Nick G. Toth
Nov 4, 2019
Math 461
Assignment IV

Exercise 1 (4.4)

Five men and 5 women are ranked according to their scores on an examination. Assume that no two scores are alike and all 10! rankings are equally likely. Let X denote the highest ranking achieved by a woman. (For instance, X = 1 if the top-ranked person is female.) Find P(X =i) for i =1,2,...,10.

Immediately notice that $\mathbb{P}(X = i) = 0$ whenever $i > 6$. We know this because there are five men, so if the five lowest scores are obtained by men, then the lowest score for any female will be ranked 6th among all of the scores. For values of i between 1 and 6, it's given that the denominator of the probability will be 10!. We also know that the probability will have a factor of 5, since there are 5 ways to choose 1 female from 5 females. We can also permute i-1 females of the 5, so we have a factor of $\frac{5!}{(5-(i-1))!}$. And finally we have a factor of (10-i)! for the arrangements of the remaining people. Putting these pieces together, we arrive at the following formula for $\mathbb{P}(X = i)$:

$$\mathbb{P}(X = i) = \begin{cases} \dfrac{5\left(\frac{5!}{(5-(i-1))!}\right)(10-i)!}{10!} & \text{if } 1 \leq i \leq 6 \\ 0 & \text{otherwise} \end{cases}$$

Exercise 2 (4.5)

Let X represent the difference between the number of heads and the number of tails obtained when a coin is tossed n times. What are the possible values of X?

Suppose we toss the coin n times. In doing so, we produce two quantities H and T for the numbers of heads and tails, respectively. Each of these values can be any integer between 0 and n, and their sum must be n. Hence, we can extract the range of X by taking the signed difference between elements of the cartesian square of $\mathbb{W}_{n+1}$. This is illustrated in the following table:

-	0	1	2	3	…	n
0	0	-1	-2	-3	…	-n
1	1	0	-1	-2	…	1-n
2	2	1	0	-1	…	2-n
3	3	2	1	0	…	3-n
⋮	⋮	⋮	⋮	⋮	⋱	⋮
n	n	n-1	n-2	n-3	…	0

It's clear by looking at this table that the range of X is $\{n - 2k : k \in \mathbb{W}_{n+1}\}$.

Problem 3 (4.7)

Suppose that a die is rolled twice. What are the possible values that the following random variables can take on:

(a) X is the maximum value to appear in the two rolls.

Each die could give any integer value between 1 to 6, so we are taking the maximum value between two numbers in the range 1 to 6. We can construct the range by taking the cartesian square of $\{1,2,3,4,5,6\}$, and taking the maximum of the pairs, as in the following able:

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

It's clear by observing the table that the range of X is $\{1,2,3,4,5,6\}$.

(b) X is the minimum value to appear in the two rolls.

In this case, the range of X can be obtained as it was in part a, by producing a table which determines the minimum values of the cartesian square of $\{1,2,3,4,5,6\}$.

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

It's clear by observing the table that the range of X is still $\{1,2,3,4,5,6\}$.

(c) X is the sum of the two rolls.

In this case, the range of X is composed of all numbers which can be constructed as the sum of two numbers, each in the domain $\{1,2,3,4,5,6\}$. We obtain the range of X by summing over pairs in the cartesian square of $\{1,2,3,4,5,6\}$, as illustrated in the following table:

$\times$	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Cells in the red region contain the values which can be obtained as the sum of two die. The blue-highlighted cells show the values which are in the range of X so that duplicates may be ignored. It's clear by observation that the range of X is $\{2,3,4,5,6,7,8,9,10,11,12\}$.

(d) X is the value of the first roll minus the value of the second roll.

In this case, the range of X is obtained by taking any and every element of $\{1,2,3,4,5,6\}$, and subtracting it from every element of $\{1,2,3,4,5,6\}$. We can use the another table to visualize the range of X, as follows:

–	1	2	3	4	5	6
1	0	-1	-2	-3	-4	-5
2	1	0	-1	-2	-3	-4
3	2	1	0	-1	-2	-3
4	3	2	1	0	-1	-2
5	4	3	2	1	0	-1
6	5	4	3	2	1	0

Cells in the red region contain the values which can be obtained by subtracting the value of one die from another. The blue-highlighted cells show the values which are in the range of X so that remaining duplicates may be ignored. It's clear by observation that the range of X is $\{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$.

