Chapter 3: Probability

1. (a) This problem asks you to derive the probability of the union of two events, as implied by the fact that the question uses the word "or." The first event A is "the randomly drawn member of Congress is a Republican" and the second event B is "the randomly drawn member of Congress is a Senator." The probability of the union of two events is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

There are 533 total members of Congress, and 233+46=279 Republicans in Congress, so the probability of A is

$$P(A) = \frac{279}{533} = .523.$$

There are 100 Senators, so the probability of B is

$$P(B) = \frac{100}{533} = .188.$$

The intersection of A and B refers to the number of Republican Senators, of which there are 46. The probability of the intersection is

$$P(A \cap B) = \frac{46}{533} = .086.$$

So the probability that a randomly drawn member of Congress is a Republican or a Senator is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = .523 + .188 - .086 = .625.$$

(b) The question asks how many ways are there to choose a group of 13 out of a larger group of 433. Order doesn't matter because if we're not considering rank and chairmanship all committee members have equal standing. That means we are to find the following combination:

$$\binom{433}{13} = \frac{433!}{13!(433-13)!} = \frac{433!}{13!420!}$$

$$= \frac{433 \times 432 \times 431 \times 430 \times 429 \times 428 \times 427 \times 426 \times 425 \times 424 \times 423 \times 422 \times 421}{13 \times 12 \times 11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}$$

$$\approx 2,518,055,000,000,000,000,000,000 \ (2.5 \ \textit{Septillion}!)$$

2. (a) A student whose parents belong to a country club probably comes from a family with some wealth. Wealthier families can afford private SAT lessons for their children. Therefore these events are conditionally independent, and are related only insofar as they both depend on the wealth of the student's family.

(b) These events are probably dependent since better schools and libraries would directly cause a lower rate of illiteracy.

(c) Two mutually exclusive events share no outcomes. That means that if one event occurs, then the other cannot occur. Therefore the two events are dependent.

This is a tricky question. Be sure not to confuse mutually exclusive with independent. An example of two mutually exclusive events might be $A =$"a person votes for the Democrat" and $B =$"a person votes for the Republican." These events are not independent because if a person votes for the Democrat he or she by definition does not vote for the Republican. In contrast, if two events are independent, then whether one occurs or not has no bearing on whether the other will occur or not.

3. There are many, many examples. One example is
$A =$"there are more than 200 visitors to Monticello (Thomas Jefferson's home, now a historical museum near Charlottesville, VA) today", and
$B =$"the Japanese stock market rises 100 points."
These events are almost certainly independent. Unless, of course, one of the visitors to Monticello is a Japanese stock broker on vacation who realizes that the elderberry wine produced in central Virginia is delicious and places an unprecedented order for elderberries on the Japanese commodities exchange and sets off a chain of events that results in the recovery of the Japanese and Virginian economies and the invention of Bluegrass Kabuki theater.

4. By the definition of a combination we know that

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

If we plug $n-k$ in for k in the above equation, we get

$$\binom{n}{n-k} = \frac{n!}{(n-k)!k!},$$

which is the same thing as before. Therefore

$$\binom{n}{k} = \binom{n}{n-k}.$$

5. (a) The event A is "at least 2 people in a class of 30 share a birthday," so the complement $\tilde{A}$ is "no one in a class of 30 shares a birthday with anyone else in the class." The sample space S consists of "every way the 30 people in the class can have a birthday."

(b) There are 365 days in a year. Consider the 30 students as if they are lined up. The first student can have any of 365 birthdays. The second student cannot share a birthday with the first student, but can have any of the 364 remaining possibilities. Likewise the 3rd student has 363 possibilities, and so on, until the 30th student has 336 possibilities. All together, multiplying these stages together, there are

$$365 \times 364 \times 363 \times \ldots \times 336 \approx 2.17 \times 10^{76} \text{ possibilities.}$$

An alternative way to think about this problem is that we are trying to choose 30 birthdays out of the set of 365 possible birthdays. Order matters here since if two students switch their birthdays, that's a different outcome. We use a permutation:

$${}_{365}P_{30} = \frac{365!}{(365-30)!} = 365 \times 364 \times 363 \times \ldots \times 336 \approx 2.17 \times 10^{76} \text{ possibilities.}$$

(c) For the sample space, the event with all possible outcomes, we don't care whether two students share a birthday or not. Each student has 365 possible birthdays. So the 30 students have

$$365^{30} \approx 7.4 \times 10^{76}.$$

possible combinations of birthdays.

