HCMC University of Technology

Dung Nguyen

Probability and Statistics

Probability (review)

Terminology

- Trial
- Outcome
- Sample space: discrete, continuous.
- Event

Operations

Conjunctions	NOT	AND	OR	BUT
Operators	$\overline{A}$	AB	$A+B$	$A-B$

Proposition (De Morgan's Rules)

1. $\overline{A+B} = \overline{A}\ \overline{B}$.
2. $\overline{AB} = \overline{A} + \overline{B}$.

Properties of Probability

1. $0 \leq P(A) \leq 1$.
2. $P(A|B) = \dfrac{P(AB)}{P(B)}$.
3. Disjoint $\neq$ Independent:

Disjoint	**Independent**
$AB = \emptyset$	$P(AB) = P(A)P(B)$
	$P(A\|B) = P(A)$
	$P(B\|A) = P(A)$

4. Mutually independent $\Longrightarrow$ Pairwise independent

Pairwise independent	**Mutually independent**
$P(A_iA_j) = P(A_i)P(A_j)$	$P(A_iA_j) = P(A_i)P(A_j)$ $P(A_iA_jA_k) = P(A_i)P(A_j)P(A_k)$ ... $P(A_1A_2\dots A_n) = P(A_1)P(A_2)\cdots P(A_n)$

for $i, j, k, \dots$ distinct.

5. Complement: $\boxed{P(\overline{A}) = 1 - P(A)}$

6. Addition:

The total probability formula & Bayes's formula

	2 events	n events
The total probability	$P(A) = P(A \mid B)P(B) + P(A \mid \overline{B})P(\overline{B})$	$P(A) = \sum_{i=1}^{k} P(A \mid B_i)P(B_i)$
Bayes's Formula	$P(B \mid A) = \dfrac{P(A \mid B)P(B)}{P(A \mid B)P(B) + P(A \mid \overline{B})P(\overline{B})}$	$P(B_j \mid A) = \dfrac{P(A \mid B_j)P(B_j)}{\sum_{i=1}^{k} P(A \mid B_i)P(B_i)}$

The total probability formula & Bayes's formula under condition C

	2 events	n events
The total probability	$P(A \mid C) = P(A \mid BC)P(B \mid C) + P(A \mid \overline{B}C)P(\overline{B} \mid C)$	$P(A \mid C) = \sum_{i=1}^{k} P(A \mid B_iC)P(B_i \mid C)$
Bayes's Formula	$P(B \mid AC) = \dfrac{P(A \mid BC)P(B \mid C)}{P(A \mid BC)P(B \mid C) + P(A \mid \overline{B}C)P(\overline{B} \mid C)}$	$P(B_j \mid AC) = \dfrac{P(A \mid B_jC)P(B_j \mid C)}{\sum_{i=1}^{k} P(A \mid B_iC)P(B_i \mid C)}$

Example 1 - Passengers

Phoenix is a hub for a large airline. Suppose that on a particular day, 8000 passengers arrived in Phoenix on this airline. Phoenix was the final destination for 1800 of these passengers. The others were all connecting to flights to other cities. On this particular day, several inbound flights were late, and 480 connecting passengers missed their connecting flight and were delayed in Phoenix. Of the 480 who were delayed, 75 were delayed overnight and had to spend the night in Phoenix. Consider the chance experiment of choosing a passenger at random from these 8000 passengers. Compute the probability that the selected passenger:

- (a) had Phoenix as a final destination.
- (b) did not have Phoenix as a final destination.
- (c) was connecting and missed the connecting flight.
- (d) was a connecting passenger and did not miss the connecting flight.
- (e) either had Phoenix as a final destination or was delayed overnight in Phoenix.

Example 2 - Accident reports

A certain factory operates three different shifts. Over the last year, 200 accidents have occurred at the factory. Some of these can be attributed at least in part to unsafe working conditions, whereas the others are unrelated to working conditions. The accompanying table gives the percentage of accidents falling in each type of accident-shift category.

		Unsafe Conditions	Unrelated to Conditions
Shift	Day	10%	35%
	Swing	8%	20%
	Night	5%	22%

Suppose one of the 200 accident reports is randomly selected from a file of reports, and the shift and type of accident are determined. What is the probability that the selected accident

- (b) did not occur on the day shift?

