The Rules for Reasoning with Probabilities

1. **Restricted Conjunction (or "AND") Rule**: P(A and B) = P(A) x P(B)
- used when A and B are **independent** outcomes

2. **General Conjunction (or "AND") Rule**: P(A and B) = P(A) x P(B/A)

3. **Restricted Disjunction (or "OR") Rule**: P(A or B) = P(A) + P(B)
- used when A and B are **mutually exclusive** outcomes

4. **General Disjunction (or "OR") Rule**: P(A or B) = P(A) + P(B) – P(A and B)

5. **Negation Rule**: P(not-A) = 1 – P(A)

6. **Conditional Probability Rule**: $P(A/B) = \frac{P(A \text{ and } B)}{P(B)}$

7. **Bayes' Rule**: $P(A/B) = \frac{P(A) \times P(B/A)}{P(A) \times P(B/A) + P(\text{not-}A) \times P(B/\text{not-}A)}$
- just a different version of rule 6.

Some Worked Examples

1. In a box there are 3 red pens, 5 blue pens, and 2 black pens. If a person selects a pen at random, what is the probability that the pen is

a. **a blue or a red pen?**
b. **a red or a black pen?**

These are *mutually exclusive* events so we can use the restricted disjunction ("OR") rule and simply add the probabilities:

a) P(blue or red) = P(blue) + P(red) = 5/10 + 3/10 = 8/10 = 4/5

b) P(red or black) = P(red) + P(black) = 3/10 + 2/10 = 5/10 = 1/2

2. At a political rally there are 8 Democrats and 10 Republicans. 6 of the Democrats are females and 5 of the Republicans are females. If a person is selected at random, what is the probability that the person is either a female or a Democrat?

Being a female and being a Democrat are not mutually exclusive events, so we need to use the general disjunction ("OR") rule for this problem.

Let P(F) = probability that the person is female
Let P(D) = probability that the person is a Democrat

The general rule: P(F or D) = P(F) + P(D) – P(F and D)

P(F) = (6 + 5)/18 = 11/18 [11 of the 18 people at the rally are female]

P(D) = 8/18 [8 of the 18 people at the rally are Democrats]

P(F and D) = 6/18 [6 of the Democrats are females]

Thus,

P(F and D) = 11/18 + 8/18 – 6/18 = 13/18

3. What is the probability of tossing a coin ten times and getting ten heads in row?

Coin tosses are independent events, so we can use the restricted conjunction ("AND") rule.

P(H and H and H and H and H and H and H and H and H and H)
= P(H) x P(H) x P(H) x P(H) x P(H) x P(H) x P(H) x P(H) x P(H) x P(H)
= $(1/2)^{10}$
= 9.7×10^{-4}

You should expect to roll ten heads in a row about once every thousand trials.

4. What is the probability of getting *this particular* sequence -- H, H, T, H, T, T, H, T, H, T -- in ten coin tosses?

Exactly the same: 9.7×10^{-4}. Any particular sequence of heads or tails of a given length is no more or less likely to occur than any other sequence of the same length. Though most people are inclined to assume that **H, H, H, H, H, H, H, H, H, H** is a less probable sequence than **H, H, T, H, T, T, H, T, H, T** (which seems like a more "typical" sequence), this is a fallacy.

5. What is the probability of drawing *two* aces from a regular deck of cards in *two* draws:

a) if the first card is replaced before the second is drawn?
b) if the first card is NOT replaced before the second is drawn?

We've got 52 cards, 4 of which are aces. So $P(Ace_1) = 4/52$ on the first draw. If we replace the card then we still have all 4 aces and 52 cards, so the probability of drawing an ace on the second draw, $P(Ace_2)$, is the same, i.e. these are *independent* events. So

$P(Ace_1 \text{ AND } Ace_2) = P(Ace_1) \times P(Ace_2) = 4/52 \times 4/52 = 16/2704 = 1/169$, which is 0.0059.

So almost six times in a thousand. Not very likely.

