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:

Replication	Number of Doughnuts 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.)