Purchasing Fishes

Let the path he shall walk through is from x_0 to x_1. Suppose that N fishes are uniformly distributed along the path while on sale at the same price. The true value v of these fishes can be approximated by a random variable, whose cumulative density function (CDF) are denoted with $F(v)$. The strategy is to examine the values of the first n fishes and determine the highest value v_{m} so discovered. Then he shall buy from the last $N-n$ fishes the first fish we see whose value is higher than v_m. Now let us find out the optimal n^* in terms of N and $F(\cdot)$.

Denoting v_i the value of the ith fish and $\bar{v}_{\mathrm{m}}(n)$ the expected maximal value of the first n fishes, it is obvious that we are to solve

$$\max_{n} \quad \bar{v}_{\mathrm{m}}(n) \cdot \Pr(\max_{n<i\leq N} v_i \geq \bar{v}_{\mathrm{m}}(n)). \tag{1}$$

Using the CDF, this problem is rewritten using the so-defined function $g(n)$ as

$$\max_{n} \quad g(n) \equiv \bar{v}_{\mathrm{m}}(n) \cdot \left(1 - F^{N-n}(\bar{v}_{\mathrm{m}}(n))\right). \tag{2}$$

The derivative of $g(n)$ is given by

$$\frac{\mathrm{d}g(n)}{\mathrm{d}n} = \left[1 - F^{N-n}(\bar{v}_{\mathrm{m}}) - (N-n)\bar{v}_{\mathrm{m}} F^{N-n-1}(\bar{v}_{\mathrm{m}})\right] \frac{\mathrm{d}\bar{v}_{\mathrm{m}}}{\mathrm{d}n}, \tag{3}$$

where the dependence of $\bar{v}_{\mathrm{m}}$ on n is omitted for clarity. The expected maximal value of n fishes is given by the integral

$$\bar{v}_{\mathrm{m}}(n) = \int v \frac{\mathrm{d}}{\mathrm{d}v} F^n(v) \mathrm{d}v. \tag{4}$$

This is a monotonically increasing function, as is evidenced by

$$\frac{\mathrm{d}\bar{v}_{\mathrm{m}}(n)}{\mathrm{d}n} = \int \frac{n \ln F(v) + 1}{n} v \frac{\mathrm{d}}{\mathrm{d}v} F^n(v) \mathrm{d}v > 0. \tag{5}$$

Thus, the behavor of $g(n)$ in variation with n depends solely on the terms in the brackets in Eq.3. The maximum of $g(n)$ is reached when the brackets become zero:

$$1 - F^{N-n}(\bar{v}_{\mathrm{m}}(n^*)) - (N-n^*)\bar{v}_{\mathrm{m}}(n^*) F^{N-n-1}(\bar{v}_{\mathrm{m}}(n^*)) = 0. \tag{6}$$

To obtian the value of n^* requires the value of N, the form of $F(\cdot)$, and a modern computer for numerically solving this equation, which is unfortunately not available in his time.

However, a crude analysis can be given as follows. For a given n, the second term in the bracket in Eq.3 decrease exponentially with N. This is also true for the third term when N is large enough. Thus, the sum of the two terms can be less than one for some n and the expected value of fish $g(n)$ can increase in some interval. However, this interval does not extend too much beyond the beginning of the whole range. To see this, it is convinent to express in the brackets the n in terms of N as $n = \alpha N$ with $\alpha \in [0,1]$ and obtain

$$h(\alpha, N) \equiv 1 - F^{(1-\alpha)N}(\bar{v}_{\mathrm{m}}) - (1-\alpha)N\bar{v}_{\mathrm{m}} F^{(1-\alpha)N-1}(\bar{v}_{\mathrm{m}}), \tag{7}$$

Here we choose a representative distribution for v: it is uniformly distributed in $[0,1]$ and hence $F(v) = v$. It is easy to find from Eq.4 $F(\bar{v}_{\mathrm{m}}) = \bar{v}_{\mathrm{m}} = \frac{n}{n+1} = \frac{\alpha N}{\alpha N+1}$. Upon substitution of this, Eq.7 becomes

$$h(\alpha, N) = 1 - [1 + (1-\alpha)N] \left(\frac{\alpha N}{\alpha N + 1}\right)^{(1-\alpha)N}. \tag{8}$$

This expression is always negative when N is large enough:

$$h(\alpha, N) = 1 - [1 + (1-\alpha)N] \, e^{1-\frac{1}{\alpha}} < 0, \tag{9}$$

which implies that the best n^* does not increase proportionally with N but lags somewhat behind. Therefore, it is optimal to inspect the very beginning part of the candidates and subsequently make a superior selection.