Conditional Probability & Bayes' Rule

- **Q**: Should the following probabilities be different?
  - ➢ Event = a pitcher wins at least 16 games in a season
    - P=?? in the beginning of the season
    - P=?? in the middle of the season
    - P=?? close to the end of the season
  - ➢ Event = rain tomorrow
    - P=?? if no information about where you are staying
    - P=?? if you are staying in a desert
    - P=?? if a typhoon will hit the place you stay tomorrow
- **Q**: What causes the differences?
  - ➢ For an event, *new information* (i.e., some other event has occurred) could change its probability
  - ➢ We call the altered probability a conditional probability

- Mathematical Definition: If A and B are two events in a sample space Ω and $P(A)>0$, then

$$P(B|A) \equiv \frac{P(A \cap B)}{P(A)}$$

is called the *conditional probability* of B given A.

  - ➢ In the classical probability,

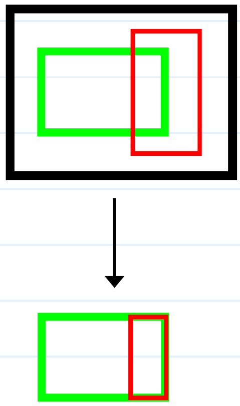
$$P(A) = \#A/\#\Omega \quad \text{and} \quad P(A \cap B) = \#(A \cap B)/\#\Omega$$

$$\Rightarrow P(B|A) = \frac{\#(A \cap B)/\#\Omega}{\#A/\#\Omega} = \frac{\#(A \cap B)}{\#A}.$$

    - Example: A family is known to have 2 children, *at least one of whom is a girl*. **Q**: Probability that the other is a boy=??
      - ▫ $\Omega = \{bb, bg, gb, gg\}$
      - ▫ $A=\{bg, gb, gg\}$ and $B=\{bb, bg, gb\}$
      - ▫ $P(B|A) = \#(A \cap B)/\#A = 2/3$
    - Note: $\#\Omega$ is reduced to $\#A$.

- In effect by conditioning,
  - we are restricting the sample space from Ω to A, i.e.,

$$\Omega \to A,$$



  - and, for an arbitrary event B (in Ω) to occur when A has occurred, we need that both A and B occur together, i.e.,

$$B \to B \cap A.$$

- The division by $P(A)$ in the definition above rescales all probabilities from the entire sample space Ω to being relative to the new sample space A
- $P(B|A)$ is a probability measure defined on B, but not A.
  - $P(\cdot|A)$ satisfies the 3 axioms of probability (exercise)

$P(\cdot)$  $P(\cdot|A)$ 

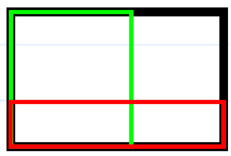
  - Any propositions developed in Chapter 2 for probability measures can be applied on $P(\cdot|A)$.
    (For example, $P(B^c|A)=1-P(B|A)$.)

- 3 Useful Formulas for Calculating Probabilities (for 2-events case)

1. If $P(A)>0$, then $P(A \cap B) = P(A)\ P(B|A)$.

2. If $0<P(A)<1$, then

$$P(B) = P(A)P(B|A) + P(A^c)P(B|A^c).$$



3. If $0<P(A)<1$ and $P(B)>0$, then

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

- Example (Urn Problem)
  - The Story. n balls sequentially and randomly chosen, *without replacement*, from an urn containing R red and $N-R$ white balls ($n \le N$). **Q**: Given that k of the n balls are red ($k \le R$), probability that the 1st ball chosen is red = ??
  - Let $A=\{k \text{ of the } n \text{ balls are red}\}$

    $B=\{1^{\text{st}} \text{ ball chosen is red}\}$
  - Method 1: $P(B|A) = \binom{n-1}{k-1} \Big/ \binom{n}{k} = k/n$
  - Method 2:
    - $P(A) = \left[\binom{R}{k} \times \binom{N-R}{n-k}\right] \Big/ \binom{N}{n}$

- $P(A \cap B) = P(B)P(A|B) = \dfrac{R}{N} \times \dfrac{\binom{R-1}{k-1} \times \binom{N-R}{n-k}}{\binom{N-1}{n-1}}$
- $P(B|A) = P(A \cap B)/P(A) = k/n$

