Abstract—The well known fuzzy measures, λ-measure and P-measure, have only one formulaic solution. A multivalent fuzzy measure with infinitely many solutions based on P-measure was proposed by our previous work, called completed L-measure. In this paper, a further improved fuzzy measure, called extensional completed L-measure, is proposed. This new fuzzy measure is proved that it is not only an extension of completed L-measure but also can be considered as an extension of the λ-measure and P-measure. For evaluating the Choquet integral regression models with our proposed fuzzy measure and other different ones, a real data experiment by using a 5-fold cross-validation mean square error (MSE) is conducted. The performances of Choquet integral regression models with fuzzy measure based on extensional completed L-measure, completed L-measure, L-measure, λ-measure, and P-measure, respectively, a ridge regression model, and a multiple linear regression model are compared. Experimental result shows that the Choquet integral regression models with respect to extensional completed L-measure based on γ-support outperforms others forecasting models.

Index Terms—λ-measure, P-measure, L-measure, completed L-measure, extensional completed L-measure

I. INTRODUCTION

When interactions among independent variables exist in forecasting problems, the performance of the multiple linear regression models are not good enough. The traditional improved methods exploited the ridge regression models [1]. Recently, the Choquet integral regression models [6-8,10-12] based on some univalent or multivalent fuzzy measures were suggested [2-8] to improve this situation. The well known fuzzy measures, λ-measure [2,3] and P-measure [4] are both univalent fuzzy measures, both of them have only one formulaic solution of fuzzy measure, the former is not a closed form, and the later is not sensitive. A multivalent fuzzy measure with infinitely many solutions of closed form based on P-measure was proposed by our previous work, called L-measure [7], but L-measure is not a completed measure, and then, the completed L-measure [10] was proposed by our next previous work. In this paper, an improved fuzzy measure, called extensional completed L-measure, is proposed. This new fuzzy measure is proved that it is not only an extension of completed L-measure but also can be considered as an extension of the λ-measure and P-measure. Thereby, we can obtain an improved Choquet integral regression model with respect to this new fuzzy measure.

This paper is organized as followings: The multiple linear regression and ridge regression are introduced in section II; two well known fuzzy measure, λ-measure, P-measure, L-measure and completed L-measure are introduced in section III; Our two multivalent fuzzy measures L-measure and completed L-measure are introduced in section IV, our new measure, Extensional completed L-measure, is introduced in section V; the fuzzy support, γ-support is described in section VI; the Choquet integral regression model based on fuzzy measures are described in section VII; experiment and result are described in section VIII; and final section is for conclusions and the future works.

II. THE MULTIPLE LINEAR REGRESSION, RIDGE REGRESSION

Let \( \mathbf{Y} = X\beta + \epsilon, \epsilon \sim N(0, \sigma^2 I_n) \) be a multiple linear model, \( \hat{\beta} = (X'X)^{-1}X'Y \) be the estimated regression coefficient vector, and \( \hat{\beta}_k = (X'X + kl)^{-1}X'Y \) be the estimated ridge regression coefficient vector, Hoerl., Kenard and Baldwin [1] suggested

\[
\hat{k} = \frac{n\sigma^2}{\beta'\beta} \quad (1)
\]

III. FUZZY MEASURES

The well known fuzzy measures, the λ-measure proposed by Sugeno in 1974, and P-measure proposed by Zadah in 1978, and L-measure proposed by our previous work in 2007 are concise introduced as follows.

A. Definition of fuzzy measures [2,3]

A fuzzy measure \( \mu \) on a finite set \( X \) is a set function \( \mu : 2^X \rightarrow [0,1] \) satisfying the following axioms:

\[
\mu(\emptyset) = 0, \mu(X) = 1 \quad \text{(boundary conditions)} \quad (2)
\]

\[
A \subseteq B \Rightarrow \mu(A) \leq \mu(B) \quad \text{(monotonicity)} \quad (3)
\]
B. Fuzzy density function (fdf) [2, 3]

A fuzzy density function of a fuzzy measure \( \mu \) on a finite set \( X \) is a function \( s: X \rightarrow [0,1] \) satisfying:

\[
s(x) = \mu(\{x\}), \quad x \in X
\]

(4)

\( s(x) \) is called the density of singleton \( x \).

