Columbia B9136 April 8, 2025

Lecture 9 — 7 Proofs of Data-driven Newsvendor, part 2

Lecturer: Will Ma *Scribe: Xiaoyu Fan*

In last class, we study the data-driven newsvendor problem, where the regret is bounded as

$$\begin{aligned}\mathbb{E}\left[L(A)-L(a^*)\right] &= \mathbb{E}\left[\underbrace{L(A)-\hat{L}(A)}_{\text{Generalization Error}}\right]+\mathbb{E}\left[\hat{L}(A)-\hat{L}(a^*)\right]\\ &\le \mathbb{E}\left[\underbrace{\sup_a |L(a)-\hat{L}(a)|}_{\text{Uniform Error}}\right]+\mathbb{E}\left[\underbrace{\hat{L}(A)-\hat{L}(\hat{a})}_{\text{In-Sample Suboptimality}}\right].\end{aligned}$$

We summarize the 7 different proofs of the data-driven newsvendor problem in Table 1 according to the algorithm, the analysis and whether it can be generalized to other settings.

Proofs of Data-Driven Newsvendor	**Algorithm A**	**Analysis**	**Generality**
(1) Direct	$\hat{a}$	Specialized to Newsvendor	
(2) Finite Hypothesis Class +Covering Number	$\hat{a}$ rounded to multiple of $\frac{1}{N}$	Uniform Error + In-Sample Error	Worked for convex functions $O_{\alpha,B}(\frac{1}{\sqrt{N}})^{(*)}$
(3) VC Theory	$\hat{a}$	Uniform Error	
(4) Regularized Stability	$\text{argmin}_a \hat{L}(a)+\frac{\lambda a^2}{2}$	Generalization Error +In-Sample Error	
(5) On-Average Stability	$\hat{a}$	Generalization Error	Specialized
(6) Direct 2	$\hat{a}$	Specialized to Newsvendor	
(7) Stochastic Gradient Descent	average of iterates	Worked for convex functions	

Table 1: 7 Proofs of Data-Driven Newsvendor

$(*)$: Proofs $(2),(3),(4)$ work for any convex function and can achieve the same regret bound $O_{\alpha,B}(\frac{1}{\sqrt{N}})$. More precisely, $l(\cdot,\xi)$ needs to be convex and α-Lipschitz over a bounded interval $[0,B]$, given any ξ. Specially, in the newsvendor setting, $l(\cdot,\xi)$ is convex and $\max\{b,h\}$-Lipschitz (see Figure 1 for an illustration).

1 Remaining Proofs of the Data-Driven Newsvendor Problem

In this class, we allow unbounded demand (with CDF F) for the newsvendor problem. To achieve any meaningful regret bound, we require $\mathbb{E}_{\xi\sim F}[\xi]\le 1$. Given a fixed quantile q, since $\mathbb{P}[\xi\ge a^*]\ge 1-q$, we have

$$(1-q)a^*\le \mathbb{P}[\xi\ge a^*]\cdot a^*\le \mathbb{E}_{\xi\sim F}[\xi]\le 1 \quad\Rightarrow\quad a^*\le\frac{1}{1-q}.$$

Therefore, this condition immediately implies an upper bound on the decision space: $B=\frac{1}{1-q}$ (which is also illustrated in Figure 2).

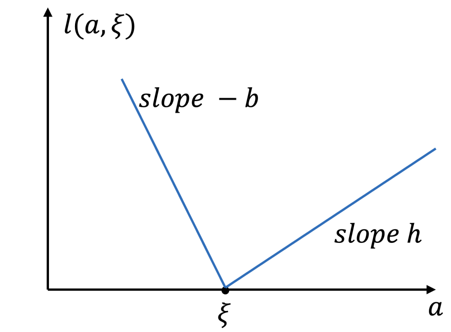

Figure 1: Illustration of $l(\cdot, \xi)$ being $\max\{b, h\}$-Lipschitz

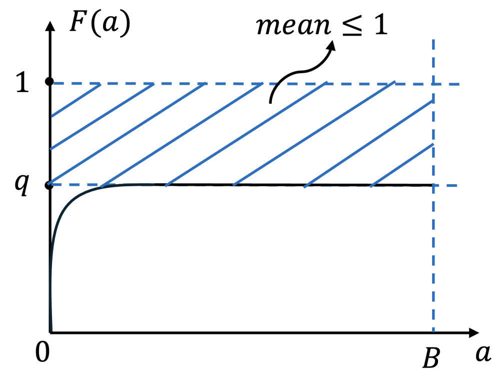

Figure 2: Illustration of $B = \frac{1}{1-q}$

1.1 Proof (6): Direct Proof 2

Recall that from the analysis in last class, we have

$$L(A) - L(a^*) = \int_{a^*}^{A} (F(\xi) - q) d\xi.$$

Set $A = \hat{a}$, and we now further take expectation over $\hat{a}$:

$$\begin{aligned} L(\hat{a}) - L(a^*) &= \int_0^{a^*} (q - F(\xi)) \mathbb{1}\{\xi \geq \hat{a}\} d\xi + \int_{a^*}^{B} (F(\xi) - q) \mathbb{1}\{\xi < \hat{a}\} d\xi \\ \Rightarrow \quad \mathbb{E}\left[L(\hat{a}) - L(a^*)\right] &= \int_0^{a^*} (q - F(\xi)) \mathbb{P}[\xi \geq \hat{a}] d\xi + \int_{a^*}^{B} (F(\xi) - q) \mathbb{P}[\xi < \hat{a}] d\xi. \end{aligned} \tag{1}$$

