Fall 2015 Qualifier Exam:
Decision Science and Operations Research

September 21, 2015

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book the area of the exam (DSOR), your name, and the question answered in that book. On *one* of your books list the numbers of *all* the questions answered.
3. Return all answer books in the folder provided. Additional answer books are available if needed.
4. You are allowed *one* 8.5×11 sheet of paper with formulae.

SPECIFIC INSTRUCTIONS:

You must answer 5 out of 6 questions.

POLICY ON MISPRINTS AND AMBIGUITIES:

The Exam Committee tries to proofread the exam as carefully as possible. Nevertheless, the exam sometimes contains misprints and ambiguities. If you are convinced a problem has been stated incorrectly, mention this to the proctor. In any case, you should indicate your interpretation of the problem in your written answer. Your interpretation should be such that the problem is nontrivial.

1. Given rational numbers $c_j, j = 1, \ldots, n$, $b_i, i = 1 \ldots, m$, and $a_{ij}, i = 1, \ldots, m, j = 1, \ldots, n$, the primal form of a linear program is

$$\max \quad \sum_{j=1}^{n} c_j x_j \tag{PLP}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j \le b_i \quad \forall i = 1, \ldots, m$$

$$x_j \ge 0 \quad \forall j = 1, \ldots, n.$$

   (a) Write the dual linear program of (PLP).

   (b) State the weak duality theorem.

   (c) Prove the weak duality theorem you stated in part (b).

   (d) State the strong duality theorem.

   (e) State the complementary slackness theorem.

   (f) Use the strong duality theorem to prove the complementary slackness theorem.

---

2. Lazy Lane is a straight highway that extends for 10km. There are n people who live on Lazy Lane at locations $p_j, j = 1, \ldots, n$, where p_j is the distance in km from the beginning of Lazy Lane to the j^{th} person's house. There is one giant alarm clock on Lazy Lane that wakes up all of its residents at exactly the same time, at which time they immediately start walking to the bus stop to go to work. Person j walks to the bus stop at rate r_j and needs to arrive to the bus stop by time $t_j, i = 1, \ldots, n$. The alarm clock goes off exactly once per day. The Lazy Lane Bus Company (LLBC) is placing a single bus stop on Lazy Lane.

   (a) LLBC would like to know where to place the bus stop so as to allow the residents of Lazy Lane to sleep as late as possible. Formulate this problem as a linear program.

   (b) Now suppose that LLBC is going to place m bus stops on Lazy Lane. Person j will get to work on time if (s)he gets to *any* of the bus stops by time t_j. LoserLinderoth, Lord of Lazy Lane, loves Linear Programming. He claims that you can also model this decision problem as a linear program? Is LoserLinderoth correct? If so, model the problem as a linear program. If not, explain why.

---

3. Given a polyhedron $P = \{x \in \mathbb{R}^n : Ax \leq b\}$, we denote by $P_I := \text{conv.hull}(P \cap \mathbb{Z}^n)$ the integer hull of P, and by P' the Chvátal closure of P. Formally,

$$P' := \{x \in \mathbb{R}^n : ax \leq \lfloor \beta \rfloor, \text{ for every } ax \leq \beta \text{ valid for } P \text{ with } a \in \mathbb{Z}^n\}.$$

Moreover for every $i \in \mathbb{N}$ we denote by $P^{(i)}$ the i-th Chvátal closure of P, i.e. $P^{(0)} := P$ and $P^{(i)} := (P^{(i-1)})'$ for $i \geq 1$.

Prove or disprove:

(a) For every two polyhedra $P, Q \subseteq \mathbb{R}^n$ with $P \subseteq Q$, $P_I \subseteq Q_I$.

(b) For every two polyhedra $P, Q \subseteq \mathbb{R}^n$ with $P \subseteq Q$, $P' \subseteq Q'$.

Given a polyhedron $P = \{x \in \mathbb{R}^n : Ax \leq b\}$, we denote with $\mathcal{R}(A, b)$ the family of the polyhedral relaxations of P that consist of the intersection of the half-spaces corresponding to a family of linearly independent inequalities $\bar{A}x \leq \bar{b}$ of the system $Ax \leq b$. Prove or disprove:

(c) $P = \cap_{R \in \mathcal{R}(A,b)} R$;

(d) $P' = \cap_{R \in \mathcal{R}(A,b)} R'$;

(e) $P_I = \cap_{R \in \mathcal{R}(A,b)} R_I$.

---

4. Due to start-up problems, a pump in a nuclear-power plant has a 0.5 chance of failing during its first hour of operation. If the pump survives the first hour of operation, then the probability of surviving each succeeding hour of operation is 0.9.

The maintenance crew checks the condition of the pump at the end of each hour. If the pump is not working at the beginning of any given hour, there is a 0.7 chance that the crew will find the fault and have the machine ready to turn on again at the end of the hour, and a 0.3 chance that the pump will still not be working at the end of the hour.

(a) Formulate as a discrete-time Markov process, and draw the transition diagram. (Hint: Your process should have three states.)

(b) Is this process periodic? Explain why or why not.

(c) Determine the limiting-state probabilities of this Markov process. According to those results, what fraction of all hours is spent in maintenance (in the long run)?

(d) For a randomly selected hour, what is the joint probability that the pump will be not working at the beginning of that hour **and** transition to a working state by the end of the hour? What is the overall probability that the Markov process will transition from one state to another in a randomly selected hour?

(e) What is the distribution for the number of hours that the pump spends in maintenance before returning to a working state, for any given visit to the maintenance state?

(f) Now, assume that the probability that the maintenance crew will find the fault in the nth hour of a failure is not 0.7, but $0.7 + 0.3(1 - e^{-(n-1)})$, so that a repair becomes more and more likely the longer the pump has not been working. Can the resulting problem still be represented as a three-state Markov process? If not, please describe at least one way in which the problem could be modeled as a stochastic process (rather than just simulating it).

---

5. A local government facility maintains its power systems. Failures to the main system occur every F days, where F is exponentially distributed with an average of $1/\lambda = 100$ days. When such a failure occurs, the facility uses its backup generator to maintain power and immediately order a repair. The backup generator lasts for B days, where B is exponentially distributed with mean $1/\gamma = 2$, and the repair occurs after R days, where R is exponentially distributed with mean $1/\mu = 1.5$ days. A repair costs \$1500. If the repair does not occur before the backup generator fails, then the facility incurs a total failure cost of $\$14,000$

(a) Find the probability that the repair occurs before the backup generator fails.