C. \( \lambda \)-measure [3]

For given fuzzy density function, \( s \), a \( \lambda \)-measure, \( g_\lambda \), is a fuzzy measure on a finite set \( X \), satisfying:

\[
A, B \in 2^X, \quad A \cap B = \emptyset, \quad A \cup B \neq X
\]

\[
\Rightarrow g_\lambda(A \cup B) = g_\lambda(A) + g_\lambda(B) + \lambda g_\lambda(A) g_\lambda(B)
\]

(5)

\[
\prod_{i=1}^{n}[1 + \lambda s(x_i)] = \lambda + 1 > 0, \quad s(x_i) = g_\lambda(\{x_i\})
\]

(6)

where, the real number, \( \lambda \), is also called the determined coefficient of \( \lambda \)-measure.

Note that once the fuzzy density function is known, we can obtain the values of \( \lambda \) uniquely by using the previous polynomial equation. In other words, \( \lambda \)-measure has a unique solution without closed form.

D. P-measure [4]

For each given fuzzy density function, \( s \), a P-measure, \( g_p \), is a fuzzy measure on a finite set \( X \), satisfying:

\[
\forall A \in 2^X \Rightarrow g_p(A) = \max_{x \in A} s(x) = \max_{x \in A} g_p(\{x\})
\]

(7)

IV. MULTIVALENT FUZZY MEASURES

A. Definition of multivalent fuzzy measures

Definition 1. A fuzzy measure is called a multivalent measure, if it has more than one fuzzy measure solution.

B. Definition of B- measures

Definition 2. B- measures [10,12]

For any given fuzzy density function, \( s(x) \), on a finite set, \( X \), the B-measure is a set function,

\[
g_B: 2^X \rightarrow [0,1],
\]

satisfying:

\[
g_B(A) = \begin{cases} 0 & A = \emptyset \\ 1 & |A| > 1, A \subseteq X \end{cases}
\]

(8)

C. Comparison between two fuzzy measures [10]

Definition 3. \( \mu_1 \)-measure \( \leq \) \( \mu_2 \)-measure, \( \mu_2 \)-measure \( \geq \) \( \mu_1 \)-measure

For any given fuzzy density function, \( s(x) \), on a finite set, \( X \), If \( \mu_1 \) and \( \mu_2 \) are two fuzzy measures, satisfying \( g_{\mu_1}(A) \leq g_{\mu_2}(A), \forall A \subseteq X \), then we say that \( \mu_1 \)-measure is not larger than \( \mu_2 \)-measure, or \( \mu_2 \)-measure is not smaller than \( \mu_1 \)-measure, denoted as \( \mu_1 \)-measure \( \leq \) \( \mu_2 \)-measure, or \( \mu_2 \)-measure \( \geq \) \( \mu_1 \)-measure.

Theorem 1. For any given fuzzy density function, \( s(x) \), on a finite set, \( X \),

1) P-measure is not larger than any other normalized fuzzy measure, \( \mu \), that is

\[
P\text{-measure} \leq \mu\text{-measure}
\]

2) B-measure is not smaller than any other normalized fuzzy measure, \( \mu \), that is

\[
\mu\text{-measure} \leq B\text{-measure}
\]

Proof. It is trivial.

D. completed-measure

Definition 4. Completed measure [10,12]

For any given fuzzy density function, \( s(x) \), on a finite set, \( X \), a multival ent fuzzy measure, \( \mu \)-measure, with determined coefficient, \( \mu \), is called a completed measure, if it satisfies following conditions

1) \( \mu \)-measure is a monotone increasing function of its determined coefficient, \( \mu \)
2) if \( \mu = 0 \) then \( \mu \)-measure is just the P-measure
3) If the upper limit normalized fuzzy measure of \( \mu \)-measure is just the B-measure.

E. L-measure

Definition 5. L-measure [7]

For each given fuzzy density function, \( s \), a L-measure, \( g_L \), is a fuzzy measure on a finite set \( X \), \( |X| = n \), satisfying:

\[
\forall A \subseteq X, \quad n - |A| + (|A| - 1)L > 0 \Rightarrow
\]

\[
g_L(A) = \max_{x \in A} s(x) + \frac{|A| - 1}{n - |A| + (|A| - 1)L} \sum_{x \in A} s(x)
\]

(9)

the real number, \( L \), is also called the determined coefficient of L-measure.

Theorem 2. For any given fuzzy density function, \( s(x) \), on a finite set, \( X \),

1) L-measure is an increasing function on \( L \).
2) \( L \in [0, \infty) \), L-measure has infinitely many fuzzy measure solutions
3) if \( L = 0 \) then L-measure is just the P-measure
4) if \( L \rightarrow \infty \) then L-measure is not the B-measure
5) L-measure is not a completed measure.

