Robustness of t-test

Name : Shreya Saha

Roll No. : 442

Registration No. : A01-2122-0855-21

Department : Statistics

Session : 2021-’24

Supervisor : Dr. Surupa Chakraborty (Roy)

St. Xavier’s College (Autonomous), Kolkata

Declaration:

I affirm that I have identified all my sources and that no part of my dissertation paper uses unacknowledged materials.

Student’s Signature:

PAPER NAME

dissertation 442.pdf

AUTHOR

Shreya Saha

WORD COUNT

2352 Words

CHARACTER COUNT

12186 Characters

PAGE COUNT

20 Pages

FILE SIZE

1.2MB

SUBMISSION DATE

Apr 4, 2024 11:44 AM GMT+5:30

REPORT DATE

Apr 4, 2024 11:44 AM GMT+5:30

3% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 3% Internet database
- 2% Publications database
- Crossref database
- Crossref Posted Content database

Excluded from Similarity Report

- Bibliographic material
- Quoted material
- Cited material
- Abstract
- Methods and Materials
- Small Matches (Less then 15 words)

4% words excluded by Custom Sections

Table of Content

Introduction

A t-test is widely used statistical tool offering practical significance in various research fields, especially when the sample size is sufficiently small.

t-test helps researchers to draw conclusions based on data analysis. Integrative uses of t-test make it so popular. It is used in various fields such as medical research, education, marketing research as well as in quality control. Thus, it is not a test that solely comes from statistical background.

Student t-test or one sample t-test is one of the branches that compares the sample mean to the hypothesized mean value relative to the variability in the sample. The result indicates whether the difference between the two values is statistically significant. t-test takes certain assumptions which may not always attained by the provided data. Here, the property "Robustness of t-test" will be studied using analytical and graphical approach. Simulation will be done whenever needed. Thus, we will try to study the properties of Student t-test in a descriptive way.

1. Parametric Test :

Parametric test is a statistical test that makes certain assumptions about the parameters of population distribution from which the sample is drawn. Most of the time, these assumptions include that the data is normally distributed.

Our interest of study lies on one sample t-test. Here we consider one tailed test only (greater than alternatives).

2. Genesis of t-statistic :

A t-statistic with n degree of freedom is given by-

$$t = \frac{X}{\sqrt{Y^2/n}}$$

where, X is a standard normal variable and Y^2 is a chi-square with n degrees of freedom and is independent of X.

3. Student's t-test : The General Idea

t-test is a parametric test that efficiently determines population parameter, mainly the mean, from small sample size but under certain assumption about the data. The assumptions underlying one sample t-test are-

- The observations are independent and randomly drawn from a population.
- They are approximately normally distributed.

Now, let us look how t-test works.

Let us consider a hypothesis as follows-

$H_0 : \mu=\mu_0$ against $H_1 : \mu>\mu_0$

Now, suppose $X_1, X_2, \ldots, X_n$ be an independent random sample from a distribution having cumulative distribution function F_x.

Assume that, X~ Normal(μ,σ^2) , $\mu \in \mathbb{R}$, $\sigma^2>0$: both unknown

$$\Rightarrow \bar{X} \sim \text{Normal}(\mu, \frac{\sigma 2}{n})$$

Then, $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ ~N(0,1)

As σ is unknown we estimate and replace it using sample standard deviation , that is given by-

$$S = \sqrt{\frac{1}{n-1}\sum_{i=0}^{n}(X_i - \bar{X})^2}$$

Hence, by central limit theorem,

$$t = \frac{\sqrt{n}(\bar{X}-\mu)/\sigma}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}}$$

$$= \frac{\sqrt{n}(\bar{X}-\mu)}{S}$$

Considering the realizations,

Under H_0 , $t = \frac{\sqrt{n}(\bar{x}-\mu_0)}{s} \sim t_{(n-1)}$

Therefore, we reject H_0 against H_1 at α level of significance if $t_{obs.} > t_{\alpha,(n-1)}$.

4. Level :

Significance level is the probability of rejecting null hypothesis given that it is true, that is probability of making wrong decision. It is denoted by α. Also, it is the probability of occurring Type-I-error.

Level = P(reject H_0 | H_0 is true)

$$= P\left(\frac{\sqrt{n}(\bar{x}-\mu_0)}{s} > t_{(n-1)}\right)$$

5. Power :

Power of a test is the probability of rejecting null hypothesis when alternative one is true, that is probability of making correct decision. This is denoted by 1-β, where β is the probability of Type-II-error.

