Chapter 1. Experiments, Models, and Probabilities - *Getting Started with Probability*

Summary: Background and basic concepts on random experiments, models to describe outcomes of random experiments, definitions of probabilities and the fundamental theorems.

1.1 Set Theory **(Yates & Goodman 1.1)**

1.2 Applying Set Theory to Probability **(Yates & Goodman 1.2)**

1.3 Probability Axioms **(Yates & Goodman 1.3)**

1.4 Conditional Probability **(Yates & Goodman 1.4)**

1.5 The Law of Total Probability and Bayes' Theorem **(Yates & Goodman 1.5)**

1.6 Independence **(Yates & Goodman 1.6)**

1.1 Set Theory

The theory of sets serves as math basis for probability.

Def. A set is a collection of things. Each thing in the set is called an element.

Upper case letters for sets and lower case letters for elements.

Set inclusion: $\in$, to represent that an entity is in a set.

To represent that an entity is not in a set: $\notin$.

Set representation

- Tabular method (name the elements).
- Use math rules.

Examples:

$$A = \{heads, tails\}.$$

$$B = \{red, yellow, blue, Alice, Bob\}.$$

$$C = \{x | x^2 + 2x > 4\}.$$

$$a = heads \in A \qquad 5 \in C \qquad black \notin B$$

Def. ***A*** **is a <u>subset</u> of** ***B*** **if every element of** ***A*** **is also an element of** ***B*****.**

for all $x \in A \Rightarrow x \in B$

'for all' $= \forall$

Notation:

A is a subset of B: $A \subset B$ A is not a subset of B: $A \not\subset B$.

Examples:

$$D = \{n^2 - 1 \,| n \text{ is an integer and } 5 \leq n \leq 100.\}$$

$$E = \{y| -1 < y < 1\}.$$

$$D \subset C \quad E \not\subset C.$$

Def. The <u>empty set</u> or <u>null set</u> is a set with no elements.

Notation: $\emptyset$

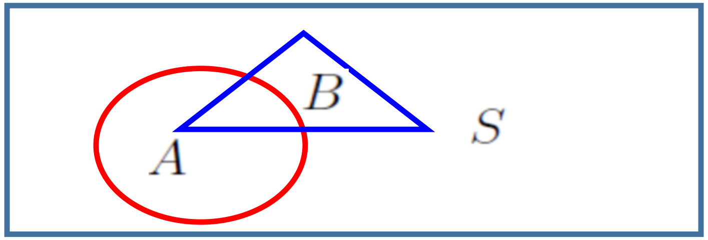
The <u>universal set</u> (S) is the largest set of all objects we consider in a given context.

Die-rolling: $S = \{1, 2, 3, 4, 5, 6\}.$

Water level: $S = \{x|x \in \mathbb{R}\}.$

Any set (in the context) is a subset of S and the empty set is a subset of any set.

Venn Diagram: To show sets and set operations with graphs.

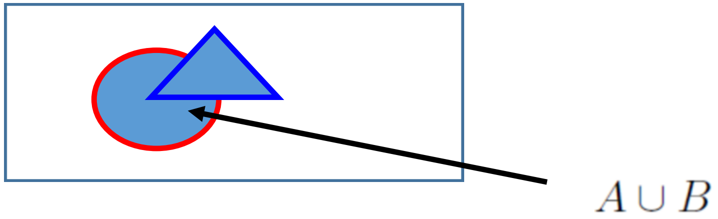


Basic set operations (Defs.)

1. Union. The union of two sets A and B is the set of elements that are either in A, or in B, or both.

Notation: $A \cup B$

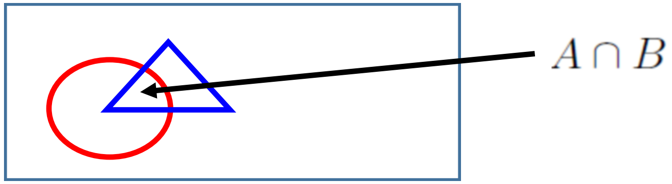
$$x \in A \cup B \text{ iff } x \in A \text{ or } x \in B$$



2. Intersection. The intersection of two sets A and B is the set of elements that are contained in both A and B.

Notation: $A \cap B$

$$x \in A \cap B \text{ iff } x \in A \text{ and } x \in B$$



3. Complement. The <u>complement</u> of a set A is the set of all elements in S that are not in A.

Notation: A^c

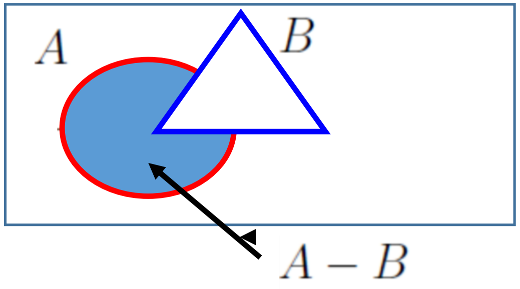
$$x \in A^c \text{ iff } x \notin A$$

<u>More on set operations</u>

<u>Def. Set equality:</u> $A = B$ iff $A \subset B$ and $B \subset A$.

<u>Def. Set difference</u> $A - B$ is the set of all elements in A that are not elements of B.

$$x \in A - B \text{ iff } x \in A \text{ and } x \notin B.$$



Example.

$S = \{1, 2, 3, 4, 5, 6\}$

$A = \{1, 2, 3, 5\} \quad B = \{2, 3, 6\}$

$A \cup B = \{1, 2, 3, 5, 6\} \quad A \cap B = \{2, 3\}$

$A^c = \{4, 6\} \quad B^c = \{1, 4, 5\} \quad A - B = \{1, 5\} \quad B - A = \{6\}$

Algebra of sets.

(a) Commutative:

$$A \cup B = B \cup A$$
$$A \cap B = B \cap A$$

(b) Distributive:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$
$$(A \cup B)^c = (A^c) \cap (B^c) \quad \text{(De Morgan's Law)}$$

Proof of De Morgan's law.

