COMPP9020 20T1
Week 9
Probability

- Textbook (R & W) - Ch. 5, Sec. 5.2; Ch. 9, Sec. 9.1
- Problem set week 9 + quiz

Assessment

Assessment isn't a "one-way street" ...

- I get to assess you in the final exam
- you get to assess me in UNSW's MyExperience Evaluation
  - go to `https://myexperience.unsw.edu.au/`
  - log on using zID@ad.unsw.edu.au and your zPass

  - or follow the direct link when on Moodle

Please fill it out ...

- give me some feedback on how you'd envision the course contents and delivery – lectures, problem sets, quizzes – in the future
- even if that is "Exactly the same. It was perfect this time."

Elementary Probability

Sample space:

$$\Omega = \{\omega_1, \ldots, \omega_n\}$$

Each point represents an outcome, each outcome ω_i equally likely:

$$P(\omega_1) = P(\omega_2) = \ldots = P(\omega_n) = \frac{1}{n}$$

This a called a **uniform probability distribution** over Ω

Examples

Tossing a coin: $\Omega = \{H, T\}$

$$P(H) = P(T) = 0.5$$

Rolling a die: $\Omega = \{1, 2, 3, 4, 5, 6\}$

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

Non-uniform Probability

Slight modification is needed to define an arbitrary (in general non-uniform) probability distribution:

$$\Omega = \{\omega_1, \ldots, \omega_n\}$$

Let

$$P(\omega_1) = p_1, P(\omega_2) = p_2, \ldots, P(\omega_n) = p_n$$

Then

$$\sum_{i=1}^{n} p_i = 1$$

Events

Definition

Event — a collection of outcomes = subset of Ω

Probability of an event:

$$P(E) = \sum_{\omega \in E} P(\omega)$$

Fact

$$P(\emptyset) = 0, \quad P(\Omega) = 1, \quad P(E^c) = 1 - P(E)$$

Exercise

$\boxed{5.2.7}$ Suppose an experiment leads to events A, B with probabilities $P(A) = 0.5, P(B) = 0.8, P(A \cap B) = 0.4$.
Find

- $P(B^c) = 1 - P(B) = 0.2$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.9$
- $P(A^c \cup B^c) = 1 - P((A^c \cup B^c)^c) = 1 - P(A \cap B) = 0.6$

Exercise

$\boxed{5.2.8}$ Given $P(A) = 0.6,\ \ P(B) = 0.7$, show $P(A \cap B) \geq 0.3$

Exercise

$\boxed{5.2.7}$ Suppose an experiment leads to events A, B with probabilities $P(A) = 0.5, P(B) = 0.8, P(A \cap B) = 0.4$.
Find

- $P(B^c) = 1 - P(B) = 0.2$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.9$
- $P(A^c \cup B^c) = 1 - P((A^c \cup B^c)^c) = 1 - P(A \cap B) = 0.6$

Exercise

$\boxed{5.2.8}$ Given $P(A) = 0.6,\ \ P(B) = 0.7$, show $P(A \cap B) \geq 0.3$

$$\begin{aligned} P(A \cap B) &= P(A) + P(B) - P(A \cup B) \\ &= 0.6 + 0.7 - P(A \cup B) \\ &\geq 0.6 + 0.7 - 1 = 0.3 \end{aligned}$$

Exercise

$\boxed{5.2.7}$ Suppose an experiment leads to events A, B with probabilities $P(A) = 0.5, P(B) = 0.8, P(A \cap B) = 0.4$.
Find

- $P(B^c) = 1 - P(B) = 0.2$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.9$
- $P(A^c \cup B^c) = 1 - P((A^c \cup B^c)^c) = 1 - P(A \cap B) = 0.6$

Exercise

$\boxed{5.2.8}$ Given $P(A) = 0.6,\ P(B) = 0.7$, show $P(A \cap B) \geq 0.3$

$$\begin{aligned} P(A \cap B) &= P(A) + P(B) - P(A \cup B) \\ &= 0.6 + 0.7 - P(A \cup B) \\ &\geq 0.6 + 0.7 - 1 = 0.3 \end{aligned}$$

Computing Probabilities by Counting

Computing probabilities with respect to a *uniform* distribution comes down to counting the size of the event.
If $E = \{e_1, \ldots, e_k\}$ then

$$P(E) = \sum_{i=1}^{k} P(e_i) = \sum_{i=1}^{k} \frac{1}{|\Omega|} = \frac{|E|}{|\Omega|}$$

Most of the counting rules carry over to probabilities wrt. a uniform distribution.

NB

The expression "selected at random", when not further qualified, means:
"subject to / according to / . . . a *uniform* distribution."

Example

$\boxed{\text{5.6.38 (supp)}}$ Of 100 problems, 75 are 'easy' and 40 'important'.
(b) n problems chosen randomly. What is the probability that all n are important?

$$p = \frac{\binom{40}{n}}{\binom{100}{n}} = \frac{40 \cdot 39 \cdots (41-n)}{100 \cdot 99 \cdots (101-n)}$$

Exercise

$\boxed{5.2.3}$ A 4-letter word is selected at random from Σ^4, where $\Sigma = \{a, b, c, d, e\}$. What is the probability that

(a) the letters in the word are all distinct?

(b) there are no vowels ("a", "e") in the word?

(c) the word begins with a vowel?

