Appendix C

Michelle Bodnar, Andrew Lohr

December 7, 2017

Exercise C.1-1

Since a substring is indexed by a pair of integers i, j such that $1 \le i \le j \le n$. Since we want the length of the substring to be k, then, $j - i = k - 1$. Rearranging this, we can solve for j and plug it into the original set of inequalities to get that we need $1 \le i$ and $i \le n - k - 1$. So, i can be precicely any of $\{1..n - k + 1\}$. So, there are $n - k + 1$ substrings of length k. To get the total number of substrings, we can just add this formula over all k. Here we are assuming that a string needs to be nonempty, a point that was left undeclared in the definition given in this section.

$$\sum_{k=1}^{n} n - k + 1 = n^2 - \frac{n(n+1)}{2} + n$$
$$= \frac{n(n+1)}{2}$$
$$= \binom{n+1}{2}$$

Such a lovely formula deserves a combinatorial explaination instead of just a bunch of algebraic symbol pushing. And it turns out that we don't need to look far. Given a string of length n, and pair of distinct indices in $\{1..n + 1\}$, we get a uniques subtring starting at the first index and ending the position before the second index. Put in other words, pairs of these indices are in bijection with substrings. Since there are n indices to begin with, there are $\binom{n+1}{2}$ pairs of indices picked from $\{1..n + 1\}$, and so $\binom{n+1}{2}$ substrings.

Exercise C.1-2

There are $2^{(2^n)}$ n-input, 1-output boolean functions and $(2^m)^{(2^n)}$ n-input, m-output boolean functions.

Exercise C.1-3

First seat one of the professors, then there are $n - 1$ choices for the person sitting immediately to his right. Then, $n - 2$ choices for the person immediately to that person's right. Continue this until there is only a single person left. Multiplying together all of these choices, we get that there are $(n - 1)(n - 2) \cdots 1 = (n - 1)!$ ways.

Exercise C.1-4

The sum of three numbers is even if and only if they are all even, or exactly two are odd and one is even. For the first case, there are $\binom{49}{3}$ ways to pick them. For the second case, there are $\binom{50}{2}\binom{49}{1}$ ways. Thus, the total number of ways to select 3 distinct numbers so that their sum is even is

$$\binom{49}{3} + \binom{50}{2}\binom{49}{1}.$$

Exercise C.1-5

We'll just do an algebraic proof, even though "nicer" combinatorial proofs do exist

$$\begin{aligned}\binom{n}{k} &= \frac{n!}{k!(n-k)!} \\ &= \frac{n}{k}\frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} \\ &= \frac{n}{k}\binom{n-1}{k-1}\end{aligned}$$

Exercise C.1-6

We can prove this directly:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n}{n-k}\frac{(n-1)!}{k!(n-k-1)!} = \frac{n}{n-k}\binom{n-1}{k}.$$

Exercise C.1-7

Distinguish one of the objects. In the case that you do select this object, then you only need to pick $k - 1$ more objects from the remaining $n - 1$ items, which there are $\binom{n-1}{k-1}$ ways of doing. If you do not select the distinguished object, you still need to pick all k items, and there are only $n - 1$ items left to pick from, since you know you can't pick the distinguished object. This gets us $\binom{n-1}{k}$ ways. Since these two cases are disjoint and cover all possibilities, we have that

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$$

2

Exercise C.1-8

The following shows Pascal's triangle with n increasing down columns and k increasing across rows.

```
1
1   1
1   2   1
1   3   3   1
1   4   6   4   1
1   5   10  10  5   1
1   6   15  20  15  6   1
```

Exercise C.1-9

There is a simple algebraic proof that can be done by induction based on the fact that $\binom{n}{2} = \frac{n^2-n}{2}$. There is a more interesting combinatorial proof as follows. Suppose that you are picking two items from a set of $n+1$ items. Then, spit into n different cases based on what the last item you pick is. Suppose that you pick position i as your last one. Notice that we must have that i is from 2 to $n + 1$. Then, for the other thing we pick, we could pick anything occurring before i, so we have $i - 1$ choices. From there, we just to a simple change of variables to get

$$\binom{n}{k} = \sum_{i=2}^{n+1} i - 1 = \sum_{i=1}^{n} i$$

Exercise C.1-10

Fix n. Then we have

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n-k+1}{k}\frac{n!}{(k-1)!(n-k+1)!} = \frac{n-k+1}{k}\binom{n}{k-1}.$$

Thus, we increase in k if and only if $\frac{n-k+1}{k} \geq 1$, which happens only when $n + 1 \geq 2k$, or $k \leq \lceil n/2 \rceil$. On the other hand, we decrease in k if and only if $\frac{n-k+1}{k} \leq 1$, so $k \geq \lfloor n/2 \rfloor$. Thus, the function is maximized precisely when k is equal to one of these.

