

Department of Computer Science & Engineering
Indian Institute of Technology Kharagpur

Class Test 1
Data Analytics : CS61061
(Autumn-2022)

Full marks: 20

Time: 40 minutes

Instructions:

- You are advised to attempt all questions.
- Use separate page to answer to next question.
- You can use non-programmable calculator and statistical table, if needed.
- Write your name, roll number clearly on the front page of the answer script.
- You may plan maximum 10 minutes to answer each question.
- Each question carry 5 marks.

1. Mr. X is looking for a safe investment that provides safe and stable returns. There are following options:
- a) **Stocks:** The volatility of the stock is 10% and the expected return is 14%.
 - b) **Mutual funds:** It offers an expected return of 13% with a volatility of 7%.
 - c) **Fixed deposits:** This scheme offers an expected return of 3% with 2% volatility.
- In order to select the most suitable investment opportunity, Mr. X should choose which investment scheme?

Solution:

Here Mr. X should calculate the coefficient of variance of each investment scheme, That is

$$cv = \frac{\sigma}{\mu} * 100 = \frac{Volatility}{Expectation} \times 100$$

$$cv \text{ (Stocks)} = (10/14) \times 100 = 71.4\%$$

$$cv \text{ (MF)} = (7/13) \times 100 = 53.8\%$$

$$cv \text{ (FD)} = (2/3) \times 100 = 66.7\%$$

Based on the above calculation, Mr. X should invest in MF scheme as it offers the lower coefficient (of variation).

2. 10 workers can do some work in 2 days, 5 workers in 10 days and 3 workers in 6 days. How many days on the average that a worker can do?

Solution:

This problem can be showed with weighted harmonic mean.

Here:

$$w1 = 10; \quad w2 = 5; \quad w3 = 3;$$

$$d1 = 2; \quad d2 = 10; \quad d3 = 6;$$

$$r = \frac{w1+w2+w3}{w1 * \frac{1}{d1} + w2 * \frac{1}{d2} + w3 * \frac{1}{d3}}$$

$$= \frac{10+5+3}{10 * \left(\frac{1}{2}\right) + 5 * \left(\frac{1}{10}\right) + 3 * \left(\frac{1}{6}\right)}$$

$$= 3 \text{ days}$$

3. A production house packs 1000 items in a box. It is believed that 1 item in 2000 on an average is defective. What is the probability that a box contains 2 or more defective items?

Solution:

We are to find $P(x \geq 2)$.

The correct distribution of the random variable is Bernoulli's Distribution. Here,

$n = 1000$; $p = 1/2000$; $q = 1999/2000$
using Bernoulli's Distribution,

$$P(x = 0) = \left(\frac{1999}{2000}\right)^{1000} = 0.60645$$

$$P(x = 1) = 1000 * \left(\frac{1999}{2000}\right)^{1000} * \left(\frac{1}{2000}\right) = 0.30338$$

$$P(x = 0) + P(x = 1) = 0.60645 + 0.30338 = 0.90983$$

$$\text{There, } P(x \geq 2) = 1 - 0.90983 = 0.0902 \text{ (Ans)}$$

4. A class test of 60 marks was conducted and the scores of the students is found to be normally distributed with a mean of 50 and standard deviation 12.
A sample of 225 students' scores are selected at random and mean score is found to be 52 with σ remains the same.
Hypothesis to be tested that mean score in the class test is valid. Calculate the p-value of this test. What it does mean?

Solution:

(a) $H_0: \mu = 50$; $H_1: \mu \neq 50$

(b) Calculate Statistical value of the sample.

$$n = 225; \quad \sigma = 12; \quad \bar{x} = 55;$$

$$z = \frac{\frac{52-50}{\frac{12}{\sqrt{225}}}}{12} = \frac{2*15}{12} = 2.5$$

(c) $p = 2 * P(z > 2.5) = 2 * 0.0062 = 0.0124$

Assuming, $\alpha = 0.05$

p-value $< \alpha$.

μ_0 is *Rejected*.

(d) Mean: The chance of committing Type – I error is 1.24%