Assignment Week 1

Soumil Hooda

January 2022

Contents

1 Answers

1.1 Question 1

Mode is the value that occurs with the b. Greatest frequency in a set of values.

1.2 Question 2

Data	
Height (in cm)	Number of Students
120-125	3
125-130	5
130-135	11
135-140	6
140-145	5
Total	30

In this dataset, modal class is 130-135 as it holds the highest frequency weight age that is 11. Therefore we obtain the following values,

L = 130
h = 5
$f_m = 11$
$f_1 = 5$
$f_2 = 6$

Putting these values in the formula,

$$Mode = L + h \times \frac{f_m - f_1}{(f_m - f_1) + (f_m + f_2)} \tag{1}$$

$$Mode = 130 + 5 \times \frac{11 - 5}{(11 - 5) + (11 + 6)} \tag{2}$$

$$Mode = 131.304 \tag{3}$$

The answer is option b.

1.3 Question 3

Median is the value occurring at the d. Mid Point of a set of ordered values.

1.4 Question 4

Data			
Marks	Number of Students	Cumulative Frequency	Percent Cumulative
0 - 20	6	6	7.5%
20 - 40	20	26	32.5%
40 - 60	37	63	78.75%
60 - 80	10	73	91.25%
80 - 100	7	80	100%

In this dataset, median class is 40-60 as it holds 50% cumulative frequency. Therefore we obtain the following values,

n = 80
h = 20
f = 37
l = 40
c = 32.5

Putting these values in the formula,

$$Median = l + h \times \frac{\frac{n}{2} - c}{f} \tag{4}$$

$$Median = 40 + 20 \times \frac{\frac{80}{2} - 32.5}{37} \tag{5}$$

$$Median = 44.05 \tag{6}$$

1.5 Question 5

The formula for mean is given as,

$$Mean = \frac{\sum_{i=1}^{n} x_i \times f_i}{\sum_{i=1}^{n} f_i} \tag{7}$$

where x_i is the midpoint of any given class which leads to,

$$Mean = \frac{10 \times 6 + 30 \times 20 + 50 \times 37 + 70 \times 10 + 90 \times 7}{6 + 20 + 37 + 10 + 7} \tag{8}$$

$$Mean = 48 \tag{9}$$

Therefore answer is option a.

1.6 Question 6

The data arranged in ascending form is as,

$$2, 6, 8, 9, 12, 15, 15, 16, 17, 22, 25, 32, 35, 42, 45 \tag{10}$$

Positions for Q1 and Q3 are 4th and 12th therefore IQR is 32-9 = 23. (option b)

1.7 Question 7

The answer is option c. 95%

1.8 Question 8

As number of trials are not too many, we can simply state the possible outcomes and find favourable events to determine probability.

The outcomes of three fair coins being tossed are as follows,

$$HHH, HTH, THH, TTH, HHT, HTT, THT, TTT \tag{11}$$

As we can clearly see, only 3 events (HHH, THH and HHT) are favourable to our problem. Therefore the probability is,

$$Probability = \frac{3}{8} \tag{12}$$

Therefore, answer is option a.

1.9 Question 9

Let us form a data table of sum of faces of 2 fair dices being rolled together. The probability can also be calculated and tabulated.

Data			
Sum	Frequency	Probability	Events
2	1	0.027	(1,1)
3	2	0.055	(1,2),(2,1)
4	3	0.083	(1,3),(3,1),(2,2)
5	4	0.111	(1,4)(2,3)(3,2)(4,1)
6	5	0.138	(1,5)(2,4)(3,3)(4,2)(5,1)
7	6	0.166	(1,6)(2,5)(3,4)(4,3)(5,2)(6,1)
8	5	0.138	(2,6)(3,5)(4,4)(5,3)(6,2)
9	4	0.111	(3,6)(4,5)(5,4)(6,3)
10	3	0.083	(4,6)(5,5)(6,4)
11	2	0.055	(5,6)(6,5)
12	1	0.027	(6,6)

The expected value is given by,

$$E(X) = \sum_{i=1}^{N} x_i \times P(x_i) \tag{13}$$

$$E(X) = 7 \tag{14}$$

Therefore, the answer is option a.

1.10 Question 10

For a continuous random variable, the expectation is given by,

$$E(X) = \int_{-\infty}^{\infty} x \times f_x \times dx \tag{15}$$

$$E(X) = \int_{1}^{\infty} x \times \frac{3}{x^4} \times dx \tag{16}$$

$$E(X) = \frac{-3}{2x^2}\Bigg|_{1}^{\infty} \tag{17}$$

$$E(X) = \frac{3}{2} \tag{18}$$

Therefore, the answer is option d.

1.11 Question 11

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \times f_x(x) \times dx \tag{19}$$

Therefore, the answer is option d.

1.12 Question 12

For mean = 2 and variance = 4 the PDF of a Gaussian RV x is,

$$f_x(x) = \frac{1}{\sigma \times \sqrt[2]{2 \times \pi}} \times e^{\frac{-(x-\mu)^2}{2 \times \sigma^2}} \tag{20}$$

$$f_x(x) = \frac{1}{\sqrt[2]{8 \times \pi}} \times e^{\frac{-(x-2)^2}{8}} \tag{21}$$

Therefore, the answer is option b.

1.13 Question 13

Using properties of the central limit theorem, the answer is

$$X \hookrightarrow \mathcal{N}(\mu, \frac{\sigma^2}{n}) \tag{22}$$

Which is option a.

1.14 Question 14

$$\widehat{f_x(x)} = \frac{\sum_{i=1}^{N} K(\frac{X_i - x}{h})}{Nh} \tag{23}$$

Therefore, the answer is option d.

1.15 Question 15

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x}) \times (y_i - \bar{y})}{\sqrt[2]{\sum_{i=1}^{n}(x_i - \bar{x})^2} \times \sqrt[2]{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{24}$$

Therefore, the answer is option c.