Probability and Computing, 2nd Edition

Solutions to Chapter 2: Discrete Random Variables and Expectation

Hahndeul Kim

June 2025

2.1

$\mathbf{E}[X] = \left(\sum_{i=1}^{k} i\right)/k = (k+1)/2.$

2.2

The probability to type "proof" is $1/26^5$. As there are $1,000,000 - 5 + 1 = 999,996$ positions to start the word "proof", the desired probability would be $999996/26^5$ by the linearity of expectations.

2.3

Take f as $f(x) = -x^2$ and X as a random variable with $\Pr(X = 1) = \Pr(X = 2) = 1/2$. Then, $-5/2 = \mathbf{E}[f(X)] < f(\mathbf{E}[X]) = -9/4$.
Take f as $f(x) = x$ and X as above. Then, $\mathbf{E}[f(X)] = f(\mathbf{E}[X]) = 3/2$.
Take f as $f(x) = x^2$ and X as above. Then, $9/4 = f(\mathbf{E}[X]) < \mathbf{E}[f(X)] = 5/2$.

2.4

Take $f(x) = x^k$, which is convex when k is an positive even integer. Then by Jensen's inequality, $\mathbf{E}[f(X)] \geq f(\mathbf{E}[X])$ holds.

2.5

Let the event that X is even be Y. Then $\Pr(Y) = \sum_{i=0,2,\cdots} \binom{n}{i}(\frac{1}{2})^n$ holds. As is known, $\sum_{i=0,2,\cdots} \binom{n}{i} = 2^{n-1}$, so $\Pr(Y) = \frac{1}{2}$ is valid.

2.6

(a) X_1 can be 2, 4 or 6. Therefore $\mathbf{E}[X|X_1 \text{ is even}] = (3+4+\cdots+8) \times \frac{1}{18} + (5+6+\cdots+10) \times \frac{1}{18} + (7+8+\cdots+12) \times \frac{1}{18} = \frac{15}{2}$.
(b) $\mathbf{E}[X|X_1 = X_2] = (2+4+6+8+10+12) \times \frac{1}{6} = 7$.
(c) $\mathbf{E}[X_1|X = 9] = (3+4+5+6) \times \frac{1}{4} = \frac{9}{2}$.
(d) $\mathbf{E}[X_1 - X_2|X = k] = 0$, since X_1 and X_2 are independent dice rolls.

2.7

(a) $\sum_{k=1}^{\infty} p(1-p)^{k-1}q(1-q)^{k-1} = pq \cdot \frac{1}{1-(1-p)(1-q)} = \frac{pq}{p+q-pq}$.

(b) $\mathbf{E}[\max(X,Y)] = \sum_{k=1}^{\infty} \Pr(X \geq k \text{ or } Y \geq k) = \sum_{k=1}^{\infty} (1 - \Pr(X < k, Y < k)) =$

$\sum_{k=1}^{\infty}(1-(1-(1-p)^{k-1})(1-(1-q)^{k-1}))$
$=\sum_{k=1}^{\infty}((1-p)^{k-1}+(1-q)^{k-1}-(1-p)^{k-1}(1-q)^{k-1})=\frac{1}{p}+\frac{1}{q}-\frac{1}{p+q-pq}$.
(c) $\Pr(\min(X,Y)=k)=\Pr(X=k)\Pr(Y\geq k)+\Pr(Y=k)\Pr(X\geq k)-\Pr(X=Y=k)=(1-p)^{k-1}(1-q)^{k-1}(p+q-pq)=(1-(p+q-pq))^{k-1}(p+q-pq)$.
(d) $\mathbf{E}[X|X\leq Y]=\mathbf{E}[\min(X,Y)]=1/(p+q-pq)$, since $\min(X,Y)\sim Geom(p+q-pq)$ from the previous exercise.

2.8

(a) Expected number of girls: $\mathbf{E}[G]=1\times\sum_{i=1}^{k}(\frac{1}{2})^i=1-2^{-k}$.

Expected number of boys: $\mathbf{E}[B]=(\frac{1}{2})^k\times k+\sum_{i=1}^{k}(\frac{1}{2})^i\times(i-1)=\frac{2^k-1}{2^k}$.
(b) The number of total children now follows $Geom(1/2)$. Thus, $\mathbf{E}[G+B]=2$ holds. Since $\mathbf{E}[G]=\lim_{k\to\infty}\frac{2^k-1}{2^k}=1$ holds using the result of the previous exercise, $\mathbf{E}[B]=1$.

