Survival analysis for customer satisfaction: A case study

M A Hadiyat, R D Wahyudi and Y Sari

Department of Industrial Engineering, University of Surabaya, Raya Kalirungkut, Surabaya 60293, Indonesia

E-mail: arbi@staff.ubaya.ac.id

Abstract. Most customer satisfaction surveys are conducted periodically to track their dynamics. One of the goals of this survey was to evaluate the service design by recognizing the trend of satisfaction score. Many researchers recommended in redesigning the service when the satisfaction scores were decreasing, so that the service life cycle could be predicted qualitatively. However, these scores were usually set in Likert scale and had quantitative properties. Thus, they should also be analyzed in quantitative model so that the predicted service life cycle would be done by applying the survival analysis. This paper discussed a starting point for customer satisfaction survival analysis with a case study in healthcare service.

Keywords: satisfaction scores; survival analysis; satisfaction life cycle; service re-design predicting

1. Introduction

Many researchers proposed to conduct the customer satisfaction survey periodically to capture the behavior of this dynamic satisfaction scores [1]. The questionnaire for this survey consists of items which influence customers' satisfaction level. Qualitatively, some researchers decide to redesign the service when there are some indications that the scores are decreasing [2]. It means that there are some limitations in predicting when the redesign process should be conducted. One needs to expert judge the redesign time base on past experiences.

However, quantitative approach in designing questionnaire for measuring customer satisfaction should be considered. Items in questionnaire for this survey are usually scaled in Likert items, so that the scores have some quantitative property [3]. Once the survey is conducted, and satisfaction score is calculated, the quantitative analysis can be done. It gives stronger interpretation in predicting the service life cycle.

As in modeling the reliability of an electric product, the similar survival analysis will be fitted for the customer satisfaction scores. The idea for choosing the product reliability models in this paper is based on [4] and [5] who has successfully modeled the customer lifetime value using survival approach. The satisfaction scores are measured from time to time to capture the dynamic behavior of what customers have perceived. Theoretically, once the new service design is deployed to customers, the satisfaction will increase due to the new service innovations. At some levels, the satisfaction scores reach the maximum level, and then decreased as the perceived service is beginning to be unsatisfied. In this point, a new service design should be made and deployed soon to maintain the customer satisfaction level.

The survival analysis helps the management predict the time when new service design should be deployed by constructing the probability that the scores are less than certain levels of satisfaction. It began by control charting the satisfaction scores over time and capturing its dynamic, fitting the best probability distribution, followed by calculating the probability of satisfaction survival. Smaller survival probability of certain level of satisfaction represents the need of service redesign.

2. Literature review

2.1. Parasuraman's satisfaction scores

Parasuraman [6] mentioned that the satisfaction of customer was the gaps measurement between expected and perceived service received by customers, defined as customer gaps. Moreover, another gap was also measured for mining the causes of customer gaps, and all those integrated gaps were summarized as the well-known method called SERVQUAL which stands for service quality.
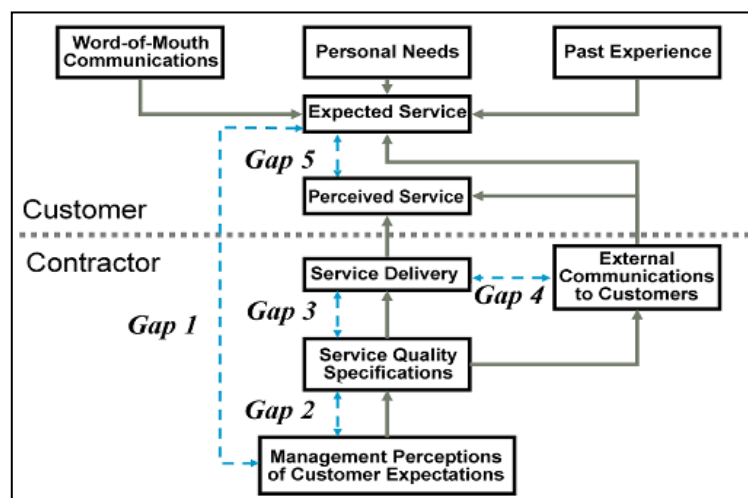

Figure 1. SERVQUAL gaps (adapted from [6]).

