৫

বিস্তৃতির পরিমাপ
Measures of Dispersion

ইউনিট ৪-এ আমরা উপাত্তরাশির কেন্দ্রীয় প্রবণতার বিভিন্ন রূপ আলোচনা করতে গিয়ে দেখেছি যে, কেন্দ্রীয় প্রবণতার সার-সংক্ষেপমূলক পরিমাপগুলো একটি বিন্যাসের প্রতিনিধিত্বশীল আদর্শ সংখ্যা মানকে নির্দেশ করে। কিন্তু একটি বিন্যাসকে বর্ণনার জন্য এই প্রতিনিধিত্বশীল আদর্শ মানগুলো যথেষ্ট নয়। কারণ, বিন্যাসের অন্তর্ভুক্ত উপাত্তরাশি একটি বিশেষ স্থানে কেন্দ্রীভূত হবার প্রবণতা প্রদর্শন করলেও প্রতিটি মান সেই কেন্দ্রীয় মানটি থেকে কিছুটা ভিন্ন। এই ভিন্নতার মাত্রা যত বেশী হবে কেন্দ্রীয় আদর্শ মানগুলোর প্রতিনিধিত্বশীলতা তত বেশী হ্রাস পাবে। অর্থাৎ, একটি উপাত্তরাশির বিন্যাসের মধ্যে সমরূপতা ও ভিন্নতার দু'টি উপাদানই বিদ্যমান থাকে। অতএব, উপাত্তরাশির বিস্তৃতি সম্পর্কে জানা না থাকলে একটি বিন্যাসের সম্পূর্ণ বর্ণনা সম্পন্ন করা যায় না। উপাত্তের প্রকৃতি ও বিশ্লেষণের প্রয়োজনীয়তার উপর ভিত্তি করে বিস্তৃতির বিভিন্ন পরিমাপ নির্ণয় করা যায়। বিস্তৃতির পরিমাপগুলো কেন্দ্রীয় প্রবণতার পরিমাপগুলোর প্রতিনিধিত্বশীলতার মাত্রা সম্পর্কে একটি সূক্ষ্ম ধারণা প্রদান করে। বিস্তৃতির দু'ধরণের পরিমাপ রয়েছে — নিরঙ্কুশ ও আপেক্ষিক। নিরঙ্কুশ পরিমাপগুরো উপাত্তের মধ্যে ভিন্ন ভিন্ন পরিমাপ এককের মাধ্যমে ভিন্নতার বর্ণনা দেয় এবং আপেক্ষিক পরিমাপগুলো কোন এককে প্রকাশিত হয় না বলে সেগুলো একটি পরিমিত মাত্রায় ভিন্নতার তুলনার জন্য উপযোগী।

এই ইউনিটে আমরা যে পাঠগুলো অধ্যয়ন করবো সেগুলো হলো:

- পাঠ - ১ : বিস্তৃতির সংজ্ঞা, প্রকৃতি ও গুরুত্ব
- পাঠ - ২ : বিস্তৃতির নিরঙ্কুশ পরিমাপ – ১: পরিসর ও চতুর্থ ব্যবধান
- পাঠ - ৩ : বিস্তৃতির নিরঙ্কুশ পরিমাপ – ২: গড় ব্যবধান ও শতহারিক পরিসর
- পাঠ - ৪ : বিস্তৃতির নিরঙ্কুশ পরিমাপ – ৩: পরিমিত ব্যবধান
- পাঠ - ৫ : বিস্তৃতির নিরঙ্কুশ পরিমাপ – ৪: ভেদাঙ্ক
- পাঠ - ৬ : পরিমিত ব্যবধান ও ভেদাঙ্কের ব্যাখ্যা এবং ব্যবহার
- পাঠ - ৭ : বিস্তৃতির আপেক্ষিক পরিমাপ – ব্যবধানাঙ্ক এবং পরিমিত মান
- পাঠ - ৮ : বিস্তৃতির পরিমাপের বিভিন্ন পদ্ধতির মধ্যে তুলনা এবং নির্বাচন

পাঠ - ১

বিস্তৃতির সংজ্ঞা, প্রকৃতি ও গুরুত্ব
Definition, Nature and Importance of Dispersion

এই পাঠ শেষে যা জানা যাবে —

- বিস্তৃতির পরিমাপ কি
- বিস্তৃতির প্রকৃতি
- বিস্তৃতির পরিমাপের গুরুত্ব
- বিস্তৃতির একটি উত্তম পরিমাপের বৈশিষ্ট্য
- বিস্তৃতির পরিমাপের প্রকারভেদ

বিস্তৃতির পরিমাপ কি (What is Measures of Dispersion)?

বিন্যাসের অন্তর্ভুক্ত উপাত্তরাশির বিস্তৃতি সম্পর্কে না জানলে একটি বিন্যাসের সম্পূর্ণ বর্ণনা সম্পন্ন করা যায় না।

কেন্দ্রীয় প্রবণতার পরিমাপগুলোর আলোচনার সময় দেখেছি যে, সেগুলো একটি বিন্যাসের অন্তর্ভুক্ত উপাত্তরাশির গুচ্ছবদ্ধতাকে বর্ণনা করে, মধ্য মানের অবস্থান নির্ণয় করে এবং অধিক পরিমাণে ঘটমান সংখ্যাকে চিহ্নিত করে। এইসব সার-সংক্ষেপমূলক পরিমাপের মাধ্যমে আমরা একটি বিন্যাসের আদর্শ মানগুলোকে পাই যা সেই বিন্যাসের প্রতিনিধিত্বশীল বৈশিষ্ট্যকে প্রকাশ করে এবং দুই বা ততোধিক বিন্যাসের মধ্যে তুলনাকে সহজতর করে। কিন্তু একটি বিন্যাসকে বর্ণনার জন্য এই প্রতিনিধিত্বশীল একক মানটি যথেষ্ট নয়। বিন্যাসের অন্তর্ভুক্ত উপাত্তরাশির বিস্তৃতি (dispersion of data) সম্পর্কে না জানলে একটি বিন্যাসের সম্পূর্ণ বর্ণনা সম্পন্ন করা যায় না। এই পাঠে আমরা উপাত্তের বিস্তৃতি সম্পর্কে আলোচনা করবো।

এ বিষয়ে আলোচনার পূর্বে কেন্দ্রীয় মানগুলোর প্রতিনিধিত্বশীলতার বিষয়টি বিবেচনা করা যাক। পরিসংখ্যানগত উপাত্তের অন্তর্নিহিত (inherent) বৈশিষ্ট্যটি হলো যে, সমরূপতার উপাদানটি (element of homogeneity) বজায় রেখেই সেগুলো বিভিন্ন মাত্রায় ভিন্নতা (variability) প্রদর্শন করে থাকে। একটি উদাহরণ দেয়া যাক। ধরা যাক, ৫০ নম্বরের একটি পরীক্ষায় ১০ জন ছাত্র-ছাত্রী নিম্নলিখিত নম্বর পেয়েছে,

২০ ২৩ ২৪ ২৬ ২৭ ২৮ ২৮ ২৯ ৩০ ৪৫

এই সংখ্যাগুলোর গাণিতিক গড় হলো,

$$\bar{x} = \frac{\sum x_i}{N}$$

$$= \frac{২০ + ২৩ + ২৪ + ২৬ + ২৭ + ২৮ + ২৮ + ২৯ + ৩০ + ৪৫}{১০}$$

$$= \frac{২৮০}{১০}$$

$$\therefore \quad \bar{x} = ২৮$$

নির্ণীত গড় সংখ্যাটি এই ১০টি রাশিমালার প্রতিনিধিত্বকারী মান হলেও লক্ষ্য করা যায় যে, মাত্র দু'জন ছাত্র-ছাত্রী ২৮ নম্বর পেয়েছে, ৩ জন পেয়েছে তার চেয়ে বেশী যার সর্বোচ্চ সংখ্যাটি হলো ৪৫ এবং ৫ জন পেয়েছে তার চেয়ে কম যার সর্বনিম্ন সংখ্যাটি হলো ২০। অর্থাৎ, প্রতিটি পরিসংখ্যানিক উপাত্তের ভিন্নতা এখানে লক্ষ্য করা যাচ্ছে। শুধু তাই নয়, দেখা যাচ্ছে যে, উপাত্ত রাশিমালা গড় থেকে একটি

বিস্তৃত পরিধি নিয়ে ছড়িয়ে রয়েছে। কেন্দ্রীয় মান থেকে উপাত্তরাশির এই ছড়িয়ে থাকার প্রবণতাকেই বিস্তৃতি বলে।

বিস্তৃতির এই পরিধি যেহেতু ২০ থেকে ৪৫ পর্যন্ত ছড়িয়ে রয়েছে, সেহেতু সহজেই প্রশ্ন তোলা যেতে পারে যে, গড় মান ২৫ কি সত্যিই এই উপাত্তরাশিকে প্রতিনিধিত্ব করে? যদি করে, তবে কতটুকু করে? একটি বিন্যাসকে সঠিকভাবে বর্ণনার জন্য এই প্রশ্ন দু'টির সমাধান অত্যন্ত জরুরী। যে সব কৌশল প্রয়োগ করে এ সকল প্রশ্নের সমাধান পাওয়া যায় সেগুলো বিস্তৃতির পরিমাপ বলে পরিচিত। সুনির্দিষ্টভাবে বলা যায় যে, বিস্তৃতির পরিমাপ হলো সে সকল পরিমাপ যা কেন্দ্রীয় মান থেকে প্রতিটি মানের ছড়িয়ে থাকার গড় মাত্রাকে নিরূপণ করে। বিস্তৃতির পরিমাপগুলো কেন্দ্রীয় প্রবণতার পরিমাপগুলোর প্রতিনিধিত্বশীলতার মাত্রা সম্পর্কে একটি সূক্ষ্ম ধারণা দিয়ে থাকে।

বিস্তৃতির পরিমাপ হলো সে সব পরিমাপ যা কেন্দ্রীয় মান থেকে প্রতিটি মানের ছড়িয়ে থাকার গড় মাত্রাকে নিরূপণ করে।

বিস্তৃতির প্রকৃতি (Nature of Dispersion)

উপাত্তরাশির মধ্যে বিস্তৃতির প্রকৃতিকে বুঝতে হলে আমাদের দু'টি বিষয় সম্পর্কে জানতে হবে। প্রথমটি হলো, বিন্যাসের অন্তর্ভুক্ত উপাত্ত রাশিমালার প্রতিটি মানের আকার কতটুকু ভিন্ন এবং প্রতিটি মান গড় থেকে কতটুকু ভিন্ন। দ্বিতীয় বিষয়টি হলো, মানগুলো গড়ের চারপাশে কিভাবে ছড়িয়ে রয়েছে? অর্থাৎ, গড় মানের উপরে বেশী মান রয়েছে, নাকি নীচে বেশী মান রয়েছে? একটি উদাহরণ দিয়ে বিষয়টি স্পষ্ট করে বোঝা যাক। ধরা যাক, ১০ জন ছাত্র-ছাত্রী দু'টি পরীক্ষায় অংশগ্রহণ করেছে। ফলাফল প্রকাশের পর দেখা গেল যে, দু'টি পরীক্ষায় প্রাপ্ত নম্বরের গড় মান একই, কিন্তু প্রাপ্ত নম্বরগুলোর মধ্যে পার্থক্য রয়েছে। সারণি ৫.১.১-এ দু'টি পরীক্ষার গড় মানসহ প্রাপ্ত নম্বরের বিন্যাসটি উপস্থাপন করা হলো।

উপাত্তরাশির মধ্যে বিস্তৃতির প্রকৃতিকে বুঝতে হলে আমাদের দু'টি বিষয় সম্পর্কে জানতে হবে। প্রথমটি হলো, বিন্যাসের অন্তর্ভুক্ত উপাত্ত রাশিমালার প্রতিটি মানের আকার কতটুকু ভিন্ন এবং প্রতিটি মান গড় থেকে কতটুকু ভিন্ন। দ্বিতীয় বিষয়টি হলো, মানগুলো গড়ের চারপাশে কিভাবে ছড়িয়ে রয়েছে? অর্থাৎ, গড় মানের উপরে বেশী মান রয়েছে, নাকি নীচে বেশী মান রয়েছে?

সারণি ৫.১.১: দু'টি পরীক্ষায় ১০ জন ছাত্র-ছাত্রীর প্রাপ্ত নম্বরের বিন্যাস

পরীক্ষা	*প্রাপ্ত নম্বর*										*গড়*
পরীক্ষা ১	২৫	২৫	২৬	২৭	২৭	২৯	২৯	৩০	৩০	৩২	২৮
পরীক্ষা ২	২০	২৩	২৪	২৬	২৭	২৮	২৮	২৯	৩০	৪৫	২৮

যেহেতু দু'টি পরীক্ষায় প্রাপ্ত নম্বরের গড় মান ২৮, সেহেতু যে কেউ এই উপসংহারে উপনীত হতে পারেন যে, দু'টি পরীক্ষায়ই ছাত্র-ছাত্রীদের কৃতকার্যতার মাত্রা একই রকম। কিন্তু দু'টি উপাত্ত বিন্যাস কিছুটা নিবিড়ভাবে পর্যবেক্ষণ করলে দেখা যায় যে, পরীক্ষা ১-এ ছাত্র-ছাত্রীদের প্রাপ্ত নম্বরগুলো গড় মানের খুব কাছাকাছি। গড় মান থেকে সর্বনিম্ন মানের পার্থক্য হলো ৩ এবং সর্বোচ্চ মানের পার্থক্য হলো ৪। এ ক্ষেত্রে বলা যায় যে, গড় মানটি বিন্যাসটিকে যথাযথভাবে প্রতিনিধিত্ব করেছে। কিন্তু পরীক্ষা ২-এ প্রতিটি নম্বর গড় মান থেকে বেশ ছড়িয়ে ছিটিয়ে রয়েছে। গড় থেকে সর্বনিম্ন মানের পার্থক্য হলো ৮ এবং সর্বোচ্চ মানের পার্থক্য হলো ১৭। গড় মান এই বিন্যাসটিকে তেমনভাবে প্রতিনিধিত্ব করে না বললেই চলে। অতএব, উপসংহারে বলা যায় যে, উপাত্তরাশির মধ্যে বিস্তৃতি একটি স্বাভাবিক ব্যাপার এবং এই বিস্তৃতির মাত্রা একটি বিন্যাস থেকে আরেকটি বিন্যাসে ভিন্ন থেকে ভিন্নতর হয়ে থাকে। কাজেই, শুধুমাত্র কেন্দ্রীয় মানের ভিত্তিতে উপাত্ত বিন্যাসকে বর্ণনা করা যথার্থ নয়। কেন্দ্রীয় মানের সাথে বিস্তৃতির বিষয়টিকেও বিবেচনায় আনতে হবে।

বিস্তৃতির পরিমাপের গুরুত্ব (Importance of Measures of Dispersion)

আমরা জেনেছি যে, কেন্দ্রীয় প্রবণতার পরিমাপগুলো দিয়ে আমরা উপাত্তরাশির কেন্দ্রীভবনের গড় প্রবণতাকে জানতে পারি এবং বিস্তৃতির পরিমাপগুলো দিয়ে কেন্দ্রীয় মান থেকে প্রতিটি মানের গড় বিস্তৃতিকে জানতে পারি। একটি বিন্যাসকে বর্ণনা বা একাধিক বিন্যাসকে তুলনার ক্ষেত্রে কেন্দ্রীয় মান যথেষ্ট নয়। বিস্তৃতির পরিমাপ হলো একটি অত্যাবশ্যকীয় পরিসংখ্যানিক প্রয়োজনীয়তা। কারণ, বিস্তৃতির পরিমাপগুলো একটি উপাত্ত রাশিমালার ভিতরে প্রতিটি মানের ভূমিকাটিকে স্পষ্ট করে তোলে; একটি উপাত্ত রাশিমালার ভিতরে সমরূপতা নির্ণয়ের উপায় হিসাবে ব্যবহৃত হয়; কেন্দ্রীয় মানের নির্ভরযোগ্যতা নির্ধারণে সাহায্য করে; উপাত্তরাশির মধ্যে ভিন্নতা নিয়ন্ত্রণের ভিত্তি হিসাবে কাজ করে; দুই বা ততোধিক বিন্যাসের বিস্তৃতিকে তুলনা করতে সাহায্য করে; এবং অন্যান্য উচ্চতর পরিসংখ্যানিক কৌশলের ব্যবহারকে সহজতর করে।

উপাত্তরাশির বিস্তৃতি পরিমাপের পর যদি দেখা যায় যে, উপাত্তরাশির মানগুলো কেন্দ্রীয় মান থেকে খুব বেশী ছড়িয়ে আছে, তখন আমরা বলতে পারি যে, বিন্যাসটির মধ্যে সমরূপতার অভাব রয়েছে। আর

যদি মানগুলো খুব কাছাকাছি থাকে তবে বলা যায় যে, বিন্যাসটি সমরূপ রয়েছে এবং কেন্দ্রীয় মান এ ক্ষেত্রে যথাযথভাবে প্রতিনিধিত্বশীল ভূমিকা পালন করেছে।

বিস্তৃতির একটি উত্তম পরিমাপের বৈশিষ্ট্য (Properties of a Good Measure of Dispersion)

বিস্তৃতির পরিমাপগুলোর প্রতিটির কিছু কার্যাবলী রয়েছে এবং কিছু সুবিধা ও সীমাবদ্ধতাও রয়েছে। বিস্তৃতির একটি উত্তম পরিমাপের নিম্নলিখিত বৈশিষ্ট্যসমূহ থাকা প্রয়োজন।

প্রথমতঃ, বিস্তৃতির পরিমাপকে সহজবোধ্য হতে হবে।

দ্বিতীয়তঃ, এটিকে সহজে নির্ণয়যোগ্য হতে হবে।

তৃতীয়তঃ, এটিকে সুনির্দিষ্টভাবে সংজ্ঞায়িত হতে হবে।

চতুর্থতঃ, বিস্তৃতির পরিমাপকে পরবর্তী বীজগাণিতিক পরিগণনার উপযোগী হতে হবে।

পঞ্চমতঃ, এটির নমুনা স্থিতিশীলতা থাকতে হবে। এবং

ষষ্ঠতঃ, এটিকে চরম মান দ্বারা অযৌক্তিকভাবে প্রভাবিত হওয়া চলবে না।

বিস্তৃতির পরিমাপের প্রকারভেদ (Types of Measures of Dispersion)

বিস্তৃতির পরিমাপ দু'ধরণের হয়ে থাকে – নিরঙ্কুশ ও আপেক্ষিক। উপাত্ত রাশিমালা যে এককে পরিমাপকৃত হয় বিস্তৃতির নিরঙ্কুশ পরিমাপগুলো সেই এককেই প্রকাশিত হয়।

ব্যাপক অর্থে বিস্তৃতির পরিমাপ দু'ধরণের হয়ে থাকে – নিরঙ্কুশ ও আপেক্ষিক। প্রদত্ত উপাত্ত রাশিমালা যে এককে পরিমাপকৃত হয় (যেমন, টাকা, বছর, কিলোগ্রাম, ইত্যাদি) বিস্তৃতির নিরঙ্কুশ পরিমাপগুলো সেই এককেই প্রকাশিত হয়। এই পরিমাপগুলো দিয়ে দুই বা ততোধিক বিন্যাসের মধ্যে তুলনা করা সম্ভব হয় যদি চলকগুলো একই এককে পরিমাপকৃত হয় এবং প্রায় একই কেন্দ্রীয় মান থাকে। যদি চলকগুলো ভিন্ন এককে পরিমাপকৃত হয় এবং ভিন্ন কেন্দ্রীয় মান দ্বারা প্রতিনিধিত্বশীল হয় সে ক্ষেত্রে বিস্তৃতির নিরঙ্কুশ পরিমাপগুলো তুলনাযোগ্য থাকে না। এ ধরণের পরিস্থিতিতে বিস্তৃতির আপেক্ষিক পরিমাপ নির্ণয় করতে হয়।

বিস্তৃতির আপেক্ষিক পরিমাপ সবসময়ই পরিমাপের একক থেকে স্বাধীন হয়ে থাকে।

বিস্তৃতির আপেক্ষিক পরিমাপ হলো গড়ের প্রেক্ষিতে বিস্তৃতির নিরঙ্কুশ পরিমাপের অনুপাত। বিস্তৃতির আপেক্ষিক পরিমাপ সবসময়ই পরিমাপের একক থেকে স্বাধীন হয়ে থাকে। এ বিষয়ে পরবর্তী পাঠগুলোতে বিস্তারিত আলোচনা করা হবে।

বিস্তৃতির নিরঙ্কুশ পরিমাপগুলো হলো:

১. পরিসর (Range)
২. আন্তচতুর্থক পরিসর বা চতুর্থক ব্যবধান (The Interquartile Range or Quartile Deviation)
৩. গড় ব্যবধান (Mean Deviation)
৪. পরিমিত ব্যবধান (Standard Deviation)
৫. ভেদাঙ্ক (Variance)

বিস্তৃতির আপেক্ষিক পরিমাপগুলো হলো:

১. চতুর্থক ব্যবধানাঙ্ক (Coefficient of Quartile Deviation)
২. গড় ব্যবধানাঙ্ক (Coefficient of Mean Deviation)
৩. ব্যবধানাঙ্ক (Coefficient of Variation)

সারাংশ

কেন্দ্রীয় মান থেকে উপাত্তরাশির ছড়িয়ে থাকার প্রবণতাকেই বিস্তৃতি বলে। বিস্তৃতির পরিমাপ দু'ধরণের হয়ে থাকে — নিরঙ্কুশ ও আপেক্ষিক। যে এককে উপাত্ত রাশিমালাকে পরিমাপ করা হয়, সেই এককেই নিরঙ্কুশ পরিমাপগুলোকে নির্ণয় করা হয়। অন্যদিকে, যদি কোন উপাত্ত ভিন্ন এককে পরিমাপকৃত হয় এবং ভিন্ন কেন্দ্রীয় মান দ্বারা প্রতিনিধিত্বশীল হয়, সে ক্ষেত্রে বিস্তৃতির আপেক্ষিক পরিমাপ নির্ণয় করতে হয় নিরঙ্কুশ পরিমাপের একক থেকে আপেক্ষিকতায় রূপান্তর করে।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক (√) চিহ্ন দিন –

১। ব্যাপক অর্থে বিস্তৃতির পরিমাপ কত ধরণের?
   ক. দুই ধরণের
   খ. তিন ধরণের
   গ. চার ধরণের
   ঘ. পাঁচ ধরণের

২। যদি উপাত্তরাশির মানগুলো কেন্দ্রীয় মান থেকে খুব বেশী ছড়িয়ে থাকে তাহলে আমরা বলতে পারি:
   ক. বিন্যাসটির মধ্যে সমরূপতার অভাব নাই।
   খ. বিন্যাসটির মধ্যে কেন্দ্রীয় প্রবণতা নাই।
   গ. বিন্যাসটির মধ্যে সমরূপতার অভাব রয়েছে।
   ঘ. বিন্যাসটি মধ্যে কেন্দ্রীয় প্রবণতা রয়েছে।

৩। কেন্দ্রীয় মান থেকে প্রতিটি মানের ছড়িয়ে থাকায় গড় মাত্রাকে নিরূপণ করে:
   ক. গড় ব্যবধান
   খ. ব্যবধানাঙ্ক
   গ. পরিমিত ব্যবধান
   ঘ. উপরের সব ক'টি

সংক্ষিপ্ত প্রশ্ন

১। বিস্তৃতির পরিমাপ কি?
২। বিস্তৃতির প্রকৃতি বলতে কি বোঝায়?

রচনামূলক প্রশ্ন

১। বিস্তৃতির পরিমাপ কি? বিস্তৃতির একটি উত্তম পরিমাপের বৈশিষ্ট্যসমূহ আলোচনা করুন।
২। বিভিন্ন ধরণের বিস্তৃতির পরিমাপসমূহ আলোচনা করুন।

পাঠ - ২

বিস্তৃতির নিরঙ্কুশ পরিমাপ – ১: পরিসর ও চতুর্থক ব্যবধান
Absolute Measures of Dispersion – 1: Range and Quartile Deviation

এই পাঠ শেষে যা জানা যাবে —

- পরিসর কি
- বিন্যস্ত উপাত্ত থেকে পরিসর নির্ণয়
- পরিসরের সুবিধা ও অসুবিধা
- আধা-আন্তচতুর্থক পরিসর বা চতুর্থক ব্যবধান
- চতুর্থক ব্যবধানের সুবিধা ও অসুবিধা

পরিসর কি (What is Range)?

