July 21 - 24, 2021 | Garki Hospital Abuja | Resource Person

Research Methodology Boot Camp

with Epi Info Training

Dr. Adamu Onu

MBBS, FWACP (FM)

MS Epidemiology & Biostatistics

PhD Public Health (Epidemiology)

Target Audience Clinical Researchers, Post-Part 1 Residents, and Others

Important Information

- Limited slots are available on a first come, first served basis
- Laptop running Windows 10 required
- Organized as morning lecture sessions and afternoon hands on coaching sessions

Highlights

- Research Methodology
- Research Design
- Data Management
- Sample Size Calculations
- Test Statistics
- Interpretation of Results
- Report Writing
- Hands-on training sessions
- Statistical consulting sessions

For further details contact

Email: epimetrix@gmail.com

Phone: +234 803 474 9930

epimetrix health systems

Normal Distribution and Hypothesis Testing

Probability Distributions

- The distribution (or shape) of a variable in a data set gives information about:
    - All the values the variable takes on in your data set, when the data are split into reasonably sized groups
    - How often each value occurs
    - The shape, center, and amount of variability in the data
- Checking the distribution of the data is always one of the first steps of data anlaysis.
- Knowing the shape of the data, provides insights into the data's statistical properties

The normal distribution

- There are some general shapes that occur so frequently in nature that these distributions are given their own names.
- The most well-known distribution has a shape similar to a bell and is called the normal distribution (or sometimes "the bell curve" or "Gaussian curve" or just "normal curve").
- The normal distribution is the most important **probability** distribution in statistics
- Many natural phenomena for example, heights, blood pressure, measurement error, and IQ scores follow the normal distribution.

The normal distribution

- The normal distribution has the following properties:
  - Symmetric.
  - The mean, median, and mode are all equal.
  - Half of the population is less than the mean and half is greater than the mean.
- The *Empirical Rule* allows you to determine the proportion of values that fall within certain distances from the mean.

Empirical Rule

This describes the percentage of the data that fall within specific numbers of standard deviations from the mean for the normal distribution.

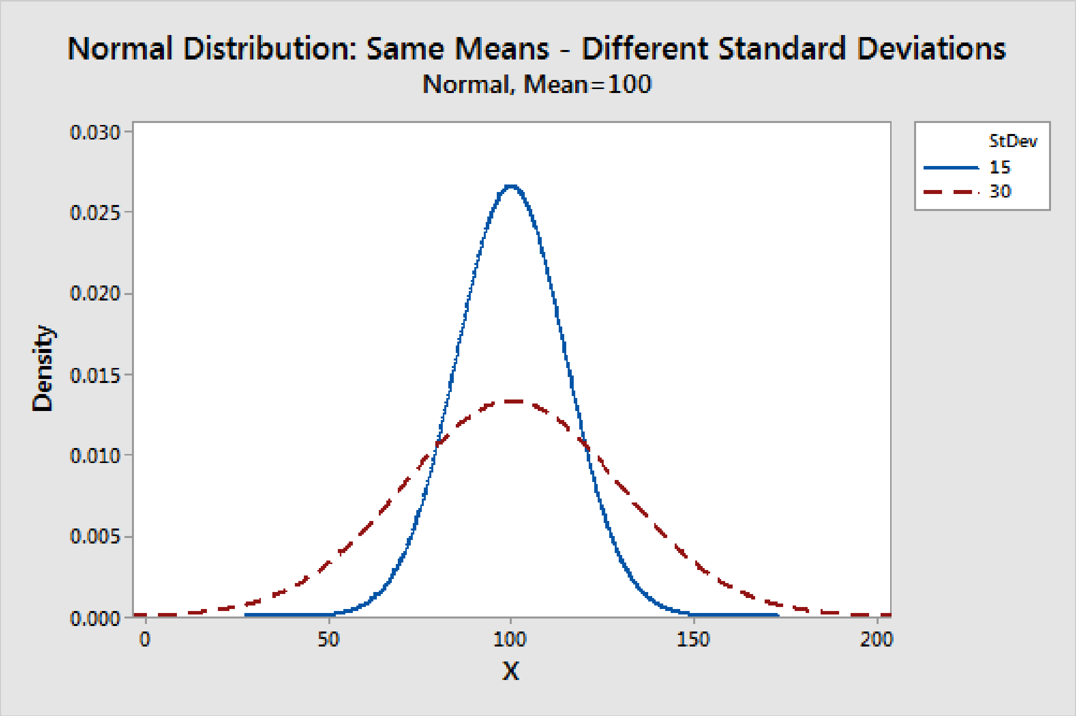
Mean ± standard deviations	Percentage of data
1	68%
2	95%
3	99.7%

