

CHAPTER 6: ALGORITHMS, ANALYTICS AND PREDICTION

1. The *Titanic* data can be downloaded from <http://biostat.mc.vanderbilt.edu/wiki/pub/Main/DataSets/titanic3.xls>.
 2. Verifying probability of precipitation: <http://www.cawcr.gov.au/projects/verification/POP3/POP3.html>.
 3. 'Electoral Precedent', *xkcd*, <https://xkcd.com/1122/>.
 4. <http://innovation.uci.edu/2017/08/husky-or-wolf-using-a-black-box-learning-model-to-avoid-adoption-errors/>.
 5. The use of COMPAS and MMR algorithms is critiqued in C. O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Penguin, 2016).
 6. NHS, Predict: Breast Cancer (2.1): http://www.predict.nhs.uk/predict_v2.1/.
-

CHAPTER 7: HOW SURE CAN WE BE ABOUT WHAT IS GOING ON?
ESTIMATES AND INTERVALS

1. UK labour market statistics, January 2018: <https://www.ons.gov.uk/releases/uklabourmarketstatisticsjan2018>. Bureau of Labor Statistics, 'Employment Situation Technical Note 2018', <https://www.bls.gov/news.release/empsit.tn.htm>.
-

CHAPTER 8: PROBABILITY - THE LANGUAGE OF
UNCERTAINTY AND VARIABILITY

1. Consider Game 1. There are many ways of winning, but only one way of losing - throwing four non-sixes in a row. It is therefore easier to find the probability of losing (this is a common trick). The chance of throwing a non-six is $1 - \frac{1}{6} = \frac{5}{6}$ (complement rule), and the chance of throwing four non-sixes in a row is $\frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} = \left(\frac{5}{6}\right)^4 = \frac{625}{1296} = 0.48$ (multiplication rule). So the probability of winning is $1 - 0.48 = 0.52$ (complement rule again). Similar reasoning for Game 2 leads to the probability of winning to be $1 - \left(\frac{35}{36}\right)^{24} = 0.49$, showing that Game 1 was slightly more favourable. These rules also show the error in the Chevalier's reasoning - he was adding the probabilities of events that were not mutually exclusive. By his reasoning