Exercises of Chapter 2

Chuang-Chieh Lin*

Department of Computer Science and Information Engineering,
National Chung Cheng University, Ming-Hsiung, Chiayi 621, Taiwan.

Exercise 2.6. *Suppose that we independently roll two standard six-sided dice. Let X_1 be the number that shows on the first die, X_2 the number on the second die, and X the sum of the numbers on the two dice.*

(a) What is $\mathbf{E}[X \mid X_1 \text{ is even}]$?

(b) What is $\mathbf{E}[X \mid X_1 = X_2]$?

(c) What is $\mathbf{E}[X_1 \mid X = 9]$?

(d) What is $\mathbf{E}[X_1 - X_2 \mid X = k]$ for k in the range $[2, 12]$?

Solution.

(a)

$$
\begin{aligned}
& \mathbf{E}[\mathbf{X} \mid \mathbf{X_1} \text{ is even}] \\
= \; & \sum_{i=1}^{12} i \cdot \mathbf{Pr}[X = i \mid X_1 \text{ is even}] \\
= \; & \sum_{i=1}^{12} i \cdot \frac{\mathbf{Pr}[\{X = i\} \cap \{X_1 \in \{2,4,6\}\}]}{\mathbf{Pr}[X_1 \in \{2,4,6\}]} \\
= \; & 1 \cdot 0 + 2 \cdot 0 + 3 \cdot \frac{1/36}{1/2} + 4 \cdot \frac{1/36}{1/2} + 5 \cdot \frac{2/36}{1/2} + 6 \cdot \frac{2/36}{1/2} + 7 \cdot \frac{3/36}{1/2} \\
& + 8 \cdot \frac{3/36}{1/2} + 9 \cdot \frac{2/36}{1/2} + 10 \cdot \frac{2/36}{1/2} + 11 \cdot \frac{1/36}{1/2} + 12 \cdot \frac{1/36}{1/2} \\
= \; & \frac{135}{18} \\
= \; & \frac{15}{2}.
\end{aligned}
$$

*Email address: lincc@cs.ccu.edu.tw

(b)

$$
\begin{aligned}
& \mathbf{Pr}[X \mid X_1 = X_2] \\
= & \sum_{i=1}^{6} (2i) \cdot \mathbf{Pr}[X = 2i \mid X_1 = X_2] \\
= & 2 \cdot \frac{1/36}{6/36} + 4 \cdot \frac{1/36}{6/36} + 6 \cdot \frac{1/36}{6/36} + 8 \cdot \frac{1/36}{6/36} + 10 \cdot \frac{1/36}{6/36} + 12 \cdot \frac{1/36}{6/36} \\
= & 7.
\end{aligned}
$$

(c)

$$
\begin{aligned}
& \mathbf{E}[X_1 \mid X = 9] \\
= & \sum_{i=1}^{6} i \cdot \mathbf{Pr}[X_1 = i \mid X = 9] \\
= & 1 \cdot \frac{0}{4/36} + 2 \cdot \frac{0}{4/36} + 3 \cdot \frac{1/36}{4/36} + 4 \cdot \frac{1/36}{4/36} + 5 \cdot \frac{1/36}{4/36} + 6 \cdot \frac{1/36}{4/36} \\
= & \frac{18}{4} \\
= & \frac{9}{2}.
\end{aligned}
$$

(d) By the linearity of conditional expectation, we have

$$\mathbf{E}[X_1 - X_2 \mid X = k] = \mathbf{E}[X_1 \mid X = k] - \mathbf{E}[X_2 \mid X = k].$$

Note that

$$
\begin{aligned}
& \mathbf{Pr}[X_1 = i \cap X_2 = j] \\
= & \mathbf{Pr}[X_1 = i] \cdot \mathbf{Pr}[X_2 = j] \text{ (since } X_1 \text{ and } X_2 \text{ are independent)} \\
= & \frac{1}{6} \cdot \frac{1}{6} \\
= & \mathbf{Pr}[X_1 = j] \cdot \mathbf{Pr}[X_2 = i] \\
= & \mathbf{Pr}[X_1 = j \cap X_2 = i].
\end{aligned}
$$

