Stat 110 Strategic Practice 2, Fall 2011

Prof. Joe Blitzstein (Department of Statistics, Harvard University)

1 Inclusion-Exclusion

1. For a group of 7 people, find the probability that all 4 seasons (winter, spring, summer, fall) occur at least once each among their birthdays, assuming that all seasons are equally likely.

2. Alice attends a small college in which each class meets only once a week. She is deciding between 30 non-overlapping classes. There are 6 classes to choose from for each day of the week, Monday through Friday. Trusting in the benevolence of randomness, Alice decides to register for 7 randomly selected classes out of the 30, with all choices equally likely. What is the probability that she will have classes every day, Monday through Friday? (This problem can be done either directly using the naive definition of probability, or using inclusion-exclusion.)

2 Independence

1. Is it possible that an event is independent of itself? If so, when?

2. Is it always true that if A and B are independent events, then A^c and B^c are independent events? Show that it is, or give a counterexample.

3. Give an example of 3 events A, B, C which are pairwise independent but not independent. Hint: find an example where whether C occurs is completely determined if we know whether A occurred and whether B occurred, but completely undetermined if we know only one of these things.

4. Give an example of 3 events A, B, C which are not independent, yet satisfy $P(A \cap B \cap C) = P(A)P(B)P(C)$. Hint: consider simple and extreme cases.

3 Thinking Conditionally

1. A bag contains one marble which is either green or blue, with equal probabilities. A green marble is put in the bag (so there are 2 marbles now), and then a random marble is taken out. The marble taken out is green. What is the probability that the remaining marble is also green?

Historical note: this problem was first posed by Lewis Carroll in 1893.

2. A spam filter is designed by looking at commonly occurring phrases in spam. Suppose that 80% of email is spam. In 10% of the spam emails, the phrase "free money" is used, whereas this phrase is only used in 1% of non-spam emails. A new email has just arrived, which does mention "free money". What is the probability that it is spam?

3. Let G be the event that a certain individual is guilty of a certain robbery. In gathering evidence, it is learned that an event E_1 occurred, and a little later it is also learned that another event E_2 also occurred.

   (a) Is it possible that individually, these pieces of evidence increase the chance of guilt (so $P(G|E_1) > P(G)$ and $P(G|E_2) > P(G)$), but together they decrease the chance of guilt (so $P(G|E_1, E_2) < P(G)$)?

   (b) Show that the probability of guilt given the evidence is the same regardless of whether we update our probabilities all at once, or in two steps (after getting the first piece of evidence, and again after getting the second piece of evidence). That is, we can either update all at once (computing $P(G|E_1, E_2)$ in one step), or we can first update based on E_1, so that our new probability function is $P_{\text{new}}(A) = P(A|E_1)$, and then update based on E_2 by computing $P_{\text{new}}(G|E_2)$.

4. A crime is committed by one of two suspects, A and B. Initially, there is equal evidence against both of them. In further investigation at the crime scene, it is found that the guilty party had a blood type found in 10% of the population. Suspect A does match this blood type, whereas the blood type of Suspect B is unknown.

   (a) Given this new information, what is the probability that A is the guilty party?

   (b) Given this new information, what is the probability that B's blood type matches that found at the crime scene?

5. You are going to play 2 games of chess with an opponent whom you have never played against before (for the sake of this problem). Your opponent is equally likely to be a beginner, intermediate, or a master. Depending on which, your chances of winning an individual game are 90%, 50%, or 30%, respectively.

   (a) What is your probability of winning the first game?

(b) Congratulations: you won the first game! Given this information, what is the probability that you will also win the second game (assume that, given the skill level of your opponent, the outcomes of the games are independent)?

(c) Explain the distinction between assuming that the outcomes of the games are independent and assuming that they are conditionally independent given the opponent's skill level. Which of these assumptions seems more reasonable, and why?

