Statistical Hypothesis Testing: Theory and Applications

Computational Statistics

November 24, 2025

1 Introduction

Hypothesis testing is a fundamental statistical method for making inferences about population parameters based on sample data. This document explores various hypothesis tests including one-sample and two-sample t-tests, paired t-tests, Analysis of Variance (ANOVA), chi-square tests for independence and goodness-of-fit, and non-parametric alternatives. We implement power analysis and effect size calculations to assess the practical significance of results and determine appropriate sample sizes for experimental design.

2 Mathematical Framework

2.1 General Hypothesis Testing

Test statistic and p-value relationship:

$$p = P(T \geq t_{obs}|H_0) \tag{1}$$

Decision rule at significance level α:

$$\text{Reject } H_0 \text{ if } p < \alpha \tag{2}$$

2.2 t-Test Statistics

One-sample t-test:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \tag{3}$$

Two-sample t-test (Welch):

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{4}$$

2.3 Effect Sizes

Cohen's d:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s_p} \tag{5}$$

where $s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$.

2.4 ANOVA F-statistic

$$F = \frac{MS_{between}}{MS_{within}} = \frac{SS_B/(k-1)}{SS_W/(N-k)} \tag{6}$$

3 Environment Setup

4 One-Sample t-Test

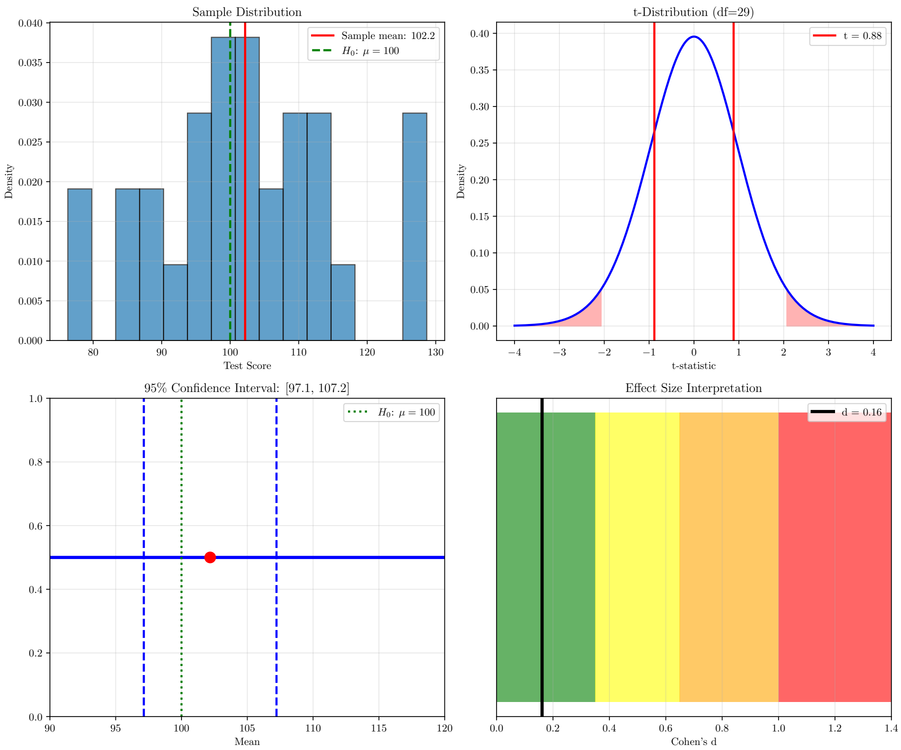

Figure 1: One-sample t-test analysis with effect size interpretation.

