Solutions to Probability and Computing (2nd Edition)

Hahndeul Kim

August 2025

1 Events and Probability

1.1

(a) Choose five coins to be the heads. $\binom{10}{5}/2^{10} = 63/256$.

(b) Choose six or more coins to be the heads. $\sum_{k=6}^{10} \binom{10}{k}/2^{10} = 193/512$.

(c) For each i, the probability that the two flips are the same is $1/2$. Therefore, the desired probability is $1/32$.

(d) $\Pr(4 \text{ consecutive heads}) = 139/2^{10}$, ruling out duplicate cases.

For more than 4 consecutive heads, the counting process is straightforward.

$\Pr(5 \text{ consecutive heads}) = 64/2^{10}$, $\Pr(6 \text{ consecutive heads}) = 28/2^{10}$,

$\Pr(7 \text{ consecutive heads}) = 12/2^{10}$, $\Pr(8 \text{ consecutive heads}) = 5/2^{10}$,

$\Pr(9 \text{ consecutive heads}) = 2/2^{10}$, $\Pr(10 \text{ consecutive heads}) = 1/2^{10}$,

In summary, the desired probability is $251/1024$.

1.2

(a) Choose a number to be same. $6/36 = 1/6$.

(b) By symmetry, $1/2 \times (1 - 1/6) = 5/12$.

(c) $(1 + 3 + 5 + 5 + 3 + 1)/36 = 1/2$.

(d) In addition to two identical rolls, $(1, 4)$ and $(4, 1)$ are the cases. $8/36 = 2/9$.

1.3

(a) The probability that no ace is included is $\binom{48}{2}/\binom{52}{2}$. Thus, the desired probability is $1 - \binom{48}{2}/\binom{52}{2}$.

(b) Similar to (a), the desired probability is $1 - \binom{48}{5}/\binom{52}{5}$.

(c) (Deferred Decision) Match the rank of the first card drawn. $3/51$.

(d) $\binom{13}{5}/\binom{52}{5}$.

(e) Choose a rank to be triplets and choose another rank to be doublets, then build the triplets and doublets. The desired probability is $(13 \times 12 \times \binom{4}{3}\binom{4}{2})/\binom{52}{5}$.

1.4

If the loser has won k games, then the total number of games played would be $n + k$, and the winner must have won the last game. Thus, there are $\binom{n+k-1}{k}$ ways to choose games that the loser won.

Let the random variable X be the number of games won by the loser. Then $\Pr(X = k) = \binom{n+k-1}{k}/2^{n+k-1}$ holds.

1.5

(a) (1,2), (1,4), (1,9), (6,2), (6,4), (6,9), (8,2), (8,4), (8,9)

Alice will win with probability 5/9.

(b) (2,3), (2,5), (2,7), (4,3), (4,5), (4,7), (9,3), (9,5), (9,7)

Alice will win with probability 5/9.
(c) (3,1), (3,6), (3,8), (5,1), (5,6), (5,8), (7,1), (7,6), (7,8)
Alice will win with probability 5/9.

1.6

Let the random variable X_n be the number of white balls in the bin for a fixed n. The claim that $\Pr(X_n = k) = 1/(n-1)$ for all k will be inductively proved.
Base Case: If n=3, $\Pr(X_3 = 1) = \Pr(X_3 = 2) = 1/2$.
Inductive Step: Suppose that $\Pr(X_n = k) = 1/(n-1)$. Then for $n+1$, $\Pr(X_{n+1} = 1) = (1/2) \times (2/3) \times \cdots \times ((n-1)/n) = 1/n$, $\Pr(X_{n+1} = n) = (1/2) \times (2/3) \times \cdots \times ((n-1)/n) = 1/n$ holds. For $2 \le k \le n-1$, $\Pr(X_{n+1} = k) = \Pr(X_n = k) \times ((n-k)/n) + \Pr(X_n = k-1) \times ((k-1)/n) = 1/n$. ■

1.7

(a) Inductive proof.
Base case: $\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2) - \Pr(E_1 \cap E_2)$. (by the axioms of probability)
Inductive Step: Suppose that there are $n+1$ events $E_1, E_2, \cdots, E_{n+1}$, and $\Pr\left(\bigcup_{i=1}^{k} E_i\right) = \sum_{i=1}^{k} \Pr(E_i) - \sum_{i<j} \Pr(E_i \cap E_j) + \cdots + (-1)^{k+1} \Pr\left(\bigcap_{i=1}^{k} E_i\right)$ for all $k \le n$. Now, $\Pr\left(\bigcup_{i=1}^{n+1} E_i\right) = \Pr\left(\left(\bigcup_{i=1}^{n} E_i\right) \bigcup E_{n+1}\right) = \Pr\left(\bigcup_{i=1}^{n} E_i\right) + \Pr(E_{n+1}) - \Pr\left(\bigcup_{i=1}^{n} (E_i \cap E_{n+1})\right)$.

$$\Pr\left(\bigcup_{i=1}^{n+1} E_i\right) = \Pr\left(\bigcup_{i=1}^{n} E_i\right) + \Pr(E_{n+1}) - \Pr\left(\bigcup_{i=1}^{n} (E_i \cap E_{n+1})\right) = \sum_{i=1}^{n+1} \Pr(E_i) - \sum_{i<j\le n} \Pr(E_i \cap E_j) + \cdots + (-1)^{n+1} \Pr\left(\bigcap_{i=1}^{n} E_i\right)$$
$$- \left(\sum_{i=1}^{n} \Pr(E_i \cap E_{n+1}) - \sum_{i<j<n+1} \Pr(E_i \cap E_j \cap E_{n+1}) + \cdots + (-1)^{n+1} \Pr\left(\bigcap_{i=1}^{n+1} E_i\right)\right)$$
$$= \sum_{i=1}^{n+1} \Pr(E_i) - \sum_{i<j\le n+1} \Pr(E_i \cap E_j) + \cdots + (-1)^{n+2} \Pr\left(\bigcap_{i=1}^{n+1} E_i\right). \blacksquare$$

(b), (c) For all $\omega \in \Omega$, if ω is included in $k \ge 1$ of the n given sets, then ω is counted once in LHS and $\sum_{i=1}^{\ell} (-1)^{i+1} \binom{k}{i}$ in RHS. Note that $\binom{k}{i} = 0$ here if $k < i$. Since $\sum_{i=1}^{\ell} (-1)^{i+1} \binom{k}{i} - 1 = \sum_{i=0}^{\ell} (-1)^{i+1} \binom{k}{i} = 0$ if $k \le \ell$, we investigate the sign of $\sum_{i=0}^{\ell} (-1)^{i+1} \binom{k}{i}$ in the case of $k > \ell$. Then it can be easily shown that this is nonnegative when ℓ is odd and nonpositive when ℓ is even, using the properties of binomial coefficients.

1.8

Let D_n be the event that the chosen integer is divisible by n. Then the desired probability is $\Pr(D_4 \cup D_6 \cup D_9) = \Pr(D_4) + \Pr(D_6) + \Pr(D_9) - \Pr(D_4 \cap D_6) - \Pr(D_6 \cap D_9) - \Pr(D_4 \cap D_9) + \Pr(D_4 \cap D_6 \cap D_9) = \Pr(D_4) + \Pr(D_6) + \Pr(D_9) - \Pr(D_{12}) - \Pr(D_{18}) - \Pr(D_{36}) + \Pr(D_{36}) = 388889/1000000$.

1.9

Let S_i be the event that the $\log_2 n + k$ flips starting from the i^{th} flip are consecutive heads. Then $\Pr(S_i) = (1/2)^{\log_2 n + k} = 1/2^k n$. By union bound, the desired probability p is bounded as $p \leq (n - \log_2 n - k + 1)/2^k n \leq 1/2^k$.

1.10

Let A be the event that a fair coin is flipped, and B the event that a biased coin is flipped. Now, let X be the result of the flip. Then the desired probability is $\Pr(B|X = H) = \Pr(B \cap (X = H))/(\Pr(X = H|A)\Pr(A) + \Pr(X = H|B)\Pr(B)) = 0.5/(0.25 + 0.5) = 2/3$.

1.11

(a) The given probability indicates the cases where the bit is flipped even times. Flipping the bit even times is the necessary and sufficient condition to receive the correct bit.
(b) $\frac{1-q_1}{2} \times \frac{1+q_2}{2} + \frac{1+q_1}{2} \times \frac{1-q_2}{2} = \frac{1-q_1 q_2}{2}$.
(c) Inductive proof.
Base case: when $n = 1$, the probability of receiving the correct bit is $1 - p$.
Inductive step: Suppose that the probability of receiving the correct bit after k relays is $\frac{1+(1-2p)^k}{2}$. Then after $k+1$ relays, the probability is $\frac{1+(1-2p)^k}{2} \times (1 - p) + \frac{1-(1-2p)^k}{2} \times p = \frac{1+(1-2p)^{k+1}}{2}$. $\blacksquare$

1.12

Without loss of generality, assume that the contestant initially chooses the first door and Monty opens the second door. Then, $\Pr(C = 1|O = 2)$ and $\Pr(C = 3|O = 2)$ are to be compared.
$\Pr(C = 1|O = 2) = \frac{\Pr(O=2|C=1)\Pr(C=1)}{\Pr(O=2|C=1)\Pr(C=1)+\Pr(O=2|C=3)\Pr(C=3)} = \frac{1}{3}$
$\Pr(C = 3|O = 2) = \frac{\Pr(O=2|C=3)\Pr(C=3)}{\Pr(O=2|C=1)\Pr(C=1)+\Pr(O=2|C=3)\Pr(C=3)} = \frac{2}{3}$
Thus, the contestant should switch curtains.

1.13

Let D be the event that an individual has the disorder, and R be the individual's test result.

$\Pr(D|R = P) = \frac{\Pr(R=P|D)\Pr(D)}{\Pr(R=P|D)\Pr(D)+\Pr(R=N|D)\Pr(D)} = \frac{0.999\times 0.02}{0.999\times 0.02+0.005\times 0.98} \approx 0.803$.

1.14

Let M be the result of the given match, E_1 the event that I am better, E_2 the event that both are equal and E_3 the event that the opponent is better.
The desired probability is $\Pr(E_3|M) = \frac{\Pr(M|E_3)\Pr(E_3)}{\Pr(M|E_1)\Pr(E_1)+\Pr(M|E_2)\Pr(E_2)+\Pr(M|E_3)\Pr(E_3)} = \frac{0.4\times 0.6^3}{0.4^3\times 0.6+0.5^4+0.4\times 0.6^3} \approx 0.461$.

1.15

By the principle of deferred decisions, consider the situation where the last roll is left. Then, there is always a unique result of the last roll to make the final sum divisible by 6. Thus, the desired probability is 1/6.