Power = P(reject H_0 | H_1 is true)

$$= P\left(\frac{\sqrt{n}(\bar{x}-\mu)}{s} > t_{(n-1)} \mid H_1 \text{ is true}\right)$$

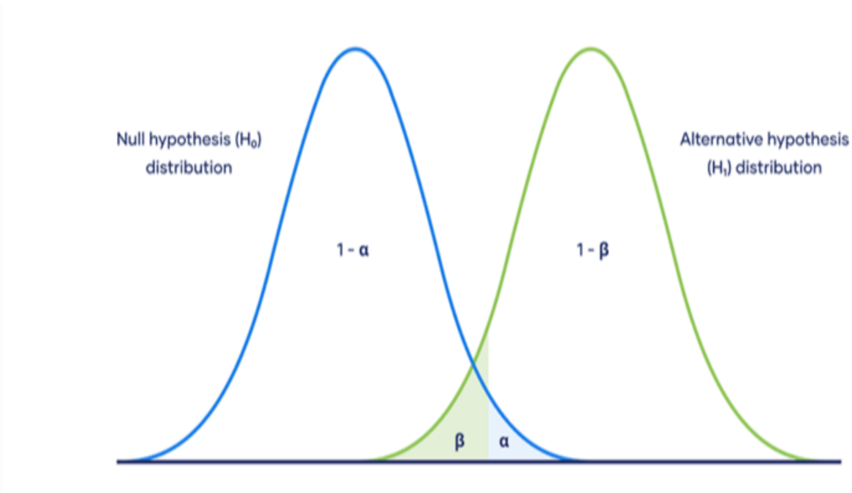

Figure 5.1

6. Simulation based study using t-test

The motive of our study is to check whether t-test is at all applicable for non-normal data. To check this we perform a simulation study.

Procedure :

Let us consider a samples of size n=20.

Let the sample is denoted by X_1, X_2, …,X_{20} ,drawn from a population distribution having cdf $F_{x.}$ We repeat this procedure 1000 times. The random samples are non-stochastic. Here we consider some population distributions just to generate the sample observations.

To test,

$$H_0 : \mu=1 \quad \text{against} \quad H_1 : \mu=2$$

In this study we will consider right tailed test only. We will choose the value of parameters of each distribution such a way that the population mean under null hypothesis is 1 and the same under alternative is 2.

We use R to generate random sample. To avoid any error due to sampling in between the data, we shall be using random seed = 1234. We proceed to calculate empirical level and power under some chosen population distribution using those 1000 samples. We continue this procedure for each population distribution.

Findings :

Distribution	Empirical Level	Empirical Power
Normal	0.048	0.993
Exponential	0.022	0.933
Uniform	0.056	0.799
Gamma	0.008	0.704
Cauchy	0.031	0.289
Logistic	0.047	0.779

Table 6.1

Interpretation :

We have taken the level of significance as $\alpha = 0.05$ to carry on the simulation based study.

Observe that, power of normal distribution is highest and its empirical level is also around 0.05 as pre-specified. Thus we can conclude that t-test gives best result for normally distributed data. Whereas, Gamma distribution has empirical level $0.008 \ll 0.05$ and Cauchy distribution has empirical power $0.289 \ll 1$ which is not at all satisfactory. Therefore, Gamma and Cauchy distribution give the worst result under t-test .

Power Curves :

Power curves depict the relationship between power of a statistical test and certain factors like sample size, significance level etc. Here, we plot the power against the specified values of parameter.

Let us draw the power curves of the above mentioned distributions to compare their behavior under one sample t-test.