F. Completed L-measure

Definition 6. Completed L-measure [10]

For each given fuzzy density function, \( s \), a L-C-measure, \( g_{LC} \), is a fuzzy measure on a finite set \( X \), \( |X| = n \), satisfying:

\[
\forall A \subseteq X, \quad n - |A| + (|A| - 1)L > 0 \Rightarrow
\]

429
\[ g_{Ln}(A) = \max \left[ s(x) \right] \]
\[ + \left( |A| - 1 \right) L \sum_{x \in d} s(x) \left[ 1 - \max \left[ s(x) \right] \right] \]
\[ \left( |A| - 1 \right) L \sum_{x \in d} s(x) + \left[ n - |A| \right] \sum_{x \in X} s(x) \]
\[ (11) \]

where, the real number, \( L \), is also called the determined coefficient of \( L_{EC} \)-measure.

**Theorem 3.** For any given fuzzy density function, \( s(x) \), on a finite set, \( X \),
1) \( L_{EC} \)-measure is an increasing function on \( L \).
2) \( L \in [0, \infty) \), \( L_{EC} \)-measure has infinitely many fuzzy measure solutions
3) if \( L = 0 \) then \( L_{EC} \)-measure is just the \( P \)-measure
4) if \( L \to \infty \) then \( L_{EC} \)-measure is just the \( B \)-measure
5) \( L_{EC} \)-measure is a completed measure.

V. EXTENSIONAL COMPLETED L-MEASURE

**A. Definition of extensional completed fuzzy measure ,**

**Definition 7.** Extensional completed fuzzy measure

For any given fuzzy density function \( s(x) \), a extensional completed \( L \)-measure, \( g_{L_{EC}} \), is a measure on a finite set

\( X, \ [X] = n \), satisfying:-

\( L \in [-1, \infty), A \subset X \Rightarrow \)

\[ \left\{ \begin{array}{l}
\sum_{x \in A} s(x) + \left( |A| - 1 \right) L \sum_{x \in d} s(x) \left[ 1 - \max \left[ s(x) \right] \right] \\
\left( |A| - 1 \right) L \sum_{x \in d} s(x) + \left[ n - |A| \right] \sum_{x \in X} s(x)
\end{array} \right. \]
\[ (12) \]

\( g_{L_{EC}}(A) = \)

\[ \max \left[ s(x) \right] \]
\[ + \left( |A| - 1 \right) L \sum_{x \in d} s(x) \left[ 1 - \max \left[ s(x) \right] \right] \]
\[ \left( |A| - 1 \right) L \sum_{x \in d} s(x) + \left[ n - |A| \right] \sum_{x \in X} s(x) \]

where, the real number, \( L \), is also called the determined coefficient of \( L_{EC} \)-measure.

**B. Properties of extensional completed fuzzy measures ,**

**Theorem 4.** Extensional completed \( L \)-measure is a fuzzy measure

Proof; If \( L \in [-1, 0] \) it is trivial, Now, to prove that if \( L \in (0, \infty) \) then it is also a fuzzy measure;

1) (To prove the boundary conditions; \( 0 \leq g_{L_{EC}}(A) \leq 1 \))

\[ 0 \leq \sum_{x \in A} s(x) + \left( |A| - 1 \right) L \sum_{x \in d} s(x) \left[ 1 - \max \left[ s(x) \right] \right] \]
\[ \left( |A| - 1 \right) L \sum_{x \in d} s(x) + \left[ n - |A| \right] \sum_{x \in X} s(x) \leq 1 \]
\[ (13) \]

Therefore

\[ 0 \leq \sum_{x \in A} s(x) + \left( |A| - 1 \right) L \sum_{x \in d} s(x) \left[ 1 - \max \left[ s(x) \right] \right] \]
\[ \left( |A| - 1 \right) L \sum_{x \in d} s(x) + \left[ n - |A| \right] \sum_{x \in X} s(x) \leq 1 \]
\[ (14) \]

That is \( 0 \leq g_{L_{EC}}(A) \leq 1, \forall A \subset X, \forall L \in (0, \infty) \)

(II) (To prove the monotonicity;
\( L \in (0, \infty), A \subset B \subset X \Rightarrow g_{L}(A) \leq g_{L}(B) \))

Let \( \sum_{x \in d} s(x) = \sum_{x \in d} s(x) + d, \) where \( 0 \leq d \leq 1 \)

Then \( g_{L_{EC}}(B) - g_{L_{EC}}(A) = \)
\[ 1 - \left( |B| - 1 \right) L \sum_{x \in d} s(x) + \left[ n - |B| \right] \sum_{x \in X} s(x) \]
\[ + L \left[ 1 - \sum_{x \in A} \right] \left[ f(B) - f(A) \right] \]
\[ (15) \]

Where \( f(B) = \)
\[ 1 - \left( |B| - 1 \right) L \sum_{x \in d} s(x) + \left[ n - |B| \right] \sum_{x \in X} s(x) \]
\[ (16) \]

Since
\[ 1 - \left( |B| - 1 \right) L \sum_{x \in d} s(x) + \left[ n - |B| \right] \sum_{x \in X} s(x) \geq 0 \]
\[ (17) \]

And
\[ \left[ f(B) - f(A) \right] \geq 0 \]
\[ (18) \]

We can obtained \( g_{L}(B) - g_{L}(A) \geq 0 \), and the proof is completed.

