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

# *t* Tests and Paired *t* Test

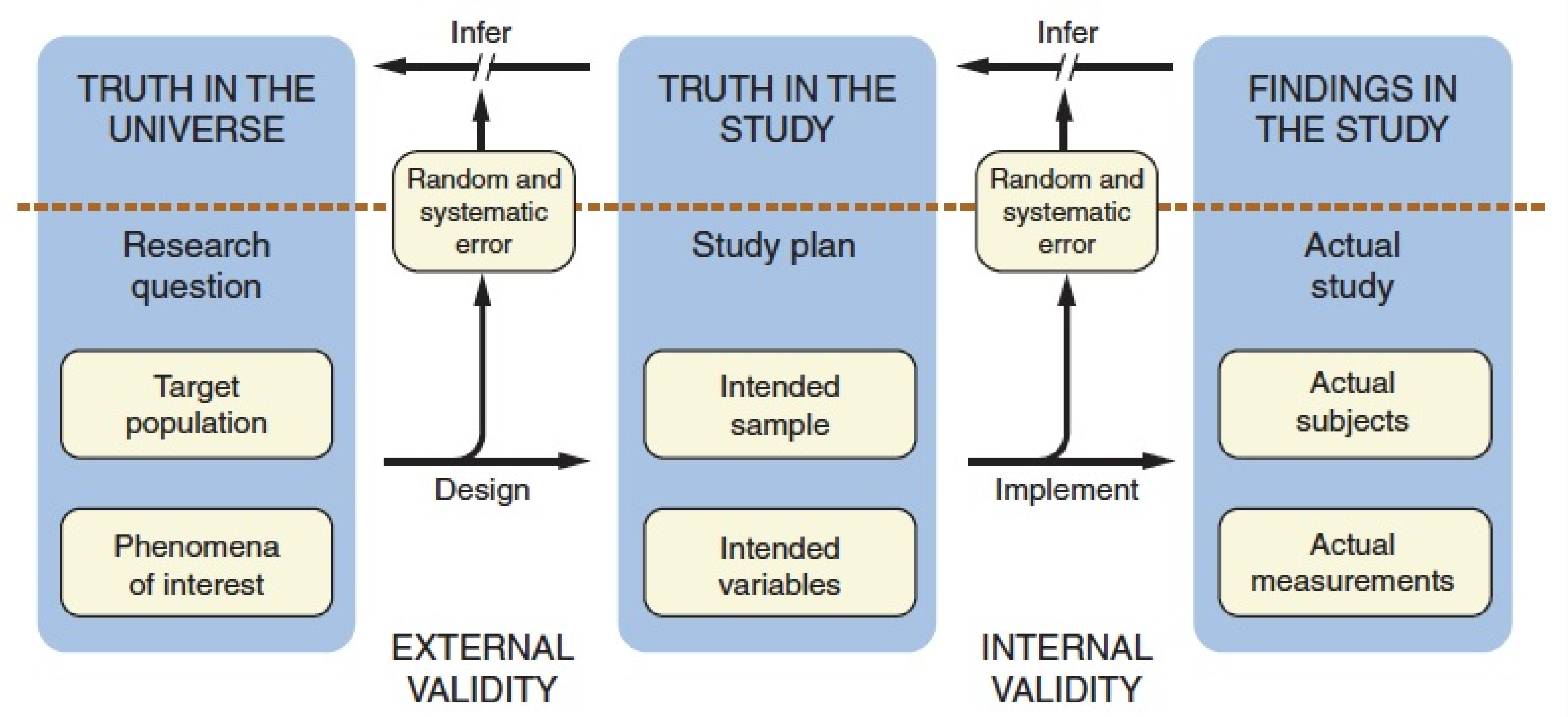
# Paraneters vs. Statistics

## Parameters

- Numerical descriptive measures corresponding to populations.
- ***Unknown*** constants referred to in Greek letters (μ, π, σ)
- Statistical inferential methods use sample data to make inferences concerning the unknown parameters.

