lecture 3 : Probability / Conditional Probability (§ 2.1, 2.2, 2.4)

Examples from § 2.1, 2.2

Example 2.12 (Devore)

Consider an experiment of testing batteries coming off an assembly line until one with a voltage within certain prescribed limits is found. If the first battery on the assembly line is found to be within the prescribed voltage limits, then the experiment is considered successful, and the event is $E_1 = \{S\}$. If the first battery was found to have a voltage outside the prescribed range, then we keep going down the assembly line until we find a battery within the limits. For example, $E_2 = \{F, S\}$, $E_3 = \{F, F, S\}$, etc.

Suppose that the probability of any battery being satisfactory is $p \Rightarrow P(S) = p$ and $P(F) = 1-p$. Then,

$$P(E_1) = p$$
$$P(E_2) = (1-p)\cdot p$$
$$P(E_3) = (1-p)^2 \cdot p$$
$$\vdots$$
$$P(E_n) = (1-p)^{n-1} \cdot p$$

This assignment of probability only makes sense if the events are independent, i.e., if the first battery fails the prescribed criteria, this says nothing about the second, third, or any other battery on the assembly line. This assignment satisfies the axioms, particularly

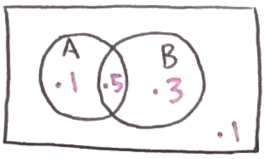
$$P(\Omega) = P(E_1 \cup E_2 \cup E_3 \cup \cdots) \leftarrow \text{suppose there's an infinite number of batteries}$$
$$= P(E_1) + P(E_2) + P(E_3) + \cdots \leftarrow E_1, E_2, \ldots \text{ disjoint}$$
$$= p + p(1-p) + p(1-p)^2 + \cdots$$
$$= p\left[1 + (1-p) + (1-p)^2 + \cdots\right] \leftarrow \text{geometric series } \sum_{n=0}^{\infty} r^n = \frac{1}{1-r} \text{ with } r = 1-p \text{ and } |r| < 1$$
$$= p\left[\frac{1}{1-(1-p)}\right] = p \cdot \frac{1}{p} = 1$$

Example 2.14 (Devore)

Local cable company : 60% of households get internet : event A
80% of households get TV : event B
50% of households get both : event $A \cap B$

A household is selected at random, what is the probability that this household gets at least one of the two services?

$$P(\text{at least one service}) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 0.6 + 0.8 - 0.5$$
$$= 0.9$$



What is the probability that a household selected at random gets exactly one of these services?

$$P(\text{exactly one service}) = P(\text{only internet}) + P(\text{only TV})$$
$$= P(A) - P(A \cap B) + P(B) - P(A \cap B)$$
$$= 0.6 - 0.5 + 0.8 - 0.5$$
$$= 0.4$$

Section 2.2 Problem 14 (Devore)

$P(\text{coffee}) = 0.55$

$P(\text{soda}) = 0.45$

$P(\text{one of the two}) = 0.7 \leftarrow P(\text{coffee} \cup \text{soda})$

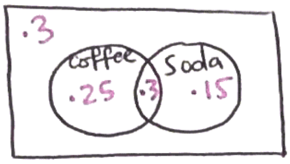
(a) $P(\text{coffee and soda}) = ?$

$$P(\text{coffee and soda}) = P(\text{coffee} \cap \text{soda})$$

$$P(\text{coffee} \cup \text{soda}) = P(\text{coffee}) + P(\text{soda}) - P(\text{coffee} \cap \text{soda})$$

$$\Rightarrow \quad 0.7 = 0.55 + 0.45 - P(\text{coffee} \cap \text{soda})$$

$$\Rightarrow \quad P(\text{coffee} \cap \text{soda}) = 0.3$$



(b) $P(\text{does not consume at least one of the two})$

↳ does not consume coffee, but consumes soda : event A

OR does not consume soda, but consumes coffee : event B

OR does not consume both : event C

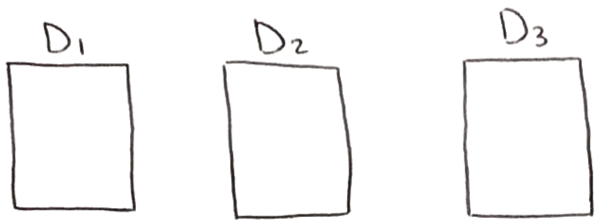
The complement of the above event is someone who consumes both coffee and water

$$P(\text{does not consume at least one of the two}) = P(A) + P(B) + P(C)$$
$$= 0.15 + 0.25 + 0.3$$
$$= 0.7$$

