Heavy Tails in a Robust Newsvendor Model

Bikramjit Das

SUTD

June 22, 2018

with A. Dhara (U Mich), K. Natarajan (SUTD)

Newsvendor model

Newsvendor model

- Decide on # of units to order before demand is observed.
- Unit purchase cost: c, unit revenue is $p > c > 0$.
- Unsold items have zero salvage value.

Newsvendor model

- Decide on $\#$ of units to order before demand is observed.
- Unit purchase cost: c, unit revenue is $p > c > 0$.
- Unsold items have zero salvage value.
- Demand $D \sim F$ for some non-negative d.f. F.

Newsvendor model

- Decide on # of units to order before demand is observed.
- Unit purchase cost: c, unit revenue is $p > c > 0$.
- Unsold items have zero salvage value.
- Demand $D \sim F$ for some non-negative d.f. F.
- Optimal order quantity is found by maximizing expected profit:

$$\max_{q \in \mathbb{R}_+} \quad (p\mathbb{E}_F[\min(q, D)] - cq)$$

Newsvendor model

- Decide on # of units to order before demand is observed.
- Unit purchase cost: c, unit revenue is $p > c > 0$.
- Unsold items have zero salvage value.
- Demand $D \sim F$ for some non-negative d.f. F.
- Optimal order quantity is found by maximizing expected profit:

$$\max_{q \in \mathbb{R}_+} \quad (p\mathbb{E}_F[\min(q, D)] - cq)$$

- Let $\alpha = 1 - c/p \in (0, 1)$ denote the *critical ratio*, then the solution is

$$q^* = F^{-1}(\alpha).$$

Distributionally Robust Newsvendor (DRN)

- F is unknown.

Distributionally Robust Newsvendor (DRN)

- F is unknown.
- DRN set up: F is ambiguous and assumed to lie in a set of possible distributions denoted by $\mathcal{F}$.

Distributionally Robust Newsvendor (DRN)

- F is unknown.
- DRN set up: F is ambiguous and assumed to lie in a set of possible distributions denoted by $\mathcal{F}$.
- Maximize worst-case expected profit.

$$\text{(DRN)} \quad \max_{q \in \mathbb{R}_+} \quad \inf_{F \in \mathcal{F}} \quad (p\mathbb{E}_F[\min(q, D)] - cq)$$

Distributionally Robust Newsvendor (DRN)

- F is unknown.
- DRN set up: F is ambiguous and assumed to lie in a set of possible distributions denoted by $\mathcal{F}$.
- Maximize worst-case expected profit.

$$\text{(DRN)} \quad \max_{q\in\mathbb{R}_+} \quad \inf_{F\in\mathcal{F}} \quad (p\mathbb{E}_F[\min(q,D)] - cq)$$

- Suppose the mean of demand be specified in $\mathcal{F}$. Then

$$\text{(DRN)} \quad \min_{q\in\mathbb{R}_+} \quad \sup_{F\in\mathcal{F}} \quad (\mathbb{E}_F[D-q]^+ + (1-\alpha)q)$$

Distributionally Robust Newsvendor (DRN)

- F is unknown.
- DRN set up: F is ambiguous and assumed to lie in a set of possible distributions denoted by $\mathcal{F}$.
- Maximize worst-case expected profit.

$$\text{(DRN)} \quad \max_{q \in \mathbb{R}_+} \quad \inf_{F \in \mathcal{F}} \quad (p\mathbb{E}_F[\min(q, D)] - cq)$$

- Suppose the mean of demand be specified in $\mathcal{F}$. Then

$$\text{(DRN)} \quad \min_{q \in \mathbb{R}_+} \quad \sup_{F \in \mathcal{F}} \quad (\mathbb{E}_F[D - q]^+ + (1 - \alpha)q)$$

- Quantity of interest for now: $\mathbb{E}_F[D - q]^+$.

Scarf's Model (1958)

Scarf's Model (1958)

- Set of demand distributions in Scarf's model (1958) is:

$$\mathcal{F}_{1,2} = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w \mathrm{d}F(w) = m_1, \int_0^\infty w^2 \mathrm{d}F(w) = m_2 \right\}$$

Scarf's Model (1958)

- Set of demand distributions in Scarf's model (1958) is:

$$\mathcal{F}_{1,2} = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w\mathrm{d}F(w) = m_1, \int_0^\infty w^2 \mathrm{d}F(w) = m_2 \right\}$$

- Given an order quantity $q > m_2/2m_1$, the worst-case demand distribution for $\sup_{F \in \mathcal{F}_{1,2}} \mathbb{E}_F[D - q]_+$ is two-point:

$$D_q^* = \begin{cases} q - \sqrt{q^2 - 2m_1 q + m_2}, & \text{w.p. } \frac{1}{2}\left(1 + \frac{q - m_1}{\sqrt{q^2 - 2m_1 q + m_2}}\right) \\ q + \sqrt{q^2 - 2m_1 q + m_2}, & \text{w.p. } \frac{1}{2}\left(1 - \frac{q - m_1}{\sqrt{q^2 - 2m_1 q + m_2}}\right) \end{cases}$$

Scarf's Model (1958)

