Discrete Math in CS	**Problem Set 7**
CS 280 Fall 2005	**Due November 4, 2005**

Reading: Rosen, Section 5.3.

---

(1) In class, we discussed Markov's Inequality: if X is a random variable that only takes non-negative values and such that $E[X] > 0$, and if $c > 1$, then

$$\Pr[X \geq c \cdot E[X]] \leq \frac{1}{c}.$$

Show that this bound is sometimes "exact"; specifically, prove that for every $c > 1$, there exists a non-negative random variable X with $E[X] > 0$, for which

$$\Pr[X \geq c \cdot E[X]] = \frac{1}{c}.$$

(So the main step in your proof should consist of the construction of an appropriate random variable X, depending on c, that has this property.)

(2) In performing blood donation, it is important to be careful that the *type* of the blood being donated is compatible with the blood type of the recipient. Concretely, this works as follows. A person's own blood supply has certain *antigens* present (we can think of antigens as a kind of molecular signature); and a person cannot receive blood with a particular antigen if their own blood does not have this antigen present. This principle underpins the division of blood into four *types*: A, B, AB, and O. Blood of type A has the A antigen, blood of type B has the B antigen, blood of type AB has both, and blood of type O has neither. Thus, patients with type A can receive only blood types A or O in a transfusion, patients with type B can receive only B or O, patients with type O can receive only O, and patients with type AB can receive any of the four types.[1] Notice that blood of type O has the property that it can be donated to anyone, so hospitals are particularly interested in having a reasonable quantity of type-O blood in supply.

Roughly, the breakdown of blood types among people in the U.S. is as follows: 45% have type O, 42% have type A, 10% have type B, and 3% have type AB.

Now, consider the following scenario. A large blood donation clinic gets 10,000 random visitors, each of whom donates one unit of blood. They want to claim that with high probability, the amount of type-O blood they receive will be at least reasonably large.

Let's model this as the following random experiment: each of the 10,000 donors who arrive has a blood type chosen independently with probability .45 of O, probability .42 of

[1]The Austrian scientist Karl Landsteiner received the Nobel Prize in 1930 for his discovery of the blood types A, B, O, and AB.

A, probability .10 of B, and probability .03 of AB, Let X be a random variable equal to the number of people, from among this group, who have type-O blood.

(a) Determine $E[X]$ and $\text{Var}(X)$. Show the reasoning and/or calculations you use to get these values.

(b) Show that

$$\Pr[X < 4000] \leq .01.$$

In other words, with high probability there will be at least 4000 type-O donors in this group. Explain your reasoning and/or calculations in arriving at this answer.

(3) Suppose you're managing the storage for a data-intensive satellite survey of the Earth's surface. Each day the survey generates an amount of data that is a whole number of terabytes drawn uniformly at random from the integers between 5 and 15 inclusive. (In other words, the amount of data generated in one day is equally likely to be one of the values 5 TB, 6 TB, 7 TB, ..., 14 TB, 15 TB.) We'll assume that the (random) amount of data generated on one day is independent of the amount generated on any other day.

The survey is going to last for 100 days, and the project's budget allows you to buy 1200 terabytes of disk storage. Naturally, you're interested in knowing how likely it is that the survey will generate so much data that it overflows the disks you have available.

Show that with probability at least 95%, the survey generates an amount of data that fits on your available disk storage. Give an explanation for your answer.