Problem 4 (4.8)

If the die in Problem 4.7 is assumed fair, calculate the probabilities associated with the random variables in parts (a) through (d).

For each of the following problems, I will use the tables produced in my solution to problem 3 (book: 4.7). My strategy will be to produce explicit functions by taking the ratio of particular values which appear in the table, to the number of elements in the table. The process is trivial and equivalent for all parts of the problem, so I will simply provide the answers below. I will, however, attach my scratch work to the end of the assignment as evidence of work.

(a) X is the maximum value to appear in the two rolls.

$$\mathscr{P}_X(k) = \begin{cases} \frac{2k-1}{36} & \text{if k is in } \{1,2,3,4,5,6\} \\ 0 & \text{otherwise} \end{cases}$$

(b) X is the minimum value to appear in the two rolls.

$$\mathscr{P}_X(k) = \begin{cases} \frac{13-2k}{36} & \text{if k is in } \{1,2,3,4,5,6\} \\ 0 & \text{otherwise} \end{cases}$$

(c) X is the sum of the two rolls.

$$\mathscr{P}_X(k) = \begin{cases} \frac{k-1}{36} & \text{if k is in } \{2,3,4,5,6,7\} \\ \frac{13-k}{36} & \text{if k is in } \{8,9,10,11,12\} \\ 0 & \text{otherwise} \end{cases}$$

(d) X is the value of the first roll minus the value of the second roll.

$$\mathscr{P}_X(k) = \begin{cases} \frac{6-|k|}{36} & \text{if k is in } \{-5,-4,-3,-2,-1,0,1,2,3,4,5\} \\ 0 & \text{otherwise} \end{cases}$$

Exercise 5 (4.11b)

I'm not going to paste the question here. It's too long.

Let k be large, and let N be a random number chosen from the set $\mathbb{N}' = \left\{n \in \mathbb{N} : n \leq 10^k\right\}$. We want to find $\mathbb{P}\big(\mu(N) = 0\big)$. To begin, observe that $\mu(N) = 0$ means that for all prime numbers p, p^2 does not divide N. The following sequence of equivalences will show that $\mathbb{P}\big(\mu(N) \neq 0\big) = \frac{6}{\pi^2}$. Note that I will use $\boldsymbol{P}$ to denote the prime numbers mod $N+1$.

$$\begin{aligned} \mathbb{P}\big(\mu(N) \neq 0\big) &= \prod_{i \in \boldsymbol{P}} \mathbb{P}\big(i^2 \nmid N\big) \\ &= \prod_{i \in \boldsymbol{P}} \Big(1 - \mathbb{P}\big((i^2 \nmid N)^c\big)\Big) \\ &= \prod_{i \in \boldsymbol{P}} \Big(1 - \mathbb{P}\big(i^2 \mid N\big)\Big) \\ &= \prod_{i \in \boldsymbol{P}} \Big(1 - \frac{1}{i^2}\Big) \\ &= \prod_{i \in \boldsymbol{P}} \Big(\frac{i^2 - 1}{i^2}\Big) \\ &= \frac{6}{\pi^2}. \end{aligned}$$

Technically, since $\mathbb{N}'$ is bounded by 10^k, the identity from which obtained $\frac{6}{\pi^2}$ doesn't hold, but we are taking k to infinity, and so N may be arbitrarily large. Thus, the probability converges to $\frac{6}{\pi^2}$. Finally, take the complement to obtain $\mathbb{P}\big(\mu(N) = 0\big) = 1 - \frac{6}{\pi^2} = \frac{\pi^2 - 6}{\pi^2}$.

Exercise 6 (4.20)

I'm not going to paste the question here. It's too long.

(a) Find$P(X > 0)$.

Let W denote a win, and let L denote a loss. We can string these together to represent the results of a series of roulette games. Now we can describe $\mathbb{P}(X > 0)$ as the sum of $\mathbb{P}(W)$ and $\mathbb{P}(LWW)$, since these are the events where are a profit is made.

$$\mathbb{P}(W) = \tfrac{18}{38}$$
$$\mathbb{P}(LWW) = \tfrac{20}{38}\left(\tfrac{18}{38}\right)^2$$

So $\mathbb{P}(X > 0) = \mathbb{P}(WNN) + \mathbb{P}(LWW) = \frac{4059}{6859} \approx 0.59.$

(b) Are you convinced that the strategy is indeed a winning one? Explain.

