# The Co-operative University of Kenya

**END OF SEMESTER EXAMINATIONS**

**EXAMINATION FOR THE DEGREE OF BACHELOR OF AGRIBUSINESS MANAGEMENT**

**COURSE CODE: BSTM 2112**

**UNIT TITLE: BUSINESS STATISTICS**

DATE: DECEMBER, 2019 Time: 2 Hours

## INSTRUCTIONS

1. Answer question **ONE** and any other **TWO (2)** questions
2. Show all your workings
3. Scientific Calculators and non-programmable calculators may be used
4. Use the provided statistical tables where applicable

---

1. (a) Data is collected in various forms. Identify the various categorization/classification of data variables and give specific examples of such variables (**6** *marks*)

   (b) The following table gives probability distribution of marks obtained by some students.

| Range of marks | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 |
|---|---|---|---|---|---|---|
| Probability | 0.093 | 0.1163 | 0.186 | 0.2326 | 0.2093 | 0.1628 |

   What is the probability that a randomly picked student from this class scored

   i. More than 24 marks? (**2** *marks*)

   ii. Between 12 and 24.5 marks? (**2** *marks*)

   (c) Ten teenagers were asked how many hours they sleep at night and their responses recorded. 9,3,4,2,9,5,8,4,7,4. Compute the least number of sleeping hours for the top 20% of the teenagers. (**4** *marks*)

   (d) Differentiate between Descriptive Statistics and Inferential Statistics (**4** *marks*)

   (e) Let X be a random variable representing the amount of water in litres used (lavatory, drinking and hand washing) by an individual University staff per day. Suppose that $X \sim N(37, 7.76)$ and that a sample of size 60 is taken from this distribution.

   i. Find the probability that if a staff member is randomly picked from the sample, he/she consumes between 36.6 and 39.3 litres (**3** *marks*)

   ii. What is the probability that the mean consumption of this sample exceeds 37 litres? (**3** *marks*)

   (f) Emily hits 60% of her free throws in basketball games. She had 25 free throws in last weekâs game.

   i. What is the expected number and the standard deviation of Emilyâs hit ?

   ii. Suppose Emily had 7 free throws in yesterdayâs game.What is the probability that she made at least 5 hits?

   (**6** *marks*)

2. (a) The following table gives the result of experience of three varieties of wheat grown in 10 plots

| | | | | |
|---|---|---|---|---|
| Variety 1 (x1): | 17 | 28 | 24 | 25 |
| Variety 2 (x2): | 21 | 24 | 26 | |
| Variety 3 (x3): | 23 | 30 | 20 | |

Test at $\alpha = 0.5$ whether the varieties are significantly different with regard to yield level. (**10** *marks*)

(b) A researcher is interested in voting preferences on change of the governing constitution in a certain country controlled by two main parties A and B. A questionnaire was developed and sent to a random sample of voters. The cross tabs are as follows

| | | Favour | Neutral | Oppose |
|---|---|---|---|---|
| Membership: | Party A | 60 | 100 | 90 |
| | Party B | 120 | 60 | 70 |

Test at $\alpha = 0.1$ whether party membership and voting preference are associated and state the conditions required for chi-square test results to be valid (**10** *marks*)

3. (a) The table below gives the total number of hours (X) in a day spent by a number of salesmen in marketing a certain product and the corresponding total sales (Y in thousands) due to the salesmen.

| x | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| y | 9.8 | 11.3 | 12.2 | 13.2 | 14.1 | 14.5 | 16.1 |

i. Calculate the product moment correlation coefficient between number of hours and total sales

ii. Fit a regression equation of the form Y = a + bX to this data

iii. Use the equation in (3(b)ii) above to predict the total sales at x = 4.5, hours and x = 10.5 hours hence identify with a reason which of the predicted estimates is more reliable. (**15** *marks*)

(b) The following measures were computed for a frequency distribution:Mean = 50, coefficient of Variation = 35% and Karl Pearson's Coefficient of Skewness $SK_p$ - 0.25. Compute Standard Deviation, Mode and Median of the distribution. (**5** *marks*)

4. (a) For the given data: 9, 3, 4, 2, 9, 5, 8, 4, 7, 4, calculate

i. The Mean

ii. The Median

iii. The Interquartile range

iv. The Mean Absolute Deviation

v. The third quartile

vi. The sixth decile

vii. The eightieth percentile (**14** *marks*)

(b) Below are the scores of two cricketers in 10 innings. By calculating their respective coefficient of variation, compare and comment on the the consistency of the scores of the two cricketers.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Cricketer A | 204 | 68 | 150 | 30 | 70 | 95 | 60 | 76 | 24 | 19 |
| Cricketer B | 99 | 190 | 130 | 94 | 80 | 89 | 69 | 85 | 65 | 40 |

(**6** *marks*)

5. (a) The following table gives probability distribution of amount of money in Ksh withdrawn from an MPESA shop on a certain day.

| Amount (x) | 1300 | 1500 | 2000 | 2300 | 2500 | 3000 |
|---|---|---|---|---|---|---|
| P(X =x) | 0.102 | 0.1429 | 0.2245 | 0.2041 | 0.1837 | 0.1429 |

What is the probability that a randomly picked person from the population deposits

i. More than Ksh. 2300 ?

ii. At least Kshs. 2000?

iii. Between Kshs. 1500 and Kshs. 2500?

iv. What is the expected value and the standard deviation? (**10** *marks*)

(b) Define a time series and briefly explain the main components. (**10** *marks*)

# Solutions to Exam

**1a**. Qualitative variables are variables whose values fall into groups or categories. They are called categorical variables and are further divided into 2 classes namely nominal and ordinal variables. Quantitative variables are numeric variables and are further divided into 2 classes namely discrete and continuous variables

**1c**.

- Required: $P_{20}$. Sorted data: 2,3,4,4,4,5,7,8,9,9
$P_{20} = \frac{20}{100}(\text{n+1})^{th}value$
$\frac{20}{100}(11)^{th}value = 2.2^{nd}value$
$2.2^{nd}value = 2^{nd}value + 0.23^{rd}value - 2^{nd}value = 3 + 0.2(4-3) = 3.2 \therefore The least number of sleeping hours for t$
$3.2hrsr3hours12min.$

**5b**. A time series is any statistical data that is arranged according to the time it was recorded or observed (chronologically).

Components:

- Trend

  These are the long term movements which give the general way in which the data move over a long period of time. A graph of observed value versus time may show small ups and downs, but the long term movement eliminates these minor variations and looks at the big picture.

- Cyclic Movements

  These are movements that happen in regular long-term cycles. In business, for example, cycles consists of alternating periods of recession and inflation, recovery and prosperous times. Cyclic movements are large scale and should be very clear in the data.

- Seasonal Movements

  These are also movements that happen in regular cycles, but repeat yearly. The patterns are caused by either natural conditions such as weather fluctuations or man made conditions such as business, administrative, political procedures, start and end of semester, Easter Holidays, festive seasons etc.

- Random or irregular fluctuations

  These are ups and downs in a time series that do not correspond to trend, seasonal variation or auto correlation. They are unpredictable and as a result off chance events such as strikes, floods, earthquakes, plane clash, post election violence etc.