ভূমিকা, নমুনা চয়নের প্রারম্ভিক অংশ ও প্রাক্কলন

(Introduction, Preliminaries in sample survey and estimates)

প্রশ্ন: উদাহরণসহ সমগ্রকের (Population) সংজ্ঞা দাও (জা.বি.-১১)/তথ্যবিশ্ব কি?

উত্তর: কোন একটি পরীক্ষণে বা পর্যবেক্ষণে সুনির্দিষ্ট কিছু বৈশিষ্ট্যের অধিকারী সম্ভাব্য সকল উপাদানের সমষ্টিকে তথ্যবিশ্ব বা সমগ্রক বলে। সমগ্রকের মোট একক সংখ্যা হলো সমগ্রকের আকার (Size) এবং একে সাধারণত N দিয়ে প্রকাশ করা হয়।

যেমন: কোন সিটি কর্পোরেশনের মেয়র নির্বাচনে উক্ত সিটি কর্পোরেশনের সকল রেজিস্টার্ড ভোটার নিয়ে হবে একটি সমগ্রক। এখানে প্রত্যেক রেজিস্টার্ড ভোটার হলো একক।

অথবা, আমরা যদি কোন এলাকায় বসবাসকারী লোকদের জীবন যাত্রার মান সম্পর্কে গবেষণা করতে চাই, তবে ঐ এলাকার সকল লোক নিয়ে হবে একটি সমগ্রক।

আকারের ভিত্তিতে সমগ্রককে দু'ভাগে ভাগ করা যায়। যথা-(ক) সসীম সমগ্রক, (খ) অসীম সমগ্রক।

(ক) সসীম সমগ্রক (Finite Population): সমগ্রকে মোট একক সংখ্যা নির্দিষ্ট অর্থাৎ গণনার সীমার মধ্যে থাকলে তাকে সসীম সমগ্রক বলে।

যেমন: কোন কলেজের মোট শিক্ষার্থীর সমগ্রক, কোন জেলায় মোট গ্রামের সমগ্রক।

(খ) অসীম সমগ্রক (Infinite Population): সমগ্রকে মোট একক সংখ্যা অনির্দিষ্ট বা অসীম অর্থাৎ গণনা সীমার মধ্যে না থাকলে সে সমগ্রককে অসীম সমগ্রক বলে।

যেমন: কোন পুকুরে মাছের সমগ্রক, মানুষের শরীরের রক্তে লোহিত কনিকা সমগ্রক।

প্রশ্ন: নমুনা কাকে বলে ? (জা.বি.-১১)/ নমুনা (Sample) কি ?

অথবা, উদাহরণসহ নমুনার সংজ্ঞা দাও (জা.বি.-১৫)

উত্তর: তথ্যবিশ্বের প্রতিনিধিত্বশীল ক্ষুদ্রতম অংশকে নমুনা বলে।

যেমন: কোন কলেজের অনার্স ১ম বর্ষের শিক্ষার্থীর সংখ্যা ১০০০। শিক্ষার্থীদের গড় উচ্চতা অনুসন্ধানে ১০০০ জন থেকে ১০০ জন শিক্ষার্থী দৈবভাবে নির্বাচন করা হলে উক্ত ১০০ জন শিক্ষার্থী নিয়ে গঠিত হয় নমুনা।

প্রশ্ন: নমুনায়ন ইউনিট কি ? (জা.বি.-১১)/ নমুনা একক (Sampling Unit) কি ?

উত্তর: তথ্যবিশ্বের প্রতিটি ক্ষুদ্রতম একক যাদের ব্যবহার করে নমুনা চয়ন করা হয়, তাকে নমুনা একক বলে। নমুনা এককগুলো ভিন্ন ও পরস্পর বর্জনশীল।

যেমন: কোন কলেজের শিক্ষার্থীদের মধ্য হতে কোন বৈশিষ্ট্যের ভিত্তিতে কোন নমুনা চয়ন করা হলে উক্ত কলেজের প্রতিজন শিক্ষার্থী হবে নমুনা একক।

প্রশ্ন: পরামান (Parameter)/ পরামিতি কাকে বলে ? (জা.বি.-১২)

উত্তর: তথ্যবিশ্বের সকল একক হতে তথ্য নিয়ে কোন বৈশিষ্ট্য পরিমাপ করা হলে যে মান পাওয়া যায়, তাকে পরামিতি বা পরামান বলে। অর্থাৎ পরামান হলো তথ্যবিশ্বের বৈশিষ্ট্য এবং অনেকক্ষেত্রেই পরামানগুলো অজানা থাকে।

যেমন: সমগ্রকের গড়, মধ্যমা, ভেদাংক, পরিমিত ব্যবধান, বিভেদাংক, বঙ্কিমতাংক, সূঁচালতাংক ইত্যাদি হলো পরামান।

বি. দ্র: কোন তথ্যবিশ্বকে গাণিতিকরূপে প্রকাশ করতে যে অজানা মানগুলো জানা প্রয়োজন, তাকেও পরামিতি বলে।

যেমন: দ্বিপদী বিন্যাসের পরামিতি n ও p।

প্রশ্ন: নমুনাজ মানের সংজ্ঞা দাও (জা.বি.-১১)/ নমুনাজমান (Statistic) বলতে কি বুঝ?

উত্তর: যে মানগুলো নমুনার বৈশিষ্ট্যকে নির্দেশ করে, তাদেরকে নমুনাজমান বলে। নমুনাজমান নমুনা ভেদে ভিন্ন ভিন্ন হয় বলে নমুনাজমান ধ্রুবক নয়। নমুনাজমান সব সময় পরামানের নিরূপক নাও হতে পারে। নমুনাজমান পরামানের মান প্রাক্কলনে কিংবা যথার্থতা যাচাইয়ে ব্যবহৃত হয়।

যেমন: নমুনা গড়, নমুনা ভেদাংক ইত্যাদি হলো নমুনাজমান।

প্রশ্ন: পরামিতি ও নমুনাজ মানের মধ্যে পার্থক্য লিখ।

উত্তর: পরামিতি ও নমুনাজ মানের মধ্যে পার্থক্যসমূহ নিম্নরূপ-

পার্থক্যের বিষয়	পরামিতি	নমুনাজমান
(i) সংজ্ঞা	তথ্যবিশ্বের সকল এককের ভিত্তিতে কোন বৈশিষ্ট্যের পরিমাপ করা হলে যে মান পাওয়া যায়, তাকে পরামিতি বলে।	তথ্যবিশ্ব হতে প্রতিনিধিত্বশীল নমুনা নির্বাচন করে নমুনা মানগুলোর ভিত্তিতে কোন বৈশিষ্ট্য পরিমাপ করা হলে তাকে নমুনাজ মান বলে।
(ii) বৈশিষ্ট্য	ইহা তথ্যবিশ্বের বৈশিষ্ট্য।	ইহা নমুনার বৈশিষ্ট্য।
(iii) প্রাক্কলন	ইহা সাধারণত অজানা থাকে এবং নমুনাজ মান দ্বারা প্রাক্কলন করা হয়।	ইহা পরামিতি প্রাক্কলনে বা নিরুপনে ব্যবহৃত হয়।
(iv) উদাহরণ	উদাহরণ-তথ্যবিশ্বের গড় μ , তথ্যবিশ্বের ভেদাংক σ^2 ইত্যাদি।	উদাহরণ-নমুনা গড় $\bar{x}$, নমুনার ভেদাংক s^2 ইত্যাদি।

প্রশ্ন: সমগ্রক ও নমুনার মধ্যে পার্থক্যসমূহ লিখ।

উত্তর: সমগ্রক ও নমুনার মধ্যে পার্থক্যসমূহ নিম্নরূপ-

পার্থক্যের বিষয়	তথ্যবিশ্ব	নমুনা
(i) সংজ্ঞা	কোন একটি পরীক্ষণে বা পর্যবেক্ষণে সুনির্দিষ্ট কিছু বৈশিষ্ট্যের অধিকারী সম্ভাব্য সকল উপাদানের সমষ্টিকে তথ্যবিশ্ব বা সমগ্রক বলে।	তথ্যবিশ্বের প্রতিনিধিত্বশীল ক্ষুদ্রতম অংশকে নমুনা বলে।

(ii) একক সংখ্যা	তথ্যবিশ্বের একক বা উপাদান সংখ্যা সসীম বা অসীম হতে পারে।	নমুনার একক সংখ্যা সাধারণত সসীম।
(iii) পরামান ও নমুনাজমান	তথ্যবিশ্বের বিভিন্ন বৈশিষ্ট্যের পরিমাপকে পরামান বলা হয়।	নমুনার বিভিন্ন বৈশিষ্ট্যের পরিমাপকে নমুনাজমান বলা হয়।
(iv) বিবিধ	তথ্যবিশ্বের পরমানগুলো সাধারণত অজানা থাকে।	নমুনার সাহায্যে সমগ্রকের পরামানগুলো প্রাক্কলন করা হয়।

প্রশ্ন: প্রাক্কলক (Estimator)/ নিরূপক বলতে কি বুঝ ? (জা.বি.-১২)

উত্তর: পরামানের প্রকৃত মান সাধারণত: অজানা থাকে। পরামানের মান নির্ণয়ে নমুনাজমান ব্যবহৃত হয়। যখন নমুনাজমান পরামানের মান নির্ণয়ে ব্যবহৃত হয়, তখন সে নমুনাজমানকে প্রাক্কলক বলা হয়। প্রাক্কলক একটি দৈব চলক যার মান এক নমুনা হতে অন্য নমুনায় পরিবর্তনশীল।

যেমন: নমুনা গড় হলো সমগ্রক গড়ের প্রাক্কলক।

প্রশ্ন: প্রাক্কলিত মান (Estimate) কি?

উত্তর: প্রাক্কলকের একটি বিশেষ মানকে বলা হয় প্রাক্কলিত মান।

যেমন: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ এর মাধ্যমে $\bar{x} = 50$ পাওয়া গেল। এখানে 50 হলো প্রাক্কলিত মান।

প্রশ্ন: প্রাক্কলকের নিরপেক্ষতা (unbiasedness of an estimator) বলতে কি বুঝ ? (জা.বি.-১১)

অথবা, নির্ঝুঁকি প্রাক্কলক (Unbiased Estimator) বলতে কি বুঝ?

উত্তর: কোন প্রাক্কলকের প্রত্যাশিত মান পরামিতির মানের সমান হলে তাকে নির্ঝুঁকি প্রাক্কলক বলা হয়।

গাণিতিকভাবে, একটি প্রাক্কলক $\hat{\theta}$ কে পরামান θ এর নিরপেক্ষ নিরূপক বলা হয়, যদি $\hat{\theta}$ এর প্রত্যাশিত মান θ এর মান সমান হয় অর্থাৎ $E(\hat{\theta}) = \theta$.

যেমন: নমুনা গড় $\bar{x}$, তথ্যবিশ্ব গড় μ এর নির্ঝুঁকি প্রাক্কলক কেননা $E(\bar{x}) = \mu$

প্রশ্ন: প্রাক্কলকের ঝোঁক এর সংজ্ঞা দাও (জা.বি.-১১)

অথবা, ঝোঁক (Bias) বলতে কি বুঝ ?

উত্তর: প্রাক্কলকের প্রত্যাশিত মান ও পরামিতির মানের ব্যবধানকে বলা হয় ঝোঁক।

গাণিতিকভাবে, একটি প্রাক্কলক $\hat{\theta}$ কে পরামান θ এর ঝুঁকি নিরূপক বলা হয়, যদি $\hat{\theta}$ এর প্রত্যাশিত মান θ এর মান সমান না হয় অর্থাৎ $E(\hat{\theta}) \neq \theta$। প্রাক্কলক $\hat{\theta}$ এর ঝোঁক, $E(\hat{\theta}) - \theta$. নির্ঝুঁকি প্রাক্কলকের ঝোঁক শূন্য।

প্রশ্ন: আদর্শ বিচ্যুতি (Standard Error) কি?

উত্তর: কোন প্রাক্কলকের (estimator) ভেদাংকের ধনাত্মক বর্গমূলকে আদর্শ বিচ্যুতি বলে। নমুনা গড় $(\bar{x})$ হলো সমগ্রক গড়ের প্রাক্কলক। যদি কোন তথ্যবিশ্বের পরিমিত ব্যবধান জানা থাকে বা অনুমান (estimate) বা নির্ণয় করা যায়, তবে নমুনা গড় $(\bar{x})$ এর আদর্শ বিচ্যুতি হলো,

$$SE(\bar{x})=\frac{\sigma}{\sqrt{n}}=\frac{\sqrt{\frac{\sum_{i=1}^{N}(X_i-\mu)^2}{N}}}{\sqrt{n}}$$

; যেখানে, $\sigma=$ তথ্যবিশ্বের পরিমিত ব্যবধান, $\mu=$ তথ্যবিশ্বের গড়,
$N=$ তথ্যবিশ্বের আকার, $n=$ নমুনার আকার

যদি তথ্যবিশ্বের পরিমিত ব্যবধান অজানা থাকে, তবে নমুনা গড় $(\bar{x})$ এর আদর্শ বিচ্যুতি হলো,

$$SE(\bar{x})=\frac{s}{\sqrt{n}}=\frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}}{\sqrt{n}}$$

; যেখানে, $s=$ নমুনার পরিমিত ব্যবধান, $\bar{x}=$ নমুনার গড়,
$n=$ নমুনার আকার

NB (Nota Bene):

(i) $$V(\bar{x})=V\left(\frac{\sum x_i}{n}\right)=\frac{1}{n^2}\sum V(x_i)=\frac{1}{n^2}\sum\sigma^2=\frac{n\sigma^2}{n^2}=\frac{\sigma^2}{n}$$

$$\therefore\ SE(\bar{x})=\sqrt{\frac{\sigma^2}{n}}=\frac{\sigma}{\sqrt{n}}$$

(ii) **Standard error** measures the variability of sample means across multiple samples of the same population whereas **Standard deviation** measures the spread or variability of individual data points within a single sample. In statistics, a sample mean deviates from the actual mean of a population; this deviation is the standard error of the mean.

For example: Imagine you are trying to estimate the average age of residents in Chawkbazar, Chattogram. You take a sample of 100 residents, and you find their average age is 35 years. The standard deviation of the ages in your sample is 10 years.

To calculate the standard error of the mean (SEM), you would use the following formula:

$$SE(\bar{x})=\frac{s}{\sqrt{n}}$$ = standard deviation / square root of sample size

$$= 10 / \sqrt{100} = 10 / 10 = 1$$

This means that the standard error of the mean (SEM) is 1 year. This tells that if you were to take multiple samples of 100 residents, the average age of those samples would vary by about 1 year, on average, from your initial sample mean of 35 years.

(iii) Sample variance, $s^2 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$. Here $(n-1)$ is used instead of n to correct for bias when estimating the population variance from a sample. This is known as **Bessel's correction**.

Sample variance is used to estimate the population variance. The deviations $(x_i - \bar{x})$ are smaller than the deviations $(x_i - \mu)$ which means $\sum_{i=1}^{n}(x_i - \bar{x})^2$ is also smaller than $\sum_{i=1}^{N}(x_i - \mu)^2$. Thus, the result of $\dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$ tends to be too small compared to $\sigma^2 = \dfrac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$. So, $\dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$ underestimates (ছোট করে কমিয়ে দেখা, অবমূল্যায়ন করা) the actual population variance.

To fix (ঠিক করা) this underestimation, we divide $\sum_{i=1}^{n}(x_i - \bar{x})^2$ by $(n-1)$ instead of n. This slightly increases the variance and makes it an unbiased estimator of the population variance i.e., $E(s^2) = \sigma^2$.

প্রশ্ন: প্রাক্কলকের সঙ্কিতা (accuracy) এর সংজ্ঞা দাও (জা.বি.-১১)

অথবা, প্রাক্কলকের সঠিকতা (Accuracy of estimator) বলতে কি বুঝ?

উত্তর: প্রাক্কলকের সঠিকতা হলো পরামানের সাথে প্রাক্কলকের নৈকট্যতা (closeness). গাণিতিকভাবে, একটি পরামান θ এর প্রাক্কলক $\hat{\theta}$ কে সঠিক বলা হয়, যদি $\hat{\theta} - \theta$ খুব ক্ষুদ্র বা অগ্রাহ্য (negligible) হয়।

প্রশ্ন: প্রাক্কলকের যথার্থতার সংজ্ঞা দাও (জা.বি.-১১)

অথবা, যথার্থতা (Precision) কি?

উত্তর: যথার্থতা হলো প্রাক্কলকের সাথে তার প্রত্যাশিত মানের নৈকট্যতা। ধরা যাক, একটি প্রাক্কলক $\hat{\theta}$ এর প্রত্যাশিত মান $E(\hat{\theta})$। প্রাক্কলক $\hat{\theta}$ কে যথার্থ বলা হয় যদি $\hat{\theta} - E(\hat{\theta})$ খুব ক্ষুদ্র বা অগ্রাহ্য হয়।

এখানে উল্লেখ্য যে, যখন প্রাক্কলক $\hat{\theta}$, পরামান θ এর নিরপেক্ষ নিরূপক হয়, তখন সঠিকতা ও যথার্থতা সমান হয়ে থাকে।

প্রশ্ন: নমুনায়ন কাঠামো (Sampling frame) কি? (জা.বি.-১৫)

উত্তর: যে তথ্যবিশ্ব হতে নমুনা নেয়া হবে সে তথ্যবিশ্বের এককগুলোর তালিকাকে নমুনায়ন কাঠামো বলে। এককগুলোকে এমনভাবে সন্নিবিষ্ট ও সাজানো হয় যেন তালিকায় প্রত্যেকটি একক কেবল একবার অন্তর্ভুক্ত হয় এবং কোন একক যেন তালিকা হতে বাদ না যায়।

যেমন: ধরি, কোন কলেজে সম্মান ২য় বর্ষের শিক্ষার্থীদের সংখ্যা ২,০০০। তাদের গড় ওজন অনুসন্ধানে শিক্ষার্থীদের ১ থেকে ২০০০ পর্যন্ত ক্রমানুসারে তালিকাভুক্ত করা হয়। এই তালিকাই নমুনা কাঠামো।

প্রশ্ন: নমুনাজ বা নমুনায়ন ত্রুটি (Sampling Error) **বলতে কি বুঝ ? (জা.বি.-১১) ইহাকে কিভাবে কমানো হয়?**

উত্তর: নমুনায়নে সমগ্রকের সকল একক হতে তথ্য সংগ্রহ না করে সমগ্রকের একটি নমুনা হতে তথ্য সংগ্রহ করা হয়। তাই পরামান ও এর প্রাক্কলকের মধ্যে পার্থক্য দেখা যায়। এ পার্থক্য বা ত্রুটিকে নমুনায়ন ত্রুটি বলা হয়।

গাণিতিকভাবে, ধরি $\hat{\theta}$, পরামান θ এর একটি প্রাক্কলক। তবে নমুনায়ন ত্রুটির পরিমাণ $|\hat{\theta}-\theta|$ । শুমারীর ক্ষেত্রে কোন নমুনাজ ত্রুটি ঘটে না।

যেমন: কোন কলেজের ছাত্রদের গড় উচ্চতা পরিমাপ করতে ১০ জন ছাত্রের একটি নমুনা নেয়া হল। নমুনাজ গড় $\bar{x}$ নির্ণয় করা হলো। সুতরাং ঐ কলেজের ছাত্রদের গড় উচ্চতা $\bar{x}$ এর নমুনাজ বিচ্যুতি $= |\bar{x}-\mu|$

নমুনাজ ত্রুটি কমানোর উপায়:

(ক) সাধারণত নমুনার আকার বৃদ্ধি করে নমুনাজ ত্রুটি কমানো যায়।

(খ) সঠিক নমুনায়ন পদ্ধতি ব্যবহার করে নমুনায়ন ত্রুটি কমানো যায়।

(গ) সঠিক প্রাক্কলক ব্যবহার করেও নমুনাজ ত্রুটি কমানো যায়। যেমন: সরল দৈব নমুনায়নের ক্ষেত্রে ভেদাংকের নিরূপক হিসেবে $\frac{\sum(x_i-\bar{x})^2}{n}$ এর পরিবর্তে $\frac{\sum(x_i-\bar{x})^2}{n-1}$ ব্যবহার করলে কোন ধরণের ত্রুটি ঘটবে না।

প্রশ্ন: নমুনায়ন ত্রুটির কারণসমূহ আলোচনা কর।

অথবা, নমুনাজ ত্রুটির কারণসমূহ (sources of sampling error) বর্ণনা কর (জা.বি.-১৬)

উত্তর: নমুনায়ন ত্রুটির কারনসমূহ নিম্নরূপ-

(ক) নমুনায়ন ডিজাইনের ত্রুটিপূর্ণ নির্বাচন: সঠিক নমুনায়ন ডিজাইন না নিয়ে নমুনা নির্বাচন করলে এ ত্রুটি দেখা যায়। যেমন: যে ক্ষেত্রে সরল দৈব নমুনায়ন ডিজাইন সঠিক হতো, সে ক্ষেত্রে ধারাবাহিক নমুনায়ন ডিজাইন নির্বাচন করলে এ ত্রুটি দেখা দেয়।

(খ) নমুনার আকার নির্ধারণে ত্রুটি: পূর্ব নির্ধারিত বিচ্যুতির ভিত্তিতে সর্বনিম্ন কত আকারের নমুনা নির্বাচন করতে হবে তা পূর্বেই নির্ধারণ করতে হয়। নমুনা আকার অপেক্ষাকৃত খুব ছোট হলে এ ত্রুটি দেখা যায়।

(গ) নমুনা একক নির্বাচনে ত্রুটি: নির্বাচিত একক থেকে তথ্য সংগ্রহ করতে অসুবিধা হলে ঐ এককের পরিবর্তে অন্য কোন একক থেকে তথ্য সংগ্রহ করা হলে এ ধরণের ত্রুটি সৃষ্টি হয়।

(ঘ) ত্রুটিপূর্ণ সীমাযুক্ত নমুনা একক নির্ধারণ: নমুনা একক হলো একটি উপাদান বা অনেকগুলো উপাদানের সমাহার। যেমন- গ্রামের সকল লোক একটি নমুনা একক হলে যদি কোন পরিবারকে বাদ দিয়ে তা তৈরি করা হয়, সেক্ষেত্রে এ ত্রুটি দেখা দেয়।

(ঙ) ত্রুটিপূর্ণ প্রাক্কলক ব্যবহার: পরামানের ঝুঁকি প্রাক্কলক ব্যবহার করে পরামান প্রাক্কলন করা হলে নমুনায়ন ত্রুটি হয়ে থাকে।

প্রশ্ন: অনমুনায়ন বা অনমুনাজ ত্রুটি (Non-sampling Error) বলতে কি বুঝ ? ইহা কিভাবে কমানো যায় ?/ কিভাবে অনমুনাজ ত্রুটি নিয়ন্ত্রণ করা যায় ? (জা.বি.-১২)

উত্তর: নমুনাজ ত্রুটি ব্যতীত প্রাক্কলনের জন্য অন্য সব ত্রুটির কারণে পরামিতি ও নমুনাজ মানের মধ্যে যে পার্থক্য পরিলক্ষিত হয়, এ পার্থক্যকে অনমুনায়ন ত্রুটি বলে।

ত্রুটিপূর্ণ নমুনায়ন পদ্ধতি, ভুল তথ্য পরিবেশন, নির্বাচিত নমুনা থেকে তথ্য সংগ্রহ না করা ও প্রক্রিয়াকরণে ভুল করা ইত্যাদি কারণে অনমুনাজ ত্রুটি ঘটতে পারে। অনমুনাজ ত্রুটি শুমারী ও নমুনা জরিপ উভয় ক্ষেত্রেই ঘটতে পারে।

অনমুনাজ ত্রুটি কমানো বা নিয়ন্ত্রণের উপায়:

(ক) সঠিক নমুনায়ন কাঠামোর ব্যবহার এবং সঠিকভাবে নমুনা নির্বাচন দ্বারা এ ত্রুটি কমানো যেতে পারে।

(খ) প্রশিক্ষণপ্রাপ্ত ও অভিজ্ঞ তথ্য সংগ্রহকারী নিয়োগ এবং তথ্য সংগ্রহকারীদের বিচক্ষণতা ও নিরপেক্ষতার দিকে খেয়াল রাখলে এ ত্রুটি কমানো যায়।

(গ) নমুনার বৈশিষ্ট্য পরিমাপের সময় সঠিক সূত্রের প্রয়োগ করতে হবে।

(ঘ) মূল জরিপের পূর্বে পাইলট জরিপ পরিচালনা করে মূল জরিপের ত্রুটি সনাক্ত করতে হবে।

প্রশ্ন: অনমুনায়ন ত্রুটির কারণসমূহ আলোচনা কর।

অথবা, অনমুনায়ন বিচ্যুতির কারণসমূহ বর্ণনা দাও (জা.বি.-১৬)

উত্তর: অনমুনায়ন ত্রুটির কারণসমূহ নিম্নরূপ-

(ক) কিছু একক থেকে তথ্য সংগ্রহে ব্যর্থতা: কখনো কখনো নির্বাচিত একক থেকে তথ্য সংগ্রহ করা সম্ভব হয় না। আবার ডাক পদ্ধতিতে উত্তরদাতা প্রশ্নপত্র পূরণ করে না পাঠালে বা জরিপের উদ্দেশ্যের সাথে সম্পর্ক নেই এমন কাউকে প্রশ্নপত্র পাঠালে নমুনা প্রতিনিধিত্বকারী হয় না। ফলে এ ত্রুটির সৃষ্টি হয়।

(খ) ভুল তথ্য পরিবেশন: উত্তরদাতা প্রশ্ন না বুঝে কিংবা উত্তরদাতা কিংবা অনুসন্ধানকারী কিংবা উভয়ই পক্ষপাতদুষ্ট হলে বা উত্তরদাতা স্বেচ্ছায় ভুল উত্তর দিতে পারে। প্রশ্ন স্বার্থবিরোধী হলে উত্তরদাতা অনেক সময় ভুল তথ্য দিতে পারে।

(গ) উত্তরবিহীন ত্রুটি: সামাজিক, অর্থনৈতিক বা রাজনৈতিক জরিপে কোন কোন উত্তরদাতা উত্তর দেয় না। ভয়ে বা অন্য কোন কারণে অনেক উত্তরদাতা উত্তর দেয় না। বার বার গিয়ে উত্তরদাতাকে নাও পাওয়া যেতে পারে। ফলে এ জাতীয় ত্রুটির সৃষ্টি হয়।

(ঘ) ত্রুটিপূর্ণ যন্ত্রপাতি: ত্রুটিপূর্ণ যন্ত্রপাতি ব্যবহার করে তথ্য সংগ্রহ করা হলে এ জাতীয় ত্রুটির সৃষ্টি হয়।

(ঙ) ত্রুটিপূর্ণ প্রশ্নমালা: ত্রুটিপূর্ণ প্রশ্নমালা দিয়ে তথ্য সংগ্রহ করলে সে তথ্য ত্রুটিপূর্ণ হয় এবং এতে অনমুনায়ন ত্রুটির সৃষ্টি হয়।

(চ) ত্রুটিপূর্ণ নমুনায়ন কাঠামো: নমুনায়ন কাঠামো ভুল, অসম্পূর্ণ, অপর্যাপ্ত ও পুরানো হলে এ ধরণের ত্রুটির সৃষ্টি হয়।

(ছ) প্রাক-বিশ্লেষণ বিচ্যুতি: সংগৃহীত তথ্যকে বিশ্লেষণের জন্য প্রস্তুত করা হয়। বিশ্লেষণের আগে তথ্যের সম্পাদনা, কোড করা বা সারণি তৈরি করা হয়। এ সব কাজে ভুল হলে অনমুনায়ন ত্রুটির সৃষ্টি হয়।

প্রশ্ন: নমুনাজ ত্রুটি ও অনমুনাজ ত্রুটির পার্থক্য লিখ (জা.বি.-১২)

উত্তর: নমুনাজ ত্রুটি ও অনমুনাজ ত্রুটির মধ্যে পার্থক্যসমূহ নিম্নরূপ-

পার্থক্যের বিষয়	নমুনাজ ত্রুটি	অনমুনাজ ত্রুটি
(i) জরিপ	ইহা কেবলমাত্র নমুনা জরিপের ক্ষেত্রে ঘটে।	ইহা শুমারী ও নমুনা জরিপ উভয় ক্ষেত্রেই ঘটে।
(ii) নমুনার আকার	অপেক্ষাকৃত ছোট আকারের নমুনার ক্ষেত্রে এ ত্রুটির মান বেশি হয়।	ছোট আকারের নমুনার ক্ষেত্রে এ ত্রুটির মান কম হয়।
(iii) ত্রুটি পরিমাপ	সম্ভাবনা নমুনায়ন দ্বারা নমুনা সংগ্রহ করলে এ ত্রুটি পরিমাপ করা সম্ভব।	অনমুনাজ ত্রুটি পরিমাপ করা যায় না।
(iv) ত্রুটি কমানো	নমুনার আকার বৃদ্ধি করে নমুনাজ ত্রুটি কমানো যায়।	কেবল মাত্র জরিপকালে সতর্কতা অবলম্বনের দ্বারা অনমুনাজ ত্রুটি কমানো যায়।

প্রশ্ন: প্রশ্নমালা বলতে কি বুঝ (জা.বি.-১২) ? একটি আদর্শ প্রশ্নমালার বৈশিষ্ট্যগুলো উল্লেখ কর।

অথবা, প্রশ্নমালা কি ? একটি আদর্শ প্রশ্নমালার আবশ্যকীয় গুণগুলি উল্লেখ কর (জা.বি.-১২)

অথবা, একটি আদর্শ প্রশ্নমালার আবশ্যকীয় গুণগুলি উল্লেখ কর (জা.বি.-১৬)

উত্তর: কোন পরিসংখ্যানিক গবেষণার তথ্য সংগ্রহের জন্য উক্ত গবেষণা সংশ্লিষ্ট এক গুচ্ছ মুদ্রিত প্রশ্নের সমাহারকে প্রশ্নমালা বলে। অর্থাৎ প্রশ্নমালা হচ্ছে, কতগুলো প্রশ্ন যাদের মাধ্যমে জরিপের প্রয়োজনীয় উপাত্ত সংগ্রহ করা হয়। জরিপের উদ্দেশ্যের সাথে সামঞ্জস্যপূর্ণ প্রশ্ন দ্বারা প্রশ্নমালা প্রণয়ন করা হয়।

একটি ভাল প্রশ্নমালার প্রধান প্রধান বৈশিষ্ট্যগুলো নিম্নরূপ-

(ক) প্রশ্নমালায় সংশ্লিষ্ট প্রতিষ্ঠান ও ব্যক্তির পরিচিতিমূলক অংশ থাকতে হবে।

(খ) জরিপের উদ্দেশ্যেরভিত্তিতে প্রশ্নমালা প্রণয়ন করা উচিত।

(গ) প্রশ্নগুলো সংক্ষিপ্ত ও বোধগম্য হওয়া উচিত।

(ঘ) প্রশ্নের সংখ্যা এবং প্রতিটি প্রশ্ন অনাবশ্যক দীর্ঘ হওয়া উচিত নয়।

(ঙ) প্রশ্নমালার প্রশ্নগুলোর উত্তর যতদূর সম্ভব ছোট হওয়া বাঞ্চনীয়।

(চ) প্রশ্নমালার প্রশ্নগুলোর উত্তর কোন এককে প্রকাশ করতে হলে তা উল্লেখ করতে হবে।

(ছ) যতদূর সম্ভব ব্যক্তিগত প্রশ্ন বর্জন করা উচিত।

(জ) ধর্মীয় বা সাম্প্রদায়িক অনুভূতিতে আঘাত হানে এমন প্রশ্নাবলী সন্নিবেশ করা উচিত নয়।

(ঝ) প্রশ্নগুলো এমন হওয়া উচিত যেন সরাসরি ও সুস্পষ্ট তথ্য পাওয়া যায়।

(ঞ) উত্তর দাতাকে হিসাব বা গণনা করে জবাব দিতে হয় এমন প্রশ্ন বর্জন করা উচিত।

(ট) প্রশ্নমালায় উত্তর দাতাকে তথ্যের গোপনীয়তা রক্ষার জন্য আশ্বাস থাকা উচিত।

প্রশ্নঃ প্রশ্নমালা ও তালিকা (বা অনুসূচী) বলতে কি বুঝ ? উহাদের মধ্যে পার্থক্যসমূহ লিখ/ প্রশ্নমালা (questionnaire) ও সিডিউল (schedule) এর মধ্যে পার্থক্য লিখ (জা.বি.-১১)।

উত্তরঃ প্রশ্নমালা- প্রশ্নমালা হচ্ছে পরিসংখ্যানিক গবেষণার জন্য তথ্য সংগ্রহের এমন একটি হাতিয়ার যেখানে সুনির্দিষ্ট কতগুলো প্রশ্ন সন্নিবেশ করে সাধারণতঃ ডাকযোগে উত্তর দাতার নিকট প্রেরণ করা হয় এবং উত্তরদাতা নিজেই তা পূরণ করে গবেষকের কাছে ফেরত পাঠায়।

তালিকা বা অনুসূচীঃ তালিকা হলো সাক্ষাতকার গ্রহণের মাধ্যমে তথ্য সংগ্রহের জন্য কতগুলো প্রশ্নের সমাহার যার উত্তর সাক্ষাতকার গ্রহণকারী নিজেই লিপিবদ্ধ করে।

প্রশ্নমালা ও তালিকার মধ্যে পার্থক্যসমূহ নিম্নরূপ-

পার্থক্যের বিষয়	প্রশ্নমালা	তালিকা
(i) তথ্য সংগ্রহ	প্রশ্নমালা সাধারণত উত্তর দাতার নিকট ডাকযোগে পাঠানো হয় এবং উত্তর দাতা তা পূরণ করে গবেষকের কাছে ফেরত পাঠায়।	সরাসরি বা মুখোমুখি সাক্ষাৎকারের মাধ্যমে তথ্য সংগ্রহের জন্য তালিকা ব্যবহৃত হয় এবং সাক্ষাৎকার গ্রহণকারী মৌখিক উত্তর এতে লিপিবদ্ধ করেন।
(ii) খরচ ও সময়	স্বল্প ঝামেলায়, কম খরচে ও কম সময়ে তথ্য সংগ্রহ করা যায়।	অধিক ব্যয়বহুল, সময় সাপেক্ষ ও জটিল পদ্ধতি।
(iii) উত্তর দাতার ধরণ	কেবলমাত্র শিক্ষিত উত্তর দাতার পক্ষে উত্তর দেওয়া সম্ভব।	শিক্ষিত ও নিরক্ষর সকলের নিকট হতে তথ্য সংগ্রহ করা সম্ভব।
(iv) প্রশ্নের ধরণ	প্রশ্নমালার প্রশ্নগুলো সহজ, সরল এবং নৈর্ব্যক্তিক আকারে থাকে।	অপেক্ষাকৃত জটিল ও উন্নত প্রশ্নমালা নিয়ে তালিকার মাধ্যমে তথ্য সংগ্রহ করা যায়।
(v) গোপনীয়তা	ব্যক্তি একান্তে উত্তর দেওয়ার সুযোগ পায়, তাই এখানে অত্যন্ত ব্যক্তিগত ও গোপন তথ্য পাওয়ার সম্ভাবনা বেশি।	মুখোমুখি সাক্ষাৎকারের জন্য ব্যক্তি তার গোপনীয় প্রশ্নের উত্তর প্রদানে অনাগ্রহ প্রকাশ করে এবং অনেক সময় ভ্রান্ত উত্তর প্রদান করে।
(vi) উত্তর দাতার সংখ্যা	এক্ষেত্রে উত্তর না পাওয়ার সংখ্যা বেশি।	এক্ষেত্রে উত্তর পাওয়ার সংখ্যা বেশি।
(vii) গবেষকের পক্ষপাতিত্ব	গবেষকের ব্যক্তিগত পক্ষপাতিত্বের সুযোগ ও সম্ভাবনা থাকে না।	গবেষকের ব্যক্তিগত পক্ষপাতিত্বের সুযোগ ও সম্ভাবনা থাকে।

প্রশ্ন: জরিপ (Survey) কি?

উত্তর: তথ্যবিশ্বের কোন বৈশিষ্ট্য বা প্রকৃতি বের করার জন্য তথ্যবিশ্ব বা নমুনার এককগুলো হতে ঐ বৈশিষ্ট্যের উপর তথ্য সংগ্রহ করে অধ্যয়ন করা হয়। এ অনুসন্ধান কাজকে জরিপ বলে।

জরিপকে মূলতঃ দুটি ভাগে ভাগ করা যায়: (ক) শুমারি জরিপ ও (খ) নমুনা জরিপ।

প্রশ্ন: নমুনা জরিপ ও শুমারী জরিপ বলতে কি বুঝ? (জা.বি.-১৫)

(ক) শুমারি জরিপ (Census)/ শুমারী জরিপ বলতে কি বুঝ ? (জা.বি.-১১): তথ্যবিশ্বের কোন বৈশিষ্ট্য বা প্রকৃতি বের করার জন্য তথ্যবিশ্বের প্রতিটি একক হতে উক্ত বৈশিষ্ট্যের উপর তথ্য সংগ্রহ করে অধ্যয়ন করা যায়। এ অনুসন্ধান কাজকে শুমারি জরিপ বলে।

যেমন: কোন কলেজের অনার্স প্রথম বর্ষের শিক্ষার্থীদের গড় উচ্চতা অনুসন্ধানে ঐ শ্রেণীর সকল শিক্ষার্থীর উচ্চতা পরিমাপ করে অধ্যয়ন করাকে বলা হয় শুমারি জরিপ।

(খ) নমুনা জরিপ (Sample Survey)/ নমুনা জরিপ কি ? (জা.বি.-১২): যে অনুসন্ধানে তথ্যবিশ্ব হতে প্রতিনিধিত্বশীল নমুনা নির্বাচন করে নির্বাচিত নমুনা একক হতে তথ্য নিয়ে তথ্যবিশ্বের কোন বৈশিষ্ট্য পরিমাপ করা হয়, তাকে নমুনা জরিপ বলে।

যেমন: ধরি, কোন কলেজের অনার্স প্রথম বর্ষের শিক্ষার্থীর সংখ্যা ১০০০। শিক্ষার্থীদের গড় উচ্চতা অনুসন্ধানে উক্ত সমগ্রক থেকে ১০০ জন শিক্ষার্থীর উচ্চতা সংগ্রহ করে অধ্যয়ন করাকে বলা হয় নমুনা জরিপ।

প্রশ্ন: Pilot জরিপের সংজ্ঞা দাও (জা.বি.-১১)/ প্রাক জরিপ (Pilot Survey) কি ? প্রাক জরিপের উদ্দেশ্য বা কার্যাবলী উল্লেখ্য কর।

উত্তর: অনেক সময় মূল জরিপের পূর্বে তথ্যবিশ্বের একটি ক্ষুদ্র অংশের উপর ছোট আকারের জরিপ পরিচালনা করা হয়, ইহাকে প্রাক জরিপ বা পাইলট জরিপ বলা হয়। শুমারী জরিপ ও নমুনা জরিপ উভয় ক্ষেত্রেই প্রাক জরিপ করা হয়। বড় ধরণের অনুসন্ধানে প্রাক জরিপ অনেক সময় কার্যকরী ভূমিকা পালন করে থাকে।

প্রাক জরিপের কার্যাবলী নিম্নরূপ:

(ক) মূল জরিপের প্রশ্নমালা সংশোধন ও উন্নয়ন করার জন্য প্রাক জরিপ করা হয়।

(খ) নমুনা একক সম্বন্ধে ধারণা লাভ করা যায়।

(গ) মূল জরিপের সময় ও ব্যয় নির্ধারণ করা যায়।

(ঘ) মাঠের কাজ পরিচালনার পরিকল্পনা করা যায়।

প্রশ্ন: শুমারীর তুলনায় নমুনা জরিপের সুবিধাসমূহ বর্ণনা/ আলোচনা কর (জা.বি.-১২, ১৬)

অথবা, শুমারী জরিপের তুলনায় নমুনা জরিপের সুবিধাবলী আলোচনা কর।

উত্তর: নিম্নলিখিত কারণে নমুনা জরিপ শুমারি জরিপের চাইতে অধিক পছন্দনীয়:

(ক) স্বল্প ব্যয়-নমুনা জরিপে শুমারি জরিপের তুলনায় তথ্য সংগ্রহ ও বিশ্লেষণ করতে খরচ খুব কম হয়।

(খ) স্বল্প সময়- নমুনায় একক সংখ্যা সীমিত হওয়ার কারণে এ পদ্ধতিতে তথ্য সংগ্রহ ও বিশ্লেষণে স্বল্প সময়ের প্রয়োজন হয়।

(গ) তথ্যের নির্ভরযোগ্যতা: নমুনা জরিপ ছোট হওয়ায় অধিক দক্ষ ও অভিজ্ঞতা সম্পন্ন জনশক্তি নিয়োগ করা যায় এবং তথ্যের পরিমাণ কম হওয়ায় এতে ভুল ত্রুটি অনেক কম থাকে। ফলে বিশ্লেষিত ফলাফল অধিক সঠিক ও নির্ভরযোগ্য হয়।

(ঘ) অসীম তথ্যবিশ্ব: বড় তথ্যবিশ্বের ক্ষেত্রে শুমারি জরিপ প্রায় অসম্ভব। যেমন: সুন্দর বনের অসংখ্য সুন্দরী গাছের প্রতিটি সম্পর্কে তথ্য সংগ্রহ করা সম্ভব নয় বলে নমুনা জরিপই একমাত্র পথ।

(ঙ) প্রশাসনিক সুবিধা: নমুনা জরিপ পরিচালনায় জনবলের স্বল্পতার কারণে প্রশাসনিক সুবিধা বেশি পাওয়া যায়।

(চ) পুনঃ পুনঃ পরিচালনা: নমুনা জরিপের ব্যয় ও সময় স্বল্প হওয়ার কারণে একে পুনঃ পুনঃ পরিচালনা করা যায় এবং যথাসম্ভব ভুল ত্রুটি পরিহার করা যায়।

(ছ) নমুনা একক ক্ষতিগ্রস্থ হলে: যে সমস্ত ক্ষেত্রে শুমারী করা হলে নমুনা একক ক্ষতিগ্রস্থ হয়, সে সকল ক্ষেত্রে নমুনা জরিপই একমাত্র উপায়।

(জ) তাৎক্ষণিক সিদ্ধান্ত গ্রহণ: কোন জরুরি বিষয়ে তাৎক্ষণিক সিদ্ধান্ত গ্রহণের জন্য নমুনা জরিপ ব্যাপক ব্যবহৃত হয়।

(ঝ) সঠিকতা যাচাই: শুমারি জরিপের ফলাফলের সঠিক কিনা তা যাচাই করতে নমুনা জরিপ পরিচালনা করা হয়।

প্রশ্ন: নমুনা জরিপের উপর শুমারি জরিপের সুবিধাবলি আলোচনা কর।

উত্তর: যদিও শুমারি জরিপের চেয়ে নমুনা জরিপ অধিক পছন্দনীয়, তবুও কিছু কিছু ক্ষেত্রে শুমারি জরিপ অপরিহার্য হয়ে পড়ে, যা নিম্নে আলোচনা করা হল-

(ক) দেশের মোট জনসংখ্যা গণনার ক্ষেত্রে আদমশুমারি করা হয়। নমুনা জরিপ এ ক্ষেত্রে সম্পূর্ণ অনুপযোগী।

(খ) সমগ্রকের বিভিন্ন অংশের বিস্তারিত ফলাফল পাওয়ার জন্য নমুনা জরিপ অপেক্ষা শুমারি জরিপই উৎকৃষ্ট।

(গ) যদি তথ্যবিশ্বের আকার খুব ছোট হয় এবং হাতে যথেষ্ট সময় থাকে, তবে নমুনা জরিপের চেয়ে শুমারি জরিপই অধিক কার্যকরী।

(ঘ) পর্যাপ্ত অর্থ, সময় ও জনবল থাকলে নমুনা জরিপ না করে শুমারি জরিপ করলে ভাল ফলাফল পাওয়া যায়।

(ঙ) শুমারি জরিপের ফলাফল নমুনা জরিপের ফলাফল অপেক্ষা অধিক স্থায়ী এবং স্থিতিশীল। তাই অধিক স্থায়ী এবং স্থিতিশীল ফলাফল অত্যাবশ্যক হলে শুমারি জরিপের প্রয়োজন হয়।

প্রশ্ন: শুমারি জরিপ ও নমুনা জরিপের মধ্যে পার্থক্য লিখ (জা.বি.-১৫)

উত্তর: শুমারি জরিপ ও নমুনা জরিপের মধ্যে পার্থক্য নিম্নরূপ-

পার্থক্যের বিষয়	শুমারি জরিপ	নমুনা জরিপ
(i) জরিপ	যে জরিপ প্রক্রিয়ায় সমগ্রকের প্রতিটি একক হতে তথ্য সংগ্রহ করা হয়, তাকে শুমারি জরিপ বলে।	যে জরিপ প্রক্রিয়ায় সমগ্রকের কোন নমুনার প্রতিটি একক হতে তথ্য সংগ্রহ করা হয়, তাকে শুমারি জরিপ বলে।
(ii) সময় ও অর্থ	শুমারি জরিপ পরিচালনায় অধিক সময় ও প্রচুর অর্থের প্রয়োজন।	নমুনা জরিপের ক্ষেত্রে স্বল্প সময় ও অল্প অর্থের প্রয়োজন হয়।
(iii) সময়	স্বল্প সময় ব্যবধানে শুমারি জরিপ পরিচালনা সম্ভব হয় না।	যে কোন নমুনা জরিপ খুব অল্প সময়ের পরিসরে সম্পন্ন করা যায়।
(iv) নিয়ন্ত্রণ	শুমারি জরিপের পরিসর বৃহৎ হওয়ায় সার্বিক বিষয় নিয়ন্ত্রণ রাখা কষ্টকর।	নমুনা জরিপের কার্যক্ষেত্র ছোট হওয়ায় এর সার্বিক বিষয় নিয়ন্ত্রণ করা বেশ সুবিধাজনক।
(v) অভিজ্ঞ লোক	তথ্য সংগ্রহ ও বিশ্লেষণে অনেক অভিজ্ঞ লোক নিয়োগ করতে হয়।	কিছু সংখ্যক অভিজ্ঞ লোক হলে নমুনা জরিপ পরিচালনা করা যায়।

(vi) পরিধি	শুমারি জরিপের আওতা বা পরিধি ব্যাপক।	নমুনা জরিপের পরিধি সীমিত।
(vii) সঠিকতা	শুমারি জরিপের সঠিকতা যাচাই করতে নমুনা জরিপ পরিচালনা করা হয়।	নমুনা জরিপের সঠিকতা যাচাই করতে প্রকল্প যাচাই (Testing hypothesis) করা হয়।
(viii) ব্যবহার	কোন কোন বিশেষ ক্ষেত্রে শুমারি জরিপ ব্যবহার করা যায় না, যেমনঃ বাল্বের জীবনী শক্তি পরীক্ষা।	নমুনা জরিপ যে কোন অনুসন্ধানের জন্য ব্যবহার করা যায়।
(ix) অনমুনাজ ক্রটি	শুমারি জরিপে অনমুনাজ ক্রটি বেশি হয়।	নমুনা জরিপে অনমুনাজ ক্রটি কম হয়।

প্রশ্ন: একটি নমুনা জরিপ পরিচালনার জন্য বিভিন্ন ধাপ সংক্ষেপে আলোচনা কর (জা.বি.-১১)

অথবা, নমুনা জরিপের প্রধান ধাপসমূহ কি কি? এদের সংক্ষিপ্ত বর্ণনা দাও (জা.বি.-১৬)

অথবা, নমুনা জরিপের বিভিন্ন ধাপগুলো আলোচনা কর।

উত্তর: কোন নমুনা জরিপ পরিকল্পনা ও বাস্তবায়নে যে সব ধাপগুলো অনুসরণ করা হয়, তা নিম্নে আলোচনা করা হল-

(১) **জরিপের উদ্দেশ্য:** প্রথমে জরিপের উদ্দেশ্য স্পষ্ট সঠিকভাবে সংজ্ঞায়িত করতে হবে। জরিপের উদ্দেশ্যের সংগে প্রাপ্ত সম্পদ, জনশক্তি ও সময় সম্পর্কযুক্ত। তাছাড়া জরিপের উদ্দেশ্যের উপর ভিত্তি করে উপাত্ত সংগ্রহ ও বিশ্লেষণ করা হয়। তাই জরিপের উদ্দেশ্য সঠিকভাবে সংজ্ঞায়িত করা নমুনা জরিপের একটি গুরুত্বপূর্ণ ধাপ।

(২) **তথ্যবিশ্ব বা সমগ্রকের সংজ্ঞা:** অত্যন্ত স্পষ্ট ও সঠিকভাবে তথ্যবিশ্ব সংজ্ঞায়িত করতে হবে। তথ্যবিশ্বের ভৌগোলিক, ডেমোগ্রাফিক এবং অন্যান্য সীমারেখা স্পষ্টভাবে নির্ধারণ করতে হবে যেন কোনরূপ অস্পষ্টতা না থাকে।

(৩) **কাঠামো এবং নমুনা একক:** তথ্যবিশ্ব যে সব বস্তু বা ব্যক্তি নিয়ে গঠিত হয়, তাদের প্রত্যেকটিকে উপাদান বা নমুনা একক বলা হয়। নমুনা এককগুলো ভিন্ন ও পরস্পর বর্জনশীল এবং তাদের সমষ্টি নিয়ে তথ্যবিশ্ব গঠিত হয়। এ এককগুলো একজন মানুষ, একটি পরিবার, একটি খামার, নির্দিষ্ট আকারের জমির প্লট ইত্যাদি হতে পারে। তথ্যবিশ্বের অন্তর্ভূক্ত সমস্ত উপাদান বা এককগুলোর তালিকাকে কাঠামো বলে। কাঠামোতে উপাদানগুলোর নাম, ঠিকানা অর্থাৎ চিহ্নিতকরণ বৈশিষ্ট্য সুস্পষ্ট থাকে। কাঠামোর উপর ভিত্তি করেই নমুনা চয়ন করা হয়। কাঠামো পুরাতন বা অসম্পূর্ণ হলে ইহাকে সংশোধন করে তথ্য সংগ্রহ করা হয়। কাঠামো হালনাগাদ করা নমুনা জরিপের একটি উল্লেখযোগ্য ধাপ।

(৪) **উপাত্তের প্রকৃতি:** কি ধরণের উপাত্ত বা তথ্য সংগ্রহ করতে হবে তা সম্পূর্ণরূপে নির্ভর করে জরিপের উদ্দেশ্যের উপর। উপাত্ত সংগ্রহের সময় অত্যন্ত সতর্কতার সংগে লক্ষ্য রাখতে হবে যেন কোন প্রয়োজনীয় তথ্য বাদ না পড়ে। কারণ এ উপাত্তের উপরই নির্ভর করবে সমস্ত জরিপের সাফল্য।

(৫) **প্রশ্নমালা প্রণয়ন:** প্রশ্নমালা হচ্চে কতগুলো প্রশ্ন যাদের মাধ্যমে জরিপের প্রয়োজনীয় তথ্য বা উপাত্ত সংগ্রহ করা হয়। প্রশ্নগুলো সাধারণত সহজ, সরল, সংক্ষেপ ও স্পষ্ট হওয়া বাঞ্চনীয় যেন উত্তরদাতা অতি সহজে এবং সংক্ষেপে উত্তর দিতে পারে। অনেক সময় বিশেষ করে জাতীয় পর্যায়ে মূল জরিপের পূর্বে প্রশ্নমালা যাচাই করে সংশোধন করা হয়।

(৬) **তথ্য সংগ্রহ:** প্রশ্নমালা প্রণয়নের পর ইহার মাধ্যমে তথ্য সংগ্রহ করা হয়। সাধারণত দু'ভাবে তথ্য সংগ্রহ করা হয়। যথা- ক) সরাসরি জিজ্ঞাসাবাদ পদ্ধতি, খ) ডাকযোগে প্রশ্নমালা প্রেরণ পদ্ধতি।

(ক) সরাসরি জিজ্ঞাসাবাদ পদ্ধতি: এ পদ্ধতিতে তথ্য সংগ্রহকারী বাড়ি গিয়ে সরাসরি জিজ্ঞাসাবাদের মাধ্যমে তথ্য সংগ্রহ করে লিপিবদ্ধ করে।

(খ) ডাকযোগে প্রশ্নমালা প্রেরণ পদ্ধতি: এ পদ্ধতিতে প্রশ্নমালা ডাকযোগে উত্তরদাতার নিকট পাঠানো হয়। উত্তরদাতা প্রশ্নমালা পূরণ করে ডাকযোগে অনুসন্ধানকারীর নিকট ফেরত পাঠান।

(৭) নিরবতা: অনেক সময়ে নমুনার সকল একক থেকে উপাত্ত সংগ্রহ করা সম্ভব হয় না। এটা নানা কারণে হতে পারে। যেমন: কাঠামোতে ভুল থাকা, উত্তরদাতার কোন তথ্য দিতে আপত্তি ইত্যাদি। এগুলোকে উত্তরদাতার নিরবতা হিসেবে গণ্য করা হয়। উপাত্ত সংগ্রহকারীকে উত্তর দাতার নিরবতার কারণ সঠিকভাবে লিপিবদ্ধ করতে হবে এবং সম্ভব হলে এটা কমাতে হবে।

(৮)যথাযথ নমুনায়ন পদ্ধতি নির্বাচন: উপযুক্ত নমুনায়ন পদ্ধতি নির্বাচন, নমুনার আকার নির্ধারণ ও তথ্যবিশ্বের পরামান প্রাক্কলন নমুনা জরিপের গুরত্বপূর্ণ ধাপ। কোন্ ক্ষেত্রে কি ধরণের নমুনায়ন পদ্ধতি নির্ধারণ করা হবে তা সম্পূর্ণরূপে নির্ভর করবে তথ্যবিশ্বের উপাদানগুলোর বৈশিষ্ট্যের উপর। গুরুত্বপূর্ণ নমুনায়ন পদ্ধতি হলো সরল দৈব নমুনায়ন, স্তরিত দৈব নমুনায়ন, ধারাবাহিক নমুনায়ন, গুচ্ছ নমুনায়ন ইত্যাদি।

(৯) মাঠের কাজ পরিচালনা: কোন জরিপের মাঠে কাজ কলতে বুঝায়, তথ্য সংগ্রহকারীদের দ্বারা উপাত্ত সংগ্রহ করা এবং পরিদর্শকদের দ্বারা তদারকি করা। যেহেতু তথ্য সংগ্রহ একটি গুরুত্বপূর্ণ কাজ, তাই তথ্য সংগ্রহকারীদের যথেষ্ট প্রশিক্ষণ দেওয়া একান্ত প্রয়োজন যাতে নমুনা এককগুলো সঠিকভাবে নির্ধারণ করে প্রয়োজনীয় তথ্য সংগ্রহ করতে পারে।

(১০) প্রাক জরিপ (Pilot Survey): অনেক সময় মূল জরিপের পূর্বে তথ্যবিশ্বের একটি ক্ষুদ্র অংশের উপর ছোট আকারের জরিপ পরিচালনা করা হয়। এটাকে প্রাক জরিপ বা পাইলট জরিপ বলা হয়। বৃহৎ আকারের অনুসন্ধানে প্রাক জরিপ গুরুত্বপূর্ণ ভূমিকা পালন করে। প্রাক জরিপের মাধ্যমে মূল জরিপের প্রশ্নমালা সংশোধন ও উন্নয়ন, নমুনা একক সম্বন্ধে ধারণা, মূল জরিপের সময় ও ব্যয় নির্ধারণ এবং মাঠ পর্যায়ের কাজ পরিচালনার পরিকল্পনা করা হয়।

(১১) তথ্য বিশ্লেষণ ও রিপোর্ট প্রণয়ন: সংগৃহিত তথ্য প্রথমে পুঙ্খানুপঙ্খভাবে পরীক্ষা করা হয়, যাতে বড় ধরণের ভুল ও অসামঞ্জস্যতা দূর করা হয়। ইহাকে সমীক্ষা বা সম্পাদনা বলা হয়। অত:পর শ্রেণীকরণ ও সারণীকরণের মাধ্যমে উপাত্ত সংক্ষিপ্তকরণ করা হয়। জরিপ বড় আকারের হলে সারণীকরণ মেশিনের সাহায্যে করা হয়। এ পর্যায়ে তথ্যবিশ্বের পরামানগুলো প্রাক্কলন করা হয় এবং যাচাই করা হয়। জরিপের বিভিন্ন ধাপের সংগে সামঞ্জস্য রেখে প্রাপ্ত ফলাফলের ভিত্তিতে একটি রিপোর্ট তৈরী করা হয়। এ রিপোর্টের উপর ভিত্তি করে জরিপের সার্বিক মূল্যায়ন করা হয়।

(১২) ভবিষ্যত জরিপের পথ নির্দেশক: যে কোন জরিপই ভবিষ্যতে ঐ ধরণের জরিপের ভিত্তি হিসেবে কাজ করে। একটি জরিপ বিশ্লেষণ করে যে সব ফলাফল বা প্রাক্কলিত মান পাওয়া যায়, তা ভবিষ্যত জরিপের পথ নির্দেশ করে।

প্রশ্ন: একটি নমুনা জরিপের মূলনীতিসমূহ (basic principles of a sample survey) আলোচনা কর (জা.বি.-১২)

উত্তর: নমুনায়ন তত্ত্বের উদ্দেশ্য হলো নমুনায়নকে বেশি কার্যকরী করা যেন কোন একটি নির্দিষ্ট সমস্যার সমাধান যথার্থ, পর্যাপ্ত ও অর্থনৈতিক দিক থেকে লাভজনকভাবে দেয়া যায়। নমুনায়ন তত্ত্ব তিনটি মূলনীতির উপর ভিত্তি করে প্রতিষ্ঠিত। মূলনীতিগুলো নিম্নরূপ-

(ক) যথার্থতার মূলনীতি (Principle of validity)

(খ) পরিসংখ্যানিক ধারাবাহিকতার মূলনীতি (Principle of statistical regularity)

(গ) সর্বোত্তম হবার মূলনীতি (Principle of optimization)

(ক) যথার্থতার মূলনীতি (Principle of validity): এ মূলনীতির ধারণা হচ্ছে নমুনায়ন নকশা যেন তথ্যবিশ্ব পরামিতির যথার্থ নিরূপক প্রদান করে। যথার্থতা বলতে আমরা বুঝি, নমুনা নির্বাচন এমনভাবে করতে হবে যেন নিরূপকগুলোকে উদ্দেশ্যপূর্ণভাবে এবং সম্ভাবনারভিত্তিতে ব্যাখ্যা করা যায়। অর্থাৎ উক্ত মূলনীতি নিশ্চিত করে যে, নমুনা নকশার প্রতিটি এককের একটি নির্দিষ্ট ও পূর্ব নির্ধারিত সম্ভাবনা থাকে।

(খ) পরিসংখ্যানিক ধারাবাহিকতার মূলনীতি (Principle of statistical regularity): এ মূলনীতি সম্ভাবনা তত্ত্বের উপর ভিত্তি করে প্রতিষ্ঠিত এবং ইহাকে নিম্নরূপে বর্ণনা করা যায়:

"একটি বড় আকারের তথ্যবিশ্ব হতে তুলনামূলকভাবে বড় আকারের দৈব নমুনা চয়ন করা হলে উক্ত নমুনাটি তথ্যবিশ্বের বৈশিষ্ট্যের পূর্ণ অধিকারী হবে অর্থাৎ নমুনাটি প্রতিনিধিত্বশীল হবে।"

(গ) সর্বোত্তম হবার মূলনীতি (Principle of optimization): এ মূলনীতির ধারণা এমন নমুনায়ন নকশা নির্বাচন করা যা সর্বোত্তম ফলাফল প্রদান করে। অন্যভাবে, সর্বোত্তম বলতে বুঝায় এমনভাবে নমুনা নির্বাচন করা এবং পরামিতির মান প্রাক্কলন করা যেন-

-সর্বনিম্ন সম্ভাবন্য সম্পদের মাধ্যমে একটি নির্দিষ্ট পরিমাণ দক্ষ নিরূপক পাওয়া যায় অথবা

-সর্বোচ্চ সম্ভাব্য দক্ষতার মাধ্যমে একটি নির্দিষ্ট পরিমাণ খরচে ভাল নিরূপক পাওয়া যায়।

এভাবে এ মূলনীতি নমুনা নকশার ঝুঁকি বা ক্ষতি কমাতে সাহায্য করে। এ মূলনীতি খরচ ও গড় বর্গ বিচ্যুতির ভিত্তিতে মোট ক্ষতিকে সর্বনিম্ন করে সর্বোত্তম ফলাফর পাওয়ার উপর গুরত্ব আরোপ করে।

প্রশ্ন: নমুনায়ন কাঠামো কি? একটি নমুনায়ন কাঠামোর কি কি প্রয়োজনীয় গুণাবলী থাকা প্রয়োজন আলোচনা কর (জা.বি.-১৫)

