PhD THESIS

SERVICE QUALITY MEASUREMENT: A NEW METHODOLOGY

Author: ANDRÉS REDCHUK

Advisors: Dr. JAVIER MARTÍNEZ MOGUERZA
Dr. CLARA LAURA CARDONE RIPORTELLA

Department of Statistics and Operations Research
Rey Juan Carlos University

Madrid - Spain
April of 2010

PhD THESIS

SERVICE QUALITY MEASUREMENT: A NEW METHODOLOGY

By ANDRÉS REDCHUK

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT
REY JUAN CARLOS UNIVERSITY
Madrid - Spain
April of 2010

Don Javier Martínez Moguerza, Profesor Titular de Universidad del Departamento de Estadística e Investigación Operativa de la Universidad Rey Juan Carlos y **Da. Clara Laura Cardone Riportella**, Profesora Titular de Universidad del Departamento de Economía de la Empresa de la Universidad Carlos III de Madrid.

AUTORIZAN:

La presentación de la Tesis Doctoral titulada

SERVICE QUALITY MEASUREMENT: A NEW METHODOLOGY

Realizada por **Don Andrés Redchuk** bajo su inmediata dirección y supervisión en el Departamento de Estadística e Investigación Operativa y que presenta para la obtención del grado de Doctor por la Universidad Rey Juan Carlos.

En Fuenlabrada, a 12 de abril de 2010.

Fdo.: Javier Martínez Moguerza

Fdo.: Clara Laura Cardone Riportella

Don Antonio Alonso Ayuso, Profesor Titular de Universidad y Director del Departamento de Estadística e Investigación Operativa de la Universidad Rey Juan Carlos.

INFORMA:

Que la presente Tesis Doctoral titulada

SERVICE QUALITY MEASUREMENT: A NEW METHODOLOGY

Ha sido realizada por Don Andrés Redchuk bajo la dirección y supervisión de Don Javier Martínez Moguerza y Da. Clara Laura Cardone Riportella y que el Departamento de Estadística e Investigación Operativa ha dado su conformidad para que sea presentada ante la Comisión de Doctorado.

En Fuenlabrada, a 12 de abril de 2010.

Fdo.: Antonio Alonso Ayuso

The undersigned hereby certify that they have read and recommend to the School of Telecommunication Engineering for acceptance a thesis entitled

"Service Quality Measurement: A New Methodology"

by Andrés Redchuk

in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Fuenlabrada, April of 2010

Javier Martínez Moguerza

Advisor

Clara Laura Cardone Riportella

Advisor

Table of Contents

List of Tables

List of Figures

Abstract

The aim of this work is to present a new methodology to measure the quality of a service. A nonparametric model is developed in which customers evaluate the overall service quality and a set of dimensions or attributes that determine this service quality.

The model assumes that overall service quality is determined by a linear combination of attributes evaluations with some unknown weights and that different customers may have different weights for the attributes.

The nonparametric techniques are based in Nearest Neighbours combined with Restricted Least Squared methods.

The model is applied to several simulated data sets where we know the true value of the parameters of the model.

Then we have applied the methodology to a specific set of data from CABINTEC ("Intelligent cabin truck for road transport").

Finally, the methodology is applied to the measurement of the quality of the postgraduate courses of a public Spanish University.

The methodology, that we call ALR Adaptive Local Regression, have demonstrate be able to treat these kind of data.

ALR permits to calculate the weight that customer assigns to each quality attribute of the service.

Resumen

En esta tesis doctoral presentamos una nueva metodología para la medición de la calidad de los servicios. Se desarrolla un modelo no paramétrico partiendo de la información aportada por los clientes, que evalúan la calidad total de un servicio y la de un conjunto de dimensiones de la calidad o atributos del mismo.

El modelo utilizado asume que la calidad total del servicio está determinada por una combinación de los atributos con un peso desconocido y que cada cliente puede asignar diferentes pesos a cada uno de esos atributos. La metodología resultante se ha denominado ALR (Adaptive Local Regression), regresión local adaptativa, y está basada en técnicas de remuestreo (resample) y de los K vecinos más próximos (Nearest K Neighbours) combinado con Mínimos Cuadrados con Restricciones (Restricted Least Squared methods).

Para conocer y validar la bondad de la metodología ALR, hemos aplicado dicha metodología a sendos conjuntos de datos simulados en los cuales se conocen a priori los verdaderos valores de los parámetros del modelo.

Luego aplicamos la metodología a un conjunto específico de datos provenientes de CABINTEC (“Intelligent cabin truck for road transport”).

Por último la metodología es aplicada a la medición de la calidad de los cursos de postgrado de una universidad pública Española.

Los resultados obtenidos demuestran que ALR es apta para el tratamiento de ese tipo de datos.

Antecedentes

El concepto de calidad ha evolucionado a lo largo del tiempo. A lo largo de la historia, la metodología de elaboración de productos y servicios y el concepto de calidad han ido evolucionando de forma paralela.

Época/Sistema de Gestión	**Concepto de Calidad**
Época artesanal	Hacer las cosas bien a cualquier costo
Industrialización	Producción
Segunda Guerra Mundial	Eficacia + Plazo = Calidad
Posguerra (Japón)	Hacer las cosas bien a la primera
Posguerra (resto de países)	Producción
Control de Calidad	Verificación de las características del producto
Gestión de la Calidad	Aptitud del producto/servicio al uso
Gestión de Calidad Total	Integrar la calidad en todo el proceso
Taguchi	Coste mínimo para la sociedad

Las empresas más comprometidas en materia de calidad han incorporado sistemas de gestión basados en la Gestión de Calidad Total. Este proceso supone integrar el concepto de calidad en todas las fases del proceso y a todos los departamentos que tienen alguna influencia en la calidad final del proceso y/o servicio prestado al cliente.

Con diversas motivaciones, implantaciones de normativas del tipo ISO, despliegues de modelos de excelencia del tipo EFQM, las organizaciones necesitan medir y monitorizar en nivel de calidad de sus servicios/productos.

Definición de Medición

Medir es comparar con una magnitud que se utiliza como patrón de referencia (ISO 31). Las métricas adoptadas para estas mediciones pueden ser objetivas o subjetivas, pueden depender en mayor o menor medida del proceso de medición; del método utilizado, de la clase del instrumento, del que mide y del objeto medido.

Cuando medimos un producto, estas métricas, en general, son objetivas (p.e. la dureza, la longitud, tiempos, etc). Si utilizamos métricas objetivas tenemos la ventaja de contar con técnicas muy maduras y validadas para reducir posibles errores y obtener medidas fiables, podemos fácilmente acotar los errores sistemáticos de método, los errores de clase del instrumento, etc (entre estas técnicas podemos citar el R&R técnica de repetibilidad y reproducibilidad, el MSA análisis del Sistema de Medición, el FMEA análisis Modal de fallos y efectos, el APQP, entre otros).

Diferencias entre Productos y servicios

Existen diferencias notables entre un producto y un servicio. Se intenta hacer coincidir las estrategias y herramientas utilizadas, adaptando técnicas y métodos, pero los resultados no son del todo aceptables. El principal elemento diferenciador radica en la intangibilidad de los servicios, que no permite que podamos percibirlo mediante los sentidos.

A continuación detallamos algunas de las características diferenciadoras:

- La *no estandarización de los servicios*, es casi imposible que se repitan dos servicios iguales, sumado a que sobre dos servicios equivalentes es muy difícil que se repita la misma valoración.
- *No se pueden* probar, no podemos devolverlo si no nos gusta.
- *La inseparabilidad*, no podemos separar el servicio de quien lo presta, sumado a que la producción del servicio va unida al consumo del mismo.
- Los servicios no se pueden *ni almacenar ni transportar* y por tanto tampoco se pueden intercambiar.
- Los servicios, en principio, *son perecederos* ya que deben utilizarse para el momento que fueron previstos y no posteriormente.
- En los servicios, la empresa está en *contacto directo con el cliente*, en los productos, difícilmente se da este hecho.
- El *cliente participa en la producción* de los servicios, en los productos muy raramente.

- Los servicios son *difíciles de valorar*, por ello en la presentación de ofertas pueden existir grandes diferencias, principalmente de precios.
- La gestión de los servicios es más *problemática* que la de los productos.

En esta tesis doctoral nosotros desarrollamos y aplicamos una nueva metodología para la *medición* de la *calidad* de los servicios.

Objetivos

Los objetivos que se han planteado son los siguientes:

- Revisar la *literatura existente* sobre medición de la calidad de los servicios.
- Revisar las *técnicas existentes*, técnicas clásicas, para la medición de la calidad de los servicios.
- Desarrollar una nueva metodología para la medición de la calidad de los servicios, que permita medir la importancia que *cada cliente asigna a cada aspecto (atributo) del servicio* y que sea universal (que no dependa de las características particulares de los datos recogidos).
- *Aplicar y validar* la metodología desarrollada.

Metodología

Supongamos que tenemos una población de clientes. Esta población incluye nuestros actuales clientes, podríamos pensar también en nuestros clientes potenciales. Asumimos que el tamaño de la población, n, es grande.

Denominamos Q al vector cuyos elementos son los Q_i, esto es, la calidad percibida de un servicio dado por el i_{TH} cliente de la población.

Es común asumir que la evaluación del cliente será una función de diferentes, k, atributos de calidad $X_1,\ldots;X_k$ que son los que determinan la evaluación global del servicio.

Denominamos X_i al vector cuyos elementos son los $X_{i1},\ldots;X_{ik}$, esto es, la evaluación de los atributos de calidad realizada por el i_{TH} cliente.

Entonces,

$$Q_i = f(X_{i1},\ldots,X_{ik}) \ .$$

Podemos aproximar este indicador lineal por:

$$Q_i = \sum_{j=1}^{k} w_{ij} X_{ij} \quad ,$$

donde los coeficientes w_{ij}, en la matriz W, son los pesos definidos por cada cliente. Estos pesos son todos positivos y deben sumar uno:

$$w_{ij} \geq 0 \qquad \forall i, \forall j,$$

$$\sum_{j=1}^{k} w_{ij} = 1 \quad \forall i.$$

Estos pesos pueden ser interpretados como la importancia relativa del atributo X_j en la determinación de la evaluación de la calidad del servicio del i_{TH} cliente.

ALR (Adaptative Local Regression)

Para desarrollar nuestro modelos de calidad, necesitamos:

- La lista completa de los atributos de calidad
- Los pesos asignados a cada atributos

La tarea más importante es la obtención de los pesos, ya que nosotros siempre vamos a poder escribir una lista muy exhaustiva de atributos de calidad aunque algunos de ellos resulten tener pesos igual a cero.

Asumimos ciertas hipótesis iniciales:

- H1: Existe una función $f \mid f(X) = Q$.
- H2: Relajamos H1 asumiendo que f es una función lineal a trozos. Dejamos las funciones no lineales para investigaciones futuras.

Entonces nuestro modelo, localmente, es $w^T X \approx Q$, debido a que $w^T X$ es una aproximación lineal de f.

Con el siguiente algoritmo, podemos estimar cada una de las componentes de la matriz W mediante $\widehat{W}$:

Para cada cliente $i = 1,\dots,n$

Paso 1.

Calculamos sus l ε-vecinos razonables, esto es $X_{(1)},\dots,X_{(l)}$.

donde $(1),\dots,(l)$ es un reordenamiento apropiado de los k índices en el conjunto $\{1,\dots,n\}$ y $l \geq k$.

Paso 2.

Construir X^i y Q^i:

$$X^i = \begin{bmatrix} X_{i1} & \dots & \dots & X_{ik} \\ X_{(1)1} & \dots & \dots & X_{(1)k} \\ \dots & \dots & \dots & \dots \\ X_{(l)1} & \dots & \dots & X_{(l)k} \end{bmatrix}, \quad Q^i = \begin{bmatrix} Q_i \\ Q_{(1)} \\ \dots \\ Q_{(l)} \end{bmatrix}$$

Paso 3.

Resolver las posibles redundancias numéricas en la matriz $\left[X^i \,|\, Q^i\right]$.

Paso 4.

Estimar el vector W_i como $\widehat{W}_i = \left[\widehat{w}_{i1}, \widehat{w}_{i2},\dots,\widehat{w}_{ik}\right]$, resolviendo los sistemas $X^i \widehat{w}_i = Q^i$ mediante el método de mínimos cuadrados con restricciones (least squares method with linear constraints).

La diagnosis del modelo local obtenido y sus utilización posterior se realiza con la técnicas tradicionales de la estadística multivariante

Conclusiones

En tesis doctoral hemos alcanzado todos los objetivos propuestos. Los resultados obtenidos demuestran que ALR es apta para el tratamiento de ese tipo de datos.

Cabe destacar que esta nueva metodología propuesta para la medición de la calidad de los servicios (ALR) presenta varias ventajas respecto a las técnicas clásicas:

- Podemos utilizar computación en paralelo para resolver los sistemas de ecuaciones que surgen a lo largo de los cálculos (Paso 4 del algoritmo).
- Cuando el encargado de la toma de decisión, precise un indicador único, siempre se podrá definir:

$$\overline{w_j} = \frac{\sum_{j=1}^{n} \widehat{w}_{ij}}{n} .$$

  Aunque esta reducción de todas las dimensiones de la evaluación a un solo número puede dar lugar a múltiples críticas, sin embargo, es suele ser necesario. No deja de ser una alternativa para la estimación de los pesos que provee resultados equivalentes a los métodos clásicos cuando se trata de trabajar con un grupo único y homogéneo de clientes.
- Podemos estimar cada componente $\widehat{w}_{ij}$. Una vez realizado estos cálculos podemos aplicar cualquier técnica multivariante para determinar nuevos grupos de clientes e inferir sobre ellos, en los que podríamos definir como segmentación a posteriori.
- Podemos trabajar directamente con cada uno de los pesos de cada cliente asigna a cada atributo de calidad. De hecho, no aceptamos a la media de los pesos como un buen estimador. Recordemos que la media solamente será un buen estadístico descriptivo cuando tengamos una muestra homogénea y lo será muy malo cuando tengamos una mezcla de segmentos muy distintos de clientes.
- Nosotros estimamos los pesos que cada cliente asigna a cada atributos e calidad con la información que obtenemos de clientes similares, recurrimos a clientes que presentan características similares. Podemos decir que recurrimos a un criterio muy intuitivo y natural.

Acknowledgements

First and foremost, I would like to thank my advisors, Professor Javier Moguerza and Professor Clara Cardone, for their invaluable directions, comments, stimulating discussions and continuous support throughout my PhD research. I feel extremely lucky to have Javier and Clara as my advisors.

I am grateful to Professor Antonio Alonso, Head of the Department of Statistics and Operations Research at Rey Juan Carlos University, for his kind help during my research.

I would also like to extend special thanks to the members of the Department of Statistics and Operations Research at Rey Juan Carlos University.

I am also grateful to Javier Prieto, for his collaboration during the first steps of my research life.

I would like to thanks to the members of CABINTEC project (Spanish MEC project PSE37010020072).

Last but not least, my family deserves particular recognition for their greatest love and support.

Specially, I am grateful to my parents for their encouragement and to my wife and my three princesses, for their love and belief in my studies, without which all that I have achieved would not have been possible.

Without any doubt, this project has been the longest and the most complicated of all in my life. In this project, all the management variables have been changed, and I am sure it would not have arrived to a happy ending without Javier Moguerza's participation.

Fuenlabrada, april of 2010.

Ten women have put order to my way and have leaded my life.

Following the order of appearing in scene, they are:

Mamá,

Gricelda,

Edy,

Lorena,

Liliana,

Chola,

Rita,

Sophie,

Vicky and

Érika.

For them is and will be my eternal gratitude.

She:	I only went through 11th grade.
She:	How far did you go in school?.
He:	I went all the way.
She:	Your folks must be really proud.
He:	... (*Silence*)

Pretty Woman. Paramount Films (1990).

Agradecimientos

Primero y principal quiero agradecer a mis directores, el Profesor Javier Moguerza y la profesora Clara Cardone, por sus brillantes aportaciones, su claridad y su continua ayuda en mi tesis doctoral. Me siento muy afortunado de tener a Javier y a Clara como directores.

Quiero agradecer al profesor Antonio Alonso, director del Departamento de Estadística e Investigación Operativa de la Universidad Rey Juan Carlos.

Me gustaría extender mi agradecimiento a todo el Departamento de Estadística e Investigación Operativa de la Universidad Rey Juan Carlos.

Estaré siempre agradecido con Javier Prieto, catedrático del departamento de Estadística de la Universidad Carlos III de Madrid por toda su ayuda en mis primeros pasos en el mundo de la investigación.

Me gustaría agradecer a los miembros del proyecto CABINTEC (proyecto del MEC PSE37010020072).

Por último pero no menos importante, quiero agradecer a mi familia por amor y apoyo incondicional. En particular, mi agradecimiento más profundo es para mis padres, mi mujer y mis tres princesas, por todo su amor, su confianza e mis estudios y su tiempo.

Sin lugar a dudas, este proyecto es el más largo y complicado de toda mi vida. En este proyecto se han modificado, prácticamente, todas las variables de gestión y estoy seguro que no hubiese llegado a buen término sin la entrada en escena de Javier Moguerza.

Fuenlabrada, abril de 2010.

Chapter 1

Introduction

In this chapter, the research motivation, objective problem of the thesis and the notation used throughout of this work are presented.

1.1 Research Motivation

Quality service has come to be recognized as a strategic tool for attaining operational efficiency and improved business performance (Anderson and Zeithaml, 1984; Babakus and Boller, 1992; Garvin, 1983; Garvin, 1984; Garvin, 1987; Phillips, Chang and Buzzell, 1983; Hendrick et al, 2001; Seth, Deshmukh and Vrat, 2005). Several authors have discussed the importance of quality to service firms (Cook Goh and Chung, 1999; Normann, 1984; Shaw, 1978; Horovitz, 2001; Parasuraman, 2002) and have demonstrated its strong relationship with profits, increased market share, return on investment (ROI), customer satisfaction, and future purchase intentions (Anderson, Fornell and Lehmann 1994; Boulding et al., 1993; Buzzell and Gale, 1987; Rust and Oliver, 1994; Llusar et al, 2001). One direct conclusion of these studies is that firms with superior quality outperform those marketing inferior quality (Zemke, 1999; Brogowicz et al, 2001; Gustafsson et al, 2003; ASCI, 2008).

Notwithstanding the recognized importance of service quality, there have been methodological issues and application problems with regard to its operationalization (Stamatis, 2003). Quality in the context of services has been conceptualized differently and based on different conceptualizations, alternative methodologies have been proposed for service quality measurement (see Brady, Cronin and Brand, 2002; Cronin and Taylor, 1992, 1994; Dabholkar, Shepherd and Thorpe, 2000; Parasuraman, Zeithaml and Berry, 1985, 1988).

Despite considerable work undertaken in the area, there is no consensus yet as to which one of the methodology is robust enough for measuring and comparing service quality. One major problem with past studies has been their preoccupation with assessing psychometric and methodological soundness of service scales in the context of service industries in the developed countries (Cronbach 1951). Virtually no empirical efforts have been made to evaluate the diagnostic ability of the scales in providing managerial insights for corrective actions in the event of quality shortfalls. Furthermore, little work has been done to examine the applicability of these scales to the services in developing countries (Fullerton 2005).

1.1.1 Evolution of quality

Quality has been defined differently by different authors. Some prominent definitions include:

Conformance to requirements	Crosby (1984)
Fitness for use	Juran (1988, 2001)
One that satisfies the customer	Eiglier and Langeard (1987)
Zero defects	Japanese Philosophy (Taguchi 2005)

Table 1.1: Quality definitions

Though initial efforts in defining and measuring service quality emanated largely from the goods sector, a solid foundation for research work in the area was laid down in the mid eighties by Parasuraman, Zeithaml and Berry (1985). They were amongst the earliest researchers to emphatically point out that the concept of quality prevalent in the goods sector is not extendable to the services sector (ISO 175001: 2004 and ISO 66992: 2001). Being inherently and essentially intangible, heterogeneous, perishable, and entailing simultaneity and inseparability of production and consumption, services require a distinct framework for quality explication and measurement (Hamer, 2003; Mitra, 2003). As against the goods sector where tangible cues exist to enable consumers to evaluate product quality, quality in the service context is explicated in terms of parameters that largely come under the domain of 'experience' and 'credence' properties and are as such difficult to measure and evaluate (Parasuraman, Zeithaml and Berry, 1985; Zeithaml and Bitner, 2001; Kang, 2004).

1.1.2 Service quality

One major contribution of Parasuraman, Zeithaml and Berry (1988) was to provide a terse definition of service quality. They defined service quality as *"a global judgment, or attitude, relating to the superiority of the service"* and explicated it as involving evaluations of the outcome (i.e., what the customer actually receives from service) and process of service act (i.e., the manner in which service is delivered) (ISO10002: 2004). In line with the propositions put forward by Gronroos (1982) and Smith and Houston (1982), posited and operationalized service quality as a difference between consumer expectations of 'what they want' and their perceptions of 'what they get.' Based on this conceptualization and operationalization, they proposed a service quality measurement scale called "SERVQUAL".

The SERVQUAL scale constitutes an important landmark in the service quality literature and has been extensively applied in different service settings. Over time, a few variants of the scale have also been proposed.

The 'SERVPERF' scale is one such scale that has been put forward by Cronin and Taylor (1992) in the early nineties. Numerous studies have been undertaken to assess the superiority of the two scales, but consensus continues to elude as to which one is a better scale.

1.1.3 SERVQUAL scale

The foundation for the SERVQUAL scale is the gap model proposed by Parasuraman, Zeithaml and Berry (1985, 1988) and several amplifications (Clement and Selvam, 2006). With roots in disconfirmation paradigm, the gap model maintains that satisfaction is related to the size and direction of disconfirmation of a person's experience vis-à-vis his/her initial expectations (Churchill and Surprenant, 1982; Parasuraman, Zeithaml and Berry, 1985; Smith and Houston, 1982). As a gap or difference between customer 'expectations' and 'perceptions,' service quality is viewed as lying along a continuum ranging from 'ideal quality' to 'totally unacceptable quality,' with some points along the continuum representing satisfactory quality. Parasuraman, Zeithaml and Berry (1988) held that when perceived or experienced service is less than expected service, it implies less than satisfactory service quality. But, when perceived service is less than expected service, the obvious inference is that service quality is more than satisfactory.

