When the experimental material is not homogeneous the LSD is better than RBD. In RBD local control is used according to one way grouping. i.e according to blocks but in LSD local control is used according to two way grouping i.e. rows and columns. Hence it is used when two sources of errors are to be controlled simultaneously. In this design number of treatments are equal to the number of replication and the treatment occurs once and only once in each row and column. In this design, Latin alphabet are used to denote the treatments, and shape is square due to equal number of treatments and replication so called Latin Square design. It is based upon the all principles of design namely replication, randomization and local control.

Let us consider m treatments with m replication each so that there are $N = m^2$ experiments unit.

Let us divide the experimental material into m^2 experimental units arranged in square so that each row as well as column contains m units. In this design none of treatments are replicated along row wise or column wise. In this case we study the variation between treatments, the variation between rows and variations between columns. It has only m^2 experimental treatments. Hence it is the case of incomplete three ways ANOVA. for complete three way ANOVA, we need m^3 experimental unit.

Let us consider t = 3(A, B, C,) then 3x3 LSD is as shown below.

Example Layout (3×3 Latin Square):

	Column 1	Column 2	Column 3
Row 1	A	B	C
Row 2	B	C	A
Row 3	C	A	B

Critical Values of Student's t-distribution

d.f.	Level of significance for one-tailed test					
	0.10	0.05	0.025	0.01	0.005	0.0005
	Level of significance for two-tailed test					
	0.20	0.10	0.05	0.02	0.01	0.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.599
3	1.638	2.353	3.182	4.541	5.841	12.924
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.869
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.408
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221

Where A, B, C represent different treatments.

3. Why Number of Treatments Should Not Be Less Than 3

1.Statistical Power: With fewer than 3 treatments, it's difficult to establish meaningful patterns or draw valid conclusions.2.**Variability Control**: Latin Squares are designed to control two sources of variability (rows and columns). With only 2 treatments, the design reduces to a simple block design, losing its advantage.3.**Balance**: The design requires each treatment to appear once in each row and column. With 2 treatments, the design becomes trivial and less effective in controlling variability.

Application Example:Scenario: Testing the effect of 3 different fertilizers (A, B, C) on crop yield, while controlling for soil type (rows) and irrigation method (columns).**Latin SquareLayout:**

Soil Type \ Irrigation	Drip	Sprinkler	Flood
Sandy	A	B	C
Loamy	B	C	A
Clay	C	A	B

Critical Values of T in the Wilcoxon Matched-Pairs Signed-Ranks Test.

n	Level of significance for one-tailed test			
	0.05	.025	.01	.005
	Level of significance for two-tailed test			
	0.10	.05	.02	.01
5	1	–	–	–
6	2	1	–	–
7	4	2	0	–
8	6	4	2	0
9	8	6	3	2
10	11	8	5	3
11	14	11	7	6
12	17	14	10	7
13	21	17	13	10
14	26	21	16	13
15	30	25	20	16

5. Advantages of Latin Square Design1. Controls two sources of nuisance variability2.Requires fewer experimental runs3.Provides balanced and orthogonal comparisons

6. Limitations1.Not suitable for interactions between factors2.Requires equal number of rows, columns, and treatments3.Becomes impractical for large numbers of treatments

Critical Values of T in the Wilcoxon Matched-Pairs Signed-Ranks Test.

n	Level of significance for one-tailed test			
	0.05	.025	.01	.005
	Level of significance for two-tailed test			
	0.10	.05	.02	.01
5	1	–	–	–
6	2	1	–	–
7	4	2	0	–
8	6	4	2	0
9	8	6	3	2
10	11	8	5	3
11	14	11	7	6
12	17	14	10	7
13	21	17	13	10
14	26	21	16	13
15	30	25	20	16

Conclusion: The Latin Square Design is a powerful tool when you need to control two sources of variation simultaneously, but it requires at least 3 treatments to be effective. This design is particularly useful in agricultural, industrial, and clinical experiments where multiple sources of variability need to be accounted for.

Type I Error (False Positive):

Type I error occurs when a null hypothesis that is actually true is incorrectly rejected. In other words, it's a false positive result.

