ENEE 324 Engineering Probability Lecture 4

Applications of Bayes' Theorem

Example: There are 10 urns, 9 of which are of type I and 1 of type II. Urn of type I carries 2 white balls and 2 black balls. Urn of type II carries 5 white balls and 1 black ball.

If a ball drawn randomly from a randomly chosen urn turns out to be *white*, then what is the probability that the chosen urn is of type II? This is a model of an inference problem.

Solution

$A :=$ ball drawn is white
$B_1 :=$ urn is of type I
$B_2 :=$ urn is of type II
B_1 and B_2 are disjoint events and define a partition $\Omega = B_1 \cup B_2$.

$$\begin{aligned} P(B_2|A) &= \tfrac{P(A|B_2)\ P(B_2)}{P(A|B_1)\ P(B_1)+P(A|B_2)\ P(B_2)} \\ P(B_1) &= 9/10\ ;\ P(B_2) = 1/10 \qquad \text{Prior probabilities} \\ P(A|B_1) &= 2/4 = 1/2 \\ P(A|B_2) &= 5/6 \end{aligned}$$

$$\begin{aligned} P(B_2|A) &= \frac{5/6 \cdot 1/10}{1/2 \cdot 9/10 + 5/6 \cdot 1/10} \\ &= \frac{5}{27+5} = \frac{5}{32} \qquad \square \end{aligned}$$

Statistical Independence; The idea that two phenomena have nothing to do with each other has a key role in probability theory.
Definition We say that in an experiment $\mathcal{E}$, two events A and B are *statistically independent* if,

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

Imagine a long series of trials, each of which involves carrying out two experiments $\mathcal{E}_1$ and $\mathcal{E}_2$, where only $\mathcal{E}_1$ leads to A_1 and only $\mathcal{E}_2$ leads to A_2.

If n = total number of trials, $n(A_1 \cap A_2)$ = number of trials leading to occurence of A_1 and A_2, then

$$\begin{aligned} P(A_1 \cap A_2) &\sim \frac{n(A_1 \cap A_2)}{n} \\ P(A_2) &\sim \frac{n(A_2)}{n} \\ P(A_1) &\sim \frac{n(A_1)}{n}. \end{aligned}$$

On the other hand

$$\begin{aligned} P(A_1 \cap A_2) &\sim \frac{n(A_1 \cap A_2)}{n} \\ &= \frac{n(A_1 \cap A_2)}{n(A_2)} \cdot \frac{n(A_2)}{n} \\ &\sim P(A_1) \cdot P(A_2) \end{aligned}$$

The following example illustrates statistical independence and related subtleties. Throw two dice resulting in the outcomes (X, Y).
Let A_1 : event that X is odd
A_2 : event that Y is odd
A_3 : event that $X + Y$ is odd.

Clearly, A_1 and A_2 are independent.

$$\begin{aligned} P(A_1) &= \tfrac{1}{2} \qquad\qquad\qquad\qquad = P(A_2) \\ P(A_3) &= \text{Prob }\{\text{X odd and Y even}\} \\ &\quad + \text{Prob }\{\text{X even and Y odd}\} \\ &= \tfrac{1}{2}\cdot\tfrac{1}{2} + \tfrac{1}{2}\cdot\tfrac{1}{2} \\ &= \tfrac{1}{2} \end{aligned}$$

$$\begin{aligned} P(A_3|A_1) &= \text{Prob }\{Y \text{ even}\} \\ &= \tfrac{1}{2} \\ P(A_3|A_2) &= \text{Prob }\{X \text{ even}\} \\ &= \tfrac{1}{2} \\ \Rightarrow \quad P(A_3|A_1) &= P(A_3) \qquad = P(A_3|A_2) \end{aligned}$$

Thus A_3 and A_1 are independent *and* A_3 and A_2 are independent. □

Definition: Given events $A_1, A_2, \ldots, A_n$, we say these are *mutually independent* if:

$$\begin{aligned} P(A_i \cap A_j) &= P(A_i) \cdot P(A_j) \\ P(A_i \cap A_j \cap A_k) &= P(A_i)P(A_j)P(A_k) \\ &\vdots \\ P(A_1 \cap A_2 \cap \cdots \cap A_n) &= P(A_1)P(A_2)\cdots P(A_n). \end{aligned}$$

In the previous example, the events A_1, A_2, A_3 are *not* mutually independent, even though they are pairwise independent, because

$$P(A_1 \cap A_2 \cap A_3) = 0$$

but

$$P(A_1)P(A_2)P(A_3) = \left(\tfrac{1}{2}\right)^3.$$

Probability Trees: When an experiment is of a *sequential nature*, it is often convenient, especially for purposes of calculation, to represent the experiment graphically by a *probability tree*. It is a *rooted tree* and the vertices represent *outcomes/events* of the experiment. The edges are labelled by the *conditional probabilities* required to *descend* from a given vertex to an adjacent one. The probability associated with the event corresponding to a vertex is obtained by taking under consideration the product of the probabilities labelling the edges forming the unique path between the vertex, and the root of the tree.

