STATISTICS FOR DATA SCIENCE

POWER OF TEST AND SIMPLE LINEAR REGRESSION

SIVASANKARI V
Department of Science & Humanities

STATISTICS FOR DATA SCIENCE

Unit 5 : Power of test and Simple linear regression

Session : 3

Sub Topic : Factors affecting Power of a test

SIVASANKARI V

Department of Science & Humanities

Power of a test - Recall

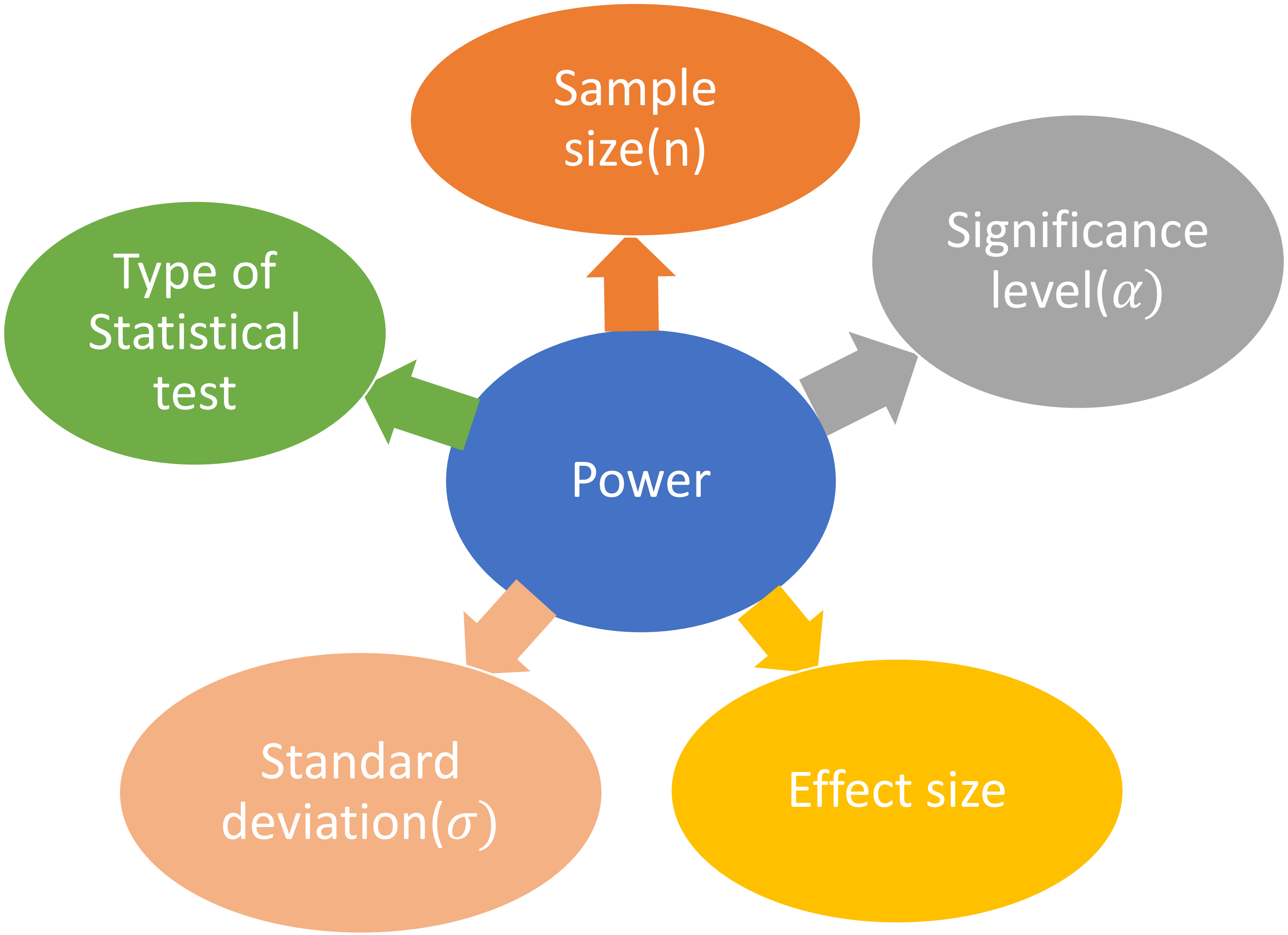
Power of a Hypothesis Test :

The power of a test is the probability of rejecting H_0 when it is false.

Power = 1 - P(type II error)
= 1 – β.

Note: *Statistical power has relevance only when the null is false.*

Factors affecting Statistical Power of test



Factors affecting Statistical Power of test – Sample size

Example:

A random sample of n people's wight whose mean and standard deviation are 168 lbs and 7.2 lbs. Can we conclude that the mean of the population is 165lb?

$$H_0: \mu = 165$$
$$H_1: \mu \neq 165$$

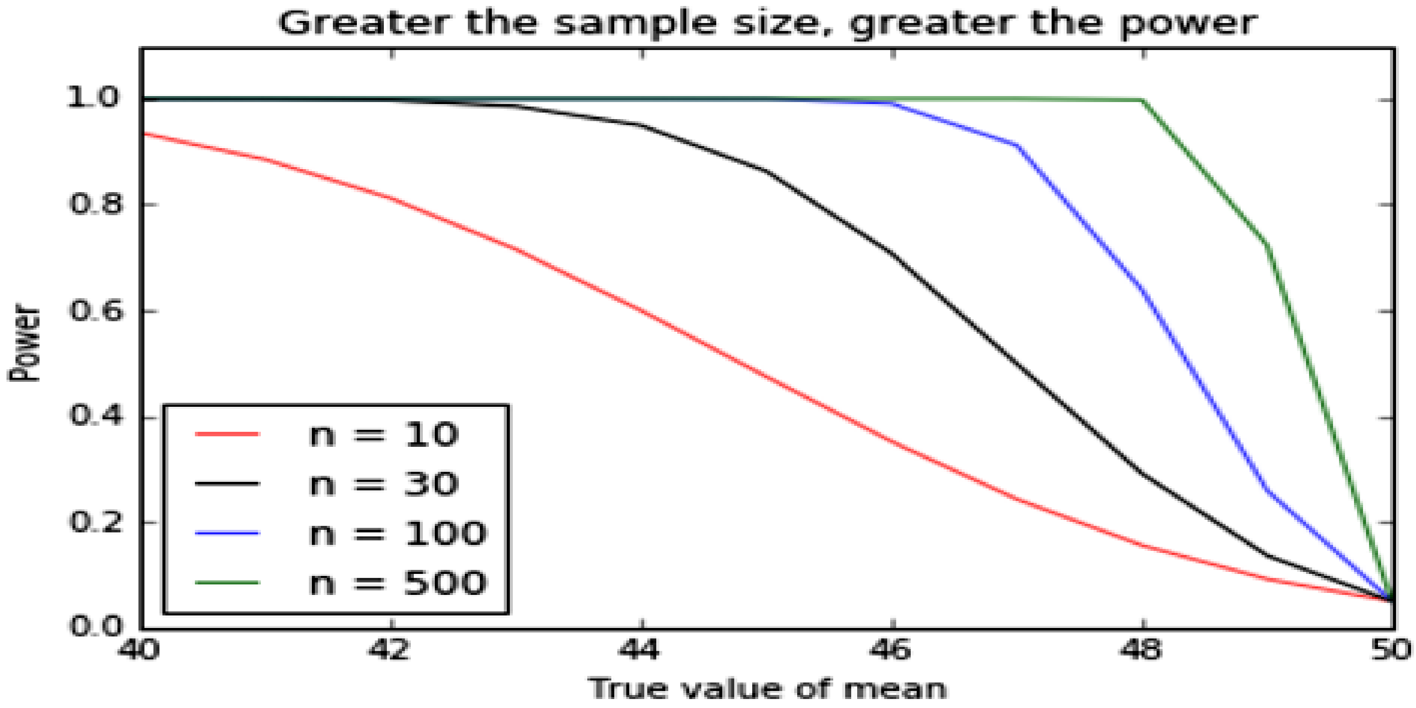
$$z = \frac{168 - 165}{7.2/\sqrt{n}} = \frac{(168 - 165)\sqrt{n}}{7.2}$$

