# Statistical Inference of Medical Data : COVID-19

Bhoj Rani Soopal, Mrinalini LNU

April 28, 2020

**Abstract**

Finding an effective Medicine for Corona Virus(COVID-19) is the most important task for scientists and researchers right now because of the severity of this pandemic.The pandemic is affecting businesses and interrupting lives, including ours. Since there is no news of an actual medicine in sight, therefore we have simulated a random sample of data for our testing purposes. This translates to the following *Hypothesis Testing* problem: Does the medicine reduce number of recovery days for a patient i.e., is the medicine effective against the COVID-19 virus? In this project, we simulate and showcase the hypothesis testing for our problem statement

## 1 Introduction

We assume that without a medicine, patients recover in a mean of 30 days.With our COVID-19 medicine, patients recover in a mean of 27 days. Our purpose is to define a Hypothesis Test to validate from a sample size of 100 patients, if our medicine is effective against COVID-19. We perform T-test and Z-test for our Hypothesis to reach a conclusion about our data.

**Definition 1.1.** Hypothesis Test[2] The purpose of statistical inference is to draw conclusions about a population on the basis of data obtained from a sample of that population. Hypothesis testing is the process used to evaluate the strength of evidence from the sample and provides a framework for making determinations related to the population, i.e., it provides a method for understanding how reliably one can extrapolate observed findings in a sample under study to the larger population from which the sample was drawn. The investigator formulates a specific hypothesis, evaluates data from the sample, and uses these data to decide whether they support the specific hypothesis

In statistics "population" refers to the total set of observations that can be made. For e.g., if we want to calculate average height of humans present on the earth, "population" will be the "total number of people actually present on the earth". A sample, on the other hand, is a set of data collected/selected from a pre-defined procedure. For our example above, it will be a small group of people selected randomly from some parts of the earth.
When "population" is infinitely large it is improbable to validate any hypothesis by calculating

the mean value or test parameters on the entire population. In such cases, a population is assumed to be of some type of a distribution. The most common forms of distributions are Normal,Binomial, Poisson and Discrete.
A critical value is a point (or points) on the scale of the test statistic beyond which we reject the null hypothesis, and, is derived from the level of significance of the test. Critical value can tell us, what is the probability of two sample means belonging to the same distribution. Higher, the critical value means lower the probability of two samples belonging to same distribution. P-value on the other hand is defined as the probability to the right of respective statistic (Z, T or chi).

Further we demonstrate Power of a Test[1] and Confidence Interval[4].

# 2 The model

The first step in testing hypotheses is the transformation of the research question into a null hypothesis, , which is referred to as the level of significance.
Just as hypothesis testing can reject a true null hypothesis (referred to as a type I error), it can fail to reject H0 when the predictor and outcome are associated (type II error). The probability of such a false-negative conclusion is called . The quantity (1) is called the power of the test and is simply the probability of drawing the correct conclusion (i.e., rejecting H0) when an association between predictor and outcome actually does exist.

# 3 Main: our results

## 3.1 Setup

Null Hypothesis : The medicine has no effect on recovery time and is ineffective
Alternate Hypothesis : The medicine reduces recovery time and is effective

-Your data should be normally distributed, iid.
-When we know standard deviation of Population] 1. Z-Test
When do we use a Z-test?

-Your data should be normally distributed, iid.
-When we know standard deviation of Population

Steps for a Z-test:

