১০

উপাত্তের বিশ্লেষণ, ব্যাখ্যা ও উপস্থাপন

Analysis, Interpretation and Presentation of Data

একটি গবেষণা প্রক্রিয়ার শেষ প্রান্তে এসে মূল্যায়ন করে দেখতে হবে যে, সংগৃহীত উপাত্তরাশি গবেষণার শুরুতে গৃহিত উদ্দেশ্যগুলোর সাথে সঙ্গতিপূর্ণ হয়েছে কি না, বা সেগুলো আদৌ কোন অর্থ বহন করে কি না। এ প্রক্রিয়ায়, উপাত্তকে বিশ্লেষণের মাধ্যমে চলকগুলোর বিভিন্ন বৈশিষ্ট্যের বর্ণনা করতে হয়, সে সকল বৈশিষ্ট্য চলকগুলোর মধ্যে সম্পর্কের সাধারণ রূপগুলো প্রকাশ করে কি না, তা পরীক্ষা করে দেখতে হয়, সেগুলোর পরিসংখ্যানগত তাৎপর্য পরীক্ষা করে দেখতে হয় এবং সেই সাধারণ রূপগুলোর কার্য-কারণ সম্পর্ক ব্যাখ্যা করতে হয়। এক-চলক, দ্বি-চলক ও বহু-চলক বিশিষ্ট বিশ্লেষণের মাধ্যমে উপাত্তের এ সকল বিশ্লেষণ করা হয়ে থাকে। সে জন্য উপাত্তের বিশ্লেষণ ও ব্যাখ্যার কতগুলো নীতিমালা অনুসরণ করতে হয়। বিশ্লেষণ ও ব্যাখ্যাকৃত গবেষণার ফলাফলকে যদি প্রকাশ না করা হয়, তবে পুরো গবেষণা প্রক্রিয়াটি অর্থহীন হয়ে পড়ে। অতএব, গবেষণার ফলাফলকে পাঠকের উদ্দেশ্যে উপস্থাপনের লক্ষ্যে একটি কার্যকর প্রতিবেদন প্রণয়নে একজন সফল গবেষকের দক্ষতার প্রকৃত প্রতিফলন ঘটে। এ ইউনিটে, এ সকল বিষয়ের আলোচনার মাধ্যমে আমরা এই গ্রন্থের সমাপ্তি টানবো।

এই ইউনিটে আমরা যে পাঠগুলো অধ্যয়ন করবো, সেগুলো হলো:

- পাঠ - ১ : উপাত্তের বিশ্লেষণ ও ব্যাখ্যার নীতিমালা
- পাঠ - ২ : এক-চলক বিশিষ্ট বিশ্লেষণ
- পাঠ - ৩ : দ্বি-চলক বিশিষ্ট বিশ্লেষণ
- পাঠ - ৪ : বহু-চলক বিশিষ্ট বিশ্লেষণ
- পাঠ - ৫ : গবেষণা প্রতিবেদন

পাঠ - ১

উপাত্তের বিশ্লেষণ ও ব্যাখ্যার নীতিমালা
Principles of Analysis and Interpretation of Data

এই পাঠ শেষে যা জানা যাবে —

- উপাত্তের বিশ্লেষণ ও ব্যাখ্যার মধ্যে সম্পর্ক
- অপরিমাণগত উপাত্তের ব্যবহার
- শ্রেণীবদ্ধকরণের নীতিমালা
- অকাঠামোবদ্ধ বিষয়বস্তু শ্রেণীবদ্ধকরণের নীতিমালা
- উপাত্তের পরিসংখ্যানগত বিশ্লেষণ

উপাত্তের বিশ্লেষণ ও ব্যাখ্যার মধ্যে সম্পর্ক (Relationship between Analysis and Interpretation of Data)

বিভিন্ন গবেষণা নক্শায়, বিশ্লেষণ ও ব্যাখ্যা সম্পর্কিত সমস্যাগুলোর ভিন্নতার কারণে উপাত্তের বিশ্লেষণ ও ব্যাখ্যার মধ্যে যে সম্পর্ক, তা একটি গবেষণা থেকে অন্য গবেষণায় ভিন্ন হয়ে থাকে। সাধারণভাবে, ব্যাখ্যামূলক বা পরীক্ষণমূলক গবেষণার তুলনায় বর্ণনামূলক গবেষণায়, উপাত্তের বিশ্লেষণ ও ব্যাখ্যার ক্ষেত্রে অধিকতর সমস্যা দেখা দেয়। তবে, যেহেতু বিশ্লেষণের মৌলিক নীতিমালাগুলো সব ধরণের গবেষণার ক্ষেত্রে একইভাবে প্রযোজ্য, সেহেতু গবেষণা নক্শার ভিত্তিতে ভিন্ন ভিন্নভাবে বিশ্লেষণ ও ব্যাখ্যার নীতিমালা সম্পর্কিত আলোচনার প্রয়োজন নেই। উপাত্তের বিশ্লেষণ ও ব্যাখ্যা সম্পর্কে দু'টি দৃষ্টিভঙ্গি রয়েছে। একটি দৃষ্টিভঙ্গি অনুযায়ী, বিশ্লেষণ ও ব্যাখ্যাকে দু'টি ভিন্ন প্রক্রিয়া হিসাবে দেখা হয় এবং দ্বিতীয় দৃষ্টিভঙ্গি অনুযায়ী, বিশ্লেষণ ও ব্যাখ্যাকে অঙ্গাঙ্গীভাবে জড়িত একটি প্রক্রিয়ার দু'টি অবিচ্ছেদ্য উপাদান হিসাবে দেখা হয়। আমরা আমাদের আলোচনায় দ্বিতীয় দৃষ্টিভঙ্গিটিকে গ্রহণ করবো।

যেহেতু উপাত্তের বিশ্লেষণ ও ব্যাখ্যা পারস্পরিকভাবে ঘনিষ্ট, সেহেতু ব্যাখ্যা আলাদা কোন কর্মকান্ড না হয়ে বিশ্লেষণেরই একটি বিশেষ বিষয়ে পরিণত হয়। অতএব, ব্যাখ্যার প্রক্রিয়াকে স্পষ্ট করে নেয়ার জন্য, বিশ্লেষণ পদ্ধতির পূর্বে ব্যাখ্যার আলোচনাটি করা যেতে পারে। উপাত্তের ব্যাখ্যা হলো, গবেষণায় প্রাপ্ত ফলাফলসমূহের একটি ব্যাপক অর্থের অন্বেষণ। এই অন্বেষার দু'টি প্রধান দিক রয়েছে। প্রথমতঃ, অন্য গবেষণায় প্রাপ্ত ফলাফলকে এই গবেষণার সাথে যুক্ত করে সামাজিক গবেষণার একটি ধারাবাহিকতা প্রতিষ্ঠা করা। ভিন্নভাবে বলা যায় যে, উপাত্তের ব্যাখ্যা বর্ণনামূলক অনুসন্ধান থেকে পরীক্ষণমূলক অনুসন্ধানে উত্তরণ ঘটানোর প্রক্রিয়ার সাথে সম্পর্কিত। এ ক্ষেত্রে, বর্ণনামূলক গবেষণায় প্রাপ্ত ফলাফলের ব্যাখ্যা পরীক্ষণমূলক গবেষণার জন্য অনুকল্প তৈরি করে। দ্বিতীয়তঃ, উপাত্তের ব্যাখ্যা উপাত্তের ব্যাখ্যামূলক প্রত্যয়ের জন্ম দেয়। উপাত্তের বিশ্লেষণ ও ব্যাখ্যার আলোচনায় আমরা অপরিমাণগত উপাত্তের ব্যবহার, বিশ্লেষণ হিসাবে শ্রেণীবদ্ধকরণ পদ্ধতি, উত্তর-দফার সারণিকরণ, উপাত্তের পরিসংখ্যানগত বিশ্লেষণ, কার্য-কারণ সম্পর্কের উপর সিদ্ধান্তমূলক উপসংহার টানা, ইত্যাদি বিষয়গুলোর উপর জোর দেবো।

অপরিমাণগত উপাত্তের ব্যবহার (Use of Non-Quantified Data)

সামাজিক গবেষণায় অপরিমাণগত উপাত্তের ব্যবহার প্রসঙ্গে পন্ডিতদের মধ্যে দু'ধরণের মতামত রয়েছে। একটি মত অনুযায়ী, সামাজিক বিজ্ঞানের সকল তথ্যকে নীতিগতভাবে পরিমাপ বা শ্রেণীবদ্ধ করা যায়। এই দৃষ্টিভঙ্গিটিকে 'দৃষ্টবাদী' দৃষ্টিভঙ্গি বলা হয়। অতএব, যখনই আমরা অপরিমাণগত উপাত্তের মুখোমুখি হই, তখনই আমাদের প্রধান কাজটি হলো, বিশ্লেষণের মাধ্যমে সেগুলোকে পরিশুদ্ধ করা, যাতে করে সেগুলোকে পরিমাণসূচকভাবে শ্রেণীবদ্ধ করা যায়। কিন্তু দ্বিতীয় মতটি প্রথম মতের বিপরীত মেরুতে অবস্থান করে। এই মত অনুযায়ী, প্রতিটি ঘটনাকেই স্বতন্ত্রভাবে অর্থবহ বলে মনে করা হয়। এই দৃষ্টিভঙ্গিটিকে 'প্রপঞ্চবাদী' দৃষ্টিভঙ্গি বলা হয়। যখন আমাদের পরিমাপ প্রক্রিয়া, মাপনী, বা শ্রেণীবদ্ধকরণের প্রক্রিয়ার পর্যাপ্ততা নিয়ে সন্দেহ দেখা দেয়, তখনই আমরা প্রপঞ্চবাদী দৃষ্টিভঙ্গির

অবস্থান গ্রহণ করে থাকি। অর্থাৎ, আমরা একটি নির্দিষ্ট পর্যবেক্ষণের অনন্যতাকে স্বীকার করে নিই। সামাজিক বিজ্ঞানের গবেষণায় দু'টি দৃষ্টিভঙ্গিরই প্রতিফলন দেখা যায়।

সামাজিক গবেষণা প্রতিবেদনের প্রতিটি পাঠকই একটি বিষয়ের সাথে পরিচিত যে, অপরিমাণগত অপরিশোধিত উপাত্তকে (যেভাবে সেগুলোকে সংগ্রহ করা হয়েছে) প্রায়শঃ উচ্চতর মাত্রায় পরিমাপকৃত পরিমাণগত উপাত্তের সাথে ব্যবহার করা হয়ে থাকে। অর্থাৎ, পরিমাণগত বিশ্লেষণের সাথে সাথে উত্তরদাতা কর্তৃক প্রদত্ত উক্তিকে হুবহু ব্যবহার করা হয়ে থাকে। এ বিষয়ে কোন সন্দেহ নেই যে, এ ধরণের হুবহু উক্তির অন্তর্ভুক্তি সামাজিক বিজ্ঞানের গবেষণার নির্ভরশীলতার মাত্রাকে বৃদ্ধি করে। তার অর্থ এই নয় যে, হুবহু উক্তির অন্তর্ভুক্তির পিছনে কোন জোর বৈজ্ঞানিক যুক্তি রয়েছে। এ বিষয়ে কোন বিতর্কে না গিয়ে, বিশ্লেষণের পর্যায়ে এ ধরণের উপাদান ব্যবহারের বৈজ্ঞানিক কর্মকান্ডটি আলোচনা করা আমাদের উদ্দেশ্য, বৈজ্ঞানিক যুক্তিটি নয়।

অতএব, অপরিশোধিত উপাত্তকে পরিমাণগতভাবে রূপান্তর করা হয়েছে কি না, সেটি বিতর্কের বাইরে রেখে বলা যেতে পারে যে, বিশ্লেষণ ও ব্যাখ্যায় অপরিশোধিত উপাত্তকে ব্যবহার করা যেতে পারে। একটি উক্তির সকল বিষয়কে পরিমাণগতভাবে শ্রেণীবদ্ধকরণ করার পাশাপাশি, সেই উক্তির হুবহু ব্যবহার করা যায়, যা পুরো প্রেক্ষাপটকে অনুধাবন করতে সাহায্য করে। তবে, এটি মনে রাখা প্রয়োজন যে, একটি বর্ণনামূলক উক্তির সকল বিষয়কেই পরিমাণগতভাবে শ্রেণীবদ্ধ করা যায় না। অপরিমাণগত অপরিশোধিত উপাত্ত ব্যবহারের একটি অতিরিক্ত সুবিধা হলো যে, পরিমাণগত বিশ্লেষণে যে বিষয়গুলোকে অবহেলা বা উপেক্ষা করা হয়, সে বিষয়গুলোও সামনে চলে আসে। বিশ্লেষণের পর্যায়ে, অপরিশোধিত উপাত্তের ব্যবহার দু'টি স্বতন্ত্র কর্মকান্ড সম্পাদন করে। একটি হলো, নতুন অন্তর্দৃষ্টিকে উদ্দীপিত করা এবং অন্যটি হলো, একটি বিশেষ শ্রেণী-দফার সাথে সংশ্লিষ্ট অর্থের পরিধিকে স্পষ্ট করে তোলা।

নতুন অন্তর্দৃষ্টিকে উদ্দীপিতকরণ (Stimulating New Insights): একটি গবেষণার নক্শা যত কঠোর নিয়মানুবর্তীতা অনুসরণ করেই নির্মাণ করা হোক না কেন এবং এর পরিসংখ্যানগত বিশ্লেষণের পদ্ধতিগুলো যত পরিশীলিতই হোক না কেন, অপরিশোধিত উপাত্তকে পরীক্ষা নিরীক্ষা করে দেখলে, তা থেকে অনেক গুরুত্বপূর্ণ অন্তর্দৃষ্টির জন্ম নিতে পারে। এ সকল অন্তর্দৃষ্টি বিশ্লেষণকে সহজতর করার মাধ্যমে চলকগুলোর মধ্যে বিদ্যমান সম্পর্কের প্রকৃতিকে স্পষ্টভাবে বুঝতে সাহায্য করে এবং ভবিষ্যৎ গবেষণার জন্য অনুকল্পের জন্ম দেয়। পরিসংখ্যানগত বিশ্লেষণের শেষ পর্যায় পর্যন্ত অপেক্ষা না করে, যদি অপরিমাণগত উপাত্তের পরীক্ষা-নিরীক্ষা পুরো গবেষণা প্রক্রিয়ার মধ্যে চালানো হয়, সে ক্ষেত্রে এটি সবচেয়ে বেশী ফলদায়ক হয়। একজন সামাজিক বিজ্ঞানী যদি এ বিষয়টি সম্পর্কে সচেতন থাকেন, তবে তিনি নিঃসন্দেহে সংগৃহীত উপাত্তের ব্যাখ্যায় অধিকতর উপাদান যুক্ত করতে পারবেন এবং তার গবেষণার উপসংহারের মানকে সমৃদ্ধ করতে পারবেন।

সামাজিক বিজ্ঞানীদের অধিকাংশ বিশ্লেষণী প্রয়াস নিযুক্ত থাকে, একদল ব্যক্তির বস্তুগত বৈশিষ্ট্য এবং তাদের ভাবগত প্রতিক্রিয়ার মধ্যে সম্পর্ক স্থাপনের উপর। কিন্তু বিদ্যমান সম্পর্কটির অস্তিত্ব যে সকল উপাদানের সাথে জড়িত, শুধুমাত্র সেগুলোকে উপলব্ধি করাই যথেষ্ট নয়। অপরিশোধিত উপাত্তের পরীক্ষা নিরীক্ষা নতুন দিগন্তের উন্মোচনও করে থাকে। তবে, এটি মনে রাখা প্রয়োজন যে, চলকগুলোর মধ্যে সম্পর্কের ব্যাখ্যার জন্য, দিক নির্দেশনার উৎস হিসাবে অপরিশোধিত উপাত্তের ব্যবহার করা গেলেও সে ব্যাখ্যাগুলো সবসময় সঠিক নাও হতে পারে। যতক্ষণ পর্যন্ত না এই প্রস্তাবিত ব্যাখ্যাগুলো পরিমাপ করা যায় এবং সেগুলো পরিসংখ্যানগত পরীক্ষাকে সম্ভব করে তোলে, ততক্ষণ পর্যন্ত সেই গবেষণার জন্য অপরিশোধিত উপাত্ত থেকে আহরিত ব্যাখ্যাগুলো সাময়িক হয়ে থাকবে। কিন্তু গবেষণা যদি একটি পারস্পরিক সম্পর্কের অন্বেষায় সীমাহীন প্রক্রিয়া হয়ে থাকে, তবে এ ধরণের ব্যাখ্যাগুলো, অন্যের গবেষণায় সম্পর্কের পরিসংখ্যানগত ব্যাখ্যার জন্য একটি গুরুত্বপূর্ণ পদক্ষেপ হিসাবে কাজ করে। অপরিশোধিত উপাত্তের মধ্যে ব্যাখ্যামূলক প্রত্যয়ের সন্ধান, ঐ বিশেষ সম্পর্কের জন্য একমাত্র সম্ভাব্য ব্যাখ্যা হয়ে উঠতে পারে। এ ধরণের পরিস্থিতিতে, সেই ব্যাখ্যাগুলো এত ব্যতিক্রমধর্মী হতে পারে যে, তার জন্য পরিশীলিত পরিসংখ্যানগত বিশ্লেষণের প্রয়োজন পড়ে না।

শ্রেণী-দফাগুলোর অর্থের পরিধিকে স্পষ্টকরণ (Illustrating the Meaning of Categories): বিভিন্ন শ্রেণী-দফা সংজ্ঞায়নের ক্ষেত্রে, অপরিশোধিত উপাত্তের ব্যবহার সংজ্ঞায়িত শ্রেণী-দফাগুলোর অর্থকে উপলব্ধির জন্য খুবই গুরুত্বপূর্ণ হয়ে উঠে। আমরা প্রায়শঃ লক্ষ্য করি যে, যে তীক্ষ্মতা নিয়ে বিভিন্ন শ্রেণী-দফাগুলোকে সংজ্ঞায়িত করা হয়, সেগুলো অপরিশোধিত উপাত্তের প্রকৃতি, গবেষণা সমস্যার ধরণ এবং যে পরিস্থিতি থেকে উপাত্ত সংগ্রহ করা হয়, তা থেকে ভিন্নতা প্রদর্শন করে। ধরা

যাক, একজন গবেষক পর্যবেক্ষণ পদ্ধতির মাধ্যমে একটি কারখানার শ্রমিক ও ব্যবস্থাপনা কর্তৃপক্ষের মধ্যেকার মিথস্ক্রিয়ার প্রকৃতির উপর গবেষণা করতে চান। এ ক্ষেত্রে, বিশ্লেষণের প্রত্যাশিত মাত্রার উপর নির্ভর করে, পর্যবেক্ষণের বিভিন্ন শ্রেণী-দফার ধরণগুলো তৈরি করা যেতে পারে। যেমন, একটি দলের সদস্যদের সাথে অন্য দলের সদস্যদের মধ্যে স্বতঃস্ফূর্ত বাক্য বিনিময়ের ঘটনসংখ্যা গণনা করা যেতে পারে, বা তাদের আসন গ্রহণের বিন্যাসটি পর্যবেক্ষণ করা যেতে পারে। এ ধরণের শ্রেণী বিভক্তিকরণের সংজ্ঞাগুলো এত সুস্পষ্ট যে, তাদের কোন ব্যাখ্যার প্রয়োজন পড়ে না।

কিন্তু এর চেয়ে জটিল পর্যবেক্ষণকে শ্রেণীবদ্ধকরণ করার ক্ষেত্রে (যেমন, দু'টি দলের মধ্যে মিথস্ক্রিয়ার রূপ, যা উচ্চতর, নিম্নতর বা সমপর্যায়ের মর্যাদাকে নির্দেশ করে), বিষয়টি অস্পষ্ট হয়ে যেতে পারে। যেমন, ব্যবস্থাপনা কর্তৃপক্ষের কোন সিদ্ধান্তের বিরোধীতা করতে গিয়ে একজন শ্রমিক বলতে পারেন, 'আমি আপনার বক্তব্যের সাথে দ্বিমত পোষণ করি'। একজন ব্যবস্থাপনা কর্তৃপক্ষের প্রতিনিধি কোন শ্রমিককে বলতে পারেন, 'আপনি ঠিক বলেন নি'। এই দু'টি উক্তিকে মর্যাদা সম্পর্কে সচেতনতার দু'টি ভিন্ন শ্রেণী-দফায় বিভক্ত করতে গিয়ে, উক্তি দু'টির অর্থ সেই শ্রেণী-দফাগুলোর অর্থের সাথে মেলাতে হবে, নইলে ব্যাখ্যাটি স্পষ্ট হবে না।

শ্রেণীবদ্ধকরণের নীতিমালা (Principles of Classification)

যদি কয়েকশ' সম্ভাব্য উত্তর-দফাকে গবেষণা প্রশ্নের উত্তর হিসাবে ব্যবহার করতে হয়, তবে সেগুলোকে অবশ্যই কিছু সীমিত সংখ্যক শ্রেণীতে বিভক্ত করে নিতে হবে। প্রাসঙ্গিক শ্রেণীগুলো কি হবে, সে বিষয়ে সিদ্ধান্ত গ্রহণের ক্ষেত্রে কিছু নীতিমালা নির্বাচন করে নিতে হয়। গবেষণা প্রশ্ন, বা যদি কোন অনুকল্প গঠিত হয় তা, শ্রেণীবদ্ধকরণের নীতিমালা নির্বাচনের ভিত্তি প্রদান করে। ধরা যাক, একটি গবেষণার বিষয়বস্তু হলো, সরকারী প্রতিষ্ঠানে দূর্নীতির মাত্রা সম্পর্কে পরিমাপ করা। এ ক্ষেত্রে, শ্রেণীবদ্ধকরণের যথাযথ নীতিমালার ভিত্তিটি হবে, সুস্পষ্টভাবে দুর্নীতির উপস্থিতির উল্লেখ করা। এ রকম নীতিমালা, তাৎক্ষণিকভাবে দু'টি উত্তর-দফাকে নির্দেশ করে: একটি হলো, দুর্নীতির উল্লেখ এবং অন্যটি হলো, দুর্নীতির অনুলেখ। এ দু'টি দফা একটি 'শ্রেণী-দফা সেট' (category set) তৈরি করে। একটি শ্রেণী-দফা সেটকে কতগুলো মৌলিক শর্তকে পূরণ করতে হয়। যেমন,

- একটি শ্রেণী-দফার সেটকে একটি একক শ্রেণীবিভক্তিকরণ নীতিমালা থেকে আহরিত হতে হবে।
- একটি শ্রেণী-দফার সেটকে সামগ্রিক বা পুঙ্খানুপুঙ্খ হতে হবে। অর্থাৎ, একটি শ্রেণী-দফায় অন্তর্ভুক্ত উত্তর-দফার কোনটিকে বাদ না দিয়ে সবগুলো উত্তর-দফাকে অন্তর্ভুক্ত হতে হবে।
- একটি সেটের মধ্যেকার শ্রেণী-দফাগুলোকে হতে হবে পারস্পরিকভাবে বিশিষ্ট। কোন একটি উত্তরকে একের অধিক শ্রেণী-দফায় স্থান দেওয়া যাবে না।

যদি কোন নিয়ম ভাঙ্গা না হয়, তবে একটি সেটে দুইয়ের অধিক শ্রেণী-দফা থাকতে পারে। সে ক্ষেত্রে, কতটুকু ব্যতিক্রমকে গ্রহণ করা হবে, সে বিষয়ে একটি সিদ্ধান্ত গ্রহণ করতে হবে। নীতিগতভাবে, শ্রেণী-দফার সেট নির্মাণের জন্য সবগুলো বৈশিষ্ট্যকে ব্যবহার করা সম্ভব হলেও বাস্তবে এটি প্রায়শঃ মিতব্যয়ী ও ফলদায়ক হয় না। কারণ, সকল শ্রেণীবিভক্তিকরণের নীতিই অনুসন্ধানের উদ্দেশ্যকে কার্যকর করে না। কিছু কিছু ক্ষেত্রে, অপরিশোধিত উপাত্তকে শ্রেণীবদ্ধ করে বিভিন্ন শ্রেণীতে বিভক্ত করা কঠিন হয়ে পড়ে। যখন উত্তরগুলো মোটামুটিভাবে সরল ও স্পষ্ট হয়, তখন শ্রেণী-দফার সেট নির্মাণ অপেক্ষাকৃতভাবে সহজ এবং সেগুলোকে সুস্পষ্টভাবে সংজ্ঞায়িত করা যায়। কিন্তু কিছু কিছু বিষয়বস্তুর ক্ষেত্রে, কাজটি কঠিন ও সময় সাপেক্ষ হয়ে পড়ে।

ধরা যাক, একটি প্রতিষ্ঠানের নীতি নির্ধারকরা কর্মকর্তা ও কর্মচারীদের জন্য একটি আবাসিক এলাকা নির্মাণের পরিকল্পনা নিয়েছেন। পরিকল্পনার প্রস্তাবনায় বলা হয়েছে যে, "আবাসিক এলাকা নির্মাণের জন্য স্থানের সঙ্কট নিরসন, কর্মস্থলে যাবার জন্য যানবাহনের যথাযথ ব্যবহার, কর্মকর্তা কর্মচারীদের মধ্যে পারস্পরিক সম্পর্কের উন্নয়ন, ইত্যাদি বিষয়গুলোকে বিবেচনায় রেখে, কর্মস্থল সংলগ্ন একটি নির্দিষ্ট এলাকায় কর্মকর্তা ও কর্মচারীদের জন্য একটি বহুতলবিশিষ্ট আবাসিক এলাকা গড়ে তোলা হবে"। সকল কর্মকর্তা কর্মচারীদের মধ্যে এই প্রস্তাবনা গ্রহণযোগ্য হবে কি না, সে লক্ষ্যে একটি মতামত জরিপ পরিচালনা করা হয়। সে জরিপে, অন্যান্য প্রশ্নের সাথে কর্মকর্তাদের জিজ্ঞাসা করা হয়েছিলো যে, "কর্মকর্তাদের সাথে কর্মচারীরা একই আবাসিক এলাকায় বসবাস করলে কর্মচারীদের অনুভূতিটি কি হবে বলে আপনি মনে করেন"? এবং কর্মচারীদের জিজ্ঞাসা করা হয়েছিলো যে, "কর্মচারীদের সাথে কর্মকর্তারা একই আবাসিক এলাকায় বসবাস করলে কর্মকর্তাদের অনুভূতিটি কি হবে বলে আপনি মনে করেন"? এই প্রশ্ন দু'টির উত্তরে দেখা গিয়েছে যে, উত্তরগুলো সবচেয়ে ঋণাত্মক মনোভাব থেকে সবচেয়ে ধনাত্মক মনোভাব পর্যন্ত বিস্তৃত রয়েছে। যেমন,

কর্মকর্তারা বলেছেন	*কর্মচারীরা বলেছেন*
• কর্মচারীরা মনে করবে যে, তারা সমমর্যাদাসম্পন্ন।	• তাদের তো বন্ধুত্বপূর্ণ বলেই মনে হয়।
• তারা ধারণাটিকে গণতান্ত্রিক মূল্যবোধ প্রতিষ্ঠার মাইল ফলক বলে মনে করবে।	• আমার মনে হয় না তারা কিছু মনে করবেন।
• আমি তাদের সাথে মেলামেশা করি না কাজেই আমি জানি না তারা কি মনে করবে।	• তারা ধারণটিকে মোটেই পছন্দ করেন নি।
• তারা নিজেরাও অস্বস্তিবোধ করবেন।	• তারা মনে করবে যে, তাদের মর্যাদাকে ক্ষুণ্ন করা হয়েছে।

এই উত্তরগুলোকে দিয়ে 'সবচেয়ে অনুকূলে' এবং 'সবচেয়ে প্রতিকূলে' মাত্রায় শ্রেণীবদ্ধকণের নীতিমালা অনুসরণ করে, একটি শ্রেণী-দফা সেট তৈরি করা যায়। অর্থাৎ, শ্রেণী-দফাগুলো হবে 'সবচেয়ে অনুকূলে', 'সবচেয়ে প্রতিকূলে', 'নিরপেক্ষ' এবং 'অন্যান্য'। উপরের প্রতিটি উত্তরের মধ্যে একটি স্বতন্ত্র অর্থ স্পষ্টভাবে বোঝা যায়। কিছু উত্তরের সাথে একটি মূল্যবোধ জড়িত রয়েছে, যা একই আবাসিক এলাকায় বসবাসের ফলে কর্মকর্তা ও কর্মচারীদের কিছু সুবিধা ও অসুবিধাকে নির্দেশ করে এবং কিছু বক্তব্যের মধ্যে সেই মূল্যবোধের অনুপস্থিতি লক্ষ্য করা যায়।

যে কোন ধরণের শ্রেণী-দফার তালিকা ব্যবহারের পূর্বে, প্রতিটি শ্রেণী-দফার বিষয়বস্তুকে যতটা সম্ভব সঠিকভাবে নির্দিষ্ট করে নেওয়া প্রয়োজন। শ্রেণী-দফার অর্থকে বোঝানোর জন্য, প্রতিটি শ্রেণী-দফার সাথে উদাহরণের মাধ্যমে একটি বা দু'টি ব্যাখ্যামূলক বাক্য যুক্ত করে দেওয়া উচিৎ। যেমন, এই উত্তরগুলোর অর্থ হলো যে, "কর্মচারীরা কর্মকর্তাদের সাথে একই আবাসিক এলাকায় বসবাস করে কিছু সুবিধা ভোগ করবে"। উদাহরণ হিসাবে বলা যায় যে, "কর্মচারীরা মনে করবে যে, তারা সমমর্যাদাসম্পন্ন"। এ ধরণের সুনির্দিষ্টভাবে সংজ্ঞায়িত শ্রেণী-দফাগুলোর ব্যবহারে কোন সমস্যা দেখা না দিলেও, কিছু কিছু উত্তরকে শ্রেণীবদ্ধ করার ক্ষেত্রে জটিলতা দেখা দেয়। যেমন, কোন কোন উত্তর যদি এমন হতো যে, "তারা ধারণাটিকে পছন্দ করেছে, কারণ তারা জানে কেন", বা "তারা একই আবাসিক এলাকায় বসবাস করতে চায় না, কারণ, তারা জানে কেন"। সে ক্ষেত্রে, যত যত্নই নেয়া হোক না কেন, এ ধরণের উত্তরকে বিশ্লেষণের জন্য শ্রেণীবদ্ধ করা বেশ জটিল ও কষ্টকর হয়ে পড়ে। কারণ, উপরের উত্তরটির অর্থ বিভ্রান্তিকর। সেটি কোন সুবিধা বা অসুবিধা সম্পর্কিত মূল্যবোধের সাথে জড়িত অর্থকে নির্দেশ করে কি না, তা নির্ধারণ করা কঠিন। এ ধরণের উত্তরকে কিভাবে শ্রেণীবদ্ধ করা হবে, সে বিষয়ে অতিরিক্ত নীতিমালা প্রণয়ন করতে হবে।

