Cheat Sheet: Hypothesis Testing for Population Proportions

Essential Concepts

One-Sample Z-Test of Proportions

1. Write out the null and alternative hypotheses.
    - The null hypothesis, H_0, is what we assume to be true to begin with. It is often a statement of no change from the previous value or from what is expected (e.g., we expect a coin to be fair).
        - The null hypothesis, H_0, is always given in the form: H_0: $p = a$.
    - The alternative hypothesis, H_A, is what we consider to be plausible if the null hypothesis is false. Often, it is a change from the null hypothesis that we would like to test the accuracy of.
    - The alternative hypothesis, H_A, is always given as an inequality:
        - H_A: $p < a$, H_A: $p > a$, or H_A: $p \neq a$
2. Check the conditions for the hypothesis test. For testing a one-sample z-test for proportions, we require:
    - Large Counts: Check that $np \geq 10$ and $n(1-p) \geq 10$.
    - Random Samples/Assignment: Check that the sample is a random sample.
    - 10% Population Size: Check that the sample size, n, is less than 10% of the population size, N: $n < 0.10(N)$
3. If not using technology, calculate a test statistic.
    - A test statistic measures the distance between the sample statistic and the null hypothesis value in terms of the standard error of the null hypothesis value.

$$\text{test statistic} = \frac{\text{sample statistic} - \text{null hypothesis value}}{\text{standard error of the null hypothesis value}}$$

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

- where $\hat{p}$ is the sample statistics and p is the null hypothesis value.

4. Use technology to calculate a P-value.
    - We define the P-value as the probability of obtaining a test statistic at least as extreme (in the direction of the alternative hypothesis) as the one that is actually seen if the null hypothesis is true.
    - The smaller the probability, the less likely it is that the sample occurred by chance alone, the more evidence we have against the null hypothesis.
    - Compare the P-value to the significance level, α, to make a decision.
        - The significance level, α, is the cut-off for P-values at which we have enough evidence to reject the null hypothesis. Typically, small significance levels such as 1%, 5%, or 10% are used in hypothesis testing.

Decision	Conclusion
If P-value $\leq \alpha$, there is enough evidence to reject the null hypothesis.	At the $\alpha \times$100% significance level, the data provide convincing evidence in support of the alternative hypothesis.
If P-value $> \alpha$, there is not enough evidence to reject the null hypothesis.	At the $\alpha \times$100% significance level, the data do not provide convincing evidence in support of the alternative hypothesis.

- We never accept the null hypothesis

5. Write a conclusion in context (e.g., we do/do not have convincing evidence...).
    - If a hypothesis test results in rejecting the null hypothesis because the P-value is less than the significance level, we say we have statistical significance.
    - If the results are meaningful, we say that the results have practical significance. Having practical significance usually means the results show a significant improvement!
    - Sometimes, due to chance, the result of the hypothesis test does not align with reality. If we reject a correct null hypothesis, we are committing a type I error. If we do not reject a null hypothesis that is actually incorrect, we are committing a type II error.

Null Hypothesis	Reject the null hypothesis	Do not reject the null hypothesis
is correct	Type I error	No error
is incorrect	No error	Type II error

Two-Sample Z-Test of Proportions

1. Write out the null and alternative hypotheses.
    - Null hypothesis: $H_0 : p_1 = p_2$ or $H_0 : p_1 - p_2 = 0$
    - Alternative hypothesis:

      $H_A:p_1<p_2$ or $H_A:p_1-p_2<0$

      $H_A : p_1 > p_2$ or $H_A : p_1 - p_2 > 0$

      $H_A : p_1 \neq p_2$ or $H_A : p_1 - p_2 \neq 0$

2. Check the conditions for the hypothesis test. For testing a one-sample z-test for proportions, we require:
    - Large Counts: For $\hat{p}_c = \dfrac{x_1 + x_2}{n_1 + n_2}$, check that: $n_1\hat{p}_c \geq 10$, $n_2\hat{p}_c \geq 10$, $n_1(1 - \hat{p}_c) \geq 10$, and $n_2(1 - \hat{p}_c) \geq 10$.
    - Random Samples/Assignment: Check that the two samples are independent and random samples or that they come from randomly assigned groups in an experiment.
    - 10%: Check that $n_1 < 0.10(N_1)$ and $n_2 < 0.10(N_2)$.

3. Calculate a test statistic.

4. Calculate a P-value.

5. Compare the P-value to the significance level, α, to make a decision.

6. Write a conclusion in context (e.g., we do/do not have convincing evidence...).

    - Typically, the conclusion drawn from a two-tailed confidence interval is usually the same as the conclusion drawn from a two-tailed hypothesis test. If a confidence interval contains the hypothesized parameter, a hypothesis test at the 0.05 level will almost always fail to reject the null hypothesis. If the 95% confidence interval does not contain the hypothesized parameter, a hypothesis test at the 0.05 level will almost always reject the null hypothesis. While this does not always hold for tests of proportions, a confidence interval typically provides more information about reasonable values of the parameter.

Key Equations

pooled sample proportion

$$\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2}$$

test statistic

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

Glossary

hypothesis testing

the process of forming hypotheses, collecting data, and using the data to draw a conclusion about the hypotheses

outcomes of hypothesis tests

reject the null hypothesis, fail to reject the null hypothesis

practical significance

the results of a hypothesis test are meaningful

***P*-value**

the probability of obtaining a test statistic at least as extreme (in the direction of the alternative hypothesis) as the one that is actually seen if the null hypothesis is true

significance level

the cut-off for P-values at which we have enough evidence to reject the null hypothesis

statistical significance

enough evidence against the null hypothesis to convince us to reject the null hypothesis

type I error

rejecting a correct null hypothesis

type II error

not rejecting an incorrect null hypothesis