Factors affecting Statistical Power of test- Sample size

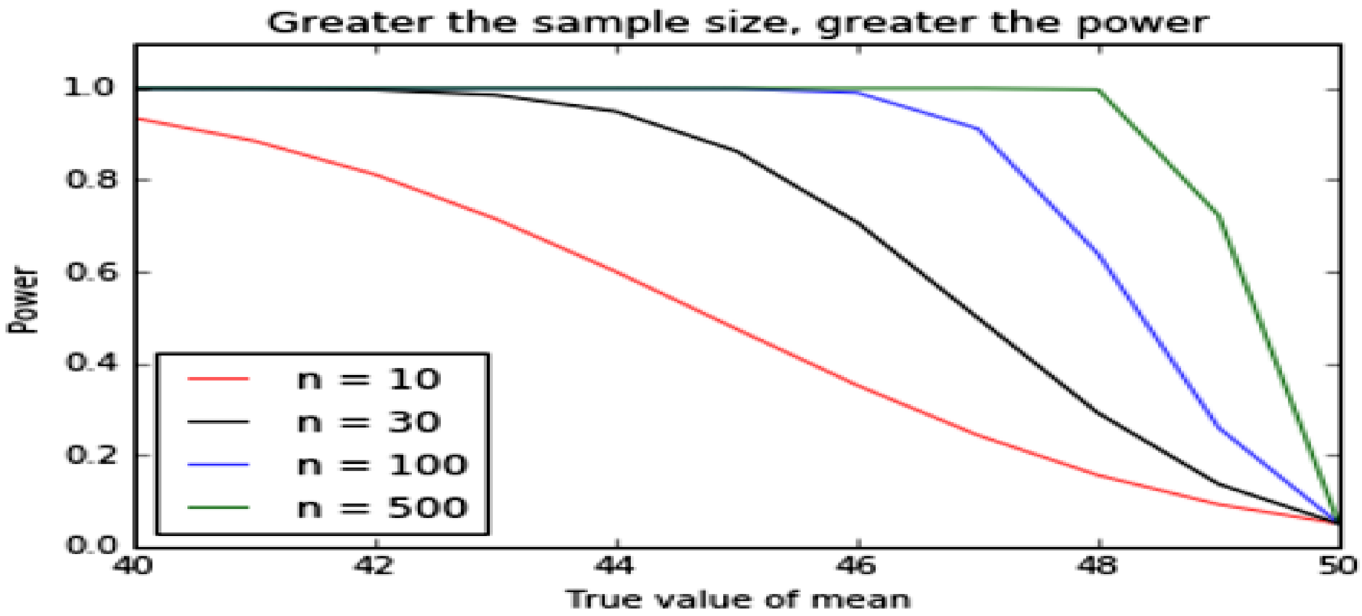
Power

Sample size(n)

The larger the sample size, the higher the power.



Factors affecting Statistical Power of test – Sample size



The above figure shows that the larger the sample size, the higher the power. Since sample size is typically under an experimenter's control, increasing sample size is one way to increase power. However, it is sometimes difficult and/or expensive to use a large sample size.

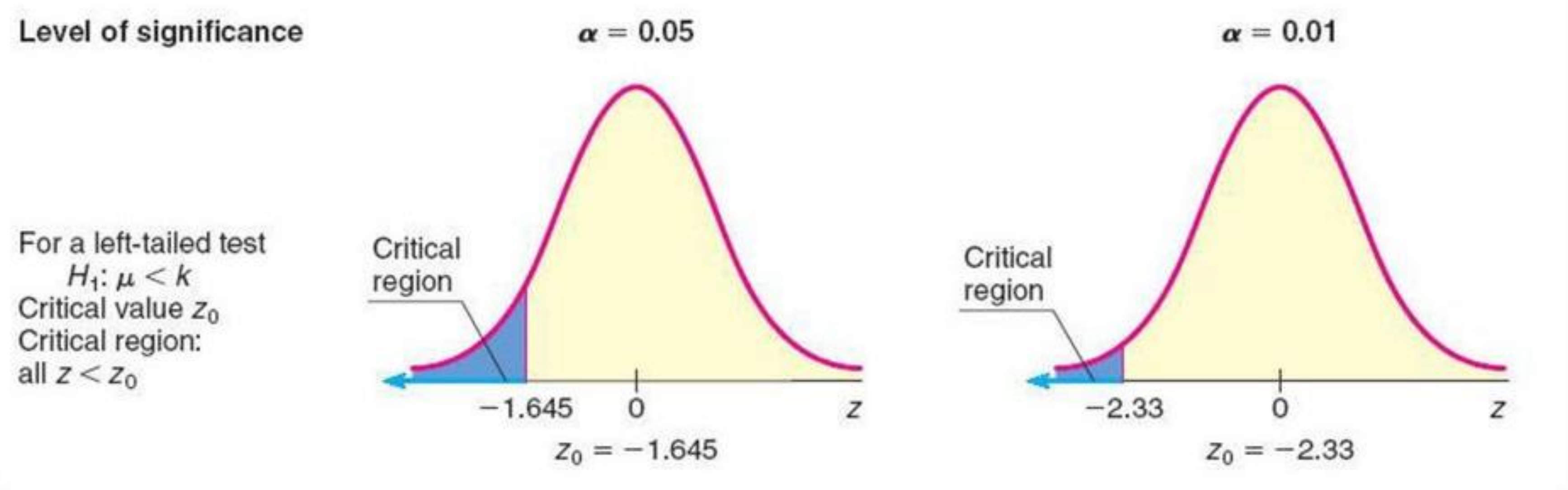
Factors affecting Statistical Power of test – Sample size

Power and Sample Size (N)

- Power increases as N increases.
- The more independent scores that are measured or collected, the more likely it is that the sample mean represents the true mean.
- Prior to a study, researchers rearrange the power calculation to determine how many scores (subjects or N) are needed to achieve a certain level of power (usually 80%).

