Grenander-type Density Estimation under Myerson Regularity

Haitian Xie

PhD, UC San Diego

Assistant Professor, Peking University

Canadian Economics Association Conference 2023

May 30, 2023

Density estimation for auctions

- Data: iid valuations $V_1, V_2, \cdots, V_n$ from second-price auctions of some product
- Want to estimate the underlying density f of V_i
- Kernel density estimator
  - motivated by smoothness assumptions
  - requires bandwidth/kernel selection
  - difficult to conduct inference at the optimal convergence rate due asymptotic bias

Research question

- Is there an estimation procedure that
  - can be motivated by economic theory
  - is fully data-driven (avoid bandwidth selection)
  - no asymptotic bias under the optimal convergence rate
- Myerson (1981) regularity condition

$$\text{virtual valuation } \varphi(v) = v - \frac{1 - F(v)}{f(v)} \text{ nondecreasing,}$$

  where F is the cumulative distribution function of V_i.

- This is a shape restriction on the density f. We are going to utilize it to design an estimator for f.

Main results

- Grenander-type estimation procedure motivated by Myerson regularity
  - equivalent representation of Myerson regularity as a convexity constraint
  - no tuning parameters needed in the construction
- Consistency, convergence rate, and asymptotic distribution (non-normal)
- Minimax optimal convergence rate

Roadmap

1. Introduction
2. Myerson regularity
3. Estimation procedure
4. Asymptotic properties
5. Conclusion

Myerson regular distributions

- Important results in mechanism design require Myerson regularity
  - e.g., the second-price auction with reserve price maximizes the revenue
- Examples:
  - any log-concave distribution: $U[0,1]$, $N(0,1)$, Exponential, logistic
  - student t, Cauchy, F (parameters > 2)
  - violations: U-shape / very heavy tails

Economic interpretation

A simple economic model (Bulow and Klemperer, 1996):

- selling a product with valuation V
- price of the product: p
- quantity demanded: $q = \mathbb{P}(V > p) = 1 - F(p)$
- inverse demand function: $p = F^{-1}(1-q)$
- revenue function: $R(q) = pq = F^{-1}(1-q)q$
- marginal revenue:

$$
\begin{aligned}
R'(q) &= F^{-1}(1-q) - \frac{q}{F'(F^{-1}(1-q))} \\
&= p - \frac{1-F(p)}{f(p)} = \varphi(p)
\end{aligned}
$$

Economic interpretation (cont'd)

- Myerson regularity: $\varphi(p)$ nondecreasing in p
  $\iff$ marginal revenue $R'(q)$ nonincreasing in q
  $\iff$ revenue function $R(q)$ concave in q
- Concave revenue functions are common
  - diminishing marginal returns/utility
  - risk aversion
  - equilirium considerations

Equivalent representation

- Difficult to directly apply the restriction [φ nondecreasing] to estimation because it involves f.
- Consider an equivalent condition (Ewerhart, 2013):
    - define $\Lambda(\cdot) = (1 - F(\cdot))^{-1}$, which only involves F
    - under very mild conditions (continuity):

$$\varphi(\cdot) \text{ nondecreasing} \iff \Lambda(\cdot) \text{ convex}$$

$$\begin{aligned} \varphi(\cdot) \text{ nondecreasing} &\iff \Lambda(\cdot) = (1 - F(\cdot))^{-1} \text{ convex} \\ &\iff \lambda(\cdot) = \Lambda'(\cdot) = f(\cdot)(1 - F(\cdot))^{-2} \\ &\quad\;\; \text{nondecreasing} \end{aligned}$$

- Take derivative and see that

$$\varphi', \lambda' \text{ share the same sign}$$

Equivalent representation (cont'd)

- λ as the derivative of Λ

$$\lambda(\cdot) = \Lambda'(\cdot) = f(\cdot)(1 - F(\cdot))^{-2}$$

- Take derivative

$$\varphi' = \frac{2f^2 + (1-F)f'}{f^2},$$
$$\lambda' = \frac{2f^2 + (1-F)f'}{(1-F)^3}.$$

- φ' and λ' share the same sign

Roadmap

1. Introduction
2. Myerson regularity
3. Estimation procedure
4. Asymptotic properties
5. Conclusion

Basic idea for estimation

- Since $f = \lambda(1 - F)^2$, we can first estimate λ then use empirical cdf F_n to replace F.
- How to estimate λ?
  - λ is the derivative of the convex function $\Lambda = (1 - F)^{-1}$
  - replace F by empirical cdf F_n
  - "convexify" this estimate of Λ and take derivative

The estimator

Step 1 Estimate Λ by Λ_n (replace cdf by empirical cdf):

$$\Lambda_n = (1 - F_n)^{-1}$$

Step 2 Let $\hat{\Lambda}_n$ be the *greatest convex minorant* (gcm) of Λ_n. Take $\hat{\lambda}_n$ as the left-derivative of $\hat{\Lambda}_n$.