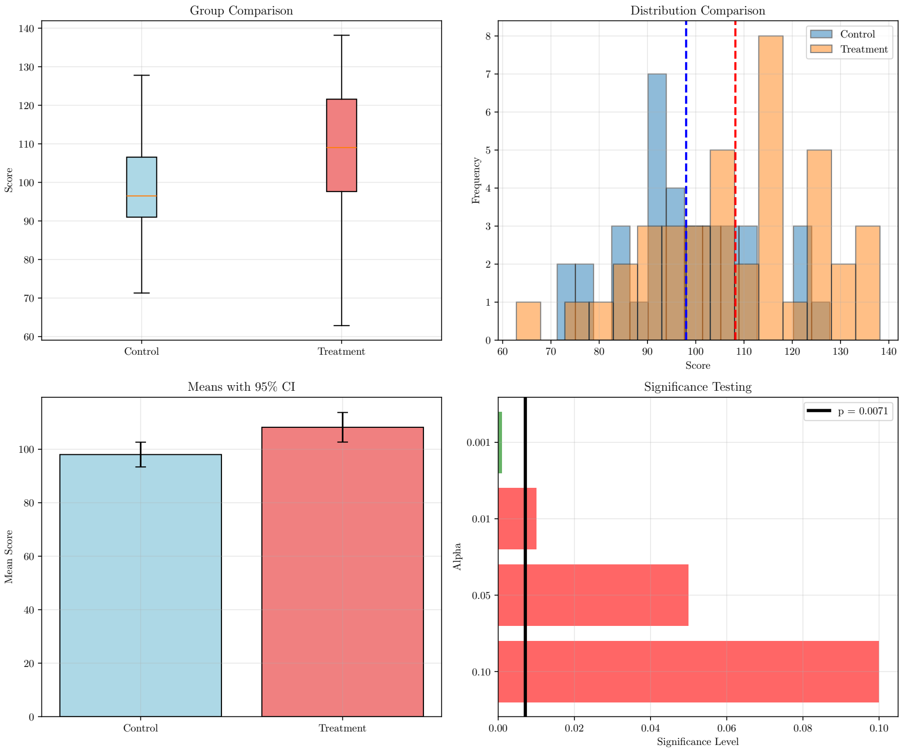

Figure 2: Two-sample t-test comparing treatment and control groups.

5 Two-Sample t-Test

6 Paired t-Test

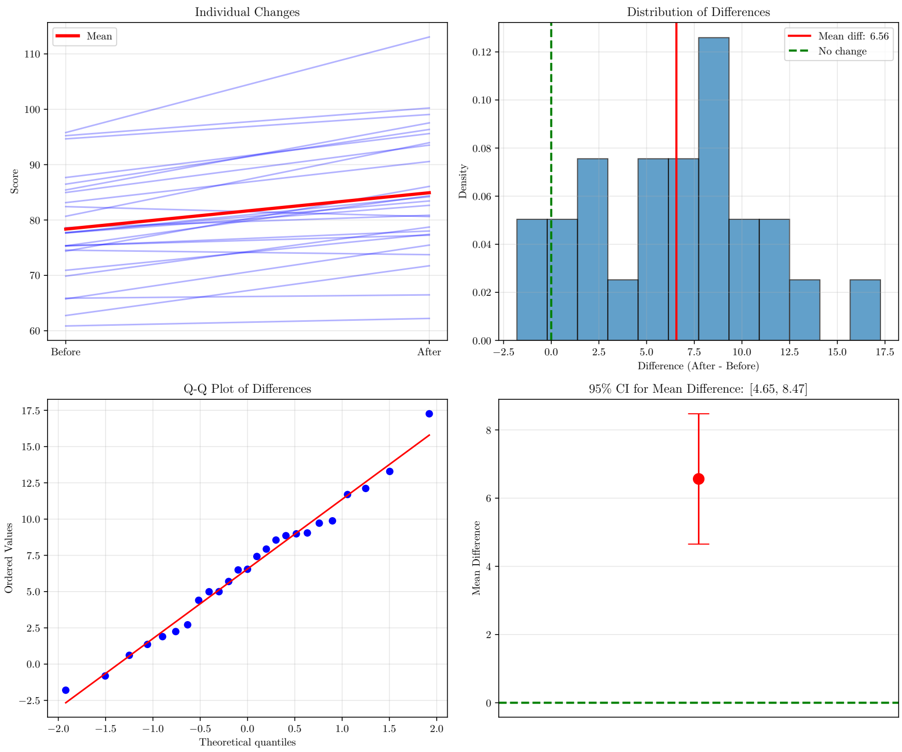

Figure 3: Paired t-test for before-after treatment comparison.

