Probability Notes

Examples

Example 1: A man has six different pairs of socks, from which he selects two socks at random. What is the probability that the selected socks will match?

Answer: $\frac{6 \cdot 1}{C(12,2)} = \frac{6}{66} = \frac{1}{11}$

I usually like to begin with counting the denominator possibilities, for often that is the easier concept to count (the sample space; all possible outcomes). Here the denominator or count of the sample space is to count all of the ways that we can pull two socks out of the drawer. When I pick a pair of socks I don't care about the order of the draw, I simply want a subset of size two, so combinations is the way to go. That is I see this mathematically as counting all of the subsets of size two that are possible from a set with 12 elements. Now in the numerator we only want to count those match-ups that are actual pairs of socks. So I thought about this as a sequence of picks where the first draw may be from any of six pairs followed by only one choice for the second draw, for it must match the pair that I drew from.

The author took order into account and came up with this way of computing the probability: $\frac{12 \cdot 1}{12 \cdot 11}$. In the denominator we are counting the number of ways that we can pick two socks out of the drawer, but we are counting order here, so a blue sock followed by picking a red sock, is considered different from picking a red sock followed by a blue sock. In my way of thinking these two would not be considered different subsets of size two, and thus only counted once. But as long as you are consistent you can sometimes count using subsets (combinations) or ordered pairs (permutations) and arrive at the same probability. Simply count the numerator and denominator using the same concept. Don't mix your metaphors or conceptualizations. Go with the method that makes the most sense to you. In the numerator under the author's scheme we look at picking any of twelve socks on the first pick, but we only have one choice (the mate) on the second draw.

There is even a third way that you could look at this problem in the larger picture of chapter 6. Take another look at the author's answer. We could rewrite this as $\frac{12 \cdot 1}{12 \cdot 11} = \frac{12}{12} \cdot \frac{1}{11} = 1 \cdot \frac{1}{11}$. These could be seen as a product of probabilities, which means that we could conceptualize this as the probability of the intersection of two events. We might look at this problem from this vantage point as the probability of drawing a sock on the first draw and drawing its mate on the second. When we find probabilities of intersections ("and") of two events we must consider if the events are independent or not. I they are independent then the probability of the intersection becomes simply the product of the probabilities of the individual events. If they are not independent events, then we must involve a conditional probability. In our case the probability of the second event "drawing the mate of the first draw" is definitely dependent upon the first event "drawing a sock," so we have to apply the general multiplication rule: $P(A \cap B) = P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B)$. I wrote it both ways because I don't think that authors emphasize this point enough. You have the power to choose which event to take first! Pick the one that makes the probability easiest to compute.

So we can look at this problem as computing the probability of "drawing a sock" and "drawing its mate," which becomes the probability of "drawing a sock" times the probability of "drawing its mate" given that we "drew a sock" under the multiplication principle. The probability of drawing a sock on the first draw is 12 choices over 12 choices or 1. This is a certain event. The conditional probability is dependent upon our first selection, so we have only 1 choice out of the 11 potential picks left.

This may seem an outrageous viewpoint, but it drives home the idea that you have the power to see any of these word problems in a variety of ways. You may choose to use "brute force" to count the numerator and denominator, you may use counting principles to find the probability, or you may use the addition or multiplication formulas (conceptualizing your complex events as unions or intersections of simpler events), or you can find your probability going backdoor [1 – P(of what you don't want)].

Example 2: A man, a woman, and their three children randomly stand in a row for a family picture. What is the probability that the parents will be standing next to each other?

Answer: $\dfrac{4 \cdot 2!3!}{5!}$

The denominator counts all of the ways that the five family members can be arranged. In the numerator the factor of 4 represents all of the ways that the two parents can be side by side. To see these fours ways, here is a picture of how I see the possibilities:

The 2! Represents the number of ways that the parents can be arranged once we have them placed, and the 3! Represents the number of ways that we can arrange the children in the remaining 3 places in the line.

Example 3: The figure below shows a partial map of the streets in a certain city. A tourist starts at pint A and selects at random a path to point B. (We shall assume that he walks only south and east.)

(a) He passes through point C and point D.