1.16

(a) The desired probability is $6/6^3 = 1/36$.
(b) The desired probability is $(6 \cdot 5 \cdot 3)/6^3 = 5/12$.
(c) Under the given condition, the player will lose if he/she fails twice to roll the other die to match the other two dice. The desired probability is $1 - (5/6)^2 = 11/36$.
(d) The desired probability is $\frac{1}{36} + \frac{5}{12} \times \frac{11}{36} + \frac{5}{9} \times (\frac{1}{36} + \frac{5}{12} \times \frac{1}{6} + \frac{5}{9} \times \frac{1}{36}) = \frac{197}{972}$.

1.17

If the vector $\overline{r}$ is chosen uniformly from $\{0, 1, \cdots, k-1\}^n$, then
$\Pr(\mathbf{AB}\overline{r} = \mathbf{C}\overline{r}) \leq 1/k$.
Thus, if the identity test is run p times, then the true positive probability is $1 - \frac{1}{k^{p+1}+1}$.

1.18

First, uniformly choose an integer x from $\{0, \cdots, n-1\}$. Then evaluate $F(z)$ as $F(z) = F((z-x)+x) = (F(z-x)+F(x)) \mod m$. At the lookup table, each value $F(z-x)$ and $F(x)$ will be of the correct value with probability 4/5. Considering that we take modulo m, the output will be equal to $F(z)$ with probability at least $(4/5)^2 = 16/25$.
If the algorithm is repeated three times, then it is possible to take the majority if it exists, or otherwise simply take the first output. For the algorithm to fail, at least two of the three outputs need to be wrong, and that occurs with probability less than $\binom{3}{2} \times (\frac{9}{25})^2 \times \frac{16}{25} + \binom{3}{3} \times (\frac{9}{25})^3 \approx 0.295$. Thus, the final output is correct with probability at least 0.705.

1.19

Let A be the event that the sum of two dice rolls is 2, and let B be the event that the first roll is 2. Then $0 = \Pr(A|B) < \Pr(A) = 1/36$.
Let A be the event that a roll of a blue die is 1, and let B be the event that a roll of a red die is 1. Then $\Pr(A|B) = \Pr(A) = 1/36$.
Let A be the event that the sum of two dice rolls is 2, and let B be the event that the first roll is 1. Then $1/6 = \Pr(A|B) > \Pr(A) = 1/36$.

1.20

The goal is to show that $\Pr\left(\bigcap_{i\in I} \overline{E_i}\right) = \prod_{i\in I} \Pr(\overline{E_i})$.

$\Pr\left(\bigcap_{i\in I} \overline{E_i}\right) = \Pr\left(\overline{\bigcup_{i\in I} E_i}\right) = 1 - \Pr\left(\bigcup_{i\in I} E_i\right)$

$= 1 - \left(\sum_{i\in I} \Pr(E_i) - \sum_{i<j\in I} \Pr(E_i \cap E_j) + \cdots + (-1)^{|I|+1} \Pr\left(\bigcap_{i\in I} E_i\right)\right)$

$= 1 - \sum_{i\in I} \Pr(E_i) + \sum_{i<j\in I} \Pr(E_i)\Pr(E_j) + \cdots + (-1)^{|I|} \prod_{i\in I} \Pr(E_i)$

$= \prod_{i\in I}(1 - \Pr(E_i)) = \prod_{i\in I} \Pr(\overline{E_i})$. ■

1.21

Suppose that a fair four-sided die is rolled. Let $X = \{1, 2\}$, $y = \{1, 3\}$,$Z = \{1, 4\}$. Then the events are pairwise independent but not mutually independent $(1/4 = \Pr(X \cap Y \cap Z) \neq \Pr(X)\Pr(Y)\Pr(Z) = 1/8$.

1.22

(a) For any $X \subset \{1, \cdots, n\}$, each element in X is chosen with probability $1/2$, and each element not in X is dropped with probability $1/2$. Thus, the probability of X to be generated is $1/2^{|X|} \times 1/2^{n-|X|} = 1/2^n$. As there are 2^n possible subsets, X is equally likely to be any one of the possible subsets.
(b) For $X \subseteq Y$, each element in X must be in Y, which happens with probability $3/4$. Thus, $\Pr(X \subseteq Y) = (3/4)^n$.
For $X \cup Y = \{1, \cdots, n\}$, each element in $\{1, \cdots, n\}$ must be in at least one of two sets X and Y, which happens with probability $3/4$. Thus, $\Pr(X \cup Y = \{1, \cdots, n\}) = (3/4)^n$.

1.23

After executing the randomized min-cut algorithm, there can be at most $\binom{n}{2}$ edges between the two vertices in the reduced graph, as n vertices are assumed. As each edge is a candidate for min-cut sets, there can be at most $\binom{n}{2} = \frac{n(n-1)}{2}$ distinct min-cut sets.

1.24

1.25

2 Discrete Random Variables and Expectation

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

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 $(1 - (\frac{1}{2})^{\delta - 1})^{\lfloor n/\delta \rfloor}$.
$(1 - (\frac{1}{2})^{\delta-1})^{\lfloor n/\delta \rfloor} \leq (1 - (\frac{1}{2})^{\log_2 n - 2\log_2 \log_2 n})^{\lfloor n/\delta \rfloor} = (1 - \frac{(\log_2 n)^2}{n})^{\lfloor n/\delta \rfloor}$
$\leq (1 - \frac{(\log_2 n)^2}{n})^{n/\log_2 n} \leq e^{-\log_2 n} = n^{-\log_2 e} \leq n^{-1}$ $(1 - x \leq e^{-x})$. ■

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 $i \geq 1$. 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 $\mathbf{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbf{E}[X_i] = n \times \frac{1}{n}$.

2.21

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

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 $\mathbf{E}[X] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Pr(a_i > a_j) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{1}{2}$, since all numbers are distinct and the input is a random permutation.
Thus, $\mathbf{E}[X] = \sum_{i=1}^{n} \frac{1}{2}(n - i) = \frac{n(n-1)}{4}$.

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 $\mathbf{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbf{E}[X_i] = \frac{n(n-1)}{4}$.

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).

3 Moments and Deviations

3.1

$\mathbf{E}[X^2] = \sum_{i=1}^{n} \frac{1}{n} \times i^2 = \frac{(n+1)(2n+1)}{6}$, and $\mathbf{E}[X] = \frac{n+1}{2}$.
Thus, $\mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{n^2-1}{12}$.

3.2

$\mathbf{E}[X] = 0$, and $\mathbf{E}[X^2] = \sum_{i=1}^{k} \frac{2}{2k+1} \times i^2 = \frac{k(k+1)}{3}$.
Thus, $\mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{k(k+1)}{3}$.

3.3

The variance of a single die roll is $\frac{35}{12}$ from problem 3.1. Since all rolls are independent, $\Pr(|X - 350| \geq 50) \leq \frac{1}{50^2} \times \frac{35}{12} \times 100 = \frac{7}{60}$.

3.4

$\mathbf{Var}[cX] = \mathbf{E}[(cX - \mathbf{E}[cX])^2] = \mathbf{E}[c^2X^2 - 2cX\mathbf{E}[cX] + (\mathbf{E}[cX])^2]$
$= c^2(\mathbf{E}[X^2] - (\mathbf{E}[X])^2) = c^2\mathbf{Var}[X]$. ■

3.5

$\mathbf{Var}[X - Y] = \mathbf{E}[((X - Y) - \mathbf{E}[X - Y])^2] = \mathbf{E}[((X - \mathbf{E}[X]) - (Y - \mathbf{E}[Y]))^2]$
$= \mathbf{E}[(X - \mathbf{E}[X])^2] - 2\mathbf{E}[(X - \mathbf{E}[X])(Y - \mathbf{E}[Y])] + \mathbf{E}[(Y - \mathbf{E}[Y])^2]$
$= \mathbf{Var}[X] - \mathbf{Cov}[X, Y] + \mathbf{Var}[Y] = \mathbf{Var}[X] + \mathbf{Var}[Y]$ $(X \perp\!\!\!\perp Y)$. ■

3.6

Let X_i $(1 \leq i \leq k)$ be the number of flips after $(i-1)$th head until ith head. Since all flips are independent, the desired variance could be computed as $\sum_{i=1}^{k} \mathbf{Var}[X_i]$. As $X_i \sim Geom(p)$, $\mathbf{Var}[X_i] = (1-p)/p^2$ for all i. Thus, the desired variance is $k(1-p)/p^2$.

3.7

Let X be the number of increases. Then $\Pr(X = k) = \binom{d}{k}p^k(1-p)^{d-k}$.
Let the price of the stock after d days be V.
Then $\mathbf{E}[V] = \sum_{k=0}^{d} qr^k(\frac{1}{r})^{d-k}\binom{d}{k}p^k(1-p)^{d-k} = \sum_{k=0}^{d} q\binom{d}{k}(pr)^k(\frac{1-p}{r})^{n-k}$.
Let $M = pr + (1-p)/r = (1 - p + pr^2)/r$. Then
$\mathbf{E}[V] = M^d \sum_{k=0}^{d} q\binom{d}{k}(\frac{pr}{M})^k(\frac{1-p}{rM})^{d-k} = M^d \sum_{k=0}^{d} q\binom{d}{k}(\frac{pr^2}{rM})^k(\frac{1-p}{rM})^{d-k} = M^dq$.

Now we compute $\mathbf{E}[V^2] = \sum_{k=0}^{d} q^2 r^{2k} (\frac{1}{r})^{2d-2k} \binom{d}{k} p^k (1-p)^{d-k}$.
$\mathbf{E}[V^2] = q^2 \sum_{k=0}^{d} \binom{d}{k} (pr^2)^k (\frac{1-p}{r^2})^{d-k} = q^2 \left(pr^2 + \frac{1-p}{r^2}\right)^d$ (similar to $\mathbf{E}[V]$).
Thus, $\mathbf{Var}[V] = q^2 \left((pr^2 + \frac{1-p}{r^2})^d - (pr + \frac{1-p}{r})^{2d}\right)$.
By plugging $q = 1$ in, we get the desired result.

3.8

Let X be the running time of the given algorithm on input strings of size n. Now, let M be the longest running time of the algorithm among the input strings of size n. Then $\Pr(X \geq M) \geq 1/2^n$ by definition.
By Markov's inequality, $1/2^n \leq \Pr(X \geq M) \leq \frac{\mathbf{E}[X]}{M}$, which leads to $M \leq 2^n \mathbf{E}[X]$. Since $\mathbf{E}[X] = O(n^2)$, we get $M = O(n^2 2^n)$.