No. The potential losses are higher than the potential winnings. It's hard to justify this well without the expected value, so let's find it!

(c) Find $\mathbb{E}(X)$.

This method has three potential outcomes. The first is that I win a dollar. The second is that I lose a dollar, and the third is that I lose three dollars. So we need to consider these probabilities. Since the only positive outcome is $\mathbb{P}(X = 1)$, we know that $\mathbb{P}(X = 1) = \mathbb{P}(X > 0)$ from part a. Now we want to determine $\mathbb{P}(X = -1)$ and $\mathbb{P}(X = -3)$.

$$\mathbb{P}(X = -1) = 2\left(\tfrac{20}{38}\right)^2 \tfrac{18}{38} = \tfrac{1800}{6859} \approx 0.26$$
$$\mathbb{P}(X = -3) = \left(\tfrac{20}{38}\right)^3 = \tfrac{1000}{6859} \approx 0.146$$

Now we can compute $\mathbb{E}(X)$:

$$\mathbb{E}(X) = (1)\frac{4059}{6859} + (-1)\frac{1800}{6859} + (-3)\frac{1000}{6859} \approx -0.1.$$

So the expected value of X is an anti-dime.

Problem 7 (4.21)

Four buses carrying 148 students from the same school arrive at a football stadium. The buses carry, respectively, 40, 33, 25, and 50 students. One of the students is randomly selected. Let X denote the number of students who were on the bus carrying the randomly selected student. One of the 4 bus drivers is also randomly selected. Let Y denote the number of students on her bus.

(a) Which of E(X) or E(Y) do you think is larger? Why?

$\mathbb{E}(X)$ is intuitively greater. Suppose we group the four bus drivers in a set B, and group the students in a set S. Now let's choose a random bus driver b from B. Clearly, b is equally likely to be the bus driver of any of the four busses, since there is a 1-1 correspondence between drivers and busses. If we choose a random student s from S, however, it is considerably more likely that s will be on the bus with 50 students than the bus with any other number of students. More generally, the probability that s will be on any particular bus is given by taking the ratio of the number of students on that bus to the sum of all students; $\mathbb{P}(s \text{ on bus } b) = \frac{|b|}{148}$.

(b) Compare E(X) and E(Y).

The expected value of X depends on unbalanced weights.

$$E(X) = \sum_{k=1}^{4} \frac{(\text{students on bus k})^2}{148} = \frac{40^2 + 33^2 + 25^2 + 50^2}{148} = \frac{5814}{148} = \frac{2907}{74}$$

$$E(Y) = \sum_{k=1}^{4} \frac{(\text{students on bus k})}{4} = \frac{40 + 33 + 25 + 50}{4} = 37$$

$\frac{2907}{74} > 37$ is just another way of saying $2907 > 2738$, which is clearly true, so $E(X) > E(Y)$.

Problem 8 (4.24)

I'm not going to paste the question here. It's too long.

(a) In the event that A writes 1, p is the probability that B guesses A's number. Then the expected profit is $p - \frac{3(1-p)}{4}$.

(b) In the event that A writes 2, p becomes the probability that B does not guess A's number. Now the expected profit is $2(1-p) - \frac{3p}{4}$.

(ab and a half)

We first solve for p by relating our solutions from parts a and b. i.e.

$$\begin{aligned} p - \frac{3(1-p)}{4} = 2(1-p) - \frac{3p}{4} &\longleftrightarrow \frac{4p-3+3p}{4} = 2 - \frac{8p+3p}{4} \\ &\longleftrightarrow \frac{7p-3}{4} + \frac{11p}{4} = 2 \\ &\longleftrightarrow 18p - 3 = 8 \\ &\longleftrightarrow p = \frac{11}{18} \end{aligned}$$

Plugging p back into $p - \frac{3(1-p)}{4}$, we obtain $\frac{23}{72}$.

(c) In the event that B guesses 1, q is the probability that A loses a dollar, and $1-q$ is the probability that A wins 75 cents. The expected loss, then, is $q - \frac{3(1-q)}{4}$.

(d) In the event that B guesses 2, $q-1$ is the probability that A loses two dollars, and q is the probability that A wins a dollar and fifty cents. The expected loss is $2(1-q) - \frac{3q}{4}$.

