A Computational Study on the Entropy of Interval-Valued Datasets from the Stock Market

Chenyi Hu[1(✉)] and Zhihui H. Hu[2]

[1] University of Central Arkansas, Conway, AR, USA
`chu@uca.edu`
[2] Edge Pursuit LLC, San Francisco, CA, USA

Abstract. Using interval-valued data and computing, researchers have reported significant quality improvements of the stock market annual variability forecasts recently. Through studying the entropy of interval-valued datasets, this work provides both information theoretic and empirical evidences on that the significant quality improvements are very likely come from interval-valued datasets. Therefore, using interval-valued samples rather than point-valued ones is preferable in making variability forecasts. This study also computationally investigates the impacts of data aggregation methods and probability distributions on the entropy of interval-valued datasets. Computational results suggest that both min-max and confidence intervals can work well in aggregating point-valued data into intervals. However, assuming uniform probability distribution should be a good practical choice in calculating the entropy of an interval-valued dataset in some applications at least.

Keywords: Interval-valued dataset · Stock market variability forecasting · Data aggregation · Probability distribution · Information entropy

1 Introduction

Recently, researchers have very successfully applied interval-valued data in information processing and uncertainty management. Related works on applications of interval-valued data include [13, 21–25], and many more. With broad applications of interval computing, the IEEE Standard Association has released the IEEE Standards for Interval Arithmetic [19] and [20] recently.

This work is a continuation of the stock market interval-valued annual variability forecasts reported in [10, 11, 13, 14, 16], and [17]. In which, a real world six-dimensional point-valued monthly dataset is first aggregated into an interval-valued annual sample. Then, interval-valued annual predictions are made with interval least-squares (ILS) regression [15]. Comparing against the commonly used point-valued confidence interval predictions with ordinary least-squares (OLS), the interval approach increased the average accuracy ratio of annual stock market forecasts from 12.6% to 64.19%, and reduced the absolute mean error from 72.35% to 5.17% [14] with the same economical model [4] and the same raw dataset.

The quality improvements are significant. However, several questions arising from previous results still need to be answered. Among them are:

1. What is the theoretic reason for such a significant quality improvements?
2. What are the impacts of data aggregation methods on the results? and
3. What are the impacts of probability distributions on the entropy of an interval-valued dataset?

In this paper, we investigate these questions from the perspective of information theory [9]. To be able to calculate and compare entropies of interval-valued datasets, it is necessary to establish the concepts and algorithms on probability and entropy for interval-valued datasets. In our work [18], also published in this volume, we lay down both theoretic and algorithmic foundations for the investigation reported in this work. In which, point-valued statistic, probabilistic, and entropy measures for interval-valued datasets are established in details with practical algorithms. Interested readers should refer that article for a solid theoretical foundation.

In the rest of this paper, we briefly review related previous work, such as the stock market annual variability forecasting, the dataset, and information entropy in Sect. 2. We try to answer the question why interval-valued data leading better quality forecasts through comparing information entropy of interval-valued samples against point-valued ones in Sect. 3. We calculate and compare the impacts of two aggregation methods (min-max and confidence intervals) associated together with commonly used probability distributions (uniform, normal, and beta) in Sect. 4. We summarize the main results and possible future work in Sect. 5.

2   Related Previous Works

We first briefly review the dataset and the stock market annual variability forecasts; and then introduce related concepts and algorithms of calculating entropies of a point-valued dataset and of an interval-valued dataset.

2.1   The Stock Market Annual Variability Forecasting and the Dataset

The S & P 500 index is broadly used as a indicator for the overall stock market. The main challenge in studying the stock market is its volatility and uncertainty. Modeling the relationship between the stock market and relevant macroeconomic variables, Chen, Roll, and Ross [4] established a broadly accepted model in economics to forecast the overall level of the stock market. According to their model, the changes in the overall stock market value (SP_t) are linearly determined by the following five macroeconomic factors:

IP_t: the growth rate variations of adjusted Industrial Production Index,
DI_t: changes in expected inflation,

UI_t: and changes in unexpected inflation,
DF_t: default risk premiums, and
TM_t: unexpected changes in interest rates.

