Math 55: Discrete Math

G.S.I. Loren Looger

November 4, 1997

Homework Assignment #12

Note: Due to the difficulty in the problem set, and the unavailability of an answer key, some GSI's (myself included) have inadvertantly disseminated incorrect information regarding the solutions to some of the problems. These issues will be straightened out in section, but I will try to make light of them in the solutions as well, so as to avoid any possible confusion.

Lenstra 11

We are given that f is a binomial random variable, so that for some n, p, the probability of f taking a given value k is determined by:

$$P(f = k) = \binom{n}{k} p^k q^{n-k}$$

where, as usual, $q = (1-p)$. Thus we may quickly write down the probabilities that f takes on an even or an odd value, respectively:

$$P(f\ is\ even) = \sum_{k=0, k\, even}^{n} P(f = k) = \sum_{k=0, k\, even}^{n} \binom{n}{k} p^k q^{n-k}$$

$$P(f\ is\ odd) = \sum_{k=0, k\, odd}^{n} P(f = k) = \sum_{k=0, k\, odd}^{n} \binom{n}{k} p^k q^{n-k}$$

Now we may also quickly substitute to find the generating function g for f:

$$g(z) = \sum_{k=0,k}^{n} P(f = k) z^k = \sum_{k=0,k}^{n} \binom{n}{k} p^k q^{n-k} z^k = (pz + q)^n$$

Thus,

$$g(-1) = P(f\ is\ even) - P(f\ is\ odd) = (-p + q)^n = (1 - 2p)^n$$

And of course,

$$P(f\ is\ even) + P(f\ is\ odd) = 1$$

Solving, we have:

$$P(f\ is\ even) = \frac{1 + (1-2p)^n}{2} \qquad P(f\ is\ odd) = \frac{1 - (1-2p)^n}{2}$$

In particular, we see that as n grows large and $p < 1$, both probabilities approach $\frac{1}{2}$.

Lenstra 12

For this problem, define the random variable f to count the number of children that get the letter corresponding to their name. Then:

$$f = \sum_{k=1}^{52} f_k,$$

where f_k is the random variable that is 1 if the k^{th} child gets the appropriate letter, and 0 if not. Then, no matter what the f_k's correspond to, we have: $E(f) = \sum_{k=1}^{52} E(f_k) = 52 * E(f_1) = 52 * 2/52 = 2$, as every child has 2 letters out of the 52 total that begin their name. Now, the question of variance is a thorny one. The f_k's are not independent, as if two children have names beginning with the same letter, then the fact that one child gets the appropriate letter decreases the available number of letters for the second child, thus decreasing the probability that the child gets the appropriate letter. Thus, we may not easily compute the variance of f from the variances of the individual f_k's. In general, the variance of f depends completely on the actual names of the children. Consider this following worst case: all 52 children are named Ken. Then $f = 2$; it is always true that exactly 2 children get the right letter. So in this case, $V(f) = 0$. But if there are 2 children whose names begin with each letter, then the variance will be much higher. So the variance may not be computed without complete knowledge of the names. I stated in section that f was a binomial random variable; this is simply not true. The limited number of each type of letter causes the f_k's to not be independent.

Lenstra 13

This problem is obviously an application of Chebyshev's Inequality; it is the only inequality regarding the distribution of the values of f around its expected value that we know. Thus:

Let the random variable f count the number of elements in a subset of the set $\{1, 2, 3, ...100\}$. So,

$$f = \sum_{k=1}^{100} f_k,$$

where f_k is the random variable taking 1 if k is in the subset, and 0 otherwise. Now, a given element k is either in or out of the subset each with probability $\frac{1}{2}$. Thus,

$$E(f) = \sum_{k=1}^{100} E(f_k) = 100 * \frac{1}{2} = 50.$$

Also, the random variable f is a binomial random variable. If a subset has k elements, then a chosen k of the 100 total elements are in the subset, and the other $100 - k$ are not. This event occurs with probability:

$$P(f = k) = \binom{100}{k} (1/2)^k (1/2)^{100-k}.$$

So f is binomial, and from Lenstra, $V(f) = n * p * q = 100 * 1/2 * 1/2 = 25$. Now, from Chebyshev's Inequality:

$$P(|f(x) - 50| \geq 10) = P(|f(x) - E(f)| \geq 10) \leq V(f)/10^2 = 25/10^2 = 0.25$$

Thus, the desired event occurs with probability $\geq (1 - 0.25) = 0.75$.

Lenstra 15

For this problem, all we must do is give one example of a random variable f where equality does hold in Chebyshev's inequality. We take the random variable

$$f = 1$$

Then, $V(f) = E(f^2) - (E(f))^2 = 1 - 1 = 0$. Also,

$$P(|f(x) - 1| \geq r) = 0,$$

for any positive r. Thus equality holds in Chebyshev's Inequality.