Example 3 – Engine

Each of 311 people who purchased a Honda Civic was classified according to gender and whether the car purchased had a hybrid engine or not.

	Hybrid	Not hybrid
Male	77	117
Female	34	83

Suppose one of these 311 individuals is to be selected at random. Find the following probabilities:

(a) P(male).

(b) P(hybrid).

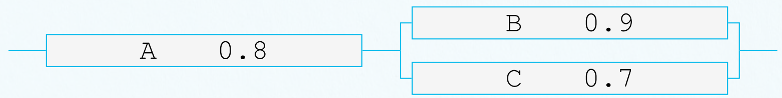
(c) P(hybrid | male).

(d) P(hybrid | female).

(e) P(female | hybrid).

Example 4 – Electrical system

An electrical system consists of three components as illustrated in the figure. The reliability (probability of working) of each component is also shown in the figure.



Find the probability that

(a) the entire system works.

(b) the component A works, given that the entire system works. Assume that the components work independently.

Example 5 – Ace

An ordinary deck of 52 playing cards is randomly divided into 4 piles of 13 cards each. Compute the probability that each pile has exactly 1 ace.

Example 6 –

Disks of polycarbonate plastic from a supplier are analyzed for scratch and shock resistance. The results from 100 disks are summarized as follows:

	Shock Resistance	
Scratch Resistence	High	Low
High	70	9
Low	16	5

Let A denote the event that a disk has high shock resistance, and let B denote the event that a disk has high scratch resistance. Are events A and B independent?

Example 7 – Independence

Let $\Omega = \{1,2,3,4\}$ and $p(i) = \frac{1}{4}$ where $i = 1,\ldots,4$. Denote $A = \{1,2\}$, $B = \{1,3\}$, $C = \{2,3\}$. Then A, B, C are pairwise independent but are not independent.

Example 8 – E-mail filter

An e-mail filter is planned to separate valid e-mails from spam. The word "free" occurs in 60% of the spam messages and only 4% of the valid messages. Also, 20% of the messages are spam. Determine the following probabilities:

(a) The message contains "free".

(b) The message is spam given that it contains "free".

(c) The message is valid given that it does not contain "free".

Example 9 – Who did the work?

A construction company employs two sales engineers. Engineer 1 does the work of estimating cost for 70% of jobs bid by the company. Engineer 2 does the work for 30% of jobs bid by the company. It is known that the error rate for engineer 1 is such that 0.02 is the probability of an error when he does the work, whereas the probability of an error in the work of engineer 2 is 0.04. Suppose a bid arrives and a serious error occurs in estimating cost. Which engineer would you guess did the work?

Example 10 – Short bolts & long bolts

In box 1, there are 60 short bolts and 40 long bolts. In box 2, there are 10 short bolts and 20 long bolts.

(a) Take a box at random, and pick a bolt. What is the probability that you chose a short bolt?

(b) A bolt is randomly selected from box 1 and transferred to box 2. Then select at random a bolt from box 2. What is the probability that the bolt chose from box 2 is a short bolt?

Example 11 – Deer ticks

Deer ticks can be carriers of either Lyme disease or human granulocytic ehrlichiosis (HGE). Based on a recent study, suppose that 16% of all ticks in a certain location carry Lyme disease, 10% carry HGE, and 10% of the ticks that carry at least one of these diseases in fact carry both of them. Select a tick at random.

(a) What is the probability that the tick is found to have carried HGE?

(b) If the tick is found to have carried HGE, what is the probability that the selected tick is also a carrier of Lyme disease?

Example 12 –

Repeat rolling a fair die. What is the probability that

(a) a run of 5 consecutive #1 occurs?

(b) a run of 5 consecutive #1 occurs before a run of 7 consecutive something else?

Example 13 –

Two different airlines have a flight from Los Angeles to New York that departs each weekday morning at a certain time. Let E denote the event that the first airline's flight is fully booked on a particular day, and let F denote the event that the second airline's flight is fully booked on that same day. Suppose that $P(E) = 0.8$, $P(F) = 0.6$, and $P(EF) = 0.49$.

