Probability and Mathematical Statistics in Data Science

Lecture 23: Section 9.1: Testing Hypotheses

Hypothesis Testing

- A researcher develops a theory (through observation or previous exploratory analysis) about how the world works.
- They wish to formally test their theory through the collection and analysis of sample data
- Oftentimes, they may be challenging currently accepted knowledge (about how the world works) or have developed a new treatment for a condition or disease they want to compare to existing treatments and/or a placebo control
- They can conduct a hypothesis test to test their theory

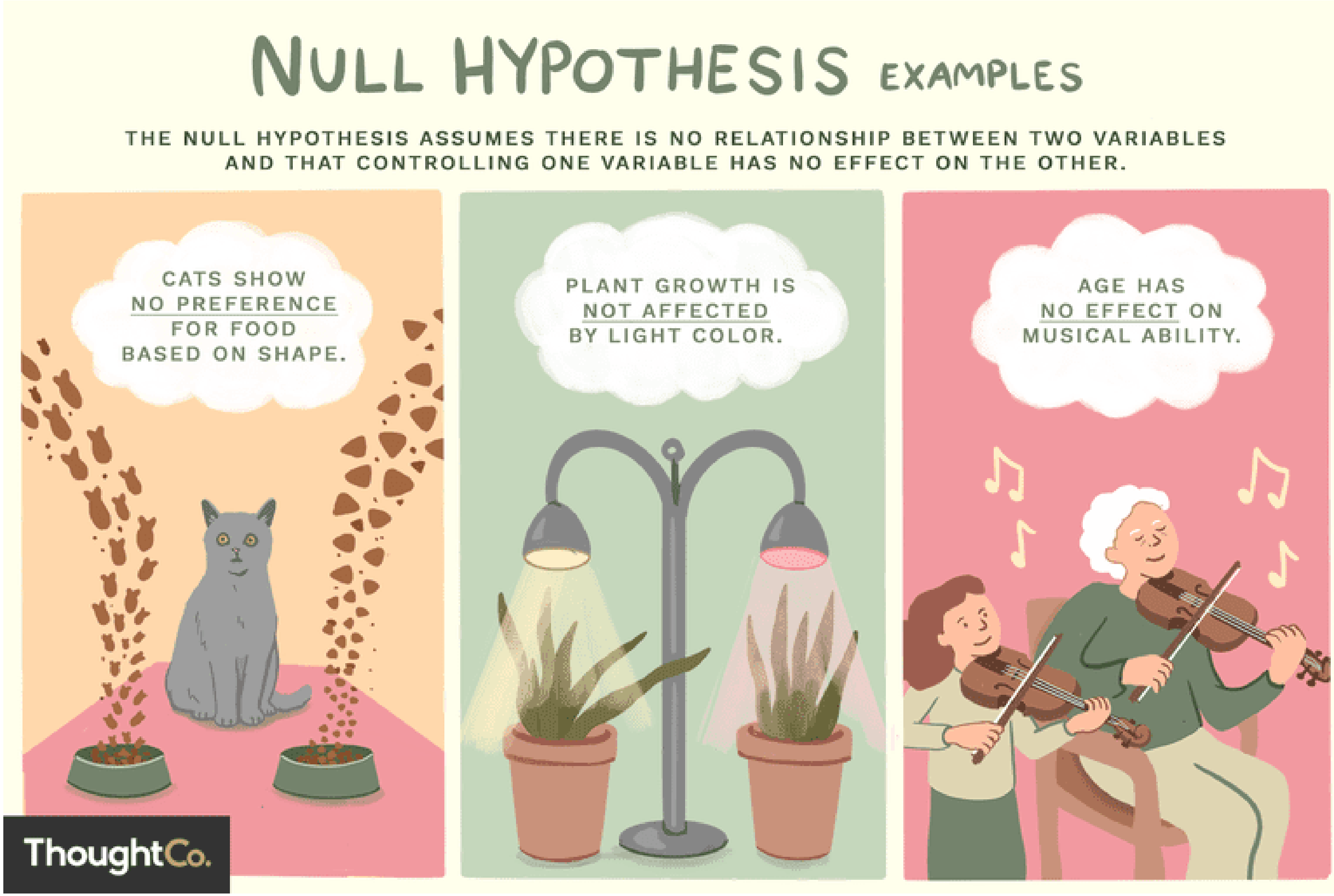
Null Hypothesis

- Hypothesis testing is a much used approach to making statistical decisions about the value of population parameters

