Hindawi Publishing Corporation
International Journal of Stochastic Analysis
Volume 2012, Article ID 697458, 8 pages
doi:10.1155/2012/697458

Research Article

Bayes' Model of the Best-Choice Problem with Disorder

Vladimir Mazalov and Evgeny Ivashko

Institute of Applied Mathematical Research, Karelian Research Centre of RAS, IAMR KRC RAS 11, Pushkinskaya Street, Petrozavodsk, Karelia 185910, Russia

Correspondence should be addressed to Evgeny Ivashko, ivashko@krc.karelia.ru

Received 10 December 2010; Accepted 21 March 2011

Academic Editor: Onesimo Hernandez Lerma

We consider the best-choice problem with disorder and imperfect observation. The decision-maker observes sequentially a known number of i.i.d random variables from a known distribution with the object of choosing the largest. At the random time the distribution law of observations is changed. The random variables cannot be perfectly observed. Each time a random variable is sampled the decision-maker is informed only whether it is greater than or less than some level specified by him. The decision-maker can choose at most one of the observation. The optimal rule is derived in the class of Bayes' strategies.

1. Introduction

In the papers we consider the following best-choice problem with disorder and imperfect observations. A decision-maker observes sequentially n iid random variables $\xi_1, \dots, \xi_{\theta-1}, \xi_\theta, \dots, \xi_n$. The observations $\xi_1, \dots, \xi_{\theta-1}$ are from a continuous distribution law $F_1(x)$ (state S_1). At the random time θ, the distribution law of observations is changed to continuous distribution function $F_2(x)$ (i.e., the disorder happen—state S_2). The moment of the disorder has a geometric distribution with parameter $1 - \alpha$. The observer knows parameters α, $F_1(x)$, and $F_2(x)$, but the exact moment θ is unknown.

At each time in which a random variable is sampled, the observer has to make a decision to *accept* (and stop the observation process) or *reject* the observation (and continue the observation process). If the decision-maker decided to *accept* at step k $(1 \le k \le n)$, she receives as the payoff the value of the random variable discounted by the factor λ^{k-1}, where $0 < \lambda < 1$. The random variables cannot be perfectly observed. The decision-maker is only informed whether the observation is greater than or less than some level specified by her.

The aim of the decision-maker is to maximize the expected value of the accepted discounted observation.

We find the solution in the class of the following strategies. At each moment k ($1 \leq k \leq n$), the observer estimates the *a posterior* probability of the current state and specifies the threshold $s = s_{n-k}$. The decision-maker accepts the observation x_k if and only if it is greater than the corresponding threshold s.

This problem is the generalization of the best-choice problem [1, 2] and the quickest determination of the change-point (disorder) problem [3–5]. The best-choice problems with imperfect information were treated in [6–8]. Only few papers related to the combined best-choice and disorder problem are published [9–11]. Yoshida [9] considered the full-information case and found the optimal stopping rule which maximizes the probability that accepted value is the largest of all $\theta + m - 1$ random variables for a given integer m. Closely related work to this study is Sakaguchi [10] where the optimality equation for the optimal expected reward is derived for the full-information model. In [11], we constructed the solution of the combined best-choice and disorder problem in the class of single-level strategies, and, in this paper, we search the Bayes' strategy which maximizes the expected reward in the model with imperfect observation.

2. Optimal Strategy

According to the problem the observer does not know the current state (S_1 or S_2). But she can estimate the state using the Bayes' formula:

$$\pi_s = \pi(s) = P\{S_1 \mid x \leq s\} = \frac{P(S_1)P(x \leq s \mid S_1)}{P(x \leq s)} = \frac{\alpha\pi F_1(s)}{F_\pi(s)}. \tag{2.1}$$

Here, $s = s_i$ is the threshold specified by the decision-maker within i steps until the end (i.e., at the step $n - i$), π is the *a prior* probability of the state S_1 (i.e., before getting the information that $x \leq s$), $F_\pi(s) = \pi F_1(s) + \overline{\pi} F_2(s)$, and $\overline{\pi} = 1 - \pi$.