Example: Flipping a coin three times:

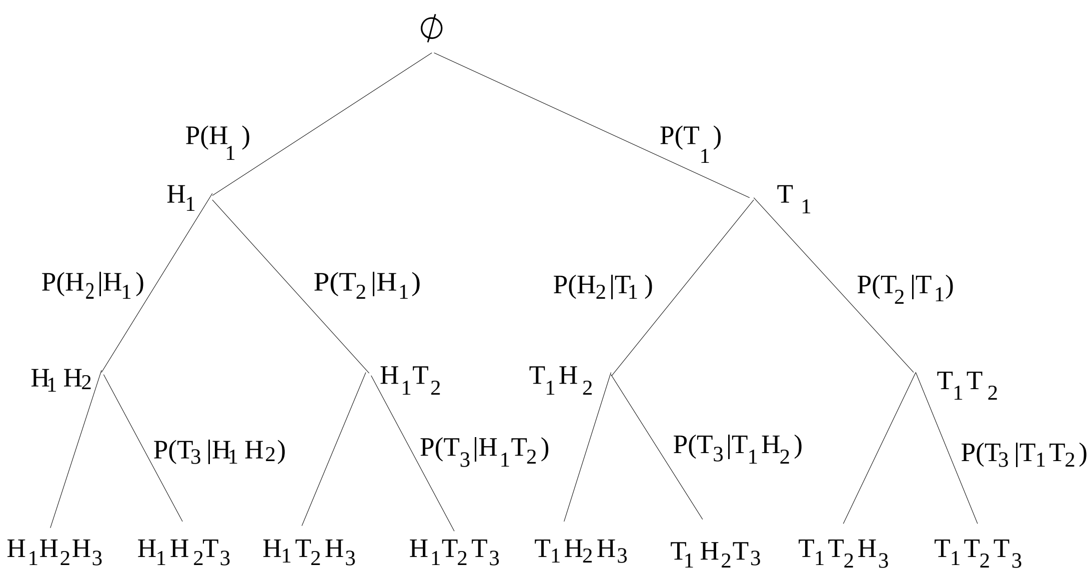

fig. 1

Corresponding Pascal's Triangle

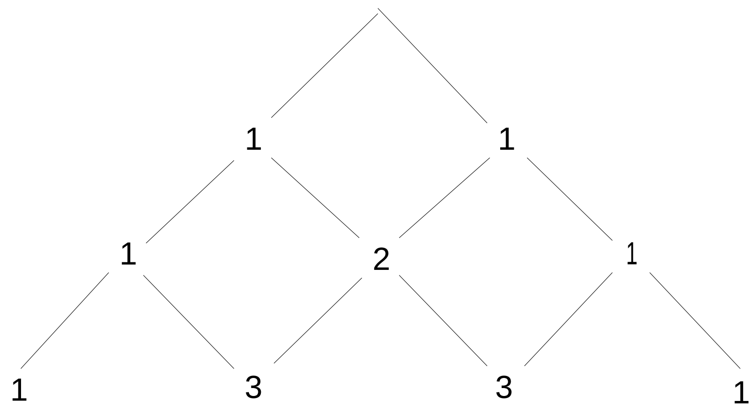

fig. 2

Probability trees may also be infinite. We give an example below.

Example: Player A flips a fair coin. If the outcome is a head, he wins; if the outcome is a tail, player B flips. If B's flip is a head, he wins; if not, player A flips the coin again. This process is repeated (*ad infinitum*, if necessary) until somebody wins. What is the probability that A wins?