- Hypothesis testing begins from the point of view of what is known as the **null hypothesis**. It could be a statement of the currently accepted belief about the value of a population parameter – say population mean height of men is equal to 70 inches

- When comparing treatments, it could be a statement that the both treatments are equally effective on average.

Hypothesis Testing

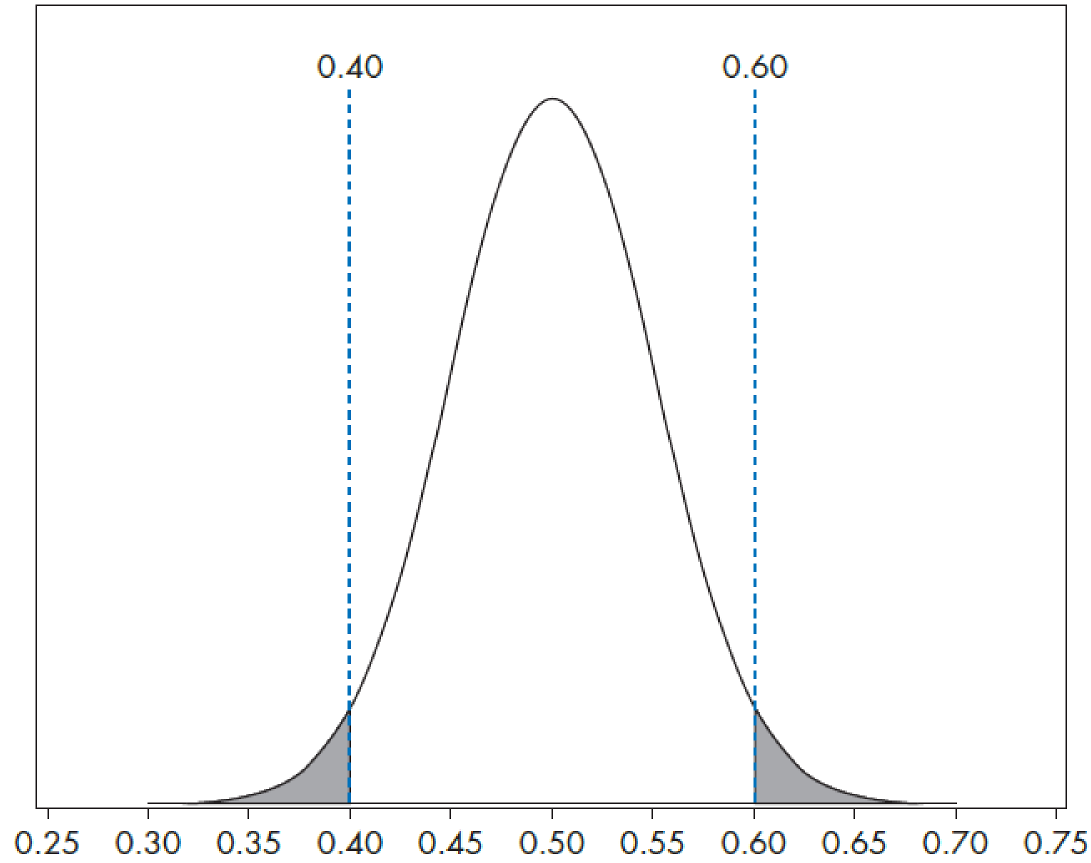


Courtroom Jury Trial Analogy

- **Null Hypothesis**: *Defendant is innocent*
- **Alternative Hypothesis**: *Defendant is guilty*

source: www.defenselitigationinsider.com.com

Hypothesis Testing: Tossing Coin 100 Times



- **Null Hypothesis**: *Coin is fair*
- **Alternative Hypothesis**: *Coin is not fair*

Examples

Ex. A factory claims no more than 10% of the components they produce are defective. A consumer group is skeptical of the claim and checks a random sample of 300 components and finds that 39 are defective. Is there evidence that more than 10% of all components made at the factory are defective?