উত্তর: নমুনায়ন কাঠামো-যে সমগ্রক হতে নমুনা নেয়া হবে সে সমগ্রকের এককগুলোর তালিকাকে নমুনায়ন কাঠামো বলে। কাঠামোতে এককগুলোকে ক্রমিক নম্বর দিয়ে চিহ্নিত করা হয় এবং নাম, ঠিকানা ও অন্যান্য চিহ্নিতকরণ বৈশিষ্ট্য সুস্পষ্ট থাকে। কাঠামোর উপর ভিত্তি করে নমুনা চয়ন করা হয়। তাই কাঠামো সঠিকভাবে করা না হলে নমুনায়নের আসল উদ্দেশ্য ব্যাহত হবে। একটি নমুনায়ন কাঠামোর কিছু প্রয়োজনীয় গুণাবলী থাকা আবশ্যক। নিম্নে তা সংক্ষেপে আলোচনা করা হল:

(ক) নমুনা কাঠামো হালনাগাদ হতে হবে। অতি পুরাতন নমুনা কাঠামো ব্যবহার করে নমুনা নির্বাচন করলে প্রতিনিধিত্বশীল নমুনা পাওয়া যাবে না।

(খ) নমুনা কাঠামো পূর্ণাঙ্গ হতে হবে। তথ্যবিশ্বের কোন একক যেন নমুনা কাঠামো থেকে বাদ না পড়ে, আবার কোন একক যেন একাধিকবার অন্তর্ভূক্ত না হয় সেদিকেও খেয়াল রাখতে হবে।

(গ) নমুনা কাঠামো সঠিক হতে হবে। তথ্যবিশ্বের অন্তর্ভূক্ত নয় এমন কোন একক যেন কাঠামোতে না থাকে সেদিকে লক্ষ্য রাখতে হবে।

(ঘ) নমুনা কাঠামো পর্যাপ্ত হতে হবে। ইহাতে যেন তথ্যবিশ্বের সকল শ্রেণীর সকল উপাদান অন্তর্ভূক্ত থাকে।

একটি ভাল নমুনা কাঠামো তৈরী করা খুব কষ্টসাধ্য কাজ। তবে অভিজ্ঞ গবেষকের দ্বারা তা সম্ভব। কাঠামো তৈরী বা হালনাগাদ করা নমুনা জরিপের একটি উল্লেখযোগ্য ধাপ।

প্রশ্নঃ নমুনায়ন (Sampling) বলতে কি বুঝ ? নমুনায়নের প্রকারভেদ আলোচনা কর।

অথবা, দৈব নমুনায়ন ও ঐচ্ছিক নমুনায়ন বলতে কি বুঝ ? (জা.বি.-১২)

উত্তরঃ সমগ্রক হতে প্রতিনিধিত্বশীল নমুনা নির্বাচন, প্রাক্কলক উদ্ভাবন এবং প্রাক্কলিত মানের ভিত্তিতে সমগ্রক সম্বন্ধে সিদ্ধান্ত নেয়ার বৈজ্ঞানিক পদ্ধতিকে নমুনায়ন বলে।

নমুনায়নের প্রকারভেদ (Types of Sampling):

তথ্যের প্রকৃতি ও অনুসন্ধানের উপর ভিত্তি করে নমুনা নির্বাচনের কৌশল নির্ধারণ করা হয়। নমুনা নির্বাচনের পদ্ধতিকে প্রধানত: তিন ভাগে ভাগ করা যায়। যথা:

(১) ঐচ্ছিক বা উদ্দেশ্যমূলক বা বিচারভিত্তিক নমুনায়ন (Purposive Sampling)

(২) দৈব বা সম্ভাবনা নমুনায়ন (Random Sampling)

(৩) মিশ্র নমুনায়ন (Mixed Sampling)

(১) ঐচ্ছিক নমুনায়ন (Purposive Sampling): কোন বিশেষ উদ্দেশ্যে তথ্য অনুসন্ধানকারী নিজের বিচার বুদ্ধি ও অভিজ্ঞতার আলোকে নমুনা নির্বাচন করেন। এরূপ নমুনায়নকে ঐচ্ছিক নমুনায়ন বলে।

যেমন: কোন ফলের আড়ত থেকে ক্রেতা যখন ফল বাছাই করেন, তখন তিনি ফলের গুনগত মান, ওজন, রং ইত্যাদির উপর গুরুত্ব আরোপ করেন। এরূপ নমুনায়ন হল ঐচ্ছিক নমুনায়ন।

অথবা, ক্রিকেট খেলার একটি নির্দিষ্ট ম্যাচ কোন কোন খেলোয়াড় খেলবে তা অধিনায়ক নিজে তার অভিজ্ঞতার উপর নির্ভর করে নির্বাচন করেন, ইহাকে উদ্দেশ্যমূলক বা বিচারভিত্তিক নমুনায়ন বলে।

বি. দ্রঃ দ্রুত এবং সস্তায় তথ্য পাওয়ার সুবিধা থাকাতে বৃহৎ আকারের সমগ্রক থেকে পরিসংখ্যানবিদগণ এ পদ্ধতিতে নমুনা নির্বাচন করতে পছন্দ করেন। এ ধরণের নমুনায়ন পদ্ধতি খুব কম ব্যবহৃত হতে দেখা যায় এবং এতে ব্যক্তিগত ঝুঁকি থাকে।

(২) দৈব বা সম্ভাবনা নমুনায়ন (Random Sampling)/ সম্ভাবনাযুক্ত নমুনায়নের সংজ্ঞা দাও (জা.বি.-১১): যে নমুনায়ন পদ্ধতিতে সমগ্রকের প্রতিটি এককের নমুনায় অন্তর্ভুক্ত হওয়ার একটি নির্দিষ্ট সম্ভাবনা থাকে, তাকে দৈব বা সম্ভাবনা নমুনায়ন বলে।

বি. দ্রঃ দৈব নমুনায়নে-

(ক) প্রতিটি এককের নির্বাচিত হওয়ার সম্ভাবনা সমান হতে পারে, যেমন: সরল দৈব নমুনায়ন।

(খ) এককগুলোর নির্বাচিত হওয়ার সম্ভাবনা বিভিন্ন হতে পারে, যেমনঃ স্তরিত নমুনায়ন।

(গ) প্রতিটি এককের নির্বাচিত হওয়ার সম্ভাবনা এককের আকারের সমানুপাতিক, যেমনঃ গুচ্ছ নমুনায়ন।

(৩) মিশ্র নমুনায়ন (Mixed Sampling): যে নমুনায়নে নমুনার কিছু অংশ সম্ভাবনাভিত্তিক এবং বাকি অংশ বিশেষ কোন নমুনায়ন পদ্ধতি দ্বারা চয়ন করা হয়, তাকে মিশ্র নমুনায়ন বলে।

যেমনঃ ধারাবাহিক নমুনায়ন পদ্ধতি। এখানে প্রথম এককটি দৈবভাবে এবং বাকি এককগুলো ধারাবাহিকভাবে চয়ন করা হয়।

প্রশ্নঃ দৈব নমুনায়ন ও ঐচ্ছিক নমুনায়ন-এদের মধ্যে পার্থক্য কর (জা.বি.-১২, ১৫)

অথবা, উদ্দেশ্যমূলক ও সম্ভাবনা নমুনায়নের মধ্যে পার্থক্য লিখ।

উত্তরঃ উদ্দেশ্যমূলক ও সম্ভাবনা নমুনায়নের মধ্যে পার্থক্যসমূহ নিম্নরূপ-

পার্থক্যের বিষয়	উদ্দেশ্যমূলক নমুনায়ন	দৈব বা সম্ভাবনা নমুনায়ন
(i) সম্ভাবনা	এ নমুনায়ন সম্ভাবনার উপর নির্ভরশীল নয়।	এ নমুনায়ন সম্ভাবনার উপর নির্ভরশীল।
(ii) পক্ষপাতিত্ব	এতে ব্যক্তিগত পক্ষপাতিত্বের ঝোঁক থাকে।	এতে ব্যক্তিগত পক্ষপাতিত্বের কোন ঝোঁক থাকে না।
(iii) নমুনা কাঠামো	নমুনা নির্বাচনে নমুনা কাঠামোর প্রয়োজন হয় না।	নমুনা নির্বাচনে নমুনা কাঠামোর প্রয়োজন হয়।
(iv) দৈব সংখ্যা সারণী বা লটারী	নমুনা নির্বাচনে দৈব সংখ্যার সারণী ব্যবহার অথবা লটারী করতে হয় না।	নমুনা নির্বাচনে দৈব সংখ্যার সারণী ব্যবহার অথবা লটারী করতে হয়।
(v) অনমুনাজ ক্রুটি	এ নমুনায়নে অনমুনাজ ক্রুটি বেশি হয়।	এ নমুনায়নে অনমুনাজ ক্রটি বেশি হয়।
(vi) বিচ্যুতি	এ নমুনায়নে বিচ্যুতির পরিমাপ করা যায় না।	এ নমুনায়নে বিচ্যুতির পরিমাপ করা যায়।

সম্ভাবনা নমুনায়নের প্রকারভেদঃ নমুনা জরিপে ব্যবহৃত কয়েকটি দৈব নমুনায়ন নিম্নরূপ-

(ক) সরল দৈব নমুনায়ন (Simple Random Sampling)

(খ) স্তরিত দৈব নমুনায়ন (Stratified Random Sampling)

(গ) ধারাবাহিক বা নিয়মমাফিক নমুনায়ন (Systematic Sampling)

(ঘ) গুচ্ছ নমুনায়ন (Cluster Sampling)

সংক্ষিপ্ত প্রশ্ন ও উত্তরঃ

(১) এমন একটি ক্ষেত্র উল্লেখ কর, যেখানে নমুনা জরিপ অপেক্ষা শুমারী জরিপ অধিক উপযুক্ত ? (জা.বি.-১২)

উত্তরঃ সমগ্রকের বিভিন্ন অংশের বিস্তারিত ফলাফল পাওয়ার জন্য নমুনা জরিপ অপেক্ষা শুমারী জরিপই উৎকৃষ্ট।

(২) লা-জবাব (non-response) বলতে কি বুঝ ? (জা.বি.-১২)

উত্তরঃ সামাজিক, অর্থনৈতিক বা রাজনৈতিক জরিপে কোন কোন উত্তরদাতা উত্তর দেয় না। ভয়ে বা অন্য কোন কারণে অনেক উত্তরদাতা উত্তর প্রদান থেকে বিরত থাকে। আবার বার বার গিয়েও উত্তরদাতাকে নাও পাওয়া যেতে পারে। উক্তরূপ উত্তর বা জবাব না পাওয়াই লা-জবাব হিসেবে বিবেচিত।

(৩) কাঠামো (frame) কি ? (জা.বি.-১১)

উত্তরঃ পরিসংখ্যানিক তথ্য সংশ্লিষ্ট কোন এককগুলোর তালিকাকে কাঠামো বলে। এককগুলোকে এমনভাবে সন্নিবিষ্ট ও সাজানো হয় যেন তালিকায় প্রত্যেকটি একক কেবল একবার অন্তর্ভুক্ত হয় এবং কোন একক যেন তালিকা হতে বাদ না যায়।

যেমনঃ ধরি, কোন কলেজে সম্মান ১ম বর্ষের শিক্ষার্থীদের সংখ্যা ২,০০০। যদি শিক্ষার্থীদেরকে ১ থেকে ২০০০ পর্যন্ত ক্রমানুসারে তালিকাভুক্ত করা হয়, তবে এই তালিকাই হচ্ছে কাঠামো।

(৪) সমগ্রক ইউনিট (populaton unit) কি ? (জা.বি.-১১)

উত্তরঃ সমগ্রকের প্রতিটি একককে সমগ্রক ইউনিট বলে। সমগ্রকের মোট একক সংখ্যা হলো সমগ্রকের আকার (Size) এবং একে সাধারণত N দিয়ে প্রকাশ করা হয়।

যেমনঃ কোন সিটি কর্পোরেশনের মেয়র নির্বাচনে উক্ত সিটি কর্পোরেশনের সকল রেজিস্টার্ড ভোটার নিয়ে হবে একটি সমগ্রক। এখানে প্রত্যেক রেজিস্টার্ড ভোটার হলো একক।

(৫) নমুনা জরিপের মূলনীতিগুলো কি কি ? (জা.বি.-১১)

উত্তরঃ নমুনা জরিপের মূলনীতিসমূহ নিম্নরূপ-

(ক) যথার্থতার মূলনীতি

(খ) পরিসংখ্যানিক ধারাবাহিকতার মূলনীতি

(গ) সর্বোত্তম হবার মূলনীতি

(৬) Mixed নমুনা বলতে কি বুঝ ? (জা.বি.-১১)

উত্তরঃ যে নমুনার কিছু অংশ সম্ভাবনাভিত্তিক এবং বাকি অংশ বিশেষ কোন নমুনায়ন পদ্ধতি দ্বারা চয়ন করা হয়, তাকে মিশ্র নমুনা বলে।

যেমনঃ ধারাবাহিক নমুনায় প্রথম এককটি দৈবভাবে এবং বাকি এককগুলো ধারাবাহিকভাবে চয়ন করা হয়।

(৭) নমুনা আকার (sample size) বলতে কি বুঝ ? (জা.বি.-১১)

উত্তরঃ নমুনার মোট এককের সংখ্যাকে নমুনা আকার বলে। নমুনার আকারকে সাধারণত n (যেখানে, n একটি ধনাত্মক পূর্ণ সংখ্যা এবং $n \geq 2$) দ্বারা সূচিত করা হয়।

যেমন: সাবানের নমুনায় 20 টি সাবান থাকলে নমুনার আকার হবে 20।

বি. দ্রঃ নমুনার আকার 30 বা তার কম হলে ঐ নমুনাকে ক্ষুদ্র নমুনা বলা হয়। আবার নমুনার আকার 30 এর বেশি হলে ঐ নমুনাকে বৃহৎ নমুনা বলা হয়।

(৮) নমুনা জরিপে বিচ্যুতিসমূহের উৎস (sources of errors in sample survey) কি ? (জা.বি.-১১)

উত্তর: নমুনা জরিপে বিচ্যুতিসমূহের উৎস নিম্নরূপ-

(ক) ত্রুটিপূর্ণ নমুনায়ন ডিজাইন নির্বাচন;

(খ) নমুনার আকার নির্ধারণে ত্রুটি;

(গ) নমুনা একক নির্বাচনে ত্রুটি;

(ঘ) ত্রুটিপূর্ণ সীমাযুক্ত নমুনা একক নির্ধারণ;

(ঙ) ত্রুটিপূর্ণ প্রাক্কলক ব্যবহার।

(৯) নমুনা সংগ্রহের বিভিন্ন পদ্ধতিগুলো কি কি ? (জা.বি.-১১)

উত্তর: নমুনা সংগ্রহের বিভিন্ন পদ্ধতিগুলো নিম্নরূপ:

(১) ঐচ্ছিক বা উদ্দেশ্যমূলক বা বিচারভিত্তিক নমুনায়ন

(২) দৈব বা সম্ভাবনা নমুনায়ন

(৩) মিশ্র নমুনায়ন

(১০) পুন:স্থাপন করে ও পুন:স্থাপন না করে নমুনায়নের মধ্যে পার্থক্য লিখ (Write the differences between sampling with replacement and sampling without replacement)

অথবা, পুনঃস্থাপনসহ ও পুনঃস্থাপন ছাড়া নমুনায়নের মধ্যে পার্থক্য লিখ (জা.বি.-১৬)

উত্তর: পুন:স্থাপন করে ও পুন:স্থাপন না করে নমুনায়নের মধ্যে পার্থক্য নিম্নরূপ-

(ক) পুন:স্থাপন করে নমুনায়নে যে কোন উত্তোলনে নির্বাচিত একককে পরবর্তী উত্তোলনের পূর্বে সমগ্রকে ফেরত দেয়া হয়। অন্যদিকে পুন:স্থাপন না করে নমুনায়নে কোন উত্তোলনেই নির্বাচিত একককে পরবর্তী উত্তোলনের পূর্বে সমগ্রকে ফেরত দেয়া হয় না।

(খ) পুন:স্থাপন করে নমুনায়নে প্রতিটি উত্তোলনে সমগ্রকের আকার অপরিবর্তিত থাকে। অন্যদিকে পুন:স্থাপন না করে নমুনায়নে প্রতিটি উত্তোলনের পর সমগ্রকের আকার এক করে কমতে থাকে।

(গ) পুন:স্থাপন করে নমুনায়নে একই একক একাধিকবার নমুনায় অন্তর্ভূক্ত হতে পারে। অন্যদিকে পুন:স্থাপন না করে নমুনায়নে কোন একটি একক নমুনায় সর্বোচ্চ একবার অন্তর্ভূক্ত হতে পারে।

(ঘ) পুনস্থাপন করে নমুনায়নে উত্তোলনগুলো পরস্পর স্বাধীন। অন্যদিকে পুনস্থাপন না করে নমুনায়নে উত্তোলনগুলো পরস্পর অধীন।

(ঙ) পুন:স্থাপন করে নমুনায়নে N আকারের সমগ্রক হতে n আকারের N^n টি সরল দৈব নমুনা নির্বাচন করা যায়। অন্যদিকে পুন:স্থাপন না করে নমুনায়নে N আকারের সমগ্রক হতে n আকারের ${}^{N}C_{n}$ টি সরল দৈব নমুনা নির্বাচন করা যায়।

(চ) সাধারণত বড় আকারের সমগ্রক হতে পুন:স্থাপন করে নমুনায়নে নমুনা নির্বাচন করা হয়। অন্যদিকে পুন:স্থাপন না করে নমুনায়নে নমুনা গড়ের আদর্শ বিচ্যুতি ছোট হয়ে থাকে।

(১১) প্রতিস্থাপন ও অপ্রতিস্থাপন পদ্ধতির নমুনা চয়ন বলতে কি বুঝ? কোন পদ্ধতিটি উত্তম এবং কেন? (জা.বি.-১১)

উত্তর: প্রতিস্থাপন পদ্ধতিতে নমুনা চয়ন-

প্রতিস্থাপন পদ্ধতির নমুনা চয়নে যে কোন উত্তোলনে নির্বাচিত একককে পরবর্তী উত্তোলনের পূর্বে সমগ্রকে ফেরত দেয়া হয়। এ পদ্ধতিতে নমুনায়নে প্রতিটি উত্তোলনে সমগ্রকের আকার অপরিবর্তিত থাকে এবং একই একক একাধিকবার নমুনায় অন্তর্ভূক্ত হতে পারে। প্রতিস্থাপন পদ্ধতিতে নমুনা চয়নে উত্তোলনগুলো পরস্পর স্বাধীন। এ পদ্ধতিতে N আকারের সমগ্রক হতে n আকারের N^n টি সরল দৈব নমুনা নির্বাচন করা যায়।

অপ্রতিস্থাপন পদ্ধতির নমুনা চয়ন-

অপ্রতিস্থাপন পদ্ধতির নমুনা চয়নে কোন উত্তোলনেই নির্বাচিত একককে পরবর্তী উত্তোলনের পূর্বে সমগ্রকে ফেরত দেয়া হয় না। এ পদ্ধতিতে নমুনায়নে প্রতিটি উত্তোলনের পর সমগ্রকের আকার এক করে কমতে থাকে এবং কোন একটি একক নমুনায় সর্বোচ্চ একবার অন্তর্ভূক্ত হতে পারে। অপ্রতিস্থাপন পদ্ধতিতে নমুনা চয়নে উত্তোলনগুলো পরস্পর অধীন। এ পদ্ধতিতে N আকারের সমগ্রক হতে n আকারের $^{N}C_{n}$ টি সরল দৈব নমুনা নির্বাচন করা যায়।

প্রতিস্থাপন ও অপ্রতিস্থাপন পদ্ধতির মধ্যে অপ্রতিস্থাপন পদ্ধতি উত্তম কেননা অপ্রতিস্থাপন পদ্ধতিতে নমুনা চয়নে নমুনা গড়ের আদর্শ বিচ্যুতি ছোট হয়ে থাকে।

(১২) নমুনা জরিপে উত্তর না দেয়া (non-response in sample survey) বলতে কি বুঝ ? (জা.বি.-১১)

উত্তর: নমুনা জরিপে অনেক সময়ে নমুনার সকল একক থেকে উপাত্ত সংগ্রহ করা সম্ভব হয় না। এটা নানা কারণে হতে পারে। তবে এর মধ্যে অন্যতম হচ্ছে উত্তরদাতার কোন তথ্য দিতে আপত্তি। এগুলোকে উত্তরদাতার নিরবতা হিসেবে গণ্য করা হয়। উপাত্ত সংগ্রহকারীকে উত্তর দাতার নিরবতার কারণ সঠিকভাবে লিপিবদ্ধ করতে হবে এবং সম্ভব হলে এটা কমাতে হবে।

সরল দৈব নমুনায়ন (Simple Random Sampling)

প্রশ্ন: সরল দৈব নমুনায়ন (Simple Random Sampling) কি? (জা.বি.-১১, ১৫)

উত্তর: যে নমুনায়ন পদ্ধতিতে সমগ্রকের প্রতিটি এককের নমুনায় অন্তর্ভূমক্ত হওয়ার স্বাধীন ও সমান সম্ভাবনা থাকে, তাকে সরল দৈব নমুনায়ন বলে। নির্বাচিত নমুনাকে সরল দৈব নমুনা বলা হয়। এ পদ্ধতিতে কোন নির্দিষ্ট আকারের প্রতিটি নমুনা নির্বাচিত হওয়ার সম্ভাবনা সমান।

এককের নমুনায় অন্তর্ভুক্তির ভিত্তিতে সরল দৈব নমুনায়নকে দু' ভাগে ভাগ করা যায়। যথাঃ

(ক) পুনঃস্থাপন করে সরল দৈব নমুনায়ন (Simple random sampling with replacement)

(খ) পুনঃস্থাপন না করে সরল দৈব নমুনায়ন (Simple random sampling without replacement)

(ক) পুনঃস্থাপন করে সরল দৈব নমুনায়ন (Simple random sampling with replacement):

যে সরল দৈব নমুনায়নে তথ্যবিশ্ব হতে একটি একক নির্বাচন করে সেটিকে পুনরায় তথ্যবিশ্বে ফেরত দিয়ে পরবর্তী এককটি নির্বাচন করা হয়, তাকে পুনঃস্থাপন সহকারে সরল দৈব নমুনায়ন বলে।

এক্ষেত্রে N আকারের তথ্যবিশ্ব থেকে n আকারের মোট N^n টি নমুনা পাওয়া যায়। তবে N^n টি নমুনা মধ্যে সবগুলো নমুনা ভিন্ন নয়।

(খ) পুনঃস্থাপন না করে সরল দৈব নমুনায়ন (Simple random sampling without replacement):

যে সরল দৈব নমুনায়নে তথ্যবিশ্ব হতে একটি একক নির্বাচন করে সেটিকে পুনরায় তথ্যবিশ্বে ফেরত না দিয়ে পরবর্তী এককটি নির্বাচন করা হয়, তাকে পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়ন বলে।

এক্ষেত্রে N আকারের তথ্যবিশ্ব থেকে n আকারের মোট ${}^{N}c_{n}$ টি ভিন্ন ভিন্ন নমুনা পাওয়া যায়। মূলতঃ সরল দৈব নমুনায়ন বলতে প্রতিস্থাপন না করে সরল দৈব নমুনায়নকেই বুঝায়।

উদাহরণ: একটি সমগ্রকের তিনটি উপাদান $y_1,\ y_2,\ y_3$. দুই আকার বিশিষ্ট নমুনা (i) পুনঃস্থাপন না করে, (ii) পুনঃস্থাপন করে চয়ন কর।

সমাধান: সমগ্রকের আকার, $N=3$ এবং নমুনার আকার, $n=2$

(i) পুনঃস্থাপন না করেঃ পুনঃস্থাপন না করে ভিন্ন ভিন্ন নমুনা চয়ন করা যায় ${}^{N}c_{n}={}^{3}c_{2}=3$ টি. নমুনাগুলো হলো $(y_1,\ y_2)$, $(y_1,\ y_3)$, $(y_2,\ y_3)$.

(ii) পুনঃস্থাপন করে ঃ পুনঃস্থাপন করে নমুনা চয়ন করা যায় $N^n=3^2=9$ টি। নমুনাগুলো হলো $(y_1,y_1),\ (y_1,y_2),\ (y_1,y_3),\ (y_2,y_1),\ (y_2,y_2),\ (y_2,y_3),\ (y_3,y_1),\ (y_3,y_2),\ (y_3,y_3)$ যেখানে সবগুলো নমুনা ভিন্ন নয়।

বি. দ্রঃ

(i) N আকারের সমগ্রক হতে পুনঃস্থাপন করে বা পুনঃস্থাপন না করে যে কোন উত্তোলনে যে কোন একক নির্বাচিত হওয়ার সম্ভাবনা $\frac{1}{N}$.

(ii) N আকারের সমগ্রক হতে পুনঃস্থাপন না করে দুটি নির্দিষ্ট এককের দুটি নির্দিষ্ট উত্তোলনে নির্বাচিত হওয়ার সম্ভাবনা $\frac{1}{N(N-1)}$.

(iii) পুনঃস্থাপন করে ও পুনঃস্থাপন না করে নমুনায়নে যে কোন একটি নমুনা নির্বাচিত হওয়ার সম্ভাবনা যথাক্রমে $\frac{1}{N^n}$ ও $\frac{1}{{}^N c_n}$.

প্রশ্নঃ পুনঃস্থাপন সহকারে ও পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নের মধ্যে পার্থক্য লিখ।

উত্তরঃ পুনঃস্থাপন সহকারে ও পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নের মধ্যে পার্থক্যসমূহ নিম্নরূপ-

পার্থক্যের বিষয়	পুনঃস্থাপন সহকারে	পুনঃস্থাপন ব্যতীত
(i) নির্বাচিত একক	এ পদ্ধতিতে নির্বাচিত এককগুলো পুনরায় তথ্যবিশ্বে ফেরত দেয়া হয়।	এ পদ্ধতিতে নির্বাচিত এককগুলো পুনরায় তথ্যবিশ্বে ফেরত দেয়া হয় না।
(ii) তথ্যবিশ্বের একক	এক্ষেত্রে তথ্যবিশ্বের একটি একক নমুনায় একাধিকবার আসতে পারে।	এক্ষেত্রে তথ্যবিশ্বের একটি একক নমুনায় কেবলমাত্র একবারই আসতে পারে।
(iii) নমুনার সংখ্যা	এ পদ্ধতিতে মোট নমুনার সংখ্যা হবে N^n. এখানে, $N =$ তথ্যবিশ্বের সংখ্যা, $n =$ নমুনার আকার।	এ পদ্ধতিতে মোট নমুনার সংখ্যা হবে ${}^N c_n$. এখানে, $N =$ তথ্যবিশ্বের সংখ্যা, $n =$ নমুনার আকার।
(iv) তথ্যবিশ্বের আকার	এক্ষেত্রে নমুনা একক নির্বাচনের সময় তথ্যবিশ্বের আকার সর্বদাই একই থাকে।	এক্ষেত্রে নমুনা একক নির্বাচনের সময় তথ্যবিশ্বের আকার ক্রমশঃ কমতে থাকে।
(v) পর্যায়ক্রমিক ঘটনা	এক্ষেত্রে নমুনা একক নির্বাচনের সময় পর্যায়ক্রমিক ঘটনা বা চেষ্টাগুলো পরস্পর স্বাধীন।	এক্ষেত্রে নমুনা একক নির্বাচনের সময় পর্যায়ক্রমিক ঘটনা বা চেষ্টাগুলো পরস্পর অধীন।
(vi) সম্ভাবনা	এক্ষেত্রে প্রতিটি নমুনা আসার সম্ভাবনা $= \frac{1}{N^n}$	এক্ষেত্রে প্রতিটি নমুনা আসার সম্ভাবনা $= \frac{1}{{}^N c_n}$

সরল দৈব নমুনা চয়নের বিভিন্ন পদ্ধতি (Different methods of drawing Simple Random Sample):

সরল দৈব নমুনায়ন পদ্ধতিতে প্রতি উত্তোলনে সমগ্রকের প্রতিটি এককের নির্বাচিত হওয়ার সম্ভাবনা সমান। সরল দৈব নমুনা সাধারণতঃ দু'ভাবে সংগ্রহ করা হয়:

(ক) লটারি পদ্ধতি (Lottery Method) ও

(খ) দৈব সংখ্যা সারণি পদ্ধতি (Random Number Method)

নিচে পদ্ধতি দুটি আলোচনা করা হল-

(ক) লটারি পদ্ধতি (Lottery Method): এ পদ্ধতিটি সবচেয়ে সহজ। ধরা যাক, একটি সমগ্রকে N টি একক আছে। প্রথম এককগুলোকে 1 হতে N পর্যন্ত ক্রমিক নম্বর দেয়া হয়। সমান আকার এবং একই জাতীয় N টি কাগজের টুকরায় নম্বরগুলো লেখা হয়। তারপর কাগজের টুকরাগুলোকে একইভাবে ভাজ দিয়ে ভালোভাবে মিশিয়ে একটি বাক্সে রাখা হয়। তারপর বাক্স হতে একটি একটি করে কাগজের টুকরা নিয়ে প্রয়োজনীয় আকারের একটি নমুনা নেয়া হয়।

পুন:স্থাপন না করে নমুনায়নে একটি কাগজের টুকরা উত্তোলনের পর সেটি বক্সে ফেরত না দিয়ে অবশিষ্ট টুকরাগুলো থেকে পরবর্তী কাগজের টুকরা উত্তোলন করা হয়। পুন:স্থাপন করে নমুনায়নে একটি কাগজের টুকরা উত্তোলনের পর কাগজের টুকরার নম্বর লিখে রেখে সেটি বক্সে ফেরত দিয়ে পরবর্তী কাগজের টুকরা উত্তোলন করা হয়।

সমগ্রকের আকার ছোট হলে এ পদ্ধতিতে দৈব নমুনা সংগ্রহ করা সহজ। তবে সমগ্রকের আকার বড় হলে এ পদ্ধতি মোটেই উপযোগী নয়।

(খ) দৈব সংখ্যা সারণি পদ্ধতি (Random Number Method)/ দৈব সংখ্যা সারণী ব্যবহার করে কিভাবে সরল দৈব নমুনা নির্বাচন করা হয় ? (জা.বি.-১১)

অথবা, দৈব সংখ্যা সারণী ব্যবহার করে কিভাবে সরল দৈব নমুনা নির্বাচন করা হয়, ব্যাখ্যা কর **(জা.বি.-১৫)**

উত্তর: দৈব সংখ্যা সারণিতে 0 হতে 9 পর্যন্ত অঙ্কগুলো পুনরাবৃত্তি সহকারে আয়তাকারে সাজানো থাকে। সারণিতে অঙ্কগুলো দৈব উপায়ে বসানো হয়। এখন, ধরি N আকারের সমগ্রক হতে n আকারের সরল দৈব নমুনা নেয়া হবে। এখানে $n < N$. এক্ষেত্রে নিচের ধাপগুলো অনুসরণ করা হয়-

i) প্রথমে সমগ্রকের এককগুলোকে 1 হতে N পর্যন্ত ক্রমিক নম্বর দেয়া হয়।

ii) একটি দৈব সংখ্যা সারণি নেয়া হয়।

iii) যদি N তথা সমগ্রক k অঙ্কের হয়, তাহলে k অঙ্কের N এর সর্বোচ্চ গুণিতকের মান N' (ধরি)। তবে 01 হতে N' পর্যন্ত সংখ্যাসমূহ সারি বা কলাম বা কোণাকুণিভাবে দৈব সংখ্যা সারণি হতে নেয়া যাবে।

iv) ধরা যাক, প্রথম নেয়া দৈব সংখ্যাটি r. যদি $r \le N$ হয়, তবে r ক্রমিকের এককটি নমুনায় নেয়া হবে। যদি $N < r \le N'$ হয় অর্থাৎ r যদি N এর চেয়ে বড় কিন্তু N' এর চেয়ে ছোট বা এর সমান হয়, এক্ষেত্রে r কে N দিয়ে ভাগ করে ভাগশেষের সমান ক্রমিকের এককটি নমুনায় নেয়া হয়, তবে ভাগশেষ 0 হলে N-তম ক্রমের (অর্থাৎ সর্বশেষ ক্রমের) এককটি নমুনায় নেয়া হয়। r যদি 0 হয় বা N' এর চেয়ে বড় হলে তা বাদ দিয়ে পরবর্তী দৈব সংখ্যা নেয়া হয়।

v) তথ্যবিশ্ব হতে n আকারের নমুনা না পাওয়া পর্যন্ত এ পদ্ধতি চলতে থাকে।

পুন:স্থাপন না করে নমুনায়নে কোন একক নমুনায় একবার নেয়া হলে সেটি আর নমুনায় নেয়া হয় না। তবে পুন:স্থাপন করে নমুনায়নে যে কোন একক যে কোন সংখ্যক বার নমুনায় আসতে পারে।

NB: Draw a simple random sample of size 10 from the following population:

37	25	20	23	33	50	45	40	42	35
33	43	40	45	38	37	48	39	40	42
41	38	47	52	33	36	37	43	50	46

Solution: Here, Population size, $N = 30$

Sample size, $n = 10$

At first, we identify the 30 units with the numbers from 1 to 30:

Table-1

Serial No.	Units	**Serial No.**	Units	**Serial No.**	Units
1	37	**11**	33	**21**	41
2	25	**12**	43	**22**	38
3	20	**13**	40	**23**	47
4	23	**14**	45	**24**	52
5	33	**15**	38	**25**	33
6	50	**16**	37	**26**	36
7	45	**17**	48	**27**	37
8	40	**18**	39	**28**	43
9	42	**19**	40	**29**	50
10	35	**20**	42	**30**	46

Population size, $N = 30$ which is two-digit number i.e., $k = 2$. So, for two-digit number, the largest multiple of 30 is $N' = 30 \times 3 = 90$. Now we select random numbers without replacement from 01 to 90 randomly according to column from a random number table:

Table-2

Random numbers (r)	**Serial no. of selected sample**	**Sample values** (y_i)
10	10	35
22	22	38
24	24	52
42	12	43
37	07	45
77	17	48
89	29	50
85	25	33
60	30	46
05	05	33

পরিশিষ্ট -১ দৈবসংখ্যা সারণি

Table 1 Random number table

1	2	3	4	5	6	7	8	9	10	11	12	13	14
10480	15011	01536	02011	81647	91646	69179	14194	62590	36207	20969	99570	91291	90700
22368	46573	25595	85393	30995	89198	27982	53402	93965	34095	52666	19174	39615	99505
24130	48360	22527	97265	76393	64809	15179	24830	49340	30281	30680	19655	63348	58629
42167	93060	06243	61680	07856	16376	39440	53537	71341	57004	00849	74917	97758	16379
37570	39975	81837	16656	06121	91782	60468	81305	49684	60672	14110	06927	01263	54613
77921	06907	11008	42751	27756	53498	18602	70659	90655	15053	21916	81825	44394	42880
99562	72905	56420	69994	98872	31016	71194	18738	44013	48840	63213	21069	10634	12952
96301	91977	05463	07972	18876	20922	94595	56869	69014	60045	18425	84903	42508	32307
89579	14342	63661	10281	17453	18103	57740	84378	25331	12566	58678	44947	05585	56941
85475	36857	53342	53988	53060	59533	38867	62300	08158	17983	16439	11458	18593	64952
28918	69578	88231	33276	70997	79936	56865	05859	90106	31595	01547	85590	91610	78188
63553	40961	48235	03427	49626	69445	18663	72695	52180	20847	12234	90511	33703	90322
09429	93969	52636	92737	88974	33488	36320	17617	30015	08272	84115	27156	30613	74952
10365	61129	87529	85689	48237	52267	67689	93394	01511	06358	85104	20285	39975	89868
07119	97336	71048	08178	77233	13916	47564	81056	97735	85977	29372	74461	28551	90707
51085	12765	51821	51259	77452	16308	60756	92144	49442	53900	70960	63990	75601	40719
02368	21382	52404	60268	89368	19885	55322	44819	01188	65255	64835	44919	05944	55157
01011	54092	33362	94904	31273	04146	18594	29852	71585	85030	51132	01915	92747	64951
52162	53916	46369	58586	23216	14513	83149	98736	23495	64350	94738	17752	35156	35749
07056	97628	33787	09998	42698	06691	76988	13602	51851	46104	88916	19509	25625	58104
48663	9125	85828	14346	09172	30168	90229	04734	59193	22178	30421	61666	99904	32812
54164	58492	22421	74103	47070	25306	76468	26384	58151	06646	21524	15227	96909	44592
32639	32363	05597	24200	13363	38005	94342	28728	35806	06912	17012	64161	18296	22851
29334	27001	87637	87308	58731	00256	45834	15398	46557	41135	10367	07684	36188	18510
02488	33062	28834	07351	19731	92420	60952	61280	50001	67658	32586	86679	50720	94953
81525	72295	04839	96423	24878	82651	66566	14778	76797	14780	13300	87074	19666	95725
29676	20591	68086	26432	46901	20849	89768	81536	86645	12659	92259	57102	80428	25280
00742	57392	39069	66432	84673	40027	32832	61362	98947	96067	64760	64584	96096	98253
05366	04213	25669	26422	44407	44048	37937	63904	45766	66134	75470	66520	34693	90449
91921	26418	64117	94305	26766	25940	39972	22209	71500	64568	91402	42416	07844	69618
00582	04711	87917	77341	42206	35126	74087	99547	81817	42607	43808	76655	62028	76630
							[illegible]	[illegible]	[illegible]	76038	65855	77919	88006

সরল দৈব নমুনায়নে ব্যবহৃত প্রতীক ও পরিভাষা (Notations and Terminology used in SRS):

মনে করি, কোন তথ্যবিশ্বে N সংখ্যক উপাদান বা একক আছে। এ উপাদানগুলোর কোন একটি বৈশিষ্ট্যকে চলক Y দ্বারা চিহ্নিত করি। মনে করি, $Y_1, Y_2, \cdots, Y_N$ হচ্ছে তথ্যবিশ্বের N একক থেকে প্রাপ্ত চলকের মান। এখন তথ্যবিশ্ব হতে n আকার বিশিষ্ট একটি দৈব নমুনা চয়ন করা হলো। মনে করি, $y_1, y_2, \cdots, y_n$ হচ্ছে নমুনার n সংখ্যক মান যা তথ্যবিশ্ব থেকে চয়ন করা হয়েছে। অতএব, তথ্যবিশ্ব ও নমুনার বিভিন্ন বৈশিষ্ট্য যেমনঃ গড়, ভেদাংক, সমষ্টি নিম্নরূপ-

তথ্যবিশ্ব	নমুনা
(১) মোট একক সংখ্যা $= N$	(১) মোট একক সংখ্যা $= n$
(২) তথ্যবিশ্ব সমষ্টি, $Y = Y_1 + Y_2 + \cdots + Y_N = \sum_{i=1}^{N} Y_i$	(২) নমুনা সমষ্টি, $y = y_1 + y_2 + \cdots + y_n = \sum_{i=1}^{n} y_i$
(৩) তথ্যবিশ্বের গড়, $\overline{Y} = \frac{\sum_{i=1}^{N} Y_i}{N}$	(৩) নমুনার গড়, $\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n}$
(৪) তথ্যবিশ্বের ভেদাংক, $\sigma^2 = \frac{\sum_{i=1}^{N} \left(Y_i - \overline{Y}\right)^2}{N}$	(৪) নমুনার ভেদাংক, $s'^2 = \frac{\sum_{i=1}^{n} \left(y_i - \overline{y}\right)^2}{n}$
(৫) সমগ্রকের গড় বর্গ বিচ্যুতি, $S^2 = \frac{\sum_{i=1}^{N} \left(Y_i - \overline{Y}\right)^2}{N-1}$	(৫) নমুনার গড় বর্গ বিচ্যুতি, $s^2 = \frac{\sum_{i=1}^{n} \left(y_i - \overline{y}\right)^2}{n-1}$

বি. দ্রঃ জানি, $S^2 = \dfrac{\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2}{N-1} \Rightarrow \sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2 = (N-1)\,S^2$ ....(i)

আবার, $\sigma^2 = \dfrac{\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2}{N} \Rightarrow \sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2 = N\sigma^2$ ....(ii)

এখন, (i) ও (ii) হতে লিখা যায়, $(N-1)\,S^2 = N\sigma^2$

উপপাদ্যঃ যদি কোন তথ্যবিশ্বে N সংখ্যক উপাদান থাকে, তবে সরল দৈব নমুনায়ন পদ্ধতিতে ইহার যে কোন একটি উপাদান r-তম চয়নে একটি নমুনাতে নির্বাচিত হবার সম্ভাবনা সমান এবং ইহার মান $\frac{1}{N}$ যেখানে, $r = 1, 2, \cdots, n$

[If there are N elements in a population then in simple random sampling, the probability of including any unit in a sample in r-th draw is equal and its value is $\frac{1}{N}$ where, $r = 1, 2, \cdots, n$]

প্রমাণঃ পুনঃস্থাপন পূর্বক এবং পুনঃস্থাপন ব্যতীত উভয় নমুনায়ন পদ্ধতিতেই করা যায়-

(ক) পুনঃস্থাপন পূর্বকঃ ধরি, তথ্যবিশ্বের N সংখ্যক উপাদান হতে n আকার বিশিষ্ট একটি নমুনা চয়ন করা হয়। তথ্যবিশ্বের একটি উপাদান, $r\ (r = 1, 2, \cdots, n)$-তম চয়নে একটি নমুনাতে নির্বাচিত হবার সম্ভাবনা নিম্নরূপঃ

১ম চয়নে নমুনায় নির্বাচিত হবার সম্ভাবনা $\frac{1}{N}$

২য় চয়নে নমুনায় নির্বাচিত হবার সম্ভাবনা $\frac{1}{N}$ (যেহেতু ১ম নির্বাচিত এককটি তথ্যবিশ্বে পুনঃস্থাপন করা হয়)

৩য় চয়নে নমুনায় নির্বাচিত হবার সম্ভাবনা $\frac{1}{N}$ (যেহেতু ১ম ও ২য় নির্বাচিত একক দুটি তথ্যবিশ্বে পুনঃস্থাপন করা হয়)

সুতরাং r-তম চয়নে তথ্যবিশ্বের একটি উপাদান নমুনায় নির্বাচিত হবার সম্ভাবনা $\frac{1}{N}$.

(খ) পুনঃস্থাপন ব্যতীতঃ ধরি, তথ্যবিশ্বের N সংখ্যক উপাদান হতে n আকার বিশিষ্ট একটি নমুনা চয়ন করা হয়। তথ্যবিশ্বের একটি উপাদান, $r\ (r = 1, 2, \cdots, n)$-তম চয়নে একটি নমুনাতে নির্বাচিত হবার সম্ভাবনা নিম্নরূপঃ

১ম চয়নে নমুনায় নির্বাচিত হবার সম্ভাবনা $\frac{1}{N}$

২য় চয়নে নমুনায় নির্বাচিত হবার সম্ভাবনা $\frac{1}{N-1}$ (যেহেতু ১ম নির্বাচিত এককটি তথ্যবিশ্বে পুনঃস্থাপন করা হয় না)

৩য় চয়নে নমুনায় নির্বাচিত হবার সম্ভাবনা $\frac{1}{N-2} = \frac{1}{N-(3-1)}$ (যেহেতু ১ম ও ২য় নির্বাচিত একক দুটি তথ্যবিশ্বে পুনঃস্থাপন করা হয় না)

$(r-1)$-তম চয়নে নমুনায় নির্বাচিত হবার সম্ভাবনা $\frac{1}{N-(r-1-1)}=\frac{1}{N-(r-2)}$

একইভাবে, r-তম চয়নে তথ্যবিশ্বের একটি উপাদান নমুনায় নির্বাচিত হবার সম্ভাবনা $\frac{1}{N-(r-1)}=\frac{1}{N-r+1}$

ধরি, $E_r =$ তথ্যবিশ্বের একটি বিশেষ উপাদান r-তম চয়নে নির্বাচিত হবার ঘটনা। যেহেতু চয়নগুলো স্বাধীন, তবে

$P(E_r) =$ P [তথ্যবিশ্বের বিশেষ উপাদানটি পূর্বের (r-1) তম চয়নে নির্বাচিত হয়নি কিন্তু উহা r-তম চয়নে নির্বাচিত হয়]

= P [১ম চয়নে নির্বাচিত না হওয়া] $\times$ P[২য় চয়নে নির্বাচিত না হওয়া] $\times \ldots \times$ P[(r-1)তম চয়নে নির্বাচিত না হওয়া] $\times$ P[r -তম চয়নে নির্বাচিত হওয়া]

$$\Rightarrow P(E_r)=\left(1-\frac{1}{N}\right)\left(1-\frac{1}{N-1}\right)\cdots\left\{1-\frac{1}{N-(r-2)}\right\}\left\{\frac{1}{N-r+1}\right\}$$

$$=\frac{N-1}{N}\times\frac{N-1-1}{N-1}\times\cdots\times\frac{N-r+2-1}{N-r+2}\times\frac{1}{N-r+1}$$

$$=\frac{N-1}{N}\times\frac{N-2}{N-1}\times\cdots\times\frac{N-r+1}{N-r+2}\times\frac{1}{N-r+1}$$

$=\frac{1}{N}$ (প্রমাণিত)

বিশেষ দ্রষ্টব্য:

$P(E_r) =$ P[তথ্যবিশ্বের বিশেষ উপাদানটি পূর্বের $(r-1)$ তম চয়নে নির্বাচিত হয়নি কিন্তু উহা r-তম চয়নে নির্বাচিত হয়]

যদি $r=3$ হয়, তবে তথ্যবিশ্বের বিশেষ উপাদানটি পূর্বের $(r-1)=3-1=2$ অর্থাৎ ২য় চয়ন পর্যন্ত নির্বাচিত হয়নি কিন্তু উহা 3-তম চয়নে নির্বাচিত হয়েছে।

উদাহরণ: ধরি, 4 টি সংখ্যা যথাক্রমে 5, 6, 7 এবং 8. পুনঃস্থাপন ব্যতিরেকে উক্ত 4 টি সংখ্যা হতে 3 টি সংখ্যা দৈবভাবে চয়ন করা হল। এখানে $N=4$ এবং $n=3$. পুনঃস্থাপন ব্যতিরেকে 4 টি হতে 3 টি সংখ্যা ${}^{N}c_n = {}^{4}c_3 = 4$ ভাবে চয়ন করা যায়, যেমন-

ক্রম	নমুনাসমূহ
1	5, 6, 7
2	5, 6, 8
3	5, 7, 8
4	6, 7, 8

এখন, ক্রম-1 নমুনার উপাদানসমূহ অর্থাৎ 5, 6, 7 এই প্রত্যেকটি যে কোন চয়নে আসার সম্ভাবনা বের করব-

পুনঃস্থাপন ব্যতিরেকে তথ্যবিশ্ব $(N=4)$ হতে ১ম, ২য় এবং ৩য় চয়নে নমুনার উপাদানসমূহ অর্থাৎ 5, 6 এবং 7 আসার পর্যায়ক্রমিক সম্ভাবনা যথাক্রমে $\frac{1}{4}$, $\frac{1}{3}$ এবং $\frac{1}{2}$ (যেহেতু নমুনার প্রত্যেকটি একক, তথ্যবিশ্ব হতে চয়ন করা হয় এবং এখানে তথ্যবিশ্ব $N=4$).

(ক) ক্রম-1 নমুনার এককসমূহের মধ্যে, ১ম চয়নে 5 আসার সম্ভাবনা $\frac{1}{4}=\frac{1}{N}$

(খ) ক্রম-1 নমুনার এককসমূহের মধ্যে ২য় চয়নে 6 আসার সম্ভাবনা হচ্ছে-

ধরি, $E_2 =$ তথ্যবিশ্বের 6 সংখ্যাটি 2-তম চয়নে নির্বাচিত হবার ঘটনা। যেহেতু চয়নগুলো স্বাধীন, তবে

জানি, $P(E_2)=$ P[তথ্যবিশ্বের 6 সংখ্যাটি পূর্বের $(2-1)$ তম অর্থাৎ ১ম চয়নে নির্বাচিত হয়নি কিন্তু উহা 2-তম চয়নে নির্বাচিত হয়]

= P[তথ্যবিশ্বের 6 সংখ্যাটি পূর্বের ১ম চয়নে নির্বাচিত না হওয়া] $\times$ P[২য় চয়নে নির্বাচিত হওয়া]

$$=\left(1-\frac{1}{4}\right)\times\frac{1}{3}$$

$$=\frac{3}{4}\times\frac{1}{3}=\frac{1}{4}=\frac{1}{N} \quad [\because\ N=4]$$

[বি.দ্র: ক্রম-১ এ নমুনার এককসমূহের মধ্যে ২য় চয়নে ৬ আসার সম্ভাবনা $\frac{1}{3}$, যেহেতু ১ম চয়নে ৬ আসেনি এবং ২য় চয়নটি ১ম চয়নের উপর নির্ভরশীল। কিন্তু ১ম চয়নে ৬ আসলে, তবে ৬ আসার সম্ভাবনা হতো $\frac{1}{4}$. কাজেই ১ম চয়নে ৬ না আসার সম্ভাবনা $\left(1-\frac{1}{4}=\right)\frac{3}{4}$. কারণ, জানি $P(\bar{A})=1-P(A)$.]

(গ) ক্রম-1 নমুনার এককসমূহের মধ্যে ৩য় চয়নে 7 আসার সম্ভাবনা হচ্ছে-

জানি, $P(E_3)=$ P[তথ্যবিশ্বের 7 সংখ্যাটি পূর্বের (3-1) তম অর্থাৎ ২য় চয়ন পর্যন্ত নির্বাচিত হয়নি কিন্তু উহা 3-তম চয়নে নির্বাচিত হয়]

= P[১ম চয়নে নির্বাচিত না হয়নি] $\times$ P[২য় চয়নে নির্বাচিত হয়নি] $\times$P[৩য় চয়নে নির্বাচিত হয়েছে]

$$=\left(1-\frac{1}{4}\right)\times\left(1-\frac{1}{3}\right)\times\frac{1}{2}$$

$$=\frac{3}{4}\times\frac{2}{3}\times\frac{1}{2}=\frac{1}{4}=\frac{1}{N} \quad [\because\ N=4]$$

সুতরাং পুনস্থাপন ব্যতিরেকে দৈবভাবে নমুনা-1 এর এককসমূহ 5, 6, 7 এর প্রত্যেকটি আসার ঘটনার সম্ভাবনা সমান অর্থাৎ $\frac{1}{4}\left(=\frac{1}{N}\right)$

NB: We note that the probability of selecting a specified unit on the first draw is $\frac{1}{N}$.

$$P_1=\frac{1}{N}$$

The second draw is conditional upon the first draw, since the sample is being drawn without replaecment. Suppose a specified unit is selected at second draw, which indicates that the specified unit is not selected at the first draw.

Since the draws are independent, so by compound probability theorm, the probability that the specified unit is selected at the second draw (P_2) is clearly the product of

1. the probability of the event A, that is not selected at the first draw and
2. the conditional probability of the event B, that is selected at the second draw

$$P_2 = P(A) \times P(B) = \left(1 - \frac{1}{N}\right)\frac{1}{N-1} = \frac{N-1}{N} \times \frac{1}{N-1} = \frac{1}{N} = P_1$$

Similarly, the third draw is conditional upon the two previous draws. It is the product of

(a) the probability of the event A and event B that they are not selected before and

(b) the conditonal probabilty of the event C, that is selected on the third draw

$$P_3 = P(A) \times P(B) \times P(C) = \left(1 - \frac{1}{N}\right)\left(1 - \frac{1}{N-1}\right)\frac{1}{N-2} = \frac{N-1}{N} \times \frac{N-2}{N-1} \times \frac{1}{N-2} = \frac{1}{N} = P_1$$

P-26, An introduction to sampling methods by M. Nurul Islam.

উপপাদ্য: তথ্যবিশ্বের একটি বিশেষ উপাদানের নমুনায় নির্বাচিত হবার মোট সম্ভাবনা $\frac{n}{N}$

অথবা, N আকারের তথ্যবিশ্ব হতে n আকারের সরল দৈব নমুনা চয়ন করা হলে, দেখাও যে একটি নির্দিষ্ট এককের নমুনায় অন্তর্ভূক্ত হবার সম্ভাবনা $\frac{n}{N}$ **(জা.বি.-১২)**

[If a simple random sample of size n is drawn from a population of size N, then show that the probability of including a specified unit in the sample is $\frac{n}{N}$.]

প্রমাণ: ধরি, তথ্যবিশ্বের উপাদান সংখ্যা N এবং নমুনার উপাদান সংখ্যা n. তবে n আকার বিশিষ্ট একটি নমুনাতে তথ্যবিশ্বের একটি বিশেষ উপাদান r $(r = 1, 2, \cdots, n)$-তম চয়নে নমুনাতে আসতে পারে। ধরি, E_r = তথ্যবিশ্বের একটি বিশেষ উপাদান r-তম চয়নে নির্বাচিত হবার ঘটনা। আবার, তথ্যবিশ্বের যে কোন একটি উপাদান r-তম চয়নে একটি নমুনাতে নির্বাচিত হবার সম্ভাবনা $\frac{1}{N}$ অর্থাৎ,

$$P(E_r) = \frac{1}{N}; \quad r = 1, 2, \cdots, n$$

সুতরাং, সম্ভাবনার যোগসূত্রানুসারে তথ্যবিশ্বের একটি বিশেষ উপাদান নমুনাতে নির্বাচিত হবার মোট সম্ভাবনা

= P(১ম চয়নে নির্বাচিত হওয়া) + P(২য় চয়নে নির্বাচিত হওয়া) $+\cdots+$ P(n-তম চয়নে নির্বাচিত হওয়া)

$$=\frac{1}{N}+\frac{1}{N}+\cdots+\frac{1}{N}$$
$$=\frac{n}{N}$$

সুতরাং সরল দৈব নমুনায়নে একটি নির্দিষ্ট এককের নমুনায় অন্তর্ভূক্ত হবার সম্ভাবনা $\frac{n}{N}$

উদাহরণ: ধরি, 4 টি সংখ্যা যথাক্রমে 3, 5, 7 এবং 9. পুন:স্থাপন ব্যতিরেকে উক্ত 4 টি সংখ্যা হতে 3 টি সংখ্যা দৈবভাবে চয়ন করা হল। এখানে $N=4$ এবং $n=3$. পুন:স্থাপন ব্যতিরেকে 4 টি হতে 3 টি সংখ্যা ${}^{N}c_n={}^{4}c_3=4$ ভাবে চয়ন করা যায়। সুতরাং নমুনাগুলো হচ্ছে $(3, 5, 7)$, $(3, 5, 9)$, $(3, 7, 9)$ ও $(5, 7, 9)$. এখন যে কোন একটি নমুনা বিবেচনা করি। ধরি, $(3, 5, 9)$ নমুনাটিতে 5 সংখ্যাটি ২য় চয়ন ছাড়াও ১ম বা ৩য় চয়নেও আসতে পারে। প্রত্যেকটি চয়ন, তার নমুনারূপ এবং সম্ভাবনা নিম্নরূপ:

ক্রম	নমুনাতে 5 সংখ্যাটি নির্বাচিত হয়, যদি	নমুনারূপ হবে	তথ্যবিশ্বের যে কোন একটি উপাদান r-তম চয়নে একটি নমুনাতে নির্বাচিত হবার সম্ভাবনা
1	১ম চয়নে	5, 3, 9	1/4
2	২য় চয়নে	3, 5, 9	1/4
3	৩য় চয়নে	3, 9, 5	1/4

সুতরাং, সম্ভাবনার যোগসূত্রানুসারে তথ্যবিশ্বের 5 সংখ্যাটি নমুনাতে নির্বাচিত হবার মোট সম্ভাবনা

= P(১ম চয়নে নির্বাচিত হওয়া) + P(২য় চয়নে নির্বাচিত হওয়া) + P(৩য় চয়নে নির্বাচিত হওয়া)

$$=\frac{1}{4}+\frac{1}{4}+\frac{1}{4}=\frac{3}{4}=\frac{n}{N}$$

প্রশ্ন: দেখাও যে, পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নে N-আকারের সমগ্রক হতে দুটি নির্দিষ্ট এককের n আকারের নমুনায় অন্তর্ভূক্তির সম্ভাবনা $\frac{n(n-1)}{N(N-1)}$ (জা. বি.-১৬)

প্রমাণ: ধরি, তথ্যবিশ্বের উপাদান সংখ্যা N এবং নমুনার উপাদান সংখ্যা n. তবে n আকার বিশিষ্ট একটি নমুনাতে তথ্যবিশ্বের একটি বিশেষ উপাদান r $(r=1, 2,\cdots,n)$-তম চয়নে নমুনাতে আসতে পারে। ধরি, $E_r=$ তথ্যবিশ্বের একটি বিশেষ উপাদান r-তম চয়নে নির্বাচিত হবার ঘটনা। আবার, তথ্যবিশ্বের যে কোন একটি উপাদান r-তম চয়নে একটি নমুনাতে নির্বাচিত হবার সম্ভাবনা $\frac{1}{N}$, অর্থাৎ

$$P(E_r)=\frac{1}{N};\quad r=1, 2,\cdots,n$$

সুতরাং, সম্ভাবনার যোগসূত্রানুসারে তথ্যবিশ্বের একটি নির্দিষ্ট উপাদান নমুনাতে নির্বাচিত হবার মোট সম্ভাবনা

= P(১ম চয়নে নির্বাচিত হওয়া) + P(২য় চয়নে নির্বাচিত হওয়া) + $\cdots$ + P(n-তম চয়নে নির্বাচিত হওয়া)

$$= \frac{1}{N} + \frac{1}{N} + \cdots + \frac{1}{N} = \frac{n}{N}$$

সুতরাং সরল দৈব নমুনায়নে একটি নির্দিষ্ট এককের নমুনায় অন্তর্ভূক্ত হবার সম্ভাবনা $\frac{n}{N}$.

পুনঃস্থাপন ব্যতীত বিধায় তথ্যবিশ্বের অবশিষ্ট উপাদান সংখ্যা $(N-1)$ এবং নমুনার অবশিষ্ট উপাদান সংখ্যা $(n-1)$.

অনুরূপভাবে, সরল দৈব নমুনায়নে ২য় নির্দিষ্ট এককের নমুনায় অন্তর্ভূক্ত হবার সম্ভাবনা $\frac{n-1}{N-1}$

যেহেতু সবগুলো চয়ন স্বাধীন, সেহেতু পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নে N-আকারের সমগ্রক হতে দুটি নির্দিষ্ট এককের n আকারের নমুনায় অন্তর্ভূক্তির সম্ভাবনা $\frac{n}{N}.\frac{n-1}{N-1} = \frac{n(n-1)}{N(N-1)}$.

উদাহরণঃ ধরি, 4 টি সংখ্যা যথাক্রমে 3, 5, 7 এবং 9. পুনঃস্থাপন ব্যতিরেকে উক্ত 4 টি সংখ্যা হতে 3 টি সংখ্যা দৈবভাবে চয়ন করা হল। এখানে $N=4$ এবং $n=3$. পুনঃস্থাপন ব্যতিরেকে 4 টি হতে 3 টি সংখ্যা ${}^{N}c_{n} = {}^{4}c_{3} = 4$ ভাবে চয়ন করা যায়। সুতরাং নমুনাগুলো হচ্ছে $(3, 5, 7)$, $(3, 5, 9)$, $(3, 7, 9)$ ও $(5, 7, 9)$. এখন যে কোন একটি নমুনা বিবেচনা করি। ধরি, $(3, 5, 9)$ নমুনাটিতে 5 সংখ্যাটি ২য় চয়ন ছাড়াও ১ম বা ৩য় চয়নেও আসতে পারে। প্রত্যেকটি চয়ন, তার নমুনারূপ এবং সম্ভাবনা নিম্নরূপঃ

ক্রম	নমুনাতে 5 সংখ্যাটি নির্বাচিত হয়, যদি	নমুনারূপ হবে	তথ্যবিশ্বের যে কোন একটি উপাদান r-তম চয়নে একটি নমুনাতে নির্বাচিত হবার সম্ভাবনা
1	১ম চয়নে	5, 3, 9	1/4
2	২য় চয়নে	3, 5, 9	1/4
3	৩য় চয়নে	3, 9, 5	1/4

সুতরাং, সম্ভাবনার যোগসূত্রানুসারে তথ্যবিশ্বের 5 সংখ্যাটি নমুনাতে নির্বাচিত হবার মোট সম্ভাবনা

= P(১ম চয়নে নির্বাচিত হওয়া) + P(২য় চয়নে নির্বাচিত হওয়া) + P(৩য় চয়নে নির্বাচিত হওয়া)

$$= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4} = \frac{n}{N}$$

পুনঃস্থাপন ব্যতীত বিধায় এখন 5 কে বাদ দিলে সমগ্রকের অবশিষ্ট উপাদানগুলো হবে 3, 7, 9 এবং নমুনা $(3, 5, 9)$ এর অবশিষ্ট উপাদানগুলো হবে $(3, 9)$ অর্থাৎ $N-1=3$ এবং $n-1=2$. এখন $(3, 9)$ কে নমুনা বিবেচনা করে 9 সংখ্যাটি ১ম বা ২য় চয়নে নিম্নরূপভাবে আসতে পারেঃ

ক্রম	নমুনাতে 9 সংখ্যাটি নির্বাচিত হয়, যদি	নমুনারূপ হবে	তথ্যবিশ্বের যে কোন একটি উপাদান r-তম চয়নে একটি নমুনাতে নির্বাচিত হবার সম্ভাবনা
1	১ম চয়নে	9, 3	1/3
2	২য় চয়নে	3, 9	1/3

সুতরাং, সম্ভাবনার যোগসূত্রানুসারে $3\,(=N-1)$ আকারের তথ্যবিশ্বের 9 সংখ্যাটি, $2\,(=n-1)$ আকারের নমুনাতে নির্বাচিত হবার মোট সম্ভাবনা

$$= P(\text{১ম চয়নে নির্বাচিত হওয়া}) + P(\text{২য় চয়নে নির্বাচিত হওয়া})$$

$$=\frac{1}{3}+\frac{1}{3}=\frac{2}{3}=\frac{n-1}{N-1}$$

যেহেতু সবগুলো চয়ন স্বাধীন, সেহেতু পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নে $N=4$ আকারের সমগ্রক হতে 5 ও 9 দুটি নির্দিষ্ট একক, $n=3$ আকারের নমুনায় অন্তর্ভূক্তির সম্ভাবনা $\frac{3}{4}.\frac{2}{3}=\frac{n}{N}.\frac{n-1}{N-1}=\frac{n(n-1)}{N(N-1)}$

উপপাদ্য: পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়ন পদ্ধতিতে N আকার বিশিষ্ট তথ্যবিশ্ব থেকে n আকারের $^{N}c_n$ সংখ্যক নমুনা চয়ন করা যায়। প্রমাণ কর যে, নমুনায়নে প্রতিটি নমুনার নির্বাচিত হবার সম্ভাবনা $\frac{1}{^{N}c_n}$.

