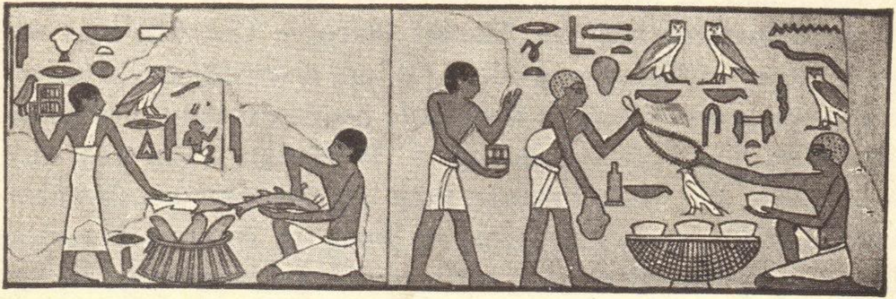
Dynamic pricing and learning

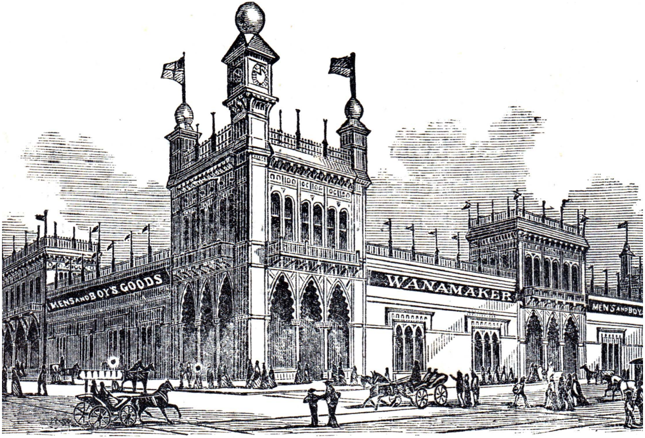
Arnoud V. den Boer
University of Amsterdam

Winter school on Mathematical Finance
Lunteren, 2020



EGYPTIAN MARKET SCENES

From an Egyptian tomb. Selling and cleaning fish; bartering a necklace for pots of perfume. In the upper and central parts of each scene are examples of Egyptian writing in hieroglyphics.



'One price and goods returnable'

Engraver unknown. From J.D. McCabe, *The Illustrated History of the Centennial Exposition*, 1876

Dynamic pricing:

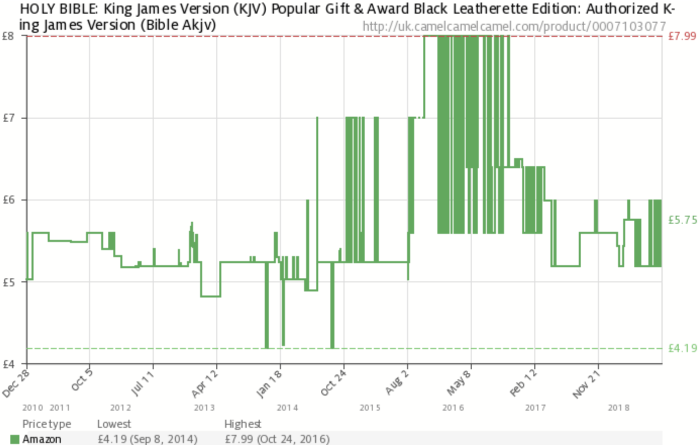
Adapting selling prices to changing circumstances

- Remaining inventory
- Competitor's prices
- Time and date
- Expiry date
- Customer profile
- Temperature
- ...

