Conditional Probability

Example

AI

- Roll two dice, yielding values D_1 and D_2 (주사위 1, 2)
- Let E be event: $D_1 + D_2 = 4$
- What is Pr(E)?
  - $|\Omega| = \overset{6\times 6}{36}$, $E = \{(1, 3), (2, 2), (3, 1)\}$ $\rightarrow |E| = 3$
  - $Pr(E) = 3/36 = 1/12$

    $Pr(E) = \frac{|E|}{|\Omega|} = \frac{3}{36} = \frac{1}{12}$
- Let F be event: $D_1 = 2$
- Pr(E | F)?
  - $F = \{(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6)\}$
  - $E \cap F = \{(2, 2)\}$
  - $Pr(E \mid F) = 1/6$

    $Pr(E|F) = \frac{|E \cap F|}{|F|}$

Chain Rule

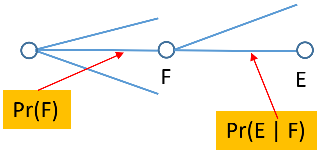
KENTECH Korea Institute of Energy Technology AI

$= P_r(F|E) \cdot P_r(E)$

- From $\Pr(E|F) = \frac{\Pr(E \cap F)}{\Pr(F)}$ ➔ $\boxed{\Pr(E \cap F) = \Pr(E \mid F) \bullet \Pr(F)}$

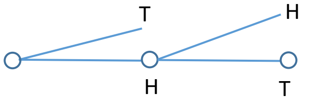
Chain Rule

- Description as a Sequential Tree



Pr(E∩F) = Pr(E | F) • Pr(F)

- If E and F are independent ➔ Pr(E|F) = Pr(E) $P_r(E \cap F) = P_r(E) \cdot P_r(F)$
  - Example: Given the first coin flip is heads, the second coin flip is tails

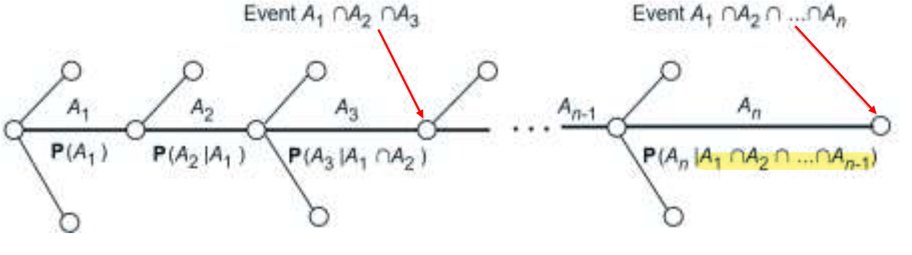


Multiplication Rule

KENTECH Korea Institute of Energy Technology AI

- Generalized chain rule (Or **Multiplication rule**)

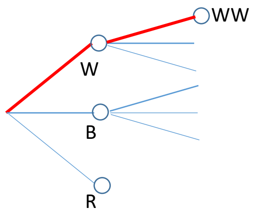
$\Pr(A_1 \cap A_2 \cap \ldots \cap A_n)$
$= \Pr(A_1)\Pr(A_2 \mid A_1)\Pr(A_3 \mid A_1 \cap A_2)\ldots\Pr(A_n \mid A_1 \cap A_2 \cap \ldots \cap A_{n-1})$



Example: Conditional Prob.

- Model a sequence of experiments as a tree
- Example
  - A bag contains 4 blue balls, 3 red balls and 5 white balls
  - Pick two balls from the bag. What is the prob. that picked balls are all white?

  ➔ Pick one ball and w/o replacement pick another ball sequentially. And two ball are all white.

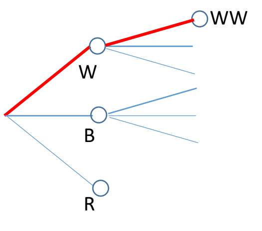


Event W1: First ball is White
Event W2: Second ball is White

Prob. of red edges = Pr(W1 ∩ W2)
= Pr(W1) ·Pr(W2 | W1)
= (5/12)·(4/11)

Example: Independent

- Example
  - A bag contains 4 blue balls, 3 red balls and 5 white balls
  - Pick one ball for the first time. Return the ball and pick one ball again. What is the prob. that all two balls are White?



Event W1: First ball is White
Event W2: Second ball is White

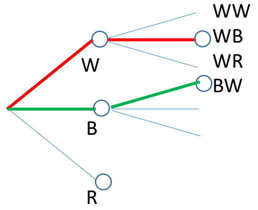
Prob. of red edges = Pr(W1 ∩ W2)
= Pr(W1)·Pr(W2)
= (5/12)·(5/12)

Example: Multiple Paths

KENTECH Korea Institute of Energy Technology AI

- Model a sequence of experiments as a tree
- Example
  - A bag contains 4 blue balls, 3 red balls and 5 white balls
  - Pick two balls from the bag. What is the prob. that 1 ball is white and 1 ball is blue?