(c) $\boxed{A - B = A \cap B^c}$ $\quad A - A^c = A \quad A - A = \emptyset$

(d) Duality principle: by replacing

$\cup$ with $\cap$
$\cap$ with $\cup$
$\mathcal{S}$ with $\emptyset$
$\emptyset$ with $\mathcal{S}$
any set identity is preserved. $A \cup \emptyset = A \Rightarrow A \cap \mathcal{S} = A$

Def. Two sets are disjoint iff their intersection set is empty.

$$A \text{ and } B \text{ are disjoint } \Leftrightarrow A \cap B = \emptyset$$

Def. A collection of sets $A_1, A_2, \cdots, A_N$ **is mutually exclusive iff any two sets in the collection are disjoint.**

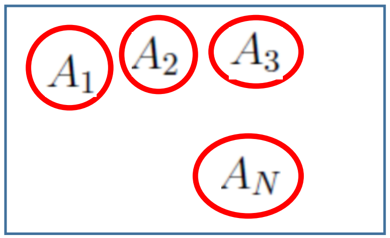
$$A_1, A_2, \cdots, A_N \text{ are mutually exclusive } \Leftrightarrow A_i \cap A_j = \emptyset \text{ for } i \neq j$$

Def. A collection of sets $A_1, A_2, \cdots, A_N$ **is collectively exhaustive iff**

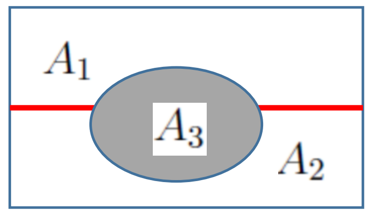
$$A_1 \cup A_2 \cup \cdots \cup A_N = S$$

Notation: $A_1 \cap A_2 \cap \cdots \cap A_N = \bigcap_{i=1}^{N} A_i \qquad A_1 \cup A_2 \cup \cdots \cup A_N = \bigcup_{i=1}^{N} A_i$

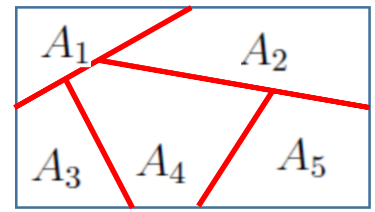
A collection of sets is a <u>partition</u> if it is both mutually exclusive and collectively exhaustive.



Mutually exclusive



Collectively exhaustive



Partition

Examples.

$S = \{1, 2, 3, 4, 5, 6\}$

$A_1 = \{1, 3, 5\} \quad A_2 = \{2, 4, 6\}$

$B_1 = \{1\} \quad B_2 = \{2\} \quad B_3 = \{4, 5\}$

$C_1 = \{1, 2, 3, 4\} \quad C_2 = \{3, 4, 5, 6\} \quad C_3 = \{6\}$

A_1, A_2 form a partition. B_1, B_2, B_3 are mutually exclusive. C_1, C_2, C_3 are collectively exhaustive.

1.2 Applying Set Theory to Probability

Repeatable experiments (with uncertain results), e.g., coin flipping, die rolling, random walk, digital communications.

*Probability theory is to **model and understand** random experiments.*

An experiment consists of a procedure and observations.

Example 1. Coin flipping.

Procedure: flip a coin. Observation: which side is up (H or T).

Example 2. Roll two dice.

Procedure: roll the dice. Observation: the points on the dice.

Example 3. Waiting time for bus.

Procedure: Go to bus stop and wait for a bus to come.

Observation: how many minutes is the waiting.

Def. An outcome is any possible observation of the experiment.

Def. The sample space S of an experiment is the set of all possible outcomes.

Def. An event is a set of outcomes.

Example 1. Coin flipping.

Outcomes: heads, tails.

Sample space: $S = \{heads, tails\}.$

One event is: 'heads' is shown. $E_1 = \{heads\}.$

Example 2. Roll two dice.

Outcomes: *(1,1), (1,2)*, …

Sample space: {*(1,1), (1,2), (1,3), …, (2,1), (2,2), …, (6,6)*}.

One event: The sum is at least 11. $E_1 = \{(5,6), (6,5), (6,6)\}$

Another event: The second die is 6.

$$E_2 = \{(i,6)\ |i = 1,2,3,4,5,6\}$$

Example 3. Waiting time for bus.

Possible outcomes: 0, 0.3, 2.5, pi, … 100.

Sample space: $S = \{x | x \geq 0\} = \mathbb{R}^{+}$

One event: Waiting time is no more than 10 minutes.

$$E_1 = \{x \, | x \in [0, 10]\}$$

Set Theory	Probability Theory
Element	Outcome
Universal set	Sample space
Set	Event
Partition	Event space

Def. An <u>event space</u> is a collectively exhaustive and mutually exclusive set of events. A partition of the sample space *S*.

Example 2 cont'd. An event space is $\{B_1, B_2, B_3\}$, where

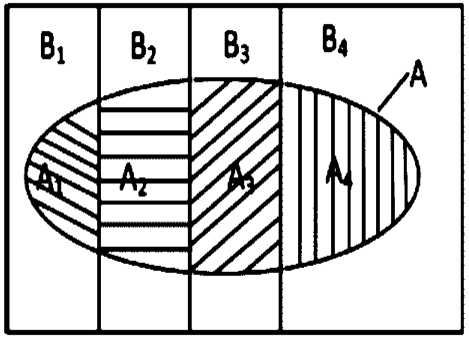
$$B_1 = \{(i,j)|i > j,\ i,j = 1,2,3,4,5,6\} = \{\text{The 1st die has a larger number.}\}$$

$$B_2 = \{(i,j)|i < j,\ i,j = 1,2,3,4,5,6\} = \{\text{The 2st die has a larger number.}\}$$

$$B_3 = \{(i,j)|i = j,\ i,j = 1,2,3,4,5,6\} = \{\text{The 2 dice have the same number.}\}$$

<u>Theorem.</u> For an event space (partition) $\{B_1, B_2, \cdots, B_N\}$ and any event A of the sample space, define $A_i = A \cap B_i$ for $i = 1, \cdots, N$. We have

$$A_i \cap A_j = \emptyset \text{ for } i \neq j \quad \text{and} \quad A = \bigcup_{i=1}^{N} A_i$$



That is, $\{A_1, A_2, \cdots, A_N\}$ defines a partition of A .

1.3 Probability Axioms

*Probability is used to **quantitatively** measure (random) events, a branch of measure theory.*

A probability model assigns a number between 0 and 1 to every event, to measure how likely the event occurs.

Relative frequency:

Repeat on an experiment n times, and denote the number of times that event A occurs by $N_A(n)$. The relative frequency of A is defined as

$$f_A = \frac{N_A(n)}{n}.$$

The probability of the union of mutually exclusive events is the sum of the probabilities of the events in the union.