This relationship can be expressed as:

$$SP_t = a_t + I_t(IP_t) + U_t(UI_t) + D_t(DI_t) + F_t(DF_t) + T_t(TM_t) \qquad (1)$$

By using historic data, one may estimate the coefficients of (1) to forecast changes of the overall stock market. The original dataset used in [14] and [17] consists of monthly data from January 1930 to December 2004 in 75 years for the six variables. Here are few sample lines of the data:

```
Yr-mth     UI          DI       SP           IP            DF       TM
30-Jan -0.00897673  0        0.014382062 -0.003860512   0.0116 -0.0094
30-Feb -0.00671673 -0.0023   0.060760088 -0.015592832  -0.0057  0.0115
30-Mar -0.00834673  0.0016   0.037017628 -0.00788855    0.0055  0.0053
30-Apr  0.00295327  0.0005   0.061557893 -0.015966279   0.01   -0.0051
30-May -0.00744673 -0.0014  -0.061557893 -0.028707502  -0.0082  0.0118
30-Jun -0.00797673  0.0005  -0.106567965 -0.046763234   0.0059  0.0025


...     ......      ...      ...          ...            ...      ...


04-Jul -0.00182673  0.0002  -0.024043354  0.00306212    0.0029  0.0147
04-Aug  0.00008127  0.0002  -0.015411102 -0.002424198   0       0.0385
04-Sep  0.00156327  0.0001   0.026033651  0.007217235   0.0005  0.0085
04-Oct  0.00470327  0        0.000368476  0.002001341   0.001   0.0143
04-Nov -0.00002273  0        0.044493038  0.006654848   0.0034 -0.0245
04-Dec -0.00461673  0.0004   0.025567309  0.001918659   0.0007  0.0235
```

To make an annual stock market forecast, a commonly used approach is to make a point-valued annual sample first, such as the end of each year, i.e., December data, or annual minimum for predicting the min, or annual maximum for estimating the max. Applying OLS to estimate the coefficients in (1), people are able to make a point-valued prediction. By adding and subtracting a factor (usually denoted as Z) of the standard deviation to the point-valued prediction, one form a confidence interval as an annual variability forecast. However, such confidence interval forecasting methods have never been widely used in the literature because of the poor forecasting quality [2] and [7] in forecasting the stock market. Normally, the forecasting intervals are so narrow that there is only a 50% chance, or even less, that a future point lies inside the interval [5] and [6]. In other cases, the forecasting intervals can be so wide that the forecasts are meaningless. This poor forecasting quality is deeply rooted in the methodology of point-based confidence interval forecasting.

Instead of commonly used point-valued approach, an interval-valued method has been proposed and applied for the annual stock market variability forecasts [14]. In which, the annual minimum and maximum form an interval-valued (min-max) sample of the year. By applying an interval least-squares algorithm [13]

with the interval-valued sample, significant quality improvements of predictions are obtained. Figure 1 illustrates the interval-valued annual forecasts comparing against the actual variations of S & P 500 from 1940–2004. In which, a ten-year sliding window was used to make an out of sample forecast.
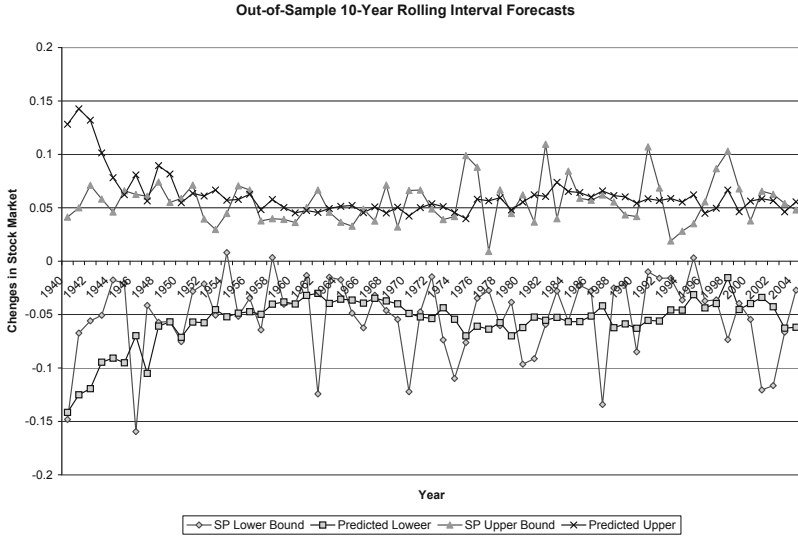

Fig. 1. Annual interval forecasts vs. actual market variations from 1940–2004

Further studies on forecasting the stock market [10] and [11], variability of mortgage rates [12], crude oi price prediction [29], and others, have consistently reported that the quality of variability forecasts with interval-valued samples and interval least-squares are significantly better than that of with point-valued samples and OLS.