(a) $|E| = \Pi(5,4)$, $\quad P(E) = \frac{5 \cdot 4 \cdot 3 \cdot 2}{5^4} = \frac{120}{625} \approx 19\%$

(b) $|E| = 3^4$, $\quad P(E) = \frac{81}{625} \approx 13\%$

(c) $|E| = 2 \cdot 5^3$, $\quad P(E) = \frac{2}{5}$

Example

$\boxed{\text{5.6.38 (supp)}}$ Of 100 problems, 75 are 'easy' and 40 'important'.
(b) n problems chosen randomly. What is the probability that all n are important?

$$p = \frac{\binom{40}{n}}{\binom{100}{n}} = \frac{40 \cdot 39 \cdots (41-n)}{100 \cdot 99 \cdots (101-n)}$$

Exercise

$\boxed{5.2.3}$ A 4-letter word is selected at random from Σ^4, where $\Sigma = \{a, b, c, d, e\}$. What is the probability that
(a) the letters in the word are all distinct?
(b) there are no vowels ("a", "e") in the word?
(c) the word begins with a vowel?

(a) $|E| = \Pi(5,4), \quad P(E) = \frac{5 \cdot 4 \cdot 3 \cdot 2}{5^4} = \frac{120}{625} \approx 19\%$
(b) $|E| = 3^4, \quad P(E) = \frac{81}{625} \approx 13\%$
(c) $|E| = 2 \cdot 5^3, \quad P(E) = \frac{2}{5}$

Exercise

$\boxed{5.2.5}$ An urn contains 3 red and 4 black balls. 3 balls are removed without replacement. What are the probabilities that
(a) all 3 are red
(b) all 3 are black
(c) one is red, two are black

All probabilities are computed using the same sample space: all possible ways to draw three balls without replacement.

The size of the sample space is $\dfrac{7 \cdot 6 \cdot 5}{3!} = 35$

(a) $E =$ All balls are red: 1 combination
(b) $E =$ All balls are black: $\binom{4}{3} = 4$ combinations
(c) $E =$ One red and two black: $\binom{3}{1} \cdot \binom{4}{2} = 18$ combinations

Exercise

$\boxed{5.2.5}$ An urn contains 3 red and 4 black balls. 3 balls are removed without replacement. What are the probabilities that
(a) all 3 are red
(b) all 3 are black
(c) one is red, two are black

All probabilities are computed using the same sample space: all possible ways to draw three balls without replacement.
The size of the sample space is $\dfrac{7 \cdot 6 \cdot 5}{3!} = 35$
(a) $E =$ All balls are red: 1 combination
(b) $E =$ All balls are black: $\binom{4}{3} = 4$ combinations
(c) $E =$ One red and two black: $\binom{3}{1} \cdot \binom{4}{2} = 18$ combinations

Exercise

$\boxed{5.2.11}$ Two dice, a red die and a black die, are rolled.
What is the probability that

(a) the sum of the values is even?

$P(R+B \in \{2,4,\ldots,12\}) = \frac{18}{36} = \frac{1}{2}$

(b) the number on the red die is bigger than on the black die?

$P(R > B) = P(R < B)$; also $P(R = B) = \frac{1}{6}$

Therefore $P(R < B) = \frac{1}{2}(1 - P(R = B)) = \frac{5}{12}$

(c) the number on the black die is twice the one on the red die?

$P(R = 2 \cdot B) = P(\{(2,1),(4,2),(6,3)\}) = \frac{3}{36} = \frac{1}{12}$

$\boxed{5.2.12}$ (a) the maximum of the numbers is 4? $P(E_1) = \frac{7}{36}$

(b) their minimum is 4? $P(E_2) = \frac{5}{36}$

Check:

$P(E_1 \cup E_2) = \frac{7}{36} + \frac{5}{36} - P(E_1 \cap E_2) = \frac{7+5-1}{36} = \frac{11}{36}$

$P(\text{at least one '4'}) = 1 - P(\text{no '4'}) = 1 - \frac{5}{6} \cdot \frac{5}{6} = \frac{11}{36}$

Exercise

$\boxed{5.2.11}$ Two dice, a red die and a black die, are rolled.
What is the probability that
(a) the sum of the values is even?
$P\big(R + B \in \{2, 4, \ldots, 12\}\big) = \frac{18}{36} = \frac{1}{2}$

(b) the number on the red die is bigger than on the black die?
$P(R > B) = P(R < B)$; also $P(R = B) = \frac{1}{6}$
Therefore $P(R < B) = \frac{1}{2}(1 - P(R = B)) = \frac{5}{12}$

(c) the number on the black die is twice the one on the red die?
$P(R = 2 \cdot B) = P(\{(2,1), (4,2), (6,3)\}) = \frac{3}{36} = \frac{1}{12}$

$\boxed{5.2.12}$ (a) the maximum of the numbers is 4? $P(E_1) = \frac{7}{36}$
(b) their minimum is 4? $P(E_2) = \frac{5}{36}$

Check:

$P(E_1 \cup E_2) = \frac{7}{36} + \frac{5}{36} - P(E_1 \cap E_2) = \frac{7+5-1}{36} = \frac{11}{36}$

$P(\text{at least one '4'}) = 1 - P(\text{no '4'}) = 1 - \frac{5}{6} \cdot \frac{5}{6} = \frac{11}{36}$

Exercise

$\boxed{5.2.11}$ Two dice, a red die and a black die, are rolled.
What is the probability that

(a) the sum of the values is even?

$P(R + B \in \{2, 4, \dots, 12\}) = \frac{18}{36} = \frac{1}{2}$

(b) the number on the red die is bigger than on the black die?

