How to participate?
joinpd.com
code: see above

Statistics for Data Science-I

Week 3 Solve with Instructor (graded)

How to participate?
joinpd.com
code: see above

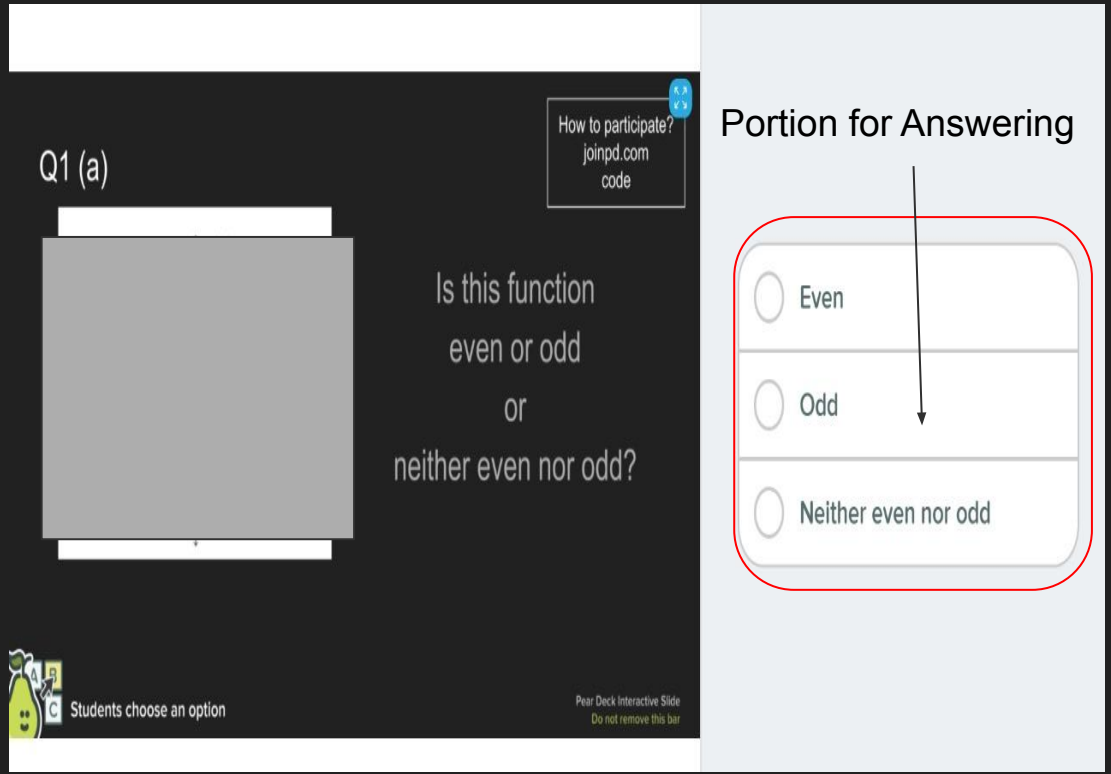
Statistics I: Week 3 Solve with Instructor

- Keep a notebook and pen ready for solving problems
- How to join?
  - Audio/screenshare on zoom - click on link sent to you
    - Doubts? Use zoom chat. Do not answer questions on webex chat.
  - Join on pear deck - joinpd.com (enter code seen on top right)
    - Answer questions only here
- For every question - 5 to 15 minutes allotted
  - Question will be shown in a slide for solving
  - If you are done solving, enter your answer at joinpd.com
  - Presenter will provide a solution
  - Questions and discussion

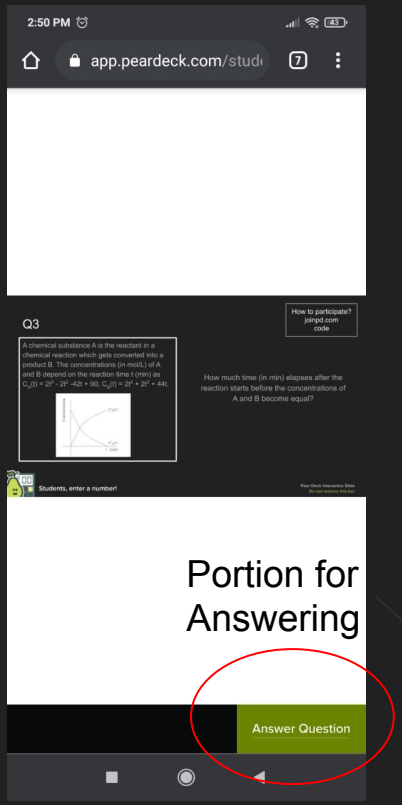
Example Screenshots

How to participate?
joinpd.com
code: see above

Laptop/Desktop



Mobile



Q1

The die is rolled 20 times and frequency of outcomes is tabulated below:

Face number	Frequency
1	2
2	5
3	3
4	4
5	2
6	4

What is the relative frequency corresponding to face number 5?