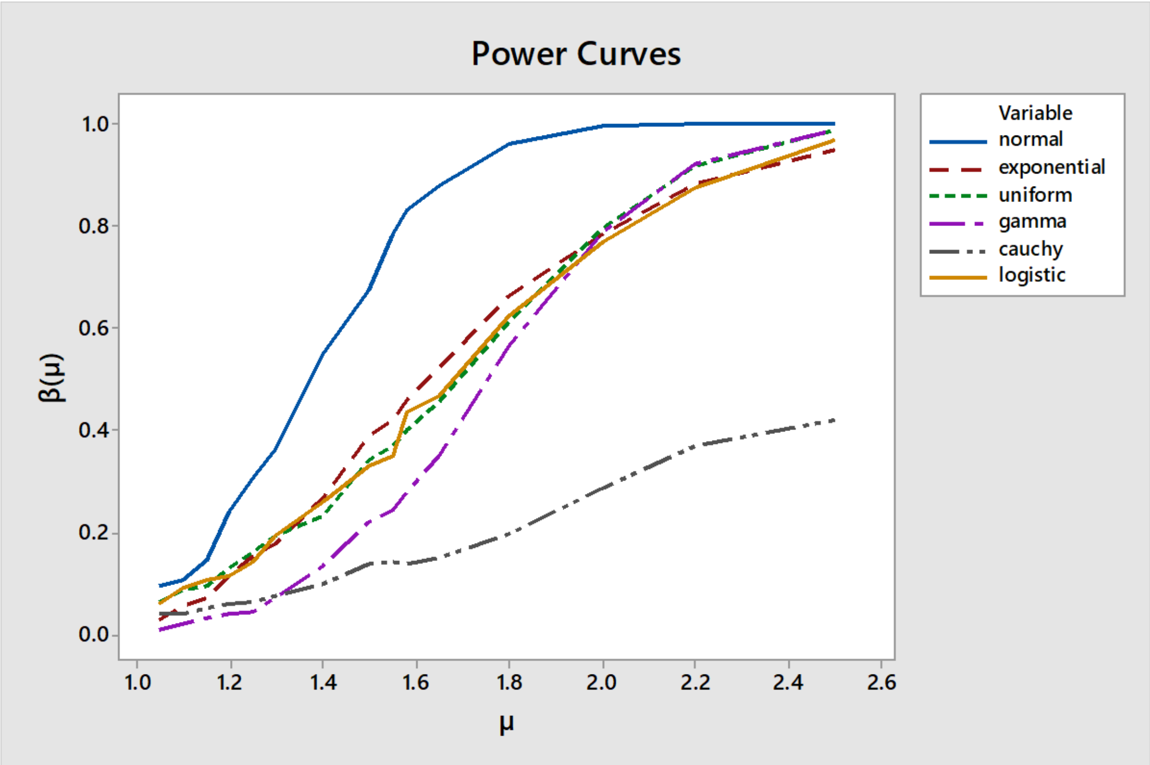

Figure 6.1

From the above plot we observe that, applying t-test the power curve of normal distribution is much higher than any other distributions. Thus, we can conclude that t-test is robust under normality assumption.

7. Non-parametric test :

In the previous context, we were concerned with the univariate population distribution from which an independent random sample was drawn. But, if we are provided with a data, the normality assumption may not be valid always. Thus, the problems are more suitable to be viewed under non-parametric setup. Because, it does not take into account any assumption regarding the parameters of the population from which the sample is drawn rather the form of the population. Non-parametric techniques are convenient particularly when location parameters are of prime importance under non-normal population distributions. We shall study Sign test which is analogous to Student t-test for testing $H_0 : \mu=\mu_0$ against H1 **:** $\mu>\mu_0$.

8. Sign Test : The General Idea :

Before coming to the analysis, let us discuss the idea of Sign test.

Suppose, $X_1,X_2,\dots,X_n$ be a random sample from a distribution with cdf F which is continuous.

Suppose, $\theta_p=\theta_p(F)$ be the p^{th} quantile of the distribution i.e.,$P_F(X\leq \theta_p) = p$

Now, to test, $H_0 : \theta_p = \theta_p^0$ against $H_1 : \theta_p > \theta_p^0$

Let us define, $z_i = \begin{cases} 1, \text{if } X_i > \theta_p^0 \\ 0, \text{otherwise} \end{cases}$, $\forall$ i=1(1)n

Now, $z_i \sim$ Bernoulli(π) ,where $\pi = P_F(X_i> \theta_p^0)$

Note that, z_i's are independent and $k=\sum_{i=0}^{n} z_i \sim$ Binomial(n, π)

Under H_0, k ~ Binomial(n,1-p)

The test statistic k is distribution free under H_0. Note that, k is actually the plus signs among the n difference $(X_i- \theta_p^0)$, $\forall$ i=1(1)n.

Rejection rule: we reject H_0 against H_1 at level α if $k\geq k_\alpha$,where k_α is chosen to be the smallest integer which satisfies $P(k\geq k_\alpha \mid H_0) \leq\alpha$.

9. Simulation based study using Sign test

We run the simulation study again to calculate the empirical level and power applying the procedure of Sign test and using the same data generated under parametric setup.

Procedure:

In the context of Sign test, we need to consider the hypotheses in terms of median. As we consider the same data generated before, the mean under null and alternative hypothesis remaining same , the median will change from one distribution to other.