বিস্তৃতির পরিমাপের সবচেয়ে সহজ, সাধারণ এবং স্থূল পরিমাপটি হলো পরিসর। উপাত্তরাশির সর্বোচ্চ ও সর্বনিম্ন মানের মধ্যকার পার্থক্যই হলো পরিসর।

বিস্তৃতির পরিমাপের সবচেয়ে সহজ, সাধারণ এবং স্থূল পরিমাপটি হলো পরিসর। উপাত্তরাশির সর্বোচ্চ ও সর্বনিম্ন মানের মধ্যকার পার্থক্যই হলো পরিসর। পরিসর নির্ণয়ের সূত্রটি হলো,

পরিসর = সর্বোচ্চ মান – সর্বনিম্ন মান

প্রতীকের মাধ্যমে,

পরিসর = L – S

যেখানে, L = সর্বোচ্চ মান

S = সর্বনিম্ন মান

ধরা যাক, ১০ জন ব্যক্তির ওজন (মানের উচ্চক্রম অনুসারে)

১১০ ১১৫ ১২০ ১২৭ ১৩০ ১৩২ ১৩৭ ১৪০ ১৪২ ১৫২ পাউন্ড

এখানে সর্বোচ্চ মান হলো ১৫২ এবং সর্বনিম্ন মান হলো ১১০ পাউন্ড। সূত্র অনুযায়ী এই উপাত্তরাশির পরিসর হলো,

পরিসর = L – S

= ১৫২ – ১১০

= ৪২

অর্থাৎ, ১০ জন ব্যক্তির ওজনের ব্যবধান হলো ৪২ পাউন্ড। কিন্তু এভাবে পরিসর নির্ণয়ের একটি সমস্যা রয়েছে। কারণ, বর্ণিত সীমায় (stated limit) সর্বোচ্চ এবং সর্বনিম্ন মান একটি পূর্ণ সংখ্যায় প্রকাশিত হবার ফলে যথার্থ সীমাটি (true limit) এখানে বিবেচিত হয়নি। আমরা জানি যে, যদিও ব্যক্তির ওজন বৃদ্ধির ক্ষুদ্রতম পরিমাণ ১ পাউন্ড-এর এককে প্রকাশিত হয়েছে, কিন্তু যথার্থ সীমাটি হলো মূলতঃ সর্বনিম্ন সীমার ০.৫ পাউন্ড নীচে এবং সর্বোচ্চ সীমার ০.৫ পাউন্ড উপরে। অর্থাৎ, আমাদের উদাহরণের ক্ষেত্রে যদি যথার্থ সীমায় সর্বোচ্চ মান ১৫২.৫ এবং সর্বনিম্ন মান ১০৯.৫ হয়, তবে পরিসর হবে,

পরিসর = ১৫২.৫ – ১০৯.৫ = ৪৩.০

এখানে পরিসরের মান ১ বেড়ে গেলেও এটিই হলো পরিসরের প্রকৃত এবং সূক্ষ্ম মান। এ বিষয়টিকে সূত্রের মধ্যে অন্তর্ভুক্ত করা যায় এভাবে,

পরিসর = (সর্বোচ্চ মান – সর্বনিম্ন মান) + পরিমাপ এককের ক্ষুদ্রতম বৃদ্ধির পরিমাণ

আমাদের উদাহরণ অনুযায়ী, ওজন বৃদ্ধির ক্ষুদ্রতম বৃদ্ধির পরিমাণ হলো ১ পাউন্ড (যেহেতু পূর্ণ সংখ্যায় প্রকাশিত হয়েছে)। অতএব,

পরিসর = (১৫২ – ১১০) + ১
= ৪২ + ১
= ৪৩

এখানে সাধারণ জ্ঞানের একটি বিষয়ও চমৎকারভাবে বিবেচিত হয়েছে যা একটি ভ্রান্তিকে দূর করেছে আমাদের অগোচরেই। সাধারণভাবে, আমরা যখন কোন ব্যবধান নির্ণয় করি তখন সর্বনিম্ন সীমার প্রথম মানটিকে বাদ দিয়ে করি। যেমন, পরিসর নির্ণয়ের ক্ষেত্রে ১৫২ থেকে ১১০-এর ব্যবধান নির্ণয় করতে গিয়ে গণনা শুরু করেছি ১১০-কে বাদ দিয়ে ১১১ থেকে, যার ফলে প্রাপ্ত ফলাফল হয়েছে ৪২। কিন্তু ১১০-কে গণনা থেকে বাদ দেবার ফলে আমরা মূলতঃ উপাত্তরাশি থেকে আমাদের অগোচরেই ১১০ সংখ্যা মানটিকে বাদ দিয়ে ফেলেছি এবং এর ফলে প্রকৃতপক্ষে উপাত্তরাশির মোট সংখ্যা গিয়ে দাঁড়িয়েছে ৯ জন, ১০ জন নয়। কাজেই, পরিসর নির্ণয়ের ক্ষেত্রে গতানুগতিক সূত্রটি ব্যবহার না করে পরিশীলিত সূত্রটি ব্যবহার করা শ্রেয়। যদি চলকগুলো ক্রমসূচক মাত্রায় পরিমাপকৃত হয়, সে ক্ষেত্রে গতানুগতিক সূত্রটিই যথেষ্ট।

এখন একটি প্রশ্ন আসতে পারে যে, সবসময়ই কি আমরা সর্বোচ্চ ও সবনিম্ন মানের ব্যবধানের সাথে ১ যোগ করবো? উত্তরটি হলো 'না'। কারণ, পরিমাপ এককের ক্ষুদ্রতম বৃদ্ধির পরিমাণ সবসময় একই রকম থাকে না। পরিমাপ এককের ক্ষুদ্রতম বৃদ্ধির পরিমাপটি নির্ধারিত হয় একটি একক থেকে অন্য এককের দূরত্বের উপর। সারণি ৫.২.২-এ বিভিন্ন পরিমাপ এককের ক্ষুদ্রতম বৃদ্ধির পরিমাণ-এর কিছু উদাহরণ দেয়া হলো।

সারণি ৫.২.২: পরিমাপ এককের ক্ষুদ্রতম বৃদ্ধির পরিমাণ এবং পরিসর নির্ণয়ের কিছু উদাহরণ

পর্যবেক্ষণের নমুনা মান					*সর্বোচ্চ মান*	*সর্বনিম্ন মান*	*ক্ষুদ্রতম বৃদ্ধির পরিমাণ*	*পরিসর*
৫.৩	৩.৭	৮.২	৪.৫	৬.১	১০.২	২.১	০.১	৮.২
৫.৩২	৩.৭৫	৮.২৬	৪.৫৩	৬.১৮	১০.২৮	২.১৭	০.০১	৮.১২
৩০০০	৫০০০	৪০০০	২০০০	৭০০০	১০০০০	২০০০	১০০০	৯০০০
৪৫০	৩৫০	২৫০	৭৫০	৬০০	১০০০	২৫০	৫০	৮০০
২৮	২১	৪৫	৭১	৩৮	১২৫	১৮	১	১০৮
	৭০ ১/৫	৮০ ৪/৫	৯৫ ২/৫	৬২ ৩/৫	১৮৫ ৩/৫	৪৫ ১/৫	১/৫	১৪০ ৩/৫

বিন্যস্ত উপাত্ত থেকে পরিসর নির্ণয় (Computing Range from Grouped Data)

বিন্যস্ত উপাত্ত থেকে সূক্ষ্মভাবে পরিসর নির্ণয় করা যায় না। কারণ, একটি গণসংখ্যা নিবেশনে মূল উপাত্তরাশিকে উপস্থাপিত না করে সেগুলো কতগুলো শ্রেণীতে বিভক্ত করে গণসংখ্যা নিবেশনে প্রকাশ করা হয়।

বিন্যস্ত উপাত্ত থেকে সূক্ষ্মভাবে পরিসর নির্ণয় করা যায় না। কারণ, একটি গণসংখ্যা নিবেশনে মূল উপাত্তরাশিকে উপস্থাপিত না করে সেগুলো কতগুলো শ্রেণীতে বিভক্ত করে গণসংখ্যা নিবেশনে প্রকাশ করা হয়। আর এই গণসংখ্যা নিবেশন থেকে পরিসরের সূক্ষ্ম মান নির্ণয় করা না গেলেও এর সম্পর্কে একটি ধারণা পাওয়া যায়। ধরা যাক, ১০ জন ব্যক্তির মাসিক আয় নিম্নরূপ,

৪০০০ ৪৪০০ ৪৫০০ ৫০০০ ৫২০০ ৬১০০ ৬৫০০ ৭৫০০ ৭৮০০ ৮০০০ টাকা

সারণি ৫.২.৩-এ এই ১০ জন ব্যক্তির মাসিক আয়ের গণসংখ্যা বিন্যাসটি উপস্থাপিত হলো।

সারণি ৫.২.৩ঃ ১২ জন ব্যক্তির মাসিক আয়ের বিন্যাস

শ্রেণী সীমা (টাকায়)	*ঘটনসংখ্যা (ব্যক্তি)*
৪০০০ – ৫০০০	৩
৫০০০ – ৬০০০	২
৬০০০ – ৭০০০	২
৭০০০ – ৮০০০	৩
মোট	১০

বিন্যস্ত উপাত্ত থেকে পরিসর নির্ণীত হয় সর্বোচ্চ শ্রেণীর উর্ধ্ব সীমার মান থেকে সর্বনিম্ন শ্রেণীর নিম্ন সীমার মানের ব্যবধান নির্ণয়ের মাধ্যমে।

বিন্যস্ত উপাত্ত থেকে পরিসর নির্ণীত হয় সর্বোচ্চ শ্রেণীর উর্ধ্ব সীমার মান থেকে সর্বনিম্ন শ্রেণীর নিম্ন সীমার মানের ব্যবধান নির্ণয়ের মাধ্যমে। পরিসর নির্ণয়ের জন্য অবিন্যস্ত উপাত্তের ক্ষেত্রে যে সূত্রটি ব্যবহৃত হয় বিন্যস্ত উপাত্তের ক্ষেত্রেও একই সূত্র ব্যবহৃত হয়। অর্থাৎ,

পরিসর = (সর্বোচ্চ মান – সর্বনিম্ন মান) + পরিমাপ এককের ক্ষুদ্রতম বৃদ্ধির পরিমাণ

এখানে সর্বোচ্চ আয়ের পরিমাণ ৮০০০ টাকা এবং সর্বনিম্ন আয়ের পরিমাণ ৪০০০ টাকা এবং ক্ষুদ্রতম বৃদ্ধির পরিমাণ ১০০ টাকা। অতএব,

পরিসর = (৮০০০ – ৪০০০) + ১০০

= ৪০০০ + ১০০

= ৪১০০

∴ নির্ণীত পরিসর হলো ৪১০০ টাকা

মনে রাখতে হবে যে, এটি আনুমানিক মান, প্রকৃত মান নয়। যদি উপাত্তরাশির মূল মানগুলো জানা না যায় এবং সরাসরি গণসংখ্যা নিবেশন থেকে পরিসর নির্ণয় করতে হয়, তবে পরিমাপ এককের ক্ষুদ্রতম বৃদ্ধির পরিমাণ নির্ণয়ের ক্ষেত্রে একটি সমস্যা তৈরী হয়। কারণ, পরিমাপ এককের ক্ষুদ্রতম বৃদ্ধির প্রকৃত পরিমাণটি জানা যায় না। সে ক্ষেত্রে পূর্ণ সংখ্যার জন্য পরিমাপ এককের ক্ষুদ্রতম বৃদ্ধির পরিমাণ হিসাবে অন্তত ১ যোগ করলে ভ্রান্তির কোন অবকাশ থাকে না।

পরিসরের সুবিধা এবং অসুবিধা (Advantages & Disadvantages of Range)

বিস্তৃতির পরিমাপ হিসাবে পরিসরের সারল্য (simplicity) হলো এর প্রধান সুবিধা। খুব সহজে এবং দ্রুত উপাত্তের বিস্তৃতি পরিমাপের জন্য পরিসরই সবচেয়ে উপযোগী পরিমাপ। এ ক্ষেত্রে উপাত্তরাশির সকল মান জানবার প্রয়োজন হয় না। শুধুমাত্র সর্বোচ্চ ও সর্বনিম্ন মান জানাই যথেষ্ট। উপরন্তু, পরিসংখ্যান সম্পর্কে যারা তেমন জ্ঞান রাখেন না, তাদের জন্য পরিসরই বিস্তৃতির সবচেয়ে উপযোগী পরিমাপ। যখন উপাত্তসমূহ ক্রমসূচক মাত্রায় উপস্থাপিত হয় তখন পরিসর সবচেয়ে সুবিধাজনক পরিমাপ। সংখ্যা মানগুলোকে মানের ক্রম অনুসারে সাজিয়ে সর্বোচ্চ ও সর্বনিম্ন মানের পার্থক্য নির্ণয় করে আমরা খুব সহজেই পরিসরের মান পেতে পারি। এখানে উল্লেখ করা প্রয়োজন যে, নামসূচক মাত্রায় পরিমাপকৃত উপাত্ত যেহেতু মানের ক্রম অনুসারে বিন্যস্ত করা যায় না, সেহেতু সেখানে পরিসর নির্ণয় করা সম্ভব হয় না।

বিস্তৃতির পরিমাপ হিসাবে পরিসরের প্রত্যয়গত সারল্য যেমন এর সুবিধা তেমনি এটি এর অসুবিধাও। পরিসর কোন পরিশীলিত পরিমাপ নয়। এটি মাত্র দু'টি মানের ভিত্তিতে নির্ণীত হয়, যা মূলতঃ উপাত্তরাশির চরম দু'টি মান। যেহেতু চরম মান সাধারণতঃ দূর্লভ এবং অস্বাভাবিক হয়, সেহেতু চরম

মান দ্বারা নির্ণীত কোন পরিমাপ আদর্শ হতে পারে না। অর্থাৎ, পরিসর চরম মান বা অসাধারণ সংখ্যা মান দ্বারা খুব বেশী প্রভাবিত হয়। ধরা যাক, ১০ জন ব্যক্তির মাসিক আয় নিম্নরূপ,

৫০০০ ৫৭০০ ৬০০০ ৬৩০০ ৮০০০ ৭৮০০ ৬৭০০ ৮৫০০ ৮৭০০ ২০০০০টাকা

অতএব, পরিসর = (২০০০০ – ৫০০০) + ১০০

= ১৫০০০ + ১০০

= ১৫১০০

এ ক্ষেত্রে দেখা যায় যে, একটি মাত্র চরম মানের কারণে পরিসরের মান সর্বোচ মানটি ছাড়া প্রতিটি মানের চেয়ে অস্বাভাবিকভাবে বড় হয়েছে, যা পুরো বিন্যাসটি সম্পর্কে একটি ভ্রান্তিমূলক ধারণা দেয়। শ্রেণী বিন্যস্ত উপাত্তের ক্ষেত্রে উন্মুক্ত শ্রেণী ব্যাপ্তি হলে পরিসর নির্ণয় করা সম্ভব নয়। পরিসর নমুনা বিচ্যুতি দ্বারা বেশী প্রভাবিত হয়। যেহেতু পরিসর নির্ণয়ে শুধুমাত্র সর্বনিম্ন ও সর্বোচ্চ মানের প্রয়োজন হয়, সেহেতু মধ্যবর্তী মানগুলো সম্পর্কে পরিসর কোন ধারণা দিতে পারে না। সুতরাং, বিন্যাসের সকল মানের জন্য বিস্তৃতির পরিমাপ হিসাবে পরিসর আদৌ নির্ভরশীল নয়।

আধা-আন্তচতুর্থক পরিসর বা চতুর্থক ব্যবধান (Semi-Interquartile Range or Quartile Deviation)

আমরা জানি যে, পরিসর মাত্র দু'টি মানের উপর ভিত্তি করে নির্ণীত এবং উপাত্তরাশির মানগুলো সম্পর্কে কোন ধারণা দিতে সক্ষম হয় না। পরিসরের এই সীমাবদ্ধতা দূর করার উদ্দেশ্যে চতুর্থক ব্যবধান বা আধা-আন্তচতুর্থক পরিসর নির্ণয় করা হয়। এটিও এক প্রকারের পরিসর, তবে উপাত্তরাশির উচ্চতম ও নিম্নতম মানের পার্থক্যের নয়। এখানে উপাত্তরাশির প্রথম ও তৃতীয় চতুর্থকের মধ্যকার পার্থক্য বিবেচনা করা হয়। উপাত্তরাশিকে মানের উচ্চ ক্রম অনুসারে সাজিয়ে মোট মানগুলোকে সমান চারভাগে ভাগ করে নিতে হয়। এর একেকটি ভাগকে চতুর্থক বলে। আন্তচতুর্থক পরিসর হলো তৃতীয় চতুর্থক মান থেকে প্রথম চতুর্থক মানের ব্যবধান। প্রতীকের মাধ্যমে সূত্রটি হলো,

পরিসর মাত্র দু'টি মানের উপর ভিত্তি করে নির্ণীত এবং উপাত্তরাশির মানগুলো সম্পর্কে কোন ধারণা দিতে সক্ষম হয় না। পরিসরের এই সীমাবদ্ধতা দূর করার উদ্দেশ্যে চতুর্থক ব্যবধান বা আধা-আন্তচতুর্থক পরিসর নির্ণয় করা হয়।

$$\text{আন্তচতুর্থক পরিসর} = Q_৩ - Q_১$$

আন্তচতুর্থক পরিসরকে ২ দিয়ে ভাগ করে আধা-আন্তচতুর্থক পরিসর বা চতুর্থক ব্যবধানে পরিণত করা হয়। অর্থাৎ, চতুর্থক ব্যবধান হলো প্রথম ও তৃতীয় চতুর্থকের মধ্যকার ব্যবধানের অর্ধেক। প্রতীকের মাধ্যমে সূত্রটি হলো,

$$\text{চতুর্থক ব্যবধান} = \frac{Q_৩ - Q_১}{২}$$

একটি উদাহরণ দেয়া যাক। ধরা যাক, ১২ জন ব্যক্তির মাসিক আয় (মানের ক্রম অনুসারে) নিম্নরূপ,

৩৯০০ ৪০০০ ৪২০০ ৫৩০০ ৫৭০০ ৬০০০
৬২০০ ৬২০০ ৬৭০০ ৭২০০ ৮০০০ ৮২০০ টাকা

এই রাশিমালার তৃতীয় চতুর্থক ($Q_৩$) হবে ৬৭০০ এবং প্রথম চতুর্থক ($Q_১$) হবে ৪২০০। অতএব,

$$\text{চতুর্থক ব্যবধান} = \frac{Q_৩ - Q_১}{২}$$

$$= \frac{৬৭০০ - ৪২০০}{২}$$

$$= \frac{২৫০০}{২}$$

$$= ১২৫০$$

∴ নির্ণীত চতুর্থক ব্যবধান হলো ১২৫০ টাকা

চতুর্থক ব্যবধানের সুবিধা ও অসুবিধা (Advantages and Disadvantages of Quartile Deviation)

মধ্যমা যেমন কেন্দ্রীয় প্রবণতার একটি অবস্থানমূলক পরিমাপ (positional measure), চতুর্থক ব্যবধান তেমনি বিস্তৃতির পরিমাপের একটি অবস্থানমূলক পরিমাপ। চতুর্থক ব্যবধান চরম মান দ্বারা কম প্রভাবিত হয় বলে এটি তুলনামূলকভাবে অধিকতর স্থিতিশীল পরিমাপ। এটি ক্রমসূচক মাত্রায় পরিমাপকৃত উপাত্তের ক্ষেত্রে বিশেষভাবে প্রযোজ্য। চতুর্থক ব্যবধান মারাত্মকভাবে বঙ্কিম বিন্যাসের ক্ষেত্রে অধিকতর উপযোগী পরিমাপ। উন্মুক্ত শ্রেণী বিন্যাসের ক্ষেত্রে বিস্তৃতি পরিমাপের জন্য এটির বিশেষ উপযোগীতা রয়েছে। চতুর্থক ব্যবধান অনেক ক্ষেত্রে পরিসর অপেক্ষা বেশী উন্নততর একটি পরিমাপ বলে বিবেচিত হয়।

বিস্তৃতির পরিমাপ হিসাবে চতুর্থক ব্যবধানের কিছু অসুবিধাও রয়েছে। এটি একটি বিন্যাসের ৫০ ভাগ উপাত্তরাশিকে অবজ্ঞা করে। অর্থাৎ, প্রথম ২৫ শতাংশ এবং শেষ ২৫ শতাংশ উপাত্তরাশিকে কখনই বিবেচনায় আনে না। চতুর্থক ব্যবধান যেহেতু বিন্যাসের প্রতিটি সংখ্যা মান দ্বারা নির্ণীত হয় না, সেহেতু এটিকে বিস্তৃতির একটি উত্তম পরিমাপ হিসাবে বিবেচনা করা যায় না। চতুর্থক ব্যবধান বীজগাণিতিক পরিগণনার উপযোগী নয়। নমুনা বিচ্যুতি দ্বারা এটি খুব বেশী প্রভাবিত হয়। অবস্থানমূলক পরিমাপ হবার কারণে অনেক পরিসংখ্যানবিদ চতুর্থক ব্যবধানকে বিস্তৃতির পরিমাপ হিসাবে স্বীকার করতে চান না। এটিকে বিস্তৃতির পরিমাপ না বলে অবস্থানের পরিমাপ বলে অভিহিত করে থাকেন। কারণ, এটি গড় থেকে প্রতিটি মানের ছড়িয়ে থাকার বিষয়টি প্রদর্শন না করে একটি মান থেকে আরেকটি মানের দূরত্বকে নির্দেশ করে।

সারাংশ

বিস্তৃতির পরিমাপের সবচেয়ে সহজ পরিমাপটি হলো পরিসর। পরিসর জানার জন্য উপাত্তরাশির সকল মান জানার প্রয়োজন হয় না। শুধুমাত্র সর্বোচ্চ ও সর্বনিম্ন মান জানাই যথেষ্ট। যেহেতু পরিসর মাত্র দু'টি মানের উপর ভিত্তি করে নির্ণীত, সেহেতু উপাত্তরাশির অন্যান্য মানগুলো সম্পর্কে কোন ধারণা দিতে সক্ষম হয় না। এই সীমাবদ্ধতা দূর করার জন্য চতুর্থক ব্যবধান বা আধা-আন্তচতুর্থক পরিসর নির্ণয় করা হয়।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক (√) চিহ্ন দিন –

১। পরিসর নির্ণয়ের সূত্র হলো:

ক. সর্বোচ্চ মান – মধ্যম মান
খ. সর্বোচ্চ মান – সর্বনিম্ন মান
গ. সর্বনিম্ন মান + সর্বোচ্চ মান
ঘ. মধ্যম মান × নিম্ন মান

২। এককের ক্ষুদ্রতম বৃদ্ধির পরিমাপটি নির্ধারিত হয়:

ক. একটি এককের সাথে অন্য এককের যোগফলের উপর
খ. একটি এককের সাথে অন্য এককের গুণফলের উপর
গ. একটি এককের সাথে অন্য এককের পার্থক্যের বর্গফলের উপর
ঘ. একটি একক থেকে অন্য এককের দূরত্বের উপর

৩। পরিসংখ্যানের সীমাবদ্ধতা দূর করার জন্য:

ক. উপাত্তরাশির গড় নির্ণয় করতে হয়
খ. উপাত্তরাশির প্রথম ও তৃতীয় চতুর্থকের যোগফল নির্ণয় করতে হয়
গ. উপাত্তরাশির প্রথম ও তৃতীয় চতুর্থকের মধ্যকার পার্থক্য বিবেচনা করা হয়
ঘ. উপরের সব ক'টি

সংক্ষিপ্ত প্রশ্ন

১। পরিসরের সংজ্ঞা দিন ।

২। উদাহরণসহ আন্ত-চতুর্থক নির্ণয় পদ্ধতি আলোচনা করুন।

রচনামূলক প্রশ্ন

১। পরিসরের সুবিধা এবং অসুবিধাগুলো আলোচনা করুন।

২। চতুর্থক ব্যবধানের সুবিধা ও অসুবিধাগুলো আলোচনা করুন।

পাঠ - ৩

বিস্তৃতির নিরঙ্কুশ পরিমাপ – ২: গড় ব্যবধান ও শতহারিক পরিসর
Absolute Measures of Dispersion – 2: Mean Deviation and Percentile Range

এই পাঠ শেষে যা জানা যাবে —

- গড় ব্যবধান কি
- অবিন্যস্ত উপাত্ত থেকে গড় ব্যবধান নির্ণয়
- বিন্যস্ত উপাত্ত থেকে গড় ব্যবধান নির্ণয়
- গড় ব্যবধানের সুবিধা ও অসুবিধা
- শতহারিক পরিসর

গড় ব্যবধান কি (What is Mean Deviation)?

