KENYA METHODIST UNIVERSITY

END OF FIRST TRIMESTER 2008 EXAMINATIONS

FACULTY	**:**	**SCIENCE AND SOCIAL STUDIES**
DEPARTMENT	**:**	**COMPUTER AND INFORMATION SCIENCE**
COURSE CODE	**:**	**MATH 231**
COURSE TITLE	**:**	**BIOSTATISTICS**
TIME	**:**	**3 HOURS**

Instructions:

- *Answer question **ONE** (compulsory) and any other **TWO** questions.*

Question 1

a) The following are the weights in pounds of children at a day-care centre.

Weight (lb)	10-19	20-29	30-39	40-49	50-59	60-69	70-79
Frequency	5	10	13	7	4	4	3

Calculate the following
(i) Mean
(ii) Standard deviation
(iii) Median (7 mks)

b) Assume that among diabetics the fasting blood level of glucose is approximately normally distributed with a mean of 105mg per 100ml and a standard deviation of 9mg per 100ml. Calculate
(i) Proportion of diabetics having levels between 90 and 125mg per 100ml.
(ii) Level cut off the lower 10% of diabetics.
(iii) Levels encompassing 95 % of diabetics.

c) The following data is for the percentage saturation of bile for male patients

40	86	111	86	106	66	123	90	112	52
88	137	88	88	65	79	87	56	110	78
80	47	74	58	88	73	118	67	57	

Construct a 99% confidence interval for the true mean. (5 mks)

d) Each of the hypertensive patients were administered several drugs on different occasions. The results of concern are for a placebo drug compared with hydrochlorothiazide. Each patient first received the placebo and then one month later they each received hydrochlorothiazide. Blood pressure was recorded on these two occasions and recorded.

Patient:	1	2	3	4	5	6	7	8	9	10	11
Placebo:	211	210	210	203	196	190	191	177	173	170	163
Hydrochlorothiazide:	181	172	196	191	167	161	178	160	149	119	156

On the basis of these experimental results, is there any evidence of a difference in mean systolic blood pressure during the use of these two drugs? (5 mks)

e) Clearly explain the four requirements for life-table calculations. (6 mks)

Question 2

a) Some investigators were interested in studying changing patterns in soft tissue sarcomas over time. There are three principle types of these sarcomas, one of which is fibroid, which is characterised by a muscle cell origin. To study this question the investigators utilized data on soft-tissue sarcomas of the arms and legs from Tumor Registry given below.

	Decade		
Type of tissue	1935-44	1945-54	1955-64
Fibroids	40	70	93
Others	33	42	85

Test whether the tissue type is independent of the decade at 95% confidence level. (7 mks)

b) An experiment was conducted at a particular university to study the psychological environment effect on the anatomy of the brain. A group of 19 rats was randomly divided into two groups. Twelve animals in the treatment group lived together in a large cage furnished with play things that were changed daily. While animals in the control group lived isolation with no toys. After a month, the experimental animals were killed and dissected. The following table gives the cortex weights (th thinking part of the brain) in milligrams.

Treatment	707	740	745	652	649	676	699	696	712	708	749	690
control	669	650	651	627	656	642	698					

Test whether the true means for the two groups are significant at 99% confidence level. (8 mks)

Question 3

a) Among susceptible individuals exposed to a particular infectious agent, 36% generally develop clinical disease. A random sample of size 144 people suspected of exposure to the agent, only 35 developed clinical diseases. Does this data support the claim at 95% confidence? (5 mks

b) An important characteristic of glaucoma, an eye disease, is the presence of classic visual field loss. Tonometry is a common form of glaucoma screening, wherein, for example, an eye is classified as positive if it has an intraocular pressure of 21mmhg of higher at a single reading. Given the following data

	Test	
Field loss	Positive	Negative
Yes	13	7
No	413	4567

Calculate this screening test;
(i) Sensitivity.
(ii) Specificity. (4 mks)

Question 4

The following table gives four measures of academic performance for a sample of 12 medical students: GPA in the preclinical years and the National Board Scores.

GPA	x	2.1	3.0	2.2	3.4	3.6	2.5	2.9	2.9	2.0	3.4	2.5	2.5
NBS	y	84	89	78	92	94	85	88	86	82	90	84	83

a) Plot a scatter diagram to represent the data. (3 mks)
b) Calculate the sample correlation coefficient between GPA and National Board Scores. (6 mks)
c) Fit a least squares regression line to the data. (6 mks)

FORMULAE

1. **Descriptive statistics.**

$$\bar{x} = \frac{\sum fx}{N}$$

$$\sigma^2 = \frac{1}{N}\sum fx^2 - \bar{x}^2 \quad or \quad \sigma^2 = \frac{1}{N}\sum f\left(x-\bar{x}\right)^2$$

$$M_d = l + \frac{h}{f}\left(\frac{N}{2} - C\right)$$

2. $A(1-\alpha)100\%$ CI **for mean**

$$\mu = \bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$$

3. **Test of hypothesis paired differences.**
   a) **Paired differences**

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

   b) **Independent samples**

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

$$s^2{}_d = \frac{(n_1-1)s_1{}^2 + (n_2-1)s_2{}^2}{n_1+n_2-2}$$

4. **Chi square test.**

$$x^2 = \sum\frac{(O-E)^2}{E}$$

$$E_{ij} = \frac{R_i C_j}{n}$$

5. **Test about proportion, p**.

$$Z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}, \quad \hat{p} = \frac{x}{n}$$

6. **Regression and correlation.**

a) $$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}} \text{ or } r = \frac{\sum\left(x-\bar{x}\right)\left(y-\bar{y}\right)}{\sqrt{\sum\left(x-\bar{x}\right)^2\left(y-\bar{y}\right)^2}}$$

b) Line y = a + bx

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2} \quad or \quad b = \frac{\sum\left(x-\bar{x}\right)\left(y-\bar{y}\right)}{\sum\left(x-\bar{x}\right)^2}$$

$$a = \frac{\sum y - b\sum x}{n}$$