2.9

(a) $\mathbf{E}[\max(X_1,X_2)]=\sum_{i=1}^{k}\frac{i^2-(i-1)^2}{k^2}\times i=\frac{4k^2+3k-1}{6k}$.

$\mathbf{E}[\min(X_1,X_2)]=\sum_{i=1}^{k}\frac{(k+1-i)^2-(k-i)^2}{k^2}\times i=\frac{2k^2+3k+1}{6k}$.
(b) Since two dice are independent, $\mathbf{E}[X_1]=\mathbf{E}[X_2]=\frac{k+1}{2}$. Therefore, the claim holds.
(c) By the linearity of expectations, $\mathbf{E}[\max(X_1,X_2)]+\mathbf{E}[\min(X_1,X_2)]=\mathbf{E}[\max(X_1,X_2)+\min(X_1,X_2)]$ holds. Since $\{\max(X_1,X_2),\min(X_1,X_2)\}=\{X_1,X_2\}$, $\mathbf{E}[\max(X_1,X_2)+\min(X_1,X_2)]=\mathbf{E}[X_1+X_2]=\mathbf{E}[X_1]+\mathbf{E}[X_2]$ holds, again by the linearity of expectations. Thus, the claim in the previous exercise must be true.

2.10

(a) Base case: when $n=1,2$, it is trivial from the definition of convexity.

Inductive step: Suppose that the claim holds for $n=k$. Now, let $\sum_{i=1}^{k+1}\lambda_i=1$ and $x_1,...,x_{k+1}\in\mathbb{R}$. Then, by the definition of convexity,
$f(\sum_{i=1}^{k+1}\lambda_i x_i)\leq(1-\lambda_{k+1})f(\frac{1}{1-\lambda_{k+1}}(\sum_{i=1}^{k}\lambda_i x_i))+\lambda_{k+1}f(x_{k+1})$ holds. Now, from the inductive hypothesis, $(1-\lambda_{k+1})f(\frac{1}{1-\lambda_{k+1}}(\sum_{i=1}^{k}\lambda_i x_i))\leq(1-\lambda_{k+1})\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}f(x_i)=$

$\sum_{i=1}^{k} \lambda_i f(x_i)$ holds. Therefore, $f(\sum_{i=1}^{k+1} \lambda_i x_i) \leq \sum_{i=1}^{k+1} \lambda_i f(x_i)$. ■

(b) If X takes on only finitely many values, we can denote the set of possible values as $\{x_1, ..., x_n\}$. Then, since $\sum_i \Pr(X = x_i) = 1$, $f(\sum_{i=1}^{n} \Pr(X = x_i)x_i) \leq \sum_{i=1}^{n} \Pr(X = x_i)f(x_i)$ holds from the previous exercise. This is equivalent to $f(\mathbf{E}[X]) \leq \mathbf{E}[f(X)]$.

2.11

Inductive proof.

Base case: It is trivial on $n = 1$.

When $n = 2$, $\mathbf{E}[X_1 + X_2|Y = y] = \sum_i \sum_j (i + j) \Pr(X_1 = i, X_2 = j|Y = y)$

$= \sum_i \sum_j i \Pr(X_1 = i, X_2 = j|Y = y) + \sum_i \sum_j j \Pr(X_1 = i, X_2 = j|Y = y)$.

Now, by the law of total probability, above equation is equivalent to

$\sum_i i \Pr(X_1 = i|Y = y) + \sum_j j \Pr(X_2 = j|Y = y) = \mathbf{E}[X_1|Y = y] + \mathbf{E}[X_2|Y = y]$.