Parasuraman, Zeithaml and Berry (1988) posited, inspired in Kano model of preference analysis, that while a negative discrepancy between perceptions and expectations — a 'performance-gap' as they call it —causes dissatisfaction, a positive discrepancy leads to consumer delight (Naumann, 1995; Karsak et al, 2002).

Based on their empirical work, they identified a set of 22 variables/items tapping five different dimensions of service quality construct. Since they operationalized service quality as being a gap between customer's expectations and perceptions of performance on these variables, their service quality measurement scale is comprised of a total of 44 items (22 for expectations and 22 for perceptions). Customers' responses to their expectations and perceptions are obtained on a 7-point Likert scale and are compared to arrive at (P-E) gap scores. The higher (more positive) the perception minus expectation score, the higher is perceived to be the level of service quality. In an equation form, their operationalization of service quality can be expressed as follows:

$$SQ_i = \sum_{j=1}^{k} \left(P_{ij} - E_{ij} \right), \qquad (1.1)$$

where:

SQi = perceived service quality of individual i_{TH} .

k = number of service attributes/items.

P = perception of individual i_{TH} with respect to performance of a service firm attribute j_{TH} .

E = service quality expectation for attribute j_{TH} that is the relevant norm for individual i_{TH} .

The importance of Parasuraman, Zeithaml and Berry's (1988) scale is evident by its application in a number of empirical studies across varied service settings (Brown and Swartz, 1989; Carman, 1990; Kassim and Bojei, 2002; Lewis, 1987 and 1991, 1991; Pitt, Gosthuizen and Morris, 1992; Witkowski and Wolfinbarger, 2002; Young, Cunningham and Lee, 1994). Despite its extensive application, the SERVQUAL scale has been criticized on various conceptual and operational grounds. Some major objections against the scale relate to the use of (P-E) gap scores, length of the questionnaire, predictive power of the instrument, and validity of the five-dimension structure (e.g., Babakus and Boller, 1992; Cronin and Taylor, 1992; Dabholkar, Shepherd and Thorpe, 2000; Teas, 1993, 1994).

Several issues have been raised with regard to the use of (P-E) gap scores, i.e., disconfirmation model. Most studies have found a poor fit between service quality as measured through Parasuraman, Zeithaml and Berry's (1988) scale and the overall service quality measured directly through a single-item scale (Babakus and Boller, 1992; Babakus and Mangold, 1989; Carman, 1990; Finn and Lamb, 1991; Spreng and Singh, 1993). Though the use of gap scores is intuitively appealing and conceptually sensible, the ability of these scores to provide additional information beyond that already contained in the perception component of service quality scale is under doubt (Babakus and Boller, 1992; Iacobucci, Grayson and Ostrom, 1994). Pointing to conceptual, theoretical, and measurement problems associated with the disconfirmation model, Teas (1993, 1994) observed that a (P-E) gap of magnitude '-1' can be produced in six ways: P=1, E=2; P=2, E=3; P=3, E=4; P=4, E=5; P=5, E=6 and P=6, E=7 and these tied gaps cannot be construed as implying equal perceived service quality shortfalls. In a similar way, the empirical study (by Peter et al, 1993; Brown et al 1993) found difference scores being beset with psychometric problems and, therefore, cautioned against the use of (P-E) scores (Sureshchandar, Rajendran and Anantharaman 2002).

Validity of (P-E) measurement framework has also come under attack due to problems with the conceptualization and measurement of expectation component of the SERVQUAL scale. While perception (P) is definable and measurable in a straightforward manner as the consumer's belief about service is experienced, expectation (E) is subject to multiple interpretations and as such has been operationalized differently by different authors (e.g., Babakus and Inhofe, 1991; Brown and Swartz, 1989; Dabholkar et al., 2000; Gronroos, 1990; Teas, 1993, 1994). Initially, Parasuraman, Zeithaml and Berry (1985, 1988) defined expectation close on the lines of Miller (1977) as “desires or wants of consumers”, i.e., what they feel a service provider should offer rather than would offer. This conceptualization was based on the reasoning that the term ‘expectation’ has been used differently in service quality literature than in the customer satisfaction literature where it is defined as a prediction of future events, i.e., what customers feel a service provider would offer. Parasuraman, Berry and Zeithaml (1990) labelled this “should be’” expectation as “normative expectation”, and posited it as being similar to “ideal expectation’” (Zeithaml and Parasuraman, 1991). Later, realizing the problem with this interpretation, they themselves proposed a revised expectation (E*) measure, i.e., what the customer would expect from ‘excellent’ service (Parasuraman, Zeithaml and Berry, 1994). It is because of the vagueness of the expectation concept that some researchers like Babakus and Boller (1992), Bolton and Drew (1991a, 1991b), Brown, Churchill and Peter (1993), and Carman (1990) stressed the need for developing a methodologically more precise scale.

The SERVPERF scale, developed by Cronin and Taylor (1992), is one of the important variants of the SERVQUAL scale. For being based on the perception component alone, it has been conceptually and methodologically posited as a better scale than the SERVQUAL scale which has its origin in disconfirmation paradigm.

1.1.4 Servperf scale

Cronin and Taylor (1992) were amongst the researchers who levelled maximum attack on the SERVQUAL scale. They questioned the conceptual basis of the SERVQUAL scale and found it confusing with service satisfaction. They, therefore, opined that expectation (E) component of SERVQUAL would be discarded and instead only performance (P) component would be used. They proposed what is referred to as the ‘SERVPERF’ scale. Besides theoretical arguments, Cronin and Taylor (1992) provided empirical evidence across four industries (namely banks, pest control, dry cleaning, and fast food) to corroborate the superiority of their 'performance-only' instrument over disconfirmation-based SERVQUAL scale.

Being a variant of the SERVQUAL scale and containing perceived performance component alone, 'performance only' scale is comprised of only 22 items. A higher perceived performance implies higher service quality. In equation form, it can be expressed as:

$$SQ_i = \sum_{j=1}^{k} P_{ij},$$

(1.2)

where:

SQi = perceived service quality of individual 'i_{TH} .

k = number of service attributes/items.

P = perception of individual i_{TH} with respect to performance of a service firm attribute j_{TH} .

Methodologically, the SERVPERF scale represents marked improvement over the SERVQUAL scale. Not only it is the scale more efficient in reducing the number of items to be measured by 50 per cent but, it has also been empirically found superior to the SERVQUAL scale for being able to explain greater variance in the overall service quality measured through the use of single-item scale. This explains the considerable support that has emerged over time in favour of the SERVPERF scale (Babakus and Boller, 1992; Bolton and Drew, 1991b; Boulding et al., 1993; Churchill and Surprenant, 1982; Gotlieb, Grewal and Brown, 1994; Hartline and Ferrell, 1996; Mazis, Antola and Klippel, 1975; Woodruff, Cadotte and Jenkins, 1983). Though still lagging behind the SERVQUAL scale in application, researchers have increasingly started making use of the performance measure of service quality (Andaleeb and Basu, 1994; Babakus and Boller, 1992; Boulding et al., 1993; Brady et al., 2001; Cronin et al., 2000; Cronin and Taylor, 1992, 1994). Also when applied in conjunction with the SERVQUAL scale, the SERVPERF measure has outperformed the SERVQUAL scale (Babakus and Boller, 1992; Brady, Cronin and Brand, 2002; Cronin and Taylor, 1992; Dabholkar et al., 2000). Seeing its superiority, even Zeithaml in a recent study observed that "…Our results are incompatible with both the one-dimensional view of expectations and the gap formation for service quality. Instead, we find that perceived quality is directly influenced only by perceptions (of performance) (Boulding et al., 1993). This admittance cogently lends a testimony to the superiority of the SERVPERF scale.

1.1.5 Unweighted and weighted paradigms

The significance of various quality attributes used in the service quality scales can considerably differ across different types of services and service customers. Security, for instance, might be a prime determinant of quality for bank customers but may not mean much to customers of a beauty parlour.

Since service quality attributes are not expected to be equally important across service industries, it has been suggested to include importance weights in the service quality measurement scales (Cronin and Taylor, 1992; Parasuraman, Zeithaml and Berry, 1994a, 1994b, 1996; Parasuraman, Berry and Zeithaml, 1991a, 1991b, 1991c; Zeithaml, Parasuraman and Berry, 1990; Ozment et al, 1994). While the unweighted measures of the SERVQUAL and the SERVPERF scales have been described above vide equations (1.1) and (1.2), the weighted versions of the SERVQUAL and the SERVPERF scales as proposed by Cronin and Taylor (1992) are as follows:

$$SQ_i = \sum_{j=1}^{k} I_{ij}\left(P_{ij} - E_{ij}\right), \tag{1.3}$$

$$SQ_i = \sum_{j=1}^{k} I_{ij} P_{ij}, \tag{1.4}$$

where:

I_{ij} is the weighting factor, i.e., importance of attribute j to an individual i.

The addition of weights in the two scales was reported in several works (Bolton and Drew, 1991a). Between weighted versions of the two scales, weighted SERVPERF scale has been theoretically posited to be superior to weighted SERVQUAL scale (Bolton and Drew, 1991a). As pointed out earlier, one major problem with the past studies has been their preoccupation with assessment of psychometric and methodological soundness of the two scales. The diagnostic ability of the scales has not been explicitly explicated and empirically investigated. The psychometric and methodological aspects of a scale are no doubt important considerations but one cannot overlook the assessment of the diagnostic power of the scales. From the strategy formulation point of view, it is rather the diagnostic ability of the scale that can help managers in ascertaining where the quality shortfalls prevail and what possibly can be done to close down the gaps (Bou, 2000).

1.2 Notation and Problem Definition

Suppose that we have a population of customers. This population includes our current customers, and it could also include future or potential customers (former customers, to study “recovery customer”, require a specific adaptation) . We assume that the size of the customer's population, N , is large.

Let us call Q_i to the perceived quality of a given service by the i_{TH} customer from this population. The customer compares his expectations towards a certain service with its perceived performance (see Parasuraman et al., 1988, 1991, 1994a, 1994b; Zeithaml et al., 1990). The judgment of quality is built up on the basis of this theoretical construct. Good service quality evaluation develops when perceptions exceed or are equal to expectations. Consequently, most approaches try to measure this gap directly (Liljander and Strandvik, 1993).

On the other hand, the models explaining quality use the concept of importance (Kawlath 1969, Hüttenrauch 1994, Behrens, Schneider and Weisberg 1978). The customer determines all characteristics he expects the ideal service to receive. Because not all of them are equally important, he weighs the importance of each. He builds his quality judgment on his perception of each characteristic multiplied with its specific significance. Summing up all evaluated criteria gives the total quality score (ISO 26362: 2009).

It is common to assume that customer's evaluation will be a function of several attributes $X_1,\ldots;X_k$ which determine the global evaluation of the service. Let us call $X_{i1},\ldots;X_{ik}$ to the evaluations of these attributes made by the i_{TH} customer. Then,

$$Q_i = f(X_{i1},\ldots,X_{ik}), \tag{1.5}$$

A linear quality indicator (Behrens, Schneider and Weisberg 1978) assume that the function (1.5) can be approximated by

$$Q_i = \sum_{j=1}^{k} w_{ij} X_{ij} \ , \tag{1.6}$$

where the coefficients w_{ij} are weights, so that they must be positive and they must add up to one:

$$w_{ij} \geq 0 \qquad \forall i, \forall j$$

$$\sum_{j=1}^{k} w_{ij} = 1 \quad \forall i \tag{1.7}$$

These weights can be considered as measures of the relative importance of attribute X_j in determining the evaluation of the quality of the service for the i_{TH} customer.

In order to deploy this quality model we need:

- The complete list of attributes.
- The weights.

The most important part is to obtain the weights, because we can always write a long list of attributes but some of them may have weights equal to zero.

In this thesis, part of our effort is dedicated to present and developed a methodology to calculate the weights considering that different customers may have different weights for the attributes.

1.3 Methods to determine the weights. Classical Tools

Several methods of measuring service quality have been developed and discussed over the last few years. Reviewing the service quality literature, most of these models work with expectations (see Parasuraman et al., 1988, 1991, 1994a, 1994b; Zeithaml, 1988).

Expectations are already integrated in the evaluation of the perceptions. When a customer judges a certain characteristic to be good, he expresses that it exceeds either his predictive or his service expectations. However, the customer often has only a vague idea about the latter. For this reason, the measurement of expectation had been rejected. Instead, it is common to work with the perceptions and the importance of the attributes.

1.3.1 Expected Quality

We assume a linear quality indicator from the function (1.6) and we assume that the weights w_{ij} used by the i_{TH} customer for the j_{TH} attribute are independent of the evaluation made by this customer for this attribute X_{ij}. The justification of this assumption is that the evaluation of an attribute represents how the level of service in this attribute compares to an ideal or standard performance. For instance, suppose that the service is a restaurant and the attribute is the speed of the service measured by the time the customer has to wait to receive his order. Then, the evaluation of the waiting time depends on previous experiences of the customer on similar situations and will normally depend on the type of restaurant. We assume that the evaluation of this attribute in a particular restaurant is independent from the importance that the speed in the service has in his judgment of the quality of the service.

We define the service quality as the expected value of Q_i in the customer's population

$$Q = E[Q_i] = \frac{\sum_{i=1}^{N} Q_i}{N} \quad . \tag{1.8}$$

The service quality can readily be obtained from equations (1.6) by using the independence of the variables w_{ij} and X_{ij}. Then this global measure of service quality will be given by:

$$Q = \sum_{j=1}^{k} E[w_{ij}] \;\; E[X_{ij}] = \sum_{j=1}^{k} w_j m_j , \tag{1.9}$$

where we have called m_j to the average evaluation of attribute j_{TH} in the population, and w_j is the mean of the distribution of the weight of this attribute in the population.

The estimation of service quality can be obtained from (1.8), by taking the average of the evaluation provided by a sample of customers, or by (1.9), by estimating the mean weights of each attribute in the population and the average of the evaluations for the attributes. Although both procedures must lead to the same final number the quality index model (1.9) provides a decomposition of the sources of the service quality with the following advantages:

(1) A quality index allows comparing the average value of the attributes of our service to the values of other companies and can reveal our relative strengths and weaknesses in a SWOT analysis.

(2) Knowing the attribute weights allows the ordering of the attributes according to their relative importance to the customer, showing the key factors in order to improve quality.

(3) Customers can be segmented by their weighting function, obtaining a market segmentation function directly linked to our quality objectives.

(4) If the attributes can be related to some objective measures of performance it is possible to substitute the subjective evaluations of the attributes by objective measurements, allowing a simple monitoring of the quality index.

The objective of many quality evaluations is to build a quality index (scalar measure) to summarize the performance of the service. Reduction of all the dimensions of an evaluation to a single number can be subject to many criticisms. However, the presence of one quality index is required for decision making.

It must be stressed that the key information is the distribution of the perceived quality in the population and the mean given by (1.3) and (1.4) is a first summary measure of this distribution. It would be useful to compute also a measure of the variability of the perceived service quality in the population and this can be done by

$$\sigma^2 = \sum_{i=1}^{n} \frac{\left(Q - \overline{Q}\right)^2}{n-1}, \tag{1.10}$$

where n is the sample size and using the average of the deviations from the mean as usual.

It would be useful to decompose this variability by its sources as it has been done with the mean. However, this is a complicated task because in order to do so we need to know the variances of each attribute evaluation in the population, the variance of the weights and also the covariances between weights and between attributes.

The operational definition of service quality presented has some limitations. First, we may have a good service quality on average, but a very bad service quality for some groups of customers. This may happen either in two ways:

- because some segments of the customers have a very different weighting function for the quality attributes, we call this situation “implicated population”
- because they have a different evaluation of the attributes, we call this situation “explicated population”.

These two situations should be identified because we can provide a better service if we identify clusters of customers with different values or opinions about quality. Then, it is more informative to measure service quality in these different populations.

It must be remembered that the mean is only a good descriptive measure when we have an homogeneous sample and that it can be non representative when the data comes from a mixture of very different populations.

1.3.2 Direct evaluation of weights

Several authors have recommended estimating the weights in a linear indicator of quality like (1.9), for consensus, asking directly to the customers. For instance, Zeithaml et al.(1990) in their model of service quality SERVQUAL identify five attributes of service quality:

Attribute
Tangibles
Reliability
Responsiveness
Assurance
Empathy

Table 1.2. Servqual service quality attributes

And then, to determine the weights of these attributes by asking to a sample of customer or small group of experts (see ASCI 1994). It is common to find different methods to determine the weights directly asking the customers. In the following table, the three most used methods are presented:

Method	**Description**
(M1)	To distribute 100 points among these attributes.
(M2)	To indicate the relative importance of each attribute in a 0 - 10 scale.
(M3)	To indicate which of the attributes is considered the most important.

Table 1.3. Direct weights evaluation methods

The first method, M1, provides directly the weights. It has the problem that many customers do not have the habit of making these kinds of assessment and the results may be very unreliable (Dutka, 1994). Also this procedure is very difficult to apply when the number of attributes is large. Even for small number of attributes customers usually indicate that they are insecure about the weights. It is not clear how to save bad assignment values and it is difficult to find a Poka Yoke alternative, fool proof methodology of anticipation (Mahapatra et al, 2006).

The second method, M2, can be used to find weights by dividing the importance of the attributes by the sum of the importance for all attributes. It can be shown (Peña, 1997a, 1999a, 1999b) that if the number of attributes is k, the weights obtained by this method are very similar of size $1/k$, leading to an almost uniform weighting for all the attributes. Then we face the following problematic situation "when everything is important, then nothing is important".

The third method, M3, can be used to obtain weights by taking as weights the proportion of customers who consider each attribute as the most important. This method leads to a very asymmetric distribution of weights. For instance, consider the case in which all the members of the population agree on that attribute number 1 is the most important, number 2 is also important and all the others are not. Then we will get a weight of 100 for attribute number 1 and zero for the rest that is clearly unsatisfactory. A modification of this method will be to assign a rank order to the attributes, obtain the mean of these orders and try to use this mean rank to build weights. The problem with this procedure is again that it does not take into account that a rank scale will not define well in general an interval or continuous scale for the weights. The difference in importance, and therefore in weights, between the 1st and the 2nd attribute will not be in general the same than the one between attributes 3rd and 4th and so on.

Example:

We can present one example to illustrate the methods:
We chose 50 customers and applied a Focus Group technique. We asked them about the importance of each SERVQUAL quality attribute.

Each customer answered in three ways:

- M1: To distribute 100 points among these attributes.
- M2: To indicate the relative importance of each attribute in a 0 - 10 scale.
- M3: To indicate which of the attributes is considered the most important.

In the following table, results are presented:

Attribute	Importance	M1	M2	M3
Tangibles	8.25	18.25	38.23	26
Reliability	9.12	20.17	2.16	11
Responsiveness	9.24	20.44	15.21	6
Assurance	9.18	20.31	2.33	14
Empathy	9.42	20.84	42.17	43

Table 1.4. Computation of weights by different methods for the Servqual model

The first column of the table presents the importance of the attribute in a 0-10 scale given directly by customers in the sample.

The second column includes the weights derived from a procedure that we have called M1. This procedure uses the result of dividing the importance of the attribute in the first column by the sum of the importance of the five attributes. For instance, 18.25 = 8.25 / (8.25+9.12+9.24+9.18+9.42).

The third column corresponds to a method we have called M2, in which the weights are taken equal to the percentage of answers that indicated that the attribute is the most important.

The fourth column corresponds to a method we have called M3, in which the customers estimate directly the weights.

The main conclusions from the table are:

- The method M1 leads to a similar and almost uniform weighting for all the attributes. This uniformity of the values increases with the dimension.
- The method M2 leads to a very asymmetric distribution of weights. This situation is expected.
- The method M3 leads to values are approximately half way between M1 and M2.

1.3.3 Indirect evaluation of the weights

For indirect evaluation of the attributes and the quality from a sample of experts from some population of experts, members of a representative sample. The weights are obtained by statistical analysis.

There are two ways to do so:

a) Fix the values of the attributes and ask for a global evaluation (value of Q). Then fit a linear model and determine the weights. This is conjoint analysis; and then we can use fractional factorials to build a model and estimate the weights.

b) To evaluate both the attributes and the global performance (or global quality) and then use several linear regression methods to build a model and estimate the weights.

Conjoint Analysis

Methods oriented to multidimensional quality measurements are usually based on Conjoint Analysis (Luce and Tukey, 1965). See Carroll and Green (1995) for a survey of the state of this methodology and Lynch et al

(1994), Wedel and DeSarbo (1994) and Ostrom and Iacobucci (1995) for interesting applications to the evaluation of service quality.

When we asked directly to consumers which attributes are the most important ones, the response may be that “they all are important” (Gustafsson, 2007).
Furthermore, individual attributes in isolation are perceived differently than in the combinations found in a product (Bagozzi and Fornell 1982).

It is difficult for a survey respondent to take a list of attributes and mentally construct the preferred combinations of them. The task is easier if the respondent is presented with the combinations of attributes that can be visualized as different product offerings. However, such a survey becomes impractical when there are several attributes that result in a very large number of possible combinations.

Conjoint analysis can facilitate this task. Conjoint analysis is a tool that allows a subset of the possible combinations of product/service features to be used to determine the relative importance of each feature in the purchasing decision/opinion. Conjoint analysis is based on the fact that the relative values of attributes considered jointly can better be measured than when considered in isolation.

