30 2025-04-07 | Week 12 | Lecture 29

30.1 Hypothesis Testing

Textbook section 8.1

The hypothesis testing framework: suppose you observe some random quantity, which is assumed to depend on some unknown numerical parameter θ. You make a *hypothesis*, which can be any claim about the value of some parameter.

We are interested in testing the following two hypotheses:

- The *null hypothesis:* whatever effect you observed was due to random chance. This is denoted H_0.
- The *alternative hypothesis:* the claim you made about the parameter. This is usually denoted H_1

Based on our observation(s) we will make a decision to either reject H_0, or fail to reject H_0.

Suppose the observed data is X. The *p-value* of X is defined as

$$\text{p-value} := \mathbb{P}\left[\text{observing a result at least as extreme as } X \mid H_0\right]$$

When the p-value is small, we **reject** H_0. When the p-value is not small, we **fail to reject** H_0**.** The meaning of "small" here is subjective.

This is the basic story. Let's do an example:

Example 90. Suppose you have a coin but you don't know if it is a fair coin or not. You think maybe the coin is biased in favor of flipping heads. As an experiment, you flip the coin 10 times and you get 8 heads. Is this convincing evidence that the coin is biased?

In this case, the data is an IID sample

$$X_1, \ldots, X_{10}$$

where

$$X_i = \begin{cases} 1 & : \text{ with probability } \theta \\ 0 & : \text{ with probability } 1-\theta \end{cases}$$

The null hypothesis is

$$H_0 : \theta = \frac{1}{2}$$

The alternative hypothesis is that heads is more likely than tails:

$$H_1 : \theta > \frac{1}{2}$$

The p-value is

$$\begin{aligned} p &= \mathbb{P}\left[\text{at least 8 heads} \mid \theta = \frac{1}{2}\right] \\ &= \mathbb{P}\left[S_n = 8 \mid \theta = \frac{1}{2}\right] + \mathbb{P}\left[S_n = 9 \mid \theta = \frac{1}{2}\right] + \mathbb{P}\left[S_n = 10 \mid \theta = \frac{1}{2}\right] \\ &= \binom{10}{8}\theta^8(1-\theta)^2 + \binom{10}{9}\theta^9(1-\theta)^1 + \binom{10}{10}\theta^{10}(1-\theta)^0 \text{ with } \theta = \frac{1}{2} \\ &= \binom{10}{8}\left(\frac{1}{2}\right)^{10} + \binom{10}{9}\left(\frac{1}{2}\right)^{10} + \binom{10}{10}\left(\frac{1}{2}\right)^{10} \\ &= \frac{7}{128} \\ &\approx 0.0546. \end{aligned}$$

In words, if the coin were fair, we would expect it to flip 8 or more heads (out of ten) about 5% of the time. That's not actually super rare, so while this is consistent with the coin being biased, it's also consistent with the null hypothesis. Hence, I wouldn't find this to be convincing evidence that the coin isn't fair, so I wouldn't reject the null hypothesis.

Maybe you would. The cutoff for what is "convincing" is necessarily subjective.

End of Example 90. □

Example 91. Now, suppose that we flip the coin 100 times and we get 80 of them to be heads. Is this convincing evidence that the coin isn't fair?

To answer this question, we'll compute the p-value. Let $X_1, \ldots, X_{100}$ be our IID random sample, where

$$X_i = \begin{cases} 1 & : \text{ if } i\text{th coin is heads} \\ 0 & : \text{ otherwise} \end{cases}$$

Then the number of heads is the random variable

$$S_n = X_1 + \ldots + X_n.$$

Observe that $\mu = \mathbb{E}[X_i] = \frac{1}{2}$ and Assuming the coin is a fair coin, we have

$$\mu = \mathbb{E}[X_i] = \frac{1}{2}$$

and

$$\sigma^2 = \mathbb{E}\left[X_i^2\right] - \frac{1}{4} = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

so that

$$\sigma = \frac{1}{2}.$$

By the central limit theorem, we know that

$$\frac{S_n}{n} \approx \mu + \frac{\sigma}{\sqrt{n}} Z,$$

where Z is a standard normal random variable. For this specific problem, we have:

$$\frac{S_{100}}{100} \approx \frac{1}{2} + \frac{1}{20} Z.$$

We can now compute the p-value:

$$\begin{aligned}
\text{p-value} &= \mathbb{P}\left[\text{at least 80 heads out of 100 tosses} \mid \text{coin is fair}\right] \\
&= \mathbb{P}\left[S_{100} \geq 80 \mid \text{coin is fair}\right] \\
&= \mathbb{P}\left[\frac{S_{100}}{100} \geq .8 \mid \text{coin is fair}\right] \\
&\approx \mathbb{P}\left[\frac{1}{2} + \frac{Z}{20} \geq .8\right] && \text{Key step. Follows by CLT.} \\
&= \mathbb{P}\left[\frac{Z}{20} \geq .3\right] \\
&= \mathbb{P}\left[Z \geq 6\right] \\
&\approx .000000001
\end{aligned}$$

In other words, if the coin were a fair coin, observing heads at least 80 out of 100 times would be extraordinarily unlikely. Therefore we reject H_0.

End of Example 91. □

30.2 The Matching Problem

Suppose there are n people in class. Everyone writes their name on one playing card. The cards are then suffled, and dealt out again, so that each person gets a new (random) card. What is the probability that at least one person gets their original card?

By enumerating possibilities, we calculated that this has probability $2/3$ when there are $n = 3$ people.

We will return to this problem later.