Research Question: Hypotheses

Prof Kumar T Bhowmik
Additional Medical Superintendent
Safdarjang Hospital & VMMC
New Delhi

ICRO- Patna- 07/09/2013

- Have you ever made a guess or a prediction?
- Have you ever predicted who was going to win the IPL?
- Have you ever predicted who was going to win the KBC?

If you answered yes than you have made a ***hypothesis***.

Definition

- the word hypothesis is derived form the Greek words
  - ✓ "hypo" means under
  - ✓ "tithemi" means place

- Hypotheses

..... is a statement subject to verification

..... a guess but experienced guess based on some facts

-

Definition

- **..is a hunch, assumption, suspicion, assertion or an idea about a phenomena, relationship, or situation, the reality of truth of which one do not know**
- **..a researcher calls these assumptions, assertions, statements, or hunches hypotheses and they become the basis of an inquiry.**

Definition

- **In most cases, the hypothesis will be based upon either previous studies or the researcher's own or someone else's observations**
- **Hypothesis is a conjectural statement of relationship between two or more variable**

Kerlinger, Fried N, Foundations of Behavioural Research , 3rd edition, New York: Holt, Rinehart and Winston, 1986)

Definition

- **Hypothesis is proposition, condition or principle which is assumed, perhaps without belief, in order to draw its logical consequences and by this method to test its accord with facts which are known or may be determined** (Webster's New International Dictionary of English)
- **A tentative statement about something, the validity of which is usually unknown** (Black, James A & Dean J Champion, Method and Issues in Social Research, New York: John Wiley & Sons, Inc, 1976)
- **Hypothesis is proposition that is stated is a testable form and that predicts a particular relationship between two or more variable. In other words, id we think that a relationship exists, we first state it is hypothesis and then test hypothesis in the field** (Baily, Kenneth D, Methods of Social Research, 3rd edition, New York: The Free Press, 1978)

Definition

- **A hypothesis is written in such a way that it can be proven or disproven by valid and reliable data – in order to obtain these data that we perform our study**

(Grinnell, Richard, Jr. Social Work Research and Evaluation, 3rd edition, Itasca, Illinois, F.E. Peacock Publishers, 988)

- **A hypothesis may be defined as a tentative theory or supposition set up and adopted provisionally as a basis of explaining certain facts or relationships and as a guide in the further investigation of other facts or relationships**

(Crisp, Richard D, Marketing Research, New York: McGraw Hill Book Co., 1957)

Characteristics

- **Hypotheses has the following characteristics:**
  - ✓ **a tentative proposition**
  - ✓ **unknown validity**
  - ✓ **specifies relation between two or more variables**

Functions

- **Bringing clarity to the research problem**
- **Serves the following functions**

✓ **provides a study with focus**

✓ **signifies what specific aspects of a research problem is to investigate**

✓ **what data to be collected and what not to be collected**

✓ **enhancement of objectivity of the study**

✓ **formulate the theory**

✓ **enable to conclude with what is true or what is false**

Characteristics

- Simple, specific, and contextually clear
- Capable of verification
- Related to the existing body of knowledge
- Operationalisable

Typologies

- Three types
  - ✓Working hypothesis
  - ✓Null hypothesis
  - ✓Alternate hypothesis

Working Hypotheses

- The working or trail hypothesis is provisionally adopted to explain the relationship between some observed facts for guiding a researcher in the investigation of a problem.
- A Statement constitutes a trail or working hypothesis (which) is to be tested and conformed, modifies or even abandoned as the investigation proceeds.

