Sem2 Statistics2

Section Id :	64065338383
Section Number :	2
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	12
Number of Questions to be attempted :	12
Section Marks :	40
Display Number Panel :	Yes
Group All Questions :	No
Enable Mark as Answered Mark for Review and Clear Response :	Yes
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	64065380775
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Number : 23 Question Id : 640653565379 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 0

Question Label : Multiple Choice Question

THIS IS QUESTION PAPER FOR THE SUBJECT "FOUNDATION LEVEL : SEMESTER 2: STATISTICS FOR DATA SCIENCE 2 (COMPUTER BASED EXAM)"

ARE YOU SURE YOU HAVE TO WRITE EXAM FOR THIS SUBJECT?

CROSS CHECK YOUR HALL TICKET TO CONFIRM THE SUBJECTS TO BE WRITTEN.

(IF IT IS NOT THE CORRECT SUBJECT, PLS CHECK THE SECTION AT THE TOP FOR THE SUBJECTS REGISTERED BY YOU)

Options :

6406531889853. ✔ YES

6406531889854. ✖ NO

Question Number : 24 Question Id : 640653565380 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 0

Question Label : Multiple Choice Question

1. **Markov's inequality:** Let X be a discrete random variable taking non-negative values with a finite mean μ. Then,
$$P(X \geq c) \leq \frac{\mu}{c}$$

2. **Chebyshev's inequality:** Let X be a discrete random variable with a finite mean μ and a finite variance σ^2. Then,
$$P(\mid X - \mu \mid \geq k\sigma) \leq \frac{1}{k^2}$$

3. **Weak Law of Large numbers:** Let $X_1, X_2, \ldots, X_n \sim$ iid X with $E[X] = \mu, \text{Var}(X) = \sigma^2$.
Define sample mean $\overline{X} = \dfrac{X_1 + X_2 + \ldots + X_n}{n}$. Then,
$$P(|\overline{X} - \mu| > \delta) \leq \frac{\sigma^2}{n\delta^2}$$

4. **Using CLT to approximate probability:** Let $X_1, X_2, \ldots, X_n \sim$ iid X with $E[X] = \mu, \text{Var}(X) = \sigma^2$.
Define $Y = X_1 + X_2 + \ldots + X_n$. Then,
$$\frac{Y - n\mu}{\sqrt{n}\sigma} \approx \text{Normal}(0, 1).$$

5. **Bias of an estimator**: $\text{Bias}(\hat{\theta}, \theta) = E[\hat{\theta}] - \theta$.

6. **Method of moments:** Sample moments, $M_k(X_1, X_2, \ldots, X_n) = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k$
Procedure: For one parameter θ
   - Sample moment: m_1
   - Distribution moment: $E(X) = f(\theta)$
   - Solve for θ from $f(\theta) = m_1$ in terms of m_1.
   - $\hat{\theta}$: replace m_1 by M_1 in the above solution.

7. **Likelihood of i.i.d. samples**: Likelihood of a sampling $x_1, x_2, \ldots, x_n$, denoted
$$L(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_X(x_i; \theta_1, \theta_2, \ldots)$$

8. **Maximum likelihood (ML) estimation**:
$$\theta_1^*, \theta_2^*, \ldots = \arg\max_{\theta_1^*, \theta_2^*, \ldots} \prod_{i=1}^{n} f_X(x_i; \theta_1, \theta_2, \ldots)$$

9. **Bayesian estimation:** Let $X_1, \ldots, X_n \sim$ i.i.d.X, parameter Θ.
   Prior distribution of $\Theta : \Theta \sim f_\Theta(\theta)$.
   Samples, $S : (X_1 = x_1, \ldots, X_n = x_n)$
   Posterior: $\Theta \mid (X_1 = x_1, \ldots, X_n = x_n)$
   Bayes' rule: Posterior $\propto$ Prior $\times$ Likelihood
   Posterior density $\propto f_\Theta(\theta) \times P(X_1 = x_1, \ldots, X_n = x_n \mid \Theta = \theta)$

10. **Normal samples with unknown mean and known variance:**
    $X_1, \ldots, X_n \sim$ i.i.d. Normal(M, σ^2).
    Prior $M \sim$ Normal(μ_0, σ_0^2).
    Posterior mean: $\hat{\mu} = \overline{X}\left(\dfrac{n\sigma_0^2}{n\sigma_0^2 + \sigma^2}\right) + \mu_0\left(\dfrac{\sigma^2}{n\sigma_0^2 + \sigma^2}\right)$

