ECE 302 Spring 2012.
Practice problems: Counting, discrete random variables, expectation, variance.
Ilya Pollak

These problems have been constructed over many years, using many different sources. If you find that some problem or solution is incorrectly attributed, please let me know at ipollak@ecn.purdue.edu.

Suggested reading: Sections 1.6–2.4 in the recommended text [3]. Equivalently, Sections 2.3, 2.6 (except for the multinomial law), 3.1, 3.4 (discrete random variables only), 3.5, and 3.6 in the Leon-Garcia text [7].

Problem 1. Gambling in Florence. (Ilya Pollak, after the discussion at the beginning of Chapter 4 of [8].)
Consider the game of independently throwing three fair six-sided dice. There are six ways in which the three resulting numbers can sum up to nine, and also six ways in which they can sum up to 10:

\[
\begin{align*}
9 &= 1 + 2 + 6 = 1 + 3 + 5 = 1 + 4 + 4 = 2 + 3 + 4 = 2 + 2 + 5 = 3 + 3 + 3 \\
10 &= 1 + 3 + 6 = 1 + 4 + 5 = 2 + 2 + 6 = 2 + 3 + 5 = 2 + 4 + 4 = 3 + 3 + 4.
\end{align*}
\]

This seems to suggest that the chances of throwing nine and 10 should be equal. Yet, a certain gambler in Florence in the early XVII century (most likely, the Grand Duke of Tuscany Cosimo II de Medici) noticed that in practice it is more likely to throw 10 than nine. Show that the probability to get a total of 10 is, in fact, larger than the probability to get a total of nine. (This problem was solved for the duke by Galileo Galilei.)

Solution. (Ilya Pollak.)
The key to solving this problem is carefully considering the sample space and enumerating the outcomes that constitute the two events. The sample space for the three throws consists of \(6^3 = 216\) outcomes. To see this, note that there are six possibilities for the first die. Each of these six possibilities can be combined with any one of the six possibilities for the second die, to give 36 different possibilities for the first two dice. Finally, each of these 36 possibilities can be combined with one of the six possibilities for the third die, to give a total of 216. Since the three throws are independent, the probability of each triple is \((1/6)^3 = 1/216\).

How many of these 216 triples sum to ten? If the first die results in “1,” the second and third will have to sum to nine. There are four ways for the second and third dice to sum to nine: (3,6), (4,5), (5,4), (6,3). If the first die results in “2,” the second and third will have to sum to eight. There are five ways in which the second and third dice can sum to eight: (2,6), (3,5), (4,4), (5,3), (6,2). If the first die results in “3,” the second and third will have to sum to seven, and this can be accomplished in six ways: (1,6), (2,5), (3,4), (4,3), (5,2), (6,1). If the first die results in “4,” the second and third will have to sum to six, which can be done in five ways: (1,5), (2,4), (3,3), (4,2), (5,1). If the first die results in “5,” the second and third will have to sum to five, which can be done in four ways: (1,4), (2,3), (3,2), (4,1). If the first die results in “6,” there are three acceptable possibilities for the remaining dies: (1,3), (2,2), (3,1). Thus, there is a total of \(4+5+6+5+4+3 = 27\) ways of obtaining ten in three throws, which means that the probability of throwing ten is \(27/216\).
How many ways are there of obtaining nine? If the first die results in 1, 2, 3, 4, 5, the sum of the second and third dice must be, respectively 8, 7, 6, 5, and 4. The numbers of ways to get these sums have been worked out in the previous paragraph: 5, 6, 5, 4, 3, respectively. If the first die results in 6, the sum of the second and third dice must be 3 which corresponds to either (2,1) or (1,2). Thus, the total number of outcomes resulting in the sum of nine is: 5+6+5+4+3+2 = 25. The probability of this event is 25/216.

**Problem 2. Benefits of Independent Retesting.** (Ilya Pollak, after Chapter 4 of [1].) Suppose your company is trying to decide whether or not to buy a certain software package. The vendor has given you a copy of the software for testing. You ask two people, Pat and Sam, to independently test the software. After the testing is done, Pat reports 20 flaws and Sam 30. (For example, a flaw could be either an outright bug, or the lack of a feature that your company needs.) Estimate the total number of flaws in the software

(a) if the total number of flaws they detect is 40 (i.e., their lists of flaws have 10 flaws in common);

(b) if the total number of flaws they detect is 49 (i.e., their lists of flaws have one flaw in common).

Assume that Pat and Sam’s detections of flaws can each be modeled as a sequence of $f$ independent Bernoulli trials where $f$ is the actual number of flaws, with probabilities of success (i.e., detection of a flaw) $p$ and $s$, respectively. Also assume that neither Pat nor Sam produce any “false alarms”—i.e., that they do not report a flaw where there is none. Estimate $p$ in terms of $f$ as the relative frequency of the event “Pat detects a flaw” in the $f$ Bernoulli trials. Similarly estimate $s$ and $ps$, then solve for $f$.

**Solution.** (Ilya Pollak.)

Let the total number of flaws be $f$, the number of flaws detected by Pat be $f_P$, the number of flaws detected by Sam be $f_S$, and the number of flaws detected by both be $f_{PS}$. The relative frequency estimates of $p$ and $s$ are then $p = f_P/f$ and $s = f_S/f$. The probability of any flaw to be detected by both Pat and Sam is $ps$, and so the relative frequency estimate of $ps$ is $ps = f_{PS}/f$. Substituting $p = f_P/f$ and $s = f_S/f$ into $ps = f_{PS}/f$, we get: $(f_P/f)(f_S/f) = f_{PS}/f$, and therefore $f = f_Pf_S/f_{PS}$. For Part (a), this is $f = 20 \cdot 30/10 = 60$. For Part (b), this is $f = 20 \cdot 30/1 = 600$.

If lots of their findings do not overlap, we know that they are not very good at detecting flaws and therefore there must be at lot more flaws than they detected. If their findings mostly overlap, and if they worked truly independently, we have no reason to suspect that there are many flaws beyond what they found.

**Problem 3. How to Construct Multiple-Choice Tests.** (Ilya Pollak.) Suppose 200 students take an exam consisting of $M$ multiple-choice questions, each with $N$ possible answers. All students decide to answer questions randomly and independently, i.e., each student’s probability to answer any question correctly is $1/N$, and all the answers are independent. Each question is worth one point, and a score of $M - 3$ or more is an “A”. What is the probability that at least one student will get an A? Express your answer in terms of $M$ and $N$. Evaluate this expression for two cases: 20 true-false questions (i.e., $M = 20$ and $N = 2$), and 20 questions with 10 options each (i.e., $M = 20$ and $N = 10$).
**Solution. (Ilya Pollak).**

Note that the complement of the event “at least one student gets an A” is “every student gets a B or less” which is equivalent to “every student has at least four incorrect answers”:

\[ P(\text{at least one student gets A}) = 1 - P(\text{every student has at least 4 wrong answers}) \]

Because the students come up with their answers independently, the probability that every student has at least four incorrect answers is equal to the product, over all 200 students, of each student’s probability to have at least four incorrect answers. Since the students’ correct answer probabilities are identical, the probability for any one student to have at least four incorrect answers is the same as the probability for any other student.

\[ P(\text{every student has at least 4 wrong answers}) = (P(\text{one student has at least 4 wrong answers}))^{200} \]

The event that one student has at least four incorrect answers is the complement of the event that one student has at most three incorrect answers:

\[ P(\text{one student has at least 4 wrong answers}) = 1 - P(\text{one student has at most 3 wrong answers}) \]

For a specific student, each sequence of \( k \) incorrect and \( M - k \) correct answers has probability \( (1/N)^{M-k}(1-1/N)^k \) because he/she answers all the questions independently. Exactly one sequence corresponds to getting all the correct answers, and the probability of this event is \( (1/N)^{M} \). \( M \) sequences correspond to getting exactly one incorrect answer because the incorrect answer could be for any one of the \( M \) questions. Therefore, the probability for a student to get exactly one answer incorrectly is \( M \cdot (1/N)^{M-1} \cdot (1-1/N) \). The number of ways to get two incorrect answers out of \( M \) is

\[
\binom{M}{2} = \frac{M!}{(M-2)!2!} = \frac{M(M-1)}{2}.
\]

The number of ways to get three incorrect answers out of \( M \) is

\[
\binom{M}{3} = \frac{M!}{(M-3)!3!} = \frac{M(M-1)(M-2)}{6}.
\]

Therefore, the probability for a student to get at most three incorrect answers is

\[
\sum_{k=0}^{3} \binom{M}{k} \left( \frac{1}{N} \right)^{M-k} \left( \frac{N-1}{N} \right)^{k} = \left( 1 + M(N-1) + \frac{M(M-1)(N-1)^2}{2} + \frac{M(M-1)(M-2)(N-1)^3}{6} \right) \cdot \left( \frac{1}{N} \right)^{M}.
\]

Plugging this back into the above formulas gives:

\[
P(\text{at least one student gets an A})
= 1 - \left[ 1 - \left( 1 + M(N-1) + \frac{M(M-1)(N-1)^2}{2} + \frac{M(M-1)(M-2)(N-1)^3}{6} \right) \cdot \left( \frac{1}{N} \right)^{M} \right]^{200}
\]
For $M = 20$,

$$P(\text{at least one student gets an A}) = 1 - \left[ 1 - (1 + 20(N - 1) + 190(N - 1)^2 + 1140(N - 1)^3) \left( \frac{1}{N} \right)^{20} \right]^{200}$$

For $N = 2$, this simplifies to

$$P(\text{at least one student gets an A}) = 1 - \left[ 1 - \left( \frac{1}{2} \right)^{20} \right]^{200} \approx 0.227.$$  

For $N = 10$, we have

$$P(\text{at least one student gets an A}) = 1 - \left[ 1 - (1 + 20 \cdot 9 + 190 \cdot 81 + 1140 \cdot 729) \left( \frac{1}{10} \right)^{20} \right]^{200}$$

$$= \left[ 1 - 846631 \left( \frac{1}{10} \right)^{20} \right]^{200} \approx 1.7 \times 10^{-12}.$$  

Thus, if the test consists of true-false questions, there is a sizeable chance of about 23% that at least one student will get an A if all students simply guess the answers. On the other hand, if each question has $N = 10$ choices, then the probability that random answers will produce an A in a class of 200 students is tiny.

