



উপাত্ত প্রক্রিয়াকরণ *Data Processing*

উপাত্ত সংগ্রহের পর, বিশাল উপাত্তরাশিমালাকে বিশ্লেষণের যোগ্য করে তুলতে হলে, সেগুলোকে বিধিবদ্ধ নিয়মের মাধ্যমে সংগঠিত করতে হয়। উপাত্তকে সংগঠিত করতে হলে, দু'টি গুরুত্বপূর্ণ কর্ম সম্পাদন করতে হয়। প্রথমটি হলো, উপাত্তের সম্পাদনা ও যাচাই এবং দ্বিতীয়টি হলো, উপাত্তের সংক্ষিপ্তকরণ। উপাত্তের সম্পাদনা ও যাচাই প্রক্রিয়ায়, পূরণকৃত প্রশ্নমালা, বা সাক্ষাৎকার অনুসূচীর মধ্যে সম্ভাব্য ভুল-ভ্রান্তিকে চিহ্নিত করার লক্ষ্যে উপাত্তের সম্পূর্ণতা, সঠিকতা, সমরূপতা, বোধগম্যতা, সঙ্গতিপূর্ণতা, ইত্যাদি পরীক্ষা করতে হয়। বিশাল উপাত্তরাশিকে সংক্ষিপ্তকরণের দু'টি প্রক্রিয়া রয়েছে। একটি হলো, সাংকেতীকরণ এবং অন্যটি হলো, সারণিকরণ। সাংকেতীকরণের মাধ্যমে প্রাপ্ত উত্তর-দফাকে অর্থপূর্ণভাবে শ্রেণীবদ্ধ করার প্রক্রিয়ায়, পারস্পরিক বিশিষ্টতা, সামগ্রিকতা ও পুঙ্খানুপুঙ্খতার উপাদানগুলো গুরুত্বপূর্ণ। সাংকেতীকরণের নিয়মাবলী অনুসরণ করে প্রতিটি চলকের জন্য সাংকেতীকরণ ছক নির্মাণের পর, সেগুলোকে একটি সংকেত-বইয়ে গ্রহিত করতে হয়, যা মূলতঃ সাংকেতীকরণ পরিকল্পনার একটি রূপরেখা। সারণিকরণের মাধ্যমে উপাত্ত সংক্ষিপ্তকরণ প্রক্রিয়া উপাত্তের বিশ্লেষণ ও ব্যাখ্যার পথ সুগম করে। এ ইউনিটে, এ সকল বিষয়কে বিস্তারিতভাবে আলোচনা করা হয়েছে। তবে, বিষয়বস্তুর মিল থাকার কারণে পাঠ-৩ এর কিছু কিছু ক্ষেত্রে “সামাজিক পরিসংখ্যান পরিচিতি” (সমাজতত্ত্ব-৪) গ্রন্থের প্রাসঙ্গিক অংশের সাথে সাদৃশ্য থাকবে।

এই ইউনিটে আমরা যে পাঠগুলো অধ্যয়ন করবো, সেগুলো হলো:

- ◆ পাঠ - ১ : সম্পাদনা ও যাচাই
- ◆ পাঠ - ২ : উপাত্ত সংক্ষিপ্তকরণ ১ – সাংকেতীকরণ
- ◆ পাঠ - ৩ : উপাত্ত সংক্ষিপ্তকরণ ২ – সারণিকরণ

সম্পাদনা ও যাচাই Editing and Verification

এই পাঠ শেষে যা জানা যাবে —

- উপাত্ত প্রক্রিয়াকরণ কী
- সম্পাদনা ও যাচাই

উপাত্ত প্রক্রিয়াকরণ কী (What is Data Processing)?

উপাত্ত সংগ্রহ সমাপ্তির পর, গবেষকের প্রধান কাজটি হলো, যে উদ্দেশ্য নিয়ে গবেষণাটি শুরু হয়েছিলো, সে উদ্দেশ্যগুলো অর্জনের লক্ষ্যে উপাত্ত বিশ্লেষণে মনোনিবেশ করা। মাঠ থেকে নিয়ে আসা অপরিশোধিত উপাত্তকে সরাসরি বিশ্লেষণ করা যায় না বলে, সেগুলোকে বিশ্লেষণের যোগ্য করে তোলার জন্য প্রক্রিয়াকরণের মাধ্যমে প্রস্তুত করে নিতে হয়। কিন্তু উপাত্ত সংগ্রহের পর, সামাজিক গবেষক, বিশেষ করে নব্য গবেষক, কিছুটা কিংকর্তব্যবিমূঢ় হয়ে পড়েন এই ভেবে যে, সংগৃহীত বিশাল উপাত্তরাশি নিয়ে তিনি কি করবেন। উপাত্তের গুণগত মান সর্বোচ্চ মানের হলেও গবেষক সব উপাত্তকে তার প্রতিবেদনে উপস্থাপন করতে পারেন না। বিশাল উপাত্তরাশিকে যাচাই-বাছাই করে সেগুলোকে সংক্ষিপ্তকরণের পর, বিশ্লেষণ ও ব্যাখ্যা করতে হবে, যাতে করে তাৎপর্যপূর্ণ কিছু উপসংহার উপস্থাপন করা যায়। উপাত্তের যাচাই-বাছাই থেকে বিশ্লেষণের পূর্ব পর্যন্ত যে সকল কর্মকাণ্ড সম্পাদনা করা হয়, সে সকল কর্মকাণ্ডের সম্মিলিত রূপকে উপাত্ত প্রক্রিয়াকরণ বলে।

উপাত্তের যাচাই-বাছাই থেকে বিশ্লেষণের পূর্ব পর্যন্ত যে সকল কর্মকাণ্ড সম্পাদনা করা হয়, সে সকল কর্মকাণ্ডের সম্মিলিত রূপকে উপাত্ত প্রক্রিয়াকরণ বলে।

উপাত্ত প্রক্রিয়াকরণের দু'টি প্রধান কর্মকাণ্ড রয়েছে। একটি হলো, উপাত্তের সম্পাদনা ও যাচাই (editing and verification) এবং অন্যটি হলো, উপাত্তের সংক্ষিপ্তকরণ (summarization)। উপাত্ত সংক্ষিপ্তকরণ দু'ভাবে করা যায় – সাংকেতীকরণ (coding) ও সারণিকরণের (tabulation) মাধ্যমে। এই কর্মকাণ্ডগুলো একে অপরের সাথে সম্পর্কযুক্ত হলেও সেগুলোর মধ্যে পার্থক্য রয়েছে। সম্পাদনা ও যাচাই হলো পূরণকৃত প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীগুলোর মধ্যে কোন ভুল-ভ্রান্তি রয়েছে কি না, তা সনাক্ত করে বের করা এবং সেগুলোকে দূরীভূত করা। বিশাল তথ্যরাশিকে সংক্ষিপ্তকরণের মাধ্যমে নিয়ন্ত্রণযোগ্য পরিমাণে নিয়ে আসার প্রক্রিয়াটি হলো সাংকেতীকরণ। কোন কোন ক্ষেত্রে, সাংকেতীকরণকে একটি সাধারণ কর্মকাণ্ড বলে মনে হলেও সাংকেতীকরণ প্রক্রিয়ায় প্রচুর পরিমাণে অভিজ্ঞতালব্ধ ও বিবেচনাপ্রসূত সিদ্ধান্ত গ্রহণের দক্ষতা প্রদর্শনের প্রয়োজন হয়। উপাত্ত বিশ্লেষণের পূর্বে, উপাত্তের মূল বৈশিষ্ট্যগুলোর স্বকীয়তা প্রদর্শনের জন্য তথ্যরাশিকে সারণিবদ্ধ করতে হয়। ক্ষুদ্র গবেষণার ক্ষেত্রে, সারণিকরণ প্রক্রিয়াটি হাতে হাতে সম্পন্ন করা যায়, তবে তাতে কারিগরী দক্ষতার পরিবর্তে প্রচুর পরিমাণে যত্ন, সঠিকতা ও ধৈর্যের প্রয়োজন হয়। যদি গবেষণাটি বৃহৎ হয়, তবে তা কম্পিউটারের মাধ্যমে যথাযথ পূর্ব পরিকল্পনা ও দক্ষতার সাথে সম্পাদনা করতে হয়। এই পাঠে আমরা সম্পাদনা ও যাচাই নিয়ে বিস্তারিতভাবে আলোচনা করবো এবং পাঠ ২ ও ৩-এ উপাত্ত সংক্ষিপ্তকরণের অন্য দু'টি উপায় (সাংকেতীকরণ ও সারণিকরণ) নিয়ে আলোচনা করবো।

সম্পাদনা ও যাচাই (Editing and Verification)

পূরণকৃত প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীর মধ্যে সম্ভাব্য ভুল-ভ্রান্তিকে চিহ্নিত করে সেগুলোকে যতটুকু সম্ভব সংশোধন করে উপাত্তের গুণগত মান বৃদ্ধি হলো সম্পাদনা ও যাচাইকরণের মূল উদ্দেশ্য। এটি গবেষণা প্রক্রিয়ায় সবচেয়ে বিরক্তিকর, নৈমিত্তিক, ধীরগতিসম্পন্ন, পুনরাবৃত্তি-প্রবণ এবং এক্ষেত্রে একটি কাজ। কিন্তু উত্তরদাতার কাছ থেকে, বা মাঠ থেকে ফেরৎ আসা পূরণকৃত প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীর মধ্যে যে কেউ এক বালক দৃষ্টি দিলেই অনুধাবন করতে পারবেন যে, সযত্ন সম্পাদনা ও

যাচাই-এর নিরঙ্কুশ প্রয়োজনীয়তাটি কত জরুরী। একজন সাক্ষাৎকার গ্রহণকারী যত দক্ষই হোক না কেন, তিনি কিছুতেই নিশ্চিত প্রতিশ্রুতি দিতে পারবেন না যে, সাক্ষাৎকার গ্রহণের সময় তিনি কোন ভুল করবেন না। এমন কি, একজন শ্রেষ্ঠ সাক্ষাৎকার গ্রহণকারীও ভুল করে থাকেন।

বিভিন্ন উৎস থেকে ভুল-ভ্রান্তির জন্য হতে পারে। একজন সাক্ষাৎকার গ্রহণকারী সাক্ষাৎকার গ্রহণের নির্দেশাবলীকে ভুল বুঝতে পারেন, প্রশ্ন জিজ্ঞাসা করতে ভুলে যেতে পারেন, উত্তর লিপিবদ্ধ করতে ভুলে যেতে পারেন, উত্তরকে অস্পষ্টভাবে লিপিবদ্ধ করতে পারেন, বা উত্তর লিপিবদ্ধ করতেও ভুল করতে পারেন। সাক্ষাৎকার গ্রহণকারী সম্পাদনা ও যাচাই-এর কাজটি শুরু করেন মাঠ পর্যায়ে সাক্ষাৎকার গ্রহণ শেষ করার সাথে সাথেই। কারণ, ভুল-ভ্রান্তি শোধরানো সবচেয়ে সহজ হয়, যদি তা সাক্ষাৎকার গ্রহণের পরপরই চিহ্নিত করা যায়। তবে, এর অর্থ এই নয় যে, মাঠ থেকে ফেরার পর সম্পাদনার আর প্রয়োজন নেই। সাক্ষাৎকার গ্রহণকারী মাঠ পর্যায়ে যত যত্নের সাথেই পূরণকৃত সাক্ষাৎকার অনুসূচী সম্পাদনা করুন না কেন, চূড়ান্ত সম্পাদনাটি করা প্রয়োজন গবেষকের কার্যালয়ে, তার নিজস্ব তত্ত্বাবধানে। পূরণকৃত ডাক প্রশ্নমালার ক্ষেত্রে, সম্পাদনা ও যাচাই-এর কাজটি মাঠে করার কোন সুযোগ নেই। সেটি গবেষকের কার্যালয়েই সম্পাদনা করতে হয়।

পূরণকৃত প্রশ্নমালা বা সাক্ষাৎকার অনুসূচী সম্পাদনা ও যাচাই-এর ক্ষেত্রে, একটি গুরুত্বপূর্ণ বিষয়কে বিবেচনায় রাখা প্রয়োজন। সেটি হলো, প্রারম্ভিক পর্যায়ে কি সবগুলো সাক্ষাৎকার অনুসূচী বা প্রশ্নমালাকে সম্পাদনা করা হবে, না কি এর থেকে কিছু সংখ্যক দ্বৈবভাবে নির্বাচন করে করা হবে? এ প্রসঙ্গে কোন বাধাধরা নিয়ম নেই, তবে সিদ্ধান্তটি নির্ভর করে বিভিন্ন ধরনের ভ্রান্তির ঝুঁকি কতটুকু রয়েছে, তার উপর। যদি ত্রুটি হওয়ার সম্ভাবনা কম থাকে, তবে সবগুলো প্রশ্নমালা বা সাক্ষাৎকার অনুসূচী সম্পাদনার মাধ্যমে সময় ও সম্পদ অপচয়ের কোন যৌক্তিকতা নেই। কিন্তু কিছু সংখ্যক প্রশ্নমালা বা সাক্ষাৎকার অনুসূচী সম্পাদনার পর, যদি দেখা যায় যে, ত্রুটি-বিচ্যুতির পরিমাণ তুলনামূলকভাবে অনেক বেশী রয়েছে, সে ক্ষেত্রে সবগুলো প্রশ্নমালা বা সাক্ষাৎকার অনুসূচী সম্পাদনা করাই শ্রেয়।

কি পরিমাণ প্রশ্নমালা বা সাক্ষাৎকার অনুসূচী সম্পাদনা ও যাচাই করা হবে, সে সিদ্ধান্তের সাথে আরেকটি প্রাসঙ্গিক বিষয়ে সিদ্ধান্ত গ্রহণের প্রয়োজন রয়েছে। সেটি হলো, একটি প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীর পুরোটিকেই কি সম্পাদনা করা হবে, না কি কোন একটি বিশেষ অংশ, বা প্রশ্নকে সম্পাদনা করা হবে? যদি একটি বিশেষ অংশ, বা প্রশ্নকে পরীক্ষা করে দেখা হয়, সে ক্ষেত্রে সুবিধাটি হলো যে, একজন সম্পাদনাকারী একটি বা দু'টি প্রশ্নের উপর দৃষ্টি নিবদ্ধ করে কার্যকরভাবে সম্পাদনার কাজটি করতে পারেন। অন্যদিকে, একটি পুরো প্রশ্নমালা বা সাক্ষাৎকার অনুসূচী পরীক্ষা করে দেখার মাধ্যমে বিভিন্ন প্রশ্ন ও উত্তরের মধ্যে সম্পূর্ণতা, সঠিকতা ও সঙ্গতিপূর্ণতার সম্পর্কের একটি সামগ্রিক চিত্র পেতে পারেন। এতে করে, একজন সাক্ষাৎকার গ্রহণকারীর দক্ষতাকেও যাচাই করে দেখা যায়।