Def. A <u>probability measure or probability</u> $\mathrm{P}[\cdot]$ is a function that maps events in the sample space to real numbers such that

Axiom 1. For any event A , $\mathrm{P}[A] \geq 0$.

Axiom 2. $\mathrm{P}[S] = 1$.

Axiom 3. For any countable collection $A_1, A_2, \cdots$ of *mutually exclusive* events,

$$\mathrm{P}[A_1 \cup A_2 \cup \cdots] = \mathrm{P}[A_1] + \mathrm{P}[A_2] + \cdots$$

If A_1 and $A_2$are **disjoint**, $\mathrm{P}[A_1 \cup A_2] = \mathrm{P}[A_1] + \mathrm{P}[A_2]$.

Event A $\xrightarrow{\mathrm{P}[\cdot]}$ $\mathrm{P}[A] \in$ [0,1]

$\boxed{\text{Set of all events}}$ $\xrightarrow{\mathrm{P}[\cdot]}$ $\boxed{[0,1]}$

Example. Roll a 6-sided die.

Sample space: $S = \{1, 2, 3, 4, 5, 6\}$

A natural probability measure (a map from sample space to real numbers):

$$\boxed{\begin{array}{lll} \{1\} \to 1/6 & \{2\} \to 1/6 & \{3\} \to 1/6 \\ \{4\} \to 1/6 & \{5\} \to 1/6 & \{6\} \to 1/6 \end{array}}$$

$$\boxed{\mathrm{P}[\text{the result is } i] = \mathrm{P}[\{i\}] = 1/6, \; i = 1, 2, \cdots, 6.}$$

The probability of any other event can be calculated from the above.

Find the probabilities of the following events.

$E_i = \{i\}, i = 1, \cdots, 6$

$A_1 = \{2, 4, 6\} = \{\text{The number is even.}\}$

$A_2 = \{1, 3, 5\} = \{\text{The number is odd.}\}$

$A_3 = \{3, 6\} = \{\text{A multiple of } 3\}$

$$\mathrm{P}[E_i] = 1/6$$

$$\mathrm{P}[A_1] = \mathrm{P}[E_2 \cup E_4 \cup E_6] = \mathrm{P}[E_2] + \mathrm{P}[E_4] + \mathrm{P}[E_6] = 1/2$$

$$\mathrm{P}[A_2] = \cdots = 1/2$$

$$\mathrm{P}[A_3] = \cdots = 1/3$$

Connection to relative frequency.

- Roll the die 5 times and find the relative frequency.

Observation	1	2	3	4	5	6
Frequency	1	0	1	0	1	2
Relative freq.	1/5	0	1/5	0	1/5	2/5

- Roll the die one million times and find the relative frequency.
  Computer simulation results:

Observation	1	2	3	4	5	6
Frequency	166223	166399	166922	167059	166504	166893
Relative freq.	0.166	0.166	0.167	0.167	0.167	0.167

- What can we say from this?

$$\lim_{n\to\infty} \frac{N_A(n)}{n} = \mathrm{P}[A].$$

Note: a lot more on the convergence speed and finite *n* behavior of the relative frequency.

Example. Use Axioms of probability to show:

$$\text{(a) } \mathrm{P}[\emptyset]=0 \qquad \text{(b) } \mathrm{P}[A^c]=1-\mathrm{P}[A] \qquad \text{(c) } A\subset B \to \mathrm{P}[A]\le \mathrm{P}[B]$$

Solution:

(a) $\forall A, A\cap\emptyset=\emptyset$, disjoint

$\mathrm{P}[A\cup\emptyset]=\mathrm{P}[A]+\mathrm{P}[\emptyset]$,

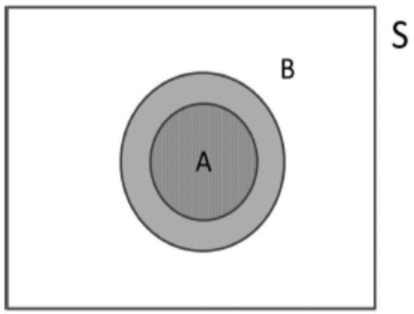
$\mathrm{P}[A]=\mathrm{P}[A]+\mathrm{P}[\emptyset] \Rightarrow \mathrm{P}[\emptyset]=0$

(b) $A\cap A^c=\emptyset \Rightarrow \mathrm{P}[A]+\mathrm{P}[A^c]=\mathrm{P}[\mathcal{S}] \Rightarrow \mathrm{P}[A^c]=1-\mathrm{P}[A]$

(c) Consider 2 events (sets) A and $B-A$.

$$A\cap(B-A)=A\cap(B\cap A^c)=(A\cap A^c)\cap B=\emptyset\cap B=\emptyset.$$

Thus A and $B-A$ are disjoint.



From Axiom 3,

$$\mathrm{P}[B]=\mathrm{P}[A\cup(B-A)]=\mathrm{P}[A]+\mathrm{P}[B-A]\ge \mathrm{P}[A],$$

where the last step is due to Axiom 1, $\mathrm{P}[B-A]\ge 0$.

Example. Flip two coins (The coins may not be fair).

Given $\mathrm{P}[\{HH\}] = 1/4, \mathrm{P}[\{HT\}] = 1/6, \mathrm{P}[\{TT\}] = 1/4.$

Find the probabilities of the following.

a) The result is *TH*.

b) At least one is tails.

c) There is 1 more tails than heads.

d) There is the same number of tails and heads.

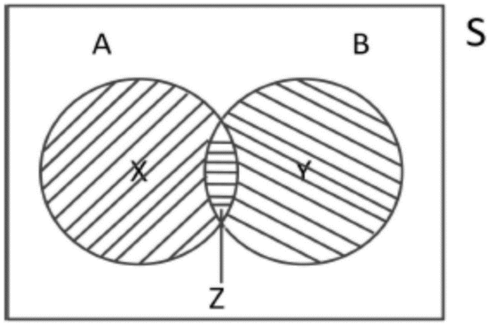
Solution: $S = \{HH, HT, TH, TT\}.$

$$a). \ \mathrm{P}[\{TH\}] = 1 - \mathrm{P}[\{HH, HT, TT\}] = 1 - (1/4 + 1/6 + 1/4) = 1/3.$$

$$b). \ \mathrm{P}[\text{At least 1 tail}s] = \mathrm{P}[\{HT, TH, TT\}] = 1/6 + 1/3 + 1/4 = 3/4.$$

$$c). \ \mathrm{P}[\text{1 more tails than heads}] = \mathrm{P}[\emptyset] = 0.$$

$$d). \ \mathrm{P}[\text{same number of tails and heads}] = \mathrm{P}[\{HT, TH\}] = 1/6 + 1/3 = 1/2.$$



Example. Show that for any events A and B,

$$\mathrm{P}[A \cup B] = \mathrm{P}[A] + \mathrm{P}[B] - \mathrm{P}[A \cap B]$$

Solution: Define $X = A - B, Y = B - A, Z = A \cap B$

Step 1. *X*, *Y*, and *Z* are mutually exclusive.

$$Y \cap Z = (B - A) \cap (A \cap B) = B \cap A^c \cap A \cap B = B \cap \emptyset = \emptyset$$

Similarly, it can be shown that X and Y are disjoint, X and Z are disjoint.

