Ch. 2. Basic Concepts of Probability Theory

- Set theory : Sample space, events.
- Axioms of Probability
- Conditional Probability : independence
- Sequential random experiments: simple subexperiments

Specifying Random Experiments

- Experimental procedure
  - Unambiguous statement of exactly what is measured or observed

- A set of one or more measurements or observations

Sample Space

- Outcome or sample point
  - cannot be decomposed into other results: mutually exclusive
- The sample space S
  - The set of all possible outcomes
  - Set notation, tables, diagrams, intervals of the real line, regions of the plane
  - Finite, Countably infinite, Uncountably infinite

Sample Space (cont'd)

- Discrete sample space if S is countable
  - Countable : one-to-one correspondence with the positive integers
- Continuous sample space if S is not countable
- Multi-dimensional sample space (one or more observations or measurements)
  - Dimension of outcome
- Impossible outcome
  - infinite life-times of a given computer memory

Events

- **Event:** A subset of the sample space S satisfying the conditions of interest

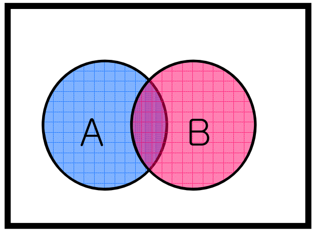
- Two special events
  - The certain event = S
  - The impossible or null event = Φ

Set Operations

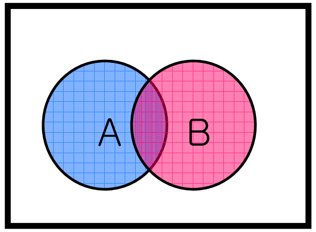
- The union:
  - $A \cup B$: the set of outcomes either in A or in B or both
- The intersection:
  - $A \cap B$: the set of outcomes in both A and B
  - If $A \cap B = \Phi$, then A and B are mutually exclusive
- The complement
  - A^c = The set of all outcomes not in A

Diagrams of Set Operations

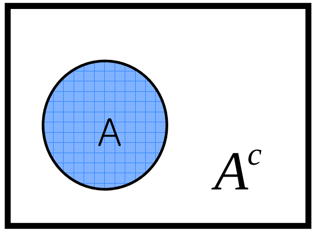
$A \bigcup B$



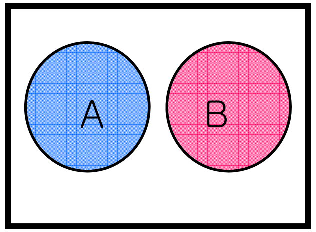
$A \bigcap B$



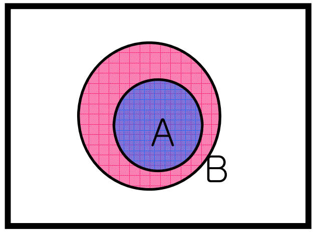
A^c



$A \bigcap B = \phi$



$A \subset B$



Cf.) If $A \subset B$, then A implies B

The Properties of Set Operations

- Commutative Properties
  - $A \cup B = B \cup A$ and $A \cap B = B \cap A$
- Associative Properties
  - $A \cup (B \cup C) = (A \cup B) \cup C$ and $A \cap (B \cap C) = (A \cap B) \cap C$
- Distributive Properties
  - $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
  - $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- De Morgan's Rules
  - $(A \cap B)^c = A^c \cup B^c$
  - $(A \cup B)^c = A^c \cap B^c$