Answer: $\frac{C(3,2)\cdot C(2,2)\cdot C(3,1)}{C(8,5)} = \frac{9}{56} \approx .161$

The tourist has to go East 5 blocks and South 3 blocks for a total of 8 blocks, so the denominator counts all of the ways the tourist can go from point A to point B. You could just as easily have C(8,3) in the denominator as well (it's the same number as C(8,5)). The idea is that we are choosing the places in the sequence of eight choices that the tourist moves East in the case of C(8,5), and South in the case of C(8,3). By default the remaining places are filled in by the other direction, so only one choice for those as well. Again I think to get the overall picture we should describe the count in the denominator as $C(8,5)\cdot 1\cdot 1\cdot 1\cdot 1\cdot 1\cdot 1\cdot 1\cdot 1$. This emphasizes that the combination only dictates where the E's go in the sequence of eight letters, it does not count the number of sequences of E's and S's for a given configuration of so many E's and S's. That is the job of the eight slots after the combination; however, since there is only one choice for each place the count collapses to simply C(8,5). If we allowed for more directions, then the places (slots) that do not go East could have more choices for them. This is similar to the coin flipping problems and die rolling problems in chapter 5. Now in the numerator we have to account for going through points C and D along the way to point B, so the C(3,2) counts the ways to go from A to C, by choosing the two places where the 2 E's go in the sequence of 3 blocks that must be traveled, and to account for the number of ways to go from C to D we have C(2,2). There is only one way to go East two blocks. Finally to finish we must count the ways to go from D to B; thus the final factor of C(3,1). That is the number of ways to place one E in the sequence of three letters. These are all multiplied because the number of choices for each successive point-to-point movement remains the same regardless of the previous choices made. That is regardless of the blocks the tourist travels from point A to C, the tourist still has C(2,2) choices to go from C to D (think tree structure, regardless of the branch taken at one stage, all branches from that stage separate into the same number of following branches). Or in terms of verbalizing the situation, to get from point A to point B we must choose a route from A to C and from C to D and from D to B. We join the point-to-point choices using the word "and." Using the word "or" would allow us to count only going from point A to C and stopping. It doesn't force the entire journey.

(b) He passes through point C or point D.

Answer: $\frac{C(3,2)\cdot C(5,3)+C(5,4)\cdot C(3,1)-C(3,2)\cdot C(2,2)\cdot C(3,1)}{C(8,5)} = \frac{36}{56} = \frac{9}{14} \approx .643$

The author uses the word "or" here, so we can count the number of ways to go from point A to B via point C separately from counting the number of ways to go from point A to B via point D, and then add these two counts together; however, whenever we add we have to be careful about double counting when the events are not mutually exclusive (the old disjoint set idea from chapter 5, but called "mutually exclusive" in probability language). In our case the two events, "going from A to B via C" and "going from A to B via D" are not mutually exclusive, for some of the A to B via D routes go through C, so we have to subtract out the "and" case, which is the routes that go from A to B via C and D. These routes are double counted. That is the structure of our numerator above. The $C(3,2)\cdot C(5,3)$ part counts the routes from A to B via C, the $C(5,4)\cdot C(3,1)$ counts the number of routes from A to B via D, and $C(3,2)\cdot C(2,2)\cdot C(3,1)$ counts the number of routes from A to B via both C and D This last part is what we computed in 27c above.

Notice that we could restructure the answer as $\frac{C(3,2)\cdot C(5,3)}{C(8,5)}+\frac{C(5,4)\cdot C(3,1)}{C(8,5)}-\frac{C(3,2)\cdot C(2,2)\cdot C(3,1)}{C(8,5)}$. This views the problem from the point of finding the probability of the union of two events: the event that the tourist travels from A to B via C or the event that the tourist travels from A to B via D. Again because the two events are not mutually exclusive we must subtract out the intersection of the two events (the overlap of the two events, or the "and" case). Seen in this light the first fraction above gives us the probability of taking a route from A to B via C, the second fraction yields the probability of going from A to B via D, and the third fraction represents the probability of going from A to B via both C and D. If a probability seems too overwhelming try breaking it into parts (subsets of the event to be computed) and sum the parts. If you can break the event into mutually exclusive (non-overlapping) events, then you don't have to worry about double counting and needing to subtract out the overlap. This is called partitioning the set or in our case the event. I see it visually this way:

Where S is the sample space, A is the event space, and E, F, and G form a partitioning of A (that is they are non-overlapping subsets of A, that when unioned together form A). Since E, F, and G are mutually exclusive we have $P(A)=P(E)+P(F)+P(G)$. So if you want to compute the probability of a highly complicated event, you may have an easier time if you can think of a way to partition your event. This is the strategy that we developed for the "at least" phrase. Rather than compute the probability of the entire phrase we busted it into its "exactly" sub pieces that were all mutually exclusive, and so could simply be summed together.

