Institut for Matematik og Datalogi
Syddansk Universitet, Odense

October 2, 2023
JBJ

DM551/MM851 – 1. Exam assignment

Hand in by Friday Nov 27, 2023 09:00.

Rules

This is the first of two sets of problems which together with the oral exam in January constitute the exam in DM551/MM851. This first set of problems may be solved in groups of up top three. Any collaboration between different groups will be considered as exam fraud. Thus you are not allowed to show your solutions to fellow students, not from your group and you may not discuss the solutions with other groups. On the other hand, you can learn a lot from discussing the problems with each other so you may do this to some extend, such as which methods can be used or similar problems from the book or exercise classes.

It is important that you **try to be as concise as possible but still argue so that the reader can follow your calculations and explanations.** You must use combinatorial arguments to solve the problems. For example in a counting problem it is not enough to generate all solutions and count them. **It is also not enough just to say that the solution follows from an example in the book or similar**. In such a case you should repeat the argument in your own words.

Remember that this (and the second set of assignment to follow later) counts as part of your exam, so do a good job and try to answer all questions carefully.

How to hand in your report

Your report, which should be written in Danish or English, must be handed in on It-slearning by Friday October 27 at 09:00

On the first page you must write your **name(s)** and the first 6 digits of your **CPR-number(s)**. **Do not write the last 4 digits!**.

Exam problems

Solve the following problems. **Remember to justify all answers.**

Problem 1 (6p)

You have been in Bilka and bought 17 distinct items. When you come home you look at the receipt and wonder how you can apply the tools from DM551 to the information on the receipt. You soon realize that you may be able to find an application of the pigeon hole principle to the info on the receipt. Prove that that, no matter how the 17 things listed on the receipt there will always be 5 items on the receipt (top to bottom) so that their price is either increasing (not smaller than the previous) or decreasing (not larger than the previous) in the order from top to bottom.

Let $c_1, c_2, \ldots, c_{17}$ be a given top to bottom ordering of the items. For each $i \in [17]$ denote by u_i (d_i) the length of the longest increasing (decreasing) subsequence starting with c_i and going down. If there was no subsequence of 5 items which either increases or decreases, then we have $u_i, d_i \in [4]$ for $i \in [17]$. Then, by the pigeonhole principle, there are indices $i < j$ such that $(u_i, d_i) = (u_j, d_j)$. Now, if item i is not more expensive than item j we have $u_i \geq u_j + 1 > u_j = u_i$, contradiction and if item i is more expensive than item j, then $d_i \geq d_j + 1 > d_j = d_i$, contradiction again. Thus there must exist a sequence as claimed.

Problem 2 (14p)

At a party there are 12 men and 8 women.

(a) How many different pairs (m, w), where m is a man and w is a woman, can one make?
There are 12 choices for m and 8 for w, so by the product rule, there are $12 \cdot 8 = 96$ different pairs.

(b) In how many ways can we form 8 pairs $(m_1, w_1), \ldots, (m_6, w_6)$ where m_i is a man, $m_i \neq m_j$ for $i \neq j$, w_i is a woman, $w_i \neq w_j$ for $i \neq j$ and the order of these pairs is not important (so that all permutations of the same 8 pairs count as one solution)? Hint: how many solutions are there for a fixed set of 8 distinct men?
There are $\binom{12}{8} = 495$ ways to choose 8 different men and for each such choice we can permute the 8 women to be matched with these in $8! = 40320$ ways. Hence the answer is $495 * 40320 = 19958400$.

(c) In how many ways can we arrange the 8 women in a circle if we consider two arrangements identical when each woman has the same two women next to her (so for example $w_1w_2w_3w_4w_5w_6w_7w_8w_1$ is the same as $w_1w_8w_7w_6w_5w_4w_3w_2w_1$)?
If we fix the position of woman w_1 there are $7! = 5040$ ways to permute the remaining 7 women but by the rule above, for each such permutation there is another that gives the same ordering. Hence the answer is $5040/2 = 2520$.

