EECS 70   Discrete Mathematics and Probability Theory
Fall 2014   Anant Sahai
Homework 9

This homework is due November 3, 2014, at 12:00 noon.

1. **Section Rollcall!**

   In your self-grading for this question, give yourself a 10, and write down what you wrote for parts (a) and (b) below as a comment. Put the answers in your written homework as well.

   (a) What discussion did you attend on Monday last week? If you did not attend section on that day, please tell us why.

   (b) What discussion did you attend on Wednesday last week? If you did not attend section on that day, please tell us why.

2. **Intro to Randomness Lab (cont.)**

   In this week's lab, we will continue our coin tossing example, but see it from a different perspective. Make sure you review the lab solution from Homework 8 before moving on.

   From this Virtual Lab onward, students who want to can also choose to completely rewrite the question. Basically, for parts (a) through (g) in this lab, read them, and then if you want, come up with your own formulation of how to do a series of experiments that result in the same discoveries. Then, write up the results nicely using plots as appropriate to show what you observed. Similarly for part (h), but this is its own separate thing.

   Please download the IPython starter code from Piazza or the course webpage, and answer the following questions.

   (a) Let's change gears a little bit from last time. Consider the following visualization of a sequence of coin flips. We start at zero. For every head we get, we add one. For every tail we get, we subtract one. Hence, a sequence of 1000 coin tosses would be a path that starts at (0,0), and then goes to either $(1,1)$ or $(1,-1)$, and continues wandering till $(1000, y)$ somewhere. Plot 20 such paths on the same plot based on randomly flipped coins. Each sample path should have 1000 coin tosses.

   What do you observe about the paths?

   Hint: First, implement the `rand_one` function, which generates -1 and 1 randomly with roughly 50% probability each.

   Then, implement the `path(n)` function, which returns a list of *n* elements that starts at 0 and every element thereafter is either one more or one less than the previous one. In Python, you can access the last element in a list with the syntax `lst[-1]`.

   Solutions: The path is quite random, but it spreads out more as *k* increases. This is consonant with our common-units scaling in part (d) of last week's lab; with more coin tosses, the deviation gets larger.

20 random walks of 1000 coin flips
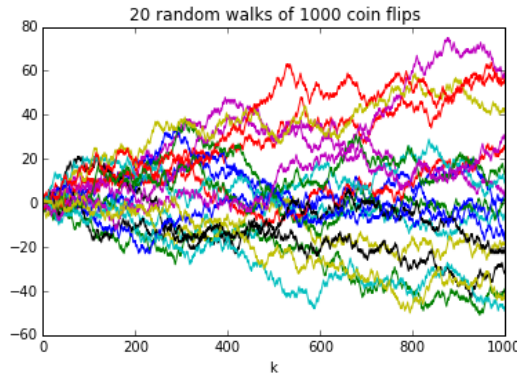
(b) Notice that the histograms you were plotting earlier were effectively looking at vertical slices in this picture and asking how many sample paths were crossing through a particular y coordinate. (If we are looking at k tosses, then having exactly h heads is the same as this sample path crossing through $(k, h - (k - h)) = (k, 2h - k)$)

Now, let's see what the rescalings we were doing correspond to. The common-set-of-units scaling is what the previous part corresponded to. How would you change the scaling to correspond to the normalized set of units in part (e) of last week's lab? (in this plot, a sample path that consists of all heads should basically be a straight line that stays at the upper-limit — say 1. And a sample path that consists of all tails should be a straight line that stays at the lower limit – say -1).

Give this new scaling (it will depend on k — so it will change the visual shape of a path) and plot 100 sample paths of 1000 coin tosses each.

Comment on what this suggests relative to the earlier plots.

Solutions: If we have taken k steps, we want to rescale our position by $\frac{1}{k}$. This is similar to what we did in part (e) from VL8. The paths are clustering closer and closer as k increases. With more coin tosses, the deviation is smaller, in contrast to the previous part.