(b) Find the *expected* cost, including possible repair and total failure costs.

(c) Suppose an expedited repair can be ordered at a cost of C, where the time until the expedited repair E is exponential with mean $1/\beta = 0.5$ days. Find the most you would be willing to pay for the expedited repair.

(d) Suppose that expedited repairs are no longer available, so the facility purchases a second backup generator that is used if the first backup generator fails. Both generators have independent and identically distributed failure times. The repair order is placed when a failure to the main system occurs (same as before), but now two backup generators must fail before a total failure occurs. Redo (b) in this scenario.

(e) Refer to (d). Find the probability that the expedited repair arrives while the second backup generator is still working.

---

6. A large parts supplier to the automobile industry wants to determine the proper number of dedicated retriever cranes and total cost for a planned Automated Storage and Retrieval System (AS/RS). [Each retriever crane only services one aisle of the AS/RS.] Some design parameters and assumptions are as follows:

    - The company needs 6,000 storage openings (52"w x 36"h x 48"d) in the AS/RS, and each opening costs $250.
    - Retriever cranes can travel at 300 ft/min (horizontal) and 150 ft/min (vertical).
    - The time to position and either place a pallet load into or take a load out of a storage opening is 4 seconds. The same positioning and movement times apply to load transfers at the input/output station of the AS/RS.
    - The maximum height of a storage rack is 65 feet.
    - The maximum length of any aisle is 500 feet.
    - Each retriever crane costs $225,000 with an annual maintenance cost of $30,000.
    - The company operates two 8-hours shifts per day.
    - The cost of a load waiting for service is estimated to be $35.00/hour.

    You have been asked by a company to develop a simulation model for this AS/RS before they purchase the AS/RS. Before they issue a purchase order to you for the cost of developing the model, they want you to address the following questions:

    (a) How would you model the arrival process for demands for cranes services (to either store and/or retrieve a load)? Address how the data would be collected and modeled.

    (b) How would you model the service process distribution given the physical size the planned AS/RS?

    (c) How would you use the simulation model to determine how many retriever cranes the company should plan for given that they wish to minimize the cost of the system?

    (d) What considerations and assumptions would you need to include?

Fall 2016 Qualifier Exam:
Decision Science and Operations Research

September 19, 2016

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book the area of the exam (DSOR), your name, and the question answered in that book. On *one* of your books list the numbers of *all* the questions answered.
3. Return all answer books in the folder provided. Additional answer books are available if needed.
4. You are allowed *one* 8.5×11 sheet of paper with formulae.

SPECIFIC INSTRUCTIONS:

You must answer 5 out of 6 questions.

POLICY ON MISPRINTS AND AMBIGUITIES:

The Exam Committee tries to proofread the exam as carefully as possible. Nevertheless, the exam sometimes contains misprints and ambiguities. If you are convinced a problem has been stated incorrectly, mention this to the proctor. In any case, you should indicate your interpretation of the problem in your written answer. Your interpretation should be such that the problem is nontrivial.

1. Suppose that

$$g(x) := \min_y c^T y \text{ subject to } Ay = b + Dx,\ y \geq 0$$

where $x \in \mathbf{R}^n$, $y \in \mathbf{R}^p$, c is a given p vector, A a given $m \times p$ matrix, b is a given m vector and D a given $m \times n$ matrix. Assume that

$$0 = \min_y c^T y \text{ subject to } Ay = 0,\ y \geq 0$$

and

$$\{z : z = Ay,\ y \geq 0\} = \mathbf{R}^m.$$

   (a) Give an example of a matrix A and a vector c that satisfy the assumptions.

   (b) For a given x, write down the dual of the problem defining $g(x)$.

   (c) Under the assumptions, show that $g(x)$ is finite for all $x \in \mathbf{R}^n$.

   (d) Under the assumptions, show that g is convex on $\mathbf{R}^n$.

   (e) What is another property that g has?

---

2. Let $D = (V, A)$ be a directed graph, and let $w : A \to \mathbb{R}$ be a weight vector. The weight of a subset B of A is defined as $w(B) := \sum_{a \in B} w_a$. Consider the problem of finding a maximum-weight subset $B \subseteq A$ such that no node of V is at the same time the head of an arc in B and the tail of another arc in B.

   (a) Formulate this problem as a $0, 1$ linear program.

   (b) Is the polyhedron defined by the natural linear programming relaxation of your $0, 1$ linear set integral? Provide a justification of your answer either way.

   (c) If the answer to part (b) is "no", give an additional class of inequalities, which is not implied by the inequalities of your natural linear programming relaxation, but which is valid for all its $0, 1$ solutions. (Hint: Think about certain directed cycles in D.)

---

3. The city of Sol, wanting to stay true to its name, is assessing the feasibility of using solar panels to provide power for its inhabitants during the warm summer months, when air conditioning costs are the highest. A multi-year survey of the inhabitants' energy habits as well as the hourly solar availability revealed the following aggregate data representing a typical summer day:

Hour of day:	1	2	3	$\cdots$	24
Hourly demand (MWh):	d_1	d_2	d_3	$\cdots$	d_{24}
Solar intensity:	s_1	s_2	s_3	$\cdots$	s_{24}

Unfortunately, demand doesn't always align itself with the solar cycle. For example, there is less sun in the evening, when the demand for energy is highest. The strategy is as follows:

- Purchase some number (N_p) of solar panels. Each solar panel costs C_p dollars and provides a maximum amount of MWh equal to the solar intensity. For example, 150 solar panels can provide up to $150s_1$ MWh during the first hour. Note that this is the *maximum* yield of the panels. We can always have the panels produce less if need be.
- Purchase some number (N_b) of batteries. Each battery costs C_b dollars and can be used to store up to 1 MWh of power, which may be used at a later time. Each battery has an hourly efficiency of 98%, so each hour the batteries lose 2% of their total stored energy. It is required that the batteries end the day with the same level of charge that they had when the day started.

Now, the problems:

(a) Write an optimization model (define and explain the variables, constraints, and objective) that would solve the problem of figuring out how many solar panels N_p and batteries N_b should be purchased so that total cost is minimized and the city will generate enough power to meet the hourly demands with a 5% buffer. In other words, we would like to be able to provide up to $1.05d_t$ of power at time t. What sort of optimization problem is this?

(b) Sol followed your recommendation and purchased N_b batteries and N_p solar panels. To reduce wear and tear on the batteries, Sol would like to operate their system in a way that minimizes the number of times the batteries are charged or discharged. This time, we will run our system with no buffer (we must exactly meet the hourly energy demands). How can this be accomplished? What sort of optimization problem is this?