The Properties of Set Operations (cont'd)

$$\bigcup_{k=1}^{n} A_k = A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n$$

$$\bigcap_{k=1}^{n} A_k = A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n$$

where n can approach infinite

2.2 The Axioms of Probability

- Probability law
  - A function that assigns a number to sets(events)
  - A rule that assigns a number $P[A]$ to each event A for the experiment E
- Axiom I: $0 \leq P[A]$
- Axiom II: $P[S] = 1$
- Axiom III: If $A \cap B = \Phi$, then $P[A \cup B] = P[A] + P[B]$
- Axiom III'(infinite sample space): If $A_1, A_2, \ldots$ is a sequence of events such that $A_i \cap A_j = \Phi$ for all $\mathrm{i} \neq \mathrm{j}$, then

$$P\left[\bigcup_{k=1}^{\infty} A_k\right] = \sum_{k=1}^{\infty} P[A_k]$$

Corollaries of Probability

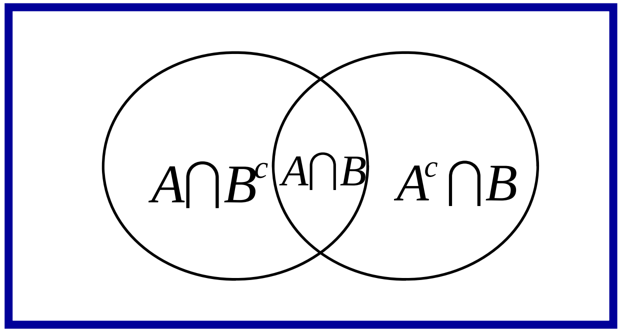
- Corollary 1: $P[A^c] = 1 - P[A]$
- Corollary 2: $P[\mathrm{A}] \leq 1$
- Corollary 3: $P[\Phi] = 0$
- Corollary 4:

  If $A_1, A_2, \ldots A_n$ are pairwise mutually exclusive, then

  $$P\left[\bigcup_{k=1}^{n} A_k\right] = \sum_{k=1}^{n} P[A_k] \qquad \text{for } n \geq 2.$$

  (proved by mathematical induction)

- Corollary 5: $P[A \cup B] = P[A] + P[B] - P[A \cap B]$
  - Proof)



  - Generalization to three events: $P[A \cup B \cup C] = P[A] + P[B] + P[C] - P[A \cap B] - P[B \cap C] - P[C \cap A] + P[A \cap B \cap C]$
  - $P[A \cup B] \le P[A] + P[B]$

- Corollary 6:

$$P\left[\bigcup_{k=1}^{n} A_k\right] = \sum_{j=1}^{n} P[A_j] - \sum_{j<k} P[A_j \cap A_k] + \cdots + (-1)^{n+1} P[A_1 \cap \cdots \cap A_n]$$

- Corollary 7: If $A \subset B$, then $P[A] \le P[B]$
  - Proof: $P[B] = P[A] + P[A^c \cap B] \ge P[A]$

Discrete Sample Spaces

- Probability law for an experiment with a countable sample space can be specified by giving the probabilities of the elementary events
- Distinct elementary events are mutually exclusive
- $$P[B] = P[\{a_1', a_2', \ldots, a_n'\}]$$
  $$= P[\{a_1'\}] + P[\{a_2'\}] + \cdots + P[\{a_n'\}]$$

Discrete Sample Spaces

- Probability assignment of equally likely outcomes for sample space of $S = \{a_1, a_2, \ldots, a_n\}$

$$P[\{a_1\}] = P[\{a_2\}] = \cdots = P[\{a_n\}] = \frac{1}{n}$$

- Ex. 2.6: Selecting a ball from urn containing 10 identical balls numbered 0,1,…,9.
  - A = number of ball selected is odd:

$$P[A] = P[\{1\}] + P[\{3\}] + P[\{5\}] + P[\{7\}] + P[\{9\}] = \frac{5}{10}$$

  - B = number of ball selected is a multiple of 3:

$$P[B] = P[\{3\}] + P[\{6\}] + P[\{9\}] = \frac{3}{10}$$

- $P[A \cup B] = P[A] + P[B] - P[A \cap B] = \dfrac{6}{10}$

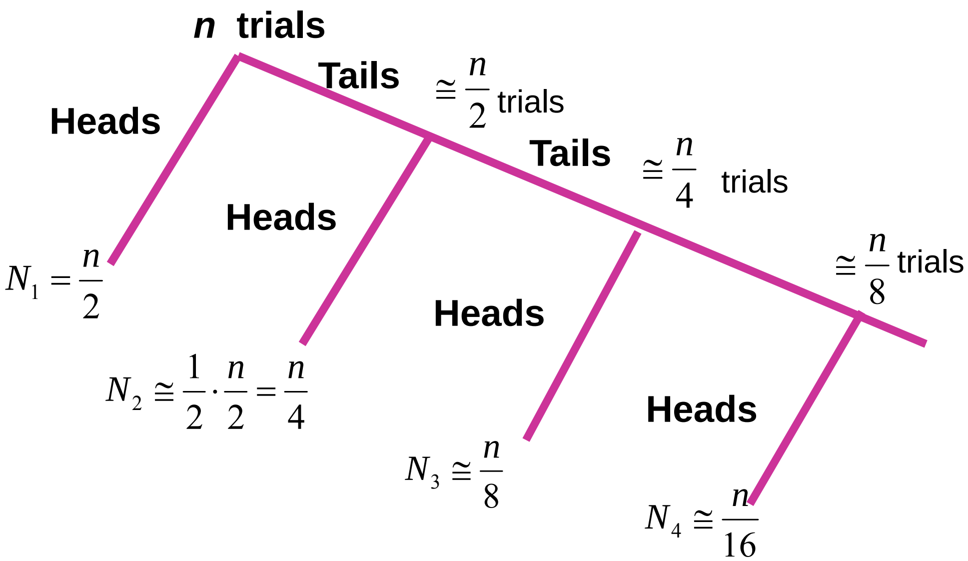
- C = number of ball selected is < 5:

$$P[C] = P[\{0\}] + P[\{1\}] + P[\{2\}] + P[\{3\}] + P[\{4\}] = \frac{5}{10}$$

$$\begin{aligned} P[A \cup B \cup C] &= P[A] + P[B] + P[C] - P[A \cap B] \\ &\quad - P[B \cap C] - P[C \cap A] + P[A \cap B \cap C] \\ &= \frac{9}{10} \end{aligned}$$

- Ex. 2.7
  - More than one reasonable probability assignment → experimental evidence is required to decide on the appropriate assignment

- Ex. 2.8
  - A fair coin is tossed repeatedly until the first heads shows up: the sample space $S=\{1,2,3, \ldots\}$
  - n times of experiments



- N_j : the number of trials in which the j th toss results in the first heads
- n is very large

- Relative frequency: $f_j \approx \frac{N_j}{n} = \left(\frac{1}{2}\right)^j, \qquad j = 1, 2, \ldots.$

- $P[j$ tosses till first heads$] = \left(\frac{1}{2}\right)^j, \qquad j = 1, 2, \ldots.$

Continuous Sample Spaces

- Outcomes are numbers
- The events of interest
  - Intervals of the real line
  - Rectangular regions in the plane
  - Complements, unions, intersections of these events
- Probability law is a rule for assigning numbers to intervals of the real line and rectangular regions in the plane

❖ Ex. 2.9

- Experiment : "Pick a number of the real line between 0 and 1"
- P[the outcome being exactly equal to ½] = 0

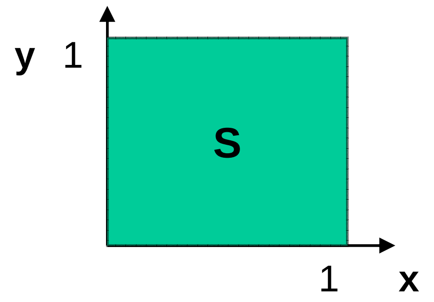
since there are uncountably infinite number of equally likely outcomes. $\because$ relative frequency $\dfrac{1}{\infty} = 0$

- The probability that the outcome falls in a subinterval of S

$$P[[a,b]] = (b-a) \quad \text{for} \quad 0 \le a \le b \le 1$$

❖ Ex. 2.11

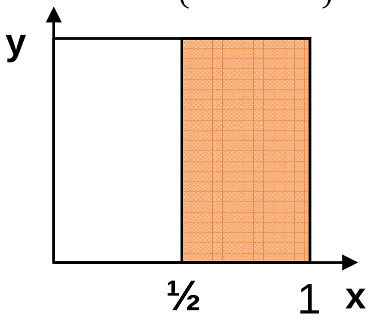
- Experiment E_{12} to pick two numbers x and y at random between zero and one.
- Sample space: the unit square