Example 4: In the Illinois Lottery Lotto game, the player chooses six different integers from 1 to 40. If the six match (in any order) the six different integers drawn by the lottery, the player wins the grand prize jackpot, which stars at $1 million and grows weekly until won. Multiple winners split the pot equally. For each $1 bet, the player must pick two (presumably different) sets of six integers. What is the probability of winning the Illinois Lottery Lotto with a $1 bet?

Answer: $\frac{2}{C(40,6)}=\frac{2}{3838380}=\frac{1}{1919190}\approx .0000005211$

In the denominator we have all of the possible lottery choices, again we don't care about the order of the selections, only that these six numbers are together (subset idea; thus a combination). In the numerator there are two out of the 3,838,380 subsets of size six that can match the official draw, namely the two subsets that you make up when you buy the ticket.

Example 5: The U.S. Senate consists of two senators from each of the 50 states. Five senators are to be selected at random to form a committee. What is the probability that no two members of the committee are from the same state?

Answer: $\dfrac{100 \cdot 98 \cdot 96 \cdot 94 \cdot 92}{P(100,5)} \approx 0.900552$

This is one where you could look at including order or not including order in the selection. When I first tried to not include order, I found it very difficult, so I went back and looked at including order, and arrived at the above structure. The denominator counts all of the subsets of five senators that can be selected out of the entire Senate (100 senators); however, because I used permutations instead of combinations, order is implied on each of these subsets. Or another way to think about it using the multiplication principle is to visualize five slots, where you have 100 choices available for the fist slot, and then you have 99 choices for the second, then 98 for the third, then 97 for the fourth, and finally 96 choices for the fifth. Multiplied together yields the same as P(100,5). In the numerator, we must select and arrange the senators in such a way that no two from the same state are chosen. So using the multiplication principle and the slots idea again, we have 100 choices for the first pick, but only 98 for the second (for not only is the chosen senator not available, but neither is his/her companion senator from that state, and so on as seen above. So for each slot we reduce by two instead of one.

The author stayed with the unordered way of counting. Here is what they had for an answer: $\dfrac{\dfrac{100 \cdot 98 \cdot 96 \cdot 94 \cdot 92}{5!}}{C(100,5)}$. It yields the same probability as my structure. In the denominator they count the number of subsets of size five, but no order is put on the subset. In essence this keeps it a set idea whereas P(100,5) converts the subsets into 5-tuples (ordered pair idea rather than a set concept). This makes the numerator count more difficult, because the 100x98x96x94x92 implies order and we cannot mix our metaphors! So, how do we un-order the 100x98x96x94x92 count? In a similar way to the arrangements of non-distinct groups in chapter five we can remove the order by dividing by the number of ways to arrange our five senators – which can happen in 5! ways. That is why you see the 5! in the denominator portion of the numerator. Now we are counting the same types of structures in both the numerator and the denominator.

After further thought I did think of another way to count this problem using combinations:

$$\frac{\dfrac{C(50,1) \cdot C(2,1) \cdot C(49,1) \cdot C(2,1) \cdot C(48,1) \cdot C(2,1) \cdot C(47,1) \cdot C(2,1) \cdot C(46,1) \cdot C(2,1)}{5!}}{C(100,5)}$$

The C(50,1) picks a state, then the C(2,1) picks a senator from that state, then so on for the next four choices of state and a senator from that state. The only problem again is the multiplication implies order on this choosing process, so to remove the ordering of the senators I also had to divide by 5!.

There is another way to conceptualize this problem. If you reconfigure what I first had you get: $\frac{100}{100} \cdot \frac{98}{99} \cdot \frac{96}{98} \cdot \frac{94}{97} \cdot \frac{92}{96}$. This can be viewed as a sequence of probabilities that are multiplied, so as an earlier example we can see this as the probability of five events that are intersected ("and"-ed together in verbal form) where the events are dependent, so we must multiply by conditional probabilities. The probability of the first event is certain, for we have all 100 senators to choose from, so the probability of picking a senator from the Senate is 100/100 or in other words 1. The second event is choosing another senator from a different state, so we now have only 98 choices left out of the remaining 99 senators. For our next step we have the probability of picking another senator different from the first two, so this leaves us 96 choices out of 98 remaining senators, and so on.

Pick your poison as to how you would like to determine this probability! You may even have found another way. Try to find a visualization that works for you.

Example 6: What is the probability that in a group of 25 people at least one person has a birthday on June 13? Why is your answer different from the probability displayed in the table below for $r = 25$?