  ➔ Pick one ball and w/o replacement pick another ball sequentially. What is the probability to the first pick is white ball and the second is blue ball?



Event W1: First ball is White
Event B2: Second ball is Blue

Prob. of red edges = Pr(W1 ∩ B2)
= Pr(W1) ·Pr(B2 | W1)
= (5/12)·(4/11)
Prob. of green edges = (4/12)·(5/11)

➔ What is the prob. that the first ball is White and the second is Blue?

Polynomial Identities: Revisit

KENTECH Korea Institute of Energy Technology AI

- Let F(x) ≠ G(x)
- Randomized algorithm: Perform k trials and decide F(x)=G(x) if all trials claim F(x)=G(x)

- With replacement
  - Select r_i uniformly at random repeatedly from **R** = {1, 2, …, 100d}
  - Return r_i to **R** after the trial
  - Let Fi be an event that i-th trial fails ➔ $F(r_i) = G(r_i)$
  - $Pr(F_1) = Pr(F_2) = … = Pr(F_k) \le 1/100$
  - Pr(Randomized algorithm fails) = $Pr(F_1 \cap F_2 \cap … \cap F_k) \le (\frac{1}{100})^k$

Polynomial Identities: Revisit

- Without replacement
  - Discard r_i after the i-th trial
  - After the i-th trial, there are 100d-i elements in **R** and at most d-i roots in **R**

  ➔ $\Pr(F_i \mid F_1 \cap F_2 \cap \ldots \cap F_{i-1}) \le \frac{d-(i-1)}{100d-(i-1)}$
  - Pr(Randomized algorithm fails)

  $= \Pr(F_1 \cap F_2 \cap \ldots \cap F_k)$

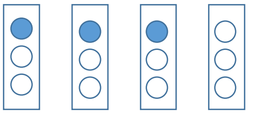
  $= \Pr(F_1)\Pr(F_2 \mid F_1)\Pr(F_3 \mid F_1 \cap F_2)\ldots\Pr(F_k \mid F_1 \cap F_2 \cap \ldots \cap F_{k-1})$

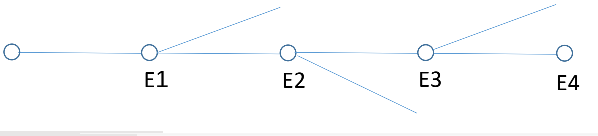
  $\le \prod_{i=1}^{k} \frac{d-(i-1)}{100d-(i-1)} < (\frac{1}{100})^k$

Only **SLIGHTLY** better than with replacement algorithm

Example: Project Team

- With 12 students and make four project teams each of which consists of three randomly selected students
- Family name distribution: 4 Kim’s (Let AKim, BKim, CKim, DKim) and 8 other surnames
- Probability that each team has exactly one Kim
- Solution
  - E1: AKim is in any one team
  - E2: AKim and BKim in different teams
  - E3: AKim, BKim and CKim in different teams
  - E4: AKim, Bkim, Ckim, and DKim in different teams
  - $\Pr(E4 \mid E1 \cap E2 \cap E3) = 3/9$

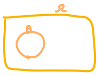




Total Probability Theorem

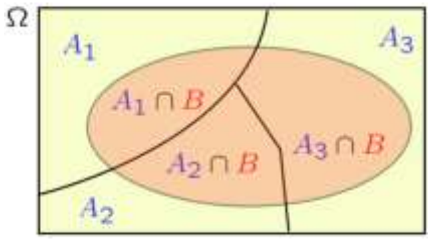
KENTECH Korea Institute of Energy Technology AI

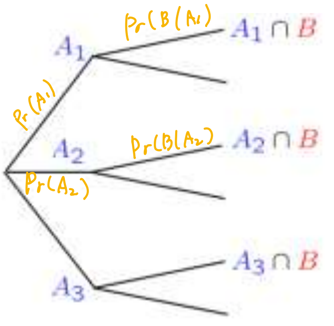
- Consider events, $\boxed{A_1, A_2, \ldots, A_n}$ Partition
  - Mutually disjoint: For any, $1 \leq j \neq k \leq n$, $A_j \cap A_k = \emptyset$
  - Exhaustive: $A_1 \cup A_2 \cup \ldots \cup A_n = \boxed{\Omega}$
    오메가



$B = B \cap \Omega$
$= B \cap (A_1 \cup A_2 \cup \cdots \cup A_n)$
$= (B \cap A_1) \cup (B \cap A_2) \cup \cdots (B \cap A_n)$
$= Pr(B \cap A_1) + \cdots + Pr(B \cap A_n)$

- $\Pr(B) = \Pr(B \cap A_1) + \Pr(B \cap A_2) + \cdots + \Pr(B \cap A_n)$
  $= \Pr(B|A_1) \cdot \Pr(A_1) + \Pr(B|A_2) \cdot \Pr(A_2) + \cdots + \Pr(B|A_n) \cdot \Pr(A_n)$





