The Undergraduate Statistician's Reading List

Dylan Armbruster

March 26, 2025

Mission Statement:

This collection is a curated set of resources I find essential for students of Statistics who want to go beyond the standard curriculum. At this level, it's easy to stick to what's taught in lectures, but there's so much more out there—papers, books, and blogs that provide unique perspectives, in-depth analyses, and practical insights you won't often encounter in class.

Each entry in this list has been chosen because it offers something extra: a clearer understanding, a fresh angle, or just the kind of wisdom that sticks with you as you grow in the field. This is the collection I wish I'd had when I started in Statistics, not only for the curious but for those who want to dig deeper and see where these concepts can really take them.

If I can offer any advice, I would say, get really comfortable with Mathematics and Computer Science. I've met too many students who fall behind in this major because they didn't have a strong enough Math and or Computer Science background.

Recommended Starting Path For The Casual Reader (High School Level Math. No Calculus Required)

I found these two books by Freedman to do a much better job than other introductory statistics books at this level. For example, how to properly understand Standard Deviation, random variable or R^2.

Statistics by David Freedman

Note: There are exercises with solutions in this text. The text is very motivating, as it includes real world examples and wisdom from experienced authors.

Mathematical Methods in Statistics a Workbook by David Freedman

Note: Supplement this text with Freedman's workbook or work on it after you've finished the text above. It is a bit more challenging because it includes

exercises where you will write computations for algebraic proofs (Don't worry! You don't need to know how to write proofs). This workbook mainly covers mathematical concepts relevant to statistics. Solutions to odd-numbered questions are included.

Reading List For Other Topics In Statistics

Bayesian Statistics:

- **Statistical Rethinking, McElreath** Note: This book is very unique compared to other texts. It has become more popular to include either R code or Python code, but this book offers a work flow along with it. You also get exposed to very relevant and useful topics today, such as Causal Inference and how to deal with Missing data. Of course, you also get exposed to Bayesian Statistics which you won't really see unless you take a Masters level class.
- **Bayesian Data Analysis Third Edition 3rd Edition Chapman HallCRC Texts in Statistical Science**

Probability

- **An Introduction to Probability and Inductive Logic**
- **Introduction to Probability, Second Edition by Blitzstein**
- **Lecture Video Series for the Blitzstein book** `https://projects.iq.harvard.edu/stat110/home`

Mathematical Statistics [Requires at least calculus 3

- **Mathematical Statistics with Applications, 7th Edition**

Statistical Inference

- **Improving Your Statistical Inferences, Lakens**
- **Confidence, Likelihood, Probability: Statistical Inference with Confidence Distributions, Schweder et al**

Non-Parametric Statistics

- **All of Nonparametric Statistics, Wasserman**

Mathematical Statistics [Started Real Analysis / Advanced Calculus]

- **Statistical Inference** Note: Graduate Level Text

DOE (Design of Experiments) and Design of Clinical Trials:

- **Statistics for Experimenters: Design, Innovation, and Discovery, 2nd Edition**
- **Introduction to Randomized Controlled Clinical Trials, Matthews**
- **Statistical Issues in Drug Development, Senn**
- **The Design Of Experiments, Fisher**
- **Clinical Trials: A Practical Approach, Pocock**
- **Design of Observational Studies, Rosenbaum**

The Bootstrap

- **An Introduction to the Bootstrap, Efron et al**

Likelihoodism

- **In All Likelihood: Statistical Modelling and Inference Using Likelihood, Pawitan**

Large-Scale Inference

- **Large-Scale Inference, Efron**

Applied Linear Regression

- **Applied Linear Regression Models**
- **Statistical Regression and Classification: From Linear Models to Machine Learning**

Generalized Linear Models

- Generalized Linear Models

Categorical Data Analysis

- Categorical Data Analysis, Agresti

Machine Learning

- An Introduction to Statistical Learning: with Applications in Python, James et al
- Probabilistic Machine Learning: An Introduction (Adaptive Computation and Machine Learning series, Murphy)