We use the dynamic programming approach to derive the optimal strategy. Let $v_i(\pi)$ be the payoff that the observer expects to receive using the optimal strategy within steps until the end. The optimality equation is as follows:

$$\begin{gathered} v_i(\pi) = \max_s E\left[\lambda v_{i-1}(\pi_s) I_{\{x \leq s\}} + x I_{\{x > s\}}\right], \quad i \geq 1, \\ v_0(\pi) = 0 \quad \forall \pi. \end{gathered} \tag{2.2}$$

Simplifying (2.2), we get

$$\begin{gathered} v_i(\pi) = \max_s [\lambda v_{i-1}(\pi_s) F_\pi(s) + \pi E_1(s) + \overline{\pi} E_2(s)], \quad i \geq 1, \\ v_0(\pi) = 0 \quad \forall \pi. \end{gathered} \tag{2.3}$$

Here, $E_k(s) = \int_s^\infty x dF_k(x)$, $k = 1, 2$.

The following theorem gives the presentation of the expected payoff in linear form on π.

Theorem 2.1. *For any i the function $v_i(\pi)$ can be written in the form*

$$v_i(\pi) = \pi A_i(s_1, \dots, s_i) + B_i(s_1, \dots, s_i), \tag{2.4}$$

where

$$s_i = s_i(\pi) = \arg\max_s [\lambda v_{i-1}(\pi_s) F_\pi(s) + \pi E_1(s) + \overline{\pi} E_2(s)], \quad i \geq 1, \ 0 \leq \pi \leq 1. \tag{2.5}$$

Proof. Using the formula (2.3), one can show that

$$v_1(\pi) = \max_s [\pi(E_1(s) - E_2(s)) + E_2(s)] = \pi A_1(s_1) + B_1(s_1), \tag{2.6}$$

where $A_1(s_1) = E_1(s_1) - E_2(s_1)$, $B_1 = E_2(s_1)$ and

$$s_1 = s_1(\pi) = \arg\max_s [\pi(E_1(s) - E_2(s)) + E_2(s)], \quad 0 \leq \pi \leq 1. \tag{2.7}$$

Threshold $s_1 = s_1(\pi)$ is the solution of (2.3) for $0 \leq \pi \leq 1$ for $i = 1$.

Assume the theorem is correct for certain $i = k$. Then, for $i = k + 1$

$$\begin{aligned} v_{k+1}(\pi) &= \max_s [\lambda(\pi_s A_k(s_1, \dots, s_k) + B_k(s_1, \dots, s_k)) F_\pi(s) + \pi E_1(s) + \overline{\pi} E_2(s)] \\ &= \max_s [\pi(\lambda \alpha F_1(s) A_k(s_1, \dots, s_k) + \lambda B_k(s_1, \dots, s_k)(F_1(s) - F_2(s)) + E_1(s) - E_2(s)) \\ &\qquad + \lambda B_k(s_1, \dots, s_k) F_2(s) + E_2(s)] \\ &= \pi A_{k+1}(s_1, \dots, s_{k+1}) + B_{k+1}(s_1, \dots, s_{k+1}), \end{aligned} \tag{2.8}$$

where

$$\begin{gathered} A_{k+1}(s_1, \dots, s_{k+1}) = \lambda \alpha F_1(s) A_k(s_1, \dots, s_k) + \lambda B_k(s_1, \dots, s_k)(F_1(s) - F_2(s)) + E_1(s) - E_2(s), \\ B_{k+1}(s_1, \dots, s_{k+1}) = \lambda B_k(s_1, \dots, s_k) F_2(s) + E_2(s), \\ s_i = s_i(\pi) = \arg\max_s [\lambda v_{i-1}(\pi_s) F_\pi(s) + \pi E_1(s) + \overline{\pi} E_2(s)], \quad i \geq 1, \ 0 \leq \pi \leq 1. \end{gathered} \tag{2.9}$$

The theorem is proved. □

The following lemma takes place.

Lemma 2.2. *Assuming $E_k < \infty$, $k = 1, 2$ as $i \to \infty$, there is a limit of the expected payoff $v_i(\pi) \to v(\pi)$.*

Proof. It is obvious that the sequence $v_i(\pi)$ is increasing by i.