পূর্বের পাঠে বিস্তৃতির যে দু'টি নিরঙ্কুশ পরিমাপ (পরিসর ও চতুর্থক ব্যবধান) সম্পর্কে আলোচনা করা হয়েছে তার কোনটিই উপাত্তরাশির সবগুলো মানকে বিবেচনায় এনে বিস্তৃতির পরিমাপ করে না। অথচ আমরা জানি যে, বিস্তৃতির একটি উত্তম পরিমাপ হিসাবে বিবেচিত হতে হলে উপাত্তরাশির সবগুলো সংখ্যা মানকে বিবেচনায় আনা প্রয়োজন। যদি আমরা সব মানকে ব্যবহার করতে চাই, তাহলে সাধারণ জ্ঞান আমাদেরকে এটিই বলবে যে, দু'টি প্রান্তিক বিন্দুর মধ্যকার ব্যবধানের দিকে দৃষ্টি নিবদ্ধ না করে আমরা একটি পরিমাপের মাত্রার (measurement scale) কোন একটি নির্দিষ্ট বিন্দু থেকে প্রতিটি মান কতটুকু ব্যবধানে ছড়িয়ে রয়েছে তার ভিত্তিতে বিস্তৃতির বিষয়টি বোঝার চেষ্টা করতে পারি। উপাত্তরাশির মানগুলো যত বেশী ছড়িয়ে ছিটিয়ে থাকবে সেই নির্দিষ্ট বিন্দু থেকে মানগুলোর ব্যবধান তত বড়বে। কিন্তু সেই নির্দিষ্ট বিন্দুটি কি হবে এবং মানগুলো কিভাবে ছড়িয়ে রয়েছে তা আমরা কিভাবে নির্ধারণ করবো? সমাধান হিসাবে বলা যায় যে, কেন্দ্রীয় প্রবণতার যে কোন একটি পরিমাপ থেকে উপাত্তরাশির প্রতিটি মানের অনপেক্ষ বিচ্যুতির (absolute difference) সমষ্টি নির্ণয় করে তার একটি গড় নির্ণয় করতে পারি।

কেন্দ্রীয় প্রবণতার পরিমাপ হিসাবে প্রচুরক ও মধ্যমা ব্যবহার করা গেলেও আমরা সাধারণতঃ গাণিতিক গড়কেই অনপেক্ষ বিচ্যুতি নির্ণয়ের জন্য ব্যবহার করে থাকি। কারণ, অধিকাংশ ক্ষেত্রে গাণিতিক গড়ই হলো কেন্দ্রীয় প্রবণতার সবচেয়ে উত্তম পরিমাপ।

কেন্দ্রীয় প্রবণতার পরিমাপ হিসাবে প্রচুরক ও মধ্যমা ব্যবহার করা গেলেও আমরা সাধারণতঃ গাণিতিক গড়কেই অনপেক্ষ বিচ্যুতি নির্ণয়ের জন্য ব্যবহার করে থাকি। কারণ, *প্রথমতঃ*, অধিকাংশ ক্ষেত্রে গাণিতিক গড়ই হলো কেন্দ্রীয় প্রবণতার সবচেয়ে উত্তম পরিমাপ। *দ্বিতীয়তঃ*, গড় থেকে প্রতিটি মানের ব্যবধানের যোগ ফল '০' হবার কারণে গাণিতিক গড়কে উপাত্তরাশির মাধ্যাকর্ষণের কেন্দ্র (center of gravity) বলা হয়। এবং *তৃতীয়তঃ*, গাণিতিক গড় থেকে প্রতিটি মানের বর্গকৃত ব্যবধানের যোগফল সবসময়ই অন্য যে কোন মান থেকে প্রতিটি মানের বর্গকৃত ব্যবধানের যোগফলের তুলনায় ক্ষুদ্রতর হয়। বিস্তৃতির পরিমাপ নির্ণয়ে এই বৈশিষ্ট্যগুলো গুরুত্বপূর্ণ।

এখানে প্রশ্ন আসতে পারে যে, অনপেক্ষ বিচ্যুতি নেয়া হয় কেন? এর উত্তরে বলা যায় যে, যেহেতু গড় থেকে প্রতিটি মানের ধনাত্মক এবং ঋণাত্মক বিচ্যুতিগুলোর যোগফল '০' হয়, সেহেতু '০' যোগফলকে এড়াতে হলে ঋণাত্মক চিহ্নকে কোন না কোনভাবে উপেক্ষা করতে হবে। ঋণাত্মক চিহ্ন উপেক্ষা করার একটি পদ্ধতি হলো অনপেক্ষ বিচ্যুতি নির্ণয় করা। অতএব, গড় ব্যবধান হলো গাণিতিক গড় থেকে প্রতিটি মানের অনপেক্ষ বিচ্যুতির গড়।

অবিন্যস্ত উপাত্ত থেকে গড় ব্যবধান নির্ণয় (Computing Mean Deviation from Ungrouped Data)

সংজ্ঞা অনুযায়ী গড় ব্যবধান নির্ণয়ের সূত্রটিকে প্রতীকের মাধ্যমে নিম্নলিখিত উপায়ে প্রকাশ করা যায়,

$$\text{M.D.} = \frac{\sum\left|\left(x_i - \bar{x}\right)\right|}{N}$$

যেখানে, M.D. = গড় ব্যবধান

x_i = উপাত্তরাশির প্রতিটি মান

$\bar{x}$ = গাণিতিক গড়

N = উপাত্তরাশির মোট সংখ্যা

$|\ \ |$ = অনপেক্ষ দন্ড

$\left|\left(x_i - \bar{x}\right)\right|$ = গাণিতিক গড় থেকে প্রতিটি মানের অনপেক্ষ বিচ্যুতি

$\sum$ = সমষ্টি চিহ্ন

একটি উদাহরণের মাধ্যমে সূত্রটি প্রয়োগ করে অবিন্যস্ত উপাত্ত থেকে গড় ব্যবধান নির্ণয় করে দেখা যাক। ধরা যাক, ১০ জন ব্যক্তির বয়স

১৮ ২১ ২৩ ২৭ ৩০ ৩৬ ৪০ ৪৩ ৪৯ ৫০ বছর

আমরা জানি যে, গড় ব্যবধান নির্ণয় করতে হলে প্রথমে উপাত্তরাশির গাণিতিক গড় নির্ণয় করে নিতে হয়। অতএব, ১০ জন ব্যক্তির বয়সের গড় হলো,

$$\bar{x} = \frac{\sum x_i}{N}$$

$$= \frac{১৮ + ২১ + ২৩ + ২৭ + ৩০ + ৩৬ + ৪০ + ৪৩ + ৪৯ + ৫০}{১০}$$

$$= \frac{৩৩৭}{১০}$$

$$= ৩৩.৭$$

$$\therefore \quad \bar{x} = ৩৩.৭$$

এখন নির্ণীত গাণিতিক গড় থেকে প্রতিটি মানের অনপেক্ষ বিচ্যুতি নির্ণয় করে সেই অনপেক্ষ বিচ্যুতিগুলোর সমষ্টি নির্ণয় করতে হবে, যা সারণি ৫.৩.১-এ দেখানো হলো।

সারণি ৫.৩.১: ১০ জন ব্যক্তির বয়সের গড় ব্যবধান নির্ণয়

ক্রমিক নং	x_i	$\lvert (x_i - \bar{x}) \rvert$
১	১৮	১৫.৭
২	২১	১.৭
৩	২৩	১০.৭
৪	২৭	৬.৭
৫	৩০	৩.৭
৬	৩৬	২.৩
৭	৪০	৬.৩
৮	৪৩	৯.৩
৯	৪৯	১৫.৩
১০	৫০	১৬.৩
মোট	**৩৩৭**	***৯৯.০***

এখন অনপেক্ষ বিচ্যুতির সমষ্টি মানটিকে সূত্রে প্রয়োগ করে আমরা গড় ব্যবধান নির্ণয় করতে পারি।

$$\text{M.D.} = \frac{\sum \lvert (x_i - \bar{x}) \rvert}{N}$$

$$= \frac{৯৯}{১০}$$

$$= ৯.৯$$

$\therefore$ নির্ণীত গড় ব্যবধান হলো ৯.৯ বছর

অতএব, আমরা বলতে পারি যে, ১০ জন ব্যক্তির বয়সের প্রতিটি মান গড় থেকে অনপেক্ষভাবে গড়ে ৯.৯ বছর ছড়িয়ে রয়েছে।

বিন্যস্ত উপাত্ত থেকে গড় ব্যবধান নির্ণয় (Computing Mean Deviation from Grouped Data)

শ্রেণী বিন্যস্ত উপাত্তের ক্ষেত্রে উপাত্তের মূল রাশিমালা আমাদের হাতে থাকে না বলে অনপেক্ষ বিচ্যুতি নির্ণয়ের জন্য প্রতিনিধিত্বমূলক মান পাবার ক্ষেত্রে সমস্যার সৃষ্টি হয়। একটি প্রতিনিধিত্বশীল মান পাবার জন্য প্রতিটি শ্রেণী ব্যাপ্তির অন্তর্ভুক্ত মানগুলো সংশ্লিষ্ট শ্রেণী ব্যাপ্তির মধ্যে সমানভাবে বিন্যস্ত ধরে নিয়ে প্রতিটি শ্রেণী ব্যাপ্তির মধ্য-বিন্দু নির্ণয় করতে হয়। অতএব, প্রথমে প্রতিটি শ্রেণী ব্যাপ্তির মধ্য-বিন্দু নির্ণয়ের পর সেই গণসংখ্যা নিবেশনের গড় নির্ণয় করতে হয়। তারপর, নির্ণীত গড় থেকে প্রতিটি শ্রেণীর মধ্য-বিন্দুগুলোর অনপেক্ষ বিচ্যুতি নির্ণয় করে প্রতিটি শ্রেণীর অনপেক্ষ বিচ্যুতিকে সেই শ্রেণীর ঘটনসংখ্যা দিয়ে গুণ করে অনপেক্ষ বিচ্যুতির সমষ্টি নির্ণয় করতে হয়। সবশেষে, প্রাপ্ত মানটিকে মোট ঘটনসংখ্যা দিয়ে ভাগ করতে হয়। শ্রেণী বিন্যস্ত উপাত্তের ক্ষেত্রে গড় ব্যবধান নির্ণয়ের সূত্রটি হলো,

$$\text{M.D.} = \frac{\sum \lvert f_i (x_i - \bar{x}) \rvert}{N}$$

এখানে, f_i = প্রতি শ্রেণীর গণসংখ্যা

x_i = শ্রেণী মধ্য-বিন্দু

$\bar{x}$ = গাণিতিক গড়

N = মোট ঘটনসংখ্যা

$|(x_i - \bar{x})|$ = শ্রেণী মধ্য-বিন্দু থেকে গাণিতিক গড়ের অনপেক্ষ বিচ্যুতি

গড় ব্যবধান নির্ণয়ের একটি উদাহরণ সারণি ৫.৩.২-এ উপস্থাপন করা হল।

সারণি ৫.৩.২: ২০ জন ব্যক্তির মাসিক আয়ের উপাত্ত ব্যবহার করে গড় ব্যবধান নির্ণয়

মাসিক আয় (টাকা)	লোক সংখ্যা (f_i)	মধ্য-বিন্দু (x_i)	$f_i x_i$	$\sum\|(x_i - \bar{x})\|$	$\sum\|f_i (x_i - \bar{x})\|$
৫০০০ - ৬০০০	১	৫৫০০	৫৫০০	২৮০০	২৮০০
৬০০০ - ৭০০০	৩	৬৫০০	১৯৫০০	১৮০০	৫৪০০
৭০০০ - ৮০০০	৬	৭৫০০	৪৫০০০	৮০০	৪৮০০
৮০০০ - ৯০০০	৩	৮৫০০	২৫৫০০	২০০	৬০০
৯০০০ - ১০০০০	৪	৯৫০০	৩৮০০০	১২০০	৪৮০০
১০০০০ - ১১০০০	২	১০৫০০	২১০০০	২২০০	৪৪০০
১১০০০ - ১২০০০	১	১১৫০০	১১৫০০	৩২০০	৩২০০
মোট	২০		১৬৬০০০		২৬০০০

প্রদত্ত উপাত্তের গাণিতিক গড় হলো,

$$\bar{x} = \frac{\sum f_i x_i}{N}$$

$$= \frac{১৬৬০০০}{২০}$$

$$= ৮৩০০$$

$\therefore$ $\bar{x}$ = ৮৩০০ টাকা

এবং $\sum f_i |x_i - \bar{x}|$ = ২৬০০০

এখন প্রাপ্ত মানটিকে সূত্রে প্রয়োগ করে আমরা গড় ব্যবধান নির্ণয় করতে পারি। অতএব,

$$\text{M.D.} = \frac{\sum f_i |x_i - \bar{x}|}{N}$$

$$= \frac{২৬০০০}{২০}$$

$$= ১৩০০$$

$\therefore$ নির্ণীত গড় ব্যবধান হলো ১৩০০ টাকা

গড় ব্যবধানের সুবিধা ও অসুবিধা

গড় ব্যবধানের প্রধান সুবিধাটি হলো এটি উপাত্তরাশির সবগুলো মানকে বিবেচনায় আনে।

গড় ব্যবধানের প্রধান সুবিধাটি হলো এটি উপাত্তরাশির সবগুলো মানকে বিবেচনায় আনে। তা ছাড়া এটি যথেষ্ট সহজবোধ্য একটি পরিমাপ। গড় সম্পর্কে ধারণা আছে এমন ব্যক্তির পক্ষে অতি সহজেই গড় ব্যবধানকে বোঝা সম্ভব। তবে গড় থেকে প্রতিটি সংখ্যার বিচ্যুতি নির্ণয়ে আমরা যে অনপেক্ষতা ব্যবহার করি তা গড় ব্যবধানকে পরবর্তী বীজগাণিতিক পরিগণনার জন্য অনুপযোগী করে তোলে। উন্মুক্ত শ্রেণী ব্যাপ্তি সম্বলিত উপাত্তের ক্ষেত্রে গড় ব্যবধান নির্ণয় করা যায় না। যেহেতু উপাত্তরাশির সবগুলো মানই এখানে ব্যবহৃত হয়, সেহেতু কোন একটি মান জানা না থাকলে গড় ব্যবধান নির্ণয় করা সম্ভব নয়। গড় ব্যবধানকে তত্ত্বগতভাবে ব্যাখ্যা করা যায় না বলে সামাজিক গবেষণায় এটির ব্যবহার খুব একটা লক্ষ্য করা যায় না।

শতহারিক পরিসর (Percentile Range)

বিস্তৃতি পরিমাপের অপর একটি অবস্থানভিত্তিক পরিমাপ হলো শতহারিক পরিসর (pecrentile range)। কখনও কখনও একটি বিন্যাসের কি পরিমাণ উপাত্তরাশি একটি নির্দিষ্ট পরিসরের মধ্যে অবস্থান করে তা আমরা জানতে আগ্রহী হই। এ ধরণের একটি পরিমাপ হলো শতহারিক পরিসর, যা সাধারণতঃ শিক্ষামূলক পরিমাপের ক্ষেত্রে ব্যবহৃত হয়।

এই পরিমাপটি বিন্যাসের সর্বনিম্ন ১০ শতাংশ এবং সর্বোচ্চ ১০ শতাংশ উপাত্তকে বাদ দিয়ে মধ্যবর্তী ৮০ শতাংশ উপাত্তরাশি নিয়ে কাজ করে। শতহারিক-এর সাথে চতুর্থক, মধ্যমা এবং দশহারিক (decile)-এর একটি সম্পর্ক রয়েছে। যেমন,

$D_১$ = প্রথম দশহারিক = ১০ম শতহারিক

$D_২$ = দ্বিতীয় দশহারিক = ২০তম শতহারিক

$Q_১$ = প্রথম চতুর্থক = ২৫তম শতহারিক

$D_৩$ = তৃতীয় দশহারিক = ৩০তম শতহারিক

$D_৪$ = চতুর্থ দশহারিক = ৪০তম শতহারিক

$Q_২$ = $D_৫$ = দ্বিতীয় চতুর্থক= পঞ্চম দশহারিক = মধ্যমা = ৫০তম শতহারিক

$D_৬$ = ষষ্ঠ দশহারিক = ৬০তম শতহারিক

$D_৭$ = সপ্তম দশহারিক = ৭০তম শতহারিক

$Q_৩$ = তৃতীয় চতুর্থক = ৭৫তম শতহারিক

$D_৮$ = অষ্টম দশহারিক = ৮০তম শতহারিক

$D_৯$ = নবম দশহারিক = ৯০তম শতহারিক

এখানে লক্ষ্যণীয় যে, ১০ম শতহারিক মানটি হলো প্রথম দশহারিক এবং ৯০তম শতহারিক হলো নবম দশহারিক। সে কারণে, এই পরিমাপটিকে ১০-৯০ শতহারিক পরিসর বলে। এটি সাধারণ পরিসরের মত চরম মান দ্বারা প্রভাবিত হয় না। কারণ, চরম মানগুলো প্রথম ও শেষ ১০ শতাংশের মধ্যে থাকে এবং শতহারিক পরিসর সেগুলো বিবেচনায় আনে না। যা হোক, সেই সুবিধাটি এই পরিমাপটির একটি মারাত্মক সীমাবদ্ধতা তৈরী করে। যেহেতু এটি বিন্যাসের সব মানকে ব্যবহার করে না, সেহেতু ১০ম শতহারিকের নীচে বা ৯০তম শতহারিকের উপরের সংখ্যা মানগুলো ঘনত্বের সাথে রয়েছে নাকি ছড়িয়ে ছিটিয়ে রয়েছে তা আমরা জানতে পারি না। ফলে, বিস্তৃতির উপর উপাত্তরাশির পুরো প্রভাবটিকে জানা যায় না।

সারাংশ

একটি উত্তম পরিমাপ হিসাবে বিবেচিত হতে হলে উপাত্তরাশির সবগুলো সংখ্যা মানকে বিবেচনায় আনা প্রয়োজন। উপাত্তরাশির মানগুলো যত বেশী ছড়িয়ে ছিটিয়ে থাকবে, নির্দিষ্ট বিন্দু থেকে মানগুলোর ব্যবধান তত বাড়বে। মানগুলো কিভাবে ছড়িয়ে রয়েছে তা নির্ধারণের জন্য কেন্দ্রীয় প্রবণতার যে কোন একটি পরিমাপ থেকে উপাত্তরাশির প্রতিটি মানের অনপেক্ষ বিচ্যুতির সমষ্টি নির্ণয় করে তার একটি গড় নির্ণয় করে দেখতে হয়। এ ক্ষেত্রে সাধারণতঃ গাণিতিক গড়কেই সর্বাধিক ব্যবহার করা হয়ে থাকে। বিস্তৃতির এই পরিমাপটিকে গড় ব্যবধান বলে। বিস্তৃতি পরিমাপের অপর একটি পরিমাপ হলো শতহারিক পরিসর। শতহারিক পরিসর যেহেতু সব মানকে ব্যবহার করে না, সেহেতু ১০ম শতহারিকের নীচে এবং ৯০তম শতহারিকের উপরের সংখ্য মানগুলো ঘনত্বের সাথে রয়েছে, নাকি ছড়িয়ে ছিটিয়ে রয়েছে তা আমরা জানতে পারি না।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক (√) চিহ্ন দিন –

১। গাণিতিক গড় থেকে প্রতিটি মানের অনপেক্ষ বিচ্যুতির গড়কে:

ক. ভেদাঙ্ক বলে
খ. পরিমিত ব্যবধান বলে
গ. গড় ব্যবধান বলে
ঘ. পরিসর বলে

২। বিন্যস্ত উপাত্ত থেকে গড় ব্যবধান নির্ণয়ের সূত্রটি হলো:

ক. $M.D. = \frac{\sum f_i \left| x_i - \bar{x} \right|}{N}$

খ. $M.D. = \frac{\sum \left(x_i - \bar{x}\right)^{২}}{N}$

গ. $M.D. = \frac{\sum \left(x_i - \bar{x}\right)^{২}}{N\text{-}১}$

ঘ. $M.D. = \frac{\sum \left(x_i^{২} - \bar{x}\right)}{N}$

৩। বিস্তৃতি পরিমাপের অবস্থানভিত্তিক পরিমাপ হলো:

ক. শতহারিক পরিসর
খ. প্রথম চতুর্থক
গ. তৃতীয় চতুর্থক
ঘ. উপরের সব ক'টি

সংক্ষিপ্ত প্রশ্ন

১। সংজ্ঞা দিন ও বৈশিষ্ট্য আলোচনা করুন:

ক. শতহারিক পরিসর
খ. বিন্যস্ত উপাত্ত থেকে গড় ব্যবধান নির্ণয়

রচনামূলক প্রশ্ন

১। গড় ব্যবধান কি? সমাজ গবেষণায় এর ব্যবহার আলোচনা করুন।

২। গড় ব্যবধানের সুবিধা ও অসুবিধা আলোচনা করুন।

পাঠ - ৪

বিস্তৃতির নিরঙ্কুশ পরিমাপ – ৩: পরিমিত ব্যবধান
Absolute Measures of Dispersion – 3: Standard Deviation

এই পাঠ শেষে যা জানা যাবে —

- পরিমিত ব্যবধান কি
- পরিমিত ব্যবধানের সূত্র
- অবিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে পরিমিত ব্যবধান নির্ণয়
- অবিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে পরিমিত ব্যবধান নির্ণয়
- বিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে পরিমিত ব্যবধান নির্ণয়
- বিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে পরিমিত ব্যবধান নির্ণয়
- পরিমিত ব্যবধানের সুবিধা ও অসুবিধা
- পরিমিত ব্যবধানের ব্যবহার

পরিমিত ব্যবধান কি (What is Standard Deviation)?

আমরা জেনেছি যে, গড় ব্যবধান উপাত্তরাশির সব মানকে বিবেচনায় এনে বিস্তৃতির বর্ণনা করলেও এটি নির্ণয়ে ঋণাত্মক চিহ্ন উপেক্ষা করার জন্য অনপেক্ষ বিচ্যুতি ব্যবহৃত হয় বলে এটিকে তত্ত্বগতভাবে ব্যাখ্যা করা যায় না এবং পরবর্তী বীজগাণিতিক পরিগণনায় ব্যবহার করা যায় না। আমরা এও জেনেছি যে, গড় মান থেকে প্রতিটি মানের বিচ্যুতির '০' যোগফলকে উপেক্ষা করার জন্যই গড় ব্যবধানে অনপেক্ষ বিচ্যুতি নির্ণয় করা হয়। কিন্তু ঋণাত্মক চিহ্নকে উপেক্ষা করার আরেকটি পদ্ধতি রয়েছে যার মাধ্যমে নির্ণীত মানটি বিস্তৃতির পরিমাপের সবচেয়ে গুরুত্বপূর্ণ তত্ত্বগত ব্যাখ্যা প্রদান করে। অনপেক্ষ বিচ্যুতি নির্ণয় না করে বিচ্যুতিগুলোর বর্গ করে নিলেই ঋণাত্মক চিহ্নগুলো সহজেই অবজ্ঞা করা যায় এবং বর্গকৃত বিচ্যুতির যোগফলের গড় নির্ণয় করে তার বর্গমূল নিলে তত্ত্বগতভাবে ব্যাখ্যাযোগ্য একটি পরিমাপ নির্ণয় করা যায়। সেই পরিমাপটি হলো পরিমিত ব্যবধান।

পরিমিত ব্যবধান হলো গাণিতিক গড় থেকে প্রতিটি মানের বর্গকৃত বিচ্যুতির গড়ের বর্গমূল।

সুনির্দিষ্টভাবে বলা যায় যে, পরিমিত ব্যবধান হলো গাণিতিক গড় থেকে প্রতিটি মানের বর্গকৃত বিচ্যুতির গড়ের বর্গমূল। এটি বিভিন্ন প্রতীকের মাধ্যমে উপস্থাপন করা হয়। এই প্রতীকগুলো খুবই গুরুত্বপূর্ণ হয়ে উঠে যখন আমরা বর্ণনামূলক পরিসংখ্যানকে অতিক্রম করে সিদ্ধান্তমূলক পরিসংখ্যানে অন্তর্ভুক্ত নমুনা ও সমগ্রক নিয়ে আলোচনা শুরু করি। বর্ণনামূলক পরিস্থিতিতে যখন শুধুমাত্র একটি বিন্যাসের বিস্তৃতির বর্ণনা করতে হয়, তখন ইংরেজী বর্ণমালার ছোট অক্ষরের 's'-কে পরিমিত ব্যবধানের প্রতীক হিসাবে ব্যবহার করা হয়। যখন নমুনা উপাত্ত থেকে সমগ্রকের বৈশিষ্ট্যকে প্রাক্কলন করা হয় তখন সিদ্ধান্তমূলক পরিসংখ্যানে নমুনার পরিমিত ব্যবধানকে চিহ্নিত করার ক্ষেত্রে পরিমিত ব্যবধানের প্রতীক হিসাবে '$\hat{s}$'-কে ব্যবহার করা হয়। এটি হলো প্রাক্কলিত পরিমিত ব্যবধান। '$\hat{s}$' প্রতীকটিকে 's-hat' বলে। সিদ্ধান্তমূলক পরিসংখ্যানে সমগ্রকের পরিমিত ব্যবধানকে চিহ্নিত করতে গিয়ে গ্রীক বর্ণমালার ছোট অক্ষরের σ (সিগ্‌মা)-কে ব্যবহার করা হয়।

পরিমিত ব্যবধান নির্ণয়ের সূত্র (Formula for Computing Standard Deviation)

পরিমিত ব্যবধানের সংজ্ঞাটিকে বিশ্লেষণ করলে এর সূত্রটি খুব সহজেই দৃশ্যমান হয়ে উঠে। আমরা জানি যে, গাণিতিক গড় থেকে প্রতিটি মানের বর্গকৃত বিচ্যুতির গড়ের বর্গমূল হলো পরিমিত ব্যবধান।

এই সংজ্ঞাটির প্রতিটি অংশকে বুঝতে পারলেই সূত্রের উপাদানগুলো পাওয়া যাবে। যেমন,

বর্গমূল $= \sqrt{}$

গড়ের $= \dfrac{\sum}{N}$

গাণিতিক গড় থেকে প্রতিটি মানের বিচ্যুতির বর্গ $= (x_i - \bar{x})^২$

পরিমিত ব্যবধান $= s$

এই টুকরো অংশগুলো সংজ্ঞা অনুযায়ী যার যার স্থানে স্থাপন করলে আমরা পূর্ণ সূত্রটি পাই।

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^২}{N}}$$

যেখানে, s = পরিমিত ব্যবধান

x_i = উপাত্তরাশির প্রতিটি মান

$\bar{x}$ = গাণিতিক গড়

N = মোট ঘটনসংখ্যা

$\sqrt{}$ = বর্গমূল

$(x_i - \bar{x})^২$ = গড় থেকে প্রতিটি মানের বিচ্যুতির বর্গ

$\sum$ = সমষ্টি চিহ্ন

এই সূত্রটি হলো সংজ্ঞাগত সূত্র। সংজ্ঞাগত সূত্রের মাধ্যমে পরিসংখ্যান নির্ণয় পদ্ধতিকে সরাসরি পদ্ধতি বলে।

অবিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে পরিমিত ব্যবধান নির্ণয় (Computing Standard Deviation from Ungrouped Data Using Definitional Formula)

একটি উদাহরণের মাধ্যমে সংজ্ঞাগত সূত্র ব্যবহার করে অবিন্যস্ত উপাত্ত থেকে পরিমিত ব্যবধান নির্ণয় করে দেখা যাক। ধরা যাক, ৫ জনের একটি দলের সদস্যদের ওজন

১৮২ ২০০ ২০৩ ২০৫ ২১০ পাউন্ড

এবং ৫ জনের অন্য একটি দলের সদস্যদের ওজন

১৪০ ১৬৮ ১৯০ ২০০ ৩০২ পাউন্ড

এই দু'টি দলের পরিমিত ব্যবধান নির্ণয় পদ্ধতি সারণি ৫.৪.১-এ দেখানো হলো।

সারণি ৫.৪.১: সংজ্ঞাগত সূত্রের মাধ্যমে ৫ সদস্য বিশিষ্ট দু'টি দলের পরিমিত ব্যবধান নির্ণয়

