Information Aggregation and the Information Content of Order Statistics

By Ilan Kremer and Andrzej Skrzypacz [1]

ABSTRACT

We show how auction theory can be used in the analysis of order statistics. In particular, we use a thought experiment in which we allocate signals to bidders who compete in a $k+1$-price auction. We use recent results about information aggregation in auction with many bidders to show that the amount of information in the k-th order statistics is increasing in k in the limit when the number of signals increases and k is fixed. Still, even the first-order statistic contains non-trivial information.

[1] We would like to thank George Shanthikumar, Matt Jackson and Peter DeMarzo for helpful comments. Contact information: Graduate School of Business, Stanford, CA 94305-5015 E-mail: ikremer@stanford.edu, andy@gsb.stanford.edu.

1 Introduction

Recently there has been a growing interest in the properties of common value auctions with a large number of bidders. This continues a classic line of research that examines markets with many strategic agents. Such analysis provides insights to the way prices aggregate private information and hence provides foundation to the concept of Rational Expectation Equilibrium (REE).

The properties of prices in these auctions are strongly related to the statistical properties of order statistics. For example, the price in a $k+1$-price auction is a function of the $k+1$ order statistic. Hence, it is no surprise that results from statistics about order statistics are the basis to much of the auction literature on this topic. In this paper we show how one can use this link in the opposite direction. We obtain results about the statistical properties of order statistics using results about information aggregation in auctions with many bidders. In particular we show how the results from Milgrom (1981), Pesendorfer and Swinkels (1997), Kremer (2000) and Jackson and Kremer (2004) can be used for that purpose. To the best of our knowledge our results were not previously known in statistics.

We derive new results about the informativeness of extreme order statistics when there are many signals. In particular, we consider a case when there are n signals about an unknown random variable V which can be thought of as the state of nature; the signals are distributed i.i.d. conditional on V. The general question is how much we can infer about the realization of V from observing different order statistics. In particular, we examine the information contained in order statistics in the limit as n grows to infinity.

We first show that under fairly standard assumptions, all order statistics are informative about V, even in the limit. That is, the expected value of V conditional on observing a k-th order statistic (for any k) does not collapse to the unconditional average value of V as n grows large. To the contrary, even in the limit there is a strict positive relationship between that conditional expectation and the actually realized value.

Second, we rank the information content of order statistics in the limit as n grows to infinity. We show that in the limit k-th order statistic is more informative than a $k+1$ order statistic, for any k. We provide formal definitions below, but in words, more informative means that if we want to infer the value of V, once we condition our inference on the realization of the $k+1$ order statistic, a further conditioning on the k-th order statistic does not improve the estimate any further (we prove that the ranking is strict, in the sense that once we condition on the k-th order statistic, we can learn more from the $k+1$ order).

Such results are of general statistical interest (apart of economics) as is evident by the large volume of research devoted to this topic (for example, Herbert (1981)). One motivation comes from the analysis of extreme or catastrophic events. For a simple example in which our analysis may be useful, suppose that a new engineering solution is incorporated in many bridges. The effect of this solution on stability of the bridges is not completely known, especially in the event of extreme weather conditions. Over time, some of the bridges show visible signs of structural damage. The question becomes how much can we infer about the invisible damage from these extreme statistics.

In econometrics, order statistics are especially important when we work with censored data. For example, in auctions we may see only the winning bids or price. The question is then how much we can infer from such data about bidders' preferences. Similarly, data on discrete decisions/selections present us only with a selection of the first-order (or a few more) statistic from a large set of alternatives. The question is then how much we can infer about the underlying state of the world from such censored data.

Our results may be also useful in examining agents who try to estimate a demand function. Consider a seller who wants to figure out the underlying value of a new product. He considers selling a few samples to estimate the demand and later plans to produce and sell more only if the distribution of buyers' valuations for his product is sufficiently high. Our results imply that selling more units in the trial run is better.

Finally, our results can be useful in analysis of models of advertising, like a recent paper by Bernhardt and Duggan (2005) on political campaigns. To illustrate, consider the following simple model of advertising in political campaigns. A political candidate has a quality V (with voters agreeing that a higher V is better) and has n attributes that are affiliated with V (and distributed i.i.d. conditional on V). Suppose n is large and the candidate can reveal to the public only $k << n$ of attributes (say, because communication is costly). Suppose that he cannot commit to a strategy of revealing a particular order statistic. The optimal strategy is then to reveal the best k of attributes. Our results on order statistics can then be used to characterize the equilibrium, in particular the inferences that voters can make about V.