3.9

(a) By linearity of expectations, $\mathbf{E}[X^2] = \mathbf{E}[\sum_{i=1}^{n} X_i X] = \sum_{i=1}^{n} \mathbf{E}[X_i X]$.
Since X_i are Bernoulli random variables, $\mathbf{E}[X_i X] = \Pr(X_i = 0) \times 0 + \Pr(X_i = 1) \times \mathbf{E}[X | X_i = 1]$. $\blacksquare$
(b) Using the equation proven in (a), $\mathbf{E}[X^2] = \sum_{i=1}^{n} p \times (1 + (n-1)p) = np + n(n-1)p^2$. Thus, $\mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = np + n(n-1)p^2 - n^2p^2 = np(1-p)$.

3.10

Let $X \sim \mathit{Geom}(p)$, and let $Y = 1$ if and only if $X = 1$ and $Y = 0$ otherwise. Then, by Lemma 2.5, $\mathbf{E}[X^3] = \Pr(Y = 1)\mathbf{E}[X^3 | Y = 1] + \Pr(Y = 0)\mathbf{E}[X^3 | Y = 0]$ $= p \times 1 + (1-p) \times \mathbf{E}[X^3 | Y = 0] = p + (1-p) \times \mathbf{E}[X^3 | X > 1]$.
Now, by the memoryless property of geometric distributions, $\mathbf{E}[X^3 | X > 1] = \mathbf{E}[(X+1)^3]$. Thus, $\mathbf{E}[X^3] = p + (1-p)(\mathbf{E}[X^3] + 3\mathbf{E}[X^2] + 3\mathbf{E}[X] + 1)$.
This leads to $\mathbf{E}[X^3] = (p^2 - 6p + 6)/p^3$.
Similarly, we can find $\mathbf{E}[X^4] = (-p^3 + 14p^2 - 36p + 24)/p^4$.

3.11

Let $X = \sum_{i<j} X_{i,j}$, where $X_{i,j}$ is an indicator variable that is 1 if a_i and a_j are inverted. Then, we compute $\mathbf{E}[X^2]$ as:
$\mathbf{E}[X^2] = \mathbf{E}[\sum_{i<j} X_{i,j}^2 + \sum_{|\{i,j,k,l\}|=4} X_{i,j}X_{k,l} + \sum_{i<j<k} 2(X_{i,j}X_{j,k} + X_{i,j}X_{i,k} + X_{i,k}X_{j,k})]$
$= \binom{n}{2} \cdot \frac{1}{2} + \binom{n}{4} \cdot \binom{4}{2} \cdot \frac{1}{4} + 2 \cdot \binom{n}{3} \cdot \left(\frac{1}{6} + \frac{1}{3} + \frac{1}{3}\right) = \frac{n(n-1)(9n^2-5n+10)}{144}$.
Since $\mathbf{E}[X] = \frac{n(n-1)}{4}$, $\mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{n(n-1)(2n+5)}{72}$.

3.12

Let X be a random variable with probability distribution like $\Pr(X = n) = \frac{1}{\zeta(3)n^3}$.
Then $\mathbf{E}[X] = \sum\limits_{n=1}^{\infty} n \times \frac{1}{\zeta(3)n^3} = \frac{1}{\zeta(3)} \times \frac{\pi^2}{6} < \infty$.
However, since $\mathbf{E}[X^2] = \sum\limits_{n=1}^{\infty} n^2 \times \frac{1}{\zeta(3)n^3} = \sum\limits_{n=1}^{\infty} \frac{1}{n} \times \frac{1}{\zeta(3)} = \infty$ (harmonic series), the variance of X is unbounded.

3.13

Let X be a random variable such that $\Pr(X = n) = \frac{1}{\zeta(k+2)n^{k+2}}$.
Then, similar to problem 3.12, $\mathbf{E}[X^k]$ converges and $\mathbf{E}[X^{k+1}]$ diverges.

3.14

$\mathbf{Var}[\sum\limits_{i=1}^{n} X_i] = \mathbf{E}\left[\left(\sum\limits_{i=1}^{n}(X_i - \mathbf{E}[X_i])\right)^2\right] = \sum\limits_{i=1}^{n} \mathbf{Var}[X_i] + \sum\limits_{i \neq j} \mathbf{Cov}[X_i, X_j] = \sum\limits_{i=1}^{n} \mathbf{Var}[X_i] + 2\sum\limits_{i=1}^{n}\sum\limits_{i<j} \mathbf{Cov}[X_i, X_j]$. ■

3.15

$\mathbf{E}[X_iX_j] = \mathbf{E}[X_i]\mathbf{E}[X_j]$ indicates that $\mathbf{Cov}[X_i, X_j] = \mathbf{E}[X_iX_j] - \mathbf{E}[X_i]\mathbf{E}[X_j] = 0$.
Thus, $\mathbf{Var}[X] = \sum\limits_{i=1}^{n} \mathbf{Var}[X_i]$ holds.

3.16

Suppose that we want the expectation to be μ. Then the desired X should satisfy $\Pr(X = k\mu) = 1/k$ and $\Pr(X = 0) = 1 - 1/k$.

3.17

Suppose that we want the expectation to be μ. Then in order to satisfy $\Pr(|X - \mathbf{E}[X]| \geq a) = \frac{\mathbf{Var}[X]}{a^2}$, X should satisfy:
$\Pr(X = \mu) = p$, $\Pr(X = \mu + a) = (1-p)/2$ and $\Pr(X = \mu - a) = (1-p)/2$.

3.18

(a) $\Pr(X - \mathbf{E}[X] \geq t\sigma[X]) = \Pr[t(X - \mathbf{E}[X]) + \sigma[X] \geq (t^2+1)\sigma[X]]$
$\leq \Pr[(t(X - \mathbf{E}[X]) + \sigma[X])^2 \geq (t^2+1)^2\mathbf{Var}[X]]$
$\leq \mathbf{E}[(t(X - \mathbf{E}[X]) + \sigma[X])^2]/(t^2+1)^2\mathbf{Var}[X]$ (Markov's inequality)
$=(t^2\mathbf{Var}[X] + \mathbf{Var}[X])/(t^2+1)^2\mathbf{Var}[X] = 1/(t^2+1)$. ■

(b) Since probabilities cannot be greater than 1, we only consider $t \geq 1$.
By Chebyshev's inequality, $\Pr(|X - \mathbf{E}[X]| \geq t\sigma[X]) \leq 1/t^2 \leq 2/(1+t^2)$. ■

3.19

(i) If $\mu = m$, then the claim is trivially valid.
(ii) If $\mu < m$, then let $t = |\mu - m|/\sigma$. Then using the result of exercise 3.18(a), $\Pr(X - \mu \geq |\mu - m|) \leq \frac{\sigma^2}{(\mu-m)^2+\sigma^2}$.
Now, from $1/2 \leq \Pr(X \geq m) \leq \Pr(X - \mu \geq |\mu - m|) \leq \frac{\sigma^2}{(\mu-m)^2+\sigma^2}$, we conclude $(\mu - m)^2 \leq \sigma^2$, thus $|\mu - m| \leq \sigma$.
(iii) If $\mu > m$, then let $t = |\mu - m|/\sigma$. Now, substituting X into $-X$ in the result of exercise 3.18(a), we get $\Pr(-X + \mathbf{E}[X] \geq t\sigma[X]) \leq 1/(t^2+1)$.
This leads to $\Pr(-X + \mu \geq |\mu - m|) \leq \frac{\sigma^2}{(\mu-m)^2+\sigma^2}$.
Similarly, from $1/2 \leq \Pr(X \leq m) \leq \Pr(-X + \mu \geq |\mu - m|) \leq \frac{\sigma^2}{(\mu-m)^2+\sigma^2}$, we conclude $(\mu - m)^2 \leq \sigma^2$, thus $|\mu - m| \leq \sigma$. ■

3.20

We first prove $\Pr(Y \neq 0) \leq \mathbf{E}[Y]$. This is trivial since
$\mathbf{E}[Y] - \Pr(Y \neq 0) = \sum_{i=1}^{\infty} i \Pr(Y = i) - \sum_{i=1}^{\infty} \Pr(Y = i) = \sum_{i=1}^{\infty} (i-1) \Pr(Y = i) \geq 0$.
Now we prove $\frac{\mathbf{E}[Y]^2}{\mathbf{E}[Y^2]} \leq \Pr(Y \neq 0)$. Let $X = Y|Y \neq 0$ such that $\Pr(X = x) = \Pr(Y = x|Y \neq 0)$. Since $(\mathbf{E}[X])^2 \leq \mathbf{E}[X^2]$, $\mathbf{E}[Y|Y \neq 0]^2 \leq \mathbf{E}[Y^2|Y \neq 0]$ holds.
Now, consider that $\mathbf{E}[X] = \mathbf{E}[0]\Pr(X = 0) + \mathbf{E}[X|X \neq 0]\Pr(X \neq 0) = \mathbf{E}[X|X \neq 0]\Pr(X \neq 0)$ is valid for any random variable X.
Combined with $\mathbf{E}[Y|Y \neq 0]^2 \leq \mathbf{E}[Y^2|Y \neq 0]$, $\left(\frac{\mathbf{E}[Y]}{\Pr(Y \neq 0)}\right)^2 \leq \frac{\mathbf{E}[Y^2]}{\Pr(Y \neq 0)}$ holds. ■

3.21

(a) Let $Y = |X - \mathbf{E}[X]|$. Then by Markov's inequality,
$\Pr(Y > t\sqrt[k]{\mathbf{E}[Y^k]}) = \Pr(Y^k > t^k\mathbf{E}[Y^k]) \leq \Pr(Y^k \geq t^k\mathbf{E}[Y^k]) \leq 1/t^k$.
(b) If k is odd, then $Y^k \neq (X - \mathbf{E}[X])^k$ and $(X - \mathbf{E}[X])^k$ may not always be positive. Therefore, we cannot apply Markov's inequality in this case.

3.22

Let X_i be indicator variables that are 1 if $\pi(i) = i$. Then $X_i \sim \mathit{Bernoulli}(1/n)$.
Then $\mathbf{Var}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbf{Var}[X_i] + 2\sum_{i=1}^{n}\sum_{i<j} \mathbf{Cov}[X_i, X_j]$.
Now, since $X_i \sim \mathit{Bernoulli}(1/n)$, $\mathbf{Var}[X_i] = \frac{1}{n}\left(1 - \frac{1}{n}\right)$.
$\mathbf{Cov}[X_i, X_j] = \mathbf{E}[X_iX_j] - \mathbf{E}[X_i]\mathbf{E}[X_j]$. Since $\mathbf{E}[X_iX_j] = \frac{1}{n(n-1)}$, $\mathbf{Cov}[X_i, X_j] = \frac{1}{n(n-1)} - \frac{1}{n^2}$.
Thus, $\mathbf{Var}[\sum_{i=1}^{n} X_i] = 1 - \frac{1}{n} + n(n-1)\left(\frac{1}{n(n-1)} - \frac{1}{n^2}\right) = 1$.