**Theorem 5.** For any given fuzzy density function, \( s(x) \), on a finite set, \( X \),

1) \( L_{EC} \)-measure is an increasing function on \( L \).
2) \( L \in [-1, \infty) \), \( L_{EC} \)-measure has infinitely many fuzzy measure solutions
3) if \( L = -1 \) then \( L_{EC} \)-measure is just the \( P \)-measure
4) if \( -1 \leq L < 0 \) then \( L_{EC} \)-measure is a subadditive measure
5) if \( L = 0 \) then \( L_{EC} \)-measure is just the additive measure
6) if \( 0 < L \leq \infty \) then \( L_{EC} \)-measure is a supradditive measure
7) if \( L \to \infty \) then \( L_{EC} \)-measure is just the \( B \)-measure
8) \( L_{EC} \)-measure is a completed measure.
Proof; omitted.

VI. SUPPORT, \( \Gamma \) - SUPPORT

**A. Definition of Support**

**Definition 8.** standard fuzzy density function, Support \([2,3]\)

For a given fuzzy density function, \( s(x) \), of a fuzzy measure \( \mu \) on a finite set \( X \), if \( \sum_{x \in X} s(x) = 1 \), then \( s \) is called a standard fuzzy density function, or a support of \( \mu \).

**B. Property about Gamma-support**

**Theorem 6.** For any given standard fuzzy density function, \( s(x) \), on a finite set, \( X \), \( P \)-measure and \( \lambda \)-measure are special cases of \( L_{EC} \)-measure

Proof;

For any given standard fuzzy density function, \( s(x) \), on a finite set, \( X \), that is \( \sum_{x \in X} s(x) = 1 \), from the property of \( \lambda \)-measure, we know that the \( \lambda \)-measure is just the additive measure, and from Theorem 5, we obtain that \( \lambda \)-
measure and P-measure are special cases of L(EC) – measure.

C. Definition of Gamma-support

**Definition 9.** Gamma-support [7]

Let \( \mu \) be a fuzzy measure on a finite set, \( X = \{ x_1, x_2, \ldots, x_n \} \), \( y_i \) be global response of subject \( i \) and \( f_i(x_j) \) be the evaluation of subject \( i \) for singleton \( x_j \), satisfying:

\[
0 < f_i(x_j) < 1, \quad i = 1, 2, \ldots, N, \quad j = 1, 2, \ldots, n
\]

\[
\gamma(x_j) = \frac{1 + r(f(x_j))}{\sum_{i=1}^{n} [1 + r(f(x_i))]}, \quad j = 1, 2, \ldots, n
\]

satisfying \( 0 \leq \gamma(x_j) \leq 1 \) and \( \sum_{j=1}^{n} \gamma(x_j) = 1 \) (21)

Where \( r(f(x_j)) \) is the Pearson correlation coefficient of dependent variable \( y \) and independent variable \( f(x_j) \).

**Definition 10.** Choquet integral [2, 3, 9]

Let \( \mu \) be a fuzzy measure on a finite set \( X \). The Choquet integral of \( f(x_j) \) with respect to \( \mu \) for individual \( i \) is denoted by

\[
\int_{c} f(x) d\mu = \sum_{j=1}^{n} [f_i(x_{j(i)}) - f_i(x_{j(i)-1})] \mu(A_{j(i)}), i = 1, 2, \ldots, N
\]

where \( f_i(x_{j(i)}) = 0 \), \( f_i(x_{j(i)}) \) indicates that the indices have been permuted so that

\[
0 \leq f_i(x_{j(i)}) \leq f_i(x_{j(2)}) \leq \cdots \leq f_i(x_{j(n)})
\]

\[
A_{j(i)} = \{ x_{j(i)}, x_{j(2)}, \ldots, x_{j(n)} \}
\]

B. Choquet integral regression models

**Definition 11.** Choquet integral regression model [10-12]

Let \( y_1, y_2, \ldots, y_N \) be global or response evaluations of \( N \) sample points and \( f_i(x_j), f_j(x_j), \ldots, f_N(x_j) \), \( j = 1, 2, \ldots, n \), be their evaluations of independent variables \( x_j \), where \( f_i: X \rightarrow R_+ \), \( i = 1, 2, \ldots, N \).