- Set of demand distributions in Scarf's model (1958) is:

$$\mathcal{F}_{1,2} = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w\mathrm{d}F(w) = m_1, \int_0^\infty w^2 \mathrm{d}F(w) = m_2 \right\}$$

- Given an order quantity $q > m_2/2m_1$, the worst-case demand distribution for $\sup_{F \in \mathcal{F}_{1,2}} \mathbb{E}_F[D-q]_+$ is two-point:

$$D_q^* = \begin{cases} q - \sqrt{q^2 - 2m_1 q + m_2}, & \text{w.p. } \frac{1}{2}\left(1 + \frac{q-m_1}{\sqrt{q^2-2m_1q+m_2}}\right) \\ q + \sqrt{q^2 - 2m_1 q + m_2}, & \text{w.p. } \frac{1}{2}\left(1 - \frac{q-m_1}{\sqrt{q^2-2m_1q+m_2}}\right) \end{cases}$$

- Support points and probabilities depend on q.

Scarf's Model (1958)

- Set of demand distributions in Scarf's model (1958) is:

$$\mathcal{F}_{1,2} = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w\mathrm{d}F(w) = m_1, \int_0^\infty w^2 \mathrm{d}F(w) = m_2 \right\}$$

- Given an order quantity $q > m_2/2m_1$, the worst-case demand distribution for $\sup_{F \in \mathcal{F}_{1,2}} \mathbb{E}_F[D - q]_+$ is two-point:

$$D_q^* = \begin{cases} q - \sqrt{q^2 - 2m_1 q + m_2}, & \text{w.p. } \frac{1}{2}\left(1 + \frac{q - m_1}{\sqrt{q^2 - 2m_1 q + m_2}}\right) \\ q + \sqrt{q^2 - 2m_1 q + m_2}, & \text{w.p. } \frac{1}{2}\left(1 - \frac{q - m_1}{\sqrt{q^2 - 2m_1 q + m_2}}\right) \end{cases}$$

- Support points and probabilities depend on q.
- For $0 \le q \le m_2/2m_1$, the worst-case distribution is $D^*_{m_2/2m_1}$.

Scarf's Model (1958)

- The worst-case bound is:

$$\sup_{F \in \mathcal{F}_{1,2}} \mathbb{E}_F[D-q]_+ = \begin{cases} \frac{1}{2}\left(\sqrt{q^2 - 2m_1 q + m_2} - (q - m_1)\right), & \text{if } q > \frac{m_2}{2m_1} \\ m_1 - \frac{q m_1^2}{m_2}, & \text{if } 0 \leq q \leq \frac{m_2}{2m_1} \end{cases}$$

Scarf's Model (1958)

- The worst-case bound is:

$$\sup_{F \in \mathcal{F}_{1,2}} \mathbb{E}_F[D-q]_+ = \begin{cases} \frac{1}{2}\left(\sqrt{q^2 - 2m_1 q + m_2} - (q - m_1)\right), & \text{if } q > \frac{m_2}{2m_1} \\ m_1 - \frac{q m_1^2}{m_2}, & \text{if } 0 \le q \le \frac{m_2}{2m_1} \end{cases}$$

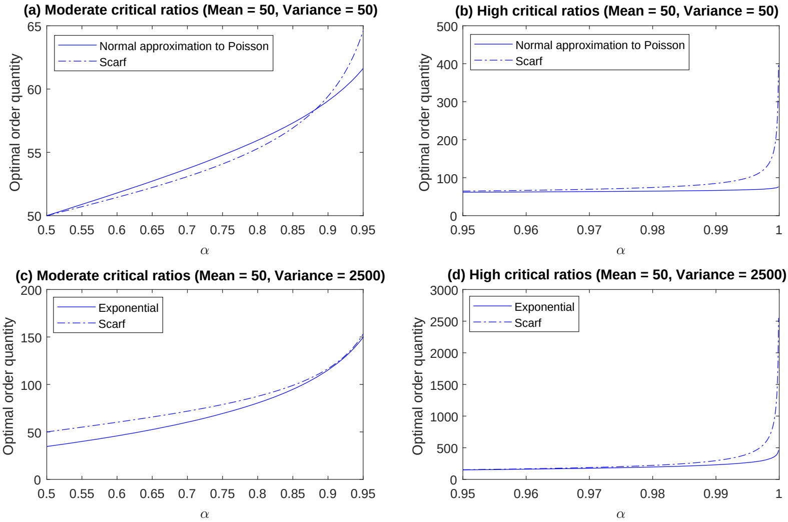
- The optimal order quantity has a closed form solution:

$$q^* = \begin{cases} m_1 + \frac{\sqrt{m_2 - m_1^2}}{2} \frac{2\alpha - 1}{\sqrt{\alpha(1-\alpha)}}, & \text{if } \frac{m_2 - m_1^2}{m_2} < \alpha < 1 \\ 0, & \text{if } 0 \le \alpha \le \frac{m_2 - m_1^2}{m_2} \end{cases}$$