$P(R > B) = P(R < B)$; also $P(R = B) = \frac{1}{6}$
Therefore $P(R < B) = \frac{1}{2}(1 - P(R = B)) = \frac{5}{12}$

(c) the number on the black die is twice the one on the red die?

$P(R = 2 \cdot B) = P(\{(2,1), (4,2), (6,3)\}) = \frac{3}{36} = \frac{1}{12}$

$\boxed{5.2.12}$ (a) the maximum of the numbers is 4? $P(E_1) = \frac{7}{36}$
(b) their minimum is 4? $P(E_2) = \frac{5}{36}$

Check:

$P(E_1 \cup E_2) = \frac{7}{36} + \frac{5}{36} - P(E_1 \cap E_2) = \frac{7+5-1}{36} = \frac{11}{36}$

$P(\text{at least one '4'}) = 1 - P(\text{no '4'}) = 1 - \frac{5}{6} \cdot \frac{5}{6} = \frac{11}{36}$

Asymptotic Estimate of Relative Probabilities

Example

Event $A \stackrel{\text{def}}{=}$ one die rolled n times and you obtain two 6's
Event $B \stackrel{\text{def}}{=}$ n dice rolled simultaneously and you obtain one 6

$$P(A) = \frac{\binom{n}{2} \cdot 5^{n-2}}{6^n} \qquad P(B) = \frac{\binom{n}{1} \cdot 5^{n-1}}{6^n}$$

Therefore $\frac{P(A)}{P(B)} = \frac{\binom{n}{2}}{\binom{n}{1}} \cdot \frac{1}{5} = \frac{n(n-1)}{2} \cdot \frac{1}{5n} = \frac{n-1}{10} \in \Theta(n)$

n	1	2	3	4	...	11	...	20	...
$P(A)$	0	$\frac{1}{36}$	$\frac{5}{72}$	$\frac{25}{216}$	...	0.296	...	0.198	...
$P(B)$	$\frac{1}{6}$	$\frac{10}{36}$	$\frac{25}{72}$	$\frac{125}{324}$	...	0.296	...	0.104	...

Inclusion-Exclusion and Probability

Inclusion-Exclusion is a very common method for deriving probabilities from other probabilities.

Two sets

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Three sets

$$\begin{aligned}
P(A \cup B \cup C) &= P(A \cup B) + P(C) - P((A \cup B) \cap C) \\
&= P(A) + P(B) - P(A \cap B) + P(C) \\
&\qquad - P((A \cap C) \cup (B \cap C)) \\
&= P(A) + P(B) - P(A \cap B) + P(C) \\
&\qquad - \big(P(A \cap C) + P(B \cap C) - P(A \cap C \cap B \cap C)\big) \\
&= P(A) + P(B) + P(C) \\
&\qquad - P(A \cap B) - P(A \cap C) - P(B \cap C) \\
&\qquad + P(A \cap B \cap C)
\end{aligned}$$

Example

A four-digit number n is selected at random (i.e. randomly from $[1000 \dots 9999]$). Find the probability p that n has each of 0, 1, 2 among its digits.

Let $q = 1 - p$ be the complementary probability and define

$$A_i = \{n : \text{no digit } i\}, A_{ij} = \{n : \text{no digits } i, j\}, A_{ijk} = \{n : \text{no } i, j, k\}$$

Then define

$T = A_0 \cup A_1 \cup A_2 = \{n : \text{ missing at least one of } 0, 1, 2\}$
$S = (A_0 \cup A_1 \cup A_2)^c = \{n : \text{ containing each of } 0, 1, 2\}$

Example (cont'd)

Once we find the cardinality of T, the solution is

$$q = \frac{|T|}{9000}, \; p = 1 - q$$

To find $|A_i|, |A_{ij}|, |A_{ijk}|$ we reflect on how many choices are available for the first digit, for the second etc. A special case is the leading digit, which must be $1, \ldots, 9$

Example (cont'd)

$$|A_0| = 9^4, \quad |A_1| = |A_2| = 8 \cdot 9^3$$
$$|A_{01}| = |A_{02}| = 8^4, \quad |A_{12}| = 7 \cdot 8^3$$
$$|A_{012}| = 7^4$$

$$\begin{aligned} |T| &= |A_0 \cup A_1 \cup A_2| \\ &= |A_0| + |A_1| + |A_2| - |A_0 \cap A_1| - |A_0 \cap A_2| - |A_1 \cap A_2| \\ &\qquad + |A_0 \cap A_1 \cap A_2| \\ &= 9^4 + 2 \cdot 8 \cdot 9^3 - 2 \cdot 8^4 - 7 \cdot 8^3 + 7^4 \\ &= 25 \cdot 9^3 - 23 \cdot 8^3 + 7^4 = 8850 \end{aligned}$$

$$q = \frac{8850}{9000}, \quad p = 1 - q \quad \approx 0.01667$$

Previous example generalised: Probability of an r-digit number having all of 0,1,2,3 among its digits.
We use the previous notation: A_i — set of numbers n *missing* digit i, and similarly for all $A_{ij\dots}$
We aim to find the size of $T = A_0 \cup A_1 \cup A_2 \cup A_3$, and then to compute $|S| = 9 \cdot 10^{r-1} - |T|$.

$$\begin{aligned} |A_0 \cup A_1 \cup A_2 \cup A_3| &= \text{sum of } |A_i| \\ &- \text{sum of } |A_i \cap A_j| \\ &+ \text{sum of } |A_i \cap A_j \cap A_k| \\ &- \text{sum of } |A_i \cap A_j \cap A_k \cap A_l| \end{aligned}$$

Probability of Sequential Outcomes

Example

Team A has probability $p = 0.5$ of winning a game against B.
What is the probability P_p of A winning a best-of-seven match if
(a) A already won the first game?
(b) A already won the first two games?
(c) A already won two out of the first three games?

