# The best choice problem for random number of objects with a refusal probability

M Horiguchi[1] M Yasuda[2]

[1]Kanagawa University

[2]Chiba University

July 2009, APS at Cornell

# Abstract:

The best choice problem by Markov Potential and Differential equation method.
Here we dare to try a classical analysis to obtain its optimal value for the problem on Markov processes.

(1) Optimaility equation(Non-linear Differentila equation) with random number object.

# Abstract:

The best choice problem by Markov Potential and Differential equation method.
Here we dare to try a classical analysis to obtain its optimal value for the problem on Markov processes.

(1) Optimaility equation(Non-linear Differentila equation) with random number object.

(2) Infinitesimal approach(scaling limit).

# Abstract:

The best choice problem by Markov Potential and Differential equation method.
Here we dare to try a classical analysis to obtain its optimal value for the problem on Markov processes.

(1) Optimaility equation(Non-linear Differentila equation) with random number object.
(2) Infinitesimal approach(scaling limit).
(3) In generally case, it is hard to solve analitically (Lambert W-function).

- Markov Decision Processes
- One-Look-Ahead Policy
- Markov Potential Theory
- Infinitesimal Analysis(Scaling limit)
- Variations from the classical problem Random Number object, Refusal probability, Zero-sum Game version, etc.

- Chow/Robbins/Siegmund(1971), Siryaev(1973), Dynkin/Yushkevitch(1969): monotone stopping problem, the martingale theory, the potential theory,
- Ross(1983); One-Look ahead policy, the optimality principle in the mathematical programming,
- Ferguson(2006), Bruss(2003), Tamaki(2002): the sum-the-odds theorem.

The last thing one knows when writing a book is what to put first.
Blaise Pascal(1623-1662)
*from J.Havil "Gamma" exploring euler's constant with a foreword by Freeman Dyson*

Let $x_n, n = 0, 1, 2, \cdots$ be a Markov Chain with a transition probability $P = P(i, j), i, j \in S$ over a state space $S$ in $R^1$.

$$v(i) = \sup_{\tau<\infty} v(i,\tau) = \sup_{\tau<\infty} E\left[ r(x_\tau) - \sum_{n=0}^{\tau} c(x_n) \;\middle|\; x_0 = i \right]$$

The relation between Markov potential theory and Dynamic Programming or Markov decision processes is well known. The optimality equation* is reduced as follows.

$$v(i) = \max\{r(i), Pv(i) - c(i)\}, i \in S.$$

Consider the set of states for which stopping immediately is at least as gppd as stopping after exactly one more period. Denote this set by

$$B = \{i \in S \mid r(i) \geq Pr(i) - c(i)\}$$

where $Pr(i) = \sum_{j \in S} P(i,j)r(j)$.

The policy, defined by hitting time of this set $B$, is called a *One-stage Loog Ahead*(abridged as OLA) policy, refer to Ross(1983).
It is sufficient that the OLA policy is optimal if $B$ is closed and eventually hits to the set. That the set is closed means if

$$P(i,j) = 0 \quad \text{for} \quad i \in B, j \in \overline{B}$$

where $\overline{B}$ denotes the complement of the set $B$.

Markov stopping theory as the excessive function, the martingale theory and so on.
However the general framework does not cover for specific solution in the exploit structure or interesting features.
The aim of this note is to obtain the explicit optimal value associated with OLA policy and various problems in the asymptotic form.

Firstly the motivated case for the classical secreraly choise problem and then for the variation of the case with the refusal probability. Let $P_A$ denote the restriction to a subset $A$ of S for the transition probability $P$, that is,

$$P_A(i,j) = P(i,j)1_A(i)$$

where the indicator of $A$ is

$$1_A(i) = \begin{cases} 1 & (i \in A) \\ 0 & (i \notin A) \end{cases}$$

A potential $N$ of $P$ for a function $f$:

$$Nf(i) = \lim_n \left[ \{I + P + P^2 + \cdots + P^k\} f \right](i) \quad i \in S.$$