2 The information in order statistics

Consider a random variable, V, distributed on $[0,1]$ according to a non-degenerate distribution $F_V(v)$. We try to estimate V using signals that are given by $\{S_i\}_{i=1}^{\infty}$ that are also distributed on $[0,1]$; conditional on V signals are distributed i.i.d. according to a conditional density $f(s|v)$. We make the following assumptions:

(A1) For any $\varepsilon > 0$ there exists $\delta > 0$ such that if $|s - s'| < \delta$ then for almost every v

$$\left| \frac{f(v|s)}{f(v|s')} - 1 \right| < \varepsilon.$$

(A2) Monotone likelihood ratio (MLRP) $\frac{f(s|v)}{f(s'|v)} > \frac{f(s|v')}{f(s'|v')}$ for $s > s', v > v'$.

(A3) There exists some $a, b > 0$ so that for any s and v we have that: $a < f(s|v) < b$.

(A1) is a uniform continuity condition in signals that applies across v. It implies that nearby signals provide similar information about the realization of V. (A2) means that s and v are strictly affiliated. This is a standard assumption in the auction literature and it implies a positive correlation between signals and V. Lastly, assumption (A3) implies that there is a limited amount of information in every signal. While in the limit the collection of all signals is sufficient to infer V precisely, a single signal conveys only partial information. Technically, (A3) implies a finite likelihood ratio of V given S_i. See Pesendorfer and Swinkels (1997) for a more detailed discussion of the implications of such bounds.

We examine what can be inferred about V when observing the k-th order statistic; we denote by $Y_n(k)$ the $k-th$ order statistic among the first n signals, $\{S_i\}_{i=1}^{n}$ and consider

$$E[V|Y_n(k)]$$

We use the notion of conditional expectation to define what we mean by saying that in the limit one random variable is more informative than another. One may worry that this definition of information is restricted to first moments. Note however that $E[V|Y_n(k)]$ is actually a random variable; it is the expectation of V conditional on the $\sigma-$algebra generated by $Y_n(k)$.[2] As a result, one can show that if $E[V|Y_n(k), Y_n(k')] - E[V|Y_n(k)] \to 0$ (in probability) then $E[f(V)|Y_n(k), Y_n(k')] - E[f(V)|Y_n(k)] \to 0$

[2]See Durrett (1996) for a definition of conditional expectation.

(in probability) for any function f, hence, one can replace the first moment with any other moment. Formally,

Definition 1 *(i) We say that in the limit the $k-th$ order statistic is at least as informative as the $k'-th$ order statistic if* $E\left[V|Y_n(k), Y_n(k')\right] - E\left[V|Y_n(k)\right] \to 0$ *in probability*
(ii) We say that in the limit the $k-th$ order statistic is more informative than the $k'-th$ order statistic if it is at least as informative and $E\left[V|Y_n(k')\right] - E\left[V|Y_n(k)\right] \nrightarrow 0$ *in probability.*
(iii) We say that in the limit the $k-th$ order statistic contains non-trivial information if $E[V] - E\left[V|Y_n(k)\right] \nrightarrow 0$ *in probability.*

In words, part (i) of the definition states that (in the limit) if we want to infer V, once we condition on $Y_n(k)$, there is no additional information in $Y_n(k')$. Part (ii) states that (in the limit) conditioning on $Y_n(k)$ and $Y_n(k')$ leads to different amount of information, which combined with part (i) allows us to strictly rank informativeness of the order statistics. Finally, part (iii) states that even in the limit conditioning on $Y_n(k)$ we can estimate V better than just by using its unconditional mean.

Our main result is that using auction theory we can rank the limiting information in order statistics. We conduct the following thought experiment: we assign signals $\{S_i\}_{i=1}^n$ to agents and run a $k+1$-price auction in which bidders compete for k identical goods whose (common) value is given by V. In such an auction, the k highest bidders win one unit each and all of them pay the $k+1$ highest price. This is a generalization of what is known as the second price auction which corresponds to $k=1$.

Our results are based on some known properties of common value auctions. Milgrom (1981) analyzed such an auction and showed that an agent with a signal s bids in a symmetric equilibrium:

$$b_n(s) = E\left[V|Y_n(k) = Y_n(k+1) = s\right]$$

The above expression implies that:

- The bidding strategy is monotone, as a result the price equals the bid of the agent with the $k+1$ highest signal, that is $P_n = b_n(Y_n(k+1))$.
- $P_n \leq E\left[V|Y_n(k+1)\right]$ almost surely; this follows from our affiliation assumption (A2),

$$E\left[V|Y_n(k)=Y_n(k+1)=s\right] \leq E\left[V|Y_n(k+1)=s, Y_n(k+1)\leq Y_n(k)\right] = E\left[V|Y_n(k+1)=s\right] \quad (1)$$