Factors affecting Statistical Power of test – Significance level

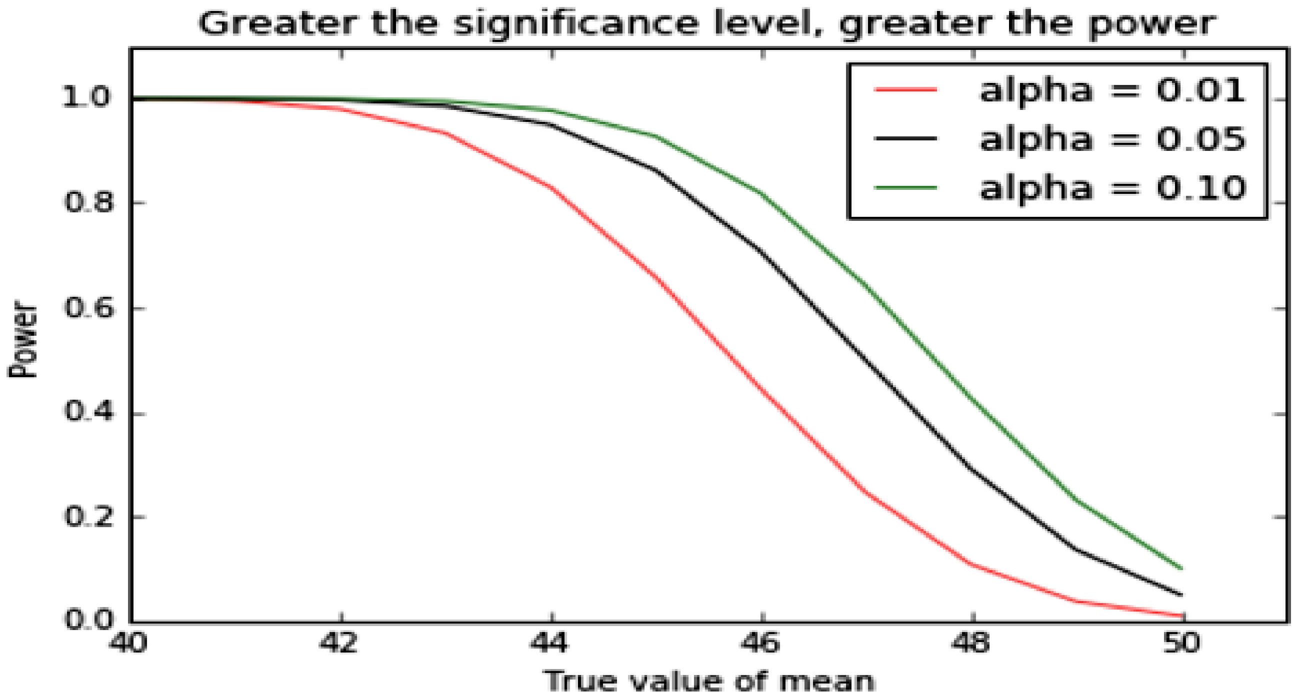
Critical Values z_0 for $\alpha = 0.05$ and $\alpha = 0.01$: Left-tailed Test



Source:Houghton Miffin Company

Factors affecting Statistical Power of test- Significance level

Power → Significance level(α) → **The larger the significance level, the higher the power.**



Factors affecting Statistical Power of test – Significance level

Power and Alpha (α)

- An increase in alpha, say from .05 to .1, artificially increases the power of a study.
- Increasing alpha reduces the risk of making a type II error, but increases that of a type I.
- Increasing the risk of making a type I error, in many cases, may be worse than making a type II error.
- Example- replacing an effective chemotherapy drug with one that is, in reality, less effective.

Factors affecting Statistical Power of test – Significance level

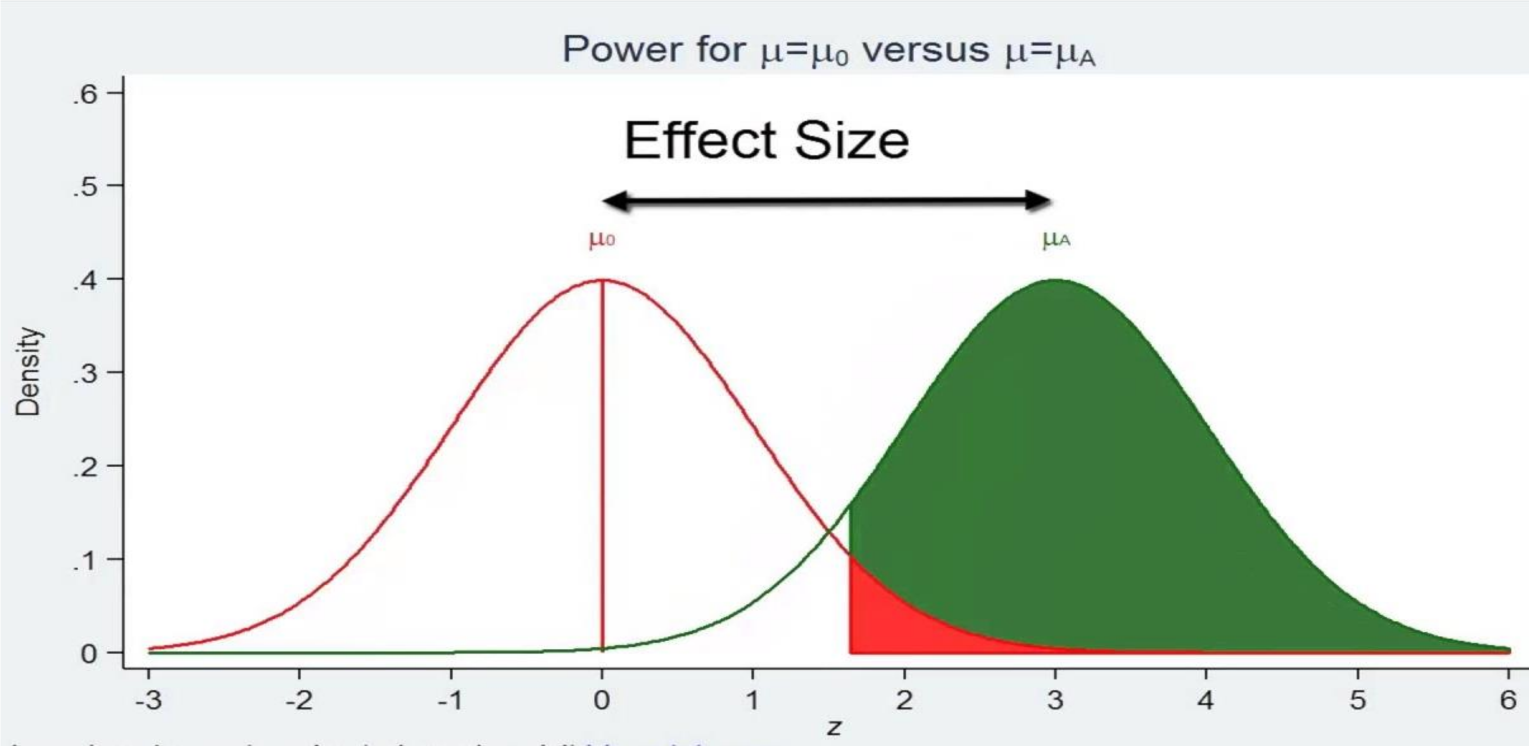
Power and Significance level (α)