[For example, if X ~ Exponential(λ) , the probability density function will be –

$$f(x) = \lambda e^{-\lambda x}, x > 0, \ \lambda > 0$$

Now, under t-test we consider the hypotheses as -

$H_0 : \frac{1}{\lambda} = 1$ against $H_1 : \frac{1}{\lambda} = 2$; $\frac{1}{\lambda}$ being the mean of Exponential(λ)

Median of Exponential(λ) is $\frac{ln\,2}{\lambda}$. Therefore, for Sign test the hypothesis will be-

$H_0 : \theta_{0.5} = 0.6931$ against $H_1 : \theta_{0.5} = 1.38$; $\theta_{0.5}$ denoting the median]

Findings :

Distribution	Empirical Level	Empirical Power
Normal	0.013	0.908
Exponential	0.012	0.440
Uniform	0.019	0.286
Gamma	0.021	0.978
Cauchy	0.016	0.623
Logistic	0.019	0.535

Table 9.1

Interpretation:

Observe that, unlike the parametric setup, the data following normal distribution possesses lower power under non-parametric setup and also empirical level is coming out as $0.013 << 0.05$. That is, we can conclude for normally distributed data t-test is appropriate.

Coming to the other distributions, if we compare the empirical power of Sign test (Table 9.1) with that of t-test (Table 6.1), the data generated from Gamma and Cauchy distribution shows a substantial change, which is so obvious. But, in case of Exponential, Uniform and Logistic distribution the empirical power decreases.

Let us look at the power curves of each population distribution to infer whether we should use t-test or non-parametric test for a given sample data.

Power Curves:

i) Normal Distribution :

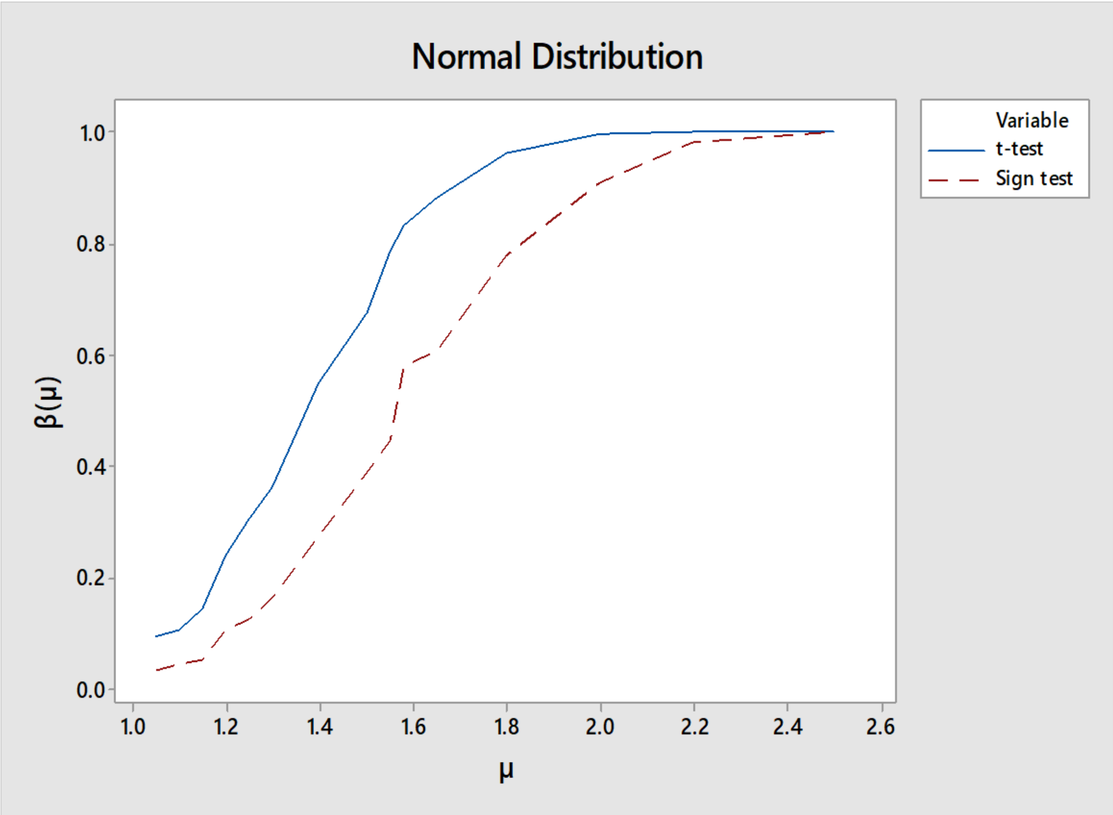

Figure 9.1

Observe that, the power curve we have got using non-parametric test is below the power curve of Student's t-test which is expected. Also, the empirical level under Sign test is 0.013 which is not at all satisfied. As we know normality is the principal assumption of one sample t-test, hence we should go for the same for normally distributed data.