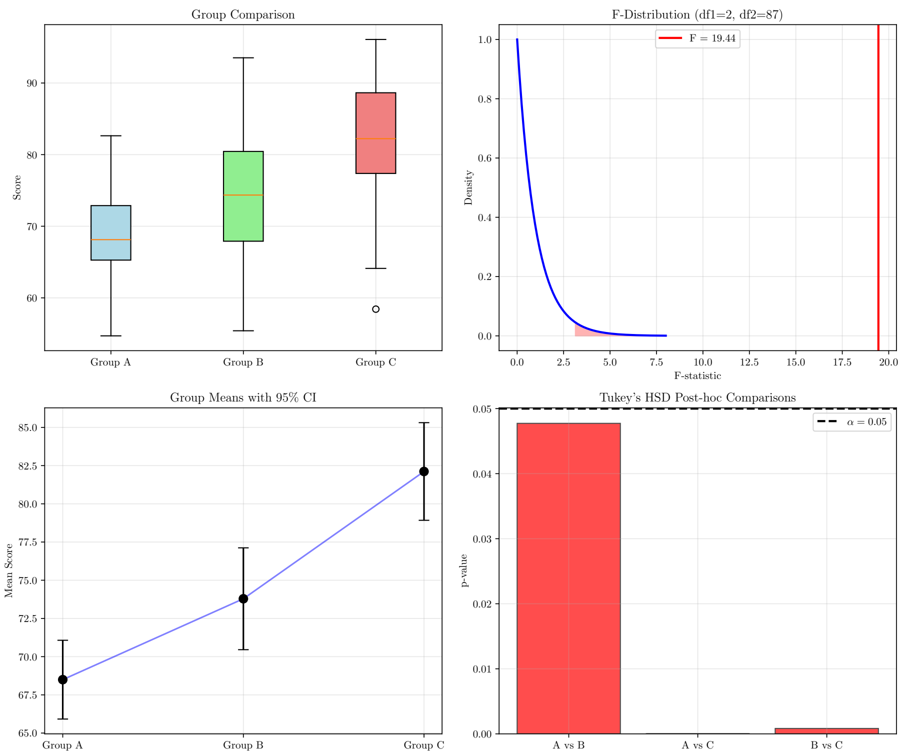

Figure 4: One-way ANOVA with Tukey's HSD post-hoc tests.

Table 1: Summary of t-Test Results

Test Type	t-statistic	p-value	Effect Size	Decision
One-sample	0.88	0.3842	d = 0.16	Fail to reject
Two-sample	-2.77	0.0071	d = 0.63	Reject H_0
Paired	7.09	0.0000	—	Reject H_0

Table 2: ANOVA and Chi-Square Test Results

Test	Statistic	df	p-value	Effect Size
One-way ANOVA	F = 19.44	(2, 87)	0.0000	$\eta^2 = 0.309$
Chi-square	$\chi^2 = 2.68$	2	0.2620	V = 0.116

