Total number of printed pages–12

3 (Sem–3/CBCS) STA HG/RC

2021

(Held in 2022)

STATISTICS

(Honours Generic/Regular)

Paper : STA–HG/RC–3016

(Basics of Statistical Inference)

Full Marks : 60

Time : Three hours

The figures in the margin indicate full marks for the questions.

Answer ***either*** *in English* ***or*** *in Assamese.*

1. Choose the correct option/Fill in the blank : 1×7=7

   শুদ্ধ উত্তৰটো বাছি উলিওৱা/খালী ঠাই পূৰ কৰা ঃ

   (a) When using sign test, if two scores are tied, then

   'চিহ্ন' পৰীক্ষা ব্যৱহাৰ কৰোতে যদি দুটা সংখ্যা একে হয়, তেন্তে

   (i) we count them

   আমি গণনা কৰো

Contd.

(ii) we discard them

আমি বাদ দিওঁ

(iii) it depends upon the scores

ইয়ে সংখ্যা দুটাৰ ওপৰত নিৰ্ভৰ কৰে

(iv) None of the above

ওপৰৰ এটাও নহয়

(b) If a Chi-square goodness of fit test has 6 categories and $N = 30$, then the correct number of degree of freedom is

যদি এটা কাই-বৰ্গৰ সুসমিত পৰীক্ষাত 6টা বিভিন্ন ভাগ থাকে আৰু $N = 30$ হয়, তেন্তে স্বতন্ত্ৰ মাত্ৰা হ'ব

(i) 5

(ii) 6

(iii) 28

(iv) 29

(c) Compared to parametric methods, the non-parametric methods

প্ৰাচলযুক্ত পদ্ধতিৰ তুলনাত অ-প্ৰাচলযুক্ত পদ্ধতিসমূহ

(i) are less accurate

কম সঠিক

(ii) are less efficient

কম দক্ষ

(iii) are computationally easier

গণনাগতভাৱে সহজ

(iv) (ii) and (iii) are correct but not (i)

(ii) আৰু (iii) শুদ্ধ কিন্তু (i) নহয়

(d) If $AB < \frac{(A)(B)}{N}$, then the association between two attributes A and B is

যদি $AB < \frac{(A)(B)}{N}$ হয়, তেন্তে দুটা গুণ A আৰু B-ৰ মাজৰ সম্বন্ধ হ'ব

(i) 0

0

(ii) positive

ধনাত্মক

(iii) symmetric

সমমিত

(iv) negative

ঋণাত্মক

(e) ANOVA is a statistical method of comparing the ________ of several populations. *(Fill in the blank)*

কেবাটাও সমষ্টিৰ ________ তুলনা কৰিবলৈ ব্যৱহাৰ কৰা পাৰিসাংখ্যিক পদ্ধতিটো হৈছে ANOVA।

(খালী ঠাই পূৰ কৰা)

(f) Consider k independent samples each containing $n_1, n_2, \dots, n_k$ items such that $n_1 + n_2 + \dots + n_k = n$. In ANOVA, we use F-distribution with degree of freedom

ধৰা হওঁক kটা স্বতন্ত্ৰ প্ৰতিদৰ্শৰ প্ৰতিটোৱে $n_1, n_2, \dots, n_k$ টা গোট থাকে, য'ত $n_1 + n_2 + \dots + n_k = n$ হয়। এইবোৰ ANOVA-ৰ কাৰণে F-বণ্টনৰ স্বতন্ত্ৰ মাত্ৰা হ'ব

(i) $k-1,\ n-k$

(ii) $k-1,\ n-1$

(iii) $k-n,\ n-k$

(iv) $n-k,\ k-1$

(g) The probability of type II error when $\alpha = 0{\cdot}05$ is

$\alpha = 0{\cdot}05$-ৰ কাৰণে দ্বিতীয় ধৰণৰ ত্ৰুটিৰ সম্ভাৱিতাৰ মান হ'ব

(i) 0·025

(ii) 0·05

(iii) 0·95

(iv) cannot be determined without more information

অধিক তথ্য নহলে মান উলিওৱা সম্ভৱ নহয়

2. Answer the following questions briefly : 2×4=8

তলৰ প্ৰশ্নবসমূহৰ উত্তৰ চমুকৈ লিখা ঃ

(a) Define type I and type II errors with examples.

প্ৰথম ধৰণৰ ত্ৰুটি আৰু দ্বিতীয় ধৰণৰ ত্ৰুটিৰ উদাহৰণৰ সৈতে সংজ্ঞা দিয়া।

(b) Mention *three* basic principles of design of experiments.

