ECE 361 Probability for Engineers (Fall, 2016)
Lecture 1b

Some background for homework 1

Homework 1 problem 1 asks to find the probability of each outcome of the following experiment.

> Consider flipping a *fair* coin. The coin is repeatedly flipped until a head occurs, and the outcome of the experiment is the number of times the coin is flipped.

A straightforward argument establishes p_n, the probability that n flips are required, is $(1/2)^n$. But problem 2 is significantly harder, and requires more thought. As a help with problem 2, consider the following alternative approach to problem 1.

Define the event A_n (for $n \in \mathbb{N}$) as consisting of all outcomes where *strictly* more than n flips are required, which means the first n flips were tails. Then $\mathbb{P}(A_n) = (1/2)^n$. Now consider A_n^c, the complement of A_n, which has probability $\mathbb{P}(A_n^c) = 1 - (1/2)^n$. In words, A_n^c is the event that n or fewer flips are required.

Define the event B_n (for $n \in \mathbb{N}$) as the outcome where exactly n flips are required. Thus $p_n = \mathbb{P}(B_n)$ is the quantity of interest in the problem. But observe the following fundamental facts:

$$A_n^c = B_n \cup A_{n-1}^c, \; B_n \cap A_{n-1}^c = \emptyset. \tag{1}$$

In words, i) the event that n or fewer flips are required equals the union of the events that exactly n flips are required with the event that $n-1$ or fewer flips are required, and ii) the event that exactly n flips are required and the event that $n-1$ or fewer flips are required are disjoint. The truth of these two statements is self-evident. Also observe $A_1^c \subset A_2^c \subset A_3^c \subset \cdots$, as is also self-evident when the symbols are translated into English. The point is that, by the additivity axiom, we can compute:

$$\begin{aligned}
\mathbb{P}(A_n^c) &= \mathbb{P}(B_n) + \mathbb{P}(A_{n-1}^c) \\
p_n &= \mathbb{P}(A_n^c) - \mathbb{P}(A_{n-1}^c) \\
&= (1-(1/2)^n) - (1-(1/2)^{n-1}) \\
&= (1/2)^{n-1} - (1/2)^n \\
&= (1/2)^{n-1}(1-1/2) \\
&= (1/2)^n
\end{aligned} \tag{2}$$

The logic in this argument is standard in probability, and can be used to solve problem 2.

§1.2 Probabilistic models

Properties of probability laws

Many further results follow from the probability axioms. Here are some examples involving one event:

- $\mathbb{P}(\emptyset) = 0$. Proof:
$$1 = \mathbb{P}(\Omega) = \mathbb{P}(\Omega \cup \emptyset) = \mathbb{P}(\Omega) + \mathbb{P}(\emptyset) = 1 + \mathbb{P}(\emptyset). \tag{3}$$
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$. Proof: note that $A \cap A^c = \emptyset$ and $A \cup A^c = \Omega$.
$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c). \tag{4}$$
- $\mathbb{P}(A) \leq 1$. Proof: use the above result and non-negativity axiom.
$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c) \leq 1. \tag{5}$$

Here are some examples involving two events:

- If $A \subset B$ then $\mathbb{P}(A) \leq \mathbb{P}(B)$. Proof:

$$\mathbb{P}(B) = \mathbb{P}(A \cup (B \setminus A)) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A). \tag{6}$$

- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ for all events A, B. Proof: there are two steps. First step: observe $B = (A \cap B) \cup (B \setminus A)$, where $A \cap B$ and $B \setminus A$ are disjoint. Thus

$$\mathbb{P}(B) = \mathbb{P}((A \cap B) \cup (B \setminus A)) = \mathbb{P}(A \cap B) + \mathbb{P}(B \setminus A), \tag{7}$$

and so $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$. Second step: observe $A \cup B = A \cup (B \setminus A)$, where A and $B \setminus A$ are disjoint. Thus

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \cup (B \setminus A)) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B). \tag{8}$$

- $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$. Proof: use non-negativity on the above result.

These results are of fundamental importance, and so we list them again.

Basic relationships among probabilities of events. For any two events A, B:

- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
- If $A \subseteq B$ then $\mathbb{P}(A) \leq \mathbb{P}(B)$.
- If $A \cap B = \emptyset$ (A, B are disjoint) then $\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B)$.
- $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$, with equality when $A \cap B = \emptyset$.