The lower the significance level, the lower the power of the test. If you reduce the significance level (e.g., from 0.05 to 0.01), the region of acceptance gets bigger. As a result, you are less likely to reject the null hypothesis. This means you are less likely to reject the null hypothesis when it is false, so you are more likely to make a Type II error. In short, the power of the test is reduced when you reduce the significance level; and vice versa.

Factors affecting Statistical Power of test- Effect Size

Effect size = True Mean - Hypothesized Mean

$$= \mu_A - \mu_0$$



Example of a power calculation

Assume that a new chemical process has been developed that may increase the yield over that of the current process. The current process is known to have a mean yield of 80 and a standard deviation of 5, where the units are the percentage of a theoretical maximum. If the mean yield of the new process is shown to be greater than 80, the new process will be put into production.

Let μ denote the mean yield of the new process. It is proposed to run the new process 50 times and then to test the hypothesis

$H_0 : \mu \leq 80$ versus $H_1 : \mu > 80$ at a significance level of 5%.

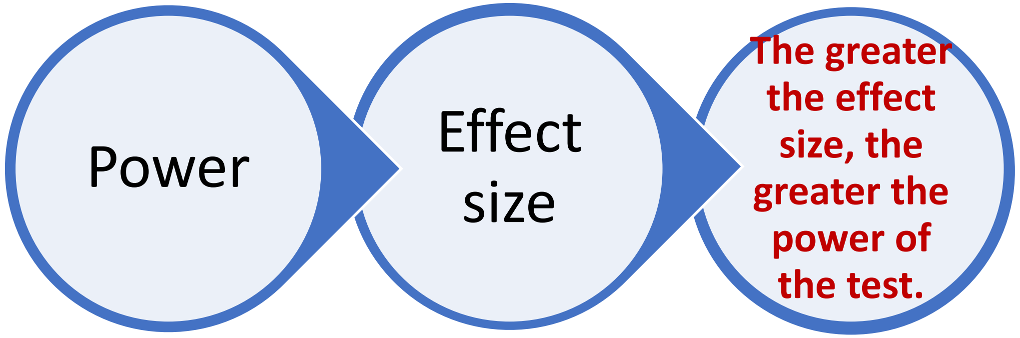
if μ is close to μ_0 : the power will be small

(when μ =81, Power=0.4090)

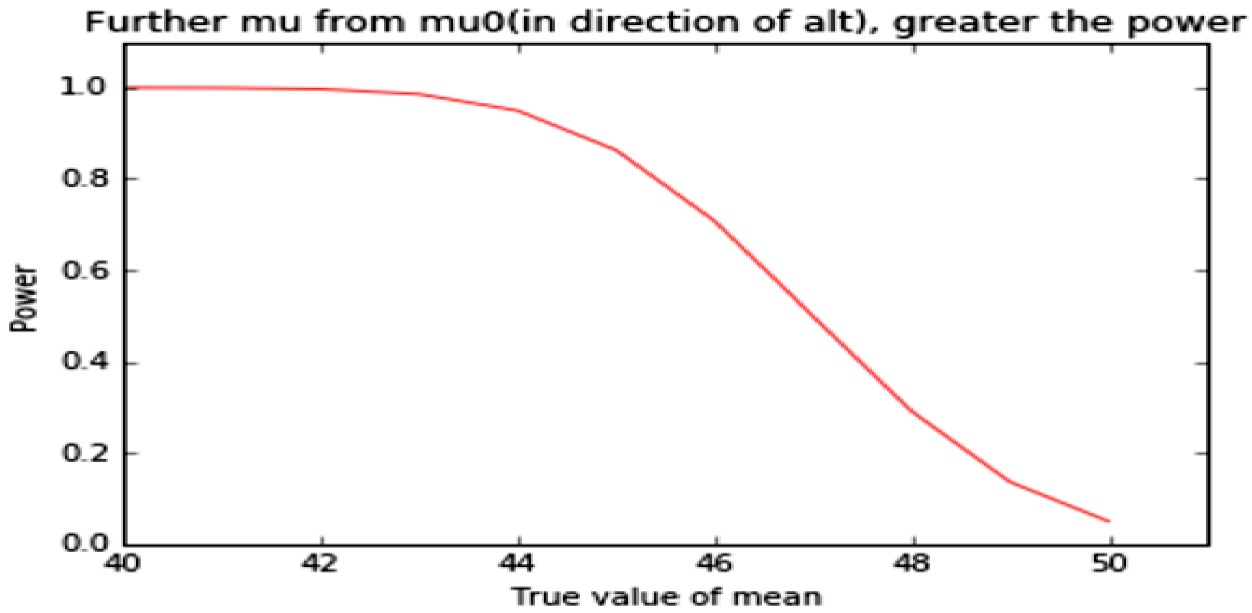
if μ is far from μ_0 : the power will be large

(when μ =82, Power=0.8830)

Factors affecting Statistical Power of test- Effect Size



Effect size = True value - Hypothesized value



Factors affecting Statistical Power of test – Effect size

- The **effect size** is the difference between the true value and the value specified in the null hypothesis.
- Effect size = True value - Hypothesized value
- For example, suppose the null hypothesis states that a population mean is equal to 100. A researcher might ask: What is the probability of rejecting the null hypothesis if the true population mean is equal to 90? In this example, the effect size would be 90 - 100, which equals -10.
- The "true" value of the parameter being tested. The greater the difference between the "true" value of a parameter and the value specified in the null hypothesis, the greater the power of the test. That is, the greater the effect size, the greater the power of the test

Factors affecting Statistical Power of test – Effect size

Power and Effect Size

- Effect size is a measure of the difference between the means of two groups of data.
- For example, the difference in mean jump height between samples of volleyball and basketball players.
- As effect size increases, so does power.
- For example, if the difference in mean jump height was very large, then it would be very likely that a Z or t-test on the two samples would detect that true difference.