কার্যকর নিরীক্ষা নিশ্চিত করার লক্ষ্যে, সম্পাদনাকারীর জন্য নির্দেশাবলী সম্বলিত একটি নির্দেশিকা থাকতে হবে, যা প্রতিটি সম্পাদনাকারী সম্পাদনার সময় অনুসরণ করবেন। সম্পাদনা শুরুর পূর্বে সম্পাদনাকারীদের উপযুক্ত প্রশিক্ষণ দিয়ে নিতে হবে। গবেষণার উদ্দেশ্য, বিভিন্ন চলক ও সূচকের পরিমাপ এবং সেগুলোর গুরুত্ব সম্পর্কে সম্পাদনাকারীর যথাযথ জ্ঞান থাকতে হবে। একটি পূরণকৃত প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীকে সাংকেতিকরণ, সারণিকরণ ও বিশ্লেষণের জন্য যোগ্য বলে বিবেচিত করার পূর্বে, গবেষককে কতগুলো গুরুত্বপূর্ণ বিষয় নিশ্চিত করে নিতে হবে। সে বিষয়গুলো হলো পূরণকৃত প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীর সম্পূর্ণতা, সঠিকতা, একরূপতা, স্পষ্টভাবে পাঠযোগ্যতা, বোধগম্যতা ও সঙ্গতিপূর্ণতা।

সম্পাদনা শুরুর পূর্বে
সম্পাদনাকারীদের উপযুক্ত
প্রশিক্ষণ দিয়ে নিতে হবে।
গবেষণার উদ্দেশ্য, বিভিন্ন চলক
ও সূচকের পরিমাপ এবং
সেগুলোর গুরুত্ব সম্পর্কে
সম্পাদনাকারীর যথাযথ জ্ঞান
থাকতে হবে।

সম্পূর্ণতা (Completeness): সম্পূর্ণতা পরীক্ষার প্রধান বিষয়টি হলো যে, প্রতিটি প্রশ্নের উত্তর লিপিবদ্ধ করা হয়েছে কি না, তা পরীক্ষা করে দেখা। সাক্ষাৎকার গ্রহণকারী যদি কোন প্রশ্ন জিজ্ঞাসা করতে ভুলে যান, বা কোন প্রশ্নের উত্তর লিপিবদ্ধ করতে ভুলে যান, সে ক্ষেত্রে কতগুলো সম্ভাব্য উপায়ে তা সংশোধন করা যেতে পারে। ভুলটি যদি মাঠ পর্যায়ে ধরা পড়ে, তবে উত্তরদাতার কাছে পুনরায়

গিয়ে তা জেনে আসা যেতে পারে। কিন্তু যদি ভুলটি মাঠ থেকে ফেরার পর ধরা পড়ে, সে ক্ষেত্রে অন্যান্য প্রাসঙ্গিক প্রশ্নের উত্তর থেকে সম্ভাব্য উত্তরটি অনুমান করে, বা স্মরণশক্তি ব্যবহার করে সে শূন্যতা পূরণ করা যেতে পারে। সংশ্লিষ্ট প্রশ্নের উত্তরটি যদি খুব গুরুত্বপূর্ণ হয়, সে ক্ষেত্রে উত্তরদাতার কাছে ফিরে গিয়ে, বা যদি পরিস্থিতি সমর্থন করে, তাহলে ডাকযোগে উত্তরদাতার কাছে পুনরায় জেনে নেবার মাধ্যমে শূন্যতাটি পূরণের উদ্যোগ গ্রহণ করা উচিত।

যদি দেখা যায় যে, কোন বিশেষ সাক্ষাৎকার অনুসূচীতে অনেক বেশী পরিমাণে প্রশ্নের উত্তর নেই, সে ক্ষেত্রে ধরে নেয়া যেতে পারে যে, সাক্ষাৎকার গ্রহণকারীর অদক্ষতার কারণে উত্তরদাতা উত্তরদানে অনিচ্ছুক ছিলেন।

কিন্তু প্রশ্নের প্রেক্ষাপট, বা অন্যান্য প্রাসঙ্গিক প্রশ্নের উত্তর থেকে যদি এটি নির্ধারণ করা না যায় যে, উত্তরদাতা কি উত্তর দিতে অস্বীকৃতি জানিয়েছিলেন, না কি সাক্ষাৎকার গ্রহণকারী প্রশ্ন জিজ্ঞাসা করতে ভুলে গিয়েছিলেন, না কি উত্তর লিপিবদ্ধ করতে ভুলে গিয়েছিলেন, না কি প্রশ্নটি উত্তরদাতার জন্য প্রযোজ্য ছিল না, সে ক্ষেত্রে সমস্যাটি জটিল আকার ধারণ করে। অন্যান্য প্রাসঙ্গিক উপাত্ত থেকে প্রশ্নের প্রযোজ্য উত্তরটি হয়তো অনুমান করা যেতে পারে, কিন্তু উত্তর বাদ পড়ার অন্যান্য কারণগুলোর ব্যাখ্যা প্রদান করা যাবে না। একজন প্রশিক্ষিত ও সতর্ক সাক্ষাৎকার গ্রহণকারী সাধারণতঃ, এ ধরনের ভুল কম করে থাকেন। কিন্তু যদি দেখা যায় যে, কোন বিশেষ সাক্ষাৎকার অনুসূচীতে অনেক বেশী পরিমাণে প্রশ্নের উত্তর নেই, সে ক্ষেত্রে ধরে নেয়া যেতে পারে যে, সাক্ষাৎকার গ্রহণকারীর অদক্ষতার কারণে উত্তরদাতা উত্তরদানে অনিচ্ছুক ছিলেন। এ ধরনের অসম্পূর্ণ সাক্ষাৎকার অনুসূচীকে বিশ্লেষণ থেকে বাদ দিয়ে দিতে হয়। প্রান্ত-খোলা প্রশ্নের জন্য সম্পূর্ণতার পরীক্ষা করা কার্যতঃ অসম্ভব। সে ক্ষেত্রে, উত্তরটি স্পষ্টভাবে লিপিবদ্ধ করা হয়েছে কি না এবং তা মোটামুটিভাবে অর্থ বহন করে কি না, এটি নির্ধারণ করা ছাড়া আর বেশী কিছু করার থাকে না।

সাক্ষাৎকার গ্রহণকারীর অসতর্কতা থেকে, বা উত্তরদাতা কর্তৃক ভ্রান্ত উত্তর দেবার সচেতন প্রয়াস থেকে, বা সৌজন্যমূলক উত্তর দেবার প্রবণতা থেকেও ভ্রান্তির জন্ম হতে পারে।

সঠিকতা (Accuracy): পূরণকৃত সাক্ষাৎকার অনুসূচী বা প্রশ্নমালা সম্পাদনা ও যাচাই-এর ক্ষেত্রে শুধু প্রতিটি প্রশ্নের উত্তর রয়েছে কি না, সে বিষয়টি দেখলেই চলবে না, উত্তরগুলো সঠিক কি না, বা উত্তরগুলোর মধ্যে কোন অসঙ্গতি রয়েছে কি না, তাও পরীক্ষা করে দেখতে হবে। যেমন, যদি একটি প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীতে দেখা যায় যে, একজন সরকারী কর্মকর্তার বয়স ৩২ বছর এবং পদমর্যাদার দিক থেকে যুগ্ম সচিবের পদটিতে টিক চিহ্ন দেয়া রয়েছে, তবে ধরে নিতে হবে যে, দু'টি তথ্যের মধ্যে কোন একটি ভুল। হয় উত্তরদাতার বয়স ভুল লেখা হয়েছে, নতুবা তার পদমর্যাদা ভুল লেখা হয়েছে। এ ধরনের ভুলগুলোকে দূর করার জন্য সর্বাঙ্গিক প্রচেষ্টা থাকতে হবে। অনেক সময় সঠিক উত্তরকেও আপাতঃদৃষ্টিতে ভ্রান্তিপূর্ণ মনে হতে পারে। সে ক্ষেত্রে, সতর্কতার সাথে লক্ষ্য রাখতে হবে যে, সংশোধনগুলো যেন শুধুমাত্র ত্রুটিপূর্ণ উত্তরগুলোর ক্ষেত্রেই করা হয়েছে। সাক্ষাৎকার গ্রহণকারীর অসতর্কতা থেকে, বা উত্তরদাতা কর্তৃক ভ্রান্ত উত্তর দেবার সচেতন প্রয়াস থেকে, বা সৌজন্যমূলক উত্তর দেবার প্রবণতা থেকেও ভ্রান্তির জন্ম হতে পারে।

সাক্ষাৎকার গ্রহণের সময়, সাক্ষাৎকার গ্রহণকারী সাধারণতঃ একটি মানসিক চাপের মধ্যে থাকেন। কারণ, তাকে একই সাথে প্রশ্ন জিজ্ঞাসা করতে হয়, উত্তর লিপিবদ্ধ করতে হয়, এবং প্রয়োজনবোধে উত্তরদাতাকে, প্রোৎসাহিতও করতে হয়। এতে করে, তিনি একটি ভুল উত্তর-দফাকে চিহ্নিত করে ফেলতে পারেন। অনেক সময়, উত্তর লিপিবদ্ধ করতে গিয়ে সাক্ষাৎকার গ্রহণকারীকে ছোট-খাট নির্ণয়মূলক হিসাব করতে হয় বলে সেই সব হিসাব থেকেও ভ্রান্তির জন্ম নিতে পারে। যেমন, সাক্ষাৎকার গ্রহণকারীকে যদি বিভিন্ন আয়ের উৎস থেকে প্রাপ্ত আয়ের যোগফল নির্ণয় করতে হয়, সে ক্ষেত্রে তার অসতর্কতার জন্য যোগফলটি ভুল হতে পারে। এ ছাড়াও, কোন প্রশ্নের উত্তরের ক্ষেত্রে, এক একক থেকে অন্য এককে (যেমন, গজ থেকে মিটার, বা বিঘা থেকে হেক্টরে) রূপান্তরের সময় ভুল হতে পারে। নির্ণয়মূলক হিসাব করতে হয়, এমন উত্তর লিপিবদ্ধ করার সময়, যোগফল বা একক পরিবর্তনের কাজটি সাক্ষাৎকার গ্রহণের সময় না করে মাঠ থেকে ফিরে এসে করা উচিত। প্রান্ত-খোলা প্রশ্নের ক্ষেত্রে, এ ধরনের ভ্রান্তি বেশী হয়ে থাকে। সবছ সম্পাদনা ও নিরীক্ষার মাধ্যমে এই ভ্রান্তিগুলোকে দূর করে সঠিকতাকে নিশ্চিত করা যায়।

সঙ্গতিপূর্ণতা (Consistency): একটি প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীর মধ্যে প্রচুর পরিমাণে অসঙ্গতির উপস্থিতি শুধুমাত্র সাংকেতিকরণকারীর জন্যই সমস্যা তৈরি করে না, সেগুলো উপাত্ত সংগ্রহ

বা লিপিবদ্ধকরণের সমস্যাকেও ইঙ্গিত করে। যেমন, উত্তরদাতার যদি কোন আয় এবং সম্পদ না থাকে, এবং সেই উত্তরদাতা যদি গ্রামে বসবাস করেন এবং তিনি যদি জানান যে, তার আয়ের অধিকাংশ ব্যয় হয় বৈদ্যুতিক বিল পরিশোধ করার জন্য, তাহলে বুঝতে হবে যে, উত্তরটি সঙ্গতিপূর্ণ নয়। এ ধরনের পরিস্থিতিতে, তাৎক্ষণিকভাবে এই অসঙ্গতির উৎস অনুসন্ধান করে তা দূরীভূত করতে হবে। যদি তা করা না যায়, তবে সেই নমুনাটি বাতিল করে নতুন নমুনার সাক্ষাৎকার নিতে হবে। আর যদি তথ্যটি তেমন গুরুত্বপূর্ণ না হয়, সে ক্ষেত্রে তা বাতিল করে দেয়াই শ্রেয়। নতুন নমুনার সাক্ষাৎকার নেবার কোন প্রয়োজন নেই।

একরূপতা (Uniformity): সাক্ষাৎকার গ্রহণের সময়, সাক্ষাৎকার গ্রহণকারী প্রশ্ন এবং নির্দেশাবলীকে একইভাবে ব্যাখ্যা ও মূল্যায়ন করেছেন কি না, সম্পাদনকারীকে তা সতর্কতার সাথে পরীক্ষা করে দেখতে হবে। ধরা যাক, পারিবারিক আয়-ব্যয় সংক্রান্ত একটি গবেষণায়, একটি প্রশ্ন রয়েছে যে, “প্রতি মাসে পত্র-পত্রিকা ও সাময়িকীর জন্য আপনি কত টাকা ব্যয় করেন”? এ প্রশ্নটি কেবলমাত্র সে সকল পরিবারের ক্ষেত্রেই প্রযোজ্য হবে, যে পরিবারগুলো বাড়িতে নিয়মিত দৈনিক পত্র-পত্রিকা ও সাময়িকী রাখেন। সম্পাদনার পর্যায়ে দেখা গেল যে, কোন কোন সাক্ষাৎকার গ্রহণকারী এমন উত্তর-দফায় টিক চিহ্ন দিয়েছেন, যা সেই উত্তরদাতার ক্ষেত্রে প্রযোজ্য নয়। কারণ, সে পরিবারে কোন পত্র-পত্রিকাই রাখা হয় না। আরও একটি উদাহরণ দেয়া যেতে পারে। যেমন, পারিবারিক আয় সংক্রান্ত একটি প্রশ্নের নির্দেশাবলীতে বলা হয়েছে যে, “সকল উৎস থেকে অর্জিত আয়ের তথ্য লিপিবদ্ধ করতে হবে”। কিন্তু দেখা গেল যে, কোন কোন সাক্ষাৎকার গ্রহণকারী পরিবার প্রধানের আয়ের বিভিন্ন উৎস থেকে অর্জিত অর্থের পরিমাণ উল্লেখ করেছেন, কিন্তু স্ত্রীর অনানুষ্ঠানিক খাত থেকে আয়ের পরিমাণ লিপিবদ্ধ করেন নি। এ ধরনের ভ্রান্তি, সম্পাদনার প্রারম্ভিক পর্যায়ে চিহ্নিত করে দূর করা সহজ হয়। অতএব, সম্পাদনার কাজটি যত বেশী প্রারম্ভিক পর্যায়ে শুরু করা যায়, ত্রুটিগুলো সংশোধনের সুযোগ থাকে তত বেশী।

