MAT 271E: PROBABILITY AND STATISTICS

PROF. DR. CANAN SARICAM

WEEK 5

CONDITIONAL PROBABILITY AND STATISTICAL INDEPENDENCE

Conditional probability

The probability of event *B* when it is known that event *A* has occurred is defined as the **conditional probability**, denoted by P(B/A) and read as probability of B given A.
The conditional probability P(B/A) and the conditional probability *P*(*A*/*B*) are, respectively, defined as follows:

$$P(B/A) = \frac{P(A \cap B)}{P(A)}, \qquad P(A) \neq 0$$

$$P(A/B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B) \neq 0$$

Conditional probability is a measure of the probability of an event occurring, given that another event has already occurred.

Example 2

This question uses the following contingency table:

	Have pets	Do not have pets	Total
Male	0.41	0.08	0.49
Female	0.45	0.06	0.51
Total	0.86	0.14	1

What is the probability a randomly selected person is male, given that they own a pet?

Example 2 - solution

	Have pets	Do not have pets	Total
Male	0.41	0.08	0.49
Female	0.45	0.06	0.51
Total	0.86	0.14	1

	Have pets	Do not have pets	Total
Male	0.41	0.08	0.49
Female	0.45	0.06	0.51
Total	0.86	0.14	1

$P(M|PO) = P(M \cap PO) / P(M) = 0.41 / 0.86 = 0.47$

Example 1

Suppose that in an adult population the proportion of people who are both overweight and suffer hypertension is 0.09; the proportion of people who are not overweight but suffer hypertension is 0.11; the proportion of people who are overweight but do not suffer hypertension is 0.02; and the proportion of people who are neither overweight nor suffer hypertension is 0.78.
An adult is randomly selected from this population.

a) Find the probability that the person selected suffers hypertension given that he is overweight.
b) Find the probability that the selected person suffers hypertension given that he is not overweight.

Example 1 - solution

a.

$$P(H|O) = \frac{P(H \cap O)}{P(O)} = \frac{0.09}{0.09 + 0.02} = 0.8182$$

b.

$$P(H|O) = \frac{P(H \cap O^c)}{P(O^c)} = \frac{0.11}{0.11 + 0.78} = 0.1236$$

Statistical independence

In probability, we say two events are independent if knowing one event occurred doesn't change the probability of the other event.

Two events are statistically independent if and only if:

$$\boxed{P(A \cap B) = P(A)P(B)}$$

If A and B are independent:

$$\boxed{P(A|B) = P(A)} \qquad \text{if } P(B)>0$$

$$\boxed{P(B|A) = P(B)} \qquad \text{if } P(A)>0$$

Statistical independence

To extend this idea to three or more events, let us suppose that we have n statistically independent events·, A1, A2, A3 , ... , An-ı, An.
The probability of the intersection of all of the Ai , i = _ 1, . .. , n, . is

$$P(\cap_{i=1}^{n} A_i) = P(A_1 \cap A_2 \cap \cdots \cap A_{n-1} \cap A_n) = \prod_{i=1}^{n} P(A_i)$$

Statistical independence

Mutually exclusive events are those that cannot happen simultaneously, whereas **independent events** are those whose probabilities do **not** affect one another.

Example 3:

A jar contains 10 marbles, 7 black and 3 white. Two marbles are drawn without replacement, which means that the first one is not put back before the second one is drawn.

a) What is the probability that both marbles are black?
b) What is the probability that exactly one marble is black?
c) What is the probability that at least one marble is black?

Example 3 - solution

$$P(B_1 \cap B_2) = \frac{7}{10} \cdot \frac{6}{9} = 0.47$$

$$P(B_1 \cap W_2) = \frac{7}{10} \cdot \frac{3}{9} = 0.23$$

$$P(W_1 \cap B_2) = \frac{3}{10} \cdot \frac{7}{9} = 0.23$$

$$P(W_1 \cap W_2) = \frac{3}{10} \cdot \frac{2}{9} = 0.07$$

a. 0.47

b. 0.23+0.23 = 0.46

c. 0.23+0.23 + 0.47 = 0.93

Example 4:

The chance of a flight being delayed is 0.2 (=20%),

a) What is the chance of no delays on a round trip?

b) What is the chance of at least one of the flight is delayed?

c) What is the chance of both flights are delayed?

Example 4:

The chance of a flight being delayed is 0.2 (=20%),

a) What is the chance of no delays on a round trip?

b) What is the chance of at least one of the flight is delayed?

c) What is the chance of both flights are delayed?

