WEAK AGGREGATING ALGORITHM FOR THE DISTRIBUTION-FREE PERISHABLE INVENTORY PROBLEM

TATSIANA LEVINA, YURI LEVIN, JEFF MCGILL, MIKHAIL NEDIAK
SCHOOL OF BUSINESS, QUEEN'S UNIVERSITY, 143 UNION ST., KINGSTON, ON, K7L 3N6, CANADA
E-MAIL:{TLEVIN,YLEVIN,JMCGILL,MNEDIAK}@BUSINESS.QUEENSU.CA
AND VLADIMIR VOVK
DEPARTMENT OF COMPUTER SCIENCE, ROYAL HOLLOWAY, UNIVERSITY OF LONDON,
EGHAM, SURREY TW20 0EX, UK
E-MAIL:VOVK@CS.RHUL.AC.UK

Abstract. We formulate the multiperiod, distribution-free perishable inventory problem as a problem of prediction with expert advice and apply an online learning method (the Weak Aggregating Algorithm) to solve it. We show that the asymptotic average performance of this method is as good as that of any time-dependent stocking rule in a given parametric class.

Keywords: online learning; newsvendor problem; aggregating algorithm.

1. Introduction

In the classical perishable inventory (also called "newsvendor") problem, a decision-maker must choose the starting inventory level for a product in a future selling period when demand for the product is uncertain. No replenishment is possible during the selling period, and the product is perishable, *i.e.* it loses part or all of its value at the end of the period. The newsvendor aims to achieve maximum return by balancing the risk of lost sales because of understocking against that of inventory spoilage because of overstocking.

This problem is a practical inventory control problem faced by many firms whose products are perishable (including newspaper publishers), but it is also a component of other important problems; for example, bank cash management, water reservoir management, and airline revenue management. The problem has a simple solution when the probability distribution of demand is known but becomes significantly more challenging when distributional information is limited, inaccurate, or unavailable. We deal with the case in which there is no prior knowledge of the distribution.

In the distribution-free case, if the newsvendor faces only a single decision period, the problem falls into the category of decision analysis under uncertainty typified by the min-max approach of [10]. If there are multiple selling periods, it is possible to gain information about the demand distribution over time and adapt ordering policies accordingly. (We will use the term multi-period to mean multiple selling periods rather than multiple opportunities to order, an interpretation that has been used in some prior research.) Recent examples of work on the multi-period version include both Bayesian (see [8]) and nonparametric approaches (see [11]). The latter paper contains an up-to-date review of the extensive literature in this area. Other articles proposing non-parametric approaches and/or bound for inventory problems include [1] which proposes distribution-free upper and lower bounds for the order quantity and reorder point in a service-constrained non-perishable inventory system. [7] and [3] consider perishable inventory system where the functional form of demand distribution is known and develop operational statistics approach to find a decision rule that maximizes the performance uniformly for all

Corresponding author: Mikhail Nediak. E-mail: mnediak@business.queensu.ca. Address: School of Business, Queen's University, 143 Union St.,Kingston, ON, K7L 3N6, CANADA.

possible values of the unknown demand parameters. Finally, [6] considers a sampling-based approach in which it is possible to obtain bounds on the number of samples needed to attain a specified accuracy level.

In this article, we propose a novel approach to the distribution-free, multi-period problem that utilizes recent advances in the theory of prediction and learning with expert advice (see Chapter 2 of [2]). This approach leads to an algorithm with performance guarantees under more general assumptions than those previously achieved (existing non-parametric results require, at least, independence of demands over time). The 'experts' in this treatment are passive predictors of best starting inventory levels over time out of the continuum of possible levels suggested by functions in a given parametric class with a bounded finite-dimensional parameter space; thus, we consider an infinite pool of experts. (The case of a finite number of stocking levels, and hence experts, is a straightforward variant of this.) The algorithm progresses, in essence, by forming successive weighted averages of the expert predictions, where the weights are adjusted according to the success of the experts in previous periods.

We make the following contributions:

(1) We cast the newsvendor problem as online learning with expert advice and show that the Weak Aggregating Algorithm (WAA) of [5] can be applied to the problem.
(2) We prove that the performance of the algorithm is asymptotically as good as the performance of the best time-dependent strategy in a given parametric class with a bounded parameter space. Thus, in the setting of the newsvendor problem, we obtain stronger results than the existing analysis of the WAA for a finite collection of experts.
(3) The performance bound holds in the absence of any statistical assumptions about the demand sequence.

In the next section we provide a formal statement of the WAA. The newsvendor problem, the explicit WAA specialized to the newsvendor problem and its analysis are given in §3.

2. The General Weak Aggregating Algorithm

The Aggregating Algorithm (AA) [12] is a general approach to online learning that involves combining or 'merging' advice from a pool of experts (typically finite). The objective is to minimize the losses from a sequence of decisions that must be made in a stochastic environment. The convergence of the AA is moderated with a learning rate parameter that can be adjusted for each particular application but is otherwise constant. The Weak Aggregating Algorithm (WAA) is similar to the AA but uses a learning rate parameter that is proportional to $\sqrt{n}$. In this section we describe a version of the WAA that utilizes a continuum of experts. For convenience, we modify the description in obvious ways to allow for maximization of gains instead of minimization of losses.