Typologies
Null Hypothesis

A null hypothesis is formulated against the working hypothesis; opposes the statement of the working hypothesis

....it is contrary to the positive statement made in the working hypothesis; formulated to disprove the contrary of a working hypothesis

When a researcher rejects a null hypothesis, he/she actually proves a working hypothesis

In statistics, to mean a null hypothesis usually Ho is used. For example, Ho →Q = O, where Q is the property of the population under investigation and O is hypothetical

Typologies
Alternate hypothesis

An alternate hypothesis is formulated when a researcher totally rejects null hypothesis

He/she develops such a hypothesis with adequate reasons

The notion used to mean alternate hypothesis is $H_1 \rightarrow Q>O$ i.e., Q is greater than O

Example

- **Working hypothesis: Population influences the number of bank branches in a town**
- **Null hypothesis (Ho): Population do not have any influence on the number of bank branches in a town.**
- **Alternate hypothesis (H_1): Population has significant effect on the number of bank branches in a town. A researcher formulates this hypothesis only after rejecting the null hypothesis.**

Hypotheses Definitions

- **Hypotheses are predictions about the relationship among two or more variables or groups based on a theory or previous research (Pittenger, 2003)**
- **Hypotheses are assumptions or theories that a researcher makes and tests.**

Importance of Hypotheses

- **Hypotheses:**
  - **Direct our observations**
    - **Identifies the variables examined and data to be collected**
    - **Describe a relationship among variables**
    - **Can state that as one variable increases, the other will decrease; as one variables increases, the other will increase, and so on.**
    - **Refer to populations**
      - **Hypotheses help researchers infer that results of a sample will translate to a population**

4 Functions of Hypotheses

- **Hypotheses can:**
  - **Estimate Population Characteristics**
  - **Correlate Variables**
  - **Display Differences among Two or more populations**
  - **Show possible Cause and Effect**

  What research designs relate to each of these 4 functions?

Hypothesis

- **A hypothesis is basically a guess or prediction based on what you think might happen or what you may think will be true.**
- **It can then be either proved or disproved my observation or collecting data**

Hypotheses

Shape and guide a research study in terms of:

- identification of study sample size
- what issues should be involved in data collection
- the proper analysis of the data
- data interpretation

Hypothesis Formulation

--- **Formulate a hypothesis**

--- **Frame the hypothesis in a format that is testable**

--- **Test the hypothesis**

Hypothesis Formulation

- **Observations from:**
  - **Literature (review PubMed on topic area)**
  - **Natural experiments (e.g. migrant studies)**
  - **Multi-national comparisons**
  - **Descriptive studies (assessment of person, place, and time characteristics)**
  - **Creativity**

Hypotheses Formulation

- **Infectious and chronic diseases show great variation from one country to another.**
- **Some differences may be attributed to:**
  - **---** **Climate**
  - **---** **Cultural factors**
  - **---** **Diet**
  - **---** **Genetics**

Descriptive Study Designs

Used to help formulate hypotheses

Case Series Approach

- Identify the experience of a group of patients with a similar diagnosis, or
- Identify the experience of a group of individuals with an exposure in common
  - Patients or individuals may be identified from a single or multiple sources

Population Survey Approach

- Describe issues related to disease or exposure in populations
- Usually rely upon routinely collected data from established surveillance or notifiable disease systems

Unique Component: usually identify the characteristics of an issue from a representative sample of the population

Three essential characteristics that we look to measure in descriptive studies are...

- Person
- Place
- Time

Person

Since disease not does occur at random:

What kinds of people tend to develop a particular disease, and who tends to be spared? What's unusual about those people?

Person Factors

- **Age, gender, race, ethnicity**
- **Genetic predisposition**
- **Concurrent disease**
- **Diet, exercise, smoking**
- **Risk taking behavior**
- **SES, education, occupation**

Place

Since disease not does occur at random:

Where is the disease especially common or rare, and what is different about those places?

Place Factors

- **Geographic place**
  - **residence**
  - **occupation**
  - **climate**
  - **geology**
  - **population density**
  - **economic development**
  - **nutritional practices**
  - **medical practices**

Time

Since disease not does occur at random:

How does disease frequency change over time, and what other factors are temporally associated with those changes?

Time Factors

- **Calendar Time / Time of Day**
- **Time since an event**
- **Date of onset**
- **Age (time since birth in the young)**
- **Seasonality**
- **Temporal trends**

Hypotheses Formation

Discoveries or hypotheses are sometimes resisted because they seem counter-intuitive

Hypothesis Framing

Traditionally…..

