Some Inferential Techniques on Responses of Clinical Trial Data

Priyanka Chakraborty
CU Registration Number:- 041-1211-0326-19
CU Roll Number:- 193041-11-0080
B.Sc Final Semester (6th Semester)
Department of Statistics
Basanti Devi College

Disha Dey
CU Registration Number:- 041-1211-0332-19
CU Roll Number:- 193041-11-0084
B.Sc Final Semester (6th Semester)
Department of Statistics
Basanti Devi College

14th July, 2022

Abstract

In this dissertation, we perform various inferential procedures (parametric and non-parametric) and compare them on a three-stage repeated measurement design (Matthews, 1987).

Keywords : Parallel design, Crossover design, Hypotheses testing, Interval estimation

ACKNOWLEDGEMENT

We have done this piece of project work entitled as
"Some Inferential Techniques on Responses of Clinical Trial Data" as DSE-B2 paper of Semester VI examination. This wonderful opportunity helps us to do a lot of research, sharpening our statistical and analytical skills to become an expert in our domain of study. We would like to express our sincere thanks and gratitude to our supervisor Dr. Joydeep Basu for encouraging us to explore different topics for project work and completion of the project. We would like to thank him for his available suggestions and reference books which helped us a lot in finalizing this project within the given time frame. Besides our advisor, we would also like to thank other teachers in our department. We also wish to express our sincere thanks to the University of Calcutta for giving us the chance for doing this project. We highly benefitted from their constant advice, sharing their extensive knowledge of Statistics. Last but not the least, we are grateful to our parents and friends for their immense support throughout this project.

____

Priyanka Chakraborty
Disha Dey

CONTENTS

I. Introduction

Clinical trials are research studies performed on human volunteer which aims to evaluate medical, surgical, or behavioral intervention. Usually, two treatment combinations (procedures) are compared to find the efficacy of one over the other (Pocock, 1983).

The way the groups get compared varies, depending on the study design.
The most common design is called a parallel study, where participants are randomly assigned to treatment arms. Each treatment arm could include a particular dose of the study drug, a placebo, or a standard of care treatment. Patients then remain in that same treatment arm throughout the course of the study.

A crossover design is a repeated measurements design such that each experimental unit (patient) receives different treatments during different time periods, i.e., the patients cross over from one treatment to another during the course of the trial. This is in contrast to a parallel design in which patients are randomized to treatment and remain on that treatment throughout the duration of the trial.

The reason to consider a crossover design when planning a clinical trial is that it could yield a more efficient comparison of treatments than a parallel design, i.e., fewer patients might be required in the crossover design in order to attain the same level of statistical power or precision as a parallel design. Intuitively, this seems reasonable because each patient serves as his/her own matched control. Every patient receives both treatments A and B.

Response Variable is the result of the experiment where the explanatory variable is manipulated. It is a factor whose variation is explained by the other factors. The response Variable is often referred to as the Dependent Variable or the Outcome Variable.

In section 2, we describe the inferential procedures we adopted in our study. Section 3 has the description of data. In section 4, various parametric & non-parametric methods of hypothesis testing and interval estimation have been furnished. Finally, we have drawn the conclusion in section 5.

II. Inferential Procedures

Since a sample is only part of a population, the features of the former will generally differ from those of the latter. In case of random sampling, we can know about the properties of the population from a knowledge of the properties of the sample with the help of probability theory. In sampling theory, it is a bit complicated to do the same. The process of going from the known sample to the unknown population is called Statistical Inference.

Statistical Inference can be divided into two parts:

i) Estimation
ii) Testing of Hypotheses

When some information of a tentative nature regarding the feature of the population may be available to the enquirer, and he may wish to check whether the information is tenable in the light of the random sample taken from the population, that is when he takes the help of the *methods of testing of hypotheses*.
Now the methods of testing of hypotheses can be of two types:

a) Parametric Methods: Parametric tests are those tests for which we have prior knowledge of the population distribution (i.e., normal), or if not then we can easily approximate it to a normal distribution which is possible with the help of the Central Limit Theorem (CLT). These tests include:
   i) p-value test
   ii) Z-test
   iii) Fisher's t-test
   iv) Testing using Variance Stabilizing Transformation
   v) Paired t test

b) Non-parametric Methods: The important place ascribed to the normal distribution in statistical theory is well justified on the basis of the CLT. However, often it is not known whether the basic distribution is such that the central limit theorem applies or whether the approximation to the normal distribution is good enough that the resulting confidence intervals and tests of hypotheses based on normal theory are as accurate as desired. In cases where it is known that the conventional methods based on the assumption of a normal distribution are not applicable, an alternative method is desired. If the basic distribution is known, one may be able to derive exact tests of hypotheses and confidence intervals based on that distribution. In many cases an experimenter does not know the form of the basic distribution and needs statistical techniques which are applicable regardless the form of the density. These techniques are called nonparametric methods.
   i) Median test
   ii) Mann-Whitney U test
   iii) Wilcoxon Signed-Rank Test
   iv) Paired Sign Test

A. Parametric Methods

➢ Test for two sample binomial proportion by p-value test:

Assumptions: Let us consider two independent population in which p_1 and p_2 be two population proportion of a certain population characteristics with sample size n_1 and n_2 drawn from first and second population respectively.
Hypotheses: Here we want to test, $H_0 : p_1 = p_2$ vs. $H_1 : p_1 \neq p_2$.
Test Statistic: Under H_0, $p_1 = p_2 = p$.
X_i = No. of person possessing the characteristic in the i^{th} population, i = 1, 2.
$X_1 \sim Bin(n_1, p_1)$
$X_2 \sim Bin(n_2, p_2)$.
X_1 and X_2 are independent. Under H_0, $X_1 \sim Bin(n_1, p)$ and $X_2 \sim Bin(n_2, p)$ independently distributed, i.e., $X_1 + X_2 \sim Bin(n_1 + n_2, p)$.
We then calculate the conditional probability of $X_1 | X_1 + X_2$ (refer to Appendix 1)
Rejection Criteria: Either a large value or a small value of $X_1|X_1 + X_2$ indicates the rejection of H_0.
$\therefore$ p-value $= 2 \times \min\{P_{H_0}[X \leq x_{10}|X_1 + X_2 = t], P_{H_0}[X \geq x_{10}|X_1 + X_2 = t]\}$; x_{10} = observed valued of X_1

➢ Test for two independent Binomial Proportions by Z test:

Assumptions: Let, p_1 and p_2 denote respectively the proportion of individual possessing a certain attribute A in two independent populations (p_1, p_2: unknown). n_1 and n_2 no. of samples are drawn from first and second population respectively.
Hypotheses: To test, H_0: $p_1 = p_2 = p$ vs. H_1: $p_1 \neq p_2$.
Test Statistic: Let,

$$X_i = \begin{cases} 1 & \text{if the } i^{th} \text{ member of the 1st sequence possesses A} \\ 0 & \text{otherwise} \end{cases}$$

$$Y_i = \begin{cases} 1 & \text{if the } j^{th} \text{ member of the 2nd sequence possesses B} \\ 0 & \text{otherwise} \end{cases}$$

