UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

OPTIMAL STOPPING IN MECHANISM DESIGN

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN GESTIÓN DE OPERACIONES
MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL MATEMÁTICO

PATRICIO TOMÁS FONCEA ARANEDA

PROFESOR GUÍA:
JOSÉ CORREA HAEUSSLER

MIEMBROS DE LA COMISIÓN:
JUAN ESCOBAR CASTRO
JAIME SAN MARTÍN ARISTEGUI
TJARK VREDEVELD

Este trabajo ha sido parcialmente financiado por CONICYT y el Núcleo Milenio Información y Coordinación en Redes.

SANTIAGO DE CHILE
2017

RESUMEN DE LA MEMORIA PARA OPTAR
AL GRADO DE MAGÍSTER EN GESTIÓN DE OPERACIONES
AL TÍTULO DE INGENIERO CIVIL MATEMÁTICO
POR: PATRICIO TOMÁS FONCEA ARANEDA
FECHA: JULIO 2017
PROF. GUÍA: JOSÉ CORREA HAEUSSLER

OPTIMAL STOPPING IN MECHANISM DESIGN

En este trabajo estudiamos un par de problemas de la teoría de paradas óptimas, y mostramos cómo aplicar estos resultados en el diseño de mecanismos. Consideramos dos versiones modificadas de la famosa desigualdad del profeta [10, 16, 17]: una *no-adaptativa* donde la regla de parada debe ser decidida de antemano, y una *adaptativa* — que corresponde a la configuración clásica de la desigualdad del profeta —, pero en el caso restringido cuando las distribuciones de las variables aleatorias están idénticamente distribuidas [13]. Para la primera situación, encontramos un factor de garantía para la regla de parada con respecto al máximo esperado de la secuencia de variables aleatorias y demostramos que es la mejor posible; para el segundo, probamos que una conjetura sobre cuál es el mejor factor posible es verdadera [14]. Cerramos esta tesis extendiendo estos resultados para resolver el problema de un vendedor que enfrenta a muchos compradores potenciales y debe diseñar una subasta secuencial para maximizar sus ingresos. El tipo de mecanismos que consideramos para estudiar este problema de pricing son los *mecanismos posted price*, y los resultados que obtenemos toman la forma de factores de aproximación con respecto al valor de la subasta óptima [19].

RESUMEN DE LA MEMORIA PARA OPTAR
AL GRADO DE MAGÍSTER EN GESTIÓN DE OPERACIONES
AL TÍTULO DE INGENIERO CIVIL MATEMÁTICO
POR: PATRICIO TOMÁS FONCEA ARANEDA
FECHA: JULIO 2017
PROF. GUÍA: JOSÉ CORREA HAEUSSLER

OPTIMAL STOPPING IN MECHANISM DESIGN

In this work we study a pair of problems in optimal stopping theory, and show how to apply these results in mechanism design. We consider two modified versions of the famous prophet inequality [10, 16, 17]: a *non-adaptive* where the stop rule must be decided beforehand, and an *adaptive* one — which corresponds to the classical prophet inequality setting —, but when the distributions of the random variables are identical [13]. For the first set-up, we find a new factor guarantee with respect to the expected maximum of the random variables sequence and prove it is the best possible; for the second, we prove that a conjecture about the best possible factor achievable is true [14]. We close this dissertation by extending these results to solve the problem of a seller that faces many potential buyers and must design a sequential auction in order to maximize its revenue. The type of mechanisms we consider to study this pricing problem are the *posted price mechanisms*, and the results we get are in the form of approximation factors guarantees with respect to the optimal auction [19].

Acknowledgements

This thesis would not have been possible without the support, guidance, and fellowship of José Correa, Ruben Hoeksma, Tim Oosterwijk and Tjark Vredeveld. This work is the result of our shared contributions, and I am therefore profoundly grateful to them.

Contents

Chapter 1

Introduction

Optimal stopping theory is concerned with the problem of choosing a time to take a particular action, in order to maximize an expected reward or minimize an expected cost. A key example of an optimal stopping problem are the *prophet inequalities*, whose study was initiated by Gilbert and Mosteller in the sixties [10]. The classic prophet inequality states that, when faced to a finite sequence of non-negative independent random variables, a gambler who knows their distribution and is allowed to stop the sequence at any time, can obtain, in expectation, at least half as much reward as a *prophet* who knows the values of each random variable and can choose the largest one. The fraction $1/2$ is also best possible [16, 17].

Independently, in 1982 Myerson [19] gave a characterization of the optimal auction, in the context of what today is known as mechanism design. The auction he studied modeled the problem faced by a monopolist who sells a single item to a set of known potential buyers, whose value for the item are random variables. The objective was to design a mechanism maximizing the revenue of the seller. Although Myerson's solution is, in some situations, a remarkably simple mechanism, for reasons that will be clarified later in this work, in many others the mechanism is hard to implement. An alternative are the simple posted price mechanisms, where prices are offered to each customer in a sequential order. As it will be discussed in the last chapter, this mechanism shares a lot of features with some optimal stopping problems.

The contributions of this work are twofold. In one hand, we find new results in the area of optimal stopping for two different settings that resemble the prophet inequality problem. In the other hand, we provide new theoretical guarantees for posted price mechanisms using the previous results, trough a mathematical tool that allows us to reduce any problem in optimal stopping based on thresholds to a posted price strategy in the pricing problem. We remark that most of the contents of this work can be found in [8].

Before starting, we need to introduce some notation. For a random variable X we will usually denote by $F(x) = \mathbb{P}(X \leq x)$ its cumulative distribution function. If S_X represents the support of X, then $F^{-1}(y) = \inf\{x \in S_X \mid F(x) \geq y\}$ is the generalized inverse of F. All functions considered will be Lebesgue-measurable functions. Finally, we denote by $[n] := \{1, \ldots, n\}$ the set of the first n natural numbers.

Chapter 2

Non-adaptive Stopping Rule

In this chapter we study the existence of non-adaptive threshold to solve the problem of selecting a value given by a uniformly random sequence of independent random variables. We provide a constructive proof of a rule that achieves a factor of $\frac{e-1}{e}$ over the expected maximum, an algorithm to implement this rule, and an example that shows that these results are tight. Along the way, we prove a basic result about Bernoulli random variables that we believe can be of independent interest.

2.1 Preliminaries

In its simpler form, the problem we consider can be described as follows. Suppose you can choose one of n prizes whose values are random variables $X_1, \ldots, X_n$ distributed according to n possibly different independent distribution $F_1, \ldots, F_n$. The prizes arrive in random order, and, upon arrival, we must decide whether we keep that prize, or simply discard it and wait for the next. The goal is to maximize the expected value of the selected prize. We are interested in a *non-adaptive* setting, i.e., we look for an acceptance criterion that is set beforehand, and only based on the distributions. In particular, we will investigate the performance of a threshold rule were we accept the first random variable whose value is above a certain threshold.

This problem resembles the already well known problems as the famous prophet inequality, the secretary problem, or the more recently prophet secretary problem [9], although the non-adaptive nature of our setting changes the analysis and the results, and can use to interpret a different kind of scenarios, as we will see in the last chapter.

2.2 Bernoulli Selection Lemma

We begin by proving an intermediate result that will be the key to our analysis.

Lemma 2.1 (Bernoulli Selection Lemma) *Given n independent Bernoulli random variables $X_1, \ldots, X_n$, where $X_i = 1$ with probability q_i and 0 otherwise, and associated prizes $b_1, \ldots, b_n$. The following inequalities hold:*

$$\frac{e}{e-1} \max_{S \subseteq [n]} \mathbb{E}\left[\frac{\sum_{i \in S} b_i X_i}{\sum_{i \in S} X_i}\right] \geq \max_{z_i \leq q_i} \left\{ \sum_{i \in N} b_i z_i \,\Big|\, \sum z_i \leq 1 \right\} \geq \mathbb{E}[\max\{b_1 X_1, \ldots, b_n X_n\}].$$

Here, when evaluating the leftmost term, we define $0/0 = 0$.

The result states that if we are given a set of nonhomogeneous independent Bernoulli random variables with associated prizes, then there is a subset of variables so that the expected average prize of the successes is at least a factor $1 - 1/e$ of the expectation of the maximum prize over all random variables.

The second inequality of Lemma 2.1 is trivial, as the expectation of the maximum is a sum over all values b_i weighed by the probability with which that value is the maximum. Since these probabilities sum to at most one, the inequality follows.

The proof for the first inequality has two main ingredients. First, we reformulate the left hand side in an appropriate way, and lower bound it by another function using KKT-conditions. Then, we show that this function is bounded from below by $1 - 1/e$.

Proof. We start the proof by rewriting the optimization problem:

$$\max_{S \subseteq [n]} \left\{ \mathbb{E}\left[\frac{\sum_{i \in S} b_i X_i}{\sum_{i \in S} X_i}\right] \right\}. \tag{P}$$

Instead of choosing a subset of $[n]$, we set for each $i \in [n]$ a value $\chi_i \in [0,1]$, which represents the probability with which we actually choose i. Now, let $\pi_i = \chi_i q_i$ denote the probability of i being picked and having $X_i = 1$. So we can consider the following maximization problem, with decision variables π, as a relaxation of (P):

$$\max_{0 \leq \pi_i \leq q_i} \sum_{S \subseteq [n]} \left(\frac{b(S)}{|S|} \left(\prod_{i \in S} \pi_i \right) \left(\prod_{i \notin S} (1 - \pi_i) \right) \right),$$

where $b(S) = \sum_{i \in S} b_i$. Note that the previous objective is linear in each variable so that there is an extreme optimal solution [7]. Thus, the previous problem is in fact equivalent to (P). Now, by changing the order of the summations, we obtain

$$\max_{0 \leq \pi_i \leq q_i} \sum_{i \in [n]} b_i \pi_i \sum_{S \subseteq [n] \setminus \{i\}} \frac{1}{1 + |S|} \prod_{j \in S} \pi_j \prod_{j \in [n] \setminus (S \cup \{i\})} (1 - \pi_j). \tag{2.1}$$

With (P) in this equivalent form, we now proceed to guess a feasible solution. To this end, consider an optimal solution z^* to

$$\max \left\{ \sum_{i \in [n]} b_i z_i \,\Big|\, \sum_{i \in [n]} z_i \leq 1, z_i \leq q_i \text{ for all } i \in [n] \right\},$$

and set $\pi_i = \frac{2z_i^*}{2+(e-2)z_i^*}$,[1][2] such that $1-\pi_i = \frac{2-(4-e)z_i^*}{2+(e-2)z_i^*}$. Note that this is a feasible choice of π_i for all $i \in [n]$, since for this choice $\pi_i \leq z_i^* \leq q_i$. We plug this back into (2.1), and obtain that (P) is lower bounded by

$$\sum_{i\in[n]} 2b_i z_i^* \left(\prod_{j\in[n]} \frac{1}{2+(e-2)z_j^*}\right) \sum_{S\subseteq[n]\setminus\{i\}} \frac{2^{|S|}}{1+|S|} \prod_{j\in S} z_j^* \prod_{j\in[n]\setminus(S\cup\{i\})} (2-(4-e)z_j^*)\,. \tag{2.2}$$

We proceed to lower bound this quantity, where we use the following technical result.

Proposition 2.2 *Consider the problem* $\min_{x\in\mathbb{R}_+^M}\{f_M(x) : \sum_{i\in M} x_i \leq a\}$*, where* $a \leq 1$ *and*

$$f_M(x) = \left(\prod_{j\in M} \frac{1}{2+(e-2)x_j}\right) \sum_{S\subseteq M} \frac{2^{|S|}}{1+|S|} \prod_{j\in S} x_j \prod_{j\in M\setminus S} (2-(4-e)x_j)\,.$$

An optimal solution satisfies that all nonzero variables have to be equal and $\sum_{i\in M} x_i = a$.

Proof. Consider an optimal solution x^*, and assume its support is $M' \subseteq M$. Let y^* be the restriction of x^* to M'. Then, $f_M(x^*) = f_{M'}(y^*)$ and y^* minimizes $f_{M'}$. Consider the function $f(y_1, y_2)$ as the function $f_{M'}$ restricted to the first two variables, while the others are fixed to y_i^*. Clearly, y_1^*, y_2^* minimize $f(y_1, y_2)$ subject to the constraints that $y_1, y_2 > 0$, and $y_1 + y_2 \leq a - \sum_{i\in M'\setminus\{1,2\}} y_i^*$. Now $f(y_1, y_2)$ can be written as

$$f(y_1,y_2) = \frac{A + By_1 + By_2 + Cy_1y_2}{(2+(e-2)y_1)(2+(e-2)y_2)}\,,$$

where

$$A = 4 \prod_{j\in[n]'} \tfrac{1}{2+(e-2)y_j^*} \sum_{S\subseteq[n]'} \tfrac{2^{|S|}}{1+|S|} \prod_{j\in S} y_j^* \prod_{j\in[n]'\setminus S} (2-(4-e)y_j^*)\,,$$

$$B = \tfrac{e-4}{2}A + 2 \prod_{j\in[n]'} \tfrac{1}{2+(e-2)y_j^*} \sum_{S\subseteq[n]'} \tfrac{2^{|S|+1}}{2+|S|} \prod_{j\in S} y_j^* \prod_{j\in[n]'\setminus S} (2-(4-e)y_j^*)\,,$$

$$C = \tfrac{(e-4)^2}{4}A + \tfrac{e-4}{2}B + \prod_{j\in[n]'} \tfrac{1}{2+(e-2)y_j^*} \sum_{S\subseteq[n]'} \tfrac{2^{|S|+2}}{3+|S|} \prod_{j\in S} y_j^* \prod_{j\in[n]'\setminus S} (2-(4-e)y_j^*)\,,$$

with $[n]' = M' \setminus \{1,2\}$. Since the constraint $y_1 + y_2 \leq a - \sum_{i\in[n]'} y_i^*$ is the only active constraint and it is symmetric with respect to y_1 and y_2, the KKT conditions dictate that a minimum of $f(y_1, y_2)$ satisfies

$$\frac{\partial f(z_1^*, z_2^*)}{\partial z_1^*} = \frac{\partial f(z_1^*, z_2^*)}{\partial z_2^*}\,. \tag{2.3}$$

[1]Because of the choice of π_i, we actually prove the slightly stronger bound where we maximize over $z_i \leq \frac{2q_i}{2-(e-2)q_i}$.

[2]The choice of π_i suggests that the random variables are not picked deterministically, but with probability less than 1, since $\pi_i < z_i^*$ if $z_i^* > 0$. However, as noted in the beginning of the proof, because of linearity of the objective in each variable, there is always an extreme optimal solution where the random variables are picked deterministically.