In the classical problem of prediction with expert advice, there are three entities or players: Master, Nature, and Pool. The players have perfect information in the usual sense that payoffs, decision sets and decisions in the previous stages are known to all participants. Master is the entity whose actions or decisions we want to improve, and Nature is the source of uncertainty, with outcomes which are unknown at the time of Master's decision. Pool and its experts represent alternative actions to learn from and to benchmark against.

Each expert in Pool is uniquely identified by an index θ, which could be a simple scalar, or in more general settings, a more complex mathematical object. For example, if the experts are various regression models, the θ's could be vectors of different possible regression coefficients. The collection of indices form a set Θ. We assume that it is possible to define a probability measure over Θ; that is, Θ is assumed to be measurable.

The decisions made by both the experts and Master are chosen from a set Γ (assumed to be a convex subset of a linear space). Nature chooses the actual outcomes from a set Ω. The gain associated with decision $\gamma \in \Gamma$ and outcome of Nature $\omega \in \Omega$ is given by the function $\pi(\omega, \gamma)$.

Here is the standard perfect-information protocol of prediction with expert advice:

Prediction with expert advice

FOR $n = 1, 2, \ldots$:

  Pool announces a measurable function mapping $\theta \in \Theta$ to $\gamma_n^\theta \in \Gamma$

  Master announces $\gamma_n \in \Gamma$

  Nature announces $\omega_n \in \Omega$

  Master obtains gain $\pi(\omega_n, \gamma_n)$

  Each expert $\theta \in \Theta$ obtains gain $\pi(\omega_n, \gamma_n^\theta)$

END FOR

The WAA tracks the relative performances of experts with a distribution of weights, with relatively greater weights assigned to relatively more successful experts. The algorithm starts from some initial (prior) probability measure and, at each iteration, uses a weighted average to merge the recommendations of the experts to produce a decision that is used by Master.

Let $q(\mathrm{d}\theta)$ be a prior probability measure on Θ. Define

$$g_n^{(\theta)} := \pi\left(\omega_n, \gamma_n^\theta\right), \quad G_N^{(\theta)} := \sum_{n=1}^{N} g_n^{(\theta)}$$

to be, respectively, the instantaneous gain of the θth expert in the nth round and his cumulative gain over the first N rounds. For all $n = 1, 2, \ldots$ define

$$w_n(\mathrm{d}\theta) := \exp\left(\frac{G_{n-1}^{(\theta)}}{\sqrt{n}}\right) q(\mathrm{d}\theta).$$

The terms $w_n(\mathrm{d}\theta)$ are the weights of the experts to use in round n. These weights are the prior weights adjusted by an exponential term that is increasing in the cumulative gain associated with each θ. The exponential form for the re-weighting coefficient ensures nonnegative weights regardless of the sign of the gains. For convenience in subsequent expressions, we use $\beta_n := \exp\left(\frac{1}{\sqrt{n}}\right)$. This is consistent with the use of β as the fixed learning rate in the (Strong) Aggregating Algorithm.

Now define the normalized weights:

$$p_n(\mathrm{d}\theta) := \frac{w_n(\mathrm{d}\theta)}{w_n(\Theta)}.$$

The WAA's prediction in round n is then

$$\gamma_n := \int_\Theta \gamma_n^\theta p_n(\mathrm{d}\theta). \tag{1}$$

Let $g_n := \pi(\omega_n, \gamma_n)$ be the WAA's gain in round n and $G_N := \sum_{n=1}^{N} g_n$ be its cumulative gain over the first N rounds.

Lemma 1. *If the gain function $\pi(\omega, \gamma)$ is concave in γ, the WAA guarantees that, for all $N = 1, 2, \ldots$,*

$$G_N \geq \sum_{n=1}^{N} \int_\Theta g_n^{(\theta)} p_n(\mathrm{d}\theta) - \sum_{n=1}^{N} \log_{\beta_n} \int_\Theta \beta_n^{g_n^{(\theta)}} p_n(\mathrm{d}\theta) + \log_{\beta_N} \int_\Theta \beta_N^{G_N^{(\theta)}} q(\mathrm{d}\theta). \tag{2}$$

Proof. In the proof, we will use the following consequence of the concavity of the gain function:

$$g_n \geq \int_\Theta g_n^{(\theta)} p_n(\mathrm{d}\theta). \tag{3}$$

This property follows from Jensen's inequality.