$H_0 : p = 0.10 \quad H_a : p > 0.10$

Ex. We are interested in height of all Berkeley students. In a sample of 12 students, the sample mean is 66.30 inches, and the sample s.d. is 4.35 inches. Should we reject the null hypothesis?

$H_0 : \mu = 68$ vs $H_a : \mu \neq 68$?

The Four Steps of Hypothesis Testing

- There are four steps to a hypothesis test:

1. **Hypotheses**
2. **The Model**
3. **Calculations**
4. **Conclusion**

Hypotheses tests: Review of steps

1. State the *null* hypothesis – that is, what is the assumption we are going to make. This will determine the distribution that we will use to compute probabilities (called the *null* distribution or null model).

2. State an appropriate *alternative* hypothesis. Note that this should not overlap with the null hypothesis. You should state both the hypotheses in informal terms and in terms of random variables.

3. Decide on a test statistic to use that will help you decide which of the two hypotheses is supported by the evidence (data). Usually there is a natural choice. The null hypothesis will specify the distribution of the test statistic.

4. Find the observed value of the test statistic, and see if it is consistent with the null hypothesis. That is, compute the chance that we would see such an observed value, or more extreme values of the statistic **(*P-value*)**

5. If this probability is too small, then we reject the null hypothesis

Vocabulary review

- null hypothesis (H_0)

- alternative hypothesis (H_1)

- *P-value or observed significance level*
  - P-value is **not** the chance of null being true. The null is either true or not.
  - The P-value is a *conditional probability* since it is computed *assuming* that the null hypothesis is true.
  - The smaller the P-value, the stronger the evidence ***against*** the null and ***towards*** the alternative (in the direction of the alternative)
  - "Small" is for you to decide. Traditionally, below 5% ("result is statistically significant") and 1% ("result is highly significant") are what have been used. Significant means the p-value is small, not that the result is important.

Example: The pill

A pharmaceutical company advertises for their birth control pill has an efficacy of 99.5% in preventing pregnancy. However, under typical use, the real efficacy is only about 95%. That is, 5% of the women taking that pill for a year will experience an unplanned pregnancy that year. A gynecologist looks back at a random sample of 200 medical records from patients who have been prescribed this pill one year before.

She finds that 14 women had become pregnant within 1 year while taking the pill. Is this surprising?

Example 9.5.1

- All the patients at a doctor's office come in annually for a check-up when they are not ill. The temperatures of the patients at these check-ups are independent and identically distributed with unknown mean μ.

- The temperatures recorded in 100 check-ups have an average of 98.2 degrees and an SD of 1.5 degrees. Do these data support the hypothesis that the unknown mean μ is 98.6 degrees, commonly known as "normal" body temperature? Or do they indicate that μ is less than 98.6 degrees?

Large-sample tests for proportions

Null hypothesis: H_0: $p = p_0$

Test statistic value: $z = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$

Alternative Hypothesis	P-Value Determination
H_a: $p > p_0$	Area under the standard normal curve to the right of z
H_a: $p < p_0$	Area under the standard normal curve to the left of z
H_a: $p \neq p_0$	2·(Area under the standard normal curve to the right of $\lvert z \rvert$)

These test procedures are valid provided that $np_0 \geq 10$ and $n(1 - p_0) \geq 10$.

They are referred to as *upper-tailed*, *lower-tailed*, and *two-tailed*, respectively, for the three different alternative hypotheses.

Example

- **Abnormalities** In the 1980s, it was generally believed that congenital abnormalities affected about 5% of the nation's children. Some people believe that the increase in the number of chemicals in the environment has led to an increase in the incidence of abnormalities. A recent study examined 384 children and found that 46 of them showed signs of an abnormality. Is this strong evidence that the risk has increased?

  a) Write appropriate hypotheses.

  b) Perform the mechanics of the test. What is the P-value?

  c) Explain carefully what the P-value means in context.

  d) What's your conclusion?

-