- Example (Sampling Experiments): An urn contains R red balls and $N-R$ white balls. Sample 2 balls from the urn.
  - $A = \{\text{red on the first draw}\}$

    $B = \{\text{red on the second draw}\}$

  - Sampling *Without Replacement*:

$$P(A) = \frac{R(N-1)}{N(N-1)} = \frac{R}{N}, \quad P(A \cap B) = \frac{R(R-1)}{N(N-1)},$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{(R(R-1))/(N(N-1))}{R/N} = \frac{R-1}{N-1}.$$

Similarly, $P(B|A^c) = \dfrac{R}{N-1}$. (exercise)

$$P(B) = P(A)P(B|A) + P(A^c)P(B|A^c)$$
$$= \frac{R}{N} \cdot \frac{R-1}{N-1} + \frac{N-R}{N} \cdot \frac{R}{N-1}$$
$$= \frac{R^2 - R + NR - R^2}{N(N-1)} = \frac{R(N-1)}{N(N-1)} = \frac{R}{N}.$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{(R(R-1))/(N(N-1))}{R/N} = \frac{R-1}{N-1}.$$

- Notes:
  - The probabilities are proportional to # of red balls left
  - $P(A|B) = P(B|A) \Rightarrow$ Symmetry.

- Sampling *With Replacement*:

$$P(A) = \frac{R \cdot N}{N \cdot N} = \frac{R}{N}, \qquad P(B) = \frac{N \cdot R}{N \cdot N} = \frac{R}{N},$$
$$P(A \cap B) = \frac{R \cdot R}{N \cdot N} = \frac{R^2}{N^2},$$
$$P(B|A) = \frac{R^2/N^2}{R/N} = \frac{R}{N} = P(B).$$

➢ Example (Diagnostic Tests)

- The Story. A diagnostic test for a *rare* disease (e.g., an Xray for lung cancer) is part of a routine physical exam.
- Let Ω = {the whole population of Taiwan}

  D = {disease is present}

  E = {test indicates disease present}
- Suppose that $P(D)=0.001$, $P(E|D)=0.98$, $P(E|D^c)=0.01$

  Q: Do you think the test is effective?
- Let us examine it from an alternative viewpoint. Suppose that you are tested as positive. What is the probability that you *actually have* the disease?

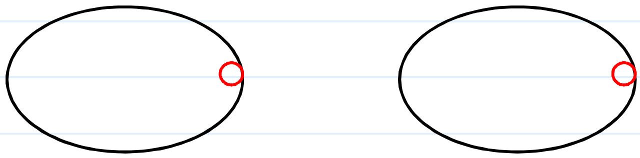
$$P(E) = P(D)P(E|D) + P(D^c)P(E|D^c)$$
$$= 0.001 \times 0.98 + 0.999 \times 0.01 = 0.01097.$$

$$P(D|E) = \frac{P(D)P(E|D)}{P(D)P(E|D) + P(D^c)P(E|D^c)} = \frac{0.00098}{0.01097} = 0.0893.$$

$$P(D^c|E) = 1 - P(D|E) = 0.9107$$

Q: Now, do you still think the test is effective?

- The probability of D increased by a factor of roughly 90 ($0.001 \rightarrow 0.0893$) when E occurs, but 0.0893 is still **small**
- The $P(E|D^c)$ (=0.01) and $P(E^c|D)$ (=$1-P(E|D)$=0.02) are called the *false positive* and *false negative rates*, respectively.
- **Q**: Why the 2 interpretations so different? What causes it?



