

Week2 Quiz1

Most of the questions in this quiz are theoretical therefore there is no solved solution for them. For the solvable questions, here is the solution.

Below the solution, there is the complete quiz and answers.

Question 1:-

Handwritten solution for Question 1:

1) for set 1:-
$$u_1 = \frac{1+5+7+9}{4} = 5.5$$
$$\text{variance}_1 = \frac{(1-5.5)^2 + (5-5.5)^2 + (7-5.5)^2 + (9-5.5)^2}{4}$$
$$= \frac{20.25 + 0.25 + 2.25 + 12.25}{4}$$
$$= 8.75$$

for set 2:-
$$u_2 = -5$$
$$\text{variance}_2 = \frac{(-20+5)^2 + (-10+5)^2 + (0+5)^2 + (10+5)^2}{4}$$
$$= \frac{225 + 25 + 25 + 225}{4}$$
$$= 125$$

for set 3:-

$$u_3 = 101.5$$

$$\text{variance 3} = \frac{(100-101.5)^2 + (101-101.5)^2 + (102-101.5)^2 + (103-101.5)^2}{4}$$

$$= \frac{2.25 + 0.25 + 0.25 + 2.25}{4}$$
$$= 1.25$$

for set 4:-

$$u_4 = -5$$

$$\text{variance 4} = \frac{(-5)^2 + (0)^2 + 5^2 + 0}{4}$$

$$= 12.5$$

Question 3:-

3):-

$$Z = X + Y$$

as

$$Z \sim N(a u_x + b u_y, a^2 \sigma_x^2 + b^2 \sigma_y^2)$$

$$\sim N(5, \overset{5}{\textcircled{5}})$$

$$\sigma = \sqrt{5}$$
$$u = 5$$

1):-

for set 1:-

$$u = \frac{1+5+7+9}{4} = 5.5$$

Question 8:-



Theoretical v.

8

~~Median~~

Bmedian = 85

American = 75

Let's find IOR_A and IOR_B as per the rough values.

$$IOR_A = 78 - 67 = 11$$

$$IOR_B = 90 - 80 = 10$$

Congratulations! You passed!

 Grade received **100%** To pass 80% or higher

[Go to next item](#)


1.

Set	Values			
1	1	5	7	9
2	-20	-10	0	10
3	100	101	102	103
4	-10	-5	0	-5

1 / 1 point

 Consider the four sets of samples above. Which one has the smallest **variance**?

- ☐ 1
☐ 2
☒ 3
☐ 4


 **Correct**

Correct! The variance measures how much a sample is spread. We can easily look at all the samples and check that this one has the smallest spread among all.

 2. About the **mean**, **mode** and **median**, it is correct to say (check all that apply).


1 / 1 point

- ☐ The mean is always greater than or equal to the median.
☒ The median is more stable to outliers than the mean.

 **Correct**


 Correct! Since the median is a **positional** measure, it is more stable to extreme values!

- ☒ The mode is the most frequent value in a sample.

 **Correct**

Correct! This is precisely the definition of mode!

- ☒ There can be more than one mode in a sample.

 **Correct**

 Correct! The most frequent value might not be unique. For instance, in the sample $\{1, 1, 2, 2, 3, 4\}$, the modes are 1 and 2. Distributions which there are more than one mode are called **multimodal distributions**!

 3. Consider the following **independent** random variables:

1 / 1 point

$$\begin{aligned}
 X &\sim \text{Normal}(3, 1^2) \\
 Y &\sim \text{Normal}(2, 2^2)
 \end{aligned}$$

 Then $Z = X + Y \sim \text{Normal}(\mu, \sigma^2)$, where μ, σ are equal to:

- ☐ $\mu = \sqrt{5}, \sigma = \sqrt{3}$

- ☒ $\mu = 5, \sigma = \sqrt{5}$
- ☐ $\mu = 5, \sigma = \sqrt{3}$
- ☐ $\mu = 5, \sigma = 5$

✓ Correct

Correct! Using the formula $\mu_Z = \mu_X + \mu_Y$ and $\sigma_Z = \sqrt{\sigma_X^2 + \sigma_Y^2}$ you get the result!

4. What is the advantage of looking at the standard deviation instead of the variance?

1 / 1 point

- ☐ The standard deviation is less affected by outliers than the variance.
- ☒ The standard deviation has the same dimension as the sample.
- ☐ The standard deviation may be negative.
- ☐ There are no advantages. They mean the same thing.

✓ Correct

Correct! The variance has the sample's dimension squared, whereas the standard deviation has the same dimension as the sample. This makes it easier to interpret.

5. Which of the following best describes the purpose of a QQ plot?

1 / 1 point

- ☐ To visualize the distribution of a dataset
- ☐ To visualize the relationship between two datasets
- ☒ To test for normality of a dataset
- ☐ To test for independence between two datasets

✓ Correct

Correct! The main purpose of a QQ plot is to test for normality of a dataset.

