Events and probabilities

- Sample space (樣本空間): the set containing all possible outcomes of a random experiment. We denote a sample space by Ω. In the textbook, a sample space is often denoted by $\mathcal{C}$.
  - Events are represented by subsets of the sample space Ω.
  - We say that an event occurs if the outcome of the random experiment belongs to the event set.
- Example 1. Consider the experiment of rolling a die. The sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$ and the event of getting an even number is $\{2, 4, 6\}$.
- Notation. Suppose that Ω is a sample space and A and B are subsets of Ω.
  - $A \cup B$: the union (聯集) of A and B, representing the event that at least one of A, B occurs.
  - $A \cap B$: the intesection (交集) of A and B, representing the event that both A and B occur.
  - A^c: the complement (補集) of A. $A^c = \{x : x \in \Omega \text{ and } x \notin A\}$, representing the event that A does not occur.
- Disjoint sets (互斥集合).
  - Two sets A and B are disjoint (互斥) if $A \cap B = \emptyset$.
  - Suppose that $\mathcal{C}$ is a collection of sets such that any two sets $\mathcal{C}$ are disjoint. Then we say that sets in $\mathcal{C}$ are disjoint or $\mathcal{C}$ is a collection of disjoint sets.
- Given a sample space Ω, the probability of an event A is denoted by $P(A)$. P can be viewed as a function defined on $\mathcal{F}$: a collection of some subsets of Ω. The collection $\mathcal{F}$ is expected to satisfy the following properties:

  (i) $\mathcal{F}$ contains $\emptyset$ and Ω.

  (ii) $\mathcal{F}$ is closed under complements. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.

  (iii) $\mathcal{F}$ is closed under countable unions. If $A_n \in \mathcal{F}$ for $n = 1, 2, \ldots$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

  If $\mathcal{F}$ satisfies (i)–(iii), we say that $\mathcal{F}$ is a σ-field on Ω.
- Remark. A σ-field is also closed under countable intersections. That is, for a σ-field $\mathcal{F}$, if $A_n \in \mathcal{F}$ for $n = 1, 2, \ldots$, then $\cap_{n=1}^{\infty} A_n \in \mathcal{F}$.
- Example 2. Suppose that Ω is a sample space. Take $\mathcal{F}$ to be the collection of all subsets of Ω, then $\mathcal{F}$ is a σ-field on Ω. The collection $\mathcal{F}$ is called the power set (冪集) of Ω and is denoted by 2^{Ω}.

- σ-field generated by a collection. Suppose that Ω is a sample space and $\mathcal{C}$ is a collection of some subsets of Ω. The smallest σ-field on Ω that contains $\mathcal{C}$ is called the σ-field generated by $\mathcal{C}$ and denoted by $\sigma(\mathcal{C})$.

- Example 3. Suppose that $\Omega = \{1,2,3,4,5,6\}$, $A = \{1,3,5\}$ and $B = \{1,2,3\}$. Let $\mathcal{C} = \{A, B\}$. Find $\sigma(\mathcal{C})$: the σ-field (on Ω) generated by $\mathcal{C}$.

  A sketch of solution: Let $C_1 = A \cap B = \{1,3\}$, $C_2 = A \cap C_1^c = \{5\}$, $C_3 = B \cap C_1^c = \{2\}$, and $C_4 = (\cup_{i=1}^3 C_i)^c = \{4,6\}$. Take $\Lambda = \{1,2,3,4\}$ and

  $$\mathcal{F} = \{C_i : i \in \Lambda\} \cup \{C_i \cup C_j : i,j \in \Lambda\} \cup \{C_i \cup C_j \cup C_k : i,j,k \in \Lambda\} \cup \{\emptyset, \Omega\}.$$

  It is clear that the every element in $\mathcal{F}$ is in $\sigma(\mathcal{C})$ and $\mathcal{F}$ is a σ-field on Ω, so $\mathcal{F} = \sigma(\mathcal{C})$.

- Borel σ-field. The Borel σ-field on R^k is the σ-field $\sigma(\mathcal{C})$, where $\mathcal{C}$ is the collection of all open sets in R^k, denoted by $\mathcal{B}(R^k)$.