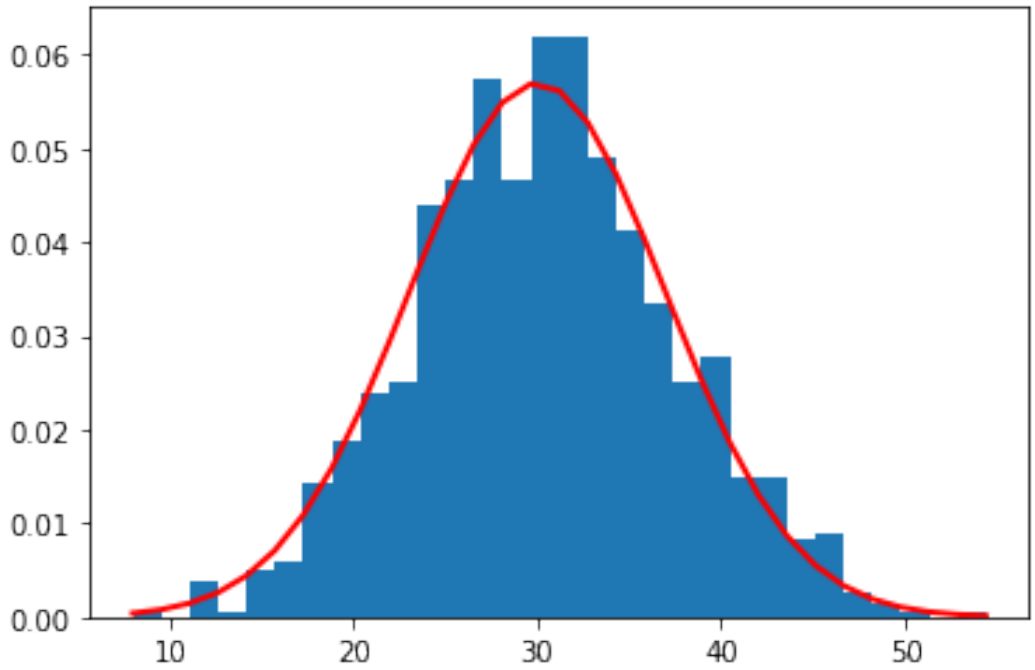
Question Statement:

We claim that the medicine reduces the recovery time. From a sample of 100 patients we found average recovery time to be 27 days.Is there sufficient evidence to support our claim? Population mean recovery days is 30 with standard deviation of 7.

$$PopulationMean : \mu_0 = 30$$

$$PopulationStandardDeviation : \sigma = 7$$

$$PopulationSize : N = 1000$$




Population Distribution $\sim N(30, 7^2)$

$$SampleMean : \overline{X} = 27$$

$$SampleStandardDeviation : SE = \sigma/(\sqrt{n})$$

$$SampleSize : n = 100$$

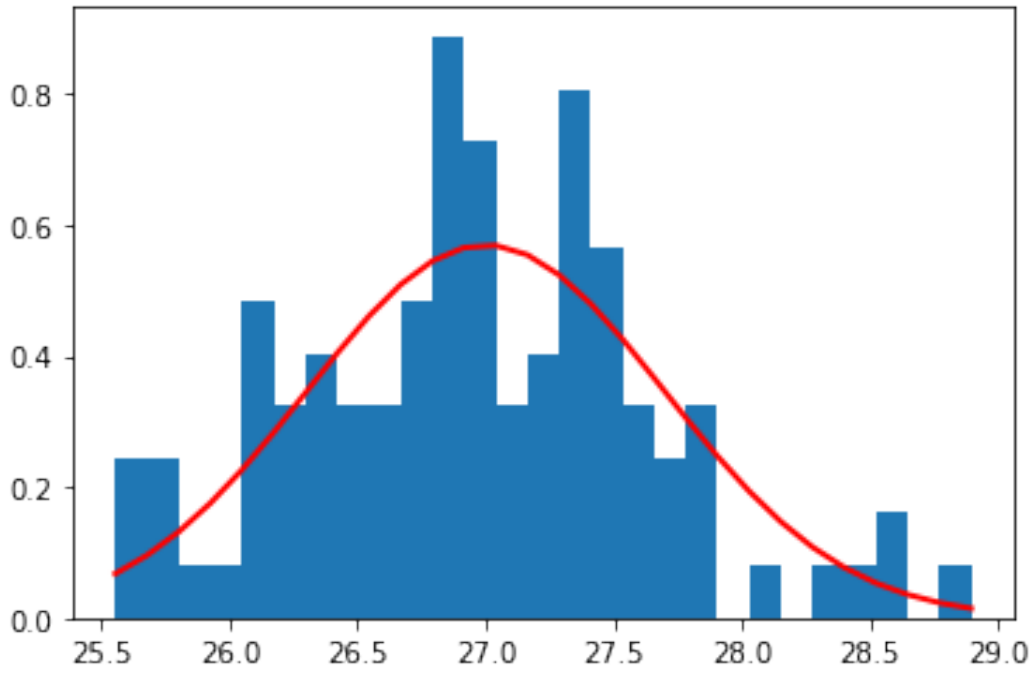



Figure 2: Sample Distribution $\sim N(27, StandardError^2)$

1.State the null hypothesis and alternate hypothesis.

$$H_0 := \mu = \mu_0$$

$$H_A := \mu < \mu_0$$

2.Choose an alpha level.

$$\alpha = 0.05$$

3.Find the critical value of z in a z table.

$$zcritical_\alpha = -1.645$$

4.Calculate the Z -test statistic.

$$Z = \frac{\overline{X} - \mu_o}{\sigma/\sqrt{n}}$$

$$\Rightarrow Z = -4.29$$

5.Compare the test statistic to the critical z value and decide if you should support or reject the null hypothesis.