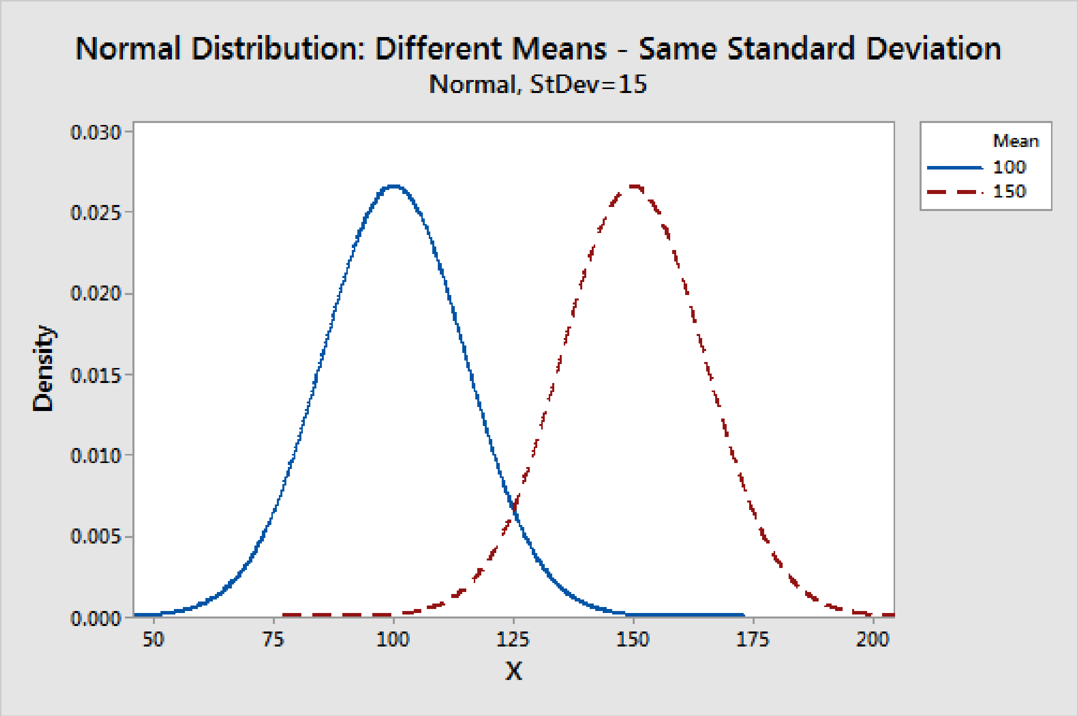
Parameters of the normal distribution

- The parameters for the normal distribution define its shape and probabilities entirely.
- The normal distribution has two parameters, the **mean (μ)** and **standard deviation (σ)**.