We define,

$$T_1 = \sum_{i=1}^{n_1} Xi \text{ , } T_2 = \sum_{i=1}^{n_2} Yj$$

$\therefore \frac{T_1}{n_1} \rightarrow$ Sample proportion of individuals possessing A in the 1^{st} sample
$\therefore \frac{T_2}{n_2} \rightarrow$ Sample proportion of individuals possessing A in the 2^{nd} sample
Since n_1, n_2 are large

$$\frac{T_1}{n_1} \overset{A}{\sim} N\left(p_1, \frac{p_1(1-p_1)}{n_1}\right) \text{ independently of}$$
$$\frac{T_2}{n_2} \overset{A}{\sim} N\left(p_2, \frac{p_2(1-p_2)}{n_2}\right)$$

Define,

$$Z = \frac{T_1}{n_1} - \frac{T_2}{n_2} \overset{A}{\sim} N\left(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right)$$

Under H_0,

$$Z \overset{A}{\sim} N\left(0, p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)$$

$$\frac{Z}{\sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \overset{A}{\sim} N(0,1)$$, where p is the common value of p_1 and p_2 under H_0

Since p is unknown, we estimate it as –

$$\hat{p} = \frac{T_1 + T_2}{n_1 + n_2}$$

Under H_0,

$$\tau = \frac{Z}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \overset{A}{\sim} N(0,1)$$

Critical region: $\left\{|\tau_{obs}| > \tau_{\frac{\alpha}{2}}\right\}$ where α is level of significance.

- Comparison of mean of two normal population (Fisher's t test):

Assumptions: Let $X_{11}, X_{12},\ldots, X_{1n_1} \sim N(\mu_1,\sigma^2)$

$X_{21}, X_{22},\ldots, X_{2n_2} \sim N(\mu_2,\sigma^2)$

Hypotheses: Here we want to test, $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$.

Test Statistic: Let us define,

$$\bar{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij};\ s_i^2 = \frac{1}{n_i-1}\sum_{j=1}^{n_i}\left(X_{ij} - \bar{X}_i\right)^2;\ s^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

Then

$$\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

i.e., under H_0,

$$T = \left|\frac{\bar{X}_1-\bar{X}_2}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right| \sim t_{n_1+n_2-2}$$

Rejection Criteria: A large value of T indicates the rejection of H_0, i.e. we reject H_0 if $T > C$ or $T < C$.

C is so chosen that the test is of size α

$$\text{i.e. } P_{H_0}[T > C \text{ or } T < C] = \alpha$$
$$\Rightarrow C = t_{\frac{\alpha}{2};\, n_1+n_2-2},$$

i.e., we reject H_0 at a level of α, if

$$\left|\frac{\bar{X}_1 - \bar{X}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right| > t_{\frac{\alpha}{2};\, n_1+n_2-2}$$

- Testing two populations by Variance Stabilizing Transformation:

Assumptions: Let, p_1 and p_2 denote respectively the proportion of individual possessing a certain attribute A in two independent populations (p_1, p_2: unknown). n_1 and n_2 no. of samples are drawn from first and second population respectively.

Hypotheses: To test, H_0: $p_1 = p_2 = p$ vs. H_1: $p_1 \neq p_2$.

Test Statistic: Let, $\widehat{p_1}$ = estimate of p_1

$\widehat{p_2}$ = estimate of p2

We define,

$$T = \left|\frac{2\left(\sin^{-1}\sqrt{\widehat{p_{ab}}}-\sin^{-1}\sqrt{\widehat{p_{ba}}}\right)}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right| \text{ (refer to Appendix 2)}$$

Rejection Criteria: We reject H_0 at level α, if $T > \tau_{\frac{\alpha}{2}}$.

- Test for comparison of population (Paired t-test):

Assumptions: $(X_1, Y_1), (X_2, Y_2), \ldots (X_n, Y_n) \sim BVN(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$.
Hypotheses: To test $H_0: \mu_x = \mu_y$ vs. $H_1: \mu_x \neq \mu_y$.
Test Statistic: We define, $Z_i = X_i - Y_i \Rightarrow \overline{Z} = \overline{X} - \overline{Y}$
i.e.

$$\frac{\sqrt{n}(\overline{Z}-\mu_Z)}{s_Z} \sim t_{n-1}, \text{ where } s_z^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Z_i - Z)^2,\ \mu_z = \mu_x - \mu_y$$

Under H_0,

$$T = \left|\frac{\sqrt{n}Z}{s_z}\right| \sim t_{n-1}$$

as under H_0, $\mu_z = \mu_x - \mu_y = 0$.
Now, a large value of T indicates the rejection of H_0, i.e. a both tailed test based on T is appropriate.
Rejection Criteria: We reject H_0 if T > C or T < C, C is so chosen that,

$$P_{H_0}[T > C \text{ or } T < C] = \alpha$$
$$\Rightarrow C = t_{\frac{\alpha}{2};\, n-1}$$

We reject H_0, if

$$\left|\frac{\sqrt{n}Z}{s_z}\right| > t_{\frac{\alpha}{2};\, n-1}$$

B. Non-parametric Methods

➢ Median Test:

Median test is a statistical procedure for testing if two independent ordered samples differ in their central tendencies. In other words, it gives information if two independent samples are likely to have been drawn from the populations with the same median.
Assumptions: Let there be two independent ordered samples of sizes m and n, say, x_1, x_2, …, x_m and y_1, y_2, …, y_n, from the populations with p.d.f.'s $f_1(.)$ and $f_2(.)$ respectively. The measurements must be at least ordinal.
We arrange the $N = m + n$ observations in the two samples combined in ascending order of magnitude. Let $z_{(1)}$, $z_{(2)}$, …, $z_{(m+n)}$ be the combined ordered sample.
We determine the combined sample median, say, M.
Then we count the number of observations in each of the two samples that lie below M and the number of those which do not lie below M.
Hypotheses : We want to test if the samples come from the same population or from different populations with the same median, i.e.,

$$H_0 : f_1(.) = f_2(.) \qquad \text{vs} \qquad H_1 : f_1(.) \neq f_2(.)$$

Test Statistic: Let V be the number of x_i's which are $< M$, U be the number of y_i's which are $< M$. So V' $(= m - v)$ is the number of x_i's which are $\geq M$, and U' $(= n - u)$ is the number of x_i's which are $\geq M$. Thus we obtain the p-value as

$$\text{p-value} = = \frac{m!n!(u+v)!(N-u-v)!}{v!(m-v)!u!(n-u)!N!}$$

Whereas for moderately large m, n (say each greater than 10), we may use the χ^2 statistic with 1 degree of freedom, where

$$\chi^2 = \frac{(v(n-u)-(m-v)u)^2 N}{mn(u+v)(N-u-v)}$$

Rejection Criterion : If the observed value of V is significantly small or quite large, then one might suspect that $\xi_{\frac{1}{2}(x)}$ is not equal to $\xi_{\frac{1}{2}(y)}$. If either observed $V < C$ or $V > C$, it indicates the rejection of $H_0 : f_1(.) = f_2(.)$ in favour of $H_1 : f_1(.) \neq f_2(.)$.

