African Journal of Business Management Vol.6 (44), pp. 11088-11094, 7 November 2012
Available online at http://www.academicjournals.org/AJBM
DOI: 10.5897/AJBM12.1001
ISSN 1993-8233 ©2012 Academic Journals

Full Length Research Paper

The importance of having mobile terminal samples for analyzing and verifying customer issues

Jarno Kankaanranta[1], Andi Mwegerano[2]*, and Ossi Hämeenoja[2]

[1]University of Turku, Ylhäistentie 2, 24130 Salo, Finland.
[2]Nokia Corporation, P.O. BOX 86, 24101 Salo, Finland.

Accepted 2 October, 2012

It may be intuitive to assume that the quality of issue corrective actions (iCA) for customers with mobile terminal (MT) issues can be improved by collecting samples of the faulty products in question for verifying the issues reported from the field. This study was created to test this hypothesis using three different established statistical analysis methods: binary logistic regression, Kruskal-Wallis statistical tests and Mood's Median Test. The methods are used because the data collected from in-house database tool is categorical. In this study, these methods were employed to test if the presence or absence of corrective actions process (CAP) samples had any effect on the perceived quality of issue corrective actions (P-QoiCA) or any effect on the perceived quality of issue resolution time (P-QoiRT) for service to the customers. This study also examined the effect of the quality of the collected samples (QoSa) on the absolute issue resolution time (iRT). The study checked if the frequency of requests for samples differs between MT products of different software platforms (SW_P). Additionally the paper investigated how the sample collection turnaround time (SC-TAT) differed in different sales areas (SA). The main findings were that the collected samples had no measurable effect upon the P-QoiCA or on P-QoiRT. Also it has been shown that QoSa and the SC-TAT had no effect on the absolute time (iRT) to resolve customers' issues. In addition to the aforementioned findings, the study noted as an aside that the Moods' median test found significant differences in iRT between different software platforms (SW_ P).

Key words: Corrective actions process (CAP)-samples, non-parametric tests, log-normal distribution, logistic regression, perceived quality.

INTRODUCTION

Several authors have written about service quality. Scholars like Cronin and Taylor (1992), Kettinger and Lee (1994) and Parasuraman et al. (1985, 1994) have created an understanding of the perceived quality of services, focusing often on the relationship of a company offering and service and the service quality as seen by the customer. The concept of service quality has also been extended to the after-sales of consumer goods (Shostack, 1977). Kasper and Lemmink (1989) have focused on understanding the differences on perceived service quality of customer and service managers. In this, perceived quality is defined as a customer's perception of the overall quality or services with intended purposes, relative to alternatives. As customers differ in their perceptual abilities, personal judgments and experience level, perceived quality will also vary accordingly. The response time and repair time are significant features in creating quality; this will ultimately lead to customer retention. The quality of the post-sales service supply chain is an important factor influencing overall competitiveness (Cohen et al., 2006). A good service quality is one which matches or exceeds customer

*Corresponding author E-mail: andi.mwegerano@nokia.com

expectations. Judgments of high or low service quality depending on how consumers perceive the actual service performance in the context they expected (Bergman and Klefsjö, 1994).

Debanjan and Golder (2006) have studied the most important factor about long-term success of products and firms. They argue that based on several other studies it is not quality itself; it is customers perceptions of quality. Golder et al. (2012) have proposed quality framework, called an integrative quality framework. In this framework there are three processes, the quality production process, the quality experience process and the quality evaluation process. As Golder et al. (2012) studied other academics in their framework they contended that in quality experience process we differentiate between perceived and delivered attributes. This means that what is delivers to customers and what customers perceive is not the same. Martínenez and Martínenez (2010) have done a collection of different service quality models. These models mathematically modeled the service quality as a function of several factors like expected service quality, perceived quality etc.