প্রথম দল				দ্বিতীয় দল			
x_i	$\bar{x}$	$(x_i-\bar{x})$	$(x_i-\bar{x})^২$	x_i	$\bar{x}$	$(x_i-\bar{x})$	$(x_i-\bar{x})^২$
১৮২	২০০	–১৮	৩২৪	১৪০	২০০	–৬০	৩৬০০
২০০	২০০	০	০	১৬৮	২০০	–৩২	১০২৪
২০৩	২০০	৩	৯	১৯০	২০০	–১০	১০০
২০৫	২০০	৫	২৫	২০০	২০০	০	০
২১০	২০০	১০	১০০	৩০২	২০০	১০২	১০৪০৪
১০০০			**৪৫৮**	**১০০০**			**১৫১২৮**

আমাদের দু'টি দলের সদস্যদের ওজনের গড় নির্ণয় করে নিতে হবে। অতএব,

$$\bar{x}_১ = \frac{\sum x_i}{N} = \frac{১০০০}{৫} = ২০০$$

$\therefore$ নির্ণীত গাণিতিক গড় $\bar{x}_১ = ২০০$

$$\bar{x}_২ = \frac{\sum x_i}{N} = \frac{১০০০}{৫} = ২০০$$

$\therefore$ নির্ণীত গাণিতিক গড় $\bar{x}_২ = ২০০$

এখন নির্ণীত গাণিতিক গড় ব্যবহার করে দু'টি দলের পরিমিত ব্যবধান নির্ণয় করা যায়।

$$s_১ = \sqrt{\frac{\sum (x_i-\bar{x})^২}{N}} = \sqrt{\frac{৪৫৮}{৫}} = \sqrt{৯১.৬} = ৯.৫৭$$

$\therefore$ নির্ণীত পরিমিত ব্যবধান $s_১ = ৯.৫৭$

$$s_২ = \sqrt{\frac{\sum (x_i-\bar{x})^২}{N}} = \sqrt{\frac{১৫১২৮}{৫}} = \sqrt{৩০২৫.৬} = ৫৫.০১$$

$\therefore$ নির্ণীত পরিমিত ব্যবধান $s_২ = ৫৫.০১$

অবিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে পরিমিত ব্যবধান নির্ণয় (Computing Standard Deviation from Ungrouped Data Using Short-Cut Method)

এর পূর্বে যে সূত্রটি ব্যবহার করে পরিমিত ব্যবধান নির্ণয় করা হয়েছে সেটি হলো সংজ্ঞাগত সূত্র। সংজ্ঞাগত সূত্রটি পরিমিত ব্যবধান নির্ণয়ের যে পদ্ধতি নির্দেশ করে তাকে প্রত্যক্ষ পদ্ধতি বলে। সংজ্ঞাগত সূত্রের সুবিধা হলো, একটি পরিসংখ্যানের জন্ম যেভাবে হয় তা এই সূত্রের মাধ্যমে দেখা যায়। কিন্তু অসুবিধাটি হলো, এই পদ্ধতিতে পরিমিত ব্যবধান নির্ণয় অধিকতর শ্রমসাধ্য। যখন উপাত্তরাশির সংখ্যা অনেক বেশী হয় তখন এটি সময়সাধ্যও বটে। সে কারণে পরিসংখ্যানবিদগণ পরিমিত ব্যবধান নির্ণয়ের সংক্ষিপ্ত পদ্ধতি উদ্ভাবন করেছেন। সংক্ষিপ্ত পদ্ধতিতে প্রতিটি মানের বিচ্যুতি নির্ণয় করতে হয় না। উপাত্তরাশির সমষ্টি এবং প্রতিটি মানের বর্গের সমষ্টি নির্ণয় করলেই চলে।

সংক্ষিপ্ত পদ্ধতিতে পরিমিত ব্যবধান নির্ণয়ের সূত্রটি হলো,

$$s = \sqrt{\frac{\sum x_i^{২}}{N} - \left(\frac{\sum x_i}{N}\right)^{২}}$$

যেখানে, x_i = উপাত্তরাশির প্রতিটি মান

$\sum x_i$ = উপাত্তরাশির সমষ্টি

$\sum x_i^{২}$ = প্রতিটি মানের বর্গের সমষ্টি

s = পরিমিত ব্যবধান

সারণি ৫.৪.১-এ প্রদত্ত প্রথম দলের সদস্যদের ওজনের উপাত্তরাশি ব্যবহার করে সারণি ৫.৪.২-এ সংক্ষিপ্ত পদ্ধতিতে পরিমিত ব্যবধান নির্ণয় করে দেখানো হলো।

সারণি ৫.৪.২: অবিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে পরিমিত ব্যবধান নির্ণয়

x_i	$x_i^{২}$
১৮২	৩৩১২৪
২০০	৪০০০০
২০৩	৪১২০৯
২০৫	৪২০২৫
২১০	৪৪১০০
$\sum x_i$ = ১০০০	$\sum x_i^{২}$ = ২০০৪৫৮

প্রাপ্ত মানগুলো সূত্রে প্রয়োগ করে আমরা পরিমিত ব্যবধান নির্ণয় করতে পারি।

$$s = \sqrt{\frac{\sum x_i^{২}}{N} - \left(\frac{\sum x_i}{N}\right)^{২}}$$

$$= \sqrt{\frac{২০০৪৫৮}{৫} - \left(\frac{১০০০}{৫}\right)^{২}}$$

$$= \sqrt{৪০০৯১.৬ - (২০০)^{২}}$$

$$= \sqrt{৪০০৯১.৬ - ৪০০০০}$$

$$= \sqrt{৯১.৬}$$

$$= ৯.৫৭$$

$\therefore$ নির্ণীত পরিমিত ব্যবধান হলো ৯.৫৭

বিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে পরিমিত ব্যবধান নির্ণয় (Computing Standard Deviation from Grouped Data Using Definitional Formula)

আমরা জানি যে, সংজ্ঞাগত সূত্রের মাধ্যমে পরিসংখ্যান নির্ণয়ের পদ্ধতিকে প্রত্যক্ষ পদ্ধতি বলে। বিন্যস্ত উপাত্ত থেকে প্রত্যক্ষ পদ্ধতিতে পরিমিত ব্যবধান নির্ণয়ের সূত্রটি হলো,

$$s = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{N}}$$

যেখানে, s = পরিমিত ব্যবধান

f_i = প্রতিটি শ্রেণীর ঘটনসংখ্যা

x_i = শ্রেণী ব্যাপ্তির মধ্য-বিন্দু

$\bar{x}$ = গাণিতিক গড়

N = মোট ঘটনসংখ্যা

$(x_i - \bar{x})^2$ = গড় থেকে প্রতিটি শ্রেণী মধ্য-বিন্দুর বিচ্যুতির বর্গ

সূত্রের প্রয়োজনীয় মানগুলো পেতে হলে প্রথমে উপাত্তরাশির গড় নির্ণয় করে নিতে হবে। দ্বিতীয় ধাপে, নির্ণীত গড় থেকে প্রতিটি শ্রেণী ব্যাপ্তির মধ্য-বিন্দুর বিচ্যুতি নির্ণয় করতে হবে। এরপর, বিচ্যুতিগুলোর বর্গ করে নিয়ে প্রতিটি শ্রেণীর ঘটনসংখ্যা দিয়ে সংশ্লিষ্ট বিচ্যুতির বর্গগুলোকে গুণ করে প্রাপ্ত গুণফলের সমষ্টি নির্ণয় করতে হবে। সারণি ৫.৪.৩-এ সেগুলো ধাপে ধাপে নির্ণয় করে দেখানো হলো।

সারণি ৫.৪.৩: বিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে পরিমিত ব্যবধান নির্ণয়

শ্রেণী ব্যাপ্তি (C.I.)	*শ্রেণী ব্যাপ্তির মধ্য-বিন্দু* (x_i)	*ঘটনসংখ্যা* (f_i)	($f_i x_i$)	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^2$
১-৯	৫	৪০০	২০০০	–১২.১	১৪৬.৪১	৫৮৫৬৪.০০
১০-১৮	১৪	৫০০	৭০০০	–৩.১	৯.৬১	৪৮০৫.০০
১৯-২৭	২৩	৩০০	৬৯০০	৫.৯	৩৪.৮১	১০৪৪৩.০০
২৮-৩৬	৩২	১৫০	৪৮০০	১৪.৯	২২২.০১	৩৩৩০১.৫০
৩৭-৪৫	৪১	১০০	৪১০০	২৩.৯	৫৭১.২১	৫৭১২১.০০
মোট		***১৪৫০***	***২৪৮০০***			***১৬৪২৩৪.৫০***

$$\therefore \quad \bar{x} = \frac{\sum f_i x_i}{N}$$

$$= \frac{২৪৮০০}{১৪৫০}$$

$$= ১৭.১০$$

$$\therefore \quad \bar{x} = ১৭.১০$$

প্রাপ্ত মানগুলো সূত্রে প্রয়োগ করে আমরা পরিমিত ব্যবধানের মানটি পাই।

অতএব, $s = \sqrt{\frac{\sum f_i (x_i - \bar{x})^{২}}{N}}$

$= \sqrt{\frac{১৬৪২৩৪.৫}{১৪৫০}}$

$= \sqrt{১১৩.২৭}$

$= ১০.৬৪$

$\therefore$ নির্ণীত পরিমিত ব্যবধান হলো ১০.৬৪

বিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে পরিমিত ব্যবধান নির্ণয় (Computing Standard Deviation from Grouped Data Using Short-Cut Method)

শ্রেণী বিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে পরিমিত ব্যবধান নির্ণয়ের ক্ষেত্রে প্রথমে শ্রেণী ব্যাপ্তির মধ্য-বিন্দু নির্ণয়ের পর তার যে কোন একটিকে অনুমিত গড় ধরে নিয়ে সেই অনুমিত গড় থেকে প্রতিটি মধ্য-বিন্দুর স্তর বিচ্যুতি নির্ণয় করতে হয়। দ্বিতীয় ধাপে, স্তর বিচ্যুতিগুলোকে প্রতিটি শ্রেণীর ঘটনসংখ্যা দিয়ে গুণ করতে হবে। তৃতীয় ধাপে, স্তর বিচ্যুতিগুলোর বর্গ করে নিতে হয় এবং সব শেষে, বর্গকৃত বিচ্যুতিগুলোকে প্রতিটি শ্রেণীর ঘটনসংখ্যা দিয়ে গুণ করে প্রাপ্ত গুণফলগুলোর সমষ্টি নির্ণয় করতে হয়। সংক্ষিপ্ত পদ্ধতিতে পরিমিত ব্যবধান নির্ণয়ের বীজগাণিতিক সূত্রটি হলো,

$$s = \sqrt{\frac{\sum fd^{২}}{N} - \left(\frac{\sum fd}{N}\right)^{২}} \times C.I.$$

যেখানে, s = পরিমিত ব্যবধান

f_i = শ্রেণীভূক্ত ঘটনসংখ্যা

d = স্তর বিচ্যুতি $= \frac{x_i - A}{C.I.}$

x_i = শ্রেণী ব্যাপ্তির মধ্য-বিন্দু

A = অনুমিত গড়

N = মোট ঘটনসংখ্যা

$C.I.$ = শ্রেণী ব্যাপ্তি

সারণি ৫.৪.৩-এ প্রদত্ত উপাত্ত ব্যবহার করে সারণি ৫.৪.৪-এ সংক্ষিপ্ত পদ্ধতিতে পরিমিত ব্যবধান নির্ণয় করে দেখানো হলো।

সারণি ৫.৪.৪: বিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে পরিমিত ব্যবধান নির্ণয়

শ্রেণী ব্যাপ্তি (C.I.)	*শ্রেণী ব্যাপ্তির মধ্য-বিন্দু* (x_i)	*ঘটনসংখ্যা* (f_i)	*স্তর বিচ্যুতি* $d = \frac{x_i - A}{C.I.}$	fd	$d^২$	$fd^২$
১ - ৯	৫	৪০০	–২	–৮০০	৪	১৬০০
১০ - ১৮	১৪	৫০০	–১	–৫০০	১	৫০০
১৯ - ২৭	২৩ = A	৩০০	০	০	০	০
২৮ - ৩৬	৩২	১৫০	১	১৫০	১	১৫০
৩৭ - ৪৫	৪১	১০০	২	২০০	৪	৪০০
মোট		***১৪৫০***		***–৯৫০***		***২৬৫০***

পরিমিত ব্যবধান নির্ণয়ের সূত্রে ব্যবহারের জন্য প্রয়োজনীয় মানগুলো আমরা পেয়ে গিয়েছি। এবার সেগুলো সূত্রে প্রয়োগ করলে আমরা পরিমিত ব্যবধানের মানটি পেতে পারি। অতএব,

$$s = \sqrt{\frac{\sum fd^{২}}{N} - \left(\frac{\sum fd}{N}\right)^{২}} \times C.I.$$

$$= \sqrt{\frac{২৬৫০}{১৪৫০} - \left(\frac{-৯৫০}{১৪৫০}\right)^{২}} \times ৯$$

$$= \sqrt{১.৮৩ - (-০.৬৬)^{২}} \times ৯$$

$$= \sqrt{১.৮৩ - ০.৪৪} \times ৯$$

$$= \sqrt{১.৩৯} \times ৯$$

$$= ১.১৮ \times ৯$$

$$= ১০.৬২$$

∴ নির্ণীত পরিমিত ব্যবধান হলো ১০.৬২

দু'টি পদ্ধতিতেই একই মান পাওয়া যায়, তবে দশমিকের পরে দু'টি করে সংখ্যা নেওয়াতে এই পদ্ধতিতে পরিমিত ব্যবধানের মান .০২ কম হয়েছে।

পরিমিত ব্যবধানের সুবিধা এবং অসুবিধা (Advantages and Disadvantages of Standard Deviation)

উপাত্তরাশির সকল মানকে অনপেক্ষভাবে না ভেবে প্রকৃতভাবে বিবেচনা করাই হলো পরিমিত ব্যবধানের প্রধান সুবিধা। অর্থাৎ, গড় ব্যবধান নির্ণয়ের সময় ঋণাত্মক চিহ্নকে উপেক্ষা করার জন্য যে অনপেক্ষ বিচ্যুতি নির্ণয় করা হয় তা একটি কৃত্রিম উপায়। তত্ত্বগতভাবে গড় ব্যবধানের কোন ব্যাখ্যা দেয়া যায় না এবং পরবর্তী বীজগাণিতিক পরিগণনায়ও প্রয়োগ করা যায় না। ঋনাত্মক চিহ্ন উপেক্ষা করার এই কৃত্রিম পদ্ধতিতে সৃষ্ট সীমাবদ্ধতাগুলো একমাত্র পরিমিত ব্যবধানের মাধ্যমেই দূর করা সম্ভব। পরিমিত ব্যবধানকে খুব সহজেই পরবর্তী বীজগাণিতিক পরিগণনার ক্ষেত্রে প্রয়োগ করা যায়। অন্যান্য পরিমাপের তুলনায় পরিমিত ব্যবধান নমুনা বিচ্যুতি দ্বারা কম প্রভাবিত হয় বলে এটি অনেক বেশী স্থীতিশীল একটি

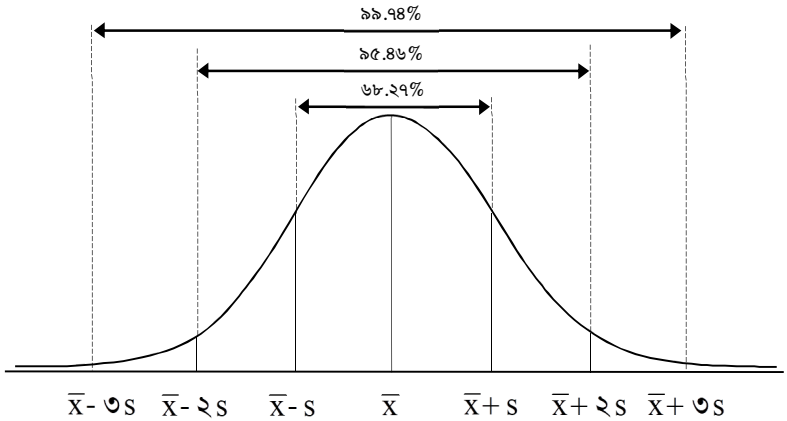
পরিমাপ। দুই বা ততোধিক উপাত্তরাশির পরিমিত ব্যবধান থেকে আমরা সম্মিলিত পরিমিত ব্যবধান নির্ণয় করতে পারি।

তবে যেহেতু পরিমিত ব্যবধান নির্ণয়ে উপাত্তরাশির সকল মানকে বিবেচনায় আনা হয়, সেহেতু এটি চরম বা অসাধারণ মান দ্বারা খুব সহজেই প্রভাবিত হয়। পরিমিত ব্যবধান নির্ণয়ের ক্ষেত্রে গাণিতিক গড়কে ব্যবহার করা হয় বলে গাণিতিক গড়ের সকল সীমাবদ্ধতা পরিমিত ব্যবধানের মধ্যে থাকে। গড় ব্যবধানের মতো এটিও উন্মুক্ত শ্রেণী সীমা সম্বলিত উপাত্তের ক্ষেত্রে প্রয়োগ করা যায় না। পরিমিত ব্যবধানকে বুঝতে এবং নির্ণয় করতে হলে যথেষ্ট গাণিতিক জ্ঞানের প্রয়োজন হয়।

পরিমিত ব্যবধানের ব্যবহার (Use of the Standard Deviation)

নমুনার আকার নির্ধারণে পরিমিত ব্যবধান একটি নিয়ামক ভূমিকা পালন করে।

বিস্তৃতির পরিমাপগুলোর মধ্যে পরিমিত ব্যবধান সবচেয়ে উপযোগী এবং উচ্চতর পরিসংখ্যানিক বিশ্লেষণে এটিকে ব্যাপকভাবে ব্যবহার করা হয়। নমুনায়ন এবং সহ-সম্পর্ক বিশ্লেষণের ভিত্তিটি দাঁড়িয়ে আছে পরিমিত ব্যবধানের উপর। নমুনার আকার নির্ধারণে পরিমিত ব্যবধান একটি নিয়ামক ভূমিকা পালন করে। পরিমিত রেখার বিশ্লেষণ পরিমিত ব্যবধানের পরিপ্রেক্ষিতে হয়ে থাকে। যেমন, একটি বিন্যাস যদি সুষম হয়, তবে তা মসৃণ স্বাভাবিক রেখার (smooth normal curve) রূপ ধারণ করে এবং একটি সুষম বিন্যাসের অধীনে ছয়টি (গড়ের ডান দিকে তিনটি এবং বাম দিকে তিনটি) পরিমিত ব্যবধান রয়েছে। এই ছয়টি পরিমিত ব্যবধানের মোট আয়তন হলো ৯৯.৭৪ শতাংশ। অন্য কথায়, গড়ে তিন পরিমিত ব্যবধানের উপরে এবং নীচে ৯৯.৭৪ শতাংশ ঘটনসংখ্যা অন্তর্ভুক্ত; দুই পরিমিত ব্যবধানের উপরে এবং নীচে ৯৫.৪৫ শতাংশ; এবং এক পরিমিত ব্যবধানের উপরে এবং নীচে ৬৮.২৭



শতাংশ ঘটনসংখ্যা অন্তর্ভুক্ত থাকে (চিত্র ৫.৪.১ দ্রষ্টব্য)।

চিত্র ৫.৪.১: স্বাভাবিক রেখার অধীনে আয়তন

পরিমিত ব্যবধানের অন্যতম প্রধান ব্যবহারটি হলো একটি বিন্যাসের মধ্যে প্রতিটি মানের আপেক্ষিক অবস্থান কি তা নির্ণয় করা। দু'টি উপাত্তরাশির মধ্যে সমরূপতা বা ভিন্নতার তুলনা করার জন্য পরিমিত ব্যবধান একটি বাস্তব ভিত্তি প্রদান করে থাকে।

সারাংশ

পরিমিত ব্যবধান হলো গাণিতিক গড় থেকে প্রতিটি মানের বর্গকৃত বিচ্যুতির গড়ের বর্গমূল। পরিমিত ব্যবধান অবিন্যস্ত এবং বিন্যস্ত উভয় প্রকার উপাত্ত থেকে নির্ণয় করা সম্ভব। উপাত্তের বর্ণনা এবং বিশ্লেষণে পরিমিত ব্যবধানকে ব্যাপকভাবে ব্যবহার করা হয়। নমুনার আকার নির্ধারণেও পরিমিত ব্যবধান একটি নিয়ামক ভূমিকা পালন করে। এর অন্যতম ব্যবহার হলো একটি বিন্যাসের মধ্যে প্রতিটি মানের আপেক্ষিক অবস্থানের চিত্র নির্ণয় করা।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক (√) চিহ্ন দিন –

১। গাণিতিক গড় থেকে প্রতিটি মানের বর্গকৃত বিচ্যুতির গড়ের বর্গমূল হল:

ক. পরিমিত গড়
খ. স্থূল বিচ্যুতি
গ. পরিমিত ব্যবধান
ঘ. অনপেক্ষ বিচ্যুতি

২। ..... নির্ণয়ে প্রতিটি মানের বর্গের সমষ্টির প্রয়োগ হয়।

ক. গড়
খ. পরিমিত ব্যবধান
গ. প্রচুরক
ঘ. মধ্যমা

৩। অবিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে পরিমিত ব্যবধান নির্ণয়ের সূত্র হলো:

ক. $s = \sqrt{\frac{\sum x_i}{N} - \left(\frac{\sum x_i}{N}\right)^{২}}$

খ. $s = \sqrt{\frac{\sum x_i^{২}}{N} - \left(\frac{\sum x_i}{N}\right)^{২}}$

গ. $s = \sqrt{\frac{\sum x_i^{২}}{N-১} - \left(\frac{\sum x_i}{N-১}\right)^{২}}$

ঘ. $s = \sqrt{\frac{\sum x_i^{২}}{N} - \left(\frac{\sum x_i}{N-১}\right)^{২}}$

সংক্ষিপ্ত প্রশ্ন

১। পরিমিত ব্যবধান কি?

২। অবিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে পরিমিত ব্যবধান নির্ণয়ের সুবিধাগুলো সংক্ষেপে আলোচনা করুন।

রচনামূলক প্রশ্ন

১। পরিমিত ব্যবধানের ব্যবহার উদাহরণসহ আলোচনা করুন।

২। পরিমিত ব্যবধান ব্যবহারের সুবিধা এবং অসুবিধা আলোচনা করুন।

পাঠ - ৫

বিস্তৃতির নিরঙ্কুশ পরিমাপ – ৪: ভেদাঙ্ক
Absolute Measures of Dispersion – 4: Variance

এই পাঠ শেষে যা জানা যাবে —

- ভেদাঙ্ক কি
- অবিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে ভেদাঙ্ক নির্ণয়
- অবিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে ভেদাঙ্ক নির্ণয়
- বিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে ভেদাঙ্ক নির্ণয়
- বিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে ভেদাঙ্ক নির্ণয়

ভেদাঙ্ক কি (What is Variance)?

