


ভেদাংক বিশ্লেষণ Analysis of Variance

ভূমিকা

পরিসংখ্যান অনুসন্ধানে আমরা এক বা একাধিক সমগ্রক থেকে দুই বা ততোধিক নমুনা সংগ্রহ করে থাকি। এই নমুনা সমূহের গড়ের পার্থক্য কতটুকু তাৎপর্যপূর্ণ এবং সংগৃহীত নমুনাসমূহ একই গড় বিশিষ্ট সমগ্রক থেকে গ্রহণ করা হয়েছে কিনা সেই সম্পর্কিত সিদ্ধান্ত পাওয়া যায় ভেদাঙ্ক বিশ্লেষণ (Analysis of variance) থাকে। দুই বা ততোধিক নমুনা গড়ের পার্থক্যের তাৎপর্য যাচাই করতে ‘ভেদাঙ্ক বিশ্লেষণ’ বা F-Test করা হয়।

ভেদাঙ্ক বিশ্লেষণের সূত্রপাত হয় কৃষি উৎপাদন সম্পর্কিত গবেষণা থেকে। কিন্তু বর্তমানে সমাজ বিজ্ঞান, প্রাকৃতিক বিজ্ঞান ইত্যাদি প্রতিটি সেक्टरে ভেদাঙ্ক বিশ্লেষণ প্রক্রিয়া ব্যবহৃত হচ্ছে। বর্তমানে পরিকল্পিত গবেষণার একটি ফান্ডামেন্টাল অংশ হল ভেদাঙ্ক বিশ্লেষণ বা ANOVA। গবেষণার ডিজাইন তৈরী করতে, নতুন প্রযুক্তির প্রভাব ব্যাখ্যা করতে, নীতি নির্ধারণ করতে ইত্যাদি প্রতিটি ক্ষেত্রে ANOVA ব্যবহার করা হয়।

 ইউনিট সমাপ্তির সময়	ইউনিট সমাপ্তির সর্বোচ্চ সময় ০২ সপ্তাহ
এ ইউনিটের পাঠসমূহ	
পাঠ ৯.১	: ভেদাংক বিশ্লেষণ
পাঠ ৯.২	: একমুখী শ্রেণীবিভাগ
পাঠ ৯.৩	: দ্বিমুখী শ্রেণীবিভাগ
পাঠ ৯.৪	: দ্বিচলক বিশিষ্ট নির্ভরণ মডেল
পাঠ ৯.৫	: বহুধা নির্ভরণ মডেল



উদ্দেশ্য

এ পাঠ শেষে আপনি-

- ভেদাংক বিশ্লেষণের অনুমিতিসমূহ বলতে পারবেন।
- ভেদাংক বিশ্লেষণের ব্যবহার ব্যাখ্যা করতে পারবেন।
- ভেদাংক বিশ্লেষণের প্রকারভেদ সম্পর্কে লিখতে পারবেন।

ভেদাংক বিশ্লেষণের অনুমিতিসমূহ

Assumptions of Analysis of Variance

ভেদাংক বিশ্লেষণের অনুমিতিসমূহ-

- প্রত্যেক নমুনা পরিমিত সমগ্রক হতে দৈবভাবে উত্তোলন করা হবে।
- নমুনা সমূহ সমগ্রকের বৈশিষ্ট্যকে প্রতিফলিত করবে।
- নমুনাগুলো যে সকল সমগ্রক হতে নেয়া হবে উহাদের গড় ও ভেদাংকসমূহ একই হবে। অর্থাৎ $\mu_1 = \mu_2 = \dots = \mu_n$ এবং $\sigma_1 = \sigma_2 = \dots = \sigma_n$ ।
- সমগ্রকগুলো পরস্পর স্বাধীন।

যদি কোন নির্দিষ্ট সমস্যায় উপরের অনুমিতিসমূহ মেনে না চলে তাহলে ভেদাংক বিশ্লেষণ পদ্ধতি ব্যবহার করা যায় না। সেক্ষেত্রে অপারামাত্রিক (Non-Parametric) কৌশল ব্যবহৃত হবে।

ভেদাংক বিশ্লেষণের ব্যবহার

Uses of Analysis of Variance

ভেদাংক বিশ্লেষণের ব্যবহার নিম্নে উল্লেখ করা হলঃ

- পরীক্ষণের উৎকর্ষতা বৃদ্ধিতে এটি ব্যবহৃত হয়।
- গবেষক তাঁর পরীক্ষার তাৎপর্য যাচাই করতে এটি ব্যবহার করেন।
- সমাজবিজ্ঞান, রাষ্ট্রবিজ্ঞান, কৃষি, অর্থনীতি, শিক্ষা, ব্যবসা-বাণিজ্য, চিকিৎসা শাস্ত্র প্রভৃতি ক্ষেত্রে পরীক্ষামূলক গবেষণায় এ পদ্ধতি ব্যাপকভাবে ব্যবহৃত হয়।
- এ পদ্ধতি উন্নতমানের পরীক্ষণ নকশা প্রণয়নে সাহায্য করে।
- পরীক্ষালব্ধ তথ্য বিশ্লেষণ ইহা ব্যবহৃত হয়।

ভেদাংক বিশ্লেষণের শ্রেণীবিভাগ

Classification of Analysis of Variance

ভেদাংক বিশ্লেষণের পদ্ধতিকে বিভিন্নভাবে শ্রেণীবিভাগ করা যায়। যথা :

- একমুখী শ্রেণীবিভাগ, খ. দ্বিমুখী শ্রেণীবিভাগ



সারসংক্ষেপ

সংগৃহীত তথ্যাবলির মোট ভেদাংককে ভিন্ন ভিন্ন শ্রেণিকরণের উপাদানের জন্য পৃথক পৃথকভাবে ভাগ করার পদ্ধতিকে ভেদাংক বিশ্লেষণ বলে। ANOVA-এর পূর্ণরূপটি হলো Analysis of Variance। ভেদাংক বিশ্লেষণ পরিসংখ্যানিক যাচাইয়ের মধ্যে এফ যাচাইয়ের সাথে সম্পর্কিত।