Exercise C.1-11

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

$$= \frac{n!}{j!(n-j)!} \frac{j!(n-j)!}{(j+k)!(n-j-k)!}$$

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

$$= \binom{n}{j} \frac{(n-j)!}{(j+k)(j+k-1)\cdots(j+1)(n-j-k)!}$$

$$\leq \binom{n}{j} \frac{(n-j)!}{k!(n-j-k)!}$$

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

For a combinatorial (not algebraic) proof, we can see that if we fisrt pick j items, and then from the remaining pick k more, we have the quantity on the right. However, for each of the selections of $j+k$ items on the right, there is at least one way of picking j then picking k, giving us our inequality.

To see it is not tight, consider when $n = 2$, and $j = k = 1$. Then, we have that the left is $\binom{2}{2} = 1$. The right, however, is $\binom{2}{1}\binom{1}{1} = 2$ which is strictly greater.

Exercise C.1-12

We'll prove inequality C.6 for $k \leq n/2$ by induction on k. For $k = 0$ we have $\binom{n}{0} = 1 \leq \frac{n^n}{0^0(n-0)^{(n-0)}} = 1$. Now suppose the claim holds for k, and that $k < n/2$. Then we have

$$\binom{n}{k+1} = \frac{n-k}{k+1}\binom{n}{k}$$

$$\leq \frac{n-k}{k+1}\frac{n^n}{k^k(n-k)^{(n-k)}}$$

$$= \frac{n^n}{(k+1)k^k(n-k)^{(n-k-1)}}.$$

To show that this is bounded from above by $\frac{n^n}{(k+1)^{k+1}(n-k-1)^{(n-k-1)}}$ we need only verify that $\left(\frac{k+1}{k}\right)^k \leq \left(\frac{n-k}{n-k-1}\right)^{n-k-1}$. This follows from the fact that the left hand side, when viewed as a function of k, is increasing, and $k < n/2$ which implies that $k + 1 \leq n - k$. By induction, the claim holds. Using equation $C.3$, we see that the claim extends to all $0 \leq k \leq n$ since the right hand side of the inequality is symmetric in k and $n - k$.

Exercise C.1-13

We start with $\binom{2n}{n} = \frac{(2n)!}{n!n!}$, then, applying Stirling's approximation, we get that this is

$$\frac{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}\left(1+\Theta\left(\frac{1}{n}\right)\right)}{\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\left(1+\Theta\left(\frac{1}{n}\right)\right)\right)\cdot\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\left(1+\Theta\left(\frac{1}{n}\right)\right)\right)} = \frac{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}\left(1+\Theta\left(\frac{1}{n}\right)\right)}{\left(2\pi n\left(\frac{n}{e}\right)^{2n}\right)\cdot\left(\left(1+\Theta\left(\frac{1}{n}\right)\right)\right)}$$

$$= \frac{\sqrt{4\pi n}}{2\pi n}2^{2n}\left(1+\Theta\left(\frac{1}{n}\right)\right)$$

$$= \frac{2^{2n}}{\sqrt{\pi n}}\left(1+\Theta\left(\frac{1}{n}\right)\right)$$

Exercise C.1-14

Differentiating the entropy function and setting it equal to 0 we have

$$H'(\lambda) = \lg(1-\lambda) - \lg(\lambda) = 0,$$

or equivalently $\lg(1-\lambda) = \lg(\lambda)$. This happens when $\lambda = 1/2$. Moreover, $H''(1/2) = \frac{-4}{\ln(2)} < 0$, so this is a local maximum. We have $H(1/2) = 1$, and since $H(0) = H(1) = 0$, this is in fact a global maximum for H.

Exercise C.1-15

$$\sum_{k=0}^{n}\binom{n}{k}k = \sum_{k=0}^{n}\frac{n!}{k!(n-k)!}k$$

$$= \sum_{k=0}^{n}\frac{(n-1)!}{(k-1)!((n-1)-(k-1))!}n$$

$$= n\sum_{k=0}^{n}\frac{(n-1)!}{(k-1)!((n-1)-(k-1))!}$$

$$= n\sum_{k=0}^{n}\binom{n-1}{k-1}$$

$$= n\sum_{k=0}^{n-1}\binom{n-1}{k}$$

$$= n2^{n-1}$$

Exercise C.2-1

The only way that Rosencrantz could of obtained more heads than Guildenstern is if he got a hears and Guildenstern got no heads, so, two tails. This

5

event has proscribed outcomes for all three coin flips, and so it happens with probability $\frac{1}{2^3} = \frac{1}{8}$.

