Optimal strategies in the best choice problem with disorder

Vladimir V. Mazalov, Evgeny E. Ivashko

September 13, 2010

Best choice problem with disorder and imperfect observation

- A production system generates a sequence of n i.i.d. random variables $X_1, X_2, ..., X_{\theta-1}, X_\theta, X_{\theta+1}, ..., X_n$;

Best choice problem with disorder and imperfect observation

- A production system generates a sequence of n i.i.d. random variables $X_1, X_2, ..., X_{\theta-1}, X_\theta, X_{\theta+1}, ..., X_n$;
- in the random time θ the *disorder* happens and the distribution law changes:
  $X_1, X_2, ..., X_{\theta-1}$ have absolutely continuous CDF $F_1(x)$ (state S_1),
  $X_\theta, X_{\theta+1}, ..., X_n$ have absolutely continuous CDF $F_2(x)$ (state S_2);

Best choice problem with disorder and imperfect observation

- A production system generates a sequence of n i.i.d. random variables $X_1, X_2, ..., X_{\theta-1}, X_\theta, X_{\theta+1}, ..., X_n$;
- in the random time θ the *disorder* happens and the distribution law changes:
  $X_1, X_2, ..., X_{\theta-1}$ have absolutely continuous CDF $F_1(x)$ (state S_1),
  $X_\theta, X_{\theta+1}, ..., X_n$ have absolutely continuous CDF $F_2(x)$ (state S_2);
- the moment of disorder θ has geometric distribution with parameter $1-\alpha$ (i.e. $P(\theta = k) = \alpha^{k-1}(1-\alpha)$);

Best choice problem with disorder and imperfect observation

- A production system generates a sequence of n i.i.d. random variables $X_1, X_2, ..., X_{\theta-1}, X_\theta, X_{\theta+1}, ..., X_n$;
- in the random time θ the *disorder* happens and the distribution law changes:
  $X_1, X_2, ..., X_{\theta-1}$ have absolutely continuous CDF $F_1(x)$ (state S_1),
  $X_\theta, X_{\theta+1}, ..., X_n$ have absolutely continuous CDF $F_2(x)$ (state S_2);
- the moment of disorder θ has geometric distribution with parameter $1-\alpha$ (i.e. $P(\theta = k) = \alpha^{k-1}(1-\alpha)$);
- a decision-maker observes sequentially one-by-one the random variables with the aim of choosing the largest;

Best choice problem with disorder and imperfect observation

- A production system generates a sequence of n i.i.d. random variables $X_1, X_2, \dots, X_{\theta-1}, X_\theta, X_{\theta+1}, \dots, X_n$;
- in the random time θ the *disorder* happens and the distribution law changes:
  $X_1, X_2, \dots, X_{\theta-1}$ have absolutely continuous CDF $F_1(x)$ (state S_1),
  $X_\theta, X_{\theta+1}, \dots, X_n$ have absolutely continuous CDF $F_2(x)$ (state S_2);
- the moment of disorder θ has geometric distribution with parameter $1-\alpha$ (i.e. $P(\theta = k) = \alpha^{k-1}(1-\alpha)$);
- a decision-maker observes sequentially one-by-one the random variables with the aim of choosing the largest;
- after the each sampling the decision-maker have to decide: to STOP or to CONTINUE:

Best choice problem with disorder and imperfect observation

- A production system generates a sequence of n i.i.d. random variables $X_1, X_2, ..., X_{\theta-1}, X_\theta, X_{\theta+1}, ..., X_n$;
- in the random time θ the *disorder* happens and the distribution law changes:
  $X_1, X_2, ..., X_{\theta-1}$ have absolutely continuous CDF $F_1(x)$ (state S_1),
  $X_\theta, X_{\theta+1}, ..., X_n$ have absolutely continuous CDF $F_2(x)$ (state S_2);
- the moment of disorder θ has geometric distribution with parameter $1-\alpha$ (i.e. $P(\theta = k) = \alpha^{k-1}(1-\alpha)$);
- a decision-maker observes sequentially one-by-one the random variables with the aim of choosing the largest;
- after the each sampling the decision-maker have to decide: to STOP or to CONTINUE:
  STOP: *accept* the value and stop the observation process;