A customer support department in any business organization focuses on helping customers, answering questions or resolving issues with the products they sell and by providing services. Unfortunately, product support is often treated purely as overhead, a cost to be minimized- so customer support personnel measured on how fast they can “close” (not necessarily resolve) an incident. They are seldom measured on how well the incident was resolved or how happy the customer is. Customer support is one element that potentially affects on customer satisfaction. Support has a tremendous opportunity to influence the customer experience. It has the opportunity to turn frustrated and angry customers into loyal customers. It has also a tremendous opportunity to gain customer trust and gather information about the customer. The information gained in the support process may be most important for the support organization to improve their understanding of their customers (Kincaid, 2002). Customers consider response time as the most important among the many other items that affect customer satisfaction (Kasper and Lemmink, 1989). The faster the issues are resolved with high quality of correction actions, the greater the chance of winning over a customer (s). When attempting to resolve customers’ issues it is important to act on facts rather than on unverified claims or assumptions (Goetz and Davis, 2004). Based on Goetz and Davis (2004), if the mobile terminal resolvers cannot reproduce an issue reported by the customer, they should request mobile terminal samples having the issue. In this paper, the customer would be the authorized service vendor (ASV). In order to reproduce the issue, the sample should be untouched that is, neither unrepaired nor tried to be repaired. In addition to the sample a step by step description of how to reproduce the issue (s) is requested.

Research objectives

This paper analyzes the role (effect) of whether the collected CAP-samples:

1. Help to resolve service issues raised by the customer,
2. Decrease or increase issue resolution time (iRT) compared to the issues without samples,
3. Have an effect on the customer’s perceived quality of issue corrective actions (P-QoiCA) and the perceived quality of issue resolution time (P-QoiT) compared to the issues without samples, or
4. Have a different collection time in different sales areas (SA).

Also analyzed was the effect of the quality of samples (QoSa) on the issue resolution time (iRT) and the hypothesis if the frequency of the requests differs between products (mobile phones) of different software platform (SW_P). Additionally it was investigated how sample collection turnaround time (SC-TAT) differs depending on the sales area (SA).

Data gathering

For receiving feedback from the field, the mobile terminals’ manufacturer might request different kinds of samples for different purposes. This paper focuses only on corrective action process (CAP) samples. The process of collecting CAP-samples is described by the stream line block diagram (Figure 1).

Figure 1 show the process for collecting CAP samples for verifying issues provided by the customers

Where:

L1 – L5 are different levels in the chain of collecting sample
L1 Authorized service vendor (ASV)
L2Technical staff in the sales area
L3 Technical staffs in the regions area
L4 Care project managers working closely with R&D
L5 R&D developers

When an issue cannot be resolved by the lower level, the issue is escalated to a higher level for example, from L1 to L2. If the higher level cannot reproduce the issue with products that are at hand, it might request samples for verifying the issue from the lower level. Most of the samples come from L1. If L4 needs samples it will escalate its request to L2 stating how many samples are required and in which condition they must be. L2 re-escalates the request to L1 which in turn collects the