Scarf's Model (1958): Observations

- Scarf (1958): A Poisson approximation works well for a wide range of values; albeit for very high critical ratios, the model prescribes higher order quantities than the Poisson distribution.
- Gallego and Moon (1993): Optimal order quantities from Scarf's model and optimal order quantity for normally distributed demands and concluded numerically that for a large range of critical ratios, the loss in profit is not significant.
- Wang, Glynn and Ye (2015): *"In the distributionally robust optimization approach, the worst-case distribution for a decision is often unrealistic. Scarf (1958) shows that the worst-case distribution in the newsvendor context is a two-point distribution. This raises the concern that the decision chosen by this approach is guarding under some overly conservative scenarios, while performing poorly in more likely scenarios. Unfortunately, these drawbacks seem to be inherent in the model choice and cannot be remedied easily."*

Scarf's Model



Distributionally Robust Newsvendor (DRN)

- Shapiro and Kleywegt (2002) and Shapiro and Ahmed (2004) show that DRN can be equivalently formulated as a classical newsvendor problem through the construction of a new probability demand distribution.

Distributionally Robust Newsvendor (DRN)

- Shapiro and Kleywegt (2002) and Shapiro and Ahmed (2004) show that DRN can be equivalently formulated as a classical newsvendor problem through the construction of a new probability demand distribution.
- Given a set $\mathcal{F}$, there exists a nonnegative random variable D^* with probability distribution F^* such that:

$$\sup_{F \in \mathcal{F}} \mathbb{E}_F[D - q]_+ = \mathbb{E}_{F^*}[D^* - q]_+, \ \forall q \in \mathbb{R}_+$$

Distributionally Robust Newsvendor (DRN)

- Shapiro and Kleywegt (2002) and Shapiro and Ahmed (2004) show that DRN can be equivalently formulated as a classical newsvendor problem through the construction of a new probability demand distribution.
- Given a set $\mathcal{F}$, there exists a nonnegative random variable D^* with probability distribution F^* such that:

$$\sup_{F \in \mathcal{F}} \mathbb{E}_F[D - q]_+ = \mathbb{E}_{F^*}[D^* - q]_+, \ \forall q \in \mathbb{R}_+$$

- F^* dominates $F \in \mathcal{F}$ in an increasing convex order sense Müller and Stoyan (2002)).

Distributionally Robust Newsvendor (DRN)

- Shapiro and Kleywegt (2002) and Shapiro and Ahmed (2004) show that DRN can be equivalently formulated as a classical newsvendor problem through the construction of a new probability demand distribution.
- Given a set $\mathcal{F}$, there exists a nonnegative random variable D^* with probability distribution F^* such that:

$$\sup_{F \in \mathcal{F}} \mathbb{E}_F[D - q]_+ = \mathbb{E}_{F^*}[D^* - q]_+, \ \forall q \in \mathbb{R}_+$$

- F^* dominates $F \in \mathcal{F}$ in an increasing convex order sense Müller and Stoyan (2002)).
- F^* may not have an explicit characterization and might not even lie in the original set of distributions $\mathcal{F}$.

Revisiting Scarf's Model

- The distribution F^* from Scarf's model

$$F^*(d) = \mathbb{P}(D^* \leq d) = \begin{cases} \frac{1}{2}\left(1 + \frac{d-m_1}{\sqrt{d^2-2m_1d+m_2}}\right), & \text{if } d > \frac{m_2}{2m_1} \\ 1 - \frac{m_1^2}{m_2}, & \text{if } 0 \leq d \leq \frac{m_2}{2m_1} \end{cases}$$

Revisiting Scarf's Model

- The distribution F^* from Scarf's model

$$F^*(d) = \mathbb{P}(D^* \leq d) = \begin{cases} \frac{1}{2}\left(1 + \frac{d - m_1}{\sqrt{d^2 - 2m_1 d + m_2}}\right), & \text{if } d > \frac{m_2}{2m_1} \\ 1 - \frac{m_1^2}{m_2}, & \text{if } 0 \leq d \leq \frac{m_2}{2m_1} \end{cases}$$

- F^* defines a censored student-t random variable with d.o.f 2:

$$D^* = \begin{cases} \tilde{t}_2\left(m_1, (m_2 - m_1^2)/2\right), & \text{if } \tilde{t}_2\left(m_1, (m_2 - m_1^2)/2\right) > \frac{m_2}{2m_1} \\ 0, & \text{otherwise} \end{cases}$$

Revisiting Scarf's Model

- The distribution F^* from Scarf's model

$$F^*(d) = \mathbb{P}(D^* \le d) = \begin{cases} \frac{1}{2}\left(1 + \frac{d-m_1}{\sqrt{d^2-2m_1d+m_2}}\right), & \text{if } d > \frac{m_2}{2m_1} \\ 1 - \frac{m_1^2}{m_2}, & \text{if } 0 \le d \le \frac{m_2}{2m_1} \end{cases}$$

- F^* defines a censored student-t random variable with d.o.f 2:

$$D^* = \begin{cases} \tilde{t}_2\left(m_1, (m_2 - m_1^2)/2\right), & \text{if } \tilde{t}_2\left(m_1, (m_2 - m_1^2)/2\right) > \frac{m_2}{2m_1} \\ 0, & \text{otherwise} \end{cases}$$

- This characterizes a "heavy tail optimality" property of Scarf's model.