পৰীক্ষণ অভিলেখৰ *তিনিটা* মূল নীতি উল্লেখ কৰা।

(c) Explain sign test for symmetry.

সমমিত ক্ষেত্ৰত চিহ্ন পৰীক্ষাটো ব্যাখ্যা কৰা।

(d) Define association of attributes.

গুণৰ সম্বন্ধৰ সংজ্ঞা দিয়া।

3. Answer ***any three*** questions from the following : 5×3=15

তলত দিয়াবোৰৰ *যিকোনো তিনিটা* প্ৰশ্নৰ উত্তৰ কৰা ঃ

(a) Write down all the conditions for the validity of χ^2-test of goodness of fit.

আসঞ্জনৰ শ্ৰেষ্ঠতা পৰীক্ষাৰ ক্ষেত্ৰত ব্যৱহৃত হোৱা χ^2-বৰ্গৰ প্ৰয়োগৰ চৰ্তসমূহ উল্লেখ কৰা।

(b) In a large city A, 20% of a random sample of 900 school children had defective eye-sight. In other large city B, 15% of random sample of 1600 children had the same defect. Write down the required hypothesis and hence verify whether the difference between two proportions are significant.

এখন ডাঙৰ চহৰ A-ত, 900 ছাত্ৰ-ছাত্ৰীৰ ভিতৰত 20% যাদৃচ্ছিক প্ৰতিদৰ্শৰ ত্ৰুটিযুক্ত দৃষ্টি শক্তিৰ প্ৰমাণ লোৱা হ'ল। আন এখন চহৰ B-ত, 1600 ছাত্ৰ-ছাত্ৰীৰ ভিতৰত 15% যাদৃচ্ছিক প্ৰতিদৰ্শৰ একে ধৰণৰ ত্ৰুটি দেখা পোৱা গ'ল। উপযুক্ত প্ৰকল্পটো লিখা আৰু দুয়োটা অনুপাতৰ ব্যৱধানৰ সাৰ্থকতা পৰীক্ষা কৰা।

(c) What do you mean by Yates' correction? Can we apply the same to every 2×2 table even if no theoretical cell frequency is less than 5?

4+1=5

'ইয়েটৰ শুধৰণি' বুলিলে কি বুজা? 2×2 সৰণিত ওই শুধৰণি আমি প্ৰয়োগ কৰিব পাৰিমনে য'ত এটাও প্ৰত্যাশিত বাৰংবাৰতা 5-ত কৈ কম নহয়?

(d) Given the frequency distribution —

$$f(x, \theta) = \begin{cases} 1/\theta, & 0 \le x \le \theta \\ 0, & \text{otherwise} \end{cases}$$

and that you are testing the null hypothesis $H_0 : \theta = 1$ against $H_1 : \theta = 2$ by means of a single observed value of x. What would be the sizes of type I and type II errors if you choose the interval as follows?

(i) $0 \cdot 5 \le x$

(ii) $1 \le x \le 1 \cdot 5$
as the critical region

বাৰংবাৰতা বণ্টনটো তলত দিয়া ধৰণে দিয়া আছে —

$$f(x, \theta) = \begin{cases} 1/\theta, & 0 \le x \le \theta \\ 0, & \text{otherwise} \end{cases}$$

এটা ৰিক্ত প্ৰকল্প $H_0 : \theta = 1$-ৰ বিপক্ষে $H_1 : \theta = 2$ পৰীক্ষা কৰিব বিচৰা হৈছে য'ত x-ৰ কেৱল এটা প্ৰেক্ষিত মান দিয়া আছে। প্ৰথম ধৰণৰ ত্ৰুটি আৰু দ্বিতীয় ধৰণৰ ত্ৰুটিৰ আকাৰ কিমান হ'ব যদিহে ব্যৱধানটো তলত দিয়া ধৰণে হয়?

(i) $0 \cdot 5 \le x$

(ii) $1 \le x \le 1 \cdot 5$
ক্ৰান্তিক ক্ষেত্ৰ ৰূপে

(e) Fill in the blanks of the following table :

তলত দিয়া সৰণিখনত খালী ঠাইসমূহ পূৰণ কৰা ঃ

Source উৎস	d.f. স্বতন্ত্ৰ মাত্ৰা	Sum of squares বৰ্গৰ যোগফল	Mean squares গড় বৰ্গ	F F
X	6	–	–	–
Y	10	–	1095	–
Error (ত্ৰুটি)	–	10160	–	
Total (মুঠ)	60	18225		