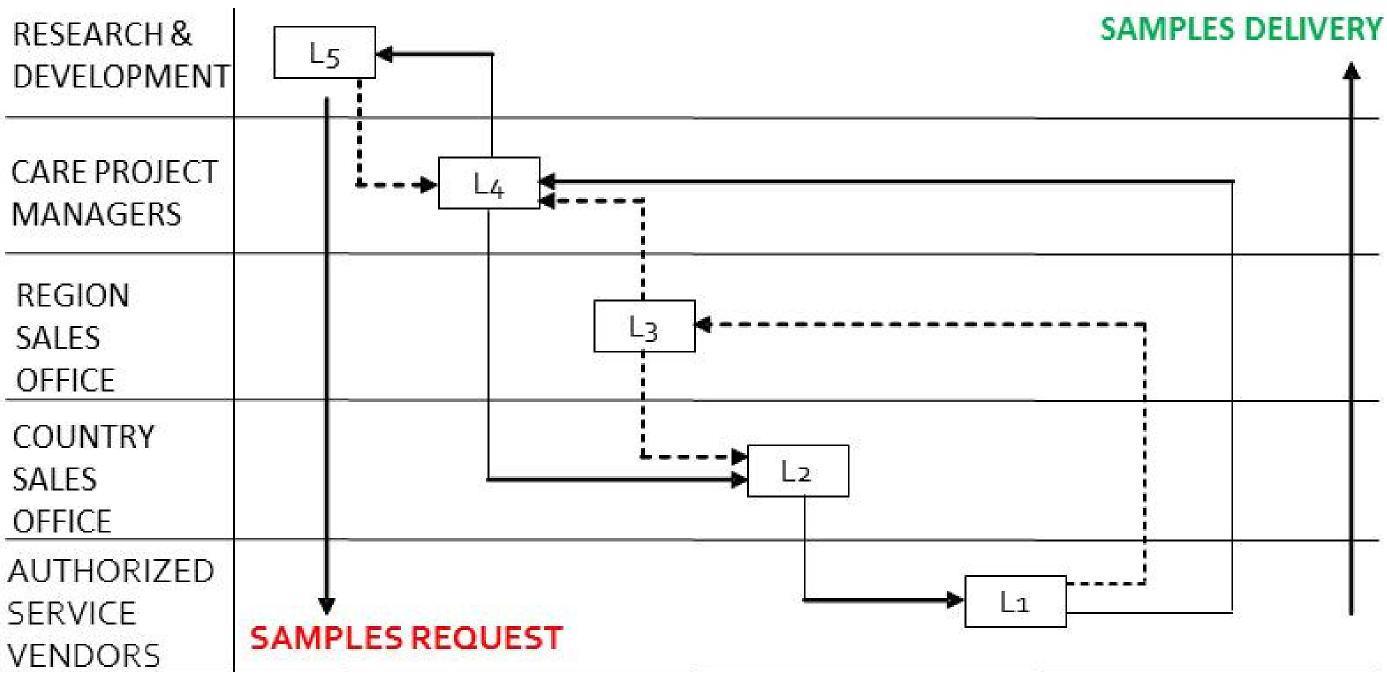

Figure 1. Process for collecting CAP samples for verifying issues provided by the customers.

samples and ships them to L4 for analysis. L3 is an option to some regions. L5 might request samples from L4.

In this work, a two years' worth of data concerning reported customer issues with or without samples was collected from the in-house database tool. Over 11000 customer issues without samples and over 800 customer issues with samples are used in this study work. The customer issues came from all over the world. All the collected issues were limited only to technical issues, that is, user interface issues, for example, accessory connectivity with mobiles terminals are not included in this study. The asymmetry proportion of the collected data (that is, data with samples and those without) has been addressed by the statistical methods applied in this paper.

The perceived quality of issue corrective actions (P-QoiCA) and perceived quality of time to resolve issue (P-QoT) are evaluated by customers, normally L1 (Figure 1), who presents the issue(s) before closing it. The evaluation is based on the following agreed statements: The P-QoiCA is evaluated as: 1) solved the customer's issue, 2) did not solve customer's issue and the P-QoT is evaluated as: 1) met the customer's expectation; 2) did not meet the customer's expectation.

METHODOLOGY AND THEORETICAL BACKROUND

This study approached the research questions through empirical data gathered from the in-house database tool. Data was analyzed using different statisticalanalysis methods. The method called binary logistic regression wasused to find out the possible association between the sample and perceived quality of corrective actions. The non-parametric One-Way Analysis of Variance (Kruskal and Wallis, 1952) and Mood's median tests were used to analyze the relationships of iRT and QoSa, SC-TAT and SW-Platform

Binary logistic regression

Regression analysis methods are fundamental part of any data analysis about describing the relationship between a response variable and one or more explanatory variables. The goal of logistic regression analysis is the same as the any other model-building technique used in statistics: to find the best fitting and parsimonious model, that describes the relationship between response variable and explanatory variable or variables (Hosmer and Lemeshow, 2000).