Exercise C.2-2

Let $B_i = A_i \setminus (\cup_{k=1}^{i-1} A_i)$. Then $B_1, B_2, \ldots$ are disjoint and $A_1 \cup A_2 \cup \ldots = B_1 \cup B_2 \cup \ldots$. Moreover, $B_i \subseteq A_i$ for each i, so $Pr(B_i) \leq Pr(A_i)$. By third axiom of probability we have

$$Pr(A_1 \cup A_2 \cup \ldots = Pr(B_1 \cup B_2 \cup \ldots) = \sum_{i \geq 1} Pr(B_i) \leq \sum_{i \geq 1} Pr(A_i).$$

Exercise C.2-3

It is a 1/6 likelyhood. Here are two ways that you can show this fact. The first kind of messy, the second much more elegant.

This first is a direct counting argument. There are $10 \cdot 9 \cdot 8 = 720$ differerent sequences of cards picked, since we can at each step pick any card that is left. To count the number of sequences that are in increasing order, we choose the first card(call its value i), then any number larger than that for the second(say j) and any number larger than the second for the third(say k). This can be expressed by the summation:

$$\sum_{i=1}^{10} \sum_{j=i+1}^{10} \sum_{k=j+1}^{10} 1 =$$

$$\sum_{i=1}^{10} \sum_{j=i+1}^{10} 10 - j =$$

$$\sum_{i=1}^{10} 10(10 - i) - \frac{10 \cdot 11}{2} + \frac{(i)(i+1)}{2} =$$

$$\sum_{i=1}^{10} 45 - 9.5i + i^2/2 =$$

$$450 - \frac{9.5 \cdot 10 \cdot 11}{2} + \frac{10 \cdot 11 \cdot 21}{2 \cdot 6} =$$

$$450 - 522.5 + 192.5 =$$

$$120$$

Where we have arrived at this count only using the formula for sum of kth powers. Then this gives us our final probability of $120/720 = 1/6$. Note also that it is unimportant to rule out the cases i=9, 10 and j=10. They are impossible cases, but if they occur then the sum will come out to be equal to zero anyways, I just figured that having it all go to 10 was more aesthetically pleasing.

The second way to show this probability is to first fix the three elements that you happened to pick and use symmetry. The set of values that are among

the cards has no bearing on whether the cards are in order. Since the numerical value on these cards is independent of the random the selection process, all possible orders for those three elements are equally likely. There are six different orderings on three elements, so the likelyhood they are ordered numerically is 1/6.

Thank you to Eric Richardson for pointing out that the originally posted solution for this exercise was misinterpreting the question.

Exercise C.2-4

We can verify this directly from the definition of conditional probability as follows:

$$Pr(A|B) + Pr(\overline{A}|B) = \frac{Pr(A \cap B)}{Pr(B)} + \frac{Pr(\overline{A} \cap B)}{Pr(B)} = \frac{Pr(B)}{Pr(B)} = 1.$$

Exercise C.2-5

We will proceed by induction on n. If $n = 2$, this is equation (C.16). Suppose that $n > 2$. Let $C = \cap i = 1^{n-1} A_i$. Then, we will use equation (C.16) followed by the inductive hypothesis to get

$$
\begin{aligned}
\Pr(A_1 \cap A_2 \cap \cdots \cap A_n) &= \Pr(C \cap A_n) \\
&= \Pr(A_n|C) \Pr(C) \\
&= \Pr(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \Pr(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \\
&= \Pr(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \cdots \Pr(A_3|A_1 \cap A_2) \Pr(A_2|A_1) \Pr(A_1)
\end{aligned}
$$

Exercise C.2-6

Let $.a_1 a_2 a_3 \ldots$ be the binary representation of a/b. Flip the fair coin repeatedly, associating 1 with heads and 0 with tails, until the first time that the value of the ith flip differs from a_i. If the value is greater than a_i, output tails. If the value is less than a_i, output heads. Every number between 0 and $.99999\ldots = 1$ has the possibility to be represented, and is created by randomly choosing its binary representation. The probability that we output tails is the probability that our number is less than a/b, which is just a/b. The expected number of flips is $\sum_{i=1}^{\infty} i2^{-i} = 2 = O(1)$.

Exercise C.2-7

Select n random colinear points. Seeing any one of them, you know nothing new about the distribution of the others. However, given two of them, you know the line that the third point would need to lie on.

Exercise C.2-8

Suppose we have a biased coin which always comes up heads, and a fair coin. We pick one at random and flip it twice. Let A be the event that the first coin is heads, B be the event that the second coin is heads, and C be the event that the fair coin is the one chosen. Then we have $Pr(A \cap B) = 5/8$ and $Pr(A) = Pr(B) = 3/4$ so $Pr(A \cap B) \neq Pr(A)Pr(B)$. Thus, A and B are not independent. However,

$$Pr(A \cap B|C) = \frac{Pr(A \cap B \cap C)}{Pr(C)} = \frac{(1/2)(1/2)(1/2)}{1/2} = 1/4$$

and

$$Pr(A|C) \cdot Pr(B|C) = \frac{Pr(A \cap C)}{Pr(C)} \frac{Pr(B \cap C)}{Pr(C)} = \frac{1/4}{1/2} \frac{1/4}{1/2} = 1/4.$$

Exercise C.2-9

This problem has many explanations. I'll give a concise one, as there are versions online that really hold your hand through the whole thing.

