From Harnett & Soni, Statistical Methods for Business and Economics, Fourth Edition

Appendix A

Subscripts, Summations, Variables and Functions, Calculus Review

Education is . . . hanging around until you've caught on.

ROBERT FROST

Subscripts and Summations

A.1

Throughout this book we use certain symbols to distinguish between the numbers in a set of data, and to indicate the sum of such numbers. For example, we may wish to distinguish between the monthly sales of a certain business, and then sum these monthly sales to get the yearly sales. To do this, suppose that we let the symbol x denote the monthly sales of this firm. Furthermore, we will add a subscript to this symbol to denote which month is being represented. Thus, x_1 = sales in the first month, x_2 = sales in the second month, and so forth, with x_{12} = sales in the twelfth month. That is, if sales in the sixth month were 120 units, then we would write $x_6 = 120$. The notation x_i thus stands for "sales in the ith month," where i can be any number from 1 to 12; that is, $i = 1, 2, \ldots, 12$. The dots in this last expression are used to indicate "and so on."

Now, assume that we want to sum the sales for all 12 months in a year, which is

$$x_1 + x_2 + \cdots + x_{12}$$

Another way of writing this sum is to use the Greek letter $\sum$ (capital sigma). This symbol is read as "take the sum of." At the bottom of this $\sum$ sign we usually place the first value of i that is to be included in the sum. The last value of i to be summed is usually placed at the top of the sum sign. Thus,

$$\sum_{i=1}^{12} x_i$$

is read as "sum the values of x_i starting from $i = 1$ and ending with $i = 12$." That is,

$$\sum_{i=1}^{12} x_i = x_1 + x_2 + \cdots + x_{12}$$

Similarly, suppose that we want the sum of only the last seven months in the year. This sum is written as follows:

$$\sum_{i=6}^{12} x_i = x_6 + x_7 + \cdots + x_{12}$$

In statistics we often do not know in advance what the final value in a summation will be. For example, we know that we want to sum a set of sales values, but we do not know how many values there are to be summed. To designate this situation, we will let the symbol n represent the last number in the sum (where n can be any integer value, such as 1, 2, 3, . . .). The notation

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$$

is thus read as "the sum of n numbers, where the first number is x_1, the second is x_2, and the last is x_n." In summing monthly sales over a year, we would thus let $n = 12$, so that $\sum_{i=1}^{n} x_i = \sum_{i=1}^{12} x_i$.

Perhaps we should mention that, in some chapters in this book, we have sometimes omitted the limits of summation, and simply written $\sum x_i$. This notation should be interpreted to mean "sum all relevant values of x_i." In these instances we have made sure that the reader always knows what the relevant values of x_i are. Also, we might point out that the choice of symbols in designating a sum of numbers is often quite arbitrary. For example, we might have used the letter y to denote monthly sales (instead of x), and used the letter j as a subscript (instead of i). In this case $\sum_{j=1}^{12} y_j$ would denote the sum of the 12 monthly values.

Double Summations

In a number of chapters in this book we have found it convenient to use *two* subscripts instead of just one. In these instances the first subscript indicates one characteristic under study, and the second subscript some other characteristic. For example, suppose that we let x_{ij} = sales in the ith month by the jth salesman. The notation $x_{6,2} = 15$ would indicate that in the sixth month

($i = 6$), salesman number 2 ($j = 2$) sold 15 units. Using the same procedure as described above, we can denote the total sales over 12 months by the jth salesman as the sum of x_{1j} (sales in the 1st month by the jth salesman) plus $x_{2j}, \ldots,$ plus $x_{12,j}$ (sales in the 12th month by the jth salesman). That is,

$$\textit{Total sales by salesman } j\text{:}\quad \sum_{i=1}^{12} x_{ij} = x_{1j} + x_{2j} + \cdots + x_{12j}$$

Another example of a similar type of sum is the sum of sales in the ith month (where i is some number between 1 and 12) by all the sales representatives in the company. If we let m = total number of sales reps, then this sum is x_{i1} (sales in month i by sales rep 1) plus $x_{i2}, \ldots,$ plus x_{im} (sales in month i by sales rep m). That is,

$$\textit{Total sales in month } i\text{:}\quad \sum_{j=1}^{m} x_{ij} = x_{i1} + x_{i2} + \cdots + x_{im}$$

Finally, we might wish to sum over all months ($i = 1, 2, \ldots, 12$) and all salesmen ($j = 1, 2, \ldots, m$). This sum could be written as

Total sales over all months and all salesmen:

$$\sum_{\text{All } j}\sum_{\text{All } i} x_{ij} = \left\{\begin{array}{l} x_{11} + x_{12} + \cdots + x_{1m} \\ + x_{21} + x_{22} + \cdots + x_{2m} \\ \quad\vdots \\ + x_{12,1} + x_{12,2} + \cdots + x_{12,m} \end{array}\right\}$$

Variables and Functions

A.2

Variables

Variables and the relationship between variables represent an important part of statistics. Hence, it is important that we define these concepts carefully.

A **variable** is a quantity that may assume any one of a set of values. For example, we might describe the worth of a common stock by the variable "current worth on the stock market." The values of this variable are the different prices the stock can assume. Or, we might be interested in describing how well a specific brand of alkaline battery works by defining the variable "the length of time before failure when in constant use." The values of this variable are the various times it might take before the battery fails.

Variables are often classified according to whether their values are *discrete* or *continuous.* The values of a discrete variable are individually distinct; that is, they are separable from one another. The price of a common stock, for instance, represents a discrete variable because the prices a stock can assume are all separate values, distinguishable from one another. The following examples also represent discrete variables:

1. the number of defectives in a production lot,