Common questionnaires for measuring these customer gaps consist of at least 20 Likert scale questions. All items are generated based on 5 dimensions i.e. tangibles, empathy, reliability, responsiveness, and assurance. The overall satisfaction score of a customer is calculated as the average of each item scores. Scores from all customers measured over time form a dataset to be modeled with survival analysis. In this paper, all satisfaction scores were converted to positive value-low positive mean unsatisfied while high one means satisfied.

2.2. Survival analysis

Similar to the calculation of product reliability, the survival analysis in this paper was based on [7] applied to satisfaction dataset. The steps in survival analysis for the product reliability were as follows:

a. Measure the reliability data t(age of component, or satisfaction scores),
b. Fit the dataset with some probability density function (PDF)$f(t)$, this function could be a product of two or more probability distribution in regards to its dynamically behavior.
c. Use the fitted PDF to calculate the reliability of component or satisfaction. Fitted probability distribution could be calculated using simulated Bayesian approach due to its multimodality.

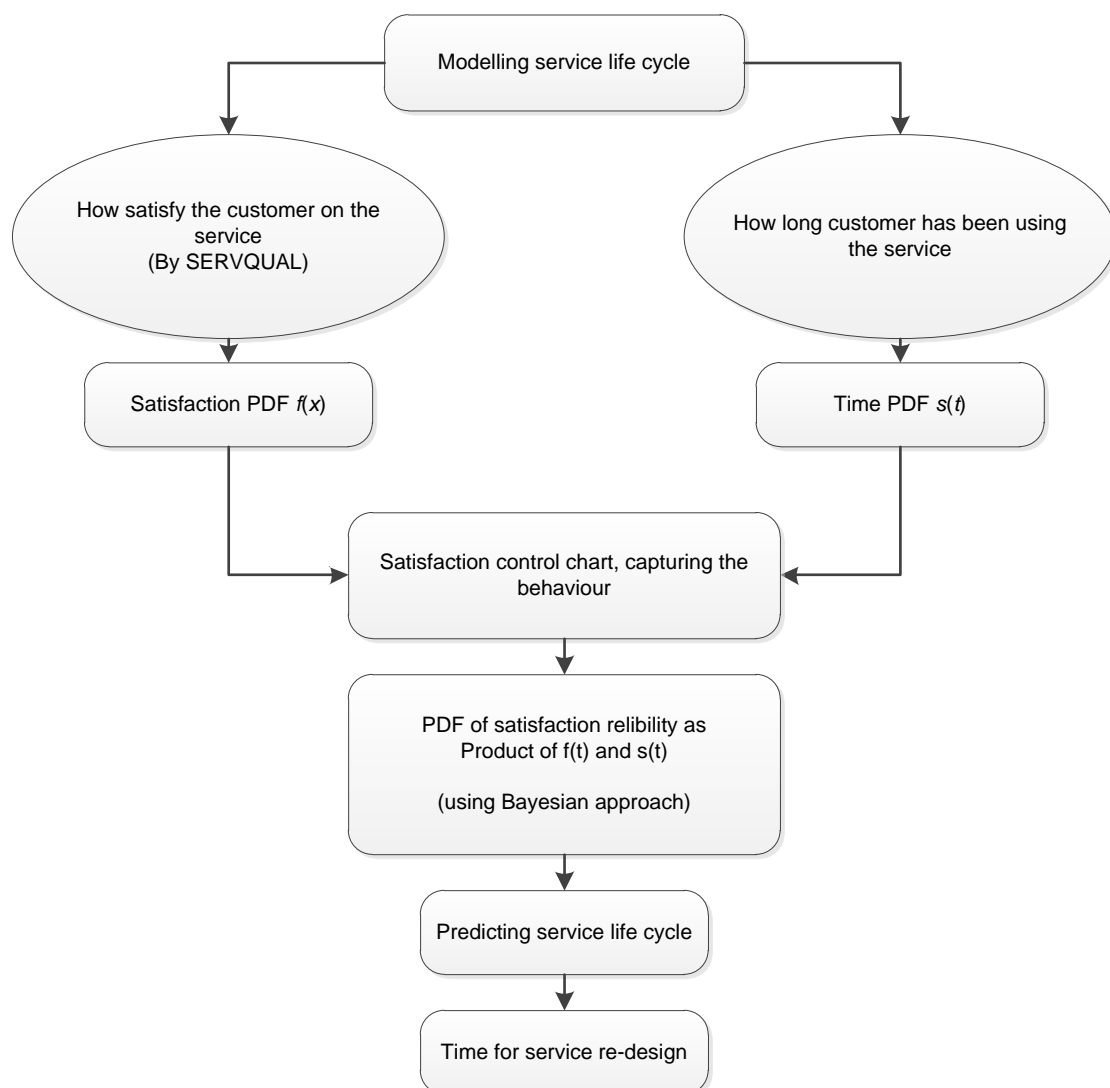$$\text{Reliability function } R(t) = P(t > T) = 1 - F(t), \qquad (1)$$

