STATISTICS IS THE GRAMMAR OF SCIENCE

PROBABILITY AND STATISTICS

LECTURE – 12

BAYE'S THEOREM

PREPARED BY
HAZBER SAMSON
FAST NUCES ISLAMABAD

BAYE'S THEOREM

Bayesian statistics is a collection of tools that is used in a special form of statistical inference which applies in the analysis of experimental data in many practical situations in science and engineering. Bayes' rule is one of the most important rules in probability theory. It is the foundation of Bayesian inference,

PARTITION OF AN EVENT

Sometimes, it is possible to partition an event, say *A*, into the union of two or more mutually exclusive events. Let S be a sample space, we want to partition S. Events $B_1, B_2 ..., B_k$ are said to partition a sample space *S* if the following two conditions are satisfied:

$$1 - B_i \cap B_j = \phi \quad \forall \; i \; and \; j$$

$$2 - B_1 \cup B_2 \cup B_3 \cup ... \cup B_K = S$$

LAW OF TOTAL PROBABILITY

If the events $B \; and \; B^c$ constitute a partition of the sample space *S* and $P(B) \neq 0$ then for any event A in S, we can write A as the union of two mutually exclusive and collectively exhaustive events $A \cap B \quad and \quad A \cap B^c$

$$A = (A \cap B) \; \cup (A \cap B^c)$$

$$P(A) = P[(A \cap B) \; \cup (A \cap B^c)]$$

$$P(A) = P(A \cap B) \; + P(A \cap B^c) - P[(A \cap B) \cap (A \cap B^c)]$$

$$P(A) = P(A \cap B) \; + P(A \cap B^c) - 0$$

$$P(A) = P(A \cap B) \; + P(A \cap B^c)$$

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

This result is known as **Law of Total Probability.**

GENERALIZED LAW OF TOTAL PROBABILITY

If the events $B_1, B_2 ..., B_k$ constitute a partition of the sample space *S* and $P(B_i) \neq 0$ $for \; i = 1,2,...,k$ then for any event A in S

$$P(A) = P(A|B_1) \, P(B_1) + P(A|B_2) \, P(B_2) + ... + P(A|B_K) P(B_K)$$

$$P(A) = \sum_{i=1}^{K} \; P(A|B_i) \, P(B_i)$$

EXAMPLES OF LAW OF TOTAL PROBABILITY

EXAMPLE-1 The completion of a construction job may be delayed because of a strike. The probabilities are 0.60 that there will be a strike, 0.85 that the construction job will be completed on time if there is no strike, and 0.35 that the construction job will be completed on time if there is a strike. What is the probability that the construction job will be completed on time?

SOLUTION Let A be the event that the construction job will be completed on time and B be the event that there will be a strike.

We are given

$$P(B)=0.60, \quad So \quad P(B^c)=0.40, \quad P(A|B^c)=0.85 \quad and \quad P(A|B)=0.35$$

We have to find $P(A)$.

Using **Law of Total Probability**

$$P(A)=P(A|B)\,P(B)+P(A|B^c)\,P(B^c)$$
$$P(A)=(0.35)(0.60)+(0.85)(0.40)=0.55$$

EXAMPLE-2 The members of a consulting firm rent cars from three rental agencies: 60 percent from agency 1, 30 percent from agency 2, and 10 percent from agency 3. If 9 percent of the cars from agency 1 need an oil change, 20 percent of the cars from agency 2 need an oil change, and 6 percent of the cars from agency 3 need an oil change, what is the probability that a rental car delivered to the firm will need an oil change?

SOLUTION Let A be the event that the car needs an oil change, and $B1$, $B2$, and $B3$ are the events that the car comes from rental agencies 1, 2, or 3 respectively.

We are given

$$P(B_1)=0.60, \quad So \quad P(B_2)=0.30, \quad P(B_3)=0.10$$

Also

$$P(A|B_1)=0.09, \quad P(A|B_2)=0.20 \quad and \quad P(A|B_3)=0.06$$

We have to find $P(A)$.

