ZIMBABWE EZEKIEL GUTI UNIVERSITY

Faculty of Heritage, Humanities and Societal Advancement

EXAMINATION PAPER

COURSE CODE	:	BSSW223
COURSE TITLE	:	SOCIAL RESEARCH METHODS 2
SPECIAL REQUIREMENTS	:	Formulae Sheets
		Graph Paper
DURATION	:	3 Hours
LEVEL	:	
DATE	:	12 FEB 2025

INSTRUCTIONS TO CANDIDATES:

1. No cell phones are allowed in the examination venue.
2. Section A is compulsory and answer any THREE (3) questions from Section B.
3. The number of marks for each question or part question is shown in brackets ( )
4. Use of non-programmable calculators is allowed.
5. Show all your workings in order to gain full marks.
6. Begin each answer on a new page.
7. **DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR INSTRUCTS YOU.**

QUESTION ONE

SECTION A (ANSWER ALL QUESTIONS FROM THIS SECTION)

QUESTION ONE

The ZEGU Authorities wants to gather opinions from students on the quality of food served at the Campus Dining Hall.

i. Discuss reasons why a sample survey is preferred to a census. (10 marks)

ii. Identify and select with justification an appropriate sampling method for the study. (5 marks)

iii. Identify and select with justification an appropriate methodology for the study. (4 marks)

iv. Identify and select with justification appropriate data collection tools/instruments for the study. (3 marks)

v. Suggest any three ethical considerations that the ZEGU Authorities may want to take into account in conducting the survey. (3 marks)

SECTION B (ANSWER ANY THREE QUESTIONS FROM THIS SECTION)

QUESTION TWO

Answer what is asked in the problem.

a. In order to assist management negotiations with the trade unions over piecework rates, the management services department of a factory is asked to obtain information on how long it takes for a certain operation to be completed. Consequently, the members of the department measure the time it takes to complete 30 repetitions of the operation, at random occasions during a month. The times are recorded to the nearest tenth of a minute.

19	21	24	18	19	15	20	22	19	17	28	23
20	18	19	16	22	18	20	21	19	25	22	17
21	20	19	21	15	23						

i. Calculate the mean for these data (3 marks)

ii. Calculate the median for these data (3 marks)

iii. Find the mode for these data (3 marks)

b. Table below gives the frequency distribution of gasoline prices frequency for petrol prices at 48 stations in a town.

Price ($)	Frequency
1.00-1.05	4
1.05-1.10	6
1.10-1.15	10
1.15-1.20	15
1.20-1.25	8
1.25-1.30	5

i. Complete the frequency distribution table by finding the midpoint (r), fx and cumulative frequency. (4 marks)

ii. Work on the mean of the data (4 marks)

iii. Work out the median of the data (4 marks)

iv. Work out the mode of the data (4 marks)

QUESTION THREE

Suppose that college students are asked to identify their preferences in political affiliation (ZANU PF, MDC T) and in ice cream (chocolate, vanilla, or strawberry). Suppose that their responses are represented in the following table. Test the independence between political party affiliation and ice cream type preference. (25 marks)

	Chocolate	**Vanilla**	**Strawberry**
ZANU PF	40	50	30
MDC T	25	43	24
ZAPU	43	46	30

QUESTION FOUR

A researcher wants to study the opinions of high school students in a large city on a new school mental health policy. Which sampling techniques could be used and why would one be more appropriate than the others. (25 marks)

QUESTION FIVE

Knowing that load shedding of electricity has an impact on students' exam scores, Dr Shumba hypothesized that students who live in Chiwaridzo with the most of average hours of load shedding are likely to have low ratings on quality of work produced (thus predicting a negative or inverse relationship between these variables in the population.

Test the hypothesis that there is no rank order relationship between the average hours of load shedding (loadshed) and the quality rating of students' work (work quality). (25 marks)

Student	Load shed (hours)	Work quality (on scale of 1 poor to 10 (good)
1	13	3
2	5	6
3	8	3
4	24	3
5	21	7
6	30	9
7	20	10
8	20	6
9	12	5
10	7	7
11	8	4
12	25	7

QUESTION SIX

Using practical example, write short notes on any **FIVE (5)** of the following:

i. Measures of Central Tendency (5 marks)
ii. Nominal data (5 marks)
iii. Descriptive statistics (5 marks)
iv. Survey research (5 marks)
v. Variance (5 marks)
vi. Sample (5 marks)

THE END

STATISTICAL FORMULAS

MEASURES OF CENTRAL TENDENCY

Ungrouped data

Population mean, $\mu = \frac{\sum x}{N}$

Sample mean, $\bar{x} = \frac{\sum x}{n}$

Median, $M_e = \frac{n+1}{2}$

Grouped data

Population mean, $\mu = \frac{\sum fx}{N}$

Sample mean, $\bar{x} = \frac{\sum fx}{n}$

Median, $Me = L_m + \frac{(\frac{n}{2} - F)C}{fm}$

Mode, $Mo = L_m + \frac{(d_1)\, c}{d_1 + d_2}$

MEASURES OF DISPERSION

Ungrouped data

Population average deviation, $AD = \frac{\sum|x-\mu|}{N}$

Sample average deviation, $AD = \frac{\sum|x-\bar{x}|}{n}$

Population variance $= \sigma^2 = \frac{1}{N}(\sum x^2 - \frac{1}{N}(\sum x)^2)$

Sample variance $= s^2 = \frac{1}{n-1}(\sum x^2 - \frac{1}{n}(\sum x)^2)$

Population standard deviation, $\sigma = \sqrt{\frac{1}{N}(\sum x^2 - \frac{1}{N}(\sum x)^2)}$