**Problem 4. How to rate mutual funds? (Ilya Pollak).**

There are over 8000 mutual funds in the United States. Suppose we measure the performance of mutual funds by whether they overperform or underperform some market average (say, the S&P 500 index). Let us consider the possibility that all the funds are independent, and have equal chances to be above or below the market in any given year. Assume that the total number of funds is exactly 8000. Based on these assumptions,

(a) what is the probability that at least one mutual fund out of 8000 will perform above the market average for 10 consecutive years, starting with the next year?

(b) what is the probability that at least one fund will perform above average for 15 consecutive years, starting with the next year?

**Solution.** To look at the 10-year performance, we are modeling the 8000 mutual funds as 8000 independent sequences of 10 Bernoulli trials. The reasoning is the same as in the problem on true-false test questions. The probability, for any one fund, to have at least one below-average year, is $1 - (1/2)^{10}$. The probability that each of them will have at least one below-average year is $(1 - (1/2)^{10})^{8000}$. The probability that at least one of them will only have above-average years is $1 - (1 - (1/2)^{10})^{8000} \approx 0.9996$. Similarly, the probability for at least one to have 15 above-average years is $1 - (1 - (1/2)^{15})^{8000} \approx 0.22$.

In other words, even if the funds’ performance is random, there are so many of them that it is virtually guaranteed that some will consistently outperform the market for 10 years in a row,
starting at any given year. For outperformance during 15 consecutive years starting in a given year, there is still a very considerable chance of about 22%.

(c) what is the probability that there will be at least one mutual fund that outperforms the market for each of 15 consecutive calendar years during some 15-year interval during the next 50 years? To answer this question, let $p(s, n)$ be the probability that one mutual fund will achieve a streak of $s$ years during the next $n$ years, and go through the following steps.

(i) Argue that the probability that we are looking for is $1 - (1 - p(15, 50))^{8000}$.

Solution. The argument is the same as above. The probability that a specific fund will not achieve a streak of 15 years during the next 50, is $1 - p(15, 50)$. The probability that none of them will is $(1 - p(15, 50))^{8000}$. The probability that at least one will is $1 - (1 - p(15, 50))^{8000}$.

(ii) Argue that $p(s, n) = 0$ for $n < s$ and that $p(s, n) = (1/2)^s$ for $s = n$.

Solution. In order to achieve a streak of $s$ wins, one needs to play at least $s$ games, therefore $p(s, n) = 0$ for $n < s$. When $n = s$, then $p(s, n)$ is the probability to have $n$ above-average years out of $n$, which is $p(s, n) = (1/2)^s$, since we are modeling the process as a sequence of $n$ independent Bernoulli trials with probability of success $1/2$.

(iii) Argue that, for $n > s$, $p(s, n) = p(s, n - 1) + (1 - p(s, n - s - 1)) \cdot (1/2)^{s+1}$.

Hint. Define the following events:

$A$ = “a streak of length $s$ happens during years 1 through $n$”

$B$ = “a streak of length $s$ happens during years 1 through $n - 1$”

$C$ = “a streak of length $s$ happens in years $n - s + 1, n - s + 2, \ldots, n$”

$D$ = “year $n - s$ is a below-average year”

$E$ = “a streak of length $s$ does not happen at any time during years 1 through $n - s - 1$”

Argue that events $B$ and $C \cap D \cap E$ form a partition of the event $A$. Further argue that the events $C$, $D$, and $E$ are independent. Then argue that $P(A) = p(s, n)$, $P(B) = p(s, n - 1)$, $P(C) = (1/2)^s$, $P(D) = 1/2$, and $P(E) = 1 - p(s, n - s - 1)$.

Solution. The event “streak of length $s$ in $n$ years” can happen in the following two mutually exclusive ways. Either the streak happened in the first $n - 1$ years, or the streak reached length $s$ during the $n$-th year. The probability of the event that the streak happened during the first $n - 1$ years is $p(s, n - 1)$. The event that the first occurrence of the streak was completed on the $n$-th year means that there was no streak during the first $n - s - 1$ years, then year $n - s$ was a below-average year, then the remaining $s$ years were above-average years. The probability of this is $(1 - p(s, n - s - 1)) \cdot (1/2) \cdot (1/2)^s$.

(iv) Use the following Matlab code to compute the final answer:

```matlab
n=50; s=15;
probs = zeros(1,n);
probs(s) = (1/2)^s;
probs(s+1) = (1/2)^s + (1/2)^(s+1);
for k=s+2:n
    probs(k) = probs(k-1) + (1-probs(k-s-1))*(1/2)^(s+1);
end;

% After this calculation, probs(50) is p(15,50).
```
The answer is 0.99. The probability \( p(15, 50) \) for a single fund is about 0.00056. Thus, for a specific fund chosen \textit{a priori}, the probability to hit an above-average streak of 15 years some time during the next 50 years is extremely small. However, simply because the total number of funds is so large, the chances are overwhelming that there will be at least one fund that will have a 15-year streak. In fact, if the total number of funds is 8000, the expected number of funds that will have such a streak some time during the next 50 years is about 4.5.

This problem is inspired by the notorious 15-year streak of Legg Mason Value Trust mutual fund (LMVTX on Yahoo Finance) which has been discussed in many articles and books, e.g., [8, 12]. It beat the S&P 500 index (\(^\ast\)GSPC on Yahoo Finance) for each calendar year from 1991 to 2005. Its cumulative return from 1991 to 2005 was about 480% compared to about 278% for S&P 500. Since then, however, LMVTX lost 41% while S&P 500 gained 7.7%, as of close on Friday, February 3, 2012. The cumulative result is that a dollar invested in LMVTX just before the beginning of 1991 would become only $3.42 on 2/3/12, whereas a dollar invested in S&P 500 would become $4.07. This is something to keep in mind when you look for a mutual fund to invest in. The fact that someone has consistently beaten the market over the last few years just \textit{might} be the result of a random fluctuation, simply due to the fact that there are thousands of mutual funds out there. If in fact the past performance \textit{was} solely due to random chance then it is irrelevant in trying to predict the future performance. (In fact, even if the past good performance was due to the great skills of the fund manager rather than luck, there are still many compelling reasons why the future may still be unpredictable, such as (a) the particular market anomaly that the fund manager has been exploiting may be discovered by many other market participants and therefore cease to be profitable; (b) the fund manager may get another job and leave, and be replaced by a less-competent person.)

**Problem 5. On the importance of default dependencies in the valuation of mortgage-backed securities.** (Ilya Pollak).

A typical mortgage-backed security (MBS) pools many mortgages together. The security has several tranches with different levels of seniority. The most junior tranche, usually called the equity tranche, is the riskiest and has the highest rate of return, in order to compensate the holders of this tranche for the risk they are taking. The senior tranche is the least risky but has the smallest rate of return. The interest payments are made to the senior tranche first. If there is anything left over after the promised rate of return for the senior tranche is satisfied, the next tranche receives its interest payments, etc. If the underlying mortgages default and cause the loss of capital, the equity tranche usually suffers losses first: the capital of the senior tranche is untouched until all the tranches below it are completely wiped out. If you are interested in more specifics on how mortgage-backed securities work, please ask questions through the forum or during the office hours or the help session. You do not actually need to know anything about mortgage-backed securities in order to solve this
This problem explores the importance of correctly accounting for the dependencies among the underlying mortgages in order to be able to properly value the MBS that they comprise. Suppose an MBS of duration one year is based on 10 1-year underlying mortgages, $200,000 each. Assume that, for each mortgage, one of two things can happen:

- Either the mortgage is paid back in full with 5% interest at the end of the year,
- or it defaults with 100% loss.

For each mortgage, the probability of default is 0.1. Suppose that the MBS consists of two tranches: the senior tranche which has 60% of the capital and the equity tranche which has 40% of the capital. The senior tranche earns 3% interest, the equity tranche earns 8%.

(a) What is the probability that the senior tranche will lose some capital, if the defaults of individual mortgages are independent? Note that the event that the senior tranche will lose some capital is the event that five or more out of the ten mortgages will default.

**Solution.** If the defaults are independent, the probability that exactly \( k \) mortgages out of 10 will default is:

\[
\binom{10}{k} 0.1^k 0.9^{10-k}
\]

For \( k = 0 \), this is \( 0.9^{10} \approx 0.348678 \). For \( k = 1 \), this is \( 10 \cdot 0.1 \cdot 0.9^9 \approx 0.387420 \). For \( k = 2 \), this is \( 45 \cdot 0.1^2 \cdot 0.9^8 \approx 0.193710 \). For \( k = 3 \), this is

\[
\frac{10!}{3! \cdot 7!} \cdot 0.1^3 \cdot 0.9^7 = \frac{8 \cdot 9 \cdot 10}{6} \cdot 0.1^3 \cdot 0.9^7 \approx 0.057396.
\]

For \( k = 4 \), this is

\[
\frac{10!}{4! \cdot 6!} \cdot 0.1^4 \cdot 0.9^6 = \frac{7 \cdot 8 \cdot 9 \cdot 10}{24} \cdot 0.1^4 \cdot 0.9^6 \approx 0.011160.
\]

Therefore, the probability that there will be zero, one, two, three, or four defaults is 0.348678 + 0.387420 + 0.193710 + 0.057396 + 0.011160 \( \approx 0.998 \). Hence, the probability of at least five defaults is approximately \( 1 - 0.998 = 0.002 \).

(b) What is the probability that the senior tranche will lose some capital if the defaults of individual mortgages are perfectly dependent (i.e., either all ten default or none of them defaults)?

**Solution.** If defaults happen in perfect unison, then the probability of at least five defaults is the same as the probability of a single default, which is 0.1.

(c) What is the probability that the senior tranche is completely wiped out if the defaults are independent? Note that this happens when all ten mortgages default.

**Solution.** Under the independence assumption, the probability of a complete wipe-out is \( 0.1^{10} = 0.0000000001 \) which is very, very small!
(d) What is the probability that the senior tranche is completely wiped out if the defaults are perfectly dependent?

**Solution.** Under the perfect unison assumption, the probability of a complete wipe-out is the same as the probability of one default, which is 0.1.

[The remaining text of this solution is optional reading.]

The point of this problem is this. Suppose that, based on the historical data, we estimate that the individual default probability for a mortgage is 0.1. Even if our estimate is very accurate, this is not enough to determine the value of the security, because the value is very much tied to the dependency structure of the underlying mortgages. In our example, if the mortgages move in perfect unison, the probability that the senior tranche loses some capital is 50 times as big as the same probability under the independence assumption. The probability of a complete wipe-out under the independence assumption is one-billionth of the same probability under the perfect unison assumption!