স্পষ্টভাবে পাঠযোগ্যতা (Legibility): সাক্ষাৎকার গ্রহণকারী, পর্যবেক্ষক, বা প্রশ্নমালা পূরণকারী যে সব প্রতীক ব্যবহার করেছেন, বা তাদের হাতের লেখা যদি পড়া না যায়, তবে সাংকেতিকরণ অসম্ভব হয়ে পড়ে। সে জন্য, প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীগুলো যখন মাঠ থেকে ফেরৎ আসে, তখন সেগুলোকে ভালোভাবে পরীক্ষা করে দেখা দরকার যে, হাতের লেখা স্পষ্টভাবে পড়া যায় কি না। প্রয়োজনবোধে, তাৎক্ষণিকভাবে পুনরায় স্পষ্টভাবে লিখিয়ে নিতে হবে। কিন্তু এতে সাংকেতিকরণকারীর জন্য একটি বাড়তি ঝামেলা তৈরি হয়। কারণ, সাক্ষাৎকার গ্রহণকারীকে খুঁজে পাওয়া অনেক সময় সমস্যা হয়ে দাঁড়ায় এবং সাংকেতিকরণ প্রক্রিয়া শুরু হয়ে যাবার পর, তা প্রায় অসম্ভব হয়ে পড়ে।

বোধগম্যতা (Comprehensibility): উত্তরদাতা কর্তৃক প্রদত্ত উত্তর সাক্ষাৎকার গ্রহণকারী, বা পর্যবেক্ষককারীর কাছে সম্পূর্ণভাবে বোধগম্য হতে পারে, কিন্তু সাংকেতিকরণকারী, বা অন্য ব্যক্তির কাছে তা নাও হতে পারে। কারণ, যে পরিস্থিতিতে উত্তরদাতা একটি উত্তর দেন বা আচরণ করেন, সাক্ষাৎকার গ্রহণকারী, বা পর্যবেক্ষকের কাছে সেই পরিস্থিতি খুবই পরিচিত, যা সাংকেতিকরণকারীর কাছে নয়। ফলে, সেই উত্তরের মাধ্যমে উত্তরদাতা প্রকৃতপক্ষে কি বোঝাতে চেয়েছেন, তা সাংকেতিকরণকারী অনুধাবন করতে পারেন না। সে জন্য, নিয়মতান্ত্রিকভাবে প্রশ্ন জিজ্ঞাসা করা, বা পর্যবেক্ষণ করা উপাত্তের গুণগত মান বৃদ্ধির জন্য অত্যন্ত জরুরী। উপাত্তের মধ্যে যে কোন ধরনের অস্পষ্টতা বা বিভ্রান্তি যেন কোনভাবে ঢুকে পড়তে না পারে, সে বিষয়ে সাক্ষাৎকার গ্রহণকারীকে সার্বক্ষণিকভাবে সতর্ক থাকতে হবে।

সারাংশ

উপাত্ত সংগ্রহের পর, গবেষক প্রথম যে কাজটি করেন তা হলো, যে উপাত্ত সংগৃহীত হয়েছে, তার যথার্থতা নিরূপণ করা। উপাত্তের যথার্থতা নিরূপণে গবেষক দু’টি প্রধান কাজ করে থাকেন — সম্পাদনা ও যাচাই। সম্পাদনা ও যাচাই পর্ব শেষ হওয়ার পর, প্রাপ্ত উপাত্তকে সাংকেতিকরণ ও সারণীকরণের মাধ্যমে বর্ণনা ও বিশ্লেষণের পর্যায়ে নেয়া হয়। এই প্রক্রিয়াটি সাধারণতঃ পরিমাণগত উপাত্তের ক্ষেত্রে বেশী অনুসরণ করা হয়। গুণগত উপাত্তের ক্ষেত্রে, একই যুক্তি ব্যবহার করা হয়, কিন্তু

এস এস এইচ এল

সারণিকরণ ততটা প্রধান্য পায় না। উপাত্ত কিভাবে প্রক্রিয়াজাত করা হবে, তা মূলতঃ গবেষণার ধরনের উপর নির্ভর করে। এই পাঠে, পরিমাণগত উপাত্তের কথা বিবেচনায় রেখে সম্পাদনা ও যাচাই-এর উপর আলোচনা করা হয়েছে।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক (✓) চিহ্ন দিন –

১। উত্তরের ক্ষেত্রে ভ্রান্তির জন্ম হতে পারে:

- ক. ভ্রান্ত উত্তর দেবার প্রয়াস থেকে
- খ. সৌজন্যমূলক উত্তর দেবার প্রয়াস থেকে
- গ. সাক্ষাৎকার গ্রহণকারীর অসতর্কতা থেকে
- ঘ. উপরের সব।

২। প্রাপ্ত উত্তরের ত্রুটিগুলো সংশোধনের সুযোগ বেশী থাকে:

- ক. সাক্ষাৎকার গ্রহণের শুরুতে
- খ. সাক্ষাৎকার গ্রহণের শেষে
- গ. প্রশ্নমালা সম্পাদনার শুরুতে
- ঘ. উপরের সব।

৩। একটি পূরণকৃত সাক্ষাৎকার অনুসূচী তখনই বিশ্লেষণের জন্য যোগ্য বলে বিবেচিত হবে, যখন এর _____ নিশ্চিত বলে স্বীকৃতি পাবে।

- ক. সঠিকভাবে পাঠযোগ্যতা
- খ. একরূপতা
- গ. সঙ্গতিপূর্ণতা
- ঘ. উপরের সব

সংক্ষিপ্ত প্রশ্ন

টীকা লিখুন:

- ক. পূরণকৃত সাক্ষাৎকার অনুসূচীতে সম্পূর্ণতা
- খ. প্রশ্নমালার বোধগম্যতা

রচনামূলক প্রশ্ন

- ১। উপাত্ত প্রক্রিয়াকরণের গুরুত্ব উদাহরণসহ আলোচনা করুন।
- ২। উপাত্ত প্রক্রিয়াকরণের ক্ষেত্রে, সম্পাদনা ও যাচাই -এর ভূমিকা আলোচনা করুন।

উপাত্ত সংক্ষিপ্তকরণ ১ – সাংকেতিকরণ Data Reduction 1 – Coding

এই পাঠ শেষে যা জানা যাবে —

- সাংকেতিকরণ কী
- সাংকেতিকরণের নিয়মাবলী
- অবরোহীমূলক ও আরোহীমূলক সাংকেতিকরণ
- সংকেত-বই নির্মাণ
- সাংকেতিকরণের সাধারণ নির্দেশনা
- উপাত্ত অন্তর্ভুক্তির কৌশল
- সাংকেতিকরণের নির্ভরশীলতা
- উপাত্ত বিশুদ্ধকরণ

সাংকেতিকরণ কী (What is Coding)?

আজকালকার দিনে এটি ধরে নেয়া হয় যে, বিশ্লেষণের জন্য সংগৃহীত উপাত্ত সাংকেতিকরণ, সংরক্ষণ, পরিশুদ্ধকরণ ও বিশ্লেষণের কাজটি কম্পিউটারের মত যন্ত্র ব্যবহারের মাধ্যমে সম্পন্ন করা হবে। যেহেতু কম্পিউটার বর্ণনামূলক বক্তব্যকে বুঝতে পারে না, সেহেতু কম্পিউটারের ভাষার সাথে সঙ্গতিপূর্ণ করে তুলতে হলে, উপাত্তরাশিকে সংখ্যায় রূপান্তর করতে হবে। একটি চলকের পর্যবেক্ষণকৃত বৈশিষ্ট্যের প্রতি বরাদ্দকৃত সংখ্যাসূচক প্রতীককে 'সংকেত' (code) বলে এবং বর্ণনামূলক বক্তব্যকে সংখ্যাসূচক প্রতীক বা সংকেতে রূপান্তরের প্রক্রিয়াটিকে 'সাংকেতিকরণ' (coding) বলা হয়। অর্থাৎ, সাংকেতিকরণ হলো উপাত্ত প্রক্রিয়াকরণের সেই কলা কৌশল, যা কোন পর্যবেক্ষণের জন্য প্রাপ্ত বর্ণনামূলক উত্তরগুলোকে শ্রেণীবদ্ধ করে সুনির্দিষ্ট উত্তর-দফায় পরিণত করে এবং সেই উত্তর-দফাগুলোকে সংখ্যাসূচক বা পরিমাণগত সংকেতে রূপান্তরিত করে। যেমন, লিঙ্গ চলকটির জন্য বরাদ্দকৃত উত্তর-দফা রয়েছে 'নারী' এবং 'পুরুষ'। এই 'নারী' এবং 'পুরুষ' -এর শব্দগত প্রতীক দু'টিকে সংখ্যাসূচক প্রতীকে রূপান্তরিত করতে হলে, এর বিপরীতে যথাক্রমে, '১' এবং '২' সংখ্যা বরাদ্দ করতে হবে। কম্পিউটার লিঙ্গ চলকের দু'টি বৈশিষ্ট্যকে চিহ্নিত করবে '১' এবং '২' দিয়ে, 'নারী' ও 'পুরুষ' দিয়ে নয়।

সাংকেতিকরণ হলো উপাত্ত প্রক্রিয়াকরণের সেই কলা কৌশল, যা কোন পর্যবেক্ষণের জন্য প্রাপ্ত বর্ণনামূলক উত্তরগুলোকে শ্রেণীবদ্ধ করে সুনির্দিষ্ট উত্তর-দফায় পরিণত করে এবং সেই উত্তর-দফাগুলোকে সংখ্যাসূচক বা পরিমাণগত সংকেতে রূপান্তরিত করে।

চলকের বৈশিষ্ট্যের বিপরীতে সংখ্যাসূচক সংকেত বরাদ্দের ক্ষেত্রে, যে বিষয়টি গুরুত্বের সাথে বিবেচনা করতে হবে, সেটি হলো, যে কোন সংকেতকে সমরূপ পরিষ্টিত্বিত্তে, ঘটনা বা বিশ্লেষণের একক অনুযায়ী সামঞ্জস্যপূর্ণ হতে হবে। যেমন, যখন আমরা '১' দিয়ে 'নারী' ও '২' দিয়ে 'পুরুষকে' সাংকেতিকরণ করবো, তখন লিঙ্গ চলকের সাথে সম্পর্কিত সকল নারীকেই সংকেত '১' -এর মাধ্যমে চিহ্নিত করতে হবে। শুধু ব্যক্তির ক্ষেত্রেই নয়, গোষ্ঠীর বৈশিষ্ট্যকেও সংকেত বরাদ্দ করা যেতে পারে। ধরা যাক, একজন গবেষক কয়েকশ' ব্যক্তির পেশার উপর তথ্য সংগ্রহ করে বিভিন্ন পেশার ধরণগুলো লিপিবদ্ধ করেছেন। যেমন, কৃষক, জেলে, কামার, কুমার, নাপিত, ছুতার, মুচি, সেবিকা, দিন-মজুর, উকিল, চিকিৎসক, প্রকৌশলী, শিক্ষক, নির্বাহী কর্মকর্তা, সাংবাদিক, পিয়ন, পুলিশ, সামরিক কর্মকর্তা, বিদ্যুৎ বিভাগের কর্মী, ইত্যাদি। পেশার এই বিভিন্ন ধরণগুলো দিয়ে সরাসরি উপাত্ত বিশ্লেষণ করা যায় না। কারণ, পেশার ধরণগুলোর মধ্যে অনেক বেশী বৈচিত্র রয়েছে। সে জন্য, এগুলোকে সংক্ষিপ্ত করে কিছু নির্দিষ্ট শ্রেণীতে বিভক্ত করতে হয়। গবেষকদের কাছে গ্রহণযোগ্য একটি শ্রেণীবদ্ধকরণ প্রক্রিয়ার মাধ্যমে উপরোক্ত পেশার ধরণগুলোকে সংক্ষিপ্ত করে বিশ্লেষণযোগ্য করে তোলা যেতে পারে। যেমন,

পেশাজীবী: উকিল, প্রকৌশলী, শিক্ষক, চিকিৎসক, সাংবাদিক, সামরিক কর্মকর্তা।

সেবা প্রদানকারী: সেবিকা, পুলিশ, বিদ্যুৎ বিভাগের কর্মী।

দক্ষ শ্রমিক: কামার, কুমার, নাপিত, ছুতার, মুচি।

অদক্ষ শ্রমিক: কৃষক, দিন-মজুর, জেলে।

সাংকেতীকরণের নিয়মাবলী (Rules of Coding)

যেহেতু সাংকেতীকরণ হলো উত্তর-দফাকে অর্থবোধক শ্রেণীতে শ্রেণীবদ্ধ করার প্রক্রিয়া, সেহেতু সাংকেতীকরণের প্রাথমিক নিয়মটি হলো যে, যে সব সংখ্যাগুলো বরাদ্দ করা হবে, সেগুলোর যেন স্বজ্ঞাগত (intuitive) অর্থ থাকে। অর্থাৎ, একটি চলকের উচ্চমানকে উচ্চ সংকেত বরাদ্দ করতে হবে। যে ব্যক্তির বয়স অন্যদের তুলনায় বেশি, তার বয়সের ক্ষেত্রে উচ্চতর সংকেত প্রদান করতে হবে। যেমন, একজনের বয়স ২৮ বছর হলে, তার সংখ্যাসূচক সংকেত হবে ২৮ এবং ৩৫ বছর বয়স্ক ব্যক্তির সংকেত হবে ৩৫। এমনকি যদি বয়সের শ্রেণীগুলো ক্রমমূলকভাবে শ্রেণীবদ্ধ করা হয়, সে ক্ষেত্রে বেশী বয়স্ক ব্যক্তির উচ্চ বয়সসম্পন্ন দলের সাথে সম্পর্কিত হবে। ব্যাপ্তিমূলক চলকের ক্ষেত্রে, এই নিয়মটি খুব সহজে প্রয়োগ ও প্রদর্শন করা যায়। কিন্তু স্বজ্ঞাগত কারণে কিছু চলকের জন্য, বিশেষ করে নামসূচক মাত্রায় পরিমাপকৃত চলকগুলোর ক্ষেত্রে, সংখ্যাসূচক সংকেত বরাদ্দের কোন স্বজ্ঞাগত সন্তোষজনক যৌক্তিক ভিত্তি নেই। যেমন, লিঙ্গের ক্ষেত্রে পুরুষকে '২' এবং নারীকে '১' সংকেত প্রদানের মাধ্যমে পুরুষের লিঙ্গগত মর্যাদা দ্বিগুণ হয়ে যায় না। উপরন্তু, তাদেরকে যদি যথাক্রমে ৩ এবং ৪, অথবা ৪ এবং ৩ সংখ্যা বরাদ্দ করা হয়, তাতেও কোন তফাৎ হবে না। যখন স্বজ্ঞাগত অর্থের শর্তটি পূরণ হয়ে যায়, তখন পারস্পরিক বিশিষ্টতা (mutual exclusivity), সামগ্রিকতা (exhaustiveness) এবং পূঙ্খানুপূঙ্খতার (detail) উপাদানগুলো সাংকেতীকরণের সিদ্ধান্তের মধ্যে বিবেচনা করতে হয়।

নামসূচক মাত্রায় পরিমাপকৃত চলকগুলোর ক্ষেত্রে, সংখ্যাসূচক সংকেত বরাদ্দের কোন স্বজ্ঞাগত সন্তোষজনক যৌক্তিক ভিত্তি নেই।