ii) Exponential Distribution :

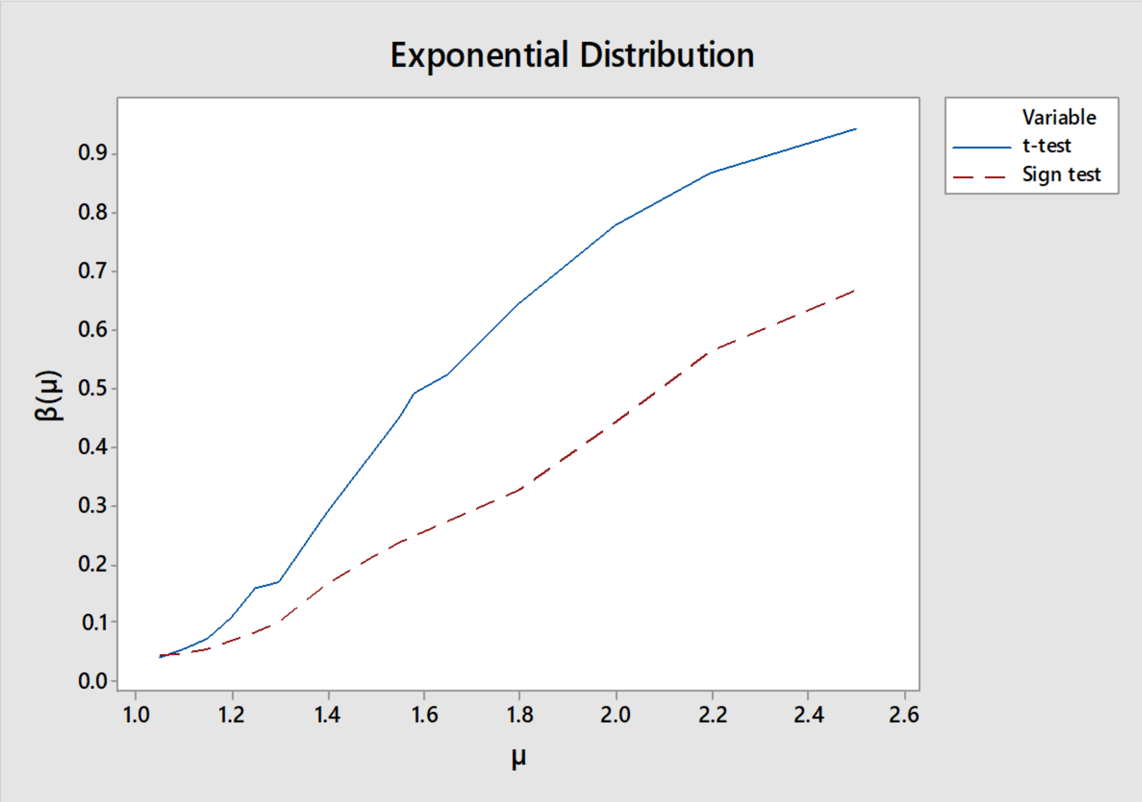

Figure 9.2

For Exponential distribution, the power curve using t-test is above the power curve of Sign test. That is t-test gives better result for a data following Exponential distribution.

Additionally, from Table 6.1 and Table 9.1, the empirical level under Sign test being 0.012 is less than that of t-test which is coming out as 0.022. Though both the values are less than 0.05 as pre-specified, still we should prefer t-test for exponential distribution to test for mean.

iii) Uniform Distribution :