Stat 110 Strategic Practice 2 Solutions, Fall 2011

Prof. Joe Blitzstein (Department of Statistics, Harvard University)

1 Inclusion-Exclusion

1. For a group of 7 people, find the probability that all 4 seasons (winter, spring, summer, fall) occur at least once each among their birthdays, assuming that all seasons are equally likely.

   Let A_i be the event that there are no birthdays in the ith season. The probability that all seasons occur at least once is $1 - P(A_1 \cup A_2 \cup A_3 \cup A_4)$. Note that $A_1 \cap A_2 \cap A_3 \cap A_4 = \emptyset$. Using the inclusion-exclusion principle and the symmetry of the seasons,

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3 \cup A_4) &= \sum_{i=1}^{4} P(A_i) - \sum_{i=1}^{3} \sum_{j>i} P(A_i \cap A_j) \\ &\quad + \sum_{i=1}^{3} \sum_{j>i} \sum_{k>j} P(A_i \cap A_j \cap A_k) \\ &= 4P(A_1) - 6P(A_1 \cap A_2) + 4P(A_1 \cap A_2 \cap A_3). \end{aligned}$$

   We have $P(A_1) = (3/4)^7$. Similarly,

$$P(A_1 \cap A_2) = \frac{1}{2^7} \text{ and } P(A_1 \cap A_2 \cap A_3) = \frac{1}{4^7}.$$

   Therefore, $P(A_1 \cup A_2 \cup A_3 \cup A_4) = 4(\frac{3}{4})^7 - \frac{6}{2^7} + \frac{4}{4^7}$. So the probability that all 4 seasons occur at least once is $1 - \left(4(\frac{3}{4})^7 - \frac{6}{2^7} + \frac{4}{4^7}\right) \approx 0.513$.

2. Alice attends a small college in which each class meets only once a week. She is deciding between 30 non-overlapping classes. There are 6 classes to choose from for each day of the week, Monday through Friday. Trusting in the benevolence of randomness, Alice decides to register for 7 randomly selected classes out of the 30, with all choices equally likely. What is the probability that she will have classes every day, Monday through Friday? (This problem can be done either directly using the naive definition of probability, or using inclusion-exclusion.)

   Direct Method: There are two general ways that Alice can have class every day: either she has 2 days with 2 classes and 3 days with 1 class, or she has 1 day with 3 classes, and has 1 class on each of the other 4 days. The number of possibilities for the former is $\binom{5}{2}\binom{6}{2}^2 6^3$ (choose the 2 days when she has 2 classes, and then select 2 classes on those days and 1 class for the other days). The number of possibilities for the latter is $\binom{5}{1}\binom{6}{3}6^4$. So the probability is

$$\frac{\binom{5}{2}\binom{6}{2}^2 6^3 + \binom{5}{1}\binom{6}{3}6^4}{\binom{30}{7}} = \frac{114}{377} \approx 0.302.$$

Inclusion-Exclusion Method: we will use inclusion-exclusion to find the probability of the complement, which is the event that she has at least one day with no classes. Let $B_i = A_i^c$. Then

$$P(B_1 \cup B_2 \cdots \cup B_5) = \sum_i P(B_i) - \sum_{i<j} P(B_i \cap B_j) + \sum_{i<j<k} P(B_i \cap B_j \cap B_k)$$

(terms with the intersection of 4 or more B_i's are not needed since Alice must have classes on at least 2 days). We have

$$P(B_1) = \frac{\binom{24}{7}}{\binom{30}{7}}, P(B_1 \cap B_2) = \frac{\binom{18}{7}}{\binom{30}{7}}, P(B_1 \cap B_2 \cap B_3) = \frac{\binom{12}{7}}{\binom{30}{7}}$$

and similarly for the other intersections. So

$$P(B_1 \cup \cdots \cup B_5) = 5\frac{\binom{24}{7}}{\binom{30}{7}} - \binom{5}{2}\frac{\binom{18}{7}}{\binom{30}{7}} + \binom{5}{3}\frac{\binom{12}{7}}{\binom{30}{7}} = \frac{263}{377}.$$

Therefore,

$$P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) = \frac{114}{377} \approx 0.302.$$

2 Independence

1. Is it possible that an event is independent of itself? If so, when?

   Let A be an event. If A is independent of itself, then $P(A) = P(A \cap A) = P(A)^2$, so $P(A)$ is 0 or 1. So this is only possible in the extreme cases that the event has probability 0 or 1.