Now, we prove that the sequence of the expected payoffs has an upper bound.

$$\begin{aligned} v_1(\pi) &\le \pi E_1 + \overline{\pi} E_2, \\ E_k &= \int_0^\infty x dF_k(x), \quad k = 1,2 \\ v_2(\pi) &= \max_s [\lambda v_1(\pi_s) F_\pi(s) + \pi E_1(s) + \overline{\pi} E_2(s)] \\ &\le \lambda(\pi E_1 + \overline{\pi} E_2) + \pi E_1 + \overline{\pi} E_2. \end{aligned} \tag{2.10}$$

Further one can show using the induction that for any $i \ge 1$ and any $0 \le \pi \le 1$ the expected payoff at the step i has the upper bound

$$v_i(\pi) \le \frac{\pi E_1 + \overline{\pi} E_2}{1-\lambda}. \tag{2.11}$$

The lemma is proved. □

Corollary 2.3. *Theorem 2.1 and the lemma yield that there are such A and B that*

$$\lim_{i\to\infty} v_i(\pi) = \lim_{i\to\infty} (\pi A_i(s_1,\ldots,s_i) + B_i(s_1,\ldots,s_i)) = \pi A + B = v(\pi). \tag{2.12}$$

As $i \to \infty$ the expected payoff satisfies the following equation:

$$v(\pi) = \lim_i v_i(\pi) = \max_s [\lambda v(\pi_s) F_\pi(s) + \pi E_1(s) + \overline{\pi} E_2(s)]. \tag{2.13}$$

To find the components of the expected payoff for a case of huge number of observation we should solve the following equation:

$$\pi A + B = \max_s [\pi(\lambda\alpha F_1(s) A + \lambda B(F_1(s) - F_2(s)) + E_1(s) - E_2(s)) + \lambda B F_2(s) + E_2(s)], \tag{2.14}$$

therefore,

$$\begin{aligned} A &= \lambda\alpha F_1(s) A + \lambda B(F_1(s) - F_2(s)) + E_1(s) - E_2(s), \\ B &= \lambda B F_2(s) + E_2(s). \end{aligned} \tag{2.15}$$

The solution of the system is as follows

$$\begin{aligned} A &= \frac{E_1(s)(1-\lambda F_2(s)) - E_2(s)(1-\lambda F_1(s))}{(1-\lambda F_2(s))(1-\lambda\alpha F_1(s))}, \\ B &= \frac{E_2(s)}{1-\lambda F_2(s)}. \end{aligned} \tag{2.16}$$

The expected payoff is

$$v(\pi) = \max_{s}(\pi A + B) \tag{2.17}$$

and the optimal threshold is

$$s = s(\pi) = \arg\max_{s}(\pi A + B). \tag{2.18}$$

The above results are summarized in the following theorem.

Theorem 2.4. *For $i \to \infty$, the solution of* (2.3) *is defined as*

$$v(\pi) = \max_{s}(\pi A + B), \tag{2.19}$$

where

$$\begin{gathered} s = s(\pi) = \arg\max_{s}(\pi A + B), \\ A = \frac{E_1(s)(1-\lambda F_2(s)) - E_2(s)(1-\lambda F_1(s))}{(1-\lambda F_2(s))(1-\lambda\alpha F_1(s))}, \\ B = \frac{E_2(s)}{1-\lambda F_2(s)}. \end{gathered} \tag{2.20}$$

3. Examples

Consider the examples of using the Bayes' strategy B defined by the formula (2.18) comparing with two strategies with constant thresholds that do not depend on π.

3.1. Normal Distribution

Consider the example of the normal distribution of the random variables where functions $F_1(x)$ and $F_2(x)$ have the variance $\sigma^2 = 1$ and the expectation $\mu_1 = 10$ and $\mu_2 = 9$, respectively.

Strategies A_1 and A_2 with constant thresholds defined by the following formula:

$$s = \frac{E(s)}{1-\lambda F(s)}, \tag{3.1}$$

where $F(s) \equiv F_1(s)$ and $E(s) \equiv E_1(s)$ for the strategy A_1; $F(s) \equiv F_2(s)$ and $E(s) \equiv E_2(s)$ for the strategy A_2.

The values of the thresholds of strategies A_1 and A_2 depending on discount rate are tabulated in Table 1.