$$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

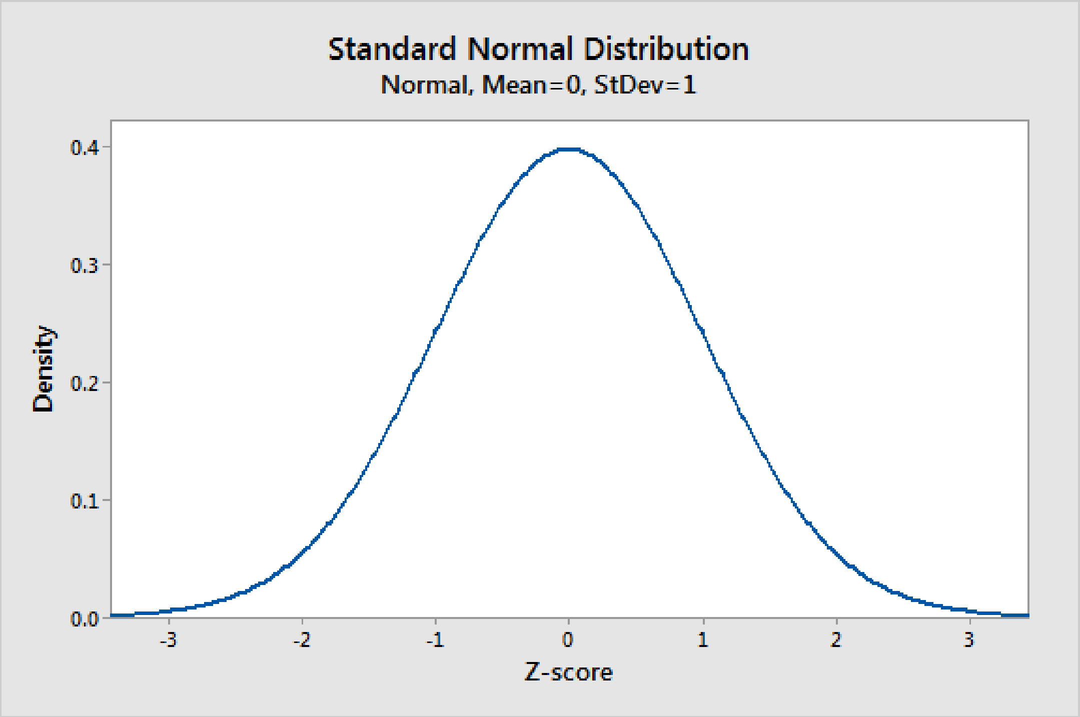
The shape of the normal distribution changes based on the parameter values





Standard normal distribution

- This is a special case of the normal distribution
- The mean is zero and the standard deviation is 1.
- This distribution is also known as the Z-distribution.



Standard normal distribution

- A value on the standard normal distribution is known as a standard score or a Z-score.
- A standard score represents the number of standard deviations above or below the mean that a specific observation falls.
- For example, a standard score of 1.5 indicates that the observation is 1.5 standard deviations above the mean.
- On the other hand, a negative score represents a value below the average. The mean has a Z-score of 0.

Standard normal distribution

$$Z = \frac{X - \mu}{\sigma}$$

- X = raw value of measurement of interest
- μ = population mean
- σ = population standard deviation

Parametric vs. non-parametric tests

- Many statistical tests are based upon the assumption that the data are sampled from a *normal* distribution.
- These tests are referred to as **parametric** tests.
- Commonly used parametric tests are:
  - Mean, SD, Student t test, unpaired t test, paired t test, ANOVA, Pearson correlation, Linear regression
- Non-parametric tests include
  - Median, IQR, Wilcoxon test, Mann-Whitney test, Kruskall-Wallis, Spearman correlation, Non-parametric regression

Parametric vs. non-parametric tests

Large samples

> What happens when you use a parametric test with data from a non-normally distributed population?

- The **central limit theorem** ensures that parametric tests work well with large samples even if the population is not normaly distributed.
- Unless the population distribution is really weird, you are probably safe choosing a parametric test when there are at least two dozen data points in each group.

Parametric vs. non-parametric tests

Large samples

> What happens when you use a **nonparametric** test with data from a non-normally distributed population?

