A method of value measurement based on conditional probability theory

Gang Xi, 1 Xiao Yi Yang, 2 Ming Xi 3

Value is one of the most fundamental concepts in economics. The existing main definitions of value have certain limitations and are difficult to be unified and quantified. Thus, this article presents a method of quantifying value based on the conditional probability theory; we set value as a random variable, and thus uses conditional probability distribution to measure value. Furthermore, the mean and variance of random variables are used to describe the weighted average of the possible values and the dispersion of values distribution. This method provides a new perspective for the measurement of value.

St. Peterburg Paradox sowed a well-founded seed for the application of probability theory in economics about 300 years ago. 1 With the rising economic practices in recent years, it came with the urgent needs of understanding the inevitable laws and regularities implied in various uncertain phenomena, and discovering the possibility of various results using mathematical methods. The essential feature of modern economic theories is that it is analytical and mathematical. 2 According to the given assumptions and logical derivation, probability theory and mathematical statistics can obtain corresponding measurement conclusions, thus becoming effective tools to forecast and analyze economic operation.

“The study of how societies use scarce resources to produce valuable goods and services and distribute them among different individuals.” 3 The term value that will be referred in this context only directs to the value of commodities, including goods and services. So what is the value of commodities? How do we exactly measure it?

Throughout the history, economists concluded various definitions of value based on the criterion of their hypothesis. When Adam Smith utilizes exchange and usage as the criteria, he determined in the book The Wealth of Nation, “the word value, it is to be observed, has two different meanings, and sometimes expresses the utility of some particular objects, and sometimes the power of purchasing other goods which the possession of that object conveys.” 4 However, if the commodity has not been exchanged or used, then Smith’s theory would not fit when examining a commodity’s value. Meanwhile, Karl Marx defined value through the labour-time socially necessary for a commodity’s production; Marx wrote in the Capital “the common substance that manifests itself in the exchange-value of commodities, whenever they are exchanged, is their value” and “as values, all commodities are only definite masses of congealed labour-time.” 5 In fact, socially necessary labour time is too broad and abstract in the real world social interaction, causing it very difficult to manifest the result. It is hardly possible to examine value through socially necessary labour time as every person’s ability varies, and consequently difficult to provide a unified measure and standard. From a different angle, Alfred Marshall uses time and place as the precedent and determines value based on their relationship; his concern was “the term value is relative, and expresses the relation between two things at a particular place and time.” 6 It is without doubt that if the conditions of time and space have been met, we could determine the values of two things when successfully seeking a visible connection between them. However, if we were only using time and place as the two criteria, it would be an illogical study, indeed, an incomplete one. Since the criterion and definition of value varies in the perspectives of well-known economists throughout the time, is it possible that we could construct an uniform method, that revolves around mathematics to measure commodity’s value and eliminate any ambiguity within values? This is the purpose of this article.

As the perfection and maturity of mathematics tools such as probability theory and mathematical statistics were completed only a century ago, even if previous economists had acutely realized that value is the essential feature in all economic activities, there were no chances of applying probability theory to actual studies. Thus, all the discussions regarding value remained relatively on the abstract or philosophical level and had certain historical limitations. Reversely, in this study, we define value as random variables, and provide a unified value measurement method by using conditional probability theory and mathematical statistics. This method quantifies the measurement of value, so that we can understand the value more clearly and comprehensively, and make the measurement of value operable in economic activities.

Precondition

Almost all economic activities are conducted under certain historic and social background. When measuring their values, we need to take into consideration of their context and conditions, including but not limited to scarcity, homoeconomicus or rational man, time, place, usage, exchange, labour, and conditions, etc. Whether consciously or not, when we are provided with certain conditions, we need to base them on the existing criteria, the value’s historical records, quantitative statistics and human experience to derive a probability distribution of values interval and provide an accurate or approximate mathematical description.