As the main objective of this work, we want to investigate the major reason for such significant quality improvements through computing and comparing the entropies of point- and interval-valued samples.

2.2   Information Entropy of a Point- and an Interval-Valued Dataset

Our investigations are carried out through calculating and comparing information entropy, i.e., the average rate at which information produced by a stochastic source of data [28].

Shannon defines the entropy for a discrete dataset $X = \{x_1, x_2, \ldots, x_n\}$ in his seminal paper "A mathematical theory of communication" [26] as:

$$H(x) = -\sum_{i=1}^{n} p(x_i) \log p(x_i) \tag{2}$$

where $p(x_i)$ is the probability of event x_i. In information theory, Shannon's entropy has been referred as information entropy, and it has been used as a measure of information in data. Viewing the stock market as a stochastic source of data, we try to measure and compare the amount of information contained in datasets.

For a point-valued dataset X, we may estimate its entropy practically with the algorithm below:

Algorithm 1: (*Calculating entropy of a discrete point-valued dataset*)

```
Input: a point-valued length n dataset X
    m, an integer for the number of bins in constructing a histogram
Output: entropy(X)

# Finding empirical probability of the dataset X
# Constructing an m-bin histogram for X
For i from 1 to m:
    c_i = the frequency count of x in the i-th bin
    p_i = c_i/|X|
# Calculating the entropy of X
entropy(X) = 0 # initialization
For i from 1 to m:
    entropy(X)  - = p_i log p_i.

Return entropy(X)
```

Applying available software tools, one can easily implement the steps in Algorithm 1 above. For example, calling the `histogram` method in Python `numpy` module returns the counts and bins in a histogram of a dataset. The rests are straightforward to implement.

However, it is not that straightforward to calculate information entropy of an interval-valued dataset. By the term interval, we mean a connected subset of $\mathbb{R}$. An interval-valued dataset is a collection of intervals. Using a boldfaced lowercase letter to denote an interval, and a boldfaced uppercase letter to specify an interval-valued dataset, we have $\mathbf{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ as an interval-valued dataset consisting of n intervals $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_n$. Applying (2) to calculate the entropy of $\mathbf{X}$ demands a probability distribution of $\mathbf{X}$. Our paper [18] provides the theoretic and algorithmic foundations needed for calculating a point-valued probability of an interval-valued dataset. For readers' convenience, here are two related definitions and a theorem from that paper:

Definition 1. *A function $f(x)$ is called a probability density function, pdf of an interval-valued dataset $\mathbf{X}$ if and only if $f(x)$ satisfies all of the conditions:*

$$\begin{cases} f(x) \geq 0 \ \forall x \in (-\infty, \infty); \\ \sum_{i=1}^{n} \int_{\boldsymbol{x}_i \in \boldsymbol{X}} f(t)dt = 1. \end{cases} \tag{3}$$

Using pdf_i to denote the probability density function for $\boldsymbol{x}_i \in \mathbf{X}$, we have the theorem below to obtain a *pdf* for $\mathbf{X}$ practically.

Theorem 1. *Let* $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ *be an interval-valued dataset; and* $pdf_i(x)$ *be the pdf of* $\boldsymbol{x}_i$ *provided* $i \in \{1, 2, \ldots, n\}$. *Then,*

$$f(x) = \frac{\sum_{i=1}^{n} pdf_i(x)}{n} \tag{4}$$

is a pdf of $\mathbf{X}$.

With (4), we define the entropy for an interval-valued dataset $\mathbf{X}$ as

Definition 2. *Let* $\mathcal{P}$ *be an interval partition of the real axis and* $pdf(x)$ *be the probability density function of* $\mathcal{P}$. *Then, the probability of an interval* $\boldsymbol{x}^{(j)} \in \mathcal{P}$ *is* $p_j = \int_{\boldsymbol{x}^{(j)}} pdf(t)dt$, *and the entropy of* $\mathcal{P}$ *is*

$$entropy(X) = -\sum_{\mathcal{P}} p_j \log p_j \tag{5}$$

Example 1. Find a *pdf* and entropy for the interval-valued sample dataset $\mathbf{X}_0 = \{[1, 5], [1.5, 3.5], [2, 3], [2.5, 7], [4, 6]\}$.