➢ Mann Whitney U Test:

Assumptions: Let $(X_1, X_2, ..., X_{n_1})$ be a random sample of size n_1 from a population having absolutely continuous cdf $F_X(.)$ and $(Y_1, Y_2, ..., Y_{n_2})$be another sample of size n_2 from a population having absolutely continuous cdf $F_Y(.)$. Both the samples are independent of each other. *Hypotheses:* We compare the location of the two populations:

$$H_0 : F_X(z) = F_Y(z) \text{ vs } H_1 : F_X(z) \neq F_Y(z)$$

i.e., the location of the X-population is not equal to that of the Y-population under the the condition $F_X(z) \neq F_Y(z)$, i.e., under H_1, X is stochastically larger than Y.

Test Statistic: We define the following:

$$\phi(\mathrm{Xi},\mathrm{Yj}) = \begin{cases} 1 & \text{if } \mathrm{Xi} < \mathrm{Yj} \\ 0 & otherwise \end{cases}$$

$$\psi(\mathrm{Xi},\mathrm{Yj}) = \begin{cases} 1 & \text{if } \mathrm{Xi} > \mathrm{Yj} \\ 0 & otherwise \end{cases}$$

Let us define,

$$U^* = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \phi\,(X_i, Y_j)$$

$$U^{**} = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \psi\,(X_i, Y_j)$$

The Mann-Whitney U statistic is given by,

$$U = \min\,(U^*, U^{**})$$

For moderately large n_1, n_2 (say each greater than 20), we may use the Z statistic, where

$$Z = \frac{U - \mu_U}{\sigma_U} \sim N(0, 1)$$

where,

$$\mu_U = \frac{n_1 n_2}{2} \text{ and } \sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

Rejection Criterion: When we drift away from H_0, it is likely to get more X-values preceeding or exceeding the Y-values, i.e., a larger value of U indicates the rejection of H_0. We reject H_0 if $U < c$ (say), now c is so chosen that,

$$P_{H_0}[U > c] \le \alpha$$

➢ Paired-Sample Wilcoxon Sign Rank Test:

Assumptions: Here the population is continuous and symmetric. Let (X_1, Y_1), $(X_2, Y_2), \ldots, (X_n, Y_n)$ be iid bivariate observations from an absolutely continuous population which is symmetric about the median (θ_x, θ_y).

Hypotheses: Here we want to test,

$$H_0 : \theta_x = \theta_y \quad \text{vs} \quad H_1 : \theta_x \neq \theta_y$$

Test Statistic: We define, $d_i = X_i - Y_i$. Thus median of d, $\theta_d = \theta_x - \theta_y$. So the test boils down to,

$$H_0 : \theta_d = 0 \quad \text{vs} \quad H_1 : \theta_d \neq 0$$

It is to be noted that d_i's are independent.

Note that, $d_i - 0 \overset{d}{\simeq} 0 - d_i$ under H_0, i.e., $d_i \overset{d}{\simeq} -d_i$ under H_0

So d_i's are symmetrically distributed with respect to 0 under H_0. So positive and negative differences of the same absolute value have equal probabilities of occurrence. After calculating $|d_i|\ \forall$ I and ranking them from 1 to n keeping the track of the sign of the i-th rank, we define

T^+: Sum of rank of positive d_i values

T^-: Sum of rank of negative d_i values

Under H_0, $T^+ = T^-$ approximately, so $T = T^+ + T^- = \frac{n(n+1)}{2}$ is our test statistic.

In practice though, smaller of T^+ and T^- is used as test statistic.

Rejection Criterion: If we deviate away from the null then $d_i \neq 0$ would be lesser, i.e., the rank of $|d_i|$ for which $d_i \neq 0$ would be few.

Hence, T^- or T^+ is likely to take smaller values and a both tailed test based on T^- or T^+ is appropriate here, i.e., we reject H_0 if $T^- < c_1$ or $T^+ < c_2$

➢ Paired Sign Test:

Assumptions: Paired sign test is the non-parametric counter-part of the paired t-test. Let $(X_1, Y_1), (X_2, Y_2),...,(X_n, Y_n)$ be a paired data drawn from an absolutely continuous joint cdf F( ). Let $\tilde{\theta}_x$ and $\tilde{\theta}_y$ respectively be the median of X population and Y population.

Hypotheses: Here we want to test the following,

$$H_0 : \tilde{\theta}_x = \tilde{\theta}_y \quad \text{vs} \quad H_1 : \tilde{\theta}_x \neq \tilde{\theta}_y$$

Test Statistic: Let us define, $Z_i = X_i – Y_i$. So median of Z is, $\tilde{\theta}_z = \tilde{\theta}_x - \tilde{\theta}_y$.

So the test boils down to,

$$H_0: \tilde{\theta}_z = 0 \text{ vs } H_1: \tilde{\theta}_z \neq 0.$$

Now note that $Z_1, Z_2,..., Z_n$ are independent. In order to test the above hypothesis, we carry out the sign test based on $Z_1, Z_2,..., Z_n$. We define the following random variable.

$$C_i = \begin{cases} 1 & \text{if } Z_i < \theta_0 \\ 0 & otherwise \end{cases}$$

Let us define the test statistic, $S = \sum_{i=1}^{n} C_i$; under H_0 : $S \sim \text{Bin}\left(n, \frac{1}{2}\right)$ $[\because P[Z_i > 0] = \frac{1}{2}]$ as 0 is the median.

Rejection Criterion: A very small value or a very large value of S indicates the rejection of H_0.

We reject H_0 if $S < c_1$ or $S > c_2$; c_1 and c_2 are so chosen that

$P_{H_0}[S < c_1 \text{ or } S < c_2] = \alpha$

If s be the observed value of S,

$$\text{p-value} = 2 \times \min\{P_{H_0}[S \leq s], P_{H_0}[S \geq s]\}$$

We reject H_0 if p-value $\leq \alpha$.