- Nonparametric tests work well with large samples from normally distributed populations.
- The p values tend to be a bit too large, but the discrepancy is small.
- In other words, nonparametric tests are *only slightly less powerful* than parametric tests with large samples.

Parametric vs. non-parametric tests

Small samples

> What happens when you use a parametric test with data from non-normal populations?

- You can't rely on the central limit theorem, so the p value may be inaccurate.
- When you use a nonparametric test with data from a normal population, the p values tend to be too high.
- The nonparametric tests lack statistical power with small samples.

Hypothesis testing

Glossary of terms

1. **Null hypothesis (H_0)**
   - A statement that declares the observed difference is due to "chance." It is the hypothesis the researcher hopes to *reject*
2. **Alternative hypothesis (H_1 or H_A)**
   - The opposite of the null hypothesis. The hypothesis the researcher hopes to bolster.
3. **Alpha (α)**
   - The probability the researcher is willing to take in (falsely) rejecting a true null hypothesis.

Glossary of terms

4. **Test statistic**
   - A statistic used to test the null hypothesis.
5. **P value**
   - A probability statement that answers the question "If the null hypothesis were true, *what is the probability of observing the current data or data that is more extreme than the current data?*."
   - It is the probability of the data conditional on the truth of H_0.
   - It is ***not*** the probability that the null hypothesis is true.

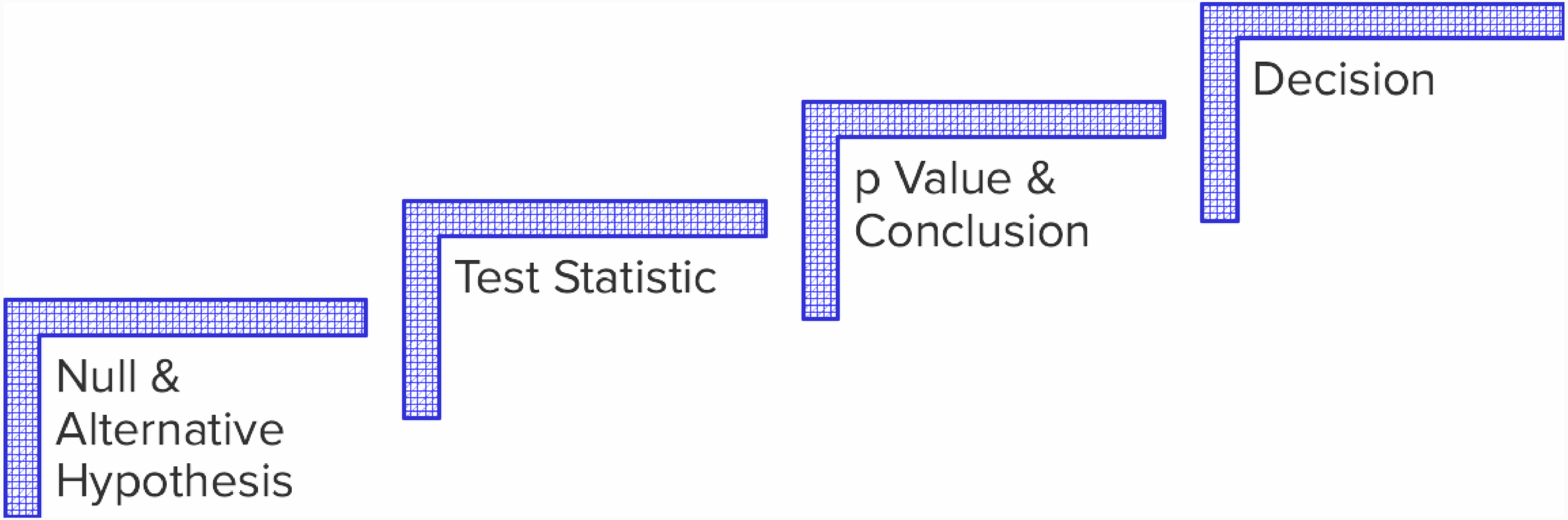
Glossary of terms

6. **Type I error**
   - A rejection of a true null hypothesis; a “false alarm,” i.e. a “false positive”
7. **Type II error**
   - A retention of an incorrect null hypothesis; “failure to sound the alarm,” i.e. a “false negative”