➢ Example (引自“快思慢想”, Kahneman)

- 在1970年代，湯姆是美國某州裡重要大學的研究生，請將下面九個研究所領域排序，標出湯姆就讀這些領域的可能性，1代表可能性最高，9代表最低。

  A. 企業管理, B. 電腦, C. 工程, D. 人文與教育, E. 法律, F. 醫學, G. 圖書館學, H. 物理和生命科學, I. 社會學和社會工作

- 下面是湯姆念高三時，心理學家根據心理測驗結果對湯姆的人格素描：
  - 湯姆是個很聰明的學生，但缺少真正的創造力。
  - 他喜歡整潔和秩序，他的每一樣東西，不管多少，都有條有理的擺放在恰當位置上。
  - 他的作文有點無趣呆板和機械式，偶爾會出現一些陳舊過時的雙關語和類似科幻想像的句子。
  - 他的好勝心很強，對人冷淡，没什麼同情心，也不喜歡跟別人來往。
  - 雖然以自我為中心，卻有很強的道德感。

  把剛剛那幾個領域再排序一下，你認為湯姆最可能是哪個領域的研究生，1代表可能性最高，9代表最低。

- 在1970年代，受試者的排序大致如下：

  1. 電腦, 2. 工程, 3. 企業管理, 4. 物理和生命科學, 5. 圖書館學, 6. 法律, 7. 醫學, 8. 人文與教育, 9. 社會學和社會工作

- Extensions of the 3 Useful Formulas (for m-events case)
  1. (Multiplication Law) If $A_1, \ldots, A_m$ are events for which $P(A_1 \cap \cdots \cap A_{m-1}) > 0$, then

$$P(A_1 \cap \cdots \cap A_m) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)\cdots P(A_m|A_1 \cap \cdots \cap A_{m-1}).$$

- Example (Birthday Problem, LNp.3-4). Let A_k be the event that the k^{th} birthday differs from the first $k-1$. Then, $P(A_1)=1$,
  - $P(A_k|A_1 \cap \cdots \cap A_{k-1}) = \dfrac{365-k+1}{365}$
  - $$P(A_1 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)\cdots P(A_n|A_1 \cap \cdots \cap A_{n-1})$$ $$= \prod_{k=1}^{n} \frac{365-k+1}{365} = \frac{365!}{365^n \cdot (365-n)!} = \frac{(365)_n}{365^n}.$$

- Example (Matching Problem, LNp.3-14). **Q**: Probability that exactly k of n members ($k \le n$) have matches = ??
  - Let Ω be all permutations $\omega = (i_1, \ldots, i_n)$ of 1, 2, …, n.
  - Let $A_j = \{\omega: i_j = j\}$ and

$$A = \bigcup_{1 \le j_1 < \cdots < j_k \le n} (A_1^c \cap \cdots \cap A_{j_1-1}^c \cap A_{j_1} \cap A_{j_1+1}^c \cap \cdots \cap A_n^c)$$

  - By symmetry,

$$P(A) = \binom{n}{k} \times P(A_1 \cap \cdots \cap A_k \cap A_{k+1}^c \cap \cdots \cap A_n^c)$$

  - Let $E = A_1 \cap \cdots \cap A_k$ and $G = A_{k+1}^c \cap \cdots \cap A_n^c$
  - Then, $P(E \cap G) = P(E)P(G|E)$, where

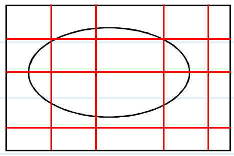
$$P(E) = P(A_1)P(A_2|A_1)\cdots P(A_k|A_1 \cap \cdots \cap A_{k-1})$$
$$= \frac{1}{n} \times \frac{1}{n-1} \times \cdots \times \frac{1}{n-k+1} = \frac{(n-k)!}{n!} = \frac{1}{(n)_k}$$

and, $P(G|E) = \sum_{i=0}^{n-k} (-1)^i \frac{1}{i!} \equiv p_{n-k}$

- $P(A) = \binom{n}{k} \frac{1}{(n)_k} p_{n-k} = \frac{p_{n-k}}{k!} \approx \frac{e^{-1}}{k!}$, when n is large

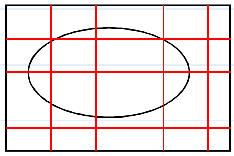
2. (Law of Total Probability) Let $A_1, \ldots, A_m$ be a partition of Ω and $P(A_i)>0$, $i=1, \ldots, m$, then for any event $B \subset \Omega$,

$$P(B) = \sum_{i=1}^{m} P(A_i)P(B|A_i).$$

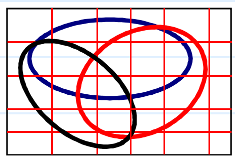


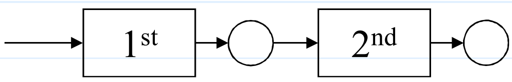
3. (Bayes' Rule) Let $A_1, \ldots, A_m$ be a partition of Ω and $P(A_i)>0$, $i=1, \ldots, m$. If B is an event such that $P(B)>0$, then for $1 \leq j \leq m$,

$$P(A_j|B) = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^{m} P(A_i)P(B|A_i).}$$