(cd and a half)

Notice that the quantities obtained in parts c and d are the same as those obtained in parts a and b, but with q in place of p. Consequently, q is also $\frac{11}{18}$. Finally, plug $\frac{11}{18}$ into $p - \frac{3(1-p)}{4}$, and obtain the minmax value $\frac{23}{72}$.

Problem 9 (4.26)

One of the numbers 1 through 10 is randomly chosen. You are to try to guess the number chosen by asking with "yes-no" answers. Compute the expected number of questions you will need to ask in each of the following two cases:

(a) Your ith question is to be "Is it i?" i = 1,2,...,10.

Let X denote the number of questions which we need to ask. Then the expected value of X is simply $\sum_{k=1}^{10} \frac{k}{10}$, since $\mathbb{P}(X = i)$ is constantly $\frac{1}{10}$. Now we can factor out the ten and use Gauss' formula to obtain $\mathbb{E}(X) = \frac{55}{10} = \frac{11}{2}$.

(b) With each question you try to eliminate one-half of the remaining numbers, as nearly as possible.

Consider the binary search algorithm, which has logarithmic time complexity. To implement the algorithm in this case, we first ask if the number is less than or equal to the middle value, 5. At that point, our problem is reduced from finding a number in the set {1,2,3,4,5,6,7,8,9,10}, to finding a number in one of the sets {1,2,3,4,5} or {6,7,8,9,10}. The procedure is repeated until we are left with a set containing only one element. We can apply the algorithm to each element to determine the expected number of questions that we must ask.

1: {1,2,3,4,5,6,7,8,9,10} -> {1,2,3,4,5} -> {1,2,3} -> {1,2} -> {1}
2: {1,2,3,4,5,6,7,8,9,10} -> {1,2,3,4,5} -> {1,2,3} -> {1,2} -> {2}
3: {1,2,3,4,5,6,7,8,9,10} -> {1,2,3,4,5} -> {1,2,3} -> {3}
4: {1,2,3,4,5,6,7,8,9,10} -> {1,2,3,4,5} -> {4,5} -> {4}
5: {1,2,3,4,5,6,7,8,9,10} -> {1,2,3,4,5} -> {4,5} -> {5}
6: {1,2,3,4,5,6,7,8,9,10} -> {6,7,8,9,10} -> {6,7,8} -> {6,7} -> {6}
7: {1,2,3,4,5,6,7,8,9,10} -> {6,7,8,9,10} -> {6,7,8} -> {6,7} -> {7}
8: {1,2,3,4,5,6,7,8,9,10} -> {6,7,8,9,10} -> {6,7,8} -> {8}
9: {1,2,3,4,5,6,7,8,9,10} -> {6,7,8,9,10} -> {9,10} -> {9}
10: {1,2,3,4,5,6,7,8,9,10} -> {6,7,8,9,10} -> {9,10} -> {10}

So it takes three steps to find 3,4,5,8,9,10, and four steps to find 1,2,6,7.

With these values, we can finally calculate that the expected number of questions that we must ask is:

$$3\left(\frac{6}{10}\right) + 4\left(\frac{4}{10}\right) = \frac{18}{10} + \frac{16}{10} = \frac{34}{10}.$$

Problem 10 (4.30)

A person tosses a fair coin until a tail appears for the first time. If the tail appears on the nth flip, the person wins 2^n dollars. Let X denote the player's winnings. Show that E (X) = ∞.

Winnings always come in powers of 2, so the range of X is $\left\{2^n : n \in \mathbb{N}\right\}$. Note that we need not include 0, since tails will appear no sooner than on the first flip, so the minimum prize is 2. The probability of getting heads k -1 times in a row is 2^{1-k}, and we also need to consider k^th flip. From these facts, we obtain the following sum for the expected value of X:

$$\mathbb{E}(X) = \sum_{k=1}^{n} 2^k 2^{1-k} \frac{1}{2} = \sum_{k=1}^{n} 2^k 2^{-k} = \sum_{k=1}^{n} 1 = n.$$

So $\mathbb{E}(X) = \lim_{n\to\infty} \mathbb{E}(X) = \lim_{n\to\infty} n = \infty$.

(a) **Would you be willing to pay $1 million to play this game once?**

No. Why on earth would you think I have a million dollars? Also, it is clearly far more likely that I would lose money on this deal. For example, the coin could land on heads immediately, and then I'm down two less than a million dollars. I wouldn't even consider playing if the coin was less than 21 times as likely to land on heads as on it was to land on tails. On the other hand…

(b) **Would you be willing to pay $1million for each game if you could play for as long as you liked and only had to settle up when you stopped playing?**

Are you offering? Can I pay after I'm finished playing?