Using **Law of Total Probability**

$$P(A)=P(A|B_1)\,P(B_1)+P(A|B_2)\,P(B_2)+P(A|B_3)P(B_3)$$
$$P(A)=(0.09)(0.60)+(0.20)(0.30)+(0.06)(0.10)=0.12$$

Thus, 12 percent of all the rental cars delivered to this firm will need an oil change

BAYE'S THEOREM

If the events B and B^c constitute a partition of the sample space S and $P(B) \neq 0$ then for any event A in S

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

Similarly

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

Proof

$$\begin{aligned} L.H.S &= P(B|A) \\ &= \frac{P(A \cap B)}{P(A)} \\ &= \frac{P(A|B).P(B)}{P(A)} \\ &= \frac{P(A|B).P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)} \\ &= R.H.S \end{aligned}$$

GENERALIZED BAYE'S THEOREM

If the events $B_1, B_2 ..., B_k$ constitute a partition of the sample space S and $P(B_i) \neq 0$ $for\ i = 1,2,...,k$ then for any event A in S

$$P(B_j|A) = \frac{P(A|B_j).P(B_j)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}$$

$for\ j = 1,2,...,k$

Proof The proof follows directly from the definition of conditional probability and the law of total probability. Note that

$$P(B_j|A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}.$$

EXAMPLES OF BAYE'S THEOREM

EXAMPLE-1 A recent graduate has submitted his application to the World Bank for a position in the young professional program. He knows the bank hires **4%** of its application. Only some of the applicants receive an interview. It is known that among all the applicants hired **98%** receive interviews, the other **2%** being related to division chiefs. Furthermore, among all applicants not hired only 1% are interviewed. Our recent graduate was just called for an interview and asks you tell him the probability that he will be hired, what do you tell him?

SOLUTION Let H denote the event that applicant is hired.
Let I denote the event that applicant is interviewed.
Now it is given that
$P(H)=0.04, \quad So \quad P(H^c)=0.96$
Also $P(I/H)=0.98, \quad So \quad P(I/H^c)=0.01$, We want to know that $P(H/I)=?$

Using **Baye's Theorem**

$$P(H/I)=\frac{P(I/H).P(H)}{P(I/H).P(H)+P(I/H^c).P(H^c)}$$
$$P(H/I)=\frac{(0.98)(0.04)}{(0.98)(0.04)+.(0.01)(0.96)}=0.803$$

Thus there is a chance better than 80%.

EXAMPLE-2 A laboratory blood test is 95 percent effective in detecting a certain disease when it is, in fact, present. However, the test also yields a "false positive" result for 1 percent of the healthy persons tested. (That is, if a healthy person is tested, then, with probability **0.01**, the test result will imply that he or she has the disease.) If **0.5** percent of the population actually has the disease, what is the probability that a person has the disease given that the test result is positive?

SOLUTION Let D be the event that the person tested has the disease and E the event that the test result is positive. Then the desired probability is $P(D/E)$.

Now using **Baye's Rule**

$$P(D/E)=\frac{P(E/D).P(D)}{P(E/D).P(D)+P(E/D^c).P(D^c)}$$
$$P(H/I)=\frac{(0.95)(0.005)}{(0.95)(0.005)+.(0.01)(0.995)}=0.323$$

Thus, only 32 percent of those persons whose test results are positive actually have the disease.

EXAMPLE-3 The probability a beginning golfer makes a good shot if he selects the correct club is $1/3$. The probability the shot is good with wrong club is $1/5$. In his bag there are four different clubs, only one of them is correct for the shot is about to make. Since the beginner knows practically nothing about the choice of the proper club, he selects a club at random. He chooses a club and takes a stroke.
(a) What is the probability he got off a good shot?
(b) Given that he got off a good shot what is the probability that he choose a wrong club.

