

Article

Decision-Making Model of Performance Evaluation Matrix Based on Upper Confidence Limits

Teng-Chiao Lin ¹, Hsing-Hui Chen ², Kuen-Suan Chen ^{3,4,5,*}, Yen-Po Chen ⁶ and Shao-Hsun Chang ² 

¹ Department of Graduate Institute of Technological & Vocational Education, National Taipei University of Technology, Taipei 106344, Taiwan; chiao@mail.moe.gov.tw

² Department of Industrial Education and Technology, National Changhua University of Education, Changhua 500208, Taiwan; d0531012@gm.ncue.edu.tw (H.-H.C.); chess1@gm.ncue.edu.tw (S.-H.C.)

³ Department of Industrial Engineering and Management, National Chin-Yi University of Technology, Taichung 411030, Taiwan

⁴ Department of Business Administration, Chaoyang University of Technology, Taichung 413310, Taiwan

⁵ Department of Business Administration, Asia University, Taichung 413305, Taiwan

⁶ Department of Foreign Languages and Literatures, National Chung Hsing University, Taichung 402202, Taiwan; doraemonsan1997@gmail.com

* Correspondence: kschen@ncut.edu.tw

Abstract: A performance evaluation matrix (PEM) is an evaluation tool for assessing customer satisfaction and the importance of service items across various services. In addition, inferences based on point estimates of sample data can increase the risk of misjudgment due to sampling errors. Thus, this paper creates a decision-making model for a performance evaluation matrix based on upper confidence limits to provide various service operating systems for performance evaluation and decision making. The concept is that through the gap between customer satisfaction and the level of importance of each service item, we are able to identify critical-to-quality (CTQ) service items requiring improvement. Many studies have indicated that customer satisfaction and the importance of service items follow a beta distribution, and based on the two parameters of this distribution, the proposed indices for customer satisfaction and the importance of service items represent standardization. The vertical axis of a PEM represents the importance index; the horizontal axis represents the satisfaction index. Since these two indices have unknown parameters, this paper uses the upper confidence limit of the satisfaction index to find out the CTQ service items and the upper confidence limit of the importance index to determine the order of improvement priority for each service item. This paper then establishes a decision-making model for a PEM based on the above-mentioned decision-making rules. Since all decision-making rules proposed in this paper are established through upper confidence limits, the risk of misjudgment caused by sampling errors can be reduced. Finally, this article uses a practical example to illustrate how to use a PEM to find CTQ service items and determine the order of improvement priority for these service items that need to be improved.

Keywords: performance evaluation matrix; beta distribution; satisfaction index; importance index; upper confidence limit

MSC: 62C05



Citation: Lin, T.-C.; Chen, H.-H.; Chen, K.-S.; Chen, Y.-P.; Chang, S.-H. Decision-Making Model of Performance Evaluation Matrix Based on Upper Confidence Limits. *Mathematics* **2023**, *11*, 3499. <https://doi.org/10.3390/math11163499>

Academic Editors: Yuhlong Lio and Tzong-Ru Tsai

Received: 4 July 2023

Revised: 9 August 2023

Accepted: 11 August 2023

Published: 13 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Many studies have indicated that the performance evaluation matrix (PEM) is an evaluation tool for customer satisfaction and the importance of service items within various service operation systems [1–3]. Some studies have pointed out that evaluations using sample point estimation can increase the risk of misjudgment due to sampling errors [4,5]. Therefore, this paper will create a decision-making model for a performance evaluation

matrix based on upper confidence limits to provide various service operating systems for performance evaluation and decision making for improvement. The concept is that through the gap between customer satisfaction and the level of importance of each service item, we can identify critical-to-quality (CTQ) service items requiring improvement. Meanwhile, many studies have worked on the research of the PEM to evaluate various service systems to see whether the performance meets the requirements and make further decisions about a service item: whether improvements are needed, whether the status quo should be maintained, or whether resource transfer is recommended [6–10]. Similar to many other studies and to maintain generality, this paper assumes that there are q service items provided by a service system. Then, $2q$ questions are designed to conduct a survey of learners about satisfaction and importance [11]. As noted by Chen et al. [12], Lambert and Sharma [13] used the Likert scale to collect data on customer satisfaction and the importance of service items with q service items. The horizontal axis of a PEM represents customer satisfaction; the vertical axis represents the importance of service items. The PEM then is divided into three areas: improvement, maintenance, and resource transfer. Hung, Huang, and Chen [1] revised the placement of the three performance areas in the above-mentioned PEM, making the evaluation rule more reasonable. Since the value of the beta distribution falls between one and zero, based on the beta distribution, Hung et al. [1] asserted that customer satisfaction and the importance of service items should follow a beta distribution, and based on the two parameters of this distribution, they proposed indices of importance and satisfaction, which represent standardization. These two index values are both between zero and one. The index value of satisfaction ranges from zero, representing zero percent satisfaction, to one, representing one hundred percent satisfaction. Likewise, the index value of importance ranges from zero, representing zero percent importance, to one, representing one hundred percent importance.