This error is often denoted as α (alpha), which represents the level of significance, and it indicates the probability of making a Type I error.

Type I errors are considered more serious in situations where making an incorrect rejection of the null hypothesis has significant consequences or costs. For example, in medical testing, a Type I error might mean diagnosing a healthy person as having a disease.**Type II Error (False Negative)**:

Type II error occurs when a null hypothesis that is actually false is incorrectly not rejected. It's a false negative result.

This error is often denoted as β (beta) and is used to represent the probability of making a Type II error.

Type II errors can be problematic when failing to reject the null hypothesis means that a real effect or relationship goes undetected. For example, in medical testing, a Type II error might mean failing to diagnose a person who actually has a disease.

Critical Values of Student's t-distribution

d.f.	Level of significance for one-tailed test					
	0.10	0.05	0.025	0.01	0.005	0.0005
	Level of significance for two-tailed test					
	0.20	0.10	0.05	0.02	0.01	0.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.599
3	1.638	2.353	3.182	4.541	5.841	12.924
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.869
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.408
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221

P-tests	NP-tests
The tests whose models specify certain conditions about parameters of the population from which the sample has been drawn.	The tests whose models don't specify any conditions about parameters of population from which sample has been drawn.
Mostly applied only to the data measured in ratio and interval scale.	Mostly only applied to data measured in nominal and ordinal scale.
Most powerful.	Weaker.
Designed to test the statistical hypothesis of one or more parameters of population.	Designed to test statistical hypothesis which don't involve any parameters of population.

Critical Values of T in the Wilcoxon Matched-Pairs Signed-Ranks Test.

n	Level of significance for one-tailed test			
	0.05	.025	.01	.005
	Level of significance for two-tailed test			
	0.10	.05	.02	.01
5	1	–	–	–
6	2	1	–	–
7	4	2	0	–
8	6	4	2	0
9	8	6	3	2
10	11	8	5	3
11	14	11	7	6
12	17	14	10	7
13	21	17	13	10
14	26	21	16	13
15	30	25	20	16

300 Statistics - II

Critical Values of T in the Wilcoxon Matched-Pairs Signed-Ranks Test.

n	Level of significance for one-tailed test			
	0.05	.025	.01	.005
	Level of significance for two-tailed test			
	0.10	.05	.02	.01
5	1	–	–	–
6	2	1	–	–
7	4	2	0	–
8	6	4	2	0
9	8	6	3	2
10	11	8	5	3
11	14	11	7	6
12	17	14	10	7
13	21	17	13	10
14	26	21	16	13
15	30	25	20	16

Stochastic process.

A stochastic process, also known as a random process, is a mathematical concept used in probability theory and statistics to model and analyze the evolution of a system or a phenomenon over time. It's essentially a collection of random variables, each of which represents a state or an outcome at a specific point in time. Here are some key points about stochastic processes:

Random Variables Over Time: A stochastic process involves random variables that evolve with time. Each random variable represents the state of the system or process at a particular time point.

Sequential Nature: Stochastic processes are inherently sequential in nature. You can think of them as a sequence or a collection of random events or observations that unfold over time.

Critical values for total number of runs 'r' at α = 0.05 for two tailed test.

The smaller critical value for a left-hand critical region, the larger for a right-hand critical region. For a one tailed test $\alpha = 0.025$ and use only-one of the critical values of r.

	The larger of n_1 and n_2															
	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2								2 6	2 6	2 6	2 6	2 6	2 6	2 6	2 6	2 6
3		2 8	2 8	2 8	2 8	2 8	2 8	2 8	2 8	2 8	3 8	3 8	3 8	3 8	3 8	3 8
4	2 9	2 9	2 10	3 10	3 10	3 10	3 10	3 10	3 10	3 10	3 10	4 10	4 10	4 10	4 10	4 10
5	2 10	3 10	3 11	3 11	3 12	3 12	4 12	4 12	4 12	4 12	4 12	4 12	4 12	5 12	5 12	5 12
6		3 11	3 12	3 12	4 13	4 13	4 13	4 13	5 14	5 14	5 14	5 14	5 14	5 14	6 14	6 14
7			3 13	4 13	4 14	5 14	5 14	5 14	5 15	5 15	6 15	6 16	6 16	6 16	6 16	6 16
8				4 14	5 14	5 15	5 15	6 16	6 16	6 16	6 16	6 17	7 17	7 17	7 17	7 17
9					5 15	5 16	6 16	6 16	6 17	7 17	7 18	7 18	7 18	7 18	7 18	7 18
10						6 16	6 17	7 17	7 18	7 18	7 18	8 19	8 19	8 19	8 20	9 20
11							7 17	7 18	7 19	8 19	8 19	8 20	9 20	9 20	9 21	9 21
12								7 19	8 19	8 20	8 20	9 21	9 21	9 21	10 22	10 22