পাঠ ৯.২

একমুখী শ্রেণীবিভাগ

One Way Classification



উদ্দেশ্য

এ পাঠ শেষে আপনি-

- ভেদাংক বিশ্লেষণের একমুখী শ্রেণীবিভাগ বর্ণনা করতে পারবেন।
- ভেদাংক বিশ্লেষণের একমুখী শ্রেণীবিভাগ সম্পর্কিত সমস্যার সমাধান করতে পারবেন।

একমুখী শ্রেণীবিভাগ

One Way Classification

কোন তথ্যসারিকে একটি বিশেষ উপাদানের বা বৈশিষ্ট্যের ভিত্তিতে শ্রেণীবদ্ধ করা হলে তাকে একমুখী শ্রেণীবিভাগ বলে। যে মডেলে একমুখী শ্রেণীবিভাগ এর সাহায্যে ভেদাংক বিশ্লেষণ করা হয় তাকে একমুখী শ্রেণীবিভাগ মডেল বলে।

উদাহরণস্বরূপ : সমগুণ সম্পন্ন কিছু জমিতে তিন ধরনের ধান চাষ করা হল। তিন ধরনের ধান সম বা অসম সংখ্যক জমিতে চাষ করে উৎপাদনের পরিমাণ সংগ্রহ করা হল। এ ধরনের তথ্যই একমুখী শ্রেণীকৃত তথ্য।

ধরি কোন তথ্যসারিকে k সংখ্যক শ্রেণীতে বিভক্ত করা হল এবং উক্ত শ্রেণীগুলো হতে যথাক্রমে n_1, n_2, \dots, n_k সংখ্যক নমুনার তথ্য নেয়া হল।

নিম্নের ধাপসমূহের সাহায্যে একমুখী শ্রেণীকৃত তথ্যের ভেদাংক বিশ্লেষণ করা হলঃ-

১ম ধাপ : অনুমান নির্ণয় – নাস্তি অনুমান ও বিকল্প অনুমান নির্ণয় করতে থাকে। এক্ষেত্রে নাস্তি অনুমান, $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$

২য় ধাপ : ভেদাংক বিশ্লেষণ সারণী (ANOVA) নির্ণয়ঃ

(i) নমুনা সমূহের সমষ্টি (T) নির্ণয় করতে হবে।

$$\text{এক্ষেত্রে } T = \sum X_1 + \sum X_2 + \dots + \sum X_k$$

এবং শুদ্ধি সংখ্যা (Correction Factor) নির্ণয় করতে হবে। এক্ষেত্রে শুদ্ধি সংখ্যা হচ্ছে $\frac{T^2}{n}$

এখানে $n = n_1 + n_2 + \dots + n_k$

(ii) মোট বর্গ সমষ্টি (Total Sum of Squares বা SST) নির্ণয় করতে হবে।

$$\begin{aligned} \text{এক্ষেত্রে } SST &= \sum X_1^2 + \sum X_2^2 + \dots + \sum X_k^2 - \frac{T^2}{n} \\ &= \sum X_i^2 - \frac{T^2}{n} \end{aligned}$$

(iii) আন্তঃশ্রেণী বর্গ সমষ্টি (Between Samples Sum of Squares বা SSB) নির্ণয় করতে হবে।

$$\begin{aligned} \text{এক্ষেত্রে } SSB &= \frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \dots + \frac{(\sum X_k)^2}{n_k} - \frac{T^2}{n} \\ &= \sum \frac{(\sum X_i)^2}{n_i} - \frac{T^2}{n} \end{aligned}$$

(iv) অন্তঃশ্রেণী বর্গ সমষ্টি (Within Samples Sum of Squares বা SSW) নির্ণয় করতে হবে।

$$\text{এক্ষেত্রে } SSW = SST - SSB$$

এমবিএ প্রোগ্রাম

$$= \sum X_i^2 - \frac{(\sum X_i)^2}{n_1} \text{ এখানে } i = 1, 2, \dots, k$$

অতএব ভেদাংক বিশ্লেষণ সারণী (Anova Table) –

ভেদের উৎস	বর্গ সমষ্টি (Sum of Square) SS	স্বাধীনতার মাত্রা (v)	গড় বর্গ সমষ্টি (Mean Square) MS	F
আন্তঃশ্রেণী (Between Samples)	$SSB = \sum \frac{(\sum X_i)^2}{n_i} - \frac{T^2}{n}$	k - 1	$MSB = \frac{SSB}{k - 1}$	$F = \frac{MSB}{MSW}$
অন্তঃশ্রেণী (Within Samples)	$SSW = SST - SSB$ $= \sum X_i^2 - \sum \frac{(\sum X_i)^2}{n_i}$	n - k	$MSW = \frac{SSW}{n - k}$	
মোট (Total)	$SST = \sum X_i^2 - \frac{T^2}{n}$	n-1		

৩য় ধাপ- সিদ্ধান্ত: উপরের সারণীতে প্রাপ্ত F এর নির্ণীত মান এবং তাত্ত্বিক (Critical) মান তুলনার মাধ্যমে নাস্তি অনুমান (H₀) গ্রহণীয় হবে, নাকি বর্জনীয় হবে তা নির্ধারণ করা যায়। যেমন F_{নির্ণীত} > F_{তাত্ত্বিক} হলে H₀ বর্জনীয় হবে। অন্যথায় H₀ গ্রহণীয় হবে।

উদাহরণ : একজন গবেষক চারটি ব্রান্ডের বাব্বের জীবনকাল নিয়ে গবেষণা করছেন, দৈবচয়ন ভিত্তিতে প্রত্যেক ব্রান্ডের তিনটি করে বাব্ব নির্বাচন করে তাদের জীবনকাল জানা গেল যে-

A	B	C	D
20	25	24	23
19	23	20	20
21	21	22	20

চারটি ব্রান্ডের বাব্বের গড় জীবনকাল সমান কিনা 5% গুরুত্বের স্তরে পরীক্ষা করুন।

সমাধান: ধরি, X₁, X₂, X₃, এবং X₄ হচ্ছে যথাক্রমে A, B, C এবং D ব্রান্ডের বাব্বের জীবনকাল নির্দেশ করে।