Step 2. $A \cup B = X \cup Y \cup Z$ and $A = X \cup Z, B = Y \cup Z$

Show that any element in the LHS set is in the RHS set, and vice versa, or show from the definition of *X,Y,Z*.

Thus, from Axiom 3,

$$\begin{aligned}
\mathrm{P}[A \cup B] &= \mathrm{P}[X \cup Y \cup Z] = \mathrm{P}[X] + \mathrm{P}[Y] + \mathrm{P}[Z] \\
&= \mathrm{P}[X] + \mathrm{P}[Z] + \mathrm{P}[Y] + \mathrm{P}[Z] - \mathrm{P}[Z] \\
&= \mathrm{P}[X \cup Z] + \mathrm{P}[Y \cup Z] - \mathrm{P}[Z] \\
&= \mathrm{P}[A] + \mathrm{P}[B] - \mathrm{P}[A \cap B].
\end{aligned}$$

Example. A person bought 2 tickets for 2 lotteries. The 1st ticket is a winning one with probability 0.1 and the second is a winning one with probability 0.05. The probability that both are winning ones is 0.01. What is the probability that he does not win?

Solution:

$$S = \{WW, WL, LW, LL\}.$$

$$A = \{\text{1st ticket wins}\} = \{WW, WL\}.$$

$$B = \{\text{2nd ticket wins}\} = \{WW, LW\}.$$

$$E = \{\text{He wins}\} = A \cup B. \quad A \cap B = \{WW\}.$$

$$\mathrm{P}[E] = \mathrm{P}[A \cup B] = \mathrm{P}[A] + \mathrm{P}[B] - \mathrm{P}[A \cap B] = 0.1 + 0.05 - 0.01 = 0.14.$$

$$\mathrm{P}[\text{no win}] = \mathrm{P}[E^c] = 1 - 0.14 = 0.86.$$

Equally likely outcomes. An experiment has sample space $S = \{s_1, s_2, \cdots, s_n\}$

The outcomes are equally likely if no one outcome is any more likely than any other. In this case,

$$\mathrm{P}[\{s_i\}] = 1/n, \text{or in simplified notation } \mathrm{P}[s_i] = 1/n.$$

If an event *A* is composed of *m* outcomes,

$$\mathrm{P}[A] = \frac{m}{n} = \frac{|A|}{|S|}.$$

Notation: $|A|$ is the cardinality of *A*. The number of elements in the set *A*.

Example. Roll a fair die. Each outcome is equally likely.

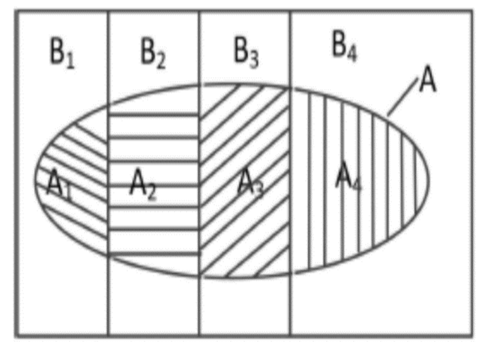
$$S = \{1, 2, 3, 4, 5, 6\}$$

$$\mathrm{P}[\{1\}] = \mathrm{P}[\{2\}] = \cdots = 1/6.$$

$$\mathrm{P}[\text{ 4 or higher}] = \mathrm{P}[\{4, 5, 6\}] = 1/2.$$

Theorem. For an event space (partition) $\{B_1, B_2, \cdots, B_N\}$ and any event A of the sample space, define $A_i = A \cap B_i$ for $i = 1, \cdots, N$. We have

$$\mathrm{P}[A] = \sum_{i=1}^{N} \mathrm{P}[A_i] = \sum_{i=1}^{N} \mathrm{P}[A \cap B_i]$$



Example. Mobile phones experience handoff as they move from cell to cell. During a call, a phone either performs zero handoff (B_1), one handoff (B_2) or more than one handoff (B_3). In addition, each call is either long (more than 3 minutes) (A_1) or short (not more than 3 minutes) (A_2).

The related probabilities are given in the following table. For example, the probability that a long call experience 0 handoff is 0.1. Find the probabilities of a long call, a short call, and a call that is long or has one handoff .

Notice that B_1 B_2 B_3 form an event space.

	B_1	B_2	B_3
A_1	0.1	0.1	0.2
A_2	0.4	0.1	0.1

$$(a) \quad \mathrm{P}[A_1] = \mathrm{P}[A_1 \cap B_1] + \mathrm{P}[A_1 \cap B_2] + \mathrm{P}[A_1 \cap B_3]$$

$$= 0.1 + 0.1 + 0.2 = 0.4.$$

$$(b) \quad \mathrm{P}[A_2] = \mathrm{P}[A_2 \cap B_1] + \mathrm{P}[A_2 \cap B_2] + \mathrm{P}[A_2 \cap B_3]$$

$$= 0.4 + 0.1 + 0.1 = 0.6.$$

$$\text{or } \mathrm{P}[A_2] = 1 - \mathrm{P}[A_1] = 1 - 0.4 = 0.6.$$

$$(c) \quad \mathrm{P}[A_1 \cup B_2] = \mathrm{P}[A_1] + \mathrm{P}[B_2] - \mathrm{P}[A_1 \cap B_2]$$

$$= 0.4 + 0.2 - 0.1 = 0.5,$$

$$\text{where } \mathrm{P}[B_2] = \mathrm{P}[B_2 \cap A_1] + \mathrm{P}[B_2 \cap A_2] = 0.2.$$

<u>1.4 Conditional Probability</u>

A modified probability model that reflects partial information about the outcome of an experiment. The modified model has a smaller sample space than the original model.