In a conjoint analysis, the respondent may be asked to arrange a list of combinations of product attributes in decreasing order of preference. Once this ranking is obtained, a data analysis is used to find the utilities of different values of each attribute that would result in the respondent's order of preference. This method is efficient in the sense that the survey does not need to be conducted using every possible combination of attributes. The utilities can be determined using a subset of possible attribute combinations. From these results one can predict the desirability of the combinations that were not tested.

Developing a conjoint analysis involves the following steps:

1. Choose product/service attributes. For example, appearance, size, or price.
2. Choose the values or options for each attribute. For example, for the attribute of size, one may choose the levels of 5", 10", or 20". The higher the number of options used for each attribute, the more burden that is placed on the respondents.
3. Define products as a combination of attribute options. The set of combinations of attributes that will be used will be a subset of the possible universe of products.
4. Choose the form in which the combinations of attributes are to be presented to the respondents. Options include verbal presentation, paragraph description, and pictorial presentation.
5. Decide how responses will be aggregated. There are three choices - use individual responses, pool all responses into a single utility function, or define segments of respondents who have similar preferences.
6. Select the technique to be used to analyze the collected data. The part-worth model is one of the simpler models used to express the utilities of the various attributes. There are also vector (linear) models and ideal-point (quadratic) models.

Conjoint analysis has become an important marketing research tool. It is well-suited for defining a new product or improving an existing one. In summary, in these procedures the customers are given several hypothetical services defined by certain levels of the quality attributes and are asked to provide quality evaluation or preferences for these services. The method assumes that the quality attributes can be given an objective interpretation and that the levels of the attributes, when presented to the customers for evaluation, have a clear meaning to them.

However, these methods are less useful in situations in which the quality attributes do not have objective standards and therefore it is not clear how to create a series of hypothetical quality situations for the customers to evaluate which have the same meaning for all customers.

In most application of quality service measurement we do not have objective weights. Note that in other indexes of standard statistical and economic use the weights are objective, as for instance, Cost of living indexes (the importance of a product depends on its contribution to the total cost of a familiar unit)

Generalized Least Squared Method

We assume that a random sample of size n from the customers' population has provided evaluations for the global service quality, y_i, $(i = 1,\ldots,n)$ as well as evaluations of the attributes that determine the quality of the service, x_{ij}, for certain well defined attributes X_j, $(i = 1,\ldots,n; j = 1,\ldots,k)$. From now on and without loss of generality we assume that these evaluations are done in a 1 - 10 scale.

The following hypotheses are made:

(H1) The quality of the service for the i_{TH} customer is an unknown continuous variable that is measured in a discrete way by the evaluation y_i provided by him. This implies some rounding error. In addition to this error we assume that the evaluation includes an additional measurement error due to the fact that the customer when ask in different moment or situation may give slightly different answers. This two factors means that the evaluation y_i is related to the service quality Q_i by

$$y_i = Q_i + u_i \ . \tag{1.11}$$

Note that in this assumption the variable u_i includes all the factors, which determine that the same customer asked about the quality of the service may give different evaluation in different moments of time. It also includes the error due to the scale of measurement. This variable will change from customer to customer, but assuming that it includes the effect of many independent factors we may suppose, by the central limit theorem, that this variable follows a normal distribution in the population. We also assume that this random error has a zero mean, that is, there are not systematic biases on the evaluation and that the variability is roughly the same for all customers in the population. Then u_i follows a normal $N(0,\sigma_u^2)$ distribution in the population of customers

We believe that H1 is quite general and it can be considered to hold in most situations.

(H2) Customers made evaluations $x_i = (x_{i1},\ldots,x_{ik})$ without error.

This assumption will be approximately true when the errors in evaluating the attributes are small compared to the error in evaluating the service quality. In practice there will always be some measurement error on the evaluation of the attributes that, besides, can be different for different attributes. However, we assume this hypothesis as a first approximation and for simplicity. Dropping it increases very much the technical complication of the model because then it is transformed into an error in variables model (see Fuller, 1987).

(H3) The service quality Q_i is a linear function of the attributes X_i, as

$$Q_i = w_i^{'} x_i \ , \tag{1.12}$$

where $w_i = (w_{i1},\ldots,w_{ik})^{'}$ are the weights. These weights are positive coefficients that must add up to one, as explained previously. We assume that the variables weights, w_{ij}, and attribute evaluations, x_{ij}, are independent.

In case the vector of attributes does not include all the relevant attributes, equation (1.6) is modified as

$$Q_i = w_0 + w_i^{'} x_i \quad ,$$

(1.13)

Where now w_0 includes the effect of all the attributes not included in the evaluation.

The assumption of linearity is a strong one, but it can be tested after estimating the model. In some cases some nonlinear effects and interaction between attributes are expected. For instance, a bad performance in two important attributes can lead to a lower perceived quality that the one implied by adding up the effects of each attribute. Then, we say that there is interaction between these two attributes and this feature should be included in the model as a product term. Again, this hypothesis can be tested when the model is estimated.

(H4) The weights w_{ij} are random variables in the customer's population and follow a normal distribution with expected value w_j and variance σ_u^2, that is the same for the k attributes.

This hypothesis is rather restrictive because the variability of the weights in the population will be, in many cases, different for some attributes. It can be eliminated, but again the complexity of the model increases.

With these four assumptions the distribution of the random variables $y_i(i = 1, \ldots, n)$ given the x will be normal with mean

$$E[y_i] = E[Q_i] = w' x \quad ,$$

(1.14)

Note that this equation is equivalent to (1.6). The variability of the observation y_i is not constant, as in the standard regression model, and it can be decomposed into two components:

The first one is the variability due to the measurement error, which has variance σ_u^2. This source of variability can not be avoided, because it depends on the precision used by the customers to indicate their evaluations.

The second source of variability is due to the variability of the different weights used by different customers, as measured by σ_w^2, and to the variability on the evaluation of the attributes. In fact, it can be shown (Peña, 1997) that the variance of y_i is given by

$$Var[y_i] = \sigma_i^2 = \sigma_u^2 \left(\theta k s_i^2 + 1\right) , \tag{1.15}$$

where $\left(\theta k s_i^2 + 1\right)$ is an inflation factor that takes into account the inequality of the variances of the observations.

This inflation factor depends on three terms:

The first is $\theta = {\sigma_w^2}/{\sigma_u^2}$, the ratio between the common variability of the weights in the population and the measurement error.

The second is k, the number of attributes considered.

The third is s_i^2, the variance of the attribute evaluations made by the customer.

Note that when σ_w^2 is small compared to σ_u^2, so that θ is very small, the variability of each observation is approximately constant.

In summary, the variability increases with the number of attributes, k, the variability of the weights, and the variability of the evaluations made by the customer, and therefore we have a heterokedastic regression model subject to linear restriction over the parameters w.

This model can be estimated, including the linear constrain that the weights must add up to one. However, as there is always the possibility that an important attribute has been overlooked it's common to propose fitting the model (1.12) with a constant term and checking if the intercept is different from zero. If it is not, then the model can be re-estimated imposing the restriction that the weights add up to one. If the intercept is statistically different from zero, this fact can be taken as an indication that an attribute is missing in the evaluations.

In order to estimate the model by generalized least squares the inflation factors of the variances of the observations need to be known. Note that θ is the ratio of the variances and it depends on the units of measurement. As the weights must add up to one, the standard deviation of a weight must be around 0.1 and can not be larger than 1. This implies that the variance will be around 0.01. The variability of the evaluations in the 0-10 scale will be a minimum of 0.5 and can be as large as 1. This means that the ratio of the variances will be smaller than one and, as a first approximation; we may assume that it is included in the interval (0.05 to 0.001). We propose to fix the value θ to 0.01 and carry out a sensitivity analysis to check if the results depend on the θ value assumed. In all cases we have found in practice that the result are quite robust to the particular value chosen in the interval.

Assuming that θ is known, the variability due to the measurement error σ_u^2 can be estimated, by

$$\hat{\sigma}_u^2 = \frac{1}{n}\sum\frac{(y_i - \hat{w}' x_i)^2}{\theta k s_i^2 + 1}.1 \ , \tag{1.16}$$

Note that we can use most of the standard regression methods to check the validity of this random coefficients model. In particular the restriction of the weights adding up to one can be tested by comparing the constrained and the unconstrained estimates. Estimating a model without these restrictions can also test the key assumption of equal variances in the distribution of the weights.

Then, for each person we have the explanatory variables, X, the response of global performance, Y, and the regression coefficients will be the weights. But the weights will be different for different judges or referees and we want to estimate the distribution of weights in the population and the average weights to measure the quality service.

A point of special interest is determining groups of customers with different weighting structure. For instance, sometimes the distribution of customer's weights can be thought of as a mixture of two or more distributions corresponding to two or more different type of customers. This should be taken into account to avoid serious misspecification errors in the model. For instance, a small set of customers with evaluations very different from the others may determine completely the weighting function if they have more extreme (either good or bad) evaluations that the bulk of the other customers. This problem has already been researched (See Peña and Yohai 1995, Wedel and DeSarbo 1994) in other occasions.

1.4 Dissertation Overview

The rest of the dissertation is organized as follows:

Chapter 2 presents a methodology in which the weights are estimated from the observed relationship between the customer's evaluations of overall quality and the evaluations of the attributes by a nonparametric procedure. Also, summarizes computer results and discusses the problems with previous approaches to estimate the weights in the quality service measurement.

Chapter 3 describes the application, in a real case, of the methodology presented for measuring the quality of CABINTEC ("Intelligent cabin truck for road transport").

Chapter 4 describes the application, in a real case, of the methodology presented for measuring the quality of postgraduate education in a Spanish public university.

Finally, chapter 5 presents our conclusions and discusses avenues to future research.

1.5 Conclusions of Chapter 1

Several methods of measuring service quality have been developed and discussed over the last few years. The objective of many quality evaluations is to build a quality index to summarize the performance of the service. Reduction of all the dimensions of an evaluation to a single number can be subject to many criticisms.

It is common to assume that customer's evaluation will be a function of several attributes $X_1,\ldots;X_k$ which determine the global evaluation of the service. Let us call $X_{i1},\ldots;X_{ik}$ to the evaluations of these attributes made by the i_{TH} customer. Then,

$$Q_i = f(X_{i1},\ldots,X_{ik}),$$

A linear quality indicator (Behrens, Schneider and Weisberg 1978) assume that the function (1.5) can be approximated by

$$Q_i = \sum_{j=1}^{k} w_{ij} X_{ij} \ ,$$

where the coefficients w_{ij} are weights, so that they must be positive and they must add up to one:

$$w_{ij} \geq 0 \qquad \forall i, \forall j$$

$$\sum_{j=1}^{k} w_{ij} = 1 \quad \forall i$$

These weights can be considered as measures of the relative importance of attribute X_j in determining the evaluation of the quality of the service for the i_{TH} customer.

In order to deploy this quality model we need:

- The complete list of attributes.
- The weights.

The most important part is to obtain the weights, because we can always write a long list of attributes but some of them may have weights equal to zero.

It must be stressed that the key information is the distribution of the perceived quality in the population and the mean is a first summary measure of this distribution, but this definition of service quality index has some limitations. First, we may have a good service quality on average, but a very bad service quality for some groups of customers.

This may happen either in two ways:

- because some segments of the customers have a very different weighting function for the quality attributes, we call this situation "implicated population"
- because they have a different evaluation of the attributes, we call this situation "explicated population".

These two situations should be identified because we can provide a better service if we identify clusters of customers with different values or opinions about quality. Then, it is more informative to measure service quality in these different populations.

It must be remembered that the mean is only a good descriptive measure when we have an homogeneous sample and that it can be non representative when the data comes from a mixture of very different populations.

For indirect evaluation of the quality from a sample of experts from some population of experts, members of a representative sample. The weights are obtained by statistical analysis.

There are two ways to do so:

a) Fix the values of the attributes and ask for a global evaluation (value of Q). Then fit a linear model and determine the weights. This is conjoint analysis; and then we can use fractional factorials to build a model and estimate the weights.

b) To evaluate both the attributes and the global performance (or global quality) and then use several linear regression methods to build a model and estimate the weights.

Classical tools are less useful in situations in which the quality attributes do not have objective standards and therefore it is not clear how to create a series of hypothetical quality situations for the customers to evaluate which have the same meaning for all customers.

For each person we have the explanatory variables, X, the response of global performance. We can calculate the weights with classical tools, but the weights will be different for different judges or referees and we want to estimate the distribution of weights in the population and the average weights to measure the quality service.

A point of special interest is determining groups of customers with different weighting structure. This should be taken into account to avoid serious misspecification errors in the model. For instance, a small set of customers with evaluations very different from the others may determine completely the weighting function if they have more extreme (either good or bad) evaluations that the bulk of the other customers.

In this thesis, part of our effort is dedicated to present and develop a methodology to measure the quality of services, calculating the weights considering that different customers may have different weights for the attributes.

Chapter 2

Adaptive Local Regression (ALR)

We have developed a methodology to measure the quality of services. In this chapter, we present this methodology.

We propose a non parametric quality model in which the individual weights for each customer can be estimated. See Gumpertz and Pantula (1998) for a review of these models and their applications and Mallet (1986) for a non parametric approach to estimate the distribution of the coefficients.

2.1 Basic Concepts

Suppose that we have a population of customers. This population includes our present customers, and it could also include future or potential customers. We assume that the size of the customer's population, n, is large.

Let us call Q the vector composed by Q_i, that is, the perceived quality of a given service by the i_{TH} customer from this population.

It's common to assume that customer's evaluation will be a function of several, k, attributes $X_1,\ldots;X_k$ which determine the global evaluation of the service. Let us call X_i the vector composed by $X_{i1},\ldots;X_{ik}$, that is, the evaluations of the quality attributes made by the i_{TH} customer.

Then,

$$Q_i = f\left(X_{i1},\ldots,X_{ik}\right) . \tag{2.1}$$

A linear quality indicator can be approximated by

$$Q_i = \sum_{j=1}^{k} w_{ij} X_{ij} \quad , \tag{2.2}$$

where the coefficients w_{ij}, in the matrix W, are weights defined for each customer, so that they must be positive and they must add up to one:

$$\begin{aligned} & w_{ij} \geq 0 \qquad \forall i, \forall j, \\ & \sum_{j=1}^{k} w_{ij} = 1 \quad \forall i. \end{aligned} \tag{2.3}$$

These weights can be considered as measures of the relative importance of attribute X_j in determining the evaluation of the quality of the service for the i_{TH} customer (Kunst and Lemmink 1996 and 1997).

Definition: ε-reasonable neighbours

Given an element X we will say that y is an ε-reasonable neighbour if $y \in B(X,\varepsilon)$, where ε denotes the size of the neighbourhood (Silverman 1986, Muñoz and Moguerza 2006).

Remark:

Notice that we may build ε-reasonable neighbours not in the sample. For instance, given an element X, if $|\xi| < \varepsilon$, the element $X + \xi\varepsilon$ is an ε-reasonable neighbours .

2.2 Estimation of GLSM with linear constraints

In this section, we deploy the generalized least squared method (GLSM) for the estimation of the parameters $\widehat{w}$ (see Peña 1997a, 1999b).

We require the covariance matrix of the vector of variables w_i. This vector follows a multivariate normal distribution with expected value $w = (w_i, \ldots, w_k)$ and covariance matrix

$$\sum_w = \sigma_w^2 \begin{bmatrix} 1 & -\frac{1}{k-1} & \cdots & -\frac{1}{k-1} \\ -\frac{1}{k-1} & 1 & \cdots & \\ \vdots & \vdots & & \vdots \\ -\frac{1}{k-1} & \cdots & \cdots & 1 \end{bmatrix} = \sigma_w^2 \quad A \quad , \tag{2.4}$$

where the A matrix is given by

$$\frac{1}{k-1}(k \quad \mathrm{I} - 1 \quad 1') \quad . \tag{2.5}$$

This covariance matrix can be obtained from the assumption that all marginal variables have the same variance, the same covariances and they add up to one (Afifi, 2004).

As:

$$(w_{i1} - w_1) + \ldots + (w_{ik} - w_k) = 0 \quad . \tag{2.6}$$

Taking the square of this expression and then expected values

$$k\sigma_w^2 + 2\binom{k}{2}\gamma_w = 0 , \tag{2.7}$$

where γ_w are the covariances between the attributes that are assumed to be equal. Then

$$k\sigma_w^2 + 2\frac{k(k-1)}{2}\gamma_w = 0 , \tag{2.8}$$

which implies that the covariances are equal to $-\sigma_w^2/(k-1)$, and the covariance matrix is obtained. Then, the distribution of the random variables $y_i(i = 1\ldots, n)$ given the X variables is normal with expected value $E(y_i) = w' x_i$ and the variance

$$\sigma_i^2 = Var(y_i) = Var(Q_i + \mu_i) = x_i' \ \sum\nolimits_w \ x_i + \sigma_u^2 , \tag{2.9}$$

that can be written, calling $\theta = \sigma_w^2 / \sigma_u^2$, as

$$\sigma_i^2 = \sigma_u^2 (\theta r_i + 1), \tag{2.10}$$

where $r_i = x_i' A x_i$.

The estimation requires that the likelihood must be maximized with the restrictions

$$\sigma_w^2 \geq 0, \quad \sigma_w^2 \geq 0, \quad w \geq 0, \quad w'1 = 1 \ . \tag{2.11}$$

We can write the equation including the Lagrange multipliers for these restrictions and maximize it.

The solution of this equation requires a nonlinear optimization algorithm. Note that the structure of this system is simple because we can fix θ and determine σ_u^2 and $\hat{w}$. Then we compute a new value for θ, which will lead to new estimates for σ_u^2 and $\hat{w}$, and so on.

Then, the vector of mean weights in the population is given by:

$$\hat{w} = \hat{w}_0 - (X'DX)^{-1} \ 1 \ (1'\hat{w}_0 - 1) \ , \tag{2.12}$$

where 1 is a k-dimensional vector of ones, the constant is a scaling factor equal to $\left(1'(X'DX)^{-1}1\right)^{-1}$, so that constraint is verified. Note that if we multiply (2.12) by $1'$ we obtain $1'\hat{w} = 1$.

The value $\hat{w}_0 = (X'DX)^{-1} X'DY$ is the generalized least squares estimate without constraints.

D is the variance-covariance matrix of the observations y_i that is calculated in this section.

The equation (2.12) shows that if $\widehat{w}_0 1' = 1$ no constraints are applied. Otherwise, the estimator is corrected to fulfil this restriction.

Note that this estimator is different from the trivial one, $\widehat{w}_0 / 1'\widehat{w}_0$ that is obtained by dividing each component of $\widehat{w}_0$ by the sum of all the components in order to fulfil the restriction.

In the equation (2.12) each component of $\widehat{w}_0$ is corrected by an amount that depends on its variance and its covariance with the other components, as measured by the matrix $(X'DX)^{-1}$.

Using this estimator, and assuming that θ is known, σ_u^2, can be estimated by

$$\sigma_u^2 = \frac{1}{n}\sum \frac{(y_i - \widehat{w}' x_i)^2}{\theta k s_i^2 + 1}$$

(2.13))

Usually θ is known and presents a method to estimate it. However, in some applications we have a priori a set of possible values for this parameters. Then, the simplest way to deal with it is to compute the residual variance with different values of this parameter and take as estimate the value minimize it. The advantage of this procedure is that it does not require a special software and it can be carried out with any standard statistical package that includes weighted least squared.

Note that we can use most standard regression methods to check the validity of this random coefficient model. In particular the restriction of the weights adding one can be tested by comparing the constrained and the unconstrained estimates. The key assumption of equal variances in the distribution of the weights can also be tested by estimating a model without these restrictions.

2.3 Constrained Least Squares

We deploy the algorithm to solve the least squared problem with quadratic constrains, in particular the LS problem over a sphere (Schott, 1992).

In the Least Squares setting it is natural to minimize $\|Xw - Q\|_2$ over a proper subset of $\Re^n$. For example, we may wish to predict Q as best we can with Xw subject to the constraint that w is a unit vector. Or, perhaps the solution defines a fitting function which is to have prescribed values and a finite number of points. This can lead to an equality constrained least squares problem. These problems can be solved using the QR factorization and the SVD.

The Least Squares minimization with Quadratic Inequality constraints (LSQI) is defined by

$$\begin{aligned} &\min \|Xw - Q\|_2 \\ &s.t. \|Bw\|_2 \le \alpha \end{aligned} \tag{2.14}$$

where $X \in \Re^{mxn}$ ($m > n$), $Q \in \Re^m$, $B \in \Re^{nxn}$ (non singular), and $\alpha \ge 0$. The constraint defines a hyperellipsoid in $\Re^n$. More generally, we have the problem:

$$\begin{aligned} &\min \|Xw - Q\|_2 \\ &s.t. \|Bw - d\|_2 \le \alpha \end{aligned}, \tag{2.15}$$

where $X \in \Re^{mxn}$ ($m > n$), $Q \in \Re^m$, $B \in \Re^{pxn}$ (non singular), $d \in \Re^p$ and $\alpha \ge 0$.

The generalized singular value decomposition sheds light on the solvelibility of (2.15). Indeed if,

$$U^T XA = diag(c_1,...,c_n), \qquad U^T U = I_m, \qquad q = \min\{p, n\},$$
$$VBA^T = diag(d_1,...,d_q), \qquad VV^T = I_p,$$

(2.16)

where $c_i \geq 0$, $d_i \geq 0$, is the generalized singular value decomposition of X and B, then (2.15) transforms to

$$\min\left\|D_W y - \widetilde{Q}\right\|_2$$
$$s.t.\left\|D_B y - \widetilde{d}\right\|_2 \leq \alpha,$$

(2.17)

where $\widetilde{Q} = U^T Q$, $\widetilde{d} = V^T d$ and $y = W^{-1} w$.