Problem 11 (4.37)

Find Var(X) and Var(X) for X and Y given in Problem 4.21.

We want to compute $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$, and $\text{Var}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2$.
We can obtain $E(X^2)$ and $E(Y^2)$ by taking the sums from problem 21, and raising the power of each term, as follows:

$$E(X^2) = \sum_{k=1}^{4} \frac{\left(\text{students on bus k}\right)^3}{148} = \frac{40^3 + 33^3 + 25^3 + 50^3}{148} = \frac{120281}{74},$$

$$E(Y^2) = \sum_{k=1}^{4} \frac{\left(\text{students on bus k}\right)^2}{4} = \frac{40^2 + 33^2 + 25^2 + 50^2}{4} = \frac{2907}{2}.$$

And we know from problem 4.21, that $E(X) = \frac{2907}{74}$ and $E(Y) = 37$, so by taking the squares of these values, we obtain the final result:

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{120281}{74} - \left(\frac{2907}{74}\right)^2 = \frac{450145}{5476} \approx 82.2,$$

$$\text{Var}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = \frac{2907}{2} - 37^2 = \frac{169}{2} = 84.5.$$

Problem 12 (4.38)

If E(X)=1 and Var(X)=5, find

(a) E[(2 + X)^2]

Var(X) $= 5 = \mathbb{E}(X^2) - \mathbb{E}(X)^2$, and $\mathbb{E}(X)^2 = 1^2 = 1$, so $\mathbb{E}(X^2) = 5 + \mathbb{E}(X)^2 = 5 + 1 = 6$.

Now we can expand $E[(2+X)^2]$, as follows

$$\begin{aligned}
E[(2+X)^2] &= E[4+4X+X^2] \\
&= E[4] + E[4X] + E[X^2] \\
&= 4 + 4E[X] + E[X^2] \\
&= 4 + 4 + E[X^2] \\
&= 8 + E[X^2].
\end{aligned}$$

And finally, we can plug in the value of $\mathbb{E}(X^2)$ to obtain $E[(2+X)^2] = 8 + 6 = 14$.

(b) Var(4+3X).

We can immediately reduce Var($4 + 3X$) to Var($3X$), since constants aren't known for changing. Then we can factor out the 3, remembering to take the square, to obtain 9Var(X). And finally, Var(X) $= 5$, so Var($4 + 3X$) $= 9(5) = 45$.

SCRATCH WORK FOR PROBLEM 4 (4.8)

1: 1/all
2: 3/all
3: 5/all
4: 7/all
5: 9/all
6: 11/all

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

Derive a function:
(1,1), (2,3), … (6,11) <-> (n,2n-1)
So P(x=i) is (2i-1)/all = (2i-1)/36

1: 11/all
2: 9/all
3: 7/all
4: 5/all
5: 3/all
6: 1/all

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

Derive a function:
(1,11), (2,9), … (6,1) <-> (n,12-(2n-1))
= 12-2n+1 = 13-2n
So P(x=i) is (2i-1)/all = (13-2i)/36

2: 1/all
3: 2/all
4: 3/all
5: 4/all
6: 5/all
7: 6/all
8: 5/all
9: 4/all
10: 3/all
11: 2/all
12: 1/all

×	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Derive a function:
(2,1), (3,2), …, (7,6) <-> (n,n-1)
and (8,5),…,(12,1) <-> (n,13-n)
= 12-2n+1 = 13-2n
So P(x=i) is (i-1)/all if i in 2..7
and (13-i)/all if i in 8..12

-5: 1/all
-4: 2/all
-3: 3/all
-2: 4/all
-1: 5/all
0: 6/all
1: 5/all
2: 4/all
3: 3/all
4: 2/all
5: 1/all

−	1	2	3	4	5	6
1	0	-1	-2	-3	-4	-5
2	1	0	-1	-2	-3	-4
3	2	1	0	-1	-2	-3
4	3	2	1	0	-1	-2
5	4	3	2	1	0	-1
6	5	4	3	2	1	0

Derive a function:
(±5,1), (±4,2), … <-> (n,6-|n|)
So P(x=i) is (6-|i|)/all