1. (2 pts)
A derangement is a permutation with no fixed points, in other words, permutations in which no elements are in their original positions. ([http://en.wikipedia.org/wiki/Derangement](http://en.wikipedia.org/wiki/Derangement))

A derangement is denoted by \( !n \). It is defined by a recurrence relation:

\[
!n = (n-1)(!(n-1)+!(n-2))
\]

where \( !n \), known as the subfactorial, represents the number of derangements, with the starting values \( !0 = 1 \) and \( !1 = 0 \).

(a) Given a substitution cipher on six letters, how many of them are derangements?

(b) If a random permutation is chosen on six letters, what is the probability that the permutation is a derangement?

SOLUTION:

(a) Apply the recurrence definition of derangement:

- Base cases: \( !0 = 1 \) and \( !1 = 0 \)
- \( !2 = !1 + !0 = 1 \)
- \( !3 = 2(!2+!1) = 2 \)
- \( !4 = 3(!3+!2) = 9 \)
- \( !5 = 4(!4+!3) = 44 \)
- \( !6 = 5(!5+!4) = 265 \)

Another way of doing this is by using the approximation \( !n \approx \text{round}\left( \frac{n!}{e} \right) \) (refer to FN text, page 12 - example 11 for explanation). The answer is 265.

You can also directly compute derangement on wolframalpha.com. For example, \( !10 \) is http://www.wolframalpha.com/input/?i=720%(10!)

(b) \( !6 / 6! = 265/720 \) which is to three decimal places 0.368

Correct Answers:

- 265
- 265/720

2. (1 pt)

How many ways are there to factorize 30030 into two or more numbers that are greater than 1? Note that the prime factorization of 30030 = \( 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \).

hint: you can compute Stirling numbers of the second kind and Bell numbers on wolframalpha.com. For example,

- \( S(6,2) \) is http://goo.gl/XrOXVo
- \( B_5 \) is http://goo.gl/0xRRXB

SOLUTION:

The prime factors of 30030 are \( \{2, 3, 5, 7, 11, 13\} \). Any grouping of them corresponds to one way of factorizing 30030 into numbers greater than 2. For example, \( (2)(3,5)(7,11,13) \) corresponds to \( 2 \cdot 15 \cdot 1001 \); \( (2)(3)(5)(7)(11)(13) \) corresponds to \( 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \). Hence the problem reduces to counting all possible partitions of \( \{2,3,5,7,11,13\} \) into subsets. Note that we did not count the one partition that groups all prime factors together: \( (2,3,5,7,11,13) \), because this partition corresponds to 30030 itself, or 30030 \cdot 1 which is not valid because the factors are required to be greater than one.

Recall that the number of ways to partition a set of \( n \) elements into \( k \) subsets can be computed using the Stirling numbers of the second kind. It is defined by the formula:

\[
S(n,k) = \frac{1}{k!} \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} j^n
\]

In this question, we need to find

\[
\sum_{k=2}^{6} S(6,k) = S(6,2) + S(6,3) + S(6,4) + S(6,5) + S(6,6)
\]

= \( 31 + 90 + 65 + 15 + 1 \)

= 202.

An alternative is to use the Bell numbers, defined as

\[
B_n = \sum_{k=1}^{n} S(n,k)
\]

Clearly, our answer is \( B_6 - S(6,1) = B_6 - 1 = 203 - 1 = 202 \).

Correct Answers:

- 202
3. (5 pts)
Let us consider the random graph $G(8, 0.3)$. That is a random graph with 8 nodes, and each edge has probability 0.3 to be present. (part b-e does not depend on part a, so even if you do not understand part a, please still attempt part b-e.)

(a) The number of edges in this random graph is a random variable. We denote it by $X$. In the previous HW, we have calculated the expected value of the number of edges in this random graph, $E[X] = C(n, 2) \cdot p$. In this question, we compute the variance of $X$.

To do that, we first introduce a binary random variable $Y$ that has value 1 with probability 0.3, and has value 0 otherwise. It turns out that, $\text{VAR}[X] = C(n, 2)\text{VAR}[Y]$. Computing $\text{VAR}[Y]$ is easy. You should be able to do this by applying the definition of variance. When you have the result, answer:

$$\text{VAR}[X] = C(n, 2)\text{VAR}[Y] =$$

(b) Suppose Bob is a node in the graph, what is the expected number of his friends?

(c) Suppose Mary is also a node, what is the probability that Bob and Mary have exactly one common friend?

d) What is the probability that Bob and Mary have more than two common friends?

(e) What is the expected number of common friends of Bob and Mary?