First, supposing that you don't switch, then the likelihood of winning will just be $\frac{1}{3}$, since the original decision was made before any additional information leaked in by way of lifting a curtain, so it must just be uniformly random.

Now, suppose that you do switch, your likelihood of winning will increase to $\frac{2}{3}$. To see this, we split into two cases. Suppose you guessed correctly initially, which happens with probability $\frac{1}{3}$, then you will fail. However, if you picked incorrectly originally, which happens with probability $\frac{2}{3}$, you will then win by switching. Therefore, you will win with probability $\frac{2}{3}$.

So, your chances of winning double by adopting the switching strategy instead of the staying strategy.

Exercise C.2-10

His chances are still $1/3$, because at least one of Y or Z would be executed, so hearing which one changes nothing about his own situation. However, the probability that Z is going free is now $2/3$. To see this, note that the probability that the free prisoner is among Y and Z is $2/3$. Since we are told that it is not Y, the $2/3$ probability must apply exclusively to Z.

Exercise C.3-1

The expectation of a single dice roll is $\frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3.5$. So, by linearity of expectation, the expectation of their sum is 7.

We have that the probability that the maximum of the two dice results is $\leq i$ is the probability that both dice results lie in the range $1, 2, .., i$. This happens with probability $\frac{i^2}{36}$. Then, to get the probability that a particular

value of the maximum occurs, we take the difference of successive values of this cumulative density function. This gets us that the probability of a max of i is $\frac{i^2 - (i-1)^2}{36} = \frac{2i-1}{36}$. So, the expected value of this max is

$$\frac{1}{36}(1 + 6 + 15 + 28 + 45 + 66) = \frac{161}{36} \approx 4.47$$

Exercise C.3-2

The probability that the maximum or minimum element is in a particular spot is $1/n$ since the ordering is random. Thus, the expected index of the maximum and minimum are the same, and given by:

$$E[X] = \sum_{i=1}^{n} i(1/n) = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Exercise C.3-3

The probability that you loose your dollar is $(5/6)^3 = \frac{125}{216}$. The probability you gain exactly one dollar is $\binom{3}{1}(1/6)(5/6)^2 = \frac{75}{216}$. The probability you gain exactly two dollars is $\binom{3}{2}(1/6)^2(5/6) = \frac{15}{216}$. The probability that you win 3 dollars is $(1/6)^3 \frac{1}{216}$. putting it all together, the payoff is

$$\frac{3}{216} + \frac{30}{216} + \frac{75}{216} - \frac{125}{216} = -\frac{17}{216}$$

So, you are expected to loose $\frac{17}{216}$ of a dollar.

Exercise C.3-4

Let X and Y be nonnegative random variables. Let $Z = X + Y$. Then Z is a random variable, and since X and Y are nonnegative, we have $Z \geq \max(X, Y)$ for any outcome. Thus, $E[\max(X, Y)] \leq E[Z] = E[X] + E[Y]$.

Exercise C.3-5

the value of $f(X)$ is only a function of X, and, so it may only leak information about the state of X. Similarly $g(Y)$ can only depend on and leak information about the state of Y. Since the states of X and Y don't affect each other, knowing the value of functions of the two random varaibles also will not affect each other.

Exercise C.3-6

We can verify directly from the definition of expectation:

$$E[X] = \sum_{i=0}^{\infty} i \cdot Pr(X = i)$$

$$\geq \sum_{i=t}^{\infty} i \cdot Pr(X = i)$$

$$\geq t \sum_{i=t}^{\infty} Pr(X = i)$$

$$= t \cdot Pr(X \geq t).$$

Dividing both sides by t gives the result.

Exercise C.3-7

Let $S' \subseteq S$ be the states of the sample space for whcih X' takes a value that is $\geq t$. For any of these states, we will also have that X takes a value that is at least t. This gets us that the set of states for which X is at least t is a superset of S'. Then, by monotonicity of probability, we have the desired inequality on the probabilities.

Exercise C.3-8

The expectation of the square of a random variable is larger. To see this, note that by (C.27) we have $E[X^2] - E^2[X] = E[(X - E[X])^2] \geq 0$ since the square of a random variable is a nonnegative random variable, so its expectation must be nonnegative.