Bayses' Theorem (Law/Rule)

KENTECH Korea Institute of Energy Technology AI

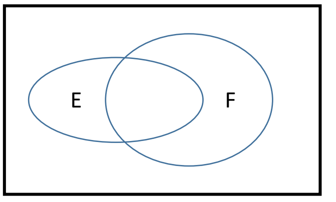
- Rev. Thomas Bayes (1702-1761) was a British minister

"Probability & Computing"

Bayes' Theorem

AI

Theorem 1.6:
Law of Total Probability



$\Omega = F \cup \bar{F}$

- $E = (E \cap F) \cup (E \cap \bar{F})$ Note $(E \cap F) \cap (E \cap \bar{F}) = \emptyset$
- $\Pr(E) = \Pr(E \cap F) + \Pr(E \cap \bar{F})$

  $= \Pr(E \mid F)\Pr(F) + \Pr(E \mid \bar{F})\Pr(\bar{F})$
- $\Pr(F \mid E) = \dfrac{Pr(E \cap F)}{\Pr(E)} = \dfrac{Pr(E \mid F)\Pr(F)}{\Pr(E)}$

  $= \dfrac{Pr(E \mid F)\Pr(F)}{\Pr(E \mid F)\Pr(F) + \Pr(E \mid \bar{F})\Pr(\bar{F})}$
- More generally,
  - Let $F_1, F_2, \ldots F_n$ be mutually exclusive and exhaustive events
  - Given E observed, want to determine which of F_j also occurred

  $\Pr(F_j \mid E) = \dfrac{Pr(E \mid F_j)\Pr(F_j)}{\sum_{i=1}^{n} Pr(E \mid F_i)\Pr(F_i)}$

Spam Email

AI

- Frequently used words and phrases in spam email
  - "Dear Friend", "Prize", "Make Money Fast(MMF)", "Hot", "Million", …
- 60% of all emails are spam → $P(F)=0.6$, $P(\bar{F})=0.4$
  - 50% of spams have MMF → $Pr(E \mid F)=0.5$
  - 10% of non-spams have MMF → $Pr(E \mid \bar{F})=0.1$
- An email has MMF. What is the probability that the email is spam?
  - E: Email has MMF
  - F: Email is spam

$\boxed{F / \bar{F}}\ \Omega$

$\Pr(F|E) = \dfrac{Pr(E \mid F)\Pr(F)}{\Pr(E \mid F)\Pr(F) + \Pr(E \mid \bar{F})\Pr(\bar{F})}$

Learn: Naïve Bayesian Filtering (NBF)

Another Example

KENTECH Korea Institute of Energy Technology AI

- Three coins
  - Two of them are un-biased and one is biased such that Pr(Heads) = 2/3
  - Flip three coins in a random order and found that first and second coins are heads and third is tails
  - Compute the probability that the first coin is the biased coin
- Solution
  - Observed event: (H,H,T)
  - F1: First coin is biased, (similarly F2, F3)
  - Pr(F1 | (H, H, T)) = ?

Important:
- **Sample space**
- **How to partition the sample space?**
- **Define events with proper symbols**

YAE: Mamma Mia

- Child is born with (A, a) gene pair
  (Event (A,a))
  - Mother has (A, A) gene pair
  - Two possible fathers:
    - Adam: (a,a), Bob: (A,a)

      → (A,a)(A,a) ⇒ 1 → (A,a) (A,A) ⇒ $\frac{1}{2}$
  - Mother's belief: Pr(Adam) = p, Pr(Bob) = (1-p)
  - What is probability that the father is Adam?
  - Pr(Adam | (A,a)) = ?

$Pr((A,a) \mid Adam) = 1$

$Pr((A,a) \mid Bob) = \frac{1}{2}$

Probability Inference

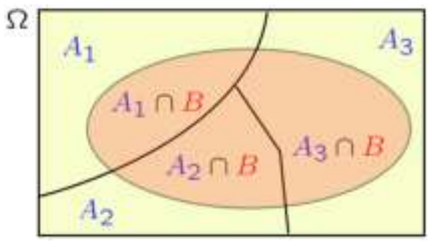
AI

- Bayes' Theorem

Posterior — Likelihood — Prior

$$\Pr(F \mid E) = \frac{Pr(E \mid F)\Pr(F)}{\Pr(E)}$$

- Probability changes after evidences (E) are observed



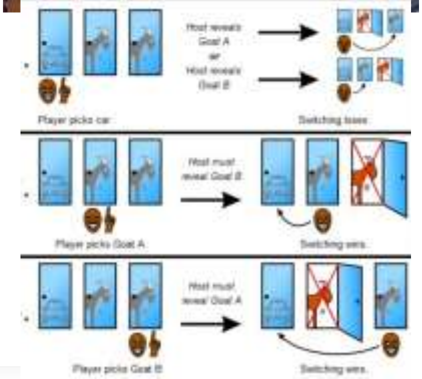
Inference: There are multiple **"causes"** (A1, A2, …) that may result in an **"effect"** (B)
Given an effect, what is the prob. that a certain cause trigger the effect?