নাস্তি অনুমান, H₀ : μ₁ = μ₂ = μ₃ = μ₄ অর্থাৎ চারটি ব্রান্ডের বাব্বের গড় জীবনকাল সমান।

বিকল্প অনুমান, H_A : μ₁ ≠ μ₂ ≠ μ₃ ≠ μ₄ অর্থাৎ চারটি ব্রান্ডের বাব্বের গড় জীবনকাল সমান নয়।

প্রদত্ত তথ্যানুসারে নিম্নের সারণীটি তৈরী করা হলোঃ

X ₁	X ₁ ²	X ₂	X ₂ ²	X ₃	X ₃ ²	X ₄	X ₄ ²
20	400	25	625	24	576	23	529
19	361	23	529	20	400	20	400
21	441	21	441	22	484	20	400
ΣX ₁ =60	Σ X ₁ ² =1202	ΣX ₂ =69	Σ X ₂ ² =1595	ΣX ₃ = 66	Σ X ₃ ² =1460	ΣX ₄ = 63	Σ X ₄ ² =1329

এখানে নমুনার সংখ্যা, N = 12

এবং শ্রেণীর সংখ্যা, K = 4

নিম্নে 6 টি ধাপের সাহায্যে অনুমান যাচাই করা হলঃ

$$\begin{aligned} \text{ধাপ-১ : নমুনাসমূহের সমষ্টি, } T &= \Sigma X_1 + \Sigma X_2 + \Sigma X_3 + \Sigma X_4 \\ &= 60 + 69 + 66 + 63 \\ &= 258 \end{aligned}$$

$$\text{শুদ্ধিকরণ সংখ্যা, } \frac{T^2}{n} = \frac{(258)^2}{12} = 5547$$

ধাপ ২ : মোট বর্গ সমষ্টি,

$$\begin{aligned} \text{SST} &= \Sigma X_i^2 - \frac{T^2}{n} \\ &= \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2 + \Sigma X_4^2 - \frac{(258)^2}{12} \\ &= 1202 + 1595 + 1460 + 1329 - 5547 \\ &= 5586 - 5547 \\ &= 39 \end{aligned}$$

ধাপ-৩ : আন্তঃ শ্রেণী বর্গ সমষ্টি,

$$\begin{aligned} \text{SSB} &= \Sigma \frac{(\Sigma X_i)^2}{n_i} - \frac{T^2}{n} \\ &= \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \frac{(\Sigma X_2)^2}{n_2} + \frac{(\Sigma X_2)^2}{n_2} - \frac{T^2}{n} \\ &= \frac{(60)^2}{3} + \frac{(69)^2}{3} + \frac{(66)^2}{3} + \frac{(63)^2}{3} - \frac{(258)^2}{12} \\ &= \frac{1}{3} [3600 + 4761 + 4356 + 3969] - 5547 \\ &= \frac{16686}{3} - 5547 \\ &= 5562 - 5547 \\ &= 15 \end{aligned}$$

ধাপ-৪ : অন্তঃশ্রেণী বর্গসমষ্টি,

$$\begin{aligned} \text{SSW} &= \text{SST} - \text{SSB} \\ &= 39 - 15 \\ &= 24 \end{aligned}$$

ধাপ-৫ : ভেদাংক বিশ্লেষণ (ANOVA) সারণী-

ভেদের উৎস	বর্গ সমষ্টি (Sum Square) SS	স্বাধীনতার মাত্রা (v)	গড় বর্গ সমষ্টি (Mean Square) MS	$F = \frac{MSB}{MSW}$
আন্তঃ শ্রেণী (Between Sample)	SSB = 15	k-1 = 4-1 = 3	MSB = $\frac{15}{3} = 5$	$F = \frac{5}{3}$ = 1.67
অন্তঃ শ্রেণী (Within Sample)	SSW = 24	n-k = 12-4 = 8	MSW = $\frac{24}{8} = 3$	
মোট (Total)	SST = 39	n-1 = 12-1 = 11		

∴ F এর নির্ণীত মান = 1.67

এমবিএ প্রোগ্রাম

ধাপ-৫ : সিদ্ধান্ত-

দেওয়া আছে,

গুরুত্বের স্তর, $\alpha = 5\% = 0.05$

এখানে, $v_1 = k-1 = 4-1 = 3$ এবং $v_2 = n-k = 12 - 4 = 8$

\therefore F এর তাত্ত্বিক মান = 4.07

যেহেতু F এর নির্ণীত মান (1.67) উহার তাত্ত্বিক মান (4.07) অপেক্ষা ছোট। অতএব নাস্তি অনুমান গ্রহণীয়। অর্থাৎ চারটি ব্রান্ডের বাব্বের গড় জীবনকাল সমান।



সারসংক্ষেপ

কোনো পরীক্ষণ হতে প্রাপ্ত তথ্য যদি কেবলমাত্র একটি বৈশিষ্ট্যের ভিত্তিতে শ্রেণিকরণ এবং সাজানো হয়, তাকে একমুখী শ্রেণিকৃত তথ্য বলে।

পাঠ ৯.৩

দ্বি-মুখী শ্রেণিবিভাগ

Two-Way Classification



উদ্দেশ্য

এ পাঠ শেষে আপনি-

- ভেদাংক বিশ্লেষণের দ্বিমুখী শ্রেণিবিভাগ ব্যাখ্যা করতে পারবেন।
- ভেদাংক বিশ্লেষণের দ্বিমুখী শ্রেণিবিভাগ সম্পর্কিত সমস্যার সমাধান করতে পারবেন।

