Chapter One

The summation sign and the rules of summation

1. The summation Sign

It is frequently necessary in statistical and psychometric calculations to take the sum of a number of values. The symbol used to indicate this operation of adding up a group of numbers is a capital Greek Sigma - $\sum$.

However, the instruction 'to take the sum of' is rather vague without an indication of what it is that is to be summed. It is necessary to have a system of notation to specify precisely which values are to be summed. Let us suppose that we have a set of four scores:

$$2, 4, 6, 8,$$

and that we let X be a general symbol for any one of these scores. The set of scores now consists of four *X's* which are 2, 4, 6 and 8. If we now assign a subscript to each of the *X's*, we can assign an X which a given subscript to each score thus:

$$X_1 = 2; \quad X_2 = 4; \quad X_3 = 6; \quad X_4 = 8.$$

In the case of four scores the subscripts will naturally run from 1 to 4, but there are usually more than four scores involved, and so it is desirable to have a generalised notation, so that we can apply the system to any group of scores. A general symbol for the number of scores is N. Any collection of scores will consist of N scores. If we want a general reference to a single score without specifying exactly which we can use the subscript i. Thus X is the *i*th score. Consider the following set of scores:

$$X_1 = 6; \quad X_2 = 7 \quad X_3 = 8; \quad X_4 = 9; \quad X_5 = 10.$$

What is the value of X_i when $i = N$? The answer is 10. There are five scores, therefore $N = 5$. X_i when $i = N$ must be X_5, which is the symbol for the fifth score, which is 10. In similar fashion the value of X_1 where $i = 2$ is 7; where $i = 1$ it is 6 and so on.

These symbols are often used in connection with the summation sign to indicate exactly which scores are to be summed.

For example:

$$\sum_{i=1}^{N} X_1 = \text{take the sum of all scores from } X_1 \, to \, X_N$$

i.e. $X_1 + X_2 + X_3 + \ldots X_N$

The symbols above and below the summation sign are called the limits of the summation. The value of i under the summation sign tells you where to start the addition, and the values above the summation sign tells you where to stop. The starting and stopping places can be anywhere in the set of scores.

$$\sum_{i=2}^{4} X_i = \text{take the sum of scores 2 to 4}$$

i.e. $X_2 + X_3 + X_4$

Often when there is no danger of confusion the symbol $\sum X$
Is used without subscripts or limits. This should be read as though it was:

$$\sum_{i=1}^{N} X_i = \sum X = \text{Take the sum of all the numbers}$$

i.e. $X_1 + X_2 + X_3 + \ldots X_N$

$\sum X$ will be used frequently in this book, but always check to see if there are any subscripts with it.

Sometimes more than one summation sign is used. Suppose that our scores are classified into groups, and let us use the symbol J for the number of groups, and n_j for the number of cases in the jth group. A particular X_i will now be found in the jth group, so we can put a double subscript under the X, to make it clear that we are talking about the ith score in the jth group, hence X_{ij}. Suppose now that we want to find the total score for one of the groups. We will need to sum all of the X_{ij} in that group and there will be $n_j X_{ij}$'s. So the instruction to sum all the scores in a group can be written:

$$\sum_{i=1}^{n_j} X_{ij} = \text{take the sum of all the scores in a group.}$$

If we now want to sum these totals of groups to find the grand total for all scores we can write:

$$\sum_{j=1}^{J}\sum_{i=1}^{n_j} X_{ij} = \quad \text{find the groups' totals}$$

and add them all together.

It will not be necessary in the following chapters to use more than double summation, but as many summation signs as necessary may be used.

The use of brackets is also important. If confronted with instructions inside a bracket always follow these instructions before following the instructions outside the brackets.

$$\sum_{i=1}^{N} X_1^2 + X_2^2 + X_3^2 + \ldots X_N^2$$

but

$$\left(\sum_{i=1}^{N} X_i\right)^2 = \left(X_1 + X_2 + X_3 + \ldots X_N\right)^2$$

which is not the same as $\sum_{i=1}^{N} X_{i/}^2$

Similarly

$$\sum_{j=1}^{J}\sum_{i=1}^{n_j} X_{ij}^2 =$$ Square every number in a group and find the total of the squared numbers, repeat for all groups and then sum the group totals.

While

$$\sum_{j=1}^{J}\left(\sum_{i=1}^{n_j} X_{ij}\right)^2 =$$ Find the total score for a group, square this total, repeat for each group, then sum the squared group totals.

So

$$\sum_{j=1}^{J}\sum_{i=1}^{n_j} X_{ij}^2 \neq \sum_{j=1}^{J}\left(\sum_{i=1}^{n_j} X_{ij}\right)^2$$

Problems

A. Given the following set of scores

$X_1 = 4$; $X_2 = 5$; $X_3 = 6$; $X_4 = 7$; $X_5 = 8$; $X_6 = 9$; $X_7 = 10$;
$X_8 = 11$; $X_9 = 12$; $X_{10} = 13$; $X_{11} = 14$; $X_{12} = 15$.

(1) What is the value of N?
(2) What is the value of X_i when I = 6?
(3) What is the value of X_i when I = N?