Monty Hall Problem

AI

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to change your choice?" Is it to your advantage to switch your choice?

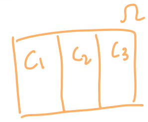
Let's Make a Deal



Marilyn Savant
vs
Erdös

Monty Hall Problem

AI



- Without loss of generality, assume the player picks door 1
- Define events
  - C1: Car is behind door 1 (Similarly C2, C3) ➔ Pr(C1) = 1/3
  - X1: Player pick door 1
  - H3: Host open door 3
  - Pr(H3 | C1 ∩ X1) = ½, Pr(H3 | C2 ∩ X1) = 1, Pr(H3 | C3 ∩ X1) = 0

    ↳ 2 choices (door 2,3) ↳ 1 choice (door 3) ↳ (door 2)
- Probability of win after switching = Pr(C2 | H3 ∩ X1)

  ➔ Show that it is 2/3

Refer to Wikipedia

Random Bit Generator

AI

- A random number bit generator produces a series of random bits, with probability p of producing a 1
  - Each bit generated is an independent trial
  - E: First n nits are all 1's, followed by a single 0
- Pr(E)?
  - Pr(first n 1's) = Pr(1st bit = 1)·Pr(2nd bit = 1) · · · Pr(n-th bit = 1)

    = p^n
  - Pr(E) = Pr(first n 1's)·Pr(n+1st bit = 0)

    = p^n(1-p)
- Let F: k out of n random bits are 1
  - Pr(First k bits are 1, then n-k 0's) = $p^k(1-p)^{n-k}$
  - Pr(k out of n random bits are 1) = $\binom{n}{k} p^k(1-p)^{n-k}$

Search, Hashing and Bitcoin

- A fundamental operation in data analysis is to **find (search)** an object in a big dataset
- Many search algorithms
  - BST (Binary Search Tree)
  - Hashing
  - Usually, hashing is the simplest, yet the most efficient algorithm
    - Complexity = O(1)

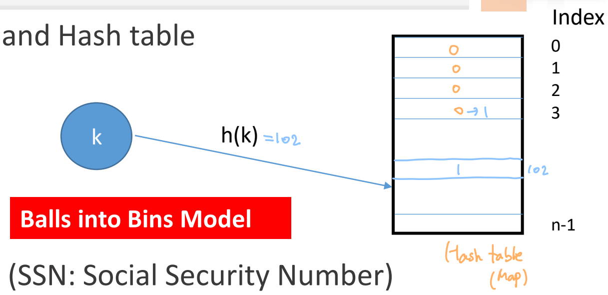
$h(x) = y$
difficult

- A hash function maps a large number to a smaller number, deterministically
  - One-way function
  - Given an input it is easy to compute its output, but the reverse is difficult
- Bitcoin
  - POW(Proof Of Work)
  - Given an output, find inputs that are close enough
  - SHA256 (256 bit Secure Hashing Algorithm)

Hash Table

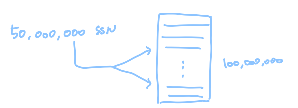
- Key, Hash function, and Hash table

K h(k)



- Example: 주민번호 (SSN: Social Security Number)
  - Each person has a unique key of 13 digits
    - Key space (K) = 10^{13}
  - There are < 10^8 unique keys

1억명 이하



- Hash function h: K → (0, 1, …, n-1)
  - Simple uniform hash function: Each key is equally likely to hash to any of n slots (buckets)
- Collision
  - Two different keys are mapped to the same slot

Hash Table

KENTECH Korea Institute of Energy Technology AI

- m keys are hashed into a hash table of n slots
  - Each key hashing is an independent trial
  - E: At least one key hashed to the first slot
  - Pr(E)? ↳ $\bar{E}$: No key hashed to the slot

Hint: Think out independent events. Then AND (intersection) of them.

- Solution
  - Fi: Key i ***not*** hashed to the first slot ($\cancel{0}$ 1 $\leq$ i $\leq$ m)
  - Pr(Fi) = 1-1/n = (n-1)/n, for all 0 $\leq$ i $\leq$ m → $\frac{n-1}{n} = 1-(\frac{1}{n})$
  - Pr(no keys hashed to the fist slot) = Pr($F_1 \cap F_2 \cap \ldots \cap F_m$)
  - P(E) = 1-Pr($F_1 \cap F_2 \cap \ldots \cap F_m$)

