Probability Theory

Solutions to Selected Exercises

1 Review Questions in Combinatorics

1. There are two points to consider: how many arrangements are there for 5 digits and 5 letters, and how many choices of digits and letters are there for any arrangement. Consider the second point: there are 10 different digits, and repetitions are allowed, so there are 10^5 possibilities. There are 26 letters, so altogether there are $10^5 \cdot 26^5$ possibilities. The only difference between the parts of the question is the number of arrangements. Denote a place for a digit by d, and a place for a letter by l.

(a) The single arrangement is *dddddlllll*. The number of possibilities is $10^5 \cdot 26^5$.

(b) In general the digits and letters should alternate. However, a single pair of adjacent digits is still possible. Observe that each letter (except perhaps for the last) is followed by a digit, so we have five objects: four ld pairs, and a single l. The remaining d may be anywhere in between, or to the left of, or to the right of these five objects − altogether 6 possibilities.

$$\begin{array}{ccc} ldldldldld & ldldldlddl & ldldlddldl \\ ldlddldldl & lddldldldl & dldldldldl \end{array}.$$

The overall number is $6 \cdot 10^5 \cdot 26^5$.

(c) The number of arrangements is $\binom{10}{5}$, so the overall number is $\binom{10}{5} \cdot 10^5 \cdot 26^5$.

2.

(a) Every combination is a sequence of length 7 over $\{1, \ldots, 5\}$, in which every digit must appear at least once. By inclusion and exclusion, the number of such sequences is

$$\sum_{k=0}^{5}(-1)^k\binom{5}{k}(5-k)^7 = 16,800.$$

(b) The new combinations are sequences as in case (a), but the first two elements have to be distinct and their order is immaterial. Thus, from the result of the preceding part we have first to subtract the number of those sequences in which the first two entries are equal, which is

$$5\left(5!+4\cdot\binom{5}{2,1,1,1}\right) = 1800,$$

and then divide by 2. Altogether, we obtain

$$\frac{16800-1800}{2} = 7500$$

possibilities.

3.

(a) k^n.

(b) Let A_i be the event that the ith processor is assigned at least one job, $i=1,2,\ldots,k$, and A the event in the question. Clearly, $A=\bigcap_{i=1}^{k} A_i$. By the principle of inclusion and exclusion, and using the symmetry of the events A_i, we have

$$\begin{aligned} P(\overline{A}) &= P\left(\bigcup_{i=1}^{k}\overline{A_i}\right) \\ &= kP(\overline{A_1}) - \binom{k}{2}P(\overline{A_1A_2}) + \ldots + (-1)^{k-1}\binom{k}{k}P\left(\bigcap_{i=1}^{k}\overline{A_i}\right). \end{aligned}$$

Thus

$$\begin{aligned} P(A) &= 1-P(\overline{A}) \\ &= 1-k\left(1-\frac{1}{k}\right)^n + \binom{k}{2}\left(1-\frac{2}{k}\right)^n + \ldots + (-1)^{k-2}\binom{k}{k-1}\cdot\frac{1}{k^n}. \end{aligned}$$

4.

(a) r^n.

(b) The first letter of the word may be any of the letters in Σ. In each of the other $n-1$ places we may put any of the $r-1$ letters distinct from the one in the preceding place. Hence there are $r(r-1)^{n-1}$ possibilities in all.

(c) Let a_n be the number of words not containing two consecutive occurrences of σ. The set of words in question consists of two disjoint subsets – those starting with σ, and those starting with some other letter. In the first set, each word may have any letter distinct from σ in the second place, and the remaining $n-2$ letters must form a word of length $n-2$ satisfying our condition. In the second set, each word may have any letter distinct from σ as the first letter, and the remaining $n-1$ letters must form a word of length $n-1$ satisfying our condition. Thus we have:

$$a_n = (r-1)a_{n-2} + (r-1)a_{n-1}, \qquad a_0 = 1, a_1 = r.$$

Solving the above recurrence relation, we obtain:

$$x^2 + (r-1)x - (r-1) = 0.$$

Hence

$$a_n = c_1 b_1^n + c_2 b_2^n,$$

where $b_1 = \frac{r-1+\sqrt{(r-1)(r+3)}}{2}$ and $b_2 = \frac{r-1-\sqrt{(r-1)(r+3)}}{2}$. Substituting in the initial conditions $a_0 = 1, a_1 = r$, we obtain:

$$c_1 = \frac{b_2 - r}{b_2 - b_1}, \qquad c_2 = \frac{r - b_1}{b_2 - b_1}.$$

(d) $\binom{n}{n_1, n_2, \ldots, n_r}$.

(e) $r^{\left[\frac{n+1}{2}\right]}$.

5.

(b) The generating function is $f(x) = \frac{1}{1-2x}$, and we obtain $a_n = 2^n,\ n \geq 0$.

8. Since $n! = e^{\sum_{i=1}^{n} \ln i}$ the inequality

$$e\left(\frac{n}{e}\right)^n \leq n! \leq e\left(\frac{n+1}{e}\right)^{n+1}$$

is equivalent to

$$n \ln n - n + 1 \leq \sum_{i=1}^{n} \ln i \leq (n+1)\ln(n+1) - n.$$

Since $\int \ln x dx = x \ln x - x + c$ and the function $\ln x$ is increasing, we have

$$\int_1^n \ln x dx \leq \sum_{i=1}^{n} \ln i \leq \int_1^{n+1} \ln x dx,$$

which gives the required result.

9.

(a)

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2} \approx \frac{\sqrt{2\pi 2n}\left(\frac{2n}{e}\right)^{2n}}{\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right)^2} = \frac{2^{2n}}{\sqrt{\pi n}}\,.$$

(b) Consider the $2n$th row of Pascal's triangle. The sum of all entries is 2^{2n}, and therefore each of them, in particular the middle entry $\binom{2n}{n}$, is less than 2^{2n}. On the other hand, it is easy to check that the binomial coefficients $\binom{2n}{j}$ increase as j increases from 0 to n, and decrease from that place on. In particular, $\binom{2n}{n}$ is the maximal entry in the row. Consequently:

$$\frac{2^{2n}}{2n+1} \le \binom{2n}{n} \le 2^{2n}\,.$$

10. The set of binary trees on n vertices splits into $n+1$ disjoint subsets, depending on the number vertices in the left subtree. For each k, the number of trees with k vertices at the left subtree (and thus $n-k-1$ vertices at the right subtree) is $C_k \cdot C_{n-1-k}$. Hence:

$$C_n = \sum_{k=0}^{n-1} C_k C_{n-1-k}, \qquad n \ge 1,$$

with the initial condition $C_0 = 1$. It follows that

$$f(x) = c_0 + x\sum_{n=1}^{\infty}\left(\sum_{k=0}^{n-1} C_k \cdot C_{n-1-k}\right)x^{n-1} = xf(x)^2,$$

and therefore

$$xf(x)^2 - f(x) + 1 = 0.$$

Consequently

$$f(x) = \frac{1 \pm \sqrt{1-4x}}{2x} = \frac{1 \pm \left(1 - 2\sum_{k=1}^{\infty}\binom{2(k-1)}{k-1}\frac{x^k}{k}\right)}{2x}.$$

Choosing the possibility with the minus sign, we obtain

$$f(x) = \sum_{n=0}^{\infty}\binom{2n}{n}\frac{x^n}{n+1},$$

so that $C_n = \frac{1}{n+1}\binom{2n}{n}$.

11.