Jackson and Kremer (2004) prove that such an auction is 'competitive' in the sense that as $n \to \infty$, bidders' profits to converge to zero, i.e.[3]

$$E[V] - E[P_n] \to 0$$

Some of the above properties extend also to a first-price auction (for a single item). In particular, while the bidding function takes a different form (see Milgrom and Weber (1982)), it is still the case that $P_n \leq E[V|Y_n(1)]$ and $E[V]-E[P_n] \to 0$. Based on this one can conclude that for any $k \geq 0$ the price in the $k+1$ price auction (where we take $k=0$ to be a first-price auction) satisfies:[4]

$$P_n - E\left[V|Y_n(k+1)\right] \to 0 \text{ in probability} \quad (2)$$

Furthermore, for $k \geq 1$ (1) implies that in a k+1 price auction $P_n \leq E\left[V|Y_n(k), Y_n(k+1)\right]$. Hence, we by the same reasoning:

$$P_n - E\left[V|Y_n(k), Y_n(k+1)\right] \to 0 \text{ in probability} \quad (3)$$

As we shall see, the combination of the above results enables us to rank the information in order statistics. First we show that extreme order statistics contain non-trivial information. One may conjecture that it is not the case as in our setup regardless of the realization of V the highest signals converge to one. However, using auction theory we argue that:

Lemma 1 *In the limit every k-th order statistic $Y_n(k)$ contains non-trivial information.*

Proof. We argued that $P_n - E[V|Y_n(k)] \to 0$ in probability. Assume, by contradiction, that $P_n - E[V] \to 0$. An individual signal contains non trivial information (even in the

[3]The convergence of $E[P_n]$ to $E[V]$ can be interpreted as a simple convergence as both $E[P_n]$ and $E[V]$ are constants and not random variables.

[4]If we have two sequences of bounded random variables where $X_n \leq Y_n$ almost surely and $E[X_n] - E[Y_n] \to 0$, then $X_n - Y_n \to 0$ in probability. Kremer (2002) uses a similar argument to compare first and second-price auctions.

limit) as $E[V|S_i]$ does not depend on n; in particular for s close to one we have that $E[V|S_i = s] > E[V]$. This leads to a contradiction as there is a positive probability that there is an agent with a high signal such that $E[V|S_i] > E[V] + \varepsilon$ for some $\varepsilon > 0$. This agent can bid $E[V] + \varepsilon/2$ and ensure a profit that is bounded away from zero. This would contradict that in such an auction the expected bidder payoffs converge to zero as $n \to \infty$. ■

Next, we show that in the limit there is a strict ranking of the information contained in different order statistics:

Theorem 1 *For any $k \geq 1$ in the limit the $k+1$ order statistic, $Y_n(k+1)$, is more informative than the $k-th$ order statistic, $Y_n(k)$.*

The formal proof appears in the appendix but we provide here a short outline. An immediate implication of (2) and (3) is that for any $k \geq 1$ in the limit the $k+1$ order statistic, $Y_n(k+1)$, is at least as informative as the $k-th$ order statistic, $Y_n(k)$.Hence, we need only to verify that $E[V|Y_n(k')] - E[V|Y_n(k)] \nrightarrow 0$. Since signals are in $[0,1]$, the maximal conditional expectations for the value of the asset are given by: $E[V|Y_n(k+1) = 1]$ and $E[V|Y_n(k) = 1]$, respectively. It can be shown that both $E[V|Y_n(k+1) = 1]$and $E[V|Y_n(k) = 1]$ do not depend on n and $E[V|Y_n(k+1) = 1] > E[V|Y_n(k) = 1]$.[5] The claim then follows from the fact that we can find a threshold s_n^* close to one so that (i) for any $s > s_n^*$, $E[V|Y_n(k+1) = s] > E[V|Y_n(k) = 1]$, and (ii) $\Pr(Y_n(k+1) > s_n^*)$ is (uniformly) bounded away from zero.

We conclude the analysis by illustrating our results using a numerical example:

Example 1 *Consider the case where $V = \{0,1\}$ with equal probabilities and*

$$f(s|v) = \begin{cases} 1 & if\ v = 0 \\ 2s & if\ v = 1 \end{cases}$$

Suppose there are $n = 1000$ signals and we observe the highest or the fifth-highest signal. To what extent can we tell the value of V? Using Bayes rule, we can express the expected value of V conditional on observing the $l-th$ highest signal, $E[V|Y_n(l) = s]$:

$$\frac{\int vF(S_i = s|v)^{n-l} f(S_i = s|v)(1 - F(S_i = s|v))^{l-1} f(v)\,dv}{\int F(S_i = s|v)^{n-l} f(S_i = s|v)(1 - F(S_i = s|v))^{l-1} f(v)\,dv}$$

[5] Consider for example $k = 1$. Conditioning on the first signal being one, $S_1 = 1$, is the same as conditioning on the first-order statistic being one. Since signals are in $[0,1]$, the S_1 must be the highest signal. This argument is used in Pesendorfer and Swinkels (1997) p.1257.