Zelfs de kapper varieert nu met zijn stoelprijzen

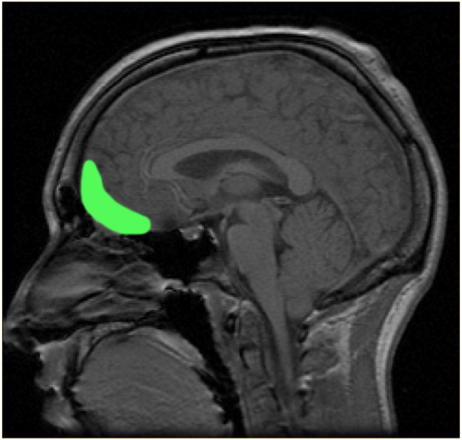
Digitalization makes price changes (practically) costless

Even ‘stable’ products are dynamically priced:



Price experiments to learn customer’s willingness-to-pay?

Learn the willingness-to-pay distribution / price-demand relation

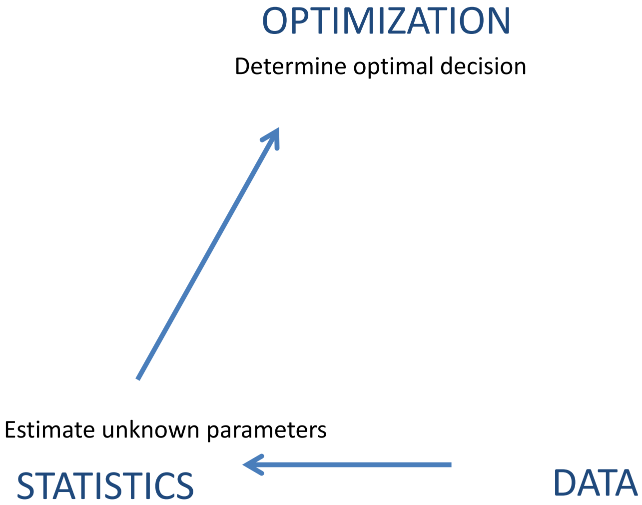


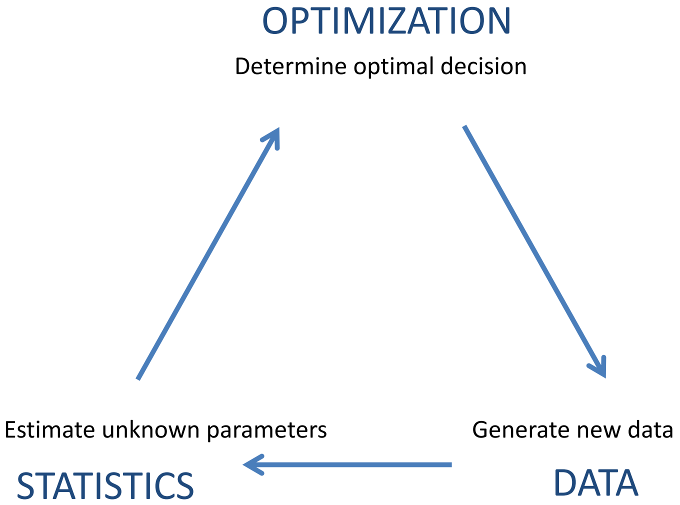
Plassmann H. et. al, Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J Neurosci.* 2007

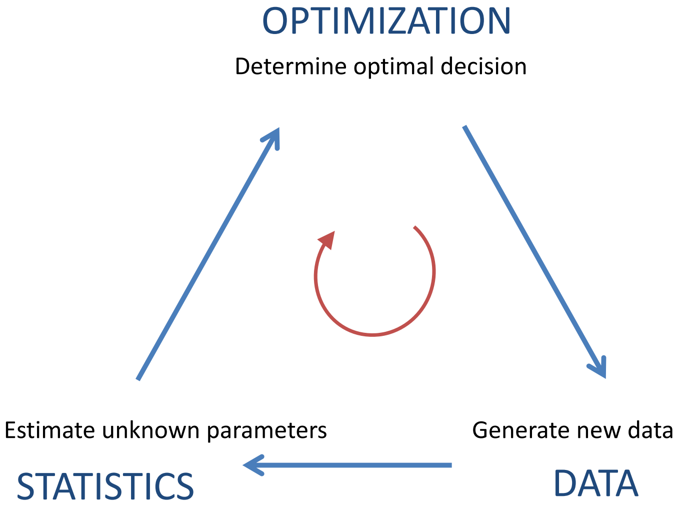
DATA

Estimate unknown parameters

STATISTICS ← DATA


OPTIMIZATION
Determine optimal decision
Estimate unknown parameters
STATISTICS
DATA