## Statistics

- Numerical descriptive measures corresponding to samples.
- Referred to in lowercase letters ($\bar{x}$, $\hat{p}$, s).
- Since samples are "random subsets" of the population, statistics are ***random*** variables.



TRUTH IN THE UNIVERSE
Research question
Target population
Phenomena of interest
Infer
Random and systematic error
Design
EXTERNAL VALIDITY
TRUTH IN THE STUDY
Study plan
Intended sample
Intended variables
Infer
Random and systematic error
Implement
INTERNAL VALIDITY
FINDINGS IN THE STUDY
Actual study
Actual subjects
Actual measurements

# Inferential statistics

- Inferential statistics provides the tools to help answer these questions
  - Whether the estimates made in the study can be generalized beyond the relatively small number of the sample - external validity or generalizability.
  - Whether the differences or associations found can be explained by chance, and so may not be real.
- An inference is a generalization made about a population from the study of a subset or sample of that population → Inferential statistics.

# Choosing a statistical test

- What is the scale of measurement used to collect the data?
  - (nominal, ordinal, or interval)
- Are you comparing means or proportions?
- How many groups are being compared? ($k = 2$, $k \geq 3$)
- Are the variables independent, correlated or paired?
- Is the number of observations sufficient to use a particular statistic?
- Do the frequencies follow a ***normal*** distribution?

# Choosing a statistical test

- The choice of an appropriate statistical test is depends on
  - Goals & objectives of the research
  - Research design
  - Types of variables
  - Number of groups of observations
  - Distribution of data
- Test statistics are fundamental to statistical inferences

> Good answers come from good questions not from esoteric analysis
>
> — Schoolman et al, (1968)

# Uncertainty

- If the study sample is not representative of the population, the inference we make from the result will be misleading.
    - Inferential statistics will not help if the sample is not representative.
    - Inferential statistics cannot correct mistakes in designing the study.
- Results from a single sample are still subject to some degree of uncertainty, or chance.
- This ***sampling error*** cannot be eliminated completely, but its probable magnitude can be calculated.

# Generalizability

- Generalization from the sample to the population from which the sample was drawn depends mainly on:
  - Size of the sample
  - Variability in the results

## Size of the sample

- If we have examined 100% of the community, then the result will represent the finding in the whole community.
- The smaller the sample drawn, the less likely its findings can be generalized.

## Variability in the results

- Marked variation between measurements means different results are more likely to be obtained from different samples.
- If the results fall within a wide range (high variability) then a small sample will be less likely to represent the parameter.

## Standard error

$$se = \frac{sd}{\sqrt{n}}$$

- The standard error (SE) is a statistical measure about the **probability** that the finding in the sample will reflect the finding in the population.
- The SE depends on 2 factors:
    - Size of the sample,
    - Variation in the sample (SD).

## Standard error

- The SE can be used to produce a **confidence interval**.
    - A confidence interval is a range of values which includes the population parameter at a specified level of probability.
    - For example, the sample mean ($\bar{x}$) ± (1.96 × SE) gives the 95% confidence interval, meaning that there is only a 5% chance that this interval does not include the mean of the population.

# Standard error

- The SE can be calculated not only on a mean, but also for
    - the difference between two means
    - on a percentage
    - on a difference between two percentages
    - on a correlation coefficient.

## Standard error

- The standard error should be clearly differentiated from the standard deviation (*SD*).
    - The *SD* is a measure of the variability in the sample studied.
    - The *SE* is a measure of the uncertainty in a sample statistic.
    - The *SE* depends on both the SD & the sample size, and is a recognition that ***a sample is unlikely to determine the population value exactly***.

# Confidence intervals

- Statistical significance of a result does not give us an indication of the magnitude of that difference in the population from which the sample was studied.
- CIs can be calculated on the result of just about any statistical test.
- CIs allow us to estimate the strength of the evidence.
  - If the confidence interval is narrow, the strength of evidence will be strong.
  - Wide CIs indicate greater uncertainty about the true value of the result.

