2 Subscript and Summation Notation

2.1 Variable and Subscript Notation

Most of the calculations we perform in statistics are repetitive operations on lists of numbers. For example, we compute the sum of a set of numbers, or the sum of the squares of the numbers, in many statistical formulas. We need an efficient notation for talking about such operations in the abstract.

2.1.1 Single Subscript Notation

In the simplest situations, we have one or two (or perhaps three) lists, and we wish to refer to particular numbers in those lists. This is the kind of situation you have probably already dealt with repeatedly in your undergraduate course in statistics. In this case, we represent numbers in a list with a notation of the form

$$X_i$$

The symbol X is the "list name," or the name of the variable represented by the numbers on the list. The symbol i is a "subscript," or "position indicator." It indicates which number in the list, starting from the top, you are referring to.

For example, if the X list consists of the numbers 11, 3, 12, 7, 19 the value of X_3 would be 12, because this is the third number (counting from the top) in the X list. Single subscript notation extends naturally to a situation where there are two or more lists. For example suppose a course has 4 students, and they take two exams. The first exam could be given the variable name X, the second Y, as in the table below.

Student	X	Y
Smith	87	85
Chow	65	66
Benedetti	83	90
Abdul	92	97

In the above list, Chow's grade of 66 on the second exam is Y_2. Abdul's 92 on the first exam is X_4.

Using different variable names to stand for each list works well when there are only a few lists, but it can be awkward for two reasons.

- In some cases the number of lists can become large. This arises quite frequently in some branches of psychology, when personality inventory data are recorded. In such cases, there might be literally hundreds of variables for each subject.

- When general theoretical results are being developed, we often wish to express the notion of some operation being performed "over all of the lists." It is difficult to express such ideas efficiently when each list is represented by a different letter, and the available list of letters is limited.

2.1.2 Double Subscript Notation

To combat the difficulties that arise when more than one list is being discussed, it is often more convenient to use double subscript notation. In this notation, data are presented in a rectangular array. The data are indicated with a single variable name, and two subscripts, like this

$$X_{ij}$$

The first subscript refers to the row that the particular value is in, the second subscript refers to the column. For example here

$$\begin{array}{ccc} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{array}$$

is a matrix, a rectangular array containing 3 rows and 3 columns. You count down to get to a particular row, and you count across from left to right to get to a particular column.

Example 2.1.1 Test your understanding of double subscript notation by finding X_{23} and X_{31} in the table below.

$$
\begin{array}{ccc}
1 & 6 & 32 \\
3 & 23 & 112 \\
12 & 21 & 34 \\
53 & 8 & 64 \\
4 & 14 & 5 \\
\end{array}
$$

Solution Go down to the second row and over to the third column to find $X_{23} = 112$. Go down to the third row and stay in the first column to find $X_{31} = 12$.

Example 2.1.2 Give the row and column subscript form for the number 14 in the table in Example 2.1.1.

Solution We find the number 14 in the 5th row and the 2nd column. Hence it is X_{52}.

When there are more than 9 elements in a row or column, this notation can be ambiguous. Suppose, for example, you wanted the element from the 11th row and the 2nd column of a 20 by 20 data array. If you write X_{112}, it could mean the element in row 1 and column 12. How do you handle this?

Oddly enough, you hardly ever see this question addressed in textbooks! Textbook "training" examples tend to have only a few rows and/or columns. Obviously you've got to do something. Generally, anything goes in these kinds of situations so long as it is very unlikely that anyone will be confused. We have several options. One is to separate the subscripts with spaces, like this

$$X_{11\ 2}$$

Another option is to surround each subscript with brackets, like this

$$X_{[11][2]}$$

Unfortunately, this choice produces ambiguities of its own when adopted as a general choice, because in some types of expressions, the notation might imply multiplication of subscripts, while in other situations it would be perfectly acceptable. Here is a notation that works well across a wide variety of situations

$$X_{11,2}$$

This is the notation we will employ in situations where there are more than 9 rows or columns in a two-dimensional data array.