Lenstra 17

For this problem, we first notice that the order of the 20 total numbers may be determined by first placing the first i numbers, and then the remaining $20 - i$ numbers. So we count the number of sequences with the i^{th} number checked. We first choose the i numbers that come first. We may do this in $\binom{20}{i}$ ways. Given this choice, there is a unique largest element among the i, and this number must go in the i^{th} position, or else the i^{th} number will not be checked. But the remaining $(i-1)$ may be placed in any of the $(i-1)!$ ways possible. Similarly, the final $(20 - i)$ numbers on the list may be placed in any of the $(20 - i)!$ ways possible. Thus the total number of sequences with the i^{th} number checked is:

$$\binom{20}{i} (i-1)!(20-i)!$$

The probability that the i^{th} number is checked is then:

$$\binom{20}{i}(i-1)!(20-i)!/20! = \frac{20!(i-1)!(20-i)!}{i!(20-i)!20!} = \frac{(i-1)!}{i!} = \frac{1}{i}.$$

Now, if f is the random variable that counts the number of checked numbers, then:

$$f = \sum_{k=1}^{20} f_k,$$

where f_k is the random variable that is 1 if the k^{th} number is checked, and 0 otherwise. Thus,

$$E(f) = \sum_{k=1}^{20} E(f_k) = \sum_{k=1}^{20} \frac{1}{k} \sim 3.6$$

I stated in office hours that we must first demonstrate that the f_k's are independent before making this claim. It is actually a surprising fact that the f_k's are independent, but this is irrelevant to the problem at hand.

Lenstra 21

The way to solve this problem is to pay careful attention to the hint, and follow it. We write x in binary:

$$x = 0.a_1a_2a_3...$$

I contend that if $a_k = 1$, then Karel will be the person throwing the coin on the k^{th} turn of the game. Otherwise, it will be Klaas. We can see this in the following manner: The statement hols true for the first turn: If $a_1 = 0$, then Klaas begins the game by tossing the coin. Otherwise, Karel is the first to throw. Now, assume the statement is true for a_k. Now, if $a_k = a_{k+1} = 1$, then by assumption, Karel flipped the coin on the k^{th} turn, and since the value of x on the k^{th} turn is $\leq 1/2$, either $x = 1/2$, and the game ends before the $(k+1)^{st}$ throw, or else the binary expansion of x begins with $0.a'_k a'_{k+1}...$, where a'_i denotes the opposite binary digit from a_k. Thus, the binary expansion of x begins with $0.00...$, and after the k^{th} round ends, the new doubled value of x still beings with $0.0...$, and thus Karel retains the coin to throw on the $(k+1)^{st}$ turn. The verification of the other three cases is similar. Now, the amount that Karel expects to win from the game is the sum of the amounts he expects to win if the game ends in exactly n turns. The probability that the game ends in exactly n turns is $\frac{1}{2^n}$, as the game must have consisted of $(n-1)$ occurrences of tails, and a single final occurrence of heads, each with probability $1/2$. Now, Karel wins something only if he is the one that tossed the winning coin. That is, only if $a_n = 1$. If he tosses the winning coin, he wins 1 coin. Thus, the expected winnings for Karel if the game ends in exactly n turns is:

$$a_n/2^n.$$

Thus his total expected winnings are:

$$\sum_{n=1}^{\infty} \frac{a_n}{2^n} = x.$$

Thus the game is fair.

Lenstra 26

(a) To get 2 different letters, one must first get one, and then given that we have one, we must get another. The expected number of boxes to yield a single letter is obviously 1. Now, given that we already have one letter, the probability of getting a new letter is $5/6$ with each new box. Thus the expected number of boxes needed to give the second letter is $1\frac{1}{5/6} = 6/5$. The total number of boxes expected is thus:

$$1 + 6/5$$

(b) To get all 6 different letters, we must first get one, then another, then another, etc.. Given that we have k different letters, $0 \leq k \leq 5$, the probability of getting a new different letter is $\frac{6-k}{6}$. The expected number of boxes needed to get the $(k+1)^{st}$ letter is thus $\frac{1}{(6-k)/6} = \frac{6}{6-k}$. Thus the total number of boxes expected to yield all six letters is:

$$6(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6}) = 14.7$$

(c) The random variable that is referenced is the random variable f that counts how long it took us to get all 6 letters in a given sequence of cereal-buying. Thus the desired generating function is

$$\sum_{n=1}^{\infty} P(f = n) z^n,$$

where $P(f = n)$ is the probability that it took us exactly n boxes to get the sixth letter. Now, what is this probability? The first thing to choose is the number of different last letters we can purchase; this is 6. Now, if we get the sixth letter on the n^{th} turn, then we got only the other 5 letters up to that point. But we must have gotten them all at least once, or else we would not be done. Thus, we have $(n-1)$ letters chosen from a set of 5, with each one used at least once. The number of such sequences is $Onto(n-1, 5)$, the number of onto functions from a set with $(n-1)$ elements to a set with 5 elements. Thus the total number of sequences giving the sixth letter on the n^{th} box is:

$$6 * Onto(n-1, 5)$$

The probability that the sixth letter comes on the n^{th} turn is thus

$$6 * Onto(n-1, 5)/6^n = Onto(n-1, 5)/6^{n-1},$$

as all sequences of n boxes are equally likely. The generating function is thus

$$\sum_{n=1}^{\infty} \frac{Onto(n-1, 5)}{6^{n-1}} z^n$$

I am told that this is equivalent to:

$$\frac{5! z^6}{(6-z)(6-2z)(6-3z)(6-4z)(6-5z)}.$$