According to Hosmer and Lemenshow (2000) the binary logistic regression model could be presented using the following equation.

$$g(x) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x$$

In this equation the logarithm part is the link-function called logit. This link-function is linear in its parameters and maybe continuous and has value between minus infinity to infinity depending on the range of x (Hosmer and Lemeshow, 2000)

Measures of association

Part of our statistical analysis in this paper is to check the measures of association. In Minitab there are two options for that, Goodman-Kruskal λ and Goodman-Kruskal τ (Goodman and Kruskal, 1954). In Logistic regression the measures of association are Somers' D (Somers, 1962), Goodman-Kruskal γ (Goodman and Kruskal, 1954) and Kendall's τ-a (Kendall, 1938).

Goodman and Kruskal (1954) give two guidelines about measures of association and the kind of properties these measures should have. These guidelines are:

i) The measure of association takes values between -1 and 1. The value -1 or 1 means "complete association" and 0 means independence.
ii) The measure of association takes values between 0 and 1. The value 1 means "complete association" and 0 means independence.

Previously mentioned measures of associations could be classified into two classes. In Goodman-Kruskal λ and Goodman-Kruskal τ values are between 0 and 1. In Somers' D, Goodman-Kruskal γ and Kendall's τ-values are between -1 and 1. (Kendall, 1938; Goodman and Kruskal, 1954; Somers, 1962).

Test of homogeneity of variance

In Minitab there are two possibilities to test the homogeneity of variance. These tests are Bartlett's test and Levene's test. M. S. Bartlett introduced his testin 1937 and H. Levene introduced his test in 1960. (Bartlett, 1992; Levene, 1960). In Bartlett's test the statistic whose sampling distribution is closely approximated by the chi-square distribution with a-1 degrees of freedom when random samples are from independent normal populations. This test is very sensitive to the assumption of normality. Consequently, when the data is not normally distributed this test should not be used. Levene's test is not sensitive to departures from normality. So this test could be used in a situation when the data is continuous but not normally distributed (Tabachnick and Fidel, 2006; Montgomery, 2008).

Non-parametric one-way analysis of variance

A common problem in statistics is to conclude whether several samples should be regarded as coming from same population. Usually this kind of problem is solved using analysis of variance; a method developed by Fisher in the 1920s which provides an F-test for this problem. This method contains assumptions about the examined data: populations should be normally distributed and have the same variance (Kruskal and Wallis, 1952; Snedecor and Cochran, 1989).

Kruskal and Wallis (1952) explained several reasons why the use of ranks instead of the original observations is sometimes useful. One of these reasons is the fact that when we use ranks only very general assumptions about the distributions from which the observations come are done. In their article there were only some assumptions. They assumed that all observations are independent and the populations are approximately the same shape. Based on these facts the H test is introduced.

The interpretation of an H test is not a simple thing to do. A similar interpretation with an F test is not possible because this needs some information about the power of the test. For the H test, like many other nonparametric tests, the power is hard to examine. Thoroughly interpreted, it is possible to conclude that the significant value of H denotes that populations differ, though in fact this may not be the case (Kruskal and Wallis, 1952).

There is also another possibility for non-parametric analysis of variance: Mood's median test. This test is sometimes called a median test or sign scores test. Mood's median test is a kind of χ^2-test. In this test the null hypothesis is that medians are equal and the alternative hypothesis is that medians are not equal. In Mood's median test it is assumed that independent random samples taken from different populations have the same continuous distribution in shape. This test is more robust to the outliers than Kruskal-Wallis one-way analysis of variance but it is less powerful for data from many distributions (Breyfogle, 2003).