In Figure 1 we use a Monte-Carlo simulation to plot the density function of the random variable $E[V|Y_n(1)]$ *, that is the expected value of* V *conditional on the first-order statistic. Figure 2 shows the results for the fifth order statistic, that is,* $E[V|Y_n(5)]$.[6] *In both cases we plot two curves that represent the conditional densities of the random variables when we conditional on the realization of* V *being zero or one.*

Two conclusions can be made:

1. The two conditional distributions are not identical, so even the first order statistic contains some information. In fact, if the first order statistic contained no information both conditional distributions would converge simply to a spike at 0.5.

2. The two conditional distributions are more separated when we use the fifth order statistic. It reflects the fact that the fifth order statistic is more informative than the first order statistic (although dispersion of conditional distributions is not the formal definition of informativeness that we use, the separation of the conditional distributions is indicative of the information contained in the order statistic).

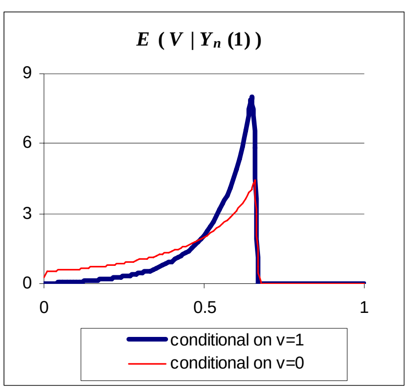

Figure 1

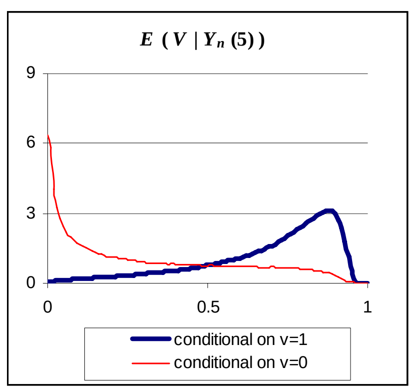

Figure 2

3 Appendix

<u>Proof of Theorem 1:</u>

[6] We use $1,000,000$ iterations.

An immediate implication of (2) and (3) is that for any $k \geq 1$ in the limit the $k+1$ order statistic, $Y_n(k+1)$, is at least as informative as the $k-th$ order statistic, $Y_n(k)$. Hence, we need only to verify that $E[V|Y_n(k')] - E[V|Y_n(k)] \nrightarrow 0$.

We first note a few properties of the expected value of V conditional on the $l-th$ order statistic being equal to s, $E[V_n|Y_n(l)=s]$, which is given by:

$$\frac{\int vF(S_i=s|v)^{n-l} f(S_i=s|v)(1-F(S_i=s|v))^{l-1} f(v)\,dv}{\int F(S_i=s|v)^{n-l} f(S_i=s|v)(1-F(S_i=s|v))^{l-1} f(v)\,dv} \tag{4}$$

Because $\lim_{s\uparrow 1} \frac{1-F(s|v)}{(1-s)f(s|v)} = 1$ (by continuity of $f(s|v)$, for details see Pesendorfer and Swinkels (1997) p.1257), taking the limit of the above expression we get:

$$E[V_n|Y_n(l)=1] = \frac{\int vf(S_i=1|v)^l f(v)\,dv}{\int f(S_i=1|v)^l f(v)\,dv}$$

Hence, $E[V_n|Y_n(l)=1]$ does not depend on n.

For $s \in [0,1)$ define

$$g(s,l) = \frac{\int vf(S_i=s|v)(1-F(S_i=s|v))^{l-1} f(v)\,dv}{\int f(S_i=s|v)(1-F(S_i=s|v))^{l-1} f(v)\,dv}$$

Note that $g(s,l)$ is continuous in s and by strict affiliation (assumption (A2)) it is strictly increasing in l. Using again $\lim_{s\uparrow 1} \frac{1-F(s|v)}{(1-s)f(s|v)} = 1$, we can define $g(1,l)$ as the limit as $s \uparrow 1$:

$$g(1,l) \equiv \lim_{s\uparrow 1} g(s,l) = E[V_n|Y_n(l)=1]$$

which is strictly increasing in l.