One of the factors significantly contributing to the recent meltdown of the credit markets is the fact that, for a security that is a derivative of a pool of assets, the dependency structure of the underlying assets is very difficult to accurately forecast based on historical data. The reason for this is that, as we saw in class, conditioning affects independence. Conditioned on a normal state of the real estate market, the defaults of someone in New York and someone in California are essentially independent. Conditioned or rapidly falling real estate prices, many people in New York, California, and many other places suddenly owe more on their house than the house is worth, and therefore start intentionally defaulting in unison. Thus, if the dependency structure is estimated during normal market times, the investors would have an unrealistically optimistic view of the risk associated with their MBS.

A word of caution: the scenario outlined in this problem is a toy example designed to be worked out with the limited tools that we have covered in this course so far. While it does give you some flavor for what has happened with the credit markets, the reality is much more complicated, in many respects, some of which are:

- Mortgages do not last for a fixed amount of time due to the possibility of default or prepayment at any time.
- Pools of mortgages comprising an MBS usually involve hundreds or thousands of mortgages, not 10.
- An MBS typically has more than two tranches.
- When a mortgage defaults, the loss is not 100% but rather is equal to the difference between the amount still owed on the mortgage and the price of the house.
- In addition to the MBSs themselves, a major contributor to the credit crisis were the collateralized debt obligations (CDOs) built from the MBSs. CDOs are built from MBSs the same way as MBSs are built from mortgages: by pooling many MBSs together and creating several tranches with different levels of seniority.

**Problem 6. (Ilya Pollak, after Gilbert and Mosteller [6].)**

Suppose you are interviewing people for the position of ECE 302 instructor. You have four candidates
who are interviewed consecutively: 1, 2, 3, and 4. If you interview any two of them, you will be able to determine which one of the two is better. The problem is, you are required to tell the decision to each candidate at the end of his/her interview: you either reject the candidate and go on to the next candidate, or you hire the candidate and reject all the remaining candidates without interviewing them. If you reject the first three candidates, you must hire the fourth. Before the interviews start, you have no information about the relative strength of the candidates, and so you assume that the best candidate is equally likely to be candidate 1, 2, 3 or 4; the second-best candidate is also equally likely to be candidate 1, 2, 3, or 4, etc.

(a) Suppose your strategy is to hire the candidate 1, regardless of what happens during his/her interview. What is the probability that you will hire the best candidate?

(b) Suppose you interview and reject candidate 1, and then continue the interviews and hire the first candidate you see who is better than candidate 1. I.e., if candidate 2 is better than candidate 1, you hire candidate 2; if candidate 2 is worse than candidate 1 then you interview candidate 3. If you interview candidate 3, then: if he is better than candidate 1, you hire candidate 3, otherwise you hire candidate 4. What is the probability that you will hire the best candidate?

(c) Suppose you interview and reject first two candidates, and then hire candidate 3 if he is better than both candidates 1 and 2, and hire candidate 4 otherwise. What is the probability that you will hire the best candidate?

Solution. Let A be the best candidate, B second best, C third best, and D the worst.

(a) Since A has 1/4 probability to be interviewed first, the answer is 1/4.

(b) A tree model for all possible sequences of the four interviewed candidates is shown in Fig. 1. The outcomes which correspond to rejecting A are marked with a +, the outcomes which correspond to rejecting A are marked with a −. If A is the first one to be interviewed, which happens with probability 1/4, then he is certain to be rejected. This is the top branch in the diagram. If B is the first one to be interviewed, which also happens with probability 1/4, then A is certain to be hired since he is the only one of the remaining three candidates who is better than B. This is the second branch in the diagram. If C is the first one to be interviewed, then there are several possibilities which are all indicated in the diagram and which come down to the following: if B is interviewed before A then A is rejected because B is better than A, and otherwise A is hired. The conditional probability of hiring A given that C is first is therefore 1/2. If D is the first to be interviewed, the only case when A gets hired is if he is the second to be interviewed, which happens with conditional probability 1/3. Using the total probability theorem, we therefore have the following result for the probability of hiring A:

\[
\frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{3} = \frac{11}{24}.
\]

Another way of arriving at the same result is to notice that A never gets hired if he is the first one to be interviewed; he always gets hired if he is the second one; he gets hired with conditional probability 1/2 if he is the third one (since the second candidate has 1/2 probability to be worse than the first), and he gets hired with conditional probability 1/3 if he is the last one (since the first candidate has a 1/3 probability to be worse than the second and the third).
(c) Now if A is among the first two candidates, which happens with probability 1/2, he will be rejected. If he is the third candidate, which happens with probability 1/4, he will be hired. If he is the fourth candidate, which also happens with probability 1/4, he will only be rejected if the third candidate is better than the first two—i.e., if the third candidate is B, which happens with conditional probability 1/3. The total probability of hiring A is therefore:

\[
\frac{1}{2} \cdot 0 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot \frac{2}{3} = \frac{5}{12}
\]

Note that Part (b) describes the best one out of the three strategies.

Now suppose we have \( N \) candidates. It can be shown that the best strategy, in the sense of maximizing the probability to hire the best candidate, is to reject the first \( r \) and then hire the first one among the remaining candidates who is better than the \( r \) initial candidates, where \( r \) is the smallest integer for which

\[
\frac{1}{r+1} + \frac{1}{r+2} + \frac{1}{r+3} + \ldots + \frac{1}{N-1} < 1.
\]

It turns out that the reasoning above can be generalized to yield the following probability of hiring the best candidate:

\[
\frac{1}{N} \left( 1 + \frac{r}{r+1} + \frac{r}{r+2} + \frac{r}{r+3} + \ldots + \frac{r}{N-1} \right).
\]

Interestingly, this turns out to converge from above to \( e^{-1} \approx 0.368 \) in the limit \( N \to \infty \). In other words, no matter how many candidates you have, the optimal strategy yields a better than 1-in-3 chance to hire the best
one! This and several related problems are solved in [6]. See pp. 38–39 of this paper for the formulation of the problem and the solution for an arbitrary positive integer \( N \).

![Diagram of communication network for Problem 7](image)

**Figure 2: Communication network for Problem 7.**

**Problem 7.** *(After Drake [4], Problem 1.23.)*

In the communication network of Fig. 2, link failures are independent, and each link has a probability of failure of \( p \). Consider the physical situation before you write anything. \( A \) can communicate with \( B \) as long as they are connected by at least one path which contains only in-service links. *(Note: \( a, b, c, d, e, f, \) and \( g \), are just the names of the links, NOT probabilities.)*

(a) Given that exactly four links have failed, determine the conditional probability that \( A \) can communicate with \( B \).

(b) Given that exactly four links have failed, determine the probability that either \( g \) or \( f \) *(but not both)* is still operating properly.

(c) Given that \( a, b, \) and \( c \) have failed *(but no information about the condition of other links)*, determine the probability that \( A \) can communicate with \( B \).

**Solution.** *(Ilya Pollak.)*

(a) Since the failures of all groups of exactly four links are equally likely *(namely, the probability that any specific four links have failed and the other three are working, is \( p^4(1-p)^3 \)*, our conditional probability law is uniform. Therefore, we need to simply count the number of outcomes in the event “\( A \) can communicate with \( B \)” given that four links have failed, and divide it by the number of the outcomes in the conditioning event “four links have failed”. The total number of groups of four links out of seven that could have failed is \( \binom{7}{4} = \frac{7!}{3!4!} = \frac{5 \cdot 6 \cdot 7}{1 \cdot 2 \cdot 3} = 35 \). The only situations when four links are out but \( A \) can still communicate with \( B \) are when \( a, b, \) and \( c \) are up, or when \( e, f, \) and \( g \) are up. The conditional probability is therefore \( 2/35 \).

(b) The number of four-link combinations which involve \( g \) but not \( f \) is \( \binom{5}{3} = \frac{5!}{2!3!} = 10 \); the same number of four-link combinations involve \( f \) but not \( g \). The answer is therefore \( 20/35 = 4/7 \).
(c) The conditional probability that $A$ can communicate with $B$ is the probability that $e$, $f$, and $g$ are working (and it does not matter whether $d$ has failed or not), which is $(1-p)^3$.

Problem 8. (Bertsekas-Tsitsiklis supplementary problems to Chapter 1, Problem 19 [2].)
An internet access provider (IAP) owns two servers. Each server has a 50% chance of being “down” independently of the other. Fortunately, only one server is necessary to allow the IAP to provide service to its customers, i.e., only one server is needed to keep the IAP’s system up. Suppose a customer tries to access the internet on four different occasions, which are sufficiently spaced apart in time, so that we may assume that the states of the system corresponding to these four occasions are independent. What is the probability that the customer will only be able to access the internet on three out of the four occasions?

Solution. The probability that the customer does not receive service on any one occasion is the probability that both servers are down, which is $1/4$. The probability that he does receive service on any one occasion is therefore $1-1/4=3/4$. The probability that the customer receives service three out of four times is given by the binomial formula:

$$\binom{4}{3} \cdot \left(\frac{3}{4}\right)^3 \cdot \left(\frac{1}{4}\right) = \frac{27}{64}.$$ 

A subway train made up of $n$ cars is boarded by $r$ passengers ($r \leq n$), each passenger choosing a car at random (using the discrete uniform probability law) and independently of other passengers. What is the probability of the passengers all ending up in different cars?

Solution. (Ilya Pollak.)
Since each passenger can be in one of $n$ cars, the total number of outcomes in the sample space is $n^r$, and all of them are equally likely because we assume uniform probability law. The number of outcomes in the event that all passengers end up in different cars is $n(n-1)(n-2)\cdots(n-r+1)$ since the first passenger can be in any one of the $n$ cars, the second passenger can choose any of the $n-1$ remaining cars, etc. The answer is therefore:

$$\frac{n(n-1)(n-2)\cdots(n-r+1)}{n^r}.$$ 

A wooden cube with painted faces is sawed up into 1000 little cubes, all of the same size. The little cubes are then mixed up, and one is chosen at random. What is the probability of its having exactly two painted faces?

Solution. (Ilya Pollak.)
In order for a little cube to have exactly two painted faces, it would have to have come from an edge of the big cube but not from a corner. There are 12 edges each of which has 10-2=8 little cubes which are not in the corners, for a total of 96 little cubes. This gives the probability of $96/1000 = 0.096$.