Glossary of terms

9. **Confidence (1 – α)**
   - The complement of alpha.
10. **Beta (β)**
    - The probability of a type II error
    - The probability of retaining a false null hypothesis.
11. **Power (1 – α)**
    - The complement of β
    - The probability of avoiding a type II error
    - The probability of rejecting a false null hypothesis.

Four steps of hypothesis testing



Step 1: Null & alternative hypothesis

- The first step of hypothesis testing is to convert the research question into **null** and **alternative** hypotheses.
- The null hypothesis (H_0) is a claim of "no difference."
- The opposing hypothesis is the alternative hypothesis (H_1 or H_A).
  - The alternative hypothesis is a claim of "a difference in the population."
  - This is the hypothesis the researcher often hopes to bolster.

> It is important to keep in mind that the null and alternative hypotheses reference population values, and not observed statistics.

Step 2: Test statistic

- We calculate a test statistic from the data.
- There are different types of test statistics depending on:
  - the nature of the data (types of variables) and,
  - the null and alternative hypotheses.
- The test statistic will *compare* the *observed sample statistic* to an *expected population parameter*.
- Large test statistics indicate data are far from expected, providing evidence against the null hypothesis and in favor of the alternative hypothesis.

Step 2: Test statistic

	Two groups		≥ Three groups	
Scale of measurement	Independent	Correlated	Independent	Correlated
Dichotomous/Nominal	χ^2, Fisher's exact	McNemar	χ^2	Cochran's Q
Ordinal	Mann-Whitney	Sign test, Wilcoxon	Kruskall-Wallis	Friedman
Continuous/Normal	*t* test	Paired *t* test	ANOVA	Repeated measures ANOVA
Continuous/skewed	Mann-Whitney	Wilcoxon	Kruskall-Wallis	Friedman

Step 3: p value & conclusion

- The test statistic is converted to a **conditional probability** called a *p value*.
- The p value answers the question:

> If the null hypothesis were true, what is the probability of observing the current data or data that is more extreme?

- Small p values provide evidence against the null hypothesis because they say ***the observed data are unlikely when the null hypothesis is true***.

Step 3: p value & conclusion

- We apply the following conventions:
  - When p value > .10 → the observed difference is “not significant”
  - When p value ≤ .10 → the observed difference is “marginally significant”
  - When p value ≤ .05 → the observed difference is “significant”
  - When p value ≤ .01 → the observed difference is “highly significant”
- Use of “significant” in this context means “the observed difference is not likely due to chance.” It does not mean of “important” or “meaningful.”

