A Brief Review of Probability Theory

The finite case. Suppose that the set $S$ of states consists of a finite set of real numbers, $S = \{s_1 < s_2 ... < s_n\}$. A probability density function $f : S \rightarrow \mathbb{R}$ assigns a probability to each possible state. The density function must have the following properties:

1. $0 \leq f(s_i) \leq 1$ for each state $s_i \in S$;
2. $\sum_{i=1}^{n} f(s_i) = 1$.

The (cumulative) distribution function $F : S \rightarrow \mathbb{R}$ associated with the density $f$ calculates for each state $s_i$ the probability that the state is less than or equal to $s_i$:

$$F(s_i) = \sum_{j=1}^{i} f(s_j).$$

The uniform distribution is the special case in which each of the $n$ states $s_i$ occurs with the same probability,

$$f(s_i) = \frac{1}{n},$$

where the $n$ reflects the total number of values in the set $S$. The expected value of the states is the sum of the states with each weighted according to its likelihood of occurrence:

$$\mathbb{E}[s_i \in S] = \sum_{i=1}^{n} f(s_i) \cdot s_i.$$  \hspace{1cm} (6)

The continuum case. We want to be able to use calculus in addressing issues of probability, and so we now assume that the set of states $S$ is a subinterval $(\xi, \pi)$ of the real line, with $\xi < \pi$ (and each possibly infinite in magnitude). It does not matter whether or not the endpoints are included. Let’s extend the above ideas to this case.

A probability density function is again a function $f : S \rightarrow \mathbb{R}$. We want to interpret $f(s)$ as the probability that the state $s$ occurs. There is a continuum of states, however, and we cannot assign positive probability to each. We interpret

$$\int_{s'}^{s''} f(s) \, ds$$

as the probability that the state $s$ lies in the subinterval $(s', s'') \subset (\xi, \pi)$ (think of the integral as summing up the probabilities of all the states $s \in (s', s'')$). Intuitively, think of the density function as specifying the probability of any given state. The density function $f$ now has the following properties:

1. $0 \leq f(s)$ for all $s \in (\xi, \pi)$.
2. $\int_{\xi}^{\pi} f(s) \, ds = 1$.

The second condition is that $S$ represents all possible states (i.e., whatever state occurs lies within $S$ with probability 1). Notice that we do not require $f(s) \leq 1$, as this is not necessary to (7) to hold. The cumulative distribution $F : S \rightarrow \mathbb{R}$ again determines the
The probability that the state is below the specified value $s'$,

$$ F(s') = \int_{s}^{s'} f(s) \, ds. \quad (8) $$

The uniform distribution is the special case in which the probability that the state $s$ lies in the subinterval $(s', s'')$ is the length of that subinterval as a fraction of the entire interval of states,

$$ \int_{s'}^{s''} f(s) \, ds = \frac{s'' - s'}{s - s} $$

Because this holds for all subintervals $(s', s'')$, it follows that

$$ f(s) = \frac{1}{s - s}, $$

i.e., the density is constant in the uniform case.

Notice also that in general,

$$ F'(s) = f(s), $$

i.e., the derivative of the distribution function is the density function. This follows from (8) by the Fundamental Theorem of Integral Calculus.

Finally, the expected value of the state is

$$ \mathbb{E}[s \in S] = \int_{s}^{s} (s \cdot f(s)) \, ds. $$

This can be interpreted as summing up (i.e., the integral) the different possible states $s$, weighting each with its probability $f(s) \, ds$.

### 0.0.12.3 Our Model of Auctions.

We begin by assuming that there are $n$ possible bidders and a seller with a single, indivisible item to sell. Each of the $n$ bidders privately knows his own reservation value. Bidder $i$'s utility is

$$ v_i \delta_i - p_i, $$

where $v_i$ is his reservation value, $\delta_i = 0$ when he does not trade, $\delta_i = 1$ when he trades, and $p_i$ is any monetary transfer or payment that he makes to the seller. We allow the possibility of payments even when a buyer fails to buy the item. For now, we will assume that the seller's reservation value is 0 and all reservation values of bidders are nonnegative.

This is a private value model in the sense that each bidder fully and privately knows at the time he bids the value of the item to him. In many situations, however, a bidder may revise or update his assessment of the item upon seeing the bids of the others. An oil company, for instance, may estimate on its own the value of the oil in a particular tract; the bids of other companies for the tract, however, reflect their estimates and perhaps might cause the first company to revise its estimate. We'll deal with these issues later in the course.

---

4 One example is the "all-pay" auction in which the highest bidder receives the item but all bidders must pay their bids to the seller (even the losers). This is sometimes used a model of firms competing for a defense contract: all the firms pay a large investment in designing the new missile or jet, but only the winning bidder receives the contract – all other bidders lose their investments.
0.0.12.4 The Vickrey or Second-Price Auction.

