An Exact Solution for the Expectation of Mutually Nearest Hamsters on a Road

Ham H.L. & Robo G.G.

4 December 2012

1 Introduction

We consider the important problem of n hamsters i.i.d. uniformly distributed along a one-dimensional road of fixed and finite length. Define any pair of hamsters that are nearest to each other than they are to any other hamsters as both being *mutually nearest*. We wish to determine the expected proportion of mutually nearest hamsters.

Consider that the hamsters' positions along the road are $h_1, h_2, ..., h_{n-1}, h_n$ in increasing order from the origin, and that the *neighbours* of a hamster h_i are h_{i-1} and h_{i+1}, if they exist. Then, every hamster must have at least one and at most two neighbour hamsters, for all non-trivial cases $n > 1$. It follows that every hamster must have exactly one *nearest neighbour* hamster, and can belong to at most one pair of mutually nearest hamsters.

Next consider that the *distances* between the above positions are $D_1, ..., D_{n-1}$, with D_i denoting the distance between h_i and h_{i+1}. Then $E(D_i) = \frac{h_n - h_1}{n-1}$, i.e. the expected distance between any neighbouring pair of hamsters is identical.

2 Approximation For Large n

A näive approximation of the desired expectation for large n could then be obtained by considering the case of any four hamsters $h_{i-1}, h_i, h_{i+1}, h_{i+2}$ with corresponding distances D_{i-1}, D_i, D_{i+1}. Define a *magnitude permutation* as some ordering of distances, such that if D_a occurs before D_b in the ordering, $D_a < D_b$. Clearly, the order of a distance D_i in a magnitude permutation may not correspond to its order by distance from the origin of the road (which is i)

Now consider the six *magnitude permutations* of $\{D_{i-1}, D_i, D_{i+1}\}$. Since $E(D_a) = E(D_b) \ \forall \ D_a, D_b$, the expectations of these magnitude permutations are exactly equal. Then, considering only the three permutations where h_i is closer to h_{i+1} than it is to h_{i-1}, i.e. $D_i < D_{i-1}$, we note that in two of these three permutations, $D_i < D_{i+1}$, and $\{h_i, h_{i+1}\}$ are a mutually nearest pair. Therefore, for all situations where this assumption holds, the expectation $E(h_i)$ that hamster h_i has the property of being part of a mutually nearest pair of hamsters is $E(h_i) = \frac{2}{3}$, and therefore for n hamsters, the expectation is $\frac{2n}{3}$.

3 An Exact Solution

The simple intuition expressed above is however but an approximation, since the assumption required does not apply at the boundaries, although this imprecision will be dominated when n is large. An exact proof is slightly more involved. We now consider, for n hamsters, all the (equally-probable) $(n-1)!$ magnitude permutations of the $n-1$ distances $D_1, ..., D_{n-1}$.

Define a *magnitude set* M_s on D_i as the set of all magnitude permutations where D_i has an order s, i.e. where D_i is the s^{th} shortest distance among all distances. Then, for any D_i, there are always $(n-1)$ magnitude sets, each with $(n-2)!$ permutations, covering all $(n-1)!$ permutations of magnitudes.

Here, we make a distinction between the two boundary cases $\{D_1, D_{n-1}\}$, and the $(n-3)$ intermediate cases $\{D_2, ..., D_{n-2}\}$.

3.1 Boundary Cases

Consider the hamsters h_1, h_2, h_3 with corresponding distances D_1, D_2 starting from the origin at the left boundary of the road. For h_1 and h_2 to be mutually nearest, it is necessary and sufficient that $D_1 < D_2$.

Now consider M_1 for D_1. In all permutations within this magnitude set of permutations, D_1 is the shortest, and therefore h_1 and h_2 are mutually nearest for all these permutations.

In M_2, D_1 is the second-shortest, and therefore h_1 and h_2 are mutually nearest for all permutations *but* the permutations where D_2 is ordered first in the permutation. In similar vein, in M_3, D_1 is the third-shortest, and h_1 and h_2 are mutually nearest for all permutations *but* those where D_2 is ordered first *or* second, and so on. Indeed, for M_s, the number of permutations where h_1 and h_2 are mutually nearest is $(n-1-s)(n-3)!$, and therefore for all M_s, the total number of such permutations is $\sum_{s=1}^{n-1}(n-1-s)(n-3)! = \frac{(n-1)!}{2}$. Since there are two boundary cases, the total number of permutations with mutually nearest $\{h_i, h_{i+1}\}$ at the boundaries is then simply:

$$T(boundary) = (n-1)! \tag{1}$$

3.2 Intermediate Cases

The analysis for the intermediate cases is similar, and the form has in fact been suggested by the approximation in Section 2. We again consider four consecutive hamsters $h_{i-1}, h_i, h_{i+1}, h_{i+2}$ with three corresponding distances D_{i-1}, D_i, D_{i+1} between them. Note that since the analysis is for intermediate cases only, and $2 \leq i \leq n-2$, it is always valid this time.

