Computational Statistics and Probability

Finalterm Assignment

Name: Nafiseen Leo

Id : 20-42195-1

| Serial No | Division | Gender | Age | Height | Weight | Education Level | Occupation | Covid-19 information |
|---|---|---|---|---|---|---|---|---|
| 1 | Dhaka | M | 46 | 174 | 82 | 16 | Teacher | Pos |
| 2 | Dhaka | F | 17 | 162 | 93 | 11 | Student | Neg |
| 3 | Dhaka | F | 43 | 168 | 53 | 12 | Housewife | Neg |
| 4 | Dhaka | M | 15 | 162 | 54 | 10 | Student | Neg |
| 5 | Dhaka | M | 12 | 150 | 46 | 7 | Student | Neg |
| 6 | Dhaka | M | 49 | 168 | 79 | 16 | Shopkeeper | Pos |
| 7 | Dhaka | F | 15 | 158 | 40 | 10 | Student | Neg |
| 8 | Dhaka | F | 41 | 163 | 51 | 12 | Housewife | Neg |
| 9 | Dhaka | F | 10 | 146 | 41 | 5 | Student | Pos |
| 10 | Dhaka | M | 38 | 165 | 69 | 15 | Teacher | Neg |
| 11 | Dhaka | M | 35 | 160 | 48 | 12 | Shopkeeper | Neg |
| 12 | Dhaka | F | 7 | 135 | 27 | 2 | Student | Pos |
| 13 | Mymensigh | F | 7 | 132 | 25 | 2 | Student | Neg |
| 14 | Mymensingh | M | 10 | 143 | 36 | 5 | Student | Neg |
| 15 | Mymensingh | M | 40 | 161 | 67 | 16 | Banker | Pos |
| 16 | Mymensingh | F | 36 | 153 | 44 | 12 | shopkeeper | Neg |
| 17 | khulna | F | 5 | 115 | 20 | 1 | student | Neg |
| 18 | Khulna | F | 30 | 158 | 63 | 16 | Businessman | Neg |
| 19 | Khulna | M | 7 | 94 | 24 | 1 | Student | Pos |
| 20 | Khulna | M | 33 | 169 | 68 | 16 | Teacher | Neg |
| 21 | Rangpur | F | 36 | 152 | 49 | 8 | Housewife | Neg |
| 22 | Rangpur | M | 40 | 157 | 58 | 12 | Businessman | Pos |
| 23 | Rangpur | M | 9 | 132 | 36 | 4 | Student | Neg |
| 24 | Rangpur | F | 8 | 120 | 28 | 3 | Student | Neg |
| 25 | Sylhet | M | 50 | 165 | 59 | 16 | Teacher | Neg |

Name: Nasinun Leo

Id: 20-42195-1

| Serial No | Division | Gender | Age | Height | Weight | Education level | Occupation | Covid-19 Information |
|---|---|---|---|---|---|---|---|---|
| 26 | Sylhet | M | 22 | 160 | 48 | 14 | Student | Neg |
| 27 | Sylhet | F | 47 | 159 | 51 | 12 | Housewife | Pos |
| 28 | Sylhet | F | 20 | 152 | 47 | 13 | Student | Pos |
| 29 | Chittagong | M | 56 | 170 | 70 | 8 | Bank Manager | Neg |
| 30 | Chittagong | F | 54 | 167 | 65 | 5 | Teacher | Pos |
| 31 | Chittagong | M | 24 | 162 | 60 | 16 | Civil officer | Pos |
| 32 | Chittagong | F | 22 | 160 | 56 | 14 | Student | Neg |
| 33 | Rajshahi | F | 60 | 164 | 64 | 15 | Housewife | Neg |
| 34 | Rajshahi | M | 64 | 166 | 68 | 16 | Businessman | Neg |
| 35 | Rajshahi | M | 30 | 169 | 72 | 15 | Teacher | Pos |
| 36 | Rajshahi | F | 28 | 160 | 63 | 15 | Housewife | Neg |
| 37 | Barisal | M | 48 | 172 | 76 | 12 | labour | Pos |
| 38 | Barisal | F | 45 | 171 | 72 | 10 | Housewife | Neg |
| 39 | Barisal | F | 18 | 163 | 55 | 12 | Student | Neg |
| 40 | Barisal | M | 15 | 158 | 62 | 10 | Student | Pos |