# Confidence intervals

- In reporting CIs, a **point estimate** of the result is given together with a **range of values** that are consistent with the data, and within which we expect the true value in the population to lie.
- The CI thus **provides a range of possibilities** for the population value.
- This is in contrast to statistical significance which only indicates whether or not the finding can be **explained by chance**.

# Confidence intervals

- As in statistical tests, the investigators must select the degree of confidence or certainty they accept to be associated with a confidence interval.
- **95%** is the most common choice, just as a **5%** level of statistical significance is widely used.
- In general, when a 95% CI contains a **zero difference**, it means that one is unable to reject the null hypothesis at the 5% level.

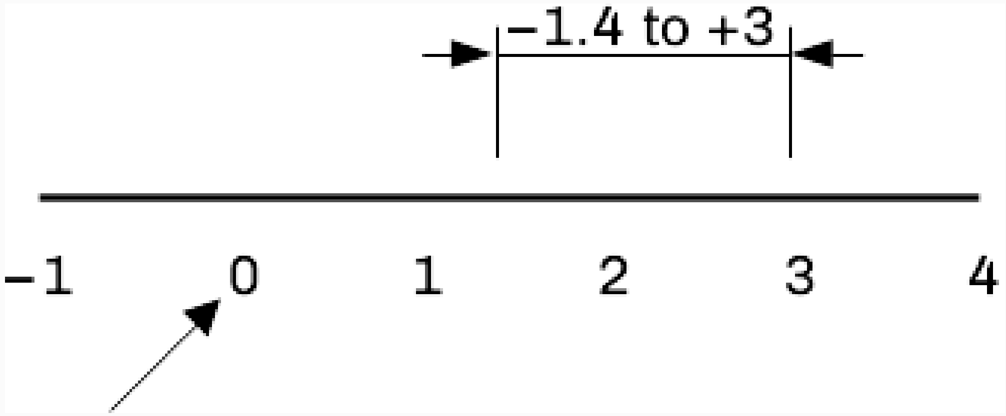
# Confidence intervals

**Mean haemoglobin**

| Male | Female |
|---|---|
| 13.2 g/dl | 11.7 g/dl |

- CI for the difference for males and female is −0.4 to +3 g/dL
  - We fail to reject the null hypothesis that there is no difference because the confidence interval ***includes 0***.

- CI for the difference males and female is 1.4 to +3 g/dL
    - We reject the null hypothesis that there is no difference because the confidence interval ***excludes 0***.



# Student *t* test

- Used in all areas of medicine
- Similar in shape to the z distribution
- Test was developed by William Gosset, who used the pseudonym Student
- The *t* distribution has a similar shape to the normal distribution, but is more widely spread out and flatter.
- The degree of spread and flatness changes according to the sample size.
- As the sample size gets larger, the *t* distribution becomes virtually identical to the normal distribution.

# *t* test

- One-sample *t* test
- Two-sample independent *t* test
- Paired *t* test

**Non-parametric equivalents**

- Wilcoxon rank-sum test or Mann-Whitney U test
- Wilcoxon signed-rank test

# When to use the *t* test

- Quantitative continuous variables
- Samples are randomly selected
- Variable is normally distributed
- Small sample (size ≤ 30)

$$t = \frac{\bar{x} - \mu_0}{sd/\sqrt{n}}$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

## One-sample *t* test

> Suppose we want to test the hypothesis that mothers with low socioeconomic status (SES) deliver babies whose birthweights are lower than "normal."

- To test the hypothesis, a list is obtained of birthweights from 100 consecutive, full-term live-born deliveries from the maternity ward of a hospital in a low-SES area.

## One-sample *t* test

- The mean birhweight ($\bar{x}$) is found to be 115 oz with a sample standard deviation ($s$) of 24 oz.
- From national surveys we know that the mean birthweight is 20 oz.
- Can we say the underlying mean birthweight from this hospital is lower than the national average?

