Chapter 2 Axioms of Probability

2.1 Introduction:

2.2 Sample Space and Events

S — Sample space: the set of all possible outcomes of an experiment

1. Flip a coin, $S = \{\text{tail}, \text{head}\}$
2. Flip two coins, $S = \{(H,H), (T,T), (H,T), (T,H)\}$.
3. Toss two dices, $S = \{(i,j): i,j = 1,2,3,4,5,6\}$
4. The order of finish in a race among 7 horses.
   $S = \{\text{all } 7! \text{ permutations of } (1,2,3,4,5,6,7)\}$
5. Measuring (in hours) the life time of a transistor,
   $S = \{x : 0 \le x < +\infty\}$

Event — Any subset E of the sample space is known as an event.
That is, an event is a set of some possible outcomes.
If the outcome of the experiment is contained in E, then we say the E has occured.

1. $E = \{\text{head}\}$, $F = \{\text{tail}\}$
2. $E = \{(H,H), (H,T)\}$ —— The first coin is head
   $F = \{(T,H), (H,T)\}$ —— The two coins have different appearances.

For E, obviously, if (H,H) happens, we can say it happens that the first coin is head. This explains that if the outcome of the experiment is contained in E, then we say the E has occured.

3. $E=\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$ — The sum of the dice is 7.
4. $E=\{$all permutations starting with 3$\}$ — Horse 3 wins the race.
5. $E=\{x: 0\le x\le 5\}$ — The transistor doesn't last longer than 5 hours.

We can define $E\cup F$ — the union of two events.

$EF\,(E\cap F)$ — the intersection of two events.

$$E_1\cup E_2\cup E_3\cdots=\bigcup_{n=1}^{\infty}E_n$$

$$E_1\cap E_2\cap E_3\cdots=\bigcap_{n=1}^{\infty}E_n$$

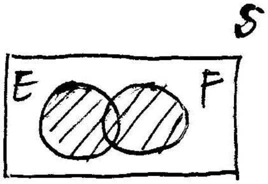
What if E_1 and E_2 have no common outcomes? $E_1E_2=\emptyset$

$E^c=\{$all outcomes not in $E\}$. $S^c=\emptyset$.

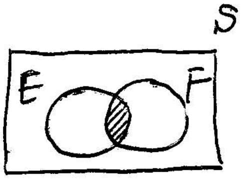
$E\subset F$ implies that all elements in E are contained by F. It implies that if E occurs, we can say F occurs.

$E\subset F \Leftrightarrow F\supset E$

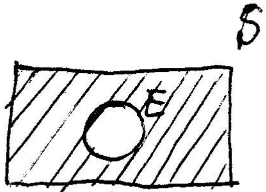
If $E\subset F$ and $F\subset E$, then $E=F$.



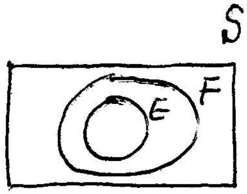
$E \cup F$



$EF = E \cap F$



E^c



$E \subset F$

Venn Diagram

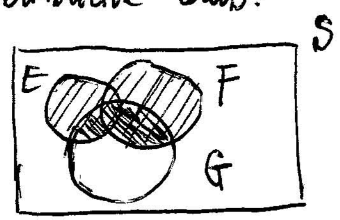
Commutative laws $\quad E \cup F = F \cup E, \quad EF = FE$

Associative laws $\quad (E \cup F) \cup G = E \cup (F \cup G), \quad (EF)G = E(FG)$

Distributive laws $\quad (E \cup F)G = (EG) \cup (FG), \quad (EF) \cup G = (E \cup G)(F \cup G)$.

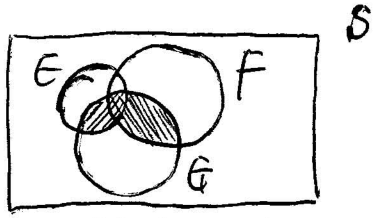
where the intersection "$\cap$" and union "$\cup$" have the same priority.

Proof of Distributive laws:



$(E \cup F)G$

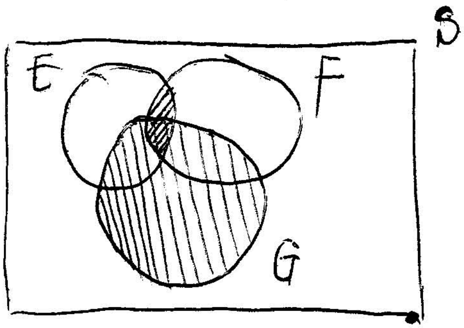
(the shaded area)



$(EG) \cup (FG)$

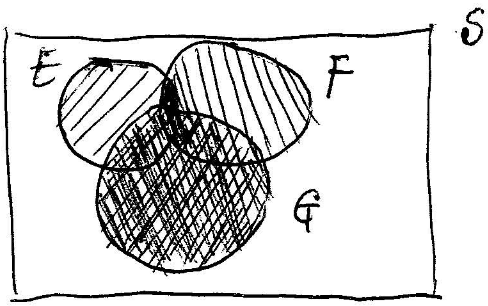
(the two striped areas).