দ্বি-মুখী শ্রেণী বিভাগ

Two-Way Classification

কোন তথ্য সারিকে- দুটি বিশেষ উপাদান বা বৈশিষ্ট্যের ভিত্তিতে শ্রেণীবদ্ধ করা হলে তাকে দ্বি-মুখী শ্রেণীবিভাগ বলে। এক্ষেত্রে প্রতিটি তথ্যমানকে দুটো-বৈশিষ্ট্যের যুক্তফল হিসাবে বিবেচনা করা হয়। মনে করি ক, উপাদানের p-সংখ্যক শ্রেণী এবং খ উপাদানের q সংখ্যক শ্রেণী আছে। ফলে প্রত্যেক তথ্যকে X_{ij} দ্বারা চিহ্নিত করা যায়। যেখান $i = 1, 2, \dots, p$ এবং $j = 1, 2, \dots, q$ আরো মনে করি নমুনার সংখ্যা n। নিম্নের ধাপসমূহের সাহায্যে কোন দ্বিমুখী শ্রেণীবিভাগের ক্ষেত্রে অনুমান-পরীক্ষা করা যায়।।

ধাপ-১ : নমুনাসমূহের সমষ্টি,

$$T = \sum_{i=1}^p \sum_{j=1}^q X_{ij}$$

$$= X_{11} + X_{12} + \dots + X_{pq}$$

$$\therefore \text{শুদ্ধি পদ} = \frac{T^2}{n}$$

ধাপ-২ : মোট বর্গ সমষ্টি, SST

$$= \sum_{i=1}^p \sum_{j=1}^q X_{ij}^2 - \frac{T^2}{n}$$

$$= X_{11}^2 + X_{12}^2 + \dots + X_{pq}^2 - \frac{T^2}{n}$$

ধাপ - ৩ : ক উপাদানগুলোর মধ্যকার বর্গসমষ্টি,

$$SS_1 = \frac{\sum_{i=1}^p X_i^2}{q} - \frac{T^2}{n}$$

$$= \frac{X_1^2 + X_2^2 + \dots + X_p^2}{q} - \frac{T^2}{n}$$

ধাপ-৪ : উপাদানগুলোর মধ্যকার বর্গ সমষ্টি,

$$SS_2 = \frac{\sum_{j=1}^q X_j^2}{p} - \frac{T^2}{n}$$

$$= \frac{X_1^2 + X_2^2 + \dots + X_q^2}{p} - \frac{T^2}{n}$$

ধাপ-৫: অবশিষ্ট (ত্রুটি) বর্গ সমষ্টি,

এমবিএ প্রোগ্রাম

$$SS_E = SST - SS_1 - SS_2$$

ধাপ-৬ : ভেদাংক বিশ্লেষণ (ANOVA) সারণী-

ভেদের উৎস	বর্গ সমষ্টি (Sum Square) SS	স্বাধীনতার মাত্রা (v)	গড় বর্গ সমষ্টি (Mean Square) MS	$F = \frac{MS_i}{MS_E}$
ক উপাদানের প্রকারভেদ	$SS_1 = \frac{\sum_{i=1}^p X_i^2}{q} - \frac{T^2}{n}$	$v_1 = p - 1$	$MS_1 = \frac{SS_1}{v_1}$	$F = \frac{MS_1}{MS_E}$
খ উপাদানের প্রকারভেদ	$SS_2 = \frac{\sum_{j=1}^q X_j^2}{p} - \frac{T^2}{n}$	$v_2 = q - 1$	$MS_2 = \frac{SS_2}{v_2}$	$F = \frac{MS_2}{MS_E}$
অবশিষ্ট (ক্রটি)	$SS_E = SST - SS_1 - SS_2$	$v_3 = (p - 1)(q - 1)$	$MS_E = \frac{SS_E}{v_3}$	
মোট	$SST = \sum_{i=1}^p \sum_{j=1}^q X_{ij}^2 - \frac{T^2}{n}$	$pq - 1$		

ধাপ-৭ : সিদ্ধান্ত-

উপরের সারণীতে প্রাপ্ত F-এর নির্ণীত মান উহার তাত্ত্বিক মান অপেক্ষা বড় হলে নাস্তি অনুমান (H_0) বর্জনীয়। অন্যথায় H_0 গ্রহণীয় হবে।

উদাহরণ : চারজন শ্রমিক কর্তৃক তিনটি ভিন্ন ধরণের মেশিনের সাহায্যে দৈনিক উৎপাদনের পরিমাণ নিম্নে উল্লেখ করা হলোঃ

মেশিনের প্রকারভেদ	শ্রমিক			
	A	B	C	D
1	4.5	6.4	7.2	6.7
2	8.8	7.8	9.6	7.0
3	5.9	6.8	5.7	5.2

1% গুরুত্বের স্তরে পরীক্ষা করুন-

ক. মেশিনগুলোর গড় উৎপাদন ক্ষমতা সমান কি না?

খ. শ্রমিকদের গড় উৎপাদন ক্ষমতা সমান কি না?

সমাধানঃ ধরি মেশিন ও শ্রমিকের উৎপাদন যথাক্রমে X_i এবং X_j দ্বারা নির্দেশ করা হয়েছে।

ক)

নাস্তি অনুমান, $H_0 : \mu_1 = \mu_2 = \mu_3$ অর্থাৎ মেশিনগুলোর গড় উৎপাদন ক্ষমতা সমান।

বিকল্প অনুমান, $H_A : \mu_1 \neq \mu_2 \neq \mu_3$ অর্থাৎ মেশিনগুলোর গড় উৎপাদন ক্ষমতা সমান নয়।

খ) নাস্তি অনুমান, $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ অর্থাৎ শ্রমিকদের গড় উৎপাদন ক্ষমতা সমান।

বিকল্প অনুমান, $H_A : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$ অর্থাৎ শ্রমিকদের গড় উৎপাদন ক্ষমতা সমান নয়।

এখানে

নমুনার সংখ্যা, $n = 12$

মেশিনের ক্ষেত্রে শ্রেণীর সংখ্যা, $p = 3$

শ্রমিকের ক্ষেত্রে সংখ্যা, $q = 4$

প্রদত্ত তথ্যানুসারে সারণী নিম্নরূপ :-

মেশিনের প্রকারভেদ	শ্রমিক				মোট (X_i)
	A	B	C	D	
1	4.5	6.4	7.2	6.7	24.8
2	8.8	7.8	9.6	7.0	33.2
3	5.9	6.8	5.7	5.2	23.6
মোট (X_j)	19.2	21.0	22.5	18.9	$\Sigma\Sigma X_{ij} = 81.6$