Through Figure 3, we can see that $\mathbb{P}[\xi \geq \hat{a}] = \mathbb{P}[\hat{F}(\xi) \geq q]$ and similarly $\mathbb{P}[\xi < \hat{a}] = \mathbb{P}[\hat{F}(\xi) < q]$. Applying Hoeffding's inequality, we have

$$\mathbb{P}[\hat{F}(\xi) \geq q] = \mathbb{P}\left[\hat{F}(\xi) - F(\xi) \geq q - F(\xi)\right] \leq exp\left(-2N(q - F(\xi))^2\right) \quad \forall \xi < a^*$$
$$\mathbb{P}[\hat{F}(\xi) < q] = \mathbb{P}\left[\hat{F}(\xi) - F(\xi) < q - F(\xi)\right] \leq exp\left(-2N(q - F(\xi))^2\right) \quad \forall \xi \geq a^*.$$

Taking this back to Equation (1) then implies

$$\begin{aligned}\mathbb{E}\left[L(\hat{a}) - L(a^*)\right] &\leq \int_0^B |q - F(\xi)| exp\left(-2N(q - F(\xi))^2\right) d\xi \\ &\leq B \cdot \frac{1}{2\sqrt{eN}} = \frac{1}{2(1-q)\sqrt{eN}}\end{aligned}$$

where in the second inequality we use the fact that $g(x) = xexp(-2Nx^2)$ is maximized at $x = \frac{1}{2\sqrt{N}}$ with maximum value $\frac{1}{2\sqrt{eN}}$.

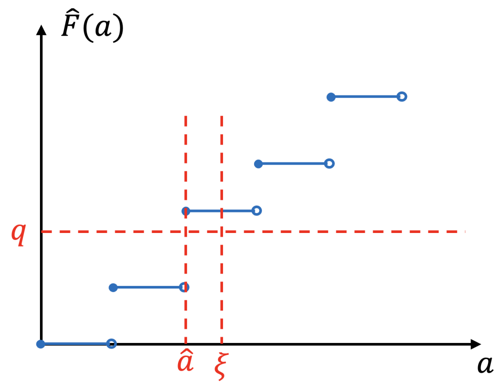

Figure 3: Illustration of $\mathbb{P}[\xi \geq \hat{a}] = \mathbb{P}[\hat{F}(\xi) \geq q]$ when $N = 4$

1.2 Proof (7): Stochastic Gradient Descent (SGD)

The SGD algorithm can be applied to the data-driven newsvendor problem as follows. It processes one sample at a time, and the final solution is the average over all the iterates.

- Select $a_1 \in [0, B]$ arbitrarily;
- Define $a_{i+1} = Proj_{[0,B]}\left(a_i - \eta_i \nabla_1 l(a_i, \xi_i)\right)$ for all $i = 1, \ldots, N$;
- Set $A = \frac{1}{N}\sum_{i=1}^{N} a_i$.

In the definition of a_{i+1}, $Proj_{[0,B]}(\cdot)$ is the projection to the interval $[0, B]$, which is formally defined as

$$Proj_{[0,B]}(a) = \begin{cases} a & if\ a \in [0, B] \\ 0 & if\ a < 0 \\ B & if\ a > B. \end{cases}$$

As illustrated by Figure 4, $\nabla_1 l(a, \xi)$, the subgradient of $l(\cdot, \xi)$ at a, can be defined as

$$\nabla_1 l(a, \xi) = -b\mathbb{1}\{a \leq \xi\} + h\mathbb{1}\{a > \xi\}.$$

Since $l(\cdot, \xi)$ is convex, it satisfies the following relation

$$(\tilde{a} - a) \cdot \nabla_1 l(a, \xi) \leq l(\tilde{a}, \xi) - l(a, \xi) \quad \forall \tilde{a}, a. \tag{2}$$

We set the step size $\eta_i = \frac{B}{\alpha\sqrt{i}}$, which is needed to achieve the optimal regret bound. Note that

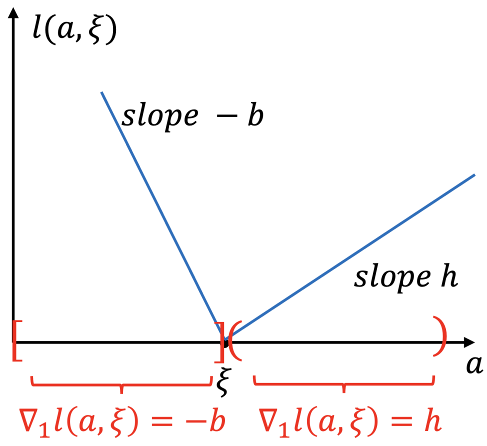

Figure 4: Illustration of $\nabla_1 l(a, \xi)$

this step size is anytime, meaning that it does not require prior knowledge of N.

