Putting a Claim on Trial

Barry the Statypus-in-Training says:

Hey guys! I was looking over some of your notes and I saw that the very first candy bag from the warehouse had 31 white candies and 21 black candies ($n = 52$). The factory blueprint says the machine is set to mix them perfectly 50/50 ($p = 0.50$). But my sample proportion is $\hat{p} = 31/52 \approx 0.596$. Since 0.596 is clearly not 0.50, doesn't that instantly prove the machine is broken? We should call Ed, that echidna can fix anything!

Sally the Statypus says: Barry, slow down and put down your phone! You can't just jump straight to a conclusion like that. Random samples fluctuate naturally. If you flip a fair coin 52 times, you aren't going to get exactly 2 head-and-tail splits every single time. To handle this properly, we use a framework called a **Hypothesis Test**.

Bill the Statypus says: Think of it like a trial in a courtroom, kid. In our legal system, a defendant has a presumption of innocence. Your charge is that the machine is faulty, but our baseline rule is that we must operate under the assumption that the machine is entirely innocent—meaning it is functioning exactly as advertised until the data forces us to believe otherwise.

Task 1: Formalizing the Assumption
The Null Hypothesis (H_0) represents our **assumption** of innocence. It states that the machine is fair, and any deviation in Barry's bag is just a common fluctuation of pure chance.

a. State the null assumption about the population parameter p in words. Section 9.3.1 in your textbook can help with this.

b. Write this assumption down in terms of symbols, we refer to it as H_0. Refer to section 9.3.1 of your book to help with this.

Bill the Statypus says: From here on, we **must** use that value of p in any of our calculations.

Sally the Statypus says: Ok, Barry, now that we have our groundwork laid, you can make your claim.

Task 2: Making a Claim

The Alternative Hypothesis (H_a) represents the **claim** Barry is bringing against the factory—the claim we are trying to find evidence to prove.

a. State our alternative hypothesis about the population parameter p in words. Remember that Barry said he claims the machine is broken, but he didn't say it necessarily favors white candies.

b. Write this claim down in terms of symbols, we refer to it as H_a. Section 9.3.2 of your book can help with this.

Barry the Statypus-in-Training says:

Okay, we have our assumption and my claim, but look at the data. Nearly 60% of my candies are white! That evidence greatly agrees with my claim!

Sally the Statypus says: You are right that it is not the value in H_0, but we can't expect $\hat{p}$ to be exactly 0.5. We need to be careful and measure exactly how rare this bag of candies is. Don't forget, we expect and welcome variation, but we also want to understand it and learn from it.

Task 3: Finding the Right Model

a. Look in Section 9.1 of your textbook and find the theorem that tells us the model we can use to analyze the distribution of sample proportions. Which theorem is it?

b. Explain why every condition of the theorem is true.

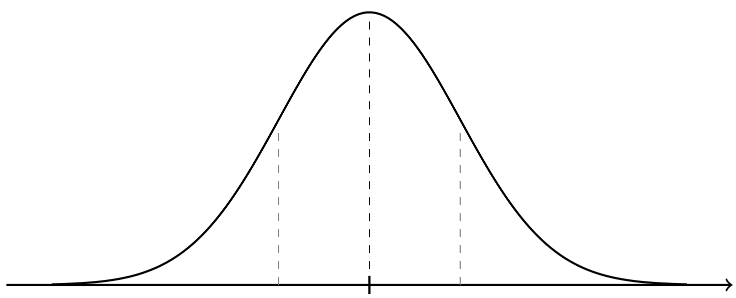
Barry the Statypus-in-Training says:

Wait! Can we actually use $\sigma_{\hat{p}}$ now? We are assuming $p = 0.5$.

Sally the Statypus says: That is an excellent observation, Barry! We were forced to use the standard error for confidence intervals before, but now that we are assuming we know the value of p, there is no approximation needed.

Task 4: Drawing the Model

To help us draw the model, use Theorem 9.3 to find the values of $\mu_{\hat{p}}$ and $\sigma_{\hat{p}}$. Then label the three dashed proportions in the generic Normal distribution plotted below.



Task 5: Shading the Real Meaning of Malfunction

Based off the values labeled on the plot above, label where our sample proportion, $\hat{p} \approx 0.596$, is on the graph and shade the "tail" from that value to the right.

Bill the Statypus says: Barry, don't forget that you would be just as alarmed if you saw 21 white candies and 21 blue candies. We should be careful about investigating your claim.

Barry the Statypus-in-Training says:

Got it! I shaded both tails because any extreme drift—high or low—counts as evidence that the machine is broken. Now I want to calculate the probability of those shaded areas.

Task 6: Finding the Area and thus the Probability

Use the `pnorm` function to find the area of each of the tails using the values from Task 4.

By identifying our baseline anchor ($H_0 : p = 0.50$) and establishing our exact measurement rule ($\sigma_{\hat{p}}$), we have built the reference distribution needed to judge Barry's data. The area we shaded during Task 5 is the sum of the values found in Task 6.

Sally the Statypus says: This area represents the probability that a sample of 52 candies would give us a sample proportion that was at least as extreme as our sample that had 31 white candies and 21 blue candies. In fact, this picture is drawn **assuming** that the machine is not broken and that $p = 0.5$.

Reflection: The Math Engine Pivot

You should have gotten a value in Task 6 that was roughly equal to one sixth. This means that we should expect to see samples which are at least as extreme as the one we have once in every 6 samples. At this point, we have to make a choice between two options as our **conclusion**:

1. Our assumption, $H_0 : p = 0.5$, is correct and our sample occurred by chance.
2. Our sample did not occur by chance alone, and, in fact, the assumption was wrong.

Assuming the machine is operating perfectly, we have a sample that is roughly as rare as rolling a 6 on a standard six sided dice. Based on this, which conclusion would you choose and why?

Statypus Insight: But...

The really canny reader may wonder if there are other conclusions we could reach. We want to make it clear here that we are also assuming that we have done our data collection in a way that does not allow any other conclusion.