Exercises for Section 2.1

X	Y	Z	W
11	21	19	17
4	10	12	24
16	15	14	13

Table 2.1: Sample Data on 4 Variables

(1) Given the data in Table 2.1, what is the proper variable and subscript designation of the value 24? ⫸

(2) In the same data set as in Exercise 1, what is the proper variable and subscript designation of the value 16? ⫸

$$X = \begin{bmatrix} 1 & 7 & 4 \\ 5 & 8 & 6 \\ 6 & 3 & 9 \end{bmatrix}$$

Table 2.2: Sample Data Matrix X

(3) In Table 2.2, which gives numerical values for a matrix X what is the double-subscript designation for the number whose value is 6? ⫸

(4) In Table 2.2, what is the numerical value of X_{22}? ⫸

2.2 Single Summation Notation

Many statistical formulas involve repetitive summing operations. Consequently, we need a general notation for expressing such operations. You may be already familiar with this notation from an undergraduate course, but you may not be aware of its full potential. We shall begin with some simple examples, and work through to some that are more complex and challenging. Many summation expressions involve just a single summation operator. They have the following general form

$$\sum_{i=1}^{n} X_i$$

In the above expression, the i is the summation index, 1 is the start value, n is the stop value. Summation notation works according to the following rules.

(a) The summation operator governs everything to its right. up to a natural break point in the expression. The break point is usually obvious from standard rules for algebraic expressions, or other aspects of the notation, and we will discuss this point further below.

(b) To evaluate an expression, begin by setting the summation index equal to the start value. Then evaluate the algebraic expression governed by the summation sign.

(c) Increase the value of the summation index by 1. Evaluate the expression governed by the summation sign again, and add the result to the previous value.

(d) Keep repeating step 3 until the expression has been evaluated and added for the stop value. At that point the evaluation is complete, and you stop.

Example 2.2.1 Suppose our list has just 5 numbers, and they are 1,3,2,5,6. Evaluate

$$\sum_{i=1}^{5} X_i^2$$

Solution In this case, we begin by setting i equal to 1, and evaluating X_1^2. Since $X_1 = 1$, our first evaluation produces a value of 1.

Next, we set i equal to 2, and evaluate X_2^2, obtaining 9, which we add to the previous result of 1. We continue in this manner, obtaining

$$\left[1^2 + 3^2 + 2^2 + 5^2 + 6^2\right] = 75$$

The order of operations is as important in summation expressions as in other mathematical notation. In the following example, we compute the square of the sum of the numbers in our list.

Example 2.2.2 Using the same numbers as in Example 2.2.1, evaluate the following expression:

$$\left[\sum_{i=1}^{5} X_i\right]^2$$

Solution In this case, we add up all the numbers, then square the result. We obtain

$$[1 + 2 + 3 + 5 + 6]^2 = 17^2 = 289$$

2.3   The Algebra of Summations

Many facts about the way lists of numbers behave can be derived using some basic rules of summation algebra. These rules are simple yet powerful. In this section, we develop these rules and employ them immediately to prove our first (very simple) statistical result.

2.3.1   The First Constant Rule

The first rule is based on a fact that you first learned when you were around 8 years old — multiplication is simply repeated addition. That is, to compute 3 times 5, you compute $5 + 5 + 5$.

Another way of viewing this fact is that, if you add a constant a certain number

of times, you have multiplied the constant by the number of times it was added. Symbolically, we express the result as

$$\sum_{i=x}^{y} a = (y - x + 1)a \tag{2.1}$$

This rule, which we will refer to as "The First Constant Rule of Summation Algebra," is used in many derivations to eliminate summation signs and make an expression simpler. Note that, if the summation index runs from x to y, the constant is added $y - x + 1$ times, not $y - x$ times! For example, if the summation index runs from 2 to 3, you go through 2 cycles, not 1. Even experienced practitioners forget this on occasion, and assume that a summation index running from x to y results in $y - x$ cycles. This "off by one" error plagues computer programmers in a number of contexts. It is unlikely that you have seen the First Constant Rule of Summation Algebra stated in the form of Equation 2.1. It is much more likely that you have seen the following less general version which applies when the starting index value is 1.

$$\sum_{i=1}^{n} a = na \tag{2.2}$$

One of the problems beginners experience with this rule and its application is that the form of Equation 2.2 is deceptively simple. The equation actually says more than it appears to at first glance. This happens frequently in statistics, and so we will examine the phenomenon carefully here. The most important thing to realize is that the symbol a in Equations 2.1 and 2.2is a placeholder. It actually stands for any expression, no matter how complicated, that does not vary as a function of the summation index (i in this case). Any algebraic expression that does not contain an i or can be reduced to such an expression falls under the rubric of the First Constant Rule. In our first example, the application of the rule is straightforward, because the expression governed by the summation operator is so simple that it is difficult not to notice that it is a constant.