For simplicity, we assume a uniform distribution for each $\boldsymbol{x}_i \in \mathbf{X}_0$, i.e.,

$$pdf_i(x) = \begin{cases} \dfrac{1}{\overline{x}_i - \underline{x}_i} & \text{if } x \in \boldsymbol{x}_i, \text{ and } \underline{x}_i \neq \overline{x}_i \\ \infty & \text{if } \underline{x}_i = \overline{x}_i \\ 0, & \text{otherwise.} \end{cases}$$

$$f(X_0) = \frac{\sum_{i=1}^{5} pdf_i(x)}{5} = \begin{cases} 0.05 & \text{if } x \in [1, 1.5] \\ 0.15 & \text{if } x \in (1.5, 2] \\ 0.35 & \text{if } x \in (2, 2.5] \\ 0.39 & \text{if } x \in (2.5, 3] \\ 0.19 & \text{if } x \in (3, 3.5] \\ 0.09 & \text{if } x \in (3.5, 4] \\ 0.19 & \text{if } x \in (4, 5] \\ 0.14 & \text{if } x \in (5, 6] \\ 0.044 & \text{if } x \in (6, 7] \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

The *pdf* of the example in (6) is a stair function. This is because of the uniform distribution assumption on each $\boldsymbol{x}_i \in \mathbf{X}_0$. The five intervals in $\mathbf{X}_0$ form a partition of $\mathbb{R}$ in eleven intervals including $(-\infty, 1)$ and $(7, \infty)$. Using (5), we have the entropy of the interval-valued sample dataset $entropy(X_0) = 2.019$ □

Example 1 illustrates the availability of a point-valued *pdf* for an interval-valued dataset. For more theoretic and algorithmic details, please refer [18]. We are ready now to investigate the question: why does the interval-valued approach significantly improve the quality of variability forecasts?

3   Why Does the Interval-Valued Approach Significantly Improve the Quality of Variability Forecasts?

Previous results have evidenced that the interval-valued approach can significantly improve the quality of forecasts in different areas (such as the stock market annual variability, the variability of the mortgage rate [12], and the variability of crude oil price [15]). However, using the same economical model and the same original dataset but point-valued samples, the quality of forecasts are much worse. To investigate the possible cause, we should examine the entropies of interval-valued and point-valued input datasets evidently.

Applying Algorithm 1 on point-valued annual samples of the six-dimensional financial dataset, we calculate their attribute-wise entropy. The four point-valued annual samples are December only, annual minimum, annual maximum, and annual midpoint[1]. With the Algorithm 3 in [18], we calculate the attribute-wise entropy of the annual min-max interval-valued sample. Table 1 summarizes the results. In which, the first row lists each of the six attributes in the dataset. The second to the last rows provide values of attribute-wise entropy of five different samples: December only, Annual minimum, Annual maximum, Annual midpoint, and Annual min-max interval, respectively.

Table 1. Entropy comparisons of different samples

	UI	DI	SP	IP	DF	TM
December only	2.32855	2.01183	2.12941	2.05978	2.39706	2.33573
Annual minimum	2.33076	2.16933	2.28035	2.09871	2.19422	2.62452
Annual maximum	1.88469	2.30266	1.53328	1.88045	2.34693	2.35843
Annual mean	2.04877	2.55961	2.31651	2.07323	2.09817	2.47341
Annual min-max intvl.	4.34192	3.06851	3.95838	4.30213	3.95359	4.31941

Figure 2 provides a visualized comparison of these entropy. From which, we can observe the followings:

The attribute-wise information entropies vary along with different samples. However, the attribute-wise entropies of the interval-valued sample are clearly much higher than that of any point-valued ones. Comparatively, the entropies of point-valued samples do not differ significantly. This indicates that the amount of information in these point-valued samples measured with entropies are somewhat similar. But, they are significantly less than that of the interval-valued ones. The greater the entropy is, the more information may possibly be extracted from. This is why the interval-valued forecasts can produce significantly better forecasts in [10,11,14], and others.

In both theory and practice, meaningless noises and irregularities may increase the entropy of a dataset too. However, it is not the case here in this study.