(a) Sample space S — 6-sequences, formed from wins (W) and losses (L)

$$|S| = 2^6 = 64$$

Favourable sequences F — those with three to six W

$$|F| = \binom{6}{3} + \binom{6}{4} + \binom{6}{5} + \binom{6}{6} = 20 + 15 + 6 + 1 = 42$$

Therefore $P_{0.5} = \frac{42}{64} \approx 66\%$

Probability of Sequential Outcomes

Example

Team A has probability $p = 0.5$ of winning a game against B.
What is the probability P_p of A winning a best-of-seven match if
(a) A already won the first game?
(b) A already won the first two games?
(c) A already won two out of the first three games?

(a) Sample space S — 6-sequences, formed from wins (W) and losses (L)

$$|S| = 2^6 = 64$$

Favourable sequences F — those with three to six W

$$|F| = \binom{6}{3} + \binom{6}{4} + \binom{6}{5} + \binom{6}{6} = 20 + 15 + 6 + 1 = 42$$

Therefore $P_{0.5} = \frac{42}{64} \approx 66\%$

Exercise

(b) Sample space S? — 5-sequences of W and L

$$|S| = 2^5 = 32$$

Favourable sequences F? — those with two to five W

$$|F| = \binom{5}{2} + \binom{5}{3} + \binom{5}{4} + \binom{5}{5} = 10 + 10 + 5 + 1 = 26$$

Therefore $P_{0.5} = \frac{26}{32} \approx 81\%$

(c)

$$|S| = 2^4 = 16$$

$$|F| = \binom{4}{2} + \binom{4}{3} + \binom{4}{4} = 6 + 4 + 1 = 11$$

Therefore $P_{0.5} = \frac{11}{16} \approx 69\%$

Exercise

(b) Sample space S? — 5-sequences of W and L

$$|S| = 2^5 = 32$$

Favourable sequences F? — those with two to five W

$$|F| = \binom{5}{2} + \binom{5}{3} + \binom{5}{4} + \binom{5}{5} = 10 + 10 + 5 + 1 = 26$$

Therefore $P_{0.5} = \frac{26}{32} \approx 81\%$

(c)

$$|S| = 2^4 = 16$$

$$|F| = \binom{4}{2} + \binom{4}{3} + \binom{4}{4} = 6 + 4 + 1 = 11$$

Therefore $P_{0.5} = \frac{11}{16} \approx 69\%$

Exercise

(b) Sample space S? — 5-sequences of W and L

$$|S| = 2^5 = 32$$

Favourable sequences F? — those with two to five W

$$|F| = \binom{5}{2} + \binom{5}{3} + \binom{5}{4} + \binom{5}{5} = 10 + 10 + 5 + 1 = 26$$

Therefore $P_{0.5} = \frac{26}{32} \approx 81\%$

(c)

$$|S| = 2^4 = 16$$

$$|F| = \binom{4}{2} + \binom{4}{3} + \binom{4}{4} = 6 + 4 + 1 = 11$$

Therefore $P_{0.5} = \frac{11}{16} \approx 69\%$

Example (cont'd)

Redo for arbitrary p

(a)

$$P_p = \binom{6}{3} p^3(1-p)^3 + \binom{6}{4} p^4(1-p)^2 + \binom{6}{5} p^5(1-p) + \binom{6}{6} p^6$$

(b)

$$P_p = \binom{5}{2} p^2(1-p)^3 + \binom{5}{3} p^3(1-p)^2 + \binom{5}{4} p^5(1-p) + \binom{5}{5} p^5$$

(c)

$$P_p = \binom{4}{2} p^2(1-p)^2 + \binom{4}{3} p^3(1-p) + \binom{4}{4} p^4$$

Use of Recursion in Probability Computations

Question

Given n tosses of a coin, what is the probability of two HEADS *in a row? Compute for $n = 5, 10, 20, \ldots$*

Approaches:

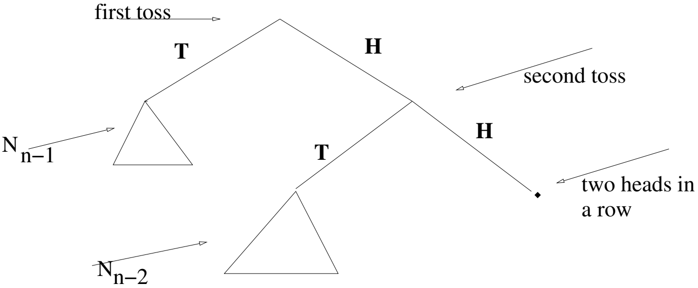
I. Write down all possibilities — 32 for $n = 5$, 1024 for $n = 10$, . . .

II. Write a program; running time $\mathcal{O}(2^n)$ — why?

III. Inter-relate the numbers of relevant possibilities
$N_n \stackrel{\text{def}}{=}$ No. of sequences of n tosses *without* . . . HH. . . pattern
Initial values:
$N_0 = 1,\ N_1 = 2,\ N_2 = 3$ (all except "HH")
$N_3 = 5$ (why?) $\quad N_4 = 8$ (why?)