According to Liu et al. [2], the random variable X_h represents the h th service item of satisfaction and then X_h is distributed as a beta distribution, denoted by $X_h \sim \text{Beta}(\alpha_h, \beta_h)$. Similarly, let Y_h indicate the h th service item of importance and then Y_h is distributed as a beta distribution, denoted by $Y_h \sim \text{Beta}(\delta_h, \gamma_h)$. Thus, these two indices can be expressed as follows:

$$\theta_h = \frac{\alpha_h}{\alpha_h + \beta_h} \text{ (satisfaction index)} \tag{1}$$

$$\tau_h = \frac{\delta_h}{\delta_h + \gamma_h} \text{ (importance index)} \tag{2}$$

Based on the above information, the values of these two indices are both between zero and one. Additionally:

1. When $\alpha_h > \beta_h$, then $\theta_h > 0.5$ and $p(X_h > 50\%) > 0.5$, which indicates that the distribution of highly satisfied customers, $p(X_h > 50\%)$, is of a high proportion.
2. When $\alpha_h = \beta_h$, then $\theta_h = 0.5$ and $p(X_h \geq 50\%) = p(X_h \leq 50\%) = 0.5$, which indicates that the distributions of highly satisfied customers, $p(X_h > 50\%)$, and customers with low satisfaction, $(X_h < 50\%)$, are of equal proportions.
3. When $\alpha_h < \beta_h$, then $\theta_h < 0.5$ and $p(X_h < 50\%) > 0.5$, which indicates that the distribution of customers with low satisfaction, $(X_h < 50\%)$, is of a high proportion.

We can see that when the proportion of highly satisfied customers is higher, the value of parameter a_l is higher than that of parameter b_l , and the value of the satisfaction index is higher as well. Likewise, when the proportion of customers with low satisfaction is higher, the value of parameter a_l is lower than that of parameter b_l , and the value of the satisfaction index is lower. And, when the value of parameter a_l is equal to that of parameter b_l , the value of the satisfaction index is exactly one-half. Additionally, the importance index has the same characteristic as the satisfaction index. We can see that these two indices can reasonably reflect the distributions of customer satisfaction and the importance of service items.

There are some papers that set the satisfaction index as the x-axis and the importance index as the y-axis to form the performance evaluation matrix; at the same time, based on the concepts of total quality management and a continuous improvement process, they use the mean value θ_0 as the standard of evaluation [14,15]. Then, based on the critical value derived from the statistical test, they form the evaluation area and the evaluation rule of the PEM. The critical value is affected by sample size n and by the value of the variance of the index estimator, causing the critical values of the q service items to be different. As a result, the minimum value is chosen to represent these q -critical values. Although this method has overcome the problem of inconsistency in evaluation criteria, another problem has arisen, i.e., inconsistency in the level of significance. To tackle this issue, this paper instead uses the upper confidence limit of the satisfaction index of the q service items to test whether the value of the satisfaction index of the service items is larger than the mean value θ_0 , and then to decide whether to improve the satisfaction level of the service items. Next, we use the upper confidence limit of the importance index of q service items to test whether the value of the importance index of the service items is larger than the mean value τ_0 , and to determine the order of improvement priority. This method is based on the statistical test and hence can lower the risk of misjudgment, which can be caused by sampling errors.

The remainder of this paper is organized as follows. In Section 2, this paper will derive the $100(1 - \alpha)\%$ upper confidence limit of the two indices based on sample data. In Section 3, this paper allows the horizontal axis to represent the mean value of the satisfaction index θ_0 , and the vertical axis to represent the mean value of the importance index τ_0 ; these two axes divide the PEM into four evaluation quadrants. Then, we use the upper confidence limits of the two indices to mark the evaluation coordinates of each service item (x_h, y_h) , and based on where the coordinates (x_h, y_h) are located in the PEM, we can evaluate whether a service item needs improvement. When resources are limited, we can set the order of improvement priority. In Section 4, we use a case study to illustrate an application of the model proposed in this paper, showing how to establish the evaluation coordinates and identify service items requiring improvement, and when resources are limited, how to set the order of improvement priority. Section 5 presents the conclusions.