2. Is it always true that if A and B are independent events, then A^c and B^c are independent events? Show that it is, or give a counterexample.

   Yes, because we have

$$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - (P(A) + P(B) - P(A \cap B));$$

   since A and B are independent, this becomes

$$1 - P(A) - P(B) + P(A)P(B) = (1 - P(A))(1 - P(B)) = P(A^c)P(B^c).$$

3. Give an example of 3 events A, B, C which are pairwise independent but not independent. Hint: find an example where whether C occurs is completely determined if we know whether A occurred and whether B occurred, but completely undetermined if we know only one of these things.

   Consider two fair, independent coin tosses, and let A be the event that the first toss is Heads, B be the event that the second toss is Heads, and C be the event that the two tosses have the same result. Then A, B, C are dependent since $P(A \cap B \cap C) = P(A \cap B) = P(A)P(B) = 1/4 \neq 1/8 = P(A)P(B)P(C)$, but they are pairwise independent: A and B are independent by definition; A and C are independent since $P(A \cap C) = P(A \cap B) = 1/4 = P(A)P(C)$, and similarly B and C are independent.

4. Give an example of 3 events A, B, C which are not independent, yet satisfy $P(A \cap B \cap C) = P(A)P(B)P(C)$. Hint: consider simple and extreme cases.

   Consider the extreme case where $P(A) = 0$. Then it's automatically true that $P(A)P(B)P(C) = 0$, and also $P(A \cap B \cap C) = 0$ since $A \cap B \cap C$ is a subset of A (and in general, if event A_1 is a subset of event A_2, then $P(A_1) \leq P(A_2)$). Then take B and C to be dependent, e.g., take any example with $B = C$ and $0 < P(B) < 1$.

3 Thinking Conditionally

1. A bag contains one marble which is either green or blue, with equal probabilities. A green marble is put in the bag (so there are 2 marbles now), and then a random marble is taken out. The marble taken out is green. What is the probability that the remaining marble is also green?

   Historical note: this problem was first posed by Lewis Carroll in 1893.

Let A be the event that the initial marble is green, B be the event that the removed marble is green, and C be the event that the remaining marble is green. We need to find $P(C|B)$. There are several ways to find this; one natural way is to condition on whether the initial marble is green:

$$P(C|B) = P(C|B, A)P(A|B) + P(C|B, A^c)P(A^c|B) = 1P(A|B) + 0P(A^c|B).$$

To find $P(A|B)$, use Bayes' Rule:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{1/2}{P(B|A)P(A) + P(B|A^c)P(A^c)} = \frac{1/2}{1/2 + 1/4} = \frac{2}{3}.$$

So $P(C|B) = 2/3$.

2. A spam filter is designed by looking at commonly occurring phrases in spam. Suppose that 80% of email is spam. In 10% of the spam emails, the phrase "free money" is used, whereas this phrase is only used in 1% of non-spam emails. A new email has just arrived, which does mention "free money". What is the probability that it is spam?

   Let S be the event that an email is spam and F be the event that an email has the "free money" phrase. By Bayes' Rule,

$$P(S|F) = \frac{P(F|S)P(S)}{P(F)} = \frac{0.1 \cdot 0.8}{0.1 \cdot 0.8 + 0.01 \cdot 0.2} = \frac{80/1000}{82/1000} = \frac{80}{82} \approx 0.9756.$$

3. Let G be the event that a certain individual is guilty of a certain robbery. In gathering evidence, it is learned that an event E_1 occurred, and a little later it is also learned that another event E_2 also occurred.

