

HYPOTHESIS TESTING

Psychology 3000

Tom Malloy

THE PLAUSIBLE COMPETING HYPOTHESIS OF CHANCE

Scientific hypothesis: The effects found in the study are due to...

Scientific skepticism: There are no effects ...

Example: Suppose you develop a pill that should lower blood pressure

Sci Hyp: Pill (IV) lowers Blood Pressure (DV)

Sci Skepticism: Pill has no effect on BP

DESIGN OF STUDY:

Random assignment to two groups:

Experimental group: Give Pill for 3 months then measure BP

Control group: Give nothing for 3 months then measure BP

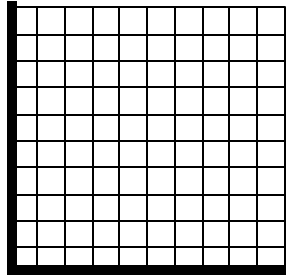
The **IV** is...

The **DV** is...

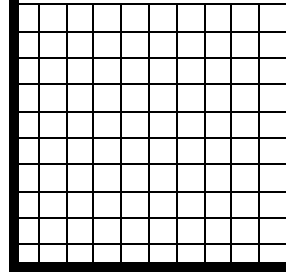
PATTERN OF RESULTS:

You **MUST** get a pattern of results that favor the scientific hyp...

Example if you don't:



Example if you do:



If you do get a pattern that supports the ...

Then Scientific Skepticism invents:

PLAUSIBLE COMPETING HYPOTHESES (PCH's):

Must be another way to explain...

Must be plausible within...

Example of a PCH: Placebo

CHANCE--Another Plausible Competing Hypothesis (PCH)

PCH OF CHANCE: The pattern of results is due to

Example: Divide any classroom...

Statistical Conclusion Validity is the strength of our argument that the pattern of results in the experiment was not due to chance.

CAVEAT:

Just because we eliminate Chance as a PCH does not mean we have eliminated other PCH's....

In the study we described, demonstrated that the average BP in the two groups was unlikely to be due to chance doesn't eliminate the Placebo PCH

Jargon: A statistically significant result is one....

A reliable result is one...

NULL AND ALTERNATIVE STATISTICAL HYPOTHESES

Simple example of Statistical Conclusion Validity logic: Flipping a single coin

FIRST, CONSIDER AN EXAMPLE AT THE LEVEL OF SCIENCE:

Say that you go down to New Orleans, down to the French quarter, down onto Bourbon Street, downstairs into a somewhat less than legal gambling house....
(You're pretty sure the bouncers have guns.)

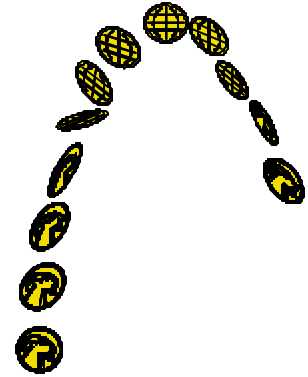
Sci Hyp:

Your friend's cautious skepticism:

So you collect data: Watch the coin flipped once...

DATA:

PCH of Chance:



NOW, CONSIDER THE EXAMPLE AT THE LEVEL OF STATISTICS:

Translate Scientific hypotheses into statistical hypotheses

PCH Chance ! H_0 :

Sci Hyp ! H_1 :

Continue the example:

More DATA (Flip the coin a second time):

PCH of Chance !

Sci Hyp !

More DATA (flip the coin a third time):

H_0 :

H_1 :

More DATA (flip the coin a fourth time):

And so on...

AT SOME POINT YOU REJECT THE IDEA THAT THIS DATA PATTERN IS DUE TO CHANCE

DATA:

HHHHHHHHHH
.....

Reject H_0

There is always some probability that you are...

Alpha level " =

Significance level =

p value =

STATISTICAL CONCLUSION VALIDITY

The validity with which you can make the argument that...

THE LOGIC OF STATISTICAL CONCLUSION VALIDITY

VACCINE EXAMPLE

Say you have developed a cold vaccine which you think works. You are ready to do some research to find out if it really does work.

Sci Hyp:

Skeptical Hypothesis:

If your research were to yield a Pattern of Results that support...
The skeptic replies with

PCH Chance:

DV & IV

The DV (what we measure about each participant) is whether or not the person gets a cold. We give the participant a 0 if no cold is caught, a 1 if a cold is caught

$$DV = X = \{$$

The IV is the cold vaccine

DIRECTIONAL SCIENTIFIC HYPOTHESES

In this case it is a Directional Sci Hyp:

A scientific hypothesis is directional if...