SOLUTION:

(a) The number of edges $X$ in the random graph $G(n, p)$ follows a binomial distribution $X \sim \text{binom}(C(n, 2), p)$. The variance of this distribution is $\text{VAR}[X] = C(n, 2)p(1 - p)$.

(b) This is equivalent to the expected number of heads in $n - 1$ coin tosses. Hence $(n - 1)p$.

(c) There are $n - 2$ choices for the only common friend is. Then there is $p^2$ probability that this friend is connected to both of them, and $(1 - p^2)^{n-3}$ probability that no other nodes are connected to both of them. Hence the final answer is $(n - 2)p^2(1 - p^2)^{n-3}$.

(d) To address the constraint that they must have more than two common friends, we consider the complement: they have either zero, one or two common friends.

Suppose we toss two coins for each of the nodes other than Bob and Mary. The first coin determines whether this node is connected with Bob, and the second coin determines whether this node is connected with Mary. It is easy to see that this node is connected to both Bob and Mary with probability $p^2$. In other word, $p^2$ is the probability that this node is a common friend of Bob and Mary. Therefore, the probability that Bob and Mary have $k$ common friends is equivalent to the probability of attaining $k$ successes in $n - 2$ trials with each trial having “success” probability $p^2$. Denote the number of common friends by the random variable $X$, its probability follows the binomial distribution:

$$P(X = k) = C(n - 2, k)(p^2)^k(1 - p^2)^{n-2-k}$$

With the general formula above, it is easy to compute the three cases we need:

- Zero common friend: $P(X = 0) = C(n - 2, 0)(1 - p^2)^{n-2}$
- One common friend: $P(X = 1) = C(n - 2, 1)(p^2)(1 - p^2)^{n-3}$
- Two common friends: $P(X = 2) = C(n - 2, 2)(p^2)^2(1 - p^2)^{n-4}$

It follows that the probability of more than two friends is the complement of the sum of the three probabilities above.

(e) By definition of the expected value,

$$\mathbb{E}(X) = \sum_{k=0}^{n-2} P(X = k) \cdot k = \sum_{k=0}^{n-2} C(n - 2, k)(p^2)^k(1 - p^2)^{n-2-k}k = (n - 2)p^2$$

Correct Answers:

- $C([8], 2) \cdot [0.3] \cdot (1 - [0.3])$
- $[8-1] \cdot [0.3]$
- $[[8-2]] \cdot [0.3] \cdot (1 - [0.3]) \cdot ([8-3])$
- $[0.0118348119899997]$ 
- $[[8-2]] \cdot [0.3] \cdot 2$

4. (3 pts)
A substitution cipher has the following two-line notation:
Convert this two-line notation to cycle notation:

Encrypt the message NEGATIVE using this cipher (write your answer in one-line notation):

Decrypt the code LDDUAWWK using this cipher:

Important notation guidelines:
- For all cycle notations in this homework set, order the letters in each group in alphabetic order, and order the groups by the first letter in each group in alphabetical order.
- If a group has only one element in it (i.e. the element stays unchanged), do not include this group. For example, instead of (ABC)(D), answer (ABC).
- If you are asked to write cycle notation, in you answer include the parenthesis, and leave no spaces between letters. e.g. (ACD)(BEF)
- If you are asked to write one-line notation, do NOT include parenthesis, and also leave no spaces between letters. e.g. CDABFE
- Any system message other than “Correct” means your answer is wrong.

SOLUTION:

The cycle notation is (BZGJHPNELURQXIMWOKVT)(CS).

The encrypted code is LDJABMTD.

The decoded message is NEEDAMMO.

Correct Answers:
- (BZGJHPNELURQXIMWOKVT)(CS)
- LDJABMTD
- NEEDAMMO

5. (2 pts)
Given two permutations \( \pi_1 = 2143 \) (written in one-line notation) and \( \pi_2 = 4132 \) (written in one-line notation), compose \( \pi_2 \) with \( \pi_1 \). This is formally written as \( \pi_2 \circ \pi_1 \), namely, \( x \) will be mapped to \( \pi_2(\pi_1(x)) \)

- Write the composition in one-line notation:
- Write the composition in cycle notation (note if your answer is \( (1)(2)(3)(4) \), simply enter an empty parenthesis ()):

Important notation guidelines:
- For all cycle notations in this homework set, order the letters in each group in alphabetic order, and order the groups by the first letter in each group in alphabetical order.
- If a group has only one element in it (i.e. the element stays unchanged), do not include this group. For example, instead of (ABC)(D), answer (ABC).
- If you are asked to write cycle notation, in you answer include the parenthesis, and leave no spaces between letters. e.g. (ACD)(BEF)
- If you are asked to write one-line notation, do NOT include parenthesis, and also leave no spaces between letters. e.g. CDABFE
- Any system message other than “Correct” means your answer is wrong.