(d) The probability of event A, "at least 2 people in a class of 30 share a birthday," is

$$P(A) = 1 - P(\tilde{A}) = 1 - \frac{|\tilde{A}|}{|S|}.$$

In part (b) we derived

$$|\tilde{A}| = 365 \times 364 \times 363 \times \ldots \times 336 \approx 2.17 \times 10^{76},$$

and in part (c) we derived

$$|S| = 365^{30} \approx 7.4 \times 10^{76}.$$

So the probability of A is

$$1 - \frac{365 \times 364 \times 363 \times \ldots \times 336}{365^{30}} = 0.706,$$

which is a surprisingly high result.

6. (a) A TRUE response (X) happens when the respondent has taken a bribe (W) **AND** has spun A (Y), **OR** has never taken a bribe ($\tilde{W}$) **AND** has spun a B ($\tilde{Y}$).

If we replace the events with their symbols, the word AND with intersections, and the word OR with a union, the statement becomes

$$X = (W \cap Y) \cup (\tilde{W} \cap \tilde{Y}).$$

(b) The events W and Y are independent because the probability of taking a bribe does not affect the probability of spinning an A or B, and the probability of spinning A or B does not affect the probability of taking a bribe.

The events X and Y are NOT independent because the probability of stating TRUE depends on the probability of spinning A or B.

(c) We start with our answer in part (a):

$$X = (W \cap Y) \cup (\tilde{W} \cap \tilde{Y}).$$

Then the probability of X is

$$P(X) = P\Big((W \cap Y) \cup (\tilde{W} \cap \tilde{Y})\Big).$$

If we apply the formula for the probability of a union, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, we get

$$P(X) = P(W \cap Y) + P(\tilde{W} \cap \tilde{Y}) - P\Big((W \cap Y) \cap (\tilde{W} \cap \tilde{Y})\Big).$$

First consider the last term. The event $W \cap Y$ represents officials who've taken a bribe and spun A. The event $\tilde{W} \cap \tilde{Y}$ represents officials who've never taken a bribe and who've spun B. These two events are mutually exclusive because they share no elements. That means that their intersection is the empty set, and the probability of the empty set is 0:

$$P\Big((W \cap Y) \cap (\tilde{W} \cap \tilde{Y})\Big) = 0.$$

That leaves us with

$$P(X) = P(W \cap Y) + P(\tilde{W} \cap \tilde{Y}).$$

Finally, recall that in part (b) we found that events W and Y are independent. The same logic must apply to $\tilde{W}$ and $\tilde{Y}$. We apply the formula for the probability of the intersection of two independent events, $P(A \cap B) = P(A)P(B)$ and find

$$P(X) = P(W)P(Y) + P(\tilde{W})P(\tilde{Y}).$$

(d) We start with

$$P(X) = P(W)P(Y) + P(\tilde{W})P(\tilde{Y}).$$

We apply the rule for the probability of a complement, $P(\tilde{A}) = 1 - P(A)$, and find

$$P(X) = P(W)P(Y) + \Big(1 - P(W)\Big)\Big(1 - P(Y)\Big).$$

Multiplying this equation out, it becomes

$$P(X) = P(W)P(Y) + \Big(1 - P(W) - P(Y) + P(W)P(Y)\Big),$$

$$P(X) = 2P(W)P(Y) + 1 - P(W) - P(Y).$$

In order to solve for $P(W)$, we bring every term with $P(W)$ over to one side, and bring every term with $P(W)$ to the other side,

$$P(X) + P(Y) - 1 = 2P(W)P(Y) - P(W),$$

we factor out $P(W)$,

$$P(X) + P(Y) - 1 = P(W)\Big(2P(Y) - 1\Big),$$

and divide to isolate $P(W)$:

$$P(W) = \frac{P(X) + P(Y) - 1}{2P(Y) - 1}.$$

(e) If the spinner had equal sized areas for A and B, then the probability of spinning A would be $P(Y) = 0.5$. But plugging this value into

$$P(W) = \frac{P(X) + P(Y) - 1}{2P(Y) - 1}$$

places 0 in the denominator. Therefore it is only possible to solve for $P(W)$ in this way when the spinner has unequal areas for A and B.