2. The $100(1 - \alpha)\%$ Upper Confidence Limits

As stated, as the satisfaction index and importance index possess unknown parameters, they need to be estimated from the sample data of respondents. Let $(X_{h,1}, X_{h,2}, \dots, X_{h,n})$ be a random sample of customer satisfaction with a sample size of n , where $h = 1, 2, \dots, q$. Then, the unbiased estimator of the satisfaction index θ_h is written as follows:

$$\hat{\theta}_h = \frac{1}{n} \times \sum_{j=1}^n X_{hj}. \tag{3}$$

This is the expected value of $\hat{\theta}_h$, which is equal to θ_h , which is expressed as follows:

$$\begin{aligned} E[\hat{\theta}_h] &= \frac{1}{n} \sum_{j=1}^n E[X_{h,j}] \\ &= \int_0^1 x \times \text{Beta}(\alpha_h, \beta_h) dx \left(\frac{1}{n} \times \sum_{j=1}^n 1 \right) \\ &= \frac{\Gamma(\alpha_h + \beta_h) \Gamma(\alpha_h + 1)}{\Gamma(\alpha_h + \beta_h + 1) \Gamma(\alpha_h)} \int_0^1 \text{Beta}(\alpha_h + 1, \beta_h) dx \\ &= \theta_h. \end{aligned} \tag{4}$$

In addition, the sample standard deviation is written as follows:

$$S_{Xh} = \sqrt{\frac{1}{n - 1} \times \sum_{j=1}^n (X_{hj} - \bar{X}_h)^2} \tag{5}$$

Let the random variable Z_l be written as follows:

$$Z_h = \frac{\hat{\theta}_h - \theta_h}{S_{Xh} / \sqrt{n}} \tag{6}$$

Based on the Central Limit Theorem (CLT), the random variable Z_h is approximately distributed as a standard normal distribution for a large sample size n [16], that is,

$$Z_h \xrightarrow{n \rightarrow \infty} N(0, 1) \tag{7}$$

Based on the above information, this paper derives the $100(1 - \alpha)\%$ upper confidence limit of the satisfaction index θ_h , as follows:

$$\begin{aligned} 1 - \alpha &= p(Z_h \geq -z_\alpha) \\ &= p\left(\frac{\hat{\theta}_h - \theta_h}{S_{Xh} / \sqrt{n}} \geq -z_\alpha\right) \\ &= p\left(\theta_h \leq \hat{\theta}_h + z_\alpha \frac{S_{Xh}}{\sqrt{n}}\right) \end{aligned} \tag{8}$$

Therefore, the $100(1 - \alpha)\%$ upper confidence limit of the satisfaction index is

$$U\theta_h = \hat{\theta}_h + z_\alpha \frac{S_{Xh}}{\sqrt{n}} \tag{9}$$

where z_α is the upper α quantile of the standard normal distribution.

Similarly, let $(Y_{h,1}, Y_{h,2}, \dots, Y_{h,n})$ be a random sample of customer importance with a sample size of n , where $h = 1, 2, \dots, q$. Then, the unbiased estimator of the importance index τ_h is written as follows:

$$\hat{\tau}_h = \frac{1}{n} \times \sum_{j=1}^n Y_{h,j} \tag{10}$$

This is the expected value of $\hat{\tau}_h$, which is equal to τ_h ($E[\hat{\tau}_h] = \tau_h$). In addition, the sample standard deviation is written as follows:

$$S_{Yh} = \sqrt{\frac{1}{n-1} \times \sum_{j=1}^n (Y_{h,j} - \bar{Y}_h)^2} \tag{11}$$

Let the random variable T_h be expressed as follows:

$$T_h = \frac{\hat{\tau}_h - \tau_h}{S_{Yh} / \sqrt{n}} \tag{12}$$

Based on the Central Limit Theorem (CLT), the random variable T_h is approximately distributed as the standard normal distribution for a large sample size n [16], that is,

$$T_h \xrightarrow{n \rightarrow \infty} N(0, 1) \tag{13}$$

Based on the above information, this paper derives the $100(1 - \alpha)\%$ upper confidence limit of the importance index τ_h , as follows:

$$\begin{aligned} 1 - \alpha &= p(T_h \geq -z_\alpha) \\ &= p\left(\frac{\hat{\tau}_h - \tau_h}{S_{Yh} / \sqrt{n}} \geq -z_\alpha\right) \\ &= p\left(\tau_h \leq \hat{\tau}_h + z_\alpha \frac{S_{Yh}}{\sqrt{n}}\right) \end{aligned} \tag{14}$$

Therefore, the $100(1 - \alpha)\%$ upper confidence limit of the importance index τ_h is

$$U\tau_h = \hat{\tau}_h + z_\alpha \frac{S_{Yh}}{\sqrt{n}}, \tag{15}$$

where z_α is the upper α quantile of the standard normal distribution.