The proof is by induction on N. To establish the base case of $N = 1$, we observe that $p_1(d\theta) = w_1(d\theta) = q(d\theta)$ since $G_0^{(\theta)} = 0$. Thus, the last two terms in (2) cancel since $G_1^{(\theta)} = g_1^{(\theta)}$, and the whole inequality becomes equivalent to (3) for $n = 1$. Assuming (2) holds for N, we obtain

$$G_{N+1} = G_N + g_{N+1} \geq G_N + \int_\Theta g_{N+1}^{(\theta)} p_{N+1}(\mathrm{d}\theta)$$
$$\geq \sum_{n=1}^{N+1} \int_\Theta g_n^{(\theta)} p_n(\mathrm{d}\theta) - \sum_{n=1}^{N} \log_{\beta_n} \int_\Theta \beta_n^{g_n^{(\theta)}} p_n(\mathrm{d}\theta) + \log_{\beta_N} \int_\Theta \beta_N^{G_N^{(\theta)}} q(\mathrm{d}\theta)$$

(the first "$\geq$" used the property (3)). Therefore, it remains to prove

$$\log_{\beta_N} \int_\Theta \beta_N^{G_N^{(\theta)}} q(\mathrm{d}\theta) \geq -\log_{\beta_{N+1}} \int_\Theta \beta_{N+1}^{g_{N+1}^{(\theta)}} p_{N+1}(\mathrm{d}\theta) + \log_{\beta_{N+1}} \int_\Theta \beta_{N+1}^{G_{N+1}^{(\theta)}} q(\mathrm{d}\theta).$$

By the definition of $p_n(\mathrm{d}\theta)$ this can be rewritten as

$$\log_{\beta_N} \int_\Theta \beta_N^{G_N^{(\theta)}} q(\mathrm{d}\theta) \geq -\log_{\beta_{N+1}} \frac{\int_\Theta \beta_{N+1}^{G_N^{(\theta)}} \beta_{N+1}^{g_{N+1}^{(\theta)}} q(\mathrm{d}\theta)}{\int_\Theta \beta_{N+1}^{G_N^{(\theta)}} q(\mathrm{d}\theta)} + \log_{\beta_{N+1}} \int_\Theta \beta_{N+1}^{G_{N+1}^{(\theta)}} q(\mathrm{d}\theta),$$

which after cancellation becomes

$$\log_{\beta_N} \int_\Theta \beta_N^{G_N^{(\theta)}} q(\mathrm{d}\theta) \geq \log_{\beta_{N+1}} \int_\Theta \beta_{N+1}^{G_N^{(\theta)}} q(\mathrm{d}\theta). \tag{4}$$

The last inequality follows from the general result about comparison of different means ([4], Theorem 85), but we can also check it directly. Let $\beta_{N+1} = \beta_N^a$, where $0 < a < 1$ (assuming $N \neq 0$). Then (4) can be rewritten as

$$\left(\int_\Theta \beta_N^{G_N^{(\theta)}} q(\mathrm{d}\theta) \right)^a \geq \int_\Theta \beta_N^{aG_N^{(\theta)}} q(\mathrm{d}\theta),$$

and the last inequality follows from the concavity of the function $t \mapsto t^a$. □

The first two terms on the right-hand side of (2) are sums over the first N rounds of different kinds of mean of the experts' gains (see, for example, [4], Chapter III, for generalized definitions of the mean); we will see later that they nearly cancel each other out. If those two terms are ignored, the remaining part of (2) is similar to the analogous bounding term of the Aggregating Algorithm (see, e.g., [12], Lemma 1), except that, here, the learning rate β depends on n. Proof of Lemma 1 can be found in the Appendix.

The following lemma (also proved in the Appendix) provides the main performance bound for the sequence of predictions generated by the WAA for the case of the bounded gain function. For consistency with the original formulation of the WAA as a loss-minimizing algorithm, we express gains π as negative.

Lemma 2. *Let $\pi \in [-L, 0]$. The WAA guarantees that, for all N,*

$$G_N \geq \left(\ln \int_\Theta e^{G_N^{(\theta)}/\sqrt{N}} q(\mathrm{d}\theta) - L^2 \right) \sqrt{N}. \tag{5}$$

Proof. From (2), we obtain:

$$
\begin{aligned}
G_N &\geq \sum_{n=1}^{N} \int_\Theta g_n^{(\theta)} p_n(\mathrm{d}\theta) - \sum_{n=1}^{N} \sqrt{n} \ln \int_\Theta \exp\left(\frac{g_n^{(\theta)}}{\sqrt{n}}\right) p_n(\mathrm{d}\theta) + \log_{\beta_N} \int_\Theta \beta_N^{G_N^{(\theta)}} q(\mathrm{d}\theta) \\
&\geq \sum_{n=1}^{N} \int_\Theta g_n^{(\theta)} p_n(\mathrm{d}\theta) - \sum_{n=1}^{N} \sqrt{n} \left(\int_\Theta \left(1 + \frac{g_n^{(\theta)}}{\sqrt{n}} + \frac{\left(g_n^{(\theta)}\right)^2}{2n}\right) p_n(\mathrm{d}\theta) - 1 \right) + \log_{\beta_N} \int_\Theta \beta_N^{G_N^{(\theta)}} q(\mathrm{d}\theta) \\
&= -\frac{1}{2} \sum_{n=1}^{N} \frac{1}{\sqrt{n}} \int_\Theta \left(g_n^{(\theta)}\right)^2 p_n(\mathrm{d}\theta) + \sqrt{N} \ln \int_\Theta e^{G_N^{(\theta)}/\sqrt{N}} q(\mathrm{d}\theta) \\
&\geq -\frac{L^2}{2} \sum_{n=1}^{N} \frac{1}{\sqrt{n}} + \sqrt{N} \ln \int_\Theta e^{G_N^{(\theta)}/\sqrt{N}} q(\mathrm{d}\theta) \\
&\geq -\frac{L^2}{2} \int_0^N \frac{\mathrm{d}t}{\sqrt{t}} + \sqrt{N} \ln \int_\Theta e^{G_N^{(\theta)}/\sqrt{N}} q(\mathrm{d}\theta) \\
&= -L^2 \sqrt{N} + \sqrt{N} \ln \int_\Theta e^{G_N^{(\theta)}/\sqrt{N}} q(\mathrm{d}\theta)
\end{aligned}
$$