- Equally likely outcome of a pair of (x, y)
- Event $A = \{x > 0.5\}$, Event $B = \{y > 0.5\}$, Event $C = \{x > y\}$

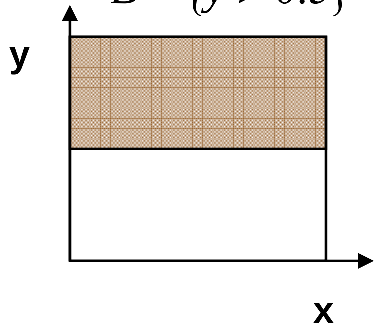
Equally likely outcome of a pair of x and y

$A = \{x > 0.5\}$



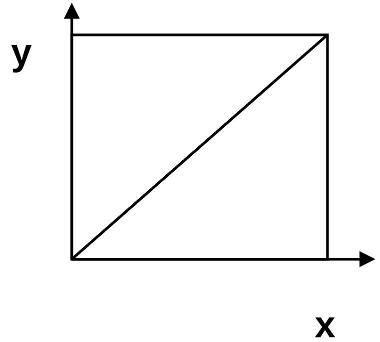
$P[A]=1/2$

$B = \{y > 0.5\}$



$P[B]=1/2$

$C = \{x > y\}$



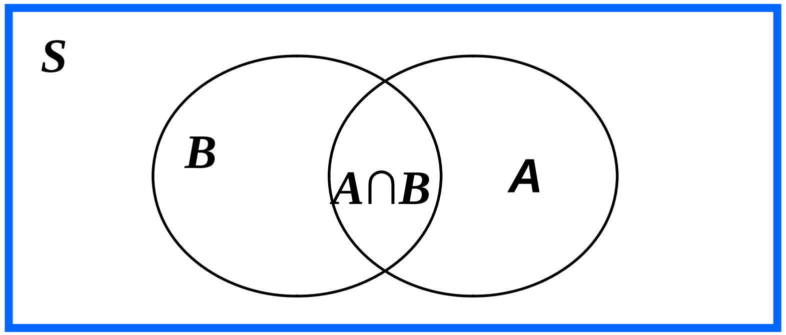
$P[C]=1/2$

H.W.: P2-7,14,15,16,18,24,28

2.4 Conditional Probability

- $P[A|B]$: Conditional Probability of event A given that event B has occurred

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} \qquad \text{for} \quad P[B] > 0$$



$\Rightarrow$ Renormalize the probability of events that occur jointly with B

cf) relative frequency

- n times of experiment, n_B times of event B.

  $n_{A\cap B}$ times of event $A\cap B$.

- The relative frequency of interest

$$\frac{n_{A\cap B}}{n_B} = \frac{n_{A\cap B}/n}{n_B/n} \rightarrow \frac{P[A\cap B]}{P[B]}$$

- $P[A\cap B] = P[A \mid B]P[B]$

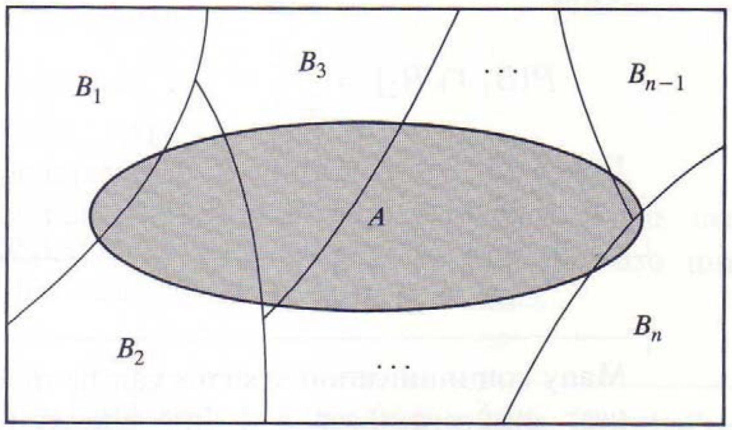
$$= P[B \mid A]P[A]$$

Partition

- Let $B_1, B_2, \ldots, B_n$ be mutually exclusive events.