Problem 11. (Rozanov [11] Chapter 1, Problem 6.)
Suppose $n$ people sit down at random and independently of each other in an auditorium containing $n+k$ seats. What is the probability that $m$ seats specified in advance ($m \leq n$) will be occupied?
**Solution.** (Ilya Pollak.)

There are $\binom{n+k}{n}$ different sets of $n$ seats that can be occupied. If $m$ people are occupying the $m$ specified seats, the remaining $n-m$ people can be occupying any of the $\binom{n+k-m}{n-m}$ subsets of the remaining $n+k-m$ seats. The probability is therefore

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

Note that the same answer is obtained if we use permutations instead of combinations. Indeed, there are $(n+k)!/k!$ $n$-permutations of the $n+k$ seats. When $m$ people are occupying the $m$ specified seats, there are $n!/(n-m)!$ $m$-permutations of such people out of $n$, and there are $(n+k-m)!/k! (n-m)$-permutations of the remaining $n+k-m$ seats to be occupied by the remaining people. The probability is therefore

$$\frac{n!/(n-m)!((n+k-m)!/k!)}{(n+k)!/k!} = \frac{n!(n+k-m)!}{(n-m)!(n+k)!},$$

which is the same answer as before.

**Problem 12.** (Ilya Pollak and Bin Ni.)

In blackjack, the objective is to get as close as possible to 21 without going over 21. The cards are valued as follows:

- An Ace can count as either 1 or 11, whichever makes the best hand.
- The cards from 2 through 10 are valued as indicated.
- Jacks, Queens, and Kings are all valued at 10.

Suppose we use a 52-card deck.

(a) We draw two cards at random from a full deck, so that each pair of cards is equally likely to be drawn. What is the probability to get a blackjack (i.e., a total of 21)?

**Solution.** The two cards must be an Ace and a 10-valued card—i.e., a ten, a Jack, a Queen, or a King. Since there are four Aces and $4 \cdot 4 = 16$ 10-valued cards in a deck, the number of such combinations is: $4 \cdot 16 = 64$. The total number of combinations of any two cards is: $\binom{52}{2} = 1326$. Therefore the probability that we get a blackjack is:

$$\frac{64}{1326} = \frac{32}{663} \approx 0.0483.$$
(b) Suppose we started with a full deck and drew the 9 of clubs and the 9 of diamonds. What is the conditional probability that, if we draw another card, we bust (i.e., go over 21)? What is the conditional probability that this third card we draw will be a 3? Assume that, after the 9 of clubs and the 9 of diamonds are drawn, each of the remaining 50 cards is equally likely to be drawn.

**Solution.** Given the two 9’s are out, we are left with 50 cards. The first question is equivalent to the conditional probability of getting a card greater than 3. What we can do for both questions is to count how many cards are greater than (equal to) 3 among the 50 cards left and then divide by 50, the total number of choices.

The number of cards from 4 to King is $10 \cdot 4 - 2 = 38$ excluding the two 9’s that are already out. Therefore the probability of a bust is:

$$\frac{38}{50} = 0.76.$$ 

The number of cards equal to 3 is 4. So the probability to get a 21 is:

$$\frac{4}{50} = 0.08.$$ 

(c) Suppose we started with a full deck and drew a 7 of clubs and an 8 of spades. What is the conditional probability that the next card leads to a bust?

**Solution.** We need a card greater than or equal to 7 to bust. The number of such cards given that one 7 and one 8 are out is: $7 \cdot 4 - 2 = 26$. The total number of choices of the third card is 50. Hence the probability of bust is:

$$\frac{26}{50} = 0.52.$$ 

(d) Suppose we started with a full deck and drew a 5 of hearts and a 10 of clubs. What is the conditional probability that the next card leads to a bust?

**Solution.** Again, we need a card greater than or equal to 7 to bust. The number of such cards given that one 5 and one 10 are out is: $7 \cdot 4 - 1 = 27$. The total number of choices of the third card is 50. Hence the probability of a bust is:

$$\frac{27}{50} = 0.54.$$ 

**Problem 13. (Ilya Pollak.)**

An important milestone in the development of the theory of probability was the exchange of letters between Blaise Pascal (1623-1662) and Pierre Fermat (1601-1665), two French mathematicians. One motivation for these letters was problems proposed to Pascal by his friend Chevalier de Méré, in particular the problem of points, described in Problem 38 of Chapter 1 in [3]. Another of de Méré’s problems is Problem 49 in Chapter 1 [3], and here is yet another one ( [5], page 56).
De Méré bet that at least one 6 would appear during a total of four independent rolls of a fair six-sided die. From past experience, he knew that he was more successful than not with this game of chance. Tired of his approach, he decided to change the game. He bet that a double 6 would appear at least once during twenty-four independent rolls of two dice. Soon he realized that his old approach to the game resulted in more money. Please help him figure out why his new approach was not as profitable. (Hint. First, count the total number of outcomes of the first game. How many of them result in no 6’s?)

**Solution.** (Ilya Pollak and Bin Ni.)

In the first game, each roll of the die results in 6 equally like numbers. The total number of outcomes of four rolls is therefore $6^4$. If there are no 6’s in four rolls, each roll can only vary from 1 to 5, and so the total number of such results is $5^4$. Therefore, the probability that no 6’s appear in 4 rolls of a die is:

$$\frac{5^4}{6^4} \approx 0.48225,$$

and the probability that de Méré wins the bet is:

$$1 - \frac{5^4}{6^4} \approx 1 - 0.48225 = 0.51775.$$

In the second game, each roll of two dice can result in $6^2 = 36$ equally likely outcomes. The total number of outcomes of 24 rolls is therefore $36^{24}$. If there is not a single double 6 in the 24 rolls, each roll can only have 35 possible outcomes and the total number of such results is $35^{24}$. Therefore, the probability that no double 6’s appear in 24 rolls of two dice is:

$$\frac{35^{24}}{36^{24}} \approx 0.50860,$$

and the probability that de Méré wins the bet is:

$$1 - \frac{35^{24}}{36^{24}} \approx 1 - 0.50860 = 0.49140.$$

We can see that the probability of winning the second game is slightly less than that of the first game. This is reason why he found the old game was more profitable. (Note that if he bets even money, he is likely to increase his wealth by playing the first game many times and to decrease his wealth by playing the second game many times. This is because the probabilities of winning are above 0.5 and below 0.5, respectively.)

**Problem 14.** (Mosteller [9], Problem 1(a).)

A drawer contains red socks and black socks. When two socks are drawn at random, the probability that both are red is 1/2. What is the smallest possible number of socks in the drawer?

**Solution.** (Ilya Pollak and Bin Ni.)

Suppose there are $n$ socks in the drawer and $m$ are red. According to the given condition, the number
of red socks should be at least 2, i.e. \( n > m \geq 2 \). The total number of combinations of two red socks is:

\[
\binom{m}{2} = \frac{m!}{2!(m-2)!} = \frac{m(m-1)}{2}.
\]

The total number of combinations of any two socks is:

\[
\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2}.
\]

Hence the probability of getting a pair of red socks should be:

\[
\frac{m(m-1)/2}{n(n-1)/2} = \frac{m(m-1)}{n(n-1)} = \frac{1}{2}.
\]

(1)

In order to find the smallest possible \( n \) satisfying Eq. (1), we can set up a table of small integers \( n \) with their values of \( n(n-1) \) as follows:

<table>
<thead>
<tr>
<th>( n )</th>
<th>( n(n-1) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
</tr>
</tbody>
</table>

The smallest \( m, n \) satisfying Eq. (1) are \( n = 4, m = 3 \). Therefore there should be at least 4 socks in the drawer.

**Problem 15.** *(Drake [4], Problem 1.28. Solutions by Ilya Pollak and Bin Ni.)*

Companies A, B, C, D and E each send three delegates to a conference. A committee of four delegates, selected by a lot, is formed. Determine the probability that:

(a) Company A is not represented on the committee.

**Solution.** Forming a committee of four without delegates from company A is equivalent to forming a committee from the 12 representatives of the remaining companies. The number of ways to do this is \( \binom{12}{4} \). The total number of combinations of 4 out of all the 15 delegates is \( \binom{15}{4} \). Therefore, the probability of a committee with no delegates from company A is:

\[
\frac{\binom{12}{4}}{\binom{15}{4}} = \frac{12!}{4!8!} \cdot \frac{15!}{13\cdot14\cdot15} = \frac{3\cdot11}{13\cdot7} = \frac{33}{91} \approx 0.36264.
\]

(b) Company A has exactly one representative on the committee.

**Solution.** Again, we first need to determine the number of ways of forming a committee with exactly one representative from company A, and then divide by \( \binom{15}{4} \). To form such a
committee, we first need to choose a person out of the three from company A. Of course there are 3 choices. For each choice, there are \( \binom{12}{3} \) ways to determine the other 3 committee members from the 12 representatives of companies B, C, D, and E. Therefore, the total number of ways to form a committee with only one member from company A is \( 3 \cdot \binom{12}{3} \). The probability of such a committee is:

\[
\frac{3 \cdot \binom{12}{3}}{\binom{15}{4}} = \frac{3 \cdot \frac{12!}{3!9!}}{\frac{15!}{4!11!}} = \frac{10 \cdot 11 \cdot 3 \cdot 4}{13 \cdot 14 \cdot 15} = \frac{11 \cdot 4}{91} \approx 0.48352.
\]

(c) Neither company A nor company E is represented on the committee.

**Solution.** In this case, all the committee members are from companies B, C, D. The number of combinations of 4 out of the 9 delegates from these companies is \( \binom{9}{4} \), and so the probability of such a committee is:

\[
\frac{\binom{9}{4}}{\binom{15}{4}} = \frac{\frac{9!}{4!5!}}{\frac{15!}{4!11!}} = \frac{6 \cdot 7 \cdot 8 \cdot 9}{12 \cdot 13 \cdot 14 \cdot 15} = \frac{2 \cdot 9}{13 \cdot 15} = \frac{6}{65} \approx 0.09231.
\]

**Problem 16.** *(Drake [4], Problem 1.30. Solution by Ilya Pollak and Bin Ni.)*

The Jones family household includes Mr. and Mrs. Jones, four children, two cats and three dogs. Every six hours there is a Jones family stroll. The rules for a Jones family stroll are:

- Exactly five beings (people + dogs + cats) go on each stroll.
- Each stroll must include at least one parent and at least one pet.
- There can never be a dog and a cat on the same stroll unless both parents go.
- All acceptable stroll groupings are equally likely.

Given that exactly one parent went on the 6pm stroll, what is the conditional probability that Rover, the oldest dog, also went?

