Bayesian Statistics

Bayesian allows us to update our beliefs in the light of new evidence.

Example (F1 Race)

Suppose out four F1 races between Hamilton and Vettel, Hamilton won three times and Vettel once. They are going to race again. What is the probability that Hamilton ~~was~~ ~~will~~ wins? Our prior belief is $\frac{3}{4}$.

What if they both one one time when it was raining and Hamilton ~~&~~ won twice when it was sunny?

You know it will rain in the next race.

Our new belief is $\frac{1}{2}$.

Let's make this kind of argument precise.

Definition

The conditional probability of event A ~~gev~~ given event B for which $P(B)>0$, is

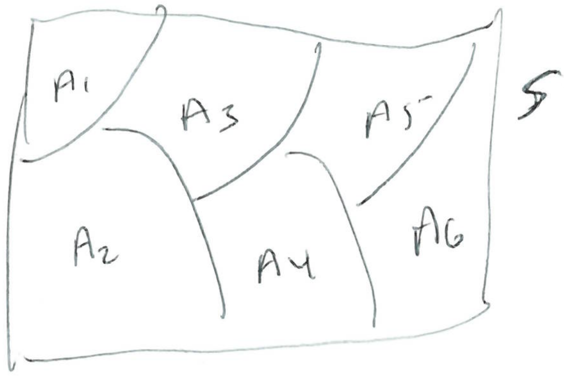
$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

A sample space S is the set of outcomes of an experiment.

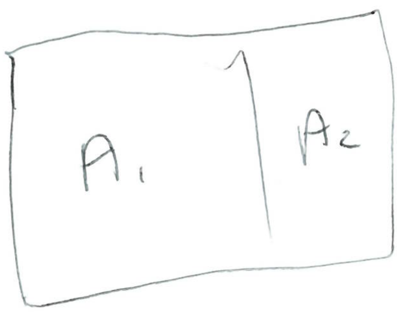
We say events $A_1, \dots, A_n$ partition S if

1. $A_1 \cup A_2 \cup \dots \cup A_n = S$
2. $A_i \cap A_j = \emptyset$ if $i \neq j$

picture



n=6



n = 2

$A_2 = A_1^c$ complement

Theorem (Bayes)

Let $B_1, B_2, \dots, B_n$ partition sample space S.

Let A be any event with $P(A) > 0$.

Then,

$$P(B_i \mid A) = \frac{P(A \mid B_i)\, P(B_i)}{P(A)}$$

$$= \frac{P(A \mid B_i)\, P(B_i)}{\sum_{j=1}^{n} P(A \mid B_j)\, P(B_j)}$$

Law of Total Probability

Proof

$$P(B_i \mid A) = \frac{\cancel{P(A)}\, P(B_i \cap A)}{P(A)} = \frac{\frac{P(B_i \cap A)}{P(B_i)}\, P(B_i)}{P(A)}$$

$$= \cancel{P(A \mid B_i} \; \frac{P(A \mid B_i)\, P(B_i)}{P(A)}$$

Example #1 continued

$B_1 = \{\text{Vettel wins}\}$ $B_2 = \{\text{Hamilton wins}\}$

$B_1 \cup B_2 = S$ they are just racing each other

$A = \{\text{it rains}\}$

$P(B_1) = \frac{1}{4}$ $P(B_c) = \frac{3}{4}$

$P(A) = \frac{2}{4} = \frac{1}{2}$

prob. it rained given Vettel won

$P(A|B_1) = \cancel{\ \ }$

$= \frac{1}{1} = 1$

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)} = \frac{1 \times \frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$

probability Vettel wins given that it rained

Example (Faulty Garment)

Three manufactures supply clothes to a retailer. Some clothes are faulty and the rest are not.

10% of clothes from manufacturer 1 are faulty.

5% " " 2 " "

15% " " 3 " "

60% of clothes are from 1

30% " " 2

10% " " 3

If a piece of clothing is faulty, what is the probability it came from manufacturer 3?

$B_i = \{\text{garment came from manufacturer } i\}$

$A = \{\text{garment is faulty}\}$

$$P(B_3|A) = \frac{P(A|B_3)P(B_3)}{\sum_{i=1}^{3} P(A|B_i)P(B_i)}$$

$$= \frac{0.15 \times 0.1}{0.10 \times 0.6 + 0.05 \times 0.3 + 0.15 \times 0.1}$$

$$= 0.167$$

$$> 0.10 = P(B_3)$$

When we apply ~~Bayesian~~ Bayes' Theorem, the random variables will ~~be continuous~~ usually be continuous.