## One-sample *t* test

- The hypothesis being considered can be formulated in terms of null and alternative hypotheses:
  1. The null hypothesis ($H_0$) is that the mean birthweight in the low-SES-area hospital ($\mu$) is equal to the mean birthweight in the country ($\mu_0$).
  2. The alternative hypothesis ($H_1$) is that the mean birthweight in this hospital ($\mu$) lower than the mean birthweight in the country ($\mu_0$).

$$H_0: \mu = \mu_0 \text{ vs. } H_1: \mu < \mu_0$$

## One-sample *t* test

- The statistical model in this case is that the birthweights come from a normal distribution with mean $\mu$ and unknown variance $\sigma^2$.
- We wish to test the null hypothesis, $H_0$, that $\mu$ = 120 oz vs. the alternative, $H_1$, that $\mu$ < 120 oz.

## One-sample *t* test

- There are four possible outcomes

| | Truth | |
|---|---|---|
| **Decision** | $H_0$ | $H_1$ |
| Accept $H_0$ | $H_0$ is true and $H_0$ is accepted | $H_1$ is true and $H_0$ is accepted |
| Reject $H_0$ | $H_0$ is true and $H_0$ is rejected | $H_1$ is true and $H_0$ is rejected |

## Type I error

| | Truth | |
|---|---|---|
| **Decision** | $H_0$ | $H_1$ |
| Accept $H_0$ | $H_0$ is true and $H_0$ is accepted | $H_1$ is true and $H_0$ is accepted |
| Reject $H_0$ | $H_0$ is true and $H_0$ is rejected | $H_1$ is true and $H_0$ is rejected |

## Type II error

| | Truth | |
|---|---|---|
| **Decision** | $H_0$ | $H_1$ |
| Accept $H_0$ | $H_0$ is true and $H_0$ is accepted | $H_1$ is true and $H_0$ is accepted |
| Reject $H_0$ | $H_0$ is true and $H_0$ is rejected | $H_1$ is true and $H_0$ is rejected |

## One-sample *t* test

- To test the hypothesis we compute the test statistic ***t*** based on the *t* distribution:

$$t = \frac{\bar{x} - \mu_0}{s/n} \qquad \begin{cases} t < t_{n-1,\alpha} & \text{then we reject } H_0 \\ t \geq t_{n-1,\alpha} & \text{then we accept } H_0 \end{cases}$$