**Solution.** Let us first determine how many acceptable stroll groupings there are with one parent and Rover. We have two choices for the parent and for each choice, we need to select the other 3 members from the 4 children, 2 cats and the other 2 dogs. According to rule 3, no cat is acceptable, and so we are left with 6 candidates. Therefore, the number of possible groupings with one parent and Rover is:

\[
2 \cdot \binom{6}{3} = 2 \cdot \frac{6 \cdot 5 \cdot 4}{3 \cdot 2 \cdot 1} = 40.
\]

Now we need to determine the total number of acceptable groupings with one parent. We still have two choices for the parent. For each choice, we have to determine the other 4 group members from
the 4 children, 2 cats and 3 dogs. Based on the rules, there are 2 possibilities regarding those four members: (1) At least one dog and some children, no cat; (2) At least one cat and some children, no dog. The number of combinations for the first case is:

\[
\binom{7}{4} - 1 = \frac{7 \cdot 6 \cdot 5 \cdot 4}{4 \cdot 3 \cdot 2 \cdot 1} - 1 = 34.
\]

This is because we are choosing 4 out of 7 candidates, and we need to exclude the combination of four children. Similarly, the number of combinations for the second case is:

\[
\binom{6}{4} - 1 = \frac{6 \cdot 5 \cdot 4 \cdot 3}{4 \cdot 3 \cdot 2 \cdot 1} - 1 = 14.
\]

Therefore, the total number of acceptable groupings with exactly one parent is:

\[
2 \cdot (34 + 14) = 96.
\]

(There are many other ways to get this number. All you need to do is partition all the possibilities into several disjoint cases and carefully count the number of combinations for each case.) Since we know that 40 out of the 96 acceptable groupings have Rover in them, we end up with the required probability equal to:

\[
\frac{40}{96} = \frac{5}{12} \approx 0.41667.
\]

**Problem 17. (Ilya Pollak and Bin Ni.)**

A poker hand consists of five cards. The different possible hands, from the lowest to the highest, are:

- One pair: two cards of different suits which have the same denominations, e.g., two Queens or two 10’s etc.
- Two pair: e.g., two 5’s and two 10’s.
- Three of a kind: e.g., three 10’s.
- Straight: five consecutive cards, which are not all of the same suit. An Ace can play either high or low, as in A-2-3-4-5 or 10-J-Q-K-A, but not both—i.e., for example, Q-K-A-2-3 is not a straight.
- Flush: five cards of the same suit which are not all consecutive.
- Full house: three of a kind and one pair.
- Four of a kind.
- Straight flush: five consecutive cards in the same suit.

Suppose we draw five cards at random from a deck of 52 cards.
(a) What is the probability to get four of a kind?

**Solution.** We need to count the number of combinations of five cards containing four of a kind, and then divide by total number of combinations. Note that there are only 13 choices for the four cards of the same denomination, i.e., four Aces, four 2's, etc. For each of the choices, we can pick any one of the \( 52 - 4 = 48 \) remaining cards as the fifth card. So the number of different hands containing four of a kind is \( 13 \cdot 48 \). The total number of combinations of any five cards is:

\[
\binom{52}{5} = \frac{52!}{47!5!} = \frac{48 \cdot 49 \cdot 50 \cdot 51 \cdot 52}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 48 \cdot 49 \cdot 5 \cdot 17 \cdot 13.
\]

Therefore, the probability of getting four of a kind with five cards is:

\[
\frac{13 \cdot 48}{\binom{52}{5}} = \frac{13 \cdot 48}{48 \cdot 49 \cdot 5 \cdot 17 \cdot 13} = \frac{1}{49 \cdot 5 \cdot 17} = \frac{1}{4165} \approx 0.00024.
\]

(b) What is the probability to get a full house?

**Solution.** The number of ways to form three of kind is \( 13 \cdot \binom{4}{3} \). This is because we have 13 different denominations of cards, and for each denomination we can pick any 3 out of the 4 suits. Given that three of a kind are drawn, the number of ways to form a pair using the \( 52-3=49 \) remaining cards is \( 12 \cdot \binom{4}{2} \). This is because we are left with 12 possible denominations of pairs, and for each denomination we can select any two of the four suits. Therefore, the total number of full houses is \( 13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2} = 13 \cdot 4 \cdot 12 \cdot 6 = (13 \cdot 48) \cdot 6 \), which is exactly six times as many as the number of hands containing four of a kind found in the Part (a). The probability of a full house is therefore six times the probability of four of a kind:

\[
\frac{6}{4165} \approx 0.00144.
\]

(c) What is the probability to get three of a kind? (**Hint.** Keep in mind that a hand like 10-10-10-K-K is not three of a kind: it is a full house. Similarly, 10-10-10-10-K is not three of a kind, but four of a kind.)

**Solution.** We first count the total number of hands that contain three of a kind. Some of these, however, contain four of a kind, and others contain a full house. We therefore need to avoid counting the full houses and fours of a kind as part of our tally. In Part (b), we saw that the number of possible ways of choosing three of a kind is \( 13 \cdot \binom{4}{3} \). Given that three of a kind are drawn, we can avoid having four of a kind by selecting any two out of 49-1=48 remaining cards to be the fourth and fifth cards (the one remaining card that we cannot use has the same denomination as our three of a kind). There are \( \binom{48}{2} \) ways to do this. Therefore, the total number of combinations that contain three of a kind but not four of a kind is \( 13 \cdot \binom{4}{3} \cdot \binom{48}{2} \).
Note, however, that some of these are full houses. Subtracting the number of full houses which was found in Part (b), we are left with:

\[ 13 \cdot \binom{4}{3} \cdot \binom{48}{2} - 13 \cdot 48 \cdot 6 = 13 \cdot (4 \cdot 47 \cdot 48/2 - 48 \cdot 6) = 13 \cdot 48 \cdot 88. \]

The probability to get three of a kind is therefore:

\[
\frac{13 \cdot 48 \cdot 88}{\binom{52}{5}} = \frac{13 \cdot 48 \cdot 88}{48 \cdot 49 \cdot 5 \cdot 17 \cdot 13} = \frac{88}{4 \cdot 5 \cdot 17} = \frac{88}{4165} \approx 0.021128.
\]

Note that, as usual, there are many alternative ways to count the number of all hands that contain three of a kind.

Method 2. As previously, there are $13 \cdot \binom{4}{3}$ ways to choose three of a kind. The denominations of the remaining two cards have to be chosen from the 12 remaining denominations, and the two denominations have to be different. The number of ways to do this is therefore $\binom{12}{2}$. Once the denominations of the two remaining cards are chosen there are four ways to choose the suit for each of them, to give a total of $\binom{12}{2} \cdot 4 \cdot 4$ different ways of choosing the two remaining cards. The total $13 \cdot \binom{4}{3} \cdot \binom{12}{2} \cdot 4 \cdot 4 = 13 \cdot 4 \cdot 11 \cdot 12/2 \cdot 4 \cdot 4 = 13 \cdot 48 \cdot 88$ is still the same as above.

Method 3. Choose the two different cards first, by selecting their denominations in $\binom{13}{2}$ different ways, and then selecting their suits in $4 \cdot 4$ ways. Then choose the denomination for three of a kind from the 11 remaining denominations, and choose the three suits in $\binom{4}{3} = 4$ ways, to get a total of $\binom{13}{2} \cdot 4 \cdot 4 \cdot 11 \cdot 4 = 12 \cdot 13/2 \cdot 4 \cdot 4 \cdot 44 = 13 \cdot 48 \cdot 88$.

(d) What is the probability to get a straight flush?

**Solution.** Since a straight can only start from an Ace, 2, 3, . . ., 10, we have that for each suit, there are 10 straight flushes, for a total of $10 \cdot 4 = 40$. Hence the probability to get a straight flush is:

\[
\frac{40}{\binom{52}{5}} = \frac{40}{48 \cdot 49 \cdot 5 \cdot 17 \cdot 13} = \frac{1}{6 \cdot 49 \cdot 17 \cdot 13} \approx 0.000015.
\]

(e) What is the probability to get one pair? (**Hint.** First, note that a hand like 10-10-2-2-3 is not a pair but two pair, 10-10-2-2-2 is a full house, 10-10-2-2-3 is three of a kind, and 10-10-10-2 is four of a kind. Conclude therefore that, on order for a hand to be one pair, the pair and each of the three remaining cards must have four distinct denominations. Count the number of ways to choose these four denominations out of 13 possible denominations; then count the number of ways to choose a pair in one of the denominations and a single card in each of the remaining three.)

**Solution.** We need to count the number of combinations of five cards containing two of a kind, and then divide by total number of combinations. There are $\binom{13}{4}$ ways to choose the four
denominations for the pair and the three remaining cards. Given these four denomination, there are four ways of choosing which one denomination to use for the pair and which three for the three remaining cards. Given any denomination, there are 4 ways of choosing a single card from it, and \( \binom{4}{2} \) ways to choose a pair. Therefore, the total number of hands that contain one pair is:

\[
\left( \binom{13}{4} \right) \cdot 4 \cdot 4^3 \cdot \binom{4}{2} = \frac{13!}{4!9!} \cdot 4^4 \cdot \frac{4!}{2!2!} = \frac{13 \cdot 12 \cdot 11 \cdot 10 \cdot 4^4}{2\cdot2!} = 13 \cdot 12 \cdot 11 \cdot 10 \cdot 4^3.
\]

The total number of combinations of any five cards is:

\[
\binom{52}{5} = \frac{52!}{47!5!} = 48 \cdot 49 \cdot 50 \cdot 51 \cdot 52 \cdot 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 48 \cdot 49 \cdot 5 \cdot 17 \cdot 13.
\]

Therefore, the probability of getting one pair with five cards is:

\[
\frac{11 \cdot 12 \cdot 11 \cdot 10 \cdot 4^3}{48 \cdot 49 \cdot 5 \cdot 17 \cdot 13} = \frac{11 \cdot 10 \cdot 4^2}{49 \cdot 5 \cdot 17} = \frac{11 \cdot 2 \cdot 4^2}{49 \cdot 17} = \frac{352}{833} \approx 0.423.
\]

(f) What is the probability to get two pair?