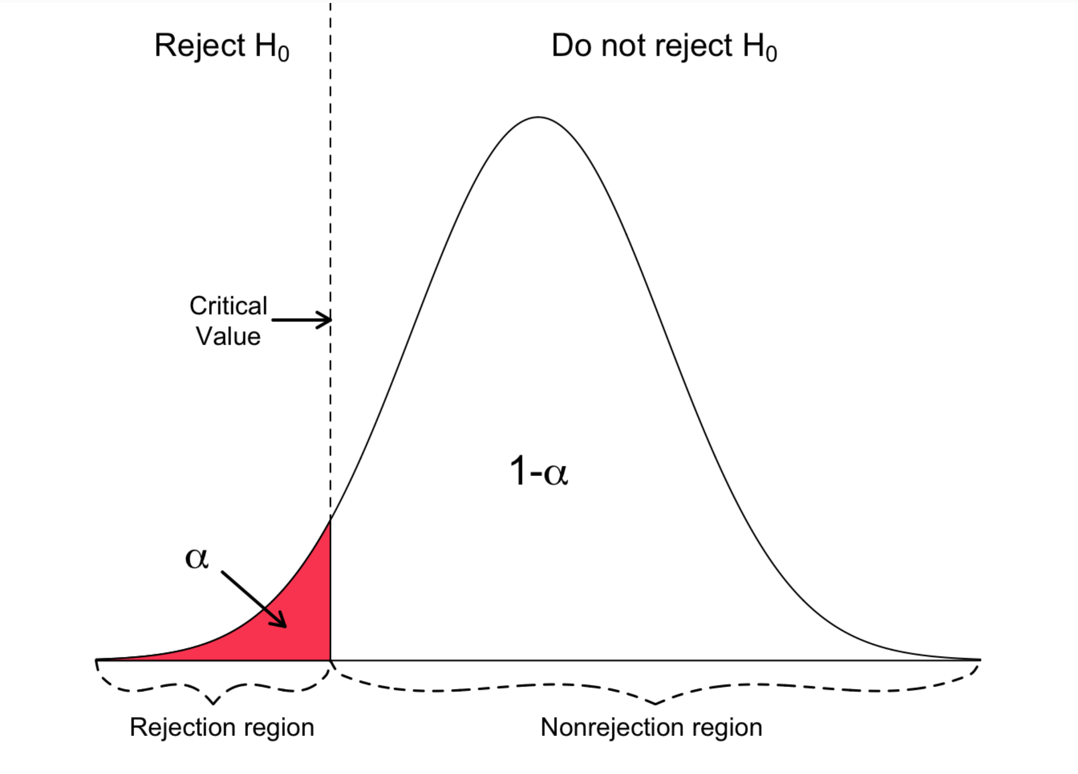
One-sided or two-sided p value?

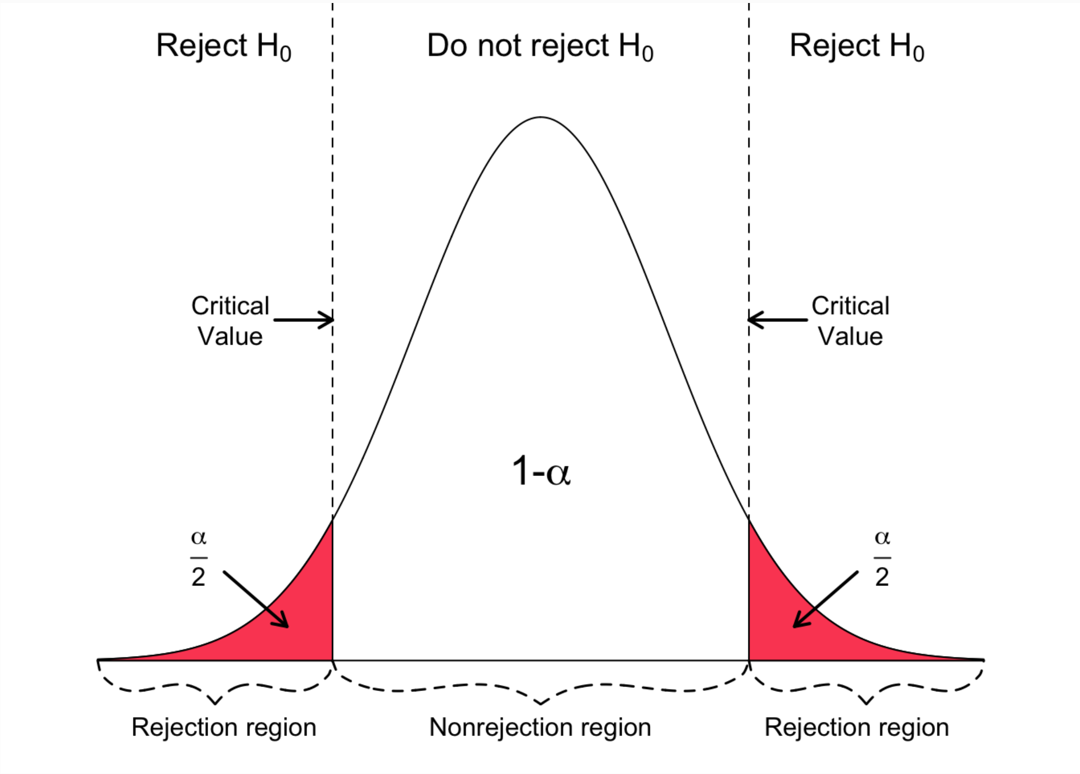
- You must choose whether you wish to calculate a one- or two-sided p value ***before collecting any data***.
- The p value is calculated for the null hypothesis that the two population parameters are equal, and any discrepancy between the two sample statistics is due to chance.