নিম্নের ধাপসমূহের সাহায্যে অনুমান পরীক্ষা করা হলঃ

$$\text{ধাপ-১ : নমুনাসমূহের সমষ্টি, } T = \sum_{i=1}^3 \sum_{j=1}^4 X_{ij} \\ = 81.6$$

$$\therefore \text{শুদ্ধি সংখ্যা} = \frac{T^2}{n} \\ = \frac{(81.6)^2}{12} \\ = \frac{6658.56}{12} \\ = 554.88$$

ধাপ - ২ : মোট বর্গের সমষ্টি,

$$\text{SST} = \sum_i \sum_j X_{ij}^2 - \frac{T^2}{n} \\ = \{(4.5)^2 + (8.8)^2 + (5.9)^2 + (6.4)^2 + (7.8)^2 + (6.8)^2 + (7.2)^2 + (9.6)^2 + (5.7)^2 + (6.7)^2 + (7.0)^2 + (5.2)^2\} - \\ \frac{(81.6)^2}{12} \\ = \{20.25 + 77.44 + 34.81 + 40.96 + 60.84 + 46.24 + 51.84 + 92.16 + 32.49 + \\ 44.89 + 49 + 27.04\} - 554.88 \\ = 577.96 - 554.88 \\ = 23.08$$

ধাপ - ৩ : মেশিনের প্রকারভেদের মধ্যকার বর্গসমষ্টি

$$\text{SS}_1 = \frac{\sum_{i=1}^p X_i^2}{q} - \frac{T^2}{n} \\ = \frac{(24.8)^2 + (33.2)^2 + (23.6)^2}{4} - \frac{(81.6)^2}{12} \\ = \frac{615.04 + 1102.24 + 556.96}{4} - 554.88$$

এমবিএ প্রোগ্রাম

$$\begin{aligned}
 &= \frac{2274.24}{4} - 554.88 \\
 &= 568.56 - 554.88 \\
 &= 13.68
 \end{aligned}$$

ধাপ - ৪ : শ্রমিকের প্রকারভেদ মধ্যকার বর্গসমষ্টি,

$$\begin{aligned}
 SS_2 &= \frac{\sum_{j=1}^q X_j^2}{p} - \frac{T^2}{n} \\
 &= \frac{(19.2)^2 + (21)^2 + (22.5)^2 + (18.9)^2}{3} - \frac{(81.6)^2}{12} \\
 &= \frac{368.64 + 441 + 506.25 + 357.21}{3} - 554.88 \\
 &= \frac{1673.10}{3} - 554.88 \\
 &= 557.70 - 554.88 \\
 &= 2.82
 \end{aligned}$$

ধাপ - ৫ : অবশিষ্ট (ক্রটির) বর্গ সমষ্টি,

$$\begin{aligned}
 SS_E &= SST - SS_1 - SS_2 \\
 &= 23.08 - 13.68 - 2.82 \\
 &= 6.58
 \end{aligned}$$

ধাপ - ৬ : ভেদাংক বিশ্লেষণ সারণী-

ভেদের উৎস	বর্গ সমষ্টি (Sum Square) SS	স্বাধীনতার মাত্রা (v)	গড় বর্গ সমষ্টি (Mean Square) MS	F
মেশিন	$SS_1 = 13.68$	$v_1 = p - 1$ $= (3-1)$ $= 2$	$MS_1 = \frac{SS_1}{v_1}$ $= \frac{13.68}{2}$ $= 6.84$	$F = \frac{MS_1}{MS_E}$ $= \frac{6.84}{1.10}$ $= 6.22$
শ্রমিক	$SS_2 = 2.82$	$v_2 = q - 1 = (4-1) = 3$	$MS_2 = \frac{SS_2}{v_2}$ $= \frac{2.82}{3}$ $= 0.94$	$F = \frac{MS_2}{MS_E}$ $= \frac{0.94}{1.10}$ $= 0.85$
অবশিষ্ট (ক্রটি)	$SS_E = 6.58$	$v_3 = (p - 1)(q - 1)$ $= (3-1)(4-1) = 6$	$MS_E = \frac{SS_E}{v_3}$ $= \frac{6.58}{6} = 1.10$	
মোট	$SST = 23.08$	$pq - 1 = 3 \times 4 - 1 = 11$		

ধাপ-৭ : সিদ্ধান্ত-

ক) দেওয়া আছে,

গুরুত্বের স্তর, $\alpha = 1\% = 0.01$

স্বাধীনতার মাত্রা দ্বয় : $v_1 = 2$ এবং $v_3 = 6$

\therefore F এর তাত্ত্বিক মান = 10.9

যেহেতু F এর নির্ণীত মান (6.22) উহার তাত্ত্বিক মান (10.09) অপেক্ষা ছোট। অতএব নাস্তি অনুমান (H_0) গ্রহণীয়। অর্থাৎ মেশিনগুলোর গড় উৎপাদন ক্ষমতা সমান।

খ) দেওয়া আছে-

গুরুত্বের স্তর, $\alpha = 1\% = 0.01$

স্বাধীনতার মাত্রা দ্বয় : $v_2 = 3$ এবং $v_3 = 6$

\therefore F এর তাত্ত্বিক মান = 9.78

যেহেতু F এর নির্ণীত মান (0.85) উহার তাত্ত্বিক মান (9.78) অপেক্ষা ছোট। অতএব নাস্তি অনুমান (H_0) গ্রহণীয়। অর্থাৎ শ্রমিকদের গড় উৎপাদন ক্ষমতা সমান।



সারসংক্ষেপ

কোনো পরীক্ষণ হতে প্রাপ্ত তথ্য যদি দুটি বৈশিষ্ট্যের ভিত্তিতে শ্রেণিকরণ এবং সাজানো হয় তাকে দ্বিমুখী শ্রেণিকরণ বলে।



উদ্দেশ্য

এ পাঠ শেষে আপনি-

- দ্বিচলক বিশিষ্ট নির্ভরণ মডেল লিখতে পারবেন।
- দ্বিচলক বিশিষ্ট নির্ভরণ মডেল সম্পর্কিত সমস্যার সমাধান করতে পারবেন।