(in the second "$\geq$" we used the inequalities $e^t \leq 1 + t + \frac{t^2}{2}$, where $t \leq 0$, and $\ln t \leq t - 1$, where $t > 0$). ☐

The following corollary provides a specialized bound for the case of discrete prior distributions:

Corollary 1. *Let $\pi \in [-L, 0]$ and the prior q be discrete. The WAA guarantees that, for all N and $\theta \in \Theta$,*

$$G_N \geq G_N^{(\theta)} + \left(\ln q(\{\theta\}) - L^2\right)\sqrt{N}.$$

Proof. Replace the $\int_\Theta e^{G_N^{(\theta)}/\sqrt{N}} q(\mathrm{d}\theta)$ in (5) by its trivial lower bound $e^{G_N^{(\theta)}/\sqrt{N}} q(\{\theta\})$. ☐

3. The Weak Aggregating Algorithm for the Newsvendor Problem

Let p and c be the unit selling price and cost of a product and assume that the value of unsold inventory at the end of the selling period is zero. The case of a positive unit salvage value s reduces to the basic case by redefining $c := c - s$ and $p := p - s$.

The newsvendor may face a stocking decision indefinitely many times, but the case of a finite horizon, defined by a terminal period N, is also covered by our result. His decision in each selling period $n = 1, 2 \ldots$ is $y_n \in [0, B]$, where B is a known upper bound on the possible stocking level. The demand realization D_n of period n is unknown at the time of the stocking decision. The newsvendor profit in a period, given decision y and demand D, is equal to $\pi(D, y) := p \min(y, D) - cy$, a concave function in y.

The performance of the newsvendor is judged in terms of the *total gain*, which, in period n, is $G_n := \sum_{k=1}^{n} \pi(D_k, y_k)$. It is compared to the gain $G_n^{(\theta)}$ of strategies (experts) in a given collection Θ: $G_n^{(\theta)} := \sum_{k=1}^{n} \pi(D_k, \gamma_n^\theta)$.

With the newsvendor problem, Master is the newsvendor, and Nature is the demand generating process. We use Pool which is the continuum of nonnegative functions γ_n^θ parameterized by d-dimensional vector θ in a hyperrectangle $\Xi = X_1 \times \ldots \times X_d$ (that is, X_i is an interval in $\mathbb{R}$ of length $|X_i|$, $i = 1, \ldots, d$). We assume that

(1) set Ξ has finite volume $V = \prod_{i=0}^{d} |X_i|$ and diameter $R = \sqrt{\sum_{i=0}^{d} |X_i|^2}$;
(2) γ_n^θ's are Cauchy-continuous with respect to θ with the same constant K; and
(3) there exists θ^0 such that $\gamma_n^{\theta^0} \equiv 0$ for all $n \geq 0$.

A very simple example of a strategy in this general class is a fixed stock level in the interval $[0, B]$. For this simple case, $\theta^0 = 0$, $R = B$ and $K = 1$. In essence, the game is played between just two players, the newsvendor and Nature, according to the following protocol:

Newsvending protocol

$G_0 := 0$
FOR $n = 1, 2, \dots$:
  Newsvendor announces $y_n \in [0, B]$
  Nature announces $D_n \in [0, \infty]$
  $G_n := G_{n-1} + \pi(D_n, y_n)$
END FOR

3.1. **Explicit algorithm for the case of "fixed-stock" experts.** The key to specialization of WAA is computation of the integral (1). For this specialization, we assume that the expert corresponding to $\theta = y \in [0, B]$ always recommends $\gamma_n^\theta = y$ and consider the period n decision. Let the (truncated to $[0, B]$) order statistics of demand for the first $n-1$ periods be $D^n_{(1)}, \dots, D^n_{(n-1)}$. We augment this ordered sequence by the endpoints of the interval: $D^n_{(0)} := 0$ and $D^n_{(n)} := B$. The upper index n will be omitted for brevity. We can transform the integral (1) as follows:

$$\gamma_n = \int_\Theta \gamma_n^\theta p_n(\mathrm{d}\theta) = \frac{\int_\Theta \gamma_n^\theta e^{G^{(\theta)}_{n-1}/\sqrt{n}} q(\mathrm{d}\theta)}{\int_\Theta e^{G^{(\theta)}_{n-1}/\sqrt{n}} q(\mathrm{d}\theta)} = \frac{\int_0^B y e^{G^{(y)}_{n-1}/\sqrt{n}} q(\,\mathrm{d}y)}{\int_0^B e^{G^{(y)}_{n-1}/\sqrt{n}} q(\,\mathrm{d}y)} = \frac{a_n}{\tilde{a}_n}. \tag{6}$$