To solve this problem, we use the method of Lagrange multipliers. Defining

$$h(\lambda, y) = \left\|D_X y - \widetilde{Q}\right\|_2^2 + \lambda\left(\left\|D_B - \widetilde{d}\right\|_2^2 - \alpha^2\right),,$$

we see that the equations $\dfrac{\partial h}{\partial y_i} = 0$, $i = 1,...,n$, lead to the linear system

$$(D_X^T D_X + \lambda D_B^T D_B) y = D_B^T \widetilde{Q} + \lambda D_B^T \widetilde{d} \ .$$

(2.18)

Assuming that the matrix of coefficient is non-singular, this has a solution $y(\lambda)$ where

$$y_i(\lambda) = \begin{cases} \dfrac{\alpha_i \widetilde{b}_i + \lambda \beta_i \widetilde{d}_i}{\alpha_i^2 + \lambda \beta_i^2} & i = 1,\ldots,q \\ \dfrac{\widetilde{b}_i}{\alpha_i} & i = q+1,\ldots,n. \end{cases} \tag{2.19}$$

To determine the Lagrange parameter we define

$$\Phi(\lambda) \equiv \|D_B y(\lambda)\|_2^2 = \sum_{i=1}^{r} \alpha_i \frac{\beta_i \widetilde{b}_i + \alpha_i \widetilde{d}_i^{\,2}}{\alpha_i^2 + \lambda \beta_i^2} + \sum_{i=r+1}^{p} \widetilde{d}_i^{\,2} , \tag{2.20}$$

and then the solution to $\Phi(\lambda) = \alpha^2$. Now $\Phi(\lambda)$ is monotone decreasing for $\lambda > 0$, and then it implies the existence of a unique positive λ^* for which $\Phi(\lambda^*) = \alpha^2$. It is easy to show that this is the desired root. It can be found through the application of any standard root-finding technique, such as Newton's method. The solution of the original LSQI problem is then $w = Xy(\lambda^*)$

For this important case over a sphere with $B = I_n$, $d = 0$, we have the following procedure:

Algorithm. Given $X \in \Re^{mxn}$ with $m > n$, $Q \in \Re^m$ $\alpha \geq 0$, the following algorithm computes a vector $w \in \Re^n$ such that find the minimum of $\|Xw - Q\|_2$, subject to the constraint that $\|w\|_2 \leq \alpha$.

Compute the SVD $X = U\Sigma V^T$

$Q = U^T Q$

$r = rank(X)$

if $\sum_{i=1}^{r} (\frac{Q_i}{\alpha_i})^2 > \alpha^2$

Find λ^* such that $\sum_{i=1}^{r} (\frac{\sigma_i Q_i}{\sigma_i^2 + \lambda^*})^2 = \alpha^2$

$$w = \sum_{i=1}^{r} (\frac{\sigma_i Q_i}{\sigma_i^2 + \lambda^*})\ v_i$$

else

$$w = \sum_{i=1}^{r} (\frac{Q_i}{\sigma_i})\ v_i$$

end

Figure 2.1. Constrained Least Squares Algorithm

2.4 Proposed Methodology

In order to deploy this quality model we need:

- The complete list of attributes.
- The weights.

The most important part is to obtain the weights, because we can always write a long list of attributes but some of them may have weights equal to zero.

Initial hypothesis:

HH1: There is exists a function $f \mid f(X) = Q$.

HH2: We relax HH1 assuming that f is a piecewise linear function. Non linear functions could be one further research of this thesis.

Then, our model is locally $w^T X \approx Q$, because $w^T X$ is a linear approximation of f.

With the following methodology we estimate each component of the matrix W with $\widehat{W}$:

Algorithm to estimate $\widehat{W}_{ij}$

For each customer $i = 1,\dots,n$

Step 1.

Calculate its l ε-reasonable neighbours, that is $X_{(1)},\dots,X_{(l)}$.

where $(1),\dots,(l)$ is a appropriate rearrangement of the k indexes in the set $\{1,\dots,n\}$ and $l \geq k$.

Step 2.

Build X^i and Q^i:

$$X^i = \begin{bmatrix} X_{i1} & \dots & \dots & X_{ik} \\ X_{(1)1} & \dots & \dots & X_{(1)k} \\ \dots & \dots & \dots & \dots \\ X_{(l)1} & \dots & \dots & X_{(l)k} \end{bmatrix}, \quad Q^i = \begin{bmatrix} Q_i \\ Q_{(1)} \\ \dots \\ Q_{(l)} \end{bmatrix}$$

Step 3.

Solve possible numerical redundancies in the matrix $\left[X^i \,|\, Q^i\right]$.

Step 4.

Estimate the vector W_i as $\widehat{W}_i = \left[\widehat{w}_{i1}, \widehat{w}_{i2}, \dots, \widehat{w}_{ik}\right]$, solving the systems $X^i \widehat{w}_i = Q^i$ using a least squares method with linear constraints.

Figure 2.2. Algorithm to estimate $\hat{w}_{ij}$

The proposed methodology presents several advantages respect to the "classical tools":

A1:

We can use parallel computation to solve the linear systems in the step 4 (Kepner 2009).

A2:

When the decision maker needs a single index, a scalar measure to summarize the performance, we can define:

$$\overline{w_j} = \frac{\sum_{j=1}^{n} \widehat{w}_{ij}}{n} \quad . \tag{2.21}$$

That reduction of all dimensions of an evaluation to a single number can be subject to many criticism, however, it is required for decision making.
It is an alternative use of the estimation of the weights and it provides an equivalent result with the "classical methods" when we are working with one group of customers.

A3:

We have estimated each component $\widehat{w}_{ij}$. Now, we can use any kind of multivariate method to determine new groups of customers, such us a posterior customer segmentation, and then to prepare inferences about it.

A4:

We can work, then, directly with weights that each customer assigns to each quality attribute. In fact, we don't accept the mean of the weights as a good representative estimator. It must be remembered that the mean is only a good descriptive measure when we have a homogeneous sample and that it can be very non representative when the data comes from a mixture of very different populations of customers.

A5:

We estimate the weights that each customer assigns to each quality attribute with the information obtained from its similar customers. We choose the set of "similarities" based on the nearest neighbourhood estimate.

Remarks: We define a vector of weights for each customer, therefore we are implicitly defining the importance given by the customer to each quality attribute. Notice that working with these weights as data we may define new relations among the data.

2.4.1 Model validation

The model can be validated by comparing the observed residuals after fitting the model to the residuals computed with artificial samples generated by using the estimated parameters (Peña and Yohai 2006).

If the model is appropriate the observed residuals should have a similar distribution as the simulated residuals. We compute the residuals and its empirical distribution function F_n.

We generate V artificial samples for Q^*, where each variable used is replaced by the estimated parameters. Then the Q-Q plot between F_n and F_n^* will be a diagnostic tool for detecting discrepancies between the model and the data.

In particular the Q–Q plot may detect outliers corresponding to respondents with atypical views or recording errors. However some groups of outliers may go undetected because of a masking effect, although this effect is not expected to be large, because the data must be between zero and one. In any case, it is safer to check that a robust estimate for regression is similar to LS.

We can study the influence of the discrete scale, in fact in most applications these variables are measured on a discrete scale. Suppose that the observed evaluation is made on a discrete rating scale, for example 0–10. Then the response variable in the model is not observed exactly but rounded off to the closest integer. In order to check the effect of this discrete scale we can simulate the model and, using the estimated values of the parameters, generate two types of samples (1) samples of continuous values, and (2) samples of discrete values obtained by rounding off the continuous values. Then we can estimate the model in both samples and compare the results. The average discrepancy found in many replications of this analysis will provide an estimate of the expected bias due to the discrete scale effect.

2.4.2 Diagnostic check

Predicting the weights for each respondent is important if we want to relate these weights to the personal characteristics of the respondents (such as gender, income, education and so on). In our methodology this task is trivial.

We can estimate the vector of weights (a k-dimensional vector), finding its neighbours and then calculating the weights. Predicted weights can be used as an additional diagnostic tool for checking the model: we can compare the distribution of the predicted weights of the observed data with those of artificial samples generated by using the proposed model with the estimated parameters. Suppose that N samples are generated and we calculate the empirical distribution function of the weights in the sample, then we can use a Q–Q plot to compare the empirical distribution functions of the predicted weights with the empirical distribution function of the real data.

2.5 Considerations

2.5.1 The Least Squares Method Motivation

Theorem 2.1. The linear least squares problem

$$\min_{x \in \Re^n} \| y - Ax \|$$

has at least one minimum point x_0. If x_1 is another minimum point, then $Ax_0 = Ax_1$. The residual $r = y - Ax_0$ is uniquely determined and satisfies the equation $A^T r = 0$. Every minimum point x_0 is also a solution of the normal equations $A^T Ax = A^T y$ and conversely.

Proof. Let $L \subseteq \Re^n$ be the linear subspace

$$L = \{ Ax \mid x \in \Re^n \},$$

which is spanned by the columns of A, and let $L^{\perp}$ be the orthogonal complement

$$L^{\perp} = \{ r \mid r^T z = 0 \quad for \quad all\, z \in L \} = \{ r \mid r^T A = 0 \} \tag{2.22}$$

Because $\Re^m = L \oplus L^{\perp}$, the vector $y \in \Re^m$ can be written uniquely in the form

$$y = s + r\,, \quad s \in L\,, \quad r \in L^{\perp}\,, \tag{2.23}$$

and there is at least one x_0 with $Ax_0 = s$. Because $A^T r = 0$, x_0 satisfies $A^t y = A^t s = A^t Ax_0$, that is, x_0 is a solution of the normal equations. Conversely, each solution x_1 of the normal equations corresponds to a representation (2.23)

$$y = s + r\,, \quad s = Ax_1\,, \quad r = y - Ax_1\,, \quad s \in L\,, \quad r \in L^{\perp}\,. \tag{2.24}$$

Because this representation is unique, it follows that $Ax_0 = Ax_1$ for all solutions x_0, x_1 of the normal equations. Further, each solution x_0 of the normal equations is a minimum points for the problem

$$\min_{x \in \Re^n} \|y - Ax\| ,$$

To see this, let x be arbitrary, and set

$$z = Ax - Ax_0 \ , \quad r = y - Ax_0 \ .$$

Then, since $r^T z = 0$,

$$\|y - Ax\|^2 = \|r - z\|^2 = \|r\|^2 + \|z\|^2 \geq \|r\|^2 = \|y - Ax_0\|^2 \ ,$$

that is, x_0 is a minimum point. This establishes Theorem 2.1

Δ

2.5.2 Stability of Least Squares Method

We have investigated how a minimum point x for the linear least squares problem

$$\min_{x \in \Re^n} \|y - Ax\| .$$

changes if the matrix A. and the vector y. are perturbed (Lawson, 1974).

We assume that columns of A are linearly independent. If the matrix A is replaced by $A + B$, and y is replaced by $y + \Delta y$, then the solution

$$x = \left(A^T A\right)^{-1} A^T y \ ,$$

changes to

$$x+\Delta x=\left((A+B)^T(A+B)\right)^{-1}(A+B)^T(y+\Delta y)$$

If B is small relative to A, then $\left((A+B)^T(A+B)\right)^{-1}$ exists and satisfies

$$\begin{aligned}\left((A+B)^T(A+B)\right)^{-1} &= \left(A^TA(I+(A^TA)^{-1}\left(A^TB+B^TA\right)\right)^{-1}\\ &= \left(I-(A^TA)^{-1}\left(A^TB+B^TA\right)\right)(A^TA)^{-1}\end{aligned}$$

To a first approximation $(I+F)^{-1}=I-F$ if the matrix F is small relative to I. Thus it follows that

$$\begin{aligned}x+\Delta x &= \left(A^TA\right)^{-1}A^Ty-\left(A^TA\right)^{-1}\left(A^TB+B^TA\right)\left(A^TA\right)^{-1}A^Ty\\ &\quad+\left(A^TA\right)^{-1}B^Ty+\left(A^TA\right)^{-1}A^T\Delta y\end{aligned}$$

And, nothing that

$$x=\left(A^TA\right)^{-1}A^Ty\ , \qquad r=y-Ax\ ,$$

it follows immediately that

$$\Delta x=-\left(A^TA\right)^{-1}A^TBx+\left(A^TA\right)^{-1}B^TBr+\left(A^TA\right)^{-1}A^T\Delta y\,.$$

Given, that we expect $\Delta y\to 0$ and $B\to 0$, it follows that $\Delta x\to 0$. That is, small perturbations on y and A produce small perturbations on x. Explicit formulas for bounds on $\|\Delta x\|$ can be consulted in (Stoer and Bulirsch 1980, Golub and Van Loan 1989, Rao and Toutenbourg 1999).

2.6 Statistical Interpretation

When we have applied our methodology and then we have weigh values, we can calculate several statistical parameters (Agresti, 2002).

For each costumer i, we are performing a multivariate regression, focused on the local neighbourhood of the costumer.

Therefore, our local model is:

$$y_i = \sum_{j=1}^{k} w_{ij} X_{ij} + \mu_i , \qquad i = 1,\ldots,l \ ,$$

with $\mu_i \propto N\left(0,\sigma^2\right)$

$$w_{ij} \geq 0 \qquad \forall i, \forall j$$

$$\sum_{j=1}^{k} w_{ij} = 1 \quad \forall i$$

where:

i denotes the index corresponding to the customer,

k denotes the number of quality attributes,

l denotes the number of neighbours chosen to estimate the customer weights $w_{i1},\ldots,w_{ik}$.

μ_i is the error term.

The error term captures the effects of all possible omitted variables. We suppose that the term verify the following hypothesis:

a) its expectation is equal to zero. This means that on average the errors balance out.

b) its variance is constant, σ^2. So, they are *homoscedastic*, this means that the variance of the disturbance is the same for each observation.

c) the disturbances are uncorrelated each other

d) its distribution is normal

We call U the vector $(u_i,...,u_n)$ and can summarize these hypotheses in:

$$U \propto NM\left(0, \sigma^2 I_n\right)$$

Also, we define the following conditions:

e) the number of data is, at least, k.

f) we assume that the independent variables, X_{ij}, are linearly independent. That is, no independent variable can be expressed as a (non-zero) linear combination of the remaining independent variables. The failure of this assumption, known as *multicollinearity*, clearly makes it infeasible to disentangle the effects of the supposedly independent variables.

g) we assume that the output variables, y_{ij}, are independent each other.

h) the distribution of the output variable is normal.

2.6.1 Properties of the local estimators

If the hypotheses above are satisfied:

- then the estimator $\hat{w}$ will be unbiased. Unbiasedness means that if we draw many different samples, the average value of the estimator based on each sample will be the true parameter value w.

- then it can be shown that the variance of the estimator $\hat{w}$ is given by: $Var(\hat{w}) = \sigma^2 \left(X^{'}X\right)^{-1}$. If the independent variables are highly intercorrelated, then the matrix $\left(X^{'}X\right)$ will be nearly singular and the element of $\left(X^{'}X\right)^{-1}$ will be large, indicating that the estimates of beta may be imprecise.

There are two important theorems about the properties of the estimators. The *Gauss-Markov theorem* states that under the assumptions above, the estimator $\hat{w}$ is best linear unbiased. That is, the estimator has smaller variance than any other linear unbiased estimator. (One covariance matrix is said to be larger than another if their difference is positive semi-definite.) If we add the assumption that the disturbances u_i have a joint normal distribution, then the estimator has minimum variance among all unbiased estimators.

Although the preceding theorems provide strong justification for using the estimator, it should be realized that least squares method is rather sensitive to departures from the assumptions. A few outliers (stray observations generated by a different process) can strongly influence the least squares estimates (Härdle, 1990).

2.6.2 Confidence Interval and Regions

Confidence interval for individual w_{ij} can be obtained from:

$$\hat{w}_{ij} \pm t_{\frac{\alpha}{2};g}\,\hat{s}\left(\hat{w}_{ij}\right),$$

where:

$\hat{s}\left(\hat{w}_{ij}\right)$ is the estimation of the distribution of the standard deviation.

$t_{\frac{\alpha}{2};g}$ corresponding to the value of the t distribution.

We can rewrite the formula above and then:

$$\hat{w}_{ij} \pm t_{n-k-1;\frac{\alpha}{2}} \hat{s}_R \sqrt{q_{ii}} \ ,$$

where:

q_{ii} is the element of the diagonal of $\left(X^{'}X\right)^{-1}$.

The coefficients of $\hat{w}$ are dependents, the values of the individual confidence interval can introduce mistakes of interpretation. We prefer always to build the confidence region of w from:

$$\frac{(\hat{w}-w)(X'X)(\hat{w}-w)}{\sigma^2} \propto \chi^2_{(k)} \ .$$

2.7 Computational Experiments with Simulated Data Sets

We have experimentally evaluated several simulated examples. In this section we present the results of applying the ALR methodology to a series of simulated data sets (Driscoll, 2009; Van Loan, 2010). The goal is to show the in what situations one might expect it to provide better performance than the existing methodologies (Hayes, 1998; Bober, 2009).

The comparison have been done with the mean quadratic error:

$$MCE = \frac{\sum_{j=1}^{n}\left[\sum_{i=1}^{k}\left(w_{ij}-\hat{w}_{ij}\right)^2\right]}{nk}, \quad j=1,\dots,n\,, \quad i=1,\dots,k$$

where:

w_{ij}: true weight of customer i_{TH} in the attribute j_{TH}.

$\hat{w}_{ij}$: estimated weight of customer i_{TH} in the attribute j_{TH} .

2.7.1 Data where ALR works similar than the existing methodologies

Example 1:

We have generated a data set with the following characteristics:

- All the customers surveyed provided similar answers,
- One hundred customers,
- Five quality dimensions or attributes,
- Variables scale from 0 to 100, with one decimal point.
- Every customer assigns approximately the following weights: w_1=0.5, w_2=0.1, w_3=0.3, w_4=0.05, w_5=0.05.

One traditional method, for example the least squares method with linear constraints, estimates the following weights:

$\hat{w}_1$=0.542, $\hat{w}_2$=0.093, $\hat{w}_3$=0.286, $\hat{w}_4$=0.059, $\hat{w}_5$=0.020

$MCE = 0{,}000598$

Our method, ALR, using a weighted Euclidean metric and adaptive neighbourhood with less than 8 neighbours, estimates the following weights:

	$\hat{w}_1$	$\hat{w}_2$	$\hat{w}_3$	$\hat{w}_4$	$\hat{w}_5$
Mean	0.50	0.10	0.30	0.05	0.05
Median	0.50	0.10	0.30	0.05	0.05
Mode	0.50	0.10	0.30	0.05	0.05
Trimmed mean (0.05)	0.50	0.10	0.30	0.05	0.05
Trimmed mean (0.10)	0.50	0.10	0.30	0.05	0.05

Table 2.1. Estimated weights for Example 1 by ALR methodology

$MCE \approx 10^{-9}$

In the following figure you can find the estimation of the weights for example 1 by ALR methodology.

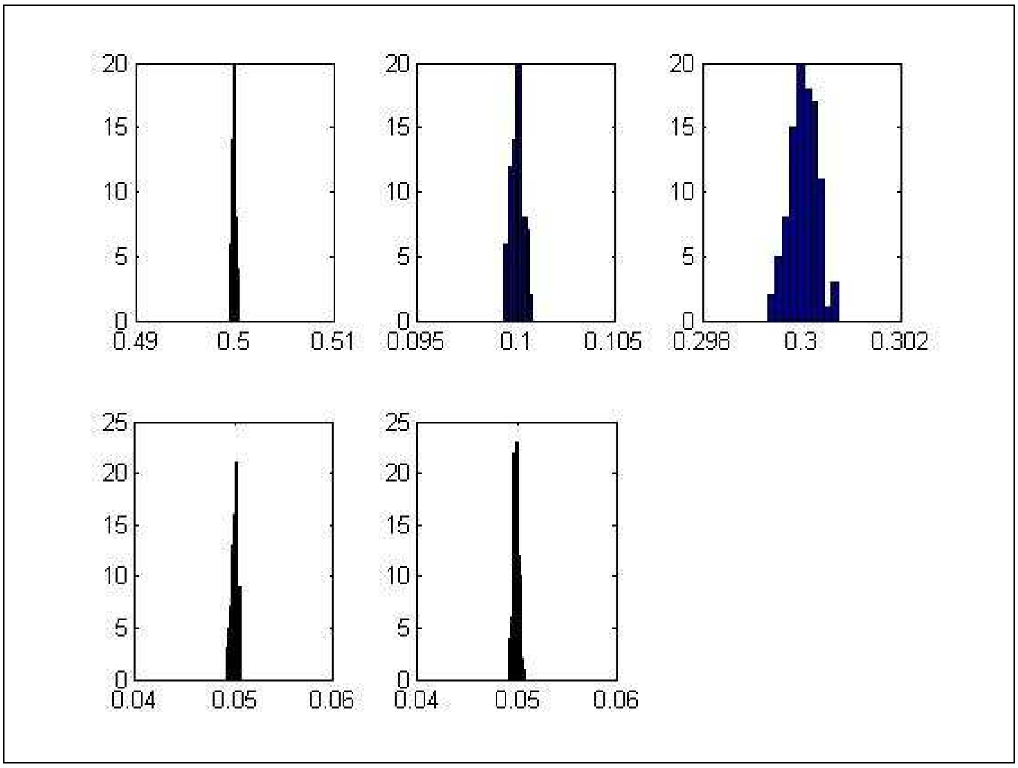


Figure 2.3. Estimated weights for Example 1 by ALR methodology

Example 2:

We have generated data with the following characteristics:

- All the customers surveyed provided similar answers,
- One hundred customers,
- Five quality dimensions or attributes,
- Variables scale from 0 to 100, with one decimal point.
- Every customer assigns approximately the following weights: w_1=0.6, w_2=0.4, w_3=0, w_4=0, w_5=0.

