Elementary Probability

Basic Concepts

- sample space: S
- events: $A, B, C, ... \subset S$
- elementary probability: $\mathbb{P}(A) = |A|/|S|$

Example Toss <u>one</u> regular, unloaded, 6-faced die. $A = \{$obtain an even number$\}$. Easy to see that

$$S = \{1, 2, ..., 6\}, \quad A = \{2, 4, 6\} \subset S, \quad \mathbb{P}(A) = |A|/|S| = 3/6.$$

Exercise Toss <u>two</u> (regular, ...) dice. $B = \{$obtain a <u>sum</u> of 10$\}$.

Answer

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Counting

Observation To use $\mathbb{P}(A) = |A|/|S|$, no need to "spell out" the sets $S = \{...\}, A = \{...\}$; just care about their sizes $|S|, |A|$.

Example Draw 5 cards from a regular deck of 52.

$\spadesuit$:	2, 3, ..., 10, J, Q, K, A
$\heartsuit$:	2, 3, ..., 10, J, Q, K, A
$\diamondsuit$:	2, 3, ..., 10, J, Q, K, A
$\clubsuit$:	2, 3, ..., 10, J, Q, K, A

Probability of obtaining a "full house", e.g., {K, K, K, 8, 8}?

Remark Can be hard and tricky, but will skip in this course.

Some Basic Rules

Notation

- S = sample space, ϕ = empty set, $A, B, C, \ldots$ = events
- A^c = the event "not A" [often denoted by $\bar{A}$ as well]
- $A \cup B$ = the event "A or B"
- $A \cap B$ = the event "A and B"

Rules

- $\mathbb{P}(S) = 1, \quad \mathbb{P}(\phi) = 0, \quad 0 \leq \mathbb{P}(A) \leq 1$ [obvious]
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ [obvious]
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ [Venn diagram]

Example

(*The Birthday Problem*)

Setting There are n persons in a room.

Question Probability of some common birthdays?

Exercise How large need n be for this probability to be $> 50\%$?

Conditional Probability

Definition

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

Intuition Again, use Venn diagram.

Example Toss two (regular, ...) dice. What's the probability of obtaining a sum of 10 given that the two dice did not have identical outcomes?

Sum of 10

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Given Non-identical Outcomes

~~(1,1)~~	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	~~(2,2)~~	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	~~(3,3)~~	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	~~(4,4)~~	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	~~(5,5)~~	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	~~(6,6)~~

Independence

Definition A and B are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Intuition Want $\mathbb{P}(A|B) = \mathbb{P}(A)$, i.e., the fact B happened does not change the probability of A. So

$$\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A) \quad \Rightarrow \quad \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Remark Generally, the joint probability of A and B must be calculated as $\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B)$ or $\mathbb{P}(B|A)\mathbb{P}(A)$.

Exercise 2.2 (p. 15)

Setting Ellen and Frank have a meeting. Let

$$E = \{\text{Ellen is late}\} \quad \text{and} \quad F = \{\text{Frank is late}\}.$$

Questions Suppose $\mathbb{P}(E) = 0.1$ and $\mathbb{P}(F) = 0.3$. What's the probability they can meet on time if

(a) E is independent of F?

(b) $\mathbb{P}(F|E) = 0.5$?

(c) $\mathbb{P}(F|E) = 0.1$?

Think In which case, (a), (b) or (c) above, is the probability (of meeting on time) the highest? Does this make intuitive sense?