   (a) Is it possible that individually, these pieces of evidence increase the chance of guilt (so $P(G|E_1) > P(G)$ and $P(G|E_2) > P(G)$), but together they decrease the chance of guilt (so $P(G|E_1, E_2) < P(G)$)?

   Yes, this is possible. In fact, it is possible to have two events which separately provide evidence in favor of G, yet which together preclude G! For example, suppose that the crime was committed between 1 pm and 3 pm on a certain day. Let E_1 be the event that the suspect was at a nearby coffeeshop from 1 pm to 2 pm that day, and let E_2 be the event that the suspect was at the nearby coffeeshop from 2 pm to 3 pm that day.

Then $P(G|E_1) > P(G), P(G|E_2) > P(G)$ (assuming that being in the vicinity helps show that the suspect had the opportunity to commit the crime), yet $P(G|E_1 \cap E_2) < P(G)$ (as being in the coffeehouse from 1 pm to 3 pm gives the suspect an alibi for the full time).

(b) Show that the probability of guilt given the evidence is the same regardless of whether we update our probabilities all at once, or in two steps (after getting the first piece of evidence, and again after getting the second piece of evidence). That is, we can either update all at once (computing $P(G|E_1, E_2)$ in one step), or we can first update based on E_1, so that our new probability function is $P_{\text{new}}(A) = P(A|E_1)$, and then update based on E_2 by computing $P_{\text{new}}(G|E_2)$.

This follows from the definition of conditional probability:

$$P_{\text{new}}(G|E_2) = \frac{P_{\text{new}}(G, E_2)}{P_{\text{new}}(E_2)} = \frac{P(G, E_2|E_1)}{P(E_2|E_1)} = \frac{P(G, E_1, E_2)/P(E_1)}{P(E_1, E_2)/P(E_1)} = P(G|E_1, E_2).$$

4. A crime is committed by one of two suspects, A and B. Initially, there is equal evidence against both of them. In further investigation at the crime scene, it is found that the guilty party had a blood type found in 10% of the population. Suspect A does match this blood type, whereas the blood type of Suspect B is unknown.

(a) Given this new information, what is the probability that A is the guilty party?

Let M be the event that A's blood type matches the guilty party's and for brevity, write A for "A is guilty" and B for "B is guilty". By Bayes' Rule,

$$P(A|M) = \frac{P(M|A)P(A)}{P(M|A)P(A) + P(M|B)P(B)} = \frac{1/2}{1/2 + (1/10)(1/2)} = \frac{10}{11}.$$

(We have $P(M|B) = 1/10$ since, given that B is guilty, the probability that A's blood type matches the guilty party's is the same probability as for the general population.)

(b) Given this new information, what is the probability that B's blood type matches that found at the crime scene?

Let C be the event that B's blood type matches, and condition on whether B is guilty. This gives

$$P(C|M) = P(C|M, A)P(A|M) + P(C|M, B)P(B|M) = \frac{1}{10} \cdot \frac{10}{11} + \frac{1}{11} = \frac{2}{11}.$$

5. You are going to play 2 games of chess with an opponent whom you have never played against before (for the sake of this problem). Your opponent is equally likely to be a beginner, intermediate, or a master. Depending on which, your chances of winning an individual game are 90%, 50%, or 30%, respectively.

(a) What is your probability of winning the first game?

Let W_i be the event of winning the ith game. By the law of total probability,

$$P(W_1) = (0.9 + 0.5 + 0.3)/3 = 17/30.$$

(b) Congratulations: you won the first game! Given this information, what is the probability that you will also win the second game (assume that, given the skill level of your opponent, the outcomes of the games are independent)?