4. (a) A wookiee and a droid decide to play a chess tournament. The droid is better than the wookiee; the droid has a 0.85 chance of winning each match. The two play until one player is ahead by three matches, when the losing player must buy the winning player dinner. What is the probability that the wookiee wins the tournament?

   Hint: use a Markov chain.

   (b) After playing chess, the wookiee takes the bus home. Buses arrive at a certain stop according to a Poisson process with rate λ. If the wookiee takes the bus from that stop it takes R time time units, measured from the time at which the wookiee enters the bus until the wookiee arrives at home. If the wookiee walks then it takes time W time units to walk home. The wookiee is impatient, especially after losing at chess. Suppose the wookiee waits T time units (a fixed value), and if a bus has not arrived by that time then the wookiee walks home. Find the expected time for the wookiee to arrive home from the time the wookiee arrives at the bus stop. Simplify your answer as much as possible.

---

5. Consider a single server queue where arrivals occur according to a Poisson process with rate λ and service times are exponentially distributed with parameter, μ. Some customers may be dissatisfied with the service they received and must be served again. Suppose that on completion of service, a customer is dissatisfied with probability $1 - \beta$, for some $0 < \beta < 1$, independent of whether that customer had been dissatisfied (one or more times) before. Subsequent service times on the same customer, if any, are also independent and exponentially distributed with parameter, μ.

   (a) Let S denote the total time spent by the server in service of a customer, until the customer is satisfied. Determine the distribution of S and estimate the mean of S.

   (b) Suppose the dissatisfied customers are served again immediately, until satisfied, and that the order of service is FCFS. What are the conditions for the queue to be stable? Under these conditions, find the expected total waiting time of a customer in the system.

---

6. Prof. Luedtke's doughnut selling business is growing, and so he has decided to start making his own doughnuts. He is considering two doughnut production processes, called A and B, and has created a discrete event simulation model to estimate the average number of doughnts that can be produced in one hour using these two processes. The data from five independent replications of the simulation on each of the processes is shown below:

	Number of Doughnuts	
Replication	Process A	Process B
1	175	160
2	142	121
3	143	130
4	149	143
5	132	128
sample mean	148.2	136.4

(a) Help Prof. Luedtke conduct an analysis to see if he can conclude one process is better than the other at the 95% confidence level. Your analysis should include the calculation of an apporpiate confidence interval. (You will need to use the table of critical values on the next page for this problem.)

(b) Prof. Luedtke would like to obtain a more precise estimate of the average number of doughnuts produced using Process A. Provide an estimate of how many total replications Prof. Luedtke should make if he wishes to know the average number of doughnuts produced by Process A to within plus or minus three doughnuts (at 95% confidence).

(c) Prof. Luedtke used common random numbers to generate the data in the replications above. More generally, let X_i^A and X_i^B be the observed values of the metric of interest for configuration A and configuration B of a system, respectively, under simulation replication $i = 1, \ldots, n$ (all replications are independent of each other). Use the fact that $\text{VAR}(X_i^A - X_i^B) = \text{VAR}(X_i^A) + \text{VAR}(X_i^B) - \text{COV}(X_i^A, X_i^B)$ to explain how using common random numbers helps when trying to determine if one configuration is bettter than the other.

(d) Discuss good implementation practices when attempting to use common random numbers in a discrete event simulation. (If it helps to make your discussion concrete, you may consider simulating a simple single-server queueing system having random customer inter-arrival times and random service times.)

Fall 2017 Qualifier Exam:
Decision Science and Operations Research

September 18, 2017

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book the area of the exam (DSOR), your name, and the question answered in that book. On *one* of your books list the numbers of *all* the questions answered.
3. Return all answer books in the folder provided. Additional answer books are available if needed.
4. You are allowed *one* 8.5×11 sheet of paper with formulae.

SPECIFIC INSTRUCTIONS:

You must answer 5 out of 6 questions.

POLICY ON MISPRINTS AND AMBIGUITIES:

The Exam Committee tries to proofread the exam as carefully as possible. Nevertheless, the exam sometimes contains misprints and ambiguities. If you are convinced a problem has been stated incorrectly, mention this to the proctor. In any case, you should indicate your interpretation of the problem in your written answer. Your interpretation should be such that the problem is nontrivial.

1.

Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ and $Q = \{x \in \mathbb{R}^n : Cx \leq d\}$ be two non-empty polyhedra.

(a) Write a linear programming formulation that solves the problem:

$$\min\{\|x - y\|_1 : x \in P, y \in Q\}$$

where $\|z\|_1 = \sum_{i=1}^{n} |z_i|$ is the 1-norm.

(b) Write the dual of the formulation you wrote in part (a).

(c) Justify that both the primal and dual problems have an optimal solution (you may use the strong duality theorem).

(d) Using the above primal/dual pair of linear programs, show that if $P \cap Q = \emptyset$, then there exists a vector $p \in \mathbb{R}^n$ such that $p^\top x < p^\top y$ for all $x \in P$ and $y \in Q$. [Hint: the vector p can be defined using an optimal dual solution.]

---

2. Suppose that women acquire human papillomavirus (HPV) (which can eventually lead to cervical cancer) according to a Markov chain. Most HPV infections resolve on their own, but some HPV infections lead to LSIL and HSIL, two preliminary disease stages before cancer is acquired. Each stage in the Markov chain represents one year. Women can also die from other causes (the Death state). The resulting Markov chain has six states with the following transition probability matrix:

P=

	Norm	HPV	LSIL	HSIL	Cancer	Death
Norm	0.94	0.05	0	0	0	0.01
HPV	0.5	0.39	0.09	0.01	0	0.01
LSIL	0.3	0.4	0.24	0.05	0	0.01
HSIL	0.01	0.05	0	0.92	0.01	0.01
Cancer	0	0	0	0	1	0
Death	0	0	0	0	0	1

You do NOT have to give numerical answers. Instead describe how you would solve for answers symbolically.

(a) Is this Markov chain ergodic? Explain your answer.

(b) Data suggests that 60% of patients start in the Norm state, 35% in the HPV state, and 5% in the LSIL state. What is the probability that a randomly selected patient will be in the HSIL state in (exactly) three years?

(c) Given that a patient is in the Norm state today, what is the probability that the patient will reach state the LSIL state at least once in the next three years?