- The value $t_{n-1,\alpha}$ is called a **critical value**

## One-sample *t* test

$$t = \frac{\bar{x} - \mu_0}{s/n}$$

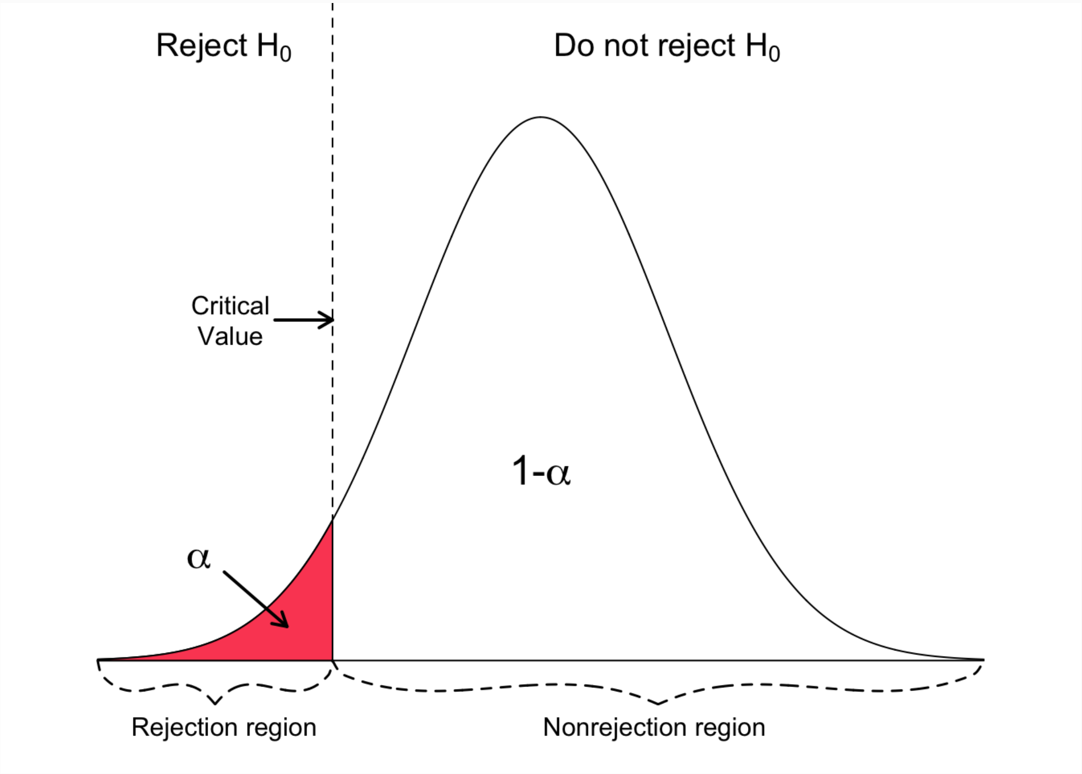
$$= \frac{115 - 120}{24/\sqrt{100}}$$

$$= \frac{-5}{2.4} = -2.08$$

# One-sample *t* test

## Decision

- $t_{99,0.05} = -1.66$
- Because −2.08 < −1.66, it follows that we can reject $H_0$ at the significance level of 0.05.
- The **probability** of $t < t_{n-1,\alpha}$ is the p value.



## Two-sample independent *t* test

- A sample of eight 35- to 39-year-old nonpregnant, premenopausal OC users have a mean systolic blood pressure (SBP) of 132.86 mm Hg and sample standard deviation of 15.34 mm Hg are identified.
- Another sample of 21 nonpregnant, premenopausal, non-OC users in the same age group are similarly identified who have mean SBP of 127.44 mm Hg and sample standard deviation of 18.23 mm Hg.

> What can be said about the underlying mean difference in blood pressure between the two groups?

# Two-sample independent *t* test

## Assumptions

- SBP is normally distributed in the first group with mean $\mu_1$ and variance $\sigma_1^2$ and in the second group with mean $\mu_1$ and variance $\sigma_1^2$.
- We wish to test the hypothesis $H_0$: $\mu_1 = \mu_2$ vs. $H_1$: $\mu_1 \neq \mu_2$

# Two-sample independent *t* test

- Assuming equal variances:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

- Where:

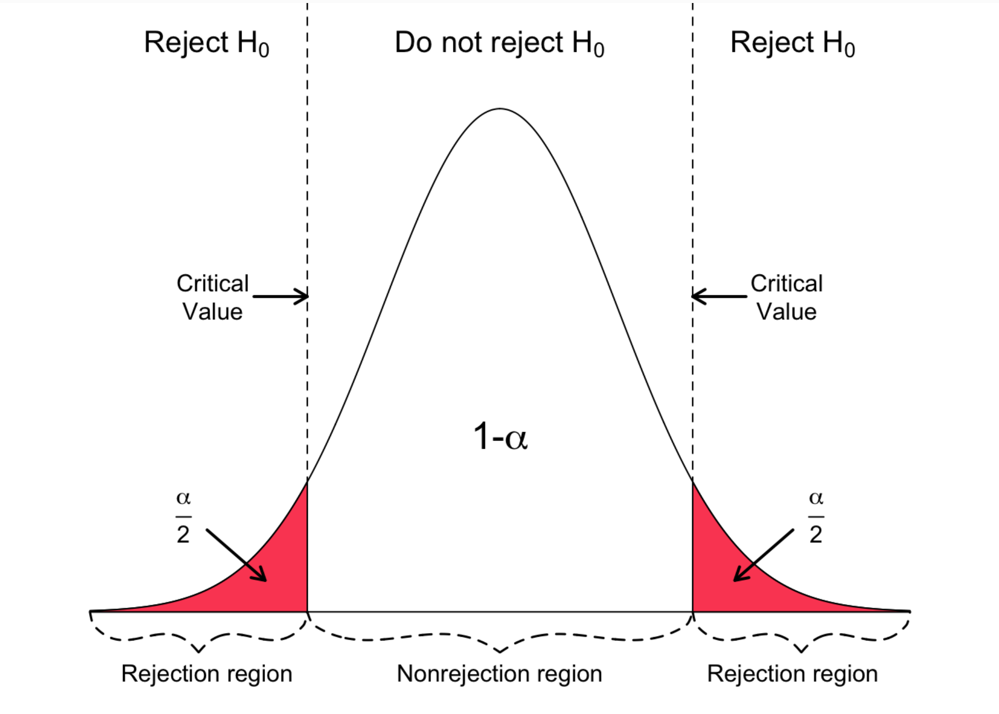
$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

