

Uncertainty Visualization

Matthew Kay

2018-10-15

Contents

1	Introduction	5
2	Why uncertainty visualization is hard	7
3	Foundations: Uncertainty as we'll see it	9
4	Foundations: Uncertainty and the grammar of graphics	11
5	Discrete data	13
6	Continuous data	15
7	Dealing with multiple uncertainties	17

Chapter 1

Introduction

This book¹ aims to provide the *why* and the *how* of uncertainty visualization. The *why* is based on what we currently know about how people interpret and make decisions from uncertainty visualizations, and the *how* is structured around generating uncertainty visualizations using the R programming language and the grammar of graphics.²

¹or what it will be—it is a work in progress

²as implemented in the *ggplot2* package (?)

Chapter 2

Why uncertainty visualization is hard

- Efficient encodings for uncertainty can be hard to find
- try putting mean, variance, and interval estimation in one plot + doing this when useful channels are already used up
- Even if you get the encoding right have to contend with making sure people understand it
- explain a CDF to someone
- Even if people understand have to deal with how people perceive uncertainty
- linear-in-log-odds and other perceptual models of probability
- Even if people correctly perceive the uncertainty have to deal with how people make decisions under uncertainty
- Even if you solve all these problems in one context a new context may cause your solution to break
- see e.g. deterministic construal errors
- Plus, you still have to build the damn thing
- But some strategies for building uncertainty vis using the grammar of graphics helps a lot here (the **how** of this book)

Chapter 3

Foundations: Uncertainty as we'll see it

- primarily Bayesian perspective
- unifies aleatory and epistemic
- gives us one framework to work in
- simplifies building things
- simplifies communicating things (see e.g. confidence fallacy)
- will still occasionally touch on frequentist, usually as a subsection at the end of each chapter:
- tips on how to translate the visualization techniques presented into that mode
- at times, visualizations unique to the frequentist approach that might be useful

Chapter 4

Foundations: Uncertainty and the grammar of graphics

- basic strategies for applying uncertainty visualization within the grammar of graphics
- visual channels, marks, etc. (and mapping onto ggplot)
- simple prediction grid example

Chapter 5

Discrete data

- categorical
- proportions
 - product plots / mosaic plots
- icon arrays
 - unit visualizations
- HOPs-like stuff
- entropy-based approaches
- ordinal / integer
- proportions
- histograms
- CDFs
- dotplots
- ...

Chapter 6

Continuous data

not sure of the organization here, especially within multivariate...

- univariate
- densities, gradients
- CDFs
- intervals
- quantile dotplots
- HOPs
- multivariate
- the joint case
 - gradients
 - contour plots
- the conditional case
 - gradients
 - bands
- generalizing quantile dotplots to multiple dimensions: representative sampling approaches
 - scatterplots
 - spaghetti plots
- HOPs
- frequentist addendum
- interpreting in a Bayesian mode: advantages and pitfalls (and what it reveals)
 - *this maybe goes elsewhere*
- those double-sided p value plot things

Chapter 7

Dealing with multiple uncertainties

- a note on comparisons
- univariate marginals might bite you...
- aleatory / epistemic (or parameter space / data space)
- frankenplots maybe (product plots / mosaic plots / unit visualizations)
- animation
 - NYT needle
 - Galton boards
- mixing discrete and continuous
- color with densities
- HOPs
- some hurricane stuff

Bibliography