ভেদাংক বিশ্লেষণ : দ্বি-চলক বিশিষ্ট নির্ভরণ মডেল

Analysis of Variance : Two Variable Regression Model

ধরি দ্বিচলক বিশিষ্ট নির্ভরণ মডেল-

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

∴ দ্বিচলক বিশিষ্ট নির্ভরণ সমীকরণ-

$$\hat{Y} = \beta_1 + \beta_2 X$$

নির্ভরণ মডেলের ঢালের পরীক্ষা ভেদাংক বিশ্লেষণের সাহায্যে করা যায়। নিম্নে ব্যাখ্যা করা হলঃ

ধাপ-১ : নির্ভরণ সমীকরণের ঢাল নির্ণয়

$$\begin{aligned} \hat{\beta}_2 &= \frac{\Sigma(X - \bar{X})(Y - \bar{Y})}{\Sigma(X - \bar{X})^2} \\ &= \frac{\Sigma XY - \frac{\Sigma X \Sigma Y}{n}}{\Sigma X^2 - \frac{(\Sigma X)^2}{n}} \end{aligned}$$

ধাপ - ২ : মোট ভেদ নির্ণয় :

$$\begin{aligned} SST &= \Sigma(Y - \bar{Y})^2 \\ &= \Sigma Y^2 - \frac{(\Sigma Y)^2}{n} \end{aligned}$$

ধাপ - ৩ : ব্যাখ্যায়িত ভেদ নির্ণয় :

$$\begin{aligned} SSR &= \Sigma(\hat{Y} - \bar{Y})^2 \\ &= \hat{\beta}_2^2 \Sigma(X - \bar{X})^2 \\ &= \hat{\beta}_2^2 \left\{ \Sigma X^2 - \frac{(\Sigma X)^2}{n} \right\} \end{aligned}$$

ধাপ - ৪ : অব্যাক্ষায়িত ভেদ নির্ণয় :

$$\begin{aligned} SSE &= \Sigma(Y - \hat{Y})^2 \\ &= SST - SSR \end{aligned}$$

ধাপ - ৫ : ভেদাংক বিশ্লেষণ (ANOVA) সারণীঃ

ভেদের উৎস	বর্গ সমষ্টি (Sum Square) SS	স্বাধীনতার মাত্রা (v)	গড় বর্গ সমষ্টি (Mean Square) MS	F
নির্ভরণ (R)	$SSR = \hat{\beta}_2^2 \Sigma (X - \bar{X})^2$	$k-1$ $= 2-1$ $= 1$	$MSB = \frac{SSR}{1}$	$F = \frac{MSR}{MSE}$
ত্রুটি (E)	$SSE = \Sigma (Y - \hat{Y})^2$	$n-k$ $= n-2$	$MSE = \frac{SSE}{n-2}$	
মোট (T)	$SST = \Sigma (Y - \bar{Y})^2$	$n-1$		

ধাপ - ৬ : সিদ্ধান্ত-

যদি F এর নির্ণীতমান উহার তাত্ত্বিক অপেক্ষা বড় হয় তাহলে নাস্তি অনুমান বর্জনীয়। অন্যথায় নাস্তি অনুমান গ্রহণীয় হবে।

উদাহরণ : নিম্নে আয় ও ভোগের তথ্য দেয়া হলঃ

আয়	৮	১০	১২	১৪	১৬	১৮	২০	২২	২৪	২৬
ভোগ	৭	৭	৯	১০	১১	১২	১২	১৪	১৬	১৫

নির্ভরণ মডেলের ঢাল শূন্য কিনা Anova টেবিলের সাহায্যে পরীক্ষা করুন (যখন $\alpha = 5\%$)

সমাধান : ধরি, আয় ও ভোগকে যথাক্রমে X এবং Y দ্বারা নির্দেশ করা হয়েছে।

নাস্তি অনুমান, $H_0 : \beta_2 = 0$ অর্থাৎ নির্ভরণ মডেলের ঢাল শূন্য।

বিকল্প অনুমান, $H_A : \beta_2 \neq 0$ অর্থাৎ নির্ভরণ মডেলের ঢাল শূন্য নয়।

প্রদত্ত তথ্যানুসারে সারণী নিম্নরূপ-

আয় (X)	ভোগ (Y)	XY	X ²	Y ²
8	7	56	64	49
10	7	70	100	49
12	9	108	144	81
14	10	140	196	100
16	11	176	256	121
18	12	216	324	144
20	12	240	400	144
22	14	308	484	196
24	16	384	576	256
26	15	390	676	225
$\Sigma X = 170$	$\Sigma Y = 113$	$\Sigma XY = 2088$	$\Sigma X^2 = 3220$	$\Sigma Y^2 = 1365$

নিম্নের ধাপসমূহের সাহায্যে অনুমান পরীক্ষা করা হলোঃ-

ধাপ - ১ : মোট ভেদ নির্ণয় :

$$\begin{aligned}
 SST &= \Sigma Y^2 - \frac{(\Sigma Y)^2}{n} \\
 &= 1365 - \frac{(113)^2}{10} \\
 &= 1365 - 1276.9 \\
 &= 88.1
 \end{aligned}$$

এমবিএ প্রোগ্রাম

ধাপ - ২ : ব্যাখ্যায়িত ভেদ নির্ণয় :

$$SSR = \hat{\beta}_2^2 \left\{ \sum X^2 - \frac{(\sum X)^2}{n} \right\}$$

এখন নির্ভরণ সমীকরণের ঢাল,

$$\begin{aligned} \text{এখ } \hat{\beta}_2 &= \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}} \\ &= \frac{2088 - \frac{170 \times 113}{10}}{3220 - \frac{(170)^2}{10}} \\ &= \frac{2088 - 1921}{3220 - 2890} \\ &= \frac{167}{330} \\ &= 0.51 \end{aligned}$$

$$\begin{aligned} \therefore SSR &= (0.51)^2 \left\{ 3220 - \frac{(170)^2}{10} \right\} \\ &= 0.26 (3220 - 2890) \\ &= 0.26 (330) \\ &= 85.8 \end{aligned}$$