With a traditional method we can estimate the following weights:

$\hat{w}_1$=0.599, $\hat{w}_2$=0.401, $\hat{w}_3$=0.00, $\hat{w}_4$=0.00, $\hat{w}_5$=0.00

$MCE \approx 10^{-7}$

Our method, ALR, using a weighted Euclidean metric and adaptive neighbourhood with less than 8 neighbours, estimates the following weights:

	$\hat{w}_1$	$\hat{w}_2$	$\hat{w}_3$	$\hat{w}_4$	$\hat{w}_5$
Mean	0.60	0.40	0.00	0.00	0.00
Median	0.60	0.40	0.00	0.00	0.00
Mode	0.60	0.40	0.00	0.00	0.00
Trimmed mean (0.05)	0.60	0.40	0.00	0.00	0.00
Trimmed mean (0.10)	0.60	0.40	0.00	0.00	0.00

Table 2.2. Estimated weights for Example 2 by ALR methodology

$MCE \approx 10^{-12}$

In the following figure you can find the estimation of the weights for example 2 by ALR methodology.

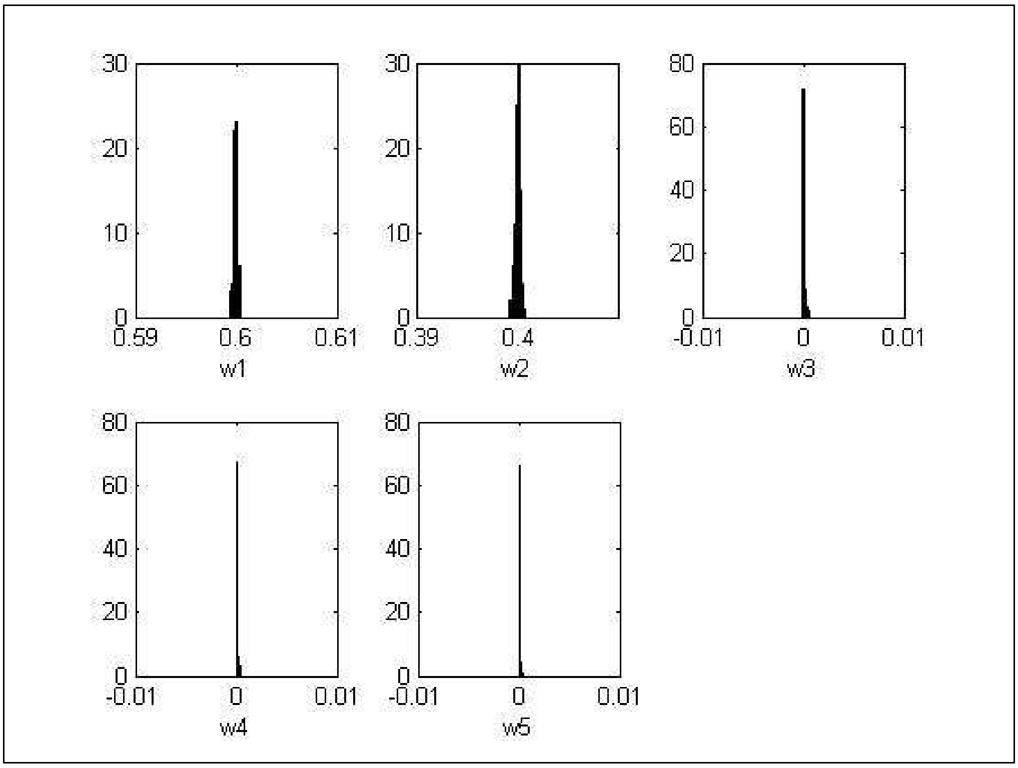


Figure 2.4. Estimated weights for Example 2 by ALR methodology

Example 3:

We have generated data with the following characteristics:

- All the customers surveyed provided similar answers,
- One hundred customers,
- Five quality dimensions or attributes,
- Variables scale from 0 to 100, with one decimal point.
- Every customer assigns exactly the following weights: w_1=0.2, w_2=0.2, w_3=0.2, w_4=0.2, w_5=0.2. “When everything I important, then nothing is important”.

With a traditional method we can estimate the following weights:

$\hat{w}_1$=0.203, $\hat{w}_2$=0.196, $\hat{w}_3$=0.199, $\hat{w}_4$=0.196, $\hat{w}_5$=0.206

$MCE \approx 10^{-5}$

Our method, ALR, using a weighted Euclidean metric and adaptive neighbourhood with less than 8 neighbours, estimates the following weights:

	$\hat{w}_1$	$\hat{w}_2$	$\hat{w}_3$	$\hat{w}_4$	$\hat{w}_5$
Mean	0.20	0.20	0.20	0.20	0.20
Median	0.20	0.20	0.20	0.20	0.20
Mode	0.20	0.20	0.20	0.20	0.20
Trimmed mean (0.05)	0.20	0.20	0.20	0.20	0.20
Trimmed mean (0.10)	0.20	0.20	0.20	0.20	0.20

Table 2.3. Estimated Weights for Example 3 by ALR methodology

$MCE \approx 10^{-13}$

In the following figure you can find the estimation of the weights for example 3 by ALR methodology.

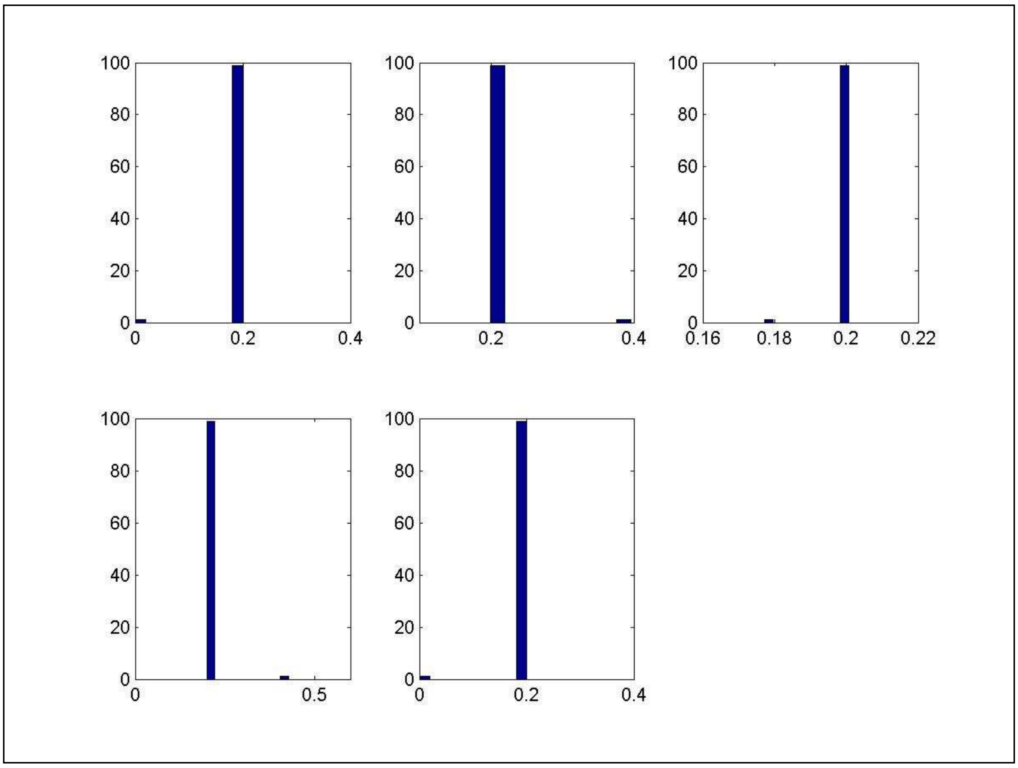


Figure 2.5. Estimated Weights for Example 3 by ALR methodology

2.7.2 Data where ALR works better than the existing methodologies

Example 4:

We have generated data with the following characteristics:

- All the customers surveyed provides similar answers,
- Two different groups provided similar answers, population 1 and population 2,
- One hundred customers, fifty customers in each population,
- Five quality dimensions or attributes,
- Variables scale from 0 to 100, with one decimal point.
- Every customer, in the population 1, assigns exactly the following weights: w_{11}=0.5, w_{12}=0.5, w_{13}=0, w_{14}=0, w_{15}=0.
- Every customer, in the population 2, assigns exactly the following weights: w_{21}=0, w_{22}=0, w_{23}=0, w_{24}=0.5, w_{25}=0.5.

With a traditional method we can estimate the following weights:

$$\hat{w}_1=0.571,\ \hat{w}_2=0.213,\ \hat{w}_3=0.000,\ \hat{w}_4=0.073,\ \hat{w}_5=0.143$$

$$MCE \approx 10^{-2}$$

Our method, ALR, using a weighted Euclidean metric and adaptive neighbourhood with less than 10 neighbours, estimates the following weights:

	$\hat{w}_{11}$	$\hat{w}_{12}$	$\hat{w}_{13}$	$\hat{w}_{14}$	$\hat{w}_{15}$
Mean	0.50	0.50	0.00	0.00	0.00
Median	0.50	0.50	0.00	0.00	0.00
Mode	0.50	0.50	0.00	0.00	0.00

Table 2.4. Estimated Weights for Example 4, population 1 by ALR methodology

	$\hat{w}_{21}$	$\hat{w}_{22}$	$\hat{w}_{23}$	$\hat{w}_{24}$	$\hat{w}_{25}$
Mean	0.00	0.00	0.00	0.50	0.50
Median	0.00	0.00	0.00	0.50	0.50
Mode	0.00	0.00	0.00	0.50	0.50

Table 2.5. Estimated Weights for Example 4, population 2 by ALR methodology

$MCE \approx 10^{-8}$

In the following figure you can find the estimation of the weights for example 4 by ALR methodology.

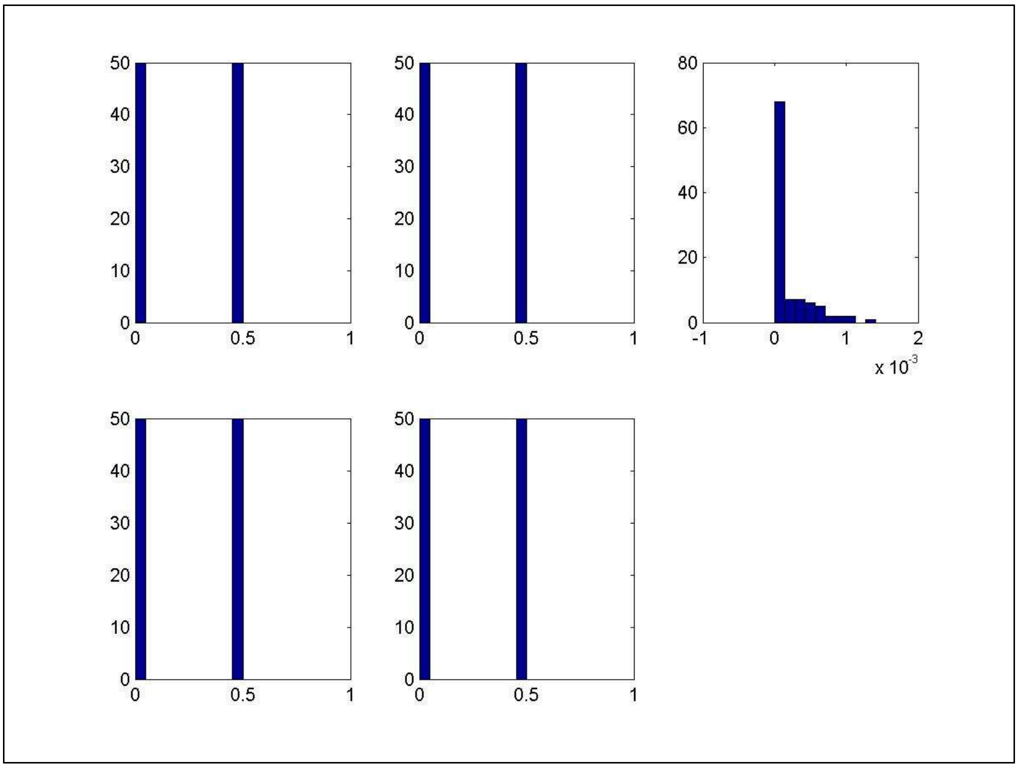


Figure 2.6. Estimated Weights for Example 4 by ALR methodology

Example 5:

We have generated data with the following characteristics:

- Two different groups provided similar answers, population 1 and population 2,
- One hundred customers, fifty customers in each population,
- Five quality dimensions or attributes,
- Variables scale from 0 to 100, with one decimal point.
- Every customer in the population 1 assigns exactly the following weights: w_{11}=0.2, w_{12}=0.2, w_{13}=0.2, w_{14}=0.2, w_{15}=0.2.
- Every customer in the population 2 assigns exactly the following weights: w_{21}=0.1, w_{22}=0.1, w_{23}=0.1, w_{24}=0.35, w_{25}=0.35.

With a traditional method we can estimate the following weights:

$$\hat{w}_1=0.0892,\ \hat{w}_2=0.2077,\ \hat{w}_3=0.0198,\ \hat{w}_4=0.2660,\ \hat{w}_5=0.4173$$

$$MCE \approx 10^{-1}$$

Our method, ALR, using a weighted Euclidean metric and adaptive neighbourhood with less than 10 neighbours, estimates the following weights:

	$\hat{w}_{11}$	$\hat{w}_{12}$	$\hat{w}_{13}$	$\hat{w}_{14}$	$\hat{w}_{15}$
Mean	0.200	0.200	0.200	0.200	0.200
Median	0.200	0.200	0.200	0.200	0.200
Mode	0.200	0.200	0.200	0.200	0.200

Table 2.6. Estimated Weights for Example 5, population 1 by ALR methodology

	$\hat{w}_{21}$	$\hat{w}_{22}$	$\hat{w}_{23}$	$\hat{w}_{24}$	$\hat{w}_{25}$
Mean	0.100	0.100	0.100	0.350	0.350
Median	0.100	0.100	0.100	0.350	0.350
Mode	0.100	0.100	0.100	0.350	0.350

Table 2.7. Estimated Weights for Example 5, population 2 by ALR methodology

$MCE \approx 10^{-7}$

In the following figure you can find the estimation of the weights for example 5 by ALR methodology.

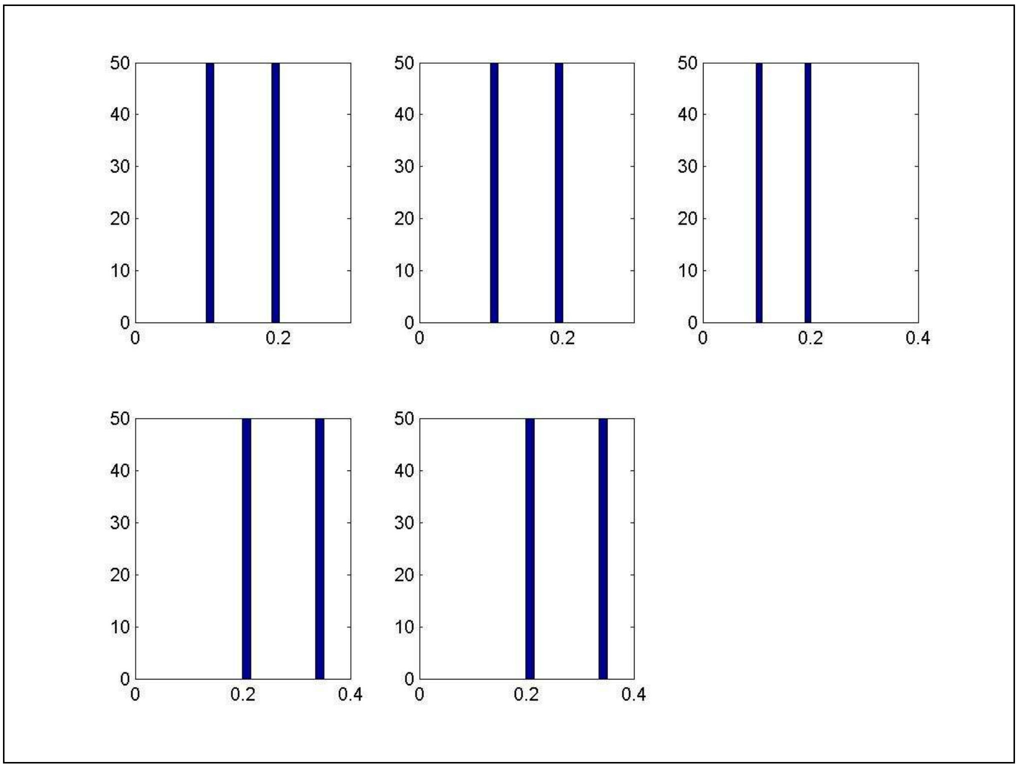


Figure 2.7. Estimated Weights for Example 5 by ALR methodology

Example 6:

We have generated data with the following characteristics:

- Three different groups provided completely different answers, population 1 and population 2 and population 3,
- Thirty, thirty and forty customers, respectively, in each population,
- Five quality dimensions or attributes,
- Variables scale from 0 to 100, with one decimal point.

- Every customer in the population 1 assigns approximately the following weights: w_{11}=0.5, w_{12}=0.5, w_{13}=0, w_{14}=0, w_{15}=0.
- Every customer in the population 2 assigns approximately the following weights: w_{21}=0, w_{22}=0, w_{23}=0, w_{24}=0.5, w_{25}=0.5.
- Every customer in the population 3 assigns approximately the following weights: w_{31}=0.2, w_{32}=0.2, w_{33}=0.2, w_{34}=0.2, w_{35}=0.2.

With a traditional method we can estimate the following weights:

$$\hat{w}_1=0.4680,\ \hat{w}_2=0.3923,\ \hat{w}_3=0.0000,\ \hat{w}_4=0.1397,\ \hat{w}_5=0.0000$$

$$MCE \approx 10^{-1}$$

Our method, ALR, using a weighted Euclidean metric and adaptive neighbourhood with less than 10 neighbours, estimates the following weights:

	$\hat{w}_{11}$	$\hat{w}_{12}$	$\hat{w}_{13}$	$\hat{w}_{14}$	$\hat{w}_{15}$
Mean	0.502	0.498	0.000	0.000	0.000
Median	0.500	0.500	0.000	0.000	0.000
Mode	0.500	0.500	0.000	0.000	0.000

Table 2.8. Estimated Weights for Example 6, population 1 by ALR methodology

	$\hat{w}_{21}$	$\hat{w}_{22}$	$\hat{w}_{23}$	$\hat{w}_{24}$	$\hat{w}_{25}$
Mean	0.006	0.001	0.002	0.521	0.470
Median	0.000	0.000	0.000	0.500	0.500
Mode	0.000	0.000	0.000	0.500	0.500

Table 2.9. Estimated Weights for Example 6, population 2 by ALR methodology

	$\hat{w}_{31}$	$\hat{w}_{32}$	$\hat{w}_{33}$	$\hat{w}_{34}$	$\hat{w}_{35}$
Mean	0.193	0.206	0.208	0.211	0.182
Median	0.200	0.200	0.200	0.200	0.200
Mode	0.200	0.200	0.200	0.200	0.200

Table 2.10. Estimated Weights for Example 6, population 3 by ALR methodology

$MCE \approx 10^{-5}$

In the following figure you can find the estimation of the weights for example 6 by ALR methodology.

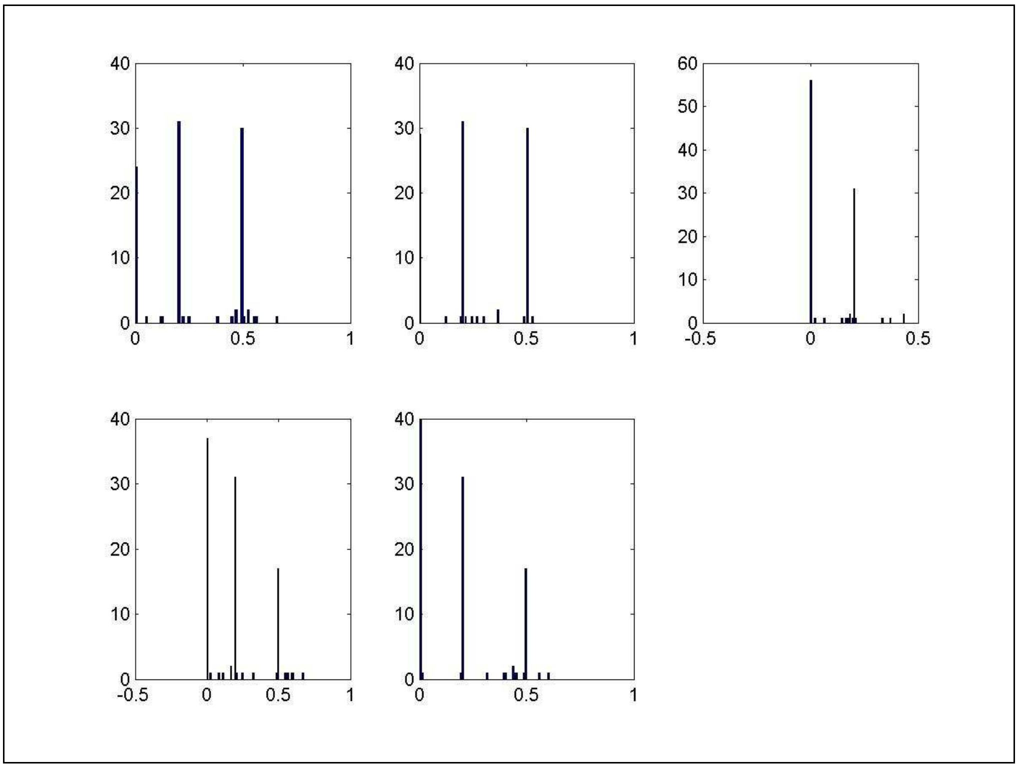


Figure 2.8. Estimated Weights for Example 6 by ALR methodology

Example 7:

We have generated data with the following characteristics:

- Four different groups provided completely different answers, population 1 and population 2, population 3 and population 4,
- One thousand customers, two hundred and fifty customers in each population,
- Five quality dimensions or attributes,
- Variables scale from 0 to 100, with one decimal point.
- Every customer in the population 1 assigns approximately the following weights: w_{11}=0.2, w_{12}=0.2, w_{13}=0.2, w_{14}=0.2, w_{15}=0.2.
- Every customer in the population 2 assigns approximately the following weights: w_{21}=0.3, w_{22}=0.4, w_{23}=0, w_{24}=0.1, w_{25}=0.2.
- Every customer in the population 3 assigns approximately the following weights: w_{31}=0, w_{32}=0.2, w_{33}=0.3, w_{34}=0, w_{35}=0.5.
- Every customer in the population 4 assigns approximately the following weights: w_{41}=0.5, w_{42}=0, w_{43}=0.4, w_{44}=0.1, w_{45}=0.