The numerator of (6) can be evaluated as

$$\begin{aligned} a_n := \int_0^B y e^{G^{(y)}_{n-1}/\sqrt{n}} q(\,\mathrm{d}y) &= \sum_{k=0}^{n-1} \int_{D_{(k)}}^{D_{(k+1)}} y \exp\left(n^{-1/2} G^{(y)}_{n-1}\right) q(\,\mathrm{d}y) \\ &= \sum_{k=0}^{n-1} \int_{D_{(k)}}^{D_{(k+1)}} y \exp\left(n^{-1/2} \sum_{i=1}^{n-1} \left(p \min(y, D_{(i)}) - cy\right)\right) q(\,\mathrm{d}y) \\ &= \sum_{k=0}^{n-1} \int_{D_{(k)}}^{D_{(k+1)}} y \exp\left(n^{-1/2}\left(p \sum_{i=1}^{k} D_{(i)} + p(n-k-1)y - c(n-1)y\right)\right) q(\,\mathrm{d}y). \end{aligned}$$

A similar expression (without y in front of exp) is available for denominator $\tilde{a}_n$. For general prior distribution $q(\mathrm{d}y)$ it is possible to use numerical integration to evaluate a_n and $\tilde{a}_n$. Specific forms of priors $q(\mathrm{d}y)$ may be chosen to capture initial perceptions of Newsvendor about the market. In the absence of some strong perceptions/initial information about the market, Newsvendor may consider a uniform prior on $[0, B]$.

When the prior distribution $q(\,\mathrm{d}y)$ is uniform on $[0, B]$, the numerator and denominator have explicit forms resulting in the following expression for suggested stock levels:

$$\gamma_n = \frac{\sum_{k=0}^{n-1} \exp\left(n^{-1/2} p \sum_{i=1}^{k} D_{(i)}\right) I_k^n}{\sum_{k=0}^{n-1} \exp\left(n^{-1/2} p \sum_{i=1}^{k} D_{(i)}\right) \tilde{I}_k^n}, \tag{7}$$

where, for $k = 1, \dots, n-1$, we let $\alpha_k^n := \frac{n^{1/2}}{p(n-k-1)-c(n-1)}$ and

$$I_k^n \quad := \quad \begin{cases} \alpha_k^n \left[\exp\left(\frac{y}{\alpha_k^n}\right)(y-\alpha_k^n)\right]_{D_{(k)}}^{D_{(k+1)}} & \text{if } p(n-k-1) \neq c(n-1), \\ \frac{1}{2}[y^2]_{D_{(k)}}^{D_{(k+1)}} & \text{if } p(n-k-1) = c(n-1). \end{cases}$$

$$\tilde{I}_k^n \quad := \quad \begin{cases} \alpha_k^n \left[\exp\left(\frac{y}{\alpha_k^n}\right)\right]_{D_{(k)}}^{D_{(k+1)}} & \text{if } p(n-k-1) \neq c(n-1), \\ [y]_{D_{(k)}}^{D_{(k+1)}} & \text{if } p(n-k-1) = c(n-1). \end{cases}$$

The WAA for this case consists in selecting the initial stock level in period n according to (7). This calculation can be accomplished in at most order of n time. The algorithm requires storage of the sorted demand history (order of n storage space).

Next, we show that the WAA for Newsvendor problem converges to the optimal stock level for the case of independent identically distributed demands. This issue is important from the practical point of view and serves as a sanity check for the proposed algorithm. (We use the assumption of i.i.d. demands for Proposition 1 only – *not* for the general performance bound presented in the next section.) It is well known that if demands follow a given cumulative distribution function $F(x)$, then Newsvendor can determine the optimal order quantity y^* as a solution to equation $F(y^*) = \frac{p-c}{p}$. However, in the setting of this paper, Newsvendor does not know $F(x)$ and cannot compute y^*. Nevertheless, the following statement holds:

Proposition 1. *Let D_n, $n \geq 1$ be independent identically distributed continuous random variables, y^* – the (unknown) optimal stock level, and suppose there exists a neighborhood of y^* in which $F(y)$ is continuously differentiable and $f(y) = F'(y)$ is strictly positive. The sequence of stocking levels γ_n produced by WAA with Pool $\Theta = [0, B]$ and uniform prior $q(\mathrm{d}y) = \frac{\mathrm{d}y}{B}$ converges to y^* in probability, i.e. for any $\epsilon > 0$,*

$$\lim_{n\to\infty} P(|\gamma_n - y^*| < \epsilon) = 1.$$

Proof. Fix sufficiently small $\epsilon > 0$ so that $\epsilon \leq \min\{y^*, B - y^*\}$ and $f(y) > 0$ for $y \in U_\epsilon(y^*)$, and define $l(n) = \lfloor nF(y^* - \frac{\epsilon}{2})\rfloor$, $m(n) = \lfloor n\frac{p-c}{p} + \frac{c}{p}\rfloor = \lfloor nF(y^*) + \frac{c}{p}\rfloor$, and $u(n) = \lceil nF(y^* + \frac{\epsilon}{2})\rceil$. Observe that order statistic $D_{(m(n))}$ of a size $n-1$ sample of demands delivers the maximum of $G_{n-1}^{(y)}$ as a function of y. Consider also size $n-1$ sample order statistics $D_{(l(n))}$, $D_{(u(n))}$ and event $A_n = \{D_{(l(n))} \in U_{\epsilon/8}(y^* - \frac{\epsilon}{2})\} \cap \{D_{(m(n))} \in U_{\epsilon/8}(y^*) \cap \{D_{(u(n))} \in U_{\epsilon/8}(y^* + \frac{\epsilon}{2})\}$. Theorem 4.1.3 on page 110 of [9] implies that $\lim_{n\to\infty} P(A_n) = 1$. (Intuitively, each of the considered order statistics is a consistent estimator of the corresponding percentile of the demand distribution.) We show that there exists a constant n_ϵ such that $A_n \subseteq \{|\gamma_n - y^*| < \epsilon\}$ for all $n \geq n_\epsilon$, proving the proposition.

Indeed, inequality $|\gamma_n - y^*| < \epsilon$ is implied by $\int_0^B |y - y^*| e^{G_{n-1}^{(y)}/\sqrt{n}} q(\mathrm{d}y) < \epsilon \int_0^B e^{G_{n-1}^{(y)}/\sqrt{n}} q(\mathrm{d}y)$. The latter inequality is, in turn, implied by the pair of inequalities

$$\int_0^{y^*} (y^* - y) e^{G_{n-1}^{(y)}/\sqrt{n}} q(\mathrm{d}y) \quad < \quad \epsilon \int_0^{y^*} e^{G_{n-1}^{(y)}/\sqrt{n}} q(\mathrm{d}y), \tag{8}$$

$$\int_{y^*}^B (y - y^*) e^{G_{n-1}^{(y)}/\sqrt{n}} q(\mathrm{d}y) \quad < \quad \epsilon \int_{y^*}^B e^{G_{n-1}^{(y)}/\sqrt{n}} q(\mathrm{d}y). \tag{9}$$

Since $G_{n-1}^{(y)}$ has a maximum at $D_{(m(n))}$ and its slope is equal to $p(n-k) - c(n-1)$ for $D_{(k-1)} < y < D_{(k)}$, condition A_n implies that $G_{n-1}^{(y)}$ is increasing for $y \leq y^* - \frac{\epsilon}{8}$ with a slope of at least $np(F(y^*) - F(y^* - \frac{\epsilon}{2}))$ for

$y < y^* - \frac{5\epsilon}{8}$, and decreasing for $y \geq y^* + \frac{\epsilon}{8}$ with a slope of at most $np(F(y^*) - F(y^* + \frac{\epsilon}{2}))$ for $y > y^* + \frac{5\epsilon}{8}$. By the conditions of the proposition, the quantities $F(y^*) - F(y^* - \frac{\epsilon}{2})$ and $F(y^*) - F(y^* + \frac{\epsilon}{2})$ are strictly positive. We now show that there exists n'_ϵ such that A_n implies (8) for $n \geq n'_\epsilon$. Assuming A_n, we have

$$\int_0^{y^*} (y^* - y) e^{G_{n-1}^{(y)}/\sqrt{n}} q(\,\mathrm{d}y) = \int_0^{y^* - \frac{7\epsilon}{8}} (y^* - y) e^{G_{n-1}^{(y)}/\sqrt{n}} q(\,\mathrm{d}y) + \int_{y^* - \frac{7\epsilon}{8}}^{y^*} (y^* - y) e^{G_{n-1}^{(y)}/\sqrt{n}} q(\,\mathrm{d}y)$$

$$\leq \frac{y^*}{B} e^{G_{n-1}^{(y^* - \frac{7\epsilon}{8})}/\sqrt{n}} + \frac{7\epsilon}{8} \int_{y^* - \frac{7\epsilon}{8}}^{y^*} e^{G_{n-1}^{(y)}/\sqrt{n}} q(\,\mathrm{d}y) \leq \frac{\epsilon^2}{16B} e^{G_{n-1}^{(y^* - \frac{5\epsilon}{8})}/\sqrt{n}} + \frac{7\epsilon}{8} \int_{y^* - \frac{7\epsilon}{8}}^{y^*} e^{G_{n-1}^{(y)}/\sqrt{n}} q(\,\mathrm{d}y)$$