Taking the derivatives

$$
\begin{aligned}
\frac{\partial f(y_1,y_2)}{\partial y_1} &= \frac{2B+2y_2C-(\mathrm{e}-2)A-(\mathrm{e}-2)y_2B}{(2+(\mathrm{e}-2)y_1)^2(2+(\mathrm{e}-2)y_2)}\,,\\
\frac{\partial f(y_1,y_2)}{\partial y_2} &= \frac{2B+2y_1C-(\mathrm{e}-2)A-(\mathrm{e}-2)y_1B}{(2+(\mathrm{e}-2)y_1)(2+(\mathrm{e}-2)y_2)^2}\,,
\end{aligned}
$$

we see that (2.3) holds if and only if

$$
\big((4C-(\mathrm{e}-2)^2A)+(2(\mathrm{e}-2)C-(\mathrm{e}-2)^2B)(y_2+y_1)\big)\,(y_2-y_1)=0\,.
$$

So either $y_1=y_2$, or at least one is strictly positive and

$$
y_1+y_2=\frac{(\mathrm{e}-2)^2A-4C}{2(\mathrm{e}-2)C-(\mathrm{e}-2)^2B}\,.
$$

We evaluate the right-hand side of the latter, using the formulae for A, B, and C. Note first that $A\geq 0$, and observe that

$$
\begin{aligned}
B &= \tfrac{\mathrm{e}-4}{2}A+2\prod_{j\in[n]'}\tfrac{1}{2+(\mathrm{e}-2)y_j^*}\sum_{S\subseteq[n]'}\tfrac{2^{|S|+1}}{2+|S|}\prod_{j\in S}y_j^*\prod_{j\in[n]'\setminus S}(2-(4-\mathrm{e})y_j^*)\\
&\leq 0+4\prod_{j\in[n]'}\tfrac{1}{2+(\mathrm{e}-2)y_j^*}\sum_{S\subseteq[n]'}\tfrac{2^{|S|}}{2+|S|}\prod_{j\in S}y_j^*\prod_{j\in[n]'\setminus S}(2-(4-\mathrm{e})y_j^*)\\
&\leq A\,.
\end{aligned}
$$

Finally, we have

$$
\begin{aligned}
C &= \tfrac{(\mathrm{e}-4)^2}{4}A+\tfrac{\mathrm{e}-4}{2}B+\prod_{j\in[n]'}\tfrac{1}{2+(\mathrm{e}-2)y_j^*}\sum_{S\subseteq[n]'}\tfrac{2^{|S|+2}}{3+|S|}\prod_{j\in S}y_j^*\prod_{j\in[n]'\setminus S}(2-(4-\mathrm{e})y_j^*)\\
&\geq \tfrac{(\mathrm{e}-4)^2}{4}A+\tfrac{\mathrm{e}-4}{2}B+\prod_{j\in[n]'}\tfrac{1}{2+(\mathrm{e}-2)y_j^*}\sum_{S\subseteq[n]'}\tfrac{2^{|S|+1}}{2+|S|}\prod_{j\in S}y_j^*\prod_{j\in[n]'\setminus S}(2-(4-\mathrm{e})y_j^*)\\
&\geq \tfrac{(\mathrm{e}-4)^2}{4}A+\tfrac{\mathrm{e}-4}{2}B+\tfrac{1}{2}B-\tfrac{(\mathrm{e}-4)}{4}A\geq\left(\tfrac{(\mathrm{e}-4)^2}{4}+\tfrac{\mathrm{e}-4}{4}+\tfrac{1}{2}\right)A\geq\tfrac{1}{2}A\,.
\end{aligned}
$$

Thus, $(\mathrm{e}-2)^2A-4C\leq 0$ and $2(\mathrm{e}-2)C-(\mathrm{e}-2)^2B\geq 0$. Therefore, $y_1^*+y_2^*$ is negative which contradicts the constraint $y_1,y_2>0$. As the choice of the index $\{1,2\}$ is arbitrary we conclude that all coordinates of y^* have to be equal.

To finish the proof, we still need to show that in an optimal solution, the constraint $\sum_{\mathrm{i}\in M}x_{\mathrm{i}}^*=\sum_{\mathrm{i}\in M'}y_{\mathrm{i}}^*\leq a$ is tight. As all coordinates of y^* are equal, we know that $y_{\mathrm{i}}^*=\bar{y}$, for some $\bar{y}$. Let k denote the number of nonzero variables in x^*, i.e., $k=|M'|$. Then, abusing notation, we let

$$
\begin{aligned}
f_{M'}(\bar{y}) &= \prod_{j\in M'}\frac{1}{2+(\mathrm{e}-2)\bar{y}}\sum_{S\subseteq M'}\frac{2^{|S|}}{1+|S|}\prod_{j\in S}\bar{y}\prod_{j\in M'\setminus S}(2-(4-\mathrm{e})\bar{y})\\
&= (2+(\mathrm{e}-2)\bar{y})^{-k}\sum_{\ell=0}^{k}\binom{k}{\ell}\frac{2^{\ell}}{1+\ell}\bar{y}^{\ell}(2-(4-\mathrm{e})\bar{y})^{k-\ell}\\
&= \frac{(2+(\mathrm{e}-2)\bar{y})^{k+1}-(2-(4-\mathrm{e})\bar{y})^{k+1}}{(2+(\mathrm{e}-2)\bar{y})^{k}\cdot 2\bar{y}(k+1)}\,.
\end{aligned}
$$

We claim that $f_{M'}(\bar{y}+\varepsilon) \leq f_{M'}(\bar{y})$, for small $\varepsilon > 0$ and $\bar{y} < \frac{a}{k}$. Hereto, we take the derivative of $f_{M'}(\bar{y})$ w.r.t. $\bar{y}$, and show that this is nonpositive for $\bar{y} \geq 0$, from which the claim follows. We see that

$$\frac{\partial f_{M'}(\bar{y})}{\partial \bar{y}} = \frac{(2+(\mathrm{e}-2)\bar{y})^{-(k+1)}}{(k+1)\bar{y}^2}\left((2-(4-\mathrm{e})\bar{y})^k(2+(\mathrm{e}-2)\bar{y}+2k\bar{y}) - (2+(\mathrm{e}-2)\bar{y})^{k+1}\right).$$

As $\bar{y} \geq 0$, it is easy to see that the sign of the derivative is equal to the sign of $(2-(4-\mathrm{e})\bar{y})^k(2+(\mathrm{e}-2)\bar{y}+2k\bar{y}) - (2+(\mathrm{e}-2)\bar{y})^k(2+(\mathrm{e}-2)\bar{y})$. Therefore, to show that the derivative is nonpositive for $\bar{y} \geq 0$, we need to show that

$$(2+(\mathrm{e}-2)\bar{y})^{k+1} \geq (2-(4-\mathrm{e})\bar{y})^k(2+(\mathrm{e}-2)\bar{y}+2k\bar{y}). \tag{2.4}$$

We will prove this inequality by induction on k. For $k=1$, we have

$$\begin{aligned}(2+(\mathrm{e}-2)\bar{y})^2 &= (2-(4-\mathrm{e})\bar{y})(2+(\mathrm{e}-2)\bar{y}+2\bar{y}) + 4\bar{y}^2\\ &\geq (2-(4-\mathrm{e})\bar{y})(2+(\mathrm{e}-2)\bar{y}+2\bar{y}).\end{aligned}$$

Assume that (2.4) is true for given k. Then,

$$\begin{aligned}(2+(\mathrm{e}-2)\bar{y})^{k+2} &\geq (2-(4-\mathrm{e})\bar{y})^k(2+(\mathrm{e}-2)\bar{y}+2k\bar{y})(2-(4-\mathrm{e})\bar{y}+2\bar{y})\\ &= (2-(4-\mathrm{e})\bar{y})^{k+1}(2+(\mathrm{e}-2)\bar{y}+2k\bar{y})\\ &\quad + (2-(4-\mathrm{e})\bar{y})^k(2-(4-\mathrm{e})\bar{y}+2(k+1)\bar{y})2\bar{y}\\ &\geq (2-(4-\mathrm{e})\bar{y})^{k+1}(2+(\mathrm{e}-2)\bar{y}+2k\bar{y}) + (2-(4-\mathrm{e})\bar{y})^{k+1}2\bar{y}\\ &= (2-(4-\mathrm{e})\bar{y})^{k+1}(2+(\mathrm{e}-2)\bar{y}+2(k+1)\bar{y}),\end{aligned}$$

where the first inequality is due to the induction hypothesis. Hence, (2.4) is true. For each $k \geq 1$ and $\bar{y} \geq 0$, the derivative is nonpositive, and $f_{M'}(\bar{y})$ is minimized for $\bar{y}$ as large as possible, that is, $\sum_{\mathrm{i}\in M'} \bar{y} = a$. □

Using Proposition 2.2, we lower bound (2.2) as follows. Consider the term

$$\left(\prod_{j\in[n]\setminus\{\mathrm{i}\}} \frac{1}{2+(\mathrm{e}-2)z_j^*}\right) \sum_{S\subseteq[n]\setminus\{\mathrm{i}\}} \frac{2^{|S|}}{1+|S|} \prod_{j\in S} z_j^* \prod_{j\in[n]\setminus(S\cup\{\mathrm{i}\})} (2-(4-\mathrm{e})z_j^*).$$

Note that this is equal to $f_{N\setminus\{\mathrm{i}\}}(z^*_{-\mathrm{i}})$.[3] So, Proposition 2.2 can be applied with $a = 1 - z^*_{\mathrm{i}}$. Thus,

$$f_{N\setminus\{\mathrm{i}\}}(z^*_{-\mathrm{i}}) \geq f_{N\setminus\{\mathrm{i}\}}(x^*),$$

with x^* the optimal solution to $\min_{x\in\mathbb{R}_+^{N\setminus\{\mathrm{i}\}}}\{f_{N\setminus\{\mathrm{i}\}}(x) : \sum_{j\in N\setminus\{\mathrm{i}\}} x_j \leq a\}$.

Proposition 2.2 states that $x_j^* = (1-z_{\mathrm{i}}^*)/k$, where $k \leq n-1$ is the number of nonzero variables in x^*. Conditioning on the cardinality of the set S, and using the Binomial Theorem, it is straightforward to show that

$$f_{[n]\setminus\{\mathrm{i}\}}(x^*) = \frac{2k+(\mathrm{e}-2)(1-z_{\mathrm{i}}^*)}{2(k+1)(1-z_{\mathrm{i}}^*)}\left(1-\left(1-\frac{2(1-z_{\mathrm{i}}^*)}{2k+(\mathrm{e}-2)(1-z_{\mathrm{i}}^*)}\right)^{k+1}\right).$$

[3] $x_{-\mathrm{i}}$ denotes the vector x with coordinate i eliminated.

As this quantity only depends on k and z_{i}^*, we may define

$$\varphi_k(z_{\mathrm{i}}^*) = \frac{2}{(2+(\mathrm{e}-2)z_{\mathrm{i}}^*)} f_{[n]\setminus\{\mathrm{i}\}}(x^*)\,,$$

to conclude that expression (2.2) (and in turn (P)) is lower bounded by

$$\sum_{\mathrm{i}\in[n]} b_{\mathrm{i}} z_{\mathrm{i}}^* \varphi_{k(\mathrm{i})}(z_{\mathrm{i}}^*)\,.$$

where the index $k(\mathrm{i})$, denoting the number of nonzero variables in x^*, is always at least 1, yet may vary, depending on i.

The remainder of the proof establishes that $\varphi_{k(\mathrm{i})}(z_{\mathrm{i}}^*) \geq 1 - \frac{1}{\mathrm{e}}$. Indeed, we show that, for all $y \in [0,1]$ and for all $n \geq 1$, we have that $\varphi_n(y) \geq 1 - \frac{1}{\mathrm{e}}$. To prove it, We start with a lemma that rephrases this claim.

Lemma 2.3 *Let*

$$f_n(x) := \frac{1}{n+1} - \frac{(1-x)^{n+1}}{n+1} - \frac{\mathrm{e}-1}{2}x + \frac{(\mathrm{e}-1)(\mathrm{e}-2)n}{\mathrm{e}(2-(\mathrm{e}-2)x)}x^2\,.$$

Then $\varphi_n(y) \geq 1 - \frac{1}{\mathrm{e}}$ *for all* $y \in [0,1]$ *and all* $n \geq 2$ *if and only if* $f_n(x) \geq 0$ *for all* $n \geq 1$ *and* $x \in [0, \bar{x}]$*, where* $\bar{x} = 1/(n-1+\mathrm{e}/2)$.

Proof. Consider the variable change

$$x = \frac{2(1-y)}{2(n-1)+(\mathrm{e}-2)(1-y)} \implies y = \frac{2-(2(n-1)+\mathrm{e}-2)x}{2-(\mathrm{e}-2)x}\,.$$

As y ranges from 0 to 1, x ranges from 0 to $\frac{1}{n-2+\mathrm{e}/2}$. Note that

$$\frac{2}{2+(\mathrm{e}-2)y} = \frac{2(2-(\mathrm{e}-2)x)}{\mathrm{e}(2-(\mathrm{e}-2)x)-2(\mathrm{e}-2)(n-1)x}\,.$$

Substituting this, we see that $\varphi_n(y) \geq \frac{\mathrm{e}-1}{\mathrm{e}}$ for all $y \in [0,1]$ and $n \geq 2$ is equivalent to

$$\frac{1}{n}(1-(1-x)^n) \geq \frac{\mathrm{e}-1}{\mathrm{e}}x\frac{\mathrm{e}(2-(\mathrm{e}-2)x)-2(\mathrm{e}-2)(n-1)x}{2(2-(\mathrm{e}-2)x)}\,,$$

for all $x \in [0, \frac{1}{n-2+\mathrm{e}/2}]$ and $n \geq 2$. Moving the index of n by 1, the result follows. □

To prove $f_n(x) \geq 0$, we use the following result.

Proposition 2.4 *If* $c \in [0, \frac{1}{2}]$ *then* $f(x) = \left(1 - \frac{1}{x+c}\right)^x$ *is a non-decreasing function of* x *in* $(1, \infty)$.

Proof. Define $g(x) = \ln(f(x)) = x \ln\left(1 - \frac{1}{x+c}\right)$. We prove $f(x)$ is nondecreasing by proving that $g'(x) \geq 0$. Note that

$$g'(x) = \ln\left(\frac{x+c-1}{x+c}\right) + \frac{x}{(x+c-1)(x+c)}\,,$$

which is nonnegative if and only if

$$\frac{x}{(x+c-1)(x+c)} \geq \ln\left(1 + \frac{1}{x+c-1}\right).$$

We substitute $\frac{1}{x+c-1} = z$. Then, the left-hand side becomes

$$\frac{x}{(x+c-1)(x+c)} = \frac{\frac{1}{z} - c + 1}{\frac{1}{z} + 1} z = \frac{1 + z - cz}{1+z} z = \frac{z}{1+z}(1 + (1-c)z).$$

We expand $\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} \pm \ldots$, so it is sufficient to prove

$$\frac{z}{1+z}(1 + (1-c)z) \geq z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} \pm \ldots.$$

We multiply both sides by $\frac{1+z}{z}$ to retrieve

$$1 + (1-c)z \geq 1 + \frac{z}{2} - \frac{z^2}{6} + \frac{z^3}{12} - \frac{z^4}{20} \pm \ldots.$$

As $c \leq \frac{1}{2}$, it suffices to prove $-\frac{z^2}{6} + \frac{z^3}{12} - \frac{z^4}{20} \pm \ldots \leq 0$, i.e.,

$$\sum_{\mathrm{i}=2}^{\infty} \frac{(-1)^{\mathrm{i}} z^{\mathrm{i}}}{\mathrm{i}(\mathrm{i}+1)} \geq 0.$$

We rewrite this as

$$\sum_{\mathrm{i}=2}^{\infty} \frac{1}{z} \frac{(-1)^{\mathrm{i}} z^{\mathrm{i}+1}}{\mathrm{i}(\mathrm{i}+1)} = \sum_{\mathrm{i}=2}^{\infty} \frac{1}{z} \int_0^z \frac{(-1)^{\mathrm{i}} t^{\mathrm{i}}}{\mathrm{i}} \mathrm{d}t = \frac{1}{z} \int_0^z \sum_{\mathrm{i}=2}^{\infty} \frac{(-1)^{\mathrm{i}} t^{\mathrm{i}}}{\mathrm{i}} \mathrm{d}t.$$

However,

$$\sum_{\mathrm{i}=2}^{\infty} \frac{(-1)^{\mathrm{i}} t^{\mathrm{i}}}{\mathrm{i}} = -\sum_{\mathrm{i}=1}^{\infty} \frac{(-1)^{\mathrm{i}+1} t^{\mathrm{i}}}{\mathrm{i}} + t = -\ln(1+t) + t \geq 0,$$

where the last inequality follows from $t \geq \ln(1+t)$ for $t \geq 0$. □

Lemma 2.5 *$f_n(x) \geq 0$ for all $n \geq 1$ and $x \in [0, \bar{x}]$.*

Proof. We break the proof in the following parts which together imply the result. All derivatives are with respect to x.

