Hypothesis Testing for a Population Proportion: A Step-by-Step Guide with Theoretical Justification

Purpose

This guide walks you through how to perform a hypothesis test for a population proportion. At each step, you'll not only see what to do, but also understand the statistical logic that justifies it. Concepts are introduced carefully, one at a time, so you can build your understanding from the ground up.

Step-by-Step Procedure (with Theoretical Rationale)

Step 1: Choose a significance level α.

Before testing any hypothesis, we must decide how much evidence we require to reject it. The *significance level*, denoted α, is a threshold that represents how much risk we're willing to take of making a false positive — that is, rejecting a hypothesis that is actually true.

Theoretical rationale: Setting α gives us control over the Type I error rate. It establishes a standard for how unusual a result must be, under the assumption that the null hypothesis is true, to be considered statistically significant.

Step 2: Define the parameter of interest.

We begin by identifying what we are trying to learn about. In this case, we define p as the true population proportion of interest — for example, the proportion of voters who support a policy.

Theoretical rationale: In statistical inference, we use sample data to make reasoned claims about unknown population parameters. Clearly specifying p tells us what we are trying to estimate or test.

Step 3: State the null and alternative hypotheses.

The *null hypothesis*, denoted H_0, is the benchmark value we assume to be true unless the data provide strong evidence against it. The *alternative hypothesis*, H_A, represents the competing claim we want to test.

$$
\begin{aligned}
&H_0 : p = p_0 \\
&H_A : p < p_0, \quad p > p_0, \quad \text{or} \quad p \neq p_0
\end{aligned}
$$

Theoretical rationale: Hypothesis testing compares two competing statements about the population. H_0 serves as the default claim, while H_A expresses the specific deviation we're interested in testing.

Step 4: Check if the normal approximation is valid.

Because the sampling distribution of a sample proportion is not exactly normal, we only use the normal model if:

$$n \cdot p_0 \geq 15 \quad \text{and} \quad n \cdot (1 - p_0) \geq 15$$

These are called the *normality conditions*.

Theoretical rationale: The sample proportion $\hat{p}$ follows a binomial distribution. The Central Limit Theorem tells us this distribution becomes approximately normal when the sample size is large enough — specifically, when we expect at least 15 "successes" and 15 "failures" under H_0.

Step 5: Compute the test statistic.

Once we have sample data, we compute the observed sample proportion $\hat{p}$, and use it to calculate the standardized test statistic:

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

Theoretical rationale: If H_0 is true, then $\hat{p}$ has a known sampling distribution centered at p_0 with standard error $\sqrt{p_0(1 - p_0)/n}$. The Z-statistic measures how many standard errors our observed $\hat{p}$ is away from the null value.

Step 6: Calculate the p-value.

The p-value is the probability of observing a sample result as extreme as (or more extreme than) ours, assuming H_0 is true. To find it, we use the standard normal distribution.

Theoretical rationale: If H_0 is true, then the Z-statistic follows a standard normal distribution. The p-value tells us how surprising our observed result is under that assumption.

Step 7: Compare the p-value to α and make a decision.

- If p-value $\leq \alpha$: we reject H_0.
- If p-value $> \alpha$: we fail to reject H_0.

Theoretical rationale: We use the p-value to quantify evidence against H_0. If that evidence exceeds our pre-set threshold α, we reject H_0 in favor of H_A.

Step 8: State your conclusion in context.

Your final conclusion should interpret the result in real-world terms. For example: "We reject the null hypothesis at the 5% level. The data provide evidence that the support rate is greater than 50%."

Theoretical rationale: Statistical analysis always begins and ends in context. We use the result of the test to inform our substantive question about the population parameter p.

Summary Insight

Hypothesis testing is not just a checklist — it's a logical framework grounded in probability. We simulate what the world would look like if H_0 were true, and then use sample evidence to evaluate whether that world is still plausible. Every step is about connecting our observed data to the theory of sampling distributions under uncertainty.