20 normalized random walks of 1000 coin flips
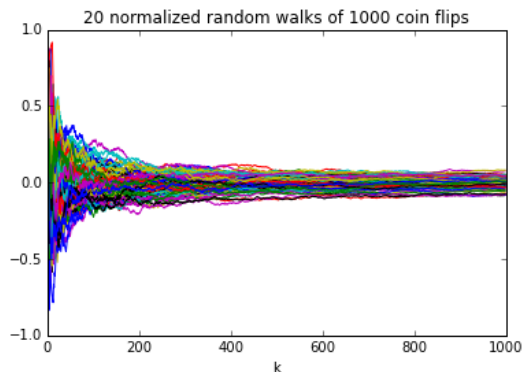
(c) Shifting gears one more time, we are now going to look at the same basic experiment — tossing a fair coin k times — in a third way. Let R for a given run be the ratio of heads.

Fix $k = 1000$ to be the number of coin tosses in a run. Let $m = 1000$ be the number of runs. Plot how often $R \le q$ as a function of q. The vertical axis should be (in linear scale) the fraction of the m runs in which $R \le q$, while the horizontal scale should have q ranging from 0 to 1.

What do you notice about this curve?

Hint: Implement the function `q_curve`, which returns the sorted fraction of heads for each of the m runs. You may find Python's built-in `sorted` function helpful here.

**Solutions:** By construction, this is an increasing curve.
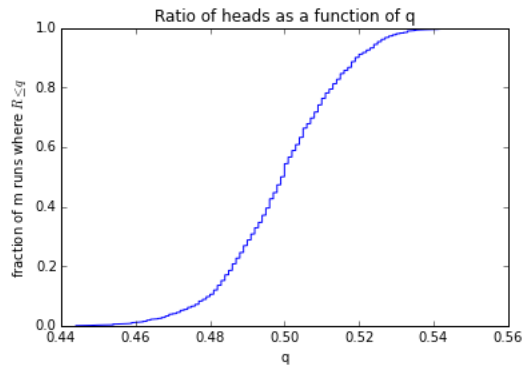
It's fine if your x-axis's range is from 0 to 1. This isn't the desired behavior (since you will only see a sharp increasing curve), but the spec was not clear about this.

(d) Repeat the previous part for different values of k and put them all on the same plot.

Try $k = 2, 10, 50, 100, 500, 10000$. (Recall that k is the number of tosses.) What do you see? Is this consonant with what you had observed in earlier plots?

Solutions: For large values of k, the slope becomes very sharp. This makes sense because as k increases, more runs are clustered around the half-heads, half-tails range. For large values of k, we don't see any run at the beginning, then suddenly we see most of the runs when q is around 0.5, leading to a sharp increase.
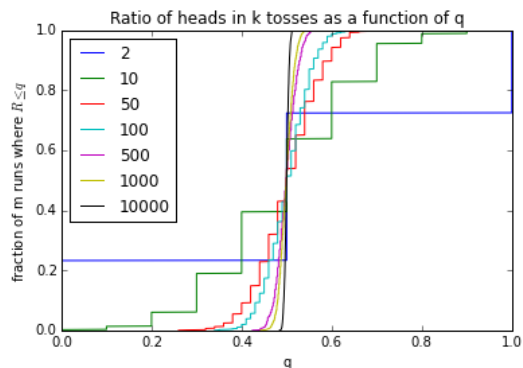
(e) Now, think about rescaling the plots in the previous parts to see if there is something common about this shape. For each k, read off the q values where the curves seem to cross horizontal lines at $0.25, 0.5, 0.75$. Call these the quartile markers. Compute these qs for your experiment. Plot them as a function of k. What do you observe?

Solutions: As k increases, the quartiles get closer and closer to 0.5. This suggests that the heads/tails distribution tend toward $50/50$. Again, this is consonant with what we observed in the previous part.

q quartile markers as a function of k
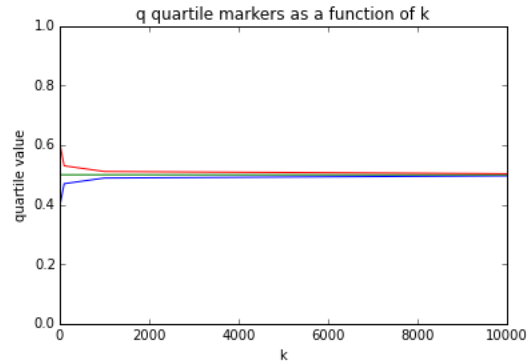
(f) Notice that the gap between the 0.75 marker and the 0.25 marker is getting smaller as k gets larger. Notice also that the 0.5 marker seems to be sticking around $q = \frac{1}{2}$. As a scientific problem, suppose you wanted to discover how indeed this was scaling with k.