Related DRN Models

Our focus is on sets where demand might take any value in $[0, \infty)$ and the high service level regime.

- Closed form solution: Ben-Tal and Hochman (1972, 1976) - mean, mean absolute deviation, Natarajan, Sim and Uichanco (2017) - mean, variance and semivariance.

Related DRN Models

Our focus is on sets where demand might take any value in $[0, \infty)$ and the high service level regime.

- Closed form solution: Ben-Tal and Hochman (1972, 1976) - mean, mean absolute deviation, Natarajan, Sim and Uichanco (2017) - mean, variance and semivariance.
- Bertsimas and Popescu (2002, 2005) and Lasserre (2002) - SDP techniques to compute the worst-case bounds given:

$$\mathcal{F}_{1,2,\dots,n} = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w^i \mathrm{d}F(w) = m_i,\ i = 1, 2, \dots, n \right\}$$

Related DRN Models

Our focus is on sets where demand might take any value in $[0, \infty)$ and the high service level regime.

- Closed form solution: Ben-Tal and Hochman (1972, 1976) - mean, mean absolute deviation, Natarajan, Sim and Uichanco (2017) - mean, variance and semivariance.
- Bertsimas and Popescu (2002, 2005) and Lasserre (2002) - SDP techniques to compute the worst-case bounds given:

$$\mathcal{F}_{1,2,\ldots,n} = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w^i \mathrm{d}F(w) = m_i, \ i = 1, 2, \ldots, n \right\}$$

- While some attempts has been made to solve this problem analytically for $n = 3$ and $n = 4$, the tight worst-case bounds have complicated expressions involving roots of cubic and quartic equations (see Jansen, Haezendonck, and Goovaerts (1986) Zuluaga, Peña and Du (2009)).

Related DRN Models

- Additional structural properties such as symmetry and unimodality have been added to the ambiguity sets with moments - SDP and SOCP methods (Popescu (2005), Van Parys, Goulart and Kuhn (2016), Li, Jiang and Mathieu (2017)).

Related DRN Models

- Additional structural properties such as symmetry and unimodality have been added to the ambiguity sets with moments - SDP and SOCP methods (Popescu (2005), Van Parys, Goulart and Kuhn (2016), Li, Jiang and Mathieu (2017)).
- Lam and Mottet (2017) analyzed the worst-case behavior with information on the tail probability of the random variable for a given threshold under shape constraints.

Related DRN Models

- Additional structural properties such as symmetry and unimodality have been added to the ambiguity sets with moments - SDP and SOCP methods (Popescu (2005), Van Parys, Goulart and Kuhn (2016), Li, Jiang and Mathieu (2017)).
- Lam and Mottet (2017) analyzed the worst-case behavior with information on the tail probability of the random variable for a given threshold under shape constraints.
- Blanchet and Murthy (2016) computed the worst-case tail probabilities where the ambiguity set is defined as the set of all distributions within a ball of distance δ around a reference distribution using Kullback-Leibler and Renyi divergence measures.

Related DRN Models

- Additional structural properties such as symmetry and unimodality have been added to the ambiguity sets with moments - SDP and SOCP methods (Popescu (2005), Van Parys, Goulart and Kuhn (2016), Li, Jiang and Mathieu (2017)).
- Lam and Mottet (2017) analyzed the worst-case behavior with information on the tail probability of the random variable for a given threshold under shape constraints.
- Blanchet and Murthy (2016) computed the worst-case tail probabilities where the ambiguity set is defined as the set of all distributions within a ball of distance δ around a reference distribution using Kullback-Leibler and Renyi divergence measures.
- Engelke and Ivanovs (2017) provide robust asymptotic bounds on tail probabilities in multivariate extremes.

Related DRN Models

- A related moment model that has been studied in option pricing but remains unexplored in the newsvendor model (see Grundy (1991)):

$$\mathcal{F}_n = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w^n \mathrm{d}F(w) = m_n \right\}$$

Related DRN Models

- A related moment model that has been studied in option pricing but remains unexplored in the newsvendor model (see Grundy (1991)):

$$\mathcal{F}_n = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w^n \mathrm{d}F(w) = m_n \right\}$$

- For $n \geq 1$, given a value $q > (n-1)m_n^{1/n}/n$, the worst-case distribution is two-point:

$$D_q^* = \begin{cases} \frac{(n-1)q}{n}, & \text{w.p. } \frac{(n-1)^n m_n}{n^n q^n} \\ 0, & \text{otherwise} \end{cases}$$

Related DRN Models

- A related moment model that has been studied in option pricing but remains unexplored in the newsvendor model (see Grundy (1991)):

$$\mathcal{F}_n = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w^n \mathrm{d}F(w) = m_n \right\}$$

- For $n \geq 1$, given a value $q > (n-1)m_n^{1/n}/n$, the worst-case distribution is two-point:

$$D_q^* = \begin{cases} \frac{(n-1)q}{n}, & \text{w.p. } \frac{(n-1)^n m_n}{n^n q^n} \\ 0, & \text{otherwise} \end{cases}$$

- For $0 \leq q \leq (n-1)m_n^{1/n}/n$, the worst-case demand distribution is degenerate with the mass at the point $m_n^{1/n}$.

