



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
Statistical Tables
Graph Paper
DURATION : 3 Hours
LEVEL : 2.2
DATE : 29 JUL 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.**

SECTION A (ANSWER ALL QUESTIONS FROM THIS SECTION)

QUESTION ONE

For each research situation below:

- (a) Identify the independent and dependent variables and
- (b) Classify the level of measurement of each variable.
- (c) Decide the statistical applications to be used: descriptive statistics (single variable), descriptive statistics (two or more variables), or inferential statistics

- i. A graduate student is studying sexual harassment on college campuses and asks 500 female students if they personally have experienced any such incident. Each student is asked to estimate the frequency of these incidents as either often, sometimes, rarely, or never. The researcher also gathers data on age and major to see if there is any connection between these variables and frequency of sexual harassment. **[5 marks]**
- ii. The adult bookstore near campus has been raided and closed by the police. Your social research class has decided to poll the student body and get their reactions and opinions. The class decides to ask each student if he or she supports or opposes the closing of the store, how many times each one has visited the store, and if he or she agrees or disagrees that "pornography is a direct cause of sexual assaults on women." The class also collects information on the sex, age, religious and political philosophy, and major of each student to see if opinions are related to these characteristics. **[5 marks]**
- iii. For a research project in a social science course, a student has collected information about the quality of life and the degree of political democracy in 50 nations. Specifically, she used infant mortality rates to measure quality of life and the percentage of all adults who are permitted to vote in national elections as a measure of democratization. Her hypothesis is that quality of life is higher in more democratic nations. **[5 marks]**
- iv. A highway engineer wonders if a planned increase in speed limit on a heavily traveled local avenue will result in any change in number of accidents. He plans to collect information on traffic volume, number of accidents, and number of fatalities for the six-month periods before and after the change. **[5 marks]**

- v. Students are planning a program to promote safe sex and awareness of a variety of other health concerns for college students. To measure the effectiveness of the program, they plan to give a survey measuring knowledge about these matters to a random sample of the student body before and after the program. **[5 marks]**

SECTION B (ANSWER ANY THREE QUESTIONS FROM THIS SECTION)

QUESTION TWO

- a) Nationally, it is claimed that the average annual salary of social workers is more than \$24 509. A random sample of 144 social workers in Harare yielded a sample average annual salary of \$25 782 with a standard deviation of \$622. Conduct a test at the 5% level of significance to determine if the average annual salary is more than \$24 509. **[20 marks]**
- b) A parkade attendant would like to select a sample of car drivers to take part in a study commissioned by the city council. Give five reasons why the city council opted to use a sample of drivers in the study instead of the whole population of drivers in the city. **[5 marks]**

QUESTION THREE

- a) Suppose that ZEGU students are asked to identify their major personality types of lecturers (Agreeableness, Conscientiousness and Neuroticism) and lecturer warmth-competence (Warm-Competent, Cold-Competent or Cold Incompetent) and indicate their preferences. Suppose that their responses are represented in the following table.

	Warm-Competent	Cold-Competent	Cold Incompetent
Agreeableness	50	37	26
Conscientiousness	40	45	24
Neuroticism	28	46	48

Test the independence of personality type and warmth competence. **[20 marks]**

- b) Social workers are interested in understanding the level of satisfaction among clients with the services provided by a community support center. Responses from a random

sample of clients were as follows: 69 were very satisfied, 55 were moderately satisfied, 5 had no opinion, 3 were moderately dissatisfied, and 2 were very dissatisfied. Using a suitable chart, show the above-mentioned information. **[5 marks]**

QUESTION FOUR

- a) An examiner and a moderator marked 7 examination scripts during a standardization process and awarded the following percentage scores:

Script Number	1	2	3	4	5	6	7
Examiner	67	54	38	70	42	70	80
Moderator	58	60	38	67	44	69	76

Calculate Spearman's rank correlation coefficient and comment on the result obtained. **[10 marks]**

- b) Social workers are interested in understanding the distribution of farms across different provinces to better plan community support and agricultural programs. The frequency distribution shown represents the number of farms per province for the 50 provinces based on information from the Department of Agriculture:

Number of Farms per Province	Frequency
0 – under 200	16
200 – under 400	11
400 – under 600	11
600 – under 800	6
800 – under 1000	4
1000 – under 1200	2

- i. Determine the average number of farms and standard deviation per province from these data. **[10 marks]**
- ii. The mean computed from the original ungrouped data was 417.96 and the standard deviation was 388.56. Compare and contrast the answers for grouped data and ungrouped data. Why might they differ? **[5 marks]**

QUESTION FIVE

A social worker is interested in knowing whether the number of social work cases in a community can be predicted by the unemployment rate. The following data represents a sample of a month's social work cases and the unemployment rate for that month:

Unemployment Rate (%)	Social Work Cases
8.5	220
3.2	60
5.5	110
6.0	130
4.0	70
7.5	190
6.8	160
5.2	115

- Plot the scatter diagram and comment on the result. **[5 marks]**
- Develop a simple regression line to predict the number of social work cases in a month by the unemployment rate for that month. **[8 marks]**
- What would be the predicted number of social work cases for an unemployment rate of 7%? **[2 marks]**
- Calculate the coefficient of determination of the data and comment. **[5 marks]**
- Calculate the correlation coefficient of the data and comment. **[5 marks]**