**Conclusion: As we can see:**

$$Z < zcritical_{\alpha}$$

Therefore we Reject Null Hypothesis: $H_0$

**We can conclude that the medicine reduces number of days of recovery, hence it is effective.**

**2. T-Test**
**When do we use a T-test?**
**-Data has Normal Distribution**
**-A t-test is used when the population parameter(standard deviation) is not known.**
**-Or when we have a small sample size.**

Steps for a T-test:

We claim that the medicine reduces the recovery time. From a sample of 100 patients we found average recovery time to be 27 days with standard deviation of 3.Is there sufficient evidence to support our claim? Population mean recovery days is 30.]Question Statement:

We claim that the medicine reduces the recovery time. From a sample of 100 patients we found average recovery time to be 27 days with standard deviation of 3.Is there sufficient evidence to support our claim? Population mean recovery days is 30.

$$PopulationMean : \mu_0 = 30$$

$$SampleMean : \overline{X} = 27$$

$$SampleStandardDeviation : S = 3$$

$$SampleSize : n = 100$$

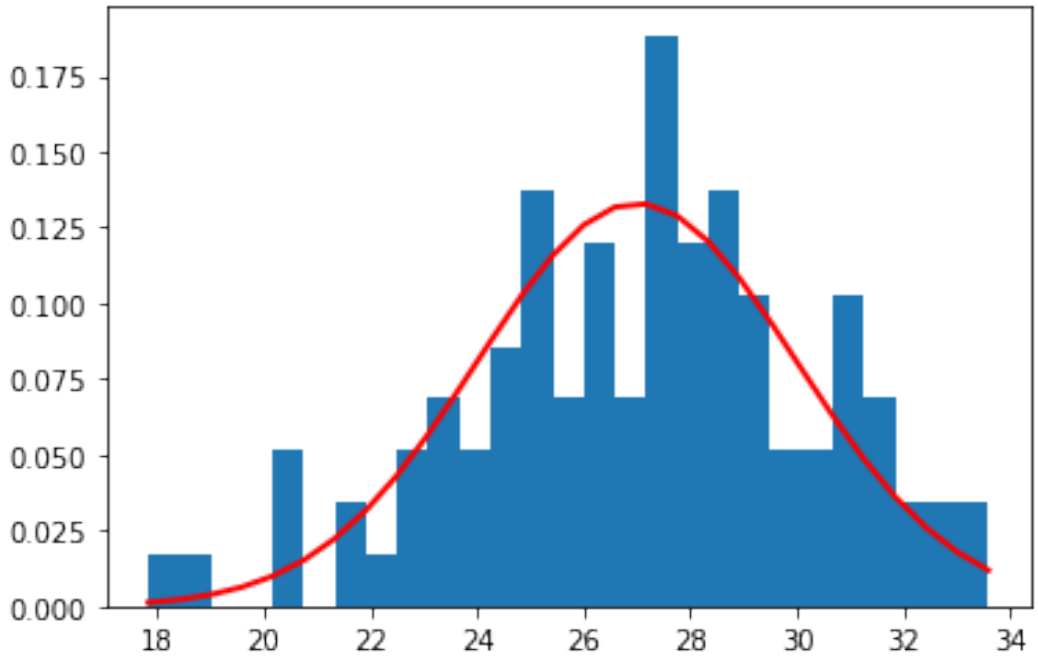



Figure 3: Sample Distribution $\sim N(27, 3^2)$

1.State the null hypothesis and alternate hypothesis.

$$H_0 := \mu = \mu_0$$

$$H_A := \mu < \mu_0$$

2.Choose an alpha level.

$$\alpha = 0.05$$

3.Find Degree of Freedom.

$$df = n - 1 = 99$$

4.Find the critical value of t in a t table.

$$tcritical_\alpha = -1.66039$$

5.Calculate the T-test statistic.

$$T = \frac{\overline{X} - \mu_o}{S/\sqrt{n}}$$

$$\Rightarrow T = -10.0$$

6.Calculate p-value.the p-value or probability value is the probability of obtaining test results at least as extreme as the results actually observed during the test, assuming that the null hypothesis is correct.

$$p = Pr(T < t < -10.0)$$

$$\Rightarrow p = 5.469878503996061e - 17$$

7.Interpret results.