Next, consider a fixed $\delta \in \left(0, \frac{1}{b}\right)$ and some $s > \underbrace{1 - \frac{\delta}{n}}_{s_n}$. We argue that the probability of the $l-th$ order statistic being above s_n is uniformly bounded away from zero and the expectations conditional on the two order statistics differ in this range (for some δ and uniformly for all n).

By assumption (A3), $F(s|v) > 1 - \frac{b\delta}{n}$ for all v, which implies that:[7]

$$F(s|v)^{n-l} > F(s|v)^n > \left(1 - \frac{b\delta}{n}\right)^n > 1 - b\delta \tag{5}$$

[7] Note that $\left(1 - \frac{x}{n}\right)^n > 1 - x$ for $x \in (0,1)$.

We claim that for any fixed l, $\Pr\left(Y_n(l) > 1 - \frac{\delta}{n}\right)$ is bounded away from zero. This probability is higher than the probability that exactly l signals are above $1 - \frac{\delta}{n}$. We bound the probability of this event by conditioning on $V = v$:

$$\Pr\left(Y_n(l) > 1 - \frac{\delta}{n}\right) \geq \binom{n}{l} F\left(S_i = 1 - \frac{\delta}{n}|v\right)^{n-l}\left(1 - F\left(S_i = 1 - \frac{\delta}{n}|v\right)\right)^l.$$

(A3) implies that $1 - F\left(S_i = 1 - \frac{\delta}{n}|v\right) > \frac{a\delta}{n}$. As in (5), we note that $F\left(S_i = 1 - \frac{\delta}{n}|v\right)^{n-l} > 1 - b\delta$ which provides a bound:

$$\Pr\left(Y_n(l) > 1 - \frac{\delta}{n}\right) \geq (1 - b\delta)\binom{n}{l}\left(\frac{a\delta}{n}\right)^l$$

For large enough n, we know that $\binom{n}{l} > \frac{n^l}{2^l l!}$. Since l is fixed, we can combine these inequalities to conclude that for any l there exists $\alpha > 0$ so that $\Pr\left(Y_n(l) > 1 - \frac{\delta}{n}\right) > \alpha$ for all n large enough.

Now, keeping $s > 1 - \frac{\delta}{n}$,using (4) and $F(s|v)^{n-l} \in (1 - b\delta, 1)$, we can conclude that:

$$E[V|Y_n(l) = s] \geq (1 - b\delta)\, g(s, l)$$

For n is large enough, continuity of $g(s, l)$ and $s > 1 - \frac{\delta}{n}$ implies that $g(s, l) > (1 - b\delta)\, g(1, l)$. Therefore:

$$E[V|Y_n(l) = s] > (1 - b\delta)^2\, g(1, l)$$

Finally, consider the k-th and $k+1$ order statistics. As we noted above, $g(1, k+1) > g(1, k)$. Set δ^* so that $(1 - b\delta^*)^2\, g(1, k+1) > g(1, k)$. The theorem then follows as we know that the probability that $Y_n(k+1)$ is above $\underbrace{1 - \frac{\delta^*}{n}}_{s_n^*}$ is bounded away from zero and for any $s > 1 - \frac{\delta^*}{n}$ we have that the conditional expectations differ: $E[V_n|Y_n(k+1) = s] > (1 - b\delta^*)^2\, g(1, k+1) > g(1, k) = E[V_n|Y_n(k) = 1] \geq E[V_n|Y_n(k) \geq s]$.

References

Bernhardt, D. and J. Duggan (2005), "Why are political campaigns so d–ned negative?" mimeo, University of Rochester.

Durrett R. (1996), "Probability: Theory and Examples," *Duxbuty Press*, Second Edition

Herbert, A. David (1981), "Order Statistics," *John Wiley & Sons, second edition*.

Jackson, M. and I. Kremer (2004), "The Relationship between the Allocation of Goods and a Seller's Revenue," *Journal of Mathematical Economics*, forthcoming.

Kremer I. (2002), "Information Aggregation in Common Value Auctions," *Econometrica*, 70, 1675-1682.

Milgrom, P. (1981), "Rational Expectations, Information Acquisition and Competitive Bidding," *Econometrica,* 49, 921–943.

Milgrom, P. and R. Weber (1982), "A Theory of Auctions and Competitive Bidding," *Econometrica,* 50, 1089–1122.

Pesendorfer, W. and J. Swinkels (1997), "The Loser's Curse and Information Aggregation in Common Value Auctions," *Econometrica,* 65, 1247–1282.