Solution. The number of different ways to get the three denominations for the two pairs and the remaining card is \( \binom{13}{3} \). Given these three denominations, the number of ways to choose two for the two pairs and one for the remaining card is 3. Given any denomination, there are 4 ways of choosing a single card from it, and \( \binom{4}{2} \) ways to choose a pair. Therefore, the total number of hands that contain two pair is:

\[
\left( \binom{13}{3} \right) \cdot 3 \cdot 4 \cdot \binom{4}{2}^2 = \frac{11 \cdot 12 \cdot 13}{2 \cdot 3} \cdot 3 \cdot 4 \cdot \left( \frac{3 \cdot 4}{2} \right)^2 = \frac{11 \cdot 12^3 \cdot 13 \cdot 3}{6} = 11 \cdot 12^2 \cdot 13 \cdot 6.
\]

Therefore, the probability of getting two pair with five cards is:

\[
\frac{11 \cdot 12^2 \cdot 13 \cdot 6}{48 \cdot 49 \cdot 5 \cdot 17 \cdot 13} = \frac{11 \cdot 12 \cdot 6}{4 \cdot 49 \cdot 5 \cdot 17} = \frac{11 \cdot 18}{49 \cdot 5 \cdot 17} = \frac{198}{4165} \approx 0.0475.
\]

(g) What is the probability to get a straight? (Hint. Keep in mind that, for example, 2-3-4-5-6 of clubs is not a straight but a straight flush.)

Solution. There are \( 4^5 \) sequences of consecutive cards starting with an Ace (i.e., A-2-3-4-5), since there are four possibilities for the Ace, four possibilities for a 2, etc. Four of these, however, are straight flushes, one in each suit. Therefore, there are \( 4^5 - 4 \) straights starting with an Ace. A straight can start with any one of the following ten cards: Ace, 2, 3, ..., 10. Therefore, the total number of straights is \( 10(4^5 - 4) = 10240 - 40 = 10200 \). The probability of a straight is

\[
\frac{10200}{48 \cdot 49 \cdot 5 \cdot 17 \cdot 13} = \frac{3 \cdot 17 \cdot 2 \cdot 4 \cdot 25}{48 \cdot 49 \cdot 5 \cdot 17 \cdot 13} = \frac{5}{2 \cdot 49 \cdot 13} = \frac{5}{1274} \approx 0.0039.
\]

21
(h) What is the probability to get a flush? (Hint. Again, keep in mind that, for example, 2-3-4-5-6 of clubs is not a flush but a straight flush.)

Solution. There are \( \binom{13}{5} \) ways to choose five cards of a certain suit, and, since there are four suits, a total of \( 4 \binom{13}{5} \) ways of choosing five cards of the same suit. As established in Part (g), 40 of these are straight flushes, and therefore the number of flushes is

\[
4 \binom{13}{5} - 40 = 4 \left( \frac{13 \cdot 12 \cdot 11 \cdot 10 \cdot 9}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} - 10 \right) = 4 \cdot 1277.
\]

The probability of a flush is:

\[
\frac{4 \cdot 1277}{48 \cdot 49 \cdot 5 \cdot 17 \cdot 13} = \frac{1277}{12 \cdot 49 \cdot 5 \cdot 17 \cdot 13} = \frac{1277}{649740} \approx 0.00197.
\]

As you can see, the rankings of the various poker hands are directly related to their probabilities: the higher the hand, the less likely it is to appear.

Problem 18. (Deepan Palguna and Ilya Pollak.)

The letters of the word \textsc{probability} are shuffled such that each distinct permutation is equally likely. An \( n \)-letter sub-word is defined as a contiguous sequence of \( n \) letters, read left to right. What is the probability that the seven-letter sub-word \textsc{ability} occurs in such a shuffling?

Solution. (Deepan Palguna, Brendan Claussen, and Ilya Pollak.)

The total number of ways to permute \textsc{probability} distinctly is \( 11!/(2! \times 2!) \). This is because there are 11 letters, and two of them—\textsc{b} and \textsc{i}—repeat twice. The number of permutations containing the sub-word \textsc{ability} is the number of ways to permute five distinct objects: \textsc{p}, \textsc{r}, \textsc{o}, \textsc{b}, and \textsc{ability}. The number of permutations of five distinct objects is \( 5! \). Since all permutations are equally likely, we can divide the number of favorable permutations by the total number of possible permutations to get the required probability as

\[
\frac{4 \times 5!}{11!} = \frac{4}{6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 \cdot 11} = \frac{1}{6 \cdot 7 \cdot 2 \cdot 9 \cdot 10 \cdot 11} = \frac{1}{83160} \approx 0.000012.
\]

Another method (from Brendan Claussen, ECE 302 Spring 2012) is to note that \textsc{ability} will appear in positions 1-7 in the permuted word with probability

\[
\frac{1}{11} \cdot \frac{2}{10} \cdot \frac{2}{9} \cdot \frac{1}{8} \cdot \frac{1}{7} \cdot \frac{1}{6} \cdot \frac{1}{5}.
\]

This product is obtained from the multiplication rule, by considering the sequential selection of letters to form \textsc{ability} in the first seven letters of the permuted word. For the first letter, we have 11 choices, and therefore the probability that \textsc{a} will get selected is 1/11. Once \textsc{a} is selected, we have 10 choices left for the second letter, of which two are \textsc{b}, yielding 2/10 for the conditional probability of selecting \textsc{b} for the second letter given that \textsc{a} was selected for the first letter. Similarly, once \textsc{ab} are selected for the first two letters, there are nine letters left and two of them are \textsc{i}, resulting in conditional probability of 2/9, etc. To get the overall probability of \textsc{ability} appearing in the
permuted word, the above product must be multiplied by five, because \textit{ABILITY} has equal chances to occur in positions 1-7, 2-8, 3-9, 4-10, and 5-11. The answer is therefore
\[
\frac{1}{11} \cdot \frac{2}{10} \cdot \frac{2}{9} \cdot \frac{1}{8} \cdot \frac{1}{7} \cdot \frac{1}{6} \cdot \frac{1}{5} \cdot 5 = \frac{4}{11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6} = \frac{1}{83160} \approx 0.000012,
\]
which is the same as the answer obtained from the first method.

\textbf{Problem 19. (After Feller \cite{feller}, Chapter II, Section 5, page 38.)}
Suppose \( N \) and \( K \) are two positive integers such that \( N > K \). Suppose \( N \) identical balls are distributed among \( K \) different boxes, and let \( n_k \) be the number of balls in the \( k \)-th box, for \( k = 1, \ldots, K \). Let \((n_1, \ldots, n_K)\) be the outcome of this experiment.

(a) Consider the set \( S \) of all outcomes such that each box has at least one ball (in other words, such that \( n_k \geq 1 \) for \( k = 1, \ldots, K \)). How many outcomes are in this set \( S \)? Your answer will be an expression involving \( K \) and \( N \).

(b) Now consider the set \( T \) of all possible outcomes, including all outcomes in which some of the boxes are empty. How many outcomes are in this set \( T \)?

\textbf{Solution.}\ (Cheng Lu and Ilya Pollak.)

(a) A convenient way to think about this problem is to imagine the \( N \) balls aligned in a row:

\[
\begin{array}{ccc}
\circ & \text{(slot)} & \circ & \ldots & \circ & \text{(slot)} & \circ \\
\end{array}
\]

In order to find all the possible outcomes for separating these \( N \) balls into \( K \) different non-empty groups (boxes), we now focus on the slots between each pair of neighboring balls. In the picture above, there are \( N - 1 \) slots, separating the balls into \( N \) distinct groups, one ball in each group. Note that removing one of these slots and keeping the remaining \( N - 2 \) slots will reduce the number of groups to \( N - 1 \). Similarly, removing \( N - K \) slots and keeping the remaining \( K - 1 \) slots will reduce the number of groups to \( K \). Therefore, the question is equivalent to finding the number of possible ways to keep \( K - 1 \) slots among all the \( N - 1 \) slots which leads to the answer
\[
\binom{N - 1}{K - 1}.
\]

(b) The only difference from Part (a) is that now any box can be empty. However, this problem can be reduced to Part (a) by adding one ball to each box and making the total number of balls \( N + K \). The question is therefore equivalent to:

\textit{Suppose \( N \) and \( K \) are two positive integers such that \( N > K \). Suppose \( N + K \) identical balls are distributed among \( K \) different boxes, and let \( n_k \) be the number of balls in the \( k \)-th box, for \( k = 1, \ldots, K \). Let \((n_1, \ldots, n_K)\) be the outcome of this experiment. Consider the set \( T \) of all outcomes such that each box has at least one ball (in other words, such that \( n_k \geq 1 \) for \( k = 1, \ldots, K \)). How many outcomes are in this set \( S \)?}
We therefore need to select \( K - 1 \) slots to keep out of all the \( N + K - 1 \) slots. The answer to this question is the answer from Part (a), with \( N \) replaced by \( N + K \): \( \binom{N + K - 1}{K - 1} \).

**Problem 20.** *(Drake [4], Problem 2.01. Solutions by Ilya Pollak and Bin Ni.)*

The geometric PMF for discrete random variable \( X \) is defined to be

\[
p_X(k) = \mathbb{P}(X = k) = \begin{cases} \frac{C}{1 - P} & \text{if } k = 1, 2, 3, \ldots, \\ 0 & \text{for all other values of } k, \end{cases}
\]

where \( P \) is a positive real number which is less than 1.

(a) Determine the value of \( C \).

**Solution.** Since the probabilities of all \( k \)'s must add up to 1, we have:

\[
\sum_{k=1}^{\infty} \frac{C}{1 - P} k^{-1} = 1
\]

\[
\Rightarrow \quad C \sum_{n=0}^{\infty} (1 - P)^n = C \frac{1}{1 - (1 - P)} = 1
\]

\[
\Rightarrow \quad C = P.
\]

Note that we made a change of variable \( n = k - 1 \), to make use of the following identity:

\[
\sum_{n=0}^{\infty} q^n = \frac{1}{1 - q} \quad (|q| < 1)
\]

(b) Let \( N \) be a positive integer. Determine the probability that an experimental value of \( X \) will be greater than \( N \).

**Solution.** In this part, we just sum the probabilities of the numbers that are greater than \( N \), that is:

\[
\mathbb{P} \{ X > N \} = \sum_{k>N} \mathbb{P}(\{ X = k \})
\]

\[
= \sum_{k>N} p_X(k) = \sum_{k=N+1}^{\infty} P(1 - P)^{k-1}
\]

\[
= P(1 - P)^N \sum_{n=0}^{\infty} (1 - P)^n = P(1 - P)^N \frac{1}{1 - (1 - P)}
\]

\[
= P(1 - P)^N \frac{1}{P} = (1 - P)^N,
\]

where \( n = k - N - 1 \).

