Introduction to Bayesian Inference

January 13th, 2010

Reading: Hoff Chapter 1-2

Probability

- The outcome of an experiment is called an event
  - Tossing a coin
  - There will be a major terrorist attack this week
  - A women will develop ovarian cancer in the next 5 years
- The probability of an event A, $P(A)$, is a non-negative real-valued function of events that satisfies certain axioms that permit algebraic manipulation of probabilities.
  - Frequency and Long-Run Probability
  - Subjective Probability

Probability provides a natual language for communicating uncertainty

Probability: Frequentist's Views

If an event can occur in a finite number of ways then the frequency probability of the event is ratio of the number of ways that the event can occur to the total number of possible events.

- Counting: equally likely outcomes (card and dice games, sampling)
- Long-Run Frequency: The long-run or empirical probability of an event is the limit of the proportion of times that the event has occurred as the number of trials increases.
  Hurricane in September? Relative frequency from "similar" years

Not all events are repeatable! How should we proceed?

Subjective Probability

Subjective probability is expresses as a measure of belief; may combine

- expert opinion
- computer simulations
- historic data from not necessarily identical conditions

applies to unique events – probability of terrorist attacks – where long run frequencies are unavailable

Belief Functions

A *belief* function is a function that assigns numbers to statements such that the larger the number, the higher the degree of belief. Belief functions allow you to combine evidence from different sources and arrive at a degree of belief (represented by a belief function) that takes into account all the available evidence.
Statements F, G, and H.

- $\text{Be}(F) > \text{Be}(G)$ implies that we would prefer to bet that F is true over that G is true
- $\text{Be}(F \mid H) > \text{Be}(G \mid H)$ implies that if we know that H is true, then we would prefer to bet that F is also true over that G is true.

Axioms of Beliefs

B1 $\text{Be}(\text{not } H \mid H) \leq \text{Be}(F \mid H) \leq \text{Be}(H \mid H)$

B2 $\text{Be}(F \text{ or } G \mid H) \geq \max\{\text{Be}(F \mid H), \text{Be}(G \mid H)\}$

B3 $\text{Be}(F \text{ and } G \mid H)$ can be derived from $\text{Be}(F \mid H)$ and $\text{Be}(G \mid F \text{ and } H)$.

Probability Axioms (Conditional Version)

P1 $0 = \text{P}(H^C \mid H) \leq \text{P}(F \mid H) \leq \text{P}(H \mid H) = 1$

P2 $\text{P}(F \cup G \mid H) = \text{P}(F \mid H) + \text{P}(G \mid H)$ if $F \cap G = \emptyset$

P3 $\text{P}(F \cap G \mid H) = \text{P}(F \mid H)\text{P}(G \mid F \cap H)$

Probability functions satisfy belief axioms

Dempster-Shafer theory generalization of Bayes.

Probability

Events E, and a partition of H into disjoint sets H_i

$$H_i \cap H_j = \emptyset \tag{1}$$

$$\cup_i H_i = H \tag{2}$$

- $P(H_j)$
- $P(E \mid H_j)$

If E occurs, then how do we update our beliefs about H_i?

Bayes Theorem

Bayes' Theorem

Bayes' Theorem reverses the conditioning:
What is probability of H_i given that E has occurred?

$$P(H_i \mid E) = \frac{P(H_i \cap E)}{P(E)} \tag{3}$$

$$= \frac{P(E \mid H_i)P(H_i)}{P(E)} \tag{4}$$

Law of Total Probability:

$$P(E) = \sum_j P(E \mid H_j)P(H_j)$$

Twin Example

- Twins are either identical (single egg that splits) or fraternal (two eggs)
- Sexes: Girl/Girl GG, Boy/Boy BB, Girl/Boy GB or Boy/Girl BG
- GG or BB twins could be identical (I) or fraternal (F)

Given that I have GG twins, what is the probability that they are identical: $P(I \mid GG)$?
Easy to calculate the probability in the other way $P(GG \mid I)$

Identical Twins

$$P(I \mid GG) = \frac{P(GG \mid I)P(I)}{P(GG \mid I)P(I) + P(GG \mid F)P(F)}$$

- $P(I)$ is the overall probability of having identical twins (among pregnancies with twins). Real-world data shows that about $1/3$ of all twins are identical twins. This is our prior information.
- If the twins are identical, they are either two boys or two girls* It is reasonable to assume that the two cases are equally probable* $P(GG \mid I) = 1/2$
- In the case of non-identical twins, the four possible outcomes are assumed to be equally likely, so $P(GG \mid F) = 1/4$

Result

Substituting the probabilities

$$P(GG) = 1/2 * 1/3 + 1/4 * 2/3 = 1/3 \tag{5}$$

$$P(I \mid GG) = \frac{P(GG \mid I)P(I)}{P(GG)} = \frac{1/2 * 1/3}{1/3} = 1/2 \tag{6}$$

The probability that the twins are identical twins, given that they are both girls is $1/2$

The result combines experimental data (two girls) with prior information (1/3 of all twins are identical twins).

Bayes Inference

- 1763 – paper authored by Thomas Bayes was published “Bayes Theorem”
- how to make statistical inferences that build upon earlier understandings of a phenomenon and how to formally combine that earlier knowledge with currently measured data in a way that updates the degree of belief of the experimenter
- This procedure of updating is now called Bayesian Inference

Statistical Analysis

Statistical induction is the process of learning about the general characteristics of a population from a subset (sample) of its members

- "Characteristics" often expressed in terms of parameters "θ"
- measurements on the subset of members given by numerical values Y
- Before the data are observed, both Y and θ are unknown
- probability model for observed data if we knew θ is the truth
- What if we have prior information about θ?

Infection rate

Interest is in the prevalence of a disease (say H1N1 flu) in a region. Rather than using a census of the population, take a random sample of n individuals.

- θ: fraction of infected individuals
- Y_i indicator that ith individuals in the sample of n is infected
- Prior information on θ
- Given data Y_i, what the new beliefs about θ?

Bayesian Inference

Bayesian inference provides a formal approach for updating prior beliefs with the observed data to quantify uncertainty *a posteriori* about θ

- Prior Distribution $p(\theta)$
- Sampling Model $p(y \mid \theta)$
- Posterior Distribution:

$$p(\theta \mid y) = \frac{p(y \mid \theta)\, p(\theta)}{\int_{\Theta} p(y \mid \theta)\, p(\theta)\, d\theta}$$

(for discrete support for θ replace integral with sum)

Analysis Goals

Bayesian methods go beyond the formal updating of the prior distribution to obtain a posterior distribution

- Estimation of uncertain quantities (parameters) with good statistical properties
- Prediction of future events
- Tests of hypotheses
- Making Decisions

Software

- R
  - EDA
  - Frequentists analysis
  - Simulation
  - Graphics
- WinBUGS - scripting langauge for “Bayes Using Gibbs Sampling”
  - simulate priors
  - Posterior inference using several methods
  - easy to modify models/prior distributions
  - May be called from R

Install R for next Tuesday