Statistical significance is determined by looking at the p-value. The p-value gives the probability of observing the test results under the null hypothesis. The lower the p-value, the lower the probability of obtaining a result like the one that was observed if the null hypothesis was true. Thus, a low p-value indicates decreased support for the null hypothesis. However, the possibility that the null hypothesis is true and that we simply obtained a very rare result can never be ruled out completely. The cut-off value for determining statistical significance is ultimately decided on by the researcher, but usually a value of .05 or less is chosen. This corresponds to a 5 percent (or less) chance of obtaining a result like the one that was observed if the null hypothesis was true.

**Conclusion: As we can see from our results that:**

$$p < \alpha$$

and

$$T < tcritcal$$

We can conclude that we have enough evidence against the Null Hypothesis. Therefore, We Reject the Null Hypothesis

## 3.2 Power of a Test

**Definition: Statistical power of a hypothesis test is simply the probability that the given test correctly rejects the null hypothesis (which means the same as accepting the H1), when the alternative is in fact true. Higher statistical power of an experiment means lower probability of committing a Type II error. It also means higher probability of detecting an effect when there is an effect to detect (true positive). This can be illustrated by the following formula:**

$$Power = Pr(rejectH_0|H_Aistrue) = 1 - Pr(failtorejectH_0|H_0isfalse)$$

**In practice, results from experiments with too little power will lead to wrong conclusions, which in turn will affect the decision-making process. That is why only**

**results with an acceptable level of power should be taken into consideration. It is quite common to design experiments with power level of 80%, which translates to a 20% probability of committing a Type II error.**

**Power analysis:** Power analysis is built from the following building blocks:

- Significance level
- Effect size
- Power
- Sample Size

The idea of power analysis can be brought down to the following: by having three out of four metrics, we estimate the missing one. This comes in handy in two ways:

1.When we are designing an experiment, we can assume what level of significance, power and effect size is acceptable to us and — as a result — estimate how big a sample we need to gather for such an experiment to yield valid results.

2.When we are validating an experiment, we can see if, given the used sample size, effect size and significance level, the probability of committing a Type II error is acceptable from the business perspective.

**Student's T-Test Power Analysis**

1.Estimate sample size through power analysis: If we take power at 80%, significance level 5%, expected effect size 80% then we will have a required sample size of 25.

**Generating Power Curves** We choose [0.2,0.5,0.8] as the considered effect size values as they correspond to the thresholds for small, medium, large in the case of Cohen's d.