VERSUS

NON-DIRECTIONAL SCIENTIFIC HYPOTHESES

Back to the example:

Suppose through lots of data you know that the Prob of a cold during January and February is .5 (or $\frac{1}{2}$). $P(\text{Cold}) = .5$

You take $n = 10$ **volunteers** and inject them with the vaccine at the end of December. Then you watch them and observe which subjects get at least one cold during Jan. and Feb.

NOW LET'S MOVE from the Realm of Science to the Realm of Statistics

Translate the Sci hyp and PCH of chance into statistical hyp's

Realm of Science

Realm of Statistics

Sci Hyp goes with

H₁:

PCH of chance goes with

H₀:

These are more precise ways of stating the Sci Hyp and the PCH of Chance

AT THIS POINT WE HAVE TWO STATISTICAL HYPOTHESES.

NOW WE HAVE TO CHOOSE BETWEEN THEM.

A "**TEST STATISTIC**" WILL BE OUR DECISION MAKER.

We will learn about many test statistics: **t, F, and Chi-square**

For now we will examine the probability theory supporting these stats..

In this example we will use a simple, common sense test statistic:

The number of subjects who get "at least one cold" be our **test statistic**. This test statistic is easy to calculate. We already have 1 or a 0 for each subject, depending on whether the subject caught a cold or not. So all we have to do is add up the measures we took as a DV.

$$DV = X = \{$$

$$TS = \text{sum of DV measures} = \sum X$$

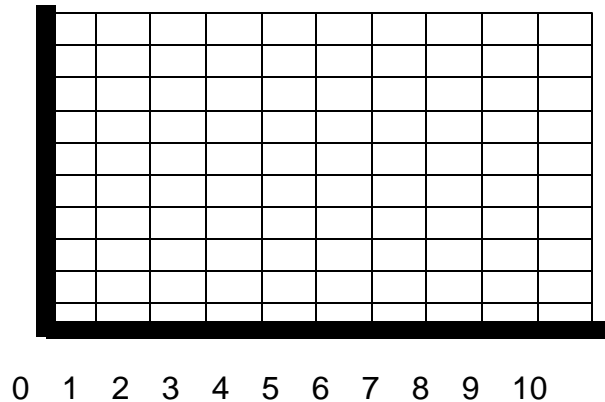
Range of TS:

WE DIVIDE THE RANGE OF THE TS INTO TWO KINDS OF "REGIONS"

Reject H_0 region

Do not reject H_0 region

Range of TS:



But where should we put the critical value?

Between 2 and 3?

Between 4 and 5?

Between

Between

We need probability theory to...

One vs two-tailed tests:

Binomial Sampling Distribution

1

4

2

3

If Sci hyp is....

If Sci hyp is...

If Sci hyp is...

HOW DO WE SET THE CRITICAL VALUES THAT DEFINE REJECTION REGIONS?

We assume H_0 is true

Prob (cold) = 1/2....like a fair coin (Bernoulli Trial)

Then Use Binomial Distribution to find Probabilities such as alpha

If H_0 is true,

the $P(\text{Cold}) = .5$, (and a Cold is defined as a "success," r)
and we have sampled 10 tries, then...

Binomial Table or StatCenter Binomial Tool gives us:

$P(r \text{ given } p = .5, N = 10)$

$P(0 \text{ C}) = .0010$

$P(1 \text{ C}) = .0098$

$P(2 \text{ C}) = .0439$

$P(3 \text{ C}) = .1172$

$P(4 \text{ C}) = .2051$

$P(5 \text{ C}) = .2461$

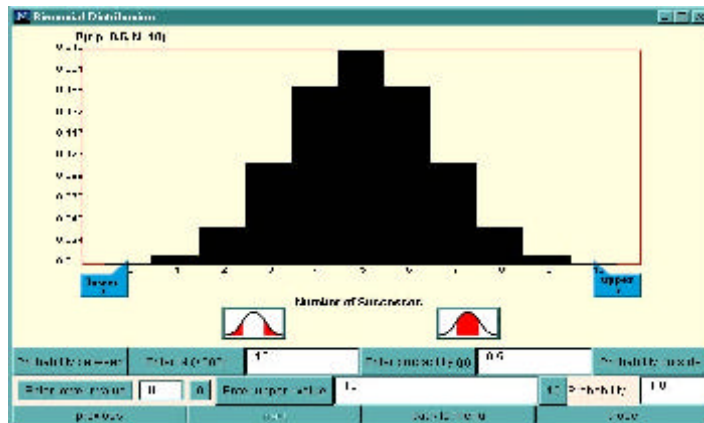
$P(6 \text{ C}) = .2051$

$P(7 \text{ C}) = .1172$

$P(8 \text{ C}) = .0439$

$P(9 \text{ C}) = .0098$

$P(10 \text{ C}) = .0010$

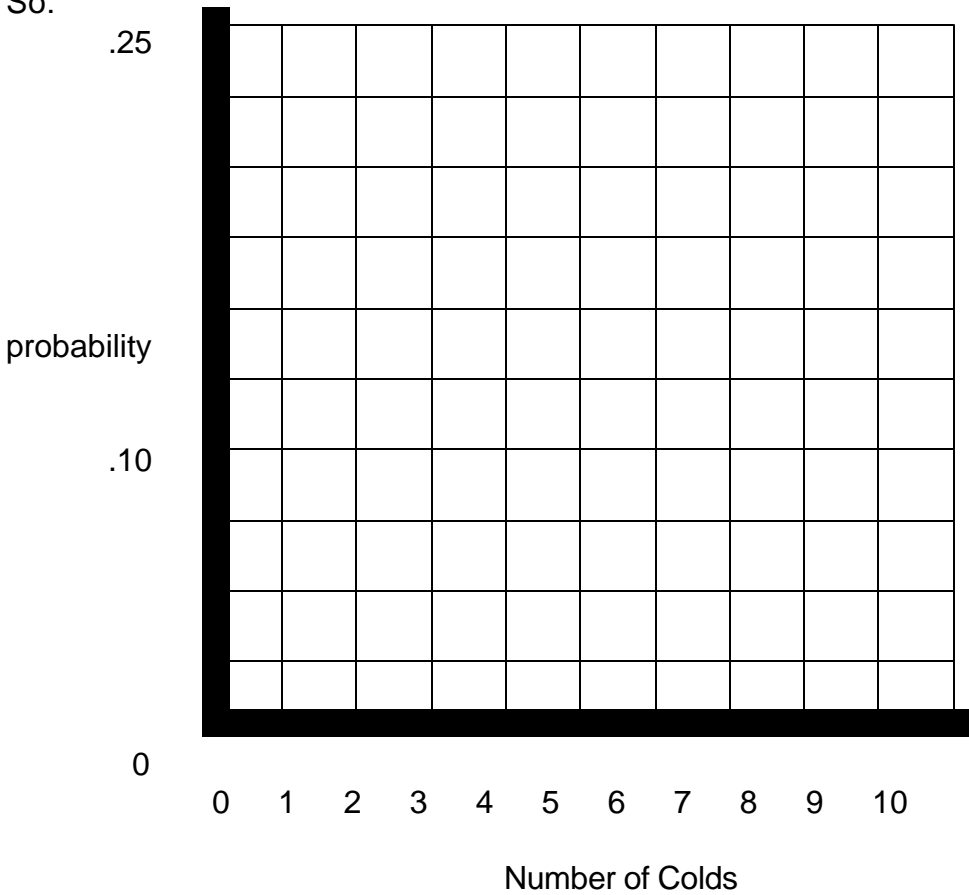


Probability Distribution Assuming H_0 is TRUE

What do the probabilities look like if we set the Critical Value at:

Critical Value set between 2 and 3:

So:

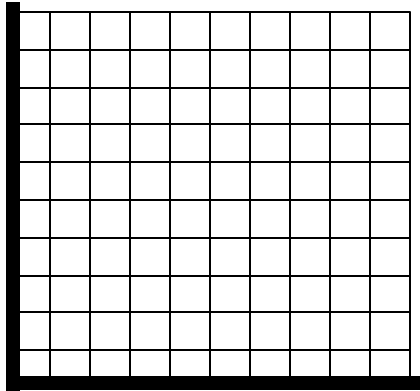


If H_0 is true

What is α , the probability of falling in the rejection region by chance alone?

This assumes Skeptic is right: Vaccine has NO EFFECT
Volunteers are getting a chance number of colds

Set

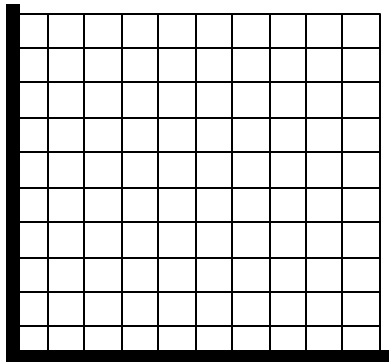


critical value between 1 and 2:

If H_0 is true

What is α , probability of falling in the rejection region by chance alone?

Set the critical value between 0 and 1:



If H_0 is true

What is α , probability of falling in the rejection region by chance alone?

To calculate α :

Assume H_0 is ...

Set location of...

Set n

Statistical Conclusion Validity

THREATS TO STATISTICAL CONCLUSION VALIDITY

1. Low power

small n

alpha too small

sloppy research, particularly unreliable DV

2. Violated assumptions of test

wrong t-test

non-independent observations of chi-square

etc.

3. Fishing for results and multiple t's

4. Reliability of measures.

5. Reliability of treatment implementation

6. Random irrelevancies in the experimental setting

7. Random heterogeneity of respondents.