অকাঠামোবদ্ধ বিষয়বস্তু শ্রেণীবদ্ধকরণের নীতিমালা (Principles of Classifying Unstructured Material)

অকাঠামোবদ্ধ বিষয়বস্তু শ্রেণীবদ্ধকরণে (যেমন, পর্যবেক্ষণের নিয়ম, ঘটনার ইতিহাস, বক্তৃতা, বা অকাঠামোবদ্ধ সাক্ষাৎকার, ইত্যাদির ক্ষেত্রে), কিছু বিশেষ সমস্যা দেখা দেয়। অকাঠামোবদ্ধ উপাদান নিয়ে কাজ করতে গেলে, প্রথমে যে সমস্যাটি দেখা দেয়, তা হলো বিষয়বস্তুর কোন দিকগুলোকে শ্রেণীবদ্ধ করা হবে এবং শ্রেণী-দফার সেট নির্মাণে কোন শ্রেণীবদ্ধকরণ নীতি ব্যবহৃত হবে, সে বিষয়ে সিদ্ধান্ত গ্রহণের সমস্যা। যেহেতু অনুসন্ধানমূলক গবেষণা সংজ্ঞাগত কারণে কোন সুস্পষ্ট অনুকল্প নিয়ে শুরু করে না, সেহেতু এ ধরণের গবেষণার ক্ষেত্রে, শ্রেণীবদ্ধকরণ নীতিমালা নির্ধারণ করা বিশেষভাবে কঠিন। উপাত্ত সংগ্রহের সময় গবেষক জানেন না যে, কোন বিষয়টি সবচেয়ে গুরুত্বপূর্ণ হয়ে দেখা দেবে। কারণ, বিশ্লেষণের সময় গবেষক শুধু অকাঠামোবদ্ধ উপাদান নিয়েই কাজ করেন না, তাকে অনেক বেশী পরিমাণে উপাত্ত নিয়েও কাজ করতে হয়, যার অধিকাংশই হয়তো পরবর্তীতে অপ্রাসঙ্গিক বলে বিবেচিত হতে পারে।

কোন অনুসন্ধানমূলক গবেষণায় তথ্য বিশ্লেষণের প্রথম ধাপটি হলো, কার্যকরী অনুকল্প গঠন করা, যা শ্রেণীবদ্ধকরণ নীতিমালার জন্ম দেয়। সে লক্ষ্যে, গবেষক তার গবেষণায় প্রাপ্ত উপাত্তের সকল উপাদানকে অধ্যয়ন করতে শুরু করেন এবং শ্রেণীবদ্ধকরণের সূত্র পাওয়া যায় কি না, সে জন্য সতর্ক থাকেন। কার্যকরী অনুকল্প নির্মাণের বেশ কিছু পদ্ধতি রয়েছে। প্রথমতঃ, যে দল বা গোষ্ঠীর উপর তিনি গবেষণা করছেন, তার ঠিক বিপরীত বৈশিষ্ট্যসম্পন্ন একটি গোষ্ঠীর উপর বিদ্যমান উপাদানগুলোকে অধ্যয়ন করবেন, যাতে করে তিনি এ দু'টি গোষ্ঠীর মধ্যে পার্থক্যগুলোকে বুঝতে পারেন। যেমন, কিশোর অপরাধীদের উপর পরিচালিত গবেষণায়, সমাজকল্যাণ দপ্তরের নথি থেকে অপরাধী নয়, এমন

কিশোরদের কেস স্টাডিও অধ্যয়ন করতে পারেন। দ্বিতীয়তঃ, গবেষক তার ঘটনাগুলোকে একইরকম সাধারণ বৈশিষ্ট্যের ভিত্তিতে বিভিন্ন দলে ভাগ করতে পারেন। তারপর, গবেষক নিজেকেই প্রশ্ন করবেন যে, তিনি বিভিন্ন শ্রেণী-দফাগুলোকে এক রকম মনে করেছেন কিসের ভিত্তিতে অনুপ্রাণিত হয়ে? তৃতীয়তঃ, গবেষক এমন বিষয়গুলোকে পর্যবেক্ষণ ও সেগুলোর সম্ভাব্য ব্যাখ্যার সন্ধান করবেন, যা সাধারণ জ্ঞান, বা তাত্ত্বিক প্রত্যাশার দিক থেকে আশ্চর্যজনক প্রপঞ্চ বলে মনে হতে পারে। এমন কি, সুস্পষ্ট অনুকল্প নিয়ে অকাঠামোবদ্ধ বিষয়বস্তুর বিশ্লেষণও অনেক সময় সমস্যা সৃষ্টি করে। কারণ, কোন একটি দলিল থেকে কোন একটি নির্দিষ্ট বিষয়ের তথ্য বাদ পড়ে যেতে পারে এবং অনুকল্পের সাথে সরাসরি সম্পর্কিত নয়, এমন বিষয়বস্তুও প্রচুর পরিমাণে পাওয়া যেতে পারে।

উপাত্তের পরিসংখ্যানগত বিশ্লেষণ (Statistical Analysis of Data)

উপাত্ত সংগঠিত করা, সেগুলোকে অর্থবহভাবে উপস্থাপন করা এবং পর্যবেক্ষণগুলোকে অনুমিত অনুকল্পের ভিত্তিতে বর্ণনা ও ব্যাখ্যার জন্য, আমাদের পরিসংখ্যানগত পদ্ধতি ব্যবহারের প্রয়োজন। কারণ, আমরা যে সকল প্রপঞ্চকে নিয়ে গবেষণা করি, সে সব প্রপঞ্চের মধ্যে বিদ্যমান সাধারণ রূপগুলো, পরিসংখ্যানগত কৌশলের প্রয়োগ ছাড়া অভিজ্ঞতালব্ধভাবে বুঝতে পারি না। পরিসংখ্যানকে যখন শুধুমাত্র বর্ণনা ও বিশ্লেষণের জন্য ব্যবহার করা হয়, তখন সেই পরিসংখ্যানগত পদ্ধতিকে বর্ণনামূলক পরিসংখ্যান পদ্ধতি বলে। উপাত্তকে কার্যকর এবং অর্থবহভাবে সংগঠিত এবং সংক্ষিপ্তকরণের জন্য, বর্ণনামূলক পরিসংখ্যান গবেষককে বিভিন্ন পদ্ধতি ও কৌশল সরবরাহ করে। অন্যদিকে, পরিসংখ্যানকে যখন অভিজ্ঞতামূলক পর্যবেক্ষণের ভিত্তিতে, কোন বস্তু বা ঘটনার উপর প্রাপ্ত সাধারণ রূপের ব্যাখ্যার মাধ্যমে সিদ্ধান্ত গ্রহণ, উপসংহার টানা এবং ভবিষ্যদ্বাণীকরণের জন্য ব্যবহার করা হয়, তখন সেই পরিসংখ্যানগত পদ্ধতিকে সিদ্ধান্তমূলক পরিসংখ্যান পদ্ধতি বলে। এ সকল সিদ্ধান্ত, উপসংহার ও ভবিষ্যদ্বাণীর পূর্বে অনুমিত অনুকল্প পরীক্ষার মাধ্যমে গৃহীত হয়। সিদ্ধান্তমূলক পরিসংখ্যান পদ্ধতি অনুকল্প পরীক্ষা ও কার্য-কারণ ব্যাখ্যার সকল প্রকার পদ্ধতি ও কৌশল সরবরাহ করে।

উপাত্তের বর্ণনায় আমরা একটি ঘটনা, বস্তু বা জনগোষ্ঠীর সাধারণ বৈশিষ্ট্যগুলো কি তা জানতে চাই। এই সাধারণ বৈশিষ্ট্যগুলো জানবার জন্য যে পরিসংখ্যানগত পরিমাপগুলো ব্যবহার করা হয়, সেগুলোকে কেন্দ্রীয় প্রবণতার পরিমাপ বলে। কেন্দ্রীয় প্রবণতার বিভিন্ন পরিমাপ রয়েছে, যার প্রতিটিকে উপাত্তের প্রকৃতি সম্পর্কে পূর্ব অনুমানের নির্দিষ্ট শর্ত পূরণ করতে হয়। সেই অনুমানগুলো যদি উপযুক্ত না হয়, তবে কেন্দ্রীয় প্রবণতার পরিমাপ বিভ্রান্তিকর হতে পারে। একটি গাণিতিক গড় এমন উপাত্তের জন্য প্রয়োগ করতে হবে, যেখানে ব্যাপ্তিমূলক মাত্রা বা ব্যাপ্তির সমতাকে নির্দেশ করবে। যেমন, 'একজন ব্যক্তি মাসে কতবার প্রেক্ষাগৃহে ছবি দেখতে যান', এ ধরণের উপাত্তের জন্য গাণিতিক গড় ব্যবহার করা যথোপযুক্ত হবে। কারণ, এই মাত্রায় 'কতবার' - এর মত সংখ্যা, ব্যাপ্তির সমতাকে ধারণ করে। অন্যদিকে, বিভিন্ন ধরণের ছবি দেখার সংখ্যার ভিত্তিতে পছন্দের ক্রম নির্ণয় করার জন্য মধ্যমা প্রয়োগ করাই শ্রেয় হবে। কারণ, যেহেতু পছন্দের মাত্রার অবস্থানগুলো একটি থেকে আরেকটি সমান দূরত্বের নয়, সেহেতু গাণিতিক গড়ের ব্যবহার, এ ক্ষেত্রে প্রযোজ্য হবে না। কোন ধরণের ছবি বেশী দেখেন, প্রশ্নের উত্তরের ক্ষেত্রে প্রচুরক হবে উপযুক্ত পরিমাপ।

উপাত্তের বর্ণনায় আমরা এও জানতে চাই যে, একটি নির্দিষ্ট ঘটনা, বস্তু বা ব্যক্তি সেই ঘটনা, বস্তু বা ব্যক্তির সাধারণ বৈশিষ্ট্য থেকে কতটুকু ভিন্ন। অর্থাৎ, প্রতিটি ব্যক্তির ছবি দেখার পছন্দ কি একই রকম, না কি সে ক্ষেত্রে ভিন্নতা রয়েছে, আমরা তাও জানতে চাই। কেন্দ্রীয় প্রবণতার পরিমাপসমূহের মত, ভিন্নতা পরিমাপেরও বিভিন্ন ধরণ রয়েছে এবং উপাত্তের প্রকৃতির সাথে নির্দিষ্ট পূর্ব-অনুমানের শর্ত পূরণ করলেই কেবল একটি পরিমাপকে প্রয়োগ করা যাবে। সাধারণ বৈশিষ্ট্য থেকে প্রতিটি উপাত্ত কতটুকু বিস্তৃত, তা বর্ণনার জন্য যে পরিমাপগুলো ব্যবহৃত হয়, সেই পরিমাপগুলোকে বিস্তৃতির পরিমাপ বলে। সেগুলোর মধ্যে পরিসর, চতুর্থক ব্যবধান, গড় ব্যবধান ও পরিমিত ব্যবধান গুরুত্বপূর্ণ।

আমরা জানি যে, পরিসংখ্যানগত পদ্ধতির সুষ্ঠু ব্যবহারে, বিন্যাসের রূপ ও আকৃতি সম্পর্কে ধারণা অর্জন অপরিহার্য। সাধারণভাবে অনুমান করে নেওয়া হয় যে, যে কোন চলকের জন্য বিন্যাসের রেখাটি স্বাভাবিক। কিন্তু প্রকৃতপক্ষে বহু অভিজ্ঞতালব্ধ বিন্যাস তাৎপর্যপূর্ণভাবে স্বাভাবিকতা থেকে আলাদা হয়ে থাকে। অর্থাৎ, যে চলকটিকে পরিমাপ করা হয়েছে, সেই চলকের পরিপ্রেক্ষিতে গোষ্ঠীর সদস্যরা কিভাবে বিন্যস্ত রয়েছে, উপাত্ত বর্ণনা করতে গিয়ে আমরা তাও জানতে চাই। যেমন, প্রেক্ষাগৃহে ছবি একবারেই দেখতে যান না, এমন লোকের সংখ্যা কি যারা মাসে তিনবার যান, তাদের প্রায় সমান? না কি তুলনামূলকভাবে বেশি মানুষ মাসে তিনবার দেখতে যান, বা অপেক্ষাকৃত কম মানুষ একেবারেই যান না এবং অপেক্ষাকৃত আরও কম মানুষ মাসে ছয় বা তার অধিকবার যান?

বিভিন্ন চলকের মধ্যে সম্পর্ক পরীক্ষা করে দেখার জন্য সম্পর্কের পরিমাপগুলো ব্যবহার করা হয়ে থাকে। কিন্তু এই পরিমাপগুলো শুধুমাত্র সম্পর্কের বর্ণনা করতে সক্ষম হয়, সম্পর্কের কার্য-কারণ ব্যাখ্যা করতে পারে না। দুই বা

ততোধিক ব্যক্তি বা গোষ্ঠীর মধ্যকার পার্থক্য বর্ণনার জন্য, আমরা বিভিন্ন পরীক্ষা ব্যবহার করি (যেমন, t-test, z-test, F-test, χ^2-test, ইত্যাদি)। সংগৃহীত তথ্য থেকে প্রাপ্ত ফলাফলের সাথে প্রকৃত ঘটনার তুলনার জন্য, আমরা বিভিন্ন পরীক্ষা করতে চাই। নমুনা থেকে প্রাপ্ত ফলাফলের ভিত্তিতে, যে সমগ্রক থেকে নমুনা চয়ন করা হয়েছে, সেই সমগ্রকের উপর সাধারণীকরণের জন্য অনুকল্প পরীক্ষা করা হয়। কার্য-কারণ সম্পর্কের ব্যাখ্যার লক্ষ্যে, বহু-চলক বিশিষ্ট বিশ্লেষণের মাধ্যমে বিভিন্ন পরিসংখ্যানগত পদ্ধতি ব্যবহার করা হয়।
সাধারণভাবে, উপাত্ত সংগ্রহ ও প্রক্রিয়াকরণের পর, গবেষক বিশ্লেষণের জন্য প্রস্তুত হন। উপাত্ত বিশ্লেষণের তিনটি প্রধান প্রকরণ রয়েছে — এক-চলক বিশিষ্ট বিশ্লেষণ, দ্বি-চলক বিশিষ্ট বিশ্লেষণ এবং বহু-চলক বিশিষ্ট বিশ্লেষণ। যদিও সামাজিক বিজ্ঞানের অধিকাংশ পরিসংখ্যানগত বিশ্লেষণ বহু-চলক বিশিষ্ট বিশ্লেষণের মাধ্যমে সম্পন্ন করা হয়, তবুও এক-চলক এবং দ্বি-চলক বিশিষ্ট বিশ্লেষণ প্রক্রিয়া সম্পর্কে ধারণা থাকা প্রয়োজন। তা না হলে, বহু-চলক বিশিষ্ট বিশ্লেষণকে পুরোপুরিভাবে অনুধাবন করা যায় না। এক-চলক বিশিষ্ট বিশ্লেষণ হলো, একটি চলকের বিন্যাসকে শুধুমাত্র একটি চলক দিয়েই পরীক্ষা করার প্রক্রিয়া এবং দ্বি-চলক বিশিষ্ট বিশ্লেষণ দু'টি চলকের মধ্যেকার সম্পর্ককে বর্ণনা করে। এই তিন ধরণের পরিসংখ্যানগত বিশ্লেষণ পদ্ধতির অধীনে বিভিন্ন প্রকার পরিসংখ্যানগত কৌশল রয়েছে, যেগুলো পাঠ ২, ৩ ও ৪ -এ বিস্তারিতভাবে আলোচনা করা হবে।

সারাংশ

উপাত্তকে সারণিবদ্ধ করার পর, তা বর্ণনা ও ব্যাখ্যার প্রয়োজনীয়তা দেখা দেয়। উপাত্তের ব্যাখ্যা ও বিশ্লেষণের ক্ষেত্রে, দু'টি দৃষ্টিভঙ্গির পরিচয় আমরা পাই। একটি দৃষ্টিভঙ্গি অনুযায়ী, বিশ্লেষণ ও ব্যাখ্যা দু'টি ভিন্ন প্রক্রিয়া এবং অপর দৃষ্টিভঙ্গি অনুযায়ী, এ দু'টি অঙ্গাঙ্গীভাবে জড়িত একটি প্রক্রিয়ার দু'টি অবিচ্ছেদ্য উপাদান। অপরিমাণগত উপাত্ত সম্পর্কেও দু'টি দৃষ্টিভঙ্গির পরিচয় পাওয়া যায়। প্রথমটি অনুযায়ী, সামাজিক বিজ্ঞানের সকল তথ্যকে নীতিগতভাবে পরিমাপ বা শ্রেণীবদ্ধ করা যায়। অপর দৃষ্টিভঙ্গি অনুযায়ী, প্রতিটি ঘটনাকেই স্বতন্ত্রভাবে অর্থবহ বলে মনে করা হয়। বস্তুতঃ, সমাজ গবেষণায় যত বিতর্ক বা দৃষ্টিভঙ্গিই থাকুক না কেন, প্রাপ্ত উপাত্তকে অর্থবহ করে ব্যাখ্যা ও বর্ণনা করার উপর নির্ভর করে গবেষণার যথার্থতা।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক (√) চিহ্ন দিন –

১। 'প্রতিটি ঘটনাকেই স্বতন্ত্রভাবে অর্থবহ বলে মনে করা হয়' — কাদের মতে?
ক. দৃষ্টবাদীদের মতে খ. ক্রিয়াবাদীদের মতে
গ. প্রপঞ্চবাদীদের মতে ঘ. মার্কসবাদীদের মতে।

২। সুস্পষ্ট অনুকল্প নিয়ে অকাঠামোবদ্ধ বিষয়বস্তুর বিশ্লেষণও সমস্যার সৃষ্টি করতে পারে। কারণ:
ক. দলিল থেকে একটি নির্দিষ্ট বিষয়ের তথ্য বাদ পড়তে পারে
খ. অনুকল্পের সাথে জড়িত নয় এমন বিষয়বস্তু প্রচুর পরিমাণে থাকতে পারে
গ. ক ও খ উভয়ই
ঘ. অকাঠামোবদ্ধ বিষয়বস্তুতে বর্ণনা থাকে পরিমিত।

৩। সাধারণভাবে, উপাত্ত বিশ্লেষণের ____________ প্রধান প্রকরণ রয়েছে।
ক. দু'টি খ. তিনটি
গ. চারটি ঘ. পাঁচটি

সংক্ষিপ্ত প্রশ্ন

১। শ্রেণীবদ্ধকরণের নীতিমালা কী?
২। অপরিমাণগত উপাত্তের বৈশিষ্ট্য কী?

রচনামূলক প্রশ্ন

১। শ্রেণীবদ্ধকরণের নীতিমালাগুলো উদাহরণসহ আলোচনা করুন।
২। উপাত্তকে অর্থবহ করে উপস্থাপনের ক্ষেত্রে, পরিসংখ্যানের ব্যবহার উদাহরণসহ আলোচনা করুন।

পাঠ - ২

এক-চলক বিশিষ্ট বিশ্লেষণ
Univariate Analysis

এই পাঠ শেষে যা জানা যাবে —

- এক-চলক বিশিষ্ট বিশ্লেষণের মাধ্যমে উপাত্তের বর্ণনা
- কেন্দ্রীয় প্রবণতার পরিমাপ
- বিস্তৃতির পরিমাপ
- বিন্যাসের রূপ ও আকৃতি
-

এক-চলক বিশিষ্ট বিশ্লেষণের মাধ্যমে উপাত্তের বর্ণনা (Description of Data through Univariate Analysis)

এক-চলক বিশিষ্ট বিশ্লেষণ হলো, একটি মাত্র চলকের উপর ভিত্তি করে সেই চলকের বৈশিষ্ট্যগুলোর বিন্যাসকে পরীক্ষা করে দেখা। এক-চলক বিশিষ্ট বিশ্লেষণ প্রক্রিয়ায়, অনেক ধরণের পরিসংখ্যানগত পদ্ধতি ও কৌশল ব্যবহার করা হয়ে থাকে। এর মধ্যে উল্লেখযোগ্য পদ্ধতি ও কৌশলগুলো হলো, গণসংখ্যা নিবেশন, কেন্দ্রীয় প্রবণতা ও বিস্তৃতির বিভিন্ন পরিমাপসহ বিন্যাসের রূপ ও আকৃতির বিভিন্ন পরিমাপ। যেহেতু ইউনিট ৯-এর পাঠ ৩-এ উপাত্ত সংক্ষিপ্তকরণের প্রক্রিয়া হিসাবে সারণিকরণ ও গণসংখ্যা নিবেশনের উপর বিস্তারিত আলোচনা করা হয়েছে, সেহেতু এই পাঠে তা নিয়ে আমরা আলোচনা করবো না। তবে, শুধু এটুকু গুরুত্বের সাথে উল্লেখ করা যায় যে, একগুচ্ছ পর্যবেক্ষণকে একটি গণসংখ্যা সারণিতে উপস্থাপন করা হলো, সাধারণতঃ যে কোন প্রকার উপাত্ত বিশ্লেষণের প্রাথমিক পদক্ষেপ। কখনো এটি শেষ পদক্ষেপও হতে পারে, যদি গণসংখ্যা নিবেশনের অতিরিক্ত কোন বিশ্লেষণ করা না হয়।

একটি গণসংখ্যা সারণির চাক্ষুষ পর্যবেক্ষণই একটি জনগোষ্ঠী বা সম্প্রদায়ের কিছু চমকপ্রদ বৈশিষ্ট্য সম্পর্কে জানার জন্য যথেষ্ট হতে পারে। একটি গণসংখ্যা নিবেশন পর্যবেক্ষণগুলোর মধ্যে ব্যবধান এবং শ্রেণীবদ্ধ গোষ্ঠীগুলোর মধ্যে বিদ্যমান পার্থক্যকে নির্দেশ করে। একটি গণসংখ্যা সারণি দেখে সহজেই বোঝা যায়, কোন শ্রেণী বা গোষ্ঠীটি পর্যবেক্ষণগুলোর মধ্যে প্রতিনিধিত্বকারী। অবিন্যস্ত উপাত্ত রাশিকে শ্রেণীবদ্ধ করে গণসংখ্যা নিবেশনের মাধ্যমে উপস্থাপন, একটি বিন্যাসকে বর্ণনার উপযোগী প্রক্রিয়া হলেও, বর্ণনা ও বিশ্লেষণের লক্ষ্যে আরো সংক্ষিপ্ত পরিমাপের প্রয়োজন হয়, যা একটি একক সংখ্যা দিয়ে একটি পুরো বিন্যাসকে বর্ণনা করবে। এ ধরণের একক সংখ্যা দিয়ে বিন্যাসের বর্ণনার জন্য কতগুলো সংক্ষিপ্ত পরিমাপ আলোচনা করে দেখা যাক। তবে, উল্লেখ করা প্রয়োজন যে, উপাত্তের বিশ্লেষণে এ সকল পরিসংখ্যানগত পদ্ধতি ও কৌশল প্রয়োগের যৌক্তিকতাটিকেই শুধুমাত্র এই পাঠে আলোচনা করা হবে, সেগুলোর নির্ণয় পদ্ধতি নয়। কারণ, সেগুলো "সামাজিক পরিসংখ্যান পরিচিতি" (সমাজতত্ত্ব-৪) গ্রন্থে বিস্তারিতভাবে উপস্থাপন করা হয়েছে। প্রয়োজনবোধে, সংশ্লিষ্ট পরিমাপের নির্ণয় পদ্ধতিগুলো বোঝার জন্য, সেই গ্রন্থ অনুসরণ করা যেতে পারে।

কেন্দ্রীয় প্রবণতার পরিমাপ (Measures of Central Tendency)

উপাত্তরাশির একটি বিন্যাসে বিভিন্ন ধরণের প্রবণতা লক্ষ্য করা যায়। যেমন, কোন উপাত্তরাশির মানগুলো একটি বিস্তৃত পরিসরে বিন্যস্ত থাকলেও অধিকাংশ সংখ্যামান একটি স্থানে কেন্দ্রীভূত হবার প্রবণতা প্রদর্শন করে। বিশেষ মানকে কেন্দ্র করে অধিকাংশ উপাত্তরাশির কেন্দ্রীভূত হবার এই প্রবণতাটিকে একটি একক সংখ্যামান পরিমাপের মাধ্যমে বর্ণনা করা যায়। এ ধরণের পরিমাপগুলোকে কেন্দ্রীয় প্রবণতার পরিমাপ বলে। কেন্দ্রীয় প্রবণতার বিভিন্ন পরিমাপের মধ্যে প্রচুরক (mode), মধ্যমা (median) ও গাণিতিক গড়ের (arithmetic mean) ব্যবহার সামাজিক গবেষণায় বেশী দেখা যায়।

একটি বিন্যাসের মধ্যে যে মানটি সর্বাধিকবার ঘটে, সেই মানটিকে প্রচুরক বলে।

প্রচুরক (Mode): কেন্দ্রীয় প্রবণতার সবচেয়ে সহজতর পরিমাপটি হলো প্রচুরক। উপাত্ত সংগ্রহের পর যখন উপাত্তরাশিকে একটি গণসংখ্যা নিবেশনে সাজানো হয়, তখন দেখা যায় যে, কিছু সংখ্যামান কোন কোন শ্রেণীতে অধিকবার আবির্ভূত হয়েছে। একটি বিন্যাসের মধ্যে যে মানটি সর্বাধিকবার ঘটে সেই মানটিকে প্রচুরক বলে এবং সর্বাধিক সংখ্যামান সম্বলিত সেই শ্রেণী বা শ্রেণী-দফাকে প্রচুরক শ্রেণী বলা হয়। অবিন্যস্ত উপাত্তের ক্ষেত্রে প্রচুরক নির্ণয় খুব সহজ। যেমন, ২৫, ৪৫, ২২, ৪৫, ৩০, ২০, ৪৫, ১৭, ৩৫ এবং ৪৫ উপাত্ত রাশিমালাটিকে পর্যবেক্ষণ করলে লক্ষ্য করা যায় যে, ৪৫ সংখ্যাটি সর্বাধিকবার ঘটেছে। এই ৪৫ সংখ্যাটিই হলো এই উপাত্তরাশির প্রচুরক। কিন্তু বিন্যস্ত উপাত্তে ক্ষেত্রে একটি বিশেষ সূত্র ব্যবহার করে প্রচুরক নির্ণয় করতে হয়, যা হলো:

$$\text{প্রচুরক} = L_১ + \frac{\Delta_১}{\Delta_১ + \Delta_২} \times C.I.$$

যেখানে, $L_১$ = প্রচুরক শ্রেণীর নিম্ন সীমা

$\Delta_১$ = প্রচুরক ও প্রচুরক-পূর্ব শ্রেণীর ঘটনসংখ্যার মধ্যে পার্থক্য

$\Delta_২$ = প্রচুরক ও প্রচুরক-পরবর্তী শ্রেণীর ঘটনসংখ্যার মধ্যে পার্থক্য

C.I. = শ্রেণী ব্যাপ্তি।

মধ্যমা (Median): ক্রমসূচক মাত্রায় পরিমাপকৃত চলকের ক্ষেত্রে, অনেক সময় আমাদের একটি বিন্যাসের মধ্য-মানটির অবস্থান নির্ণয়ের প্রয়োজন হতে পারে। যেমন, মেধাক্রম নির্ণয়ের পর, ছাত্র-ছাত্রীদের কৃতকার্যতার আপেক্ষিক গুরুত্ব বোঝার জন্য আমরা জানতে চাইতে পারি যে, নির্দিষ্ট একটি সংখ্যামানের নীচে বা উপরে কতজন, বা কত শতাংশ ছাত্র-ছাত্রী রয়েছে। এ ধরণের আপেক্ষিক অবস্থান নির্ণয়ের জন্য, কেন্দ্রীয় প্রবণতার পরিমাপ হিসাবে মধ্যমা নির্ণয় করা হয়ে থাকে। মধ্যমা একটি উপাত্তরাশিকে সমান দুই ভাগে ভাগ করে। অর্থাৎ, মানের উর্ধ্বক্রম বা নিম্নক্রম অনুসারে সাজানো উপাত্ত রাশিমালাকে যে সংখ্যামানটি সমান দু'ভাগে ভাগ করে, সেই সংখ্যামানটিকে মধ্যমা বলে। উপাত্তরাশির মোট সংখ্যা যদি বেজোড় হয়, সে ক্ষেত্রে ঠিক মাঝখানের রাশিটি হলো মধ্যমা। কিন্তু যদি উপাত্তরাশির মোট সংখ্যা জোড় সংখ্যায় হয়, সে ক্ষেত্রে উপাত্তরাশির মাঝখানের দু'টি মানের গড় মানটি মধ্যমাকে নির্দেশ করে।

