JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF HEALTH SCIENCES

UNIVERSITY EXAMINATION FOR DEGREE OF BACHELOR OF SCIENCE

PUBLIC HEALTH / COMMUNITY HEALTH AND DEVELOPMENT

2ND YEAR 1ST SEMESTER 2022/2023 ACADEMIC YEAR

MAIN CAMPUS

COURSE CODE: SBB 1409

COURSE TITLE: BIOSTATISTICS II

EXAM VENUE: **STREAM: (BSc. Env. Hlth/ Comm Hlth & Dev)**

DATE: **EXAM SESSION: DECEMBER 2022**

TIME: 2.00 HOURS

_____

Instructions:

1. **Answer the questions**

    Section A consists of 10 multiple-choice Question (MCQ) and amounts to 10 marks.
    Section B consists of 4 short answer questions (SAQ) totalling to 20 marks.
    Section C consists of 3Long essay questions (SAQ) totalling to 40 marks.

2. **Candidates are advised not to write on the question paper.**

3. **Candidates must hand in their answer booklets to the invigilator while in the examination room.**

Section A, Answer all questions in this section

1. The stages of a malignant disease (cancer) are recorded using the symbols 0, I, II, III, IV. We say that the scale used is?
   A. Alphanumeric
   B. Numerical
   C. Ordinal
   D. Nominal

2. If the average of a series of values is 10 and their variance is 4, then the coefficient of variation
   (= the ratio standard deviation / average) is:
   A.40%
   B.20%
   C.80%
   D.10%
   *Standard deviation is square root of variance $\sqrt{4} = 2$, so the coefficient of variation is 2/10=0.2=20%

3. The first quartile of a series of values is:
   A. The value in the ordered series located at 25% above the median
   B. The value of the ordered series located at 75% of the number of values in the series
   C. The numeric value for which a quarter of the series' values are lower
   D. The numeric value for which a quarter of the series' values are higher

4. In a health care utilization journal, results are reported from a study performed on a random sample of 100 deliveries at a large teaching hospital. The sample mean birth weight is reported as 120 ounces, and the sample standard deviation is 25 ounces. The researchers neglected to report a 95% confidence interval for the population birth weight (i.e.: mean birthweight for all deliveries in the hospital). You decide to do so, and find the 95% confidence interval for the population mean birth weight to be:
   A. 119.5 ounces to 120.5 ounces
   B. 115 ounces to 125 ounces
   C. 70 ounces to 170 ounces
   D. 117.5 ounces to 122.5 ounces

5. The standard error of a statistic is?
   A. the mean of the sampling distribution
   B. the standard deviation of the sampling distribution
   C. the statistic divided by the square root of the sample size

6. If on a group of 457 patients, for a risk factor we calculated a Relative Risk RR= 12.74, the possibility of developing the disease being investigated is?
   A. very high when exposed to the factor
   B. very small when exposed to the factor
   C. the same in the case of exposure in the case of non-exposure
   D. lower in the exposed than in the unexposed, RR being less than 100

7. Suppose a random sample of 100 12-year-old boys were chosen and the heights of these 100 boys recorded. The sample mean height is 64 inches, and the sample standard deviation is 5 inches. You may assume heights of 12-year-old boys are normally distributed. Calculate the 95% CI of the heights of 12-year-old boys [1 mark]

   A. 59 inches to 69 inches
   B. 64 inches
   C. 54 inches to 74 inches
   D. None of the above

8. A regression line is a straight line which:
   A. is located as close as possible to all the points of a scatter chart

B. is defined by an equation having 2 parameters: the slope and the intercept
C. provides an approximate relationship between the values of two parameters
D. All of the above

9. Best measure for central tendency in a highly skewed distribution is ?
   A. Median
   B. Mean
   C. Mode
   D. Range

10. Females and males are examples of what kind of variables?
    A. Continuous
    B. Ratio
    C. Ordinal
    D. Nominal

Section B: Answer all questions in this section

1. A study is conducted concerning the blood pressure of 60-year-old women with glaucoma. In the study 200 60-year-old women with glaucoma are randomly selected and the sample mean systolic blood pressure is 140 mm Hg and the sample standard deviation is 25 mm Hg. Calculate a 95% confidence interval for the true mean systolic blood pressure among the population of 60-year-old women with glaucoma. [5 marks]

2. Indicate whether the following statements are true/false. You do not need to give reasons for your choice **[5 marks]**

   a) Gauss curve is defined as: Symmetrical to the vertical axis, which passes through**.**
   b) Categorical date is the name is given to data which can be ranked
   c) Type I is an error made when one fails to reject the null hypothesis when it is false is:
   d) Mean of statistics is unaffected by outliers?
   e) The mean is, the middlemost score.

3. Most studies conducted on groups of people, uses both descriptive and inferential statistics to analyze results and draw conclusions.
   Briefly define the following terms: **[5 marks]**
   a) Descriptive statistic:
   b) Inferential statistics:

4. There are 2.5 million people aged 75-84 in Country X, which has a total population of 25 million. 4000 people died of influenza and pneumonia in Country X in 2011. 2,540 of these deaths occurred in the 75-84 age group. What is the 2011 mortality rate for influenza and pneumonia among people in Country X aged 75-84?

Section C: Answer any two questions in this section

1. In a study 20 'normal' people take the standard treatment for bad breath (drug A). 20 garlic eaters take drug B. The results of the study indicate that the people taking drug A have better breath. Is drug A better than drug B? [1 mark]

   A. What are the potential sources of bias? [3 marks]
   B. Solutions to the problem of bias [8 marks]
   C. How do you Control for potential confounders? [8 marks]

2. Use the information below to answer question 3A-E

HIV Acquisition During Pregnancy and Postpartum is Associated with Genital Infections and Partnership Characteristics: A Cohort Study. John KINUTHIA *et al., AIDS 2015 September 24: 29(15): 2025*

Objective—To determine the risk and cofactors for HIV acquisition during pregnancy andpostpartum.