    $= 1-(\frac{n-1}{n})^m$

    $P(E) = 1-P(\bar{E}) = 1-(\frac{n-1}{n})^m$

    $P(\bar{E}) \equiv F_1 \cap F_2 \cap F_3 \cap \cdots \cap F_m$ independent

    $= Pr(F_1)\cdot Pr(F_2)\cdots Pr(F_m)$

- Like the *birthday problem*
  - Among m friends, at least one friend has the same birthday as you (n = 365)

Hash Table

KENTECH Korea Institute of Energy Technology AI

- m keys are hashed into a hash table of n slots
- E: At least one of slots (1 to k) has keys hashed to it

- Solution
  - Ei: At least one key hashed into the i-th slot

  - Pr(E) = Pr(E1 U E2 U…U Ek)

    = 1-Pr($\overline{E1 \cup E2 \cup \ldots \cup Ek}$)

    = 1-Pr($\overline{E1} \cap \overline{E2} \cap \ldots \cap \overline{Ek}$)

Ei & Ej independent ?

$= 1-(\frac{n-k}{n})^m$

Odds

KENTECH Korea Institute of Energy Technology AI

ex) 기아 vs SSG

기아 wins the game.

- **Odds** of an event (H) is defined as

$$\frac{\Pr(H)}{\Pr(\bar{H})} = \frac{\Pr(H)}{1-\Pr(H)}$$

SSG wins the game

- Odds of H given evidence $\boxed{E}$ ← result of the first game. (total 7 games)

$$\frac{\Pr(H \mid E)}{\Pr(\bar{H} \mid E)} = \frac{\Pr(H)\Pr(E \mid H)\,/\Pr(E)}{\Pr(\bar{H})\Pr(E \mid \bar{H})\,/\Pr(E)}$$

Bases's theory

$$= \frac{\Pr(H)\Pr(E \mid H)}{\Pr(\bar{H})\Pr(E \mid \bar{H})} = \underbrace{\boxed{\frac{\Pr(H)}{\Pr(\bar{H})}}}_{\text{odds}} \frac{\Pr(E \mid H)}{\Pr(E \mid \bar{H})}$$

- After observing E, update odds by $\frac{\Pr(E \mid H)}{\Pr(E \mid \bar{H})}$

Lee Sedol vs AlphaGo

KENTECH Korea Institute of Energy Technology AI

- Let H: Lee is better than AG
- Before the match, Pr(H) = 0.9

$\Pr(\bar{H}) = 0.1$

- If Lee is better than AG (H), then Lee wins game (E) with 0.8 probability → $\Pr(\bar{E} \mid H) = 0.8$, $\Pr(E \mid H) = 0.2$
- If AG is better than Lee ($\bar{H}$), then AG wins game (E) with 0.9 probability → $\Pr(E \mid \bar{H}) = 0.9$, $\Pr(\bar{E} \mid \bar{H}) = 0.1$
- E: AG won a game
- What is updated odds after the game? $X = \underbrace{\boxed{\frac{\Pr(H)}{\Pr(\bar{H})}}}_{\text{odds}} = \frac{0.9}{0.1} = 9$ $\qquad \frac{\Pr(E \mid H)}{\Pr(E \mid \bar{H})} = \frac{0.2}{0.9} = \frac{2}{9}$
- What if AG wins two games in a row? $\Rightarrow 9 \times \frac{2}{9} = 2$

Coins & Urns

KENTECH Korea Institute of Energy Technology AI

- An urn contains 2 coins: A and B
  - A comes up heads with probability ¼
  - B comes up heads with probability ¾
  - Pick coin randomly and flip it, and it comes up heads

$Pr(A) = \frac{1}{2}$ $Pr(A|H) \geq \frac{1}{2}$?
$= \frac{1}{2}$
$\leq \frac{1}{2}$

$Pr(A) \rightarrow Pr(A|H)$
$= \frac{1}{2}$ $= \frac{1}{2}$보다
increase

- What are the odds that A was picked?
  - Before the experiment Pr(A) = Pr($\bar{A}$) = Pr(B) = ½

$$\rightarrow \frac{\Pr(A)}{\Pr(\bar{A})} = \frac{\Pr(A)}{1-\Pr(A)} = 1$$

  - After the experiment

$$\rightarrow \frac{\Pr(A \mid heads)}{\Pr(\bar{A} \mid heads)} = \frac{\Pr(A)\ \Pr(heads \mid A)}{\Pr(\bar{A})\ \Pr(heads \mid \bar{A})} \quad \frac{1}{4},\ \frac{3}{4}$$

$$= \frac{1/2}{1/2}\frac{1/4}{3/4} = 1/3$$

Verifying Matrix Multiplication

KENTECH Korea Institute of Energy Technology AI

- Given three n×n matrices **A**, **B**, and **C**
- Want to verify that **AB** = **C**
- Complexity of matrix multiplication
  - $\Theta(n^3)$
  - $\Theta(n^{2.37})$ (Best Algorithm)
- Randomized algorithm
  - Select a vector $\bar{r} = (r_1, r_2, \ldots, r_n) \in \{0,1\}^n$
  - Compute **AB**$\bar{r}$ (First compute **B**$\bar{r}$ and then **A**(**B**$\bar{r}$), Complexity = $\Theta(n^2)$)