Plot the distance between the 0.25 and 0.75 marker as a function of k as a scatter plot. Try all of the traditional axes combinations: linear-linear, log-linear, linear-log, and log-log. Which one seems to offer some insight?

Hint: In Python, the plotting functions with the aforementioned axes combinations are `plt.plot`, `plt.semilogx`, `plt.semilogy`, and `plt.loglog`.

Solutions: The plots all look pretty random, except for the loglog one, which somewhat resembles a straight line.
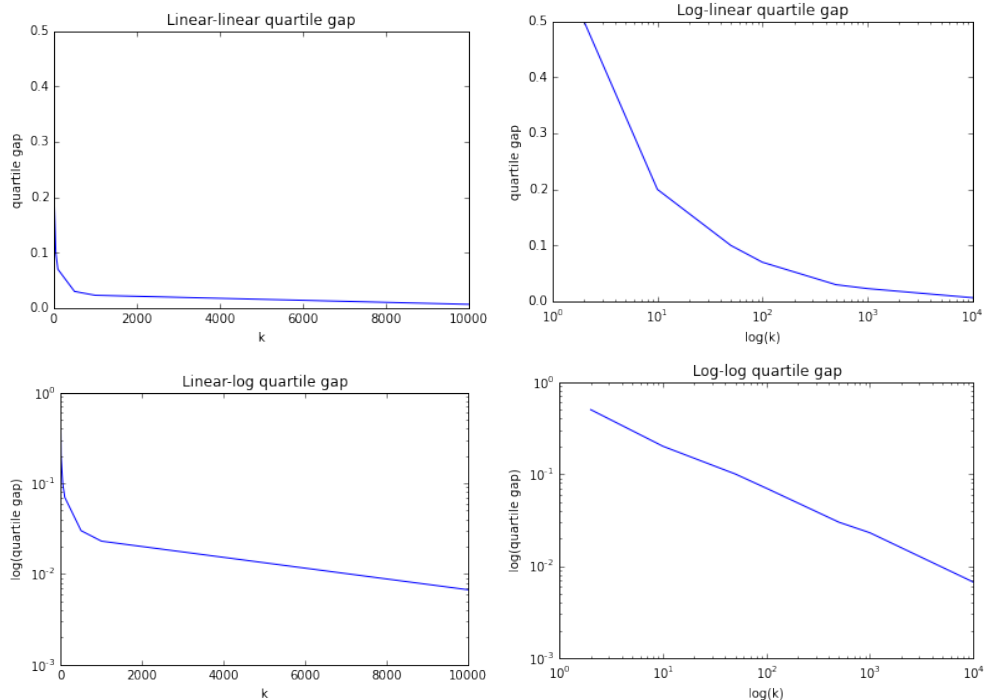
(g) Based on what you observed in the previous set of plots, conjecture a scaling rule that lets you calculate the gap between the 0.75 marker and the 0.25 marker as a function of k for the fair coin tosses case. Explain your derivation in your writeup.

Use this rule to rescale the horizontal axis of the plots from three parts ago. What do you now observe

about the curves for different values of k? By construction, they should be very close to each other in terms of where they are crossing $0.25, 0.5, 0.75$, but what about elsewhere?

Hint: Implement the function `q_curve_norm`, which does the same as `q_curve`, except now every point is normalized using your scaling rule.

Think about the previous part and how it can help you come up with a scaling rule.

Solutions: First, we know that the distance between the 0.75 marker and the 0.25 marker is just twice the distance between the 0.75 and the 0.5 marker, since they are symmetric, based on part (e).

We observe that the slope of the line in the log-log scale is approximately 0.5 (it moves two powers of x before changing one power of y), which implies that in a normal scale, $y \approx \frac{1}{\sqrt{x}}$.

Therefore, to scale the gap, we first subtract 0.5 to remove the symmetry, scale by the square root of k, and then add 0.5 back to the final result to adjust for the subtraction earlier.

