
DATA SCIENCE BRAIN
@datasciencebrain

# HYPOTHESIS TESTING Cheat Sheet

Save for later reference

Our website: deepakjosecodes.com

## 01 DEFINE HYPOTHESES

- Null Hypothesis (H0):
  - A statement of no effect or no difference.
- Alternative Hypothesis (H1 or Ha):
  - States the effect, difference, or relationship you are testing.

## 02 SELECT SIGNIFICANCE LEVEL

- Common choices: 0.05, 0.01.
- Represents the probability of rejecting the null hypothesis when it is true.

Our website: deepakjosecodes.com

## 03 CHOOSE THE TEST

- **Z-Test for Population Mean:**
  - Known population standard deviation, large sample size.
- **T-Test for Population Mean:**
  - Unknown population standard deviation, small sample size.
- **Z-Test for Proportions:**
  - Binary outcome, large sample size.
- **Independent Samples T-Test:**
  - Compare means of two independent groups.
- **Paired Samples T-Test:**
  - Compare means of two related groups.
- **Z-Test for Difference in Proportions:**
  - Compare proportions of two independent groups.

Our website: deepakjosecodes.com

# 04 CHOOSE THE TEST

- **One-Way ANOVA:**
  - Compare means of three or more independent groups.
- **Chi-Square Goodness of Fit Test:**
  - Compare observed and expected frequencies for one categorical variable.
- **Chi-Square Test for Independence:**
  - Assess independence between two categorical variables.
- **Mann-Whitney U Test:**
  - Non-parametric alternative to independent samples t-test.
- **Wilcoxon Signed-Rank Test:**
  - Non-parametric alternative to paired samples t-test.
- **Kruskal-Wallis Test:**
  - Non-parametric alternative to one-way ANOVA.
- **Pearson Correlation Coefficient:**
  - Measure linear relationship between two continuous variables.
- **Spearman Rank Correlation:**
  - Non-parametric measure of correlation.
- **Simple Linear Regression Test:**
  - Assess significance of the relationship between one independent and one dependent variable.
- **Multiple Linear Regression Test:**
  - Assess overall significance of a multiple regression model.
- **Logistic Regression Wald Test:**
  - Assess significance of coefficients in logistic regression.
- **Log-Rank Test:**
  - Compare survival distributions between two or more groups in survival analysis.
- **Multivariate Analysis of Variance (MANOVA):**
  - Extension of ANOVA to multiple dependent variables.

Our website: deepakjosecodes.com

## 05 CONSIDERATIONS

- Normality Assumption:
  - Many parametric tests assume normality, check this assumption.
- Sample Size:
  - Larger sample sizes increase the power of the test.
- Equal Variances:
  - Some tests assume equal variances, check for homogeneity of variances.

## 06 INTERPRETATION

**P-Value:**

- If $p<\alpha$, reject H0.

**Effect Size:**

- Consider the practical significance of results.

Our website: deepakjosecodes.com

## 07 ADVANCED TECHNIQUES

**Bayesian T-Test:**

- If you want to incorporate Bayesian principles.

**Bootstrap Confidence Intervals:**

- If you prefer non-parametric methods and want robust confidence intervals.

**Never Miss a Post!**
Turn on the Notifications

# Was it helpful?

Follow Us For More Amazing Data Science & Programming Related Posts

DATA SCIENCE BRAIN
@datasciencebrain

LIKE TO SUPPORT

COMMENT

SHARE

SAVE FOR LATER

# Checkout Our Other Posts

DATA SCIENCE BRAIN
@datasciencebrain

## 07 Killer Data Science Project ideas

With Description

DATA SCIENCE BRAIN
@datasciencebrain

## Actual Projects That Data Scientists Work

On In Companies

DATA SCIENCE BRAIN
@datasciencebrain

## Data Science Concepts Explained

Overfitting & Underfitting

DATA SCIENCE BRAIN
@datasciencebrain

## Data Science Interview

Questions & Answers

**Save for later reference**