C. Estimation of Confidence Interval

Adjusted Wald Confidence Interval for a Difference of Binomial Proportions Based on Paired Data:

The analysis of paired data is an important topic in introductory statistics courses. Designs that use paired data have many interesting special cases including repeated-measure designs, pretest-posttest designs, matched-pairs designs, within-subject experiments, and rater-agreement designs.
For the case of two independent samples, the $100(1 - \alpha)\%$ Wald confidence interval for $\pi_1 - \pi_2$ is

$$\widehat{\pi_1} - \widehat{\pi_2} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\widehat{\pi_1}(1 - \widehat{\pi_1})}{n_1} + \frac{\widehat{\pi_2}(1 - \widehat{\pi_2})}{n_2}}$$

where $\widehat{\pi_j} = \frac{f_j}{n_j}$. Agresti and Caffo (2000) show that replacing n_j with $n_j + 2$ and replacing $\widehat{\pi_j} = \frac{f_j}{n_j}$ with Laplace estimates $\widehat{\pi_j} = \frac{f_j+1}{n_j+2}$ dramatically improves the small-sample performance of the stated Wald confidence interval.

III. Collected Data

From the three stage repeated measurement design, we consider the first two stages of the treatment combination. We consider systolic blood pressures after applying the treatment :

$$\begin{cases} 1 \text{ if } \leq 139 \text{ considered as success} \\ 0 \text{ if } > 139 \text{ considered as failure} \end{cases}$$

Systolic blood pressures (mmHg) analysed in 2 sequences are: 1. AB ; 2. BA

Table 1 : For sequence 1. AB

Sl. No.	Period	
	A	B
1	159	140
2	153	172
3	160	156
4	160	200
5	170	170
6	174	132
7	175	155
8	154	138
9	160	170
10	160	160
11	145	140
12	148	154
13	170	170
14	125	130
15	140	112
16	125	140
17	150	150
18	154	145
19	160	140
20	210	190
21	110	112
22	130	140
23	180	190
24	155	120
25	170	164
26	170	140
27	155	130
28	115	110
29	180	136
30	130	120
31	135	140
32	148	148
33	180	180
34	190	155

Table 2 : For sequence 2. BA

Sl. No.	Period	
	B	A
1	165	154
2	160	165
3	140	150
4	140	125
5	158	160
6	180	165
7	170	160
8	140	158
9	126	170
10	130	125
11	144	140
12	140	160
13	120	145
14	145	150
15	155	130
16	168	168
17	150	160
18	160	145
19	156	152
20	195	195
21	130	126
22	130	136
23	140	140
24	160	160
25	140	180
26	140	135
27	100	129
28	148	164
29	150	170
30	205	240
31	140	140
32	154	180
33	150	130
34	140	130

IV. Computational Procedures

A. Crossover Design

i. Parametric Methods

➢ Test for two independent Binomial Proportions by p-value test:

$X_{AB} \sim \text{Bin}(34, p_{AB})$, where $p_{AB} = \frac{10}{34}$ & $n_{AB} = 34$

$X_{BA} \sim \text{Bin}(34, p_{BA})$, where $p_{BA} = \frac{9}{34}$ & $n_{BA} = 34$

Under H_0, $p_{AB} = p_{BA} = p$

Thus

$X_{AB} + X_{BA} \sim \text{Bin}(n_{AB} + n_{BA}, p)$

i.e., $X_{AB} + X_{BA} \sim \text{Bin}(68, p)$

Now, $P[X_{AB} = x | X_{AB} + X_{BA} = t] = \frac{\binom{34}{x}\binom{34}{t-x}}{\binom{68}{t}}$

As we can clearly see that,

the above conditional probability is absolutely free of parameter p (Tending to Hypergeometric)

To test,

H_0: $p_{AB} = p_{BA}$ vs. H_1: $p_{AB} \neq p_{BA}$, where x_{10} = observed valued of X_{AB}

$\therefore$ p-value $= 2\times\min\{P_{H_0}[X_{AB} \geq 9 | X_{AB} + X_{BA} = 19], P_{H_0}[X_{AB} < 9 | X_{AB} + X_{BA} = 19]\}$

$= 2 \times \min\{[\sum_{x:x_{AB}>9} \frac{\binom{34}{x_{AB}}\binom{34}{19-x_{AB}}}{\binom{68}{19}}], [\sum_{x:x_{AB}\leq 9} \frac{\binom{34}{x_{AB}}\binom{34}{19-x_{AB}}}{\binom{68}{19}}]\}$

$= 2 \times \min\{0.75, 0.5\}$

$= 2 \times 0.5$

$= 1.$

$\because$ The p-value is larger than 0.05, we cannot reject the null hypothesis.

➢ Test for two independent Binomial Proportions by Z test:

$\widehat{p_{AB}} = \frac{X_{AB}}{n_{AB}} = \frac{10}{34}$

$\widehat{p_{BA}} = \frac{X_{BA}}{n_{BA}} = \frac{9}{34}$

$X_{AB} \sim \text{Bin}(34, \frac{10}{34})$

$X_{BA} \sim \text{Bin}(34, \frac{9}{34})$

$\alpha = 0.05$

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

H_0: $p_{AB} = p_{BA} = p$ vs. H_1: not H_0

Cumulative area (From left) = 0.025+0.95=0.975

the Z-values are $\pm(1.9 + 0.06) = \pm 1.96$ (From positive Z-score table)

Here, p = pooled proportion $= \frac{10+9}{34+34} = \frac{19}{68}$

Under H_0 : $Z_c = \frac{\left(\frac{10}{34} - \frac{9}{34}\right) - 0}{\sqrt{\frac{19}{68}\left(1 - \frac{19}{68}\right)\left(\frac{1}{34} + \frac{1}{34}\right)}} = 0.2702$

So (+0.2702) lies somewhere between 0 to +1.96.

Since the Z-value doesn't exceed 1.96, i.e it falls in the acceptance region.

Therefore, we can conclude that, we cannot reject the null hypothesis.

- Comparison of mean of two normal population (Fisher's t test):

Here, $n_1 = n_2 = 34$
$\overline{X}_1 = \overline{c} = 148.5$
$\overline{X}_2 = \overline{d} = 154.029$
$s^2_{AB} = 517.7727$
$s^2_{BA} = 544.393$
$\therefore s^2 = 531.0825$
$\therefore s = 23.0453487$

Then, $T = \left| \frac{(\overline{X}_1 - \overline{X}_2)}{s\sqrt{\frac{1}{n_{AB}} + \frac{1}{n_{BA}}}} \right|$
$= 0.9892132207$

and $t_{0.025;66} = 1.997$, where $\alpha = 0.05$
$\therefore$ We can see that,

$$T > t_{0.025;66} = 1.997$$

So we can't reject H_0.

- Testing two populations by Variance Stabilizing Transformation:

Here, $n_{AB} = 34 = n_{BA}$

$p_{AB} = \frac{10}{34}$, $p_{BA} = \frac{9}{34}$

Under H_0, $T = \frac{2(\sin^{-1}\sqrt{p_{AB}} - \sin^{-1}\sqrt{p_{BA}})}{\sqrt{\frac{1}{n_{AB}} + \frac{1}{n_{BA}}}} \xrightarrow{d} N(0,1)$

$\therefore T = 15.4896784$

To test, $H_0: p_{AB} = p_{BA}$ vs H_1: not H_0
We reject H_0 at level α, if $|T| > \tau_{\alpha/2} = 1.960$
$\because |T| > \tau_{\alpha/2}$
so we simply reject H_0.

ii. Non-parametric Methods

➢ Median Test:

We have two independent samples x_i's (results after applying treatment sequence AB) and y_i's (results after applying treatment sequence BA).

