Journal of Statistical Physics, Vol. 58, Nos. 3/4, 1990

The Statistics of Derangement—A Survey

J. Gillis[1]

Received July 14, 1989

Given a finite set X of elements, divided into disjoint subsets, we define a derangement of X as a permutation which leaves none of the elements in their original subsets. The probability of a random permutation being a derangement is discussed, particularly its asymptotic value as the cardinality of X and the number of subsets tend, under certain conditions, to infinity. Finally, the problem is extended to studying the number of elements which are transferred by a general permutation to a subset other than their initial one.

KEY WORDS: Combinatorial analysis; derangements; Laguerre polynomials.

Suppose that $S_1, S_2,..., S_k$ are mutually exclusive finite sets of elements and $X = \bigcup_{i=1}^{k} S_i$. A *derangement* of X is a permutation of its elements in which no element is left in its original set. If n_i denotes the number of elements in S_i, $(1 \leqslant i \leqslant k)$, we define $D(n_1, n_2,..., n_k)$ as the total number of possible derangements of X. The probability that a random permutation is in fact a derangement is then given by

$$P(n_1, n_2,..., n_k) = D(n_1, n_2,..., n_k) \Big/ \left(\sum_i n_i\right)!$$

The starting point for the derivation of most of the results given below is the formula[1,3]

$$D(n_1, n_2,..., n_k) = (-1)^{\sum n_i} \int_0^\infty \left\{\prod_{i=1}^{k} L_{n_i}(x)\right\} e^{-x}\, dx \qquad (1)$$

This paper is dedicated to Cyril Domb, in friendship.

[1] Department of Applied Mathematics and Computer Science, Weizmann Institute of Science, Rehovot 76100, Israel.

0022-4715/90/0200-0575$06.00/0

where $L_n(x)$ is the Laguerre polynomial whose expression is

$$L_n(x) = \sum_{r=0}^{n} (-1)^r \binom{n}{r} \frac{x^r}{r!}$$

It follows that

$$P(n_1, n_2, \ldots, n_k) = \left[(-1)^{\sum_{i=1}^{k} n_i} \Big/ \left(\sum_{i=1}^{k} n_i \right)! \right] \int_0^\infty \left\{ \prod_{i=1}^{k} L_{n_i}(x) \right\} e^{-x}\, dx \quad (2)$$

In the special case $n_1 = n_2 = \cdots - n_k = n$, it will be more convenient to use the simpler notation $P_k(n)$, $D_k(n)$. The explicit formula (2) can be exploited to prove for the function $P(n_1, n_2, \ldots, n_k)$ a variety of inequalities, recurrence relations, and asymptotic estimates, not all of which seem to be derivable directly from combinatorial considerations.

The trivial case of $P_k(1)$ is traditionally modeled by the problem of the secretary who typed k letters, with an appropriate envelope for each, but then absentmindedly put letters into envelopes simply at random. The probability that none of the letters reaches its proper destination is given by $P_k(1)$. It is easily proved by an inclusion-exclusion argument that

$$P_k(1) = \sum_{r=0}^{k} (-1)^r/r! = e^{-1} + O\left(\frac{1}{(k+1)!} \right) \quad (3)$$

This result can also be obtained directly from (2) by noting that $L_1(x) = 1 - x$.

A less trivial problem arises if each of the k letters consists of n pages and the secretary now shuffles the entire pack of nk pages before putting a random n pages into each of the k envelopes. What is now the probability that not a simple page is correctly addressed? From (2) we have

$$P_k(n) = \frac{(-1)^{nk}}{(nk)!} \int_0^\infty \{L_n(x)\}^k e^{-x}\, dx \quad (4)$$

For large k, the integral can be estimated by the Laplace method and we obtain

$$P_k(n) = e^{-n} + O(k^{-1}) \quad (5)$$

It follows incidentally, from (3) and (5), that for large k,

$$P_k(n) \sim \{P_k(1)\}^n + O(1/k) \quad (6)$$