3. The Decision-Making Model

As noted in some studies, the PEM is widely used to evaluate the performance level of q service items across various service systems [17,18]. Similar to these papers, the satisfaction index θ_h is set as the x -axis and the importance index τ_h is set as the y -axis to form the PEM. Let

$$\theta_0 = \frac{1}{q} \sum_{h=1}^q \hat{\theta}_h \tag{16}$$

and

$$\tau_0 = \frac{1}{q} \sum_{l=1}^q \hat{\tau}_h \tag{17}$$

Then, the hypothesis of the statistical test for the satisfaction index h is written as follows:

$$\text{null hypothesis } H_0 : \theta_h \geq \theta_0 \tag{18}$$

$$\text{alternative hypothesis } H_1 : \theta_h < \theta_0 \tag{19}$$

Similarly, the hypothesis of the statistical test for importance index h is written as follows:

$$\text{null hypothesis } H_0 : \tau_h \geq \tau_0 \tag{20}$$

$$\text{alternative hypothesis } H_1 : \tau_h < \tau_0 \tag{21}$$

Then, this paper uses the horizontal line $y = \tau_0$ and the vertical line $x = \theta_0$ to divide the performance evaluation matrix into four performance quadrants, as follows:

$$\text{Performance Quadrant I : } Q_I = \{(x, y) | \theta_0 \leq x \leq 1, \tau_0 \leq y \leq 1\} \tag{22}$$

$$\text{Performance Quadrant II : } Q_{II} = \{(x, y) | 0 \leq x \leq \theta_0, \tau_0 \leq y \leq 1\} \tag{23}$$

$$\text{Performance Quadrant III : } Q_{III} = \{(x, y) | 0 \leq x \leq \theta_0, 0 \leq y \leq \tau_0\} \tag{24}$$

$$\text{Performance Quadrant IV : } Q_{IV} = \{(x, y) | \theta_0 \leq x \leq 1, 0 \leq y \leq \tau_0\} \tag{25}$$

According to Equations (22)–(25), the performance evaluation matrix and four performance quadrants can be depicted, as shown in Figure 1, the red dots representing items that need improvement, and the blue dots representing items that do not need improvement.

Next, this paper uses the upper confidence limits of these two indices to serve as the evaluation coordinates of each service item (x_h, y_h) , as follows:

$$(x_h, y_h) = (U\theta_h, U\tau_h) = \left(\hat{\theta}_h + z_\alpha \frac{S_{Xh}}{\sqrt{n}}, \hat{\tau}_h + z_\alpha \frac{S_{Yh}}{\sqrt{n}} \right) \tag{26}$$

For the satisfaction index θ_h , based on the statistical test rules, the rules for whether improvements should be made are as follows:

1. When the $100(1 - \alpha)\%$ upper confidence limit of the satisfaction index θ_h is greater than or equal to θ_0 , that is, $x_h = U\theta_h \geq \theta_0$, we can conclude that $\theta_h \geq \theta_0$, indicating that the value of the satisfaction level of the service item h is higher than the mean value, and thus no improvement is needed.

- When the $100(1 - \alpha)\%$ upper confidence limit of the satisfaction index θ_h is smaller than θ_0 , that is, $x_h = U\theta_h < \theta_0$, we can conclude that $\theta_h < \theta_0$, indicating that the value of the satisfaction level of the service item h is lower than the mean value, and thus improvements are needed.

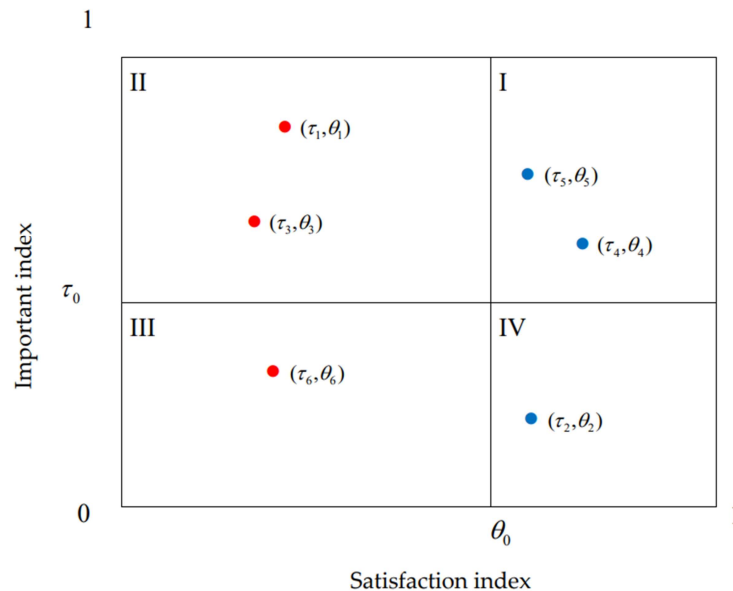


Figure 1. Performance evaluation matrix (sketch map).