The Hypotheses: $H_0 : f_1(.) = f_2(.)$ vs. $H_1 : f_1(.) \neq f_2(.)$

Here a both tailed test is needed.

Calculations: Let us order the data.

Ordered Data: 110, 112, 112, 120, 120, 125, 125, 126, 129, 130, 130, 130, 130, 130, 132, 135, 136, 136, 138, 140, 140, 140, 140, 140, 140, 140, 140, 140, 140, 145, 145, 145, 148, 150, 150, 150, 152, 154, 154, 155, 155, 156, 158, 160, 160, 160, 160, 160, 160, 164, 164, 165, 165, 168, 170, 170, 170, 170, 170, 172, 180, 180, 180, 190, 190, 195, 200, 240.

Combined Sample Median: No. of total observations = 68

Since 68 is even number, then

$$\text{Combined Sample Median} = \text{Average of } \left(\frac{68}{2}\right)^{\text{th}} \text{ and } \left(\frac{68}{2}+1\right)^{\text{th}} \text{ values}$$
$$= \text{Average of } 34^{\text{th}} \text{ and } 35^{\text{th}} \text{ values}$$
$$= \frac{150+150}{2}$$
$$= 150$$

Exact Test: $P = \frac{34!34!33!35!}{19!15!14!20!68!} = 0.0934$

Since $P > 0.05$, we accept H_0.

Approximate Test: Here since m, n are moderately large, then we can use the χ^2 statistic with degree of freedom 1, where

$$\chi^2_{\text{obs}} = \frac{(19.20-15.14)^2 68}{34.34.33.35}$$
$$= 1.47186$$

From χ^2-distribution table, critical value for 5% level α and degrees of freedom 1 is 3.841.

As $-3.841 < \chi^2_{\text{obs}} = 1.472 < 3.841$, χ^2_{obs} falls in the acceptance region.

Hence, we accept $H_0 : f_1(.) = f_2(.)$.

- Mann Whitney U Test:

We have two independent samples x_i's (results after applying treatment sequence AB) and y_i's (results after applying treatment sequence BA).
The Hypotheses: $H_0 : F_X(z) = F_Y(z)$ vs. $H_1 : F_X(z) \neq F_Y(z)$ for some z
Calculations:

Table for necessary calculations

Sl. No.	AB (X_i)	BA (Y_j)	No. of X_i's < Y_j	No. of X_i's > Y_j
1	140	154	20	13
2	172	165	26	8
3	156	150	19	14
4	200	125	5	29
5	170	160	24	9
6	132	165	26	8
7	155	160	24	9
8	138	158	24	10
9	170	170	26	5
10	160	125	5	29
11	140	140	10	17
12	154	160	24	9
13	170	145	17	16
14	130	150	19	14
15	112	130	5	27
16	140	168	26	8
17	150	160	24	9
18	145	145	17	16
19	140	152	20	14
20	190	195	33	1
21	112	126	5	29
22	140	136	8	25
23	190	140	10	17
24	120	160	24	9
25	164	180	30	3
26	140	135	8	26
27	130	129	5	29
28	110	164	25	8
29	136	170	26	5
30	120	240	34	0
31	140	140	10	17
32	148	180	30	3
33	180	130	5	27
34	155	130	5	27
Total	–	–	3985	490

Here $n_1 = n_2 = 34$
$U^* = \Sigma$ (# X_i's < Y_j) = 3985
$U^{**} = \Sigma$ (# X_i's > Y_j) = 490

Our test statistic, U = min (U*, U**) = 490
As n_1 and n_2 are both > 20, we will use the large sample approximation of the test statistic U.

$$\mu_U = \frac{34+34}{2}$$
$$= 34$$
$$\sigma_U = \sqrt{\frac{34.34(34+34+1)}{12}}$$
$$= 81.5291$$
$$Z = \frac{490 - 34}{81.5291}$$
$$= 5.5931$$

From the Z-table, we can see that the criticial Z-value for 5% level of significance is 1.96.
Since our observed value of Z = 5.5931 > 1.96, i.e., the critical Z-value we reject H_0.

B. Parallel Design

i. Parametric Methods

➢ Test for two sample binomial proportion by p-value test:

$X_A \sim$ Bin (34, p_A), where $p_A = \frac{7}{34}$ & $n_A = 34$
$X_B \sim$ Bin (34, p_B), where $p_B = \frac{6}{34}$ & $n_B = 34$
Under H_0, $p_A = p_B = p$
Thus, $X_A + X_B \sim$ Bin (68, p)
Now, $P[X_A = x | X_A + X_B = t] = \frac{\binom{34}{x}\binom{34}{t-x}}{\binom{68}{t}}$

As we can clearly see that, the above conditional probability is absolutely free of parameter p (Tending to Hypergeometric)
To test, H_0: $p_A = p_B$ vs. H_1: $p_A \neq p_B$, where x_{10} = observed valued of X_A
$\therefore$ p-value $= 2 \times \min\{P_{H_0}[X_A \geq 7 | X_A + X_B = 13], P_{H_0}[X_A < 7 | X_A + X_B = 13]\}$

$$= 2 \times \min\{[\textstyle\sum_{x:x_A > 7} \frac{\binom{34}{x_A}\binom{34}{13-x_A}}{\binom{68}{13}}], [\textstyle\sum_{x:x_A \leq 7} \frac{\binom{34}{x_A}\binom{34}{13-x_A}}{\binom{68}{13}}]\}$$
$$= 2 \times \min\{0.5, 0.7306\}$$
$$= 2 \times 0.5 = 1.$$

$\because$ The p-value is larger than 0.05, we cannot reject H_0.

➢ Test for two sample binomial proportion by Z test:

$\widehat{p_A} = \frac{7}{34}$
$\widehat{p_B} = \frac{6}{34}$
$X_A \sim$ Bin (34, $\frac{7}{34}$)
$X_B \sim$ Bin (34, $\frac{6}{34}$)
To test, H_0: $p_A = p_B = p$ vs. H_1: not H_0
Here, p=pooled proportion $= \frac{7+6}{34+34} = \frac{13}{68}$

So, our critical Z-value, $Z_c = \frac{(\widehat{p_A} - \widehat{p_B}) - 0}{\sqrt{p(1-p)\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}}$, Under H_0: $p_A = p_B \Rightarrow (p_A - p_B) = 0$

$$= \frac{\left(\frac{7}{34} - \frac{6}{34}\right) - 0}{\sqrt{\frac{13}{68}\left(1 - \frac{13}{68}\right)\left(\frac{1}{34} + \frac{1}{34}\right)}}$$

$$= 0.3084$$

So (+0.3084) lies somewhere between 0 to +1.96.