(a) The set of all permutations which may be obtained by the system in the question decomposes into a union of disjoint subsets as follows. Consider the number of the step, between 2 and $2n$ at which the number 1 was moved from S to Q_2. The step number may be any even number in this range. Consider the set of permutations in which this move took place at the $2j$th step. Thus, we moved the number 1 from Q_1 to S at the first step, moved the next $j-1$ integers $2, 3, \ldots, j$ from Q_1 to Q_2 (via S) in the next $2j-2$ steps, moved 1 from S to Q_2 at the $2j$th step, and then moved the remaining integers $j+1, j+2, \ldots, n$ from Q_1 to Q_2 in the next $2(n-j)$ steps. We had P_{j-1} possibilities for moving $2, 3, \ldots, j$ and P_{n-j} possibilities for moving $j+1, j+2, \ldots, n$. Consequenly, the sequence (P_n) satisfies the recurrence:

$$P_{n+1} = P_0P_n + P_1P_{n-1} + P_2P_{n-2} + \ldots + P_{n-1}P_1 + P_nP_0.$$

The sequence is completely determined by this recurrence and the initial condition $P_0 = 1$.

(b) Let $f(x) = \sum_{i=0}^{\infty} P_i x^i$. Solving the recurrence relation in the preceding part we have:

$$f(x) - P_0 = x \sum_{n=0}^{\infty} \sum_{i=0}^{n} P_i P_{n-i} x^n = x f^2(x),$$

which yields the solutions

$$f_1(x) = \frac{1 + \sqrt{1-4x}}{2x}$$

and

$$f_2(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

Now $\sqrt{1-4x} = \sum_{n=0}^{\infty} \binom{1/2}{n}(-4x)^n = 1 - \sum_{n=1}^{\infty} \frac{1}{2n-1}\binom{2n}{n} x^n$. Since the P_n's are positive, we have to select $f_2(x)$ as the relevant solution. The generating function is then

$$f(x) = \frac{1}{2x}\left(1 - \left(1 - \sum_{n=1}^{\infty} \frac{1}{2n-1}\binom{2n}{n} x^n\right)\right) = \sum_{n=0}^{\infty} \frac{1}{n+1}\binom{2n}{n} x^n,$$

which gives

$$P_n = \frac{1}{n+1}\binom{2n}{n}.$$

12.

(a) We need to make $l+m$ moves, out of which l should be to the right and the other m upwards. Hence the number of possibilities is $\binom{m+l}{l}$.

(b) The problem is equivalent to the ballot problem, so that the number of possibilities is $\binom{m+l}{l} - \binom{m+l}{l+1}$ if $m \leq k$ (and 0 otherwise).

(c) Clearly, the number of all walks going from $(0,0)$ to (l,m) is:

$$\begin{aligned} W_{l,m} &= W_{l,m-1} + W_{l-1,m} + W_{l-1,m-1}, \\ W_{l,0} &= W_{0,m} = 1, \quad l,m \geq 0. \end{aligned}$$

Thus

$$W_{l,m}x^l y^m = W_{l,m-1}x^l y^m + W_{l-1,m}x^l y^m + W_{l-1,m-1}x^l y^m$$

and

$$\sum_{m=1}^{\infty}\sum_{l=1}^{\infty} W_{l,m}x^l y^m = y\sum_{m=1}^{\infty}\sum_{l=1}^{\infty} W_{l,m-1}x^l y^{m-1} + x\sum_{m=1}^{\infty}\sum_{l=1}^{\infty} W_{l-1,m}x^{l-1}y^m + xy\sum_{m=1}^{\infty}\sum_{l=1}^{\infty} W_{l-1,m-1}x^{l-1}y^{m-1},$$

or equivalently

$$f(x,y) - \sum_{l=1}^{\infty} W_{l,0}x^l - \sum_{m=1}^{\infty} W_{0,m}y^m - W_{0,0} = y\left(f(x,y) - \sum_{m=0}^{\infty} W_{0,m}y^m\right) + x\left(f(x,y) - \sum_{m=0}^{\infty} W_{l,0}x^l\right) + xyf($$

Since $W_{l,0} = W_{0,m} = 1$ we obtain

$$f(x,y) - \frac{x}{1-x} - \frac{y}{1-y} - 1 = y\left(f(x,y) - \frac{1}{1-y}\right) + x\left(f(x,y) - \frac{1}{1-x}\right) + xyf(x,y),$$

which provides the result: $f(x,y) = 1/(1-x-y-xy)$.