Answer: $1 - \frac{364^{25}}{365^{25}} \approx .066$

Backdoor approach is easiest here. Compute the probability of what we don't want, which is that no birthdays match June 13, and then subtract this from one. In the denominator you could conceptualize 25 slots, one for each person, and in each slot you have 365 choices for a birthday; thus, you have 365 multiplied by itself 25 times. In the numerator we also can conceptualize this with 25 slots, only here we have 364 choices for the birthday not to match June 13; thus, 364 multiplied by itself 25 times. What is different here from the regular birthday problem is that we only have one day not to match. With the regular birthday problem the numerator number dropped by one for each slot because we could not repeat any previous birthdays, so there were successively fewer days to pick from. This number is smaller because it is more difficult to get matches for one day as opposed to the greater possibilities in the regular birthday problem.

Example 7: Fred is having a dinner party and is limited to 10 guests. He has 16 men friends and 12 women friends, including Mary and Laura.

(a) If Fred chooses his guests at random, what is the probability that Mary and Laura are invited?

Answer: $\dfrac{C(2,2)\cdot C(26,8)}{C(28,10)} \approx 0.11905$

In the denominator we have the number of ways that Fred can choose any ten of his 28 friends. No order implied, so combinations are used to count (again think of it mathematically as subsets of size ten out of a set with 28 elements, that is what combinations count). The numerator must have two particular women in the selected group, so the C(2,2) picks them (only one way to do this task). Follow this with choosing the other 8 guests, which can be chosen from any of the remaining 26 guests (after Laura and Mary have been removed).

(b) If Fred decides to invite 5 men and 5 women, what is the probability that Mary and Laura are invited?

Answer: $\dfrac{C(2,2)\cdot C(10,3)\cdot C(16,5)}{C(16,5)\cdot C(12,5)} \approx 0.15152$

The denominator changes here because Fred is going to choose 5 men and 5 women, so we reduce down to the elements that we want to select for each factor. Thinking of an assembly line process, Fred could first pick the 5 men out of the 16 possible choices followed by selecting the five women out of 12 choices. Again, these are subsets of size 5, so we use combinations. C(16,5) counts all of the subsets of size 5 from 16 objects (men), and C(12,5) does the same for the women. We multiply because we verbalized it as 5 men ***and*** 5 women. The multiplication principal associates for each set of 5 men a set of 5 women. That is what the multiplication principle does. Here is a possible instance of what is being counted, structured mathematically: one such association out of the C(16,5)xC(12,5) possibilities is ({Joe, Tom, Dick, Harry, Benny}, {Rachael, Mary, Debbie, Katie, Amy}). It is an ordered pair with components that are sets with 5 elements, and there are 3,459,456 such structures for this problem, but that is how I see it mathematically, and the more that I can see the words of problems in this mathematical way, the easier it is for me to connect to the counting formulas. That is, there are only a few mathematical structures (basically subsets or ordered tuples) and thereby only a few counting techniques to count them, but there are a whole bunch of word problems, so if I can reduce word problems down to a few mathematical structures, then all of this detail can be broken down to just a few basic forms, and I can then match up the few counting techniques to the few mathematical structures that really exist. It makes life easier by generalizing all of the specific details, so that there is less to remember. This is a huge part of mathematics, generalizing, but it isn’t easy to cut through the details to see the general pattern. It takes practice and experience. Hopefully you folks will be able to begin that process in this course.

Example 8: In the American League, the East, Central, and West divisions consist of 5, 5, and 4 baseball teams, respectively. A sportswriter predicts that winner of each of the three divisions by choosing a team completely at random in each division. If the sportswriter eliminates from each division one team that clearly has no chance of winning and predicts a winner at random from the remaining teams, what is the writer's chance of predicting at least one winner? Assume that the eliminated teams don't end up surprising anyone.

Answer: $1-\frac{3\cdot3\cdot2}{4\cdot4\cdot3}=1-\frac{18}{48}=1-\frac{3}{8}=\frac{5}{8}=0.625$

The backdoor approach (using the complement) is easiest here. To compute what we don't want (no winners selected in any division) we start with the denominator where we count all of the ways that a team can be picked to win, 4x4x3 ways. In the numerator we want only those selections that do not win, so there are 3 choices of a non-winner in two of the divisions and 2 choices of non-winner in the other division; thus, the 3x3x2 numerator.

Example 9: A doctor studies the known cancer patients in a certain town. The probability that a randomly chosen resident has cancer is found to be .001. It is found that 30% of the town works for Ajax Chemical Company. The probability that an employee of Ajax has cancer is equal to .003. Are the events "has cancer" and "works for Ajax" independent of one another?

