Optimal Search for the Best Alternative

Martin Weitzman

Econometrica, 1979

Outline

Outline

Introduction

- The uncertainty about the reward from a source can be eliminated, at a fee, by searching or sampling.

Introduction

- The uncertainty about the reward from a source can be eliminated, at a fee, by searching or sampling.
- Each source has its own independent probability distribution for the reward, search cost, and search time.

Introduction

- The uncertainty about the reward from a source can be eliminated, at a fee, by searching or sampling.
- Each source has its own independent probability distribution for the reward, search cost, and search time.
- Sources are sampled sequentially, in whatever order is desired.

Introduction

- The uncertainty about the reward from a source can be eliminated, at a fee, by searching or sampling.
- Each source has its own independent probability distribution for the reward, search cost, and search time.
- Sources are sampled sequentially, in whatever order is desired.
- When it has been decided to stop searching, only one opportunity is accepted, the maximum sampled reward.

Introduction

- The uncertainty about the reward from a source can be eliminated, at a fee, by searching or sampling.
- Each source has its own independent probability distribution for the reward, search cost, and search time.
- Sources are sampled sequentially, in whatever order is desired.
- When it has been decided to stop searching, only one opportunity is accepted, the maximum sampled reward.
- What sequential search strategy maximizes expected present discounted value?

Introduction

- Each source is assigned a *reservation price*.

Introduction

- Each source is assigned a *reservation price*.
- The selection rule is to search next that unsampled source with highest reservation price.

Introduction

- Each source is assigned a *reservation price.*
- The selection rule is to search next that unsampled source with highest reservation price.
- The stopping rule is to terminate search whenever the maximum sampled reward is above the reservation price of every unsampled source.

Outline

Example

- Finding a new and cheaper way to produce some commodity.

Example

- Finding a new and cheaper way to produce some commodity.
- Two substitute technologies are being considered.

Example

- Finding a new and cheaper way to produce some commodity.
- Two substitute technologies are being considered.
- No more than one technology would actually be used even if both were developed.

Example

- Finding a new and cheaper way to produce some commodity.
- Two substitute technologies are being considered.
- No more than one technology would actually be used even if both were developed.
- Research and development must be done to remove the uncertainty, and it is costly.

Example

TABLE I

Project	α		ω	
Cost	15		20	
Duration	1		2	
Reward	100	55	240	0
Probability	$1/2$	$1/2$	$1/5$	$4/5$

Example

- The expected value of researching α is:

$$-15+\left(\frac{1}{1.1}\right)\times\left(\frac{1}{2}(100)+\frac{1}{2}(55)\right)=55.5$$

- The expected value of researching ω is:

$$-20+\left(\frac{1}{1.1}\right)^{2}\times\left(\frac{1}{5}(240)+\frac{4}{5}(0)\right)=19.7$$

Example

Suppose α is developed first.

- If the payoff turns out to be 55, the expected value of researching ω is:

$$-20 + \left(\frac{1}{1.1}\right)^2 \times \left(\frac{1}{5}(240) + \frac{4}{5}(55)\right) = 56$$

- If the payoff turns out to be 100, the expected value of researching ω is:

$$-20 + \left(\frac{1}{1.1}\right)^2 \times \left(\frac{1}{5}(240) + \frac{4}{5}(100)\right) = 85.8$$

Example

- The expected value of an optimal policy beginning with developing α is:

$$-15 + \left(\frac{1}{1.1}\right) \times \left[\frac{1}{2}(100) + \frac{1}{2}\left(-20 + \left(\frac{1}{1.1}\right)^2 \times \left(\frac{1}{5}(240) + \frac{4}{5}(55)\right)\right)\right] = 55.9$$

Example

- The expected value of an optimal policy beginning with developing α is:

$$-15+\left(\frac{1}{1.1}\right)\times\left[\frac{1}{2}(100)+\frac{1}{2}\left(-20+\left(\frac{1}{1.1}\right)^{2}\times\left(\frac{1}{5}(240)+\frac{4}{5}(55)\right)\right)\right]=55.9$$

- The expected value of an optimal policy beginning with developing ω is:

$$-20+\left(\frac{1}{1.1}\right)^{2}\times\left[\frac{1}{5}(240)+\frac{4}{5}\left(-15+\left(\frac{1}{1.1}\right)\times\left(\frac{1}{2}(100)+\frac{1}{2}(55)\right)\right)\right]=56.3$$

Example

- The expected value of an optimal policy beginning with developing α is:

$$-15 + \left(\frac{1}{1.1}\right) \times \left[\frac{1}{2}(100) + \frac{1}{2}\left(-20 + \left(\frac{1}{1.1}\right)^2 \times \left(\frac{1}{5}(240) + \frac{4}{5}(55)\right)\right)\right] = 55.9$$

- The expected value of an optimal policy beginning with developing ω is:

$$-20 + \left(\frac{1}{1.1}\right)^2 \times \left[\frac{1}{5}(240) + \frac{4}{5}\left(-15 + \left(\frac{1}{1.1}\right) \times \left(\frac{1}{2}(100) + \frac{1}{2}(55)\right)\right)\right] = 56.3$$

- ω is researched first. (a counter-intuitive property)

Outline

Model

- There are n closed boxes, and $\mathcal{I} = \{1, 2, \ldots, n\}$.
- Box i contains a potential reward x_i with pdf $F_i(x_i)$ independent of the other rewards.
- It cost c_i to open box i.
- Box i become know only after a time lag t_i.
- An initially amount x_0 is available.
- All costs and benefits are converted to present values by the discount rate r.