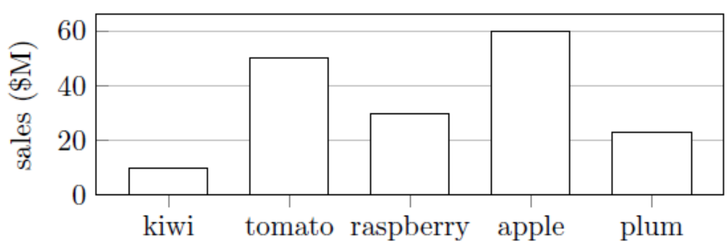
(d) Given that a patient starts in the Norm state, what is the expected time (in years) spent in the HPV state?

(e) Given that a patient starts in the Norm state, what is the probability that the HSIL state is ever reached?

3.

It is common knowledge that words/objects/entities have color associations. For example, *anger* is often associated with the color red. These associations are not one-to-one mappings, e.g. *strawberry* is also associated with the color red. The associations are not unique either; *apple* can be associated with red or green, and if we're talking about the company Apple Inc., the associations will be different still!

You are given a bar graph where each bar represents a different entity and your task is to choose colors to use for each of the bars. For example, the graph might look like the one below:



Your task is to choose colors for the bars in the graph so that each chosen color has a strong association with the category it represents. Suppose the labels for the bars in the graph are $\{b_1, \ldots, b_m\}$ and the colors at your disposal are $\{c_1, \ldots, c_n\}$. You have access to a dataset of color-category association strengths. The data is in the form of a table:

Category \ Color	c_1	c_2	$\cdots$	c_n
b_1	a_{11}	a_{12}	$\ldots$	a_{1n}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
b_m	a_{m1}	a_{m2}	$\ldots$	a_{mn}

So a_{ij} is association strength between category b_i and color c_j. We'll assume all the data are normalized so that $0 \leq a_{ij} \leq 1$ and $\sum_j a_{ij} = 1$. In other words, you can think of the i^{th} row of the table as a distribution over colors for the category b_i. We'll assume $n \geq m$, so there are more colors than categories.

(a) Suppose we would like to assign the colors to the categories in a way that maximizes the total association strength of all pairs. For example, if we associate b_1 with c_6 and b_2 with c_1, then the total association strength is $a_{16} + a_{21}$. Note: you cannot assign the same color to two different categories. Formulate this optimization problem as a linear program that depends on the data $\{a_{ij}\}$. Be sure to explain why your model is correct and describe the variables, constraints, and objective function.

(b) The approach of minimizing total association strength doesn't work as well when several categories all have similar color association profiles. Instead, we'll look for a way to assign colors to categories such that the chosen pair has a high association strength and the non-chosen pairs have a low association strength. To this effect, define a new objective for all i, j:

$$h_{ij} = a_{ij} - \tau \max_{k \neq i} a_{kj}$$

Here, $\tau \geq 0$ is a parameter and the max is taken over all $k \in \{1, \ldots, i-1, i+1, \ldots, m\}$. The net effect of using such an objective is that b_i should be strongly associated with c_j and at the same time, c_j should not be strongly associated with any of the other $b_k' s$ for $k \neq i$. How should you modify your linear program to account for this new objective?

(c) Picking a different τ in the formula for h_{ij} generally leads to a different optimal assignment of colors to categories. Prove that when $m = 2$ and $n = 2$ (two categories and two colors), all values of $\tau \geq 0$ lead to the same solution.

4.

Given an undirected graph $G = (V, E)$ and a weight function $w : E \to \mathbb{R}$ on the edges, a *matching* $M \subseteq E$ is a subset of pairwise disjoint edges of G (i.e., every node of G is contained in at most one edge of M). The weight $w(M)$ of a matching M is defined as the sum of the weights of the edges in M, namely

$$w(M) = \sum_{e \in M} w_e.$$

In this setting the *maximum weight matching problem* asks to find a matching in the graph with maximum weight.

(a) Explain how the maximum weight matching problem can be solved in polynomial time if G is bipartite.

The *greedy algorithm* for the maximum weight matching problem proceeds as follows:

- Set $M := \emptyset$.
- Set $A := E$.
- While $A \neq \emptyset$, do:
  - Let e be an edge in A with highest weight.
  - Add e to M.
  - Remove from A all edges adjacent to e.
- Return M.

(b) Show an example in which the greedy algorithm does not find a matching with maximum weight.

(c) We now consider a restricted version of the maximum weight matching problem in which the weights of all edges are 1, hence a maximum matching M is simply a matching with maximum cardinality $|M|$. Notice that the greedy algorithm in this case chooses an arbitrary edge from A in every iteration.

Let OPT be the cardinality of the optimal solution and let M_g be the output of the greedy algorithm. Show that $\frac{|M_g|}{OPT} \geq 0.5$ (In other words, the matching that the greedy algorithm finds is at least half the size of an optimum one).

(d) For every $n \in \mathbb{Z}_+$ give a graph with at least n vertices for which $\frac{|M_g|}{OPT} = 0.5$.

---

5. Jobs arriving to a queueing system from a single customer, in a Poisson manner with rate λ. They are served with mean service time T. The system has a maximum capacity of four jobs—including both any job being served, and those jobs in the queue. Any jobs arriving when the system is full will just be rejected, and never enter the queue.

a) If service times are exponential, and only one job can be handled at a time, please draw a diagram showing the possible states of this system, and write down equations for the steady-state probabilities. (You do not need to solve those equations.)

b) The owner of the facility is not convinced that exponential service times are a good assumption, since there are fewer extremely short service times than would be predicted by the exponential model. Please describe another way that the service times could be modeled. In particular, is there a way of modeling the service times that would still leave the resulting system satisfying the Markovian property?

c) To simplify the process, the owner has decided to accept the exponential model after all. However, she is concerned that the average time in queue may be too long. Rather than accept all jobs and have many of them experience long wait times, the owner is considering turning away half the potential jobs, to keep average wait times short for those she accepts.

If she rejects every other job, is the resulting system still Markovian? What is the distribution for the time between accepted jobs in this case?

If she rejects half of all jobs at random by flipping a coin, is the resulting system still Markovian? What is the distribution for the time between accepted jobs in this case?

d) Another change that the owner is considering is having her customer submit jobs in batches of four. In this case, jobs would be accepted only when the system is empty. If batches of four jobs arrive in a Poisson manner, is the resulting system still Markovian? Is it reasonable to expect that jobs would arrive in a Poisson manner if the customer was asked to send jobs in batches?

