Exercise 1 (Due: 7.12.2008)

Grading: Every question states its maximum grade. The sum of the question grades throughout the semester forms the basis on which the final course grade is calculated, so higher maximum grades imply higher weights.

Random linear equations (9 points)

Consider the following system of m random homogeneous linear equations over $\mathbb{R}^n$: For every $1 \le i \le m$ and $1 \le j \le n$, we choose $a_{i,j}$ uniformly and independently at random from $\{0, 1\}$. Then for every i the i'th equation is $\sum_{j=1}^{n} a_{i,j} x_j = 0$. Show that for $m = n + 1$, the probability that the system has no solution but the all-zero one is at least $\frac{1}{2}$.

Hint: Try proving this over $(\mathbb{Z}_2)^n$ first. In fact, *you will get 6 out of the 9 available points for doing just that.*

Games with envelopes (12 points)

Consider the following game: A random variable X gets its value from a known distribution over the positive integers, and then an amount of $\$X$ is placed in one envelope and an amount of $\$2X$ is placed in the other. A player gets to open one of the two envelopes, chosen uniformly at random, and then faces the choice of either keeping its contents or trading it for the other unopened envelope.

- Show that if the expectation $\mathrm{E}[X]$ is finite then there exists n such that it is better (on average) for the player not to trade whenever the first envelope contains this amount.
- Show that there exists a distribution (without a finite expectation) over the positive integers for which the player would prefer to trade regardless of the amount in the first envelope. This shows the futility of "on the average" calculations for a distribution that had an infinite expectation to begin with.
- Does there exists a distribution with a finite expectation for which there exist infinitely many amounts (but not all of them) which cause the player to trade? Prove your answer.

Exercise 2 (Due: 11.1.2009)

Satisfying parts of a CNF (9 points)

Suppose that we have a CNF formula with m clauses, each being an OR of some literals. All we are told is that there are no empty clauses (which are by themselves unsatisfiable), and that for no variable x_i do we have both clauses "x_i" and "$\neg x_i$" (which means that every pair of clauses can in itself be simultaneously satisfied). Prove that there is an assignment that satisfies at least $\frac{1}{2}(\sqrt{5}-1)m$ of the clauses.

Hint: It is easier to first show this for a CNF formula where each clause consists of either one non-negated literals or an OR of two negated literals, but repeated clauses are allowed. *You will get 6 out of the 9 available points for showing just that.* Note that there is a 1981 article proving the above result, only it did not explicitly use the probabilistic method and hence was longer.

Isolating multisets (12 points)

Suppose that $\mathcal{F}$ is a family of *multisets* taken from the set $A = \{1, \ldots, n\}$, where now every member of A can appear up to r times in a member of $\mathcal{F}$. Again we pick uniformly and independently a weight $w(a) \in \{1, \ldots, c\}$ for every $a \in A$, and calculate the weights of the members of $\mathcal{F}$ with multiplicities (i.e. if $E \in \mathcal{F}$ "contains" t copies of $b \in A$ then $w(b)$ is added t times). Prove that there exist a unique member of $\mathcal{F}$ with minimum weight with probability at least $1 - \frac{rn}{c}$.

Also, give a counter example (it is allowed to do so for specific values of n, r and c) for which the probability if having a unique member with minimum weight is *less* than $1 - \frac{n}{c}$.

You get 8 points for the proof and 4 points for the counter example. You can also get a partial credit of 4 points for a proof of a weaker statement that gives a probability of $1 - \frac{f(r)n}{c}$ *for some function* $f(r)$.

Exercise 3 (Due: 15.2.2009)

Fixed points in a permutation (6 points)

Let $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ denote a permutation chosen uniformly at random (from the set of all $n!$ possible permutations), and let $f(\sigma)$ be the number of its fixed points – a fixed point is $1 \le i \le n$ for which $\sigma(i) = i$. Prove that the probability of having no fixed points at all is less than $\frac{19}{20}$ (it is enough to prove this for any constant number lesser than 1, although it is hard to think of a method that will do worse than $\frac{19}{20}$).

A martingale inequality (9 points)

Show that for every martingale $X_0, X_1, \ldots, X_n$, where X_0 is not guaranteed to be a constant, the following holds:

$$\sum_{i=1}^{n} \mathrm{E}[(X_i - X_{i-1})^2] = \mathrm{E}[X_n^2] - \mathrm{E}[X_0^2] \le \mathrm{Var}[X_n]$$

You will get 6 points for the equality and 3 points for the inequality (it is enough to prove this over finite probability spaces).

Exposing a permutation (6 points)

Again let σ denote a permutation chosen uniformly at random. Let $c(\sigma)$ be the number of cycles in its decomposition to unique disjoint cycles. We define an exposure martingale $X_0, \ldots, X_n$ by letting X_i be the expectation of $c(\sigma)$ conditioned on $\sigma(1), \ldots, \sigma(i)$. In other words we use $\mathcal{D}_i = \{1, \ldots, i\}$. Note that $X_{n-1} = X_n$ with probability 1. Prove that the martingale satisfies the Lipschitz condition, $|X_i - X_{i-1}| \le 1$ with probability 1 for every $1 \le i \le n-1$.

Exercise 4 (Due: 15.3.2009)

Kleitman for multisets (9 points)

Suppose that $\mathcal{A}$ and $\mathcal{B}$ are sets (or families) of multisets (remember those?) over some S, where every member of S can appear at most r times in a member of $\mathcal{A}$ or $\mathcal{B}$. Also assume that these are monotone nondecreasing: If $C \in \mathcal{A}$ and $C \subseteq C'$ then $C' \in \mathcal{A}$, and the same happens for $\mathcal{B}$ (here containment is defined that for every $t \in S$ the multiset C' has at least as many instances of t as C does).

Show that in this case $|\mathcal{A}||\mathcal{B}| \le (r+1)^{|S|}|\mathcal{A} \cap \mathcal{B}|$ (this is a standard intersection of sets of objects, do not confuse it with $\sqcap$). It is allowed to use theorems proved in class.

Walking on hypercubes (9 points)

Here we concern ourselves with the following graph G: The set of vertices is $V = \{0,1\}^n$, (i.e. the set of all binary strings of size n, note that $|V| = 2^n$), and the set of edges E is the set of all pairs of binary strings that differ on exactly one coordinate. This graph is called the hypercube graph of dimension n.

Now we define the following random sequence $\underline{\mathrm{X}} = X_0, X_1, \ldots$: We start with $X_0 = (0, \ldots, 0)$, and given X_k we set $X_{k+1} = X_k$ with probability $\frac{1}{2}$, and with probability $\frac{1}{2}$ we set X_{k+1} to be a uniformly random neighbor of X_k. Note that this is the "random walk with stops" that was defined in class, for the special case of the graph G given above.

Prove that for every ϵ there exists c such that if $k \ge n(\ln n + c)$ then the variation distance between the unconditional distribution of X_k and the uniform distribution over V is at most ϵ. You may assume that n is larger than some constant, such as 2.