Philosophy Of Statistics

- Statistical Inference as Severe Testing: How to Get Beyond the Statistics Wars, Mayo

Statistics History

- Fisher, Neyman, and the Creation of Classical Statistics, Lehmann
- A History of Mathematical Statistics from 1750 to 1930, Hald
- The History of Statistics: The Measurement of Uncertainty before 1900, Stigler
- Logic of Statistical Inference, Hacking et al
- The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference, Hacking

Casual Reading on Statistics

- Dicing with Death (Living by Data), Senn
- Probability, Statistics and Truth, Von Mises et al
- The Seven Pillars of Statistical Wisdom, Stigler
- The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century, Salsburg

Hypothesis Testing

- **Statistical Hypothesis Testing in Context: Volume 52: Reproducibility, Inference, and Science (Cambridge Series in Statistical and Probabilistic Mathematics, Series Number 52), Fay et al**

- **Testing Statistical Hypotheses, Lehmann et al**

Programming:

R Programming:

- **The Art of R Programming: A Tour of Statistical Software Design**
- **Building reproducible analytical pipelines with R**

Python

- **Murach's Python Programming**

Data Structures

- **Data Structures** This uses C++ which I highly recommend to pick up as a language before you start any other language.

Causal Inference

- **The Book of Why, Pearl et al**
- **The Effect: An Introduction to Research Design and Causality, Huntington-Klein**
- **Observation and Experiment: An Introduction to Causal Inference, Rosenbaum**
- **A First Course in Causal Inference, Ding**
- **Causal Inference, Rothman**
- **Statistical Models and Causal Inference: A Dialogue with the Social Sciences, Freedman**
- **Causal Inference: What If, Robins, Hernan**
- **Causality: Models, Reasoning and Inference, Pearl**

Papers, Blogs, Youtube Videos

Papers

- **Ecological inference and the ecological fallacy** by David A. Freedman
- **Pay No Attention to the Model Behind the Curtain** by Philip B. Stark
- **Frequentist Probability and Frequentist Statistics** by J. Neyman
- **The Nature of Probability** by R.A. Fisher
- **In Defense of the Neyman-Pearson Theory of Confidence Intervals** by Mayo and Cox
- **Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations** by Greenland et al
- **Must We Randomize Our Experiment?, Box**
- **Null Hypothesis Significance Testing Defended and Calibrated by Bayesian Model Checking, Bickel**
- **Defending the P-value, Pawitan**
- **Frequentist statistical inference without repeated sampling, Vos et al**

Online Articles:

- **S. Senn: Randomisation is not about balance, nor about homogeneity but about randomness (Guest Post)**
- **Mathematics and Statistics**
- **Objections to Frequentism**

Blogs/Websites

- **Data Methods**
- **Andrew Gelman's Blog**

YouTube Videos

- **Bayesian or Frequentist, Which Are You?** By Michael I. Jordan (Part 1 of 2) (Jordan explains the difference between Frequentists and Bayesians)
- **Deborah Mayo: Statistical Inference as Severe Testing** (Philosophy of Statistics)

- **Mathematical Statistics** (Supplemental video for Undergrad Mathematical Statistics courses)
- **Causal Inference and Clinical Trials: Myths of Randomization** (Stephen Senn talks about Misunderstandings Practitioners have about Randomization)

Calculus / Real Analysis/ Linear Algebra

- **Paules Oneline Math Notes For Calculus(1-3)** (Study notes and practice problems)
- **Real Analysis: A Long-Form Mathematics Textbook, Cummings** (Formal textbook on Real Analysis)
- **Calculus: A Rigorous First Course** (Formal textbook on Calculus)

- ## Math Proofs
  - **Book of Proof, Hammack**
  - **Mathematical Maturity via Discrete Mathematics, Ponomarenko**