A Primer on Summation Notation

George H Olson, Ph. D.
Doctoral Program in Educational Leadership
Appalachian State University
Spring 2010

Summation Operator

The summation operator (∑) {Greek letter, capital *sigma*} is an instruction to sum over a series of values. For instance, if we have the set of values for the variable, $X = \{X_1, X_2, X_3, X_4, X_5\}$, then

$$\sum_{i=1}^{n=5} X_i = X_1 + X_2 + X_3 + X_4 + X_5$$

Literally, the expression, $\sum_{i=1}^{n=5} X_i$, says: beginning with i=1 and ending with i=5, sum over the *variables* X_i. As an example, let

$X_i = 8$, $X_2 = 10$, $X_3 = 11$, $X_4 = =15$, $X_5 = 16$.

Then n = 5 {the number of *cases*}, and

$$\sum_{i=1}^{n=5} X_i = 8 + 10 + 11 + 15 + 16$$
$$= 60.$$

In many contexts, it is clear that the summation is over all cases and we do not need the superscript over the summation operator. Furthermore, in most contexts it is assumed that the summation begins with i = 1. Hence, the notation, $\sum_i X_i$ is taken to imply $\sum_{i=1}^{n=5} X_i$. In most situations, where the variable has only one subscript, as in X_i, the subscript can be omitted. In these situations, $\sum X$ implies $\sum_{i=1}^{n=5} X_i$.

In other contexts, the variable X may have more than one subscript, e.g., X_{ij}. This occurs, for instance, when individuals belong to two or more subgroupings or cross-classifications. We might have a situation as shown below in Table 1.

Table 1

Group 1	Group 2	Group 3
$X_{11}, X_{21}, X_{31}, X_{41}$	$X_{12}, X_{22}, X_{32}, X_{42}, X_{52}, X_{62}$	$X_{13}, X_{23}, X_{33}, X_{43}, X_{53}$

Here we have three groups, each with a different number of cases. We denote the *i*th case in the *j*th group with the symbol, X_{ij}. To sum all the cases, over all three groups, we would use the following, double summation operator,

$$\sum_{j}^{3}\sum_{i}^{n_j} X_{ij} \,,$$

which instructs us to sum over the three groups (j=1, 2, and 3) and, within each group, sum over the number of cases in the group (i=1, 2, 3, 4 for Group 1; i=1, 2, 3, 4, 5, 6 for Group 2; i=1, 2, 3, 4, 5 for Group 3). For simplicity, we often write the summation expression as,

$$\sum\sum X_{ij} \,,$$

where it is assumed that we are to sum over all groups and all cases within each group. For example, lets substitute the following numbers for the symbolic values given above.

Table 2

Group 1	Group 2	Group 3
10, 8, 12, 13	6, 11, 8, 10, 8, 12	14, 6, 6, 10, 9

Then,

$$\sum\sum X_{ij} = \sum\left[\sum X_{ij}\right] = [10+8+12+13]$$

$$+[6+11+8+10+8+12]$$

$$+[14+6+6+10+9]$$

$$= 43+55+45$$

$$= 143.$$

A more complex situation occurs when cases are grouped into cross-classifications. Table 3 represents a situation where cases are cross-classified by sex and age-category.

Table 3

		Age Category		
		Child	Adolescent	Adult
Sex	Male	X_{111}, X_{211}, X_{311}, X_{411}	X_{112}, X_{212}, X_{312}, X_{412}	X_{113}, X_{213}, X_{313}, X_{413}
	Female	X_{121}, X_{221}, X_{321}, X_{421}	X_{122}, X_{222}, X_{322}, X_{422}	X_{123}, X_{223}, X_{323}, X_{423}

In the table the notation, X_{ijk}, denotes the *i*th individual in the *j*th row (Sex category) and *k*th column (Age Category). Hence, X_{423} is the last case in the Adult-Female cell. To indicate summation over all the cases in the table, we would use the notation,

$$\sum_k \sum_j \sum_i X_{ijk} ,$$

where it is assumed that the summation is over all *N* cases, *i*, over all *J* rows, *j*, and all *K* columns, *k*.