There are a couple of ways to check for independence between events. One way is to see if $P(A\cap B)=P(A)\cdot P(B)$, and another way is to check to see if $P(A\mid B)=P(A)$ or $P(B\mid A)=P(B)$. I see the conditional check as meaning that having knowledge about an event does not change the likelihood of the other event. For example, with $P(A\mid B)=P(A)$, having knowledge about B doesn't change our perception of the likelihood of A. With or without the knowledge the probability is the same. In essence I see this as meaning that the proportion of event A in B is the same as A is in S, so independence is a proportion idea. Mutually exclusive events are a whole different concept that deals with the intersection of two events being empty (disjoint sets). The proportion idea can be seen from the definitions: $P(A\mid B)=\frac{n(A\cap B)}{n(B)}$. This is the reduced sample space viewpoint. Since we know that we are in B the likelihood of A is that part of A that is in B relative to B. The probability of A is defined as $P(A)=\frac{n(A)}{n(S)}$, where S is the sample space. So, if A and B are independent, then $P(A\mid B)=P(A)$ implies that $\frac{n(A\cap B)}{n(B)}=\frac{n(A)}{n(S)}$. This means that A has to be distributed in B in the same proportion as A is in S, the sample space, so having knowledge about B doesn't tell us anything or change the likelihood of A. The knowledge doesn't make A more likely or less likely.

So one way to check for independence here is to check to see if the P(cancer) = P(cancer | works for Ajax). Does the knowledge that a worker works for Ajax change the likelihood of having cancer? P(cancer) = 0.001 whereas P(cancer | works for Ajax) = P(cancer and works for Ajax) / P(works for Ajax) = 0.003 / 0.3 = 0.01. So the two events are not independent, for the conditional probability is different from the probability of the single event. In this case, having knowledge that a worker works for Ajax increases the likelihood of having cancer.

The other way to test it is to see if P(cancer and works for Ajax) = P(cancer) x P(works for Ajax). Again this shows that the two events are not independent, for 0.003 is not equal to (0.001)(0.3) = 0.0003.

Example 10: The probabilities that a person A and a person B will live an additional 15 years are .8 and .7, respectively. Assuming that their lifespans are independent, what is the probability that A or B will live an additional 15 years?

In the last problem we checked to see if we had independence of events, sometimes we know or assume that we have independent events and use it to help us compute another probability. That is the case here. We want to find P(A lives an extra 15 years or B lives an extra 15 years). Since we do not know if the events are mutually exclusive (remember just because you have independence does not mean that you have mutually exclusive events, and vice versa) we must use the longer formula to compute this probability. We have P(A lives an extra 15 years or B lives an extra 15 years) = P(A lives an extra 15 years) + P(B lives an extra 15 years) – P(A lives an extra 15 years and B lives an extra 15 years) = 0.8 + 0.7 – (0.8)(0.7) = 0.94. Notice how we used independence to compute the "and" case. Since we had independence all I had to do is multiply the probabilities of the two events, I did not have to use any conditional probabilities like I would have if we did not have independence.

Example 11: Use the inclusion-exclusion principle for (nonconditional) probabilities to show that if E, F, and G are events in S, then $P(E \cup F \mid G) = P(E \mid G) + P(F \mid G) - P(E \cap F \mid G)$

$$
\begin{aligned}
P(E \cup F) \mid G) &= \frac{P[(E \cup F) \cap G)}{P(G)} = \frac{P[(E \cap G) \cup (F \cap G)]}{P(G)} \\
&= \frac{P(E \cap G) + P(F \cap G) - P[(E \cap G) \cap (F \cap G)]}{P(G)} \\
&= \frac{P(E \cap G) + P(F \cap G) - P(E \cap F \cap G)}{P(G)} \\
&= \frac{P(E \cap G)}{P(G)} + \frac{P(F \cap G)}{P(G)} - \frac{P[(E \cap F) \cap G]}{P(G)} \\
&= P(E \mid G) + P(F \mid G) - P(E \cap F \mid G)
\end{aligned}
$$

All done with smoke and mirrors. Not really, simply applying definitions to get to the next step until you get the final desired result. Isn't math fun?

Example 12: Out of 250 third-grade boys, 120 played baseball, 140 played soccer, and 50 played both. Find the probability that a boy chosen at random did not play soccer, given that he did not play baseball.

A Venn diagram may be a helpful tool for organizing the probabilities in this problem.

$$P(S' \mid B') = \frac{P(S' \cap B')}{P(B')} = \frac{0.16}{0.36 + 0.16} \approx 0.31$$

Where S represents the boys who play soccer and B represents the boys who play baseball.