Recall that our goal is to show

$$\mathbb{E}\left[L(A) - L(a^*)\right] = \mathbb{E}\left[L(A) - \hat{L}(a^*)\right] \leq \mathbb{E}\left[L(A) - \hat{L}(\hat{a})\right] = O(\frac{1}{\sqrt{N}}).$$

We first argue that Claim 1 suffices. It is straightforward to see $\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} l(\hat{a}, \xi_i)\right] = \mathbb{E}[\hat{L}(\hat{a})]$. On the other hand, $\mathbb{E}[L(A)] \leq \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} l(a_i, \xi_i)\right]$ since

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} l(a_i, \xi_i)\right] &= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[l(a_i, \xi_i)\right] \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[\mathbb{E}\left[l(a_i, \xi_i)\middle|\xi_1, \ldots, \xi_{i-1}\right]\right] \\
&\overset{(**)}{=} \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[L(a_i)\right] \\
&\overset{(***)}{\geq} \mathbb{E}\left[L(\frac{1}{N}\sum_{i=1}^{N} a_i)\right] = \mathbb{E}\left[L(A)\right].
\end{aligned}$$

Note that the critical step $(**)$ holds because conditioning on $\xi_1, \ldots, \xi_{i-1}$, a_i is deterministic and ξ_i is independent of $\xi_1, \ldots, \xi_{i-1}$. Finally, in $(***)$ we use Jensen's inequality and the fact that $L(\cdot)$ is convex.

Claim 1. $\frac{1}{N}\sum_{i=1}^{N} l(a_i, \xi_i) - \frac{1}{N}\sum_{i=1}^{N} l(\hat{a}, \xi_i) \leq \frac{3\alpha B}{2\sqrt{N}}$ *for any* $\boldsymbol{\xi} \in \mathbb{R}^N_{\geq 0}$.

Proof. According to the update rule of SGD,

$$\begin{aligned}
(a_{i+1} - \hat{a})^2 &= \left(Proj_{[0,B]}(a_i - \eta_i \nabla_1 l(a_i, \xi_i)) - \hat{a}\right)^2 \\
&\leq (a_i - \eta_i \nabla_1 l(a_i, \xi_i) - \hat{a})^2 \\
&\leq (a_i - \hat{a})^2 - 2\eta_i \nabla_1 l(a_i, \xi_i)(a_i - \hat{a}) + \eta_i^2 \alpha^2
\end{aligned}$$

where the first inequality is because projection is a contraction mapping ($|Proj_{[0,B]}(a) - Proj_{[0,B]}(\tilde{a})| \leq |a - \tilde{a}|$ for any $a, \tilde{a}$), and in the second inequality we use $|\nabla_1 l(\cdot,\cdot)| \leq \alpha$. Rearranging the terms in the above equation then gives

$$2\nabla_1 l(a_i, \xi_i)(a_i - \hat{a}) \leq \frac{(a_i - \hat{a})^2 - (a_{i+1} - \hat{a})^2}{\eta_i} + \eta_i \alpha^2. \tag{3}$$

By Equation (2),

$$l(a_i, \xi_i) - l(\hat{a}, \xi_i) \leq \nabla_1 l(a_i, \xi_i)(a_i - \hat{a}). \tag{4}$$

Combining Equation (3) with Equation (4), we have

$$2\left(l(a_i, \xi_i) - l(\hat{a}, \xi_i)\right) \leq \frac{(a_i - \hat{a})^2 - (a_{i+1} - \hat{a})^2}{\eta_i} + \eta_i \alpha^2.$$

Summing i from 1 to N on both sides,

$$\begin{aligned} 2\left(\sum_{i=1}^N l(a_i, \xi_i) - \sum_{i=1}^N l(\hat{a}, \xi_i)\right) &\leq \sum_{i=1}^N \frac{(a_i - \hat{a})^2 - (a_{i+1} - \hat{a})^2}{\eta_i} + \sum_{i=1}^N \eta_i \alpha^2 \\ &= \sum_{i=1}^N (a_i - \hat{a})^2 \left(\frac{1}{\eta_i} - \frac{1}{\eta_{i-1}}\right) - \frac{(a_{N+1} - \hat{a})^2}{\eta_N} + \sum_{i=1}^N \eta_i \alpha^2 \quad \left(\frac{1}{\eta_0} := 0\right) \\ &\leq B^2 \sum_{i=1}^N \left(\frac{1}{\eta_i} - \frac{1}{\eta_{i-1}}\right) + \sum_{i=1}^N \eta_i \alpha^2 \\ &= B^2 \frac{1}{\eta_N} + \sum_{i=1}^N \eta_i \alpha^2 = B\alpha\sqrt{N} + B\alpha \sum_{i=1}^N \frac{1}{\sqrt{i}} \leq 3B\alpha\sqrt{N}. \end{aligned}$$