Example 2.3.1 Application of the First Constant Rule.

$$\sum_{i=2}^{3} 2 = \ ?$$

Solution In this case, we are adding the number 2 twice, so the answer is 4. We can also solve by direct application of Equation 2.1. This leads to

$$\sum_{i=2}^{3} 2 = (3 - 2 + 1)\, 2 = (2)2 = 4$$

In our second example, the expression governed by the summation operator is much more complex than in Example 2.3.1, but it is still a constant.

Example 2.3.2 Applying the First Constant Rule to a Complex Expression.

$$\sum_{i=1}^{n} (2X_j - 5)^2 = \ ?$$

Solution In this case, the expression is visually more complex then in Example 2.3.1. This tends to obscure the fact that it is constant with respect to the summation index i. However, the expression does not contain an i, so the First Constant Rule applies. The expression reduces to

$$n \left(2X_j - 5\right)^2$$

2.3.2 The Second Constant Rule

The second rule of summation algebra, like the first, derives from a principle we learned very early in our educational careers. When we were first learning algebra, we discovered that a common multiple could be factored out of additive expressions. For example,

$$2x + 2y = 2\left(x + y\right)$$

or

$$8w + 8x + 8y = 8\left(w + x + y\right)$$

Consequently, we may write

$$\sum_{i=1}^{n} aX_i = a \sum_{i=1}^{n} X_i \tag{2.3}$$

Again, the rule actually says more than it appears to. At first glance, it appears to be a rule about multiplication. You can move a factorable constant outside of a summation operator. However, the term a could also stand for a fraction, and so the rule also applies to factorable divisors in the summation expression. The following two examples explore these applications of the rule.

Example 2.3.3 Factoring a Common Multiple.

$$\sum_{i=1}^{n} 6yX_i = \ ?$$

Solution In this case the term $6y$ can be factored from the expression governed by the summation operator. Consequently, it may be moved outside the summation sign as follows:

$$6y \sum_{i=1}^{n} X_i$$

Example 2.3.4 Factoring a Common Divisor.

$$\sum_{i=1}^{n} \left(\frac{X_i}{2}\right) = \ ?$$

Solution Here the common divisor may be factored out, yielding

$$\frac{1}{2} \sum_{i=1}^{n} X_i$$

2.3.3 The Distributive Rule

The third rule of summation algebra relates to a third notion that we learned early in our mathematics education — when numbers are added or subtracted, the ordering of addition and/or subtraction doesn't matter. For example

$$(1+2) + (3+4) = (1+2+3+4)$$

Consequently, in summation notation, we have

$$\sum_{i=1}^{n} (X_i + Y_i) = \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} Y_i \tag{2.4}$$

Again, the rule might appear to be saying less at first glance than it is, since the terms on the left may be either positive or negative. Hence Equation 2.4 implies also that

$$\sum_{i=1}^{n} (X_i - Y_i) = \sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i \tag{2.5}$$

2.3.4 Proving a Result With Summation Algebra

In this section, we prove a basic Theorem using the rules of summation algebra. Before we begin the proof, we must introduce some basic definitions.

Definition The **sample mean**, or arithmetic average of a list of n numbers is defined as

$$\overline{X}_\bullet = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{2.6}$$

Definition The **deviation score** dx_i corresponding to the **raw score** X_i is defined as

$$dx_i = X_i - \overline{X}_\bullet \tag{2.7}$$

i.e.,the position of X_i relative to the group average.

Frequently we shall refer to the scores X_i as we originally find them as the **raw scores**. The deviation score corresponding to a particular raw score expresses where that score is relative to the sample mean, or group average. So, for example, if a person's deviation score on an examination is $+2$, it means the person was two points above the group average. In many contexts, the group average will vary for reasons that are in some sense irrelevant to the measurement task at hand. For example, one professor's exam may be easier than another's, and so the class average may be higher. In such situations, if we have to compare grades for students in the two classes, their deviation scores may be a better performance indicator than their raw scores. By referencing the scores to the mean within their group, you factor out group mean differences.

Example 2.3.5 Calculating Deviation Scores. If you work with deviation scores for any length of time, you quickly notice that they seem to inevitably add up to zero. For example, consider the following simple set of scores X_i and their corresponding deviation scores dx_i. We obtain the deviation scores by subtracting the mean of the X scores, 3, from each X.