একটি উদাহরণ দিয়ে বিষয়টি বোঝার চেষ্টা করা যাক। নয়টি সংখ্যার একটি উপাত্তরাশি হলো, ৩, ৮, ১১, ১২, ১৫, ১৮, ২২, ২৫ এবং ৩৬। সংজ্ঞা অনুযায়ী, এই নয়টি সংখ্যা মানের মাঝখানের সংখ্যাটি হলো মধ্যমা, অর্থাৎ, ১৫। কিন্তু অতিরিক্ত একটি সংখ্যা, ১৬, এই উপাত্ত রাশিমালার সাথে যোগ করা হলে, তখন সেখানে আর কোন মধ্য-মান থাকে না। ক্রম অনুসারে সাজানো সেই অতিরিক্ত মানযুক্ত দশটি মানের উপাত্ত রাশিমালাটি হলো, ৩, ৮, ১১, ১২, ১৫, ১৬, ১৮, ২২, ২৫ এবং ৩৬। এ ক্ষেত্রে মধ্যমা নির্ণয় করতে হলে, উপাত্ত রাশিমালার মাঝখানের সংখ্যা মান দু'টির গড় নির্ণয় করতে হবে।

$$\text{অতএব, মধ্যমা} = \frac{(১৫ + ১৬)}{২} = ১৫.৫$$

বিন্যস্ত উপাত্ত থেকে মধ্যমা নির্ণয়ের সূত্রটি হলো,

$$\text{মধ্যমা} = L_১ + \frac{\frac{n}{২} - F}{f} \times C.I.$$

যেখানে, $L_১$ = মধ্যমা সম্বলিত শ্রেণীর নিম্ন সীমা

F = মধ্যমা সম্বলিত শ্রেণীর নিম্ন সীমার ক্রমযোজিত গণসংখ্যা

f = মধ্যমা সম্বলিত শ্রেণী ব্যাপ্তির অন্তর্ভুক্ত গণসংখ্যা

n = মোট গণসংখ্যা

C.I. = মধ্যমা সম্বলিত শ্রেণীর ব্যাপ্তি।

উপাত্ত রাশিমালার সকল মানের যোগফলকে উপাত্তরাশির মোট সংখ্যা দিয়ে ভাগ করলে যে সংখ্যামানটি পাওয়া যায়, সেটিই হলো গাণিতিক গড়।

গাণিতিক গড় (Arithmetic Mean): গাণিতিক গড় হলো, সামাজিক গবেষণায় সবচেয়ে বেশী জনপ্রিয় এবং বহুল ব্যবহৃত কেন্দ্রীয় প্রবণতার পরিমাপ । উপাত্ত রাশিমালার সকল মানের যোগফলকে উপাত্তরাশির মোট সংখ্যা দিয়ে ভাগ করলে যে সংখ্যামানটি পাওয়া যায়, সেটিই হলো গাণিতিক গড়। সংজ্ঞা অনুযায়ী, গড় নির্ণয়ের বীজগাণিতিক সূত্রটি হলো,

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N}$$

যেখানে, $\bar{x}$ = x চলকের গড়

x_i = x চলকের প্রতিটি মান

N = উপাত্তরাশির মোট সংখ্যা।

ধরা যাক, ৫টি পরিবারের সদস্য সংখ্যা যথাক্রমে ৪, ৩, ২, ৫ ও ৬ জন। অতএব, সূত্র অনুযায়ী,

$$\bar{x} = \frac{\sum x_i}{N} = \frac{৪+৩+২+৫+৬}{৫} = \frac{২০}{৫} = ৪$$

অর্থাৎ, ৫টি পরিবারের গড় আকার হলো ৪।

বিন্যস্ত উপাত্ত থেকে গাণিতিক গড় নির্ণয়ের সূত্রটি হলো,

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

যেখানে, $\bar{x}$ = x চলকের গড়

f_i = ঘটনসংখ্যা

x_i = x চলকের প্রতিটি মান

N = উপাত্তরাশির মোট সংখ্যা।

গাণিতিক গড় অনেক বেশী নির্ভরযোগ্য এবং প্রতিনিধিত্বশীল পরিমাপ।

উপাত্তকে বিশ্লেষণ করতে গিয়ে গবেষক কোন পরিমাপটি ব্যবহার করবেন, তা নির্ভর করে উপাত্তরাশির প্রকৃতি, গবেষণার বিষয়বস্তু, উপাত্ত বিশ্লেষণের প্রকৃতি এবং সুনির্দিষ্ট উদ্দেশ্যে ব্যবহৃত কেন্দ্রীয় প্রবণতার প্রত্যয়টির উপর। উপাত্ত রাশিমালার বিন্যাসটি যদি সুষম হয়, সে ক্ষেত্রে তিনটি পরিমাপের যে কোন একটি ব্যবহার করলেই চলে। তবে গাণিতিক গড়ের ব্যবহার সবসময়ই প্রধান বিবেচনায় থাকা উচিৎ। কারণ, গাণিতিক গড় অনেক বেশী নির্ভরযোগ্য এবং প্রতিনিধিত্বশীল পরিমাপ। গাণিতিক গড়কে উচ্চতর পরিসংখ্যানিক ব্যবহারে প্রয়োগ করা যায়। নমুনা বিচ্যুতি দ্বারা কম প্রভাবিত হয় বলে গাণিতিক গড় কেন্দ্রীয় প্রবণতার সবচেয়ে স্থিতিশীল একটি পরিমাপ। কিন্তু সব ধরণের পরিস্থিতিতে, গাণিতিক গড় আদর্শ কেন্দ্রীয় পরিমাপ নয়। উপাত্তরাশি যখন নামসূচক মাত্রায় পরিমাপকৃত হয়, তখন প্রচুরক হলো কেন্দ্রীয় প্রবণতার একমাত্র পরিমাপ। যখন উপাত্তরাশি গুণবাচক হয়, বা ক্রমসূচক মাত্রায় পরিমাপকৃত হয়, তখন মধ্যমা হলো কেন্দ্রীয় প্রবণতার শ্রেষ্ঠ পরিমাপ। যখন একটি বিন্যাস খুব বেশী বঙ্কিম হয়, তখন মধ্যমা একটি অর্থবহ কেন্দ্রীয় পরিমাপ হিসাবে কাজ করে। ব্যাপ্তিমূলক মাত্রায় পরিমাপকৃত চলকের ক্ষেত্রে, গাণিতিক গড় নির্ণয় হলো যথাযথ। কিন্তু উপাত্তরাশির মধ্যে প্রান্তিক মানের উপস্থিতি থাকলে, এটি একটি বিন্যাস সম্পর্কে ভ্রান্ত ধারণা দিয়ে থাকে। যদি তিনটি পরিমাপের মানের মধ্যে খুব বেশী ভিন্নতা দেখা যায়, তবে সে ক্ষেত্রে তিনটি পরিমাপকেই নির্ণয় করে উপস্থাপন করা বাঞ্ছনীয়।

বিস্তৃতির পরিমাপ (Measures of Dispersion)

কোন বিন্যাসের অন্তর্ভুক্ত উপাত্তরাশি একটি বিশেষ স্থানে কেন্দ্রীভূত হবার প্রবণতা প্রদর্শন করলেও প্রতিটি মান সেই কেন্দ্রীয় মানটি থেকে ভিন্ন। বিস্তৃতির পরিমাপগুলো, কেন্দ্রীয় প্রবণতার পরিমাপগুলোর প্রতিনিধিত্বশীলতার মাত্রা সম্পর্কে একটি সূক্ষ্ম ধারণা প্রদান করে। উপাত্তরাশির মানগুলো যদি কেন্দ্রীয় মান থেকে খুব বেশী ছড়িয়ে থাকে, 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 বিভিন্ন পরিমাপ রয়েছে। যেমন, পরিসর (range), চতুর্থক ব্যবধান (quartile deviation), গড় ব্যবধান (mean deviation), পরিমিত ব্যবধান (standard deviation) এবং ভেদাঙ্ক (variance)।

পরিসর (Range): বিস্তৃতির পরিমাপের সবচেয়ে সহজ, সাধারণ এবং স্থূল পরিমাপটি হলো পরিসর। সংখ্যামানগুলোকে মানের ক্রম অনুসারে সাজিয়ে, সর্বোচ্চ ও সর্বনিম্ন মানের পার্থক্য নির্ণয় করে আমরা খুব সহজেই পরিসরের মান পেতে পারি। ধরা যাক, ১০ জন ব্যক্তির ওজন ৫৫, ৫৭, ৬০, ৬৫, ৬৩, ৫৬, ৬৭, ৭০, ৭৫ এবং ৭১ কিলোগ্রাম। এখানে সর্বোচ্চ মান হলো ৭৬ এবং সর্বনিম্ন মান হলো ৫৫

কিলোগ্রাম। এই উপাত্তরাশির পরিসর হলো, ৭৫ − ৫৫ = ৪২ কিলোগ্রাম। অর্থাৎ, ১০ জন ব্যক্তির ওজনের ব্যবধান হলো ৪২ কিলোগ্রাম। অর্থাৎ, উপাত্তরাশির সর্বোচ্চ ও সর্বনিম্ন মানের মধ্যকার পার্থক্যই হলো পরিসর। পরিসংখ্যান সম্পর্কে যারা তেমন জ্ঞান রাখেন না, তাদের জন্য পরিসরই বিস্তৃতির সবচেয়ে উপযোগী পরিমাপ।

চতুর্থক ব্যবধান (Quartile Deviation): পরিসর কোন পরিশীলিত পরিমাপ নয়। এটি মাত্র দু'টি মানের ভিত্তিতে নির্ণীত হয়, যা মূলতঃ উপাত্তরাশির প্রান্তিক দু'টি মান। যেহেতু পরিসর নির্ণয়ে শুধুমাত্র সর্বনিম্ন ও সর্বোচ্চ মানের প্রয়োজন হয়, সেহেতু মধ্যবর্তী মানগুলো সম্পর্কে পরিসর কোন ধারণা দিতে পারে না। পরিসরের এই সীমাবদ্ধতা দূর করার জন্য চতুর্থক ব্যবধান নির্ণয় করা হয়। এটিও এক প্রকারের পরিসর, তবে উপাত্তরাশির উচ্চতম ও নিম্নতম মানের পার্থক্যের ভিত্তিতে 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। চতুর্থক ব্যবধান হলো, প্রথম ও তৃতীয় চতুর্থকের মধ্যকার ব্যবধানের অর্ধেক। মধ্যমা যেমন কেন্দ্রীয় প্রবণতার একটি অবস্থানমূলক পরিমাপ (positional measure), চতুর্থক ব্যবধান তেমনি বিস্তৃতির পরিমাপের একটি অবস্থানমূলক পরিমাপ। চতুর্থক ব্যবধান নির্ণয়ের সূত্রটি হলো,

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

গড় ব্যবধান হলো, গাণিতিক গড় থেকে প্রতিটি মানের অনপেক্ষ বিচ্যুতির গড়।

গড় ব্যবধান (Mean Deviation): পরিসর ও চতুর্থক ব্যবধান উপাত্তরাশির সবগুলো মানকে বিবেচনায় এনে বিস্তৃতির পরিমাপ করে না। বিস্তৃতির একটি উত্তম পরিমাপ হিসাবে বিবেচিত হতে হলে, উপাত্তরাশির সবগুলো সংখ্যা মানকে বিবেচনায় আনা প্রয়োজন। যদি আমরা সব মানকে ব্যবহার করতে চাই, তাহলে দু'টি প্রান্তিক বিন্দুর মধ্যকার ব্যবধানের দিকে দৃষ্টি নিবদ্ধ না করে, আমরা কেন্দ্রীয় প্রবণতার যে কোন একটি পরিমাপ থেকে উপাত্তরাশির প্রতিটি মানের অনপেক্ষ বিচ্যুতির (absolute difference) সমষ্টি নির্ণয় করে, তার একটি গড় নির্ণয় করতে পারি। অনপেক্ষ বিচ্যুতি নির্ণয়ের জন্য, কেন্দ্রীয় প্রবণতার পরিমাপ হিসাবে গাণিতিক গড়কেই ব্যবহার করা হয়ে থাকে। 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। সংজ্ঞা অনুযায়ী, গড় ব্যবধান নির্ণয়ের বীজগাণিতিক সূত্রটি হলো,

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

যেখানে, M.D. = গড় ব্যবধান

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

N = উপাত্তরাশির মোট সংখ্যা।

শ্রেণী বিন্যস্ত উপাত্তের ক্ষেত্রে গড় ব্যবধান নির্ণয়ের সূত্রটি হলো,

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

যেখানে, M.D. = গড় ব্যবধান

f_i = প্রতি শ্রেণীর গণসংখ্যা

N = মোট ঘটনসংখ্যা

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

পরিমিত ব্যবধান হ[...] গাণিতিক গড় থেকে [...] মানের বর্গকৃত বিচ্যু[...] বর্গমূল।

পরিমিত ব্যবধান (Standard Deviation): গড় ব্যবধান নির্ণয়ের সময়, ঋনাত্মক চিহ্ন উপেক্ষা করার জন্য যে অনপেক্ষ বিচ্যুতি নির্ণয় করা হয়, তা একটি কৃত্রিম উপায়। এর ফলে, তত্ত্বগতভাবে গড় ব্যবধানের কোন ব্যাখ্যা দেয়া যায় না এবং পরবর্তী বীজগাণিতিক পরিগণনায়ও প্রয়োগ করা যায় না। ঋণাত্মক চিহ্নকে উপেক্ষা করার এই কৃত্রিম পদ্ধতির পরিবর্তে, বিচ্যুতিগুলোর বর্গ করে নিলেই ঋণাত্মক

চিহ্নগুলো সহজেই এড়ানো যায় এবং বর্গকৃত বিচ্যুতির যোগফলের গড় নির্ণয় করে তার বর্গমূল নিলে, তত্ত্বগতভাবে ব্যাখ্যাযোগ্য একটি পরিমাপ নির্ণয় করা যায়। সেই পরিমাপটি হলো পরিমিত ব্যবধান। অর্থাৎ, পরিমিত ব্যবধান হলো, গাণিতিক গড় থেকে প্রতিটি মানের বর্গকৃত বিচ্যুতির গড়ের বর্গমূল। পরিমিত ব্যবধ্যান নির্ণয়ের বীজগাণিতিক সূত্রটি হলো,

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

যেখানে, s = পরিমিত ব্যবধান
$(x_i - \bar{x})^2$ = গড় থেকে প্রতিটি মানের বিচ্যুতির বর্গ
N = মোট ঘটনসংখ্যা।

বিন্যস্ত উপাত্ত থেকে 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
$(x_i - \bar{x})^2$ = গড় থেকে প্রতিটি শ্রেণী মধ্য-বিন্দুর বিচ্যুতির বর্গ
N = মোট ঘটনসংখ্যা।

উপাত্তরাশির গড় থেকে সংখ্যা মানগুলোর বিচ্যুতির বর্গের গড়ই হলো ভেদাঙ্ক।

ভেদাঙ্ক (Variance): উচ্চতর পরিসংখ্যানিক কাজে ব্যবহারের ক্ষেত্রে গাণিতিক গড় থেকে প্রতিটি মানের বর্গকৃত বিচ্যুতির গড়ের বর্গমূল না নিয়ে বর্গকৃত বিচ্যুতির গড় নির্ণয় করা হয়, যা ভেদাঙ্ক নামে পরিচিত। অর্থাৎ, উপাত্তরাশির গড় থেকে সংখ্যা মানগুলোর বিচ্যুতির বর্গের গড়ই হলো ভেদাঙ্ক। ভেদাঙ্ককে গড় বর্গও (mean square) বলা হয়ে থাকে। পরিমিত ব্যবধান ও ভেদাঙ্কের মধ্যে পার্থক্য হলো পরিমিত ব্যবধান অশোধিত মানের এককে থাকে, অথচ ভেদাঙ্ক থাকে বর্গকৃত এককে। বিস্তৃতির পরিমাপ হিসাবে পরিমিত ব্যবধান অধিক ব্যবহৃত হলেও পরিমিত ব্যবধানের তুলনায় ভেদাঙ্কের অধিকতর তাত্ত্বিক মূল্য রয়েছে। ভেদাঙ্ক দু'ভাবে নির্ণয় করা যায়। *প্রথমতঃ*, পরিমিত ব্যবধানের বর্গ করে এবং *দ্বিতীয়তঃ*, ভেদাঙ্ক নির্ণয়ের সংজ্ঞাগত সূত্র ব্যবহার করে। পরিমিত ব্যবধান ও ভেদাঙ্ক নির্ণয়ের পদ্ধতি একই রকম, শুধু একটি পার্থক্য ছাড়া। পরিমিত ব্যবধানের ক্ষেত্রে বিচ্যুতির বর্গের গড়ের বর্গমূল নিতে হয়, ভেদাঙ্কের ক্ষেত্রে তা করা হয় না। ভেদাঙ্ক নির্ণয়ের সংজ্ঞাগত সূত্রটি হলো,

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

যেখানে, s^2 = ভেদাঙ্ক
$(x_i - \bar{x})^2$ = গড় থেকে প্রতিটি মানের বিচ্যুতির বর্গ
N = উপাত্তরাশির মোট সংখ্যা

প্রশ্ন হলো যে, বিস্তৃতির শ্রেষ্ঠ পরিমাপ কোনটি? এক কথায় এর জবাব দেয়া সম্ভব নয়। কারণ, পরিমাপগুলোর প্রত্যয়গত অনন্যতা, গবেষণার উদ্দেশ্য, উপাত্তের প্রকৃতি এবং উপাত্ত বর্ণনার ক্ষেত্রে সেগুলোকে কিভাবে ব্যবহার করা হবে সে সকল বিষয়ের উপর সম্যক ধারণা নেবার পরই কেবলমাত্র এর জবাব দেয়া সম্ভব। বিস্তৃতির পরিমাপ হিসাবে পরিসরের সারল্য হলো এর প্রধান সুবিধা। চতুর্থক ব্যবধান প্রান্তিক মান দ্বারা কম প্রভাবিত হয় বলে এটি তুলনামূলকভাবে অধিকতর স্থিতিশীল পরিমাপ। কিন্তু বিস্তৃতির পরিমাপ হিসাবে চতুর্থক ব্যবধানের অসুবিধাটি হলো যে, এটি একটি বিন্যাসের প্রথম ২৫ শতাংশ এবং শেষ ২৫ শতাংশ উপাত্তরাশিকে কখনই বিবেচনায় আনে না। গড় ব্যবধান উপাত্তরাশির সবগুলো মানকে বিবেচনায় আনলেও এটি তত্ত্বগতভাবে ব্যাখ্যা করা যায় না বলে সামাজিক গবেষণায় এটির ব্যবহার খুব একটা লক্ষ্য করা যায় না।

ক্ষেত্র বিশেষে বিশেষ একটি পরিমাপ উপযুক্ত হলেও সার্বিক বিচারে বিস্তৃতির পরিমাপের অন্তর্নিহিত যুক্তি অনুযায়ী, একটি বিন্যাসের বিস্তার বর্ণনার জন্য পরিমিত ব্যবধান সবচেয়ে বেশী কার্যকর, উপযোগী এবং গ্রহণযোগ্য পরিমাপ। এ ক্ষেত্রে প্রধানতঃ চারটি কারণ উল্লেখ করা যেতে পারে। *প্রথমতঃ*, মূল উপাত্ত যে

এককে পরিমাপকৃত হয় পরিমিত ব্যবধান 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 হয়; এবং *চতুর্থতঃ*, পরিমিত ব্যবধানের বীজগাণিতিক বৈশিষ্ট্যগুলো এটিকে জটিল পরিসংখ্যানিক কর্মকান্ডে ব্যবহারের জন্য সুযোগ করে দেয়। তবে যেহেতু পরিমিত ব্যবধান নির্ণয়ে উপাত্তরাশির সকল মানকে বিবেচনায় আনা হয়, সেহেতু এটি প্রান্তিক 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।

কোন উপাত্তরাশির গড় এবং পরিমিত ব্যবধানের শতকরা অনুপাতকে সেই উপাত্তরাশির ব্যবধানাঙ্ক বলে।

ব্যবধানাঙ্ক (Coefficient of Variation): একটি বিন্যাসের গড় থেকে সংখ্যা মানগুলোর বিস্তৃতি পরিমাপের জন্য নিরঙ্কুশ বিস্তার পরিমাপগুলো উপযোগী হলেও, সেগুলো দুই বা ততোধিক উপাত্তরাশির মধ্যে তুলনার জন্য উপযোগী নয়। অতএব, দু'টি উপাত্তরাশির মধ্যে ভিন্নতার তুলনার জন্য নিরঙ্কুশ ব্যবধানগুলোকে আপেক্ষিক পদে রূপান্তরিত করতে হয়। আপেক্ষিক বিস্তার পরিমাপের একাধিক পদ্ধতি রয়েছে। তবে, এর মধ্যে সবচেয়ে গুরুত্বপূর্ণ এবং সর্বাধিক ব্যবহৃত পদ্ধতিটি হলো ব্যবধানাঙ্ক (coefficient of variation)। কোন উপাত্তরাশির গড় এবং পরিমিত ব্যবধানের শতকরা অনুপাতকে সেই উপাত্তরাশির ব্যবধানাঙ্ক বলে। অর্থাৎ,

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

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। বিপরীতভাবে, যে উপাত্তরাশির ব্যবধানাঙ্কের মান কম হবে সেই উপাত্তরাশির বিন্যাসটিকে কম বিস্তৃত এবং অধিক সমরূপ বলে গণ্য করা হয়। ধরা যাক, ক ও খ দু'টি গ্রামের নারীদের আয়ের গড় ও পরিমিত ব্যবধান যথাক্রমে $\bar{x}_ক$ = ৩,৫২০ টাকা ও $s_ক$ = ২,৮০০ টাকা এবং $\bar{x}_খ$ = ৩,৩২০ টাকা ও $s_খ$ = ২,০১০ টাকা। প্রদত্ত উপাত্ত ব্যবহার করে ব্যবধানাঙ্ক নির্ণয়ের মাধ্যমে দু'টি উপাত্তরাশির মধ্যে বিদ্যমান বিস্তৃতির তুলনা করা যেতে পারে। সূত্র অনুযায়ী, গ্রাম 'ক'-এর নারীদের আয়ের ব্যবধানাঙ্ক হলো,

$$C.V._ক = \frac{s}{\bar{x}} \times ১০০ = \frac{২,৮৪০}{৩,৫২০} \times ১০০ = ৮০.৬৮$$

এবং গ্রাম 'খ' এর মহিলাদের আয়ের ব্যবধানাঙ্ক হলো,

$$C.V._খ = \frac{s}{\bar{x}} \times ১০০ = \frac{২,০৪০}{৩,৫২০} \times ১০০ = ৬০.৫৪$$

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

বিন্যাসের রূপ ও আকৃতি (Forms and Shapes of a Distribution)

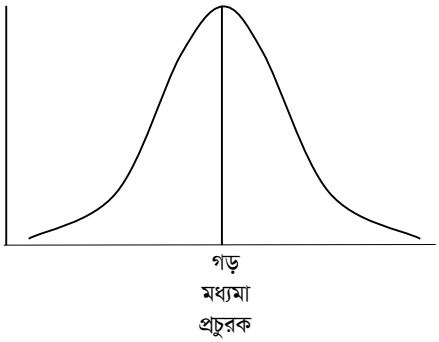
একটি বিন্যাসের রূপ সম্পর্কে ধারণা পাবার বঙ্কিমতা ও সূচালতার পরিমাপ করা হয়।

পরিসংখ্যান পদ্ধতির মাধ্যমে উপাত্ত রাশিমালার কেন্দ্রীভূত হবার প্রবণতা ও বিস্তৃতির মাত্রা পরিমাপের মাধ্যমে, একটি বিন্যাসের অন্তর্ভুক্ত মানগুলোর মধ্যে সমরূপতা ও ভিন্নতার চিত্রটি বোঝা যায়। কিন্তু বিন্যাসটির রূপ ও আকৃতি সম্পর্কে কোন স্পষ্ট ধারণা পাওয়া যায় না। অর্থাৎ, বিন্যাসটি সুষমভাবে বিন্যস্ত, না কি ডান বা বাম দিকে বঙ্কিম, তা জানা যায় না। অভিজ্ঞতালব্ধ বিন্যাসগুলো বিন্যাসের কতটুকু সন্নিকটবর্তী হয়েছে তা জানতে পারলে, নমুনা মানগুলোর ভিত্তিতে পরামানকে প্রাক্কলনের ক্ষেত্রে আমাদের আস্থা অনেক বেশী বাড়িয়ে দেয়। একটি বিন্যাসের রূপ ও আকৃতি সম্পর্কে ধারণা পাবার জন্য, বঙ্কিমতা ও সূচালতার মাত্রা পরিমাপ করা হয়।

বঙ্কিমতা (Skewness): একটি গণসংখ্যা বিন্যাস সম্পর্কে বর্ণনা দেবার জন্য কেন্দ্রীয় প্রবণতা ও বিস্তৃতির পরিমাপগুলো যথেষ্ট নয়। উপাত্ত রাশিমালা কিভাবে গড়ের চারপাশে বিন্যস্ত রয়েছে অথবা

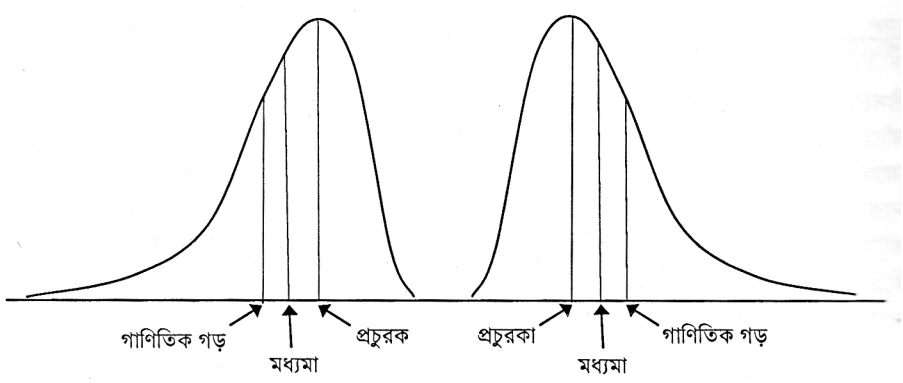
গড়ের উপরে বেশী উপাত্তরাশি রয়েছে, নাকি নীচে বেশী রয়েছে তা জানার প্রয়োজন রয়েছে। দু'টি বিন্যাসের একই গড় ও পরিমিত ব্যবধান থাকা সত্ত্বেও বিন্যাস দু'টির রূপ ও আকৃতি ভিন্ন হতে পারে। উপাত্তরাশির রূপ বর্ণনার জন্য বঙ্কিমতা একটি গুরুত্বপূর্ণ উপাদান। বঙ্কিমতা বলতে একটি বিন্যাসের অসমরূপতাকে (asymmetry) বোঝায়। একটি বিন্যাসের মধ্যে সুষমতার অভাব থাকলে সেই বিন্যাসকে বঙ্কিম বিন্যাস বলে। অর্থাৎ, উপাত্তরাশির মানগুলো গড়ের চারপাশে সুষমভাবে বিন্যস্ত রযেছে কি না, বঙ্কিমতা সেটিকেই বর্ণনা করে।

ধরা যাক, একটি বৃহৎ সংখ্যার উপাত্তরাশির গাণিতিক গড়, মধ্যমা এবং প্রচুরক-এর মান একই হয়েছে। প্রাপ্ত বিন্যাসটিকে একটি গণসংখ্যা রেখার (frequency curve) মধ্যে উপস্থাপন করলে দেখা যাবে যে, রেখাচিত্রটি বিন্যাসটিকে সর্বোচ্চ বিন্দুতে ঠিক সমান দু'ভাগে ভাগ করেছে। অর্থাৎ, প্রচুরকের উপরে এবং নীচে সমান সংখ্যক উপাত্তরাশি বিদ্যমান। শুধু তাই নয়, মধ্যমা ও গাণিতিক গড়ের উপরে এবং নীচেও সমান সংখ্যক সংখ্যা মান পাওয়া যাবে। অর্থাৎ, উপাত্তরাশির মানগুলো প্রচুরক, মধ্যমা ও গাণিতিক গড়ের দু'পাশে সুষমভাবে বিন্যস্ত রয়েছে (চিত্র ১০.২.১ দ্রষ্টব্য)।



চিত্র ১০.২.১: সুষম বিন্যাস

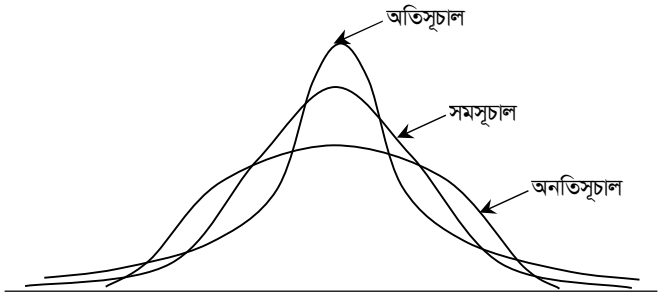
একটি ঋণাত্মকভাবে বঙ্কিম বিন্যাসে, প্রচুরক গাণিতিক গড়ের চেয়ে বড় হয় এবং মধ্যমা গাণিতিক গড় ও প্রচুরকের মধ্যে কোন এক স্থানে অবস্থান করে। গাণিতিক গড় নিম্নমান সম্পন্ন সংখ্যার দিকে অবস্থান করে। একটি বিন্যাসে যদি গাণিতিক গড় প্রচুরকের চেয়ে বড় হয় এবং মধ্যমা গড় ও প্রচুরকের মাঝখানে কোন এক স্থানে অবস্থান করে, তবে বলা যাবে যে, বিন্যাসটি ধনাত্মকভাবে বঙ্কিম হয়েছে। প্রচুরক, মধ্যমা ও গাণিতিক গড়ের এই সম্পর্কটি বঙ্কিমতার মাত্রা পরিমাপের ভিত্তি প্রদান করে থাকে (চিত্র ১০.২.২ক ও ১০.২.২খ দ্রষ্টব্য)।



