LECTURE 2
STATISTICAL NOTATIONS

PSY2002

Hye Won Suk

NOTATION FOR A VARIABLE

- In statistics, usually scores for a particular variable are represented by the letter X.
  - X : midterm score (out of 100)
- A set of scores can be presented in a column that is headed by X.

X (midterm)
65
68
62
59
71

SUBSCRIPT NOTATION

- To denote a particular person's score on the variable, a subscript is often used.
  - X_i : the score on the variable X for the ith person

- $X_1 = 65$
- $X_2 = 68$
- $X_3 = 62$
- $X_4 = 59$
- $X_5 = 71$

X (midterm)
65
68
62
59
71

SUMMATION NOTATION

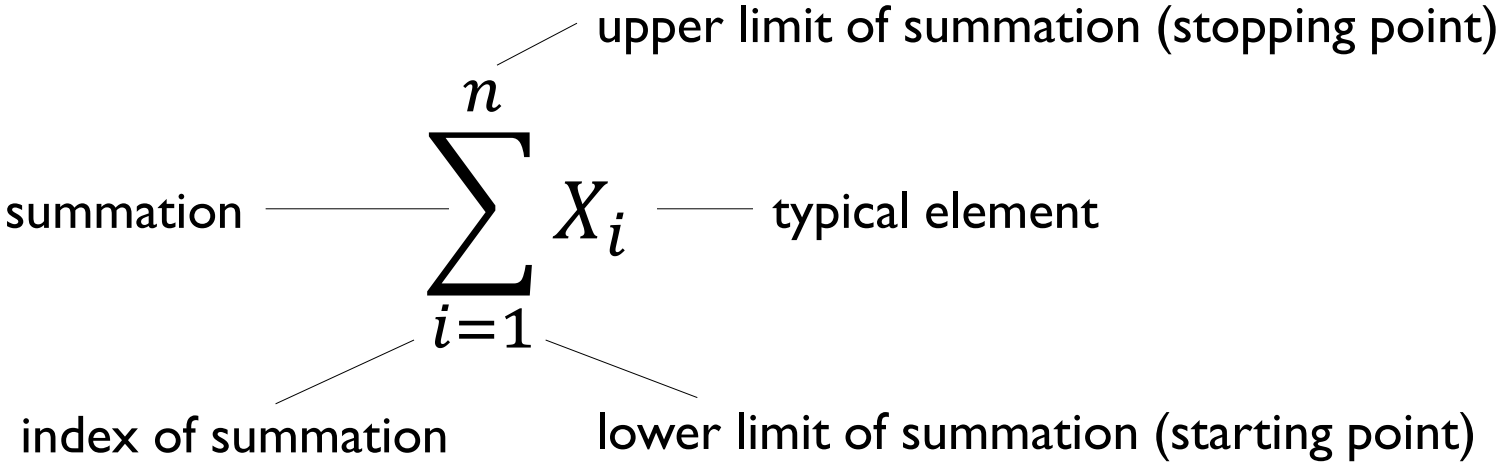
- Σ (read "*sigma*"), the Greek upper case letter S, is used to stand for summation.

- Σ instructs us to sum the elements of a sequence of numbers. A typical element of the sequence which is being summed appears to the right of the summation sign.

- The variable of summation is represented by an index which is placed beneath the summation sign. The index is often represented by i.

http://www.columbia.edu/itc/sipa/math/summation.html

SUMMATION NOTATION



$$\sum_{i=1}^{n} X_i = X_1 + X_2 + \cdots + X_n$$

http://www.columbia.edu/itc/sipa/math/summation.html

SUMMATION NOTATION

- Examples

X (midterm)
65
68
62
59
71

$$\sum_{i=1}^{3} X_i = X_1 + X_2 + X_3 = 65 + 68 + 62 = 195$$

$$\sum_{i=2}^{5} X_i = X_2 + X_3 + X_4 + X_5 = 68 + 62 + 59 + 71 = 260$$

A COUPLE OF VARIABLES

- When a couple of variables are measured for each person, the observed scores (or data) can be presented as two lists labeled X and Y.
  - X : midterm (out of 100)
  - Y : assignment (out of 180)

X (midterm)	Y (assignment)
65	150
68	144
62	132
59	125
71	174

MULTIPLE VARIABLES

- When there are multiple variables, the following double-subscript notation can be used.

  - X_1 :Variable 1, X_{1i} : The score of person i on Variable 1
  - X_2 :Variable 2, X_{2i} : The score of person i on Variable 2
  - X_3 :Variable 3, X_{3i} : The score of person i on Variable 3
  - X_4 :Variable 4, X_{4i} : The score of person i on Variable 4

  - And so on…
  - If there are more than 10 variables or 10 persons, a comma can be used to separate the two subscripts. E,g.,

$$X_{2,15} \quad X_{10,7}$$

ARITHMETIC OPERATIONS WITH Σ

- Sometimes, the summation process is included with several other arithmetic operations.

  - Order of arithmetic operations

  1. Any calculation contained within parentheses
  2. Squaring
  3. Multiplying and/or dividing
  4. Summation using the notation Σ
  5. Any other addition and/or subtraction

EXERCISE QUESTIONS

- Find the answer for each expression.

$$\sum_{i=1}^{5} 6 =$$

$$\sum_{i=1}^{10} 1 =$$

EXERCISE QUESTIONS

- For the following data, find the answer for each expression.

X	Y
5	1
8	4
6	2
3	5
7	3

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

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

EXERCISE QUESTIONS

- For the following data, find the answer for each expression.

X	Y
5	1
8	4
6	2
3	5
7	3

$$\sum_{i=1}^{5}(X_i + 3)$$

$$\sum_{i=1}^{5}(2X_i + 3)$$

EXERCISE QUESTIONS

- For the following data, find the answer for each expression.

X	Y
5	1
8	4
6	2
3	5
7	3

$$\sum_{i=1}^{5} X_i + 3$$

EXERCISE QUESTIONS

- For the following data, find the answer for each expression.

X	Y
5	1
8	4
6	2
3	5
7	3

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

$$\left(\sum_{i=1}^{5} X_i \right)^2$$

EXERCISE QUESTIONS

- For the following data, find the answer for each expression.

X	Y
5	1
8	4
6	2
3	5
7	3

$$\sum_{i=1}^{5}(2X_i + 3Y_i)$$

$$2\sum_{i=1}^{5} X_i + 3\sum_{i=1}^{5} Y_i$$

EXERCISE QUESTIONS

- For the following data, find the answer for each expression.

X	Y
5	1
8	4
6	2
3	5
7	3

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

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

EXERCISE QUESTIONS

- For the following data, find the answer for each expression.

X	Y
5	1
8	4
6	2
3	5
7	3

$$\sum_{i=1}^{5}(X_i - 6)^2$$

SUMMARY

- Notation for variables
- Subscript notation
- Summation notation
- Arithmetic operations with summation notation

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

$$\sum_{i=1}^{n} (aX_i + b) = a\sum_{i=1}^{n} X_i + nb$$

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

$$\sum_{i=1}^{n} (aX_i + bY_i) = a\sum_{i=1}^{n} X_i + b\sum_{i=1}^{n} Y_i$$

$$\sum_{i=1}^{n} X_i Y_i \neq \left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} Y_i\right)$$