3.23

(a) Since each coin is fair, the pair of coins to decide the value of Y_i can be one of $(H,T),(T,H),(H,H),(T,T)$ with equal probability. Thus, $\Pr(Y_i = 0) = \Pr(Y_i = 1) = 1/2$.
(b) Let the ith coin be denoted as C_i. Then if the first pair is C_1, C_2, the second pair is C_1, C_3 and the third pair is C_2, C_3,
$\Pr(Y_1 = Y_2 = Y_3 = 0 \neq 1/8 = \Pr(Y_1 = 1)\Pr(Y_2 = 1)\Pr(Y_3 = 1)$.
(c) Let ith pair be C_a, C_b and jth pair be C_c, C_d.
If $|\{a,b,c,d\}| = 4$, then Y_i and Y_j are independent. The claim trivially holds.
If $|\{a,b,c,d\}| = 3$, then $\mathbf{E}[Y_iY_j] = \Pr(Y_i = Y_j = 1) = 1/4 = \mathbf{E}[X_i]\mathbf{E}[X_j]$.
(d) $\mathbf{Var}[Y] = \mathbf{Var}[\sum\limits_{i=1}^{m} Y_i] = \sum\limits_{i=1}^{m} \mathbf{Var}[Y_i]$ holds from the result of exercise 3.15.
Since $\mathbf{Var}[Y_i] = 1/4$, $\mathbf{Var}[Y] = \frac{n(n-1)}{8}$.
(e) By Chebyshev's inequality, $\Pr(|Y - \mathbf{E}[Y]| \geq n) \leq \frac{\mathbf{Var}[Y]}{n^2} = \frac{n-1}{8n} \leq \frac{1}{8}$.

3.24

3.25

3.26

By Chebyshev's inequality, $\Pr\left(\left|\frac{X_1+X_2+\cdots+X_n}{n} - \mu\right| > \epsilon\right) \leq \frac{\mathbf{Var}[\sum\limits_{i=1}^{n} X_i/n]}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$.
Since $0 \leq \lim\limits_{n\to\infty} \Pr\left(\left|\frac{X_1+X_2+\cdots+X_n}{n} - \mu\right| > \epsilon\right) \leq \lim\limits_{n\to\infty} \frac{\sigma^2}{n\epsilon^2} = 0$, the desired result is obtained by the squeeze theorem.

4 Chernoff and Hoeffding bounds

4.1

Let the number of games that Alice wins be X, where $X \sim B(n, 0.6)$. Alice will lose the tournament with probability $\Pr(X \leq \frac{n-1}{2})$. Now, let δ s. t. $(1-\delta) \times \frac{3n}{5} = \frac{n-1}{2}$ to obtain the tightest bound.
$\Pr(X \leq \frac{n-1}{2}) = \Pr(X \leq (1-\delta)\mathbf{E}[X]) \leq \exp(-\frac{3n}{5} \cdot \delta^2 \cdot \frac{1}{2})$
$= \exp(-\frac{1}{10}(\frac{1}{12}n + \frac{5}{6} + \frac{25}{12n})) \leq \exp(-\frac{1}{8})$ (AM-GM inequality).

4.2

With Markov's inequality, $\Pr(X \geq n/4) \leq (n/6)/(n/4) = 2/3$.
With Chebyshev's inequality, $\Pr(X \geq n/4) \leq \Pr(|X - n/6| \geq n/12) \leq \frac{\mathbf{Var}[X]}{(n/12)^2}$
$= \frac{144}{n^2} \times (n \cdot \frac{1}{6} \cdot \frac{5}{6}) = 20/n$.
To use Chernoff bounds, let $\delta = 1/2$. Then $\Pr(X \geq n/4) = \Pr(X \geq (1+\delta)\mathbf{E}[X])$
$\leq \left(\frac{e^{0.5}}{1.5^{1.5}}\right)^{n/6} = \left(\frac{e}{1.5^3}\right)^{n/12}$.

4.3

(a) Let $X \sim B(n,p)$. Then $M_X(t) = \mathbf{E}[e^{tX}] = \sum_{i=0}^{n} e^{it} \Pr(X = i)$
$= \sum_{i=0}^{n} e^{it} \binom{n}{i} p^i (1-p)^{n-i} = \sum_{i=0}^{n} \binom{n}{i} (pe^t)^i (1-p)^{n-i} = (pe^t + 1 - p)^n$.
(b) $M_{X+Y}(t) = \mathbf{E}[e^{t(X+Y)}] = \mathbf{E}[e^{tX}e^{tY}] = \mathbf{E}[e^{tX}]\mathbf{E}[e^{tY}] = (pe^t + 1 - p)^{m+n}$.
(c) Since moment generating function uniquely determines the distribution, $X + Y \sim B(m+n, p)$.

4.4

Let the total number of heads be X, where $X \sim B(100, \frac{1}{2})$. Then we find $\Pr(X \geq 55) \approx 0.1841$.
From Chernoff bound, we find that $\Pr(X \geq (1 + \frac{1}{10})50) \leq \exp(-\frac{50}{3} \cdot \frac{1}{10^2}) = \exp(-\frac{1}{6}) \approx 0.8465$.
For $Y \sim B(1000, \frac{1}{2})$, $\Pr(Y \geq 550) \approx 0.0009$.
From Chernoff bound, we find that $\Pr(Y \geq (1 + \frac{1}{10})500) \leq \exp(-\frac{500}{3} \cdot \frac{1}{10^2}) = \exp(-\frac{5}{3}) \approx 0.1889$.

4.5

Let $Y = NX$, so that we aim to satisfy $\Pr(|Y - Np| > N\epsilon p) \leq \delta$. Consider that $\Pr(Y > Np(1+\epsilon)) < \exp(-Np \cdot \frac{\epsilon^2}{3})$, and $\Pr(Y < Np(1-\epsilon)) < \exp(-Np \cdot \frac{\epsilon^2}{2})$.
Thus, we aim to satisfy $\exp(-Np \cdot \frac{\epsilon^2}{3}) + \exp(-Np \cdot \frac{\epsilon^2}{2}) \leq 2\exp(-Np \cdot \frac{\epsilon^2}{3}) \leq \delta$.
$\therefore N \geq \frac{3}{p\epsilon^2} \ln \frac{2}{\delta}$. With $\epsilon = 0.1$, $\delta = 0.05$ and $0.2 \leq p \leq 0.8$, $N \geq 1500 \ln 40 \approx 5533$.

4.6

(a) Let $X \sim B(1000000, 0.02)$. Then $\Pr(X \geq 40000) \leq e^{-20000/3}$.
(b) Set X and Y as given and choose k, l such that $l \leq k - 10000$ so that bounding $\Pr((X > k) \cap (Y < l))$ suffices. As examples, we choose $k = 15300$ and $l = 4900$ here. Since $X \sim B(510000, 0.02)$, $Y \sim B(490000, 0.02)$ and $X \perp\!\!\!\perp Y$, $\Pr((X > k) \cap (Y < l)) = \Pr(X > k)\Pr(Y < l) \leq e^{-10200/12} \times e^{-9800/8} = e^{-2025}$.

4.7

Recall that $M_X(t) = \prod_{i=1}^{n}(p_i e^t + (1 - p_i)) = \prod_{i=1}^{n}(1 + p_i(e^t - 1)) \leq \prod_{i=1}^{n} e^{p_i(e^t-1)}$
$= e^{\mu(e^t-1)}$ holds when X is the sum of Poisson trials ($\Pr(X_i = 1) = p_i$).
Let $t = \ln(1 + \delta)$ and follow the derivation of Chernoff bounds.
$\Pr(X \geq (1 + \delta)\mu_H) \leq \frac{\mathbf{E}[e^{tX}]}{e^{t(1+\delta)\mu_H}} \leq \frac{e^{\mu(e^t-1)}}{e^{t(1+\delta)\mu_H}} \leq \left(\frac{e^{e^t-1}}{e^{t(1+\delta)}}\right)^{\mu_H} = \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu_H}$.
Similarly, let $t = \ln(1 - \delta)$ and prove the latter inequality.
$\Pr(X \leq (1 - \delta)\mu_L) \leq \frac{\mathbf{E}[e^{tX}]}{e^{t(1-\delta)\mu_L}} \leq \frac{e^{\mu(e^t-1)}}{e^{t(1-\delta)\mu_L}} \leq \left(\frac{e^{e^t-1}}{e^{t(1-\delta)}}\right)^{\mu_L} = \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu_L}$. ■

4.8

For any permutation π produced with the given approach, $\Pr(f = \pi) = \prod_{i=1}^{n} \frac{1}{k+1-i}$ holds. Since the number of possible permutations is $\frac{k!}{(k-n)!} = \frac{1}{\Pr(f=\pi)}$, the given approach produces a permutation chosen uniformly at random from all permutations.
Now, let X_j be the number of black box calls to determine $f(j)$. Then $X_j \sim Geom(\frac{k+1-j}{k})$ holds. Thus, $\mathbf{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbf{E}[X_i] = \sum_{i=1}^{n} \frac{k}{k+1-i}$.
When $k = n$, $\mathbf{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \frac{n}{i} = nH(n) \approx n \ln n$.
Similarly, when $k = 2n$, $\mathbf{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \frac{2n}{n+i} = 2n(H(2n) - H(n)) \approx 2n \ln 2$. In this case, $\frac{2n+1-j}{2n} \geq \frac{2n+1-n}{2n} \geq \frac{1}{2}$.
Now, to derive the desired Chernoff bound, we first compute the moment generating function of $X = \sum_{i=1}^{n} X_j$. Let $p_i = \frac{2n+1-i}{2n}$. Since X_i are independent,
$\mathbf{E}[e^{tX}] = \prod_{i=1}^{n} \mathbf{E}[e^{tX_i}] = \prod_{i=1}^{n}\left(\prod_{j=1}^{\infty}(e^{tj}p_i(1 - p_i)^{j-1}\right) = \prod_{i=1}^{n}\left(\frac{p_i}{1-p_i}\prod_{j=1}^{\infty}(e^t(1 - p_i))^j\right)$.
Suppose that we choose t s. t. $0 < t < \ln 2$ when deriving the Chernoff bound.
Then $\mathbf{E}[e^{tX}] = \prod_{i=1}^{n} \frac{p_i e^t}{1-e^t(1-p_i)}$. Since $t > 0$, $\frac{\partial}{\partial p_i}\left(\frac{p_i e^t}{1-e^t(1-p_i)}\right) = \frac{1-e^t}{(1-e^t(1-p_i))^2} < 0$.
This leads to $\mathbf{E}[e^{tX}] = \prod_{i=1}^{n} \frac{p_i e^t}{1-e^t(1-p_i)} \leq \left(\frac{\frac{1}{2}e^t}{1-\frac{1}{2}e^t}\right)^n$.
Now derive the desired Chernoff bound with $\Pr(X \geq 4n) \leq \frac{\mathbf{E}[e^{tX}]}{e^{4nt}} \leq \left(\frac{1}{(2-e^t)e^{3t}}\right)^n$.