H_0: "Null" hypothesis (assumed)
H_1: "Alternative" hypothesis

Hypothesis Framing

H_0: There is no association between the exposure and disease of interest

H_1: There is an association between the exposure and disease of interest (beyond what might be expected from random error alone)

Hypothesis Framing

Another Type of Framing:

What is the best estimate of the risk of disease in those who are exposed compared to those who are unexposed (i.e. exposed are at XX times higher risk of disease).

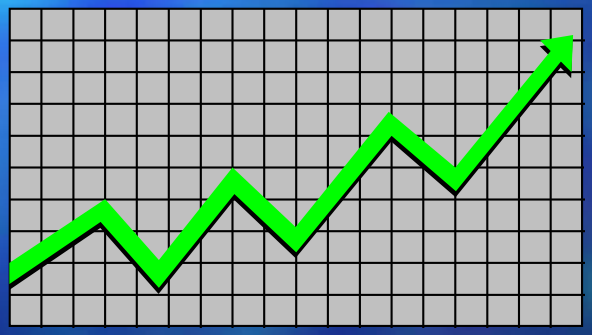
This moves away from the simple dichotomy of yes or no for an exposure/disease association – to the estimated magnitude of effect irrespective of whether it differs from the null hypothesis.

Hypothesis Framing

Ways to Express Hypotheses:

1. Suggest possible events…

The rate of survival will increase after surgery.



Hypothesis Framing

Ways to Express Hypotheses:

2. Suggest relationship between specific exposure and health-related event…

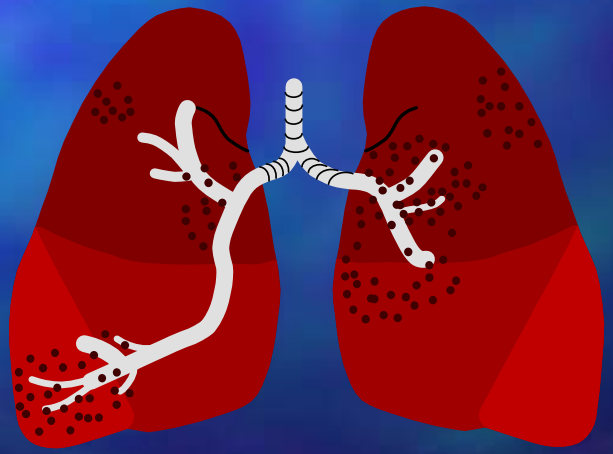
A high cholesterol intake is associated with the development (risk) of coronary heart disease.

Hypothesis Framing

Ways to Express Hypotheses:

3. Suggest cause-effect relationship….

Cigarette smoking is a cause of lung cancer



Hypothesis Framing

Ways to Express Hypotheses:

4. "One-sided" vs. "Two-sided"

One-sided example:

***Helicobacter pylori* infection is associatedwith increased risk of stomach ulcer**

Two-sided example:

Weight-lifting is associated with risk of lower back injury

Hypothesis Framing

Guidelines for Framing Hypotheses:

1. **State the exposure to be measured as specifically as possible.**

2. **State the health outcome as specifically as possible.**

Strive to explain the smallest amount of ignorance

Hypothesis Framing

Example Hypotheses:

POOR

Eating junk food is associated with the development of cancer.

GOOD

The human papilloma virus (HPV) subtype 16 is associated with the development of cervical cancer.