**Plots:**

- **Power Vs. Number of Observations**

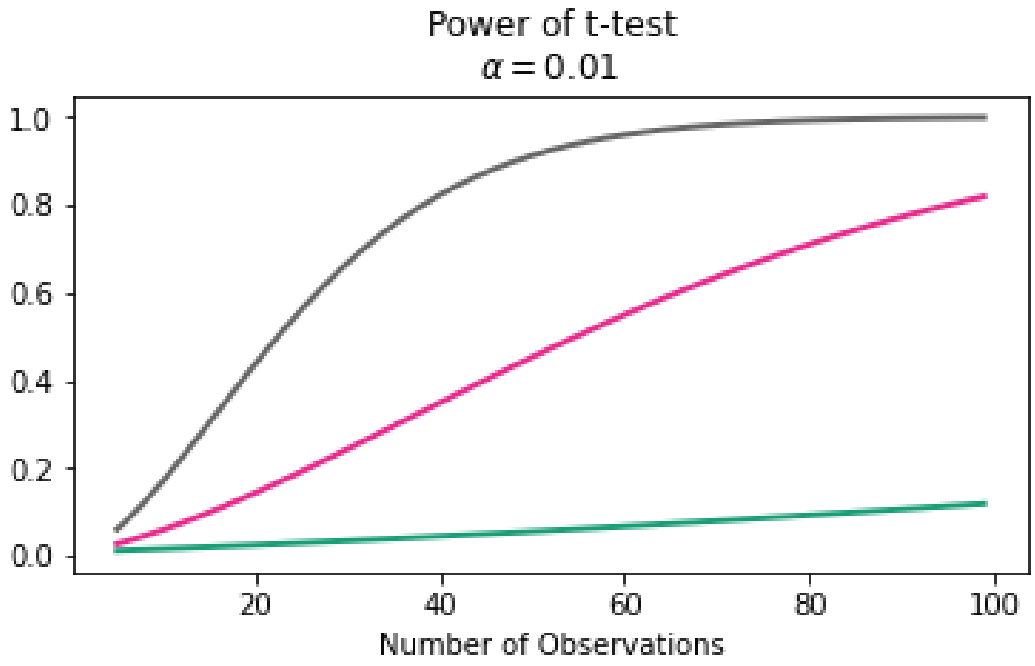



Figure 4: Power Vs. No. of Observation, $\alpha = 0.01$

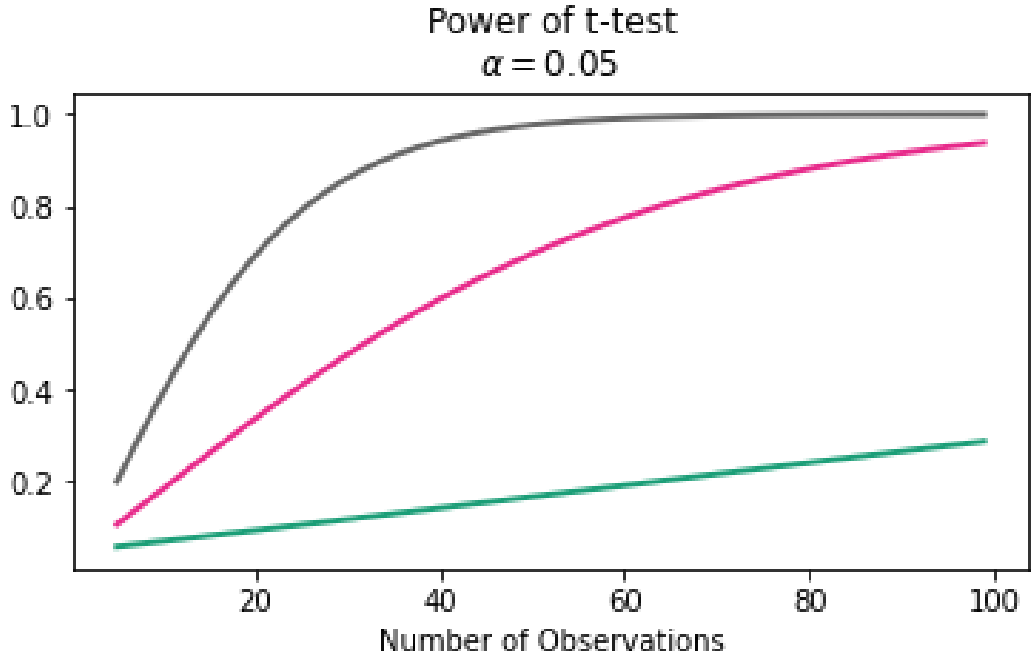



Figure 5: Power Vs. No. of Observation, $\alpha = 0.05$
We observe that an increase in sample size results in an increase in power. That is, the bigger the sample, the higher the power, keeping other parameters constant.