Assume that

$$\lim_k \left[ (P_{\overline{B}})^k r \right](i) = 0$$

$$[N_{\overline{B}}(P_B r - c)](i) < \infty \quad \text{for} \quad i \notin B.$$

Then

(i) the OLA policy $\tau_B$ is optimal,

(ii) the optimal value $v(i), i \in S$ is given by

$$\begin{aligned} v(i) &= r(i) + N(Pr - r - c)^+(i) \\ &= \begin{cases} r(i) & \text{on} \quad B \\ N_{\overline{B}}(P_B r - c)(i) & \text{on} \quad \overline{B} \end{cases} \end{aligned}$$

The secretary problem is an optimal stopping problem based on relative ranks for objects for objects arriving in a random fashion. The objective is to find the stopping rule that maximizes the probability of stopping at the best object among the sequence. The formulation of the problem as Markov processes is given by Dynkin/Yushkevitch(1969).

The state space is $S = \{1, 2, 3, \cdots, n\}$ and the reward, its win of probability is $r(i) = i/n, i \in S$.
The transition probability is

$$P(i,j) = \frac{i}{j(j-1)},$$

on $1 \leq i < j \leq n$, and $= 0$, otherwise.
The optimal value $v(i)$ satisfies, for $i = 1, 2, \cdots, n-1,$,

$$v(i) = \max\left\{\frac{i}{n}, \sum_{j=i+1}^{n} \frac{i}{j(j-1)} v(j)\right\},$$
$$v(n) = 1.$$

Taking a scale limit of $\frac{i}{n} \to x$ as $i, n \to \infty$, it becomes that

$$v(x) = \max\left\{x, x\int_x^1 \frac{v(y)}{y^2}\,dy\right\}, \quad x \in S = [0,1],$$
$$v(1) = 1.$$

The solution is well known as

$$v(x) = \begin{cases} e^{-1} & on \quad [0, e^{-1}], \\ x & on \quad [e^{-1}, 1]. \end{cases}$$

By extending the reward $r(x)$ in general, that is,

$$v(x) = \max\{r(x), x \int_x^1 \frac{v(y)}{y^2}\, dy\}, \quad x \in S = [0,1],$$
$$v(1) = 1.$$

It is also the OLA policy is still optimal if the following assumption holds.

Assumption:

The function

$$h(x) = r(x) - x \int_x^1 \frac{r(y)}{y^2}\, dy$$

is finite on $[0, 1]$ for a given $r(x)$ and
it changes its sign from $-$ to $+$ only once as $x$ increases $0$ to $1$.
So there exists a unique solution $\alpha$ of the equation $h(x) = 0$, that is,

$$\alpha = \inf\{x; h(x) \geq 0\}.$$

Theorem:

Using this solution $\alpha$, the assumption implies

$$v(x) = \begin{cases} r(\alpha) & on \quad [0, \alpha], \\ r(x) & on \quad [\alpha, 1]. \end{cases}$$

Because of using Markov potential, it is calculated as

$$P_B r(x) = r(\alpha)\frac{x}{\alpha},$$

$$\left(P_{\overline{B}}\right)^n P_B r(x) = r(\alpha)\frac{x}{\alpha}\log^n\left(\frac{\alpha}{x}\right)/n!$$

and

$$N_{\overline{B}} P_B r(x) = r(\alpha)$$

directly. Here $B = [0, \alpha]$.

The basic idea based on the infinitesimal operator of Markov processes.
It is equivalent to,

$$-x\frac{d\,v(x)}{dx} = \left(r(x) - v(x)\right)^{+}, \quad x \in [0,1], \qquad v(1) = 1.$$

This method is derived by Mucci(1973).

A variant for the classical best choice problem is considered by Smith(1975) inducing a refusal probability.
A given $p$ $(0 < p < 1)$ means the probability of not refused case and the quality $1 - p$ is a probability of the refusal. When there occues no refusal$(1 - p = 0)$, that is, the usual standard problem corresponds to $p = 1$.

