MATH3011

Probabilistic Models and Inference

Tutorial: Week 1 (Autumn 2025)

Shavika Rastogi Oliver Obst

Exercises

(Original from David Sontag's lecture)

Task 1

A fair coin is tossed 4 times. Define X to be the number of heads in the first 2 tosses, and Y to be the number of heads in all 4 tosses.

- Calculate the table of the joint probability $p(X, Y)$.
- Calculate the tables of marginal probabilities $p(X)$ and $p(Y)$.
- Calculate the tables of conditional probabilities $p(X|Y)$ and $p(Y|X)$.
- What is the distribution of $Z = Y - X$?

Task 2

You go for your yearly checkup and have several lab tests performed. A week later your doctor calls you and says she has good and bad news. The bad news is that you tested positive for a marker of a serious disease, and that the test is 99% accurate (i.e., the probability of testing positive given that you have the disease is 0.99, as is the probability of testing negative given that you do not have the disease). The good news is that this is a rare disease, striking only 1 in 20,000 people. Why is it good news that the disease is rare? What are the chances that you actually have the disease?

Task 3

Show that the statement

$$p(A, B|C) = p(A|C)p(B|C)$$

is equivalent to the statement

$$p(A|B, C) = p(A|C)$$

and also to

$$p(B|A, C) = p(B|C)$$

(you need to show both directions, i.e., that each statement implies the other).

Task 4

We want to understand how knowing about conditional independence can reduce the amount of information needed to calculate probabilities.

Consider three random variables:

$$H, \quad E_1, \quad \text{and} \quad E_2.$$

Here, $p(H)$ represents the probability for every possible value of H.

(a) Assume we want to compute the probability $p(H \mid E_1, E_2)$ and we do not have any information about conditional independence. Which set of numbers would be enough to perform this calculation? Consider the following options:

1. $p(E_1, E_2), \quad p(H), \quad p(E_1 \mid H), \quad p(E_2 \mid H)$.
2. $p(E_1, E_2), \quad p(H), \quad p(E_1, E_2 \mid H)$.
3. $p(E_1 \mid H), \quad p(E_2 \mid H), \quad p(H)$.

Explain your reasoning for choosing one of these sets.

(b) Now assume we know that E_1 and E_2 are conditionally independent given H. With this extra information, which of the three sets listed above would be enough to calculate $p(H \mid E_1, E_2)$? Explain why the answer might change when conditional independence is known.

Probabilistic Models and Inference

Tutorial: Week 1 (Autumn 2025)

Shavika Rastogi Oliver Obst

Solutions

Task 1

We consider an experiment in which a fair coin is tossed 4 times. An *outcome* is a sequence of 4 coin results (each being Head, H, or Tail, T), for example, `HTTH`. Since the coin is fair and tosses are independent, each outcome has probability 1/16.

Now, we define

$$X = \text{number of heads in the first two tosses}$$

and

$$Y = \text{number of heads in all four tosses.}$$

Notice that the remaining heads, $Z = Y - X$, come from the last two tosses.

For two tosses, the binomial coefficient $\binom{2}{k}$ (read as "2 choose k") counts the number of ways to obtain k heads. In our example, the computed values are:

$$\binom{2}{0} = 1, \quad \binom{2}{1} = 2, \quad \binom{2}{2} = 1.$$

(a) Joint Probability Table $p(X, Y)$
For the first two tosses, X can be 0, 1, or 2 with probabilities:

$$P(X = x) = \binom{2}{x}\left(\frac{1}{2}\right)^2 \quad \text{for } x = 0, 1, 2.$$

For the last two tosses, if we let $k = Y - X$ (i.e. the number of heads among tosses 3 and 4), then:

$$P(k) = \binom{2}{k}\left(\frac{1}{2}\right)^2, \quad k = 0, 1, 2.$$

Thus, the joint probability is given by:

$$p(X = x, Y = y) = \begin{cases} \binom{2}{x}\binom{2}{y-x}\left(\frac{1}{2}\right)^4, & \text{if } 0 \le x \le 2 \text{ and } 0 \le y - x \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

Substituting the computed values for $\binom{2}{k}$, the joint table becomes:

$X \backslash Y$	0	1	2	3	4
0	$1 \cdot 1 \cdot \frac{1}{16} = \frac{1}{16}$	$1 \cdot 2 \cdot \frac{1}{16} = \frac{2}{16}$	$1 \cdot 1 \cdot \frac{1}{16} = \frac{1}{16}$	0	0
1	0	$2 \cdot 1 \cdot \frac{1}{16} = \frac{2}{16}$	$2 \cdot 2 \cdot \frac{1}{16} = \frac{4}{16}$	$2 \cdot 1 \cdot \frac{1}{16} = \frac{2}{16}$	0
2	0	0	$1 \cdot 1 \cdot \frac{1}{16} = \frac{1}{16}$	$1 \cdot 2 \cdot \frac{1}{16} = \frac{2}{16}$	$1 \cdot 1 \cdot \frac{1}{16} = \frac{1}{16}$

(b) Marginal Probabilities $p(X)$ and $p(Y)$
Marginal probabilities are the probabilities of one variable irrespective of the other. They are obtained by summing (or integrating) the joint probabilities over the other variable.