পারস্পরিক বিশিষ্টতা (Mutual Exclusivity): নির্মিত সাংকেতীকরণ ছকে অন্তর্ভুক্ত সংকেত দফাগুলোকে অবশ্যই পারস্পরিকভাবে বিশিষ্ট এবং সামগ্রিক হতে হবে। পারস্পরিক বিশিষ্টতার নিয়ম অনুযায়ী প্রতিটি চলকের সংকেত দফাগুলোকে এমনভাবে নির্মাণ করতে হবে, যাতে করে প্রতিটি ঘটনা, বা বিশ্লেষণের একক কেবলমাত্র সংশ্লিষ্ট চলকের একটি দফায়ই অন্তর্ভুক্ত হয়। যেমন, ব্যক্তির ধর্মাবলম্বীতার ক্ষেত্রে, মুসলমান ও হিন্দু ধর্ম দু'টিকে দু'টি আলাদা দফায় চিহ্নিত করতে হবে। গবেষককে আরও নিশ্চিত হতে হবে যে, তার বাছাইকৃত উত্তর-দফাগুলো যেন এত বিস্তৃত না হয়, যাতে করে উত্তর-দফার মধ্যে গুরুত্বপূর্ণ পার্থক্যগুলো অস্পষ্ট হয়ে যায়।

সামগ্রিকতা (Exhaustiveness): সামগ্রিকতার অর্থ হলো, প্রত্যাশিত উত্তর-দফার সবগুলো দফাকেই আলাদা আলাদাভাবে চিহ্নিত করতে হবে, কোনটিকে বাদ দেয়া যাবে না। যেমন, বৈবাহিক মর্যাদা শ্রেণীবদ্ধকরণের ক্ষেত্রে, আমরা যদি 'অবিবাহিত', 'বিবাহিত', 'তালাকপ্রাপ্ত/তালাকপ্রাপ্তা' এবং 'বিধবা/বিপত্নীক' এই চারটি দফার মধ্যে সীমাবদ্ধ থাকি, তাহলে এই শ্রেণীবদ্ধকরণটি সামগ্রিকতার বৈশিষ্ট্য হারিয়ে ফেলে। কারণ, 'স্বামী পরিত্যক্তা' বা 'বিচ্ছিন্নভাবে বসবাসকারী' ব্যক্তিকে এর মধ্যে অন্তর্ভুক্ত করা হয় না। অতএব, এ ধরনের আরও দু'টি উত্তর-দফা সাংকেতীকরণ ছকের মধ্যে অন্তর্ভুক্ত করা উচিত। যদি একটি চলকের অধীনে প্রত্যাশিত উত্তর-দফাগুলো সামগ্রিক না হয়, সে ক্ষেত্রে 'অন্যান্য' উত্তর-দফার মধ্যে সবচেয়ে বেশি ঘটনা অন্তর্ভুক্ত হয়ে পড়বে। পাশাপাশি, 'উত্তরদানে অস্বীকৃতি', 'উত্তর না দেয়া' এবং 'প্রযোজ্য নয়' সংকেত দফাকেও উক্ত চলকের সংকেত দফার মধ্যে অন্তর্ভুক্ত করতে হবে।

পূঙ্খানুপূঙ্খতা (Detail): একটি সাংকেতীকরণ ছকে অন্তর্ভুক্ত উত্তর-দফাগুলোর পূঙ্খানুপূঙ্খতা গবেষণা সমস্যার উপর নির্ভর করলেও এর কিছু সাধারণ নিয়ম রয়েছে। প্রথমতঃ, উত্তর-দফা চিহ্নিতকরণ সম্পর্কে কোন সন্দেহ থাকলে, একটি নতুন উত্তর-দফা যুক্ত করা যেতে পারে। দ্বিতীয়তঃ, পর্যালোচনার

পর, যদি কিছু কিছু উত্তর-দফার মধ্যে খুব বেশী পার্থক্য নিরূপণ করা না যায় এবং উত্তর-দফার সংখ্যা অনেক বেশী হয়ে যায়, সে ক্ষেত্রে বিভিন্ন উত্তর-দফাকে সঙ্কুচিত করে প্রত্যাশিত উত্তর-দফা নির্মাণ করা যেতে পারে। তবে মনে রাখতে হবে যে, একবার উত্তর-দফা চিহ্নিত হয়ে গেলে, সেই উত্তর-দফাগুলোকে আর পরে ভেঙ্গে নতুন উত্তর-দফা তৈরি করা যায় না। তৃতীয়তঃ, গবেষণা সমস্যার সাথে সম্পর্কিত তত্ত্ব, গবেষণা সমস্যা ও গবেষণাধীন জনগোষ্ঠীর উপর গবেষকের জ্ঞানের ভিত্তিতে উত্তর-দফাগুলোকে সংশোধন ও নির্মাণ করা উচিত।

অবরোহীমূলক ও আরোহীমূলক সাংকেতীকরণ (Deductive and Inductive Coding)

উত্তরদাতার কাছ থেকে গবেষক কি ধরণের উত্তর-দফা প্রত্যাশা করবেন, সে সম্পর্কে তিনি তত্ত্ব থেকেই ধারণা গ্রহণ করে থাকেন। অর্থাৎ, যে সব গবেষক পরিমাণমূলক বিশ্লেষণের সাথে সংশ্লিষ্ট, তারা সাধারণতঃ তত্ত্ব থেকে আহরিত অনুকল্প পরীক্ষা করেন এবং তাদের ব্যবহৃত সাংকেতীকরণ ব্যবস্থাটির মাধ্যমে তারা সেই তত্ত্বকে সমর্থন, বা ভুল প্রমাণ করার চেষ্টা করেন। যেমন, গবেষক যদি উদারনৈতিকতা সম্পর্কে গবেষণা করতে চান, তবে তত্ত্বই তাকে এই প্রত্যয়টির বহুমাত্রিকতার বিষয়টি সম্পর্কে দিক নির্দেশনা প্রদান করবে। একজন ব্যক্তি সামাজিকভাবে উদারনৈতিক হতে পারেন, কিন্তু অর্থনৈতিকভাবে নাও হতে পারেন। অর্থাৎ, একজন ব্যক্তি নারীর সম্মান নেবার, বা না নেবার অধিকারকে সমর্থন করলেও জন্ম নিয়ন্ত্রণ পদ্ধতির ব্যবহার বিস্তারের লক্ষ্যে সরকারের অর্থ ব্যয়কে সমর্থন নাও করতে পারেন। এ ক্ষেত্রে, সামাজিকভাবে উদারনৈতিকতার উচ্চ মান অর্থনৈতিক উদারনৈতিকতার উচ্চ মানের সাথে সম্পর্কিত হবে না। অতএব, উপাত্ত সংগ্রহের হাতিয়ার গঠনের ক্ষেত্রে, সম্ভাব্য উত্তর-দফা নির্মাণে তত্ত্বকে ব্যবহার করে অবরোহীমূলক সাংকেতীকরণ করা যায়। এ ক্ষেত্রে, উত্তর-দফাগুলো পূর্ব-নির্ধারিত হয়ে থাকে। তবে পূর্ব-পরীক্ষার (pre-test) মাধ্যমে তত্ত্ব থেকে আহরিত প্রত্যাশিত উত্তর-দফাগুলোকে আরো তীক্ষ্ণ করে তোলা যায়।

যখন উত্তরদাতার কাছ থেকে কি ধরণের উত্তর-দফা আশা করা যায় সে বিষয়ে তত্ত্ব গবেষককে খুব একটা দিক নির্দেশনা দিতে পারে না, তখন আরোহীমূলক সাংকেতীকরণ পদ্ধতি ব্যবহার করা হয়।

যখন একটি গবেষণা বর্ণনামূলক হয়, বা যখন উত্তরদাতার কাছ থেকে কি ধরণের উত্তর-দফা আশা করা যায় সে বিষয়ে তত্ত্ব গবেষককে খুব একটা দিক নির্দেশনা দিতে পারে না, তখন আরোহীমূলক সাংকেতীকরণ পদ্ধতি ব্যবহার করা হয়। আরোহীমূলক সাংকেতীকরণে, গবেষক দলিলপত্র থেকে প্রাপ্ত তথ্য, বা বিভিন্ন প্রশ্নের উত্তর (বিশেষ করে প্রান্ত-খোলা প্রশ্নের) থেকে প্রাপ্ত তথ্য, বা অংশগ্রহণমূলক পর্যবেক্ষণের মাধ্যমে প্রাপ্ত তথ্যের প্রতিনিধিত্বমূলক নমুনার ভিত্তিতে সাংকেতীকরণ ছক নির্মাণ করেন। একটি আরোহীমূলক সাংকেতীকরণ ছকে যে সব উত্তর বেশী আসে, উপাত্ত বিশ্লেষণের জন্য সেগুলোকে সাংকেতীকরণ ছকে অন্তর্ভুক্ত করা হয়। একবার যখন গবেষক একটি সাংকেতীকরণ ছক নির্মাণ করে ফেলেন, তখন তা অবশিষ্ট উপাত্তের উপর প্রয়োগ করা হয়। একটি উদাহরণ দেওয়া যাক। ধরা যাক, একজন নারী তার স্বামী কর্তৃক নির্যাতিত হবার পর তিনি কি করতে পারেন, এ প্রশ্নের উত্তরে প্রাপ্ত উত্তরদফাগুলোর জন্য একটি সাংকেতীকরণ ছক নির্মাণ করা প্রয়োজন। সে ক্ষেত্রে, সম্ভাব্য উত্তরগুলো যা হতে পারে, তা হলো নিম্নরূপ:

১. স্বামীকে ছেড়ে বাড়ি থেকে চলে যাবো।
২. স্বামীর সাথে থেকেই সমস্যাটি সমাধানের চেষ্টা করবো।
৩. কোন বন্ধু বা আত্মীয়ের কাছে সাহায্য চাইবো।
৪. কোন সামাজিক সেবামূলক সংস্থার কাছে পরামর্শের জন্য যাবো।
৫. নির্যাতনকারী স্বামীর বিরুদ্ধে একটি সাময়িক নিষেধাজ্ঞা আরোপ করার জন্য আদালতে যাবো।
৬. পুলিশে খবর দেবো।

দেখা যাবে যে, বাস্তবে অধিকাংশ নারীই এ ধরণের প্রতিক্রিয়া ব্যক্ত করবেন। এই উত্তরগুলো বেশী উল্লিখিত হয় বলে, এগুলো মূল উত্তর-দফায় স্থান পায়। কিন্তু কিছু কিছু নারী ভিন্ন উত্তরও দিতে পারেন।

কেউ উত্তরদানে অস্বীকৃতি জানাতে পারেন। কোন কোন নারীর ক্ষেত্রে, প্রশ্নটি প্রযোজ্য নাও হতে পারে। কেউ কেউ হয়তো জানেন না যে, তারা কি করবেন। আবার কোন কোন প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীতে প্রশ্নটির ক্ষেত্রে, কোন উত্তর লিপিবদ্ধ করা নাও থাকতে পারে। এ সকল সম্ভাবনার কথা চিন্তা করে আরোহীমূলক সাংকেতীকরণের মাধ্যমে প্রধান উত্তর-দফাগুলো চিহ্নিত করার পর, গবেষক চূড়ান্ত সাংকেতীকরণ ছকে নিম্নলিখিত দফাগুলোকে যুক্ত করেন।

১. অন্যান্য (নির্দিষ্ট করুন)।
২. জানি না।
৩. উত্তরদানে অস্বীকৃতি
৪. প্রযোজ্য নয় (কারণ উত্তরদাতা বিবাহিত নন)
৫. কোন উত্তর নেই।

আরোহীমূলক পদ্ধতির প্রধান সুবিধা হলো এর নমনীয়তা এবং সমৃদ্ধতা, যা গবেষণায় প্রাপ্ত তথ্য থেকে ব্যাখ্যা গঠনে সহায়তা করে। তাছাড়া, এটি গবেষককে একটি পর্যবেক্ষণের ওপর বিভিন্ন ধরনের সাংকেতীকরণ ছক নির্মাণের সুযোগ করে দেয় এবং প্রায়ই এতে নতুন নতুন উত্তর-দফার প্রস্তাব করা যায়। এই পদ্ধতির অসুবিধা হলো যে, এতে বিশাল উপাত্তরাশির ব্যাখ্যার প্রচেষ্টায় গবেষক দিক বিদিক হারিয়ে ফেলতে পারেন। উত্তর-দফা চিহ্নিত করা সবসময় সহজ হয় না, বলে একটি সর্বাঙ্গিক সাংকেতীকরণ ছক নির্মাণ প্রক্রিয়া সম্পন্ন করতে গবেষকের দীর্ঘ সময় লেগে যায়।

সংকেত-বই নির্মাণ (Code-Book Construction)

সংকেত-বই হলো, কোন চলক বা প্রশ্নের বিপরীতে কি কি উত্তর-দফা থাকবে এবং সে সকল উত্তর-দফার বিপরীতে কি কি সংকেত প্রদান করা হবে, তার একটি পরিকল্পনা। সেই পরিকল্পনা অনুযায়ী গবেষক সাংকেতীকরণ প্রক্রিয়াকে কার্যকর করেন। এই পরিকল্পনার রূপরেখাটিকে সাংকেতীকরণ ছক বলে। সাংকেতীকরণের নিয়মাবলী অনুসরণ করে প্রতিটি চলকের জন্য সাংকেতীকরণ ছক নির্মাণের পর, এগুলোকে একটি সংকেত-বইয়ে গ্রহিত করা হয়।

সংকেত-বই বিভিন্ন প্রকার হতে পারে। তবে, যে কোন একটি আদর্শ সংকেত-বইয়ে চলকের নাম ও নম্বর, প্রতিটি চলকের অবস্থান, সাংকেতীকরণ ছক এবং হারিয়ে যাওয়া উপাত্তের জন্য সংকেত সম্পর্কিত তথ্যাবলী থাকে। এই সংকেত-বইটি সাংকেতীকরণকারীর জন্য (যারা অপরিশোধিত উপাত্তকে কম্পিউটারে পরিসংখ্যানগত বিশ্লেষণে ব্যবহারের জন্য রূপান্তর করবেন) একটি নির্দেশিকা হিসাবে কাজ করে। এটি প্রধান গবেষকসহ অন্যান্য গবেষকের জন্য, যারা এই উপাত্তরাশিকে ব্যবহার করতে চান, একটি সহায়িকা হিসাবে কাজ করবে। এ ছাড়া, অনেক সংকেত-বইয়ে উপাত্ত সংগ্রহের উদ্দেশ্যে পরিচালিত মূল প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীকেও অন্তর্ভুক্ত করা হয়ে থাকে। একটি সংকেত-বইয়ের কি আঙ্গিক হওয়া উচিত, তার সংক্ষিপ্ত উদাহরণ সারণি ৯.২.১-এ দেখানো হলো।

সংকেত-বই হলো কোন চলক বা প্রশ্নের বিপরীতে কি কি উত্তর-দফা থাকবে এবং সে সকল উত্তর-দফার বিপরীতে কি কি সংকেত প্রদান করা হবে তার একটি পরিকল্পনা।