Example. Die-rolling. What is the probability 1 occurs? What is the probability if the outcome is known to be odd?

Example. The chance of developing diabetes by age 70 is 1%. What about the chance if a direct relative has developed diabetes?

Notation:

Probability of event *A*: $\mathrm{P}[A]$

Probability of event *A* given that event *B* has occurred: $\mathrm{P}[A|B]$

Def. The <u>conditional probability</u> of event *A* given the occurrence of event *B* is

$$\mathrm{P}[A|B] = \frac{\mathrm{P}[A \cap B]}{\mathrm{P}[B]}$$

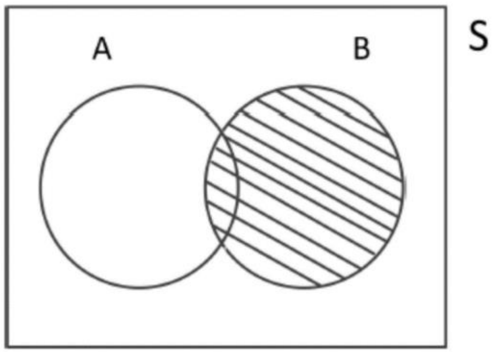
provided that $\mathrm{P}[B] \neq 0$ **.**

Example: In a die-rolling, the probability of getting 1 given that the result is odd is

$$\mathrm{P}[\{1\}|\{1,3,5\}] = \frac{\mathrm{P}[\{1\} \cap \{1,3,5\}]}{\mathrm{P}[\{1,3,5\}]} = \frac{1}{3}$$

To understand conditional probability:
Instead of S, working on B.
Normalized by the probability of B.

$$\mathrm{P}[A|B] = \frac{\mathrm{P}[A \cap B]}{\mathrm{P}[B]}$$



Results on conditional probability (from definition and axioms).

(a) $\mathrm{P}[A|B] \geq 0$

(b) $\mathrm{P}[A \cap B] = \mathrm{P}[A|B]\,\mathrm{P}[B]$

(c) *S* is the sample space.

$$\mathrm{P}[S|B] = \frac{\mathrm{P}[S \cap B]}{\mathrm{P}[B]} = 1.$$

(d) If $A_1, A_2, \ldots, A_m$ are mutually exclusive and $A = \cup_{i=1}^{m} A_i$,

$$\mathrm{P}[A|B] = \mathrm{P}[A_1|B] + \mathrm{P}[A_2|B] + \ldots + \mathrm{P}[A_m|B]$$

Theorems for regular probability work for conditional probability, e.g.,

$$\mathrm{P}[A_1 \cup A_2] = \mathrm{P}[A_1] + \mathrm{P}[A_2] - \mathrm{P}[A_1 \cap A_2]$$

$$\mathrm{P}[A_1 \cup A_2|B] = \mathrm{P}[A_1|B] + \mathrm{P}[A_2|B] - \mathrm{P}[A_1 \cap A_2|B]$$

Example. A dealer has 100 cars. There are only two car brands (L and A) and only three colors red (R), black (B), and white (W). The numbers of different cars are as in the table. Assume that a car is selected randomly All cars have the same likelihood of being chosen.

Find the probabilities of the follows.

	Red	Black	White
Brand L	10	25	5
Brand A	12	8	40

1) The selected car is black.
2) The selected car is Brand A and white.
3) You were told that Brand L is chosen. What is the probability it is red?
4) You were told that Brand A is chosen. What is the probability it is red?

Solution: Define Event L as cars being Brand L, Event A as cars being Brand A, Event R as cars being red, Event B as cars being black, Event W as cars being white.

	Red	Black	White	Total
Brand L	10	25	5	40
Brand A	12	8	40	60
Total	22	33	45	100

$$\mathrm{P}[B] = \frac{33}{100} = 0.33.$$

$$\mathrm{P}[A \cap W] = \frac{40}{100} = 0.4.$$

$$\mathrm{P}[R|L] = \frac{\mathrm{P}[R \cap L]}{\mathrm{P}[L]} = \frac{10/100}{40/100} = 0.25 \text{ or } \mathrm{P}[R|L] = \frac{10}{40} = 0.25.$$

$$\mathrm{P}[R|A] = \frac{\mathrm{P}[R \cap A]}{\mathrm{P}[A]} = \frac{12/100}{60/100} = 0.2 \text{ or } \mathrm{P}[R|A] = \frac{12}{60} = 0.2.$$

Example: A fair die is rolled twice.

(a) What is the probability that the sum is larger than 7?

Solution: Define $A = \{\text{sum} > 7\}$

$$\mathcal{S} = \{(1,1),(1,2),\ldots,(6,6)\} \rightarrow |\mathcal{S}| = 36$$

$$A = \{(2,6),(3,6),(3,5),(4,6),(4,5),(4,4),(5,3),(5,4),(5,5),(5,6),$$
$$(6,2),(6,3),(6,4),(6,5),(6,6)\} \rightarrow |A| = 15$$
$$\therefore \mathrm{P}[A] = \frac{15}{36}$$

(b) Given that the 1st roll is 3, what is the probability that the sum is larger than 7?

Solution:

$$A = \{\text{sum} > 7\}, \quad B = \{\text{first roll} = 3\}$$
$$\mathrm{P}[A|B] = \frac{\mathrm{P}[A \cap B]}{\mathrm{P}[B]} = \frac{\mathrm{P}[\{(3,5),(3,6)\}]}{\mathrm{P}[\{(3,1),(3,2),\cdots,(3,6)\}]} = \frac{2/36}{6/36} = 1/3.$$

(c) Given that at least one roll is less than 4, what is the probability that the sum is larger than 7?