(i) $f_n(0) = 0$ for all $n \geq 1$,

(ii) $f_n(\bar{x}) \geq 0$ for all $n \geq 1$,

(iii) $f_n'(0) > 0$ for all $n \geq 1$,

(iv) $f_n'(\bar{x}) < 0$ for all $n \geq 1$,

(v) $f_n'''(x) > 0$ for all $x \in [0, \bar{x}]$ and $n \geq 1$.

First we show how the lemma follows from these parts. We prove $f_n(x) \geq 0$ by contradiction. Assume that for some n there exists an $x_1 \in [0, \bar{x}]$ such that $f_n(x_1) < 0$. As $f_n(x)$ is differentiable and $f_n(\bar{x}) \geq f_n(0) = 0$, there exists an x_1 such that $f_n'(x_1) = 0$. Since the function increases from a negative value in x_1 to a nonnegative value in $\bar{x}$, there exists some $x_2 \in (x_1, \bar{x})$ such that $f_n'(x_2) > 0$. However, as $f_n'(\bar{x}) < 0$ and $f_n'(x_2) > 0$, there exists an $x_3 \in (x_2, \bar{x})$ with $f_n''(x_3) = 0$. By symmetry, the same analysis holds in the interval $(0, x_1)$ and therefore there also exists an $x_4 < x_1$ with $f_n''(x_4) = 0$. However, this contradicts (v) as f'' is strictly increasing in x.

To prove the statements, we compute the first three derivatives of f with respect to x.

$$
\begin{aligned}
f_n'(x) &= (1-x)^n - \frac{\mathrm{e}-1}{2} + \frac{n(\mathrm{e}-1)(\mathrm{e}-2)}{\mathrm{e}} \left(\frac{4x - (\mathrm{e}-2)x^2}{(2-(\mathrm{e}-2)x)^2} \right) , \\
f_n''(x) &= n \left(-(1-x)^{n-1} + \frac{8(\mathrm{e}-1)(\mathrm{e}-2)}{\mathrm{e}(2-(\mathrm{e}-2)x)^3} \right) , \\
f_n'''(x) &= n \left((n-1)(1-x)^{n-2} + \frac{24(\mathrm{e}-1)(\mathrm{e}-2)^2}{(2-(\mathrm{e}-2)x)^4} \right) .
\end{aligned}
$$

We finish the proof by proving the five statements.

(i) $f_n(0) = 0$ for all $n \geq 1$ by a direct calculation.

(ii) $f_n(\bar{x}) \geq 0$ for all $n \geq 1$ is equivalent to $\varphi_n(z_{\mathrm{i}}^*) \geq 1 - 1/\mathrm{e}$ for $z_{\mathrm{i}}^* = 0$ and all $n \geq 2$. By direct evaluation, we see this is true for $n = 2, 3, 4$. Thus, we need to prove that for all $n \geq 5$,

$$
\frac{n-1+\frac{\mathrm{e}-2}{2}}{n} \left(1 - \left(1 - \frac{1}{n-1+\frac{\mathrm{e}-2}{2}} \right)^n \right) \geq 1 - \frac{1}{\mathrm{e}} ,
$$

or equivalently, that for all $n \geq 4$,

$$
\frac{n-1+\frac{\mathrm{e}}{2}}{n+1} \left(1 - \left(1 - \frac{1}{n-1+\frac{\mathrm{e}}{2}} \right)^{n+1} \right) \geq 1 - \frac{1}{\mathrm{e}} .
$$

We write this as

$$
\left(1 - \frac{1}{n-1+\frac{\mathrm{e}}{2}} \right)^{n+1} \leq 1 - \frac{(n+1)(1-\frac{1}{\mathrm{e}})}{n-1+\frac{\mathrm{e}}{2}} = \frac{\frac{n+1}{\mathrm{e}} + \frac{\mathrm{e}}{2} - 2}{n-1+\frac{\mathrm{e}}{2}} ,
$$

and multiplying by $(1 - 1/(n-1+\mathrm{e}/2))^{(\mathrm{e}-5)/2}$ yields

$$
\begin{aligned}
\left(1 - \frac{1}{n-1+\frac{\mathrm{e}}{2}} \right)^{n+\frac{\mathrm{e}}{2}-\frac{3}{2}} &\leq \frac{\frac{n+1}{\mathrm{e}} + \frac{\mathrm{e}}{2} - 2}{n-2+\frac{\mathrm{e}}{2}} \left(1 - \frac{1}{n-1+\frac{\mathrm{e}}{2}} \right)^{\frac{\mathrm{e}}{2}-\frac{3}{2}} \\
&= \frac{(\frac{n+1}{\mathrm{e}} + \frac{\mathrm{e}}{2} - 2)(n-1+\frac{\mathrm{e}}{2})^{\frac{3-\mathrm{e}}{2}}}{(n-2+\frac{\mathrm{e}}{2})^{\frac{5-\mathrm{e}}{2}}} .
\end{aligned}
$$

By invoking Proposition 2.4 in the left-hand side, we see that this is a non-decreasing function in n with limit $1/\mathrm{e}$. For the right-hand side, note that the limit for n to infinity is also $1/\mathrm{e}$, and its derivative with respect to n is

$$
4\frac{3-\mathrm{e}}{\mathrm{e}}(1-(3-\mathrm{e})n)\frac{(2n+\mathrm{e}-2)^{\frac{1-\mathrm{e}}{2}}}{(2n+\mathrm{e}-4)^{\frac{7-\mathrm{e}}{2}}} ,
$$

which is negative for $n \geq 4$. The proof of (ii) is complete.

(iii) $f'_n(0) = 1 - (\mathrm{e}-1)/2 > 0$.

(iv) For $n = 1$ and $n = 2$, direct evaluation of $f'_1(\bar{x})$ and $f'_2(\bar{x})$ gives negative values. For $n \geq 3$, proving that

$$f'_n(\bar{x}) = \left(1 - \frac{1}{n + \frac{\mathrm{e}}{2} - 1}\right)^n + \frac{(\mathrm{e}-1)((\mathrm{e}-2)^2 + 2n(\mathrm{e}-4))}{4\mathrm{e}n} < 0\,,$$

is equivalent to proving that

$$\mathrm{e}\left(1 - \frac{1}{n + \frac{\mathrm{e}}{2} - 1}\right)^n + \frac{(\mathrm{e}-1)(\mathrm{e}-2)^2}{4n} < \frac{(\mathrm{e}-1)(4-\mathrm{e})}{2}\,.$$

By taking the limit of, and invoking Proposition 2.4 for the first term, and using $n \geq 3$ for the second term, we get

$$\mathrm{e}\left(1 - \frac{1}{n + \frac{\mathrm{e}}{2} - 1}\right)^n + \frac{(\mathrm{e}-1)(\mathrm{e}-2)^2}{4n} < 1 + \frac{(\mathrm{e}-1)(\mathrm{e}-2)^2}{12} < \frac{(\mathrm{e}-1)(4-\mathrm{e})}{2}\,.$$

(v) Since $0 \leq x \leq \bar{x} < 1$ for all n, $f'''_n(x)$ consists of sums, products, and quotients of only strictly positive terms.

□

From Lemmata 2.3 and 2.5 we conclude that $\varphi_n(z^*_\mathrm{i}) \geq 1 - \frac{1}{\mathrm{e}}$, $\forall z^*_\mathrm{i} \in [0,1], n \geq 2$. □

2.3 Main Theorem

Consider n random variables $X_1, \ldots, X_n$ with corresponding distributions $F_1, \ldots, F_n$. Let $F_\mathrm{i}^{-1}(q) = \inf\{x \geq 0 \mid F(x) \geq q\}$ be the generalized inverse of F_i and $\tau_\mathrm{i}(q) = F_\mathrm{i}^{-1}(1-q)$ be the *threshold function*.

Note that for values of q for which $\tau(\cdot)$ is not constant (i.e., where F is continuous), $\tau(q)$ is such that $\mathbb{P}(X > \tau(q)) = q$. Thus, we can interpret $\tau(q)$ as a value for which the *acceptance probability* equals q (i.e., the probability of X being above $\tau(q)$ equals q). In general, $\tau(q)$ may have mass and one can only assert that $\mathbb{P}(X \geq \tau(q)) \geq q \geq \mathbb{P}(X > \tau(q))$.

We will construct the threshold rule choosing a priori the probability of acceptance q_i for every random variable X_i, $\mathrm{i} = 1, \ldots, n$. Therefore, when facing a random variable X_i, we will select it whenever $X_\mathrm{i} > \tau_\mathrm{i}(q_\mathrm{i})$, so the acceptance probability is at most q_i. If $\tau_\mathrm{i}(q_\mathrm{i})$ has mass, there might be no value that accomplish the previous condition. We circumvent this allowing our rule to randomize its decision: with probability s_i the rule will choose the random variable if $X_\mathrm{i} \geq \tau_\mathrm{i}(q_\mathrm{i})$, and with probability $1 - s_\mathrm{i}$, if $X_\mathrm{i} > \tau_\mathrm{i}(q_\mathrm{i})$, where $s_\mathrm{i} = [q_\mathrm{i} - \mathbb{P}(X_\mathrm{i} > \tau_\mathrm{i}(q_\mathrm{i})] / \mathbb{P}(X_\mathrm{i} = \tau_\mathrm{i}(q_\mathrm{i}))$. In this way, the probability of selecting a certain random variable, conditioned in this variable being faced, is fixed to be equal to the probability of acceptance that defines the corresponding threshold.

We now derive the main theorem of this section using the Bernoulli Selection Lemma.

Theorem 2.6 *Given n independent non-negative random variables $X_1, \dots, X_n$, there exist thresholds $\tau_1, \dots, \tau_n$ such that*

$$\mathbb{E}\left[\max\{X_1, \dots, X_n\}\right] \leq \frac{\mathrm{e}}{\mathrm{e}-1}\mathbb{E}\left[\frac{\sum_{\mathrm{i}\in R} X_\mathrm{i}}{|R|}\right],$$

where $R = \{\mathrm{i} \in [n] \mid X_\mathrm{i} > \tau_\mathrm{i}\}$ is the set of random variables above their corresponding thresholds.

Furthermore, the bound of $\mathrm{e}/(\mathrm{e}-1)$ is the best possible.

Note that the quantity on the right exactly corresponds to the expected value of the first X_i above τ_i, when the X_i's are ordered uniformly at random.

Before proving the theorem, we state and prove the following technical result.

Lemma 2.7 *Let $X_1, \dots, X_n$ be non-negative random variables with corresponding distribution function $F_1, \dots, F_n$. Set $q_\mathrm{i} = \mathbb{P}(X_\mathrm{i} = \max_{j\in[n]} X_j)$ and α_i a value for which $1 - F_\mathrm{i}(\alpha_\mathrm{i}) = q_\mathrm{i}$. Then,*

$$\mathbb{E}[X_\mathrm{i} \mid X_\mathrm{i} = \max_{j\in[n]} X_j] \leq \mathbb{E}[X_\mathrm{i} \mid X_\mathrm{i} > \alpha_\mathrm{i}].$$

If $F_\mathrm{i}^{-1}(\{1-q_\mathrm{i}\})$ is empty for some $\mathrm{i} = 1, \dots, n$, the result still holds via randomization.

Proof. For $x > \alpha_\mathrm{i}$, we have

$$\mathbb{P}(X_\mathrm{i} > x \mid X_\mathrm{i} > \alpha_\mathrm{i}) = \int_x^\infty \frac{1}{q_\mathrm{i}} \mathrm{d}F_\mathrm{i}(t),$$

while, if $x \leq \alpha_\mathrm{i}$, the previous probability equals 1. On the other hand,

$$\mathbb{P}(X_\mathrm{i} > x \mid X_\mathrm{i} = \max_{j\in[n]} X_j) = \int_x^\infty \frac{\prod_{j\neq \mathrm{i}} F_j(t)}{q_\mathrm{i}} \mathrm{d}F_\mathrm{i}(t).$$

From this, it follows that $\mathbb{P}(X_\mathrm{i} > x \mid X_\mathrm{i} > \alpha_\mathrm{i}) \geq \mathbb{P}(X_\mathrm{i} > x \mid X_\mathrm{i} = \max_{j\in[n]} X_j)$ for all $x \geq 0$. Thus, $X_\mathrm{i} \mid (X_\mathrm{i} > \alpha_\mathrm{i})$ stochastically dominates $X_\mathrm{i} \mid (X_\mathrm{i} = \max_{j\in[n]} X_j)$, and the conclusion follows.

When some F_i are not continuous, it could be the case that there is no α_i such that $1 - F_\mathrm{i}(\alpha_\mathrm{i}) = q_\mathrm{i}$ or that $\sum q_\mathrm{i} > 1$. If the former happens, the result still holds provided α_i is chosen randomly between the biggest and lowest values which distribution is closer to $1 - q_\mathrm{i}$. The latter case is solved by slightly perturbing the support of the random variables in a way that the probability that two or more are the maximum simultaneously is negligible. □

In the remaining of the chapter, we will consider that the random variables are slightly perturbed such that the probability of two or more being the maximum simultaneously is negligible. As well, to simplify notation and the understanding of the proof, whenever $\tau_i(q)$ has mass , we will not condition on the decision of the threshold rule. We will refer to the threshold rule as to stop every time a random variable is strictly above its corresponding threshold. In virtue of Lemma 2.7, all arguments will still apply to the case when we randomize between strict inequality and weak inequality.

Proof of Theorem 2.6. If $\mathbb{E}[X_i] = \infty$ for some i, the result is trivial. Otherwise, let q_i be the probability that $X_i = \max_{j\in[n]} X_j$. Consider $b_i = \mathbb{E}[X_i \mid X_i > \tau_i(q_i)]$ and the Bernoulli random variables $Z_1, \ldots, Z_n$ where Z_i has parameter q_i. We apply the Bernoulli Selection Lemma to this instance, and thus let $S \subseteq [n]$ be a set for which the inequality of the lemma holds.