Example 13: Show that if events E and F are independent of each other, then so are E’ and F’.

Example 14: A card is drawn from a 52-card deck. We continue to draw until we have drawn a king or until we have drawn five cards, whichever comes first. Draw a tree diagram that illustrates the experiment. Put the appropriate probabilities on the tree. Find the probability that the drawing ends before the fourth draw.

Example 15: There are three sections of a mathematics course available at convenient times for a student. There is a 20% chance that Professor Jones gives a final exam, a 10% chance that Professor Cates gives a final exam, and a 5% chance that Professor Smithson give a final. At other times there are two biology sections, and in those the probabilities of a final are 20% and 13%. Find the probability that a student who randomly chooses one mathematics course and one biology course has to take at least one final examination.

Example 16: Three ordinary quarters and a fake quarter with two heads are placed in a hat. One quarter is selected at random and tossed twice. If the outcome is “HH,” what is the probability that the fake quarter was selected?

Example 17: Suppose that the reliability of a test for hepatitis is specified as follows. Of people with hepatitis, 95% have a positive reaction and 5% have a negative reaction; of people free of hepatitis, 90% have a negative reaction and 10% have a positive reaction. From a large population of which .05% of the people have hepatitis, a person is selected at random and given the test. If the test is positive, what is the probability that the person actually has hepatitis?

Example 18: Thirteen cards are dealt from a deck of 52 cards. Suppose the experiment in part (b) is repeated a total of 10 times (replacing the card looked at each time), and the ace of spades is not seen. What is the probability that the ace of spades actually is one of the 13 cards? The experiment in part (b) is: Suppose one of the 13 cards is chosen at random and found not to be the ace of spades. What is the probability that none of the 13 cards is the ace of spades?

Example 19: An over-the-counter pregnancy test claims to be 99% accurate. Actually, what the insert says is that if the test is performed properly, it is 99% sure to detect a pregnancy. Let us assume that the probability is 98% that the test result is negative for a woman who is not pregnant. If the woman estimates that her chances of being pregnant are about 40% and the test result is positive, what is the probability that she is actually pregnant?

Example 20: It is estimated that 10% of Olympic athletes use steroids. The test currently being used to detect steroids is said to be 93% effective in correctly detecting steroids in users. It yields false positives in only 2% of the tests. A country's best weightlifter tests positive. What is the probability that he actually takes steroids?

Example 21: An archer can hit the bull's-eye of the target with probability $\frac{1}{3}$. She shoots until she hits the bull's-eye or until four shots have been taken. The number of shots is observed. Determine the probability distribution for this experiment.

Here is a tree that I constructed to help me set up the probability distribution table for this problem.

The tree helped me to establish the following probability distribution table.

X	P(X)
1	$P(H_1)=\frac{1}{3}$
2	$P(H_1'\cap H_2)=\frac{2}{3}\cdot\frac{1}{3}=\frac{2}{9}$
3	$P(H_1'\cap H_2'\cap H_3)=\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}=\frac{4}{27}$
4	$P[(H_1'\cap H_2'\cap H_3'\cap H_4)\cup(H_1'\cap H_2'\cap H_3'\cap H_4')]=$ $\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{1}{3}+\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{2}{3}=\frac{24}{81}=\frac{8}{27}$

Sum $\frac{1}{3}+\frac{2}{9}+\frac{4}{27}+\frac{24}{81}=\frac{27+18+12+24}{81}=1$

Example 22: In a certain carnival game the player selects two balls at random from an urn containing two red balls and four white balls. The player receives \$5 if he draws two red balls and \$1 if he draws one red ball. He loses \$1 if no red balls are in the sample. Determine the probability distribution for the experiment of playing the game and observing the player's earnings.

Here is my distribution in table form.

X (net earnings in dollars)	Number of red balls (underlying random variable)	P(X)
5	2	$\frac{C(2,2)\cdot C(4,0)}{C(6,2)}$
1	1	$\frac{C(2,1)\cdot C(4,1)}{C(6,2)}$
-1	0	$\frac{C(2,0)\cdot C(4,2)}{C(6,2)}$

Example 23: Here is the probability distribution of the random variable U.

K	P(*U*=k)
0	3/15
1	2/15
2	4/15
3	5/15
4	?