| No | 1 | 4 | 7 | 11 | 13 | 19 | 25 | 31 | 36 | 38 |
|---|---|---|---|---|---|---|---|---|---|---|
| Height | 174 | 162 | 158 | 160 | 132 | 143 | 165 | 162 | 160 | 171 |

Answer to the Question No(1).

Estimate of sample variance,

$$s^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{(\sum x)^2}{n}\right]$$

Name: Natinuan Leo Id: 20-42195-1

$$= \frac{1}{10-1}\left(253267 - \frac{2519569}{10}\right)$$

$$= 156.678$$

Estimate of the variance of sample means,

$$V(\bar{x}) = \frac{N-n}{Nn} S^2$$

$$= \frac{40-10}{40 \times 10} \times 156.678$$

$$= 11.75085$$

Estimate standard error of sample mean number of height. $s.e.(\bar{x}) = \sqrt{V(\bar{x})}$

$$= \sqrt{11.75085}$$

$$= 3.43$$

Name: Nafinan Leo

Id: 20-42195-1

Answer to the Question No (2)

Let, $x \sim N(\mu, \sigma^2)$

$\mu_0 = 172$

Here, we need to test, $H_0: \mu = \mu_0$ against $H_1: \mu \neq \mu_0$

$$\bar{x} = \frac{1}{n} \sum x$$

$$= \frac{1}{10} \times 1587$$

$$= 158.7$$

$$S^2 = \frac{1}{n-1} \left| \sum x^2 - \frac{(\sum x)^2}{n} \right|$$

$$= 156.678$$

Test statistics: $t = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$

$$= \frac{158.7 - 172}{12.52/\sqrt{10}}$$

$$|t| = 3.359$$

$$= -3.359$$

Name: Nasiman Leo. Id: 20-42135-1

$t_{10-1} = t_9 = 2.262$

Hence, $|t| \not< t_9$

not

So, $H_0$ is accepted

Now, from the sample we can conclude that the

not

average height of the population is 172 cm.

Answer to the Question No. (3)

| x (height) | y (weight | xy | $x^2$ | $y^2$ |
|---|---|---|---|---|
| 174 | 82 | 14268 | 30276 | 6724 |
| 162 | 59 | 8748 | 26244 | 2916 |
| 158 | 40 | 6320 | 24964 | 1600 |
| 160 | 48 | 7680 | 25600 | 2304 |
| 132 | 25 | 3300 | 17424 | 625 |
| 143 | 36 | 5148 | 20449 | 1296 |
| 165 | 54 | 8910 | 27225 | 2916 |
| 162 | 60 | 9720 | 26244 | 3600 |
| 160 | 63 | 10080 | 25600 | 3969 |
| 171 | 72 | 12312 | 29241 | 5184 |
| $\Sigma x = 1587$ | $\Sigma y = 539$ | $\Sigma xy = 86986$ | $\Sigma x^2 = 253267$ | $\Sigma y^2 = 31139$ |

Name: Nafiur Co. Id: 20-42196-1

$$SS(x) = 253267 - \frac{(1587)^2}{10}$$

$$= 1910.1$$

$$SS(y) = 33134 - \frac{(539)^2}{10}$$

$$= 2618.9$$

$$SP(xy) = 86486 - \frac{1587 \times 539}{10}$$

$$= 1740.2$$

$$\therefore r = \frac{1740.2}{\sqrt{1910.1 \times 2618.9}}$$

$$= 0.905$$

∴ The variables height (x) and weight (y) are positively correlated.