---

[1] The arithmetic average of annual min and annual max.
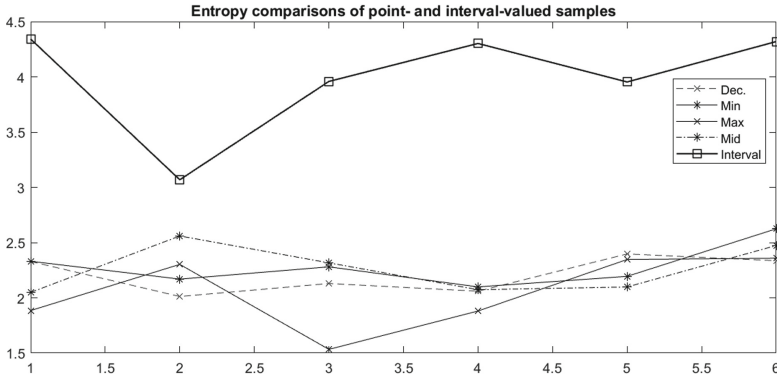
Fig. 2. Attribute-wise entropy comparison of point- and interval-valued samples

The interval rolling least-squares algorithm [16] has successfully extracted the additional information and made significant quality improvements. The advantages of using interval-valued samples instead of point-valued ones have also been observed in predicting variations of the mortgage rate [12], the crude oil price [15], and others. The interval-valued samples indeed contain more meaningful information. Therefore, in making variability forecasts like the stock market, it is preferable of using interval-valued samples rather than point-valued ones.

Here is an additional note. The attribute-wise entropies of the annual min-max interval-valued sample in Table 1 and the sum of entropies of the point-valued annual minimum and maximum are similar. If one uses the point-valued annual minimum and annual maximum separately, can he obtain quality forecasts similar to that of using the min-max interval-valued sample? Unfortunately, an empirical study show that is not the case. In [11], a comparison of the following two approaches is reported. One of the two is of applying the point-valued annual minimum and maximum samples to predict annual lower and upper bounds of the market with the OLS separately. Then, confidence intervals are constructed as annual variability forecasts. The other applies the ILS with the min-max interval-valued sample. The quality of forecasts produced in the later approach is still much better than that of the former approach. In [10], using the sample of annual midpoints is studied for the same reason of performance comparison. The ILS with interval-valued annual sample still significantly outperform the point-valued approach in terms of higher average accuracy ratio, lower mean error, and a higher stability in terms of less standard deviation. This suggests that, to extract information from an interval-valued sample, one should use the ILS instead of OLS.

4   Impacts of Data Aggregation Strategies and Probability Distributions on Calculating the Entropy of an Interval-Valued Dataset

Yes, an interval-valued sample may contain more information than a point-valued sample does. But, there are various strategies, such as in [1,8] and others, to aggregate data other than the min-max method. What are the impacts of different aggregation strategies on the entropy of resulting interval-valued dataset? Furthermore, in calculating the entropy of an interval-valued dataset, Eq. (4) requires the pdf_i for each $\boldsymbol{x}_i \in \mathbf{X}$. What are the impacts of these pdf_is on calculating the entropy of $\mathbf{X}$? We now investigate these two questions computationally again.

In studying probability distribution of interval-valued annual stock market forecasts, point-valued data are aggregated with confidence intervals instead of annual min-max intervals [17]. In which, the points within a year are first fit with a normal distribution attribute-wise. Then, confidence intervals are formed at a selected level of probabilistic confidence with an intention of filtering out possible outliers. With different levels of confidence (by adjusting the Z-values), the interval-valued samples vary. So do the variability forecasts. However, we have observed that the variations are not very significant at all when Z is between 1.25 to 2, see [17]. Specifically, the average accuracy ratio associated with the Z-values are: 61.75% with $Z = 1.25$, 64.23% with $Z = 1.50$, 64.55% with $Z = 1.75$, and 62.94% with $Z = 2.00$. These accuracy ratios are very similar to 64.19% reported in [14] with the min-max aggregation.

In calculating the attribute-wise entropy of the annual min-max interval-valued sample with Algorithm 3 in [18] earlier, we have assumed a uniform distribution for each interval. In addition to uniform distribution, we consider both normal and beta distributions in this work because of their popularity in applications. In this study, we computationally investigate the impacts of a combination of an aggregation strategy associated with a probabilistic distribution on the entropy of resulting interval-valued data. We report our numerical results on each of the following four combinations:

(a) Min-max interval with uniform distribution;
(b) Fitting data with a normal distribution then forming confidence interval with $Z = 1.5$, using normal distribution in entropy calculation;
(c) Fitting data with a normal distribution then forming confidence interval with $Z = 1.5$, then assuming uniform distribution on each interval in entropy calculation; and
(d) Min-max interval fitting with a beta distribution.