Related DRN Models

- A related moment model that has been studied in option pricing but remains unexplored in the newsvendor model (see Grundy (1991)):

$$\mathcal{F}_n = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w^n \mathrm{d}F(w) = m_n \right\}$$

- For $n \geq 1$, given a value $q > (n-1)m_n^{1/n}/n$, the worst-case distribution is two-point:

$$D_q^* = \begin{cases} \frac{(n-1)q}{n}, & \text{w.p. } \frac{(n-1)^n m_n}{n^n q^n} \\ 0, & \text{otherwise} \end{cases}$$

- For $0 \leq q \leq (n-1)m_n^{1/n}/n$, the worst-case demand distribution is degenerate with the mass at the point $m_n^{1/n}$.
- The distribution F^* in this model defines a Pareto random variable:

$$D^* = \text{Pareto}\left(\frac{(n-1)m_n^{1/n}}{n}, n \right)$$

Model

- Consider an ambiguity set defined by a known value of the first and the nth moment:

$$\mathcal{F}_{1,n} = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w \mathrm{d}F(w) = m_1, \int_0^\infty w^n \mathrm{d}F(w) = m_n \right\}$$

Model

- Consider an ambiguity set defined by a known value of the first and the nth moment:

$$\mathcal{F}_{1,n} = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w \mathrm{d}F(w) = m_1, \int_0^\infty w^n \mathrm{d}F(w) = m_n \right\}$$

- We allow for any real number $n > 1$ (not necessarily an integer).

Model

- Consider an ambiguity set defined by a known value of the first and the nth moment:

$$\mathcal{F}_{1,n} = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w \mathrm{d}F(w) = m_1, \int_0^\infty w^n \mathrm{d}F(w) = m_n \right\}$$

- We allow for any real number $n > 1$ (not necessarily an integer).
- Scarf's model: $n = 2$.

Model

- Consider an ambiguity set defined by a known value of the first and the nth moment:

$$\mathcal{F}_{1,n} = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w\mathrm{d}F(w) = m_1, \int_0^\infty w^n \mathrm{d}F(w) = m_n \right\}$$

- We allow for any real number $n > 1$ (not necessarily an integer).
- Scarf's model: $n = 2$.
- To the best of our knowledge, no closed form solution for $\sup_{F \in \mathcal{F}_{1,n}} \mathbb{E}_F[D - q]_+$ is currently known for general values of n and seems to be hard to find.

Model

- Consider an ambiguity set defined by a known value of the first and the nth moment:

$$\mathcal{F}_{1,n} = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w\mathrm{d}F(w) = m_1, \int_0^\infty w^n \mathrm{d}F(w) = m_n \right\}$$

- We allow for any real number $n > 1$ (not necessarily an integer).
- Scarf's model: $n = 2$.
- To the best of our knowledge, no closed form solution for $\sup_{F \in \mathcal{F}_{1,n}} \mathbb{E}_F[D - q]_+$ is currently known for general values of n and seems to be hard to find.
- For rational values of n, the DRN problem under this ambiguity set can be solved as a SDP.

Model

- Consider an ambiguity set defined by a known value of the first and the nth moment:

$$\mathcal{F}_{1,n} = \left\{ F \in \mathbb{M}(\mathbb{R}_+) : \int_0^\infty \mathrm{d}F(w) = 1, \int_0^\infty w \mathrm{d}F(w) = m_1, \int_0^\infty w^n \mathrm{d}F(w) = m_n \right\}$$

- We allow for any real number $n > 1$ (not necessarily an integer).
- Scarf's model: $n = 2$.
- To the best of our knowledge, no closed form solution for $\sup_{F \in \mathcal{F}_{1,n}} \mathbb{E}_F[D - q]_+$ is currently known for general values of n and seems to be hard to find.
- For rational values of n, the DRN problem under this ambiguity set can be solved as a SDP.
- We develop lower and upper bounds that are approximately optimal for large values of q, which helps us characterize F^*.

Main Results

Proposition

Given an ambiguity set $\mathcal{F}_{1,n}$, there exists a positive quantity $\underline{q}(m_1, m_n, n)$ which depends on the parameters m_1, m_n and n, such that:

$$\sup_{F \in \mathcal{F}_{1,n}} \mathbb{E}_F[D - q]_+ \quad \geq \quad \frac{(m_n - m_1^n)}{n^n q^{n-1}}(n-1)^{n-1}, \quad \forall q \geq \underline{q}(m_1, m_n, n).$$

Main Results

Proposition

Consider the ambiguity set $\mathcal{F}_{1,n}$.