Scaled ratio of heads in k tosses as a function of q
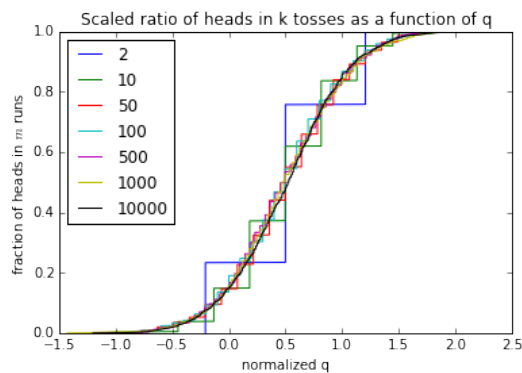
(h) In the skeleton, you will find a simple implementation that simulates the Monty Hall problem. There are n doors, and only one contains the prize. The contestant first picks a door, and then the host Monty will open all $n-2$ doors that don't contain the prize. The contestant is then given a choice to switch or stay with his current choice.

Your task will be to simulate 10000 trials of the Monty Hall problem. What is the probability of winning when the contestant switches? How about when he/she stays? Does it match your expectation and what was described in lecture and the lecture note? Please report the result in your answer.

Solutions: A sample output from running 10000 trials.

Simulating 10000 trials.
Switching won 6596 times out of 10000 (65.96% of the time)
Staying won 3404 times out of 10000 (34.04% of the time)

This matches our expectation that the contestant wins via switching with probability $\frac{2}{3}$, and wins via staying with probability $\frac{1}{3}$.

Reminder: When you finish, don't forget to convert the notebook to pdf and merge it with your written homework. Please also zip the `ipynb` file and submit it as `hw9.zip`.

3. **To Be Fair**

Suppose you have a biased coin with $P(heads) \neq 0.5$. How could you use this coin to simulate a fair coin? (Hint: Think about pairs of tosses.)

Solutions: Let H be the event that the coin flip comes up heads and T be the event that the coin flip comes up tails. We do not know the probability of getting heads since we've been told the coin is biased, but let's denote it by the variable p (i.e., $p = P(H)$).

We conduct a mini experiment in which we flip the same coin twice. Let HT be the event that we get heads and then tails in two consecutive flips. Similarly, let TH be the event that we get tails and then heads in two consecutive flips. Also, we'll let HH be the event that we get two heads in a row, and TT be the event that we get two tails in a row. Our sample space for the experiment is then $\Omega = \{HH, HT, TH, TT\}$.

Since we know $P(H) = p$, we can write the probabilities of the events in our sample space as follows:

$$P(HH) = P(\text{First toss H})P(\text{Second toss H}|\text{First toss H})$$

Because the nature of coins tells us that the second toss doesn't remember the first toss so

$$P(HH) = P(H)P(H) = p^2.$$

Similarly,

$$P(HT) = P(\text{First toss H})P(\text{Second toss T}|\text{First toss H}) = P(H)(1 - P(H)) = p(1 - p),$$
$$P(TH) = P(\text{First toss T})P(\text{Second toss H}|\text{First toss T}) = (1 - P(H))P(H) = p(1 - p),$$
$$P(TT) = P(\text{First toss T})P(\text{Second toss T}|\text{First toss T}) = (1 - P(H))(1 - P(H)) = (1 - p)^2.$$

We notice that the probability $P(HT)$ and the probability $P(TH)$ are equal, i.e., they are both $p(1 - p)$. By symmetry, these two probabilities must be the same since the coin doesn't know what it came up before. Since these two are the same, we can simply condition on the fact that something was returned to get that the resulting simulated coin toss is fair.

Therefore, we can simulate a fair coin using the following process. We toss the coin twice. If the outcome turns out to be heads both times or tails both times, we throw away the result and repeat the whole process again. Otherwise, if the outcome is HT, we return "heads", and if the outcome is TH, we return "tails".

How do we know this procedure will return a result at all? How do we know that this can't go on forever? We are not yet in a position to prove this rigorously because we haven't built the tools yet. But intuitively, this would require getting an infinite sequence of HH's or TT's. Why would the coin always agree with itself? This seems like it defies the nature of coin tossing: that the coin doesn't know what it did before.