চিত্র ১০.২.২ক: ঋণাত্মক বঙ্কিমতা *চিত্র ১০.২.২খ: ধনাত্মক বঙ্কিমতা*

সূচালতা একটি গণসংখ্যা রেখার প্রচুরক থেকে সূচ্যগ্রতার মাত্রার বর্ণনা দেয়।

সূচালতা (Kurtosis): সূচালতা একটি গণসংখ্যা রেখার প্রচুরক থেকে সূচ্যগ্রতার (peakedness) মাত্রার বর্ণনা দেয়। একটি গণসংখ্যা রেখা যদি স্বাভাবিক রেখার তুলনায় বেশী সূচ্যগ্র হয়, তবে সেই রেখাটিকে অতিসূচাল (leptokurtic) রেখা বলে। যদি রেখাটি স্বাভাবিক রেখার তুলনায় কম সূচ্যগ্র হয়, তবে তাকে অনতিসূচাল (platykurtic) রেখা বলে। একটি সুষম রেখাকে সমসূচাল (mesokurtic) রেখা বলে। এই রেখাটি একটি এক প্রচুরক সম্পন্ন বিন্যাসের রূপ নেয়, যেখানে রেখার মাঝখানটি উঁচু থাকে এবং প্রান্ত দু'টি ঢালু হয়ে অনুভূমিক রেখায় নেমে যায় তবে স্পর্শ করে না। অর্থাৎ, স্বাভাবিক রেখার সকল বৈশিষ্ট্য সমসূচাল রেখার মধ্যে থাকে (চিত্র ১০.২.৩ দ্রষ্টব্য)।



চিত্র ১০.২.৩: সূচালতার বিভিন্ন রূপ

আমরা জানি যে, পরিঘাত একটি বিন্যাসের আকার ও প্রকৃতি সম্পর্কে বর্ণনা দেয়। দ্বিতীয়, তৃতীয় ও চতুর্থ পরিঘাত থেকে নির্ণীত দু'টি সহগ ব্যবহার করে বঙ্কিমতা ও সূচালতার মাত্রা পরিমাপ করা হয়। সেই সহগ দু'টি হলো, $\beta_১$ (বিটা$_১$) এবং $\beta_২$ (বিটা$_২$)। $\beta_১$ বঙ্কিমতার পরিমাপ হিসাবে ব্যবহৃত হয়। $\beta_১$ নির্ণয়ের সূত্রটি হলো,

$$\beta_১ = \frac{\mu_৩^{২}}{\mu_২^{৩}}$$

একটি সুষম বিন্যাসে $\beta_১$ সবসময় '০' হবে। তবে বঙ্কিমতার পরিমাপ হিসাবে $\beta_১$-এর সীমাবদ্ধতাটি হলো যে, এটি বঙ্কিমতার দিক (direction) সম্পর্কে আমাদের কোন ধারণা দেয় না। এই সীমাবদ্ধতাটি দূরীভূত হয় যদি আমরা কার্ল পিয়ারসন-এর $\gamma_১$ (গ্যামা$_১$) নির্ণয় করি। $\gamma_১$-কে $\beta_১$-এর বর্গমূল হিসাবে সংজ্ঞায়িত করা যায়। অর্থাৎ,

$$\gamma_১ = \sqrt{\beta_১} = \frac{\mu_৩}{\mu_২^{৩/২}} = \frac{\mu_৩}{\sigma^৩}$$

যেহেতু বঙ্কিমতার চিহ্নটি নির্ভর করে তৃতীয় পরিঘাতের ($\mu_৩$) মানের উপর, সেহেতু যদি $\mu_৩$-এর মান ধনাত্মক হয়, তবে বিন্যাসটি ধনাত্মকভাবে বঙ্কিম হবে। আর যদি ঋণাত্মক হয়, তবে বিন্যাসটি

ঋণাত্মকভাবে বঙ্কিম হবে। $\beta_২$-কে সূচালতার পরিমাপ হিসাবে ব্যবহার করা হয়। $\beta_২$ নির্ণয়ের বীজগাণিতিক সূত্রটি হলো,

$$\beta_২ = \frac{\mu^৪}{\mu_২^২}$$

$$\gamma_২ = \beta_২ - ৩$$

একটি সুষম ও সমসূচাল বিন্যাসের ক্ষেত্রে $\beta_১ = ০$ এবং $\beta_২ = ৩$ হয়। যদি $\beta_২ < ৩$ হয়, তবে বিন্যাসটির আকৃতি অতিসূচাল হবে, এবং $\beta_২ > ৩$ হলে, অনতিসূচাল হবে।

এক গুচ্ছ পর্যবেক্ষণের মধ্যে একটি পর্যবেক্ষণ কিভাবে এবং কতটুকু ভিন্ন, তা আমাদের জানার প্রয়োজন হতে পারে। পরিমিত মান নির্ণয়ের মাধ্যমে আমরা তা জানতে পারি।

পরিমিত মান (Standard Score): পরিমিত ব্যবধান আমাদেরকে গড় থেকে সংখ্যা মানগুলোর গড় বিস্তৃতি সম্পর্কে ধারণা দেয়। কিন্তু এক গুচ্ছ পর্যবেক্ষণের মধ্যে একটি পর্যবেক্ষণ কিভাবে এবং কতটুকু ভিন্ন তা আমাদের জানার প্রয়োজন হতে পারে। পরিমিত মান নির্ণয়ের মাধ্যমে আমরা তা জানতে পারি। পরিমিত মানকে z-মানও বলা হয়ে থাকে।

একটি বিন্যাসের প্রতিটি সংখ্যা মানেরই একটি পরিমিত মান রয়েছে, যা সেই বিন্যাসের গড় ও পরিমিত ব্যবধানের মানের উপর নির্ভর করে। 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 উদাহরণ দিয়ে গড় থেকে যে কোন সংখ্যা মানের মধ্যেকার আয়তন নির্ণয় করে দেখা যাক। ধরা যাক, একটি স্বাভাবিক রেখার গড় হলো, ৫০ এবং পরিমিত ব্যবধান হলো, ১০। এখন, ৫০ থেকে ৬৬ অক্ষের মধ্যে কি পরিমাণ পর্যবেক্ষণ রয়েছে বা কতটুকু আয়তন রয়েছে তা দেখা যাক। প্রথমে আমাদের দেখতে হবে যে, ৫০ থেকে ৬৬ কত পরিমিত ব্যবধান দূরত্বে রয়েছে। এটি নির্ণয় করতে গিয়ে নিম্নোক্ত সূত্রটি ব্যবহার করা হয়,

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

যেখানে, z = পরিমিত মান

x_i = প্রতিটি অশোধিত মান

$\bar{x}$ = গাণিতিক গড়

s = পরিমিত ব্যবধান।

অতএব, $z = \frac{৬৬ - ৫০}{১০} = \frac{১৬}{১০} = ১.৬$

অর্থাৎ, ৫০ থেকে ৬৬ মানের দূরত্ব হলো, ১.৬ পরিমিত ব্যবধান। স্বাভাবিক রেখার অধীনে, ১.৬ পরিমিত ব্যবধানের আয়তন হলো, ০.৪৪৫২ (সারণি পরি ২ দ্রষ্টব্য)। অর্থাৎ, ৫০ থেকে ৬৬ মানের মধ্যে ৪৪.৫২ শতাংশ পর্যবেক্ষণ অন্তর্ভুক্ত রয়েছে। যেহেতু ৫০ থেকে ৬৬ মানটি ডান দিকে ১.৬ পরিমিত ব্যবধান দূরত্বে রয়েছে, সেহেতু মানটি ধনাত্মক হয়েছে। কিন্তু যদি ৫০ থেকে ৩৪ মানের মধ্যেকার আয়তন নির্ণয় করতে হয়, সে ক্ষেত্রে প্রাপ্ত পরিমিত ব্যবধানের মানটি ঋণাত্মক হবে, অর্থাৎ, –১.৬। অন্য কথায়, ৫০ থেকে ৬৬ মানটি গড় থেকে ১.৬ পরিমিত ব্যবধান উপরে এবং ৫০ থেকে ৩৪ মানটি –১.৬ পরিমিত ব্যবধান নীচে অবস্থান করছে।

সারাংশ

উপাত্ত প্রক্রিয়াকরণ ও সংক্ষিপ্তকরণের পর, সেগুলো বর্ণনার জন্য প্রথম পদক্ষেপ হিসাবে এক-চলক বিশিষ্ট বিশ্লেষণের উদ্যোগ নেয়া হয়। উপাত্তের বৈশিষ্ট্যগুলো বর্ণনায় একটি মাত্র চলককে ব্যবহার করা হয়। এক-চলক বিশিষ্ট বর্ণনায় যে সব পরিসংখ্যানগত পরিমাপ ব্যবহৃত হয়, সেগুলো হলো, কেন্দ্রীয় প্রবণতার বিভিন্ন পরিমাপ, বিস্তৃতির পরিমাপ এবং বিন্যাসের রূপ ও আকৃতির পরিমাপ। বর্তমান পাঠে, এক-চলক বিশিষ্ট বিশ্লেষণের মাধ্যমে উপাত্তের বর্ণনার ক্ষেত্রে বিভিন্ন পরিসংখ্যানের ব্যবহারগুলো বিস্তারিতভাবে আলোচনা করা হয়েছে। মূল বক্তব্যগুলোকে সহজ করে তোলার জন্য এই বিস্তারিত আলোচনার আশ্রয় নেয়া হয়েছে।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক ($\surd$) চিহ্ন দিন –

১। যে কোন প্রকার উপাত্ত বিশ্লেষণের প্রাথমিক পদক্ষেপ হলো:
   ক. উপাত্তের তালিকা করা
   খ. উপাত্তের সারমর্ম করা
   গ. উপাত্তের গণসংখ্যা সারণি তৈরি করা
   ঘ. উপাত্তের সংখ্যা বর্ণনা করা।

২। উপাত্ত বিশ্লেষণ করতে গিয়ে গবেষক কোন পরিমাপটি ব্যবহার করবেন, তা নির্ভর করে:
   ক. উপাত্তরাশির প্রকৃতির উপর
   খ. গবেষণার বিষয়বস্তুর উপর
   গ. উপাত্ত বিশ্লেষণের প্রকৃতির উপর
   ঘ. উপরের সব।

৩। অনপেক্ষ বিচ্যুতি নির্ণয়ের জন্য কেন্দ্রীয় প্রবণতার পরিমাপ হিসাবে:
   ক. গাণিতিক গড়কে ব্যবহার করা হয়
   খ. গড় ব্যবধানকে ব্যবহার করা হয়
   গ. পরিমিত ব্যবধান ব্যবহার করা হয়
   ঘ. ভেদাঙ্ককে ব্যবহার করা হয়।

সংক্ষিপ্ত প্রশ্ন

টীকা লিখুন:

   ক. গাণিতিক গড় কী?
   খ. ব্যবধানাঙ্ক কী?

রচনামূলক প্রশ্ন

১। গবেষণার ফলাফল বর্ণনায় এক-চলক বিশিষ্ট বিশ্লেষণের গুরুত্ব উদাহরণসহ আলোচনা করুন।

২। উপাত্ত বিশ্লেষণে কেন্দ্রীয় প্রবণতার পরিমাপ ও বিস্তৃতির পরিমাপের গুরুত্ব আলোচনা করুন।

পাঠ - ৩

দ্বি-চলক বিশিষ্ট বিশ্লেষণ
Bivariate Analysis

এই পাঠ শেষে যা জানা যাবে —

- দ্বি-চলক বিশিষ্ট বিশ্লেষণ কি
- দ্বি-চলক বিশিষ্ট সারণির মাধ্যমে সম্পর্কের বিশ্লেষণ
- বিক্ষেপ চিত্রের মাধ্যমে সম্পর্কের প্রদর্শন
- সহ-সম্পর্কের সহগ নির্ণয়
- নির্ভরণ বিশ্লেষণ

দ্বি-চলক বিশিষ্ট বিশ্লেষণ কি (What is Bivariate Analysis)?

'ক' ও 'খ' এর মধ্যে সম্পর্ক রয়েছে এমন প্রস্তাবনা তখনই করা যাবে যখন দেখা যাবে যে, 'ক' চলকের কিছু বৈশিষ্ট্য নিয়মতান্ত্রিকভাবে 'খ' চলকের কিছু বৈশিষ্ট্যের সাথে সহগামী হয়েছে।

উপাত্তের বিশ্লেষণে এক-চলক বিশিষ্ট বিন্যাসের বর্ণনার পর, গবেষণা সমস্যার সাথে সংশ্লিষ্ট চলকগুলোর মধ্যে সম্পর্কের রূপটিকে পরীক্ষা করার প্রয়োজন হয়ে পড়ে। আমরা জানি যে, আমাদের চারপাশের জগতে যা কিছু রয়েছে, তার সব কিছু কোন না কোনভাবে সম্পর্কিত। যেমন, একটি শিশু বেড়ে উঠার সাথে সাথে তার ওজন বৃদ্ধি পায়, গ্রামের তুলনায় শহরে পরিবেশ দূষণ বেশী হয়ে থাকে, শিক্ষা বৃদ্ধির সাথে সাথে আয় বৃদ্ধি পায়, ইত্যাদি। এ ধরণের প্রতিটি পর্যবেক্ষণ দু'টি ঘটনার মধ্যে একটি সম্পর্ককে নির্দেশ করে ড় বয়স এবং ওজনের মধ্যে, নগরায়ণ ও পরিবেশ দূষণের মধ্যে এবং শিক্ষা ও আয়ের মধ্যে। অর্থাৎ, 'ক' ও 'খ' এর মধ্যে সম্পর্ক রয়েছে, এমন প্রস্তাবনা তখনই করা যাবে, যখন দেখা যাবে যে, 'ক' চলকের কিছু বৈশিষ্ট্য নিয়মতান্ত্রিকভাবে 'খ' চলকের কিছু বৈশিষ্ট্যের সাথে সহগামী হয়েছে। যখন চলকগুলো একটির সাথে আরেকটি সম্পর্কিত হয় না, তখন আমরা বলি যে, তারা একটি থেকে আরেকটি স্বাধীন। অর্থাৎ, তারা সহগামী নয়। যদিও গবেষণার গভীরতর বিশ্লেষণে, বহু চলকের মধ্যে সম্পর্ক পরীক্ষা করে দেখা হয়, এ পর্যায়ে আমরা সম্পর্কের বিশ্লেষণকে দু'টি চলকের মধ্যেই সীমাবদ্ধ রাখবো। অতএব, এক কথায় বলা যায় যে, দ্বি-চলক বিশিষ্ট বিশ্লেষণ হলো, দু'টি চলকের মধ্যে সম্পর্কের উপস্থিতি ও অনুপস্থিতির বর্ণনা এবং ব্যাখ্যা।

দ্বি-চলক বিশিষ্ট বিশে- ষণ হলো, দু'টি চলকের মধ্যে সম্পর্কের উপস্থিতি ও অনুপস্থিতির বর্ণনা এবং ব্যাখ্যা।

দ্বি-চলক বিশিষ্ট সারণির মাধ্যমে সম্পর্কের বিশ্লেষণ (Analysing Relationship through a Bivariate Table)

একটি দ্বি-চলক বিশিষ্ট সারণির প্রতিটি স্তম্ভের শ্রেণী-দফাকে আলাদাভাবে এক একটি এক-চলক বিশিষ্ট ঘটনসংখ্যার বিন্যাস হিসাবে বিবেচনা করা যায়।

দু'টি চলকের মধ্যে সম্পর্ক রয়েছে কি না, তা পরীক্ষার জন্য প্রথম পদক্ষেপটি হলো, একটি দ্বি-চলক বিশিষ্ট সারণি গঠন। এ ধরণের বিশ্লেষণকে সাপেক্ষ সারণি বিশ্লেষণও (contingency table analysis) বলা হয়ে থাকে। আমরা ইউনিট ৯ -এর পাঠ ৩ -এ জেনেছি যে, একটি দ্বি-চলক বিশিষ্ট সারণিতে, একটি চলকের শ্রেণী-দফাগুলো সারিভিত্তিক এবং অন্য চলকের শ্রেণী-দফাগুলোকে স্তম্ভভিত্তিক আড়া-আড়িভাবে উপস্থাপন করা হয়। স্তম্ভভিত্তিক সাজানো চলকটিকে স্বাধীন চলক (যা সারণির উপরে উপস্থাপন করা হয়) এবং সারিভিত্তিক চলকটিকে নির্ভরশীল চলক (সারণির বাম পাশে উপস্থাপন করা হয়) বলে। একটি দ্বি-চলক বিশিষ্ট সারণির প্রতিটি স্তম্ভের শ্রেণী-দফাকে আলাদাভাবে এক একটি এক-চলক বিশিষ্ট ঘটনসংখ্যার বিন্যাস হিসাবে বিবেচনা করা যায়।

ধনাত্মক সম্পর্কের অর্থ হলো যে, একটি চলকের মান বৃদ্ধির সাথে সাথে অপর চলকটিরও মান বৃদ্ধি পাওয়া এবং ঋণাত্মক সম্পর্ক বলতে দু'টি চলকের মধ্যে বিপরীতমূখী সম্পর্ককে বোঝায়।

সারণি ১০.৩.১-এ, ১৪০ জন উচ্চপদস্থ সরকারী কর্মকর্তার লিঙ্গভিত্তিক বৈবাহিক মর্যাদার একটি বিন্যাসকে উপস্থাপন করা হয়েছে। এই সারণিতে, লিঙ্গ চলকের একটি মাত্রায় রয়েছে, দু'টি শ্রেণী-দফা — নারী ও পুরুষ এবং বৈবাহিক মর্যাদার অন্য মাত্রায় রয়েছে, চারটি শ্রেণী-দফা — বিবাহিত, অবিবাহিত, তালাকপ্রাপ্ত/তালাকপ্রাপ্তা এবং বিধবা/বিপত্নীক। অর্থাৎ, এটিতে পর্যবেক্ষণগুলোকে চারটি সারি এবং দু'টি স্তম্ভে সাজানো হয়েছে, যার প্রতিটি এক একটি শ্রেণী-দফাকে প্রতিনিধিত্ব করে এবং যার প্রতিটিকে, এক একটি এক-চলক বিশিষ্ট ঘটনসংখ্যার বিন্যাস হিসাবে গণ্য করা যায়। পঞ্চম সারিটি স্তম্ভের মোট সংখ্যা এবং তৃতীয় স্তম্ভটি সারির মোট সংখ্যাকে প্রতিনিধিত্ব করে। এই সারণিটিকে ৪×২ সারণি বলা হয়। সারণি ১০.৩.১ অনুযায়ী, উচ্চপদস্থ সরকারী কর্মকর্তাদের লিঙ্গ এবং বৈবাহিক মর্যাদার মধ্যে একটি ভিন্নতার সম্পর্ককে স্পষ্টভাবে প্রদর্শন করে। এখানে দেখা যাচ্ছে যে, দু'টি এক-চলক বিশিষ্ট বিন্যাসের রূপ (নারী ও পুরুষ) একটির সাথে আরেকটি ভিন্নতা প্রদর্শন করে। যেমন, নারীরা পুরুষদের তুলনায় কম বিবাহিত এবং অধিকহারে তালাকপ্রাপ্তা বা বিধবা।

সারণি ১০.৩.১: উচ্চ পদস্থ সরকারী কর্মকর্তাদের লিঙ্গভিত্তিক বৈবাহিক মর্যাদার বিন্যাস।

বৈবাহিক মর্যাদা	লিঙ্গ		মোট
	নারী	পুরুষ	
অবিবাহিত	২	২	৪
বিবাহিত	৪৯	৫৯	১০৮
তালাকপ্রাপ্ত/তালাকপ্রাপ্তা	১০	৫	১৫
বিধবা/ বিপত্নীক	৯	৪	১৩
মোট	৭০	৭০	১৪০

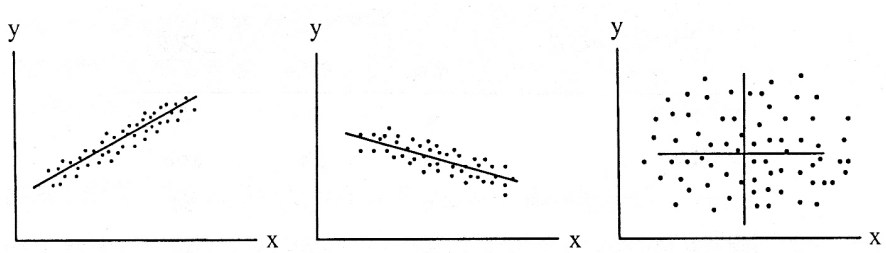
দু'টি চলকের প্রতি জোড়া পর্যবেক্ষণ লেখচিত্রে উপস্থাপন করে সহ-সম্পর্কের বিক্ষেপ চিত্র পাওয়া যায় এবং সহ-সম্পর্কের সহগ নির্ণয় করে চলকদ্বয়ের সহ-সম্পর্কের দিক ও মাত্রা সুনির্দিষ্টভাবে নির্ণয় করা যায়।

এতক্ষণ পর্যন্ত, দু'টি চলকের মধ্যে সহগামীতার মাত্রাটি মূল্যায়িত হয়েছে, একটি দ্বি-চলক বিশিষ্ট সারণির দু'টি এক-চলক বিশিষ্ট বিন্যাসকে মূল্যায়নের মাধ্যমে। কিন্তু যদি চলকগুলোর শ্রেণী-দফার পরিমাণ অনেক বেশী হয় (যেমন, একটি ১২দ্ধ৮ সারণি), সে ক্ষেত্রে সহগামীতার মাত্রার এ ধরণের মূল্যায়ন অনেক সময় পাঠককে নিরুৎসাহিত করতে পারে। কারণ, এতে করে সহগমনের প্রকৃত চিত্রটি ঝাপসা হয়ে যায়। এ ছাড়া, সারণির মাধ্যমে দ্বি-চলক বিশিষ্ট বিশ্লেষণের দু'টি চলকের মধ্যে সম্পর্কের একটি স্থূল ধারণা পাওয়া গেলেও সম্পর্কের দিক ও মাত্রা সম্পর্কে সুস্পষ্ট কোন ধারণা পাওয়া যায় না। এ সব সমস্যা থেকে পরিত্রাণ পাবার উপায় হিসাবে সহগামীতার একটি একক সংক্ষিপ্ত পরিমাপের প্রয়োজন। চলকসমূহের মধ্যে সম্পর্ককে অনুধাবন, পরিমাপ ও বিশ্লেষণের জন্য বিভিন্ন পদ্ধতি ব্যবহৃত হয়। এই পদ্ধতিগুলোর মধ্যে সবচেয়ে বেশী পরিচিত দু'টি পদ্ধতি হলো: বিক্ষেপ চিত্র অঙ্কন (scatter diagram) এবং সহ-সম্পর্কের সহগ নির্ণয় (computation of coefficient of correlation)। দু'টি চলকের প্রতি জোড়া পর্যবেক্ষণ লেখচিত্রে উপস্থাপন করে সহ-সম্পর্কের বিক্ষেপ চিত্র পাওয়া যায় এবং সহ-সম্পর্কের সহগ নির্ণয় করে, চলকদ্বয়ের সহ-সম্পর্কের দিক ও মাত্রা সুনির্দিষ্টভাবে নির্ণয় করা যায়।

বিক্ষেপ চিত্রের মাধ্যমে সহ-সম্পর্কের প্রদর্শন (Demonstration of Correlationship through Scatter Diagram)

যে লেখচৈত্রিক কৌশল, দু'টি চলকের মধ্যে বিদ্যমান সম্পর্ককে চাক্ষুষভাবে পর্যবেক্ষণের একটি প্রত্যক্ষ সুযোগ প্রদান করে, তাকে বিক্ষেপ চিত্র বলে। বিক্ষেপ চিত্র অঙ্কনের জন্য দু'টি চলকের কিছু জোড়া পর্যবেক্ষণের (paired observations) প্রয়োজন হয়, যেখানে সাধারণতঃ একটি চলককে x এবং অপরটিকে y ধরা হয়। যদি প্রতিটি x মানের জন্য একটি করে y মান থাকে, তবে x ও y চলকের জোড়া সংখ্যা মানগুলোকে বিন্দুর মাধ্যমে চিহ্নিত করে লেখচিত্রে উপস্থাপন করলে, একটি বিক্ষেপ চিত্র অঙ্কিত হয়। জোড়া বিন্দুগুলো কিভাবে গুচ্ছবদ্ধ রয়েছে, বা বিক্ষিপ্ত রয়েছে, তা যাচাই করে দু'টি চলকের মধ্যেকার সম্পর্কের প্রকৃতি সম্পর্কে বর্ণনা করা যায়।

x ও y চলকের জোড়া মানের বিন্দুগুলি যদি এমনভাবে বিন্যস্ত হয় যে, সেগুলো যোগ করলে একটি সরল রেখা অঙ্কিত হয়, তবে তা চলক দু'টির মধ্যে নিখুঁত, বা পূর্ণ সহ-সম্পর্ক নির্দেশ করবে। জোড়া মানের বিন্দু নির্দেশক সরল রেখাটি যদি ডান দিকে উর্ধ্বগামী হয়, তবে তা ধনাত্মক সম্পর্ক, এবং যদি ডান দিকে নিম্নগামী হয়, তবে তা ঋণাত্মক সম্পর্ককে নির্দেশ করে। অর্থাৎ, ধনাত্মক সম্পর্কের অর্থ হলো যে, একটি চলকের মান বৃদ্ধির সাথে সাথে অপর চলকটিরও মান বৃদ্ধি পাওয়া এবং ঋণাত্মক সম্পর্ক বলতে দু'টি চলকের মধ্যে বিপরীতমূখী সম্পর্ককে বোঝায়। লেখচিত্রে উপস্থাপিত বিন্দুগুলি যদি কোন সুনির্দিষ্ট রূপ প্রদর্শন না করে এলোমেলো বা বিক্ষিপ্তভাবে অবস্থান করে, অথবা বিন্দুগুলি যোগ করে যদি x -অক্ষ বা y -অক্ষের সমান্তরাল একটি রেখা পাওয়া যায়, তবে তা চলকদ্বয়ের সম্পর্কহীনতাকে নির্দেশ করে। অর্থাৎ, একটি চলকের মানের পরিবর্তনের সাথে সাথে যদি অপর চলকের মানের মধ্যে কোন পরিবর্তন না হয়, তবে বুঝতে হবে যে, চলক দু'টি সম্পর্কহীন (চিত্র ১০.৩.১ক, খ ও গ দ্রষ্টব্য)।



চিত্র ১০.৩.১ ক, খ ও গ: বিক্ষেপ চিত্রে প্রদর্শিত ধনাত্মক, ঋণাত্মক ও শূন্য সম্পর্ক সম্পর্ক

বিক্ষেপ চিত্রের সবচেয়ে বড় সুবিধা হলো যে, এটি দু'টি চলকের মধ্যে সম্পর্ক প্রদর্শনের জন্য সবচেয়ে সহজ, সাধারণ ও অগাণিতিক একটি পদ্ধতি। খুব সাধারণ পরিসংখ্যানগত জ্ঞান থাকলে, যে কেউ বিক্ষেপ চিত্রের সাহায্যে দুটি চলকের সম্পর্ক সম্বন্ধে একটি প্রাথমিক ধারণা নিতে পারেন। এ ছাড়া, অধিকাংশ গাণিতিক পদ্ধতির মত, বিক্ষেপ চিত্র একটি বিন্যাসের চরম মান দ্বারা প্রভাবিত হয় না। বিক্ষেপ চিত্রের অসুবিধাটি হলো যে, এটির মাধ্যমে দু'টি চলকের মধ্যে সহ-সম্পর্কের দিক ও মাত্রা সম্পর্কে একটি সাধারণ ধারণা পাওয়া গেলেও, সূক্ষ্মভাবে তা নির্ণয় করা যায় না। সে ক্ষেত্রে, গাণিতিক সহগ নির্ণয় পদ্ধতি ব্যবহার করা অপরিহার্য হয়ে পড়ে।

সহ-সম্পর্কের সহগ নির্ণয় (Computing Coefficient of Correlation)

একটি চলকের সাহায্যে আরেকটি চলকের ভবিষ্যদ্বাণীকরণে ত্রুটির মাত্রা কতটুকু কমে, বা ভবিষ্যদ্বাণীর উন্নয়নে কতটুকু ভূমিকা রাখে, ল্যাম্বডা সহগ তার মাত্রাকে প্রকাশ করে।

ল্যাম্বডা সহগ (Lambda Coefficient): নামসূচক মাত্রায় পরিমাপকৃত দু'টি চলকের মধ্যে সম্পর্ক নির্ণয়ের জন্য একটি জনপ্রিয় পরিমাপ হলো 'ভবিষ্যদ্বাণীমূলক গাটম্যান সহগ' (Guttman coefficient of predictibility), বা ল্যাম্বডা সহগ (Lambda coefficient), যা গ্রীক বর্ণমালার 'λ' অক্ষরটি দিয়ে চিহ্নিত করা হয়। ল্যাম্বডা সহগ একটি অপ্রতিসাম্য (asymmetrical) পরিমাপ এবং এটি 'ভ্রান্তির সমানুপাতিক হ্রাসকরণ' (proportional reduction in error) নীতিমালার উপর প্রতিষ্ঠিত। একটি চলকের সাহায্যে আরেকটি চলকের ভবিষ্যদ্বাণীকরণে ত্রুটির মাত্রা কতটুকু কমে, বা ভবিষ্যদ্বাণীর উন্নয়নে কতটুকু ভূমিকা রাখে, ল্যাম্বডা সহগ তার মাত্রাকে প্রকাশ করে। ল্যাম্বডার মান '০' থেকে '১' পর্যন্ত বিস্তৃত। যদি একটি চলক থেকে আরেকটি চলকের ভবিষ্যদ্বাণীকরণে ত্রুটির মাত্রা না কমে, তখন ল্যাম্বডার মান '০' হয়। ল্যাম্বডার সীমাবদ্ধতাটি হলো যে, যখন ঘটনসংখ্যাগুলো স্বাধীন চলকের একটি শ্রেণী-দফায় ঘনীভূত হয়, তখন সহ-সম্পর্কের পরিমাপ হিসাবে এটি উপযোগীতা হারিয়ে ফেলে। কারণ, এ ধরণের পরিস্থিতিতে প্রকৃতপক্ষে দু'টি চলক সম্পর্কিত হলেও ল্যাম্বডার মান '০' হবে। ল্যাম্বডা নির্ণয়ের সূত্রটি হলো,

$$\lambda = \frac{\sum m_y - M_y}{N - M_y}$$

যেখানে, λ = ল্যাম্বডা সহগ

M_y = নির্ভরশীল চলক, y -এর সার্বিক প্রচুরক ঘটনসংখ্যা

m_y = স্বাধীন চলক, y -এর ভিন্ন ভিন্ন শ্রেণী-দফার মধ্যে নির্ভরশীল চলক, y -এর উপর প্রচুরক ঘটনসংখ্যা