(4) What are the values of:

(a) $\sum_{i=1}^{4} X_i$ (b) $\sum_{i=10}^{N} X_i$ (c) $\sum_{i=5}^{8} X_i$

B. In the following groups indicate which values will be the same as another.

(1) (a) $\left(\sum_{i=1}^{N} X^2\right)$ (b) $\left(\sum_{i=1}^{N} X\right)^2$ (c) $\sum_{i=1}^{N} X^2$

(2) (a) $\left(\sum_{j=1}^{J}\sum_{i=1}^{n_j} X_{ij}\right)^2$ (b) $\sum_{j=1}^{J}\left(\sum_{i=1}^{n_j} X_{ij}\right)^2$

(c) $\sum_{j=1}^{J}\sum_{i=1}^{n_j} X_{ij}^2$ (d) $\sum_{j=1}^{J}\left(\sum_{i=1}^{n_j} X_{ij}^2\right)$

Answers

A. (1) 12; (2) 9; (3) 15; (4) (a) 22; (b) 42; (c) 38.
B. (1) (a) and (c); (2) (c) and (d).

2. Some Rules of Summation

On several occasions later in this book, proofs will be presented which require knowledge of some of the rules of summation. In this section the rules will be stated and proofs of the rules provided. The proofs are easy and well within the competence of any reader of this book. The reader is therefore urged to read the proof as well as the rule.

Summation Rule 1: *The sum of the sums of two or more Variables is equal to the sum of their Summations.*

i.e. $$\sum_{i=1}^{N}\left(X_i+Y_i+Z_i\right)=\sum_{i=1}^{N}X_i+\sum_{i=1}^{N}Y_i+\sum_{i=1}^{N}Z_i$$

Proof

(1) $$\sum_{i=1}^{N}\left(X_i+Y_i+Z_i\right)=\left(X_1+Y_1+Z_1\right)$$

$$+\left(X_2+Y_2+Z_2\right)+\left(X_3+Y_3+Z_3\right)$$

$$+\ldots\left(X_N+Y_N+Z_N\right)$$

(2) Removing the brackets leaves

$X_1 + Y_1 + Z_1 + X_2 + Y_2 + Z_2 + X_3 + Y_3 + Z_3 \ldots + X_N + Y_N + Z_N$

(3) This equals all the *X*'s plus all of the *Y*'s plus all of the *Z*'s.

(4) Which is the same as

$$\sum_{i=1}^{N}X_i+\sum_{i=1}^{N}Y_1+\sum_{i=1}^{N}Z_i$$

Summation Rule 2: *The sum of a constant times the values of a variable is equal to the constant times the sum of the variable.*

i.e. $$\sum_{i=1}^{N}(cX_1) = c\sum_{i=1}^{N} X_i$$ where *c* is a constant.

Proof

(1) $$\sum_{i=1}^{N} cX_1 = cX_1 + cX_2 + cX_3 \ldots + cX_N$$

(2) As everything is multiplied by *c* this can be written as:
$c(X_1 + X_2 + X_3 \ldots X_N)$

(3) The term inside the brackets is $\sum_{i=1}^{N} X_i$ therefore

$$C(X_1 + X_2 + X_3 \ldots + X_N) + c\sum_{i=1}^{N} X_i$$

(4) The brackets can be removed to give $c\sum_{i=1}^{N} X_i$

Summation Rule 3: *The sum of a constant taken N times is The constant times N.*

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

Proof

(1) $$\sum_{i=1}^{N} c = c + c + c \ldots c$$

(2) It can be seen that *N* constants are added together. This is the same as taking the constant *N* times which equals *Nc*.

Summation Rule 4: *The sum of the values of a variable plus A constant, is equal to the sum of the Values of the variable plus N times the constant.*

i.e. $$\sum_{i=1}^{N}(X_i + c) = \sum_{i=1}^{N} X_i + Nc$$

Proof

(1) $\sum_{i=1}^{N}(X_i + c) = \sum_{i=1}^{N} X_i + \sum_{i=1}^{N} c$ (by Summation Rule 1).

(2) $\sum_{i=1}^{N} c = Nc$ (by Summation Rule 3).

(3) So we obtain $\sum_{i=1}^{N} X_i + Nc$

Summation Rule 5: *The sum of the values of a variable minus a Constant is equal to the sum of the values Of the variable minus N times the constant.*

i.e. $$\sum_{i=1}^{N}(X_i - c) = \sum_{i=1}^{N} X_i - Nc$$

Proof

This can be left as an exercise.

Problems

A. Simplify the following expression:

$$\sum_{i=1}^{N}\left(X_i + Y_i - c - d\right)$$ where c and d are constants.

B. Check your answer using the following data:

Subject	*X Score*	*Y Score*
(1)	1	5
(2)	2	6
(3)	3	7
(4)	4	8

c = 2 d = 3

C. In the above example what is the value of

(a) $\sum_{i=1}^{N} d$ (b) $\sum_{i=1}^{N} cX_i$ (c) $\sum_{i=1}^{N}\left(X_i + Y_i\right)$?

D. Are the answers you gave to question C consistent with the rules of summation given above?

Answers

A. $\sum_{i=1}^{N} X_i + \sum_{i=1}^{N} Y_i - N(c+d)$

B. $\left(X_i + Y_i - c - d\right)$ for subject (1) is 1; for (2) is 3; for (3)is 5 and for (4) is 7. The sum of these is 16.

$$\sum_{i=1}^{N} X_i = 10; \sum_{i=1}^{N} Y_i = 26; N = 4; c + d = 5$$

Inserting these values in A gives 16 which is the same as for:

$$\sum_{i=1}^{N}\left(X_i + Y_i - c - d\right)$$

C. (a) 12. (b) 20. (c) 36.