4. **Intuition vs. Logic**

(a) I have a bag containing either a \$5 bill (with probability $1/3$) or a \$10 bill (with probability $2/3$). I then add a \$5 bill to the bag, so it now contains two bills. The bag is shaken, and you randomly draw a bill from the bag without looking into the bag. Suppose it turns out to be a \$5 bill. If a second student draws the remaining bill from the bag, what is the probability that it, too, is a \$5 bill? Show your calculations.

Solutions: Let A denote the event that the bag originally contained 5 dollars and let B denote the event that the bag originally contained 10 dollars. We are given that $P(A) = 1/3$ and $P(B) = 2/3$.

Now, we add another 5 dollar bill to the bag, and perform an experiment. You draw the first bill, and another student draws the second bill. Let A_1 be the event that the first bill is \$5, let A_2 be the event that the second bill is \$5, let B_1 be the event that the first bill is \$10 and let B_2 be the event that the second bill is \$10. Our sample space for this experiment is $\Omega = \{(A_1, A_2), (A_1, B_2), (B_1, A_2), (B_1, B_2)\}$. Note that $A = A_1 \cap A_2$ and $B = (B_1 \cap A_2) \cup (A_1 \cap B_2)$.

The question asks us to find the conditional probability $P(A_2|A_1)$. We know by the definition of conditional probability that

$$P(A_2|A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)}. \tag{1}$$

The only way that both the first and second bill are both 5 dollar bills (the event $A_1 \cap A_2$) is if the bag originally contained a 5 dollar bill. By the problem statement, we know the probability that the bag originally contained a 5 dollar bill is 1/3. Therefore, $P(A_1 \cap A_2) = \frac{1}{3}$.

We now look at the denominator $P(A_1)$. To calculate the marginal probability of event A_1, we use the law of total probability. This law states that

$$P(A_1) = P(A_1|A) \cdot P(A) + P(A_1|B) \cdot P(B) \tag{2}$$

for two disjoint events A and B.

Let's first look at $P(A_1|A)$. This is the probability that the first bill drawn is a 5 dollar bill, given that the bag initially contained a 5 dollar bill. Since both bills inside the bag are now 5 dollar bills, this probability is just 1! How about the second term, $P(A_1|B)$? This is the probability that the first bill drawn is a 5 dollar bill, given that the bag initially contained a 10 dollar bill. Since the first draw has an equal chance of drawing either bill, this probability is just 1/2.

Plugging in the appropriate terms in (2), we get $P(A_1) = 1 \cdot \left(\frac{1}{3}\right) + \left(\frac{1}{2}\right) \cdot \left(\frac{2}{3}\right) = \frac{2}{3}$. Then plugging into the conditional probability formula (1), we have $P(A_2|A_1) = \frac{\left(\frac{1}{3}\right)}{\left(\frac{2}{3}\right)} = \boxed{\frac{1}{2}}$.

(b) Your gambling buddy found a website where he could buy trick coins that are either heads or tails on both sides. He puts three coins into a bag: one coin that is heads on both sides, one coin that is tails on both sides, and one that is heads on one side and tails on the other side. You shake the bag, draw out a coin at random, put it on the table without looking at it, then look at the side that is showing. Suppose you notice that the side that is showing is heads. What is the probability that the other side is heads? Show your calculations.

Solutions: The experiment in this case is that we pull 1 of 3 coins out of the bag and flip it. Let A_1 be the event that the side face-up is heads and let B_1 be the event that the side face up is tails. Let A_2 be the event that the side face-down is heads and let B_2 be the event that the side face down is tails. Our sample space for this experiment is $\Omega = \{(A_1, A_2), (A_1, B_2), (B_1, A_2), (B_1, B_2)\}$.

Moreover, let H be the event the coin has both sides heads; i.e., $H = A_1 \cap A_2$. Let T be the event the coin has both sides tails; i.e., $T = B_1 \cap B_2$. Let F be the event the coin has one side heads and one side tails; i.e., $F = (A_1 \cap B_2) \cup (B_1 \cap A_2)$. We assume that there is a uniform probability of any of the 3 coins being chosen from the bag, i.e., $P(H) = P(T) = P(F) = 1/3$.

