Lecture 12

We use hypothesis testing to test possible values for the population parameter. We will be focusing on testing for two population parameters:

population mean μ estimated by the statistic $\bar{X}$
population proportion p estimated by the statistic $\hat{p}$

Hypothesis Tests About Proportions

When our variables are categorical, the parameter of interest are usually poplation proportions in the categories. The astrology example from Lecture 11 is situation where we are testing hypotheses about the proportion of correct predictions from an astrologer.

Let's go through the steps again but in terms of proportions.

Step 1: Assumptions

- The data are obtained using randomization
- The sample size is large enough that the sampling distribution of $\hat{p}$ is approximately normal

NOTE: Remember that $\hat{p} = \frac{X}{n}$ where X is a binomial random variable representing a count. Approximate normality occurs when the expected number of successes (np) and failures ($n(1-p)$) are both at least 15 at the null hypothesis value for p.

For the astrology example, $n = 116$. Our null hypothesis is that $p = \frac{1}{3}$ so we expect $116(\frac{1}{3}) = 38.7 \approx 39$ correct guesses and $116(\frac{2}{3}) = 77.3 \approx 77$ wrong guesses if the true p for astrologers really is $\frac{1}{3}$. Both of these are greater than 15 so our normality assumption is satisfied.

As with other statistical inference methods, without randomization the validity of the results is questionable. A survey should use random sampling. An experiment should use principles of randomization and blinding with study subjects, as was done in the astrology study. In that study, the astrologers were randomly selected, but the subjects evaluated were people (mainly students) who volunteered for the study. Because of this, any inference applies to the population of all astrologers but only to the particular subjects in the study. If the study could have randomly chosen the subjects as well, then the inference would extend more broadly to all people.

Step 2: Hypotheses

The null hypothesis about a proportion always has the form

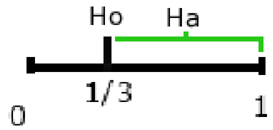
$$H_o : p = p_o \tag{1}$$

where p_o represents a particular hypothesis value for the proportion, between 0 and 1. In the astrology example, the null hypothesis of no effect states that the astrologers' predictions correspond to random guessing so we write $H_o : p = \frac{1}{3}$.

The alternative hypothesis refers to the alternative parameter value that are not the null hypothesized value. There are two possible forms of an alternative hypothesis.

One-sided alternative hypothesis:

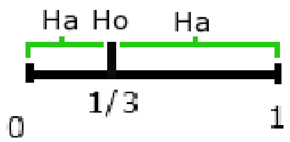
$$H_a : p > p_o \tag{2}$$



This is used when a test is designed to detect whether p is larger than the number in the null hypothesis. Another possible one-sided alternative hypothesis is $H_a : p < p_o$.

Two-sided alternative hypothesis:

$$H_a : p \neq p_o \tag{3}$$



This includes all the other possible values, both below and above the value p_o in H_o. It states that the population proportion differs from the number in the null hypothesis.

Step 3: Test statistic

The test statistic measures how far the sample proportion falls from the null hypothesize value p_o, relative to what we'd expect if H_o were true. Remember that the sample distribution of the sample proportion, $\hat{p}$, has mean equal to the population proportion p and standard deviation equal to $\sqrt{\frac{p(1-p)}{n}}$.

When H_o is true $p = p_o$, so the sampling distribution of $\hat{p}$ has mean p_o and standard deviation $\sqrt{\frac{p_o(1-p_o)}{n}}$.

Our test statistic is

$$Z = \frac{\hat{p} - p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}} = \frac{\textit{sample proportion} - \textit{null hypothesis proportion}}{\textit{standard deviation when null hypothesis is true}} \tag{4}$$

This Z-score measures the number of standard errors between the sample proportion $\hat{p}$ and the null hypothesis p_o.

For the astrology example the standard deviation is $\sqrt{\frac{\frac{1}{3}(\frac{2}{3})}{116}}$=0.0438. The astrologers were correct with 40 out of their 116 predictions which gives a sample proportion of $\hat{p}$. The test statistic is

$$Z = \frac{\hat{p}-p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}} = \frac{0.345-\frac{1}{3}}{0.0438} = 0.26$$

The sample proportion of 0.345 is only 0.26 standard deviations above the null hypothesis value of $\frac{1}{3}$.

Step 4: P-value

Does Z=0.26 give much evidence against $H_o : p = \frac{1}{3}$ an in suport of $H_a : p > \frac{1}{3}$? The P-value summaries the evidence. It describes how unusual the data would be if H_o were true. The **P-value** is the probability that the test statistic takes a value like the observed test statistic or even more extreme, if the null hypothesis is true.

Draw a standard normal distribution showing the P-value for Z=0.26.

Step 5: Conclusion

We summarize the test by reporting and interpreting the P-value. The P-value of 0.40 is not very small and therefore does <u>not</u> provide strong evidence the null hypothesis. We would not conclude that astrologers have special predictive powers.