(a) Calculate $P(E|F)$ the probability that the first airline's flight is fully booked given that the second airline's flight is fully booked.

(b) Calculate $P(F|E)$.

Example 14 –

Approximately 30% of the calls to an airline reservation phone line result in a reservation being made.

(a) Suppose that an operator handles 10 calls. What is the probability that none of the 10 calls results in a reservation?

(b) What assumption did you make in order to calculate the probability in Part (a)?

(c) What is the probability that at least one call results in a reservation being made?

Example 15 –

Consider the summary data for the men who participated in the study

Supplier	Survived	Died	Total
Treatment A	120	80	200
Treatment B	20	20	40
Total	140	100	

ⓐ Find $P(S)$

ⓑ Find $P(S|A)$

ⓒ Find $P(S|B)$

ⓓ Which treatment appears to be better?

Example 16 – Recreational equipment

A recreational equipment supplier finds that among orders that include tents, 40% also include sleeping mats. Only 5% of orders that do not include tents do include sleeping mats. Also, 20% of orders include tents. Determine the following probabilities:

ⓐ The order includes sleeping mats.

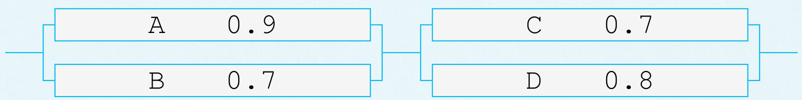
ⓑ The order includes a tent given it includes sleeping mats.

Example 17 –

Two Web colors are used for a site advertisement. If a site visitor arrives from an affiliate, the probabilities of the blue or green colors being used in the advertisement are 0.8 and 0.2, respectively. If the site visitor arrives from a search site, the probabilities of blue and green colors in the advertisement are 0.4 and 0.6, respectively. The proportions of visitors from affiliates and search sites are 0.3 and 0.7, respectively. What is the probability that a visitor is from a search site given that the blue ad was viewed?

Example 18 – Electrical system

An electrical system consists of three components as illustrated in the figure. The reliability (probability of working) of each component is also shown in the figure.



Find the probability that

ⓐ the entire system works.

ⓑ the component A works, given that the entire system works. Assume that the components work independently.

Example 19 –

Samples of emissions from three suppliers are classified fro conformance to air-quality specifications. The results from 100 samples are summarized as follows

	Conforms	
Supplier	Yes	No
1	22	8
2	25	5
3	30	10

Let A denote the event that a sample is from supplier 1, and let B denote the event that a sample conforms to a specifications. Are events A and B indepenent?

Example 20 – Independence

Suppose two fair coins are flipped. Let

- A be the events "first coin shows heads"
- B be the event "second coin shows heads"
- C be the event "both coins show heads or both coins show tails."

Prove that A, B, C are not independent.

Example 21 – Independence

Now suppose that the die is loaded such that the four outcomes 2, 3, 5, and 30 have probabilities 11/24, 7/24, a, and b respectively. A, B, and C are events that the outcome is a multiple of 2, 3, and 5 respectively.

a. Are A and B independent?
b. Assume that $a = 4/24$. Are A and C independent?
c. Determine a such that A and C are independent? Are A, B, C independent?

Example 22 – Cost overrun

A certain federal agency employs three consulting firms (A, B, and C) with probabilities 0.350, 0.40, and 0.25, respectively. From past experience it is known that the probabilities of cost overruns for the firms are 0.05, 0.25, and 0.15, respectively.

a. What is the probability of cost overrun?
b. Suppose a cost overrun is experienced by the agency. What is the probability that the consulting firm involved is company C?

Example 23 - Disposable flashlight

A bin contains 3 different types of disposable flashlights. The probability that a type 1 flashlight will give over 100 hours of use is 0.7, with the corresponding probabilities for type 2 and type 3 flashlights being 0.4 and 0.3, respectively. Suppose that $(3m+n+10)$ percent of the flashlights in the bin are type 1, $(40-3m-n)$ percent are type 2, and 50 percent are type 3. What is the probability that a randomly chosen flashlight will give more than 100 hours of use? Given that a flashlight lasted over 100 hours, what is the conditional probability that it was a type 1 flashlight?