Hypothesis testing; Type I and Type II error

1 What is a hypothesis?

A hypothesis can be defined as a proposed explanation for a phenomenon. It is not the absolute truth, but a provisional, working assumption, perhaps best thought of as a potential suspect in a criminal case.

2 Null and alternate hypothesis in a non-"mathy" way

In scientific research, we often start with two competing claims: the *null hypothesis* and the *alternative hypothesis*. Let's break these down in simple terms.

The null hypothesis is what we are willing to assume is the case until proven otherwise. It is relentlessly negative, denying all progress and change. But this does not mean that we actually believe the null hypothesis is literally true. The alternative hypothesis is the suspected claim that contradicts or attempts to replace the null hypothesis.

Example: There is a strong analogy with criminal trials in the legal system. It is said that 'innocent until proven guilty.' Here, the null hypothesis is the person is not guilty and alternate hypothesis is that the person is guilty. Unless there are compelling arguments to refute the null hypothesis, court-of-law believes that the person is innocent.

Example: Assume you are in the forest and you hear a rustling in the bushes. Now you have to decide whether you think there is a tiger behind the bushes or not. In this case, the null hypothesis is that there is no tiger behind the bushes. The alternate hypothesis is that there is a tiger behind the bushes.

3 Null and alternate hypothesis in a "mathy" way

1. Null Hypothesis (H_0) - Currently accepted or default assumption
2. Alternate Hypothesis (H_A) - The suspected claim that contradicts or attempts to replace the null hypothesis.

Example: Data from the Center for Disease Control estimates that about 30% of American teenagers were overweight in 2008. A professor in public health at a major university wants to determine whether the proportion has changed since 2008. Write down the null and alternate hypothesis.

$H_0 : p = 0.3$

$H_a : p \neq 0.3$

Example: An article recently published in a local newspaper states that 80% of all human population is right handed.You are skeptical of it and want to determine whether the proportion ios less than stated. Write down the null and alternate hypothesis.

$H_0 : p = 0.9$

$H_a : p < 0.9$

4 Hypothesis testing

To actually test a null hypothesis and establish evidence against it, we can perform what is known as a 'Test of Significance.' We do not ever try to "prove" an alternative hypothesis in hypothesis testing. We are trying to refute the null hypothesis; therefore we have two possible outcomes:

1. Reject the null hypothesis
2. Fail to reject the null hypothesis

Example: Height and Starting Salaries. Suppose your best friend believes there is a correlation between one's Height and their starting salary in their first full-time job. You, being a Statistics student, are skeptical of the claim and wish to test this claim.

H_0 : There is some correlation between Height and Starting Salary (claim)

H_a : There is no correlation between Height and Starting Salary (suspicion)

Once we've performed a Significance Test, depending on the result we may:

- Reject the claim that there is a correlation between height and starting salary
- Fail to reject the claim that there is a correlation between height and starting salary

Note: We never try to prove or establish either claim. Tests of significance are inherently tests of refutation.

Example: Suppose you and your friend like to flip a coin to decide who pays when you go out to eat. Your friend has a "lucky" coin that they love to use. If the coin lands on heads, they pay. If the coin lands on tails, you pay. The last 7 times you've gone out to eat, the coin has come up as tails every time. Your friend claims this is a perfectly fair coin, but you have your suspicions.

a) Identify H_0 and H_a :

H_0 : Your friend's cain is not biased (claim)

H_a : Your friend's coin is bias (suspicion)

b) Suppose a significance test finds that the past 7 outcomes being tails is too unlikely to occur purely by chance (by comparing its likelihood to a judgment threshold i.e. significance level). What, then, should the conclusion of your test be?

We reject the null hypothesis that the coin is unbiased.

5 What if we are wrong?

A Hypothesis Test or Significance Test is ultimately founded upon likelihoods, not definitive answers about reality. Because of that, there is a small chance of arriving at the incorrect conclusion.

Example: Consider a patient is undergoing routine screening for cancer. In reality, the person may or may not have cancer. The screening test may indicate a positive result (presence of cancer) or it may indicate a negative result (absence of cancer).

a) Identify H_0 and H_a :

H_0 : The patient does not have cancer (default assumption)

H_a : The patient does have cancer (potential suspicion)

b) Identify all possible outcomes of the test:

Belief / Reality	Negative Test (H_0)	Positive Test (H_a)
Doesn't Have Cancer	True Negative	False Positive
Has Cancer	False Negative	True Positive

6 Type I and type II error

Type I error is made when we reject a null hypothesis when it is true, and a Type II error is made when we do not reject a null hypothesis when in fact the alternative hypothesis holds.

False Positives are known as Type I errors and false negatives are known as Type II errors.

In this course, we will study Type I errors and Type II errors in terms of false positives and false negatives.

Example: Suppose your spouse has been behaving suspiciously as of late. As a statistician, naturally, you decide to perform a hypothesis test.

H_0 : Your spouse is not cheating.

H_A : Your spouse is cheating.

Draw the two way table (confusion matrix) and identify the Type 1 and Type 2 error scenarios.

Belief / Reality	Not cheating (H_0)	Cheating (H_a)
Not cheating	True Negative	False Positive (Type I error)
Cheating	False Negative (type II error)	True Positive

Example: Let's finish the chapter with the example we started our discussion with. You hear a rustling in the bushes and you have to make a decision whether to run or not. You believe there is no tiger behind the bushes, but your friend believes that there is a tiger behind the bushes.

H_0 : There is no tiger (claim).

H_A : There is a tiger (suspicion).

Draw the two way table (confusion matrix) and identify the Type 1 and Type 2 error scenarios.

Belief / Reality	No tiger (H_0)	Tiger (H_a)
No tiger	True Negative	False Positive (Type I error)
Tiger	False Negative (type II error)	True Positive

7 Extra stuff. Accuracy and Precision

Precision and accuracy are two important concepts in the field of measurement and are often used interchangeably, but they do have different meanings. Precision is a measure of how consistent the results of a measurement are when the measurement is repeated multiple times. Accuracy, on the other hand, refers to the degree to which a measurement reflects the true value of the quantity being measured.

We can relate accuracy and precision to Type I error and Type II error as well.

Recall,

		Actual	
		Positive	*Negative*
Predicted	*Positive*	*True Positive (TP)*	*False Positive (FP) (Type I Error)*
	Negative	*False Negative (FN) (Type II error)*	*True Negative (TN)*

Accuracy is calculated as:

$$\frac{TP + FP}{TP + FP + FN + FP}$$

Similarly, precision is calculated as:

$$\frac{TP}{TP + FP}$$