OPTIMIZATION
Determine optimal decision
Generate new data
DATA
Estimate unknown parameters
STATISTICS


OPTIMIZATION
Determine optimal decision
Estimate unknown parameters
STATISTICS
Generate new data
DATA

How to learn the optimal selling price?

den Boer and Zwart, *Management Science* 60(3), 2014

- A firm sells a single product in T discrete time periods $t = 1, \ldots, T$.

- A firm sells a single product in T discrete time periods $t = 1, \ldots, T$.
- No competition, infinite supply, marginal costs zero.

- A firm sells a single product in T discrete time periods $t = 1, \ldots, T$.
- No competition, infinite supply, marginal costs zero.
- Each period t:

  (i) choose selling price p_t;

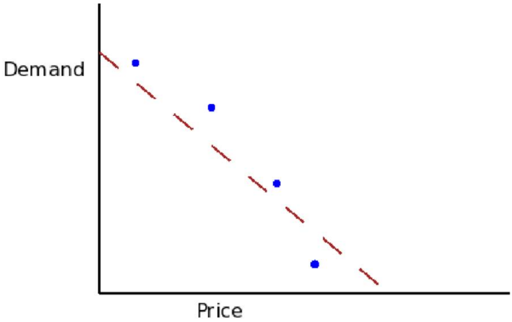
  (ii) observe demand

$$d_t = \theta_1 + \theta_2 p_t + \epsilon_t,$$

  where $\theta = (\theta_1, \theta_2)$ are unknown parameters in some known set Θ, ϵ_t unobservable random disturbance term with zero mean;

  (iii) collect revenue $p_t d_t$.


Demand
Price



Which non-anticipating prices $p_1, \ldots, p_T$ maximize cumulative expected revenue $\inf_{\theta\in\Theta} \mathbb{E}_\theta\left[\sum_{t=1}^{T} p_t d_t\right]$?

Which non-anticipating prices $p_1, \ldots, p_T$ maximize cumulative expected revenue $\inf_{\theta \in \Theta} \mathbb{E}_\theta \left[\sum_{t=1}^{T} p_t d_t \right]$?

Intractable problem

Myopic pricing

An intuitive solution

- Choose arbitrary initial prices $p_1 \neq p_2$.
- For each $t \geq 2$:

  (i) determine least-square estimate $\hat{\theta}_t$ of θ, based on available sales data;

  (ii) set

$$p_{t+1} = \arg\max_p p \cdot (\hat{\theta}_{t1} + \hat{\theta}_{t2} p)$$

Myopic pricing

An intuitive solution

- Choose arbitrary initial prices $p_1 \neq p_2$.
- For each $t \geq 2$:

  (i) determine least-square estimate $\hat{\theta}_t$ of θ, based on available sales data;

  (ii) set

$$p_{t+1} = \arg\max_p p \cdot (\hat{\theta}_{t1} + \hat{\theta}_{t2} p) \quad \text{perceived optimal decision}$$

Myopic pricing

An intuitive solution

- Choose arbitrary initial prices $p_1 \neq p_2$.
- For each $t \geq 2$:

  (i) determine least-square estimate $\hat{\theta}_t$ of θ, based on available sales data;

  (ii) set

  $$p_{t+1} = \arg\max_p p \cdot (\hat{\theta}_{t1} + \hat{\theta}_{t2} p) \quad \text{perceived optimal decision}$$

- 'Always choose the perceived optimal action'.