11. **Hypothesis Testing**

- **Test for mean**
  Case (1): When population variance σ^2 is known (z-test)

Test	H_0	H_A	Test statistic	Rejection region
right-tailed	$\mu = \mu_0$	$\mu > \mu_0$	$T = \overline{X}$ $Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$	$\overline{X} > c$
left-tailed	$\mu = \mu_0$	$\mu < \mu_0$	$T = \overline{X}$ $Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$	$\overline{X} < c$
two-tailed	$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \overline{X}$ $Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$	$\lvert\overline{X} - \mu_0\rvert > c$

Case (2): When population variance σ^2 is unknown (t-test)

Test	H_0	H_A	Test statistic	Rejection region
right-tailed	$\mu = \mu_0$	$\mu > \mu_0$	$T = \overline{X}$ $t_{n-1} = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$	$\overline{X} > c$
left-tailed	$\mu = \mu_0$	$\mu < \mu_0$	$T = \overline{X}$ $t_{n-1} = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$	$\overline{X} < c$
two-tailed	$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \overline{X}$ $t_{n-1} = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$	$\lvert\overline{X} - \mu_0\rvert > c$

- χ^2-test for variance:

Test	H_0	H_A	Test statistic	Rejection region
right-tailed	$\sigma = \sigma_0$	$\sigma > \sigma_0$	$T = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$	$S^2 > c^2$
left-tailed	$\sigma = \sigma_0$	$\sigma < \sigma_0$	$T = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$	$S^2 < c^2$
two-tailed	$\sigma = \sigma_0$	$\sigma \neq \sigma_0$	$T = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$	$S^2 > c^2$ where $\frac{\alpha}{2} = P(S^2 > c^2)$ or $S^2 < c^2$ where $\frac{\alpha}{2} = P(S^2 < c^2)$

- Two samples z-test for means:

Test	H_0	H_A	Test statistic	Rejection region
right-tailed	$\mu_1 = \mu_2$	$\mu_1 > \mu_2$	$T = \overline{X} - \overline{Y}$ $\overline{X} - \overline{Y} \sim \text{Normal}\left(0, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$ if H_0 is true	$\overline{X} - \overline{Y} > c$
left-tailed	$\mu_1 = \mu_2$	$\mu_1 < \mu_2$	$T = \overline{Y} - \overline{X}$ $\overline{Y} - \overline{X} \sim \text{Normal}\left(0, \frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}\right)$ if H_0 is true	$\overline{Y} - \overline{X} > c$
two-tailed	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$T = \overline{X} - \overline{Y}$ $\overline{X} - \overline{Y} \sim \text{Normal}\left(0, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$ if H_0 is true	$\lvert\overline{X} - \overline{Y}\rvert > c$

- Two samples F-test for variances

Test	H_0	H_A	Test statistic	Rejection region
one-tailed	$\sigma_1 = \sigma_2$	$\sigma_1 > \sigma_2$	$T = \frac{S_1^2}{S_2^2} \sim F_{(n_1-1,n_2-1)}$	$\frac{S_1^2}{S_2^2} > 1 + c$
one-tailed	$\sigma_1 = \sigma_2$	$\sigma_1 < \sigma_2$	$T = \frac{S_1^2}{S_2^2} \sim F_{(n_1-1,n_2-1)}$	$\frac{S_1^2}{S_2^2} < 1 - c$
two-tailed	$\sigma_1 = \sigma_2$	$\sigma_1 \neq \sigma_2$	$T = \frac{S_1^2}{S_2^2} \sim F_{(n_1-1,n_2-1)}$	$\frac{S_1^2}{S_2^2} > 1 + c_R$ where $\frac{\alpha}{2} = P(T > 1 + c_R)$ or $\frac{S_1^2}{S_2^2} < 1 - c_L$ where $\frac{\alpha}{2} = P(T < 1 - c_L)$

Use the following values if required:
$F_{t_{15}}(2) = 0.968, F_{t_{15}}(\frac{3}{2}) = 0.923, F_Z(1.83) = 0.9664$

Options :

6406531889855. ✔ Useful Data has been mentioned above.

6406531889856. ✖ This data attachment is just for a reference & not for an evaluation.