বর্গমূল না নিয়ে বর্গকৃত বিচ্যুতির গড় নির্ণয় করা হয়, যা ভেদাঙ্ক নামে পরিচিত। অর্থাৎ, উপাত্তরাশির গড় থেকে সংখ্যামানসমূহের বিচ্যুতির বর্গের গড়ই হলো ভেদাঙ্ক।

খুব সহজেই আমাদের মনে প্রশ্ন আসতে পারে যে, বিস্তৃতির পরিমাপ নির্ণয়ের ক্ষেত্রে বর্গমূল নেবার কোন প্রয়োজনীয়তা রয়েছে কি? এর সহজ উত্তরটি হতে পারে যে, পরিমিত ব্যবধানকে এভাবেই সংজ্ঞায়িত করা হয়েছে বলে বর্গমূল নিতে হয়। কিন্তু এটি আদৌ কোন সন্তোষজনক উত্তর নয়। বরং বলা যায়, অনপেক্ষতার সমস্যা দূর করার জন্য যেহেতু আমরা গাণিতিক গড় থেকে প্রতিটি মানের বিচ্যুতির বর্গ করি, সেহেতু সেই বর্গের ক্ষতিপূরণ হিসাবে বর্গকৃত বিচ্যুতির গড়ের বর্গমূল নিয়ে থাকি। তবে উচ্চতর পরিসংখ্যানিক কাজে ব্যবহারের ক্ষেত্রে বর্গমূল না নিয়ে বর্গকৃত বিচ্যুতির গড় নির্ণয় করা হয়, যা ভেদাঙ্ক নামে পরিচিত। অর্থ্যাৎ, উপাত্তরাশির গড় থেকে সংখ্যা মানসমূহের বিচ্যুতির বর্গের গড়ই হলো ভেদাঙ্ক। পরিমিত ব্যবধানের সাথে ভেদাঙ্কের পার্থক্য হলো পরিমিত ব্যবধানে বিচ্যুতির গড়ের বর্গমূল নির্ণয়ের প্রয়োজন হয়, ভেদাঙ্কে তার প্রয়োজন হয় না। ভেদাঙ্ককে গড় বর্গও (mean square) বলা হয়ে থাকে। ভেদাঙ্ককে ইংরেজী বর্ণমালার ছোট অক্ষরের $s^২$ প্রতীকের মাধ্যমে চিহ্নিত করা হয়। সমগ্রকের ভেদাঙ্ককে গ্রীক বর্ণমালার ছোট অক্ষরের $\sigma^২$ প্রতীক দিয়ে চিহ্নিত করা হয়।

ভেদাঙ্ক দু'ভাবে নির্ণয় করা যায়। *প্রথমতঃ*, নির্ণীত পরিমিত ব্যবধানের বর্গ করে। যেমন, পরিমিত ব্যবধান নির্ণয়ের উদাহরণে পরিমিত ব্যবধানের প্রাপ্ত মানটি ছিল ১০.৬৪। এই মানটিকে বর্গ করে নিলেই আমরা ভেদাঙ্কের মান পাবো। অতএব,

$$s^২ = (১০.৬৪)^২$$
$$= ১১৩.২১$$

একটি বিষয় মনে রাখা প্রয়োজন যে, নির্ণয়ের সময় rounding-এর কারণে দশমিকের পরে ভগ্নাংশের ক্ষেত্রে মানের সামান্য তারতম্য ঘটতে পারে। তবে কম্পিউটার বা উচ্চ ক্ষমতাসম্পন্ন ক্যালক্যুলেটর ব্যবহার করলে সেই তারতম্যটি আর থাকে না।

দ্বিতীয় পদ্ধতিটি হলো, ভেদাঙ্ক নির্ণয়ের সংজ্ঞাগত সূত্র ব্যবহার করে। এখন আমরা ভেদাঙ্ক নির্ণয়ের সংজ্ঞাগত সূত্র ব্যবহার করে এর নির্ণয় পদ্ধতি আলোচনা করবো।

অবিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে ভেদাঙ্ক নির্ণয় (Computing Variance from Ungrouped Data Using Definitional Formula)

অবিন্যস্ত উপাত্ত থেকে ভেদাঙ্ক নির্ণয়ের জন্য নিম্নলিখিত সংজ্ঞাগত সূত্রটি ব্যবহার করা হয়।

$$s^২ = \frac{\sum (x_i - \bar{x})^২}{N}$$

যেখানে, $s^২$ = ভেদাঙ্ক

x_i = উপাত্তরাশির প্রতিটি মান

$\bar{x}$ = গাণিতিক গড়

N = উপাত্তরাশির মোট সংখ্যা

$(x_i - \bar{x})^২$ = গড় থেকে প্রতিটি মানের বিচ্যুতির বর্গ

সারণি ৫.৪.৫-এ অবিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্র ব্যবহার করে ভেদাঙ্ক নির্ণয় পদ্ধতি দেখানো হলো।

সারণি ৫.৪.৫: অবিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে ভেদাঙ্ক নির্ণয়

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^২$
৬	– ১৭	২৮৯
১০	– ১৩	১৬৯
১৫	– ৮	৬৪
২৭	৪	১৬
৩১	৮	৬৪
৩২	৯	৮১
৪০	১৭	২৮৯
$\sum x_i =$ ১৬১		$\sum (x_i - \bar{x})^২ =$ ৯৭২

প্রথমে আমাদের উপাত্তরাশির গাণিতিক গড় নির্ণয় করে নিতে হবে। অতএব,

$$\bar{x} = \frac{\sum x_i}{N} = \frac{১৬১}{৭} = ২৩$$

∴ নির্ণীত গড় হলো ২৩

এবার আমাদের নির্ণীত গড় থেকে প্রতিটি মানের বিচ্যুতির বর্গ করে সেগুলোর সমষ্টি নির্ণয় করতে হবে। প্রাপ্ত বিচ্যুতির বর্গের সমষ্টির মানটি সূত্রে প্রয়োগ করে উপাত্তরাশির মোট সংখ্যা দিয়ে ভাগ করলে আমরা ভেদাঙ্কের মান পাই। অতএব,

$$s^২ = \frac{\sum (x_i - \bar{x})^২}{N} = \frac{৯৭২}{৭} = ১৩৮.৮৬$$

∴ নির্ণীত ভেদাঙ্ক হলো ১৩৮.৮৬

অবিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে ভেদাঙ্ক নির্ণয় (Computing Variance from Ungrouped Data using Short-Cut Method)

সংজ্ঞাগত সূত্র ব্যবহার করে ভেদাঙ্ক নির্ণয় একটু কষ্টসাধ্য এবং সময় সাপেক্ষ। তাই পরিসংখ্যানবিদগণ ভেদাঙ্ক নির্ণয়ের জন্য একটি সহজ সূত্র উদ্ভাবন করেছেন। সেটি হলো,

$$s^২ = \frac{\sum x_i^২}{N} - \left(\frac{\sum x_i}{N}\right)^২$$

যেখানে, $s^২$ = ভেদাঙ্ক

x_i = উপাত্তরাশির প্রতিটি মান

$x_i^২$ = প্রতিটি মানের বর্গ

N = উপাত্তরাশির মোট সংখ্যা

সারণি ৫.৪.৫-এ প্রদত্ত উপাত্ত ব্যবহার করে সারণি ৫.৪.৬-এ সংক্ষিপ্ত পদ্ধতিতে ভেদাঙ্ক নির্ণয় করে দেখানো হলো।

সারণি ৫.৪.৬: অবিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে ভেদাঙ্ক নির্ণয়

x_i	$x_i^২$
৬	৩৬
১০	১০০
১৫	২২৫
২৭	৭২৯
৩১	৯৬১
৩২	১০২৪
৪০	১৬০০
$\sum x_i = ১৬১$	$\sum x_i^২ = ৪৬৭৫$

প্রাপ্ত মানগুলো সূত্রে প্রয়োগ করে আমরা ভেদাঙ্কের মান পাই।

$$s^২ = \frac{\sum x_i^২}{N} - \left(\frac{\sum x_i}{N}\right)^২$$

$$= \frac{৪৬৭৫}{৭} - \left(\frac{১৬১}{৭}\right)^২$$

$$= ৬৬৭.৮৬ - ৫২৯$$

$$= ১৩৮.৮৬$$

$\therefore$ নির্ণীত ভেদাঙ্ক হলো ১৩৮.৮৬

বিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে ভেদাঙ্ক নির্ণয় (Computing Variance from Grouped Data Using Definitional Formula)

বিন্যস্ত উপাত্ত থেকে ভেদাঙ্ক নির্ণয়ের সংজ্ঞাগত সূত্রটি হলো,

$$s^২ = \frac{\sum f_i (x_i - \bar{x})^২}{N}$$

যেখানে, $s^২$ = ভেদাঙ্ক

f_i = প্রতিটি শ্রেণীর ঘটনসংখ্যা

x_i = প্রতিটি শ্রেণীর মধ্য-বিন্দু

$\bar{x}$ = গাণিতিক গড়

N = মোট ঘটনসংখ্যা

$(x_i - \bar{x})^২$ = গড় থেকে প্রতিটি মানের বিচ্যুতির বর্গ

$\sum$ = সমষ্টি চিহ্ন

সারণি ৫.৪.৭-এ উপস্থাপিত ১০০ জন ছাত্রের ব্যয়ের উপাত্ত বিন্যাসকে ব্যবহার করে উপরোক্ত সূত্রের মাধ্যমে ভেদাঙ্ক নির্ণয় করে দেখানো হলো।

সারণি ৫.৪.৭: বিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে পদ্ধতিতে ভেদাঙ্ক নির্ণয়

শিক্ষাবছর সমাপ্ত	*ঘটন সংখ্যা* f_i	*শ্রেণী ব্যাপ্তির মধ্য-বিন্দু* x_i	$f_i x_i$	*গড় থেকে প্রতিটি মধ্য-বিন্দুর বিচ্যুতি* $(x_i - \bar{x})$	*বিচ্যুতির বর্গ* $(x_i - \bar{x})^২$	$f_i (x_i - \bar{x})^২$
০ - ২	২০	১	২০	–৫.২৩	২৭.৩৫	৫৪৭.০০
২ - ৪	২৮	৩	৮৪	–৩.২৩	১০.৪৩	২৯২.০৪
৬ - ৮	১৯	৫	৯৫	–১.২৩	১.৫১	২৮.৬৯
৮ - ১০	৯	৭	৬৩	০.৭৭	০.৫৯	৫.৩১
১০ - ১২	৩৫	৯	৩১৫	২.৭৭	৭.৬৭	২৬৮.৪৫
১২ - ১৪	১০	১১	১১০	৪.৭৭	২২.৭৫	২২৭.৫০
১৪ - ১৬	৬	১৩	৭৮	৬.৭৭	৪৫.৮৩	২৭৪.৯৮
১৬ - ১৮	৩	১৫	৪৫	৮.৭৭	৭৬.৯১	২৩০.৭৩
মোট	***১৩০***		***৮১০***			***১৮৭৪.৭০***

ভেদাঙ্ক নির্ণয়ের পূর্ব শর্ত হিসাবে আমাদের গাণিতিক গড় নির্ণয় করে নিতে হবে। অতএব,

$$\bar{x} = \frac{\sum f_i x_i}{N}$$

$$= \frac{৮১০}{১৩০}$$

$$= ৬.২৩$$

∴ নির্ণীত গড় হলো ৬.২৩

এখন গড় থেকে প্রতিটি শ্রেণী ব্যাপ্তির মধ্য-বিন্দুর বিচ্যুতি নির্ণয় করতে হবে এবং বিচ্যুতিগুলোর বর্গ করে নিয়ে সেই বর্গকৃত বিচ্যুতিকে ঘটনসংখ্যা দিয়ে গুণ করে তার সমষ্টি নির্ণয় করতে হবে। প্রাপ্ত মানগুলো সূত্রে প্রয়োগ করলে আমরা ভেদাঙ্কের মান পাই। অতএব,

$$s^২ = \frac{\sum f_i (x_i - \bar{x})^২}{N} = \frac{১৮৭৪.৭}{১৩০} = ১৪.৪২$$

∴ নির্ণীত ভেদাঙ্ক হলো ১৪.৪২

বিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে ভেদাঙ্ক নির্ণয় (Computing Variance from Grouped Data Using Short-Cut Method)

বিন্যস্ত উপাত্তের ক্ষেত্রে সংক্ষিপ্ত পদ্ধতিতে ভেদাঙ্ক নির্ণয়ের জন্য সাধারণতঃ নিম্নের সূত্রটি ব্যবহৃত হয়।

$$s^{২} = \left\{\frac{\sum fd^{২}}{N} - \left(\frac{\sum fd}{N}\right)^{২}\right\} \times C.I.$$

যেখানে, $s^{২}$ = ভেদাঙ্ক

C.I. = শ্রেণী ব্যাপ্তি

d = স্তর বিচ্যুতি = $\frac{x_i - A}{C.I.}$

x_i = শ্রেণীর ব্যাপ্তির মধ্য-বিন্দু

A = অনুমিত গড়

f_i = ঘটনসংখ্যা

N = মোট ঘটনসংখ্যা

সারণি ৫.৪.৭-এ প্রদত্ত উপাত্ত ব্যবহার করে সারণি ৫.৪.৮-এ সংক্ষিপ্ত পদ্ধতিতে ভেদাঙ্ক নির্ণয় করে দেখানো হলো।

সারণি ৫.৪.৮: বিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে ভেদাঙ্ক নির্ণয়

শিক্ষার বছর	*ঘটনসংখ্যা* f_i	শ্রেণী ব্যাপ্তির মধ্য-বিন্দু (x_i)	$d = \frac{x_i - A}{C.I.}$	d^2	fd^2	fd
০ - ২	২০	১	–৩	৯	১৮০	–৬০
২ - ৪	২৮	৩	–২	৪	১১২	–৫৬
৪ - ৬	১৯	৫	–১	১	১৯	–১৯
৬ - ৮	৯	৭ = A	০	০	০	০
৮ - ১০	৩৫	৯	১	১	৩৫	৩৫
১০ - ১২	১০	১১	২	৪	৪০	২০
১২ - ১৪	৬	১৩	৩	৯	৫৪	১৮
১৪ - ১৬	৩	১৫	৪	১৬	৪৮	১২
মোট	***১৩০***				***৪৮৮***	***–৫০***

প্রাপ্ত মানগুলো সূত্রে প্রয়োগ করে ভেদাঙ্কের মান নির্ণয় করতে পারি। অতএব,

$$s^{২} = c^{২} \times \left\{\frac{\sum fd^{২}}{N} - \left(\frac{\sum fd}{N}\right)^{২}\right\}$$

$$= ২^{২}\left\{\frac{৪৮৮}{১৩০} - \left(\frac{-৫০}{১৩০}\right)^{২}\right\}$$

$$= \quad ৪\left\{ ৩.৭৫৪ - \left(-০.৩৮৫\right)^{২} \right\}$$
$$= \quad ৪\left\{ ৩.৭৫৪ - \left(০.১৪৮\right) \right\}$$
$$= \quad ৪\left\{ ৩.৬০৬ \right\}$$
$$= \quad ১৪.৪২$$

$\therefore$ নির্ণীত ভেদাঙ্ক হলো ১৪.৪২

সারাংশ

উচ্চতর পরিসংখ্যানিক কাজে ব্যবহারের ক্ষেত্রে গাণিতিক গড় থেকে প্রতিটি মানের বর্গকৃত বিচ্যুতির বর্গমূল না নিয়ে বর্গকৃত বিচ্যুতির গড় নির্ণয় করা হয় যা ভেদাঙ্ক নামে পরিচিত। অর্থ্যাৎ, উপাত্তরাশির গড় থেকে সংখ্যা মানসমূহের বিচ্যুতির বর্গের গড়ই হলো ভেদাঙ্ক। পরিমিত ব্যবধানের সাথে ভেদাঙ্কের পার্থক্য হলো পরিমিত ব্যবধানে বিচ্যুতির গড়ের বর্গমূল নির্ণয়ের প্রয়োজন হয়, ভেদাঙ্কে তার প্রয়োজন হয় না। ভেদাঙ্ককে গড় বর্গও (mean square) বলা হয়ে থাকে। ভেদাঙ্ক দু'ভাবে নির্ণয় করা যায়। *প্রথমতঃ*, নির্ণীত পরিমিত ব্যবধানের বর্গ করে। দ্বিতীয় পদ্ধতিটি হলো, ভেদাঙ্ক নির্ণয়ের সংজ্ঞাগত সূত্র ব্যবহার করে।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক (√) চিহ্ন দিন –

১। নীচের কোনটিকে গড় বর্গ বলে?

ক. পরিমিত ব্যবধান
খ. গড় ব্যবধান
গ. ভেদাঙ্ক
ঘ. শতহারিক পরিসর

২। ভেদাঙ্ক নির্ণয়ের অন্যতম পদ্ধতি হলো:

ক. পরিমিত ব্যবধানের বর্গ করে
খ. গড় ব্যবধানের বর্গ করে
গ. পরিসরের বর্গ করে
ঘ. উপরের কোনটিই নয়

৩। বিন্যস্ত উপাত্ত থেকে ভেদাঙ্ক নির্ণয়ের সূত্র হলো:

ক. $s^2 = \frac{\sum f_i(x_i - \bar{x})}{N}$

খ. $s^2 = \frac{\sum f_i(x_i - \bar{x})^2}{N}$

গ. $s^2 = \frac{\sum (x_i - \bar{x})^2}{N}$

ঘ. $s^2 = \sum f_i(x_i - \bar{x})^2$

সংক্ষিপ্ত প্রশ্ন

১। ভেদাঙ্ক কি?

২। ভেদাঙ্ক নির্ণয়ের পদ্ধতি দু'টি কি?

রচনামূলক প্রশ্ন

১। অবিন্যস্ত উপাত্ত থেকে সংক্ষিপ্ত পদ্ধতিতে কিভাবে ভেদাঙ্ক নির্ণয় করা হয় তা উদাহরণসহ আলোচনা করুন।

২। বিন্যস্ত উপাত্ত থেকে সংজ্ঞাগত সূত্রের মাধ্যমে কিভাবে ভেদাঙ্ক নির্ণয় করা হয় তা উদাহরণসহ আলোচনা করুন।

পাঠ - ৬

পরিমিত ব্যবধান ও ভেদাঙ্কের ব্যাখ্যা এবং ব্যবহার
Interpretation and Use of Standard Deviation and Variance

এই পাঠ শেষে যা জানা যাবে —

- পরিমিত ব্যবধান ও ভেদাঙ্কের মধ্যে তুলনা
- পরিমিত ব্যবধান ও ভেদাঙ্কের ব্যাখ্যা
- পরিমিত ব্যবধানের ব্যবহার

পরিমিত ব্যবধান ও ভেদাঙ্কের মধ্যে তুলনা (Comparison between Standard Deviation and Variance)

আমরা জানি যে, গাণিতিক গড়ের মত পরিমিত ব্যবধান ও ভেদাঙ্ক নির্ণয়ের ক্ষেত্রেও উপাত্তরাশির প্রতিটি মান ব্যবহৃত হয়। অতএব, উপাত্তরাশির যে কোন একটি মান পরিবর্তিত হলে এই পরিমাপ দু'টির মানের পরিবর্তন ঘটবে। এই পরিবর্তন নমুনা বিচ্যুতির মধ্যেও প্রভাব ফেলবে। কারণ, একই সমষ্টির একেকটি নমুনার ক্ষেত্রে একেকটি পরিমিত ব্যবধান ও ভেদাঙ্কের মান পাওয়া যায়। তবে পরিসর বা চতুর্থক ব্যবধানের তুলনায় এই পরিমাপ দু'টির ক্ষেত্রে নমুনা বিচ্যুতি কম ঘটে। যেহেতু সংখ্যা মানগুলো বর্গ করে সমষ্টি নির্ণয় করা হয়, সেহেতু এই পরিমাপ দু'টির মান কখনোই ঋণাত্মক হবে না। যদি বিন্যাসের সব মানগুলো একই হয় সে ক্ষেত্রে পরিমাপ দু'টির মান খুব কম হলে '০' হতে পারে। কারণ, সব মানগুলো '০' হলে সব বিচ্যুতিই হবে '০'। অতএব, বিচ্যুতির বর্গগুলো এবং তাদের সমষ্টিও '০' হবে। কিন্তু বাস্তবে তা ঘটে না। আর ঘটে না বলেই পরিসংখ্যান বিলাসিতা না হয়ে তা গবেষণা কার্যে প্রয়োজনীয়তা হিসাবে দেখা দেয়।

তবে দু'টি পরিমাপের মধ্যে পার্থক্য রয়েছে। যেমন, পরিমিত ব্যবধান মূল অশোধিত মানের এককে (original raw score units) থাকে, অথচ ভেদাঙ্ক থাকে বর্গকৃত এককে (squared units)। পরিমিত ব্যবধানকে লৈখিকভাবে উপস্থাপিত করা যায়, কিন্তু ভেদাঙ্ককে করা যায় না। পরিমিত ব্যবধান হলো মানের মাত্রায় একটি দূরত্ব নির্দেশ করে, আর ভেদাঙ্ক হলো বর্গকৃত দূরত্ব। পরিমিত ব্যবধান পরিসর, গাণিতিক গড়, মধ্যমা, ইত্যাদির মত একই এককে প্রকাশিত হয়, অথচ ভেদাঙ্ক তা করে না। সে কারণে, পরিমিত ব্যবধান বিস্তৃতির পরিমাপ হিসাবে অধিক ব্যবহৃত হয়। তবে পরিমিত ব্যবধানের চেয়ে ভেদাঙ্কের অধিকতর তাত্ত্বিক মূল্য রয়েছে। কারণ, পরিমিত ব্যবধানের ষজ্ঞাগত অর্থ (intuitive meaning) ততক্ষণ প্রতিভাত হয় না যতক্ষণ না তা পরিমিত রেখার অধীনে আয়তন (area under normal curve) নির্ণয়ে ব্যবহার করি। অন্য কথায়, পরিমিত ব্যবধান হলো একটি বিমূর্ত সংখ্যা যার পরিপূর্ণ ব্যবহার ও ব্যাখ্যা পরবর্তী পাঠে পরিমিত মান নির্ণয়ের আলোচনায় স্পষ্ট হয়ে উঠবে।

পরিমিত ব্যবধানকে লৈখিকভাবে উপস্থাপিত করা যায়, কিন্তু ভেদাঙ্ককে করা যায় না।

পরিমিত ব্যবধান ও ভেদাঙ্কের ব্যাখা (Interpretation of Standard Deviation and Variance)

এই পর্যায়ে এসে আমরা ভেদাঙ্ক ও পরিমিত ব্যবধান নির্ণয়ের সূত্রগুলো ব্যবহার করে একটি বিন্যাসের জন্য সংখ্যামূলক মান নির্ণয় করতে পারি এবং পরিমিত ব্যবধান ও ভেদাঙ্কের নির্ণীত মানগুলোকে প্রায়শঃ যাদুকরী সংখ্যা হিসেবে মনে করা হয়। কিন্তু এই সংখ্যা মানগুলোর নির্ণয়ই যথেষ্ট নয়। কারণ, কোন পরিসংখ্যানগত প্রত্যয়ের কার্যকর ব্যবহার করতে হলে আমাদের শুধু সেই প্রত্যয়গুলোর নির্ণয় পদ্ধতি জানলেই চলবে না, সেগুলোকে অন্যান্য প্রত্যয়ের সাথে সম্পর্কিত করতে হবে। সুনির্দিষ্টভাবে আমাদের জানতে হবে, কেন আমরা পরিমিত ব্যবধান নির্ণয় করবো? কেন প্রায় সব গবেষণা প্রতিবেদনে

কোন পরিসংখ্যানগত প্রত্যয়ের কার্যকর ব্যবহার করতে হলে আমাদের শুধু সেই প্রত্যয়গুলোর নির্ণয় পদ্ধতি জানলেই চলবে না, সেগুলোকে অন্যান্য প্রত্যয়ের সাথে সম্পর্কিত করতে হবে।

এটি অন্তর্ভুক্ত করা হয়ে থাকে? এবং কখন এটি নির্ণয় করতে হবে এবং কখন করতে হবে না? এর জন্য আমাদের পরিমিত ব্যবধান ও ভেদাঙ্কের অর্থবহ ব্যাখ্যা জানা প্রয়োজন।

প্রথম যে ব্যাখ্যাটি আমরা দিয়ে থাকি সেটি হলো, যদি পরিমিত ব্যবধান ও ভেদাঙ্কের মানটি ছোট হয় তবে আমরা বলি যে বিন্যাসটি সমরূপ, আর যদি বড় হয় তবে আমরা বলি বিন্যাসটির বিস্তৃতি অনেক বেশী। এই ব্যাখ্যাটি সরাসরি গাণিতিক গড়ের সাথে সম্পর্কিত। কারণ, পরিমিত ব্যবধানের সংজ্ঞার মধ্যে গাণিতিক গড়ের দ্বিতীয় বৈশিষ্ট্যটি (property of least squares) অন্তর্নিহিত রয়েছে। বিন্যাসের মানগুলো গড় থেকে বেশী দূরত্বে অবস্থান করলে বিচ্যুতিগুলো বড় হবে, বিচ্যুতির বর্গগুলো বড় হবে এবং পরিমিত ব্যবধান ও ভেদাঙ্কের মান বড় হবে। বিন্যাসের মানগুলো যদি গড় থেকে কম দূরত্বে অবস্থান করে তবে বিচ্যুতিগুলো ছোট হবে, তার বর্গগুলো ছোট হবে এবং পরিমিত ব্যবধান ও ভেদাঙ্কের মান ছোট হবে। সারণি ৫.৫.১-এ প্রদত্ত দু'টি রাশিমালাকে পর্যবেক্ষণ করলেই বিষয়টি স্পষ্ট হয়ে উঠবে।

সারণি ৫.৫.১: দু'টি বিন্যাসের বিস্তৃতির ভিন্নতা

বিন্যাস 'ক'		*বিন্যাস 'খ'*	
সংখ্যা মান =	২১ ২২ ২২ ২৩ ২৩	সংখ্যা মান =	০ ৩ ৬ ৮ ৯
	২৩ ২৪ ২৪ ২৫		১০ ১২ ১৫ ১৮
দূরত্ব =	–২ –১ –১ ০ ০	দূরত্ব =	–৯ –৬ –৩ –১ ০
	০ ১ ১ ২		১ ৩ ৬ ৯
বিচ্যুতির বর্গ =	৪ ১ ১ ০ ০	বিচ্যুতির বর্গ =	৮১ ৩৬ ৯ ১ ০
	০ ১ ১ ৪		১ ৯ ৩৬ ৮১
ভেদাঙ্ক$_{ক}$ =	$\frac{\text{বিচ্যুতির বর্গের সমষ্টি}}{\text{মোট সংখ্যা}}$	ভেদাঙ্ক$_{খ}$ =	$\frac{\text{বিচ্যুতির বর্গের সমষ্টি}}{\text{মোট সংখ্যা}}$
=	$\frac{১২}{৯}$	=	$\frac{২৫৪}{৯}$
=	১.৩৩	=	২৮.২২
পরিমিত ব্যবধান$_{ক}$		পরিমিত ব্যবধান$_{খ}$	
=	$\sqrt{১.৩৩}$	=	$\sqrt{২৮.২২}$
=	১.১৫	=	৫.৩১

উপরের সারণিতে দেখা যাচ্ছে যে, বিন্যাস 'ক'-এর সংখ্যা মানগুলোর মধ্যে দূরত্ব কম বলে বিচ্যুতিগুলো ক্ষুদ্র এবং বিচ্যুতির বর্গগুলোও ক্ষুদ্র। ফলে, বিন্যাস 'ক'-এর ভেদাঙ্ক ও পরিমিত ব্যবধানের মান ক্ষুদ্র হয়েছে। অন্যদিকে, বিন্যাস 'খ'-এর সংখ্যা মানগুলো বেশ দূরত্বের সাথে ছড়িয়ে রয়েছে বলে বিচ্যুতির মানগুলো বড় হয়েছে এবং বিচ্যুতির বর্গগুলোও বড় হয়েছে। ফলে, ভেদাঙ্ক ও পরিমিত ব্যবধানের মান দু'টিও বড় হয়েছে।

কেন্দ্রীয় প্রবণতার পরিমাপগুলো যেমন বিস্তৃতির মান সম্পর্কে আমাদের কোন ধারণা দিতে পারে না, তেমনি পরিমিত ব্যবধান বা ভেদাঙ্কও কেন্দ্রীয় প্রবণতার মান সম্পর্কে কোন ধারণা দিতে পারে না।