No delay=1-0.2=0.8
No delays ona round trip=0.8*0.8=0.64

Both of trips are delayed= 0.2*0.2=0.04
First trip is delayed= 0.2*0.8=0.16
Second trip is delayed=0.16
At least one of the flight is delayed=0.04+2*0.16=0.36

Both flights are delayed=0.2*0.2=0.04

Example 4 - solution:

$0.8 \times 0.8 =$	**0.64** chance of **no delays**
$0.2 \times 0.8 =$	**0.16** chance of 1st flight delayed
$0.8 \times 0.2 =$	**0.16** chance of return flight delayed
$0.2 \times 0.2 =$	**0.04** chance of both flights delayed

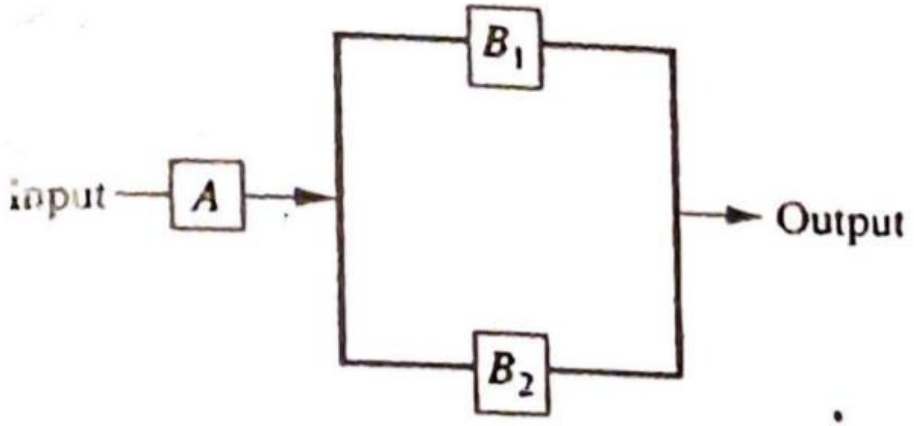
a. 0.64

b. $0.16 + 0.16 + 0.04 = 0.36$

c. 0.04

Example 5:

Example 4.13 An electrical device requires that two linked subsystems function. The following schematic shows that A must function and that at least one of the two B's must function.



Assume that the B components function independently of A and of each other. The *reliability* (probability of functioning) of A is 0.9 and that of each B is 0.8. Calculate the reliability of the device.

Example 5 - solution:

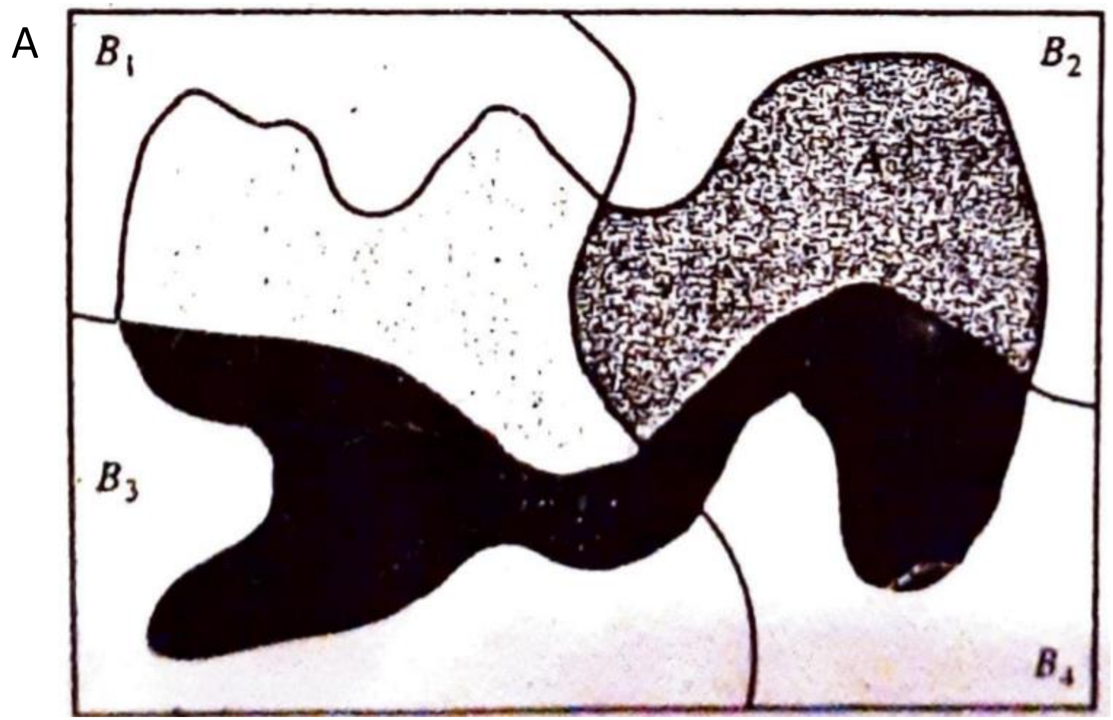
Solution. There are several ways to compute the device reliability. The most straightforward approach is to list the outcomes in the sample space and add the probabilities of the outcomes that allow the system to function. Let $\overline{A}(\overline{B})$ designate that component $A(B)$ has failed. The sample space is made up of eight outcomes as follows.