  - When $k = 1$, the Borel σ-field on R is denoted by $\mathcal{B}(R)$.
  - Let $\mathcal{C}_1 = \{(-\infty, t] : t \in R\}$, then $\sigma(\mathcal{C}_1) = \mathcal{B}(R)$.
  - Let $\mathcal{C}_2 = \{(a,b) : a, b \in R, a < b\}$, then $\sigma(\mathcal{C}_2) = \mathcal{B}(R)$.

- Definition of a probability function. Suppose that P is a function defined on $\mathcal{F}$: a σ-field on Ω. Then P is a probability function defined on $\mathcal{F}$ if P satisfies the following properties.

  (a) For $A \in \mathcal{F}$, $P(A) \geq 0$.

  (b) $P(\Omega) = 1$.

  (c) Additivity. Suppose that $\{A_n\}_{n \in I}$ is a sequence of disjoint events in $\mathcal{F}$. Then
  $$P\left(\bigcup_{n \in I} A_n\right) = \sum_{n \in I} P(A_n).$$
  Here the index set I can be $\{1, 2, \ldots\}$ or a finite set $\{1, 2, \ldots, N\}$.

- Note that the results listed below follow from Properties (a)–(c).

  - $P(\emptyset) = 0$.
  - $P(A) + P(A^c) = 1$.
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
  - For $A \subset B$, $P(A) \leq P(B)$.
  - $0 \leq P(A) \leq 1$.

- Example 4. Consider the $\mathcal{F}$ in Example 2. Let $A = \{1,3,5\}$, $B = \{1,2\}$ and $C = \{2\}$. Suppose that P is a function defined on $\mathcal{F}$. In which of the following cases, P cannot be a probability function?

(a) $P(A) = 0.2$, $P(B) = 0.3$, $P(C) = 0.2$
(b) $P(A) = 0.2$, $P(B) = 0.5$, $P(C) = 0.2$
(c) $P(A) = 0.3$, $P(B) = 0.4$, $P(C) = 0.2$

Ans. (b)

- Given a finite sample space $\Omega = \{o_1, \ldots, o_m\}$, one way to define a probability function P on $\mathcal{F}$: the σ-field of all subsets of Ω is to specify $P(\{o_i\})$ for each $i \in \{1, \ldots, m\}$. Then

$$P(A) = \sum_{o_i \in A} P(\{o_i\}). \tag{1}$$

In the special case where $P(\{o_i\}) = 1/m$ for $i = 1, \ldots, m$ (the equilikely case), (1) can be simplified to

$$P(A) = \frac{\text{number of elements in } A}{m}.$$

- Example 5. Consider the experiment of tossing a coin twice and record the result of heads and tails as elements in $\Omega = \{HH, HT, TH, TT\}$. Suppose that $P(\{HH\})$, $P(\{HT\})$, $P(\{TH\})$, $P(\{TT\})$ are p^2, $p(1-p)$, $p(1-p)$, and $(1-p)^2$ respectively, where $p \in (0,1)$. Then $P(A)$ can be determined for $A \subset \Omega$ when p is given. When $p = 0.5$, we have the equilikely case.

- Permutation and combination.
  - The number of ways of choosing k out of n objects is $C_k^n = n!/(k!(n-k)!)$.
  - The number of ways of choosing k out of n objects and order them is $P_k^n = n!/(n-k)!$.

- Example 6. Suppose that a group of 20 students have received 4 tickets of a special event. Suppose that 4 students will be selected by random to win the tickets. Suppose that among the 20 students, 15 are male and 5 are female. What is the probility that among the 4 selected students, 3 are male and 1 is female? Here it is assumed that the random selection is fair.

Ans.

$$\frac{C_3^{15} C_1^5}{C_4^{20}} = \frac{455}{969} \approx 0.4695562$$

- Example 7. 假設某彩券遊戲規則如下：玩家從1–365中選出一數字，開獎時也會從1–365中開出一數字，若所選數字和開出數字相同即爲中獎。假設該遊戲每期開獎爲獨立，且1–365中每個數字被開出的機率皆相同。求n期中所開出中獎數字均不重覆之機率。

Ans.

$$\frac{P_n^{365}}{365^n} = \frac{C_n^{365} \cdot n!}{365^n}$$

- R command for computing C_n^m is `choose(m,n)`.
- R command for computing $n!$ is `factorial(n)`.

