Random Numbers: Expectation of the Largest

fashionablemathematician

May 27, 2008

Problem: Given a set of N random numbers between 0 and 1, determine the expected value of the maximum of the N number.

Solution, N=2 case: We have two numbers, call them x and y. Suppose that the first number is x. Then, if $x > y$, the maximum value is itself x, and this occurs with probability x (as this is the probability that y will be between 0 and x). Otherwise, $y > x$ (occurring with probability $1 - x$). If $y > x$, than the expected value of y is $\frac{1+x}{2}$. We can then integrate these values over the interval of possible x values, $[0, 1]$:

$$< E_{max} > = \int_0^1 x^2 dx + \int_0^1 \frac{(1+x)(1-x)}{2} dx \tag{1}$$

$$= \int_0^1 \frac{x^2}{2} + \frac{1}{2} dx \tag{2}$$

$$= \frac{2}{3} \tag{3}$$

Now, this approach works just fine in the $N = 2$ case, but we can see that it will quickly become messy, and I can't see how to generalize it easily. While above we integrated over all possible values of only the first number, let's try integrating over all N values:

Solution, Arbitrary N: In general, if we integrate the value of a function over a volume and divide by the volume, we'll get the average value. Let's do that for the function $max(x_1, x_2, ..., x_N)$ over the N-dimensional unit cube (conveniently having volume 1):

$$< E_{max} >= \int_0^1 \int_0^1 ... \int_0^1 max(x_1, x_2, ..., x_N) dx_1 \; dx_2 \; ... \; dx_N \tag{4}$$

We can then break up this cube into regions, in which each of the N numbers is the maximum. Of course, if x_i is the maximum number, then $max(x_1, x_2, ...x_N) =$

x_i, giving us:

$$
\begin{aligned}
< E_{max} > &= \sum_{i=1}^{N} \int_0^1 \int_0^{x_i} \dots \int_0^{x_i} x_i dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_N dx_i && (5)\\
&= N \int_0^1 x^N dx && (6)\\
&= \frac{N}{N+1} && (7)
\end{aligned}
$$