The problem wants us to find $P(A_2|A_1)$, or the probability that the face-down side is heads given that the face-up side is heads. We can again use the conditional probability formula

$$P(A_2|A_1) = \frac{P(A_2 \cap A_1)}{P(A_1)}. \tag{3}$$

Similarly to part (a), we use the law of total probability to find $P(A_1)$. We obtain

$$P(A_1) = P(A_1|H) \cdot P(H) + P(A_1|T) \cdot P(T) + P(A_1|F) \cdot P(F). \tag{4}$$

Now, let's look at the numerator. $P(A_2 \cap A_1)$ denotes the event both face-up and face-down sides are heads. This event is equivalent to event H. Thus, $P(A_2 \cap A_1) = P(H) = 1/3$.

We now look at each term in the denominator. $P(A_1|H)$ is the event that the face-up side is a head, given the coin has both heads on each side. This is just 1! Similarly, $P(A_1|T)$ is 0 since if the coin has both tails on each side, there is no way we will get a head. Now $P(A_1|F)$ is the probability that we get a head facing up, given that we have a fair coin. This is just $1/2$ since flipping a head or a tail is equally likely.

Thus plugging in all the terms into (4), we have

$$P(A_1) = P(A_1|H) \cdot P(H) + P(A_1|T) \cdot P(T) + P(A_1|F) \cdot P(F) = 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2}.$$

We can then plug all terms into (3) to compute

$$P(A_2|A_1) = \frac{P(A_2 \cap A_1)}{P(A_1)} = \frac{\left(\frac{1}{3}\right)}{\left(\frac{1}{2}\right)} = \boxed{\frac{2}{3}}.$$

5. Playing Pollster

As an expert in probability, the staff members at the Daily Californian have recruited you to help them conduct a poll to determine the percentage p of Berkeley undergraduates that plan to participate in the student sit-in. They've specified that they want your estimate $\hat{p}$ to have an error of at most ε with confidence $1 - \delta$. That is,
$$P(|\hat{p} - p| \leq \varepsilon) \geq 1 - \delta.$$

Assume that you've been given the bound

$$P(|\hat{p} - p| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2},$$

where n is the number of students in your poll.

(a) Using the formula above, what is the smallest number of students n that you need to poll so that your poll has an error of at most ε with confidence $1 - \delta$?

Solutions: We know we need to have

$$P(|\hat{p} - p| \leq \varepsilon) \geq 1 - \delta.$$

Subtracting both sides from 1, it follows that we must have

$$P(|\hat{p} - p| > \varepsilon) \leq \delta.$$

Therefore if we choose n such that

$$\frac{1}{4n\varepsilon^2} \leq \delta,$$

we will have

$$P(|\hat{p} - p| \geq \varepsilon) \leq \delta,$$

and since $P(|\hat{p} - p| > \varepsilon) \leq P(|\hat{p} - p| \geq \varepsilon)$, this will meet the requirement that

$$P(|\hat{p} - p| > \varepsilon) \leq \delta.$$

Thus we must have that

$$\frac{1}{4n\varepsilon^2} \leq \delta$$
$$\frac{1}{n} \leq 4\varepsilon^2\delta$$
$$n \geq \frac{1}{4\varepsilon^2\delta}.$$

(b) At Berkeley, there are about 26,000 undergraduates and about 10,000 graduate students. Suppose you only want to understand the frequency of sitting-in for the undergraduates. If you want to obtain an estimate with error of at most 5% with 98% confidence, how many undergraduate students would you need to poll? Does your answer change if you instead only want to understand the frequency of sitting-in for the graduate students?

Solutions: Plugging in to the bound you found above, you get that $n \geq 5000$. The answer is the same for graduate students; the size of the population does not affect the number of samples you need.

(c) It turns out you just don't have as much time for extracurricular activities as you thought you would this semester. The writers at the Daily Californian insist that your poll results are reported with at least 95% confidence, but you only have enough time to poll 500 students. Based on the bound above, what is the worst-case error with which you can report your results?

Solutions: If you only have time to poll 500 people and want to report your results with 95% confidence, you must report that the error in your estimate is at most 10%. You can find this by plugging in $\frac{1}{4\cdot500\cdot\varepsilon^2} = .05$ and solving for ε.

