MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.436J/15.085J Fall 2018
Problem Set 3

Readings:
Notes from Lecture 4 and 5.

Supplementary readings:
[GS], Sections 3.1-3.7.
[C], Section 2.1
[BT], for background on counting (which we don't cover) go through the last section of Ch. 1 of [BT], available at:
http://athenasc.com/Prob-2nd-Ch1.pdf.

Exercise 1. Suppose that $X_1, X_2, \ldots, X_n, \ldots$ are random variables defined on the same probability space. Show that $\max\{X_1, X_2\}$, $\sup_n X_n$, and $\limsup_{n\to\infty} X_n$ are random variables, using only Definition 1 in the notes for Lecture 4, and first principles, without quoting any other known facts about measurability.

Exercise 2. Given a distribution function F_X show that $F_X(x)$ is continuous at x_0 if and only if $\mathbb{P}(X = x_0) = 0$.

Exercise 3. The probabilistic method. A party of $n = 20$ people is gathered. A host selects some of the $n(n-1)/2$ pairs of people and introduces them to each other. Show that the host can do the introduction in such a way that for every group of 7 people there are at least two who are introduced to each other and there are at least two who are not.

Hint: Consider introducing people randomly and independently.

Note: The probabilistic method is a general method for proving existence: if you can prove that a randomly selected structure has certain desired properties with some positive probability (no matter how small), then a structure with these properties is guaranteed to exist.

Exercise 4. Let X be a nonnegative integer random variable. Show that

$$\mathbb{E}[X] = \sum_{n=1}^{\infty} \mathbb{P}(X \geq n).$$

Be careful in citing whatever results from the lecture notes are needed to justify the steps in your derivation.

Exercise 5. Let F_1 and F_2 be two CDFs, and suppose that $F_1(t) < F_2(t)$, for all t. Assume that F_1 and F_2 are continuous and strictly increasing. Show that there exist random variables X_1 and X_2, with CDFs F_1 and F_2, respectively, defined on the same probability space such that $X_1 > X_2$. *Hint:* Think of simulating X_1 and X_2 using a common "random number generator".

Exercise 6. Let $\{X_n\}$ be a sequence of independent non-negative random variables. Show that sequence X_n is almost surely bounded if and only if $\sum_{n=1}^{\infty} \mathbb{P}(X_n > c) < \infty$ for some c. (Hint: X_n a.s. bounded simply means $\mathbb{P}(\sup_n X_n = \infty) = 0$.)

Exercise 7. (Another application of the probabilistic method.) Let G be an undirected graph with neither loops nor multiple edges, and write d_v for the degree of vertex v (i.e., the number of edges incident on v). An independent set is a set of vertices no pair of which is joined by an edge. Let (G) be the size of the largest independent set of G. Use the probabilistic method to show that $(G) \geq \sum_v 1/(1 + d_v)$. *Hint:* Order the nodes at random, and examine the nodes one at a time, putting them in the independent set as long as there are no conflicts with previously examined nodes. Find the expected value of the resulting set.

In addition, you need to be sure that you can solve elementary problems. As a check, make sure you are able to solve the next problem (not to be handed in).

Drill problem: At his workplace, the first thing Oscar does every morning is to go to the supply room and pick up one, two, or three pens with equal probability 1/3. If he picks up three pens, he does not return to the supply room again that day. If he picks up one or two pens, he will make one additional trip to the supply room, where he again will pick up one, two, or three pens with equal probability 1/3. (The number of pens taken in one trip will not affect the number of pens taken in any other trip.) Calculate the following:

(a) The probability that Oscar gets a total of three pens on any particular day.

(b) The conditional probability that he visited the supply room twice on a given day, given that it is a day in which he got a total of three pens.

(c) $\mathbb{E}[N]$ and $\mathbb{E}[N \mid C]$, where $\mathbb{E}[N]$ is the unconditional expectation of N, the total number of pens Oscar gets on any given day, and $\mathbb{E}[N \mid C]$ is the conditional expectation of N given the event $C = \{N > 3\}$.

(d) $\sigma_{N|C}$, the conditional standard deviation of the total number of pens Oscar gets on a particular day, where N and C are as in part (c).

(e) The probability that he gets more than three pens on each of the next 16 days.

(f) The conditional standard deviation of the total number of pens he gets in the next 16 days given that he gets more than three pens on each of those days.