Table 2 lists attribute-wise entropies for each of the four cases above. Figure 3 provides a visual comparison. Python modules `numpy` and `scipy` are used as the main software tools in carrying out the computational results.

We now analyze each of the outputs from (a)–(d) in Fig. 3.

Table 2. Entropy comparison of data aggregation methods and *pdf* selection

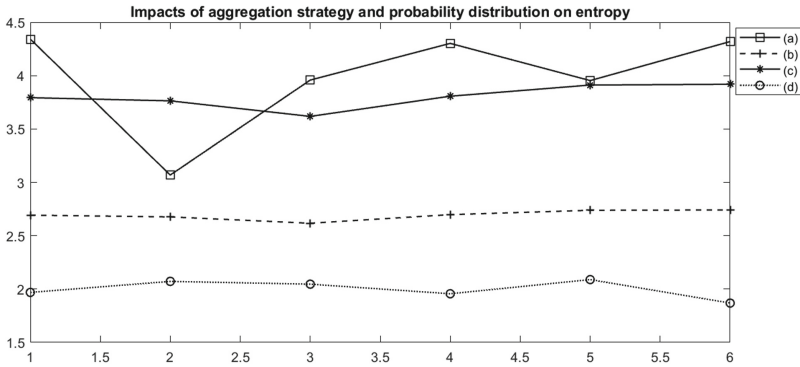|	UI	DI	SP	IP	DF	TM
(a) Min-max, unif.	4.34192	3.06851	3.95838	4.30213	3.95359	4.31941
(b) Conf. intvl, normal	2.69246	2.67623	2.61681	2.69736	2.73824	2.74129
(c) Conf. intvl, unif.	3.79327	3.76349	3.61804	3.80710	3.91177	3.91903
(d) Min-max, beta	1.96865	2.07197	2.04587	1.95605	2.08885	1.86871

Fig. 3. Entropy comparison of data aggregation methods with *pdf* selection

The line (a) is exactly the same as the min-max interval line in Fig. 2. This is because of that we have already assumed uniform distribution in calculating the attribute-wise entropy for each of the min-max intervals.

The line (b) indicates that the entropies of the interval-valued sample formed with the method (b) are much less than that of the interval-valued one, i.e., the line (a). This is not by an accident. Equation (4) uses the arithmetic average of $\sum_i pdf_i$ as the *pdf* of an interval-valued dataset **X**. As we know, the sum of normal random variables follows a normal distribution. Therefore, the resulting interval-valued dataset obtained with (b) follows a normal distribution, which is determined only by its mean and standard deviation with much less irregularity. Therefore, the calculated entropy is much less than that of (a). However, one should not abolish confidence interval aggregation at all. The only thing causing the relatively less entropy is the entropy calculation, in which, we assumed normal distribution for each pdf_i. This is further explained on the line (c) below.

The line (c) shows the results obtained with the same confidence intervals in (b) but then assuming a uniform distribution for each interval in calculating the entropy. The Corollary 2 in [18] makes this practically doable. Notice that the lines (c) and (a) are fairly close to each other comparing against (b) and (d). This means that using a confidence interval to aggregate points can still be a valid practical approach. Computational results in [17] repeated below further verify the claim as an evidence. By adjusting the *Z*-values of normal distribution, several interval-valued annual samples are formed at different levels of probabilistic

confidence. Using them, that work reports some changes in overall quality of the stock market annual forecasts. The average accuracy ratio associated with the Z-values are: 61.75% with $Z = 1.25$, 64.23% with $Z = 1.50$, 64.55% with $Z = 1.75$, and 62.94% with $Z = 2.00$. They are very close to 64.19% reported in [14] with the min-max intervals. The relatively overall closeness of line (c) and (a) can be an explanation for the similarity of the average accuracy ratios. The closeness of (a) and (c) also implies that adjusting the Z-value in data aggregation may slightly improve the quality of forecasts but not significantly. Lastly, the ILS algorithm [15] does not depend on any specific probability distribution but the calculation of entropy does. Therefore, in calculating entropy of samples formed with confidence intervals, assuming a uniform distribution can be a good choice like in the reported case study of stock market forecasting. Unless, each attribute follows a normal distribution indeed.