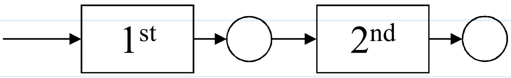
- From Bayesians' viewpoint,
  $P(A_j)$=probability of A_j *before* B occurs $\to$ *prior* prob.
  $P(A_j|B)$=probability of A_j *after* B occurs $\to$ *posterior* prob.
  $\Rightarrow$ The Bayes' rule tells how to update the probabilities of A_j in light of the new information (i.e., B occurs)
- An Application of Bayes' Rule. Suppose that a random experiment consists of two random stages





  - The probabilities of the 2nd-stage results depend on what happened in the 1st stage
  - We never see the result of the 1st stage, only the final result
  - We may be interested in finding the probability for outcomes in the 1st stage given the final result

- Example (Gold Coins):



  - The Story.
    - Box 1 contains 2 silver coins.
    - Box 2 contains 1 gold and 1 silver coin.
    - Box 3 contains 2 gold coins.
    - Experiment: (i) Select a box at random and, (ii) Examine the 2 coins in order (assuming all choices are equally likely at each stage)
  - **Q**: Given that 1st coin is gold, what is the probability that Box k is selected, k=1, 2, 3?
    - Let A_k = {Box k is selected}, B = {1st coin is gold},
    - $$P(B|A_k) = \begin{cases} 0, & \text{if } k=1, \\ 1/2, & \text{if } k=2, \\ 1, & \text{if } k=3. \end{cases}$$

$$P(B) = \frac{1}{3}\cdot 0 + \frac{1}{3}\cdot\frac{1}{2} + \frac{1}{3}\cdot 1 = \frac{1}{2}.$$

- $$P(A_1|B) = \frac{(1/3)\cdot 0}{1/2} = 0.$$

  Similarly, $P(A_2|B) = 1/3$, $P(A_3|B) = 2/3$.

- **Q**: Given that 1st coin is gold, what is the probability that 2nd coin is gold?
  - Let C = {2nd coin is gold}. $P(B\cap C|A_k) = \begin{cases} 0, & \text{if } k=1, \\ 0, & \text{if } k=2, \\ 1, & \text{if } k=3. \end{cases}$
  - $$P(B\cap C) = \frac{1}{3}\cdot 0 + \frac{1}{3}\cdot 0 + \frac{1}{3}\cdot 1 = \frac{1}{3}.$$

$$P(C|B) = \frac{P(B\cap C)}{P(B)} = \frac{1/3}{1/2} = \frac{2}{3}.$$

- Example (TV Game Show: Let's Make A Deal)

  - The story.
    1. The contestant is given an opportunity to select one of three doors.
    2. Behind one of the doors is a great prize (say, a car) and there is nothing behind the other two doors.
    3. The host knows which door contains the car, but the contestant does not.

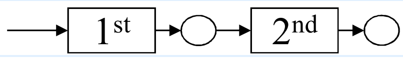
4. After the contestant select a door, the host opens an *empty door* that the contest *did not* pick.
5. After opening an empty door, the host always offers the contestant the opportunity to switch to the other remaining *unopened* door.

- **Q**: Should the contestant switch to the other door or not?
- Argument 1 (*The Drunkar's Walk* by L. Mlodinow): "Two doors are available --- open one and you win; open the other and you lose …, your chances of wining are 50/50."
- Argument 2. Without loss of generality, assume that the contestant select door 3. Let

  A_i = {the car is behind the door i}, i=1, 2, 3.

  B = {door 1 is opened}

  - $P(A_1) = P(A_2) = P(A_3) = 1/3$



  - $P(B|A_1) = 0, \quad P(B|A_2) = 1,$
    $P(B|A_3) = 1/2$

  - $$P(A_3|B) = \frac{(1/3)\times(1/2)}{(1/3)\times 0 + (1/3)\times 1 + (1/3)\times(1/2)} = 1/3$$
  - $P(A_2|B) = 2/3$
  - Similar result obtained if B={door 2 is opened}
- Intuitive interpretation.

