Review: Basics

- Sample space S: the set of all possible outcomes of an experiment.
- Events:
  - Event E: a subset of sample space S, $E \subset S$. A set of possible outcomes when an experiment is performed.
  - If E and F are events then so are: $E \cup F$, $E \cap F$, E^c, F^c
- Axioms:
  - $0 \leq P(E) \leq 1$
  - $P(S) = 1$
  - If E and F are mutually exclusive events ($E \cap F = \emptyset$) then $P(E \cup F) = P(E) + P(F)$. Otherwise $P(E \cup F) = P(E) + P(F) - P(E \cap F)$.
- $P(E^c) = 1 - P(E)$
- $E \subset F$ implies that $P(E) \leq P(F)$
- If all outcomes equally likely then $P(E) = \frac{\text{Number of outcomes in E}}{\text{Number of outcomes in S}} = \frac{|E|}{|S|}$

Example

Roll a 6-sided die:

- Sample space $S = \{1, 2, 3, 4, 5, 6\}$
- Event that comes up a 6 is $E = \{6\}$.
- $E = \{6\}$ and $F = \{2, 6\}$ are events and so are $E \cup F = \{2, 6\}$, $E \cap F = \emptyset$, $E^c = \{1, 2, 3, 4, 5\}$, $F^c = \{1, 3, 4, 5\}$
- $P(S) = 1$, $P(\emptyset) = 0$.
- $P(E \cup F) = P(\{2, 6\}) = P(\{2\}) + P(\{6\})$
- $P(\{1, 2, 3, 4, 5\}) = 1 - P(\{6\})$
- $\{6\} \subset \{2, 6\}$ implies that $P(\{6\}) \leq P(\{2, 6\})$
- All outcomes are equally likely so $P(\{6\}) = \frac{1}{6}$, $P(\{2, 6\}) = \frac{2}{6}$

Conditional Probability

- Deinition: $P(E|F) = \frac{P(E \cap F)}{P(F)}$ when $P(F) > 0$.
- $P(E|F)$ is a probability (it satisfies all the axioms) with sample space $S \cap F$, event $E \cap F$.
- Example:

1st toss	2nd toss
H	H
H	T
T	H
T	T

$$P(\text{2nd toss is H }|\text{1st toss is T}) = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$

$$P(\text{1st toss is T}|\text{got at least one T}) = \frac{\frac{2}{4}}{\frac{3}{4}} = \frac{2}{3}$$

Bayes Theorem

$$P(E|F) = \frac{P(F|E)P(E)}{P(F)}$$

posterior *likelihood* *prior*

Suppose the event E is that it rains tomorrow, and F is the event that it is cloudy today.

- Prior. Our guess for the chance of rain tomorrow, with no extra info.
- Likelihood. The probability of a cloudy day before rain.
- Posterior. Our updated probability of rain tomorrow after observing clouds today
- Evidence $P(F)$ is the chance of a cloudy day, with no extra info.

Example

- Suppose two users submit jobs to a server. Jobs can be big or small. For user 1 jobs are big with probability 0.5 and for user 2 with probability 0.1. The fraction of jobs submitted by user 1 is 0.1.
- A big job is submitted. What is the probability that it came from user 1 ?
- Bayes:

$$
\begin{aligned}
P(\text{user } 1|\text{big}) &= \frac{P(\text{big}|\text{user } 1)P(\text{user } 1)}{P(\text{big})} \\
&= \frac{P(\text{big}|\text{user } 1)P(\text{user } 1)}{P(\text{big}|\text{user } 1)P(\text{user } 1) + P(\text{big}|\text{user } 2)P(\text{user } 2)} \\
&= \frac{0.5 \times 0.1}{0.5 \times 0.1 + 0.1 \times (1 - 0.1)} \approx 0.36
\end{aligned}
$$

Independence

- Two events E and F are **independent** if $P(E \cap F) = P(E)P(F)$
- Two events E and F are called **conditionally independent given G** if $P(E \cap F|G) = P(E|G)P(F|G)$

Making Predictions

Suppose have prior $P(E)$. Then observe event F. And now calculate posterior probability $P(E|F)$. Basket analysis example:

- From past data observe that 60% of baskets containing beer also contain ice cream and 20% of baskets without beer contain ice cream. We also observe that overall 10% of baskets contain beer.
- This is based on a large number of baskets, so we assume probability $P(E|F) = 0.6$ and $P(E|F^c) = 0.2$, where F is the event that the basket contains beer and E the event that it contains ice cream. Also that $P(F) = 0.1$.
- We observe a customer putting beer in their basket. What's the probability that they will add ice cream ?

$$P(F|E) = \frac{P(E|F)P(F)}{P(E)} = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^c)(1 - P(F))}$$

$$\approx \frac{0.6 \times 0.1}{0.6 \times 0.1 + 0.2 \times (1 - 0.1)} = 0.25$$

Somewhat higher than $P(F) = 0.1$, so worth offering a discount ?

Monty Hall[1]

Three doors.

- Behind one door is a prize (equally likely to be any door)
- Behind the other two is nothing
- Choose one door.
- The host opens one of the two remaining doors, always revealing nothing
- We have the option to change to the other remaining door, should we ?
- Note: if we don't switch the probability of winning is $\frac{1}{3}$.

[1]See https://en.wikipedia.org/wiki/Monty_Hall_problem

Monty Hall

Demo

Monty Hall

Possible outcomes:

- Door we picked is winner, this is with prob $\frac{1}{3}$. Remaining door loses.
- Door we picked is loser, this is with prob $\frac{2}{3}$. Remaining door wins.

We should always switch.

Prosecutors Fallacy

- Probability that two DNA profiles match by chance is 1 in 10,000.
- DNA found at a crime scene is compared against a database of 20,000 people. A match is found.
- Does this mean that probability innocent is 1 in 10,000 ?
- Let G be the event that person is guilty. Bayes to the rescue:

$$P(G|\text{match}) = \frac{P(\text{match}|G)P(G)}{P(\text{match})}$$
$$= \frac{P(\text{match}|G)P(G)}{P(\text{match}|G)P(G) + P(\text{match}|G^C)(1 - P(G))}$$

- Suppose $P(\text{match}|G) = 1$. Then:

$$P(G|\text{match}) = \frac{P(G)}{P(G) + P(\text{match}|G^C)(1 - P(G))}$$

Prosecutors Fallacy

$$P(G|\text{match}) = \frac{P(G)}{P(G) + P(\text{match}|G^C)(1 - P(G))}$$

- Probability that one person doesn't have DNA match is $(1 - \frac{1}{10000})$.
- Probability that none out of 20,000 people have a DNA match is $(1 - \frac{1}{10000})^{20000}$ (assuming independence).
- So probability of a match when innocent $P(\text{match}|G^C) = 1 - (1 - \frac{1}{10000})^{20000} \approx 0.86$.
- That is,

$$P(G|\text{match}) \approx \frac{P(G)}{P(G) + 0.86(1 - P(G))}$$

If probability of guilt $P(G)$ is small then $P(G|\text{match}) \approx \frac{P(G)}{0.86}$ is also small.

Prosecutors Fallacy (Again)

- M50 toll bridge cameras correctly recognise car registrations 99% of the time. Probability of error is 0.01.
- 100,000 cars cross the bridge each day.
- You're sent a penalty notice. What is the probability that it is valid ?