treatment.

The null hypothesis that the $k$ population means are equal against the alternative that at least two of the means are unequal may now be replaced by the equivalent hypothesis

$$H_0\text{: } \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0,$$
$$H_1\text{: At least one of the } \alpha_i \text{ is not equal to zero.}$$

## Resolution of Total Variability into Components

Our test will be based on a comparison of two independent estimates of the common population variance $\sigma^2$. These estimates will be obtained by partitioning the total variability of our data, designated by the double summation

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2,$$

into two components.

**Theorem 13.1:** **Sum-of-Squares Identity**

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2 = n\sum_{i=1}^{k}(\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2$$

It will be convenient in what follows to identify the terms of the sum-of-squares identity by the following notation:

Three Important Measures of Variability

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2 = \text{total sum of squares},$$

$$SSA = n\sum_{i=1}^{k}(\bar{y}_{i.} - \bar{y}_{..})^2 = \text{treatment sum of squares},$$

$$SSE = \sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2 = \text{error sum of squares}.$$

The sum-of-squares identity can then be represented symbolically by the equation

$$SST = SSA + SSE.$$

The identity above expresses how between-treatment and within-treatment variation add to the total sum of squares. However, much insight can be gained by investigating the **expected value of both *SSA* and *SSE***. Eventually, we shall develop variance estimates that formulate the ratio to be used to test the equality of population means.

**Theorem 13.2:**

$$E(SSA) = (k-1)\sigma^2 + n\sum_{i=1}^{k}\alpha_i^2$$

The proof of the theorem is left as an exercise (see Review Exercise 13.53 on page 556).

If $H_0$ is true, an estimate of $\sigma^2$, based on $k-1$ degrees of freedom, is provided by this expression:

Treatment Mean Square

$$s_1^2 = \frac{SSA}{k-1}$$

If $H_0$ is true and thus each $\alpha_i$ in Theorem 13.2 is equal to zero, we see that

$$E\left(\frac{SSA}{k-1}\right) = \sigma^2,$$

and $s_1^2$ is an unbiased estimate of $\sigma^2$. However, if $H_1$ is true, we have

$$E\left(\frac{SSA}{k-1}\right) = \sigma^2 + \frac{n}{k-1}\sum_{i=1}^{k}\alpha_i^2,$$

and $s_1^2$ estimates $\sigma^2$ plus an additional term, which measures variation due to the systematic effects.

A second and independent estimate of $\sigma^2$, based on $k(n-1)$ degrees of freedom, is this familiar formula:

Error Mean Square

$$s^2 = \frac{SSE}{k(n-1)}$$

It is instructive to point out the importance of the expected values of the mean squares indicated above. In the next section, we discuss the use of an ***F*-ratio** with the treatment mean square residing in the numerator. It turns out that when $H_1$ is true, the presence of the condition $E(s_1^2) > E(s^2)$ suggests that the $F$-ratio be used in the context of a **one-sided upper-tailed test**. That is, when $H_1$ is true, we would expect the numerator $s_1^2$ to exceed the denominator.

## Use of *F*-Test in ANOVA

The estimate $s^2$ is unbiased regardless of the truth or falsity of the null hypothesis (see Review Exercise 13.52 on page 556). It is important to note that the sum-of-squares identity has partitioned not only the total variability of the data, but also the total number of degrees of freedom. That is,

$$nk - 1 = k - 1 + k(n - 1).$$

## *F*-Ratio for Testing Equality of Means

When $H_0$ is true, the ratio $f = s_1^2/s^2$ is a value of the random variable $F$ having the $F$-distribution with $k-1$ and $k(n-1)$ degrees of freedom (see Theorem 8.8). Since $s_1^2$ overestimates $\sigma^2$ when $H_0$ is false, we have a one-tailed test with the critical region entirely in the right tail of the distribution.

The null hypothesis $H_0$ is rejected at the $\alpha$-level of significance when

$$f > f_\alpha[k - 1, k(n - 1)].$$

Another approach, the $P$-value approach, suggests that the evidence in favor of or against $H_0$ is

$$P = P\{f[k - 1, k(n - 1)] > f\}.$$

The computations for an analysis-of-variance problem are usually summarized in tabular form, as shown in Table 13.3.

Table 13.3: Analysis of Variance for the One-Way ANOVA

| **Source of Variation** | **Sum of Squares** | **Degrees of Freedom** | **Mean Square** | **Computed *f*** |
|---|---|---|---|---|
| Treatments | $SSA$ | $k-1$ | $s_1^2 = \dfrac{SSA}{k-1}$ | $\dfrac{s_1^2}{s^2}$ |
| Error | $SSE$ | $k(n-1)$ | $s^2 = \dfrac{SSE}{k(n-1)}$ | |
| Total | $SST$ | $kn-1$ | | |

**Example 13.1:** Test the hypothesis $\mu_1 = \mu_2 = \cdots = \mu_5$ at the 0.05 level of significance for the data of Table 13.1 on absorption of moisture by various types of cement aggregates.

***Solution*:** The hypotheses are

$$H_0: \ \mu_1 = \mu_2 = \cdots = \mu_5,$$
$$H_1: \ \text{At least two of the means are not equal.}$$
$$\alpha = 0.05.$$

Critical region: $f > 2.76$ with $v_1 = 4$ and $v_2 = 25$ degrees of freedom. The sum-of-squares computations give

$$SST = 209{,}377, \quad SSA = 85{,}356,$$
$$SSE = 209{,}377 - 85{,}356 = 124{,}021.$$

These results and the remaining computations are exhibited in Figure 13.1 in the *SAS* ANOVA procedure.

```
                          The GLM Procedure
Dependent Variable: moisture

                                                Sum of
Source                     DF       Squares     Mean Square    F Value    Pr > F
Model                       4     85356.4667      21339.1167       4.30    0.0088
Error                      25    124020.3333       4960.8133
Corrected Total            29    209376.8000

  R-Square     Coeff Var      Root MSE    moisture Mean
  0.407669      12.53703      70.43304         561.8000

Source                     DF     Type I SS     Mean Square    F Value    Pr > F
aggregate                   4    85356.46667     21339.11667       4.30    0.0088
```

Figure 13.1: *SAS* output for the analysis-of-variance procedure.

Decision: Reject $H_0$ and conclude that the aggregates do not have the same mean absorption. The $P$-value for $f = 4.30$ is 0.0088, which is smaller than 0.05.

In addition to the ANOVA, a box plot was constructed for each aggregate. The plots are shown in Figure 13.2. From these plots it is evident that the absorption is not the same for all aggregates. In fact, it appears as if aggregate 4 stands out from the rest. A more formal analysis showing this result will appear in Exercise 13.21 on page 531.

During experimental work, one often loses some of the desired observations. Experimental animals may die, experimental material may be damaged, or human subjects may drop out of a study. The previous analysis for equal sample size will still be valid if we slightly modify the sum of squares formulas. We now assume the $k$ random samples to be of sizes $n_1, n_2, \ldots, n_k$, respectively.

Sum of Squares, Unequal Sample Sizes

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2, \ SSA = \sum_{i=1}^{k} n_i(\bar{y}_{i.} - \bar{y}_{..})^2, \ SSE = SST - SSA$$