Solution:

$$A = \{\text{sum} > 7\} \qquad C = \{\text{at least 1 roll} < 4\}$$

$$A \cap C = \{(2,6), (3,5), (3,6), (5,3), (6,3), (6,2)\}, |A \cap C| = 6$$

$$C^c = \{\text{both rolls} \geq 4\} = \{(x,y)|x,y = 4,5,6\}. \Rightarrow |C^c| = 9.$$

$$\mathrm{P}[C^c] = 9/36 \Rightarrow \mathrm{P}[C] = 1 - 9/36 = 27/36$$

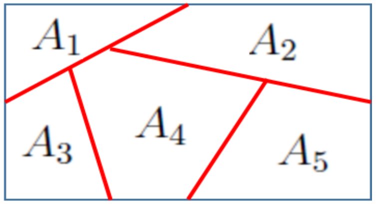
$$\mathrm{P}[A \cap C] = 6/36$$

$$\therefore \mathrm{P}[A|C] = \frac{\mathrm{P}[A \cap C]}{\mathrm{P}[C]} = \frac{6}{27} = \frac{2}{9}$$

1.5 The Law of Total Probability and Bayes' Theorem

Law of total probability expresses the probability of an event as the sum of probabilities of outcomes that are in the separate sets of a partition.

Review on partition: A set of mutually exclusive and collectively exhaustive sets/events.



Example: Communications of 10 bits between two devices.

A_1 = {No more than 4 bits are wrong.}

A_2 ={more than 4 bits are wrong.}

$\{A_1, A_2\}$ is a partition.

Law of Total Probability:

For a partition $\{B_1, B_2, \cdots, B_m\}$ with $\mathrm{P}[B_i] > 0$

$$\mathrm{P}[A] = \sum_{i=1}^{m} \mathrm{P}[A|B_i]\,\mathrm{P}[B_i]$$

for any event *A*.

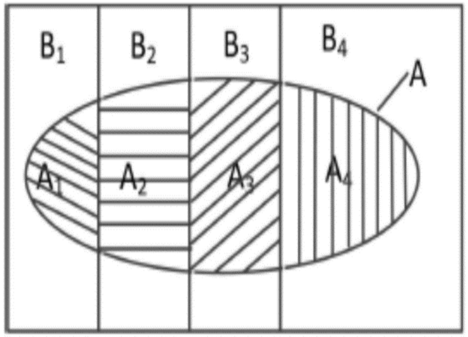
Proof. Recall the Theorem in Section 1.3. For $A_i = A \cap B_i$,

$$\mathrm{P}[A] = \sum_{i=1}^{m} \mathrm{P}[A_i] = \sum_{i=1}^{m} \mathrm{P}[A \cap B_i]$$

From the definition of conditional probability:

$$\mathrm{P}[A \cap B_i] = \mathrm{P}[A|B_i]\mathrm{P}[B_i]$$

Thus the law is proved.



Example: In a certain city, three car brands A,B, and C have 20%, 30%, and 50% of the market share, respectively. The probabilities that a car needs major repair during its 1st year of purchase for the three brands are 5%, 10%, 15%, respectively. What is the probability that a car needs repair during the 1st year of purchase?

Solution: Define

$$E = \{\text{Cars that need major repair in 1st year of purchase}\}$$

$$E_A = \{\text{Cars of brand A}\}$$

$$E_B = \{\text{Cars of brand B}\}$$

$$E_C = \{\text{Cars of brand C}\}$$

$\{E_A, E_B, E_C\}$ is a partition. From the law of total probability:

$$\begin{aligned}\mathrm{P}(E) &= \mathrm{P}(E|E_A)\mathrm{P}(E_A) + \mathrm{P}(E|E_B)\mathrm{P}(E_B) + \mathrm{P}(E|E_C)\mathrm{P}(E_C) \\ &= 5\% \times 20\% + 10\% \times 30\% + 15\% \times 50\% = 11.5\%.\end{aligned}$$

Example. A guy has a car and a bike. 80% of the time he goes to work by bike and 20% of the time he drives the car. If he drives the car, he is at work on time with probability 0.9. If he uses the bike, he is at work on time with probability 0.6. What is the probability that he is on time for one morning?

Solution: Define the following events:

C=He drives car. B=He rides bike. O=He is on time.

We have

$$\mathrm{P}(C) = 0.2,\ \mathrm{P}(B) = 0.8,\ \mathrm{P}(O|C) = 0.9,\ \mathrm{P}(O|B) = 0.6.$$

$\{C, B\}$ is a partition. From the law of total probability,

$$\mathrm{P}(O) = \mathrm{P}(O|B)\mathrm{P}(B) + \mathrm{P}(O|C)\mathrm{P}(C) = 0.6 \times 0.8 + 0.9 \times 0.2 = 0.66.$$

Bayes' theorem:

$$\mathrm{P}[B|A] = \frac{\mathrm{P}[A|B]\mathrm{P}[B]}{\mathrm{P}[A]}.$$

Proof. Directly from the definition of conditional probability.

If the conditional probability (likelihood, advance information) is known, to calculate the a-posteriori probability (backward information).

Let $\{B_1, B_2, \cdots, B_N\}$ *be a partition that include all possible state of interests but we cannot observe directly. If we know the priori probability* $\mathrm{P}[B_i]$ *and the likelihood of the observation* A *given that* B_i *is the actual state, we can calculate the probability of* B_i *based on the condition that* A *is observed (after* A *is observed):*

$$\mathrm{P}[B_i|A] = \frac{\mathrm{P}[A|B_i]\mathrm{P}[B_i]}{\mathrm{P}[A]}$$

Example. For the car example, If a car in this city needs major repair during its first year of purchase and you are required to make a guess on the brand of this car, what is your guess and why?

Solution: Using Bayes' theorem,

$$\mathrm{P}(E_A|E) = \frac{\mathrm{P}(E|E_A)\mathrm{P}(E_A)}{\mathrm{P}(E)} = \frac{5\% \times 20\%}{11.5\%} \approx 8.70\%.$$

$$\mathrm{P}(E_B|E) = \frac{\mathrm{P}(E|E_B)\mathrm{P}(E_B)}{\mathrm{P}(E)} = \frac{10\% \times 30\%}{11.5\%} \approx 26.09\%.$$

$$\mathrm{P}(E_C|E) = 1 - \mathrm{P}(E_A|E) - \mathrm{P}(E_B|E) \approx 65.21\%.$$

My guess is Brand C since it has the highest probability.

Example. For the going to work example, One particular morning, you are told that he is on time at work. What is the probability that he used the bike?

Solution: From Bayes' theorem,

$$\mathrm{P}(B|O) = \frac{\mathrm{P}(O|B)\mathrm{P}(B)}{\mathrm{P}(O)} = \frac{0.8 \times 0.6}{0.66} = 72.7\%.$$

Example. A data packet is transmitted on the Internet over three possible paths. The probabilities it gets dropped when using Path 1, 2, or 3 are 0.1, 0.05, and 0.12 respectively.