X	dx
5	+2
4	+1
3	0
2	−1
1	−2

You may verify easily that, in the above numerical example, the deviation scores sum to zero. In the next Theorem, we use the rules of summation algebra to prove that, for any list of numbers, the deviation scores must always add up to zero. At first, you may find the notion of having to prove results in statistics an intimidating departure from the approach taken in your undergraduate training. However, we present proofs in this text to try to get you to accept the idea that you can actually understand where statistical techniques came from and why, rather than having

to simply accept these techniques as something to memorize. Rest assured that a generation of graduate and undergraduate students in my courses have mastered all of these proofs. You can do it too. Follow the first few proofs closely, and you should get the hang of it quickly.

Theorem 2.3.6 (The Sum of Deviation Scores)   *For any list of n numbers, The sum of deviation scores is always zero. That is*

$$\sum_{i=1}^{n} dx_i = 0 \tag{2.8}$$

Proof *The proof that follows is typical of many of the summation algebra proofs that appear in this book. It begins by restating an equality relationship that is to be proven in mathematical notation. Next, the left side of the equation is expanded in straightforward fashion. Then, available results are applied to simplify the expression and obtain the quantity on the right side of the original equation. Most of the proofs in this book involve mathematical insight or creativity only at one or two key points. I will try to make clear where the difficult points are in each proof.*

We begin by restating the quantity we wish to prove is zero.

$$\sum_{i=1}^{n} dx_i = ?$$

Next we substitute the definition of a deviation score (Equation 2.7) into the formula, obtaining

$$\sum_{i=1}^{n} dx_i = \sum_{i=1}^{n} \left(X_i - \overline{X}_\bullet \right)$$

At this point, we employ the Distributive Rule of Summation Algebra (Equation 2.4) to distribute the summation sign to the two terms on the right side of

the expression.

$$\sum_{i=1}^{n} \left(X_i - \overline{X}_\bullet \right) = \sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \overline{X}_\bullet$$

One of the two terms on the right involves summing an expression that does not contain an i, and so the First Constant Rule (Equation 2.2) applies. The expression becomes

$$\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \overline{X}_\bullet = \sum_{i=1}^{n} X_i - n\overline{X}_\bullet$$

Next we need to recall the definition of the sample mean (Equation 2.6). Multiply both sides of the definition by n, and one finds that

$$n\overline{X}_\bullet = \sum_{i=1}^{n} X_i$$

so we obtain

$$\sum_{i=1}^{n} X_i - n\overline{X}_\bullet = \sum_{i=1}^{n} X_i - \sum_{i=1}^{n} X_i = 0$$

This completes our proof.

2.4   Double and Multiple Summation Expressions

So far, we have developed single and double subscript notation, and an algebra of summations. We have proved a simple but important statistical result using the notation and algebra. However, the algebra of summations has been presented and developed with respect to summation expressions involving only a single summation sign and a single subscript index. These examples are useful, but they fail to convey the full power and complexity of summation and subscript notation. In this section, we explore the use of summation and subscript notation in more complex expressions, beginning with simple expressions you probably saw in your undergraduate course, and working through to more complicated expressions.

2.4.1   Interpreting Double and Multiple Summation Expressions

First, we explore the basics. How does one interpret what one sees in an expression involving two or more summation indices? In the pure mathematical sense, we actually know already. That is, the four rules for summation operators given in 2.2 actually cover the case of multiple summation indices. However, there are many simplification strategies that we can employ when examining summation expressions, so simply stating the mathematical truth would be inadequate in this case. Consequently, after examining multiple summation expressions from the strict mathematical standpoint, we will go on to describe some of the simplification strategies you will find useful in practice for interpreting statistical formulas that use summations.

Example 2.4.1 (A Simple Double Summation Expression.)   Consider the following expression

$$\sum_{i=1}^{3}\sum_{j=1}^{3} X_{ij}$$

What does this expression mean?