(c) Given that an experimental value of random variable \( X \) is greater than integer \( N \), what is the conditional probability that it is also larger than \( 2N \)?
Solution. By the definition of conditional probability, we have:

\[ P(\{X > 2N\} | \{X > N\}) = \frac{P(\{X > 2N\} \cap \{X > N\})}{P(\{X > N\})} = \frac{P(\{X > 2N\})}{P(\{X > N\})} = \left(1 - P\right)^{2N} \left(1 - P\right)^{N} = (1 - P)^N. \]

(d) What is the probability that an experimental value of \(X\) is equal to an integer multiple of 3?

Solution. In this part, we need to sum up all the probabilities of multiples of 3, i.e.

\[
\sum_{k=3,6,9,\ldots} P(1 - P)^{k-1} = \sum_{n=1}^{\infty} P(1 - P)^{3n-1}
\]

\[
= P(1 - P)^2 \sum_{n'=0}^{\infty} (1 - P)^{3n'} = P(1 - P)^2 \sum_{n'=0}^{\infty} [(1 - P)^3]^{n'}
\]

\[
= \frac{P(1 - P)^2}{1 - (1 - P)^3} = \frac{P^2 - 2P + 1}{P^2 - 3P + 3},
\]

where \(n = k/3\) and \(n' = n - 1\).

Problem 21. (After Problem 16 from Chapter 2 of [3].)
Let \(X\) be a random variable with PMF

\[ p_X(x) = \begin{cases} \frac{|x|}{C}, & \text{if } x = -4, -3, -2, -1, 0, 1, 2, 3, 4 \\ 0, & \text{otherwise.} \end{cases} \]

(a) Find \(C\).

Solution. According to the normalization axiom, the probabilities of all the experimental outcomes of a discrete random variable must add to one, i.e.

\[
1 = \sum_{x=-4}^{4} p_X(x) = \frac{1}{C} \left( 4 + 3 + 2 + 1 + 0 + 1 + 2 + 3 + 4 \right) = \frac{20}{C}
\]

\[
\Rightarrow C = 20.
\]

(b) Find \(E[X]\).

Solution. By definition, for a discrete random variable, we have:

\[
E[X] = \sum_{x} x \cdot p_X(x) = \sum_{x=-4}^{4} x \cdot \frac{|x|}{20}
\]

\[
= \frac{1}{20} \left[ -4 \cdot 4 + (-3) \cdot 3 + (-2) \cdot 2 + (-1) \cdot 1 + 0 \cdot 0 + 1 \cdot 1 + 2 \cdot 2 + 3 \cdot 3 + 4 \cdot 4 \right]
\]

\[
= 0.
\]
In general, any random variable \( X \) with an even \( 1 \) PMF must have \( E[X] = 0 \).

(c) Find the PMF of the random variable \( Z = (X - E[X])^2 \).

**Solution.** Since \( E[X] = 0 \), we have \( Z = X^2 \). It follows from definition that:

\[
p_Z(z) = P\{Z = z\} = P\{X^2 = z\}.
\]

This is the probability that \( X^2 \) equals \( z \). Since \( X \) can only be \( \pm 1, \pm 2, \pm 3, \pm 4 \) (with non-zero probability), \( X^2 \) can only be \( 1, 4, 9, \) or \( 16 \). Therefore for \( z \) not equal to these four values, we must have \( p_Z(z) = 0 \), and

\[
p_Z(z) = P\{X^2 = z\} = p_X(-\sqrt{z}) + p_X(\sqrt{z})
\]

\[
= \frac{|-\sqrt{2}|}{20} + \frac{|\sqrt{2}|}{20} = \frac{\sqrt{2}}{10} \quad \text{for } z = 1, 4, 9, 16.
\]

(d) Using Part (c), compute the variance of \( X \).

**Solution.** By definition, the variance of \( X \) is:

\[
\text{var}(X) = E[(X - E[X])^2] = E[Z]
\]

\[
= \sum_{z=1,4,9,16} z \cdot p_Z(z)
\]

\[
= 1 \cdot \frac{\sqrt{1}}{10} + 4 \cdot \frac{\sqrt{4}}{10} + 9 \cdot \frac{\sqrt{9}}{10} + 16 \cdot \frac{\sqrt{16}}{10} = 10.
\]

(e) Compute the variance of \( X \) using the identity \( \text{var}(X) = \sum_x (x - E[X])^2 p_X(x) \).

**Solution.** We only need to sum over \( x \)'s which have non-zero probability:

\[
\text{var}(X) = \sum_{x=\pm 1, \pm 2, \pm 3, \pm 4} (x - E[X])^2 p_X(x)
\]

\[
= (-4)^2 p_X(-4) + (-3)^2 p_X(-3) + (-2)^2 p_X(-2) + (-1)^2 p_X(-1)
\]

\[
+ (1)^2 p_X(1) + (2)^2 p_X(2) + (3)^2 p_X(3) + (4)^2 p_X(4)
\]

\[
= 16 \cdot \frac{4}{20} + 9 \cdot \frac{3}{20} + 4 \cdot \frac{2}{20} + 1 \cdot \frac{1}{20} + 1 \cdot \frac{1}{20} + 4 \cdot \frac{2}{20} + 9 \cdot \frac{3}{20} + 16 \cdot \frac{4}{20}
\]

\[
= \frac{(64 + 27 + 8 + 1) \cdot 2}{20} = 10.
\]

**Problem 22.** (Mosteller [9], Problem 4.)

(a) On the average, how many independent throws of a fair six-sided die must be performed until one gets a 6?

---

1 A function \( f \) is called even if \( f(x) = f(-x) \).
(b) On the average, how many independent tosses of a fair coin must be performed until one gets an H?

Solution. *(Ilya Pollak and Bin Ni.)*

Intuitively, we can expect the answer to Part (a) to be 6. Because the six sides are equally likely, the fraction of 6’s in a long sequence of independent rolls will be approximately 1/6, i.e. we would be getting a 6 every 6-th roll, on average.

To show that this is indeed the correct expected value, let random variable \(N\) be the number of throws required to get a 6 for the first time. Since each throw is a Bernoulli trial with probability of success \(p = 1/6\), \(N\) is a geometric random variable with parameter \(p = 1/6\). As shown in class, the expected value of this random variable is \(1/p = 6\).

Similarly, in Part (b) each throw is a Bernoulli trial with probability of success \(p = 1/2\) and therefore, the number of tosses is geometric with parameter \(p = 1/2\). Therefore, the answer to Part (b) is 2.

Problem 23. Suppose that \(X\) is a Bernoulli random variable, with \(p_X(0) = 1 - p\) and \(p_X(1) = p\). Suppose further that \(E[X] = 3\text{var}(X)\). Find \(p\).

Solution. We have: 

\[
E[X] = E[X^2] = p, \quad \text{and so } \text{var}(X) = E[X^2] - (E[X])^2 = p - p^2.
\]

Since we are given that \(E[X] = 3\text{var}(X)\), we have: \(p = 3p - 3p^2\), and therefore \(3p(2/3 - p) = 0\) which means that either \(p = 0\) (in this case, \(E[X] = 3\var(X) = 0\)) or \(p = 2/3\) (in this case, \(E[X] = 2/3\) and \(\text{var}(X) = 2/9\)).

Problem 24. *(Ross [10], Chapter 4, Example 3d.)*

A class of 120 students are driven in 3 buses to a concert. There are 36 students in one of the buses, 40 in another, and 44 in the third bus. When the buses arrive, one of the 120 students is chosen, according to a discrete uniform probability law. Let \(X\) denote the number of students on the bus of that randomly chosen student. Find \(E[X]\).

Solution. *(Ross [10], Chapter 4, Example 3d.)*

Since the randomly chosen student is equally likely to be any of the 120 students, it follows that \(p_X(36) = 36/120\), \(p_X(40) = 40/120\), \(p_X(44) = 44/120\), and hence

\[
E[X] = 36 \cdot \frac{36}{120} + 40 \cdot \frac{40}{120} + 44 \cdot \frac{44}{120} = \frac{54}{5} + \frac{40}{3} + \frac{242}{15} = \frac{162 + 200 + 242}{15} = \frac{604}{15} \approx 40.2667.
\]

On the other hand, the average number of students on a bus is \(120/3 = 40\), showing that the expected number of students on the bus of a randomly chosen student is larger than the average number of students on a bus. This is a general phenomenon and occurs because the more students there are on a bus, the more likely a randomly chosen student to have been on that bus. As a result, buses with many students are given more weight than those with fewer students.

Problem 25. *(After Ross [10], Chapter 4, Example 6g.)*

In a U.S. presidential election the candidate who gains the maximum number of votes in a state is awarded the total number of electoral college votes allocated to that state. The number of electoral college votes of a given state is roughly proportional to the population of the state. Let us determine the average power in a close presidential election of a citizen in a state with \(n\) voters and \(c\) electoral
votes, where by *average power* of a voter we mean the expected number of electoral votes he/she will affect.

Assume that there are two presidential candidates in the election. Assume that there are \( n = 2k + 1 \) voters in a state: Joe and \( 2k \) other voters. The objective of this problem is to calculate Joe’s average power.

(a) Assuming that each of the other \( 2k \) voters acts independently and is equally likely to vote for either candidate, find the probability that Joe’s vote will be the deciding vote. In other words, find the probability that the \( 2k \) remaining votes are split evenly between the two candidates. (Hint. The number of votes for one of the candidates is a binomial random variable with parameter \( p = 1/2 \).)

**Solution.** Using the hint, we have:

\[
P(\text{Joe's vote is deciding}) = \binom{2k}{k} \left( \frac{1}{2} \right)^k \left( \frac{1}{2} \right)^k = \frac{(2k)!}{(k!)^2 2^{2k}}.
\]

(b) Simplify the expression obtained in Part (a) using Stirling’s approximation: \( m! \approx m^{m+1/2} e^{-m} \sqrt{2\pi} \) (this approximation is valid for any large integer \( m \).)

**Solution.** Using Stirling’s approximation with \( m = 2k \) for the numerator and with \( m = k \) for the denominator, the above expression simplifies to:

\[
P(\text{Joe's vote is deciding}) \approx \frac{(2k)^{2k+1/2} e^{-2k} \sqrt{2\pi}}{k^{2k+1} e^{-2k} (2\pi)^{2k}} = \frac{1}{\sqrt{k} \pi} \approx \frac{1}{\sqrt{n/2}}
\]

(c) If Joe’s vote is the deciding vote in the state, then it will affect \( c \) electoral votes; otherwise, it will not affect any electoral votes. Find Joe’s average power—i.e., the expected number of electoral votes that his vote will affect. You can use \( n/2 \approx k \). Calculate an approximate numerical answer for \( c = 25 \) (the number of Florida’s electoral votes in 2000) and \( n = 6,000,000 \) (approximate number of people who voted in the presidential elections in Florida in 2000).

