MATH 461 LECTURE NOTE
WEEK 2

DAESUNG KIM

1. Axioms of Probability I (Sec 2.2-3)

Sample Spaces and Events. We consider an experiment whose outcome is unpredictable. A sample space is the set of all possible outomes, which is usually denoted by S. For example,

(i) If the outcome of an experiment consists of the determination of the sex of a newborn child, then $S = \{g, b\}$.
(ii) If the experiment consists of flipping two coins, then $S = \{(H, H), (H, T), (T, H), (T, T)\}$.
(iii) If the experiment consists of tossing two dice, then $S = \{(i, j) : 1 \le i, j \le 6\}$.

Let S be a sample space. An event is a subset of the sample space S. For example, $\varnothing$ and S itself are also events.

(i) If $S = \{g, b\}$, then the events are $\varnothing$, $\{g\}$, $\{b\}$, and $\{g, b\}$.
(ii) If $S = \{(H, H), (H, T), (T, H), (T, T)\}$, then $\{(H, H), (H, T)\}$ (the first coin is head) is an event.
(iii) If $S = \{(i, j) : 1 \le i, j \le 6\}$, then $\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$ is an event.

Operations on Events

Let S be a sample space and E, F events.

(i) The union of E and F is an event, denoted by $E \cup F$, that either E or F occurs.
(ii) The intersection of E and F is an event, denoted by EF or $E \cap F$, that both E and F occur.
(iii) The complement of E is an event, denoted by E^c, that E does not occur.
(iv) If all outcomes in E are also in F, we say that E is contained in F, denoted by $E \subset F$.
(v) Let $E_1, E_2, \cdots$ be countably many events. The union of these events is an event, denoted by $\cup_{i=1}^{\infty} E_i$, that at least one of these event occurs. The intersection of these events is an event, denoted by $\cap_{i=1}^{\infty} E_i$, that all the events occur.

Example 1. Let $S = \{(H, H), (H, T), (T, H), (T, T)\}$. Let $E = \{\text{the first coin is head.}\} = \{(H, H), (H, T)\}$ and $F = \{\text{the first is different from the second.}\} = \{(H, T), (T, H)\}$. Then, $E \cup F = \{(H, H), (H, T), (T, H)\}$, $E \cap F = \{(H, T)\}$, and $E^c = \{(T, T), (T, H)\} = \{\text{the first is tail}\}$.

Example 2. We have $S^c = \varnothing$, $\varnothing^c = S$, and $\varnothing \subset E \cap F \subset E, F \subset E \cup F \subset S$.

Laws for operations

Let S be a sample space and E, F, G events.

(i) (Commutativity) $E \cup F = F \cup E$ and $EF = FE$.
(ii) (Associativity) $(E \cup F) \cup G = E \cup (F \cup G)$ and $(E \cap F) \cap G = E \cap (F \cap G)$.
(iii) (Distribution laws) $(E \cup F) \cap G = (E \cap G) \cup (F \cup G)$ and $(E \cap F) \cup G = (E \cup G) \cap (F \cup G)$.
(iv) (De Morgan's laws) For events $E_1, \cdots, E_n$,

$$\left(\bigcup_{k=1}^{n} E_k\right)^c = \bigcap_{k=1}^{n} E_k^c, \qquad \left(\bigcap_{k=1}^{n} E_k\right)^c = \bigcup_{k=1}^{n} E_k^c.$$

Example 3. Let E, F, G be events in a sample space S. Find expressions of the following events.

(i) Both E and G, but not F, occur: $E \cap G \cap F^c$.
(ii) At least two of the events occur: $(E \cap F) \cup (F \cap G) \cup (G \cap E)$.

(iii) At most one of them occur: the complement of the event in (ii). That is,

$$\begin{aligned}((E\cap F)\cup(F\cap G)\cup(G\cap E))^c &= (E\cap F)^c\cap(F\cap G)^c\cap(G\cap E)^c\\ &= (E^c\cup F^c)\cap(F^c\cup G^c)\cap(G^c\cup E^c).\end{aligned}$$

Axioms of Probability. An intuitive way of defining the probability of an event E is to consider its relative frequency. Suppose we perform the same experiments n times and count the number of occurance of the event E, denotes by $\phi_n(E)$. Then, we consider

$$\lim_{n\to\infty}\frac{\phi_n(E)}{n}$$

as a definition of probability of E. However, this definition leads to several delicate questions: how do we know that the limit exists? Even if so, how do we know the limit is consistent? And so on.

Modern probability theory starts from the axioms of probability and show that the relative frequency converges to its probability.

Axioms of Probability
Let S be a sample space. For each event E, the probability $\mathbb{P}(E)$ is an assignment so that the following axioms are satisfied:

(i) For all events, $0\le\mathbb{P}(E)\le 1$.
(ii) $\mathbb{P}(S)=1$.
(iii) For any sequence of mutually exclusive events $E_1,E_2,\cdots$, (meaning tthat $E_iE_j=\varnothing$ for $i\neq j$),

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty}E_i\right)=\sum_{i=1}^{\infty}\mathbb{P}(E_i).$$

The pair of a sample space and probability $(S,\mathbb{P})$ satisfying the three axioms is called a probabilty space.

Direct consequences are the following:

(i) $\mathbb{P}(\varnothing)=0$.
(ii) If $E_1,E_2,\cdots,E_n$ are mutually exclusive events, then

$$\mathbb{P}\left(\bigcup_{i=1}^{n}E_i\right)=\sum_{i=1}^{n}\mathbb{P}(E_i).$$

In particular if $A\cap B=\varnothing$, then $\mathbb{P}(A\cup B)=\mathbb{P}(A)+\mathbb{P}(B)$.