Name: Nafiur Leo, Id: 20-42196-1

Answer to the Question No. (9)

$$b = \frac{SP(XY)}{SS(X)}$$

$$= \frac{1740.2}{1410.1}$$

$$= 1.239$$

$$a = \bar{Y} - b\bar{X}$$

$$= \frac{\Sigma Y}{n} - \frac{b\Sigma X}{n}$$

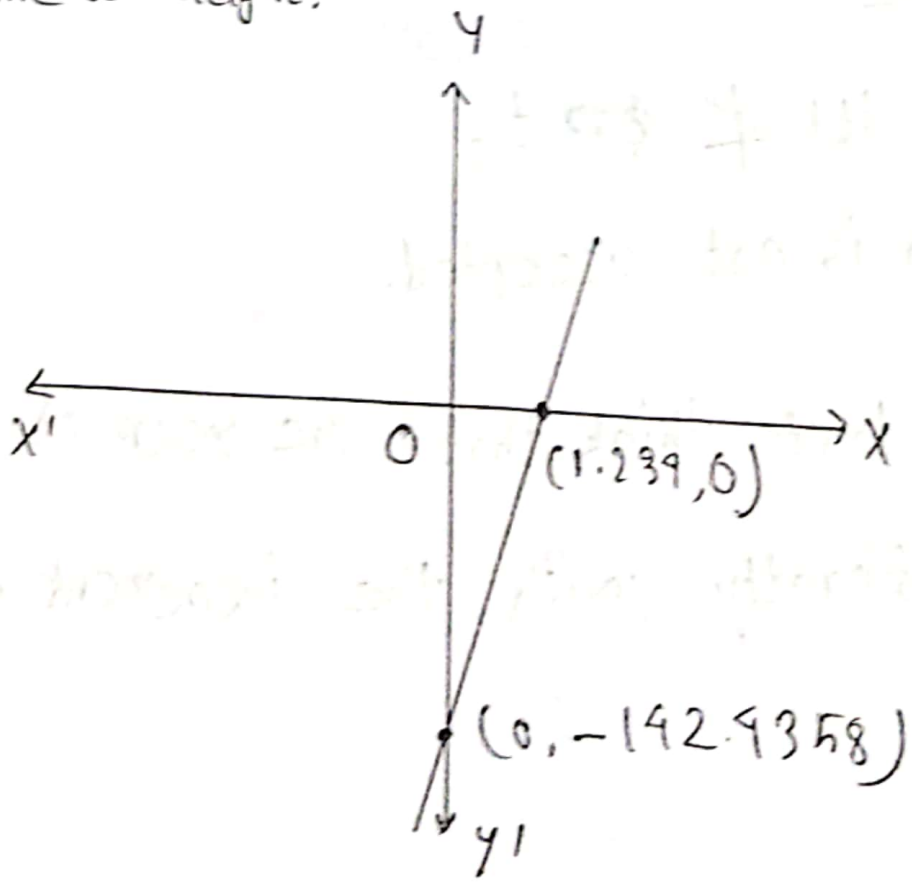
$$= \frac{539}{10} - \frac{1.239 \times 1587}{10}$$

$$= -142.4358$$

$$\therefore \hat{Y} = a + bX$$

$$= -142.4358 + 1.239X$$

Regression line of height:



Name: Nafiseur Leo Id: 20-42395-1

Answer to the Question No(5)

Here,

We need to test, $H_0 = \rho = 0$ vs $H_1: \rho \neq 0$

Test statistics: $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$

$= \frac{0.905\sqrt{10-2}}{\sqrt{1-(0.905)^2}}$

$= 6.017$

$t_{10-2} = t_8 = 2.306$

So, $|t| \not< t_8$

$H_0$ is not accepted.

So, I don't think that the year of schooling increase significantly with the increment of age.