$$\text{Where } F(t) = P(t < T) = \int_{-\infty}^{T} f(t)dt$$

d. Predict the probability that a component or satisfaction will decreased at a certain level

The PDF used in this paper referred to the common survival analysis such as Weibull, exponential, gamma, and similar distribution covered in a normal family distribution, multiplied by similar distribution, so that $f(t)$ could be interpreted.

3. Research framework and methodology

This paper was a part of preliminary research in modeling customer satisfaction data and predicting the service life cycle. Commonly, the satisfaction level of a customer is related with how long he has been using the service, the longer using the service then the satisfaction level usually decreased, as mentioned in [8]. Meanwhile, since each customer has different satisfaction level variation over time, so there will be two random variables affecting the life-cycle of service, that are the satisfaction level (measured using questionnaire) and how long a customer has been using the service (measured in time). Each of these two variables has different statistical PDF, and both of them should be used to model the survival analysis for prediction the service life-cycle. The main framework included the integration of two probability distribution fitted by Bayesian approach. The satisfaction PDF was called $f(x)$, and how long customers have been using service was denoted by $s(t)$.
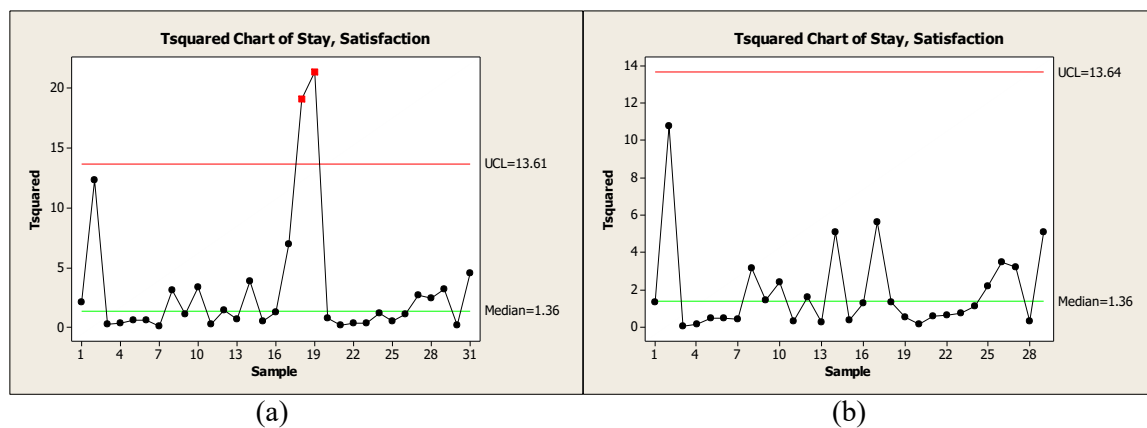
Figure 2. Research framework.

