CHENE 4670 Chemical Engineering Data Analysis

Instructor

Prof. Kyle Bishop
kyle.bishop@columbia.edu, 212.854.7260, https://bishop.cheme.columbia.edu/

Teaching Assistant

Varies from year to year

Time

Varies from year to year

Location

Varies from year to year

Course Description

The ability to understand and design chemical engineering processes relies on mathematical models, which explain existing observations and make testable predictions. This course describes how experimental and chemical-process data can be used to estimate model parameters (parameter estimation), assess the relative plausibility of competing models (model selection), and design effective experiments (experimental design). These topics are connected within a common framework of *probability theory as extended logic*, which provides the unique consistent rules for conducting inference of any kind. Knowledge of these fundamental concepts will be of immediate and lasting value to all chemical engineers regardless of their future pursuits. Theory is reduced to practice through example problems drawn from chemical engineering practice and implemented in Python.

Course Prerequisites

This course is aimed at senior undergraduate and graduate students in Chemical Engineering. Familiarity with mathematical models in the context of thermodynamics, chemical kinetics, and transport phenomena is assumed. Prior knowledge of Python programming will be helpful but is not required.

Required Textbook

D. Sivia, J. Skilling, *Data Analysis: A Bayesian Tutorial*, 2nd Edition, Oxford University Press (2006)

Recommended Textbooks

A. Gelman *et al. Bayesian Data Analysis*, 3rd Edition, Chapman & Hall (2013)
P. C. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences*, Cambridge University Press (2005)
E. T. Jaynes, *Probability Theory: The Logic of Science,* Cambridge University Press (2003)
C. M. Bishop, *Pattern Recognition & Machine Learning,* Springer (2006)

Exams

One Midterm Exam. The exam is open book and open notes.

Homeworks

There will be five (5) problem sets, one due every two weeks. Late assignments will not be accepted.

Final Project

The course will culminate in a team project focused on a real-world problem in chemical engineering data analysis. The final project will include both a written report and a virtual poster presentation.

Course Grade

The following table provides an **approximate** conversion from points to letter grades.

	Points
Midterm	30
Homework	40
Project	30
Total	100

Grade	**Points**
-A+	85:90:95
-B+	70:75:80
-C+	65:70:75
-D+	50:55:60
F	<50

Course Schedule

Subject to modification from year to year.

Lecture No.	**Topic**	**Reading**
L00	Introduction to Bayesian data analysis	Sivia, Ch. 1
L01	Estimation of a single parameter	Sivia, Ch. 2.1
L02	Reliabilities in parameter estimation	Sivia, Ch. 2.2
L03	Gaussian noise model (known variance)	Sivia, Ch. 2.3, 2.4
L04	Multivariate parameter estimation	Sivia, Ch. 3.1
L05	Multivariate reliabilities	Sivia, Ch. 3.2
L06	Gaussian noise model (unknown variance)	Sivia, Ch. 3.3, 3.4

L07	Least squares regression	Sivia, Ch. 3.5
L08	Error propagation and variable transformations	Sivia, Ch. 3.6
L09	Bayesian model selection	Sivia, Ch. 4.1
L10	Hypothesis testing	Gregory, Ch. 7
L11	Model criticism	
L12	Midterm review session	
L13	Markov Chain Monte Carlo (MCMC) sampling	
L14	Logistic regression	
L15	Outliers	Sivia, Ch. 8.1-8.3
L16	Correlated noise	Sivia, Ch. 8.5
L17	Spectral analysis	Gregory, Ch. 13
L18	Assigning prior probabilities	Sivia, Ch. 5
L19	Experimental design	
L20	Parameter estimation in systems of ODEs	
L21	Gaussian mixture models	Bishop, Ch. 9
L22	Variational inference	Bishop, Ch. 10
L23	Bayesian neural nets	
L24	More numerics: autodifferentiation	

Disability-related academic accommodations

In order to receive disability-related academic accommodations for this course, students must first be registered with their school Disability Services (DS) office. Detailed information is available online for both the Columbia and Barnard registration processes. Refer to the appropriate website for information regarding deadlines, disability documentation requirements, and drop-in hours (Columbia)/intake session (Barnard).

Students registered with the Columbia DS office can refer to the Master TARF section of the DS Testing Accommodations page for more information regarding disability-related academic accommodations for this course.