- $B_1 \cup B_2 \cup \ldots \cup B_n = S$

- $B_1, B_2, \ldots, B_n$ form a partition of S



- $$A = A \cap S = A \cap (B_1 \cup B_2 \cup \cdots \cup B_n)$$
$$= (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_n)$$
$$\therefore P[A] = P[A \cap B_1] + P[A \cap B_2] + \cdots + P[A \cap B_n]$$
$$= P[A \mid B_1]P[B_1] + P[A \mid B_2]P[B_2] + \cdots + P[A \mid B_n]P[B_n]$$

- Ex)
  - Two black balls and three white balls
  - Two balls are sequentially selected at random without replacement
  - Probability of the event W_2 that the second ball is white.
  - B_1 = {(b,b), (b,w)}, W_1 = {(w,b), (w,w)}

    B_1, W_1 form a partition of S

$$\therefore P[W_2] = P[W_2 \mid B_1]P[B_1] + P[W_2 \mid W_1]P[W_1]$$

$$= \frac{3}{4}\cdot\frac{2}{5} + \frac{1}{2}\cdot\frac{3}{5} = \frac{3}{5}$$

조건부 확률 **(** ∵첫번째 **ball**이 **black**인 것이 가정된 상태**)**

Bayes' Rule

- $B_1, B_2, \ldots B_n$: a partition of a sample space S.

$$P[B_j \mid A] = \frac{P[B_j \cap A]}{P[A]} = \frac{P[A \mid B_j]P[B_j]}{\sum_{k=1}^{n} P[A \mid B_k]P[B_k]}$$

  - The "a priori probabilities" of these events, $P[B_j]$, are the probabilities of the events before the experiment is performed.
  - We are informed that event A occurred.
  - "a posteriori probabilities" are the probabilities of the events in the partition $P[B_j|A]$ given this additional information.

❖ Ex) Communication System

Which input is more probable given that the receiver has output a 1.

Assume that, "a priori", the input is equally likely to be 0 or 1.

Sol) A_k : the event having input k where k=0 or 1

$\Rightarrow A_0, A_1$: partition of the sample space of input-output pairs.

B_1 : event with output=1

$$P[B_1] = P[B_1 \mid A_0]P[A_0] + P[B_1 \mid A_1]P[A_1]$$

$$= \varepsilon\left(\frac{1}{2}\right) + (1-\varepsilon)\left(\frac{1}{2}\right) = \frac{1}{2}$$

❖ "a posteriori" probabilities

$$P[A_0 \mid B_1] = \frac{P[B_1 \mid A_0]P[A_0]}{P[B_1]} = \frac{\varepsilon/2}{1/2} = \varepsilon$$

$$P\left[A_1 \mid B_1\right] = \frac{P[B_1 \mid A_1]P[A_1]}{P[B_1]} = \frac{(1-\varepsilon)/2}{1/2} = (1-\varepsilon)$$

2.5 Independence of Events

- If $P[A \cap B] = P[A]P[B]$ then A and B are independent.
  - Since $P[A \mid B] = \dfrac{P[A \cap B]}{P[B]} = P[A]$
  - Implies that $P[A \mid B] = P[A]$, $P[B \mid A] = P[B]$

- Two events A and B
  - $P[A] \neq 0$, $P[B] \neq 0$, $P[A \cap B]=0$

  $\Rightarrow$ A and B cannot be independent.