□

2 Statistical Lower Bound

We show that the regret upper bound $\mathbb{E}[L(A) - L(a^*)] \leq O(\frac{1}{\sqrt{N}})$ is tight by providing a matching lower bound:

Theorem 2.

$$\inf_{A:[0,1]^N \to [0,1]} \sup_{F \in \Delta([0,1])} \mathbb{E}_{\boldsymbol{\xi} \sim F^N} \left[L_F(A(\boldsymbol{\xi})) - L_F(a_F^*)\right] = \Omega(\frac{1}{\sqrt{N}}).$$

Note that here we go back to consider bounded demand distributions. This is because to prove the lower bound, it suffices to show for any algorithm A, we can construct a distribution $F \in \Delta([0,1])$ such that $\mathbb{E}_{\boldsymbol{\xi} \sim F^N} \left[L_F(A(\boldsymbol{\xi})) - L_F(a_F^*)\right] \geq \frac{C}{\sqrt{N}}$. Also, we do not need to prove the result for randomized algorithms $A : [0,1]^N \to \Delta([0,1])$. This is because since the cost function is convex, no randomized algorithm can do better than its mean:

$$L_F\left(\mathbb{E}_{a \sim A(\boldsymbol{\xi})}[a]\right) \leq \mathbb{E}_{a \sim A(\boldsymbol{\xi})}\left[L_F(a)\right].$$

(Intuitively, it is also because the adversary can only select the distribution F, but does not have control over the samples observed. Therefore, the algorithm does not need to be randomized to trick the adversary.)

2.1 The Total Variational Distance

We first introduce the total variational distance between two distributions P and Q, a key concept for understanding the statistical learning limit of a problem:

$$TV(P,Q) = \frac{1}{2}\sum_{\xi} |p(\xi) - q(\xi)|.$$

Here we consider discrete distributions P, Q, and $p(\cdot), q(\cdot)$ are the corresponding probability mass functions. (In general, we can define $TV(P,Q) = \frac{1}{2}\int |\frac{dP}{d\lambda} - \frac{dQ}{d\lambda}| d\lambda$ for some measure λ such that $Q, P \ll \lambda$). Intuitively, $TV(P,Q)$ reflects how different the two distributions are, and the larger $TV(P,Q)$ is, the easier to differentiate between P and Q. To have a better understanding of this concept, we consider the following example.

Suppose $P = Ber(\frac{1}{3})$ and $Q = Ber(\frac{2}{3})$. Given samples iid drawn from one of the distributions, we wish to tell which distribution it is. We make our decisions under the Bayesian framework, where we assume as a prior that each distribution happens with equal probability. In this case, given one sample ξ_1, if $\xi_1 = 0$ then we would guess it is drawn from P, and if $\xi_1 = 1$ then we would guess it is drawn from Q. The probability for us to be correct is $\mathbb{P}_{\xi_1 \sim P}[\xi_1 = 0] = \mathbb{P}_{\xi_1 \sim Q}[\xi_1 = 1] = \frac{2}{3}$. If we are given two samples ξ_1, ξ_2, then we would guess P if $\xi_1 = \xi_2 = 0$, guess Q if $\xi_1 = \xi_2 = 1$, and make a random guess if $\xi_1 \neq \xi_2$. The probability for us to be correct is $\mathbb{P}_{\xi_1,\xi_2 \sim P}[\xi_1 = \xi_2 = 0] + \frac{1}{2}\mathbb{P}_{\xi_1,\xi_2 \sim P}[\xi_1 \neq \xi_2] = \frac{2}{3}$ (a similar probability can be computed when the underlying environment is Q). Therefore, observing the second sample does not increase our probability of being correct.

The above example can be explained through the total variational distance. As illustrated by Figure 5 and Figure 6, we can compute $TV(P,Q) = TV(P^2, Q^2) = \frac{1}{3}$. Therefore, the information we can obtain from two samples is the same to that of one sample. More precisely, Figure 6 shows that the event $\xi_1 \neq \xi_2$ (the colored blocks) happens with the exact same probability and does not provide any information about the underlying distribution.

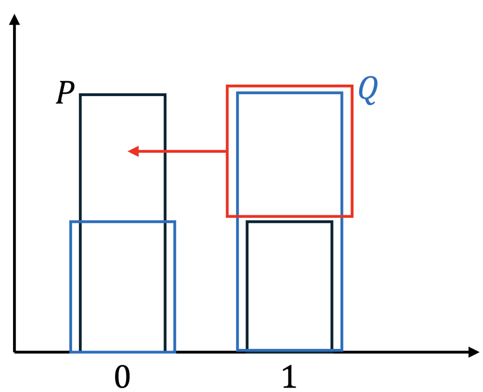

Figure 5: Illustration of $TV(P,Q)$

2.2 Proof of Theorem 2

We use the following claim to prove the lower bound.