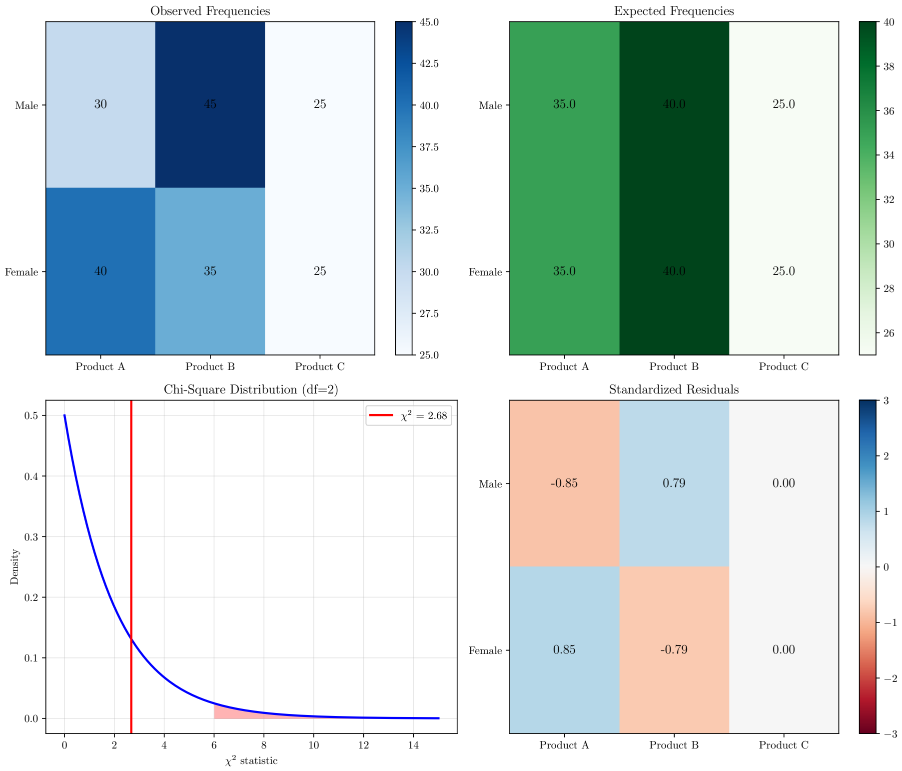

Figure 5: Chi-square test of independence for categorical variables.

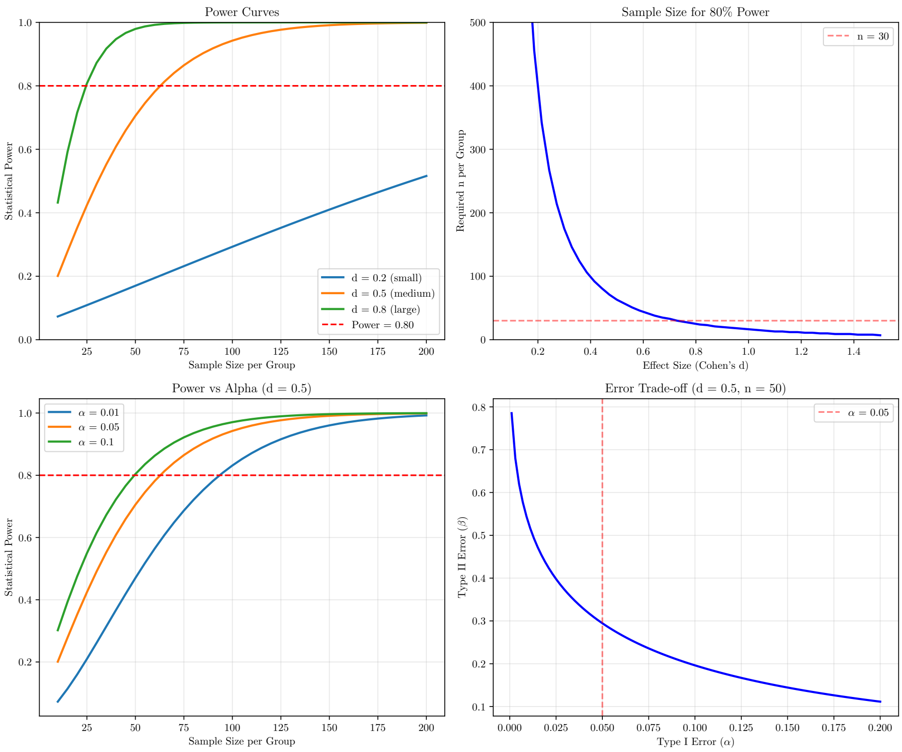

Figure 6: Power analysis showing sample size requirements and error trade-offs.

