Statistics Notes

Created by Mr. Francis Hung

Last updated: April 22, 2011

Law of summation

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

2. $\sum_{i=3}^{5} x_i = x_3 + x_4 + x_5$

3. $\sum_{i=1}^{n} x_{2i-1} = x_1 + x_3 + x_5 + \ldots + x_{2n-1}$

4. $\sum_{i=1}^{n} a = a + a + \ldots + a = na$

5. $\sum_{i=1}^{n} (ax_i + by_i) = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} y_i$

6. $\sum_{k=1}^{n} k = 1 + 2 + \ldots + n = \frac{n(n+1)}{2}$

7. $\sum_{k=1}^{n} k^2 = 1^2 + 2^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$

8. $\sum_{r=1}^{\infty} \frac{1}{3^{r-1}} = 1 + \frac{1}{3} + \frac{1}{9} + \cdots \text{to infinity} = \frac{1}{1-\frac{1}{3}} = \frac{3}{2}$

9. $\sum_{i=1}^{n}\sum_{j=1}^{m} x_{i,j} = x_{1,1} + x_{1,2} + \ldots + x_{1,m}$
$+ x_{2,1} + x_{2,2} + \ldots + x_{2,m}$
$+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots$
$+ x_{n,1} + x_{n,2} + \ldots + x_{n,m}$

10. $\sum_{i=1}^{n}\sum_{j=1}^{i} x_{i,j} = x_{1,1}$
$+ x_{2,1} + x_{2,2}$
$+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots$
$+ x_{n,1} + x_{n,2} + \ldots + x_{n,n}$

11. Let $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$, $\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}$

$$\sigma^2 = \frac{\sum_{i=1}^{n} \left(x_i^2 - 2\bar{x}x_i + \bar{x}^2\right)}{n}$$

$$= \frac{\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \bar{x}^2}{n}$$

$$= \frac{\sum_{i=1}^{n} x_i^2 - 2n\bar{x}\cdot\bar{x} + n\bar{x}^2}{n} = \frac{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}{n}$$

$$= \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2 = \text{square mean} - \text{mean square} \quad (\text{SM} - \text{MS})$$

Combined mean, combined variance and combined standard deviation

e.g.

	Class A	Class B
Number of students	40	30
Mean	$\overline{A} = 75$	$\overline{B} = 72$
Standard deviation	$\sigma_A = 12$	$\sigma_B = 10$

$\overline{A} = \frac{\sum a_i}{n}$ $\overline{B} = \frac{\sum b_i}{m}$

$75 = \frac{\sum a_i}{40}$ $72 = \frac{\sum b_i}{30}$

$\sum a_i =$ $\sum b_i =$

Combined mean $= \frac{\sum a_i + \sum b_i}{m+n}$

$=$

$\sigma_A^2 = SM - MS$ $\sigma_B^2 = \frac{\sum b_i^2}{m} - \overline{B}^2$

$\sigma_A^2 = \frac{\sum a_i^2}{n} - \overline{A}^2$

$12^2 = \frac{\sum a_i^2}{40} - 75^2$

$\sum a_i^2 =$ $\sum b_i^2 =$

Combined variance $\sigma^2 = \frac{\sum a_i^2 + \sum b_i^2}{m+n} -$ (combined mean)^2

$=$

Combined standard deviation =

Exercise 1 The means and standard deviations of two distributions of 100 and 150 measurements are 56, 54 and 24, 20 respectively. Find the mean and standard deviation of all 250 measurements taken as a group of measurements. ans. 54.8, 21.71

Exercise 2 A distribution consists of three parts, characterised as follows:

Parts	Number of items	mean	standard deviation
1	200	5	3
2	250	10	4
3	300	15	5

Find the mean and standard deviation of the whole distribution. ans. 10.67, 5.83

Corrections of means and standard deviation for mistakes

Example A set of 10 numbers gave a mean of 13 and a standard deviation 2. Later it was found that the number 12 in the set should be 21. Find the correct mean and correct standard deviation.

Solution: $n =$ _____, $\bar{x} =$ _____, $\sigma =$ _____

incorrect $\sum x_i =$

correct $\sum x_i =$ ( ) – ( ) + ( )

correct mean = _____

$$\sigma^2 = \frac{\sum x_i^2}{n} - \bar{x}^2$$

incorrect $\sum x_i^2 =$

correct $\sum x_i^2 =$ ( ) – ( ) + ( )

correct $\sigma^2 =$ ( ) – ( )^2

correct $\sigma = \sqrt{} =$

Exercise 1 Mean and standard deviation of 200 items are found to be 60 and 20. If at the time of calculations, two items are wrongly taken as 3 and 67 instead of 13 and17, find the correct mean and correct standard deviation. ans. 59.8, 20.09

Exercise 2 Given $n = 12$, $\bar{x} = 7$, $\sigma = 1.8$. If two numbers 4, 9 are deleted from the set, find the new mean and the new standard deviation. ans. 7.1, 1.61

An example of wrong question

The marks of the mathematics examination is given below:

Number of students $n = 40$

	Paper 1(120 marks)	Paper 2 (80 marks)
Mean	$\bar{x} = 75.0$	$\bar{y} = 52.2$
Standard deviation	$S_x = 25.2$	$S_y = 16.3$

Find the combined mean and the combined standard deviation.

Solution:

$\bar{x} = \frac{\sum x}{n}$ $\bar{y} = \frac{\sum y}{n}$

$\sum x =$ $\sum y =$

Combined mean $= \frac{\sum x + \sum y}{n} =$

$S_x^2 = \frac{\sum x^2}{n} - \bar{x}^2$ $S_y^2 = \frac{\sum y^2}{n} - \bar{y}^2$

$\sum x^2 =$ $\sum y^2 =$

Combined variance $= \frac{\sum x^2 + \sum y^2}{n} - (\quad)^2$

$=$

Combined standard deviation =

Explain what is wrong with the question.