For instance, Smith considers “nothing is more useful than water; but it will purchase scarce anything; scarce anything can be had in exchange for it. A diamond, on the contrary, has scarce any value in use; but a very great quantity of other goods may frequently be had en exchange for it.” 7 Indeed, when observing the value over water and diamonds, the default is comparing the value under urban conditions, rarely under special conditions such as in the desert. Normally, people have the consensus that a diamond’s value is much more than water. However, in special circumstances, the value relationship between the two items could change drastically. In the probability theory perspective, this could be seen as a rare event. When people compare water and diamond’s values, the hidden precedent is that people are in a familiar place, which is the big probability condition. But what if the environment that the commodity is situated becomes a desert where human rarely lives, which is the small probability condition, the water’s value would thrust up the roof. In the probability theory viewpoint, when a rare event happens on the population distribution interval, the random variable in this rare event’s probability distribution is greater than the probability in population distribution.

1EAL, Vancouver Community College, 1155 East Broadway, Vancouver, B.C. V5T 4V5, Canada. 2School of Economics & Management, Xidian University, 266 Xinglong Section of Xifeng Road, Xi’an, Shaanxi 710126, China. 3School of Humanities, Shenzhen Polytechnic, Liu Xian Avenue 7098, Nanshan Distrect, Shenzhen, Guangdong, 518055, China. Correspondence and requests for materialsshould be addressed to G.X. (email: carlx@live.com)
and some condition’s probability distribution would way exceed the population one.

In fact, the probability theory explanation of various problems in economics is the conditional probability explanation based on the basic hypothesis. Therefore, before discussing the preconditions of value measurement, we need to understand the main hypotheses of economics.

**Hypothesis of Scarce Resource.** In the event of extremely abundant resources, “like sand in desert or seawater at the beach,” each person’s disposable resources have mere differences, and the produced commodities should have no difference either. Therefore, it is not necessary to use economic behaviours to distribute commodities among people. This hypothesis supports the idea that scarce resource always exists when we think about value distribution. If this idea does not exist, then all commodities’ value would be identical, and no consideration is needed.

**Hypothesis of Homoeconomicus or Rational Man.** If everyone is aiming for profit maximization, then economic activities for the society would have motive and goals. As Smith describes “it is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities but of their advatages.” This is saying that only this kind of people are rational or economical man; otherwise, they are irrational. The hypothesis of rational man is based on the reality of human economic interaction, which is a concrete practice that fit human descriptions. If we have validated this hypothesis, then the measurement of value in economic practices would not stem and value would not be the centre of economic activities; no one nor the society would care about the size of value anymore. The concept of value would lose its needs for research or measurement. Therefore this hypothesis is necessary in order to conduct the study on measuring value.

In economics, we define value as a random variable, in other words, the value numeric is a random variable. \( P(E) \) represents the probability that the commodity’s value is \( E \), that is to say, the degree of confirmation or statistical frequency that the commodity’s value is \( E \). In addition, the various preconditions of value are collectively referred to as \( F \). Certainly, \( F \) may be a compound condition, that is, a combination of multiple conditions, which is called condition \( F \). Furthermore, condition \( F \) must include the hypotheses of resource scarcity and homoeconomicus. These two conditions are only regarded as implicit conditions when considering the value measurement, and are often not specially marked.

In order to facilitate the discussion of value measurement, the classical definition of probability is given: (See reference 7 for specific definitions)

We shall assume that for each event \( E \) of an experiment having a sample space \( S \), there is a number, denoted by \( P(E) \), that is in accord with the following three axioms.

**AXIOM 1**

\[
0 \leq P(E) \leq 1
\]

**AXIOM 2**

\[
P(S) = 1
\]

**AXIOM 3**

For any sequence of mutually exclusive events \( E_1, E_2, \ldots \) (that is, events for which \( E_iE_j = \emptyset \) when \( i \neq j \)),

\[
P(U_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i), \quad n = 1, 2, \ldots, \infty
\]

We call \( P(E) \) the probability of the event \( E \).

Thus, Axiom 1 states that the probability that the outcome of the experiment is contained in \( E \) is some number between 0 and 1. Axiom 2 states that, with probability 1, the outcome will be a member of the sample space \( S \). Axiom 3 states that for any set of mutually exclusive events the probability that at least one of these events occurs is equal to the sum of their respective probabilities.

