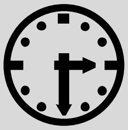
ইউনিট

৮

কল্পনা যাচাই
Test of Hypothesis

ভূমিকা

তথ্যবিশ্ব থেকে নমুনায়নের মাধ্যমে নমুনা সংগ্রহ ও বিশ্লেষণ করে তথ্যবিশ্বের প্রকৃতি সম্পর্কে ধারণা পাওয়া যায় অর্থাৎ তথ্যবিশ্বের পরামান সমূহ (গড়, প্রচুরক, ভেদাংক, ইত্যাদি) নিরূপণ করা যায়। তথ্যবিশ্বের পরামানসমূহের নিরূপিত মান এদের প্রকৃত মানের কাছাকাছি কিনা এটা পরীক্ষা করার পদ্ধতিই হচ্ছে যথার্থতা যাচাই। যথার্থতা যাচাই করার পূর্বে তথ্যবিশ্বের পরামান সম্পর্কে কল্পনা করে নিতে হয়। এরপর নমুনায়নের মাধ্যমে প্রাপ্ত তথ্যমানের ভিত্তিতে ঐ কল্পনা গৃহীত বা বাতিল হবে এ সম্পর্কে সিদ্ধান্ত নেয়ার পরিসংখ্যানিক পদ্ধতিই হল যথার্থতা যাচাই পদ্ধতি। এ ইউনিটে যথার্থতা যাচাই সম্পর্কিত কতিপয় শব্দাবলীর সংগে, যথার্থতা যাচাইয়ের গুরুত্বপূর্ণ ধাপ, গড়, ভেদাংক, অনুপাত, ইত্যাদির যথার্থতা যাচাই পদ্ধতি আলোচনা করা হবে।



ইউনিট সমাপ্তির সময়	ইউনিট সমাপ্তির সর্বোচ্চ সময় ০২ সপ্তাহ

এ ইউনিটের পাঠসমূহ

পাঠ ৮.১	ঃ	কল্পনা যাচাইয়ের সংজ্ঞা ও এর কতিপয় ধারণা
পাঠ ৮.২	ঃ	গড় সম্পর্কিত কল্পনা যাচাই
পাঠ ৮.৩	ঃ	অনুপাত সম্পর্কিত কল্পনা যাচাই
পাঠ ৮.৪	ঃ	সংশ্লেষাংক সম্পর্কিত কল্পনা যাচাই
পাঠ ৮.৫	ঃ	ভেদাংক সম্পর্কিত কল্পনা যাচাই

পাঠ ৮.১ কল্পনা যাচাইয়ের সংজ্ঞা এবং এর কতিপয় ধারণা
Definition and Some Concepts Related to Hypothesis Test

উদ্দেশ্য

এ পাঠ শেষে আপনি-

- কল্পনা যাচাই সম্পর্কিত কতিপয় বিষয়ের সংজ্ঞা লিখতে পারবেন।
- কল্পনা যাচাইয়ের ধাপসমূহ বর্ণনা করতে পারবেন।
- কল্পনা যাচাই সম্পর্কিত গুরুত্বপূর্ণ যাচাই ব্যাখ্যা করতে পারবেন।

কল্পনা যাচাইয়ের সংজ্ঞা
Definition of Hypothesis Test

সমগ্রক সম্বন্ধে যে কোন বিবৃতিকে বলা হয় কল্পনা (Hypothesis) সাধারণ অর্থে কল্পনা হলো পূর্ব অনুমান। অর্থাৎ গবেষণার শুরুতে গবেষণার বিষয় সম্পর্কে যে পূর্ব ধারণা পোষণ করা হয়, তাই হল কল্পনা (Hypothesis) অনুমান বা কল্পনা হলো এক ধরনের সিদ্ধান্ত যা যাচাই সাপেক্ষ। সমগ্রকের নিরুপিত মানের ওপর ভিত্তি করে পরামান সম্পর্কে কোন ব্যাখ্যা, মন্তব্য বা সিদ্ধান্ত যাচাই গ্রহণ বা বর্জন করতে হয়। এই ব্যাখ্যা, মন্তব্য বা সিদ্ধান্তকে কল্পনা বলা হয়।

পরামিতিক কল্পনা যাচাই (Parametiric Hypothesis Test) ঃ কোন সমগ্রকের বিন্যাসের পরামিতি সম্পর্কে যে কোন কল্পনাকে বলা হয় পরামিতিক কল্পনা (Paramtric hypotheis)। যেমনঃ ধরা যাক, কোন কলেজের সমস্ত ছাত্রছাত্রীদের বয়সের গড় 20 বৎসর। গড় সম্পর্কে এ ধরনের কল্পনাকে পরামিতিক কল্পনা (Parametric hypothesis) বলা হয়।

অপরামিতিক কল্পনা যাচাই (Non Parametric Hypothesis Test) ঃ কোন সমগ্রকের বিন্যাস সম্পর্কে যে কোন কল্পনাকে বলা হয় অপরামিতিক কল্পনা। যেমনঃ একজন কোম্পানির মালিক দাবি করলেন যে, তার উৎপাদিক দ্রব্যের ওজন পরিমিত বিন্যাস মেনে চলে। তবে উৎপাদিত দ্রব্যের ওজনের বিন্যাস সম্পর্কে কল্পনাকে বলা হয় অপরামিতিক কল্পনা (Non parametric hypothesis)

কল্পনা যাচাই সম্পর্কিত কতিপয় ধারণা
Some Concepts Related to Hypothesis Test

যথার্থতা যাচায়ে যাবার আগে যথার্থতা যাচাইয়ের সাথে সম্পর্কিত কতিপয় বিষয় সম্বন্ধে জ্ঞান থাকা আবশ্যক। এ জাতীয় কিছু বিষয় নিয়ে নিচে আলোচনা করা হলো–

১. নমুনার আকার (Sample size) ঃ কোন সমগ্রক হতে নির্বাচিত ও গৃহীত নমুনায় নমুনা উপাদান (Sample unit) এর মোট সংখ্যাকে ঐ নমুনার নমুনা আকার বলা হয়। নমুনা আকার যথার্থতা যাচাইয়ের একটি তাৎপর্যপূর্ণ বিষয়, কারণ নমুনা আকার ছোট ($n < 30$) বা বড় ($n \geq 30$) হওয়ার কারণে যথার্থতা যাচায়ে পদ্ধতি ভিন্ন হয়।
এবার আমরা যথার্থতা যাচাইয়ের জন্য প্রয়োজনীয় ও আবশ্যক কিছু বিষয় নিয়ে আলোচনা করব–

২. পরিসংখ্যানিক কল্পনা (Statistical Hypothesis) ঃ পরিসংখ্যানিক কল্পনা হলো দৈব চলকের সম্ভাবনা বিন্যাস সম্বন্ধে একটি অনুমিত বিবৃত (statement)। পরিসংখ্যানিক সিদ্ধান্ত গ্রহণের জন্য সমগ্রক বা পরামিতি সম্পর্কে কল্পনা করতে হয়। নমুনা তথ্যের ভিত্তিতে এ কল্পনা গ্রহণ বা বাতিলের সিদ্ধান্ত গ্রহণ করা হয়। এ কল্পনাকেই পরিসংখ্যানিক কল্পনা বলে।
পরিসংখ্যানিক কল্পনা দুই ধরনের হয়ে থাকে–

(i) নাস্তি কল্পনা (Null Hypothesis) ঃ যে পরিসংখ্যানিক কল্পনায় ধরে বা অনুমান করে নেয়া হয় যে সব কিছু সাম্যবস্থায় আছে বা একাধিক বস্তু বা ঘটনার মধ্যে কোনরকম ব্যবধান নেই (অর্থাৎ তারা সমান) তবে তাকে নাস্তি কল্পনা বলা হয়। নাস্তি কল্পনাকে H_0 দ্বারা প্রকাশ করা হয়।

(ii) বিকল্প কল্পনা (Alternative Hypothesis) ঃ যে পরিসংখ্যানিক কল্পনা নাস্তি কল্পনার বিপরীত ধারণা ব্যক্ত করে তাকে বিকল্প কল্পনা বলে। যেমন– $(H_0) : \mu = 0$ হলে বিকল্প কল্পনা (H_1): $\mu \neq 0$ অথবা $\mu > 0$ বা $\mu < 0$ হবে। বিকল্প কল্পনাকে H_1 বা H_A দ্বারা প্রকাশ করা হয়।

৩. প্রথম প্রকার ভুল ও দ্বিতীয় প্রকার ভুল (Type I error and Tyep II error)

প্রথম ধরনের ভুল (Type I Error) ঃ কোন যথার্থতা যাচায়ে নাস্তি কল্পনা (H_0) সত্য হওয়ার পরও বাতিল বা নাকোচ করলে যে ভুলের উদ্ভব হয় তাকে প্রথম ধরনের ভুল বলে। প্রথম ধরনের ভুলের সম্ভাবনাকে α দ্বারা প্রকাশ করা হয়। α–কে যথার্থতা মাত্রা (Level of Significance) ও বলা হয়।

গাণিতিকভাবে, P (প্রথম ধরনের ভুল) = P (H_0 বাতিল । H_0 সত্য) = α

দ্বিতীয় ধরনের ভুল (Type II Error) ঃ কোন যথার্থতা যাচায়ে নাস্তি কল্পনা (H_0) সত্য না হলেও তা গ্রহণ করলে যে ভুলের উদ্ভব হয় তাকে দ্বিতীয় ধরনের ভুল বলে। দ্বিতীয় ধরনের ভুলের সম্ভাবনাকে β (বিটা) দ্বারা প্রকাশ করা হয়।

গাণিতিকভাবে, P (দ্বিতীয় ধরনের ভুল) = P (H_0 গৃহীত । H_0 সত্য নয়) = β

প্রথম এবং দ্বিতীয় ধরনের ভুলকে নিচের সারণীর মাধ্যমে সহজে ব্যক্ত করা যায়–

প্রকৃতির ধরণ	সিদ্ধান্ত	
	(H_0) গ্রহণ	(H_0) বর্জন
H_0 সত্য	সঠিক সিদ্ধান্ত	প্রথম ধরনের ভুল (Type I Error)
H_0 সত্য নয় (অর্থাৎ H_1 সত্য)	দ্বিতীয় ধরনের ভুল (Type I Error)	সঠিক সিদ্ধান্ত

৪. যথার্থতা যাচাই (Test of Significance) ঃ যথার্থতা যাচাই এমন একটি পরিসংখ্যানিক পদ্ধতি যেখানে নমুনা হতে প্রাপ্ত তথ্যের বিশ্লেষণ করে তথ্যবিশ্বের পরমান বা তথ্যবিশ্ব সম্বন্ধে সিদ্ধান্ত নেয়া হয়। এই পদ্ধতিতে আমরা কোন কল্পনাকে সঠিকভাবে ভুল বা সত্য প্রমাণ করি না। প্রাপ্ত নমুনা হতে সম্ভাবনার ভিত্তিতে আমরা কল্পনা সম্বন্ধে সিদ্ধান্ত নিয়ে থাকি। তাই এই পদ্ধতিতে আমরা কিছু ভুল স্বীকার করে সিদ্ধান্ত নিয়ে থাকি। যেমনঃ ধরা যাক, ঢাকা সিটি কলেজের কোন শ্রেণিতে 540 ছাত্র-ছাত্রী আছে। তাহলে তথ্য বিশ্বে 540 টি উপাদান বা একক আছে। মনে করি এদের গড় বয়স 22 বৎসর। এখন এ তথ্যবিশ্ব হতে 12 আকার বিশিষ্ট নমুনা দৈবক্রমে নির্বাচন করা হল, যাদের গড় বয়স 18 বৎসর পাওয়া গেল। যথার্থতা যাচাই করে দেখা যায় যে ঐ শ্রেণির ছাত্র-ছাত্রীদের বয়স 18 বৎসর হওয়ার সম্ভাবনা খুবই কম। তাই নমুনার প্রাপ্ত মানগুলোর ভিত্তিতে কল্পনাটি গ্রহণযোগ্য হবে না। অথচ উক্ত তথ্যবিশ্ব থেকে চয়ন করা হয়েছে। তাই এ পদ্ধতিতে আমরা সঠিক সিদ্ধান্ত নিতে পারি না। তবে সিদ্ধান্ত নেওয়ার সম্ভাবনা যত বেশি হবে কল্পনাটি তত বেশি সঠিক হবে।

৫. যথার্থতা মাত্রা বা যথার্থতার সীমা, যাচাইয়ের শক্তি এবং যাচাই নমুনাজমান (Level of significance, power of test and test statistic)

যথার্থতার মাত্রা (Level of Significance)ঃ

কোন যথার্থতা যাচায়ে প্রথম ধরনের ভুল করার সর্বোচ্চ সম্ভাবনাকে যথার্থতার মাত্রা বলা হয়। কোন যাচায়ে যথার্থতার মাত্রা যদি উল্লেখ করা না থাকে তবে এর মান 0.05 ধরা হয়। যথার্থতার মাত্রাকে α (আলফা) দ্বারা প্রকাশ করা হয়। গাণিতিকভাবে α = P (H_0 বাতিল । H_0 সত্য)। কোন যাচাইয়ে যথার্থতার মাত্রা উল্লেখ না থাকলে $\alpha = 0.05$ হয়। $\alpha = 0.05$ হলে প্রথম ধরনের ভুল করার সম্ভাবনা 0.05 বা 5%; অর্থাৎ গৃহীত সিদ্ধান্তের যথার্থতা সম্পর্কে 0.95 বা 95% নিশ্চিত।

যাচাইয়ের শক্তি (Power of a test)ঃ

নাস্তি কল্পনা প্রকৃতপক্ষে সত্য নয়, কিন্তু নমুনা তথ্যের ভিত্তিতে তা বাতিল হওয়ার সম্ভাবনাকে যাচাইয়ের শক্তি বলা হয়। কোন যথার্থতা যাচাইয়ে দ্বিতীয় ধরনের ভুল করার সম্ভাবনা অর্থাৎ নাস্তি কল্পনা সত্য না হওয়া সত্ত্বেও তা গৃহীত হওয়ার সম্ভাবনা β হলে $1-\beta$ কে যাচাইয়ের শক্তি বলা হয়।

গাণিতিকভাবে, P (H_0 সত্য নয় | H_1 সত্য) = β হলে যাচাইয়ের শক্তি = P (H_0 বাতিল | H_1 সত্য) = $1-\beta$

যাচাই নমুনাজমান (Test Statistic)ঃ

কোন নমুনার অন্তর্গত উপাদান সমূহের যে কোন গাণিতিক ক্রিয়াকান্ড বা ফাংশন (Mathematical function)-কে ঐ নমুনার নমুনাজমান বলা হয়। ধরা যাক, $(X_1, X_2, \ldots\ldots\ldots, X_N)$ সমগ্রক হতে গৃহীত $(x_1, x_2, \ldots\ldots\ldots, x_n)$ একটি নমুনা, তাহলে নমুনা $(x_1, x_2, \ldots\ldots\ldots, x_n)$ এর যে কোন ফাংশন, যথা ঃ $T = f(x_1, x_2, \ldots\ldots\ldots, x_n)$ বা $\sum_{i=1}^{n} x_i$ বা, $\bar{x} = \frac{\Sigma x_i}{n}$ ইত্যাদি $(x_1, x_2, \ldots\ldots\ldots, x_n)$ নমুনাটির নমুনাজমান হবে। নমুনাজমানের সাহায্যে সাধারণত পরামিতির নিরূপিত (estimated) মান নির্ণয় করা হয়।