Answer

We can summarise all possible outcomes in a **recursive tree**



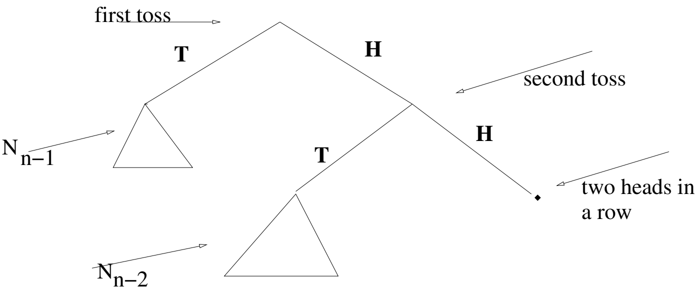
$N_n = N_{n-1} + N_{n-2}$ — Fibonacci recurrence: $N_n = \text{FIB}(n+2)$

$N_n \approx \frac{1}{\sqrt{5}} \left(\frac{\sqrt{5}+1}{2} \right)^{n+2} \approx 0.72 \cdot (1.6)^{n+1}$

$p_n = \frac{2^n - \text{FIB}(n+2)}{2^n} \approx 1 - 0.72 \cdot (0.8)^{n+1}$

Answer

We can summarise all possible outcomes in a **recursive tree**



$$N_n = N_{n-1} + N_{n-2} \text{ — Fibonacci recurrence: } N_n = \textsc{fib}(n+2)$$

$$N_n \approx \frac{1}{\sqrt{5}} \left(\frac{\sqrt{5}+1}{2}\right)^{n+2} \approx 0.72 \cdot (1.6)^{n+1}$$

$$p_n = \frac{2^n - \textsc{fib}(n+2)}{2^n} \approx 1 - 0.72 \cdot (0.8)^{n+1}$$

Example

Question

Given n tosses, what is the probability q_n of at least one HHH?

$q_0 = q_1 = q_2 = 0; q_3 = \frac{1}{8}$

Then recursive computation:

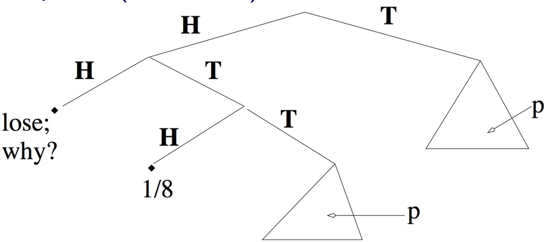
$$\begin{aligned} q_n =& \frac{1}{2} q_{n-1} && \text{(initial: T)} \\ &+ \frac{1}{4} q_{n-2} && \text{(initial: HT)} \\ &+ \frac{1}{8} q_{n-3} && \text{(initial: HHT)} \\ &+ \frac{1}{8} && \text{(start with: HHH)} \end{aligned}$$

Example

Question

A coin is tossed 'indefinitely'. Which pattern is more likely (and by how much) to appear first, HTH or HHT?

let $p = P(HTH \text{ first})$



$p = \frac{1}{8} + \frac{1}{8}p + \frac{1}{2}p \Rightarrow \frac{3}{8}p = \frac{1}{8} \Rightarrow p = \frac{1}{3}$

NB

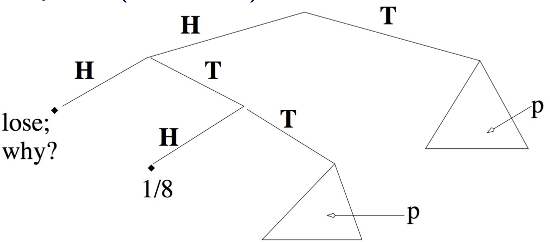
Probability that either pattern would appear at a given, *prespecified* point in the sequence of tosses is, obviously, the same.

Example

Question

A coin is tossed 'indefinitely'. Which pattern is more likely (and by how much) to appear first, HTH or HHT?

let $p = P(HTH \text{ first})$



$$p = \frac{1}{8} + \frac{1}{8}p + \frac{1}{2}p \;\Rightarrow\; \frac{3}{8}p = \frac{1}{8} \;\Rightarrow\; p = \frac{1}{3}$$

NB

Probability that either pattern would appear at a given, *prespecified* point in the sequence of tosses is, obviously, the same.

Example

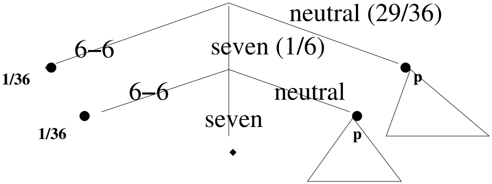
Question

Two dice are rolled repeatedly. What is the probability that '6–6' will occur before two consecutive (back-to-back) 'totals seven'?

NB

The probability of either occurring at a given roll is the same: $\frac{1}{36}$.

Let $p = P(6\text{–}6 \text{ first})$



$$p = \frac{1}{36} + \frac{1}{6}\cdot\frac{1}{36} + \frac{1}{6}\cdot\frac{29}{36}p + \frac{29}{36}p \;\Rightarrow\; 216p = 7 + 203p \;\Rightarrow\; p = \frac{7}{13}$$

Example

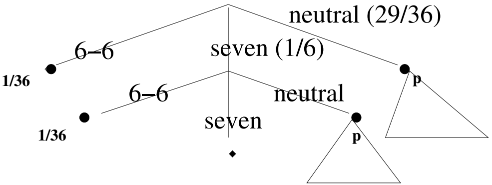
Question

Two dice are rolled repeatedly. What is the probability that '6–6' will occur before two consecutive (back-to-back) 'totals seven'?