In fact the last three are consequences (special cases) of the first. A key generalization of the last result above is the so-called union bound:

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i). \tag{9}$$

In words: the probability of a union of events is upper bounded by the sum of the probabilities of those events. Note the union bound holds with equality when the events are *disjoint*.

§1.3 Conditional probability

Conditional probability captures partial information in a random experiment. Here are some examples:

- In rolling two dice, what is the probability the first die is a 6 given you are told the sum is 9?
- How likely is it that a person has a disease given a test is negative?

Formally, partial information is indicated by restricting the set of outcomes of the experiment − the more complete the information the more restrictive the set of outcomes. Write $\mathbb{P}(A|B)$ to denote the probability of the event A given knowledge that the outcome will be in B. To be clear: $\mathbb{P}(A|B) = \mathbb{P}(\omega \in A | \omega \in B)$.

Conditional probability definition.

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad \text{for all events } B \text{ with } \mathbb{P}(B) > 0. \tag{10}$$

Conditional probabilities specify a probability law

Conditioning on any event B with $\mathbb{P}(B) > 0$ induces a new probability law that satisfies the probability axioms:

- Nonnegativity: $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \geq 0$.

- Additivity: Fix disjoint A_1, A_2. Then:

$$\begin{aligned}\mathbb{P}(A_1 \cup A_2 | B) &= \frac{\mathbb{P}((A_1 \cup A_2) \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}((A_1 \cap B) \cup (A_2 \cap B))}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_1 \cap B) + \mathbb{P}(A_2 \cap B)}{\mathbb{P}(B)} \\ &= \frac{\mathbb{P}(A_1 \cap B)}{\mathbb{P}(B)} + \frac{\mathbb{P}(A_2 \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A_1|B) + \mathbb{P}(A_2|B). \end{aligned} \quad (11)$$

- Normalization: $\mathbb{P}(\Omega|B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$.

Since conditioning induces a valid probability law, it follows that all the consequences of the probability axioms also hold under conditioning. For example:

$$\mathbb{P}(A \cup C|B) \leq \mathbb{P}(A|B) + \mathbb{P}(C|B). \quad (12)$$

Moreover, if all outcomes in Ω are equally likely, we have the conditional discrete uniform probability law:

$$\mathbb{P}(\omega) = \frac{1}{|\Omega|} \forall \omega \in \Omega \Rightarrow \mathbb{P}(A|B) = \frac{|A \cap B|}{|B|}. \quad (13)$$

Here are three examples of this law.

1. Toss a fair coin three times successively. What is the probability that more heads than tails come up (A) given that the first toss is a head (B)? Note $B = \{HHH, HHT, HTH, HTT\}$ and $A \cap B = \{HHH, HHT, HTH\}$ and thus $\mathbb{P}(A|B) = 3/4$.
2. A fair 4-sided die is tossed twice, where (X, Y) are the results of the two rolls. Find the probability the max is m (A) given the min is 2 (B). Note $B = \{(2,2), (2,3), (2,4), (3,2), (4,2)\}$ and thus

$$\mathbb{P}(\max(X,Y) = m|B) = \begin{cases} 2/5, & m = 4 \\ 2/5, & m = 3 \\ 1/5, & m = 2 \\ 0, & m = 1 \end{cases} \quad (14)$$

   Observe:

$$B = \begin{array}{cc} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} & & & \\ & \times & \times & \times \\ & \times & & \\ & \times & & \end{bmatrix} \end{array}, \; A_2 = \begin{array}{cc} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} & \times & & \\ \times & \times & & \\ & & & \\ & & & \end{bmatrix} \end{array}, \; A_3 = \begin{array}{cc} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} & & \times & \\ & & \times & \\ \times & \times & \times & \\ & & & \end{bmatrix} \end{array}, \; A_4 = \begin{array}{cc} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} & & & \times \\ & & & \times \\ & & & \times \\ \times & \times & \times & \times \end{bmatrix} \end{array} \quad (15)$$

3. A conservative design team (C) and an innovative design team (N) design a new product, and each team is either successful (S) or unsuccessful (F), with the probability that C succeeds as 2/3, that N succeeds as 1/2, and that at least one succeeds as 3/4. Given exactly one successful design is produced (B), what is the probability it was by N (A)?