With a traditional method we can estimate the following weights:

$$\hat{w}_1=0.2444,\ \hat{w}_2=0.2906,\ \hat{w}_3=0.1861,\ \hat{w}_4=0.1642,\ \hat{w}_5=0.1147$$

$$MCE \approx 0{,}3$$

Our method, ALR, using a weighted Euclidean metric and adaptive neighbourhood with less than 10 neighbours, estimates the following weights:

	$\hat{w}_{11}$	$\hat{w}_{12}$	$\hat{w}_{13}$	$\hat{w}_{14}$	$\hat{w}_{15}$
Mean	0.223	0.197	0.261	0.196	0.123
Median	0.207	0.202	0.214	0.198	0.197
Mode	0.200	0.200	0.200	0.200	0.200

Table 2.11. Estimated Weights for Example 7, population 1 by ALR methodology

	$\hat{w}_{21}$	$\hat{w}_{22}$	$\hat{w}_{23}$	$\hat{w}_{24}$	$\hat{w}_{25}$
Mean	0.304	0.426	0.0601	0.1129	0.097
Median	0.305	0.402	0.000	0.105	0.200
Mode	0.300	0.400	0.000	0.100	0.200

Table 2.12. Estimated Weights for Example 7, population 2 by ALR methodology

	$\hat{w}_{31}$	$\hat{w}_{32}$	$\hat{w}_{33}$	$\hat{w}_{34}$	$\hat{w}_{35}$
Mean	0.062	0.212	0.302	0.045	0.379
Median	0.000	0.207	0.296	0.000	0.476
Mode	0.000	0.200	0.300	0.000	0.500

Table 2.13. Estimated Weights for Example 7, population 3 by ALR methodology

	$\hat{w}_{41}$	$\hat{w}_{42}$	$\hat{w}_{43}$	$\hat{w}_{44}$	$\hat{w}_{45}$
Mean	0.492	0.024	0.415	0.035	0.044
Median	0.494	0.000	0.411	0.097	0.000
Mode	0.500	0.000	0.400	0.100	0.000

Table 2.14. Estimated Weights for Example 7, population 4 by ALR methodology

$MCE \approx 10^{-3}$

In the following figure you can find the estimation of the weights for example 6 by ALR methodology.

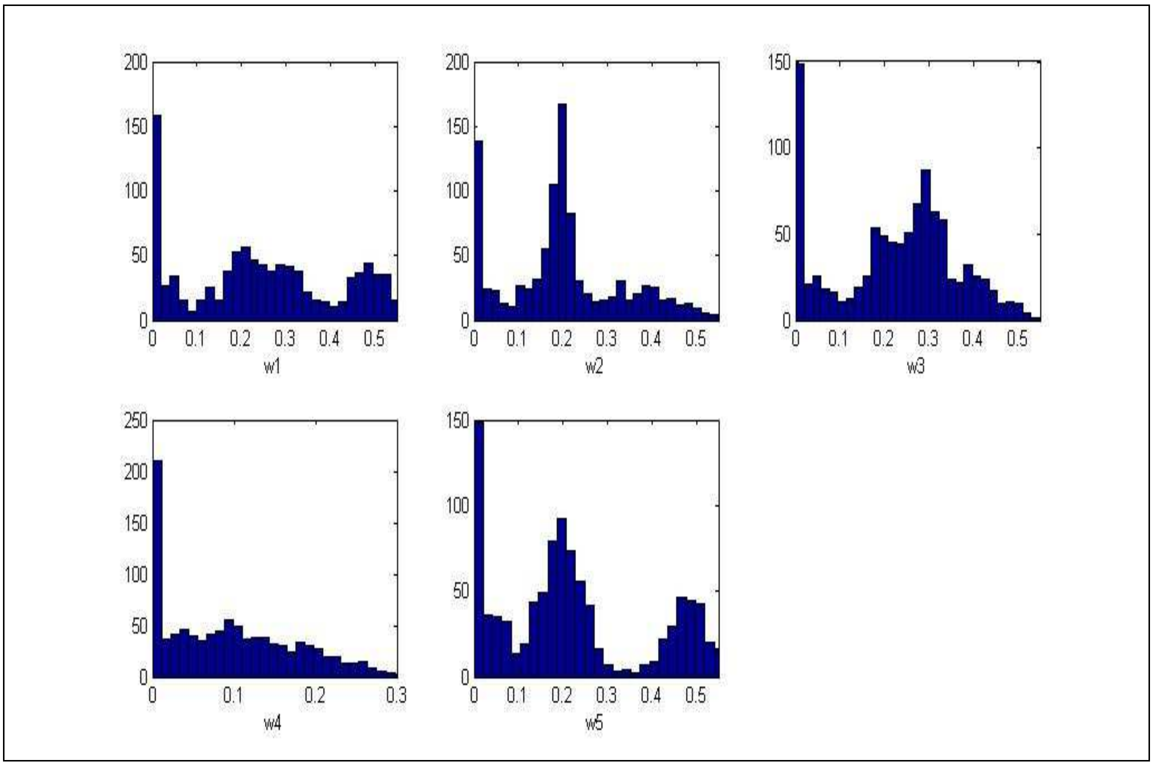


Figure 2.9. Estimated Weights for Example 7 by ALR methodology

2.8 Conclusions of Chapter 2

The classical point of view is not able to detect the correct weights for heterogeneous populations.

We have developed a methodology to measure the quality of services. We propose a non parametric quality model based on resample and nearest K-neighbours techniques in which the individual weights for each customer can be estimated.

We have experimentally evaluated several simulated examples. We presented the results of applying the ALR methodology to a series of simulated data sets. The goal was to show the in what situations one might expect it to provide better performance than the existing methodologies. Results were satisfactory, ALR is able to manage this king of data.

The proposed methodology, ALR, presents several advantages respect to the “classical tools”:

- We can use parallel computation.
- When the decision maker needs a single index, a scalar measure to summarize the performance, we can define it.
- When we have estimated each weight, we can use any kind of multivariate method to determine new groups of customers, such us a posterior customer segmentation, and then to prepare inferences about it.
- We can work, then, directly with weights that each customer assigns to each quality attribute. In fact, we don’t accept the mean of the weights as a good representative estimator. It must be remembered that the mean is only a good descriptive measure when we have a homogeneous sample and that it can be very non representative when the data comes from a mixture of very different populations of customers.
- We estimate the weights that each customer assigns to each quality attribute with the information obtained from its similar customers. We choose the set of “similarities” based on the nearest neighbourhood estimate.

Chapter 3

Real Case: Measuring the Driving Quality of CABINTEC the "Intelligent cabin truck for road transport"

ALR methodology was applied to measure the quality of CABINTEC the "Intelligent cabin truck for road transport".

We have a video from a truck simulator where several internal truck magnitudes were stored. Three experts were asked to evaluate the driving risk using a Visual Analog Scale (VAS). We have used the evaluation of the three experts to find the weights that they assign to every dimension of the simulation.

Results show that the risks correspond to abnormal behaviours of the driver and these risks are related to variables as speed and the angle of the steering wheel (SWA).

3.1. Introduction

Mobility is a key factor in the European economy. In general, the transport sector employs more than 10 million people and is responsible for more than 10% of the European Gross Domestic Product (GDP). Approximately, road traffic absorbs 44% of the total freight transport. Insurance of traffic safety is one of the main state priorities. The price paid for the mobility in Europe, mainly in social costs, is still too high. Distraction at the steering-wheel is responsible of 42% of the road fatalities. In spite of the reduction of traffic accidents in Europe since 2003, it is still alarming (Trezise et al, 2006).

The relation between road fatalities and distractions it is well known. US research estimates that distracted driving accounts for 25−50% of all vehicle crashes (Wang et al, 1996). Nevertheless, given the high number of driver distraction factors there is not a general definition of driving risk.

3.2 CABINTEC

CABINTEC ("Intelligent cabin truck for road transport") (cabintec, 2010) is an ongoing project funded by the Spanish Ministry of Science and Innovation involving 16 partners (universities, research centres and private companies). This project is focused on risk reduction for traffic safety. Three main aspects of traffic safety are considered:

- road,
- vehicle,
- driver.

All the signals and events in the vehicle are recorded and studied and many sensors are added to the truck. The drivers are surveyed to measure their lack of attention (ISO/TR 21707:2008).

The objectives of the project are:

- to design a new safe truck cabin,
- to develop an architecture to integrate the new components,
- to increase traffic safety,
- to identify overturning risk,
- to develop a system for the reconstruction of accidents,
- to identify unsuitable behaviour and lack of attention.

Unsuitable behaviour is the one that causes risk (for instance, drivers hands located far from the steering wheel and particularly in areas such as GPS, radio, or mobile phone). In order to make decisions in real time and to analyze the driver behaviour to be labelled as suitable (no risk) or unsuitable (risk), computer vision processes data and a high level knowledge algorithm is applied. The key idea is to develop a system which will prove assistance for the driver, a type of virtual co-driver. It should be able to notify the main driver when there is a driving risk.

The categorization of risk or safe in traffic driving is hard to measure. Due to its subjective entity and the great number of elements involved, it is a multidimensional measurement. The goal is to get objectivity in a fundamentally subjective phenomenon and with the added characteristic of a high individual variability. So, the thresholds of safety are hard to evaluate. Our aim will be to calculate some of these thresholds combining the information from experts (Martín de Diego, Conde and Cabello, 2009).

In the literature, there are examples of analysis of risk based on video images (Dingus et al, 2005; Lauro, 2002; Klauer et al, 2006). The results indicate that the speed, driving while drowsy, drivers eyes off the forward roadway for more than two seconds and aggressive driving behaviours are related to an increased driving risk. In addition, Fuzzy Logic has been used to predict the driver behaviour (Inkamon et al 2008, Tronci et al 2007) an index measure to be used for the selection of the experts is presented.

3.3 Simulator and Evaluation of Risk by Experts

The CABINTEC project will be tested in a truck simulator. The simulator is a real truck cabin placed over mechanical actuators so the effects of driving are very realistic. All the visual field of the driver is covered by a detailed 3D scene simulation. The scene and the actuators are coordinated by a computer, so the driving feeling is near real (the actuators move the cabin in case of bumps or hits). Driver feels like driving in a real truck few minutes after the start. The scene involves real traffic and interactions with other vehicles. So, the driver attention will be conducted to the driving fact.

Figure 3.1. Truck simulator

A video was recorded from the truck simulator. To do this, a professional driver was employed. No previous information was given to the driver. So, a natural driver behaviour was expected. Information on several variables was collected at the simulator:

- the speed of the truck,
- the revolutions per minute (RPM),
- the angle of the steering wheel (SWA),
- the position of the truck on the road,
- and images of the truck cabin.

A graphical interface is used in order to get the risk evaluation from the experts. To collect this information, a Visual Analog Scale (VAS) is employed. This method has been considered the best for subjective measurements, recommended by World Health Organization to measure pain (Cork et al, 2004).

Figure 3.2. Graphical interface

Three safety experts of the Spanish Automobile Royal Club ("RACE") evaluated the risk level at each point assigning a value in the VAS line according to their individual perception. The measures were taking from the zero dot to 100. This method was chosen because it is a simple method well correlated with other descriptive scales, it has good sensibility and liability, and it is easily repeated.

The experts evaluate the driver's behaviour during 10 minutes, that corresponds to 35000 clock cycles (Singh, 1986). Pre-processing of data decide eliminate the firsts 5000 clock cycles, because the warm up and start up of the measure process. Thus, three evaluations of the risk, one for each expert, from 0 to 100 were acquired. These evaluations are presented in figure 3.3.

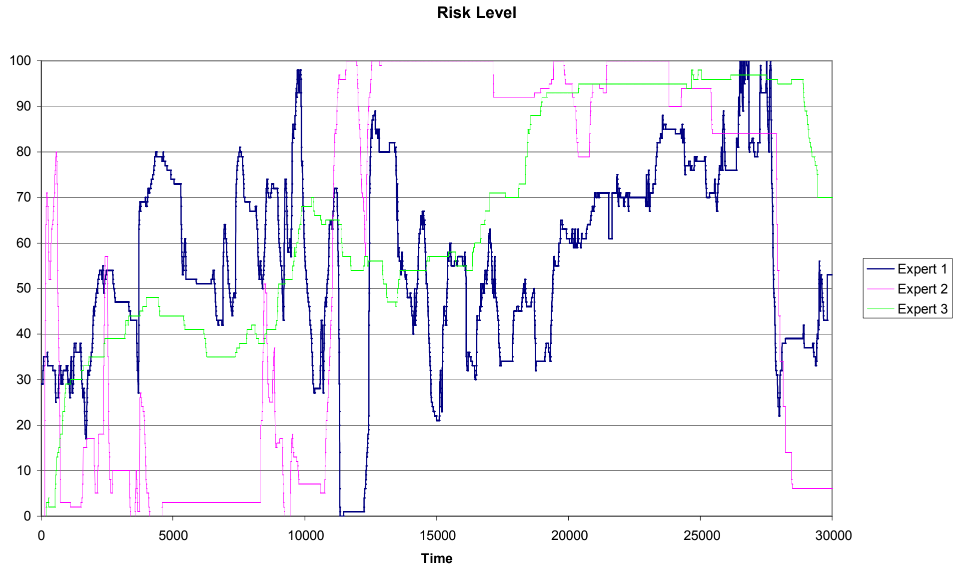


Figure 3.3. Quantitative evaluation of the risk by the experts

Notice that the expert number one (the blue line) presents an evaluation of the risk more stable (with lowest variability) than the evaluation of the risk built by the other two experts (the red green and pink line).

3.4 Driving Risk

We could enumerate several driving risk definitions (see Martín de Diego, Conde and Cabello, 2009). In our case, we have defined the output variable driving risk, "*RiskLevel*" with a weighted combination of three expert evaluations.

$$RiskLevel = \sum_{j=1}^{n} \alpha_i Risk_i \ , \tag{3.1}$$

where:

n : number of experts.

α_i : prior weight. It depends on the background of the expert i.

$Risk_i$: individual risk level evaluation.

In the following figure we can see the values of the combination selected, for starting we have used prior weights $\alpha_i = \frac{1}{3}$, $i = 1,\dots,3$. (see Barron and Barrett, 1996).

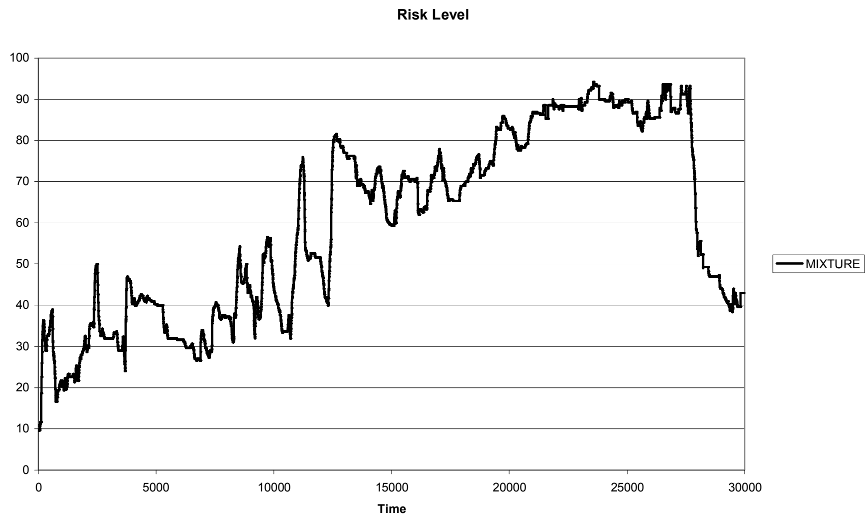


Figure 3.4. Risk level for the combination of experts result

3.5 Results

In the following section we present the results of applying our methodology in the CABINTEC project datasets. The goal is to determine the relative importance of each input variable to predict the risk level and then to develop a system that will provide assistance for the driver (Christensen, 2001). Results for each individual expert and for the mixture of experts are presented.

We have defined two subsets of the data:

- Training subset: It is used to estimate the weights ("train"). The size of this subset is called n.
- Predict subset: It is used to validate de model established. The size of this subset is called Tn.

3.5.1 Qualitative results

We have observed the video to do a qualitative classification of data. We may conclude that risk level always is high in the instant (Fuchs 1998; Rencher, 2002; Spicer, 2005):

- when the driver is using the mobile phone with one hand, and simultaneously using the gearshift with the other hand. So no hands at the driving wheel were present.
- when the driver is using the mobile phone while crossing the traffic lines.
- when the driver is talking to another passenger while parking.
- when the driver is using the mobile phone while turning at low speed.

3.5.2 Results for the mixture of experts

In the figures 3.5, 3.6, 3.7 and 3.8 we can see the results for the mixture of experts.

3.5.2.1 Results for n=20000 and Tn=10000

Figures 3.5 and 3.6 show the results with training subset size of n=20000. Figure 3.6 shows the difference between positive and negative errors in the estimation. Errors are symmetric.

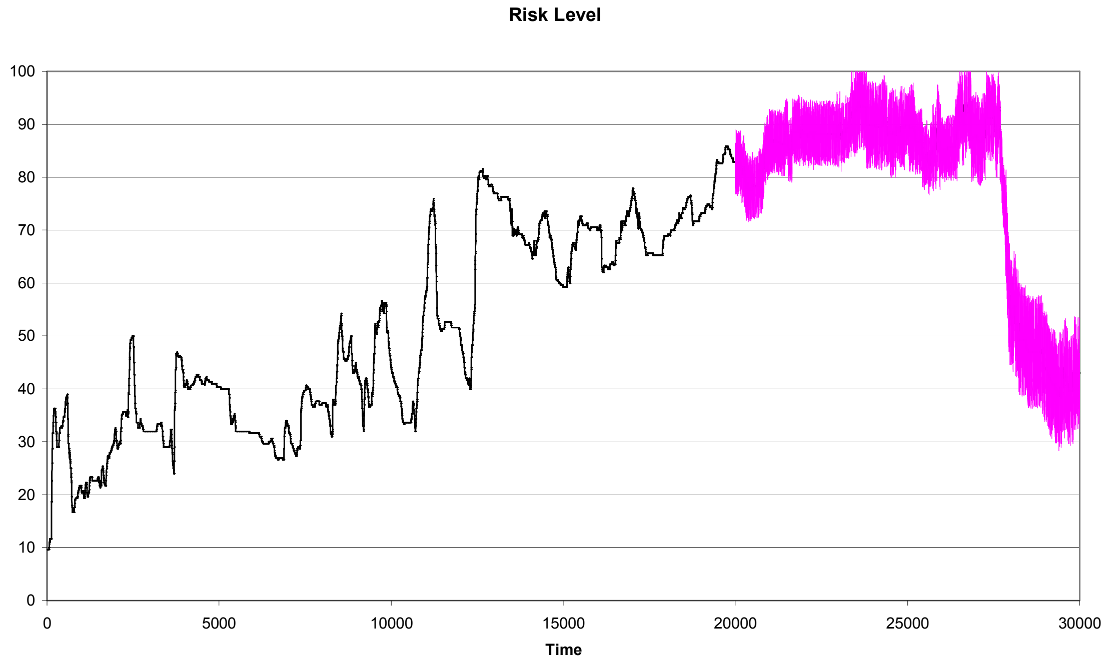


Figure 3.5. Driving risk evaluation given by mixture of experts.
n = 20000 and Tn = 10000

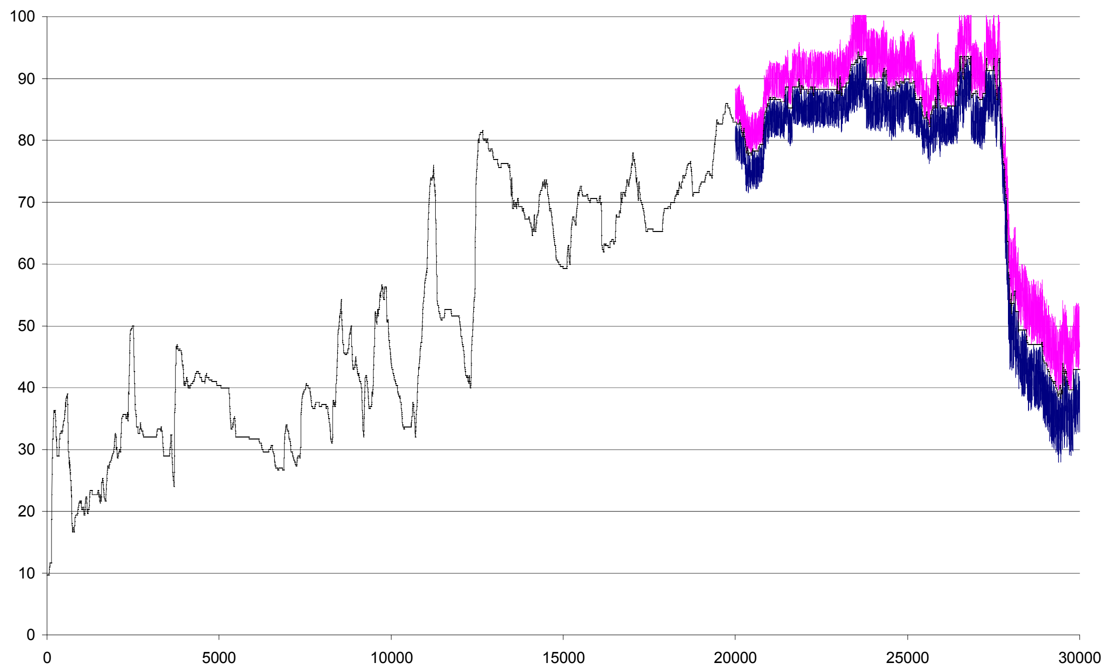


Figure 3.6. Driving risk evaluation given by mixture of experts.
n = 20000 and Tn = 10000. Positive and negative errors.