  Proof) If they are independent

  $0 = P[A \cap B] = P[A]P[B] \Rightarrow$ contradiction

- $P[\mathrm{A}] = P[A \mid B]$
  - When the proportion of outcomes in S that lead to the occurrence of A is equal to the proportion of outcomes in B that lead to A.
  - Knowledge of the occurrence of B does not alter the probability of the occurrence of A

Independence of Three Events

- Two conditions for independence of three events:
- Pairwise independent
  - $P[A \cap B] = P[A]P[B],\ P[A \cap C] = P[A]P[C],\ P[B \cap C] = P[B]P[C]$

- The joint occurrence of any two should not affect the probability of the third.
  - i.e., $P[C \mid A \cap B] = P[C]$

$$\Rightarrow P[C \mid A \cap B] = \frac{P[A \cap B \cap C]}{P[A \cap B]} = P[C]$$

$$\Rightarrow P[A \cap B \cap C] = P[A \cap B]P[C] = P[A]P[B]P[C]$$

- "The three events A, B and C are independent if the probability of the intersection of any pair or triplet of events is equal to the product of the probabilities of the individual events."

cf) Pairwise independence of three events does not always mean that $P[A \cap B \cap C] = P[\mathrm{A}]P[B]P[C]$

- Ex) Two numbers are randomly selected from the unit interval
- $B=\{y>1/2\}$, $D=\{x<1/2\}$,
- $F=\{x<1/2 \;\&\; y<1/2\} \cup \{x>1/2 \;\&\; y>1/2\}$

- $P[B\cap D]=1/4= P[B]P[D]$,
- $P[B\cap F]=1/4= P[B]P[F]$,
- $P[D\cap F]=1/4= P[D]P[F]$, BUT
- $P[B\cap D\cap F]= P[\odot]= 0 \neq P[B]P[D]P[F]=1/8$

Independence of n Events

- A_1, A_2, … A_n are independent if for $k = 2,\ldots,$n.

  $$P[A_{i_1} \cap A_{i_2} \cap \cdots A_{i_k}] = P[A_{i_1}]P[A_{i_2}]\ldots P[A_{i_k}]$$

  where $1 \le i_1 < i_2 < \cdots < i_k \le n$

Sequential Experiments

- Sequences of independent experiments
  - Experiments $E_1, E_2, \ldots, E_n$

    Outcome $s = (s_1, s_2, \ldots, s_n)$

    Sample space $S = S_1 \times S_2 \times \ldots \times S_n$

  - If subexperiments are independent, outcomes of the subexperiments are independent.

    $$\mathrm{P}[A_1 \cap A_2 \cap \ldots \cap A_n] = P[A_1]P[A_2]\ldots P[A_n]$$

The Binomial Probability Law

- n : number of independent Bernoulli trials
- k : number of successes

❖ cf) Bernoulli trial : Performing an experiment once and noting a success or failure.

$$p_n(k) = \binom{n}{k} p^k (1-p)^{n-k}$$ $\Rightarrow$ **Probability of k successes in n trials**

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$ $\Rightarrow$ **Binomial coefficient**

$$N_n(k) = \binom{n}{k} = C_k^n$$ $\Rightarrow$ **Picking k positions out of n for the success**

❖ cf) $n!$ grows quickly with n. To avoid the evaluation of $n!$

$$\boxed{p_n(k+1) = \frac{(n-k)p}{(k+1)(1-p)} p_n(k)}$$

Recursive expression

➢ Proof)

$$p_n(k+1) = \binom{n}{k+1} p^{k+1} (1-p)^{n-k-1}$$

Multinomial Probability Law

- $B_1, B_2, \ldots, B_M$: a partition of the sample space S
- $P[B_j] = p_j$
  then $p_1 + p_2 + p_3 + \ldots + p_M = 1$
- $(k_1, k_2, \ldots k_M) \Rightarrow k_j$: the number of times event B_j occurs out of *n* independent repetitions of the experiment.