Now define $\tau_i = \tau_i(q_i)$ for $i \in S$ and $\tau_i = \infty$ otherwise, and consider the random variables Y_i as the indicator of the event $\{X_i > \tau_i\}$. Note that for $i \notin S$, we have $Y_i = 0$ almost surely, and for $i \in S$, we have $\mathbb{P}(X_i > \alpha_i) = \mathbb{P}(Y_i = 1) = q_i$. It follows that

$$\begin{aligned}
\mathbb{E}\left[\frac{\sum_{i\in R} X_i}{|R|}\right] &= \mathbb{E}\left[\frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n Y_i}\right] \\
&= \sum_{i\in S} \mathbb{E}\left[\frac{X_i Y_i}{\sum_{j\in S} Y_j}\right] \\
&= \sum_{i\in S} \mathbb{E}[X_i \mid Y_i = 1]\mathbb{E}\left[(1 + \sum_{j\in S\setminus\{i\}} Y_j)^{-1} \middle| Y_i = 1\right] \mathbb{P}(Y_i = 1) \\
&= \sum_{i\in S} \mathbb{E}[X_i \mid X_i > \alpha_i]\mathbb{E}\left[\frac{Y_i}{\sum_{j\in S} Y_j}\right] \\
&= \mathbb{E}\left[\frac{\sum_{i\in S} \mathbb{E}[X_i \mid X_i > \alpha_i] Z_i}{\sum_{i\in S} Z_i}\right] \\
&\geq \frac{e-1}{e} \max_{z_i \leq q_i} \left\{ \sum_{i=1}^n \mathbb{E}[X_i \mid X_i > \alpha_i] z_i \,\middle|\, \sum_{i=1}^n z_i \leq 1 \right\} \\
&\geq \frac{e-1}{e} \sum_{i=1}^n \mathbb{E}[X_i \mid X_i > \alpha_i] q_i \,,
\end{aligned}$$

where the second to last inequality follows from the Bernoulli Selection Lemma, while the last holds since $\sum_{i=1}^n q_i = 1$. Now note that $\mathbb{E}[\max\{X_1, \ldots, X_n\}] = \sum_{i=1}^n \mathbb{E}[X_i \mid X_i = \max_{j\in[n]} X_j] q_i$. Lemma 2.7 leads to conclude. □

Note that since the expected reward of the strategy is linear for each probability of acceptance, it is always the case that at least one of the choices of the threshold rule (whether apply strict inequality or weak inequality) reports a conditional expected value larger or equal to $\mathbb{E}(X_i \mid X_i = \max_{j\in[n]} X_j)$, and thus if we can determine which decision is better, no randomization in the threshold rule decision is needed.

Tightness

We now provide a family of instances that show that the $1-1/\mathrm{e}$ bound in the Bernoulli Selection Lemma is actually best possible. Since the context of the lemma is a particular case for the hypothesis of Theorem 2.6, this example is also a lower bound for the general case and proves the tightness of our result.

Consider n^2 independent identically distributed Bernoulli random variables with parameter $1/n$ and prizes $b_1 = n/(\mathrm{e}-2)$ and $b_\mathrm{i} = 1$ for $2 \leq \mathrm{i} \leq n^2$. The expectation of the maximum prize is given by

$$\mathbb{E}\left[\max_{\mathrm{i}\in[n]} \{b_\mathrm{i} X_\mathrm{i}\}\right] = \frac{1}{\mathrm{e}-2} + \left(1-\frac{1}{n}\right)\left(1-\left(1-\frac{1}{n}\right)^{n^2-1}\right) \longrightarrow \frac{1}{\mathrm{e}-2} + 1\,.$$

In this particular setting, where the Bernoulli random variables are i.i.d., the best strategy is to sort by prize and take some subset with those of higher prize. This means to choose the first random variable and a subset of size $k-1$ of the rest for some $1 \leq k \leq n^2$. This yields an expected average value that can be upper bounded by

$$\left(1-\left(1-\frac{1}{n}\right)^k\right)\frac{\frac{n}{\mathrm{e}-2}+k-1}{k} \leq \left(1-\left(1-\frac{1}{n}\right)^k\right)\left(\frac{n}{k(\mathrm{e}-2)}+1\right).$$

This in turn is less or equal to

$$\max_{0\leq x\leq n} (1-\mathrm{e}^{-x})\left(\frac{1}{x(\mathrm{e}-2)}+1\right),$$

where $x = \frac{k}{n}$. To solve this maximization problem, ee compute the first two derivatives of the objective:

$$f'(x) = \frac{-\mathrm{e}^x + (\mathrm{e}-2)x^2 + x + 1}{(\mathrm{e}-2)x^2}\mathrm{e}^{-x}\,,$$

$$f''(x) = \frac{2\mathrm{e}^x - ((\mathrm{e}-2)x^3 + x^2 + 2x + 2)}{(\mathrm{e}-2)x^3}\mathrm{e}^{-x}\,.$$

We see that $f'(1) = 0$. To show that $x = 1$ is a global maximum, we prove that $f'(x) > 0$ for $x < 1$ and $f'(x) < 0$ for $x > 1$. To see this, first note that $f''(x)$ has the same sign as the function

$$g(x) = 2\mathrm{e}^x - \big((\mathrm{e}-2)x^3 + x^2 + 2x + 2\big)\,.$$

Note further that $g(0) = 0$. Since this is an exponential function with a positive coefficient minus a polynomial with only positive coefficients, $g(x)$ first decreases until some point because of the polynomial, after which it is increasing, because of the exponential term that starts to dominate the polynomial. So there exists some $x^* > 0$ such that $g(x) < 0$ for $x < x^*$, $g(x^*) = 0$ and $g(x) > 0$ for $x > x^*$. Because $g(1) = \mathrm{e} - 3 < 0$, we know $x^* > 1$. Therefore, $f''(x) < 0$ up to $x^* > 1$ and $f''(x) > 0$ afterwards, and hence, $f'(x)$ is decreasing up to $x^* > 1$ and increasing afterwards.

Since $f'(1) = 0$ and $f'(x)$ is decreasing for $x \leq 1$, we know $f'(x) > 0$ for $x < 1$. Furthermore, $f'(x) < 0$ for $1 < x \leq x^*$, since $f'(1) = 0$ and $f'(x)$ is decreasing for $1 < x < x^*$. Since

$\lim_{x\to\infty} f'(x) = 0$ and $f'(x)$ is increasing from x^* onwards, we know $f'(x) < 0$ for $x > x^*$, and hence, $f'(x) < 0$ for all $x > 1$. Therefore, $x = 1$ is a global maximum. This yields the value

$$(1 - \mathrm{e}^{-1})\left(\frac{1}{\mathrm{e}-2} + 1\right) = (1 - 1/\mathrm{e})\mathbb{E}\left[\max_{\mathrm{i}\in N}\{b_\mathrm{i} X_\mathrm{i}\}\right].$$

Algorithm

Although Theorem 2.6 states only the existence of thresholds that accomplish the result, the proof is constructive and provides a way to compute their values as a function of the probability distribution function of the random variables. When distributions are continuous (and after proper perturbations of the support) the algorithm takes the following form:

Algorithm 2.1: Non-adaptive thresholds rule

Input: Distributions $F_1, \dots, F_n$, values $X_1, \dots, X_n$ draw from the corresponding distribution.

Initialize $R = \emptyset$, $j = 0$, $X_0 = 0$
for i $= 1$ **to** n **do**
  Compute $q_\mathrm{i} = \mathbb{P}(X_\mathrm{i} = \max_{j\in[n]} X_j)$
  Set threshold $\tau_\mathrm{i} = \begin{cases} F^{-1}(1-q_\mathrm{i}) & \text{w.p. } \frac{2}{2+(\mathrm{e}-2)q_\mathrm{i}} \\ \infty & \text{otherwise.} \end{cases}$
  if $X_\mathrm{i} > \tau_\mathrm{i}$ **then**
    $R \leftarrow R \cup \{\mathrm{i}\}$
  end if
end for
if $R \neq \emptyset$ **then**
  Select j uniformly at random from R
end if
return X_j

In the general case, it suffices to change the condition $X_\mathrm{i} > \tau_\mathrm{i}$ by the randomized version of the threshold rule, where with some probability the condition changes to $X_\mathrm{i} \geq \tau_\mathrm{i}$. Of course, we can derandomized if we can assert which of the two conditions reports larger value.

Remark 2.1 Notice that with probability $1 - \frac{2q_\mathrm{i}}{2+(\mathrm{e}-2)q_\mathrm{i}}$ random variable X_i is never selected. Therefore, the probability that the algorithm does not chose any of the random variables (and get value zero) can be lower bounded by

$$\prod_{\mathrm{i}=1}^{n}\left(1 - \frac{2q_\mathrm{i}}{2+(\mathrm{e}-2)q_\mathrm{i}}\right) \geq \left(1 - \frac{2}{\mathrm{e}}\right)^n \geq 1 - \frac{2}{\mathrm{e}}.$$

In consequence, the probability of actually accomplish a non-zero value is at most $2/\mathrm{e} \approx 0.74$. The intuition to this is that if we shoot for an algorithm that selects too frequently, we risk ending up with a value too low.

Chapter 3

Prophet Inequality for I.I.D. Distributions

In this chapter we study the existence of thresholds to solve the problem of selecting a value given by a uniformly random sequence of independent identically distributed random variables. We provide a constructive proof of a rule that achieves a factor of approximately 0.745 over the expected maximum, an algorithm to implement this rule, and an example that shows that these results are tight. By doing so, we answer a more than 30 years old open problem in the area of optimal stopping.

3.1 Preliminaries

For fixed $n > 1$, let $X_1, \ldots, X_n$ be non-negative, independent and identically distributed random variables and T_n their set of stopping rules. In 1982, Hill and Kertz [13] proved that

$$\mathbb{E}(\max\{X_1, \ldots, X_n\}) \leq a_n \sup\{\mathbb{E}(X_t) : t \in T_n\},$$

where a_n is a constant depending on n and such that for all $n > 1$, $1.1 < a_n < 1.6$. Although they proved the constant a_n is tight for every n, 1.6 is not the best bound. Kertz [14] conjectured that the best upper bound for any n was $\beta^* \approx 1.341$, where β^* is the unique solution to

$$1 + \int_0^1 \frac{1}{y(\ln(y) - 1) - (\beta - 1)} \mathrm{d}y = 0, \tag{3.1}$$

and furthermore proved that $a_n \to \beta^*$.

More recently, Abolhassan et al. [1] improved this bound to 1.355, getting closer to the answer, but leaving the question still open. In this part of the work, we prove that β^* is actually the best possible bound, closing the by now 30 year old conjecture. To accomplish this result, we consider, as in [20], a *threshold rule*, where we choose to stop the first time we see a variable above certain value. Unlike them, however, where the same threshold is set for all the variables, we define a decreasing sequence of threshold for every one of the variables X_i, i $= 1, \ldots, n$. Since any threshold rule defines a stopping time, our result answers positively the three decades conjecture.

This type of inequalities are commonly called *prophet inequalities*, since the factor represents the odds a player aiming to get the largest value must be paid when betting against a prophet who knows exactly which the largest value is.

3.2 Properties of the Threshold Rule

For $X_1, \ldots, X_n$ non-negative i.i.d. random variables, we take F as their probability distribution function and for simplicity, we will often refer to X as a random variable with the same common distribution. As in the previous chapter, let F^{-1} be the generalized inverse of F and $\tau(q) = F^{-1}(1-q)$.

Again, we can interpret $\tau(q)$ as a value for which the acceptance probability equals q. In general, $\tau(q)$ may have mass and one can only assert that $\mathbb{P}(X \geq \tau(q)) \geq q \geq \mathbb{P}(X > \tau(q))$.

We will construct the threshold rule choosing the probability of acceptance on every step, in the same way in Section 2.3. For a random variable X, we select a proper acceptance probability q and will choose to stop if $X > \tau(q)$. If $\tau(q)$ has mass, we randomize. Again, the probability of stopping in a certain step, conditioned in having reached that step, is fixed to be equal to the probability of acceptance that defines the corresponding threshold.

Let $R(q)/q$ be the expected value of the random variable X given the threshold rule decided to stop when its acceptance probability was at most q. Note that if $q = \mathbb{P}(X > \tau(q))$, then $R(q) = \mathbb{E}(X \mid X > \tau(q))q$.

We begin with the following lemma:

Lemma 3.1

$$\mathbb{E}(\max\{X_1, \ldots, X_n\}) = n \int_0^1 (n-1)(1-q)^{n-2} R(q) \mathrm{d}q. \tag{3.2}$$

Proof. Let $\bar{q} = \mathbb{P}(X > \tau(q))$ and note that $\tau(\bar{q}) = \tau(q)$. It is straightforward to check that if $q = \bar{q}$, then $R(q) = \mathbb{E}(X \mid X > \tau(q))q = \int_0^q \tau(\theta)\mathrm{d}\theta$. Otherwise (which may occur when $\tau(q)$ has mass), we have that

$$\begin{aligned}
R(q) &= s \int_0^\infty \mathbb{P}(X > t, X \geq \tau(q))\mathrm{d}t + (1-s) \int_0^\infty \mathbb{P}(X > t, X > \tau(q))\mathrm{d}t \\
&= s \int_0^\infty \mathbb{P}(X > t, X = \tau(q))\mathrm{d}t + \int_0^\infty \mathbb{P}(X > t, X > \tau(q))\mathrm{d}t \\
&= s \int_0^{\tau(q)} \mathbb{P}(X = \tau(q))\mathrm{d}t + \int_0^\infty \mathbb{P}(X > t, X > \tau(\bar{q}))\mathrm{d}t \\
&= (q - \bar{q})\tau(q) + R(\bar{q})
\end{aligned}$$

$$= \int_{\bar{q}}^{q} \tau(\theta)\mathrm{d}\theta + \int_{0}^{\bar{q}} \tau(\theta)\mathrm{d}\theta$$
$$= \int_{0}^{q} \tau(\theta)\mathrm{d}\theta.$$

Where the second to last inequality follows because $F^{-1}(1-\theta)$ is constant in the interval $(\bar{q}, q)$. Then,

$$\begin{aligned}
\mathbb{E}(\max\{X_1,\ldots,X_n\}) &= \int_0^\infty 1 - F^n(t)\mathrm{d}t \\
&= \int_0^1 F^{-1}(\sqrt[n]{z})\mathrm{d}z \\
&= n\int_0^1 (1-q)^{n-1}\tau(q)\mathrm{d}q \\
&= n\int_0^1 (n-1)(1-q)^{n-2}\left(\int_0^q \tau(\theta)\mathrm{d}\theta\right)\mathrm{d}q \\
&= n\int_0^1 (n-1)(1-q)^{n-2}R(q)\mathrm{d}q\,.
\end{aligned}$$

Note that the second inequality is justified by Fubini's Theorem. □

In the remaining of the chapter, to simplify notation and the understanding of the proof, whenever $\tau(q)$ has mass, we will not condition on the decision of the threshold rule. We will refer to the threshold rule as to stop every time a random variable is strictly above its corresponding threshold. In virtue of Lemma 3.1, all arguments will still apply to the case when we randomize between strict inequality and weak inequality. Adittionally, as we will see later, it is always the case that at least one of the choices of the threshold rule (whether apply strict inequality or weak inequality) reports a conditional expected value $R(q)$ larger or equal to $\int_0^q \tau(\theta)\mathrm{d}\theta$, and thus if we can determine which decision is better, no randomization in the threshold rule decision is needed.