Sub-Section Number : 2

Sub-Section Id : 64065380776

Question Shuffling Allowed : Yes

Is Section Default? : null

Question Number : 25 Question Id : 640653565381 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

Sushant first throws a fair die, then throws as many fair coins as the number that showed on the die. If the die showed 5, then find the conditional probability that 3 heads are obtained. Enter the answer correct to three decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.310 to 0.316

Question Number : 26 Question Id : 640653565383 Question Type : SA Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

Suppose $X \sim \text{Uniform}[0, \theta]$, where θ is an unknown constant. Consider a sample (4, 7, 8, 11, 5). Let $\hat{\theta}_{MME}$ and $\hat{\theta}_{ML}$ be the method of moments estimator and the maximum likelihood estimator of θ, respectively. Find $|\hat{\theta}_{MME} - \hat{\theta}_{ML}|$.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

3

Sub-Section Number :	3
Sub-Section Id :	64065380777
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 27 Question Id : 640653565384 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3 Selectable Option : 0

Question Label : Multiple Select Question

Which of the following statement(s) is/are correct?

Options :

6406531889863. ✔ If the P-value of a test is 0.025, then the corresponding test will reject the null hypothesis at the significance level of 0.03.

6406531889864. ✖ $\alpha + \beta = 1$ for all statistical tests, where α is the significance level and β is the probability of correctly rejecting the null hypothesis.

6406531889865. ✖ The probability of accepting the null hypothesis when it is false is equal to the power of the test.

6406531889866. ✔ The probability of rejecting the Null hypothesis when it is true is called the level of significance.

Sub-Section Number : 4

Sub-Section Id : 64065380778

Question Shuffling Allowed : Yes

Is Section Default? : null

Question Number : 28 Question Id : 640653565382 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Multiple Choice Question

Suppose X_1, X_2, X_3 are i.i.d. samples from a distribution X with an unknown mean μ and variance σ^2. Let $\hat{\mu}_1 = X_1 + X_2 - X_3$, $\hat{\mu}_2 = 3X_1 + 2X_2 - 4X_3$ and $\hat{\mu}_3 = \frac{X_1}{2} + X_2 - \frac{X_3}{2}$ be three unbiased estimators of μ. Which of the following is true?

Options :

6406531889858. ✖ $Var(\hat{\mu}_1) < Var(\hat{\mu}_3) < Var(\hat{\mu}_2)$

6406531889859. ✖ $Var(\hat{\mu}_2) > Var(\hat{\mu}_3) > Var(\hat{\mu}_1)$

6406531889860. ✔ $Var(\hat{\mu}_3) < Var(\hat{\mu}_1) < Var(\hat{\mu}_2)$

6406531889861. ✖ $Var(\hat{\mu}_3) > Var(\hat{\mu}_2) > Var(\hat{\mu}_1)$

Question Number : 29 Question Id : 640653565385 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Multiple Choice Question

Consider two independent samples $X_1, X_2, \ldots, X_{50} \sim$ i.i.d. Normal$(\mu_1, 25)$ and $Y_1, Y_2, \ldots, Y_{20} \sim$ i.i.d. Normal$(\mu_2, 70)$. Let the null and alternative hypothesis be

$$H_0 : \mu_1 = \mu_2$$

$$H_A : \mu_1 \neq \mu_2$$

Suppose $T = \overline{Y} - \overline{X}$, where $\overline{Y} = \dfrac{Y_1 + Y_2 + \cdots + Y_{20}}{20}$ and $\overline{X} = \dfrac{X_1 + X_2 + \cdots + X_{50}}{50}$. Consider a test that rejects H_0 if $|T| > c$ for some constant c. What is the size of the test in terms of 'c'?

Options :

6406531889867. ✖ $1 - F_Z\left(\dfrac{-c}{2}\right)$

6406531889868. ✔ $2F_Z\left(\dfrac{-c}{2}\right)$

6406531889869. ✖ $1 - 2F_Z\left(\dfrac{c}{2}\right)$

6406531889870. ✖ $F_Z\left(\dfrac{c}{2}\right)$

Sub-Section Number :	5
Sub-Section Id :	64065380779
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Id : 640653565386 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (30 to 31)

Question Label : Comprehension

Suppose $X \sim$ Binomial $(6, p)$, then

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 30 Question Id : 640653565387 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Find the value of p for
which $9P(X = 4) = P(X = 2)$.