N = নমুনার মোট সংখ্যা।

কাই-বর্গ পরীক্ষা (Chi-Square Test): একটি দ্বি-চলক বিশিষ্ট সারণিতে উপস্থাপিত দু'টি চলকের একটি যখন নামসূচক মাত্রায় এবং অন্যটি ক্রমসূচক মাত্রায় পরিমাপকৃত হয়, তখন সে দু'টি চলকের মধ্যে সম্পর্কহীনতার অনুমানের ভিত্তিতে, পর্যবেক্ষণকৃত ঘটনসংখ্যার সাথে প্রত্যাশিত ঘটনসংখ্যার তুলনা করে, একটি স্বাধীনতার তাৎপর্য পরীক্ষা করে দেখতে হবে যে, চলক দু'টোর একটি থেকে আরেকটি স্বাধীন কি না। যে সম্ভাবনা বিন্যাস ব্যবহার করে পরীক্ষাটি পরিচালনা করা হয়, সেটি কাই-বর্গ (χ^2) বিন্যাস নামে পরিচিত। বিন্যাস নামে পরিচিত। কাই-বর্গ বিন্যাস হলো, ডানদিকে বঙ্কিম একটি অবিচ্ছিন্ন সম্ভাবনা বিন্যাস, যা ধনাত্মক দিকে শূন্য থেকে অসীম মান পর্যন্ত বিস্তৃত। এর কোন ঋণাত্মক মান নেই। কারণ, কাই-বর্গ মানটি প্রত্যাশিত এবং পর্যবেক্ষণমূলক ঘটনসংখ্যার পার্থক্যকে বর্গ করে নির্ণীত হয়। অর্থাৎ, প্রত্যাশা ও পর্যবেক্ষণের মধ্যে যে ব্যবধান তৈরি হয়, সেটি কি পর্যবেক্ষণকৃত মানের সাথে তত্ত্বের সামঞ্জস্যতার অভাবে হয়েছে, না কি দৈবক্রমে ঘটেছে, কাই-বর্গ পরীক্ষার মাধ্যমে আমরা তা জানতে পারি। যদি কাই-বর্গের মান '০' হয়, তবে বুঝতে হবে যে, প্রত্যাশিত ও পর্যবেক্ষণকৃত ঘটনসংখ্যার মধ্যে হুবহু মিল হয়েছে। কাই-বর্গের মান যত বড় হবে, প্রত্যাশিত ও পর্যবেক্ষণকৃত ঘটনসংখ্যার মধ্যে ব্যবধান তত বেশী রয়েছে বলে ধরে নিতে হবে। কাই-বর্গ নির্ণয়ের বীজগাণিতিক সূত্রটি হলো,

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e}$$

যেখানে, χ^2 = কাই-বর্গ

f_0 = পর্যবেক্ষণকৃত ঘটনসংখ্যা
f_e = প্রত্যাশিত ঘটনসংখ্যা।

ফাই সহগ (Phi Coefficient): কাই-বর্গের উপর ভিত্তি করে সম্পর্কের মাত্রাকে বোঝার জন্য ফাই সহগ নির্ণয় করা হয়। একটি ২×২ সারণিতে উপস্থাপিত দু'টি দ্বি-বিভাজিত (dichotomous) চলকের মধ্যে সম্পর্কের মাত্রা পরিমাপের জন্য ফাই সহগ হলো একটি উপযোগী পরিমাপ। ফাই সহগ নির্ণয়ের সূত্রটি হলো,

$$\varphi = \left(\frac{\chi^2}{N}\right)^{\frac{১}{২}}$$

যেখানে, φ = ফাই সহগ
χ^2 = কাই-বর্গ
N = মোট ঘটনসংখ্যা।

ফাই সহগের মান '০' থেকে '১' পর্যন্ত বিস্তৃত। যদি দু'টি চলকের মধ্যে কোন সম্পর্ক না থাকে, তবে $\varphi = ০$ হয় এবং পূর্ণ ধনাত্মক সম্পর্ক হলে $\varphi = ১$ হয়। ফাই সহগের সমস্যা হলো যে, যেহেতু ২×২ সারণির কোন উচ্চসীমা নেই, সেহেতু ২×২ সারণির চেয়ে বড় সারণির ক্ষেত্রে এটি প্রয়োগ করা যায় না।

ক্রমমূলক মাত্রায় পরিমাপকৃত দু'টি চলকের মধ্যে সম্পর্ক পরিমাপের জন্য যে পরিমাপটি বহুল প্রচলিত সেটি হলো গ্যামা সহগ।

গ্যামা সহগ (Gamma Coefficient): সামাজিক বিজ্ঞানের গবেষণায় আমরা প্রায়শঃ এমন অনেক চলক নিয়ে কাজ করি, যেগুলোকে শুধুমাত্র ক্রমমূলক মাত্রায় বিন্যস্ত করা যায়। যেমন, মনোভাব, পদমর্যাদা, শ্রেণী অবস্থান, ইত্যাদি। ক্রমমূলক মাত্রায় পরিমাপকৃত দু'টি চলকের মধ্যে সম্পর্ক পরিমাপের জন্য যে পরিমাপটি বহুল প্রচলিত, সেটি হলো, গ্যামা সহগ (Gamma coefficient), যা গ্রীক বর্ণমালার 'γ' অক্ষরটি দিয়ে চিহ্নিত করা হয়। এটি একটি প্রতিসাম্য (symmetrical) পরিমাপ। যেহেতু গবেষক একটি চলকের পর্যবেক্ষণের জোড়ার অবস্থান, অন্য চলকের পর্যবেক্ষণের জোড়ার অবস্থানকে ভবিষ্যদ্বাণী করার ক্ষেত্রে উপযোগী কি না, তা জানতে আগ্রহী হন, সেহেতু গ্যামা সহগ নির্ণয়ের ক্ষেত্রে, দু'টি ক্রমমূলক চলকের পর্যবেক্ষণগুলোর জোড়ার ক্রম অবস্থানের প্রতি গুরুত্ব দেওয়া হয়। গ্যামা সহগের মান '০' থেকে যে ± ১ পর্যন্ত বিস্তৃত। একটি চলকের পর্যবেক্ষণের জোড়ার ক্রম অবস্থান, অন্য চলকের পর্যবেক্ষণের জোড়ার ক্রম অবস্থান নির্ণয়ে যদি কোন ভূমিকা না রাখে, তখন গ্যামা সহগের মান '০' হয়। যখন সমমাত্রার জোড়াগুলো আধিপত্য বিস্তার করে, তখন গ্যামা সহগের মানটি হয় ধনাত্মক এবং যখন অসমমাত্রার জোড়াগুলোর পরিমাণ বেশী হয়, তখন সহগটি ঋণাত্মক মানের হয়। গ্যামা সহগ নির্ণয়ের সূত্রটি হলো,

$$\gamma = \frac{N_c - N_d}{N_c + N_d}$$

যেখানে, γ = গ্যামা সহগ
N_c = সমমাত্রার জোড়া
N_d = অসমমাত্রার জোড়া।

যদি $N_c = ০$, তবে $\gamma = -১$, যদি $N_d = ০$, তবে $\gamma = ১$ এবং যদি $N_c = N_d$ হয়, তবে $\gamma = ০$ হয়।

কেন্ডালের টাউ-বি সহগ (Kendall's τ_b Coefficient): গ্যামা সহগের প্রধান দূর্বলতাটি হলো যে, এটি নির্ণয় করতে গিয়ে পর্যবেক্ষণের যে জোড়াগুলো একই রকম (tied pairs) হয়, সেগুলোকে বাদ দেয়া হয়। বলে দু'টি চলকের মধ্যে সম্পর্ক না থাকা সত্ত্বেও গ্যামা সহগের মান +১ থেকে − ১ হয়ে থাকে। এই সমস্যাটি দূর করার জন্য একটি সংশোধনী যুক্ত করে কেন্ডালের টাউ-বি (Kendall's τ_b) সহগ নির্ণয় করা যায়, যার বীজগাণিতিক অভিব্যক্তিটি নিম্নরূপ:

$$\tau_b = \frac{N_c - N_d}{\sqrt{(N_a + N_d + N_x)(N_c + N_d + N_y)}}$$

যেখানে, τ_b = কেন্ডালের টাউ-বি

N_c = সমমাত্রার জোড়া

N_d = অসমমাত্রার জোড়া

N_x = স্বাধীন চলক, x -এর উপর একই রকম জোড়ার (tied pairs) সংখ্যা

N_y = নির্ভরশীল চলক, y -এর উপর একই রকম জোড়ার সংখ্যা।

x ও y চলক দু'টির মানসমূহের মধ্যে সহ-সম্পর্ক স্থাপন না করে, স্পিয়ারম্যান rho বরং তাদের ক্রমবিন্যস্ত অবস্থানসমূহের মধ্যে সহ-সম্পর্ক প্রদর্শন করে।

স্পিয়ারম্যান rho (Spearman rho): ক্রমসূচক মাত্রায় পরিমাপকৃত চলকের ক্ষেত্রে আরেকটি পরিসংখ্যানিক পদ্ধতির মাধ্যমে সহ-সম্পর্কের সহগ নির্ণয় করা হয়, তা হলো Spearman Rank-Difference Correlation Coefficient, সংক্ষিপ্তভাবে যা স্পিয়ারম্যান rho, বা ρ নামে পরিচিত। এটিকে জনপ্রিয়ভাবে Rank-Correlationও বলা হয়ে থাকে। স্পিয়ারম্যান rho-এর নির্ণয় পদ্ধতিতে, x ও y -চলকের মানকে ব্যবহার না করে, চলক দু'টির ক্রমমানমূলক পদসোপানিক অবস্থানের (hierarchical rank-order) পার্থক্যকে ব্যবহার করে। x ও y চলক দু'টির মানসমূহের মধ্যে সহ-সম্পর্ক স্থাপন না করে, স্পিয়ারম্যান rho বরং তাদের ক্রমবিন্যস্ত অবস্থানসমূহের (ranks) মধ্যে সহ-সম্পর্ক প্রদর্শন করে। প্রথমে মানের ক্রমমানমূলক অবস্থানের পার্থক্য গ্রহণ করি, তারপর, সেই পার্থক্যগুলোর বর্গের যোগফলকে একটি সূত্রে প্রয়োগ করে স্পিয়ারম্যান rho-এর মানটি নির্ণয় করি। স্পিয়ারম্যান rho-এর ব্যাখ্যা সহ-সম্পর্কের অন্যান্য সহগের সাধারণ ব্যাখ্যারই অনুরূপ। অর্থাৎ, দু'টি চলকের সহ-সম্পর্কের নির্ণীত মান (+) কিংবা (–) হলে তা যথাক্রমে চলকদ্বয়ের মধ্যে ধনাত্মক বা ঋণাত্মক সম্পর্ককে নির্দেশ করে। যদি দু'টি চলকের প্রতিটির ক্রমমানমূলক অবস্থান প্রতিটির সাথে মিলে যায়, তবে স্পিয়ারম্যান rho-এর মান হবে +১, যদি প্রতিটির সাথে প্রতিটির অমিল ঘটে, তবে মানটি হবে –১ এবং যদি কোন সম্পর্কই না থাকে, তবে মানটি হবে '০'। স্পিয়ারম্যান rho নির্ণয়ের বীজগাণিতিক সূত্রটি হলো,

$$\rho = ১ - \frac{৬\sum D^২}{N(N^২ - ১)}$$

যেখানে, ρ = স্পিয়ারম্যান rho-এর সহগ

$D^২$ = দু'টি চলকের মানসমূহের ক্রমবিন্যস্ত অবস্থানের পার্থক্যের বর্গ

৬ = একটি ধ্রুবক

N = জোড় বাঁধা পর্যবেক্ষণের মোট সংখ্যা।

আমরা একটি অবিচ্ছিন্ন চলক এবং একটি দ্বি-বিভাজিত নামসূচক বা ক্রমসূচক চলকের মধ্যে সহ-সম্পর্ক নির্ণয়ের জন্য পয়েন্ট বাইসেরিয়াল সহ-সম্পর্কের সহগ ব্যবহার করতে পারি।

পয়েন্ট বাইসেরিয়াল সহ-সম্পর্কের সহগ (Point Biserial Correlation Coefficient): সামাজিক গবেষণায় আমরা যে সকল চলক নিয়ে কাজ করি, তার অধিকাংশই নামসূচক বা ক্রমসূচক মাত্রায় পর্যবেক্ষণ করা হয়। ক্রমসূচক মাত্রায় পরিমাপকৃত চলকসমূহের ক্ষেত্রে, Spearman rank-difference correlation coefficient, বা স্পিয়ারম্যান rho নির্ণয়ের কৌশলটি একটি যৌক্তিক নির্বাচন হিসাবে বিবেচিত হয়েছে। কিন্তু সহ-সম্পর্ক নির্ণয়ের ক্ষেত্রে, যদি দু'টি চলকের মধ্যে একটি নামসূচক, বা ক্রমসূচক মাত্রায় পরিমাপকৃত হয় এবং সেই চলকটি যদি দ্বি-বিভাজিত (dichotomous) হয়, সে ক্ষেত্রে স্পিয়ারম্যান rho নির্ণয় পদ্ধতিটি প্রয়োগ করা যায় না। পয়েন্ট বাইসেরিয়াল সহ-সম্পর্কের সহগ নির্ণয়ের মাধ্যমে এ ধরণের সঙ্কট থেকে উত্তোরিত হওয়া যায় এবং এর মাধ্যমে সামাজিক সম্পর্কের বহুবিধ রূপ বর্ণনা করা যায়। অর্থাৎ, আমরা একটি অবিচ্ছিন্ন চলক এবং একটি দ্বি-বিভাজিত নামসূচক বা ক্রমসূচক চলকের মধ্যে সহ-সম্পর্ক নির্ণয়ের জন্য পয়েন্ট বাইসেরিয়াল সহ-সম্পর্কের সহগ (point biserial correlation coefficient) ব্যবহার করতে পারি, যাকে সংক্ষিপ্তভাবে r_{pb} প্রতীকের মাধ্যমে চিহ্নিত করা হয়।

ধরা যাক, একজন গবেষক জন্মহারের সাথে আধুনিক ভোগ্যপণ্য (যেমন, টেলিভিশন, ভিসিআর, রেডিও, রেফ্রিজারেটর, ইত্যাদি) ব্যবহারের সম্পর্ক পরিমাপ করে দেখতে চান। সে ক্ষেত্রে, জন্মহার একটি ব্যাপ্তিমূলক বা অনুপাতমূলক চলকের রূপ নেবে এবং আধুনিক ভোগ্যপণ্যের ব্যবহার দ্বি-বিভাজিত (হ্যাঁ/না) নামসূচক চলকের রূপ নেবে। যদি দেখা যায় যে, উচ্চ হারে আধুনিক ভোগ্যপণ্য ব্যবহারকারীদের অধিকাংশেরই জন্মহার নিম্ন, তবে বুঝতে হবে যে, আধুনিক ভোগ্যপণ্য ব্যবহারের সাথে জন্মহারের একটি উচ্চ ধনাত্মক সম্পর্ক রয়েছে। আর যদি দেখা যায় যে, আধুনিক ভোগ্যপণ্য ব্যবহারকারী ব্যক্তিদের অনেকের জন্মহার উচ্চ, তবে বোঝা যাবে যে, সম্পর্কটি দুর্বল এমন কি,

ঋণাত্মকও হতে পারে। কাজেই, r_{pb} নির্ণয় করে বিষয়টি পরীক্ষা করে দেখা যেতে পারে। পয়েন্ট বাইসেরিয়াল সহ-সম্পর্কের সহগ নির্ণয়ের বীজগাণিতিক সূত্রটি হলো,

$$r_{pb} = \frac{\bar{x}_p - \bar{x}_q}{s_x}\sqrt{pq}$$

যেখানে, $\bar{x}_p$ = যারা আধুনিক ভোগ্যপণ্য ব্যবহারের প্রশ্নে 'হ্যাঁ' উত্তর দিয়েছেন, তাদের জন্মহারের গড়

$\bar{x}_q$ = যারা আধুনিক ভোগ্যপণ্য ব্যবহারের প্রশ্নে 'না' উত্তর দিয়েছেন, তাদের জন্মহারের গড়

s_x = জন্মহার চলকের পরিমিত ব্যবধান

p = যারা আধুনিক ভোগ্যপণ্য ব্যবহারের প্রশ্নের উত্তরে 'হ্যাঁ' বলেছেন, তাদের সমানুপাত

q = যারা আধুনিক ভোগ্যপণ্য ব্যবহারের প্রশ্নের উত্তরে 'না' বলেছেন, তাদের সমানুপাত।

পিয়ারসন r (Pearson r): ব্যাপ্তিমূলক বা অনুপাতমূলক চলকের মধ্যে সহ-সম্পর্কের সহগ নির্ণয়ের জন্য যে পরিসংখ্যানগত কৌশলটি ব্যবহৃত হয়, সেটি হলো Karl Pearson কর্তৃক প্রদত্ত product-moment correlation coefficient। এটিকে সংক্ষেপে পিয়ারসন r, বা শুধু r বলা হয়। পিয়ারসন r নির্ণয়ের জন্য আমাদেরকে কতগুলো পূর্ব-শর্ত অবশ্যই পূরণ করতে হবে। সেগুলো হলো, গবেষণাধীন চলক দু'টিকে অবশ্যই ব্যাপ্তিমূলক, বা অনুপাতমূলক মাত্রায় পরিমাপকৃত হতে হবে, চলক দু'টির মধ্যে সরলরৈখিক সম্পর্ক হতে হবে এবং পিয়ারসন r নির্ণয়ের ক্ষেত্রে চলক দু'টির ভেদাঙ্কের সমবিস্তৃতি (homoscedasticity of variances) থাকতে হবে। দু'টি চলক যদি কমপক্ষে ব্যাপ্তিমূলক মাত্রায় পরিমাপকৃত না হয়, বা তাদের মধ্যে সরলরৈখিক সম্পর্ক না থাকে, বা চলক দু'টির সম্পর্কের মধ্যে ভেদাঙ্কের সমবিস্তৃতি না থাকে, তবে সে ক্ষেত্রে নির্ণীত পিয়ারসন r, দু'টি চলকের মধ্যে সম্পর্কের সঠিক চিত্রটি প্রদান করবে না।

Pearson r এর মান '০' থেকে ±১ পর্যন্ত বিস্তৃত। Pearson r এর মান '০' থেকে যতোই ক্রমশঃ –১, বা +১-এর দিকে যেতে থাকবে, ততই দু'টি চলকের সম্পর্কের মাত্রা শূন্য থেকে হ্রাস পেতে পেতে ঋণাত্মক, বা বৃদ্ধি পেতে পেতে ধনাত্মক পূর্ণ সহ-সম্পর্কের দিকে যাবে। ধরা যাক, একটি গবেষণায় নারীর শিক্ষাগত যোগ্যতা ও বিয়ের বয়সের মধ্যে সহ-সম্পর্কের সহগের মান পাওয়া গিয়েছে +০.৮৭। সুতরাং, আমরা বলতে পারি যে, এ দু'টি চলকের মধ্যে উচ্চমাত্রায় ধনাত্মক সম্পর্ক বিদ্যমান। অর্থাৎ, নারী শিক্ষা বৃদ্ধির সাথে সাথে, তাদের বিয়ের বয়স বৃদ্ধি পায়। পিয়ারসন r-কে নিম্নলিখিত সূত্রের মাধ্যমে নির্ণয় করা যায়।

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}}$$

যেখানে, x_i = x চলকের প্রতিটি মান

y_i = y চলকের প্রতিটি মান

$\bar{x}$ = x চলকের মানের গড়

$\bar{y}$ = y চলকের মানের গড়

r = পিয়ারসন r সহগ।

নির্ভরণ বিশ্লেষণ (Regression Analysis)

একটি চলকের জানা মান থেকে অন্য একটি চলকের অজানা মানকে প্রাক্কলনের মাধ্যমে ভবিষ্যদ্বাণী করতে পারি, সেই কৌশলটিকে নির্ভরণ বিশ্লেষণ বলে।

সহ-সম্পর্ক আমাদের কেবলমাত্র দু'টি চলকের মধ্যে সম্পর্ক, বা সহগামীতার প্রবণতাকে বর্ণনা করে, তাদের মধ্যে কার্য-কারণ সম্পর্ককে ব্যাখ্যা করে না। অর্থাৎ, সহ-সর্ম্পকের মাধ্যমে আমরা দু'টি চলকের সম্পর্কের দিক ও মাত্রা সম্পর্কে জানতে পারি, কিন্তু কোন চলকটি কোন চলকের পরিবর্তনের কারণ, তা জানতে পারি না। যদিও সহ-সম্পর্ক কার্য-কারণ সম্পর্ক নয়, তবুও সহ-সম্পর্ক প্রত্যয়ের মধ্যে ভবিষ্যদ্বাণীকরণের ধারণাটি অন্তর্নিহিত রয়েছে। দু'টি চলক যখন সহগামী হয়, তখন একটি চলক দিয়ে

অন্যটিকে ভবিষ্যদ্বাণী করা যায়। যদি চলক দু'টি সহগামী না হয়, তবে একটির উপরে প্রাপ্ত তথ্য অন্যটির ভবিষ্যদ্বাণীকরণে সাহায্য করে না। যে পরিসংখ্যানগত কৌশলের সাহায্যে আমরা একটি চলকের জানা মান থেকে অন্য একটি চলকের অজানা মানকে প্রাক্কলনের মাধ্যমে ভবিষ্যদ্বাণী করতে পারি, সেই কৌশলটিকে নির্ভরণ বিশ্লেষণ বলে।

নির্ভরণ বিশ্লেষণ মূলতঃ পিয়ারসন r-এর ধারণার উপর ভিত্তি করে নির্মিত। পিয়ারসন r নির্ণয়ের ক্ষেত্রে, দু'টি চলককে যে সকল পূর্ব-শর্ত পূরণ করতে হয়, নির্ভরণ বিশ্লেষণের ক্ষেত্রেও সেই সকল শর্ত পূরণ করতে হয়। অর্থাৎ, দু'টি চলককেই কমপক্ষে ব্যাপ্তিমূলক মাত্রায় পরিমাপকৃত হতে হবে, দু'টি চলকেরই ভেদাঙ্কের সমবিস্তৃতি থাকতে হবে এবং দু'টি চলকের সম্পর্কটি সরলরৈখিক হতে হবে। নির্ভরণ বিশ্লেষণ শুধুমাত্র দু'টি চলকের মধ্যেই সীমাবদ্ধ নয়। একটি নির্ভরশীল চলককে ব্যাখ্যা করার জন্য বহু স্বাধীন চলক দায়ী থাকতে পারে। বিশ্লেষণের মধ্যে যখন দুই-এর অধিক চলক অন্তর্ভুক্ত হয়, তখন তাকে বহুধা নির্ভরণ বিশ্লেষণ (multiple regression analysis) বলে, যা পরবর্তী পাঠে আলোচনা করা হবে।

আমরা জানি যে, দু'টি চলকের মধ্যে সহ-সম্পর্ক থাকলে, তা চলক দু'টির মধ্যে সম্পর্কের একটি পথ নির্দেশ করে। এই পথকে বর্ণনা করার জন্য একটি রেখা ব্যবহার করা হয়, যে রেখাটিকে নির্ভরণ রেখা বলে। সম্পর্কের প্রকৃতি অনুসারে, এই রেখা সরল বা বক্র হতে পারে। নির্ভরণ রেখার বীজগাণিতিক অভিব্যক্তিই হলো, নির্ভরণ সমীকরণ। ভবিষ্যদ্বাণীর জন্য যে চলকটির অজানা মানকে প্রাক্কলন করা হয়, সে চলকটিকে নির্ভরশীল চলক (y) এবং যে চলকটির জানা মান ব্যবহার করে অজানা মানটিকে প্রাক্কলিত করা হয় সে চলকটিকে স্বাধীন চলক (x) বলা হয়। সরলরৈখিক নির্ভরণ রেখা নির্ণয়ের সাধারণ সূত্রটি হলো,

$$\hat{Y} = a + bX$$

যেখানে, $\hat{Y}$ = প্রাক্কলিত Y চলক
X = স্বাধীন চলক
a = ছেদক
b = নির্ভরাঙ্ক।

a ও b (ছেদক ও নির্ভরাঙ্ক) হলো দু'টি ধ্রুবক, যা নির্ভরণ রেখার অবস্থানকে নির্ধারণ করে এবং এই ধ্রুবক দু'টিকে নির্ভরণ রেখার চরিত্র-নির্ধারক বৈশিষ্ট্য বলে চিহ্নিত করা হয়। a এবং b ধ্রুবকের মান নির্ণয় করা গেলে নির্ভরণ রেখাটি সম্পূর্ণভাবে নির্ধারণ করা সম্ভব। ক্ষুদ্রতম বর্গের পদ্ধতির মাধ্যমে এই ধ্রুবক দু'টির মান নির্ণয় করা হয়। ধ্রুবক দু'টির, বা এর যে কোন একটির মান পরিবর্তিত হলে, নতুন একটি নির্ভরণ রেখার জন্ম নেবে। a ধ্রুবকটি নির্ভরণ রেখার মাত্রাকে (অর্থাৎ রেখার উৎপত্তিস্থলের উপরের বা নীচের দূরত্ব) এবং b ধ্রুবকটি নির্ভরণ রেখার তির্যকতাকে (অর্থাৎ x চলকে এক একক পরিবর্তনের ফলে y চলকে পরিবর্তনের পরিমাণ) নির্ধারণ করে। a ও b ধ্রুবক দু'টি নির্ণয়ের বীজগাণিতিক সূত্র দু'টি হলো,

$$a = \frac{\sum Y - b\sum X}{N}$$

$$b = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - (\sum X)^2}$$

যেখানে, a = ছেদক
b = নির্ভরাঙ্ক
X = স্বাধীন চলকের প্রতিটি মান
Y = নির্ভরশীল চলকের প্রতিটি মান
N = মোট ঘটনসংখ্যা।

সারাংশ

দু'টি চলকের সম্পর্কের ভিত্তিতে যে বিশ্লেষণ করা হয়, তাকে দ্বি-চলক বিশিষ্ট বিশ্লেষণ বলা হয়। দ্বি-চলক বিশিষ্ট বিশ্লেষণে বর্ণনার পাশাপাশি, ব্যাখ্যাও করা হয়ে থাকে। তবে মনে রাখতে হবে যে,

ব্যাখ্যার ক্ষেত্রেও দু'টি চলককেই ব্যবহার করা হয়। অর্থাৎ, একটি চলকের সাথে সাথে অন্য চলকের সম্পর্ক এবং একটি চলক দিয়ে অন্য চলকের মানের পরিবর্তনকে ব্যাখ্যা করা। এই সম্পর্কের মাত্রা নির্ধারণের জন্য পরিসংখ্যানের বিভিন্ন গুরুত্বপূর্ণ পদ্ধতি ব্যবহার করা হয়। দ্বি-চলক বিশিষ্ট বিশ্লেষণে যে সব পরিসংখ্যানগত পদ্ধতি ব্যবহৃত হয়ে থাকে, সেগুলো হলো, কাই-বর্গ পরীক্ষা, গ্যামা সহগ, পিয়ারসন r, এবং নির্ভরণ বিশ্লেষণ। বর্তমান পাঠে, দ্বি-চলক বিশিষ্ট বিশ্লেষণের সাথে সম্পর্কিত পরিসংখ্যান পদ্ধতিগুলো আলোচনা করা হয়েছে।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক ($\surd$) চিহ্ন দিন –

১। দু'টি চলকের মধ্যে সম্পর্ক রয়েছে কিনা তা পরীক্ষার জন্য প্রথম পদক্ষেপ হলো:
   ক. উপাত্তের গড় বের করা
   খ. দু'টি চলকের সম্পর্ক দেখিয়ে লেখচিত্র তৈরি করা
   গ. দ্বি-চলক বিশিষ্ট সারণি গঠন করা
   ঘ. উপরের সব।

২। বিক্ষেপ চিত্রের অসুবিধাটি হলো যে:
   ক. সহ-সম্পর্ক সম্পর্কে সাধারণ জ্ঞান অর্জন করা যায় না
   খ. সহ-সম্পর্ক সম্পর্কে দৃঢ়-ধারণা পাওয়া যায় না
   গ. সহ-সম্পর্কের মাত্রা সম্পর্কে গাণিতিক মান পাওয়া যায় না
   ঘ. উপরের সব।

৩। যদি দু'টি চলকের প্রতিটির ক্রমমানমূলক অবস্থান প্রতিটির সাথে মিলে যায়, তবে স্পিয়ারম্যান rho-এর মান হবে:
   ক. +১
   খ. –১
   গ. +০
   ঘ. –০।

সংক্ষিপ্ত প্রশ্ন

টীকা লিখুন:

   ক. বিক্ষেপ চিত্র।
   খ. পয়েন্ট বাইসেরিয়াল সহ-সম্পর্কের সহগ।

রচনামূলক প্রশ্ন

১। দ্বি-চলক বিশিষ্ট বিশ্লেষণ কী? দ্বি-চলক বিশিষ্ট সারণির ব্যবহার আলোচনা করুন।

২। দ্বি-চলক বিশিষ্ট সারণি ব্যাখ্যায় যে পরিসংখ্যানগত বিশ্লেষণ ব্যবহার করা হয়, সেগুলো আলোচনা করুন।

পাঠ - ৪

বহু-চলক বিশিষ্ট বিশ্লেষণ
Multivariate Analysis

এই পাঠ শেষে যা জানা যাবে —

- বহু-চলক বিশিষ্ট বিশ্লেষণ কী
- বহুধা নির্ভরণ বিশ্লেষণ
- পথ বিশ্লেষণ
- উপাদান বিশ্লেষণ
- ভেদাঙ্কের বিশ্লেষণ
- পরিসংখ্যানগত তাৎপর্যের পরীক্ষা

বহু-চলক বিশিষ্ট বিশ্লেষণ কী (What is Mulivariate Analysis)?