(f) If 35% of respondents select TRUE, then $P(X) = 0.35$, and we know that $P(Y) = 0.75$. We simply plug these probabilities into the equation for $P(W)$:

$$P(W) = \frac{P(X) + P(Y) - 1}{2P(Y) - 1},$$

$$P(W) = \frac{0.35 + 0.75 - 1}{2(0.75) - 1} = \frac{0.1}{0.5} = 0.2.$$

So given these results, we conclude that 20% of government officials have taken bribes.

7. Let H be the event in which I have a headache and let F be the event in which I have the flu. The problem tells us that

- $P(H) = .1$, since I have a headache 1 out of 10 days in any event,
- $P(H|F) = .5$ since half of all flu sufferers have headaches,
- and $P(F) = 0.02$ since 2% of the population will come down with the flu.

What we are trying to find is $P(F|H)$, the probability of coming down with the flu given a headache. Bayes' rule tells us that

$$P(F|H) = \frac{P(H|F)P(F)}{P(H)}.$$

So we simply plug in the corresponding values:

$$P(F|H) = \frac{0.5 \times 0.02}{0.1} = 0.1.$$

8. (a) Let N be the event that an email contains the word "Nigeria" and let S be the event that an email is spam. The question tells us that

   - $P(N|S) = 0.05$ since 5% of all spam messages contain the word "Nigeria,"
   - $P(S) = 0.35$ since 35% of emails are spam,
   - $P(\tilde{S}) = 0.65$ since this is the complement event to S,
   - and $P(N|\tilde{S}) = 0.001$ since only 0.1% of legitimate emails contain the word "Nigeria."

   In order to find $P(S|N)$, the probability that an email that contains the word "Nigeria" is spam, we apply version 2 of Bayes' rule:

$$P(S|N) = \frac{P(N|S)P(S)}{P(N|S)P(S) + P(N|\tilde{S})P(\tilde{S})} = \frac{0.05 \times 0.35}{0.05 \times 0.35 + 0.001 \times 0.65} = 0.96.$$

   (b) This question requires us to find the proportions of non-spam with the word "Nigeria," $P(N|\tilde{S})$, such that the probability that an email with "Nigeria" is spam is less than 0.95:

$$P(S|N) < 0.95.$$

   First, substitute the Bayes' rule expression we found in part (a) in for $P(S|N)$:

$$\frac{P(N|S)P(S)}{P(N|S)P(S) + P(N|\tilde{S})P(\tilde{S})} < 0.95.$$

   Next, plug in what we still know to be true – $P(S) = 0.35$, $P(\tilde{S}) = 0.65$, and $P(N|S) = 0.05$:

$$\frac{0.05 \times 0.35}{(0.05 \times 0.35) + 0.65P(N|\tilde{S})} < 0.95,$$

$$\frac{0.0175}{0.0175 + 0.65P(N|\tilde{S})} < 0.95.$$

   Finally, we solve the inequality for $P(N|\tilde{S})$:

$$0.0175 < 0.95\Big(0.0175 + 0.65P(N|\tilde{S})\Big),$$

$$0.0175 < 0.0166 + 0.62P(N|\tilde{S}),$$

$$P(N|\tilde{S}) > \frac{0.0175 - 0.0166}{0.62} = .0015.$$

   So if at least 0.15% of non-spam emails contain the word "Nigeria," then emails with this word will no longer be filtered.

   (c) We know that, regardless of which version of the email is used, the probability that someone is gullible is $P(G) = 0.05$ and the probability that someone is not is $P(\tilde{G}) = 0.95$. Our goal is to calculate $P(G|R)$, the probability that someone is gullible given that they respond to the email. Let's first consider version 1 of the email. In this version, we know that

   - $P(R|G) = 0.4$ since 40% of gullible people will respond,

- and $P(R|\tilde{G}) = 0.2$ since 20% of non-gullible people respond.