The relationship between two random variables can often be clarified by consideration of the conditional distribution of one given the numeric of the other.

Recall that for any two events \( E \) and \( F \), the conditional probability of \( E \) given \( F \) is defined, provided that \( P(F) > 0 \), by

\[
P(E \mid F) = \frac{P(EF)}{P(F)}
\]

**Methods**

We use the conditional probability distribution of random variables to measure the value numeric of commodity and joint commodity. For the sake of clarity, we give examples of commodity \( C1 \) and \( C2 \) and the joint commodity \( C1 + C2 \).

**Value Variables.** In this paper, we use discrete random variables to represent random variables of value (value variables). For a discrete random variable \( X \), we define the probability mass function \( p(a) \) of \( X \) by

\[
p(a) = P(X = a)
\]

(1)

The probability mass function \( p(a) \) is positive for at most a countable number of numerics of \( a \). That is, if \( X \) must assume one of the numerics \( x_1, x_2, \ldots \), then

\[
p(x_i) > 0, \quad i = 1, 2, \ldots
\]

\[
p(x) = 0, \quad \text{all other values of } x
\]

Since \( X \) must take on one of the numerics \( x_i \), we have
\[ \sum_{i=1}^{\infty} P(x_i) = 1 \]

symbolized by \( X \), also known as the \( X \)'s sequence of distribution of all possible numeric \( x_1, x_2, \ldots, x_n, \ldots \), \( p_1, p_2, \ldots \), \( p_n, \ldots \)'s \( X \)'s distribution

\[ p_n = P(X = x_n) \quad (n = 1, 2, \ldots) \quad (2) \]

The most precise way to determine a population’s conditional distribution is taking into accounts of all elements in the population. However, we know in the real social and economic activities for most cases, due to the limitation of cost and other realistic conditions, it is impossible and unnecessary to collect the values of all elements. We can only collect a range of elements for approximate measurement. In experiment, we usually collect sampling statistics, which are obtained from various sources to form a sampling distribution. When samples are surveyed, the observed sample distribution can be used to approximate the population distribution. There are mature methods to determine the accuracy of approximation in mathematical statistics, so we will not reiterate the specifics here.

As shown below, Table 1 and 2 are frequency tables for a data set consisting of C1 and C2’s value numeric. \( p_X \) and \( p_Y \) are the probability mass functions of \( X \) and \( Y \). \( X \) and \( Y \) are the value variables of commodity C1 and C2, such as -2, 0, 5, etc. \( p_n \) is the probability, when value numeric is \( x_n \), such as the probability is 1/32, when that C1’s value numeric is 1.

Table 1 Distribution series of C1’s value numeric

| \( X \) | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 |
|------|----|----|---|---|---|---|---|---|
| \( P(X = x_n) \) | 1/32 | 1/32 | 1/16 | 1/8 | 1/4 | 1/8 | 1/16 | 1/16 |

Table 2 Distribution series of C2’s value numeric

| \( Y \) | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|------|----|---|---|---|---|---|---|---|
| \( P(Y = y_k) \) | 1/64 | 1/64 | 1/32 | 1/16 | 1/8 | 1/4 | 1/4 | 1/8 |

For straightforward illustration, we use the following histogram to portray C1’s value numeric distribution (when under certain conditions; we could ignore the preceding conditions when they are known to people, and they do not affect our discussion of relevant problems.)

![Histogram of C1 and C2 distributions](image)

**Figure 1.** All possible value numeric of C1’s value variable \( X \) is -2, -1, 0, 1, 2, 3, 4, 5 and corresponding probability distribution \( P(X) \) is 1/32, 1/32, 1/16, 1/8, 1/2, 1/8, 1/16, 1/16. All possible value numeric of C2’s value variable \( Y \) is -1, 0, 1, 2, 3, 4, 5, 6 and corresponding probability distribution \( P(Y) \) is 1/64, 1/32, 1/32, 1/16, 1/4, 1/4, 1/4, 1/8.