Convergence

Does $\hat{\theta}_t$ converge to θ as $t \to \infty$?

Convergence

Does $\hat{\theta}_t$ converge to θ as $t \to \infty$?

No

$\hat{\theta}_t$ always converges, but w.p. zero to the true θ.

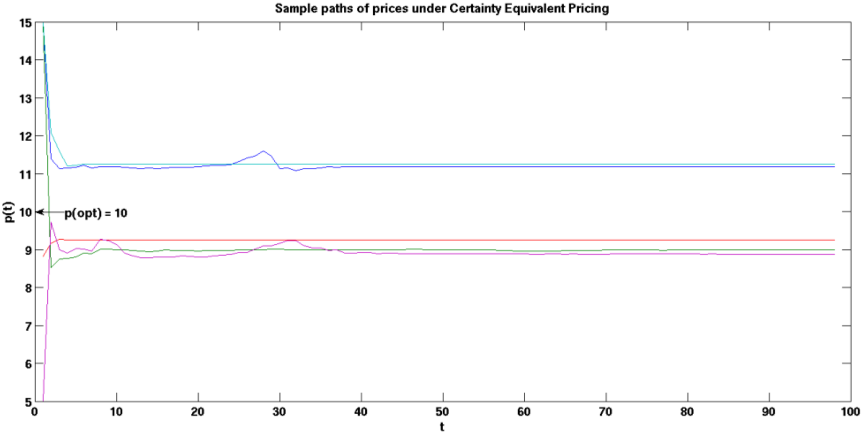


Figure: $D = 10 - 0.5p + N(0,1)$

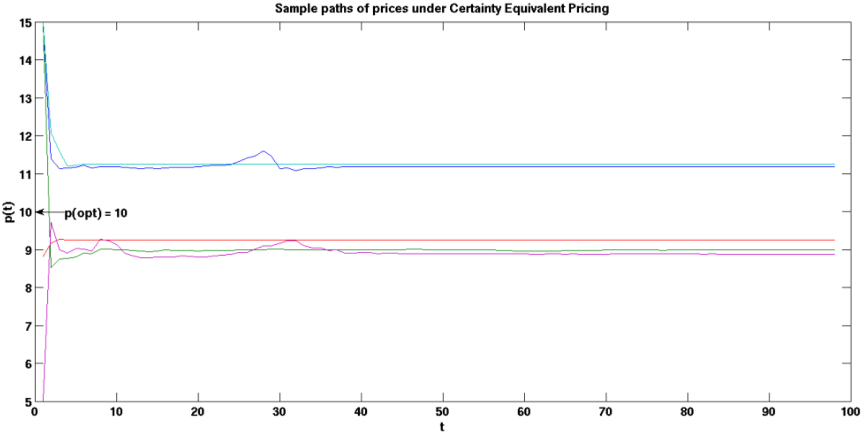


Figure: $D = 10 - 0.5p + N(0,1)$

Caused by indeterminate equilibria: what-you-see is what-you-predict

Indeterminate equilibria

If $\hat{\theta}$ suff. close to θ, then $\arg\max_p p \cdot (\hat{\theta}_1 + \hat{\theta}_2 p) = -\hat{\theta}_1/(2\hat{\theta}_2)$.

Then:

$$\text{'True' expected demand: } \theta_1 + \theta_2 \frac{-\hat{\theta}_1}{2\hat{\theta}_2}. \tag{1}$$

$$\text{'Predicted' expected demand: } \hat{\theta}_1 + \hat{\theta}_2 \frac{-\hat{\theta}_1}{2\hat{\theta}_2}. \tag{2}$$

Indeterminate equilibria

If $\hat{\theta}$ suff. close to θ, then $\arg\max_p p \cdot (\hat{\theta}_1 + \hat{\theta}_2 p) = -\hat{\theta}_1/(2\hat{\theta}_2)$.