- **Power Vs. Effect**

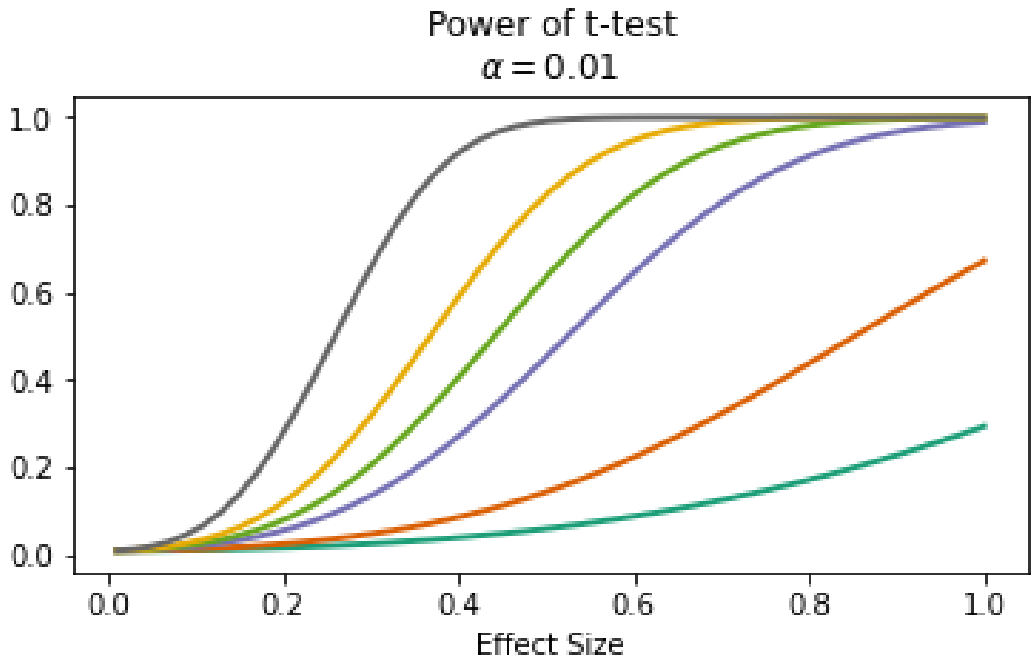



Figure 6: Power Vs. Effect, $\alpha = 0.01$

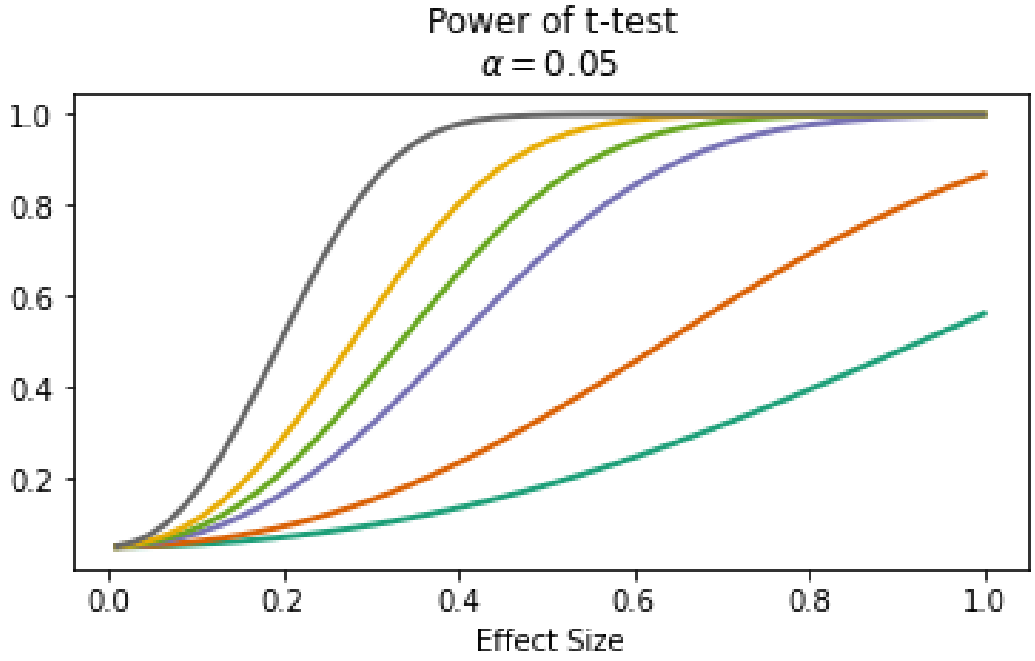



Figure 7: Power Vs. Effect, $\alpha = 0.05$

We observe that an increase in effect size results in an increase in power. That is, the bigger the sample, the higher the power, keeping other parameters constant.