SOLUTION Let G denote the event that golfer makes a good shot.
Let C denote the event that he selects the correct club.
We are given

$$P(C)=\frac{1}{4}, \quad So \quad P(C^c)=\frac{3}{4}$$

Also

$$P(G|C)=\frac{1}{3}, \quad P(G|C^c)=\frac{1}{5}$$

(a) Here we have to find $P(G)$.

Using **Law of Total Probability**

$$P(G)=P(G|C)P(C)+P(G|C^c)P(C^c)$$
$$P(G)=(1/3)(1/4)+(1/5)(3/4)=7/30$$

(b) Here we have to find $P(G)$.

$$P(C^c|G)=\frac{P(G|C^c)P(C^c)}{P(G)}$$
$$P(C^c|G)=\frac{(1/5)(3/4)}{(7/30)}=0.643$$

EXAMPLE-4 In a certain assembly plant, three machines, B_1, B_2 *and* B_3, make 30%, 45%, and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by each machine, respectively, are defective. Now, suppose that a finished product is randomly selected.
(a) What is the probability that it is defective?
(b) If a product was chosen randomly and found to be defective, what is the probability that it was made by machine B_3.
SOLUTION Let us define the events

A: the product is defective,
B_1: the product is made by machine B_1,
B_2: the product is made by machine B_2,
B_3: the product is made by machine B_3.

(a) Here we have to find $P(A)$. By Law of Total Probability we have

$$P(A) = P(A/B_1)P(B_1) + P(A/B_2)P(B_2) + P(A/B_3)P(B_3)$$
$$P(A) = (0.02)(0.3) + (0.03)(0.45) + (0.02)(0.25) = 0.0245$$

(b) Here we have to find $P(B_3/A)$.

Now using **Baye's Rule**

$$P(B_3/A) = \frac{P(A/B_3).P(B_3)}{P(A/B_1).P(B_1) + P(A/B_2).P(B_2) + P(A/B_3)P(B_3)}$$
$$P(B_3/A) = \frac{(0.02.)(0.25)}{0.0245} = \frac{0.005}{0.0245} = \frac{10}{49}$$

EXAMPLE-5 A manufacturing firm employs three analytical plans for the design and development of a particular product. For cost reasons, all three are used at varying times. In fact, plans 1, 2, and 3 are used for 30%, 20%, and 50% of the products, respectively. The defect rate is different for the three procedures as follows:

$$P(D|P_1) = 0.01, \qquad P(D|P_2) = 0.03, \qquad P(D|P_3) = 0.02,$$

where $P(D|P_j)$ is the probability of a defective product, given plan j. If a random product was observed and found to be defective, which plan was most likely used and thus responsible?

SOLUTION From the statement of the problem

$$P(P_1) = 0.30, \quad P(P_2) = 0.20, \quad \text{and} \quad P(P_3) = 0.50,$$

We must find $P(B_j/D)$ *for* $j = 1,2,3$.

Before that lets find $P(D)$.

$$P(D) = P(D/P_1)P(P_1) + P(D/P_2)P(P_2) + P(D/P_3)P(P_3)$$
$$P(A) = (0.01)(0.30) + (0.03)(0.20) + (0.02)(0.50) = 0.019$$

Now using **Baye's Rule**

$$P(P_1/D) = \frac{P(D/P_1).P(P_1)}{P(D/P_1).P(P_1) + P(D/P_2).P(P_2) + P(D/P_3)P(P_3)} = \frac{(0.01)(0.30)}{0.019} = 0.158$$

$$P(P_2/D) = \frac{P(D/P_2).P(P_2)}{P(D/P_1).P(P_1) + P(D/P_2).P(P_2) + P(D/P_3)P(P_3)} = \frac{(0.03)(0.20)}{0.019} = 0.316$$

$$P(P_2/D) = \frac{P(D/P_2).P(P_2)}{P(D/P_1).P(P_1) + P(D/P_2).P(P_2) + P(D/P_3)P(P_3)} = \frac{(0.02)(0.50)}{0.019} = 0.526$$

The conditional probability of a defect given plan 3 is the largest of the three; thus a defective for a random product is most likely the result of the use of plan 3.