Log-normal distribution

A random variable X is said to be Log-normally distributed if log(X) is normally distributed (log(X) ~N(μ,σ)). In log-normal distribution only positive values are possible so the distribution is positively or right skewed. Two parameters are needed to specify log-normal distribution. These parameters are mean (μ) and standard deviation (σ) or variance (σ^2). The probability density function of this random variable X is possible to write according to following equation (Limpert et al., 2001).

$$f(x) = \frac{1}{x \cdot \sigma\sqrt{2\pi}} exp\left(-\frac{1}{2\sigma^2}(log(x)-\mu)^2\right)$$

RESULTS

Analysis of association between sample and perceived quality of corrective actions

First, basic descriptive statistics are calculated. It would be nice to find out in how many cases there were samples and in how many cases there were no samples. In the first case P-QoiT "did not meet the customers' expectations" and P-QoiCa were analyzed. In this case 67.7% of total cases were solved. In 201 issues out of 242 issues there were samples. The examination of association between variables sample and P-QoiCA (perceived quality of issues corrective actions) is done with Goodman-Kruskal λ and τ The value of λ is zero and τ is almost zero (0.0001578). So based on Goodman and Kruskal (1954), the interpretation of these both association measures is that there is no association between variables sample and P-QoiCA. Based on our analysis 89% of cases were solved though there were no samples available and all together there were 477 issues. The association between variables sample and P-QoiCA were similarly analyzed. The value of λ is zero and τ is almost zero (0.0000519). Based on association measures in these two previously mentioned cases there is no statistically significant association. According to these results it can be concluded that there is no need for sample collection.

Fitting the logistic regression model

Next the binary logistic regression model between variables sample, P-QoiCA and P-QoiT is fitted. The link-function used was logit. P-QoiCA and P-QoiT were predictors and sample variable was treated as the response. In our response (Sample), there were 12303

Table 1. Statistical analysis results.

Log-normal distribution		Levene's test	Kruskall-Wallis test or Moods Median test
QoSa	**p-value**	**p-value**	**p-value**
N/A	0.266	0.416	0.46
Not OK	0.633		
OK	0.022		
SC-TAT	p-value	p-value	p-value
SA1	0.065	0.404	<0.005
SA2	0.062		
SA3	<0.005		
SA4	0.069		
SA5	0.298		
SA6	0.008		
SA7	0.041		
SA8	0.123		
iRT	p-value	p-value	p-value
SW_P1	<0.005	<0.005	<0.005
SW_P2	<0.005		
SW_P3	<0.005		
SW_P4	<0.005		
SW_P5	<0.005		
SW_P6	<0.005		

cases altogether and in 11465 instances of these cases there were no samples. In the preliminary fitted model it was noticed that a 95% confidence interval of the odds ratio for P-QoiCA was from 0.79 to 1.27. Because the number 1 is included in this interval it can be concluded that P-QoiCA can be removed from the model. So in the fitted new model there were two variables: sample variable as a response and P-QoiT as a predictor.

The small (<0.05) p-value (0.002) for the second model shows that the model is significant. Based on the p-value, it can be said that one of the estimated coefficients is different from zero. According to our analysis degree of freedom (DF) for the second model is 1. So it is not possible to calculate any goodness of fit statistics. The reason for this is that the model takes all the degrees of freedom and there is no more left for goodness of fit statistics.

The association between response variable and predictor P-QoiT were also inspected. All association statistics (Somers' D, Goodman-Kruskal γ and Kendall's τ-a) are near zero. Based on Goodman and Kruskal (1954) it can be concluded that there is no statistically significant association between response variable and predictor P-QoiT.

This point out the fact that response sample has no effect upon P-QoiT. According to this result it can be concluded that there is no need for sample collection.

Differences between issue resolution time in dissimilar sample qualities, differences between different software platforms and sales units

Next it was analyzed how the quality of samples (QoSa) (N/A, NOT OK, OK) affects the issue resolution time (iRT). It is found that there are statistically significant differences between sample collection turnaround times (SC-TAT) in different sales areas (SA1…SA8). Also examined are the statistically significant differences in issue resolution time (iRT) between different software platforms (SW_P1…SW_8).\

First, the distribution of each dataset in the previously mentioned cases is analyzed. Because data is expressed in terms of times in all previously mentioned cases (issue resolution time=iRT and sample collection turnaround time=SC-TAT, there is no possibility that these values could be smaller than 0. Therefore it is possible to conclude that the suitable distribution might be Log-normal. Probability plot is used to analyze whether the data is Log-normally distributed and the results are provided in Table 1.