উদাহরণে প্রদত্ত সংকেত-বইয়ে দেখা যাচ্ছে যে, যে সব সংকেত-দফা ব্যবহার করা হয়েছে তার মান, যে স্তরে প্রতিটি চলকের সাংকেতীক মান পাওয়া যাবে তার অবস্থান, কোন মানটি হারানো মান হিসাবে চিহ্নিত হবে সেটি এবং প্রতিটি চলকের ক্ষেত্রে অন্য কি বিশেষ সাংকেতীক নিয়ম প্রয়োগ করা হবে, তা সহজেই চিহ্নিত করা যায়। পরিসংখ্যানমূলক কম্পিউটার প্রোগ্রাম অনুযায়ী উপাত্তকে বিস্তারিতপত্রের (spreadsheet) আকারে বিন্যস্ত করে চলকগুলোকে স্তর অনুসারে এবং ঘটনাগুলোকে সারি অনুসারে সাজানো হয়। এই উপাত্তরাশি ব্যবহার করতে আশ্রয়ী যে কোন গবেষককে প্রতিটি স্তর সংখ্যা বলে দেয় যে, কোন স্তরে কোন নির্দিষ্ট চলকের মান রয়েছে। সংকেত-বইয়ে ধারণকৃত তথ্য থেকেই যে কোন গবেষক সম্পূর্ণ তথ্য সমগ্রককে পুনর্গঠন করতে পারবেন।

সারণি ৯.২.১: একটি সংকেত-বইয়ের সংক্ষিপ্ত উদাহরণ

চলকের নাম	সাংকেতীকরণ ছক	স্তম্ভ নম্বর
পরিচয়বহনকারী নম্বর	উত্তরদাতার পরিচিতি নম্বর (প্রকৃত নম্বরটি লিখুন)	১-৩
৮০০১	লিঙ্গ ১ = নারী ২ = পুরুষ	৪
৮০০২	উত্তরদাতার বয়স (প্রকৃত বছর লিখুন)	৫-৬
৮০০৩	শিক্ষাগত যোগ্যতা: সম্পন্নকৃত সর্বোচ্চ শ্রেণী ১ = ১ - ৮ ২ = ৯ - ১১ ৩ = ১২ ৪ = ১৩ - ১৫ ৫ = ১৬ ৬ = ১৭	৭
৮০০৪	মাসিক আয়: বর্তমান/পূর্বের চাকরী (প্রকৃত আয়ের পরিমাণ লিখুন)	৮-১২
৮০০৫	সাধারণ স্বাস্থ্যগত অবস্থা ১ = খুবই ভাল ২ = খুব ভাল ৩ = ভাল ৪ = মোটামুটি ৫ = খারাপ ৬ = খুবই খারাপ ৮ = জানি না ৯ = উত্তর নেই	১৩
৮০০৬	শিক্ষা গ্রহণই সামনে এগিয়ে যাওয়ার উপায় ১ = পুরোপুরি একমত ২ = একমত ৩ = মোটামুটি একমত ৪ = মোটামুটি দ্বিমত পোষণ করি ৫ = দ্বিমত পোষণ করি ৬ = পুরোপুরি দ্বিমত পোষণ করি ৮ = জানি না ৯ = উত্তর নেই	১৪

সাংকেতীকরণের সাধারণ নির্দেশনা (General Coding Instructions)

সাংকেতীকরণ লাল কালি এবং মাঠ বিভাগের শুদ্ধিকরণগুলো সবুজ কালি দিয়ে আলাদা করা উচিত। সাক্ষাৎকার গ্রহণকারী কর্তৃক বৃত্তকৃত কোন সংকেত, বা মতামত কখনোই মুছে ফেলা উচিত নয়। প্রশ্নমালা বা সাক্ষাৎকার অনুসূচী সংশোধনের সময়, ভুল হিসাবে বৃত্তকৃত সংকেতটির মধ্য দিয়ে এমনভাবে একটি দাগ টানতে হবে, যাতে করে পূর্বে যা করা হয়েছিলো তা যেন উদ্ধার করা অসম্ভব হয়ে না পড়ে। প্রতিটি স্তম্ভে মাত্র একটি করে সংকেত থাকতে হবে এবং কোন স্তম্ভেই একের অধিক স্পষ্টভাবে বৃত্তকৃত সংকেত থাকবে না।

যে সব প্রশ্নে, কোন বিশেষ সংকেত সমস্যা অনুমান করা হয়েছে, সে সব ক্ষেত্রে, সাংকেতিকরণকারীদের কিছু ছোট ছোট পতাকা সরবরাহ করা হয়, যা সমস্যা সম্বলিত পাতায় আটকাতে হবে। সাংকেতিকরণকারীরা অন্যান্য অংশকেও পতাকায়ুক্ত করতে পারবেন, বিশেষতঃ যেখানে তা অসম্পূর্ণ বা অস্পষ্ট, বা যাতে তত্ত্বাবধায়কের কিছুটা মনযোগ প্রয়োজন হবে। যেসব প্রশ্নে 'LIST' বলে চিহ্নিত করা থাকবে, সে সব প্রশ্নের ক্ষেত্রে একটি ভিন্ন তালিকা থাকতে হবে। যদি কোন প্রশ্নের একাধিক তালিকাবদ্ধ সংকেত থাকে, তবে প্রতিটি সংকেতের জন্য আলাদা আলাদা তালিকা রাখতে হবে। প্রতিটি ফর্মের উপরে গবেষণা নম্বর, প্রশ্ন নম্বর, ডেক ও স্তম্ভ নম্বর এবং তালিকাবদ্ধ সংকেত লিপিবদ্ধ করতে হবে।

যখন উত্তরদাতা উত্তর দেন না, বা যখন সাক্ষাৎকার গ্রহণকারী কোন প্রশ্ন করতে, বা উত্তর লিপিবদ্ধ করতে ভুলে যান, বা যখন লিখিত তথ্য পরস্পর বিরোধী, বা দ্ব্যর্থমূলক হয়, তখন 'উত্তর নেই' সংকেতটিকে গ্রহণ করা হয়। প্রশ্নমালায় 'উত্তর নেই' এবং 'উত্তরদানে অসম্মতি' সংকেতের জন্য, এক স্তম্ভের ক্ষেত্রে ব্যবহৃত সংখ্যা মানটি হলো ৭, দুই স্তম্ভের ক্ষেত্রে ৯৭, তিন স্তম্ভের ক্ষেত্রে ৯৯৭, ইত্যাদি। 'জানি না' সংকেত যদি আগে থেকে নির্ধারিত না হয়ে থাকে, তবে এক স্তম্ভের ক্ষেত্রে, এটির জন্য ৮ সংখ্যাটি বরাদ্দ করা হয়, দুই স্তম্ভের ক্ষেত্রে ৯৮, তিন স্তম্ভের ক্ষেত্রে ৯৯৮ বরাদ্দ করা হয়। একটি প্রশ্ন কোন উত্তরদাতাকে জিজ্ঞাসা করার প্রয়োজন না হলে, 'প্রযোজ্য নয়' সংকেতটি ব্যবহৃত হয়। সে জন্য, এক স্তম্ভের ক্ষেত্রে সংকেত-মানটি ধরা হয় ৯, দুই স্তম্ভের ক্ষেত্রে ৯৯, তিন স্তম্ভের ক্ষেত্রে ৯৯৯, ইত্যাদি।

উপাত্ত অন্তর্ভুক্তির কৌশল (Data Entry Devices)

সংকেত-বই তৈরি হয়ে যাবার পর, সাংকেতিকরণের মাধ্যমে উপাত্তকে এমনভাবে স্থানান্তর করতে হবে, যাতে করে যে কেউ তা সংরক্ষণ ও বিশ্লেষণের জন্য কম্পিউটারে অন্তর্ভুক্ত করতে পারেন। অর্থাৎ, প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীতে বৃত্তকৃত সংখ্যাগুলোকে যেন যে কেউ সংকেত-বইয়ে সংজ্ঞায়িত চলকের যথাযথ স্তম্ভ বা ক্ষেত্রে প্রবেশ করাতে পারেন। অপরিশোধিত উপাত্তকে সাংকেতিকরণ করে উপাত্ত অন্তর্ভুক্তির বিভিন্ন কৌশল রয়েছে। নিম্নে, আমরা উপাত্ত অন্তর্ভুক্তির বিভিন্ন কৌশল আলোচনা করবো।

স্থানান্তরপত্র (Transfer Sheet): শুরুতে, কম্পিউটার কার্ডে যন্ত্রের মাধ্যমে ছিদ্র করে সকল ধরণের উপাত্ত অন্তর্ভুক্ত করা হতো। এই কম্পিউটার কার্ডে ৮০টি সংখ্যাসূচক প্রতীক সংরক্ষণ করা যায়। স্থানান্তরপত্রও কম্পিউটার কার্ডের মত ৮০টি স্তম্ভ রয়েছে। সাংকেতিকরণকারী সংকেত-বইয়ে নির্ধারিত স্তম্ভ ও ক্ষেত্র অনুযায়ী, স্থানান্তরপত্রের স্তম্ভ ও ক্ষেত্রে উপাত্তের উৎস থেকে উপাত্ত স্থানান্তর করেন। স্থানান্তরপত্রে, উপাত্তের প্রতিটি লাইন ঘটনার পরিচিতি নম্বর দিয়ে শুরু করতে হয়। যদি উপাত্তের জন্য একটি লাইনের বেশী প্রয়োজন হয়, তবে পরিচিতি নম্বর দিয়ে দ্বিতীয় লাইনটিতে উপাত্ত স্থানান্তর করতে হবে। স্থানান্তরপত্রে উপাত্ত লিপিবদ্ধ করা হয়ে গেলে, সেই উপাত্তের লাইনগুলো floppy disk, hard disk, বা magnetic tape-এ সংরক্ষণ করা হয়। তবে, এটি মনে রাখতে হবে যে, যেহেতু স্থানান্তরপত্র ব্যবহারের জন্য বিভিন্ন ব্যক্তিকে বিভিন্ন ধাপে কাজ করতে হয়, সেহেতু প্রতিটি ধাপে ভ্রান্তি হবার সুযোগ তৈরি হয়। ফলে, উপাত্তের নির্ভরশীলতা কমে যাবার সম্ভাবনা থাকে। সে জন্য, প্রতিটি পদক্ষেপে সঠিক চলকের জন্য সঠিক সংকেতটি অন্তর্ভুক্ত করা হয়েছে কি না, তা যাচাই করা উচিত।

প্রান্ত সাংকেতিকরণ (Edge Coding): প্রান্ত সাংকেতিকরণ কৌশলের ব্যবহার স্থানান্তরপত্রের প্রয়োজনীয়তাকে দূরীভূত করে। এই পদ্ধতিতে সাংকেতিকরণকারী তথ্যগুলোকে প্রশ্নমালা বা সাক্ষাৎকার অনুসূচীর বাইরের প্রান্তে যে ফাঁকা জায়গা থাকে, সেই স্থানে স্থানান্তর করেন। যখন একটি উপাত্ত সংগ্রহের হাতিয়ারকে সাংকেতিকরণ করা হয়, তখন একজন উপাত্ত অন্তর্ভুক্তিকরণ কর্মী প্রান্তে লিপিবদ্ধকৃত তথ্যকে সরাসরি floppy disk, hard disk, বা magnetic tape-এ সংরক্ষণ করতে পারেন।

আলোক অভিবীক্ষণ (Optical Scanning): সাংকেতীকরণকারী আলোক অভিবীক্ষণ পদ্ধতির মাধ্যমেও উপাত্তকে বিশ্লেষণের জন্য সংরক্ষণ করতে পারেন। একজন পরীক্ষার্থী যেভাবে সরাসরি একটি OMR Form -এর উপর তার প্রশ্নের উত্তর বৃত্তাকার করে, একজন সাংকেতীকরণকারীও সেভাবে উপাত্তকে আলোক অভিবীক্ষণপত্রের (optical scanning sheet) উপর স্থানান্তর করতে পারেন। অভিবীক্ষণকারী যন্ত্র সেই কালো কালির বৃত্তগুলোকে চিহ্নিত করে একটি উপাত্তের ফাইল তৈরি করে। আলোক অভিবীক্ষণের ব্যবহার এত ব্যাপক যে, উত্তরদাতা সরাসরি একটি অভিবীক্ষণপত্র পূরণের মাধ্যমে তার উত্তর প্রদান করতে পারেন।

সরাসরি উপাত্ত অন্তর্ভুক্তিকরণ (Direct Data Entry): সাংকেতীকরণের সবচেয়ে উদ্ভাবনীমূলক পদ্ধতিটি সম্ভবতঃ সরাসরি উপাত্ত অন্তর্ভুক্তিকরণ। দু'ধরণের সরাসরি উপাত্ত অন্তর্ভুক্তিকরণ কৌশল রয়েছে। একটি হলো, উপাত্ত সংগ্রহের হাতিয়ার থেকে সাংকেতীকরণ এবং অন্যটি হলো, টেলিফোন সাক্ষাৎকারের মাধ্যমে সাংকেতীকরণ। দু'টি ধরণই কম্পিউটার প্রোগ্রামের উপর ভিত্তি করে করা হয়, যা কম্পিউটারের পর্দায় উপাত্ত সংগ্রহের হাতিয়ারের প্রতিটি দফাকে প্রদর্শন করে এবং সাংকেতীকরণকারী বা সাক্ষাৎকার গ্রহণকারীকে প্রয়োজনীয় উত্তর-দফাকে অন্তর্ভুক্ত করার জন্য নির্দেশ করে। তবে, প্রশ্নমালা বা সাক্ষাৎকার অনুসূচী থেকে সাংকেতীকরণকৃত বিষয়কে সতর্কতার সাথে সম্পাদনা করতে হয়। অন্যদিকে, কম্পিউটারের সাহায্যে টেলিফোন সাক্ষাৎকার অনেক বেশী পরিশীলিত পদ্ধতি বলে, সে ক্ষেত্রে সাংকেতীকরণে ভ্রান্তি হবার সুযোগ থাকে অনেক কম।

উপাত্ত বিশুদ্ধকরণ (Data Cleaning)