Classification: Stochastic processes can be classified into different types based on various characteristics. Common classifications include discrete-time vs. continuous-time processes, stationary vs. non-stationary processes, and Markov vs. non-Markov processes, among others.

Critical values for total number of runs 'r' at α = 0.05 for two tailed test.

The smaller critical value for a left-hand critical region, the larger for a right-hand critical region. For a one tailed test $\alpha = 0.025$ and use only-one of the critical values of r.

	The larger of n_1 and n_2															
	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2								2 6	2 6	2 6	2 6	2 6	2 6	2 6	2 6	2 6
3		2 8	2 8	2 8	2 8	2 8	2 8	2 8	2 8	2 8	3 8	3 8	3 8	3 8	3 8	3 8
4	2 9	2 9	2 10	3 10	3 10	3 10	3 10	3 10	3 10	3 10	3 10	4 10	4 10	4 10	4 10	4 10
5	2 10	3 10	3 11	3 11	3 12	3 12	4 12	4 12	4 12	4 12	4 12	4 12	4 12	5 12	5 12	5 12
6		3 11	3 12	3 12	4 13	4 13	4 13	4 13	5 14	5 14	5 14	5 14	5 14	5 14	6 14	6 14
7			3 13	4 13	4 14	5 14	5 14	5 14	5 15	5 15	6 15	6 16	6 16	6 16	6 16	6 16
8				4 14	5 14	5 15	5 15	6 16	6 16	6 16	6 16	6 17	7 17	7 17	7 17	7 17
9					5 15	5 16	6 16	6 16	6 17	7 17	7 18	7 18	7 18	7 18	7 18	7 18
10						6 16	6 17	7 17	7 18	7 18	7 18	8 19	8 19	8 19	8 20	9 20
11							7 17	7 18	7 19	8 19	8 19	8 20	9 20	9 20	9 21	9 21
12								7 19	8 19	8 20	8 20	9 21	9 21	9 21	10 22	10 22

Applications: Stochastic processes find applications in various fields, such as finance, physics, engineering, biology, and economics. For example, in finance, they are used to model stock prices, in physics to simulate random motion, and in biology to study population dynamic.

Critical values for total number of runs 'r' at α = 0.05 for two tailed test.

The smaller critical value for a left-hand critical region, the larger for a right-hand critical region. For a one tailed test $\alpha = 0.025$ and use only-one of the critical values of r.

	The larger of n_1 and n_2															
	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2								2 6	2 6	2 6	2 6	2 6	2 6	2 6	2 6	2 6
3		2 8	2 8	2 8	2 8	2 8	2 8	2 8	2 8	2 8	3 8	3 8	3 8	3 8	3 8	3 8
4	2 9	2 9	2 10	3 10	3 10	3 10	3 10	3 10	3 10	3 10	3 10	4 10	4 10	4 10	4 10	4 10
5	2 10	3 10	3 11	3 11	3 12	3 12	4 12	4 12	4 12	4 12	4 12	4 12	4 12	5 12	5 12	5 12
6		3 11	3 12	3 12	4 13	4 13	4 13	4 13	5 14	5 14	5 14	5 14	5 14	5 14	6 14	6 14
7			3 13	4 13	4 14	5 14	5 14	5 14	5 15	5 15	6 15	6 16	6 16	6 16	6 16	6 16
8				4 14	5 14	5 15	5 15	6 16	6 16	6 16	6 16	6 17	7 17	7 17	7 17	7 17
9					5 15	5 16	6 16	6 16	6 17	7 17	7 18	7 18	7 18	7 18	7 18	7 18
10						6 16	6 17	7 17	7 18	7 18	7 18	8 19	8 19	8 19	8 20	9 20
11							7 17	7 18	7 19	8 19	8 19	8 20	9 20	9 20	9 21	9 21
12								7 19	8 19	8 20	8 20	9 21	9 21	9 21	10 22	10 22