ধাপ - ৩ : অব্যাক্ষায়িত ভেদ নির্ণয় :

$$\begin{aligned} SSE &= SST - SSR \\ &= 88.1 - 85.8 \\ &= 2.3 \end{aligned}$$

ধাপ - ৪ : ভেদাংক বিশ্লেষণ (ANOVA) সারণী:

ভেদের উৎস	বর্গ সমষ্টি (Sum Square) SS	স্বাধীনতার মাত্রা (v)	গড় বর্গ সমষ্টি (Mean Square) MS	F
নির্ভরণ (R)	SSR = 85.8	$v_1 = k - 1$ $= 2 - 1$ $= 1$	$MSR = \frac{SSR}{v_1}$ $= 85.8$	$F = \frac{MSR}{MSE}$ $= \frac{85.8}{0.29}$ $= 295.86$
ত্রুটি (E)	SSE = 2.3	$v_2 = n - k$ $= 10 - 2$ $= 8$	$MSE = \frac{SSE}{v_2}$ $= 0.29$	
মোট (T)	SST = 88.1	$n - 1 = 10 - 1$ $= 9$		

∴ F এর নির্ণীত মান = 295.86

ধাপ-৫ : সিদ্ধান্ত-

দেওয়া আছে - গুরুত্বের স্তর, $\alpha = 5\% = 0.05$

এখানে স্বাধীনতার মাত্রা $v_1 = 1$ এবং $v_2 = 8$

∴ F এর তাত্ত্বিক মান = 5.32

যেহেতু F এর নির্ণীত মান (295.85) উহার তাত্ত্বিক মান (5.32) অপেক্ষা বড়। অতএব নাস্তি অনুমান (H_0) বর্জনীয়। অর্থাৎ নির্ভরণ মডেলের ঢাল শূন্য নয়।



সারসংক্ষেপ

নির্ভরণ মডেলের ঢালের পরীক্ষা ভেদাংক বিশ্লেষণের সাহায্যে করা যায়।



উদ্দেশ্য

এ পাঠ শেষে আপনি-

- বহুধা নির্ভরণ মডেল বর্ণনা করতে পারবেন।
- বহুধা নির্ভরণ মডেল সম্পর্কিত সমস্যার সমাধান করতে পারবেন।

ভেদাংক বিশ্লেষণ : বহুধা নির্ভরণ মডেল

Analysis of Variance : Multiple Regression Model

যে পরিসংখ্যানিক পদ্ধতিতে দুই বা ততোধিক স্বাধীন চলকের সাথে অধীন চলকের গাণিতিক সম্পর্ক সমীকরণের মাধ্যমে প্রকাশ করা হয় এবং স্বাধীন চলকগুলোর প্রদত্ত মানের জন্য অধীন চলকের গড় মান জানা যায় তাকে বহুধা নির্ভরণ মডেল বলে।

ধরি X_1 ও X_2 দুটি স্বাধীন চলক এবং Y অধীন চলক।

∴ বহুধা নির্ভরণ মডেল-

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i$$

∴ বহুধা নির্ভরণ সমীকরণ-

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2$$

নির্ভরণ মডেলের ঢালগুলোর মান শূন্য কিনা ভেদাংক বিশ্লেষণের সাহায্যে পরীক্ষা করা যায়। নিম্নে ব্যাখ্যা করা হলো :

ধাপ - ১ : মোট ভেদ, $SST = \sum y^2$ এখানে $y = Y - \bar{Y}$

ধাপ - ২ : ব্যাখ্যায়িত ভেদ, $SSR = \hat{\beta}_1 \sum yx_1 + \hat{\beta}_2 \sum yx_2$

$$\text{এখানে } \hat{\beta}_1 = \frac{\sum yx_1 + \sum x_2^2 - \sum x_1 x_2 + \sum yx_2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} \quad x_1 = x_1 - \bar{x}_1$$

$$\text{এবং } \hat{\beta}_2 = \frac{\sum yx_2 + \sum x_1^2 - \sum x_1 x_2 + \sum yx_1}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} \quad x_2 = x_2 - \bar{x}_2$$

ধাপ - ৩ : অব্যাক্ষায়িত ভেদ, $SSE = SST - SSR$

$$\therefore \sum \hat{u}^2 = \sum y^2 - \hat{\beta}_1 \sum yx_1 - \hat{\beta}_2 \sum yx_2$$

ধাপ- ৪ : ভেদাংক বিশ্লেষণ সারণী-

ভেদের উৎস	বর্গ সমষ্টি (Sum Square) SS	স্বাধীনতার মাত্রা (v)	গড় বর্গ সমষ্টি (Mean Square) MS	F
নির্ভরণ (R)	$SSR = \hat{\beta}_1 \sum yx_1 + \hat{\beta}_2 \sum yx_2$	$v_1 = k - 1$	$MSR = \frac{SSR}{v_1}$	$F = \frac{MSR}{MSE}$
ত্রুটি (E)	$SSE = \sum \hat{u}^2$	$v_2 = n - k$	$MSE = \frac{SSE}{v_2}$	
মোট (T)	$SST = \sum y^2$	$n - 1$		

এখানে, k = পরিমাপকৃত পরামিত্রের সংখ্যা

n = তথ্যের সংখ্যা

ধাপ - ৫ : সিদ্ধান্ত :

যদি F এর নির্ণীত মান উহার তাত্ত্বিক মান অপেক্ষা বড় হয় তাহলে নাস্তি অনুমান বর্জনীয়। অন্যথায় নাস্তি অনুমান গ্রহণীয় হবে।

উদাহরণ : একটি বহুধা নির্ভরণ মডেলের তথ্যসমূহ নিম্নরূপ-

$$\Sigma yx_1 = 775.22, \Sigma y x_1^2 = 6086.10, \Sigma y x_2^2 = 192.20, \Sigma x_1x_2 = -1019.62, \Sigma yx_2 = -125.61, \Sigma y^2 = 104.75, n = 10$$