2 Elementary Probability Calculations

18.

(a) For the event in question to occur, the first $\lceil n/2 \rceil$ tosses may have any outcomes, and then the other $\lfloor n/2 \rfloor$ tosses are uniquely determined by them. Hence the required probability is $2^{\lceil n/2 \rceil}/2^n = 1/2^{\lfloor n/2 \rfloor}$.

(b) There are only two sequences satisfying the property, and hence the probability is $\frac{2}{2^n} = \frac{1}{2^{n-1}}$.

(c) For a non-negative integer n and a word w over $\{0,1\}$, denote by $a_{n,w}$ the number of words of length n starting with w and satisfying the requiremnts. It is easy to see that

$$a_{n,0} = a_{n,00} + a_{n,01} = a_{n,001} + a_{n,01} = a_{n-2,1} + a_{n-1,1}.$$

Due to symmetry, $a_{n,w} = a_{n,\overline{w}}$, where $\overline{w}$ is the word obtained from w upon replacing each 0 by 1 and each 1 by 0. Therefore:

$$a_{n,0} = a_{n-2,0} + a_{n-1,0}.$$

The initial conditions are $a_{0,0} = a_{1,0} = 1$. It follows that $a_{n,0}$ is Fibonacci's sequence. Therefore the required probability is $\frac{2F_n}{2^n} = \frac{F_n}{2^{n-1}}$.

(d) If n is even the probability is $= \binom{n}{n/2}/2^n$, while if n is odd there are no sequences satisfying the condition, so that the probability is 0.

25. The sets A_1 and A_2 may be chosen in $2^n \cdot 2^n = 4^n$ ways altogether. To satisfy the condition $A_1 \bigcap A_2 = \emptyset$, we have to require that each $j \in \{1, 2, \ldots, n\}$ belongs to at most one of the sets A_1 and A_2. Thus we have 3 possibilities for each j, namely either $j \in A_1 \bigcap \overline{A}_2$ or $j \in \overline{A}_1 \bigcap A_2$ or $j \in \overline{A}_1 \bigcap \overline{A}_2$. Hence the number of possibilities satisfying the requirement is 3^n. It follows that the probability of the event in question is $(3/4)^n$.

30. Due to symmetry, all $3! = 6$ possible orderings of X_1, X_2 and X_3 are equi-probable, whence each has probability 1/6.

32. The number of possibilities for choosing the cards is $\binom{52}{13}$ (order does not matter). This constitutes the denominator for all parts.

(a) There are 4 possible full hands, so the probability is $4/\binom{52}{13}$.

(b) All the 13 cards should be chosen from the 48 non-ace cards: $\binom{48}{13}/\binom{52}{13} = \frac{39\cdot38\cdot37\cdot36}{52\cdot51\cdot50\cdot49}$.

(c) There are $\binom{4}{1}$ ways to choose a king, the same for a queen, and the other 11 should be chosen from the remaining 40 cards: $\binom{4}{1}\binom{4}{1}\binom{52-4-4-4}{11}/\binom{52}{13} = \frac{\binom{4}{1}^2\cdot\binom{40}{11}}{\binom{52}{13}}$.

(d) There are $\binom{4}{1}$ ways to choose each card, and hence the probability is $\binom{4}{1}^{13}/\binom{52}{13}$.

41.

(a) Since all the events in the union are disjoint, the probability is the sum of probabilities. Consequently:

$$P\left(\bigcup_{i=1}^{\infty}\left[\frac{1}{2i+1}, \frac{1}{2i}\right]\right) = \sum_{i=1}^{\infty}\left(\frac{1}{2i} - \frac{1}{2i+1}\right) = 1 - \ln 2$$

.

(b) For any n, the set in question is contained in the set of numbers whose infinite decimal expansion does not contain the digit 7 in any of the first n places. The latter set is clearly of probability $(9/10)^n$. Thus the probability of our set is less than $(9/10)^n$ for each n, and therefore it vanishes.