কোন যথার্থতা যাচায়ে যাচাই নমুনাজমান এমন একটি দৈব চলক যার মান নমুনা থেকে নির্ণয় করা হয় এবং যেটি নাস্তি কল্পনা গ্রহণ বা বর্জনে ব্যবহৃত হয়। নাস্তি কল্পনা গ্রহণ করার যোগ্যতা যাচাইয়ের জন্য নমুনা তথ্যে ভিত্তিতে একটি উপযুক্ত নমুনাজমান গঠন করা হয়। অতএব এরূপ নমুনাজমানের উপর ভিত্তি করে নাস্তি কল্পনা গ্রহণ বা বাতিল করার জন্য পরিসংখ্যানিক সিদ্ধান্ত গ্রহণ করা হয়। এ ধরনের নমুনাজমানকে যাচাই নমুনাজমান বলে।

৬. গ্রহণ অঞ্চল বা আস্থা এলাকা এবং বর্জন অঞ্চল বা সংশয় এলাকা

(Acceptance Region and Rejection Region/Critical Region)

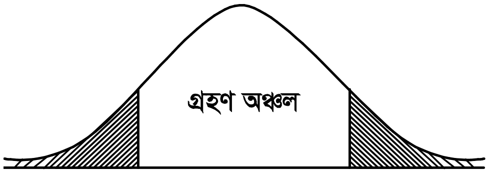
গ্রহণ অঞ্চল (Acceptance Region)

কোন যথার্থতা যাচাইয়ে নাস্তি কল্পনা (H_0) সত্য হলে এবং যাচাই নমুনাজমানের যে সকল মানের জন্য নাস্তি কল্পনা গ্রহণ বলে বিবেচিত হয়, সেই সকল মানের সমন্বয়ে গঠিত অঞ্চলকে গ্রহণ অঞ্চল বলা হয়।

ডান প্রান্তিক যাচাইয়ে গ্রহণ অঞ্চল বিন্যাসের বামদিকে অবস্থান করে, যা নিম্নে নিত্রের সাহায্যে দেখানো হলঃ

বাম প্রান্তিক যাচাইয়ে গ্রহণ অঞ্চল বিন্যাসের ডানদিকে অবস্থান করে, যা নিম্নরূপঃ

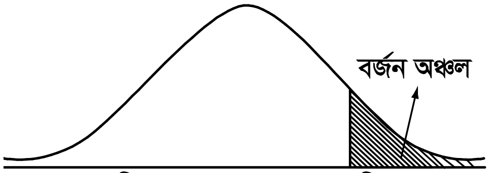
দ্বিপ্রান্তিক যাচাইয়ে গ্রহণ অঞ্চল বিন্যাসের মধ্যবর্তী স্থানে নির্দেশ করে, যা নিম্নরূপঃ



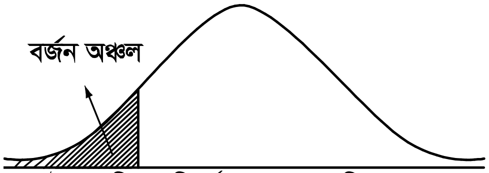
বর্জন অঞ্চল বা সংশয় এলাকা (Rejection Region/Critical Region)ঃ

কোন যথার্থতা যাচাইয়ে নাস্তি কল্পনা (H_0) সত্য হলে এবং যাচাই নমুনাজমানের যে সকল মানের জন্য নাস্তি কল্পনা বাতিল বলে বিবেচিত হয়, সেই সকল মানের সমন্বয়ে গঠিত অঞ্চলকে বর্জন অঞ্চল বলা হয়।

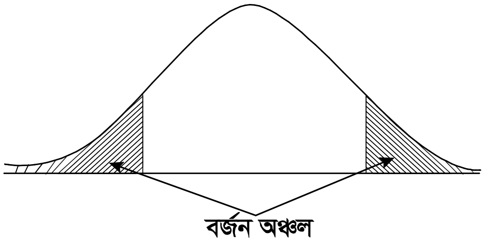
ডান প্রান্তিক যাচাইয়ে বর্জন অঞ্চল বিন্যাসের ডানদিকে অবস্থান করে। যা নিম্নে চিত্রের সাহায্যে বর্জন এলাকা দেখানো হলঃ



বাম প্রান্তিক যাচাইয়ে বর্জন অঞ্চল বিন্যাসের বামদিকে অবস্থান করে, যা নিম্নরূপঃ



দ্বিপ্রান্তিক যাচাইয়ে বর্জন অঞ্চল বিন্যাসের উভয় দিকে নির্দেশ করে, যা নিম্নরূপঃ



৭. এক প্রান্তিক যাচাই ও দ্বি-প্রান্তিক যাচাই (One tailed and two tailed test)

এক প্রান্তিক যাচাই (One Sided Test or One Tailed Test)ঃ

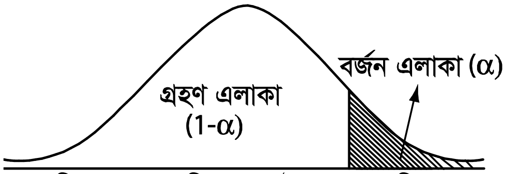
কোন যথার্থতা যাচাযে বর্জনক্ষেত্র যদি নমুনাজ বিন্যাসের যে কোন একপার্শে, অর্থাৎ ডানপাশে বা বামপাশে থাকে তবে তাকে এক প্রান্তিক যাচাই বলে। এক প্রান্তিক যাচাই পরীক্ষা দু'ধরনের হতে পারে। যথা:

(i) ডান প্রান্তিক যাচাই (Right or upper tailed test)

(ii) বাম প্রান্তিক যাচাই (Left or lower tailed test)

(i) ডান প্রান্তিক যাচাই (Right or upper tailed test) ঃ

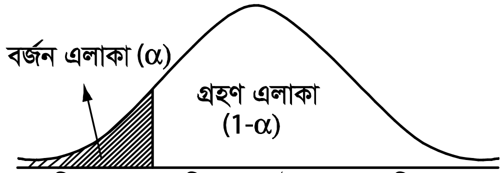
যদি বর্জন এলাকা নমুনা বিন্যাসের ডান প্রান্ত অর্থাৎ উর্ধ্ব প্রান্তে অবস্থিত হয় তবে এরূপ যাচাইকে ডান প্রান্তিক যাচাই (Right tailed test) বলা হয। উদারহণস্বরূপ, কোন যথার্থতা যাচাইয়ে নাস্তি কল্পনা H_0: $\mu = \mu_0$ এবং বিকল্প কল্পনা H_1: $\mu > \mu_0$ হয়, তবে বর্জন এলাকা বিন্যাসের ডানদিকে পড়বে। নিম্নে চিত্রের মাধ্যমে বর্জন এলাকা দেখানো হলঃ



চিত্র: এক প্রান্তিক যাচাই (ডান প্রান্তিক)

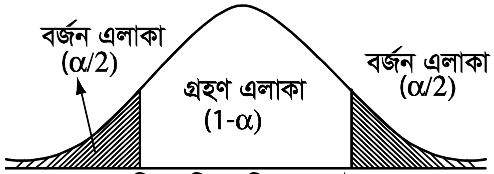
(ii) বাম প্রান্তিক যাচাই (Left or lower tailed test)

যদি বর্জন এলাকা নমুনা বিন্যাসের বাম প্রান্তে অর্থাৎ নিম্ন প্রান্তে অবস্থিত হয় তবে এরূপ যাচাইকে বাম প্রান্তিক যাচাই (Left tailed test) বলা হয। উদাহরণস্বরূপ, কোন যথার্থতা যাচাইয়ে নাস্তি কল্পনা H_0: $\mu = \mu_0$ এবং বিকল্প কল্পনা H_1: $\mu < \mu_0$ হয়, তবে বর্জন এলাকা বিন্যাসের বামদিকে পড়বে। নিম্নে চিত্রের মাধ্যমে বর্জন এলাকা দেখানো হলঃ



চিত্র: এক প্রান্তিক যাচাই (বাম প্রান্তিক)

দ্বি-প্রান্তিক যাচাই (Two Sided or Two Tailed Test) ঃ
কোন যথার্থতা যাচায়ে বর্জন ক্ষেত্র যদি নমুনাজ বিন্যাসের উভয় পার্শেই থাকে, তবে তাকে দ্বি-প্রান্তিক বা দ্বি-প্রান্তিক যাচাই বলে। উদারহণ, কোন যথার্থতা যাচাইয়ে নাস্তি কল্পনা H_0: $\mu = \mu_0$ এবং বিকল্প কল্পনা H_1: $\mu \neq \mu_0$ হয়, তবে বর্জন এলাকা বিন্যাসের উভয় প্রান্তে পড়বে। নিম্নে চিত্রের মাধ্যমে বর্জন এলাকা দেখানো হলঃ



চিত্র: দ্বি-প্রান্তিক যাচাই

৮. স্বাধীনতার মাত্রা (Degree of freedom)
নমুনাজ মান নির্ণয়ের ক্ষেত্রে যে কয়েকটি মান স্বাধীনভাবে নেয়া যায় তাদের সংখ্যাকে স্বাধীনতার মাত্রা বলে। যেমন– ধরা যাক, আমরা পাঁচটি সংখ্যা নিতে চাই যাদের যোগফল হবে 120-এক্ষেত্রে 4 টি মান আমরা ইচ্ছা মতো নিতে পারি কিন্তু পঞ্চম মান স্বাধীনভাবে নেয়া সম্ভব নয়। এখানে যে প্রতিবন্ধকতা (Restriction) রয়েছে তা হলো – সংখ্যা পাঁচটির যোগফল হবে 120,

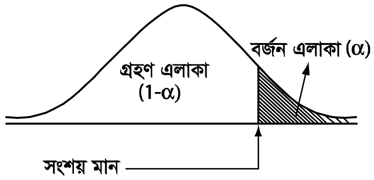
$\therefore$ স্বাধীনতার মাত্রা, $v = n - r = 5 - 1 = 4$

যেখানে, r = প্রতিবন্ধকতার সংখ্যা (Number of constraints) = 1

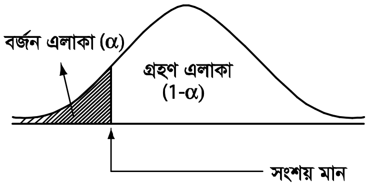
n = কার্যরত সংখ্যা = 5

৯. p-মান (P-value) ঃ যথার্থতা যাচায়ে p-মান হলো শূন্য বা নাস্তি কল্পনা সত্য হবার সর্বোচ্চ সম্ভাবনা। এটি হল প্রথম ধরনের ভুল সংগঠিত হওয়ার প্রকৃত ঝুঁকি। যখন p-মান যথার্থতার মাত্রা (α) থেকে ছোট হয় তখন নাস্তি কল্পনাকে বর্জন করা হয়। সাধারণত, p-মান 0.05 বা 0.01 থেকে ছোট হলে নাস্তি কল্পনা বর্জিত হয়।
ব্যবহার : যথার্থতা যাচাইয়ের সিদ্ধান্ত গ্রহণে, অর্থাৎ, নাস্তি কল্পনা গ্রহণ বা বর্জন করতে p-মান ব্যবহৃত হয়।

১০. সংশয় মান (Critical value) ঃ কল্পনা (Hypothesis) যাচাই করতে যাচাই নমুনাজমান ব্যবহৃত হয়। কল্পনা আবার দুই ধরনের হয়। যথা: (i) নাস্তি কল্পনা (H_0) ও (ii) বিকল্প কল্পনা (H_1)। এখন নাস্তি কল্পনা (H_0) এর ভিত্তিতে যাচাই নমুনাজমান নির্ধারিত হয়। আর এ ধরনের নমুনাজমানের বিন্যাসকে নমুনাজমান বিন্যাস বলা হয়। এ নমুনাজমান বিন্যাসের রেখাকে যে মান দ্বারা গ্রহণ এলাকা এবং বর্জন এলাকায় বিভক্ত করা হয়, তাকে সংশয় মান (Critical value) বলা হয়।
উদাহরণস্বরূপ ঃ (i) এক প্রান্তিক যাচাইয়ের ক্ষেত্রে,

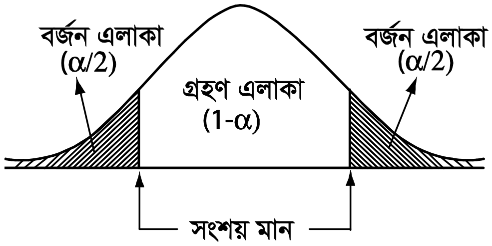


চিত্র: এক প্রান্তিক যাচাই (ডান প্রান্তিক)



চিত্র: এক প্রান্তিক যাচাই (বাম প্রান্তিক)

(ii) দ্বি-প্রান্তিক যাচাইয়ের ক্ষেত্রে,



চিত্র: দ্বি-প্রান্তিক যাচাই

যথার্থতা যাচায়ের ধাপসমূহ

Test of Significance Approach

কোন পরিসংখ্যানিক অনুমান সমূহ পরীক্ষার সাহায্যে সিদ্ধান্ত গ্রহণের জন্য প্রয়োজনীয় ধাপসমূহ নিম্নে বর্ণিত হল–

১ম ধাপ-অনুমান নির্ণয় : পরিসংখ্যানিক পরীক্ষা করতে প্রথমে নাস্তি অনুমান ও বিকল্প অনুমান নির্ধারণ করতে হবে। যেমন : কর্মচারীদের গড় আয় 3000 টাকা কিনা পরীক্ষা করতে বলা হলে–

নাস্তি অনুমান, $H_0 : \mu = 3000$ অর্থাৎ কর্মচারীদের গড় আয় 3000।

বিকল্প অনুমান, $H_A : \mu \neq 3000$ অর্থাৎ কর্মচারীদের গড় আয় 3000 টাকা নহে।

২য় ধাপ– Test Statistic নির্ণয়: নমুনা হতে প্রাপ্ত মানের ভিত্তিতে test Statistic কি হবে তা নির্ধারণ করতে হবে। যেমন σ জানা আছে, নমুনার আকার বড় এবং সমগ্রক পরিমিতিভাবে বিন্যাস্ত হলে test Statistic হবে Z। **৩য় ধাপ–Test Statistic এর মান নির্ণয় :** নমুনা হতে প্রাপ্ত মানের ভিত্তিতে test Statistic এর মান নির্ণয় করতে হবে। যেমন: z এর মান নির্ণয় করতে হবে।

৪র্থ ধাপ– Test Statistic এর Critical মান নির্ণয়: গুরুত্বের স্তরের সাহায্যে পরিসংখ্যানিক টেবিল হতে Test Statistic এর তাত্ত্বিক (Critical) মান নির্ণয় করতে হবে। যেমন : z এর Critical মান নির্ণয় করতে হবে।

৫ম ধাপ- সিদ্ধান্ত : নমুনা Statistic এর নির্ণীত মান এবং Critical মানের তুলনার মাধ্যমে নাস্তি অনুমান গ্রহণীয় হবে নাকি বর্জনীয় হবে তা নির্ধারণ করতে হবে। যেমন: যদি Test Statistic এর ধনাত্মক নির্ণীত মান উহার ধনাত্মক Critical মান অপেক্ষা বড় হয় তাহলে H_0 বর্জনীয় হবে। অন্যথায় H_0 গ্রহণীয় হবে।

কতিপয় গুরুত্বপূর্ণ কল্পনা যাচাই

Some Important Test of Hypothsis

পরিসংখ্যানে বহুল ব্যবহৃত এবং গুরুত্বপূর্ণ যাচাইগুলো হচ্ছে

১. গড় সম্পর্কিত কল্পনা যাচাই (Test of Hypothesis about Mean) :
২. অনুপাত সম্পর্কিত কল্পনা যাচাই (Test of Hypothesis about Proportion) :
৩. সংশ্লেষাংক সম্পর্কিত কল্পনা যাচাই (Test of Hypothesis about Coefficient of Correlation)
৪. ভেদাংক সম্পর্কিত কল্পনা যাচাই (Test of Hypothesis about Variance)