Similarly,



$(EF) \cup G$

(the two striped areas)



$(E \cup G)(F \cup G)$

(the shaded area)

De Morgan's laws.

$$\left(\bigcup_{i=1}^{n} E_i\right)^c = \bigcap_{i=1}^{n} E_i^c$$

$$\left(\bigcap_{i=1}^{n} E_i\right)^c = \bigcup_{i=1}^{n} E_i^c$$

Proof: Suppose $x \in \left(\bigcup_{i=1}^{n} E_i\right)^c$

then $x \notin \bigcup_{i=1}^{n} E_i$,

then $x \notin E_i$, for $\overset{(all)}{any}$ $i = 1, 2, \cdots, n$.

then $x \in E_i^c$, for $\underset{(all)}{any}$ $i = 1, 2, \cdots, n$.

then, $x \in \bigcap_{i=1}^{n} E_i^c$

$\therefore \left(\bigcup_{i=1}^{n} E_i\right)^c \subset \bigcap_{i=1}^{n} E_i^c$,

Similarly, we can get $\bigcap_{i=1}^{n} E_i^c \subset \left(\bigcup_{i=1}^{n} E_i\right)^c$.

$\therefore$ the two sets are contained by each other. and if $A \subset B$, $B \subset A$, then $A = B$.

Therefore, $\left(\bigcup_{i=1}^{n} E_i\right)^c = \bigcap_{i=1}^{n} E_i^c$

2.3 Axioms of Probability

S — sample space. E — Event

Axiom 1. $0 \leq P(E) \leq 1$, for any event E.

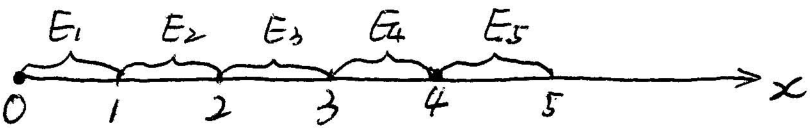
Axiom 2. $P(S) = 1$

Axiom 3. For any sequence of mutually exclusive events $E_1, E_2, \cdots$ (that is, $E_iE_j = \phi$, when $i \neq j$), $P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$

Example: Measuring (in hours) the lifetime of a transistor

$$S = \{ x : 0 \le x \le \infty \}$$

let $E_i = \{ x ; \ i-1 \le x < i \}$ $\quad i = 1, 2, \cdots$



$P(E_1) = \frac{1}{2}$, $P(E_2) = \frac{1}{4}$, $P(E_3) = \frac{1}{8}$, $\cdots$ $P(E_i) = \frac{1}{2^i}$

$$P(\bigcup_{i=1}^{\infty} E_i) = P(S) = 1$$

$$\sum_{i=1}^{\infty} P(E_i) = \sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 1$$

$\therefore$ $P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$, which confirms the axiom 3.

$$P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i) \Longrightarrow P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i)$$

Choose $E_1 = S$, $E_2 = \emptyset$, $E_3 = \emptyset$, $E_4 = \emptyset$, $\cdots$ ($E_i E_j = \emptyset$, when $i \neq j$).

then, $P(\bigcup_{i=1}^{\infty} E_i) = P(S)$.

$$\sum_{i=1}^{\infty} P(E_i) = P(S) + P(\emptyset) + P(\emptyset) + P(\emptyset) + \cdots$$

By axiom 3, $\quad P(S) = P(S) + P(\emptyset) + P(\emptyset) + \cdots$

$\therefore$ $P(\emptyset) = 0$.

To prove $P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i)$,

let $E_{n+1} = \emptyset$, $E_{n+2} = \emptyset$, $\cdots$

then $\bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{n} E_i$, $P(E_{n+1}) = 0$, $P(E_{n+2}) = 0$, $\cdots$

$\therefore$ $P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$ implies $P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i)$

Example: Toss a coin, if a head is as likely as a tail, then

$$P(\{H\}) = P(\{T\}) = \frac{1}{2}.$$

If the coin is biased, $P(\{H\}) = \frac{2}{3}$, $P(\{T\}) = \frac{1}{3}$.

Example: Roll a die, if all sides are equally likely,

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6}$$

$$P(\{2,4,6\}) = P(\{2\} \cup \{4\} \cup \{6\}) = P(\{2\}) + P(\{4\}) + P(\{6\}) = 3 \times \frac{1}{6} = \frac{1}{2}$$

2.4 Some Simple Propositions.

Proposition 4.1. $P(E^c) = 1 - P(E)$.

Proof: $E^c \cup E = S$, $E^c \cap E = \emptyset$

$P(S) = P(E) + P(E^c)$, due to $P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i)$

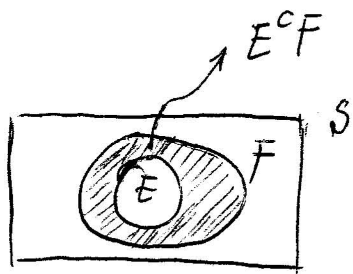
$\therefore$ $P(E) + P(E^c) = 1$,

$\therefore$ $P(E^c) = 1 - P(E)$.

