The Power of the Sample Size

Bill the Statypus says: Imagine you are at a carnival. The game is simple: "Make 2 shots to win a prize." You get 3 attempts. You shoot and quickly realize the rim is bent and the ball is overinflated—it is nothing like the hoop at the school gym. You miss all 3.

Sally the Statypus says: The carney running the booth says, "Tell you what, try again for free." You shoot 3 more times and only make 1. You decide this game is a scam and start to walk away.

Bill the Statypus says: But then the carney makes his final pitch: "What if I give you 5 shots to make 2, instead of just 3? Still the same price."

Should you take the deal? Let's say your actual probability of making a shot on this rigged hoop is $p = 0.40$. Using our binomial probability tools (`pbinom`), the chance of making at least 2 shots out of 3 is roughly 35%. But with 5 shots, the chance of making at least 2 jumps to over 66%!

Sally the Statypus says: This carnival game isn't the exact same mathematical structure as building a confidence interval, but the moral of the story is perfectly aligned: We are mathematically more confident in our outcomes when we are given a larger sample size to work with.

Task 1: The Formal Architecture (Theorem 9.5)
Locate Theorem 9.5 in the text. Copy down the formula used to calculate the minimum sample size required for a specific margin of error and clearly label what each of the terms in that formula represents in plain English.

Task 2: The 99% Standard
Suppose you are auditing a massive shipment of our wedding candies. You want to be 99% confident that your calculated sample proportion of white candies is within 1% ($MOE = 0.01$) of the true population proportion. Calculate the required sample size using Theorem 9.5. Write your final answer as a full sentence starting with "We would need a sample size of..."

Sally the Statypus says: That formula from Theorem 9.5 is mathematically precise, but what if we just want a standard 95% confidence interval and have absolutely no idea what our sample proportion will be? Check out Theorem 9.6 in your book for a fantastic, simple shortcut!

Task 3: The 95% Shortcut
Now, use the simple rule from Theorem 9.6 to calculate the sample size needed to be 95% confident to within a slightly more relaxed precision of 3 percentage points, $MOE = 0.03$. Write your final answer as a full sentence starting with "We would need a sample size of..."

Sally the Statypus says: We don't need Bill to tell us that those two sample sizes are very different!

Bill the Statypus says: That may be true, but let's make sure we understand the practical interpretation of the sample size difference.

Reflection: The Friction of Reality

You just calculated two very different required sample sizes.

- To be **99% confident** with a very tight **1% margin of error**, your math shows you need roughly **16,587 samples**.
- To be **95% confident** with a more relaxed **3% margin of error**, the simple rule shows you need **1,112 samples**.

In the real world, data is rarely free. Assume that every survey participant requires an interviewer to spend **1 minute** to collect their response.

1. How many total **hours** of interviewing time are required for the high-precision strategy (16,587 samples)?

2. How many total **hours** of interviewing time are required for the practical strategy (1,112 samples)?

If your company pays an interviewer \$25.00 per hour, the jump from the practical strategy to the high-precision strategy represents a massive increase in financial expenditure. Is that roughly 15-fold increase in labor time and cost worth the gain in precision? Or, if this was your own work, would you be willing to work for 7 weeks to gain the improved numbers as compared to the simpler case where it only takes a bit over 2 days of work?

Task 4: Precision Planning in Public Polling

Imagine you are working for a regional polling agency tracking a closely contested local ballot initiative. The director needs to know how many random voters to contact over the upcoming weekend. Your goal is to determine the field requirements for the study, collect the data, and build the final interval.

Part 1: Determining Sample Size

The agency wants a standard regional check. They need to achieve a **90% confidence level** paired with a desired **4% margin of error** ($MOE = 0.04$).

1. Find the critical value (z^*) required for a 90% confidence level:

2. Calculate the exact minimum sample size (n) required showing your work.

Part 2: Field Execution & Inference

You go to work over the weekend and interview a random sample of 450 people to make sure you are above the minimum level. Exactly 241 voters state they intend to vote "Yes" on the ballot initiative.

1. Calculate the sample proportion ($\hat{p}$) of "Yes" votes from the field data:

2. Find the formal 90% confidence interval for the true population proportion based on these field results. *(Use the Code Template from Section 9.2.2).*

Sally the Statypus says: Look at how the loop closes perfectly! We used Theorem 9.5 to budget our labor and count how many doors to knock on, executed the data collection, and turned those raw counts right back into a precise confidence statement for our client.

Bill the Statypus says: True, but does this mean we are confident that the initiative will pass?

Sally the Statypus says: That's a completely different kind of question...

Bill the Statypus says: Yep, it's time to test our hypotheses!