Best choice problem with disorder and imperfect observation

- A production system generates a sequence of n i.i.d. random variables $X_1, X_2, ..., X_{\theta-1}, X_\theta, X_{\theta+1}, ..., X_n$;
- in the random time θ the *disorder* happens and the distribution law changes:
  $X_1, X_2, ..., X_{\theta-1}$ have absolutely continuous CDF $F_1(x)$ (state S_1),
  $X_\theta, X_{\theta+1}, ..., X_n$ have absolutely continuous CDF $F_2(x)$ (state S_2);
- the moment of disorder θ has geometric distribution with parameter $1-\alpha$ (i.e. $P(\theta = k) = \alpha^{k-1}(1-\alpha)$);
- a decision-maker observes sequentially one-by-one the random variables with the aim of choosing the largest;
- after the each sampling the decision-maker have to decide: to STOP or to CONTINUE:
  STOP: *accept* the value and stop the observation process;
  CONTINUE: *reject* the observation and observe the next r.v.

Best choice problem with disorder and imperfect observation

- A production system generates a sequence of n i.i.d. random variables $X_1, X_2, ..., X_{\theta-1}, X_\theta, X_{\theta+1}, ..., X_n$;
- in the random time θ the *disorder* happens and the distribution law changes:
  $X_1, X_2, ..., X_{\theta-1}$ have absolutely continuous CDF $F_1(x)$ (state S_1),
  $X_\theta, X_{\theta+1}, ..., X_n$ have absolutely continuous CDF $F_2(x)$ (state S_2);
- the moment of disorder θ has geometric distribution with parameter $1-\alpha$ (i.e. $P(\theta = k) = \alpha^{k-1}(1-\alpha)$);
- a decision-maker observes sequentially one-by-one the random variables with the aim of choosing the largest;
- after the each sampling the decision-maker have to decide: to STOP or to CONTINUE:
  STOP: *accept* the value and stop the observation process;
  CONTINUE: *reject* the observation and observe the next r.v.
- The rejected observation cannot be recalled later;

Best choice problem with disorder and imperfect observation

Best choice problem with disorder and imperfect observation

- λ $(0 < \lambda < 1)$ is the discount factor: if the decision-maker accepts the r.v. X_k she gets the value $\lambda^{k-1}X_k$;

Best choice problem with disorder and imperfect observation

- λ $(0 < \lambda < 1)$ is the discount factor: if the decision-maker accepts the r.v. X_k she gets the value $\lambda^{k-1} X_k$;
- the r.v. cannot be perfectly observed. Each time a r.v. is sampled the sampler is informed only whether it is greater than or less than some level specified by her;

Best choice problem with disorder and imperfect observation

- λ $(0 < \lambda < 1)$ is the discount factor: if the decision-maker accepts the r.v. X_k she gets the value $\lambda^{k-1}X_k$;
- the r.v. cannot be perfectly observed. Each time a r.v. is sampled the sampler is informed only whether it is greater than or less than some level specified by her;
- a decision-maker knows parameters λ,α and n, but the real state of the system is unknown;

Best choice problem with disorder and imperfect observation

- λ $(0 < \lambda < 1)$ is the discount factor: if the decision-maker accepts the r.v. X_k she gets the value $\lambda^{k-1}X_k$;
- the r.v. cannot be perfectly observed. Each time a r.v. is sampled the sampler is informed only whether it is greater than or less than some level specified by her;
- a decision-maker knows parameters λ,α and n, but the real state of the system is unknown;

Goal: maximize the expected value of the accepted observation.