Then:

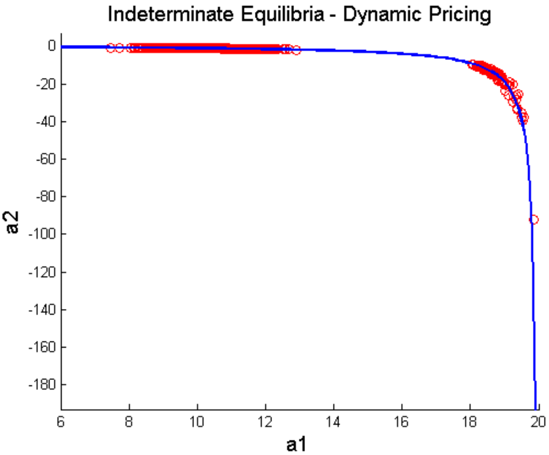
$$\text{`True' expected demand: } \theta_1 + \theta_2 \frac{-\hat{\theta}_1}{2\hat{\theta}_2}. \tag{1}$$

$$\text{`Predicted' expected demand: } \hat{\theta}_1 + \hat{\theta}_2 \frac{-\hat{\theta}_1}{2\hat{\theta}_2}. \tag{2}$$

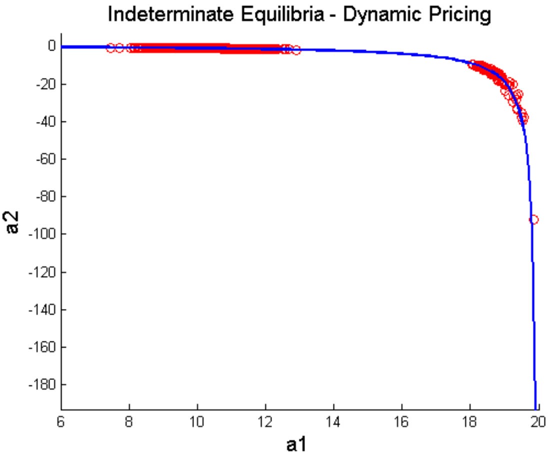
If (1) equals (2), then $\hat{\theta}$ is an IE.

Model output 'confirms' correctness of the (incorrect) estimates.

Indeterminate equilibria: example



Indeterminate equilibria: example



Moral: do not always do what seems best, but deviate in order to learn more about the dynamics of your system

Back to original problem

Which non-anticipating prices $p_1, \ldots, p_T$ maximize

$$\inf_{\theta \in \Theta} \mathbb{E}\Big[\sum_{t=1}^{T} p_t d_t\Big],$$

or, equivalently, minimize the Regret(T)

$$\sup_{\theta \in \Theta} \mathbb{E}\Big[T \cdot \max_{p} p \cdot (\theta_1 + \theta_2 p) - \sum_{t=1}^{T} p_t d_t\Big]$$

Back to original problem

Which non-anticipating prices $p_1, \ldots, p_T$ maximize

$$\inf_{\theta \in \Theta} \mathbb{E}\Big[\sum_{t=1}^{T} p_t d_t\Big],$$

or, equivalently, minimize the Regret(T)

$$\sup_{\theta \in \Theta} \mathbb{E}\Big[T \cdot \max_p p \cdot (\theta_1 + \theta_2 p) - \sum_{t=1}^{T} p_t d_t\Big]$$

- Exact solution intractable

Back to original problem

Which non-anticipating prices $p_1, \ldots, p_T$ maximize

$$\inf_{\theta \in \Theta} \mathbb{E}\Big[\sum_{t=1}^{T} p_t d_t\Big],$$

or, equivalently, minimize the Regret(T)

$$\sup_{\theta \in \Theta} \mathbb{E}\Big[T \cdot \max_p p \cdot (\theta_1 + \theta_2 p) - \sum_{t=1}^{T} p_t d_t\Big]$$