---

6. A local package delivery company is exploring new policies for pricing package deliveries based on weight and volume of the package and the number of zones the package will travel. The table below shows the values of these parameters for 5 packages. When answering parts (b) and (c), be sure to show all your intermediate calculations, so that we can see the process you used, and also so that, in case you make a calculation error, we can identify it and still give the majority of credit if the process is correct.

Package	Weight (kg)	Volume (L)	Number of zones
1	8	6	1
2	7	10	2
3	10	12	1
4	15	16	2
5	5	7	3

(a) Use an appropriate plot to assess whether the Weight and Volume of packages can be modeled as statistically independent random variables. Draw the plot (label the axes) and make the best conclusion you can from the data.

(b) The current pricing policy is to charge $P_1 = 2 + \min\{W, V\} * N$ dollars, where W is the weight, V is the volume, and N is the number of zones. Use the data above to construct a 95% confidence interval for the expected value of a package delivery when pricing using policy P_1.

(c) The company is considering a new pricing policy , $P_2 = 1 + \max\{W, V\} * N$ dollars for a package of weight W, volume V and travelign N zones. Conduct an appropriate confidence interval (at the 95% confidence level) to determine how the expected value of the price using policy P_2 compares to the expected value of the price using P_1. Comment on whether your analysis suggests a (i) the statistically significant difference, and (ii) a practically significant differnece, between the two policies.

Fall 2018 Qualifier Exam:
Decision Science and Operations Research

September 17, 2018

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book the area of the exam (DSOR), your name, and the question answered in that book. On *one* of your books list the numbers of *all* the questions answered.
3. Return all answer books in the folder provided. Additional answer books are available if needed.
4. You are allowed *one* 8.5×11 sheet of paper with formulae.

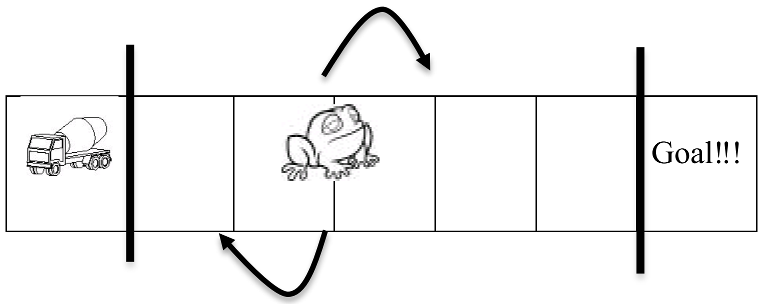
SPECIFIC INSTRUCTIONS:

You must answer 5 out of 6 questions.

POLICY ON MISPRINTS AND AMBIGUITIES:

The Exam Committee tries to proofread the exam as carefully as possible. Nevertheless, the exam sometimes contains misprints and ambiguities. If you are convinced a problem has been stated incorrectly, mention this to the proctor. In any case, you should indicate your interpretation of the problem in your written answer. Your interpretation should be such that the problem is nontrivial.

1. A new video game called Frogette tracks the activity of a frog crossing a 5 lane highway toward her goal in a series of hops. Frogette moves one lane closer to her goal to the right with a probability of 0.7, otherwise she moves one lane to the left with a probability of 0.3. Each lane crossing is independent. The game ends when she either reaches the finish line or ends up at the start line (where she gets run over by a semi-truck. Poor Frogette).



You do NOT have to give numerical answers. Instead describe how you would solve for answers symbolically using formulas.

(a) Formulate this as a Markov chain. What are the states? What is the transition probability matrix P?

(b) Is this Markov chain ergodic? Explain your answer.

(c) If she starts two lanes toward her goal (with 3 lanes to go), what is the probability she is three lanes toward her goal (one lane to the right in the picture above) in exactly three moves?

(d) If she starts two lanes toward her goal (with 3 lanes to go), what is the probability that she eventually reaches her goal on the right side of the highway?

(e) How many moves does she take, on average, before the game ends?

2. Assume that external virus attacks on a computer arrive according to a Poisson process with arrival rate λ. The probability that any particular virus attack will be successfully prevented by virus protection software is equal to P.

- Determine the probability that exactly k virus attacks will be successfully prevented in T time units.
- If we pick an interval of T time units, what is the probability that in that interval, there will be exactly R virus attacks that are successfully prevented and S virus attacks that are NOT prevented?
- Starting at time 0, a researcher studying computer security plans to observe the computer until at least one virus attack that is not prevented, and then a LATER attack that IS successfully prevented. Determine the expected amount of time that the researcher will have to observe the computer.
- Determine the probability mass function for the TOTAL number of virus attacks (regardless of whether they are successfully prevented) up to and including the third virus attack that is successfully prevented.
- Determine the expected value and probability mass function for the number of virus attacks that will be successfully prevented out of exactly N virus attacks in total.
- Estimate the arrival rate of virus attacks if the inter-arrival times between successive attacks (in days) are given as follows:

    10 days
    99 days
    15 days
    98 days
    20 days

A Simulation Question

Dumbledore has heard that you are a wizard of simulation, and he needs your help. Death Eaters are attacking Hogwarts, and he needs to build a simulation models of the attacks. As a first step, he needs to understand how often to expect an attack. Dumbledore enlisted Neville Longbottom to keep statistics about the number of attacks her hour. For the last 221 hours, Neville observed the frequency of attacks and collected the following data:

# Attacks	**Observed Frequency**
0 or 1	29
2 or 3	43
4 or 5	66
6 or 7	30
8 or 9	25
10 or 11	7

So 29 out of 200 times, either 0 or 1 attack occurred, 25 out of the last 200 hours, 8 or 9 attacks occurred, etc. Dumbledore believes that the Death-eaters attacks arrive according to a Poisson process. You will help him know if he can dispel this notion. Recall that the probability density function for a Poisson-distributed random variable is

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

Hermione Granger went to the restricted section of the Library, and retrieved 4 pages of tables that you may find useful when answering the following questions.

1. Explain what Dumbledore is assuming if the attack process is Poisson. Conversely, under what circumstances might a Poisson distribution for the number of attacks in an interval *not* be appropriate.

2. Estimate the average arrival rate parameter λ using the data given.

3. Employ an appropriate goodness of fit test to determine if one can reject the hypothesis that the arrival pattern is Poisson. Explain any estimates or approximations you are using, and what confidence level you have in your statement.