সারসংক্ষেপ

কোন যথার্থতা যাচায়ে প্রথম ধরনের ভুল করার সর্বোচ্চ সম্ভাবনাকে যথার্থতার মাত্রা বলা হয়। এছাড়া নাস্তি কল্পনা প্রকৃতপক্ষে সত্য নয়, কিন্তু নমুনা তথ্যের ভিত্তিতে তা বাতিল হওয়ার সম্ভাবনাকে যাচাইয়ের শক্তি বলা হয়। নমুনাজমানের উপর ভিত্তি করে নাস্তি কল্পনা গ্রহণ বা বাতিল করার জন্য পরিসংখ্যানিক সিদ্ধান্ত গ্রহণ করা হয়। এ ধরনের নমুনাজমানকে যাচাই নমুনাজমান বলে।

পাঠ ৮.২ গড় সম্পর্কিত কল্পনা যাচাই
Test of Hypothesis about Mean

উদ্দেশ্য

এ পাঠ শেষে আপনি-

- গড়ের উপর একটি নমুনার পরীক্ষা পদ্ধতি বর্ণনা করতে পারবেন।
- গড়ের উপর দুটি নমুনার পরীক্ষা ব্যাখ্যা করতে পারবেন।

গড় সম্পর্কিত কল্পনা যাচাই

Test of Hypothesis about Mean

গড় সম্পর্কিত কল্পনা যাচাই নিম্নোক্তভাবে করা যেতে পারে :

ক. গড়ের উপর একটি নমুনার পরীক্ষা (Test for single population mean)

খ. গড়ের উপর দুটি নমুনার পরীক্ষা (Test for two population mean)

ক. গড়ের উপর একটি নমুনার পরীক্ষা (Test for single population mean)

গড়ের উপর একটি নমুনার পরীক্ষা নিম্নোক্তভাবে করা যেতে পারে:

i. পরিমিত ব্যবধান (σ বা S) জানা নেই এবং নমুনার আকার ছোট (n<30) (t Test)

ii. পরিমিত ব্যবধান (σ বা S) জানা আছে এবং নমুনার আকার ছোট (n<30) (Z Test)

iii. পরিমিত ব্যবধান (σ বা S) জানা নেই এবং নমুনার আকার বড় (n>30) (Z Test)

iv. পরিমিত ব্যবধান (σ বা S) জানা আছে এবং নমুনার আকার বড় (n>30) (Z Test)

i. পরিমিত ব্যবধান (σ বা S) জানা নেই এবং নমুনার আকার ছোট (n<30) (t Test)

পরিমিত ব্যবধান (σ বা S) জানা নেই এবং নমুনার আকার ছোট (n<30) এবং সমগ্রক পরিমিতভাবে বিন্যস্ত হলে Test Statistic হবে:

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

$\bar{X}$ = নমুনার গড়

μ = সমগ্রকের গড়

n = নমুনার আকার

এক্ষেত্রে, নমুনার পরিমিত ব্যবধান, $S = \frac{\sqrt{\Sigma X^2 - \frac{(\Sigma X)^2}{n}}}{n-1}$

উদাহরণ : কোন প্রতিষ্ঠানের ম্যানেজার মনে করেন গড় উৎপাদন 8000 ইউনিটের কম। দৈবচয়ন ভিত্তিতে 25 দিনের উৎপাদনকে নির্বাচন করে দেখা গেল যে, গড় উৎপাদন 7750 ইউনিট এবং পরিমিত ব্যবধান 345 ইউনিট। ম্যানেজারের দাবী সত্য কিনা 5% গুরুত্বের স্তরে পরীক্ষা করুন।

সমাধান : দেয়া আছে, নমুনার আকার, n = 25

নমুনার গড়, $\bar{X}$ = 7750

নমুনার পরিমিত ব্যবধান, S = 345

প্রথম ধাপ ঃ প্রতিপাদ্য বা কল্পনা প্রণয়ন ঃ

নাস্তি অনুমান, $H_0 : \mu = 8000$ (অর্থাৎ ম্যানেজারের দাবী সত্য নয়)

বিকল্প অনুমান, $H_A : \mu < 8000$ (অর্থাৎ ম্যানেজারের দাবী সত্য)

দ্বিতীয় ধাপ ঃ t এ মান নির্ণয়ঃ

যেহেতু নমুনার আকার ছোট এবং পরিমিতভাবে বিন্যস্ত সমগ্রকের পরিমিত ব্যবধান (σ) জানা নেই,
তাই Test Statistic হবে,

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

$$= \frac{7750 - 8000}{\frac{345}{\sqrt{25}}}$$

$$= -3.62$$

∴ t-এর নির্ণীত মান = – 3.62

তৃতীয় ধাপ ঃ ঃ t এর তাত্ত্বিক মান নির্ণয়ঃ

যেহেতু বিকল্প $H_A : \mu < 8000$ অর্থাৎ 8000 এর কম (বাম প্রান্তিক)

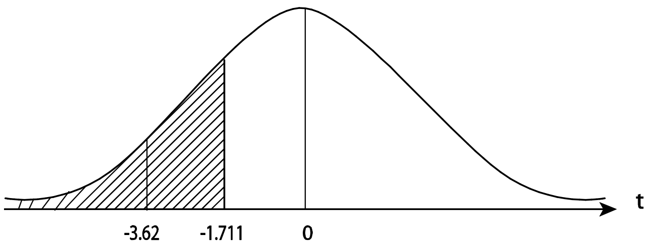
তাই এক্ষেত্রে এক পার্শ্বিক পরীক্ষা (One tailed test) প্রযোজ্য হবে।

এখানে গুরুত্বের স্তর $\alpha = 5\% = 0.05$

স্বাধীনতার মাত্রা $v = n - 1 = 25 - 1 = 24$

∴ t এর critical মান = - 1.711

চতুর্থ ধাপ ঃ সিদ্ধান্ত গ্রহণঃ



যেহেতু t এর নির্ণীত মান (-3.62) উহার তাত্ত্বিক মান (-1.711) এর চাইতে ছোট। তাই নাস্তি কল্পনা H_0 বর্জনীয়। অতএব, ম্যানেজারের দাবী সত্য।

উদাহরণ : কোন প্রতিষ্ঠানের নির্বাচিত 10 জন শ্রমিকের সাপ্তাহিক মজুরি হচ্ছে ঃ
578, 572, 570, 568, 572, 578, 570, 572, 596, 584
শ্রমিকের গড় মজুরি 580 টাকা কিনা 1% তাৎপর্য স্তরে পরীক্ষা করুন।

সমাধানঃ

প্রথম ধাপ ঃ প্রতিপাদ্য প্রণয়নঃ

নাস্তি অনুমান, $H_0 : \mu = 580$ (অর্থাৎ শ্রমিকদের গড় মজুরি 580 টাকা)

বিকল্প অনুমান, $H_A : \mu \neq 580$ (অর্থাৎ শ্রমিকদের গড় মজুরি 580 টাকা নয়)

দ্বিতীয় ধাপ ঃ t এ মান নির্ণয়ঃ

যেহেতু নমুনার আকার ছোট এবং পরিমিতভাবে বিন্যস্ত সমগ্রকের পরিমিত ব্যবধান (σ) জানা নেই,

তাই এক্ষেত্রে, Test Statistic হবে, $t = \dfrac{\bar{x}-\mu}{\dfrac{s}{\sqrt{n}}}$

$\bar{x}$ = নমুনার গড়
n = নমুনার আকার
S = নমুনার পরিমিত ব্যবধান।

নমুনার গড় ও পরিমিত ব্যবধান নির্ণয়ঃ

মজুরি (x)	x^2
578	334084
572	327184
570	324900
568	322624
572	327184
578	334084
570	324900
572	327184
596	355216
584	341056
$\Sigma x = 5760$	$\Sigma x^2 = 3318416$

নমুনার গড়, $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{5760}{20} = 576$

নমুনার পরিমিত ব্যবধান, $S = \dfrac{\sqrt{\Sigma x^2 - \dfrac{(\Sigma x)^2}{n}}}{n-1}$

$= \dfrac{\sqrt{3318416 - \dfrac{(5760)^2}{10}}}{10-1}$

$= \dfrac{\sqrt{656}}{9} = 8.54$

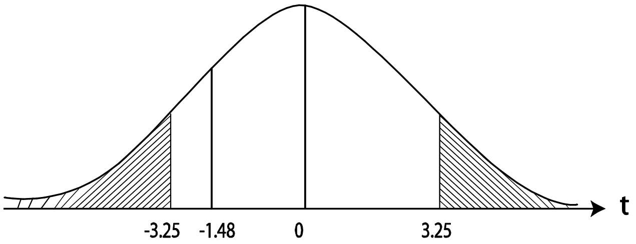
$\therefore$ t- $= \dfrac{\bar{x}-\mu}{\dfrac{s}{\sqrt{n}}} = \dfrac{576-580}{\dfrac{8.56}{\sqrt{10}}} = \dfrac{-4}{2.70} = 1.48$

তৃতীয় ধাপ ঃ t এর তাত্ত্বিক মান নির্ণয়ঃ

যেহেতু বিকল্প $H_A : \mu \neq 580$ অর্থাৎ 580 এর সমান নয়, তাই দ্বি-পার্শ্বিক পরীক্ষা (two tailed test) প্রযোজ্য হবে।
এখানে তাৎপর্য স্তর $\alpha = 1\% = 0.01$

স্বাধীনতার মাত্রা $v = n - 1 = 10 - 1 = 9$
$\therefore$ t এর তাত্ত্বিক মান = ± 3.25

চতুর্থ ধাপ ঃ সিদ্ধান্ত গ্রহণঃ



যেহেতু t এর নির্ণীত মান (-1.48) উহার তাত্ত্বিক মান -3.25 থেকে + 3.25 এর মধ্যে অবস্থান করে তাই নাস্তি কল্পনা H_0 গ্রহণীয়। অতএব, শ্রমিকদের গড় মজুরি 580 টাকা।

ii.পরিমিত ব্যবধান (σ বা S) জানা আছে এবং নমুনার আকার ছোট (n<30): (Z Test)

পরিমিত ব্যবধান (σ বা S) জানা আছে এবং নমুনার আকার ছোট (n<30) এবং সমগ্রক পরিমিতভাবে বিন্যস্ত হলে Test Statistic হবেঃ

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

উদাহরণ : কোন প্রতিষ্ঠানের উৎপাদন পরিমিতভাবে বিন্যাস্ত, যার পরিমিত ব্যবধান 5.4। দৈবচয়ন ভিত্তিতে ঐ প্রতিষ্ঠান হতে 25টি দ্রব্য নির্বাচন করা হর যার গড় ওজন 128 গ্রাম। উৎপাদিত দ্রব্যের গড় ওজন 130 গ্রাম কিনা 5% গুরুত্বের স্তরে পরীক্ষা করুন।

সমাধানঃ দেয়া আছে– সমগ্রকের পরিমিত ব্যবধান, σ = 5.4
নমুনার গড়, $\bar{x}$ =128
নমুনার আকার, n = 25

এখানে-

নাস্তি অনুমান, $H_0 : \mu = 130$ অর্থাৎ উৎপাদিত দ্রব্যের গড় ওজন 130 গ্রাম।

বিকল্প অনুমান, $H_A : \mu \neq 130$ অর্থাৎ উৎপাদিত দ্রব্যের গড় ওজন 130 গ্রাম নয়।

যেহেতু σ জানা আছে, নমুনার আকার ছোট এবং সমগ্রক পরিমিতভাবে বিন্যস্ত।

অতএব Test Statistic হবে

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$\therefore \quad z = \frac{128 - 130}{\frac{5.4}{\sqrt{25}}}$$

$$= \frac{-2}{\frac{5.4}{5}}$$

$$= \frac{-2}{1.08} = -1.85$$

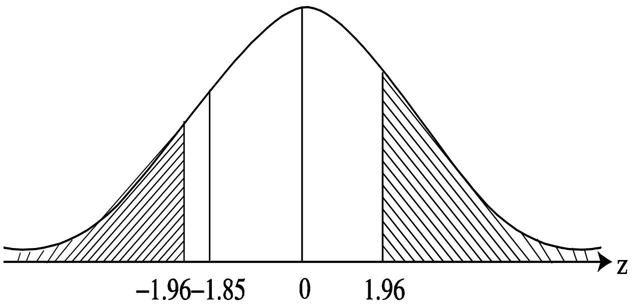
∴z-এর নির্ণীত মান = -1.85

সেহেতু $H_A : \mu \neq 130$

অতএব পরীক্ষাটি দুই দিক বিশিষ্ট।
এখানে গুরুত্বের স্তর, $\alpha = 5\% = 0.05$

$\therefore$ z এর তাত্ত্বিক (Critical) মান = $\pm$ 1.96

সিদ্ধান্তঃ



যেহেতু z এর নির্ণীত মান (-1.85) উহার তাত্ত্বিক (Critical) মানদ্বয়ের মধ্যে আছে। তাই নমুনা গ্রহণীয় এলাকায় পড়ে। অতএব H_0 গ্রহণীয়। অর্থাৎ উৎপাদিত দ্রব্যের গড় ওজন 130 গ্রাম।

iii. পরিমিত ব্যবধান (σ বা S) জানা নেই এবং নমুনার আকার বড় (n>30): (Z Test)

পরিমিত ব্যবধান (σ বা S) জানা নেই এবং নমুনার আকার বড় (n>30) এবং সমগ্রক পরিমিতভাবে বিন্যস্ত হলে Test Statistic হবেঃ

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

উদাহরণ : একটি কোম্পানীর তৈরী ফ্লোরোসেন্ট আলোর 100 টি বাল্বের কোন নমুনার গড় স্থায়িত্ব 1570 ঘন্টা, পরিমিত ব্যবধান 120 ঘন্টা হতে পারে বলে নিরূপন করা হয়েছে। কোম্পানীর তৈরী সব বাল্বের গড় স্থায়ীত্ব যদি μ হয় (অ) 0.05 এবং (আ) 0.01 তাৎপর্য মাত্রা ব্যবহার করে বিকল্প কল্পনা $\mu \neq 1600h$ এর বিপরীতে $\mu = 1600h$ কল্পনাটি যাচাই করুন।

সমাধান :

দেয়া আছে–

নমুনার আকার, $n = 100$
নমুনার গড়, $\bar{x} = 1570$ ঘন্টা
নমুনার পরিমিত ব্যবধান, $S = 120$ ঘন্টা
সমগ্রকের গড়, $\mu = 1600$ h
সমগ্রকের পরিমিত ব্যবধান, $\sigma = 120$ h

এখানে-

নাস্তি অনুমান, $H_0 : \mu = 1600$ অর্থাৎ কোম্পানীর বাল্বের গড় স্থায়িত্ব 1600 ঘন্টা।
বিকল্প অনুমান, $H_A : \mu \neq 1600$ অর্থাৎ কোম্পানীর বাল্বের গড় স্থায়িত্ব 1600 ঘন্টা নয়।
যেহেতু σ জানা নেই, নমুনার আকার বড় এবং সমগ্রক পরিমিতভাবে বিন্যস্ত।
অতএব Test Statistic হবে-

