COUNTING (RECAP)& PROBABILITY

Summer 21

Asst. Prof. Nattapol Aunsri, Ph.D.
IATE Research Unit
CCC Research Group
School of Information Technology, MFU

PREPARED FOR DIGITAL INDUSTRY INTEGRATION: DII, CAMT, CMU

COUNTING RECAP

- *If an experiment E has k subexperiments E1, E2,… , Ek where Ei has n_i outcomes, then E has $\prod_{i=1}^{k} n_i$*

- *The number of k-permutations of n distinguishable objects is*

$$nPk = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$

- ***Sampling WITHOUT replacement (Combination, order is not matter)***

The number of ways to choose k objects out of n distinguishable objects is

$$nCk = \binom{n}{k} = \frac{nPk}{k!} = \frac{n!}{k!\,(n-k)!}$$

- ***Sampling WITH replacement***

Each object can be chosen repeatedly

COUNTING RECAP

Examples

- *Number of picking 3 students from 5 students is* $\binom{5}{3}$
- *A baseball team has 15 field players and 10 pitchers. Each field player can take any of the 8 non pitching positions and one pitcher. Therefore, the number of possible starting lineups is* $N = \binom{15}{8}\binom{10}{1}$*=64350.*
- *The number of ways to draw 7 cards from a deck is* $\binom{52}{7}$*.*
- *The number of ways to draw 7 cards from a deck without any queens is* $\binom{48}{7}$*.*

COUNTING RECAP

- *Given m distinguishable objects, there are m^n ways to choose with replacement an ordered sample of n objects*

- *For n repetitions of a subexperiment with sample space $S = \{s_1, s_2, \dots, s_m\}$, there are m^n possible observation sequences.*

- *The number of observation sequences for n subexperiments with sample space $S = \{0,1\}$ with 0 appears n_0 times and 1 appears n_1 times is*

$$\binom{n}{n_0} = \binom{n}{n - n_1}$$

Example

For five subexperiments with sample space $S = \{0,1\}$, how many observation sequences are there in which 0 appears $n_0 = 2$ times and 1 appears $n_1 = 3$ times?

Ans *5C3 or 5C2*

COUNTING RECAP

- *For n repetitions of a subexperiment with sample space $S = \{s_1, s_2, \dots, s_m\}$, the number of length $n = n_1 + n_2 + \cdots n_m$ observation sequences with s_i appearing n_i times is*

$$\binom{n}{n_1, \cdots, n_m} = \frac{n!}{n_1! n_2! \cdots, n_m!}$$

Example

Consider a binary code with 4 bits (0 or 1) in each code word. An example of a code word is 1001.

a. *How many different code words are there?*

b. *How many code words have exactly two zeroes?*

c. *How many code words begin with a zero?*

d. *In a constant-ration binary code, each code word has N bits. In very word, M of the N bits are 1 and the other $N - M$ bits are 0. How many different code words are in the code with $N = 8$ and $M = 3$?*

PROBABILITY AXIOMS

- ***Definition*** *Axioms of Probability*

A probability measure $P[\cdot]$ is a function that maps events in the sample space to real numbers such that

1. *for any event A, $P[A] \geq 0$.*
2. $P[S] = 1$.
3. *For any countable collections $A_1, A_2, \ldots$ of mutually exclusive events*

$$P[A_1 \cup A_2 \cup \cdots] = P[A_1] + P[A_2] + \cdots.$$

Consequent theorems

- *If $A = A_1 \cup A_2 \cup \cdots \cup A_m$ and$A_i \cap A_j = \emptyset$ for $i \neq j$. Then*

$$P[A] = \sum_{i=1}^{m} P[A_i].$$

PROBABILITY

Example

- *Roll a six-sided die in which all faces are equally likely. What is the probability of each outcome? Find the probabilities of the events: "Roll 4 or higher." and "Roll the square of an integer"*

Solution

$$P[i] = \tfrac{1}{6}, for\ i = 1,2,3,4,5,6$$

$$P[\text{Roll 4 or higher}] = P[4] + P[5] + P[6] = \frac{3}{6} = \frac{1}{2}.$$

$$P[\textit{Roll the square of an integer}] = P[1] + P[4] = \tfrac{2}{6} = \tfrac{1}{3}$$