একটি নির্ভরশীল চলকের মধ্যে পরিবর্তনের জন্য দায়ী বহু চলকের যুগপৎ প্রভাবের পরীক্ষাই হলো বহু-চলক বিশিষ্ট বিশ্লেষণ।

পাঠ ২-এ আমরা গণসংখ্যা নিবেশন, কেন্দ্রীয় প্রবণতা ও বিস্তৃতির পরিমাপের মাধ্যমে একটি চলকের স্বরূপ ও অস্তিত্বের বর্ণনা করেছি এবং পাঠ ৩-এ দু'টি চলকের মধ্যে সম্পর্কের বর্ণনা ও ব্যাখ্যার বিভিন্ন পরিসংখ্যানগত কৌশল আলোচনা করেছি। যদি একজন গবেষকের মূল আগ্রহটি শুধুমাত্র দু'টি চলকের মধ্যে সম্পর্কের বর্ণনা ও ব্যাখ্যার মধ্যেই সীমাবদ্ধ থাকে, তবে দ্বি-চলক বিশিষ্ট বিশ্লেষণই যথেষ্ট। কিন্তু সামাজিক প্রপঞ্চসমূহের ব্যাখ্যায় শুধুমাত্র দু'টি চলকের মধ্যেই একজন গবেষক নিজেকে সীমাবদ্ধ রাখতে পারেন না। কারণ, দ্বি-চলক বিশিষ্ট বিশ্লেষণে শুধু দু'টি চলকের মধ্যে সম্পর্ক বর্ণনা করা হয় এবং সেই দু'টি চলকের একটিকে দিয়ে অন্যটিকে ব্যাখ্যা করা হয়। কিন্তু সামাজিক বাস্তবতায় দেখা যায় যে, বিভিন্ন পরিস্থিতিতে একটি নির্ভরশীল চলকের মধ্যে যে পরিবর্তন ঘটে, তা ব্যাখ্যার জন্য একটি স্বাধীন চলকই যথেষ্ট নয়। সে কারণে গবেষক তার বিশ্লেষণে আরো অনেক চলক অন্তর্ভুক্ত করেন। সামাজিক গবেষণায়, একটি নির্ভরশীল চলকের মধ্যে পরিবর্তনের জন্য দায়ী বহু চলকের যুগপৎ প্রভাবের পরীক্ষাই হলো, বহু-চলক বিশিষ্ট বিশ্লেষণ।

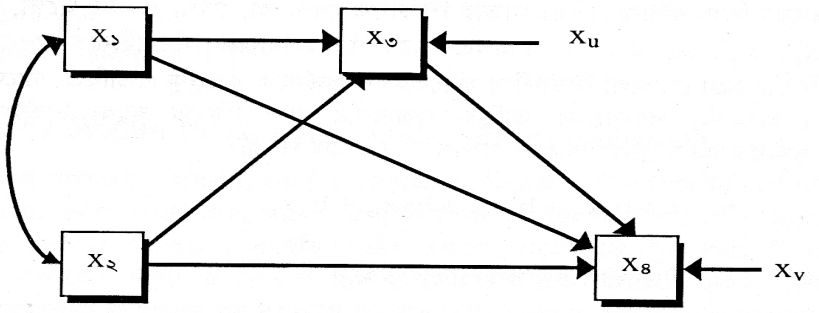
বহু চলক বিশিষ্ট বিশ্লেষণের বেশ কিছু পরিসংখ্যানগত কৌশল রয়েছে। এর মধ্যে যে কৌশলগুলো বেশী ব্যবহৃত হয়, সেগুলো হলো বহুধা নির্ভরণ বিশ্লেষণ, পথ বিশ্লেষণ, উপাদান বিশ্লেষণ, ভেদাঙ্কের বিশ্লেষণ, ইত্যাদি। এ সব কৌশল সম্পর্কে আলোচনার পূর্বে এটি উল্লেখ করা প্রয়োজন যে, বহু-চলক বিশিষ্ট বিশ্লেষণ উচ্চতর পরিসংখ্যানিক বিশ্লেষণে প্রয়োগ করা হয় এবং এগুলো কলাকৌশলগত দিক থেকে কিছুটা জটিল। অতএব, এ সব কৌশলের খুব বেশী গভীরে না গিয়ে, শুধুমাত্র একটি প্রাথমিক ধারণা দেবার জন্যই সেগুলোর উপস্থাপন করা হবে।

বহুধা নির্ভরণ বিশ্লেষণ (Multiple Regression Analysis)

পাঠ ৩ - এ আমরা জেনেছি যে, নির্ভরণ বিশ্লেষণের মাধ্যমে একটি চলকের উপর অর্জিত জ্ঞানের ভিত্তিতে অন্য একটি চলককে ব্যাখ্যা করা যায়। যে চলকটির জানা মান ব্যবহার করে অন্য একটি চলকের অজানা মানকে প্রাক্কলিত করা হয়, সে চলকটিকে স্বাধীন চলক এবং যে চলকটির অজানা মানকে প্রাক্কলন করা হয়, সে চলকটিকে নির্ভরশীল চলক বলা হয়। অন্য কথায়, যে চলকটি প্রভাবিত হয়, সেটিকে নির্ভরশীল চলক (y) এবং যে চলকটি প্রভাবিত করে, সেটিকে স্বাধীন চলক (x) বলে। এ ক্ষেত্রে লক্ষ্যণীয় বিষয়টি হলো যে, শুধুমাত্র একটি চলক দিয়ে আরেকটি চলকের মধ্যে পরিবর্তনকে ব্যাখ্যা করা হয়। কিন্তু আমরা এও জানি যে, বহু-চলক বিশিষ্ট বিশ্লেষণের অন্তর্নিহিত যুক্তি অনুযায়ী, একটি নির্ভরশীল চলককে ব্যাখ্যা করার জন্য বহু স্বাধীন চলক দায়ী থাকে। অতএব, ব্যাখ্যাকারী একাধিক স্বাধীন চলককে $x_১$, $x_২$, $x_৩$, ..., x_n প্রতীকের মাধ্যমে চিহ্নিত করা হয় এবং প্রতিটি স্বাধীন চলকের প্রভাবের মাত্রা নির্ধারণকারী নির্ভরাঙ্কগুলোকে যথাক্রমে $b_১$, $b_২$, $b_৩$, ..., b_n প্রতীকের মাধ্যমে চিহ্নিত করা হয়।

ধরা যাক, একজন গবেষক বিভিন্ন তত্ত্ব ও গবেষণা সাহিত্য পর্যালোচনা এবং নিজস্ব অভিজ্ঞতার আলোকে দেখলেন যে, সন্তানের শিক্ষাগত যোগ্যতার মাত্রাটি নির্ধারিত হয়, প্রধানতঃ চারটি নিয়ামক উপাদানের মাধ্যমে। সেই উপাদানগুলো হলো সন্তানের লিঙ্গ-পরিচয়, পিতা-মাতার শিক্ষাগত যোগ্যতা, পেশা ও আয়। অর্থাৎ, সন্তান পুরুষ হলে, পিতা-মাতা সেই সন্তানের লেখা পড়ার বিষয়ে বেশী যত্নবান হন, পিতা-মাতা যদি উচ্চ শিক্ষিত হন, উচ্চতর পেশায় কর্মরত থাকেন এবং উচ্চ আয় সম্পন্ন হন, তবে স্বাভাবিকভাবেই সে সকল পিতা-মাতার সন্তান উচ্চ শিক্ষিত হবে। অতএব, গবেষকের তাত্ত্বিক

প্রস্তাবনা অনুযায়ী, এ ক্ষেত্রে সন্তানের শিক্ষাগত যোগ্যতা হলো, নির্ভরশীল চলক 'y' এবং সন্তানের লিঙ্গ-পরিচয়, পিতা-মাতার শিক্ষাগত যোগ্যতা, পেশা ও আয় হলো, স্বাধীন চলক, যা যথাক্রমে $x_১$, $x_২$, $x_৩$ এবং $x_৪$। এখানে মনে রাখা প্রয়োজন যে, স্বাধীন চলকগুলোর গুরুত্বের ক্রম অবস্থান নির্ভর করে সেগুলো নির্ভরশীল চলকের পরিবর্তনকে কতটুকু প্রভাবিত করতে পারে, তার মাত্রার উপর। অতএব, গবেষকের তাত্ত্বিক প্রস্তাবনার পরিপ্রেক্ষিতে বহুধা নির্ভরণ বিশ্লেষণের বীজগাণিতিক সমীকরণটি নিম্নলিখিতভাবে প্রকাশ করা যেতে পারে:



চিত্র ১০.৪.১: মোট জন্মহার ব্যাখ্যার একটি পথ নক্শা

যেখানে, $\hat{Y}$ = নির্ভরশীল চলকের প্রাক্কলিত মান

a = ছেদক

$b_১, b_২, b_৩, b_৪$ = স্বাধীন চলকগুলোর নির্ভরাঙ্ক

$x_১, x_২, x_৩$ ও $x_৪$ = স্বাধীন চলকসমূহ

e = ভ্রান্তিমূলক প্রভাব।

প্রতিটি স্বাধীন চলকের সাথে তার প্রভাবের মাত্রা নির্ধারণকারী সংশ্লিষ্ট নির্ভরাঙ্ককে ($b_১, b_২, b_৩, b_৪$) আংশিক নির্ভরাঙ্ক বলা হয়ে থাকে। যে স্বাধীন চলকটির প্রভাবের মাত্রা নির্ণীত হচ্ছে, সেটি ছাড়া অন্য স্বাধীন চলকগুলোকে নিয়ন্ত্রণ করে আংশিক নির্ভরাঙ্কগুলো নির্ণীত হয়। অর্থাৎ, আমাদের উদাহরণের ক্ষেত্রে, সন্তানের লিঙ্গের প্রভাবটি নির্ধারিত হয় পিতামাতার শিক্ষাগত যোগ্যতা, পেশা ও আয়ের প্রভাবকে নিয়ন্ত্রিত করে। এই আংশিক নির্ভরাঙ্কগুলোকে সংশ্লিষ্ট স্বাধীন চলকের নীট প্রভাবও বলা হয়ে থাকে। প্রতিটি আংশিক নির্ভরাঙ্ক বা নীট প্রভাবকে যোগ করে পূর্ণ প্রভাবের মাত্রাটি পাওয়া যায়। অতএব, বহুধা নির্ভরণ সমীকরণটি হলো যোজনমূলক (additive)।

কিন্তু যেহেতু সামাজিক প্রপঞ্চের ব্যাখ্যায় একটি প্রপঞ্চের পরিবর্তনে যে সকল উপাদান প্রভাবিত করে তার সবগুলোকেই নির্ভরণ সমীকরণের অন্তর্ভুক্ত করা যায় না, সেহেতু নিভরশীল চলকের মধ্যে পরিবর্তনের কিছু মাত্রা অব্যাখ্যাত থেকে যায়। এই অব্যাখ্যাগত পরিবর্তনের পরিমাণটিকেই আমরা ভ্রান্তিমূলক প্রভাব (e) বলে চিহ্নিত করি। এই ব্যাখ্যা করতে না পারার বিষয়টি নির্ভর করে উপাত্তের অপ্রাপ্যতা ও নির্ভরশীল চলকের ব্যাখ্যায় অসম্পূর্ণ জ্ঞানের কারণে। তবে যেহেতু পরিসংখ্যানগত সিদ্ধান্তের জন্য নির্ভরণ বিশ্লেষণের ব্যবহার পিয়ারসন r-এর অন্তর্নিহিত ধারণার উপর প্রতিষ্ঠিত, সেহেতু সামাজিক গবেষণায় পিয়ারসন r-এর অন্তর্নিহিত ধারণার সকল শর্তগুলো পূরণ করা সবসময় সম্ভব হয় না বলে নির্ভরণ বিশ্লেষণের ফলাফল মূল্যায়নে সতর্কতা অবলম্বন করতে হয়।

পথ বিশ্লেষণ (Path Analysis)

পথ বিশ্লেষণ হলো, বিভিন্ন চলকের মধ্যে সম্পর্ককে বোঝার জন্য একটি কার্য-কারণ সম্পর্কিত মডেল।

বহু-চলক বিশিষ্ট বিশ্লেষণের কৌশল হিসাবে সাম্প্রতিককালে সমাজবিজ্ঞানে, পথ বিশ্লেষণ বেশ জনপ্রিয়তা পেয়েছে। পথ বিশ্লেষণ হলো, বিভিন্ন চলকের মধ্যে সম্পর্ককে বোঝার জন্য একটি কার্য-কারণ সম্পর্কিত মডেল। একটি নির্ভরশীল চলকের উপর স্বাধীন চলকগুলোর প্রত্যক্ষ, পরোক্ষ ও মেকি (spurious) প্রভাব নির্ণয়ের জন্য পথ বিশ্লেষণ খুবই উপযোগী। অন্য যে কোন উপায়ের তুলনায় পথ বিশ্লেষণের মাধ্যমে বিভিন্ন চলকের সম্পর্কের গাঁথুনির একটি সুস্পষ্ট চিত্র পাওয়া যায়। একটি উদাহরণের মাধ্যমে বিষয়টি বোঝা যাক। ধরা যাক, একজন গবেষক মোট জন্মহারের মাত্রাকে ধার্মিকতা,

নগরায়ণের হার ও নারীর বিয়ের বয়স, এ তিনটি চলক দিয়ে ব্যাখ্যা করতে চান। সে ক্ষেত্রে, এই চলকগুলোকে তিনি একটি পথ নক্শার (path diagramme) মধ্যে উপস্থাপন করবেন, যা নিম্নরূপ:

যেখানে, $x_১$ = ধার্মিকতা
$x_২$ = নগরায়ণের হার
$x_৩$ = নারীর বিয়ের বয়স
$x_৪$ = মোট জন্মহার
x_u ও x_v = ভ্রান্তিমূলক প্রভাব।

এই পথ নক্শার মাধ্যমে, তীর চিহ্নগুলো দিয়ে চলকগুলোর মধ্যেকার কার্য-কারণ সম্পর্কের জালটি সুস্পষ্টভাবে প্রর্দশন করা যায়। শুধু তাই নয়, এই পথ নক্শায় মোট জন্মহারের উপর কোন চলকগুলোর প্রত্যক্ষ, ও পরোক্ষ প্রভাব রয়েছে, তা দৃশ্যমান হয়ে উঠে। কার্য-কারণ সম্পর্কের মাত্রা, বা শক্তিটি নির্ণীত হয় একটি চলকের উপর অন্য চলকের প্রভাবের মাত্রা নির্দেশক নির্ভরাঙ্কের মাধ্যমে, যা 'পথ সহগ' (path coefficient) বলে পরিচিত। বহুধা নির্ভরণ বিশ্লেষণের উপর ভিত্তি করে পথ সহগগুলো নির্ণয় করা হয়ে থাকে। সাধারণভাবে, পথ বিশ্লেষণ নির্ভরশীল চলকের উপর প্রতিটি স্বাধীন চলকের প্রভাব সম্বলিত কতগুলো নির্ভরণ সমীকরণের সমন্বয়। যেহেতু পথ বিশ্লেষণের বিস্তারিত আলোচনা এই গ্রন্থের পরিধির বাইরে, সেহেতু প্রভাবের পথ নির্ধারণের জন্য বিভিন্ন সমীকরণ আহরণের ধাপগুলো এখানে আলোচনা করা হবে না।

উপাদান বিশ্লেষণ (Factor Analysis)

উপাদান বিশ্লেষণ হলো, বহু-চলক বিশিষ্ট বিশ্লেষণের আর একটি ভিন্ন কৌশল। এ পর্যন্ত আমরা যে সকল কৌশল আলোচনা করেছি, সে সব কৌশল থেকে উপাদান বিশ্লেষণের পরিসংখ্যানগত ভিত্তি বেশ ভিন্ন এবং জটিল। উপাদান বিশ্লেষণ বিভিন্ন চলকের মানের ভিন্নতার মধ্যে নির্দিষ্ট রূপ আবিষ্কারের জন্য ব্যবহার করা হয়ে থাকে। কিছু কৃত্রিম মাত্রা (উপাদান) উদ্ভাবনের মাধ্যমে উপাদান বিশ্লেষণ করা হয়, যেগুলো প্রকৃত চলকের সাথে উচ্চ মাত্রায় সহ-সম্পর্কিত হয় এবং যেগুলো একটি থেকে আরেকটি স্বাধীন। ধরা যাক, একটি উপাত্তরাশির মধ্যে ব্যক্তির পক্ষপাতমূলক, বা সংস্কারাচ্ছন্ন আচরণের বেশ কিছু সূচক রয়েছে। এ সব সূচকের প্রতিটি দফা সংস্কারাচ্ছন্ন আচরণের কিছু ইঙ্গিত দেয়, কিন্তু কোনটিই নিখুঁত নির্দেশনা দেয় না এবং দফা-গুলো পারস্পরিকভাবে উচ্চ মাত্রায় সম্পর্কযুক্ত। এ ক্ষেত্রে, গবেষক উপাদান বিশ্লেষণের মাধ্যমে একটি কৃত্রিম মাত্রা (উপাদান) তৈরি করবেন, যা সংস্কারাচ্ছন্ন আচরণ পরিমাপকারী প্রতিটি দফার সাথে উচ্চ মাত্রায় সম্পর্কিত থাকবে। চলকগুলোর মধ্যে পর্যবেক্ষণকৃত সম্পর্ক এবং উপাদানগুলোর সাথে চলকগুলোর সহ-সম্পর্ক থেকে প্রাপ্ত এই কৃত্রিম মাত্রাগুলোর মানকে factor loadings বলে। প্রাপ্ত factor loading গুলো থেকে একজন গবেষক সহজেই বুঝতে পারেন যে, সেগুলো কোন উপাদানকে প্রতিনিধিত্ব করে।

উপাদান বিশ্লেষণের মাধ্যমে যে উপাদানগুলোকে (কৃত্রিম মাত্রা) উদ্ভাবন করা হয়, সেগুলোর কোন প্রকৃত অর্থ থাকে না। উপাদানগুলোকে ব্যাখ্যার মাধ্যমে অর্থ প্রদান করতে হয়। সে জন্য এই উপাদানগুলোর সূচক নির্বাচন খুবই গুরুত্বপূর্ণ। কারণ, সঠিক সূচক নির্বাচন না করলে, উদ্ভাবিত কৃত্রিম মাত্রাগুলো বাস্তব অবস্থার সাথে সঙ্গতিপূর্ণ হয় না। তবে, এটি মনে রাখতে হবে যে, যেহেতু উপাদানগুলো প্রকৃত অর্থ ছাড়াই তৈরি হয়, সেহেতু গবেষক তুলনা করা যায় না, এমন কিছু চলকের জন্য উচ্চ মাত্রার factor loading সম্পন্ন উপাদান পেতে পারেন, যা উপাদান বিশ্লেষণের মূল উপযোগীতাকে নষ্ট করে ফেলে।

ভেদাঙ্কের বিশ্লেষণ (Analysis of Variance)

যখন নামসূচক চলক এবং ব্যাপ্তিমূলক চলকের মধ্যে কার্য-কারণ সম্পর্ক ব্যাখ্যার জন্য একটি পরিসংখ্যানগত পরীক্ষা করা হয়, তখন ভেদাঙ্কের বিশ্লেষণ একটি গুরুত্বপূর্ণ কৌশল হিসাবে কাজ করে।

যখন নামসূচক চলক এবং ব্যাপ্তিমূলক চলকের মধ্যে কার্য-কারণ সম্পর্ক ব্যাখ্যার জন্য একটি পরিসংখ্যানগত পরীক্ষা করা হয়, তখন ভেদাঙ্কের বিশ্লেষণ একটি গুরুত্বপূর্ণ কৌশল হিসাবে কাজ করে। ভেদাঙ্কের বিশ্লেষণকে সংক্ষেপে ANOVA বলা হয়। ভেদাঙ্কের বিশ্লেষণের মাধ্যমে পরীক্ষণ পরিচালনার সুবিধাটি হলো যে, এটি দু'টির বেশী দলের ক্ষেত্রে প্রয়োগ করা যায়। অর্থাৎ, দু'টির বেশী দলের মধ্যে পরীক্ষণ পরিচালনার মাধ্যমে নির্ভরশীল ও স্বাধীন চলকের মধ্যে সম্পর্কের এমন কিছু বিষয় উন্মোচিত হয়, যা দু'টি দল বিশিষ্ট পরীক্ষণের মাধ্যমে পাওয়া যায় না। পরীক্ষণের পূর্বে এবং পরে দল দু'টি নামসূচক চলকের দু'টি শ্রেণী-দফাকে প্রতিনিধিত্ব করে এবং ব্যাপ্তিমূলক চলকটি প্রতিটি শ্রেণী-দফায় একটি গড় মানকে প্রতিনিধিত্ব করে। শ্রেণী-দফা দু'টির প্রতিটি ক্ষেত্রে গড় মান, প্রতিটি দলের জন্য গড় প্রবণতাকে প্রকাশ করে।

একটি উদাহরণ দেয়া যাক। ধরা যাক, একজন গবেষক মায়ের বুকের দুধ খাওয়ানোর সাথে শিশুর উচ্চতা বৃদ্ধির সম্পর্ক পরীক্ষা করে দেখতে চান। সে ক্ষেত্রে, পরীক্ষণের সাধারণ পদ্ধতি অনুযায়ী দু'টি দল তৈরি করা হয়। একটি পরীক্ষণ দল, যেখানে জন্মের পর থেকে শিশুকে নিয়মিত মায়ের বুকের দুধ খাওয়ানো হয় এবং অন্যটি নিয়ন্ত্রণ দল, যেখানে শিশুকে মায়ের বুকের দুধ খাওয়ানো হয় না। একটি নির্দিষ্ট সময়ের পরে, দু'টি দলের শিশুদের ওজন নেয়া হয়। যদি দেখা যায় যে, দু'টি দলের মধ্যে পরীক্ষণের পূর্বের তুলনায় পরের ওজনের মধ্যে পার্থক্য তৈরি হয়েছে, সে ক্ষেত্রে গবেষক বলতে পারেন যে, মায়ের বুকের দুধ খাওয়ানোর একটি ধনাত্মক প্রভাব শিশুর ওজন বৃদ্ধির উপর পড়েছে। এখন প্রশ্ন হলো, পর্যবেক্ষণকৃত এই তফাৎটি কি মায়ের বুকের দুধ খাওয়ানোর ফলে হয়েছে, না কি শুধুমাত্র নমুনাগত ভিন্নতার কারণে হয়েছে? আরো অতিরিক্ত দলের উপর পরীক্ষণ চালানোর মাধ্যমে ভেদাঙ্কের বিশ্লেষণ আমাদের এ প্রশ্নের জবাব পেতে সাহায্য করে।

ভেদাঙ্কের বিশ্লেষণ পরিমাপের মাত্রাটি যে বিন্যাসের মাধ্যমে মূল্যায়িত হয়, সেই বিন্যাসটিকে F-বিন্যাস বলে। ভেদাঙ্কের বিশ্লেষণ দলগুলোর মধ্যে বিদ্যমান ভিন্নতাকে প্রতিটি দলের মধ্যেকার ভিন্নতা দিয়ে ভাগ করে একটি অনুপাত নির্ণয় করে, যা F-অনুপাত নামে পরিচিত। t-বিন্যাস ও কাই-বর্গ বিন্যাসের মত, F-বিন্যাসও কতগুলো নমুনা বিন্যাসের পরিবার, একটি একক বিন্যাস নয়। সে কারণে স্বাধীনতার মাত্রার উপর ভিত্তি করে কাই-বর্গ বিন্যাস ও t-বিন্যাসের মত, F-বিন্যাসেরও এক একটি বিন্যাস তৈরি হয়। F-অনুপাত নির্ণয়ের সূত্রটি হলো:

$$F = \frac{s_b^{২}}{s_w^{২}}$$

যেখানে, F = F-অনুপাতের মান

$s_b^{২}$ = দলগুলোর মধ্যে ভেদাঙ্ক বা ভিন্নতার পরিমাণ

$s_w^{২}$ = দলগুলোর অভ্যন্তরের ভেদাঙ্ক বা ভিন্নতার পরিমাণ।

এ ক্ষেত্রে, F-অনুপাতের মানটিকে স্বাধীনতার মাত্রার ভিত্তিতে F-বিন্যাসের তাত্ত্বিক মানের সাথে তুলনা করা হয়। দলগুলোর মধ্যেকার ভিন্নতার জন্য স্বাধীনতার মাত্রা হলো, $k-১$ এবং দলগুলোর অভ্যন্তরের ভিন্নতার জন্য স্বাধীনতার মাত্রা হলো $N-k$। যদি দলগুলোর মধ্যে কোন ভিন্নতা দেখা না যায়, তাহলে F-অনুপাত শূন্য হবে। অনুপাতটি যত বড় হবে, দলগুলো হবে তত বেশী ভিন্নরূপ। যখন একটি স্বাধীন চলক নিয়ে পরীক্ষণটি পরিচালনা করা হয়, তখন সেই ভেদাঙ্কের বিশ্লেষণকে এক-মুখী ভেদাঙ্কের বিশ্লেষণ বলে। যখন দু'টি স্বাধীন চলক নিয়ে বিশ্লেষণ করে, তখন সেটিকে দ্বি-মুখী ভেদাঙ্কের বিশ্লেষণ বলে। এই মূলনীতিকে অনুসরণ করে বহু-চলক বিশিষ্ট ভেদাঙ্কের বিশ্লেষণ করা যায়, যা সংক্ষিপ্তভাবে MANOVA (multiple analysis of variance) বলে পরিচিত।

পরিসংখ্যানগত তাৎপর্যের পরীক্ষা (Statistical Tests of Significance)

বর্ণনামূলক পরিসংখ্যানগত পরিমাপগুলো থেকে উপসংহারমূলক সিদ্ধান্ত গ্রহণ করতে হলে, সেই পরিমাপগুলোর প্রাপ্ত মানগুলো পরিসংখ্যানগতভাবে তাৎপর্যপূর্ণ কি না, সেটি নির্ধারণ করা প্রয়োজন। কারণ, পরিসংখ্যানগুলো যদি একটি নির্দিষ্ট মাত্রায় তাৎপর্যপূর্ণ হয়, কেবলমাত্র তখনই সেগুলোর ভিত্তিতে সাধারণীকরণ করা সম্ভব। আমরা বিভিন্ন মাত্রায় পরিমাপকৃত দু'টি চলকের মধ্যে সহ-সম্পর্কের সহগ নির্ণয়ের মাধ্যমে সহ-সম্পর্কের দিক ও মাত্রা সম্পর্কে জেনেছি। কিন্তু যেহেতু সমগ্রক থেকে সংগৃহীত নমুনার মাধ্যমে সহ-সম্পর্কের সহগগুলোকে নির্ণয় করা হয়েছে, সেহেতু স্বাভাবিকভাবেই প্রশ্ন তোলা যেতে পারে যে, দু'টি চলকের মধ্যে নির্ণীত সহ-সম্পর্কের সহগটি কি পরিসংখ্যানগতভাবে তাৎপর্যপূর্ণ, যার ভিত্তিতে সাধারণীকরণ করা যায়? অথবা বহু-চলক বিশিষ্ট বিশ্লেষণের কার্য-কারণ সম্পর্ক নির্ণয়ের ভিত্তিতে যে সব প্রাক্কলন পাওয়া যায়, সেগুলো কি একটি নির্দিষ্ট মাত্রায় পরিসংখ্যানগতভাবে তাৎপর্যপূর্ণ? বিভিন্ন বৈশিষ্টের ভিত্তিতে বলা যায় কি যে, বিভিন্ন গোষ্ঠীর মধ্যে বিদ্যমান পার্থক্য কি সত্যিকার, না কি সেগুলো নমুনাগত পার্থক্যের কারণে ঘটেছে? পরিসংখ্যানগত অনুকল্প পরীক্ষার মাধ্যমে এ প্রশ্নগুলোর উত্তর খোঁজা যেতে পারে।

ভেদাঙ্কের বিশ্লেষণ পরিমাপের মাত্রাটি যে বিন্যাসের মাধ্যমে মূল্যায়িত হয়, সেই বিন্যাসটিকে F- বিন্যাস বলে।

অনুকল্প পরীক্ষার বিভিন্ন পদ্ধতি রয়েছে। সে পদ্ধতিগুলোকে প্রধানতঃ দু'টি বৃহৎ পরিবারে বিভক্ত করা যায় – পরামাত্রিক (parametric) ও অপরামাত্রিক (non-parametric)। পরামাত্রিক পরীক্ষাগুলোর মধ্যে বহুল ব্যবহৃত পদ্ধতিগুলো হলো z-পরীক্ষা, t-পরীক্ষা এবং F-পরীক্ষা। অপরামাত্রিক পরীক্ষাগুলোর মধ্যে সবচেয়ে বেশী ব্যবহৃত পদ্ধতিটি হলো, কাই-বর্গ ($\chi^{২}$) পরীক্ষা। এ সকল অনুকল্প পরীক্ষার অন্তর্নিহিত যৌক্তিকতা ও নির্ণয় পদ্ধতি "সামাজিক পরিসংখ্যান পরিচিতি" (সমাজতত্ত্ব-৪) গ্রন্থে বিস্তারিতভাবে আলোচনা করা হয়েছে বলে, এখানে কেবলমাত্র সেগুলোর উল্লেখ করা হলো। আমরা