সারণি ৫.৫.১-এর পর্যবেক্ষণ আরও একটি বিষয়কে স্পষ্ট করে তোলে। সেটি হলো, কেন্দ্রীয় প্রবণতার পরিমাপগুলো যেমন বিস্তৃতির মান সম্পর্কে আমাদের কোন ধারণা দিতে পারে না, তেমনি পরিমিত ব্যবধান বা ভেদাঙ্কও কেন্দ্রীয় প্রবণতার মান সম্পর্কে কোন ধারণা দিতে পারে না। যেমন, বিন্যাস 'ক'-এর গড় মান (২৩) বড় হলেও ভেদাঙ্ক ও পরিমিত ব্যবধানের মান (যথাক্রমে ১.৩৩ এবং ১.১৫) ছোট হয়েছে, অথচ বিন্যাস 'খ'-এর গড় মান (৯) ছোট হয়েও ভেদাঙ্ক ও পরিমিত ব্যবধানের মান (যথাক্রমে ২৮.২২ ও ৫.৩১) বড় হয়েছে। বিন্যাস 'ক'-এর বড় গড় মান থেকে এটি মনে হতে পারে যে, পরিমিত ব্যবধান বা ভেদাঙ্কের মানও বড় হবে এবং বিস্তৃতিও বড় হবে। বিন্যাস 'খ' থেকে ঠিক এর বিপরীত ধারণাটি হতে পারে। অর্থাৎ, ভেদাঙ্ক ও পরিমিত ব্যবধানের বড় মান দেখে মনে হতে পারে যে,

বিন্যাসটির গড় মান বড় হবে। এখানে দু'টি গুরুত্বপূর্ণ বিষয় স্পষ্ট হয়ে উঠেছে। *প্রথমতঃ*, গড় মানটি বড় বা ছোট হতে পারে কেবলমাত্র সংখ্যার মানগুলোর বড় বা ছোট আকারের কারণে। *দ্বিতীয়তঃ*, ভেদাঙ্ক ও পরিমিত ব্যবধানের বড় বা ছোট মান সংখ্যা মানগুলোর আকারের উপর নির্ভর করে না, করে সংখ্যা মানগুলো একটি আরেকটি থেকে কত দূরত্বে ছড়িয়ে রয়েছে তার উপর। পরিমিত বিন্যাস ও ভেদাঙ্ক ব্যাখ্যার ক্ষেত্রে এই দু'টি বিষয় গুরুত্বের সাথে মনে রাখা প্রয়োজন।

ভেদাঙ্ক ও পরিমিত ব্যবধানের বড় বা ছোট মান সংখ্যা মানগুলোর আকারের উপর নির্ভর করে না, করে সংখ্যা মানগুলো একটি আরেকটি থেকে কত দূরত্বে ছড়িয়ে রয়েছে তার উপর।

একটি বিন্যাসের মধ্যে উপাত্তরাশির মোট সংখ্যা বৃদ্ধি পেলে ভেদাঙ্ক ও পরিমিত ব্যবধানের মান বৃদ্ধি বা হ্রাস পেতে পারে। যতবার একটি করে সংখ্যা বিন্যাসের মধ্যে যুক্ত হবে ততবার সূত্রের 'হর'-এর (denomentor) মান বৃদ্ধি পাবে এবং একই সাথে, 'লব'-এও (numerator) কয়েকটি নতুন বিচ্যুতির বর্গ যুক্ত হবে। উপাত্তরাশির মোট সংখ্যা বাড়লে বিস্তৃতির পরিমাপ দু'টির মান কতটুকু বৃদ্ধি বা হ্রাস পাবে সেটি নির্ভর করবে বিন্যাসে অন্তর্ভুক্ত পূর্বের সংখ্যা মানগুলো থেকে নতুন সংখ্যা মানটি কতটুকু ভিন্ন তার উপর। যদি সংখ্যাটি গড় মানের সমান বা খুব কাছাকাছি হয়, তবে বিচ্যুতির বর্গটি '০' বা তার কাছাকাছি হবে এবং বিচ্যুতির বর্গের গড় মানটি কিছুটা হলেও হ্রাস পাবে এবং নির্ণীত পরিমিত ব্যবধান ও ভেদাঙ্কের মানও হ্রাস পাবে। আর যদি নতুন মানটি গড় থেকে অনেক দুরত্বে অবস্থান করে তবে বিচ্যুতির বর্গটি অস্বাভাবিকভাবে বড় হবে এবং বিচ্যুতির বর্গের গড়ও বৃদ্ধি পাবে। ফলে, পরিমিত ব্যবধান ও ভেদাঙ্কের মান বৃদ্ধি পাবে।

ধরা যাক, বিন্যাস 'ক'-এ দু'টি সংখ্যা যোগ করা হয়েছে: ২০ এবং ২৬। এর ফলে গড় এবং ভেদাঙ্কের মধ্যে কি পরিবর্তন হচ্ছে তা দেখা যাক।

$$\bar{x} = \frac{২০ + ২১ + ২২ + ২২ + ২৩ + ২৩ + ২৩ + ২৪ + ২৪ + ২৫ + ২৬}{১১}$$

$$= \frac{২৫৩}{১১}$$

$$= ২৩$$

$$s^২ = \frac{৯ + ৪ + ১ + ১ + ০ + ০ + ০ + ১ + ১ + ৪ + ৯}{১১}$$

$$= \frac{৩০}{১১}$$

$$= ২.৭৩$$

২০ ও ২৬ সংখ্যা দু'টি যোগ করার ফলে গড় মানের কোন পরিবর্তন না হলেও ভেদাঙ্কের মান ১.৩৩ থেকে বেড়ে গিয়ে ২.৭৩ হয়েছে। এর কারণটি হলো, সংখ্যা দু'টি 'হর' এবং 'লব' দু'টিরই মান বাড়িয়ে দিয়েছে এবং সংখ্যা দু'টি প্রান্তিকভাবে যুক্ত হবার ফলে বিচ্যুতির বর্গগুলো বড় আকার ধারণ করে বিচ্যুতির সমষ্টিকে বেশ কিছুটা বাড়িয়ে দিয়েছে। কিন্তু যদি বিন্যাসটিতে আরো তিনটি সংখ্যা (যেমন, ২২, ২৩ ও ২৪) যুক্ত করা হয় তখন কি হবে? গড় মানের কোন পরিবর্তন ঘটবে না এবং ভেদাঙ্কের মানের পরিবর্তন ঘটলেও তা খুব বেশী বাড়বে না। কারণ, নতুন সংখ্যাগুলো গড় মানের কাছাকাছি হওয়াতে বিচ্যুতির বর্গগুলো ছোট হবে এবং ভেদাঙ্কের মান কিছুটা বৃদ্ধি পাবে। এর অর্থ এই নয় যে, নতুন মান যুক্ত করলে গড় মানের কোন পরিবর্তন হবে না। বেশ বড় বা ছোট একটি সংখ্যা যুক্ত করলে গড় মানেরও পরিবর্তন ঘটবে।

কেন্দ্রীয় প্রবণতার ক্ষেত্রে যেমন গাণিতিক গড় হলো একটি ভারসাম্যের বিন্দু, বিস্তৃতির পরিমাপের ক্ষেত্রে পরিমিত ব্যবধান তেমনি একটি ভারসাম্যের বিন্দু।

এখানে একটি বিষয় উল্লেখ করা প্রয়োজন যে, এ ধরণের একটি ছোট বিন্যাসের ক্ষেত্রে বিচ্যুতির বর্গের বিষয়টি তাৎক্ষণিকভাবে প্রমাণ করা গেলেও বৃহৎ উপাত্তরাশি সম্বলিত বিন্যাসের ক্ষেত্রে তা সম্ভব নয়। সে ক্ষেত্রে একটি পরিমিত সূচক প্রয়োজন যার বিপরীতে নতুন সংখ্যা মান যুক্ত করার প্রভাবটি দেখে বিন্যাসের বিস্তৃতির প্রকৃতিটি বোঝা যেতে পারে। পরিমিত ব্যবধান এই পরিমিত সূচকের কাজটি করে। কেন্দ্রীয় প্রবণতার ক্ষেত্রে যেমন গাণিতিক গড় হলো একটি ভারসাম্যের বিন্দু (point of balance), বিস্তৃতির পরিমাপের ক্ষেত্রে পরিমিত ব্যবধান তেমনি একটি ভারসাম্যের বিন্দু। যখন কোন বিন্যাসে নতুন সংখ্যা মান যুক্ত করা হয় এবং তা যদি এক পরিমিত ব্যবধানের নীচে হয়, তবে ভেদাঙ্কের মান কমবে। কারণ, আমরা জানি যে বিচ্যুতির বর্গগুলো ছোট হয় বলে তার গড়টিও ছোট হয়। বিষয়টি একটি উদাহরণ দিয়ে বোঝা যাক। ধরা যাক, ১০০টি সংখ্যা রাশির গড় ৫০ এবং পরিমিত ব্যবধান ১০। যদি এই ১০০টি সংখ্যা মানের সাথে আরো ৫টি ৪৮ মানের এবং ৫টি ৫২ মানের সংখ্যা যোগ করি, তাহলে ফলাফলটি কি দাঁড়ায় দেখা যাক। যেহেতু গড় মান হলো ৫০, সেহেতু ৪৮ সংখ্যামানের বিচ্যুতি হবে – ২ এবং ৫২ সংখ্যা মানের বিচ্যুতির হবে +২। এই বিচ্যুতিগুলোর বর্গ করলে হবে ৪ এবং ৪। অতএব, এই নতুন ১০টি সংখ্যার বিচ্যুতির সমষ্টি হবে (৫×৪)+(৫×৪)=২০+২০=৪০।

যেহেতু পরিমিত ব্যবধান হলো ১০, সেহেতু ভেদাঙ্ক অবশ্যই হবে ১০০ (কারণ, ভেদাঙ্ক =পরিমিত ব্যবধানের বর্গ)। ভেদাঙ্কের উপর এই নতুন মানগুলোর প্রভাব কি হবে? উত্তরে বলা যায় যে, ভেদাঙ্কের মান কমবে। কারণ, যে মানগুলো যুক্ত হয়েছে (৪৮ এবং ৫২) সেগুলো পরিমিত ব্যবধানের মানের নীচে রয়েছে। অর্থাৎ, পরিমিত ব্যবধান = ১০-এর অর্থ হলো বিন্যাসের মানগুলো গড় থেকে গড়ে এক পরিমিত ব্যবধান ছড়িয়ে রয়েছে। আমাদের মূল বিন্যাসের ভেদাঙ্ক হলো ১০০। এবার দেখা যাক, নতুন মানগুলো যুক্ত করলে ভেদাঙ্কের মান কি দাঁড়ায়?

মূল বিন্যাসের বিচ্যুতির বর্গের সমষ্টি হলো,

$$s^{২} = \frac{\sum(x_i - \bar{x})^{২}}{১০০} = ১০০$$

অতএব, $\sum(x_i - \bar{x})^{২} = ১০০ \times ১০০$

$= ১০০০০$

যেখানে, $N = ১০০$

$s^{২} = ১০০$

নতুন বিন্যাসের বিচ্যুতির বর্গের সমষ্টি হলো,

$$s^{২} = \frac{\text{মূল বর্গের সমষ্টি} + \text{নতুন বর্গের সমষ্টি}}{\text{মূল মোট সংখ্যা} + \text{নতুন সংখ্যা}}$$

$$= \frac{১০০০০ + ৪০}{১০০ + ১০}$$

$$= \frac{১০০৪০}{১১০}$$

$$= ৯১.৩$$

$$\therefore \quad s^{২} = ৯১.৩$$

অতএব, প্রমাণিত হলো যে, ১০টি নতুন সংখ্যার মান যুক্ত হওয়াতে মূল ভেদাঙ্কের মান ১০০ হ্রাস পেয়ে ৯১.৩ হয়েছে। এবার যদি ৫টি ৩৫ সংখ্যা মান এবং ৫টি ৬৫ সংখ্যা মান যুক্ত করি তবে দেখবো যে,

ভেদাঙ্কের মান বৃদ্ধি পেয়েছে। কারণ, নতুন সংখ্যাগুলো গড় থেকে ১৫ কম এবং ১৫ বেশী হওয়াতে প্রতিটি বিচ্যুতির বর্গ হয়েছে ২২৫। ১০টি মানের বিচ্যুতির বর্গের সমষ্টি হবে ১০ × ২২৫ =২২৫০। এই নতুন বিচ্যুতির বর্গের সমষ্টিটি মূল বিন্যাসের বিচ্যুতির বর্গের সমষ্টির সাথে যোগ করলে দেখা যাবে যে ভেদাঙ্কের মান বৃদ্ধি পেয়েছে। যেমন, নতুন বিন্যাসের ভেদাঙ্ক হলো,

$$s^{২} = \frac{\text{মূল বর্গের সমষ্টি + নতুন বর্গের সমষ্টি}}{\text{মূল মোট সংখ্যা + নতুন সংখ্যা}}$$

$$= \frac{১০০০০ + ২২৫০}{১০০ + ১০}$$

$$= \frac{১২২৫০}{১১০}$$

$$= ১১১.৪$$

$$\therefore \quad s^{২} = ১১১.৪$$

অতএব, দেখা যাচ্ছে যে, গড় থেকে এক পরিমিত ব্যবধানের চেয়ে বড় বা ছোট সংখ্যা যোগ করলে ভেদাঙ্কের মান বৃদ্ধি পায়। আর ভেদাঙ্কের মান বৃদ্ধি পাবার অর্থ হলো পরিমিত ব্যবধানের মান বৃদ্ধি পাওয়া। তবে এটি মনে রাখা প্রয়োজন যে, পরিমিত ব্যবধানকে পরিমিত মানে (standard score) প্রকাশিত করার পূর্বে এটিকে যথাযথভাবে ব্যাখ্যা করা যায় না এবং এর ষজ্ঞাগত অর্থটি অনুধাবন করা যায় না।

সারাংশ

পরিমিত ব্যবধান এবং ভেদাঙ্ক হলো ঘনিষ্টভাবে সম্পর্কিত বিস্তৃতির দু'টি পরিমাপ। এই দু'টি পরিমাপকে বিস্তৃতির প্রধান পরিমাপ হিসাবে অধিকাংশ গবেষণায় ব্যবহার করা হয়। শুধু বিস্তৃতির পরিমাপ হিসাবে নয়, অন্যান্য উচ্চতর পরিসংখ্যানিক ও বীজগাণিতিক পরিগণনার ক্ষেত্রে এই দু'টি পরিমাপের অন্তর্নিহিত ক্ষুদ্রতম বর্গের বৈশিষ্ট্যটি গুরুত্বপূর্ণ ভূমিকা পালন করে। পরিমিত ব্যবধানের সাথে ভেদাঙ্কের পার্থক্য হলো পরিমিত ব্যবধান বিচ্যুতির গড়ের বর্গমূল নির্ণয়ের প্রয়োজন হয়, ভেদাঙ্কে তার প্রয়োজন হয় না।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক (√) চিহ্ন দিন –

১। ..... লৈখিকভাবে উপস্থাপন করা যায় না।

ক. পরিমিত ব্যবধান
খ. গড় ব্যবধান
গ. প্রমিত ব্যবধান
ঘ. ভেদাঙ্ক

২। একটি বিন্যাসকে সমরূপ বলা যায় যখন:

ক. পরিমিত ব্যবধান ছোট এবং ভেদাঙ্ক বড় হয়
খ. পরিমিত ব্যবধান বড় এবং ভেদাঙ্ক ছোট হয়
গ. পরিমিত ব্যবধান ও ভেদাঙ্কের মান ছোট হয়
ঘ. পরিমিত ব্যবধান ও ভেদাঙ্কের মান বড় হয়

৩। ভেদাঙ্কের মান বৃদ্ধি বা হ্রাস পেতে পারে যদি:

ক. উপাত্তরাশির মোট সংখ্যা বৃদ্ধি পায়
খ. উপাত্তরাশির মোট সংখ্যা হ্রাস পায়
গ. উপাত্তরাশির মোট সংখ্যা হ্রাস বা বৃদ্ধি পায়
ঘ. উপাত্তরাশির মোট সংখ্যা সমান থাকে

সংক্ষিপ্ত প্রশ্ন

১। পরিমিত ব্যবধান ও ভেদাঙ্কের মধ্যে পার্থক্য কি?

২। মানের আপেক্ষিক অবস্থান নির্ণয়ে পরিমিত ব্যবধানের ব্যবহার আলোচনা করুন।

রচনামূলক প্রশ্ন

১। উপাত্ত বর্ণনায় পরিমিত ব্যব্যধান ও ভেদাঙ্কের মধ্যে কোনটিকে আপনি অধিক যুক্তিযুক্ত বলে মনে করেন? আপনার উত্তরের স্বপক্ষে যুক্তি দিন।

২। উপযুক্ত উদাহরণসহ পরিমিত ব্যবধান ও ভেদাঙ্কেও ব্যাখ্যা দিন।

পাঠ - ৭

বিস্তৃতির আপেক্ষিক পরিমাপ – ব্যবধানাঙ্ক এবং পরিমিত মান

Relative Measures of Dispersion – Coefficient of Variation and Standard Score

এই পাঠ শেষে যা জানা যাবে —

- বিস্তৃতির আপেক্ষিক পরিমাপ
- ব্যবধানাঙ্ক
- ব্যবধানাঙ্ক নির্ণয়
- পরিমিত মান
- পরিমিত মানের বৈশিষ্ট্য
- পরিমিত মান নির্ণয়
- পরিমিত মানের সুবিধা

বিস্তৃতির আপেক্ষিক পরিমাপ (Relative Measures of Dispersion)

একটি নিবেশনের গড় থেকে সংখ্যা মানগুলোর বিস্তৃতির পরিমাপের জন্য নিরঙ্কুশ বিস্তার পরিমাপ (absolute measures of dispersion) উপযোগী হলেও দুই বা ততোধিক উপাত্তসারির মধ্যে তারতম্যের ভিত্তিতে তুলনা করার জন্য উপযোগী নয়। নিরঙ্কুশ বিস্তার পরিমাপগুলো উপাত্তের মধ্যে ভিন্ন ভিন্ন পরিমাপ এককের (যেমন, টাকা, বছর, কিলোগ্রাম, ইত্যাদি) মাধ্যমে ভিন্নতার বর্ণনা দেয়। দু'টি উপাত্ত বিন্যাসের পরিমাপ একক যদি ভিন্ন ভিন্ন হয় তবে নিরঙ্কুশ পরিমাপের মাধ্যমে ভিন্নতার তুলনা করা যায় না। অতএব, ভিন্ন পরিমাপ একক সম্বলিত দু'টি উপাত্তরাশির মধ্যে ভিন্নতার তুলনার জন্য নিরঙ্কুশ ব্যবধানগুলোকে আপেক্ষিক পদে রূপান্তরিত করার প্রয়োজন হয়। আপেক্ষিক বিস্তার পরিমাপগুলো বিস্তৃতির নিরঙ্কুশ পরিমাপগুলোর পরিবর্তিত রূপ মাত্র। বিস্তৃতির আপেক্ষিক পরিমাপগুলো কোন এককে প্রকাশিত হয় না। সেগুলো একক বিহীন বলে তুলনার ক্ষেত্রে বিশেষভাবে উপযোগী। আপেক্ষিক বিস্তার পরিমাপের একাধিক পদ্ধতি রয়েছে। যেমন,

১. চতুর্থক ব্যবধানাঙ্ক (Coefficient of Quartile Deviation)

২. গড় ব্যবধানাঙ্ক (Coefficient of Mean Deviation)

৩. ব্যবধানাঙ্ক (Coefficient of Variation)

এর মধ্যে সবচেয়ে গুরুত্বপূর্ণ এবং সর্বাধিক ব্যবহৃত পদ্ধতি হলো ব্যবধানাঙ্ক। সেটি আমরা এখন আলোচনা করবো।

ব্যবধানাঙ্ক (Coefficient of Variation)

কোন উপাত্তরাশির গড় এবং পরিমিত ব্যবধানের শতকরা অনুপাতকে সেই উপাত্তরাশির ব্যবধানাঙ্ক বলা হয়। অর্থাৎ,

কোন উপাত্তরাশির গড় এবং পরিমিত ব্যবধানের শতকরা অনুপাতকে সেই উপাত্তরাশির ব্যবধানাঙ্ক বলা হয়।

$$C.V. = \frac{s}{\bar{x}} \times ১০০$$

যেখানে, C.V. = ব্যবধানাঙ্ক

s = পরিমিত ব্যবধান

$\bar{x}$ = গাণিতিক গড়

যখন দু'টি উপাত্তরাশির মধ্যে বিস্তার পরিমাপের জন্য ব্যবধানাঙ্ক ব্যবহৃত হয়, তখন যে উপাত্তরাশির ব্যবধানাঙ্কের মান বড় হয়, সেই উপাত্তরাশির বিন্যাসটিকে অধিক বিস্তৃত বা অসমরূপ বলে গণ্য করা হয়। অন্যদিকে, যে উপাত্তরাশির ব্যবধানাঙ্কের মান কম হবে সেই উপাত্তরাশির বিন্যাসটিকে কম বিস্তৃত এবং অধিক সমরূপ বলে গণ্য করা হয়। বিস্তৃতির আপেক্ষিক পরিমাপ নির্ণয়ের জন্য বিস্তৃতির পরিমাপের যে কোন নিরঙ্কুশ পদ্ধতি (পরিসর, চতুর্থক ব্যবধান, গড় ব্যবধান ও পরিমিত ব্যবধান) সংশ্লিষ্ট কেন্দ্রীয় প্রবণতার পরিমাপের (মধ্যমা, গাণিতিক গড়) সাহায্যে ব্যবহার করা যায়। কিন্তু পরিসংখ্যানবিদগণ বিস্তৃতির আপেক্ষিক পরিমাপের ক্ষেত্রে প্রায় সবসময়ই পরিমিত ব্যবধান ও গাণিতিক গড় ব্যবহার করে থাকেন।

ব্যবধানাঙ্ক নির্ণয় (Computing Coefficient of Variation)

আমরা জানি যে, কোন উপাত্তরাশির গড় ও পরিমিত ব্যবধানের শতকরা অনুপাতকে ব্যবধানাঙ্ক বলে। সুতরাং, ব্যবধানাঙ্ক নির্ণয়ের পূর্বে উপাত্তরাশির গড় ও পরিমিত ব্যবধান নির্ণয় করা আবশ্যক। সারণি ৫.৬.১-এ প্রদত্ত উপাত্ত ব্যবহার করে ব্যবধানাঙ্ক নির্ণয়ের মাধ্যমে দু'টি উপাত্তরাশির মধ্যে বিদ্যমান বিস্তৃতির তুলনা করা হলো।

সারণি ৫.৬.১: দু'টি গ্রামের মহিলাদের মাসিক আয়ের বিন্যাস

মাসিক আয় (টাকায়)	*গ্রাম 'ক' (মহিলার সংখ্যা)*	*গ্রাম 'খ' (মহিলার সংখ্যা)*
০ - ২০০০	৪২	২২
২০০০ - ৪০০০	২৩	৫৬
৪০০০ - ৬০০০	১৫	১২
৬০০০ - ৮০০০	১০	৪
৮০০০ - ১০০০০	৮	৬
১০০০০ - ১২০০০	১	০
১২০০০ - ১৪০০০	১	০
মোট	*১০০*	*১০০*

সূত্র অনুযায়ী, প্রথমে গ্রাম 'ক'-এর ব্যাবধানাঙ্ক নির্ণয়ের জন্য প্রয়োজনীয় মানগুলো (অর্থাৎ, গাণিতিক গড় ও পরিমিত ব্যবধান) সারণি ৫.৬.২-এ নির্ণয় করে দেখানো হলো।

সারণি ৫.৬.২: গ্রাম 'ক'-এর মহিলাদের মাসিক আয়ের গড় ও পরিমিত ব্যবধান নির্ণয়

মাসিক আয় (টাকা)	*মহিলার সংখ্যা* f_i	*শ্রেণী মধ্য-বিন্দু* x_i	$d = \frac{x_i - A}{C.I.}$	$d^২$	fd	$fd^২$
০ - ২০০০	৪২	১০০০	–৩	৯	–১২৬	৩৭৮
২০০০ - ৪০০০	২৩	৩০০০	–২	৪	–৪৬	৯২
৪০০০ - ৬০০০	১৫	৫০০০	–১	১	–১৫	১৫
৬০০০ - ৮০০০	১০	৭০০০ = A	০	০	০	০
৮০০০ - ১০০০০	৮	৯০০০	১	১	৮	৮
১০০০০ - ১২০০০	১	১১০০০	২	৪	২	৪
১২০০০ - ১৪০০০	১	১৩০০০	৩	৯	৩	৯
মোট	*১০০*				*–১৭৪*	*৫০৬*

$$\bar{x}_{ক} = A + \frac{\sum fd}{N} \times C.I.$$

$$= ৭০০০ + \left(\frac{-১৭৪}{১০০}\right) \times ২০০০$$

$$= ৭০০০ + (-১.৭৪) \times ২০০০$$

$$= ৭০০০ - ৩৪৮০$$

$$= ৩৫২০ \text{ টাকা}$$

অর্থাৎ, গ্রাম 'ক'-এর মহিলাদের মাসিক আয়ের গড় হলো ৩৫২০ টাকা।

এবার আমাদের গ্রাম 'ক'-এর মহিলাদের আয়ের পরিমিত ব্যবধান নির্ণয় করতে হবে।

$$\therefore \quad s_{ক} = C.I.\sqrt{\frac{\sum fd^২}{N} - \left(\frac{\sum fd}{N}\right)^২}$$

$$= ২০০০\sqrt{\frac{৫০৬}{১০০} - \left(\frac{-১৭৪}{১০০}\right)^২}$$

$$= ২০০০\sqrt{৫.০৬ - (-১.৭৪)^২}$$

$$= ২০০০\sqrt{৫.০৬ - ৩.০৩}$$

$$= ২০০০ \times \sqrt{২.০৩}$$

$$= ২০০০ \times ১.৪২$$

$$= ২৮৪০$$

অর্থাৎ, গ্রাম 'ক'-এর মহিলাদের আয়ের পরিমিত ব্যবধান ২৮০০ টাকা

অতএব, গ্রাম 'ক'-এর মহিলাদের আয়ের ব্যবধানাঙ্ক হলো,

$$C.V_{ক} = \frac{s}{\bar{x}} \times ১০০$$

$$= \frac{২৮৪০}{৩৫২০} \times ১০০$$

$$= ৮০.৬৮$$

অর্থাৎ, গ্রাম 'ক'-এর মহিলাদের আয়ের ব্যবধানাঙ্ক হলো ৮০.৬৮।

সারণি ৫.৬.৩-এ গ্রাম 'খ'-এর মহিলাদের আয়ের বিন্যাসের গড় ও পরিমিত ব্যবধান নির্ণয় করে দেখানো হলো।