প্রমাণ: পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়ন পদ্ধতিতে N আকার বিশিষ্ট তথ্যবিশ্ব থেকে n আকারের $^{N}c_n$ সংখ্যক নমুনা চয়ন করা যায়। ১ম চয়নে তথ্যবিশ্বের একটি একক নির্বাচিত হবার সম্ভাবনা $\frac{1}{N}$. ২য় চয়নে একটি নমুনা একক নির্বাচিত হবার সম্ভাবনা $\frac{1}{N-1}$, কারণ প্রথম চয়নের পর নমুনা এককটি তথ্যবিশ্বে ফেরত দেয়া হয় না। আবার, ৩য় চয়নে একটি নমুনা একক নির্বাচিত হবার সম্ভাবনা $\frac{1}{N-2}=\frac{1}{N-(3-1)}$. অনুরূপভাবে, i-তম চয়নে একটি নমুনা একক নির্বাচিত হবার সম্ভাবনা $\frac{1}{N-(i-1)};\quad i=1,2,\cdots,n$

যেহেতু সবগুলো চয়ন স্বাধীন, তবে n আকার বিশিষ্ট একটি নমুনা একটি বিশেষ পর্যায়ক্রমে নির্বাচিত হবার সম্ভাবনা

$$\frac{1}{N}.\frac{1}{N-1}.\frac{1}{N-2}\cdots\frac{1}{N-(n-1)}$$

$$=\frac{1}{N\,(N-1)\,(N-2)\cdots(N-n+1)}$$

যেহেতু নমুনার এককগুলোর $n!$ সংখ্যক বিন্যাস আছে, তাই সম্ভাবনার যোগসূত্রানুযায়ী n আকার বিশিষ্ট নমুনা পাওয়ার সম্ভাবনা

$$\frac{n!}{N\,(N-1)\,(N-2)\cdots(N-n+1)}$$

$$=\frac{n!\,(N-n)\,(N-n-1)\,(N-n-2)\cdots1}{N\,(N-1)\,(N-2)\cdots(N-n+1)\,(N-n)\,(N-n-1)\,(N-n-2)\cdots1}$$

$$=\frac{n!\;(N-n)!}{N!}$$

$$= \frac{1}{\frac{N!}{n!\,(N-n)!}} = \frac{1}{{}^{N}c_{n}}$$

উদাহরণঃ ধরি, 4 টি সংখ্যা যথাক্রমে 5, 6, 7 এবং 8. পুনঃস্থাপন ব্যতিরেকে উক্ত 4 টি সংখ্যা হতে 3 টি সংখ্যা দৈবভাবে চয়ন করা হল। এখানে $N=4$ এবং $n=3$. পুনঃস্থাপন ব্যতিরেকে 4 টি হতে 3 টি সংখ্যা ${}^{N}c_{n} = {}^{4}c_{3} = 4$ ভাবে চয়ন করা যায়, যেমন-

ক্রমিক	নমুনাসমূহ
1	5, 6, 7
2	5, 6, 8
3	5, 7, 8
4	6, 7, 8

এখন **১**ম নমুনার ক্ষেত্রে, পুনঃস্থাপন ব্যতিরেকে তথ্যবিশ্ব $(N=4)$ হতে দৈবভাবে পর্যায়ক্রমে 5 আসার সম্ভাবনা $= \frac{1}{4}$, 6 আসার সম্ভাবনা $= \frac{1}{3}$, 7 আসার সম্ভাবনা $= \frac{1}{2}$. যেহেতু সবগুলো চয়ন স্বাধীন, সুতরাং নমুনা-1 এর এককসমূহের পর্যায়ক্রমে নির্বাচিত হবার সম্ভাবনা $= \frac{1}{4}\cdot\frac{1}{3}\cdot\frac{1}{2} = \frac{1}{4.3.2}$

আবার, 1 নং ক্রমিকের নমুনা এককসমূহের বিন্যাস $(5,6,7)$, $(5,7,6)$, $(6,5,7)$, $(6,7,5)$, $(7,5,6)$, $(7,6,5)$ অর্থাৎ $3!=6$ সংখ্যক বিন্যাস আছে। আবার প্রতিটি ক্রমের সম্ভাবনা $\frac{1}{4.3.2}$. সুতরাং নমুনা-1 এর মোট সম্ভাবনা $= \frac{6}{4.3.2} = \frac{3!}{4.3.2}$

সুতরাং 3 আকার বিশিষ্ট একটি নমুনা পাওয়ার সম্ভাবনা $= \frac{3!}{4.3.2} = \frac{1}{{}^{4}c_{3}} = \frac{1}{{}^{N}c_{n}} \quad [\because N=4,\ n=3]$

উপপাদ্যঃ সরল দৈব নমুনায়নের মাধ্যমে প্রাপ্ত নমুনা গড়, তথ্যবিশ্ব গড়ের নির্ঝুঁকি (unbiased) প্রাক্কলক।

অথবা, সমগ্রক গড় $\overline{Y}$ ও নমুনা গড় $\overline{y}$ হলে প্রমাণ কর যে, $E(\overline{y}) = \overline{Y}$

অথবা, পুনঃস্থাপনবিহীন সরল দৈব নমুনায়নের ক্ষেত্রে সমগ্রকের গড় ও সমগ্রকের সমষ্টির নির্ঝুঁকি নিরূপক নির্ণয় কর (জা. বি.-**১১**)

অথবা, সচরাচর ব্যবহৃত প্রতীকসহকারে দেখাও যে, $E(\overline{y}) = \overline{Y}$ (জা. বি.-**১১**)

অথবা, পুনঃস্থাপন ছাড়া সরল দৈব নমুনায়নের ক্ষেত্রে সচরাচর সংকেতমালায় দেখাও যে, $E(\overline{y}) = \overline{Y}$ (জা. বি.-**১৫**)

প্রমাণঃ নমুনা গড়, তথ্যবিশ্ব গড়ের নির্ঝুঁকি প্রাক্কলক-

মনে করি, কোন একটি সমগ্রকের N সংখ্যক উপাদানসমূহ যথাক্রমে $Y_1, Y_2, \cdots, Y_N$ এবং সরল দৈব নমুনায়নের মাধ্যমে প্রাপ্ত n সংখ্যক নমুনা উপাদানসমূহ যথাক্রমে $y_1, y_2, \cdots, y_n$ ।

সংজ্ঞানুযায়ী, সমগ্রক গড়, $\overline{Y} = \dfrac{\sum_{i=1}^{N} Y_i}{N}$ এবং নমুনা গড়, $\overline{y} = \dfrac{\sum_{i=1}^{n} y_i}{n}$

অতএব, $E(\overline{y}) = E\left(\dfrac{\sum_{i=1}^{n} y_i}{n}\right)$

$$= \frac{1}{n} E\left(\sum_{i=1}^{n} y_i\right)$$

$$\Rightarrow E(\overline{y}) = \frac{1}{n} \sum_{i=1}^{n} E(y_i) \text{............(i)}$$

সরল দৈব নমুনায়নের শর্তানুসারে, নমুনার i-তম মান y_i , এর সম্ভাব্য মান হতে পারে সমগ্রকের $Y_1, Y_2, \cdots, Y_N$ এবং সমগ্রকের প্রতিটি উপাদান নমুনায় আসার সম্ভাবনা $\dfrac{1}{N}$ অর্থাৎ $P(Y_1) = \dfrac{1}{N},\ P(Y_2) = \dfrac{1}{N}, \cdots, P(Y_N) = \dfrac{1}{N}$.

প্রত্যাশার সংজ্ঞানুযায়ী, $E(y_i) = \sum_{i=1}^{N} Y_i\, P(Y_i) = \sum_{i=1}^{N} Y_i \times \dfrac{1}{N} = \dfrac{\sum_{i=1}^{N} Y_i}{N} = \overline{Y}$

এখন, $E(y_i)$ এর মান (i) নং সমীকরণে বসিয়ে পাই,

$$E(\overline{y}) = \frac{1}{n} \sum_{i=1}^{n} \overline{Y} = \frac{1}{n} \times n\,\overline{Y} = \overline{Y}$$

$$\therefore\ E(\overline{y}) = \overline{Y}$$

অতএব, সরল দৈব নমুনায়নের মাধ্যমে প্রাপ্ত নমুনা গড়, তথ্যবিশ্ব গড়ের নিঝুঁকি (unbiased) প্রাক্কলক (প্রমাণিত)

২য় অংশ: প্রমাণ করতে হবে যে, $\hat{Y} = N\ \overline{y}$ তথ্যবিশ্ব সমষ্টি Y এর নিঝুঁকি প্রাক্কলক।

প্রমাণ: জানি, সরল দৈব নমুনায়নে নমুনা গড় তথ্যবিশ্ব গড়ের নিঝুঁকি প্রাক্কলক অর্থাৎ $E(\overline{y}) = \overline{Y}$

এখন, $E(\hat{Y}) = E(N\,\overline{y}) = N\ E(\overline{y}) = N\,\overline{Y} = Y =$ তথ্যবিশ্বের সমষ্টি

অর্থাৎ, $\hat{Y}$ হলো তথ্যবিশ্ব সমষ্টি Y এর নিঝুঁকি প্রাক্কলক।

অনুসিদ্ধান্ত: $E(y) = E(n\ \overline{y}) = n\ E(\overline{y}) = n\ \overline{Y} \neq Y$

অর্থাৎ, নমুনা সমষ্টি কখনও তথ্যবিশ্ব সমষ্টির নিঝুঁকি প্রাক্কলক নয়।

উপপাদ্য: পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নে নমুনা গড়ের ভেদাংক হলো

$$V(\bar{y})=\frac{N-n}{N}\cdot\frac{S^2}{n}=\frac{N-n}{N-1}\cdot\frac{\sigma^2}{n}$$, যেখানে $S^2=\frac{\sum_{i=1}^{N}\left(Y_i-\bar{Y}\right)^2}{N-1}$

অথবা, সচরাচর ব্যবহৃত প্রতীকসহকারে দেখাও যে, $V(\bar{y})=\frac{N-n}{N-1}\left(\frac{\sigma^2}{n}\right)$ **(জা.বি.-১১)**

অথবা, পুন ঃ স্থাপন ছাড়া সরল দৈব নমুনায়নের ক্ষেত্রে সচরাচর সংকেতমালায় দেখাও যে, $V(\bar{y})=\frac{N-n}{N-1}\cdot\frac{\sigma^2}{n}$ (জা.বি.-১৫)

প্রমাণ: মনে করি, কোন একটি সমগ্রকের N সংখ্যক উপাদানসমূহ যথাক্রমে $Y_1, Y_2,\cdots,Y_N$ এবং সরল দৈব নমুনায়নের মাধ্যমে নির্বাচিত n সংখ্যক নমুনা উপাদানসমূহ যথাক্রমে $y_1, y_2,\cdots,y_n$.

সংজ্ঞানুযায়ী, সমগ্রক গড়, $\bar{Y}=\frac{\sum_{i=1}^{N}Y_i}{N}$ এবং নমুনা গড়, $\bar{y}=\frac{\sum_{i=1}^{n}y_i}{n}$

$$V(\bar{y})=E\left\{\bar{y}-E(\bar{y})\right\}^2$$

$$=E\left(\bar{y}-\bar{Y}\right)^2$$

$$=E\left(\frac{\sum_{i=1}^{n}y_i}{n}-\bar{Y}\right)^2$$

$$=E\left(\frac{\sum_{i=1}^{n}y_i-n\bar{Y}}{n}\right)^2$$

$$=\frac{1}{n^2}E\left(\sum_{i=1}^{n}y_i-\sum_{i=1}^{n}\bar{Y}\right)^2$$

$$=\frac{1}{n^2}E\left\{\sum_{i=1}^{n}\left(y_i-\bar{Y}\right)\right\}^2$$

$$=\frac{1}{n^2}E\left\{\sum_{i=1}^{n}\left(y_i-\bar{Y}\right)^2+\sum_{i\neq}^{n}\sum_{j=1}^{n}\left(y_i-\bar{Y}\right)\left(y_j-\bar{Y}\right)\right\}\quad\left[\because\left(\sum_{i=1}^{n}y_i\right)^2=\sum_{i=1}^{n}{y_i}^2+\sum_{i\neq}^{n}\sum_{j=1}^{n}y_iy_j\right]$$

$$\therefore\quad V(\bar{y})=\frac{1}{n^2}\left\{\sum_{i=1}^{n}E\left(y_i-\bar{Y}\right)^2+\sum_{i\neq}^{n}\sum_{j=1}^{n}E\left(y_i-\bar{Y}\right)\left(y_j-\bar{Y}\right)\right\}\text{.............(i)}$$

নমুনার i-তম মান $\left(y_i-\overline{Y}\right)^2$ এর সম্ভাব্য মানগুলো হলো সমগ্রকের $\left(Y_1-\overline{Y}\right)^2,\ \left(Y_2-\overline{Y}\right)^2,\cdots,\left(Y_N-\overline{Y}\right)^2$ এবং এই সকল মানের যে কোন একটি নমুনায় আসার সম্ভাবনা $\frac{1}{N}$. তবে

$$E\left(y_i-\overline{Y}\right)^2=\sum_{i=1}^{N}\left(Y_i-\overline{Y}\right)^2.\frac{1}{N}$$

$$=\frac{N-1}{N}.\frac{\sum_{i=1}^{N}\left(Y_i-\overline{Y}\right)^2}{N-1}=\frac{N-1}{N}S^2\ \ldots\ldots\ldots\ldots\text{(ii)}$$

একইভাবে, $\left(y_i-\overline{Y}\right)\left(y_j-\overline{Y}\right)$ এর সম্ভাব্য মানগুলো হলো সমগ্রকের $\left(Y_1-\overline{Y}\right)\left(Y_2-\overline{Y}\right),\ \left(Y_2-\overline{Y}\right)\left(Y_3-\overline{Y}\right),\cdots,\left(Y_{N-1}-\overline{Y}\right)\left(Y_N-\overline{Y}\right)$, যাদের মোট সংখ্যা $N(N-1)$ এবং এই সকল মানের যে কোন একটি নমুনায় আসার সম্ভাবনা $\frac{1}{N\left(N-1\right)}$.

$$E\left(y_i-\overline{Y}\right)\left(y_j-\overline{Y}\right)=\sum_{i\neq}^{N}\sum_{j=1}^{N}\left(Y_i-\overline{Y}\right)\left(Y_j-\overline{Y}\right).\frac{1}{N(N-1)}$$

$$=\frac{1}{N\left(N-1\right)}\left[\left\{\sum_{i=1}^{N}\left(Y_i-\overline{Y}\right)\right\}^2-\sum_{i=1}^{N}\left(Y_i-\overline{Y}\right)^2\right]\quad\left[\because\ \left(\sum_{i=1}^{n}y_i\right)^2-\sum_{i=1}^{n}y_i^2=\sum_{i\neq}^{n}\sum_{j=1}^{n}y_iy_j\right]$$

$$=\frac{1}{N\left(N-1\right)}\left[0-\sum_{i=1}^{N}\left(Y_i-\overline{Y}\right)^2\right]\quad\left[\because\ \sum_{i=1}^{N}\left(Y_i-\overline{Y}\right)=0\right]$$

$$=-\frac{\sum_{i=1}^{N}\left(Y_i-\overline{Y}\right)^2}{N\ (N-1)}$$

$$\Rightarrow E\left(y_i-\overline{Y}\right)\left(y_j-\overline{Y}\right)=-\frac{S^2}{N}\qquad\left[\because\quad S^2=\frac{\sum_{i=1}^{N}\left(Y_i-\overline{Y}\right)^2}{N-1}\right]\ldots\ldots\ldots\ldots\text{(iii)}$$

এখন, (ii) নং ও (iii) নং হতে মান (i) নং এ বসিয়ে পাই,

$$V\left(\overline{y}\right)=\frac{1}{n^2}\left\{\sum_{i=1}^{n}\frac{N-1}{N}S^2+\sum_{i\neq}^{n}\sum_{j=1}^{n}\left(-\frac{S^2}{N}\right)\right\}$$

$$=\frac{1}{n^2}\left\{\frac{n\left(N-1\right)}{N}S^2+n\left(n-1\right)\left(-\frac{S^2}{N}\right)\right\}$$

$$=\frac{n\left(N-1\right)S^2}{n^2\,N}-\frac{n\left(n-1\right)S^2}{n^2\,N}$$

$$=\frac{(N-1)S^2}{nN}-\frac{(n-1)S^2}{nN}$$

$$=\frac{S^2}{nN}(N-1-n+1)$$

$$=\frac{(N-n)\,S^2}{nN}$$

$$\therefore\quad V(\bar{y})=\frac{N-n}{N}.\frac{S^2}{n}$$

$$=\left(1-\frac{n}{N}\right)\frac{S^2}{n}=(1-f)\frac{S^2}{n};\quad \text{যেখানে } f=\frac{n}{N}$$

আবার, $V(\bar{y})=\frac{N-n}{N}.\frac{S^2}{n}=\frac{N-n}{N(N-1)}.\frac{(N-1)S^2}{n}$

$$=\frac{N-n}{N(N-1)}.\frac{N\sigma^2}{n}\qquad\left[\because (N-1)S^2=N\sigma^2\right]$$

$$=\frac{N-n}{N-1}.\frac{\sigma^2}{n}$$

বিশেষ দ্রষ্টব্য:

(ক) $\sum_{i\neq}^{3}\sum_{j=1}^{3}(Y_i-\bar{Y})(Y_j-\bar{Y})$

$$=(Y_1-\bar{Y})(Y_2-\bar{Y})+(Y_1-\bar{Y})(Y_3-\bar{Y})+(Y_2-\bar{Y})(Y_3-\bar{Y})+(Y_2-\bar{Y})(Y_1-\bar{Y})+(Y_3-\bar{Y})(Y_1-\bar{Y})+(Y_3-\bar{Y})(Y_2-\bar{Y})$$

= Summation of $6=3\,(3-1)$ terms.

Similarly, $\sum_{i\neq}^{n}\sum_{j=1}^{n}(Y_i-\bar{Y})(Y_j-\bar{Y})=$ Summation of $n(n-1)$ terms.

(খ) $\frac{n}{N}=f$ কে নমুনায়ন ভগ্নাংশ (Sampling fraction) এবং $\frac{N-n}{N}=1-\frac{n}{N}=1-f$ কে সসীম তথ্যবিশ্ব শুদ্ধি (Finite Population Correction-FPC) বলা হয়। এখন $\frac{n}{N}=f$ ব্যবহার করে পাই,

$$V(\bar{y})=\frac{N-n}{N}.\frac{S^2}{n}=\left(1-\frac{n}{N}\right)\frac{S^2}{n}=(1-f)\frac{S^2}{n}$$

(গ) নমুনা গড় $\bar{y}$ এর আদর্শ বিচ্যুতি (Standard Error-SE)

$$SE(\bar{y})=\sqrt{\frac{N-n}{N}.\frac{S^2}{n}}$$

(ঘ) **নমুনা থেকে $V(\bar{y})$ এবং $SE(\bar{y})$ নির্ণয়:** যেহেতু S^2 এর মান জানা থাকে না, তাই S^2 এর পরিবর্তে এর নিঝুকি প্রাক্কলক s^2 এর মান নমুনা থেকে নির্ণয় করে ভেদাংক এবং আদর্শ বিচ্যুতির প্রাক্কলিত মান নিম্নরূপে পাওয়া যায়-

$$\text{Est. } V(\bar{y})=\frac{N-n}{N}.\frac{s^2}{n}$$

$$\text{Est. } SE(\bar{y})=\sqrt{\frac{N-n}{N}.\frac{s^2}{n}}$$, যেখানে $s^2=\dfrac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}$

(ঙ) $V(x)=E\{x-E(x)\}^2$

(চ) সমগ্রকের গড় বর্গ বিচ্যুতি, $S^2=\dfrac{\sum_{i=1}^{N}\left(Y_i-\bar{Y}\right)^2}{N-1}$

$$\Rightarrow \sum_{i=1}^{N}\left(Y_i-\bar{Y}\right)^2=(N-1)S^2$$

প্রশ্ন ঃ দেখাও যে, তথ্যবিশ্ব সমষ্টির প্রাক্কলকের ভেদাংক $V(\hat{Y})=N^2(1-f)\dfrac{S^2}{n}$

প্রমাণ: জানি, $V(\bar{y})=\dfrac{N-n}{N}.\dfrac{S^2}{n}$

তথ্যবিশ্বের সমষ্টি Y এর প্রাক্কলক $\hat{Y}=N\,\bar{y}$ এর ভেদাংক,

$$V(\hat{Y})=V(N\bar{y})=N^2\,V(\bar{y})=N^2.\frac{N-n}{N}.\frac{S^2}{n}=N^2(1-f)\frac{S^2}{n}$$

$$SE(\hat{Y})=\sqrt{N^2(1-f)\frac{S^2}{n}}$$

উপপাদ্য: পুন:স্থাপন সহকারে সরল দৈব নমুনার ক্ষেত্রে নমুনা গড়ের ভেদাংক $V(\bar{y})=\dfrac{\sigma^2}{n}=\dfrac{N-1}{N}.\dfrac{S^2}{n}$

প্রমাণ: মনে করি, কোন একটি সমগ্রকের N সংখ্যক উপাদানসমূহ যথাক্রমে $Y_1, Y_2, \cdots, Y_N$ এবং সরল দৈব নমুনায়নের মাধ্যমে নির্বাচিত n সংখ্যক নমুনা উপাদানসমূহ যথাক্রমে $y_1, y_2, \cdots, y_n$.

সংজ্ঞানুযায়ী, সমগ্রক গড়, $\bar{Y}=\dfrac{\sum_{i=1}^{N}Y_i}{N}$ এবং নমুনা গড়, $\bar{y}=\dfrac{\sum_{i=1}^{n}y_i}{n}$

এখন, $V(\bar{y}) = V\left(\frac{\sum_{i=1}^{n} y_i}{n}\right)$

$$= \frac{1}{n^2} V\left(\sum_{i=1}^{n} y_i\right)$$

$$= \frac{1}{n^2} V(y_1 + y_2 + \cdots + y_n)$$

$$= \frac{1}{n^2}\left[V(y_1) + V(y_2) + \cdots + V(y_n) + 2Cov(y_1, y_2) + 2Cov(y_2, y_3) + \cdots + 2Cov(y_{n-1}, y_n)\right]$$

পুনঃস্থাপন করে সরল দৈব নমুনায়নে $y_1, y_2, \cdots, y_n$ পরস্পর স্বাধীন বলে, $Cov(y_i, y_j) = 0$

$$\therefore \quad V(\bar{y}) = \frac{1}{n^2}\left[V(y_1) + V(y_2) + \cdots + V(y_n) + 0 + 0 + \cdots + 0\right]$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} V(y_i)$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} E\{y_i - E(y_i)\}^2$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} E(y_i - \bar{Y})^2 \quad \left[\because E(y_i) = \sum_{i=1}^{N} Y_i\, P(Y_i) = \sum_{i=1}^{N} Y_i \cdot \frac{1}{N} = \frac{1}{N}\sum_{i=1}^{N} Y_i = \bar{Y}\right]$$

$$= \frac{1}{n^2} \sum_{i=1}^{n}\left[\sum_{j=1}^{N} (Y_j - \bar{Y})^2 \cdot \frac{1}{N}\right]$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2 \quad \left[\because \sigma^2 = \frac{\sum_{j=1}^{N} (Y_j - \bar{Y})^2}{N}\right]$$

$$= \frac{1}{n^2} \times n\, \sigma^2$$

$$\therefore \ V(\bar{y}) = \frac{\sigma^2}{n}$$

আবার, $V(\bar{y}) = \frac{(N-1)S^2}{Nn} \quad \left[\because \ (N-1)S^2 = N\sigma^2 \ \Rightarrow \sigma^2 = \frac{(N-1)S^2}{N}\right]$

$$= \frac{N-1}{N} \cdot \frac{S^2}{n}$$

বি. দ্রঃ $V(x+y) = V(x) + V(y) + 2Cov(x, y)$

প্রশ্ন: পুনঃস্থাপন ব্যতীত এবং পুনঃস্থাপন পূর্বক নমুনায়নের তুলনা কর।

অথবা, প্রমাণ কর যে, $V(\bar{y})_{wor} \le V(\bar{y})_{wr}$.

উত্তর: জানি, পুনঃস্থাপন ব্যতীত দৈব নমুনার ক্ষেত্রে, $V(\bar{y})_{WOR} = \frac{N-n}{N}.\frac{S^2}{n}$

পুনঃস্থাপন পূর্বক দৈব নমুনার ক্ষেত্রে, $V(\bar{y})_{WR} = \frac{N-1}{N}.\frac{S^2}{n}$

এখন, $V(\bar{y})_{WR} - V(\bar{y})_{WOR} = \frac{N-1}{N}.\frac{S^2}{n} - \frac{N-n}{N}.\frac{S^2}{n}$

$$= \frac{S^2}{Nn}(N-1-N+n)$$

$$= \frac{S^2}{Nn}(n-1)$$

$$\therefore\ V(\bar{y})_{WR} - V(\bar{y})_{WOR} \ge 0$$

$$\Rightarrow V(\bar{y})_{WR} \ge V(\bar{y})_{WOR}$$

$$\Rightarrow V(\bar{y})_{WOR} \le V(\bar{y})_{WR}$$

অর্থাৎ তথ্যবিশ্ব গড়ের জন্য পুনঃস্থাপন ব্যতীত নমুনা গড়, পুনঃস্থাপন পূর্বক নমুনা গড়ের চেয়ে অধিক দক্ষ প্রাক্কলক।

প্রশ্ন: n আকারবিশিষ্ট একটি সরল দৈব নমুনার জন্য দেখাও যে, প্রতিস্থাপন ব্যতিরেকে নমুনা গড়ের পরিমিত বিচ্যুতি প্রতিস্থাপন সহকারে নমুনা গড়ের পরিমিত বিচ্যুতি অপেক্ষা ছোট (জা.বি.-১১)

(For a simple random sampling of size n, show that the standard error of sample mean in sampling without replacement is smaller than that in sampling wtih replacement)

উত্তর: জানি, পুনঃস্থাপন ব্যতীত দৈব নমুনার ক্ষেত্রে,

$$V(\bar{y})_{WOR} = \frac{N-n}{N}.\frac{S^2}{n}$$

$$\Rightarrow SE(\bar{y})_{WOR} = \sqrt{\frac{N-n}{N}.\frac{S^2}{n}}$$

পুনঃস্থাপন পূর্বক দৈব নমুনার ক্ষেত্রে, $V(\bar{y})_{WR} = \frac{N-1}{N}.\frac{S^2}{n}$

$$\Rightarrow SE(\bar{y})_{WR} = \sqrt{\frac{N-1}{N}.\frac{S^2}{n}}$$

এখন, $SE(\bar{y})_{WR} - SE(\bar{y})_{WOR} = \sqrt{\frac{N-1}{N}\cdot\frac{S^2}{n}} - \sqrt{\frac{N-n}{N}\cdot\frac{S^2}{n}}$

$$= \frac{S}{\sqrt{Nn}}\left(\sqrt{N-1} - \sqrt{N-n}\right)$$

$\therefore\ SE(\bar{y})_{WR} - SE(\bar{y})_{WOR} > 0$ (যখন $n > 1$)

$\Rightarrow SE(\bar{y})_{WR} > SE(\bar{y})_{WOR}$

$\Rightarrow SE(\bar{y})_{WOR} < SE(\bar{y})_{WR}$

অর্থাৎ n আকারবিশিষ্ট একটি সরল দৈব নমুনার জন্য, প্রতিস্থাপন ব্যতিরেকে নমুনা গড়ের পরিমিত বিচ্যুতি প্রতিস্থাপন সহকারে নমুনা গড়ের পরিমিত বিচ্যুতি অপেক্ষা ছোট ।

উপপাদ্য: পুনঃস্থাপনবিহীন সরল দৈব নমুনায়নের ক্ষেত্রে প্রমাণ কর যে, $E(s^2) = S^2$ (প্রতীকগুলো চিরাচরিত) **জা.বি.-১২**

অথবা, পুনঃস্থাপন ব্যতীত সরল দৈব নমুনার গড় বর্গ বিচ্যুতি (s^2), তথ্যবিশ্বের গড় বর্গ বিচ্যুতি (S^2) এর নিঝুঁকি প্রাক্কলক অর্থাৎ $E(s^2) = S^2$

অথবা, দেখাও যে সরল দৈব নমুনায়নে পুনঃস্থাপন না করে পদ্ধতিতে $E(s^2) = S^2$ কিন্তু $E(S^2) \neq \sigma^2$, যেখানে সংকেতগুলি প্রচলিত (জা.বি.-১৬)

প্রমাণ: জানি, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ এবং $S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$

এখন, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \bar{Y} + \bar{Y} - \bar{y}\right)^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \bar{Y} - \bar{y} + \bar{Y}\right)^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left[(y_i - \bar{Y}) - (\bar{y} - \bar{Y})\right]^2$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}(y_i - \bar{Y})^2 - 2\sum_{i=1}^{n}(y_i - \bar{Y})(\bar{y} - \bar{Y}) + \sum_{i=1}^{n}(\bar{y} - \bar{Y})^2\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}(y_i - \bar{Y})^2 - 2(\bar{y} - \bar{Y})\sum_{i=1}^{n}(y_i - \bar{Y}) + n(\bar{y} - \bar{Y})^2\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}(y_i - \bar{Y})^2 - 2(\bar{y} - \bar{Y})(n\bar{y} - n\bar{Y}) + n(\bar{y} - \bar{Y})^2\right]$$

$$=\frac{1}{n-1}\left[\sum_{i=1}^{n}\left(y_i-\overline{Y}\right)^2-2n\left(\overline{y}-\overline{Y}\right)^2+n\left(\overline{y}-\overline{Y}\right)^2\right]$$

$$=\frac{1}{n-1}\left[\sum_{i=1}^{n}\left(y_i-\overline{Y}\right)^2-n\left(\overline{y}-\overline{Y}\right)^2\right]$$

$$\therefore\quad E\left(s^2\right)=\frac{1}{n-1}\left[\sum_{i=1}^{n}E\left(y_i-\overline{Y}\right)^2-n\,E\left(\overline{y}-\overline{Y}\right)^2\right]\text{..........(i)}$$

নমুনার i-তম মান $\left(y_i-\overline{Y}\right)^2$ এর সম্ভাব্য মানগুলো হলো সমগ্রকের $\left(Y_1-\overline{Y}\right)^2,\left(Y_2-\overline{Y}\right)^2,\cdots,\left(Y_N-\overline{Y}\right)^2$ এবং এই সকল মানের যে কোন একটি নমুনায় আসার সম্ভাবনা $\frac{1}{N}$.

$$E\left(y_i-\overline{Y}\right)^2=\sum_{i=1}^{N}\left(Y_i-\overline{Y}\right)^2.\frac{1}{N}$$

$$=\sigma^2\text{...........(ii)}$$

এবং $E\left(\overline{y}-\overline{Y}\right)^2=E\left[\overline{y}-E\left(\overline{y}\right)\right]^2\qquad\left[\because\ E\left(\overline{y}\right)=\overline{Y}\right]$

$$=V\left(\overline{y}\right)$$

$$=\frac{N-n}{N}.\frac{S^2}{n}\text{.............(iii)}$$

এখন, (ii) নং ও (iii) নং হতে মান (i) নং এ বসিয়ে পাই,

$$E\left(s^2\right)=\frac{1}{n-1}\left[\sum_{i=1}^{n}\sigma^2-n.\frac{N-n}{N}.\frac{S^2}{n}\right]$$

$$=\frac{1}{n-1}\left[n\,\sigma^2-\frac{(N-n)S^2}{N}\right]$$

$$=\frac{1}{n-1}\left[n.\frac{(N-1)S^2}{N}-\frac{(N-n)S^2}{N}\right]\qquad\left[\because (N-1)S^2=N\sigma^2\Rightarrow\sigma^2=\frac{(N-1)S^2}{N}\right]$$

$$=\frac{1}{n-1}.\frac{S^2}{N}\left[n(N-1)-(N-n)\right]$$

$$=\frac{1}{n-1}.\frac{S^2}{N}\left[nN-n-N+n\right]$$

$$=\frac{1}{n-1}.\frac{S^2}{N}.N\left(n-1\right)$$

$$=S^2$$

$\therefore\quad E\left(s^2\right)=S^2$ (প্রমাণিত)

উপপাদ্যঃ পুনঃস্থাপন সহকারে সরল দৈব নমুনার গড় বর্গ বিচ্যুতি (s^2), তথ্যবিশ্বের ভেদাংকের (σ^2) নিঝুঁকি প্রাক্কলক অর্থাৎ $E(s^2)=\sigma^2$

প্রমাণঃ জানি, $s^2=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\overline{y})^2$ এবং $S^2=\frac{1}{N-1}\sum_{i=1}^{N}(Y_i-\overline{Y})^2$

এখন,
$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\overline{y})^2$$
$$=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\overline{Y}+\overline{Y}-\overline{y})^2$$
$$=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\overline{Y}-\overline{y}+\overline{Y})^2$$
$$=\frac{1}{n-1}\sum_{i=1}^{n}\left[(y_i-\overline{Y})-(\overline{y}-\overline{Y})\right]^2$$
$$=\frac{1}{n-1}\left[\sum_{i=1}^{n}(y_i-\overline{Y})^2-2\sum_{i=1}^{n}(y_i-\overline{Y})(\overline{y}-\overline{Y})+\sum_{i=1}^{n}(\overline{y}-\overline{Y})^2\right]$$
$$=\frac{1}{n-1}\left[\sum_{i=1}^{n}(y_i-\overline{Y})^2-2(\overline{y}-\overline{Y})\sum_{i=1}^{n}(y_i-\overline{Y})+n(\overline{y}-\overline{Y})^2\right]$$
$$=\frac{1}{n-1}\left[\sum_{i=1}^{n}(y_i-\overline{Y})^2-2(\overline{y}-\overline{Y})(n\overline{y}-n\overline{Y})+n(\overline{y}-\overline{Y})^2\right]$$
$$=\frac{1}{n-1}\left[\sum_{i=1}^{n}(y_i-\overline{Y})^2-2n(\overline{y}-\overline{Y})^2+n(\overline{y}-\overline{Y})^2\right]$$
$$=\frac{1}{n-1}\left[\sum_{i=1}^{n}(y_i-\overline{Y})^2-n(\overline{y}-\overline{Y})^2\right]$$
$$\therefore\ E(s^2)=\frac{1}{n-1}\left[\sum_{i=1}^{n}E(y_i-\overline{Y})^2-n\,E(\overline{y}-\overline{Y})^2\right]\text{...........(i)}$$

নমুনার i-তম মান $(y_i-\overline{Y})^2$ এর সম্ভাব্য মানগুলো হলো সমগ্রকের $(Y_1-\overline{Y})^2,(Y_2-\overline{Y})^2,\cdots,(Y_N-\overline{Y})^2$ এবং এই সকল মানের যে কোন একটি নমুনায় আসার সম্ভাবনা $\frac{1}{N}$.

$$E(y_i-\overline{Y})^2=\sum_{i=1}^{N}(Y_i-\overline{Y})^2.\frac{1}{N}$$
$$=\sigma^2\text{............(ii)}$$

এবং $E(\overline{y}-\overline{Y})^2=E\left[\overline{y}-E(\overline{y})\right]^2$ $\quad\left[\because\ E(\overline{y})=\overline{Y}\right]$

$$= V(\bar{y})$$

$$= \frac{\sigma^2}{n} \text{.............(iii)}$$

এখন, (ii) নং ও (iii) নং হতে মান (i) নং এ বসিয়ে পাই,

$$E(s^2) = \frac{1}{n-1}\left[\sum_{i=1}^{n}\sigma^2 - n.\frac{\sigma^2}{n}\right]$$

$$= \frac{1}{n-1}\left[n\,\sigma^2 - \sigma^2\right]$$

$$= \frac{1}{n-1}\times \sigma^2(n-1)$$

$$= \sigma^2$$

$\therefore \quad E(s^2) = \sigma^2$ (প্রমাণিত)

আস্থা সীমা বা নিশ্চয়তার সীমা (Confidence limits):

$\bar{y}$ ও $\hat{Y}$ হলো যথাক্রমে তথ্যবিশ্ব গড় $\bar{Y}$ ও তথ্যবিশ্ব সমষ্টি Y এর প্রাক্কলক। এছাড়া s^2 হচ্ছে সমগ্রকের গড় বর্গ বিচ্যুতি S^2 এর প্রাক্কলক। সাধারণতঃ অনুমান করা হয় যে, $\bar{y}$ ও $\hat{Y}$ এর বিন্যাস পরিমিত বিন্যাস, যাদের গড় যথাক্রমে $\bar{Y}$ ও $N\bar{Y}$ এবং ভেদাংকের প্রাক্কলক যথাক্রমে $(1-f)\frac{s^2}{n}$ ও $N^2(1-f)\frac{s^2}{n}$ ।

(ক) ছোট আকারের নমুনার (নমুনার আকার 30 এর কম) জন্য-

সমগ্রকের গড়ের আস্থা ব্যাপ্তি (Confidence interval):

$$\left(\bar{y} - t_{\alpha/2,n-1}\ SE(\bar{y}),\quad \bar{y} + t_{\alpha/2,n-1}\ SE(\bar{y})\right)$$

$$\Rightarrow \left(\bar{y} - t_{\alpha/2,n-1}\sqrt{\frac{1-f}{n}s^2},\quad \bar{y} + t_{\alpha/2,n-1}\sqrt{\frac{1-f}{n}s^2}\right)$$

একইভাবে, সমগ্রকের সমষ্টির আস্থা ব্যাপ্তি (Confidence interval) :

$$\left(\hat{Y} - t_{\alpha/2,n-1}.\ SE(\hat{Y}),\quad \hat{Y} + t_{\alpha/2,n-1}.\ SE(\hat{Y})\right)$$

$$\Rightarrow \left(N\bar{y} - t_{\alpha/2,\,n-1}.\ N\sqrt{\frac{1-f}{n}s^2},\quad N\bar{y} + t_{\alpha/2,\,n-1}.\ N\sqrt{\frac{1-f}{n}s^2}\right)$$

এখানে, $t_{\alpha/2,\,n-1}$ হলো $(n-1)$ স্বাধীনতার মাত্রা ও α সংশয় মাত্রায় (দ্বিপ্রান্তিক যাচাইয়ে) t বিন্যাসের সারণির মান।

(খ) বড় আকারের নমুনার (নমুনার আকার 30 বা তার বেশি) জন্য-

সমগ্রকের গড়ের আস্থা ব্যাপ্তি (Confidence interval) :

$$\left(\bar{y} - z_{\alpha/2}\ \sqrt{\frac{1-f}{n}s^2},\quad \bar{y} + z_{\alpha/2}\ \sqrt{\frac{1-f}{n}s^2}\right)$$

একইভাবে, সমগ্রকের সমষ্টির আস্থা ব্যাপ্তি (Confidence interval):

$$\left(N\bar{y} - z_{\alpha/2}.\ N\sqrt{\frac{1-f}{n}s^2},\quad N\bar{y} + z_{\alpha/2}.\ N\sqrt{\frac{1-f}{n}s^2}\right)$$

এখানে, $z_{\alpha/2}$ হলো α সংশয় মাত্রায় (দ্বিপ্রান্তিক যাচাইয়ে) আদর্শ পরিমিত বিন্যাসের সারণির মান।

বিশেষ দ্রষ্টব্য:

(ক) পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নে নমুনা গড়ের ভেদাংক

$$V(\bar{y}) = \frac{N-n}{N}.\frac{S^2}{n} = \left(1-\frac{n}{N}\right)\frac{S^2}{n} = (1-f)\frac{S^2}{n} = \frac{1-f}{n}S^2$$

এখানে, $\frac{n}{N} = f$ হচ্ছে নমুনায়ন ভগ্নাংশ (Sampling fraction)

$$\therefore\quad SE(\bar{y}) = \sqrt{\frac{1-f}{n}S^2}$$

(খ) আবার, জানি $V(\bar{y}) = \frac{N-n}{N}.\frac{S^2}{n}$

তথ্যবিশ্ব সমষ্টি Y এর প্রাক্কলক $\hat{Y} = N\bar{y}$ এর ভেদাংক,

$$V(\hat{Y}) = V(N\bar{y}) = N^2\ V(\bar{y}) = N^2.\frac{N-n}{N}.\frac{S^2}{n} = N^2(1-f)\frac{S^2}{n} = \frac{N^2(1-f)}{n}S^2$$

$$\therefore\quad SE(\hat{Y}) = \sqrt{\frac{N^2(1-f)}{n}S^2} = N\ \sqrt{\frac{1-f}{n}S^2}$$

অনুপাতের জন্য সরল দৈব নমুনায়ন (Simple Random Sampling for Proportion):

মনে করি কোন সমগ্রকে N টি একক আছে, যাদের কোন গুণগত বৈশিষ্ট্য অনুযায়ী দুটি শ্রেণীতে ভাগ করা হয়েছে। ধরি সমগ্রকে N টি এককের মধ্যে উল্লেখিত গুণের অধিকারী এককের সংখ্যা A। সুতরাং উল্লেখিত গুণের অধিকারী নয়, এমন এককের সংখ্যা $N-A$.

ধরি উক্ত সমগ্রক হতে n টি একক সরল দৈব নমুনার মাধ্যমে সংগ্রহ করা হয়েছে, যাদের মধ্যে উল্লেখিত গুণের অধিকারী এককের সংখ্যা a এবং $n-a$ টি একক উল্লেখিত গুণের অধিকারী নয়।

সমগ্রকে উল্লেখিত বৈশিষ্ট্যের অধিকারী এককগুলোর অনুপাত, $P=\frac{A}{N}$

নমুনায় উল্লেখিত বৈশিষ্ট্যের অধিকারী এককগুলোর অনুপাত, $p=\frac{a}{n}$

সমগ্রকে উল্লেখিত বৈশিষ্ট্যের অধিকারী নয় এমন এককগুলোর অনুপাত, $Q=\frac{N-A}{N}=1-\frac{A}{N}=1-P$

নমুনায় উল্লেখিত বৈশিষ্ট্যের অধিকারী নয় এমন এককগুলোর অনুপাত, $q=\frac{n-a}{n}=1-\frac{a}{n}=1-p$

ধরি, Y একটি গুণবাচক চলক যেখানে

$$Y_i=\begin{cases}1; & \text{যদি একককটি উল্লেখিত গুণের অধিকারী হয়।}\\ 0; & \text{যদি একককটি উল্লেখিত গুণের অধিকারী না হয়।}\end{cases}$$

সুতরাং সমগ্রক গড়, $\overline{Y}=\frac{\sum_{i=1}^{N}Y_i}{N}=\frac{A}{N}=P$ এবং নমুনার গড়, $\overline{y}=\frac{\sum_{i=1}^{n}y_i}{n}=\frac{a}{n}=p$

উপপাদ্যঃ সরল দৈব নমুনায়নের ক্ষেত্রে দেখাও যে, নমুনার অনুপাত (p) তথ্যবিশ্ব অনুপাত (P) এর নিঝুঁকি প্রাক্কলক এবং $V(p)=\frac{N-n}{N-1}.\frac{PQ}{n}$.

প্রমাণঃ জানি, সরল দৈব নমুনায়নে নমুনা গড় $\overline{y}$ হলো তথ্যবিশ্ব গড় $\overline{Y}$ এর নিঝুকি প্রাক্কলক অর্থাৎ $E(\overline{y})=\overline{Y}$.

ধরি, Y একটি গুণবাচক চলক যেখানে

$$Y_i=\begin{cases}1; & \text{যদি একককটি উল্লেখিত গুণের অধিকারী হয়।}\\ 0; & \text{যদি একককটি উল্লেখিত গুণের অধিকারী না হয়।}\end{cases}$$

ধরি তথ্যবিশ্বে N টি এককের মধ্যে উল্লেখিত গুণের অধিকারী এককের সংখ্যা A. ধরি উক্ত তথ্যবিশ্ব হতে n টি একক সরল দৈব নমুনার মাধ্যমে সংগ্রহ করা হয়েছে, যাদের মধ্যে উল্লেখিত গুণের অধিকারী এককের সংখ্যা a.

তথ্যবিশ্বে উল্লেখিত বৈশিষ্ট্যের অধিকারী এককগুলোর অনুপাত, $P=\frac{A}{N}$

নমুনায় উল্লেখিত বৈশিষ্ট্যের অধিকারী এককগুলোর অনুপাত, $p=\frac{a}{n}$

এখন তথ্যবিশ্ব গড়, $\overline{Y}=\frac{\sum_{i=1}^{N}Y_i}{N}=\frac{A}{N}=P$ এবং নমুনার গড়, $\overline{y}=\frac{\sum_{i=1}^{n}y_i}{n}=\frac{a}{n}=p$

সুতরাং, $E(p)=P \quad \left[\because E(\overline{y})=\overline{Y}\right]$

অর্থাৎ নমুনার অনুপাত (p) তথ্যবিশ্ব অনুপাত (P) এর নিঝুঁকি প্রাক্কলক।

জানি, সরল দৈব নমুনায়নে নমুনা গড়ের ভেদাংক হলো

$$V(\bar{y}) = \frac{N-n}{N}.\frac{S^2}{n}\text{, যেখানে } S^2 = \frac{\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^2}{N-1}$$

$$\Rightarrow V(p) = \frac{N-n}{Nn}S^2 \quad [\because\ \bar{y} = p]...........(i)$$

এখন, $S^2 = \frac{\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^2}{N-1}$

$$= \frac{1}{N-1}\left[\sum_{i=1}^{N} Y_i^2 - N\bar{Y}^2\right]$$

$$= \frac{1}{N-1}\left[NP - NP^2\right] \quad [\because\ \bar{Y} = P]$$

$$= \frac{NP(1-P)}{N-1}$$

$$\Rightarrow S^2 = \frac{NPQ}{N-1} \quad [\because\ P + Q = 1]$$

এখন (i) নং কে লিখা যায়,

$$V(p) = \frac{N-n}{Nn}S^2 = \frac{N-n}{Nn}.\frac{NPQ}{N-1} = \frac{N-n}{N-1}.\frac{PQ}{n} \text{ (প্রমাণিত)}$$

NB: যেহেতু, $Y_i = 0$ or 1 তবে $Y_i^2 = 0$ or 1.

সুতরাং $Y_i = Y_i^2$

$$\Rightarrow \sum_{i=1}^{N} Y_i = \sum_{i=1}^{N} Y_i^2$$

জানি, $\bar{Y} = \frac{\sum_{i=1}^{N} Y_i}{N} = P$

$$\Rightarrow \frac{\sum_{i=1}^{N} Y_i^2}{N} = P \quad \left[\because\ \sum_{i=1}^{N} Y_i = \sum_{i=1}^{N} Y_i^2\right]$$

$$\Rightarrow \sum_{i=1}^{N} Y_i^2 = NP$$

সমগ্রকে উল্লেখিত বৈশিষ্ট্যের অধিকারী এককগুলোর অনুপাত, $P = \frac{A}{N}$

সমগ্রকে উল্লেখিত বৈশিষ্ট্যের অধিকারী নয় এমন এককগুলোর অনুপাত, $Q = \frac{N-A}{N}$

$$\therefore\ P+Q=\frac{A}{N}+\frac{N-A}{N}=\frac{A+N-A}{N}=1$$

প্রশ্ন: সরল দৈব নমুনায়নের সুবিধা ও অসুবিধাসমূহ আলোচনা কর (জা.বি.-১১)

উত্তর: সরল দৈব নমুনায়নের সুবিধা ও অসুবিধাসমূহ নিম্নে আলোচনা করা হল-

সুবিধা (Advantages):

(ক) সরল দৈব নমুনায়নে সমগ্রকের প্রত্যেকটি এককের নমুনায় অন্তর্ভুক্ত হওয়ার সম্ভাবনা সমান। তাই এ নমুনায়নে ব্যক্তিগত পছন্দের কোন ঝুঁকি নেই। ফলে নমুনাটি সমগ্রকের প্রতিনিধিত্বকারী হয়ে থাকে।

(খ) এ দৈব নমুনায়নে সমগ্রকের পরামানগুলো প্রাক্কলন করা যায়।

(গ) সমগ্রকের উপাদানগুলো সমজাতীয় হলে এ পদ্ধতিতে ছোট আকারের নমুনা চয়ন করে ভাল ফলাফল পাওয়া সম্ভব। এতে ব্যয় ও সময় কম লাগে।

অসুবিধা (Disadvantages):

(ক) সরল দৈব নমুনা নির্বাচনের জন্য নমুনা কাঠামো প্রয়োজন। সঠিক এবং আধুনিক নমুনা কাঠামো সব সময় পাওয়া যায় না। ফলে এ নমুনায়ন পদ্ধতি সব সময় যথোপযুক্ত হয় না।

(খ) নমুনা এককগুলো বিস্তৃত এলাকায় ছড়িয়ে ছিটিয়ে থাকলে এ পদ্ধতিতে নমুনা সংগ্রহ করতে অর্থ ও সময় বেশি ব্যয় হয়।

(গ) এ নমুনায়নে নমুনার আকার যথেষ্ট বড় না হলে নমুনা এককগুলো সমগ্রকের এককগুলোর প্রতিনিধিত্বকারী নাও হতে পারে। ফলে নমুনা হতে প্রাপ্ত প্রাক্কলক যথার্থতা হারিয়ে ফেলে।

সংক্ষিপ্ত প্রশ্ন ও উত্তর ঃ

(১) সরল দৈব নমুনায়নে একটি নমুনা নির্বাচিত হবার সম্ভাবনা কত ?

উত্তর: পুন:স্থাপন ব্যতীত সরল দৈব নমুনায়ন পদ্ধতিতে N আকার বিশিষ্ট তথ্যবিশ্ব থেকে n আকারের ${}^{N}c_{n}$ সংখ্যক নমুনা চয়ন করা যায়। নমুনায়নে একটি নমুনা নির্বাচিত হবার সম্ভাবনা $\frac{1}{{}^{N}c_{n}}$.

(২) নমুনায়ন ভগ্নাংশ কি?

উত্তর: $f=\frac{n}{N}$ কে নমুনায়ন ভগ্নাংশ (Sampling fraction) বলা হয়।

স্তরিত দৈব নমুনায়ন (Stratified Random Sampling)

প্রশ্ন: স্তরিত দৈব নমুনায়নে স্তর বলতে কি বুঝ? (জা.বি.-১১)/ স্তর (Stratum) কি?

উত্তর: যে সমগ্রক থেকে নমুনা নেয়া হবে সেই সমগ্রকের এককগুলো যদি বিভিন্ন বৈশিষ্ট্যের হয়, তবে বৈশিষ্ট্যের সমমাত্রিকতা অনুসারে এককগুলোকে বিভিন্ন শ্রেণীতে বা দলে বিভক্ত করা হলে প্রতিটি শ্রেণীকে এক একটি স্তর বলা হয়। সুতরাং প্রতিটি স্তরের এককগুলো হবে যথাসম্ভব সমমাত্রিক এবং স্তরে স্তরে ব্যবধান হবে অসমমাত্রিক।

যেমন: কোন মেট্রোপলিটন এলাকা থেকে যদি ১০০ টি পরিবারের নমুনা সংগ্রহ করতে হয়, তবে প্রথমে সংশ্লিষ্ট এলাকার পরিবারগুলোকে নিম্নবিত্ত, মধ্যবিত্ত ও উচ্চবিত্ত এ তিন শ্রেণীতে বিভক্ত করা যেতে পারে। এখানে প্রতিটি শ্রেণীকে এক একটি স্তর বলা হবে।

প্রশ্ন: স্তরীকরণের নীতিগুলো (Principles of Stratification) কি?

উত্তর: কোন সমগ্রককে স্তরে বিভক্ত করার জন্য কিছু নীতি অনুসরণ করা হয়। এ নীতিগুলো হলো:

(ক) সমগ্রকের প্রতিটি একক যেন কোন একটি স্তরের অন্তর্ভূক্ত হয়।

(খ) প্রতিটি স্তরের অন্তর্গত এককগুলোর মধ্যে তারতম্য যেন যতদূর সম্ভব কম হয় এবং স্তরে স্তরে ব্যবধান যেন যতদূর সম্ভব বেশি হয়।

প্রশ্ন: স্তরীভূত নমুনায়নে কি কি উপাদানের ভিত্তিতে সমগ্রককে স্তরিত করা হয়? (What are the stratifying factors in stratified sampling?)

উত্তর: যে নির্ণায়কের উপর ভিত্তি করে তথ্যবিশ্বকে স্তরিতকরণ করা হয়, তাকে স্তরিতকরণের উপাদান (Stratifying factor) বলা হয়। সাধারণত স্তরিতকরণের সময় যে সব উপাদান বিবেচনা করা হয়, সেগুলো হল-বয়স, লিঙ্গ, শিক্ষা, আয়ের স্তর, ভৌগোলিক আয়তন, অর্থনৈতিক স্তর ইত্যাদি।

প্রশ্ন: স্তরিত নমুনায়ন বলতে কি বুঝ? (জা.বি.-১১, ১৫)

অথবা, স্তরিত দৈব নমুনায়ন পদ্ধতি বর্ণনা কর **(জা.বি.-১২)**

উত্তর: সমগ্রকের এককগুলো সমসত্ব না হলে বৈশিষ্ট্যের সমমাত্রিকতা অনুসারে এককগুলোকে বিভিন্ন স্তরে ভাগ করে প্রতি স্তর থেকে জানা আকারের সরল দৈব নমুনা স্বাধীনভাবে চয়ন করা হলে সমস্ত পদ্ধতিকে স্তরিত দৈব নমুনায়ন বলে।

স্তরে ভাগ করার সময় লক্ষ্য রাখতে হবে যেন স্তরের ভেদাংক যতদূর সম্ভব কম এবং স্তরে স্তরে ভেদাংক বেশি হয়। অর্থাৎ প্রত্যেক স্তরের অন্তর্গত এককগুলোর বৈশিষ্ট্য যতদূর সম্ভব সমগুণ (homogeneous) সম্পন্ন হয় এবং দুটি ভিন্ন স্তরের এককগুলোর বৈশিষ্ট্য যতদূর সম্ভব অসমগুণ (heterogeneous) সম্পন্ন হয়।

যে আকারের নমুনা চয়ন করতে হবে তা স্থির রেখে কোন্ স্তর থেকে কি আকারের নমুনা নিতে হবে তাও পূর্বে নির্ধারণ করতে হয়। এরপর প্রতি স্তর থেকে নির্ধারিত আকারের নমুনা, সরল দৈব নমুনায়ন পদ্ধতিতে স্বাধীনভাবে নেয়া হলেই স্তরিত দৈব নমুনা পাওয়া যাবে।

উদাহরণ-যদি কোন সরকারি কলেজের শিক্ষকদের আয়-ব্যয় জানার জন্য একটি জরিপ করা হয়, তবে সেক্ষেত্রে শিক্ষকদের প্রভাষক, সহকারী অধ্যাপক, সহযোগী অধ্যাপক ও অধ্যাপক এ চারটি স্তরে ভাগ করে সমানুপাতিক হারে প্রতিটি স্তর থেকে দৈব নমুনা নিয়ে আয়-ব্যয়ের প্রাক্কলিত মান নির্ণয় করা হলে তা হবে স্তরিত দৈব নমুনায়ন।

প্রশ্নঃ বাস্তব ক্ষেত্রে স্তরিত দৈব নমুনায়ন কি কারণে ব্যবহারের করা হয়?