(c) As in the preceding part, the probability is 0.

43.

(a) Let us show that:

$$\limsup_{n\to\infty} A_n = [0,2], \qquad \liminf_{n\to\infty} A_n = [1/2,1].$$

Indeed, if $x \in [0,1]$, then $x \in A_n$ for each even n, while if $x \in [1,2]$, then $x \in A_n$ for each odd n, so that $\limsup_{n\to\infty} A_n \supseteq [0,2]$. On the other hand, if $x < 0$, then $x \notin A_n$ for any n, while if $x > 2$ then $x \notin A_n$ for $n > \frac{1}{x-2}$. This gives the inverse inclusion $\limsup_{n\to\infty} A_n \subseteq [0,2]$.
If $x \in [1/2,1]$, then $x \in A_n$ for each n, and in particular $\liminf_{n\to\infty} \supseteq [1/2,1]$. If $x < 1/2$, then $x \notin A_n$ for any odd n, while if $x > 1$ then $x \notin A_n$ for odd $n > \frac{1}{x-1}$. Therefore $\liminf_{n\to\infty} \subseteq [1/2,1]$.

(b) A point belongs to $\limsup_{n\to\infty} A_n$ if it belongs to infinitely many of the events A_n, which happens if and only if it belongs to the union $\cup_{i=k}^{\infty} A_i$ for each k. It follows that

$$\limsup_{n\to\infty} A_n = \cap_{k=1}^{\infty} \cup_{i=k}^{\infty} A_i,$$

which representation proves that $\limsup_{n\to\infty} A_n$ is an event. Similarly

$$\liminf_{n\to\infty} A_n = \cup_{k=1}^{\infty} \cap_{i=k}^{\infty} A_i,$$

which proves that $\liminf_{n\to\infty} A_n$ is an event.

46.

(a) The number of all subsets of A is of size 2^n. Thus, equivalently, we have to calculate the sum of those binomial coefficients $\binom{n}{k}$ with even k. Since the expression $\frac{1+(-1)^k}{2}$ takes the value 1 for even k and vanishes for odd k, we have:

$$\begin{aligned}\sum_{2|k} \binom{n}{k} &= \sum_{k=0}^{n} \frac{1+(-1)^k}{2} \binom{n}{k} \\ &= \frac{1}{2}\sum_{k=0}^{n} \binom{n}{k} + \frac{1}{2}\sum_{k=0}^{n} (-1)^k \binom{n}{k}\end{aligned}$$

$$= \frac{1}{2} \cdot 2^n + \frac{1}{2} \cdot (1-1)^n = 2^{n-1}.$$

Consequently the required probability is $\frac{1}{2}$.

(b) A simple calculation yields:

$$1 + \omega^k + \omega^{2k} = \begin{cases} 3, & 3|k, \\ 0, & 3 \nmid k. \end{cases}$$

Consequently:

$$\begin{aligned} \sum_{3|k} \binom{n}{k} &= \sum_{k=0}^{n} \frac{1 + \omega^k + \omega^{2k}}{3} \binom{n}{k} \\ &= \frac{1}{3} \left[2^n + (1+\omega)^n + (1+\omega^2)^n \right] \\ &= \frac{2^n + (-\omega^2)^n + (-\omega)^n}{3} . \end{aligned}$$

Hence the probability for $|R|$ to be divisible by 3 is

$$\frac{1}{3} \left[1 + \frac{(-\omega^2)^n + (-\omega)^n}{2^n} \right] .$$

Similarly, to find the probability for $|R|$ to be 1 modulo 3, we calculate:

$$1 + \omega^2 \omega^k + \omega \omega^{2k} = \begin{cases} 3, & k \equiv 1 (\bmod 3), \\ 0, & \text{otherwise.} \end{cases}$$

Hence

$$\begin{aligned} \sum_{k \equiv 1 (\bmod 3)} \binom{n}{k} &= \sum_{k=0}^{n} \frac{1 + \omega^2 \omega^k + \omega \omega^{2k}}{3} \binom{n}{k} \\ &= \frac{1}{3} \left[2^n + \omega^2 (1+\omega)^n + \omega (1+\omega^2)^n \right] \\ &= \frac{2^n + \omega^2 (-\omega^2)^n + \omega (-\omega)^n}{3} . \end{aligned}$$

and the probability is $\frac{1}{3} \left[1 - \frac{(-\omega^2)^{n+1} + (-\omega)^{n+1}}{2^n} \right]$.