The Next Step

- **Formally test the identified hypotheses in a research study**
- **The study should follow a specific plan or protocol (the study design)**
- **Study designs direct how the investigation is conducted and allows for the translation of a conceptual hypothesis into an operational one**

Symbols used in Hypotheses

- ***M*= mean**
- **μ (mu: mew)= *population mean***
- **Roman Letters (e.g., A, B, C, D) are used to represent statistics**
- **Greek Letters (e.g., α, β) are used to represent parameters**
- **α= significance level; probability of committing a Type I Error (α= .05)**
- ***p*= probability value (*p*= .05)**
- **Null Hypothesis= (H_0: $\mu_1 - \mu_2 = 0$ or H_0: $\mu_1 = \mu_2$)**
- **Alternative Hypothesis= (H_1: $\mu_1 - \mu_2 \neq 0$ or H_1: $\mu_1 \neq \mu_2$)**
  - **Sometimes you may see it noted as H_A**

Types of Hypotheses

- **Research Hypotheses**
- **Statistical Hypotheses**

Research Hypotheses

- **Research Hypothesis: A statement of the relationship among two or more variables or groups.**
- **The acceptance or non-acceptance of which is based on resolving a logical alternative with a null hypothesis.**

Example: Graduate students who read the text in research methods will score higher on their comprehensive exams than graduate students who did not read their research methods text.

Research Hypotheses Cont.

- **Research hypotheses can be stated as *Directional* or *Non-directional*.**
- **Directional hypotheses predict the specific relationship among two or more variables or groups:**
  - **Graduate students who read the text in research methods will score higher on their comprehensive exams than graduate students who did not read their research methods text.**

$H_0: \mu_1 \leq \mu_2$ $H_1: \mu_1 > \mu_2$

Research Hypotheses Cont.

- **Non-Directional Hypotheses predict that there will be differences among two or more groups, but *do not specify the direction of the differences***
  - **Men and Women will differ in their recall of phone numbers**
  - **The scores on the Geriatric Depression Scale will differ between people with Stroke and people with Alzheimer's disease**
  - **IQ scores will correlate with Self Esteem scores**

$$H_0: \mu_1 = \mu_2 \qquad H_1: \mu_1 \neq \mu_2$$

Statistical Hypotheses

- **Statistical Hypotheses are mathematical, or logical statements that help researchers interpret the results of research**

- **Statistical hypotheses consist of the *Null Hypothesis* (H_0), the hypothesis of no difference and the *Alternative Hypothesis* (H_1 or H_A) which is similar in form to the research hypothesis.**

  - **Null: (H_0: $\mu 1 - \mu 2 = 0$)**
  - **Alternative: (H_1: $\mu 1 - \mu 2 \neq 0$)**

Statistical Hypotheses Cont.

- **Remember, *and this is important*:**

  The null hypothesis always implies that there is no relation or statistical difference between variables or groups

  - **The alternative hypothesis implies that there is a meaningful relationship among variables or groups**

Type of studies

- **Superiority studies**
  - **A is better than B (A = active and B = placebo or gold-standard)**
  - **Conventional one-sided hypothesis test**
- **Equivalence studies**
  - **A is more or less like B (A = active and B = standard)**
  - **Two-sided interval hypothesis**
- **Non-inferiority studies**
  - **A is not worse than B (A = active and B = standard with adverse effects)**
  - **One-sided interval hypothesis**

Equivalence

- **We are interested in verifying (instead of rejecting) the null hypothesis of a conventional hypothesis test**
- **We have to redefine the alternative hypothesis as a range of values with an equivalent effect**
- **The differences within this range are considered clinically irrelevant**
- **Problem: it is very difficult to define the maximum difference without clinical relevance for the Cmax and AUC of each drug**
- **Solution: 20% based on a survey among physicians**

Estimation with confidence intervals in a superiority trial

It is not statistically significant!

Because the CI includes the d=0 value

Confidence interval 90% - 95%

d < 0
Negative effect

d = 0
No difference

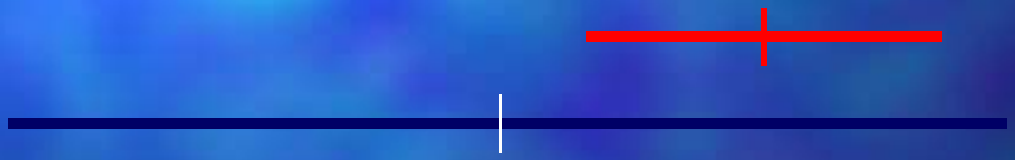
d > 0
Positive effect

Estimation with confidence intervals in a superiority trial

It is statistically significant!