Exercise C.3-9

Since X only takes the values 0 and 1, we will always have that $X^2 = X$. Then, the probability that X takes the value of 1 is equal to the expected value of X. So,

$$Var[X] = E[X^2] - (E[X])^2$$
$$= E[X] - (E[X])^2$$
$$= E[X](1 - E[X])$$
$$= E[X](E[1] - E[X])$$
$$= E[X](E[1 - X])$$

Exercise C.3-10

Proceeding from (C.27) we have

$$\begin{aligned} Var[aX] &= E[(aX)^2] - E^2[aX] \\ &= E[a^2 X^2] - a^2 E^2[X] \\ &= a^2 E[X^2] - a^2 E^2[X] \\ &= a^2 (E[X^2] - E^2[X]) \\ &= a^2 Var[X]. \end{aligned}$$

Exercise C.4-1

Axiom 2 states that the probability of the entire space should be one. This is saying that if we add up the probabilities of each i occurring over all i, we should get one. This is easy to see because

$$\sum_{i=1}^{\infty} (1-p)^{i-1} p = \frac{p}{1-(1-p)} = \frac{p}{p} = 1$$

Exercise C.4-2

Let X be the number of times we must flip 6 coins before we obtain 3 heads and 3 tails. The probability that we obtain 3 heads and 3 tails is $\binom{6}{3}(1/2)^6 = 5/16$, so the probability that we don't is $11/16$. Moreover, X has a geometric distribution, so by (C.32) we have

$$E[X] = 1/p = 16/5.$$

Exercise C.4-3

To go from the first line to the second, we use the fact that picking k items is identical to selecting the $n-k$ items you aren't picking.

$$\begin{aligned} b(k; n, p) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &= \binom{n}{n-k} p^k (1-p)^{n-k} \\ &= b(n-k; n, 1-p) \end{aligned}$$

Exercise C.4-4

Using Stirling's approximation we have $b(k; n, p) \approx \frac{\sqrt{n} n^n}{\sqrt{2\pi}\sqrt{k(n-k)} k^k (n-k)^{n-k}} p^k q^{n-k}$. The binomial distribution is maximized at its expectation. Plugging in $k = np$ gives

$$b(np; n, p) \approx \frac{\sqrt{n} n^n p^{np} (1-p)^{n-np}}{\sqrt{np(n-np)}(np)^{np}(n-np)^{n-np}} = \frac{1}{\sqrt{2\pi npq}}.$$

Exercise C.4-5

The probability that there will be no successes is equal to

$$\lim_{n\to\infty} b(0; n, 1/n) = \lim_{n\to\infty} \binom{n}{0} \frac{1}{n}^0 (1 - \frac{1}{n})^{n-0}$$
$$= \lim_{n\to\infty} (1 - \frac{1}{n})^n$$
$$= \lim_{n\to\infty} ((1 + \frac{1}{-n})^{-n})^{-1}$$
$$= e^{-1} = \frac{1}{e}$$

Similarly, the probability of one success is equal to

$$\lim_{n\to\infty} b(1; n, 1/n) = \lim_{n\to\infty} \binom{n}{1} \frac{1}{n}^1 (1 - \frac{1}{n})^{n-1}$$
$$= \lim_{n\to\infty} \frac{n}{n(1 - \frac{1}{n})} (1 - \frac{1}{n})^n$$
$$= \left(\lim_{n\to\infty} \frac{n}{n(1 - \frac{1}{n})} \right) \cdot \left(\lim_{n\to\infty} (1 - \frac{1}{n})^n \right)$$
$$= 1 \cdot \frac{1}{e} = \frac{1}{e}$$

Exercise C.4-6

There are $2n$ total coin flips amongst the two professors. They get the same number of heads if Professor Guildenstern flips k heads and Professor Rosencrantz flips $n - k$ tails. If we imagine flipping a head as a success for Professor Rosencrantz and flipping a tail as a success for Professor Guildenstern, then the professors get the same number of heads if and only if the total number of successes achieved by the professors is n. There are $\binom{2n}{n}$ ways to select which coins will be a successes for their flipper. Since the outcomes are equally likely, and there are $2^{2n} = 4^n$ possible flip sequences, the probability is $\binom{2n}{n}/4^n$.

To verify the identity, we can imagine counting successes in two ways. The right hand side counts via our earlier method. For the left hand side, we can imagine first choosing k successes for Professor Guildenstern, then choosing the remaining $n - k$ successes for Professor Rosencrantz. We sum over all k to get the total number of possibilities. Since the two sides of the equation count the same thing they must be equal.

Exercise C.4-7
We apply (C.7) with $\lambda = \frac{k}{n}$, then,

$$b(k; n, 1/2) = \binom{n}{k}(1/2)^k(1 - 1/2)^{n-k}$$
$$= \binom{n}{k}2^{-n}$$
$$\leq 2^{nH(k/n)2^{-n}}$$
$$= 2^{nH(k/n)-n}$$

Exercise C.4-8

Let $a_1, a_2, \ldots, a_n$ be uniformly chosen from $[0, 1]$. Let X_i be the indicator random variable that $a_i \leq p_i$ and Y_i be the indicator random variable that $a_i \leq p$. Let $X' = \sum_{i=1}^n X_i$ and $Y' = \sum_{i=1}^n Y_i$. Then for any event we have $X' \leq Y'$, which implies $P(X' < k) \geq P(Y' < k)$. Let Y be the number of successes in n trials if each trial is successful with probability p. Then X has the same distribution as X' and Y has the same distribution of Y', so we conclude that $P(X < k) \geq P(Y < k)$.