যেহেতু পাঠ ৩-এ দ্বি-চলক বিশিষ্ট বিশ্লেষণে কাই-বর্গ পরীক্ষা এবং এই পাঠে ভেদাঙ্কের বিশ্লেষণে F-পরীক্ষার উল্লেখ করেছি, সেহেতু এই দু'টি পরিসংখ্যানগত পরীক্ষার পুনরাবৃত্তি আমরা করবো না।

অনুকল্প পরীক্ষা করতে গিয়ে, আমরা পরস্পর বিপরীতমুখী দু'টি অনুকল্প নির্মাণ করি এবং পরিসংখ্যানগত পরীক্ষার মাধ্যমে প্রাপ্ত ফলাফলের ভিত্তিতে কোন অনুকল্পটি সঠিক, তা নির্ধারণ করি। প্রথম ধরণের অনুকল্পটিকে নাস্তি অনুকল্প বলে, যেখানে অনুমান করা হয় যে, একটি সমগ্রকের দু'টি চলকের মধ্যে পার্থক্য নেই, বা পরামানের সাথে নমুনা মানের কোন পার্থক্য নেই। নাস্তি অনুকল্পের বিকল্প হিসাবে দ্বিতীয় অনুকল্পটি নির্মাণ করা হয়, যা গবেষণা অনুকল্প, বা বিকল্প অনুকল্প হিসাবে বিবেচিত হয়। এর সাথে, বিভিন্ন নমুনা বিন্যাস ব্যবহার করে বিভিন্ন তাৎপর্য মাত্রায় অনুকল্পগুলোর সঠিকতার সম্ভাবনার মাত্রা পরীক্ষা করা হয়। এই তাৎপর্যের মাত্রার ভিত্তিতে আমরা নাস্তি অনুকল্পকে হয় প্রত্যাখান করি, নতুবা গ্রহণ করি। তবে নাস্তি অনুকল্প প্রত্যাখ্যানের ক্ষেত্রে, সিদ্ধান্ত গ্রহণে ভুল হতে পারে। যদি একটি সত্য নাস্তি অনুকল্পকে প্রত্যাখান করা হয়, তবে সেই সিদ্ধান্তকে টাইপ-১ ভ্রান্তি, বা আলফা (α) ভ্রান্তি বলে। কিন্তু যদি একটি মিথ্যা নাস্তি অনুকল্পকে প্রত্যাখান করতে ব্যর্থ হই, তখন সেই সিদ্ধান্তকে টাইপ-২ ভ্রান্তি, বা বিটা (β) ভ্রান্তি বলে। এ দু'ধরণের ভ্রান্তির ঝুঁকি বিপরীতভাবে সম্পর্কযুক্ত। অর্থাৎ, টাইপ-১ ভ্রান্তির ঝুঁকি যত কম হবে, টাইপ-২ ভ্রান্তির সম্ভাবনা তত বেড়ে যাবে। কারণ, আমরা জানি না যে, নাস্তি অনুকল্পটি প্রকৃতপক্ষেই সত্য, বা মিথ্যা কি না।

z-পরীক্ষা (z-test): বৃহৎ নমুনা সম্বলিত অনুকল্প পরীক্ষার ক্ষেত্রে, z-পরীক্ষা পরিচালনা করা হয়। z-পরীক্ষা পরিচালনার জন্য অন্তর্নিহিত ধারণাগুলো হলো, একটি নমুনাকে দৈব-চয়িত ভাবে নির্ধারিত হতে হবে, নমুনার বিন্যাসটি স্বাভাবিক হতে হবে এবং চলকগুলোকে ব্যাপ্তিমূলক মাত্রায় পরিমাপকৃত হতে হবে। একটি নমুনার জন্য z-পরীক্ষার সূত্রটি হলো,

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

যেখানে z = স্বাভাবিক মান
$\bar{x}$ = নমুনা গড়
μ = সমগ্রক গড়
$\sigma_{\bar{x}}$ = গড়ের পরিমিত ভ্রান্তি।

দু'টি নমুনার গড়ের মধ্যে পার্থক্যের পরিসংখ্যানগত তাৎপর্য পরীক্ষার জন্য z-পরীক্ষার সূত্রটি হলো,

$$z = \frac{\bar{x}_১ - \bar{x}_২}{\hat{\sigma}_{\bar{x}_১ - \bar{x}_২}}$$

যেখানে, z = স্বাভাবিক মান
$\bar{x}_১$ = প্রথম নমুনার গড়
$\bar{x}_২$ = দ্বিতীয় নমুনার গড়
$\hat{\sigma}_{\bar{x}_১ - \bar{x}_২}$ = পার্থক্যের প্রাক্কলিত পরিমিত ভ্রান্তি।

t-পরীক্ষা (t-test): z-পরীক্ষা নির্ণয়ের ক্ষেত্রে আমরা দেখেছি যে, সমগ্রকটিকে স্বাভাবিকভাবে বিন্যস্ত হতে হয় এবং সমগ্রকের পরিমিত ব্যবধানের মানটি জানা থাকে। কিন্তু সামাজিক গবেষণায়, খুব কম ক্ষেত্রেই সমগ্রকের পরিমিত ব্যবধান, বা ভেদাঙ্কের মানটি জানা থাকে। যখন সমগ্রকের পরিমিত ব্যবধানের মানটি জানা থাকে না এবং নমুনার আকারটি ছোট হয়, তখন আমরা t-বিন্যাসের মাধ্যমে সমগ্রকের গড় সম্পর্কিত অনুকল্প পরীক্ষা, দু'টি গড়ের মধ্যে পার্থক্য সম্পর্কিত অনুকল্প পরীক্ষা, নির্ভরশীল চলক সম্বলিত দু'টি গড়ের মধ্যে পার্থক্য সম্বলিত অনুকল্প পরীক্ষা এবং সহ-সম্পর্কের সহগ সম্পর্কিত অনুকল্প পরীক্ষা পরিচালনা করে থাকি। t-বিন্যাসের সাথে সম্পর্কিত একটি গুরুত্বপূর্ণ প্রত্যয় হলো, স্বাধীনতার মাত্রা। এই স্বাধীনতার মাত্রার ভিত্তিতেই t-পরীক্ষায় প্রাপ্ত মানের তাৎপর্যের বিষয়ে সিদ্ধান্ত গ্রহণ করা হয়ে থাকে। একটি নমুনার জন্য অনুকল্প পরীক্ষার ক্ষেত্রে, t-পরীক্ষা পরিসংখ্যান নির্ণয়ের সূত্রটি হলো,

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n - ১}}}$$

যেখানে, t = t-বিন্যাসের পর্যবেক্ষণকৃত মান

$\bar{x}$ = নমুনা গড়

μ = সমগ্রক গড়

$\frac{s}{\sqrt{n-১}}$ = নমুনা থেকে প্রাক্কলিত পরিমিত ব্যবধান।

দু'টি নমুনার মধ্যে পার্থক্যের তাৎপর্য পরীক্ষার জন্য t -পরীক্ষা পরিসংখ্যান নির্ণয়ের সূত্রটি হলো,

$$t = \frac{\bar{x}_১ - \bar{x}_২}{\sqrt{\frac{n_১ s_১^২ + n_২ s_২^২}{n_১ + n_২ - ২}} \sqrt{\frac{n_১ + n_২}{(n_১)(n_২)}}}$$

যেখানে, t = t -বিন্যাসের পর্যবেক্ষণকৃত মান

$\bar{x}_১$ = প্রথম নমুনার গড় মান

$\bar{x}_২$ = দ্বিতীয় নমুনার গড় মান

$n_১$ = প্রথম নমুনার আকার

$n_২$ = দ্বিতীয় নমুনার আকার

$s_১^২$ = প্রথম নমুনার ভেদাঙ্ক

$s_২^২$ = দ্বিতীয় নমুনার ভেদাঙ্ক

$n_১ + n_২ - ২$ = দু'টি নমুনার স্বাধীনতার মাত্রা।

সহ-সম্পর্কের সহগের তাৎপর্য পরীক্ষার বীজগাণিতিক সূত্রটি হলো,

$$t = r\frac{\sqrt{n-২}}{\sqrt{১-r^২}}$$

যেখানে, t = t -পরিসংখ্যান

r = পিয়ারসন r সহগ

$n-২$ = স্বাধীনতার মাত্রা।

সারাংশ

দুই-এর অধিক চলক নিয়ে যে বিশ্লেষণ করা হয়, তাকে বহু-চলক বিশিষ্ট বিশ্লেষণ বলে। বহু-চলক বিশিষ্ট বিশ্লেষণে, চলকসমূহের মধ্যে কার্য-কারণ সম্পর্ক ছাড়াও সম্পর্কের মাত্রাকে ব্যাখ্যা করা সম্ভব। কার্য-কারণ সম্পর্ক ব্যাখ্যার ক্ষেত্রে, পরিসংখ্যানের একাধিক পদ্ধতি ব্যবহৃত হয়ে থাকে। যেমন, বহুধা নির্ভরণ বিশ্লেষণ, পথ বিশ্লেষণ, উপাদান বিশ্লেষণ, ইত্যাদি। প্রতিটি ক্ষেত্রেই, গবেষককে গবেষণার সমস্যা এবং উপাত্তের ব্যবহার সম্পর্কে অনেক বেশী সাবধান থাকতে হয়। কোন উপাত্তের ক্ষেত্রে কোন ধরণের পরিসংখ্যান পদ্ধতি ভালো ফল দেবে, তা নির্ভর করে উপাত্তের যথার্থতা এবং নির্ভরশীলতার উপর।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক ($\surd$) চিহ্ন দিন –

১। পথ বিশ্লেষণ হলো:
   ক. গবেষককে প্রদর্শিত গবেষণার পথ
   খ. দ্বি-চলক বিশিষ্ট একটি মডেল
   গ. চলকগুলোর মধ্যে সম্পর্ককে বোঝার একটি কার্য-কারণ সম্পর্কিত মডেল
   ঘ. উপরের সব।

২। উপাদান বিশ্লেষণ ব্যবহৃত হয়:
   ক. বিভিন্ন চলকের মধ্যে সম্পর্কের মাত্রা চিহ্নিত করার জন্য
   খ. বিভিন্ন চলকের মানের ভিন্নতার মধ্যে নির্দিষ্ট রূপ চিহ্নিত করার জন্য
   গ. বিভিন্ন চলকের সংখ্যা ঠিক করার জন্য
   ঘ. বিভিন্ন চলকের মান চিহ্নিত করার জন্য।

৩। যখন সমগ্রকের পরিমিত ব্যবধানের মানটি জানা থাকে না, তখন আমরা _____ মাধ্যমে সমগ্রকের গড় সম্পর্কিত অনুকল্প পরীক্ষা পরিচালনা করে থাকি।
   ক. কাই-বর্গের
   খ. ভেদাঙ্ক বিশ্লেষণের
   গ. t-বিন্যাসের
   ঘ. z-পরীক্ষার

সংক্ষিপ্ত প্রশ্ন

টীকা লিখুন:

   ক. উপাদান বিশ্লেষণ
   খ. বহু-চলক বিশিষ্ট বিশ্লেষণ

রচনামূলক প্রশ্ন

১। বহু-চলক বিশিষ্ট বিশ্লেষণের গুরুত্ব উদাহরণসহ আলোচনা করুন।

২। উপাদান বিশ্লেষণ কী? আমাদের দেশে আধুনিকায়নের প্রতি মনোভাব নির্ধারণে উপাদান বিশ্লেষণ কেমন ফল দেবে বলে আপনি মনে করেন? আলোচনা করুন।

পাঠ - ৫

গবেষণা প্রতিবেদন
Research Report

এই পাঠ শেষে যা জানা যাবে —

- গবেষণা প্রতিবেদন
- একটি প্রতিবেদনে কী থাকা উচিৎ
- প্রতিবেদনের রচনাশৈলী
- সংক্ষিপ্ত প্রতিবেদন

গবেষণা প্রতিবেদন (Research Report)

যতক্ষণ পর্যন্ত না প্রতিবেদন লেখা হয়, ততক্ষণ পর্যন্ত গবেষণা কর্ম শেষ হয়েছে বলে মনে করা যাবে না। যত বুদ্ধিদীপ্ত অনুকল্প নির্মাণ করা হোক না কেন, যত সতর্কতার সাথে গবেষণা নক্শা প্রণয়ন এবং কার্যকর করা হোক না কেন, যত চমকপ্রদ উপাত্তই সংগ্রহ করা হোক না কেন, এ সবের কোন গুরুত্বই থাকবে না, যদি সেগুলো প্রকাশ না করা হয়। গবেষণার ফলাফল প্রচার গবেষকের দায়িত্বের মধ্যে একটি গুরুত্বপূর্ণ কাজ। কারণ, এর মাধ্যমে প্রাপ্ত ফলাফলকে জ্ঞান ভান্ডারের অংশ হিসাবে পরিণত করা যায়। কিন্তু অনেক সামাজিক বিজ্ঞানী মনে করেন যে, গবেষণা প্রতিবেদন প্রণয়নের মত একটি অপ্রীতিকর কাজকে গবেষণার সাথে যুক্ত করে দেওয়া হয়, এটি গবেষণার অবিচ্ছেদ্য অংশ নয়। উপাত্তের বিশ্লেষণ থেকে যখন দৃষ্টিটি প্রতিবেদন প্রণয়নের দিকে নিবদ্ধ করা হয়, তখন আবিষ্কারের সকল উত্তেজনা অনেক গবেষকের মধ্য থেকে উবে যায়।

প্রতিবেদন প্রণয়নের মূল উদ্দেশ্যটি নিজের সাথেই নিজের কথা বলা নয় বরং পাঠকের সাথে একটি যোগাযোগ স্থাপন করা।

অন্যান্য পর্যায়ের তুলনায়, গবেষণার এই পর্যায়ে ভিন্ন ধরণের দক্ষতার প্রয়োজন হয়। গবেষণা প্রতিবেদনের মাধ্যমে ফলাফল প্রকাশ, গবেষণার অন্যান্য পর্যায়ের মতোই যত্নের দাবী রাখে। প্রতিবেদন প্রণয়নে যদি পর্যাপ্ত সময় দেয়া হয়, তবে বিভিন্ন টুকরো টুকরো তথ্যগুলোকে একটি বৃহত্তর প্রেক্ষাপটে প্রয়োগের সুযোগটি গবেষক উপভোগ করতে পারেন। শুধু তাই নয়, ফলাফলের প্রতিক্রিয়া, এর মাধ্যমে বিদ্যমান জ্ঞানের শূন্যতা পূরণ, নতুন প্রশ্নের উত্থাপন ও ভবিষ্যৎ গবেষণার কি ক্ষেত্র তৈরি হতে পারে, তাও গবেষককে যথেষ্ট আন্দোলিত করতে পারে। প্রতিবেদন প্রণয়নের সময়, যে বিষয়টি সবচেয়ে গুরুত্বের সাথে মনে রাখা প্রয়োজন, সেটি হলো প্রতিবেদনের কার্যাবলী। প্রতিবেদন প্রণয়নের মূল উদ্দেশ্যটি নিজের সাথেই নিজের কথা বলা নয় বরং পাঠকের সাথে একটি যোগাযোগ স্থাপন করা। যদিও এ উদ্দেশ্যটি সকল গবেষকের মনেই থাকে, কিন্তু সামাজিক গবেষণার বিভিন্ন প্রতিবেদন পর্যালোচনা করলে দেখা যায় যে, অধিকাংশ প্রতিবেদনে গবেষক, বা প্রতিবেদন প্রণয়নকারী তার নিজস্ব চিন্তা-ভাবনার ব্যাখ্যা দিতে গিয়েই গলদঘর্ম হয়েছেন। সে ধরণের গবেষণা প্রতিবেদন পাঠকের মনে কোন ঔৎসুক্যের জন্ম দেয় না।

পাঠকের সাথে যোগাযোগ প্রতিষ্ঠার মূল শর্তটি হলো, প্রতিবেদনটি কোন ধরণের পাঠকের জন্য তৈরি করা হচ্ছে, সে সম্পর্কে প্রতিবেদন প্রণয়নকারীর একটি স্বচ্ছ ধারণা থাকা প্রয়োজন। গবেষণার ফলাফলের ভিত্তিতে বিভিন্ন কর্মকান্ড বাস্তবায়নকারী প্রসাশকের উদ্দেশ্যে প্রণীত প্রতিবেদনের তুলনায়, সহকর্মী সামাজিক বিজ্ঞানীদের উদ্দেশ্যে প্রণীত গবেষণা প্রতিবেদন অনেক ক্ষেত্রেই ভিন্ন হবে। আবার সাধারণ মানুষকে অবগতকরণের উদ্দেশ্যে প্রণীত গবেষণা প্রতিবেদন, উপরিল্লিখিত দু'ধরণের প্রতিবেদন থেকে ভিন্ন হবে। যে ধরণের পাঠকই হোক না কেন, প্রতিবেদন প্রণয়নের পরিকল্পনায় দু'টি প্রধান বিষয় গুরুত্বের সাথে বিবেচনায় নিতে হবে। প্রথমতঃ, গবেষণার কোন দিকটি সম্পর্কে পাঠক জানতে চান, বা তাদের জানা প্রয়োজন এবং দ্বিতীয়তঃ, উপাত্তকে কতটুকু উত্তমভাবে উপস্থাপন করা যায়। অর্থাৎ, একটি গবেষণা প্রতিবেদনে কি থাকা উচিৎ এবং প্রতিবেদনের রচনাশৈলীটি কি হবে?

একটি প্রতিবেদনে কী থাকা উচিৎ (What a Report Should Contain)?

একটি গবেষণা প্রতিবেদনে, গবেষণা সমস্যার একটি স্পষ্ট ও বিস্তারিত বর্ণনা থাকতে হবে; গবেষণা পদ্ধতি এবং চলকসমূহের জন্য বিভিন্ন মাত্রায় ব্যবহৃত পরিমাপ পদ্ধতির পর্যাপ্ততার যৌক্তিকতা তুলে ধরতে হবে; নমুনা নির্বাচন এবং উপাত্ত সংগ্রহ কত সতর্কতা ও যত্নের সাথে করা হয়েছে, তার বিবরণ থাকতে হবে; উপাত্ত বিশ্লেষণ ও ব্যাখ্যার মাধ্যমে ফলাফলের যৌক্তিকতা ও যথার্থতার প্রমাণ উপস্থাপন করতে হবে; এবং ফলাফলের অন্তর্নিহিত অর্থ ও এর সম্ভাব্য প্রভাবগুলোকে আলোচনা করতে হবে। বিষয়গুলো কিছুটা বিস্তারিতভাবে আলোচনা করা যাক।

সমস্যার বর্ণনা (Statement of the Problem): আমরা জানি যে, গবেষণা প্রক্রিয়ার প্রথম পদক্ষেপটি হলো, গবেষণা সমস্যার সূক্ষ্ম সূত্রবদ্ধকরণ করা। সাধারণভাবে, গবেষণা প্রতিবেদনও শুরু হয় যে সমস্যা নিয়ে গবেষণা করা হয়েছে, তার বর্ণনা দিয়ে। কেন এই সমস্যাটিকে নিয়ে গবেষণার জন্য বিবেচনা করা হয়েছে সে সম্পর্কে পাঠককে স্পষ্ট ধারণা দেবার জন্য পর্যাপ্ত পটভূমিমূলক তথ্য দিতে হবে। যেহেতু একজন সমাজবিজ্ঞানী একটি নির্দিষ্ট বাস্তব সমস্যা সমাধানের চাইতে, মানব আচরণের সাধারণ জ্ঞানভান্ডারে গবেষণাটির অবদান সম্পর্কে বেশী আগ্রহী হয়ে থাকেন, সেহেতু পাঠকের জন্য সমাজতাত্ত্বিক তত্ত্বের সাথে গবেষণা সমস্যার প্রাসঙ্গিকতাকে প্রতিষ্ঠা করতে হবে। কিন্তু এটিও স্বীকার করতে হবে যে, সব গবেষণার উপরই সরাসরি তত্ত্বগত বিষয়ের প্রভাব থাকে না। প্রাসঙ্গিকতাটুকু তখনই প্রতীয়মান হয়, যখন গবেষক তার ফলাফলকে কোন তাত্ত্বিক কাঠামোর ভিত্তিতে বুঝতে চান। সামাজিক বিজ্ঞানে, বহু গবেষণা পরিচালিত হয় একটি সুসংবদ্ধ তত্ত্বের নির্দেশনা ছাড়াই। সে ক্ষেত্রে, তাত্ত্বিক প্রাসঙ্গিকতা প্রতিষ্ঠার জন্য অহেতুক ভান করার কোন প্রয়োজন নেই। তাত্ত্বিক ও বাস্তবসম্মত গুরুত্বের পাশাপাশি, অন্যান্য প্রাসঙ্গিক গবেষণার একটি পর্যালোচনা সমস্যার বর্ণনায় অন্তর্ভুক্ত করা উচিৎ। এতে করে, গবেষণার পরিপ্রেক্ষিতটি বোঝা যায়। যদি কোন অনুকল্প নির্মাণ করা হয়, তবে সে অনুকল্পগুলো এবং অনুকল্পে ব্যবহৃত প্রধান প্রধান প্রত্যয়গুলোর সংজ্ঞাকেও সুস্পষ্টভাবে বর্ণনা করতে হবে।

গবেষণা পদ্ধতি (Research Methodology): কিভাবে গবেষণাটি পরিচালনা করা হয়েছে, একজন বিজ্ঞানসম্মত পাঠক তা জানতে চান। এ সম্পর্কিত তথ্য সরবরাহ করতে গিয়ে গবেষককে গবেষণা নক্শার প্রকৃতিটি কি ছিলো, তা বর্ণনা করতে হবে। গবেষণা নক্শাটি যদি পরীক্ষণমূলক হয়ে থাকে, তবে পরীক্ষণগুলো কিভাবে পরিচালনা করা হয়েছিলো, পরীক্ষণের সাথে সম্পর্কিত পরিমাপ প্রক্রিয়াটি কখন এবং কিভাবে পরিচালনা করা হয়েছিলো, তার পদ্ধতিগত প্রক্রিয়ার উল্লেখ করতে হবে। যদি প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীর মাধ্যমে উপাত্ত সংগ্রহ করা হয়ে থাকে, তবে প্রকৃতপক্ষে কি কি প্রশ্ন করা হয়েছিলো, পাঠককে তার একটি ধারণা দিতে হবে। এ ক্ষেত্রে, প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীর একটি অনুলিপি পরিশিষ্টে সংযোজিত করা যেতে পারে। সাক্ষাৎকার গ্রহণকারীদের নির্বাচন প্রক্রিয়া, তাদের যোগ্যতা নির্ধারণ ও প্রশিক্ষণ প্রক্রিয়ার বর্ণনা দিতে হবে।

পরিমাপগুলো যদি পর্যবেক্ষণের ভিত্তিতে হয়ে থাকে, তবে পর্যবেক্ষককে কি নির্দেশাবলী দেওয়া হয়েছিলো এবং পর্যবেক্ষণ বা প্রশ্নের উত্তরগুলো কিভাবে গবেষণার সাথে সম্পর্কিত চলকগুলোর পরিমাপে রূপান্তর করা হয়েছিলো, সে সকল বিষয়ে জানার জন্য পাঠক উৎসাহী হয়ে থাকেন। কাদেরকে নমুনা হিসাবে নির্বাচন করা হয়েছিলো, নমুনা সংখ্যা কি ছিলো এবং তাদের কিভাবে নির্বাচন করা হয়েছিলো, এ সম্পর্কিত তথ্যগুলো ফলাফলের সাধারণীকরণ যোগ্যতার মাত্রা নিরূপণে খুবই গুরুত্বপূর্ণ। অল্প সংখ্যক নমুনার উপর ভিত্তি করে ব্যাপক উপসংহার টানা হয়েছিলো কি না, হয়ে থাকলে তার যৌক্তিকতা কি ছিলো, এ সব তথ্য পাঠকের জন্য জানা গুরুত্বপূর্ণ। কারণ, এ বিষয়গুলো গবেষণায় অন্তর্ভুক্ত নমুনার প্রতিনিধিত্বশীলতাকেই নির্দেশ করে। নমুনায় অন্তর্ভুক্ত জনগোষ্ঠীর বৈশিষ্ট্য সম্পর্কে পর্যাপ্ত বর্ণনা থাকা প্রয়োজন, যাতে করে অন্যান্য জনগোষ্ঠীর ক্ষেত্রে এই ফলাফল প্রয়োগ করা যাবে কি না, পাঠক নিজেই সে সম্পর্কে উপসংহার টানতে পারেন। উপাত্তের পরিসংখ্যানগত বিশ্লেষণের কৌশলগুলো কি ছিলো, বিজ্ঞানসম্মত পাঠক সে বিষয়েও জানতে চাইবেন। অর্থাৎ, বিভিন্ন গোষ্ঠীর বৈশিষ্ট্যের মধ্যে পার্থক্য নির্ধারণের জন্য কি কৌশল ব্যবহার করা হয়েছিলো তা বর্ণনা করতে হবে। সাধারণভাবে, কৌশলগুলো এবং এর সাথে সংশ্লিষ্ট আস্থার মাত্রা উল্লেখ করলেই চলে, কিন্তু যদি কৌশলটি নতুন বা উদ্ভাবনীমূলক হয়ে থাকে, সে ক্ষেত্রে তার ব্যাখ্যা প্রয়োজন।

বিজ্ঞানসম্মত লেখক কি লিখবেন এবং কি লিখবেন না, এর উপর সিদ্ধান্ত নেওয়ার ব্যাপারে অন্যান্য লেখকদের মত তিনি স্বাধীন নন।

ফলাফল (Results): গবেষণার ফলাফল উপস্থাপনের মূল নীতিমালাটি হলো, গবেষণা সমস্যার সাথে সম্পর্কিত গবেষণার সব প্রমাণকে উপস্থাপন করতে হবে, তা গবেষকের মতের সাথে সঙ্গতিপূর্ণ হোক বা না হোক। বিজ্ঞানসম্মত প্রতিবেদন প্রণয়নের এটি হলো একটি অপরিহার্য শর্ত। বিজ্ঞানসম্মত লেখক কি লিখবেন এবং কি লিখবেন না, এর উপর সিদ্ধান্ত নেওয়ার ব্যাপারে অন্যান্য লেখকদের মত তিনি স্বাধীন নন। কিন্তু তারপরেও গবেষককে নির্বাচন করতে হয় যে, প্রতিবেদনে উপস্থাপন করার জন্য কি প্রাসঙ্গিক, কি নয়। উপাত্ত বিশ্লেষণের পর্যায়ে প্রস্তুতকৃত প্রতিটি সারণি প্রতিবেদনে স্থান পায় না, বা পাওয়া উচিৎ নয়। সে ক্ষেত্রে, প্রতিবেদনে উপস্থাপনের জন্য কি কি সারণি ব্যবহার করা হবে, তা সুচিন্তিতভাবে নির্ধারণ করতে হবে। তাহলে প্রশ্ন হলো, কোন সারণিটি প্রাসঙ্গিক সে বিষয়ে গবেষক কিভাবে সিদ্ধান্ত নেবেন? এ প্রসঙ্গে, গবেষণা সমস্যা, গবেষণার উদ্দেশ্য, উপাত্তের প্রকৃতি এবং যদি কোন অনুকল্প নির্মাণ করা হয়ে থাকে, সেগুলোর ভিত্তিতে নির্দেশনাটি আসতে পারে।

যখন গবেষণা সমস্যাটিকে সূত্রবদ্ধকরণ করা হয়েছিলো, তখন গবেষণাটি স্ফটিকের মত স্বচ্ছ ও স্পষ্ট ছিলো না। গবেষণা পরিচালনা করতে গিয়ে সমস্যার বর্ণনাটি পর্যাপ্ততার দিকে এগুতে থাকে, নতুন অনুকল্পের জন্ম হয় এবং পূর্বে চিন্তা করা হয় নি, এমন সম্পর্কও ধরা পড়ে। অতএব, সমস্যার মূল সূত্রবদ্ধকরণটি প্রতিবেদনের মৌলিক স্মারক হিসাবে কাজ করলেও, পরবর্তী পর্যায়ের বিকাশগুলোকে এর মধ্যে অন্তর্ভুক্ত করতে হবে। চলকের মধ্যে সম্পর্ক, বা বিভিন্ন গোষ্ঠীর মধ্যে তুলনার সাথে সম্পর্কিত প্রতিটি ফলাফলের বিশ্লেষণ ও ব্যাখ্যা এবং এর সাথে সংশ্লিষ্ট পরিসংখ্যানগত তাৎপর্যের মাত্রাকে উল্লেখ করতে হবে। সাধারণভাবে, একটি নির্দিষ্ট পাঠকদলের জন্য প্রস্তুতকৃত বিস্তারিত প্রতিবেদনে, যে সকল ফলাফল গুরুত্বপূর্ণ মনে হয় সেগুলোর সাথে সম্পর্কিত সারণি, লেখচিত্র বা ছক হয় মূল প্রতিবেদনে, বা পরিশিষ্টে অন্তর্ভুক্ত করতে হবে।