Since the function $(2-e^t)e^{3t}$ has its maximum at $t = \ln\frac{3}{2}$ and $0 < \ln\frac{3}{2} < \ln 2$, we choose $t = \ln\frac{3}{2}$ for the tightest possible bound.
The desired bound would be $\Pr(X \geq 4n) \leq \left(\frac{1}{(2-e^t)e^{3t}}\right)^n \Big|_{t=\ln\frac{3}{2}} = (\frac{16}{27})^n$.

4.9

(a) By Chebyshev's inequality, $\Pr[|\sum_{i=1}^{t} X_i - \mathbf{E}[X]| \geq \epsilon\mathbf{E}[X]] \leq \frac{\mathbf{Var}[X]}{t(\epsilon\mathbf{E}[X])^2} = \frac{r^2}{t\epsilon^2}$.
Thus, setting t to satisfy $\frac{r^2}{t\epsilon^2} \leq \delta$ suffices. This leads to $t \geq \frac{r^2}{\epsilon^2\delta}$, which proves the claim.
(b) Set $\delta = 1 - 3/4 = 1/4$. Then we get $t \geq \frac{4r^2}{\epsilon^2}$, which proves the claim.
(c) Let Y_i be indicator variables that are 1 if $|X_i - \mathbf{E}[X]| \geq \epsilon\mathbf{E}[X]$. Then let the median of Y_is be m, and bound the probability $\Pr(|m - \mathbf{E}[X]| \geq \epsilon\mathbf{E}[X])$.
Note that $\mathbf{E}[\sum_{i=1}^{t} Y_i] \leq t/4$ by definition, and $|m - \mathbf{E}[X]| \geq \epsilon\mathbf{E}[X]$ holds only if $\sum_{i=1}^{t} Y_i \geq t/2$. Then, $\Pr(|m - \mathbf{E}[X]| \geq \epsilon\mathbf{E}[X]) \leq \Pr\left(\sum_{i=1}^{t} Y_i \geq t/2\right)$. Let $Y = \sum_{i=1}^{t} Y_i$. Then $\Pr(Y \geq t/2) = \Pr\Big(Y \geq (1 + (\frac{t}{2\mathbf{E}[Y]} - 1))\mathbf{E}[Y]\Big)$
$\leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{\mathbf{E}[Y]} = (\frac{2e}{t})^{t/2} \times e^{-\mathbf{E}[Y]}\mathbf{E}[Y]^{t/2}$.
Since $\frac{\partial}{\partial\mathbf{E}[Y]}\left((\frac{2e}{t})^{t/2}e^{-\mathbf{E}[Y]}\mathbf{E}[Y]^{t/2}\right) = (\frac{2e}{t})^{t/2}e^{-\mathbf{E}[Y]}\mathbf{E}[Y]^{t/2-1}(t/2 - \mathbf{E}[Y]) > 0$, substitute $t/4$ for $\mathbf{E}[Y]$ to derive our bound. Thus, $\Pr(Y \geq t/2) \leq (\frac{e}{4})^{t/4}$.
Here we need t that satisfies $(\frac{e}{4})^{t/4} \leq \delta$, which leads to $t \geq \frac{4}{\ln\frac{4}{e}}\ln\frac{1}{\delta}$. Therefore, together with 4.9.(b), we only need $O(\log(1/\delta))$ estimates constructed from $O(r^2\log(1/\delta)/\epsilon^2)$ samples.

4.10

Let $X = \sum_{i=1}^{1000000} X_i$ where X_i denotes the winnings of the ith game.
Then by the Chernoff bound, $\Pr(X \geq 10000) \leq \frac{\mathbf{E}[e^{tX}]}{e^{10000t}} = \left(\frac{\mathbf{E}[e^{tX_i}]^{100}}{e^t}\right)^{10000}$
$= \left(\frac{(167/200)e^{-t} + (4/25)e^{2t} + (1/200)e^{99t}}{e^{0.01t}}\right)^{1000000}$. Using graph software, you can choose $t = 0.0006$ and derive $\Pr(X \geq 10000) \leq 0.0001606$.

4.11

Since $\mathbf{E}[X_i] = 1$, $\mathbf{E}[X] = n$. Thus, we bound $\Pr(X \geq (1+\delta)n)$ as
$\Pr(X \geq (1+\delta)n) \leq \frac{\mathbf{E}[e^{tX}]}{e^{t(1+\delta)n}}$ with $t > 0$. $\mathbf{E}[e^{tX}] = \prod_{i=1}^{n}\mathbf{E}[e^{tX_i}] = \left(\frac{1}{3}(1 + e^t + e^{2t})\right)^n$ leads to $\Pr(X \geq (1+\delta)n) \leq \left(\frac{1+e^t+e^{2t}}{3e^{t(1+\delta)}}\right)^n$. Although $t = \frac{\delta+\sqrt{4-3\delta^2}}{1-\delta}$ minimizes $\frac{1+e^t+e^{2t}}{3e^{t(1+\delta)}}$, it is too complex to be used as a generalized bound. Thus, we put

$t = \ln(1+\delta)$ for simplicity and derive $\Pr(X \geq (1+\delta)n) \leq \left(\frac{3+3\delta+\delta^2}{3(1+\delta)^{(1+\delta)}}\right)^n$.
The Chernoff bound for $\Pr(X \leq (1-\delta)n)$ can also be derived in a similar way.

4.12

(a) We can think of X_i as the number of tails between $i-1$th head and ith head. Now, let Y_i be indicator variables that are 1 if ith flip is head. Then let $Y = \sum_{i=1}^{(1+\delta)2n} Y_i$, and derive the Chernoff bound as $\Pr(X \geq (1+\delta)2n) = \Pr(Y \leq n)$. Since $\mathbf{E}[Y] = (1+\delta)n$, $\Pr(Y \leq n) = \Pr(Y \leq (1-\frac{\delta}{1+\delta})\mathbf{E}[Y]) \leq e^{-\frac{1}{2}\mathbf{E}[Y](\frac{\delta}{1+\delta})^2} = e^{-\frac{n\delta^2}{2(1+\delta)}}$.
(b) Here, the moment generating function for X can be derived as $\mathbf{E}[e^{tX}] = \left(\frac{e^t}{2-e^t}\right)^n$ for $0 < t < \ln 2$ (refer to the solution for exercise 4.8).
Thus, $\Pr(X \geq (1+\delta)2n) \leq \frac{(\frac{e^t}{2-e^t})^n}{e^{t(1+\delta)2n}} = \left(\frac{1}{e^{t(1+2\delta)}(2-e^t)}\right)^n$. Since $e^{t(1+2\delta)}(2-e^t)$ is maximized at $t = \ln(\frac{1+2\delta}{1+\delta}) < \ln 2$, we choose it to derive the tightest bound. Therefore, $\Pr(X \geq (1+\delta)2n) \leq \left((\frac{1+\delta}{1+2\delta})^{1+2\delta}(1+\delta)\right)^n$.
(c) To compare two bounds, we inspect the sign of $e^{-\frac{\delta^2}{2(1+\delta)}} - (\frac{1+\delta}{1+2\delta})^{1+2\delta}(1+\delta)$. The simpler equivalent would be $(1+2\delta)\ln(1+2\delta) - (2+2\delta)\ln(1+\delta) - \frac{\delta^2}{2(1+\delta)}$. The computation can be performed numerically using $\ln(x+1) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n} x^n$, or with the help of graph software. In either way, it can be shown that the bound derived in (b) is better.

4.13

(a) From the Chernoff bound, $\Pr(X \geq xn) \leq \mathbf{E}[e^{tX}]/e^{txn}$.
Since $\mathbf{E}[e^{tX}] = \prod_{i=1}^{n} \mathbf{E}[e^{tX_i}] = (1-p+pe^t)^n$, $\Pr(X \geq xn) \leq \left(\frac{1-p+pe^t}{e^{xt}}\right)^n = ((1-p)e^{-xt} + pe^{(1-x)t})^n$. To derive the tightest bound, we solve for $\frac{\partial}{\partial t}((1-p)e^{-xt} + pe^{(1-x)t}) = 0$, which gives $t = \ln(x(1-p)) - \ln((1-x)p)$. Since $(1-p)e^{-xt} + pe^{(1-x)t}$ is convex w. r. t. t with given conditions, this gives the minimum. By plugging this in, we can show that $\Pr(X \geq xn) \leq e^{-nF(x,p)}$.
(b) Since $\frac{\partial^2}{\partial x^2}(F(x,p) - 2(x-p)^2) = \frac{1}{x} + \frac{1}{1-x} - 4 = \frac{(2x-1)^2}{x(1-x)} \geq 0$, $F(x,p) - 2(x-p)^2$ is convex w. r. t. x when $0 < x, p < 1$. Considering that $\frac{\partial}{\partial x}(F(x,p) - 2(x-p)^2) = 0$ yields $x = p$ and $(F(x,p) - 2(x-p)^2)\Big|_{x=p} = 0$, we get $F(x,p) - 2(x-p)^2 \geq 0$.
(c) $\Pr(X \geq (p+\epsilon)n) \leq e^{-nF(p+\epsilon,p)}$ holds by (a), and $e^{-nF(p+\epsilon,p)} \leq e^{-n\times 2(p+\epsilon-p)^2} = e^{-2n\epsilon^2}$ holds by (b).
(d) Take $Y_i = 1 - X_i$, and let $Y = n - X$. Then, $\Pr(X \leq (p-\epsilon)n) =$

$\Pr(Y \geq ((1-p)+\epsilon)n) \leq e^{-2n\epsilon^2}$ holds by (c). Combined with (c), we get $\Pr(|X - pn| \geq \epsilon n) = \Pr(X \leq (p-\epsilon)n) + \Pr(X \geq (p+\epsilon)n) \leq 2e^{-2n\epsilon^2}$.