The asymptotic form of the transition probability is given as

$$P(x, dy) = \begin{cases} \dfrac{p}{y}\left(\dfrac{x}{y}\right)^p dy, & 0 < x < y < 1, \\ 0, & otherwise. \end{cases}$$

Similarly before

Assumption: The function

$$h_p(x) = r(x) - p \int_x^1 \frac{r(y)}{y} (x/y)^p \, dy$$

is finite on $[0, 1]$ for a given $r(x)$ and it changes its sign from $-$ to $+$ only once as $x$ increases $0$ to $1$. Its unique solution is denoted by $\alpha_p$ of the equation $h_p(x) = 0$, i.e.,

$$\alpha_p = \inf\{x; h_p(x) \geq 0\}.$$

Theorem: Using this solution $\alpha_p$, the assumption implies that

$$v(x) = \begin{cases} r(\alpha_p) & on \quad [0, \alpha_p], \\ r(x) & on \quad [\alpha_p, 1]. \end{cases}$$

It is also known the technique of using a differential equation by Mucci(1973).
Let

$$V(x) = p \int_x^1 \frac{r(y)}{y} (x/y)^p \, dy, \qquad x \in [0, 1]$$

which means a contditional optimal value.
This satisfies

$$\begin{cases} -x \dfrac{dV(x)}{dx} = p\big(r(x) - V(x)\big)^+, & x \in [0, 1], \\ V(1) = 0. \end{cases}$$

The optimal value at the begining

$$v^* = V(0) = r(\alpha_p).$$

Confirming that the case of $r(x) = x$, the equation $h_p(x) = 0$ solves

$$x = \alpha_p = p^{1/(1-p)}$$

which consist with the result by Smith(1975).

The discussion on the stopping choice problem with a random number of objects is firstly by Presman/Sonin(1972). Then Rasmussen/Robbins(1975), Tamaki(1979), Irle(1980), Petruccelli(1983) etc.
If we restrict ourselves that the support of distribution for a number of objects is bounded, the technique of the previous scaling limit could be applied.
This assumption links to the technique of the scaling limit. It should be noted that Presman/Sonin discussed the number of objects is Poisson distribution, so we could not apply this technique in that case. The next is a fundamental assumption of the discussion in here.

Assumption:

A random number of objects $N$ is bounded and let

$$n = \inf\{k \geq 1; P(N > k) = 0\}.$$

By using the notation of $p_i = P(N = i)$ and $\pi_i = P(N \geq i) = \sum_{k \geq i} p_i$, the transition probability matrix for this case is as follows:

$$p(i,j) = \frac{i\,\pi_j}{j(j-1)\,\pi_i}, \quad 1 \leq i < j \leq n,$$

$$p(i,n) = \sum_{k=i+1}^{n} \frac{i\,p_k}{k\,\pi_i}, \quad 1 \leq i < n, \quad p(n,n) = 1,$$

$$r(i) = \sum_{k=i}^{n} \frac{i\,p_k}{k\,\pi_i}, \qquad c(i) = 0, \ \forall i.$$

If we denote the distribution by $\Phi(x), x \in [0,1]$ of its random number for objects, and taking the scale limit of $k, n \to \infty$ with $k/n = x$, then $v(i) \to V(x)$ and $p_k = P(N = i) \to d\Phi(x)$, $\pi_k = \sum_{k \geq i} p_i \to 1 - \Phi(x)$, so

$$p_k/\pi_k \to d\Phi(x)/(1 - \Phi(x)),$$

$$r(k+1) = \sum_{k=i}^{n} \frac{i\, p_k}{k\, \pi_i} = \frac{i}{\pi_i} \sum_{k=i}^{n} \frac{p_k}{k}$$

$$\to R(x) = \frac{x}{1 - \Phi(x)} \int_x^1 y^{-1}\, d\Phi(y).$$

The resulting the differential equation is given as

$$\begin{cases} -x\,dV(x) + x\,\dfrac{V(x)}{1-\Phi(x)}\,d\Phi(x) \\ \qquad\qquad = (R(x) - V(x))^{+}\,dx, \quad x \in [0,1] \\ V(1) = 0 \end{cases}$$

where

$$R(x) = \frac{x}{1-\Phi(x)} \int_x^1 y^{-1}\,d\Phi(y).$$

The details refer to Yasuda(1984).

