My favorite lesson introduces what is arguably one of the most important topics in Introductory Statistics: hypothesis testing. This low cost lesson includes a beneficial review of binomial probabilities.

I first remind my students of the basics of binomial probability. An experiment will result in a binomial distribution when there are a fixed number of trials, each trial independent of the previous one. Additionally, there must be only two possible outcomes of each trial, with respective probabilities of $p$ (success) and $1 - p$ (failure), where $p$ is constant. In such an experiment, the probability of exactly $x$ successes in $n$ consecutive trials is

$$P(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

We then perform the experiment. I pull a U.S. quarter from my pocket and instruct a student to flip it exactly 7 times. Amazingly, the 7 flips come up 7 heads in a row! How likely is this unusual occurrence? We compute

$$P(7) = \binom{7}{7} .5^7 (1-.5)^{7-7} \approx .0078125$$

Thus, the likelihood of getting 7 heads in a row with 7 flips of a coin is less than 1%. I chose 7 flips in order to obtain a probability slightly less than 1%.

In our quick review of binomial probability calculations, I set the stage for introducing the concept of a null hypothesis. We consider a null hypothesis, $H_0$, “The coin is a fair coin,” and an alternative hypothesis, $H_1$, “The coin is a two-headed coin.” I frame the conversation around putting our coin “on trial.” The quarter is either innocent (a fair coin) or guilty (a two-headed coin).

A criminal defendant on trial is either innocent (null hypothesis) or guilty (alternative hypothesis). The jury takes a risk in declaring the final verdict. An innocent defendant could be found guilty. In this case, the jury rejects the null hypothesis (innocence) in favor of the alternative hypothesis (guilt), called a Type I error. Instead, a guilty defendant could be found innocent. In this case, the jury votes for the null hypothesis (innocence) when in fact the alternative hypothesis (guilt) is true. This is a Type II error. Our quarter is either innocent (a fair coin) or guilty (a two-headed coin). We must analyze the evidence we collected: the probability of getting 7 heads in a row is approximately .0078125. Is this too rare of an event to have occurred by chance? This is the moment to introduce alpha levels.

An alpha level sets a level of tolerance for accepting the null hypothesis in an experiment. If the result is below the alpha level, we vote instead for the alternative hypothesis; the experimental result is too rare to have occurred by chance. Typical alpha levels are .01, .05, and .10. Here, if we choose an alpha level of .01, we would be saying that any event with a probability less than .01 is considered too rare to have occurred by chance. Thus, in the trial of our quarter, we select the alternative hypothesis—it is a guilty, two-headed coin. We have decided that the quarter is two-sided, even though 7 heads in a row can happen with an innocent, fair coin. We run the risk of having made a Type I error, accepting the
alternative hypothesis when the null hypothesis is true: the coin could actually be fair, and we just witnessed a rare event.

In this case we can actually verify whether the alternative hypothesis is in fact correct. I encourage the student look more closely at the coin, and it is in fact a two-headed quarter. My small investment in this two-headed quarter - - found on the internet - - pays off every year when I use this binomial experiment to introduce the concept and vocabulary of hypothesis testing.