Now, if the first card is NOT replaced, the situation is a bit different. We draw an ace on the first draw; we know that $P(Ace_1)$ is 4/52. What are the chances of drawing an ace on the next draw, *if we don't put the first ace back*? This *isn't an independent event anymore*; drawing an ace on the first draw *alters* the probability of drawing an ace on the second draw, if we don't put the ace back. We have to use the general conjunction rule:

$P(Ace_1 \text{ AND } Ace_2) = P(Ace_1) \times P(Ace_2/Ace_1)$

Here, $P(Ace_2/Ace_1)$ reads "the probability of drawing on ace on the second draw, *given that* you've drawn an ace on the first draw". So what is this probability? Well, there are only 3 aces left in the deck, and the deck now has only 51 cards. So $P(Ace_2/Ace_1) = 3/51$.

Thus, $P(Ace_1 \text{ AND } Ace_2) = P(Ace_1) \times P(Ace_2/Ace_1) = 4/52 \times 3/51 = 1/221$ which is 0.0045.

So less than 5 times in a thousand. Not replacing the card makes this event slightly less probable event than if you replaced it.

6. Imagine two urns, each containing red and green balls. Urn A has 80% red balls and 20% green balls. Urn B has 40% red and 60% green. You pick an urn at random by tossing a coin, and then draw a ball from the urn you chose (remember, you don't know which urn you've drawn from). The aim is to guess which urn you've drawn from, based on the color of the ball.

You draw a red ball. What is the probability that you've drawn the ball from urn A?

This is a typical conditional probability argument. Let A = the ball was drawn from urn A, B = the ball was drawn from urn B, R = the drawn ball is red, G = the drawn ball is green. We want to know the probability of drawing from urn A, given that the ball is red, i.e. P(A/R).

It's most convenient to use Bayes' rule for problems like this:

$$P(A/R) = \frac{P(A) \times P(R/A)}{P(A) \times P(R/A) + P(B) \times P(R/B)}$$

We're given all the information we need to fill in these numbers.

What is P(A)? It's just the unconditional probability of having drawn the ball from urn A. But the urn was chosen at random, so P(A) = P(B) = 0.5.

What is P(R/A)? It's the probability of drawing a red ball, given that it's from urn A. But we know this, it's 80%, or 0.8. And we know that P(R/B) = 40%, or 0.4. So we have

$$P(A/R) = \frac{(0.5 \times 0.8)}{(0.5 \times 0.8) + (0.5 \times 0.4)}$$

$$= \frac{0.4}{0.4 + 0.2}$$

$$= 0.4/0.6$$

$$= 2/3$$

There's a 66% chance that you've drawn the red ball from urn A.

7. Taxi Hit and Run

A town has two taxi companies, Blue Cab and Yellow Cab, named after the color the taxis are painted. 95% of the taxis in town are owned by Yellow Cab; Blue Cab owns the rest.

A taxi dents a parked car on the street and speeds away. A witness reports to the police that the taxi was blue.

Psychological studies have shown that witnesses to a single-car accident are 80% likely to be able to report the color of the car correctly.

***Question*: Given the eye witness report, and assuming that this witness is typical, *how likely is it that a blue taxi was really the culprit*?**

We might be inclined to think the odds of it being a blue taxi are close to 80%, since that's the reliability of the eye witness report, but this is a mistake; it ignores the relevant information about the "base rate" frequencies, and the error rate for eye witness reports. But it's easily solved using Bayes' rule.

Let B = a blue taxi hit the car
Let Y = a yellow taxi hit the car
Let E_B = an eye witness reports the color of the car as blue

We want to know $P(B/E_B)$. Bayes' rule gives us

$$P(B/E_B) = \frac{P(B) \times P(E_B/B)}{P(B) \times P(E_B/B) + P(Y) \times P(E_B/Y)}$$

The "base rate" frequencies are
$P(B) = 0.05$
$P(Y) = 0.95$

Now, if witnesses are only able to identify the correct color 80% of the time, then

$$P(E_B/B) = P(E_Y/Y) = 0.8$$

but then we can infer that the *error* rate is

$$P(E_B/Y) = P(E_Y/B) = 0.2$$

Filling in these numbers, we get

$$P(B/E_B) = \frac{(0.05)(0.8)}{(0.05)(0.8) + (0.95)(0.2)}$$

$$= \frac{0.04}{0.04 + 0.19}$$

$$= 0.04/0.23$$

$$= 0.17$$

Thus, even with the eye witness report, there's only a 17% chance that the culprit was a blue taxi.