MATH 50: READING NOTES

Discrete models and simulations (week 1)

- [1, Section 2.4]
- [2, Section 2.8]
- [2, Section 2.7]
  - This is where joint and marginal distributions are defined.

Expectation, Variance, Continuous models and Normal distribution (week 2)

- [2, Chapter 3]
  - Examples 3.1.6 and 3.17 were covered in class, the other examples are useful to know about, but the calculations are beyond what we do in this course.
  - I introduced the result of Theorem 3.1.1 as the definition of $E[g(X)]$, but they state it as a Theorem.
  - Theorem 3.1.2 and Theorem 3.1.3 were covered in class.
  - Exercise 3.1.2 is very good practice for the exams.
- [2, Section 2.4]
  - I didn't explicitly use their definitions. It might still be useful to skim through and get a different perspective, but don't worry too much about the technical details.
  - Examples 2.4.3 and 2.4.4 were covered in class.
  - The Normal distribution is introduced on pages 56-59
  - Exercises 2.4.1 and 2.4.2 are good practice.

Central limit theorem, multivariate Normal models, and regression basics (week 3)

- [2, Section 3.3]
  - Here is where they cover covariance.
- [2, Section 4.4.1]
  - Theorem 4.4.3 is the CLT, although stated slightly differently than I stated it.
  - Exercises 4.4.12-4.4.14 are good practice.
- [2, Section 2.5.4]
  - The mixture distributions they discuss are another way to see the conditional models we look at in class.
- [2, Section 10.1]
  - Example 10.1.1. This is where they introduce the linear regression model
- [2, Section 10.3]
  - Theorem 10.3.1 gives the formulas for a and b derived in class.
  - Theorem 10.3.4 will be discussed in class, time permitted.
  - The other results are mostly beyond the scope of this course or will be derived later in Week 5.
- [1, Section 3.1.1]
  - Their presentation of linear regression is quite different than mine. In particular, I take a much more probabilistic approach where the regression model is defined in terms of conditional probabilities.
  - Formulas 3.4 are the formulas for β_1 and β_0 from class (equivalent of [2, Theorem 10.3.1]).

Statistical inference and hypothesis testing (week 4)

- [2, Section 5.1]
  - This provides a nice conceptual introduction to statistical inference.
  - Read Example 5.1.1
- [2, Section 5.3]
  - Read Example 5.3.4
- [2, Section 4.4.2]
- [2, Section 6.3.1]
  - Bias and consistency are defined here.
  - Section 6.1 and 6.2 are also useful, but they go into much more technical detail than we do.
- [2, Section 6.3.1]
  - Example 6.3.7 is a nice supplement to class
- [1, Section 3.1.2]
  - This is a discussion of statistical inference for the regression model. The themes are the same as my lecture notes, but again, I take a much more probabilistic perspective.

Regression with multiple predictors (week 5-6)

- [1, Section 10.1]
  - Example 10.1.2. This is the linear regression model with multiple predictors.
  - This entire section is useful and all the exercises are good practice.

- [2, Section 10.2]
- [2, Section 10.5]
- [1, Section 3.2]

Interactions, nonlinear models and overfitting (week 7-8)

- A great discussion of overfitting and bias variance tradeoff [3]
- Cross validation. I don't really discuss all the difference types of cross validation [1, Section 5.1]
- Interactions terms are discussed here [1, Section 3.3]
- Discussion of nonlinear models, although with a different emphasis [1, Chapter 7]

Prediction and model evaluation (extra)

- [2, Section 10.5]

Logistic regression (extra)

- [2, Section 10.5]
- [1, Section 4.3]

References

[1] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, et al. *An introduction to statistical learning (python version)*, volume 112. Springer, 2013.

[2] John Tabak. *Probability and statistics: The science of uncertainty*. Infobase Publishing, 2014.

[3] Pankaj Mehta, Marin Bukov, Ching-Hao Wang, Alexandre GR Day, Clint Richardson, Charles K Fisher, and David J Schwab. A high-bias, low-variance introduction to machine learning for physicists. *Physics reports*, 810:1–124, 2019.