We could observe that C1’s minimum value numeric is -2, maximum value numeric is 5, smallest probability is 1/32, and biggest probability is 1/2. C2’s minimum value numeric is -1, maximum value is 6, the smallest probability is 1/64, the biggest probability is 1/4. In real economic activities, just illustrating the probability distribution is not sufficient enough to discuss the problem. Hence, we need to introduce a few numerical features of the probability distribution of random variables, and these features are beneficial in terms of enhancing the understanding towards value.

**Mean measure.** In probability theory, when we discuss the numerical value of a random variable, we need to grasp its numerical features such as mathematical expectation and variance, etc. These could reflect the random variable’s certain characteristics in order to better comprehend and compare the random variables. For instance, mathematical expectation illustrates a random variable’s mean, so called the expectation or mean. When we measure the value distribution’s mean value, it is the population mean, or expectation. In other words, the expected value of \( X \) is a weighted average of the possible values that \( X \) can take on, each value being
weighted by the probability that $X$ assumes it.

If $X$ is a discrete random variable taking on the possible numerics $x_1, x_2, \ldots$, then the expectation or expected numeric of $X$, denoted by $E[X]$, is defined by

$$E(X) = \sum_i x_i P(X = x_i)$$

(3)

Computer C1’s mean $E(X)$ and C2’s mean $E(Y)$:

$E(X) = \frac{31}{32}$

$E(Y) = \frac{57}{64}$

The presented data more accurately reflect the recognition level or confirmation degree of C1 and C2 values. By comparing the mean values, we can see that under normal conditions, the average value of C1 is 1.96875, which is less than that of C2 3.890625. It can also be said that under given conditions, the value of C2 is greater than that of C1. The value size here usually refers to the average value; of course, this does not mean that the value of C2 must be greater than C1.

Variance Measure. We use variance to inspect the stability or degree of dispersion of individuals to the value of a commodity. Variance reflects the characteristic of the dispersion of any random variables around the mean. From knowing the features of the random variable, we could compare the measures of dispersion of commodities’ value from comparing probability functions of different random variables.

Variance is to describe the degree of sample separation by calculating the distance between the sample and the average value. The greater the difference between the data, the greater the variance; vice versa, the more convergence between the data, the smaller the variance. If the variance is zero, the value will be scalar. There will be no difference in the value in economic activities, and the value measurement will lose its practical significance. We use variance to investigate the dispersion degree of commodity values relative to the mean value, so as to compare the concentration or dispersion of different commodity values.

If $X$ is a random variable with mean $\mu$, then the variance of $X$, denoted by $\text{Var}(X)$, is defined by

$$\text{Var}(X) = E[(X - \mu)^2]$$

An alternative formula for $\text{Var}(X)$ can be derived as follows:

$$\text{Var}(X) = E[X^2] - E[(X)^2]$$

(4)

Compute $\text{Var}(X)$ or $\text{Var}(Y)$ when $X$ or $Y$ represents the outcome of C1 or C2’s value variable.

$$\text{Var}(X) = \frac{95}{1024} = \frac{380}{4096}$$

$$\text{Var}(Y) = \frac{655}{4096}$$

Jointly Value Measure. Real economic activities are complicated and the combined value of two or more commodities should also be considered. Here, we consider the combination value measure of two commodities, and the combination of multiple commodities can be analogized. Therefore, it is necessary to introduce the definition of two-dimensional joint probability function.

For a given experiment, we are often interested not only in the probability distribution functions of individual random variables but also in the relationships between two or more random variables.

To specify the relationship between two random variables, we define the joint cumulative probability distribution function of
\( X \) and \( Y \) by

\[
F(x, y) = P[X \leq x, Y \leq y]
\]

In the case where \( X \) and \( Y \) are both discrete random variables whose possible values are, respectively, \( x_1, x_2, \ldots, \) and \( y_1, y_2, \ldots \), we define the joint probability mass function of \( X \) and \( Y \), \( p(x, y) \), by

\[
p(x_i, y_j) = P[X = x_i, Y = y_j]
\]

The random variables \( X \) and \( Y \) are said to be independent if for any two sets of real numbers \( A \) and \( B \)

\[
P(X \in A, Y \in B) = P(X \in A)P(Y \in B) \tag{5}
\]

In other words, \( X \) and \( Y \) are independent if, for all \( A \) and \( B \), the events \( E_A = \{X \in A\} \) and \( E_B = \{Y \in B\} \) are independent.