Introduction to conditional Probability (some examples)

* Monty Hall Problem (excerpt from The Drunkard's Walk)

Suppose the contestants on a game show are given the choice of three doors: Behind one door is a car; behind the others, goats. After a contestant picks a door, the host, who knows what's behind all the doors, opens one of the unchosen doors, which reveals a goat. He then says to the contestant, "Do you want to switch to the other unopened door?" Is it to the contestant's advantage to make the switch?



Suppose the contestant has chosen D_1.
In this case, the host will open either D_2 or D_3, but the host does not choose randomly. Here are the scenarios:

* Scenario 1 (Lucky Guess scenario):
In this case, your initial guess is correct. The host will now randomly open D_2 or D_3, and if you choose to switch, you'll be stuck with a goat! In the Lucky Guess Scenario, you are better off not switching - but the probability of landing in the Lucky Guess scenario is $1/3$.

* scenario 2 (Wrong Guess scenario)
In this case, your initial choice was wrong. The chances you guessed wrong are 2 out of 3, so the Wrong Guess scenario is twice as likely to occur as the Lucky Guess scenario. Unlike the Lucky Guess scenario, in this case the host does not randomly open an unchosen door. Since he does not want

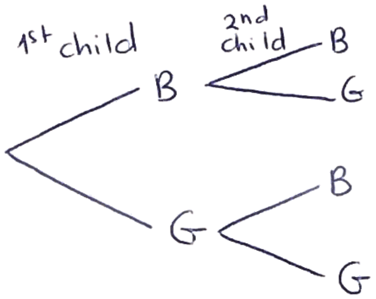
to reveal the car, he chooses to open precisely the door that does not have the car behind it. In other words, in the Wrong Guess scenario the host intervenes in what until now has been a random process. So, the process is no longer random.

To summarize: if you are in the Lucky Guess scenario (probability 1 in 3), you'll win if you stick to your choice. If you are in the Wrong Guess scenario (probability 2 in 3), you'll win if you switch your choice.

Two Daughter Problem (excerpt from The Drunkard's Walk)

In a family with two children, if one of the children is a girl, what are the chances that both children are girls?

→ If the "if clause" (underlined) was not present, the answer would be 1 in 4, because in a family with two children, the sample space is



$$\Rightarrow \Omega = \{BB, BG, GB, GG\}$$

$$P(GG) = 1/4$$

If we know that one of the two children is a girl, the the "new" sample space (pruned sample space) is: $\{BG, GB, GG\}$, and therefore, the chances that both children are girls is 1 in 3.

A variant of this problem is this: in a family with two children, what are the chances, if one of the children is a girl named Florida, that both children are girls?

In this problem, our information concerns not just the gender of the children, but also, for the girls, the name. Since our original sample space should be a list of all the possibilities, in this case it is a list of both gender and name. Denote "girl named Florida" by girl-F and "girl not named Florida" by girl-NF.

Then, the sample space is:
(boy, boy), (boy, girl-F), (boy, girl-NF), (girl-F, boy)
(girl-NF, boy), (girl-NF, girl-F), (girl-F, girl-NF),
(girl-NF, girl-NF), and (girl-F, girl-F)

Pruning the sample space so that one of the children is girl-F, our new sample space becomes:
(boy, girl-F), (girl-F, boy), (girl-NF, girl-F), (girl-F, girl-NF),
and (girl-F, girl-F)

we can neglect this possibility because
1. parents tend to shy away from giving their kids the same name.
2. probability of naming a girl Florida is like 1 in a million, so naming the second Florida is very unlikely

Since 2 out of 4 elements in the pruned sample space are families with two girls (circled), the answer is ½.