3.5.2.2 Results for n=15000 and Tn=15000

Figures 3.7 and 3.8 show the results with training subset size of n=15000. Figure 3.8 shows the difference between positive and negative errors in the estimation. Errors are symmetric and became greater than the above situation.

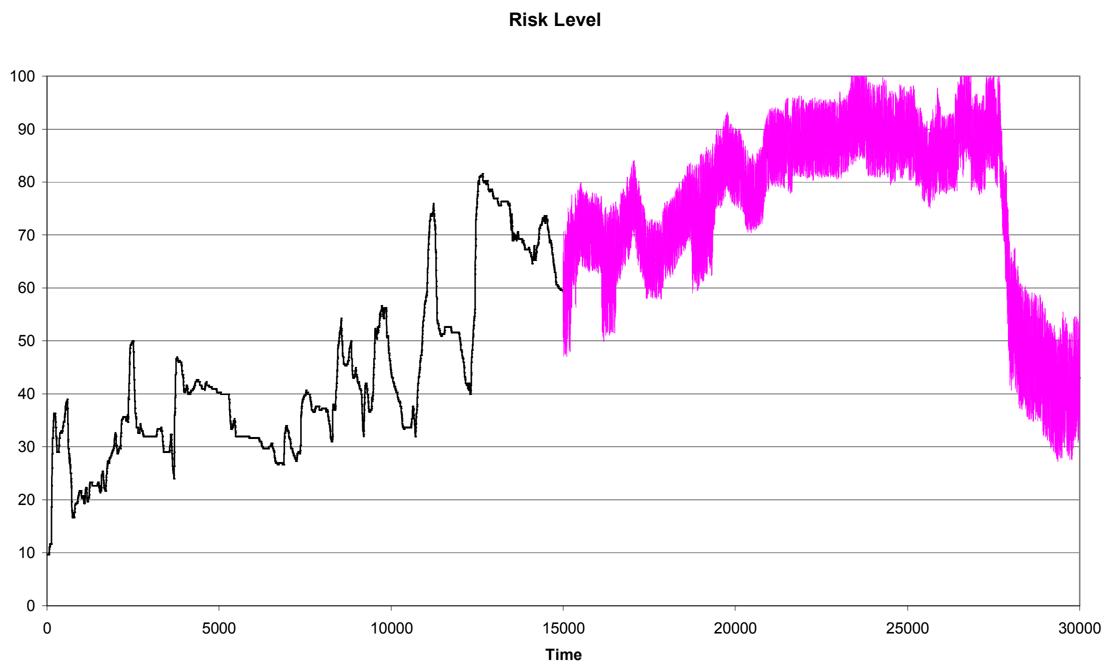


Figure 3.7. Driving risk evaluation given by mixture of experts.
n = 15000 and Tn = 15000

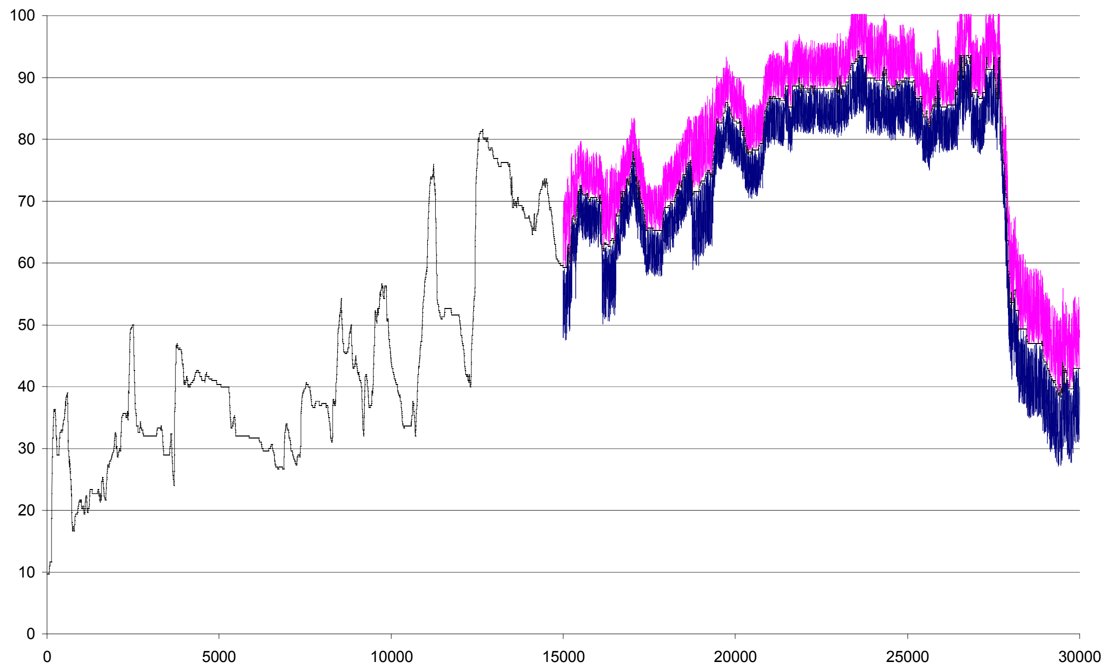


Figure 3.8. Driving risk evaluation given by mixture of experts.
n = 15000 and Tn = 15000. Positive and negative errors.

In the following table a comparison of results for mixture of experts are presented:

	n = 20000	n = 15000
max (absolute error)	7.12	9.66
mean squared error	0.72	0.83

Table 3.1. Results for the mixture of experts with n=20000 and n=15000.

3.5.3 Results for the expert 1

In the figures 3.9, 3.10 and 3.11 we can see the results for the experts 1.

Figures 3.9 show 4 results with several training subset size between n=25000 and n=10000.

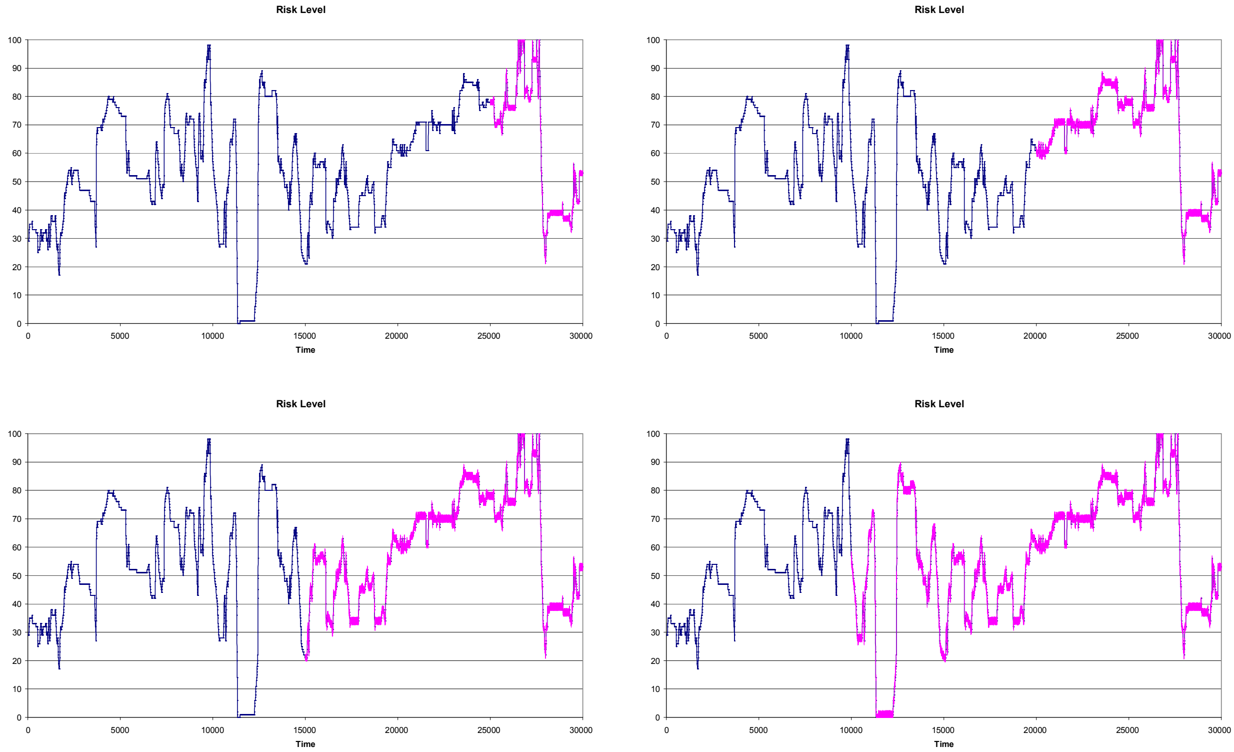


Figure 3.9. Driving risk evaluation given by expert 1 (N=25000, 20000, 15000 and 10000).

Figures 3.10 and 3.11 shows, again, the result for n=20000 and n=25000.

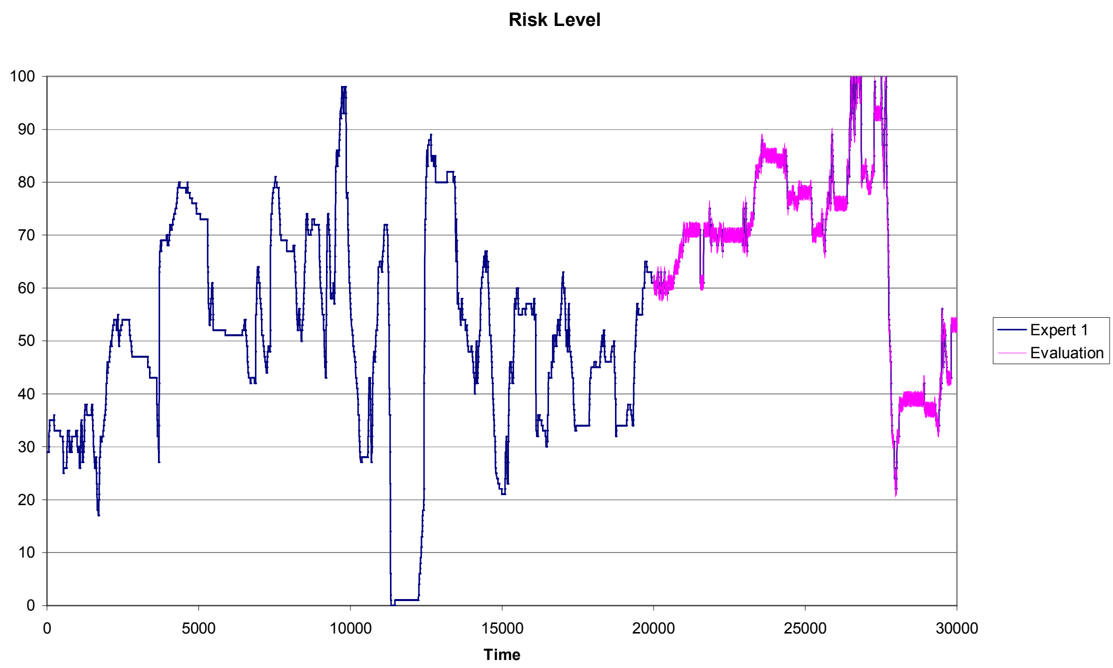


Figure 3.10. Driving risk evaluation given by expert 1.
n = 20000 and Tn = 10000

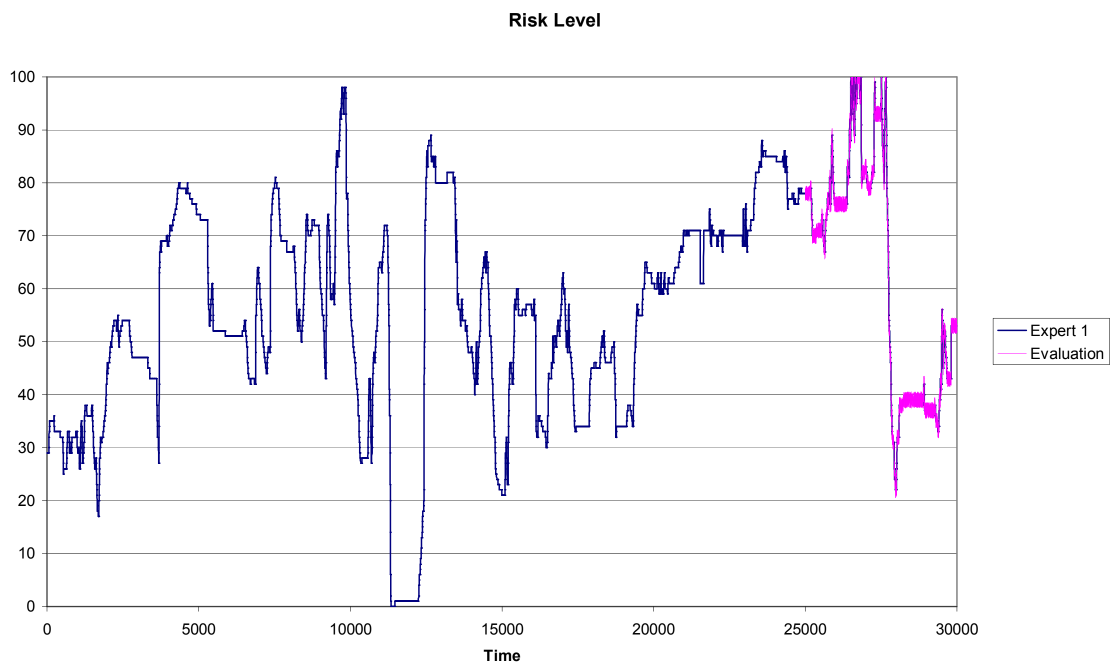


Figure 3.11. Driving risk evaluation given by expert 1.
n = 25000 and Tn = 5000

In the following table a comparison of results for expert 1 are presented:

	n = 25000	n = 20000	n = 15000	n = 10000
max (absolute error)	1.052	1.286	1.422	1.466
mean squared error	0.362	0.552	0.689	0.732

Table 3.2. Results for the expert 1 with n=25000; n=20000; n=15000 and n=10000.

3.5.4 Results for the expert 2

In the figures 3.12, 3.13 and 3.14 we can see the results for the experts 2.

Figures 3.12 show 4 results with several training subset size between n=25000 and n=10000.

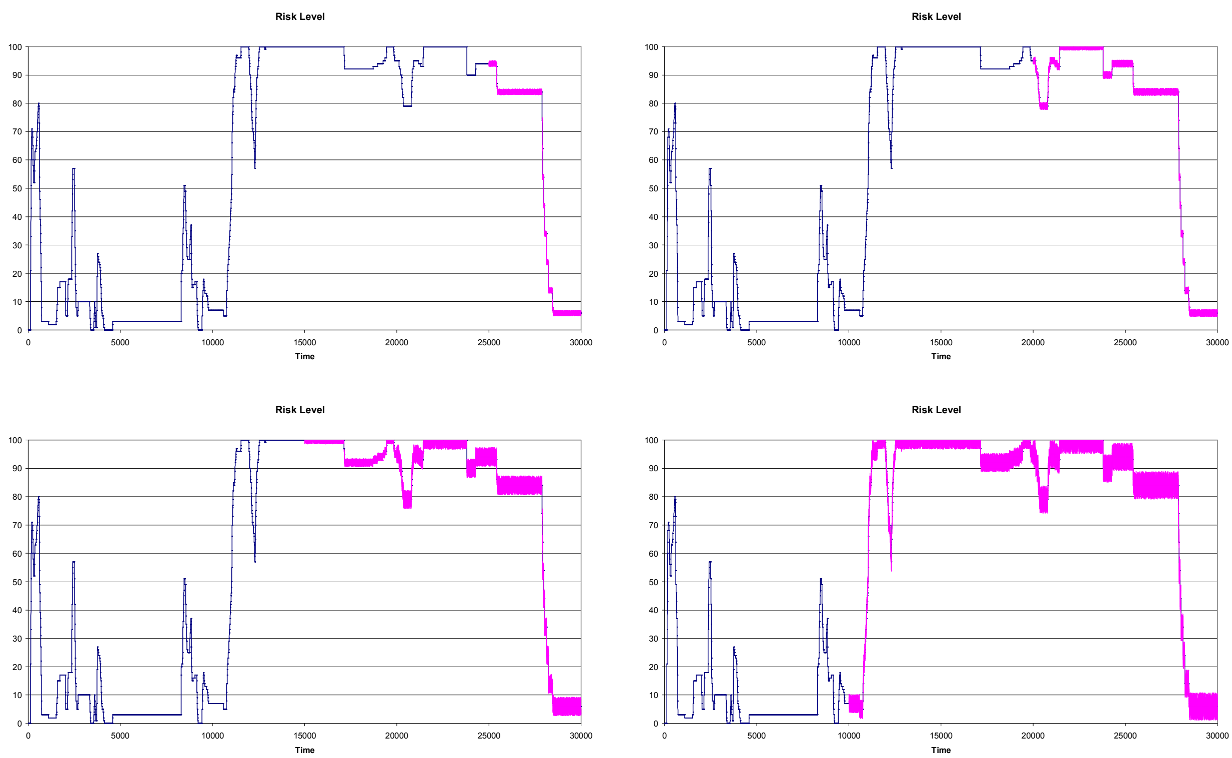


Figure 3.12. Driving risk evaluation given by expert 2 (N=25000, 20000, 15000 and 10000).

Figures 3.13 and 3.14 shows, again, the result for n=20000 and n=25000.

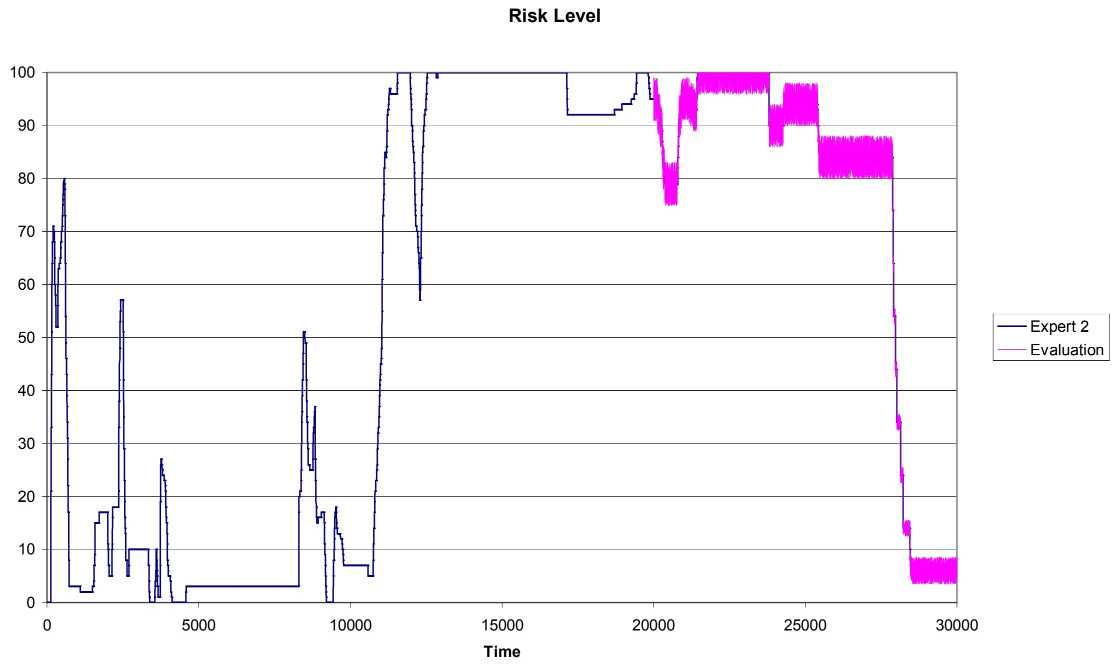


Figure 3.13. Driving risk evaluation given by expert 2.

n = 20000 and Tn = 10000

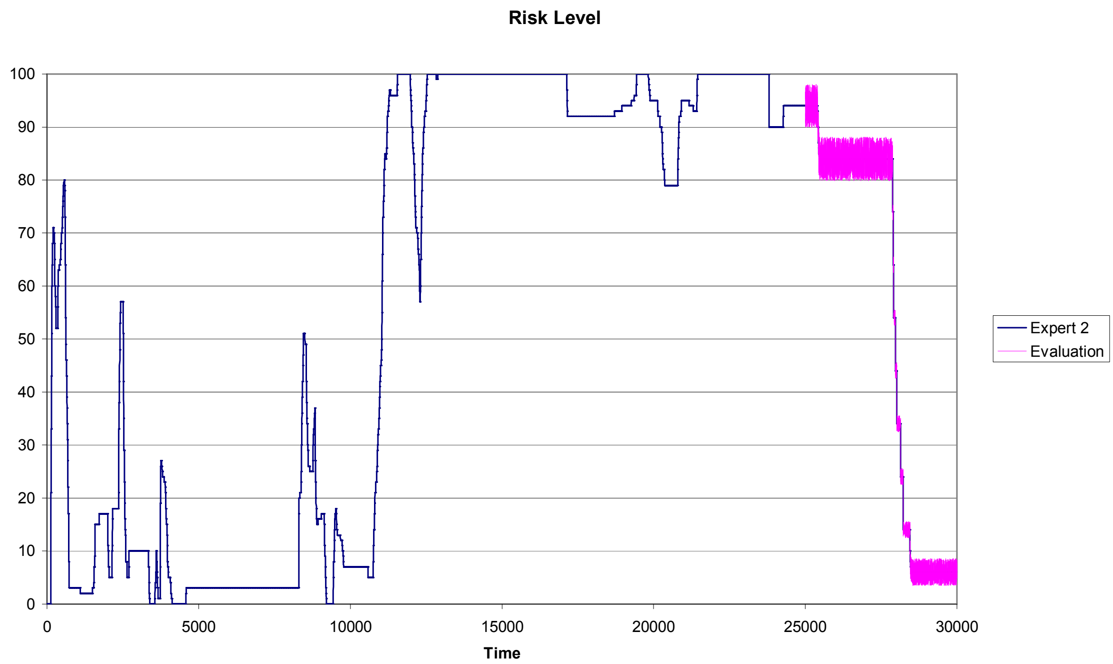


Figure 3.14. Driving risk evaluation given by expert 2.

n = 25000 and Tn = 5000

In the following table a comparison of results for expert 2 are presented:

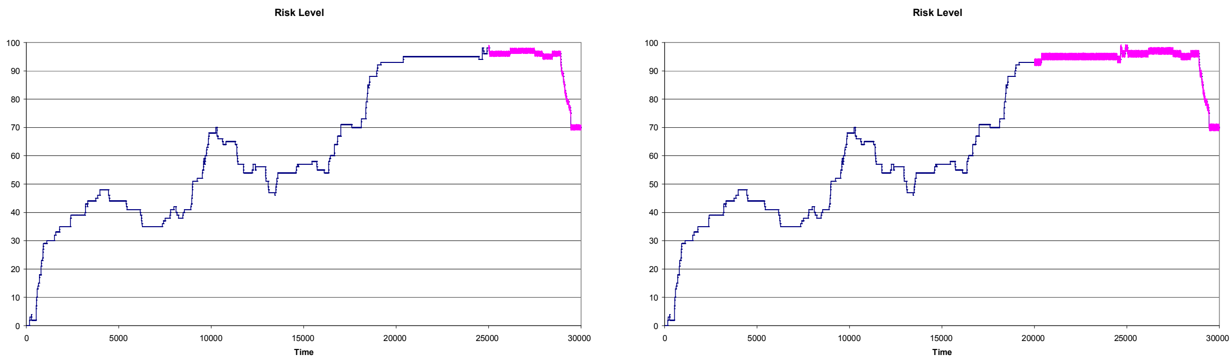
	n = 25000	n = 20000	n = 15000	n = 10000
max (absolute error)	1.262	2.146	4.116	5.635
mean squared error	0.366	0.715	2.727	5.854

Table 3.3. Results for the expert 2 with n=25000; n=20000; n=15000 and n=10000.