Strategy

We find the solution in the class of the following strategies.
Each moment k $(1 \leq k \leq n)$ the observer estimates the *a posterior* probability of the current state and specifies the threshold $s_k = s_k(x_1, \dots, x_{k-1})$.
The decision-maker accepts the observation x_k if and only if it is greater than the corresponding threshold s_k.

Gain function

The observer estimates the current state of the system using the Byes' formula:

Gain function

The observer estimates the current state of the system using the Byes' formula:

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

Gain function

The observer estimates the current state of the system using the Byes' formula:

$$\pi_s = \pi(s) = P\{S_1 | x \le s\} = \frac{P(S_1)P(x \le s | S_1)}{P(x \le s)} = \frac{\alpha\pi F_1(s)}{F_\pi(s)}.$$

Here

- $s = s_i$ is the threshold specified by the decision-maker within i steps till 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 $x \le s$);
- π_s is the *a posterior* probability of the state S_1 (i.e. *after* getting the information that $x \le s$);
- $F_\pi(s) = \pi F_1(s) + \overline{\pi} F_2(s)$;
- $\overline{\pi} = 1 - \pi$.

Gain function

We use the dynamic programming approach to derive the optimal strategy.

Gain function

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

$$\begin{cases} v_i(\pi) & = & \max\limits_s E\left[\lambda v_{i-1}(\pi_s) I_{x\leq s} + x I_{x>s}\right] \\ & = & \max\limits_s \left[\lambda v_{i-1}(\pi_s) F_\pi(s) + \pi E_1(s) + \overline{\pi} E_2(s)\right], \; i \geq 1, \\ v_0(\pi) & = & 0 \; \forall \pi. \end{cases} \quad (1)$$

Here
$E_k(s) = \int\limits_s^\infty x dF_k(x)$, $k = 1, 2$ and
$I_{a<b} = \begin{cases} 1, & \text{if } a < b \\ 0, & \text{otherwise} \end{cases}$

Gain function – Theorem 1

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

Theorem 1. *For any i the function $v_i(\pi)$ could be written if the form*

$$v_i(\pi) = \pi A_i(s_1, ..., s_i) + B_i(s_1, ..., s_i),$$

where

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

Gain function – Theorem 1

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

Theorem 1. *For any i the function $v_i(\pi)$ could be written if the form*

$$v_i(\pi) = \pi A_i(s_1, ..., s_i) + B_i(s_1, ..., s_i),$$

where
$s_i = s_i(\pi) = \arg\max_s \left[\lambda v_{i-1}(\pi_s) F_\pi(s) + \pi E_1(s) + \overline{\pi} E_2(s)\right],\ i \geq 1, 0 \leq \pi \leq 1.$
The theorem can be proved by induction.

Gain function – Theorem 2

We prove the following lemma.

Lemma. *As* $i \to \infty$ *there is a limit of the expected payoff* $v_i(\pi) \to v(\pi)$.

Gain function – Theorem 2

We prove the following lemma.

Lemma. *As $i \to \infty$ there is a limit of the expected payoff $v_i(\pi) \to v(\pi)$.*

Corollary. *From the theorem 1 and lemma one can show that there are such A and B that*

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

Gain function – Theorem 2

Theorem 2. *For $i \to \infty$ the solution of the full-information best choice problem with disorder is defined as*

$$v(\pi) = \max_s(\pi A + B),$$

where

$$s = s(\pi) = \arg\max_s(\pi A + B)$$

and

$$A = \frac{E_1(s)(1-\lambda F_2(s)) - E_2(s)(1-\lambda F_1(s))}{(1-\lambda F_2(s))(1-\lambda\alpha F_1(s))}$$
$$B = \frac{E_2(s)}{1-\lambda F_2(s)}.$$

Example – Normal distribution

Consider the examples of using the Bayes' strategy B comparing with two strategies with constant thresholds that not depend on π.