Our proof not only assures the existence of thresholds for which a constant of β^* is achieved, but also provides an algorithm to compute the values of these threshold in terms of the probability distribution function. The intuition behind the threshold rule is to start with a high threshold and decrease it for the remaining variables, such to mitigate the risk of not stopping.

We partition the interval $A = [0,1]$ into n intervals $A_\mathrm{i} = [\varepsilon_{\mathrm{i}-1}, \varepsilon_\mathrm{i}]$, with $0 = \varepsilon_0 < \varepsilon_1 < \ldots < \varepsilon_{n-1} < \varepsilon_n = 1$. To implement this idea we use Expression (3.2) as a guide and construct the the threshold for variable i by drawing a q_i from the interval A_i according to probability density function $f_\mathrm{i}(q) = \frac{\psi(q)}{\alpha_\mathrm{i}}$, where $\psi(q) = (n-1)(1-q)^{n-2}$ and α_i is a normalization parameter equal to $\alpha_\mathrm{i} = \int_{q\in A_\mathrm{i}} \psi(q)\mathrm{d}q$. This q_i is meant to be the acceptance probability of variable X_i, so that under the assumption of F being continuous with positive density, the corresponding threshold at step i is $\tau_\mathrm{i} = \tau(q_\mathrm{i})$.

With the thresholds set, we next establish the factor guarantee. To this end, let $r := \min\{\mathrm{i} \in \{1,\ldots,n\} : X_\mathrm{i} > \tau_\mathrm{i}\}$. We prove in Lemma 3.2 that the expected value of the first random variable above its threshold, X_r, satisfies

$$\mathbb{E}(X_r) = \sum_{\mathrm{i}=1}^{n} \rho_{\mathrm{i}} \int_{\varepsilon_{\mathrm{i}-1}}^{\varepsilon_{\mathrm{i}}} (n-1)(1-q)^{n-2} R(q) q \mathrm{d}q,$$

where $\rho_1 = \frac{1}{\alpha_1}$ and $\rho_{\mathrm{i}+1} = \frac{\rho_{\mathrm{i}}}{\alpha_{\mathrm{i}+1}} \int_{\varepsilon_{\mathrm{i}-1}}^{\varepsilon_{\mathrm{i}}} \psi(q)(1-q)\mathrm{d}q$ for $\mathrm{i} = 1, \ldots, n-1$.

By choosing the values of $\varepsilon_1, \ldots, \varepsilon_{n-1}$ in such a way that $\rho_1 = \rho_2 = \ldots = \rho_n$, we have $\mathbb{E}(\max\{X_1, \ldots, X_n\}) \leq n\alpha_1 \mathbb{E}(X_r)$, and thus we wrap-up by proving that the term $n\alpha_1$ is bounded by β^*. For the latter we set up a recursion whose solution determines α_1 and then approximate the recursion with an ordinary differential equation.

Relating the maximum and the threshold rule

Lemma 3.2 *Let* $\rho_1 = \frac{1}{\alpha_1}$ *and* $\rho_{\mathrm{i}+1} = \frac{\rho_{\mathrm{i}}}{\alpha_{\mathrm{i}+1}} \int_{\varepsilon_{\mathrm{i}-1}}^{\varepsilon_{\mathrm{i}}} \psi(q)(1-q)\mathrm{d}q$ *for* $\mathrm{i} = 1, \ldots, n-1$*. Then the expected value of the value of the first random variable above its threshold,* X_r*, satisfies*

$$\mathbb{E}(X_r) = \sum_{\mathrm{i}=1}^{n} \rho_{\mathrm{i}} \int_{\varepsilon_{\mathrm{i}-1}}^{\varepsilon_{\mathrm{i}}} (n-1)(1-q)^{n-2} R(q) q \mathrm{d}q\,.$$

Proof. Let q_{i} denote the drawn acceptance probability for customer i. The expected value obtained in step i correspond to the expected value of X_{i} given it is above its threshold times the probability this occurs. This is is exactly $R(q)q$.

Now, for $j = 1, \ldots, n-1$, let q_j denote the drawn acceptance probability of customer j. Then, the probability that we actually get to step i is equal to $\prod_{j=1}^{\mathrm{i}-1} 1 - q_j$. Hence, the expected value of X_r, the first variable above its threshold is

$$\begin{aligned}
\mathbb{E}(X_r) \quad &= \sum_{\mathrm{i}=1}^{n} \mathbb{E}(R(q_{\mathrm{i}})) \prod_{j=1}^{\mathrm{i}-1} \mathbb{E}(1-q_j) \\
&= \sum_{\mathrm{i}=1}^{n} \int_{\varepsilon_{\mathrm{i}-1}}^{\varepsilon_{\mathrm{i}}} (n-1)(1-q)^{n-2} R(q) \mathrm{d}q \frac{\prod_{j=1}^{\mathrm{i}-1} \int_{\varepsilon_{j-1}}^{\varepsilon_j} \psi(q)(1-q)\mathrm{d}q}{\prod_{j=1}^{\mathrm{i}} \gamma_{\mathrm{i}}} \\
&= \sum_{\mathrm{i}=1}^{n} \rho_{\mathrm{i}} \int_{\varepsilon_{\mathrm{i}-1}}^{\varepsilon_{\mathrm{i}}} (n-1)(1-q)^{n-2} R(q) \mathrm{d}q\,.
\end{aligned}$$

□

If we denote $\bar{q}_{\mathrm{i}} = \mathbb{P}(X \geq \tau(q_{\mathrm{i}}))$ and $\underline{q}_{\mathrm{i}} = \mathbb{P}(X > \tau(q_{\mathrm{i}}))$, then $F^{-1}(1-q)$ is constant in $[\underline{q}_{\mathrm{i}}, \bar{q}_{\mathrm{i}}]$. Thus, $R(q) = \int_0^q F^{-1}(1-\theta)\mathrm{d}\theta$ is linear in that interval, implying that $E(X_r)$ is linear as a function of q_{i} when $\tau(q_{\mathrm{i}})$ has mass. This means that one of the following, the strategy that stops in step i whenever $X \geq \tau(q_{\mathrm{i}})$, or the strategy that does so when $X > \tau(q_{\mathrm{i}})$, attains a larger expected reward than $\mathbb{E}(X_r)$. Thus, if we can assert which of the two gives higher expected reward, then we can derandomize this step.

By choosing $\varepsilon_1, \ldots, \varepsilon_{n-1}$ appropriately, we can express the expected value of the thresholds rule in terms of that of the maximum.

Lemma 3.3 *If we choose $\varepsilon_1, \ldots, \varepsilon_{n-1}$ such that $\rho_1 = \rho_2 = \ldots = \rho_n$, then*

$$\mathbb{E}(\max\{X_1, \ldots, X_n\}) = n\alpha_1 \mathbb{E}(X_r)\,.$$

PROOF. If we choose $\varepsilon_1, \ldots, \varepsilon_{n-1}$ such that $\rho_{\mathrm{i}} = \rho_1 = \frac{1}{\alpha_1}$ for all i, then by Lemmatta 3.1 and 3.2 we can express the expected value of the maximum by

$$\begin{aligned}
\mathbb{E}(\max\{X_1, \ldots, X_n\}) &= n \int_0^1 (n-1)(1-q)^{n-2} R(q) q \, \mathrm{d}q \\
&= n\alpha_1\rho_1 \sum_{\mathrm{i}=1}^{n} \int_{\varepsilon_{\mathrm{i}-1}}^{\varepsilon_{\mathrm{i}}} (n-1)(1-q)^{n-2} R(q) q \, \mathrm{d}q \\
&= n\alpha_1 \sum_{\mathrm{i}=1}^{n} \rho_{\mathrm{i}} \int_{\varepsilon_{\mathrm{i}-1}}^{\varepsilon_{\mathrm{i}}} (n-1)(1-q)^{n-2} R(q) q \, \mathrm{d}q \\
&= n\alpha_1 \mathbb{E}(X_r).
\end{aligned}$$

□

Bounding through a recursion

Since $\rho_{\mathrm{i}+1} = \frac{\rho_{\mathrm{i}}}{\alpha_{\mathrm{i}+1}} \int_{\varepsilon_{\mathrm{i}-1}}^{\varepsilon_{\mathrm{i}}} \psi(q)(1-q)\mathrm{d}q$ for all i, choosing $\varepsilon_1, \ldots, \varepsilon_{n-1}$ such that all ρ_{i} are the same amounts to choosing them such that $\alpha_{\mathrm{i}+1} = \int_{\varepsilon_{\mathrm{i}-1}}^{\varepsilon_{\mathrm{i}}} \psi(q)(1-q)\mathrm{d}q$ for all i. By the definition of $\alpha_{\mathrm{i}+1}$ and $\psi(q)$, this is equivalent to choosing them such that for all i

$$\int_{\varepsilon_{\mathrm{i}-1}}^{\varepsilon_{\mathrm{i}}} \psi(q)(1-q)\, \mathrm{d}q = \frac{n-1}{n} \left((1-\varepsilon_{\mathrm{i}-1})^n - (1-\varepsilon_{\mathrm{i}})^n\right)$$

is equal to

$$\int_{\varepsilon_{\mathrm{i}}}^{\varepsilon_{\mathrm{i}+1}} \psi(q)\, \mathrm{d}q = (1-\varepsilon_{\mathrm{i}})^{n-1} - (1-\varepsilon_{\mathrm{i}+1})^{n-1}\,.$$

Now, substitute $x_{\mathrm{i}} = 1 - \varepsilon_{\mathrm{i}}$. Then, from Lemma 3.3, we obtain the following recursion on x_{i}:

$$\frac{{x_{\mathrm{i}-1}}^n}{n} - \frac{{x_{\mathrm{i}}}^n}{n} = \frac{{x_{\mathrm{i}}}^{n-1}}{n-1} - \frac{{x_{\mathrm{i}+1}}^{n-1}}{n-1}\,, \tag{3.3}$$

where $x_0 = 1$ and $x_n = 0$. Moreover,

$$\alpha_1 = \int_0^{\varepsilon_1} \psi(q)\, \mathrm{d}q = 1 - {x_1}^{n-1}\,.$$

Combining this with Lemma 3.3, we see that if $n(1 - {x_1}^{n-1}) \leq \beta$, for some $\beta \geq 1$, then the expected value of the maximum is at most the expected value of the threshold rule times β.

Note that $n(1-{x_1}^{n-1}) \leq \beta$ implies that $x_1 \geq (1-\frac{\beta}{n})^{1/(n-1)}$. Thus, if we find the minimum value for β such that $x_1 < (1-\frac{\beta}{n})^{1/(n-1)}$ implies $x_n < 0$, we know that $x_1 \geq (1-\frac{\beta}{n})^{1/(n-1)}$ for that value of β.

Hence, we proceed by showing an upper bound on the value of x_n. For this, we use the following lemma.

Lemma 3.4 *For values $x_0, x_1, \ldots, x_n$ satisfying* (3.3) *and $x_0 = 1$ and $x_n = 0$, we have for* $\mathrm{i} = 1, \ldots, n-1$,

$$x_{\mathrm{i}+1} = \left(\frac{n-1}{n}{x_{\mathrm{i}}}^n + {x_1}^{n-1} - \frac{n-1}{n}\right)^{1/(n-1)} . \tag{3.4}$$

Proof. We prove this lemma by induction. For $\mathrm{i} = 1$ equation (3.3) gives

$$x_2 = \left({x_1}^{n-1} + \frac{n-1}{n}{x_1}^n - \frac{n-1}{n}\right)^{1/(n-1)} .$$

Now, suppose the claim is true for $\mathrm{i} = 1, \ldots, j$. From (3.3), we know that

$$\begin{aligned} {x_{j+1}}^{n-1} &= {x_j}^{n-1} + \frac{n-1}{n}{x_j}^n - \frac{n-1}{n}{x_{j-1}}^n \\ &= \frac{n-1}{n}{x_{j-1}}^n + {x_1}^{n-1} - \frac{n-1}{n} + \frac{n-1}{n}{x_j}^n - \frac{n-1}{n}{x_{j-1}}^n \\ &= \frac{n-1}{n}{x_j}^n + {x_1}^{n-1} - \frac{n-1}{n} , \end{aligned}$$

where the second equality is due to the induction hypothesis. $\square$

In the following, we show that each of the terms x_{i} in the recursion can be bounded with a function $y(t) : [0,1] \to \mathbb{R}$, defined through the following differential equation. All derivatives of y are with respect to t.

$$\begin{aligned} y' &= y(\ln(y) - 1) - (\beta - 1) , \\ y(0) &= 1 . \end{aligned} \tag{ODE}$$

Furthermore, we define $y(1) = \lim_{t \uparrow 1} y(t)$ as the continuous extension of $y(t)$.

Later on, we will choose $\beta = \beta^* \approx 1.34$. For this β, we have $y \in [0,1]$, so we restrict our analysis of (ODE) to this interval. We assume $\beta > 1.25$ and $y \in [0,1]$. We validate these assumptions at the end of our analysis.

Lemma 3.5 *Differential equation* (ODE) *has a unique solution $y(t)$, which is a decreasing and strictly convex function on the interval* $[0,1]$. *Furthermore, $y'''(t) > 0$ for $y \in (0,1)$.*

Proof. Note that $y'(0) = -\beta < 0$ because $y(0) = 1$. For $y \in (0,1]$, we know $\ln(y) \leq 0$. Also, as $\beta > 1$, we conclude $y'(t) < 0$. Furthermore, $y(t)$ is convex as for $y \in [0,1)$,

$$y'' = y'(\ln(y) - 1) + y\frac{y'}{y} = y' \ln(y) > 0 ,$$

and $y'' = 0$ for $y = 1$. Finally,

$$y''' = y'' \ln(y) + y' \frac{y'}{y} = y' \ln^2(y) + \frac{(y')^2}{y} = y' \left(\ln^2(y) + \ln(y) - 1 - \frac{\beta - 1}{y} \right) .$$

We show that $\ln^2(y) + \ln(y) - 1 - \frac{\beta-1}{y} < 0$ for $y \in (0,1)$ or, equivalently, that $g(y) = y \ln^2(y) + y \ln(y) - y - \beta + 1 < 0$ for $y \in (0,1)$. To determine the maximum value of $g(y)$, observe that

$$\begin{aligned} \frac{\mathrm{d}g(y)}{\mathrm{d}y} &= \ln^2(y) + 2y \ln(y) \frac{1}{y} + \ln(y) + y \frac{1}{y} - 1 \\ &= \ln^2(y) + 3 \ln(y) = \ln(y) \left(\ln(y) + 3 \right) . \end{aligned}$$

Note that $\frac{\mathrm{d}g(y)}{\mathrm{d}y} \geq 0$ on $y \in (0, \mathrm{e}^{-3})$ and $g'(y) < 0$ on $y \in (\mathrm{e}^{-3}, 1)$. Hence, since $g(y)$ is continuous, its maximum is attained at $y = \mathrm{e}^{-3}$, and $g(\mathrm{e}^{-3}) = 5\mathrm{e}^{-3} - \beta + 1 < 0$ as $\beta > 1.25$.

Moreover, note that if $y \in (0,1)$, then $|y''|$ is bounded, and hence y' is Lipschitz continuous. Therefore, by the Picard-Lindelöf Theorem [18], $y(t)$ is unique on $(0,1)$. As $y(0)$ is given, and we defined $y(1)$ as the continuous extension of $y(t)$, the solution $y(t)$ is unique on $[0,1]$. □

We now proceed to prove that the solution of (ODE) dominates the terms of the recurrence. Before continuing, we require a technical result.