NB

The probability of either occurring at a given roll is the same: $\frac{1}{36}$.

Let $p = P(\text{6–6 first})$



$$p = \frac{1}{36} + \frac{1}{6}\cdot\frac{1}{36} + \frac{1}{6}\cdot\frac{29}{36}p + \frac{29}{36}p \;\Rightarrow\; 216p = 7 + 203p \;\Rightarrow\; p = \frac{7}{13}$$

NB

The majority of problems in probability and statistics *do not have* such elegant solutions. Hence the use of computers for either precise calculations or approximate simulations is mandatory. However, it is the use of recursion that simplifies such computing or, quite often, makes it possible in the first place.

Conditional Probability

Definition

Conditional probability of E **given** S:

$$P(E|S) = \frac{P(E \cap S)}{P(S)}, \quad E, S \subseteq \Omega$$

It is defined only when $P(S) \neq 0$

NB

$P(A|B)$ and $P(B|A)$ are, in general, not related — one of these values predicts, by itself, essentially nothing about the other. The only exception, applicable when $P(A), P(B) \neq 0$, is that $P(A|B) = 0$ iff $P(B|A) = 0$ iff $P(A \cap B) = 0$.

Conditional Probability

Definition

Conditional probability of E **given** S:

$$P(E|S) = \frac{P(E \cap S)}{P(S)}, \quad E, S \subseteq \Omega$$

It is defined only when $P(S) \neq 0$

NB

$P(A|B)$ and $P(B|A)$ are, in general, not related — one of these values predicts, by itself, essentially nothing about the other. The only exception, applicable when $P(A), P(B) \neq 0$, is that $P(A|B) = 0$ iff $P(B|A) = 0$ iff $P(A \cap B) = 0$.

If P is the uniform distribution over a finite set Ω, then

$$P(E|S) = \frac{\frac{|E \cap S|}{|\Omega|}}{\frac{|S|}{|\Omega|}} = \frac{|E \cap S|}{|S|}$$

This observation can help in calculations...

Example

$\boxed{9.1.6}$ A coin is tossed four times. What is the probability of
(a) two consecutive HEADS
(b) two consecutive HEADS *given* that ≥ 2 tosses are HEADS

T	T	T	T	H	T	T	T
T	T	T	H	H	T	T	H
T	T	H	T	H	T	H	T
T	T	H	H	H	T	H	H
T	H	T	T	H	H	T	T
T	H	T	H	H	H	T	H
T	H	H	T	H	H	H	T
T	H	H	H	H	H	H	H

(a) $\frac{8}{16}$ (b) $\frac{8}{11}$

Exercise

$\boxed{9.1.12}$ What is the probability of a flush given that all five cards in a Poker hand are red?

Red cards = $\diamondsuit$'s + $\heartsuit$'s
flush = all cards of the same suit, not of sequential rank

$$P(\text{flush} \mid \text{all five cards are Red}) = \frac{2 \cdot \left(\binom{13}{5} - 10\right)}{\binom{26}{5}} \approx 4\%$$

Exercise

$\boxed{9.1.12}$ What is the probability of a flush given that all five cards in a Poker hand are red?

Red cards $= \diamondsuit$'s $+ \heartsuit$'s
flush $=$ all cards of the same suit, not of sequential rank

$$P(\text{flush} \,|\, \text{all five cards are Red}) = \frac{2 \cdot (\binom{13}{5} - 10)}{\binom{26}{5}} \approx 4\%$$

Some General Rules

Fact

- $A \subseteq B \Rightarrow P(A|B) \geq P(A)$
- $A \subseteq B \Rightarrow P(B|A) = 1$
- $P(A \cap B|B) = P(A|B)$
- $P(\emptyset|A) = 0 \textit{ for } A \neq \emptyset$
- $P(A|\Omega) = P(A)$
- $P(A^c|B) = 1 - P(A|B)$

NB

- $P(A|B)$ and $P(A|B^c)$ are not related
- $P(A|B), P(B|A), P(A^c|B^c), P(B^c|A^c)$ are not related

Example

Two dice are rolled and the outcomes recorded as b for the black die, r for the red die and $s = b + r$ for their total.
Define the events $B = \{b \geq 3\}$, $R = \{r \geq 3\}$, $S = \{s \geq 6\}$.

$$P(S|B) = \frac{4+5+6+6}{24} = \frac{21}{24} = \frac{7}{8} = 87.5\%$$

$$P(B|S) = \frac{4+5+6+6}{26} = \frac{21}{26} = 80.8\%$$

The (common) numerator $4 + 5 + 6 + 6 = 21$ represents the size of the $B \cap S$ — the common part of B and S, that is, the number of rolls where $b \geq 3$ and $s \geq 6$. It is obtained by considering the different cases: $b = 3$ and $s \geq 6$, then $b = 4$ and $s \geq 6$ etc.

The denominators are $|B| = 24$ and $|S| = 26$

NB

Bayes' Formula: $P(S|B) \cdot P(B) = P(B|S) \cdot P(S)$

Example (cont'd)

Recall: $B = \{b \geq 3\}, \ R = \{r \geq 3\}, \ S = \{s \geq 6\}$

$P(B) = P(R) = 2/3 = 66.7\%$

$P(S) = \frac{5+6+5+4+3+2+1}{36} = \frac{26}{36} = 72.22\%$

$P(S|B \cup R) = \frac{2+3+4+5+6+6}{32} = \frac{26}{32} = 81.25\%$

The set $B \cup R$ represents the event 'b or r'.
It comprises all the rolls except for those with *both* the red and the black die coming up either 1 or 2.