For the importance index τ_h , based on the statistical test rules, the rules for determining the order of improvement priority are written as follows:

- When the $100(1 - \alpha)\%$ upper confidence limit of the importance index τ_h is greater than or equal to τ_0 , that is, $y_h = U\tau_h \geq \tau_0$, we can conclude that $\tau_h \geq \tau_0$, indicating that the value of the importance level of the service item h is higher than the mean value, and improvements of this service item should be prioritized when resources are limited.
- When the $100(1 - \alpha)\%$ upper confidence limit of the importance index τ_h is smaller than τ_0 , that is $y_h = U\tau_h < \tau_0$ we can conclude that $\tau_h < \tau_0$, indicating that the value of the importance level of the service item h is higher than the mean value, and improvements of this service item can be made later when resources are limited.

Based on the above concept, the evaluation rules determining which service items should be improved and the order of improvement priority are established as follows:

- If the evaluation coordinates of the h th service item $(x_h, y_h) \in Q_I$, we can conclude that $\theta_h \geq \theta_0$. Thus, service item h does not need improvements.
- If the evaluation coordinates of the h th service item $(x_h, y_h) \in Q_{II}$, we can conclude that $\theta_h < \theta_0$ and $\tau_h \geq \tau_0$. Thus, service item h needs to be improved and with high priority.
- If the evaluation coordinates of the h th service item $(x_h, y_h) \in Q_{III}$, we can conclude that $\theta_h < \theta_0$ and $\tau_h < \tau_0$. Thus, service item h needs to be improved and with lower priority.
- If the evaluation coordinates of the h th service item $(x_h, y_h) \in Q_{IV}$, we can conclude that $\theta_h \geq \theta_0$. Thus, service item h does not need improvements.

4. Case Study

As stated, the decision-making model of the performance evaluation matrix based on upper confidence limits, developed in this paper, can assist managers of various service systems in quickly identifying CTQ service items requiring improvements and in making improvements. For example, in the field of e-learning, many researchers have been working

on research about using teaching apps and computer-assisted language learning systems to increase learning effectiveness [19–22]. The proposed method can also be applied to the performance evaluation and model analysis of a real teaching setting. In technological and vocational education, students in practical courses learn through hands-on experience to combine theory and practice. So, this paper uses the case of practical courses as an example to illustrate the application of the proposed model. The questionnaires for the level of teaching satisfaction designed by Chen and Yureference Yeh et al. [23] and Yu et al. [24]. The questionnaires comprise an adequate number of questions and are quite comprehensive. Based on the version of Chen and Yu, this paper has designed a questionnaire to survey the level of satisfaction with the practical course, which includes seventeen questions and five dimensions: (1) Teaching preparation, (2) Teaching attitude, (3) Teaching ability, (4) Teaching management, and (5) Assessment. For each question, respondents are asked to provide both their level of satisfaction and the level of importance, and there are, in fact, thirty-four ($q = 17, 2q = 34$) questions to be answered. The questionnaire is as follows:

Dimension 1: Teaching preparation

1. The degree of difficulty of the practical material prepared by the teacher is moderate.
2. The learning amount of the practical material is moderate.
3. The practical material is helpful in enhancing students’ practical abilities.
4. The teacher is well-prepared for the practical course.

Dimension 2: Teaching attitude

5. The teacher pays attention to students’ hands-on experience.
6. The teacher considers students’ opinions important.
7. The teacher maintains a good relationship with each student.
8. The teacher is willing to assist students with problems.
9. The teacher treats every student fairly.

Dimension 3: Teaching ability

10. The teacher speaks clearly.
11. The teacher speaks coherently.

Dimension 4: Teaching management

12. The teacher implements a diversity of teaching methods in the practical course.
13. The practical course stimulates my interest in real teaching practice.
14. The teaching pace of the practical course is moderate.

Dimension 5: Assessment

15. The number of assignments and examinations is adequate.
16. The degree of difficulty of assignments and examinations is adequate.
17. The assignments or assessments can help enhance my practical abilities.

This paper used the above-shown questionnaire to collect data on the level of satisfaction and level of importance of students participating in practical courses. The respondents were students participating in a practical course at Taiwan’s University of Technology and Science. In total, 367 questionnaires were issued and 301 were recovered (with a recall rate of 82.02%). Next, we calculated the values of the sample statistics of each service item, as follows:

According to Equations (16) and (17), we obtain

$$\theta_0 = \frac{1}{17} \sum_{h=1}^{17} \hat{\theta}_h = 0.7864 \tag{27}$$

and

$$\tau_0 = \frac{1}{17} \sum_{l=1}^{17} \hat{\tau}_l = 0.8337 \tag{28}$$

Then, the four performance quadrants of the performance evaluation matrix are as follows:

$$\text{Performance Quadrant I : } Q_I = \{(x, y) | 0.7864 \leq x \leq 1, 0.8337 \leq y \leq 1\} \quad (29)$$

$$\text{Performance Quadrant II : } Q_{II} = \{(x, y) | 0 \leq x \leq 0.7864, 0.8337 \leq y \leq 1\} \quad (30)$$

$$\text{Performance Quadrant III : } Q_{III} = \{(x, y) | 0 \leq x \leq 0.7864, 0 \leq y \leq 0.8337\} \quad (31)$$

$$\text{Performance Quadrant IV : } Q_{IV} = \{(x, y) | 0.7864 \leq x \leq 1, 0 \leq y \leq 0.8337\} \quad (32)$$

Based on the data from Table 1, we plotted the 17 evaluation coordinates (x_h, y_h) as dots on the PEM below:

Table 1. Sample statistics of each service item.