4. Answer ***any three*** questions from the following : 10×3=30

তলত দিয়াবোৰৰ *যিকোনো তিনিটা* প্ৰশ্নৰ উত্তৰ লিখা ঃ

(a) Discuss the analysis of randomized block design (RBD). 10

যাদৃচ্ছিক খণ্ড পৰীক্ষা অভিলেখৰ বিশ্লেষণ আলোচনা কৰা।

(b) State and prove Neyman-Pearson fundamental lemma. 10

নিমেন-পিয়েৰছন-ৰ মৌলিক সূত্ৰটো লিখা আৰু প্ৰমাণ কৰা।

(c) (i) Two sample polls of votes for two candidates A and B for a public office are taken, one from among the residents of rural areas. The results are given in the following table :

এটা চৰকাৰী কাৰ্যালয়ৰ বাবে দুজন প্ৰাৰ্থী A অথবা B নিৰ্বাচিত কৰিবলৈ এটা প্ৰতিদৰ্শ ভোট গ্ৰহণ কৰা হয়। ফলাফলসমূহ তলৰ তালিকাত দিয়া হৈছে ঃ

	Vote for (ভোট)		Total (মুঠ)
Area (এলেকা)	A	B	
Rural (গ্ৰাম্য)	620	380	1000
Urban (চহৰ)	550	450	1000
Total (মুঠ)	1170	830	2000

Examine whether the nature of the area is related to voting preference in the election. 6

(Table value of χ^2 for 1.d.f. is 3·841)

পৰীক্ষা কৰা যে ভোটদানৰ ধৰণটো এলেকাৰ প্ৰকাৰৰ ওপৰত নিৰ্ভৰ কৰে নে নকৰে।

(χ^2-বৰ্গৰ তালিকাভুক্ত মান 1 স্বতন্ত্ৰ মাত্ৰাৰ কাৰণে দিয়া আছে 3·841)

(ii) In a series of houses actually invaded by smallpox, 70% of the inhabitants are attacked and 85% have been vaccinated. What is the lowest percentage of the vaccinated that must have been attacked? 4

সাৰু আইৰ দ্বাৰা আক্ৰান্ত কিছু আৱাসৰ 70% আৱাসী আক্ৰান্ত হৈছে আৰু 85% আৱাসী প্ৰতিষেধক গ্ৰহণ কৰিছে। প্ৰতিষেধক গ্ৰহণ কৰা কিমান নূন্যতম শতাংশ আৱাসী আক্ৰান্ত হৈছে উলিওৱা।

(d) (i) Write down the hypothesis that is tested in completely randomised design (CRD). Hence write down the AOV table. 2+3=5

সম্পূৰ্ণ যাদৃচ্ছিক পৰীক্ষণ অভিলেখৰ পৰীক্ষা কৰিবলগীয়া প্ৰকল্পসমূহ লিখা। ইয়াৰ সহায়ত প্ৰসৰণ বিশ্লেষণ তালিকাখন প্ৰস্তুত কৰা।

(ii) Estimate the missing value in case of CRD. 5

সম্পূৰ্ণ যাদৃচ্ছিক পৰীক্ষণ অভিলেখৰ বাবে নিখোঁজ গোটি পদ্ধতিটো আলোচনা কৰা।

(e) An insurance agent has claimed that the average age of policy holder who insure through him is less than the average for all agents which is 30·5 years. A random sample of 100 policy holders who had insured through him gave the following age distribution : 10

এজন বীমা এজেণ্টে দাবী কৰে যে তেওঁৰ দ্বাৰা বীমা কৰা গ্ৰাহকসকলৰ গড় বয়স আন গ্ৰাহক সকলৰ গড় বয়স (30·5 বছৰ)-তকৈ কম। এটা 100 জন গ্ৰাহকৰ প্ৰতিদৰ্শ যিসকলে উল্লেখিত এজেণ্ট জনৰ জৰিয়তে বীমা কৰাইছিল, বয়সৰ বণ্টন তলত দিয়া হ'ল ঃ

Age at last birthday : (যোৱা জন্মদিনত বয়স)	16–20	21–25	26–30	31–35	36–40
No. of persons : (মানুহৰ সংখ্যা)	12	22	20	30	16

Calcualte the arithmatic mean and standard deviation of this distribution and use these values to test his claim at the 5% level of significance.
[You are given that $Z(1\cdot65) = 0\cdot95$]

বণ্টনটোৰ সমান্তৰ মাধ্য আৰু মানক বিচলন উলিওৱা আৰু এই তথ্য ব্যৱহাৰ কৰি বীমা এজেণ্ট জনৰ দাবী 5% সাৰ্থকতা স্তৰত পৰীক্ষা কৰা।
[দিয়া আছে $Z(1{\cdot}65) = 0{\cdot}95$]