One-sided or two-sided p value?

- If this null hypothesis is true, the one-sided P value is the probability that the two sample statistics would differ as much as was observed (or further) ***in the direction specified*** by the hypothesis just by **chance**, even though the means of the overall populations are actually equal.

One-sided or two-sided p value?





One-sided or two-sided p value?

- A one-sided p value is appropriate when you can state with certainty (and before collecting any data) that:
  - there either will be no difference between the means or
  - that the difference will go in a direction you can specify in advance (i.e., you have specified which group will have the larger mean).
- If you cannot specify the direction of any difference before collecting data, then a two-sided p value is more appropriate.
- If in doubt, select a two-sided P value.

Step 4: Decision

- Alpha (α)) is a probability threshold for a decision.
- If $p \leq \alpha$, we will reject the null hypothesis.
- Otherwise it will be retained for want of evidence.

Type I error (α)

- The chance of finding a difference when none exists, α (type I error) = p value = our lack of confidence in the result. This is what we refer to as a **false positive**.
- Confidence level, (1 – α) = the probability of a true difference = our confidence in the result.

Type II error (β)

- The chance of finding no difference when there is indeed a difference (β) is called a **type II error**.
- This is what we refer to as a **false negative**.
- Power = (1 – β) = the chance of finding a difference if there really is a difference.

Type I and Type II errors (α & β)

	Defendant innocent	Defendant guilty
Guilty verdict	Type I error	Correct
Not guilty verdict	Correct	Type II error

Fallacies of statistical hypothesis testing

1. Failure to reject the null hypothesis leads to its acceptance.
    - WRONG! Failure to reject the null hypothesis implies insufficient evidence for its rejection.

Fallacies of statistical hypothesis testing

2. The p value is the probability that the null hypothesis is incorrect.
    - WRONG! The p value is the probability of the current data or data that is more extreme assuming H0 is true.

Fallacies of statistical hypothesis testing

3. α = .05 is a standard with an objective basis.
   - WRONG! α = .05 is merely a convention that has taken on unwise mechanical use.
   - There is no sharp distinction between “significant” and “insignificant” results, only increasingly strong evidence as the p value gets smaller.

> Surely God loves p = .06 nearly as much as p = .05

Fallacies of statistical hypothesis testing

4. Small p values indicate large effects.
    - WRONG! p values tell you next to nothing about the size of an effect.
5. Data show a theory to be true or false.
    - WRONG! Data can at best serve to bolster or refute a theory or claim.
6. Statistical significance implies importance.
    - WRONG! WRONG! WRONG! Statistical significance says very little about the importance of a relation.

> An approximate answer to the right question is worth a great deal more than a precise answer to the wrong question

– John Tukey


I CAN'T BELIEVE SCHOOLS ARE STILL TEACHING KIDS ABOUT THE NULL HYPOTHESIS.
I REMEMBER READING A BIG STUDY THAT CONCLUSIVELY DISPROVED IT *YEARS* AGO.