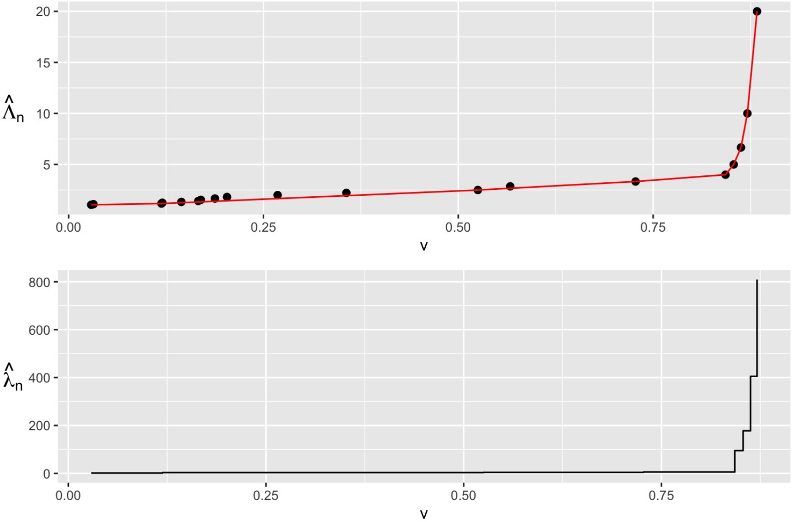
Step 3 The density estimator is $\hat{f}_n = \hat{\lambda}_n (1 - F_n)^2$.

- No tuning parameters because this estimator is not local. It uses the *global* shape restriction to estimate f.

Grenander-type estimator

- Estimators representable as the left derivative of the greatest convex minorant or least concave majorant of an estimator of a primitive function.
- Literature
  - monotone density: Grenander (1956)
  - concave distribution function: Beare and Fang (2017)
  - isotonic regression: Robertson and Wright (1975)
  - monotone hazard rate: Marshall and Proschan (1965)
  - General framework: Westling and Carone (2020), Durot et al. (2012), Durot (2007)
- This paper is the first one that uses Myerson regularity for Grenander-type estimation.

Demonstration: uniform, $n = 20$



Roadmap

1. Introduction
2. Myerson regularity
3. Estimation procedure
4. Asymptotic properties
5. Conclusion

Consistency

- Assume the density f is continuous and Myerson regular.
- Due to inconsistency of Grenander-type of estimators at boundary points (Woodroofe and Sun, 1993; Kulikov and Lopuhaä, 2006; Balabdaoui et al., 2011), we focus on $v \in [a, b]$, where $[a, b]$ is in the interior of the support of V_i

Theorem 1

1. $\hat{f}_n(v)$ is consistent for $f(v)$.
2. If further assume that f is uniformly continuous, then $\hat{f}_n$ is uniformly consistent.

Asymptotic distribution

Theorem 2

If f is continuously differentiable, then for any $v \in [a, b]$,

$$n^{1/3}(\hat{f}_n(v) - f(v)) \xrightarrow{d} C(v) \operatorname*{argmax}_{t \in \mathbb{R}} \{\mathbf{B}(t) - t^2\},$$

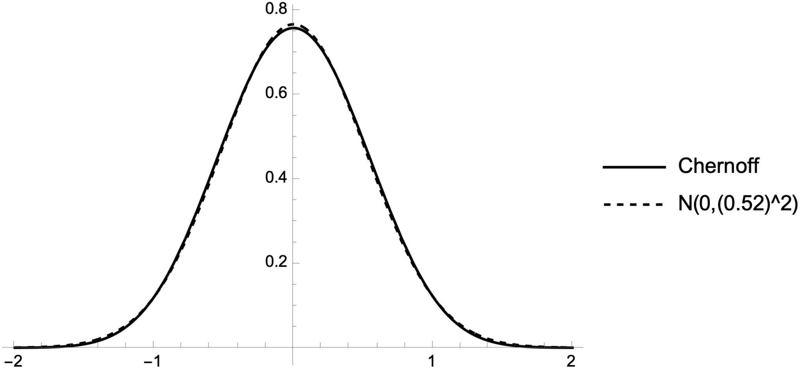
where $\mathbf{B}(t)$ is a two-sided Brownian motion and

$$C(v) = \left(\frac{8f(v)^3}{1 - F(v)} + 4f(v)f'(v) \right)^{1/3}.$$

Chernoff's distribution

- $\mathrm{argmax}_{t \in \mathbb{R}} \{\mathbf{B}(t) - t^2\}$
- it first arose in Chernoff (1964) on mode estimation, also appeared in Venter (1967) for another mode estimator
- log-concave
- symmetric around zero (has mean zero)
- variance ≈ 0.26
- can be approximated by $N(0, (0.52)^2)$ (Dykstra and Carolan, 1999)

Density plot



Inference

- obtain consistent estimates of f' to construct the test statistic (which is possible under our assumption that f' is continuous)
- the test statistics converges at the cube root rate
- for kernel estimates under the same condition, we usually need to undersmooth to kill the bias
- therefore, the local power function for our test would be "infinitely more efficiency" than the kernel-based test

Minimax convergence rate

Question: is it possible to do better (in terms of convergence rate) under our assumptions? No.

Theorem 3

Let $\mathcal{F}$ be the set of distributions that have a.e. continuously differentiable densities and satisfy Myerson regularity. For any $v \in [a, b]$, there exists $c > 0$ such that

$$\inf_{\tilde{f}_n} \sup_{F \in \mathcal{F}} \mathbb{E}_F |\tilde{f}_n(v) - f(v)| \geq c n^{-1/3},$$

where $\mathbb{E}_F$ denotes the expectation with respect to the distribution F. The infimum $\inf_{\tilde{f}_n}$ is taken over the set of all estimators.

Conclusion

- A tuning-parameter-free nonparametric density estimator.
- Asymptotic properties
  - (uniform) consistency
  - cube root convergence rate
  - non-normal asymptotic distribution
- Conjecture: $\hat{\lambda}_n$ is the nonparametric maximum likelihood estimator.
- Future work: simulation and empirical application.

Thank You!

Haitian Xie

Guanghua School of Management
Peking University

Website: haitianxie.org