COS 341: Discrete Mathematics

Homework #7 — Fall 2006
Probability II — Due: Friday, December 1

See instructions on the "assignments" web-page on how and when to turn in homework, and be sure to read the collaboration and late policy for this course. Approximate point values are given in brackets. *Be sure to show your work and justify your answers.*

1. [7] Consider the following Java code to compute the maximum number stored in an array of size n:

```
max = a[0];
for (int i = 1; i < n; i++)
{
    if (max < a[i])
    {
        max = a[i];
    }
}
```

Assume that the array `a` is a permutation of the integers $\{1, 2, \ldots, n\}$, and each permutation is equally likely. What is the expected number of times the line `max = a[i];` is executed?

2. Melanie the meteorologist gets a job predicting the weather at a TV station that pays her a bonus based on the accuracy of her predictions. Specifically, let Y be a random variable that indicates whether or not it rains tomorrow; that is, $Y = 1$ if it rains, and $Y = 0$ if it does not rain. Suppose the "true" (but generally unknown) probability of rain tomorrow is p; thus, $Y = 1$ with probability p, and $Y = 0$ with probability $1 - p$. Melanie is asked to make a prediction of the form, "The probability it will rain tomorrow is q."

a. [5] Suppose that the TV station pays Melanie a bonus of $1000 \cdot (1 - |Y - q|)$ dollars based on the outcome Y and her prediction q; the idea is that the closer her prediction q is to the actual outcome Y, the higher is her bonus. (For instance, if she predicts 70% chance of rain, and it does rain, she would get paid a bonus of $1000(1-|1-.7|) = 700$ dollars.) Suppose Melanie somehow knows the true probability of rain p. How should she choose her prediction q (as a function of p) to maximize her expected bonus? What will her expected bonus be if she chooses q in this manner?

b. [5] Answer the same questions as above when the TV station instead pays Melanie a bonus of $1000 \cdot (1 - (Y - q)^2)$ dollars based on the outcome Y and her prediction q.

c. [2] Which pay scheme seems more appropriate for this situation, and why?

3. The hat-check staff has had a long day, and at the end of the party they decide to return people's hats at random. Suppose that n people have their hats returned at random. We previously showed that the expected number of people who get their own hat back is 1, irrespective of the total number of people. In this problem, we will calculate the variance in the number of people who get their hat back.

Let $X_i = 1$ if the i-th person gets his or her own hat back and 0 otherwise. Let

$$S = \sum_{i=1}^{n} X_i$$

be the total number of people who get their own hat back.

a. [7] Show that $\mathrm{Var}(S) = 1$. (*Hint:* First calculate $\mathrm{E}[X_i X_j]$ for all i, j, and then calculate $\mathrm{E}[S^2]$.)

b. [2] Explain why you cannot use the variance of sums formula to calculate $\mathrm{Var}(S)$.

c. [4] Show that the probability that at least 11 people get their own hats back is at most 0.01.

4. A real-valued function f defined on the real numbers (or an interval of the real numbers) is said to be *convex* if for any two numbers x and y in its domain, and any number $0 \leq a \leq 1$,

$$f(ax + (1-a)y) \leq af(x) + (1-a)f(y).$$

A useful fact, which you may use below without proof, states that f is convex if its second derivative exists and is nonnegative everywhere, that is, if $f''(x) \geq 0$ for all x in f's domain.

a. [7] *Jensen's inequality* states that if X is any real-valued random variable and f is any convex function (whose domain includes the range of X), then

$$f(\mathrm{E}[X]) \leq \mathrm{E}[f(X)].$$

Prove Jensen's inequality in the special case that the range of X is finite.

b. [5] Give *two* short proofs — one based on Jensen's inequality, and the other based on properties of variance — that for any real-valued random variable X,

$$\mathrm{E}[X^2] \geq (\mathrm{E}[X])^2.$$

5. Let $p_1, \ldots, p_n$ and $q_1, \ldots, q_n$ be sequences of strictly positive numbers for which

$$\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1.$$

Thus, the p_i's and q_i's define probability distributions P and Q over n items. Let us define the *entropy* of P, denoted $H(P)$, to be

$$H(P) = -\sum_{i=1}^{n} p_i \lg p_i,$$

and let us define the *relative entropy* between P and Q, denoted $D(P\|Q)$, to be

$$D(P\|Q) = \sum_{i=1}^{n} p_i \lg(p_i/q_i).$$

(First side note: In class, we defined $H(p)$, where $0 \leq p \leq 1$ is a real number rather than a distribution. In fact, this notation is just shorthand for the entropy of a (Bernoulli) distribution over $\{0, 1\}$ in which the probability of 1 is equal to p. Thus, $H(p) = H((p, 1-p))$. Similarly, the notation used in class $D(p\|q)$, where p and q are numbers between 0 and 1, is shorthand for $D((p, 1-p)\|(q, 1-q))$.)
(Second side note: Although you only need to prove the results in this problem when the p_i's and q_i's are strictly positive, these results also hold even when some of them may be zero, provided we define $0 \lg 0$ to be 0.)

a. [7] Use Jensen's inequality to prove that $D(P\|Q) \geq 0$. Also show that $D(P\|Q) = 0$ if P and Q are identical. (*Hint:* In Jensen's inequality, consider a random variable X which is equal to q_i/p_i with probability p_i; then find an appropriate convex function f to prove the inequality.)

b. [5] Use your answer to the previous part to show that $0 \leq H(P) \leq \lg n$.