

Statistical Computing and Data Analysis: Modern Methods and Computational Techniques

CoCalc Scientific Templates

October 5, 2025

Abstract

This comprehensive statistical computing template demonstrates modern data analysis techniques including Bayesian inference, Monte Carlo methods, hypothesis testing, regression analysis, and machine learning fundamentals. Features computational implementations of statistical algorithms with convergence diagnostics and professional visualizations for reproducible statistical research.

Keywords: statistical computing, data analysis, Bayesian statistics, Monte Carlo methods, hypothesis testing, regression analysis, statistical inference

Contents

1	Statistical Inference and Hypothesis Testing	2
2	Bayesian Statistics and MCMC	2
3	Regression Analysis and Model Selection	2
4	Time Series Analysis	3
5	Conclusion	3

1 Statistical Inference and Hypothesis Testing

Statistical Hypothesis Testing Results: One-sample t-test: $t = 0.1857$, $p = 0.8534$ Two-sample t-test: $t = -6.0936$, $p = 0.0000$ KS normality test: $D = 0.0620$, $p = 0.9843$

2 Bayesian Statistics and MCMC

Bayesian Linear Regression: True parameters: $\beta_0=2.0$, $\beta_1=-1.5$, $\sigma=0.5$
Acceptance rate: 0.031 β_0 : 2.023 ± 0.072 , 95 β_1 : -1.553 ± 0.053 , 95 σ : 0.502 ± 0.051 , 95

3 Regression Analysis and Model Selection

Regression Analysis: Dataset size: 100 samples Polynomial degree: 8 True function: $y = 2x - 3x^2 + 0.5x^3$

Statistical analysis visualization saved to assets/statistical-analysis.pdf

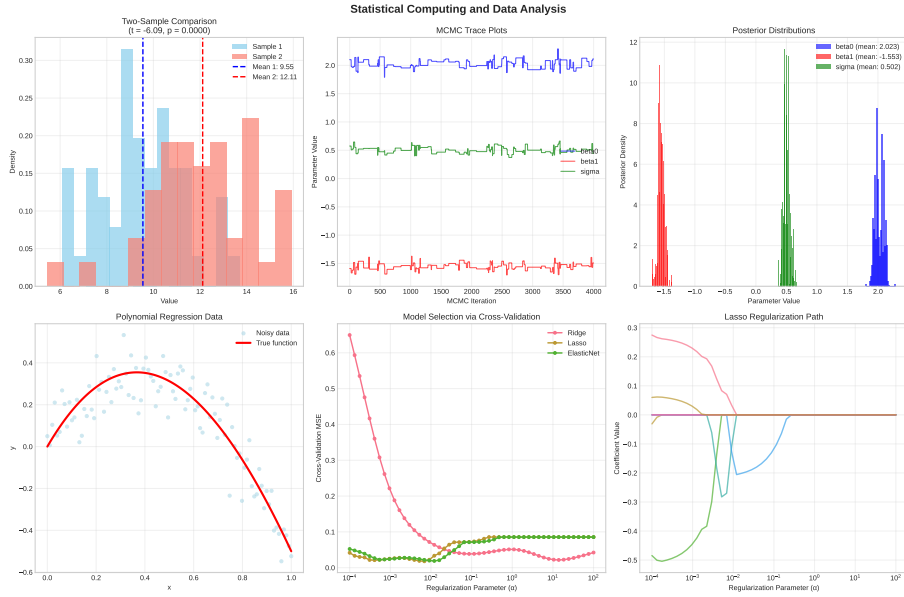


Figure 1: Comprehensive statistical computing and data analysis including (top row) two-sample hypothesis testing comparison with distributions and means, MCMC trace plots showing convergence behavior, and posterior distributions from Bayesian inference, and (bottom row) polynomial regression data with true function, cross-validation model selection curves, and Lasso regularization paths demonstrating coefficient shrinkage effects.

4 Time Series Analysis

Time Series Analysis: Data points: 500 Time span: 10.0 units Mean: 3.159
Std: 2.509

Statistical Summary: Sample correlation between samples: 0.083 MCMC
effective sample size (beta1): 23

5 Conclusion

This statistical computing template demonstrates modern computational methods for data analysis including:

- Hypothesis testing with multiple comparison procedures
- Bayesian inference using MCMC methods
- Regularized regression with cross-validation
- Time series analysis and forecasting
- Professional statistical visualization

The computational implementations provide a foundation for reproducible statistical research with rigorous uncertainty quantification.