Since the Z-value doesn't exceed 1.96, i.e it falls in the acceptance region.

Therefore, we can conclude that, we cannot reject the null hypothesis.

- Comparison of mean of two normal population (Fisher's t test):

Here, $n_1 = n_2 = 34$
$\overline{X}_1' = \overline{c'} = 155.8824$
$\overline{X}_2' = \overline{d}' = 149.0882$
$s_A^2 = 467.5615$
$s_B^2 = 407.6586$
$\therefore s^2 = 437.6100836$
$\therefore s = 20.919132$
Then

$$\frac{(\overline{X}_1' - \overline{X}_2') - (p_A - p_B)}{s\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \sim t_{n_A + n_B - 2}$$

$\because$ under H_0 $p_A = p_B$

$$T = \left| \frac{(\overline{X}_1' - \overline{X}_2')}{s\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \right|$$
$$= 1.3391$$

and $t_{0.025;66} = 1.997$, where $\alpha = 0.05$

$\therefore$ We can see that, $T < t_{0.025;66} = 1.997$
So we can't reject H_0.

- Testing two populations by Variance Stabilizing Transformation:

Here, $n_A = n_B = 34$

$p_A = \frac{7}{34}$
$p_B = \frac{6}{34}$

$\therefore T = 14.1579$

To test, $H_0: p_a = p_b$ vs H_1: not H_0
We reject H_0 at level α, if $|T| > \tau_{\alpha/2} = 1.96$
$\because$ We can see that

$$T > \tau_{\alpha/2}$$

so we simply reject H_0.

ii. Non-parametric Methods

➢ Median Test

We have two independent samples x_i's (results after applying treatment A) and y_i's (results after applying treatment B).

The Hypotheses: $H_0 : f_1(.) = f_2(.)$ vs. $H_1 : f_1(.) \neq f_2(.)$

Here a both tailed test is needed.

Calculations: Let us order the data.

Ordered Data: 100, 110, 115, 120, 125, 125, 126, 130, 130, 130, 130, 130, 135, 140, 140, 140, 140, 140, 140, 140, 140, 140, 140, 144, 145, 145, 148, 148, 148, 150, 150, 150, 150, 153, 154, 154, 154, 155, 155, 155, 156, 158, 159, 160, 160, 160, 160, 160, 160, 160, 160, 165, 168, 170, 170, 170, 170, 170, 174, 175, 180, 180, 180, 180, 190, 195, 205, 210.

Combined Sample Median: No. of total observations = 68

Since 68 is even number,

Then,

Combined Sample Median = Average of $\left(\frac{68}{2}\right)^{\text{th}}$ and $\left(\frac{68}{2}+1\right)^{\text{th}}$ values

$= \text{Average of } 34^{\text{th}} \text{ and } 35^{\text{th}} \text{ values}$

$= \frac{153+154}{2}$

$= 153.5$

Exact Test: $P = \frac{34!34!34!34!}{13!21!21!13!68!} = 0.0302$

Since $P < 0.05$, we reject H_0.

Approximate Test: Here since m, n are moderately large, then we can use the χ^2 statistic with degree of freedom 1, where

$$\chi^2_{\text{obs}} = \frac{(13.13-21.21)^2 68}{34.34.34.34} = 3.7647$$

From χ^2-distribution table (refer to appendix), critical value for 5% level α and degrees of freedom 1 is 3.841.

As $-3.841 < \chi^2_{\text{obs}} = 3.765 < 3.841$, χ^2_{obs} falls in the acceptance region.

Hence, we accept $H_0 : f_1(.) = f_2(.)$.

- Mann Whitney U Test:

We have two independent samples x_i's (results after applying treatment sequence AB) and y_i's (results after applying treatment sequence BA).
The Hypotheses: $H_0 : F_X(z) = F_Y(z)$ vs. $H_1 : F_X(z) \neq F_Y(z)$ for some z
Calculations:

Table for necessary calculations

Sl. No.	A (X_i)	B (Y_j)	No. of X_i's < Y_j	No. of X_i's > Y_j
1	159	165	23	11
2	153	160	18	11
3	160	140	7	26
4	160	140	7	26
5	170	158	17	17
6	174	180	29	2
7	175	170	23	7
8	154	140	7	26
9	160	126	4	30
10	160	130	4	28
11	145	144	8	26
12	148	140	7	26
13	170	120	2	32
14	125	145	8	25
15	140	155	15	17
16	125	168	23	11
17	150	150	11	22
18	154	160	18	11
19	160	156	17	17
20	210	195	33	1
21	110	130	4	28
22	130	130	4	28
23	180	140	7	26
24	155	160	18	11
25	170	140	7	26
26	170	140	7	26
27	155	100	0	34
28	115	148	9	23
29	180	150	11	22
30	130	205	33	1
31	135	140	7	26
32	148	154	13	19
33	180	150	11	22
34	190	140	7	26
Total	–	–	419	690

Here $n_1 = n_2 = 34$
$U^* = \Sigma$ (# X_i's < Y_j) = 419
$U^{**} = \Sigma$ (# X_i's > Y_j) = 690

Our test statistic, U = min (U*, U**) = 419

As n_1 and n_2 are both > 20, we will use the large sample approximation of the test statistic U.

$$\mu_U = \frac{34+34}{2}$$

$$= 34$$

$$\sigma_U = \sqrt{\frac{34.34(34+34+1)}{12}}$$

$$= 81.5291$$

$$Z = \frac{419 - 34}{81.5291}$$

$$= 4.7222$$

From the Z-table, we can see that the criticial Z-value for 5% level of significance is 1.96.
Since our observed value of Z = 4.7222 > 1.96, i.e., the critical Z-value we reject H_0.

C. Paired Sample Tests

i. Parametric Methods

- Test for comparison of population (Paired t-test):
  - For treatment combination AB:

We define, $s_s^2 = 341.8797$, where $n_A = 34$

We can observe that (reference to spreadsheet),

$$T = \left|\frac{\sqrt{n}\bar{s}}{s_s}\right| = 2.3281 > t_{\alpha/2;\, nB-1} = t_{0.025;\, 33} = 2.035$$

i.e. we reject H_0.

  - For treatment combination BA:

We define, $s_t^2 = 298.8449$, where $n_B = 34$

We can observe that (reference to spreadsheet),

$$T = \left|\frac{\sqrt{n}\bar{t}}{s_t}\right| = |-1.6666| = 1.6666 < t_{\alpha/2;\, nB-1} = t_{0.025;\, 33} = 2.035$$

i.e. we fail to reject H_0.

ii. Non-parametric Methods

- For treatment combination AB:

➢ Paired Sign Test:

Let us define the following random variable.

$$C_i = \begin{cases} 1 & \text{if } Z_i < \theta_0 \\ 0 & \text{otherwise} \end{cases}$$

Let us also define, $S = \sum_{i=1}^{n} C_i$.