Dimensions	Item	$\hat{\theta}_h$	S_{Xh}	$\hat{\tau}_h$	S_{Yh}	$x_h = U\theta_h$	$y_h = U\tau_h$	Quadrant
Dimension 1	1	0.7684	0.2255	0.7983	0.2043	0.7898	0.8176	I
	2	0.7617	0.2230	0.8023	0.1961	0.7829	0.8209	IV *
	3	0.7498	0.2352	0.8262	0.1856	0.7721	0.8438	II **
	4	0.7737	0.2310	0.8302	0.1984	0.7956	0.8490	I
Dimension 2	5	0.7963	0.2171	0.8275	0.1925	0.8169	0.8458	I
	6	0.8036	0.2199	0.8587	0.1630	0.8244	0.8742	I
	7	0.8116	0.2133	0.8375	0.1824	0.8318	0.8548	I
	8	0.8209	0.2107	0.8680	0.1650	0.8408	0.8837	I
	9	0.8474	0.2004	0.8700	0.1615	0.8664	0.8853	I
Dimension 3	10	0.7870	0.2313	0.8262	0.1995	0.8089	0.8451	I
	11	0.7956	0.2193	0.8561	0.1731	0.8164	0.8725	I
Dimension 4	12	0.7557	0.2295	0.8189	0.1956	0.7775	0.8374	II **
	13	0.7584	0.2459	0.8322	0.1942	0.7817	0.8506	II **
	14	0.7744	0.2335	0.8282	0.2037	0.7965	0.8475	I
Dimension 5	15	0.7896	0.2244	0.8401	0.1841	0.8109	0.8576	I
	16	0.7943	0.2148	0.8069	0.2047	0.8147	0.8263	I
	17	0.7810	0.2215	0.8454	0.1873	0.8020	0.8632	I

Remark: items need improvement with a high priority **; items need improvement with a low priority *.

Results and Discussion

In Figure 2, the evaluation coordinates of these 17 service items are plotted in the four evaluation quadrants, the red dots representing items that need improvement, and the blue dots representing items that do not need improvement.

Quadrant 1: items 4, 5, 6, 7, 8, 9, 10, 11, 14, 15, and 17

According to the evaluation rule determining which service items (larger than $\theta_0 = 0.7864$) should not be improved, service items 4, 5, 6, 7, 8, 9, 10, 11, 14, 15, and 17 do not need improvements.

Quadrant 2: items 3, 12, and 13

According to the evaluation rules determining which service items should be improved (smaller than $\theta_0 = 0.7864$) and the order of improvement priority (larger than $\tau_0 = 0.8337$), service items 3, 12, and 13 need improvements with a high priority.

Quadrant 3: item 2

According to the evaluation rules determining which service items should be improved (smaller than $\theta_0 = 0.7864$) and the second order of improvement priority (smaller than $\tau_0 = 0.8337$), service item 2 needs improvement with a low priority.

Quadrant 4: items 1 and 16

According to the evaluation rule determining which service items should not be improved (larger than $\theta_0 = 0.7864$) and the second order of improvement priority (smaller than $\tau_0 = 0.8337$), service items 1 and 16 do not need improvements.