Our prior beliefs are represented by a PDF on parameters $\underline{\theta} = (\theta_1, \dots, \theta_p)$. The parameters become random variables. This PDF is denoted by $f(\underline{\theta})$ and is called the prior distribution

We collect data $\underline{y} = (y_1, \dots, y_n)$ whose distribution depends on $\underline{\theta}$ through a likelihood function $f(\underline{y} \mid \underline{\theta})$

We will form a new p.d.f. on $\underline{\theta}$ the denoted by $f(\underline{\theta} \mid \underline{y})$ called the posterior distribution.

Comments

1. Let $p=1$. Θ is a r.v. so we could write it Θ and its density $f_{\Theta}(\theta)$, but we write $f(\theta)$.
2. Even if the Y_i are discrete random veriables $f(\theta)$ can be a p.d.f. E.G. $Y_i \sim \text{Poisson}(\theta)$, $\theta \in \mathbb{R}$.
3. Later we will let $f(\theta)$ be different from a p.d.f.
4. We write $f(\underline{y} \mid \underline{\theta})$ instead of $f(\underline{y}; \underline{\theta})$ because $\underline{\theta}$ is a random variable.

The posterior is defined by

$$f(\underline{\theta} \mid \underline{y}) = \frac{f(\underline{y} \mid \underline{\theta}) f(\underline{\theta})}{f(\underline{y})}$$

where

$$f(\underline{y}) = \int f(\underline{y} \mid \underline{\theta}) f(\underline{\theta}) \, d\underline{\theta}$$

Note that $\int f(\underline{\theta} \mid \underline{y}) \, d\underline{\theta} = 1$

4-step procedure for determining the posterior

Step 1 Select a prior $f(\underline{\theta})$

Step 2 Select a likelihood $f(\underline{y} \mid \underline{\theta})$

Step 3 Form $f(\underline{y} \mid \underline{\theta}) f(\underline{\theta})$ and determine

the posterior $\propto f(\underline{y} \mid \underline{\theta}) f(\underline{\theta})$

proportional as a function of $\underline{\theta}$

Any factors not dependent on $\underline{\theta}$

can be ignored.

Step 4 Name the distribution

A) either from its θ dependence

B) or calculate $f(\underline{y})$ and $f(\theta|\underline{y})$ directly.

Example (The number of claims in an insurance policy)

A new policy is sold to customers.
The number of claims in the first 6 months is 10. Assume the number of claims in different months are independent and Poisson(θ) distributed.

The ~~pre~~ prior on θ is Gamma(16, 4)

The prior was chosen to have expectation $\frac{16}{4} = 4$

and variance $\frac{16}{4^2} = 1$.

The prior $\cancel{f(\theta) = \frac{\lambda^{\alpha}}{\Gamma(\lambda)} y^{\alpha-1} e^{-\lambda y}}$

$$f(\theta) = \frac{\lambda^{\alpha}}{\Gamma(\lambda)} \theta^{\alpha-1} e^{-\lambda\theta}, \quad \theta > 0$$

$$f(\theta) \propto \theta^{16-1} e^{-4\theta}$$

$$\theta^{15} \cancel{e^{14}} e^{-4\theta}$$

$$f(\underline{y}|\underline{\theta}) = \prod_{i=1}^{6} \frac{e^{-\theta}\theta^{y_i}}{y_i!} = \frac{e^{-6\theta} \theta^{\sum_{i=1}^{n} y_i}}{\prod_{i=1}^{n} y_i!}$$

$$\propto e^{-6\theta} \theta^{10}$$

Therefore,

$$f(\theta|\underline{y}) \propto f(\theta) f(\underline{y}|\theta)$$

$$\propto \theta^{15} e^{-4\theta} e^{-6\theta} \theta^{10}$$

$$\propto \theta^{25} e^{-10\theta}$$

The posterior distribution is Gamma(26, 10)

The posterior expectation is $\frac{26}{10} = 2.6$

The posterior variance is $\frac{26}{10^2} = 0.26$

10

Example [Short life/Long life Batteries]

You bought 10 batteries.
Batteries have lifetimes which are Exponential(λ) where $\lambda = \frac{1}{1000}$ or $\lambda = \frac{1}{100}$

You don't which type the batteries are, but they are all of the same type.
After 80 hours, all the batteries are still working.

You have no prior opinion about which type of batteries you bought.

$$P\left(\lambda = \frac{1}{1000}\right) = P\left(\lambda = \frac{1}{100}\right) = \frac{1}{2}.$$

Let Y be the lifetime of one battery.

$$P(Y > 80) = e^{-80\lambda}$$

so $$\prod_{i=1}^{10} P(Y_i > 80) = \left(e^{-80\lambda}\right)^{10} = e^{-800\lambda}$$