

Indian Statistical Institute
Bangalore Centre
B.Math (Hons.) II Year 2018-2019
Mid-semester Examination

Statistics I

Date : 12.09.18

Answer as many questions as possible. The maximum you can score is 60.

The notation used have their usual meaning unless stated otherwise.

1. A market investigator returns with the following data. Of 100 people consulted, 83 like chocolates, 76 liked toffee and 56 liked both. Show that this data can not be correct. [3]
2. (a) Define (i) sample median and (ii) median of a continuous probability distribution. Suppose X is a continuous random variable with a density. Let M denote the median of X . Prove that

$$\text{for any real number } C, E(|X - C|) \geq E(|X - M|).$$

[(2 + 6)]

3. As part of a health study the blood glucose levels of 150 students were measured. The results, in mg/ml correct to 1 decimal place, are summarized in the following table.

Glucose level	< 3.0	3.0-3.9	4.0-4.9	5.0-5.9	6.0-6.9	≥ 7.0
Frequency	7	55	72	10	4	2

(a) Estimate the percentage of students with blood glucose level greater than 5.2. Why is an estimate ?

(b) The number of students with blood glucose level greater than 5.2 is equal to the number with blood glucose level less than a . Find a .

[Your answer may be an algebraic expression involving real numbers; computation is not needed.]

[(5 + 1) + 5 = 11]

4. (a) Define mode of a probability distribution.
- (b) When is a unimodal probability density function said to be symmetric, positively skewed or negatively skewed ? Illustrate with graphs.
- (c) If a unimodal probability density function is negatively skewed, show that both mean and median are smaller than the mode. [1 + 3 + (3 + 4) = 11]

5. (a) Define sample first and third quartiles.

In a study the noise levels of 18 power lawn mowers were recorded. The data obtained are
95 120 117 99 110 107 125 98 85 127 105 114 103 112 92 101 122 120.

Find the 1st, 2nd and 3rd quartiles of the noise levels.

[3 + 4 = 7].

6. In order to study the effects of insulin on reducing blood sugar level in rats, the following experiment was conducted. n rats were selected at random, each one was injected a different dose of insulin (X) and the reduction in blood sugar level (Y) was noted.

Assume the following model.

$$Y_i = \alpha + \beta X_i + \epsilon_i, i = 1, 2, \dots, n.$$

Here ϵ_i 's are independent random variables satisfying $E(\epsilon_i) = 0$ and $Var(\epsilon_i) = \sigma^2$.

- (a) Find least square estimates ($\hat{\alpha}$) of α and ($\hat{\beta}$) of β .
- (b) Show that $E[\hat{\alpha}] = \alpha$ and $E[\hat{\beta}] = \beta$.
- (c) Find the variances of $\hat{\alpha}$ and $\hat{\beta}$. Find $Cov[\hat{\alpha}, \hat{\beta}]$.
- (d) Suppose ϵ_i 's are i.i.d $N(0, \sigma^2)$ variables. Let $\hat{Y}_x = \hat{\alpha} + \hat{\beta}x$, where x is a dose of insulin.

Find the distribution of \hat{Y}_x .

[6 + 7 + 9 + 6 = 28]