COS 341, September 27, 2000
Handout Number 2

The Dating Problem and the Lottery Problem

Probability

A *probability space* Ω is a pair (U, p) where U is a finite set, and $p : U \to [0, 1]$ is a function such that $\sum_{u \in U} p(u) = 1$. An *event* T is a subset of U. The *probability for* the event T *to occur* is defined as $p(T) \equiv \sum_{u \in T} p(u)$.

The notion of probability space is used to model a random process, in which U is the set of all possible configurations, and configuration u is to be the realized configuration with probability $p(u)$. We restrict U to be a finite set for simplicity (in general U could be infinite). The notion of event models a boolean predicate that is of interest.

The Dating Problem

In the Dating Problem with parameters n, k, the probability space is $\Omega = (U, p)$, where U is the set of all permutations of $\{1, 2, \cdots, n\}$ (hence $|U| = n!$), and $p(u) = 1/|U|$ for all $u \in U$. It is easy to see that a permutation $u = (i_1, i_2, \cdots, i_n) \in U$ is in T if and only if the following conditions are satisfied:

C1: $1 \in \{i_{k+1}, i_{k+2}, \cdots, i_n\}$;
C2: Let $i_j = 1$ where $k < j \leq n$, then the minimum of $i_1, i_2, \cdots, i_{j-1}$ is among the first k of these numbers.

Let T_j denote the set of all u's satisfying C1, C2 with j being the j in C2. Then T is the disjoint union of such T_j. By the Addition Principle, we have

$$|T| = \sum_{k < j \leq n} |T_j|. \tag{1}$$

We assert that for each $k < j \leq n$,

$$|T_j| = \frac{k}{j-1}(n-1)!. \tag{2}$$

To see this, note that T_j is itself the disjoint union of $T_{j,1}, T_{j,2}, \cdots, T_{j,k}$, where $T_{j,s}$ consists of $u \in T_j$ with the minimum of $\{i_1, i_2, \cdots, i_{j-1}\}$ occurring at i_s. Now each element of $T_{j,s}$ can be specified by first choosing an $(n-j)$-permuation of $\{2, 3, \cdots, n\}$ (to fix $i_{j+1}, i_{j+2}, \cdots, i_n$), and then a $(j-2)$-permutation of the set L, where L is defined as the set $\{2, 3, \cdots, n\} - \{i_{j+1}, i_{j+2}, \cdots, i_n\}$ minus its minimum element (to fix $(i_1, i_2, \cdots, i_{j-1})$). Therefore, by the Multiplication Principle, we have

$$\begin{aligned} |T_{j,s}| &= P(n-1, n-j) \cdot (j-2)! \\ &= \frac{(n-1)!}{(n-1-(n-j))!}(j-2)! \\ &= \frac{(n-1)!}{j-1}. \end{aligned}$$

This proves (2). (Alternatively, one can argue that it is equally likely for a random u with $i_j = 1$ to have the minimum of $i_1, i_2, \cdots, i_{j-1}$ to occur at s for any $1 \leq s \leq j-1$. Since there are $(n-1)!$ u's with $i_j = 1$, the number of such permutations with the minimum occurring in the first k locations is equal to $\frac{k}{j-1} \cdot (n-1)!$.)

It follows from (1) and (2) that

$$\begin{aligned} |T| &= \sum_{k<j\leq n} \frac{k}{j-1} \cdot (n-1)! \\ &= n! \frac{k}{n}\left(\frac{1}{k} + \frac{1}{k+1} + \cdots + \frac{1}{n-1}\right). \end{aligned}$$

Since $|U| = n!$, this means

$$p(T) = \frac{k}{n}\left(\frac{1}{k} + \frac{1}{k+1} + \cdots + \frac{1}{n-1}\right). \tag{3}$$

This completes the analysis of the probability of success under the strategy used with parameters n, k.

For example, if $n = 8, k = 4$,

$$p(T) = \frac{4}{8}\left(\frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}\right) = 319/840 = 0.38.$$

Lottery

We considered in class a simplified version of California Lottery. Let $S = \{1, 2, \cdots, 51\}$. The probability space is $\Omega = (U, p)$, where U is the set of all 6-combinations of S, and p is uniform on U (i.e., $p(u) = 1/|U|$ for all $u \in U$). Assume that you have picked a particular 6-combination $I = \{i_1, i_2, \cdots, i_6\}$. When the lottery drawing is to be done, the event T corresponding to your winning 5 dollars is the set of all 6-combinations that intersect $\{i_1, i_2, \cdots, i_6\}$ in exactly three elements. (If the intersection is larger than 3, you win more than 5 dollars.) What is $p(T)$?

Clearly, $p(T) = |T|/|U|$. As $|U| = \binom{51}{6}$, we only need to calculate $|T|$. For each subset $J \subseteq I$ of size 3, let T_J be the set of all 6-combinations of S that have J as their intersection with I. Then by the Addition Principle,

$$|T| = \sum_J |T_J|.$$

As each $|T_J| = \binom{|S|-|I|}{3} = \binom{45}{3}$, we have $|T| = \binom{6}{3}\binom{45}{3}$. (Alternatively, the above formulas follows from the Multiplication Principle as demonstrated in class.) Thus,

$$p(T) = \binom{6}{3}\binom{45}{3}/\binom{51}{6} = 283800/18009460 = 1.5\%.$$