So we have

$$
\begin{aligned}
& \mathbf{Pr}[X_1 = i \mid X = k] \\
= \ & \mathbf{Pr}[\{X_1 = i\} \cap \{X_2 = k - i\} \mid X = k] \\
= \ & \frac{\mathbf{Pr}[\{X_1 = i\} \cap \{X_2 = k - i\} \cap \{X = k\}]}{\mathbf{Pr}[X = k]} \\
= \ & \frac{\mathbf{Pr}[\{X_1 = i\} \cap \{X_2 = k - i\}]}{\mathbf{Pr}[X = k]} \\
= \ & \frac{\mathbf{Pr}[\{X_1 = k - i\} \cap \{X_2 = i\}]}{\mathbf{Pr}[X = k]} \\
= \ & \frac{\mathbf{Pr}[\{X_1 = k - i\} \cap \{X_2 = i\} \cap \{X = k\}]}{\mathbf{Pr}[X = k]} \\
= \ & \mathbf{Pr}[X_2 = i \mid X = k].
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& \mathbf{E}[X_1 - X_2 \mid X = k] \\
= \ & \mathbf{E}[X_1 \mid X = k] - \mathbf{E}[X_2 \mid X = k] \\
= \ & \sum_{i=1}^{6} i \cdot \mathbf{Pr}[X_1 = i \mid X = k] - \sum_{i=1}^{6} i \cdot \mathbf{Pr}[X_2 = i \mid X = k] \\
= \ & 0.
\end{aligned}
$$

□

Exercise 2.7. *Let X and Y be independent geometric random variables, where X has parameter p and Y has parameter q.*

(a) What is the probability that $X = Y$?

(b) What is $\mathbf{E}[\max(X, Y)]$?

(c) What is $\mathbf{Pr}[\min(X, Y) = k]$?

(d) What is $\mathbf{E}[X \mid X \leq Y]$?

Solution.

(a) Note that $\{X = Y\} = \bigcup_{i \geq 1}(\{X = i\} \cap \{Y = i\})$, and $(\{X = 1\} \cap \{Y = 1\}), (\{X =$

$2\} \cap \{Y = 2\}), \ldots$ are mutually disjoint. Then we have the following deduction:

$$\begin{aligned}
\mathbf{Pr}[X = Y] &= \mathbf{Pr}\left[\bigcup_{i\geq 1}(\{X = i\} \cap \{Y = i\})\right] \\
&= \sum_{i\geq 1} \mathbf{Pr}[\{X = i\} \cap \{Y = i\}] \\
&= \sum_{i\geq 1} \mathbf{Pr}[\{X = i\}] \cdot \mathbf{Pr}[\{Y = i\}] \text{ (since } X, Y \text{ are independent)} \\
&= \sum_{i\geq 1} (1-p)^{i-1}p \cdot (1-q)^{i-1}q \\
&= pq \sum_{i\geq 1} ((1-p)(1-q))^{i-1} \\
&= pq \cdot \frac{1}{1-(1-p)(1-q)} \text{ (since } (1-p)(1-q) < 1) \\
&= \frac{pq}{p+q-pq}.
\end{aligned}$$

(b)

$$\begin{aligned}
& \mathbf{E}[\max(X, Y)] \\
= & \sum_{x=1}^{\infty}\sum_{y=1}^{x-1} x\mathbf{Pr}[\{X = x\} \cap \{Y = y\}] + \sum_{y=1}^{\infty}\sum_{x=1}^{y-1} y\mathbf{Pr}[\{X = x\} \cap \{Y = y\}] \\
& + \sum_{x\geq 1} x\mathbf{Pr}[\{X = x\} \cap \{Y = x\}] \\
= & \sum_{x=1}^{\infty}\sum_{y=1}^{x-1} x(1-p)^{x-1}p(1-q)^{y-1}q + \sum_{y=1}^{\infty}\sum_{x=1}^{y-1} y(1-p)^{x-1}p(1-q)^{y-1}q \\
& + \sum_{x=1}^{\infty} x(1-p)^{x-1}p(1-q)^{x-1}q \\
= & pq\sum_{x=1}^{\infty} x(1-p)^{x-1}\sum_{y=1}^{x-1}(1-q)^{y-1} + pq\sum_{y=1}^{\infty} y(1-q)^{y-1}\sum_{x=1}^{y-1}(1-p)^{x-1} \\
& + pq\sum_{x=1}^{\infty} x((1-p)(1-q))^{x-1} \\
= & p\sum_{x=1}^{\infty} x(1-p)^{x-1}(1-(1-q)^{x-1}) + q\sum_{y=1}^{\infty} y(1-q)^{y-1}(1-(1-p)^{y-1}) \\
& + \frac{pq}{(p+q-pq)^2} \\
= & A + B + C,
\end{aligned}$$

where

$$
\begin{aligned}
A &= p\left(\sum_{x=1}^{\infty} x(1-p)^{x-1} - \sum_{x=1}^{\infty} x[(1-p)(1-q)]^{x-1}\right) \\
B &= q\left(\sum_{y=1}^{\infty} y(1-q)^{y-1} - \sum_{y=1}^{\infty} y[(1-p)(1-q)]^{y-1}\right) \\
C &= \frac{pq}{(p+q-pq)^2}.
\end{aligned}
$$

$$
\begin{aligned}
A &= \sum_{x=1}^{\infty} x(1-p)^{x-1}p - p \cdot \frac{1}{(p+q-pq)^2} \\
&= \frac{1}{p} - \frac{p}{(p+q-pq)^2}.
\end{aligned}
$$