(a) *Given $n \in (2, \infty)$, there exists a positive quantity $\overline{q}(m_1, m_n, n)$ which depends on the parameters m_1, m_n and n, such that:*

$$\sup_{F \in \mathcal{F}_{1,n}} \mathbb{E}_F[D - q]_+ \leq \frac{(m_n - m_1^n)}{n^n q^{n-1} - n^2 m_1^{n-1}(n-1)^{n-1}}(n-1)^{n-1} \quad \forall q \geq \overline{q}(m_1, m_n, n).$$

(b) *Given $n \in (1, 2)$, for any $\epsilon > 0$, there exists a positive quantity $\overline{q}(m_1, m_n, n, \epsilon)$ which depends on the parameters m_1, m_n, n and ϵ such that:*

$$\sup_{F \in \mathcal{F}_{1,n}} \mathbb{E}_F[D - q]_+ \leq \frac{(m_n - m_1^n)}{n^n q^{n-1} - n^2 m_1^{n-1}(n-1)^{n-1-\epsilon}}(n-1)^{n-1} \quad \forall q \geq \overline{q}(m_1, m_n, n, \epsilon).$$

Outline of Proof: Lower Bound

- Primal feasible - Consider a three point random variable D with a distribution defined as follows:

$$D = \begin{cases} \frac{qn}{n-1}, & \text{w.p. } \frac{(m_n - m_1^n)}{n^n q^n}(n-1)^n \\ d, & \text{w.p. } p \\ 0, & \text{w.p. } 1 - p - \frac{(m_n - m_1^n)}{n^n q^n}(n-1)^n \end{cases}$$

- Find the values d and p as a function of q such that moment constraints for m_1 and m_n are met.
- Show that for large values of q, the value d is less than q and the corresponding p defines a valid probability measure.

Outline of Proof: Upper Bound

- Dual feasible - The dual formulation is:

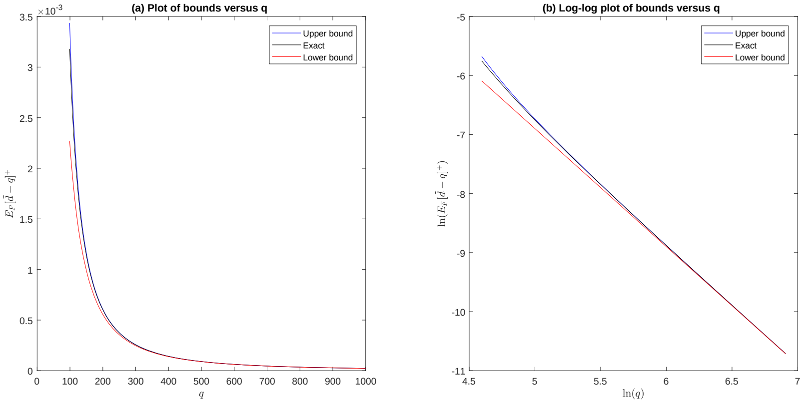
$$\begin{array}{lll} \inf & y_0 + y_1 m_1 + y_n m_n & \\ \text{s.t.} & y_0 + y_1 d + y_n d^n \geq 0, & \forall d \geq 0, \\ & y_0 + y_1 d + y_n d^n \geq d - q, & \forall d \geq 0, \end{array}$$

- Set:

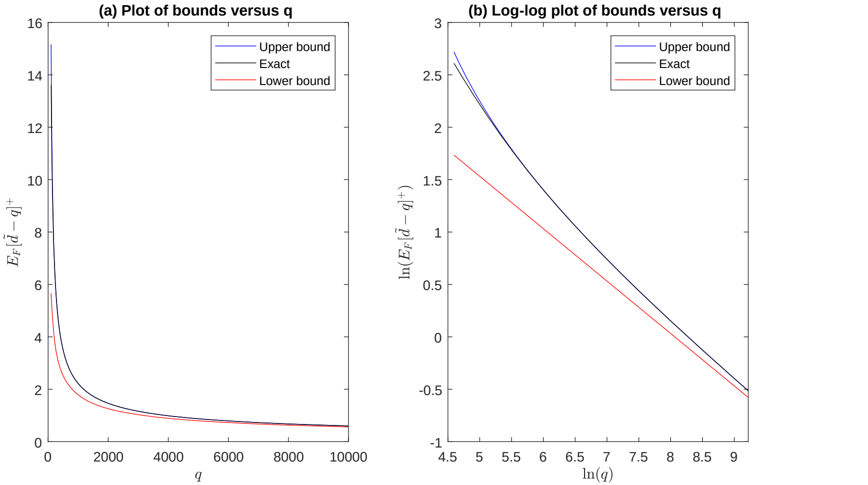
$$\begin{aligned} y_0 &= \frac{(n-1) m_1^n (n-1)^{n-1}}{n^n (q^{n-1} - K)}, \\ y_1 &= \frac{-n m_1^{n-1} (n-1)^{n-1}}{n^n (q^{n-1} - K)}, \\ y_n &= \frac{(n-1)^{n-1}}{n^n (q^{n-1} - K)}, \end{aligned}$$

- Choose K carefully for each of the cases with $n \in (2, \infty)$ and $n \in (1, 2)$ which ensures dual feasibility for large values of q.