PROBABILITY

- ***Theorem*** *consequence of axioms of Probability*

A probability measure $P[\cdot]$ satisfies

1. $P[\emptyset] = 0.$
2. $P[A^C] = 1 - P[A].$
3. *For any A and B (not necessarily disjoint),*

$$P[A \cup B] = P[A] + P[B] - P[A \cap B].$$

4. *If $A \subset B$, then $P[A] \leq P[B]$.*

PROBABILITY

Example

Monitor a phone call. Classify as a voice call (V) if someone is speaking, or a data call(D) if using data (fb, line chat, for eg.) Classify the call as long (L) if the call lasts for more than three minutes; otherwise classify as brief (B). Based on the data collected by the telephone company, we use the following probability model: $P[V] = 0.7, P[L] = 0.6, P[VL] = 0.35$. *Find the following probabilities.*

- $P[DL]$
- $P[D \cup L]$
- $P[VB]$
- $P[V \cup L]$
- $P[D \cup V]$
- $P[LB]$

Solution

	V	D
L	0.35	?
B	?	?

CONDITIONAL PROBABILITY

- ***Definition*** *The conditional probability of the event A given the occurrence of the event B is*

$$P[A|B] = \frac{P[AB]}{P[B]}$$

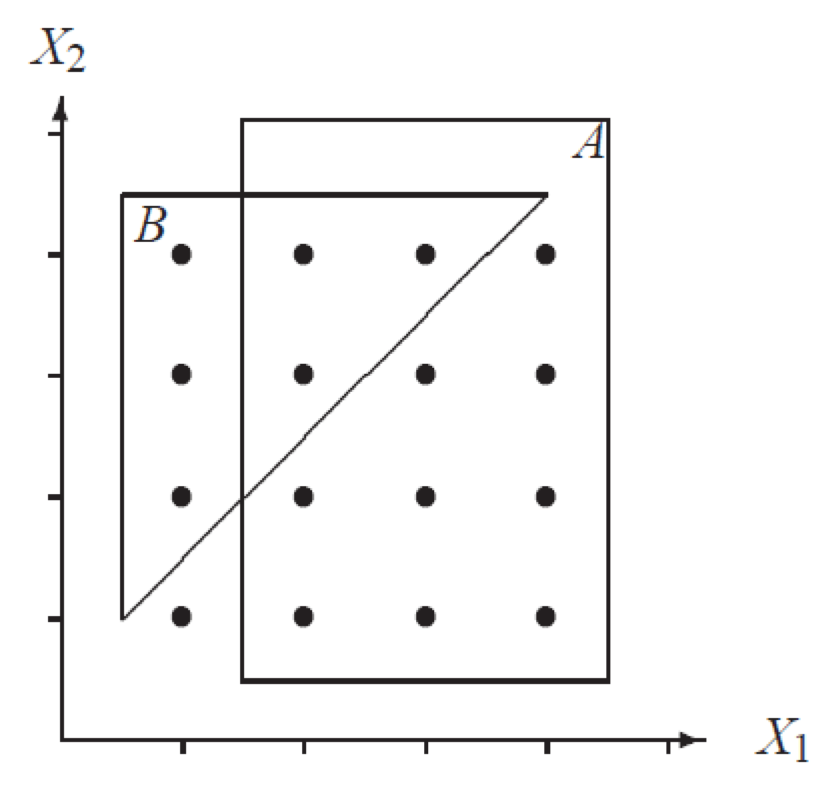
Example

Roll two fair four-sided dice. Let X_1 and X_2 denote the number of dots that appear on die 1 and die 2, respectively. Let A be the event $X_1 \geq 2$. What is $P[A]$? Let B denote the event$X_2 > X_1$. What is $P[B]$ and $P[A|B]$

CONDITIONAL PROBABILITY

Solution

It is easy to solve the problem using the diagram below.



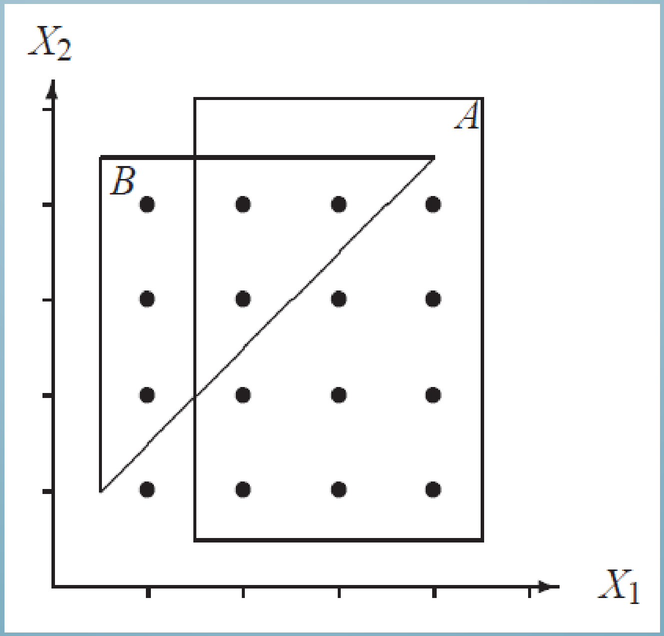
CONDITIONAL PROBABILITY

Solution (cont'd)

