August 10, 2013

Re: M/J 2 Solution
To Allan Gottlieb, Technology Review Puzzle Corner:

The problem of having 2N socks of N colors and withdrawing n of them, one at a time, without replacement, is an example of hypergeometric distribution with *a*=2 socks of each color. For a particular color, among the n socks, there may be x=0, 1, or 2 of that color. The hypergeometric distribution (used when *a* objects of a certain kind are identifiable among 2N total objects and samples are drawn without replacement) then tells us that the distribution of x is
$f(x) = \frac{C(2,x)C(2N-2,n-x)}{C(2N,n)}$. For each value of x, the probabilities are
$p_0(n) = \frac{(2N-n)(2N-n-1)}{2N(2N-1)}$, $p_1(n) = \frac{n(2N-n)}{N(2N-1)}$, $p_2(n) = \frac{n(n-1)}{2N(2N-1)}$. It can be shown that these three probabilities add up to 1, and also the sum of expectation values for N colors $0*N*p_0(n)+1*N*p_1(n)+2*N*p_2(n)=n$. The terms show expectation values, after n socks are withdrawn, of number of socks of missing colors (0), one color (1) and matched pairs (2). We get the answer to the problem in the 2^{nd} term ($Np_1(n)$) or by taking $n-2Np_2(n)$. Each of these is equal to the expectation value of the number of socks on the bed after n socks are withdrawn,

$$E_{sox}(n) = \frac{2nN-n^2}{2N-1}. \quad (1)$$

From the derivative, the maximum should be at N-½, but when we check N=10 to N=20 on an Excel sheet, we find that n=N produces the maximum each time, for an expectation value of $N^2/(2N-1)$ at n=N. Note how this gradually approaches N/2. Number of unmatched socks for N=10 is 5.263... and for N=20 is 10.2564...

The M/J 2 problem statement, however, asks for the "expected value of the maximum" number of unmatched socks on the bed. The Puzzle Corner solution, now published online, addresses that. The above Eq. (1) gives the expected value of the number of unmatched socks, and looks for a maximum, a slightly different concept. But (1) is a closed form solution and it does check exactly with the state diagram as published in Sept/Oct 2013 Puzzle Corner (but please note that the (3,1) node should be 3/5, not 2/5) for 2N=6.

The hypergeometric distribution formula above is as stated by the late, great MIT Professor George Wadsworth, in his probability book that we used in 18.10. Every time I re-read a section of that book, I am enchanted by his concise and precise writing.

Yours sincerely,

Timothy J. Maloney '71
Palo Alto, CA