➢ **Q**: Why are the 3 formulas useful in calculating probabilities? (Note: They all benefit from conditional probabilities.)

Ans: (i)繁→簡&簡&⋯&簡;

(ii)簡=conditioning because the sample space is reduced from Ω to a smaller set. (e.g., in many previous examples, $P(B|A)$'s are known or easier to evaluate)

• Odds and Conditional Odds

➢ The odds of an event B:

$$o(B) \equiv \frac{P(B)}{P(B^c)} = \frac{P(B)}{1-P(B)}$$

- The odd of event B given A:

$$o(B|A) \equiv \frac{P(B|A)}{P(B^c|A)}$$

and

$$o(B|A) = o(B) \times \frac{P(A|B)}{P(A|B^c)}$$

❖ **Reading**: textbook, Sec 3.1, 3.2, 3.3, 3.5

Independence

- Definition (independence for 2-events case): Two *events* A and B are said to be *independent* if and only if

$$P(A \cap B) = P(A)P(B).$$

Otherwise, they are said to be *dependent*.

- Notes. If $P(A) > 0$, events A and B are independent if and only if

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

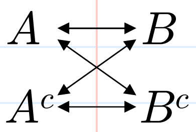
similarly, if $P(B) > 0$, if and only if $P(A|B) = P(A)$.

Q: How to interpret the equality?

- Example (Sampling 2 balls, LNp.4-6~7). Events A and B were "independent" for sampling with replacement, but "dependent" for sampling without replacement.
- Example (Cards): If a card is selected from a standard deck, let
  - A = {ace} and B={spade}. Then,
  - $P(A) = \frac{4}{52} = \frac{1}{13}$, $P(B) = \frac{13}{52} = \frac{1}{4}$,

    $$P(A \cap B) = \frac{1}{52} = P(A)P(B)$$
  - Face and Suit are independent
- Theorem (Independence and Complements, 2-events case). If A and B are independent, then so are A^c and B.

- Corollary: If A and B^c are independent, so are A^c and B^c
- Corollary: If A and B are independent and $0<P(A)<1$, $0<P(B)<1$, then



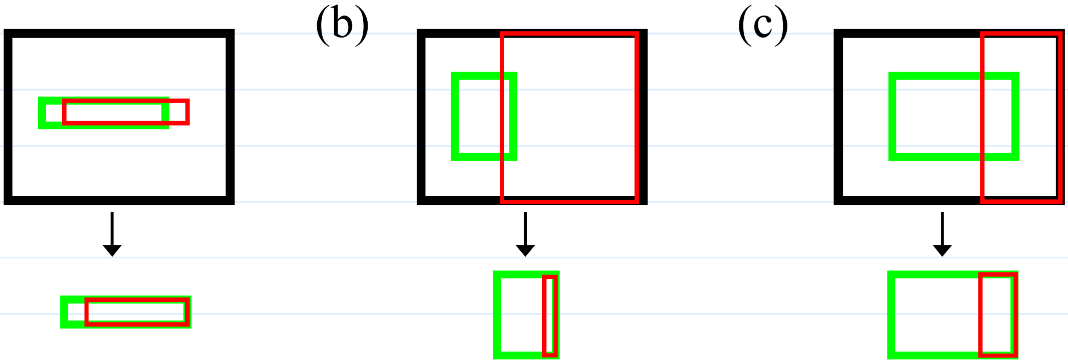
$$P(B) = P(B|A) = P(B|A^c),$$
$$P(B^c) = P(B^c|A) = P(B^c|A^c),$$
$$P(A) = P(A|B) = P(A|B^c),$$
$$P(A^c) = P(A^c|B) = P(A^c|B^c).$$

Q: What do these equalities say?

➢ Example.

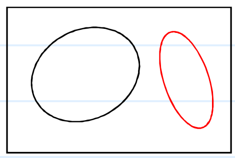
- **Q**: Which of the following graphs represents "green and red events are independent" (assume probability ∝ area)?