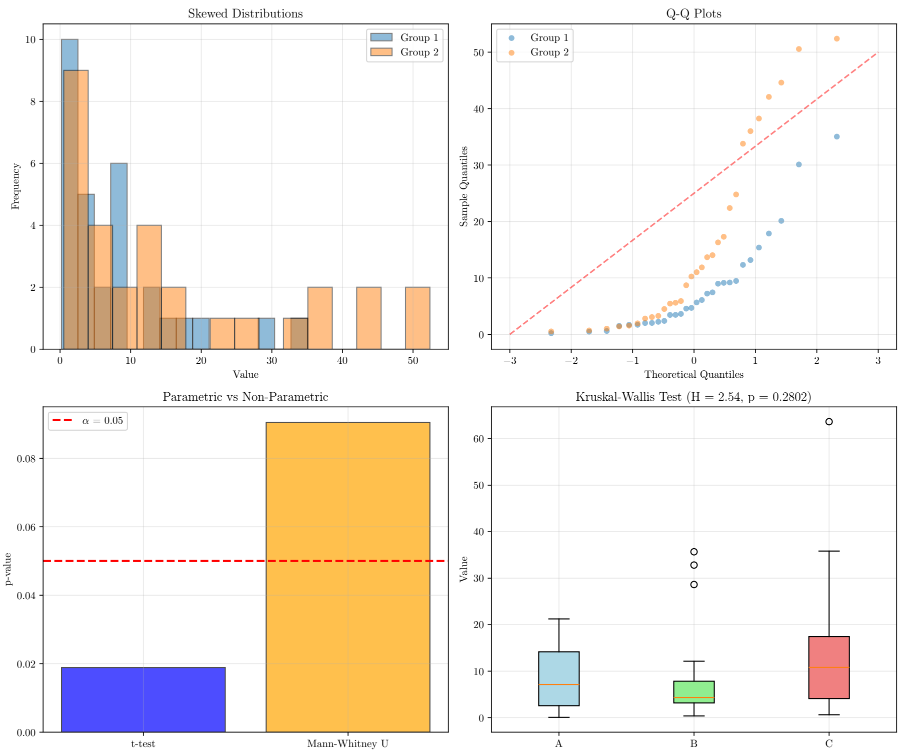

Figure 7: Non-parametric tests for skewed and non-normal distributions.

7 One-Way ANOVA

8 Chi-Square Test of Independence

9 Power Analysis

10 Non-Parametric Alternatives

11 Results Summary

11.1 t-Test Results

11.2 ANOVA and Chi-Square Results

11.3 Power Analysis Results

Table 3: Power Analysis Summary

Parameter	Value
Power (n=30, d=0.5)	0.491
Required n for d=0.5 (power=0.80)	63 per group
Required n for d=0.8 (power=0.80)	25 per group

12 Statistical Summary

Key hypothesis testing findings:

- One-sample t-test: p = 0.3842, d = 0.16
- Two-sample t-test: p = 0.0071, d = 0.63
- ANOVA: F = 19.44, $\eta^2 = 0.309$
- Chi-square: $\chi^2 = 2.68$, Cramer's V = 0.116
- Required sample size (d=0.5, power=0.80): 63 per group

13 Conclusion

This computational analysis demonstrates a comprehensive framework for statistical hypothesis testing. Effect sizes provide practical significance beyond p-values, with Cohen's d and eta-squared quantifying the magnitude of differences. Power analysis reveals that detecting small effects requires substantially larger samples than medium or large effects. Non-parametric alternatives offer robust options when distributional assumptions are violated. The combination of statistical significance, effect size, and confidence intervals provides a complete picture for scientific inference and decision-making.