Solution   If we review the rules for evaluating summation expressions in 2.2, we discover that these rules handle the situation quite nicely. Rule 1 says that a summation operator governs everything to its right. That means that, starting from the left of the expression, the first summation operator, the one with i as its index, governs the entire subexpression to its right. To remind us of that, I will surround this expression with large parentheses.

$$\sum_{i=1}^{3}\left(\sum_{j=1}^{3} X_{ij}\right)$$

Now examine the second summation operator, with summation index j. This operator governs the entire subexpression to its right. To remind us of that, I will surround this subexpression with brackets.

$$\sum_{i=1}^{3}\left(\sum_{j=1}^{3}\left[X_{ij}\right]\right)$$

We are now ready to evaluate the expression. We begin, according to rule 2 of 2.2, by setting the summation index i to one. According to rule 3, we then evaluate the entire expression governed by the first summation operator. This is the expression contained in large parentheses. This expression is itself a summation expression. Evaluating this expression therefore involves examining and evaluating a second summation operator. We must evalute the expression

$$\left(\sum_{j=1}^{3} \left[X_{ij} \right] \right)$$

while the summation index controlled by the first operator is held constant at the value $i = 1$. We must, therefore, evaluate the expression

$$\left(\sum_{j=1}^{3} \left[X_{1j} \right] \right)$$

This expression is a routine single summation expression, much like the ones we have already evaluated. To evaluate it, we run the second summation index j over the values from 1 to 3, each time evaluating the expression governed by the second summation sign, and adding the result. We obtain

$$\left(\left[X_{11} \right] + \left[X_{12} \right] + \left[X_{13} \right] \right)$$

Note that, to help you follow the logic of the summation operation, I have deliberately left the brackets around the subexpressions. At this point, we have evaluated the entire subexpression governed by the first summation operator, so we are ready for the next step.

Next, we set the index of the first summation operator i to its next value. We set $i = 2$ and evaluate the subexpression governed by the first summation index again. At this point, we are evaluating the expression

$$\left(\sum_{j=1}^{3} \left[X_{2j} \right] \right)$$

This evaluates to

$$\left(\left[X_{21}\right] + \left[X_{22}\right] + \left[X_{23}\right]\right)$$

and so at this point, the entire expression has been evaluated as

$$\left(\left[X_{11}\right] + \left[X_{12}\right] + \left[X_{13}\right]\right) + \left(\left[X_{21}\right] + \left[X_{22}\right] + \left[X_{23}\right]\right)$$

If we continue the process for $i = 3$, we obtain

$$\begin{aligned}\left(\left[X_{11}\right] + \left[X_{12}\right] + \left[X_{13}\right]\right) \quad &+ \quad \left(\left[X_{21}\right] + \left[X_{22}\right] + \left[X_{23}\right]\right) \\ &+ \quad \left(\left[X_{31}\right] + \left[X_{32}\right] + \left[X_{33}\right]\right)\end{aligned}$$

Both summation indices have reached their upper limits of 3, so the evaluation of the expression is now complete. To summarize, evaluation of expressions involving multiple summation signs involves the same rules as those that govern evaluation of single summation signs. Each summation operator governs everything in the expression to its right, including all summation signs.

2.4.2   The Odometer Metaphor

The mathematical interpretation of multiple summation operations is, in principle at least, straightforward. However, the notation would not be very useful if it always took us as much time to decode an expression as the previous example required. We want to devise some strategies that will allow us to decode the simpler expressions quickly without sacrificing accuracy. If an expression is challenging, one can always apply the strategy of the previous section and work through the expression slowly and carefully.

There are several metaphors which beginners find useful for decoding expressions involving several summations. The first such metaphor is one I call the "odometer metaphor." Consider a typical trip odometer on an automobile. As you drive down the road, the digits move from 0 to 9, with the rightmost digit moving first. As the rightmost digit reaches a "stop value" of 9, the next digit to the left increases by 1,

and the right digit resets to a "starting value" of zero. For example,

$$
\begin{array}{cc|c}
0 & 0 & 8 \\
0 & 0 & 9 \\
0 & 1 & 0
\end{array}
$$

If we examine the behavior of the indices i and j in the summation expression we evaluated in Example 2.4.1 we see that they behaved very similarly to the numbers on an odometer. The only differences are that the "start value" and "stop value" for the summation indices are 1 and 3, while the "start value" and "stop value" on the odometer are 0 and 9, respectively. Simply review the values taken on by i and j, as shown in the table below.

i	1	1	1	2	2	2	3	3	3
j	1	2	3	1	2	3	1	2	3

As j hits its stop value, it resets to 1 and i "clicks off" to the next value, just like an odometer. The odometer metaphor is particularly useful when you are evaluating simple summation expressions.

2.4.3   The Cartesian Product Metaphor

Many of my students have found another metaphor particularly useful for evaluating summation expressions.

