Chapter 1.3

# Intro to Bayes

# A visit to the doctor's office

- Say you have a scratchy throat and so you go to the doctor to be tested from strep throat
- Denote your true disease status as $\theta = 1$ if you have strep throat and $\theta = 0$ otherwise
- An unknown quantity such as $\theta$ that we hope to estimate is called a **parameter**
- Unless the doctor's test is perfect, we will never know $\theta$ exactly

# A visit to the doctor's office

- Let the **data** $Y$ be the result of the rapid strep test with $Y = 1$ if you test positive and $Y = 0$ otherwise

- The distribution of the data given the parameters is called the **likelihood**

- In this example the likelihood is defined by the false positive rate

  $$\text{Prob}(Y = 1|\theta = 0) = p$$

  and the false negative rate

  $$\text{Prob}(Y = 0|\theta = 1) = q$$

- For now, we assume that we know $p$ and $q$ based on previous analyses

# A visit to the doctor's office

- Say you test positive and $Y = 1$, how likely is it that you have strep?
- To formalize this problem statistically, we must first decide whether $Y$, $\theta$ or both are random variables
- Should we treat $Y$ as random?

# A visit to the doctor's office

- Should we treat $\theta$ as random?

# Bayesian learning

- Bayesians quantify uncertainty about fixed but unknown parameters by treating them as random variables
- This requires that we set a **prior distribution** $\pi(\theta)$ to summarize uncertainty before observing the data
- The distribution of the observed data given the model parameters is the **likelihood function**, $f(Y|\theta)$
- The likelihood function is the most important piece of a Bayesian analysis because it links the data and parameters

# Bayesian learning

- The **posterior distribution** $p(\theta|Y)$ summarizes uncertainty about the parameters given the prior and data
- The reduction in uncertainty from prior to posterior represents **Bayesian learning**
- **Bayes' Theorem** (Bayes' Rule) converts the likelihood and prior to the posterior
- Bayes' Theorem:

$$p(\theta|Y) = \frac{f(Y|\theta)\pi(\theta)}{m(Y)}$$

where $m(Y) = \int f(Y|\theta)\pi(\theta)d\theta$ is the marginal distribution of the data and can usually be ignored

# How to select the prior?

- There is no “true” or “correct” prior
- In some cases expert opinion or similar studies can be used to specify an informative prior
- It would be a waste to discard this information
- If prior information is unavailable, then the prior should be uninformative
- The prior is best viewed as an initial value to a statistical procedure

# How to select the prior?

- As we'll see, as Bayesian learning continues and more and more data are collected, the posterior concentrates around the true value for any reasonable prior
- However, in a finite sample the prior can have some effect
- We will study several systematic ways to select priors
- However, there is inherent **subjectivity** to selecting the prior
- That is, different analysts may pick different priors and thus have different results

# How to select the likelihood?

- The likelihood is the same as in a freqentist analysis
- For example, in a linear regression analysis we might say

$$Y_i|\boldsymbol{\beta}, \sigma^2 \sim \text{Normal}\left(\sum_{j=1}^{p} X_{ij}\beta_j, \sigma^2\right)$$

- Is there “true” or “correct” likelihood?
- Is specification of the likelihood subjective like the prior?

# How to select the likelihood?

Subjective decisions required to specify the likelihood:

- The errors are independent
- $\log(Y_i)$ follows a normal distribution
- The mean is a function of $X_{i1}$, $X_{i3}$ and $X_{i3}^2$
- $Y_{43}$ is an outlier and removed
- Many, many more

# My opinions about subjectivity

- Perhaps we should aspire to objectivity, but in most real-life analyses we are forced to accept some subjectivity
- A notable exception is a tightly controlled experiment, although even this is debatable
- However, not all subjective decisions (assumptions) are equal
- If readers disagree with your assumptions they will reject your findings
- It is your job to justify your assumptions theoretically and empirically
- Determining the sensitivity to key assumptions is an important step