Proposition 3.6 *If $x \in (0,1]$ and $n \geq 2$, then*

$$x + \frac{x(\ln(x) - 1)}{n} + \frac{x(\ln(x) - 1) - (\beta - 1)}{2n^2} \ln(x) \geq \frac{n-1}{n} x^{\frac{n}{n-1}} .$$

Proof. Fix a value for n. Since $-(\beta - 1)\ln(x)$ is positive, and since $x > 0$, it suffices to prove that

$$f(x) := 1 + \frac{\ln(x) - 1}{n} + \frac{\ln(x)(\ln(x) - 1)}{2n^2} - \frac{n-1}{n} x^{\frac{1}{n-1}} \geq 0 .$$

As $f(1) = 0$ for all n, showing that f is nonincreasing completes the proof. All derivatives are with respect to x. We see that

$$f'(x) = \frac{1}{nx} + \frac{\ln(x) - 1}{2n^2 x} + \frac{\ln(x)}{2n^2 x} - \frac{1}{n} x^{\frac{1}{n-1} - 1} = \frac{1}{nx} \left(1 - x^{\frac{1}{n-1}} + \frac{1}{2n} (2 \ln(x) - 1) \right) .$$

For $x \in (0,1]$, f' has the same sign as $g(x) := 1 - x^{\frac{1}{n-1}} + \frac{1}{2n}(2\ln(x) - 1)$. We prove that g has a maximum $x^* \in (0,1]$ with $g(x^*) \leq 0$. This implies that both g and f' are nonpositive. Indeed,

$$\begin{aligned} g'(x) &= \frac{1}{nx} - \frac{x^{\frac{1}{n-1} - 1}}{n-1} , \\ g''(x) &= -\frac{1}{nx^2} + \frac{n-2}{(n-1)^2} x^{\frac{1}{n-1} - 2} . \end{aligned}$$

Observe that $g'(x^*) = 0$ only when $x^* = (\frac{n-1}{n})^{n-1}$. Furthermore, g'' has the same sign as $h(x) := -\frac{1}{n} + \frac{n-2}{(n-1)^2} x^{\frac{1}{n-1}}$, which is an increasing function in x for all $n \geq 2$. As $h(1) = -\frac{1}{n(n-1)^2} < 0$, g'' is negative, g is concave and attains its maximum at x^*. Finally,

$$\begin{aligned} g(x^*) &= 1 - \frac{n-1}{n} + \frac{n-1}{n} \ln\left(\frac{n-1}{n}\right) - \frac{1}{2n} \\ &= \frac{1}{2n} + \frac{n-1}{n} \ln\left(1 - \frac{1}{n}\right) \leq \frac{1}{2n} + \frac{n-1}{n}\left(-\frac{1}{n}\right) \\ &= \frac{1}{n^2} - \frac{1}{2n} = \frac{1}{n}\left(\frac{1}{n} - \frac{1}{2}\right) \leq 0\,, \end{aligned}$$

where the last inequality follows from $n \geq 2$. This concludes the proof. □

Using this proposition, we bound the recurrence by the function $y(t)$ in the following way.

Lemma 3.7 *If $x_1 < (1 - \frac{\beta}{n})^{\frac{1}{n-1}}$, then ${x_{\mathrm{i}}}^{n-1} < y(\frac{\mathrm{i}}{n})$ for $\mathrm{i} = 1, \ldots, n$, where $y(t)$ is the unique solution of* (ODE).

PROOF. First note that $x_0 = y(0) = 1$, by definition. Moreover, we already saw that $y'(0) = -\beta$. As $y(t)$ is strictly convex, we know that $y(1/n) > y(0) - \frac{1}{n}\beta > {x_1}^{n-1}$, where the last inequality follows by assumption. Now assume that ${x_{\mathrm{i}}}^{n-1} < y(\frac{\mathrm{i}}{n})$, then we prove ${x_{\mathrm{i}+1}}^{n-1} < y(\frac{\mathrm{i}+1}{n})$. First observe that the Taylor expansion of $y(\frac{\mathrm{i}+1}{n})$ around $\frac{\mathrm{i}}{n}$ is

$$y(\tfrac{\mathrm{i}+1}{n}) = y(\tfrac{\mathrm{i}}{n}) + \frac{1}{n} y'(\tfrac{\mathrm{i}}{n}) + \frac{1}{2n^2} y''(\tfrac{\mathrm{i}}{n}) + \frac{1}{6n^6} y'''(\zeta)\,,$$

with $\zeta \in [\frac{\mathrm{i}}{n}, \frac{\mathrm{i}+1}{n}]$. As $y''' > 0$, it follows that

$$\begin{aligned} y(\tfrac{\mathrm{i}+1}{n}) &> y(\tfrac{\mathrm{i}}{n}) + \frac{1}{n} y'(\tfrac{\mathrm{i}}{n}) + \frac{1}{2n^2} y''(\tfrac{\mathrm{i}}{n}) \\ &= y(\tfrac{\mathrm{i}}{n}) + \frac{y(\frac{\mathrm{i}}{n})(\ln(y(\frac{\mathrm{i}}{n})) - 1) - (\beta - 1)}{n} + \frac{y(\frac{\mathrm{i}}{n})(\ln(y(\frac{\mathrm{i}}{n}) - 1) - (\beta - 1))}{2n^2} \ln(y(\tfrac{\mathrm{i}}{n})) \\ &\geq \frac{n-1}{n} y(\tfrac{\mathrm{i}}{n})^{\frac{n}{n-1}} - \frac{\beta - 1}{n} \\ &> \frac{n-1}{n} x_{\mathrm{i}}^n - \frac{\beta - 1}{n} \\ &> x_{\mathrm{i}+1}^{n-1}\,, \end{aligned}$$

where the last inequality follows from Lemma 3.4 and the assumption that $x_1^{n-1} < 1 - \frac{\beta}{n}$. □

3.3 Main Theorem

We are now ready to prove the main theorem of this chapter.

Theorem 3.8 *For any $n > 1$ and non-negative i.i.d. random variables $X_1, \ldots, X_n$, there exist thresholds $\tau_1, \ldots, \tau_n$, such that*

$$\mathbb{E}(\max\{X_1, \ldots, X_n\}) \leq \beta^* \mathbb{E}(X_r)$$

where β^* is the unique solution to (3.1) and $r := \min\{\mathrm{i} \in \{1, \ldots, n\} \mid X_\mathrm{i} > \tau_\mathrm{i}\}$. If the latter set is empty, we take $X_r = 0$. If some of the distribution have mass in their corresponding thresholds, the definition of r must change accordingly.

Proof. Choosing $0 = \varepsilon_0 < \varepsilon_1 < \ldots < \varepsilon_{n-1} < \varepsilon_n = 1$ such that $\int_{\varepsilon_{\mathrm{i}-1}}^{\varepsilon_\mathrm{i}} \psi(q)(1-q)\,\mathrm{d}q = \alpha_{\mathrm{i}+1}$, we know by Lemma 3.3 that

$$\mathbb{E}(\max\{X_1, \ldots, X_n\}) \leq n\alpha_1 \mathbb{E}(X_r)\,,$$

where $\alpha_1 = 1 - (1-\varepsilon_1)^{n-1} = 1 - x_1^{n-1}$. Hence, we want to show $n(1 - x_1^{n-1}) \leq \beta^*$ for $\beta^* \approx 1.3415$.

We prove by contradiction and assume $x_1^{n-1} < 1 - \frac{\beta}{n}$. Then Lemma 3.7 yields $x_n < y(1)$, so we choose β such that $y(1) = 0$ to reach a contradiction with the fact that $x_n = 0$. Note that this indeed implies $y \in [0,1]$ as we assumed. Hereto, note that $y(t)$ is invertible by Lemma 3.5. Hence, we can consider t as a function of y, for which we know $t(1) = 0$, and we want to choose β such that $t(0) = 1$. In virtue of the inverse function theorem, we have that

$$\begin{aligned} t(1) = t(0) + \int_0^1 \frac{\mathrm{d}t}{\mathrm{d}y}\mathrm{d}y &= 1 + \int_0^1 \frac{1}{\frac{\mathrm{d}y}{\mathrm{d}t}}\mathrm{d}y \\ &= 1 + \int_0^1 \frac{1}{y(\ln(y) - 1) - (\beta - 1)}\mathrm{d}y\,. \end{aligned}$$

So β^* is the value such that the last integral equals -1. This yields $\beta^* \approx 1.3415$. □

We finish this section by proving that the sequence a_n defined in [13] equals exactly ours $n\alpha_1$. Our recurrence was given by $x_0 = 1$, $x_n = 0$ satisfied the relation given in Lemma 3.4 for $\mathrm{i} = 1, \ldots, n$:

$$x_{\mathrm{i}+1} = \left(\frac{n-1}{n}{x_\mathrm{i}}^n + {x_1}^{n-1} - \frac{n-1}{n}\right)^{1/(n-1)}. \tag{3.5}$$

Set $\gamma_n = n - 1 - nx_1^{n-1}$. Now, equation (3.5) can be rewritten as

$$x_\mathrm{i}^n = \frac{n}{n-1}x_{\mathrm{i}+1}^{n-1} + \frac{\gamma_n}{n-1}\,. \tag{3.6}$$

Consider the variables $z_\mathrm{i} = x_{n-\mathrm{i}}^n$. These new variables satisfy $z_0 = 0$, $z_n = 1$ and the recurrence

$$z_{\mathrm{i}+1} = \frac{n}{n-1}z_\mathrm{i}^{\frac{n-1}{n}} + \frac{\gamma_n}{n-1}. \tag{3.7}$$

We recover the Hill-Kertz recursion by noting that $z_1 = \frac{\gamma_n}{n-1}$. Then, our approximation factor $n(x_1^{n-1} - 1)$ is equal to theirs $1 + \alpha_n$. Note here that our α_1 does depend on n, though we have omited this dependency for simplicity of notation. Thus our result implies that $a_n \leq \beta^*$ and and by the work of Hill and Kertz [13] and Kertz [14] we know that this bound is tight.

Algorithm

As in the previous chapter, the proof of Theorem 3.8 provides a construction to compute the thresholds as a function of the probability distribution function of the random variables. When the distribution is continuous the algorithm takes the following form:

Algorithm 3.1: Adaptive i.i.d. thresholds rule

Input: Distribution F, values $X_1, \dots, X_n$ draw from F.

Initialize $r = 0$, $X_0 = 0$

partition $[0, 1]$ in n disjoint intervals $A_i = [a_{i-1}, a_i]$, $i = 1, \dots, n$ where $x_i = 1 - a_i$ satisfy (3.3).

for $i = 1$ **to** n **do**

    draw q_i from A_i with density proportional to $(n-1)(1-q)^{n-2}$

    set threshold $\tau_i = F^{-1}(1 - q_i)$

    if $X_i > \tau_i$ **then**

        $r \leftarrow i$

        break

    end if

end for

return X_r

In the general case, it suffices to change the condition $X_i > \tau_i$ by the randomized version of the threshold rule, where with some probability the condition changes to $X_i \geq \tau_i$. This again can be derandomized using the same technique as in Algorithm 2.1.

Chapter 4

Applications to Mechanism Design

In this chapter we discuss applications of the previous advances in optimal stopping theory in mechanisms design. In particular, we will use these tools to produce Bayesian near-optimal solutions to mechanism design problems in the form of approximations algorithms to the optimal solution. Broadly speaking, we analyze the problem faced by a seller who wishes to sell an item to potential customers and whose objective is to maximize her revenue. We also point a direct application in social welfare maximization.

4.1 Preliminaries

The problem we consider is in general labeled as a problem of pricing. A monopolist sells a single item to a set of n buyers. The seller places no value on the item, while the buyers are assumed to have independent random valuations $V_1, \dots, V_n$ for the item, following continuous and increasing distributions $F_1 \dots, F_n$ (in particular, distributions have positive density over their support). The main question is to design a mechanism maximizing the revenue of the seller. In other words, the problem of pricing is to find an allocation rule and prices to charge for all possible values customers declare to have over the item. This question was answered in a seminal paper by R. Myerson in 1981 [19].

For each customer $\mathrm{i} \in [n]$, define $\phi_\mathrm{i}(v) = v - \frac{1-F_\mathrm{i}(v)}{F_\mathrm{i}'(v)}$ as the *virtual value function*. We will say that F_i is a *regular* distribution if ϕ_i is a non-decreasing function. Otherwise, we will say it is *non-regular*. In this case, we can transform ϕ_i into a non-decreasing function with a process called *ironing*. Define G_i as the convex hull of the function $H_\mathrm{i}(q) = -(1-q)F_\mathrm{i}^{-1}(\theta)$:

$$G(q) = \min \{xH(q_1) + (1-x)H(q_2) : xq_1 + (1-x)q_2 = q \text{ and } x, q_1, q_2 \in [0,1]\}.$$

The function $\bar{\phi}_\mathrm{i}(v) = G'(F(v))$ is called the ironed virtual value and is non-decreasing (note that when ϕ_i is non-decreasing, both functions coincide).

Let $\bar{v}_\mathrm{i}$ be the reserve price of customer i, i.e., a value for which $\bar{\phi}_\mathrm{i}(\bar{v}_\mathrm{i}) = 0$. Myerson's optimal auction allocates the item to the customer with highest (ironed) virtual value, among the customers with non-negative virtual value. If the valuation of every customer is below their reserve price, then the seller keeps the item. Then, it can be proved that the expected revenue of the optimal auction is given by $OPT = \mathbb{E}(\max\{\bar{\phi}_1(V_1), \dots, \bar{\phi}_n(V_n), 0\})$. If we define $X_i = \bar{\phi}_i(V_i)^+$ as the positive part of the virtual valuation of customer i, then

$$OPT = \mathbb{E}(\max\{X_1, \dots, X_n\}).$$

This solution to the sellers problem is, in some situations, a remarkably simple mechanism. However, in many situations it is hard to implement, and the mechanism of choice turns out to be a simple *posted price mechanism*. Posted price mechanisms constitute an attractive and widely applicable way of selling items to strategic consumers. In this context, consumers are faced with take-it-or-leave-it offers, and therefore strategic behaviour simply vanishes. The setting then is that of a *sequential auction*: there is an order over the set of customers, and the offers are made according to that order. This type of mechanism has been vastly studied, particularly in the marketing community [5]. In recent years, there has been a significant effort to understand the expected revenue of the outcome generated by different posted price mechanisms when compared to that of the optimal auction [2, 4, 6, 21].

The objective of this chapter is to apply the techniques developed in Chapters 2 and 3 to investigate the performance of posted price mechanisms to sell a single item to a given set of customers who arrive in a random unknown order, with respect to the optimal auction. We consider two different models which share the property that each customer is offered the item at most once. Upon receiving an offer, a customer immediately decides whether to buy the item at that price or to pass and simply not buy.

The *non-adaptive* model considers the situation in which all offers have to be made simultaneously, and customers respond in random order, akin to direct mail campaigns, in which the seller contacts its potential buyers directly and offers each one a certain price for the item. The item is then sold to the first consumer who accepts the offer [5, 7]. The *adaptive* model considers a situation in which the seller may adapt the offer. Here, customers again arrive in random order. Whenever a customer arrives, she is offered the item at a price, which the seller may base on the customer he is offering to, as well as the customers who already rejected earlier offers.