6. Which of the following is true about a box plot with a long whisker on the upper end?

1 / 1 point

- ☐ The data is negatively skewed
- ☒ The data is positively skewed
- ☐ The data is normally distributed
- ☐ There are no outliers

✓ Correct

Correct! A box plot with a long whisker on the upper end is an indication of positively skewed data.

7. Which of the following is a correct statement regarding a QQ plot?

1 / 1 point

☐ If the points fall close to the diagonal, the data is not normally distributed

☒ If the points fall close to the diagonal line, the data is normally distributed

✓ Correct

Correct! If the points fall close to the diagonal line, it indicates that the data is normally distributed.

☐ If the points fall far from the diagonal line, the data is normally distributed

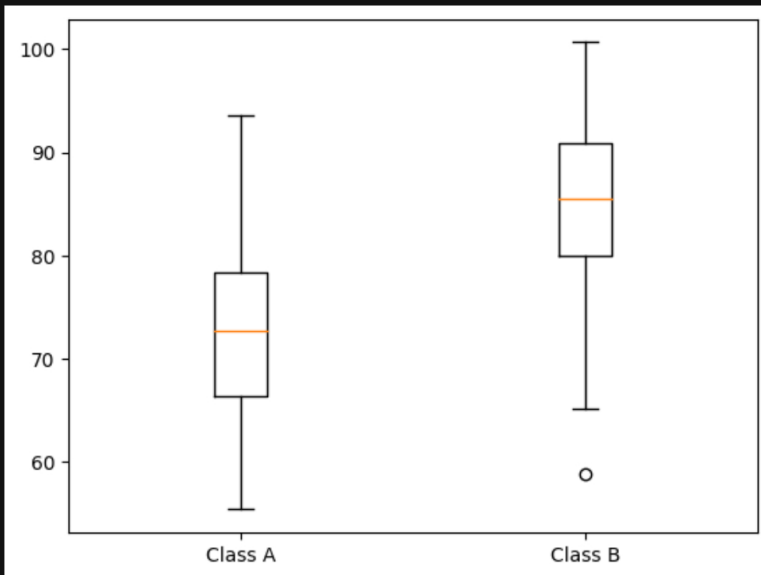
☒ If the points fall far from the diagonal line, the data is not normally distributed

✓ Correct

Correct! If the points fall far from the diagonal line, it indicates that the data is not normally distributed.

8. Consider the following box plot for the test scores of two classes, A and B:

1 / 1 point



Which of the following statements is true?

☐ The median score of Class A is higher than the median score of Class B.

☒ The median score of Class B is higher than the median score of Class A.

✓ Correct

Correct! Looking at the box plot, we can see that the median of Class A is around 75, while the median of Class B is around 85.

☒ The interquartile range of Class A is larger than the interquartile range of Class B.

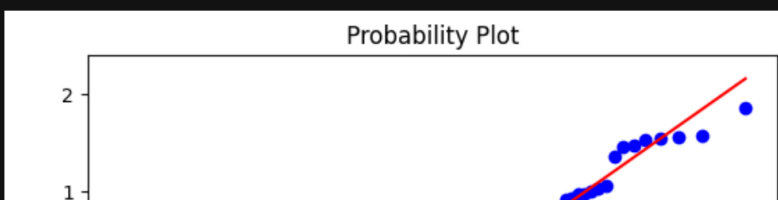
✓ Correct

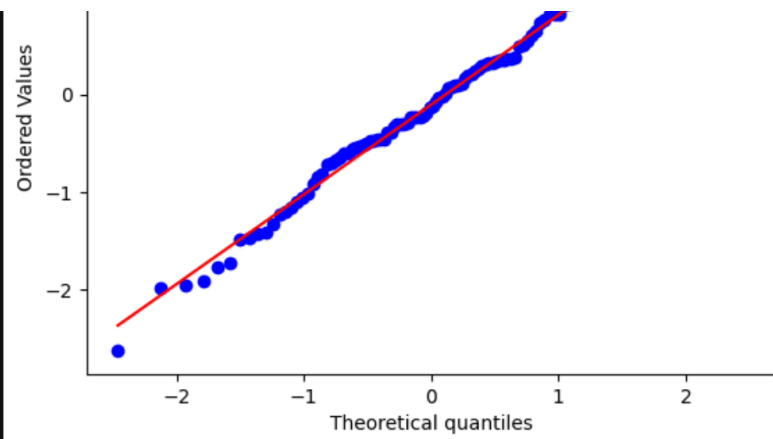
Correct! The rectangle in A is bigger than B.

☐ The interquartile range of Class B is larger than the interquartile range of Class A.

9. Consider the following QQ plot for a set of data:

1 / 1 point





Which of the following statements is true?

- ☒ The data looks normally distributed.
- ☐ The data is not normally distributed.
- ☐ The data has a higher variance than a normal distribution.
- ☐ The data has a lower variance than a normal distribution.

✓ Correct

The QQ plot compares the observed data with the theoretical quantiles of a normal distribution. If the points lie close to the diagonal line, then the data is likely normally distributed.