STAT 425: Introduction to Bayesian Analysis

Marina Vannucci

Rice University, USA

Fall 2017
Outline of the Course

1. Course overview and introduction
2. Bayes rule and example
3. The Binomial model
4. Bayesian inference for single-parameter models
5. Multiple-parameter models
6. Posterior simulation and integration
7. Markov chain Monte Carlo methods
8. Hierarchical models
9. Regression Models

Textbook:
Lecture 1: Introduction to Bayesian Statistics

- Bayes rule and example
- Bayesian inference (prior, likelihood, posterior)
Making inferences about hypotheses: the Frequentist & the Bayesian viewpoint

\[ P(\text{data} \mid H_0) \quad \quad P(H \mid \text{data}) \]
Experiment: phenomenon where outcomes are uncertain – e.g., single throws of a six-sided die.

Sample space: set of all outcomes of the experiment – $S = \{1, 2, 3, 4, 5, 6\}$.

Event: a subset of $S$ – $A = \{3\}$, $B = \{3, 4, 5, 6\}$. 
Basic properties of probability

- If $S$ is the sample space, $P(S) = 1$.
- For any event $A$, $0 \leq P(A) \leq 1$.
- For any complementary events $A$ and $A^c$,
  \[ P(A^c) = 1 - P(A) \quad P(\emptyset) = 1 - P(S) = 0 \]
- For any two events $A$ and $B$, the probability that either $A$ or $B$ will occur is given by
  \[ P(A \cup B) = P(A) + P(B) - P(A \cap B) \]
- The conditional probability of $A$ given $B$ for any two sets $A$ and $B$ is defined as
  \[ P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \text{if } P(B) \neq 0 \]
- $A$ and $B$ are independent if
  \[ P(A|B) = P(A) \quad \text{or equivalently} \quad P(A \cap B) = P(A)P(B) \]
Refresher: Conditional probability

Let $A$ and $B$ be two events. Then, we define the conditional probability of $A$ given $B$ as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- $A \cap B$ denotes the intersection of $A$ and $B$.
- Let $A^c$ denote the *complement* of $A$. 

The Bayes Theorem (two events)

- The Bayes Theorem allows us to obtain $P(A|B)$ from $P(B|A)$,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A) \times P(A)}{P(B)}$$

- Note that the denominator can be written as follows:

$$P(B) = P(B|A) \times P(A) + P(B|A^c) \times P(A^c)$$

(Law of total probability)
The Bayes Theorem (multiple events)

Let $A_1, \ldots, A_n$ be a partition of the sample space, i.e., a set of events such that

$$\bigcup_{i=1}^{n} A_i = S \quad \text{and} \quad A_i \cap A_j = \emptyset \quad \text{for} \quad i \neq j \quad \text{and} \quad P(A_i) > 0 \quad \text{for all} \quad i.$$

Then, given an event $B$,

$$P(A_j|B) = \frac{P(B|A_j) \cdot P(A_j)}{P(B)} = \frac{P(B|A_j) \cdot P(A_j)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}.$$
Example

Diagnostic test (non-controversial, widely accepted use of Bayes rule)

Say you know that HIV has a prevalence in the population of 1/1000. A particular test for HIV has a 95% sensitivity and 98% specificity\(^1\). What is the probability that someone testing positive actually has HIV?

\(B=\text{test is positive, } A=\text{have HIV}\)

\[P(A) = \frac{1}{1000}, \quad P(B|A) = 0.95, \quad P(B^c|A^c) = 0.98\]

(or \(P(B^c|A) = 0.05\) false negatives and \(P(B|A^c) = 0.02\) false positives)

\[P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} = \frac{(0.95)(0.001)}{(0.95)(0.001) + (0.02)(0.999)} = 0.045\]

Over 95% of the those testing positive will not have HIV - \(P(A^c|B)\)

(test of limited diagnostic value - low incidence disease)

\(P(A)=\text{prior belief (disease prevalence); } B=\text{data (test result)}\)

\(P(A|B)=\text{posterior probability of disease}\)

\(^1\)Sensitivity = prob of testing positive given the disease; specificity = prob of testing negative given that the individual is disease free
Despite the apparent high accuracy of the test, the incidence of the disease is so low that the vast majority of patients who test positive do not have the disease.

Nonetheless, this is 40 times the proportion before we knew the outcome of the test! The test is not useless.
More Comments

- Disease prevalence as “prior belief” that a person has the disease.
- We observe a positive result (i.e., data)
- Bayes rule tells us how the test result should change (update) our belief about the probability of disease in the presence of new evidence.
- We update our belief to a posterior probability of disease.
- More formal notation:

- $\theta =$ disease status ($\theta = 1$ is person has disease, $\theta = 0$ otherwise)
- $X =$ Random variable ($X = 1$ if test positive, $X = 0$ otherwise)
- **Probability model** for $X$: $P(X = i | \theta = j)$, $i, j = 0, 1$
- Prior belief on $\theta$: $P(\theta = 1) = 0.001$, $P(\theta = 0) = .999$
- Likelihood of $X = 1$: $P(X = 1 | \theta = 0) = 0.02$, $P(X = 1 | \theta = 1) = 0.95$
- Use Bayes rule to update our prior belief to

$$P(\theta = 1 | X = 1) = \frac{P(X=1|\theta=1)P(\theta=1)}{P(X=1|\theta=1)P(\theta=1)+P(X=1|\theta=0)P(\theta=0)} = 0.045$$
Example 2: Paternity dispute

- Suppose you are on a jury considering a paternity suit brought by Suzy Smith’s mother against Al Edged.
- Suzy’s mother has blood type O and Al Edged is type AB.
- You have other information as well. You hear testimony concerning whether Al Edged and Suzy’s mother had sexual intercourse during the time that conception could have occurred, about the timing and frequency of such intercourse, about Al Edged fertility, about the possibility that someone else is the father, and so on. You put all this information together in assessing $P(F)$, your probability that Al is Suzy’s father.
- The evidence of interest is Suzy’s blood type, which turns out to be B (if it were O, Al Edged would be excluded from paternity).
According to Mendelian genetics, $P(B|F) = \frac{1}{2}$.

You also accept the blood bank’s estimate $P(B|F^c) = 0.09$.

According to Bayes’ rule

$$P(F|B) = \frac{P(B|F)P(F)}{P(B|F)P(F) + P(B|F^c)P(F^c)}$$

The relationship between our prior probability, $P(F)$, and our posterior probability, $P(F|B)$ may be summarized:

<table>
<thead>
<tr>
<th>$P(F)$</th>
<th>0</th>
<th>0.100</th>
<th>0.250</th>
<th>0.500</th>
<th>0.750</th>
<th>0.900</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(F</td>
<td>B)$</td>
<td>0</td>
<td>0.382</td>
<td>0.649</td>
<td>0.847</td>
<td>0.943</td>
<td>0.980</td>
</tr>
</tbody>
</table>