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

$$= \frac{1570-1600}{\frac{120}{\sqrt{100}}}$$

$$= \frac{-30}{\frac{120}{10}}$$

$$= \frac{-30}{12} = -2.5$$

$\therefore z$-এর নির্ণীত মান $= -2.5$

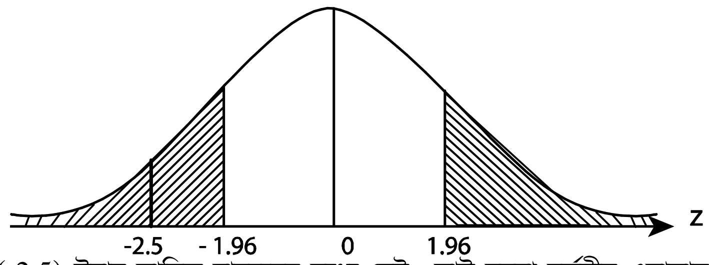
সেহেতু $H_A : \mu \neq 1600$

অতএব পরীক্ষাটি দুইদিক বিশিষ্ট।

(অ) দেওয়া আছে, তাৎপর্য মাত্রা, $\alpha = 0.05$

$\therefore$ z এর তাত্ত্বিক মান $= \pm 1.96$

সিদ্ধান্তঃ

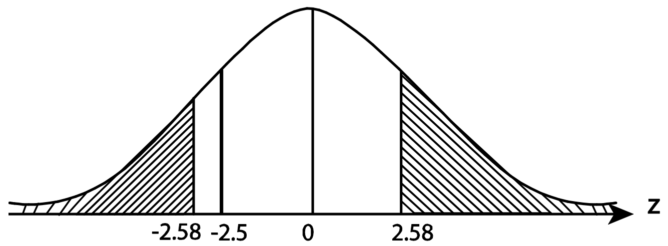


যেহেতু z এর নির্ণীত মান (-2.5) উহার তাত্ত্বিক মানদ্বয়ের মধ্যে নেই। তাই নমুনা বর্জনীয় এলাকায় পড়ে। অর্থাৎ H_0 বর্জনীয়। অতএব কোম্পানীর বাল্বের গড় স্থায়িত্ব 1600 ঘন্টা নয়।

(আ) দেওয়া আছে, তাৎপর্য মাত্রা, $\sigma = 1\% = 0.01$

$\therefore$ z এর তাত্ত্বিক মান $= \pm 2.58$

সিদ্ধান্তঃ



যেহেতু z এর নির্ণীত মান (-2.5) উহার তাত্ত্বিক মানদ্বয়ের মধ্যে আছে। তাই নমুনা গ্রহণীয় এলাকায় পড়ে। অর্থাৎ H_0 গ্রহণীয়। অতএব কোম্পানীর বাল্বের গড় স্থায়িত্ব 1600 ঘন্টা।

iv. পরিমিত ব্যবধান (σ বা S) জানা আছে এবং নমুনার আকার বড় (n>30): (Z Test)

পরিমিত ব্যবধান (σ বা S) জানা আছে এবং নমুনার আকার বড় (n>30) এবং সমগ্রক পরিমিতভাবে বিন্যস্ত হলে Test Statistic হবেঃ

$$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

উদাহরণ : একটি কোম্পানী কর্তৃক উৎপাদিত তারের গড় ভঙ্গুঁরাঙ্ক 1800N এবং পরিমিত ব্যবধান 100N। কোম্পানিটি দাবী করে যে উৎপাদন প্রক্রিয়ায় নতুন প্রযুক্তি ব্যবহার করলে তারের ভঙ্গুরাঙ্ক বাড়তে পারে, এ দাবিটি যাচাই করার জন্য 50 টি তারের একটি নমুনা পরীক্ষা করে দেখা গেল যে, গড় ভঙ্গুঁরাঙ্ক 1850N, 0.01 ভাগ তাৎপর্য স্তরে আমরা কি কোম্পানিটির দাবী সমর্থন করতে পারি।

সমাধান : দেয়া আছে– সমগ্রকের গড়, $\mu = 1800$ N
সমগ্রকের পরিমিত ব্যবধান, $\sigma = 100$ N
নমুনার আকার, $n = 50$
নমুনার গড়, $\bar{x} = 1850$ N

এখানে- নাস্তি অনুমান, $H_0 : \mu \leq 1800$
অর্থাৎ $H_0 : \mu = 1800$ অর্থাৎ কোম্পানীর দাবী সমর্থন করা যায় না।
বিকল্প অনুমান, $H_A : \mu > 1800$ অর্থাৎ কোম্পানীর দাবী সমর্থন করা যায়।

যেহেতু **σ জানা আছে, নমুনার আকার বড়** এবং সমগ্রক পরিমিতভাবে বিন্যস্ত।
অতএব Test Statistic হবে-

$$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

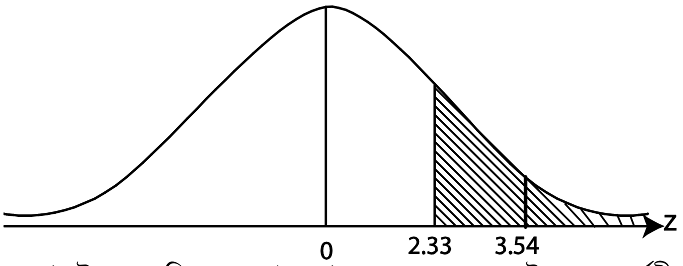
$$= \frac{1850-1800}{\frac{100}{\sqrt{50}}}$$

$$= \frac{50}{\frac{100}{7.07}}$$

$$= \frac{50}{14.14} = 3.54$$

$\therefore$ z-এর নির্ণীত মান = 3.54
সেহেতু $H_A : \mu > 1800$
অতএব পরীক্ষাটি একদিক (ডানদিক) বিশিষ্ট।
এখানে তাৎপর্য স্তর, $\alpha = 0.01$
$\therefore$ z এর তাত্ত্বিক (Critical) মান = 2.33

সিদ্ধান্তঃ



যেহেতু z এর নির্ণীত মান (3.54) উহার তাত্ত্বিক মান (2.33) অপেক্ষা বড়। তাই নমুনা বর্জনীয় এলাকায় পড়ে। অর্থাৎ H_0 বর্জনীয়। অতএব কোম্পানীর দাবি সমর্থন করা যায়।

গড়ের উপর ২টি নমুনার পরীক্ষা (Test for two population mean)ঃ

গড়ের উপর দুইটি নমুনার পরীক্ষা নিম্নোক্তভাবে করা যেতে পারে:

i. যখন পরিমিত ব্যবধান (σ বা S) জানা নেই এবং নমুনার আকার ছোট (n<30) (t Test)

ii. যখন পরিমিত ব্যবধান (σ বা S) জানা আছে এবং নমুনার আকার ছোট (n<30) (Z Test)

iii. যখন পরিমিত ব্যবধান (σ বা S) জানা নেই এবং নমুনার আকার বড় (n>30) (Z Test)

iv. যখন পরিমিত ব্যবধান (σ বা S) জানা আছে এবং নমুনার আকার বড় (n>30) (Z Test)

নিম্নে পর্যায়ক্রমে পরীক্ষাগুলো ব্যাখ্যা করা হলো :

i. পরিমিত ব্যবধান (σ বা S) জানা নেই এবং নমুনার আকার ছোট (n<30)ঃ (t Test)

পরিমিত ব্যবধান (σ বা S) জানা নেই এবং নমুনার আকার ছোট (n<30) এক্ষেত্রে গড়ের উপর ২টি নমুনার পরীক্ষা ২ভাবে হতে পারে।

ক. গড়ের উপর ২টি নমুনার (গড়ের পার্থক্যের) পরীক্ষা (স্বাধীন নমুনা)

খ. গড়ের উপর ২টি নমুনার (গড়ের পার্থক্যের) পরীক্ষা (সম্পর্কযুক্ত নমুনা)

ক. গড়ের উপর ২টি নমুনার (গড়ের পার্থক্যের) পরীক্ষা (স্বাধীন নমুনা)

Hypothesis Testing for Two Population Means (Difference) (Independent Samples)

মনে করা যাক যে, সমান পরিমিত ব্যবধান বিশিষ্ট ($\sigma_2 = \sigma_2$) দুটি স্বাভাবিক সমগ্রক (Normal population) থেকে n_1 এবং n_2 আয়তন বিশিষ্ট নমুনা দৈবচয়ন করা হয়েছে। আরও ধরা যাক যে, এ দুটি নমুনার গড় যথাক্রমে $\overline{X}_1$ ও $\overline{X}_2$ এবং পরিমিত ব্যবধান যতাক্রমে S_1 এবং S_2। এখন নমুনাগুলো একই সমগ্রক (i.e. $\mu_1 = \mu_2$ এবং $\sigma_2 = \sigma_2$) থেকে এসেছে। যদি নমুনার আকার ছোট এবং পরিমিত ভাবে বিন্যস্ত সমগ্রকের পরিমিত ব্যবধান (σ) অজানা থাকে তবে, Test Statistic

হবে, $t = \dfrac{\overline{X}_1 - \overline{X}_2}{\sigma\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ এখানে,

$$\sigma = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}}$$

$\overline{X}_1$ = প্রথম নমুনার গড়

$\overline{X}_2$ = ২য় নমুনার গড়

S_1 = ১ম নমুনার পরিমিত ব্যবধান

S_2 = ২য় নমুনার পরিমিত ব্যবধান

n_1 = ১ম নমুনার আকার

n_2 = ২য় নমুনার আকার

স্বাধীনতার মাত্রা $v = n_1 + n_2 - 2$

উদাহরণ : কোন একটি নগরীর এক অঞ্চলের 16 জন ছাত্রের বুদ্ধাংক (IQ) পরিমাপ করে দেখা গেল যে, গড় 107 এবং পরিমিত ব্যবধান 10। অন্যদিকে, অপর একটি এলাকার 14 জন ছাত্রের বুদ্ধাংকের (IQ) গড় 112 এবং পরিমিত ব্যবধান 8। দুটি অঞ্চলের ছাত্রদের বুদ্ধাংকের (IQ) মধ্যে কোন তাৎপর্যপূর্ণ পার্থক্য আছে কী? তাৎপর্য স্তর (a) 0.01 এবং (b) 0.05 ব্যবহার করুন।

সমাধান ঃ দেওয়া আছে,

	অঞ্চল-১	অঞ্চল-২
নমুনার আকার,	$n_1 = 16$	$n_2 = 14$
নমুনার গড়,	$\overline{X}_1 = 107$	$\overline{X}_2 = 112$
নমুনার পরিমিত ব্যবধান,	$S_1 = 10$	$S_2 = 8$

(a) $\alpha = 0.01$ তাৎপর্যস্তর

প্রথম ধাপ ঃ প্রতিপাদ্য / কল্পনা প্রণয়নঃ

নাস্তি অনুমান, $H_0 : \mu_1 = \mu_2$ (অর্থাৎ দুটি অঞ্চলের ছাত্রদের বুদ্ধাংক সমান)

বিকল্প অনুমান, $H_A : \mu_1 \neq \mu_2$ (অর্থাৎ দুটি অঞ্চলের ছাত্রদের বুদ্ধাংক সমান নয়)

দ্বিতীয় ধাপ ঃ t এ মান নির্ণয়ঃ

যেহেতু নমুনার আকার ছোট এবং পরিমিতভাবে বিন্যস্ত সমগ্রকের পরিমিত ব্যবধান (σ) জানা নেই,
তাই Test Statistic হবে,

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{107 - 112}{9.44\sqrt{\frac{1}{16} + \frac{1}{14}}}$$

$$= \frac{-5}{9.44 \times 0.366}$$

$$= -1.45$$

এখানে,

$$\sigma = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}}$$

$$= \sqrt{\frac{16(10)^2 + 14(8)^2}{16 + 14 - 2}}$$

$$= \sqrt{\frac{2496}{28}}$$

$$= 9.44$$

তৃতীয় ধাপ ঃ ঃ t এর তাত্ত্বিক মান নির্ণয়ঃ

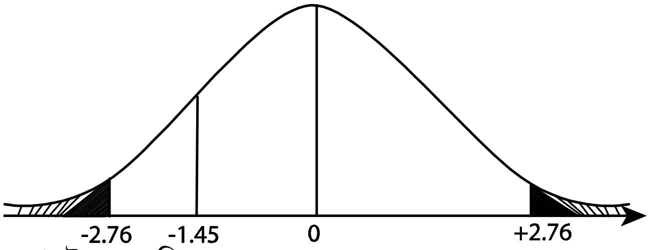
যেহেতু বিকল্প কল্পনা $H_A : \mu_1 \neq \mu_2$ তাই দ্বি-পার্শ্বিক যাচাই (two tailed test) প্রযোজ্য হবে।

এখানে তাৎপর্য স্তর $\alpha = 1\% = 0.01$

স্বাধীনতার মাত্রা $v = n_1 + n_2 - 2 = 16 + 14 - 2 = 28$

$\therefore$ t এর তাত্ত্বিক মান = ± 2.76

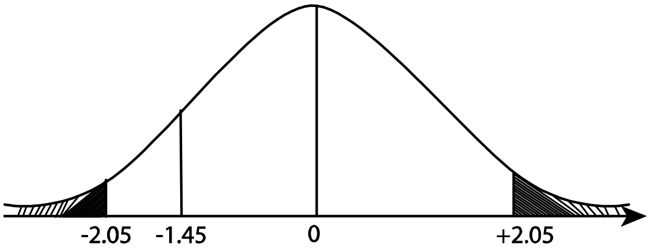
চতুর্থ ধাপ ঃ সিদ্ধান্ত গ্রহণঃ



যেহেতু t এর নির্ণীত মান (-1.45) উহার তাত্ত্বিক মান -2.76 থেকে + 2.76 এর মাঝে অবস্থান করে তাই নাস্তি কল্পনা H_0 গ্রহণীয়। সুতরাং বলা যায়, দুটি অঞ্চলের ছাত্রদের বুদ্ধাংকের মধ্যে কোন তাৎপর্যপূর্ণ পার্থক্য নেই।

(b) $\alpha = 0.05$ **তাৎপর্য স্তর**

$v = 28$ হলে এক্ষেত্রে t -এর তাত্ত্বিক মান $= \pm\, 2.05$



যেহেতু t এর নির্ণীত মান (-1.45) উহার তাত্ত্বিক মান -2.05 থেকে + 2.05 এর মাঝে অবস্থান করে তাই নাস্তি কল্পনা H_0 গ্রহণযোগ্য। সুতরাং বলা যায়, দুটি অঞ্চলের ছাত্রদের বুদ্ধাংকের মধ্যে কোন তাৎপর্যপূর্ণ পার্থক্য নেই।

উদাহরণ ঃ মনোবিজ্ঞানের উপর একটি পরীক্ষায় একটি ক্লাসের 12 জন ছাত্রের প্রাপ্ত নম্বরের গড় 78 এবং পরিমিত ব্যবধান 6। অন্যদিকে, অপর একটি ক্লাসের 15 জন ছাত্রের প্রাপ্ত নম্বরের গড় 74 এবং পরিমিত ব্যবধান 8। 0.05 তাৎপর্য স্তরে যাচাই করুন যে, ১ম ক্লাসের ছাত্ররা ২য় ক্লাসের ছাত্রদের চেয়ে মনোবিজ্ঞানে ভালো।

সমাধান ঃ দেওয়া আছে,

	ক্লাস-১	ক্লাস-২
নমুনার আকার,	$n_1 = 12$	$n_2 = 15$
নমুনার গড়,	$\overline{X}_1 = 78$	$\overline{X}_2 = 74$
নমুনার পরিমিত ব্যবধান,	$S_1 = 6$	$S_2 = 8$

প্রথম ধাপ ঃ প্রতিপাদ্য / কল্পনা প্রণয়নঃ

নাস্তি অনুমান, $H_0 : \mu_1 = \mu_2$ (অর্থাৎ দুটি ক্লাসের ছাত্রদের মনোবিজ্ঞানের দক্ষতা একই রকম)