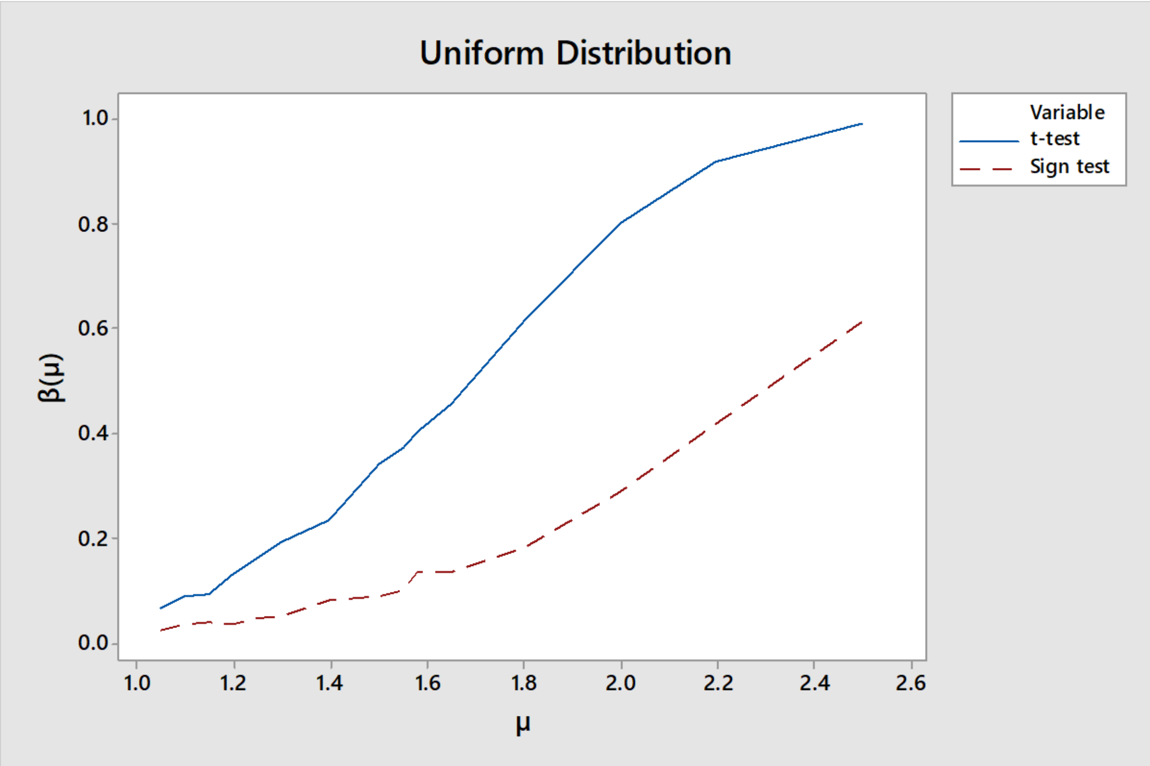

Figure 9.3

Observe that, under Uniform distribution the power curve of t-test goes above the power curve of non-parametric test. That is, though it violates the assumption of Student's t-test, it gives better result in this context.

Again, from Table 6.1 and Table 9.1, the empirical level under Sign test being 0.019 is much less than that of t-test which is 0.056. Also, under t-test empirical level is satisfactory. So, we should go for t-test for uniformly distributed data.

iv) Gamma Distribution :