- **Power Vs. Significance Level**

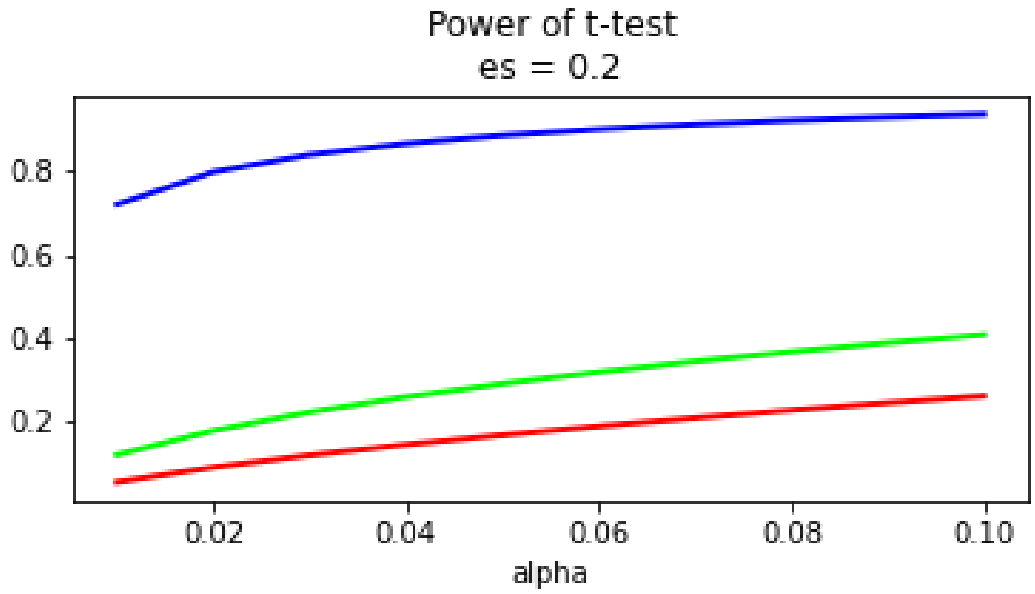



Figure 8: Power Vs. Significance Level, Effect Size = 0.02

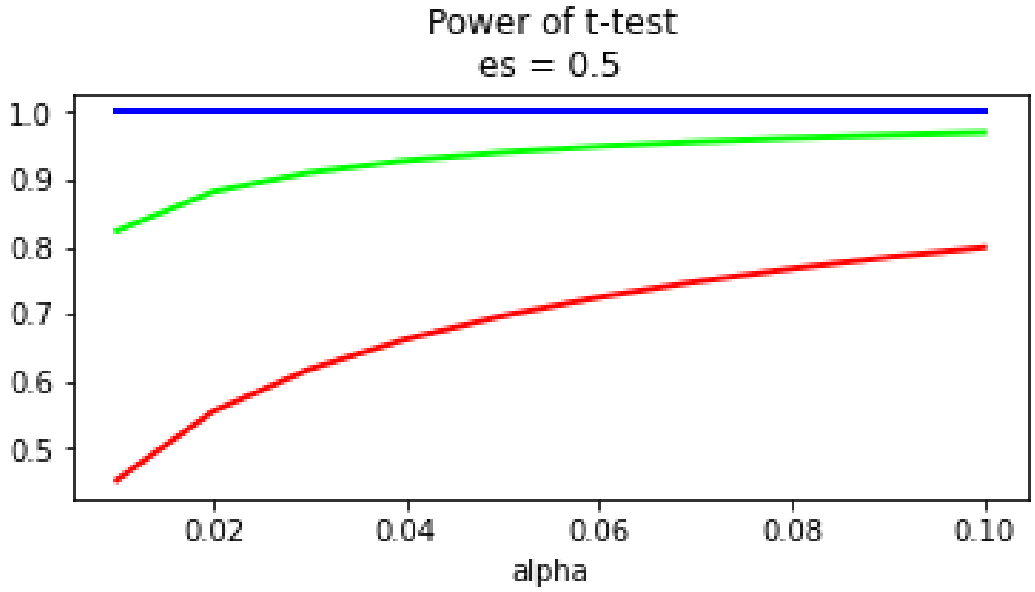



Figure 9:Power Vs. Significance Level, Effect Size = 0.05

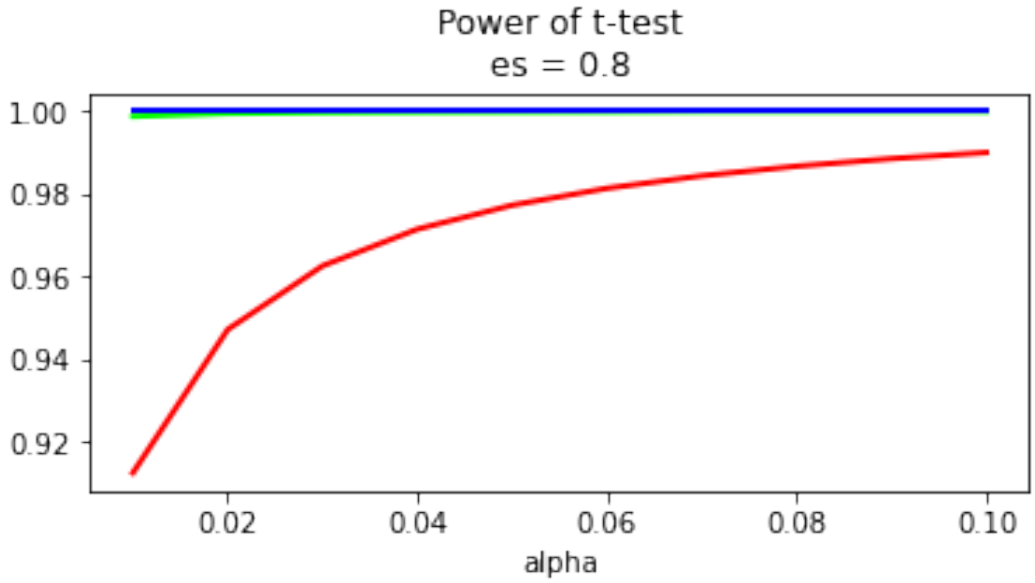



Figure 10: Power Vs. Significance Level, Effect Size = 0.08
We observe that an increase in Significance Level results in an increase in power. That is, the bigger the sample, the higher the power, keeping other parameters constant.

## 3.3 Confidence Interval

The confidence level represents the frequency (i.e. the proportion) of possible confidence intervals that contain the true value of the unknown population parameter. In other words, if confidence intervals are constructed using a given confidence level from an infinite number of independent sample statistics, the proportion of those intervals that contain the true value of the parameter will be equal to the confidence level.[3]
To estimate the range of population mean, we define the standard error of the mean:

$$StandardError(SE) = \sigma/\sqrt{n}$$

$$\sigma = SampleStandardDeviation$$

$$n = SampleSize$$

$$\mu is the sample mean$$

$$(\mu - 1.96 * SE, \mu + 1.96 * SE)$$

The above interval is called confidence interval.

We usually use 1.96 to calculate a 95% confidence interval since the sample mean follows a normal distribution.