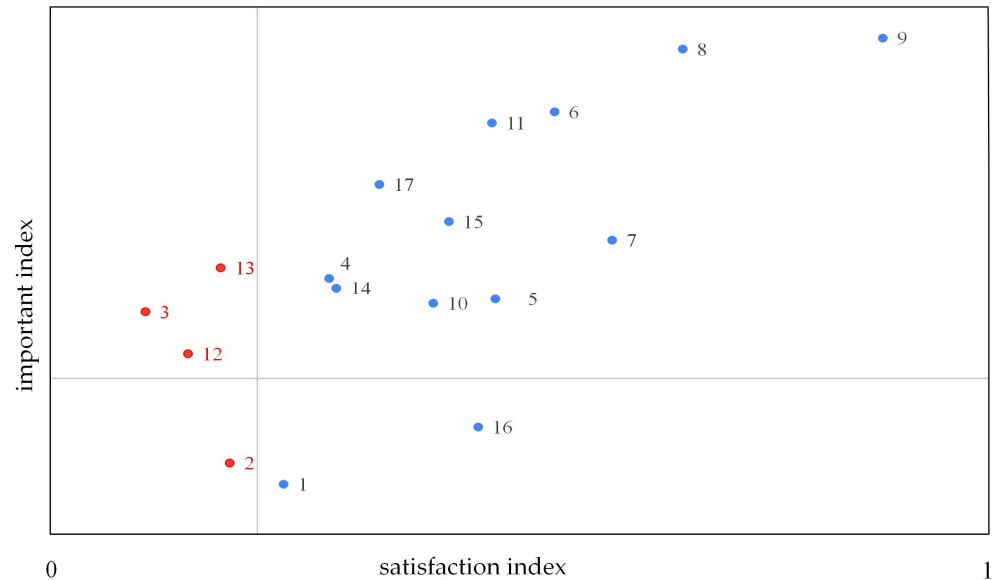


Figure 2. Performance evaluation matrix of the practical course.

Obviously, this model takes the average of the two evaluation indices as the testing standard and can improve the service direction with relatively weak performance, which is in line with the spirit of continuously improving total quality management. In addition, in the case of limited resources, priority should be given to improving service items with high importance to improve improvement efficiency. Based on this case, readers can be more aware of the application of the model proposed in this article, and it is convenient for enterprises to apply this model for evaluation and improvement.

5. Conclusions, Research Limitations, and Future Research

5.1. Conclusions

This paper has discussed the correlation between the values of the two indices and the proportion of customer satisfaction/importance of service items. We discovered that when the values of the indices are larger, the proportion of customer satisfaction/importance of service items becomes higher as well. We can see that these two indices can adequately reflect the actual distributions of customer satisfaction and the importance of service items. This paper set the mean value of the satisfaction index θ_0 and that of the importance index τ_0 as the horizontal line and the vertical of the PEM, which divided the matrix into four evaluation quadrants. This defined the set of these four evaluation quadrants, making this tool easy for managers to apply to their own systems. Since these two indices have unknown parameters, if the evaluation is performed directly by the point estimation, there will be a risk of misjudgment caused by sampling errors. As a result, based on the rule of using the upper confidence limits to conduct the statistical test, this paper used the upper confidence limits of the two indices as the evaluation coordinates of each service item (x_h, y_h) , and based on where the coordinates (x_h, y_h) are located in the PEM, we can evaluate whether a service item needs improvement, and when resources are limited, we can set the order of improvement priority. Then, we used a real case to explain how to apply the model proposed in this paper, construct evaluation coordinates, identify CTQ service items, and determine the order of priority for improvements when resources are limited. We can see that this method is based on statistical tests; as a result, the risk of misjudgment caused by sampling errors can be reduced. Obviously, this rule is in line with the spirit of continuously improving total quality management. At the same time, under

the condition of limited resources, prioritizing the improvement of high-importance service items can improve improvement efficiency. Based on this case, readers can be more aware of the application of the model proposed in this article, and it is convenient for enterprises to apply this model for evaluation and improvement.

5.2. Research Limitations and Future Research

Based on the recommendations of many studies, this paper uses the beta distribution as the distribution of customer importance and satisfaction. However, the statistical inference of the two indices must follow the central limit theorem, so the sample size must be large enough, which is the limitation of this study. In addition, the beta distribution is a continuous random variable. However, the collection of customer voices must be carried out through discontinuous scales. How to use the scale to make a reasonable discontinuity is the focus of future research.

Author Contributions: Conceptualization, T.-C.L. and K.-S.C.; methodology, T.-C.L. and K.-S.C.; software, H.-H.C. and Y.-P.C.; validation, S.-H.C.; formal analysis, T.-C.L., and K.-S.C.; resources, H.-H.C.; data curation, H.-H.C.; writing—original draft preparation, T.-C.L., H.-H.C., K.-S.C., Y.-P.C., and S.-H.C.; writing—review and editing, T.-C.L. and K.-S.C.; visualization, H.-H.C. and Y.-P.C.; supervision, K.-S.C.; project administration, T.-C.L. and H.-H.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data underlying the results are available as part of the article and no additional source data are required.