Example 4. Let $S=\{H,T\}$, then all the events are $\varnothing,\{H\},\{T\},S$. We assign the probabilities $\mathbb{P}(\varnothing)=0$, $\mathbb{P}(\{H\})=p$, $\mathbb{P}(\{T\})=1-p$, and $\mathbb{P}(S)=1$ for some $0\le p\le 1$. One can check that the three axioms are satisfied.

2. Axioms of Probability II (Sec 2.4)

Probability Space
Let S be a sample space. For each event E, the probability $\mathbb{P}(E)$ is an assignment so that the following axioms are satisfied:

(i) For all events, $0\le\mathbb{P}(E)\le 1$.
(ii) $\mathbb{P}(S)=1$.
(iii) For any sequence of mutually exclusive events $E_1,E_2,\cdots$, (meaning tthat $E_iE_j=\varnothing$ for $i\neq j$),

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty}E_i\right)=\sum_{i=1}^{\infty}\mathbb{P}(E_i).$$

The pair of a sample space and probability $(S,\mathbb{P})$ satisfying the three axioms is called a probabilty space.

Example 5. Let $S = \{1, 2, 3, 4, 5, 6\}$ and $\mathbb{P}(\{1\}) = \mathbb{P}(\{2\}) = \mathbb{P}(\{3\}) = \mathbb{P}(\{4\}) = \mathbb{P}(\{5\}) = \mathbb{P}(\{6\}) = \frac{1}{6}$. For an event E in S, we define the probability $\mathbb{P}(E)$ by

$$\mathbb{P}(E) = \sum_{i \in E} \mathbb{P}(\{i\}).$$

Then the three axioms are satisfied.

Proposition: Properties of Probability

(i) $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$
(ii) If $E \subset F$, then $\mathbb{P}(E) \leq \mathbb{P}(F)$.
(iii) For any two events E and F, $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(EF)$.

Proof. (i) Let $E_1 = E$, $E_2 = E^c$, and $E_i = \varnothing$ for all $i > 2$. The axiom (iii) yields that $\mathbb{P}(S) = 1 = \sum_{i=1}^{\infty} \mathbb{P}(E_i)$. Since $\mathbb{P}(\varnothing) = 0$ (why?), we have $1 = \mathbb{P}(E) + \mathbb{P}(E^c)$.

(ii) Let $E_1 = E$, $E_2 = F \cap E^c$, and $E_i = \varnothing$ for all $i > 2$. The axiom (iii) yields that $\mathbb{P}(F) = \sum_{i=1}^{\infty} \mathbb{P}(E_i) = \mathbb{P}(E) + \mathbb{P}(F \cap E^c)$. Since $\mathbb{P}(F \cap E^c) \geq 0$ by the axiom (i), we conclude that $\mathbb{P}(E) \leq \mathbb{P}(F)$.

(iii) Using the axiom (iii), one has

$$\begin{aligned}
\mathbb{P}(E \cup F) &= \mathbb{P}(EF^c) + \mathbb{P}(EF) + \mathbb{P}(FE^c) \\
\mathbb{P}(E) &= \mathbb{P}(EF^c) + \mathbb{P}(EF) \\
\mathbb{P}(F) &= \mathbb{P}(EF) + \mathbb{P}(FE^c),
\end{aligned}$$

which leads to

$$\mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cup F) = \mathbb{P}(EF).$$

$\square$

Example 6. J is taking two books along on her holiday vacation. With probability .5, she will like the first book; with probability .4, she will like the second book; and with probability .3, she will like both books. What is the probability that she likes neither book?

What if we have more than two events? Let E, F, G be events and consider $\mathbb{P}(E \cup F \cup G)$. Using the proposition (iii) for $(E \cup F)$ and G and for E and F, one has

$$\begin{aligned}
\mathbb{P}(E \cup F \cup G) &= \mathbb{P}(E \cup F) + \mathbb{P}(G) + \mathbb{P}((E \cup F)G) \\
&= \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(EF) + \mathbb{P}(G) + \mathbb{P}(EG \cup FG).
\end{aligned}$$

Using the same proposition for EG and FG, we have

$$\mathbb{P}(EG \cup FG) = \mathbb{P}(EG) + \mathbb{P}(FG) - \mathbb{P}(EFG).$$

Therefore, we get

$$\mathbb{P}(E \cup F \cup G) = \mathbb{P}(E) + \mathbb{P}(F) + \mathbb{P}(G) - \mathbb{P}(EF) - \mathbb{P}(FG) - \mathbb{P}(GE) + \mathbb{P}(EFG).$$

In general, we have the following.

Lemma: Inclusion-Exclusion Principle
For events $E_1, E_2, \cdots, E_n$,

$$\begin{aligned}
\mathbb{P}(E_1 \cup E_2 \cup \cdots E_n) = (\mathbb{P}(E_1) + \mathbb{P}(E_2) + \cdots + \mathbb{P}(E_n)) \\
&- (\mathbb{P}(E_1E_2) + \mathbb{P}(E_1E_3) + \cdots + \mathbb{P}(E_iE_j) + \cdots) + \cdots \\
&+ (-1)^{r+1} \sum \mathbb{P}(E_{i_1}E_{i_2} \cdots E_{i_r}) + \cdots \\
&+ (-1)^{n+1} \mathbb{P}(E_1E_2 \cdots E_n).
\end{aligned}$$

References

[SR] Sheldon Ross, *A First Course in Probability*, 9th Edition, Pearson

Department of Mathematics, University of Illinois at Urbana-Champaign
E-mail address: daesungk@illinois.edu