6. Blood Type

Consider the three alleles, A, B, and O, for human blood types. As each person inherits one of the 3 alleles from each parent, there are 6 possible genotypes: AA, AB, AO, BB, BO, and OO. Blood groups A and B are dominant to O. Therefore, people with AA or AO have type A blood. Similarly, BB and BO result in type B blood. The AB genotype is called type AB blood, and the OO genotype is called type O blood. Each parent contributes one allele randomly. Now, suppose that the frequencies of the A, B, and O alleles are 0.4, 0.25, and 0.35, respectively, in Berkeley. Alice and Bob, two residents of Berkeley are married and have a daughter, Mary. Alice has blood type AB.

(a) What is the probability that Bob's genotype is AO?

Solutions: Let B_{1A}, B_{1B} and B_{1O} be the events that Bob's first allele is A, B, and O, respectively. Let B_{2A}, B_{2B} and B_{2O} be the events that Bob's second allele is A, B, and O respectively. Bob's blood type can be AA, AB, AO, BB, or BO. Let B_{AA} be the event that Bob has type AA blood, B_{AB} be the event that Bob has type AB blood, B_{AO} be the event that Bob has type AO blood, B_{BB} be the event that Bob has type BB blood, and B_{BO} be the event that Bob has type BO blood. The sample space is $\Omega = \{B_{AA}, B_{AB}, B_{AO}, B_{BB}, B_{BO}\}$. Note that we have

$$B_{AA} = B_{1A} \cap B_{2A}$$
$$B_{AB} = (B_{1A} \cap B_{2B}) \cup (B_{1B} \cap B_{2A})$$
$$B_{AO} = (B_{1A} \cap B_{2O}) \cup (B_{1O} \cap B_{2A})$$
$$B_{BB} = B_{1B} \cap B_{2B}$$
$$B_{BO} = (B_{1B} \cap B_{2O}) \cup (B_{1O} \cap B_{2B}).$$

Since the first allele and second allele don't know about each other, the occurrence of the first allele will not affect the second, and vice versa. Therefore, using the rules that $P(A \cap B) = P(A|B)P(B)$ and $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,

$$P(B_{AA}) = P(B_{1A}|B_{2A})P(B_{2A}) = P(B_{1A})P(B_{2A}) = (.4)(.4) = .16$$
$$P(B_{AB}) = P(B_{1A}|B_{2B})P(B_{2B}) + P(B_{1B}|B_{2A})P(B_{2A}) - P((B_{1A} \cap B_{2B}) \cap (B_{1B} \cap B_{2A}))$$
$$= P(B_{1A})P(B_{2B}) + P(B_{1B})P(B_{2A}) = 2(.4)(.25) = .2$$
$$P(B_{AO}) = P(B_{1A}|B_{2O})P(B_{2O}) + P(B_{1O}|B_{2A})P(B_{2A}) - P((B_{1A} \cap B_{2O}) \cap (B_{1O} \cap B_{2A}))$$
$$= P(B_{1A})P(B_{2O}) + P(B_{1O})P(B_{2A}) = 2(.4)(.35) = .28$$
$$P(B_{BB}) = P(B_{1B}|B_{2B})P(B_{2B}) = P(B_{1B})P(B_{2B}) = (.25)(.25) = .0625$$
$$P(B_{BO}) = P(B_{1B}|B_{2O})P(B_{2O}) + P(B_{1O}|B_{2B})P(B_{2B}) - P((B_{1B} \cap B_{2O}) \cap (B_{1O} \cap B_{2B}))$$
$$= P(B_{1B})P(B_{2O}) + P(B_{1O})P(B_{2B}) = 2(.25)(.35) = .175.$$

Therefore $P(B_{AO}) = .28$.

(b) Assume that Bob's genotype is AO. What is the probability that Mary's blood type is AB?

Solutions: Since Alice has type AB and Bob has type AO, the sample space of possible genotypes for Mary is $\{AA, AO, AB, BO\}$. Since there is uniform probability of inheriting either allele from a given parent, there is a 1/4 chance that Mary will have type AB blood.

(c) Assume Mary's blood type is AB. What is the probability that Bob's genotype is AA?