A Little More on Effect Size

- While a p-value indicates the statistical significance of a test, the effect size indicates the “practical” significance.
- If the units of measurement are meaningful (e.g., jump height in cm), then the effect size can simply be portrayed as the difference between two means.
- If the units of measurement are not meaningful (questionnaire on behaviour), then a standardized method of calculating effect size is useful.

Factors affecting Statistical Power of test –Standard deviation

Example:

A random sample of 200 people's wight whose mean is 168 lbs. Can we conclude that the mean of the population is 165lb?

$$H_0: \mu = 165$$
$$H_1: \mu \neq 165$$

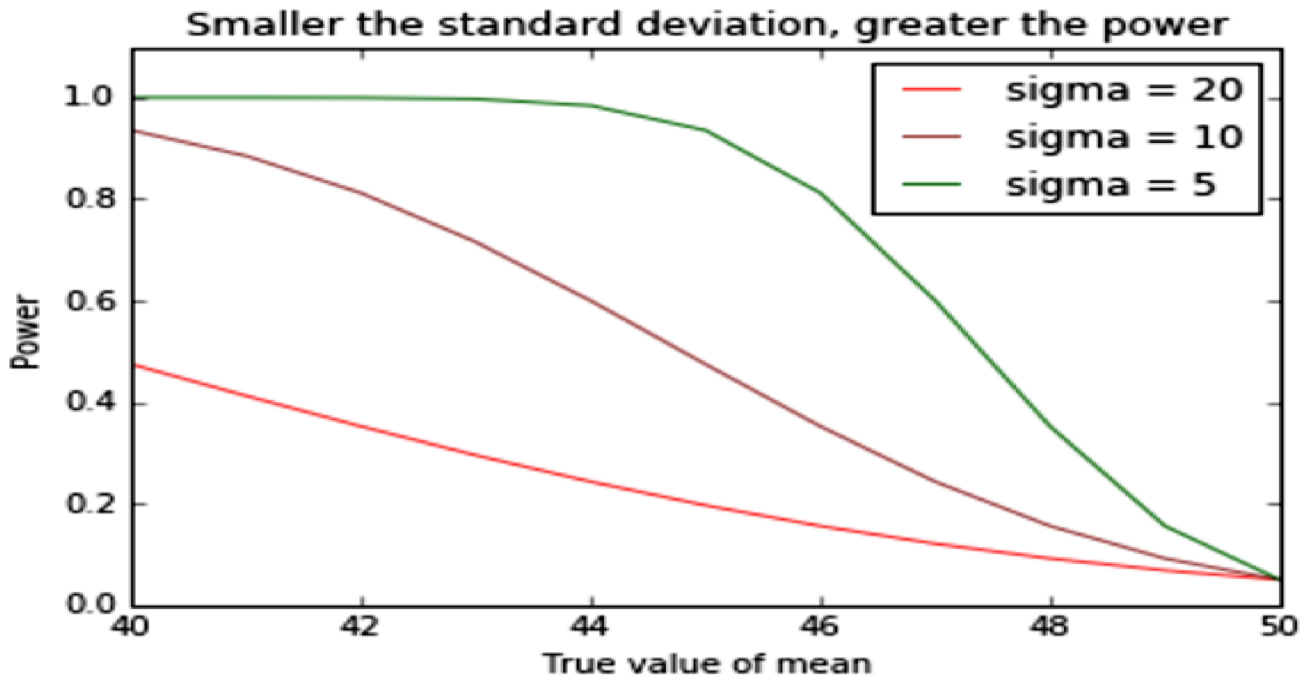
$$z = \frac{168 - 165}{\sigma/\sqrt{200}} = \frac{(168 - 165)\sqrt{200}}{\sigma}$$

Factors affecting Statistical Power of test- Standard deviation

Power

Standard deviation

Smaller the standard deviation, greater the power of the test.



