DESCRIPTIVE STATS

Measures of Central Tendency:

Mean:

$$\frac{\Sigma x}{n} \quad \text{or} \quad \frac{\Sigma xf}{\Sigma f}$$

These are the two typical formulas used for calculating the mean of a set of data

↳ If is continuous and not discrete, the midpoint of the class will be used as the value of x to estimate a value for the mean

Assumed mean:

This method can also be used to calculate the mean (and variance) of a set of data using the following formula:

$$\bar{x} = a + \frac{\Sigma d}{n}$$

where a is any mean that you assumed
d is the deviation of that mean from a value in the data set
n is the total number of elements in the data set

For example:

1, 3, 5, 3, 10, 13 → set of data

Using the typical method:

$$\bar{x} = \frac{\Sigma x}{n} = \frac{1+3+5+3+10+13}{6} = \boxed{\underline{5.8\bar{3}}}$$

mean

The assumed mean formula can also be written as:

$$\bar{x} = a + \frac{\Sigma(x-a)}{n}$$

Using assumed mean method:

$$\bar{x} = a + \frac{\Sigma d}{n}$$

and we assume that mean is 7

$$\bar{x} = 7 + \frac{\Sigma d}{n} \qquad d = (x-a)$$

x	$x-a$	
1	1 − 7	= −6
3	3 − 7	= −4
5	5 − 7	= −2
3	3 − 7	= −4
10	10 − 7	= 3
13	13 − 7	= 6
		−7 = Σd

$$\bar{x} = 7 + \frac{-7}{6} = 7 - 1.16667 = \boxed{\underline{5.8\bar{3}}}$$

mean

Combined mean:

This is used to find the combined mean of two separate sets of data

$$\text{combined mean} = \frac{\Sigma x + \Sigma y}{n_x + n_y}$$

Measures of Variation

Just like how there are different approaches to calculating the mean depending on the information given, there are also corresponding ways to calculate the variance of the data, and hence, it's standard deviation.

↳ As a rule, remember: $(\text{standard deviation})^2 = \text{variance}$

In a normal set of data:

ie. 1, 5, 9, 10, 15 Mean = 8

$$\text{Variance} = \frac{\Sigma x^2}{n} - (\bar{x})^2 \quad \text{or} \quad \frac{\Sigma(x-\bar{x})^2}{n}$$

$$\frac{1 + 25 + 81 + 100 + 225}{5} - 64$$

$$\boxed{\sigma^2 = 22.4}$$

↳ Variance

$$\frac{(1-8)^2 + (5-8)^2 + (9-8)^2 + (10-8)^2 + (15-8)^2}{5}$$

$$= \frac{(-7)^2 + (-3)^2 + (1)^2 + (2)^2 + (7)^2}{5}$$

$$= \frac{49 + 9 + 1 + 4 + 49}{5}$$

$$\boxed{\sigma^2 = 22.4}$$

↳ Variance.

Variance through assumed mean:

When using the assumed mean formula $\left(\bar{x} = a + \frac{\Sigma(x-a)}{n}\right)$, the variance can be calculated using the following formula:

$$\sigma^2 = \frac{\Sigma(x-a)^2}{n} - \left[\frac{\Sigma(x-a)}{n}\right]^2$$

Using the same data set as an example: 1, 5, 9, 10, 15 where we assume $a = 7$

x	$(x-a)$	$(x-a)^2$
1	$1-7=-6$	$(-6)^2 = 36$
5	$5-7=-2$	$(-2)^2 = 4$
9	$9-7=2$	$2^2 = 4$
10	$10-7=3$	$3^2 = 9$
15	$15-7=8$	$8^2 = 64$
	5	117

$$\sigma^2 = \frac{\Sigma(x-7)^2}{5} - \left[\frac{\Sigma(x-7)}{5}\right]^2$$

$$= \frac{117}{5} - \left[\frac{5}{5}\right]^2$$

$$= \frac{117-1}{5}$$

$$= \frac{112}{5}$$

$$\boxed{\sigma^2 = 22.4}$$ → Variance

$\Sigma(x-a)$ $\Sigma(x-a)^2$

Combined Variance :

Just like how two separate sets of data can have a combined mean, they can also have a combined variance, which is calculated using the following formula:

$$\text{combined } \sigma^2 = \frac{\Sigma x^2 + \Sigma y^2}{n_x + n_y} - (\text{Combined mean})^2$$