(a) If the path is chosen at random and equally likely, what is the probability of the packet being dropped?

$\{P_1, P_2, P_3\}$ is a partition.

$$\begin{aligned} \mathrm{P}[\text{drop}] &= \mathrm{P}[\text{drop}|P_1]\mathrm{P}[P_1] + \mathrm{P}[\text{drop}|P_2]\mathrm{P}[P_2] + \mathrm{P}[\text{drop}|P_3]\mathrm{P}[P_3] \\ &= 0.1 \times \frac{1}{3} + 0.05 \times \frac{1}{3} + 0.12 \times \frac{1}{3} = 0.09. \end{aligned}$$

(b) What if the probabilities of using the 3 paths are 2/7, 4/7, and 1/7 respectively.

$$\mathrm{P}[\text{drop}] = 0.1 \times \frac{2}{7} + 0.05 \times \frac{4}{7} + 0.12 \times \frac{1}{7} \approx 0.074.$$

(c) The probabilities of using the paths are 2/7, 4/7, and 1/7 respectively. A packet is dropped. What is the probability that Path 1 is selected?

$$\mathrm{P}[P_1|\text{drop}] = \frac{\mathrm{P}[P_1 \cap \text{drop}]}{\mathrm{P}[\text{drop}]} = \frac{\mathrm{P}[\text{drop} \cap P_1]}{\mathrm{P}[\text{drop}]} = \frac{\mathrm{P}[\text{drop}|P_1]\mathrm{P}[P_1]}{\mathrm{P}[\text{drop}]} = \frac{0.1 \times \frac{2}{7}}{0.074} = 0.386$$

Example. The probability that a person has a specific disease is 0.01. A test is proposed for the disease. A person with the disease is tested positive with probability 0.9; a person does not have the disease is tested positive with probability 0.05. Tom took the test and the result is positive. What is the probability that Tom has the disease?

Solution:

$$D = \text{a person has the disease}$$

$$N = \text{a person does not have the disease} = D^c.$$

$$P = \text{test is positive}$$

$\{D, N\}$ is a partition.

$$\begin{aligned}\mathrm{P}[D|P] &= \frac{\mathrm{P}[P|D]\mathrm{P}[D]}{\mathrm{P}[P]} = \frac{\mathrm{P}[P|D]\mathrm{P}[D]}{\mathrm{P}[P|D]\mathrm{P}[D] + \mathrm{P}[P|N]\mathrm{P}[N]} \\ &= \frac{0.9 \times 0.01}{0.9 \times 0.01 + 0.05 \times (1 - 0.01)} = \frac{0.009}{0.0585} \\ &= 0.154\end{aligned}$$

10,000 people take the test. About 585 are positive. But only about 90 of the 585 have the disease.
To increase the effectiveness of the test, need to decrease $\mathrm{P}[P|N]$.

Example. 3 cards each has 2 sides.

Card 1: Red-Red, Card 2: Yellow-yellow, Card 3: Red-yellow.

A card is drawn at random (equally likely) and placed on the table with a random top side. The top side is red; what is the probability of the other side is yellow, i.e, it is Card 3?

Solution: Define events C_1, C_2, C_3, R.

$$P[R] = P[R|C_1]\,P[C_1] + P[R|C_2]\,P[C_2] + P[R|C_3]\,P[C_3] = (1)(1/3) + 0(1/3) + (1/2)(1/3) = 1/2$$

$$P[C_3|R] = \frac{P[R|C_3]\,P[C_3]}{P[R]} = \frac{(1/2)(1/3)}{1/2} = 1/3$$

$$P[C_2|R] = 0$$

$$P[C_1|R] = 2/3$$

The probability that the other side is yellow is 1/3.

Card 2 (red-red card) is more likely since it has two red faces.

Example. A pond may have 1, 2, 3, 4, or 5 fish with equal probability. You catch a fish, tag it, and release it back to the pond. Then you catch a fish again and observe that the fish is not tagged. What are the respective probabilities that the pond has 1, 2, 3, 4, 5 fish after this observation?

Solution:

Define Events F_i to be the pond has *i* fish. $\{F_1, F_2, F_3, F_4, F_5\}$ is a partition.

Let A be the event that the fish (of the 2nd catch) is not tagged.

$$\mathrm{P}[F_i] = 1/5, \quad i = 1, 2, 3, 4, 5.$$

$$\mathrm{P}[A|F_1] = 0, \quad \mathrm{P}[A|F_2] = \frac{1}{2}, \quad \mathrm{P}[A|F_3] = \frac{2}{3}, \quad \mathrm{P}[A|F_4] = \frac{3}{4}, \quad \mathrm{P}[A|F_5] = \frac{4}{5}.$$

$$\Longleftrightarrow \mathrm{P}[A|F_i] = \frac{i-1}{i}.$$

$$\mathrm{P}[A] = \sum_{i=1}^{5} \mathrm{P}[A|F_i]\mathrm{P}[F_i] = 0 \times \frac{1}{5} + \frac{1}{2} \times \frac{1}{5} + \frac{2}{3} \times \frac{1}{5} + \frac{3}{4} \times \frac{1}{5} + \frac{4}{5} \times \frac{1}{5} \approx 0.54.$$

$$\mathrm{P}[F_i|A] = \frac{\mathrm{P}[A|F_i]\mathrm{P}[F_i]}{\mathrm{P}[A]} = \frac{\frac{i-1}{i}\frac{1}{5}}{0.54}$$

i	**1**	**2**	**3**	**4**	**5**
$\mathrm{P}[F_i\|A]$	0	0.18	0.24	0.28	0.29

1.6 Independence

Two events are independent if observing one does not change the probability of observing the other event.