The Hypotheses:

H_0 : Population median difference = 0 vs. H_1 : Population median difference ≠ 0

Here a two tailed test is necessary.

Table for necessary calculations

A (X_i)	B (Y_i)	Sign of $Z_i = X_i - Y_i$
159	140	+
153	172	-
160	156	+
160	200	-
~~170~~	~~170~~	
174	132	+
175	155	+
154	138	+
160	170	-
~~160~~	~~160~~	
145	140	+
148	154	-
~~170~~	~~170~~	
125	130	-
140	112	+
125	140	-
~~150~~	~~150~~	
154	145	+
160	140	+
210	190	+
110	112	-
130	140	-
180	190	-
155	120	+
170	164	+
170	140	+
155	130	+
115	110	+
180	136	+
130	120	+
135	140	-
~~148~~	~~148~~	
~~180~~	~~180~~	
190	155	+

T^+ : 18 and T^- : 10

As $T^- < T^+$, so we will take T^- as our test statistic here.

$\therefore S = T^- = 10$

There are 6 observations where no change is observed. So we exclude those observations.

Hence, Revised n = 34 – 6 = 28

Now, under H_0 : $S \sim \text{Bin}\left(28, \frac{1}{2}\right)$

$$P(S \geq 10) = \sum_{x=10}^{28} \binom{28}{x} \left(\frac{1}{2}\right)^{28}$$

$$= 0.9564$$

$$P(S \leq 10) = \sum_{x=0}^{10} \binom{28}{x} \left(\frac{1}{2}\right)^{28}$$

$$= 0.0925$$

Now p-value of both tailed test $= 2 \times \min\{P_{H_0}[S \leq 10], P_{H_0}[S \geq 10]\}$

$= 2 \times \min\{0.0925, 0.9564\}$

$= 2 \times 0.0925$

$= 0.185$

As p-value $= 0.185 > 0.05$

Hence we accept H_0 at 5% level α.

- Paired Sample Wilcoxon Signed Rank Test:

The Hypotheses:
H_0 : Population median difference = 0 vs. H_1 : Population median difference $\neq$ 0
Here a both tailed test is needed.

Table for necessary calculations

x_i (A)	y_i (B)	$d_i = x_i - y_i$	Signed Ranks
159	140	19	+16.5
153	172	-19	-16.5
160	156	4	+2
160	200	-40	-27
~~170~~	~~170~~		
174	132	42	+28
175	155	20	+19
154	138	16	+15
160	170	-10	-11.5
~~160~~	~~160~~		
145	140	5	+4.5
148	154	-6	-7.5
~~170~~	~~170~~		
125	130	-5	-4.5
140	112	28	+22
125	140	-15	-14
~~150~~	~~150~~		
154	145	9	+9
160	140	20	+19
210	190	20	+19
110	112	-2	-1
130	140	-10	-11.5
180	190	-10	-11.5
155	120	35	+24.5
170	164	6	+7.5
170	140	30	+23
155	130	25	+21
115	110	5	+4.5
180	136	44	+28
130	120	10	+11.5
135	140	-5	-4.5
~~148~~	~~148~~		
~~180~~	~~180~~		
190	155	35	+24.5

$T^+ = 298.5$
$T^- = 109.5$
As $T^- < T^+$, so we will take T^- as our test statistic here.
There are 6 observations where no change is observed. So we exclude those observations.
Hence, Revised n = 34 – 6 = 28

From Wilcoxon Signed Rank Table, Critical Value = 116.
Now, Critical Value = 116 > 109.5 = T^-
So H_0 is rejected.

- For treatment combination BA:

➢ Paired Sign Test:

Let us define the following random variable.

$$C_i = \begin{cases} 1 & \text{if } Z_i < \theta_0 \\ 0 & \textit{otherwise} \end{cases}$$

Let us also define, $S = \sum_{i=1}^{n} C_i$.

The Hypotheses:

H_0 : Population median difference = 0 vs. H_1 : Population median difference ≠ 0
Here a two tailed test is necessary.

<u>Table for necessary calculations</u>

B (X_i)	A (Y_i)	Sign of $Z_i = X_i - Y_i$
165	154	+
160	165	-
140	150	-
140	125	+
158	160	-
180	165	+
170	160	+
140	158	-
126	170	-
130	125	+
144	140	+
140	160	-
120	145	-
145	150	-
155	130	+
~~168~~	~~168~~	
150	160	-
160	145	+
156	152	+
~~195~~	~~195~~	
130	126	+
130	136	-
~~140~~	~~140~~	
~~160~~	~~160~~	
140	180	-
140	135	+
100	129	-
148	164	-
150	170	-
205	240	-
~~140~~	~~140~~	
154	180	-
150	130	+
140	130	+

$T^+ : 13$
$T^- : 16$
As $T^+ < T^-$, so we will take T^+ as our test statistic here.
$\therefore S = T^+ = 13$
There are 5 observations where no change is observed. So we exclude those observations.
Hence, Revised n = 34 – 5 = 29
Now, under $H_0 : S \sim \text{Bin}\left(29, \frac{1}{2}\right)$

$$P(S \geq 13) = \sum_{x=13}^{29} \binom{29}{x} \left(\frac{1}{2}\right)^{29} = 0.7709$$

$$P(S \leq 13) = \sum_{x=0}^{13} \binom{29}{x} \left(\frac{1}{2}\right)^{29}$$
$$= 0.3555$$

Now p-value of both tailed test $= 2 \times \min\{P_{H_0}[S \leq 13], P_{H_0}[S \geq 13]\}$
$$= 2 \times \min\{0.3555, 0.7709\}$$
$$= 2 \times 0.3555$$
$$= 0.7111$$

As p-value $= 0.7111 > 0.05$
Hence we accept H_0 at 5% level α.

- Paired Sample Wilcoxon Signed Rank Test:

The Hypotheses:
H_0 : Population median difference = 0 vs. H_1 : Population median difference ≠ 0
Here a both tailed test is needed.

Table for necessary calculations(for treatment combination BA)

x_i (B)	y_i (A)	$d_i = x_i - y_i$	Signed Ranks
165	154	11	+14
160	165	-5	-6.5
140	150	-10	-11.5
140	125	15	+16
158	160	-2	-1
180	165	15	+16
170	160	10	+11.5
140	158	-18	-19
126	170	-44	-29
130	125	5	+6.5
144	140	4	+3
140	160	-20	-21
120	145	-25	-23.5
145	150	-5	-6.5
155	130	25	+23.5
~~168~~	~~168~~		
150	160	-10	-11.5
160	145	15	+16
156	152	4	+3
~~195~~	~~195~~		
130	126	4	+3
130	136	-6	-9
~~140~~	~~140~~		
~~160~~	~~160~~		
140	180	-40	-28
140	135	5	+6.5
100	129	-29	-26
148	164	-16	-18
150	170	-20	-21
205	240	-35	-27
~~140~~	~~140~~		
154	180	-26	-25
150	130	20	+21
140	130	10	+11.5

$T^+ = 151.5$
$T^- = 283.5$
As $T^+ < T^-$, so we will take T^+ as our test statistic here.