উত্তরঃ যে সমগ্রক থেকে সাধারণত নমুনা নেয়া হয় তা নানান বৈশিষ্ট্যের ভিত্তিতে অসমমাত্রিক (heterogeneous) হয়। ফলে নমুনা নির্বাচনের পূর্বে সমগ্রককে বিভিন্ন স্তরে বিভক্ত করে নমুনা নেয়া হলে তা অধিক প্রতিনিধিত্বশীল হয়, ফলে নিরূপকের ভেদাংক কম হয় এবং প্রাক্কলকের যথার্থতা বৃদ্ধি করা যায়। কারণ প্রাক্কলকের যথার্থতা উহার ভেদাংকের সংগে উল্টাভাবে নির্ভরশীল।

যেমনঃ জনকল্যানমূলক নীতির প্রতি জনসাধারণের মনোভাব জানার জন্য যদি নমুনা চয়ন করা হয়, তবে সেখানে সমগ্রককে লিঙ্গের ভিত্তিতে যেমন-পুরুষ ও মহিলা ভিত্তিক স্তরে ভাগ করা যেতে পারে। কারণ পুরুষ ও মহিলাদের মনোভাবের মধ্যে যথেষ্ট পার্থক্য থাকতে পারে। আবার সমগ্রককে শিক্ষিত ও অশিক্ষিত অর্থাৎ শিক্ষার ভিত্তিতেও স্তরিতকরণ করা যেতে পারে। এভাবে স্তরিতকরণের পর নমুনা চয়ন করা হলে জরিপের জন্য নমুনাটি অধিক প্রতিনিধিত্বশীল হবে। এভাবে প্রায় প্রতিটি জরিপেই দেখা যায় যে, জরিপের উদ্দেশ্যকে সফল করার জন্য অন্যান্য নমুনায়নের চেয়ে স্তরিত দৈব নমুনায়নই অধিক সুবিধাজনক। তাই এরূপ নমুনায়ন বাস্তবে বেশি ব্যবহৃত হয়।

আবার, তথ্য সংগ্রহের জন্যও সকল এককে একই পদ্ধতি ব্যবহারের চেয়ে ভিন্ন ভিন্ন পদ্ধতি ব্যবহার করা হলে সঠিক তথ্য কম খরচে ও কম সময়ে সংগ্রহ করা যায়, যা শুধু স্তরিত দৈব নমুনায়নেই সম্ভব। তাই এরূপ নমুনায়ন ব্যবহার করা সুবিধাজনক।

স্তরিত দৈব নমুনায়নে ব্যবহৃত প্রতীক ও পরিভাষা (Notations and Terminology used in Stratified Sampling):

ধরি, তথ্যবিশ্বে k সংখ্যক স্তর আছে এবং

$N_i = i \quad (i=1, 2, \cdots, k) -$ তম স্তরের মোট একক সংখ্যা

$N = \sum_{i=1}^{k} N_i =$ তথ্যবিশ্বের মোট একক সংখ্যা

মনে করি, $Y_{ij}\ (i=1, 2, \cdots, k,\ \ j=1, 2, \cdots, N_i)$ হল $i-$ তম স্তরের $j-$ তম এককের মান।

আরও মনে করি, $\overline{Y}_i$ এবং S_i^2 হল যথাক্রমে $i-$ তম স্তরের গড় ও ভেদাংক যাদের নিম্নভাবে সংজ্ঞায়িত করা হলোঃ

$$\overline{Y}_i = \frac{1}{N_i}\sum_{j=1}^{N_i} Y_{ij} \quad \text{এবং} \quad S_i^2 = \frac{1}{N_i - 1}\sum_{j=1}^{N_i}\left(Y_{ij} - \overline{Y}_i\right)^2$$

সমগ্রকের গড় $\overline{Y}$ হলে

$$\overline{Y} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{N_i} Y_{ij} = \frac{1}{N}\sum_{i=1}^{k} N_i \overline{Y}_i = \sum_{i=1}^{k} p_i \overline{Y}_i$$

যেখানে, $p_i = \frac{N_i}{N}$ কে i – তম স্তরের গুরুত্ব (weight) বলা হয়।

এখন, মনে করি নমুনার ক্ষেত্রে

$n_i = i$ – তম স্তর থেকে পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নের মাধ্যমে চয়ন করা নমুনার আকার।

$n = \sum_{i=1}^{k} n_i =$ সমস্ত স্তর থেকে চয়ন করা মোট নমুনার আকার।

$y_{ij} = i$ – তম স্তর থেকে চয়ন করা নমুনার j – তম এককের মান।

এখন, i – তম স্তর থেকে প্রাপ্ত নমুনার গড় ও ভেদাংক $\bar{y}_i$ ও s_i^2 দ্বারা চিহ্নিত করি, যাদের নিম্নরূপভাবে সংজ্ঞায়িত করা যায়-

$$\bar{y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij} \quad \text{এবং} \quad s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_i\right)^2$$

এখন, তথ্যবিশ্ব গড় $\bar{Y}$ এর দুটি প্রাক্কলক নিম্নরূপে বিবেচনা করি-

$$\text{নমুনার গড়, } \bar{y} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij} = \frac{1}{n}\sum_{i=1}^{k} n_i\,\bar{y}_i$$

$$\text{এবং স্তরিত দৈব নমুনার গড়, } \bar{y}_{st} = \frac{1}{N}\sum_{i=1}^{k} N_i\,\bar{y}_i = \sum_{i=1}^{k} p_i\,\bar{y}_i$$

উপপাদ্যঃ প্রমাণ কর যে, $\bar{y}_{st}$ হলো তথ্যবিশ্ব গড় $\bar{Y}$ এর নিঝুঁকি প্রাক্কলক।

অথবা, স্তরিত নমুনায়নের ক্ষেত্রে সমগ্রক গড়ের নিঝুঁকি নিরূপক নির্ণয় কর **(জা.বি.-১১)**

অথবা, স্তরিত নমুনায়নের ক্ষেত্রে সমগ্রক গড়ের নিঝুঁকি নিরূপক নির্ণয় কর এবং এই প্রাক্কলকের ভেদাংক নির্ণয় কর **(জা.বি-১৫)**

প্রমাণঃ জানি, স্তরিত দৈব নমুনার গড়, $\bar{y}_{st} = \frac{1}{N}\sum_{i=1}^{k} N_i\,\bar{y}_i$

$$\text{এবং সমগ্রকের গড়, } \bar{Y} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{N_i} Y_{ij} = \frac{1}{N}\sum_{i=1}^{k} N_i\,\bar{Y}_i$$

এখানে, $\bar{y}_i$ হল সরল দৈব নমুনায়ন পদ্ধতিতে প্রাপ্ত i – তম স্তর থেকে চয়ন করা নমুনার গড়। তাই $\bar{y}_i$ হল i – তম স্তরের গড়, $\bar{Y}_i$ এর নিঝুঁকি প্রাক্কলক অর্থাৎ $E(\bar{y}_i) = \bar{Y}_i$ । ইহা যে কোন স্তরের জন্য সঠিক।

$$\text{অতএব, } E(\bar{y}_{st}) = E\left(\frac{1}{N}\sum_{i=1}^{k} N_i\,\bar{y}_i\right)$$

$$= \frac{1}{N}\sum_{i=1}^{k} N_i\, E(\bar{y}_i)$$

$$= \frac{1}{N}\sum_{i=1}^{k} N_i\,\bar{Y}_i$$

$$= \overline{Y}$$

$$\therefore\ E(\overline{y}_{st}) = \overline{Y}\ \text{(প্রমাণিত)}$$

উপপাদ্য: দেখাও যে, স্তরিত দৈব নমুনার গড় $\overline{y}_{st}$ এর ভেদাংক $V(\overline{y}_{st}) = \frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i(N_i - n_i)S_i^2}{n_i}$

অথবা, স্তরিত নমুনায়নের ক্ষেত্রে সমগ্রক গড়ের প্রাক্কলকের ভেদাংক নির্ণয় কর **(জা.বি.-১১, ১৫)**

প্রমাণ: জানি, $\overline{y}_{st} = \frac{1}{N}\sum_{i=1}^{k}N_i\,\overline{y}_i$

এখানে, $\overline{y}_i$ হল পুন:স্থাপন ব্যতীত সরল দৈব নমুনায়ন পদ্ধতিতে প্রাপ্ত i – তম স্তর থেকে চয়ন করা নমুনার গড়।

তাই $V(\overline{y}_i) = \frac{N_i - n_i}{N_i}\frac{S_i^2}{n_i};\quad i = 1, 2, \cdots, k$

এখন, $V(\overline{y}_{st}) = V\left(\frac{1}{N}\sum_{i=1}^{k}N_i\,\overline{y}_i\right)$

$$= \frac{1}{N^2}\sum_{i=1}^{k}N_i^2\ V(\overline{y}_i)$$ [যেহেতু সকল স্তর থেকে স্বাধীনভাবে নমুনা নেয়া হয়েছে, তাই সহভেদাংকগুলোর মান শুন্য]

$$= \frac{1}{N^2}\sum_{i=1}^{k}N_i^2\,\frac{N_i - n_i}{N_i}\cdot\frac{S_i^2}{n_i}$$

$$\therefore\ V(\overline{y}_{st}) = \frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i(N_i - n_i)S_i^2}{n_i}$$

বিশেষ দ্রষ্টব্য:

(ক) লক্ষ্যণীয় যে $V(\overline{y}_{st})$, S_i^2 এর উপর নির্ভরশীল। তাই স্তরগুলোর ভিতরের ভেদাংক ছোট হলে অর্থাৎ প্রতিটি স্তরের উপাদানগুলোর বৈশিষ্ট্য সমজাতীয় হলে স্তরিত দৈব নমুনার প্রাক্কলকগুলো অধিক যথার্থ হয়। আর এ নীতির উপর ভিত্তি করেই তথ্যবিশ্বকে স্তরিতকরণ করা হয়।

(খ) স্তরিত নমুনা গড়ের ভেদাংকের প্রাক্কলিত মান নির্ণয়-

জানি, $V(\overline{y}_{st}) = \frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i(N_i - n_i)S_i^2}{n_i}$

সাধারণত S_i^2 এর মান জানা থাকে না। যেহেতু প্রত্যেক স্তর থেকে সরল দৈব নমুনা চয়ন করা হয়, তাই $E(s_i^2) = S_i^2$ ।

অতএব, $V(\overline{y}_{st})$ এর নিঝুঁকি প্রাক্কলককে নিম্নরূপে লেখা যায়-

$$\text{Est. } V(\overline{y}_{st}) = \frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i(N_i - n_i)s_i^2}{n_i}$$

উপপাদ্য: প্রমাণ কর যে, যদি প্রত্যেক স্তরের নমুনার আকার সেই স্তরের আকারের সমানুপাতিক হয় অর্থাৎ যদি প্রত্যেক স্তরের নমুনার ভগ্নাংশ সমান হয়, তবে নমুনার গড় $\bar{y}$, তথ্যবিশ্ব গড় $\bar{Y}$ এর নিঝুঁকি নিরূপক।

প্রমাণ: জানি, নমুনার গড়, $\bar{y} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij} = \frac{1}{n}\sum_{i=1}^{k} n_i\,\bar{y}_i$

এবং সমগ্রকের গড়, $\bar{Y} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{N_i} Y_{ij} = \frac{1}{N}\sum_{i=1}^{k} N_i\,\bar{Y}_i$

দেওয়া আছে, প্রত্যেক স্তরের নমুনার আকার সেই স্তরের আকারের সমানুপাতিক অর্থাৎ

$$n_i \propto N_i$$

$$\Rightarrow n_i = c\ N_i \quad [\text{ c হলো ধ্রুবক }]$$

$$\Rightarrow \frac{n_i}{N_i} = c$$

$$\Rightarrow \frac{n_i}{N_i} = \frac{n}{N}$$

$$\Rightarrow \frac{n_i}{n} = \frac{N_i}{N}$$

জানি, প্রত্যেক স্তরের জন্য $E(\bar{y}_i) = \bar{Y}_i$

অতএব, $E(\bar{y}) = E\left[\frac{1}{n}\sum_{i=1}^{k} n_i\,\bar{y}_i\right]$

$$= \frac{1}{n}\sum_{i=1}^{k} n_i\ E(\bar{y}_i)$$

$$= \sum_{i=1}^{k} \frac{n_i}{n}\ \bar{Y}_i$$

$$= \sum_{i=1}^{k} \frac{N_i}{N}\ \bar{Y}_i \qquad \left[\because \ \frac{n_i}{n} = \frac{N_i}{N}\right]$$

$$= \frac{1}{N}\sum_{i=1}^{k} N_i\,\bar{Y}_i$$

$\therefore \ E(\bar{y}) = \bar{Y}$ (প্রমাণিত)

প্রশ্ন: তথ্যবিশ্ব সমষ্টি Y এর প্রাক্কলক ও এর ভেদাংক নির্ণয় কর।

উত্তর: জানি, তথ্যবিশ্ব সমষ্টি, $Y = N\bar{Y}$

যেহেতু স্তরিত দৈব নমুনায়নে $\bar{Y}$ এর নিঝুঁকি প্রাক্কলক হচ্ছে $\bar{y}_{st}$

অর্থাৎ $E(\bar{y}_{st}) = \bar{Y}$

$$\Rightarrow E(N\ \bar{y}_{st}) = N\bar{Y}$$

$$\Rightarrow E(N\ \bar{y}_{st}) = Y$$

$\therefore$ $N\ \bar{y}_{st}$ হলো তথ্যবিশ্ব সমষ্টি Y এর নিঝুকি প্রাক্কলক।

অর্থাৎ, $\hat{Y}_{st} = N\bar{y}_{st}$

এখন, $V(\hat{Y}_{st}) = V(N\ \bar{y}_{st}) = N^2\ V(\bar{y}_{st})$

$$= N^2 \frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i(N_i - n_i)S_i^2}{n_i}$$

$$= \sum_{i=1}^{k}\frac{N_i(N_i - n_i)S_i^2}{n_i}$$

এবং $Est.V(\hat{Y}_{st}) = \sum_{i=1}^{k}\frac{N_i(N_i - n_i)s_i^2}{n_i}$

বি. দ্রঃ $\bar{y}_{st}$ এর পরিমিত বিচ্যুতি (Standard Error-SE) হলো

$$S.E(\bar{y}_{st}) = \sqrt{\frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i(N_i - n_i)S_i^2}{n_i}}$$

$$Est.\ S.E(\bar{y}_{st}) = \sqrt{\frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i(N_i - n_i)s_i^2}{n_i}}$$

প্রশ্নঃ স্তরিত দৈব নমুনায়নে বণ্টন সমস্যা বলতে কি বুঝ? এ প্রেক্ষিতে সমানুপাতিক বণ্টন এবং অপটিমাম বণ্টনের বর্ণনা দাও। এ সমস্যা সমাধানের উপায় বর্ণনা কর।

অথবা, বণ্টন বলতে কি বুঝ? স্তরিত দৈব নমুনায়ন বিভিন্ন প্রকার বণ্টন পদ্ধতি বর্ণনা কর **(জা.বি.-১৬)**

উত্তরঃ স্তরিত দৈব নমুনায়নে দুটি গুরুত্বপূর্ণ সমস্যা হলো-

(ক) সমগ্রককে সঠিকভাবে স্তরে বিভক্ত করা এবং

(খ) কোন স্তর থেকে কি আকারের নমুনা চয়ন করতে হবে তা নির্ধারণ করা।

উপরোক্ত সমস্যাগুলোর মধ্যে ২য় সমস্যাটি হলো বণ্টন সমস্যা। সুতরাং স্তরিত দৈব নমুনার আকারকে স্থির রেখে কোন্ স্তর থেকে কি আকারের নমুনা চয়ন করতে হবে তা নির্ধারণ করার সমস্যাই হল বণ্টন সমস্যা।

এ সমস্যা তিনটি উপায়ে সমাধান করা যায়-

(ক) সমান বণ্টন (Equal allocation)

(খ) সমানুপাতিক বণ্টন (Proportional allocation)

(গ) সর্বোত্তম বা অপটিমাম বণ্টন (Optimum allocation)

(ক) সমান বণ্টন (Equal allocation): যে বণ্টনে প্রতিটি স্তর থেকে সমান আকারের নমুনা চয়ন করা হয়, তাকে সমান বণ্টন বলে। সমান বণ্টন অনুসারে $i-$ তম স্তরের নমুনার আকার

$$n_i = \frac{n}{k}$$

এখানে, $n =$ মোট নমুনার আকার এবং $k =$ স্তর সংখ্যা

(খ) সমানুপাতিক বণ্টন (Proportional allocation): বিভিন্ন স্তরের নমুনার আকার (n_i) নির্ধারণকে তখনই সমানুপাতিক বণ্টন বলা হয়, যখন প্রত্যেক স্তরের নমুনা ভগ্নাংশ (sample fraction) সমান হয়। অর্থাৎ

$$\frac{n_1}{N_1} = \frac{n_2}{N_2} = \cdots = \frac{n_k}{N_k} = \frac{\sum n_i}{\sum N_i} = \frac{n}{N} = c$$ হয়, যেখানে c একটি ধ্রুবক।

$$\Rightarrow \frac{n_i}{N_i} = c$$

$$\Rightarrow n_i = c\,N_i$$

$$\therefore\ n_i \propto N_i \qquad [i = 1, 2, \cdots, k]$$

এখানে, $N_i = i-$তম স্তরের আকার।

$n_i = i-$ তম স্তর থেকে যে নমুনা নেয়া হবে তার আকার।

সমানুপাতিক বণ্টনে কোন স্তরের নমুনার আকার (n_i), উক্ত স্তরের আকারের (N_i) সমানুপাতিক হয়।

বি. দ্রঃ জানি, $n_i = c\,N_i$ ....(i)

$$\Rightarrow \sum n_i = c \sum N_i$$

$$\Rightarrow n = c\,N$$

$$\Rightarrow c = \frac{n}{N}$$

$$\therefore\ n_i = \frac{n}{N} N_i$$ [(i) নং c এর মান বসিয়ে]

(গ) সর্বোত্তম বা অপটিমাম বণ্টন (Optimum allocation)/ সর্বোত্তম বণ্টন কি? (জা.বি.-১১):

এ পদ্ধতিতে $i-$ তম স্তর থেকে n_i আকারের নমুনা এমনভাবে নির্ধারণ করতে হবে যেন-

(ক) নির্দিষ্ট আকারের নমুনা 'n' এবং নির্দিষ্ট পরিমাণ খরচের জন্য তথ্যবিশ্বের গড়ের নিরূপকের ভেদাংক সবচেয়ে কম হয়।

(খ) নির্দিষ্ট পরিমাণ যথার্থতা (precision) পাওয়ার জন্য মোট খরচের পরিমাণ যেন ন্যূনতম হয়।

উপরোক্ত বণ্টন নীতি অনুসারে যদি বিভিন্ন স্তরে n_i সমূহ বণ্টন করা হয়, তবে তাকে অপটিমাম বণ্টন বলে।

বি. দ্রঃ অপটিমাম বণ্টনের মধ্যে সবচেয়ে সহজ এবং উল্লেখযোগ্য হল নেম্যানের সর্বোত্তম বণ্টন (optimum allocation of Neyman)। এ পদ্ধতিতে প্রতি স্তরের নমুনার আকার n_i, স্তরের আকার N_i এর সমানুপাতিক না হয়ে $N_i S_i$ এর সমানুপাতিক হয়। তাই নমুনার আকার বণ্টনে যদি n_i গুলি $N_i S_i$ এর সমানুপাতিক হয়, তবে তাকে নেম্যানের সর্বোত্তম বণ্টন বলে। অর্থাৎ

$$n_i \propto N_i\ S_i$$

$$\Rightarrow n_i = c\,N_i\,S_i \,............(i)$$

$$\Rightarrow \sum_{i=1}^{k} n_i = \sum_{i=1}^{k} c\,N_i\,S_i$$

$$\Rightarrow n = c\sum_{i=1}^{k} N_i\ S_i$$

$$\Rightarrow c = \frac{n}{\sum_{i=1}^{k} N_i\,S_i}$$

$$\therefore\ \ n_i = \frac{n}{\sum_{i=1}^{k} N_i\,S_i} N_i\,S_i \quad \text{[(i) নং c এর মান বসিয়ে]}$$

এখানে, $n =$ স্তরিত দৈব নমুনার আকার।

$N_i = i -$ তম স্তরের আকার।

$S_i = i -$ তম স্তরের পরিমিত ব্যবধান।

এইরূপ নমুনায়ন তখনই বেশি উপযোগী, যখন বিভিন্ন স্তরের পরিমিত ব্যবধান বেশি হয়।

প্রশ্নঃ সমানুপাতিক ও অপটিমাম বণ্টন পদ্ধতির সুবিধা-অসুবিধাগুলো লিখ।

উত্তরঃ সমানুপাতিক বণ্টনের সুবিধা-অসুবিধাসমূহ নিম্নরূপ-

সমানুপাতিক বণ্টনের সুবিধাসমূহ নিম্নরূপ-

(ক) এ পদ্ধতিতে নমুনা নির্বাচন করলে নমুনার প্রতিনিধিত্বশীলতা ও নির্ভরযোগ্যতা বৃদ্ধি পায়।

(খ) সমানুপাতিক স্তরিত দৈব নমুনায় সমগ্রকের বৈশিষ্ট্যাবলীর সত্যিকার প্রতিফলন ঘটে বলে নমুনা ভ্রান্তি কম হয় এবং

(গ) এ পদ্ধতিতে নমুনা নির্বাচনে প্রতিটি স্তরকে ভার আরোপিত করতে হয় না।

সমানুপাতিক বন্টনের অসুবিধাসমূহ নিম্নরূপ-

(ক) সরল দৈব নমুনায়ন পদ্ধতির চেয়ে সমানুপাতিক স্তরিত নমুনায়ন পদ্ধতি বেশি জটিল। ফলে এ পদ্ধতিতে নমুনা নির্বাচনে অপেক্ষাকৃত বেশি সময় লাগে।

(খ) এ পদ্ধতিতে নির্ভুলভাবে নমুনা নির্বাচন করতে হলে গবেষককে সমগ্রকের গঠন প্রকৃতি ও অন্যান্য বৈশিষ্ট্যাবলী সম্পর্কে বিস্তারিতভাবে জানতে হয়, যা সব সময় সম্ভব হয় না।

(গ) সমানুপাতিক স্তরিত নমুনায় স্তর বিন্যাসজনিত ভ্রান্তি ঘটার সম্ভাবনা থাকে।

অপটিমাম বন্টনের সুবিধা-অসুবিধাসমূহ নিম্নরূপ-

সুবিধাসমূহ-

(ক) এ পদ্ধতিতে নমুনা নেয়া হলে নিরূপকের ভেদাংক কম হয়।

(খ) বিভিন্ন স্তরের এককগুলোর মধ্যে পরিমিত ব্যবধান S_i গুলো বেশি হলে এ পদ্ধতিটি সুবিধাজনক।

অসুবিধাসমূহ-

(ক) প্রতি স্তরের পরিমিত ব্যবধান অধিকাংশ সময়ই জানা থাকে না। ফলে এ বন্টন পদ্ধতি ব্যবহার করা অসুবিধাজনক।

(খ) এ পদ্ধতিতে নমুনা ভ্রান্তি সৃষ্টি হতে পারে।

প্রশ্ন: সমানুপাতিক বন্টন ক্ষেত্রে স্তরিত দৈব নমুনার গড়ের ভেদাংক নির্ণয় কর।

উত্তর: মনে করি, N আকারের কোন সমগ্রককে k সংখ্যক স্তরে ভাগ করা হয়েছে। ধরি, N_i এবং $\overline{Y}_i$ যথাক্রমে $i-$তম স্তরের আকার এবং গড়। সমগ্রকের গড় $\overline{Y}$ হলে $\overline{Y} = \dfrac{\sum N_i \overline{Y}_i}{N}$ যেখানে, $\sum_{i=1}^{k} N_i = N$.

ধরি, $i-$তম স্তর থেকে নেয়া n_i আকারের নমুনার গড় $\overline{y}_i$। সুতরাং স্তরিত দৈব নমুনার গড় $\overline{y}_{st} = \dfrac{1}{N}\sum_{i=1}^{k} N_i \overline{y}_i$ হলো $\overline{Y}$ এর নির্ঝুকি নিরূপক।

জানি, $V(\overline{y}_{st}) = \dfrac{1}{N^2}\sum_{i=1}^{k}\dfrac{N_i(N_i-n_i)S_i^2}{n_i}$

$$\Rightarrow V(\overline{y}_{st}) = \frac{1}{N^2}\sum_{i=1}^{k} N_i\left(\frac{N_i}{n_i}-1\right)S_i^2 \text{ ....(i)}$$

কিন্তু সমানুপাতিক বণ্টনে $\dfrac{n_i}{N_i} = \dfrac{n}{N}$

$$\Rightarrow \frac{N_i}{n_i} = \frac{N}{n}$$

সুতরাং (i) নং কে লেখা যায়,

$$V(\overline{y}_{st})_{prop} = \frac{1}{N^2}\sum_{i=1}^{k} N_i\left(\frac{N}{n}-1\right)S_i^2$$

$$=\frac{1}{N^2}\sum_{i=1}^{k}N_i\left(\frac{N-n}{n}\right)S_i^2$$

$$=\frac{N-n}{N^2 n}\sum_{i=1}^{k}N_i\ S_i^2$$

$$=\frac{N-n}{N n}\sum_{i=1}^{k}\frac{N_i}{N}\ S_i^2$$

$$\therefore\quad V(\bar{y}_{st})_{prop}=\left(\frac{1}{n}-\frac{1}{N}\right)\sum_{i=1}^{k}p_i\ S_i^2\quad [\text{যেখানে},\ p_i=\frac{N_i}{N}\ ]$$

প্রশ্নঃ নেম্যানের অপটিমাম বণ্টনের ক্ষেত্রে স্তরিত দৈব নমুনার গড়ের ভেদাংক নির্ণয় কর।

অথবা, নেম্যান বণ্টনে দেখাও যে $V(\bar{y}_{st})=\frac{1}{n}\left(\sum_{i=1}^{k}P_i\ S_i^2\right)-\frac{1}{N}\sum_{i=1}^{k}P_i\ S_i^2$ **; যেখানে** $P_i=\frac{N_i}{N}$

উত্তরঃ জানি, $V(\bar{y}_{st})=\frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i(N_i-n_i)S_i^2}{n_i}$

$$=\frac{1}{N^2}\left(\sum_{i=1}^{k}\frac{N_i^2\ S_i^2}{n_i}-\sum_{i=1}^{k}\frac{N_i\ n_i\ S_i^2}{n_i}\right)$$

$$\therefore\quad V(\bar{y}_{st})=\frac{1}{N^2}\left(\sum_{i=1}^{k}\frac{N_i^2\ S_i^2}{n_i}-\sum_{i=1}^{k}N_i\ S_i^2\right)\ldots\ldots\ldots\ldots.(ii)$$

কিন্তু নেম্যানের অপটিমাম বণ্টনে

$$n_i=\frac{n\ N_i\ S_i}{\sum_{i=1}^{k}N_i\ S_i}$$

সুতরাং (ii) নং কে লেখা যায়,

$$V(\bar{y}_{st})_{opt}=\frac{1}{N^2}\left(\sum_{i=1}^{k}\frac{N_i^2\ S_i^2}{\frac{n\ N_i\ S_i}{\sum_{i=1}^{k}N_i\ S_i}}-\sum_{i=1}^{k}N_i\ S_i^2\right)$$

$$=\frac{1}{N^2}\left(\sum_{i=1}^{k}\frac{N_i^2\ S_i^2\sum_{i=1}^{k}N_iS_i}{n\ N_i\ S_i}-\sum_{i=1}^{k}N_i\ S_i^2\right)$$

$$=\frac{1}{N^2}\left(\sum_{i=1}^{k}\frac{N_i\ S_i\sum_{i=1}^{k}N_i\ S_i}{n}-\sum_{i=1}^{k}N_i\ S_i^2\right)$$

$$=\frac{1}{N^2}\left(\frac{\sum_{i=1}^{k}N_i\ S_i\ \sum_{i=1}^{k}N_i\ S_i}{n}-\sum_{i=1}^{k}N_i\ S_i^2\right)$$

$$=\frac{1}{N^2}\left(\frac{\left(\sum_{i=1}^{k}N_i\ S_i\right)^2}{n}-\sum_{i=1}^{k}N_i\ S_i^2\right)$$

$$=\frac{\left(\sum_{i=1}^{k}N_i\ S_i\right)^2}{N^2\ n}-\frac{\sum_{i=1}^{k}N_i\ S_i^2}{N^2}$$

$$=\frac{1}{n}\left(\frac{\sum_{i=1}^{k}N_i\ S_i}{N}\right)^2-\frac{1}{N}\sum_{i=1}^{k}\frac{N_i\ S_i^2}{N}$$

$$=\frac{1}{n}\left(\sum_{i=1}^{k}\frac{N_i}{N}\ S_i\right)^2-\frac{1}{N}\sum_{i=1}^{k}\frac{N_i}{N}\ S_i^2$$

$$=\frac{1}{n}\left(\sum_{i=1}^{k}p_i\ S_i\right)^2-\frac{1}{N}\sum_{i=1}^{k}p_i\ S_i^2 \quad \text{যেখানে, } p_i=\frac{N_i}{N}$$

ব্যয় অপেক্ষক (Cost function): যে কোন নমুনা জরিপে নমুনা একক থেকে নেয়া তথ্যের মূলমানকে উহা হতে তথ্য সংগ্রহের খরচের সাথে সর্বদাই সমতা আনার চেষ্টা করা হয়। স্তরিত দৈব নমুনায়নে কোন এক স্তরের একক থেকে তথ্য সংগ্রহের খরচের চেয়ে অন্য স্তরের একক থেকে তথ্য সংগ্রহের খরচ বেশি হতে পারে।

উদাহরণস্বরূপ-যাতায়াতের খরচের জন্য শহর এলাকার লোকদের নিকট হতে তথ্য সংগ্রহ করার চেয়ে গ্রামীণ এলাকার লোকদের নিকট হতে তথ্য সংগ্রহের খরচ বেশি হয়ে থাকে। স্তরিত দৈব নমুনায়নে ব্যয় অপেক্ষকের সহজতম রূপকে নিম্নের সরলরৈখিক মডেলের মাধ্যমে উপস্থাপন করা যায়-

$$C=a+\sum_{i=1}^{k}C_i\ n_i$$

যেখানে, ‘a’ হলো সার্বিক ব্যয় (overhead cost) এবং C_i হলো i – তম স্তরের একক/উপাদান প্রতি ব্যয়।

উপপাদ্য: স্থির নমুনার আকার n এর জন্য স্তরিত নমুনা গড়ের ভেদাংক $V(\bar{y}_{st})$ এর মান সর্বনিম্ন হবে যদি $n_i \propto N_i S_i$ হয়।

প্রমাণ: ধরি, স্তরিত দৈব নমুনায়নে তথ্যবিশ্বের মোট N সংখ্যক উপাদানকে k স্তরে বিভক্ত করা হয়েছে এবং উহা হতে n আকারের নমুনা সংগ্রহ করা হয়েছে। যদি $i\ (i=1, 2, \cdots, k)$ – তম স্তরের আকার N_i এবং নমুনার আকার n_i হয়, তবে জানি স্তরিত দৈব নমুনা গড়ের $(\bar{y}_{st})$ ভেদাংক

$$V(\bar{y}_{st})=\frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i\,(N_i-n_i)S_i^2}{n_i} \text{ ....(i)}$$

এখন $V(\bar{y}_{st})$ এর মান সর্বনিম্ন করতে হবে যেন

$$\sum_{i=1}^{k} n_i = n$$ (স্থির) হয়..........(ii)

(ii) নং শর্তের সাপেক্ষে (i) নং এর মান সর্বনিম্ন হবার অর্থ হলো

$\phi = V(\bar{y}_{st}) + \lambda\left(\sum_{i=1}^{k} n_i - n\right)$ সর্বনিম্ন হওয়া, যেখানে λ হলো Lagrange's গুণক।

অর্থাৎ $$\phi = \frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i(N_i - n_i)S_i^2}{n_i} + \lambda\left(\sum_{i=1}^{k} n_i - n\right)$$

$$\Rightarrow \phi = \frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i^2 S_i^2}{n_i} - \frac{1}{N^2}\sum_{i=1}^{k} N_i S_i^2 + \lambda\left(\sum_{i=1}^{k} n_i - n\right)$$

এখন, n_i এর সাপেক্ষে ϕ এর মান সর্বনিম্ন হবে যদি $\frac{\delta\phi}{\delta n_i} = 0$ হয়।

অতএব, $$\frac{\delta\phi}{\delta n_i} = -\frac{N_i^2 S_i^2}{N^2 n_i^2} + \lambda = 0$$

$$\Rightarrow \frac{N_i^2 S_i^2}{N^2 n_i^2} = \lambda$$

$$\Rightarrow \frac{N^2 n_i^2}{N_i^2 S_i^2} = \frac{1}{\lambda}$$

$$\Rightarrow n_i^2 = \frac{N_i^2 S_i^2}{N^2 \lambda}$$

$$\Rightarrow n_i = \frac{N_i S_i}{N\sqrt{\lambda}} \text{............(iii)}$$

আবার, $$\frac{\delta^2\phi}{\delta n_i^2} = \frac{2N_i^2 S_i^2}{N^2 n_i^3} > 0$$

যা দ্বারা প্রমাণিত হয় যে, (iii) নং এ উল্লেখিত n_i এর জন্য ϕ এর মান সর্বনিম্ন হবে।

এখন (iii) নং এর উভয় পক্ষে i – এর মান 1 থেকে k পরিসরে যোগ করে পাই,

$$\sum_{i=1}^{k} n_i = \frac{\sum_{i=1}^{k} N_i S_i}{N\sqrt{\lambda}}$$

$$\Rightarrow n = \frac{\sum_{i=1}^{k} N_i S_i}{N\sqrt{\lambda}}$$ [(ii) নং হতে পাই]

$$\Rightarrow \sqrt{\lambda} = \frac{\sum_{i=1}^{k} N_i S_i}{Nn}$$

$\sqrt{\lambda}$ এর মান (iii) নং এ বসিয়ে পাই,

$$\Rightarrow n_i = \frac{N_i S_i}{N \frac{\sum_{i=1}^{k} N_i S_i}{Nn}} = \frac{N_i S_i}{\frac{\sum_{i=1}^{k} N_i S_i}{n}} = \frac{n}{\sum_{i=1}^{k} N_i S_i} N_i S_i$$

$$\Rightarrow n_i = k\, N_i S_i \qquad \left[\frac{n}{\sum_{i=1}^{k} N_i S_i} = k \text{ (ধ্রুবক) ধরে } \right]$$

$$\therefore\ n_i \propto N_i S_i$$

ইহাকে নেম্যানের সর্বোত্তম বণ্টন বলা হয়।

ইহা হতে বুঝা যায় যে, যে স্তরের $N_i S_i$ এর মান বেশি সেখান থেকে বেশি সংখ্যক নমুনা একক চয়ন করা হলে স্তরিত নমুনা গড়ের যথার্থতা (precision) বৃদ্ধি পাবে।

NB: $$\phi = \frac{1}{N^2}\sum_{i=1}^{k} \frac{N_i^2 S_i^2}{n_i} - \frac{1}{N^2}\sum_{i=1}^{k} N_i S_i^2 + \lambda\left(\sum_{i=1}^{k} n_i - n\right)$$

$$\Rightarrow \phi = \frac{1}{N^2}\left(\frac{N_1^2 S_1^2}{n_1} + \frac{N_2^2 S_2^2}{n_2} + \cdots + \frac{N_k^2 S_k^2}{n_k}\right) - \frac{1}{N^2}\sum_{i=1}^{k} N_i S_i^2 + \lambda\left(n_1 + n_2 + \cdots + n_k - n\right)$$

Now, $$\frac{\delta\phi}{\delta n_1} = \frac{1}{N^2}\left(-\frac{N_1^2 S_1^2}{n_1^2}\right) - 0 + \lambda = -\frac{N_1^2 S_1^2}{N^2 n_1^2} + \lambda$$

Again, $$\frac{\delta\phi}{\delta n_2} = \frac{1}{N^2}\left(-\frac{N_2^2 S_2^2}{n_2^2}\right) - 0 + \lambda = -\frac{N_2^2 S_2^2}{N^2 n_2^2} + \lambda$$

Similarly, $$\frac{\delta\phi}{\delta n_i} = -\frac{N_i^2 S_i^2}{N^2 n_i^2} + \lambda$$

উপপাদ্য: দেখাও যে, স্তরিত দৈব নমুনায়নে প্রদত্ত ব্যয় অপেক্ষক $C = a + \sum_{i=1}^{k} C_i n_i$ **এর জন্য** $V\left(\bar{y}_{st}\right)$ **এর মান সর্বনিম্ন হবে যদি** $n_i \propto \frac{N_i S_i}{\sqrt{C_i}}$ **। উক্ত** n_i **এর জন্য নমুনা আকার** n **এর মান নির্ণয় কর।**

প্রমাণ: ধরি, স্তরিত দৈব নমুনায়নে তথ্যবিশ্বের মোট N সংখ্যক উপাদানকে k স্তরে বিভক্ত করা হয়েছে এবং উহা হতে n আকারের নমুনা সংগ্রহ করা হয়েছে। যদি $i\ (i = 1, 2, \cdots, k)$ – তম স্তরের আকার N_i এবং নমুনার আকার n_i হয়, তবে জানি স্তরিত দৈব নমুনা গড়ের $\left(\bar{y}_{st}\right)$ ভেদাংক

$$V\left(\bar{y}_{st}\right)=\frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i\left(N_i-n_i\right)S_i^{\,2}}{n_i}\;\text{...........(i)}$$

এখন $V\left(\bar{y}_{st}\right)$ এর মান সর্বনিম্ন করতে হবে যেন

$$\sum_{i=1}^{k}C_i\,n_i=C-a\;\text{হয়..........(ii)}$$

(ii) নং শর্তের সাপেক্ষে (i) নং এর মান সর্বনিম্ন হবার অর্থ হলো

$\phi=V\left(\bar{y}_{st}\right)+\lambda\left(\sum_{i=1}^{k}C_i\,n_i-C+a\right)$ এর মান সর্বনিম্ন হওয়া, যেখানে λ হলো Lagrange's গুণক।

অর্থাৎ $\phi=\frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i(N_i-n_i)S_i^{\,2}}{n_i}+\lambda\left(\sum_{i=1}^{k}C_i\,n_i-C+a\right)$

$$=\frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i^{\,2}S_i^{\,2}}{n_i}-\frac{1}{N^2}\sum_{i=1}^{k}N_i\,S_i^{\,2}+\lambda\left(\sum_{i=1}^{k}C_i\,n_i-C+a\right)$$

এখন, n_i এর সাপেক্ষে ϕ এর মান সর্বনিম্ন হবে যদি $\frac{\delta\phi}{\delta n_i}=0$ হয়।

অতএব, $\frac{\delta\phi}{\delta n_i}=-\frac{N_i^{\,2}S_i^{\,2}}{N^2\,n_i^{\,2}}+\lambda\,C_i=0$

$$\Rightarrow\frac{N_i^{\,2}S_i^{\,2}}{N^2\,n_i^{\,2}}=\lambda C_i$$

$$\Rightarrow\frac{N^2\,n_i^{\,2}}{N_i^{\,2}S_i^{\,2}}=\frac{1}{\lambda C_i}$$

$$\Rightarrow n_i^{\,2}=\frac{N_i^{\,2}S_i^{\,2}}{N^2\,\lambda C_i}$$

$$\Rightarrow n_i=\frac{N_i\,S_i}{N\sqrt{\lambda}\sqrt{C_i}}$$

$$\Rightarrow n_i=\frac{\frac{N_i\,S_i}{\sqrt{C_i}}}{N\sqrt{\lambda}}\;\text{............(iii)}$$

আবার, $\frac{\delta^2\phi}{\delta n_i^{\,2}}=\frac{2N_i^{\,2}S_i^{\,2}}{N^2\,n_i^{\,3}}>0$

যা দ্বারা প্রমাণিত হয় যে, (iii) নং এ উল্লেখিত n_i এর জন্য ϕ এর মান সর্বনিম্ন হবে।

এখন (iii) নং এর উভয় পক্ষে i – এর মান 1 থেকে k পরিসরে যোগ করে পাই,

$$\sum_{i=1}^{k} n_i = \frac{\sum_{i=1}^{k} N_i S_i / \sqrt{C_i}}{N\sqrt{\lambda}}$$

$$\Rightarrow n = \frac{\sum_{i=1}^{k} N_i S_i / \sqrt{C_i}}{N\sqrt{\lambda}}$$

$$\Rightarrow \sqrt{\lambda} = \frac{\sum_{i=1}^{k} N_i S_i / \sqrt{C_i}}{Nn}$$

$\sqrt{\lambda}$ এর মান (iii) নং এ বসিয়ে পাই,

$$\Rightarrow n_i = \frac{N_i S_i / \sqrt{C_i}}{N\frac{\sum_{i=1}^{k}\left(N_i S_i / \sqrt{C_i}\right)}{Nn}} = \frac{N_i S_i / \sqrt{C_i}}{\frac{\sum_{i=1}^{k}\left(N_i S_i / \sqrt{C_i}\right)}{n}}$$

$$\Rightarrow n_i = \frac{n}{\sum_{i=1}^{k}\left(N_i S_i / \sqrt{C_i}\right)} N_i S_i / \sqrt{C_i} \text{ ....(iv)}$$

$$\Rightarrow n_i = k\, N_i S_i / \sqrt{C_i} \qquad \left[\ \frac{n}{\sum_{i=1}^{k}\left(N_i S_i / \sqrt{C_i}\right)} = k \text{ (ধ্রুবক) ধরে } \right]$$

$$\therefore\ n_i \propto N_i S_i / \sqrt{C_i}$$

অর্থাৎ সর্বোত্তম বণ্টনে একটি পরিমাণ খরচের জন্য $n_i \propto N_i S_i / \sqrt{C_i}$. (iv) নং সূত্রানুসারে বিভিন্ন স্তরে n_i এর মান বণ্টন করা হলে তাকে সর্বোত্তম বণ্টন (optimum allocation) বলে।

নমুনার আকার 'n' এর মান নির্ণয়:

একটি পরিমাণ খরচ 'C' এর জন্য 'n' এর মান নির্ণয় করতে (iv) নং হতে প্রাপ্ত n_i এর মান, ব্যয় অপেক্ষক $C = a + \sum_{i=1}^{k} C_i n_i$ এ বসিয়ে পাই,

$$C = a + \sum_{i=1}^{k} C_i \left[\frac{n}{\sum_{i=1}^{k}\left(N_i S_i / \sqrt{C_i}\right)} N_i S_i / \sqrt{C_i}\right]$$

$$\Rightarrow C - a = \sum_{i=1}^{k}\left[\frac{C_i\, n \frac{N_i S_i}{\sqrt{C_i}}}{\sum_{i=1}^{k}\left(N_i S_i / \sqrt{C_i}\right)}\right]$$

$$\Rightarrow C-a=\sum_{i=1}^{k}\left[\frac{n\,N_i\,S_i\,\sqrt{C_i}}{\sum_{i=1}^{k}\left(N_i\,S_i/\sqrt{C_i}\right)}\right]$$

$$\Rightarrow C-a=\frac{n\sum_{i=1}^{k}N_i\,S_i\,\sqrt{C_i}}{\sum_{i=1}^{k}\left(N_i\,S_i/\sqrt{C_i}\right)}$$

$$\Rightarrow n\sum_{i=1}^{k}N_i\,S_i\sqrt{C_i}=(C-a)\sum_{i=1}^{k}\left(N_i\,S_i/\sqrt{C_i}\right)$$

$$\Rightarrow n=\frac{(C-a)\sum_{i=1}^{k}\left(N_i\,S_i/\sqrt{C_i}\right)}{\sum_{i=1}^{k}N_i\,S_i\sqrt{C_i}}\;\ldots\ldots\ldots\ldots(v)$$

যদি প্রতি স্তর থেকে নমুনা নেয়ার খরচ $C_i=C_0$ হয়, তবে নির্দিষ্ট পরিমাণ খরচের জন্য সর্বোত্তম বণ্টন স্থির নমুনার আকারের সর্বোত্তম বণ্টনে রূপান্তরিত হয়। এক্ষেত্রে

$$C=a+\sum_{i=1}^{k}C_i\,n_i$$

$$\Rightarrow C-a=\sum_{i=1}^{k}C_0\,n_i=C_0\sum_{i=1}^{k}n_i=C_0\,n$$

এখন, (iv) নং এ $C_i=C_0$ বসিয়ে পাই,

প্রতি স্তরের নমুনার আকার, $n_i=\dfrac{n\,N_i\,S_i/\sqrt{C_0}}{\sum_{i=1}^{k}N_i\,S_i/\sqrt{C_0}}=\dfrac{n\,N_i\,S_i}{\sum_{i=1}^{k}N_i\,S_i}$

$$\Rightarrow n_i=n.\frac{N_i\,S_i}{\sum_{i=1}^{k}N_i\,S_i}$$, যা নেম্যানের সূত্র।

এবং (v) নং এ $C-a=C_0\,n$ এবং $C_i=C_0$ বসিয়ে পাই,

নমুনার আকার, $n=\dfrac{C_0\,n\sum_{i=1}^{k}\left(N_i\,S_i/\sqrt{C_0}\right)}{\sum_{i=1}^{k}N_i\,S_i\sqrt{C_0}}=\dfrac{C_0\,n\sum_{i=1}^{k}N_i\,S_i}{\sqrt{C_0}\,\sqrt{C_0}\sum_{i=1}^{k}N_i\,S_i}=\dfrac{C_0\,n}{C_0}=n$ (স্থির)

বি. দ্রঃ $n_i\propto N_i\,S_i/\sqrt{C_i}$ থেকে নিম্নোক্ত সিদ্ধান্তে উপনীত হওয়া যায়-

(ক) যে স্তরের N_i এর মান বড়, সে স্তর থেকে বড় নমুনা নিতে হবে।

(খ) যে স্তরের S_i এর মান বড়, সে স্তর থেকে বড় নমুনা নিতে হবে।

(গ) যে স্তরের প্রতি একক থেকে তথ্য সংগ্রহের খরচ $\left(C_i\right)$ এর মান কম, সে স্তর থেকে বড় আকারের নমুনা নিতে হবে।

প্রশ্ন: স্তরিত দৈব নমুনায়নে নির্দিষ্ট পরিমাণ ভেদাংক এর জন্য ব্যয় অপেক্ষকের মান সর্বনিম্ন করতে হলে n_i ও n এর মান নির্ণয় কর।

সমাধান: ধরি, ব্যয় অপেক্ষক $C = a + \sum_{i=1}^{k} C_i n_i$

$$\Rightarrow \sum_{i=1}^{k} C_i n_i = C - a \ldots\ldots\ldots\ldots(i)$$

তবে n_i ও n এর মান এমনভাবে নির্ণয় করতে হবে, যেন নির্দিষ্ট পরিমাণ ভেদাংক $V\left(\bar{y}_{st}\right) = V_0$ (ধরি) এর জন্য ব্যয় সর্বনিম্ন হয়। এখন,

$$V\left(\bar{y}_{st}\right) = V_0 \ldots\ldots\ldots\ldots(ii)$$

(ii) নং শর্তের সাপেক্ষে (i) নং এর মান সর্বনিম্ন হবার অর্থ হলো

$\phi = \sum_{i=1}^{k} C_i n_i - C + a + \lambda\left[V\left(\bar{y}_{st}\right) - V_0\right]$ এর মান সর্বনিম্ন হওয়া, যেখানে λ হলো Lagrange's গুণক

$$\Rightarrow \phi = \sum_{i=1}^{k} C_i n_i - C + a + \lambda\left[\frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i\left(N_i - n_i\right)S_i^2}{n_i} - V_0\right]$$

$$\Rightarrow \phi = \sum_{i=1}^{k} C_i n_i - C + a + \lambda\left[\frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i^2 S_i^2}{n_i} - \frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i n_i S_i^2}{n_i} - V_0\right]$$

$$\Rightarrow \phi = \sum_{i=1}^{k} C_i n_i - C + a + \lambda\left[\frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i^2 S_i^2}{n_i} - \frac{1}{N^2}\sum_{i=1}^{k} N_i S_i^2 - V_0\right]$$

এখন, n_i এর সাপেক্ষে ϕ এর মান সর্বনিম্ন হবে যদি $\frac{\delta\phi}{\delta n_i} = 0$ হয়।

অতএব, $$\frac{\delta\phi}{\delta n_i} = C_i - \frac{\lambda N_i^2 S_i^2}{N^2 n_i^2} = 0$$

$$\Rightarrow \frac{\lambda N_i^2 S_i^2}{N^2 n_i^2} = C_i$$

$$\Rightarrow C_i N^2 n_i^2 = \lambda N_i^2 S_i^2$$

$$\Rightarrow n_i^2 = \frac{\lambda N_i^2 S_i^2}{C_i N^2}$$

$$\Rightarrow n_i = \frac{\sqrt{\lambda}\, N_i S_i}{\sqrt{C_i}\, N}$$

$$\Rightarrow n_i = \frac{\sqrt{\lambda}\, N_i\, S_i / \sqrt{C_i}}{N} \text{ ....(iii)}$$

আবার, $\dfrac{\delta^2 \phi}{\delta n_i^{\,2}} = -\dfrac{2\,\lambda\, N_i^{\,2}\, S_i^{\,2}}{N^2\, n_i^{\,3}} > 0$

যা দ্বারা প্রমাণিত হয় যে, (iii) নং এ উল্লেখিত n_i এর জন্য ϕ এর মান সর্বনিম্ন হবে।

এখন (iii) নং এর উভয় পক্ষে i – এর মান 1 থেকে k পরিসরে যোগ করে পাই,

$$\sum_{i=1}^{k} n_i = \frac{\sqrt{\lambda} \sum_{i=1}^{k} N_i\, S_i / \sqrt{C_i}}{N}$$

$$\Rightarrow n = \frac{\sqrt{\lambda}\, \sum_{i=1}^{k} N_i\, S_i / \sqrt{C_i}}{N}$$

$$\Rightarrow \sqrt{\lambda}\, \sum_{i=1}^{k} N_i\, S_i / \sqrt{C_i} = n\, N$$

$$\Rightarrow \sqrt{\lambda} = \frac{n\, N}{\sum_{i=1}^{k} N_i\, S_i / \sqrt{C_i}}$$

$\sqrt{\lambda}$ এর মান (iii) নং এ বসাইয়া পাই,

$$n_i = \frac{n\, N . N_i\, S_i / \sqrt{C_i}}{\sum_{i=1}^{k} N_i\, S_i / \sqrt{C_i} . N} = \frac{n\, N_i\, S_i / \sqrt{C_i}}{\sum_{i=1}^{k} N_i\, S_i / \sqrt{C_i}}$$

$$\Rightarrow n_i = \frac{n}{\sum_{i=1}^{k} N_i\, S_i / \sqrt{C_i}}\; N_i\, S_i / \sqrt{C_i} \text{ ....(iv)}$$

$\Rightarrow n_i = K.\, N_i\, S_i / \sqrt{C_i}$ ধরি, $K = \dfrac{n}{\sum_{i=1}^{k} N_i\, S_i / \sqrt{C_i}}$

$$\therefore\; n_i \propto N_i\, S_i / \sqrt{C_i}$$

অর্থাৎ একটি নির্দিষ্ট পরিমাণ ভেদাংকের জন্য ব্যয় সর্বনিম্ন হবে, যদি $n_i \propto N_i\, S_i / \sqrt{C_i}$ হয়।

নমুনার আকার 'n' এর মান নির্ণয়:

$V(\bar{y}_{st}) = V_0$ [(ii) নং হতে]

$$\Rightarrow \frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i^2 S_i^2}{n_i} - \frac{1}{N^2}\sum_{i=1}^{k} N_i S_i^2 = V_0$$

$$\Rightarrow \frac{1}{N^2}\sum_{i=1}^{k}\frac{N_i^2 S_i^2}{n_i} = V_0 + \frac{1}{N^2}\sum_{i=1}^{k} N_i S_i^2 \text{ ....(v)}$$

এখন, n এর মান নির্ণয়ের জন্য (iv) নং হতে n_i এর মান (v) নং সমীকরণে বসিয়ে পাই,

$$\frac{1}{N^2}\sum_{i=1}^{k}\left(N_i^2 S_i^2 . \frac{\sum_{i=1}^{k} N_i S_i / \sqrt{C_i}}{n N_i S_i / \sqrt{C_i}}\right) = V_0 + \frac{1}{N^2}\sum_{i=1}^{k} N_i S_i^2$$

$$\Rightarrow \frac{1}{N^2 n}\sum_{i=1}^{k} N_i S_i \sqrt{C_i} . \sum_{i=1}^{k} N_i S_i / \sqrt{C_i} = \frac{N^2 V_0 + \sum_{i=1}^{k} N_i S_i^2}{N^2}$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{k} N_i S_i \sqrt{C_i} . \sum_{i=1}^{k} N_i S_i / \sqrt{C_i} = N^2 V_0 + \sum_{i=1}^{k} N_i S_i^2$$

$$\Rightarrow n\left(N^2 V_0 + \sum_{i=1}^{k} N_i S_i^2\right) = \sum_{i=1}^{k} N_i S_i \sqrt{C_i} . \sum_{i=1}^{k} N_i S_i / \sqrt{C_i}$$

$$\Rightarrow n = \frac{\sum_{i=1}^{k} N_i S_i \sqrt{C_i} . \sum_{i=1}^{k} N_i S_i / \sqrt{C_i}}{N^2 V_0 + \sum_{i=1}^{k} N_i S_i^2}$$

আবার, যদি $C_i = C_0$ **হয়, তখন**

$$n_i = \frac{n N_i S_i / \sqrt{C_0}}{\sum_{i=1}^{k} N_i S_i / \sqrt{C_0}} = \frac{n N_i S_i}{\sum_{i=1}^{k} N_i S_i}$$, যাকে নেম্যানের সর্বোত্তম বণ্টন বলা হয়।

এবং নির্দিষ্ট V_0 এর জন্য

$$n = \frac{\sum_{i=1}^{k} N_i S_i \sqrt{C_0} . \sum_{i=1}^{k} N_i S_i / \sqrt{C_0}}{N^2 V_0 + \sum_{i=1}^{k} N_i S_i^2} = \frac{\left(\sum_{i=1}^{k} N_i S_i\right)^2}{N^2 V_0 + \sum_{i=1}^{k} N_i S_i^2}$$

উপপাদ্য: প্রচলিত সংকেতে প্রমাণ কর যে, $V(\bar{y}_{st})_{opt} \le V(\bar{y}_{st})_{prop} \le V(\bar{y})_{ran}$ **(জা.বি.-১২)**

অথবা, যদি $\frac{1}{N}$ **বা** $\frac{1}{N_i}$ $(i = 1, 2, \cdots, k)$ **নগন্য হয়, তবে দেখাও যে,** $V(\bar{y})_{ran} \ge V(\bar{y}_{st})_{prop} \ge V(\bar{y}_{st})_{opt}$

অথবা যদি $\frac{1}{N_i}$ $(i = 1, 2, \cdots, k)$ **নগন্য হয়, তবে দেখাও যে,** $V(\bar{y}_{st})_{ran} \ge V(\bar{y}_{st})_{prop} \ge V(\bar{y}_{st})_{opt}$ **(জা.বি.-১৬)**

প্রমাণঃ জানি, $V(\bar{y})_{ran} = \frac{N-n}{N}\frac{S^2}{n}$

$$V(\bar{y})_{prop} = \frac{N-n}{N^2 n}\sum N_i S_i^2$$

$$V(\bar{y})_{opt} = \frac{1}{N^2 n}\left(\sum N_i S_i\right)^2 - \frac{1}{N^2}\sum N_i S_i^2$$

জানি, স্তরিত দৈব নমুনায়নে সমগ্রকের গড় বর্গ

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{N_i}\left(Y_{ij} - \bar{Y}\right)^2$$

$$\Rightarrow (N-1)S^2 = \sum_{i=1}^{k}\sum_{j=1}^{N_i}\left(Y_{ij} - \bar{Y}_i + \bar{Y}_i - \bar{Y}\right)^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{N_i}\left[\left(Y_{ij} - \bar{Y}_i\right) + \left(\bar{Y}_i - \bar{Y}\right)\right]^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{N_i}\left(Y_{ij} - \bar{Y}_i\right)^2 + 2\sum_{i=1}^{k}\sum_{j=1}^{N_i}\left(Y_{ij} - \bar{Y}_i\right)\left(\bar{Y}_i - \bar{Y}\right) + \sum_{i=1}^{k}\sum_{j=1}^{N_i}\left(\bar{Y}_i - \bar{Y}\right)^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{N_i}\left(Y_{ij} - \bar{Y}_i\right)^2 + 2\sum_{i=1}^{k}\left(\bar{Y}_i - \bar{Y}\right)\sum_{j=1}^{N_i}\left(Y_{ij} - \bar{Y}_i\right) + \sum_{i=1}^{k}N_i\left(\bar{Y}_i - \bar{Y}\right)^2$$

$$\Rightarrow (N-1)S^2 = \sum_{i=1}^{k}(N_i - 1)S_i^2 + 2\sum_{i=1}^{k}\left(\bar{Y}_i - \bar{Y}\right)\times 0 + \sum_{i=1}^{k}N_i\left(\bar{Y}_i - \bar{Y}\right)^2 \quad \left[\because S_i^2 = \frac{1}{N_i - 1}\sum_{j=1}^{N_i}\left(Y_{ij} - \bar{Y}_i\right)^2\right]$$

$$\Rightarrow N\left(1 - \frac{1}{N}\right)S^2 = \sum_{i=1}^{k}N_i\left(1 - \frac{1}{N_i}\right)S_i^2 + \sum_{i=1}^{k}N_i\left(\bar{Y}_i - \bar{Y}\right)^2$$

$$\Rightarrow NS^2 = \sum_{i=1}^{k}N_i S_i^2 + \sum_{i=1}^{k}N_i\left(\bar{Y}_i - \bar{Y}\right)^2 \quad \left[\text{যেহেতু } \frac{1}{N} \text{ ও } \frac{1}{N_i} \text{ নগন্য}\right]$$

$$\Rightarrow S^2 = \frac{\sum_{i=1}^{k}N_i S_i^2 + \sum_{i=1}^{k}N_i\left(\bar{Y}_i - \bar{Y}\right)^2}{N}$$

এখন, সরল দৈব নমুনা গড়ের ভেদাংক

$$V(\bar{y})_{ran} = \frac{N-n}{N}\frac{S^2}{n}$$

$$= \frac{N-n}{Nn}\left[\frac{\sum_{i=1}^{k}N_i S_i^2 + \sum_{i=1}^{k}N_i\left(\bar{Y}_i - \bar{Y}\right)^2}{N}\right]$$

$$=\frac{N-n}{N^2 n}\sum_{i=1}^{k} N_i S_i^2+\frac{N-n}{N^2 n}\sum_{i=1}^{k} N_i\left(\bar{Y}_i-\bar{Y}\right)^2$$

$$\Rightarrow V(\bar{y})_{ran}=V(\bar{y}_{st})_{prop}+\frac{N-n}{N^2 n}\sum_{i=1}^{k} N_i\left(\bar{Y}_i-\bar{Y}\right)^2 \quad \left[\because V(\bar{y})_{prop}=\frac{N-n}{N^2 n}\sum N_i S_i^2\right]$$

যেহেতু, $\frac{N-n}{N^2 n}\sum_{i=1}^{k} N_i\left(\bar{Y}_i-\bar{Y}\right)^2$ একটি অঋণাত্মক রাশি, সুতরাং

$$V(\bar{y})_{ran}\geq V(\bar{y}_{st})_{prop}$$

$$\Rightarrow V(\bar{y}_{st})_{prop}\leq V(\bar{y})_{ran} \text{ ....(i)}$$

আবার, $V(\bar{y}_{st})_{prop}-V(\bar{y}_{st})_{opt}=\frac{N-n}{N^2 n}\sum N_i S_i^2-\frac{1}{N^2 n}\left(\sum N_i S_i\right)^2+\frac{1}{N^2}\sum N_i S_i^2$

$$=\frac{N}{N^2 n}\sum N_i S_i^2-\frac{n}{N^2 n}\sum N_i S_i^2-\frac{1}{N^2 n}\left(\sum N_i S_i\right)^2+\frac{1}{N^2}\sum N_i S_i^2$$

$$=\frac{1}{Nn}\sum N_i S_i^2-\frac{1}{N^2}\sum N_i S_i^2-\frac{1}{N^2 n}\left(\sum N_i S_i\right)^2+\frac{1}{N^2}\sum N_i S_i^2$$

$$=\frac{1}{Nn}\sum N_i S_i^2-\frac{1}{N^2 n}\left(\sum N_i S_i\right)^2$$

$$=\frac{1}{Nn}\left[\sum N_i S_i^2-\frac{\left(\sum N_i S_i\right)^2}{N}\right]$$

$$\Rightarrow V(\bar{y}_{st})_{prop}-V(\bar{y}_{st})_{opt}=\frac{1}{Nn}\left[\sum N_i\left(S_i-\bar{S}\right)^2\right]$$, যেখানে $\bar{S}=\frac{\sum N_i S_i}{N}$

যেহেতু উপরের সমীকরণের ডানপক্ষ একটি অঋণাত্মক মান, সুতরাং

$$V(\bar{y}_{st})_{prop}-V(\bar{y}_{st})_{opt}\geq 0$$

$$\Rightarrow V(\bar{y}_{st})_{prop}\geq V(\bar{y}_{st})_{opt}$$

$$\Rightarrow V(\bar{y}_{st})_{opt}\leq V(\bar{y}_{st})_{prop} \text{ ....(ii)}$$

(i) ও (ii) নং হতে পাই, $V(\bar{y}_{st})_{opt}\leq V(\bar{y}_{st})_{prop}\leq V(\bar{y})_{ran}$ (প্রমাণিত)

বি. দ্রঃ $\sum f_i\left(x_i-\bar{x}\right)^2$

$$=\sum f_i\left(x_i^2-2x_i\bar{x}+\bar{x}^2\right)$$

$$=\sum f_i x_i^2-2\bar{x}\sum f_i x_i+\bar{x}^2\sum f_i$$

$$= \sum f_i x_i^2 - 2\bar{x}.\ N\bar{x} + N\bar{x}^2$$

$$= \sum f_i x_i^2 - N\bar{x}^2$$

$$= \sum f_i x_i^2 - N\left(\frac{\sum f_i x_i}{N}\right)^2$$

$$= \sum f_i x_i^2 - \frac{\left(\sum f_i x_i\right)^2}{N}$$

$$\therefore \quad \sum f_i \left(x_i - \bar{x}\right)^2 = \sum f_i x_i^2 - \frac{\left(\sum f_i x_i\right)^2}{N}, \text{ যেখানে } \bar{x} = \frac{\sum f_i x_i}{N}$$

অনুরূপভাবে, $\sum N_i \left(S_i - \bar{S}\right)^2 = \sum N_i S_i^2 - \frac{\left(\sum N_i S_i\right)^2}{N}$, যেখানে $\bar{S} = \frac{\sum N_i S_i}{N}$

প্রশ্ন: সরল দৈব নমুনায়ন পদ্ধতির উপর স্তরিত দৈব নমুনায়নের সুবিধাগুলো কি কি?