Claim 3. *For any algorithm $A(\cdot)$, there exists an event E of measure $1 - TV(P^N, Q^N)$ under both P and Q (which can be considered as the event for $A(\cdot)$ to make an error). Then, under both P*

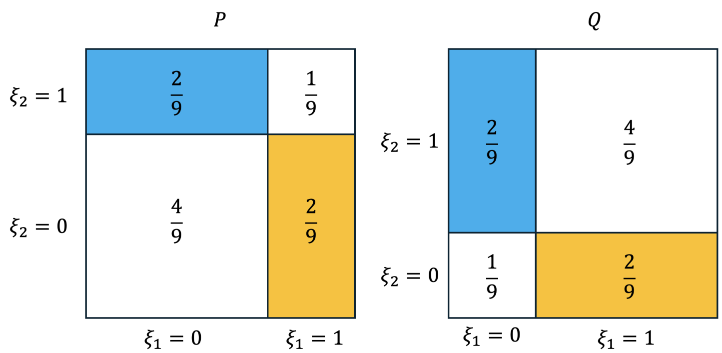

Figure 6: Illustration of $TV(P^2, Q^2)$

and Q, we have

$$\mathbb{E}\left[L(A(\boldsymbol{\xi})) - L(a^*)\right] \geq \mathbb{P}\left[E\right]\mathbb{E}\left[L(A(\boldsymbol{\xi})) - L(a^*)\Big|E\right],$$

where the distribution of $A(\boldsymbol{\xi})$ is the same under either P or Q when conditioning on E.

Proof. The proof sketch for Claim 3 is as follows. Fix any algorithm $A(\cdot)$, consider the sample path of its execution under world P, and the sample path of its execution under world Q respectively. The two worlds can be coupled in a way such that with probability $1 - TV(P^N, Q^N)$, the samples observed are identical under the two worlds. We set E to be the event of "identical observations", and E has measure $1 - TV(P^N, Q^N)$ under either of the world. Further, conditioning on E, the distribution of $A(\cdot)$ is the same under either P or Q. □

Example 4. Recall the example in Section 2.1, where $P = Ber(\frac{1}{3})$ and $Q = Ber(\frac{2}{3})$. Given two samples, E is the event where

- observation is $(1, 0)$ (which has probability mass $\frac{2}{9}$)
- observation is $(0, 1)$ (which has probability mass $\frac{2}{9}$)
- observation is $(0, 0)$ with probability mass $\frac{1}{9}$
- observation is $(1, 1)$ with probability mass $\frac{1}{9}$

Therefore, $\mathbb{P}_P\left[E\right] = \mathbb{P}_Q\left[E\right] = \frac{2}{3} = 1 - TV(P^2, Q^2)$.

For the lower bound instance, we set $B = 1$, $q = \frac{1}{2}$, $P = Ber(\frac{1}{2} - \epsilon)$ and $Q = Ber(\frac{1}{2} + \epsilon)$. Then it is straightforward to see $a_P^* = 0$ and

$$L_P(a) = (\frac{1}{2} - \epsilon)(1 - a) + (\frac{1}{2} + \epsilon)a = \frac{1}{2} - \epsilon + 2a\epsilon.$$

For any algorithm $A(\cdot)$ such that $\mathbb{E}_P\left[A(\boldsymbol{\xi})\Big|E\right] \geq \frac{1}{2}$, we have

$$\mathbb{E}_P\left[L_P(A(\boldsymbol{\xi})) - L_P(a_P^*)\Big|E\right] = L_P\left(\mathbb{E}_P\left[A(\boldsymbol{\xi})\Big|E\right]\right) - L_P(a_P^*) \geq \epsilon$$

where the equality is because $L_P(\cdot)$ is linear. Using Claim 3 we have

$$\begin{aligned}\mathbb{E}_P\left[L_P(A(\boldsymbol{\xi})) - L_P(a_P^*)\right] &\geq \mathbb{P}_P\left[E\right]\mathbb{E}_P\left[L_P(A(\boldsymbol{\xi})) - L_P(a_P^*)\Big|E\right] \\ &\geq \left(1 - TV(P^N, Q^N)\right)\epsilon.\end{aligned}$$

That is, in this case, the adversary should choose the environment as P. By a similar discussion, one can show that for any algorithm $A(\cdot)$ such that $\mathbb{E}_Q\left[A(\boldsymbol{\xi})\middle|E\right] < \frac{1}{2}$,

$$\mathbb{E}_Q\left[L_Q(A(\boldsymbol{\xi})) - L_Q(a_Q^*)\middle|E\right] = L_Q\left(\mathbb{E}_Q\left[A(\boldsymbol{\xi})\middle|E\right]\right) - L_Q(a_Q^*) \geq \epsilon$$

and using Claim 3

$$\begin{aligned}\mathbb{E}_Q\left[L_Q(A(\boldsymbol{\xi})) - L_Q(a_Q^*)\right] &\geq \mathbb{P}_Q\left[E\right]\mathbb{E}_Q\left[L_Q(A(\boldsymbol{\xi})) - L_Q(a_Q^*)\middle|E\right] \\ &\geq \left(1 - TV(P^N, Q^N)\right)\epsilon.\end{aligned}$$

