KENYATTA UNIVERSITY

UNIVERSITY EXAMINATIONS 2011/2012

FIRST SEMESTER EXAMINATION FOR THE DEGREE OF MASTER OF SCIENCE

SMA 661: NON-PARAMETRIC INFERENCE

DATE: Thursday, 1st December, 2011 TIME: 9.00 a.m. – 12.00 p.m.

INSTRUCTIONS:

Answer question **ONE** and any other **TWO** questions.

Question 1 (30 marks)

a) Explain the meaning of the following a used in Non-Parametric Statistical Inference.

i) Non-parametric hypothesis

ii) Distribution free statistic

iii) Empirical distribution function

iv) The signed – rank test statistic

(8 marks)

b) Show that the empirical distribution $s_n(x)$ is a consistent estimater of the population distribution function $F(x)$ (6 marks)

c) State and prove the Probability Integral Transform (PIT) theorem (7 marks)

e) Let $x_1, x_2, ..., x_n$ be a random sample from a continuous population with a distribution function $F(x)$.

Let M be the median of the distribution. Give the Ordinary – Sign – Test for $H_o : M = M_o$ against $H_1 : M > M_o$, where M_o is a specified number. Show the test based on the Sign – Test is consistent for the alternative H_1 (9 marks)

Question 2 (20 marks)

a) Let $x_1, x_2, ..., x_n$ be a random sample from a continuous distribution $F(x)$ with the median M. Let $D_i = X_i - M_o$, where M_o is a specified number $i = 1, 2, ..., n.$
The Wilcoxon gives the signed – rank test statistic for positive differences as

$$T^+ = \sum_{i=1}^{n} z_i \left\{rank\left(|d_i|\right)\right\},\ where\ z_i = \begin{cases} 1\ if\ D_i > 0 \\ 0\ if\ D_i < 0 \end{cases}$$

Show that when H_o is true then $E\left\{T^+\right\} = \frac{n(n+1)}{4}$ (6 marks)

b) Suppose $(x_i, y_i), i = 1, 2, ..., n$ is a random sample from a continuous bivariate distribution with the distribution function $F(x, y)$. $Let\ D_i = Y_i - x_i, i = 1, 2, ..., n.$
Assume that $D_i, i = 1, 2, ..., n$ is a random sample of differences form
$F(x, y)$, construct a $100(1-\alpha)\%$
confidence interval for the unknown median of differences M_D based on the signed – rank test statistic (8 marks)

c) Let $x_1, x_2, ..., x_n$ be a random sample from a continuous distribution function $F_1(x)$ and $y_1, y_2, ... y_n$ be a random sample form another continuous distribution function $F_2(x)$. Describe a two – sample version of the Kolmogorov – smirnov level α test for $H_0 : F_2(x) = F_1(y)\ for\ all\ x$ against $H_1 : F_2(x) \geq F_1(x)$ with strict in equality for some x. (6 marks)

Question 3 (20 marks)

The one – sided Kolmogorov – Smifnov test statistic is given by
$D_n^- = \underset{x}{Sup}\left\{F(x) - s_n(x)y\right.$ where $F(x)$ is a continuous population cumulative distribution function and s_n is the empirical distribution function.

a) Prove that D_n^- is completely distribution free (9 marks)

b) Prove that for testing:

$H_0: F(x) = F_0(x)$ for all x against

$H_1: F(x) \leq F_0(x)$ with strict inequality for some x, the test given by rejecting H_0 at α level whenever $D_n^- > D_{n,\alpha}^-$ is consistent for the alternative

(11 marks)

Question 4 (20 marks)

Let $x_1, x_2, ..., x_m$ be a random sample from a continuous distribution function $Fx(x)$ *and* $y_1, y_2, ..., y_n$ be a random sample from a continuous distribution function $F_y(x)$.

a) Give the Mann – Whitney U- statistic for comparing the two population distributions $F_x(x)$ *and* $F_y(x)$ and the linear rank test statistic W_N explaining clearly the components in each case. (6 marks)

b) Show that the statistics U *and* W_N are linearly related. (5 marks)

c) To test $H_0: F_x(x) = F_y(x)$ for all x against a suitable alternative, samples of sizes $m = 3$ *and* $n = 4$ are available. By tabulating possible configurations upto $U = 3$, determine the rejection regions for the Mann – Whiteney U – statistic with nominal significance level 0.05 and 0.10 respectively. Give the exact level in each case. (9 marks)

\- E N D -