সারণি ৫.৬.৩: গ্রাম 'খ'-এর মহিলাদের আয়ের বিন্যাসের গড় ও পরিমিত ব্যবধান নির্ণয়

মাসিক আয় (টাকা)	*মহিলার সংখ্যা*	*শ্রেণী মধ্য-বিন্দু* x_i	$d = \frac{x_i - A}{C.I.}$	$d^২$	fd	$fd^২$
০ - ২০০০	২২	১০০০	–৩	৯	–৬৬	১৯৮
২০০০ - ৪০০০	৫৬	৩০০০	–২	৪	–১১২	২২৪
৪০০০ - ৬০০০	১২	৫০০০	–১	১	–১২	১২
৬০০০ - ৮০০০	৪	৭০০০ = A	০	০	০	০
৮০০০ - ১০০০০	৬	৯০০০	১	১	৬	৬
১০০০০ - ১২০০০	০	১১০০০	২	৪	০	০
১২০০০ - ১৪০০০	০	১৩০০০	৩	৯	০	০
মোট	*১০০*				*–১৮৪*	*৪৪০*

$$\therefore \quad \bar{x}_{খ} = A + \frac{\sum fd}{N} \times C.I.$$

$$= ৭০০০ + \frac{-১৮৪}{১০০} \times ২০০০$$

$$= ৭০০০ + (-১.৮৪) \times ২০০০$$

$$= ৭০০০ - ৩৬৮০$$

$$= ৩৩২০$$

অর্থাৎ গ্রাম 'খ' এর মহিলাদের আয়ের গড় হলো ৩৩২০ টাকা।

এবার আমাদের গ্রাম 'খ'-এর মহিলাদের আয়ের পরিমিত ব্যবধান নির্ণয় করতে হবে।

$$\therefore \quad s_{খ} = C.I.\sqrt{\frac{\sum fd^২}{N} - \left(\frac{\sum fd}{N}\right)^২}$$

$$= ২০০০\sqrt{\frac{৪৪০}{১০০} - \left(\frac{-১৮৪}{১০০}\right)^২}$$

$$= \quad ২০০০\sqrt{৪.৪-(-১.৮৪)^{২}}$$

$$= \quad ২০০০\sqrt{৪.৪-৩.৩৯}$$

$$= \quad ২০০০\times\sqrt{১.০১}$$

$$= \quad ২০০০ \times ১.০০৫$$

$$= \quad ২০১০$$

অর্থাৎ, গ্রাম 'খ'-এর মহিলাদের আয়ের পরিমিত ব্যবধান হলো ২০১০ টাকা।

অতএব, গ্রাম 'খ' এর মহিলাদের আয়ের ব্যবধানাঙ্ক হলো,

$$C.V_{খ} = \frac{s}{\bar{x}}\times ১০০$$

$$= \frac{২০১০}{৩৩২০}\times ১০০$$

$$= ৬০.৫৪$$

অর্থাৎ, গ্রাম 'খ'-এর মহিলাদের আয়ের ব্যবধানাঙ্ক হলো ৬০.৫৪।

যেহেতু গ্রাম 'ক'-এর মহিলাদের মাসিক আয়ের ব্যবধানাঙ্ক (৮০.৬৮) গ্রাম 'খ'-এর মহিলাদের মাসিক আয়ের ব্যবধানাঙ্ক (৬০.৫৪) অপেক্ষা বেশী, সেহেতু উপসংহারে আমরা বলতে পারি যে, গ্রাম 'ক'-এর মহিলাদের মাসিক আয়ের বিন্যাস গ্রাম 'খ'-এর মহিলাদের মাসিক আয়ের বিন্যাস অপেক্ষা বেশী বিক্ষিপ্ত বা বিস্তৃত।

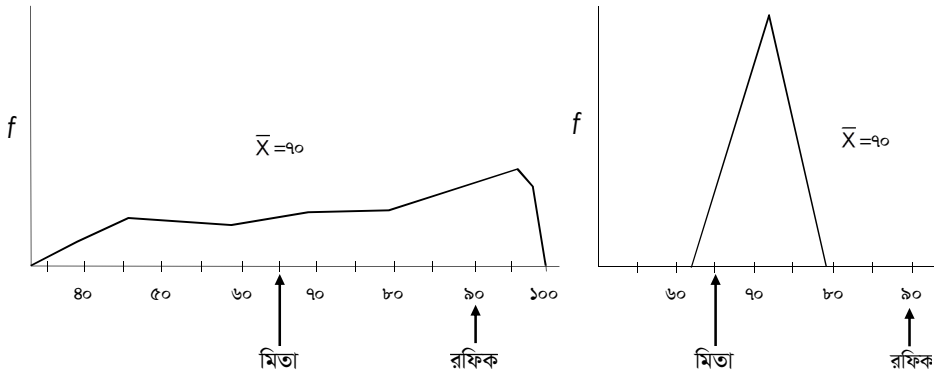
পরিমিত মান (Standard Score)

পরিমিত ব্যবধান আমাদেরকে গড় থেকে সংখ্যা মানগুলোর গড় বিস্তৃতি সম্পর্কে ধারণা দেয়। কিন্তু এছাড়াও কোন উপাত্তরাশির কোন বিশেষ একটি মানের অবস্থান আমাদের জানার প্রয়োজন হতে পারে। পরিমিত মান নির্ণয়ের মাধ্যমে আমরা তা জানতে পারি। শতহারিকের মাধ্যমেও আমরা আপেক্ষিক অবস্থান জানতে পারি। তবে শতহারিকের সীমাবদ্ধতাটি হলো যে, সেগুলোকে যোগ করা যায় না, গুণ করা যায় না, গড় করা যায় না, এবং জটিল পরিগণনার জন্য সম্মিলিত করা যায় না। এক গুচ্ছ পর্যবেক্ষণের মধ্যে একটি পর্যবেক্ষণ কিভাবে এবং কতটুকু ভিন্ন তা নির্দেশের জন্য পরিমিত মান একটি বিকল্প পদ্ধতি প্রদান করে। পরিমিত মানকে গড় করা যায় এবং জটিল পরিসংখ্যানিক পরিগণনায়ও ব্যবহার করা যায়। পরিমিত মানকে z-মানও বলা হয়ে থাকে।

পরিমিত ব্যবধান আমাদেরকে গড় থেকে সংখ্যা মানগুলোর গড় বিস্তৃতি সম্পর্কে ধারণা দেয়। কিন্তু এছাড়াও কোন উপাত্তরাশির কোন বিশেষ একটি মানের অবস্থান আমাদের জানার প্রয়োজন হতে পারে। পরিমিত মান নির্ণয়ের মাধ্যমে আমরা তা জানতে পারি।

একটি উদাহরণ দেয়া যাক। ধরা যাক, কোন একটি শ্রেণীতে দু'টি পরীক্ষা নেয়া হলো। একটি মধ্যবর্তী পরীক্ষা এবং অপরটি চূড়ান্ত পরীক্ষা। দু'টি পরীক্ষার ফলাফলই ছাত্র-ছাত্রীদের গ্রেড নির্ধারণে সমান গুরুত্ব বহন করবে। দেখা গেল, উভয় পরীক্ষায় ছাত্র-ছাত্রীদের প্রাপ্ত নম্বরের গড় ৬০। মধ্যবর্তী পরীক্ষায় মিতা ৮০ নম্বর পেয়েছে এবং চূড়ান্ত পরীক্ষায় সে পেয়েছে ৫৫ নম্বর। অন্যদিকে, রফিক মধ্যবর্তী পরীক্ষায় ৫৫ এবং চূড়ান্ত পরীক্ষায় ৮০ নম্বর পেয়েছে। আমরা কি বলতে পারি যে, রফিক ও মিতা সর্বশেষে একই রকম ফলাফল করেছে? তারা দু'জনেই কি একই গ্রেড পাবে?

যদি নিরঙ্কুশ সংখ্যা মানের উপর ভিত্তি করে পরীক্ষার ফলাফল বিশ্লেষণ করা হয়, তাহলে উত্তর হবে 'হ্যাঁ'। কারণ, তারা দু'টি পরীক্ষায় সমান নম্বর পেয়েছে। অর্থাৎ, উভয়েই মোট ১৩৫ নম্বর পেয়েছে এবং উভয়েই দু'টি পরীক্ষায় গড়ে ৬৭.৫ নম্বর পেয়েছে। কিন্তু, যদি পুরো শ্রেণীর প্রেক্ষিতে তাদের আপেক্ষিক কৃতিত্বকে বিবেচনায় আনা হয়, তাহলে উত্তরটি হবে 'না'। দেখা যাবে যে, পরীক্ষায় তাদের দু'জনের অর্জন সমান নয়। নিম্নের লেখচিত্রের মাধ্যমে দু'টি পরীক্ষায় ছাত্র-ছাত্রীদের প্রাপ্ত নম্বরের গণসংখ্যা বিন্যাসের দিকে লক্ষ্য করা যাক।



মধ্যবর্তী পরীক্ষায় প্রাপ্ত নম্বর **চূড়ান্ড় পরীক্ষায় প্রাপ্ত নম্বর**

চিত্র ৫.৬.১: দু'টি পরীক্ষায় ছাত্র-ছাত্রীদের প্রাপ্ত নম্বরের বিন্যাস

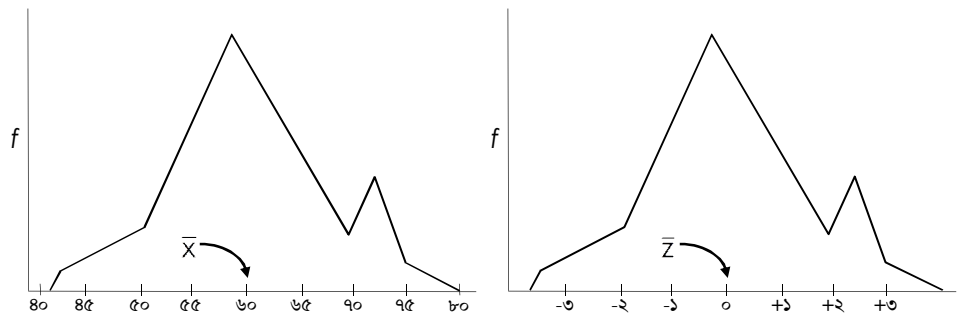
উপরের চিত্রে দেখা যায় যে, উভয়েই একটি পরীক্ষায় গড় থেকে ২০ নম্বর বেশী এবং অপর পরীক্ষায় গড় থেকে ৫ নম্বর কম পেয়েছে। কিন্তু পুরো শ্রেণীর প্রেক্ষিতে দেখা যায় যে, মধ্যবর্তী পরীক্ষায় মিতার প্রাপ্ত ৮০ নম্বরের তুলনায় রফিকের পরীক্ষার ৮০ নম্বর ভালো অবস্থান প্রদর্শন করে। কারণটি হলো, মধ্যবর্তী পরীক্ষার নম্বরগুলো অনেক বেশী বিস্তৃতভাবে ছড়িয়ে রয়েছে। অর্থাৎ, দু'টি বিন্যাসের বিস্তৃতি দু'রকম এবং ফলে দু'টি বিন্যাসের মধ্যে পরিমিত ব্যবধান ভিন্ন হয়েছে। মধ্যবর্তী পরীক্ষার পরিমিত ব্যবধান ২০ এবং চূড়ান্ত পরীক্ষার পরিমিত ব্যবধান ৫। বিস্তৃতির এই ভিন্নতাটিই প্রতিটি মানের আপেক্ষিক অবস্থান নির্ণয়ের জন্য নির্ধারক হিসাবে কাজ করে।

এ রকম পরিস্থিতিতে চূড়ান্ত পরীক্ষায় গড় থেকে ২০ নম্বর বেশী পাওয়াটি যে মধ্যবর্তী পরীক্ষার গড় থেকে ২০ নম্বর পাওয়ার চেয়ে অধিকতর শ্রেয় তা নির্ণয় করার জন্য আমাদের একটি উপায় বের করা প্রয়োজন। প্রতিটি মানের আপেক্ষিক অবস্থান নির্ণয়ের শ্রেষ্ঠ পথটি হলো সংশ্লিষ্ঠ বিন্যাসগুলোর পরিমিত ব্যবধানকে বিবেচনায় এনে প্রতিটি অশোধিত মানের জন্য পরিমিত মান নির্ণয় করা।

পরিমিত মানের বৈশিষ্ট্য (Characteristics of Standard Scores)

পরিমিত মানের কিছু সুনির্দিষ্ট বৈশিষ্ট্য রয়েছে। পরিমিত মানগুলো হলো মৌলিক সংখ্যা। কারণ, অশোধিত সংখ্যাগুলো পরিমিত সংখ্যায় রূপান্তরের সময় তাদের পরিমাপের এককগুলো বাতিল হয়ে যায়। অর্থাৎ, পরিমিত মান হলো একক বিহীন স্বাধীন সংখ্যা।

পরিমিত মানের কিছু সুনির্দিষ্ট বৈশিষ্ট্য রয়েছে। পরিমিত মানগুলো হলো মৌলিক সংখ্যা। কারণ, অশোধিত সংখ্যাগুলো পরিমিত সংখ্যায় রূপান্তরের সময় তাদের পরিমাপের এককগুলো বাতিল হয়ে যায়। অর্থাৎ, পরিমিত মান হলো একক বিহীন স্বাধীন সংখ্যা। যদি বিন্যাসের প্রতিটি মানের জন্য পৃথকভাবে পরিমিত মান নির্ণয় করা হয় এবং সেগুলোর একটি গণসংখ্যা বিন্যাস লেখচিত্রে উপস্থাপন করা হয়, তবে দেখা যাবে যে, পরিমিত মানগুলোর লেখচিত্রায়িত বিন্যাস অশোধিত মূল মানের চিত্রায়িত বিন্যাসের অনুরূপ হবে। লেখচিত্রের অনুভূমিক অক্ষে উপস্থাপিত অশোধিত মূল সংখ্যা মানগুলোর পরিবর্তে শুধুমাত্র পরিমিত মানের মানগুলো উপস্থাপিত হবে। নিম্নের লেখচিত্র দু'টির মাধ্যমে অশোধিত মূল সংখ্যা মান ও পরিমিত সংখ্যা মানের এই সম্পর্ক স্পষ্ট হয়ে উঠে।



(ক) অশোধিত মূল মানের বিন্যাস (খ) রূপান্তরিত পরিমিত মানের বিন্যাস

চিত্র ৫.৬.২: অশোধিত মূল মানের বিন্যাস (ক) এবং রূপান্তরিত পরিমিত মানের বিন্যাস (খ)

উপরোক্ত চিত্র দু'টি থেকে আমরা পরিমিত মানের আরো কিছু গুরুত্বপূর্ণ বৈশিষ্ট্য দেখতে পাই। যেমন, যে কোন বিন্যাসের জন্য পরিমিত মানসমূহের গড় ($\bar{z}$) সবসময়ই '০' হবে এবং পরিমিত ব্যবধান হবে ১। যখন বিন্যাসটি অপেক্ষাকৃত মসৃণ বা সুষম (smooth or symmetrical) হয়, তখন অধিকাংশ পরিমিত সংখ্যা মানই –৩ থেকে +৩ পরিমিত ব্যবধানের মধ্যেই বিরাজ করে। অন্য কথায়, সুষম বিন্যাসের অধিকাংশ মানই গড় থেকে ৩ পরিমিত ব্যবধান উপরে বা নীচে অবস্থান করে। তবে এটি মনে রাখা প্রয়োজন যে, এই বৈশিষ্ট্যটি কোন লৌহ কঠিন নিয়ম নয়। কারণ, অনেক সময় –৩-এর চেয়ে অনেক নীচে এবং +৩-এর অনেক উপরেও পরিমিত মান অবস্থান করে। পরিমিত মানের আরেকটি আকর্ষণীয় বৈশিষ্ট্য হলো, বর্গকৃত পরিমিত মানের সমষ্টি বিন্যাসের মোট সংখ্যা মানের সমান। অর্থাৎ, $\sum z^2 = N$ ।

আমরা আরও জানি যে, একটি সুষম বিন্যাসের মধ্যে গড়ের উপরে এবং নীচে ৩টি করে পরিমিত ব্যবধান রয়েছে। প্রতিটি পরিমিত ব্যবধানের অভ্যন্তরে কি পরিমাণ ঘটনসংখ্যা রয়েছে পরিমিত মানের মাধ্যমে তার আয়তন বের করা যায়। যেমন, $\pm$১ পরিমিত ব্যবধানের অধীনে ৬৮.২৭ শতাংশ, $\pm$২ পরিমিত ব্যবধানের অধীনে ৯৫.৪৫ শতাংশ, এবং $\pm$৩ পরিমিত ব্যবধানের অধীনে ৯৯.৭৪ শতাংশ ঘটনসংখ্যা রয়েছে। আরো সুনির্দিষ্টভাবে বলা যায় যে, একটি নিখুঁত সুষম রেখার অধীনে মোট আয়তন হলো ১ (চিত্র ৫.৪.১ দ্রষ্টব্য)।

পরিমিত মান নির্ণয় (Computing Standard Scores)

একটি উপাত্তরাশির যে কোন একটি অশোধিত মানের (x) পরিমিত মান হলো উক্ত অশোধিত মান থেকে বিন্যাসের গড়ের পার্থক্যকে $(x_i - \bar{x})$ উপাত্তরাশির পরিমিত ব্যবধান (s) দিয়ে ভাগ করলে যে সংখ্যা মানটি পাওয়া যায় তা। একটি বিন্যাসের প্রতিটি সংখ্যা মানেরই একটি পরিমিত মান রয়েছে, যা সেই বিন্যাসের গড় ও পরিমিত ব্যবধানের মানের উপর নির্ভর করে। এখন আমাদের উদাহরণটি সূত্রে প্রয়োগ করে মিতা ও রফিকের আপেক্ষিক অবস্থানটি নির্ণয় করে দেখবো। পরিমিত মান নির্ণয়ের বীজগাণিতিক সূত্রটি হলো,

$$z = \frac{x_i - \bar{x}}{s}$$

যেখানে, z = পরিমিত মান

x_i = প্রতিটি অশোধিত মান

$\bar{x}$ = গাণিতিক গড়

s = পরিমিত ব্যবধান

আমরা জেনেছি যে, উপাত্তরাশির একটি নির্দিষ্ট সংখ্যা মান গড় থেকে কত পরিমিত ব্যবধান উপরে বা নীচে অবস্থান করছে তা পরিমিত মানের মাধ্যমে জানা যায়। যদি পরীক্ষার নম্বরের গড় ৬০ এবং পরিমিত ব্যবধান ২০ হয়, তবে পরীক্ষায় প্রাপ্ত ৮০ নম্বর গড় থেকে এক পরিমিত ব্যবধান উপরে অবস্থান করবে। অপর দিকে, যদি পরীক্ষায় প্রাপ্ত নম্বরের গড় ৬০ এবং পরিমিত ব্যবধান ৫ হয়, তবে পরীক্ষায় প্রাপ্ত ৮০ নম্বর গড় থেকে চার পরিমিত ব্যবধান উপরে অবস্থান করবে। সারণি ৫.৬.১-এ আমাদের উদহারণের সংখ্যাগুলো ব্যবহার করে মিতা ও রফিকের দু'টি পরীক্ষায় প্রাপ্ত নম্বরের পরিমিত মান নির্ণয় করে দেখানো হলো।

সারণি ৫.৬.১: মিতা ও রফিকের পরীক্ষায় প্রাপ্ত অশোধিত মানের পরিমিত মান নির্ণয়

	মধ্যবর্তী পরীক্ষা $\bar{x}$ = ৬০, s = ২০	***চূড়ান্ত পরীক্ষা*** $\bar{x}$ = ৬০, s = ৫
মিতা	অশোধিত মান = ৮০ $z = \frac{৮০ - ৬০}{২০} = +১.০$	অশোধিত মান = ৫৫ $z = \frac{৫৫ - ৬০}{৫} = -১.০$
রফিক	অশোধিত মান = ৫৫ $z = \frac{৫৫ - ৬০}{২০} = -২.৫$	অশোধিত মান = ৮০ $z = \frac{৮০ - ৬০}{৫} = +৪.০$

চূড়ান্ত পরীক্ষায় রফিকের পরিমিত মান +৪.০ দ্বারা বোঝা যায় যে, সে ৮০ পেয়ে গড় থেকে ৪ পরিমিত ব্যবধান উপরে অবস্থান করছে। অপরদিকে, মধ্যবর্তী পরীক্ষায় মিতার পরিমিত মান +১.০ দ্বারা বোঝা যায় যে, সে ৮০ পেয়ে গড় থেকে ১ পরিমিত ব্যবধান উপরে অবস্থান করছে। সুতরাং, রফিকের প্রাপ্ত ৮০ নম্বর মিতার প্রাপ্ত ৮০ নম্বরের চেয়ে অধিক ভালো ফলাফল। পরিমিত মান ধনাত্মক এবং ঋণাত্মক উভয়ই হতে পারে। ঋণাত্মক পরিমিত মান অর্থ হলো যে, একটি সংখ্যা মানের গড় থেকে নীচে অবস্থান করাকে নির্দেশ করে।

পরিমিত মানের সুবিধা (Advantages of Standard Scores)

অশোধিত মানের তুলনায় পরিমিত মানের সুবিধা অনেক বেশী। একটি উদাহরণের মাধ্যমে বিষয়টি বোঝার চেষ্টা করা যাক। ধরা যাক, কোন একটি প্রতিষ্ঠানে কর্মচারী নিয়োগের জন্য প্রার্থীদের তিন ধরণের পরীক্ষায় অংশ নিতে হয় — প্রথমটি হলো ৫০০ নম্বরের লিখিত পরীক্ষা, দ্বিতীয়টি হলো ১০০ নম্বরের স্মরণশক্তি পরীক্ষা, এবং তৃতীয়টি হলো ৩ পয়েন্টের ভাষা দক্ষতার পরীক্ষা। পরীক্ষা অনুষ্ঠিত হবার পর দেখা গেলো যে, একজন প্রার্থী লিখিত পরীক্ষায় ৩৭০, স্মরণশক্তি পরীক্ষায় ৬০ এবং ভাষা দক্ষতা পরীক্ষায় ১.৬ পেয়েছে। এখন প্রশ্ন হলো, সার্বিকভাবে সংশ্লিষ্ট প্রার্থী কেমন করেছে? বিভিন্ন মাত্রায় পরিমাপ করা এই অশোধিত মানগুলোকে আমরা কিভাবে তুলনা করবো? অন্যান্য প্রার্থীর প্রাপ্ত নম্বরের চেয়ে এই প্রার্থী কত ভালো বা খারাপ করেছে?