In particular, for M_s with D_i, we note that there are $(s-1)$ distances that are shorter than D_i, and $(n-1-s)$ distances that are longer. We are interested in the number of permutations where $D_i < D_{i-1}$ and $D_i < D_{i+1}$, i.e. D_{i-1} and D_{i+1} are both longer than D_i. Then there are ${}^{(n-1)-3}P_{s-1}$ permutations for the front $(s-1)$ distances, since we can choose any distances other than D_{i-1}, D_i, D_{i+1}, and $(n-1-s)!$ permutations for the back $(n-1-s)$ distances, since in each permutation they are drawn from the $(n-1-s)$ remaining valid distances after restrictions were observed on the front distances. The consolidated number of permutations where h_i and h_{i+1} are mutually nearest for M_s is therefore:

$$T(M_s|D_i, intermediate) = (n-1-s)!^{n-4}P_{s-1} \tag{2}$$

the total number over all M_s is:

$$T(D_i, intermediate) = \sum_{s=1}^{n-1} (n-1-s)!^{n-4}P_{s-1} \tag{3}$$

and the total number over all intermediate distances is:

$$T(intermediate) = (n-3) \sum_{s=1}^{n-1} (n-1-s)!^{n-4}P_{s-1} \tag{4}$$

3.3 Synthesis

From (1) and (4), we now have an expression for the total number of distances D_i that are bounded by a pair of mutually nearest hamsters, over all possible permutations of distances (by their magnitude or otherwise):

$$T(mutual) = (n-3) \sum_{s=1}^{n-1} (n-1-s)!^{n-4}P_{s-1} + (n-1)! \tag{5}$$

out of the total number of hamsters implied by those permutations:

$$T(total) = n(n-1)! = n! \tag{6}$$

and, motivated by the approximation, we would like to show that:

$$E(.) = \frac{2T(mutual)}{T(total)} = \frac{2}{3} \tag{7}$$

with $T(mutual)$ multiplied by two as each satisfying distance implies two mutually nearest hamsters.

Rearranging (7) for convenience, and simplifying:

$$3[(n-3)\sum_{s=1}^{n-1}(n-1-s)!^{n-4}P_{s-1}+(n-1)!]=n!$$

$$3(n-3)\sum_{s=1}^{n-1}(n-1-s)!\frac{(n-4)!}{[(n-4)-(s-1)]!}=n!-3(n-1)!$$

$$3\sum_{s=1}^{n-1}(n-1-s)!\frac{(n-3)(n-4)!}{(n-3-s)!}=n(n-1)!-3(n-1)!$$

$$3(n-3)!\sum_{s=1}^{n-1}(n-1-s)(n-2-s)=(n-3)(n-1)!$$

$$\sum_{s=1}^{n-1}(n-1-s)(n-2-s)=\frac{(n-1)(n-2)(n-3)}{3} \tag{8}$$

which can be solved by splitting and applying the well-known formulae for the sum of the first n natural numbers, $\sum_{k=1}^{n}k=\frac{n(n+1)}{2}$, and the sum of the squares of the first n natural numbers, $\sum_{k=1}^{n}k^2=\frac{n(n+1)(2n+1)}{6}$. Continuing from (8):

$$\begin{aligned}
LHS &= \sum_{s=1}^{n-1}(n-2-s)^2+(n-2-s)\\
&= [(\sum_{k=1}^{n-3}k^2)+1]+[(\sum_{k=1}^{n-3}k)-1]\\
&= \frac{(n-3)(n-2)(2n-5)}{6}+\frac{(n-3)(n-2)}{2}\\
&= \frac{(2n-5)(n-2)(n-3)+3(n-2)(n-3)}{6}\\
&= \frac{2(n-1)(n-2)(n-3)}{6}=RHS \quad \square
\end{aligned} \tag{9}$$

and thus Equation (7) holds for all $n \geq 4$.

As the reasoning for the case n=3 is trivial, we have therefore proven that the expected number of hamsters that are mutually nearest to their nearest neighbour hamster is exactly $\frac{2n}{3}$ for all $n \geq 3$.