JOMO KENYATTA UNIVERSITY
OF
AGRICULTURE AND TECHNOLOGY
UNIVERSITY EXAMINATIONS 2014/2015
THIRD YEAR SUPPLEMENTARY/SPECIAL EXAMINATIONS FOR THE DEGREE OF
BACHELOR OF SCIENCE IN ACTUARIAL SCIENCE
STA 2301: TESTS OF HYPOTHOTHESIS

DATE: MARCH 2015 TIME: 2 HOURS

**INSTRUCTIONS: ANSWER QUESTION ONE (COMPULSORY) AND ANY OTHER TWO QUESTIONS.**

## QUESTION ONE (30 MARKS)

(a) Define the following terms: [6 marks]

(i) A statistical hypothesis.
(ii) A test of a statistical hypothesis.
(iii) Type I error.
(iv) Power of a test.
(v) Critical region.
(vi) Size of a test.

(b) Suppose a random variable X follows the normal distribution with unknown mean u and variance $6^2 = 9$. A random sample of size $n = 16$ is taken on X. To test the null hypothesis $H^0$: $u = 2$ against the alternative $H_1$: $u = 4.47$, the critical region is given as $C = \{(x_1, x_2, \ldots, x_n): \bar{x} > 3.23\}$.

Compute:

(i) Pr (Type I error)
(ii) Pr (Type II error)
(iii) The power of the test at $\mu = 4.47$ [7 marks]

(c) State without proof, the Neyman-Pearson Lemma, giving precisely what kind of hypothesis applies. [5 marks]

(d) (i) Define the term probability value (P-value) and explain briefly how it can be used in testing a statistical hypothesis. [2 marks]

(ii) An investor wants to determine whether the mean income of workers working in a town two miles from a proposed estate building site exceeds \$24,400. What conclusion can be made at 5% level of significance if the mean income of a random sample of 60 workers working in that town is \$24,524 and standard deviation of \$763. Use P-value to answer this question. [5 marks]

(e) State and explain briefly the various steps involved in testing a hypothesis. [5 marks]

QUESTION TWO (20 MARKS)

(a) Let $x_1, x_2, \ldots, x_n$ be a random sample of size $n = 100$ from the normal distribution $N(\mu, \sigma^2)$ where $\sigma^2 = 400$. A size $\alpha = 0.05$ test for $H_0$: $\mu = 165$ against $H_1$: $\mu = 168$ is given by a Best Critical Region (BCR) $C = \{(x_1, x_2, \ldots, x_n): \bar{x} > k\}$

(i) Determine the value of K. [5 marks]

(ii) Give the expression for the power function, hence compute the power of the test when $\mu = 172.22$. [6 marks]

(iii) Give a uniformly most powerful (UMP) test for $H_0$: $\mu = 165$ against $H_1$: $\mu > 165$. Explain how this is derived from the Best Critical Region. [5 marks]

(b) A random sample of size n = 1 is obtained on a continuous random variable X. We wish to test the simple null hypothesis

$$H_0: f(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2},\ -\infty<x<\infty$$

against the simple alternative hypothesis

$$H_1: f(x) = \frac{1}{\pi(1+x^2)},\ -\infty<x<\infty$$

Give a size $\alpha$ BCR for the test. [4 marks]

## QUESTION THREE (20 MARKS) regression

Consider a linear statistical model given by $Y = \beta_0 + \beta_1 x + e$, describing the relationship between the variables X and Y, where $\beta_0$ and $\beta_1$ are constants and e is a random variable assumed to have the normal distribution with mean zero and a constant variance $\sigma^2$.

(a) Find the Maximum Likelihood Estimators (MLEs) of $\beta_0$ and $\beta_1$ and $\sigma^2$ under $\Omega$. [5 marks]

(b) Using results of part (a), derive the Likelihood Ratio Test (LRT) of size $\alpha$ for testing $H_1$: $\beta_1 = 0$ against $H_1$: $\beta_1 \neq 0$ [5 marks]

(c) Show that the LRT is part (b) can be based on the sample correlation coefficient r. [10 marks]

## QUESTION FOUR (20 MARKS) Two sample

(a) Let X and Y be independent random variables such that $X\sim N(\mu_1, \theta)$ and $Y\sim N(\mu_2, \theta)$. Samples of sizes m and n are taken on X and Y respectively. You are required to construct a level $\alpha$ Likelihood Ratio Test (LRT) for $H_0$: $\mu_1 = \mu_2$ against $H_1$: $\mu_1 \neq \mu_2$. The following information on the maximum likelihood estimators are given:

Under the entire parameter space $\Omega$

$$\hat{\mu}_1 = \bar{x},\ \hat{\mu}_2 = \bar{y},\ \hat{\theta} = \frac{1}{m+n}\left[\sum_{i=1}^{m}(x_i - \bar{x})^2 + \sum_{i=1}^{n}(y_i - \bar{y})^2\right]$$

Under the restriction Ho: $\mu_1 = \mu_2 = \mu_0$

$$\hat{\mu}_0 = \frac{m\bar{x} + n\bar{y}}{m+n},\quad \hat{\theta}_0 = \frac{1}{m+n}\left[\sum_{i=1}^{m}(x_i - \hat{\mu}_0)^2 + \sum_{i=1}^{n}(y_i - \hat{\mu}_0)^2\right]$$

Derive the required LRT and show that the test is based on the student's t-distribution. Give the degrees of freedom for the distribution. [12 marks]

(b) Thirty 10-year-old children were selected at random for a study to compare methods of teaching children how to spell words. 15 of the children were randomly allocated to each method. After the learning period, each child was scored for spelling ability. The scores of the 30 children were summarized as below:

| $n=15$ Method I | $n=15$ Method II |
|---|---|
| $\bar{X}_1 = 65.3$ | $\bar{X}_2 = 62.4$ |
| $S_1 = 4.2$ | $S_2 = 3.4$ |

The test at $\alpha = 0.05$ level whether the mean scores for the two methods are equal. State assumptions you have made on the distribution of the scores. [8 marks]