1 Definitions

1. Conditional Probability: $P(A \mid B) =$

   $P(A \mid B) = \frac{P(A,B)}{P(B)}$

2. Product Rule: $P(A, B) =$

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

3. Bayes' Theorem: $P(A \mid B) =$

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

4. Normalization: $P(A \mid B) =$

   $P(A \mid B) = \frac{P(A,B)}{P(B)} = \frac{P(A,B)}{\sum_a P(a,B)}$

5. Chain Rule: $P(A, B, C) =$

   $P(A, B, C) = P(A \mid B, C)P(B \mid C)P(C)$

6. Law of Total Probability: [using only $P(B)$ and $P(A \mid B)$] $P(A) =$

   $P(A) = \sum_{b \in B} P(A \mid b)P(b)$

   For binary B:

   $P(A) = P(A \mid b_1)P(b_1) + P(A \mid b_2)P(b_2)$

7. Independence: A, B independent, $P(A, B) =$

   If A and B are independent, then $P(A, B) = P(A)P(B)$

8. Conditional Independence: If A and B are conditionally independent given C, then $P(A, B \mid C) =$

   If A and B are conditionally independent given C, then $P(A, B \mid C) = P(A \mid C)P(B \mid C)$

2 Warm Up

(a) State the two ways to write the chain rule (conditional probability decomposition) for $P(A, B)$

   $P(A)P(B \mid A) = P(B)P(A \mid B)$

(b) Rearrange the above equation to find $P(A \mid B)$

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

(c) Find $P(a)$ in terms of the joint $P(a, b)$ for any $a \in A, b \in B$ (Hint: these are specific values, answer should include a sum)

$P(a) = \sum_{b \in B} P(a, b)$

(d) Find $P(b \mid a)$ in terms of the joint $P(a, b)$ for any $a \in A, b \in B$

$P(b \mid a) = \frac{P(a,b)}{\sum_{b' \in B} P(a,b')}$

(e) Find $P(b \mid a)$ in terms of the distributions $P(b)$, $P(a \mid b)$, for any $a \in A, b \in B$

$P(b \mid a) = \frac{P(a|b)P(b)}{\sum_{b' \in B} P(a|b')P(b')}$

(f) Assume we had some fixed a and wanted to find each element of $P(b \mid a)$ (i.e. wanted to find $P(B \mid a)$). Would the numerator of the fraction in the previous question change for each value of b? What about the denominator? How could you use this to do the calculation with less steps?

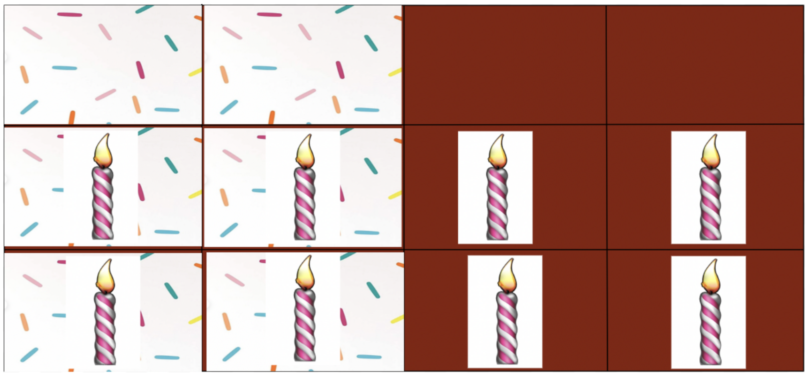
The numerator changes because the value of b changes. The denominator is constant because $P(a)$ will be the same for every value of b that we change. We can calculate all the numerators first, then normalize/equivalently compute the denominator at the end.

(g) Assume A is a random variable that can take 3 values, B is a random variable that can take 2 values, and C is a random variable that can take 1 value. What do the following probability tables sum to?