**100 meds with 95% confidence interval :('26.86', '27.137200')**
**1000 meds with 95% confidence interval :('29.57', '30.433864')** From our results, we can note that a 95% confidence interval becomes narrower that is if we increase sample size from 100 to 1000.

Assuming that N is positive infinite, then we will have $\lim_{n\to\infty} \sigma/\sqrt(n) = 0$.
$In other words, the confidence interval will become the sample mean!$

]If we want to estimate an interval of the population so that 95% of the time the interval will contain the population mean, the interval is calculated as:

$$(\mu - 1.96 * SE, \mu + 1.96 * SE)$$

The above interval is called confidence interval.

We usually use 1.96 to calculate a 95% confidence interval since the sample mean follows a normal distribution.

**100 meds with 95% confidence interval :('26.86', '27.137200')**
**1000 meds with 95% confidence interval :('29.57', '30.433864')** From our results, we can note that a 95% confidence interval becomes narrower that is if we increase sample size from 100 to 1000.

Assuming that N is positive infinite, then we will have $\lim_{n\to\infty} \sigma/\sqrt(n) = 0$.
$In other words, the confidence interval will become the sample mean!$

# References

[1] Jason Brownlee. A gentle introduction to statistical power and power analysis in python. 2020.

[2] Roger B. Davis and Kenneth J. Mukamal. Hypothesis testing means. *Statistcal Primer for CardioVascular Research*, 114(10):1078–1082, 2006.

[3] F M Dekking;. *A modern introduction to probability and statistics (NL-LeOCL)317542265.* Springer, 2005.

[4] Pav Kalinowski. Understanding confidence intervals (cis) and effect size estimation. Observer Vol.23, No.4 April 2010.