বিকল্প অনুমান, $H_A : \mu_1 > \mu_2$ (অর্থাৎ ১ম ক্লাসের ছাত্ররা মনোবিজ্ঞানে ভালো)

দ্বিতীয় ধাপ ঃ t এর মান নির্ণয়ঃ

যেহেতু নমুনার আকার ছোট এবং পরিমিতভাবে বিন্যস্ত সমগ্রকের পরিমিত ব্যবধান (σ) জানা নেই,

তাই Test Statistic হবে,

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

এখানে,

$$\sigma = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}}$$

$$= \frac{78-74}{7.46\sqrt{\frac{1}{12}+\frac{1}{15}}}$$

$$= \frac{4}{7.46 \times .38729}$$

$$= 1.38$$

$$= \sqrt{\frac{12(6)^2+15(8)^2}{12+15-2}}$$

$$= \sqrt{\frac{1392}{25}}$$

$$= \sqrt{55.68}$$

$$= 7.46$$

তৃতীয় ধাপ ঃ ঃ t এর তাত্ত্বিক মান নির্ণয়ঃ

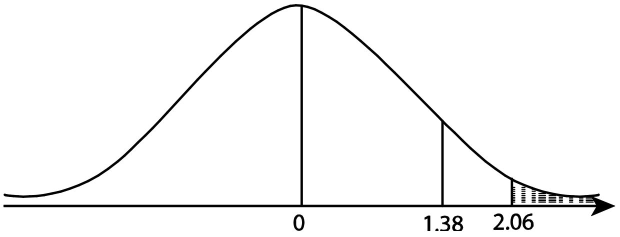
যেহেতু বিকল্প কল্পনা $H_A : \mu_1 > \mu_2$ তাই একপার্শ্বিক যাচাই (one tailed test) প্রযোজ্য হবে।

তাৎপর্য স্তর $\alpha = 0.05$

স্বাধীনতার মাত্রা $v = n_1 + n_2 - 2 = 12 + 15 - 2 = 25$

$\therefore$ t এর তাত্ত্বিক মান = 2.06

চতুর্থ ধাপ ঃ সিদ্ধান্ত গ্রহণঃ



যেহেতু t এর নির্ণীত মান 1.38 উহার তাত্ত্বিক মান 2.06 এর চেয়ে ছোট তাই নাস্তি কল্পনা H_0 গ্রহণযোগ্য। সুতরাং ১ম ক্লাসের ছাত্ররা ২য় ক্লাসের ছাত্রদের তুলনায় মনোবিজ্ঞানে ভালো একথা বলা যায় না।

খ. গড়ের উপর ২টি নমুনার (গড়ের পার্থক্যের) পরীক্ষা (সম্পর্কযুক্ত নমুনা)

Hypothesis Testing for Two Population Means (Difference) (Dependent Samples)

একই নমুনা কিন্তু দুটি ভিন্ন অবস্থার মধ্যে তুলনার ক্ষেত্রে একটি যাচাই সূত্র ব্যবহার করা হয়। উদাহরণস্বরূপঃ একদল শ্রমিকের প্রশিক্ষণের পূর্বের এবং পরের অবস্থার মধ্যে তুলনা, একদল ছাত্রের কোচিং করার পূর্বের এবং পরের অবস্থার মধ্যে তুলনা ইত্যাদি ক্ষেত্রে t যাচাই ব্যবহার করা হয়। যদি নমুনার আকার ছোট এবং পরিমিতিভাবে বিন্যস্ত সমগ্রকের পরিমিত ব্যবধানদ্বয় (σ) জানা না থাকে, এক্ষেত্রে test statistic হবে,

$$t = \frac{\bar{d}\sqrt{n}}{s}$$

যেহেতু, $S = \sqrt{\frac{\Sigma(d-\bar{d})^2}{n-1}}$

বা, $S = \sqrt{\frac{\Sigma d^2 - n(\bar{d})^2}{n-1}}$

$\bar{d}$ = দুটি পার্থক্যের গড়

s = দুটি পার্থক্যের পরিমিত ব্যবধান

n = নমুনার আকার

স্বাধীনতার মাত্রা $v = n - 1$

উদাহরণঃ পাঁচটি একই বয়সের ইঁদুরের ওজন নেওয়া হলো। অত:পর এক সপ্তাহ তাদের একটি বিশেষ খাবার দিয়ে পুনরায় ওজন নেওয়া হলো। তাদের ওজন নিচে (গ্রামে) দেওয়া হলোঃ

ইঁদুরের ক্রমিক নম্বর	1	2	3	4	5

প্রথম ওজন (X)	10.2	9.4	11.8	9.1	8.3
এক সপ্তাহ পরের ওজন (Y)	10.6	9.8	12.3	11.7	8.8

বিশেষ খাবারের সঙ্গে আগের ওজনের পার্থক্য যাচাই করুন, যখন $\alpha = 0.05$

সমাধান ঃ প্রথম ধাপঃ প্রতিপাদ্য প্রণয়নঃ

নাস্তি কল্পনা, $H_0 : \mu_x = \mu_y$ (অর্থাৎ বিশেষ খাবারের পরের এবং আগের ওজন একই রকম)

বিকল্প কল্পনা, $H_A : \mu_x \neq \mu_y$ (অর্থাৎ বিশেষ খাবারের পরের ওজনের সাথে আগের ওজনের পার্থক্য রয়েছে)

দ্বিতীয় ধাপ ঃ t এ মান নির্ণয়ঃ

যেহেতু নমুনার আকার ছোট এবং পরিমিতভাবে বিন্যস্ত সমগ্রকের পরিমিত ব্যবধান দ্বয় (σ) জানা নেই, তাই এক্ষেত্রে test statistic হবে,

$$t = \frac{\bar{d}\sqrt{n}}{s}$$

$\bar{d}$ ও s নির্ণয়ের সারণীঃ

ইঁদুরের ক্রমিক নং	প্রথম ওজন (X)	এক সপ্তাহ পরের ওজন (Y)	d = (X-Y)	d^2
1	10.2	10.6	-0.4	0.16
2	9.4	9.8	-0.4	0.16
3	11.8	12.3	-0.5	0.25
4	9.1	11.7	-2.6	6.76
5	8.3	8.8	-0.5	0.25
n = 5			Σd =-44	Σd^2 =758

$$\therefore \bar{d} = \frac{\Sigma d}{n} = \frac{-4.5}{5} = -0.88$$

$$\therefore S = \sqrt{\frac{\Sigma d^2 - n(\bar{d})^2}{n-1}} = \sqrt{\frac{7.58 - 5(-0.88)^2}{5-1}}$$

$$= \sqrt{\frac{7.58 - 3.872}{4}} = \sqrt{0.927} = 0.962$$

$$\therefore t = \frac{\bar{d}\sqrt{n}}{s} = \frac{-0.88\sqrt{5}}{0.962} = \frac{-1.968}{0.962} = -2.05$$

তৃতীয় ধাপ ঃ t এর তাত্ত্বিক মান নির্ণয়ঃ

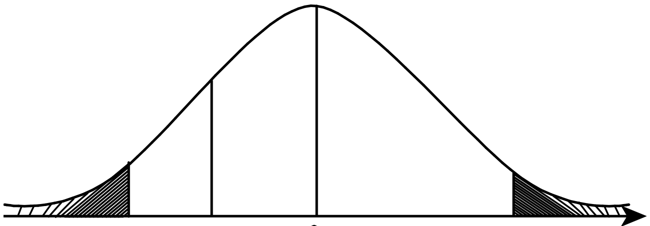
এক্ষেত্রে বিকল্প $H_A : \mu_x \neq \mu_y$ তাই দ্বি-পার্শ্বিক যাচাই (Two tailed test) প্রযোজ্য হবে।

তাৎপর্য স্তর $\alpha = 0.05$

স্বাধীনতার মাত্রা $v = n - 1 = 5 - 1 = 4$

$\therefore$ t এর তাত্ত্বিক মান = ± 2.78

চতুর্থ ধাপ ঃ সিদ্ধান্ত গ্রহণঃ



যেহেতু t এর নির্ণীত মান (-2.05) উহার তাত্ত্বিক মান -2.78 থেকে বড় তাই নাস্তি কল্পনা H_0 গ্রহণীয়। সুতরাং, বলা যায় বিশেষ খাবারের পরের ওজনের সাথে আগের ওজনের কোন তাৎপর্যপূর্ণ পার্থক্য নেই।

ii. পরিমিত ব্যবধান (σ বা S) জানা আছে এবং নমুনার আকার ছোট (n<30) (Z Test) :

পরিমিত ব্যবধান (σ বা S) জানা আছে এবং নমুনার আকার ছোট (n<30) এবং সমগ্রক পরিমিতভাবে বিন্যস্ত হলে Test Statistic হবেঃ

$$z = \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)}{\sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}}}$$

উদাহরণ : অর্থনীতি ২য় বর্ষ সম্মান পরীক্ষায় নির্বাচিত ২৬ জন ছাত্রের গড় নম্বর ৪৫ এবং সমগ্রকের ভেদাংক ৮। আবার নির্বাচিত ৩২ জন ছাত্রীর গড় নম্বর ৪৮ এবং সমগ্রকের ভেদাংক ৩.৫।

ক. ছাত্রীরা ছাত্রদের চেয়ে ভাল কিনা ১% তাৎপর্য স্তরে পরীক্ষা করুন।

সমাধান : ধরি, ছাত্রদের নম্বর x_1 এবং ছাত্রীদের নম্বর x_2

দেয়া আছে–ছাত্রদের ক্ষেত্রে, নমুনার আকার, $n_1 = 26$

নমুনার গড়, $\bar{x}_1 = 45$

সমগ্রকের ভেদাংক, $\sigma_1^2 = 8$

ছাত্রীদের ক্ষেত্রে, নমুনার আকার, $n_2 = 32$

নমুনার গড়, $\bar{x}_2 = 48$

সমগ্রকের ভেদাংক $\sigma_2^2 = 3.5$

এখানে,

নাস্তি অনুমান, $H_0 : (\mu_2 - \mu_1) \leq 0$ অর্থাৎ ছাত্রীরা ছাত্রদের চেয়ে ভাল নয়।

বিকল্প অনুমান, $H_A : \mu_2 - \mu_1 > 0$ অর্থাৎ ছাত্রীরা ছাত্রদের চেয়ে ভাল।

যেহেতু σ^2 দ্বয় জানা আছে, নমুনার আকার ছোট ($\because n_1 + n_2 = 58 < 60$) এবং সমগ্রক পরিমিতভাবে বিন্যস্ত।

অতএব Test Statistic হবে-

$$z = \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)}{\sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}}}$$

যেহেতু $H_A : (\mu_2 - \mu_1) > 0$

অতএব পরীক্ষাটি একদিক (ডানদিক) বিশিষ্ট।

এখানে তাৎপর্য স্তর, $\alpha = 1\% = 0.01$

$\therefore$ z এর তাত্ত্বিক (Critical) মান, $z_c = 2.32$

এখন, $z_c = \frac{(\bar{x}_2 - \bar{x}_1)_c - (\mu_2 - \mu_1)}{\sqrt{\frac{\sigma_2^2}{n_2} + \frac{\sigma_1^2}{n_1}}}$

বা, $2.32 = \dfrac{(\bar{x}_2 - \bar{x}_1)_c - 0}{\sqrt{\dfrac{3.5}{32} + \dfrac{8}{26}}}$

বা, $2.32 = \dfrac{(\bar{x}_2 - \bar{x}_1)_c}{\sqrt{0.109 + 0.308}}$

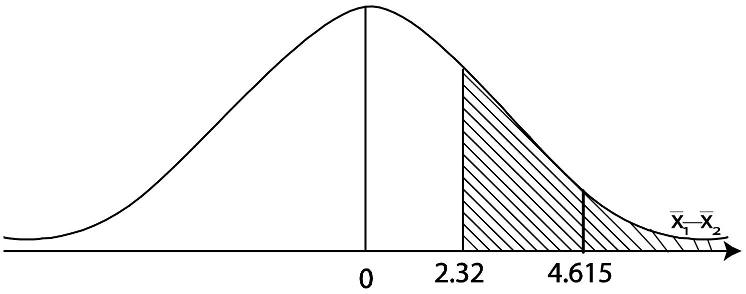
বা, $2.32 = \dfrac{(\bar{x}_2 - \bar{x}_1)_c}{.65}$

বা, $(\bar{x}_2 - \bar{x}_1)_c = 1.51$

$\therefore$ $(\bar{x}_2 - \bar{x}_1)$ এর তাত্ত্বিক মান = 1.51

এখানে $(\bar{x}_2 - \bar{x}_1)$ এর নির্ণীত মান = 48 - 45 = 3

সিদ্ধান্তঃ



যেহেতু $(\bar{x}_2 - \bar{x}_1)$ এর নির্ণীত মান (3) উহার তাত্ত্বিক মান 1.51 অপেক্ষা বড়। তাই নমুনা বর্জনীয় এলাকায় পড়ে। অতএব H_0 বর্জনীয়। অতএব ছাত্রীরা ছাত্রদের চেয়ে ভাল।

iii. পরিমিত ব্যবধান (σ বা S) জানা নেই এবং নমুনার আকার বড় (n>30): (Z Test)

পরিমিত ব্যবধান (σ বা S) জানা নেই এবং নমুনার আকার বড় (n>30) এবং সমগ্রক পরিমিতভাবে বিন্যস্ত হলে Test Statistic হবেঃ

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

উদাহরণ : মনে করুন দুটি সমগ্রক থেকে পৃথকভাবে 100 এবং 120 আকার বিশিষ্ট দু'টি দৈব নমুনা চয়ন করা হল। নমুনাগুলির গড় যথাক্রমে 104, 101 এবং ভেদাংক যথাক্রমে 90, 100। নমুনা দুটির গড়ের পার্থক্য কি 1% যথার্থ মাত্রায় তাৎপর্য পূর্ণ?