- Assuming unequal variances

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- Where approximate *df* (d′):

$$d' = \frac{\left(s_1^2/n_1 + s_1^2/n_1\right)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}$$




$$\begin{cases} t > t_{n_1+n_2-2,1-\alpha/2} \text{ or } t < t_{n_1+n_2-2,1-\alpha/2} & \text{then } H_0 \text{ is rejected} \\ -t_{n_1+n_2-2,1-\alpha/2} \leq t \leq t_{n_1+n_2-2,1-\alpha/2} & \text{then } H_0 \text{ is accepted} \end{cases}$$

# Two-sample independent *t* test

- Assuming equal variances:

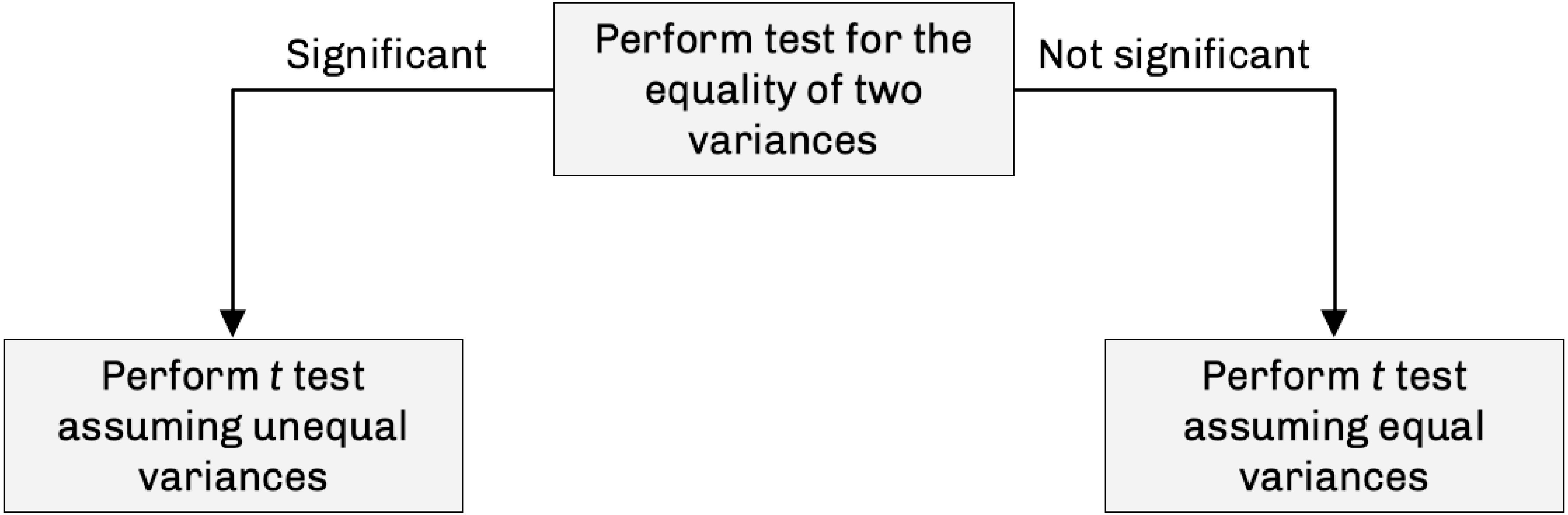
$$t = \frac{132.86 - 127.44}{17.527\sqrt{1/8 + 1/21}}$$