- Write a function in R without using `choose` and `factorial` to compute the answer in Example 7:

```
ans.fun <- function(n){
 ans <- 1
 if (n >= 2){
   for (i in 2:n){  ans <- ans*(365-i+1)/365 }
 }
 return(ans)
}

### compute ans.fun(n) when n=10
ans.fun(10)
choose(365,10)*factorial(10)/(365^(10))
```

- Limit of a monotone sequence of sets. Suppose that $\{A_n\}_{n=1}^{\infty}$ is a sequence of sets.
  - Non-decreasing case. If $A_n \subset A_{n+1}$ for all n, then

$$\lim_{n\to\infty} A_n = \cup_{n=1}^{\infty} A_n.$$

  - Non-increasing case. If $A_n \supset A_{n+1}$ for all n, then

$$\lim_{n\to\infty} A_n = \cap_{n=1}^{\infty} A_n.$$

- Example 8. Suppose that $A_n = (-n, n)$ for $n \in \{1, 2, \ldots\}$. Then $\lim_{n\to\infty} A_n = (-\infty, \infty)$.
- Example 9. Suppose that $A_n = (-1/n, 1/n)$ for $n \in \{1, 2, \ldots\}$. Then $\lim_{n\to\infty} A_n = \{0\}$.
- Continuity. Suppose that $\mathcal{F}$ is a σ-field on Ω and P is a probability function defined on $\mathcal{F}$. Suppose that $\{A_n\}_{n=1}^{\infty}$ is a sequence of events in $\mathcal{F}$ such that $A_n \subset A_{n+1}$ for all n. Then

$$\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n). \tag{2}$$

  Proof. Let $B_1 = A_1$ and $B_n = A_n \cap A_{n-1}^c$ for $n \geq 2$. Then $\{B_n\}_{n=1}^{\infty}$ is a sequence of disjoint events. The result follows from additivity of P.
- Example 10. Suppose that $\mathcal{F}$ is a σ-field on $(-\infty, \infty)$ such that $(0, n) \in \mathcal{F}$ for each $n \in \{1, 2, \ldots\}$. Suppose that P is a probability function defined on $\mathcal{F}$ and

$$P((0, n)) = 0.5 \int_0^n e^{-x} dx$$

for each $n \in \{1, 2, \ldots\}$. Find $P((0, \infty))$.

Sol. Since the sequence $\{(0, n)\}_{n=1}^{\infty}$ is increasing and $\cup_{n=1}^{\infty}(0, n) = (0, \infty)$, by the continuity of P, we have

$$P((0, \infty)) = \lim_{n\to\infty} P((0, n)) = \lim_{n\to\infty} 0.5 \int_0^n e^{-x} dx = 0.5.$$

- Note that (2) also holds if $A_n \supset A_{n+1}$ for all n. The proof is left as an exercise.

- The inclusion exclusion formula. Suppose that A_1, ..., A_k are k distinct events.

$$P(A_1 \cup A_2 \cup \cdots \cup A_k) = \sum_{i=1}^{k} (-1)^{i+1} p_i,$$

where

$$p_i = \sum_{\{n_1,\ldots,n_i\} \in S_{k,i}} P(A_{n_1} \cap \cdots \cap A_{n_i})$$

$$S_{k,i} = \{\{n_1, \ldots, n_i\} : n_1, \ldots, n_i \text{ are } i \text{ distinct numbers in } \{1, \ldots, k\}\}$$

- In general, given A_1, ..., A_k, we have for $k \geq 2$ and $m \leq k/2$,

$$p_1 - p_2 + \cdots + p_{2m-1} \geq P(A_1 \cup A_2 \cup \cdots \cup A_k) \geq p_1 - p_2 + \cdots + p_{2m-1} - p_{2m}.$$

The specical case $m = 1$:

$$\sum_{i=1}^{k} P(A_i) \geq P(A_1 \cup A_2 \cup \cdots \cup A_k) \geq \sum_{i=1}^{k} P(A_i) - \sum_{1 \leq < i < j \leq k} P(A_i \cap A_j). \quad (3)$$

- In the textbook, the statement $p_1 \geq p_2 \geq \cdots \geq p_k$ is true for $k \leq 3$, but is not true in general.