THE END

STATISTICAL FORMULAS

MEASURES OF CENTRAL TENDENCY

Ungrouped data

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

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

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

Grouped data

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

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

$$\text{Median, } M_e = L_m + \frac{\left(\frac{n}{2} - F\right)C}{f_m}$$

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

MEASURES OF DISPERSION

Ungrouped data

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

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

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

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

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

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

Grouped data

Range = maximum value – minimum value

$$\text{Percentile, } P_k = L_p + \frac{\left(\frac{kn}{100} - F\right)C}{f_p}$$

$$\text{Decile, } D_x = L_d + \frac{\left(\frac{xn}{10} - F\right)C}{f_d}$$

$$\text{Lower quartile, } Q_1 = L_q + \frac{\left(\frac{n}{4} - F\right)C}{f_q}$$

$$\text{Upper quartile } Q_3 = L_q + \frac{\left(\frac{3n}{4} - F\right)C}{f_q}$$

Interquartile range = $Q_3 - Q_1$

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

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

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

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

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

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

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

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

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

SHAPE OF FREQUENCY DISTRIBUTIONS

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

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

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

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

BASIC PROBABILITY CONCEPTS

- i. **Classical Method of Assigning Probabilities:** $P(E) = \frac{ne}{N}$
- ii. **Probability by Relative Frequency of Occurrence**

$$\frac{\text{Number of Times an Event Occurred}}{\text{Total Number of Opportunities for the Event to Occur}}$$
- iii. **Complementary rule:** $P(A') = 1 - P(A)$
- iv. **Addition rule**
 - a) If X, Y are non-mutually exclusive, $P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$
 - b) If X, Y are mutually exclusive, $P(X \cup Y) = P(X) + P(Y)$
- v. **Multiplication rule**
 - a) If X, Y are dependent (conditional probability),

$$P(X \cap Y) = P(X) \cdot P(Y|X) = P(Y) \cdot P(X|Y)$$
 - b) If X, Y are independent, $P(X \cap Y) = P(X) \cdot P(Y)$
- vi. **Law of Conditional Probability,** $P(X | Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{P(X) \cdot P(Y | X)}{P(Y)}$

PROBABILITY DISTRIBUTION

Probability distribution	Mean	Variance
Binomial distribution $P(x) = \frac{n!}{x!(n-x)!} p^x (q)^{n-x}$	np	npq
Poisson distribution $P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$	λ	λ
Standard normal distribution $Z = \frac{x - \mu}{\sigma}$	0	1

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 (ONE SAMPLE CASE)

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:

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

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
$H_0: \mu = \mu_0$	Z -distribution	$Z > Z_{\alpha/2}$ $Z < - Z_{\alpha/2}$
$H_1: \mu \neq \mu_0$	T - distribution	$T > T_{\alpha/2} (n - 1)$ $T < - t_{\alpha/2}(n-1)$
$H_0: \mu \leq \mu_0$	Z-distribution	$Z > Z_{\alpha}$
$H_1: \mu > \mu_0$	T-distribution	$T > t_{\alpha} (n-1)$
$H_0: \mu \geq \mu_0$	Z-distribution	$Z < - Z_{\alpha}$
$H_1: \mu < \mu_0$	T-distribution	$T < - t_{\alpha} (n-1)$

TESTING OF HYPOTHESIS (TWO SAMPLE CASE)

Comparing Two Population Means when samples are from independent populations

Use if sample sizes > 30 or if σ_1 and σ_2 are known: $z = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

Use if sample sizes > 30 or if σ_1 and σ_2 are unknown: $z = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

Comparing Population Means with Unknown Population Standard Deviations (the Pooled t-test)

Small sample test of means ($n < 30$)

The required assumptions are: (1) Both populations must follow the normal distribution. (2) The populations must have equal standard deviations and (3) The samples are from independent populations.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \text{ where } S_p^2 = S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Degrees of freedom (df) : $n_1 + n_2 - 2$

Two-Sample Tests of Hypothesis: Dependent Samples

Use the following test when the samples are dependent:

$$t = \frac{\bar{d}}{s_d / \sqrt{n}}$$

Degrees of freedom (df): $n - 1$

CHI-SQUARE TEST

$$\chi^2_{\text{calc}} = \sum \frac{(f_o - f_e)^2}{f_e}$$

Degrees of freedom (df) = (number of rows - 1) × (number of columns - 1)

SIMPLE REGRESSION AND CORRELATION

Fitted Trend Line: $Y = a + bX$

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

$$a = \bar{Y} - b \bar{X}$$

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

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

BUSINESS CALCULATIONS

Quadratic formula: $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

Differentiation

- i. if $f(x) = x^n$ then $f'(x) = nx^{n-1}$; if $y = x^n$ then, $\frac{dy}{dx} = nx^{n-1}$
- ii. Rule 1: The constant rule. If $h(x) = cf(x)$ then $h'(x) = cf'(x)$ for any constant c .
- iii. Rule 2: The sum rule. If $h(x) = f(x) + g(x)$ then $h'(x) = f'(x) + g'(x)$
- iv. Rule 3: The difference rule. If $h(x) = f(x) - g(x)$ then $h'(x) = f'(x) - g'(x)$
- v. Rule 4: The chain rule. If y is a function of u , which is itself a function of x , then

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$$

- vi. Rule 5: The product rule. If $y = uv$ then $\frac{dy}{dx} = u \frac{dv}{dx} + v \frac{du}{dx}$
- vii. Rule 6: The quotient rule. If $y = \frac{u}{v}$ then $\frac{dy}{dx} = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2}$
- viii. Implicit differentiation: $\frac{dy}{dx} = -\frac{f_x}{f_y}$