Dot Notation

It is often easier to denote aggregates using dot notation. In an expression such as

$$X_{\bullet} = \sum_i X_i ,$$

the dot (**·**) represents *aggregation* or summation over the missing (dotted) subscript. For instance, in the example given earlier, where we had $\sum_{i=1}^{n=5} X_i$ = 60, we could simply write

$X_{\bullet}$ = 60 {read X-dot = 60}. Using dot notation we would represent the cell aggregates in the Sex by Age Category table above as shown in Table 4.

Table 4

		Age Category		
		Child	Adolescent	Adult
Sex	Male	$X_{\bullet 11}$	$X_{\bullet 12}$	$X_{\bullet 13}$
	Female	$X_{\bullet 21}$	$X_{\bullet 22}$	$X_{\bullet 23}$

where, for the Male-Adolescent cell, $X_{\cdot 12} = \sum_i X_{i12}$. If we wanted to denote the sum of the values for all the males, we would write, $X_{\cdot 1 \cdot}$; for all females, $X_{\cdot 2 \cdot}$. Similarly, the sum of the values for all children is, $X_{\cdot\cdot 1}$; all adolescents, $X_{\cdot\cdot 2}$; and all adults, $X_{\cdot\cdot 3}$ And the sum of the values for all the cases, $X_{\cdot\cdot\cdot}$ Note that

$$X_{\cdot 1 \cdot} = \sum_k \sum_i X_{i1k} \,,$$

$$X_{\cdot\cdot 1} = \sum_j \sum_i X_{ij1} \text{, and}$$

$$X_{\cdot\cdot\cdot} = \sum_k \sum_j \sum_i X_{ijk} \,.$$

Rules of summation

Summation of a constant. Let c be some constant value. Then,

$$\sum_{i=1}^{N} c = Nc \,.$$

In other words, summing a constant N times is the same as multiplying the constant by N. Hence, if $c = 5$, then

$$\sum_{i=1}^{12} c = 12 \times 5 = 60 \,.$$

This rule can be extended to double summation. Thus,

$$\sum \sum c = \sum_j^J \left[\sum_i^{n_j} c \right] = \sum_j^J n_j c \,.$$

As an example, consider the situation involving the three groups given earlier in Table 2. If all cases, in all groups, have the constant value, 10, then

$$\sum_j^3 \sum_i^{n_j} 10 = \sum_j^3 \left[\sum_i^{n_j} 10 \right] = [10+10+10+10]$$

$$+ \; [10+10+10+10+10+10]$$

$$+ \; [10+10+10+10+10]$$

$= (4\times10) + (6\times10) + (5\times10)$

$= 15\times10$

$= 150.$

Multiplication by a constant. If all the values of a variable, *X*, are multiplied by the same constant, *c*, then,

$$\sum_{i}^{N} cX_i = c\left(\sum_{i}^{N} X_i\right) = c\sum_{i}^{N} X_i \ .$$

For example, let the set of 5 values of the variable *X* be {3, 9, 5, 7, 10}, assume that each is multiplied by the constant, 2. Then,

$$\sum_{i} 2X_i = 2(3)+2(9)+2(5)+2(7)+2(10)$$

$$= 2(3+9+5+7+10)$$

$$= 34$$

$$= 2\sum_{i} X_i$$

$$= cX_{..}$$

Again, this can be expanded to multiple summations:

$$\sum_{j}\sum_{i} cX_{ij} = c\sum_{j}\sum_{i} X_{ij} = cX_{..},$$

$$\sum_{k}\sum_{j}\sum_{i} cX_{ijk} = c\sum_{k}\sum_{j}\sum_{i} X_{ij} = cX_{...},$$

and so on.