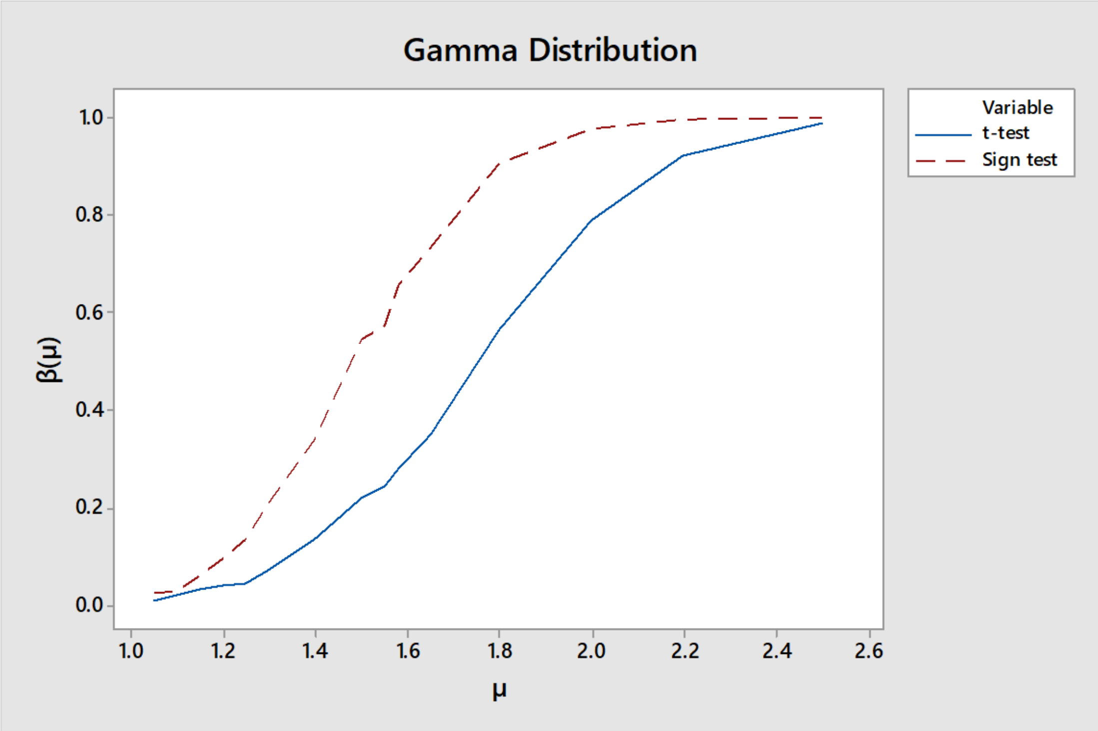

Figure 9.4

For Gamma distribution, the power curve using Sign test is above than that of t-test. From the graph it is clearly observed that Sign test yield comparatively better result for a random sample following Gamma distribution.

Moreover, from Table 6.1 and Table 9.1, the empirical level under Sign test being 0.021 is less than 0.05 as pre-specified. But, there is a substantial increment in empirical level with respect to t-test. Overall, Sign test is preferred in case of Gamma distribution.

v) <u>Cauchy Distribution</u> :

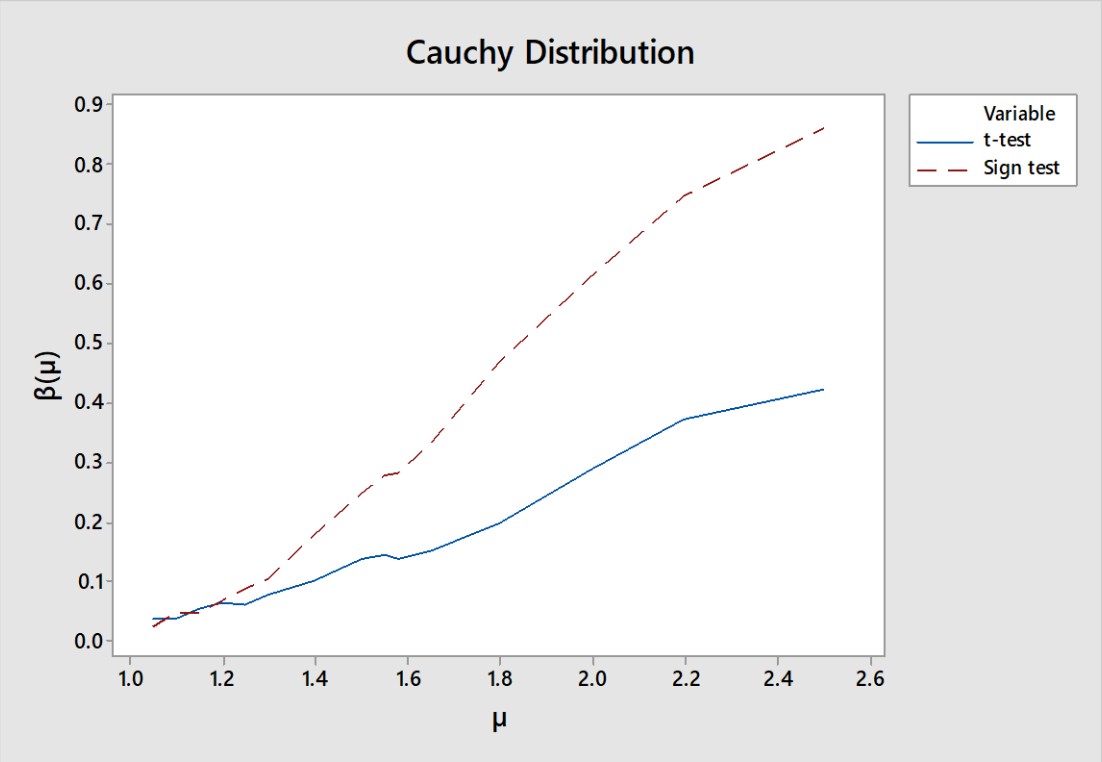

Figure 9.5

Similarly, under Cauchy distribution, the power curve of t-test is much below than that of Sign test. But, the empirical level under t-test is 0.031, whereas it is 0.016 under Sign test. Both the values are not around 0.05 as pre-determined.

