Math 170E

Lecture Notes Section 1.4 – 1.5 *†

Independent events and Bayes theorem

Instructor: Swee Hong Chan

NOTE: Materials that appear in the textbook but do not appear in the lecture notes might still be tested. Please send me an email if you find typos.

---

*Version date: Wednesday 13th January, 2021, 10:46.

†This notes is based on March Boedihardjo's and Jamie Haddock's notes from the past quarters, and I would like to thank them for their generosity. "*Nanos gigantum humeris insidentes* (I am but a dwarf standing on the shoulders of giants)".

1 Independent events

Events A and B are **independent** if

$$P(A \cap B) = P(A)P(B).$$

Otherwise A and B are called dependent events.

In particular, if A and B are independent and $P(B) > 0$,

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

2 Independent events: Theorem

Theorem 1. *If A and B are independent events, then so are:*

- A *and* B';
- A' *and* B;
- A' *and* B'.

3 Independent events: Proof

We have (BT)

$$
\begin{aligned}
P(A' \cap B') &= P((A \cup B)') \\
&= 1 - P(A \cup B) \\
&= 1 - P(A) - P(B) + P(A \cap B) \\
&= 1 - P(A) - P(B) + P(A)\,P(B) \qquad \text{(independence)} \\
&= (1 - P(A))\,(1 - P(B)) \\
&= P(A')\,P(B').
\end{aligned}
$$

We leave the other two cases as exercises.

4 Independence for three elements

A, B, and C are **pairwise independent** if

$$P(A \cap B) = P(A)P(B);$$

$$P(A \cap C) = P(A)P(C);$$

$$P(B \cap C) = P(B)P(C).$$

A, B, and C are **mutually independent** if they are pairwise independent and

$$P(A \cap B \cap C) = P(A)P(B)P(C).$$

Important: Pairwise independence **does not imply** mutual independence!

5 Bayes' theorem

Theorem 2. *Suppose that* $P(A) > 0$ *and* $P(B) > 0$. *Then*

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

Proof. We have

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

$\square$

Note: Review the Covid Example from Lecture 1.3.

6 Law of total probability

Let $B_1, B_2, \ldots, B_m$ be mutually exclusive and exhaustive events, and $P(B_k) > 0$ for $k = 1, 2, \ldots, m$. Then

$$\begin{aligned} P(A) &= \sum_{k=1}^{m} P(A \mid B_k)\, P(B_k) \\ &= P(A \mid B_1)\, P(B_1) \;+\; \ldots \;+\; P(A \mid B_m)\, P(B_m). \end{aligned}$$

Proof.

$$\begin{aligned} P(A) &= P(A \cap (B_1 \cup \ldots \cup B_k)) \\ &= P(A \cap B_1) \;+\; \ldots \;+\; P(A \cap B_k) \\ &= P(A \mid B_1)\, P(B_1) \;+\; \ldots \;+\; P(A \mid B_m)\, P(B_m). \end{aligned}$$

$\square$

7 Bayesian statistics: Example

In a screw factory, three machines I, II, III are producing screws of the same length.

In terms of accuracy, Machine I, II, and III respectively produce 1%, 2%, and 3% of defective springs.

In terms of productivity, Machine I produces 25% of the screws, Machine II produces 35%, and Machine III produces 40%.

You select a screw at random, and it turns out that the screw is defective. You are now tasked to find the machine making the defective screw.

8 Bayesian statistics: Answer

Let A be the event that a screw you select is defective. Let B_1, B_2, and B_3 be the event that Machine I, II, and III are the producer of the screw, respectively.

In the language of Bayesian statistics:

- A is the **current evidence** that something has gone wrong in the production.
- $P(A)$ is sometimes called the **marginal likelihood** or the **moral evidence**; The probability $P(A)$ is **unknown**.
- B_1, B_2, B_3 are the **hypothesis** that you are making regarding the cause of the problem.

- $P(B_1)$, $P(B_2)$, $P(B_3)$ are called **prior probabilities** *before* the evidence, the broken screw, is observed. We know that

$$P(B_1) = 0.25; \qquad P(B_2) = 0.35 \qquad P(B_3) = 0.4.$$

- $P(A \mid B_1)$, $P(A \mid B_2)$, $P(A \mid B_3)$ are **likelihoods** that the broken screw is produced by the chosen machines. We know that

$$P(A \mid B_1) = 0.01; \;\; P(A \mid B_2) = 0.02; \;\; P(A \mid B_3) = 0.03.$$

- $P(B_1 \mid A)$, $P(B_2 \mid A)$, $P(B_3 \mid A)$ are **posterior probabilities**. This is **what the question is asking**, the probability of a hypothesis given the evidence.

By the law of total probability,

$$
\begin{aligned}
P(A) =& P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + P(A \mid B_3)P(B_3) \\
=& \frac{25}{100} \times \frac{1}{100} + \frac{35}{100} \times \frac{2}{100} + \frac{40}{100} \times \frac{3}{100} \\
=& \frac{25}{1000} + \frac{70}{10000} + \frac{120}{1000} \\
=& \frac{215}{10000}.
\end{aligned}
$$

It then follows from the Bayes theorem that

$$
\begin{aligned}
P(B_1 \mid A) &= \frac{P(A \mid B_1)\, P(B_1)}{P(A)} = \frac{\frac{1}{100} \times \frac{25}{100}}{\frac{215}{1000}} = \frac{25}{215}. \\
P(B_2 \mid A) &= \frac{P(A \mid B_2)\, P(B_2)}{P(A)} = \frac{\frac{2}{100} \times \frac{35}{100}}{\frac{215}{1000}} = \frac{70}{215} \\
P(B_3 \mid A) &= 1 - P(B_1 \mid A) - P(B_2 \mid A) \\
&= 1 - \frac{25}{215} - \frac{70}{215} = \frac{120}{215}.
\end{aligned}
$$