In this case, the adversary chooses the environment as Q. Since $\mathbb{E}_Q\left[A(\boldsymbol{\xi})\middle|E\right] = \mathbb{E}_P\left[A(\boldsymbol{\xi})\middle|E\right]$, the events $\mathbb{E}_P\left[A(\boldsymbol{\xi})\middle|E\right] \geq \frac{1}{2}$ and $\mathbb{E}_Q\left[A(\boldsymbol{\xi})\middle|E\right] < \frac{1}{2}$ are complementary. Therefore, for any algorithm $A(\cdot)$, we derive

$$\begin{aligned}\sup_{F\in\Delta([0,1])}\mathbb{E}_F\left[L_F(A(\boldsymbol{\xi})) - L_F(a_F^*)\right] &\geq \sup_{F\in\{P,Q\}}\mathbb{E}_F\left[L_F(A(\boldsymbol{\xi})) - L_F(a_F^*)\right] \\ &\geq \left(1 - TV(P^N, Q^N)\right)\epsilon.\end{aligned} \tag{5}$$

In the rest of the proof, our goal is to tune ϵ based on N so that $TV(P^N, Q^N) \leq \frac{1}{2}$. Since the total variational distance can be hard to analyze, the standard technique is to upper bound it through some other divergence $D(\cdot,\cdot)$ with the nice property that $D(P^N, Q^N) \leq C \cdot \sqrt{N}D(P,Q)$. Possible candidates are the KL-divergence and the Hellinger distance. In class we choose the Hellinger distance, which is defined as:

$$H^2(P,Q) = \frac{1}{2}\sum_{\xi}\left(\sqrt{p(\xi)} - \sqrt{q(\xi)}\right)^2.$$

Through simple calculations, we have

$$\begin{aligned} H^2(P,Q) &= \frac{1}{2}\sum_{\xi}\left(p(\xi) + q(\xi) - 2\sqrt{p(\xi)q(\xi)}\right) \\ \Rightarrow \quad 1 - H^2(P,Q) &= \sum_{\xi}\sqrt{p(\xi)q(\xi)}.\end{aligned}$$

Therefore,

$$\begin{aligned}1 - H^2(P^N, Q^N) &= \sum_{\xi_1,\ldots,\xi_N}\sqrt{p(\xi_1)q(\xi_1)\cdots p(\xi_N)q(\xi_N)} \quad \text{(by independence)} \\ &= \sum_{\xi_1,\ldots,\xi_N}\sqrt{p(\xi_1)q(\xi_1)}\cdots\sqrt{p(\xi_N)q(\xi_N)} \\ &= \prod_{i=1}^{N}\left(\sum_{\xi_i}\sqrt{p(\xi_i)q(\xi_i)}\right) = \left(\sum_{\xi_1}\sqrt{p(\xi_1)q(\xi_1)}\right)^N \\ &= \left(1 - H^2(P,Q)\right)^N \geq 1 - NH^2(P,Q)\end{aligned}$$

which implies

$$H^2(P^N, Q^N) \leq NH^2(P,Q). \tag{6}$$

Next, we show the total variational distance can be upper bounded by the Hellinger distance:

$$
\begin{aligned}
TV(P,Q) &= \frac{1}{2}\sum_{\xi}|p(\xi)-q(\xi)| \\
&= \frac{1}{2}\sum_{\xi}|\sqrt{p(\xi)}-\sqrt{q(\xi)}||\sqrt{p(\xi)}+\sqrt{q(\xi)}| \\
&\le \frac{1}{2}\sqrt{\sum_{\xi}\left(\sqrt{p(\xi)}-\sqrt{q(\xi)}\right)^2}\sqrt{\sum_{\xi}\left(\sqrt{p(\xi)}+\sqrt{q(\xi)}\right)^2} \\
&= \frac{1}{\sqrt{2}}\sqrt{H^2(P,Q)}\sqrt{\sum_{\xi}\left(p(\xi)+q(\xi)+2\sqrt{p(\xi)q(\xi)}\right)} \\
&= \frac{1}{\sqrt{2}}\sqrt{H^2(P,Q)}\sqrt{2+2\sum_{\xi}\sqrt{p(\xi)q(\xi)}} \\
&\le \sqrt{2H^2(P,Q)},
\end{aligned}
$$

where the first inequality follows from the Cauchy-Schwartz inequality, and the second inequality uses the fact $\sum_{\xi}\sqrt{p(\xi)q(\xi)} = 1 - H^2(P,Q) \le 1$. This together with Equation (6) implies that

$$
TV(P^N,Q^N) \le \sqrt{2H^2(P^N,Q^N)} \le \sqrt{2NH^2(P,Q)}.
$$

Finally, we compute $H^2(P,Q)$ as

$$
\begin{aligned}
H^2(P,Q) &= H^2\left(Ber(\frac{1}{2}-\epsilon), Ber(\frac{1}{2}+\epsilon)\right) \\
&= 1 - 2\sqrt{(\frac{1}{2}-\epsilon)(\frac{1}{2}+\epsilon)} \\
&= 1 - \sqrt{1-4\epsilon^2} \\
&\le 2\epsilon^2 + 8\epsilon^4 \qquad (\sqrt{1-z} \ge 1 - \frac{z}{2} - \frac{z^2}{2}) \\
&\le 4\epsilon^2 \qquad (\epsilon \le \frac{1}{2}).
\end{aligned}
$$