We have $P(W_2|W_1) = P(W_2, W_1)/P(W_1)$. The denominator is known from (a), while the numerator can be found by conditioning on the skill level of the opponent:

$$P(W_1, W_2) = \frac{1}{3}P(W_1, W_2|\text{beginner}) + \frac{1}{3}P(W_1, W_2|\text{intermediate}) + \frac{1}{3}P(W_1, W_2|\text{expert}).$$

Since W_1 and W_2 are conditionally independent given the skill level of the opponent, this becomes

$$P(W_1, W_2) = (0.9^2 + 0.5^2 + 0.3^2)/3 = 23/60.$$

So

$$P(W_2|W_1) = \frac{23/60}{17/30} = 23/34.$$

(c) Explain the distinction between assuming that the outcomes of the games are independent and assuming that they are conditionally independent given the opponent's skill level. Which of these assumptions seems more reasonable, and why?

Independence here means that knowing one game's outcome gives no information about the other game's outcome, while conditional independence is the same statement where all probabilities are conditional on the opponent's skill level. Conditional independence given the opponent's skill level is a more reasonable assumption here. This is because winning the first game gives information about the opponent's skill level, which in turn gives information about the result of the second game.

That is, if the opponent's skill level is treated as fixed and known, then it may be reasonable to assume independence of games given this information; with the opponent's skill level random, earlier games can be used to help infer the opponent's skill level, which affects the probabilities for future games.

Stat 110 Homework 2, Fall 2011

Prof. Joe Blitzstein (Department of Statistics, Harvard University)

1. Arby has a belief system assigning a number $P_{\text{Arby}}(A)$ between 0 and 1 to every event A (for some sample space). This represents Arby's subjective degree of belief about how likely A is to occur. For any event A, Arby is willing to pay a price of $1000 \cdot P_{\text{Arby}}(A)$ dollars to buy a certificate such as the one shown below:

Certificate
The owner of this certificate can redeem it for \$1000 if A occurs. No value if A does not occur, except as required by federal, state, or local law. No expiration date.

Likewise, Arby is willing to sell such a certificate at the same price. Indeed, Arby is willing to buy or sell any number of certificates at this price, as Arby considers it the "fair" price.

Arby, not having taken Stat 110, stubbornly refuses to accept the axioms of probability. In particular, suppose that there are two *disjoint* events A and B with

$$P_{\text{Arby}}(A \cup B) \neq P_{\text{Arby}}(A) + P_{\text{Arby}}(B).$$

Show how to make Arby go bankrupt, by giving a list of transactions Arby is willing to make that will *guarantee* that Arby will lose money (you can assume it will be known whether A occurred and whether B occurred the day after any certificates are bought/sold).

2. A card player is dealt a 13 card hand from a well-shuffled, standard deck of cards. What is the probability that the hand is void in at least one suit ("void in a suit" means having no cards of that suit)?

3. A family has 3 children, creatively named A, B, and C.

(a) Discuss intuitively (but clearly) whether the event "A is older than B" is independent of the event "A is older than C."

(b) Find the probability that A is older than B, given that A is older than C.

4. Two coins are in a hat. The coins look alike, but one coin is fair (with probability 1/2 of Heads), while the other coin is biased, with probability 1/4 of Heads. One of the coins is randomly pulled from the hat, without knowing which of the two it is. Call the chosen coin "Coin C".

(a) Coin C is tossed twice, showing Heads both times. Given this information, what is the probability that Coin C is the fair coin?

(b) Are the events "first toss of Coin C is Heads" and "second toss of Coin C is Heads" independent? Explain.

(c) Find the probability that in 10 flips of Coin C, there will be exactly 3 Heads. (The coin is equally likely to be either of the 2 coins; do not assume it already landed Heads twice as in (a).)

5. A woman has been murdered, and her husband is accused of having committed the murder. It is known that the man abused his wife repeatedly in the past, and the prosecution argues that this is important evidence pointing towards the man's guilt. The defense attorney says that the history of abuse is irrelevant, as only 1 in 1000 men who beat their wives end up murdering them.