Solutions: Bob's blood type can be AA, AB, AO, BB, or BO. As in part (a), let B_{AA} be the event that Bob has type AA blood, B_{AB} be the event that Bob has type AB blood, B_{AO} be the event that Bob has type AO blood, B_{BB} be the event that Bob has type BB blood, and B_{BO} be the event that Bob has type BO blood. We already computed the probability of Bob having these blood types in part (a):

$$Pr(B_{AA}) = .16$$
$$Pr(B_{AB}) = .2$$
$$Pr(B_{AO}) = .28$$
$$Pr(B_{BB}) = .0625$$
$$Pr(B_{BO}) = .175.$$

Now, let the event that Mary has blood type AB be M_{AB}. The problem asks us to find $Pr(B_{AA}|M_{AB})$. We can compute this using Bayes' formula, which says that

$$Pr(B_{AA}|M_{AB}) = \frac{Pr(M_{AB}|B_{AA}) \cdot Pr(B_{AA})}{Pr(M_{AB})}.$$

To find $Pr(M_{AB})$, we can use the Law of Total Probability, which says that

$$Pr(M_{AB}) = Pr(M_{AB}|B_{AA}) \cdot Pr(B_{AA}) + Pr(M_{AB}|B_{AB}) \cdot Pr(B_{AB}) + Pr(M_{AB}|B_{AO}) \cdot Pr(B_{AO})$$
$$+ Pr(M_{AB}|B_{BB}) \cdot Pr(B_{BB}) + Pr(M_{AB}|B_{BO}) \cdot Pr(B_{BO}).$$

To calculate this, we must find the conditional probabilities that Mary has AB blood given Bob's blood type. Recall that Alice has type AB blood.

- If Bob has AA blood, the possible combinations of their alleles are AA, AA, AB, and AB, so $Pr(M_{AB}|B_{AA}) = 1/2$.

- If Bob has AB blood, the possible combinations of their alleles are AA, AB, AB, and BB, so $Pr(M_{AB}|B_{AB}) = 1/2$.
- If Bob has AO blood, the possible combinations of their alleles are AA, AO, AB, and BO, so $Pr(M_{AB}|B_{AO}) = 1/4$.
- If Bob has BB blood, the possible combinations of their alleles are AB, AB, BB, and BB, so $Pr(M_{AB}|B_{BB}) = 1/2$.
- If Bob has BO blood, the possible combinations of their alleles are AB, AO, BB, and BO, so $Pr(M_{AB}|B_{BO}) = 1/4$.

We now have all the information we need to plug in and solve. By the Law of Total Probability above, we have

$$Pr(M_{AB}) = Pr(M_{AB}|B_{AA}) \cdot Pr(B_{AA}) + Pr(M_{AB}|B_{AB}) \cdot Pr(B_{AB}) + Pr(M_{AB}|B_{AO}) \cdot Pr(B_{AO})$$
$$+ Pr(M_{AB}|B_{BB}) \cdot Pr(B_{BB}) + Pr(M_{AB}|B_{BO}) \cdot Pr(B_{BO})$$
$$= (.5)(.16) + (.5)(.2) + (.25)(.28) + (.5)(.0625) + (.25)(.175)$$
$$= .325,$$

and plugging in to Bayes' formula, we find that

$$Pr(B_{AA}|M_{AB}) = \frac{Pr(M_{AB}|B_{AA}) \cdot Pr(B_{AA})}{Pr(M_{AB})} = \frac{(.5)(.16)}{.325} = .246.$$

7. **Midterm question 3**

   Re-do midterm question 3.

8. **Midterm question 4**

   Re-do midterm question 4.

9. **Midterm question 5**

   Re-do midterm question 5.

10. **Midterm question 6**

    Re-do midterm question 6.

11. **Midterm question 7**

    Re-do midterm question 7.

12. **Midterm question 8**

    Re-do midterm question 8.

13. **Midterm question 9**

    Re-do midterm question 9.

14. **Midterm question 10**

    Re-do midterm question 10.

15. **Midterm question 11**

    Re-do midterm question 11.

16. **Midterm question 12**

    Re-do midterm question 12.

17. **Midterm question 13**

    Re-do midterm question 13.

18. **Midterm question 14**

    Re-do midterm question 14.

19. **Write your own problem**

    Write your own problem related to this week's material and solve it. You may still work in groups to brainstorm problems, but each student should submit a unique problem. What is the problem? How to formulate it? How to solve it? What is the solution?