Problem of Welfare

Before discussing the problem of pricing, we briefly explain an application to welfare maximization. Suppose there is a good and n agents with random valuation for the good. A central planner wishes to allocate the good in order to maximize the expected utilitarian social welfare.

If the valuations of the agents are given by non-negative random variables $Z_1, \dots, Z_n$, then the expected maximum welfare is exactly the quantity $\mathbb{E}(\max\{Z_1, \dots, Z_n\}))$, which is achieved when the good is allocated to the agent who values it the most.

Consider a sequential setting where the values of agents are learned in order, and the central planner must decide to allocate the good in the moment it learns the valuation of the agent. It is easy to observe that this setting is equivalent to the ones already discussed in the previous chapters. This means that guarantees found for the non-adaptive and the adaptive i.i.d. thresholds rule — and the algorithms to achieve them — also apply for the central planner trying to approximate the maximum social welfare.

4.2 Reduction to Pricing

The main tool of this chapter will be to derive a general principle to treat a problem of posted price as a prophet inequality in a wide spectrum of settings. One of the first works in this directions were developed by [11, 3], and particularly well known are the results by [6]. We provide alternative and novel profs to achieve this reduction, which may shed light about the applicability of this construction. We recall that all distributions considered in this chapter are continuous and increasing.

Lemma 4.1 *For a non-negative random variable V with regular distribution F, let X be the positive part of its virtual valuation. Then, for any $\tau \geq 0$,*

$$\mathbb{E}(X \mid X > \tau) = F^{-1}(1-q), \text{ with } q = \mathbb{P}(X > \tau).$$

If the distribution is non-regular, there exist $q_1, q_2, x \in [0,1]$ such that $xq_1 + (1-x)q_2 = q$ and

$$\mathbb{E}(X \mid X > \tau) = \frac{xq_1 F^{-1}(1-q_1) + (1-x)q_2 F^{-1}(1-q_2)}{q}.$$

Proof. Let $\bar{c}$ be the ironed virtual value function. Recall that $\bar{c}(v) = G'(F(v))$, where G is the convex hull of the function $H(\theta) = -(1-\theta)F^{-1}(\theta)$. Thus,

$$\mathbb{E}(X \mid X > \tau)\mathbb{P}(X > \tau) = \int_\tau^\infty \bar{c}(u)\mathrm{d}F(u) = \int_0^q G'(1-\theta)\mathrm{d}\theta = -G(1-q).$$

Additionally, when the distribution is regular, $G = H$ and the result follows. □

Since the events $X_i > \tau_i$ and $V_i \geq F_i^{-1}(1-q_i)$ are equivalent, it is a direct corollary that if the virtual value function ϕ is increasing, then $\mathbb{E}(X \mid X > \tau) = \varphi^{-1}(\tau)$.

Using Lemma 4.1 we can make a reduction from any threshold rule in a prophet inequality setting to a posted price guarantee. Consider the traditional setting where a set of random variables with known distributions are presented in order according some permutation. Suppose there is a threshold rule that when choosing the first variable above the corresponding threshold, we get an expected value of at least α times the expected value of the maximum. Then, we can get the same guarantee for the expected revenue for a posted price mechanism with the same permutation, in relation to the value of the optimal auction.

To see how, let $V_1, \ldots, V_n$ be the random valuation that the n customers have over the item. If a posted price mechanism offers prices $v_1, \ldots, v_n$, the expected revenue we get is

$$\sum_{\mathrm{i}=1}^{n} v_{\mathrm{i}} \mathbb{P}(V_{\mathrm{i}} \geq v_{\mathrm{i}}, \text{i is the first above price}).$$

Let $X_1, \ldots, X_n$ be their respective positive virtual values and consider non-negative thresholds $\tau_1, \ldots, \tau_n$ for every one of them[1]. If r denotes the first random variable whose virtual value is above its corresponding threshold, the expected virtual value we get in the regular case can be expressed as

$$\begin{aligned}
\mathbb{E}(X_r) &= \sum_{\mathrm{i}=1}^{n} \mathbb{E}(X_{\mathrm{i}} \mid \mathrm{i} = r)\mathbb{P}(\mathrm{i} = r) \\
&= \sum_{\mathrm{i}=1}^{n} \mathbb{E}(X_{\mathrm{i}} \mid X_{\mathrm{i}} > \tau_{\mathrm{i}})\mathbb{P}(X_{\mathrm{i}} > \tau_{\mathrm{i}}, \text{i is the first above threshold}) \\
&= \sum_{\mathrm{i}=1}^{n} F_{\mathrm{i}}^{-1}(1 - q_{\mathrm{i}})\mathbb{P}(V_{\mathrm{i}} \geq F_{\mathrm{i}}^{-1}(1 - q_{\mathrm{i}}), \text{i is the first above threshold})
\end{aligned}$$

where $q_{\mathrm{i}} = \mathbb{P}(X_{\mathrm{i}} > \tau_{\mathrm{i}})$. Then, the expected value of the threshold rule over the virtual values equal the expected value of the revenue of a posted price mechanism over the valuations, under a proper selection of prices. In case some distribution is regular, then we must randomized between two prices when making the offer to that customer.

Since X_{i} is positive for all $\mathrm{i} = 1, \ldots, n$, and $OPT = \mathbb{E}(\max\{X_1, \ldots, X_n\})$, choosing thresholds for this random variables from a strategy for the secretary problem that accomplish an α approximation of the maximum, and setting the proper posted prices, gets a revenue of at least α times the optimal auction. To sum up, what we get is

$$OPT = \mathbb{E}(\max\{X_1, \ldots, X_n\}) \leq \alpha\mathbb{E}(X_r) = \alpha\mathbb{E}(F_r^{-1}(1 - q_r)).$$

Remark 4.1 Since any threshold derived from a threshold rule over the non-negative virtual valuations will be non-negative, all prices induced by these thresholds will be larger or equal than the reserve price for every one of the customers valuations.

We can thus conclude the following result:

Theorem 4.2 *Consider a sequential auction of one item where customers have private valuations with known distributions. Suppose there exists a threshold rule for non-negative random variables that are presented following the same sequence as the customers in the auction, achieving an expected value of α times the expected maximum. Then there exists a posted price mechanism for such auction that achieves a revenue of α timesthe optimal auction.*

We remark that this approach has already been taken by many authors (e.g., [6, 9, 15]), however, to our knowledge, this derivation of the result is new.

[1] In general, thresholds do not necessary depend (only) on the random variable they are facing, but the notation is without loss of generality since in the proof we are only interested in the value of the threshold each random variable faces and not the process by which this value is obtained.

4.3 Non-adaptive Posted Price Mechanism

Formally stated, the problem we consider is the following. A seller has a single item to sell to a given set of n customers. We assume that the seller has no value for keeping the item. Customers have independent random valuations for the item with customer i valuing the item at V_{i}, drawn from distribution F_{i}. The customers arrive in random order, and the goal of the seller is to maximize his expected revenue, under the restriction that prices v_{i} must be set beforehand. The seller sets prices $v_{\mathrm{i}} \geq 0$ for all $\mathrm{i} \in [n]$, with the goal of maximizing his expected revenue.

The next result is a corollary of of Theorem 2.6 and Theorem 4.2.

Theorem 4.3 *For any given set of potential customers whose valuations are independet, there exists a non-adaptive posted price mechanism that achieves an expected revenue of at least a* $1 - 1/\mathrm{e}$ *fraction of that of Myerson's optimal auction.*

Algorithm

Note that the non-adaptive algorithm of Chapter 2 (Algorithm 2.1) computes the probabilities of each random variable being the largest and uses these values to set the thresholds. Since the optimal auction always allocate the item to the highest virtual value customer above its reserve price, and positive virtual thresholds reduce in prices above the reserve price, this is equivalent to compute the probability of each customer winning the optimal auction. When distributions are regular, the algorithm takes then the following implementation:

Algorithm 4.1: Non-adaptive posted price mechanism

Input: Distributions $F_1, \ldots, F_n$, valuations $V_1, \ldots, V_n$ draw from the corresponding distribution.

Initialize $R = \emptyset$, $j = 0$, $v_0 = 0$
for $\mathrm{i} = 1$ **to** n **do**
  Compute q_{i} = probability optimal auction assigns to customer i
  Set price $v_{\mathrm{i}} = \begin{cases} F^{-1}(1 - q_{\mathrm{i}}) & \text{w.p. } \frac{2}{2+(\mathrm{e}-2)q_{\mathrm{i}}} \\ \infty & \text{otherwise.} \end{cases}$
  if $V_{\mathrm{i}} \geq v_{\mathrm{i}}$ **then**
    $R \leftarrow R \cup \{\mathrm{i}\}$
  end if
end for
if $R \neq \emptyset$ **then**
  Select j uniformly at random from R
end if
return v_j

For non-regular distributions, price v_{i} must be set randomly between two different prices (although this can be derandomized choosing the one reporting larger revenue).

Alternative Proof

For completeness, we will describe how to derive the same posted price policy using only the Bernoulli Selection Lemma (Lemma 2.1) and a technical result by [6], for which we also provide an alternative simple proof.

Lemma 4.4 ([6]) *If all value distributions are regular, then the expected value of Myerson's optimal auction is bounded from above by*

$$\sum_{i\in[n]} F_i^{-1}(1-q_i)q_i\,,$$

where q_i is the probability that the optimal auction assigns the item to customer i.

For every i *(with regular or non-regular distribution) there exist two prices $\underline{p_i}$ and $\overline{p_i}$, with corresponding probabilities $\underline{q_i}$ and $\overline{q_i}$, and a number $0\le x_i\le 1$, such that $x_i\underline{q_i}+(1-x_i)\overline{q_i}=q_i$, and the expected revenue of Myerson's optimal auction is bounded from above by*

$$\sum_{i\in[n]} x_i\underline{p_i}\,\underline{q_i}+(1-x_i)\overline{p_i}\,\overline{q_i}\,.$$

Proof. Using Lemma 2.7 we can express the revenue of the optimal auction as

$$OPT=\mathbb{E}(\max\{X_1,\ldots,X_n\})=\sum_{i=1}^{n}\mathbb{E}(X_i\mid X_i=\max_{j\in[n]}X_j)q_i\le\sum_{i=1}^{n}\mathbb{E}(X_i\mid X_i>\alpha_i)q_i$$

with X_i the positive part of the virtual valuation and α_i a value for which $\mathbb{P}(X_i>\alpha_i)=q_i$. We recall Lemma 4.1 to conclude. □

Proof of Theorem 4.3. We prove the regular case first. Let q_i denote the probability with which Myerson's optimal auction assigns the item to customer i, and set $b_i=F_i^{-1}(1-q_i)$. The expected revenue of a non-adaptive posted price mechanism, that chooses to sell only to customers in $S\subseteq[n]$ while offering prices b_i, is given by

$$\begin{aligned}\sum_{i\in S}b_i\mathbb{P}[i=\operatorname*{argmin}_{j\in S}\{\sigma(j)\,|\,v_j\ge b_j\}]&=\sum_{i\in S}b_iq_i\mathbb{P}[i=\operatorname*{argmin}_{j\in S}\{\sigma(j)\,|\,v_j\ge b_j\}\mid v_i\ge b_i]\\&=\sum_{i\in S}b_iq_i\sum_{R\subseteq S\setminus\{i\}}\frac{1}{1+|R|}\prod_{j\in R}q_j\prod_{j\in S\setminus(R\cup\{i\})}(1-q_j)\\&=\sum_{i\in S}b_iq_i\mathbb{E}\left[\tfrac{1}{1+\sum_{j\in S\setminus\{i\}}X_j}\right]\\&=\mathbb{E}\left[\tfrac{\sum_{i\in S}b_iX_i}{\sum_{i\in S}X_i}\right]\,,\end{aligned}$$

where $\{X_i\}_{i\in[n]}$ are Bernoulli random variables with $X_i=1$ with probability q_i, and σ is a uniformly random permutation. By the Bernoulli Selection Lemma we can choose the set $S\subseteq[n]$ to be such that the latter is lower bounded by

$$\left(1-\frac{1}{e}\right)\max_{z_i\le q_i}\left\{\sum_{i\in[n]}b_iz_i\;\middle|\;\sum z_i\le 1\right\}\ge\left(1-\frac{1}{e}\right)\sum_{i\in[n]}F_i^{-1}(1-q_i)q_i\,.$$

Therefore, Lemma 4.4 leads to the desired conclusion.

In the non-regular case, the posted price mechanism runs a lottery between two prices to get the desired bound[2]. First, for every bidder with positive probability of winning the optimal auction, set

$$b'_{\mathrm{i}} = \frac{x_{\mathrm{i}}\underline{p_{\mathrm{i}}}\,\underline{q_{\mathrm{i}}} + (1-x_{\mathrm{i}})\overline{p_{\mathrm{i}}}\,\overline{q_{\mathrm{i}}}}{q_{\mathrm{i}}}\,,$$

where the variables are defined as in the lemma. Also consider the same Bernoulli random variables presented in the first part of the proof. The non-adaptive posted price mechanism sells only to a set S' of customers (to be defined). For every $\mathrm{i} \in S'$, it offers a random price p_{i} equal to $\underline{p_{\mathrm{i}}}$ with probability x_{i}, and $\overline{p}_{\mathrm{i}}$ otherwise. This way, the a priori probability that v_{i} is above the price offered is exactly $x_{\mathrm{i}}\underline{q_{\mathrm{i}}} + (1-x_{\mathrm{i}})\overline{q_{\mathrm{i}}} = q_{\mathrm{i}}$, while the expected revenue of the mechanism can be evaluated as

$$\begin{aligned}
&\sum_{\mathrm{i}\in S'} x_{\mathrm{i}}\underline{p_{\mathrm{i}}}\,\underline{q_{\mathrm{i}}}\mathbb{P}[\mathrm{i} = \operatorname*{argmin}_{j\in S'}\{\sigma(j) \mid v_j \geq p_j\} \mid v_{\mathrm{i}} \geq p_{\mathrm{i}},\ p_{\mathrm{i}} = \underline{p_{\mathrm{i}}}] \\
&\quad + (1-x_{\mathrm{i}})\overline{p_{\mathrm{i}}}\,\overline{q_{\mathrm{i}}}\mathbb{P}[\mathrm{i} = \operatorname*{argmin}_{j\in S'}\{\sigma(j) \mid v_j \geq p_j\} \mid v_{\mathrm{i}} \geq p_{\mathrm{i}},\ p_{\mathrm{i}} = \overline{p_{\mathrm{i}}}] \\
&\quad = \sum_{\mathrm{i}\in S'}(x_{\mathrm{i}}\underline{p_{\mathrm{i}}}\,\underline{q_{\mathrm{i}}} + (1-x_{\mathrm{i}})\overline{p_{\mathrm{i}}}\,\overline{q_{\mathrm{i}}}) \sum_{R\subseteq S'\setminus\{\mathrm{i}\}} \frac{1}{1+|R|}\prod_{j\in R} q_j \prod_{j\in S'\setminus(R\cup\{\mathrm{i}\})}(1-q_j) \\
&\quad = \sum_{\mathrm{i}\in S'} b'_{\mathrm{i}}q_{\mathrm{i}}\mathbb{E}\left[\frac{1}{1+\sum_{j\in S'\setminus\{\mathrm{i}\}} X_j}\right] \\
&\quad = \mathbb{E}\left[\frac{\sum_{\mathrm{i}\in S'} b'_{\mathrm{i}}X_{\mathrm{i}}}{\sum_{\mathrm{i}\in S'} X_{\mathrm{i}}}\right]\,.
\end{aligned}$$

As before, Lemma 2.1 implies that there exists $S' \subseteq [n]$ such that the latter is lower bounded by $(1-1/\mathrm{e})\sum_{\mathrm{i}\in[n]} b'_{\mathrm{i}}q_{\mathrm{i}}$. Lemma 4.4 implies the bound over the optimal auction. □

Tight instance with i.i.d. valuations.