Similarly we have $B = \frac{1}{q} - \frac{q}{(p+q-pq)^2}$. Thus $\mathbf{E}[\max(X,Y)] = A + B + C = \frac{1}{p} + \frac{1}{q} - \frac{1}{p+q-pq}$.

(c)

$$
\begin{aligned}
& \mathbf{Pr}[\min(X,Y) = k] \\
= \; & \mathbf{Pr}[\{X = k\} \cap \{Y \geq k+1\}] + \mathbf{Pr}[\{Y = k\} \cap \{X \geq k+1\}] \\
& + \mathbf{Pr}[\{Y = k\} \cap \{X = k\}] \\
= \; & p(1-p)^{k-1}(1-q)^k + q(1-q)^{k-1}(1-p)^k + ((1-p)^{k-1}p) \cdot ((1-q)^{k-1}q) \\
= \; & (1-p)^{k-1}(1-q)^{k-1}(p+q-pq).
\end{aligned}
$$

(d) Thanks for Dr. Ton Kloks for giving us the following arguments. First we note that

$$
\begin{aligned}
\mathbf{Pr}[\{X = x\} \cap \{Y \geq x\}] &= \sum_{y \geq x} \mathbf{Pr}[\{X = x\} \cap \{Y = y\}] \text{ (Law of total probablility)} \\
&= \sum_{y \geq x} \mathbf{Pr}[X = x] \cdot \mathbf{Pr}[Y = y] \\
&= \mathbf{Pr}[X = x] \cdot \sum_{y \geq x} \mathbf{Pr}[Y = y] \\
&= \mathbf{Pr}[X = x] \cdot \mathbf{Pr}[Y \geq x].
\end{aligned}
$$

Hence we have

$$\begin{aligned}
\mathbf{E}[X \mid X \le Y] &= \sum_{x=1}^{\infty} x \cdot \mathbf{Pr}[X = x \mid Y \ge x] \\
&= \sum_{x=1}^{\infty} x \cdot \frac{\mathbf{Pr}[\{X = x\} \cap \{Y \ge x\}]}{\mathbf{Pr}[Y \ge x]} \\
&= \sum_{x=1}^{\infty} x \cdot \frac{\mathbf{Pr}[X = x] \cdot \mathbf{Pr}[Y \ge x]}{\mathbf{Pr}[Y \ge x]} \text{ (since } X, Y \text{ are independent)} \\
&= \sum_{x=1}^{\infty} x \cdot \mathbf{Pr}[X = x] \\
&= \frac{1}{p}.
\end{aligned}$$

Actually, we can simply derive $\mathbf{E}[X \mid X \le Y] = \mathbf{E}[X] = 1/p$ since X and Y are independent random variables.

□

Exercise 2.8.

(a) Alice and Bob decide to have children until either they have their first girl or they have $k \ge 1$ children. Assume that each child is a boy or girl independently with probability $1/2$ and that there are no multiple births. What is the expected number of female children that they have? What is the expected number of male children that they have?

(b) Suppose that Alice and Bob simply decide to keep having children until they have their first girl. Assuming that this is possible, what is the expected umber of boys that they have?

Solution. Let X, X_g, X_b be random variables denoting the number of children they have, the number of girls they have, and the number of boys they have respectively (until either they have their first girl or they have $k \ge 1$ children). Hence it is clear that $X = X_g + X_b$.

(a) From the description of the problem, we know that $\mathbf{Pr}[X_g = i] = 0$ for $i \ge 2$. Let G_i be the event that their ith child is a girl, and let B_i be the event that their ith child is a boy. So we know $\{X_g = 1\} = \bigcup_{i=1}^{k}(\{X = i\} \cap G_i)$. Thus we obtain that $\mathbf{E}[X_g] = 0 \cdot \mathbf{Pr}[X_g = 0] + 1 \cdot \mathbf{Pr}[X_g = 1] = 1 \cdot (1 - 2^{-k}) = \underline{\underline{(1 - 2^{-k})}}$.