আমরা এতক্ষণ এমন সব গবেষণা প্রতিবেদনের কথা আলোচনা করেছি, যার মূল নির্দেশনাটি এসেছে গবেষণা সমস্যার প্রাথমিক সূত্রবদ্ধকরণ, ফলাফলের পরিসংখ্যানগত বিশ্লেষণ, এবং অনুকল্প থেকে। কিন্তু সব গবেষণাই গবেষণা সমস্যার সূত্রববদ্ধকরণের বিস্তারিত বিবরণ দিয়ে শুরু হয় না এবং সব গবেষণাই পরিসংখ্যানগত উপাত্ত ব্যবহার করে না। যেমন, একটি অনুসন্ধানমূলক গবেষণায়, প্রতিবেদনের বিষয়বস্তুর যথাযথ কাঠামো কি হবে, তা গবেষণা নক্শা, বা উপাত্তের বিশ্লেষণ সুস্পষ্টভাবে নির্দেশ করতে পারে না। এ ক্ষেত্রে, প্রতিবেদনের রূপরেখাটি কি হবে, সে বিষয়ে সিদ্ধান্ত গ্রহণের জন্য গবেষককে নিজের বিবেচনাপ্রসূত সিদ্ধান্তের উপরই নির্ভর করতে হবে। তবে, যে ধরণের গবেষণাই হোক না কেন, সমস্যার যে বিষয়গুলো গবেষণা পরিচালনায় প্রভাব বিস্তারকারী উপাদান হিসাবে কাজ করেছিলো, সেগুলোসহ যে পদ্ধতিতে গবেষণা পরিচালনা করা হয়েছিলো, যে উপসংহার টানা হয়েছিলো এবং চূড়ান্ত উপসংহারের ভিত্তি কি ছিলো, তা স্পষ্টভাবে উল্লেখ করতে হবে।

অন্তর্নিহিত প্রভাবের আলোচনা (Discussion of Implications): ফলাফলের প্রকৃত অর্থ বহন করার জন্য, ফলাফলের সাদামাটা বর্ণনাই যথেষ্ট নয়। মানব আচরণকে উপলব্ধির জন্য পাঠক ফলাফলের অন্তর্নিহিত অর্থ ও এর প্রভাবকে জানতে চান। এ ধরণের অন্তর্নিহিত প্রভাবের আলোচনা, কখনো কখনো উপাত্ত বিশ্লেষণ ও ব্যাখ্যার সাথে উপস্থাপন করা হয়, আবার কখনো তা ভিন্ন অনুচ্ছেদেও উপস্থাপন করা হয়। যেখানেই এই আলোচনা উপস্থাপন করা হোক না কেন, এটি সাধারণতঃ চারটি প্রধান বিষয়কে অন্তর্ভুক্ত করে। প্রথমতঃ, একটি বিশেষ পরিস্থিতির ফলাফল থেকে আহরিত সিদ্ধান্তমূলক উপসংহার, যা একই রকম অন্যান্য পরিস্থিতিতে প্রযোজ্য হবে বলে প্রত্যাশা করা যেতে পারে। এ ধরণের উপসংহার, উপাত্তের খুব কাছাকাছি মাত্রায় হতে পারে, অথবা বিমূর্ততার উচ্চমাত্রায়ও হতে পারে। দ্বিতীয়তঃ, এ সকল সিদ্ধান্তমূলক উপসংহারের ব্যাখ্যা হিসাবে, গবেষণায় প্রাপ্ত ফলাফলের সাধারণীকরণের সীমাবদ্ধতার মাত্রা যে সকল শর্ত দ্বারা প্রভাবিত হয়, গবেষককে সেগুলোর উল্লেখ করতে হবে। গবেষণায় অন্তর্ভুক্ত নমুনার বৈশিষ্ট্য বৃহত্তর জনগোষ্ঠী থেকে কতটুকু ভিন্ন, তার উল্লেখ করতে হবে। পদ্ধতিগত সীমাবদ্ধতা থাকলে, তা ব্যতিক্রমধর্মী ফলাফলে কি প্রভাব বিস্তার করেছে, সেগুলো উল্লেখ করতে হবে। তৃতীয়তঃ, যে সকল প্রাসঙ্গিক প্রশ্নের উত্তর পাওয়া যায়নি, বা যে সকল নতুন প্রশ্ন উত্থাপিত হয়েছে, সে সকল প্রশ্নের উত্তর ভবিষ্যতে কি ধরণের গবেষণার মাধ্যমে পাওয়া যাবে, ফলাফলের অন্তর্নিহিত প্রভাবের আলোচনায় তার নির্দেশনা থাকতে হবে। চতুর্থতঃ, গবেষণার ফলাফল তাত্ত্বিক বিকাশে কি মৌলিক অবদান রাখবে, বা নীতি নির্ধারণের ক্ষেত্রে কি প্রভাব বিস্তার করবে, তার সুস্পষ্ট উল্লেখ থাকতে হবে।

সারসংক্ষেপ ও উপসংহার (Summary and Conclusions): প্রথা অনুযায়ী, একটি সংক্ষিপ্ত সারসংক্ষেপের মাধ্যমে উপসংহার টানতে হবে। এ প্রসঙ্গে, সমস্যার একটি সংক্ষিপ্ত পুনর্বর্ণনা, গবেষণা পদ্ধতি, প্রধান ফলাফল এবং সে সব ফলাফল থেকে আহরিত উপসংহারগুলোকে অন্তর্ভুক্ত করতে হবে। পুরো গবেষণাটি সম্পর্কে একটি সার্বিক ধারণা নেবার জন্য ব্যস্ত পাঠক, মূলতঃ এই অংশটিই পড়বেন। সে জন্য সারসংক্ষেপ ও উপসংহার লেখার সময়, কিছুটা অতিরিক্ত গুরুত্ব দিয়ে পুরো গবেষণার একটি সঠিক এবং নির্ভুল চিত্র উপস্থাপন করা বুদ্ধিমানের কাজ হবে। এই অংশে, প্রতিবেদন প্রণয়নকারী খুব রীতিবদ্ধভাবে গবেষণার ফলাফলকে সংক্ষিপ্ত আকারে উপস্থাপন করবেন। গবেষণার শুরুতে গবেষণা প্রশ্নগুলো যেভাবে প্রস্তাব করা হয়েছে, সেই প্রস্তাবনার পরিপ্রেক্ষিতে ফলাফলগুলোকে উপস্থাপন করতে হবে। প্রয়োজনীয় উপসংহারগুলো উপস্থাপনের পর, যে সমস্যার উপর গবেষণা পরিচালিত হয়েছে সে সমস্যাটিকে কিভাবে সমাধান করা যায়, বা আরো গভীরভাবে উপলব্ধি করা যাবে, গবেষক সে সম্পর্কিত কিছু সাধারণ এবং সুনির্দিষ্ট সুপারিশ প্রদান করবেন।

গ্রন্থ-স্মারক (References): প্রতিটি গবেষণা প্রতিবেদনেই একটি গ্রন্থ-স্মারক সন্নিবেশিত হওয়া উচিৎ। যে সকল গ্রন্থ, গবেষণা প্রতিবেদন, প্রবন্ধ, নিবন্ধ, ইত্যাদি প্রতিবেদনে ব্যবহার করা হয়েছে, সেগুলোর

একটি পূর্ণ তালিকা বর্ণমালার ক্রম অনুযায়ী গবেষণা প্রতিবেদনের শেষে উপস্থাপন করতে হবে, যাতে করে পাঠক প্রতিবেদনে উল্লিখিত গ্রন্থ-স্মারকের মূল উৎসে গিয়ে, বিষয়টি সম্পর্কে আরো বিস্তারিতভাবে জানতে পারেন। গ্রন্থ-স্মারক উপস্থাপনের বিভিন্ন রীতি রয়েছে। যেমন, হার্ভার্ড রীতি, ভ্যাঙ্কুভার রীতি, ইত্যাদির কথা উল্লেখ করা যেতে পারে। যে রীতিই অনুসরণ করা হোক না কেন, তা যেন বিজ্ঞানসম্মত পাঠক সমাজের কাছে গ্রহণযোগ্য একটি রীতি হয়। গ্রন্থ-স্মারক কোন রীতিতে তৈরি করা হবে, তা নির্ভর করে প্রতিবেদনটি কার উদ্দেশ্যে লেখা হচ্ছে, তার উপর। যদি গ্রন্থ আকারে প্রকাশের জন্য লেখা হয়, সে ক্ষেত্রে প্রকাশক যে রীতিটি অনুসরণ করতে বলবেন, সেটি করতে হবে। যদি সাময়িকীতে প্রকাশের জন্য লেখা হয়, সে ক্ষেত্রে সাময়িকীর সম্পাদক যেভাবে চাইবেন, সে রীতিতে উপস্থাপন করতে হবে। আর যদি ব্যক্তিগত উদ্দ্যোগে প্রকাশ, বা কোন প্রতিষ্ঠানের উদ্দেশ্যে লেখা হয়, সে ক্ষেত্রে প্রতিবেদন প্রণয়নকারীর নিজস্ব পছন্দের রীতিটি অনুসরণ করতে পারেন। যে উদ্দেশ্যেই লেখা হোক না কেন, প্রতিবেদন প্রণয়নকারীকে লক্ষ্য রাখতে হবে যে, তা যেন সাধারণভাবে গৃহীত ও পরিচিত রীতিটিকে অনুসরণ করে।

প্রতিবেদনের রচনাশৈলী (Style of the Report)

একটি ভালো বিজ্ঞানসম্মত লেখার মৌলিক গুণাবলী হলো, সঠিকতা ও স্পষ্টতা। নান্দনিক বা সাহিত্যের দৃষ্টিভঙ্গি থেকে প্রীতিকর রচনাশৈলী হলো, এর একটি উপরি পাওনা। একটি প্রতিবেদনকে রঙীন ও রুচিশীল করে তোলার জন্য গবেষকের কোন বাধ্যবাধকতা না থাকলেও, পাঠককে দীর্ঘ, অস্পষ্ট ও বাগাড়ম্বরপূর্ণ বাক্য, বা অনুচ্ছেদের মধ্যে দিয়ে অতিক্রম করতে বাধ্য করার মাধ্যমে, তাকে দূরে সরিয়ে দেয়ার কোন যৌক্তিকতা নেই। গবেষককে সিদ্ধান্ত নিতে হবে যে, কোন তথ্য তিনি পাঠকের কাছে পৌঁছাতে চান। বিভিন্ন বিষয়গুলো পরস্পরের সাথে কিভাবে সম্পর্কযুক্ত, তার ব্যাখ্যা প্রদান করতে হবে। এ কাজটি সঠিকভাবে সম্পাদন করতে পারলে, অর্ধেক সমস্যার সমাধান হয়ে যায়। এ পর্যায়ে, প্রতিবেদনের জন্য একটি বিস্তারিত রূপরেখা প্রণয়ন করা খুবই গুরুত্বপূর্ণ। কিছু অভিজ্ঞ গবেষক রয়েছেন, যারা প্রতিবেদনের চূড়ান্ত রূপটি কি হবে, সে সম্পর্কে একটি ধারণা মাথায় নিয়ে কাগজ-কলম বা একটি কম্পিউটার নিয়ে বসে পড়লেই, তাদের গবেষণার একটি স্পষ্ট এবং সুসংগঠিত বর্ণনা লিখে ফেলতে পারেন। কিন্তু অধিকাংশ গবেষকের ক্ষেত্রে, বিশেষ করে নব্য গবেষকদের ক্ষেত্রে, এটি একটি অকার্যকর পদ্ধতি। ভাসাভাসা ধারণা থেকে প্রতিবেদন লেখার কাজ শুরু করলে, তা এমন একটি অসম্পূর্ণ, অসংলগ্ন ও অর্থহীন প্রতিবেদনে পরিণত হবে, যা আবার নতুন করে লেখার প্রয়োজন হয়ে পড়বে।

একটি বিস্তারিত রূপরেখা, একজন গবেষক বা প্রতিবেদন প্রণয়নকারীকে 'প্রতিবেদন কিভাবে লিখতে হবে', সে বিষয়ে দুশ্চিতা না করে, 'কি লিখতে হবে', তার উপর নিরঙ্কুশভাবে মনোনিবেশ করতে সাহায্য করে। এর জন্য প্রথমে, প্রতিবেদনের একটি কঙ্কালসার কাঠামো নির্মাণ করতে হবে এবং এরপর পরীক্ষা করে দেখতে হবে যে, আরো কোন গুরুত্বপূর্ণ বিষয় সেখানে যুক্ত করার প্রয়োজন রয়েছে কি না। রূপরেখার অন্তর্ভুক্ত প্রধান প্রসঙ্গ, উপ-প্রসঙ্গ এবং তার উপ-প্রসঙ্গগুলো, একজন গবেষককে সে সকল প্রসঙ্গের মধ্যে সম্পর্কের যৌক্তিকতাকে দেখতে বাধ্য করে। একজন প্রতিবেদন প্রণয়নকারী যদি একটি রূপরেখা ছাড়া লেখা শুরু করেন, তবে সে প্রতিবেদনটি হবে, খাপছাড়া ধরণের। যেমন, তিনি নমুনায়নের উপর দু'টি বা তিনটি অনুচ্ছেদ লেখার পর, উপাত্ত সংগ্রহের কৌশলের বর্ণনায় চলে যাবেন এবং চিন্তার ধারার মধ্যে কি ঘটেছে তার দিকে দৃষ্টি না দিয়েই, নমুনা সম্পর্কে হঠাৎ মনে পড়ে যাওয়া বিষয়কে লিখতে শুরু করবেন। একটি রূপরেখা, একটি প্রসঙ্গ থেকে অন্য প্রসঙ্গে কিভাবে উত্তোরণ ঘটে, তা প্রদর্শনের মাধ্যমে একটি প্রসঙ্গের মধ্যে কি কি বিষয় আলোচিত হবে, তা গোছাতে সাহায্য করে।

একটি রূপরেখা, একটি প্রসঙ্গ থেকে অন্য প্রসঙ্গে কিভাবে উত্তোরণ ঘটে, তা প্রদর্শনের মাধ্যমে একটি প্রসঙ্গের মধ্যে কি কি বিষয় আলোচিত হবে, তা গোছাতে সাহায্য করে।

একবার যখন রূপরেখাটি প্রণয়ন করা হয়ে যায়, তখন গুরুত্বপূর্ণ কোন কিছু বাদ পড়েছে কি না এবং ধারণাগুলোকে যৌক্তিকভাবে বিভক্ত করা হয়েছে কি না, তা দেখার জন্য এটিকে আবার পর্যালোচনা করে দেখা ভালো। এতে করে, লেখা শুরু করার পূর্বে প্রয়োজনীয় সংশোধনগুলো করে নেয়া যায়। প্রথম খসড়া তৈরির পর্যায়ে, একজন গবেষক প্রতিবেদনের আঙ্গিক ও রচনাশৈলী নিয়ে দুশ্চিন্তা করবেন কি না, তা ব্যক্তি বিশেষের স্বভাব ও বৈশিষ্ট্যের উপর নির্ভর করে। কোন কোন প্রতিবেদন প্রণয়নকারী প্রতিটি বাক্যের জন্য দীর্ঘ সময় ব্যয় করেন এবং সতর্কতার সাথে শব্দ চয়ন করেন, যাতে করে সেগুলোর অর্থকে সবচেয়ে উত্তমরূপে প্রকাশ করা যায়। এতে কোন দোষের কিছু নেই, কিন্তু যখন সাহিত্যিক মূল্যের চেয়ে তথ্যকে পাঠকের কাছে পৌঁছে দেয়াই প্রধান উদ্দেশ্য হয়, তখন প্রথম খসড়াটি যত তাড়াতাড়ি সম্ভব লিখে ফেলাই যুক্তিসঙ্গত। এতে করে, সময়ের সাশ্রয় হয় অনেক বেশী। যখন হাতের কাছে একটি খসড়া থাকে, তখন বাক্য বা অনুচ্ছেদকে পছন্দসই করে লেখার জন্য পরবর্তীতে পর্যাপ্ত সময় ব্যয় করা যেতে পারে।

কয়েকটি জট পাকানো বাক্যাংশ দিয়ে তৈরি একটি জটিল বাক্যের পরিবর্তে, দুই বা তিনটি সরল বাক্য একটি ধারণাকে অনেক বেশী সুস্পষ্টভাবে প্রকাশ করে।

কিন্তু একসময় না একসময়, রচনাশৈলীর প্রতি মনোযোগ দেওয়া অত্যাবশ্যকীয় হয়ে পড়ে। এ প্রসঙ্গে, যে বিষয়টিকে প্রথমেই গুরুত্ব দেয়া প্রয়োজন, তা হলো সরলতা এবং শুদ্ধ ব্যাকরণগত কাঠামো। এ পর্যায়ে, যে

সাধারণ ত্রুটিটি লেখক করেন, সেটি হলো, দাম্ভিকতার সাথে কৃতিত্বের দাবী করা। সামাজিক বিজ্ঞানে গুরুত্ব রয়েছে, এমন কলাকৌশলগত শব্দ বা পদাবলীর ব্যবহার অবশ্যই পুরোপুরিভাবে যথাযথ, কিন্তু সামাজিক বিজ্ঞানের অধিকাংশ প্রতিবেদন জটিল ও দূর্বোধ্য কলাকৌশলগত ভাষার সমন্বয়ে রচিত হয় না। এক একক সম্বলিত শব্দ (one syllable word) ব্যবহারের পরিবর্তে, অহেতুক চার একক সম্বলিত শব্দ ব্যবহার করে জটিল বাক্য গঠনের প্রয়োজন নেই। কারণ, কয়েকটি জট পাকানো বাক্যাংশ দিয়ে তৈরি একটি জটিল বাক্যের পরিবর্তে, দুই বা তিনটি সরল বাক্য একটি ধারণাকে অনেক বেশী সুস্পষ্টভাবে প্রকাশ করে।

অনেকে মনে করেন যে, ব্যাকরণগত নিয়মাবলী হলো পন্ডিতদের আওতাধীন, যা ইংরেজী শিক্ষক ছাড়া অন্য কারো দ্বারা দক্ষতার সাথে ব্যবহার করা সম্ভব নয়। এটি একটি মারাত্মক ভুল ধারণা। কারণ, ব্যাকরণগত নিয়মাবলীর অন্যতম প্রধান কাজ হলো, ভাষা ব্যবহারে আমাদের এমনভাবে সাহায্য করা, যাতে করে তা যতটুকু সম্ভব কম অস্পষ্টতার সাথে অর্থকে প্রতিভাত করে তোলে। কোন একটি নির্দিষ্ট বাক্যাংশের পর, বিরতিচিহ্ন ব্যবহার করা কোন স্বেচ্ছাচারী সিদ্ধান্ত নয়। এটি যে শব্দটিকে বিশ্লেষিত করে, তার সাথে বাক্যাংশের সম্পর্কটিকে সুনির্দিষ্ট করে। শব্দ চয়ন, বাক্য কাঠামো, অনুচ্ছেদ নির্মাণ, সংশোধন এবং বিরতিচিহ্নের যথাযথ ব্যবহারের লক্ষ্যে, একটি ভালো অভিধান ও একটি ভালো রচনাশৈলীর নির্দেশিকা, লেখকের জন্য গুরুত্বপূর্ণ হাতিয়ার হিসাবে কাজ করে। সারণি, লেখচিত্র, চিত্র ও বিভিন্ন ছক নির্মাণে যথেষ্ট যত্ন নেয়া প্রয়োজন। সকল সারণি এবং চিত্রকে সুস্পষ্টভাবে শিরোনামযুক্ত হতে হবে। শিরোনামগুলোকে সংক্ষিপ্তভাবে সারণিতে উপস্থাপিত বিষয়কে ধারণ করতে হবে। সারণির পাদটিকায় প্রয়োজনীয় ব্যাখ্যা প্রদান করতে হবে। স্তম্ভ ও সারির শিরোনামগুলোকে উপাত্তের সাথে সংক্ষিপ্ত হতে হবে। যখন উপাত্তকে শতকরা হারে উপস্থাপন করা হয়, তখন তার পাশাপাশি ঘটনসংখ্যাকে উল্লেখ করতে হবে।

যখন যথাযথ সারণি ও লেখচিত্র সংযুক্ত করে প্রতিবেদন প্রণয়নের কাজ সমাপ্ত করা হয়, তখন প্রতিবেদনটিকে পুনরায় আদ্যপান্ত পড়ে কতগুলো প্রশ্নের উত্তর পেতে হবে। যেমন, বাক্যগুলো কি স্পষ্ট হয়েছে? সেগুলো কি ব্যাকরণগতভাবে শুদ্ধ হয়েছে? যা প্রতিবেদক বলতে চেয়েছেন, তা কি বাক্যগুলোর মাধ্যমে প্রকাশিত হয়েছে? বিষয়গুলোকে কি আরো সরলভাবে ব্যক্ত করা যেত? সারণির মধ্যে উপস্থাপিত উপাত্তগুলো কি, যে উপসংহার টানা হয়েছে, তার যৌক্তিকতাকে প্রমাণ করে? ইত্যাদি। প্রতিবেদনকে চূড়ান্ত করার পূর্বে, অন্ততঃপক্ষে একজন সহকর্মীকে দিয়ে তা পড়িয়ে নেয়া উচিৎ। যে বাক্যগুলো লেখকের কাছে স্ফটিক-স্বচ্ছ মনে হয়, অন্যের কাছে তা বিভ্রান্তিকর মনে হতে পারে। লেখকের কাছে মনে হওয়া যৌক্তিক সম্পর্কগুলোকে, অন্যদের কাছে অযৌক্তিক ও পারম্পর্যহীন মনে হতে পারে। অস্বচ্ছ ও অযৌক্তিক অনুচ্ছেদগুলোকে চিহ্নিত করা ও সেগুলো শোধরানোর নির্দেশনার মাধ্যমে একটি বন্ধুত্বপূর্ণ সমালোচনা, যথাযথ যোগাযোগ প্রতিষ্ঠার জন্য অমূল্য হয়ে উঠে।

সংক্ষিপ্ত প্রতিবেদন (Shorter Report)

অনেক সময়, কিছু কিছু গবেষণার ক্ষেত্রে, অর্থ এবং সময় ব্যয় করে বিস্তারিত প্রতিবেদন প্রণয়নের প্রয়োজন পড়ে না, বা যুক্তিযুক্ত হয় না। কারণ, হয়তো গবেষণাটি ততোটা গুরুত্বপূর্ণ নয় যে, তা ব্যাপক জনগোষ্ঠীর মধ্যে প্রচার করতে হবে। বিশেষ করে, গবেষণা সাময়িকীতে প্রকাশের জন্য প্রবন্ধ আকারে লিখিত প্রতিবেদন, কয়েক পাতার মধ্যে সীমাবদ্ধ থাকলেই চলে। এতে করে, একইসাথে গবেষণার ফলাফল সম্পর্কে সংশ্লিষ্ট পাঠককে জানানোর উদ্দেশ্যটি অর্জিত হয় এবং সময় ও অর্থ দু'য়েরই সাশ্রয় ঘটে। এ ধরণের প্রতিবেদন, খুবই সংক্ষিপ্ত হয়ে থাকে। গবেষণা সাময়িকীতে প্রকাশের জন্য লিখিত প্রবন্ধে, গবেষণা সাহিত্যের বিস্তারিত উপস্থাপনের প্রয়োজন নেই। অন্য যে সব প্রকাশনায় এর সংক্ষিপ্তসার প্রকাশিত হয়েছে, সে সব প্রকাশনা, বা মূল গবেষণাগুলোর উল্লেখ করলেই চলে। যদি বর্তমান গবেষণাটি অন্য গবেষণার ধারাবাহিকতার অংশ হিসাবে পরিচালিত হয়ে থাকে, সে ক্ষেত্রে গুরুত্বপূর্ণ প্রাসঙ্গিক বিষয়গুলোর উল্লেখ করে বিস্তারিত তথ্যের জন্য, সেই গবেষণা প্রতিবেদনগুলোকে নির্দেশ করা যেতে পারে। গবেষণার ফলাফলের সাথে তাত্ত্বিক কাঠামোর সম্পর্কের আলোচনাটিও সংক্ষিপ্তভাবে করতে হবে।

উপাত্ত সংগ্রহের হাতিয়ার ও পদ্ধতির আলোচনাকে পরিহার করা যেতে পারে। তবে পাঠককে সাধারণভাবে কি পদ্ধতিতে উপাত্ত সংগ্রহ করা হয়েছিলো, সে বিষয়ে কয়েকটি বাক্যের মাধ্যমে অবহিত করা যেতে পারে। কি কি বিষয়ের উপর উপাত্ত সংগ্রহ করা হয়েছিলো, সেই বিষয়ভিত্তিক শিরোনামগুলো উল্লেখ করা যেতে পারে। যে গবেষণার উপর প্রবন্ধটি লেখা হয়েছে, সেটি যদি অন্য কোন গবেষণায় ব্যবহৃত পদ্ধতি অনুসরণ করে পরিচালনা করা হয়ে থাকে, তবে পদ্ধতিগত বিষয়ে কোন আলোচনা না করে, সেই গবেষণার স্মারক উল্লেখ করা যেতে পারে। যদি পদ্ধতির উপর কোন লিখিত বিবরণ প্রকাশিত হয়ে না থাকে, বা যদি উদ্ভাবনীমূলক পদ্ধতির ব্যবহার করা হয়ে থাকে, সে ক্ষেত্রে লেখক তার নতুন কৌশলের বিবরণ দিয়ে ভিন্নভাবে একটি প্রবন্ধ প্রকাশ করতে পারেন।

গবেষণা সাময়িকীতে স্থান সংকুলানের সীমাবদ্ধতার কারণে, সংক্ষিপ্ত প্রতিবেদনে কি কি বিষয়ে আলোচনা করা হবে, তার নির্বাচনে লেখককে খুব বেশী সতর্ক হতে হবে। তবে ধনাত্মক এবং ঋণাত্মক, দু'ধরণের ফলাফলকে উপস্থাপনের বিষয়টি নিশ্চিত করতে হবে। কিন্তু এ ক্ষেত্রে, ঋণাত্মক ফলাফলগুলো সারণি, বা লেখচিত্র ছাড়া সংক্ষিপ্তভাবে বর্ণনা করতে হবে। গবেষণাটি যদি বৃহৎ ও ব্যাপক হয়, সে ক্ষেত্রে কেবলমাত্র সবচেয়ে গুরুত্বপূর্ণ প্রসঙ্গের উপর প্রাপ্ত ফলাফলগুলো উপস্থাপন করতে হবে। যে সকল বিষয় গুরুত্বপূর্ণ এবং আলোচনার প্রয়োজন রয়েছে বলে মনে হয়, অথচ উপাত্ত কোন নির্দিষ্ট দিক নির্দেশনা দিতে পারে না, গবেষণা সাময়িকীতে প্রবন্ধ প্রকাশের জন্য সেগুলোকে বাদ দেয়া যেতে পারে। দীর্ঘ প্রতিবেদনের তুলনায়, গবেষণা সাময়িকীতে প্রকাশের জন্য লিখিত প্রবন্ধে, খুব কম সংখ্যক সারণি, লেখচিত্র, ছক, তালিকা, ইত্যাদি ব্যবহার করতে হবে। গবেষণা সাময়িকীর জন্য, গবেষণা প্রতিবেদনের শেষে উপস্থাপিত সারসংক্ষেপের পরিবর্তে প্রতিবেদনের শুরুতে, একশ' বা তার চেয়ে কম শব্দ দিয়ে একটি সংক্ষিপ্তসার সংযোজন করতে হবে।

সারাংশ

এই পাঠে, আমরা গবেষণা প্রতিবেদনের গুরুত্ব ও এর প্রয়োজনীয়তা নিয়ে আলোচনা করেছি। বিশেষ করে, একটি উত্তম গবেষণা প্রতিবেদনে কি কি থাকা উচিৎ এবং এর উপস্থাপনের রীতি কেমন হবে। উপাত্ত সংগ্রহ করে যদি প্রতিবেদন তৈরি না করা হয়, তাহলে পুরো গবেষণা প্রক্রিয়াটি অর্থহীন হয়ে পড়ে। গবেষণার শেষ পর্যায়ে এসে গবেষণা প্রতিবেদন প্রণয়ন গবেষকের জন্য একটি গুরুত্বপূর্ণ কর্মকান্ড, যেখানে গবেষক উপাত্তকে সংক্ষিপ্ত করে অনুকল্পের আলোকে বর্ণনা ও ব্যাখ্যা করেন। একটি উত্তম গবেষণা প্রতিবেদনে, সমস্যার বর্ণনা, গবেষণা পদ্ধতি, প্রাপ্ত ফলাফল, ফলাফলের অন্তর্নিহিত প্রভাব, গবেষণার সারসংক্ষেপ ও গ্রন্থ-স্মারক দিয়ে প্রতিবেদনের সমাপ্তি টানা হয়।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক ($\surd$) চিহ্ন দিন –

১। গবেষণা প্রতিবেদনের মূল উদ্দেশ্য হচ্ছে:
   ক. নিজের সাথে নিজে কথা বলা
   খ. পাঠকের সাথে যোগাযোগ স্থাপন করা
   গ. প্রতিবেদন ব্যবহার করে চাকরিতে প্রমোশন নেয়া
   ঘ. প্রতিবেদনের বক্তব্য অনুযায়ী সমাজ-নীতি তৈরি করা।

২। গবেষণার প্রথম পদক্ষেপটি হলো:
   ক. সমস্যার সূত্রবদ্ধকরণ
   খ. গবেষণা পরিচালনার অর্থ সংগ্রহ করা
   গ. উপাত্ত সংগ্রহের জন্য গবেষণা-সহকর্মী নিয়োগ দেয়া
   ঘ. সরকারের উপরের মহলের অনুমতি নেয়া।

৩। গবেষণায় নতুন অনুকল্পের জন্ম হয়, যখন:
   ক. নমুনার সংখ্যা বেশী থাকে
   খ. উপাত্তের জগতে পরিমাণগত উপাত্তের আধিক্য থাকে
   গ. সমস্যার বর্ণনাটি পর্যাপ্ততার দিকে এগুতে থাকে
   ঘ. গবেষকের স্বচ্ছ দৃষ্টিভঙ্গি থাকে।

সংক্ষিপ্ত প্রশ্ন

টীকা লিখুন:

ক. গবেষণা প্রতিবেদন
খ. প্রতিবেদন রচনাশৈলী

রচনামূলক প্রশ্ন

১। গবেষণা প্রতিবেদন কী? এর প্রয়োজনীয়তা উদাহরণসহ আলোচনা করুন।
২। একটি উত্তম প্রতিবেদনে কি কি থাকা উচিৎ? উদাহরণসহ আলোচনা করুন।