Statistically significant differences between issue resolution times (iRT) depending on differing quality of samples (QoSa) are analyzed. Because all the p-values are not large (>0.05) all the distributions are not log-normal and the parametric analysis of variance could not be used. The homogeneity of variances is assumed as well and it is satisfied because Levene's test p-value is large (>0.05) (Table 1, upper section). Similarly the distributions of sample collection turnaround time (SC-TAT) in different sales areas (SA) and homogeneity of variance are analyzed. All p-values are not large (>0.05), which means that all the distributions are not log-normal and so Kruskall-Wallis one-way analysis of variance needs to be used here as well. In this case the variances could be assumed to be equal because Levene's test p-value is large (>0.05) (Table 1, middle section). Regarding statistically significant differences between issue resolution time (iRT) in different software platforms (SW_P): The p-values are all small (<0.05) for testing the log-normal distribution and the homogeneity of variance is not satisfied, and Levene's test p-value is small (<0.005) (Table 1, lower section). Because the variances are not equal, it is not possible to use the Kruskall-Wallis one-way analysis of variance: Moods median test needs to be used instead. Moods median test is used because the only condition needed is the same shape of distributions.

There are no statistically significant differences between issue resolution times (iRT) in different quality of samples. This conclusion is based on large (>0.05) p-value in Kruskall-Wallis one-way analysis of variance (Table 1, upper section). The difference between sample collection time (SC-TAT) in different sales areas are statistically significant, because the p-value in Kruskall-Wallis one-way analysis of variance is small (<0.05) (Table 1, middle section). There are as well statistically significant differences between issue resolution times (iRT) in different software platforms (SW-P). This conclusion is based on small p-value (<0.05) in Moods median test (Table 1, lower section).

DISCUSSION AND CONCLUSIONS

Based on the Goodman and Kruskal one way analysis of variance in this study case, variable samples have no effect on perceived quality of issue correction actions (P-QoiCA) or on perceived quality of issue resolution time (P-QoiT). Also according to Kruskal-Wallis one way analysis of variance was noted that: 1) The sample quality (QoSa) has no effect on issue resolution time (iRT) 2) The sample collection turnaround time (SC-TAT) differs from sales to sales area (SA). According to Moods' median test it was noted that there are significant differences in issue resolution times (iRT) in different software platforms (SW_PL). This is based on the fact that p-value < 0.005 (Table 1, lower section). The counter- intuitive results might be due to the fact that there was a long SC-TAT from different SA globally as perceived by issue resolvers (Mwegerano et al., 2012). The long lead of SC-TAT might be caused by different custom regulations in different countries. Another factor which has not been investigated is how quick the resolvers investigate samples when they receive them. Further research could be done to investigate how fast the resolvers verify reported issues with the samples received. On top of this it could be precisely investigated the ratio of good and bad samples for verifying the issue. If bad sample are received what do resolvers do? Do they ask for more samples? Or do they resolve the issue without the sample in that case? Good samples are the ones which they could easily reproduce the issues reported from the field. In a nut shell in the future, based on the hypotheses reasons and un answered questions mentioned e earlier, this study could be expanded to research why samples collected to help resolve issues have no effect on P-QoiCA, P-QoiT and also why the sample quality QoSa has no effect on iRT.

NOMENCLATURE

CAP: Corrective action process; **iRT:** Absolute issue resolution time; **P-QoiCA:** Perceived quality of issue correction action as perceived by a customers; **P-QoT:** Perceived quality of time spent on resolving an issue as perceived by a customer; **SW_P:** Software platform-The subtype of operating system used in mobile phones e.g.S30, S40, S60 etc. where S is the Symbian operating system; **SC-TAT:** Sample Collection Turnaround Time; **SA:** Sales area-Countries where the mobile sales activities are conducted; **QoSa:** Quality of the samples.