যে কৌশল ব্যবহার করেই উপাত্ত অন্তর্ভুক্ত করা হউক না কেন, বা যত যত্ন করেই তা অন্তর্ভুক্ত করা হউক না কেন, কিছু ভুল ভ্রান্তি হবেই। কোন চলককে বাদ দেওয়া, ভুল মান অন্তর্ভুক্ত করা, ভুল স্তম্ভে মান বসানো, ভুল সাংকেতীকরণ করা, লিপিবদ্ধকৃত সংকেতকে ভুল পড়া, বিভিন্ন চিহ্নের ভুল অর্থ করা, ইত্যাদি কারণে ত্রুটি হতে পারে। বিশ্লেষণের পূর্বে, এই ত্রুটিগুলো সংশোধন করে নিতে হবে, নইলে যথাযথ পরিসংখ্যানগত পদ্ধতি প্রয়োগ করা যাবে না এবং এর ফলে ফলাফলের পরিসংখ্যানগত বিশ্লেষণ বিভ্রান্তিকর হবে। উপাত্ত বিশুদ্ধকরণের বিভিন্ন কৌশল রয়েছে, যেগুলোর নির্বাচন ও প্রয়োগ নির্ভর করে নমুনার আকার, উপাত্ত অন্তর্ভুক্তির কৌশল, লভ্য প্রযুক্তি এবং উপযুক্ত কলাকৌশলের উপর গবেষকের জ্ঞান ও প্রয়োগের দক্ষতার উপর।

যখন উপাত্ত অন্তর্ভুক্তির জন্য punch-card ব্যবহার করা হয়, তখন প্রতিটি স্তম্ভে চলকের বৈশিষ্ট্যের জন্য স্থান নির্ধারণ করা থাকে। যেমন, লিঙ্গ চলকটির জন্য punch-card-এ সম্ভাব্য তিনটি ছিদ্র করতে হয় — নারীর জন্য '১', পুরুষের জন্য '২' এবং উত্তর নেই সংকেতের জন্য '০'। কিন্তু যদি '১' -এর ঘরে একটি ছিদ্র পাওয়া যায়, তবে বুঝতে হবে যে, সেখানে ভ্রান্তি হয়েছে। তাৎক্ষণিকভাবে, সেটিকে শুধরে নিতে হবে। উপাত্তের মধ্যে যৌক্তিকতার কাঠামো নিরীক্ষা করেও ত্রুটিকে সংশোধন করা যায়। যেমন, একজন নারীর কতজন সন্তান রয়েছে প্রশ্নের উত্তরে ধরে নেওয়া যায় যে, কোন না কোন সংকেত মান তার উত্তর-দফার বিপরীতে থাকবে। সেটি '০' হতে পারে, অথবা '৫' হতে পারে। কিন্তু যদি কোন পুরুষ উত্তরদাতার ক্ষেত্রে কোন সংকেত মান পাওয়া যায়, সে ক্ষেত্রে ধরে নিতে হবে যে, সেটি একটি ভ্রান্তি। কারণ, প্রশ্নটি নারীর ক্ষেত্রে প্রযোজ্য, পুরুষের ক্ষেত্রে নয়। এ কৌশলটি, হয় বাছাই প্রক্রিয়া, নতুবা কম্পিউটার প্রোগ্রামের মাধ্যমে প্রয়োগ করা যেতে পারে।

উপাত্তরাশি যখন ছোট হয় (অর্থাৎ, যদি চলকের সংখ্যা ৫০টির মত হয়), তখন কম্পিউটারে উপাত্ত সম্পাদনার পর্দায় (data editor screen) উপাত্তকে অভীক্ষণের মাধ্যমে পরীক্ষা করা যেতে পারে। কম্পিউটারের পর্দায় সম্পাদনার পাতা প্রক্ষেপণ করলেই সম্ভাব্য ত্রুটিগুলো চোখে পড়ে। যেমন, খালি ঘর, অস্বাভাবিক রকমের বড় সংখ্যা, সংখ্যার পরিবর্তে বর্ণমালা, ইত্যাদি। তবে কোন নিয়মতান্ত্রিক পদ্ধতি ব্যবহারের মাধ্যমে এটি করা ভাল। যেমন, কম্পিউটারে ব্যবহারযোগ্য পরিসংখ্যানগত পদ্ধতি (statistical package) বা বিস্তারিতপত্রের (spreadsheet) মাধ্যমে বাছাই প্রক্রিয়া (sort

উপাত্তরাশি যখন ছোট হয় তখন কম্পিউটারে সম্পাদনার পর্দায় উপাত্তকে অভীক্ষণের মাধ্যমে পরীক্ষা করা যেতে পারে।

procedure) ব্যবহার করা যেতে পারে। বাছাই প্রক্রিয়া sort-key হিসাবে নির্বাচিত চলকের মানগুলোর ভিত্তিতে উপাত্তরাশির মধ্যে সব ঘটনাগুলোর ক্রমকে পূর্ণবিন্যস্ত করে। বর্ণমালা চলকগুলোকে বর্ণমালার ভিত্তিতে এবং সংখ্যাসূচক চলকগুলোকে সংখ্যার ভিত্তিতে ক্রম অনুযায়ী সাজানো হয়। বাছাইকৃত উপাত্তকে কম্পিউটারের পর্দায় পরীক্ষা করলেই ত্রুটিগুলো ধরা পড়বে। কিন্তু অনেক বেশী উপাত্তরাশির ক্ষেত্রে, এই পদ্ধতিটি জটিল আকার ধারণ করে। সে ক্ষেত্রে, SPSS ব্যবহারের মাধ্যমে ঘটনাসংখ্যা প্রক্রিয়া (frequency procedure) ব্যবহার করে ত্রুটি বের করা যেতে পারে।

সাংকেতীকরণের নির্ভরশীলতা (Reliability of Coding)

যথাযথভাবে নির্মিত সংকেত-বই, পূর্ব-সাংকেতীকরণকৃত প্রান্ত-বন্ধ প্রশ্ন এবং সাংকেতীকরণকারীদের যথাযথ প্রশিক্ষণ ও অভিজ্ঞতাসহ একটি গবেষণা পরিচালনা করলে, সাংকেতীকরণের নির্ভরশীলতা নিয়ে তেমন সমস্যা দেখা দেয় না। কারণ, এ ধরনের গবেষণায়, কোন উত্তরের জন্য কোন সংকেত প্রদান করতে হবে, সে বিষয়ে সাংকেতীকরণকারীদের নিজেদের বিবেচনাপ্রসূত সিদ্ধান্ত নিতে হয় খুব কম। এ ধরনের গবেষণায়, সবচেয়ে গুরুত্বপূর্ণ বিষয়টি হলো, সাংকেতীকরণকারীরা সংকেতগুলোকে সঠিক স্তম্ভে স্থাপন করছেন কি না, তা নিশ্চিত করা। এটি নিশ্চিত করার জন্য প্রতিটি সাংকেতীকরণকারীর কাজকে নমুনাভিত্তিক যাচাই করা হয়। কি ধরনের সাংকেতীকরণ কৌশল ব্যবহার করা হচ্ছে, তার উপরও সাংকেতীকরণের নির্ভরশীলতার মাত্রা নির্ভর করে।

কিন্তু প্রান্ত-খোলা প্রশ্ন, বা অন্য কোন অকাঠামোবদ্ধ উপাত্ত সংগ্রহের হাতিয়ার ব্যবহার করে পরিচালিত গবেষণার উপাত্ত সাংকেতীকরণ প্রক্রিয়ায়, উত্তর-দফার সংকেত প্রদানের ক্ষেত্রে, সাংকেতীকরণকারীদের অনেক বেশী নিজেদের বিবেচনাপ্রসূত সিদ্ধান্ত গ্রহণ করতে হয়। এর ফলে, বিভিন্ন সাংকেতীকরণকারী বিভিন্ন উত্তরের বিভিন্ন ব্যাখ্যা করে সংকেত প্রদান করতে পারেন। এ ধরনের পরিস্থিতিতে, সাংকেতীকরণ অনির্ভরযোগ্য হয়ে পড়তে পারে। এ ক্ষেত্রে, সবচেয়ে সহজ কৌশলটি হলো যে, প্রায়শঃ দু'তিনজন সাংকেতীকরণকারীর সংকেতগুলোকে যাচাই করে দেখা এবং তাৎক্ষণিকভাবে সমস্যাটির সমাধান করে বিষয়টি অন্যান্য সাংকেতীকরণকারীদের জানিয়ে দেয়া। উত্তর-দফার ব্যাখ্যার ভিন্নতা ও অসামঞ্জস্যতা সম্পর্কিত অন্য সমস্যাটি হলো যে, সাংকেতীকরণকারী ও উত্তরদাতার মধ্যে উত্তর-দফার অর্থের ব্যাখ্যার ভিন্নতা। এই ভিন্নতা পুরোপুরিভাবে কখনো দূর করা যাবে না, তবে গভীর পূর্ব-পরীক্ষা ও প্রশিক্ষণ এই অসংগতির মাত্রা অনেকটা কমিয়ে আনতে পারে।

সারাংশ

উপাত্ত সংগ্রহের পরপরই যে কাজটি গবেষককে করতে হয়, তা হলো প্রাপ্ত উপাত্তকে সংকেতে রূপান্তর করা। কারণ, সংকেতে বা সংখ্যায় রূপান্তর করার পরই, কেবল সংক্ষিপ্তভাবে উপাত্তের বর্ণনা ও বিশ্লেষণ করা সম্ভব। উপাত্ত সাংকেতীকরণের জন্য প্রথমেই প্রয়োজন, একটি সংকেত-বই নির্মাণ করা। সংকেত-বই মূলতঃ উপাত্ত প্রক্রিয়াজাতকরণের নির্দেশিকা হিসাবে কাজ করে। প্রাপ্ত উপাত্তের যে কোন সমস্যা কিভাবে সমাধান করা হবে, তা এই সংকেত-বইয়ে বর্ণিত হয়ে থাকে। বর্তমান পাঠে, সাংকেতীকরণের প্রয়োজনীয়তা, নিয়মাবলী এবং সর্বোপরি কিভাবে প্রাপ্ত উপাত্তকে বিশুদ্ধ করা যায় তা আলোচনা করা হয়েছে। তবে মনে রাখতে হবে যে, অনেক সময় প্রাপ্ত উপাত্তের মান উন্নত হলেও গবেষণায় ফলাফল ভালো হবে না, যদি উপাত্ত প্রক্রিয়াজাতকরণ যথাযথ না হয়।

পাঠোত্তর মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক (✓) চিহ্ন দিন –

১। উপাত্ত বিশুদ্ধকরণের প্রয়োগ নির্ভর করে:

- ক. উপাত্ত অন্তর্ভুক্তির কৌশলের উপর
- খ. উপযুক্ত কলাকৌশলের উপর গবেষকের জ্ঞানের মাত্রার উপর
- গ. কলাকৌশল প্রয়োগের দক্ষতার উপর
- ঘ. উপরের সব।

২। চলকের বিপরীতে সংখ্যাসূচক সংকেত বরাদ্দের ক্ষেত্রে, যে বিষয়টি গুরুত্বের সাথে বিবেচনায় নিতে হবে সেটি হলো:

- ক. সমরূপ পরিস্থিতিতে সংকেতের এককে ভিন্নতা থাকতে হবে
- খ. সংকেতের একক পরিস্থিতি অনুযায়ী নির্ধারিত হতে হবে
- গ. সমরূপ পরিস্থিতিতে সংকেতের একক সামঞ্জস্যপূর্ণ হতে হবে
- ঘ. উপরের সব।

৩। একটি আদর্শ সংকেত-বইয়ে:

- ক. প্রতিটি চলকের অবস্থান থাকে
- খ. হারিয়ে যাওয়া উপাত্তের জন্য নির্দেশনা থাকে
- গ. প্রতিটি চলকের বর্ণনা থাকে
- ঘ. উপরের সব।

সংক্ষিপ্ত প্রশ্ন

টীকা লিখুন:

- ক. সাংকেতীকরণের নির্ভরশীলতা
- খ. উপাত্ত বিশুদ্ধকরণ

রচনামূলক প্রশ্ন

- ১। সাংকেতীকরণ কি? উপাত্ত ব্যবহারে এর গুরুত্ব উদাহরণসহ আলোচনা করুন।
- ২। সাংকেতীকরণের নিয়মাবলীগুলো উদাহরণসহ আলোচনা করুন।

উপাত্ত সংক্ষিপ্তকরণ ২ – সারণিকরণ Data Reduction 2 – Tabulation

এই পাঠ শেষে যা জানা যাবে —

- সারণিকরণের গুরুত্ব
- সারণির প্রকারভেদ
- সারণি গঠনে বিবেচ্য বিষয়সমূহ
- গণসংখ্যা নিবেশন

সারণিকরণের গুরুত্ব (Importance of Tabulation)

যেহেতু বিশাল উপাত্তরাশিকে সংক্ষিপ্ত করে বিশ্লেষণের উপযোগী করে তোলা হলো পরিসংখ্যানের অন্যতম প্রধান কাজ, সেহেতু সম্পাদনা ও সাংকেতিকরণের পর উপাত্তকে কোন না কোনভাবে উপস্থাপনের প্রয়োজন হয়ে পড়ে। উপাত্তকে সাধারণতঃ চারটি উপায়ে উপস্থাপন করা যায়। প্রথমতঃ, মূল আলোচনার মধ্যে অন্তর্ভুক্ত করে। দ্বিতীয়তঃ, আধা-সারণির মধ্যে বিন্যস্ত করে। তৃতীয়তঃ, লেখচিত্রের মাধ্যমে প্রকাশ করে। চতুর্থতঃ, সারণির মধ্যে উপস্থাপন করে। আলোচনার মধ্যে উপাত্তের উপস্থাপন বেশী কার্যকর ফলাফল প্রদান করে না। একটি বর্ণনার মধ্যে উপস্থাপিত উপাত্তসমূহের অর্থ অনুধাবন করার জন্য পুরো বর্ণনাটিকে পাঠ করতে হয় এবং পুরো পাঠের পরও সব উপাত্তের অর্থ ও প্রাসঙ্গিকতা অনুধাবন করা সহজ হয়ে উঠে না। কোন নির্দিষ্ট উপাত্তের বিশেষ বৈশিষ্ট্যকে তুলে আনা পাঠকের জন্য কঠিন হয়ে পড়ে। এই পদ্ধতিতে, উপাত্ত উপস্থাপন পাঠকের মনোযোগকে একটি নির্দিষ্ট দিকে প্রবাহিত করতে পারলেও বিশেষ বিশেষ উপাত্তের দিকে আগ্রহ সৃষ্টি এবং তুলনার ক্ষেত্র প্রস্তুত করতে পারে না।

আধা-সারণির মাধ্যমে উপাত্ত উপস্থাপন প্রথম পদ্ধতির চেয়ে কিছুটা উন্নততর। যখন একটি আলোচনায় স্বল্পসংখ্যক উপাত্ত ব্যবহার করা হয়, তখন উপাত্তকে মূল পাঠ্যাংশ থেকে আলাদা করে নিয়ে দু'টি অনুচ্ছেদের মাঝখানে উপস্থাপনের মাধ্যমে উপাত্তের স্বকীয়তাকে তুলে ধরা যায়। তবে, উপাত্তের পরিমাণ বেশী হলে, এই পদ্ধতি পুরোপুরি অনুপযোগী হয়ে উঠে। এই পদ্ধতিটি সামাজিক গবেষণায় সচরাচর তেমন ব্যবহৃত হয় না। লেখচিত্র বা অন্যান্য চিত্রের মাধ্যমে উপাত্তের উপস্থাপন বেশ উপযোগী এবং কার্যকর। তবে, উপাত্তের পরিমাণ, মাত্রা ও ব্যাপ্তি বেড়ে গেলে, এই পদ্ধতিও তেমন কার্যকর থাকে না। সীমিত সংখ্যক উপাত্তের ক্ষেত্রে এটি চমৎকার ফলাফল দেয়।