4.14

We first bound the moment generating function of X to $\mathbf{E}[e^{tX}] = \prod_{i=1}^{n}(1 - p_i + e^{ta_i}p_i) = \prod_{i=1}^{n}(1 + p_i(e^{ta_i} - 1)) \leq \prod_{i=1}^{n}\exp(p_i(e^{ta_i} - 1))$. Since $0 \leq a_i \leq 1$, $\prod_{i=1}^{n}\exp(p_i(e^{ta_i} - 1)) \leq \prod_{i=1}^{n}\exp(p_i(e^{t} - 1)) = \exp(\sum_{i=1}^{n} p_i(e^t - 1)) = \exp(\mu(e^t - 1))$.
Following the proof of Theorem 4.4 in the textbook, we get $\Pr(X \geq (1+\delta)\mu) = \Pr(e^{tX} \geq e^{t(1+\delta)\mu}) \leq \frac{\exp((e^t-1)\mu)}{\exp(t(1+\delta)\mu)}$ for $t > 0$. Now take $t = \ln(1+\delta)$, and we show the desired Chernoff bound.
Similarly, $\Pr(X \leq (1-\delta)\mu) = \Pr(e^{tX} \geq e^{t(1-\delta)\mu}) \leq \frac{\exp((e^t-1)\mu)}{\exp(t(1-\delta)\mu)}$ for $t < 0$. Now for $0 < \delta < 1$, take $t = \ln(1-\delta)$, and we can show that $\Pr(X \leq (1-\delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}$.

4.15

Note that $|(1-p_i) - (-p_i)| = 1$, and $\mathbf{E}[X_i] = 0$ for all i. Applying the Hoeffding bound to X, we get $\Pr(|\frac{1}{n}\sum_{i=1}^{n} X_i| > \epsilon) = \Pr(|X| > n\epsilon) \leq 2e^{-2n\epsilon^2}$. Now take $\epsilon = \frac{a}{n}$ to get $\Pr(|X| > a) \leq 2e^{-2a^2/n}$.

4.16

Let $Y_i = a_i(X_i - p_i)$. Then we observe that $\Pr(-a_i p_i \leq Y_i \leq a_i(1-p_i)) = 1$ and $\mathbf{E}[Y_i] = 0$. Now applying the Hoeffding bound to $\sum_{i=1}^{n} Y_i$ as:
$\Pr(|\sum_{i=1}^{n} Y_i| \geq \delta\mu) \leq 2e^{-2n(\frac{\delta\mu}{n})^2 \frac{1}{\max a_i^2}} \leq 2e^{-2\delta^2\mu^2/n}$.

4.17

Let the total time (in steps) of a single processor be $X = \sum_{i=1}^{n/m} X_i$, where X_i is the number of steps for the ith job of the processor. Since $\mathbf{E}[X_i] = p + (1-p)k$, we take Y_i as $X_i = 1 + (k-1)Y_i$ so that $Y_i \sim \mathit{Bernoulli}(1-p)$. With Y_i, we get $X = (n/m) + (k-1)\sum_{i=1}^{n/m} Y_i$.

Applying the Chernoff bounds to $Y = \sum_{i=1}^{n/m} Y_i$, we get $\Pr(|Y - \mathbf{E}[Y]| \geq \delta\mathbf{E}[Y]) \leq 2e^{-\mathbf{E}[Y]\delta^2/3}$. Since $\mathbf{E}[Y] = (n/m)(1-p)$, we can bound X as:

$\Pr(|X-(n/m)(1+(1-p)(k-1))| \geq \delta(n/m)(1-p)(k-1)) \leq 2e^{-\frac{n}{m}(1-p)\frac{\delta^2}{3}}$. Using the union bound, we bound the total time T to $\Pr(|T-(n/m)(1+(1-p)(k-1))| \geq \delta(n/m)(1-p)(k-1)) \leq 2me^{-\frac{n}{m}(1-p)\frac{\delta^2}{3}}$. We can take $\delta = \sqrt{\frac{3\ln(2m/\epsilon)}{(n/m)(1-p)}}$ to derive the bound with probability of at least $1-\epsilon$.

4.18

5 Balls, Bins, and Random Graphs

5.1

As $(1+1/n)^n$ increases, we find the smallest n to reach the threshold.
$(1+1/n)^n$ first reaches $0.99e$ at $n=50$, and $0.999999e$ at $n=499982$.
Since $(1-1/n)^n$ also increases, we solve in a similar way. $(1-1/n)^n$ first reaches $0.99/e$ at $n=51$ and $0.999999/e$ at $n=499991$.

5.2

Recall the formula used in the birthday paradox: If there are N possibilities, then we solve for the smallest n that satisfies $\prod_{i=1}^{n-1}(1-\frac{i}{N}) \approx \prod_{i=1}^{n-1} e^{-i/n} = e^{-(n-1)n/2N} < 1/2$. Note that we omitted the final approximation to derive exact numerical answers.
Regardless of whether the number of Social Security number digits is 9 or 13, using the last four digits gives $N = 10000$ and this gives $n = 119$.
In the case where the number of digits is 9 ($N = 10^9$), we get $n = 37234$.
In the case where the number of digits is 13 ($N = 10^{13}$), we get $n = 3723298$.

5.3

Let the number of balls thrown be m. Then the desired probability is $\prod_{i=0}^{m-1}(1-\frac{i}{n})$.
We first determine c_1. $m = c_1\sqrt{n}$ should satisfy $\prod_{i=0}^{m-1}(1-\frac{i}{n}) \le \prod_{i=0}^{m-1} e^{-i/n} = e^{-(m-1)m/2n} \le e^{-1}$. Since $(m-1)m = c_1^2 n - c_1\sqrt{n} \ge 2n$, $(c_1^2-2)\sqrt{n} \ge c_1$.
Therefore, we choose c_1 that is greater than or equal to $\frac{1}{2}\left(\frac{1}{\sqrt{n}} + \sqrt{\frac{1}{n}+8}\right)$.
Now we determine c_2. To use the given hint, assume that $2m < n$.
$\prod_{i=0}^{m-1}(1-\frac{i}{n}) \ge \prod_{i=0}^{m-1}\exp(-\frac{i}{n}-\frac{i^2}{n^2}) = \exp(-\frac{m(m-1)}{2n} - \frac{(m-1)m(2m-1)}{6n^2})$
$= \exp(-\frac{m(m-1)}{2n}(1+\frac{2m-1}{3n})) \ge \exp(-\frac{m^2}{2n}(1+\frac{2m}{3n})) \ge \frac{1}{2}$ should be satisfied for $m = c_2\sqrt{n}$. This is equivalent to satisfying $\frac{c_2^2}{2}(1+\frac{2c_2}{3\sqrt{n}}) \le \ln 2$.
Since n is sufficiently large, choosing $c_2 = \sqrt{2\ln 2 - \frac{1}{\ln n}}$ yields the desired result.

5.4

Let event A indicate that there exist two or more people who share a birthday, and event B indicate that exactly two people share a birthday. Then our desired probability would be $\Pr(A-B) = \Pr(A) - \Pr(B)$ since $B \subset A$.
We first determine $\Pr(A)$, which is easy: $\Pr(A) = 1 - \prod_{i=1}^{100}\frac{366-i}{365}$.
We now determine $\Pr(B)$. If there are i shared birthdays in which each day is

shared by exactly two people, then the number of possible permutations would be the product of the following terms:
$\binom{365}{i}$ ways to choose i shared days, $\binom{100}{2i}$ ways to choose $2i$ people to share birthdays, $\prod_{j=1}^{i} \binom{2j}{2}$ ways to distribute i birthdays to $2i$ people and $\prod_{j=1}^{100-2i} (366 - i - j)$ ways to distribute unique birthdays to the rest.
Thus, $\Pr(B) = \sum_{i=0}^{50} \frac{365!100!}{i!(100-2i)!(265+i)!2^i} \times \frac{1}{365^{100}}$.
Therefore, we can determine our desired probability $\Pr(A-B) = \Pr(A) - \Pr(B)$.

5.5

Let $X \sim Poisson(\lambda)$. Then $M_X(t) = \mathbf{E}[e^{tX}] = e^{\lambda(e^t-1)}$ holds. By computing the second derivative of $M_X(t)$ with respect to t and plugging $t = 0$ in, we get $\mathbf{E}[X^2] = \lambda + \lambda^2$. Thus, $\mathbf{Var}[X] = \lambda$ follows.

5.6

We first show that $Y \sim Poisson(\mu p)$.
$\Pr(Y = k) = \sum_{i=k}^{\infty} \Pr(X = i)\binom{i}{k}p^k(1-p)^{i-k} = \sum_{i=k}^{\infty} \frac{e^{-\mu}\mu^i}{i!} \frac{i!}{k!(i-k)!} p^k (1-p)^{i-k}$
$= \frac{e^{-\mu}(p\mu)^k}{k!} \sum_{i=k}^{\infty} \frac{(\mu(1-p))^{i-k}}{(i-k)!} = \frac{e^{-\mu}(p\mu)^k}{k!} e^{\mu(1-p)} = \frac{e^{-\mu p}(\mu p)^k}{k!}$.
We can also similarly show that $Z \sim Poisson(\mu(1-p))$.
Now we show that $\Pr(Y = i, Z = j) = \Pr(Y = i)\Pr(Z = j)$. Note that $X = Y + Z$ by definition. This allows us to write $\Pr(Y = i, Z = j)$ as $\Pr(Y = i, X = i + j) = \Pr(X = i + j)\binom{i+j}{i}p^i(1-p)^j = \frac{e^{-\mu}\mu^{i+j}}{i!j!}p^i(1-p)^j$.
Since $\Pr(Y = i)\Pr(Z = j) = \frac{e^{-p\mu}(p\mu)^i}{i!} \frac{e^{-(1-p)\mu}((1-p)\mu)^j}{j!} = \frac{e^{-\mu}\mu^{i+j}p^i(1-p)^j}{i!j!} = \Pr(Y = i, X = i + j)$, $Y \perp\!\!\!\perp Z$. ■

5.7

We first prove that $\ln(1 + x) \leq x$, which is equivalent to $1 + x \leq e^x$.
Since $\ln(1 + x) - x = -\frac{x^2}{2} + \frac{x^3}{3} - \cdots$, this can be seen as an alternating series as $\frac{x^n}{n}$ is monotonically decreasing in $|x| \leq 1$. We can apply rearrangements to the alternating series as $\ln(1 + x) - x = -\sum_{n=1}^{\infty} \left(\frac{x^{2n}}{2n} - \frac{x^{2n+1}}{2n+1}\right)$, since the Taylor expansion of $\ln(1 + x)$ is absolutely convergent (to $e^x - 1$). The rearrangement gives $\ln(1 + x) - x \leq 0$, which is the desired result.
We now prove $x + \ln(1-x^2) \leq \ln(1+x)$, which is equivalent to $e^x(1-x^2) \leq 1+x$. Since $\ln(1 - x^2) = \ln(1 + x) + \ln(1 - x)$, $x + \ln(1 - x^2) \leq \ln(1 + x)$ is reduced to $\ln(1 - x) \leq -x$. At $|x| \leq 1$, this is equivalent to $\ln(1 + x) \leq x$, which we have previously proved. ■

5.8

(a) Since the ball is equally likely to fall in one of the three bins, the desired probability is 1/3.
(b) Since the bin 2 did not receive balls, we can simply think of this as throwing balls n into $n-1$ bins. The conditional expectation would be $n/(n-1)$.
(c) Note that the probability that bin 1 receives more balls than bin 2 is the same as that of bin 2 receiving more balls than bin 1. Thus, we first compute the probability that two bins receive the same number of balls, which is $P = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k}\binom{2k}{k}(\frac{1}{n})^{2k}(1-\frac{2}{n})^{n-2k}$. The desired probability would be $(1-P)/2$.