Exercise C.4-9

Consider having the underlying state space be n copies of the unit interval. Then, a particular state will be a sequence of n numbers in $[0, 1]$. Then, X' will be the number of times the ith coordinate is less than or equal to p_i. Similarly, X is the number of times that the ith coordinate is less than or equal to p_i'. Since $p_i \leq p_i'$, we have that for any particular state, $X \leq X'$. Then we can directly apply the results of C.3-7 to get the desired result.

Exercise C.5-1

Which one is less likely depends on the value of n.

The probability that you obtain no head when flipping n times is just 2^{-n}, straight from the definition of the binomial distribution.

The probability P that we will have fewer than n heads from $4n$ flips is exactly

$$P = \Pr\{X < n\} = \sum_{k=0}^{k=n-1} b(k, 4n, 1/2) = 2^{-4n} \sum_{k=0}^{k=n-1} \binom{4n}{k}$$

In particular, since all the terms in the sum are positive, the probability is at least the value of the largest term, so

$$P \geq \binom{4n}{n-1}2^{-4n}$$

So, if we can just show that the quantity $\binom{4n}{n-1}2^{-3n} > 1$, then we have that P is greater than the probability that we have no head when flipping n times.

This is in fact not true in general, but we will show it is true for $n \geq 20$. We can then show by explicit computation what happens for $n = 1 \ldots 19$. As in inductive step, we can confirm that $\binom{80}{19}2^{-60} \approx 1.005 > 1$.

Now, suppose that we have $\binom{4n}{n-1}2^{-3n} > 1$.

$$
\begin{aligned}
\binom{4n+4}{n}2^{-3n-3} &= \frac{(4n+4)!}{8(n)!(3n+4)!}2^{-3n} \\
&= \frac{(4n+4)(4n+3)(4n+2)(4n+1)}{8n(3n+4)(3n+3)(3n+2)}\frac{(4n)!}{(n-1)!((3n+1)!}2^{-3n} \\
&= \frac{(4n+4)(4n+3)(4n+2)(4n+1)}{8n(3n+4)(3n+3)(3n+2)}\binom{4n}{n-1}2^{-3n} \\
&> \frac{(4n+4)(4n+3)(4n+2)(4n+1)}{8n(3n+4)(3n+3)(3n+2)} \\
&= \frac{32}{27}\frac{(n+1)(n+3/4)(n+1/2)(n+1/4)}{8n(n+4/3)(n+1)(n+2/3)}
\end{aligned}
$$

Using elementary calculus, we can see that the limit of this rational expression in n is equal to $\frac{32}{27}$, so, there is some n so that for all n past there, it will be greater than 1, so, we will have $\binom{4n+4}{n}2^{-3n-3} > 1$.

More concretely, we can show that the rational function is an increasing function on n for $n > 0$ by taking it's derivative:

$$
\frac{-24 - 108n - 93n^2 + 116n^3 + 144n^4}{3n^2(2+3n)^2(4+3n)^2}
$$

and noticing it is positive for $n \geq 1$ because the numerator has real roots at $\approx -.368$ and $\approx .937$ and end behavor of $+\infty$ as $n \to \infty$. Then, since $n = 3$ causes the original rational expression to be $\frac{35}{33}$, we have that it is greater than 1 for all $n \geq 3$. Since we are only considering $n \geq 20$ here, we definitly have that the expression is > 1.

Lastly, we turn our attention to the finitely many values of n that we left out before. Since we only need to consider $n = 1 \ldots 19$, we can just perform the explicit computation for P as $\sum_{k=0}^{k=n-1} b(k, 4n, 1/2)$ at each value of n and compare it to 2^{-n}.

n	P	2^{-n}
1	1/16	1/2
2	9/256	1/4
3	79/4096	1/8
4	697/65536	1/16
5	1549/262144	1/32
6	55455/16777216	1/64
7	249589/134217728	1/128
8	4514873/4294967296	1/256
9	10249879/17179869184	1/512
10	93396401/274877906944	1/1024
11	3414035527/17592186044416	1/2048
12	31278197839/281474976710656	1/4096
13	71797952431/1125899906842624	1/8192
14	1321035185597/36028797018963968	1/16384
15	6086999773681/288230376151711744	1/32768
16	224723513577529/18446744073709551616	1/65536
17	519244699305955/73786976294838206464	1/131072
18	4805026093205835/118059162071741303424	1/262144
19	4451531435563063/18889465931478580854784	1/524288

We can stop the table here since our inductive proof that P was larger was starting at a base case of $n = 20$. So, by looking at the numerical values in the table, we can see that if $n < 18$, then it is more unlikely that you have fewer than n heads out of $4n$ flips and if $n \geq 18$, then it more unlikely to have no heads after flipping n times.