    $O(n^2)$ $O(n^2)$
  - Compute **C**$\bar{r}$
  - Decision:
    - If **AB**$\bar{r}$ = **C**$\bar{r}$ → Conclude that **AB** = **C**
    - If **AB**$\bar{r} \neq$ **C**$\bar{r}$ → Conclude that **AB** $\neq$ **C**

Theorem

- If **AB** $\neq$ **C** and if $\bar{r}$ is chosen uniformly at random from $\{0,1\}^n$, then Pr(**AB**$\bar{r}$ = **C**$\bar{r}$) $\leq$ ½
- Proof
  - First, note that selecting $\bar{r}$ uniformly at random from $\{0,1\}^n$ is equivalent to select each r_i uniformly at random from $\{0,1\}$
  - Let **D** = **AB** – **C** $\neq$ **0** → none zero 존재.

    $D = \begin{bmatrix} d_{11} \cdots & d_{1n} \\ \vdots & \vdots \end{bmatrix}$ → 0이 아닌 최소 1개 존재
  - From **AB**$\bar{r}$ = **C**$\bar{r}$, we know that **D**$\bar{r}$ = **0**
  - Because **D** $\neq$ **0**, there must be some non-zero elements in **D**
  - Let a non-zero element is d_{11}
  - $\sum_{j=1}^{n} d_{1j} \cdot r_j = 0$

    $$\rightarrow r_1 = -\frac{\sum_{j=2}^{n} d_{1j} \cdot r_j}{d_{11}} \qquad (1)$$
  - There is at most one choice of r_1 that satisfies Eq 1.
  - Because r_1 can be either 0 or 1, the probability that **AB**$\bar{r}$ = **C**$\bar{r}$ is at most ½

Randomized Algorithm

- Assume that **AB** $\neq$ **C**
- Repeat the test k times with $\bar{r}$ selected uniformly at random from $\{0,1\}^n$.

  If all k test results are **AB**$\bar{r}$ = **C**$\bar{r}$, then conclude that **AB** = **C**
- Analysis $\boxed{AB\bar{r} = C \cdot \bar{r}}$
  - Fi: Event that i-th test fails
  - $\Pr(F_1) = \Pr(F_2) = \ldots = \Pr(F_k) \leq 1/2$
  - Pr(Algorithm fails) = $\Pr(F_1 \cap F_2 \cap \ldots \cap F_k) \leq 2^{-k}$

    ↓ $O(n^2)$

Revisit: Matrix Multiplication

Application of Bayes' Theorem

KENTECH Korea Institute of Energy Technology AI

- E: Event that **AB = C**
- At the beginning, we do not know if it is true or false
  - ➔ Prior knowledge $Pr(E) = Pr(\bar{E}) = ½$
- B1: First test returns that the identity is correct
- $Pr(E \mid B1) = \dfrac{Pr(B1 \mid E) \cdot Pr(E)}{\underbrace{Pr(B1 \mid E)}_{=1} \cdot \underbrace{Pr(E)}_{=\frac{1}{2},\ \geq \frac{2}{3}} + \underbrace{Pr(B1 \mid \bar{E})}_{\leq \frac{1}{2}} \cdot \underbrace{Pr(\bar{E})}_{\frac{1}{2},\ \leq \frac{1}{3}}}$

  $\geq 2/3$
- B2: Second test returns that the identity is correct

$B_2 \to B$

$Pr(E \mid B_1) \to Pr(E)$

- $Pr(E \mid B2) \geq \dfrac{2/3}{2/3 + 1/3 \cdot 1/2} \geq 4/5$
- Assume that after i-th test, our belief is that $Pr(E) \geq 2^i / (2^i+1)$
- $Pr(E \mid Bi+1) \geq \dfrac{2^{i+1}}{2^{i+1}+1} = 1 - \dfrac{1}{2^i+1}$

Advanced Conditional Probability

KENTECH Korea Institute of Energy Technology AI

- Insurance companies have been using probabilities to make different yet proper charges to customers
  - For example, customers who are more probable to incur costs are charged more than customers with less risks
- Car insurance company problem
  - There are two types of drivers: Careful (0.6) and Careless(0.4)
  - Probabilities that careful and careless customers have accidents in one year are 0.2 and 0.4, respectively
  - Events to have accidents in each year are independent (Depends only on the driver types)
  - Given that a new customer has accidents in the first year, What is the probability that the customer have accidents in the second year?