(d) Now consider a fixed cyclic ordering $w_{i_1}w_{i_2}w_{i_3}w_{i_4}w_{i_5}w_{i_6}w_{i_7}w_{i_8}w_{i_1}$ of the eight women. We want to place the men into the circle in such a way that no two women stand next to each other. In how many ways can this be done if we do not distinguish between the men? Hint: Compare with Exercise 6.5.48.
We know that there has to be at least one man between any consecutive women. After placing these there are 4 men left to distribute. Let x_j denote the number of extra men that we place between w_{i_j} and $w_{i_{j+1}}$, where $w_{i_9} = w_{i_1}$. Then we seek the number of integer solutions to $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 4$. This is $\binom{11}{7} = 330$.

Problem 3 (10p)

(a) Suppose we choose a random letter x from the string 'RECURRENCE' and a random letter y from the string 'RELATION'. What is the probability that

$x = y$?

(b) How many different permutations are there of the string 'RECURRENCE'?

Answer to (a): there are 3 different letters that occur in both strings, namely 'R','E','N'. To find the probability that $x = y$ we thus we need to find the probability that x, y are both 'R', 'E' or 'N'. We find

- $p(x = 'R' \wedge y = 'R') = \frac{3}{10} \cdot \frac{1}{8} = \frac{3}{80}$
- $p(x = 'E' \wedge y = 'E') = \frac{3}{10} \cdot \frac{1}{8} = \frac{3}{80}$
- $p(x = 'N' \wedge y = 'N') = \frac{1}{10} \cdot \frac{1}{8} = \frac{1}{80}$

Hence we get $p(x = y) = \frac{7}{80}$.
Answer to (b): according to Rosen Theorem 3 section 6.5 the answer is $\frac{10!}{3!3!2!}$, because there are 2 sets of 3 and one set of 2 indistinguishable letters.

Problem 4 (6p)

Prove that for all non-negative integers n we have

$$\sum_{k=0}^{n} \binom{n}{k} 17^k 3^{n-k} = \sum_{k=0}^{n} \binom{n}{k} 10^n$$

By the binomial formula we have

$$\sum_{k=0}^{n} \binom{n}{k} 17^k 3^{n-k} = (17+3)^n = (10+10)^n = \sum_{k=0}^{n} \binom{n}{k} 10^n$$

Problem 5 (10p)

Consider an experiment where we roll two dice once. Let X denote the minimum value of the two dice (so if we roll 3 and then 2, we have $X = 2$).

(a) What are the different values that X can take?

(b) What is the probability of X taking the different values? That is, find for all possible values r the quantity $p(X = r)$.

(c) Determine $E(X)$

Answer to (a): X can take the values 1,2,3,4,5,6.
Answer to (b):

- $p(X = 6) = \frac{1}{36}$ as both dice must roll a 6.
- $p(X = 5) = \frac{3}{36}$ as both must roll at least 5 and we have at least one 5.

- $p(X = 4) = \frac{5}{36}$ out of the $3 \cdot 3$ rolls where both are at least 4, 4 are bad (both above 4).
- $p(X = 3) = \frac{7}{36}$ again there are 16 rolls with both at least 3 and 9 of these have both above 3.
- $p(X = 2) = \frac{9}{36}$ out of 25 rolls with both at least 2, 16 have both above 2.
- $p(X = 1) = \frac{11}{36}$ out of the 36 rolls 25 have both above 1.

Answer to (c): $E(X) = \frac{1}{36}(6 \cdot 1 + 5 \cdot 3 + 4 \cdot 5 + 3 \cdot 7 + 2 \cdot 9 + 1 \cdot 11) = \frac{91}{36}$

Problem 6 (16p)

(a) Find the number of non-negative integer solutions to $x_1 + x_2 + x_3 = 15$
This is the same as the number of ways to choose 15 '*'s among 17 symbols that are all '*' or '|', so the answer is $\binom{17}{15} = 136$.

(b) Solve the problem above with the extra condition that $x_1 \geq 4, x_3 \geq 5$.
Now 9 of the 15 '*"s are already fixed so we need to solve $x_1' + x_2' + x_3' = 6$. There are $\binom{8}{6} = 28$ solutions.

(c) In how many ways can one distribute 15 identical balls into 3 distinct boxes such that box 1 contains at most 5 balls, box 2 at most 8 balls and box 3 at most 9 balls? Hint: use inclusion-exclusion.
Here we need to subtract the number of solutions in (a) which violate at least one of the bounds. Let P_i be the property that the bound for x_i is violated. Then $N(P_1) = \binom{11}{9} = 55$, $N(P_2) = \binom{8}{6} = 28$, $N(P_3) = \binom{7}{5} = 21$, $N(P_1P_2) = 1$ and all other values are zero. Hence we need to subtract $55 + 28 + 21 - 1 = 103$ so the answer is $136 - 103 = 33$.