We plug these values into version 2 of Bayes' rule for $P(G|R)$:

$$P(G|R) = \frac{P(R|G)P(G)}{P(R|G)P(G) + P(R|\tilde{G})P(\tilde{G})} = \frac{0.4 \times 0.05}{0.4 \times 0.05 + 0.2 \times 0.95} = 0.095.$$

For the second version of the email,

- $P(R|G) = 0.1$ since 10% of gullible people will respond,
- and $P(R|\tilde{G}) = 0.001$ since only 1 in 1000 non-gullible people respond.

Again, we plug these values into version 2 of Bayes' rule for $P(G|R)$:

$$P(G|R) = \frac{P(R|G)P(G)}{P(R|G)P(G) + P(R|\tilde{G})P(\tilde{G})} = \frac{0.1 \times 0.05}{0.1 \times 0.05 + 0.001 \times 0.95} = 0.84.$$

Therefore version 2 yields a much, much higher probability that the respondents will be gullible.

9. (a) The expected utility of the juror for voting for acquittal is

$$\text{EU(acquit)} = \text{U(acquit an innocent person)}P(\tilde{G}) + \text{U(acquit a guilty person)}P(G),$$
$$= 0(1-\pi) - (1-z)\pi = -\pi(1-z).$$

The expected utility of the juror for voting to convict is

$$\text{EU(convict)} = \text{U(convict an innocent person)}P(\tilde{G}) + \text{U(convict a guilty person)}P(G).$$
$$= -z(1-\pi) + 0\pi = -z(1-\pi).$$

The juror will vote to convict when

$$\begin{aligned} \text{EU(convict)} &> \text{EU(acquit)}, \\ -z(1-\pi) &> -\pi(1-z), \\ -z + z\pi &> -\pi + z\pi, \\ -z &> -\pi, \\ z &< \pi. \end{aligned}$$

Substantively, this result means that the juror will only vote to convict when the juror's belief that the defendant is guilty outweighs her aversion to convicting an innocent person.

(b) We know from part (a) that $P(G) = \pi$ and $P(\tilde{G}) = 1 - \pi$. In addition, the problem tells us that $P(D|\tilde{G}) = p$ and $P(D|G) = q$. We simply plug these values into version 2 of Bayes' rule for $P(G|D)$:

$$P(G|D) = \frac{P(D|G)P(G)}{P(D|G)P(G) + P(D|\tilde{G})P(\tilde{G})} = \frac{q\pi}{q\pi + p(1-\pi)}.$$

(c) If $\pi = 1$ the juror is already convinced that the defendant is guilty before the trial has even taken place. Her posterior belief,

$$\frac{q(1)}{q(1) + p(1-1)} = \frac{q}{q} = 1,$$

is equal to the prior. That is, she remains convinced that the defendant is guilty. In this case, the signal didn't matter. Indeed, the fact that the defense presented a better case should have made the juror less certain about the guilt of the defendant, but because her prior biases were so strong she did not even consider the information revealed by the trial.

If $p = q$, her posterior belief becomes

$$\frac{p\pi}{p\pi + p(1-\pi)} = \frac{p\pi}{p\pi + p - p\pi} = \frac{p\pi}{p} = \pi.$$

In this case her posterior belief is equal to her prior belief in the defendant's guilt. But this result is true for any prior belief, not just the prior that $\pi = 1$. If $p = q$, then the defense would have presented the better case with the same probability regardless of whether the defendant is actually guilty or innocent. This situation can occur if the defense is far overmatched by the prosecution; regardless of the truth, the defense will lose because the lawyers are bad. Counter-intuitively, this situation can also occur if the defense is dominant; the defense will win regardless of the truth because the lawyers are excellent. In either case, the juror is receptive to the signal being sent by the trial. But because the performance of the defense appears to be independent of the truth of the case, the signal is too weak to alter the juror's belief.

Finally, if $p = 1$ and $q = 0$, the juror's posterior becomes

$$\frac{0\pi}{0\pi + 1(1-\pi)} = 0.$$

In this case, the defense is certain to present the better case if the truth is that the defendant is innocent, and is certain to present the weaker case if the defendant is guilty. Therefore the fact that the defense presented the stronger case sends the strongest possible signal to the juror about the defendant's innocence. As a result, the juror updates her belief to be certain of the defendant's innocence, regardless of her prior belief (as long as the prior is less than 1. If the prior where equal to 1 in this case, then the posterior would become 0/0, and the juror's head would simply explode).