(c)

$$P(|R| \equiv i (mod\ 4)) = \begin{cases} \frac{2^n + (1+i)^n + (1-i)^n}{2^{n+2}} & i = 0 \\ \frac{2^n - i(1+i)^n + i(1-i)^n}{2^{n+2}} & i = 1 \\ \frac{2^n - (1+i)^n - (1-i)^n}{2^{n+2}} & i = 2 \\ \frac{2^n + i(1+i)^n - i(1-i)^n}{2^{n+2}} & i = 3 \end{cases}$$

51.

(a) The problem is equivalent to the problem regarding $2n$ people, n with \$10 bills and n with \$5 bills, waiting in line to but tickets for a show, which has been solved in class. Hence the required probability is $\frac{1}{n+1}$.

(b) The number of legal expressions is the same as in the preceding part, namely $\frac{\binom{2n}{n}}{n+1}$,. However, this time the total number of possibilities is 2^{2n}. Thus the required probability is $\frac{\binom{2n}{n}}{(n+1)2^{2n}}$.

3 Conditional Probability

4 Discrete Distributions

87.

(c) For any $c > 0$ the values assumed by $p(x)$ are non-negative. The value of c is determined by the requirement that their sum be 1. First let us decompose the given rational function. Namely, we are looking for constants a, b and d for which:

$$\frac{1}{x(x+1)(x+2)} = \frac{a}{x} + \frac{b}{x+1} + \frac{d}{x+2} .$$

This gives:

$$a(x+1)(x+2) + bx(x+2) + dx(x+1) = 1.$$

Making the substitutions $x = 0$, $x = -1$ and $x = -2$ we obtain:

$$2a = 1, \qquad -b = 1, \qquad 2d = 1,$$

and therefore

$$a = \frac{1}{2}, \qquad b = -1, \qquad d = \frac{1}{2} .$$

Hence:

$$\sum_{x=1}^{\infty} \frac{1}{x(x+1)(x+2)} = \sum_{x=1}^{\infty} \left(\frac{1/2}{x} - \frac{1}{x+1} + \frac{1/2}{x+2} \right) = \frac{1/2}{1} - \frac{1}{2} + \frac{1/2}{2} = \frac{1}{4} .$$

Thus $c = 4$.

5 Expectation

96.

(b)

$$\begin{aligned} E(X) &= \sum_{k=0}^n 2^k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^n \binom{n}{k} (2p)^k (1-p)^{n-k} \\ &= (2p + (1-p))^n = (1+p)^n. \end{aligned}$$

(c)

$$\begin{aligned} E(X) &= \sum_{k=0}^n \sin k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^n \frac{e^{ik}-e^{-ik}}{2i} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \frac{1}{2i} \left(\sum_{k=0}^n \binom{n}{k} (pe^i)^k (1-p)^{n-k} \quad - \quad \sum_{k=0}^n \binom{n}{k} (pe^{-i})^k (1-p)^{n-k} \right) \\ &= \frac{1}{2i} \left((pe^i + 1 - p)^n \quad - \quad (pe^{-i} + 1 - p)^n \right). \end{aligned}$$

99.

