- 13.7 (Bowen Chen) Consider the set of all possible five-card poker hands.
    - How many atomic events are there?
    - What is the probability of each atomic event?
    - What is the probability of being dealt a royal straight flush? Four of a kind?

- 13.9 (Daryl Seah) In his letter of August 24, 1654, Pascal was trying to show how a pot of money should be allocated when a gambling game must end prematurely. Imagine a game where each turn consists of a roll of a die. Player *E* gets a point when the die is even, *O* gets a point when the die is odd. The first person to 7 points wins the pot. Suppose the game ends with *E* leading 4-2. How should the money be fairly split in this case? What is the general formula?

- 13.18 (Dani Chu) Suppose you are given a bag containing *n* unbiased coins. You are told that *n-1* coins are normal, with heads on one side and tails on the other, whereas one coin has heads on both sides.

  - Suppose you pick a coin at random, flip it, and get a head. What is the conditional probability that the coin you chose is the fake one?

  - Suppose you continue flipping for a total of *k* times and get all heads. Now what is the conditional probability that the coin you chose is the fake one?

  - Suppose you wanted to decide whether a chosen coin was fake by flipping it *k* times. The decision procedure returns *fake* if all flips come up heads, otherwise *normal*. What is the (unconditional) probability that this procedure makes an error?

- 13.14 (Chris Lee) Suppose you are given a coin that lands *heads* with probability x and *tails* with probability $1-x$. Are the outcomes of successive flips independent of each other given that you know the value of x? How about if you don't know the value of x? Justify your answer.

13.7 This is a classic combinatorics question that could appear in a basic text on discrete mathematics. The point here is to refer to the relevant axioms of probability: principally, axiom 3 on page 422. The question also helps students to grasp the concept of the joint probability distribution as the distribution over all possible states of the world.

a. There are $\binom{52}{5} = (52 \times 51 \times 50 \times 49 \times 48)/(1 \times 2 \times 3 \times 4 \times 5) = 2{,}598{,}960$ possible five-card hands.

b. By the fair-dealing assumption, each of these is equally likely. By axioms 2 and 3, each hand therefore occurs with probability 1/2,598,960.

c. There are four hands that are royal straight flushes (one in each suit). By axiom 3, since the events are mutually exclusive, the probability of a royal straight flush is just the sum of the probabilities of the atomic events, i.e., 4/2,598,960 = 1/649,740. For "four of a kind" events, There are 13 possible "kinds" and for each, the fifth card can be one of 48 possible other cards. The total probability is therefore $(13 \times 48)/2,598,960 = 1/4,165$.

13.9 Let e and o be the initial scores, m be the score required to win, and p be the probability that E wins each round. One can easily write down a recursive formula for the probability that E wins from the given initial state:

$$w_E(p,e,o,m) = \begin{cases} 1 & \text{if } e = m \\ 0 & \text{if } o = m \\ p \cdot w_E(p,e+1,o,m) + (1-p) \cdot w_E(p,e,o+1,m) & \text{otherwise} \end{cases}$$

This translates directly into code that can be used to compute the answer,

$$w_E(0.5, 4, 2, 7) = 0.7734375 .$$

With a bit more work, we can derive a nonrecursive formula:

$$w_E(p,e,o,m) = p^{m-e} \sum_{i=0}^{m-o-1} \binom{i+m-e-1}{i} (1-p)^i .$$

13.18

a. A typical "counting" argument goes like this: There are n ways to pick a coin, and 2 outcomes for each flip (although with the fake coin, the results of the flip are indistinguishable), so there are $2n$ total atomic events, each equally likely. Of those, only 2 pick the fake coin, and $2 + (n-1)$ result in heads. So the probability of a fake coin given heads, $P(\mathit{fake}|\mathit{heads})$, is $2/(2+n-1) = 2/(n+1)$.

Often such counting arguments go astray when the situation gets complex. It may be better to do it more formally:

$$\begin{aligned}\mathbf{P}(Fake|heads) &= \alpha\mathbf{P}(heads|Fake)\mathbf{P}(Fake)\\ &= \alpha\langle 1.0, 0.5\rangle\langle 1/n, (n-1)/n\rangle\\ &= \alpha\langle 1/n, (n-1)/2n\rangle\\ &= \langle 2/(n+1), (n-1)/(n+1)\rangle\end{aligned}$$

b. Now there are $2^k n$ atomic events, of which 2^k pick the fake coin, and $2^k + (n-1)$ result in heads. So the probability of a fake coin given a run of k heads, $P(\mathit{fake}|\mathit{heads}^k)$, is $2^k/(2^k + (n-1))$. Note this approaches 1 as k increases, as expected. If $k = n = 12$, for example, than $P(\mathit{fake}|\mathit{heads}^{10}) = 0.9973$.

Doing it the formal way:

$$\begin{aligned}\mathbf{P}(Fake|heads^k) &= \alpha\mathbf{P}(heads^k|Fake)\mathbf{P}(Fake)\\ &= \alpha\langle 1.0, 0.5^k\rangle\langle 1/n, (n-1)/n\rangle\\ &= \alpha\langle 1/n, (n-1)/2^k n\rangle\\ &= \langle 2^k/(2^k+n-1), (n-1)/(2^k+n-1)\rangle\end{aligned}$$

c. The procedure makes an error if and only if a fair coin is chosen and turns up heads k times in a row. The probability of this

$$P(heads^k|\neg fake)P(\neg fake) = (n-1)/2^k n\ .$$

If the probability x is known, then successive flips of the coin are independent of each other, since we know that each flip of the coin will land *heads* with probability x. Formally, if $F1$ and $F2$ represent the results of two successive flips, we have

$$P(F1 = heads, F2 = heads|x) = x * x = P(F1 = heads|x)P(F2 = heads|x)$$

Thus, the events $F1 = heads$ and $F2 = heads$ are independent.

If we do not know the value of x, however, the probability of each successive flip is dependent on the result of all previous flips. The reason for this is that each successive flip gives us information to better estimate the probability x (i.e., determining the posterior estimate for x given our prior probability and the evidence we see in the most recent coin flip). This new estimate of x would then be used as our "best guess" of the probability of the coin coming up *heads* on the next flip. Since this estimate for x is based on all the previous flips we have seen, the probability of the next flip coming up *heads* depends on how many *heads* we saw in all previous flips, making them dependent.

For example, if we had a uniform prior over the probability x, then one can show that after n flips if m of them come up heads then the probability that the next one comes up heads is $(m + 1)/(n + 2)$, showing dependence on previous flips.