Steps of hypothesis testing

State the hypotheses: First one has to state the alternative and the null hypothesis. These hypotheses shout be about a population (usually a parameter), should be exclusive (if one of the hypothesis is true the other should be false and vice-versa) and exhaustive (one of the hypothesis must be true). What we believe unless we have strong evidence against it goes into the null hypothesis (H_0) and what we will not believe unless we have strong evidence to support it goes into the alternative hypothesis (H_1). But there is a limitation to this: the null hypothesis should include equality. As we showed in class, if the null hypothesis does not include equality, then we can not conduct the hypotheses test. Thus if we strongly believe that a coin is not fair but would like to test this, then we have to take $\pi \neq 0.5$ as the null hypothesis and $\pi = 0.5$ as the alternative (where π is the probability of observing a head in a single toss). But in class we showed that we can not conduct this test.

Decide on the significance level: We denote the significance level with α. The significance level represents what we mean by "almost impossible".

Test Statistics: Design an experiment (this usually consists of taking an random sample for the population of interest) and choose a test statistic (this is a measure which summarizes the sample).

State the distribution of the test statistic: In order to decide which values (of the test statistic) are unlikely to occur, we have to know the distribution of the test statistic. This depends on the test statistic and the assumptions about the population stated in the null hypothesis.

State the decision rule and find the critical value(s): The critical values, are values such that the event that test statistic takes values as "extreme" as these is almost impossible (i.e., has a probability less than or equal to α) if H_0 is true. Hence, the critical value/s is/are value/s of the test statistic for which, the test statistic taking any value at least as "extreme" as this/these value/s is more in accord with H_1 than with H_0. The word "extreme" depends on the null hypothesis, the distribution of the test statistic and what we mean by almost impossible (the significance level).

The decision rule tells us which of the hypothesis we will accept as true based on the result of the experiment (value of the test statistics).

- If we observe a value which is "almost impossible" to observe *when* H_0 is true, then we choose H_1. In this case we say: we reject H_0.
- If we observe a value which is not "almost impossible", then we choose H_0. In this case we say: we fail to reject H_0.

Conduct the experiment and calculate the test statistic: Take a random sample and for this sample calculate the value of the test statistic.

State the conclusion: State your conclusion: do you reject of fail to reject the null hypothesis. You should do this both in a technical language (like: we reject the null hypothesis) and in a plain language (like, there is strong evidence to suggest that the mean income is ....)

Calculate the p-value: This gives us the probability of observing a value (of the test statistic) as *extreme* as what we have calculated, when the null hypothesis is true.