5.9

In the given condition, the expected number of elements in a single bucket is at most a. Since $a = O(1)$, sorting all buckets can still be done in linear time.

5.10

(a) By the Poisson approximation, the probability p is bounded as $p \le e\sqrt{n}(\frac{1}{e})^n$.
(b) $\frac{n!}{n^n}$.
(c) Since $\Pr(Z = n) = \frac{e^{-n}n^n}{n!}$ when $Z \sim Poisson(n)$, $\frac{n!}{n^n} \times \frac{e^{-n}n^n}{n!} = e^{-n}$ shows the claim. Theorem 5.6 states that the distribution $(Y_1, ..., Y_n)$ constrained on $\sum_{i=1}^{n} Y_i = n$ is equivalent to the balls and bins model. Note that each Y_i follows $Poisson(1)$ and each X_i denotes the load of the ith bin in the balls and bins model. Then using theorem 5.6, $(1/e)^n/(\frac{e^{-n}n^n}{n!}) = \frac{\Pr(\forall i, Y_i=1)}{\Pr(\sum_i Y_i=n)} = \Pr(\forall i, Y_i = 1 | \sum_i Y_i = n) = \Pr(\forall i, X_i = 1) = \frac{n!}{n^n}$.

5.11

Let X_i be the indicator variables that are 1 if there is a k-gap starting at i.
Let $X = \sum_{i=0}^{n-k} X_i$.

(a) The expected number of k-gaps would be $\sum_{i=0}^{n-k} \mathbf{E}[X_i] = (n-k+1)(1-\frac{k}{n})^m$.
(b) First, we assume that the bin loads follow the Poisson distribution to derive the Poisson-approximated Chernoff bound. We divide $\{X_i\}$ into k subsets, so that all indicator variables in the same subset are independent. First, we derive the Chernoff bound for $Y_0 = \sum_{i \ge 0} X_{ik}$ where $0 < \delta < 1$.

$\Pr(Y_0 \ge (1+\delta)\mathbf{E}[Y_0]) \le \exp(-\frac{1}{3}\frac{\mathbf{E}[X]}{k}\delta^2)$, and $\Pr(Y_0 \le (1-\delta)\mathbf{E}[Y_0]) \le \exp(-\frac{1}{2}\frac{\mathbf{E}[X]}{k}\delta^2)$ holds. Therefore, $\Pr(|Y_0 - \mathbf{E}[Y_0]| \ge \delta\mathbf{E}[Y_0]) \le 2\exp(-\frac{1}{3}\frac{\mathbf{E}[X]}{k}\delta^2)$.
With the union bound for all Y_i, we get $\Pr(|X - \mathbf{E}[X]| \ge \delta\mathbf{E}[X]) \le 2k\exp(-\frac{1}{3}\frac{\mathbf{E}[X]}{k}\delta^2)$.

Since we have used the Poisson approximation to compute the Chernoff bound, the computed upper bound should be multiplied by $e\sqrt{m}$.

5.12

Let X_i be the indicator variables that are 1 if a ball landed in bin i by itself.
(a) The expected number of balls to be *served* in this round would be $\sum_{i=1}^{n} \mathbf{E}[X_i] = \sum_{i=1}^{n} b \times \frac{1}{n}(1-\frac{1}{n})^{b-1} = b(1-\frac{1}{n})^{b-1}$. Therefore, the expected number of balls at the start of the next round would be $b(1-(1-\frac{1}{n})^{b-1})$.
(b) Note that if $n = 1$, then the number of rounds required would be trivially 1. Therefore, we only consider the cases where $n \geq 2$.
Since $x_{j+1} = x_j(1-(1-\frac{1}{n})^{x_j-1}) \leq x_j(1-(1-\frac{x_j-1}{n})) = x_j\frac{x_j-1}{n} \leq \frac{x_j^2}{n}$, the inequality given in the hint is true.
With $x_1 = n(1-(1-\frac{1}{n})^{n-1})$, cascading the inequality yields $x_k \leq n(\frac{x_1}{n})^{2^{k-1}} = n(1-(1-\frac{1}{n})^{n-1})^{2^{k-1}} \leq n(1-(1-\frac{1}{n})^{n})^{2^{k-1}}$.
Now, let k^* be the minimum k that satisfies $n(1-(1-\frac{1}{n})^n)^{2^{k-1}} \leq 1$, so that the operation ends after no more than k^*+1 rounds. Since $1-(1-\frac{1}{n})^n \leq \frac{3}{4}$ at $n \geq 2$, we calculate the minimum k that satisfies $n(\frac{3}{4})^{2^{k-1}} \leq 1$.
Taking log on both sides gives $\ln n + 2^{k^*-1}\ln\frac{3}{4} \leq 0$, which is equivalent to $2^{k^*-1} \geq \frac{\ln n}{\ln\frac{4}{3}}$. Therefore, we get $k^*-1 \geq \ln\frac{\ln n}{\ln\frac{4}{3}}$, which shows that $k^* = O(\log\log n)$.

5.13

Let the load of bin i be X_i, and let $Y_k = X_{kn/\log_2 n}$ where $k \in \mathbb{N}_0$. Then for all i, there exists $k \in \mathbb{N}$ such that $kn/\log_2 n \leq i \leq (k+1)n/\log_2 n$.
When a ball is thrown into the bin i, only one of the two bins (bin $kn/\log_2 n$ and bin $(k+1)n/\log_2 n$) must be chosen together. This means that $X_i \leq Y_k + Y_{k+1}$.
Since only one of the bins in $S = \{i | i = kn/\log_2 n\}$ gets a ball within a player's round, we can see Y_k as each bin in the model of balls and bins with $\log_2 n$ bins and $\log_2 n$ balls.
Recall that the probability that the maximum load is more than $3\ln n/\ln\ln n$ is at most $1/n$ in the model of n balls and n bins (Lemma 5.1). Therefore, we bound the probability that the maximum load is greater than $6M = 6\ln\log_2 n/\ln\ln\log_2 n$, considering $X_i \leq Y_k + Y_{k+1}$. Thus, we can derive the upper bound as $\Pr(\sum_{i=0}^{n-1} \mathbf{1}_{X_i \geq 6M} > 0) \leq \Pr(\sum_{i=0}^{\log_2 n-1} \mathbf{1}_{Y_i \geq 3M} > 0) \leq \frac{1}{\log_2 n}$. Thus, the maximum load is less than $6\ln\log_2 n/\ln\ln\log_2 n = O(\log\log n/\log\log\log n)$ with probability $1-\frac{1}{\log_2 n}$ which approaches 1 as $n \to \infty$.

5.14

(a) $\Pr(Z = \mu + h) - \Pr(Z = \mu - h - 1) = \frac{e^{-\mu}\mu^{\mu+h}}{(\mu+h)!} - \frac{e^{-\mu}\mu^{\mu-h-1}}{(\mu-h-1)!} = \frac{\mu^{\mu-h}}{(\mu+h)!}(\mu^{2h} - \sum_{i=1}^{h}(\mu^2 - i^2))$. Since $\mu^2 > \mu^2 - i^2$, $\Pr(Z = \mu + h) - \Pr(Z = \mu - h - 1) \geq 0$ holds and the claim is proved.

(b) $\Pr(Z \geq \mu) \geq \sum_{h=0}^{\mu-1} \Pr(Z = \mu + h) \geq \sum_{h=0}^{\mu-1} \Pr(Z = \mu - h - 1) = \Pr(Z < \mu) = 1 - \Pr(Z \geq \mu)$ shows that $2\Pr(Z \geq \mu) \geq 1$, proving the claim.

(c) Numerical validation can show that $\Pr(Z \geq \mu) \leq 1/2$ for all integers μ from 1 to 10.

5.15

(a) $\mathbf{E}[f(Y_1^{(m)}, ..., Y_n^{(m)})] = \sum_{k=0}^{\infty} \mathbf{E}[f(X_1^{(k)}, ..., X_n^{(k)})] \Pr(\sum_{i=1}^{n} Y_i^{(m)} = k)$ holds (recall the proof of Theorem 5.7).

If $\mu(m) = \mathbf{E}[f(X_1^{(m)}, ..., X_n^{(m)})]$ is monotonically increasing in m, then we have $\mu(k) \geq \mu(m)$ for $k \geq m$ for some m. Thus, $\mathbf{E}[f(Y_1^{(m)}, ..., Y_n^{(m)})] = \sum_{k<m} \mu(k) \Pr(\sum_{i=1}^{n} Y_i^{(m)} = k) + \sum_{k \geq m} \mu(k) \Pr(\sum_{i=1}^{n} Y_i^{(m)} = k)$ $(f \geq 0)$

$\geq \sum_{k \geq m} \mu(m) \Pr(\sum_{i=1}^{n} Y_i^{(m)} = k) = \mathbf{E}[f(X_1^{(m)}, ..., X_n^{(m)})] \Pr(\sum Y_i^{(m)} \geq m)$.

Similarly, if $\mu(m)$ is monotonically decreasing in m, then we have $\mu(k) \geq \mu(m)$ for $k \leq m$ for some m. Thus,

$\mathbf{E}[f(Y_1^{(m)}, ..., Y_n^{(m)})] = \sum_{k \leq m} \mu(k) \Pr(\sum_{i=1}^{n} Y_i^{(m)} = k) + \sum_{k > m} \mu(k) \Pr(\sum_{i=1}^{n} Y_i^{(m)} = k)$

$\geq \sum_{k \leq m} \mu(m) \Pr(\sum_{i=1}^{n} Y_i^{(m)} = k) = \mathbf{E}[f(X_1^{(m)}, ..., X_n^{(m)})] \Pr(\sum Y_i^{(m)} \leq m)$.

