Preface

> I never felt such a glow of loyalty and respect towards the sovereignty and magnificent sway of mathematical analysis as when his answer reached me confirming, by purely mathematical reasoning, my various and laborious statistical conclusions.

Regression towards Mediocrity in Hereditary Stature
Sir Francis Galton, FRS (1886)

This collection of lecture notes is an updated version of my original lecture ntoes from 2017. They are now typeset in Quarto thanks to the suggestion of my former PhD student, Dr. Katie L Burak, who is currently an assistant professor of teaching at the University of British Columbia.

These revised and enhanced notes have some nice additions. Most notably, there is embedded R code and datasets within the text, which is (perhaps) the main point of using Quarto over my old LaTex notes. I have also taken it upon myself to add new sections and expand on those I find most interesting. This mostly occurs at the end of the notes with additional sections on advanced topics like logisitic regression, LASSO, and some Bayesian regression models.

The material in these notes is now too much for a single semester course in linear regression. Of note, I personally plan to skip the sections on influential points in regression models as the formulae are both tedious and ad-hoc. Nevertheless, I did copy these bits into the new version of my notes for completeness sake.

Adam B Kashlak
Edmonton, Canada
August 2025

The following are lecture notes originally produced for an upper level undergraduate course on linear regression at the University of Alberta in the fall of 2017. Regression is one of the main, if not the primary, workhorses of statistical inference. Hence, I do hope you will find these notes useful in learning about regression.

The goal is to begin with the standard development of ordinary least squares in the multiple regression setting, then to move onto a discussion of model assumptions and issues that can arise in practice, and finally to discuss some specific instances of generalized linear models (GLMs) without delving into GLMs in full generality. Of course, what follows is by no means a unique exposition but is mostly derived from three main sources: the text, *Linear Regression Analysis*, by Montgomery, Peck, and Vining; the course notes of Dr. Linglong Kong who

lectured this same course in 2013; whatever remains inside my brain from supervising (TAing) undergraduate statistics courses at the University of Cambridge during my PhD years.

Adam B Kashlak
Edmonton, Canada
August 2017