উল্লিখিত নম্বরগুলো দেখেই আমরা এসব প্রশ্নের জবাব পেতে পারি না। কারণ, তিনটি পরীক্ষায় ব্যবহৃত মাত্রা ভিন্ন — ৫০০, ১০০, এবং ৩। সুতরাং, এ ক্ষেত্রে পরিমিত মানের মাধ্যমে আমরা এ সকল প্রশ্নের জবাব পেতে পারি। আমরা জানি যে, কোন বিন্যাসের গড় ও পরিমিত ব্যবধান জানা থাকলে আমরা সেই বিন্যাসের কোন একটি নির্দিষ্ট সংখ্যামানকে পরিমিত মানে রূপান্তরিত করতে পারি। এ ক্ষেত্রেও সকল পরীক্ষার্থীর নম্বরের গড় ও পরিমিত ব্যবধান জেনে আমরা আলোচ্য পরীক্ষার্থীর পরিমিত মান নির্ণয় করতে পারি।

ধরা যাক, বিভিন্ন পরীক্ষায় আলোচ্য পরীক্ষার্থীর প্রাপ্ত নম্বরের পরিমিত মানগুলো নিম্নরূপ,

z = –০.০৫ (লিখিত পরীক্ষা)

z = ০.১ (স্মরণশক্তি পরীক্ষা)

z = –২.০ (ভাষা দক্ষতার পরীক্ষা)

পরিমিত মানগুলো পর্যালোচনা করলে দেখা যায় যে, অন্যান্য পরীক্ষার্থীর তুলনায় আমাদের আলোচ্য পরীক্ষার্থী দু'টি পরীক্ষায় গড়ের চেয়ে নীচে অবস্থান করছে এবং একটি পরীক্ষায় গড় থেকে ১ পরিমিত ব্যবধান উপরে রয়েছে। অর্থাৎ, উক্ত পরীক্ষার্থী সার্বিকভাবে তেমন কৃতকার্যতা দেখাতে পারে নি।

সারাংশ

ভিন্ন পরিমাপ একক সম্বলিত দু'টি উপাত্তরাশির মধ্যে ভিন্নতার তুলনার জন্য নিরঙ্কুশ ব্যবধানগুলোকে আপেক্ষিক পদে রূপান্তরিত করার প্রয়োজন হয়। আপেক্ষিক বিস্তার পরিমাপের একাধিক পদ্ধতি রয়েছে। যেমন, চতুর্থক ব্যবধানাঙ্ক, গড় ব্যবধানাঙ্ক, ব্যবধানাঙ্ক, পরিমিত মান, ইত্যাদি। কোন উপাত্তরাশির গড় এবং পরিমিত ব্যবধানের শতকরা অনুপাতকে সেই উপাত্তরাশির ব্যবধানাঙ্ক বলা হয়। বিস্তৃতির আপেক্ষিক পরিমাপ নির্ণয়ের জন্য বিস্তৃতির যে কোন নিরঙ্কুশ পরিমাপ ব্যবহৃত হলেও আপেক্ষিক পরিমাপের ক্ষেত্রে প্রায় সবসময়ই আমরা পরিমিত ব্যবধান ব্যবহার করে থাকি। এক গুচ্ছ পর্যবেক্ষণের মধ্যে একটি বিশেষ পর্যবেক্ষণ কিভাবে এবং কতটুকু ভিন্ন তা নির্দেশের জন্য পরিমিত মান একটি বিশেষভাবে উপযোগী পদ্ধতি হিসাবে কাজ করে। পরিমিত মানকে z-মানও বলা হয়ে থাকে।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক (√) চিহ্ন দিন –

১। কোন উপাত্তরাশির গড় ও পরিমিত ব্যবধানের শতকরা অনুপাতকে:

ক. ভেদাঙ্ক বলে
খ. ব্যবধানাঙ্ক বলে
গ. পরিমিত ব্যবধান বলে
ঘ. পরিসর বলে

২। কোন উপাত্তরাশির কোন বিশেষ একটি মানের অবস্থান জানার জন্য:

ক. পরিমিত ব্যবধান জানা প্রয়োজন
খ. পরিমিত গড় জানা প্রয়োজন
গ. পরিমিত ভেদাঙ্ক জানা প্রয়োজন
ঘ. পরিমিত মান জানা প্রয়োজন

৩। পরিমিত মান নির্ণয়ের বীজগাণিতিক সূত্রটি হলো:

ক. $z = \frac{x_i - \overline{x}}{s}$

খ. $z = \frac{\overline{x} - x_i}{s}$

গ. $z = \frac{x_i^{২} - \overline{x}}{s}$

ঘ. $z = \frac{x_i - (\overline{x})^{২}}{s}$

সংক্ষিপ্ত প্রশ্ন

১। পরিমিত মানের সুবিধা আলোচনা করুন।

২। পরিমিত মানের বৈশিষ্ট্য আলোচনা করুন।

রচনামূলক প্রশ্ন

১। পরিমিত মানের ব্যবহার উদাহরণসহ আলোচনা করুন।

২। ব্যবধানাঙ্ক নির্ণয়ের প্রয়োজনীয়তা কি? উদাহরণসহ ব্যবধানাঙ্ক নির্ণয়ের পদ্ধতি আলোচনা করুন।

পাঠ - ৮

বিস্তৃতির পরিমাপের বিভিন্ন পদ্ধতির মধ্যে তুলনা এবং নির্বাচন
Comparison and Choice between Measures of Dispersion

এই পাঠ শেষে যা জানা যাবে —

- বিস্তৃতির পরিমাপসমূহের মধ্যে তুলনা
- বিস্তৃতির যথাযথ পরিমাপ নির্বাচন

বিস্তৃতির পরিমাপসমূহের মধ্যে তুলনা (Comparison between Measures of Dispersion)

উন্মুক্ত শ্রেণী ব্যাপ্তির ক্ষেত্রে সর্বাধিক উপযুক্ত পদ্ধতি হলো চতুর্থক ব্যবধান।

বিস্তৃতির প্রতিটি পরিমাপের নিজস্ব কিছু বৈশিষ্ট্য রয়েছে এবং প্রতিটিরই কিছু সুবিধা এবং অসুবিধা রয়েছে। এগুলো প্রয়োগের সিদ্ধান্ত নির্ভর করে উপাত্তের ধরণ এবং গবেষণার উদ্দেশ্যের উপর। প্রতিটি পদ্ধতির বৈশিষ্ট্য এবং উপযোগিতা বিবেচনা করে আমরা তাদের মধ্যে একটি তুলনামূলক বিশ্লেষণ করতে পারি এবং কোন পরিমাপটি অধিক গ্রহণযোগ্য সে বিষয়েও একটি সিদ্ধান্তে পৌঁছতে পারি।

পরিমিত ব্যবধান এবং ভেদাঙ্ককেই ব্যাপ্তিমূলক মাত্রায় বেশী ব্যবহার করা হয়।

পরিমাপের মাত্রা: আমরা জানি যে, বিভিন্ন মাত্রায় পরিমাপকৃত উপাত্তের জন্য বিস্তার পরিমাপের বিভিন্ন পদ্ধতি রয়েছে। নামসূচক উপাত্তের ক্ষেত্রে বিস্তার পরিমাপের জন্য উপযুক্ত পদ্ধতি নির্বাচন করা সম্ভব নয়। যদি ক্রমসূচক মাত্রায় পরিমাপকৃত উপাত্ত থাকে, তবে আমরা পরিসর, আন্তচতুর্থক পরিসর বা চতুর্থক ব্যবধান ব্যবহার করে থাকি। ব্যাপ্তিমূলক মাত্রায় পরিমাপকৃত উপাত্তের জন্য আমরা গড় ব্যবধান, পরিমিত ব্যবধান ও ভেদাঙ্ক ব্যবহার করে থাকি। তবে গড় ব্যবধান নির্ণয়ের ক্ষেত্রে ঋণাত্মক চিহ্ন উপেক্ষা করার জন্য যেহেতু অনপেক্ষতাকে ব্যবহার করা হয়, সেহেতু গড় ব্যবধানকে বীজগাণিতিক পরিগণনায় ব্যবহার করা যায় না এবং তত্ত্বগতভাবে ব্যাখ্যা করাও যায় না। তাই পরিমিত ব্যবধান এবং ভেদাঙ্ককেই ব্যাপ্তিমূলক মাত্রায় বেশী ব্যবহার করা হয়।

গড় ব্যবধান পরবর্তী বীজগাণিতিক পরিগণনায় ব্যবহারযোগ্য নয় এবং এটির তত্ত্বগত ব্যাখ্যা সম্ভব নয় বলে এ ক্ষেত্রে পরিমিত ব্যবধান ও ভেদাঙ্কের ব্যবহারই বেশী হয়ে থাকে।

কেন্দ্রীয় প্রবণতার পরিমাপের সাথে সম্পর্ক: কখনো কখনো বিস্তৃতির পরিমাপসমূহ কেন্দ্রীয় প্রবণতার পরিমাপসমূহের সাথে সঙ্গতি রেখে প্রয়োগ করা হয়। যদি কোন উপাত্তরাশির বর্ণনায় মধ্যমা ব্যবহার করা হয়, তবে তার বিস্তার পরিমাপের ক্ষেত্রে চতুর্থক ব্যবধান ব্যবহৃত হয়। কারণ, উভয় পদ্ধতিই অবস্থান ভিত্তিক পরিমাপ এবং শতহারিকের উপর নির্ভর করে। গড়ের মাধ্যমে উপাত্তের বিস্তৃতি পরিমাপের ক্ষেত্রে গড় ব্যবধান, পরিমিত ব্যবধান ও ভেদাঙ্ক প্রযোজ্য। কিন্তু গড় ব্যবধান পরবর্তী বীজগাণিতিক পরিগণনায় ব্যবহারযোগ্য নয় এবং এটির তত্ত্বগত ব্যাখ্যা সম্ভব নয় বলে এ ক্ষেত্রে পরিমিত ব্যবধান ও ভেদাঙ্কের ব্যবহারই বেশী হয়ে থাকে।

পরিমিত ব্যবধানের মান উপাত্তরাশির অসাধারণ মান দ্বারা সহজেই প্রভাবিত হয়। তাই বঙ্কিম বিন্যাসের ক্ষেত্রে পরিসর বা পরিমিত ব্যবধানের প্রয়োগ যথার্থ নয়। এ ক্ষেত্রে চতুর্থক ব্যবধানই অধিক নিরাপদ পরিমাপ।

বঙ্কিমতা: আমরা জানি যে, কোন উপাত্তরাশিতে চরম মান বা অসাধারণ মান থাকলে তা অধিক বঙ্কিমতা প্রদর্শন করে। আমরা এও জানি যে, পরিমিত ব্যবধানের মান উপাত্তরাশির অসাধারণ মান দ্বারা সহজেই প্রভাবিত হয়। তাই বঙ্কিম বিন্যাসের ক্ষেত্রে পরিসর বা পরিমিত ব্যবধানের প্রয়োগ যথার্থ নয়। বরং, এ ক্ষেত্রে চতুর্থক ব্যবধানই অধিক নিরাপদ পরিমাপ। কারণ, চতুর্থক ব্যবধান প্রথম ও তৃতীয় চতুর্থকের মধ্যে তুলনা করে বলে উপাত্তরাশির মধ্যে বিদ্যমান চরম মানগুলো এর বিবেচনার বাইরে থেকে যায়।

শ্রেণী ব্যাপ্তির প্রকৃতি: পরিসর উপাত্তরাশির সর্বোচ্চ ও সর্বনিম্ন মানের ব্যবধান প্রদর্শন করে। কিন্তু উন্মুক্ত শ্রেণী ব্যাপ্তির ক্ষেত্রে সর্বোচ্চ বা সর্বনিম্ন মানের যে কোন একটি বা উভয় মানই আমাদের অজানা থেকে যায়। সুতরাং, উন্মুক্ত শ্রেণী ব্যাপ্তির ক্ষেত্রে পরিসর প্রয়োগ করা যায় না। গড় ব্যবধান বা পরিমিত ব্যবধান নির্ণয়ের ক্ষেত্রে উপাত্তরাশির প্রতিটি মানকে বিবেচনায় এনে গণনা করা হয় এবং সে ক্ষেত্রে শ্রেণী ব্যবধান একটি গুরুত্বপূর্ণ বিষয়। যেহেতু উন্মুক্ত শ্রেণী ব্যাপ্তির ক্ষেত্রে কোন কোন শ্রেণীর ব্যবধান আমাদের অজানা থেকে যায়, সেহেতু সেখানে গড় ব্যবধান, পরিমিত ব্যবধান বা ভেদাঙ্ক নির্ণয় করা যায় না। সুতরাং, উন্মুক্ত শ্রেণী ব্যাপ্তির ক্ষেত্রে সর্বাধিক উপযুক্ত পদ্ধতি হলো চতুর্থক ব্যবধান। এ ক্ষেত্রে, উন্মুক্ত শ্রেণী সীমাকে প্রথম ও তৃতীয় চতুর্থকের বাইরে রেখে সহজেই চতুর্থক ব্যবধান নির্ণয় করা যায়। কিন্তু যদি উন্মুক্ত শ্রেণী ব্যাপ্তিটি প্রথম বা তৃতীয় চতুর্থকের মধ্যে অবস্থান করে, সে ক্ষেত্রে চতুর্থক ব্যবধানও প্রযোজ্য নয়।

বীজগাণিতিক পরিগণনা: যে কোন পরিসংখ্যানিক পদ্ধতি নির্বাচনের ক্ষেত্রে বীজগাণিতিক পরিগণনা একটি অন্যতম গুরুত্বপূর্ণ বিবেচ্য বিষয়। বিস্তার পরিমাপের ক্ষেত্রেও এর ব্যতিক্রম ঘটে না। এর সবগুলো পদ্ধতি বিবেচনা করে আমরা দেখতে পাই যে, একমাত্র ভেদাঙ্ক ও পরিমিত ব্যবধানই পরবর্তী বীজগাণিতিক পরিগণনার জন্য সবচেয়ে বেশী উপযুক্ত। কারণ, বিভিন্ন মাত্রায় পরিমাপকৃত উপাত্তগুলোর মধ্যে একমাত্র ব্যাপ্তিমূলক বা অনুপাতমূলক মাত্রায় পরিমাপকৃত উপাত্ত দিয়ে আমরা অধিক বীজগাণিতিক পরিগণনার দিকে অগ্রসর হতে পারি। আর এ রকম উপাত্তের জন্য পরিমিত ব্যবধানই সবচেয়ে উপযুক্ত পরিমাপ। একাধিক উপাত্তরাশির পরিমিত ব্যবধান থেকে আমরা একটি সমন্বিত পরিমিত ব্যবধান নির্ণয় করতে পারি, যার মাধ্যমে বিভিন্ন উপাত্তরাশির মধ্যে তুলনা করা সম্ভব হয়। বিস্তৃতির সবগুলো নিরঙ্কুশ পরিমাপের মধ্যে একমাত্র পরিমিত ব্যবধানই উপাত্তরাশির সবগুলো মানকে প্রকৃতভাবে বিবেচনা করে। গড় ব্যবধান সবগুলো মানকে বিবেচনা করলেও এতে ব্যবহৃত অনপেক্ষতা বীজগাণিতিক পরিগণনার ক্ষেত্রে এর উপযোগীতাকে কমিয়ে দেয়।

যদিও গবেষণার উদ্দেশ্য এবং উপাত্তের প্রকৃতি অনুযায়ী ক্ষেত্র বিশেষে অন্য পদ্ধতিগুলোকে একজন গবেষক বেশী উপযুক্ত মনে করেন, কিন্তু সার্বিক বিচারে, পরিমিত ব্যবধানই বিস্তার পরিমাপের সবচেয়ে গ্রহণযোগ্য পদ্ধতি।

সহজবোধ্যতা: সহজবোধ্যতার বিবেচনায় বিস্তার পরিমাপের পদ্ধতিগুলোর মধ্যে পরিসরই সবচেয়ে বেশী গ্রহণযোগ্য। কারণ, উপাত্তরাশির শুধুমাত্র দু'টি মান (সর্বনিম্ন ও সর্বোচ্চ) জানা থাকলেই পরিসর নির্ণয় করা সম্ভব। তাই, যখন খুব দ্রুততার সাথে বিস্তার পরিমাপ সম্পর্কে ধারণা নেবার প্রয়োজন হয় তখন পরিসরই হলো উপযুক্ত পরিমাপ।

হলো উপযুক্ত পরিমাপ।

বিস্তৃতির যথাযথ পরিমাপ নির্বাচন (Selection of Appropriate Measure of Dispersion)

বিস্তৃতির বিভিন্ন পরিমাপের মূল কথাটি হলো, যদি উপাত্তরাশি তুলনামূলকভাবে ছড়িয়ে ছিটিয়ে থাকে তবে পরিমাপের মানগুলো বড় হবে, আর যদি উপাত্তরাশি তুলনামূলকভাবে কাছাকাছি অবস্থান করে তবে সেগুলোর মান ছোট হবে। এখন প্রশ্ন হলো যে, বিস্তৃতির কোন পরিমাপটি শ্রেষ্ঠ? এক কথায় এর জবাব দেয়া সম্ভব নয়। কারণ, পরিমাপগুলোর প্রত্যয়গত অনন্যতা উপাত্তের প্রকৃতি এবং উপাত্ত বর্ণনার ক্ষেত্রে সেগুলোকে কিভাবে ব্যবহার করা হবে এই সকল বিষয়ের উপর সম্যক ধারণা নেবার পরই কেবলমাত্র এর জবাব দেয়া সম্ভব। পরিসর একটি বিন্যাসের সর্বোচ্চ ও সর্বনিম্ন মানের মধ্যে দূরত্বকে নির্দেশ করে। ক্রমসূচক ও ব্যাপ্তিমূলক মাত্রায় পরিমাপকৃত উপাত্তের বিস্তৃতি বর্ণনার জন্য এটিকে ব্যবহার করা হয়। পরিসরের সুবিধাটি হলো, এটি সহজে নির্ণয় করা যায় এবং বোঝা যায়। কিন্তু অসুবিধাটি হলো, এটি একটি অস্থিতিশীল পরিমাপ। কারণ, এটি সম্পূর্ণভাবে দু'টি পর্যবেক্ষণ মানের উপর নির্ভরশীল।

বঙ্কিম বিন্যাসের ক্ষেত্রে আন্তচতুর্থক পরিসর একটি উপযুক্ত পরিমাপ হিসাবে ব্যবহৃত হয়। আন্তচতুর্থক পরিসর তৃতীয় এবং প্রথম চতুর্থক অথবা ৭৫তম এবং ২৫তম শতহারিকের মধ্যকার পার্থক্যকে প্রদর্শন করে। যদিও পরিসরের তুলনায় অধিকতর স্থিতিশীল এবং ক্রমসূচক বা তার চেয়ে উচ্চ মাত্রায় পরিমাপকৃত উপাত্তের ক্ষেত্রে প্রয়োগ করা যায়, আন্তচতুর্থক পরিসর বিন্যাসের অন্তর্ভুক্ত সকল উপাত্তকে প্রতিফলিত করে না।

বিস্তৃতিকে বর্ণনা করার আরেকটি পদ্ধতি হলো, একটি পরিমাপের মাত্রায় কোন নির্ধারিত বিন্দু (fixed point) থেকে পর্যবেক্ষণগুলো বা উপাত্তরাশির মানগুলো কতটুকু ভিন্ন তার মাত্রাটি নির্দেশ করা। গাণিতিক যুক্তিতে একটি বিন্যাসের গড়কে এই স্মারক বিন্দু (reference point) হিসাবে নির্বাচন করা হয়। গড় থেকে প্রতিটি মানের বিচ্যুতির উপর নির্ভর করে দু'টি পরিমাপ উপাত্তের বিস্তৃতিকে নির্ণয় করে – একটি হলো গড় ব্যবধান, এবং অন্যটি হলো পরিমিত ব্যবধান। কিন্তু পরিমাপের অন্তর্নিহিত যুক্তি অনুযায়ী, পরিমিত ব্যবধানই হলো সবচেয়ে বেশী উপযোগী পরিমাপ। এই অন্তর্নিহিত যুক্তিগুলো আমরা পূর্বে আলোচনা করেছি। কাজেই, এ পর্যায়ে এসে আমরা আর সেগুলোর পুনরুক্তি করবো না। তবে, যেটি সংক্ষিপ্তভাবে উল্লেখ করা প্রয়োজন তা হলো, একটি বিন্যাসের বিস্তার বর্ণনার জন্য পরিমিত ব্যবধান সবচেয়ে বেশী কার্যকর এবং উপযোগী পরিমাপ প্রধানতঃ চারটি কারণে। *প্রথমতঃ*, মূল উপাত্ত যে এককে পরিমাপকৃত হয় এটি সেই পরিমাপেই প্রকাশিত হয়; *দ্বিতীয়তঃ*, এটি বিন্যাসের সকল মানকে প্রতিফলিত করে; *তৃতীয়তঃ*, নমুনা বিচ্যুতির দ্বারা কম প্রভাবিত হয় বলে বিস্তৃতির অন্যান্য পরিমাপের তুলনায় এটি অধিকতর স্থিতিশীল পরিমাপ; এবং *চতুর্থতঃ*, পরিমিত ব্যবধানের বীজগাণিতিক বৈশিষ্ট্যগুলো এটিকে জটিল পরিসংখ্যানিক কর্মকান্ডে ব্যবহারের জন্য সুযোগ করে দেয়। এক কথায়, যদিও গবেষণার উদ্দেশ্য এবং উপাত্তের প্রকৃতি অনুযায়ী ক্ষেত্র বিশেষে অন্য পদ্ধতিগুলোকে একজন গবেষক বেশী উপযুক্ত মনে করেন, কিন্তু সার্বিক বিচারে পরিমিত ব্যবধানই বিস্তার পরিমাপের সবচেয়ে গ্রহণযোগ্য পদ্ধতি।

শতহারিক (Percentile)

বিস্তৃতি পরিমাপের অপর একটি অবস্থান ভিত্তিক পরিমাপ হলো শতহারিক। সহজ কথায় বলা যায় যে, শতহারিক হলো একটি চলকের সেই সব মান যা সারিবদ্ধ তথ্যকে সমান ১০০ ভাগে ভাগ করে। অর্থাৎ, একটি সারিবদ্ধ তথ্যরাশির মধ্যে আমরা ৯৯টি শতহারিক পাই, শতহারিক নির্ণয়ে ক্ষেত্রে আমরা যততম শতহারিক নির্ণয় করতে চাই, $\frac{N+১}{১০০}$ -এর সাথে তত গুণ করেই আমরা সেই শতহারিকটি পাই। যেমন, একটি নিবেশনে যদি আমরা ৫ম শতহারিক বর্ণনা করতে চাই, তাহলে $L_১ + \frac{\frac{(N+১)n}{১০০} - f_c}{f_p} \times c$ তম সংখ্যাটিই হবে পঞ্চম শতহারিক। শতহারিক নির্ণয়ের প্রক্রিয়া চতুর্থক নির্ণয়ের প্রক্রিয়ার মতোই। অর্থাৎ,

$$n \text{ তম শতহারিক} = L_১ + \frac{\frac{(N+১)n}{১০০} - f_c}{f_P} \times c$$

এখানে, $L_১$ = n তম শতহারিক যে শ্রেণীতে অবস্থান করছে সেই শ্রেণীর নিম্ন সীমা

f_c = শতহারিক শ্রেণীর পূর্ববর্তী শ্রেণীর ক্রমযোজিত গণসংখ্যা

f_P = শতহারিক শ্রেণীর গণসংখ্যা

c = শ্রেণী ব্যবধান

একটি উদাহরণের মাধ্যমে বিষয়টি বোঝা যেতে পারে। পরবর্তী পৃষ্ঠায় একটি গ্রামের ৪০০ জন ব্যাক্তির আয়ের বিন্যাস নিম্নের সারণিতে উপস্থাপিত হলো:

সারণি ৫.৮.১: একটি গ্রামের ৪০০ জন ব্যাক্তির আয়ের বিন্যাস

মাসিক আয় (টাকা)	*গ্রামবাসীর সংখ্যা* (f_i)	*যোজিত গণসংখ্যা* (f_c)
৫০০০ – ১০০০০	৯২	৯২
১০০০০ – ১৫০০০	৭৯	১৭১
১৫০০০ – ২০০০০	৪৫	২১৬
২০০০০ – ২৫০০০	৩০	১৪৬
২৫০০০ – ৩০০০০	৬১	৩০৭
৩০০০০ – ৩৫০০০	২৬	৩৪৩
৩৫০০০ – ৪০০০০	৩০	৩৭৩
৪০০০০ – ৪৫০০০	৮	৩৮১
৪৫০০০ – ৫০০০০	৫	৩৮৬
৫০০০০ – ৫৫০০০	১০	৩৯৬
৫৫০০০ – ৬০০০০	৪	৪০০
$\sum f_i = ৪০০$		

উপরের শ্রেণী বিন্যাস থেকে যদি আমরা ৩০তম শতহারিক নির্ণয় করতে চাই, তাহলে প্রথমে আমাদের জানতে হবে ৩০তম শতহারিক কততম সংখ্যা এবং তা কোন শ্রেণীতে অবস্থান করছে। অর্থাৎ,

$\frac{(N+১)n}{১০০}$ তম সংখ্যা = $\frac{(৪০০+১)৩০}{১০০}$ = ১২০.৩তম সংখ্যা মানটিই হবে ৩০তম শতহারিক। এই মানটি ১০০০০ – ১৫০০০ এই শ্রেণীতে অবস্থান করছে।

সুতরাং, ৩০তম শতহারিক হবে $= ১০০০০ + \frac{১২০.৩ + ৯২}{১৭১} \times ৫০০০$

$= ১০০০০ + ০.১৬৫৪৯৭০৭৬ \times ৫০০০$

$= ১০০০০ + ৮২৭.৪৮৫৩৮$

$= ১০৮২৭.৪৮৫৩৮$

সুতরাং ৩০তম শতহারিক হলো ১০৮২৭.৪৯ টাকা

তাই, যখন খুব দ্রুততার সাথে বিস্তার পরিমাপ সম্পর্কে ধারণা নেবার প্রয়োজন হয় তখন পরিসরই হলো উপযুক্ত পরিমাপ।

সারাংশ

বিস্তৃতির পরিমাপের প্রতিটিরই কিছু নিজস্ব বৈশিষ্ট্য রয়েছে। এগুলোর যথাযথ প্রয়োগের সিদ্ধান্ত নির্ভর করে উপাত্তের বৈশিষ্ট্যের উপর। সাধারণতঃ নামসূচক উপাত্তের ক্ষেত্রে উপযুক্ত পদ্ধতি নির্বাচন করা সম্ভব নয়। কিন্তু ক্রমসূচক পরিমাপের ক্ষেত্রে পরিসর, আন্তচতুর্থক পরিসর, এবং ব্যাপ্তিমূলক পরিমাপের ক্ষেত্রে গড় ব্যবধান, পরিমিত ব্যবধান ও ভেদাঙ্ক ব্যবহার করে থাকি। বিস্তৃতির আর একটি বহুল ব্যবহৃত পদ্ধতি হচ্ছে শতহারিক পরিমাপ। শতহারিক হলো একটি চলকের সেই সব মান যা সারিবদ্ধ তথ্যকে সমান ১০০ ভাগে ভাগ করে।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক (√) চিহ্ন দিন –

১। ..... একটি বিন্যাসের সর্বোচ্চ ও সর্বনিম্ন মানের মধ্যে দূরত্বকে নির্দেশ করে।
ক. পরিমিত ব্যবধান খ. পরিমিত গড়
গ. শতহারিক ঘ. পরিসর

২। ..... হলো একটি চলকের সেই সব মান যা সারিবদ্ধ তথ্যকে সমান ১০০ ভাগে ভাগ করে।
ক. গড় খ. শতহারিক
গ. চতুর্থক ঘ. উপরের সব ক'টি

৩। ব্যাপ্তিমূলক মাত্রায় পরিমাপকৃত উপাত্তের জন্য আমরা:
ক. গড় ব্যবধান ব্যবহার করে থাকি খ. পরিমিত ব্যবধান ব্যবহার করে থাকি
গ. ভেদাঙ্ক ব্যবহার করে থাকি ঘ. উপরের সব ক'টি উত্তরই সঠিক

সংক্ষিপ্ত প্রশ্ন

১। সংজ্ঞা দিন ও পার্থক্য আলোচনা করুন:
ক. বঙ্কিমতা – শ্রেণী ব্যাপ্তির প্রকৃতি
খ. শতাহারিক – প্রথম চতুর্থক

রচনামূলক প্রশ্ন

১। বিস্তৃতির যথাযথ পরিমাপ নির্বাচনে কোন কোন বিষয়গুলিকে গুরুত্ব দেয়া উচিৎ বলে আপনি মনে করেন।

২। তথ্য ব্যাখ্যায় শতাহারিকের গুরুত্ব আলোচনা করুন।