4 Conditional Probability

Probabilities do not live in a vacuum. They are determined by information available, for instance how an experiment is carried out. Some times additional information is available that changes the probability.

$P(B|A)$ is the conditional probability of B given the knowledge that event A occurs.

Example 2.22 400 items are classified by surface flaws and as defective.

		Surface Flaws		
		Yes (event F)	No	Total
Defective	Yes (event D)	10	18	28
	No	30	342	372
	Total	40	360	400

Notice the following,

$$P(D) = \frac{28}{400}$$

$$P(D|F) = \frac{10}{400}$$

Thus we see how the additional knowledge that another event can change the probability of other related events. In other words, events D and F are dependent: knowledge of F occurring changes the probability of D.

There is a general formula for conditional probability

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \tag{1}$$

We can see how this formula is derived

$$P(B|A) = \frac{\# \text{ outcomes in } B \text{ among outcomes in } A}{\# \text{ of outcomes in} A}$$
$$= \frac{|A \cap B|}{|A|}$$
$$= \frac{|A \cap B|/|S|}{|A|/|S|}$$
$$= \frac{P(A \cap B)}{P(A)}$$

Example 2.106 Failures of AC systems have been classified below by 2 events. Let E be the event there is evidence of an electrical failure, and let G be the event there is evidence of a gas failure.

	Yes G	No G	Total
Yes E	55	17	72
No E	32	3	35
	87	20	107

$P(G) = 87/107; P(E|G) = 55/87; P(G|E) = 55/72; P(G \cap E) = 55/107$

Exercise 2.109 There are 350 chips of which 8 are defective. Let D_1 be the event the first chip is defective, and let D_2 be the event the second chip is defective. If we decide to sample 2 without replacement: what is the probability the second chip is defective given the first chip is defective? And what is the probability the first and second chip is defective?

In this example, we highlight the difference of conditional probability and the mutual probability of events.

Ans.

$$P(D_2|D_1) = 7/349$$

In this answer, we know that we already had one chip pulled out from the 350 chips which was defective. This means there are only 349 chips remaining, and there are only 7 defective chips to pull from. We could also perform this calculation the long way

$$P(D_2|D_1) = \frac{P(D_2 \cap D_1)}{P(D_1)} = \frac{\binom{8}{2}/\binom{350}{2}}{\binom{8}{1}/\binom{350}{1}} = 7/349$$

Now we can ask what is the probability of D_1 and D_2

$$P(D_1 \cap D_2) = \frac{8}{350} \cdot \frac{7}{349} = \binom{8}{2}/\binom{350}{2}$$

Notice the answers are different, conditional probability is not mutual probability!

4.1 Multiplication Rules

We can manipulate equation (1) to get one of the most useful equations in probability

$$P(A \cap B) = P(B|A)P(B) \ |OR| \ P(A \cap B) = P(A|B)P(B)$$

Example 2.28 Semi Conductor manufacturing failures and product failures are based on levels of contamination in the factory. Below outlines the probability of failure for each level of contamination.

Probability of Failure $P(F)$	Level of Contamination
0.10	High H
0.01	Medium M
0.001	Low L

Given the probability of high level contamination to be $P(H) = 0.20$, probability of medium level contamination to be $P(M) = 0.30$, and low level contamination to be $P(L) = 0.50$. Find the probability that a chip randomly selected from the production line will fail.

Ans. The probability of failure $P(F)$ is given as

$$P(F) = P(F \cap H) + P(F \cap M) + P(F \cap L)$$

$$= (0.10)(0.20) + (0.30)(0.01) + (0.001)(0.50)$$

$$= 0.0235$$

The next page outlines a tree structure to see how these probabilities are laid out.

Contamination

High 0.20 Medium 0.30 Low 0.50

P(Fail|High) = 0.10 P(Not Fail|High) = 0.90 P(Fail|Medium) = 0.01 P(Not Fail|Medium) = 0.99 P(Fail|Low) = 0.001 P(Not Fail|Low) = 0.999

0.10(0.20) = 0.02   0.90(0.20) = 0.18   0.01(0.30) = 0.003   0.99(0.30) = 0.297   0.001(0.50) = 0.0005   0.999(0.50) = 0.4995

P(Fail) = 0.02 + 0.003 + 0.0005 = 0.0235