A bid is collected from each bidder. The bids are ordered in a list

\[ b_{(1)} \leq b_{(2)} \leq \ldots \leq b_{(n)}. \]

The bidder who bid the most (\( b_{(n)} \)) wins the item and pays and pays the second highest bid \( b_{(n-1)} \) to the seller. In the case of ties at the high bid, a fair lottery among those who submitted the high bid of \( b_{(n)} \) determines the winner, and the price is again the second highest bid of \( b_{(n)} = b_{(n-1)} \).

We show below that bidding one’s reservation value weakly dominates any other strategy. It is the unique dominant strategy for each bidder. The auction was first proposed by William Vickrey in a famous 1961 paper. It’s purpose is to take the "gaming" or "strategizing" out of auctions: every bidder’s best interest is simply to truthfully report his reservation value. Campbell refers to the situation in which honest reporting is a dominant strategy as **incentive compatible**. Notice also that when bidders use their dominant strategy of bidding their true reservation values, the outcome of the auction is **efficient** in the sense that it successfully awards the item to the bidder who values it the most.

Consider bidder \( i \) with value \( v_i \) who considers the bid \( b_i \). Let \( \overline{b} \) denote the maximal bid of the other \( n-1 \) bidders. Bidder \( i \)'s utility as a function of \( v_i, b_i, \) and \( \overline{b} \) is

\[
u_i(v_i, b_i, \overline{b}) = \begin{cases} v_i - \overline{b} & \text{if } b_i > \overline{b} \\ \delta (v_i - \overline{b}) & \text{if } b_i = \overline{b} \\ 0 & \text{if } b_i < \overline{b} \end{cases}
\]

We see here the utility consequences of bidder \( i \) being the high bidder, tying as the high bidder, or losing the auction. Here, \( \delta \leq 1/2 \) reflects the randomization in the case of ties and depends upon the number of bidders who submit the bid \( \overline{b} \). For any given bids of the other bidders (which bidder \( i \) doesn’t know when he chooses his bid), bidder \( i \)'s choice of bid \( b_i \) doesn’t change the 3 possible utilities that he might receive: it only affects which of the 3 values that he receives. We can order the possible utilities as follows:

- if \( v_i > \overline{b} \), then \( v_i - \overline{b} > \delta (v_i - \overline{b}) > 0 \).
- if \( v_i = \overline{b} \), then \( v_i - \overline{b} = \delta (v_i - \overline{b}) = 0 \).
- if \( v_i < \overline{b} \), then \( 0 > \delta (v_i - \overline{b}) > v_i - \overline{b} \).

The bids of the other bidders present bidder \( i \) with a choice among 3 possible utilities. The key point to notice is that choosing \( b_i = v_i \) insures that bidder \( i \) receives the largest of these 3 payoffs for all possible profiles of bids by the other bidders: if he happens to be the high bidder with \( v_i = b_i > \overline{b} \), then \( v_i - b \) is the maximal possible utility available to him; if he ties as high bidder with \( v_i = b_i = \overline{b} \), then he is indifferent between winning and losing (0 is the best that he can do); if he loses with \( v_i = b_i < \overline{b} \), then he’s glad to lose because the bids of the other bidders simply do not present him with an opportunity to make a profit. For all possible profiles of bids by the other bidders, bidding his value maximizes \( i \)'s utility. It is therefore a dominant strategy.

It is interesting that a bidder would not benefit from learning the bids of the other bidders in advance, for bidding his value is optimal regardless of the bids of the other
bidders. How does one argue that it is bidder $i$’s unique dominant strategy? Any other bid is not optimal for some choice of $\beta$. Consider for instance, $v_i < b_i$. If $v_i < b < b_i$, then bidder $i$ wins the auction and suffers a loss of $v_i - b$. Consider $b_i < v_i$. If $b_i < b < v_i$, then bidder $i$ loses the auction and forgoes the positive profit he could have won by bidding between $b$ and $v_i$. Bidding one’s value is the only bid with the property that it maximizes one’s payoff against all possible specifications of the bids of the others.

0.0.12.5 Social Cost.

Allocating an item to one person in an auction means that some other person misses out on obtaining the item. If we order the reservation values of bidders in a list, $v(1) \leq v(2) \leq \ldots \leq v(n)$, and if the item is sold to the person with reservation value $v(n)$, then the social cost to the other $n-1$ bidders is $v(n-1)$, i.e., the maximum possible value that could have been achieved for the group if the item had not been awarded to the person with reservation value $v(n)$. This is the opportunity cost of awarding the item to a particular person (i.e., it then cannot be awarded to someone else).