Therefore, $TV(P^N,Q^N) \le \sqrt{8N\epsilon^2}$, and to set it less than $\frac{1}{2}$ we need $\epsilon = \frac{1}{\sqrt{32N}}$. Taking it back to Equation (5), we complete the proof:

$$
\sup_{F\in\Delta([0,1])} \mathbb{E}_F\left[L_F(A(\boldsymbol{\xi})) - L_F(a_F^*)\right] \ge \frac{\epsilon}{2} = \frac{1}{8\sqrt{2N}}.
$$

2.3 Remarks

The intuition for proving a statistical lower bound is to construct two distributions P, Q such that

- Indistinguishable: P and Q are close enough such that their total variational distance is small even after N observations, i.e. $TV(P^N,Q^N) \le \frac{1}{2}$;
- Separable: P and Q result in different optimal solutions such that the loss of misspecification is lower bounded by $\Omega(\frac{1}{\sqrt{N}})$.

One may immediately think of an alternative way of selecting P, Q such that the misspecification loss is $\Omega(1)$ (which corresponds to setting $\epsilon = \Theta(1)$ in our proof). However, this would result in an exponential growth in the total variational distance, and hence an invalid lower bound. Setting $\epsilon = \Theta(\frac{1}{\sqrt{N}})$ also makes sense intuitively, as the confidence interval of the mean estimator for a Bernoulli variable is of width $O(\frac{1}{\sqrt{N}})$ given N samples. Also, it can be seen from the proof that the constant in front of ϵ is critical for controlling the total variational distance.

So far we have focused on the expected regret: $\mathbb{E}\left[L(A) - L(a^*)\right] = \Theta(\frac{1}{\sqrt{N}})$. Another form of performance guarantee is the probably approximately correct (PAC) guarantee, which upper bounds the probability of $\mathbb{P}\left[L(A) - L(a^*) \geq \epsilon\right]$. A similar term is called the PAC sample complexity, which asks the question that given $\epsilon, \delta > 0$, how many samples are needed for $\mathbb{P}\left[L(A) - L(a^*) \leq \epsilon\right] \geq 1 - \delta$ to hold (the sample size is usually in the order of $O(\frac{1}{\epsilon^2} log \frac{1}{\delta})$). Some of the proofs we covered in this class (e.g., Proof (1)) can be extended to achieve the PAC guarantee.

3 Strongly Convex Newsvendor Problem

We revisit Proof (7) for the SGD algorithm under the stronger assumption that $F \in \Delta([0,1])$ is continuous with PDF f satisfying $f(\xi) \geq \gamma$ for all $\xi \in [0,1]$. In this case, the expected cost function $L(\cdot)$ is γ-strongly convex:

$$L(a) = \int_a^{a^*} (q - F(\xi)) d\xi \quad \Rightarrow \quad L'(a) = F(a) - q \quad \Rightarrow \quad L''(a) = f(a) \geq \gamma.$$

We show that a faster convergence rate of $O(\frac{logN}{N})$ can be obtained in this case.

By a similar discussion to Equation (3) (where we replace $\hat{a}$ with a^*),

$$2\nabla_1 l(a_i, \xi_i)(a_i - a^*) \leq \frac{(a_i - a^*)^2 - (a_{i+1} - a^*)^2}{\eta_i} + \eta_i \alpha^2.$$

Taking expectations on both sides,

$$2\mathbb{E}[L'(a_i)(a_i - a^*)] \leq \frac{\mathbb{E}[(a_i - a^*)^2] - \mathbb{E}[(a_{i+1} - a^*)^2]}{\eta_i} + \eta_i \alpha^2 \tag{7}$$

where we use the fact that conditioning on $\xi_1, \ldots, \xi_{i-1}$, a_i is deterministic and ξ_i is independent of $\xi_1, \ldots, \xi_{i-1}$, and hence

$$\begin{aligned} \mathbb{E}[\nabla_1 l(a_i, \xi_i)(a_i - a^*)] &= \mathbb{E}\left[\mathbb{E}\left[\nabla_1 l(a_i, \xi_i) \middle| \xi_1, \ldots, \xi_{i-1}\right](a_i - a^*)\right] \\ &= \mathbb{E}\left[L'(a_i)(a_i - a^*)\right]. \end{aligned}$$