$P(S|B \cap R) = 1 = 100\%$ — because $S \supseteq B \cap R$

Exercise

$\boxed{9.1.9}$ Consider three red and eight black marbles; draw two without replacement. We write b_1 — Black on the first draw, b_2 — Black on the second draw, r_1 — Red on first draw, r_2 — Red on second draw

Using conditional probabilities, find the probabilities

(a) both Red:

$$P(r_1 \wedge r_2) = P(r_1)P(r_2|r_1) = \frac{3}{11} \cdot \frac{2}{10} = \frac{3}{55}$$

Equivalently:

$|\text{two-samples}| = \binom{11}{2} = 55$; $|\text{Red two-samples}| = \binom{3}{2} = 3$

$P(\cdot) = \frac{\binom{3}{2}}{\binom{11}{2}} = \frac{3}{55}$

(b) both Black:

$$P(b_1 \wedge b_2) = P(b_1)P(b_2|b_1) = \frac{8}{11} \cdot \frac{7}{10} = \frac{28}{55} \quad \left(= \frac{\binom{8}{2}}{\binom{11}{2}}\right)$$

Exercise

$\boxed{9.1.9}$ Consider three red and eight black marbles; draw two without replacement. We write b_1 — Black on the first draw, b_2 — Black on the second draw, r_1 — Red on first draw, r_2 — Red on second draw

Using conditional probabilities, find the probabilities

(a) both Red:

$$P(r_1 \wedge r_2) = P(r_1)P(r_2|r_1) = \frac{3}{11} \cdot \frac{2}{10} = \frac{3}{55}$$

Equivalently:

$|\text{two-samples}| = \binom{11}{2} = 55$; $|\text{Red two-samples}| = \binom{3}{2} = 3$

$P(\cdot) = \frac{\binom{3}{2}}{\binom{11}{2}} = \frac{3}{55}$

(b) both Black:

$$P(b_1 \wedge b_2) = P(b_1)P(b_2|b_1) = \frac{8}{11} \cdot \frac{7}{10} = \frac{28}{55} \quad \left(= \frac{\binom{8}{2}}{\binom{11}{2}}\right)$$

Exercise

(c) one Red, one Black:

$$P(r_1 \wedge b_2) + P(b_1 \wedge r_2) = \frac{3 \cdot 8}{\binom{11}{2}} \quad \text{— why?}$$

By textbook (the 'hard way')

$$P(r_1 \wedge b_2) + P(b_1 \wedge r_2) = \frac{3}{11} \cdot \frac{8}{10} + \frac{8}{11} \cdot \frac{3}{10}$$

or

$$P(\cdot) = 1 - P(r_1 \wedge r_2) - P(b_1 \wedge b_2) = \frac{55 - 3 - 28}{55}$$

Exercise

(c) one Red, one Black:

$$P(r_1 \wedge b_2) + P(b_1 \wedge r_2) = \frac{3 \cdot 8}{\binom{11}{2}} \quad \text{— why?}$$

By textbook (the 'hard way')

$$P(r_1 \wedge b_2) + P(b_1 \wedge r_2) = \frac{3}{11} \cdot \frac{8}{10} + \frac{8}{11} \cdot \frac{3}{10}$$

or

$$P(\cdot) = 1 - P(r_1 \wedge r_2) - P(b_1 \wedge b_2) = \frac{55 - 3 - 28}{55}$$

Exercise

$\boxed{9.1.22}$ Prove the following:

If $P(A|B) > P(A)$ ("positive correlation") then $P(B|A) > P(B)$

$P(A|B) > P(A)$

$\Rightarrow P(A \cap B) > P(A) \cdot P(B)$

$\Rightarrow \frac{P(A \cap B)}{P(A)} > P(B)$

$\Rightarrow P(B|A) > P(B)$

Exercise

$\boxed{9.1.22}$ Prove the following:
If $P(A|B) > P(A)$ ("positive correlation") then $P(B|A) > P(B)$

$P(A|B) > P(A)$

$\Rightarrow P(A \cap B) > P(A) \cdot P(B)$

$\Rightarrow \frac{P(A \cap B)}{P(A)} > P(B)$

$\Rightarrow P(B|A) > P(B)$

Stochastic Independence

Definition

A and B are **stochastically independent** (notation: $A \perp B$) if $P(A \cap B) = P(A) \cdot P(B)$

If $P(A) \neq 0$ and $P(B) \neq 0$, all of the following are *equivalent*:

- $P(A \cap B) = P(A)P(B)$
- $P(A|B) = P(A)$ (i.e. B does not affect the probability of A)
- $P(B|A) = P(B)$ (i.e. A does not affect the probability of B)
- $P(A^c|B) = P(A^c)$ or $P(A|B^c) = P(A)$ or $P(A^c|B^c) = P(A^c)$

The last one means that

$$A \perp B \Leftrightarrow A^c \perp B \Leftrightarrow A \perp B^c \Leftrightarrow A^c \perp B^c$$

Basic non-independent sets of events (if $P(A), P(B) > 0$)

- $A \subseteq B$
- $A \cap B = \emptyset$
- Any pair of one-point events $\{x\}, \{y\}$:
  either $x = y$ and $P(x|y) = 1$
  or $x \neq y$ and $P(x|y) = 0$

Independence of $A_1, \ldots, A_n$

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdots P(A_{i_k})$$

for all possible collections $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$.
This is often called (for emphasis) a *full* independence

Basic non-independent sets of events (if $P(A), P(B) > 0$)

- $A \subseteq B$
- $A \cap B = \emptyset$
- Any pair of one-point events $\{x\}, \{y\}$:
  either $x = y$ and $P(x|y) = 1$
  or $x \neq y$ and $P(x|y) = 0$

Independence of $A_1, \ldots, A_n$

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdots P(A_{i_k})$$

for all possible collections $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$.
This is often called (for emphasis) a *full* independence

Pairwise independence is a *weaker* concept.