Let \( \mu \) be a fuzzy measure with \( \gamma \)-support, \( \alpha, \beta \in R_+ \), \( y_i = \alpha + \beta \int_{c} f(x) d\mu + \varepsilon_i, \varepsilon_i \sim N(0, \sigma^2) \), \( i = 1, 2, \ldots, N \)

\[
(\hat{\alpha}, \hat{\beta}) = \arg \min_{\alpha, \beta} \left[ \sum_{i=1}^{N} \left( y_i - \alpha - \beta \int_{c} f(x) d\mu \right)^2 \right]
\]

then \( \hat{y}_i = \hat{\alpha} + \hat{\beta} \int_{c} f(x) d\mu, i = 1, 2, \ldots, N \) is called the Choquet integral regression equation of \( \mu \) with \( \gamma \)-support, where

\[
\hat{\beta} = S_{\gamma} / S_{\gamma} \]

\[
S_{\gamma} = \sum_{i=1}^{N} \left[ y_i - \frac{1}{N} \sum_{j=1}^{N} f_i dx_j \right] \frac{1}{N-1} \int_{c} f(x) d\mu
\]

\[
S_{\gamma} = \frac{1}{N} \sum_{i=1}^{N} f_i dx_j \frac{1}{N-1} \int_{c} f(x) d\mu
\]

VIII. EXPERIMENT AND RESULT

A real data set with 60 samples from a junior high school in Taiwan including the independent variables, examination scores of four courses, and the dependent variable, the score of the Basic Competence Test of nature science of junior high school listed in Table 2 of our previous work [8] is applied to evaluate the performances of four Choquet integral regression models with P-measure, \( \lambda \)-measure, L-measure completed L-measure and maximized L-measure based on \( \gamma \)-support respectively, a ridge regression model, and a multiple linear regression model by using 5-fold cross validation method to compute the mean square error (MSE) of the dependent variable. The formula of MSE is

\[
MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2
\]

The monotone density function, \( \gamma \)-support, of the P-measure, \( \lambda \)-measure, L-measure, L-measure completed L-measure and maximized L-measure are listed as follows, which can be obtained by using the formula (20)

\[
\{0.2229, 0.2848, 0.2567, 0.2356\}
\]

For any fuzzy measure, \( \mu \)-measures, once the \( \mu \)-measure is given, all event measures of \( \mu \) can be found, and then, the Choquet integral based on \( \mu \) and the Choquet integral regression equation based on \( \mu \) can also be found by using above corresponding formulae.

**TABLE I. MSE of PREDICTION MODELS**

<table>
<thead>
<tr>
<th>Prediction models</th>
<th>5-fold CV MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choquet Integral regression model</td>
<td>47.4753</td>
</tr>
<tr>
<td>L_E(measure)</td>
<td>47.9759</td>
</tr>
<tr>
<td>L-measure</td>
<td>48.4610</td>
</tr>
<tr>
<td>( \lambda )-measure</td>
<td>49.1832</td>
</tr>
<tr>
<td>P-measure</td>
<td>53.9582</td>
</tr>
<tr>
<td>Ridge regression model</td>
<td>59.1329</td>
</tr>
<tr>
<td>Multiple linear regression model</td>
<td>65.0664</td>
</tr>
</tbody>
</table>

The experimental results of six forecasting models are listed in Table I. We can find that the Choquet integral
regression model with maximized L-measure, $L_M$ -measure, based on $\gamma$-support outperforms other forecasting regression models.

IX. Conclusions

In this paper, an improved fuzzy measure, extensional completed L-measure, is proposed. This new fuzzy measure is proved that it is not only an extension of completed L-measure but also can be considered as an extension of the $\lambda$-measure and P-measure. For evaluating the Choquet integral regression models with our proposed fuzzy measure and other different ones, a real data experiment by using a 5-fold cross-validation mean square error (MSE) is conducted. The performances of Choquet integral regression models with fuzzy measure based on extensional completed L-measure, completed L-measure, L-measure, $\lambda$-measure, and P-measure, respectively, a ridge regression model, and a multiple linear regression model are compared. Experimental result shows that the Choquet integral regression models with respect to extensional completed L-measure based on $\gamma$-support outperforms others forecasting models.

In the future, we will apply the proposed Choquet integral regression model with this new fuzzy measure based on $\gamma$-support to develop multiple classifier system.

Acknowledgement

This paper is partially supported by the National Science Council grant (NSC 98-2410-H-468-014)

References