- Exact solution intractable
- Myopic pricing not optimal

Back to original problem

Which non-anticipating prices $p_1, \ldots, p_T$ maximize

$$\inf_{\theta \in \Theta} \mathbb{E}\Big[\sum_{t=1}^{T} p_t d_t\Big],$$

or, equivalently, minimize the Regret(T)

$$\sup_{\theta \in \Theta} \mathbb{E}\Big[T \cdot \max_{p} p \cdot (\theta_1 + \theta_2 p) - \sum_{t=1}^{T} p_t d_t\Big]$$

- Exact solution intractable
- Myopic pricing not optimal
- Let's find asymptotically optimal policies: smallest growth rate of Regret(T) in T.

Asymptotically optimal policy

Important observation: Variation in controls $\Rightarrow$ better estimates.

Asymptotically optimal policy

Important observation: Variation in controls $\Rightarrow$ better estimates.

$$||\hat{\theta}_t - \theta||^2 \sim \frac{\log t}{t\mathrm{Var}(p_1, \ldots, p_t)}$$

To ensure convergence of $\hat{\theta}_t$, some amount of experimentation is necessary.

Asymptotically optimal policy

Important observation: Variation in controls ⇒ better estimates.

$$||\hat{\theta}_t - \theta||^2 \sim \frac{\log t}{t\mathrm{Var}(p_1, \ldots, p_t)}$$

To ensure convergence of $\hat{\theta}_t$, some amount of experimentation is necessary.

But, not *too* much.

‘Controlled Variance pricing’

- Choose arbitrary initial prices $p_1 \neq p_2$.
- For each $t \geq 2$:

  (i) determine least-squares estimate $\hat{\theta}_t$ of θ, based on available sales data;

  (ii) set

$$p_{t+1} = \arg\max_p p \cdot (\hat{\theta}_{t1} + \hat{\theta}_{t2} p)$$

‘Controlled Variance pricing’

- Choose arbitrary initial prices $p_1 \neq p_2$.
- For each $t \geq 2$:

  (i) determine least-squares estimate $\hat{\theta}_t$ of θ, based on available sales data;

  (ii) set

$$p_{t+1} = \arg\max_p p \cdot (\hat{\theta}_{t1} + \hat{\theta}_{t2} p) \quad \text{perceived optimal decision}$$

‘Controlled Variance pricing’

- Choose arbitrary initial prices $p_1 \neq p_2$.
- For each $t \geq 2$:

  (i) determine least-squares estimate $\hat{\theta}_t$ of θ, based on available sales data;

  (ii) set

$$p_{t+1} = \arg\max_p p \cdot (\hat{\theta}_{t1} + \hat{\theta}_{t2} p) \quad \text{perceived optimal decision}$$

$$\text{s.t. } t \cdot \mathrm{Var}(p_1, \ldots, p_{t+1}) \geq f(t), \quad \text{‘information constraint’}$$

‘Controlled Variance pricing’

- Choose arbitrary initial prices $p_1 \neq p_2$.
- For each $t \geq 2$:

  (i) determine least-squares estimate $\hat{\theta}_t$ of θ, based on available sales data;

  (ii) set

$$p_{t+1} = \arg\max_p p \cdot (\hat{\theta}_{t1} + \hat{\theta}_{t2} p) \quad \text{perceived optimal decision}$$

$$\text{s.t. } t \cdot \text{Var}(p_1, \ldots, p_{t+1}) \geq f(t), \quad \text{‘information constraint’}$$

  for some increasing $f : \mathbb{N} \to (0, \infty)$.