as long as $\sqrt{n} \ln \frac{16 y^*}{\epsilon^2} \leq G_{n-1}^{(y^* - \frac{5\epsilon}{8})} - G_{n-1}^{(y^* - \frac{7\epsilon}{8})}$. From the condition on the minimum slope, the latter holds if $\sqrt{n} \ln \frac{16 y^*}{\epsilon^2} \leq \frac{\epsilon}{4} np(F(y^*) - F(y^* - \frac{\epsilon}{2}))$ or, equivalently, if $n \geq n'_\epsilon := \left(\frac{4 \ln(16 y^*/\epsilon^2)}{\epsilon p (F(y^*) - F(y^* - \frac{\epsilon}{2}))} \right)^2$. Since we also have

$$\frac{\epsilon^2}{16B} e^{G_{n-1}^{(y^* - \frac{5\epsilon}{8})}/\sqrt{n}} \leq \frac{\epsilon}{8} \int_{y^* - \frac{5\epsilon}{8}}^{y^* - \frac{\epsilon}{8}} e^{G_{n-1}^{(y)}/\sqrt{n}} q(\,\mathrm{d}y),$$

it follows that inequality (8) holds for all $n \geq n'_\epsilon$. A similar analysis of (9) shows that there exists n''_ϵ such that A_n implies (9) for all $n \geq n''_\epsilon$. We complete the proof by taking $n_\epsilon = \max\{n'_\epsilon, n''_\epsilon\}$. □

3.2. **Main Result.** The main result of this paper is that the newsvendor has a strategy, provided by the WAA, for which the time-average returns are asymptotically as good as the best fixed-stock strategy, regardless of the sequence of demand realizations. The theorem applies to the general class of strategies parameterized by $\theta \in \Xi$.

Theorem 1. *Suppose that γ_n^θ satisfy assumptions 1-3. Newsvendor has a strategy (WAA with a uniform prior $q(\mathrm{d}\theta) = \frac{\mathrm{d}\theta}{V}$ on Pool Ξ) that guarantees*

$$G_N \geq \max_{\theta \in \Xi} G_N^{(\theta)} - \left(K^2 R^2 p^2 + KRp + \ln N^{d/2} \right) \sqrt{N}, \tag{10}$$

for all $N = 1, 2, \ldots$.

Proof. In the proof, we utilize Lemma 2. Since identical-zero function belongs to the given class of experts, and the functions are Cauchy-continuous with constant K, the maximum value of any function is bounded by $B := KR$. The lowest and highest possible gains in one period are, respectively, $-Bc$ and $B(p-c)$; that is, the gain function for the newsvendor satisfies

$$-Bc \leq \pi(y, D) \leq B(p - c).$$

Without loss of generality, we can redefine $\pi(y, D) := \pi(y, D) - B(p - c)$, and obtain:

$$-Bp \leq \pi(y, D) \leq 0.$$

Defining $L := Bp = KRp$, we put our Newsvendor problem in the framework of Lemma 2. Let $m := \max(c, p-c) \leq p$ be the least upper bound of the absolute value of the slope of the gain function as a function of y. We can bound the integral $\int_\Theta e^{G_N^{(\theta)}/\sqrt{N}} q(\mathrm{d}\theta)$ in (5) from below by replacing $\Theta = \Xi$ by the neighborhood of any given θ that is a direct product of 1-dimensional $|X_i|/\sqrt{N}$-neighborhoods of θ's individual components θ_i, $i = 1, \ldots, d$. The volume of this neighborhood is at least $\prod_{i=1}^d (|X_i|/\sqrt{N}) = V/N^{d/2}$. Indeed, in X_i, the 1-dimensional neighborhood of θ_i is an interval of length at least $|X_i|/\sqrt{N}$ (its length is exactly $|X_i|/\sqrt{N}$ if θ_i is one of the end-points of X_i). Intuitively, bounding the integral in this way is reasonable since the best-performing values of θ will asymptotically have much higher weights than the poorly performing ones. As a result, most of the weight will concentrate in the neighborhood of the best values. Since, on this neighborhood, $\|\theta' - \theta\| \leq \sqrt{\sum_{i=1}^{d} |X_i|^2/N} = R/\sqrt{N}$, $|\pi(\gamma_n^{\theta'}, D) - \pi(\gamma_n^{\theta}, D)| \leq m|\gamma_n^{\theta'} - \gamma_n^{\theta}| \leq mK\|\theta' - \theta\|$, and $G_N^{(\theta')} \geq G_N^{(\theta)} - NmKR/\sqrt{N}$, we have

$$\int_{\Xi} e^{G_N^{(\theta)}/\sqrt{N}} q(\mathrm{d}\theta) \geq \frac{1}{N^{d/2}} e^{(G_N^{(\theta)} - NKRm/\sqrt{N})/\sqrt{N}} = \frac{1}{N^{d/2}} e^{G_N^{(\theta)}/\sqrt{N} - KRm}.$$

Substituting this into (5) gives

$$G_N \geq \left(G_N^{(y)}/\sqrt{N} - KRm - \ln N^{d/2} - L^2 \right) \sqrt{N}.$$