SOLUTION:

The one-line notation is 1423.

The cycle notation is (243).

Correct Answers:
- 1423
- (243)

6. (2 pts)
A box contains 20 lightbulbs, of which 9 lightbulbs are defective. If 10 lightbulbs are drawn from the box WITHOUT REPLACEMENT, what is the probability that no more than 2 of the lightbulbs drawn are defective?
(b) what is the probability that the fifth defective lightbulb that is drawn occurs at the last draw? (note: this part is independent with part a)

SOLUTION:

(a) We can have either 0, 1 or 2 defective bulbs. Let X be a random variable that represents the number of defective bulbs drawn. We need to find

\[ P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) \]

Recall the meaning of a hypergeometric distribution: if the total number of items is \( N \), and \( K \) of those are defective, the probability that a draw of \( n \) items contains exactly \( k \) defective items is given by a hypergeometric distribution:

\[ P(X = k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}} \]

In this question,

\[ P(X = 0) = \frac{C(9, 0)C(11, 10)}{C(20, 10)} = 5.95379852345797e-05 \]
\[ P(X = 1) = \frac{C(9, 1)C(11, 9)}{C(20, 10)} = 0.0026792093355608 \]
\[ P(X = 2) = \frac{C(9, 2)C(11, 8)}{C(20, 10)} = 0.032150512026673 \]

Hence the answer is the sum of probabilities above: 0.034889253474637.

(b) If the last lightbulb is the fifth defective drawn, then there must have been 5 normal lightbulbs and four defective ones in the previous 9 draws. Apply the formula for the hypergeometric distribution:

\[ P(X = 4) = \frac{C(9, 4)C(11, 5)}{C(20, 10)} = 0.346582519647535 \]

We then multiply this by the probability that the fifth lightbulb is defective. There are 20-9 bulbs left after the first 9 draws, and among them 5 are defective, so this probability is

\[ \frac{5}{20-9} = 0.454545454545455 \]

The complete answer is

\[ \frac{C(9, 4)C(11, 5) \cdot 5}{20-9} = 0.346582519647535 \cdot 0.454545454545455 = 0.157537508930698 \]

Correct Answers:

- [0.034889253474637]
- [0.157537508930698]

7. (3 pts)
On the plugboard of the Enigma Machine, one can put wires between pairs of letters to swap each pair.

(a) If the operator must use 10 wires, what is the total number of plugboard combinations?

\[ \binom{26}{10} \cdot \frac{26!}{10!10!} = 532985208200576 \]

(b) If the operator can use any number between 0 and 13 plugboard wires, what is the total number of plugboard combinations?

\[ \sum_{k=0}^{13} \binom{26}{k} \cdot \frac{26!}{(2!)^k k!} = 532985208200576 \]

SOLUTION:

(a) If the number of wires used is \( k \), to count the possible configurations we first select which set of \( 2k \) letters are connected by wires. That is \( C(26, 2k) \) choices. Then we count the number of ways to partition these \( 2k \) letters into \( k \) pairs. Normalize the permutation count by the fact that 1) the pairs (wires) are non-distinguishable and 2) the order of the two letters in each pair does not matter. This gives \( \frac{(2k)!}{(2!)^k k!} \) ways.

Hence the number of configurations with exactly \( k \) wires is

\[ C(26, 2k) \cdot \frac{(2k)!}{(2!)^k k!} \]

(b) The answer is the answer to (a) summed over \( k = 0 \cdots 13 \),

\[ \sum_{k=0}^{13} C(26, 2k) \cdot \frac{(2k)!}{(2!)^k k!} = 532985208200576 \]

An alternative way to think of each configuration of the wires as a partition of the set \( \{A,B,C,...,Z\} \) into unlabeled subsets of size two and subsets of size one. Each subset of size two corresponds to a pair of letters that are connected by a wire; and each subset of size one corresponds to a letter that is not connected.

(c) By definition, the expected value is \( \sum_{k=0}^{13} \frac{1}{14} C(26, 2k) \cdot \frac{(2k)!}{(2!)^k k!} = 532985208200576/14 \)

Correct Answers:

- \[ C(26, 2[10]) \cdot (\{2[10]\}!) / (2^\{10\} [10]! \} \]
- 532985208200576
- 532985208200576/14