Hypothesis Testing: One-Page Reference

BMSC 620

What is hypothesis testing?

Goal: Use sample data to evaluate evidence against a specific claim about a population parameter.

- The claim we test is the **null hypothesis**
- We assume the null is true and ask: *How surprising is our data under that assumption?*

Key hypotheses

Null hypothesis (H_0)

- Represents the status quo or specific claim
- Uses an equals sign

$$H_0 : \mu = \mu_0$$

Alternative hypothesis (H_A)

- Represents what we're looking for evidence in favor of
- Uses $\neq$, $<$, or $>$

$$H_A : \mu \neq \mu_0 \quad \text{(two-sided)}$$
$$H_A : \mu < \mu_0 \quad \text{or} \quad H_A : \mu > \mu_0 \quad \text{(one-sided)}$$

Significance level (α)

- α is the threshold for "strong evidence"
- Chosen **before** seeing the data
- Most common: $\alpha = 0.05$

Interpretation: If H_0 is true, we are willing to reject it incorrectly at most $\alpha \times 100\%$ of the time.

Assumptions for a one-sample t-test

1. Observations are **independent**
2. Data are approximately **normal** OR sample size is large ($n \geq 30$, CLT applies)

Test statistic (what comes from the data)

The t-statistic measures how far the sample mean is from the null value, in standard error units:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

- Comes from the **sample**
- **Random** (varies from study to study)
- Under H_0, follows a t-distribution with $df = n - 1$

Critical value (what defines "too extreme")

- Comes from the **t-distribution**
- Depends on α and degrees of freedom
- **Fixed** before seeing the data

$$t^* = t_{1-\alpha/2,\, df} \quad \text{(two-sided)}$$

Decision rules (three equivalent ways)

1. Test-statistic approach

Reject H_0 if: $|t_{\text{obs}}| > t^*$

2. P-value approach

P-value: Probability of observing a test statistic as extreme as (or more extreme than) what we saw, assuming H_0 is true.

Reject H_0 if: p-value $< \alpha$

3. Confidence interval approach (two-sided tests)

Reject H_0 if: the $(1-\alpha) \times 100\%$ confidence interval does **not** contain μ_0

What p-values mean (and don't mean)

P-value IS:

- A measure of how surprising the data are if H_0 were true

P-value is NOT:

- The probability that H_0 is true
- The probability you made a mistake
- A measure of effect size or importance

Common language to use

"We reject the null hypothesis"
"We fail to reject the null hypothesis"

"We accept the null hypothesis"

One-sample t-test in R

```r
t.test(x, mu = mu0, alternative = "two.sided", conf.level = 0.95)
```

Key output:

- t-statistic
- degrees of freedom
- p-value
- confidence interval
- sample mean

Reporting results (example)

A one-sample t-test was conducted to assess whether mean body temperature differs from 98.6°F. The sample ($n = 130$) had a mean of 98.25°F (SD = 0.733). The test was statistically significant, $t(129) = -5.45$, $p < 0.001$, with a 95% confidence interval of [98.12, 98.38], indicating that the population mean body temperature is lower than 98.6°F.

Big picture reminder

- Hypothesis tests and confidence intervals use the **same information**
- A small p-value does **not** imply practical importance
- Always **interpret results in context** “‘