‘Controlled Variance pricing’

- Choose arbitrary initial prices $p_1 \neq p_2$.
- For each $t \geq 2$:

  (i) determine least-squares estimate $\hat{\theta}_t$ of θ, based on available sales data;

  (ii) set

$$p_{t+1} = \arg\max_p p \cdot (\hat{\theta}_{t1} + \hat{\theta}_{t2} p) \quad \text{perceived optimal decision}$$

$$\text{s.t. } t \cdot \text{Var}(p_1, \dots, p_{t+1}) \geq f(t), \quad \text{‘information constraint’}$$

  for some increasing $f : \mathbb{N} \to (0, \infty)$.
- ‘Always choose the perceived optimal action that induces sufficient experimentation’.

'Controlled Variance pricing' - performance

Information constraint:

$$(p_{t+1} - \bar{p}_t)^2 \geq f(t+1) - f(t)$$

where $\bar{p}_t = (p_1 + \ldots + p_t)/t$.

‘Controlled Variance pricing’ - performance

Information constraint:

$$(p_{t+1} - \bar{p}_t)^2 \geq f(t+1) - f(t)$$

where $\bar{p}_t = (p_1 + \ldots + p_t)/t$.

‘Taboo interval’ induces sufficient experimentation

‘Controlled Variance pricing’ - performance

Information constraint:

$$(p_{t+1} - \bar{p}_t)^2 \geq f(t+1) - f(t)$$

where $\bar{p}_t = (p_1 + \ldots + p_t)/t$.

‘Taboo interval’ induces sufficient experimentation

- Regret$(T) \sim f(T) + \sum_{t=1}^{T} \frac{\log t}{f(t)}$.

‘Controlled Variance pricing’ - performance

Information constraint:

$$(p_{t+1} - \bar{p}_t)^2 \geq f(t+1) - f(t)$$

where $\bar{p}_t = (p_1 + \ldots + p_t)/t$.

‘Taboo interval’ induces sufficient experimentation

- $\text{Regret}(T) \sim f(T) + \sum_{t=1}^{T} \frac{\log t}{f(t)}$.
- f balances between exploration and exploitation.

‘Controlled Variance pricing’ - performance

Information constraint:

$$(p_{t+1} - \bar{p}_t)^2 \geq f(t+1) - f(t)$$

where $\bar{p}_t = (p_1 + \ldots + p_t)/t$.

‘Taboo interval’ induces sufficient experimentation

- $\text{Regret}(T) \sim f(T) + \sum_{t=1}^{T} \frac{\log t}{f(t)}$.
- f balances between exploration and exploitation.
- Optimal f gives $\text{Regret}(T) \sim \sqrt{T \log T}$, and no policy beats $\sqrt{T}$.

‘Controlled Variance pricing’ - performance

Information constraint:

$$(p_{t+1} - \bar{p}_t)^2 \geq f(t+1) - f(t)$$

where $\bar{p}_t = (p_1 + \ldots + p_t)/t$.

‘Taboo interval’ induces sufficient experimentation

- $\text{Regret}(T) \sim f(T) + \sum_{t=1}^{T} \frac{\log t}{f(t)}$.
- f balances between exploration and exploitation.
- Optimal f gives $\text{Regret}(T) \sim \sqrt{T \log T}$, and no policy beats $\sqrt{T}$.

First-order optimal policy.

‘Controlled Variance pricing’ - performance

Information constraint:

$$(p_{t+1} - \bar{p}_t)^2 \geq f(t+1) - f(t)$$

where $\bar{p}_t = (p_1 + \ldots + p_t)/t$.

‘Taboo interval’ induces sufficient experimentation

- $\text{Regret}(T) \sim f(T) + \sum_{t=1}^{T} \frac{\log t}{f(t)}$.
- f balances between exploration and exploitation.
- Optimal f gives $\text{Regret}(T) \sim \sqrt{T \log T}$, and no policy beats $\sqrt{T}$.

First-order optimal policy.

In implementation, ‘second-order’ refinements possible.

