18.440 Problem Set 10

Due in class Friday December 5; late work will not be accepted. You can discuss problems with anyone, but you should write solutions entirely on your own.

1. (30 points) Markov's inequality says that if X is a non-negative random variable and $a > 0$, then

$$P\{X \geq a\} \leq E[X]/a.$$

a) Give an example of X and a for which equality holds. (Hint: What happens if X is a constant?)

b) Describe as completely as you can all possible random variables X for which equality holds.

2. (30 points) Chebyshev's inequality says that if X is a random variable with mean μ and variance σ^2, then

$$P\{|X - \mu| \geq a\} \leq \sigma^2/a^2.$$

a) Give an example of X and a for which equality holds.

b) Describe as completely as you can all possible random variables X for which equality holds.

You should do exactly one of the next two problems. (Please, be kind to the grader: don't do both! If you do, only the first will be graded.) The second one is more interesting, but it requires some basic facts from linear algebra, which you aren't supposed to need for this course. That's why the (boring) first one is offered as a possible subsitute.

3, option 1. (40 points) (Text, problem 7 on page 430.) Suppose that a fair die is rolled (independently) 100 times, and that the value obtained on the ith roll is X_i. Then each X_i is an integer between 1 and 6, so the product of all the X_i is an integer between 1 and 6^{100} Define

$$Z = \left(\prod_{i=1}^{100} X_i \right)^{1/100},$$

a random variable taking values between 1 and 6. Find a reasonable approximation for the cumulative distribution function

$$F_Z(a) = P\{Z \leq a\}, 1 < a < 6.$$

In particular, estimate the number

$$P\{2.8 \leq Z \leq 3.2\}.$$

3, option 2. (40 points) If $v = (v_0, v_1, \ldots, v_M)$ is a row vector in $\mathbb{R}^{M+1}$, then the "L^1-norm" of v is by definition

$$\|v\|_1 = |v_0| + |v_1| + \cdots + |v_M|.$$

For example, $\|(5,-2,3)\|_1 = 5+2+3 = 10$.

Suppose T is an $(M+1)\times(M+1)$ transition matrix for a Markov chain, so that the entries T_{ij} satisfy

$$T_{ij} \geq 0, \qquad \sum_{j=0}^{M} T_{ij} = 1 \quad (0 \leq i \leq M).$$

a) Write C for the $(M+1)\times 1$ column vector for which every entry is 1. Show that $TC = C$. Explain why this implies that $T - I$ has rank at most M (with I the $(M+1)\times(M+1)$ identity matrix), and therefore that $T^t - I$ has rank at most M. (Here T^t is the transpose of the matrix T.

b) Explain why there must be a non-zero row vector v such that $vT = v$. (Hint: v^t is an eigenvector of T^t with eigenvalue 1.)

c) The vector v from part (b) is a step in the direction of the "limit probabilities" (Π_i) of Theorem 9.2.1 in the text, but we haven't arrived there yet. Give an example of a transition matrix T and a row vector v having both positive and negative coordinates, satisfying $vT = v$. (Since some coordinates are negative, they can't be made into probabilities even by dividing by some constant.)

d) Assume that for some $\epsilon > 0$, every entry of T satisfies

$$T_{ij} \geq \epsilon.$$

Suppose that w is a row vector such that $\sum_{i=0}^{M} w_i = 0$. Show that

$$\|wT\|_1 \leq (1-(M+1)\epsilon)\|w\|_1.$$

(Hint: let E be the $(M+1)\times(M+1)$ matrix for which every entry is equal to 1 (not just the diagonal entries). Show that $wE = 0$, and therefore that $wT = w(T-\epsilon E)$. The matrix $T - \epsilon E$ has non-negative entries, and the sum of the entries in each row is equal to $1-(M+1)\epsilon$. Now write down the definition of $\|w(T-\epsilon E)\|_1$, and follow your nose.

e) Still assuming that $T_{ij} \geq \epsilon$ for all i and j, suppose that v is a non-zero vector with $vT = v$. Show that

$$\sum_{i=0}^{M} v_i \neq 0.$$

.

f) Still assuming that $T_{ij} \geq \epsilon$ for all i and j, fix a vector v with $vT = v$ and $\sum v_i = 1$ (as is possible by part (e)). Suppose x is any row vector with $\sum x_i = 1$. Show that

$$\|xT^n - v\|_1 \leq (1-(M+1)\epsilon)^n \|x - v\| \to 0$$

as $n \to \infty$. This implies that

$$\lim_{n\to\infty} xT^n = v.$$

This is a version of Theorem 9.2.1 in the text. Why does it follow from this statement that the entries v_i must all be non-negative?