Section A: Answer all questions. Each question weighs 10 marks

1. What **statistical test** would you use to analyse the following?
   a) To determine an **association** between two **categorical variables [2 marks]**
   b) To determine an association between two independent **categorical data** sets when counts are less than 5 **[2 marks]**
   c) To compare two independent (unrelated) observations of **continuous data** sets that are **not** normally distributed **[2 marks]**
   d) To determine whether there are any statistically significant differences between the means of two or more independent groups **[2 marks]**
   e) To **predict a value** from **several** measured variables **[2 marks]**

2. Briefly discuss the two common measures of disease incidence **[10 marks]**

3. What is meant by herd immunity? Briefly describe the three epidemiological concepts of disease **[10 marks]**

4. The Director for Laboratory Services at the Ministry of Health would like to set up an interview survey of a sample of Chief Biomedical Scientists in-charge of different hospital laboratories in Zambia. The hospitals are located in the 10 provinces of the country, and are classified into three types according to size of hospital. A sample of hospitals is to be taken, and samples of Chief Biomedical Scientists will be drawn from these hospitals.
   a) Suggest how cluster sampling could be used to choose the sample of hospitals. Discuss any potential advantages and disadvantages of this method for the proposed survey **[5 marks]**
   b) Suggest how stratified sampling could be used to choose the sample of hospitals. Discuss any potential advantages and disadvantages of this method for the proposed survey. **[5 marks]**

Herd immunity is the most probable, a group will develop the disease (epidemic threat) because the proportion of immune individuals reduces the chance of contact btw infected & susceptible individuals.

The entire population does not need to be immunized to prevent disease transmission. to become epidemic.

t-test formula

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where $s = \sqrt{\left[(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2\right]/(n_1 + n_2 - 2)}$

t-distribution table

Table 3 Percentage points of the t distribution ($t_{\alpha,\nu}$)

Degrees of freedom, ν	α								
	75	80	85	90	95	975	99	995	9995
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.598
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073

0.05 0.820

End of Examination

Section B: Answer three (3) questions. Each question weighs 20 marks.

The following is a dataset from a two-by-two longitudinal clinical research study to evaluate changes in brain glutathione (GSH), an important brain antioxidant, using proton magnetic resonance spectroscopy (¹H MRS) in HIV-infected methamphetamine drug users (HIV+Meth), methamphetamine users (Meth), HIV-infected patients (HIV) and controls (Control).

Patient ID	Group	Age at baseline (yr 1)	Sex	GSH baseline (yr 1)	GSH (yr 2)	GSH (yr 3)	GSH (yr 4)	GSH (yr 5)
10001	HIV+Meth	25	M	6.2mM	6.0mM	5.8mM	7.8mM	6.1mM
10002	HIV	32	M	4.2mM	5.2mM	4.1mM	4.0mM	3.5mM
10003	HIV	49	F	6.5mM	6.5mM	6.2mM	6.1mM	6.6mM
10004	Meth	25	F	6.2mM	7.3mM	5.5mM	5.2mM	5.1mM
10005	Meth	26	M	6.8mM	6.8mM	6.8mM	6.5mM	6.8mM
10006	HIV+Meth	32	M	8.2mM	7.2mM	6.2mM	7.5mM	6.2mM
10007	Control	57	M	6.3mM	6.0mM	6.0mM	6.7mM	6.9mM
10008	Control	35	M	9.2mM	9.5mM	9.5mM	9.4mM	9.0mM
10009	HIV+Meth	38	F	3.4mM	3.4mM	3.4mM	3.4mM	3.4mM
10010	Control	25	F	7.5mM	7.5mM	7.6mM	7.6mM	7.5mM

a) You perform a one-way ANOVA to determine if there is a statistical difference in glutathione levels across the four groups at baseline and obtain the following one-way ANOVA table below. Make a statistical decision about whether the levels of glutathione are different or not across the four groups at baseline? **[10 marks]**

ANOVA Table for GSH (yr 1)

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Group	3	7.667	2.556	820	5285	2.460	142
Residual	6	19.695	3.175				

b) Suggest a statistical test you would use to test the hypothesis that there is a decrease in levels of glutathione in HIV-infected methamphetamine users **across** the five year period **[4 marks]**

c) Name the statistical test you would use to determine if there is a relationship/association between glutathione and age. What values does the test give you and briefly discuss how the values help you interpret that relationship? **[6 marks]**

The **sample size** for any research study is affected by **four** factors. Name and briefly discuss each of the four factors and how each can affect sample size (i.e. whether you would need a smaller or larger sample size). **[20 Marks]**

University of Zambia
School of Medicine

2015/16 Deferred and Supplementary Examinations

BSc Biomedical Sciences

Fourth Year

Epidemiology and Biostatistics - BMS 4415

Date: Tuesday, 27th September, 2016 **Time:** 09:00 to 12:00 hours

Venue: Pharmacy Hall

Instructions

1. The paper consists of eight **(8)** questions. You should answer **all** the four **(4)** questions from Section A and three **(3)** questions from Section B. This paper weighs **60%** of the final examination score for those seating for Deferred and **100%** for those seating for Supplementary Examinations.

2. Each question must be answered in a **SEPARATE** answer booklet. Ask for additional booklets, if required. All Questions in Section A carry 10 marks each and all Questions from Section B carry 20 marks each.

3. It is **ESSENTIAL** that you indicate your Student Number, the section and number of each question you have attempted on the cover of every booklet.

4. Complete answer booklets should be handed in, all tied together, and will be collected BEFORE you leave your seat.

5. Time allowed: **3 hours**.

b) Which statistical test would you use to determine where the differences in glutathione levels are across the groups at baseline? **[4 marks]**

c) Suggest a statistical test you would use to test the hypothesis that there is a decrease in levels of glutathione in HIV-infected methamphetamine user across the five year period **[4 marks]**

d) The above study only had only 10 research participants, name and briefly discuss four factors that determine **sample size**. Explain how each factor will affected sample size i.e. whether you need a smaller or larger sample size **[8 marks]**

End of Examination