Assumption:

We assume that $d\Phi$ satisfies the conditions:

1. $(1-\Phi(x))^{-1}\int_x^1 y^{-1}d\Phi(y) \to 1 \quad as \quad x \to 1,$
2. $x\int_x^1 y^{-1}d\Phi(y) \to 0 \quad as \quad x \to 0.$

This assumtion implies that $R(x) \to 0\,(x \to 0)$ and $\to 1(x \to 1)$. In the usual non-random case, $R(x) = x$, these conditions are satisfied since $\dfrac{1}{1-\Phi(x)} \displaystyle\int_x^1 y^{-1}\, d\Phi(y) = 1$ for $x \neq 1$.

Example:

Example of a random number $N$:
Uniformly distributed case on a partial interval

$$\{n-m, n-m+1, \cdots, n-1, n\} \text{ of } \{1, 2, \cdots, n\}$$

Letting $\theta = m/n$ is fixed for $n, m \to \infty$,

$$\Phi(x) = \begin{cases} 0, & (0 < x < 1-\theta), \\ (x-1+\theta)/\theta, & (1-\theta < x < 1) \end{cases}$$

| case | $\alpha$: threshold | $v^*$:optimal value |
|---|---|---|
| $1-\theta \geq e^{-2}$ | $\sqrt{1-\theta}e^{-1}$ | $-\frac{\sqrt{1-\theta}}{\theta}\log(1-\theta)e^{-1}$ |
| $1-\theta \leq e^{-2}$ | $e^{-2}$ | $2e^{-2}/\theta$ |

If $\theta \to 1$ it tend to $2e^{-2}$ as discussed in Presman/Sonin(1972) etc. Alternatively, if $\theta \to 0$, this reduces to the non-random classical problem and it tend to $e^{-1}$ as is well known. However it is interesting that the threshold $\alpha = e^{-2}$ does not depent on $\theta$ in case of $1 - \theta \approx 0$, that is, case of the ambiguity is being in spread.

Here we would like to impose a refusal probability for the model. Instead of $V(x)$ in the optimality equation consider $dv(x) = dV(x) - \dfrac{V(x)}{1-\Phi(x)}\, d\Phi(x)$. Then, it holds that

$$v(x) = \max\left\{R(x),\ \frac{x}{1-\Phi(x)}\int_x^1 (1-\Phi(y))y^{-2}v(y)\,dy\right\}$$

The optimality equation of stopping problem:

$$v(x) = \max\{R(x), Pv(x)\},\ 0 \le x \le 1$$

have a modification with a refusal probability $p(x), 0 \le x \le 1$ becomes

$$v(x) = \max\{p(x)\, R(x) + (1 - p(x))\, Pv(x),\ \ Pv(x)\}.$$

By letting

$$Pv(x) = \frac{x}{1-\Phi(x)} \int_x^1 (1-\Phi(y))y^{-2}v(y)\,dy,$$

$$q(x) = \exp\left(\int_x^1 y^{-1}(1-p(x))\,dy\right)$$

and

$$h(x) = \int_x^1 y^{-1}\,d\Phi(y),$$

Assumption:

Assume that

$$H_p(x) = h(x) - q(x) \int_x^1 \frac{h(y)\,p(y)}{y\,q(y)} dy$$

changes its sign once from $-$ to $+$ as $x$ increase.

Theorem:
The solution is given as the value

$$\alpha_p^* = \begin{cases} \inf\{x; H_p(x) \geq 0\} \\ 1 \quad \textit{if empty} \end{cases}$$

determine its threshold policy and so the optimal value $v(0+)$.

However in order to display the above assertion or explicit form, it is not easy for us.
Even the simplest case of

$$p(x) = p\,(0 < p < 1)$$

and

$$\Phi(x) = x\,(0 \leq x \leq 1),$$

it needs the discussion on Lambert W function.