We test for median in case of Sign test and Cauchy distribution considers median as its parameter. Therefore, theoretically Sign test should give better result. Also, observing the power curves, it is clear that we should go for non-parametric test for a data following Cauchy distribution.

vi) Logistic Distribution :

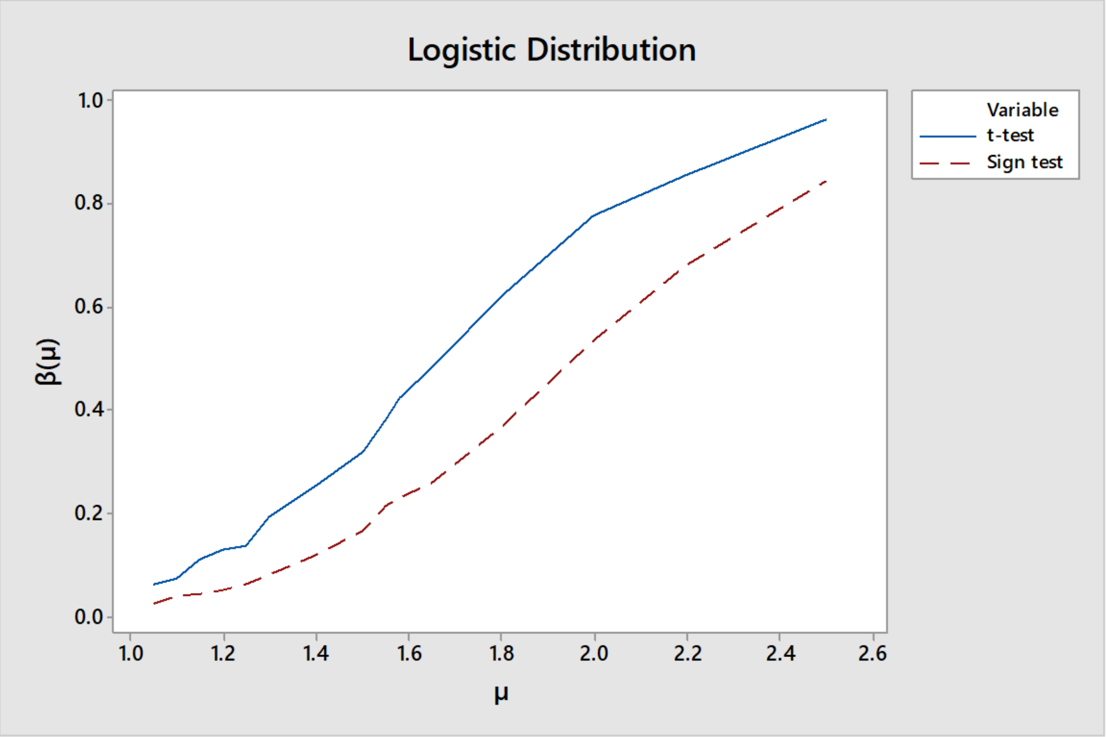

Figure 9.6

Under Logistic distribution the power curve of t-test goes above the power curve of non-parametric test. Again, from Table 6.1 and Table 9.1, the empirical level under Sign test being 0.019 is much less than that of t-test which is 0.047. Also, empirical level of t-test is satisfactory. So, we should go for t-test for a data following Logistic distribution.

10. Conclusion :

Comparing all the outcomes we have got throughout the study, we come to a satisfactory result. If we summarize our findings, we have the followings :

1) t-test is giving adequate results for Normal distribution, Exponential distribution, Uniform distribution as well as Logistic distribution.
2) Sign test is providing the same for Gamma distribution and Cauchy distribution.

Thus, it shows that though we rely upon the assumption of normality for applying t-test, it may not always true. Sometimes, t-test is also suitable or reliable for testing the hypotheses about mean of a single sample following some non-normal distributions. In that case, the parametric assumptions are violated.

Therefore, we can conclude that Student t-test is robust to the deviation from normality and moderate violation from the assumptions.

11. Reference

- *Jean Dickson Gibbons, Nonparametric Statistical Inference*
- *A.M. Gun, M.K. Gupta, B. Dasgupta, An Outline of Statistical Theory*
- *https://www.wikipedia.org*
- *https://www.scribbr.com/statistics/type-i-and-type-ii-errors/*

12. Acknowledgement

I would like to acknowledge my sincerest thanks to my supervisor, Dr. Surupa Chakraborty (Roy) for her constant guidance throughout the duration of completion of this project on “Robustness of t-test”. I would also like to express my special gratitude to authorities of St. Xavier’s College (Autonomous), Kolkata for giving me a golden opportunity to explore new and exciting field of research and implement them in this paper. Such an opportunity to prepare a dissertation paper helped me to polish my research mindset.

Shreya Saha

St. Xavier’s College(Autonomous), Kolkata