**Solution.**

\[
\text{average power} = c P(\text{Joe's vote is deciding}) \approx \frac{c}{\sqrt{n/2}},
\]

which is quite significant. For example, in Florida in 2000, \( c \) was 25, \( n \) was approximately six million, and so Joe’s average power would have been approximately 0.008 electoral votes. Conclusion: in a close election, every vote is very important. Another interesting conclusion is that, since \( c \) is roughly proportional to \( n \), the average power grows with \( n \), i.e., voters in large states have more power than do those in smaller states.

**Problem 26.** (Drake [4], Problem 2.02. Solutions by Ilya Pollak and Bin Ni.)

The probability that any particular bulb will burn out during its \( K \)-th month of use is given by the PMF for \( K \),

\[
p_K(k) = \frac{1}{5} \left( \frac{4}{5} \right)^{k-1}, \quad k = 1, 2, 3, \ldots
\]

Four bulbs are life-tested simultaneously. Their lifetimes are independent. Determine the probability that
(a) None of the four bulbs fails during its first month of use.

**Solution.** The probability that a bulb fails in the first month is 1/5. Therefore it has probability 4/5 to live longer than 1 month. All the 4 bulbs work independently, so the probability that all of them live longer than 1 month is:

\[ \left( \frac{4}{5} \right)^4 \approx 0.4096. \]

(b) Exactly two bulbs have failed by the end of the third month.

**Solution.** There are two types of bulbs at the end of the third month: Failed, Working. We use F and W to denote them. All the possible states of the 4 bulbs at that time can be described by 4 letter sequences of F and W such as WWWF (meaning that bulbs 1, 2, 3 are still working and bulb 4 failed). The total number of sequences with 2 F’s is: \( \binom{4}{2} = \frac{4!}{2!2!} = 6 \). Since all the 4 bulbs work independently, the probability of each such sequence is: \( p_3^2(1-p_3)^2 \), where \( p_3 \) is the probability that a bulb fails in the first 3 months, which is determined by:

\[
p_3 = p_K(1) + p_K(2) + p_K(3) = \frac{1}{5} \left( \frac{4}{5} \right)^0 + \frac{1}{5} \left( \frac{4}{5} \right)^1 + \frac{1}{5} \left( \frac{4}{5} \right)^2 = \frac{25}{125} + \frac{20}{125} + \frac{16}{125} = \frac{61}{125}.
\]

The total probability that two bulbs fail in the first three months is:

\[
6p_3^2(1-p_3)^2 = 6 \cdot \left( \frac{61}{125} \right)^2 \left(1 - \frac{61}{125} \right)^2
\]

\[
= 6 \cdot \left( \frac{61 \cdot 64}{125 \cdot 125} \right)^2
\]

\[
\approx 0.37457.
\]

(c) Exactly one bulb fails during each of the first three months.

**Solution.** At the end of the third month, all the bulbs can be classified into 4 types: Failed in month 1, Failed in month 2, Failed in month 3 and Working. If we use symbols 1, 2, 3, W to denote them, all the possible states of the 4 bulbs can be described by four-symbol sequences (such as 113W which means bulbs 1 and 2 failed in the first month, bulb 3 failed in the third month and bulb 4 is still working). Note that for exactly one bulb to fail during each of the first three months, the corresponding sequence would have to have four different symbols. The number of such sequences is the number of permutations of 4 letters which is equal to 4! = 24. The probability of each such sequence is: \( p_K(1)p_K(2)p_K(3)(1-p_3) = \frac{1}{5} \cdot \frac{4}{25} \cdot \frac{16}{125} \cdot \left(1 - \frac{61}{125} \right) \approx 0.0020972 \). Therefore, the overall probability is:

\[
24p_K(1)p_K(2)p_K(3)(1-p_3) \approx 0.050331.
\]
(d) Exactly one bulb has failed by the end of the second month, and exactly two bulbs are still working at the start of the fifth month.

Solution. For this part, we can classify the bulbs into the following 3 types at the beginning of the fifth month: Failed in the first 2 months, Failed in months 3 and 4, Working. Let us use S, F, W, respectively, to denote these types. We want to calculate the probability of all the sequences that contain one S, one F and two W’s. The number of such sequences is the number of ways to partition 4 objects into 2 groups of 1 and 1 group of 2, which equals:

\[
\binom{4}{1,1,2} = \frac{4!}{1!1!2!} = \frac{24}{2} = 12.
\]

The probability of each of the sequences is:

\[
[p_K(1) + p_K(2)][p_K(3) + p_K(4)] \left(1 - \sum_{k=1}^{4} p_K(k)\right)^2
\]

\[
= \left(\frac{1}{5} + \frac{4}{25}\right) \left(\frac{16}{125} + \frac{64}{625}\right) \left(1 - \frac{61}{125} - \frac{64}{625}\right)^2
\]

\[
\approx 0.013916.
\]

The total probability is: \(12 \cdot 0.013916 = 0.16699\).

Problem 27. (Rozanov [11], Chapter 4, Problem 12.)

Balls are drawn from an urn containing \(w\) white balls and \(b\) black balls until a white ball appears. Find the mean value \(m\) and variance \(\sigma^2\) of the number of black balls drawn, assuming each ball is placed back into the urn after being drawn. (In other words, for each attempt we have \(w\) white balls and \(b\) black balls in the urn.) Assume that all draws are independent.

Solution. (Ilya Pollak and Bin Ni.)

Let \(p\) be the probability to get a white ball in each trial. Then \(1 - p\) is the probability to get a black ball. Since all the \(w + b\) balls are equally likely, \(p = \frac{w}{w+b}\). Let \(X\) be the number of black balls drawn until the first white ball appears, and let \(Y = X + 1\) be the total number of balls drawn. Since each trial is a Bernoulli trial with probability of success \(p = \frac{w}{w+b}\), and since the trials are independent, it follows that \(Y\) is a geometric random variable with parameter \(p\). As shown in class and in the text, the mean and variance of \(Y\) are \(1/p\) and \(1/p^2 - 1/p\), respectively. Using linearity of expectation, we
therefore have:

\[
\]

\[
= \frac{1}{p} - 1 = \frac{b + w}{w} - 1
\]

\[
= \frac{b}{w};
\]

\[
\text{var}(X) = E[(X - E[X])^2] = E[(Y - 1 - E[Y] - 1)^2]
\]

\[
\]

\[
= \text{var}(Y) = \frac{1}{p^2} - \frac{1}{p} = \frac{(b + w)^2}{w^2} - \frac{b + w}{w}
\]

\[
= \frac{(b + w)b}{w^2}.
\]

**Problem 28.** *(After Mosteller [9], Problem 7. Solutions by Ilya Pollak and Bin Ni.)*

American roulette wheels have 38 equally likely numbers. If the player’s number comes up, he is paid 35 times his stake and gets his original stake back; otherwise he loses his stake.

Al always bets a dollar on the number 13 at roulette. Bob bets Al $20 that Al will be behind at the end of 36 roulette plays. In other words, if Al is losing to the casino after 36 roulette plays, he must pay $20 to Bob; if Al’s profit against the casino is zero or positive after 36 roulette plays, he gets $20 from Bob. Assume that all roulette plays are independent.

(a) Calculate Al’s expected gain for the 36 roulette plays only, without Bob’s $20 bet. (If Al is expected to lose money, his expected gain will be a negative number.)

**Solution.** Let \(X_i\) to be Al’s gain in the \(i\)-th play. In each play, if he wins (which happens with probability \(1/38\)), \(X_i = 35\), otherwise \(X_i = -1\). The random variable \(X_i\) therefore has the following PMF:

\[
p_{X_i}(x) = \begin{cases} 
1/38 & x = 35 \\
37/38 & x = -1 \\
0 & \text{otherwise}.
\end{cases}
\]

Therefore, the expected value of \(X_i\) is:

\[
E[X_i] = 35 \cdot \frac{1}{38} - 1 \cdot \frac{37}{38} = -\frac{2}{38}.
\]

The gain \(X\) from the first 36 plays is the sum of the gains from each play, \(X = \sum_{i=1}^{36} X_i\). Using the linearity of expectation, we can compute the expected gain \(E[X]\) from 36 plays as follows:

\[
E[X] = E\left[\sum_{i=1}^{36} X_i\right] = \sum_{i=1}^{36} E[X_i]
\]

\[
= 36 \cdot \left(-\frac{2}{38}\right) = -\frac{36}{19} \approx -1.8947.
\]
(b) Calculate the probability that Al is behind at the end of 36 roulette plays. Here and for the remaining parts of this problem, assume that all the roulette plays are independent. (Hint. Can he be behind if he won at least once in the 36 plays?)

**Solution.** If he won once and lost 35 times in the 36 plays, his gain would be: $35 - 1 \cdot 35 = 0$, and he would not be behind. If he won more than once, his gain would therefore be positive. Thus, the only way for Al to be behind at the end of the 36 plays is to lose all of them. Since the plays are independent, the probability to lose all 36 of them is:

$$p_b = \left( \frac{37}{38} \right)^{36} \approx 0.38287.$$

(c) Calculate Al’s expected gain in his bet against Bob.

**Solution.** When he ends up behind, which happens with probability $p_b$, he loses $20, otherwise he gains $20 from Bob. If $K$ is his gain in this bet, the PMF of $K$ is:

$$p_K(k) = \begin{cases} p_b & k = -20 \\ 1 - p_b & k = 20 \\ 0 & \text{otherwise}, \end{cases}$$

and the expectation of $K$ is:

$$E[K] = -20p_b + 20(1 - p_b) \approx -20 \cdot 0.38287 + 20 \cdot (1 - 0.38287) = 4.6852.$$

(d) Using the linearity of expectation and your results from Parts (a) and (c), calculate Al’s overall expected gain from roulette and the bet.

**Solution.** By linearity, we have:

$$E[X + K] = E[X] + E[K] \approx -1.8947 + 4.6852 = 2.7905,$$

and so Al gains 2 dollars and 79 cents per 36 trials; he is finally making money at roulette!

**References**


online.wsj.com/article/SB10001424052970204517204577044570430299472.html