Because the CI does not includes the d=0 value

Confidence interval 90% - 95%



d < 0
Negative effect

d = 0
No difference

d > 0
Positive effect

Estimation with confidence intervals in a superiority trial

It is statistically significant with P=0.05

Because the boundary of the CI touches the d=0 value

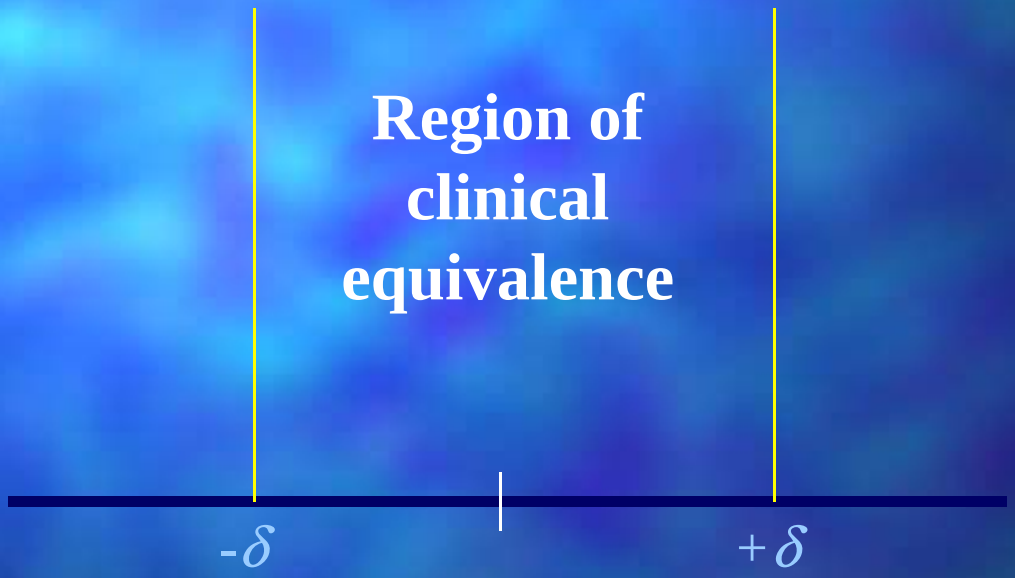
Confidence interval 90% - 95%

d < 0
Negative effect

d = 0
No difference

d > 0
Positive effect

Equivalence study

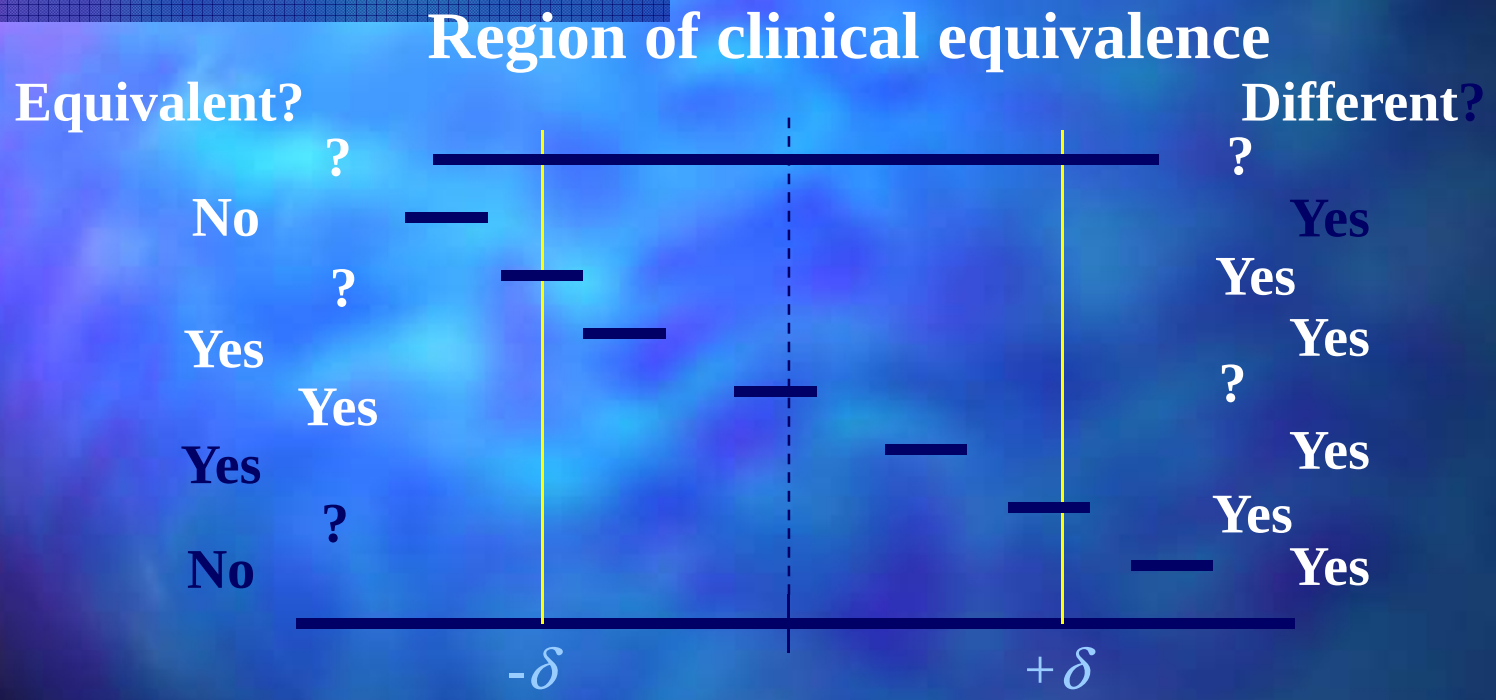


d < 0
Negative effect

d = 0
No difference

d > 0
Positive effect

Equivalence vs. difference

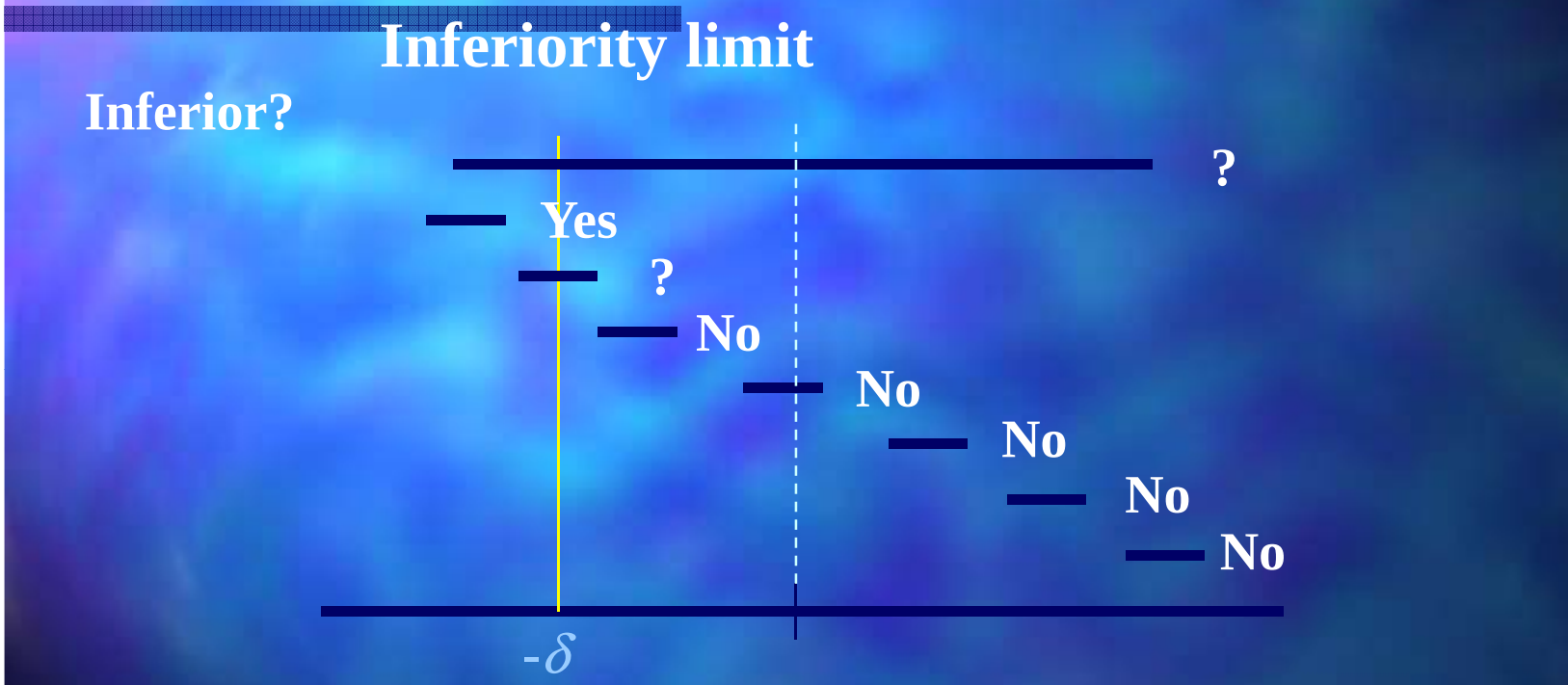


d < 0
Negative effect

d = 0
No difference

d > 0
Positive effect

Non-inferiority study

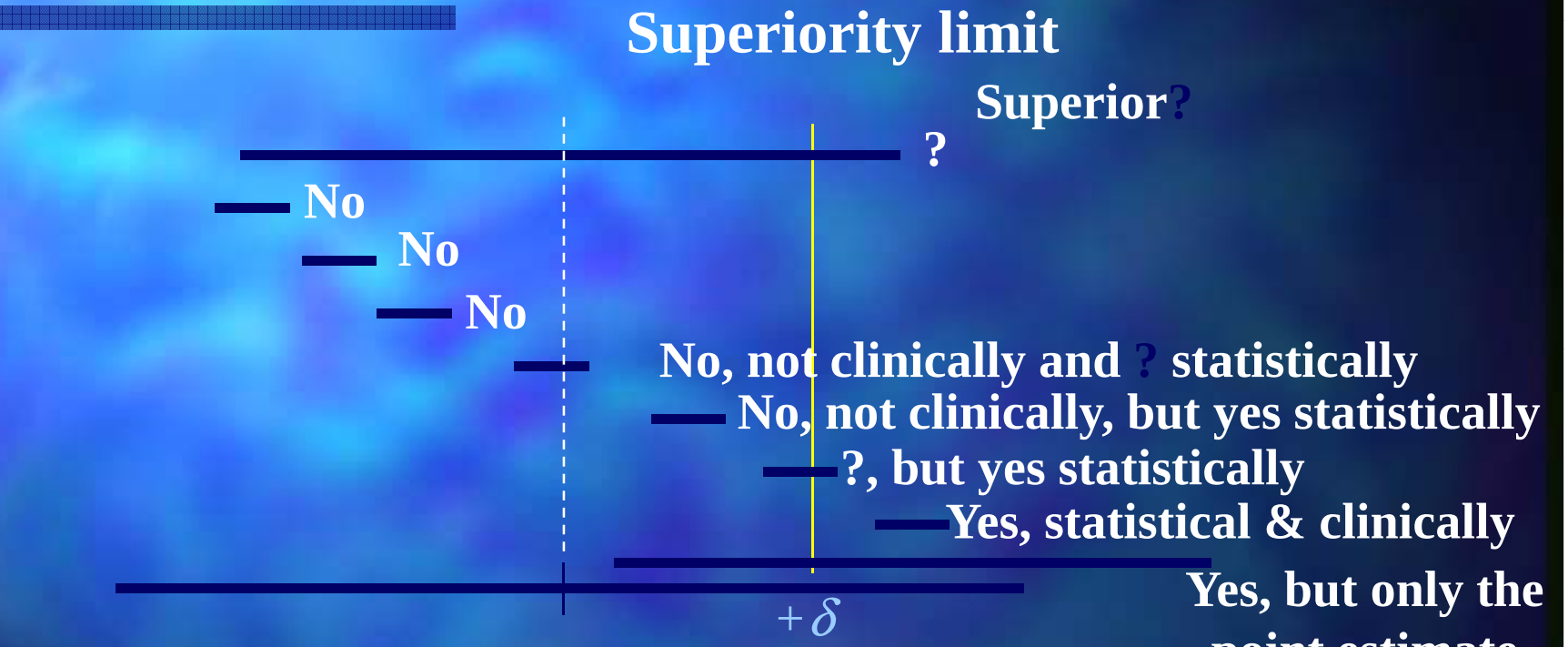


d < 0
Negative effect

d = 0
No difference

d > 0
Positive effect

Superiority study (?)



Testing Hypotheses

- **Is it possible that our observations occurred by chance?**

Testing Hypotheses Cont.