As an example, again consider the three group situation given earlier in Table 2. If all cases were multiplied by 5, then

$$\sum_{j}^{3}\sum_{i}^{n_j} 5(X_{ij}) = [(5\times10) + (5\times8) + (5\times12) + (5\times13)]$$

$$+[(5\times6) + (5\times11) + (5\times8) + (5\times10) + (5\times8) + (5\times12)]$$

$$+[(5\times14) + (5\times6) + (5\times6) + (5\times10) + (5\times9)]$$

$$= 5(43+55+45)$$

$$=5\left(\sum_{j}^{3}\sum_{i}^{n_j}(X_{ij})\right)$$

$$= 10(143)$$

$$= 715.$$

In some situations, the values in different groups are multiplied by different constants. For instance, suppose the values in Group 1 (in the example just given) were multiplied by the constant c_1, the values in Group 2 by c_2, and the values in group 3 by c_3. Then, the sum of all the cases would be given by,

$$\sum_{i} c_i X_i .$$

In this situation, it is necessary that the constants remain within the summation operator. Now, suppose that in the Sex by Age Category example given earlier, we have the values as shown below in Table 5.

Table 5

		Age Category		
		Child	Adolescent	Adult
Sex	Male	2, 3, 5, 4	5, 1, 1, 3	2, 2, 3, 0
	Female	1, 3, 5, 2	2, 0, 3, 1	5, 3, 4, 3

If all the male values are multiplied by the constant, 5, and all the female values are multiplied by the constant, 10. Then, letting c_1 = 5 and c_2 = 10, the sum over all cases, rows, and columns is given by,

$$\sum_{k}\sum_{j}\sum_{i} c_j X_{ijk} = \sum_{k}\sum_{j} c_j\left(\sum_{i} X_{ijk}\right)$$

$$= 5(14+10+7) + 10(11+6+15)$$

$= 5(31) + 10(32)$

$= 475.$

Note that the right-hand side of the summation notation above could have been written as

$$\sum_k \sum_j c_j \sum_i X_{ijk} ,$$

without the parentheses.

Order of operations. The order of operations, when using summation operators, is the same as that for arithmetic. That is, operations involving the values in a summation operation are indicated by mathematical punctuation. When indicated by punctuation, operations are to be carried out on values prior to summation. Some examples should suffice.

$$\sum_i X_i^2 = (X_1^2 + X_2^2 + X_3^2 + \cdots X_N^2) ,$$

$$\sum_i \sqrt{X_i} = \sqrt{X_1} + \sqrt{X_2} + \cdots + \sqrt{X_N} , \text{ and}$$

$$\sum_i \log X_i = \log X_1 + \log X_2 + \cdots + \log X_N .$$

On the other hand,

$$\left(\sum_i X_i \right)^2 = (X_1 + X_2 + \cdots + X_3)^2 ,$$

$$\sqrt{\sum_i X_i} = \sqrt{X_1} + \sqrt{X_2} + \cdots + \sqrt{X_N} , \text{ and}$$

$$\log\left(\sum_i X_i \right) = \log(X_1 + X_2 + \cdots X_N) .$$

Note that these rules apply even when we have multiple summations. For example, letting capital *J* represent the number of groups,

$$\sum_{j=1}^{J}\left(\sum_{i} X_i\right)^2 = \sum_{j}\left(X_{1j} + X_{2j} + \cdots + X_{nj}\right)^2 = \sum_{j} X_{\bullet j}^2 .$$

In words, within each group, *j*: sum the values over cases, *i*, then square the sum. After squaring the sums within all of the *J* groups, sum the squared sums over groups. For example, consider the data given in Table 2, earlier.