REFERENCES

Bartlett MS (1992). Properties of Sufficiency and Statistical Tests. In Kotz S. Breakthroughs in Statistics: Foundations and Basic Theory vol. 1. Springer-Verlag Berlin and Heidelberg GmbH & Co.

Bergman B, Klefsjö B (1994). Quality from Customer Needs to Customer Satisfaction. Studentliteratur AB.

Breyfogle FW (2003). Implementing Six Sigma: Smarter Solutions Using Statistical Methods. John Wiley & Sons.

Cohen MA, Graval N, Agrawal V (2006). Winning in the Aftermarket. Harvard Bus. Rev. 84(5):129-138.

Cronin JJ, Taylor SA (1992). Measuring Service Quality – Reexamination and Extension. J. Mark. 56(3):55-68.

Debanjan M, Golder PN (2006). How Does Objective Quality Affect Perceived Quality? Short-Term Effects, Long-Term Effects, and Asymmetries. Market. Sci. 25(3):230-247.

Goetz D, Davis SB (2004). Effective Customer Service. Ten Steps for Technical Profession. Pearson.

Golder PN, Debanjan M, Moorman C (2012). What is Quality? An Integrative Framework of Processes and States. J. Mark. 76(4):1-23.

Goodman LA, Kruskal WH (1954). Measures of Association for Cross Classifications. J. Am. Stat. Assoc. 49(268):732-765.

Hosmer DW, Lemenshow,S (2000). Applied Logistic Regression. Wiley-Interscience Publication.

Kasper H, Lemmink J (1989). After Sales Service Quality: Views Between Industrial Customers and Service Managers. Ind. Mark. Manag. 18(3):199-208.

Kendall MG (1938). A New Measure of Rank Correlation. Biometrika 30(1/2):81-93.

Kettinger WJ, Lee CC (1984). Perceived Service Quality and User Satisfaction with the Information-services Function. Decis. Sci. 25(5, 6):737-766.

Kincaid WJ (2002). Customer Relationship Management: Getting It Right. Prentice Hall.

Kruskal WH, Wallis WA (1952). Use of Ranks in One-Criterion Variance Analysis. J. Am. Stat. Assoc. 47(260):583-621.
Levene H (1960). Robust Test for Equality of Variances. In Olkin I. Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Stanford University Press.
Limpert E, Stahel WA, Abbt M (2001). Log-normal Distributions Across the Sciences: Keys and Clues. BioScience 51(5):341-352.
Martínenez JA, Martínenez L (2010). Some Insights on Conceptualizing and Measuring Service Quality. J. Retail. Consum. Serv. 17(1):29-42.
Montgomery DC (2008). Design and Analysis of Experiments. Wiley.
Mwegerano A, Kytösaho P, Tuominen A (2012). Applying Self-Organizing Maps Method to Analyze the Corrective Action's Quality Provided to Customers with Mobile Terminals. iBusiness 4(2):108-120.
Parasuraman A, Zeithaml VA, Berry LL (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. J. Mark. 44(4):41-50.
Parasuraman A, Zeithaml VA, Berry LL (1994). Reassessment of Expectations as a Comparison Standard in Measuring Service Quality – Implications for Further Research. J. Mark. 58(1):111-124.
Shostack GL (1977). Breaking Free from Product Marketing. J. Mark. 41(1):73-80.
Snedecor GW, Cochran WG (1989). Statistical Methods. Iowa State University Press.
Somers RH (1962). A New Asymmetric Measure of Association for Ordinal Variables. Am. Sociol. Rev. 27(6):799-811.
Tabachnick BG, Fidel LS (2006). Using Multivariate Statistics. Allyn & Bacon.