- The probability of the vector $(k_1, k_2, \ldots k_M)$

$$\boxed{P[(k_1, k_2, \ldots, k_M)] = \frac{n!}{k_1! k_2! \ldots k_M!} p_1^{k_1} p_2^{k_2} \cdots p_M^{k_M}}$$

H.W. Prove this.

where $k_1 + k_2 + \ldots + k_M = n$

Geometric Probability Law

- Sequential experiment until the occurrence of the first success.
  - m : the number of trials carried out until the occurrence of the first success.

$$\boxed{\begin{aligned} p(m) &= P[A_1^C A_2^C \ldots A_{m-1}^C A_m] \\ &= (1-p)^{m-1} p \,, \quad m = 1,2,\ldots \end{aligned}}$$

where p is the probability of success for the Bernoulli trial.

cf) $$\sum_{m=1}^{\infty} p(m) = p \sum_{m=1}^{\infty} q^{m-1} = p \frac{1}{1-q} = 1$$

where $q = 1 - p$

Sequences of Dependent Experiments

- Markov chain : the outcome of a given experiment determines which subexperiment is performed next.

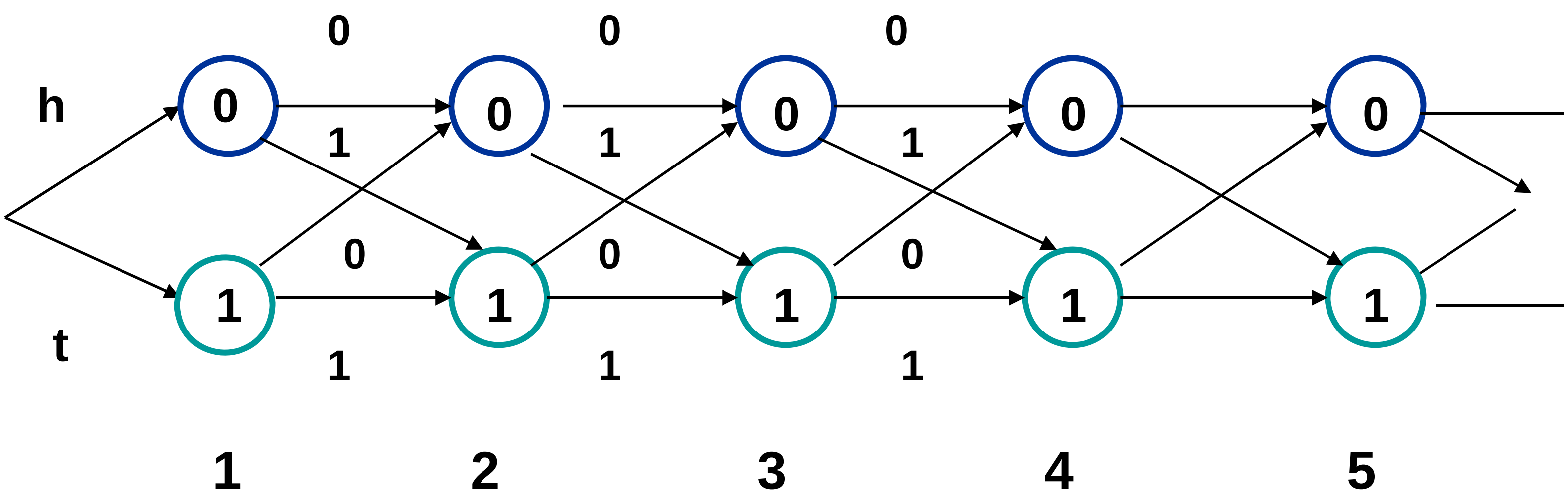
  - Ex. 2.41, 2.42.

① Toss a fair coin
- **Head →Urn 0**
- **Tails → Urn 1**

② Urn 0
- **a ball with the number 1 → Urn 1**
- **two balls with the number 0 → Urn 0**

Urn 1
- **5 balls with the number 1 → Urn 1**
- **1 ball with the number 0 → Urn 0**

③ Trellis Diagram



Each possible sequence corresponds to a path through the "trellis" diagram.

Probability of a particular sequence of outcomes, s_0, s_1, s_2.