The final result is obtained after substitution of p for m (using $m \leq p$) and KRp for L. □

A consequence of this bound is that the average performance of Newsvendor utilizing the WAA is at most an order of $\frac{\ln(N)}{\sqrt{N}}$ worse than any of the the experts in Ξ. As $N \to \infty$, the performance of Newsvendor approaches that of the best expert regardless of the demand realization sequence and the form of the demand distribution. This result holds even for dependent and adversarially selected demand sequences. For the case of $\Xi = [0, B]$ – fixed stock levels in the interval $[0, B]$, the bound (10) takes the form

$$G_N \geq \max_{y \in [0,B]} G_N^{(y)} - \left(B^2p^2 + Bp + \ln \sqrt{N} \right) \sqrt{N}.$$

A perceived limitation of Theorem 1 is the use of prior distribution on the initial collection of experts even though the algorithm may be used with other priors. However, this limitation is rather mild. If the prior distribution is not uniform, the proof will still work, but it will lead to a more awkward expression, without adding much power to the result. The reason is that bound (10) is very robust with respect to the choice of the prior distribution on Ξ. Since in Lemma 2 the term $q(d\theta)$ is within the scope of ln, increasing the prior density 100-fold for a range of experts will only lead to an additive term of $\ln 100 \approx 4.6$ in the parentheses in the statement of Theorem 1 (for y's in this range). Of course, even such a modest improvement might be desirable; however, the price of achieving it is the loss of competitiveness with respect to experts outside the chosen range. Thus, the most interesting variations of the algorithm arise when we consider a more general class of experts rather that a more general prior distribution.

We illustrate one such possibility in situations where Newsvendor may expect demand to follow S-shaped multiplicative trend indicative of market saturation. One class of S-shaped trend laws is Gompertz curve $f(n) = e^{b(1-e^{c(1-n)})}$, $n \geq 1$ where $b, c > 0$. Thus, we let $\theta = (y, b, c)$ where $y \in [0, \bar{B}]$ represents suggested first-period stock level and $b \in [\underline{b}, \bar{b}]$, $c \in [\underline{c}, \bar{c}]$ represent parameters of the trend curve so that

$$\gamma_n^{\theta} = y \times e^{b(1-e^{c(1-n)})}.$$

Parameter b, together with y, determines the maximum market size ye^b, while parameter c determines the speed of growth. To apply WAA, the prior distribution needs to be defined on $[0, \bar{B}] \times [\underline{b}, \bar{b}] \times [\underline{c}, \bar{c}]$, for example, by forming a product of independent priors on y, b and c:

$$q(\mathrm{d}\theta) = q(\mathrm{d}y)q(\mathrm{d}b)q(\mathrm{d}c).$$

The required integrals in (1) with respect to measure $q(\mathrm{d}\theta)$ can be evaluated numerically.

Acknowledgment. This work was partially supported by Natural Sciences and Engineering Research Council of Canada grant numbers 261512, 341412, and 388724 and Engineering and Physical Sciences Research Council (U.K.) grant number EP/F002998/1.

References

[1] V. Agrawal and S. Seshadri. Distribution free bounds for service constrained (q, r) inventory systems. *Naval Research Logistics*, 47(8):635–656, 2000.

[2] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, Cambridge, 2006.

[3] L. Chu, J. Shanthikumar, and Z. Shen. Solving operational statistics via a Bayesian analysis. *Oper. Res. Letters*, 36:110–116, 2008.

[4] G. Hardy, J. Littlewood, and G. Polya. *Inequalities*. Cambridge University Press, Cambridge, 1967.

[5] Yuri Kalnishkan and Michael V. Vyugin. The weak aggregating algorithm and weak mixability. *Journal of Computer and System Sciences*, 74:1228–1244, 2008. Special Issue devoted to COLT 2005.

[6] R. Levi, R. Roundy, and D. Shmoys. Provably near-optimal sampling-based policies for stochastic inventory control models. *Math. of Oper. Res.*, 32(4):821–839, 2007.

[7] L. Liyanage and J. Shanthikumar. A practical inventory control policy using operational statistics. *Oper. Res. Letters*, 33:341–348, 2005.

[8] X. Lu, J. Song, and K. Zhu. Analysis of perishable-inventory systems with censored demand data. *Oper. Res.*, 56(4):1034–1038, 2008.

[9] R. Reiss. *Approximate Distributions of Order Statistics*. Springer, New York, 1989.

[10] H. Scarf. A min-max solution of an inventory problem. In K. Arrow, S. Karlin, and H. Scarf, editors, *Studies in the Mathematical Theory of Inventory and Production*, pages 201–209. Stanford University Press, 1958.

[11] W. Tim Huh and P. Rusmevichientong. A nonparametric asymptotic analysis of inventory planning with censored demand. *Math. of Oper. Res.*, 34(1):103–123, 2009.

[12] V. Vovk. Competitive on-line statistics. *International Statistical Review*, 69:213–248, 2001.