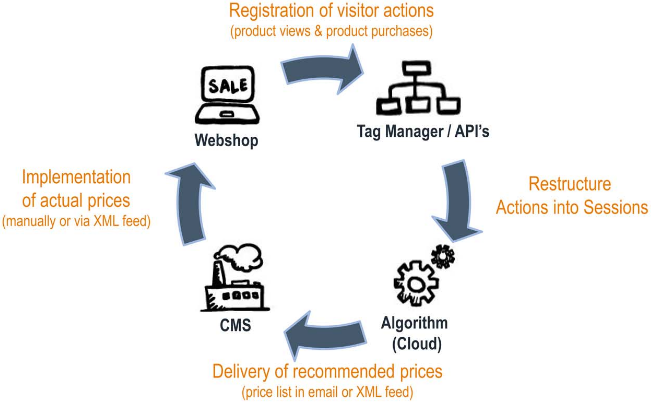
DynaPrice

Dynamic pricing at Wijnvoordeel.nl

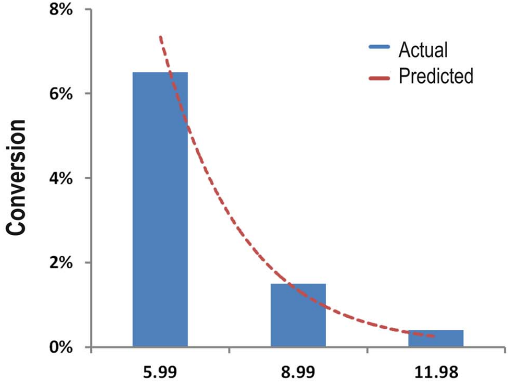
The pilot (2015-2016)

Arenoe Marketing Intelligence

Implementation of a highly structured data extraction & analysis process



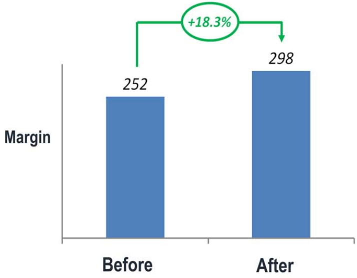
Estimating the demand models



Arenoe Marketing Intelligence

Final outtake

Margin improvement



Insight into pricing dynamics

Arenoe
Marketing Intelligence

Some extensions

This extends to multiple products; instead of $t \cdot \mathrm{Var}(p_1, \ldots, p_t)$, control

$$\lambda_{\min}\left(\sum_{i=1}^{t}\begin{pmatrix}1\\ \mathbf{p}_1(i)\\ \vdots\\ \mathbf{p}_n(i)\end{pmatrix}\begin{pmatrix}1 & \mathbf{p}_1(i) & \ldots & \mathbf{p}_n(i)\end{pmatrix}\right).$$

(den Boer, *Mathematics of Operations Research* 39(3), 2014).

Some extensions

This extends to multiple products; instead of $t \cdot \mathrm{Var}(p_1, \ldots, p_t)$, control

$$\lambda_{\min}\left(\sum_{i=1}^{t}\begin{pmatrix}1\\ \mathbf{p}_1(i)\\ \vdots\\ \mathbf{p}_n(i)\end{pmatrix}\begin{pmatrix}1 & \mathbf{p}_1(i) & \ldots & \mathbf{p}_n(i)\end{pmatrix}\right).$$

(den Boer, *Mathematics of Operations Research* 39(3), 2014).

If inventory is finite and selling seasons are repeated,
no experimentation is needed, and $O(\log^2(T))$ *regret possible.*

(den Boer and Zwart, *Operations Research* 63(4), 2015).

Some extensions

Linear demand function is robust for misspecification (Besbes and Zeevi, *Management Science* 61(4), 2015).

Some extensions

Linear demand function is robust for misspecification (Besbes and Zeevi, *Management Science* 61(4), 2015).

Behavioral aspects such as *reference effects* can be dealt with

(den Boer and Keskin, Dynamic Pricing with Demand Learning and Reference Effects, SSRN.

Thanks for your attention!