	Outcome	Probability	Outcome	Probability
•	AB_1B_2	$(0.9)(0.8)(0.8) = 0.576$	$\overline{A}B_1B_2$	$(0.1)(0.8)(0.8) = 0.064$
•	$AB_1\overline{B}_2$	$(0.9)(0.8)(0.2) = 0.144$	$\overline{A}B_1\overline{B}_2$	$(0.1)(0.8)(0.2) = 0.016$
•	$A\overline{B}_1B_2$	$(0.9)(0.2)(0.8) = 0.144$	$\overline{A}\overline{B}_1B_2$	$(0.1)(0.2)(0.8) = 0.016$
	$A\overline{B}_1\overline{B}_2$	$(0.9)(0.2)(0.2) = 0.036$	$\overline{A}\overline{B}_1\overline{B}_2$	$(0.1)(0.2)(0.2) = 0.004$

Bayes formula

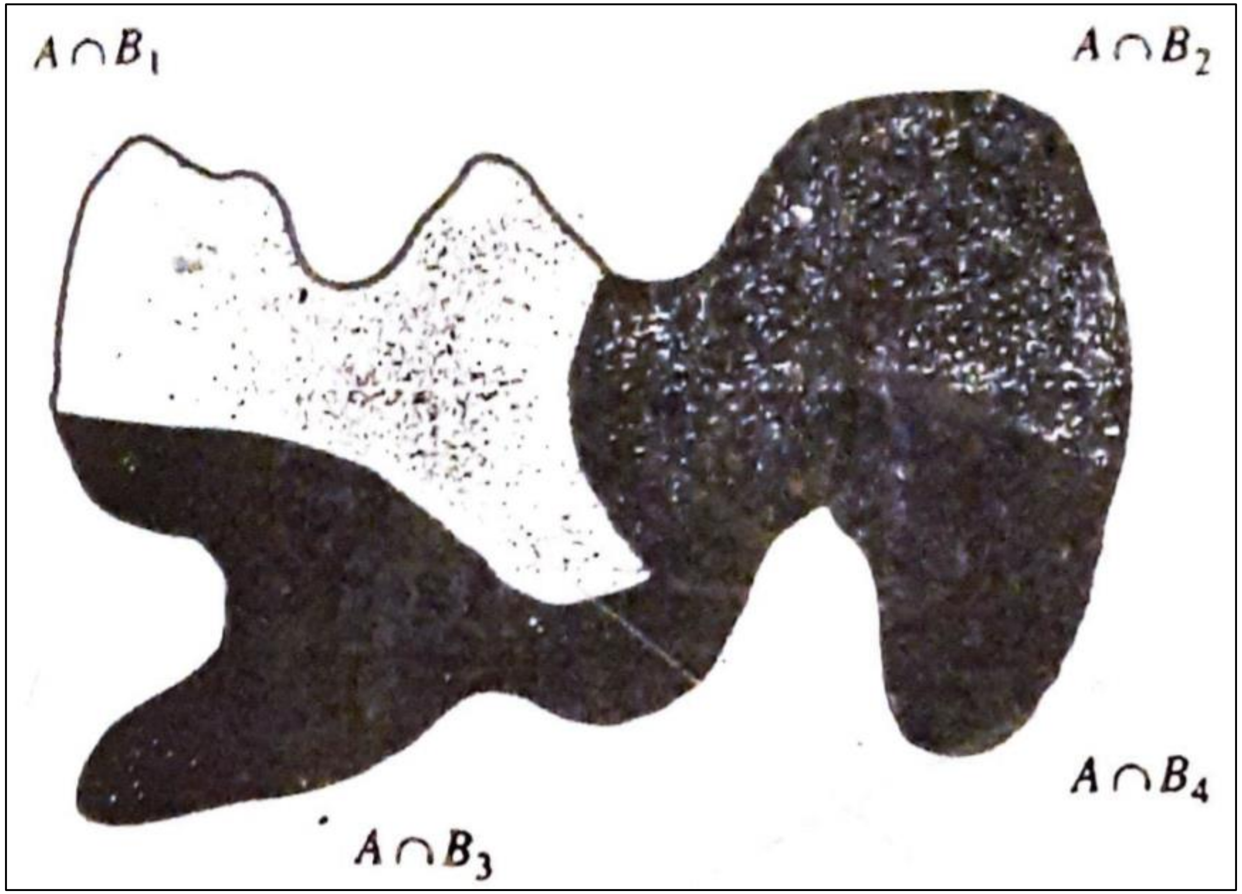
Bayes' formula is a natural outgrowth of conditional probability and the general rule of multiplication. Suppose you have a group of events, $B_1, B_2, \ldots, B_n$, that are *mutually exclusive and exhaustive*; i.e., the B_i's have no outcomes in common and, together, contain all the outcomes in the sample space. In mathematical terms, the mutually exclusive B_i's are exhaustive if and only if $\sum_{i=1}^{n} P(B_i) = 1$. Further, suppose that another event A is defined on the same sample space. Since the B_i's are exhaustive, A must intersect with one or more of the B_i's. Therefore, one way to obtain the probability of A would be to sum the probabilities of $A \cap B_i$ over all values of i: $\sum_{i=1}^{n} P(A \cap B_i) = P(A)$.