(a) $P(A \mid b)$
(b) $P(A \mid C)$
(c) $P(C \mid B)$
(d) $P(B \mid a)$
(e) $P(B \mid A)$

$P(A \mid b), P(A \mid C)$, and $P(B \mid a)$ sum to 1. $P(C \mid B)$ sums to 2 because B can take 2 values (b_1 and b_2). $P(C \mid b_1)$ and $P(C \mid b_2)$ each sum to 1, so if we add them, we get 2. $P(B \mid A)$ sums to 3 because A can take 3 values (a_1, a_2, and a_3). Each of $P(B \mid a_1)$, $P(B \mid a_2)$, and $P(B \mid a_3)$ sum to 1, so the total sums to 3.

3 Cake



Consider the above cake with 12 slices. Let s_1 indicate a slice with no sprinkles and s_2 be a slice with sprinkles. Let c_1 indicate a slice with no candles and c_2 be a slice with candles. Let S be a random variable indicating sprinkles and C be a random variable indicating candles. Calculate the following probabilities.

1. $P(C = c_1)$

   By counting the number of slices that don't have candles, we can see that the probability of getting a slice with no candles is 4/12.

2. $P(S = s_1, C = c_2)$

   By counting the number of slices that don't have sprinkles but have candles, we can see that the probability of getting a slice with candles and no sprinkles is 4/12.

3. $P(C = c_2 \mid S = s_1)$

   We first constrain our world to only include the slices that contain no sprinkles which is 6 slices. 4 of those slices contain candles, so this probability becomes 4/6. Another way to calculate this probability is by using the definition of conditional probability, $P(C = c_2 \mid S = s_1) = \frac{P(C=c_2,S=s_1)}{P(S=s_1)} = \frac{4/12}{6/12} = \frac{4}{6}$

4. $\sum_{s\in\{s_1,s_2\}} \sum_{c\in\{c_1,c_2\}} P(s, c)$

   $P(s_1, c_1) + P(s_1, c_2) + P(s_2, c_1) + P(s_2, c_2) = 1$. Because we are summing up all possible disjoint combinations of the given sample space, the answer is 1.

5. $\sum_{c\in\{c_1,c_2\}} \sum_{s\in\{s_1,s_2\}} P(s \mid c)$

   $P(s_1 \mid c_1)+P(s_2 \mid c_1)+P(s_1 \mid c_2)+P(s_2 \mid c_2) = 2/4+2/4+4/8+4/8 = 2$. Intuitively, we are summing up two different complete probability distributions, $P(S \mid c_1)$ and $P(S \mid c_2)$: one world where there are no candles, and another world where there are definitely candles.

6. $\sum_{s\in\{s_1,s_2\}} \sum_{c\in\{c_1,c_2\}} P(c \mid s)$

   $P(c_1 \mid s_1)+P(c_2 \mid s_1)+P(c_1 \mid s_2)+P(c_2 \mid s_2) = 2/6+4/6+2/6+4/6 = 2$. Intuitively, we are summing up two different complete probability distributions, $P(C \mid s_1)$ and $P(C \mid s_2)$: one world where there is no sprinkles, and another world where there is definitely sprinkles.

4 Queries on a Large Joint Distribution

Consider binary (two outcomes) random variables A, B, C, D, and the following joint distribution table of all four variables.

A	B	C	D	$P(A,B,C,D)$
$+a$	$+b$	$+c$	$+d$	$12/64$
$+a$	$+b$	$+c$	$-d$	$4/64$
$+a$	$+b$	$-c$	$+d$	$2/64$
$+a$	$+b$	$-c$	$-d$	$2/64$
$+a$	$-b$	$+c$	$+d$	$8/64$
$+a$	$-b$	$+c$	$-d$	$4/64$
$+a$	$-b$	$-c$	$+d$	$2/64$
$+a$	$-b$	$-c$	$-d$	$4/64$
$-a$	$+b$	$+c$	$+d$	$6/64$
$-a$	$+b$	$+c$	$-d$	$3/64$
$-a$	$+b$	$-c$	$+d$	$4/64$
$-a$	$+b$	$-c$	$-d$	$6/64$
$-a$	$-b$	$+c$	$+d$	$2/64$
$-a$	$-b$	$+c$	$-d$	$1/64$
$-a$	$-b$	$-c$	$+d$	$3/64$
$-a$	$-b$	$-c$	$-d$	$1/64$