- **We use a variety of statistical procedures to test null hypotheses. The choice of which procedure we use depends on a variety of factors including:**
  - **the research hypothesis,**
  - **the data,**
  - **the sampling strategy,**
  - **and what we want to be able to say as a result of our testing.**

Types of Tests

- **Statistical procedures that are commonly used for hypothesis testing include: correlation, analysis of variance (ANOVA), analysis of covariance (ANCOVA), regression, multivariate analysis of variance (MANOVA), t-tests, and Chi-Square. Each of these procedures has an associated test statistic, which is used to determine significance. For example ANOVA, ANCOVA, and regression use F statistics and their associated p-values.**
- **Multivariate procedures, like MANOVA, use a variety of test statistics with interesting names, like Wilk's lambda. These are then related to a more common test statistic, like F.**
- **The secret here, for the layperson, is that all test statistics are eventually related to a probability distribution and a p-value. These p-values mean the same thing across test statistics.**

Error Types

- **In hypothesis testing, we must contend with two types of errors -- Type I and Type II.**
  - **Errors are mistakes that we can make when judging the null hypothesis**
- **Type I error is what happens when the tested hypothesis is falsely rejected. (It is when you say you found something, but that something is really an error.) A type I error is a false positive.**
- **Type II error is what happens when a false tested hypothesis is not rejected (Hays, 1986). (It is when you don't find something that is, in fact, there.) A type II error is a false negative.**

Error Types Cont.

- **Alpha is the level of probability (pre-set by the researcher) that the tested hypothesis will be falsely rejected. Alpha is the pre-set risk of a Type I error. In other words, alpha is the degree of risk that you accept, in advance of conducting the study, that what you find will be an error.**
- **Beta is the probability (often neglected by the researcher) that a false null hypothesis will not be rejected. Beta is the probability that you won't find what you are looking for if, in fact, it is really there.**

Error Types Cont.

- Error Types Chart

Decision		H_0 is True	H_1 is True
	Reject H_0	Type I α	Correct 1- β
	Fail to Reject (decide in favor of H_0)	Correct 1- α	Type II β

Example

Do we use Null Hypotheses in the real world?

"Innocent until Proven Guilty"

	Defendant Innocent	Defendant Guilty
Reject Presumption of Innocence (Guilty Verdict)	Type I Error	Correct
Fail to Reject Presumption of Innocence (Not Guilty Verdict)	Correct	Type II Error

Steps in Hypothesis Testing for Quantitative Research Designs

- Hypothesis testing is a 4 phase procedure:
  - **Phase I: Research Hypotheses, Design, and Variables**
  - **Phase II: Statistical Hypotheses**
  - **Phase III: Hypotheses Testing**
  - **Phase IV: Decision/Interpretation**

Questions about Hypotheses

?