Numerical ($n = 3$, $m_1 = 50$, $m_3 = 125150$)



Numerical ($n = 3/2,\ m_1 = 50,\ m_{3/2} = 500, \epsilon = 0.2$)



Regular Variation

- A function $u : \mathbb{R}_+ \to \mathbb{R}_+$ is regularly varying at infinity (Bingham, Goldie and Teugels (1998), Resnick (2008)) if for all $t > 0$, we have:

  $$\lim_{x\to\infty} \frac{u(tx)}{u(x)} = t^{\beta},$$

  for some $\beta \in \mathbb{R}$. We denote this function as $u \in \mathcal{RV}_\beta$.
- A nonnegative random variable with distribution function F is regularly varying if $\overline{F} := 1 - F \in \mathcal{RV}_{-\beta}$ for some $\beta \geq 0$.
- Examples: Pareto, Cauchy, Burr, Log-gamma, Student-t.

Tail behavior

Theorem

Given an ambiguity set $\mathcal{F}_{1,n}$ and fixed parameters m_1, m_n and n, as $q \to \infty$

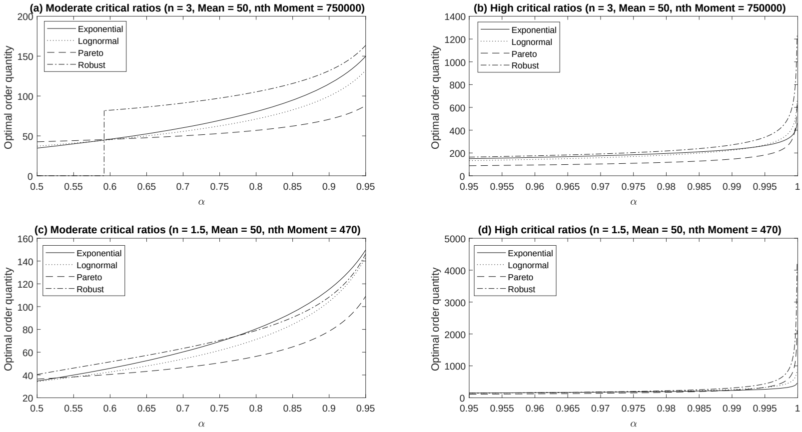
$$\mathbb{E}_{F^*}[D-q]_+ = \sup_{F \in \mathcal{F}_{1,n}} \mathbb{E}_F[D-q]_+ \quad \in \quad \mathcal{RV}_{-(n-1)}.$$

Moreover (under differentiation, Landau (1916))

$$\overline{F}^* \quad \in \quad \mathcal{RV}_{-n},$$

with $\mathbb{E}_{F^*}[D^{*k}] < \infty$ for all $0 \le k < n$ and $\mathbb{E}_{F^*}[D^{*k}] = \infty$ if $k \ge n$.

Numerical: Optimal Order Quantity



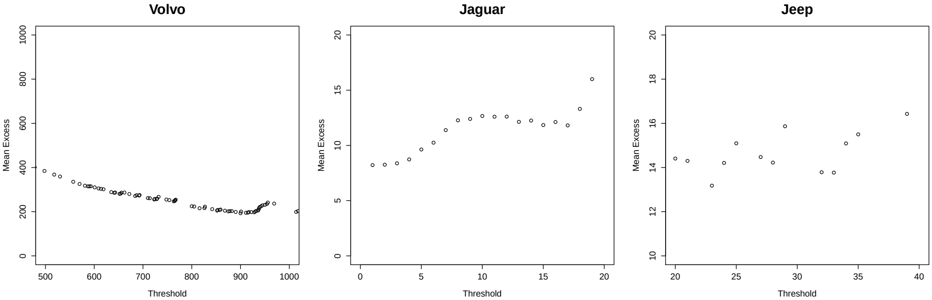
Evidence of Heavy-tailed Demand

Reference	Data	Demand
Clauset, Shalizi & Newman (2009)	Two data sets out of 24 - Number of calls, books sold	Power law is a good fit in 17 sets.
Gaffeo, Scorcu & Vici (2008)	Italian books	Power law: parameter < 2
Chevalier & Goolsbee (2003)	Sales of books on Amazon.com	Pareto: Parameter 1.2
Bimpkis & Markakis (2016)	626 products from North American retailer	Pareto: Parameter 1.38
Natarajan, Sim & Uichanco (2017)	Spare parts (IBM)	Best fit is often a heavy-tailed distribution

Norway monthly car sales: Jan 2007 – Jan 2017

- Published by *Opplysningsrådet for Veitrafikken (OFV)*.
- Obtained from `https://www.kaggle.com`.
- Monthly sales of cars by Make and Model.
- We observe monthly sales for: Volvo, Jaguar and Jeep.
- # of months observed: 121, 117, 109.
- Partition into: training set, test set.

Mean Excess Plot



Volvo

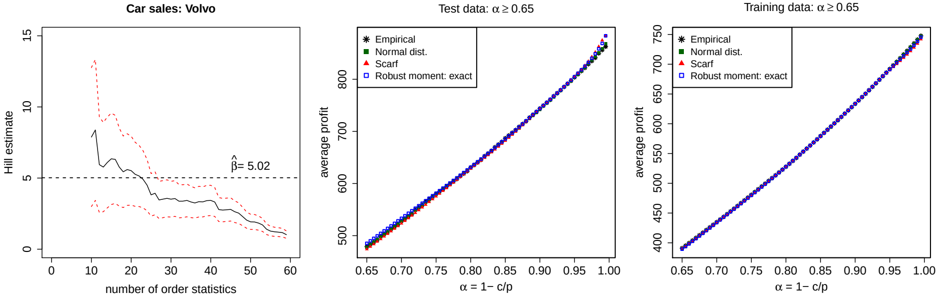
- Tail parameter estimate $\hat{\beta} = 5.02$.