When \( X \) and \( Y \) are discrete random variables, the condition of independence Equation (5) is equivalent to

\[
p(x, y) = p_X(x)p_Y(y) \quad \text{for all} \ x, y \tag{6}
\]

where \( p_X \) and \( p_Y \) are the probability mass functions of \( X \) and \( Y \). The equivalence follows because, if Equation (5) is satisfied, then we obtain Equation (6) by letting \( A \) and \( B \) be, respectively, the one-point sets \( A = \{x\} \), \( B = \{y\} \).

We use \( C_1 + C_2 \) to represent the combination of commodity \( C_1 \) and commodity \( C_2 \), because the value variable of commodity \( C_1 \) and the value variable of commodity \( C_2 \) are independent of each other:

Table 3 Joint distribution series of \( C_1 \)'s and \( C_2 \)'s value numeric

| \( P(x_n, y_k) \) | -2 | -1 | 0  | 1  | 2  | 3  | 4  | 5  | 6  |
|------------------|----|----|----|----|----|----|----|----|----|
|                  | 1/2048 | 1/2048 | 1/1024 | 1/512 | 1/128 | 1/128 | 1/256 | 1/512 | 1/1024 |
| -1               | 1/2048 | 1/2048 | 1/1024 | 1/512 | 1/128 | 1/128 | 1/256 | 1/512 | 1/1024 |
| 0                | 1/1024 | 1/1024 | 1/512 | 1/256 | 1/128 | 1/128 | 1/256 | 1/512 | 1/1024 |
| 1                | 1/2048 | 1/2048 | 1/1024 | 1/512 | 1/128 | 1/128 | 1/256 | 1/512 | 1/1024 |
| 2                | 1/128  | 1/128  | 1/64  | 1/32  | 1/16  | 1/8  | 1/4  | 1/2  | 1   |
| 3                | 1/2048 | 1/2048 | 1/1024 | 1/512 | 1/256 | 1/128 | 1/64  | 1/32  | 1/16  |
| 4                | 1/1024 | 1/1024 | 1/512 | 1/256 | 1/128 | 1/64  | 1/32  | 1/16  | 1/8   |
| 5                | 1/2048 | 1/2048 | 1/1024 | 1/512 | 1/256 | 1/64  | 1/32  | 1/16  | 1/8   |

According to the joint distribution of value variables of commodity \( C_1 \) and commodity \( C_2 \), the probability distribution of value variables of \( C_1 + C_2 \) can be calculated as follows:

Table 4 Distribution series of \( C_1+C_2 \)'s value numeric

| \( X+Y \) | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|----------|----|----|----|---|---|---|---|---|---|---|---|---|---|---|---|
| \( P(X+Y) \) | 1/2048 | 1/1024 | 5/2048 | 3 | 5/12 | 5/12 | 1/256 | 1/128 | 1/64 | 1/32 | 1/16 | 1/8 | 3/64 | 3/128 | 1/128 |

**Figure 3.** All possible value numeric of \( C_1+C_2 \)'s value variable is -3, -2, -1, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and corresponding probability distribution is 1/2048, 1/1024, 5/2048, 3/512, 11/512, 17/512, 59/1024, 49/512, 191/1024, 103/512, 49/256, 1/8, 3/64, 3/128, 1/128. \( C_1+C_2 \)'s smallest probability is 1/2048, as value numeric is -3; biggest probability is 103/512, as value numeric is 6.
Computer E(\(X\)) 1.96875 + E(Y) 3.890625 = E(\(X + Y\)) 5.859375, when \(C1 + C2\), greater than the individual mean of \(C1\) and \(C2\); Var(\(X\)) 2.092773 + Var(Y) 2.159912 = Var(\(X + Y\)) 4.252685, greater than the individual variance of \(C1\) and \(C2\).