Sample standard deviation, $s = \sqrt{\frac{1}{n-1}(\sum x^2 - \frac{1}{n}(\sum x)^2)}$

Grouped data

Range = maximum value – minimum value

Percentile, $P_k = L_p + \frac{(\frac{kn}{100} - F)C}{fp}$

Decile, $D_x = L_d + \frac{(\frac{xn}{10} - F)C}{fd}$

Lower quartile, $Q_1 = L_q + \frac{(\frac{n}{4} - F)C}{fq}$

Upper quartile $Q_3 = L_q + \frac{(\frac{3n}{4} - F)C}{fq}$

Interquartile range = $Q_3 - Q_1$

Quartile deviation = $\frac{Q_3 - Q_1}{2}$

Population average deviation, AD= $\frac{\sum f|x-\mu|}{N}$

Sample average deviation, AD = $\frac{\sum f|x-\bar{x}|}{n}$

Population variance, $\sigma^2 = \frac{1}{N}(\sum fx^2 - \frac{1}{N}(\sum fx)^2)$

Sample variance= $s^2 = \frac{1}{n-1}(\sum fx^2 - \frac{1}{n}(\sum fx)^2)$

Population standard deviation, $\sigma = \sqrt{\frac{1}{N}(\sum fx^2 - \frac{1}{N}(\sum fx)^2)}$

Sample standard deviation, $s = \sqrt{\frac{1}{n-1}(\sum fx^2 - \frac{1}{n}(\sum fx)^2)}$

Population coefficient of variation, CV= $\frac{\sigma}{\mu}$ x 100%

Sample coefficient of variation, CV = $\frac{s}{\bar{x}}$ x 100%

SHAPE OF FREQUENCY DISTRIBUTIONS

Population skewness, Sk = $\frac{3(\mu - \text{median})}{\sigma}$ or $\frac{(\mu - \text{mode})}{\sigma}$

Sample skewness, Sk = $\frac{3(\bar{x} - \text{median})}{s}$ or $\frac{(\bar{x} - \text{mode})}{s}$

Population kurtosis = $\frac{\sum f(X-\mu)^4}{\sigma^4}$

Sample kurtosis = = $\frac{\sum f(X-\bar{x})^4}{s^4}$

CONFIDENCE INTERVALS

Mean of a single normal population

i. **If sample size is small (i.e., n<30) and σ^2 Unknown**

$$\bar{x} - t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}} \text{ or simply } \bar{x} \pm t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}$$

ii. **If sample size is large and σ^2 unknown**

$$\bar{x} - Z_{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\alpha/2}\frac{s}{\sqrt{n}} \text{ or simply, } \bar{x} \pm Z_{\alpha/2}\frac{s}{\sqrt{n}}$$

iii. **If variance, σ^2 is known**

$$\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \text{ or simply, } \bar{x} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

TESTING OF HYPOTHESIS

Tests on the mean of a single population

i. If when σ^2 is unknown and sample size is small (n<30), then the t Statistic is such that,

$$t = \frac{\bar{x} - \mu_x}{\frac{s}{\sqrt{n}}}$$

ii. If when σ^2 is known and sample size is large, then the z statistic is such that:

$$z = \frac{\bar{x} - \mu_x}{\sigma_{\bar{x}}} = z = \frac{\bar{x} - \mu_x}{\frac{\sigma}{\sqrt{n}}}$$

iii. If when σ^2 is unknown and sample size is large, then the t statistic is approximately a standard normal random variable such that: $z = \frac{\bar{x} - \mu_x}{\frac{s}{\sqrt{n}}}$

Hypothesis to be tested	Distribution	Reject H_0 if
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	Z -distribution	$Z > Z_{\alpha/2}$ $Z < -Z_{\alpha/2}$
	T - distribution	$Z > Z_{\alpha/2}$ $T < -t_{\alpha/2}(n-1)$
$H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$	Z-distribution T-distribution	$Z > Z_\alpha$ $T > t_\alpha (n-1)$

$H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$	Z-distribution T-distribution	$Z < -Z_\alpha$ $T < -t_\alpha\,(n-1)$

Chi-square tests: $\chi^2_{calc} = \sum \frac{(f_o - f_e)^2}{f_e}$

SIMPLE REGRESSION AND CORRELATION

Fitted linear model: $\hat{Y} = \hat{b}_0 + \hat{b}_1 X$

$$\hat{b}_1 = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2}$$

$$\hat{b}_0 = \overline{Y} - \hat{b}_1 \overline{X}$$

Covariance, $Cov(X, Y) = \frac{\sum (XY)}{n} - \overline{X}\,\overline{Y}$

Correlation coefficient, $r = \sqrt{R^2} = \frac{n\sum XY - \sum X \sum Y}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$

Coefficient of determination, $R^2 = \left(\frac{n\sum XY - \sum X \sum Y}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}\right)^2$

1.PM