Methods—Pregnant women in western Kenya were enrolled if HIV seronegative at that visit or within 3 months. Serial HIV nucleic acid amplification tests (NAATs) were conducted at 1–3 month intervals to 9 months postpartum. Genital swabs were collected for detection of chlamydia and gonorrhea at baseline, and for trichomonas, bacterial vaginosis (BV), and yeast at baseline andfollow-up.

Results—Among 1304 pregnant women, median age was 22 years, 78% were married for a median of 4 years, 66% reported knowing partner HIV status, and 8% reported using condoms. Study retention was 98%. During 1235 person-years of follow-up, HIV incidence was 2.31/100 person-years (95% Confidence Interval [CI]:0.71–4.10). Incident HIV was associated with syphilis (Hazard Ratio [HR] 9.18, 95% CI:2.15–39.3), chlamydia (HR 4.49, 95% CI:1.34–15.0),

   A. What is the most appropriate study design used for this study **[1 mark]**
   B. What is the target population to which results can be generalized [**1 mark]**
   C. Briefly explain two possible bias that be encountered while conducting this study **[4 marks]**
   D. With examples, briefly explain strengths and limitation of this study designs **[14 marks]**

3. A biostatistics lecturer makes out his final grades for 200 students in his statistics class. He is curious to see if his grade distribution resembles the "normal curve" and notes from the college catalog that in a normal distribution of grades 45% of them would be C's, 24% of them would be B's, 24% D's, 3.5% of them would be A's, and 3.5% F's. The lecturer compared the frequency of grades given in his class to the normal curve. The frequency of each grade is given as follows: 15, 53, 87, 33,12

   Note: Null hypothesis for this exercise is: The grade distribution used by the instructor will be the same as the distribution described in the collegecatalog.

   Use this statement to answer question A

   A.   Calculate Chi-square for this exercise and interpret your results [20 marks]

Table of the chi square distribution

Level of Significance α

df	0.200	0.100	0.075	0.050	0.025	0.010	0.005	0.001	0.0005
1	1.642	2.706	3.170	3.841	5.024	6.635	7.879	10.828	12.116
2	3.219	4.605	5.181	5.991	7.378	9.210	10.597	13.816	15.202
3	4.642	6.251	6.905	7.815	9.348	11.345	12.838	16.266	17.731
4	5.989	7.779	8.496	9.488	11.143	13.277	14.860	18.467	19.998
5	7.289	9.236	10.008	11.070	12.833	15.086	16.750	20.516	22.106
6	8.558	10.645	11.466	12.592	14.449	16.812	18.548	22.458	24.104
7	9.803	12.017	12.883	14.067	16.013	18.475	20.278	24.322	26.019
8	11.030	13.362	14.270	15.507	17.535	20.090	21.955	26.125	27.869
9	12.242	14.684	15.631	16.919	19.023	21.666	23.589	27.878	29.667
10	13.442	15.987	16.971	18.307	20.483	23.209	25.188	29.589	31.421
11	14.631	17.275	18.294	19.675	21.920	24.725	26.757	31.265	33.138
12	15.812	18.549	19.602	21.026	23.337	26.217	28.300	32.910	34.822
13	16.985	19.812	20.897	22.362	24.736	27.688	29.820	34.529	36.479
14	18.151	21.064	22.180	23.685	26.119	29.141	31.319	36.124	38.111
15	19.311	22.307	23.452	24.996	27.488	30.578	32.801	37.698	39.720
16	20.465	23.542	24.716	26.296	28.845	32.000	34.267	39.253	41.309
17	21.615	24.769	25.970	27.587	30.191	33.409	35.719	40.791	42.881
18	22.760	25.989	27.218	28.869	31.526	34.805	37.157	42.314	44.435
19	23.900	27.204	28.458	30.144	32.852	36.191	38.582	43.821	45.974
20	25.038	28.412	29.692	31.410	34.170	37.566	39.997	45.315	47.501
21	26.171	29.615	30.920	32.671	35.479	38.932	41.401	46.798	49.013
22	27.301	30.813	32.142	33.924	36.781	40.289	42.796	48.269	50.512
23	28.429	32.007	33.360	35.172	38.076	41.639	44.182	49.729	52.002
24	29.553	33.196	34.572	36.415	39.364	42.980	45.559	51.180	53.480
25	30.675	34.382	35.780	37.653	40.646	44.314	46.928	52.620	54.950
26	31.795	35.563	36.984	38.885	41.923	45.642	48.290	54.053	56.409
27	32.912	36.741	38.184	40.113	43.195	46.963	49.645	55.477	57.860
28	34.027	37.916	39.380	41.337	44.461	48.278	50.994	56.894	59.302
29	35.139	39.087	40.573	42.557	45.722	49.588	52.336	58.302	60.738
30	36.250	40.256	41.762	43.773	46.979	50.892	53.672	59.704	62.164
40	47.269	51.805	53.501	55.759	59.342	63.691	66.766	73.403	76.097
50	58.164	63.167	65.030	67.505	71.420	76.154	79.490	86.662	89.564
60	68.972	74.397	76.411	79.082	83.298	88.380	91.952	99.609	102.698
70	79.715	85.527	87.680	90.531	95.023	100.425	104.215	112.319	115.582
80	90.405	96.578	98.861	101.880	106.629	112.329	116.321	124.842	128.267
90	101.054	107.565	109.969	113.145	118.136	124.117	128.300	137.211	140.789
100	111.667	118.498	121.017	124.342	129.561	135.807	140.170	149.452	153.174