$$= \frac{5.42}{17.527 \times 0.415}$$

$$= 0.74$$

- $t_{27,\ .975} = 2.052$
- $-2.052 \leq 0.74 \leq 2.052$
- $H_0$ is accepted using a two-sided test at the 5% level

## Strategy for testing the equality of means in two independent normally distributed samples



## Paired *t* test

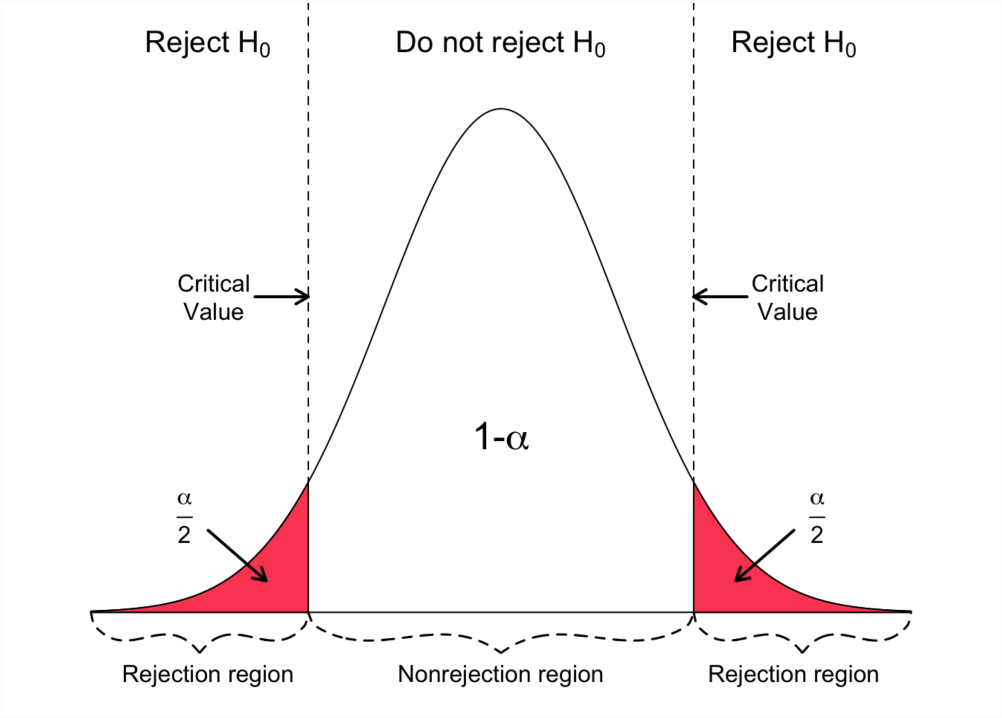
- This may be considered to be an extension of the one-sample *t* test based on the differences ($d_i$) for **matched** pairs.
- Thus the hypothesis $H_0$: $\Delta = 0$ vs. $H_1$: $\Delta \neq 0$, when variance is unknown is based on the mean difference:

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$$

# Paired *t* test

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}$$

$$s_d = \sqrt{\frac{\sum_{i=1}^{n} d_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} d_i\right)^2}{n-1}}$$