সমাধান : ধরি, দুটি নমুনার চলকদ্বয় যথাক্রমে x_1 ও x_2

দেয়া আছে– ১ম নমুনার ক্ষেত্রে, নমুনার আকার, $n_1 = 100$

নমুনার গড়, $\bar{x}_1 = 104$

নমুনার ভেদাংক, $s_1^2 = 90$

২য় নমুনার ক্ষেত্রে- নমুনার আকার, $n_2 = 120$

নমুনার গড়, $\bar{x}_2 = 101$

সমগ্রকের ভেদাংক $s_2^2 = 100$

এখানে,

নাস্তি অনুমান, $H_0 : (\mu_1 - \mu_2) = 0$ অর্থাৎ নমুনা দুটির গড়ের মধ্যে পার্থক্য নাই।

বিকল্প অনুমান, $H_A : \mu_1 - \mu_2 \neq 0$ অর্থাৎ নমুনা দুটির গড়ের মধ্যে পার্থক্য আছে।

যেহেতু σ^2 দ্বয় জানা নেই, নমুনার আকার বড় এবং সমগ্রক পরিমিতভাবে বিন্যস্ত।
অতএব Test Statistic হবে-

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$= \frac{(104 - 101) - 0}{\sqrt{\frac{90}{100} + \frac{100}{120}}}$$

$$= \frac{3}{\sqrt{0.9 + 0.83}}$$

$$= \frac{3}{1.32}$$

$$= 2.27$$

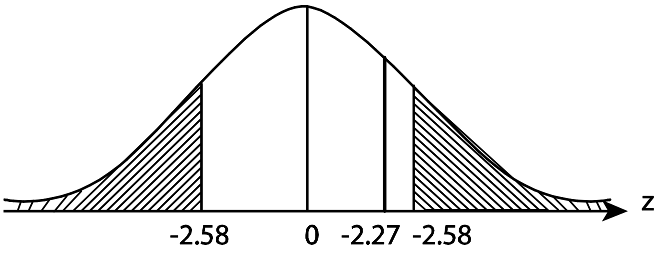
$\therefore$ z এর নির্ণীত মান = 2.27

যেহেতু $H_A : (\mu_1 - \mu_2) \neq 0$
অতএব পরীক্ষাটি দুইদিক বিশিষ্ট।
এখানে যথার্থতার মাত্রা, $\alpha = 1\% = 0.01$

$$\therefore \frac{\alpha}{2} = \frac{.01}{2} = 0.005$$

$\therefore$ z এর তাত্ত্বিক (Critical) মান = $\pm$ 2.58

সিদ্ধান্তঃ



যেহেতু z এর নির্ণীত মান (2.27) উহার তাত্ত্বিক মানদ্বয়ের মধ্যে আছে। তাই নমুনা গ্রহণীয় এলাকায় পড়ে। অতএব H_0 গ্রহণীয়। অতএব নমুনা দুটির গড়ের মধ্যে পার্থক্য নাই।

iv. পরিমিত ব্যবধান (σ বা S) জানা আছে এবং নমুনার আকার বড় (n>30): (Z Test)

পরিমিত ব্যবধান (σ বা S) জানা আছে এবং নমুনার আকার বড় (n>30) এবং সমগ্রক পরিমিতভাবে বিন্যস্ত হলে Test Statistic হবে:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

এক্ষেত্রে উভয় নমুনার আকারের সমষ্টি ৬০ এর কম হলে নমুনার আকার ছোট। অন্যথায় নমুনার আকার বড়।

উদাহরণ : দুটি দলের গাড়ী পেট্রোল দ্বারা পরিচালিত। ১ম দলের নির্বাচিত 36টি গাড়ী গড়ে 24 মাইল যায়। যার সমগ্রকের ভেদাংক 1.5 মাইল। অপর দলের নির্বাচিত 72টি গাড়ী গড়ে 22.5 মাইল যায়। যার সমগ্রকের ভেদাংক 2 মাইল। উভয় দলের গাড়ীর মধ্যে পার্থক্য আছে কিনা 1% তাৎপর্য স্তরে পরীক্ষা করুন?

সমাধান : দেয়া আছে–

১ম দল- নমুনার আকার, $n_1 = 36$

নমুনার গড়, $\bar{x}_1 = 24$ মাইল

সমগ্রকের ভেদাংক $\sigma_1^2 = 1.5$ মাইল

২য় দল- নমুনার আকার, $n_2 = 72$

নমুনার গড়, $\bar{x}_2 = 22.5$ মাইল

সমগ্রকের ভেদাংক $\sigma_2^2 = 2$ মাইল

এখানে,

নাস্তি অনুমান, $H_0 : \mu_1 - \mu_2 = 0$ অর্থাৎ উভয় দলের গাড়ীর মধ্যে পার্থক্য নাই।

বিকল্প অনুমান, $H_A : \mu_1 - \mu_2 \neq 0$ অর্থাৎ উভয় দলের গাড়ীর মধ্যে পার্থক্য আছে।

যেহেতু **σ^2 দ্বয় জানা আছে, নমুনার আকার বড়** এবং সমগ্রক পরিমিতভাবে বিন্যস্ত।

অতএব Test Statistic হবে

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$= \frac{(24 - 22.5) - 0}{\sqrt{\frac{1.5}{36} + \frac{2}{72}}}$$

$$= \frac{1.5}{\sqrt{0.042 + 0.028}}$$

$$= \frac{1.5}{0.26}$$

$$= 5.77$$

$\therefore$ z এর নির্ণীত মান $= 5.77$

যেহেতু $H_A : \mu_1 - \mu_2 \neq 0$

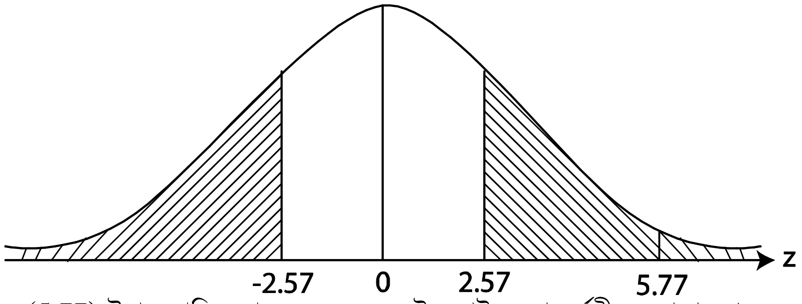
অতএব পরীক্ষাটি দুইদিক বিশিষ্ট।

এখানে গুরুত্বের স্তর, $\alpha = 1\% = 0.01$

$$\therefore \frac{\alpha}{2} = \frac{.01}{2} = 0.005$$

$\therefore$ z এর তাত্ত্বিক মান $= \pm 2.57$

সিদ্ধান্তঃ



যেহেতু z এর নির্ণীত মান (5.77) উহার তাত্ত্বিক মানদ্বয়ের মধ্যে নেই। তাই নমুনা বর্জনীয় এলাকায় পড়ে। অতএব H_0 বর্জনীয়। অতএব উভয় দলের গাড়ীর মধ্যে পার্থক্য আছে।

সারসংক্ষেপ

গড়ের উপর একটি নমুনার পরীক্ষার সূত্র হচ্ছে $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$ এবং $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

আবার গড়ের উপর ২টি নমুনার (গড়ের পার্থক্যের) পরীক্ষার সূত্র হচ্ছে (স্বাধীন নমুনা) (n<30): $t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

এছাড়া গড়ের উপর ২টি নমুনার (গড়ের পার্থক্যের) পরীক্ষার সূত্র হচ্ছে (সম্পর্কযুক্ত নমুনা) (n<30): $t = \frac{\bar{d}\sqrt{n}}{s}$ এবং

গড়ের উপর ২টি নমুনার (গড়ের পার্থক্যের) পরীক্ষার সূত্র হচ্ছে (n>30): $z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

পাঠ ৮.৩ অনুপাত সম্পর্কিত কল্পনা যাচাই
Test of Hypothesis about Proportion

উদ্দেশ্য

এ পাঠ শেষে আপনি-

- অনুপাতের উপর একটি নমুনার পরীক্ষা পদ্ধতি বর্ণনা করতে পারবেন।
- অনুপাতের উপর দুটি নমুনার পরীক্ষা ব্যাখ্যা করতে পারবেন।

অনুপাত সম্পর্কিত কল্পনা যাচাই
Test of Hypothesis about Proportion

অনুপাত সম্পর্কিত কল্পনা যাচাই নিম্নোক্তভাবে করা যেতে পারে:

ক. অনুপাতের উপর একটি নমুনার পরীক্ষা (Test for single population proportion) (z test)

খ. অনুপাতের উপর দুটি নমুনার (অনুপাতের পার্থক্য) পরীক্ষা (Test for two population proportion) (z test)

ক. সমগ্রকের অনুপাতের উপর একটি নমুনার পরীক্ষা
One Sample Test on Population Proportion

সমগ্রকের অনুপাতের ক্ষেত্রে অনুমান বা কল্পনা যাচাইয়ে Z-Test ব্যবহৃত হয়। যা হল−

$$z = \frac{\bar{p} - \bar{\wedge}}{\sqrt{\dfrac{\bar{\wedge}(1-\bar{\wedge})}{n}}}$$

এখানে n = নমুনার আকার

$\bar{p}$ = নমুনার অনুপাত

$\bar{\wedge}$ = সমগ্রকের অনুপাত

উদাহরণ ঃ একজন ছাত্র প্রতিনিধি দাবী করে যে, 80% শিক্ষার্থী ক্যাম্পাসের খাদ্য সরবরাহের উপর অসন্তুষ্ট। দৈবচয়ন ভিত্তিতে 50 জন শিক্ষার্থীকে নির্ধারণ করে জানা গেল যে, 20 জন বর্তমান খাদ্য সরবরাহকে সন্তোষ জনক বলেছে। অসন্তোষের হার কমেছে কিনা 1% তাৎপর্যস্তরে পরীক্ষা করুন।

সমাধানঃ দেওয়া আছে,

নমুনার আকার, $n = 50$

অসন্তোষের ক্ষেত্রে নমুনার অনুপাত, $\bar{p} = \frac{50-20}{50} = \frac{30}{50} = 0.6$

অসন্তোষের ক্ষেত্রে সমগ্রকের অনুপাত, $\bar{\wedge} = 80\% = 0.8$

প্রথম ধাপ ঃ প্রতিপাদ্য বা কল্পনা প্রণয়ন ঃ

নাস্তি অনুমান, H_0: $\bar{\wedge} \geq 0.80$ (অর্থাৎ অসন্তোষের হার কমছে না)

বিকল্প অনুমান, H_A: $\bar{\wedge} < 0.80$ (অর্থাৎ অসন্তোষের হার কমছে)

দ্বিতীয় ধাপ ঃ z এ মান নির্ণয়ঃ

এক্ষেত্রে Test Statistic হবে, $z = \frac{\bar{p} - \bar{\wedge}}{\sqrt{\dfrac{\bar{\Lambda}(1-\bar{\wedge})}{n}}}$

$$=\frac{0.6-0.8}{\sqrt{\frac{0.8(1-0.8)}{50}}}$$

$$=\frac{-0.2}{\sqrt{\frac{0.8\times0.2}{50}}}$$

$$=\frac{-0.2}{0.566}=3.53$$

$\therefore$ z = এর নির্ণীত মান = -3.53

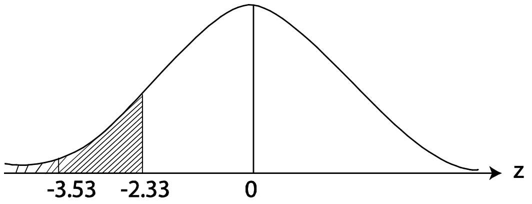
যেহেতু $H_A: \overline{\wedge} < 0.8$

অতএব পরীক্ষাটি একদিক (বামদিক) বিশিষ্ট।

এখানে তাৎপর্য স্তর, $\alpha = 1\% = 0.01$

$\therefore$ z এর তাত্ত্বিক (Critical) মান = -2.33

সিদ্ধান্ত :



যেহেতু z এর নির্ণীত মান (-3.53) উহার তাত্ত্বিক মান (-2.33) অপেক্ষা ছোট। তাই নমুনা বর্জনীয় এলাকায় পড়ে। অর্থাৎ H_0 বর্জনীয়। অতএব অসন্তোষের হার কমছে।

উদাহরণ ঃ একজন সমাজবিজ্ঞানীর ধারণা যে, শহরে বসবাসকারী প্রাপ্ত বয়স্ক লোকদের 80% লোক চা পান করেন। তার এ ধারণার সত্যতা যাচাই করতে দৈবভাবে 400 প্রাপ্তবয়স্ক লোক থেকে জানতে পারলেন যে, 350 জন লোক চা পান করেন। প্রাপ্ত তথ্য থেকে উক্ত সমাজবিজ্ঞানীর ধারণার সত্যতা যাচাই করুন যখন $\alpha = 0.01$।

সমাধানঃ দেওয়া আছে,

নমুনার আকার, n = 400

নমুনা চা পানের অনুপাত, $\overline{p}=\frac{350}{400}=0.875$

সমগ্রক চা পানের অনুপাত, $\overline{\wedge} = 80\% = 0.80$

প্রথম ধাপ ঃ প্রতিপাদ্য বা কল্পনা প্রণয়ন ঃ

নাস্তি কল্পনা, $H_0: \overline{\wedge} = 0.80$ (অর্থাৎ সমাজবিজ্ঞানীদের ধারণা সঠিক)

বিকল্প অনুমান, $H_A: \overline{\wedge} \neq 0.80$ (অর্থাৎ সমাজবিজ্ঞানীদের ধারণা সঠিক নয়)

দ্বিতীয় ধাপ ঃ z এ মান নির্ণয়ঃ

এক্ষেত্রে Test Statistic হবে, $z = \frac{\overline{p}-\overline{\wedge}}{\sqrt{\frac{\overline{\Lambda}(1-\overline{\wedge})}{n}}}$

$$= \frac{0.875-0.80}{\sqrt{\frac{0.80(1-0.80)}{400}}}$$

$$= \frac{0.075}{\sqrt{\frac{0.16}{400}}}$$

$$= 3.5$$

$\therefore$ z-এর নির্ণীত মান = 3.5

তৃতীয় ধাপ ঃ z এর তাত্ত্বিক মান নির্ণয়ঃ

যেহেতু, এক্ষেত্রে বিকল্প কল্পনা $H_A : \pi \neq 0.80$ তাই two tailed test প্রযোজ্য হবে।

$\alpha = 1\% = 0.01$ তাৎপর্য মাত্রায় Two tailed test এ পরিমিত সারণী অনুসারে z এর তালিকা মান = ± 2.58।

চতুর্থ ধাপ ঃ সিদ্ধান্ত গ্রহণঃ

যেহেতু z এর নির্ণীত মান 3.5 উহার তাত্ত্বিক মান -2.58 থেকে +2.58 এর মাঝে অবস্থান করে না, তাই নাস্তি কল্পনা H_0 বর্জনীয়। সুতরাং, বলা যায় যে, উক্ত 'সমাজবিজ্ঞানীর ধারণা সঠিক নয়।

খ. অনুপাতের উপর দুটি নমুনার (অনুপাতের পার্থক্য) পরীক্ষা:

Test on the Difference between Two Proportions

A জেলার 300 জন এবং B জেলার 200 জন ভোটারের মধ্যে যথাক্রমে 56% এবং 48% ভোট কোন প্রার্থীর পক্ষে পড়ে। 0.05 ভাগ স্তরে পরীক্ষা কর যে-

(i) উভয় জেলার ভোটারদের পছন্দনীয়তার মধ্যে পার্থক্য আছে।

(ii) A জেলার ভোটাররা ঐ প্রার্থীকে বেশী পছন্দ করে

সমাধান:দেয়া আছে,

A জেলার ক্ষেত্রে:

নমুনার আকার, $n_A = 300$

নমুনার অনুপাত, $\overline{p}_A = 56\% = 0.56$

B জেলার ক্ষেত্রে:

নমুনার আকার, $n_B = 300$

নমুনার অনুপাত, $\overline{p}_B = 48\% = 0.48$

এখানে,

নাস্তি অনুমান, $H_0 : (\pi_A - \pi_B) = 0$ অর্থাৎ উভয় জেলার ভোটারদের পছন্দনীয়তার মধ্যে পার্থক্য নেই।

নাস্তি অনুমান, $H_A : (\pi_A - \pi_B) \neq 0$ অর্থাৎ উভয় জেলার ভোটারদের পছন্দনীয়তার মধ্যে পার্থক্য আছে।