The line (d) is much lower than the rests. However, we ignore it because of the reasons explained below. In our implementation, we call the `beta.fit` in `scipy.stats` module to estimate the parameters of a beta distribution, which fits the data best. During run time, we have encountered multiple run-time warnings although our implementation returns the reported attribute-wise entropy. After checking our code carefully without finding any bugs, we examine the latest available official documentation of scipy updated on December 19, 2019. Regarding beta fit, it states "The returned answer is not guaranteed to be the globally optimal MLE (Maximum Likelihood Estimate), it may only be locally optimal, or the optimization may fail altogether" [27]. We do not have any other explanations for the numerical results. Due to the run-time warnings and current software documentation, we accept that the specific computational results on (d) are not reliable as a fact.

5   Conclusions and Possible Future Work

Applying interval-valued data rather than point-valued ones, researchers have made very significant quality improvements of variability forecasts. This work strongly suggests that the significant quality improvements in previous studies very much likely come from the interval-valued inputs. Figure 2 clearly shows that the attribute-wise entropies of an interval-valued sample are much higher than that of those point-valued samples. The more information contained in the input data, the higher quality outputs could be expected. Furthermore, the interval least-squares algorithm [15] can be applied to successfully extract information from an interval-valued sample rather than using the traditional ordinary least-squares approaches as reported in [11] and others.

Computational results also conclude that both min-max and confidence intervals can be effectively used to aggregate point-valued data into intervals. Both of them may lead to similarly well quality variability forecasts with the evidence on the stock market reported in [3] and [17]. This is because of that they may result in interval-valued samples with similar entropies as illustrated in Fig. 3 lines (a) and (c). While the interval least-squares algorithm itself does not demand probability distribution information at all, calculating the entropy of an interval-valued

dataset does. The lines (b) and (c) in Fig. 3 suggest that a uniform probability distribution on each interval can be a good choice in calculating the entropy of an interval-valued dataset.

In summary, this work provides information theoretic evidences, in addition to empirical results published previously, on the followings:

– Using interval-valued samples together with ILS is preferable than using point-valued ones with OLS in variability forecasts like predicting annual variability of the stock market and others.
– Applying min-max interval and/or confidence interval (at an appropriate level of confidence) to aggregate points into intervals may result in interval-valued samples containing similar amount of information.
– When estimating the entropy of an interval-valued dataset with (5), it can be a good choice of assuming a uniform distribution on each interval. Unless, it follows a normal distribution indeed.

The future work may consist of both sides of application and theory. With the information theoretic evidence, we have validated previously published results with interval-valued data and ILS. Therefore, applying interval methods in variability forecasts with uncertainty has a high priority. On the theoretic side, we should indicate that attribute-wise entropy is not exactly the same as the entropy of a multidimensional dataset. Investigating attribute-wise entropy in this study is not only because of its simplicity, but also because [18] only provides point-valued probability and entropy for single dimensional interval-valued datasets. Therefore, establishing point-valued probability and entropy for a multidimensional interval-valued dataset is among future works too.

References

1. Bouchon-Meunier, B.: Aggregation and Fusion of Imperfect Information. Springer, Heidelberg (2013)
2. Chatfield, C.: Prediction intervals for time-series forecasting. In: Armstrong, J.S. (ed.) Principles of Forecasting. International Series in Operations Research & Management Science, vol. 30, pp. 475–494. Springer, Boston (2001). https://doi.org/10.1007/978-0-306-47630-3_21
3. Chen, G., Hu, C.: A computational study on window-size selection in stock market RILS interval forecasting. In: 2009 World Congress on Computer Science and Information Engineering, Los Angeles, CA, pp. 297–301. IEEE (2009)
4. Chen, N.F., Roll, R., Ross, S.A.: Economic forces and the stock market. J. Bus. **59**(3), 383–403 (1986)
5. Gardner, E.: A simple method of computing prediction intervals for time series forecasts. Manage. Sci. **34**, 541–546 (1988)

6. Granger, C.: Can we improve the perceived quality of economic forecasts? J. Appl. Econom. **11**, 455–473 (1996)

7. Gooijer, J., Hyndman, R.: 25 years of time series forecasting. J. Forecast. **22**, 443–473 (2006)

8. Grabisch, M., Marichal, J., Mesiar, R., Pap, E.: Aggregation Functions. Cambridge University Press, New York (2009)

9. Gray, R.M.: Entropy and Information Theory, 1st edn., Corrected, Springer, New York (2013)

10. He, L.T., Hu, C.: Midpoint method and accuracy of variability forecasting. Empir. Econ. **38**, 705–715 (2009). https://doi.org/10.1007/s00181-009-0286-6