For X:

$$p(X = x) = \sum_y p(X = x, y) = \binom{2}{x} \left(\frac{1}{2}\right)^2 = \binom{2}{x} \frac{1}{4}.$$

Thus:

$$p(X = 0) = \frac{1}{4}, \quad p(X = 1) = \frac{2}{4} = \frac{1}{2}, \quad p(X = 2) = \frac{1}{4}.$$

For Y, we sum over x:

- $y = 0$: Only $x = 0$ contributes, so $p(Y = 0) = \frac{1}{16}$.
- $y = 1$: Contributions from $x = 0$ and $x = 1$ yield $p(Y = 1) = \frac{2}{16} + \frac{2}{16} = \frac{4}{16}$.
- $y = 2$: Contributions from $x = 0, 1, 2$ give $p(Y = 2) = \frac{1}{16} + \frac{4}{16} + \frac{1}{16} = \frac{6}{16}$.
- $y = 3$: Contributions from $x = 1$ and $x = 2$ yield $p(Y = 3) = \frac{2}{16} + \frac{2}{16} = \frac{4}{16}$.
- $y = 4$: Only $x = 2$ contributes, so $p(Y = 4) = \frac{1}{16}$.

(c) Conditional Probabilities $p(X \mid Y)$ and $p(Y \mid X)$
The conditional probability is defined as:

$$p(X = x \mid Y = y) = \frac{p(X = x, Y = y)}{p(Y = y)}.$$

For example, when $y = 2$:

$$p(X = 0 \mid Y = 2) = \frac{1/16}{6/16} = \frac{1}{6}, \quad p(X = 1 \mid Y = 2) = \frac{4/16}{6/16} = \frac{2}{3}, \quad p(X = 2 \mid Y = 2) = \frac{1/16}{6/16} = \frac{1}{6}.$$

Similarly, for $p(Y \mid X)$:

$$p(Y = y \mid X = x) = \frac{p(X = x, Y = y)}{p(X = x)}.$$

For instance, when $x = 1$ (with $p(X = 1) = \frac{1}{2}$):

$$p(Y = 1 \mid X = 1) = \frac{2/16}{1/2} = \frac{4}{16} = \frac{1}{4},$$
$$p(Y = 2 \mid X = 1) = \frac{4/16}{1/2} = \frac{8}{16} = \frac{1}{2},$$
$$p(Y = 3 \mid X = 1) = \frac{2/16}{1/2} = \frac{4}{16} = \frac{1}{4}.$$

(d) Distribution of $Z = Y - X$

The variable Z represents the number of heads in tosses 3 and 4. Its distribution is given by:

$$P(Z = z) = \binom{2}{z}\left(\frac{1}{2}\right)^2, \quad z = 0, 1, 2.$$

Thus, substituting the computed binomial coefficients:

$$P(Z = 0) = 1 \cdot \frac{1}{4} = \frac{1}{4}, \quad P(Z = 1) = 2 \cdot \frac{1}{4} = \frac{1}{2}, \quad P(Z = 2) = 1 \cdot \frac{1}{4} = \frac{1}{4}.$$

Task 2

Suppose:

$$D = \text{having the disease}, \quad \bar{D} = \text{not having the disease}, \quad T^+ = \text{testing positive}.$$

We are given:

$$\begin{aligned} P(T^+ \mid D) &= 0.99, \\ P(T^- \mid \bar{D}) &= 0.99 \quad \Rightarrow \quad P(T^+ \mid \bar{D}) = 0.01, \\ P(D) &= \frac{1}{20000} = 0.00005, \quad P(\bar{D}) = 0.99995. \end{aligned}$$

By Bayes' theorem, the probability of having the disease given a positive test is:

$$P(D \mid T^+) = \frac{P(T^+ \mid D)P(D)}{P(T^+ \mid D)P(D) + P(T^+ \mid \bar{D})P(\bar{D})}.$$

Substituting the values:

$$P(D \mid T^+) = \frac{0.99 \cdot 0.00005}{0.99 \cdot 0.00005 + 0.01 \cdot 0.99995}.$$

Calculating the numerator:

$$0.99 \cdot 0.00005 = 0.0000495.$$

Calculating the denominator:

$$0.0000495 + 0.01 \cdot 0.99995 \approx 0.0000495 + 0.0099995 = 0.010049.$$

Thus,

$$P(D \mid T^+) \approx \frac{0.0000495}{0.010049} \approx 0.004925,$$

or about 0.4925%.

Interpretation: The disease is rare (only 1 in 20,000 people), so even though the test is 99% accurate, the number of false positives (due to the large number of healthy individuals) greatly exceeds the number of true positives. Hence, a positive result still implies only about a 0.5% chance of actually having the disease, which is the good news.

Task 3

We wish to show that

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

is equivalent to

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

and to

$$p(B \mid A, C) = p(B \mid C).$$

(1) Show that $p(A, B \mid C) = p(A \mid C)p(B \mid C) \implies p(A \mid B, C) = p(A \mid C)$.