অথবা, সরল দৈব নমুনায়নের তুলনায় স্তরিত দৈব নমুনায়নের সুবিধাগুলো কি কি? **(জা.বি.-১১)**

অথবা, কোন কোন শর্তে স্তরিত দৈব নমুনায়ন সরল দৈব নমুনায়ন অপেক্ষা পছন্দনীয়? **(জা.বি.-১২, ১৬)**

উত্তর: সরল দৈব নমুনায়নের উপর স্তরিত দৈব নমুনায়নের সুবিধাগুলো নিম্নরূপ-

(ক) সরল দৈব নমুনায়নে নিরূপকের ভেদাংক কমানোর জন্য নমুনার আকার বৃদ্ধি করতে হয়। কিন্তু স্তরিত দৈব নমুনায়নে নমুনার আকার না বাড়িয়েও নিরূপকের ভেদাংক কমানো সম্ভব কারণ এ পদ্ধতিতে অধিক প্রতিনিধিত্বশীল নমুনা সংগ্রহ করা যায়।

(খ) স্তরিত দৈব নমুনায়নে সরল দৈব নমুনায়ন অপেক্ষা অধিক নির্ভরযোগ্য নিরূপক পাওয়া যায়।

(গ) সরল দৈব নমুনায়নের চেয়ে স্তরিত দৈব নমুনায়নে জরিপ পরিচালনা ও বাস্তবায়ন করতে প্রশাসনিক সুবিধা বেশি। স্তরিত দৈব নমুনায়নে কম সময়ে এবং কম খরচে প্রয়োজনীয় তথ্য সংগ্রহ করা যায়।

(ঘ) সমগ্রককে বিভিন্ন স্তরে ভাগ করা হয় বলে প্রতিটি স্তরকে এক একটি উপসমগ্রক বিবেচনা করা যায়। স্তরিত দৈব নমুনায়নে এরূপ উপসমগ্রকগুলোর জন্য আলাদা নিরূপক পাওয়া সম্ভব যা সরল দৈব নমুনায়নে সম্ভব নয়। তথ্য সংগ্রহের ক্ষেত্রেও স্তরিত দৈব নমুনায়নে বিভিন্ন পদ্ধতি ব্যবহার সম্ভব যা সরল দৈব নমুনায়নে সম্ভব নয়।

প্রশ্ন: স্তরিত নমুনায়নের সুবিধা-অসুবিধা বর্ণনা কর (জা.বি.-১১, ১৫)

উত্তর: স্তরিত দৈব নমুনায়নের অসুবিধাসমূহ নিম্নরূপ-

(ক) স্তরিত দৈব নমুনায়নে নমুনার আকার না বাড়িয়ে নিরূপকের ভেদাংক কমানো সম্ভব কারণ এ পদ্ধতিতে অধিক প্রতিনিধিত্বশীল নমুনা সংগ্রহ করা যায়।

(খ) স্তরিত দৈব নমুনায়নে জরিপ পরিচালনা ও বাস্তবায়ন করতে প্রশাসনিক সুবিধা বেশি।

(গ) স্তরিত দৈব নমুনায়নে কম সময়ে এবং কম খরচে প্রয়োজনীয় তথ্য সংগ্রহ করা যায়।

(ঘ) সমগ্রককে বিভিন্ন স্তরে ভাগ করা হয় বলে প্রতিটি স্তরকে এক একটি উপসমগ্রক বিবেচনা করা যায়। স্তরিত দৈব নমুনায়নে এরূপ উপসমগ্রকগুলোর জন্য আলাদা নিরূপক পাওয়া সম্ভব।

স্তরিত দৈব নমুনায়নের অসুবিধাসমূহ নিম্নরূপ-

(ক) সঠিকভাবে সমগ্রককে স্তরিকরণ করা না হলে প্রাপ্ত তথ্য থেকে নির্ভরযোগ্য নিরূপক পাওয়া যায় না।

(খ) কোন স্তর থেকে কি আকারের নমুনা চয়ন করতে হবে তা নির্ধারণ করাও অসুবিধাজনক।

(গ) স্তরিত নমুনায়নের পূর্ণ সুবিধা পেতে হলে স্তরের আকার N_i এর মানগুলো জানা থাকতে হবে যা সব সময়ে সম্ভব নয়।

সংক্ষিপ্ত প্রশ্ন ও উত্তরঃ

(১) স্তরিত দৈব নমুনায়নে নমুনা বণ্টনের পদ্ধতিগুলো কি কি?

উত্তরঃ স্তরিত দৈব নমুনায়নে নমুনা বণ্টনের পদ্ধতিগুলো নিম্নরূপ-

(ক) সমান বণ্টন (Equal allocation)

(খ) সমানুপাতিক বণ্টন (Proportional allocation)

(গ) অপটিমাম বণ্টন (Optimum allocation)

(২) কখন স্তরিত দৈব নমুনায়ন ব্যবহার করা হয়? (জা.বি.-১২)

উত্তরঃ যে সমগ্রক থেকে সাধারণত নমুনা নেয়া হয় তা নানান বৈশিষ্ট্যের ভিত্তিতে অসমমাত্রিক (heterogeneous) হয়। তখন স্তরিত দৈব নমুনায়ন ব্যবহার করে নমুনা নির্বাচনের পূর্বে সমগ্রককে বিভিন্ন স্তরে বিভক্ত করে নমুনা নেয়া হলে তা অধিক প্রতিনিধিত্বশীল হয়, ফলে নিরূপকের ভেদাংক কম হয় এবং প্রাক্কলকের যথার্থতা বৃদ্ধি করা যায়। কারণ প্রাক্কলকের যথার্থতা উহার ভেদাংকের সংগে উল্টাভাবে নির্ভরশীল।

(৩) সমানুপাতিক বণ্টন কি? (জা.বি.-১২)

উত্তরঃ সমানুপাতিক বণ্টনে কোন স্তরের নমুনার আকার, উক্ত স্তরের আকারের সমানুপাতিক হয়। অর্থাৎ

$$n_i \propto N_i \quad (i = 1, 2, \cdots, k)$$

এখানে, $N_i = i -$ তম স্তরের আকার।

$n_i = i -$ তম স্তর থেকে যে নমুনা নেয়া হবে তার আকার।

(৪) কখন নেম্যান বণ্টন ও সমানুপাতিক বণ্টন একই হয়? (জা.বি.-১২)

উত্তরঃ যদি সকল স্তরের ভেদাঙ্ক সমান হয়, তখন নেম্যান বণ্টন ও সমানুপাতিক বণ্টন একই হয়।

(৫) কখন সর্বোত্তম বণ্টন ও নেম্যান বণ্টন একই হয়? (জা.বি.-১১, ১৫)

উত্তর: যদি সকল স্তরের খরচ সমান হয়, তখন সর্বোত্তম বণ্টন ও নেম্যান বণ্টন একই হয়।

(৬) স্তরিত দৈব নমুনায়নের ক্ষেত্রে তথ্যবিশ্বের সমষ্টির নিঝুঁকি নিরূপকটি লিখ (জা.বি.-১২)

উত্তর: স্তরিত দৈব নমুনায়নের ক্ষেত্রে $N\,\bar{y}_{st}$ হলো তথ্যবিশ্ব সমষ্টি Y এর নিঝুকি নিরূপক অর্থাৎ, $\hat{Y}_{st} = N\,\bar{y}_{st}$ ।

(৭) স্তরিত নমুনায়নে বণ্টন সমস্যা বলতে কি বুঝ? উক্ত সমস্যা কিভাবে সমাধান করা যায়? (জা.বি.-১২)

উত্তর: স্তরিত দৈব নমুনার আকারকে স্থির রেখে কোন্ স্তর থেকে কি আকারের নমুনা চয়ন করতে হবে তা নির্ধারণ করার সমস্যাই হল বণ্টন সমস্যা।

উক্ত সমস্যা তিনটি উপায়ে সমাধান করা যায়-

(ক) সমান বণ্টন

(খ) সমানুপাতিক বণ্টন

(গ) সর্বোত্তম বা অপটিমাম বণ্টন

(৮) স্তরিকরণের কারণসমূহ লিখ (জা.বি.-১১)

উত্তর: স্তরিকরণের কারণসমূহ নিম্নরূপ-

(ক) নমুনা নির্বাচনের পূর্বে সমগ্রককে বিভিন্ন স্তরে বিভক্ত করে নমুনা নেয়া হলে তা অধিক প্রতিনিধিত্বশীল হয়,

(খ) নিরূপকের ভেদাংক কম হয় এবং

(গ) প্রাক্কলকের যথার্থতা বৃদ্ধি পায়। উল্লেখ্য প্রাক্কলকের যথার্থতা উহার ভেদাংকের সংগে উল্টাভাবে নির্ভরশীল।

ধারাবাহিক নমুনায়ন (Systematic Sampling)

প্রশ্নঃ ধারাবাহিক নমুনায়ন কাকে বলে? (জা.বি.-১১, ১৫)

উত্তরঃ যে নমুনায়ন পদ্ধতিতে নমুনার প্রথম এককটি শুধু দৈবভাবে চয়ন করা হয় এবং অবশিষ্ট এককগুলো পূর্ব নির্ধারিত নকশা অনুযায়ী ধারাবাহিকভাবে প্রথম নির্বাচিত একক অনুযায়ী চয়ন করা হয়, তাকে ধারাবাহিক নমুনায়ন বলা হয়। এ পদ্ধতিতে তখনই নমুনা চয়ন করা সম্ভব যখন তথ্যবিশ্বের পূর্ণ এবং সঠিক নমুনা কাঠামো জানা থাকে।

প্রশ্নঃ ধারাবাহিক নমুনা নির্বাচন পদ্ধতির বর্ণনা দাও (জা.বি.-১১, ১৫)

অথবা, সরল ধারাবাহিক ও চক্রাকার ধারাবাহিক নমুনা নির্বাচনের পদ্ধতি আলোচনা কর **(জা.বি.-১২)**

উত্তরঃ ধরি, কোন তথ্যবিশ্বে N সংখ্যক একক আছে। প্রথমে এককগুলোকে 1 থেকে N পর্যন্ত সংখ্যা দ্বারা চিহ্নিত করা হয়। উক্ত তথ্যবিশ্ব থেকে n আকারের নমুনা চয়ন করতে হবে। ধরা যাক, $N = nk$, যেখানে k হলো একটি পূর্ণ সংখ্যা যাকে নমুনায়ন ব্যবধান (sampling interval) বলা হয়। এখন n আকারের একটি নমুনা চয়ন করার জন্য প্রথম k একক থেকে একটি একক, দৈব পদ্ধতিতে চয়ন করা হয়। অতঃপর প্রত্যেক k-তম একক চয়ন করা হয়। সুতরাং যদি প্রথম একক i $(i \le k)$-তম নম্বরযুক্ত হয়, তবে নির্বাচিত নমুনা এককগুলো হবে

$$i,\ i+k,\ i+2k,\cdots,i+(n-1)k$$, এখানে i-কে বলা হয় Random start.

এ ধরণের নমুনাকে ধারাবাহিকভাবে নমুনা বলে। আর যে পদ্ধতিতে ধারাবাহিক নমুনা চয়ন করা হয়, তাকে **সরল ধারাবাহিক নমুনায়ন** (linear systematic sampling) বলা হয়।

যদি তথ্যবিশ্বের আকার N, নমুনার আকার n এর সঠিক গুণিতক না হয়, সেক্ষেত্রে k এর মান হবে $\frac{N}{n}$ এর নিকটতম পূর্ণ সংখ্যা। তখন ধারাবাহিক নমুনার আকার n অথবা $(n-1)$ হবে যা নির্ভর করে প্রথম নির্বাচিত এককের উপর। এ অসুবিধা দূর করার জন্য প্রথম নির্বাচিত এককটি প্রথম k-তম একক থেকে চয়ন না করে পুরো তথ্যবিশ্বের একক থেকে অর্থাৎ 1 থেকে N পর্যন্ত একক থেকে চয়ন করতে হবে এবং প্রত্যেক k-তম একক পর পর চয়ন করতে হবে যাতে নমুনার আকার n হয়। এ পদ্ধতিতে নমুনা চয়ন করা হলে তাকে **চক্র-ক্রমিক ধারাবাহিক নমুনায়ন** (circular systematic sampling) বলে। এ ক্ষেত্রে যদি প্রথম নির্বাচিত এককটি i $(i \le N)$ তম নম্বরযুক্ত হয়, তবে নির্বাচিত নমুনা এককগুলো হবে

$i,\ i+jk$, যদি $i+jk \le N$

$i,\ i+jk-N$, যদি $i+jk > N$

এখানে, $j = 1, 2, \cdots, (n-1)$

বি. দ্রঃ সরল ধারাবাহিক নমুনায়নের ক্ষেত্রে

ধরি, $N = nk$. এখানে k হলো একটি পূর্ণ সংখ্যা, যা নিম্নরূপ তিনটি বিষয়কে নির্দেশ করেঃ

(ক) নমুনায়ন ব্যবধান হবে k .

(খ) n আকারের একটি নমুনা চয়ন করার জন্য প্রথম k একক থেকে একটি একক, দৈব পদ্ধতিতে চয়ন করা হয়।

(গ) মোট ধারাবাহিক নমুনায়ন হবে k সংখ্যকটি।

উদাহরণঃ ধরি, তথ্যবিশ্বের আকার $N = 12$, নমুনার আকার $n = 4$

এক্ষেত্রে, $N = nk$

$$\Rightarrow k = \frac{N}{n} = \frac{12}{4} = 3$$

অর্থাৎ নমুনায়ন ব্যবধান, $k = 3$ যা একটি পূর্ণ সংখ্যা।

ধরি, তথ্যবিশ্বের 12 টি একক যথাক্রমে $y_1,\ y_2, \cdots, y_{12}$, তবে সম্ভাব্য নমুনাসমূহ হবে নিম্নরূপ-

প্রথম নির্বাচিত একক	নমুনার নির্বাচিত একক
y_1	$y_1,\ y_4,\ y_7,\ y_{10}$
y_2	$y_2,\ y_5,\ y_8,\ y_{11}$
y_3	$y_3,\ y_6,\ y_9,\ y_{12}$

কিন্তু যদি নমুনানয়ন ব্যবধান k পূর্ণ সংখ্যা না হয় অথবা N যদি n এর পূর্ণ গুণিতক (integral multiple) না হয়, তবে কিছু সমস্যা তৈরি হয়। যেমনঃ

ধরি, তথ্যবিশ্বের আকার $N = 11$, নমুনার আকার $n = 4$

এক্ষেত্রে, $N = nk$

$$\Rightarrow k = \frac{N}{n} = \frac{11}{4} = 2.75 \cong 3$$

অর্থাৎ নমুনায়ন ব্যবধান, $k = 2.75$ কে এর নিকটতম পূর্ণ সংখ্যা 3 ধরা হয়েছে।

ধরি, তথ্যবিশ্বের 11 টি একক যথাক্রমে $y_1,\ y_2, \cdots, y_{11}$, তবে সম্ভাব্য নমুনাসমূহ হবে নিম্নরূপ-

প্রথম নির্বাচিত একক	নমুনার নির্বাচিত একক
y_1	$y_1,\ y_4,\ y_7,\ y_{10}$
y_2	$y_2,\ y_5,\ y_8,\ y_{11}$
y_3	$y_3,\ y_6,\ y_9$

কাজেই N যদি n এর পূর্ণ গুণিতক না হয়, তবে উক্তরূপ সমস্যা চক্রক্রমিক ধারাবাহিক নমুনায়নের (circular systematic sampling) মাধ্যমে সমাধান করা যেতে পারে। এক্ষেত্রে প্রথম নির্বাচিত এককটি প্রথম k-তম একক থেকে চয়ন না করে পুরো তথ্যবিশ্বের একক থেকে অর্থাৎ 1 হতে N এর মধ্যে থেকে চয়ন করতে হবে। যেমনঃ

ধরি, তথ্যবিশ্বের আকার $N = 11$, নমুনার আকার $n = 4$

এক্ষেত্রে, $N = n\,k$

$$\Rightarrow k = \frac{N}{n} = \frac{11}{4} = 2.75 \cong 3$$

ধরি, তথ্যবিশ্বের 11 টি একক যথাক্রমে $y_1,\ y_2, \cdots, y_{11}$, তবে সম্ভাব্য নমুনাসমূহ হবে নিম্নরূপ-

প্রথম নির্বাচিত একক	নমুনার নির্বাচিত একক
y_1	y_1, y_4, y_7, y_{10}
y_2	y_2, y_5, y_8, y_{11}
y_3	y_3, y_6, y_9, y_1
y_4	y_4, y_7, y_{10}, y_2
y_5	y_5, y_8, y_{11}, y_3
y_6	y_6, y_9, y_1, y_4
y_7	y_7, y_{10}, y_2, y_5
y_8	y_8, y_{11}, y_3, y_6
y_9	y_9, y_1, y_4, y_7
y_{10}	y_{10}, y_2, y_5, y_8
y_{11}	y_{11}, y_3, y_6, y_9

NB: We have, $k = \frac{N}{n} = \frac{11}{4} = 2.75 \cong 3$

Now, $i,\ i + j\,k$, if $i + j\,k \le N$; Here $i = 1, 2, \cdots, 11;\ \ j = 1, 2, \cdots, (n-1) = 1, 2, \cdots, (4-1) = 1, 2, 3$

i	j	$i + j\,k$	Selected sample
1	1	$1 + 1 \times 3 = 4$	y_1, y_4, y_7, y_{10}
	2	$1 + 2 \times 3 = 7$	
	3	$1 + 3 \times 3 = 10$	
2	1	$2 + 1 \times 3 = 5$	y_2, y_5, y_8, y_{11}
	2	$2 + 2 \times 3 = 8$	
	3	$2 + 3 \times 3 = 11$	

Again, $i,\ i + j\,k - N$, if $i + j\,k > N$; Here $j = 1, 2, \cdots, (n-1) = 1, 2, \cdots, (4-1) = 1, 2, 3$

i	j	$i + j\,k - N$	Selected sample
9	1	$9 + 1 \times 3 - N(=11) = 1$	y_9, y_1, y_4, y_7
	2	$9 + 2 \times 3 - N(=11) = 4$	

	3	$9+3\times3-N(=11)=7$	
10	1	$10+1\times3-N(=11)=2$	$y_{10},\ y_2,\ y_5,\ y_8$
	2	$10+2\times3-N(=11)=5$	
	3	$10+3\times3-N(=11)=8$	

প্রশ্ন: পুনরাবৃত্ত ধারাবাহিক নমুনায়ন কী? এটি কখন প্রয়োজন হয়? পদ্ধতিটি একটি উদাহরণের সাহায্যে ব্যাখ্যা কর (জা.বি.-১৬)

উত্তর: যদি তথ্যবিশ্বের আকার N, নমুনার আকার n এর সঠিক গুণিতক না হয়, সেক্ষেত্রে k এর মান হবে $\frac{N}{n}$ এর নিকটতম পূর্ণ সংখ্যা। তখন ধারাবাহিক নমুনার আকার n অথবা $(n-1)$ হবে যা নির্ভর করে প্রথম নির্বাচিত এককের উপর।

এ অসুবিধা দূর করার জন্য প্রথম নির্বাচিত এককটি প্রথম k-তম একক থেকে চয়ন না করে পুরো তথ্যবিশ্বের একক থেকে অর্থাৎ 1 থেকে N পর্যন্ত একক থেকে চয়ন করতে হবে এবং প্রত্যেক k-তম একক পর পর চয়ন করতে হবে যাতে নমুনার আকার n হয়। এ পদ্ধতিতে নমুনা চয়ন করা হলে তাকে পুনরাবৃত্ত বা চক্র-ক্রমিক ধারাবাহিক নমুনায়ন (circular systematic sampling) বলে। এ ক্ষেত্রে যদি প্রথম নির্বাচিত এককটি i $(i \le N)$ তম নম্বরযুক্ত হয়, তবে নির্বাচিত নমুনা এককগুলো হবে:

$$i,\ i+jk\ , \text{ যদি } i+jk \le N$$

$$i,\ i+jk-N\ , \text{ যদি } i+jk > N$$

এখানে, $j=1,\ 2,\cdots,(n-1)$

উদাহরণ: ধরি, তথ্যবিশ্বের আকার $N=11$, নমুনার আকার $n=4$

এক্ষেত্রে, $N=nk$

$$\Rightarrow k=\frac{N}{n}=\frac{11}{4}=2.75\cong 3$$

অর্থাৎ নমুনায়ন ব্যবধান, $k=2.75$ কে এর নিকটতম পূর্ণ সংখ্যা 3 ধরা হয়েছে।

ধরি, তথ্যবিশ্বের 11 টি একক যথাক্রমে $y_1,\ y_2,\cdots,y_{11}$, তবে সম্ভাব্য নমুনাসমূহ হবে নিম্নরূপ-

প্রথম নির্বাচিত একক	নমুনার নির্বাচিত একক
y_1	$y_1,\ y_4,\ y_7,\ y_{10}$
y_2	$y_2,\ y_5,\ y_8,\ y_{11}$
y_3	$y_3,\ y_6,\ y_9$

কাজেই N যদি n এর পূর্ণ গুণিতক না হয়, তবে উক্তরূপ সমস্যা চক্রক্রমিক ধারাবাহিক নমুনায়নের (circular systematic sampling) মাধ্যমে সমাধান করা যেতে পারে। এক্ষেত্রে প্রথম নির্বাচিত এককটি প্রথম k-তম একক থেকে চয়ন না করে পুরো তথ্যবিশ্বের একক থেকে অর্থাৎ 1 হতে N এর মধ্যে থেকে চয়ন করতে হবে। যেমন:

ধরি, তথ্যবিশ্বের আকার $N = 11$, নমুনার আকার $n = 4$

এক্ষেত্রে, $N = nk$

$$\Rightarrow k = \frac{N}{n} = \frac{11}{4} = 2.75 \cong 3$$

ধরি, তথ্যবিশ্বের 11 টি একক যথাক্রমে $y_1, y_2, \cdots, y_{11}$, তবে সম্ভাব্য নমুনাসমূহ হবে নিম্নরূপ-

প্রথম নির্বাচিত একক	নমুনার নির্বাচিত একক
y_1	y_1, y_4, y_7, y_{10}
y_2	y_2, y_5, y_8, y_{11}
y_3	y_3, y_6, y_9, y_1
y_4	y_4, y_7, y_{10}, y_2
y_5	y_5, y_8, y_{11}, y_3
y_6	y_6, y_9, y_1, y_4
y_7	y_7, y_{10}, y_2, y_5
y_8	y_8, y_{11}, y_3, y_6
y_9	y_9, y_1, y_4, y_7
y_{10}	y_{10}, y_2, y_5, y_8
y_{11}	y_{11}, y_3, y_6, y_9

প্রশ্ন: অন্যান্য নমুনায়নের তুলনায় ধারাবাহিক নমুনায়নের সুবিধা-অসুবিধাগুলো আলোচনা কর (জা.বি.-১১)

উত্তর: অন্যান্য নমুনায়নের তুলনায় ধারাবাহিক নমুনায়নের সুবিধা-অসুবিধাগুলো নিম্নে আলোচনা করা হল-

সুবিধাসমূহ:

(ক) ধারাবাহিক নমুনায়নে দৈব বা স্তরিত দৈব নমুনায়নের চেয়ে সহজে নমুনা চয়ন করা যায় এবং এতে সময়ও কম লাগে।

(খ) নমুনা কাঠামো যদি হালনাগাদ হয় এবং অসম্পূর্ণ না থাকে, তাহলে এ পদ্ধতির নমুনায়ন সরল দৈব নমুনায়ন অপেক্ষা উত্তম।

(গ) ধারাবাহিক নমুনায়নে মাঠ কর্মী থেকে আরম্ভ করে তত্ত্বাবধান ও বাস্তবায়ন অন্যান্য নমুনায়নের চেয়ে সুষ্ঠু ও সঠিকভাবে পরিচালনা করা সম্ভব।

অসুবিধাসমূহ:

(ক) ধারাবাহিক নমুনায়নের মাধ্যমে নমুনা সংগ্রহ করতে হলে সম্পূর্ণ এবং সঠিক নমুনা কাঠামো জানা প্রয়োজন যা সব সময় জানা থাকে না।

(খ) এ ধরণের নমুনাটি দৈব নয়। ফলে এ নমুনা থেকে প্রাপ্ত প্রাক্কলক তার যথার্থতা হারিয়ে ফেলে।

(গ) যদি তথ্যবিশ্বের এককগুলো নিয়মিত ব্যবধানে পুনঃ পুনঃ ঘটে, তবে ধারাবাহিক নমুনা যথার্থতা হারিয়ে ফেলে এবং সেক্ষেত্রে প্রাক্কলকগুলো নির্ঝুঁকি হয় না।

(ঘ) যদি তথ্যবিশ্বের আকার N, নমুনার আকার n এর সঠিক গুণিতক না হয়, তবে প্রকৃত নমুনার আকার প্রাপ্ত নমুনার আকারের চেয়ে বড় হতে পারে। সেক্ষেত্রে নমুনার গড় তথ্যবিশ্বের গড়ের নির্ঝুঁকি প্রাক্কলক হয় না।

(ঙ) ধারাবাহিক নমুনায়নে প্রাক্কলকের ভেদাংকের নির্ঝুঁকি প্রাক্কলক পাওয়া সম্ভব নয়, যা এ নমুনায়নের একটি বড় অসুবিধা।

বি. দ্রঃ নমুনার ব্যবধান (sampling interval), $k=\frac{N}{n}$, নমুনার ভগ্নাংশ (sampling fraction) $f=\frac{n}{N}$

ধারাবাহিক দৈব নমুনায়নে ব্যবহৃত প্রতীক ও পরিভাষা (Notations and Terminology used in SRS):

ধরি, তথ্যবিশ্বের একক সংখ্যা $=N$

নমুনার একক সংখ্যা $=n$

ধরি, $N=n\times k$, অর্থাৎ তথ্যবিশ্ব থেকে k সংখ্যক ধারাবাহিক নমুনা নেয়া যেতে পারে।

ধরি, $y_{ij}\ (i=1,2,\cdots,k;\ \ j=1,2,\cdots,n)$ হল i-তম ধারাবাহিক নমুনার j-তম একক বা উপাদান।

সুতরাং, i-তম ধারাবাহিক নমুনার গড়, $\bar{y}_{i\cdot}=\frac{1}{n}\sum_{j=1}^{n}y_{ij};\quad i=1,2,\cdots,k$

ধারাবাহিক নমুনায়নে ধারাবাহিক নমুনা গড় $\bar{y}_{i\cdot}$ কে $\bar{y}_{sys}$ দ্বারা চিহ্নিত করা হয়।

তথ্যবিশ্বের গড়, $\bar{y}_{\cdot\cdot}=\frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n}y_{ij}=\frac{1}{nk}\sum_{i=1}^{k}\sum_{j=1}^{n}y_{ij}=\frac{1}{k}\sum_{i=1}^{k}\frac{1}{n}\sum_{j=1}^{n}y_{ij}=\frac{1}{k}\sum_{i=1}^{k}\bar{y}_{i\cdot}$

তথ্যবিশ্বের গড় বর্গ, $S^2=\frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_{\cdot\cdot}\right)^2=\frac{1}{nk-1}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_{\cdot\cdot}\right)^2$

ধারাবাহিক নমুনাগুলোর ভিতরের উপাদানগুলোর বর্গের গড় (mean square among units that lie within the same systematic sample), $S^2_{wsy}=\frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_{i\cdot}\right)^2$

আবার, ধারাবাহিক নমুনার মধ্যের উপাদানগুলোর সংশ্লেষাংক, $\rho_w=\frac{\sum_{i=1}^{k}\sum_{j\neq j'=1}^{n}\left(y_{ij}-\bar{y}_{\cdot\cdot}\right)\left(y_{ij'}-\bar{y}_{\cdot\cdot}\right)}{(n-1)(N-1)S^2}$

ρ_w কে শ্রেণী মধ্য (Intra-class) সংশ্লেষাংকও বলা হয়।

ধারাবাহিক নমুনার ভেদাংক, $V\left(\bar{y}_{sys}\right)=\frac{1}{k}\sum_{i=1}^{k}\left(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot}\right)^2$

বি. দ্রঃ $V\left(\bar{y}_{sys}\right)=V\left(\bar{y}_{i\cdot}\right)$

$$=E\left[\bar{y}_{i\cdot}-E\left(\bar{y}_{i\cdot}\right)\right]^2$$

$$= E\left[\bar{y}_{i.} - \bar{y}_{..}\right]^2 \quad \left[\because E\left(\bar{y}_{i.}\right) = \bar{y}_{..}\right]$$

$$= \sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2 \times \frac{1}{k}$$

$$\therefore\ V\left(\bar{y}_{sys}\right) = \frac{1}{k}\sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

নিম্নে k সংখ্যক ধারাবাহিক নমুনা এবং উহাদের গড় ও সম্ভাবনা দেয়া হল-

নমুনার সংখ্যা	প্রথম নির্বাচিত একক i (Random Start)	নমুনার নির্বাচিত একক	নমুনা গড়	সম্ভাবনা
1	1	$y_1,\ y_{1+k},\cdots,y_{1+jk},\cdots,y_{1+(n-1)k}$	$\bar{y}_{1.}$	$\frac{1}{k}$
2	2	$y_2,\ y_{2+k},\cdots,y_{2+jk},\cdots,y_{2+(n-1)k}$	$\bar{y}_{2.}$	$\frac{1}{k}$
.....	....	....	....	....
i	i	$y_i,\ y_{i+k},\cdots,y_{i+jk},\cdots,y_{i+(n-1)k}$	$\bar{y}_{i.}$	$\frac{1}{k}$
.....	....	....	....	....
k	k	$y_k,\ y_{2k},\cdots,y_{(1+j)k},\cdots,y_{nk}$	$\bar{y}_{k.}$	$\frac{1}{k}$

বি. দ্রঃ উপরোক্ত ছকটি নিম্নরূপভাবেও উপস্থাপন করা যেতে পারে-

নমুনার সংখ্যা	প্রথম নির্বাচিত একক i (Random Start)	নমুনার এককসমূহ	নমুনা গড়	সম্ভাবনা
1	1	$y_{11},\ y_{12},\cdots,y_{1j},\cdots,y_{1n}$	$\bar{y}_{1.}$	$\frac{1}{k}$
2	2	$y_{21},\ y_{22},\cdots,y_{2j},\cdots,y_{2n}$	$\bar{y}_{2.}$	$\frac{1}{k}$
....	....	....	.....	...
i	i	$y_{i1},\ y_{i2},\cdots,y_{ij},\cdots,y_{in}$	$\bar{y}_{i.}$	$\frac{1}{k}$
....	....	....	....	...
k	k	$y_{k1},\ y_{k2},\cdots,y_{kj},\cdots,y_{kn}$	$\bar{y}_{k.}$	$\frac{1}{k}$

উপপাদ্যঃ দেখাও যে, ধারাবাহিক নমুনা গড় তথ্যবিশ্ব গড়ের নির্ঝুঁকি প্রাক্কলক অর্থাৎ $E\left(\bar{y}_{sys}\right) = \bar{y}_{..}$ **এবং**

$$V\left(\bar{y}_{sys}\right) = \frac{N-1}{N}S^2 - \frac{(n-1)k}{N}S^2_{wsy}$$

যেখানে, $S^2_{wsy} = \frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)^2$ **, যা ধারাবাহিক নমুনাগুলোর ভিতরের উপাদানগুলোর বর্গের গড়।**

অথবা, সচরাচর প্রতীকে দেখাও যে, ধারাবাহিক নমুনা গড়, $\bar{y}_{sy}$ সমগ্রক গড় $\bar{y}$ এর নিঝুঁকি প্রাক্কলক এবং $V\left(\bar{y}_{sy}\right)=\frac{1}{N}\left[\left(N-1\right)S^2-k\left(n-1\right){S^2}_{wsy}\right]$ যেখানে, ${S^2}_{wsy}=\frac{1}{k\left(n-1\right)}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_{i\cdot}\right)^2$ **জা.বি.-১৬**

অথবা, ধারাবাহিক নমুনায়নের ক্ষেত্রে সমগ্রক গড় প্রাক্কলনের পদ্ধতিটি বর্ণনা কর। এর ভেদাংক বের কর **(জা.বি.-১১)**

অথবা, সচরাচর ব্যবহৃত প্রতীকে দেখাও যে, $V\left(\bar{y}_{sys}\right)=\frac{N-1}{N}S^2-\frac{k\left(n-1\right)}{N}S^2_{wsy}$ **(জা.বি.-১১)**

অথবা, ধারাবাহিক নমুনায়নের ক্ষেত্রে সমগ্রক গড়ের নিঝুঁকি প্রাক্কলক নির্ণয় কর **(জা.বি.-১৫)**

প্রমাণ: ধরি, তথ্যবিশ্বের একক সংখ্যা $=N$

নমুনার একক সংখ্যা $=n$

অতএব, $N=nk$ (k = নমুনায়ন ব্যবধান)

অর্থাৎ তথ্যবিশ্ব থেকে k সংখ্যক ধারাবাহিক নমুনা নেয়া যেতে পারে।

ধরি, y_{ij} $\left(i=1,2,\cdots,k;\quad j=1,2,\cdots,n\right)$ হল i-তম ধারাবাহিক নমুনার j-তম একক বা উপাদান।

সুতরাং, i-তম ধারাবাহিক নমুনার গড়, $\bar{y}_{i\cdot}=\frac{1}{n}\sum_{j=1}^{n}y_{ij};\quad i=1,2,\cdots,k$

তথ্যবিশ্বের গড়, $\bar{y}_{\cdot\cdot}=\frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n}y_{ij}=\frac{1}{nk}\sum_{i=1}^{k}\sum_{j=1}^{n}y_{ij}=\frac{1}{k}\sum_{i=1}^{k}\frac{1}{n}\sum_{j=1}^{n}y_{ij}=\frac{1}{k}\sum_{i=1}^{k}\bar{y}_{i\cdot}$

তথ্যবিশ্বের গড় বর্গ, $S^2=\frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_{\cdot\cdot}\right)^2=\frac{1}{nk-1}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_{\cdot\cdot}\right)^2$

যেহেতু নমুনার প্রথম এককটি তথ্যবিশ্বের প্রথম k সংখ্যক একক হতে নির্বাচিত হয়, তাই প্রথম এককটি নির্বাচিত হবার সম্ভাবনা $\frac{1}{k}$। অতএব, ধারাবাহিক নমুনায়নে k সংখ্যক ধারাবাহিক নমুনার যে কোন একটি নমুনা নির্বাচিত হবার সম্ভাবনা $\frac{1}{k}$.

সুতরাং, $E\left(\bar{y}_{sys}\right)=E\left(\bar{y}_{i\cdot}\right)$

$$=\sum_{i=1}^{k}\bar{y}_{i\cdot}\ P\left(\bar{y}_{i\cdot}\right)$$

$$=\left(\bar{y}_{1\cdot}\times\frac{1}{k}\right)+\left(\bar{y}_{2\cdot}\times\frac{1}{k}\right)+\cdots+\left(\bar{y}_{k\cdot}\times\frac{1}{k}\right)$$

$$=\frac{1}{k}\left(\bar{y}_{1\cdot}+\bar{y}_{2\cdot}+\cdots+\bar{y}_{k\cdot}\right)$$

$$=\frac{1}{k}\sum_{i=1}^{k}\bar{y}_{i\cdot}$$

$$\therefore \quad E\left(\bar{y}_{sys}\right) = \bar{y}_{..}$$

অর্থাৎ ধারাবাহিক নমুনা গড় তথ্যবিশ্ব গড়ের নিঝুঁকি প্রাক্কলক।

সংজ্ঞানুযায়ী ধারাবাহিক নমুনা গড়ের ভেদাংক হল,

$$V\left(\bar{y}_{sys}\right) = V\left(\bar{y}_{i.}\right)$$

$$= E\left[\bar{y}_{i.} - E\left(\bar{y}_{i.}\right)\right]^2$$

$$= E\left[\bar{y}_{i.} - \bar{y}_{..}\right]^2$$

$$= \sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2 \times \frac{1}{k}$$

$$\Rightarrow V\left(\bar{y}_{sys}\right) = \frac{1}{k}\sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$$\Rightarrow k\ V\left(\bar{y}_{sys}\right) = \sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2 \ \ldots\ldots\ldots\ldots(i)$$

জানি, $S^2 = \dfrac{1}{N-1}\displaystyle\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{..}\right)^2$

$$\Rightarrow (N-1)S^2 = \sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{..}\right)^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.} + \bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n}\left[\left(y_{ij} - \bar{y}_{i.}\right) + \left(\bar{y}_{i.} - \bar{y}_{..}\right)\right]^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)^2 + 2\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)\left(\bar{y}_{i.} - \bar{y}_{..}\right) + \sum_{i=1}^{k}\sum_{j=1}^{n}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)^2 + 2\sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right) + \sum_{i=1}^{k} n\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)^2 + 2\sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right) \times 0 + n\sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$\Rightarrow (N-1)S^2 = k(n-1){S^2}_{wsy} + nk\ V\left(\bar{y}_{sys}\right)$ [(i) নং হতে মান বসিয়ে]

$$\left[\because \quad {S^2}_{wsy} = \frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)^2\right]$$

$$\Rightarrow nk\, V\left(\bar{y}_{sys}\right) = (N-1)S^2 - k(n-1){S^2}_{wsy}$$

$$\Rightarrow V\left(\bar{y}_{sys}\right) = \frac{N-1}{nk}S^2 - \frac{k(n-1)}{nk}{S^2}_{wsy}$$

$$\Rightarrow V\left(\bar{y}_{sys}\right) = \frac{N-1}{N}S^2 - \frac{(n-1)k}{N}{S^2}_{wsy}$$

$$\therefore\ V\left(\bar{y}_{sy}\right) = \frac{1}{N}\left[(N-1)S^2 - k(n-1){S^2}_{wsy}\right]$$ (প্রমাণিত)

বি. দ্রঃ ${S^2}_{wsy} =$ mean square among units that lie **within the same systematic sample**.

উপপাদ্যঃ দেখাও যে, ধারাবাহিক নমুনায়ন সরল দৈব নমুনায়নের চেয়ে ভাল প্রাক্কলক প্রদান করবে, যদি ${S^2}_{wsy} > S^2$ হয় (প্রতীকগুলো চিরাচরিত) জা.বি.-১২, ১৬

অথবা, ধারাবাহিক নমুনা গড় $\bar{y}_{sys}$, দৈব নমুনা গড় $\bar{y}$ এর চেয়ে অধিক যথার্থ (precise) হবে যদি ${S^2}_{wsy} > S^2$ হয় অর্থাৎ, ধারাবাহিক নমুনা গড়ের ভেদাংক, সরল দৈব নমুনা গড়ের ভেদাংকের চেয়ে ছোট হবে যদি ${S^2}_{wsy} > S^2$ হয়।

অথবা, ধারাবাহিক নমুনায়নে নমুনা গড়ের ভেদাংকের সাথে সরল দৈব নমুনায়নে নমুনা গড়ের ভেদাংকের তুলনা কর **(জা.বি.-১৪)**

অথবা, দেখাও যে নির্দিষ্ট শর্ত সাপেক্ষে ধারাবাহিক নমুনায়ন সরল দৈব নমুনায়নের চেয়ে বেশি সঠিক **(জা.বি.-১৫)**

প্রমাণঃ জানি, সরল দৈব নমুনা গড় $\bar{y}$ এর ভেদাংক, $V(\bar{y}) = \frac{N-n}{N}\frac{S^2}{n}$

অন্যদিকে, ধারাবাহিক নমুনা গড় $\bar{y}_{sys}$ এর ভেদাংক, $V\left(\bar{y}_{sys}\right) = \frac{N-1}{N}S^2 - \frac{(n-1)k}{N}{S^2}_{wsy}$

এখন, $V(\bar{y}) - V\left(\bar{y}_{sys}\right) = \frac{N-n}{N}\frac{S^2}{n} - \frac{N-1}{N}S^2 + \frac{(n-1)k}{N}{S^2}_{wsy}$

$$= \frac{S^2}{N}\left(\frac{N-n}{n} - N + 1\right) + \frac{(n-1)k}{nk}{S^2}_{wsy}$$

$$= \frac{S^2}{N}\left(\frac{N-n-nN+n}{n}\right) + \frac{n-1}{n}{S^2}_{wsy}$$

$$= \frac{S^2}{N}\left(\frac{N-nN}{n}\right) + \frac{n-1}{n}{S^2}_{wsy}$$

$$= \frac{S^2}{N}\,\frac{N(1-n)}{n} + \frac{n-1}{n}{S^2}_{wsy}$$

$$= -\frac{S^2(n-1)}{n} + \frac{n-1}{n} S^2{}_{wsy}$$

$$= \frac{n-1}{n} S^2{}_{wsy} - \frac{n-1}{n} S^2$$

$$\Rightarrow V(\bar{y}) - V(\bar{y}_{sys}) = \frac{n-1}{n}\left[S^2{}_{wsy} - S^2\right]$$

ধারাবাহিক নমুনা গড়, সরল দৈব নমুনা গড়ের চেয়ে অধিক যথার্থ হবে, যদি

$$V(\bar{y}) - V(\bar{y}_{sys}) > 0$$

$$\Rightarrow \frac{n-1}{n}\left[S^2{}_{wsy} - S^2\right] > 0$$

$$\Rightarrow S^2{}_{wsy} - S^2 > 0$$

$$\therefore \quad S^2{}_{wsy} > S^2$$

ইহা হতে এ সিদ্ধান্ত নেয়া যায় যে, ধারাবাহিক নমুনায়ন সরল দৈব নমুনায়নের চেয়ে অধিক যথার্থ হবে যদি ধারাবাহিক নমুনার ভিতরের উপাদানগুলোর ভেদাংক তথ্যবিশ্বের ভেদাংকের চেয়ে বড় হয়। অর্থাৎ ধারাবাহিক নমুনা তখনই ভাল প্রাক্কলক দেয় যখন নমুনার ভেতরের উপাদানগুলোর অধিক অসাদৃশ হয়।

উপপাদ্য: ধারাবাহিক নমুনায়নে আন্তঃশ্রেণি সংশ্লেষাংক ρ_w হলে দেখাও যে,

$$V(\bar{y}_{sy}) = \frac{N-1}{N}.\frac{S^2}{n}\left[1+(n-1)\rho_w\right]$$ **জা.বি.-১২**

অথবা, দেখাও যে, $V(\bar{y}_{sys}) = \frac{(N-1)S^2}{Nn}\left[1+(n-1)\rho_w\right]$**; যেখানে** ρ_w **হলো আন্তঃশ্রেণি সংশ্লেষাংক (জা.বি.-১৫)**

অথবা, ধারাবাহিক নমুনা গড়ের ভেদাংক হল

$$V(\bar{y}_{sys}) = \frac{S^2}{n}\left(\frac{N-1}{N}\right)\left[1+(n-1)\rho_w\right]$$

যেখানে ρ_w হল ধারাবাহিক নমুনার মধ্যের উপাদানগুলোর সংশ্লেষাংক যাকে শ্রেণী মধ্য (Intra-class) সংশ্লেষাংক বলা হয়। ρ_w কে নিম্নরূপে সংজ্ঞায়িত করা হয়-

$$\rho_w = \frac{\sum_{i=1}^{k}\sum_{j\neq j'=1}^{n}\left(y_{ij}-\bar{y}_{..}\right)\left(y_{ij'}-\bar{y}_{..}\right)}{(n-1)(N-1)S^2}$$

প্রমাণ: জানি, $V(\bar{y}_{sys}) = \frac{1}{k}\sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$

$$= \frac{1}{k} \sum_{i=1}^{k} \left(\frac{1}{n} \sum_{j=1}^{n} y_{ij} - \bar{y}_{..} \right)^2$$

$$= \frac{1}{k} \sum_{i=1}^{k} \left(\frac{\sum_{j=1}^{n} y_{ij} - n\,\bar{y}_{..}}{n} \right)^2$$

$$= \frac{1}{n^2 k} \sum_{i=1}^{k} \left(\sum_{j=1}^{n} y_{ij} - n\,\bar{y}_{..} \right)^2$$

$$\Rightarrow n^2 k\ V\left(\bar{y}_{sys}\right) = \sum_{i=1}^{k} \left(\sum_{j=1}^{n} y_{ij} - \sum_{j=1}^{n} \bar{y}_{..} \right)^2$$

$$= \sum_{i=1}^{k} \left[\sum_{j=1}^{n} \left(y_{ij} - \bar{y}_{..} \right) \right]^2$$

$$= \sum_{i=1}^{k} \left[\sum_{j=1}^{n} \left(y_{ij} - \bar{y}_{..} \right)^2 + \sum_{j \neq j'=1}^{n} \left(y_{ij} - \bar{y}_{..} \right)\left(y_{ij'} - \bar{y}_{..} \right) \right]$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{n} \left(y_{ij} - \bar{y}_{..} \right)^2 + \sum_{i=1}^{k} \sum_{j \neq j'=1}^{n} \left(y_{ij} - \bar{y}_{..} \right)\left(y_{ij'} - \bar{y}_{..} \right)$$

$$\Rightarrow n^2 k\ V\left(\bar{y}_{sys}\right) = (N-1)S^2 + \rho_w (n-1)(N-1)S^2$$

$$\left[\because\ S^2 = \frac{1}{N-1} \sum_{i=1}^{k} \sum_{j=1}^{n} \left(y_{ij} - \bar{y}_{..} \right)^2; \quad \rho_w = \frac{\sum_{i=1}^{k} \sum_{j \neq j'=1}^{n} \left(y_{ij} - \bar{y}_{..} \right)\left(y_{ij'} - \bar{y}_{..} \right)}{(n-1)(N-1)S^2} \right]$$

$$\Rightarrow V\left(\bar{y}_{sys}\right) = \frac{(N-1)\ S^2}{n^2 k} + \frac{(n-1)\ (N-1)\ S^2}{n^2 k}\ \rho_w$$

$$= \frac{S^2}{n} \cdot \frac{N-1}{nk} + \frac{S^2}{n} \cdot \frac{(n-1)\ (N-1)}{nk}\ \rho_w$$

$$= \frac{S^2}{n} \cdot \frac{N-1}{N} + \frac{S^2}{n} \cdot \frac{(n-1)(N-1)}{N} \rho_w \qquad [\because\ N = nk]$$

$$\therefore\ V\left(\bar{y}_{sys}\right) = \frac{S^2}{n} \cdot \frac{N-1}{N} \left[1 + (n-1)\rho_w \right]$$

$$= \frac{N-1}{N} \cdot \frac{S^2}{n} \left[1 + (n-1)\rho_w \right]$$ (প্রমাণিত)

বি. দ্রঃ উক্ত সম্পর্ক থেকে নিম্নলিখিত সিদ্ধান্তে উপনীত হওয়া যায়-

(ক) ধারাবাহিক নমুনা গড় $\bar{y}_{sys}$ ভাল প্রাক্কলক হবে, যখন ρ_w এর মান সবচেয়ে ছোট হবে। অর্থাৎ ρ_w এর সবচেয়ে বড় ঋণাত্মক মানের জন্য $\bar{y}_{sys}$ এর ভেদাংকের মান ছোট হবে। সেক্ষেত্রে $\bar{y}_{sys}$ অধিক যথার্থ হবে।

(খ) যেহেতু ρ_w এর গুণিতক $(n-1)$ । তাই ρ_w এর ছোট ধনাত্মক মানের জন্যই $\bar{y}_{sys}$ এর বেশ বড় মান হবে।

(গ) জানি, $V\left(\bar{y}_{sys}\right) \geq 0$

অতএব, $1+(n-1)\rho_w \geq 0$

$$\Rightarrow \rho_w \geq -\frac{1}{n-1}$$

সুতরাং ρ_w এর সবচেয়ে ছোট মান হল $-\dfrac{1}{n-1}$, যখন $V\left(\bar{y}_{sys}\right)=0$ ।

উপপাদ্যঃ সরল দৈব নমুনার তুলনায় ধারাবাহিক নমুনার দক্ষতা যাচাই কর।

অথবা, ধারাবাহিক নমুনায়নে নমুনা গড়ের ভেদাংকের সাথে সরল দৈব নমুনায়নে নমুনা গড়ের ভেদাংকের তুলনা কর (জা.বি.-১৪)

উত্তরঃ জানি, $V(\bar{y})=\dfrac{N-n}{Nn}S^2=\dfrac{nk-n}{nk.n}S^2=\dfrac{nk-n}{n^2k}S^2$

$$V\left(\bar{y}_{sys}\right)=\frac{N-1}{N}.\frac{S^2}{n}\left[1+(n-1)\rho_w\right]$$

$$=\frac{nk-1}{nk}.\frac{S^2}{n}\left[1+(n-1)\rho_w\right]$$

সুতরাং সরল দৈব নমুনার গড় $\bar{y}$ এর সংগে ধারাবাহিক নমুনার গড় $\bar{y}_{sys}$ এর দক্ষতাকে নিম্নরূপে সংজ্ঞায়িত করা যায়-

$$E=\frac{V(\bar{y})}{V\left(\bar{y}_{sys}\right)}$$

যখন $E>1$ অর্থাৎ $\dfrac{V(\bar{y})}{V\left(\bar{y}_{sys}\right)}>1$

$$\Rightarrow V(\bar{y})>V\left(\bar{y}_{sys}\right)$$

$$\Rightarrow \frac{nk-n}{n^2k}S^2>\frac{nk-1}{nk}.\frac{S^2}{n}\left[1+(n-1)\,\rho_w\right]$$

$$\Rightarrow nk-n>\left(nk-1\right)\,\left[1+(n-1)\,\rho_w\right]$$

$$\Rightarrow nk-n>nk-1+\left(nk-1\right)(n-1)\,\rho_w$$

$$\Rightarrow nk-n-nk+1>\left(nk-1\right)(n-1)\,\rho_w$$

$$\Rightarrow -n+1 > (nk-1)\ (n-1)\ \rho_w$$

$$\Rightarrow -(n-1) > (nk-1)(n-1)\ \rho_w$$

$$\Rightarrow -1 > (nk-1)\ \rho_w$$

$$\Rightarrow (nk-1)\ \rho_w < -1$$

$$\Rightarrow \rho_w < -\frac{1}{nk-1}$$

অর্থাৎ সরল দৈব নমুনার তুলনায় ধারাবাহিক নমুনা তখনই বেশি দক্ষ (অর্থাৎ ভেদাংক কম) হবে যখন $\rho_w < -\dfrac{1}{nk-1}$ হয়।

অপরপক্ষে, যদি $\rho_w > -\dfrac{1}{nk-1}$ হয় তবে সরল দৈব নমুনা, ধারাবাহিক নমুনার চেয়ে বেশি দক্ষ হয়।

মন্তব্য:

(ক) যদি ρ_w তার সবচেয়ে ছোট মান নেয় অর্থাৎ যদি $\rho_w = -\dfrac{1}{n-1}$ হয় তবে $V(\bar{y}_{sys}) = 0$ হয়। সেক্ষেত্রে $E = \infty$. তখন দৈব নমুনার তুলনায় ধারাবাহিক নমুনার দক্ষতা 100%.

(খ) অন্যদিকে যদি ρ_w তার সবচেয়ে বড় মান নেয়, অর্থাৎ $\rho_w = 1$ হয় তবে $E = \dfrac{k-1}{nk-1} = \dfrac{k-1}{N-1}$. এক্ষেত্রে ধারাবাহিক নমুনার দক্ষতা সরল দৈব নমুনার তুলনায় সবচেয়ে কম।

প্রশ্ন: স্তরিত দৈব নমুনার তুলনায় ধারাবাহিক নমুনার দক্ষতা যাচাই কর।

উত্তর: ধরি, N আকারের তথ্যবিশ্ব থেকে n আকারের k সংখ্যক ধারাবাহিক নমুনা সংগ্রহ করা হলো। k ধারাবাহিক নমুনার n সংখ্যক কলামকে n স্তর হিসেবে বিবেচনা করা যায়। ধরি, প্রতিটি স্তর থেকে এক একক বিশিষ্ট স্তরিত দৈব নমুনা সংগ্রহ করা হলো, যা থেকে n আকারের একটি স্তরিত দৈব নমুনা পাওয়া গেল।

নিম্নে k সংখ্যক ধারাবাহিক নমুনা দেয়া হল-

নমুনার সংখ্যা	প্রথম নির্বাচিত একক	নমুনার এককসমূহ (n সংখ্যক স্তর)					
1	1	y_{11}	y_{12}	...	y_{1j}	...	y_{1n}
2	2	y_{21}	y_{22}	...	y_{2j}	...	y_{2n}
...	...	...	...	...	...	....	...
i	i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{in}
...	...	...	...	...	...	...	...
k	k	y_{k1}	y_{k2}	...	y_{kj}	...	y_{kn}
স্তরের গড়		$\bar{y}_{\bullet 1}$	$\bar{y}_{\bullet 2}$	...	$\bar{y}_{\bullet j}$	...	$\bar{y}_{\bullet n}$

এখানে, $N = nk$

অতএব, $i \ \ (i = 1, 2, \cdots, k)$ -তম ধারাবাহিক নমুনার গড়, $\bar{y}_{i\cdot} = \dfrac{1}{n}\sum_{j=1}^{n} y_{ij}$

এবং $j \ \ (j = 1, 2, \cdots, n)$ -তম স্তরের গড়, $\bar{y}_{\cdot j} = \dfrac{1}{k}\sum_{i=1}^{k} y_{ij}$

তথ্যবিশ্ব গড়, $\bar{y}_{\cdot\cdot} = \dfrac{1}{nk}\sum_{i=1}^{k}\sum_{j=1}^{n} y_{ij} = \dfrac{1}{k}\sum_{i=1}^{k}\bar{y}_{i\cdot} = \dfrac{1}{n}\sum_{j=1}^{n}\bar{y}_{\cdot j}$

স্তরের গড় বর্গ, $S_j^2 = \dfrac{1}{N_j - 1}\sum_{i=1}^{k}\left(y_{ij} - \bar{y}_{\cdot j}\right)^2$

$$\Rightarrow S_j^2 = \frac{1}{k-1}\sum_{i=1}^{k}\left(y_{ij} - \bar{y}_{\cdot j}\right)^2 \ \ldots\ldots\ldots\ldots\text{(i)}$$

$$\left[\because\ N_j = k; \quad j = 1, 2, \cdots, n\right]$$

আবার স্তরের অভ্যন্তরের এককগুলোর সমন্বিত গড় বর্গ,

$$S^2_{wst} = \frac{1}{n(k-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{\cdot j}\right)^2 \ \ldots\ldots\ldots\ldots\text{(ii)}$$

[$\because$ ধারাবাহিক নমুনাগুলোর ভিতরের উপাদানগুলোর বর্গের গড় $S^2_{wsy} = \dfrac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i\cdot}\right)^2$]

অপরপক্ষে, একই ধারাবাহিক নমুনার জোড়া জোড়া মানগুলো থেকে সেই স্তরের গড়ের ব্যবধানের সংশ্লেষাংক ρ_{wst} হলে

$$\rho_{wst} = \frac{\sum_{i=1}^{k}\sum_{j \neq j'=1}^{k}\left(y_{ij} - \bar{y}_{\cdot j}\right)\left(y_{ij'} - \bar{y}_{\cdot j'}\right)}{(n-1)\,n\,(k-1)\,S^2_{wst}} \ \ldots\ldots\ldots\ldots\text{(iii)}$$

এখন, $V\left(\bar{y}_{sys}\right) = \dfrac{1}{k}\sum_{i=1}^{k}\left(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}\right)^2$

$$= \frac{1}{k}\sum_{i=1}^{k}\left[\frac{1}{n}\sum_{j=1}^{n} y_{ij} - \frac{1}{n}\sum_{j=1}^{n}\bar{y}_{\cdot j}\right]^2$$

$$= \frac{1}{n^2 k}\sum_{i=1}^{k}\left[\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{\cdot j}\right)\right]^2$$

$$= \frac{1}{n^2 k}\sum_{i=1}^{k}\left[\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{\cdot j}\right)^2 + \sum_{j \neq j'=1}^{n}\left(y_{ij} - \bar{y}_{\cdot j}\right)\left(y_{ij'} - \bar{y}_{\cdot j'}\right)\right]$$

$$= \frac{1}{n^2 k}\left[\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\overline{y}_{\cdot j}\right)^2+\sum_{i=1}^{k}\sum_{j\neq j'=1}^{n}\left(y_{ij}-\overline{y}_{\cdot j}\right)\left(y_{ij'}-\overline{y}_{\cdot j'}\right)\right]$$

$$= \frac{1}{n^2 k}\left[n(k-1)S^2_{wst}+n(n-1)(k-1)S^2_{wst}\ \rho_{wst}\right] \text{ [(ii) ও (iii) নং হতে]}$$

$$= \frac{n(k-1)}{n^2 k}\ S^2_{wst}\left[1+(n-1)\ \rho_{wst}\right]$$

$$\Rightarrow V\left(\overline{y}_{sys}\right)=\frac{k-1}{nk}S^2_{wst}\ \left[1+(n-1)\ \rho_{wst}\right]$$

আবার, $V\left(\overline{y}_{st}\right)=\dfrac{1}{N^2}\displaystyle\sum_{j=1}^{n}N_j\ \left(N_j-n_j\right)\dfrac{S_j^{\,2}}{n_j}$

কিন্তু, এখানে $N_j = k$ এবং $n_j = 1$

সুতরাং $V\left(\overline{y}_{st}\right)=\dfrac{1}{N^2}\displaystyle\sum_{j=1}^{n}k\,(k-1)\ S_j^{\,2}$

$$= \frac{k(k-1)}{n^2 k^2}\ \sum_{j=1}^{n}S_j^{\,2}$$

$$= \frac{k-1}{n^2 k}\ \sum_{j=1}^{n}\left[\frac{1}{k-1}\sum_{i=1}^{k}\left(y_{ij}-\overline{y}_{\cdot j}\right)^2\right] \quad \text{[(i) নং হতে]}$$

$$= \frac{1}{n^2 k}\ \sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\overline{y}_{\cdot j}\right)^2$$

$$= \frac{1}{n^2 k}\times n\,(k-1)S^2_{wst} \quad \text{[(ii) নং হতে]}$$

$$\Rightarrow V\left(\overline{y}_{st}\right)=\frac{k-1}{n\,k}S^2_{wst}$$

সুতরাং স্তরিত দৈব নমুনা গড় $\overline{y}_{st}$ এর সংগে ধারাবাহিক নমুনার গড় $\overline{y}_{sys}$ এর দক্ষতাকে নিম্নরূপে সংজ্ঞায়িত করা যায়-

$$E=\frac{V\left(\overline{y}_{st}\right)}{V\left(\overline{y}_{sys}\right)}$$

যখন $E>1$, অর্থাৎ $\dfrac{V\left(\overline{y}_{st}\right)}{V\left(\overline{y}_{sys}\right)}>1$

$$\Rightarrow V\left(\overline{y}_{st}\right)>V\left(\overline{y}_{sys}\right)$$

$$\Rightarrow \frac{k-1}{nk} S^2{}_{wst} > \frac{k-1}{nk} S^2{}_{wst} \left[1+(n-1)\ \rho_{wst}\right]$$

$$\Rightarrow 1 > 1+(n-1)\ \rho_{wst}$$

$$\Rightarrow 0 > (n-1)\ \rho_{wst}$$

$$\Rightarrow (n-1)\ \rho_{wst} < 0$$

$$\therefore\ \rho_{wst} < 0$$

অর্থাৎ $\rho_{wst} < 0$ হলে ধারাবাহিক নমুনায়ন, স্তরিত দৈব নমুনায়নের চেয়ে অধিক দক্ষ হবে।

আবার, যদি $E < 1$ হয় তখন $\rho_{wst} > 0$, তবে স্তরিত দৈব নমুনায়ন, ধারাবাহিক নমুনায়নের চেয়ে বেশি দক্ষ হবে।

অন্যদিকে, যদি $\rho_{wst} = 0$ হয়, তবে উভয় নমুনায়নের দক্ষতা সমান হবে।

NB: জানি, স্তরিত নমুনায়নের ক্ষেত্রে সমগ্রক গড়ের প্রাক্কলকের ভেদাংক

$$V(\bar{y}_{st}) = \frac{1}{N^2} \sum_{i=1}^{k} N_i\ (N_i - n_i) \frac{S_i^2}{n_i}$$

উপপাদ্যঃ সরল রৈখিক গতিধারাবিশিষ্ট সমগ্রক হতে নমুনায়নের ক্ষেত্রে দেখাও যে, $V(\bar{y}_{st}) < V(\bar{y}_{sys}) < V(\bar{y}_{ran})$ **জা.বি.-১১**

অথবা, সরলরৈখিক গতিধারাবিশিষ্ট সমগ্রক হতে নমুনায়নের ক্ষেত্রে দেখাও যে, $V(\bar{y}_{st}) \le V(\bar{y}_{sys}) \le V(\bar{y}_{ran})$ **জা.বি.-১৫**

অথবা, যদি তথ্যবিশ্বের মানগুলো সরলরৈখিক ধারা (linear trend) অনুসরণ করে, তবে প্রমাণ কর যে

$$V(\bar{y}_{st}) \le V(\bar{y}_{sys}) \le V(\bar{y}_{ran})$$

প্রমাণঃ ধরি, তথ্যবিশ্বের মানগুলো নিম্নের সরলরৈখিক ধারা অনুসরণ করে।

$$Y_i = 1,\ 2, \cdots, N$$

অর্থাৎ $Y_i = i \quad (i = 1,\ 2, \cdots, N)$

তথ্যবিশ্বের গড়, $\overline{Y} = \bar{y}_{..}$

$$= \frac{1}{N} \sum_{i=1}^{N} Y_i$$

$$= \frac{1}{N} \sum_{i=1}^{N} i$$

$$= \frac{1}{N} \times \frac{N(N+1)}{2}$$

$$=\frac{N+1}{2}=\frac{nk+1}{2} \text{ ....(i)}$$

$$S^2=\frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i-\overline{Y}\right)^2$$

$$=\frac{1}{N-1}\left(\sum_{i=1}^{N}Y_i^2-N\overline{Y}^2\right)$$

$$=\frac{1}{N-1}\left(\sum_{i=1}^{N}i^2-N\overline{Y}^2\right)$$

$$=\frac{1}{N-1}\left[\frac{N(N+1)(2N+1)}{6}-N.\frac{(N+1)^2}{4}\right]$$

$$=\frac{N(N+1)}{N-1}\left[\frac{2N+1}{6}-\frac{N+1}{4}\right]$$

$$=\frac{N(N+1)}{N-1}\left[\frac{4N+2-3N-3}{12}\right]$$

$$=\frac{N(N+1)}{N-1}\times\frac{N-1}{12}$$

$$\Rightarrow S^2=\frac{N(N+1)}{12} \text{ ....(ii)}$$

$$\therefore\ S^2=\frac{nk(nk+1)}{12}$$

এখন, $V\left(\bar{y}_{ran}\right)=\frac{N-n}{N}\frac{S^2}{n}$

$$=\frac{nk-n}{nk}\times\frac{nk(nk+1)}{12n}$$

$$=\frac{n(k-1)\ (nk+1)}{12n}$$

$$=\frac{(k-1)\ (nk+1)}{12} \text{ ....(iii)}$$

যেহেতু ধারাবাহিক নমুনার সাথে তুলনা করতে হবে তাই $N=nk$ বসানো হয়েছে।

আবার, $V\left(\bar{y}_{st}\right)=\frac{1}{N^2}\sum_{j=1}^{n}N_j\left(N_j-n_j\right)\frac{S_j^2}{n_j}$