সারণিকরণ হলো উপাত্ত উপস্থাপনের শ্রেষ্ঠ উপায়। সারণির ভিতরে কি বিষয়ের উপর উপাত্ত উপস্থাপন করা হয়েছে, একটি সারণির শিরোনাম তা সহজেই বুঝতে সাহায্য করে। সারণিতে উপস্থাপিত উপাত্তের বিশ্লেষণ ও ব্যাখ্যা গবেষণায় প্রাপ্ত ফলাফলের বিশেষ দিকগুলোকে বুঝতে সাহায্য করে। একটি সারণি বৈশিষ্ট্যগতভাবেই নিবিড়, সংক্ষিপ্ত ও স্পষ্ট। সারণির স্তম্ভ ও সারির সুনির্দিষ্ট শিরোনামগুলো পাঠ্যাংশে বর্ণিত পুনরাবৃত্তিগুলোকে অপ্রয়োজনীয় করে তোলে। ফলে, উপাত্তের নিবিড়তা ও স্পষ্টতা বৃদ্ধি পায় এবং তুলনামূলক বিশ্লেষণে অনেক বেশী সহায়ক হয়ে উঠে। কিন্তু এটি মনে রাখা প্রয়োজন যে, একটি সারণি যত সুগঠিতই হোক না কেন, যারা পরিসংখ্যানগত উপাত্ত বিশ্লেষণের সাথে খুব বেশী পরিচিত নন, তাদের কাছে অনেক সময় সারণিতে বিন্যস্ত উপাত্তরাশি দুর্বোধ্য হয়ে উঠে। লেখচিত্রের মাধ্যমে উপাত্তের উপস্থাপন সে কারণে সে ধরণের পাঠকের কাছে অনেক বেশী জনপ্রিয়। লেখচিত্র সাধারণ পাঠকের কাছে উপাত্তের বৈশিষ্ট্যগুলোকে সহজেই দৃষ্টিগ্রাহ্য ও অনুধাবনযোগ্য করে তোলে।

একটি সারণি যত সুগঠিতই হোক না কেন, যারা পরিসংখ্যানগত উপাত্ত বিশ্লেষণের সাথে খুব বেশী পরিচিত নন, তাদের কাছে অনেক সময় সারণিতে বিন্যস্ত উপাত্তরাশি দুর্বোধ্য হয়ে উঠে। লেখচিত্রের মাধ্যমে উপাত্তের উপস্থাপন সে কারণে সে ধরণের পাঠকের কাছে অনেক বেশী জনপ্রিয়। লেখচিত্র সাধারণ পাঠকের কাছে উপাত্তের বৈশিষ্ট্যগুলোকে সহজেই দৃষ্টিগ্রাহ্য ও অনুধাবনযোগ্য করে তোলে।

পরিসংখ্যানগত প্রক্রিয়াকরণ, উপস্থাপন ও ব্যাখ্যার সুবিধার জন্য উপাত্ত সংক্ষিপ্তকরণের প্রক্রিয়াকে সারণিকরণ বলে। সারণিকরণ প্রক্রিয়া বহু কারণে গুরুত্বপূর্ণ। প্রথমতঃ, সংগৃহীত তথ্যকে যদি সারণির মাধ্যমে উপস্থাপন করা যায়, তবে সেগুলো সহজেই বোধগম্য হয়ে উঠে, সময়ের সাশ্রয় ঘটে, এবং বিভ্রান্তি দূর হয়। অর্থাৎ, সারণিকরণ জটিল তথ্যরাশিকে সহজবোধ্য করে তোলে। দ্বিতীয়তঃ, সারণির সাহায্যে তথ্যরাশির তুলনা সহজতর হয়। একটি কার্যকর ও উপযোগী সারণি সাধারণতঃ তথ্যরাশিকে সুসংবদ্ধভাবে স্তম্ভ ও সারির মাধ্যমে উপস্থাপন করে বলে উপাত্তের তুলনামূলক বিশ্লেষণের একটি স্পষ্ট রূপরেখা পাওয়া যায়। তৃতীয়তঃ, যদি উপাত্তকে সারণিতে উপস্থাপন না করে শুধু বর্ণনা করা হয়, তবে অনেক গুরুত্বপূর্ণ তথ্যের তাৎপর্য অনুধাবনে তা সমস্যা তৈরি করে। কারণ, উপাত্তের মধ্যে প্রচুর বিভ্রান্তিকর পুনরাবৃত্তি থাকে। উপাত্ত সারণিকরণের মাধ্যমে এ সকল পুনরাবৃত্তি দূর করা সম্ভব এবং উপাত্তকে একটি যথাযথ দৃষ্টিভঙ্গিতে ব্যাখ্যা করা সম্ভব। অর্থাৎ, সারণিবদ্ধ উপাত্ত একটি সাধারণ রূপকে প্রকাশ করে, যা বর্ণনার মাধ্যমে অর্জন করা সম্ভব নয়। চতুর্থতঃ, সারণিকরণ উপাত্তের ব্যাপকতর বিশ্লেষণকে সহজতর করে। পঞ্চমতঃ, যখন একটি শিরোনামের মাধ্যমে উপাত্তকে সুসংবদ্ধভাবে সারণিতে উপস্থাপন করা হয়, তখন সারণিতে উপস্থাপিত উপাত্তের একটি স্বাতন্ত্র্য তৈরি হয়, যা সমস্যার ব্যাখ্যায় তথ্যনির্দেশক হিসাবে কাজ করে।

সারণির প্রকারভেদ (Types of Table)

সারণি গঠনের বেশ কিছু গুরুত্বপূর্ণ নীতিমালা রয়েছে। সে সকল নীতিমালার ভিত্তিতে উপস্থাপনের উদ্দেশ্য এবং উপাত্তের প্রকৃতি অনুযায়ী বিভিন্ন ধরনের সারণি প্রস্তুত করা যায়। ব্যবহারের দিক থেকে, সারণিকে দু'ভাগে ভাগ করা যায়। একটি হলো, সাধারণ বা তথ্যনির্দেশমূলক (general or reference type) সারণি এবং অন্যটি হলো, সার-সংক্ষেপমূলক (summary or text type) সারণি। সাধারণ বা তথ্যনির্দেশমূলক সারণিগুলো তথ্যের ভান্ডার হিসাবে ব্যবহৃত হয়। এ ধরনের সারণিগুলো প্রায়শঃ বিশাল আকৃতির হয়ে থাকে। এগুলো একজন পাঠককে বিস্তারিত তথ্য প্রদান করে এবং তথ্যের তাৎক্ষণিক উৎস হিসাবে কাজ করে। এ ধরনের সারণিগুলো পরিশিষ্ট বা প্রতিবেদনের শেষে আলাদাভাবে উপস্থাপন করা হয়। সার-সংক্ষেপমূলক সারণিগুলো আকারে ছোট হয়ে থাকে এবং এতে ঘনিষ্ঠভাবে সম্পর্কযুক্ত অল্প কিছু তথ্যকে যতটুকু সম্ভব কার্যকরভাবে উপস্থাপন করা হয়। সার-সংক্ষেপমূলক সারণিকে সচরাচর একটি আলোচনার সাথে সংযুক্ত করা হয়ে থাকে। সাধারণ বা তথ্যনির্দেশমূলক সারণিগুলো জটিল হয়ে থাকে এবং সার-সংক্ষেপমূলক সারণিগুলো সহজভাবে গঠিত হয়।

বৈশিষ্ট্যের ভিত্তিতে, সারণিকে তিনভাবে ভাগ করা যায়। যেমন, (১) এক-মুখী সারণি (one-way table), (২) দ্বি-মুখী সারণি (two-way table), এবং (৩) উচ্চ-মাত্রার সারণি (high-order table)। এক-মুখী সারণি মাত্র একটি বৈশিষ্ট্যকে ব্যবহার করে। সেই বৈশিষ্ট্যের ভিত্তিতে সারণিতে ঘটনসংখ্যাকে উপস্থাপন করা হয়। নিচে প্রদত্ত সারণির মাধ্যমে বিষয়টি বোঝা যেতে পারে। এই সারণিতে একটি বৈশিষ্ট্যকে ব্যবহার করা হয়েছে। সেই বৈশিষ্ট্যটি হলো, বাংলাদেশের জনসংখ্যার ঘনত্ব।

সারণি-৯.৩.১: প্রতি বর্গ-কিলোমিটারে বাংলাদেশের জনসংখ্যার ঘনত্ব, ১৯৭৪-২০০১

আদমশুমারীর বছর	জনসংখ্যার ঘনত্ব
১৯৭৪	৫১৮
১৯৮১	৬০৯
১৯৯১	৭৫৫
২০০১	৮৮১

উৎস: বাংলাদেশ আদম শুমারী ২০০১, জাতীয় প্রতিবেদন (সাময়িক), পৃষ্ঠা-২৬, বাংলাদেশ পরিসংখ্যান ব্যুরো, জুলাই ২০০৩।

দ্বি-মুখী সারণিতে, দু'টি বৈশিষ্ট্য ব্যবহার করা হয়। এ ক্ষেত্রে, স্তম্ভগুলো একটি বৈশিষ্ট্যের বিভিন্ন শ্রেণী-দফাকে প্রতিনিধিত্ব করে এবং সারিগুলো আরেকটি বৈশিষ্ট্যের শ্রেণী-দফাগুলোকে প্রতিনিধিত্ব করে। সারণি ৯.৩.২-এ গ্রাম/শহর বাসস্থানের ভিত্তিতে ৫-২৪ বছর বয়স্ক জনগোষ্ঠীর স্কুল উপস্থিতির শতকরা বিন্যাস উপস্থাপন করে একটি দ্বি-মুখী সারণির উদাহরণ দেয়া হল।

সারণি-৯.৩.২: ৫-২৪ বছর বয়স্ক জনসংখ্যার বাসস্থানভিত্তিক স্কুল উপস্থিতির শতকরা বিন্যাস, ২০০১

বয়স (বছর)	বাসস্থান		
	বাংলাদেশ	গ্রাম	শহর
৫-৯	৪৯.৬৯	৪৮.৯৪	৫২.৮৭
১০-১৪	৬৩.৯৫	৬৩.৯৬	৬৩.৯৪
১৫-১৯	৩৮.১৬	৩৭.১১	৪১.০৪
২০-২৪	১৩.১০	১০.৫০	১৯.৪৬

সূত্র: বাংলাদেশ আদম শুমারী ২০০১, জাতীয় প্রতিবেদন (সাময়িক), পৃষ্ঠা-৭৮
বাংলাদেশ পরিসংখ্যান ব্যুরো, জুলাই ২০০৩।

যখন একটি সারণিতে তিনটি, অথবা তার অধিক উপাদান বা বৈশিষ্ট্য ব্যবহার করা হয়, তখন সেই সারণিকে উচ্চ-মাত্রার সারণি বলা হয়। যদি সারণি ৯.৩.২-এ আমরা লিঙ্গ বৈশিষ্ট্যটিকে যোগ করি, তবে সেটি উচ্চমানের সারণিতে পরিণত হবে। সারণি ৯.৩.৩-এ তিনটি মাত্রায় উপাত্তকে উপস্থাপন করা হয়েছে। মাত্রাগুলো হলো বাসস্থান, বয়স ও লিঙ্গ।

সারণি -৯.৩.৩: ৫-২৪ বছর বয়স্ক জনসংখ্যার বাসস্থান ও লিঙ্গভিত্তিক স্কুল উপস্থিতির শতকরা বিন্যাস, ২০০১

বয়স (বছর)	বাসস্থান								
	বাংলাদেশ			গ্রাম			শহর		
	মোট	পুরুষ	নারী	মোট	পুরুষ	নারী	মোট	পুরুষ	নারী
৫-৯	৪৯.৬৯	৪৯.৫৬	৪৯.৪২	৪৮.৯৪	৪৮.৭৫	৪৯.১৬	৫২.৮৭	৫৩.০৯	৫২.৬৩
১০-১৪	৬৩.৯৫	৬২.৩২	৬৫.৮০	৬৩.৯৬	৬৭.৮৯	৬৬.৩২	৬৩.৯৪	৬৩.৮৪	৬৪.০৬
১৫-১৯	৩৮.১৬	৪০.৮৮	৩৫.১৫	৩৭.১১	৪০.১০	৩৩.৭৮	৪১.০৪	৪২.৯৮	৩৮.৮৯
২০-২৪	১৩.১০	২০.০৭	৭.৫১	১০.৫০	১৭.৩২	৫.৫৪	১৯.৪৬	২৫.৭০	১৩.১৪

সূত্র: বাংলাদেশ আদম শুমারী ২০০১, জাতীয় প্রতিবেদন (সাময়িক), পৃষ্ঠা-৭৮
বাংলাদেশ পরিসংখ্যান ব্যুরো, জুলাই ২০০৩।

সারণি গঠনে বিবেচ্য বিষয়সমূহ (Issues to be Considered in Table Construction)

আমরা জানি যে, একটি সারণির প্রকৃত রূপটি কি হবে, তা কিছুটা নির্ভর করে সারণিকরণের উদ্দেশ্যের উপর এবং কিছুটা নির্ভর করে উপাত্তের প্রকৃতির উপর। যে কোন সারণি গঠনের জন্য কতগুলো বিষয় বিবেচনায় রাখা প্রয়োজন।

- প্রতিটি সারণির একটি প্রধান শিরোনাম থাকতে হবে, যা সাধারণতঃ সারণির উপরে উল্লিখিত হয়। সারণির বিষয়বস্তু নির্দেশকারী শিরোনামকে হতে হবে স্পষ্ট, সংক্ষিপ্ত এবং যথার্থ।
- সারণিকে যতদূর সম্ভব সহজভাবে গঠন করতে হবে। প্রয়োজনবোধে, একটি দীর্ঘ এবং জটিল সারণির পরিবর্তে একাধিক সারণির মাধ্যমে উপাত্তকে উপস্থাপন করতে হবে।
- স্তম্ভ ও সারির শিরোনামগুলোকে সুস্পষ্ট ও যথার্থভাবে নামায়িত করতে হবে।