Inductive step: Suppose that the claim holds for $n = k$. Then,

$\mathbf{E}[\sum_{i=1}^{k+1} X_i|Y = y] = \mathbf{E}[X_{k+1}|Y = y] + \mathbf{E}[\sum_{i=1}^{k} X_i|Y = y] = \sum_{i=1}^{k+1} \mathbf{E}[X_i|Y = y]$. ■

2.12

The expected number of cards to draw to see all n cards is equivalent to the coupon collector's problem in the textbook. Let X_i be the number of draws to perform to observe the ith card. Then $X_i \sim \mathit{Geom}(1 - \frac{i-1}{n})$ holds, deriving $\mathbf{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbf{E}[X_i] = \sum_{i=1}^{n} \frac{n}{n-i+1} = \sum_{i=1}^{n} \frac{n}{i}$.

Let Y_i be the indicator variable that is 1 if ith card was not chosen within $2n$ draws. Then the expected number of unchosen cards would be $\sum_{i=1}^{n} \mathbf{E}[Y_i] = n(\frac{n-1}{n})^{2n}$.

Using the same idea, the expected number of cards chosen only once would be $n \times \binom{2n}{1} \frac{1}{n} (\frac{n-1}{n})^{2n-1}$.

2.13

(a) The exercise is equivalent to the coupon collector's problem, since the probability of observing the ith coupon stays as $1 - \frac{2i-2}{2n} = 1 - \frac{i-1}{n}$.

(b) For any positive integer k, the result is equivalent. The probability of observing the ith coupon is $1 - \frac{ki-k}{kn} = 1 - \frac{i-1}{n}$.

2.14

The nth flip must be head. Taking this into account, there would be $\binom{n-1}{k-1}$ ways to assign the ordering of $k-1$ heads and $n-k$ tails.
Therefore, $\Pr(X = n) = \binom{n-1}{k-1}p^k(1-p)^{n-k}$.

2.15

Since it is inefficient to algebraically compute the expectation of a negative binomial distribution, simply introduce $X_1, ..., X_k$ where X_i denotes the number of flips performed after $(i-1)$th head until ith head. Then, $\mathbf{E}[\sum_{i=1}^{k} X_i] = \sum_{i=1}^{k} \mathbf{E}[X_i] = k/p$.

2.16

(a) Take $n = 2^k$, and let X_i be an indicator variable that is 1 if a streak of length $\log_2 n + 1 = k + 1$ occurred starting from the ith flip.
Then $\mathbf{E}[\sum_{i=1}^{n-k} X_i] = \sum_{i=1}^{n-k} \mathbf{E}[X_i] = (n-k)(\frac{1}{2})^k = 1 - \frac{\log_2 n}{n}$ holds.
Now, $1 - \frac{\log_2 n}{n}$ is $1 - o(1)$ since $\lim_{n\to\infty} \frac{\log_2 n}{n} = 0$.
(b) Let $\lfloor \log_2 n - 2\log_2 \log_2 n \rfloor = \delta$. Note that the desired probability is upper-bounded by the probability that all disjoint δ blocks are not a streak, which is $(1 - (\frac{1}{2})^{\delta-1})^{\lfloor n/\delta \rfloor}$.
$(1 - (\frac{1}{2})^{\delta-1})^{\lfloor n/\delta \rfloor} \le (1 - (\frac{1}{2})^{\log_2 n - 2\log_2 \log_2 n})^{\lfloor n/\delta \rfloor} = (1 - \frac{(\log_2 n)^2}{n})^{\lfloor n/\delta \rfloor}$
$\le (1 - \frac{(\log_2 n)^2}{n})^{n/\log_2 n} \le e^{-\log_2 n} = n^{-\log_2 e} \le n^{-1}$ $(1 - x \le e^{-x})$. ■

2.17

$\mathbf{E}[Y_0] = 1$, $\mathbf{E}[Y_1] = 2p$ obviously holds. Now, we have $\mathbf{E}[Y_i | Y_{i-1} = j] = 2pj$ for $i \ge 1$. Then, by the definition of conditional expectation, $\mathbf{E}[Y_i] = \mathbf{E}[\mathbf{E}[Y_i | Y_{i-1}]] = \sum_j 2pj \Pr(Y_{i-1} = j) = 2p\mathbf{E}[Y_{i-1}]$. Thus, $\mathbf{E}[Y_i] = (2p)^i$, and the expected total number of copies $\mathbf{E}[\sum_{i=0}^{\infty} Y_i]$ is bounded if $p < 1/2$.

2.18

Inductive proof.
Base case: It is trivial on $n = 1$.
Inductive step: Suppose that $\Pr(X_k = I_i) = 1/k$ for all i where X_k is the item stored after the kth item (I_k) appeared.