(a) Let A_k, $1 \le k \le n$, denote the event whereby k is the largest number in the sample. Then

$$P(A_k) = \frac{k^n - (k-1)^n}{N^n},$$

and therefore

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

It follows that:

$$E(X) = \frac{1}{N^n} \left(N^{N+1} - \sum_{k=1}^{N-1} k^n \right) = N - \sum_{k=1}^{N-1} \left(\frac{k}{N} \right)^n.$$

(b) Write $E(X)$ in the form:

$$E(X) = N - N \sum_{k=0}^{N-1} \frac{1}{N} \left(\frac{k}{N} \right)^n.$$

The sum on the right hand side is a Darboux sum corresponding to the integral $\int_0^1 x^n dx$. Hence $E(X)$ behaves asymptotically as $N - \frac{N}{n+1}$.

(c) As n becomes large, all the terms in the sum on the right hand side of the expression for $E(X)$ tend to 0, and therefore $E(X)$ tends to N.

104.

(a) If the envelope we took contains an amount x, then it is in principle possible that the value x was selected from the distribution, in which case the other envelope contains an amount of $2x$, and it is possible that only $x/2$ was selected and we received the second envelope, in which case the other envelope is the first and so contains only $x/2$. Now the probability for the first event is (using conditional probabilities)

$$\frac{p(x)\cdot\frac{1}{2}}{p(x)\cdot\frac{1}{2}+p(x/2)\cdot\frac{1}{2}}=\frac{p(x)}{p(x)+p(x/2)}\,.$$

The probability in this case that $x/2$ was selected is therefore $\frac{p(x/2)}{p(x)+p(x/2)}$. Hence our expected prize if we switch to the other envelope is

$$\frac{p(x)}{p(x)+p(x/2)}\cdot 2x+\frac{p(x/2)}{p(x)+p(x/2)}\cdot\frac{x}{2}\,.$$

Therefore we should indeed switch if and only

$$\frac{p(x)}{p(x)+p(x/2)}\cdot 2x+\frac{p(x/2)}{p(x)+p(x/2)}\cdot\frac{x}{2}>x$$

(where in the case of equality we are actually indifferent between the two options). A routine simplification yields the equivalent condition $2p(x)>p(x/2)$.

(b) For the distribution in (i) we have the condition in (a) is satisfied for every possible value of x, and we should always switch. For the distribution in (ii), the condition is satisfied only for $x=2$, so we should switch only in this case. (Notice that in this latter case we are actually certain that the other envelope contains an amount of 4, as otherwise it would mean that the value 1 had been selected, which is impossible.)

106. Denote by D the distance between v_1 and v_2. Since $P(D=1)=p$:

$$E(D)\geq 1\cdot P(D=1)+2\cdot P(D\geq 2)=p+2(1-p)=2-p.$$

On the other hand, by the solution of Problem 50 we have

$$P(D\geq 3)\leq (n-2)\cdot\left(1-p^2\right)^{n-2},$$

and consequently:

$$\begin{aligned}E(D) &\leq 1\cdot P(D=1)+2\cdot P(D\geq 2)+n\cdot P(D\geq 3)\\ &\leq p+2(1-p)+n(n-2)\cdot\left(1-p^2\right)^{n-2}\underset{n\to\infty}{\longrightarrow}2-p.\end{aligned}$$

Thus $E(D) \underset{n\to\infty}{\longrightarrow} 2-p$.

110.

(Algorithm a) Denote by "success" the event that an n-tuple forms a permutation. The probability for this event is $p = \frac{n!}{n^n}$. The number of selections of n-tuples is distributed $G(p)$. Hence the expected number of selections is

$$\frac{1}{p} = \frac{n^n}{n!}\,.$$

Since each selection consists of n integers, the expected number of random integers by this algorithm is $\frac{n^{n+1}}{n!}$. , which is approximately $\sqrt{\frac{2\pi}{n}}e^n$.

(Algorithm b) Denote by X_i the number of steps required to obtain the i-th digit, $i = 1, \ldots n$. Clearly, $X_i \sim G(1-(i-1)/n)$, and therefore

$$E(X_i) = \frac{n-i+1}{n}\,, \qquad i = 1, \ldots, n.$$

The total number of selections is $X = \sum_{i=1}^{n} X_i$, and hence:

$$E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \frac{n-i+1}{n} = n\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}\right).$$

This time the expected number of selections is only about $n \log n$.