Assume that the defense attorney's 1 in 1000 figure is correct, and that half of men who murder their wives previously abused them. Also assume that 20% of murdered women were killed by their husbands, and that if a woman is murdered and the husband is *not* guilty, then there is only a 10% chance that the husband abused her. What is the probability that the man is guilty? Is the prosecution right that the abuse is important evidence in favor of guilt?

6. A family has two children. Assume that birth month is independent of gender, with boys and girls equally likely and all months equally likely, and assume that the elder child's characteristics are independent of the younger child's characteristics).

(a) Find the probability that both are girls, given that the elder child is a girl who was born in March.

(b) Find the probability that both are girls, given that at least one is a girl who was born in March.

Stat 110 Homework 2 Solutions, Fall 2011

Prof. Joe Blitzstein (Department of Statistics, Harvard University)

1. Arby has a belief system assigning a number $P_{\text{Arby}}(A)$ between 0 and 1 to every event A (for some sample space). This represents Arby's subjective degree of belief about how likely A is to occur. For any event A, Arby is willing to pay a price of $1000 \cdot P_{\text{Arby}}(A)$ dollars to buy a certificate such as the one shown below:

Certificate
The owner of this certificate can redeem it for \$1000 if A occurs. No value if A does not occur, except as required by federal, state, or local law. No expiration date.

Likewise, Arby is willing to sell such a certificate at the same price. Indeed, Arby is willing to buy or sell any number of certificates at this price, as Arby considers it the "fair" price.

Arby, not having taken Stat 110, stubbornly refuses to accept the axioms of probability. In particular, suppose that there are two *disjoint* events A and B with

$$P_{\text{Arby}}(A \cup B) \neq P_{\text{Arby}}(A) + P_{\text{Arby}}(B).$$

Show how to make Arby go bankrupt, by giving a list of transactions Arby is willing to make that will *guarantee* that Arby will lose money (you can assume it will be known whether A occurred and whether B occurred the day after any certificates are bought/sold).

Suppose first that

$$P_{\text{Arby}}(A \cup B) < P_{\text{Arby}}(A) + P_{\text{Arby}}(B).$$

Call a certificate like the one show above, with any event C in place of A, a C-certificate. Measuring money in units of thousands of dollars, Arby is willing to pay $P_{\text{Arby}}(A) + P_{\text{Arby}}(B)$ to buy an A-certificate and a B-certificate, and is willing to sell an $(A \cup B)$-certificate for $P_{\text{Arby}}(A \cup B)$. In those transactions, Arby loses $P_{\text{Arby}}(A) + P_{\text{Arby}}(B) - P_{\text{Arby}}(A \cup B)$ and will not recoup any of that loss because if A or B occurs, Arby will have to pay out an amount equal to the amount Arby receives (since it's impossible for both A and B to occur).

Now suppose instead that

$$P_{\text{Arby}}(A \cup B) > P_{\text{Arby}}(A) + P_{\text{Arby}}(B).$$

Measuring money in units of thousands of dollars, Arby is willing to sell an A-certificate for $P_{\text{Arby}}(A)$, sell a B-certificate for $P_{\text{Arby}}(B)$, and buy a $(A \cup B)$-certificate for $P_{\text{Arby}}(A \cup B)$. In so doing, Arby loses $P_{\text{Arby}}(A \cup B) - (P_{\text{Arby}}(A) + P_{\text{Arby}}(B))$, and Arby won't recoup any of this loss, similarly to the above. (In fact, in this case, even if A and B are not disjoint, Arby will not recoup any of the loss, and will lose more money if both A and B occur.)

By buying/selling a sufficiently large number of certificates from/to Arby as described above, you can guarantee that you'll get all of Arby's money; this is called an *arbitrage opportunity*. This problem illustrates the fact that the axioms of probability are not arbitrary, but rather are *essential* for coherent thought (at least the first axiom, and the second with finite unions rather than countably infinite unions).