Exercise C.5-2

For Corollary C.6 we have $b(i; n, p)/b(i - 1; n, p) \leq \frac{(n-k)p}{kq}$. Let $x = \frac{(n-k)p}{kq}$. Note that $x < 1$, so the infinite series below converges. Then we have

$$Pr(X > k) = \sum_{i=k+1}^{n} b(i; n, p)$$

$$\leq \sum_{i=k+1}^{n} x^{i-k} b(k; n, p)$$

$$= b(k; n, p) \sum_{i=1}^{n-k} x^i$$

$$\leq b(k; n, p) \sum_{i=1}^{\infty} x^i$$

$$= b(k; n, p) \frac{x}{1-x}$$

$$= b(k; n, p) \frac{(n-k)p}{k-np}.$$

For Corollary C.7, we use Corollary C.6 and the fact that $x < 1$ as follows:

$$\frac{Pr(X > k)}{Pr(X > k-1)} = \frac{Pr(X > k)}{Pr(X > k) + Pr(X = k-1)} \leq \frac{x b(k; n, p)}{x b(k; n, p) + b(k; n, p)} < \frac{1}{2}.$$

Exercise C.5-3

Divide both sides of the expression through by $(a+1)^n$. Since $a > 0$, this quantity is > 1, so the inequalities remain the way they are. This means that it is equivalent to show that

$$\sum_{i=0}^{k-1} \binom{n}{i} a^i (a+1)^{-n} < \frac{k b(k; n, a/(a+1))}{na - k(a+1)}$$

So, we do some algebra, and apply Theorem C.4.

$$\sum_{i=0}^{k-1}\binom{n}{i}a^i(a+1)^{-n} = \sum_{i=0}^{k-1}\binom{n}{i}\left(\frac{a}{a+1}\right)^i\left(\frac{1}{a+1}\right)^{n-i}$$

$$= \sum_{i=0}^{k-1}\binom{n}{i}\left(\frac{a}{a+1}\right)^i\left(1-\frac{a}{a+1}\right)^{n-i}$$

$$= \sum_{i=0}^{k-1}b(i;n,\frac{a}{a+1})$$

$$< \frac{k\left(1-\frac{a}{a+1}\right)}{n\frac{a}{a+1}-k}b(k;n,\frac{a}{a+1})$$

$$= \frac{k\left(\frac{1}{a+1}\right)}{n\frac{a}{a+1}-k}b(k;n,\frac{a}{a+1})$$

$$= \frac{k}{na-k(a+1)}b(k;n,\frac{a}{a+1})$$

completing the proof.

Exercise C.5-4

Using Lemma C.1 and Corollary C.4 we have

$$\sum_{i=0}^{k-1}p^iq^{n-i} \le \sum_{i=0}^{k-1}\binom{n}{i}p^iq^{n-i}$$

$$= \sum_{i=0}^{k-1}b(i;n,p)$$

$$\le \frac{kq}{np-k}b(k;n,p)$$

$$\le \frac{kq}{np-k}\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}.$$

Exercise C.5-5

(note that this problem had extensive bugs prior to the third printing of the third edition, see errata.)

Since the values of p_i and q_i appear nowhere in this expression, we can swap the two. This has the effect of changing the mean to $n - \mu$. This means that the condition $r > \mu$ becomes $r > n - \mu$ which we have. Then, the right hand side becomes $\Pr\{(n-X)(b-\mu) \ge r\} = \Pr\{\mu - X \ge r\}$. Then, just replacing

the μ occurring in the left, we obtain that

$$\Pr\{\mu - X\} \leq \left(\frac{(n-\mu)e}{r}\right)^r$$

Similarly, we can apply corollary C.9 to this distribution counting the number of failures (swapping p and q). This has the result of making

$$\Pr\{np - X\} = \Pr\{n - X - n(1-p)\} = \Pr\{X - nq \geq r\} \leq \left(\frac{nqe}{r}\right)^r$$

again, the restriction on r changes exactly into what we are given.

