

DA Assignment -1

Sayam Kumar
S20180010158
Page 1

Q-1 First type of data list is on interval scale because we don't have a clear zero definition but can't perform additive operations

Second type of data is on ratio scale because it has a clear definition of zero and both additive and multiplicative operations are valid

Q-2 Data

25000, 25000, 35000, 35000, 35000, 35000, 35000,
35000, 45000, 45000, 45000, 45000, 60000, 60000,
 500×10^3 , 500×10^3

Total data points = 16

$$\text{Median}(Q_2) = \frac{35000 + 45000}{2} = 40,000$$

$$Q_1 = 35000$$

$$Q_3 = \frac{45000 + 60000}{2} = 52500$$

$$\text{IQR} = Q_3 - Q_1 = 52500 - 35000 = 17500$$

$$\text{Mean}(\bar{x}) = \frac{\sum x_i}{16} = 1560000$$

$$\text{Standard deviation}(\sigma) = \sqrt{\frac{1}{(N-1)} \sum (x_i - \bar{x})^2} = 157437.81$$

Q-3 Mean absolute deviation around mean

$$(a) AAD = \frac{1}{N} \sum_{i=1}^N |x_i - \bar{x}|$$

$$= \frac{1610000}{16} = 100625$$

Sayam

S20180010158

Page 2

(ii) Median absolute deviation around median

$$\text{Median} = 40,000$$

$$MAD = \text{median}(|x_i - \text{median}(x)|)$$

$$= \frac{5000 + 5000}{2}$$

$$\boxed{= 5000}$$

(from updated list)

Part 2 Data Categorization

Q-1 (a) Harmonic mean - Interval, Ratio

(b) Median of a sample - Ordinal, Interval, Ratio

(c) Range - Interval, Ratio

(d) IQR - Interval, Ratio

(e) Maximum - Ordinal, Interval, Ratio

because we can compare O, I, and R data

Q Part 3

Prob Distributions and Hypothesis Test

Sayam

S20180010158

Page 3

Q-1 Data

53.8, 56.1, 54.5, 54.8, 55.2, 55.1,
54.7, 55.8, 55.9, 54.5

Null Hypothesis $\rightarrow P(X \leq 55)$

$$\bar{x} = 55.04$$

Alternate Hypothesis $\rightarrow P(X > 55)$

$$\sigma = 0.3$$

$$H_0: P\left(Z \leq \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}\right) = P\left(Z \leq \frac{55.04 - 55}{0.3/\sqrt{10}}\right)$$

$$Z_{\text{stat}} = 0.4216$$

$$P\text{value} = 0.3372 \quad (\text{from the one sided } Z \text{ table})$$

Yes, we accept the null hypothesis
because $p\text{value} = 1 - 0.6628$
 $= 0.3372$

z value doesnot lie in rejection region.

we accept the null hypothesis.

Q-2 we will use Chi-Square test to verify the amount of colour to be less than 0.3g. Since chi-square test involves degrees of freedom, and variance, it is best suited here.

Sayam
S20180010158
Page 4

Q-3

Pair	Experimental Group	Control Group	Difference	Rank
		112	-2	1
1	110	120	5	
2	125		11	
3	139	128	7	
4	142	135	1	
5	127	126		

Ordered Absolute values	Signed Ranks
1	1
-2	-2
5	3
7	4
11	5

$W^+ = \text{sum of +ve signed ranks} = 13$

$W^- = \text{sum of -ve signed ranks} = 2$

Part (a) is correct

Q-3 Part-3

The differences of group 2 ID from group 1 are

$$(-2, 1, 3, 5, 4)$$

$$\bar{x} = \frac{\sum x_i}{n} = 2.2$$

$$s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 = 7.69$$

$$s = \sqrt{7.69} = 2.77$$

$$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

s = standard deviation of sample

The mean of difference of IDs from the groups as per table = 0

$$\text{So } Z = \frac{2.2}{2.77/\sqrt{5}} = 1.78$$

$$p\text{-value} | Z = 1.78 \Rightarrow 0.9625$$

So, the closest option is (b) 0.9663

Reasonable p-value is 0.9663 (b) Answer.

Sayam

20180010158

Page 5

Part 2 Q-3

Sayam
S20120010156
Page 6

H_0 : Both the groups are similar

$$\mu_1 = \mu_2$$

H_1 : IQ for group 1 is higher than group 2

$$\mu_1 > \mu_2$$

$$\alpha = 0.05$$

$$\begin{aligned}\text{Test statistics} &= \min(w_+, w_-) \\ &= \min(13, 2) \\ &= 2\end{aligned}$$

$$\text{Critical Value} = 0 \quad | \quad \text{at } \alpha = 0.05 \text{ and } n = 5$$

We fail to reject null hypothesis

because test statistics lie in acceptance region (2 > 0).

So answer = C part The experimental groups tends to have similar IQ to control group.

C part answer