There are 5 observations where no change is observed. So we exclude those observations.
Hence, Revised $n = 34 - 5 = 29$
From Wilcoxon Signed Rank Table, Critical Value = 126.
Now, Critical Value $= 126 < 151.5 = T^+$
So H_0 is accepted.

D. Estimation of Confidence Interval

➢ Adjusted Wald Confidence Interval for a Difference of Binomial Proportions Based on Paired Data:

For the case of two independent samples, the 100(1 – α)% Wald confidence interval for $\widehat{p_{BA}} - \widehat{p_{AB}}$ is

$$CI = \widehat{p_{BA}} - \widehat{p_{AB}} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\widehat{p_{AB}}(1-\widehat{p_{AB}})}{n_{AB}} + \frac{\widehat{p_{BA}}(1-\widehat{p_{BA}})}{n_{BA}}},$$

where $n_{AB} = 34$, $p_{AB} = \frac{X_{AB}}{n_{AB}} = \frac{10}{34}$ & $n_{BA} = 34$, $p_{BA} = \frac{X_{BA}}{n_{BA}} = \frac{9}{34}$, $z_{\frac{\alpha}{2}} = 1.96$

The Laplace estimates –

$$\widehat{p_{AB}} = \frac{X_{AB}+1}{n_{AB}+2} = = \frac{10+1}{34+2} = \frac{11}{36}$$
$$\text{and } \widehat{p_{BA}} = \frac{X_{BA}+1}{n_{BA}+2} = = \frac{9+1}{34+2} = \frac{10}{36}$$

dramatically improves the small-sample performance of the stated Wald confidence interval.
∴ the Wald confidence interval for $\widehat{p_{BA}} - \widehat{p_{AB}}$ = [0.1882, –0.2437]

V. Conclusion

In majority of the tests of hypothesis we carried out for crossover design, except for the test by Variance Stabilizing Transformation (VST), a parametric test and Mann Whitney U test, a non-parametric test, we did not have enogh evidence to reject the null hypothesis, i.e., the treatment A and B are equally effective.

Similarly, among the tests done for parallel design, except for the VST and Mann Whitney U test, we failed to reject the null hypothesis. Thus leading to our conclusion that the treatments A and B individually are almost similar.

On the other hand, we carried out paired tests to check which treatment is better. In case of treatment combination AB, except for the paired sample sign test, we have rejected the null hypothesis in paired t-test and paired sample Wilcoxon signed rank test. Hence we can simply conclude that in treatment combination AB, treatment B is more effective than treatment A. Meanwhile, in case of the treatment combination BA, the results show that the treatments individually are equally effective.

There is a 95% chance that the confidence interval of $[-0.2437, 0.1882]$ contains the true difference in the proportion of treatment combinations BA and AB.

In this study, we cannot effectively conclude which treatment is better than the other, that means we claim that both the drugs might give the same and equal effect to some extent.

Here, we have discussed the possible inferential testing for this data and the study provides inferential procedures for the treatment effects, we've concluded about acceptance or rejection of those two treatments A and B. But a few limitations are here: like, the difference between the results of the tests maybe due to the approximation, or there maybe some other hidden cause, which we failed to notice.
Rectifying these drawbacks for betterment of our study will be our interest in future.

VI. Appendix

1. The conditional probability of $X_1 | X_1 + X_2$ is given by

$$P[X_1 = x | X_1 + X_2 = t] = \frac{P[X_1 = x, X_2 = t - x]}{P[X_1 + X_2 = t]}$$
$$= \frac{P(X_1 = x) . P(X_2 = t - x)}{P(X_1 + X_2 = t)}$$
$$= \frac{\binom{n_1}{x} p^2 (1-p)^{n_1 - x} \binom{n_2}{t-x} p^{t-x} (1-p)^{n_2 - t + x}}{\binom{n_1+n_2}{t} p^t (1-p)^{n_1+n_2-t}}$$
$$= \frac{\binom{n_1}{x}\binom{n_2}{t-x}}{\binom{n_1+n_2}{t}}$$

i.e. the conditional distribution of $X_1 | X_1 + X_2$ is free of the parameter p [tends to Hypergeometric].

2. $$\sqrt{n_1}\left(\sin^{-1}\sqrt{\widehat{p_1}} - \sin^{-1}\sqrt{p_1}\right) \xrightarrow{d} N\left(0, \frac{1}{4}\right)$$

$$\Rightarrow \sin^{-1}\sqrt{\widehat{p_1}} \xrightarrow{d} N\left(\sin^{-1}\sqrt{p_1}, \frac{1}{4n_1}\right)$$

Similarly, $\sin^{-1}\sqrt{\widehat{p_2}} \xrightarrow{d} N\left(\sin^{-1}\sqrt{p_2}, \frac{1}{4n_2}\right)$

$$\therefore \left(\sin^{-1}\sqrt{\widehat{p_1}} - \sin^{-1}\sqrt{\widehat{p_2}}\right) \xrightarrow{d} N\left(\sin^{-1}\sqrt{p_1} - \sin^{-1}\sqrt{p_2}, \frac{1}{4}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)$$

Under H_0, $p_1 = p_2 = p$

$$\Rightarrow \left(\sin^{-1}\sqrt{\widehat{p_1}} - \sin^{-1}\sqrt{\widehat{p_2}}\right) \xrightarrow{d} N\left(0, \frac{1}{4}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)$$

i.e. $T = \dfrac{2\left(\sin^{-1}\sqrt{\widehat{p_1}} - \sin^{-1}\sqrt{\widehat{p_2}}\right)}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \xrightarrow{d} N(0,1)$, under H_0.

VII. Reference

- Matthews, J.N.S. (1987): Biometrica, Issue 74
- Pocock, Stuart J. (1983): Clinical Trials A Practical Approach
- Pearson, E. (1976): Biometrika Tables for Statisticians
- Agresti, A. & Caffo, B. (2000): The American Statistician, Volume 54, Issue 4, Article: Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Result from Adding Two Successes and Two Failures
- Goon, A.M., Gupta, M.K. and Dasgupta, B. (2002): Fundamentals of Statistics, vol. 1, 8th edition, The World Press
- https://www.heart.org/en/health-topics/high-blood-pressure/the-facts-about-high-blood-pressure
- https://www.researchgate.net/publication/258150348_Adjusted_Wald_Confidence_Interval_for_a_Difference_of_Binomial_Proportions_Based_on_Paired_Data
- Z-table: https://www.ztable.net/
- t-table: https://www.medcalc.org/manual/t-distribution-table.php
- Wilcoxon table: https://www.real-statistics.com