Example – Normal distribution

Consider the examples of using the Bayes' strategy B comparing with two strategies with constant thresholds that not depend on π.
Let r.v. $X_1, ..., X_n$ have the normal distribution 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 for the S_1 and S_2 states.

Example – Normal distribution

Consider the examples of using the Bayes' strategy B comparing with two strategies with constant thresholds that not depend on π.
Let r.v. $X_1, ..., X_n$ have the normal distribution 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 for the S_1 and S_2 states.
Strategies A_1 and A_2 with constant thresholds s:

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

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.

Example – Normal distribution

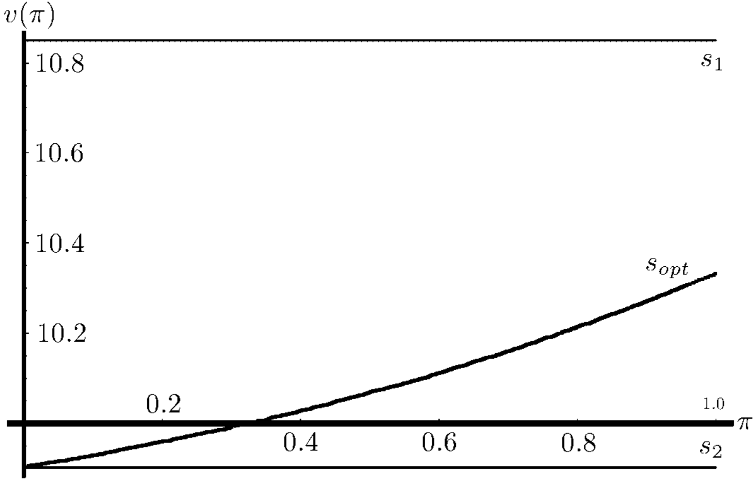

Figure: Graphics of the optimal thresholds for strategies A_1, A_2 and B for $\alpha = 0.9$, $\lambda = 0.99$

Example – Normal distribution

The following figure shows the numerical results of the expected payoffs of the observer who use the strategies A_1, A_2 and B (thresholds s_1, s_2 and s_{opt} respectively).

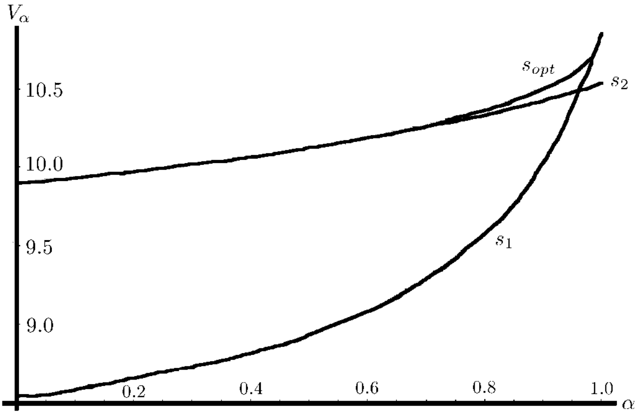

Figure: Expected payoffs of the observer who use the strategies A_1, A_2 and B α for $\lambda = 0.99$

Bibliography

- 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.
- Z. Porosinski and K. Szajowski. Random priority two-person full-information best choice problem with imperfect observation // Applicationes Mathematicae, 27 (3), 2000, pp. 251-263.
- P. Neumann, Z. Porosinski, K. Szajowski. On two-person full-information best-choice problem with imperfect observation // Nova Journal of Mathematics, Game Theory and Algebra, 1996.
- V. Mazalov, E. Ivashko. Best-choice problem with disorder // Surveys in Applied and Industrial Mathematics, V. 14, Iss. 2, 2007, pp. 215-224 (in russian).
- V. Mazalov, E. Ivashko. Bayes' model in the best-choice problem with disorder // Vestnik of SPbGU. 10, Iss. 4, 2009, pp. 142-151 (in russian).