Exercise C.5-6

As in the proof of Theorem C.8, we'll bound $E[e^{\alpha(X-\mu)}$ and substitute a suitable value for α. First we'll prove (with a fair bit of work) that if $q = 1 - p$ then $f(\alpha) = e^{\alpha^2/2} - pe^{\alpha q} - qe^{-\alpha p} \geq 0$ for $\alpha \geq 0$. First observe that $f(0) = 0$. Next we'll show $f'(\alpha) > 0$ for $\alpha > 0$. To do this, we'll show that $f'(0) = 0$ and $f''(\alpha) > 0$. We have

$$f'(\alpha) = \alpha e^{\alpha^2/2} - pqe^{\alpha q} + pqe^{-\alpha p}$$

so $f'(0) = 0$. Moreover

$$f''(\alpha) = \alpha^2 e^{\alpha^2/2} + e^{\alpha^2/2} - pq(qe^{\alpha q} + pe^{-\alpha p}).$$

Since $\alpha^2 e^{\alpha^2/2} > 0$ it will suffice to show that $e^{\alpha^2/2} \geq pq(qe^{\alpha q} + pe^{-\alpha p})$. Indeed, we have

$$pq(qe^{\alpha q} + pe^{-\alpha p}) \leq (1/4)(qe^{\alpha q} + pe^{-\alpha p} \leq (1/4)e^{-\alpha p}(e^\alpha + 1) \leq (1/4)(e^\alpha + 1)$$

so we need to show $4e^{\alpha^2/2} \geq e^\alpha + 1$. Since $e^{\alpha^2/2} > 1$, it is enough to show $3e^{\alpha^2/2} \geq e^\alpha$. Taking logs on both sides, we need $\alpha^2/2 - \alpha + \ln(3) \geq 0$. By taking a derivative we see that this function is minimized when $\alpha = 1$, where it attains the value $\ln(3) - 1/2 > 0$. Thus, the original inequality holds. Now can proceed with the rest of the proof. As in the proof of Theorem C.8, $E[e^{\alpha(X-\mu)} = \prod_{i=1}^n E[e^{\alpha(X_i - p_i)}]$. Using the inequality we just proved we have

$$E[e^{\alpha(X_i - p_i)}] = p_i e^{\alpha q_i} + q_i e^{-\alpha p_i} \leq e^{\alpha^2/2}.$$

Thus, $E[e^{\alpha(X-\mu)}] \leq \prod_{i=1}^n e^{\alpha^2/2} = e^{n\alpha^2/2}$. By this and (C.43) and (C.44) we have

$$Pr(X - \mu \geq r) \leq E[e^{\alpha(X-\mu)}]e^{-\alpha r} \leq e^{n\alpha^2/2 - \alpha r}.$$

Finally, taking $\alpha = r/n$ gives the desired result.

18

Exercise C.5-7

We put on our freshman calculus hats. With them securely affixed, we take a derivative of the expression with respect to α. This gets us $(\mu e^\alpha - r)\exp(\mu e^\alpha - \alpha r)$. Since a minima of the original expression will occur when the derivative is zero, and the end behavior of the expression is increasing without bound, we set this derivative equal to zero. The second factor is an exponential and so can never be zero. So, we are trying to solve $\mu e^\alpha - r = 0$. This means that $\alpha = \ln(r/\mu)$. at the minimum.

Problem C-1

a. For each of the n balls, we can make one of b different decisions about where to place it. Since this number of possible decisions is independent of the previous choices, the total number of possibilities is just b^n.

b. First, we pretend that we can distingush the sticks. This means that we want to arrange a total of n balls and $b - 1$ sticks, that is, $n + b - 1$ total things. There are exactly $(b + n - 1)!$ ways of doing this. Then, we realize that a bunch of these are the same, in particular, no matter how we permute the $b-1$ sticks in some arrangement, we end up with the same answer. Therefore, we can arrange n distinguishable balls and $b - 1$ indistinguishable sticks in $\frac{(n+b-1)!}{(b-1)!}$ many ways.

Once we have such an arrangement, we can relate it back to the original statement, where we imagine the sticks as being the dividing lines between bins, and the ordered balls between them being the ordered balls in each bin.

c. Taking the result from before, we notice that any of the n permutations of the balls will result in the same configuration, so, we must divide our count from the previous part by $n!$. This means that we are left with $\frac{(n+b-1)!}{n!(b-1)!} = \binom{n+b-1}{n}$.

d. Since each bin can either contain a bin or not, we are selecting a set of bins to contain balls in the part. The number of non-empty bins must be equal to the number of balls, so we are selecting n bins. That is, we are selecting a subset of size n of the bins from the set of all bins, which there are n of. This is then exactly the combinatorial definition of $\binom{b}{n}$ defined in terms of selecting subsets.

e. Since no bin may be left empty, this means that we need to put one ball into each bin. This means that we have $n - b$ balls left, and, since each bin now contains at least one, we can put them into the bins with no further restriction. This means we are in the case of part c. The only difference is that the number of balls that we have to distribute it $n-b$ not n. This means that the answer is

$$\binom{(n-b)+b-1}{n-b} = \binom{n-1}{n-b} = \binom{n-1}{(n-1)-(n-b)} = \binom{n-1}{b-1}$$