> **Definition** The **Cartesian product** of two sets A and B is the set of all ordered pairs of values i and j where i is a member of set A and j is a member of set B.

The Cartesian Product notion can serve us well in interpreting summation expressions. Consider again the expression from Example 2.4.1. Another way of evaluating the expression is as follows:

(a) Consider each summation sign and record the range of subscript values it allows. (In this case, the ranges are $1 \leq i \leq 3$ and $1 \leq j \leq 3$.)

(b) Remove the summation signs from the expression (but leave appropriate paren-theses) to yield a new expression S. Call this the "reduced expression."

(c) Evaluate and sum expression S for all combinations of the subscript indices (i and j in this case) such that i is in the range for the first index and j is in the range for the second index.

In other words, the summation expression is the sum of the versions of the reduced expression produced by the Cartesian product of the ranges. In the following example, we see how this approach works with the expression from Example 2.4.1.

Example 2.4.2 (Applying the Cartesian Product Metaphor) Evaluate

$$\sum_{i=1}^{3}\sum_{j=1}^{3} X_{ij}$$

using the "Cartesian product" approach.

Solution First create the "reduced expression" by installing the parentheses but removing the summation signs. One obtains

$$\left(\left[X_{ij}\right]\right)$$

Now simply evaluate this expression for all combinations of integer subscript values that satisfy $1 \le i \le 3$ and $1 \le j \le 3$. The result is

$$\left(\left[X_{11}\right] + \left[X_{12}\right] + \left[X_{13}\right]\right) \quad + \quad \left(\left[X_{21}\right] + \left[X_{22}\right] + \left[X_{23}\right]\right)$$
$$+ \quad \left(\left[X_{31}\right] + \left[X_{32}\right] + \left[X_{33}\right]\right)$$

2.4.4   Applying Double Summations to Rectangular Arrays

One of the most common applications of double summation expressions is to rectangular arrays of numbers, in conjunction with double subscript notation. Consider, for example the scores of 3 students on 3 exams, as presented in Table 2.3.

Student	Exam 1	Exam 2	Exam 3
Joe	87	75	92
Fred	76	78	81
Mary	98	93	99

Table 2.3: Scores for 3 Students on 3 Exams

In the following examples, we study how double summation notation can be used to pick out specified subarrays from this array.

> **Example 2.4.3 (Summing Lower Trangular Elements of an Array)** Evaluate the following expression numerically, using the data in Table 2.3
>
> $$\sum_{i=1}^{3}\sum_{j=1}^{i} X_{ij}$$

> ***Solution*** This expression introduces a new wrinkle to summation notation. Note that the outer (leftmost) summation index controls the stop value of the inner (rightmost) summation operator. Consequently, the end value of the j index varies at different points in the evaluation. If we apply the "odometer metaphor" to the expression, we begin by setting $i = 1$ and stepping j through the values from 1 to i, or in this case from 1 to 1. For the first loop of the expression, we end up with simply
>
> $$\left(\left[X_{11}\right]\right)$$
>
> Next, we set $i = 2$ and step j through the values from 1 to i, or in this case from 1 to 2. Evaluating the expression, we now have
>
> $$\left(\left[X_{11}\right]\right) + \left(\left[X_{21} + X_{22}\right]\right)$$
>
> Finally, we set $i = 3$ and step j through the values from 1 to i, or in this case from 1 to 3. Evaluating the expression, we end up with the final result
>
> $$\left(\left[X_{11}\right]\right) + \left(\left[X_{21} + X_{22}\right]\right) + \left(\left[X_{31} + X_{32} + X_{33}\right]\right)$$
>
> For this expression, the "intersection metaphor" can provide us with considerable insight. The expression is telling us to sum X_{ij} values for $1 \leq i \leq 3$, i.e., for all available values of i, and for values of j that run from 1 to the current value of i, in other words for all values of j less than or equal to i. If we remember that i is the row subscript and j the column subscript, it becomes apparent that the notation is telling us to sum all the elements for which the row number is greater than or equal to the column number.

In Table 2.4 we show the data matrix with the lower triangular elements that are summed in boldface.

The elements in boldface are sometimes called the "lower triangular" elements of the data array. They form a triangular shape.