এখানে $N_j=k$ এবং $n_j=1$

সুতরাং $V\left(\bar{y}_{st}\right)=\dfrac{1}{N^2}\sum_{j=1}^{n}k(k-1)S_j^{\,2}$

$$=\frac{k(k-1)}{n^2k^2}\sum_{j=1}^{n}S_j^{\,2}$$

$$=\frac{k-1}{n^2k}\sum_{j=1}^{n}S_j^{\,2}$$

(ii) নং হতে পাই, $S^2=\dfrac{N(N+1)}{12}$ । কিন্তু ধারাবাহিক নমুনায়নকে n স্তর বিশিষ্ট স্তরিত দৈব নমুনা বিবেচনা করা যেতে পারে, যেখানে প্রতি স্তরের একক সংখ্যা k । অতএব, $S_j^{\,2}=\dfrac{k(k+1)}{12}$.

$$\therefore\quad V\left(\bar{y}_{st}\right)=\frac{k-1}{n^2k}\sum_{j=1}^{n}\frac{k(k+1)}{12}$$

$$=\frac{k-1}{n^2k}\cdot\frac{nk(k+1)}{12}$$

$$=\frac{k^2-1}{12n}\quad\text{.............(iv)}$$

এখন ধারাবাহিক নমুনার ভেদাংক, $V\left(\bar{y}_{sys}\right)=\dfrac{1}{k}\sum_{i=1}^{k}\left(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot}\right)^2$

এখানে, i-তম ধারাবাহিক নমুনার গড়,

$$\bar{y}_{i\cdot}=\frac{1}{n}\sum_{j=1}^{n}y_{ij}$$

$$=\frac{1}{n}\left[i+(i+k)+(i+2k)+\cdots+i+(n-1)k\right]$$

$$=\frac{1}{n}\left[ni+\left\{k+2k+\cdots+(n-1)k\right\}\right]$$

$$=\frac{1}{n}\left[ni+k\left\{1+2+\cdots+(n-1)\right\}\right]$$

$$=\frac{1}{n}\left[ni+k.\frac{(n-1)(n-1+1)}{2}\right]$$

$$=\frac{1}{n}\left[ni+k.\frac{(n-1)n}{2}\right]$$

$$=\frac{ni}{n}+\frac{k\,(n-1)\,n}{2n}$$

$$= i + \frac{k(n-1)}{2} \quad ....(v)$$

$$V\left(\bar{y}_{sys}\right) = \frac{1}{k}\sum_{i=1}^{k}\left[i + \frac{k(n-1)}{2} - \frac{nk+1}{2}\right]^2$$ [(i) ও (v) নং হতে মান বসিয়ে]

$$= \frac{1}{k}\sum_{i=1}^{k}\left[i + \frac{nk-k-nk-1}{2}\right]^2$$

$$= \frac{1}{k}\sum_{i=1}^{k}\left[i + \frac{-k-1}{2}\right]^2$$

$$= \frac{1}{k}\sum_{i=1}^{k}\left[i - \frac{k+1}{2}\right]^2$$

$$= \frac{1}{k}\sum_{i=1}^{k}\left[i^2 + \left(\frac{k+1}{2}\right)^2 - 2i\frac{k+1}{2}\right]$$

$$= \frac{1}{k}\sum_{i=1}^{k} i^2 + \frac{1}{k}\sum_{i=1}^{k}\left(\frac{k+1}{2}\right)^2 - \frac{(k+1)}{k}\sum_{i=1}^{k} i$$

$$= \frac{1}{k}\frac{k(k+1)(2k+1)}{6} + \frac{1}{k}k\frac{(k+1)^2}{4} - \frac{(k+1)}{k}\frac{k(k+1)}{2}$$

$$= \frac{(k+1)(2k+1)}{6} + \frac{(k+1)^2}{4} - \frac{(k+1)^2}{2}$$

$$= \frac{(k+1)(2k+1)}{6} - \frac{(k+1)^2}{4}$$

$$= \frac{k+1}{2}\left[\frac{2k+1}{3} - \frac{k+1}{2}\right]$$

$$= \frac{k+1}{2}\left[\frac{4k+2-3k-3}{6}\right]$$

$$= \frac{k+1}{2} \times \frac{k-1}{6} = \frac{k^2-1}{12} \quad ....(vi)$$

(iii), (iv) ও (vi) নং হতে পাই,

$$V\left(\bar{y}_{st}\right) : V\left(\bar{y}_{sys}\right) : V\left(\bar{y}_{ran}\right) = \frac{k^2-1}{12n} : \frac{k^2-1}{12} : \frac{(k-1)(nk+1)}{12}$$

$$= \frac{k+1}{n} : k+1 : nk+1$$

$$= \frac{k+1}{n} : k+1 : n\left(k + \frac{1}{n}\right)$$

$$\cong \frac{1}{n} : 1 : n$$

$$\cong 1 : n : n^2$$

$$\therefore \; V(\bar{y}_{st}) \le V(\bar{y}_{sys}) \le V(\bar{y}_{ran})$$

তবে, $n > 1$ হলে $V(\bar{y}_{st}) < V(\bar{y}_{sys}) < V(\bar{y}_{ran})$

প্রশ্নঃ যদি তথ্যবিশ্বের মানগুলো $Y_i = a + bi$ আকারের সরলরৈখিক ধারা অনুসরণ করে, তবে দেখাও যে,

$$V(\bar{y}_{st}) \le V(\bar{y}_{sys}) \le V(\bar{y}_{ran})$$

সমাধানঃ দেওয়া আছে, $Y_i = a + bi; \quad (i = 1, 2, \cdots, N)$

যেখানে, a ও b দুটি ধ্রুবক।

তথ্যবিশ্বের গড়, $\bar{Y} = \bar{y}_{..} = \frac{1}{N}\sum_{i=1}^{N} Y_i$

$$= \frac{1}{N}\sum_{i=1}^{N}(a + bi)$$

$$= a + b\,\frac{\sum_{i=1}^{N} i}{N}$$

$$= a + b\,\frac{1 + 2 + \cdots + N}{N}$$

$$= a + b\,\frac{N(N+1)}{2N}$$

$$= a + \frac{b(N+1)}{2}$$

$$= a + \frac{b(nk+1)}{2} \quad \text{..........(i)}$$

এবং $S^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^2$

$$= \frac{1}{N-1}\sum_{i=1}^{N}\left[a + bi - a - \frac{b(N+1)}{2}\right]^2$$

$$= \frac{b^2}{N-1}\sum_{i=1}^{N}\left[i - \frac{N+1}{2}\right]^2$$

$$= \frac{b^2}{N-1}\sum_{i=1}^{N}\left[i^2 + \frac{(N+1)^2}{4} - 2.i.\frac{N+1}{2}\right]$$

$$= \frac{b^2}{N-1}\left[\sum_{i=1}^{N} i^2 + \frac{N(N+1)^2}{4} - 2\frac{N+1}{2}\sum_{i=1}^{N} i\right]$$

$$= \frac{b^2}{N-1}\left[\frac{N(N+1)(2N+1)}{6} + \frac{N(N+1)^2}{4} - (N+1)\frac{N(N+1)}{2}\right]$$

$$= \frac{b^2 N(N+1)}{N-1}\left[\frac{2N+1}{6} + \frac{N+1}{4} - \frac{N+1}{2}\right]$$

$$= \frac{b^2 N(N+1)}{N-1}\left[\frac{2N+1}{6} - \frac{N+1}{4}\right]$$

$$= \frac{b^2 N(N+1)}{N-1} \times \frac{4N+2-3N-3}{12}$$

$$= \frac{b^2 N(N+1)}{N-1} \times \frac{N-1}{12}$$

$$= \frac{N(N+1)b^2}{12}$$

$$\therefore S^2 = \frac{nk(nk+1)b^2}{12}$$

এখন, $V(\bar{y}_{ran}) = \frac{N-n}{N}\frac{S^2}{n}$

$$= \frac{nk-n}{nk} \times \frac{nk(nk+1)b^2}{12n}$$

$$= \frac{n(k-1)(nk+1)b^2}{12n}$$

$$= \frac{(k-1)(nk+1)b^2}{12} \quad ....(ii)$$

যেহেতু ধারাবাহিক নমুনার সাথে তুলনা করতে হবে তাই $N = nk$ বসানো হয়েছে।

আবার, $V(\bar{y}_{st}) = \frac{1}{N^2}\sum_{j=1}^{n} N_j (N_j - n_j)\frac{S_j^2}{n_j}$

এখানে $N_j = k$ এবং $n_j = 1$

সুতরাং $V(\bar{y}_{st}) = \frac{1}{N^2}\sum_{j=1}^{n} k(k-1)\, S_j^2$

$$= \frac{k(k-1)}{n^2 k^2}\sum_{j=1}^{n} S_j^2$$

$$=\frac{k-1}{n^2 k}\sum_{j=1}^{n} S_j^{\,2}$$

আমরা পূর্বে পেয়েছি, $S^2=\frac{N(N+1)\,b^2}{12}$ । কিন্তু ধারাবাহিক নমুনায়নকে n স্তর বিশিষ্ট স্তরিত দৈব নমুনা বিবেচনা করা যেতে পারে, যেখানে প্রতি স্তরের একক সংখ্যা k। অতএব, $S_j^{\,2}=\frac{k(k+1)\,b^2}{12}$.

$$\therefore\ V\left(\bar{y}_{st}\right)=\frac{k-1}{n^2 k}\sum_{j=1}^{n}\frac{k(k+1)b^2}{12}$$

$$=\frac{k-1}{n^2 k}\cdot\frac{n\,k(k+1)\,b^2}{12}$$

$$=\frac{b^2\left(k^2-1\right)}{12n}\ ....(iii)$$

এখন ধারাবাহিক নমুনার ভেদাংক, $V\left(\bar{y}_{sys}\right)=\frac{1}{k}\sum_{i=1}^{k}\left(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot}\right)^2$

এখানে, i-তম ধারাবাহিক নমুনার গড়,

$$\bar{y}_{i\cdot}=\frac{1}{n}\sum_{j=1}^{n} y_{ij}$$

$$=\frac{1}{n}\left[a+bi+a+b(i+k)+a+b(i+2k)+\cdots+a+b\left\{i+(n-1)k\right\}\right]$$

$$=\frac{1}{n}\left[na+nbi+b\left\{k+2k+\cdots+(n-1)k\right\}\right]$$

$$=\frac{1}{n}\left[na+nbi+bk\left\{1+2+\cdots+(n-1)\right\}\right]$$

$$=\frac{1}{n}\left[na+nbi+bk.\frac{(n-1)\ (n-1+1)}{2}\right]$$

$$=\frac{1}{n}\left[na+nbi+bk.\frac{(n-1)n}{2}\right]$$

$$=a+bi+\frac{bk(n-1)}{2}\ ....(iv)$$

$$V\left(\bar{y}_{sys}\right)=\frac{1}{k}\sum_{i=1}^{k}\left[a+bi+\frac{bk(n-1)}{2}-a-\frac{b(nk+1)}{2}\right]^2 \quad \text{[(i) ও (iv) হতে মান বসিয়ে]}$$

$$=\frac{b^2}{k}\sum_{i=1}^{k}\left[i+\frac{k(n-1)}{2}-\frac{nk+1}{2}\right]^2$$

$$= \frac{b^2}{k} \sum_{i=1}^{k} \left[i + \frac{nk - k - nk - 1}{2} \right]^2$$

$$= \frac{b^2}{k} \sum_{i=1}^{k} \left[i + \frac{-k-1}{2} \right]^2$$

$$= \frac{b^2}{k} \sum_{i=1}^{k} \left[i - \frac{k+1}{2} \right]^2$$

$$= \frac{b^2}{k} \sum_{i=1}^{k} \left[i^2 + \frac{(k+1)^2}{4} - 2i\frac{k+1}{2} \right]$$

$$= \frac{b^2}{k} \left[\sum_{i=1}^{k} i^2 + \frac{k(k+1)^2}{4} - (k+1) \sum_{i=1}^{k} i \right]$$

$$= \frac{b^2}{k} \left[\frac{k(k+1)(2k+1)}{6} + \frac{k(k+1)^2}{4} - (k+1)\frac{k(k+1)}{2} \right]$$

$$= \frac{k(k+1)b^2}{k} \left[\frac{(2k+1)}{6} + \frac{k+1}{4} - \frac{k+1}{2} \right]$$

$$= \frac{k(k+1)b^2}{k} \left[\frac{(2k+1)}{6} - \frac{k+1}{4} \right]$$

$$= \frac{k(k+1)b^2}{k} \times \frac{4k+2-3k-3}{12}$$

$$= \frac{(k+1)b^2(k-1)}{12}$$

$$= \frac{b^2(k^2-1)}{12} \text{.............(v)}$$

(ii), (iii) ও (v) হতে পাই,

$$V(\bar{y}_{st}) : V(\bar{y}_{sys}) : V(\bar{y}_{ran}) = \frac{b^2(k^2-1)}{12n} : \frac{b^2(k^2-1)}{12} : \frac{b^2(k-1)(nk+1)}{12}$$

$$= \frac{k+1}{n} : k+1 : nk+1$$

$$\cong \frac{1}{n} : 1 : n$$

$$\cong 1 : n : n^2$$

$$\therefore \; V(\bar{y}_{st}) \le V(\bar{y}_{sys}) \le V(\bar{y}_{ran})$$

প্রশ্নঃ ধারাবাহিক নমুনায়নে তথ্যবিশ্ব সমষ্টির নিঝুঁকি প্রাক্কলক বের কর এবং এর ভেদাংক নির্ণয় কর।

উত্তরঃ ধরি, তথ্যবিশ্বের একক সংখ্যা $=N$

নমুনার একক সংখ্যা $=n$

এখানে $N=nk$, ($k=$ নমুনায়ন ব্যবধান)

অর্থাৎ তথ্যবিশ্ব থেকে k সংখ্যক ধারাবাহিক নমুনা নেয়া যেতে পারে।

ধরি, y_{ij} $(i=1, 2, \cdots, k;\ j=1, 2, \cdots, n)$ হল $i-$তম ধারাবাহিক নমুনার $j-$তম উপাদানের মান।

$i-$তম ধারাবাহিক নমুনার গড়, $\bar{y}_{i\cdot}=\dfrac{1}{n}\sum_{j=1}^{n} y_{ij}$

তথ্যবিশ্ব গড়, $\bar{y}_{\cdot\cdot}=\dfrac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n} y_{ij}=\dfrac{1}{nk}\sum_{i=1}^{k}\sum_{j=1}^{n} y_{ij}=\dfrac{1}{k}\sum_{i=1}^{k}\bar{y}_{i\cdot}$

তথ্যবিশ্ব সমষ্টি, $Y=N\bar{y}_{\cdot\cdot}$

$S^2=\dfrac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_{\cdot\cdot}\right)^2$ এবং $S^2_{wsy}=\dfrac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_{i\cdot}\right)^2$; যেখানে $\bar{y}_{i\cdot}$ কে $\bar{y}_{sys}$ দ্বারা চিহ্নিত করা হয়।

ধরি, $\hat{Y}=N\bar{y}_{sys}$ হলো তথ্যবিশ্ব সমষ্টির নিঝুঁকি প্রাক্কলক।

এখন,
$$
\begin{aligned}
E\left(\hat{Y}\right) &= E\left(N\ \bar{y}_{sys}\right)\\
&= N\ E\left(\bar{y}_{sys}\right)\\
&= N\ E\left(\bar{y}_{i\cdot}\right)\\
&= N\sum_{i=1}^{k}\bar{y}_{i\cdot}\ P\left(\bar{y}_{i\cdot}\right)\\
&= N\sum_{i=1}^{k}\bar{y}_{i\cdot}\times\frac{1}{k}\\
&= N\bar{y}_{\cdot\cdot}\\
&= Y
\end{aligned}
$$

অতএব, $\hat{Y}=N\bar{y}_{sys}$ হল তথ্যবিশ্ব সমষ্টির নিঝুঁকি প্রাক্কলক।

এখন, তথ্যবিশ্ব সমষ্টির প্রাক্কলকের ভেদাংক,

$$V\left(\hat{Y}\right)=V\left(N\ \bar{y}_{sys}\right)=N^2\ V\left(\bar{y}_{sys}\right)\ldots\ldots\ldots\ldots(i)$$

এখন, $V\left(\bar{y}_{sys}\right)=E\left[\bar{y}_{sys}-E\left(\bar{y}_{sys}\right)\right]^2$

$$= E\left[\bar{y}_{i.} - \bar{y}_{..}\right]^2$$

$$= \frac{1}{k}\sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$$\Rightarrow k\ V\left(\bar{y}_{sys}\right) = \sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2 \ \ldots\ldots\ldots\ldots(ii)$$

জানি, $S^2 = \frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{..}\right)^2$

$$\Rightarrow (N-1)S^2 = \sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.} + \bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n}\left[\left(y_{ij} - \bar{y}_{i.}\right) + \left(\bar{y}_{i.} - \bar{y}_{..}\right)\right]^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)^2 + 2\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)\left(\bar{y}_{i.} - \bar{y}_{..}\right) + \sum_{i=1}^{k}\sum_{j=1}^{n}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)^2 + 2\sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right) + n\sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)^2 + 2\sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)\times 0 + n\sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$$\Rightarrow (N-1)S^2 = \sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)^2 + n\sum_{i=1}^{k}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$$\Rightarrow (N-1)S^2 = k(n-1)\,S^2_{wsy} + n\,k\ V\left(\bar{y}_{sys}\right)$$ [(ii) নং হতে মান বসিয়ে]

$$\left[\because\ S^2_{wsy} = \frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i.}\right)^2\right]$$

$$\Rightarrow n\,k\ V\left(\bar{y}_{sys}\right) = (N-1)S^2 - k(n-1)S^2_{wsy}$$

$$\Rightarrow V\left(\bar{y}_{sys}\right) = \frac{N-1}{n\,k}S^2 - \frac{k(n-1)}{n\,k}S^2_{wsy}$$

$$\Rightarrow V\left(\bar{y}_{sys}\right) = \frac{N-1}{N}S^2 - \frac{k(n-1)}{N}S^2_{wsy}$$

এখন (i) নং কে লেখা যায়-

$$V\left(\hat{Y}\right) = N^2\left\{\frac{N-1}{N}S^2 - \frac{k(n-1)}{N}S^2_{wsy}\right\}$$

$$= N(N-1)S^2 - k\,N(n-1)S^2_{wsy}$$

সুতরাং তথ্যবিশ্ব সমষ্টির ভেদাংক, $V(\hat{Y}) = N(N-1)S^2 - kN(n-1)S^2_{wsy}$

প্রশ্ন: ধারাবাহিক নমুনায়নকে কেন গুচ্ছ নমুনায়ন ও স্তরিত নমুনায়ন বলা হয়ে থাকে ?

উত্তর: গুচ্ছ নমুনায়নে এককের পরিবর্তে এককের গ্রুপকে নমুনার একক হিসেবে ধরা হয় এবং এককের গ্রুপকে গুচ্ছ (cluster) একক বলা হয়। যদি পরিবারের সদস্যকে একক ধরা হয়, তবে পরিবারকে গুচ্ছ একক হিসেবে ধরা যায়। ধারাবাহিক নমুনায়নে প্রথম উপাদান নির্বাচনের পর বাকিগুলো ধারাবাহিকভাবে আসে। উপরের সারণীর k সারিকে k গুচ্ছ হিসেবে গণ্য করা যায়। ধারাবাহিক নমুনায়নে এই k গুচ্ছের একটিকে দৈবভাবে চয়ন করা হয়। তাই ধারাবাহিক নমুনায়নকে গুচ্ছ নমুনায়ন বলা হয়।

আবার উপরের সারণীর কলামগুলোকে এক একটি স্তর (stratum) হিসেবে গণ্য করা যায়। যেহেতু ১ম k একক, ২য় k একক, ৩য় k একক,..., শেষ k একক থেকে একটি একটি একক নমুনায় নির্বাচিত হয়, সে কারণে এ নমুনায়ন পদ্ধতিকে প্রতি স্তর থেকে এক একক বিশিষ্ট স্তরিত দৈব নমুনায়নও বলা যেতে পারে। কিন্তু সঠিকভাবে এটাকে স্তরিত দৈব নমুনা হিসেবে গণ্য করা যায় না, কারণ ধারাবাহিক নমুনায় প্রত্যেক স্তরের এককগুলো নির্ধারিত একই অবস্থানে থাকে। কিন্তু স্তরিত দৈব নমুনায়নে এককগুলোর অবস্থান দৈবভাবে নির্বাচিত হয়।

সংক্ষিপ্ত প্রশ্ন ও উত্তর:

(১) কখন চক্রাকার ধারাবাহিক নমুনায়ন ব্যবহার করা হয়? (জা.বি.-১২)

উত্তর: যদি তথ্যবিশ্বের আকার N, নমুনার আকার n এর সঠিক গুণিতক না হয়, সেক্ষেত্রে k এর মান হবে $\frac{N}{n}$ এর নিকটতম পূর্ণ সংখ্যা। তখন ধারাবাহিক নমুনার আকার n অথবা $(n-1)$ হবে যা নির্ভর করে প্রথম নির্বাচিত এককের উপর। এ অসুবিধা দূর করার জন্য চক্রাকার ধারাবাহিক নমুনায়ন ব্যবহার করা হয়।

গুচ্ছ নমুনায়ন (Cluster Sampling)

প্রশ্নঃ গুচ্ছ নমুনায়নের সংজ্ঞা দাও (জা.বি.-১৫)। এমন একটি উদাহরণ দাও যেখানে গুচ্ছ নমুনায়ন ব্যবহৃত হয় (জা.বি.-১১)

অথবা, গুচ্ছ নমুনায়ন কি? (জা.বি.-১৫)

উত্তরঃ গুচ্ছ নমুনায়নে প্রথমে তথ্যবিশ্বের এককসমূহকে কতগুলো গ্রুপে ভাগ করা হয়, যাদের বলা হয় গুচ্ছ। গুচ্ছগুলো এমনভাবে তৈরী করা হয় যেন প্রতিটি একক কোন না কোন গুচ্ছে একবার এবং মাত্র একবার অন্তর্ভূক্ত হবে। একই গুচ্ছের এককগুলো যতদূর সম্ভব বৈসাদৃশ্যপূর্ণ হবে এবং ভিন্ন গুচ্ছের এককগুলো সাদৃশ্যপূর্ণ হবে। উক্ত গুচ্ছগুলোকে নমুনা একক হিসেবে বিবেচনা করে প্রয়োজনীয় সংখ্যক গুচ্ছ সমান অথবা অসমান সম্ভাবনার ভিত্তিতে নির্বাচন করে নির্বাচিত গুচ্ছের প্রতিটি একক থেকে তথ্য সংগ্রহ করা হলে সম্পূর্ণ পদ্ধতিকে বলা হয় গুচ্ছ নমুনায়ন।

উদাহরণ- বাংলাদেশে কত সক্ষম দম্পতি পরিবার পরিকল্পনা গ্রহণ করে তা নিরূপন করতে হবে। এক্ষেত্রে বাংলাদেশের প্রায় ৬৮ হাজার গ্রামকে গুচ্ছ বিবেচনা করে কিছু গ্রামকে দৈবায়িতভাবে চয়ন করে চয়ন করা গ্রামগুলোর সকল সক্ষম দম্পতিদের নিকট থেকে তথ্য সংগ্রহ করা হলে তা হবে গুচ্ছ নমুনায়ন।

প্রশ্নঃ গুচ্ছ নমুনায়নের নীতিগুলো আলোচনা কর (জা.বি.-১১)

উত্তরঃ গুচ্ছ নমুনায়নের নীতিগুলো নিম্নরূপ-

(ক) গুচ্ছগুলো এমনভাবে তৈরী করা হয় যেন প্রতিটি একক কোন না কোন গুচ্ছে কেবলমাত্র একবার অন্তর্ভূক্ত হবে।

(খ) একই গুচ্ছের এককগুলো যতদূর সম্ভব বৈসাদৃশ্যপূর্ণ হবে।

(গ) ভিন্ন গুচ্ছের এককগুলো সাদৃশ্যপূর্ণ হবে।

(ঘ) গুচ্ছের আকার সমান অথবা অসমান হতে পারে।

প্রশ্নঃ গুচ্ছ নমুনায়নের সুবিধা ও অসুবিধাগুলো আলোচনা কর (জা.বি.-১৬)

প্রশ্নঃ অন্যান্য নমুনায়নের তুলনায় গুচ্ছ নমুনায়নের সুবিধাগুলো আলোচনা কর **(জা.বি.-১২, ১৫)**

উত্তরঃ তথ্যবিশ্বের নমুনা কাঠামো জানা থাকলেও সরল দৈব নমুনায়ন বা স্তরিত দৈব নমুনায়ন বা ধারাবাহিক নমুনায়ন সময় এবং সম্পদের দিক থেকে সুবিধাজনক নয়। সুতরাং গুচ্ছ নমুনায়নের সুবিধাগুলো হলো-

(ক) গুচ্ছ যেহেতু নিকটবর্তী একক নিয়ে গঠিত, সে কারণে উপাত্ত সংগ্রহ করা সহজ। অতি দ্রুত এবং অল্প খরচে তা সম্পন্ন করা যায় এবং প্রশাসনিক দিক থেকেও এটি সুবিধাজনক। কারণ জরিপ কাজের তত্ত্বাবধান সহজ।

(খ) তথ্যবিশ্বের নমুনা কাঠামো জানা না থাকলেও এ পদ্ধতি সহজেই প্রয়োগ করা যায়।

প্রশ্নঃ গুচ্ছ নমুনায়নের অসুবিধাসমূহ লিখ।

উত্তরঃ গুচ্ছ নমুনায়নের অসুবিধাসমূহ নিম্নরূপ-

(ক) একই গুচ্ছের এককসমূহ সমজাতীয় হলে গুচ্ছ নমুনায়নের দক্ষতা কমে যায়।

(খ) গুচ্ছ একক সংখ্যা বেশি হলেও গুচ্ছ নমুনায়ন ভাল ফলাফল দেয় না।

বি. দ্রঃ গুচ্ছ নমুনায়ন ভাল হবে যদি গুচ্ছের সংখ্যা বেশি হয় এবং গুচ্ছের একক সংখ্যা কম হয়।

গুচ্ছ নমুনায়নে ব্যবহৃত প্রতীক ও পরিভাষা (Notations and Terminology used in Cluster Sampling):

মনে করি, একটি তথ্যবিশ্বে N গুচ্ছ আছে এবং প্রতি গুচ্ছে M সংখ্যক একক আছে। ধরি, N গুচ্ছ থেকে সরল দৈব নমুনায়নের মাধ্যমে n গুচ্ছ চয়ন করা হলো। তাহলে

তথ্যবিশ্বের গুচ্ছের সংখ্যা $= N$

নমুনায় গুচ্ছের সংখ্যা $= n$

গুচ্ছের একক (বা উপাদান) সংখ্যা $= M$

$y_{ij} =$ i-তম গুচ্ছের j-তম এককের মান $\left(i = 1, 2, \cdots, N;\ j = 1, 2, \cdots, M\right)$

তথ্যবিশ্বের i-তম গুচ্ছের গড়, $\overline{y}_{i\cdot} = \dfrac{1}{M}\sum\limits_{j=1}^{M} y_{ij}$

n আকারের নমুনা গুচ্ছের জন্য গুচ্ছ গড়ের গড়, $\overline{y}_n = \overline{y}_c = \dfrac{1}{n}\sum\limits_{i=1}^{n} \overline{y}_{i\cdot}$

তথ্যবিশ্বে গুচ্ছ গড়ের গড়, $\overline{Y}_N = \dfrac{1}{N}\sum\limits_{i=1}^{N} \overline{y}_{i\cdot}$

তথ্যবিশ্বের গড়, $\overline{Y} = \dfrac{1}{NM}\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{M} y_{ij} = \dfrac{1}{N}\sum\limits_{i=1}^{N}\dfrac{1}{M}\sum\limits_{j=1}^{M} y_{ij} = \dfrac{1}{N}\sum\limits_{i=1}^{N} \overline{y}_{i\cdot}$

তথ্যবিশ্বের গুচ্ছসমূহের গড় বর্গ বিচ্যুতি (Mean square between clusters in the population),

$$S_b^{\,2} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\overline{y}_{i\cdot} - \overline{Y}_N\right)^2$$

তথ্যবিশ্বের এককসমূহের গড় বর্গ বিচ্যুতি (Mean square between elements in the population),

$$S^2 = \frac{1}{NM-1}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(y_{ij} - \overline{Y}\right)^2$$

একটি গুচ্ছের মধ্যবর্তী এককসমূহের অন্তঃগুচ্ছ সংশ্লেষাংক (intra-cluster correlation coefficient),

$$\rho = \frac{\sum\limits_{i=1}^{N}\sum\limits_{j\neq}^{M}\sum\limits_{k=1}^{M}\left(y_{ij} - \overline{Y}\right)\left(y_{ik} - \overline{Y}\right)}{(M-1)(NM-1)S^2}$$

উপপাদ্যঃ দেখাও যে, সমান আকারের গুচ্ছের ক্ষেত্রে তথ্যবিশ্বে গুচ্ছ গড়ের গড়, তথ্যবিশ্বের গড়ের সমান অর্থাৎ $\overline{Y}_N = \overline{Y}$

প্রমাণঃ জানি তথ্যবিশ্বের গড়, $\overline{Y} = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}$

$$= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{M}\sum_{j=1}^{M} y_{ij}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\overline{y}_{i.}$$

$$= \overline{Y}_N = \text{ তথ্যবিশ্বে গুচ্ছ গড়ের গড়}$$

অর্থাৎ সমান আকারের গুচ্ছের ক্ষেত্রে তথ্যবিশ্বে গুচ্ছ গড়ের গড়, তথ্যবিশ্বের গড়ের সমান।

উপপাদ্যঃ গুচ্ছ নমুনায়নে তথ্যবিশ্ব গড়ের নিঝুঁকি প্রাক্কলক নির্ণয় কর।

অথবা, গুচ্ছ নমুনায়নে তথ্যবিশ্বটির একটি নিঝুঁকি নিরূপক নির্ণয় কর **(জা.বি.-১১)**

অথবা, দেখাও যে নমুনা গুচ্ছের জন্য গুচ্ছ গড়ের গড়, তথ্যবিশ্ব গড়ের নিঝুঁকি প্রাক্কলক অর্থাৎ $E(\overline{y}_n) = \overline{Y}$

অথবা, গুচ্ছ নমুনায়নের ক্ষেত্রে প্রচলিত সংকেতে প্রমাণ কর যে, $E(\overline{y}_n) = \overline{Y}$ **(জা.বি.-১৫)**

অথবা, ধরা যাক, N গুচ্ছবিশিষ্ট তথ্যবিশ্বের প্রতিটি গুচ্ছ M সংখ্যক একক আছে। n গুচ্ছের একটি সরল দৈব নমুনা পুনঃস্থাপন ব্যতিরেকে নেয়া হল। যদি নমুনার প্রতি এককের গড় $\overline{y}_c$ হয়, তবে দেখাও যে, $\overline{y}_c$ হল সমগ্রক গড়ের নিঝুঁকি প্রাক্কলক **(জা.বি.-১২)**

অথবা, N সংখ্যক গুচ্ছের একটি সমগ্রকের প্রতিটি গুচ্ছে M সংখ্যক উপাদান থাকলে পুনঃস্থাপন না করে সরল দৈব নমুনায়নের মাধ্যমে সংখ্যক গুচ্ছ চয়ন করা হলে দেখাও যে, নমুনা গড় $\overline{\overline{y}}$ হচ্ছে সমগ্রক গড় $\overline{\overline{Y}}$ এর নিঝুঁকি নিরূপক **(জা.বি.-১৬)**

প্রমাণঃ জানি, নমুনা গুচ্ছের জন্য গুচ্ছ গড়ের গড়, $\overline{y}_n = \frac{1}{n}\sum_{i=1}^{n}\overline{y}_{i.}$

এখন, $E(\overline{y}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n}\overline{y}_{i.}\right) = \frac{1}{n}\sum_{i=1}^{n}E(\overline{y}_{i.})$..........(i)

$\overline{y}_{i.}$ এর সম্ভাব্য মানসমূহ $\overline{y}_{1.}, \overline{y}_{2.}, \cdots, \overline{y}_{N.}$ এবং প্রতিটির সম্ভাবনা $\frac{1}{N}$

প্রত্যাশার সংজ্ঞানুযায়ী, $E(\overline{y}_{i.}) = \sum_{i=1}^{N}\overline{y}_{i.}\ P(\overline{y}_{i.})$

$$= \sum_{i=1}^{N}\overline{y}_{i.} \times \frac{1}{N}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\overline{y}_{i.}$$

$$=\frac{1}{N}\sum_{i=1}^{N}\frac{1}{M}\sum_{j=1}^{M}y_{ij}$$

$$=\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}y_{ij}$$

$$=\overline{Y}$$

এখন, $E\left(\overline{y}_{i.}\right)$ এর মান (i) নং সমীকরণে বসিয়ে পাই,

$$E\left(\overline{y}_n\right)=\frac{1}{n}\sum_{i=1}^{n}\overline{Y}=\frac{1}{n}\times n\overline{Y}=\overline{Y}$$

$$\therefore\ E\left(\overline{y}_n\right)=\overline{Y}$$

অতএব, নমুনা গুচ্ছের জন্য গুচ্ছ গড়ের গড়, তথ্যবিশ্ব গড়ের নিঝুঁকি (unbiased) প্রাক্কলক (প্রমাণিত)।

বি. দ্রঃ নমুনার প্রতি এককের গড়, $\overline{y}_c=\frac{1}{n}\sum_{i=1}^{n}\overline{y}_{i.}$

উপপাদ্যঃ সম-আকারের গুচ্ছায়নের ক্ষেত্রে এর ভেদাংকের সূত্রটি উদ্ভাবন কর (জা.বি.-১১)

অথবা, দেখাও যে, $V\left(\overline{y}_n\right)=\frac{N-n}{nMN}S^2\left[1+\left(M-1\right)\rho\right]$, যেখানে ρ হল আন্তঃগুচ্ছ সংশ্লেষাংক **(জা.বি.-১২)**

অথবা, N সংখ্যক গুচ্ছের একটি সমগ্রকের প্রতিটি গুচ্ছে M সংখ্যক উপাদান থাকলে পুনঃস্থাপন না করে সরল দৈব নমুনায়নের মাধ্যমে n সংখ্যক গুচ্ছ চয়ন করা হলে দেখাও যে, নমুনা গড় $\overline{\overline{y}}$ এর ভেদাংক , $V\left(\overline{\overline{y}}\right)=\frac{1-f}{Mn}S^2\left[1+\left(M-1\right)\rho\right]$, এখানে,

$\rho=$ আন্তশ্রেণি সংশ্লেষাংক, $f=\frac{n}{N}$ **(জা.বি.-১৬)**

অথবা, দেখাও যে গুচ্ছ নমুনায়নে নমুনা গুচ্ছ গড়ের ভেদাংক হলো

$$V\left(\overline{y}_n\right)=\frac{N-n}{N}.\frac{{S_b}^2}{n}=\left(1-f\right)\frac{{S_b}^2}{n}=\frac{1-f}{Mn}S^2\left[1+\left(M-1\right)\rho\right]$$

যেখানে ${S_b}^2=\frac{1}{N-1}\sum_{i=1}^{N}\left(\overline{y}_{i.}-\overline{Y}_N\right)^2$, অন্তঃগুচ্ছ সংশ্লেষাংক, $\rho=\frac{\sum_{i=1}^{N}\sum_{j\neq}^{M}\sum_{k=1}^{M}\left(y_{ij}-\overline{Y}\right)\left(y_{ik}-\overline{Y}\right)}{(M-1)\ (NM-1)S^2}$ এবং $f=\frac{n}{N}$

প্রমাণঃ জানি, তথ্যবিশ্বের গড়, $\overline{Y}=\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}y_{ij}$

এবং নমুনা গুচ্ছের জন্য গুচ্ছ গড়ের গড়, $\overline{y}_n=\frac{1}{n}\sum_{i=1}^{n}\overline{y}_{i.}$

এখন, $V\left(\overline{y}_n\right)=E\left[\overline{y}_n-E\left(\overline{y}_n\right)\right]^2$

$$=E\left(\overline{y}_n-\overline{Y}\right)^2\quad\left[\because\ E\left(\overline{y}_n\right)=\overline{Y}\right]$$

$$=E\left(\frac{\sum_{i=1}^{n}\overline{y}_{i\cdot}}{n}-\overline{Y}\right)^2$$

$$=E\left(\frac{\sum_{i=1}^{n}\overline{y}_{i\cdot}-n\overline{Y}}{n}\right)^2$$

$$=\frac{1}{n^2}\;E\left(\sum_{i=1}^{n}\overline{y}_{i\cdot}-\sum_{i=1}^{n}\overline{Y}\right)^2$$

$$=\frac{1}{n^2}\;E\left\{\sum_{i=1}^{n}\left(\overline{y}_{i\cdot}-\overline{Y}\right)\right\}^2$$

$$=\frac{1}{n^2}\;E\left\{\sum_{i=1}^{n}\left(\overline{y}_{i\cdot}-\overline{Y}_N\right)\right\}^2 \quad \left[\because\; \overline{Y}=\overline{Y}_N\right]$$

$$=\frac{1}{n^2}\;E\left\{\sum_{i=1}^{n}\left(\overline{y}_{i\cdot}-\overline{Y}_N\right)^2+\sum_{i\neq}^{n}\sum_{i'=1}^{n}\left(\overline{y}_{i\cdot}-\overline{Y}_N\right)\left(\overline{y}_{i'\cdot}-\overline{Y}_N\right)\right\}$$

$$\left[\because\;\left(\sum_{i=1}^{n}y_i\right)^2=\sum_{i=1}^{n}{y_i}^2+\sum_{i\neq}^{n}\sum_{i'=1}^{n}y_i\,y_{i'}\right]$$

$$\therefore\;\; V\left(\overline{y}_n\right)=\frac{1}{n^2}\left\{\sum_{i=1}^{n}E\left(\overline{y}_{i\cdot}-\overline{Y}_N\right)^2+\sum_{i\neq}^{n}\sum_{i=1}^{n}E\left(\overline{y}_{i\cdot}-\overline{Y}_N\right)\left(\overline{y}_{i'\cdot}-\overline{Y}_N\right)\right\}\text{.........(i)}$$

নমুনার i-তম মান $\left(\overline{y}_{i\cdot}-\overline{Y}_N\right)^2$ এর সম্ভাব্য মানগুলো হলো সমগ্রকের $\left(\overline{y}_{1\cdot}-\overline{Y}_N\right)^2,\left(\overline{y}_{2\cdot}-\overline{Y}_N\right)^2,\cdots,\left(\overline{y}_{N\cdot}-\overline{Y}_N\right)^2$ এবং এই সকল মানের যে কোন একটি নমুনায় আসার সম্ভাবনা $\frac{1}{N}$.

$$\therefore\;\; E\left(\overline{y}_{i\cdot}-\overline{Y}_N\right)^2=\sum_{i=1}^{N}\left(\overline{y}_{i\cdot}-\overline{Y}_N\right)^2\;\frac{1}{N}$$

$$=\frac{N-1}{N}\,\frac{\sum_{i=1}^{N}\left(\overline{y}_{i\cdot}-\overline{Y}_N\right)^2}{N-1}$$

$$=\frac{N-1}{N}\,{S_b}^2\text{...........(ii)} \quad \left[\because\; {S_b}^2=\frac{1}{N-1}\sum_{i=1}^{N}\left(\overline{y}_{i\cdot}-\overline{Y}_N\right)^2\right]$$

একইভাবে, $\left(\overline{y}_{i\cdot}-\overline{Y}_N\right)\left(\overline{y}_{i'\cdot}-\overline{Y}_N\right)$ এর সম্ভাব্য মানগুলো হলো সমগ্রকের $\left(\overline{y}_{1\cdot}-\overline{Y}_N\right)\left(\overline{y}_{2\cdot}-\overline{Y}_N\right)$, $\left(\overline{y}_{2\cdot}-\overline{Y}_N\right)\left(\overline{y}_{3\cdot}-\overline{Y}_N\right),\cdots,\left(\overline{y}_{N-1\cdot}-\overline{Y}_N\right)\left(\overline{y}_{N\cdot}-\overline{Y}_N\right)$ যাদের মোট সংখ্যা $N(N-1)$ এবং এই সকল মানের যে কোন একটি নমুনায় আসার সম্ভাবনা $\frac{1}{N(N-1)}$.

$$E\left(\bar{y}_{i.}-\bar{Y}_N\right)\left(\bar{y}_{i'.}-\bar{Y}_N\right)=\sum_{i\neq}^{N}\sum_{i'=1}^{N}\left(\bar{y}_{i.}-\bar{Y}_N\right)\left(\bar{y}_{i'.}-\bar{Y}_N\right)\ \frac{1}{N(N-1)}$$

$$=\frac{1}{N(N-1)}\left[\left\{\sum_{i=1}^{N}\left(\bar{y}_{i.}-\bar{Y}_N\right)\right\}^2-\sum_{i=1}^{N}\left(\bar{y}_{i.}-\bar{Y}_N\right)^2\right]\quad\left[\because \sum_{i\neq}^{n}\sum_{i'=1}^{n}y_i y_{i'}=\left(\sum_{i=1}^{n}y_i\right)^2-\sum_{i=1}^{n}y_i^{\,2}\right]$$

$$=\frac{1}{N(N-1)}\left[0-\sum_{i=1}^{N}\left(\bar{y}_{i.}-\bar{Y}_N\right)^2\right]\quad\left[\because \sum_{i=1}^{N}\left(\bar{y}_{i.}-\bar{Y}_N\right)=0\right]$$

$$=-\frac{\sum_{i=1}^{N}\left(\bar{y}_{i.}-\bar{Y}_N\right)^2}{N(N-1)}$$

$$=-\frac{S_b^{\,2}}{N}\ ....(iii)$$

এখন, (ii) নং ও (iii) নং হতে মান (i) নং এ বসিয়ে পাই,

$$V\left(\bar{y}_n\right)=\frac{1}{n^2}\left\{\sum_{i=1}^{n}\frac{N-1}{N}S_b^{\,2}+\sum_{i\neq}^{n}\sum_{i'=1}^{n}\left(-\frac{S_b^{\,2}}{N}\right)\right\}$$

$$=\frac{1}{n^2}\left\{\frac{n\,(N-1)}{N}S_b^{\,2}+n\,(n-1)\left(-\frac{S_b^{\,2}}{N}\right)\right\}$$

$$=\frac{n(N-1)S_b^{\,2}}{n^2\,N}-\frac{n(n-1)S_b^{\,2}}{n^2\,N}$$

$$=\frac{(N-1)S_b^{\,2}}{nN}-\frac{(n-1)S_b^{\,2}}{nN}$$

$$=\frac{S_b^{\,2}}{n\,N}\left(N-1-n+1\right)$$

$$=\frac{\left(N-n\right)S_b^{\,2}}{nN}$$

$$\therefore\quad V\left(\bar{y}_n\right)=\frac{N-n}{N}\cdot\frac{S_b^{\,2}}{n}=\left(1-f\right)\frac{S_b^{\,2}}{n}$$ (প্রমাণিত)

আবার, $V\left(\bar{y}_n\right)=\frac{N-n}{N}\cdot\frac{S_b^{\,2}}{n}$

$$=\frac{N-n}{N}\ \frac{\sum_{i=1}^{N}\left(\bar{y}_{i.}-\bar{Y}_N\right)^2}{n\left(N-1\right)}$$

$$\Rightarrow V\left(\bar{y}_n\right)=\frac{N-n}{Nn\,\left(N-1\right)}\sum_{i=1}^{N}\left(\bar{y}_{i.}-\bar{Y}_N\right)^2\ ....(iv)$$

এখন, $\sum_{i=1}^{N}\left(\bar{y}_{i.}-\bar{Y}_N\right)^2=\sum_{i=1}^{N}\left[\frac{\sum_{j=1}^{M}y_{ij}}{M}-\bar{Y}\right]^2 \quad \left[\because\ \bar{Y}_N=\bar{Y}\right]$

$$=\sum_{i=1}^{N}\left[\frac{\sum_{j=1}^{M}y_{ij}-M\bar{Y}}{M}\right]^2$$

$$=\frac{1}{M^2}\sum_{i=1}^{N}\left[\sum_{j=1}^{M}y_{ij}-M\bar{Y}\right]^2$$

$$=\frac{1}{M^2}\sum_{i=1}^{N}\left[\sum_{j=1}^{M}y_{ij}-\sum_{j=1}^{M}\bar{Y}\right]^2$$

$$=\frac{1}{M^2}\sum_{i=1}^{N}\left[\sum_{j=1}^{M}\left(y_{ij}-\bar{Y}\right)\right]^2$$

$$=\frac{1}{M^2}\sum_{i=1}^{N}\left[\sum_{j=1}^{M}\left(y_{ij}-\bar{Y}\right)^2+\sum_{j\neq}^{M}\sum_{k=1}^{M}\left(y_{ij}-\bar{Y}\right)\left(y_{ik}-\bar{Y}\right)\right]$$

$$=\frac{1}{M^2}\left[\sum_{i=1}^{N}\sum_{j=1}^{M}\left(y_{ij}-\bar{Y}\right)^2+\sum_{i=1}^{N}\sum_{j\neq}^{M}\sum_{k=1}^{M}\left(y_{ij}-\bar{Y}\right)\left(y_{ik}-\bar{Y}\right)\right]$$

$$=\frac{1}{M^2}\left[\left(NM-1\right)S^2+\rho(M-1)\left(NM-1\right)S^2\right]$$

$$\left[\because\ S^2=\frac{1}{NM-1}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(y_{ij}-\bar{Y}\right)^2 \text{ এবং } \rho=\frac{\sum_{i=1}^{N}\sum_{j\neq}^{M}\sum_{k=1}^{M}\left(y_{ij}-\bar{Y}\right)\left(y_{ik}-\bar{Y}\right)}{(M-1)(NM-1)S^2}\right]$$

$$\Rightarrow\sum_{i=1}^{N}\left(\bar{y}_{i.}-\bar{Y}_N\right)^2=\frac{\left(NM-1\right)S^2}{M^2}\left[1+(M-1)\rho\right]$$

এখন (iv) নং কে লেখা যায়-

$$V\left(\bar{y}_n\right)=\frac{N-n}{Nn\left(N-1\right)}\frac{\left(NM-1\right)S^2}{M^2}\left[1+(M-1)\ \rho\right]$$

$$\approx\frac{N-n}{NnN}\ \frac{NMS^2}{M^2}\left[1+(M-1)\ \rho\right] \quad [\ \because\ N \text{ বড় হলে } NM-1\approx NM,\ \ N-1\approx N\ ]$$

$$=\frac{N-n}{Nn}\ \frac{S^2}{M}\left[1+(M-1)\ \rho\right]$$

$$\Rightarrow V\left(\bar{y}_n\right)=\frac{N-n}{nMN}.\ S^2\left[1+(M-1)\rho\right]$$

$$= \frac{1-f}{Mn} S^2 [1+(M-1)\rho]$$

বি. দ্র : (i) $\sum_{i\neq}^{3}\sum_{j=1}^{3}\left(Y_i-\overline{Y}\right)\left(Y_j-\overline{Y}\right)$

$$=\left(Y_1-\overline{Y}\right)\left(Y_2-\overline{Y}\right)+\left(Y_1-\overline{Y}\right)\left(Y_3-\overline{Y}\right)+\left(Y_2-\overline{Y}\right)\left(Y_3-\overline{Y}\right)+\left(Y_2-\overline{Y}\right)\left(Y_1-\overline{Y}\right)+\left(Y_3-\overline{Y}\right)\left(Y_1-\overline{Y}\right)+\left(Y_3-\overline{Y}\right)\left(Y_2-\overline{Y}\right)$$

= Summation of $\{6=3(3-1)\}$ terms.

Similarly, $\sum_{i\neq}^{n}\sum_{j=1}^{n}\left(Y_i-\overline{Y}\right)\left(Y_j-\overline{Y}\right)=$ Summation of $n(n-1)$ terms.

(ii) $V\left(\overline{y}_n\right)=\left(1-f\right)\frac{{S_b}^2}{n}$ ....(i)

$$V\left(\overline{y}_n\right)=\frac{1-f}{n}\,\frac{S^2}{M}\left[1+(M-1)\rho\right] \text{.........(ii)}$$

এখন, (i) ও (ii) নং হতে লেখা যায়,

$$\left(1-f\right)\frac{{S_b}^2}{n}=\frac{1-f}{n}\,\frac{S^2}{M}\left[1+(M-1)\rho\right]$$

$$\Rightarrow {S_b}^2=\frac{S^2}{M}\left[1+(M-1)\rho\right]$$

আবার, $S^2=\dfrac{M{S_b}^2}{1+(M-1)\rho}$

পুনঃস্থাপন ব্যতীত নমুনা থেকে $V\left(\overline{y}_n\right)$ প্রাক্কলিত মান নির্ণয়:

যেহেতু ${S_b}^2$ এর মান জানা থাকে না, তাই ${S_b}^2$ এর পরিবর্তে এর নিঝুকি প্রাক্কলক ${s_b}^2$ এর মান নমুনা থেকে নির্ণয় করে ভেদাংকের প্রাক্কলিত মান নিম্নরূপে পাওয়া যায়-

$$\text{Est. } V\left(\overline{y}_n\right)=\left(1-f\right)\frac{{s_b}^2}{n}$$

প্রশ্ন: সমান আকারের গুচ্ছ নমুনায়নের ক্ষেত্রে পুনঃস্থাপনবিহীন নমুনা চয়ন করা হলে তথ্যবিশ্ব সমষ্টির প্রাক্কলক ও প্রাক্কলকের ভেদাংক নির্ণয় কর।

অথবা, গুচ্ছ নমুনায়নে তথ্যবিশ্ব সমষ্টির একটি নিঝুঁকি নিরূপক নির্ণয় কর এবং সম-আকারের গুচ্ছায়নের ক্ষেত্রে এর ভেদাংকের সূত্রটি উদ্ভাবন কর **(জা.বি.-১৫)**

অথবা, M সংখ্যক এককবিশিষ্ট N-গুচ্ছের সমগ্রক হতে n গুচ্ছের একটি সরল দৈব নমুনা পুনঃস্থাপন ছাড়া নেয়ার ক্ষেত্রে দেখাও যে $\hat{Y}_c = NM\ \bar{y}$ হলো সমগ্রক সমষ্টি (Y) এর নিঝুঁকি প্রাক্কলক এবং এই প্রাক্কলকের ভেদাংক নিম্নরূপঃ

$$V\left(\hat{Y}_c\right) = \frac{(1-f)N^2 MS^2}{n}\left[1+(M-1)\rho\right]$$, যেখানে ρ হলো আন্তঃগুচ্ছ সংশ্লেষাংক **(জা.বি.-১৫)**

উত্তরঃ তথ্যবিশ্ব সমষ্টি, $Y = NM\bar{Y}$

জানি, $E(\bar{y}_n) = \bar{Y}$

$$\Rightarrow E(NM\ \bar{y}_n) = NM\ \bar{Y}$$

$$\Rightarrow E(NM\ \bar{y}_n) = Y$$

অতএব, $\hat{Y} = NM\ \bar{y}_n$ হলো তথ্যবিশ্ব সমষ্টি Y এর নিঝুঁকি প্রাক্কলক।

এখন, $V(\hat{Y}) = V(NM\ \bar{y}_n)$

$$= N^2 M^2\ V(\bar{y}_n)$$

$$= N^2 M^2 \frac{N-n}{N}\frac{{S_b}^2}{n}$$

$$= N^2 M^2 (1-f)\frac{{S_b}^2}{n}\quad\left[\because f = \frac{n}{N}\right]$$

$$= \frac{N^2 M^2 (1-f)}{n}\frac{S^2}{M}\left[1+(M-1)\rho\right]\quad \because {S_b}^2 = \frac{S^2}{M}\left[1+(M-1)\rho\right]$$

$$\therefore\quad V(\hat{Y}) = \frac{(1-f)N^2 M}{n}S^2\left[1+(M-1)\rho\right]$$

প্রশ্নঃ সরল দৈব নমুনায়নের তুলনায় গুচ্ছ নমুনায়নের আপেক্ষিক দক্ষতা নির্ণয় কর এবং মন্তব্য কর (জা.বি.-১২, ১৬)

সমাধানঃ তথ্যবিশ্ব হতে n M একক সরল দৈব নমুনার মাধ্যমে চয়ন করা হলে

$$V(\bar{y})_{ran} = \frac{(1-f)S^2}{nM}$$

আবার, গুচ্ছ নমুনায়নের ক্ষেত্রে,

$$V(\bar{y}_n)_C = \frac{1-f}{n}{S_b}^2$$

অতএব, সরল দৈব নমুনায়নের সংগে সমান আকার বিশিষ্ট গুচ্ছ হতে গুচ্ছ নমুনায়নের দক্ষতাকে নিম্নরূপে সংজ্ঞায়িত করা যায়-

$$E = \frac{V(\bar{y})_{ran}}{V(\bar{y}_n)_C}$$

যখন $E > 1$, অর্থাৎ $\frac{V(\bar{y})_{ran}}{V(\bar{y}_n)_C} > 1$

$$\Rightarrow V(\bar{y})_{ran} > V(\bar{y}_n)_C$$

$$\Rightarrow \frac{(1-f)S^2}{nM} > \frac{1-f}{n} S_b^{\,2}$$

$$\Rightarrow \frac{S^2}{M} > S_b^{\,2}$$

$$\Rightarrow \frac{MS_b^{\,2}}{M\left[1+(M-1)\rho\right]} > S_b^{\,2} \qquad \left[\because\ S^2 = \frac{M\ S_b^{\,2}}{1+(M-1)\ \rho}\right]$$

$$\Rightarrow \frac{1}{1+(M-1)\ \rho} > 1$$

$$\Rightarrow 1 > 1+(M-1)\ \rho$$

$$\Rightarrow 1+(M-1)\rho < 1$$

$$\Rightarrow (M-1)\rho < 0$$

$$\therefore\ \rho < 0$$

অর্থাৎ যদি $E > 1$ হয় তখন $\rho < 0$, তবে সরল দৈব নমুনায়নের চেয়ে গুচ্ছ নমুনায়ন অধিক দক্ষ হবে।

যদি $E < 1$ হয় তখন $\rho > 0$, তবে সরল দৈব নমুনায়নের চেয়ে গুচ্ছ নমুনায়ন কম দক্ষ হবে।

অন্যদিকে, যদি $E = 1$ হয় তখন $\rho = 0$, তবে উভয় নমুনায়নের দক্ষতা সমান হবে।

NB: We have, $E = \frac{V(\bar{y})_{ran}}{V(\bar{y}_n)_C}$

(a) If $V(\bar{y})_{ran} = V(\bar{y}_n)_C$, then $E = 1$ i.e., both sample is equally efficient.

(b) If $V(\bar{y})_{ran} > V(\bar{y}_n)_C$, then $E > 1$ i.e., cluster sample is more efficient.

(c) If $V(\bar{y})_{ran} < V(\bar{y}_n)_C$, then $E < 1$ i.e., cluster sample is less efficient.

প্রশ্ন: দেখাও যে, গুচ্ছ নমুনায়নের অসমান গুচ্ছের ক্ষেত্রে নমুনা গুচ্ছ গড়ের গড়, তথ্যবিশ্ব গড়ের নির্ঝুঁকি প্রাক্কলক অর্থাৎ $E(\bar{y}_C) = \bar{Y}$

উত্তর: ধরা যাক, তথ্যবিশ্বে N সংখ্যক গুচ্ছ আছে এবং i-তম গুচ্ছে M_i $(i = 1,\ 2, \cdots, N)$ সংখ্যক একক আছে এবং $\sum_{i=1}^{N} M_i = M_0$

i-তম গুচ্ছের গড়, $\bar{y}_{i\cdot} = \dfrac{1}{M_i}\sum_{j=1}^{M_i} y_{ij}$

i-তম গুচ্ছের সমষ্টি, $\sum_{j=1}^{M_i} y_{ij} = M_i\,\bar{y}_{i\cdot}$

তথ্যবিশ্ব গড়, $\bar{Y} = \dfrac{\sum_{i=1}^{N}\sum_{j=1}^{M_i} y_{ij}}{M_0} = \dfrac{\sum_{i=1}^{N} M_i\,\bar{y}_{i\cdot}}{M_0}$

ধরা যাক, তথ্যবিশ্ব হতে পুনঃস্থাপনবিহীন n গুচ্ছ নির্বাচন করা হয়েছে এবং নির্বাচিত প্রতি গুচ্ছের এককসমূহ হতে উপাত্ত সংগ্রহ করা হয়েছে। অসমান গুচ্ছের ক্ষেত্রে নমুনা গুচ্ছ গড়ের গড়কে নিম্নরূপে সংজ্ঞায়িত করা যায় -

$$\bar{y}_C = \frac{\sum_{i=1}^{n} M_i\,\bar{y}_{i\cdot}}{n\,\dfrac{M_0}{N}} = \frac{N}{n\,M_0}\sum_{i=1}^{n} M_i\,\bar{y}_{i\cdot}$$

এখন, $E(\bar{y}_C) = E\left(\dfrac{N}{n\,M_0}\sum_{i=1}^{n} M_i\,\bar{y}_{i\cdot}\right)$

$$= \frac{N}{n\,M_0}\sum_{i=1}^{n} E(M_i\,\bar{y}_{i\cdot})$$

$$= \frac{N}{n\,M_0}\sum_{i=1}^{n}\left[\sum_{i=1}^{N} M_i\,\bar{y}_{i\cdot} \times \frac{1}{N}\right]$$

$$= \frac{N}{n\,M_0}\; n\sum_{i=1}^{N} M_i\,\bar{y}_{i\cdot} \times \frac{1}{N}$$

$$= \frac{\sum_{i=1}^{N} M_i\,\bar{y}_{i\cdot}}{M_0}$$

$$= \bar{Y}$$

অতএব, গুচ্ছ নমুনায়নের অসমান গুচ্ছের ক্ষেত্রে নমুনা গুচ্ছ গড়ের গড়, তথ্যবিশ্ব গড়ের নিঝুঁকি প্রাক্কলক।

বি. দ্র. (i) তথ্যবিশ্বের প্রতি গুচ্ছে গড়ে উপাদান সংখ্যা (the average number of elements per cluster in the population is)

$$\bar{M} = \frac{\sum_{i=1}^{N} M_i}{N} = \frac{M_0}{N}$$

(ii) N টি গুচ্ছে মোট একক সংখ্যা M_0

$\therefore$ 1 ” ” ” ” ” $\frac{M_0}{N}$

$\therefore$ n ” ” ” ” ” $n\frac{M_0}{N}$

(iii) ধরি, নমুনা গুচ্ছের সংখ্যা, $n = 3$

1^{st} cluster	2^{nd} cluster	3^{rd} cluster
1	4	7
2	5	8
3	6	9
Total = 6	Total = 15	Total = 24
Average = 2	Average = 5	Average = 8

n আকারের নমুনা গুচ্ছের জন্য গুচ্ছ গড়ের গড়, $\bar{y}_c = \frac{2+5+8}{3} = \frac{15}{3} = 5$

অন্যভাবে, n আকারের নমুনা গুচ্ছের জন্য গুচ্ছ গড়ের গড়, $\bar{y}_c = \frac{6+15+24}{9} = \frac{45}{9} = 5$

(iv) অসমান গুচ্ছের ক্ষেত্রে-

1	2	....	i	....	N
y_{11}	y_{21}	....	y_{i1}	....	y_{N1}
y_{12}	y_{22}	....	y_{i2}	....	y_{N2}
....	....			....	....
y_{1M_1}	y_{2M_2}	....	$y_{i\,M_i}$	....	y_{NM_N}
$\bar{y}_{1\cdot} = \frac{y_{1\cdot}}{M_1}$	$\bar{y}_{2\cdot} = \frac{y_{2\cdot}}{M_2}$	....	$\bar{y}_{i\cdot} = \frac{y_{i\cdot}}{M_i}$	....	$\bar{y}_{N\cdot} = \frac{y_{N\cdot}}{M_N}$
$y_{1\cdot} = M_1\ \bar{y}_{1\cdot}$	$y_{2\cdot} = M_2\ \bar{y}_{2\cdot}$	....	$y_{i\cdot} = M_i\ \bar{y}_{i\cdot}$	....	$y_{N\cdot} = M_N\ \bar{y}_{N\cdot}$

If we select $n = 2$ clusters out of N clusters through srswor then

$$\bar{y}_C = \frac{y_{1\cdot} + y_{2\cdot}}{2\frac{M_0}{N}} = \frac{M_1\ \bar{y}_{1\cdot} + M_2\ \bar{y}_{2\cdot}}{2\frac{M_0}{N}} = \frac{\sum_{i=1}^{2} M_i\ \bar{y}_{i\cdot}}{2\frac{M_0}{N}} = \frac{\sum_{i=1}^{n} M_i\ \bar{y}_{i\cdot}}{n\ \frac{M_0}{N}}$$

সংক্ষিপ্ত প্রশ্ন ও উত্তরঃ

(১) গুচ্ছের আকার (size of a cluster) বলতে কি বুঝ ? (জা.বি.-১২)

উত্তরঃ গুচ্ছ নমুনায়নে প্রথমে তথ্যবিশ্বের এককসমূহকে কতগুলো গ্রুপে ভাগ করা হয়, যাদের বলা হয় গুচ্ছ। গুচ্ছগুলো এমনভাবে তৈরী করা হয় যেন প্রতিটি একক কোন না কোন গুচ্ছে একবার এবং মাত্র একবার অন্তর্ভূক্ত হবে। একই গুচ্ছের এককগুলো যতদূর সম্ভব বৈসাদৃশ্যপূর্ণ হবে এবং ভিন্ন গুচ্ছের এককগুলো সাদৃশ্যপূর্ণ হবে।

প্রতিটি গুচ্ছে যতটি একক বা সংখ্যা থাকে, তাই গুচ্ছের আকার। এককের সংখ্যার দিক থেকে গুচ্ছগুলোর আকার ভিন্ন হতে পারে, আবার সকল গুচ্ছের আকার অভিন্নও হতে পারে।

(২) কখন গুচ্ছ নমুনায়ন ব্যবহার করা হয়? (জা.বি.-১১)

অথবা, গুচ্ছ নমুনায়ন ব্যবহারের কারণসমূহ কি? **(জা.বি.-১১, ১৫)**

অথবা, কোন কোন ক্ষেত্রে গুচ্ছ নমুনা অধিকতর গ্রহণযোগ্য তা আলোচনা কর **(জা.বি.-১২)**

অথবা, গুচ্ছ নমুনায়ন ব্যবহৃত হয় এমন কিছু ক্ষেত্র উল্লেখ কর **(জা.বি.-১৬)**

অথবা, কোন কোন ক্ষেত্রে গুচ্ছ নমুনায়ন ভাল?