Since $L(\cdot)$ is γ-strongly convex,

$$\begin{aligned} & L(\tilde{a}) \geq L(a) + L'(a)(\tilde{a} - a) + \frac{\gamma}{2}(\tilde{a} - a)^2 \quad \forall \tilde{a}, a \\ \Rightarrow \quad & L'(a_i)(a_i - a^*) \geq L(a_i) - L(a^*) + \frac{\gamma}{2}(a_i - a^*)^2. \end{aligned}$$

Taking this back to Equation (7), we have

$$\begin{aligned} & 2\mathbb{E}[L(a_i)] - 2L(a^*) + \gamma \mathbb{E}[(a_i - a^*)^2] \leq \frac{\mathbb{E}[(a_i - a^*)^2] - \mathbb{E}[(a_{i+1} - a^*)^2]}{\eta_i} + \eta_i \alpha^2 \\ \Rightarrow \quad & 2\mathbb{E}[L(a_i)] - 2L(a^*) \leq \frac{\mathbb{E}[(a_i - a^*)^2] - \mathbb{E}[(a_{i+1} - a^*)^2]}{\eta_i} - \gamma \mathbb{E}[(a_i - a^*)^2] + \eta_i \alpha^2. \end{aligned}$$

Summing over i from 1 to N on both sides,

$$\begin{aligned}
2\sum_{i=1}^{N}\mathbb{E}[L(a_i)] - 2NL(a^*) &\le \sum_{i=1}^{N}\frac{\mathbb{E}[(a_i-a^*)^2]-\mathbb{E}[(a_{i+1}-a^*)^2]}{\eta_i} - \sum_{i=1}^{N}\gamma\mathbb{E}[(a_i-a^*)^2] + \sum_{i=1}^{N}\eta_i\alpha^2 \\
&= \sum_{i=1}^{N}\mathbb{E}[(a_i-a^*)^2]\left(\frac{1}{\eta_i}-\frac{1}{\eta_{i-1}}-\gamma\right) - \frac{\mathbb{E}[(a_{N+1}-a^*)^2]}{\eta_N} + \sum_{i=1}^{N}\eta_i\alpha^2 \\
&\le \sum_{i=1}^{N}\eta_i\alpha^2 \qquad (\text{by setting } \eta_i = \frac{1}{\gamma i}) \\
&\le \frac{\alpha^2(1+lnN)}{\gamma}.
\end{aligned}$$

Finally, by the convexity of $L(\cdot)$, we have

$$\mathbb{E}[L(A)] = \mathbb{E}\left[L\left(\frac{1}{N}\sum_{i=1}^{N}a_i\right)\right] \le \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}[L(a_i)] \le L(a^*) + \frac{\alpha^2(1+lnN)}{2\gamma N}.$$

Remark 5. The SGD algorithm works also for censored demand feedback, that is, given an ordering quantity $a \in [0, B]$, only the sales quantity $\min\{a, \xi\}$ is observable, instead of the real demand ξ. This is because when we choose the subgradient in the following form:

$$\nabla_1 l(a, \xi) = -b\mathbb{1}\{a \le \xi\} + h\mathbb{1}\{a > \xi\},$$

it can be recovered from the censored demand. More precisely, we observe that $\mathbb{1}\{a \le \xi\} = \mathbb{1}\{\min\{a, \xi\} = a\}$ and $\mathbb{1}\{a > \xi\} = \mathbb{1}\{\min\{a, \xi\} < a\}$. Note that this nice property no longer holds if we define $\nabla_1 l(a, \xi) = -b\mathbb{1}\{a < \xi\} + h\mathbb{1}\{a \ge \xi\}$, because the indicator function $\mathbb{1}\{a \ge \xi\}$ cannot be implied from $\min\{a, \xi\}$.

Bibliographical notes. The sixth proof of Data-driven Newsvendor comes from Lin et al. (2022), although the version we present is simpler because we assume that we know an upper bound on the demand mean, and hence an upper bound of B on the decision. The seventh proof and the strongly convex variant follow from a standard SGD analysis (see Bubeck et al., 2015). The lower bound can be proved in different ways but the arguments used here mirror Chen and Ma (2024). An eighth proof of $O(1/\sqrt{N})$ regret for Data-driven Newsvendor can be derived from the exact analysis of Besbes and Mouchtaki (2023).

References

O. Besbes and O. Mouchtaki. How big should your data really be? data-driven newsvendor: Learning one sample at a time. *Management Science*, 69(10):5848–5865, 2023.

S. Bubeck et al. Convex optimization: Algorithms and complexity. *Foundations and Trends® in Machine Learning*, 8(3-4):231–357, 2015.

Z. Chen and W. Ma. Survey of data-driven newsvendor: Unified analysis and spectrum of achievable regrets. *arXiv preprint arXiv:2409.03505*, 2024.

M. Lin, W. T. Huh, H. Krishnan, and J. Uichanco. Data-driven newsvendor problem: Performance of the sample average approximation. *Operations Research*, 70(4):1996–2012, 2022.