Then, $\Pr(X_{k+1} = I_i) = \Pr(X_k = I_i) \times (1 - \frac{1}{k+1}) = \frac{1}{k+1}$ for all $1 \leq i \leq k$, and obviously $\Pr(X_{k+1} = I_{k+1}) = \frac{1}{k+1}$ which is the probability of replacement. ■

2.19

Let X_k be the item stored after the kth item appeared. Since $k = 1$ is trivial, we will solve for $k \geq 2$. Then $\Pr(X_k = i) = (\frac{1}{2})^{k+1-i}$ for all $2 \leq i \leq k$ and $\Pr(X_k = 1) = \Pr(X_k = 2)$.

2.20

Let X_i be an indicator variable that is 1 if $\pi(i) = i$. Then the expected number of fixed points would be $\mathbf{E}[\sum\limits_{i=1}^{n} X_i] = \sum\limits_{i=1}^{n} \mathbf{E}[X_i] = n \times \frac{1}{n}$.

2.21

$\mathbf{E}[\sum\limits_{i=1}^{n} |a_i - i|] = \sum\limits_{i=1}^{n} \mathbf{E}[|a_i - i|] = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} |j - i| = \sum\limits_{i=1}^{n} \frac{1}{n}(\sum\limits_{j=1}^{i-1} j + \sum\limits_{j=1}^{n-i} j)$
$= \sum\limits_{i=1}^{n} \frac{1}{n}(i^2 - i) = \frac{n^2-1}{3}$.

2.22

In bubble sort, the number of all possible pairs (i, j) that a_i and a_j are inverted is equivalent to the number of inversions that need to be corrected.
Let X be the number of inversions. Then $\mathbf{E}[X] = \sum\limits_{i=1}^{n-1} \sum\limits_{j=i+1}^{n} \Pr(a_i > a_j) = \sum\limits_{i=1}^{n-1} \sum\limits_{j=i+1}^{n} \frac{1}{2}$, since all numbers are distinct and the input is a random permutation.
Thus, $\mathbf{E}[X] = \sum\limits_{i=1}^{n} \frac{1}{2}(n - i) = \frac{n(n-1)}{4}$.

2.23

Let X_i be the number of swaps needed for the ith element. Since the input is a random permutation, $\mathbf{E}[X_i] = (i - 1)/2$.
Thus, the expected number of swaps would be $\mathbf{E}[\sum\limits_{i=1}^{n} X_i] = \sum\limits_{i=1}^{n} \mathbf{E}[X_i] = \frac{n(n-1)}{4}$.

2.24

Let X be the number of dice rolls, and X_1 be the result of the first roll.
Then $\mathbf{E}[X] = \mathbf{E}[X|X_1 = 6]\Pr(X_1 = 6) + \mathbf{E}[X+1]\Pr(X_1 \neq 6)$ holds by the memoryless property.
Thus, $\mathbf{E}[X] = \frac{1}{6}(\frac{1}{6} \times 2 + \frac{5}{6}\mathbf{E}[X+2]) + \frac{5}{6}\mathbf{E}[X+1] = \frac{35}{36}\mathbf{E}[X] + \frac{7}{6}$. $\therefore \mathbf{E}[X] = 42$.

2.25

(a) To make the test negative, all the people in the pool need to be negative, which happens with probability $(1-p)^k$. Thus, the desired probability is $1 - (1-p)^k$.
(b) Since there are n/k pools, the number of expected necessary tests would be $(n/k) \times ((1-(1-p)^k) \times 1 + (1-p)^k \times (k+1)) = n(1 + \frac{1}{k} - (1-p)^k)$.
(c) Compute the derivative of the expectation derived in (b), and numerically solve the gradient being zero.
(d) $n(1 + \frac{1}{k} - (1-p)^k) < n$ must hold for the pooling method to be better than naïve method. The inequality evaluates to $\frac{1}{k} < (1-p)^k$ for a fixed k.