উত্তরঃ নিম্নলিখিত ক্ষেত্রে গুচ্ছ নমুনা অধিকতর গ্রহণযোগ্য-

(ক) যদি তথ্যবিশ্বের এককসমূহের সম্পূর্ণ এবং হালনাগাদ নমুনা কাঠামো জানা না থাকে এবং তা তৈরী করাও ব্যয় সাপেক্ষ, সেক্ষেত্রে গুচ্ছ নমুনায়ন ব্যবহার করা হয়।

(খ) যদি তথ্যবিশ্বের এককগুলোকে সমজাতীয় স্তরে বিভক্ত করা না যায়।

(গ) বৃহৎ আকারের নমুনা জরিপের অর্থাৎ তথ্যবিশ্ব বৃহৎ হলে ব্যয় ও যথার্থতার নিরীখে গুচ্ছ নমুনায়ন সার্থকভাবে ব্যবহৃত হয়।

(৩) গুচ্ছ ও স্তর এর মধ্যে পার্থক্য নির্ণয় কর (জা.বি.-১১)

অথবা, গুচ্ছ নমুনায়ন ও স্তরিত দৈব নমুনায়নের মধ্যে পার্থক্য লিখ **(জা.বি.-১৬)**

উত্তরঃ গুচ্ছ নমুনায়নে প্রথমে তথ্যবিশ্বের এককসমূহকে কতগুলো গ্রুপে ভাগ করা হয়, যাদের বলা হয় গুচ্ছ। অন্যদিকে, যে সমগ্রক থেকে নমুনা নেয়া হবে সেই সমগ্রকের এককগুলো যদি বিভিন্ন বৈশিষ্ট্যের হয়, তবে বৈশিষ্ট্যের সমমাত্রিকতা অনুসারে এককগুলোকে বিভিন্ন শ্রেণীতে বা দলে বিভক্ত করা হলে প্রতিটি শ্রেণীকে এক একটি স্তর বলা হয়।

আবার, একই গুচ্ছের এককগুলো যতদূর সম্ভব বৈসাদৃশ্যপূর্ণ হবে এবং ভিন্ন গুচ্ছের এককগুলো সাদৃশ্যপূর্ণ হবে। কিন্তু প্রতিটি স্তরের এককগুলো হবে যথাসম্ভব সমমাত্রিক এবং স্তরে স্তরে ব্যবধান হবে অসমমাত্রিক।

জাতীয় বিশ্ববিদ্যালয়ের বিভিন্ন সালের প্রশ্নাবলী

পরিসংখ্যান-২০১২ [নতুন সিলেবাস অনুযায়ী]
বিষয় কোডঃ 3664
(Sampling Technique)
সময়-৪ ঘণ্টা
পূর্ণমান-৮০
[**দ্রষ্টব্যঃ-**ধারাবাহিকভাবে প্রশ্নের উত্তর দাও]

ক বিভাগ
(যে কোন **দশটি** সংক্ষিপ্ত প্রশ্নের উত্তর দাও)
মান- ১ x ১০ = ১০

১। (ক) পরামিতি কাকে বলে ?
(খ) নিরূপক বলতে কি বুঝ ?
(গ) প্রশ্নমালা বলতে কি বুঝ ?
(ঘ) কখন স্তরিত দৈব নমুনায়ন ব্যবহার করা হয় ?
(ঙ) নমুনায়ন ভগ্নাংশ কি ?
(চ) এমন একটি ক্ষেত্র উল্লেখ কর, যেখানে নমুনা জরিপ অপেক্ষা শুমারী জরিপ অধিক উপযুক্ত।
(ছ) লা-জবাব (non-response) বলতে কি বুঝায় ?
(জ) সমানুপাতিক বণ্টন (proportional allocation) কি ?
(ঝ) কখন নেম্যান বণ্টন (Neyman's allocation) ও সমানুপাতিক বণ্টন একই হয় ?
(ঞ) কখন চক্রাকার ধারাবাহিক নমুনায়ন (circular systematic sampling) ব্যবহার করা হয় ?
(ট) স্তরিত দৈব নমুনায়নের ক্ষেত্রে তথ্যবিশ্বের সমষ্টির নির্ঝুঁকি নিরূপকটি (unbiased estimate of population total) লিখ।
(ঠ) গুচ্ছের আকার (size of a cluster) বলতে কি বুঝ ?

খ বিভাগ
(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)
মান- ৪ x ৫ = ২০

২। দৈব নমুনায়ন ও ঐচ্ছিক নমুনায়ন বলতে কি বুঝ ? এদের মধ্যে পার্থক্য কর।

৩। নমুনা জরিপ কি ? শুমারীর তুলনায় নমুনা জরিপের সুবিধাসমূহ বর্ণনা কর।

৪। সরল ধারাবাহিক ও চক্রাকার ধারাবাহিক নমুনা নির্বাচনের পদ্ধতি আলোচনা কর।

৫। 'স্তরিত দৈব নমুনায়ন' পদ্ধতি বর্ণনা কর। কোন কোন শর্তে স্তরিত দৈব নমুনায়ন সরল দৈব নমুনায়ন অপেক্ষা পছন্দনীয় ?

৬। পুনঃস্থাপনবিহীন সরল দৈব নমুনায়নের ক্ষেত্রে প্রমাণ কর যে, $E(s^2) = S^2$; প্রতীকগুলো চিরাচরিত।

৭। গুচ্ছ নমুনা কি ? কোন কোন ক্ষেত্রে গুচ্ছ নমুনা অধিকতর গ্রহণযোগ্য তা আলোচনা কর।

৮। অনুপাত প্রাক্কলকের বৈশিষ্ট্যগুলো (properties of ratio estimator) উল্লেখ কর।

৯। সরল দৈব নমুনায়নের তুলনায় গুচ্ছ নমুনায়নের আপেক্ষিক দক্ষতা (relative efficiency) নির্ণয় কর এবং মন্তব্য কর।

গ বিভাগ

(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ১০ X ৫ = ৫০

১০। (ক) একটি নমুনা জরিপের মুলনীতিসমূহ আলোচনা কর।

(খ) নমুনাজ ত্রুটি ও অনমুনাজ ত্রুটির পার্থক্য লিখ। কিভাবে অনমুনাজ ত্রুটি নিয়ন্ত্রণ করা যায় ?

১১। (ক) স্তরিত নমুনায়নে বণ্টন সমস্যা বলতে কি বুঝ ? উক্ত সমস্যা কিভাবে সমাধান করা যায় ?

(খ) প্রচলিত সংকেতে প্রমাণ কর যে, $V\left(\bar{y}_{st}\right)_{opt} \leq V\left(\bar{y}_{st}\right)_{prop} \leq V\left(\bar{y}\right)_{ran}$

১২। (ক) দেখাও যে, ধারাবাহিক নমুনায়ন সরল দৈব নমুনায়নের চেয়ে ভাল প্রাক্কলক প্রদান করবে, যদি $S^2_{wsy} > S^2$ হয় (প্রতীকগুলো চিরাচরিত)

(খ) ধারাবাহিক নমুনায়নে আন্তঃশ্রেণি সংশ্লেষাংক ρ_w হলে, দেখাও যে

$$V\left(\bar{y}_{sy}\right) = \frac{N-1}{N} \cdot \frac{S^2}{n}\left[1+(n-1)\rho_w\right]$$

১৩। (ক) গুচ্ছ নমুনায়নের সংজ্ঞা দাও। অন্যান্য নমুনায়নের তুলনায় গুচ্ছ নমুনায়নের সুবিধাগুলো আলোচনা কর।

(খ) ধরা যাক, N গুচ্ছবিশিষ্ট তথ্যবিশ্বের প্রতিটি গুচ্ছে M সংখ্যক একক আছে। n গুচ্ছের একটি সরল দৈব নমুনা পুনঃস্থাপন ব্যতিরেকে নেয়া হল। যদি নমুনার প্রতি এককের গড় $\bar{y}_c$ হয়, তবে দেখাও যে, $\bar{y}_c$ হল সমগ্রক গড়ের নির্ঝুঁকি প্রাক্কলক এবং $V(\bar{y}_c) = \frac{N-n}{nMN} . S^2\left[1+(M-1)\rho\right]$; যেখানে ρ হল আন্তঃগুচ্ছ সংশ্লেষাংক।

১৪। (ক) নির্ভরণ নিরূপকের সংজ্ঞা দাও। অনুপাত নিরূপক ও নির্ভরণ নিরূপকের পার্থক্য লিখ।

(খ) সরল দৈব নমুনায়নে বড় আকারের নমুনার জন্য, প্রচলিত প্রতীকে প্রমাণ কর যে,

$$V(\hat{\bar{Y}}_{lr}) = \frac{N-n}{Nn} . S_y{}^2\left(1-\rho^2\right)$$

১৫। (ক) অনুপাত প্রাক্কলক বলতে কি বুঝ ? তথ্যবিশ্ব সমষ্টির অনুপাত প্রাক্কলকের ভেদাংক নির্ণয় কর।

(খ) নমুনা হতে উক্ত ভেদাংক কিভাবে প্রাক্কলন করা যায় ?

১৬। (ক) N আকারের তথ্যবিশ্ব হতে n আকারের সরল দৈব নমুনা চয়ন করা হলে, দেখাও যে, একটি নির্দিষ্ট এককের নমুনায় অন্তর্ভূক্ত হবার সম্ভাবনা $\frac{n}{N}$ ।

(খ) সরল দৈব নমুনায়নের ক্ষেত্রে দেখাও যে, নমুনার অনুপাত (p) তথ্যবিশ্ব অনুপাত (P) এর নির্ঝুঁকি প্রাক্কলক এবং

$$V(p) = \frac{N-n}{N-1} . \frac{PQ}{n}$$

১৭। (ক) প্রশ্নমালা (questionnaire) কি ? একটি আদর্শ প্রশ্নমালার আবশ্যকীয় গুণগুলি উল্লেখ কর।

(খ) খোলা প্রশ্নমালা (closed questionnaire) ও আবদ্ধ প্রশ্নমালার (open questionnaire) মধ্যে পার্থক্য দেখাও। বাংলাদেশের পরিপ্রেক্ষিতে প্রশ্নমালার মাধ্যমে তথ্য সংগ্রহের অসুবিধাগুলি সম্পর্কে আলোচনা কর।

পরিসংখ্যান-২০১৩

বিষয় কোডঃ 3664

(Sampling Technique)

সময়-৪ ঘণ্টা

পূর্ণমান-৮০

[**দ্রষ্টব্যঃ-**একই বিভাগের প্রশ্নের বিভিন্ন অংশের উত্তর ধারাবাহিকভাবে লিখতে হবে]

ক বিভাগ

(যে কোন **দশটি** সংক্ষিপ্ত প্রশ্নের উত্তর দাও)

মান- ১ x ১০ = ১০

১। (ক) নমুনাজের সংজ্ঞা দাও।

[Define statistic.]

(খ) সমগ্রকের সংজ্ঞা দাও।

[Define population.]

(গ) পাইলট জরিপ কী ?

[What is pilot survey?]

(ঘ) স্তর কি ?

[What is stratum?]

(ঙ) দৈব নমুনায়ন কি ?

[What is random sampling?]

(চ) উদ্দেশ্যমুখী নমুনায়নের একটি উদাহরণ দাও।

[Give an example of purposive sampling.]

(ছ) নমুনার আকার বলতে কি বুঝ ?

[What do you mean by sample size?]

(জ) শুমারী জরিপ কি ?

[What is census?]

(ঝ) সসীম সমগ্রক শুদ্ধি বলতে কি বুঝ ?

[What do you mean by finite population correction (fpc)?]

(ঞ) অনুপাত প্রাক্কলক কখন ব্যবহৃত হয় ?

[When ratio estimator is used?]

(ট) নমুনা জরীপ কেন করা হয় ?

[What are the reasons for sample survey?]

(ঠ) কখন গুচ্ছ নমুনায়ন ব্যবহৃত হয় ?

[When cluster sampling is used?]

খ বিভাগ

(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ৪ x ৫ = ২০

২। নমুনা কাঠামো কি ? একটি কাঠামোর কি কি ত্রুটি থাকতে পারে-সংক্ষেপে লিখ।

[What is sampling frame? Write down the defects of a frame.]

৩। প্রশ্নমালা কি ? বিভিন্ন প্রকার প্রশ্নমালা সম্পর্কে আলোচনা কর।

[What is questionnaire? Discuss different types of questionnaire.]

৪। সরল দৈব নমুনায়ন বলতে কি বুঝ ? দেখাও যে, পুনঃস্থাপন ছাড়া সরল দৈব নমুনায়নে N আকারের সমগ্রক হতে দুটি নির্দিষ্ট এককের n আকারের নমুনায় অন্তর্ভুক্তির সম্ভাবনা $\frac{n(n-1)}{N(N-1)}$ ।

[What do you mean by simple random sampling? Show that for SRSWOR, probability of inclusion of any two specified units in the sample of size n from a population of size N is $\frac{n(n-1)}{N(N-1)}$.]

৫। স্তরিত নমুনায়নের ক্ষেত্রে সমগ্রক গড়ের নিঝুঁকি প্রাক্কলক নির্ণয় কর এবং এই প্রাক্কলকের ভেদাংক সমানুপাতিক বণ্টনের ক্ষেত্রে বের কর।

[Find the unbiased estimator of population mean in case of stratified random sampling and find also the variance of this estimator in case of proportional allocation.]

৬। ধারাবাহিক নমুনায়ন বলতে কি বুঝ ? ধারাবাহিক নমুনায়ন এর সাথে স্তরিত দৈব নমুনায়ন কিভাবে সংযুক্ত করবে ?

[What is meant by systematic sampling? How systematic sampling is related to stratified random sampling?]

৭। গুচ্ছ নমুনায়ন কি ? স্তরিত নমুনায়নের সাথে গুচ্ছ নমুনায়নের তুলনা কর।

[What is cluster sampling? Compare cluster sampling with stratified sampling.]

৮। নমুনা জরিপে ত্রুটির উৎসসমূহ সংক্ষেপে আলোচনা কর।

[Discuss briefly the sources of errors in sample survey.]

৯। অনুপাত নিরূপক কি ? দেখাও যে, $B(\hat{R}) = \frac{-\operatorname{cov}(\hat{R}\,\bar{x})}{\bar{x}}$ ।

[What is ratio estimator? Show that $B(\hat{R}) = \frac{-\operatorname{cov}(\hat{R}\,\bar{x})}{\bar{x}}$.]

গ বিভাগ

(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ১০ x ৫ = ৫০

১০। (ক) নমুনা জরিপে ধাপগুলি সংক্ষেপে আলোচনা কর।

(খ) শুমারি জরিপ ও নমুনা জরিপের পার্থক্যগুলি লিখ।

[(a) Briefly discuss the steps involved in sample survey.

(b) Distinguish between census and sample survey.]

১১। (ক) পুনঃস্থাপনসহ ও পুনঃস্থাপন ছাড়া নমুনায়ন বলতে কি বুঝ ? এদের মধ্যকার পার্থক্য লিখ।

(খ) পুনঃস্থাপন ছাড়া সরল দৈব নমুনায়নের ক্ষেত্রে সচরাচর সংকেতমালায় দেখাও যে,

$$E(\bar{y}) = \bar{Y} \text{ এবং } V(\bar{y}) = \frac{N-n}{N} \cdot \frac{S^2}{n}$$

[(a) What do you mean by sampling with replacement and without replacement? Distinguish between them.

(b) In simple random sampling without replacement, with usual notations show that

$$E(\bar{y}) = \bar{Y} \text{ and } V(\bar{y}) = \frac{N-n}{N} \cdot \frac{S^2}{n}]$$

১২। (ক) ব্যবহারিক ক্ষেত্রে কিভাবে স্তরীকরণ করা হয় তা উদাহরণের সাহায্যে ব্যাখ্যা কর। সরল দৈব নমুনায়নের চেয়ে স্তরিত নমুনায়নের সুবিধাসমূহ কি কি ?

(খ) প্রমাণ কর যে, $V(\bar{Y}_{st}) = \frac{1}{n}\left[\sum_{h=1}^{L} W_N S_N\right]^2 - \frac{1}{N}\sum_{h=1}^{L} W_N S_N{}^2$

[(a) Illustrate with example how stratification can be done in practice. What are the advantages of stratified random sampling over simple random sampling?

(b) Prove that $V(\bar{Y}_{st}) = \frac{1}{n}\left[\sum_{h=1}^{L} W_N S_N\right]^2 - \frac{1}{N}\sum_{h=1}^{L} W_N S_N{}^2$]

১৩। (ক) ধারাবাহিক নমুনায়ন কি ? এ নমুনায়নের সুবিধা ও অসুবিধাগুলি লিখ।

(খ) সরল রৈখিক ধারাযুক্ত সমগ্রকের ক্ষেত্রে প্রচলিত প্রতীকে দেখাও যে, $V_{st} : V_{sys} : V_{srs} = \frac{1}{n} : 1 : n$

[(a) What is systematic sampling? Write down the advantages and disadvantages of this sampling method.

(b) In case of populations with linear trend in usual notations, show that $V_{st} : V_{sys} : V_{srs} = \frac{1}{n} : 1 : n$]

১৪। (ক) অনুপাত ও নির্ভরণ নিরূপকের ব্যবহারগুলো আলোচনা কর। অনুপাত ও নির্ভরণ নিরূপকের বৈশিষ্ট্যগুলো লিখ।

(খ) প্রমাণ কর যে, $V(\hat{Y}_R) = \frac{1-f}{n} Y^2 \left[\frac{S_y{}^2}{\bar{Y}^2} + \frac{S_x{}^2}{\bar{x}^2} - 2\frac{S_{xy}}{\bar{X}\bar{Y}}\right]$

[(a) Discuss the uses of ratio and regression estimators. Write the properties of ratio and regression estimator.

(b) Show that $V(\hat{Y}_R) = \frac{1-f}{n} Y^2 \left[\frac{S_y{}^2}{\bar{Y}^2} + \frac{S_x{}^2}{\bar{x}^2} - 2\frac{S_{xy}}{\bar{X}\bar{Y}}\right]$

১৫। (ক) একটি নিরূপকের পক্ষপাত, নির্ভুলতা এবং উৎকর্ষতার সংজ্ঞা দাও। নমুনার আকার নির্ধারণ আলোচনা কর।

(খ) নমুনা জরিপে non-response এর সংজ্ঞা দাও। নিরূপকের ক্ষেত্রে non-response এর প্রভাব আলোচনা কর।

[(a) Define bias, accuracy and precision of an estimator. Discuss the procedure of sample size determination.

(b) Define non-response in sample survey. Discuss the effects of non-response on the estimator.]

১৬। (ক) গুচ্ছ নমুনায়ন ব্যবহৃত হয় এমন কিছু ক্ষেত্র উল্লেখ কর।

(খ) M সংখ্যক একক বিশিষ্ট N গুচ্ছের সমগ্রক হতে n গুচ্ছের একটি সরল দৈব নমুনা পুনঃস্থাপন ছাড়া নেয়ার ক্ষেত্রে দেখাও যে, $\hat{Y}_c = NM\bar{y}$ হলো সমগ্রক সমষ্টি (Y) এর নিঝুঁকি প্রাক্কলক এবং এই প্রাক্কলকের ভেদাংক নিম্নরূপঃ

$$V\left(\hat{Y}_c\right) = N^2 M \frac{(1-f)}{n} S^2 \left[1+(M-1)\rho\right]$$

[(a) Mention some situations in which cluster sampling is used.

(b) In SRSWOR of n cluster each containing M elements from a population of N cluster. Show that $\hat{Y}_c = NM\bar{y}$ is an unbiased estimate of population total (Y) and its variance is given by $V\left(\hat{Y}_c\right) = N^2 M \frac{(1-f)}{n} S^2 \left[1+(M-1)\rho\right]$; where ρ is the intra-cluster correlation co-efficient.]

১৭। (ক) নির্ভরণ নিরূপকের সংজ্ঞা দাও। সরল দেব নমুনায়নে b_o প্রদত্ত ধ্রুবক হলে প্রমাণ কর যে, রৈখিক নির্ভরণ প্রাক্কলক $\bar{Y}_{lr} = \bar{Y} + b_0(\bar{X} - \bar{x})$ নিরপেক্ষ নিরূপক।

(খ) বড় আকারের নমুনার ক্ষেত্রে নির্ভরণ প্রাক্কলক, অনুপাত প্রাক্কলক ও সরল প্রাক্কলকের তুলনা কর।

[(a) Define regression estimator. Prove that for simple random sampling in which b_o is pre-assigned constant, the linear regression estimate $\bar{Y}_{lr} = \bar{Y} + b_0(\bar{X} - \bar{x})$ is unbiased.

(b) For large sample, compare regression estimator with ratio estimator and simple estimator.]

পরিসংখ্যান (তত্ত্বীয়)-২০১৪
বিষয় কোডঃ 3664
(Sampling Technique)
সময়-৪ ঘণ্টা
পূর্ণমান-৮০
[**দ্রষ্টব্যঃ-**একই বিভাগের প্রশ্নের বিভিন্ন অংশের উত্তর ধারাবাহিকভাবে লিখতে হবে]

ক বিভাগ
(যে কোন **দশটি** সংক্ষিপ্ত প্রশ্নের উত্তর দাও)
মান- ১ x ১০ = ১০

১। (ক) পরামিতির সংজ্ঞা দাও।
[Define parameter.]
(খ) নমুনা একক কি ?
[What is sampling unit ?]
(গ) নমুনায়ন বলতে কি বুঝ ?
[What do you mean by sampling ?]
(ঘ) প্রাক্কলকের সংজ্ঞা দাও ।
[Define estimator.]
(ঙ) একটি প্রাক্কলকের নিরপেক্ষতা বলতে কি বুঝ ?
[What do you mean by unbiasedness of an estimator ?]
(চ) পরিমিত ব্যবধান ও আদর্শ বিচ্যুতির মধ্যে পার্থক্য কী ?
[What is the difference between standard deviation and standard error ?]
(ছ) প্রাক্ জরিপ কি ?
[What is pilot survey ?]
(জ) ধারাবাহিক নমুনায়নের অসুবিধাগুলি কি কি ?
[What are the disadvantages of systematic sampling?]
(ঝ) নমুনায়ন কাঠামো বলতে কি বুঝ ?
[What is meant by sapling frame ?]
(ঞ) প্রশ্নমালা কি ?
[What is questionnaire ?]
(ট) নির্ভরণ প্রাক্কলকের বৈশিষ্ট্যগুলো লিখ ।
[Write down the characteristics of regression estimator]
(ঠ) কখন স্তরিত দৈব নমুনায়ন ব্যবহার করা হয় ?
[When stratified random sampling is used ?]

খ বিভাগ
(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)
মান- ৪ x ৫ = ২০

২। নমুনা জরিপ বলতে কি বুঝ ? নমুনা জরিপের মূলনীতি আলোচনা কর।

[What do you mean by sample survey ? Discuss the basic principle of sample survey.]

৩। প্রশ্নমালা ও তালিকার মধ্যে পার্থক্য কি ? একটি প্রশ্নমালা তৈরি করতে যে সকল প্রধান দিক বিবেচনায় আনা হয়, তা আলোচনা কর।

[Distinguish between questionnaire and schedule? What are the main points to be taken in consideration for preparing a questionnaire ?]

৪। অনমুনাজ ত্রুটি কি ? এটি কমানোর উপায়গুলো কি কি ? [What is non-sampling error ? What are the ways of reducing non-sampling error ?]

৫। নিরপেক্ষ নিরূপক বলতে কি বুঝ ? দেখাও যে, সরল দৈব নমুনায়নে নমুনা গড় সমগ্রক গড়ের নিরপেক্ষ নিরূপক।

[What do you mean by unbiased estimator? Show that in simple random sapling sample mean is an unbiased estimator of population mean.]

৬। দৈব নমুনায়ন বলতে কি বুঝ ? দৈব সংখ্যা সারণি ব্যবহার করে কিভাবে সরল দৈব নমুনা নির্বাচন করা যায় ?

[What do you mean by random sampling? How do a simple random sample can be selected using random number table ?]

৭। নিয়মক্রমিক নমুনায়নের সংজ্ঞা দাও। সরলরৈখিক ও চক্রক্রমিক নিয়মক্রমিক নমুনায়নের মধ্যে পার্থক্য নির্দেশ কর।

[Define systematic sampling. Distinguish between linear and circular systematic sampling.]

৮। গুচ্ছ ও স্তরের মধ্যে পার্থক্য নির্দেশ কর। গুচ্ছ নমুনায়নের সুবিধা-অসুবিধা আলোচনা কর।

[Distinguish between cluster and strata. Discuss the advantages and disadvantages of cluster sampling.]

৯। নির্ভরণ নিরূপকের চেয়ে অনুপাত নিরূপক ভাল হওয়ার শর্তটি প্রতিষ্ঠা কর।

[Establish the condition for which ratio estimator is superior to regression estimator.]

গ বিভাগ

(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ১০ X ৫ = ৫০

১০। (ক) সরল দৈব নমুনা চয়নের বিভিন্ন পদ্ধতি আলোচনা কর।

(খ) দেখাও যে, পুনঃস্থাপন ব্যতিরেকে সরল দৈব নমুনায়নে একটি নির্দিষ্ট একক যে কোন চয়নে নমুনায় অন্তর্ভূক্ত হওয়ার সম্ভাবনা, এককটি প্রথম চয়নে নমুনায় অন্তর্ভূক্ত হওয়ার সম্ভাবনার সমান।

[(a) Discuss the various method of drawing simple random sample.

(b) Show that, in simple random sampling without replacement, the probability of including a unit in the sample in any draw is equal to the probability of including the unit in the sample in first draw.]

১১। (ক) সরল দৈব নমুনা কি ? সরল দৈব নমুনায়নের জন্য দেখাও যে, নমুনা ভেদাংক s^2 সমগ্রক ভেদাংক S^2 এর নিরপেক্ষ নিরূপক।

(খ) অনুপাতের জন্য সরল দৈব নমুনায়ন বলতে কি বুঝ ? সরল দৈব নমুনায়নে সমগ্রক অনুপাতের একটি নিরপেক্ষ নিরূপক বের কর এবং তার ভেদাংক নির্ণয় কর।

[(a) What is simple random sample ? Show that, in simple random sampling sample variance s^2 is an unbiased estimate of population variance S^2.

(b) What do you mean by simple random sampling for proportions? For simple random sapling find an unbiased estimate of population proportion and determine its variance]

১২। (ক) স্তরীকরণ বলতে কি বুঝ ? স্তরিত দৈব নমুনায়নে ব্যবহৃত নমুনা বণ্টনের বিভিন্ন পদ্ধতিসমূহ বর্ণনা কর।

(খ) প্রচলিত সংকেতে দেখাও যে, $V(\bar{y}_{st})_{opt} \le V(\bar{y}_{st})_{prop} \le V(\bar{y})_{ran}$

[(a) What do you mean by stratification? Describe different types of sample allocations used in stratified random sampling.

(b) For usual notations show that, $V(\overline{y}_{st})_{opt} \leq V(\overline{y}_{st})_{prop} \leq V(\overline{y})_{ran}$]

১৩। (ক) ধারাবাহিক নমুনায়ন কি? ধারাবাহিক নমুনায়নে সমগ্রক গড়ের নিরপেক্ষ নিরূপক এবং তার ভেদাংক বের কর।

(খ) ধারাবাহিক নমুনায়নে নমুনা গড়ের ভেদাংকের সাথে সরল দৈব নমুনায়নে নমুনা গড়ের ভেদাংকের তুলনা কর।

[(a) What is systematic sampling? In systematic sampling find the unbiased estimator of population mean and find its variance.

(b) Compare the variance of sample mean by systematic sampling with that of simple random sampling.]

১৪। (ক) দেখাও যে, পুনঃস্থাপন ছাড়া সরল দৈব নমুনায়নের মাধ্যমে N গুচ্ছবিশিষ্ট সমগ্রক হতে n গুচ্ছ চয়ন করা হলে নমুনা গড় $\overline{y}_c$ সমগ্রক গড় $\overline{Y}$ এর নিরপেক্ষ নিরূপক এবং $v(\overline{y}_c) = \frac{1-f}{nM} S^2[1+(M-1)\rho]$, যেখানে ρ অন্তঃশ্রেণি সংশ্লেষাংক এবং M গুচ্ছের আকার।

(খ) সরল দৈব নমুনায়নের সাথে গুচ্ছ নমুনায়নের আপেক্ষিক দক্ষতার তুলনা কর।

[(a) Show that, if a sample of n clusters is drawn from a population of N clusters, then the sample mean $\overline{y}_c$ is an unbiased estimator of population mean $\overline{Y}$ and, $v(\overline{y}_c) = \frac{1-f}{nM} S^2[1+(M-1)\rho]$, where, ρ is intra-class correlation and M is size of cluster.

(b) Compare the relative efficiency of cluster sampling with that of simple random sampling.]

১৫। (ক) অনুপাত প্রাক্কলকের সংজ্ঞা দাও। দেখাও যে, নমুনা আকার n বড় হলে সমগ্রক অনুপাত $R = \frac{Y}{X}$ এর প্রাক্কলকের ভেদাংক, $v(\hat{R}) = \frac{1-f}{n\overline{X}^2} \cdot \frac{1}{N-1} \sum_{i=1}^{N} (y_i - Rx_i)^2$ ।

(খ) দেখাও যে, $v(\hat{R}) = (1-f)\frac{R^2}{n}\left[C_{yy} + C_{xx} - 2C_{yx}\right]$, যেখানে, $f = \frac{n}{N}$, C_{yy} ও C_{xx} যথাক্রমে y_i ও x_i এর বিভেদাংকের বর্গ এবং C_{yx} আপেক্ষিক সহ-ভেদাংক।

[(a) Define ratio estimator. Show that, if sample size n is large then the variance of the estimate of population ratio $R = \frac{Y}{X}$ is $v(\hat{R}) = \frac{1-f}{n\overline{X}^2} \cdot \frac{1}{N-1} \sum_{i=1}^{N} (y_i - Rx_i)^2$.

(b) Show that, $v(\hat{R}) = (1-f)\frac{R^2}{n}\left[C_{yy} + C_{xx} - 2C_{yx}\right]$, where $f = \frac{n}{N}$, C_{yy} and C_{xx} are the square of co-efficient of variation of y_i and x_i respectively and C_{yx} is the relative covariance.]

১৬। (ক) নমুনা জরিপে প্রাক্কলনের বিভিন্ন পদ্ধতি আলোচনা কর। কোন অবস্থায় কোন পদ্ধতি উপযুক্ত তা আলোচনা কর।

(খ) দেখাও যে, নির্ভরণ প্রাক্কলক $\overline{y}_{ir}$ সমগ্রক গড় $\overline{Y}$ এর নিরপেক্ষ প্রাক্কলক এবং $v(\overline{y}_{ir}) = \frac{1-f}{n}\left(1-\rho^2 S_y{}^2\right)$।

[(a) Discuss the various method of estimation in sample survey. Discuss which method is applicable in which situation.

(b) Show that, regression estimator $\overline{y}_{ir}$ is the unbiased estimator of population mean $\overline{Y}$ and $v(\overline{y}_{ir}) = \frac{1-f}{n}\left(1-\rho^2 S_y{}^2\right)$.]

১৭। (ক) সম্ভাবনা নমুনায়ন এবং উদ্দেশ্যমূলক নমুনায়নের মধ্যে পার্থক্য নির্দেশ কর। সরল দৈব নমুনায়নের সুবিধা-অসুবিধা আলোচনা কর।

(খ) সরল দৈব নমুনায়নের ক্ষেত্রে সমগ্রক সমষ্টির একটি নিরপেক্ষ নিরূপক বের কর এবং তার আদর্শ বিচ্যুতি বের কর।

[(a) Distinguish between probability sampling and purposive sampling. Discuss the advantages and disadvantages of simple random sampling.

(b) For simple random sampling find an unbiased estimate of population total and find its standard error.]

পরিসংখ্যান (তত্ত্বীয়)-২০১৫

(Sampling Technique)

বিষয় কোডঃ 223603

সময়-৪ ঘণ্টা

পূর্ণমান-৮০

[**দ্রষ্টব্যঃ-**একই বিভাগের প্রশ্নের বিভিন্ন অংশের উত্তর ধারাবাহিকভাবে লিখতে হবে]

ক বিভাগ

(যে কোন **দশটি** সংক্ষিপ্ত প্রশ্নের উত্তর দাও)

মান- ১ x ১০ = ১০

১। (ক) সমগ্রকের সংজ্ঞা দাও।
[Define population.]

(খ) নমুনাজ মান কি ?
[What is statistic?]

(গ) বিভিন্ন প্রকার দৈব নমুনায়নের নাম লিখ।
[Write down the types of random sampling.]

(ঘ) প্রাক্কলকের নিরপেক্ষতা বলতে কি বুঝ?
[What do you mean by unbiasedness of an estimator?]

(ঙ) সরল দৈব নমুনা চয়নের পদ্ধতিগুলো কি কি?
[What are the methods of drawing simple random sampling ?]

(চ) নমুনা জরিপ কেন করা হয়?
[What are reasons for sample survey ?]

(ছ) নমুনায়ন ত্রুটি বলতে কি বুঝ?
[What do you mean by sampling error ?]

(জ) কখন চক্রাকার ধারাবাহিক নমুনায়ন ব্যবহার করা হয়?
[When circular systematic sampling is used?]

(ঝ) স্তরিত দৈব নমুনায়নে নমুনা বন্টনের পদ্ধতিগুলো কি কি?
[What are the methods of allocating sample in stratified random sampling ?]

(ঞ) সসীম সমগ্রক শুদ্ধি বলতে কি বুঝ?
[What do you mean by finite population correaction?]

(ট) কোটা নমুনায়ন কি?
[What is quota sampling?]

(ঠ) উদ্দেশ্যমূলক নমুনায়ন কাকে বলে?
[What is meant by purposive sampling ?]

খ বিভাগ

(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ৪ x ৫ = ২০

২। নমুনায়ন কাঠামো কি? একটি নমুনায়ন কাঠামোর কি কি প্রয়োজনীয় গুণাবলি থাকা প্রয়োজন আলোচনা কর।

[What is sampling frame? Discuss the necessary characteristics which are required for a sampling frame.]

৩। সরল দৈব নমুনায়ন কি? দৈব সংখ্যা সারণী ব্যবহার করে কিভাবে সরল দৈব নমুনা নির্বাচন করা হয়, ব্যাখ্যা কর।

[What is simple random sampling? Explain how a simple random sample can be selected using random number table.]

৪। নমুনা জরিপ ও শুমারী জরিপ বলতে কি বুঝ? এদের মধ্যে পার্থক্য লিখ।

[What do you mean by sample survey and census? Distinguish between them.]

৫। একটি ভাল নমুনায়ন পরিকল্পনার আবশ্যকীয়তাসমূহ বর্ণনা কর।

[Describe the requirements of a good sampling design.]

৬। স্তরিত নমুনায়নের ক্ষেত্রে সমগ্রক গড়ের নিঝুঁকি নিরূপক নির্ণয় কর এবং এই প্রাক্কলকের ভেদাংক নির্ণয় কর।

[Find the unbiased estimator of population mean in case of stratified sampling and find the variance of this estimator.]

৭। ধারাবাহিক নমুনা নির্বাচন পদ্ধতির বর্ণনা দাও।

[Describe the method of selecting a systematic sampling.]

৮। দৈব নমুনায়ন ও ঐচ্ছিক নমুনায়ন বলতে কি বুঝ? এদের মধ্যে পার্থক্য কর।

[What do you mean by random sampling and purposive sampling? Distinguish between them.]

৯। গুচ্ছ নমুনায়ন কি? গুচ্ছ নমুনায়নের ক্ষেত্রে প্রচলিত সংকেতে প্রমাণ কর যে, $E(\bar{y}_n) = \bar{Y}$.

[What is cluster sampling? In case of cluster sampling, under usual notations prove that $E(\bar{y}_n) = \bar{Y}$.]

গ বিভাগ

(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ১০ x ৫ = ৫০

১০। (ক) উদাহরণসহ সংজ্ঞা দাও। নমুনা, নমুনায়ন কাঠামো। সম্ভাবনা নমুনায়ন ও অসম্ভাবনা নমুনায়নের মধ্যে পার্থক্য দেখাও

(খ) নমুনায়নের সঙ্গে সম্পৃক্ত বিভিন্ন ধরনের ত্রুটিসমূহ কী কী? এদের বর্ণনা দাও।

[(a) Define with example: Sample, Sampling frame. Distinguish between probability sampling and non- probability sampling.

(b) What are different types of errors associated with sampling? Describe them

১১। (ক) পুনঃস্থাপনসহ ও পুনঃস্থাপন ছাড়া নমুনায়ন বলতে কি বুঝ? প্রমাণ কর যে, $V(\bar{y})_{wor} \le V(\bar{y})_{wr}$.

(খ) পুনঃস্থাপন ছাড়া সরল দৈব নমুনায়নের ক্ষেত্রে সচরাচর সংকেতমালায় দেখাও যে,

$$E(\bar{y}) = \bar{Y} \text{ এবং } V(\bar{y}) = \frac{N-n}{N-1} \cdot \frac{\sigma^2}{n}$$

[(a) What do you mean by sampling with replacement and without replacement? Prove that $V(\bar{y})_{wor} \le V(\bar{y})_{wr}$.

(b) In SRSWOR under usual notations, show that $E(\bar{y}) = \bar{Y}$ and $V(\bar{y}) = \frac{N-n}{N-1} \cdot \frac{\sigma^2}{n}$]

১২। (ক) ধারাবাহিক নমুনায়ন কাকে বলে? কিভাবে ধারাবাহিক নমুনায়ন স্তরিত দৈব নমুনায়নের সাথে সম্পর্কিত?

(খ) সরলরৈখিক গতিধারাবিশিষ্ট সমগ্রক হতে নমুনায়নের ক্ষেত্রে দেখাও যে, $V(\bar{y}_{st}) \le V(\bar{y}_{sys}) \le V(\bar{y}_{rsn})$

[(a) What is systematic sampling? How is systematic sampling related to stratified random sampling?

(b) Show that in case of sampling from a population with linear trend, $V(\bar{y}_{st}) \le V(\bar{y}_{sys}) \le V(\bar{y}_{rsn})$.]

১৩। (ক) গুচ্ছ নমুনায়নের সংজ্ঞা দাও। অন্যান্য নমুনায়নের তুলনায় গুচ্ছ নমুনায়নের সুবিধাগুলো আলোচনা কর।

(খ) গুচ্ছ নমুনায়নে তথ্যবিশ্ব সমষ্টির একটি নিঝুঁকি নিরূপক নির্ণয় কর এবং সম-আকারের গুচ্ছায়নের ক্ষেত্রে এর ভেদাংকের সূত্রটি উদ্ভাবন কর।

[(a) Define cluster sampling. Discuss the advantages of cluster sampling in comparison to other sampling.

(b) Find an unbiased estimate of population total in case of cluster sampling with equal size of clusters and derive its variance.]

১৪। (ক) স্তরিত নমুনায়ন বলতে কি বুঝ? স্তরিত নমুনায়নের সুবিধা-অসুবিধা বর্ণনা কর।

(খ) বণ্টন কি? কখন সর্বোত্তম বণ্টন ও নেম্যান বণ্টন একই হয়? নেম্যান বণ্টনে দেখাও যে,

$V(\bar{y}_{st}) = \frac{1}{n}\left(\sum_{i=1}^{k} P_i S_i^{\,2}\right) - \frac{1}{N}\sum_{i=1}^{k} P_i S_i^{\,2}$; যেখানে $P_i = \frac{N_i}{N}$

[(a) What do you mean by stratified sampling? Describe the advantages and disadvantages of stratified sampling.

(b) What is allocation? When will optimum allocation and Neyman allocation be same? Show that for Neyman allocation $V(\bar{y}_{st}) = \frac{1}{n}\left(\sum_{i=1}^{k} P_i S_i^{\,2}\right) - \frac{1}{N}\sum_{i=1}^{k} P_i S_i^{\,2}$; where $P_i = \frac{N_i}{N}$.]

১৫। (ক) একটি ধারাবাহিক নমুনা নির্বাচনের পদ্ধতি আলোচনা কর। দেখাও যে, নির্দিষ্ট শর্ত সাপেক্ষে ধারাবাহিক নমুনায়ন সরল দৈব নমুনায়নের চেয়ে বেশি সঠিক।

(খ) ধারাবাহিক নমুনায়নের ক্ষেত্রে সমগ্রক গড়ের নিঝুঁকি প্রাক্কলক নির্ণয় কর এবং দেখাও যে,

$V(\bar{y}_{sys}) = \frac{(N-1)S^2}{Nn}\left[1+(n-1)\rho_w\right]$; যেখানে ρ_w হলো আন্তঃশ্রেণি সংশ্লেষাংক।

[(a) Discuss a method of selecting a systematic sampe. Show that under certain conditions, systematic sampling is more precise than simple random sampling.

(b) Find an unbiased estimate of population mean in case of systematic sampling and show that $V(\bar{y}_{sys}) = \frac{(N-1)S^2}{Nn}\left[1+(n-1)\rho_w\right]$; where ρ_w is intra-class correlation co-efficient.]

১৬। (ক) গুচ্ছ নমুনায়ন ব্যবহারের কারণসমূহ কি?

(খ) M সংখ্যক এককবিশিষ্ট N-গুচ্ছের সমগ্রক হতে n গুচ্ছের একটি সরল দৈব নমুনা পুন:স্থাপন ছাড়া নেয়ার ক্ষেত্রে দেখাও যে $\hat{Y}_c = NM\bar{y}$ হলো সমগ্রক সমষ্টি (Y) এর নিঝুঁকি প্রাক্কলক এবং এই প্রাক্কলকের ভেদাংক নিম্নরূপ:

$V(\hat{Y}_c) = \frac{(1-f)N^2MS^2}{n}\left[1+(M-1)\rho\right]$, যেখানে ρ হলো আন্তঃগুচ্ছ সংশ্লেষাংক।

[(a) What are the reasons for using cluster sampling?

(b) In SRSWOR of n cluster each containing M elements, from a population of N cluster, show that $\hat{Y}_c = NM\bar{y}$ is an unbiased estimate of population total (Y) and its variance is given by $V\left(\hat{Y}_c\right) = \frac{(1-f)N^2 MS^2}{n}\left[1+(M-1)\rho\right]$, where ρ is the intra-cluster correlation co-efficient.

১৭। (ক) উদাহরণসহ অনুপাত প্রাক্কলক ও নির্ভরণ প্রাক্কলকের মধ্যে পার্থক্য লিখ। কখণ নির্ভরণ প্রাক্কলক অনুপাত প্রাক্কলকে পরিণত হয়?

(খ) অনুপাত প্রাক্কলকের ঝোঁক নির্ণয় কর এবং দেখাও যে, $V\left(\hat{y}_R\right) = \frac{(1-f)}{n}\sum_{i=1}^{N}\frac{(y_i - Rx_i)^2}{N-1}$.

[(a) Write down the differences between ratio estimate and regression estimate with example. When does regression estimate tend to ratio estimate?

(b) Find the bias of ratio estimate and show that $V\left(\hat{y}_R\right) = \frac{(1-f)}{n}\sum_{i=1}^{N}\frac{(y_i - Rx_i)^2}{N-1}$

পরিসংখ্যান (তত্ত্বীয়)-২০১৬
[২০১৩-২০১৪ সালের সিলেবাস অনুযায়ী]
বিষয় কোড: 223603
(Sampling Technique)
সময়-৪ ঘণ্টা
পূর্ণমান-৮০
[**দ্রষ্টব্য:-**একই বিভাগের প্রশ্নের বিভিন্ন অংশের উত্তর ধারাবাহিকভাবে লিখতে হবে]

ক বিভাগ
(যে কোন **দশটি** প্রশ্নের উত্তর দাও)
মান- ১ x ১০ = ১০

১। (ক) পরামিতি কাকে বলে?

[What is called parameter?]

(খ) নমুনা কী?

[What is sampling?]

(গ) নিরূপক বলতে কি বুঝ?

[What do you mean by estimator?]

(ঘ) নমুনায়ন ভগ্নাংশ কী?

[What is sampling fraction?]

(ঙ) কখন স্তরিত দৈব নমুনায়ন ব্যবহার করা হয়?

[When is stratified random sampling used?]

(চ) সর্বোত্তম বন্টন কি?

[What is optimum allocation?]

(ছ) দৈব শুরু বলতে কি বুঝ?

[What do you mean by random start?]

(জ) গুচ্ছ বলতে কী বুঝ?

[What do you mean by cluster?]

(ঝ) নমুনা আকার বলতে কি বুঝ?

[What do you mean by sample size?]

(ঞ) প্রশ্নমালা কী?

What is questionnaire?

(ট) লা-জবাব বলতে কি বুঝায়?

What is meant by non-response?

(ঠ) কখন স্তরিত দৈব নমুনায়নের পরিবর্তে গুচ্ছ নমুনায়ন ব্যবহার করা হয়?

[When instead of stratified sampling, cluster sampling is used?]

খ বিভাগ

(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ৪ x ৫ = ২০

২। শুমারীর তুলনায় নমুনা জরিপের সুবিধাসমূহ আলোচনা কর।

[Discuss the advantages of sample survey over census.]

৩। নমুনাজ ত্রুটির কারণসমূহ বর্ণনা কর।

[Describe the sources of sampling error.]

৪। পুনঃস্থাপনসহ ও পুনঃস্থাপন ছাড়া নমুনায়নের মধ্যে পার্থক্য লিখ।

[Distinguish between sampling with replacement and without replacement.]

৫। কোন কোন শর্তে স্তরিত দৈব নমুনায়ন সরল দৈব নমুনায়ন অপেক্ষা পছন্দনীয়?

[Under what conditions stratified random sampling is preferred to simple random sampling?]

৬। পুনরাবৃত্ত ধারাবাহিক নমুনায়ন কী? এটি কখন প্রয়োজন হয়? পদ্ধতিটি একটি উদাহরণের সাহায্যে ব্যাখ্যা কর।

[What is repeated systematic sampling? When it is required? Elaborate the method with an example.]

৭। দেখাও যে, সমগ্রকের প্রাক্কলিত গড় $\bar{y}_{st}$ এর ভেদাংক ন্যূনতম হবে যদি $nh \propto Nh\dfrac{S_h}{\sqrt{ch}}$ হয়।

[Show that the variance of the estimator $\bar{y}_{st}$ of the population mean is minimum if $nh \propto Nh\dfrac{S_h}{\sqrt{ch}}$.]

৮। সরল দৈব নমুনায়নের তুলনায় গুচ্ছ নমুনায়নের আপেক্ষিক দক্ষতা নির্ণয় কর এবং মন্তব্য কর।

[Find the relative efficiency of cluster sampling compared to simple random sampling and comment.]

৯। অনুপাত প্রাক্কলকের বৈশিষ্ট্যগুলো উল্লেখ কর।

[Mention the properties of ratio estimator.]

গ বিভাগ

(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ১০ x ৫ = ৫০

১০। (ক) নমুনা জরিপের প্রধান ধাপসমূহ কি কি? এদের সংক্ষিপ্ত বর্ণনা দাও।

(খ) অনমুনায়ন বিচ্যুতির কারণসমূহ বর্ণনা দাও।

[(a) What are the main steps involved in sample survey? Discuss them briefly.

(b) Describe the reasons of non-sampling errors.]

১১। (ক) দেখাও যে, পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নে N-আকারের সমগ্রক হতে দুটি নির্দিষ্ট এককের n আকারের নমুনায় অন্তর্ভূক্তির সম্ভাবনা $\dfrac{n(n-1)}{N(N-1)}$.

(খ) দেখাও যে, সরল দৈব নমুনায়নে পুনস্থাপন না করে পদ্ধতিতে $E(S^2) = S^2$ কিন্তু $E(S^2) \neq \sigma^2$, যেখানে সংকেতগুলি প্রচলিত।

[(a) Show that, SRSWOR probability of inclusion of any two specified units in the sample of size n from a population of size N is $\frac{n(n-1)}{N(N-1)}$.

(b) Show that in case of random sampling without replacement $E(S^2) = S^2$ but $E(S^2) \neq \sigma^2$, where the notations are as usual.]

১২। (ক) বণ্টন বলতে কি বুঝ? স্তরিত দৈব নমুনায়ন বিভিন্ন প্রকার বণ্টন পদ্ধতি বর্ণনা কর।

(খ) যদি $\frac{1}{N_i}(i=1,2,...,k)$ নগন্য হয়, তবে দেখাও যে, $V(\bar{y}_{st})_{ran} \geq V(\bar{y}_{st})_{prop} V(\bar{y}_{st})_{opd}$.

[(a) What is meant by allocation? Describe the different methods of allocation of stratified random sampling.

(b) If $\frac{1}{N_i}(i=1,2,...,k)$ is negligible, then show that $V(\bar{y}_{st})_{ran} \geq V(\bar{y}_{st})_{prop} V(\bar{y}_{st})_{opd}$.]

১৩। (ক) অনুপাত প্রাক্কলক ও নির্ভরণ প্রাক্কলকের মধ্যকার পার্থক্যসমূহ লিখ।

(খ) N সংখ্যক গুচ্ছের একটি সমগ্রকের প্রতিটি গুচ্ছে M সংখ্যক উপাদান থাকলে পুনঃস্থাপন না করে সরল দৈব নমুনায়নের মাধ্যমে n সংখ্যক গুচ্ছ চয়ন করা হলে দেখাও যে, নমুনাগড় $\bar{\bar{y}}$ হচ্ছে সমগ্রক গড় $\bar{\bar{Y}}$ এর নিঝুঁকি নিরূপক এবং $\bar{\bar{y}}$ এর ভেদাংক, $V(\bar{\bar{y}}) = \frac{1-f}{Mn} S^2 [1+(M-1)\rho]$, এখানে, $\rho =$ আন্তশ্রেণি সংশ্লেষাংক, $f = \frac{n}{N}$

[(a) Distinguish between ratio estimator and regression estimator.

(b) In a simple random sampling without replacement of n clusters each containing M elements from a population of N clusters, show that the sample mean $\bar{\bar{y}}$ is an unbiased estimator of population mean $\bar{\bar{Y}}$ and its variance is $V(\bar{\bar{y}}) = \frac{1-f}{Mn} S^2 [1+(M-1)\rho]$, where $\rho =$ intraclass correlation co-efficient, $f = \frac{n}{N}$.]

১৪। (ক) দেখাও যে, ধারাবাহিক নমুনায়ন সরল দৈব নমুনায়নের চেয়ে ভাল প্রাক্কলক প্রদান করবে যদি $S^2_{wsy} > S^2$ হয় (প্রতীকগুলো চিরাচরিত)

(খ) সচরাচর প্রতীকে দেখাও যে, ধারাবাহিক নমুনা গড়, $\bar{y}_{sy}$ সমগ্রক গড় $\bar{y}$ এর নিঝুঁকি প্রাক্কলক এবং $V(\bar{y}_{sy}) = \frac{1}{N}\left[(N-1)S^2 - k(n-1)S^2_{wsy}\right]$ যেখানে, $S^2_{wsy} = \frac{1}{k(n-1)} \sum_{i=1}^{k} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{i\bullet})^2$

[(a) Show that systematic sampling gives more efficient estimator than simple random sampling if $S^2_{wsy} > S^2$ (Notations are as usual)

(b) Under usual notations, show that systematic sample mean $\bar{y}_{sy}$ is an unbiased estimate of population mean $\bar{y}$ and $V(\bar{y}_{sy}) = \frac{1}{N}\left[(N-1)S^2 - k(n-1)S^2_{wsy}\right]$, where,

$$S^2_{wsy} = \frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_{i\bullet}\right)^2 .]$$

১৫। (ক) গুচ্ছ নমুনায়ন ব্যবহৃত হয় এমন কিছু ক্ষেত্র উল্লেখ কর। গুচ্ছ নমুনায়ন ও স্তরিত দৈব নমুনায়নের মধ্যে পার্থক্য লিখ

(খ) গুচ্ছ নমুনায়নের সুবিধা ও অসুবিধাগুলো আলোচনা কর।

[(a) Mention some situations where cluster sampling is used. Distinguish between cluster sampling and stratified random sampling.

(b) Discuss the advantages and disadvantages of cluster sampling.]

১৬। (ক) অনুপাত প্রাক্কলক বলতে কি বুঝ? তথ্যবিশ্ব সমষ্টির অনুপাত প্রাক্কলকের ভেদাংক নির্ণয় কর।

(খ) নমুনা হতে উক্ত ভেদাংক কিভাবে প্রাক্কলন করা যায়?

[(a) what do you mean by ratio estimator? Find the variance of ratio estimator of the population total.

(b) How this variance can be estimated from sample?]

১৭। (ক) একটি আদর্শ প্রশ্নমালার আবশ্যকীয় গুণগুলি উল্লেখ কর।

(খ) বাংলাদেশের পরিপ্রেক্ষিতে প্রশ্নমালার মাধ্যমে তথ্য সংগ্রহের অসুবিধাগুলো সম্পর্কে আলোচনা কর।

[(a) Mention the necessary characteristics of an ideal questionnaire.

(b) Discuss about the disadvantages of data collection by a questionnaire in context of Bangladesh.]

পরিসংখ্যান (তত্ত্বীয়)-২০১৭

বিষয় কোড: 223603

(Sampling Technique)

সময়—৪ ঘন্টা

পূর্ণমান—৮০

[**দ্রষ্টব্য**:—একই বিভাগের বিভিন্ন প্রশ্নের উত্তর ধারাবাহিকভাবে লিখতে হবে।]

ক বিভাগ

(যে কোন **দশটি** প্রশ্নের উত্তর দাও)

মান—১×১০ =১০

১। (ক) নমুনায়ন বলতে কি বুঝ?

[What do you mean by sampling?]

(খ) নমুনা একক কী?

[What is sampling unit?]

(গ) সিডিউল কি?

[What is schedule?]

(ঘ) প্রাক্কলকের যথার্থতা বলতে কি বুঝ?

[What do you mean by precision of estimator?]

(ঙ) আদর্শ বিচ্যুতি কি?

[What is standard error?]

(চ) মিশ্র নমুনায়ন বলতে কি বুঝ?

[What do you mean by mixed sampling?]

(ছ) F.P.C কী?

[What is F.P.C?]

(জ) স্তর কী?

[What is stratum?]

(ঝ) স্তরিত নমুনায়নের ব্যয় অপেক্ষকটি লিখ।

[Write down the cost function of stratified sampling.]

(ঞ) Quota নমুনা কি?

[What is quota sample?]

(ট) ধারাবাহিক নমুনায়নে নমুনায়ন ব্যবধান কী?

[What is sampling interval in systematic sampling?]

(ঠ) কি শর্তে গুচ্ছ নমুনায়ন যথার্থ প্রাক্কলক প্রদান করে?

[Under what condition cluster sampling gives precious estimate?]

খ বিভাগ

(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)

মান—৪×৫ = ২০

২। শুমারী জরিপের সাথে নমুনা জরিপের পার্থক্য কর।

[Distinguish between sampling survey and census.]

৩। নমুনা জরিপের সুবিধাসমূহ আলোচনা কর।

[Discuss the advantages of sample survey.]

৪। পাইলট জরিপ বলতে কি বুঝ? এ ধরনের জরিপের আবশ্যকতা আলোচনা কর।

[What do you mean by pilot survey? Discuss the necessity of such type of survey.]

৫। প্রমাণ কর যে, $V(\overline{Y})_{WOR} \leq V(\overline{Y})_{WR}$.

[Prove that $V(\overline{Y})_{WOR} \leq V(\overline{Y})_{WR}$]

৬। দেখাও যে, পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নে একটি নির্দিষ্ট এককের যে কোনো চয়নে নমুনায় নির্বাচিত হবার সম্ভাবনা, এককটি প্রথম চয়নে নির্বাচিত হবার সম্ভাবনার সমান।

[Show that, in simple random sampling without replacement, the probability of selecting a specified unit in the sample in any draw is equal to the probability of selecting the unit in the first draw.]

৭। স্তরিত দৈব নমুনায়নে নমুনা গড়টি লিখ এবং দেখাও যে, নমুনা গড়টি তথ্যবিশ্ব গড়ের নির্ঝুঁকি নিরূপক নয়। কি শর্তে এটি নির্ঝুঁকি হতে পারে?

[Write down the sample mean in stratified random sampling and show that, sample mean is not an unbiased estimator for population mean. Under what condition it may be unbiased?]

৮। সরল দৈব নমুনায়ন কি? দৈব সংখ্যা সারণি ব্যবহার করে কিভাবে সরল দৈব নমুনা নির্বাচন করা হয়?

[What is simple random sampling? How do a simple random sample can be selected using random number tabele?]

৯। দেখাও যে, সরল দৈব নমুনায়নের তুলনায় ধারাবাহিক নমুনায়ন বেশী দক্ষ হবে যদি $\rho w < -\dfrac{1}{nk-1}$ হয়।

[Show that, systematic sampling will be more efficieent than simple random sampling if $\rho w < -\dfrac{1}{nk-1}$.]

গ বিভাগ

(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)

মান—১০×৫ = ৫০

১০। (ক) পার্থক্য কর:—

(i) সম্ভাবনা নমুনায়ন ও অসম্ভাবনা নমুনায়ন;

(ii) প্রশ্নমালা ও অনুসূচী।

(খ) নমুনায়ন কাঠামো কি? একটি নমুনায়ন কাঠামোর আবশ্যকীয় গুণাবলীসমূহ বর্ণনা কর।

[(a) Distinguish between:—

(i) Probability sampling and non-Probability sampling;

(ii) Questionnaire and schedule.

(b) What is sampling frame? Discuss the requisite characteristics of a sampling frame.]

১১। পুনঃস্থাপনবিহীন সরল দৈব নমুনায়নের ক্ষেত্রে সমগ্রকের গড় ও সমগ্রকের সমষ্টির নিঝুঁকি নিরূপক নির্ণয় কর। নিরূপকগুলির ভেদাংক ও পরিমিত বিচ্যুতি নির্ণয় কর।

[Determine the unbiased estimators of the population mean and population total in case of simple random sampling without replacement. Find their vairances and standard errors.]

১২। (ক) সরল দৈব নমুনায়নে তথ্যবিশ্ব অনুপাতের একটি নিঝুঁকি নিরূপক বের কর এবং নিরূপকটির ভেদাংক নির্ণয় কর।

(খ) ৫% যথার্থতা মাত্রায় সরল দৈব নমুনায়নে নমুনার আকার নির্ধারণ কর যেখানে পরামিতি $(\overline{Y})$ এর সাথে নিরূপকের বিচ্যুতির পরিমাণ d ।

[(a) In simple random sampling find an unbiased estimate for population proportion and also find the variance of the estimator.

(b) Determine the sample size in simple random sampling at level 5% of significance where the deviation between parameter $(\overline{Y})$ and estimator is d .]

১৩। (ক) স্তরীকরণের নীতি কি কি? স্তরিত দৈব নমুনায়নে তথ্যবিশ্ব সমষ্টির নিঝুঁকি প্রাক্কলক বের কর এবং প্রাক্কলকটির ভেদাংক নির্ণয় কর।

(খ) সমানুপাতিক বন্টন ও নেম্যানের বন্টনের ক্ষেত্রে স্তরিত নমুনাগড়ের ভেদাংক নির্ণয় কর।

[(a) What are the principles of stratification? In stratified random sampling find an unbiased estimator for population total and find the variance of the estimator.

(b) Find the variance of stratified sample mean for proportional allocation and Neyman allocation.]