Model

- Let $\mathcal{I}$ be partitioned into any set S such that:
  - $S \cup \bar{S} = \mathcal{I}$
  - $S \cap \bar{S} = \emptyset$
- The variable y will represent the maximum sampled reward:

$$y = \max_{i \in S \cup \{0\}} \{x_i\}$$

- The state of the system at any time is given by the statistic $(\overline{S}, y)$.
- Define $\Psi(\bar{S}, y)$ as the expected present discounted value of following an optimal policy. Ψ must satisfies the fundamental recursive relation:

Model

$$\Psi(\bar{S}, y) = \max\left\{y, \max_{i \in \bar{S}}\left\{-c_i + \beta_i\left[\int_{-\infty}^{y} \Psi(\bar{S} - \{i\}, y)\mathrm{d}F_i(x_i) + \int_{y}^{\infty} \Psi(\bar{S} - \{i\}, x_i)\mathrm{d}F_i(x_i)\right]\right\}\right\}$$

- $\Psi(\emptyset, x) = x$
- $\beta_i = \mathrm{e}^{-rt_i}$

Outline

The Optimal Strategy

Suppose for the moment there are just two boxes (one is the closed box i).

- If the searcher elects not to open box i, she receives:

The Optimal Strategy

Suppose for the moment there are just two boxes (one is the closed box i).

- If the searcher elects not to open box i, she receives:

$$z_i$$

The Optimal Strategy

Suppose for the moment there are just two boxes (one is the closed box i).

- If the searcher elects not to open box i, she receives:

$$z_i$$

- If she opens box i, the searcher can expect a net benefit:

The Optimal Strategy

Suppose for the moment there are just two boxes (one is the closed box i).

- If the searcher elects not to open box i, she receives:

$$z_i$$

- If she opens box i, the searcher can expect a net benefit:

$$-c_i + \beta_i \left[\int_{-\infty}^{z_i} z_i \mathrm{d}F_i(x_i) + \int_{z_i}^{\infty} x_i F_i(x_i) \right]$$

The Optimal Strategy

- The closed and opened boxes are heuristically "equivalent" if the searcher is just indifferent between opening box i and not opening it:

The Optimal Strategy

- The closed and opened boxes are heuristically "equivalent" if the searcher is just indifferent between opening box i and not opening it:

$$c_i = \beta_i \int_{z_i}^{\infty} (x_i - z_i) \mathrm{d}F_i(x_i) - (1 - \beta_i) z_i$$

The Optimal Strategy

- The closed and opened boxes are heuristically "equivalent" if the searcher is just indifferent between opening box i and not opening it:

$$c_i = \beta_i \int_{z_i}^{\infty} (x_i - z_i) \mathrm{d}F_i(x_i) - (1 - \beta_i) z_i$$

- The critical number z_i is called the *reservation price* of box i.

The Optimal Strategy

Proposición 1: Pandora's Rule

SELECTION RULE: *If a box is to be opened, it should be that closed box with highest reservation price.*

STOPPING RULE: *Terminate search whenever the maximum sampled reward exceeds the reservation price of every closed box.*

Outline

First Application

Gambling

- Suppose that box i contains one of two outcomes: either zero reward ("failure") with probability $1 - p_i$, or positive reward R_i ("success") with probability p_i.
- There's no discounting factor ($\beta_i = 1$), and $p_i R_i - c_i > 0$. Then:

$$z_i = R_i - \frac{c_i}{p_i}$$

First Application

Gambling

- Suppose that box i contains one of two outcomes: either zero reward ("failure") with probability $1 - p_i$, or positive reward R_i ("success") with probability p_i.
- There's no discounting factor ($\beta_i = 1$), and $p_i R_i - c_i > 0$. Then:

$$z_i = R_i - \frac{c_i}{p_i}$$

- If search is for the same object:

$$z_i = R - \frac{c_i}{p_i}$$

First Application

Gambling

- Suppose that box i contains one of two outcomes: either zero reward ("failure") with probability $1 - p_i$, or positive reward R_i ("success") with probability p_i.
- There's no discounting factor ($\beta_i = 1$), and $p_i R_i - c_i > 0$. Then:

$$z_i = R_i - \frac{c_i}{p_i}$$

- If search is for the same object:

$$z_i = R - \frac{c_i}{p_i}$$

- The searcher moves down the list until a success is encountered ;)

Second Application

Gold Mining Problem

- Mine i contains amount G_i of gold, but after digging it out the machine is liable to break down with probability q_i, preventing all further mining.

Second Application

Gold Mining Problem

- Mine i contains amount G_i of gold, but after digging it out the machine is liable to break down with probability q_i, preventing all further mining.
- In what order should the mines be exploited to maximize total expected gold?

Second Application

Gold Mining Problem

- Mine i contains amount G_i of gold, but after digging it out the machine is liable to break down with probability q_i, preventing all further mining.
- In what order should the mines be exploited to maximize total expected gold?
- $c_i = -G_i$, $x_0, x_i = 0$, $\beta_i = 1 - q_i$.

Second Application

Gold Mining Problem

- Mine i contains amount G_i of gold, but after digging it out the machine is liable to break down with probability q_i, preventing all further mining.
- In what order should the mines be exploited to maximize total expected gold?
- $c_i = -G_i$, $x_0, x_i = 0$, $\beta_i = 1 - q_i$.

$$z_i = \frac{G_i}{q_i}$$

Second Application

Gold Mining Problem

- Mine i contains amount G_i of gold, but after digging it out the machine is liable to break down with probability q_i, preventing all further mining.
- In what order should the mines be exploited to maximize total expected gold?
- $c_i = -G_i$, $x_0, x_i = 0$, $\beta_i = 1 - q_i$.

$$z_i = \frac{G_i}{q_i}$$

- The optimal policy is to exploit next that unopened mine with maximum gold per probability of machine breakdown, a classic result.