- উপাত্তের সুনির্দিষ্ট পরিমাপের একককে উল্লেখ করতে হবে। অর্থাৎ, বিষয়টি বছরে, টাকায়, শতাংশে, বা লিটারে পরিমাপ করা হয়েছে কি না, তা স্পষ্টভাবে উল্লেখ করতে হবে।
- সারণির নীচে পাদটীকার মাধ্যমে সারণিতে ব্যবহৃত প্রতীকগুলোর যথাযথ ব্যাখ্যা দিতে হবে।
- সারণিতে উপস্থাপিত উপাত্ত বিভিন্ন উৎস থেকে সংগৃহীত হলে, সারণির নীচে সে সকল উৎসের উল্লেখ করতে হবে।
- সারণির আকার বড় হলে, বা বেশী উপাত্তের সমাহার ঘটলে, উপাত্তের প্রধান প্রধান শ্রেণীগুলোকে এমনভাবে উপস্থাপন করতে হবে, যেন সেগুলো বৈশিষ্ট্যমণ্ডিত হয়ে উঠে। প্রয়োজনবোধে, স্তম্ভ ও সারিগুলোকে বিভক্ত করে সংখ্যা বা বর্ণমালা ব্যবহার করে চিহ্নিত করতে হবে।
- শতকরা বা সমানুপাতমূলক সংখ্যার ক্ষেত্রে, প্রকৃত ঘটনসংখ্যাটি বন্ধণীর মধ্যে উপস্থাপন করতে হবে।
- তুলনা ও কার্য-কারণ সম্পর্কের ক্ষেত্রে, স্বাধীন চলককে সারণির উপরে এবং নির্ভরশীল চলক সারণির বাম পার্শ্বে উপস্থাপন করতে হবে। এ ক্ষেত্রে, শতকরা নির্ণয় করতে হবে স্তম্ভের যোগফল ভিত্তিক এবং তুলনা করতে হবে সারিভিত্তিক। স্তম্ভভিত্তিক তুলনার ক্ষেত্রে, শতকরা নির্ণয় করতে হবে সারিভিত্তিক।
- সারণিতে শতকরা ব্যবহার করলে, তা কোন সংখ্যার সাথে যুক্ত তা নির্দিষ্ট করতে হবে। যেমন, 'মোট সংখ্যার শতকরা' (percent of total), বা 'বৃদ্ধি বা হ্রাসের শতকরা' (percent increase or decrease)। শুধুমাত্র 'শতকরা' (percent) শব্দটির ব্যবহার সবসময় অর্থপূর্ণ হয় না।

সারণিকরণ প্রক্রিয়া হাতে হাতে সম্পন্ন করা যায়। হাতে হাতে সম্পন্নকৃত সারণিকরণ খুবই সরল এবং এতে কোন কলাকৌশলগত জ্ঞান, বা দক্ষতার প্রয়োজন হয় না। এমন কি, দ্বি-চলক বিশিষ্ট আড়া-আড়ি সারণিকরণের (cross-tabulation) ক্ষেত্রেও এটি বেশী শ্রমসাধ্য নয়। যদি গবেষক আয় ও পেশার একটি বিন্যাসকে সারণিতে উপস্থাপন করতে চান, তবে একটি চলককে অনুভূমিকভাবে এবং অন্য চলকটিকে উলম্বভাবে উপস্থাপন করে, পুরো সাক্ষাৎকার অনুসূচী পর্যালোচনার মাধ্যমে প্রতিটি ঘটনসংখ্যাকে গণনা করতে হবে। প্রতিটি সারণি প্রস্তুত করার সময় একই প্রক্রিয়ার পুনরাবৃত্তি করতে হবে। কিন্তু প্রতিবার সাক্ষাৎকার অনুসূচীতে ফিরে যাওয়ার বিষয়টি বেশ অসুবিধার সৃষ্টি করে। সে ক্ষেত্রে, বিকল্প হিসাবে সারণিতে ঘটনসংখ্যা গণনার পূর্বে, প্রতিটি চলকের বিপরীতে ট্যালী চিহ্ন বসিয়ে উপাত্তকে একটি উপাত্ত স্থানান্তরপত্রে তুলে আনলে সারণিকরণের কাজটি সহজ হয়। কিন্তু হাতে হাতে সারণিকরণ খুবই ধীরগতিসম্পন্ন ও ক্লান্তিকর একটি কাজ। খুব সাধারণ ক্ষেত্র এবং ছোট নমুনার গবেষণা ছাড়া, হাতে হাতে সারণিকরণ কার্যকর পদ্ধতি নয়। নমুনার আকার বড় হলে, বা জটিল সারণি প্রস্তুত করার প্রয়োজন হলে, কম্পিউটার বা অন্য কোন যন্ত্রের মাধ্যমে সারণিকরণ করাই যুক্তিসঙ্গত।

গণসংখ্যা নিবেশন (Frequency Distribution)

গণসংখ্যা নিবেশন হলো, তথ্যরাশিকে অর্থবহভাবে সারণিকরণের একটি অন্যতম পদ্ধতি। যখন কিছু পর্যবেক্ষণকে গণনা করে সংক্ষিপ্তভাবে সাজানো হয়, তখন সেটি একটি গণসংখ্যা নিবেশনে পরিণত হয়। একটি চলক দিয়ে যে গণসংখ্যা নিবেশন তৈরি করা হয়, সেটিকে এক-চলক বিশিষ্ট গণসংখ্যা নিবেশন বলে। সুনির্দিষ্টভাবে বলা যায় যে, গণসংখ্যা নিবেশন হলো, কতগুলো পর্যবেক্ষণের মধ্যে একটি চলকের বিভিন্ন মান কতবার ঘটেছে তার সংখ্যাসূচক বিন্যাস। গণসংখ্যা নিবেশনকে দুই স্তম্ভের একটি সারণিও বলা যায়। একটি স্তম্ভ কোন ঘটনা বা বৈশিষ্ট্যকে তালিকাবদ্ধ করে এবং অন্য দুটি স্তম্ভ সে সকল বৈশিষ্ট্যের ঘটনসংখ্যাকে উল্লেখ করে।

নামসূচক বা ক্রমসূচক মাত্রায় পরিমাপকৃত চলকসমূহের ক্ষেত্রে, প্রতিটি পদ বা দফা একটি স্তম্ভে তালিকাভুক্ত করে প্রতি পদ বা দফার বিপরীতে সংশ্লিষ্ট গণসংখ্যাকে লিপিবদ্ধ করা হয়ে থাকে। ব্যাপ্তিমূলক বা অনুপাতমূলক মাত্রায় পরিমাপকৃত চলকের ক্ষেত্রে, একই পদ্ধতি অনুসরণ করা হয়ে থাকে। তবে, এ ক্ষেত্রে একটি বাড়তি কাজ করতে হয়। তা হলো, উপাত্তরাশিকে কয়েকটি নির্দিষ্ট শ্রেণীতে বিভক্ত করে, সে সকল শ্রেণীর বিপরীতে গণসংখ্যার বিন্যাস তৈরি করতে হয়। কিন্তু কয়টি শ্রেণী তৈরি করতে হবে এবং প্রতিটি শ্রেণীর ব্যাপ্তিটি কত বড় হবে, সে বিষয়ে পূর্বেই সিদ্ধান্ত গ্রহণ করা প্রয়োজন। শ্রেণী সংখ্যা নির্ধারণের বিষয়টি নির্ভর করে, কিছুটা উপাত্তের প্রকৃতির উপর এবং কিছুটা নির্ভর করে, মোট কতটি ঘটনসংখ্যা রয়েছে, তার উপর। যদি ঘটনসংখ্যার পরিমাণ বেশী হয়, তবে স্বাভাবিকভাবেই শ্রেণী সংখ্যা বেড়ে যাবে। আর যদি ঘটনসংখ্যার পরিমাণ কম হয়, তবে শ্রেণী সংখ্যা কম হবে। শ্রেণী সংখ্যা নির্ধারণের ক্ষেত্রে এটি মনে রাখা প্রয়োজন যে, যদি শ্রেণী সংখ্যা অনেক বেশী হয়, তবে শ্রেণীগুলো খুব কম সংখ্যক গণসংখ্যা ধারণ করবে এবং পুরো বিন্যাসটি একটি ভ্রান্ত ধারণা দেবে। অন্যদিকে, শ্রেণী সংখ্যা অনেক কম হলে এবং শ্রেণী ব্যাপ্তিটি অনেক বেশী প্রশস্ত হলে উপাত্তের প্রকৃত স্বরূপটি হারিয়ে যেতে পারে।

শ্রেণী সংখ্যাটি নির্ধারিত হয়ে গেলে, আমাদের জন্য শ্রেণী ব্যাপ্তি নির্ধারণ খুব সহজ হয়ে পড়ে। তবে, এমন কোন শ্রেণী সীমা, বা ব্যাপ্তি নির্ধারণ করা ঠিক হবে না, যা অপ্রাসঙ্গিক, বা অর্থহীন মনে হতে পারে। যেমন, ২০, ৩০ বা ৫০ শ্রেণীব্যাপ্তি যেমন যৌক্তিক হবে না, তেমনি ৪.১৫ বা ৫.৩ শ্রেণী ব্যাপ্তিও অর্থহীন হবে। আবার শার্টের মূল্যের ক্ষেত্রে, ৫ টাকার শ্রেণী ব্যাপ্তি উপযোগী হলেও মোটর গাড়ীর ক্ষেত্রে, তা প্রয়োজ্য নয়। কয়টি শ্রেণী গঠন করতে হবে সে সম্পর্কে নিশ্চিত হতে না পারলে, সাধারণতঃ ছোট ছোট ব্যাপ্তি নিয়ে শ্রেণীগুলো গঠন করা উচিত। যদি দেখা যায় যে, পুরো গণসংখ্যা নিবেশনের পর শ্রেণীর সংখ্যা অনেক বেশী হয়ে গিয়েছে, বা শ্রেণী ব্যাপ্তিটি খুব ছোট হয়ে গিয়েছে, সে ক্ষেত্রে, দুটি বা তিনটি শ্রেণীকে একীভূত করে শ্রেণী সংখ্যাকে কমিয়ে ব্যাপ্তি বাড়িয়ে দেয়া যেতে পারে। প্রথমেই যদি অনেক বড় ব্যাপ্তি নিয়ে শুরু করা হয়, তবে পরে ছোট করতে গেলে নিবেশন প্রক্রিয়াটি আবার প্রথম থেকে শুরু করতে হবে। যেহেতু শ্রেণী সংখ্যা বা ব্যাপ্তি নির্ধারণের বাঁধাধরা কোন নিয়ম নেই, সেহেতু গবেষণার ফলাফল বিশ্লেষণ ও ব্যাখ্যার উদ্দেশ্যের দিকে লক্ষ্য রেখেই এই সিদ্ধান্তটি গ্রহণ করতে হবে।

আমরা জানি যে, সারণিবদ্ধ উপাত্তের বিন্যাস একটি চলকের বিন্যস্ত বৈশিষ্ট্যগুলোকে বর্ণনায় সাহায্য করে এবং পরিসংখ্যান নির্ণয়ের ভিত্তি হিসেবে কাজ করে। কিন্তু দু'টি বিন্যাসের মধ্যে তুলনার ক্ষেত্রে, সাধারণভাবে প্রস্তুতকৃত গণসংখ্যা নিবেশনের বিন্যাসটি যথেষ্ট নয়। কারণ, বিভিন্ন বিন্যাস বিভিন্ন এককে প্রকাশিত হয় (যেমন, টাকা, বছর, কিলোগ্রাম, ঘন্টা, ইত্যাদি)। ফলে, ভিন্ন একক সম্বলিত দু'টি বিন্যাসের মধ্যে তুলনা করা সম্ভব হয় না। এমন কি, দু'টি বিন্যাস একই এককে প্রকাশিত হলেও মোট পর্যবেক্ষণ সংখ্যার ভিন্নতার কারণে তুলনামূলক বিশ্লেষণ সম্ভব হয় না। এ ছাড়া, পর্যবেক্ষণগুলো কোন শ্রেণীতে কি অনুপাতে বিন্যস্ত রয়েছে, অথবা একটি নির্দিষ্ট মানের উপরে বা নীচে কি পরিমাণ, বা শতকরা কতটি পর্যবেক্ষণ রয়েছে, কখনো কখনো আমাদের তা জানার প্রয়োজন পড়ে। নিরঙ্কুশ গণসংখ্যা নিবেশন-এর পরিবর্তে আপেক্ষিক গণসংখ্যা নিবেশন-এর মাধ্যমে উপাত্তকে উপস্থাপন করলে, এ সব সমস্যা দূরীভূত হয়। অনুপাত, সমানুপাত, শতকরা, হার, ইত্যাদি পরিসংখ্যানের মাধ্যমে আপেক্ষিক গণসংখ্যা নিবেশনকে উপস্থাপন করা হয়।

সারাংশ

এই পাঠে, উপাত্ত সংক্ষিপ্তকরণের সাথে সম্পর্কিত কয়েকটি বিষয়, যেমন, সারণির প্রকারভেদ, সারণি গঠনে বিবেচ্য বিষয়সমূহ এবং এর গুরুত্ব আলোচিত হয়েছে। পরিমাণগত উপাত্তকে সারণিতে উপস্থাপন না করলে, বিশ্লেষণ উপযোগী পরিবেশের সৃষ্টি হয় না। বস্তুতঃ, উপাত্ত সংগৃহীত হওয়ার পর, সারণি তৈরির মাধ্যমে আমরা উপাত্তের বৈশিষ্ট্যগত পরিচয় পাই। বর্ণনা ও ব্যাখ্যার জগতে বিভিন্ন রকম সারণির

এস এস এইচ এল

ব্যবহার দেখা যায়। আমরা কখনো এক-চলক, কখনো একের অধিক চলক, আবার কখনো সরল লেখ-
চিত্রের মাধ্যমে উপাত্তের উপস্থাপন করে থাকি। উপাত্ত কিভাবে উপস্থাপন করা হবে, তা মূলতঃ নির্ভর
করে গবেষণার উদ্দেশ্য ও বিশ্লেষণের প্রকৃতির উপর।

পাঠ্যের মূল্যায়ন

নৈর্ব্যক্তিক প্রশ্ন

সঠিক উত্তরের পাশে টিক (✓) চিহ্ন দিন –

১। উপাত্তকে সাধারণতঃ _____ উপায়ে উপস্থাপন করা যায়।

- ক. দুই
- খ. তিন
- গ. চার
- ঘ. পাঁচ

২। উপাত্ত উপস্থাপনের শ্রেষ্ঠ উপায় হলো:

- ক. আলোচনা
- খ. আধা-সারণি
- গ. লেখচিত্র
- ঘ. সারণি।

৩। একটি সারণি তখনই উচ্চ মাত্রার সারণিতে পরিণত হবে, যখন:

- ক. নমুনার সংখ্যা বেশী প্রতিনিধিত্ব হবে
- খ. গণসংখ্যার সংখ্যা বেশী থাকবে
- গ. উচ্চ মাত্রার বক্তব্য প্রতিনিধিত্ব হবে
- ঘ. তিন অথবা তার অধিক উপাদান ব্যবহৃত হবে।

সংক্ষিপ্ত প্রশ্ন

- ১। সারণিকরণ কী?
- ২। গণসংখ্যা নিবেশন কী?

রচনামূলক প্রশ্ন

- ১। উদাহরণসহ সারণির প্রকারভেদ আলোচনা করুন।
- ২। সারণি গঠনে বিবেচ্য বিষয়সমূহ আলোচনা করুন।