   In case the death-eaters have oblivated your memory, the Kolmogorov-Smirnov statistic is

   $$D = \max |F(x) - S_N(x)|,$$

   for the empirical distribution $S_N(x)$, and the and the chi-squared statistic is

   $$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i},$$

   where O_i and E_i are the observed and expected values in an interval.

4. **Linear Programming**

Recall the definitions of 1-norm and ∞-norm of a vector $x \in \mathbb{R}^n$:

$$\|x\|_1 := \sum_{i=1}^{n} |x_i|, \qquad \|x\|_\infty := \max\{|x_i| : i = 1, \dots, n\}.$$

Consider the following optimization problem:

$$\begin{aligned} \min\ & c^\top x \\ \text{s.t.}\ & \|Ax + b\|_1 \le 1. \end{aligned} \tag{1}$$

In this formulation, the decision variables are $x \in \mathbb{R}^n$, and the given data consists of $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.

(a) Formulate this problem as a linear program (LP) in inequality form and prove that your LP formulation is equivalent to problem (1).
Hint 1: You may use additional variables.
Hint 2: Recall that you can show that two maximization problems (A) and (B) are equivalent by showing: (i) For any feasible solution of (A) there is a feasible solution of (B) with objective value not lower, (ii) For any feasible solution of (B) there is a feasible solution of (A) with objective value not lower.

(b) Derive the dual LP, and show that it is equivalent to the problem

$$\begin{aligned} \max\ & b^\top z - \|z\|_\infty \\ \text{s.t.}\ & A^\top z + c = 0. \end{aligned} \tag{2}$$

What is the relation between the optimal value of z of problem (2) and the optimal value of the variables in the dual LP derived?

(c) Let x be feasible for (1) (i.e., $\|Ax + b\|_1 \le 1$) and let z be feasible for (2) (i.e., $A^\top z + c = 0$). Using only the weak duality theorem, what can you argue about the relation between $c^\top x$ and $b^\top z - \|z\|_\infty$?

5. **Optimization Modeling**

An *economy* consists of *sectors*. You can think of a sector as a process that consumes resources at the start of the year and produces other resources at the end of the year. We can also choose an *activity level* for each sector, which determines how much consumption and production happens in each sector.

Example. Here is an example with two sectors and three resources:

- Sectors: $\{\text{house-building}, \text{road-building}\}$
- Resources: $\{\text{wheat}, \text{brick}, \text{ore}\}$

At an activity level of 1, suppose we have the following:

- House-building consumes (1 brick, 1 ore) and produces (2 wheat, 2 brick, 2 ore).
- Road-building consumes (1 wheat, 1 brick) and produces (1 ore, 2 brick).

We can think of the consumption and production levels above as *rates*. To find the actual consumption and production, we multiply by the activity level. For example, if we choose an activity level of 100 for house-building and 50 for road-building in Year 1, then our economy behaves as follows:

- Total resources consumed by all sectors: (50 wheat, 150 brick, 100 ore)
- Total resources produced by all sectors: (200 wheat, 300 brick, 250 ore)

Every year, we must choose activity levels for the sectors, with the goal of making every sector *grow*. That is, we want the activity level of each sector to increase compared to the previous year. However, we cannot grow too fast: the consumption of a given year cannot exceed the production in the previous year. For example, if we decided to triple our activity levels for Year 2, this would require consuming (150 wheat, 450 brick, 300 ore). This is not possible because we produced an insufficient amount of brick and ore (300 and 250 respectively) in the previous year.

For this problem, we will look at optimizing growth over a two-year planning period. Specifically, we assume:

- There are m resources $i = 1, \dots, m$ and n sectors $j = 1, \dots, n$.
- For sector j, resource i is consumed at a rate b_{ij} and produced at a rate a_{ij}. These are fixed quantities known ahead of time.
- Sector j has activity level $x_j^{(1)}$ in Year 1 and $x_j^{(2)}$ in Year 2. These activity levels are things we must decide on.

- Year 1 activity levels are strictly positive and Year 2 activity levels are nonnegative.
- The consumption in Year 2 must not exceed the production in Year 1.

Define the *growth rate* of sector j as $x_j^{(2)}/x_j^{(1)}$. Our objective will be to maximize the *minimum growth rate*, which is the growth rate of the sector with the smallest growth rate. This ensures that every sector is growing. Finally, here are the problems:

(a) Formulate the above as an optimization problem. That is, specify the parameters, decision variables, constraints, and objective function.

(b) The formulation from Part (a) includes the strict inequalities $x_j^{(1)} > 0$ for $j = 1, \ldots, n$. Explain why strict inequalities are generally undesirable in an optimization model. Also explain how and why the strict inequalities can be replaced by $x_j^{(1)} \geq 1$ without any loss of generality.

(c) Suppose we want to know whether it's possible to achieve a minimum growth rate of r. Explain how the problem of Part (a) can be reformulated as a linear program where r appears as a parameter.

(d) Consider maximizing the *total annual growth* $\left(\sum_{j=1}^n x_j^{(2)}\right)/\left(\sum_{j=1}^n x_j^{(1)}\right)$ instead. In this scenario, we allow activity levels in Year 1 to be zero, so long as the total activity in Year 1 is strictly positive. Formulate this problem as a linear program.

6. **Integer Optimization**

Let $S = \{v^1, \ldots, v^T\} \subseteq \mathbb{R}^n$ be a finite (possibly huge) set of points and consider the optimization problem:

$$z^* = \min\ c^\top x \tag{3}$$
$$\text{s.t. } Ax \geq b \tag{4}$$
$$x \in S \tag{5}$$

where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$. Assume that this optimization problem has a feasible solution, so that z^* is finite. For $\lambda \in \mathbb{R}^m_+$, define:

$$z(\lambda) = \lambda^\top b + \min\ (c^\top - \lambda^\top A)x$$
$$\text{s.t. } x \in S$$

Finally, define:

$$z^{LD} = \max\{z(\lambda) : \lambda \in \mathbb{R}^m_+\}$$

The number of points that will be allocated to each part when grading are given in brackets at the beginning of each part.

(a) [2 pts] Show that $z(\lambda) \leq z^*$ for any $\lambda \in \mathbb{R}^m_+$.

(b) [4 pts] Recall that conv(S) is notation for the convex hull of S. Show that

$$z^{LD} = \min\ c^\top x$$
$$\text{s.t. } Ax \geq b$$
$$x \in \text{conv}(S).$$

[Hint: Start by formulating the problem $\max\{z(\lambda) : \lambda \in \mathbb{R}^m_+\}$ as a linear program, possibly by adding additional decision variable(s).]