We construct a family of instances for the non-adaptive pricing problem with i.i.d. customer valuations, such that, for all $\varepsilon > 0$, there is an instance from this family for which no non-adaptive strategy can achieve an expected revenue within a factor $(1+\varepsilon)(1-1/\mathrm{e})$ of the optimal expected revenue. but here we achieve this with i.i.d. valuations. Consider n^2 customers whose values are independent identically distributed such that with probability $\frac{1}{n^3}$, the value is $\frac{n}{\mathrm{e}-2}$, with probability $\frac{1}{n}$ is 1, and 0 otherwise.

Consider an auction that offers the item for price $n/(\mathrm{e}-2) - c$ (with c a small value, say $c = 2$) to any bid above that price (and assigns the item at random if more than one such offer is received), and if no such bid is received, then it runs a lottery at price 1 among all the bids above that price. As there are many buyers of value 1, a potential large value customer will prefer to make a revenue of c rather than risking to lose the item in the lottery.

[2] This lottery can be derandomized using standard techniques, since each combination of prices offered to the customers is a deterministic mechanism in itself and the random mechanism is simply a lottery over, and thus a convex combination of, those deterministic mechanisms.

Therefore the revenue the auction will generate will approach $1/(\mathrm{e}-2)+1$ as $n \to \infty$. Of course, the revenue of the optimal auction is then at least this quantity. On the other hand, the best posted price mechanism offers a price of 1 to, say, customers $1, \ldots, k$ and $n/(\mathrm{e}-2)$ to the rest of the customers. The resulting revenue can be computed as a recursion on the expected revenue of the remaining buyers.

Let $V(j)$ be the expected revenue when there are j customers left, then the total expected revenue is

$$\begin{aligned} V(n^2) &= (1-p)\frac{1}{n^3}\frac{n}{\mathrm{e}-2} + p\left(\frac{1}{n}+\frac{1}{n^3}\right) + \left(1-\frac{1}{n^3}-\frac{p}{n}\right)V(n^2-1) \\ &= \left((1-p)\frac{1}{n^2}\frac{1}{\mathrm{e}-2} + p\left(\frac{1}{n}+\frac{1}{n^3}\right)\right)\sum_{\mathrm{i}=0}^{n^2-1}\left(1-\frac{1}{n^3}-\frac{p}{n}\right)^{\mathrm{i}} \end{aligned}$$

Let $p = \frac{\beta}{n}$ for some $\beta \in [0, n]$. Then, the expected revenue can be bounded by

$$\begin{aligned} V(n^2) &= \left(\frac{1-\frac{\beta}{n}}{\mathrm{e}-2} + \beta\left(1+\frac{1}{n^2}\right)\right)\frac{1-\left(1-\frac{1}{n^3}-\frac{\beta}{n^2}\right)^{n^2}}{\frac{1}{n}+\beta} \\ &\leq \left(\frac{1}{\mathrm{e}-2} + \beta\left(1+\frac{1}{n^2}\right)\right)\frac{1-\left(1-\frac{1}{n^3}-\frac{\beta}{n^2}\right)^{n^2}}{\frac{1}{n}+\beta} = \hat{V}(n^2)\,. \end{aligned}$$

Note that β may be a function of n, as long as $\beta \in [0, n]$. We set $\hat{\beta} = \lim_{n\to\infty}\beta$ and have

$$\lim_{n\to\infty}\hat{V}(n^2) = \left(\frac{1}{\hat{\beta}(\mathrm{e}-2)}+1\right)\left(1-\mathrm{e}^{-\hat{\beta}}\right).$$

Hence, we know that for any $\varepsilon > 0$, for large enough n we can bound $V(n^2)$ by

$$V(n^2) \leq \left(\frac{1}{\hat{\beta}(\mathrm{e}-2)}+1\right)\left(1-\mathrm{e}^{-\hat{\beta}}\right)+\varepsilon\,.$$

By the proof of tightness in Chapter 2, we know that this function is maximized at $\hat{\beta} = 1$, yielding a value of

$$V(n^2) \leq \left(\frac{1}{\mathrm{e}-2}+1\right)(1-\mathrm{e}^{-1})+\varepsilon = (1-\mathrm{e}^{-1})\mathrm{OPT}+\varepsilon'\,.$$

4.4 Adaptive I.I.D. Posted Price Mechanism

In the previous section we considered the setting in which the posted price only depends on the customer, not on the order. In this section we consider the setting in which the posted price may depend both on the customer and on the customers that arrived before her. We assume that the valuations of the customers are i.i.d. with distribution F. Again, a direct corollary of Theorems 3.8 and 4.2 is the following:

Theorem 4.5 *For any given set of potential customers whose valuations are independent and identically distributed, there exists a non-adaptive posted price mechanism that achieves an expected revenue of at least a* $1/\beta^* \approx 0.745$ *fraction of that of Myerson's optimal auction, where* β^* *is the unique solution to* (3.1).

In opposition to the non-adaptive case, to derive a simple algorithm starting from the previous characterization is not that straightforward. We will in turn analyze in more detail how to construct such algorithm. We begin by characterizing the expected optimal revenue.

The expected profit of the optimal auction equals its expected virtual surplus (see, e.g., [12]), i.e., the largest non-negative virtual value . Note that $\bar{\phi}$ is an increasing function, and let v^* be the reserve price: the value at which $\bar{\phi}(v^*) = 0$, or zero, if no such value exists. Then, the latter can be evaluated as:

$$\begin{aligned} OPT &= \int_{v^*}^{\infty} nF(v)^{n-1}\bar{\phi}(v)\mathrm{d}F(v) \\ &= n\int_0^{\alpha^*} (1-q)^{n-1}G'(1-q)\mathrm{d}q \\ &= nG(1) - nG(F(v^*))F(v^*)^{n-1} - n(n-1)\int_0^{\alpha^*} (1-q)^{n-2}G(1-q)\mathrm{d}q\,. \end{aligned}$$

Where $\alpha^* = 1 - F(v^*)$. Since $\bar{\phi}(v^*) = 0$, we know that G attains a minimum at $F(v^*)$ and, therefore, equals $H(F(v^*))$ at that point. Therefore, we can conclude that

$$\begin{aligned} OPT &= -nH(F(v^*))F(v^*)^{n-1} - n(n-1)\int_0^{\alpha^*} (1-q)^{n-2}G(1-q)\mathrm{d}q \\ &= nv^*(1-F(v^*))F(v^*)^{n-1} - n(n-1)\int_0^{\alpha^*} (1-q)^{n-2}G(1-q)\mathrm{d}q \\ &= n(n-1)\int_0^1 (1-q)^{n-2}\bar{G}(1-q)\mathrm{d}q, \end{aligned}$$

with

$$\bar{G}(1-q) = \begin{cases} -G(1-q) & \text{if } 1-q > F(v^*)\,, \\ v^*(1-F(v^*)) & \text{otherwise}\,. \end{cases}$$

In the adaptive setting, the price offered to a customer also depends on the set of customers that declined the offer. However, as the customers are i.i.d., an adaptive pricing mechanism only needs to know how many customers have received an offer and not exactly which customers.

As we did in Chapter 3, we partition the interval $A = [0, 1]$ into n intervals $A_{\mathrm{i}} = [\varepsilon_{\mathrm{i}-1}, \varepsilon_{\mathrm{i}}]$, with $0 = \varepsilon_0 < \varepsilon_1 < \ldots < \varepsilon_{n-1} < \varepsilon_n = 1$. The pricing mechanism will thus choose a price for customer i such that the probability that this customer accepts the offer lies in A_{i}, making sure that no customer will ever receive an offer lower than the reservation price v^*.

First, assume that the item is offered to customer i. Let q_{i} denote the drawn acceptance probability for customer i, drawn from the interval A_{i} according to the probability density

$\frac{(n-1)(1-q_i)^{n-2}}{\alpha_i}$ with α_i the normalization constant. The expected revenue obtained from selling the item to customer i is $\bar{G}(1-q_i)$. To see this, suppose that $q_i < 1 - F(v^*)$. Then, for monotone virtual valuations, the price offered to customer i is $F^{-1}(1-q_i)$, and thus the expected revenue is $q_i F^{-1}(1-q_i) = -G(1-q_i) = \bar{G}(1-q_i)$. On the other hand, if $q_i > 1 - F(v^*)$, the price offered to customer i is v^* which is accepted with probability $1-F(v^*)$. Similar arguments hold when the virtual valuation is not monotone, where it might be the case that $q_i F^{-1}(1-q_i) = -H(1-q_i) < \bar{G}(1-q_i)$, and by offering a price $F^{-1}(1-q_i)$ we might not get the best revenue. To circumvent this problem, we can randomize between two acceptance probabilities q_{i1} and q_{i2} such that $G(1-q_i) = \gamma H(1-q_{i1}) + (1-\gamma)H(1-q_{i2})$ and $q_i = \gamma q_{i1} + (1-\gamma)q_{i2}$.

Following the same reasoning as the analogous part in the proof of Theorem 3.8 (see [8]), we can bound the expected revenue of this strategy by

$$\sum_{i=1}^{n} \rho_i \int_{\varepsilon_{i-1}}^{\varepsilon_i} (n-1)(1-q)^{n-2}\bar{G}(1-q)\mathrm{d}q \tag{4.1}$$

where $\rho_1 = \frac{1}{\alpha_1}$ and $\rho_{i+1} = \frac{\rho_i}{\alpha_{i+1}} \int_{\varepsilon_{i-1}}^{\varepsilon_i} (n-1)(1-q)^{n-1}\mathrm{d}q$ for $i = 1, \dots, n-1$. Again, if we choose $\varepsilon_1, \dots, \varepsilon_{n-1}$ such that $\rho_1 = \rho_2 = \ldots = \rho_n$, and solve the recurrence $\{\varepsilon_i\}_{i=1}^{n-1}$ satisfies, then expression (4.1) equals

$$\frac{1}{n\alpha_1} OPT \geq \frac{1}{\beta} OPT \approx 0.745 OPT.$$

Algorithm

We can now state a very simple algorithm for the adaptive posted price mechanism when valuations are i.i.d. according to a regular distribution F. Note that when the distribution is non-regular, it suffices to set the prices v_i as a lottery between two other prices (although we can derandomize if we can assert which of the two prices gives higher expected revenue).

Algorithm 4.2: Adaptive i.i.d. posted price mechanism

```
Input: Distribution F, valuations V_1, ..., V_n draw from F.
Initialize r = 0, v_0 = 0, v* = reserve price.
Partition [0,1] in n disjoint intervals A_i = [a_{i-1}, a_i], i = 1, ..., n where x_i = 1 - a_i
  satisfy recurrence (3.3).
for i = 1 to n do
    Draw q_i from A_i with density proportional to (n-1)(1-q)^{n-2}.
    Set price v_i = max{v*, F^{-1}(1 - q_i)}.
    if V_i ≥ v_i then
        r ← i
        break
    end if
end for
return v_r
```

Bibliography

[1] M. Abolhassan, S. Ehsani, H. Esfandiar, M.T. Hajiaghayi, R. Kleinberg, and B. Lucier. Beating 1-1/e for ordered prophets. In *Proceedings of the Forty-ninth ACM Symposium on the Theory of Computing*, STOC '17. ACM, 2017.

[2] M. Adamczyk, A. Borodin, D. Ferraioli, B. Keijzer, and S. Leonardi. Sequential posted price mechanisms with correlated valuations. In *Proceedings of the 11th International Conference on Web and Internet Economics*, WINE 2015, pages 1–15. Springer-Verlag New York, Inc., 2015.

[3] M. Babaioff, N. Immorlica, D. Kempe, and R. Kleinberg. Online auctions and generalized secretary. In *SIGecom Exchanges*, volume 7(2), pages 1–11. ACM, 2008.

[4] L. Blumrosen and T. Holenstein. Posted prices vs. negotiations: An asymptotic analysis. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, EC '08, pages 49–49. ACM, 2008.

[5] J. Bult and T. Wansbeek. Optimal selection for direct mail. *Marketing Science*, 14(4):378–394, 1995.

[6] S. Chawla, J. D. Hartline, D. L. Malec, and B. Sivan. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the Forty-second ACM Symposium on Theory of Computing*, STOC '10, pages 311–320. ACM, 2010.

[7] R. Cominetti, J. R. Correa, T. Rothvoß, and J. San Martín. Optimal selection of customers for a last-minute offer. *Operations Research*, 58(4):878–888, 2010.

[8] J. Correa, P. Foncea, R. Hoeksma, T. Oosterwijk, and T. Vredeveld. Posted price mechanisms for a random stream of customers. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC '17, pages 169–186. ACM, 2017.

[9] H. Esfandiari, M. Hajiaghayi, V. Liaghat, and M. Monemizadeh. Prophet secretary. In *Algorithms-ESA 2015*, pages 496–508. Springer, 2015.

[10] J. P. Gilbert and F. Mosteller. Recognizing the maximum of a sequence. *Journal of the American Statistical Association*, 61:35–76, 1966.

[11] M. T. Hajiaghayi, R. Kleinberg, and D. C. Parkes. Adaptive limited-supply online auctions. In *Proceedings of the Fifth ACM conference on Electronic Commerce*, EC '04,

pages 71–80. ACM, 2004.

[12] J. Hartline. Mechanism design and approximation, 2016.

[13] T.P. Hill and R.P. Kertz. Comparisons of stop rule and supremum expectations of i.i.d. random variables. *The Annals of Probability*, 10(2):336–345, 1982.

[14] R.P. Kertz. Stop rule and supremum expectations of i.i.d. random variables: A complete comparison by conjugate duality. *Journal of Multivariate Analysis*, 19:88–112, 1986.

[15] R. Kleinberg and S. M. Weinberg. Matroid prophet inequalities. In *Proceedings of the Forty-fourth ACM Symposium on Theory of Computing*, STOC '12, pages 123–136. ACM, 2012.

[16] U. Krengel and L. Sucheston. Semiamarts and finite values. *Bull. Amer. Math. Soc.*, 83:745–747, 1977.

[17] U. Krengel and L. Sucheston. On semiamarts, amarts, and processes with finite value. *Advances in Probability*, 4:197–266, 1978.

[18] E. Lindelöf. Sur l'application de la méthode des approximations successives aux équations différentielles ordinaires du premier ordre. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 116:454–457, 1894.

[19] R. B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

[20] E. Samuel-Cahn. Comparisons of threshold stop rule and maximum for independent nonnegative random variables. *The Annals of Probability*, 12(4):1213–1216, 1983.

[21] Q. Yan. Mechanism design via correlation gap. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, SODA '11, pages 710–719. ACM-SIAM, 2011.