Table 1 shows how much the discount rate is affect on the thresholds.

Figure 1 shows the graphics of the optimal thresholds for strategies A_1 and A_2 (s_1 and s_2, resp.) and strategy B (s_{opt}) depending on π. As the figure shows, the strategy B depends

Table 1: The values of the thresholds of strategies A_1 and A_2.

λ	Strategy A_1	Strategy A_2
0.99	10.851	9.902
0.9	9.088	8.210
0.7	7.000	6.300

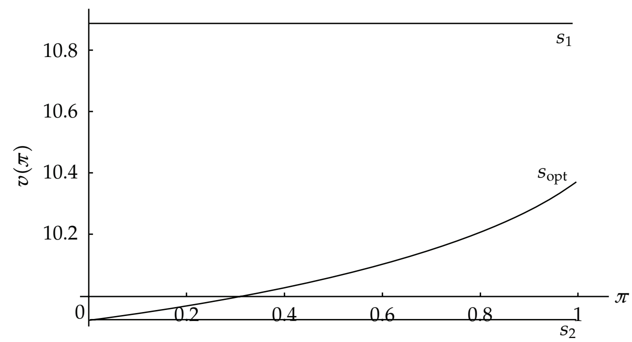

Figure 1: Graphics of the optimal thresholds for strategies A_1, A_2, and B for $\alpha = 0.9$, $\lambda = 0.99$.

Table 2: Main characteristics of the best-choice process for $\alpha = 0.9$, $\lambda = 0.99$.

Characteristic	Strategy A_1	Strategy A_2	Strategy B
Expected payoff	10.035	10.429	10.500
Average time of accepting the observation	14.526	2.472	3.072
Average number of steps after the disorder	30.406	4.503	5.031
Number of the values accepted before the disorder, %	64.100	83.066	79.738

on the *a posterior* probability of the state $S_1(\pi)$. As π tends to zero, the optimal threshold of the strategy B tends to threshold s_2.

We compare the payoffs that the observer expects to receive using different strategies. Define V_α as the expected payoff for $\pi = 1$ and depending on probability of disorder α.

Figure 2 shows the numerical results of the expected payoffs of the observer who uses the strategies A_1, A_2, and B (thresholds s_1, s_2, and s_{opt}, resp.).

The expected payoff of the observer who uses the Bayes' strategy B is greater if she uses one of the strategies A_1 or A_2. The difference is significant for $\alpha \in [0.75, 0.98]$, because of uncertainty of the current state of the system.

Table 2 shows the numerical results of the main characteristics of the best-choice process.

For the small probability of the disorder $(1 - \alpha = 0.1)$, the expected payoff according to the strategy A_2 is greater (10.429) than according to the strategy A_1 (10.035). But the Bayes' strategy B that depends on π gives the largest expected payoff (10.500).

Table 2 shows that the average time of accepting the observation is increasing with respect to the value of the threshold. Note that the strategy A_1 does not depend on the disorder and this leads to a high value of the average time of accepting the observation. Both strategies A_2 and B have a small average time of accepting the observation.

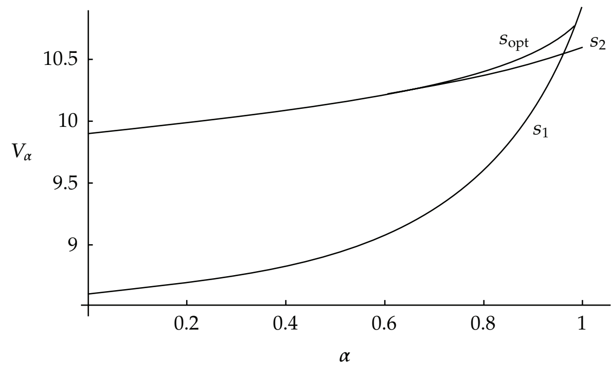

Figure 2: Expected payoffs of the observer who uses the strategies A_1, A_2 and B for $\alpha = 0.9$, $\lambda = 0.99$.

Table 3: The values of the thresholds of strategies A_1 and A_2.