   First write down the sample space and the events of interest:

   - $\Omega = \{(CS, NS), (CS, NF), (CF, NS), (CF, NF)\}$
   - $B = \{(CS, NF), (CF, NS)\}$
   - $A = \{(CS, NS), (CF, NS)\}$
   - $A \cap B = \{(CF, NS)\}$

   Alternately expressed as:

$$\Omega = \begin{array}{cc} & \begin{array}{cc} NS & NF \end{array} \\ \begin{array}{c} CS \\ CF \end{array} & \begin{bmatrix} (CS, NS) & (CS, NF) \\ (CF, NS) & (CF, NF) \end{bmatrix} \end{array} \quad (16)$$

We have four unknowns: the probabilities of the four outcomes. We adopt a shorthand notation as follows, where right column and the bottom row are the sums of the corresponding elements:

$$\mathbb{P} = \left| \begin{array}{cc|c} p_{CS,NS} & p_{CS,NF} & p_{CS} \\ p_{CF,NS} & p_{CF,NF} & p_{CF} \\ \hline p_{NS} & p_{NF} & 1 \end{array} \right. \tag{17}$$

Here the third row and third column are the sums of their entries, and the lower right one is the sum of the four uknowns. We are given that:

$$p_{CS} = 2/3, \; p_{NS} = 1/2, \; p_{CF,NF} = 1/4. \tag{18}$$

These equations can be solved to find:

$$\mathbb{P} = \begin{array}{cc|c} 5/12 & 1/4 & 2/3 \\ 1/12 & 1/4 & 1/3 \\ \hline 1/2 & 1/2 & 1 \end{array} \tag{19}$$

Then:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1/12}{1/4 + 1/12} = 1/4. \tag{20}$$

Using conditional probability for modeling

Often it is more natural to specify the probabilistic model in terms of the conditional probabilities $\mathbb{P}(A|B)$, than to specify the joint probabilities $\mathbb{P}(A \cap B)$. The joint probabilities may be obtained from the conditional probabilities via $\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B)$. This is illustrated in the following example.

Radar detection. If an aircraft is present (A) it may generate an alarm (B). There are four joint events of interest:

- Present and detected: $A \cap B$
- Present and not detected: $A \cap B^c$ (missed detection)
- Absent and detected $A^c \cap B$ (false alarm)
- Absent and not detected $A^c \cap B^c$

The probabilities are specified conditionally (this is the model):

$$\begin{array}{rcl} \text{probability of correct detection: } P(B|A) & = & 0.99 \text{ (implies } P(B^c|A) = 0.01) \\ \text{probability of false alarm: } P(B|A^c) & = & 0.10 \text{ (implies } P(B^c|A^c) = 0.90) \\ \text{probability an aircraft is present: } P(A) & = & 0.05 \text{ (implies } P(A^c) = 0.95) \end{array} \tag{21}$$

We are asked to find the above four probabilities using this data:

$$\begin{array}{rcl} \mathbb{P}(A \cap B) & = & \mathbb{P}(A)\mathbb{P}(B|A) = 0.05 \times 0.99 = 0.0495 \\ \mathbb{P}(A \cap B^c) & = & \mathbb{P}(A)\mathbb{P}(B^c|A) = 0.05 \times 0.01 = 0.0005 \\ \mathbb{P}(A^c \cap B) & = & \mathbb{P}(A^c)\mathbb{P}(B|A^c) = 0.95 \times 0.1 = 0.095 \\ \mathbb{P}(A^c \cap B^c) & = & \mathbb{P}(A^c)\mathbb{P}(B^c|A^c) = 0.95 \times 0.90 = 0.855 \end{array} \tag{22}$$

Note these four sum to one. We can write the sample space as

$$\Omega = \begin{array}{c} \\ A \\ A^c \end{array} \begin{array}{c} \begin{array}{cc} \quad B \quad & \quad B^c \quad \end{array} \\ \left[\begin{array}{cc} \text{present and detected} & \text{missed detection} \\ \text{false alarm} & \text{absent and not detected} \end{array} \right] \end{array}. \tag{23}$$

Filling in the above probabilities:

$$\mathbb{P} = \begin{array}{cc|c} 0.0495 & 0.0005 & 0.05 \\ 0.0950 & 0.8550 & 0.95 \\ \hline 0.1445 & 0.8555 & 1.00 \end{array} \tag{24}$$

References

[1] *Introduction to Probability, 2nd Edition* by Dimitri P. Bertsekas and John N. Tsitsiklis, Athina Scientific Press, 2008.