Power of test

STANDARD DEVIATION

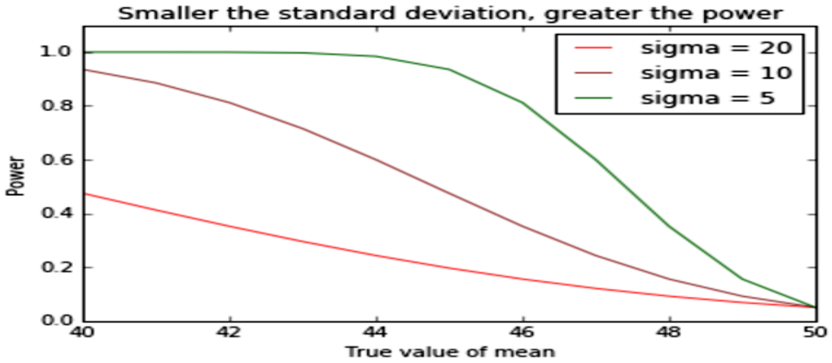
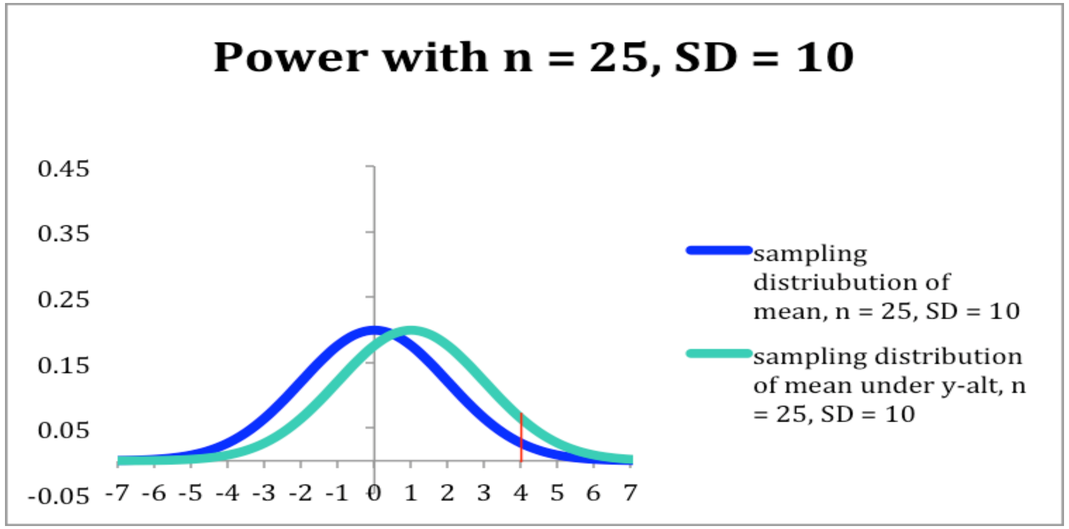
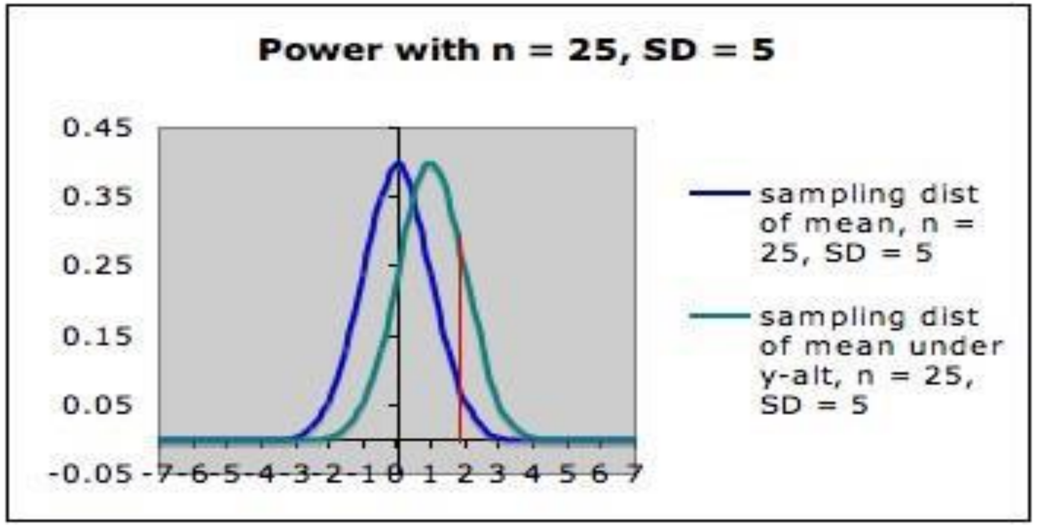


Figure also shows that power is higher when the standard deviation is small than when it is large. For all values of N, power is higher for the standard deviation of 10 than for the standard deviation of 15 (except, of course, when N = 0). Experimenters can sometimes control the standard deviation by sampling from a homogeneous population of subjects, by reducing random measurement error, and/or by making sure the experimental procedures are applied very consistently.

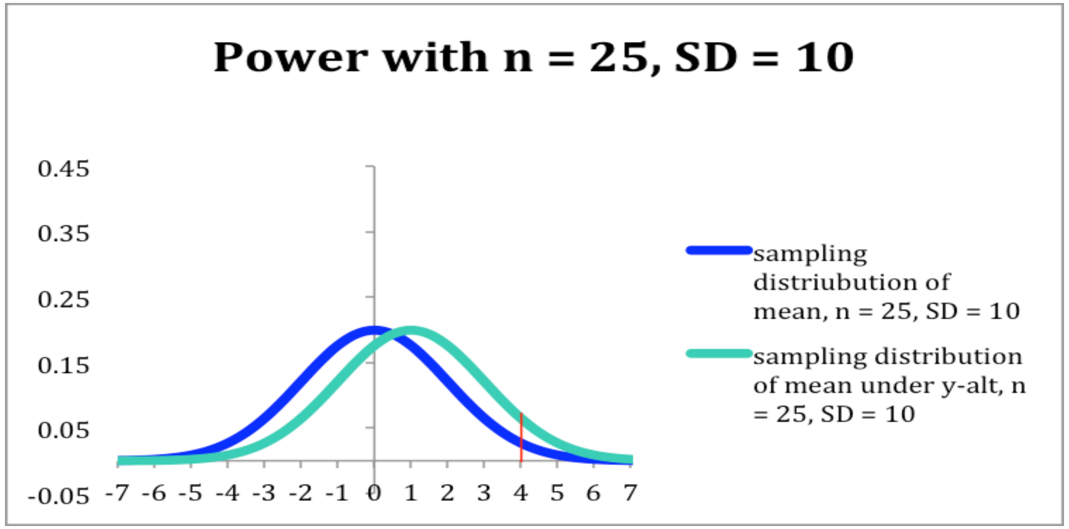
Power of test

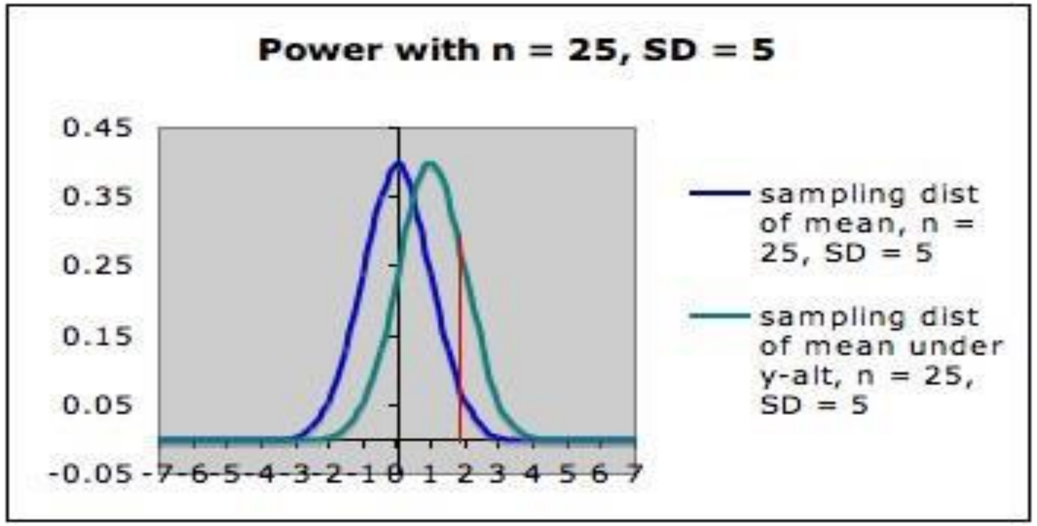




In the first picture, the standard deviation is 10; in the second picture, it is 5. Note that *both graphs are in the same scale*. In both pictures, the blue curve is centered at 0 (corresponding to the null hypothesis) and the green curve is centered at 1 (corresponding to the alternate hypothesis).