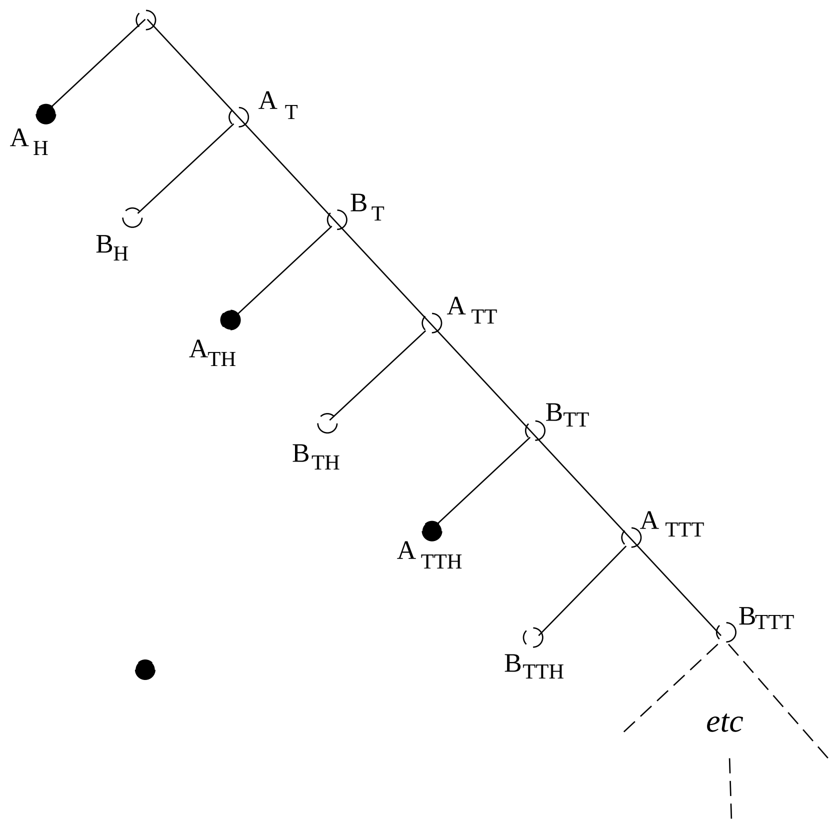

fig. 3

For the probability tree above, the darkened vertices correspond to the *elementary events* for which A wins. Since the probability represented by each branch of the tree is 1/2, we have:

$P\{A \text{ wins}\}$ calculated via sampling with replacement

$$
\begin{aligned}
&= P\{A_H\} + P\{A_{TH}\} + P\{A_{TTH}\} + \cdots \\
&= \frac{1}{2} + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^5 + \cdots \\
&= \frac{1}{2}\left[1 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^6 + \cdots\right] \\
&= \frac{1}{2}\,\frac{1}{1-\left(\frac{1}{2}\right)^2} \\
&= \frac{1}{2}\,\frac{1}{1-1/4} \\
&= \frac{2}{3}
\end{aligned}
$$

There is a *big advantage for A* to flip first.

Gambler's Ruin – (Application of Total Probability Law)

Example: (1) Toss coin. Call correctly, win 1 dollar. Call wrongly, loose 1 dollar.

Payoff Matrix

Call \ Toss	Head	Tail
Head	1	-1
Tail	-1	1

Fig. 4

Initial Capital $= x$ dollars and x is a positive integer.

STRATEGY — PLAY UNTIL EITHER :

- Win m Dollars (*i.e.* has a total of m dollars)
- Lose Shirt (**RUIN**)

Question: What is the probability $p(x)$ of ruin?

$$\begin{aligned} A &= \text{RUIN} \\ B_1 &= \text{Win first call} = p \\ B_2 &= \text{Lose first call} = (1-p) \end{aligned}$$

$$\begin{aligned} P(A) &= P(A|B_1)\,P(B_1) + P(A|B_2)\cdot P(B_2) \\ p(x) &= p(x+1)\cdot\tfrac{1}{2} + p(x-1)\,\tfrac{1}{2} \qquad 1 \le x \le m-1 \\ &= p(x+1)\,p + p(x-1)\cdot(1-p) \end{aligned}$$

$$\text{B.C.}\begin{cases} p(0) = 1 \\ p(m) = 0 \end{cases}$$

$$p(x) = C_1 + C_2 x \qquad \text{is the solution}$$
$$C_1 = 1 \qquad C_1 + C_2 m = 0$$

Hence:

$$\boxed{p(x) = 1 - x/m} \qquad 0 \le x \le m$$

> If $p \neq 1/2$ the solution is *not* linear

Example [Matching]:

n distinct items to be matched against n distinct cells. What is the probability of at least 1 match?

Solution:

$A_k :=$ event that k^{th} item is matched (we don't care about the rest)
$P^{(n)} =$ Probability of at least 1 match

$$
\begin{aligned}
&= P(\cup_{k=1}^{n} A_k) \\
&= \sum_{i=1}^{n} P(A_i) - \sum_{i<j=2}^{n} P(A_i \cap A_j) \\
&\qquad + \sum_{i<j<k=3}^{n} P(A_i \cap A_j \cap A_k \cdots) \\
&\qquad +(-1)^{n+1} P(A_1 \cap A_2 \cap \cdots A_n) \\
&= P_1 - P_2 + P_3 \cdots \pm P_n
\end{aligned}
$$

$$
P(A_{i_1} \cap A_{i_2} \cdots \cap A_{i_m}) = \frac{(n-m)!}{n!}
$$

$$
P_m = \sum_{a \le i_1 < i_2 < \cdots < i_m \le n} P(A_{i_1} \cap A_{i_2} \cdots \cap A_{i_m}) = \binom{n}{m} \frac{(n-m)!}{n!}
$$

$$
= \frac{n!}{(h-m)!m!} \frac{(n-m)!}{n!} = \frac{1}{m!}
$$

$$
P^{(n)} = 1 - \tfrac{1}{2!} + \tfrac{1}{3!} - \tfrac{1}{4!} \cdots + (-1)^{n+1} \tfrac{1}{n!}
$$