(c) [2 pts] Provide an example where $z^{LD} < z^*$. [Hint: this can be done with $n = 1$ and a set S containing two points.]

(d) [1 pt] Suppose $S = \{x \in \mathbb{Z}^n_+ : Dx \geq d\}$. Use the result from part (b) to show that the $z^{LD} \geq z^{LP}$, where z^{LP} is the optimal value of the linear programming relaxation of (3)-(5):

$$z^{LP} = \min\ c^\top x$$
$$\text{s.t. } Ax \geq b$$
$$Dx \geq d$$
$$x \in \mathbb{R}^n_+$$

(e) [1 pt] Using the definition of S from part (d), suppose that the matrix D is totally unimodular and d is an integer vector. Again using results from previous parts, argue that in this case $z^{LD} = z^{LP}$.

Fall 2019 Qualifier Exam:
Decision Science and Operations Research

September 16, 2019

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book the area of the exam (DSOR), your name, and the question answered in that book. On *one* of your books list the numbers of *all* the questions answered.
3. Return all answer books in the folder provided. Additional answer books are available if needed.
4. You are allowed *one* 8.5×11 sheet of paper with formulae.

SPECIFIC INSTRUCTIONS:

You must answer 5 out of 6 questions.

POLICY ON MISPRINTS AND AMBIGUITIES:

The Exam Committee tries to proofread the exam as carefully as possible. Nevertheless, the exam sometimes contains misprints and ambiguities. If you are convinced a problem has been stated incorrectly, mention this to the proctor. In any case, you should indicate your interpretation of the problem in your written answer. Your interpretation should be such that the problem is nontrivial.

1. **Poisson Processes**

Ice cream production at Babcock Hall can be modeled as a Poisson process. Each day, the ice cream makers make three flavors of ice cream: vanilla, chocolate, and blue moon. They start all three recipes at the same time and make one batch of each kind per day.

- The time to make a batch of vanilla or chocolate is distributed exponentially with a mean of 2.5 hours.
- Blue moon is more difficult to make. The time to make a batch is distributed exponentially with a mean of 3 hours.

(a) Find the probability that vanilla is the first batch to finish.

(b) Find the probability that blue moon is the last batch to finish.

(c) Lets say that there is an equipment failure and the ice cream makers can only make two flavors at a time on parallel. They decide to start with making vanilla and blue moon at the same time. When one finishes, they make chocolate on the equipment that is freed up. Find the probability that blue moon finishes last.

2. **Markov Processes**

A factory has two redundant pieces of equipment. When either one is working, the factory is operational. If at the beginning of a week both machines are working, then at the end of the week exactly one will be working with probability 0.3; neither will be working with probability 0.4. However, if only one machine is working at the beginning of a week, the probability that it will still be working at the end of the week is 0.4.

A repairman checks once a week to determine whether the factory is operational. If not, it takes him a week to fix both broken machines.

(a) What fraction of all weeks is spent in repair (in the long run)?

(b) How many repairs do we expect will be performed on average during a period of n weeks (in the long run)?

(c) Why might your answers to parts (a) and (b) above not apply if the machine starts in perfect working order, and we are interested in its performance during the first n weeks (rather than in the long run)? What methods could be used to characterize the short run behavior of the machine?

(d) Would this system still be Markovian if the transition probabilities were uncertain (e.g., described by probability distributions) rather than known constants? Why or why not? Explain your reasoning.

3. **Simulation**

(a) Consider an unbounded continuous random variable X having cumulative distribution function (cdf) F. Suppose you wish to generate observations of a random variable Y, which has the distribution of X but is *truncated* to the interval $[a, b]$. Specifically, Y has the cdf given by $F_Y(y) = 0$ for $y < a$,

$$F_Y(y) = \frac{F(y) - F(a)}{F(b) - F(a)}, \quad a \leq y \leq b,$$

and $F_Y(y) = 1$ for $y > b$. Assume that you have an algorithm that can efficiently generate independent and identically distributed random outcomes having the distribution of X.

i. Derive an acceptance/rejection type algorithm for generating a random variate having distribution of Y using the method for generating observations of X, and demonstrate that the method is correct. (I.e., state the algorithm, and show that the output has the correct distribution.)

ii. What is the expected number iterations the acceptance/rejection method requires?

iii. Consider the following routine:
Step 1. Generate U uniformly in $[0, 1]$ and let $V = F(a) + U[F(b) - F(a)]$.
Step 2. Let $Y = F^{-1}(V)$.
Show that the output of this routine has the distribution of Y.

iv. Discuss the potential trade-off in computation effort between the algorithm in part (i) and that in part (iii).

(b) Consider a nonterminating simulation of a service queueing system, producing a sequence of observed waiting times $X_1, X_2, X_3, \ldots$. You are interested in estimating the value of $\mu = E[X_i]$, for the system once it is in steady state (at which point the value $E[X_i]$ is independent of i). You have estimated that the system appears to be roughly in steady state after L observations.

i. Describe how you would estimate and construct a confidence interval for μ using the replication/deletion approach. (Your description should include mathematical descriptions as needed to be clear.)

ii. Describe how you would estimate and construct a confidence interval for μ using the batch means approach. (Your description should include mathematical descriptions as needed to be clear.)

iii. Assuming you simulate roughly the same number of observations of X_i in the two approaches, describe one potential advantage of the replication/deletion approach over the batch means approach, and vice versa one potential advantage of the batch means approach over the replication/deletion approach.

4. **Optimization Modeling**

You have been tasked with determining how to best pickup crates of books from five different library locations and return them all to the central library. In this model, the central library is located at the origin of an xy-grid, and the (x, y) coordinates of the other 5 locations are given in the table below, along with how many crates of books are at each location.

Location	**# Crates**	**x-coordinate**	**y-coordinate**
A	9	-2	1
B	6	2	2
C	4	3	0
D	3	2	-2
E	5	-1	-1

For simplicity, we will measure the distances between cities using the "Manhattan"-metric (or the ℓ_1 distance) where the bookmobile can only go north-south or east-west on the grid. For example, the distance between locations A and B is

$$d_{AB} = |-2-2| + |1-2| = 5.$$

The bookmobile starts the day at the central library and can carry at most 12 crates of books. Further, if the bookmobile goes to a location, it *must* take all of the books at that location. Thus, if the bookmobile goes to location A, it either has to return to the central library or go to location D, all other locations have too many books.