Student	Exam 1	Exam 2	Exam 3
Joe	**87**	75	92
Fred	**76**	**78**	81
Mary	**98**	**93**	**99**

Table 2.4: Lower Triangular Elements

Example 2.4.4 (Summing Upper Triangular Elements of an Array)

Suppose we define the upper triangular elements of a data array to be those elements for which i is less than or equal to j. Write a double summation expression to sum the upper triangular elements of the data in Table 2.3. Try not to look at the solution to this example until you have given it a good try.

Solution The expression we seek is

$$\sum_{j=1}^{3} \sum_{i=1}^{j} X_{ij}$$

You should be able to verify for yourself that this expression sums the appropriate values.

This example has a surprising number of subtle lessons to teach us. In my experience, many students find the solution to this example difficult to achieve, and the above question is typical. After all, the i comes before j in the subscripting scheme. Should not the summation signs follow the same order? I often ask such students "Where did you first get the idea that the summation indices have to follow a particular ordering?" Many students have this misconception. It apparently stems from the way students naturally incorporate mathematical ideas that are taught primarily "by example." Many students have only been exposed to examples where, i comes before j both in the summation indices and in the subscripting system. Ultimately, they form an abstraction, or "mental set," that implicitly restricts the possibilities of the notation for them. They see limitations where none exist. Along these lines, there is another valid solution to the preceding problem. It is

$$\sum_{i=1}^{3} \sum_{j=1}^{i} X_{ji}$$

Notice how changing the position of the j and the i subscripts has changed the result.

Again, many students will have difficulty seeing or understanding this solution to the problem, because somewhere along the line they developed the restrictive belief that the i subscript should always precede the j subscript. Remember, it is the position of the subscript that determines whether it is referring to a row or a column in a rectangular array. In this case, by switching the position of the i and j, we made i refer to a column and j refer to a row.

Example 2.4.5 (Calculating the Trace of a Matrix) The trace of a square matrix is the sum of the elements X_{ij} for which $i = j$. For the data in Table ??, compute the trace of the array.

Solution The first solution that comes to mind is something like the following:

$$\sum_{i=1}^{3} \sum_{j=i}^{i} X_{ij}$$

Example 2.4.6 (An Easier Way) However, there is an alternative solution to Example 2.4.5 that is much simpler, but less obvious. It requires only a single summation sign. Can you deduce this solution?

Solution The simpler solution often eludes the beginning student, again because of the mental set problem we discussed above. This solution is

$$\sum_{i=1}^{3} X_{ii}$$

If you failed to see this solution, it was probably because you were assuming, implicitly at least, that the two symbols for subscripts of X had to be different. We are so used to seeing symbols like X_{ij} that it is quite natural to adopt this assumption after a while. Of course, this assumption is incorrect. These two symbols are merely placeholders. They can, indeed, be functions of indices. They need not be indices themselves, as we see in the next example.

Example 2.4.7 (Summing Elements on a Subdiagonal Strip of an Array)
Write a summation expression, using only a single summation sign, to sum
the elements highlighted in boldface in the data set in Table 2.5.

Solution The numbers in boldface occupy the only positions in the array
where the row subscript is exactly 1 greater than the column subscript. Hence,
one answer is

$$\sum_{i=2}^{3} X_{i,i-1}$$

Note how, in this case, one subscript is actually a function of the other.

4	4	5
3	5	23
6	**1**	7

Table 2.5: Summing Elements on a Subdiagonal Strip

2.5 Variations in Summation Notation

So far, we have been examining "standard" summation notation. There are many
variations you will observe in different books, or in different sections of the same
book. Here we shall discuss just a few of these variations.

2.5.1 "Reduced" Notation

In many cases, aspects of the notation remain constant throughout a derivation, and
repeating them introduces redundancy, and makes the derivation harder to read. In
such situations, it is quite common to "reduce" the notation by removing redundant
aspects. Here are some examples

Example 2.5.1 (Reduced Notation) Instead of

$$\sum_{i=1}^{n} X_i$$

write

$$\sum_{i} X_i$$

or even

$$\sum X$$

Since you are summing all the elements of the X array, and this is usually implicit in the context of the discussion, why add unnecessary visual complication? This simplified notation is seen frequently in textbooks.