Volvo

- Tail parameter estimate $\hat{\beta} = 5.02$.
- Take $n = 5$. We also compute $m_1 = 754.7, \ m_5 \sim 5.8 \times 10^{14}$.

Volvo

- Tail parameter estimate $\hat{\beta} = 5.02$.
- Take $n = 5$. We also compute $m_1 = 754.7$, $m_5 \sim 5.8 \times 10^{14}$.



Jaguar

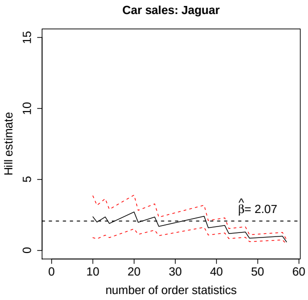
- Tail parameter estimate $\hat{\beta} = 2.07$.

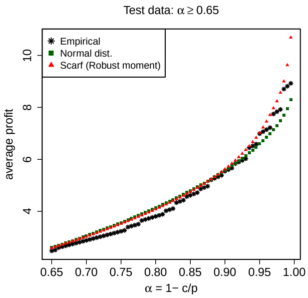
Jaguar

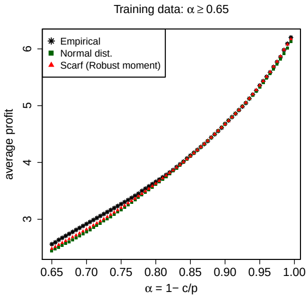
- Tail parameter estimate $\hat{\beta} = 2.07$.
- Take $n = 2$. This corresponds to Scarf's model. We compute $m_1 = 6.36,\ m_2 = 60.40$.

Jaguar

- Tail parameter estimate $\hat{\beta} = 2.07$.
- Take $n = 2$. This corresponds to Scarf's model. We compute $m_1 = 6.36$, $m_2 = 60.40$.







Jeep

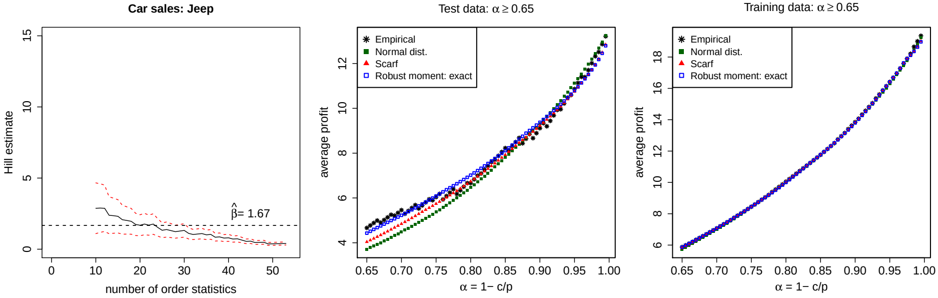
- Tail parameter estimate $\hat{\beta} = 1.672$.

Jeep

- Tail parameter estimate $\hat{\beta} = 1.672$.
- Take $n = 5/3$. We also compute $m_1 = 19.74,\ m_{5/3} \sim 207.34$.

Jeep

- Tail parameter estimate $\hat{\beta} = 1.672$.
- Take $n = 5/3$. We also compute $m_1 = 19.74$, $m_{5/3} \sim 207.34$.



Final Thoughts

- Given the first and the nth moment, for high critical ratios, the optimal order quantity from a distributionally robust newsvendor model is optimal for a $\mathcal{RV}_{-n}$ distribution. Scarf's model is a special case.

Final Thoughts

- Given the first and the nth moment, for high critical ratios, the optimal order quantity from a distributionally robust newsvendor model is optimal for a $\mathcal{RV}_{-n}$ distribution. Scarf's model is a special case.
- Are we using the best n?

Final Thoughts

- Given the first and the nth moment, for high critical ratios, the optimal order quantity from a distributionally robust newsvendor model is optimal for a $\mathcal{RV}_{-n}$ distribution. Scarf's model is a special case.
- Are we using the best n?
- More general moment conditions: Moments not completely specified but within a range. Further moments available.

Final Thoughts

- Given the first and the nth moment, for high critical ratios, the optimal order quantity from a distributionally robust newsvendor model is optimal for a $\mathcal{RV}_{-n}$ distribution. Scarf's model is a special case.
- Are we using the best n?
- More general moment conditions: Moments not completely specified but within a range. Further moments available.
- Multi-product decision.

Final Thoughts

- Given the first and the nth moment, for high critical ratios, the optimal order quantity from a distributionally robust newsvendor model is optimal for a $\mathcal{RV}_{-n}$ distribution. Scarf's model is a special case.
- Are we using the best n?
- More general moment conditions: Moments not completely specified but within a range. Further moments available.
- Multi-product decision.
- `https://arxiv.org/abs/1806.05379`.

Thank you