## Paired data

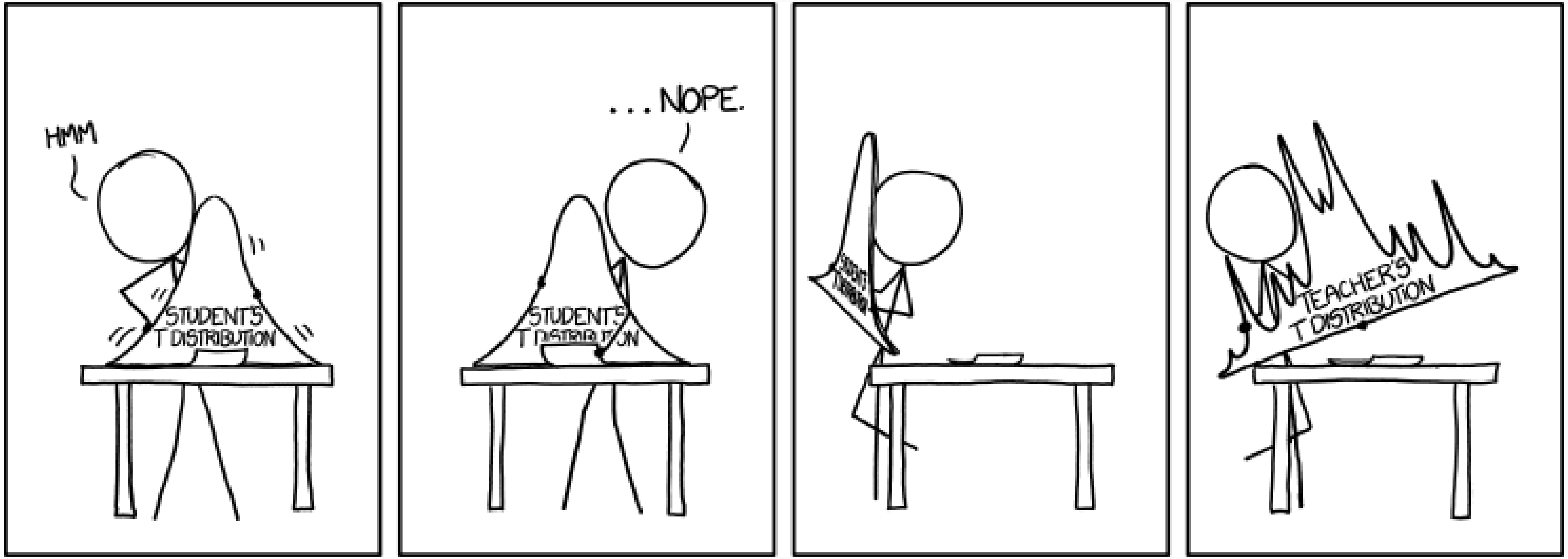
> A group of nonpregnant women, premenopausal women age 16-49 years had their systolic blood pressure measured at *baseline* while not on OC. The blood pressures were again measured after 1 year of OC use.

| ***i*** | **Baseline** | **After 1 year** | ***d*** |
|---|---|---|---|
| 1 | 115 | 128 | 13 |
| 2 | 112 | 115 | 3 |
| 3 | 107 | 106 | −1 |
| 4 | 119 | 128 | 9 |
| 5 | 115 | 122 | 7 |
| 6 | 138 | 145 | 7 |
| 7 | 126 | 132 | 6 |
| 8 | 105 | 109 | 4 |
| 9 | 104 | 102 | −2 |
| 10 | 115 | 117 | 2 |

## Paired *t* test

$$
\begin{aligned}
\bar{d} &= (13 + 3 + \cdots + 2)/10 = 4.80 \\
s_d^2 &= [(13 - 4.8)^2 + \cdots (2 - 4.8)^2]/9 = 20.844 \\
s_d &= \sqrt{20.844} = 4.566 \\
t &= 4.80/(4.566/\sqrt{10}) = 3.32
\end{aligned}
$$

- $t_{9,\ .975} = 2.262$
- Since 3.32 > 2.262 and −3.32 < −2.262
- $H_0$ is rejected using a two-sided test with α = .05


HMM
STUDENT'S T DISTRIBUTION
...NOPE.
TEACHER'S T DISTRIBUTION