$$P[A] = \frac{12}{16} = \frac{3}{4}.$$

$$P[B] = \frac{6}{16} = \frac{3}{8}.$$

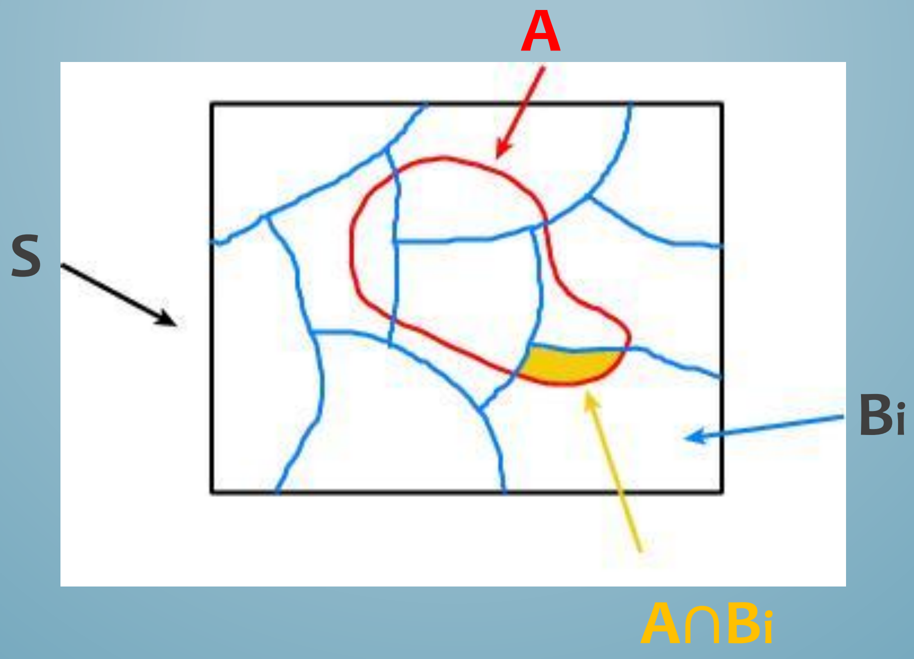
$$P[AB] = \frac{3}{16} \rightarrow P[A|B] = \frac{P[AB]}{P[B]} = \frac{3/16}{3/8} = \frac{1}{2}.$$



LAW OF TOTAL PROBABILITY

- ***Theorem*** *For an event space* $\{B_1, B_2, \ldots, B_n\}$ *with* $P[B_i] > 0$ *for all* i, *then A probability measure* $P[\cdot]$ *satisfies*

$$P[A] = \sum_{i=1}^{n} P[A|B_i]P[B_i]\,.$$



BAYES' THEOREM

- ***Bayes' Theorem***

$$P[B|A] = \frac{P[A|B]P[B]}{P[A]}$$

Using the law of total probability, we have

$$P[B_i|A] = \frac{P[A|B_i]P[B_i]}{\sum_{i=1}^{n} P[A|B_i]P[B_i]}$$

AN EXAMPLE USING BAYES' THEOREM

Suppose the probability (for anyone) to have AIDS is:

$$P(\text{AIDS}) = 0.001$$
$$P(\text{no AIDS}) = 0.999$$

← prior probabilities, i.e., before any test carried out

Consider an AIDS test: result is + or -

$$P(+|\text{AIDS}) = 0.98$$
$$P(-|\text{AIDS}) = 0.02$$

← probabilities to (in)correctly identify an infected person

$$P(+|\text{no AIDS}) = 0.03$$
$$P(-|\text{no AIDS}) = 0.97$$

← probabilities to (in)correctly identify an uninfected person

Suppose your result is +. How worried should you be?

BAYES' THEOREM EXAMPLE (CONT.)

The probability to have AIDS given a + result is

$$P(\text{AIDS}|+) = \frac{P(+|\text{AIDS})P(\text{AIDS})}{P(+|\text{AIDS})P(\text{AIDS}) + P(+|\text{no AIDS})P(\text{no AIDS})}$$

$$= \frac{0.98 \times 0.001}{0.98 \times 0.001 + 0.03 \times 0.999}$$

$= 0.032$ $\leftarrow$ posterior probability

i.e. you're probably OK!