- Events A and B are independent means $P(A \cap B) = P(A)P(B)$.
  - Events A_1, ..., A_k are independent means for any m distinct events B_1, ..., B_m in $\{A_1, \ldots, A_k\}$,

$$P(B_1 \cap \cdots \cap B_m) = P(B_1) \cdots P(B_m)$$

- Example 11. Suppose that A_1, A_2, A_3 and A_4 are four events such that A_i and A_j are independent for $i \neq j$. Suppose that $P(A_i) = a$ for $i = 1$, ..., 4. Give an upper bound and a lower bound for $P(\cup_{k=1}^4 A_k)$ when $a = 0.9$ and $a = 0.1$.

  Sol. Apply (3), then we have

$$4a \geq P(\cup_{k=1}^4 A_k) \geq a \times 4 - C_2^4 (a)^2 = 4a - 6a^2$$

  When $a = 0.1$, the lower bound is 0.34 and the upper bound is 0.4. When $a = 0.9$, the lower bound is -1.26 and the upper bound is 3.6. In such case, we can replace the lower bound and the upper bound by 0 and 1 respectively.

- Conditional probability
  - The conditional probability of event A given event B is denoted by $P(A|B)$, where $P(A|B) = P(A \cap B)/P(B)$.
  - If A and B are indepdent and $P(B) > 0$, then $P(A|B) = P(A)$.

– Bayes theorem. Suppose that $\{A_i\}_{i=1}^n$ is a sequence of disjoint events such that the sample space $\Omega = \cup_{i=1}^n A_i$, then given an event B such that $P(B) > 0$, we have

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1) + \cdots + P(B|A_n)P(A_n)}$$

for $i = 1, \ldots, n$.

- Exercise. Suppose that P is a probability function defined on $\mathcal{F}$: a σ-field on Ω and A is an event in $\mathcal{F}$ such that $P(A) > 0$. For $B \in \mathcal{F}$, define $Q(B) = P(B|A)$. Is Q a probability function defined on $\mathcal{F}$?

- Example 12. Suppose that $\mathcal{F}$ is a σ-field on Ω and P_1 and P_2 are two probability functions defined on $\mathcal{F}$. Define

$$Q(A) = \frac{P_1(A) + P_2(A)}{2}$$

for $A \in \mathcal{F}$. Show that Q is a probability function defined on $\mathcal{F}$.

Sol. We will show that Q is a probability function defined on $\mathcal{F}$ by verifying the following:

(a) $Q(A) \geq 0$ for all $A \in \mathcal{F}$.

(b) $Q(\Omega) = 1$.

(c) Suppose that $\{A_n\}_{n\in I}$ is a sequence of disjoint events in $\mathcal{F}$, then

$$Q(\cup_{n\in I} A_n) = \sum_{n\in I} Q(A_n). \tag{4}$$

– For (a), note that P_1 and P_2 are probability functions on $\mathcal{F}$, so $Q(A) = (P_1(A) + P_2(A))/2 \geq 0$ since $P_1(A) \geq 0$ and $P_2(A) \geq 0$.

– For (b), $Q(\Omega) = (P_1(\Omega) + P_2(\Omega))/2 = (1+1)/2 = 1$. Here $P_1(\Omega) = 1$ and $P_2(\Omega) = 1$ since P_1 and P_2 are probability functions on $\mathcal{F}$.

– For (c), suppose that $\{A_n\}_{n\in I}$ is a sequence of disjoint events in $\mathcal{F}$. Since P_1 and P_2 are probability functions on $\mathcal{F}$, we have

$$\begin{cases} P_1(\cup_{n\in I}(A_n)) = \sum_{n\in I} P_1(A_n) \\ P_2(\cup_{n\in I}(A_n)) = \sum_{n\in I} P_2(A_n) \end{cases} \tag{5}$$

Compute $Q(\cup_{n\in I} A_n)$ using the definition of Q and we have

$$\begin{aligned} Q(\cup_{n\in I} A_n) &= (P_1(\cup_{n\in I} A_n) + P_2(\cup_{n\in I} A_n))/2 \\ &\overset{(5)}{=} \frac{1}{2}\left(\sum_{n\in I} P_1(A_n) + \sum_{n\in I} P_2(A_n)\right) \\ &= \sum_{n\in I} \frac{P_1(A_n) + P_2(A_n)}{2} = \sum_{n\in I} Q(A_n), \end{aligned}$$

so (4) holds.