$$\sum_{j=1}^{J}\left(\sum_{i} X_i\right)^2 = 43^2 + 55^2 + 45^2 = 6899.$$

Distributive rule of summation. The summation operator is distributive when the value being operated upon is itself a sum (or difference). For example,

$$\sum_{i}\left(X_i + Y_i\right) = \sum_{i} X_i + \sum_{i} Y_i \ ,$$

$$\begin{aligned}\sum_{j}^{J}\sum_{i}\left(X_{ij} - \bar{X}_{\bullet\bullet}\right)^2 &= \sum_{j}^{J}\sum_{i}\left(X_{ij}^2 - 2\bar{X}_{\bullet\bullet}X_{ij} + \bar{X}_{\bullet\bullet}^2\right) \\ &= \sum_{j}^{J}\sum_{i} X_{ij}^2 - 2\bar{X}_{\bullet\bullet}\sum_{j}^{J}\sum_{i} X_{ij} + \sum_{j}^{J}\sum_{i}\bar{X}_{\bullet\bullet}^2 \\ &= \sum_{j}^{J}\sum_{i} X_{ij}^2 - 2\bar{X}_{\bullet\bullet}\sum_{j}^{J}\sum_{i} X_{ij} + \bar{X}_{\bullet\bullet}^2\sum_{j}^{J} n_j\end{aligned}$$

Note that the last step in this equation follows from the rule pertaining to summation over a constant, given earlier. Hence,

$$\begin{aligned}\sum_{j}^{J}\sum_{i} X_{\bullet\bullet}^2 &= \left(\sum_{i}^{n_1} X_{\bullet\bullet}^2 + \sum_{i}^{n_2} X_{\bullet\bullet}^2 + \cdots + \sum_{i}^{n_J} X_{\bullet\bullet}^2\right) \\ &= \left(n_1 X_{\bullet\bullet}^2 + n_2 X_{\bullet\bullet}^2 + \cdots + n_J X_{\bullet\bullet}^2\right) \\ &= \sum_{j}^{J} n_j X_{\bullet\bullet}^2 \\ &= X_{\bullet\bullet}^2 \sum_{j}^{J} n_j\end{aligned}$$

Summation involving two or more variables. If each case has values on two variables, X_i and Y_i, say, then,

$$\sum_{i}^{N} X_i Y_i \neq \left(\sum_i X_i\right)\left(\sum_i Y_i\right)$$

Instead,

$$\sum_i X_i Y_i = (X_1 Y_1 + X_2 Y_2 + \cdots + X_N Y_N).$$

These rules can be extended, quite logically, to situations involving more than two variables.

Exercises

In each of the following exercises, use symbols to extend the expression given in the summation operation. For example,

$$\sum_{i=3}^{7} X_i = (X_3 + X_4 + X_5 + X_6 + X_7)$$

1. $\sum_{l=2}^{6} Y_l$

2. $\sum_{i=1}^{3} X_i^2 \sum_{2}^{4} 10$

3. $\sum_{j}^{3} \sum_{i}^{2} \left(X_{ij}^2 + Y_j\right)$

4. $\sum_{i}^{5} 5\left(X_i + Y_i + Z_i\right)$

5. $\sum_{j=5}^{7} \left(\sum_{i=1}^{3} X_{ij}\right)^2$

6. $\sum_{j}^{3} n_j \left(\sum_{i}^{4} X_{ij}\right)^2$

Reduce the following extended expressions to an expression using the summation operator. For example,

$$(X_3 + X_4 + X_5 + X_6 + X_7)$$

would be written as

$$\sum_{i=3}^{7} X_i$$

7. $X_1Y_1 + X_2Y_2 + X_3Y_3 + X_4Y_4 + X_5Y_5 + X_6Y_6$

8. $X_1Y_1 + X_2Y_2 + \cdots + X_NY_N$

9. $(X_5 + Y_5)^2 + (X_6 + Y_6)^2 + (X_7 + Y_7)^2 + \cdots + (X_{10} + Y_{10})^2$

10. $\left(X_1 + X_2 + X_3 + X_4 + X_5\right)^2\left(Y_1^2 + Y_2^2 + Y_3^2 + Y_4^2 + Y_5^2\right)$