112. For $0 \le i \le n$, denote by X_i the diameter of the graph obtained after the ith stage of the construction process. With this notation, we have to bound $E(X_n)$ from below. Denote:

$$D_i = X_i - X_{i-1}, \qquad 1 \le i \le n.$$

Obviously $D_1 = 0$ and $D_2 = 1$. For $i \ge 3$ we have $D_i = 1$ if at the ith stage we connect the selected vertex to a vertex which is at a distance X_{i-1} from some vertex of the graph we have after the $(i-1)$st stage; otherwise – $D_i = 0$. Since the diameter after the $(i-1)$st stage is X_{i-1}, there are at least two vertices at that point satisfying this condition, so that

$$P(D_i = 1) \ge \frac{2}{i-1}\,.$$

Hence:

$$\begin{aligned} E(X_n) &= \sum_{i=1}^{n} E(D_i) \ge 0 + 1 + \sum_{i=3}^{n} \frac{2}{i-1} \\ &= 2(1 + 1/2 + 1/3 + \ldots + 1/(n-1)) - 1. \end{aligned}$$

6 Continuous Distributions

119. Obviously, X is distributed Cauchy.

7 Variance and Covariance

134. Let X denote the number of ones. Then $X = \sum_{i=1}^{n} X_i$, where $X_i = 1$ if the outcome of the ith roll is 1 and $X_i = 0$ otherwise. Let Y and Y_i, $1 \le i \le n$, be defined similarly for the sixes. Obviously, $X, Y \sim B(n, 1/6)$, so that:

$$E(X) = E(Y) = \frac{n}{6}\ .$$

Now

$$\begin{aligned} E(XY) &= E\left(\sum_{i=1}^{n} X_i \cdot \sum_{j=1}^{n} Y_j\right) \\ &= \sum_{i,j=1}^{n} E(X_i Y_j) = \sum_{i \ne j} E(X_i)E(Y_j) + \sum_{i=1}^{n} E(X_i Y_i) \\ &= n(n-1) \cdot \frac{1}{6} \cdot \frac{1}{6} + \sum_{i=1}^{n} 0 = \frac{n(n-1)}{6}\ , \end{aligned}$$

and therefore

$$\mathrm{Cov}(X, Y) = \frac{n(n-1)}{6} - \frac{n}{6} \cdot \frac{n}{6} = -\frac{n}{36}\ .$$

135.

(a) Obviously, $X \sim H(m, a, b)$, $Y \sim H(n, a, b)$, and therefore

$$E(X) = \frac{ma}{a+b}\ , \qquad V(X) = \frac{mab}{(a+b)^2}\left(1 - \frac{m-1}{a+b-1}\right),$$

and

$$E(Y) = \frac{na}{a+b}\ , \qquad V(Y) = \frac{nab}{(a+b)^2}\left(1 - \frac{n-1}{a+b-1}\right).$$

(b) Write $X = \sum_{i=1}^{n} X_i$, where $X_i = 1$ if the ith ball is white and $X_i = 0$ otherwise. Write $Y = \sum_{i=1}^{n} Y_i$, analogously for the second batch. Then

$$\begin{aligned} E(XY) &= E\left(\sum_{i=1}^{m} X_i \cdot \sum_{j=1}^{n} Y_j\right) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} E(X_i Y_j) = mn\frac{a(a-1)}{(a+b)(a+b-1)} , \end{aligned}$$

so that

$$\begin{aligned} \mathrm{Cov}(X,Y) &= mn\frac{a(a-1)}{(a+b)(a+b-1)} - \frac{ma}{a+b} \cdot \frac{na}{a+b} \\ &= -\frac{mnab}{(a+b)^2(a+b-1)} . \end{aligned}$$

(c) The covariance is negative since the more white balls there are in the first batch the less we should expect to have in the second.

8 Multi-Dimensional Distributions

9 Independence