১৪। (ক) সরল দৈব নমুনায়নের তুলনায় স্তরিত দৈব নমুনায়নের সুবিধা ও অসুবিধাগুলো আলোচনা কর।

(খ) দেখাও যে, স্থির নমুনার আকার n -এর জন্য স্তরিত নমুনায়নের ভেদাংক সর্বনিম্ন হবে যদি $n_i \propto N_i \; S_i$ হয়।

[(a) Discuss the advantages and disadvantages of stratified random sampling over simple random sampling.

(b) Show that, for fixed sample size n, the variance of stratified sample mean wil lbe minimum if $n_i \propto N_i \; S_i$]

১৫। (ক) ধারাবাহিক নমুনায়ন বলতে কি বুঝ? ধারাবাহিক নমুনা নির্বাচনের পদ্ধতি আলোচনা কর।

(খ) ধারাবাহিক নমুনায়নে দেখাও যে, ধারাবাহিক নমুনা গড়, $\bar{Y}sys$ হচ্ছে গড়ের, তথ্যবিশ্ব গড়ের নিঝুঁকি প্রাক্কলক এবং $V\left(\overline{Y}_{sys}\right)=\frac{S^2}{n}\cdot\frac{N-1}{N}\cdot\left[1+(n-1)\rho_w\right]$, যেখানে ρ_w হলো আন্তঃশ্রেণী সংশ্লেষাংক। এখান থেকে ρ_w এর সর্বনিম্ন মান বের কর।

[(a) What do you mean by systematic sampling? Discuss the method of drawing systematic sample.

(b) In systematic sampling show that, systematic sample mean $\bar{Y}sys$ is an unbiased estimate for population mean and $V\left(\bar{Y}_{sys}\right)=\frac{S^2}{n}\cdot\frac{N-1}{N}\cdot\left[1+(n-1)\rho_w\right]$, where ρ_w is intra-calss corelation co-efficient. Find the minimum value of ρ_w from it.]

১৬। (ক) ধারাবাহিক নমুনায়নের সুবিধা ও অসুবিধাসমূহ আলোচনা কর।

(খ) $Y_i = a + b_i$ আকারের সরলরৈখিক গতিধারাবিশিষ্ট তথ্যবিশ্ব হতে নমুনায়নের ক্ষেত্রে দেখাও যে,

$$V\left(\bar{Y}_{st}\right)\leq V\left(\bar{Y}_{sys}\right)\leq V\left(\bar{Y}_{ran}\right)$$

[(a) Discuss the advantages and disadvantages of systematic sampling.

(b) In case of sampling from a population with a linear trend of the form $Y_i = a + b_i$ show that $V\left(\bar{Y}_{st}\right)\leq V\left(\bar{Y}_{sys}\right)\leq V\left(\bar{Y}_{ran}\right)$]

১৭। (ক) কখন সরল দৈব নমুনায়নের চেয়ে গুচ্ছ নমুনায়ন পছন্দনীয়?

(খ) M সংখ্যক উপাদানবিশিষ্ট N গুচ্ছের সমগ্রক হতে পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নের মাধ্যমে n সংখ্যক গুচ্ছ নির্বচন করা হলে দেখাও যে, $Y = N\bar{Y}$ হলো তথ্যবিশ্ব সমষ্টির নির্ঝুঁকি প্রাক্কলক যেখানে $\bar{Y}$ হলো গুচ্ছ প্রতি নমুনাগড়। উক্ত প্রাক্কলকের ভেদাংক নির্ণয় কর।

[(a) When cluster sampling is preferred to simple random sampling?

(b) If n cluster are slelected by simple random sampling without replacement from a population of N clusters each containing M elements then show that, $Y = N\bar{Y}$ is an unbiased estimaor for population total, where $\bar{Y}$ is the sample mean per cluster. Find the variance of the estimator.]

পরিসংখ্যান (তত্ত্বীয়)-২০১৮

বিষয় কোড ঃ 223603

Sampling Technique

সময় - ৪ ঘন্টা

পূর্ণমান - ৮০

[**দ্রষ্টব্য ঃ –** একই বিভাগের প্রশ্নের উত্তর ধারাবাহিকভাবে লিখতে হবে।]

ক বিভাগ

(যে কোন **দশটি** প্রশ্নের উত্তর দাও)

মান – ১ × ১০ = ১০

১। (ক) নমুনা বলতে কি বোঝ?

[What do you mean by sample?]

(খ) পাইলট জরিপ কী?

[What is Pilot survey?]

(গ) প্রশ্নমালা কী?

[What is questionnaire?]

(ঘ) উত্তরহীনতা বলতে কি বুঝ?

[What do you mean by non-response?]

(ঙ) নমুনায়ন ভগ্নাংশ কী?

[What is sampling fraction?]

(চ) দৈব সংখ্যা সারণী কী?

[What is random number table?]

(ছ) নিরুপক কী?

[What is estimator?]

(জ) স্তরিতকরণ বলতে কি বুঝ?

[What do you mean by stratification?]

(ঝ) নমুনায়ন কাঠামো কী?

[What sampling frame?]

(ঞ) সমবন্টন কী?

[What is equal allocation?]

(ট) নমুনায়ন ত্রুটি কী?

[What is samphng error?]

(ঠ) গুচ্ছ কী?

[What is cluster?]

খ বিভাগ

(যে কোনো **পাঁচটি** প্রশ্নের উত্তর দাও)

মান – ৪ × ৫ = ২০

২। শুমারি জরিপ কি? এর উদ্দেশ্য বর্ণনা কর।

[What is census? Describe its objectives.]

৩। সম্ভাবনা নমুনায়ন ও অসম্ভাবনা নমুনায়নের মধ্যে পার্থক্য কর।

[Distinguish between probability sampling and non- probability sampling.]

৪। সরল দৈব নমুনায়ন বলতে কী বুঝ? দেখাও যে, পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নে N-আকারের সমগ্রক হতে দু'টি নির্দিষ্ট এককের n-আকারের নমুনায় অন্তর্ভুক্তির সম্ভাবনা $\frac{n(n-1)}{N(N-1)}$।

[What do you mean by simple random sampling? Show that for SRSWOR, probability of inclusion of any two specified units in the sample of size n from a population of size N is $\frac{n(n-1)}{N(N-1)}$.]

৫। ধারাবাহিক নমুনায়নে গড়ের ভেদাংক এবং সরল দৈব নমুনায়ে নমুনা গড়ের ভেদাংকের মধ্যে সম্পর্ক স্থাপন কর।

[Establish the relation between the variance of sample mean in systematic sampling and the variance of sample mean under SRS.]

৬। স্তরিত নমুনায়ন বলতে কী বুঝ? স্তরিত দৈব নমুনায়নের সুবিধা ও অসুবিধা বর্ণনা কর।

[What do you mean by stratified sampling? Describe the advantages and disadvantages of stratified sampling.]

৭। স্তরিত দৈব নমুনায়ন ও গুচ্ছ নমুনায়নের মধ্যে পার্থক্য কর।

[Distinguish between stratified random sampling and cluster sampling.]

৮। নমুনা জরিপে ভ্রম-উৎসসমূহ কি কি? এই ভ্রম কিভাবে কমানো যায়?

[What are sources of error in sample survey? How these error can be reduced?]

৯। সরল দৈব নমুনায়নের তুলনায় গুচ্ছ নমুনায়নের আপেক্ষিক দক্ষতা নির্ণয় কর ও মন্তব্য কর।

[Find the relative efficiency of cluster sampling compared to simple random sampling and comment.]

গ বিভাগ

(যে কোনো **পাঁচটি** প্রশ্নের উত্তর দাও)

মান – ১০ × ৫ = ৫০

১০। (ক) নমুনা জরিপের প্রধান ধাপসমূহ বর্ণনা কর।

(খ) পার্থক্য দেখাও :

(i) খোলা ও বন্ধ প্রশ্নমালা

(ii) গঠিত ও অগঠিত প্রশ্নমালা।

[(a) Describe the main steps involved in sample survey.

(b) Distinguish between:

(i) Closed and open questionnaire.

(ii) Structured and non-structured questionnaire.]

১১। (ক) সরল দৈব নমুনায়ন বলতে কি বুঝ? N আকারের তথ্য বিশ্ব হতে n আকারের সরল দৈব নমুনা চয়ন করা হলে দেখাও যে, একটি নির্দিষ্ট এককের নমুনায় অন্তর্ভুক্ত হওয়ার সম্ভাবনা $\frac{n}{N}$.

(খ) সরল দৈব নমুনায়নে নমুনার অনুপাত P হলে দেখাও যে, $V(P)=\frac{N-n}{N-1}.\frac{PQ}{n}$.

[(a) What do you mean by simple random sampling? If a simple random sample of size n is drawn from a population of size N, then show that the probability of including a specified unit in the sample is $\frac{n}{N}$.

(b) In simple random sampling if P be the sample proportion, then show that $V(P)=\frac{N-n}{N-1}.\frac{PQ}{n}$.]

১২। (ক) দেখাও যে, সরল দৈব নমুনায়নে নমুনা ভেদাংক হলো সমগ্রক ভেদাংকের নিরপেক্ষ নিরূপক।

(খ) n আকারবিশিষ্ট একটি সরল দৈব নমুনার জন্য দেখাও যে, পুনঃস্থাপন ব্যতিরেকে নমুনা গড়ের পরিমিতি বিচ্যুতি পুনঃস্থাপন সহকারে নমুনা গড়ের পরিমিত বিচ্যুতি অপেক্ষা ছোট।

[(a) Show that in SRS sample variance is an unbiased estimate of population variance.

(b) For a simple random sample of size n, show that the standard error of sample mean in sampling without replacement is smaller than that in sampling with replacement.]

১৩। (ক) বন্টন বলতে কী বুঝ? স্তরিত দৈব নমুনায়নের বিভিন্ন বন্টন পদ্ধিত বর্ণনা কর।

(খ) সর্বোত্তম বন্টনের ক্ষেত্রে নমুনা আকারের রাশিমালা উদ্ভাবন কর।

[(a) What is meant by allocation? Describe the different methods of allocation of stratified random sampling.

(b) Derive the expression for sample size in optimum allocation.]

১৪। (ক) স্তরিত নমুনায়নের ক্ষেত্রে সমগ্রক গড়ের নিৰ্ঝুঁকি নিরূপক নির্ণয় কর এবং এই প্রাক্কলকের ভেদাংক নির্ণয় কর।

(খ) যদি $\frac{1}{N_i}$ (i = 1, 2 ….,k) নগণ্য হয়, তবে দেখাও যে, $v(y_{st})_{ran} \geq v(y_{st})_{prop} \geq v(\bar{y}_{st})_{opt}$।

[(a) Find the unbiased estimator of population mean in case of stratified sampling and find the variance of this estimator.

(b) If $\frac{1}{N_i}$ (i = 1, 2 ….,k) is negligible, then show that $v(y_{st})_{ran} \geq v(y_{st})_{prop} \geq v(\bar{y}_{st})_{opt}$.]

১৫। (ক) দেখাও যে, ধারাবাহিক নমুনায়ন সরল দৈব নমুনায়নের চেয়ে ভাল প্রাক্কলক প্রদান করবে যদি $s^2{}_{wsy} > s^2$ হয় (প্রতীকগুলো চিরাচরিত)।

(খ) সচরাচর প্রতীকে দেখাও যে, ধারাবাহিক নমুনা গড় $\bar{y}_{sy}$, সমগ্রক গড় $\bar{y}$ এর নিৰ্ঝুঁকি প্রাক্কলক এবং $V(\bar{y}_{sy})=\frac{1}{N}\left[(N-1)s^2-k(n-1)s^2{}_{wsy}\right]$ যেখানে $s^2{}_{wsy}=\frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_i\right)^2$।

[(a) Show that systematic sampling gives more efficient estimator than simple random sampling if $s^2{}_{wsy} > s^2$ (Notations are as usual).

(b) Under usual notations, show that systematic sample mean $\bar{y}_{sy}$ is an unbiased estimate of population mean, $\bar{y}$ and $V(\bar{y}_{sy})=\frac{1}{N}\left[(N-1)s^2-k(n-1)s^2{}_{wsy}\right]$ where $s^2{}_{wsy}=\frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_i\right)^2$]

১৬। (ক) স্তরিত নমুনায়নের সাথে গুচ্ছ নমুনায়নের তুলনা কর।

(খ) গুচ্ছ নমুনায়নের সুবিধা ও অসুবিধাগুলো আলোচনা কর।

[(a) Compare cluster sampling with stratified sampling.

(b) Discuss the advantages and disadvantages of cluster sampling.]

১৭। (ক) একটি আদর্শ প্রশ্নমালার আবশ্যকীয় বৈশিষ্ট্যগুলো বিবৃতি কর।

(খ) বাংলাদেশের পরিপ্রেক্ষিতে প্রশ্নমালার মাধ্যমে তথ্য সংগ্রহের অসুবিধাগুলো সম্পর্কে আলোচনা কর।

[(a) State the necessary characteristics of an ideal questionnaire.

(b) Discuss about the disadvantages of data collection by a questionnaire in context of Bangladesh.]

পরিসংখ্যান (তত্ত্বীয়)- ২০১৯

বিষয় কোড ঃ 223603

Sampling Technique

সময় - ৪ ঘন্টা

পূর্ণমান - ৮০

[**দ্রষ্টব্য ঃ –** প্রতিটি বিভাগের প্রশ্নের উত্তর ধারাবাহিকভাবে লিখতে হবে।]

ক-বিভাগ

(যে কোন **দশটি** প্রশ্নের উত্তর দাও)

মান- ১ x ১০ = ১০

১। (ক) সমগ্রকের সংজ্ঞা দাও।

[Define population.]

(খ) নমুনা একক বলতে কী বুঝ?

[What do you mean by sampling unit?]

(গ) প্রাক্কলিত মান কী?

[What is an estimate?]

(ঘ) প্রাক্কলকের নিরপেক্ষতা বলতে কী বুঝ?

[What do you mean by unbiasedness of an estimator?]

(ঙ) স্তর বলতে কী বুঝ?

[What do you mean by stratum?]

(চ) উদ্দেশ্যমূলক নমুনায়ন বলতে কী বুঝ?

[What do you by purposive sampling?]

(ছ) কখন নেম্যান বন্টন ও সমানুপাতিক বন্টন একই হয়?

[When Neyman’s allocation is same to the proportional allocation?]

(জ) কখন চক্রাকার ধারাবাহিক নমুনায়ন ব্যবহার করা হয়?

[When circular systematic sampling is used?]

(ঝ) নিরূপকের সঠিকতা বলতে কী বুঝ?

[What do you mean by accuracy of estimator?]

(ঞ) অনমুনাজ ত্রুটি কী?

[What do you non-sampling error?]

(ট) গুচ্ছের আকার বলতে কী বুঝ?

[What do you mean by size of a cluster?]

(ঠ) কখন গুচ্ছ নমুনায়ন ব্যবহার করা হয়?

[When cluster sampling is used?]

খ বিভাগ

(যে কোন **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ৪ x ৫ = ২০

২। নমুনা জরিপের মূলনীতি বর্ণনা কর।

[Discuss the basic principles of sample survey.]

৩। উত্তরদাতার নিরবতা বলতে কী বুঝ? নমুনা জরিপে এর অসুবিধাগুলো আলোচনা কর।

[What do you mean by non-response? Discuss its disadvantages in sample survey.]

৪। সরল দৈব নমুনায়ন কী? দৈব সংখ্যা সারণী ব্যবহার করে কিভাবে সরল দৈব নমুনা চয়ন করা হয়? ব্যাখ্যা কর।

[What is simple random sampling? Explain how a simple random sample can be selected using random number table.]

৫। সরল দৈব নমুনায়নের ক্ষেত্রে প্রমাণ কর যে, নমুনা গড় সমগ্রক গড়ের নিরপেক্ষ নিরূপক।

[In SRS, prove that, sample mean in an unbiased estimate of population mean.]

৬। স্তরিত দৈব নমুনায়নে সমগ্রক সমষ্টির নিরপেক্ষ নিরূপক ও তার ভেদাংক নির্ণয় কর।

[In stratified random sampling find the unbiased estimate of population total and its variance.]

৭। দেখাও যে, পুনঃস্থাপন ছাড়া সরল দৈব নমুনায়নের মাধ্যমে N গুচ্ছবিশিষ্ট সমগ্রক হতে n গুচ্ছ চয়ন করা হলে $\bar{y}_c$ এর ভেদাংক হবে $v(\bar{Y}_c) = \frac{1-f}{nM}S^2[1+(M-1)\rho]$ যেখানে, ρ অন্তঃশ্রেণি সংশ্লেষাংক এবং M গুচ্ছের আকার।

[If a sample of n clusters is drawn from a population of N clusters, then show that the variance of $\bar{y}_c$ is $v(\bar{Y}_c) = \frac{1-f}{nM}S^2[1+(M-1)\rho]$, where ρ is intra-class correlation co-efficient and M is size of cluster.]

৮। ধারাবাহিক নমুনা নির্বাচন পদ্ধতি বর্ণনা কর।

[Describe the method of selecting a systematic sample.]

৯। তোমার কলেজের শিক্ষার্থীদের মাসিক গড় ব্যয় নিরূপণে গুচ্ছ পদ্ধতি বর্ণনা কর।

[Discuss cluster sampling technique to estimate the monthly average expenditure of the students of your college.]

গ বিভাগ

(যে কোনো **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ১০ X ৫ = ৫০

১০। (ক) শুমারি ও নমুনা জরিপের মধ্যে পার্থক্য কর।

(খ) নমুনাজ ত্রুটি ও অনমুনাজ ত্রুটির কারণগুলো লিখ। এসব ত্রুটি কমানোর উপায় আলোচনা কর।

[(a) Distinguish between census and sample survey.

(b) Write down the causes of sampling error and non-sampling error. Discuss the procedure to reduce such type of error.]

১১। (ক) সরল দৈব নমুনায়নে নমুনা গড়ের আদর্শ বিচ্যুতি নির্ণয় কর।

(খ) দেখাও যে, পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নে একটি নির্দিষ্ট এককের যে কোনো চয়নে নমুনায় নির্বাচিত হবার সম্ভাবনা, এককমটি প্রথম চয়নে নির্বাচিত হবার সম্ভাবনার সমান।

[(a) In simple random sampling, find the standard error of sample mean.

(b) Show that, in simple random sampling without replacement, the probability of selecting a specified unit in the sample in any draw is equal to the probability of selecting the unit in the first draw.]

১২। (ক) সরল দৈব নমুনায়নের তুলনায় স্তরিত দৈব নমুনায়নের সুবিধা ও অসুবিধাগুলো আলোচনা কর।

(খ) স্তরিত দৈব নমুনায়নে তথ্যবিশ্ব সমষ্টির নিঝুঁকি প্রাক্কলক বের কর এবং প্রাক্কলকটির ভেদাংক বের কর।

[(a) Discuss the advantages and disadvantages of stratified random sampling over simple random sampling.

(b) In stratified random sampling find an unbiased estimator for population total and find the variance of the estimator.]

১৩। (ক) সমানুপতিক বন্টন ও নেম্যানের বন্টনের ক্ষেত্রে স্তরিত নমুনা গড় $(\bar{Y}_{st})$ এর আদর্শ বিচ্যুতি নির্ণয় কর।

(খ) দেখাও যে, ব্যয় অপেক্ষক $C = a + \sum_{i=1}^{k} c_i . n_i$ বিশিষ্ট স্তরিত দৈব নমুনায়নে স্তরিত নমুনা গড় $(\bar{Y}_{st})$ এর ভেদাংক সর্বনিম্ন হবে যদি $n_i \; \alpha \; \frac{N_i S_i}{\sqrt{C_i}}$ হয় (প্রতীকগুলো চিরাচরিত)।

[(a) Find the standard error of stratified sample mean for proportional allocation and Neyman allocation.

(b) Show that in stratified random sampling with cost function $C = a + \sum_{i=1}^{k} c_i . n_i .$, the variance of stratified sample mean $(\bar{Y}_{st})$ will be minimum if $n_i \; \alpha \; \frac{N_i S_i}{\sqrt{C_i}}$ (notations are as usal)]

১৪। (ক) ধারাবাহিক নমুনায়ন বলতে কী বুঝ? ধারাবাহিক নমুনায়নে তথ্যবিশ্ব সমষ্টির নিঝুঁকি প্রাক্কলক বের কর এবং প্রাক্কলকটির আদর্শ বিচ্যুতি নির্ণয় কর।

(খ) দেখাও যে, সরল দৈব নমুনায়নের তুলনায় ধারাবাহিক নমুনায়ন বেশী দক্ষ হবে যদি $\rho_\omega < -\frac{1}{nk-1}$ হয়।

[(a) What do you mean by systematic sampling? In systematic sampling find an unbiased estimator for population total and find standard error of the estimator.

(b) Show that, systematic sampling will be more efficient than simple random sampling if $\rho_\omega < -\frac{1}{nk-1}$.]

১৫। (ক) ধারাবাহিক নমুনায়নের সুবিধা ও অসুবিধাসমূহ আলোচনা কর।

(খ) সরলরৈখিক গতিধারাবিশিষ্ট তথ্যবিশ্ব হতে নমুনায়নের ক্ষেত্রে দেখাও যে, $V(\bar{Y}_{st}) \le v(\bar{Y}_{sys}) \le V(\bar{Y}_{ran})$.

[(a) Discuss the advantages and disadvantages of systematic sampling.

(b) In case of sampling from a population with linear trend, show that $V(\bar{Y}_{st}) \le v(\bar{Y}_{sys}) \le V(\bar{Y}_{ran})$.]

১৬। (ক) উদাহরণ গুচ্ছ নমুনায়ন ব্যাখ্যা কর। কখন সরল দৈব নমুনায়নের চেয়ে গুচ্ছ নমুনায়ন পছন্দনীয়?

(খ) গুচ্ছ কাকে বলে? M সংখ্যক উপাদানবিশিষ্ট N গুচ্ছের সমগ্রক হতে সরল দৈব নমুনায়নের মাধ্যমে n সংখ্যক গুচ্ছ নির্বাচন করা হলে দেখাও যে, নমুনা গুচ্ছ গড়ের গড়, সমগ্রক গড়ের নিঝুঁকি প্রাক্কলক।

[(a) Explain cluster sampling with example. When cluster sampling is preferred to simple random sampling?

(b) What is called Cluster? If n clusters are selected by simple random sampling from a population of N clusters each with M elements then show that mean of sample cluster mean is an unbiased estimator for population mean.]

১৭। (ক) সিডিউল কী? প্রশ্নমালা ও সিডিউলের মধ্যে পার্থক্য লিখ।

(খ) একটি প্রশ্নমালা তৈরির ধাপসমূহ বর্ণনা কর।

[(a) What is Schedule? Distinguish between questionnaire and schedule.

(b) Describe the steps in constructing a questionnaire.]

পরিসংখ্যান (তত্ত্বীয়)- ২০২০

বিষয় কোড : 223603

(Sample Technique)

সময়- ৪ ঘন্টা

পূর্ণমান- ৮০

[**দ্রষ্টব্য:** প্রতিটি বিভাগের বিভিন্ন প্রশ্নের উত্তর ধারাবাহিকভাবে লিখতে হবে ।]

ক বিভাগ

(যে কোনো **দশটি প্রশ্নের** উত্তর দাও)

মান- ১ × ১০ = ১০

১। (ক) নমুনা বলতে কী বুঝ?

[What do you mean by sample?]

(খ) শুমারি কী?

[What is census?]

(গ) প্রাক-জরিপ কী?

[What is pilot survey?]

(ঘ) অসীম সমগ্রক কী?

[What is infinite population?]

(ঙ) নমুনা কাঠামো কী?

[What is sampling frame?]

(চ) নমুনা ত্রুটি কী?

[What is sampling error?]

(ছ) প্রাক্কলক বলতে কী বুঝ?

[What do you mean by estimator?]

(জ) সসীম সমগ্রক শুদ্ধি (fpc) বলতে কী বুঝ?

[What do you mean by finite population correction (fpc)?]

(ঝ) নমুনায়ন কী?

[What is Quota sampling?]

(ঞ) ধারাবাহিক নমুনায়নের অসুবিধাগুলো কী কী?

[What are the disadvantages of systematic sampling?]

(ট) সমবণ্টন বলতে কী বুঝ?

[What do you mean by equal allocation?]

(ঠ) গুচ্ছ বলতে কী বুঝ?

[What do you mean by cluster?]

খ- বিভাগ

(যে কোনো **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ৪ × ৫ = ২০

২। নমুনায়ন কী? দৈব নমুনায়ন এবং বিচারভিত্তিক নমুনায়নের মধ্যে পার্থক্য কর ।

[What is sampling? Distinguish between random sampling and judgement sampling.]

৩। পার্থক্য কর-

(ক) বদ্ধ ও খোলা প্রশ্নমালা

(খ) গঠিত ও অগঠিত প্রশ্নমালা।

[Distinguish between-

(a) Closed and open questionnaire

(b) Structured and non-structured questionnaire.]

৪। সরল দৈব নমুনায়নের ক্ষেত্রে দেখাও যে, $V(\bar{y})_{wor} \le V(\bar{y})_{wr}$.

[In case of simple random sampling show that, $V(\bar{y})_{wor} \le V(\bar{y})_{wr}$.]

৫। সরল দৈব নমুনায়নের n আকারের নমুনার জন্য নমুনা p হলে দেখাও যে, $V(p) = \frac{N-n}{n(N-1)}PQ$.

[In simple random sampling if p be the sample proportion for sample size n, then show that, $V(p) = \frac{N-n}{n(N-1)}PQ$]

৬। স্তরিত দৈব নমুনায়ন কী? এর সুবিধা ও অসুবিধা আলোচনা কর।

[What is stratified random sampling? Discuss its advantages and disadvantages.]

৭। স্তরিত দৈব নমুনায়নের নমুনা গড়টি লেখ এবং দেখাও যে, নমুনা গড়টি সমগ্রক গড়ের নির্ঝুঁকি নিরূপক নয়।

[Write down the sample mean of stratified random sampling and show that sample mean is not an unbiased estimator of population mean.]

৮। দেখাও যে, সরল দৈব নমুনায়নের চেয়ে ধারাবাহিক নমুনায়ন অধিকতর দক্ষ প্রাক্কলক প্রদান করবে যদি $S_{wsy}{}^2 > S^2$ হয় (প্রতীকগুলো চিরাচরিত)।

[Show that systematic sampling provides more efficient estimator than simple random sampling if $S_{wsy}{}^2 > S^2$ (notations are as usual).]

৯। গুচ্ছ নমুনায়নের সুবিধা ও অসুবিধাসমূহ আলোচনা কর।

[Discuss the advantages and disadvantages of cluster sampling.]

গ- বিভাগ

(যে কোনো **পাঁচটি** প্রশ্নের উত্তর দাও)

মান − ১০ × ৫ = ৫০

১০। (ক) নমুনা জরিপ কী? এর সুবিধা ও অসুবিধাগুলো আলোচনা কর।

(খ) একটি নমুনা জরিপের ধাপসমূহ বর্ণনা কর।

[(a) What is sample survey? Discuss its advantages and disadvantages.

(b) Describe the steps involved in a sample survey.]

১১। (ক) অনুসূচি বলতে কী বুঝ? প্রশ্নমালা ও অনুসূচির মধ্যে পার্থক্য কর।

(খ) একটি উত্তম প্রশ্নমালার মানদণ্ডসমূহ বিবৃত কর।

[(a) What do you mean by schedule? Distinguish between questionnaire and schedule.

(b) State the criterion of a good questionnaire.

১২। (ক) পুনঃস্থাপন সহকারে দৈব নমুনায়নের ক্ষেত্রে সমগ্রক সমষ্টির প্রাক্কলকের ভেদাংক নির্ণয় কর।

(খ) পুনঃস্থাপন না করে সরল দৈব নমুনায়নের ক্ষেত্রে দেখাও যে, $E(s^2) = E(S^2)$, কিন্তু $E(s^2) \neq \sigma^2$, যেখানে সংকেতগুলো প্রচলিত।

[(a) In case of simple random sampling with replacement determine the variance of the estimate of the population total.

(b) In case of simple random sampling without replacement show that, $E(s^2) = E(S^2)$, but $E(s^2) \neq \sigma^2$, where the notations are as usual.]

১৩। (ক) স্তরিত দৈব নমুনায়নের ক্ষেত্রে সমগ্রক গড়ের প্রাক্কলকের ভেদাংক নির্ণয় কর।

(খ) দেখাও যে, স্থির নমুনা আকারের জন্য স্তরিত দৈব নমুনায়নের $V(\bar{y}_{st})$ ক্ষুদ্রতম হবে যদি $n_h \alpha N_h S_h$ হয়, যেখানে প্রতীকসমূহ চিরাচরিত।

[(a) In case of stratified random sampling find the variance of the estimate of population total.

(b) Show that for fixed sample size $V(\bar{y}_{st})$ in stratified random sampling will be least if $n_h \alpha N_h S_h$, where the symbols are usual.]

১৪। (ক) বণ্টন বলতে কী বুঝ? স্তরিত দৈব নমুনায়নের বিভিন্ন বণ্টন পদ্ধতিসমূহ আলোচনা কর।

(খ) প্রচলিত সংকেতের অধীনে প্রমাণ কর যে,

$$V(\bar{y}_{st})_{opt} \leq V(\bar{y}_{st})_{prop} \leq V(\bar{y}_{st})_{ran}$$

[(a) What do you mean by allocation? Discuss different methods of allocation of stratified random sampling.

(b) Under usual notations prove that,

$$V(\bar{y}_{st})_{opt} \leq V(\bar{y}_{st})_{prop} \leq V(\bar{y}_{st})_{ran}$$

১৫। (ক) সচরাচর প্রতীকে দেখাও যে, ধারাবাহিক নমুনা গড় $\bar{y}_{sy}$, সমগ্রক গড় $\bar{y}$-এর নির্ঝুঁকি নিরূপক এবং

$$V(\bar{y}_{sy}) = \frac{1}{N}[(N-1)s^2 - k(n-1){s_{wsy}}^2 \text{ , যেখানে } {s_{wsy}}^2 = \frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y}_i)^2 .$$

(খ) পুনরাবৃত্ত ধারাবাহিক নমুনায়ন বলতে কী বুঝ? এ নমুনায়ন পদ্ধতিটি একটি উদাহরণের সাহায্যে ব্যাখ্যা কর।

[(a) Under usual notations show that systematic sample mean $\bar{y}_{sy}$ is an unbiased estimate of population mean $\bar{y}$ and $V(\bar{y}_{sy}) = \frac{1}{N}[(N-1)s^2 - k(n-1)s_{wsy}{}^2$; where, $s_{wsy}{}^2 = \frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y}_i)^2$.

(b) What do you mean by repeated systematic sampling? Explain this sampling method with an example.]

১৬। (ক) ধারাবাহিক নমুনায়নে নমুনা গড়ের ভেদাংক এবং সরল দৈব নমুনায়নে নমুনা গড়ের ভেদাংকের মধ্যে সম্পর্ক স্থাপন কর।

(খ) কীভাবে ধারাবাহিক নমুনায়ন স্তরিত দৈব নমুনায়নের সাথে সম্পর্কিত?

[(a) Establish the relationship between the variance of sample mean in systematic sampling and the variance of sample mean under SRS.

(b) How is systematic sampling related to stratified random sampling?]

১৭। (ক) গুচ্ছ নমুনায়নের সংজ্ঞা দাও। গুচ্ছ ও স্তরের মধ্যে পার্থক্য কর।

(খ) M সংখ্যক উপাদানবিশিষ্ট N গুচ্ছের সমগ্রক হতে পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নের মাধ্যমে n সংখ্যক গুচ্ছ নির্বাচন করা হলে দেখাও যে, $\hat{y} = N\bar{y}$ হলো তথ্যবিশ্ব সমষ্টির নিৰ্ঝুঁকি প্রাক্কলক যেখানে $\bar{y}$ হলো গুচ্ছ প্রতি নমুনা গড়। উক্ত প্রাক্কলকের ভেদাংক নির্ণয় কর।

[(a) Define cluster sampling. Distinguish between cluster and stratum.

(b) If n clusters are selected by simple random sampling without replacement from a population of N clusters each containing M elements then show that, $\hat{y} = N\bar{y}$ is an unbiased estimator for population total, where $\bar{y}$ is the sample mean per cluster. Find the variance of the estimator.]

পরিসংখ্যান- ২০২১

বিষয় কোড : 223603

(Sampling Technique)

সময়- ৪ ঘন্টা

পূর্ণমান- ৮০

[**দ্রষ্টব্য:** প্রতিটি বিভাগের বিভিন্ন প্রশ্নের উত্তর ধারাবাহিকভাবে লিখতে হবে ।]

ক বিভাগ

(যে কোনো **দশটি প্রশ্নের** উত্তর দাও)

মান- ১ × ১০ = ১০

১। (ক) নমুনাজমান কী?

[What is statistic?]

(খ) পরামিতির সংজ্ঞা দাও ।

[Define parameter.]

(গ) প্রশ্নমালা বলতে কী বুঝ?

[What do you mean by questionnaire?]

(ঘ) নমুনায়ন ভগ্নাংশ কী?

[What is sampling fraction?]

(ঙ) সরল দৈব নমুনা চয়নের পদ্ধতিগুলো কী কী?

[What are the methods of drawing simple random sample?]

(চ) উদ্দেশ্যমুখী নমুনায়নের একটি উদাহরণ দাও।

[Give an example of purposive sampling.]

(ছ) সর্বোত্তম বন্টন কী?

[What is optimum allocation?]

(জ) কখন নেম্যান বন্টন ও সমানুপাতিক বন্টন একই হয়?

[When Neyman's allocation is same to the proportional allocation?]

(ঝ) নিরূপকের সঠিকতা বলতে কী বুঝ?

[What do you mean by accuracy of estimator?]

(ঞ) পরিমিত বিচ্যুতি কী?

[What is standard error?]

(ট) গুচ্ছের আকার বলতে কী বুঝ?

[What do you mean by size of a cluster?]

(ঠ) স্তরিতকরণ বলতে কী বুঝায়?

[What is meant by stratification?]

খ- বিভাগ

(যে কোনো **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ৪ × ৫ = ২০

২। দৈব নমুনায়ন ও ঐচ্ছিক নমুনায়ন বলতে কী বুঝ? এদের মধ্যে পার্থক্য কর।

[What do you mean by random sampling and purposive sampling? Distinguish between them.]

৩। নমুনায়ন কাঠামো কী? একটি ভালো নমুনায়নে কাঠামোর গুণাবলি আলোচনা কর।

[What is sampling frame? Discuss the necessary characteristics of a good sampling frame.]

৪। সরল দৈব নমুনায়ন কী? পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নের ক্ষেত্রে দেখাও যে, নমুনা গড় সমগ্রক গড়ের নিরপেক্ষ নিরূপক।

[What is simple random sampling? In case of simple random sampling without replacement, show that, sample mean is an unbiased estimator of the population mean.]

৫। শুমারি জরিপ বলতে কী বুঝ? শুমারি জরিপের সুবিধা ও অসুবিধাসমূহ বর্ণনা কর।

[What do you mean by census? Describe the advantages and the disadvantages of census.]

৬। স্তরিত দৈব নমুনায়ন ও গুচ্ছ নমুনায়নের মধ্যে পার্থক্য দেখাও।

[Distinguish between stratified random sampling and cluster sampling.]

৭। দেখাও যে, পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নের মাধ্যমে N গুচ্ছবিশিষ্ট সমগ্রক হতে n গুচ্ছ চয়ন করা হলে $\bar{y}_c$ এর ভেদাংক হবে $v(\bar{y}_c)=\frac{1-f}{nM}s^2\left[1+(M-1)\rho\right]$ যেখানে, ρ অন্তঃশ্রেণি সংশ্লেষাংক এবং M গুচ্ছের আকার।

[If a sample of n clusters is drawn from a population of N clusters, then show that the variance of $\bar{y}_c$ is $v(\bar{y}_c)=\frac{1-f}{nM}s^2\left[1+(M-1)\rho\right]$; where ρ is intra-class correlation coefficient and M is the size of cluster.]

৮। সরল ধারাবাহিক ও চক্রাকার ধারাবাহিক নমুনা নির্বাচনের পদ্ধতি আলোচনা কর।

[Discuss about the simple systematic sample and cyclical systematic sample selection procedures.]

৯। সরল দৈব নমুনায়নের তুলনায় গুচ্ছ নমুনায়নের আপেক্ষিক দক্ষতা নির্ণয় কর এবং মন্তব্য কর।

[Find the relative efficiency of cluster sampling compared to simple random sampling and comment.]

গ- বিভাগ

(যে কোনো **পাঁচটি** প্রশ্নের উত্তর দাও)

মান − ১০ × ৫ = ৫০

১০। (ক) নমুনাজ ত্রুটি ও অনমুনাজ ত্রুটির কারণগুলো লেখ। এসব ত্রুটি কমানোর উপায় আলোচনা কর।

(খ) পাইলট জরিপ বলতে কী বুঝ? এ ধরণের জরিপের আবশ্যকতা আলোচনা কর।

[(a) Write down the causes of sampling error and non-sampling error. Discuss the procedure to reduce such type of error.

(b) What do you mean by pilot survey? Discuss the necessity of such type of survey.

১১। (ক) সরল দৈব নমুনায়নের সুবিধা ও অসুবিধাসমূহ লেখ।

(খ) পুনঃস্থাপনসহকারে সরল দৈব নমুনায়নের ক্ষেত্রে দেখাও যে, যেখানে প্রতীকগুলো প্রচলিত।

[(a) Write down the advantages and disadvantages of simple random sampling.

(b) In case of simple random sampling with replacement show that $E(s^2)=\sigma^2$, where the symbols are as usual.]

১২। (ক) সরল দৈব নমুনায়ন কী? দৈব সংখ্যা সারণি ব্যবহার করে কীভাবে সরল দৈব নমুনা চয়ন করা হয়, ব্যাখ্যা কর।

(খ) দেখাও যে, পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নে N আকারের সমগ্রক হতে দুটি নির্দিষ্ট এককের n আকারের নমুনায় অন্তর্ভুক্তির সম্ভাবনা $\frac{n(n-1)}{N(N-1)}$.

[(a) What is simple random sampling? Explain how a simple random sample can be selected using random number table.

(b) Show that, for SRSWOR probability of inclusion of any two specified units in the sample of size n from a population of size N is $\frac{n(n-1)}{N(N-1)}$.]

১৩। (ক) স্তরীকরণের নীতিসমূহ কী কী? স্তরিত দৈব নমুনায়নে সমগ্রক সমষ্টির নির্ঝুঁকি প্রাক্কলক বের কর এবং প্রাক্কলকটির আদর্শ বিচ্যুতি নির্ণয় কর।

(খ) সমানুপাতিক বণ্টন ও নেম্যান বণ্টনের ক্ষেত্রে স্তরিত নমুনা গড়ের ভেদাংক নির্ণয় কর।

[(a) What are the principles of stratification? In stratified random sampling find an unbiased estimator for population total and find the standard error of the estimator.

(b) Find the variance of stratified sample mean for proportional allocation and Neyman allocation.]

১৪। (ক) সরল দৈব নমুনায়নের তুলনায় স্তরিত দৈব নমুনায়নের সুবিধা ও অসুবিধা আলোচনা কর।

(খ) সর্বোত্তম বণ্টনের ক্ষেত্রে নমুনা আকারের রাশিমালা উদ্ভাবন কর।

[(a) Discuss the advantages and disadvantages of stratified random sampling over simple random sampling.

(b) Derive the expression for sample size in optimum allocation.]

১৫। (ক) ধারাবাহিক নমুনায়ন কী? ধারাবাহিক নমুনায়নের সুবিধা ও অসুবিধা আলোচনা কর।

(খ) সরলরৈখিক গতিধারাবিশিষ্ট সমগ্রক হতে নমুনায়নের ক্ষেত্রে দেখাও যে, $v(\bar{y}_{ran}) > v(\bar{y}_{sys}) > v(\bar{y}_{st})$.

[(a) What is systematic sampling? Discuss the advantages and disadvantages of systematic sampling.

(b) Show that, in case of sampling from a population with linear trend,

$$v(\bar{y}_{ran}) > v(\bar{y}_{sys}) > v(\bar{y}_{st}).]$$

১৬। (ক) ধারাবাহিক নমুনায়ন বলতে কী বুঝ? ধারাবাহিক নমুনায়নে তথ্যবিশ্ব সমষ্টির নিরপেক্ষ প্রাক্কলক বের কর এবং প্রাক্কলকটির আদর্শ বিচ্যুতি নির্ণয় কর।

(খ) দেখাও যে, সরল দৈব নমুনায়নের তুলনায় ধারাবাহিক নমুনায়ন বেশি দক্ষ হবে যদি $\rho_w < -\frac{1}{n_k - 1}$.

[(a) What do you mean by systematic sampling? In systematic sampling find an unbiased estimator for population total and find the standard error of the estimator.

(b) Show that systematic sampling will be more efficient than simple random sampling if $\rho_w < -\frac{1}{n_k - 1}$.]

১৭। (ক) গুচ্ছ নমুনায়ন কী? গুচ্ছ নমুনায়নের সুবিধাসমূহ আলোচনা কর।

(খ) সংখ্যক উপাদানবিশিষ্ট গুচ্ছেক সমগ্রক হতে সরল দৈব নমুনায়নের মাধ্যমে সংখ্যক গুচ্ছ নির্বাচন করা হলে, দেখাও যে, নমুনা গুচ্ছ গড়ের গড় সমগ্রক গড়ের নিরপেক্ষ নিরূপক।

[(a) What is cluster sampling? Discuss the advantages of cluster sampling.

(b) If n clusters are selected by simple random sampling from a population of N clusters each with M elements, then show that mean of sample cluster mean is an unbiased estimator for population mean.]

পরিসংখ্যান- ২০২২

বিষয় কোড : 223603

(Sampling Technique)

সময়- ৪ ঘন্টা

পূর্ণমান- ৮০

[**দ্রষ্টব্য:** প্রতিটি বিভাগের বিভিন্ন প্রশ্নের উত্তর ধারাবাহিকভাবে লিখতে হবে ।]

ক বিভাগ

(যে কোনো **দশটি প্রশ্নের** উত্তর দাও)

মান- ১ × ১০ = ১০

১। (ক) নমুনায়ন বলতে কী বুঝ?

[What do you mean by sampling?]

(খ) সমগ্রকের সংজ্ঞা দাও।

[Define population?]

(গ) পাইলট জরিপ কী?

[What is pilot survey?]

(ঘ) স্তর কী?

[What is stratum?]

(ঙ) প্রাক্কলকের নিরপেক্ষতা বলতে কী বুঝ?

[What do you mean by unbiasedness of an estimation?]

(চ) F. P. C. কী?

[What is F. P. C.?]

(ছ) মিশ্র নমুনায়নের একটি উদাহরণ দাও।

[Give an example of mixed sampling.]

(জ) স্তরিতকরণের কারণসমূহ লেখ।

[Write down the reasons of stratification.]

(ঝ) গুচ্ছ কী?

[What is cluster?]

(ঞ) কখন গুচ্ছ নমুনায়ন ব্যবহার করা হয়?

[When is cluster sampling used?]

(ট) নমুনায়ন কাঠামো বলতে কী বুঝ?

[What do you mean by sampling frame?]

(ঠ) ৩, ৪, ৫ এবং ৬ উপাদানবিশিষ্ট সমগ্রক হতে পুনঃ স্থাপন ব্যতিরেকে ২ আকারবিশিষ্ট সম্ভাব্য সরল দৈব নমুনাগুলো লেখ।

[Wrtie down the all-possible simple random samples of size 2 without replacement from a population having units 3, 4, 5 and 6.]

খ- বিভাগ

(যে কোনো **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ৪ × ৫ = ২০

২। নমুনা জরিপ ও শুমারী জরিপ কী? এদের মধ্যে পার্থক্য লেখ।

[What is sample survey and census? Distinguish between them.]

৩। একটি ভালো নমুনায়ন পরিকল্পনার আবশ্যকীয়তাসমূহ বর্ণনা কর।

[Describe the requirements of a good sampling design.]

৪। দেখাও যে, সরল দৈব নমুনাতে একটি নির্দিষ্ট এককের অন্তর্ভুক্ত হবার সম্ভাবনা $\frac{n}{N}$, যেখানে n এবং N যথাক্রমে নমুনার আকার ও সমগ্রকের আকার।

[In SRS show that the probability of a specified unit being included in the sample is $\frac{n}{N}$; where n and N are the sample size and population size respectively.]

৫। শুমারি জরিপ অপেক্ষা নমুনা জরিপের সুবিধাবলি আলোচনা কর।

[Describe the advantages of a sample survey in comparison with census.]

৬। দেখাও যে, স্থির নমুনা আকারের জন্য স্তরিত দৈব নমুনায়নের $v(\bar{y}_{st})$ ক্ষুদ্রতম হবে যদি $n_h \infty N_h S_h$ হয়, যেখানে প্রতীকসমূহ চিরাচরিত।

[Show that for fixed sample size, $v(\bar{y}_{st})$ in stratified random sampling will be least if $n_h \infty N_h S_h$ where the symbols are as usual.]

৭। গুচ্ছ নমুনায়নের সংজ্ঞা দাও। এর অসুবিধাগুলো আলোচনা কর।

[Define cluster sampling. Discuss its disadvantages.]

৮। চক্রাকার ধারাবাহিক নমুনায়ন বলতে কী বুঝ? এটি কখন প্রয়োজন হয়? পদ্ধতিটির একটি উদাহরণ দাও।

[What do you mean by cyclical systematic sampling? When it is required? Give an example of this method.]

৯। ধারাবাহিক নমুনায়ন এবং স্তরিত দৈব নমুনায়ন এর মধ্যে তুলনা কর।

[Compare systematic sampling with stratified random sampling.]

গ- বিভাগ

(যে কোনো **পাঁচটি** প্রশ্নের উত্তর দাও)

মান – ১০ × ৫ = ৫০

১০। (ক) উত্তরদাতার নিরবতা কী? নমুনা জরিপে এর অসুবিধাগুলো আলোচনা কর।

(খ) নমুনা জরিপের প্রধান ধাপসমূহ বর্ণনা কর।

[(a) What is non-response? Discuss its disadvantages in sample survey.

(b) Describe the main steps involved in sample survey.]

১১। (ক) পুনঃস্থাপনসহ ও পুনঃস্থাপনছাড়া নমুনায়ন বলতে কী বুঝ? এদের মধ্যকার পার্থক্য লেখ।

(খ) পুনঃস্থাপনছাড়া সরল দৈব নমুনায়নের ক্ষেত্রে প্রচলিত সংকেতমালায় দেখাও যে, $v(\bar{y}) = (1-f)\frac{S^2}{n}$.

[(a) What do you mean by sampling with replacement and without replacement? Distinguish between them.

(b) In simple random sampling without replacement with usual notations, show that $v(\bar{y}) = (1-f)\frac{S^2}{n}$.]

১২। (ক) সরল দৈব নমুনায়নে অনুপাত বলতে কী বুঝ? দেখাও যে, নমুনা অনুপাত সমগ্রক অনুপাতের নিরপেক্ষ নিরুপক।

(খ) নুমনা অনুপাত p হলে, $v(p)$ নির্ণয় কর।

[(a) What do you mean by proportion in simple random sampling? Show that, sample proportion is an unbiased estimate of population proportion.

(b) If sample proportion is p then calculate $v(p)$.]

১৩। (ক) বাস্তবে কীভাবে স্তরিতকরণ করা যায় তা একটি উদাহরণের সাহায্যে ব্যাখ্যা কর। সরল দৈব নমুনায়নের চেয়ে স্তরিত নমুনায়নের সুবিধাসমূহ কী কী?

(খ) দেখাও যে, সচরাচর প্রতীকে $\bar{y}_{st}$ সমগ্রক গড়ের একটি নিরপেক্ষ নিরূপক। আরও দেখাও যে,

$$V(\bar{y}_{st}) = \frac{1}{N^2}\sum_{h=1}^{L} N_h(N_h - n_h)\frac{S_h^{\,2}}{n_h}.$$

[(a) Explain with an example how stratification can be done in practice. What are the advantages of stratified sampling over simple random sampling?

(b) Show that under usual notation $\bar{y}_{st}$ is an unbiased estimator of the population mean. Also show that, $V(\bar{y}_{st}) = \frac{1}{N^2}\sum_{h=1}^{L} N_h(N_h - n_h)\frac{S_h^{\,2}}{n_h}$.]

১৪। (ক) স্তরিত দৈব নমুনায়নে ব্যবহৃত নমুনা বণ্টনের বিভিন্ন পদ্ধতিসমূহ বর্ণনা কর।

(খ) প্রচলিত সংকেতে দেখাও যে, $v(\bar{y}_{st})_{opt} \le v(\bar{y}_{st})_{prop} \le v(\bar{y}_{st})_{ran}$.

[(a) Describe the different types of sample collection used in stratified random sampling.

(b) For usual notations, show that $v(\bar{y}_{st})_{opt} \le v(\bar{y}_{st})_{prop} \le v(\bar{y}_{st})_{ran}$.]

১৫। (ক) কীভাবে ধারাবাহিক নমুনায়ন স্তরিত দৈব নমুনায়নের সাথে সম্পর্কিত?

(খ) দেখাও যে, ধারাবাহিক নমুনায়ন সরল দৈব নমুনায়নের চেয়ে ভালো প্রাক্কলক প্রদান করবে যদি $S^2_{wsy} > S^2$ হয় (প্রতীকগুলো চিরাচরিত)।

[(a) How is systematic sampling related to stratified random sampling?

(b) Show that, systematic sampling gives more efficient estimator than simple random sampling if $S^2_{wsy} > S^2$ (notations are as usual).]

১৬। (ক) ধারাবাহিক নমুনায়নের ক্ষেত্রে সমগ্রক গড়ের নিঝুঁকি প্রাক্কলক নির্ণয় কর এবং দেখাও যে,

$v\left(\bar{y}_{sys}\right)=\frac{(N-1)S^2}{Nn}\left[1+(n-1)\rho_w\right]$ যেখানে ρ_w হল অন্তঃশ্রেণি সংশ্লেষাংক।

(খ) একটি নিয়মক্রমিক নমুনা নির্বাচনের পদ্ধতি আলোচনা কর।

[(a) Find the unbiased estimator of population mean in case of systematic sampling and show that $v\left(\bar{y}_{sys}\right)=\frac{(N-1)S^2}{Nn}\left[1+(n-1)\rho_w\right]$ where ρ_w is the intra-class correlation co-efficient.

(b) Discuss a method to select a systematic sample.]

১৭। (ক) গুচ্ছ ও স্তরের মধ্যে পার্থক্য কী? গুচ্ছ নমুনায়ন ব্যবহৃত হয় এমন কিছু ক্ষেত্র উল্লেখ কর।

(খ) M সংখ্যক এককবিশিষ্ট N গুচ্ছের সমগ্রক হতে n গুচ্ছের একটি সরল দৈব নমুনা পুনঃস্থাপন ছাড়া নেয়ার ক্ষেত্রে দেখাও যে $\hat{y}_C = NM\bar{y}$ হলো সমগ্রক সমষ্টি (Y) এর নির্ঝুঁকি প্রাক্কলক এবং এই প্রাক্কলকের ভেদাংক নিম্নরূপঃ

$V\left(\hat{Y}_C\right)=N^2M\left(\frac{1-f}{n}\right)S^2\left[1+(M-1)\rho\right]$ যেখানে হলো অন্তঃগুচ্ছ সংশ্লেষাংক।

[(a) What are the differences between cluster and strata? Mention some situation where cluster sampling is used.

(b) In SRSWOR of n cluster each containing M elements from a population of N cluster. Show that $\hat{y}_C = NM\bar{y}$ is an unbiased estimate of population total (Y) and its variance is given by

$V\left(\hat{Y}_C\right)=N^2M\left(\frac{1-f}{n}\right)S^2\left[1+(M-1)\rho\right]$; where ρ is the intra-cluster correlation co-efficient.]

পরিসংখ্যান- ২০২৩

বিষয় কোড : 223603

(Sampling Technique)

সময়- ৪ ঘন্টা

পূর্ণমান- ৮০

[**দ্রষ্টব্য:** প্রতিটি বিভাগের বিভিন্ন প্রশ্নের উত্তর ধারাবাহিকভাবে লিখতে হবে।]

ক বিভাগ

(যে কোনো **দশটি প্রশ্নের** উত্তর দাও)

মান- ১ × ১০ = ১০

১। (ক) প্রাক্কলক বলতে কী বুঝ?

[What do you mean by estimator?]

(খ) প্রশ্নমালা কী?

[What is questionnaire?]

(গ) কোটা নমুনায়ন বলতে কী বুঝ?

[What do you mean by quota sampling?]

(ঘ) নমুনায়ন ভগ্নাংশ কী?

[What is sampling fraction?]

(ঙ) এমন একটি ক্ষেত্র উল্লেখ কর, যেখানে নমুনা জরিপই অত্যাবশ্যক।

[Mention a case, where sample survey is must.]

(চ) লা-জবাব বলতে কী বুঝ?

[What is meant by non-response?]

(ছ) নেম্যান বণ্টন কী?

[What is Neyman allocation?]

(জ) স্তরিত নমুনায়নের ব্যয় অপেক্ষকটি লেখ।

[Write down the cost function of stratified sampling.]

(ঝ) স্তরিত দৈব নমুনায়নের ক্ষেত্রে সমগ্রক সমষ্টির নিরপেক্ষ নিরূপকটি লেখ।

[Write down the unbiased estimate of population total in case of stratified random sampling.]

(ঞ) কখন চক্রাকার ধারাবাহিক নমুনায়ন ব্যবহার করা হয়?

[When circular systematic sampling is used?]

(ট) গুচ্ছের আকার বলতে কী বুঝ?

[What do you mean by size of a cluster?]

(ঠ) কখন স্তরিত দৈব নমুনায়নের পরিবর্তে গুচ্ছ নমুনায়ন ব্যবহার করা হয়?

[When cluster sampling is used in stead of stratified sampling?]

খ- বিভাগ

(যে কোনো **পাঁচটি** প্রশ্নের উত্তর দাও)

মান- ৪ × ৫ = ২০

২। সম্ভাবনা নমুনায়ন ও বিচারভিত্তিক নমুনায়ন বলতে কী বুঝ? এদের মধ্যে পার্থক্য কর।

[What do you mean by probability sampling and judgement sampling? Distinguish between them.]

৩। প্রাক-জরিপ সংজ্ঞায়িত কর। এর প্রয়োজনীয়তা আলোচনা কর।

[Define pilot-survey. Discuss it's usefulness.]

৪। সমগ্রকের পরামান প্রাক্কলনে একটি জরিপ পরিকল্পনার নির্বাচিত নমুনা আকার নির্ধারণের বিভিন্ন উপায় বর্ণনা কর

[Discuss different procedures for determining the sample size in planning a sample survey for estimating the population parameter.]

৫। পুনঃস্থাপন ছাড়া সরল দৈব নমুনায়নের ক্ষেত্রে প্রমাণ কর যে, $E(s^2)=S^2$; প্রতীকগুলো চিরাচরিত।

[For simple random sampling without replacement, prove that $E(s^2)=S^2$; notations are as usual.]

৬। n আকার বিশিষ্ট একটি সরল দৈব নমুনার জন্য দেখাও যে, পুনঃস্থাপন ছাড়া নমুনা গড়ের পরিমিত বিচ্যুতি পুনঃস্থাপন সহকারে নমুনা গড়ের পরিমিত বিচ্যুতি অপেক্ষা ছোট।

[For a simple random sample of size n, show that the standard error of sample mean in sampling without replacement is smaller than that in sampling with replacement.]

৭। সর্বোত্তম বণ্টনের ক্ষেত্রে নমুনা আকারের রাশিমালা উদ্ভাবন কর।

[Derive the expression for sample size in case of optimum allocation.]

৮। সচরাচর প্রতীকে দেখাও যে, $V(\bar{y}_{sys})=\frac{N-1}{N}S^2-\frac{K(n-1)}{N}S^2{}_{wsy}$.

[Under usual notation, show that $V(\bar{y}_{sys})=\frac{N-1}{N}S^2-\frac{K(n-1)}{N}S^2{}_{wsy}$.]

৯। সরল দৈব নমুনায়নের তুলনায় গুচ্ছ নমুনায়নের আপেক্ষিক দক্ষতা নির্ণয় কর এবং মন্তব্য কর।

[Find the relative efficiency of cluster sampling compared to simple random sampling and comment.]

গ- বিভাগ

(যে কোনো **পাঁচটি** প্রশ্নের উত্তর দাও)

মান − ১০ × ৫ = ৫০

১০। (ক) নমুনা পরিকল্পনা ও পরীক্ষণ পরিকল্পনার মধ্যে পার্থক্য কী? একটি নমুনা পরিকল্পনার নীতিগুলো বর্ণনা কর।

(খ) একটি ভালো নমুনা কাঠামোর বৈশিষ্ট্যগুলো লেখ।

[(a) What is the difference between sample design and expreimental design? Describe the principles of sample design.

(b) Write down the charactersitics of a good sampling frame.]

১১। (ক) একটি আদর্শ প্রশ্নমালার আবশ্যকীয় গুণাবলি উল্লেখ কর। তালিকার সাথে প্রশ্নমালার পার্থক্য কী?

(খ) বাংলাদেশের পরিপ্রেক্ষিতে প্রশ্নমালার মাধ্যমে তথ্য সংগ্রহের অসুবিধাগুলো সম্পর্কে আলোচনা কর।

[(a) Mention the necessary characteristics of an ideal questionnaire. What is the difference of questionnaire compared to schedule?

(b) Discuss about the disadvantages of data collection by a questionnaire in context of Bangladesh.]

১২। (ক) নমুনাজ ত্রুটি ও অনমুনাজ ত্রুটি বলতে কী বুঝ? এসব ত্রুটির উৎসগুলো লেখ।

(খ) এসব ত্রুটি কমানোর উপায়সমূহ আলোচনা কর।

[(a) What is meant by sampling error and non-sampling error? Write down the sources of such types of error.

(b) Discuss the procedure to reduce such types of error.]

১৩। (ক) সরল দৈব নমুনায়ন বলতে কী বুঝ? দৈব সংখ্যা সারণী ব্যবহার করে একটি সরল দৈব নমুনা চয়ন করার পদ্ধতি বর্ণনা কর।

(খ) দেখাও যে, পুনঃস্থাপন ব্যতীত সরল দৈব নমুনায়নে একটি নির্দিষ্ট এককের যে কোনো চয়নে নমুনায় নির্বাচিত হওয়ার সম্ভাবনা, একককটি প্রথম চয়নে নির্বাচিত হওয়ার সম্ভাবনার সমান।

[(a) What do you mean by simple random sampling? Describe the procedure to draw a simple random sample using random number table.

(b) Show that, in simple random sampling without replacement, the probability of selecting a specified unit in the sample in any draw is equal to the probability of selecting the unit in the first draw.]

১৪। সরল দৈব নমুনায়নের ক্ষেত্রে সমগ্রক সমষ্টির নিরপেক্ষ নিরূপক বের কর। উক্ত নিরূপকের আদর্শ বিচ্যুতি নির্ণয় কর।

[Find the unbiased estimate of population total in case of simple random sampling. Also find the standard error of this estimator.]

১৫। (ক) স্তরিত দৈব নমুনায়নে সমগ্রক সমষ্টির নিরপেক্ষ নিরূপক বের কর এবং নিরূপকটির ভেদাংক নির্ণয় কর।

(খ) সমানুপাতিক বণ্টন ও নেম্যান বণ্টনের ক্ষেত্রে স্তরিত দৈব নমুনা গড়ের ভেদাংক নির্ণয় কর।

[(a) In stratified random sampling find an unbiased estimator of population total and find its variance.

(b) Find the variance of stratified random sample mean for proportional allocation and Neyman allocation.]

১৬। (ক) ধারাবাহিক নমুনায়ন বলতে কী বুঝ? এর সুবিধা ও অসুবিধাগুলো লেখ।

(খ) সরলরৈখিক গতিধারাবিশিষ্ট সমগ্রক হতে নমুনায়নের ক্ষেত্রে দেখাও যে, $V(\bar{y}_{st}) \le V(\bar{y}_{sys}) \le V(\bar{y}_{ran})$.

[(a) What do you mean by systematic sampling? Write down its advantages and disadvantages.

(b) In case of sampling from a population with linear trend show that,

$V(\bar{y}_{st}) \le V(\bar{y}_{sys}) \le V(\bar{y}_{ran})$.]

১৭। (ক) গুচ্ছ নমুনায়ন ব্যবহারের কারণসমূহ কী কী?

(খ) ধরা যাক, N গুচ্ছবিশিষ্ট সমগ্রকের প্রতিটি গুচ্ছে M সংখ্যক একক আছে। উক্ত সমগ্রক হতে n গুচ্ছের একটি নমুনা পুনঃস্থাপন ব্যতিরেকে নেওয়া হল। যদি নমুনার প্রতি এককের গড় $\bar{y}_c$ হয়, তবে দেখাও যে, $E(\bar{y}_c) = \bar{Y}$ এবং $V(\bar{y}_c)$ নির্ণয় কর।

[(a) What are the reasons for using cluster sampling?

(b) Suppose in a population with N clusters, each cluster has M sampling units. A simple random sample of n clusters is taken without replacement. If $\bar{y}_c$ is the sample mean per unit then show that $E(\bar{y}_c) = \bar{Y}$ and also find $V(\bar{y}_c)$.]