Options :

6406531889871. ✔ $\frac{1}{4}$

6406531889872. ✖ $\frac{1}{2}$

6406531889873. ✖

$\frac{1}{8}$

6406531889874. ✖ 0

Question Number : 31 Question Id : 640653565388 Question Type : SA Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

Find $E[X^2]$. Enter the answer correct to two decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

3.35 to 3.41

Question Id : 640653565389 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (32 to 33)

Question Label : Comprehension

Let X and Y be two random variables having joint density function:

$$f_{XY}(x,y) = \begin{cases} kx^2y, & \text{if } 0 < x < 1\ ,\ 0 < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 32 Question Id : 640653565390 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Find the value of *k*.

Options :

6406531889876. ✖ $\frac{1}{6}$

6406531889877. ✖ $\frac{1}{3}$

6406531889878. ✔ 6

6406531889879. ✖ 3

Question Number : 33 Question Id : 640653565391 Question Type : SA Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

Find the value of $P(X + Y < 1)$. Enter the answer correct to one decimal place.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

0.1

Question Id : 640653565392 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (34 to 35)

Question Label : Comprehension

The density function of a continuous random variable X is given by

$$f_X(x) = \begin{cases} \frac{1+\theta x}{2}, & -1 < x < 1, \\ 0, & \text{otherwise}, \end{cases}$$

where $-1 < \theta < 1$.

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 34 Question Id : 640653565393 Question Type : SA Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Short Answer Question

Consider a random sample (−0.2, 0.3, 0.7,−0.6, 0.1). Find the method of moments estimate of θ for the given sample. Enter the answer correct to two decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.16 to 0.20

Question Number : 35 Question Id : 640653565394 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

Consider a random sample $(\frac{-1}{2}, 0, \frac{1}{3}, 0, \frac{1}{3})$. Find the maximum likelihood estimate of θ for the given sample.

Enter the answer correct to two decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.30 to 0.36

Question Id : 640653565395 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (36 to 37)

Question Label : Comprehension

We wish to estimate the probability *p* of getting the number one on a biased die using a Bayesian estimator. Consider 8 independent throws and suppose one appears two times.
Assume the prior distribution of *p* to be Uniform[0, 1].

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 36 Question Id : 640653565396 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction

Time : 0

Correct Marks : 3

Question Label : Multiple Choice Question

Find the posterior distribution of p.

Options :

6406531889883. ✖ $\text{Uniform}[0, 1]$

6406531889884. ✖ $\text{Beta}(2, 6)$

6406531889885. ✔ $\text{Beta}(3, 7)$

6406531889886. ✖ $\frac{1}{4}$ with probability 1

Question Number : 37 Question Id : 640653565397 Question Type : SA Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Short Answer Question

Find the posterior mean. Enter the answer correct to one decimal place.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

0.3

Question Id : 640653565398 Question Type : COMPREHENSION Sub Question Shuffling

Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (38 to 39)

Question Label : Comprehension

A company has provided insurance policies to several people whose age is normal with variance 36. An insurance agent of the company has claimed that the average age of policyholders who are insured through her is 30 years. We suspect that this is too low. To test the agent's claim, we have selected a random sample of 100 policyholders who purchased their policies through her. If the average age of the selected 100 policyholders is greater than 31.1 years, we reject the claim; otherwise, we accept it.

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 38 Question Id : 640653565399 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Define null hypothesis and alternative hypothesis.

Options :

6406531889888. ✔ $H_0 : \mu = 30, H_A : \mu > 30$

6406531889889. ✖ $H_0 : \mu = 30, H_A : \mu \neq 30$

6406531889890. ✖ $H_0 : \mu = 30, H_A : \mu < 30$

6406531889891. ✖ $H_0 : \mu \neq 30, H_A : \mu = 30$

Question Number : 39 Question Id : 640653565400 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

Find the significance level of the test. Enter the answer correct to two decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.02 to 0.04

Sem2 Maths2

Section Id :	64065338384
Section Number :	3
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	15
Number of Questions to be attempted :	15
Section Marks :	50
Display Number Panel :	Yes
Group All Questions :	No
Enable Mark as Answered Mark for Review and Clear Response :	Yes
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	64065380780
Question Shuffling Allowed :	No