Hypothesis Testing Overview (Framework)

1. State the null and alternative hypotheses

   Example (coin toss – is the coin biased)

   H_0: p = 0.50
   H_A: p != 0.50

   where p is the probability of getting heads.

2. Carry out an experiment to get a sample statistic (e.g., sample proportion), whose distribution is known under H_0.

3. Convert the sample statistic to a test statistic (e.g., convert to a z score), whose distribution is also known under H_0.

4. Calculate a p-value, which is the probability that the test statistic differs from its expected value by at least as much as it does. If the observation is *Z* standard deviations from the mean, this can also be calculated as the probability of an observation at least *Z* standard deviations from the mean.

   If we flip a coin 100 times, and get 62 heads, the P-value is the probability that we observe at least 12 heads more or less than what is expected by chance (assuming the null hypothesis – that the coin is fair – is true)

   The corresponding Z-statistic is 2.4 (or 2.3 with the continuity correction).
   The p-value is the probability that a normally distributed observation is at least 2.4 standard deviations from the mean.

   So the p-value in this case is the probability that under the null hypothesis, the observed proportion of heads would be more than 2.4 standard deviations away from its expected value (which we also call the mean).

5. Make a decision regarding H_0.

   If p -value < 0.05, we would reject H_0 and conclude that the coin was biased
   If p-value > 0.05., we would fail to reject H_0 and conclude that there is not sufficient evidence that the coin is biased.

In the color-blind example, the proportion of color-blind males is about 0.098 and the proportion of color blind females is about 0.0068.

This is a difference of 0.098 – 0.0068 = 0.0912.

The p-value is the probability that the difference (in absolute value) between the sample proportions is at least 0.0912 (assuming that the null hypothesis – that there is no difference in colorblind rates between males and females – is true)

The corresponding Z-statistic is 10.9. The p-value is the probability that a normally distributed observations is at least 10.9 standard deviations from the mean.

Let's do the same thing but for a 1-sample t-test

1. State the null and alternative hypotheses

   Example (does GPA of a specific class significantly differ from 3.3, which is a B+)

   H0: mu = 3.3
   H1: mu != 3.3,

   where mu is the mean GPA.

2. Carry out an experiment to get a sample statistic (e.g., sample mean), whose distribution is known under H_0 (this would be true if the standard deviation was known).

3. Convert the sample statistic to a test statistic (e.g., convert to a *t* statistic), whose distribution is also known under H_0 (this is true using the *sample* standard deviation)

4. Calculate a p-value, which is the probability that the test statistic differs from its expected value by as much as it does (or more).

   In a sample of 19 students, the mean GPA is about 3.37, which differs from the expected mean by 3.37 – 3.30 = 0.07. The P-value is the probability that our observed GPA is 0.07 more or less than what is expected by chance (assuming the null hypothesis – that the mean GPA is 3.3 – is true)

   The corresponding *t*-statistic is 0.934. The p-value is the probability that a random sample from the *t* distribution (with 19 – 1 degrees of freedom) is at least 0.07 away from 0 (or, equivalently, that a normally distributed observation is at least 0.934 *sample* standard deviations from the expected value).

5. Make a decision regarding H_0.

   Because p-value is NOT < 0.05, we fail to reject H_0. There is not sufficient evidence that the mean GPA of this class differs from 3.3.