C $\bar{C}$ A A_1 A_2

$Pr(C) = 0.6$

$Pr(\bar{C}) = 0.4$

$Pr(A \mid C) = 0.2$

$Pr(A \mid \bar{C}) = 0.4$

$Pr(A_2 \mid C) = 0.2$

$Pr(A_2 \mid \bar{C}) = 0.4$

$Pr(A_2 \mid A_1) = ?$

$Pr(A_1 \cap A_2 \mid C) = (0.2) \cdot (0.2)$

$Pr(A_1 \cap A_2 \mid \bar{C}) = (0.4) \cdot (0.4)$

$Pr(A_1 \cap A_2 \mid C) \cdot Pr(C)$

$Pr(A_1 \cap A_2 \cap C) + Pr(A_1 \cap A_2 \cap \bar{C})$

$\dfrac{Pr(A_2 \cap A_1)}{Pr(A_1)}$

$Pr(A_1 \cap C) + Pr(A_1 \cap \bar{C})$

- Note $Pr(E \mid F) = Pr(E \mid G \cap F)\, Pr(G \mid F) + Pr(E \mid \bar{G} \cap F)\, Pr(\bar{G} \mid F)$
  - $Pr(E \mid F) = Pr(E \cap G \mid F) + Pr(E \cap \bar{G} \mid F)$

    $= Pr(E \cap G \cap F) / Pr(F) + Pr(E \cap \bar{G} \cap F) / Pr(F)$

    $= Pr(E \mid G \cap F)\, Pr(G \cap F) / Pr(F) + Pr(E \mid \bar{G} \cap F)\, Pr(\bar{G} \cap F) / Pr(F)$

    $= Pr(E \mid G \cap F)\, Pr(G \mid F)\, Pr(F) / Pr(F)$

    $+ Pr(E \mid \bar{G} \cap F)\, Pr(\bar{G} \mid F)\, Pr(F) / Pr(F)$

Advanced Conditional Probability

KENTECH Korea Institute of Energy Technology AI

- Solution
  - A2: Event that the customer have accidents in the second year
  - A1: Event that the customer have accidents in the first year
  - C: Event that customer is careful ($\bar{C}$: Careless)
  - Pr(E | F) = Pr(E | G∩F) Pr (G | F) + Pr(E | $\bar{G}$∩F) Pr ($\bar{G}$ | F)
  - E ← A2, F ← A1, G ← C
  - Pr(A2 | A1) = Pr(A2 | A1∩C) Pr (C | A1) + Pr(A2 | A1 ∩ $\bar{C}$) Pr ($\bar{C}$ | A1)
  - Compute Pr(C | A1) using Bayes' Theorem
  - Pr(A2 | A1∩C) ??

- Suppose a customer have accidents in first and second years consecutively, what is the probability that the customer is a careful driver?

  ➔ Pr (C | A1∩A2)

  $= \frac{Pr(A_1 \cap A_2 | C) \cdot Pr(C)}{Pr(A_1 \cap A_2)}$

AlphaGo W/W/W/L (win ... lose)

$Pr(L | WWWL)$

Sequential Information Update

Sequential Information Update

KENTECH Korea Institute of Energy Technology AI

- A hypotheses H (such as a driver is careful driver) with an initial guess is given Pr(H is True) = p = 1- Pr(H is False)
- After an Event E is occurred, the conditional probability that H is True (Let this be T) is given as
  - $Pr(T | E) = \frac{Pr(E | T)Pr(T)}{Pr(E | T)Pr(T) + Pr(E | F)Pr(F)}$

Bayes theory

$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$

- Now, suppose we observed two successive (independent) events E1 and E2
  - $Pr(T | E1 \cap E2) = \frac{Pr(E1 \cap E2 | T)Pr(T)}{Pr(E1 \cap E2 | T)Pr(T) + Pr(E1 \cap E2 | F)Pr(F)}$
- Can we consider E2 as E and Pr(T | E1) as Pr(T)?
- Solution
  - Yes, if E1 and E2 are conditionally independent given H

    ➔ Pr(E1∩E2 | H) = Pr(E2 | H) Pr(E1 | H)
  - To show $Pr(T | E1 \cap E2) = \frac{Pr(E2 | T)Pr(T | E1)}{Pr(E2 | T)Pr(T | E1) + Pr(E2 | F)Pr(F | E1)}$

Conditional Independence

KENTECH Korea Institute of Energy Technology AI

- Events E and F are **conditionally Independent** given G iff

  $\boxed{\Pr(E \cap F \mid G) = \Pr(E \mid G)\Pr(F \mid G)}$ $Pr(E \cap F) = Pr(E) \cdot Pr(F)$

- Claim: If E and F are independent

  ➔ E given G and F given G are independent

- No, Counter example
  - Roll two dice yielding values D1 and D2
  - E: D1=1 $Pr(E) = \frac{1}{6}$
  - F: D2=6 $Pr(F) = \frac{1}{6}$
  - G: D1+D2=7   (1,6) (2,5), ... (6,1) → 6가지
  - $= Pr(E)Pr(F) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$
  - E and F are independent, Pr(E∩F) = 1/36 and Pr(E)=1/6, Pr(F)=1/6
  - Pr(E | G)=1/6, Pr(F | G)=1/6 and Pr(E∩F | G)=1/6 → not independent