Markov Property: In a Markov process, the future behavior of the process depends only on its present state and not on its past states. This property is often used to simplify the analysis of stochastic processes.

Examples: Common examples of stochastic processes include the Poisson process (used to model the arrival of events over time), Brownian motion (used to describe random motion), and autoregressive integrated moving average (ARIMA) models (used for time series analysis).

Modeling and Analysis: Stochastic processes are used to model uncertainty and randomness in real-world phenomena. They allow researchers and analysts to make probabilistic predictions and assess the likelihood of various outcomes.

Simulation: Stochastic processes are often used in Monte Carlo simulations, a technique for estimating the behavior of complex systems by repeatedly sampling random variables to understand the range of possible outcomes.

of the Mann-Whitney test.

$n_2 = 3$

	n_1		
U_0	1	2	3
0	.25	.10	.05
1	.50	.20	.10
2		.40	.20
3		.60	.35
4			.50

$n_2 = 4$

	n_1			
U_0	1	2	3	4
0	.2000	.0667	.0286	.0143
1	.4000	.1333	.0571	.0286
2	.6000	.2667	.1143	.0571
3		.4000	.2000	.1000
4		.6000	.3143	.1714
5			.4286	.2429
6			.5714	.3429
7				.4429
8				.5571

Let $\{X_n\}$ be a sequence of values describing a mutually exclusive and exhaustive system of events. Let X_n values take only discrete the union I of all possible values of X_n is then a countable set called the state space of the process. Each element $i \in I$ is called a state. The index n is of time. The number of events may be finite or infinite the values of $\{X_n\}$ is said to be a Markov chain or Markov dependent if for all $i_0, i_1, i_2, \ldots i_{n-1}, i_n \in I$ and for all n.

$P(X_n = i_n (X_0 = i_0, X_1 = i_1 \ldots\ldots X_{(n-1)} = i_{(n-1)}) = P(X_n = i_n X_{n-1} = i_{n-1})$

The conditional distribution of X_n given the values $X_0, X_1, \ldots\ldots, X_{n-1}$ depends only on X_{n-1} not on the preceding values. If the state space is finite then we have finite Markov Chain.

i. It is a mathematical system experiences transitions from one state to another according to certain probabilistic rules.

ii. No matter how the process arrived at its present state, the possible future states are fixed i.e. the probability of transitioning to any particular state is dependent solely on the current state and time elapsed.

iii. The state space, or set of all possible states, can be anything: letters, numbers, weather conditions, baseball scores, or stock performances.

	n_1								
U_0	1	2	3	4	5	6	7	8	9
0	.1000	.0182	.0045	.0014	.0005	.0002	.0001	.0000	.0000
1.	.2000	.0364	.0091	.0028	.0010	.0004	.0002	.0001	.0000
2.	.3000	.0727	.0182	.0056	.0020	.0008	.0003	.0002	.0001
3.	.4000	.1091	.0318	.0098	.0035	.0014	.0006	.0003	.0001
4.	.5000	.1636	.0500	.0168	.0060	.0024	.0010	.0005	.0002
5.		.2182	.0727	.0252	.0095	.0038	.0017	.0008	.0004
6.		.2909	.1045	.0378	.0145	.0060	.0026	.0012	.0006
7.		.3636	.1409	.0531	.0210	.0088	.0039	.0019	.0009
8.		.4545	.1864	.0741	.0300	.0128	.0058	.0028	.0014
9.		.5455	.2409	.0993	.0415	.0180	.0082	.0039	.0020