The product PDF as multiplication of $f(x)$ and $s(t)$ was done by simulating those prior distributions so that the posterior PDF was obtained. Mathematically, the posterior PDF (product PDF) for modeling satisfaction reliability survival function is written as follows:

$$L(xt) = \int f(x)s(t)\,dxdt, \qquad x \geq 0, t \geq 0 \qquad (2)$$

4. Results and discussion

The case study in this paper was taken from [9]. There were 155 data recorded from survey conducted while evaluating hospital patients' perception about its services. The dataset consisted of two variables, patient satisfaction scores (on scale of 1-7) and how long they have been treated and staying in the hospital (in days). First, the behavior of those two variables was captured using multivariate control chart, as in [10] and refers to [11].

(a) (b)

Figure 3. Multivariate control chart. (a) Before omitting unusual observations, and (b) After removing unusual observations.

Two points were out of control; indicating that some special cases should be investigated at a certain time. If there were some causes, the two points should be removed from analysis. Therefore, the PDF fitting process for those two variables could be held.

Based on parametric survival analysis using MINITAB software, the fitting process for $f(x)$ and $s(t)$ was done. PDF $f(x)$ was well fitted by Weibull distribution, and $s(t)$ was assumed to follow an exponential distribution.
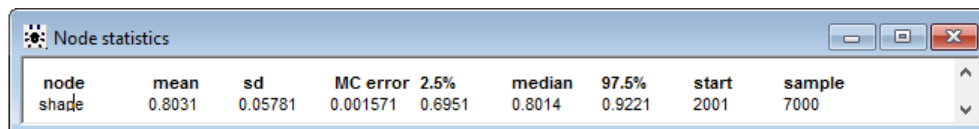
```
Goodness-of-Fit for Satisfaction

                         Anderson-Darling   Correlation
Distribution                        (adj)   Coefficient
Weibull                             2.212         0.981
Lognormal                           6.213         0.919
Exponential                        73.699             *
Loglogistic                         6.117         0.920
3-Parameter Lognormal               2.072         0.976
2-Parameter Exponential            57.498             *
3-Parameter Loglogistic             2.706         0.966
Smallest Extreme Value              1.734         0.985
Normal                              2.062         0.976
Logistic                            2.693         0.966
```

Figure 4. Minitab output; probability distribution fitting for satisfaction data.

The multiplication process in (2) could not be solved analytically, so Bayesian Markov Chain Monte Carlo [10] simulation was used to fit the PDF $L(xt)$ as the Bayesian posterior distribution. The Bayesian structure for likelihood function $f(t)$, prior distribution $s(t)$, and posterior $L(xt)$ were defined as:
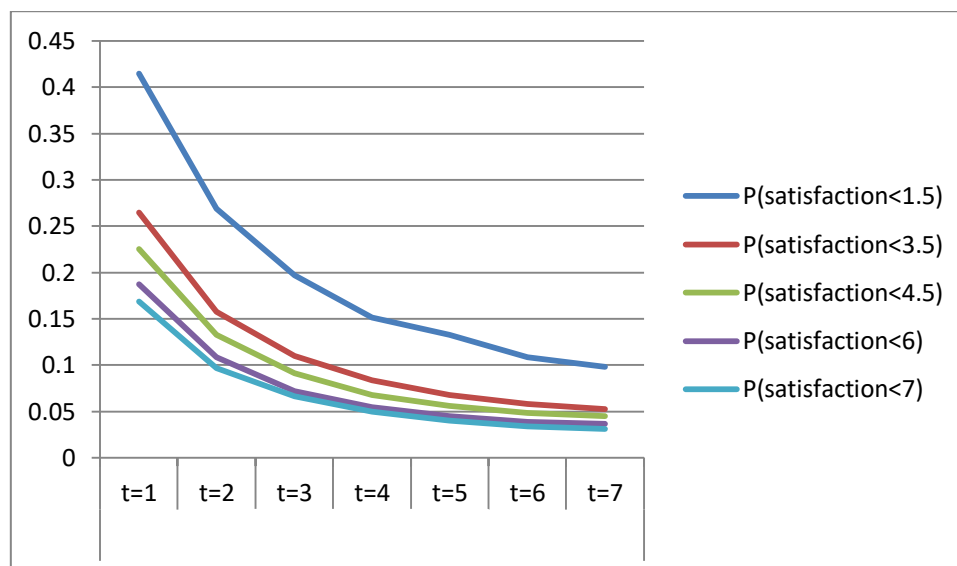
$$L(xt|\alpha, \mu) = \int f(x)s(t)\, dxdt$$
$$f(x|t, \alpha, \mu) \sim weibull(\alpha, s(t))$$
$$s(t) \sim exponential(\mu) \qquad\qquad (3)$$