Your viewpoint: my degree of belief that I have AIDS is 3.2%

Your doctor's viewpoint: 3.2% of people like this will have AIDS

SEQUENTIAL EXPERIMENTS AND TREE DIAGRAM

- The method to effectively deal with sequential experiments.
- Tree diagram is used represents the sequence of subexperiments.
- The labels of the branches of second (an so on) subexperiment are the conditional probabilities of the events in the second (and so on) subexperiments.
- Nodes at the end of final subexperiments are leaves of the tree corresponding to an outcome of the entire sequential experiment.
- Probability of each outcome is the product of the probabilities and conditional probabilities of the path from the root to the leaf.

SEQUENTIAL EXPERIMENTS AND TREE DIAGRAM

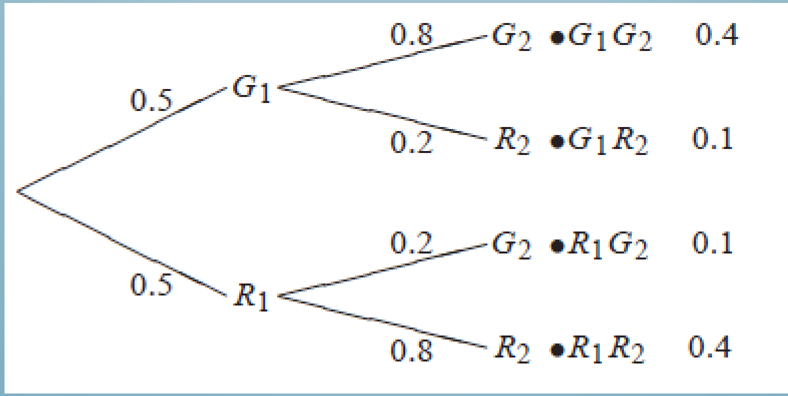
- *Example*

Suppose traffic engineers have coordinated the timing of two traffic lights to encourage a run of green lights. In particular, the timing was designed so that with probability 0.8 a driver will find the second light to have the same color as the first. Assuming the first light is equally likely to be red or green, what is the probability $P[G_2]$ that the second light is green? Also, what is $P[W]$, the probability that you wait for at least one light? Lastly, what is $P[G_1|R_2]$, the conditional probability of a green first light given a red second light?

SEQUENTIAL EXPERIMENTS AND TREE DIAGRAM

- *Example*

Suppose traffic engineers have coordinated the timing of two traffic lights to encourage a run of green lights. In particular, the timing was designed so that with probability 0.8 a driver will find the second light to have the same color as the first. Assuming the first light is equally likely to be red or green, what is the probability $P[G_2]$ that the second light is green? Also, what is $P[W]$, the probability that you wait for at least one light? Lastly, what is $P[G_1|R_2]$, the conditional probability of a green first light given a red second light?



SEQUENTIAL EXPERIMENTS AND TREE DIAGRAM

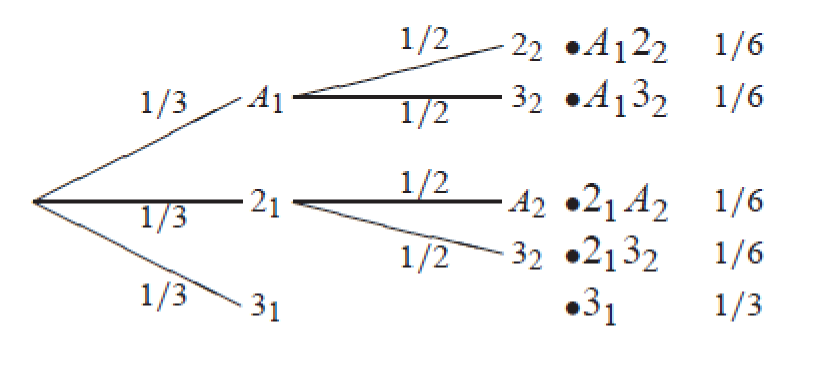
- *Example*

Consider the game of Three. You shuffle a deck of three cards: ace, 2, 3. With the ace worth 1 point, you draw cards until your total is 3 or more. You win if your total is 3. What is $P[W]$, the probability that you win?

SEQUENTIAL EXPERIMENTS AND TREE DIAGRAM

- *Example*

Consider the game of Three. You shuffle a deck of three cards: ace, 2, 3. With the ace worth 1 point, you draw cards until your total is 3 or more. You win if your total is 3. What is $P[W]$, the probability that you win?



THANKS