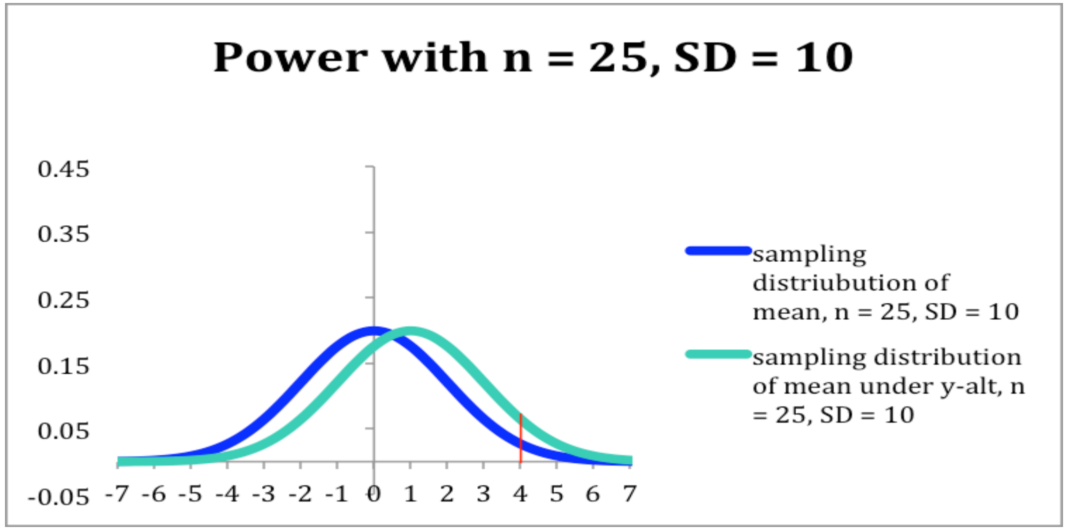
Power of test

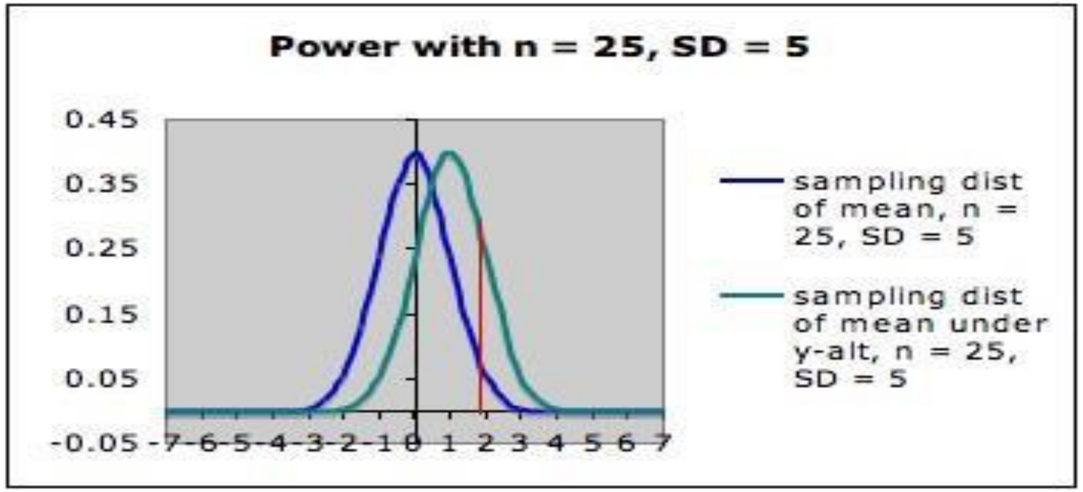




In each picture, the red line is the cut-off for rejection with alpha = 0.05 (for a one-tailed test) -- that is, in each picture, the area under the *blue* curve to the right of the red line is 0.05.

Factors affecting Statistical Power of test – Standard deviation





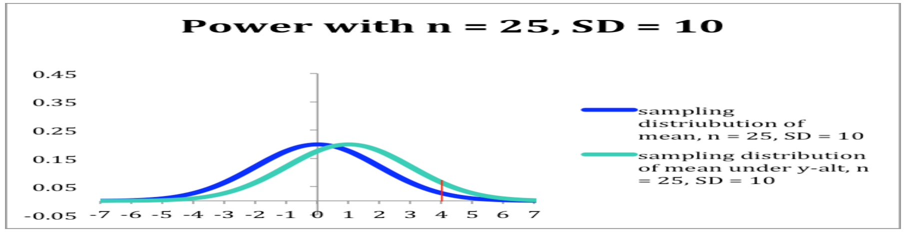
In each picture, the area under the *green* curve to the right of the red line is the power of the test against the alternate depicted. Note that this area is *larger* in the second picture (the one with smaller standard deviation) than in the first picture.

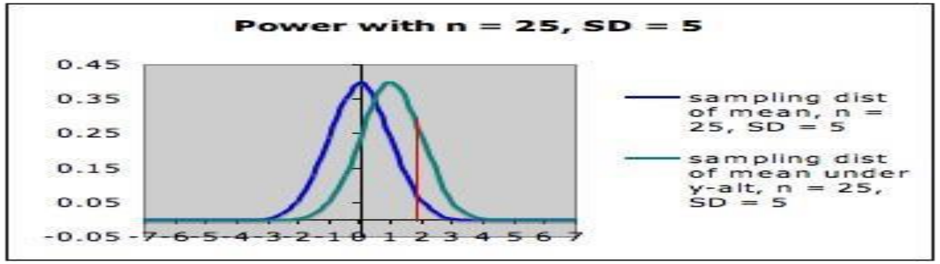
Factors affecting Statistical Power of test – Variance

Variance

Power also depends on variance: smaller variance yields higher power.

