On Bayesian Statistical Inference

1 Total Probability and Bayes Formula

Assume that the whole sample space S is partitioned by events $H_1, H_2, \ldots, H_n$. The events H_i will be called **hypotheses** and they satisfy:

$$H_1 \cup H_2 \cup \ldots \cup H_n = S$$
$$H_i \cap H_j = \emptyset, i \neq j.$$

If we are interested in probability of event A, then the **formula of total probability** gives

$$P(A) = P(A|H_1)P(H_1) + P(A|H_2)P(H_2) + \ldots + P(A|H_n)P(H_n).$$

Recall that multiplication rule claims:

$$P(A \cap H) = P(A)P(H|A) = P(H)P(A|H).$$

This simple identity is the essence of **Bayes Theorem:**

$$P(H|A) = \frac{P(A|H)P(H)}{P(A)}.$$

1.1 Problems

1. A new test has been devised for detecting a particular type of cancer. If the test is applied to the person who actually has that type of cancer, the probability that person will have a positive reaction is 0.995. If applied to a person who does not have this type of cancer, the probability that the person will have a (false) positive reaction is 0.012.

Suppose that test is given to patients at high risk of having cancer, and that one person, out of 50 in this group actually has this type of cancer.

(i) What is the probability that the randomly selected person from the group tests *positive?*

(ii) If randomly selected person tests *positive*, what is the probability that he/she **does not have** this type of cancer.

2. A student answers a multiple choice examination question that has 4 possible answers. Suppose that the probability that the student knows the answer to the question is 0.80 and the probability that the student guesses is 0.20. If student guesses, probability of correct answer is 0.25.

(i) What is the probability that the fixed question is answered correctly?

(ii) If it is answered correctly what is the probability that the student really knew the correct answer.

3. Factory has three types of machines producing an item. Probabilities that the item is I quality f it is produced on i-th machine are given in the following table:

machine	probability of I quality
1	0.8
2	0.7
3	0.9

The total production is done 30% on type I machine, 50% on type II, and 20% on type III. One item is selected at random from the production.

(i) What is the probability that it is of I quality?

(ii) If it is of first quality, what is the probability that it was produced on the machine I?

4 One out of 1000 coins has two tails. The coin is selected at random out of these 1000 coins and flipped 5 times. If tails appeared all 5 times, what is the probability that the selected coin was 'two-tailed'?

5. In Kokomo, IN, 50% are conservatives, 20% are liberals and 30% are independents.

Records show that in a particular election 82% of conservatives voted, 65% of liberals voted and 50% of independents voted.

If the person from the city is selected at random and it is learned that he/she did not vote, what is the probability that the person is liberal?

6. Businesses commonly project revenues under alternative economic scenarios. For a stylized example, inflation could be high or low and unemployment could be high or low. There are four possible scenarios, with the assumed probabilities:

Scenario	Inflation	Unemployment	Probability
1	high	high	0.16
2	high	low	0.24
3	low	high	0.36
4	low	low	0.24

a. What is the probability of high inflation?

b. What is the probability of high inflation if unemployment is high?

c. Are inflation and unemployment independent?

2 Estimation

Theorem 2.1 *Let $X \sim N(\theta, \sigma)$ and $\theta \sim N(\mu, \tau)$. Then*

$$\hat{\theta} = \frac{\sigma^2 \mu}{\sigma^2 + \tau^2} + \frac{\tau^2 x}{\sigma^2 \tau^2} = x - \frac{\sigma^2}{\sigma^2 + \tau^2}(x - \mu).$$

Example 1. A child is given an intelligence test. Assume that the test result is $N(\theta, 100)$, where θ is the true IQ level of the child, as measured by the test. Assume that in the population as a whole, θ is distributed according to a $N(100, 225)$ distribution. Using the above information estimate θ if the child scored $X = 115$.

Answer: $\theta = \frac{100 \cdot 100 + 225 \cdot x}{100 + 225} = 110.39.$

College Entrance Test. Because of the role of college aptitude test scores in college entrance decision, there are minicourses that purport to teach students how to take these tests. A particular aptitude test has been found to produce scores that are normally distributed, with mean θ and standard deviation 80. If the minicourse directed at this test is effective (on average), the mean score θ of students who take the course is larger than 500; otherwise it is not. We want to test

Figure 1: The Reverend Thomas Bayes

$$H_0 : \theta \leq 500 \quad \text{versus} \quad H_1 : \theta > 500,$$

and our prior for θ is $N(520, 20^2)$.

1. Find the prior probabilities of hypotheses H_0 and H_1, π_0 and π_1.
2. If 50 observations gives the mean 513, find the posterior probabilities of hypotheses H_0 and H_1, p_0 and p_1, and make the decision.
3. What is 90% credible set for θ.