(d) Explain why the following does not lead to the correct answer: The sum of the upper bounds is $22 = 5 + 8 + 9$ so 7 more than 15. Find the number of ways to distribute 7 balls in three boxes and return this as the answer. Here the distribution of the 7 balls would tell us how much to lower each upper bound so that we lower them by 7 in total.
There are $\binom{9}{7} = 36$ ways of selecting the balls to be taken out again. Some of these solutions (in fact exactly 3) will have x_1 larger than 5, so we would be left with a negative amount of balls in box 1.

Problem 7 (16p)

Suppose we have to form m committees, each with k persons, so that each committee represents k different skills from a set S of n skills (so $k \leq n$). The rule is that we must cover each skill by a fixed person and no person may be assigned to more than one of the skills (so there will be exactly n different persons covering the skills). We can see a committee X as a subset of S and the committees can overlap, that is, several committees may need a person with the same skill s and this person (the one who is assigned to skill s) will then belong to all those committees.

Suppose that we have many skilled employers both men and female so that we can can cover any subset $S' \subseteq S$ of the skills by different men and the remaining skills $S \setminus S'$ by different women.

Our task is now to analyse when we can form m committees, each with a prescribed sets of k skills, so that all of these committees have both a man and a woman.

Let us see an example: Suppose there are 4 skills, a, b, c, d and we want 3 committees with skills $C_1 = \{a, b, c\}$, $C_2 = \{b, c, d\}$ and $C_3 = \{a, c, d\}$, respectively. If we choose women w_1, w_2 for skills a and c and men m_1, m_2 for the skills b and d, then it is easy to check that all the three committees have both a man and a woman: committee C_1 will consist of both women and man m_1, committee C_2 will consist of both men and woman w_2 and finally committee C_3 will consist of both women and man m_2.

(a) In how many ways can we assign persons to the n skills if we just want to cover each skill by either a man or a woman?

(b) Prove, using the probabilistic method, that if the number m of committees we wish to form is less than 2^{k-1}, then there is always an assignment of men and women to the skills such that each of the m committees will have at least one man and at least one woman. Hint: Compare with the notes on Weekly note 3 (consider a random assignment of qualified men and women to the n skills).

- Answer to (a): From the information that we have enough men and women to cover all subsets by distinct men and the rest by distinct women, we get that the answer is the same as the number of subsets of an n-set, namely 2^n.

- Answer to (b): As in the hint we consider a random '2-colouring' of the skills by 'm','w'. For each committee $i \in [m]$ we let the indicator random variable X_i take the value 1 if the committee is covered only by men or only by women by our random assignment. As there are k skills in the committee we get that $p(X_i = 1) = \frac{1}{2^k} \cdot 2 = \frac{1}{2^{k-1}}$. This is also the expected value of X_i as X_i is an indicator random variable. Now the number of committees that are 'monochromatic' (that is, have only men or only women assigned to them) is given by $X = X_1 + \ldots + X_m$ and by linearity of expectation we get:

$$E(X) = E(\sum_{i=1}^{m} X_i) = \sum_{i=1}^{m} E(X_i) = \sum_{i=1}^{m} \frac{1}{2^{k-1}} = \frac{m}{2^{k-1}}$$

  So $E(X)$ is strictly less than 1 when $m < 2^{k-1}$. Now it follows from the first moment principle (or by Markov's inequality) that the probability that X takes the value 0 under the random assignment of men and women is strictly larger than 0 and hence there is a good assignment.

Problem 8 (22p)

This exercise is about Monte Carlo (MC) algorithms. You should start by recalling how the MC algorithm for the majority element problem (see Weekly note 4) works and how we can choose the parameter (the number of repetitions) so as to get the probability of a correct answer as close to 1 as we want (but still smaller than 1 of course).