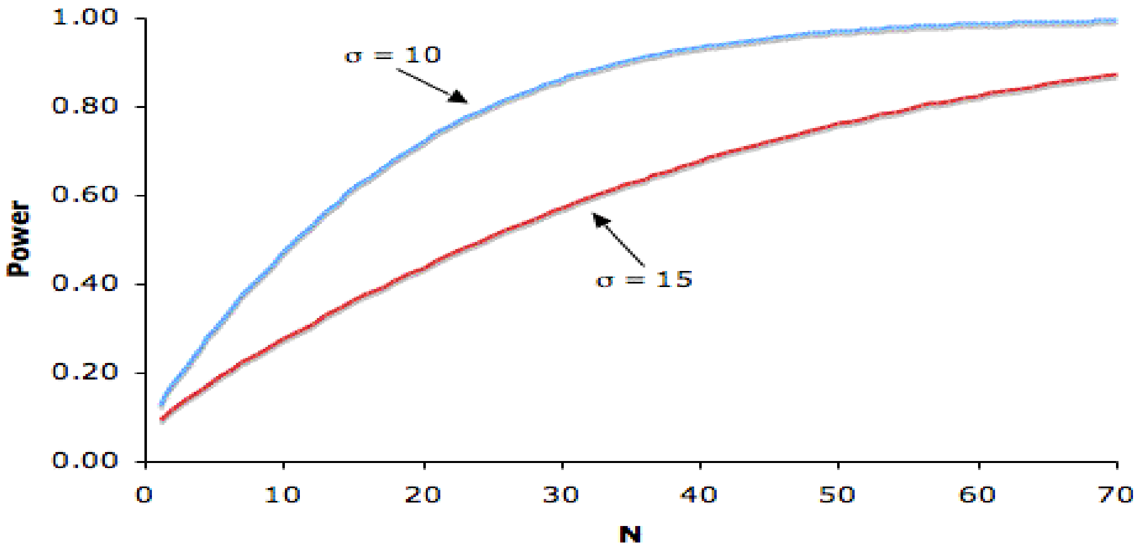
Example: The pictures below each show the sampling distribution for the mean under the null hypothesis μ = 0 (blue -- on the left in each picture) together with the sampling distribution under the alternate hypothesis μ = 1 (green -- on the right in each picture), both with sample size 25, but *for different standard deviations of the underlying distributions*. (Different standard deviations might arise from using two different measuring instruments, or from considering two different populations.)





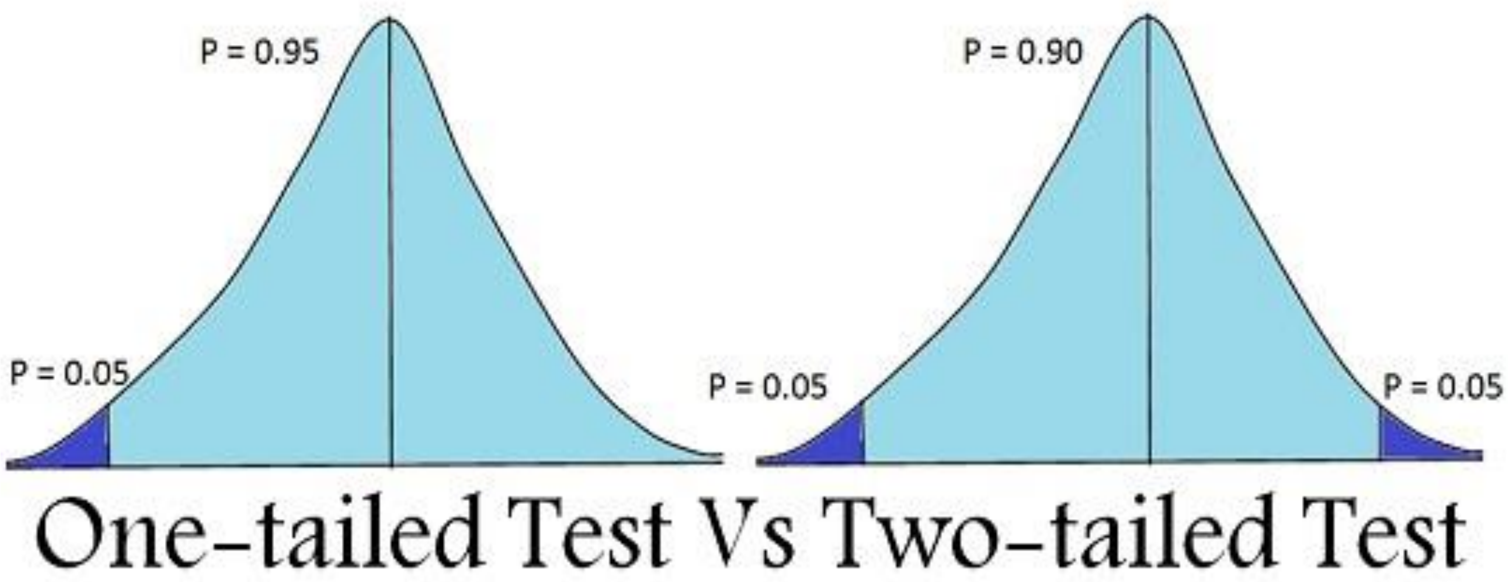
Factors affecting Statistical Power of test – Standard deviation

Smaller the standard deviation, greater the power of the test.

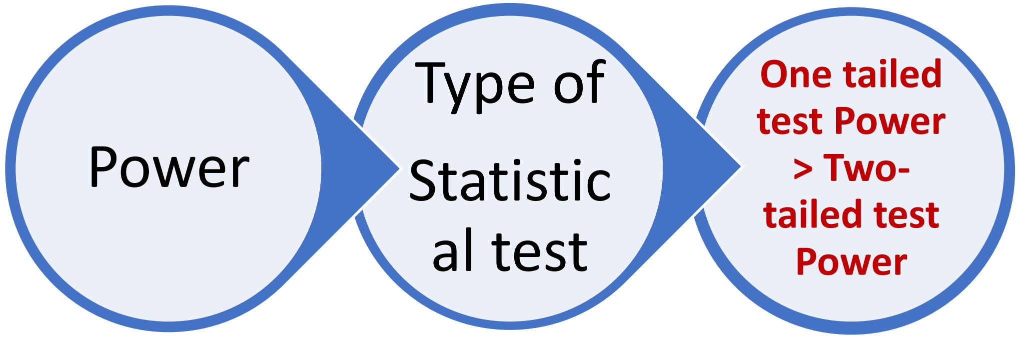


The relationship between sample size and power for H0: μ = 75, real μ = 80, one-tailed α = 0.05, for σ's of 10 and 15.

Factors affecting Statistical Power of test – Type of Statistical test



Factors affecting Statistical Power of test- Type of Statistical test



ONE- VERSUS TWO-TAILED TESTS

Power is higher with a *one-tailed* test than with a *two-tailed* test as long as the hypothesized direction is correct. A one-tailed test at the 0.05 level has the same power as a two-tailed test at the 0.10 level. A one-tailed test, in effect, raises the significance level.

Factors affecting Statistical Power of test- Summary

- Greater the sample size, the higher the power.
- The larger the significance level, the higher the power.
- The greater the effect size, the greater the power of the test..
- Smaller the standard deviation, greater the power of the test.
- One tailed test Power > Two-tailed test Power

THANK YOU

SIVASANKARI V
Department of Science & Humanities

sivasankariv@pes.edu