Acknowledgments: Thanks to Kuen-Suan Chen and Chun-Min Yu for providing the questionnaire for this paper and assisting in the collection of the questionnaires.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hung, Y.H.; Huang, M.L.; Chen, K.S. Service Quality Evaluation by Service Quality Performance Matrix. *Total Qual. Manag. Bus. Excell.* **2003**, *14*, 79–89. [[CrossRef](#)]
2. Liu, C.C.; Yu, C.H.; Chen, K.C. Using statistical test method to establish a decision model of performance evaluation matrix. *Appl. Sci.* **2023**, *13*, 5139. [[CrossRef](#)]
3. Yu, C.H.; Liu, C.C.; Chen, K.S.; Yu, C.M. Constructing fuzzy hypothesis methods to determine critical-to-quality service items. *Mathematics* **2020**, *8*, 573. [[CrossRef](#)]
4. Cheng, S.W. Practical implementation of the process capability indices. *Qual. Eng.* **1994**, *7*, 239–259. [[CrossRef](#)]
5. Chen, K.S.; Yu, C.M. Fuzzy test model for performance evaluation matrix of service operating systems. *Comput. Ind. Eng.* **2020**, *140*, 106240. [[CrossRef](#)]
6. Hossain, M.J.; Ahmed, S.M.Z. Developing a service performance assessment system to improve service quality of academic libraries. *Bus. Inf. Rev.* **2013**, *30*, 210–221. [[CrossRef](#)]
7. Li, Y.; Wang, L.; Li, F. A data-driven prediction approach for sports team performance and its application to national basketball association. *Omega* **2021**, *98*, 102123. [[CrossRef](#)]
8. Nam, J.H.; Lee, T.J. Foreign travelers' satisfaction with traditional Korean restaurants. *Int. J. Hosp. Manag.* **2011**, *30*, 982–989. [[CrossRef](#)]
9. Yu, C.M.; Zhuo, Y.J.; Lee, T.S. A Novel Social Media App Performance Evaluation Model. *J. Econ. Manag.* **2020**, *16*, 69–81.
10. Özkaya, M.; İzgi, B.; Perc, M. Axioms of Decision Criteria for 3D Matrix Games and Their Applications. *Mathematics* **2022**, *10*, 4524. [[CrossRef](#)]
11. Yu, C.M.; Chang, T.C.; Hu, I.H. Developing a discriminant index to determine critical service attributes of continuous performance improvement. *J. Serv. Sci. Res.* **2018**, *10*, 145–165. [[CrossRef](#)]
12. Chen, K.S.; Lii, P.C.; Li, S.Y.; Yu, C.M. Development and application of a performance evaluation matrix: A case study on exploring the items considered critical to quality. *J. Test. Eval.* **2020**, *48*, 3468–3478. [[CrossRef](#)]
13. Lambert, D.M.; Sharma, A. A customer-based competitive analysis for logistics decisions. *Int. J. Phys. Distrib. Logist. Manag.* **1990**, *20*, 17–24. [[CrossRef](#)]
14. Yang, C.C. Establishment and applications of the integrated model of service quality measurement. *Manag. Serv. Qual.* **2003**, *3*, 310–324. [[CrossRef](#)]
15. Yu, C.M.; Chang, H.T.; Chen, K.S. Developing a performance evaluation matrix to enhance the learner satisfaction of an e-learning system. *Total Qual. Manag. Bus. Excell.* **2018**, *29*, 727–745. [[CrossRef](#)]

16. Hogg, R.V.; Tanis, E.A. *Probability and Statistical Inference*, 6th ed.; Prentice-Hall: Hoboken, NJ, USA, 2001.
17. Wong, R.C.P.; Szeto, W.Y. An alternative methodology for evaluating the service quality of urban taxis. *Transp. Policy* **2018**, *69*, 132–140. [[CrossRef](#)]
18. Wu, J.; Wang, Y.; Zhang, R.; Cai, J. An approach to discovering product/service improvement priorities: Using dynamic importance-performance analysis. *Sustainability* **2018**, *10*, 3564. [[CrossRef](#)]
19. Chen, M.J.; Guo, C.Y.; Tang, C.W. Monitoring learning-oriented school education: Indicator construction and their application. *Stud. Educ. Eval.* **2022**, *73*, 101142. [[CrossRef](#)]
20. Amane, M.; Aissaoui, K.; Berrada, M. New perspective of learning objects in e-learning system. *Int. J. Inf. Learn. Technol.* **2023**, *40*, 269–279. [[CrossRef](#)]
21. Ma, W. Study on the evaluation method of students' english classroom performance based on big data analysis. *Int. J. Contin. Eng. Educ. Life-Long Learn.* **2022**, *32*, 19–34. [[CrossRef](#)]
22. Martínez-Caro, E.; Cegarra-Navarro, J.G.; Cepeda-Carrión, G. An application of the performance-evaluation model for e-learning quality in higher education. *Total Qual. Manag. Bus. Excell.* **2015**, *26*, 632–647. [[CrossRef](#)]
23. Yeh, L.C.; Tung, C.C.; Yang, S.Y.; Chen, J.H.; Shiau, F.H. The Development of the University Teacher Instructional Evaluation Scale. *Psychol. Test.* **2005**, *52*, 59–82.
24. Yu, C.M.; Chang, H.T.; Hsu, S.Y. An assessment of quality and quantity for foreign language training course to enhance students' learning effectiveness. *Int. J. Inf. Manag. Sci.* **2017**, *28*, 53–66.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.