এখানে Test Statistic হবে–

$$z = \frac{(\overline{p}_A - \overline{p}_B) - (\pi_A - \pi_B)}{\sqrt{\frac{\overline{p}_A(1-\overline{p}_A)}{n_A} + \frac{\overline{p}_B(1-\overline{p}_B)}{n_B}}}$$

$$= \frac{(0.56-0.48)-0}{\sqrt{\frac{0.56(1-0.56)}{300} + \frac{0.48(1-0.48)}{200}}}$$

$$= \frac{0.08}{\sqrt{\frac{0.56 \times 0.44}{300} + \frac{0.48 \times 0.52}{200}}}$$

$$= \frac{0.08}{\sqrt{0.00082 + 0.00125}}$$

$$= \frac{0.08}{0.045} = 1.78$$

$\therefore$ z = এর নির্ণীত মান = 1.78

যেহেতু $H_A : (\overline{\wedge}_1 - \overline{\wedge}_2) \neq 0$

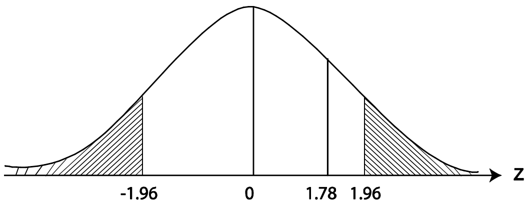
অতএব পরীক্ষাটি দুইদিক বিশিষ্ট।

এখানে গুরুত্বের স্তর, $\alpha = 0.05$

$$\therefore \frac{\alpha}{2} = \frac{0.05}{2} = 0.025$$

$\therefore$ z এর তাত্ত্বিক (Critical) মান = $\pm$ 1.96

সিদ্ধান্ত :



যেহেতু z এর নির্ণীত মান (1.78) উহার তাত্ত্বিক মানদ্বয়ের মধ্যে আছে। তাই নমুনা গ্রহণীয় এলাকায় পড়ে। অর্থাৎ H_0 গ্রহণীয়। অতএব উভয় জেলার ভোটারদের পছন্দনীয়তার মধ্যে পার্থক্য নাই।

(ii) এখানে,

নাস্তি অনুমান, $H_0 : (\overline{\wedge}_A - \overline{\wedge}_B) \leq 0$ অর্থাৎ A জেলার ভোটাররা ঐ প্রার্থীকে বেশী পছন্দ করে না।

নাস্তি অনুমান, $H_0 : (\overline{\wedge}_A - \overline{\wedge}_B) > 0$ অর্থাৎ A জেলার ভোটাররা ঐ প্রার্থীকে বেশী পছন্দ করে।

এক্ষেত্রে Test Statistic হবে–

$$z = \frac{(\bar{p}_A - \bar{p}_B) - (\overline{\wedge}_A - \overline{\wedge}_B)}{\sqrt{\frac{\bar{p}_A(1-\bar{p}_A)}{n_A} + \frac{\bar{p}_B(1-\bar{p}_B)}{n_B}}}$$

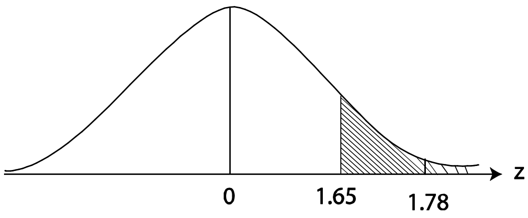
$$= \frac{(0.56 - 0.48) - 0}{\sqrt{\frac{0.56(1-0.56)}{300} + \frac{0.48(1-0.48)}{200}}}$$

$$= \frac{0.08}{0.045} = 1.78$$

$\therefore$ z = এর নির্ণীত মান = 1.78

যেহেতু $H_A : (\overline{\wedge}_A - \overline{\wedge}_B) > 0$
অতএব পরীক্ষাটি একদিক (ডানদিক) বিশিষ্ট।
এখানে গুরুত্বের স্তর, $\alpha = 0.05$
$\therefore$ z এর তাত্ত্বিক (Critical) মান = 1.56

সিদ্ধান্ত :



যেহেতু z এর নির্ণীত মান (1.78) উহার তাত্ত্বিক মান (1.56) অপেক্ষা বড়। তাই নমুনা বর্জনীয় এলাকায় পড়ে। অর্থাৎ H_0 বর্জনীয়। অতএব A জেলার ভোটার ঐ প্রার্থীকে বেশী পছন্দ করে।

সারসংক্ষেপ

অনুপাতের উপর একটি নমুনার পরীক্ষার সূত্র হচ্ছে: $z = \dfrac{\overline{p} - \overline{\wedge}}{\sqrt{\dfrac{\overline{\wedge}(1-\overline{\wedge})}{n}}}$ এবং

অনুপাতের উপর দুটি নমুনার (অনুপাতের পার্থক্য) পরীক্ষার সূত্র হচ্ছে: $z = \dfrac{(\overline{p}_A - \overline{p}_B) - (\overline{\wedge}_A - \overline{\wedge}_B)}{\sqrt{\dfrac{\overline{p}_A(1-\overline{p}_A)}{n_A} + \dfrac{\overline{p}_B(1-\overline{p}_B)}{n_B}}}$

পাঠ ৮.৪ সংশ্লেষাংক সম্পর্কিত কল্পনা যাচাই
Test of Hypothesis about Coefficient of Correlation

উদ্দেশ্য

এ পাঠ শেষে আপনি-

- শূন্য সংশ্লেষাংক যাচাই পদ্ধতি বর্ণনা করতে পারবেন।
- নির্দিষ্ট মানের সংশ্লেষাংক যাচাই পদ্ধতি ব্যাখ্যা করতে পারবেন।
- দুটি সংশ্লেষাংকের পার্থক্য যাচাই পদ্ধতি লিখতে পারবেন।

সংশ্লেষাংক সম্পর্কিত কল্পনা যাচাই
Test of Hypothesis about Coefficient of Correlation

সংশ্লেষাংক সম্পর্কিত যে সমস্ত যাচাই করা হয় সেগুলোকে তিন ভাগে ভাগ করা যায়। যথাঃ

(ক) শূন্য সংশ্লেষাংক যাচাই; (t Test)

(খ) নির্দিষ্ট মানের সংশ্লেষাংক যাচাই; (Z Test)

(গ) দুটি সংশ্লেষাংকের পার্থক্য যাচাই। (Z Test)

নিচে শূন্য সংশ্লেষাংক যাচাই, নির্দিষ্ট মানের সংশ্লেষাংক যাচাই এবং দুটি সংশ্লেষাংকের পার্থক্য যাচাই বিশ্লেষণ করা হলো:

(ক) **শূন্য সংশ্লেষাংক যাচাই** (T Test)

Zero Coefficient of Correlation Test

ধরি, দ্বি-চলক বিশিষ্ট সমগ্রক হতে n সংখ্যক জোড়া নমুনা (x_1,y_1), (x_2,y_2), ... (x_n,y_n) দৈবচয়ন করা হয়েছে। সমগ্রকের সংশ্লেষাংক হল ρ। ফলে শূন্য সংশ্লেষাংক যাচাই করার জন্য নাস্তি কল্পনা (Null hypothesis) হবে, H_0: $\rho = 0$ যার বিকল্প কল্পনা H_0: $\rho \neq 0$ কিংবা H_A: $\rho > 0$ অথবা H_0: $\rho < 0$ হতে পারে।

এক্ষেত্রে n জোড়া নমুনা তথ্য থেকে প্রাপ্ত ρ এর নিরপেক্ষ নিরূপক যদি r হয় এবং n জোড়া তথ্যমান যদি দ্বি-চলক বিশিষ্ট পরিমিত বিন্যাস থেকে চয়ন করা হয়ে থাকে তবে নাস্তি কল্পনার যাচাই তথ্যজমান (Test Statistic) হবে,

$$t = \frac{r\sqrt{(n-2)}}{\sqrt{1-r^2}}$$

যা 'Student' t বিন্যাস অনুসরণ করে এবং যার স্বাধীনতার মাত্রা হল, $v = n - 2$

উদাহরণ ঃ ১০ টি পরিবারের আয় (Y) এবং ভোগব্যয় (C) সম্পর্কিত উপাত্ত থেকে দেখা যায় C এবং Y এর মধ্যেকার সহ-সংশ্লেষাংক r = 0.93। এর উপাত্ত অনুসারে আমরা কি সিদ্ধান্ত গ্রহণ করতে পারি যে, Y এবং C এর মধ্যে রৈখিক সহ-সম্বন্ধ রয়েছে।

সমাধান ঃ নমুনার আকার, $n = 10$

নমুনার সংশ্লেষাংক, $r = 0.93$

প্রথম ধাপ ঃ প্রতিপাদ্য বা কল্পনা প্রণয়ন ঃ

নাস্তি কল্পনা, $H_0 : \rho = 0$ [অর্থাৎ আয় (Y) ও ভোগ (C) এর মধ্যে কোন সহ-সম্বন্ধ নেই।]

বিকল্প কল্পনা, $H_A : \rho \neq 0$ [অর্থাৎ আয় (Y) ও ভোগ (C) এর মধ্যে কোন সহ-সম্বন্ধ আছে।]

দ্বিতীয় ধাপ ঃ t এর মান নির্ণয়ঃ

এক্ষেত্রে যাচাই তথ্যজমান (Test Statistic) হবে,

t $= \frac{r\sqrt{(n-2)}}{\sqrt{1-r^2}}$

$= \frac{.93\sqrt{10-2}}{\sqrt{1-(0.93)^2}}$

$= \frac{.93\sqrt{8}}{\sqrt{0.14}}$

= 7.03

তৃতীয় ধাপ ঃ t এর তাত্ত্বিক মান নির্ণয়ঃ

যেহেতু এক্ষেত্রে বিকল্প কল্পনা $H_A : \rho \neq 0$ তাই দ্বিপার্শ্বিক যাচাই (two tailed test) প্রযোজ্য হবে।

তাৎপর্য স্তর $\alpha = 5\% = 0.05$

$\therefore = \frac{\alpha}{2} = \frac{0.05}{2} = 0.025$

স্বাধীনতার মাত্রা v = n-2 = 10-2= 8

t এর critical value $= \pm 2.306$

চতুর্থ ধাপ ঃ সিদ্ধান্ত গ্রহণঃ

এক্ষেত্রে দেখা যায় যে t এর নির্ণীত মান 7.03 উহার তাত্ত্বিক মান -2.306 থেকে +2.306 এর মধ্যে অবস্থান করে ন। তাই 5% যথার্থতার মাত্রায় নাস্তি কল্পনা H_0 বাতিল হয়ে যাবে। সুতরাং, বলা যায় আয় (Y) এবং ভোগব্যয় (C) এর মধ্যে তাৎপর্যপূর্ণ রৈখিক সহ-সম্বন্ধ বিরাজ করছে।

(খ) নির্দিষ্ট মানের সংশ্লেষাংক যাচাই (Test about the *p*articular correlation coefficient) (Z Test)

ধরি, X ও Y দুটি চলকের n সংখ্যক জোড়া মান যথাক্রমে $(x_1, y_1), (x_2, y_2), \ldots\ldots\ldots (x_n, y_n)$ দ্বিচলকবিশিষ্ট সমগ্রক থেকে দৈবভাবে চয়ন করা হয়েছে। যদি নমুনার সংশ্লেষাংক r সমগ্রকের সংশ্লেষাংক ρ এবং সমগ্রকের কল্পিত সংশ্লেষাংক ρ_0 হলে–

নাস্তি কল্পনা $H_0: \rho = \rho_0$

উক্ত নাস্তি কল্পনা যাচাইয়ের জন্য Test Statistic হবেঃ

$Z = (d-v)\sqrt{n-3}$

যেখানে, $d = \frac{1}{2}\log_e \frac{1+r}{1-r}$

এবং, $v = \frac{1}{2}\log_e \frac{1+\rho_0}{1-\rho_0}$

Z কে নিম্নোক্তভাবেও লেখা যায়,

$$Z = \left(\frac{1}{2}\log_e \frac{1+r}{1-r} - \frac{1}{2}\log_e \frac{1+\rho_0}{1-\rho_0}\right)\sqrt{n-3}$$

যা শূন্য গড় এবং ১ ভেদাঙ্ক বিশিষ্ট পরিমিত বিন্যাস অনুসরণ করে।

উদাহরণ ঃ দৈবভাবে চয়ন করা 20 জোড়া তথ্যমানের সংশ্লেষাংক পাওয়া গেল r = 0.884. সমগ্রকের সংশ্লেষাংক ρ = 0.92 এর সাথে নমুনা সংশ্লেষাংক r এর তাৎপর্যপূর্ণ পার্থক্য আছে কিনা তা 5% যথার্থতার মাত্রায় যাচাই করুন।

সমাধানঃ দেওয়া আছে, নমুনার আকার, $n = 20$
নমুনার সংশ্লেষাংক, $r = 0.884$
সমগ্রকের সংশ্লেষাংক, $\rho = 0.92$

প্রথম ধাপ ঃ প্রতিপাদ্য বা কল্পনা প্রণয়ন ঃ
নাস্তি কল্পনা, $H_0 : \rho = 0.92$ (অর্থাৎ নমুনা সংশ্লেষাংক ও সমগ্রকের সংশ্লেষাংক সমান)
বিকল্প কল্পনা, $H_A : \rho \neq 0.92$ (অর্থাৎ নমুনা সংশ্লেষাংকের সাথে সমগ্রকের সংশ্লেষাংকের পার্থক্য আছে)

দ্বিতীয় ধাপ ঃ z এ মান নির্ণয়ঃ

এক্ষেত্রে Test Statistic হবে,

$$z = (d-v)\sqrt{n-3}$$
$$= (1.39378-1.5890)\sqrt{20-3}$$
$$= -0.8049$$

যেখানে,

$$d = \frac{1}{2}\log_e\frac{1+r}{1-r}$$
$$= \frac{1}{2}\log_e\frac{1+0.884}{1-0.884}$$
$$= 1.39678$$
$$v = \frac{1}{2}\log_e\frac{1+\rho_0}{1-\rho_0}$$
$$= \frac{1}{2}\log_e\frac{1+0.90}{1-0.90}$$
$$= 1.5890$$

তৃতীয় ধাপ ঃ z এর তাত্ত্বিক মান নির্ণয়ঃ
এক্ষেত্রে বিকল্প কল্পনা $\rho \neq 0.92$ তাই দ্বিপার্শ্বিক যাচাই (two tailed test) প্রযোজ্য হবে। $\alpha = 5\% = 0.05$ যথার্থতার মাত্রায় two tailed test -এ পরিমিত সারণী অনুসারে z এর তালিকা মান $= \pm 1.96$।

চতুর্থ ধাপ ঃ সিদ্ধান্ত গ্রহণঃ
যেহেতু z এর নির্ণীত মান (-0.8049) উহার তাত্ত্বিক মান -1.96 থেকে +1.96 এর মাঝে অবস্থান করছে, তাই 5% যথার্থতার মাত্রায় নাস্তি কল্পনা H_0 গ্রহণযোগ্য। সুতরাং, বলা যায় যে, সমগ্রকের সংশ্লেষাংকের সাথে নমুনার সংশ্লেষাংকের কোন তাৎপর্যপূর্ণ পার্থক্য নেই।

(গ) দুটি সংশ্লেষাংকের পার্থক্য যাচাই (Z Test)
(Test of difference between two coefficients of correlation)
ধরি, যে দুটি সমগ্রক থেকে দুটি নমুনা দৈবভাবে সংগ্রহ করা হয়েছে, তাদের সংশ্লেষাংক যথাক্রমে ρ_1 ও ρ_2। নমুনা দুটির সংশ্লেষাংক যথাক্রমে r_1 ও r_2। নমুনা দুটির আকার n_1 ও n_2 হলে দুটি সংশ্লেষাংকের পার্থক্য যাচাইয়ের ক্ষেত্রে,
নাস্তি কল্পনা, $H_0 : \rho_1 = \rho_2$
বিকল্প কল্পনা, $H_A : \rho_1 \neq \rho_2$ (two tailed test)
one tailed test এর ক্ষেত্রে বিকল্প কল্পনা $H_A : \rho_1 > \rho_2$ কিংবা $\rho_1 < \rho_2$ হবে।
এক্ষেত্রে test statistic হবে,

$$Z = \frac{Z_1 - Z_2}{\sqrt{\frac{1}{(n_1-3)} + \frac{1}{(n_2-3)}}}$$

যেখানে, $Z_1 = \frac{1}{2}\log_e\frac{1+r_1}{1-r_1}$

এবং $Z_2 = \frac{1}{2}\log_e \frac{1+r_2}{1-r_2}$

উদাহরণ ঃ 12 এবং 16 আকার বিশিষ্ট দুটি দৈব নমুনা থেকে প্রাপ্ত সংশ্লেষাংক যথাক্রমে $r_1 = 0.48$ এবং $r_2 = 0.44$। এই সংশ্লেষাংক দুটির মধ্যে কোন তাৎপর্যপূর্ণ পার্থক্য আছে কিনা তা 1% যথার্থতার মাত্রায় যাচাই কর।