Now consider the following variant of the majority element problem: we still have a set $S = \{x_1, x_2, \ldots, x_n\}$ consisting of n, not necessarily distinct numbers and we want to find out whether S contains two distinct numbers $x_i \neq x_j$ such that each of x_i, x_j occur at least $K = \lfloor \frac{n}{3} \rfloor + 1$ times in S. We call such a pair a **majority pair**.

Consider the following approach.

Repeat the following up to m times:

1. Pick a random index r and check whether x_r occurs at least K times in S.
2. If is the case, then delete all copies of x_r from S and call the resulting set S'; otherwise exit the loop (go to the next round).
3. Pick a random index t among the $|S'|$ indices of S' and check whether x_t occurs at least K times in S'.
4. If this is the case then return 'true' together with the majority pair (x_r, x_t); otherwise exit the loop (go to the next round).

If no majority pair was found in any of the m rounds above, then return 'false'

Let $\mathcal{A}$ denote the randomized algorithm that follows the strategy above

(a) Argue that $\mathcal{A}$ is always correct if it returns 'true'

(b) Argue that S can have at most one majority pair.

(c) Prove that when there is a majority pair in S, then the probability that $\mathcal{A}$ finds this pair in any execution of its loop is at least $\frac{1}{3}$. Hint: assume x, y is the unique majority pair. Define events E_1, E_2 so that E_1 is the event that $x_r \in \{x, y\}$ and E_2 is the event that $x_t \in \{x, y\} - \{x_r\}$. Then calculate a lower bound for $p(E_1 \cap E_2)$.

(d) What should the value of m be if we wish to ensure that the probability of S having no majority pair is at least 99% if $\mathcal{A}$ returns 'false'?

(e) Suppose that S does have a majority pair. What is the expected number of times we need to repeat the loop of $\mathcal{A}$ before we have found the pair ?

- Answer to (a): If $\mathcal{A}$ returns 'true', then x_r occurs at least K times in S and $x_t \neq x_r$ also occurs at least K times in S. Hence $\{x_r, x_t\}$ is a majority pair in S.
- Answer to (b): Suppose S had two distinct majority pairs $\{x, y\}$ and $\{x', y'\}$ As they are not equal, the set $\{x, y, x', y'\}$ has at least 3 elements which implies that $n = |S| \geq 3 * K = 3 * (\lfloor \frac{n}{3} \rfloor + 1) \geq n + 1$, contradiction.

- Answer to (c): Suppose $\{x, y\}$ is the majority pair in S. Following the suggestion in the hint, we define E_1, E_2 as above. Since E_1 occurs when x_r is either x or y which together have at least $2K > \frac{2n}{3}$ occurences in S, we get that $p(E_1) > \frac{2}{3}$. We next seek the probability that $x_t \in \{x, y\}$ given that $x_r \in \{x, y\}$ (otherwise we $\mathcal{A}$ would continue to the next round). We know that at least K elements have been deleted from S when we create S' so $|S'| \leq \frac{2n}{3} \leq 2K$. Suppose without loss of generality that we had $x_r = x$ so all occurrences of x have been deleted and S' contains all copies of y. As there are at least K copies of y in S', the probability that we get $x_t = y$ is at least $\frac{1}{2}$ and we conclude that $p(E_2|E_1) \geq \frac{1}{2}$. Now we get that the probability of $\{x_r, x_t\} = \{x, y\}$ is given by

$$p(E_1 \cap E_2) = p(E_2|E_1)p(E_1) \geq \frac{1}{2}\frac{2}{3} = \frac{1}{3}.$$

- Answer to (d): Suppose S has a majority pair $\{x, y\}$. By (d) the probability that $\mathcal{A}$ does not find this pair in any given round is at most $\frac{2}{3}$. As these can be seen as independent experiments, the probability the $\mathcal{A}$ does not find $\{x, y\}$ in any of its m rounds is at most $(\frac{2}{3})^m$. We want this to be at most 1% so we just need m to be at least $\log_{\frac{2}{3}}(\frac{1}{100})$. This is satisfied when $m \geq 12$.

- Answer to (e): We saw in (c) that if S has a majority pair, then $\mathcal{A}$ will find this with probability $p \geq \frac{1}{3}$ in any given round. Since the rounds are independent we can see that the process of waiting for the the first success follows a geometric distribution with success probability p. Hence we expect $\frac{1}{p} \leq 3$ repetitions before $\mathcal{A}$ finds the majority pair.