In a similar vein, you will often see something like this

$$\sum_i \sum_j X_{ij}$$

instead of

$$\sum_{i=1}^n \sum_{j=1}^n X_{ij}$$

2.5.2   Single Summation Operators with a Subscript Inclusion Rule

In Example 2.4.4, we used notation like this

$$\sum_{j=1}^3 \sum_{i=1}^j X_{ij}$$

to sum the upper triangular elements of an array. Many books will economize the notation by using a single summation sign, and placing below it an "inclusion rule" for the set of i, j subscript pairs to be summed. The above expression says, in effect, "sum the elements for which j is greater than or equal to i." The "economized" notation for saying the same thing is

$$\sum_{j \geq i} X_{ij}$$

2.6   Bar-Dot Notation

In this section, we examine a notation that is commonly used to simplify expressions in the Analysis of Variance.

2.6.1   Computing Row or Column Sums in a Rectangular Array

Consider the following rectangular array of data.

$$\begin{bmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{bmatrix}$$

The following expression will sum the elements in the first row.

$$\sum_{j=1}^{3} X_{1j}$$

Notice how the expression says, in effect. "Stay within the first row, and step across the columns, summing as you go."

Suppose you wished to sum the elements in the third column. You would use

$$\sum_{i=1}^{3} X_{i3}$$

Computing a row or column sum is an operation that is repeated many times in routine Analysis of Variance calculations. Looking back at the two preceding examples, we are struck by the visual inefficiency of basic summation notation for conveying this simple repetitive operation.

2.6.2   Dot Notation

Dot notation provides a simpler way of conveying the simple notion of summing across all observations in a particular direction. We introduce the notation here in the context of two dimensional arrays only. We will return to the notation when we get to the analysis of variance.

Definition (Dot Notation) Given a rectangular data array with R rows and C columns. We define

$$X_{\bullet j} = \sum_{i=1}^{R} X_{ij}$$

and

$$X_{i\bullet} = \sum_{j=1}^{C} X_{ij}$$

Dot notation is particularly easy to read, once you get the hang of it. For example, the "third row sum" is $X_{3\bullet}$. The "second column sum" is $X_{\bullet2}$. What makes dot notation particularly useful is the way that it generalizes effortlessly to a situation that causes complications for summation notation, namely the case of missing data. Consider the data array in 2.6. Suppose this array represented 3 groups, with groups represented in columns. There are three numbers in the first and third group, but only two numbers in the second group. Suppose we wished to convey an operation that involved computing the 3 column sums for these data. In this case, the previous definition that assumes there are the same numbers of scores in each column cannot be used, because the number of numbers in a column is not a constant. Instead, we need a notational device for conveying how many numbers there are in each column. One common solution is to use the number n_j to stand for the number of numbers in the jth column. In that case, we can still use the notation $X_{\bullet j}$ to stand for the jth column sum. Here we simply revise the definition as

$$X_{\bullet j} = \sum_{i=1}^{n_j} X_{ij}$$

With this revised definition, the dot notation achieves new flexibility. One may use $X_{\bullet j}$ to stand for the jth column sum *regardless of the number of numbers in a particular column.* It simply means the sum of however many numbers there are in the column.

Dot notation generalizes to situations where there are more than two subscript indices. In these more complex situations, which we will deal with when we begin discussing the Analysis of Variance, a dot in place of a subscript indicates that all values of that subscript have been summed over.

Group 1	Group 2	Group 3
4	4	5
3	6	23
6		7

Table 2.6: Hypothetical Data for 3 Groups of Unequal Size

2.6.3   Bar-Dot Notation

Besides computing row and column sums, we frequently compute row and column means in the context of analysis of variance and other statistical operations. A simple

addition to dot notation allows us to convey this very efficiently.

Definition (Bar Notation) If a bar is placed over a dot notation expression, it means "divide by the number of numbers that were summed in the dot expression." For example, if there are n_j numbers in the jth column, we have

$$\overline{X}_{\bullet j} = \frac{1}{n_j} \sum_{i=1}^{n_j} X_{ij}$$

Bar dot notation can be used in situations where there are more or less than two dimensions in the array. For example, suppose we have only one list. The sum of the numbers is $X_{\bullet}$, the mean of the numbers $\overline{X}_{\bullet}$.

Example 2.6.1 (Column Means)   Suppose you have a rectangular data array, and you wish to compute the mean of the scores in the first column. Write the expression in bar-dot notation.

Solution
$$\overline{X}_{\bullet 1}$$

Example 2.6.2 (Averaging Row Means)   Again suppose a rectangular data array. You wish to compute the average of the means of the first 3 rows. Write the expression using a summation operator and bar-dot notation.

Solution
$$\frac{1}{3} \sum_{i=1}^{3} \overline{X}_{i\bullet}$$