Before calculate $\mathbf{E}[X_b]$, we calculate $\mathbf{E}[X_b]$ so that $\mathbf{E}[X_b]$ can be obtained by $\mathbf{E}[X_b] = \mathbf{E}[X] - \mathbf{E}[X_g]$. Note that

$$\begin{aligned}
\sum_{i=1}^{k-1} i p^{i-1} &= \frac{d}{dp}\left(\sum_{i=1}^{k-1} p^i\right) \\
&= \frac{d}{dp}\left(\frac{p(1 - p^{k-1})}{1 - p}\right) \\
&= \frac{1 - kp^{k-1} + (k-1)p^k}{(1-p)^2}.
\end{aligned}$$

When $p = 1/2$, from the above equality we have

$$\begin{aligned}
\sum_{i=1}^{k-1} i \left(\frac{1}{2}\right)^{i} &= \frac{1}{2} \cdot \sum_{i=1}^{k-1} i \left(\frac{1}{2}\right)^{i-1} \\
&= \frac{1}{2} \cdot \left(4 - k\left(\frac{1}{2}\right)^{k-3} + (k-1)\left(\frac{1}{2}\right)^{k-2}\right) \\
&= 2 - (k+1)\left(\frac{1}{2}\right)^{k-1}.
\end{aligned}$$

Now it is clear that

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{i=1}^{k} i \cdot \mathbf{Pr}[X = i] \\
&= \sum_{i=1}^{k-1} i \cdot \mathbf{Pr}[\{\text{first } i-1 \text{ are boys and the } i\text{th one is a girl}\}] \\
&\quad + k \cdot \mathbf{Pr}[\{\text{first } k-1 \text{ are boys and the } k\text{th one is a boy or a girl}\}] \\
&= \left(\sum_{i=1}^{k-1} i \cdot \left(\frac{1}{2}\right)^{i-1} \cdot \left(\frac{1}{2}\right)\right) + k \cdot \left(\frac{1}{2}\right)^{k-1} \\
&= 2 - \left(\frac{1}{2}\right)^{k-1}.
\end{aligned}$$

Hence we derive that $\mathbf{E}[X_b] = \mathbf{E}[X] - \mathbf{E}[X_g] = 2 - (1/2)^{k-1} - (1 - 2^{-k}) = \underline{\underline{1 - 2^{-k}}}$.

(b) By the assumption that it is possible for Alice and Bob to have their first girl while keeping having children, we calculate $\mathbf{E}[X_b]$ as follows.

$$\begin{aligned}
\mathbf{E}[X_b] &= \sum_{i=1}^{\infty} i \cdot \mathbf{Pr}[X_b = i] \\
&= \sum_{i=1}^{\infty} i \cdot \mathbf{Pr}[X = i+1] \\
&= \sum_{i=1}^{\infty} i \left(\frac{1}{2}\right)^{i+1} \\
&= \left(\frac{1}{2}\right)^{2} \cdot \sum_{i=0}^{\infty} i \left(\frac{1}{2}\right)^{i-1} \\
&= \left(\frac{1}{2}\right)^{2} \cdot \frac{1}{(1-1/2)^2} \\
&= 1.
\end{aligned}$$

Note that when $x < 1$, we have $0+1+2x+3x^2+\ldots = \frac{d}{dx}(1+x+x^2+x^3+\ldots) = \frac{1}{(1-x)^2}$.

We can also use the result of (a) and take its limit when $k \to \infty$, then we will have $\lim_{k\to\infty}(1 - 2^{-k}) = 1$. □

Exercise 2.18. *The following approach is often called* reservoir sampling. *Suppose we have a sequence of items passing by one at a time. We want to maintain a sample of one item with the probability that it is uniformly distributed over all the items that we have seen at each step. Moreover, we want to accomplish this without knowing the total number of items in advance or storing all of the items that we see.*

Consider the following algorithm, which stores just one item in memory at all times. when the first item appears, it is stored in the memory. When the kth item appears, it replaces the item in memory with probability $1/k$. Explain why this algorithm solves the problem.

