Name:V DHARANI SREE **RollNumber-A22126510148**

CSE-C

Class Test **Total: 5*1=5m**

UNIT–I: PROBABILITY & DISTRIBUTIONS

Q1. Apply the Gaussian distributionformulatocomputethe probability density at x = 1 for μ = 0 and σ = 1. CO1,CO2,CO3 L3

Answer:

To computethe **probability density** of a **Gaussian (Normal) distribution** at **x = 1**, with **mean μ = 0** and **standard deviation σ = 1**, use the formula:

Gaussian Probability Density Function (PDF)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Given:

- $x = 1$
- $\mu = 0$
- $\sigma = 1$

Step-by-Step Solution

1. Substitute values

$$f(1) = \frac{1}{1\sqrt{2\pi}} e^{-\frac{(1-0)^2}{2(1)^2}}$$

2. Simplify exponent

$$(1-0)2 = 1$$

$$-\frac{1}{2} = -0.5$$

So,

$$f(1) = \frac{1}{\sqrt{2\pi}} e^{-0.5}$$

3. Numerical values

$$\sqrt{2\pi} \approx 2.5066$$
$$e^{-0.5} \approx 0.6065$$

4. Final computation

$$f(1) = \frac{0.6065}{2.5066}$$
$$f(1) \approx 0.24197$$

Final Answer

$$\boxed{f(1) \approx 0.24197}$$

This is the probability density value of the standard normal distribution at x = 1.

UNIT–II: HYPOTHESIS TESTING

Q2. Apply the steps of hypothesis testing to check whether a sample mean of 52 differs significantly from population mean 50 (σ = 5, n = 25). CO3 L3

Answer:

1. Set up hypotheses

- Null hypothesis: $H_0: \mu = 50$
- Alternative hypothesis: $H_a: \mu \neq 50$ (two-tailed — we're checking if the sample mean differs)

Assumptions: population standard deviation $\sigma$is known (=5), sample is random, and either the population is approximately normal or $n$is large enough (here n = 25— borderline but OK if population is roughly normal).

---

2. Choose significance level

Use α = 0.05(common choice for two-tailed test). Critical

z-values for two-tailed $\alpha = 0.05$:

$$z\text{crit} = \pm 1.96$$

---

3. Compute test statistic (z)

Formula:

$$\boldsymbol{z} = \frac{\bar{\boldsymbol{x}} - \boldsymbol{\mu 0}}{\boldsymbol{\sigma}/\sqrt{\boldsymbol{n}}}$$

Plug values:

- $\bar{x} = 52$
- $\mu 0 = 50$
- $\sigma = 5$ n
- $= 25$so $\sqrt{n} = 5$

Denominator: $\sigma/\sqrt{n} = 5/5 = 1$.

So

$$z = \frac{52 - 50}{1} = \frac{2}{1} = 2.00.$$

---

4. Find p-value

Two-tailed p-value = $2 \times P(Z > 2.00)$.

Standard normal tail $P(Z > 2.00) \approx 0.0228$.
So p-value $\approx 2 \times 0.0228 = 0.0456$ (rounded).

---

5. Decision

- Compare test statistic to critical values: $z = 2.00$ is greater than $+1.96$, so it falls in the rejection region.
- Or compare p-value: $0.0456 < \alpha(= 0.05)$ → reject H_0.

---

6. Conclusion

At the 5% significance level, there is sufficient evidence to conclude the sample mean 52 differs significantly from the population mean 50. (Two-tailed test: reject $H0$.)

Final summary (boxed)

$$\boxed{z = 2.00, \text{p-value} \approx 0.0456, \text{Reject } H_0} \boxed{\text{at } \alpha = 0.05.}$$

UNIT–III: REGRESSION ANALYSIS

Q3. Apply the linear regression formula to calculate ŷ for x = 10, given model ŷ = 2x + 5. CO4 L3
Answer:

To calculate the predicted value $\hat{y}$ ûsing the linear regression model:

$$\hat{y}\text{^} = 2x + 5$$

Given:

$$x = 10$$

Substitute the value:

$$\hat{y}\text{^} = 2(10) + 5$$
$$\hat{y}\text{^} = 20 + 5$$

$\hat{y}$^ = 25

Final Answer:

$$\boxed{\hat{y}\text{^} = 25}$$

UNIT–IV: CLASSIFICATION & CLUSTERING

Q4. Apply the K-means algorithm to assign a point P(4,6) to its nearest cluster if centroids are C1(2,3) and C2(6,7). CO4 L3

Answer:

To assign point P(4,6) to the nearest cluster, compute the Euclidean distance to both centroids.

Step 1: Distance to C_1(2,3)

$$d(P, C_1) = \sqrt{(4-2)^2 + (6-3)^2}$$

$$= \sqrt{2^2 + 3^2}$$

$$= \sqrt{4 + 9} = \sqrt{13}$$

$$\approx 3.61$$

Step 2: Distance to C_2(6,7)

$$d(P, C_2) = \sqrt{(4-6)^2 + (6-7)^2}$$

$$= \sqrt{(-2)^2 + (-1)^2}$$

$$= \sqrt{4 + 1} = \sqrt{5}$$

$$\approx 2.24$$

Step 3: Compare distances

- Distance to $C_1 \approx 3.61$
- Distance to $C_2 \approx 2.24$

Final Answer:

$$\boxed{\text{Point } P(4,6) \text{ is assigned to Cluster } C}$$

UNIT–V: DATA VISUALIZATION

Q5. Apply a heatmap to identify patterns: If a correlation matrix shows corr(A,B)=0.9 and corr(A,C)=0.1, which pair will appear darker? CO5 L3

Answer:

Inaheatmap of a correlation matrix, darker colors usually represent stronger correlations (either strongly positive or strongly negative).

Given:

- corr(A,B) = 0.9 → very strong positive correlation
- corr(A,C) = 0.1 → very weak correlation

Answer:

$$\boxed{\boldsymbol{Thepair(A,B)willappeardarkerintheheatmap}.}$$

Because 0.9 is much stronger than 0.1, it will be shown with a darker (more intense) color.