Law of Total Probability

Theorem Suppose $B_1, B_2, ..., B_n$ is a partition of S such that

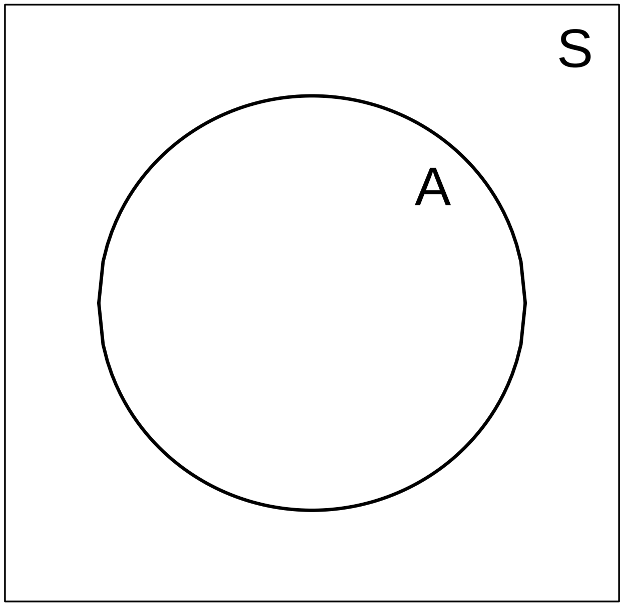
$$B_i \cap B_j = \phi \text{ for all } i \neq j \quad \text{and} \quad \bigcup_{i=1}^{n} B_i = S.$$

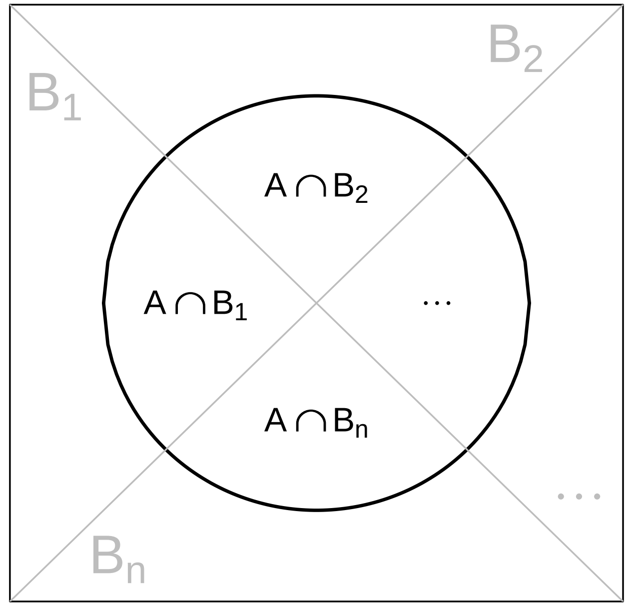
Then,

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \cap B_i) = \sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i).$$

Remark Deceptively simple but extremely useful and difficult to master! If $\mathbb{P}(A)$ is hard to determine, look for $B_1, B_2, ..., B_n$ so that each $\mathbb{P}(A|B_i)$ is easier to determine. Very fundamental technique.

Law of Total Probability





Bayes Theorem

Version 1 Since

$$\begin{aligned} \mathbb{P}(A \cap B) &= \mathbb{P}(A|B)\mathbb{P}(B) \\ \mathbb{P}(B \cap A) &= \mathbb{P}(B|A)\mathbb{P}(A), \end{aligned}$$

so

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)}.$$

Version 2 If $B_1, B_2, ..., B_n$ is a partition (of S), then

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{j=1}^{n} \mathbb{P}(A|B_j)\mathbb{P}(B_j)}.$$

Example

Setting Peter flips a coin. If he obtains "head", he draws a ticket from box A; if he obtains "tail", he draws a ticket from box B.

A:	5 red tickets	+	8 green tickets
B:	10 red tickets	+	3 green tickets

Question

(a) What's the probability Peter will obtain a red ticket?

(b) Given that Peter has obtained a red ticket, what's the probability that his initial coin flip resulted in "head"?

Example 2.4 (pp. 18–19)

Setting Consider an algorithm for predicting frauds. Suppose prevalence $= 0.5\%$, and the algorithm has a true positive rate of 98% and a false positive rate of 1%. That is,

$$\mathbb{P}(F) = 0.5\%, \quad \mathbb{P}(\oplus|F) = 98\%, \quad \text{and} \quad \mathbb{P}(\oplus|F^c) = 1\%.$$

Not a bad algorithm, right?

Question The algorithm has just flagged a transaction positive. What's the probability that it actually is fraudulent?

Example 2.3 (pp. 16–17)

Setting In a randomly shuffled deck of cards, there are n "regular" cards plus a joker. You and I take turns to draw a card from this deck until someone draws the joker. You go first.

Question What's the probability that you will get the joker?