Solution. Let p_i be the probability that the ith item is stored in the memory when k items have been seen. For $1 \leq i \leq k$, we have

$$\begin{aligned} p_1 &= 1 \cdot \frac{1}{2} \cdot \frac{2}{3} \cdots \frac{k-1}{k} = \frac{1}{k} \\ p_2 &= \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdots \frac{k-1}{k} = \frac{1}{k} \\ &\vdots \\ p_i &= \frac{1}{i} \cdot \frac{i}{i+1} \cdot \frac{i+1}{i+2} \cdots \frac{k-1}{k} = \frac{1}{k}. \end{aligned}$$

Thus each of the k items we have seen has the same probability of being stored in the memory, i.e., they are uniformly distributed over all the items we have seen. Since each time only one item is stored by the algorithm, the algorithm really solves the problem. □

Exercise 2.25. *A blood test is being performed on n individuals. Each person can be tested separately, but this is expensive. Pooling can decrease the cost. The blood samples of k people can be pooled and analyzed together. If the test is negative, this one test suffices for the group of k individuals. If the test is positive, then each of the k persons must be tested separately and thus $k+1$ total tests are required for the k people.*

Suppose that we create n/k disjoint groups of k people (where k divides n) and use the pooling method. Assume that each person has a positive result on the test independently with probability p.

(a) *What is the probability that the test for a pooled sample of k people will be positive?*

(b) *What is the expected number of tests necessary?*

(c) *Describe how to find the best value of k.*

(d) *Give an inequality that shows for what values of p pooling is better than just testing every individual.*

Solution.

(a) Since a pooled sample has a negative result of testing only when everyone in the pooled sample has a negative result of testing, we obtain the probability that the test for a pooled sample of k people is positive is $1-(1-p)^k$.

(b) Let X_i be a random variable denoting the number of tests for group i, where $i = 1, 2, \ldots, n/k$. Let X be a random variable denoting the total number of test necessary. We can derive that $\mathbf{Pr}[X_i = k+1] = 1-(1-p)^k$ and $\mathbf{Pr}[X_i = 1] = (1-p)^k$. Hence we have

$$\begin{aligned}
\mathbf{E}[X] &= \mathbf{E}\left[\sum_{i=1}^{n/k} X_i\right] \\
&= \sum_{i=1}^{n/k} \mathbf{E}[X_i] \\
&= \sum_{i=1}^{n/k} \left((k+1)[1-(1-p)^k] + 1 \cdot (1-p)^k\right) \\
&= (k+1)(n/k) - n(1-p)^k \\
&= n\left(1 + \frac{1}{k} - (1-p)^k\right).
\end{aligned}$$

(c) Let $g(p, k) = 1 + 1/k - (1-p)^k$. The derivative of g with respect to k is

$$f(p, k) = \frac{d}{dk} g(p, k) = \frac{-1}{k^2} - (1-p)^k \ln(1-p).$$

Our goal is to find the minimum of $g(p, k)$, and this can be done by setting the first derivative of $g(p, k)$ with respect to k equal to 0. Consider the case that p is very small. Recall that the Taylor series (Maclaurin series) for $\ln(1-x)$ is

$$\ln(1-x) = \sum_{i=1}^{\infty} \frac{x^i}{i}, \text{ for } |x| \leq 1, x \neq -1,$$

so $\ln(1-p)$ can be approximated by $-p$. Besides, $(1-p)^k$ is close to 1 since p is small. We can approximate $\ln(1-p)$ and $(1-p)^k$ by $-p$ and 1 respectively, so we can approximate $f(p, k)$ by $-1/k^2 + p$. Let $f(p, k) = 0$ we have the equality $k^2(1-p)^k \ln(1-p) + 1 = 0$. By the previous approximations of some terms, we have $k^2 \cdot 1 \cdot (-p) + 1 = 0$, hence we obtain that $k = 1/\sqrt{p}$. Since $-1/k^2 + p$ is smaller than 0 when $k < 1/\sqrt{p}$ and greater than 0 when $k > 1/\sqrt{p}$, it is clear that we can obtain a minimum value of $\mathbf{E}[X]$ by taking $1/\sqrt{p}$ to be the value of k. As to further discussions, for example, the cases about the value of p, please refer to [1, 2] for more detailed analysis.

(d) Let $h(k, n, p) = n\left(1 + \frac{1}{k} - (1-p)^k\right) - n = n\left(\frac{1}{k} - (1-p)^k\right)$. That is, $h(k, n, p)$ is the difference between the pooling method and just testing every individual. Let $h(k, n, p) < 0$ we have $\frac{1}{k} - (1-p)^k < 0$. Therefore, we derive that $p < 1 - (1/k)^{1/k}$.

□

References

[1] W. Feller: An Introduction to Probability Theory and Its Applications, Vol. I, 3rd edition, John Wiley, New York, 1968.

[2] D. W. Turner, F. E. Tidmore, and D. M. Young: A calculus based approach to the blood testing problem. *SIAM Review* **30** (1988) 119–122.