The Markov Chain Monte Carlo simulation was conducted using WINBUGS (Bayesian Using Gibss Sampler) software [11].

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
shape	0.8031	0.05781	0.001571	0.6951	0.8014	0.9221	2001	7000

Figure 5. WINBUGS Markov Chain Monte Carlo Result for posterior PDF.

Figure 5 shows that the estimated α was 0.803, using a similar way that the estimation of other parameters was obtained. Next, this value would be used to predict the probability of satisfaction survival. Table 1 represents the trends of survival probability.

Figure 6. Survival probability over time (satisfaction Likert scale of 1-7).

The survival probability decreased over time; Moreover, the probability at same time t for high satisfaction level gave higher probability than lower ones. Figure 6 describes that until $t=7$, customer satisfaction becomes higher. On the first day in hospital, patients felt uncomforted because of inadaptable conditions, but after several days the patients perceived more satisfaction and got more complete service.

5. Concluding remarks

This was an initial research for mathematically modeling customer satisfaction due to the service quality of a nonmanufacturing company or public service. The level of satisfaction was dynamic and interesting to study. As shown in the analysis, practically life-cycle of service can be modeled and predicted, and this information is useful in deciding when a service design should be innovate by provider to re-increase the customer satisfaction level. The development of this research leads to the innovation of robust service design.

Acknowledgement

References

[1] Hill N, Roche G and Allen R 2007 *Customer Satisfaction: The Customer Experience through the Customer's Eyes* (London: Cogent Publishing) pp 245–250

[2] Lovelock C H, Patterson P and Wirtz J 2015 *Services Marketing* (Melbourne: Pearson Australia) pp 407–410

[3] Audrezet A, Olsen S O and Tudoran A A 2016 The GRID scale: A new tool for measuring service mixed satisfaction *J. Serv. Mark.* **30** pp 29–47

[4] Chen Y, Zhang Hand Zhu P 2009 Study of customer lifetime value model based on survival analysis *Proc. of WRI World Congress on Computer Science and Information Engineering* (Los Angeles: IEEE) pp 266–270

[5] Lu J 2003 Modeling customer lifetime value using survival analysis: An application in the telecommunications industry *Proc. of SAS User Group International Conference 28* (Washington: SAS Institute) pp 120–128

[6] Parasuraman A, Zeithaml V A and Berry L L 1988 SERVQUAL: a multi-item scale for measuring consumer perceptions of the service quality *J. Retailing* **64** pp 12–40.

[7] Ebeling C 2010 *An Introduction to Reliability and Maintainability Engineering* (London: McGraw Hill) pp 20–30

[8] Mittal V, Katrichis J M and Kumar P 2001 Attribute performance and customer satisfaction over time: evidence from two field studies *J. Serv. Mark.* **15** pp 343–356

[9] Trusko B, Pexton C, Harrington J and Gupta P K 2007 *Improving Healthcare Quality and Cost with Six Sigma* (London: FT Press) pp 241–370

[10] Khorsidi H A, Nikfalazar S and Gunawan I 2016 Statistical process control applications on service quality using SERVQUAL and QFD with a case study in trains service *The TQM Journal* **28** pp 195–201

[11] Montgomery D C 2005 *Introduction to Statistical Quality Control* (New York: John Wiley & Sons) pp 487–504

[12] Lunn D, Jackson C, Best N, Thomas A and Spiegelhalter D 2013 *The BUGS Book: A Practical Introduction to Bayesian Analysis* (New York: CRC Press) pp 57–79