- $P[\{s_0\}\cap\{s_1\}\cap\{s_2\}]$, let $A=\{s_2\}$, $B=\{s_1\}\cap\{s_0\}$
- $P[A \cap B] = P[A \mid B]P[B]$
  - $P[\{s_0\}\cap\{s_1\}\cap\{s_2\}] = P[\{s_2\} \mid \{s_0\}\cap\{s_1\}]\, P[\{s_0\}\cap\{s_1\}]$
    $= P[\{s_2\} \mid \{s_0\}\cap\{s_1\}]\, P[\{s_1\} \mid \{s_0\}]\, P[\{s_0\}]$

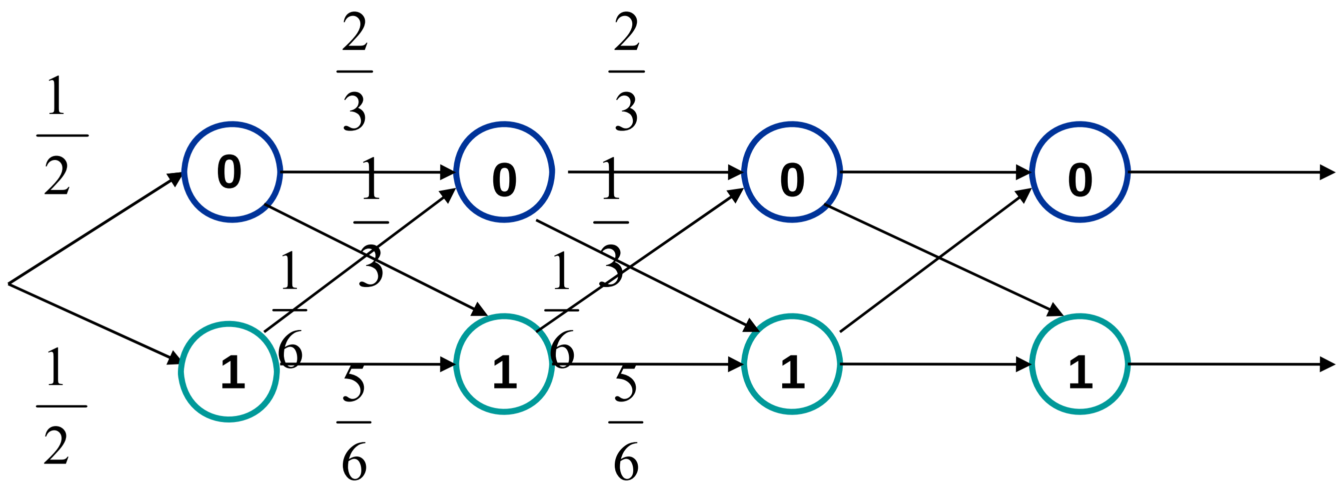
Markov Chain

cf) $\boxed{P[\{s_n\} \mid \{s_0\}\cap\{s_1\}\cap\ldots\cap\{s_{n-1}\}] = P[\{s_n\} \mid \{s_{n-1}\}]}$

Since the most recent outcome determine which subexperiment is performed in the above example.

- Therefore $P[\{s_0\}\cap\{s_1\}\cap\{s_2\}] = P[\{s_2\} \mid \{s_1\}]P[\{s_1\} \mid \{s_0\}]P[\{s_0\}]$

- For these experiments of Markov chains
  - $P[s_0, s_1, s_2, \ldots, s_n] = P[s_n \mid s_{n-1}]P\,[s_{n-1} \mid s_{n-2}]\ldots P[s_1 \mid s_0]P[s_0]$
- Ex)



$$P[0011] = P[1 \mid 1]P[1 \mid 0]P[0 \mid 0]P[0]$$

$$= \left(\frac{5}{6}\right)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)\left(\frac{1}{2}\right) = \frac{5}{54}$$

H.W.

- 1. See Appendix C.
  - ⇨ Random number generation : Pseudo-random
  - Generate a random binary number having 100 digits
  - Repeat the above 10 times with different seeds
  - Calculate the relative frequency of 0 and 1 for each random number
  - Calculate the difference between the relative frequency of 0 and 1

- 2. HW: 38,47,49,51,57,60,65,74,78,80,85