    ↑ 6

Conditional Independence

KENTECH Korea Institute of Energy Technology AI

- Claim: If E and F are conditionally independent given G then Pr(E | F∩G) = Pr(E | G)

  ↳ $Pr(E \cap F \mid G) = Pr(E \mid G) \cdot Pr(F \mid G)$

- Proof
  - $\Pr(E \mid F \cap G) = \frac{\Pr(E \cap F \cap G)}{\Pr(F \cap G)}$

$$= \frac{Pr(E \cap F \mid G) \cdot Pr(G)}{Pr(F \cap G)}$$

$$= \frac{Pr(E \cap G)}{Pr(G)} = Pr(E \mid G)$$

$$\frac{Pr(E \mid G) Pr(F \mid G) \cdot Pr(G)}{Pr(E \cap F)}$$

Another Example

- 100 person in “The Class” Bldg
  - 30 are in AI Track (Either students or faculty)
  - 20 are Faculty
  - There are 6 AI Faculty
  - Pr(AI)=0.3, Pr(F)=0.2 Pr(AI∩F)=0.06 ➔ AI and F are independent
  - Only the persons in AI or Faculty can use the ServerRoom
  - AI given ServerRoom and F given ServerRoom are independent?
- Solution
  - D: ServerRoom users = AI U F
  - |D| = 30 + 20 - 6 = 44
  - Pr(AI|D) = 30/44, Pr(F|D) = 20/44, Pr(AI∩F | D) = 6/44

    ➔ **Conditionally Dependent**

Independence & Conditioning

- Conditioning can make dependent events to independent?
- Yes, Example
  - Sample space: {M, Tu, W, Th, F, Sa, Su} 7 days
  - A: not Monday = {Tu, W, Th, F, Sa, Su} → 6 days
  - B: {Sa}
  - C: {Sa, Su}
  - Pr(A)=6/7, Pr(B)=1/7 and Pr(A ∩ B)=1/7 ≠ Pr(A) · Pr(B) $\frac{6}{7} \times \frac{1}{7} = \frac{6}{49}$

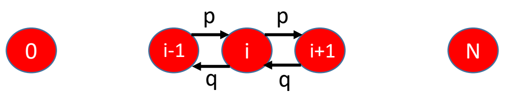
    ➔ A and B are dependent

  - Pr(A|C)=1, Pr(B|C)=1/2, P(A ∩ B|C)=1/2 = Pr(A|C) · Pr(B|C) $1 \times \frac{1}{2}$

    ➔ A|C and B|C are independent

Gambler's Ruin Problem

- Game setting
  - Gambler A and B
  - Successive coin flips. If heads, A collect one unit from B. If tails, A give one unit to B
  - Pr(heads) = p = 1-Pr(tails)
  - A starts with i units and B starts with N-i units
  - Game finishes when one of gamblers collects all
  - Probability that A wins?



- Solution
  - E: A wins
  - H: first flip is heads
  - Pi = Pr(E) = Pr(E | H)Pr(H) + Pr(E | $\overline{H}$)Pr($\overline{H}$)

Gambler's Ruin Problem

- Solution
  - Pi = Pr(E | H)•p + Pr(E | $\overline{H}$) •(1-p)

    = p•P_{i+1} + q•P_{i-1}

    ➔ p•P_i + q•P_i = p•P_{i+1} + q•P_{i-1}

    ➔ P_{i+1}- P_i = q/p (P_i – P_{i-1})
  - Obviously, P_0 = 0 and P_N = 1

    P_2 - P_1 = q/p (P_1 - P_0) = (q/p) P_1

    P_3 - P_2 = q/p (P_2 - P_1) = $(q/p)^2$ P_1

    .

    .

    Pi – P_{i-1} = $(q/p)^{i-1}$ P_1

    ➔ Pi – P_1 = P_1 [(q/p) + $(q/p)^2$ +…+ $(q/p)^{i-1}$]

    ➔ Pi = $\begin{cases} \frac{1-(^q/_p)^i}{1-(^q/_p)} \cdot P_1, & if\ p \neq {}^1/_2 \\ i \cdot P_1, & if\ p = {}^1/_2 \end{cases}$

Gambler's Ruin Problem

AI

- Solution
  - From $P_N = 1$, we obtain

$$P_1 = \begin{cases} \frac{1-(q/p)}{1-(q/p)^N} & ,if\ p \neq 1/2 \\ \frac{1}{N} & ,if\ p = 1/2 \end{cases}$$

$$\rightarrow P_i = \begin{cases} \frac{1-(q/p)^i}{1-(q/p)^N} & ,if\ p \neq 1/2 \\ \frac{i}{N} & ,if\ p = 1/2 \end{cases}$$