U_0	n_1								
	1	2	3	4	5	6	7	8	9
0	.1000	.0182	.0045	.0014	.0005	.0002	.0001	.0000	.0000
1.	.2000	.0364	.0091	.0028	.0010	.0004	.0002	.0001	.0000
2.	.3000	.0727	.0182	.0056	.0020	.0008	.0003	.0002	.0001
3.	.4000	.1091	.0318	.0098	.0035	.0014	.0006	.0003	.0001
4.	.5000	.1636	.0500	.0168	.0060	.0024	.0010	.0005	.0002
5.		.2182	.0727	.0252	.0095	.0038	.0017	.0008	.0004
6.		.2909	.1045	.0378	.0145	.0060	.0026	.0012	.0006
7.		.3636	.1409	.0531	.0210	.0088	.0039	.0019	.0009
8.		.4545	.1864	.0741	.0300	.0128	.0058	.0028	.0014
9.		.5455	.2409	.0993	.0415	.0180	.0082	.0039	.0020

iv. Markov chains may be modeled by finite state machines, and random walks provide a prolific example of their usefulness in mathematics.

v. They arise broadly in statistical and information-theoretical contexts and are widely employed in economics, game theory, queueing (communication) theory, genetics, and finance.

vi. While it is possible to discuss Markov chains with any size of state space, the initial theory and most applications are focused on cases with a finite (or countably infinite) number of states.

vii. Many uses of Markov chains require proficiency with common matrix methods.

12.			.3462	.1518	.0646	.0280	.0125	.0058	.0028	.0014
13.			.4056	.1868	.0823	.0363	.0165	.0078	.0038	.0019
14.			.4685	.2268	.1032	.0467	.0215	.0103	.0051	.0026
15.			.5315	.2697	.1272	.0589	.0277	.0133	0.066	.0034
16.				.3177	.1548	.0736	.0351	.0171	.0086	.0045
17.				.3666	.1855	.0903	.0439	.0217	.0110	.0057
18.				.4196	.2198	.1099	.0544	.0273	.0140	.0073
19.				.4725	.2567	.1317	.0665	.0338	.0175	.0093
20.				.5275	.2970	.1566	.0806	.0416	.0217	.0116
21.					.3393	.1838	.0966	.0506	.0267	.0144
22.					.3839	.2139	.1148	.0610	.0326	.0177
23.					.4296	.2461	.1349	.0729	.0394	.0216
24.					.4765	.2811	.1574	.0864	.0474	.0262
25.					.5235	.3177	.1819	.1015	.0564	.0315
26.						.3564	.2087	.1185	.0667	.0376
27.						.3962	.2374	.1371	.0782	.0446
28.						.4374	.2681	.1577	.0912	.0526
29.						.4789	.3004	.1800	.1055	.0615
30.						.5211	.3345	.2041	.1214	.0716

12.			.3462	.1518	.0646	.0280	.0125	.0058	.0028	.0014
13.			.4056	.1868	.0823	.0363	.0165	.0078	.0038	.0019
14.			.4685	.2268	.1032	.0467	.0215	.0103	.0051	.0026
15.			.5315	.2697	.1272	.0589	.0277	.0133	0.066	.0034
16.				.3177	.1548	.0736	.0351	.0171	.0086	.0045
17.				.3666	.1855	.0903	.0439	.0217	.0110	.0057
18.				.4196	.2198	.1099	.0544	.0273	.0140	.0073
19.				.4725	.2567	.1317	.0665	.0338	.0175	.0093
20.				.5275	.2970	.1566	.0806	.0416	.0217	.0116
21.					.3393	.1838	.0966	.0506	.0267	.0144
22.					.3839	.2139	.1148	.0610	.0326	.0177
23.					.4296	.2461	.1349	.0729	.0394	.0216
24.					.4765	.2811	.1574	.0864	.0474	.0262
25.					.5235	.3177	.1819	.1015	.0564	.0315
26.						.3564	.2087	.1185	.0667	.0376
27.						.3962	.2374	.1371	.0782	.0446
28.						.4374	.2681	.1577	.0912	.0526
29.						.4789	.3004	.1800	.1055	.0615
30.						.5211	.3345	.2041	.1214	.0716

ii. Properties of a good estimator1. Unbiasedness.