Special Cases: Number of permutations of n things in which there is at least 1 match $= P^{(n)} \cdot n!$.

$$
\begin{aligned}
n = 3 \qquad P^{(n)} n! &= 6 \times \left(1 - \tfrac{1}{2} + \tfrac{1}{6}\right) \\
&= \underline{\underline{4}} \\
n = 4 \qquad P^{(n)} n! &= 24 \left(1 - \tfrac{1}{2} + \tfrac{1}{6} - \tfrac{1}{24}\right) \\
&= \underline{\underline{15}}
\end{aligned}
$$

Problem: Given any n events, $A_1,\ A_2, \cdots A_n$ prove that the probability of exactly $m \le n$ events occurring is

$$
P = P_m - \binom{m+1}{m} P_{m+1} + \binom{m+2}{m} P_{m+2} \cdots \ \pm \binom{n}{m} P_n
$$

where

$$
P_k = \sum_{1 \le i_1 < i_2} , \cdots i_k \le n P\left(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\right)
$$

Good Example of Bayesian Inference

Sometimes the application of Bayes' theorem may yield results that appear counter-intuitive.

Example: A laboratory test is developed to detect *mononucleosis* (mono, for short). The probability that a person selected at random has mono is 0.005. If a person has mono, 95% of the time he test will be positive. If a person does not have mono, the test will be positive only 4% of the time. These circumstances are described by the *binary channel* shown in *Figure 5*.

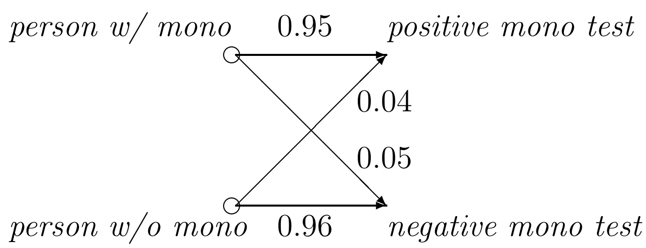

Fig. 5

What is the probability that a person has mono conditioned on the fact that his test came out positive?

$$M = \text{person has mono}$$
$$T = \text{positive mono test}$$

prior probabilities

$$\begin{cases} P(M) = 0.005 \\ P(\bar{M}) = 0.995 \end{cases}$$

conditional probabilities

$$\begin{cases} P(T|M) = 0.95 \\ P(T|\bar{M}) = 0.04 \end{cases}$$

Then, by Bayes' theorems,

$$\begin{aligned}
\text{a posteriori probability} & \\
\{P(M|T) &= \frac{P(T|M)P(M)}{P(T|M)P(M) + P(T|\bar{M})P(\bar{M})} \\
&= \frac{0.95 \times 0.005}{0.95 \times 0.005 + 0.04 \times 0.995} \\
&= \frac{0.00475}{0.00475 + 0.0398} \\
&= \frac{0.00475}{0.04455} \\
&= \underline{\underline{0.107}} \qquad !
\end{aligned}$$

Thus the test might give rise to too many false alarms. How to improve? Bring *down the probability* $P(T|\bar{M})$ from 0.04. Improve the test.

A useful form of Bayes' theorem is obtained by conditioning in more than one event.

Let H := hypothesis (e.g. *a disease* event),
Let E := evidence of data (e.g. *image data* event), and
Let C := context (e.g. *age group*). Then,

$$P(H|E \cap C) = \frac{P(E|H \cap C) \cdot P(H|C)}{P(E|C)}$$

To see this, observe that the r.h.s. above

$$\begin{aligned}
&= \frac{P(E \cap H \cap C)}{P(H \cap C)} \cdot \frac{P(H \cap C)}{P(C)} \cdot \frac{P(C)}{P(E \cap C)} \\
&= \frac{P(E \cap H \cap C)}{P(E \cap C)} \\
&= \frac{P(H \cap (E \cap C))}{P(E \cap C)}
\end{aligned}$$

$$
\begin{aligned}
&= P(H|E \cap C) \\
&= l.h.s
\end{aligned}
$$