নির্ভরণ মডেলের ঢালগুলোর মান শূন্য কিনা ৫% গুরুত্বের স্তরে পরীক্ষা করণ।

সমাধান: দেওয়া আছে-

$$\Sigma yx_1 = 775.22, \Sigma y x_1^2 = 6086.10, \Sigma y x_2^2 = 192.20, \Sigma x_1x_2 = -1019.62, \Sigma yx_2 = -125.61, \Sigma y^2 = 104.75, n = 10$$

প্রদত্ত তথ্যানুসারে নির্ভরণ মডেল হবে-

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + U$$

নাস্তি অনুমান, $H_0 : \beta_1 = \beta_2 = 0$ অর্থাৎ নির্ভরণ মডেলের ঢালগুলোর মান শূন্য।

বিকল্প অনুমান, $H_A : \beta_1 \neq \beta_2 \neq 0$ অর্থাৎ নির্ভরণ মডেলের ঢালগুলোর মান শূন্য নয়।

নিম্নের ধাপসমূহের সাহায্যে অনুমান পরীক্ষা করা হলো-

ধাপ - ১ : মোট ভেদ

$$\begin{aligned} SST &= \Sigma y^2 \\ &= 104.75 \end{aligned}$$

ধাপ - ২ : ব্যাখ্যায়িত ভেদ

$$\begin{aligned} SSR &= \hat{\beta}_1 \Sigma yx_1 + \hat{\beta}_2 \Sigma yx_2 \\ \text{এখন } \hat{\beta}_1 &= \frac{\Sigma yx_1 \Sigma x_2^2 - \Sigma x_1x_2 \Sigma yx_2}{\Sigma x_1^2 \Sigma x_2^2 - (\Sigma x_1x_2)^2} \\ &= \frac{(775.22)(192.20) - (-1019.62)(-125.61)}{(6086.10)(192.20) - (-1019.62)^2} \\ &= \frac{148997.28 - 128074.47}{1169748.42 - 1039624.94} \\ &= \frac{20922.81}{130123.48} \\ &= 0.1608 \end{aligned}$$

$$\begin{aligned} \text{এবং } \hat{\beta}_2 &= \frac{\Sigma yx_2 \Sigma x_1^2 - \Sigma x_1x_2 \Sigma yx_1}{\Sigma x_1^2 \Sigma x_2^2 - (\Sigma x_1x_2)^2} \\ &= \frac{(-125.61)(6086.10) - (-1019.62)(775.22)}{(6086.10)(192.20) - (-1019.62)^2} \\ &= \frac{-764475.02 + 790429.82}{1169748.42 - 1039624.94} \\ &= \frac{25954.80}{130123.48} \end{aligned}$$

এমবিএ প্রোগ্রাম

$$\begin{aligned} &= 0.1995 \\ \text{SSR} &= (0.1608)(775.22) + (0.1995)(-125.61) \\ &= 124.66 - 25.06 \\ &= 99.60 \end{aligned}$$

ধাপ - ৩ : অব্যাক্ষয়িত ভেদ

$$\begin{aligned} \text{SSR} &= \text{SST} - \text{SSR} \\ &= 104.75 - 99.66 \\ &= 5.15 \end{aligned}$$

ধাপ - ৪ : ভেদাংক বিশ্লেষণ সারণী-

ভেদের উৎস	বর্গ সমষ্টি (SS)	স্বাধীনতার মাত্রা (v)	গড় বর্গ সমষ্টি (MS)	$F = \frac{MSR}{MSE}$
নির্ভরণ (R)	SSR = 99.60	$v_1 = k - 1$ $= 3 - 1$ $= 2$	$MSR = \frac{SSR}{v_1}$ $= \frac{99.60}{2}$ $= 49.8$	$F = \frac{49.8}{0.74}$ $= 67.30$
ত্রুটি (E)	SSE = 5.15	$v_2 = n - k$ $= 10 - 3$ $= 7$	$MSE = \frac{SSE}{v_2}$ $= \frac{5.15}{7}$ $= 0.74$	
মোট (T)	SST = 104.75	$n - 1 = 10 - 1$ $= 9$		

∴ F এর নির্ণীত মান = 67.30

ধাপ-৫ : সিদ্ধান্ত-

দেয়া আছে - গুরুত্বের স্তর, $\alpha = 5\% = 0.05$

এখানে স্বাধীনতার মাত্রাদ্বয় : $v_1 = 2$ এবং $v_2 = 7$

∴ F এর তাত্ত্বিক মান = 4.74

যেহেতু F এর নির্ণীত মান (67.30) উহার তাত্ত্বিক মান (4.74) অপেক্ষা বড়। অতএব নাস্তি অনুমান H_0 বর্জনীয়। অর্থাৎ নির্ভরণ মডেলের ঢালগুলোর মান শূন্য নয়।



সারসংক্ষেপ

যে পরিসংখ্যানিক পদ্ধতিতে দুই বা ততোধিক স্বাধীন চলকের সাথে অধীন চলকের গাণিতিক সম্পর্ক সমীকরণের মাধ্যমে প্রকাশ করা হয় এবং স্বাধীন চলকগুলোর প্রদত্ত মানের জন্য অধীন চলকের গড় মান জানা যায় তাকে বহুধা নির্ভরণ মডেল বলে।



রচনামূলক প্রশ্ন

- ১। ভেদাংক বিশ্লেষণ বলতে কী বুঝেন ?
- ২। ভেদাংক বিশ্লেষণ অনুমিতিসমূহ লিখুন।
- ৩। ভেদাংক বিশ্লেষণের ব্যবহার লিখুন।
- ৪। একমুখী শ্রেণিবিভাগ হতে ভেদাংক বিশ্লেষণ সারণি তৈরি করার উপায় বর্ণনা করুন।
- ৫। দ্বি-মুখী শ্রেণিবিভাগের ক্ষেত্রে ভেদাংক বিশ্লেষণ সারণি কিভাবে তৈরি করা যায়?
- ৬। বহু নির্ভরণ মডেল বলতে কি বুঝা? এর ক্ষেত্রে ভেদাংক বিশ্লেষণ সারণি তৈরি করুন।