An estimator is said to be unbiased if its expected value is identical with the population parameter being estimated. That is if θ is an unbiased estimate of θ, then we must have E (θ) = θ. Many estimators are "Asymptotically unbiased" in the sense that the biases reduce to practically insignificant value (Zero) when n becomes sufficiently large. The estimator S_2 is an example.

It should be noted that bias is estimation is not necessarily undesirable. It may turn out to be an asset in some situations.**Consistency.**If an estimator, say θ, approaches the parameter θ closer and closer as the sample size n increases, θ is said to be a consistent estimator of θ. Stating somewhat more rigorously, the estimator θ is said is be a consistent estimator of θ if, as n approaches infinity, the probability approaches 1 that θ will differ from the parameter θ by no more than an arbitrary constant.The sample mean is an unbiased estimator of μ no matter what form the population distribution assumes, while the sample median is an unbiased estimate of μ only if the population distribution is symmetrical. The sample mean is better than the sample median as an estimate of μ in terms of both unbiasedness and consistency.

b. Critical values of U for a one-tailed test at 0.05 or for a two-tailed test at 0.10

n_1 / n_2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																			0	0
2					0	0	0	1	1	1	1	2	2	2	3	3	3	4	4	4
3			0	0	1	2	2	3	3	4	5	5	6	7	7	8	9	9	10	11
4			0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
5		0	1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
6		0	2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
7		0	2	4	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
8		1	3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
9		1	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
10		1	4	7	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
11		1	5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
12		2	5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
13		2	6	10	15	19	24	28	33	37	42	47	51	56	61	65	70	75	80	84
14		2	7	11	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
15		3	7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
16		3	8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
17		3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
18		4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
19	0	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
20	0	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138

3. Efficiency.

The concept of efficiency refers to the sampling variability of an estimator. If two competing estimators are both unbiased, the one with the smaller variance (for a given sample size) is said to be relatively more efficient. Stated in a somewhat different language, an estimator θ is said to be more efficient than another estimator θ_2 for θ if the variance of the first is less than the variance of the second. The smaller the variance of the estimator, the more concentrated is the distribution of the estimator around the parameter being estimated and, therefore, the better this estimator is.

b. Critical values of U for a one-tailed test at 0.05 or for a two-tailed test at 0.10

n_1 / n_2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																			0	0
2					0	0	0	1	1	1	1	2	2	2	3	3	3	4	4	4
3			0	0	1	2	2	3	3	4	5	5	6	7	7	8	9	9	10	11
4			0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
5		0	1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
6		0	2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
7		0	2	4	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
8		1	3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
9		1	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
10		1	4	7	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
11		1	5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
12		2	5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
13		2	6	10	15	19	24	28	33	37	42	47	51	56	61	65	70	75	80	84
14		2	7	11	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
15		3	7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
16		3	8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
17		3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
18		4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
19	0	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
20	0	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138

4. Sufficiency.

x = 0

n \ x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	5	•														
2	250	500	•													
3	125	500	875	•												
4	063	313	688	938	•											
5	031	188	500	812	969	•										
6	016	109	344	656	891	984	•									
7	008	062	227	500	773	938	992	•								
8	004	035	145	363	637	855	965	996	•							
9	002	020	090	254	500	746	910	980	998	•						

An estimator is said to be sufficient if it conveys much information as is possible about the parameter which is contained in the sample. The significance of sufficiency lies in the fact that if a sufficient estimator exists, it is absolutely unnecessary to considered any other estimator; a sufficient estimator ensures that all information a sample a sample can furnished with respect to the estimation of a parameter is being utilized.

Many methods have been devised for estimating parameters that may provide estimators satisfying these properties. The two important methods are the least square method and the method of maximum likelihood.

x = 0

n \ x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	5	•														
2	250	500	•													
3	125	500	875	•												
4	063	313	688	938	•											
5	031	188	500	812	969	•										
6	016	109	344	656	891	984	•									
7	008	062	227	500	773	938	992	•								
8	004	035	145	363	637	855	965	996	•							
9	002	020	090	254	500	746	910	980	998	•						

1. Test of Equality of Two Variances

This test is used to determine whether two populations have equal variances. It is commonly applied before conducting a t-test, where equal variance assumption is required.