(a) Create an integer programming instance that will minimize the total distance the bookmobile has to travel to return all of the crates of books to the library. (Note that just solving the problem by inspection, or giving a model that relies significantly on inspection, will not get very much credit. We are looking for a model that would work even if there are many more locations.)

(b) Modify your model from item (a) to maximize the number of crates of books returned to the library if the bookmobile can only make two trips.

(c) Take the dual of the linear programming relaxation of the model that you built in item (a).

(d) Now suppose you are given a set of N locations with coordinates (x_i, y_i) and a number of crates w_i $\forall i = 1, 2 \ldots, N$. The capacity of the bookmobile is Q. Describe how to modify your instance from item (a) for this general case. What is the size of your instance? If $N > 100$, how might you go about solving an instance of the model?

5. **Linear Programming**

You may use the following version of Farkas' Lemma in this problem:

Theorem 1 (Farkas' Lemma)**.** *For a linear inequality system* $\Sigma\colon \{Ax \le b\}$*, either* $Ax \le b$ *is feasible, or there exists* $y \in \mathbb{R}^m$ *such that* $y \ge 0$*,* $y'A = 0$ *and* $y'b = -1$*, but not both.*

Let $\Sigma : \{Ax \le b\}$, where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, be an infeasible system of linear inequalities. An infeasible subsystem Σ' of Σ is an *Irreducible Infeasible Subsystem (IIS)* if every proper subsystem of Σ' is feasible. By answering the question in this problem you will prove the following result, that relates the IISs of Σ to the extreme points of a given alternative polyhedron. Recall that the support of a vector x, denoted by $\mathrm{supp}(x)$, is the set of indices of the nonzero components of x.

Theorem 2. *The IISs of* Σ *are in one-to-one correspondence with the extreme points of the polyhedron*

$$P = \{y \in \mathbb{R}^m : y \ge 0, y'A = 0, y'b = -1\}.$$

Precisely, the support of any extreme point of P *indexes the inequalities of an IIS.*

(a) Consider a polyhedron $P \subseteq \mathbb{R}^n$. Give the definitions of: (i) extreme point; (ii) vertex; and (iii) basic feasible solution.

(b) The three definitions above are equivalent. State a necessary and sufficient condition for a polyhedron to contain at least a vertex.

(c) Let $\Sigma_1 : \{A_1x \le b_1\}$ be an IIS of Σ, and suppose wlog that A_1 and b_1 consist of the first m_1 rows of A and b, respectively. Let $P_1 = \{w \in \mathbb{R}^{m_1} : w \ge 0, w'A_1 = 0, w'b_1 = -1\}$. Show that there exists a vertex y of P such that $\mathrm{supp}(y) = \{1, \ldots, m_1\}$ by following the steps below:

 i. Prove that P_1 is nonempty, i.e. it contains a point $\bar{w}$.

 ii. Prove that all vectors in P_1 are strictly positive, i.e. $w > 0 \;\forall\, w \in P_1$.

 iii. Prove that P_1 contains a vertex, and that $\bar{w}$ is a vertex of P_1. In fact, $P_1 = \{\bar{w}\}$.

 iv. Construct a vertex y of P such that $\mathrm{supp}(y) = \{1, \ldots, m_1\}$.

(d) Let y be a vertex of P and assume wlog that $\mathrm{supp}(y) = \{1, \ldots, m_1\}$. Show that $\Sigma_1 : \{A_1x \le b_1\}$ is an IIS of Σ, where A_1 and b_1 consist of the first m_1 rows of A and b, respectively. *Hint: First, prove that* Σ_1 *is an infeasible subsystem of* Σ*. Then to prove that* Σ_1 *is* irreducible *proceed by contradiction: assuming it is reducible* (i) *show that there exists* $u \in P$ *with supp*$(u) \subset s(y)$*;* (ii) *prove that* $y - u$ *is a feasible direction at* y*, and use it to show that* y *is not an extreme point of* P*.*

6. **Integer Optimization**

For every $j \in \{0, \dots, n\}$, let S_j be the set of vectors $x \in \mathbb{R}^n$ such that j components of x are 1/2 and the remaining $n - j$ components are equal to 0 or 1. For example, the set S_1^2 is defined by

$$S_1^2 := \{(1/2, 0), (1/2, 1), (0, 1/2), (1, 1/2)\}.$$

(a) (1 point) Show that, for every $j \in \{0, \dots, n-1\}$, any vector $v \in S_{j+1}$ is the convex combination of two vectors in S_j.

(b) Given $j \in \{0, \dots, n-1\}$, consider any $\pi \in \mathbb{Z}^n$, $\pi_0 \in \mathbb{Z}$ such that $\pi x < \pi_0$ for every $x \in S_j$ (note the strict inequality). Show that $\pi x \leq \pi_0 - 1$ for every $x \in S_{j+1}$. To prove this result, let $v \in S_{j+1}$ and consider the following two cases:

(b1) (1 point) First prove $\pi v \leq \pi_0 - 1$ when $\pi v \in \mathbb{Z}$. [Hint: Use (a).]

(b2) (3 points) Then prove $\pi v \leq \pi_0 - 1$ when $\pi v \notin \mathbb{Z}$. [Hint: Without loss of generality, assume that $v_1 = 1/2$ and $\pi_1 \neq 0$. Then write v as the convex combination of two vectors $v^1, v^2 \in S_j$. Show that, for $i = 1, 2$, each of the line segments $[v, v^i]$ contains a point $\tilde{v}^i$ such that $\pi \tilde{v}_i \in \mathbb{Z}$. Deduce that $\pi v \leq \pi_0 - 1$.]

(c) For any $n \geq 1$, let

$$P := \left\{ x \in \mathbb{R}^n : 0 \leq x \leq 1, \quad \sum_{j \in J} x_j + \sum_{j \notin J} (1 - x_j) \geq 1/2 \quad \forall J \subseteq \{1, 2, \dots, n\} \right\}.$$

(c1) (1 point) Show that $P \cap \mathbb{Z}^n = \emptyset$.

(c2) (1 point) Show that $S_1 \subset P$.

(c3) (3 points) Use what you have proven so far to show that the Chvátal rank of P is at least n. *Hint: Use the observation that every Chvátal cut for a polyhedron Q can be written in the form $\pi x \leq \pi_0 - 1$, where $\pi \in \mathbb{Z}^n$, $\pi_0 \in \mathbb{Z}$ and $\pi x < \pi_0$ for every $x \in Q$.*