(a) (b) (c)



- **Q**: Let green event={graduate from Tsing-Hua University}, red event={your future dream will come true}. Which of the graphs would you prefer?
- **Q**: What do we prefer? independent? or dependent?

➢ Theorem (Independence and Mutually Exclusive). If A and B are mutually exclusive and $P(A)>0$, $P(B)>0$, then A and B are dependent since



$$P(B|A) = 0 \neq P(B).$$

- Definition (independence for n-events case). Events $A_1, \ldots, A_n$ are said to be *pairwise independent* iff for all $1 \le i < j \le n$;

$$P(A_i \cap A_j) = P(A_i)P(A_j),$$

$A_1, \ldots, A_n$ are said to be *mutually independent* iff for $k=2, \ldots, n$,

$$P(A_{i_1} \cap A_{i_2}) = P(A_{i_1})P(A_{i_2}), \quad \text{for } 1 \le i_1 < i_2 \le n,$$
$$P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) = P(A_{i_1})P(A_{i_2})P(A_{i_3}), \quad \text{for } 1 \le i_1 < i_2 < i_3 \le n,$$
$$\ldots,$$
$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}), \quad \text{for } 1 \le i_1 < \cdots < i_k \le n,$$
$$\ldots$$

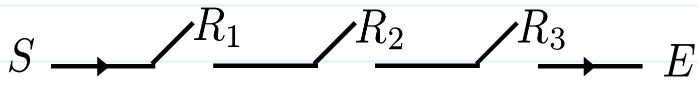
- Note:
  - Suppose $A_1, \ldots, A_n$ are mutually independent. For $1 \le r < k \le n$, and different $t_1, \ldots, t_r, t_{r+1}, \ldots, t_k \in \{1, 2, \ldots, n\}$,
$$P(A_{t_1} \cap \cdots \cap A_{t_r} | A_{t_{r+1}} \cap \cdots \cap A_{t_k}) = P(A_{t_1} \cap \cdots \cap A_{t_r}).$$
  - Mutual independence implies pairwise independence; but, the converse statement is usually not true.
  - "n events are independent" means "mutually independent"
- Example (Sampling With Replacement)
  - A sample of n balls is drawn with replacement from an urn containing R red and $N-R$ white balls
  - Let A_k={red on the k^{th} draw}, then
$$P(A_k)=R/N, \quad k=1, \ldots, n.$$
  - For all $1 \le i_1 < \cdots < i_k \le n$, where $k=2, \ldots, n$,
$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = \frac{R^k N^{n-k}}{N^n} = \left(\frac{R}{N}\right)^k = P(A_{i_1}) \cdots P(A_{i_k}),$$
$\Rightarrow A_1, \ldots, A_n$ are mutually independent

- Example. Draw one card from a standard deck.
  - Let A={Spades or Clubs},
    B={Hearts or Clubs},
    C={Diamonds or Clubs}.
  - $P(A) = 26/52 = 1/2$, similarly, $P(B) = P(C) = 1/2$.
  - $P(A \cap B) = P(\{\text{Clubs}\}) = \frac{13}{52} = \frac{1}{4} = P(A)P(B)$, similarly,

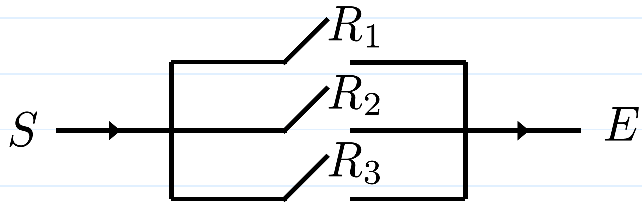
$$P(A \cap C) = 1/4 = P(A)P(C), \; P(B \cap C) = 1/4 = P(B)P(C).$$
$\Rightarrow A$, B, and C are pairwise independent
  - However,
$$P(A \cap B \cap C) = P(\{\text{Clubs}\}) = \frac{1}{4} \neq \frac{1}{8} = P(A)P(B)P(C),$$
$\Rightarrow A$, B, and C are *not* mutually independent
- Theorem (Independence and Complements, n-events case). $A_1, \ldots, A_n$ are mutually independent if and only if
$$P(B_1 \cap \cdots \cap B_n) = P(B_1) \cdots P(B_n),$$
where B_i is either A_i or A_i^c, for $i=1, \ldots, n$.