Discussion
To sum up, the value variable is a discrete random variable. If all situations are taken into account, the value variable will definitely occur, and the probability is 1, that is 100%. In fact, a value variable can not necessarily occur, there is only the probability of occurrence. Therefore, we can measure the probability of the occurrence of value variables of any commodity and commodity combination by collecting data through sampling surveys, and obtain the probability distribution function of the relevant commodities under certain conditions. In reality, the comparison of value means is difficult and can not accurately reflect the size of value variables. We can also measure the concentration of commodity value distribution by variance. The larger the variance, the more speculative it is; the smaller the variance is, the stronger the stability. Analogous to the mean’s being the center of gravity of a distribution of value mass, the variance represents the risk. The mean and variance of portfolio value can be jointed. After suming, the mean and variance increase, that is, the value increases and the speculation increases.

There are various factors and influences that could determine value. We would not take into consideration the detailed element analysis that affects value here, as we would never be able to list them explicitly. We are only considering the probabilistic description of value; this could be influenced by the probability of the degree of recognition that people have towards objects and values. By doing this, the examination of value would not be constrained by certain factors, making it more distinct and clearer to measure value.

As the cornerstone of economics' concept, the essence of value has been explored and studied by many historical economists. However, due to the limitations of historical knowledge and mathematical tools, only some empirical descriptions are made, not quantitative descriptions. According to the theory of conditional probability, this paper constructs a value measurement method for straightforward operation. With the general assumption of economics, this method transforms the problem of value measurement in economics into conditional probability distribution function in mathematics. The parameters of this probability distribution are calculated by collecting data through sampling surveys. The mean value is used to measure the comparative value, and the variance is used to evaluate the dispersion of value. This paper theoretically shows that the value of any commodities can be measured, and the value of different commodities can be compared and combined. J. M. Keynes once remarked that "the economics theory does not furnish a body of settled conclusions immediately applicable to policy. It is a method rather than a doctrine, an apparatus mind, a technique of thinking, which helps its possessor to draw conclusions." Our value measurement method evocatively develops and enriches the understanding of value, which is a thinking technique. In particular, it has an extensive application prospect in the digital economy, fictitious economy, and other fields promoted by the new generation technology represented by blockchain technology, etc.

References
1. George Zsipo. Value Judgments. 29 August 2013, Vol 500, Nature. 521, https://www.nature.com/articles/500521a (2013).
2. Akira Takayama. Mathematical Economics. The Dryden Press, 3, ISBN 0-03-086653-7 (1974).
3. Paul A. Samuelson, William D. Nunlhaus. Economics. 18th Edition, McGraw-Hill Companies, Inc. 4, ISBN 0-07-287205-5 (2005).
4. Adam Smith. An inquiry into the nature and causes of the wealth of nations. Liberty Classics, 12, 23, 25 (1776).
5. Karl Marx. Capital. (Volume I ), Progress publishers, Moscow, USSR, 28, 29 First English edition of 1887 (4th German edition changes included as indicated) (1887).
6. Alfred Marshall. Principles of Economics. (London: Macmillan and Co. 8th ed. 1920), 43, http://oll.libertyfund.org/title/1676 (1890).
7. Sheldon M. Ross. Introduction to Probability and Statistics for Engineers and Scientists. 3rd Edition, Elsevier Academic Press, 59-121, ISBN 0-12-598057-4 (2004).
8. J. M. Keynes. “Introduction” (to the series), Cambridge Economic Handbooks (the first book in the series, Supply and Demand by H. D. Henderson, appeared in 1922, published by Harcourt Brace and Co.).

Acknowledgements
We are very grateful to Zhuqing Huang (Qingdao), who was the first to ask what is value with us. We have had many a time discussions over the last six months. In particular, we thank Yinglin Xi (Vancouver) for translating. We thank the referees for their constructive comments that helped improve this paper.

Author Contributions
Gang Xi and Ming Xi initiated the research the project and wrote the paper. Xiaoyi Yang performed analytical calculations.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to G. X.
Reprints and permissions information is available at www.nature.com/reprints.
Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.