Students, enter a number!

Solution:

Total frequency of outcomes is 20.

As we know that,

$$Relative\ frequency = \frac{frequency}{Total\ frequency}$$

The frequency corresponding to face number 5 is 2.

Therefore,

$$Relative\ frequency = \frac{2}{20} = 0.1$$

Solution: Relative frequencies corresponding to outcomes of rolling a die 20 times.

Face Number	Frequency	Relative frequency
1	2	0.1
2	5	0.25
3	3	0.15
4	4	0.2
5	2	0.1
6	4	0.2

Concept of mean:

Mean:

The mean of the dataset is the sum of observations divided by the number of observations.

The sample mean for n observations is given by,

$$\overline{x} = \frac{x_1 + x_2 + x_3 + ...+ x_n}{n}$$

The population mean for N observations is given by,

$$\mu = \frac{x_1 + x_2 + x_3 + ...+ x_N}{N}$$

Prelude 1 to Q2

The number of the students in a class are divided in groups according to their weights and tabulated in the table below:

Weights (in kg)	Number of students
50-60	12
60-70	6
70-80	10
80-90	2

Q. What is the midpoint for the class interval [60-70)?

Students, enter a number!

Solution:

- The mid point of the class interval is the mean of the lowest and highest limits of class

Weights (in kg)	Number of students	Mid Point
50-60	12	55
60-70	6	65
70-80	10	75
80-90	2	85

Q.2

The number of the students in a class are divided in groups according to their weights and tabulated in the table below:

Weights (in kg)	Number of students
50-60	12
60-70	6
70-80	10
80-90	2

Q. What is the average weight of the class (in kg)?

Students, enter a number!

Solution:

- First, we will prepare a frequency table.

Weights (in kg)	Number of students f_i	Mid Point m_i	$f_i \cdot m_i$
50-60	12	55	660
60-70	6	65	390
70-80	10	75	750
80-90	2	85	170

Solution:

The mean for grouped data is given by,

$$\overline{x} = \frac{\Sigma f_i \, . \, m_i}{\Sigma f_i}$$

Therefore,

$$\overline{x} = \frac{1970}{30} = 65.66$$

Prelude 1 to Q3

The heights of five persons (in cm) are given in below table.

Name	Height (in cm)
Suresh	165
Prateek	170
Kajol	155
Mayur	170
Manaswi	160

Q. What is the average height (in cm) for this dataset?

Students, enter a number!

Solution:

The mean or average of the data is given by,

$$\overline{x} = \frac{x_1 + x_2 + x_3 + ...+ x_n}{n}$$

Therefore, the average height of persons given in the data is,

$$\overline{x} = \frac{165 + 170 + 155 + 170 + 160}{5}$$

$$\overline{x} = 164$$

Variance and Standard deviation

Population variance is given by: $$\sigma^2 = \frac{\Sigma (x-\mu)^2}{N}$$

Sample variance is given by : $$s^2 = \frac{\Sigma (x-\overline{x})^2}{n-1}$$

Population standard deviation is given by: $$\sigma = \sqrt{\frac{\Sigma (x-\mu)^2}{N}}$$

Population standard deviation is given by: $$s = \sqrt{\frac{(x-\overline{x})^2}{n-1}}$$

Prelude 2 to Q3

The heights of five persons (in cm) are given in below table.

Name	Height (in cm)
Suresh	165
Prateek	170
Kajol	155
Mayur	170
Manaswi	160

Q. What is the sample variance (in cm^2)of the data ?

Students, enter a number!

Solution:

We know, the average height of the persons = 164 cm.

The sample variance of the data is given by,

$$s^2 = \frac{\Sigma (x - \overline{x})^2}{n-1}$$

Therefore, the sample variance of the given data is,

$$s^2 = \frac{1^2 + 6^2 + 9^2 + 6^2 + 4^2}{4}$$

$$s^2 = 42.5$$

Q3

The heights of five persons (in cm) are given in below table.

Name	Height (in cm)
Suresh	165
Prateek	170
Kajol	155
Mayur	170
Manaswi	160

Q. What is the sample standard deviation (in cm) of the data ?

Students, enter a number!

Solution:

We know, the sample variance is 42.5.

Now, the sample standard deviation is given by,

$$s = \sqrt{s^2} = \sqrt{42.5} = 6.519$$

Thank You