সমাধান ঃ দেওয়া আছে

নমুনার আকার,	$n_1 = 12$	$n_2 = 16$
নমুনার সংশ্লেষাংক,	$r_1 = 0.48$	$r_2 = 0.44$

প্রথম ধাপ ঃ প্রতিপাদ্য বা কল্পনা প্রণয়ন ঃ

নাস্তি কল্পনা, $H_0 : \rho_1 = \rho_2$ (অর্থাৎ সমগ্রক দুটির সংশ্লেষাংক সমান)

বিকল্প কল্পনা, $H_A : \rho_1 \neq \rho_2$ (অর্থাৎ সমগ্রক দুটির সংশ্লেষাংকের মধ্যে তাৎপর্যপূর্ণ পার্থক্য রয়েছে)

দ্বিতীয় ধাপ ঃ Z এর মান নির্ণয়ঃ

এক্ষেত্রে Test Statistic হবে,

$$Z = \frac{Z_1 - Z_2}{\sqrt{\frac{1}{(n_1-3)} + \frac{1}{(n_2-3)}}}$$

যেখানে, $Z_1 = \frac{1}{2}\log_e \frac{1+r_1}{1-r_1}$

$$= \frac{1}{2}\log_e \frac{1+0.48}{1-0.48}$$

$$= 0.52298$$

এবং $Z_2 = \frac{1}{2}\log_e \frac{1+r_2}{1-r_2}$

$$= \frac{1}{2}\log_e \frac{1+0.44}{1-0.44}$$

$$= 0.47223$$

$$\therefore Z = \frac{Z_1 - Z_2}{\sqrt{\frac{1}{(n_1-3)} + \frac{1}{(n_2-3)}}}$$

$$= \frac{0.52298 - 0.47223}{\sqrt{\frac{1}{(12-3)} + \frac{1}{(16-3)}}}$$

$$= \frac{0.05075}{0.433629}$$

$$= 0.1170$$

তৃতীয় ধাপ ঃ Z এর তাত্ত্বিক মান নির্ণয়ঃ

এক্ষেত্রে বিকল্প কল্পনা $H_A : \rho_1 \neq \rho_2$ তাই দ্বিপার্শ্বিক যাচাই (two tailed test) প্রযোজ্য হবে। $\alpha = 1\% = 0.01$ যথার্থতার মাত্রা অনুসারে two tailed test-এ পরিমিত সারণী অনুসারে Z এর তাত্ত্বিক মান $= \pm 2.58$।

চতুর্থ ধাপ ঃ সিদ্ধান্ত গ্রহণঃ

যেহেতু Z এর নির্ণীত মান 0.1170 উহার তাত্ত্বিক মান -2.58 থেকে +2.58 এর মধ্যে অবস্থান করে, তাই 1% যথার্থতার মাত্রায় নাস্তি কল্পনা H_0 গ্রহণযোগ্য। সুতরাং, বলা যায় যে, সমগ্রকের সংশ্লেষাংক দুটির মধ্যে কোন তাৎপর্যপূর্ণ পার্থক্য নেই। ফলে, নমুনা দুটির সংশ্লেষাংকের মধ্যেও তাৎপর্যপূর্ণ পার্থক্য নেই।

সারসংক্ষেপ

শূন্য সংশ্লেষাংক যাচাইয়ের সূত্র হচ্ছেঃ $t = \frac{r\sqrt{(n-2)}}{\sqrt{1-r^2}}$ এবং

নির্দিষ্ট মানের সংশ্লেষাংক যাচাইয়ের সূত্র হচ্ছেঃ $Z = \left(\frac{1}{2}\log_e\frac{1+r}{1-r} - \frac{1}{2}\log_e\frac{1+p_0}{1-p_0}\right)\sqrt{n-3}$ এবং

দুটি সংশ্লেষাংকের পার্থক্য যাচাইয়ের সূত্র হচ্ছেঃ $Z = \frac{Z_1 - Z_2}{\sqrt{\frac{1}{(n_1-3)} + \frac{1}{(n_2-3)}}}$

পাঠ ৮.৫ ভেদাংক সম্পর্কিত কল্পনা যাচাই

Test for Variance

উদ্দেশ্য

এ পাঠ শেষে আপনি-

- ভেদাংকের উপর একটি নমুনার পরীক্ষা পদ্ধতি বর্ণনা করতে পারবেন।
- ভেদাংকের উপর দুটি নমুনার পরীক্ষা ব্যাখ্যা করতে পারবেন।

ভেদাংক সম্পর্কিত কল্পনা যাচাই

Test for Variance

ভেদাংক সম্পর্কিত কল্পনা যাচাই নিম্নোক্তভাবে করা যেতে পারে :

ক. ভেদাংকের উপর একটি নমুনার পরীক্ষা (Test for single population variance) (χ^2 Test)

খ. ভেদাংকের উপর দুইটি নমুনার পরীক্ষা (Test for two population variance) (F Test)

ক. ভেদাংকের উপর একটি নমুনার পরীক্ষা (Test for single population variance)

উদাহরণ-: 10টি দ্রব্যের ওজন নিম্নে দেয়া হল–

38, 40, 45, 53, 47, 43, 55, 48, 52, 49

সকল দ্রব্যের ওজনের ভেদাংক 20 কিনা 1% তাৎপর্য স্তরে পরীক্ষা করুন।

সমাধান ঃ

নাস্তি অনুমান, $H_0 : \sigma^2 = 20$ অর্থাৎ সকল দ্রব্যের ওজনের ভেদাংক 20

বিকল্প অনুমান, $H_A : \sigma^2 \neq 20$ অর্থাৎ সকল দ্রব্যের ওজনের ভেদাংক 20 নয়।

এক্ষেত্রে test statistic হবে-

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

নমুনার ভেদাংক (s^2) নির্ণয়:-

ওজন (X)	X^2
38	1444
40	1600
45	2025
53	2809
47	2209
43	1849
55	3025
48	2304
52	2704
49	2401
$\Sigma x=470$	$\Sigma x^2 = 22370$

এখানে নমুনার আকার, n = 10

নমুনার ভেদাংক, $S^2 = \dfrac{\Sigma x^2 - \dfrac{(\Sigma x)^2}{n}}{n-1}$

$$= \frac{22370 - \frac{(470)^2}{10}}{10-1}$$

$$= \frac{22370 - 22090}{9}$$

$$= \frac{280}{9} = 31.11$$

$$\therefore \chi^2 = \frac{(10-1)31.11}{20}$$

$$= \frac{9 \times 31.11}{20}$$

$$= 14$$

$\therefore \chi^2$ এর নির্ণীত মান = 14

যেহেতু $\sigma^2 \neq 20$

অতএব পরীক্ষাটি দুই দিক বিশিষ্ট।

এখানে তাৎপর্য স্তর, $\alpha = 1\% = 0.01$

$\therefore \frac{\alpha}{2} = \frac{0.01}{2} = 0.005$

এবং $1 - \frac{\alpha}{2} = 1 - 0.005 = 0.995$

স্বাধীনতার মাত্রা, $v = n-1$

$= 10-1 = 9$

$\therefore \chi^2$ এর বামদিকের Critical মান, $\chi^2_{0.995} = 1.74$

এবং ডানদিকের Critical মান $\chi^2_{0.005} = 23.59$

যেহেতু χ^2 এর নির্ণীত মান (14) তার Critical মানদ্বয়ের মধ্যে আছে। তাই নমুনা গ্রহণীয় এলাকায় পড়ে। অর্থাৎ H_0 গ্রহণীয়। অতএব বলা যায় যে, সকল দ্রব্যের ভেদাংক 20।

খ. ভেদাংকের উপর দুইটি নমুনার পরীক্ষা (Test for two population variance)

উদাহরণ-: 16 এবং 25 ভেদাংক বিশিষ্ট দুটি পরিমিত বিন্যাসকৃত তথ্য সমগ্রক হতে 9 এবং 12 আয়তনের দুটি নমুনা রেখা হল। নমুনার ভেদাংক গুলো যদি 20 এবং 8 হয় তবে (i) 0.05 (ii) 0.01 তাৎপর্য মাত্রায় দ্বিতীয় নমুনার চেয়ে ১ম নমুনার ভেদাংক যথেষ্ট বড় কিনা যাচাই করুন।

সমাধানঃ দেওয়া আছে,

১ম নমুনা–নমুনার আকার, $n_1 = 9$

নমুনার ভেদাংক, $s_1^2 = 20$

সমগ্রকের ভেদাংক, $\sigma_1^2 = 16$

২য় নমুনা– নমুনার আকার, $n_2 = 12$

নমুনার ভেদাংক, $s_2^2 = 8$

সমগ্রকের ভেদাংক, $\sigma_2^2 = 25$

নাস্তি অনুমান $H_0 : \sigma_1^2 = \sigma_2^2$ অর্থাৎ দ্বিতীয় নমুনার চেয়ে প্রথম নমুনার ভেদাংক বড় নয়।

বিকল্প অনুমান, $H_A : \sigma_1^2 > \sigma_2^2$ অর্থাৎ দ্বিতীয় নমুনার চেয়ে প্রথম নমুনার ভেদাংক যথেষ্ট বড়।

এক্ষেত্রে test statistic হবে−

$$F = \frac{\dfrac{n_1 s_1^2}{(n_1-1)\sigma_1^2}}{\dfrac{n_2 s_2^2}{(n_2-1)\sigma_2^2}}$$

$$= \frac{\dfrac{9\times 20}{(9-1)16}}{\dfrac{12\times 8}{(12-1)25}}$$

$$= \frac{\dfrac{180}{128}}{\dfrac{96}{275}}$$

$$= \frac{1.41}{0.35} = 4.03$$

$\therefore$ F এর নির্ণীত মান = 4.03

(i) দেয়া আছে তাৎপর্য মাত্রা, $\alpha = 0.05$

স্বাধীনতার মাত্রাদ্বয়-

$v_1 = n_1-1 = 9-1=8$

$v_2 = n_2-1 = 12-1=11$

$\therefore$ F এর Critical মান = 2.95

যেহেতু F এর নির্ণীত মান (4.03) উহার Critical মান (2.95) অপেক্ষা বড়। তাই নমুনা বর্জনীয় এলাকায় পড়ে। অর্থাৎ H_0 বর্জনীয়। অতএব বলা যায় যে, দ্বিতীয় নমুনার চেয়ে প্রথম নমুনার ভেদাংক যথেষ্ট বড়।

(ii) দেয়া আছে তাৎপর্য মাত্রা, $\alpha = 0.01$

স্বাধীনতার মাত্রাদ্বয়-

$v_1 = n_1-1 = 9-1=8$

$v_2 = n_2-1 = 12-1=11$

$\therefore$ F এর Critical মান = 4.74

যেহেতু F এর নির্ণীত মান (4.03) উহার Critical মান (4.74) অপেক্ষা ছোট। তাই নমুনা গ্রহণীয় এলাকায় পড়ে। অর্থাৎ H_0 গ্রহণীয়। অতএব বলা যায় যে, দ্বিতীয় নমুনার চেয়ে প্রথম নমুনার ভেদাংক যথেষ্ট বড় নয়।

সারসংক্ষেপ

ভেদাংকের উপর একটি নমুনার পরীক্ষার সূত্র হচ্ছে: $S^2 = \dfrac{\Sigma x^2 - \dfrac{(\Sigma x)^2}{n}}{n-1}$ এবং ভেদাংকের উপর দুইটি নমুনার পরীক্ষার সূত্র হচ্ছে: $F = \dfrac{\dfrac{n_1 s_1^2}{(n_1-1)\sigma_1^2}}{\dfrac{n_2 s_2^2}{(n_2-1)\sigma_2^2}}$

পাঠোত্তর মূল্যায়ন

রচনামূলক প্রশ্ন

১। যথার্থতা যাচাই বলতে কি বুঝেন? এর সাথে সম্পর্কযুক্ত বিভিন্ন শব্দাবলীর সংজ্ঞা লিখুন।

২। যথার্থতা যাচাই এর বিভিন্ন ধাপসমূহ উল্লেখ করুন।

৩। নাস্তি কল্পনা, বিকল্প কল্পনা, যাচাই ষ্ট্যাটিস্টিক, বাতিল এলাকা, গ্রহণীয় এলাকা, ১ম ও ২য় প্রকার ভুল, যথার্থতার মাত্রা এবং স্বাধীনতার মাত্রার উদাহরণসহ সংজ্ঞা দিন।

৪। পরিমিত যাচাই, t-যাচাই, χ২–যাচাই এবং F-যাচাই এর বর্ণনা করুন।

৫। পরিমিত যাচাই এবং t-যাচাই-এর মধ্যে পার্থক্য উল্লেখ করুন।

৭। ১৩ ভেদাংক বিশিষ্ট একটি তথ্যবিশ্ব থেকে ৩৫ আকারের একটি নমুনা নিয়ে তার গড় পাওয়া গেল ৩৭। উক্ত তথ্যবিশ্বের গড় ৩২ কিনা তা কিভাবে যাচাই করবেন?

৮। ১৫ ও ২০ আকার বিশিষ্ট দুটি নমুনার ভেদাংক পাওয়া গেল যথাক্রমে ২১ ও ২৭। নমুনা দুটি একই ভেদাংক বিশিষ্ট পরিমিত তথ্যবিশ্ব থেকে নেয়া হয়েছে কিনা যাচাই করুন।

১০। একটি ঔষধ প্রয়োগ করে দেখা গেল যে, ১৫০ জন রোগীর ৪০% রোগী আরোগ্য লাভ করে এবং ২০০ জন রোগী যাদেরকে ঔষধটি প্রয়োগ করা হয়নি তাদের ৩০% রোগী আরোগ্য লাভ করে। ঔষধটি কার্যকারিতা যথার্থ কিনা যাচাই করুন।

১১। ১৫, ২০ ও ২৫ আকার বিশিষ্ট তিনটি নমুনা পরিমিত তথ্যবিশ্ব থেকে নেওয়া হয়েছে এবং তাদের ভেদাংক যথাক্রমে ৩০, ৩৭ ও ৪১। নমুনা তিনটির ভেদাংক সমমাত্রিক কিনা যাচাই করুন।

১২। দুটি টিউটোরিয়াল কেন্দ্র থেকে যথাক্রমে ১০ ও ১৫ জন ছাত্র দৈব পদ্ধতিতে নির্বাচন করে তাদের বয়স নিম্নে দেওয়া হল।

ছাত্রদের বয়স (বৎসর)

১ম টিউটোরিয়াল কেন্দ্র	১৫, ১৪, ১৭, ১৯, ১৫, ২০, ১৯, ১৮, ১৭, ২৬
২য় টিউটোরিয়াল কেন্দ্র	১৫, ২০, ১৭, ১৯, ২১, ২২, ১৮, ১৬, ১৪, ১৮, ২০, ১৯, ১৭, ১৮, ১৫

টিউটোরিয়াল কেন্দ্র দুটির ছাত্রদের গড় বয়সের তুলনা যাচাই করুন।