1. Calculate the following probabilities:

(a) $P(+c)$

Sum all the entries that contain $+c$ to get: $P(+c) = \sum_a \sum_b \sum_d P(a, b, +c, d) = 40/64$

(b) $P(+a, -b)$

Sum all the entries that contain both $+a$ and $-b$ to get: $P(+a, -b) = \sum_c \sum_d P(+a, -b, c, d) = 18/64$

(c) $P(-b \mid +a)$

1) Sum all entries with both $+a$ and $-b$ to get: $P(+a, -b) = \sum_c \sum_d P(+a, -b, c, d) = 18/64$

2) Sum all entries with $+a$ to get: $P(+a) = \sum_b \sum_c \sum_d P(+a, b, c, d) = 38/64$

3) Use definition of conditional probability to compute: $P(-b \mid +a) = \frac{P(+a,-b)}{P(+a)} = 18/38$

(d) $P(-a, +b, -d)$

Sum all entries with $-a$, $+b$, and $-d$ to get: $P(-a, +b, -d) = \sum_c P(-a, +b, c, -d) == P(-a, +b, +c, -d) + P(-a, +b, -c, -d) = 9/64$

(e) $P(+c \mid -a, +b, -d)$

1) Find entry with $-a, +b, +c, -d$ to get: $P(-a, +b, +c, -d) = 3/64$

2) Sum all entries with $-a, +b, -d$ to get: $P(-a, +b, -d) = \sum_c P(-a, +b, c, -d) = 9/64$

3) Use definition of conditional probability to compute: $P(+c \mid -a, +b, -d) = \frac{P(-a,+b,+c,-d)}{P(-a,+b,-d)} = 3/9$

(f) $P(+c, +d \mid +a, +b)$

1) Find entry with $+a, +b, +c, +d$ to get: $P(+a, +b, +c, +d) = 12/64$

2) Sum all entries with $+a, +b$ to get: $P(+a, +b) = \sum_c \sum_d P(+a, +b, c, d) = 20/64$

3) Use definition of conditional probability to compute: $P(+c, +d \mid +a, +b) = \frac{P(+a,+b,+c,+d)}{P(+a,+b)} = 12/20$

2. What value do the following probability tables sum to?

(a) $P(B)$

The short answer is that we are considering the entries for all the possible values of B, so this should sum to 1. You could calculate both entries in this table to convince yourself, $P(+b)$ and $P(-b)$.

(b) $P(+b \mid C, +d)$

Sadly, there is no shortcut here. The two entries in this table come from two different worlds that are unrelated: one world where $+c$ and $+d$ are given; and another world where $-c$ and $+d$ are given.

Important note: There is no real reason to add these numbers together in this strange probability table. This is primarily a counterexample to show that these do *not* sum to one.

We compute these two values similar to the methods in the previous questions, and then add them together.

$P(+b \mid +c, +d) = \frac{P(+b,+c,+d)}{P(+c,+d)} = 9/14$

$P(+b \mid -c, +d) = \frac{P(+b,-c,+d)}{P(-c,+d)} = 6/11$ When we add these two values together, we get 1.1883.

(c) $P(C, D \mid +a, +b)$

The short answer is that we are considering all possible entries for a single world where $+a$ and $+b$ are given, so this should sum to 1. You could calculate all four entries in this table to convince yourself, $P(+c, +d \mid +a, +b)$, $P(+c, -d \mid +a, +b)$, $P(-c, +d \mid +a, +b)$, $P(-c, -d \mid +a, +b)$.