- **Hypotheses:**
  - H0H_0H0: σ12=σ22\sigma_1^2 = \sigma_2^2σ12=σ22
  - H1H_1H1: σ12≠σ22\sigma_1^2 \ne \sigma_2^2σ12=σ22
- **Test Statistic (F-test):**

F=s12s22F = \frac{s_1^2}{s_2^2}F=s22s12

Where s12s_1^2s12 and s22s_2^2s22 are the sample variances from two groups (with s12>s22s_1^2 > s_2^2s12>s22).

- The statistic follows an **F-distribution** with degrees of freedom (n1−1,n2−1)(n_1 - 1, n_2 - 1)(n1−1,n2−1).
- **Decision Rule:**
  Compare the calculated FFF with critical values from the F-distribution table.
  - If FFF lies outside the acceptance region, **reject H0H_0H0** (i.e., variances are unequal).

i.e. $p_0 \leq P(H > H^*)$ where $H^* = H_{cal.}$

Sample sizes			H	p	Sample sizes			H	p
n_1	n_2	n_3			n_1	n_2	n_3		
2	1	1	2.7000	.500	4	3	2	6.4444	.008
								6.3000	.011
2	2	2	3.6000	.200				5.4444	.046
								5.4000	.051
2	2	2	4.5714	.067				4.5111	.098
			3.7143	.200				4.4444	.102
3	1	1	3.2000	.300	4	3	3	6.7455	.010
								6.7455	.010
3	2	1	4.2857	.100				6.7091	.013
			3.8571	.100				5.7909	.046
3	2	2	5.3572	.029				4.7091	.092
			4.7143	.048				4.7000	.101
			4.5000	.067					
			4.4643	.105	4	4	1	6.6667	.010
								6.1667	.022
3	3	1	5.1429	.043				4.9667	.048
			4.5714	.100				4.8667	.054
			4.0000	.129				4.1667	.082
3	3	2	6.2500	.011				4.0667	.102
			5.3611	.032	4	4	2	7.0364	.006
			5.1389	.061					

2. Adjusted R2R^2R2 (Adjusted Coefficient of Determination)

Adjusted R2R^2R2 is a modified version of R2R^2R2 that adjusts for the number of predictors in a regression model. It is used to judge the goodness of fit while penalizing unnecessary variables.

- **Formula:**

Adjusted R2=1−((1−R2)(n−1)n−k−1)\text{Adjusted } R^2 = 1 - \left(\frac{(1 - R^2)(n - 1)}{n - k - 1} \right)Adjusted R2=1−(n−k−1(1−R2)(n−1))

Where:

-
  - R2R^2R2 = coefficient of determination
  - nnn = sample size
  - kkk = number of predictors**Key Points:**
  - Adjusted R2R^2R2 can **decrease** if a new predictor does not improve the model significantly.
  - It is a better measure than R2R^2R2 when comparing models with different numbers of variables.

i.e. $p_0 = P(H > H^*)$ where $H^* = H_{cal.}$

Sample sizes			H	p	Sample sizes			H	p
n_1	n_2	n_3			n_1	n_2	n_3		
2	1	1	2.7000	.500	4	3	2	6.4444	.008
								6.3000	.011
2	2	2	3.6000	.200				5.4444	.046
								5.4000	.051
2	2	2	4.5714	.067				4.5111	.098
			3.7143	.200				4.4444	.102
3	1	1	3.2000	.300	4	3	3	6.7455	.010
								6.7455	.010
3	2	1	4.2857	.100				6.7091	.013
			3.8571	.100				5.7909	.046
3	2	2	5.3572	.029				4.7091	.092
			4.7143	.048				4.7000	.101
			4.5000	.067					
			4.4643	.105	4	4	1	6.6667	.010
								6.1667	.022
3	3	1	5.1429	.043				4.9667	.048
			4.5714	.100				4.8667	.054
			4.0000	.129				4.1667	.082
3	3	2	6.2500	.011				4.0667	.102
			5.3611	.032	4	4	2	7.0364	.006
			5.1389	.061					