10. (a) Each site will be cleared with probability p, where p is .95 if the Syrian government is compliant, and .6 if the government is non-compliant. Whether or not each site gets cleared can be thought of as a series of independent experiments with the same probability of success, so we can calculate the probability of clearing 7 sites using the binomial distribution. Denote the event that the government clears 7 sites as $X = 7$. If the government is compliant, then

$$P(X = 7|C) = \binom{10}{7} .95^7 .05^3 = 0.01,$$

and if the government is non-compliant, the probability is

$$P(X = 7|\tilde{C}) = \binom{10}{7} .6^7 .4^3 = 0.215.$$

(b) We want to find the probability that the government was compliant given that 7 sites were cleared. Since there are only two possible states of the world under consideration here – the one in which the Syrian government is compliant and the one in which the government is non-compliant – we can use the second version of Bayes' rule. The equation is

$$P(C|X=7) = \frac{P(X=7|C)P(C)}{P(X=7|C)P(C) + P(X=7|\tilde{C})P(\tilde{C})}.$$

Our prior belief that the government was compliant, as given in the problem, is $P(C) = .8$, implying that $P(\tilde{C}) = .2$. In part (a) we calculated that $P(X=7|C) = .01$ and $P(X=7|\tilde{C}) = .215$. We simply plug these numbers into the above equation:

$$P(C|X=7) = \frac{P(X=7|C)P(C)}{P(X=7|C)P(C) + P(X=7|\tilde{C})P(\tilde{C})} = \frac{.01 \times .8}{.01 \times .8 + .215 \times .2} = \frac{.008}{.051} = .157.$$

So our belief that the government was compliant is reduced from .8 to .157.

11. Let T be the event in which Obama uses the word "terrorism." We are trying to update our beliefs on Obama's ideology. The probability that Obama is far left given that he spoke about terrorism is

$$P(FL|T) = \frac{P(T|FL)P(FL)}{P(T)}.$$

We don't, however, know $P(T)$. What we do know are the probabilities of T given each of the six ideologies. We can calculate these from the table:

- $P(T|FL) = \frac{2}{10000} = 0.0002$
- $P(T|EL) = \frac{5}{10000} = 0.0005$
- $P(T|CL) = \frac{70}{10000} = 0.007$
- $P(T|CR) = \frac{50}{10000} = 0.005$
- $P(T|LR) = \frac{40}{10000} = 0.004$
- $P(T|FR) = \frac{120}{10000} = 0.0120$

Because these categories are a partition, version 3 of Bayes' rule tells us that we can replace $P(T)$ in the denominator with the sum of these conditional times the prior probabilities of each category:

$$P(T) = P(T|FL)P(FL) + P(T|EL)P(EL) + P(T|CL)P(CL)$$
$$+P(T|CR)P(CR) + P(T|LR)P(LR) + P(T|FR)P(FR)$$

Plugging in the probabilities, we get

$$P(T) = (0.0002 \times 0.2) + (0.0005 \times 0.25) + (0.007 \times 0.35)$$
$$+(0.005 \times 0.1) + (0.004 \times 0.05) + (0.0120 \times 0.05) = 0.003915$$

Now it is fairly straightforward to update our belief that Obama is far left:

$$P(FL|T) = \frac{P(T|FL)P(FL)}{P(T)} = \frac{0.0002 \times 0.2}{0.003915} = 0.01.$$

For economic left

$$P(EL|T) = \frac{P(T|EL)P(EL)}{P(T)} = \frac{0.0005 \times 0.25}{0.003915} = 0.03.$$

For center left,

$$P(CL|T) = \frac{P(T|CL)P(CL)}{P(T)} = \frac{0.007 \times 0.35}{0.003915} = 0.63.$$

For center right,

$$P(CR|T) = \frac{P(T|CR)P(CR)}{P(T)} = \frac{0.005 \times 0.12}{0.003915} = 0.13.$$

For libertarian right,

$$P(LR|T) = \frac{P(T|LR)P(LR)}{P(T)} = \frac{0.004 \times 0.05}{0.003915} = 0.05.$$

And for far right,

$$P(FR|T) = \frac{P(T|FR)P(FR)}{P(T)} = \frac{0.0120 \times 0.05}{0.003915} = 0.15.$$