The Digital Production Floor

You have completed your theoretical audit using the Central Limit Theorem. Now, we will use the script **DigitalWater.r** (found under the **Course Resources** tab at `statypus.org`) to simulate 10,000 production cycles.

Step 1: Calibrating the Simulator

Open `DigitalWater.r` in RStudio. The first section defines our "Truth." We are instructing the computer to behave exactly like the manufacturer claims the machine behaves.

```
mu    <- 500.5   # Our programmed target
sigma <- 0.2     # The manufacturer's tolerance
```

Step 2: The "Lone Wolf" Simulation ($n = 1$)

Previously, you shaded the "Target Zone" in the center of the distribution. In an audit, however, we are often more interested in the **failure rate**—the bottles that land in the tails. Run the following block to see 10,000 individual bottles:

```
lone_wolves <- rnorm(10000, mean = mu, sd = sigma)
hist(lone_wolves, xlim = c(499.5, 501.5))
abline(v = c(500.2, 500.8), col = "red", lwd = 3, lty = 2)
```

To find the exact proportion of bottles that missed the target (the unshaded area from your previous work), we use the "OR" operator (`|`) to count anything too low **or** too high:

```
mean(lone_wolves < 500.2 | lone_wolves > 500.8)
```

Analysis of Individual Variation: Based on your histogram and the proportion calculated above, how often does this machine produce a bottle that misses the target zone? Does this match the "white space" in the tails of your first hand-shaded curve?

Step 3: The "Sniper Rifle" Simulation (10,000 Cases of $n = 16$)

Now, run the `replicate()` section. This command tells R to simulate **10,000 separate cases**, where each case consists of **16 bottles**.

Note: For the purpose of this audit, we have already accounted for the tare weight of the plastic and packaging required to assemble a full case; the values below represent the liquid volume only.

```
case_avgs <- replicate(10000, mean(rnorm(16, mu, sigma)))
hist(case_avgs, xlim = c(499.5, 501.5))
abline(v = c(500.2, 500.8), col = "red", lwd = 3, lty = 2)
```

Observations of Group Consistency: Does the center of this distribution appear to be in the same place as the previous histogram? Describe what happened to the "spread" of the data relative to the red target lines.

The Probability Audit: What was the result of the count `sum(case_avgs <= 500.1)`? Based on your simulation of 10,000 cases, how likely is it to find a 16-pack with an average this low?

Theory vs. Evidence: The Verdict

In your earlier work, your math suggested that a 16-pack average of 500.1 mL was a $z = -8$ event. In this digital simulation, you likely saw zero results out of 10,000 at that volume.

The Confrontation: If you are the Lead Engineer holding a physical 16-pack in your hand right now that averages 500.1 mL, you are faced with a logical choice. You must choose the most reasonable explanation for what you are holding:

A. **The Miracle:** The machine is working perfectly ($\mu = 500.5$), and you just happened to pull the one-in-a-trillion case that exists.

B. **The Shift:** The machine is not actually centered at the target ($\mu \neq 500.5$).

Your Conclusion: Identify which option you choose and justify your decision using the evidence from your simulation. If we found a second, completely different 16-pack that also averaged 500.1 mL, how would that affect your confidence in your conclusion?