UNIVERSITY EXAMINATIONS

2022/2023 ACADEMIC YEAR

FIRST YEAR FIRST SEMESTER

MAIN EXAMINATION

FOR THE DEGREE OF MASTER OF SCIENCE IN STATISTICS

COURSE CODE: **STA 820**

COURSE TITLE: LARGE SAMPLE THEORY OF STATISTICAL INFERENCE

DATE: 09/08/2023 **TIME:** 2:00 PM -5:00 PM

INSTRUCTIONS TO CANDIDATES

Answer Any THREE Questions

TIME: 3 Hours

QUESTION ONE (20 MARKS)

1. (a) Name and explain any two probability tools for establishing consistency and asymptotic normality of estimators. (2 mks)

   (b) Show that if $\{a_n\}$ and $\{b_n\}$ are symptotically equivalent then their relative error tends to zero. (2 mks)

   (c) Let a sequence $a_n = n$ and $b_n = 2n$. Show that a sequence $a_n = O(b_n)$ (2 mks)

   (d) If $X_1, X_2, ..., X_n$ is a sequence of random variables and if mean μ_n and standard deviation σ_n of X_n exists for all n and $\sigma_n \to -\infty$ as $n \to \infty$, then $x_n - \mu_n \to 0$ as $n \to \infty$ (2 mks)

   (e) Let $a_n = \frac{1}{n} + \frac{3}{n^2} + \frac{1}{n^3}$ and $b_n = \frac{1}{n}$ are equivalent (3 mks)

   (f) Calculate $P(|Y| < k\sigma)$ exactly for $k = 1, 2, 3, 4$ when Y has an Exp(1) distribution and compare this with the bounds from Chebyshevs inequality. (4 mks)

   (g) The height of the bar at $x = 2$ in the first histogram is 0.26. How many of the 500 realizations were between 1.75 and 2.25? (5 mks)

QUESTION TWO (20 MARKS)

2. (a) Discuss any two importance of the large sample theory (4 mks)

   (b) Explain the following terms: (6 mks)

      i. convergence in probability

      ii. almost sure convergence

      iii. convergence in mean square

   (c) Show that if $F_n \to F$, then the convergence is uniform;

$$\sup_x |F_n| - F(x) \to 0$$

   as $n \to \infty$ (10 mks)

QUESTION THREE (20 MARKS)

3. (a) For Geometric distribution $P(x) = 2^{-x}$; such that $x = 1, 2, \ldots$ prove that Chebyshev's inequality gives $P[|x - 2| \leq 2] > \frac{1}{2}$ while the actual probability is $\frac{15}{16}$ (10 mks)

   (b) Use the Chebyshev inequality to determine how many times a balanced coin must be tossed in order that the probability will be at least 0.80 that the ratio of the observed number of heads to the number of tosses will be between 0.4 and 0.6 (10 mks)

QUESTION FOUR (20 MARKS)

4. (a) Let $x_1, x_2, \ldots, x_n$ be i.i.d observations from the uniform distribution, $f(x;\theta) = \frac{1}{\theta}$, where θ is unknown positive parameter. Show that the MLE of θ is $\hat{\theta} = X_n = Max$ (6 mks)

   (b) An accountant wants to simplify his bookkeeping by rounding amounts to the nearest integer, for example, rounding KSh.99.53 and KSh.100.46 both to KSh.100. What is the cumulative effect of this if there are, say, 100 amounts? To study this we model the rounding errors by 100 independent $U(-0.5, 0.5)$ random variables $X_1, X_2, \ldots, X_{100}$.

      i. Compute the expectation and the variance of the X_i. (7 mks)

      ii. Use Chebyshevs inequality to compute an upper bound for the probability $P(|X_1 + X_2 + \ldots + X_{100}| > 10)$ that the cumulative rounding error $X_1 + X_2 + \ldots + X_{100}$ exceeds KSh.10. (7 mks)

QUESTION FIVE (20 MARKS)

5. (a) State and prove the Polya's Theorem (10 mks)

   (b) Suppose $X_n \to X$, show that for any continuous function $g(.)$, $g(X_n) \to g(X)$. (3 mks)

   (c) A dice is thrown 9000 times and a throw of 3 or 4 is observed 3240 times. Show that the dice cannot be regarded as an unbiased one and find the limits between which the probability of a throw of 3 or 4 lies. (7 mks)

QUESTION SIX (20 MARKS)

6. Of the voters in Nairobi County, a proportion p will vote for candidate G, and a proportion $1p$ will vote for candidate B. In an election poll a number of voters are asked for whom they will vote. Let X_i be the indicator random variable for the event the ith person interviewed

will vote for G. A model for the election poll is that the people to be interviewed are selected in such a way that the indicator random variables $X_1, X_2, \ldots$ are independent and have a Ber(p) distribution.

(a) Suppose we use $\bar{X}_n$ to predict p. According to Chebyshevs inequality, how large should n be (how many people should be interviewed) such that the probability that $\bar{X}_n$ is within 0.2 of the true p is at least 0.9? Hint: solve this first for $p = 1/2$, and use that $p(1p)1/4$ for all $0p1$. (5 mks)

(b) Answer the same question, but now $\bar{X}_n$ should be within 0.1 of p. (5 mks)

(c) Answer the question from part a, but now the probability should be at least 0.95. (5 mks)

(d) If $p > 1/2$ candidate G wins; if $\bar{X}_n > 1/2$ you predict that G will win. Find an n (as small as you can) such that the probability that you predict correctly is at least 0.9, if in fact $p = 0.6$. (5 mks)