11. He, L.T., Hu, C.: Impacts of interval computing on stock market variability forecasting. Comput. Econ. **33**(3), 263–276 (2009). https://doi.org/10.1007/s10614-008-9159-x

12. He, L., Hu, C., Casey, M.: Prediction of variability in mortgage rates: interval computing solutions. J. Risk Finance **10**(2), 142–154 (2009)

13. Kreinovich, V., Korvin, A., Baker Kearfott, R., Hu, C. (eds.): Knowledge Processing with Interval and Soft Computing. AIKP. Springer, London (2008). https://doi.org/10.1007/978-1-84800-326-2

14. Hu, C., He, L.T.: An application of interval methods to stock market forecasting. Reliable Comput. **13**, 423–434 (2007). https://doi.org/10.1007/s11155-007-9039-4

15. Hu, C., He, L.T., Xu, S.: Interval function approximation and applications. In: Kreinovich, V., Korvin, A., Baker Kearfott, R., Hu, C. (eds.) Knowledge Processing with Interval and Soft Computing. AIKP, pp. 119–134. Springer, London (2008). https://doi.org/10.1007/978-1-84800-326-2_5

16. Hu, C.: Using interval function approximation to estimate uncertainty. In: Huynh, V.N., Nakamori, Y., Ono, H., Lawry, J., Kreinovich, V., Nguyen, H.T. (eds.) Interval / Probabilistic Uncertainty and Non-Classical Logics. AINSC, vol. 46, pp. 341–352. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-77664-2_26

17. Hu, C.: A note on probabilistic confidence of the stock market ILS interval forecasts. J. Risk Finance **11**(4), 410–415 (2010)

18. Hu, C., Hu, Z.: On statistics, probability, and entropy of interval-valued datasets. In: Lesot, M.-J., et al. (eds.) IPMU 2020. CCIS, vol. 1239, pp. 407–421. Springer, Cham (2020)

19. IEEE Standard for Interval Arithmetic. IEEE Standards Association (2015). https://standards.ieee.org/standard/1788-2015.html

20. IEEE Standard for Interval Arithmetic (Simplified). IEEE Standards Association (2018). https://standards.ieee.org/standard/1788_1-2017.html

21. de Korvin, A., Hu, C., Chen, P.: Generating and applying rules for interval valued fuzzy observations. In: Yang, Z.R., Yin, H., Everson, R.M. (eds.) IDEAL 2004. LNCS, vol. 3177, pp. 279–284. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-28651-6_41

22. Marupally, P., Paruchuri, V., Hu, C.: Bandwidth variability prediction with rolling interval least squares (RILS). In: Proceedings of the 50th ACM SE Conference, Tuscaloosa, AL, USA, 29–31 March 2012, pp. 209–213. ACM (2012). https://doi.org/10.1145/2184512.2184562

23. Nordin, B., Chen, B., Sheng, V.S., Hu, C.: Interval-valued centroids in K-Means algorithms. In: Proceedings of the 11th IEEE International Conference on Machine Learning and Applications (ICMLA), Boca Raton, FL, USA, pp. 478–481. IEEE (2012). https://doi.org/10.1109/ICMLA.2012.87

24. Pękala, B.: Uncertainty Data in Interval-Valued Fuzzy Set Theory: Properties, Algorithms and Applications. SFSC, vol. 367, 1st edn. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-93910-0
25. Rhodes, C., Lemon, J., Hu, C.: An interval-radial algorithm for hierarchical clustering analysis. In: Proceedings of the 14th IEEE International Conference on Machine Learning and Applications (ICMLA), Miami, FL, USA, pp. 849–856. IEEE (2015). https://doi.org/10.1109/ICMLA.2015.118
26. Shannon, C.-E.: A mathematical theory of communication. Bell Syst. Tech. J. **27**, 379–423 (1948)
27. Scipy.stats Documentation. https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.rv_continuous.fit.html. Updated 19 Dec 2019
28. Wikipedia, Information entropy. https://en.wikipedia.org/wiki/entropy_(information_theory). Edited 23 Dec 2019
29. Xu, S., Chen, X., Han, A.: Interval forecasting of crude oil price. In: Huynh, V.N., Nakamori, Y., Ono, H., Lawry, J., Kreinovich, V., Nguyen, H.T. (eds.) Interval / Probabilistic Uncertainty and Non-Classical Logics. AINSC, vol. 46, pp. 353–363. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-77664-2_27