The dominant strategy equilibrium of the Vickrey auction has each bidder submitting his true reservation value as his bid. The bidder with the highest value therefore wins and pays the second highest value as his price. This is social cost pricing in the sense that the winning bidder is charged an amount equal to the value that he takes away from the others by virtue of winning the auction. It is analogous to charging/compensating people for the external effects that they create through their actions.

Campbell, p. 338: "...social cost pricing in general involves charging an individual a fee equal to the cost that the individual’s action has imposed on the rest of society."

Don’t spend too much time on the other examples of social cost pricing on pages 338-341 – I don’t think that they are particularly interesting or important.

0.0.13 2.3 Incentives, Efficiency, and Social Cost Pricing

We have shown that the Vickrey auction has the following two properties:

- it is incentive compatible: it is a dominant strategy for each bidder to report his true reservation value;
- it is efficient in the sense that it awards the item in equilibrium to a bidder who truly values it the most.

We next note that it satisfies the following participation constraint:

- A bidder who fails to win the item neither pays anything nor receives any payment.

This is stronger than a more typical participation constraint, namely:

- In the honest equilibrium, no bidder ever suffers a loss. In particular, a bidder who fails to buy is not required to make any positive payment.

This last condition is a participation constraint in the sense that it is always sensible
for a bidder to voluntarily participate in the Vickrey auction; if he sometimes suffered a loss, then he might not want to participate. Campbell’s participation constraint is more restrictive in the sense that it rules out a bidder receiving a positive monetary payment when he fails to win the auction.

We next clarify a sense in which the Vickrey auction is the only auction that satisfies these three properties. For incentive compatibility, we need to restrict attention to auctions in which bidders make a report concerning their reservation values. This is a **direct (or revelation) auction**. Many auctions have this property, though we will discuss at least one that does not (the English or ascending bid auction). We want to prove the following:

**Theorem 3 (Uniqueness of the Vickrey Auction)**  
The Vickrey auction is the only direct auction satisfying incentive compatibility, efficiency, and the participation constraint.

We consider an arbitrary direct auction that is incentive compatible, efficient, and satisfies the participation constraint. To prove this theorem, let’s distinguish between bidder $i$’s reservation value $v_i$ and his report $r_i$. The price and the assignment of an item in a direct auction is based on the reports, and we want to show that any direct auction that has these three properties necessarily allocates the item and collects payments based upon these reports in exactly the same way as the Vickrey auction.

- Given a list of reported values $r_1, r_2, r_3, \ldots, r_n$, the item must be allocated to the person who makes the highest report. This is true because the auction is incentive compatible and efficient: when the reports are honest, the item must be assigned to the bidder who reports the highest value. The direct auction receives reports, however, and does not know whether or not they are honest; whatever reports are submitted, it must therefore assign the item to the bidder who submitted the largest report. The direct auction is therefore identical to the Vickrey auction in its rule for allocating the item based upon the reports.

We’ve figured out that the direct auction must assign the item to the bidder who makes the largest report. Campbell’s participation constraint implies that a trader who does not trade also does not pay or receive anything. It remains to be shown that price paid by the winning bidder equals the second highest reported value, as in the Vickrey auction. Assume that $r_1$ is the largest report and let $r^*$ denote the largest report of the other bidders, 

$$r^* = \max \{r_2, r_3, \ldots, r_n\}.$$ 

Let $P(r_1, r_2, \ldots, r_n)$ denote the price paid by bidder 1 in the direct auction when $r_1$ is the largest report, i.e., $r_1 > r^*$. We want to show that

$$P(r_1, r_2, \ldots, r_n) = r^*,$$

i.e., the direct auction charges the winning bidder the second highest report, just as in the Vickrey auction. We’ll do this in three steps:

- If $v_1 > r^*$, then by reporting honestly bidder 1 wins the auction and receives 

$$v_1 - P(v_1, r_2, \ldots, r_n).$$
We must have
\[ v_1 - P(v_1, r_2, \ldots, r_n) \geq 0 \]
for all \( v_1 > r^* \) by incentive compatibility; if this were not true, then bidder 1 would change his report so that it wasn’t the highest and thereby obtain a utility of 0. Changing the \( v_1 \) to \( r_1 \), we have therefore deduced that
\[ P(r_1, r_2, \ldots, r_n) \leq r_1 \quad (9) \]
for all \( r_1 > r^* \).

* We next show that
\[ P(r_1, r_2, \ldots, r_n) \leq r^* \]
for all \( r_1 > r^* \). This is a “tighter” upper bound on \( P(r_1, r_2, \ldots, r_n) \) than \( r_1 \), and we are inching our way toward the result we want. Suppose not, i.e., \( r^* < P(r_1, r_2, \ldots, r_n) \).