(b) Since the sum of independent Poisson random variables is also a Poisson random variable, $\sum Y_i^{(m)} \sim Poisson(m)$. Using the result of exercise 5.14.(b) and 5.15.(a), $\mathbf{E}[f(Y_1^{(m)}, ..., Y_n^{(m)})] \geq \mathbf{E}[f(X_1^{(m)}, ..., X_n^{(m)})] \Pr(\sum Y_i^{(m)} \geq m)$ $\geq \mathbf{E}[f(X_1^{(m)}, ..., X_n^{(m)})] \times \frac{1}{2}$. Thus, Theorem 5.10 is proved for the monotonically increasing case.

If one can derive a formal proof on the statement of exercise 5.14.(c), Theorem 5.10 can also be proved for the monotonically decreasing case.

5.16

(a) Expectations are computed as $\mathbf{E}[X_1 X_2 \cdots X_k] = \Pr(X_1 X_2 \cdots X_k = 1) = (1 - \frac{k}{n})^n$ and $\mathbf{E}[Y_1 Y_2 \cdots Y_k] = \Pr(Y_1 Y_2 \cdots Y_k = 1) = p^k = (1 - \frac{1}{n})^{k \times n}$.

Since $1 - \frac{k}{n} \leq (1 - \frac{1}{n})^k$ by the Bernoulli inequality, $\mathbf{E}[X_1 X_2 \cdots X_k] \leq \mathbf{E}[Y_1 Y_2 \cdots Y_k]$.

(b) $\mathbf{E}[e^{tX}] = \mathbf{E}\left[\sum_{k=0}^{\infty} \frac{(tX)^k}{k!}\right] = \sum_{k=0}^{\infty} \mathbf{E}[\frac{(tX)^k}{k!}]$ holds since $\mathbb{N}$ is countable, and the same applies to $\mathbf{E}[e^{tY}]$. Therefore, we show $\mathbf{E}[X^k] \leq \mathbf{E}[Y^k]$ for $\forall k \in \mathbb{N}$.

$\mathbf{E}[X^k] = \mathbf{E}[(X_1+X_2+\cdots+X_n)^k]$ holds, and $(X_1+X_2+\cdots+X_n)^k$ is the sum of products where each product is in the form of $X_{i_1}X_{i_2}\cdots X_{i_k}$ ($|\{i_1, i_2, ...i_k\}| \le k$). Since X_i are indicator variables, $X_i^n = X_i$ holds for $\forall n \in \mathbb{N}$, so the repeats can be ignored from the product.
Suppose that $|\{i_1, i_2, ...i_k\}| = m$. Then the result of 5.16.(a) applies with $k = m$. Thus, $\mathbf{E}[X^k] \le \mathbf{E}[Y^k]$ holds and, therefore, $\mathbf{E}[e^{tX}] \le \mathbf{E}[e^{tY}]$ holds.
(c) $\Pr(X \ge (1+\delta)\mu) \le \frac{\mathbf{E}[e^{tX}]}{e^{t(1+\delta)\mu}} \le \frac{\mathbf{E}[e^{tY}]}{e^{t(1+\delta)\mu}}$ holds. Since $\mathbf{E}[e^{tY}] = (1-p+pe^t)^n$, we can choose $t = \ln(1+\delta)$ to derive $\Pr(X \ge (1+\delta)\mu) \le \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu$.

5.17

(a) Since there are $\binom{n}{5}$ possible 5-cliques and each 5-clique has $\binom{5}{2} = 10$ edges, the expected number of 5-cliques is $\binom{n}{5} \times p^{10}$. Solving $\binom{n}{5} \times p^{10} = 1$ for p gives $p = \binom{n}{5}^{-1/10}$.
(b) Since there are $\frac{1}{2}\binom{n}{6}\binom{6}{3}$ possible $K_{3,3}$ subgraphs and each $K_{3,3}$ subgraph has nine edges, the expected number of $K_{3,3}$ subgraphs is $\frac{1}{2}\binom{n}{6}\binom{6}{3} \times p^9$. Solving $\frac{1}{2}\binom{n}{6}\binom{6}{3} \times p^9 = 1$ for p gives $p = \left\{\frac{1}{2}\binom{n}{6}\binom{6}{3}\right\}^{-1/9}$.
(c) Since there are $(n-1)!/2$ possible Hamiltonian cycles and each Hamiltonian cycle has n edges, the expected number of Hamiltonian cycles is $\frac{1}{2}(n-1)! \times p^n$. Solving $\frac{1}{2}(n-1)! \times p^n = 1$ for p gives $p = \left\{\frac{1}{2}(n-1)!\right\}^{-1/n}$.

5.18

For any nonnegative function f, $\mathbf{E}[f(G_{n,p}) \mid |E| = k] = \mathbf{E}[f(G_{n,k})]$ holds.
Thus, $\mathbf{E}[f(G_{n,p})] = \sum_{k=0}^{M} \mathbf{E}[f(G_{n,k})] \Pr(|E| = k) \ge \mathbf{E}[f(G_{n,N})] \Pr(|E| = N)$ holds ($M = \binom{n}{2}$). Now we bound the probability $\Pr(|E| = N)$ for the $G_{n,p}$ model to derive the statement. Note that $p = N/M$. Using Stirling's bounds, $\Pr(|E| = N) = \binom{M}{N} p^N (1-p)^{M-N} = \frac{M!}{(M-N)!N!}(\frac{N}{M})^N(\frac{M-N}{M})^{M-N} \ge \frac{\sqrt{2\pi}}{e^2}\frac{\sqrt{M}}{\sqrt{MN-N^2}}$ $= \frac{\sqrt{2\pi}}{e^2}\frac{1}{\sqrt{N(1-p)}} \ge \frac{\sqrt{2\pi}}{e^2}\frac{1}{\sqrt{N}}$. Thus, we let f be an indicator of an event and prove that every event that happens with a small probability P in the $G_{n,p}$ model also happens with small probability (at most $\frac{e^2}{\sqrt{2\pi}}\sqrt{N} \times P$) in the $G_{n,N}$ model.

5.19

We use the result of exercise 5.18 to use the $G_{n,p}$ model. Let $M = \binom{n}{2}$ and $G \sim G_{n,p}$ where $p = \frac{r \ln n}{n}$ $(r > 2)$.
The probability that G is a disconnected graph containing a disconnected set of size $k < n/2$ is upper bounded by $\binom{n}{k}(1-p)^{k(n-k)}$ (union bound).
We can derive $\binom{n}{k} = \frac{(n-k+1)\cdots(n-1)n}{k!} \le \frac{n^k}{k!} \le (\frac{en}{k})^k$, where the last inequality is from $e^k = \sum_{i=0}^{\infty} \frac{k^i}{i!} > \frac{k^k}{k!}$. Using this, $\binom{n}{k}(1-p)^{k(n-k)} \le (\frac{en}{k})^k e^{-pk(n-k)} \le$

$(\frac{en}{k})^k e^{-pkn/2} = (\frac{en^{1-r/2}}{k})^k$.
We can then upper-bound the probability that G is a disconnected graph again by the union bound: $\sum_{k=1}^{n/2}(\frac{en^{1-r/2}}{k})^k = en^{1-r/2} + \sum_{k=2}^{n/2}(\frac{en^{1-r/2}}{k})^k = O(n^{1-r/2})$.
Since $r > 2$, $O(n^{1-r/2}) = o(1)$. Thus, G is connected with probability $1 - o(1)$.
Now, to use the result of exercise 5.18, let $N = Mp = \frac{r}{2}(n-1)\ln n$. Since $N = O(n \log n)$, $r > 3$ guarantees that $G_{n,N}$ is disconnected with probability $o(1)$, i. e. connected with probability $1 - o(1)$. Thus, any $c > 3/2$ can be the desired constant.

5.20

6 The Probabilistic Method

6.1

(a) Since all k literals must be false for a clause not to be satisfied, the probability that a clause is satisfied is $1-2^{-k}$. As the given SAT instance has m clauses, the expected number of satisfied clauses is $m(1-2^{-k})$. By the expectation argument, we know that there is an assignment that satisfies at least $m(1-2^{-k})$ clauses. Since counting the number of satisfied clauses in the given assignment can be done in $O(mk)$, we bound the probability p that a random assignment satisfies at least $m(1-2^{-k})$ clauses. Let the number of satisfied clauses be X.
Then $m(1-2^{-k}) = \mathbf{E}[X] = \sum_{i<m(1-2^{-k})} i \Pr(X=i) + \sum_{i \geq m(1-2^{-k})} i \Pr(X=i)$
$\leq \left(m(1-2^{-k})-1\right)(1-p) + mp$, as $X \leq m$.
Thus, $p \geq 1/(m2^{-k}+1)$ holds. This indicates that the expected running time of the algorithm would be $O(mk + m^2k2^{-k})$.
(b) We can process each variable one by one. For each variable x, compute the conditional expectation of X while leaving all unset variables as random, for both $x = True$ or $x = False$. Then take x as the value that gives a larger conditional expectation, breaking ties arbitrarily. Since the initial expectation is $m(1-2^{-k})$, the assignment should satisfy at least $m(1-2^{-k})$ as the conditional expectation never decreases.

6.2

(a) Consider randomly assigning a color to each edge, so that each K_4 is monochromatic with a probability of $2 \times 2^{-\binom{4}{2}} = 2^{-5}$. By the linearity of expectations, the expected number of monochromatic K_4 would be $\binom{n}{4}2^{-5}$. By the expectation argument, there exists a two-coloring of K_n where the number of monochromatic K_4 does not exceed $\binom{n}{4}2^{-5}$.
(b) Let p be the probability that a random coloring of K_n has no more than $\binom{n}{4}2^{-5}$ monochromatic K_4s. Let the number of monochromatic K_4 be X, and we bound p as $\binom{n}{4}2^{-5} = \mathbf{E}[X] = \sum_{i<\binom{n}{4}2^{-5}} i \Pr(X=i) + \sum_{i \geq \binom{n}{4}2^{-5}} i \Pr(X=i) \leq$ $\left(\binom{n}{4}2^{-5}-1\right)p + \binom{n}{4}(1-p)$. Thus, $p \geq 1/(1+\frac{31}{32}\binom{n}{4})$ holds. This indicates that the expected running time of the algorithm would be $O(n^8)$.
(c) We can color each edge one by one. For each edge e_i, compute the conditional expectation of X while leaving all unset edge colors as random for both colors. Then take the color of e_i as the value that gives a smaller conditional expectation, breaking ties arbitrarily. Since the initial expectation is $\binom{n}{4}2^{-5}$, the obtained two-coloring should have no more than $\binom{n}{4}2^{-5}$ monochromatic K_4s as the conditional expectation never increases.

6.3