- Example (Series and Parallel Connections of Relays).

Series Connection



Parallel Connection



- The Story. For n electrical relays $R_1, \ldots, R_n$, let

$$A_k = \{R_k \text{ works properly}\},$$

$k=1,\ldots,n$, and suppose that $A_1, \ldots, A_n$ are independent.

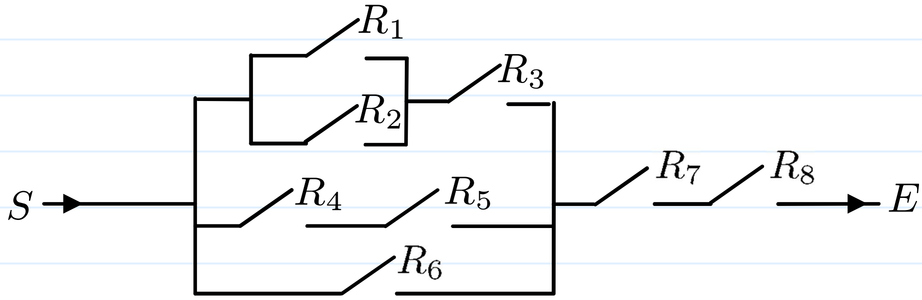
- Series Connection. The probability that current can flow from S to E (corresponding to the event $A_1 \cap \cdots \cap A_n$) is

$$P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdots P(A_n).$$

- Parallel Connection. The probability that current can flow from S to E (corresponding to the event $A_1 \cup \cdots \cup A_n$) is

$$P(A_1 \cup \cdots \cup A_n) = 1 - P(A_1^c \cap \cdots \cap A_n^c)$$
$$= 1 - P(A_1^c) \cdots P(A_n^c) = 1 - \prod_{k=1}^{n} [1 - P(A_k)]$$

- Combination of Series and Parallel Connections



➢ Theorem. If $A_1, \ldots, A_n$ are mutually independent and $B_1, \ldots, B_m$, $m \leq n$, are formed by taking unions or intersections of mutually exclusive subgroups of $A_1, \ldots, A_n$, then $B_1, \ldots, B_m$ are mutually independent.

- Definition (conditional independence): Events $B_1, \ldots, B_n$ are (*pairwise* or *mutually*) *independent* under the probability measure $P(\cdot|A)$.
  ➢ e.g., B_1 and B_2 are conditionally independent given A iff

$$P(B_1 \cap B_2|A) = P(B_1|A)P(B_2|A),$$

or, equivalently,

$$P(B_1|B_2 \cap A) = P(B_1|A) \text{ or } P(B_2|B_1 \cap A) = P(B_2|A)$$

- **Example (Gold Coins):**
  - The Story.
    - Box i contains i gold coins and $k-i$ silver coins, $i=0,1,\ldots,k$.
    - Experiment: (i) Select a box at random, (ii) Draw coins *with replacement* from the box
  - **Q**: Given that first n draws are all gold, what is the probability that $(n+1)^{\text{st}}$ draw is gold?
    - Let $A_i = \{\text{Box } i \text{ is selected}\}$, $B = \{\text{first } n \text{ draws are gold}\}$, $C = \{(n+1)^{\text{st}} \text{ draw is gold}\}$
    - By applying law of total probability on $P(\cdot|B)$,

$$P(C|B) = \sum_{i=0}^{k} P(A_i|B)P(C|A_i \cap B)$$

    - Because B and C are *conditionally independent* given A_i,

$$P(C|A_i \cap B) = P(C|A_i) = i/k$$

    - By Bayes' rule,

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=0}^{k} P(A_j)P(B|A_j)} = \frac{[1/(k+1)](i/k)^n}{\sum_{j=0}^{k}[1/(k+1)](j/k)^n}$$

    - Hence, $P(C|B) = \sum_{i=0}^{k}(i/k)^{n+1} \Big/ \sum_{j=0}^{k}(j/k)^n$

  - **Q**: Are the events B and C independent under $P(\cdot)$?

❖ **Reading**: textbook, Sec 3.4, 3.5