By the definition of conditional probability, provided $p(B, C) > 0$,

$$p(A \mid B, C) = \frac{p(A, B \mid C)}{p(B \mid C)}.$$

Assuming

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

we substitute to get

$$p(A \mid B, C) = \frac{p(A \mid C)p(B \mid C)}{p(B \mid C)} = p(A \mid C).$$

(2) Show that $p(A \mid B, C) = p(A \mid C) \implies p(A, B \mid C) = p(A \mid C)p(B \mid C)$.

Using the definition of conditional probability:

$$p(A, B \mid C) = p(A \mid B, C)p(B \mid C).$$

If $p(A \mid B, C) = p(A \mid C)$, then

$$p(A, B \mid C) = p(A \mid C)p(B \mid C).$$

(3) The equivalence with $p(B \mid A, C) = p(B \mid C)$ **follows by symmetry.**

(a) Assuming $p(A, B \mid C) = p(A \mid C)p(B \mid C)$, then by definition,

$$p(B \mid A, C) = \frac{p(A, B \mid C)}{p(A \mid C)} = \frac{p(A \mid C)p(B \mid C)}{p(A \mid C)} = p(B \mid C).$$

(b) Conversely, if $p(B \mid A, C) = p(B \mid C)$, then

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

which is the same as the original statement.

Thus, we have shown that each statement implies the other.

Task 4

We are given three random variables H, E_1 and E_2 and wish to compute

$$p(H \mid E_1, E_2) = \frac{p(H, E_1, E_2)}{p(E_1, E_2)}.$$

Since

$$p(H, E_1, E_2) = p(H)p(E_1, E_2 \mid H),$$

the essential ingredient is $p(E_1, E_2 \mid H)$. In the three options below, we ask whether the provided numbers are sufficient to compute $p(H \mid E_1, E_2)$.

(a) No conditional independence information

(a) $p(E_1, E_2)$, $p(H)$, $p(E_1 \mid H)$, $p(E_2 \mid H)$

Without any conditional independence assumption, there is no general way to combine $p(E_1 \mid H)$ and $p(E_2 \mid H)$ to obtain $p(E_1, E_2 \mid H)$. In the absence of further assumptions, the product $p(E_1 \mid H)p(E_2 \mid H)$ is not guaranteed to equal $p(E_1, E_2 \mid H)$. Hence, this set is *not sufficient*.

(b) $p(E_1, E_2)$, $p(H)$, $p(E_1, E_2 \mid H)$

Here the joint probability $p(E_1, E_2 \mid H)$ is directly provided. Therefore, one can compute:

$$p(H, E_1, E_2) = p(H)p(E_1, E_2 \mid H)$$

and then

$$p(H \mid E_1, E_2) = \frac{p(H)p(E_1, E_2 \mid H)}{p(E_1, E_2)}.$$

Thus, this set is *sufficient*.

(c) $p(E_1 \mid H)$, $p(E_2 \mid H)$, $p(H)$

As in option (a), without the assumption of conditional independence, $p(E_1, E_2 \mid H)$ cannot be recovered from $p(E_1 \mid H)$ and $p(E_2 \mid H)$ alone. In addition, note that $p(E_1, E_2)$ is not provided and would need to be computed via marginalisation (which in turn requires $p(E_1, E_2 \mid H)$). Hence, this set is also *not sufficient*.

Conclusion for (a): Only option (b) is sufficient when no conditional independence information is available.

(b) With the assumption that E_1 and E_2 are conditionally independent given H
The assumption of conditional independence implies that

$$p(E_1, E_2 \mid H) = p(E_1 \mid H)\, p(E_2 \mid H).$$

Thus, we re-examine the options:

(a) $p(E_1, E_2)$, $p(H)$, $p(E_1 \mid H)$, $p(E_2 \mid H)$

Now, since $p(E_1, E_2 \mid H)$ can be computed as $p(E_1 \mid H)p(E_2 \mid H)$, this set provides all the necessary information to compute:

$$p(H, E_1, E_2) = p(H)\, p(E_1 \mid H)p(E_2 \mid H)$$

and hence $p(H \mid E_1, E_2)$. Thus, option (a) becomes *sufficient*.

(b) $p(E_1, E_2)$, $p(H)$, $p(E_1, E_2 \mid H)$

This set is also sufficient since it explicitly provides the joint likelihood $p(E_1, E_2 \mid H)$.

(c) $p(E_1 \mid H)$, $p(E_2 \mid H)$, $p(H)$

Even though $p(E_1, E_2)$ is not directly given, we can compute it by marginalising over H:

$$p(E_1, E_2) = \sum_h p(H = h)p(E_1, E_2 \mid H = h) = \sum_h p(H = h)p(E_1 \mid H = h)p(E_2 \mid H = h).$$

Therefore, option (c) is also *sufficient*.

Conclusion for (b): With the conditional independence assumption, all three options (a), (b), and (c) provide sufficient information to calculate $p(H \mid E_1, E_2)$.