Solution

Week 38 (6/2/03)

Sum over 1

(a) **First Solution:** We will use the following fact: Given n random numbers between 0 and 1, the probability $P_n(1)$ that their sum does not exceed 1 equals $1/n!$.[1] To prove this, let us prove a slightly stronger result:

Theorem: *Given n random numbers between 0 and 1, the probability $P_n(s)$ that their sum does not exceed s equals $s^n/n!$ (for all $s \le 1$).*

Proof: Assume inductively that the result holds for a given n. (It clearly holds for all $s \le 1$ when $n = 1$.) What is the probability that $n+1$ numbers sum to no more than t (with $t \le 1$)? Let the $(n+1)$st number have the value x. Then the probability $P_{n+1}(t)$ that all $n+1$ numbers sum to no more than t equals the probability $P_n(t-x)$ that the first n numbers sum to no more than $t-x$. But $P_n(t-x) = (t-x)^n/n!$. Integrating this probability over all x from 0 to t gives

$$P_{n+1}(t) = \int_0^t \frac{(t-x)^n}{n!}\,dx = -\frac{(t-x)^{n+1}}{(n+1)!}\Big|_0^t = \frac{t^{n+1}}{(n+1)!}\,. \tag{1}$$

We see that if the theorem holds for n, then it also holds for $n+1$. Therefore, since the theorem holds for all $s \le 1$ when $n = 1$, it holds for all $s \le 1$ for any n. ■

We are concerned with the special case $s = 1$, for which $P_n(1) = 1/n!$.

The probability that it takes exactly n numbers for the sum to exceed 1 equals $1/(n-1)! - 1/n!$. This is true because the first $n-1$ numbers must sum to less than 1, and the nth number must push the sum over 1, so we must subtract off the probability that is does not.

The expected number of numbers, N, to achieve a sum greater than 1, is therefore

$$\begin{aligned} N &= \sum_2^\infty n\left(\frac{1}{(n-1)!} - \frac{1}{n!}\right) \\ &= \sum_2^\infty \frac{1}{(n-2)!} \\ &= e. \end{aligned} \tag{2}$$

Could it really have been anything else?

Second Solution: We will use the result, $P_n(s) = s^n/n!$, from the first solution. Let $F_n(s)\,ds$ be the probability that the sum of n numbers is between

[1] The number $1/n!$ is the volume of the region in n dimensions bounded by the coordinate planes and the hyperplane $x_1 + x_2 + \cdots + x_n = 1$. For example, in two dimensions we have a triangle with area 1/2; in three dimensions we have a pyramid with volume 1/6; etc.

s and $s + ds$. Then $F_n(s)$ is simply the derivative of $P_n(s)$, with respect to s. Therefore, $F_n(s) = s^{n-1}/(n-1)!$.

In order for it to take exactly m numbers for the sum to exceed 1, two things must happen: (1) the sum of the first $m-1$ numbers must equal a number, s, less than 1; this occurs with probability density $s^{m-2}/(m-2)!$. And (2) the mth number must push the sum over 1, that is, the mth number must be between $1-s$ and 1; this occurs with probability s.

The probability that it takes exactly m numbers for the sum to exceed 1 is therefore $\int_0^1 s(s^{m-2}/(m-2)!)\, ds$. The expected number of numbers, N, to achieve a sum greater than 1, therefore equals

$$\begin{aligned} N &= \sum_2^\infty m \int_0^1 \frac{s^{m-1}}{(m-2)!}\, ds \\ &= \sum_2^\infty m \frac{1}{m(m-2)!} \\ &= e. \end{aligned} \tag{3}$$

(b) **First Solution:** After n numbers have been added, the probability that their sum is between s and $s+ds$ is, from above, $F_n(s) = s^{n-1}/(n-1)!$ (for $s \leq 1$). There is a probability of s that the $(n+1)$st number pushes the sum over 1. If this happens, then (since this last number must be between $1-s$ and 1, and is evenly distributed) the average result will be equal to $s + (1 - s/2) = 1 + s/2$. The expected sum therefore equals

$$\begin{aligned} S &= \sum_1^\infty \left(\int_0^1 \left(1 + \frac{s}{2}\right) s \frac{s^{n-1}}{(n-1)!}\, ds \right) \\ &= \sum_1^\infty \frac{3n^2 + 5n}{2(n+2)!} \\ &= \sum_1^\infty \frac{3(n+2)(n+1) - 4(n+2) + 2}{2(n+2)!} \\ &= \sum_1^\infty \left(\frac{3}{2n!} - \frac{2}{(n+1)!} + \frac{1}{(n+2)!} \right) \\ &= \frac{3}{2}(e-1) - 2(e-2) + (e - 5/2) \\ &= \frac{e}{2}. \end{aligned} \tag{4}$$

Second Solution: Each of the random numbers has an average value of 1/2. Therefore, since it takes (on average) e numbers for the sum to exceed 1, the average value of the sum will be $e/2$.

This reasoning probably strikes you as being either completely obvious or completely mysterious. In the case of the latter, imagine playing a large number of games in succession, writing down each of the random numbers in one long sequence. (You can note the end of each game by, say, putting a mark after the final number, but this is not necessary.) If you play N games (with N very large), then the result from part (a) shows that there will be approximately Ne numbers listed in the sequence. Each number is a random number between 0 and 1, so the average value is $1/2$. The sum of all of the numbers in the sequence is therefore approximately $Ne/2$. Hence, the average total per game is $e/2$.