Notation: $\mathrm{P}[A \cap B] = \mathrm{P}[AB] = \mathrm{P}[A \text{ and } B]$

Def. Event A **and** B **are independent if and only if**

$$\mathrm{P}[A \cap B] = \mathrm{P}[A]\mathrm{P}[B]$$

Equivalently, $\mathrm{P}[A|B] = \mathrm{P}[A]$ **or** $\mathrm{P}[B|A] = \mathrm{P}[B]$ **.**

Knowing that B occurs does not change our information about A.

$$\boxed{\underline{\text{Disjoint} \neq \text{independent}}}$$

Example. Roll a fair die. E is the event of `even' and $O$ is the event of `odd'. They are disjoint but not independent.

$$E = \{2, 4, 6\} \quad O = \{1, 3, 5\} \quad E \cap O = \emptyset$$

$$\mathrm{P}[E] = \mathrm{P}[O] = 1/2 \quad \mathrm{P}[E \cap O] = \mathrm{P}[\emptyset] = 0 \neq 1/4 = \mathrm{P}[E] \cdot \mathrm{P}[O]$$

B is the event of `larger than 3', $M$ is the event of `multiple of 3'.
They are not disjoint but independent.

$$B = \{4, 5, 6\} \quad M = \{3, 6\} \quad \mathrm{P}[B] = 1/2 \quad \mathrm{P}[M] = 1/3$$

$$\mathrm{P}[B \cap M] = \mathrm{P}[\{6\}] = 1/6 = \mathrm{P}[B] \cdot \mathrm{P}[M]$$

Example. In an experiment with events A, B C, D, we have

Event	A	B	C	D
Probability	1/4	1/8	5/8	3/8

A and B are disjoint. C and D are independent.

(a) Find $\mathrm{P}[A \cap B], \mathrm{P}[A \cup B], \mathrm{P}[C \cup D], \mathrm{P}[C^c \cap D^c]$

$$\mathrm{P}[A \cap B] = 0.$$

$$\mathrm{P}[A \cup B] = \mathrm{P}[A] + \mathrm{P}[B] - \mathrm{P}[A \cap B] = 3/8.$$

$$\mathrm{P}[C \cap D] = \mathrm{P}[C] \cdot \mathrm{P}[D] = 15/64.$$

$$\mathrm{P}[C \cup D] = \mathrm{P}[C] + \mathrm{P}[D] - \mathrm{P}[C \cap D] = 49/64.$$

$$\mathrm{P}[C^c \cap D^c] = 1 - \mathrm{P}[(C^c \cap D^c)^c] = 1 - \mathrm{P}[C \cup D] = 15/64$$

(b) Are C^c and D^c independent?

$$\mathrm{P}[C^c] = 1 - \mathrm{P}[C] = 3/8 \quad \mathrm{P}[D^c] = 1 - \mathrm{P}[D] = 5/8$$

$$\mathrm{P}[C^c \cap D^c] = \mathrm{P}[C^c] \cdot \mathrm{P}[D^c]. \text{ Thus, independent.}$$

Example. Show that when A and B are independent, then A^c and B are independent. Similarly, when A and B^c are independent, then A^c and B^c are independent.

Solution:

$$A \text{ and } B \text{ are independent } \Rightarrow \mathrm{P}[A \cap B] = \mathrm{P}[A] \cdot \mathrm{P}[B]$$

$$B = (A^c \cap B) \cup (A \cap B), \quad (A^c \cap B) \text{ and } (A \cap B) \text{ are disjoint}$$

$$\mathrm{P}[B] = \mathrm{P}[A^c \cap B] + \mathrm{P}[A \cap B]$$

$$\begin{aligned} \Rightarrow \mathrm{P}[A^c \cap B] &= \mathrm{P}[B] - \mathrm{P}[A \cap B] \\ &= \mathrm{P}[B] - \mathrm{P}[A] \cdot \mathrm{P}[B] \\ &= \mathrm{P}[B](1 - \mathrm{P}[A]) \\ &= \mathrm{P}[A^c] \cdot \mathrm{P}[B] \end{aligned}$$

Thus, A^c and B are independent.

If observing *A* does not change the probability of observing *B*, then observing that *A* does not happen also does not change the probability of observing *B*.

Def. 3 Events *A, B, C,* are <u>mutually independent</u> if and only if

1) $\mathrm{P}[A \cap B] = \mathrm{P}[A]\,\mathrm{P}[B]$

2) $\mathrm{P}[A \cap C] = \mathrm{P}[A]\,\mathrm{P}[C]$

3) $\mathrm{P}[B \cap C] = \mathrm{P}[B]\,\mathrm{P}[C]$

4) $\mathrm{P}[A \cap B \cap C] = \mathrm{P}[A]\,\mathrm{P}[B]\,\mathrm{P}[C]$

Any 2 are independent plus Condition 4.

Def. A set of n events $A_1, A_2, \cdots, A_n$ are <u>mutually independent</u> if and only if

(a) Any collection of n-1 events chosen from the set are mutually independent.

(b) $\mathrm{P}[A_1 \cap A_2 \cap \cdots \cap A_n] = \mathrm{P}[A_1]\mathrm{P}[A_2]\cdots\mathrm{P}[A_n].$

Example. An experiment has sample space and events as follows:

$$\mathcal{S} = \{1, 2, 3, 4\}, A = \{1, 2\}, B = \{1, 3\}, C = \{1, 4\}.$$

Assume that all outcomes are equally likely. Show that A, B, C are pairwise independent, but not mutually independent.

Solution:

$$\mathrm{P}[A] = 1/2, \mathrm{P}[B] = 1/2, \mathrm{P}[C] = 1/2$$

$$A \cap B = \{1\}, \mathrm{P}[A \cap B] = 1/4, \mathrm{P}[A]P[B] = 1/4$$

$$A \cap C = \{1\}, \mathrm{P}[A \cap C] = 1/4, \mathrm{P}[A]P[C] = 1/4$$

$$B \cap C = \{1\}, \mathrm{P}[B \cap C] = 1/4, \mathrm{P}[B]P[C] = 1/4$$

Thus they are pairwise independent.

$$A \cap B \cap C = \{1\}, \mathrm{P}[A \cap B \cap C] = 1/4 \neq \mathrm{P}[A]\,\mathrm{P}[B]\,\mathrm{P}[C] = 1/8$$

Thus, not mutually independent.

If both A and B occur, the outcome is 1, thus C must occur.

Example. Urn A has 4 red balls and 3 blue balls. Urn B has 8 red balls and x blue balls. An urn is selected randomly and a ball is drawn randomly from it. Find x to make the following two events independent.

E_1: Urn A is selected.

E_2: A blue ball is drawn.

Solution:

For E_1 and E_2 to be independent, need

$$\mathrm{P}[E_2|E_1] = \mathrm{P}[E_2], \mathrm{P}[E_2|E_1^c] = \mathrm{P}[E_2]$$

We have

$$\mathrm{P}[E_2|E_1] = 3/7, \ \mathrm{P}[E_2|E_1^C] = \frac{x}{x+8}$$

$$\frac{x}{x+8} = \frac{3}{7} \Rightarrow x = 6$$