Bayes formula



Notice that A consists wholly of $A \cap B_1$, $A \cap B_2$, $A \cap B_3$, and $A \cap B_4$.
Thus $P(A) = P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3) + P(A \cap B_4)$.
that $P(A) = \sum_{i=1}^{4} P(A \cap B_i) = \sum_{i=1}^{4} P(B)_i)P(A \mid B_i)$.

Bayes formula



$$P(B_i \mid A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(B_i)P(A \mid B_i)}{\sum_{j=1}^{n} P(B_j)P(A \mid B_j)} = \frac{P(B_i)P(A \mid B_i)}{\sum_{j=1}^{n} P(B_j \cap A)} \quad (3)$$

Equation 3 is known as *Bayes' formula*.

Example 1:

Example 4.15 Vendors I, II, III, and IV provide all the bushings that the Acme Axle Co. purchases in the amounts 25%, 35%, 10%, 30%, respectively. It is known from long experience that vendors I, II, III, and IV provide 80%, 95%, 70%, and 90% good components. What is the probability that a randomly selected bushing is bad? Given that a bushing is bad, what is the probability it came from vendor III?

Example 1 - solution:

Solution. Let A represent the selection of a bad bushing, and let B_1, B_2, B_3, and B_4 represent the selection of a bushing from vendors I, II, III, and IV, respectively.

$$P(A) = \sum_{i=1}^{4} P(A \cap B_i) = \sum_{i=1}^{4} P(B_i)P(A \mid B_i)$$

Example 1 - solution:

Solution. Let A represent the selection of a bad bushing, and let B_1, B_2, B_3, and B_4 represent the selection of a bushing from vendors I, II, III, and IV, respectively.

$$P(A) = \sum_{i=1}^{4} P(A \cap B_i) = \sum_{i=1}^{4} P(B_i)P(A \mid B_i)$$

$$= 0.25(0.2) + 0.35(0.05) + 0.1(0.3) + 0.3(0.1)$$

$$= 0.1275$$

Therefore;

$$P(B_3 \mid A) = \frac{P(B_3 \cap A)}{P(A)} = \frac{0.03}{0.1275} = 0.2353$$

Example 2:

An individual has 3 different email accounts. Most of her messages, in fact 70% come into account #1, whereas 20% come into account#2 and the remaining 10% into account #3.

Of the messages into account #1, only 1% are spam, whereas the corresponding percentages for accounts #2 and #3 are 2% and 5% respectively.

What is the probability that a randomly selected message is a spam

What is the probability that a randomly selected message is from account 2 given that it is a spam

Example 2:

An individual has 3 different email accounts. Most of her messages, in fact 70% come into account #1, whereas 20% come into account#2 and the remaining 10% into account #3.

Of the messages into account #1, only 1% are spam, whereas the corresponding percentages for accounts #2 and #3 are 2% and 5% respectively.

What is the probability that a randomly selected message is a spam

$P(S)=0.7*0.01+0.2*0.02+0.1*0.05=0.016$

What is the probability that a randomly selected message is from account 2 given that it is a spam

$P(II/S)=0.2*0.02/0.016=0.25$