Example

Toss of two coins

$$\left.\begin{array}{l} A = \langle \text{first coin } H \rangle \\ B = \langle \text{second coin } H \rangle \\ C = \langle \text{exactly one } H \rangle \end{array}\right\} \begin{array}{l} P(A) = P(B) = P(C) = \frac{1}{2} \\ P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4} \\ \text{However: } P(A \cap B \cap C) = 0 \end{array}$$

One can similarly construct a set of n events where any k of them are independent, while any $k+1$ are dependent (for $k < n$).

NB

Independence of events, even just pairwise independence, can greatly simplify computations and reasoning in AI applications. It is common for many expert systems to make an approximating assumption of independence, even if it is not completely satisfied.

$$P(\text{sense}_t \mid \text{loc}_t, \text{sense}_{t-1}, \text{loc}_{t-1}, \ldots) = P(\text{sense}_t \mid \text{loc}_t)$$

Exercise

$\boxed{9.1.7}$ Suppose that an experiment leads to events A, B and C with $P(A) = 0.3$, $P(B) = 0.4$ and $P(A \cap B) = 0.1$

(a) $P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{1}{4}$

(b) $P(A^c) = 1 - P(A) = 0.7$

(c) Is $A \perp B$? No. $P(A) \cdot P(B) = 0.12 \neq P(A \cap B)$

(d) Is $A^c \perp B$? No, as can be seen from (c).

Note: $P(A^c \cap B) = P(B) - P(A \cap B) = 0.4 - 0.1 = 0.3$
$P(A^c) \cdot P(B) = 0.7 \cdot 0.4 = 0.28$

Exercise

$\boxed{9.1.7}$ Suppose that an experiment leads to events A, B and C with $P(A) = 0.3$, $P(B) = 0.4$ and $P(A \cap B) = 0.1$

(a) $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{4}$

(b) $P(A^c) = 1 - P(A) = 0.7$

(c) Is $A \perp B$? No. $P(A) \cdot P(B) = 0.12 \neq P(A \cap B)$

(d) Is $A^c \perp B$? No, as can be seen from (c).

Note: $P(A^c \cap B) = P(B) - P(A \cap B) = 0.4 - 0.1 = 0.3$
$P(A^c) \cdot P(B) = 0.7 \cdot 0.4 = 0.28$

Exercise

$\boxed{9.1.8}$ Given $A \perp B, \;\; P(A) = 0.4, \;\; P(B) = 0.6$

$P(A|B) = P(A) = 0.4$

$P(A \cup B) = P(A) + P(B) - P(A)P(B) = 0.76$

$P(A^c \cap B) = P(A^c)P(B) = 0.36$

Exercise

$\boxed{9.1.8}$ Given $A \perp B,\ \ P(A) = 0.4,\ \ P(B) = 0.6$

$P(A|B) = P(A) = 0.4$

$P(A \cup B) = P(A) + P(B) - P(A)P(B) = 0.76$

$P(A^c \cap B) = P(A^c)P(B) = 0.36$

Exercise

$\boxed{\text{9.5.5 (supp)}}$ We are given two events with $P(A) = \frac{1}{4},\ P(B) = \frac{1}{3}$.
True, false or could be either?

(a) $P(A \cap B) = \frac{1}{12}$ — possible; it holds when $A \perp B$

(b) $P(A \cup B) = \frac{7}{12}$ — possible; it holds when A, B are disjoint

(c) $P(B|A) = \frac{P(B)}{P(A)}$ — false; correct is: $P(B|A) = \frac{P(B \cap A)}{P(A)}$

(d) $P(A|B) \geq P(A)$ — possible (it means that B "supports" A)

(e) $P(A^c) = \frac{3}{4}$ — true, since $P(A^c) = 1 - P(A)$

(f) $P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$ — true

NB

Total probability: $P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$

Exercise

$\boxed{\text{9.5.5 (supp)}}$ We are given two events with $P(A) = \frac{1}{4},\ P(B) = \frac{1}{3}$.
True, false or could be either?

(a) $P(A \cap B) = \frac{1}{12}$ — possible; it holds when $A \perp B$

(b) $P(A \cup B) = \frac{7}{12}$ — possible; it holds when A, B are disjoint

(c) $P(B|A) = \frac{P(B)}{P(A)}$ — false; correct is: $P(B|A) = \frac{P(B \cap A)}{P(A)}$

(d) $P(A|B) \geq P(A)$ — possible (it means that B "supports" A)

(e) $P(A^c) = \frac{3}{4}$ — true, since $P(A^c) = 1 - P(A)$

(f) $P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$ — true

NB

Total probability: $P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$

Summary

- Sample space, probability distribution, events
- Inclusion-exclusion and probability
- Probability of sequential outcomes, recursion for probabilities
- Conditional probability $P(A|B)$, independence $A \perp B$
- Bayes' formula, total probability

Coming up . . .

- Ch. 9, Sec. 9.2-9.4 (Expectation)
- Course review