(a) Determine the probability that *U*=4.

$$P(U=4) = 1-\left(\frac{3}{15}+\frac{2}{15}+\frac{4}{15}+\frac{5}{15}\right)=1-\frac{14}{15}=\frac{1}{15}$$

(b) Find $P(U\geq 2)$

$$P(U\geq 2)=\frac{4}{15}+\frac{5}{15}+\frac{1}{15}=\frac{10}{15}=\frac{2}{3}$$

(c) Find the probability that *U* is at most 3.

$$P(U \le 3) = \frac{3}{15} + \frac{2}{15} + \frac{4}{15} + \frac{5}{15} = \frac{14}{15}$$

(d) Find the probability that *U* + 2 is less than 4.

Here the author is trying to make the connection that random variables can be treated just like the regular old variables that we've come to know and love. We can solve for these just like the other variables in order to create an equivalency in terms of our original random variable. That is here we are asked to find P(U+2 < 4), so we would like to interpret this in terms of our random variable U. Just like any other inequality let's subtract 2 from both sides to express the probability in terms of U. That is P(U+2 < 4) = P(U < 2) = 3/15 + 2/15 = 5/15 = 1/3. Another way to handle this is to rewrite your probability distribution table with the new random variable, and then use the new table to compute probabilities. The table would look like this:

U	*U*+2	P(*U*); P(*U*+2)
0	2	3/15
1	3	2/15
2	4	4/15
3	5	5/15
4	6	1/15

Example 24: A die is tossed 12 times. What is the probability of at least two 5's?

I found going backdoor easier here, so find the probability of what we don't want (the complement), and then subtract from one. Since the experiment of rolling a die 12 times is a binomial type of experiment we can use the formula for setting them up.

1 – P(0 fives or 1 five) = 1 – [P(0 fives) + P(1 five)] =

$$1 - \left[C(12,0)\left(\frac{1}{6}\right)^{0}\left(\frac{5}{6}\right)^{11} + C(12,1)\left(\frac{1}{6}\right)^{1}\left(\frac{5}{6}\right)^{11} \right] \approx 0.6187$$

Example 25: A single die is rolled 10 times and the number of sixes is observed. What is the probability that a six appears 9 times, given that it appears at least 9 times?

P(exactly 9 sixes | at least 9 sixes) =

P(exactly 9 sixes and at least 9 sixes) / P(at least 9 sixes) =

P(exactly 9 sixes) / P(at least 9 sixes) =

$$\frac{C(10,9)\left(\frac{1}{6}\right)^{9}\left(\frac{5}{6}\right)^{1}}{C(10,9)\left(\frac{1}{6}\right)^{9}\left(\frac{5}{6}\right)^{1}+C(10,10)\left(\frac{1}{6}\right)^{10}\left(\frac{5}{6}\right)^{0}} \approx 0.9804$$

Notice that the intersection of the events "exactly 9 sixes" and "at least 9 sixes" is simply the event "exactly 9 sixes." This is because it is a subset of the "at least 9 sixes" event, so the intersection produces the subset event. Those are the outcomes that are shared by the two events.

Example 26: In a carnival game, the player selects balls one at a time, without replacement, from an urn containing two red and four white balls. The game proceeds until a red ball is drawn. The player pays $1 to play the game and receives $0.50 for each ball drawn. Write down the probability distribution for the player's earnings and find its expected value.

To make a probability distribution usually means setting up a table with a column for all of the possible outcomes of the experiment (i.e. the sample space or the random variable values; I usually abbreviate this column heading as "X," and another column for the probabilities of these outcomes, which I head as "P(X)." Often the probabilities are determined not directly by the given experiment, but rather by some underlying probability distribution that the given outcomes depend on. That is the case here. The outcomes are the net earnings in dollars, but these are dependent on the number of balls drawn. To help compute the probabilities I used this tree diagram:

Here is the Probability Distribution Table for this game (experiment):
Note: I used the convention of using subscripted Letters to denote the color of the ball on a given draw. The letter denotes the color and the subscript the number of the draw. Notice how the "and"s are translated into multiplications, and that the probabilities in the products after the first draw are conditional (dependent upon the previous draws).