Proposition 4.2 If $E \subset F$, then $P(E) \le P(F)$

Proof: $F = E \cup (E^c F)$

$\therefore$ $P(F) = P(E) + P(E^c F)$.

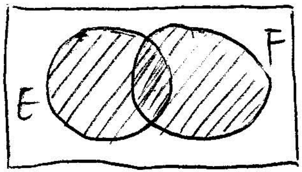


Since $P(E^c F) \geq 0$, $\therefore$ $P(E) \le P(F)$.

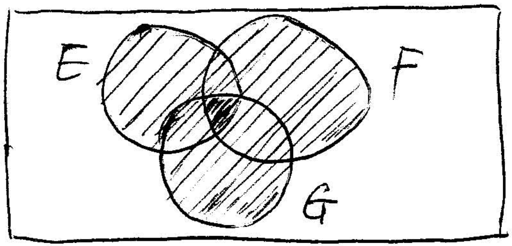
Proposition 4.3. $P(E \cup F) = P(E) + P(F) - P(EF)$

Proposition 4.4. $P(E \cup F \cup G) = P(E) + P(F) + P(G) - P(EF) - P(EG) - P(FG) + P(EFG)$

For Prop. 4.3



For Prop. 4.4



Think of $P(E \cup F)$ and $P(E \cup F \cup G)$ as the striped area respectively in the above two Venn diagrams. By calculating the areas of the two striped regions, we can get proposition 4.3 and 4.4 respectively.

$$P(E\cup F\cup G\cup H)= P(E)+P(F)+P(G)+P(H)-P(EF)-P(EG)-P(EH)-P(FG)-P(FH)-P(GH)$$
$$+P(EFG)+P(EFH)+P(EGH)+P(FGH)-P(EFGH)$$

Proposition 4.4 (Inclusion - Exclusion Identity)

$$P(E_1\cup E_2\cup\cdots\cup E_n)$$
$$=\sum_{i=1}^{n}P(E_i)-\sum_{i_1<i_2}P(E_{i_1}E_{i_2})+\cdots+(-1)^{r+1}\sum_{i_1<i_2<\cdots<i_r}P(E_{i_1}E_{i_2}\cdots E_{i_r})+\cdots+(-1)^{n+1}P(E_1\cdots E_n)$$

Example: Mike is going to take two courses A and B next term, With probability 0.8, he will like course A, with probability 0.7, he will like course B. With probability 0.6, he will like both courses.
What's the probability that he will like neither course?

Solution: $E=\{$mike will like course A$\}$; $F=\{$Mike will like course B$\}$;
then, $E\cup F=\{$Mike will like at least one of them$\}$.
Since $P(E\cup F)=P(E)+P(F)-P(EF)=0.8+0.7-0.6=0.9$.
$\therefore P\{$Mike will like neither$\}=1-P(E\cup F)=1-0.9=0.1$

2.5. Sample space having equally likely outcomes.

$S=\{1,2,\cdots,N\}$,
$P(\{1\})=P(\{2\})=\cdots=P(\{N\})=\frac{1}{N}$
Therefore, $P(E)=\dfrac{\text{number of outcomes in } E}{\text{total number of outcomes in } S}$

Example: If two dice are rolled, what's the probability that the sum of the dice will equal 7?

Solution: $E_7 = \{$the sum of the dice will equal 7$\} = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$.

The total number of outcomes in S is $6 \times 6 = 36$.

$\therefore P(E_7) = 6/36 = 1/6$

Similarly, $E_6 = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$

$\therefore P(E_6) = \frac{5}{36}$

Example: If 3 balls are "randomly drawn" from a bowl containing 6 white and 5 black balls, what is the probability that one is white and the other two are black?

Solution:

$$\frac{\binom{6}{1}\cdot\binom{5}{2}}{\binom{11}{3}} = \frac{6 \times 10}{\frac{11\times10\times9}{3\times2\times1}} = \frac{60}{165} = \frac{4}{11}$$

Example: A room of n people, what's the probability that no two of them celebrate their birthday on the same day of the year? How big should n be such that the prob. is less than $\frac{1}{2}$?

Solution: As each person can celebrate his or her birthday on any one of 365 days, there is a total of $(365)^n$ possible outcomes. Then, the desired probability is

$$P_0 = (365)(364)(363)\cdots(365-n+1)/(365)^n$$

Since P_0 is non-increasing with n, we can calculate the P_0 for the specific value of n and find out when $n \geq 23$, this probability is less than $\frac{1}{2}$. That is, if there are 23 or more people in a room, then the probability that at least two of them have the same birthday exceeds $\frac{1}{2}$, i.e. $(1-P_0) > \frac{1}{2}$. Furthermore, we can continue to get the following results:

when $n=50$, $1-P_0 \approx 97\%$,

when $n=100$, $1-P_0 \approx 1 - \frac{1}{3000000}$