2.26

Let X_i be the number of i-cycles in the graph. Then, the expected number of cycles would be $\mathbf{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbf{E}[X_i]$.
$\mathbf{E}[X_i] = \binom{n}{i}\frac{(k-1)!}{n(n-1)\cdots(n-k+1)} = \frac{n!}{(n-i)!i!}\frac{(i-1)!(n-i)!}{n!} = \frac{1}{i}$ holds. Thus, $\mathbf{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \frac{1}{i} = H(n)$ (*harnomic number*).

2.27

$\mathbf{E}[X] = \sum_{i=1}^{\infty} x\Pr(X = x) = \sum_{i=1}^{\infty}(6/\pi^2)x^{-1} = \infty$, which follows from the well-known divergence of harmonic series.

2.28

If the player won at the kth spin for the first time, the total money lost is $(1+2+\cdots+2^{k-2})$, and earned money is 2^{k-1}. Since $(1+2+\cdots+2^{k-2}) = 2^{k-1}-1$, the player eventually wins a dollar.
$\mathbf{E}[X] = \sum_{i=1}^{\infty}(\frac{1}{2})^i(2^{i-1}-1) = \sum_{i=1}^{\infty}(\frac{1}{2}-(\frac{1}{2})^i) = \infty$. This implies that this strategy is impractical and would lead to bankruptcy, since the player has a finite amount of money.

2.29

Let $S_n = \sum_{j=0}^{n} X_j$. Then from the linearity of expectations for a finite number of random variables, $\mathbf{E}[S_n] = \sum_{j=0}^{n} \mathbf{E}[X_j]$ holds. Here, RHS converges from the given absolute convergence, and thus LHS should also converge. Thus, applying $\lim_{n\to\infty}$ on each side, we get $\mathbf{E}[\sum_{j=0}^{\infty} X_j] = \sum_{j=0}^{\infty} \mathbf{E}[X_j]$.

2.30

Since a player needs to lose all previous $j-1$ bets in order to participate in the jth bet, $\mathbf{E}[X_j] = (1-(\frac{1}{2})^{j-1}) \times 0 + (\frac{1}{2})^j \times 2^{j-1} + (\frac{1}{2})^j \times (-2^{j-1}) = 0$ holds.
$\sum_{j=0}^{\infty} \mathbf{E}[X_j] = 0$ holds, thus the linearity of expectations does not hold here.
This exercise does not fall under the circumstances of exercise 2.29, since $\sum_{j=0}^{\infty} \mathbf{E}[|X_j|] = \infty$ holds.

2.31

The expected winnings would be $\sum_{k=1}^{\infty} (\frac{1}{2})^k \times \frac{2^k}{k} = \sum_{k=1}^{\infty} \frac{1}{k} = \infty$. Thus, the player should be willing to pay any amount of money to play the game.

2.32

(a) By definition, $\Pr(E_i) = 0$ for $i \le m$, and $\Pr(E) = \sum_{i=1}^{n} \Pr(E_i)$ is true.
If $i > m$, then the ith candidate must be the best among all n candidates, and the second-best candidate must be one of the first m candidates. Thus, $\Pr(E_i) = \frac{1}{n} \times \frac{m}{i-1}$.
Therefore, $\Pr(E) = \sum_{i=m+1}^{n} \frac{1}{n} \times \frac{m}{i-1} = \frac{m}{n} \sum_{j=m+1}^{n} \frac{1}{j-1}$.
(b) Since $\sum_{j=m+1}^{n} \frac{1}{j-1} \ge \int_{m+1}^{n+1} \frac{1}{x-1} dx = \ln n - \ln m$, $\Pr(E) \ge \frac{m}{n}(\ln n - \ln m)$ holds.
Also, since $\sum_{j=m+1}^{n} \frac{1}{j-1} \le \int_{m}^{n} \frac{1}{x-1} dx = \ln(n-1) - \ln(m-1)$, $\Pr(E) \ge \frac{m}{n}(\ln(n-1) - \ln(m-1))$ holds.
(c) For a fixed n, $\frac{\partial}{\partial m} \frac{m(\ln n - \ln m)}{n} = \frac{\ln n - \ln m - 1}{n} = 0$ when $m = n/e$. This choice of m is the maximizer, since the given formula has only one local maximum w.r.t. m.
Since $\frac{m(\ln n - \ln m)}{n} = 1/e$ when $m = n/e$, $\Pr(E) \ge 1/e$ holds by (b).