(X): Net earnings in dollars	Underlying random variable: # of balls drawn	P(X): probability of each outcome
-1/2	1 (because the only way to lose a half-dollar is to draw only one ball, which can only happen if a red ball is drawn on the first draw).	P(losing 50 cents) = P(R_1) = $\frac{2}{6}=\frac{1}{3}$
0	2 (because the only	

	way to break even is to draw 2 balls, and this can only happen if a white is drawn first, and then a red on the second draw).	P(breaking even) = P(W_1 and R_2) = $\frac{4}{6}\cdot\frac{2}{5}=\frac{4}{15}$
+1/2	3 (because the only way to win a half-dollar is to draw three balls, and this can only happen if a white is drawn on the first two draws and a red on the third).	P(winning 50 cents) = P(W_1 and W_2 and R_3) = $\frac{4}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}=\frac{1}{5}$
+1	4 (etc., see the connection with the tree?)	P(winning \$1) = P(W_1 and W_2 and W_3 and R_4) $=\frac{4}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{2}{3}=\frac{2}{15}$
+3/2	5	P(winning \$1.50) = P(W_1 and W_2 and W_3 and W_4 and R_5) = $\frac{4}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{1}{3}\cdot\frac{2}{2}=\frac{1}{15}$

Notice that the probabilities in the table are simply the ones that we get from the leaves of our tree. For example, P(W_1 and R_2) is found by multiplying the probabilities along the route traced from the beginning of the tree through to the R on the second draw. Symbolically these probabilities are $P(W_1) \times P(R_2|W_1)$, but by our previous definition of the probability of the intersection of two events this is the same as P(W_1 and R_2).

Now that we have the table setup we can easily compute the expected value for the random variable associated with this experiment (game). To compute E(X) we simply add together the products of each outcome with its probability. For us this yields, E(X) = (-1/2)(1/3) + 0(4/15) + (1/2)(1/5) + 1(2/15) + (3/2)(1/15) = \$0.167 or about 17 cents per game. That is in the long run you will average a gain of 17 cents per game; thus this game is not fair (the expected value is not zero), but it favors the player. Definitely a game to play!

Example 27: Using life insurance tables, a retired man determines that the probability of living 5 more years is 0.9. He decides to take out a life insurance policy that will pay \$10,000 in the event that he dies during the next 5 years. How much should he be willing to pay for this policy? (Do not take account of interest rates or inflation.)

This is a nice backwards problem. Instead of being given the probability distribution and using it to find the expected value, we are given the expected value (well sort of) and asked to define a probability distribution. We do not know the price to pay for this insurance, so the power of algebra is that we can let a letter (variable) represent this idea, so here let's use p to represent this unknown price. To make a table to help us generate the expected value we need outcomes and probabilities. Here is what I set up.

X	P(X)
-p	0.9
10,000 – p	0.1

The –p represents the cost to the individual if he lives, and 10,000 – p represents the benefit to the family if he dies (notice that the cost of the insurance is deducted). To compute the expected value we multiply each outcome by its probability, and then add these products together. This yields the equation: $E(X) = -0.9p + 0.1(10{,}000 - p)$. We have too many unknowns yet in order to solve the equation, so we need to replace the E(X) on the left-hand side of the equation with a number, then we can solve for p. The implicit assumption here is that the man is will to pay whatever makes the expected value zero. To pay more means a loss for his family. To pay less money benefits his family, so the best I can figure is that this is a maximum price. I would think that a cheaper price would certainly be okay with the man. I mean I wouldn't complain, but to pay more would make no sense. By letting E(X) = 0, we get the equation: $0 = -0.9p + 0.1(10{,}000 - p)$, which when solved for p, yields a price of $1000.

Example 28: A pair of dice is tossed and the larger of the two numbers showing is recorded. Find the expected value of this experiment.

Here is a table that helped me for this problem.

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	2	3	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

I put the larger of the row and column numbers in the intersecting cell. I see that there are 36 possible outcomes. Let's use this to set up the probability distribution table, which we can then use to compute our expected value. Below is the distribution table.

X	P(X)
1	1/36

2	3/36
3	5/36
4	7/36
5	9/36
6	11/36
Sum	36/36 =1

Example 29: A truck can carry a maximum of 75,000 pounds cargo. How many cases of cargo can it carry if half of the cases have an average (arithmetic mean) weight of 20 pounds and the other half have an average weight of 30 pounds?

Let x represent the number of cases. I conceptualize constructing an equation as saying the same thing twice, but in two different ways. We have 75,000 pounds that we can carry, so how can we say 75,000, but in a different way? Half of our load is made up of 20 pound cases, and the other half averages 30 pounds, so 20[(1/2)x] represents the weight in pounds of the 20 pound cases and 30[(1/2)x] represents the weight of the 30 pound cases. These two weights must add up to be our total load, which is 75,000 pounds. So there is our other way, and we get the equation: 20[(1/2)x] + 30[(1/2)x] = 75,000. Solve this for x, and you find that 3,000 such cases are the most that we can carry.