Arbitrary axioms allow arbitrage attacks; principled properties and perspectives on probability potentially prevent perdition.

2. A card player is dealt a 13 card hand from a well-shuffled, standard deck of cards. What is the probability that the hand is void in at least one suit ("void in a suit" means having no cards of that suit)?

Let S, H, D, C be the events of being void in Spades, Hearts, Diamonds, Clubs, respectively. We want to find $P(S \cup D \cup H \cup C)$. By inclusion-exclusion and symmetry,

$$P(S \cup D \cup H \cup C) = 4P(S) - 6P(S \cap H) + 4P(S \cap H \cap D) - P(S \cap H \cap D \cap C).$$

The probability of being void in a specific suit is $\frac{\binom{39}{13}}{\binom{52}{13}}$. The probability of being void in 2 specific suits is $\frac{\binom{26}{13}}{\binom{52}{13}}$. The probability of being void in 3 specific suits is $\frac{1}{\binom{52}{13}}$. And the last term is 0 since it's impossible to be void in everything. So the probability is

$$4\frac{\binom{39}{13}}{\binom{52}{13}} - 6\frac{\binom{26}{13}}{\binom{52}{13}} + \frac{4}{\binom{52}{13}} \approx 0.051.$$

3. A family has 3 children, creatively named A, B, and C.

(a) Discuss intuitively (but clearly) whether the event "A is older than B" is independent of the event "A is older than C."

They are not independent: knowing that A is older than B makes it more likely that A is older than C, as the if A is older than B, then the only way that A can be younger than C is if the birth order is CAB, whereas the birth orders ABC

and ACB are both compatible with A being older than B. To make this more intuitive, think of an extreme case where there are 100 children instead of 3, call them $A_1, \dots, A_{100}$. Given that A_1 is older than all of $A_2, A_3, \dots, A_{99}$, it's clear that A_1 is very old (relatively), whereas there isn't evidence about where A_{100} fits into the birth order.

(b) Find the probability that A is older than B, given that A is older than C.

Writing $x > y$ to mean that x is older than y,

$$P(A > B|A > C) = \frac{P(A > B, A > C)}{P(A > C)} = \frac{1/3}{1/2} = \frac{2}{3}$$

since $P(A > B, A > C) = P(A \text{ is the eldest child}) = 1/3$ (unconditionally, any of the 3 children is equally likely to be the eldest).

4. Two coins are in a hat. The coins look alike, but one coin is fair (with probability 1/2 of Heads), while the other coin is biased, with probability 1/4 of Heads. One of the coins is randomly pulled from the hat, without knowing which of the two it is. Call the chosen coin "Coin C".

(a) Coin C is tossed twice, showing Heads both times. Given this information, what is the probability that Coin C is the fair coin?

By Bayes' Rule,

$$P(\text{fair}|HH) = \frac{P(HH|\text{fair})P(\text{fair})}{P(HH)} = \frac{(1/4)(1/2)}{(1/4)(1/2) + (1/16)(1/2)} = \frac{4}{5}.$$

(b) Are the events "first toss of Coin C is Heads" and "second toss of Coin C is Heads" independent? Explain.

They're not independent: the first toss being Heads is evidence in favor of the coin being the fair coin, giving information about probabilities for the second toss.

(c) Find the probability that in 10 flips of Coin C, there will be exactly 3 Heads. (The coin is equally likely to be either of the 2 coins; do not assume it already landed Heads twice as in (a).)