Let \( T_1 \) denote the average of \( r^* \) and \( P(r_1, r_2, \ldots, r_n) \), i.e.,
\[ T_1 = \frac{r^* + P(r_1, r_2, \ldots, r_n)}{2} \]
We have
\( r^* < T_1 < P(r_1, r_2, \ldots, r_n) \).

Inequality (9) above implies
\[ P(T_1, r_2, \ldots, r_n) \leq T_1. \]
The inequality \( T_1 > r^* \) means that bidder 1 still wins the auction if he reports \( T_1 \) instead of \( r_1 \), and the inequality
\[ P(T_1, r_2, \ldots, r_n) \leq T_1 < P(r_1, r_2, \ldots, r_n) \]
implies that he reduces the price that he pays by reporting \( T_1 \) instead of \( r_1 \). In the case of \( r_1 = v_1 \), we see that this contradicts incentive compatibility.

* We have shown that
\[ P(r_1, r_2, \ldots, r_n) \leq r^* < r_1. \]
The last step in the proof is to assume that
\[ P(r_1, r_2, \ldots, r_n) < r^* \]
and then derive a contradiction. We can then conclude that \( P(r_1, r_2, \ldots, r_n) = r^* \) when \( r_1 > r^* \), and so the direct auction has the same pricing rule as the Vickrey auction. Consider the case in which
\[ v_1 = \frac{P(r_1, r_2, \ldots, r_n) + r^*}{2}. \]
Notice that
\[ P(r_1, r_2, \ldots, r_n) < v_1 < r^* < r_1. \]
Because \( v_1 < r^* \), bidder 1 does not trade if he reports truthfully and therefore receives a profit of 0. If he instead reports \( r_1 \), however, he wins the auction and receives
\[ v_1 - P(r_1, r_2, \ldots, r_n) > 0. \]
This contradicts incentive compatibility and completes the proof.

I won’t assign problems concerning the Vickrey auction because we worked similar problems at the beginning of the term.
0.0.14 Digression: Order Statistics.

Suppose that we draw $n$ values according to the cumulative distribution $F$. This means that the probability that any single draw is no more than some value $x$ is $F(x)$. If we take the sample of $n$ values

$$v_1, v_2, \ldots, v_n$$

and order them in a list

$$v(1) \leq v(2) \leq \ldots \leq v(n),$$

then we have created the list of order statistics for the sample. We’re interested here in such questions as, "What is the probability that the $i^{th}$ order statistic $v(i)$ is no more than some number $x$?" (i.e., what is the cumulative distribution $F_i$ of $v(i)$), and "What is the density function $f_i = F_i'$ of the $i^{th}$ order statistic?" The field of order statistics addresses this question for each of the $n$ values of $i$. In studying auctions, we are particularly interested in $F_n(x)$ (the probability that the highest value is no more than the specified number $x$) and $F_{n-1}(x)$, as $v(n)$ and $v(n-1)$ come up frequently in analyzing one’s likelihood of winning an auction (i.e., having the highest value) and the expected price that one might pay.

Let’s work a few examples to introduce this topic and to begin to feel comfortable with the mathematics. What is $F_n(x)$? We have

$$F_n(x) = \Pr \{v_n \leq x\} = \Pr \{v_1, v_2, \ldots, v_n \leq x\} = \Pr \{v_1 \leq x\} \cdot \Pr \{v_2 \leq x\} \cdot \ldots \cdot \Pr \{v_n \leq x\} = F(x)^n.$$  

The third line reflects the assumption that values are sampled independently. We therefore have

$$f_n(x) = F'_n(x) = nF(x)^{n-1}f(x).$$

I’d like to interpret this density as a probability. Interpreting

$$f_n(x) \cdot dx$$

as an approximation to the likelihood that $v(n)$ lies in a $dx$-sized interval containing $x$, we have

$$f_n(x)dx = n \cdot f(x)dx \cdot F(x)^{n-1}.$$  

There are $n$ values, and the probability that any one of them lies in the $dx$-sized interval is $f(x)dx$. The $n \cdot f(x)dx$ is therefore the likelihood that at least one of the $n$ sampled values lies in this interval. One of the $n$ values lying in this $dx$-sized interval does not make the value the largest in the sample; the other $n-1$ values must also lie below $x$. The term $F(x)^{n-1}$ is the probability that the other $n-1$ values lies below $x$. We therefore interpret the term $f_n(x)dx = n \cdot f(x)dx \cdot F(x)^{n-1}$ as the probability that at least one of the $n$ values lies in the interval times the probability that the other $n-1$ values lies below $x$.

Question: What is the expected value of the largest order statistic $v(n)$? It is

$$\int x \cdot f_n(x) \cdot dx = \int x \cdot nF(x)^{n-1}f(x)dx,$$

where the integral is evaluated over the interval in which $f(x)$ is nonzero (possibly the
whole real line).