Solution

Week 74 (2/9/04)

Comparing the numbers

(a) Let your number be n. We will average over the equally likely values of n (excluding $n = 1$) at the end of the calculation.

For convenience, let $p \equiv (n-1)/(N-1)$ be the probability that a person you ask has a number smaller than yours. And let $1-p \equiv (N-n)/(N-1)$ be the probability that a person you ask has a number larger than yours.

Let A_n be the average number of people you have to ask in order to find a number smaller than yours, given that you have the number n. A_n may be calculated as follows.

There is a probability p that it takes only one check to find a smaller number.

There is a probability $1-p$ that the first person you ask has a larger number. From this point on, you have to ask (by definition) an average of A_n people in order to find a smaller number. In this scenario, you end up asking a total of $A_n + 1$ people.

Therefore, it must be true that

$$A_n = p \cdot 1 + (1-p)(A_n + 1). \qquad (1)$$

This gives

$$A_n = \frac{1}{p} = \frac{N-1}{n-1}\,. \qquad (2)$$

All values of n, from 2 to N, are equally likely, so we simply need to find the average of the numbers A_n, for n ranging from 2 to N. This average is

$$\begin{aligned} A &= \frac{1}{N-1}\sum_{n=2}^{N}\frac{N-1}{n-1} \\ &= 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N-1}\,. \qquad (3) \end{aligned}$$

This expression for A is the exact answer to the problem. To obtain an approximate answer, note that the sum of the reciprocals of the numbers from 1 to M equals $\ln M + \gamma$, where $\gamma \equiv 0.577...$ is Euler's constant. So if N is large, you have to check about $\ln N + \gamma$ other numbers before you find one that is smaller than yours.

(b) Let your number be n. As in part (a), we will average over the equally likely values of n (excluding $n = 1$) at the end of the calculation.

Let B_n^N be the average number of people you have to ask in order to find a smaller number, given that you have the number n among the N numbers. B_n^N may be calculated as follows.

There is a probability $(n-1)/(N-1)$ that it takes only one check to find a smaller number.

There is a probability $(N-n)/(N-1)$ that the first person you ask has a larger number. From this point on, you have to ask (by definition) an average of B_n^{N-1} people in order to find a smaller number. In this scenario, you end up asking a total of $B_n^{N-1}+1$ people.

Therefore, it must be true that

$$\begin{aligned} B_n^N &= \frac{n-1}{N-1}\cdot 1 + \frac{N-n}{N-1}\Big(B_n^{N-1}+1\Big) \\ &= 1 + \left(\frac{N-n}{N-1}\right) B_n^{N-1}. \end{aligned} \tag{4}$$

Using the fact that $B_n^N = 1$ when $N = n$, we can use eq. (4) to inductively increase N (while holding n constant) to obtain B_n^N for $N > n$. If you work out a few cases, you will quickly see that $B_n^N = N/n$. We can then easily check this by induction on N; it is true for $N = n$, so we simply need to verify in eq. (4)) that

$$\frac{N}{n} = 1 + \left(\frac{N-n}{N-1}\right)\frac{N-1}{n}\,, \tag{5}$$

which is indeed true. Therefore,

$$B_n^N = \frac{N}{n}\,. \tag{6}$$

As in part (a), all values of n, from 2 to N, are equally likely, so we simply need to find the average of the numbers $B_n^N = N/n$, for n ranging from 2 to N. This average is

$$\begin{aligned} B &= \frac{1}{N-1}\sum_{n=2}^{N}\frac{N}{n} \\ &= \frac{N}{N-1}\left(\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{N}\right). \end{aligned} \tag{7}$$

This expression for B is the exact answer to the problem. If N is large, then the result is approximately equal to $\ln N + \gamma - 1$, due to the first term of "1" missing in the parentheses. This result is one person fewer than the result in part (a). So a good memory saves you, on average, one query.

Remark: The continuum version of this problem (in which case the quality of your memory in irrelevant, to leading order) is the following.

Someone gives you a random number between 0 and 1, with a flat distribution. Pick successive random numbers until you finally obtain one that is smaller. How many numbers, on average, will you have to pick?

This is simply the original problem, in the limit $N \to \infty$. So the answer should be $\ln(\infty)$, which is infinite. And indeed, from the reasoning in part (a), if you start with the number x, then the average number of picks you need to make to find a smaller number is $1/x$, from eq. (2). Averaging these waiting times of $1/x$, over the equally likely values of x, gives an average waiting time of[1]

$$\int_0^1 \frac{dx}{x} = \infty. \tag{8}$$

We can get around this infinite answer by changing the probability distribution on the unit interval. For example, let the probability distribution be proportional to $1/\sqrt{x}$. Then the probability of someone having a number smaller than x is proportional to $\int_0^x dx/\sqrt{x} \propto \sqrt{x}$. Therefore, if you start with the number x, then the average number of picks you need to make to find a smaller number is proportional to $1/\sqrt{x}$, from eq. (2). Averaging these waiting times of $1/\sqrt{x}$, over the equally likely values of x, gives an average waiting time proportional to

$$\int_0^1 \frac{1}{\sqrt{x}}\, dx \neq \infty. \tag{9}$$

In general, if the probability distribution is proportional to x^r, then $r = 0$ (that is, a flat distribution) is the cutoff case between having a finite or infinite expectation value for the number of necessary picks.

[1] However, if you play this game a few times, you will quickly discover that the average number of necessary picks is not infinite. If you find this unsettling, you are encouraged to look at Problem of the Week 6.