Let X be the number of Heads in 10 tosses. By the Law of Total Probability (conditioning on which of the two coins C is),

$$\begin{aligned} P(X=3) &= P(X=3|\text{fair})P(\text{fair}) + P(X=3|\text{biased})P(\text{biased}) \\ &= \binom{10}{3}(1/2)^{10}(1/2) + \binom{10}{3}(1/4)^3(3/4)^7(1/2) \\ &= \frac{1}{2}\binom{10}{3}\left(\frac{1}{2^{10}} + \frac{3^7}{4^{10}}\right). \end{aligned}$$

5. A woman has been murdered, and her husband is accused of having committed the murder. It is known that the man abused his wife repeatedly in the past, and the prosecution argues that this is important evidence pointing towards the man's guilt. The defense attorney says that the history of abuse is irrelevant, as only 1 in 1000 men who beat their wives end up murdering them.

Assume that the defense attorney's 1 in 1000 figure is correct, and that half of men who murder their wives previously abused them. Also assume that 20% of murdered women were killed by their husbands, and that if a woman is murdered and the husband is *not* guilty, then there is only a 10% chance that the husband abused her. What is the probability that the man is guilty? Is the prosecution right that the abuse is important evidence in favor of guilt?

The defense attorney claims that the probability of a man murdering his wife given that he abuses her is 1/1000. Note though that this conditional probability is not what is in issue. We need to *condition on all the evidence*, and here the evidence is not only that the man beat his wife, but also that the wife did get murdered. That is, we need the probability of guilt given that the husband abused his wife *and* that she was murdered. Let's introduce some notation to specify the different events.

G: the man is guilty of murdering his wife
M: the wife was murdered
A: the man abused his wife

The relevant probability is $P(G|A,M)$. We are given that $P(A|G,M) = 0.5, P(G|M) = 0.2, P(A|G^c,M) = 0.1$.

$$P(G|A,M) = \frac{P(A|G,M)P(G|M)}{P(A|M)} = \frac{P(A|G,M)P(G|M)}{P(A|G,M)P(G|M) + P(A|G^c,M)P(G^c|M)}$$

(This is Bayes' Rule, where the probabilities are all taken as conditional M.) So

$$P(G|A,M) = \frac{0.5 \cdot 0.2}{0.5 \cdot 0.2 + 0.1 \cdot 0.8} = \frac{5}{9}.$$

This means that the evidence of abuse raised the probability of guilt from 20% to 56%, certainly important evidence. Since the woman was indeed murdered, it is crucial to condition on that information when weighing other evidence.

6. A family has two children. Assume that birth month is independent of gender, with boys and girls equally likely and all months equally likely, and assume that the elder child's characteristics are independent of the younger child's characteristics).

(a) Find the probability that both are girls, given that the elder child is a girl who was born in March.

Let G_j be the event that the jth born child is a girl and M_j be the event that the jth born child was born in March, for $j \in \{1, 2\}$. Then $P(G_1 \cap G_2 | G_1 \cap M_1) = P(G_2 | G_1 \cap M_1)$, since if we know that G_1 occurs, then $G_1 \cap G_2$ occurring is the same thing as G_2 occurring. By independence of the characteristics of the children, $P(G_2 | G_1 \cap M_1) = P(G_2) = 1/2$.

(b) Find the probability that both are girls, given that at least one is a girl who was born in March.

$$
\begin{aligned}
P(\text{both girls}|\text{at least one March-born girl}) &= \frac{P(\text{both girls, at least one born in March})}{P(\text{at least one March-born girl})} \\
&= \frac{(1/4)(1 - (11/12)^2)}{1 - (23/24)^2} \\
&= \frac{23}{47} \\
&\approx 0.489.
\end{aligned}
$$

In contrast, $P(\text{both girls}|\text{at least one girl}) = 1/3$. So the seemingly irrelevant "born in March" information actually matters! By symmetry, the answer would stay the same if we replaced "born in March" by, say, "born in July"; so it's not the awesomeness of March that matters, but rather the fact that the information brings "at least one" closer to being "a specific one." The more detailed the information being conditioned on, the closer this becomes to specifying one of the children, and thus the closer the answer gets to 1/2.