λ	Strategy A_1	Strategy A_2
0.99	6.756	3.378
0.9	3.358	1.679

Table 4: Main characteristics of the best-choice process for $\alpha = 0.9$, $\lambda = 0.99$.

Characteristic	Strategy A_1	Strategy A_2	Strategy B
Expected payoff	2.355	4.438	4.499
Average time of accepting the observation	678.930	15.397	16.923
Average number of steps after the disorder	856.535	29.110	29.610
Number of the values accepted before the disorder, %	21.57	70.89	56.01

3.2. *Exponential Distribution*

Consider the example of the exponential distribution of the observations. Let $F_1(x)$ and $F_2(x)$ have the exponential distribution with parameters $\lambda_1 = 0.5$ and $\lambda_2 = 1$, respectively. As in the previous example, consider the strategies A_1 and A_2 comparing with the Bayes' strategy B,

$$s = \frac{E(s)}{1 - \lambda F(s)}, \tag{3.2}$$

where $F(s) \equiv F_1(s)$ and $E(s) \equiv E_1(s)$ for the strategy A_1; $F(s) \equiv F_2(s)$ and $E(s) \equiv E_2(s)$ for the strategy A_2.

Table 3 shows the values of the thresholds for the strategies A_1 and A_2 depending on the discount rate.

The value of the optimal threshold of the strategy B as in the case of the normal distribution of the observations is increasing by π and equal to the threshold of the strategy A_2 at $\pi = 0$. The graphics of the expected payoffs have the same view as in Figure 2. Table 4 shows the main characteristics of the best-choice process for different strategies.

As in the previous example, the Bayes' strategy gives better payoff than the strategy A_2, but it has bigger average time of accepting the observation. The strategy A_1 is the worst for all the parameters.

4. Results

In the article, we consider the best-choice problem with disorder and imperfect observations. We propose the Bayes' strategy where the threshold depends on the *a posterior* probability of the disorder. The numerical results show that this strategy gives better expected payoff than the constant strategies.

Acknowledgment

The paper is supported by grants of Russian Fund for Basic Research, Project 10-01-00089-a and Division of Mathematical Sciences, Program "Mathematical and algorithmic Problems of New Information Systems".

References

[1] J. P. Gilbert and F. Mosteller, "Recognizing the maximum of a sequence," *Journal of the American Statistical Association*, vol. 61, pp. 35–73, 1966.
[2] B. A. Berezovskiĭ and A. V. Gnedin, *The Problem of Optimal Choice*, Nauka, Moscow, Russia, 1984.
[3] A. N. Širjaev, *Statistical Sequential Analysis*, vol. 38, American Mathematical Society, Providence, RI, USA, 1973.
[4] T. Bojdecki, "Probability maximizing approach to optimal stopping and its application to a disorder problem," *Stochastics*, vol. 3, no. 1, pp. 61–71, 1979.
[5] K. Szajowski, "On a random number of disorders. Forthcoming," in *Probability and Mathematical Statistics*, 2011.
[6] E. G. Enns, "Selecting the maximum of a sequence with imperfect information," *Journal of the American Statistical Association*, vol. 70, no. 351, pp. 640–643, 1975.
[7] P. Neumann, Z. Porosiński, and K. Szajowski, "On two person full-information best choice problem with imperfect observation," in *Game Theory and Applications*, vol. 2, pp. 47–55, Nova Science Publishers, Hauppauge, NY, USA, 1996.
[8] Z. Porosiński and K. Szajowski, "Modified strategies in two person full-information best choice problem with imperfect observation," *Mathematica Japonica*, vol. 52, no. 1, pp. 103–112, 2000.
[9] M. Yoshida, "Probability maximizing approach to a secretary problem with random change-point of the distribution law of the observed process," *Journal of Applied Probability*, vol. 21, no. 1, pp. 98–107, 1984.
[10] M. Sakaguchi, "A best-choice problem for a production system which deteriorates at a disorder time," *Scientiae Mathematicae Japonicae*, vol. 54, no. 1, pp. 125–134, 2001.
[11] V. V. Mazalov and E. E. Ivashko, "Full-information best-choice problem with disorder," *Surveys in Applied and Industrial Mathematics*, vol. 14, no. 2, pp. 215–224, 2007.