3.5.5 Results for the experts 3

In the figures 3.15, 3.16 and 3.17 we can see the results for the experts 2.

Figures 3.15 show 4 results with several training subset size between n=25000 and n=10000.



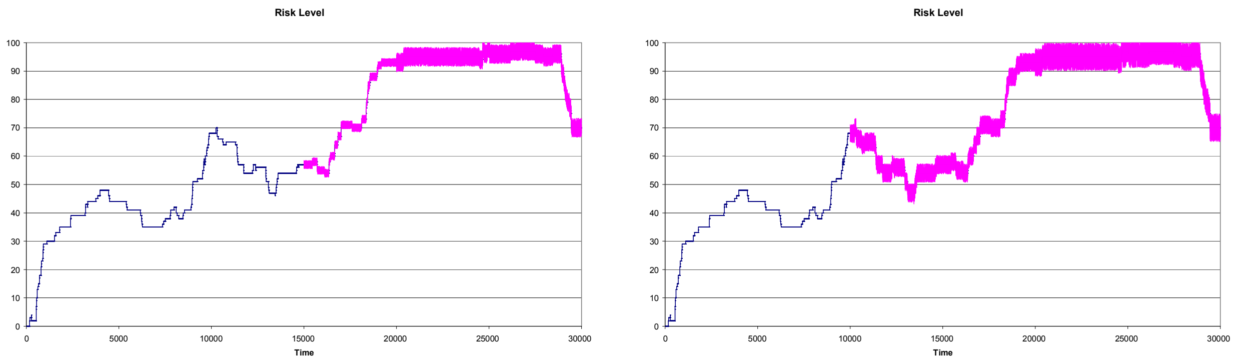


Figure 3.15. Driving risk evaluation given by expert 3 (N=25000, 20000, 15000 and 10000).

Figures 3.16 and 3.17 shows, again, the result for n=20000 and n=25000.

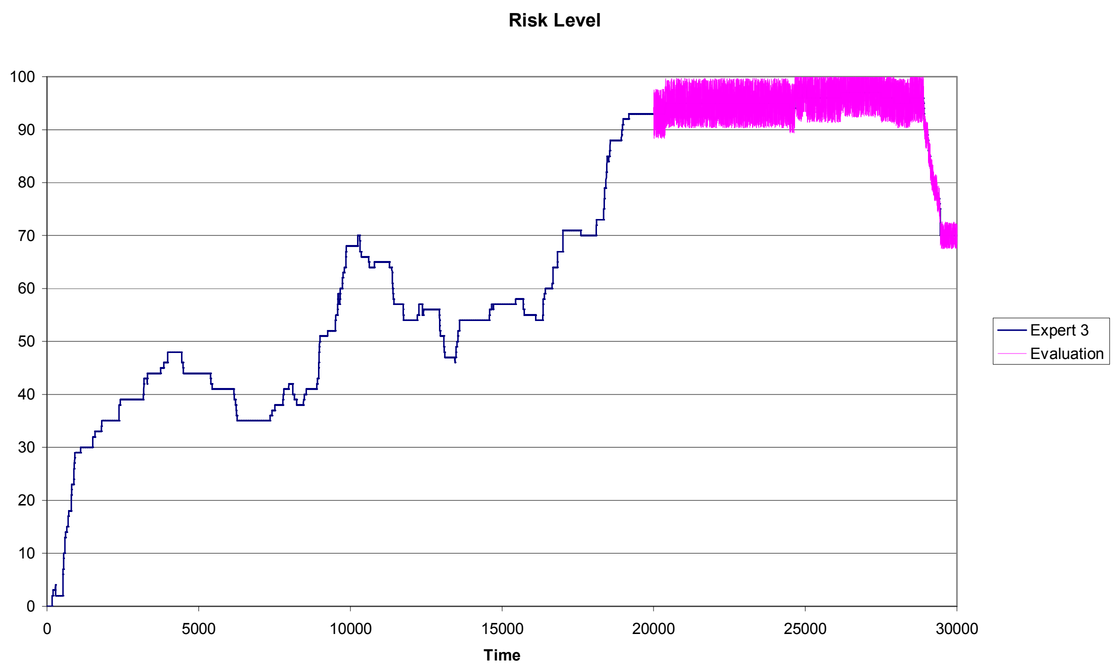


Fig. 3.16. Driving risk evaluation given by expert 3.
n = 20000 and Tn = 10000

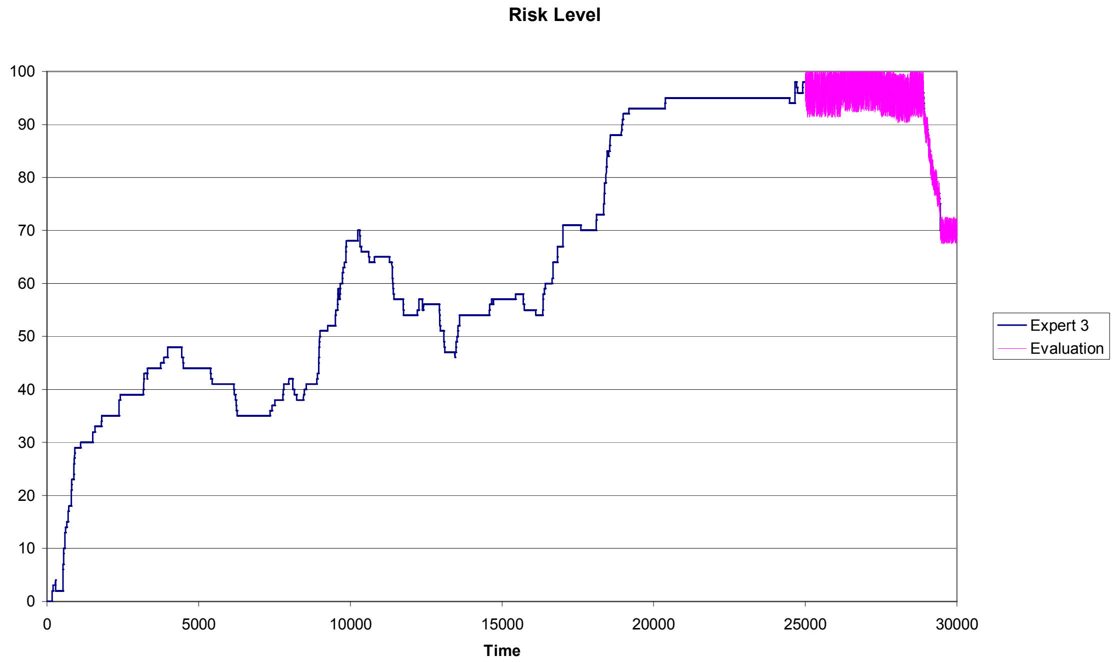


Figure 3.17. Driving risk evaluation given by expert 3.
n = 25000 and Tn = 5000

In the following table a comparison of results for expert 3 are presented:

	n = 25000	n = 20000	n = 15000	n = 10000
max (absolute error)	1.644	1.436	3.626	5.143
mean squared error	0.382	0.598	2.917	5.911

Table 3.4. Results for the expert 3 with n=25000; n=20000; n=15000 and n=10000.

3.6. Classification of the Experts

We have applied ALR methodology to CABINTEC data and we have found several results, we have used this information to classify the experts depending on their results.

To classify the experts, we have defined a following variable:

$$Ranking_i = \frac{\frac{1}{MCE_i}}{\sum_{i=1}^{3}\left(\frac{1}{MCE_i}\right)}$$

In the following table we can see the results for each expert and each training subset sizes:

	n = 25000	n = 20000	n = 15000	n = 10000
Expert 1	34.05%	37.10%	67.17%	80.07%
Expert 2	33.68%	28.65%	16.97%	10.01%
Expert 3	32.27%	34.25%	15.86%	9.92%

Table 3.5. Posterior weights for the experts

This variable can be seen like a posterior weight for the experts.

For n=25000 each expert “woks properly”, all of them have similar risk level values (posterior weight value 1/3 approximately).

For n=10000 expert 1 has better results (posterior weight over 0.8), expert 2 and expert 3 have extremely lower results.

3.7. Conclusions of Chapter 3

ALR methodology was applied to measure the quality of CABINTEC the “Intelligent cabin truck for road transport”.

We have a video from a truck simulator where several internal truck magnitudes were stored. Three experts were asked to evaluate the driving risk using a Visual Analog Scale (VAS). We have used the evaluation of the three experts to find the weights that they assign to every dimension of the simulation.

Results show that the risks correspond to abnormal behaviours of the driver and these risks are related to variables as speed and the angle of the steering wheel (SWA). Results were satisfactory.

We have several conclusions:

- Errors have a symmetric distribution with short tails.
- When training dataset size increase, then the errors decrease.
- Errors are “constant”, the algorithm does not learn.
- Probably errors are systematic errors, probably they are introduced by us during the perturbation of data.
- Previous experience in identical situations is not necessary.
- Extend ALR methodology to CABINTEC project dataset is simple conceptually, and does not present practical difficulties.
- Employing weighted Euclidean metric, our default approach, is enough to reach good results.
- ALR represents a smooth fitting for this kind of data. Visual trial and error method, for picking adaptive values was enough, but we have used the minimization of the mean squared error and crossvalidation for it.

As future work, we will apply our methodology to additional videos, experts and scenarios. We can try to find new specific definitions of risk, and of neighbourhoods for this kind of data (Stork 2004).

Chapter 4

Real Case: Measuring the Quality of Postgraduate Education

ALR methodology was applied to measure the quality of the education at a postgraduate department of a Public Spanish University.

We have used a real data set collected from a postgraduate program in a business school. We have used those data to find the weights that students assign to every dimension of the "service" (Perez, 2009; Ho, 2006; Bayo, 2003).

Our aim is to show that ALR methodology is able to treat this kind of data.

4.1. Introduction

The education is passing through a period of re-organization and re-establishment of new principles. Both at European and National levels, the issue of educational services quality is brought forward, taking into consideration the fact that universities are approached as socio-economical entities which objective is the survival in a competitive environment. Quality tools become a good option in this situation (Chua, 2004; Abdullah, 2005; Temponi, 2005).

The attention being devoted to the measurement and evaluation of the quality of postgraduate programs, particularly of Masters programs, and of students' satisfaction with these programs, is quite a new fact [Irons, 1994; Dubas, Ghani and Strong, 1998; Marks, 2001; Martin and Bray, 1997; Colbert, Levary and Shaner, 2000; Naik, 2003; Julia, 2004]. As both generic Masters programs and more specialized programs grow as a proportion of such programs in the education market, it has become increasingly important that they are evaluated for quality (see Lado, Cardone and Rivera 2003).

Masters programs must meet the demands of both students and the companies that employ graduates of the programs (stakeholders). Education and training are services provided to the student, which in turn is provided to the companies by the student (Cullen, 2003; Lomas, 2004). Therefore, the success of a program will depend on a large extent to its market orientation and on the quality and degree of satisfaction experienced by the student. Marketing research on quality of service and customer satisfaction is especially useful in this context (Rapert, Smith and Garretson, 2004; Kannan, 2005).

There are research studies that prove the applicability of:

- *factor analysis techniques* for analyzing the motivations of university students [Juric, Tood and Henry, 1997];
- *cluster analysis* to analyze student profiles [Stafford, 1994];
- *multidimensional scales* for evaluating performance in a faculty [Herche and Swenson, 1991];

- *conjoint analysis* to design the course offering Dubas and Strong, 1993]; analyses of repositioning of universities and of their Masters programs [Goldgehn and Kane, 1997; Comm and LaBay, 1996].

A dominant trend in education is based on the idea that students and their potential employers may be treated as market segments with expectations that educators must strive to know and satisfactorily meet [Anderson, Summey and Summey, 1991; Kotler et al 2003; Colbert, Levary and Shaner, 2000].

According to the European Foundation for Quality Management (EFQM, 2008), the following stakeholders can be distinguished within university education:

- *the corporate world* of potential employers;
- *families*, who contribute financial resources and demand security and information on the student's progress;
- *prospective students*, who need information on which to base their choices;
- *alumni*, who may require additional training and updating of their knowledge; and
- *society*, as a whole, which needs a competent labour force and free, educated citizens.

Therefore, it can be seen that when considering the concept of student-customer, there are diametrically opposed positions. Ritzer's approach (1996) considers the student comparable to any buyer who demands a good service, whilst Barret (1996) maintains that the final objective of education is never to satisfy the student-customer, since the person assuming this role does not know how to specify his or her needs, especially at the beginning of the degree course. Most of the remaining positions are in an intermediate area where the student is seen as a singular customer, an active participant in the process of his or her education.

The interest in quality in university education is not new (Morgan and Murgatroyd, 1994; Peña, 1997b). In the last years, as a reflection of the growing importance of quality in the corporate world and in academic research, corporate and academic concepts and methods have been extended to the public sector and university education (Sahney, Banwet and Karunes, 2004).

Experimental programs to encourage quality in university teaching are being advocated in the European Union, and in Spain the Ministry of Education and Science has promoted a program that is now in effect. According to Peña (1997b), these initiatives are based on the hypothesis that the perspective and methods of quality improvement in the business world are applicable to university teaching.

Having defined above the concept of the student as customer in order to measure its qualities, we must now consider the concept of product/service in university education (Dill, 1995). According to the EFQM (2000), the product is defined in terms of value added to the student's knowledge, skills and personal development (Shanahan and Gerber, 2004). As with the corporation, the quality of the product is linked to the quality of the process. Therefore, assessing the quality of the product in teaching entails analyzing the quality of the educational processes and identifying its key elements. The quality of the faculty is a particularly important key factor in Barnett's (1992) proposed integrative model.

4.2 Quality Of Service Models Applied To Teaching

Because the work of the teacher involves delivery of a service, there is a growing tendency to consider graduate teaching as a special case of customer service for which the theoretical fundamentals of service quality and MO are valid [Fernández and Mateo, 1992; Giacobbe and Segal, 1994; Stafford, 1994; Athiayaman, 1997; Coates and Koerner, 1996; Joseph and Joseph, 1997; Browne et al. 1993; Dubas et al., 1998; Lawton and Lundsten, 1998].

By applying this perspective, Giacobbe and Segal (1994) adapted the model for evaluating service quality put forward by Parasuraman, Zeithaml and

Berry (1985, 1988) to the case of teaching business management and administration at university level. To the original model that presented the relationship between two parties—the service provider and the service recipient—Giacobbe and Segal (1994) added a third element: the labour market, or potential employers who will evaluate the final product, knowledge [Wambsganss and Kennett, 1995]. The model they propose can be extended to social groups and institutions that also receive the services provided by the university.

Along the same lines, Colbert, Levary and Shaner (2000) measure and compare the efficiency of MBA programs by considering three outputs: the degree of student satisfaction, the degree of employer satisfaction and an index that jointly measures both levels of satisfaction (Chakrapani 1998).

Satisfaction and quality of service are two closely related concepts that have attracted the attention of many researchers. Despite this, no unanimous agreement has been reached regarding the sense of the relationship between the two concepts. While Parasuraman, et al., (1991a, b, c; 1994a, b) and Cronin and Taylor (1992) hold that the perception of quality of service is a determinant of satisfaction, Bolton and Drew (1991b) and Bitner (1990) find that satisfaction precedes a perception of quality.

When dealing with the perception of the quality of teaching service, Athiayaman (1997) asserts that perceived quality is a result of the student's satisfaction with the courses received. This author considers that the perception of quality influences the degree to which the student's expectations at the beginning of the degree course are met, or are not. Athiayaman's empirical results also indicate that the student's perceptions of quality strongly influence what the student communicates about the program to third parties. These results coincide with those found by Martin and Bray (1997) for the specific case of MBA programs.

Browne et al. (1998) and Guolla (1999) maintain that the perceived quality of the offering and of the university education services explain differences in levels of student satisfaction. Browne et al. (1998) use the SERVQUAL scale and various measures of satisfaction to empirically study the relationship between perceived quality and university students' satisfaction. In the opinion of Guolla (1999), satisfaction is a highly appropriate variable for measuring the quality of teaching, particularly when the student is observed in the role of customer.

Student satisfaction is generally measured by periodic surveys. While the use of surveys as instruments for measuring student evaluation of teaching (SET) has given rise to some controversy [Simpson and Siguaw, 2000] shows they are systematically used by 98% of universities and 99% of business schools in the United States. These authors report that teachers have perceived certain weaknesses in the surveys and have developed different practices to influence these evaluations.

Therefore, it is important to have and use complementary evaluative instruments (Ray and Jeon, 2003). Authors such as Murphy (1999) propose independent evaluation. The institutions that use this method delegate the task of making unannounced observation visits to classes to another teacher at the same level.

Despite the criticism levelled at the survey system, its utility as a measuring tool is widely recognized [Greenwald, 1997; McKeachie, 1997; Cashin and Downey 1992; Younker and Sterner, 1988; Guolla, 1999]. A review of the most widely used questionnaires can be found in Guolla (1999).

A revision of the literature on student evaluation of teachers can be found in Marsh (1987, 1991a, 1991b) and Joseph et al (2005). The conclusions suggest that SET are reliable, stable, multidimensional ("the same teacher may be enthusiastic but disorganized," Marsh, 1994), primarily a function of the teacher rather than the course being taught, relatively valid against different indicators of effective teaching, and perceived as being useful by faculty as feedback about their teaching activity. However, controversy continues over the relative importance of the different dimensions of SET, as well as whether multiple set dimensions should be applied for summative purposes, and if so, how they should be applied. Abrami and d'Apollonia (1991) defend the position that only global ratings or weighted averages of dimensions should be used for decisions about personnel. Marks (2000) alternatively suggests that “to average dissimilar items to generate an overall score is not appropriate.” Marsh and Roche (1997) argue that “for purposes of feedback to the teacher and personnel decisions it could be more useful to weight set factors according to their relative importance in a specific context.” Our article contributes to this debate since by analyzing the effects of different indicators on the student’s satisfaction, we provide a means to consider the importance of these aspects (see Lado, Cardone and Rivera 2003).

4.3 Preliminary Results

4.3.1 Sample

The data were obtained from surveys conducted from the Masters developed by the Department of Business Administration of a Spanish university. Information was gathered from questionnaires on all the subjects taught and all the teachers who taught the subjects.

Data from a survey carried out from 2003 to 2008, the unit of analysis was students of two master of a business school: MAE Master en Administración de Empresas (Spanish language version) and MBA Master of Business Administration (English language version). A total of 5769 questionnaires were administered, and the number of valid questionnaires received was 4372. The questionnaires considered valid were those in which the respondent had answered all of the questions of interest, yielding a full set of variables used in the subsequent analysis (Derek, 2000; Everitt, 2001; Johnson, 2002; Chambers, 2005).

Data from survey have been classified by years, terms and type of the subjects:

- **Years:** from 2003 to 2008.
- **Terms:** T1, T2 y T3.

- **Subject types:** 2, 1 and 0; qualitative, quantitative and mixed subject, respectively.

In the following tables we can see the evolution of received questionnaires since 2003 until 2008, separated by years, master and academic year terms.

Total (MBA and MAE)

YEAR	T1	T2	T3	Total
2003-04	537	537	672	**1746**
2004-05	484	398	482	**1364**
2005-06	213	225	265	**703**
2006-07	146	376	371	**893**
2007-08	289	370	404	**1063**
Total	**1669**	**1906**	**2194**	**5769**

Table 4.1 Evolution of received questionnaires

YEAR	MAE	MBA
2003-04	791	955
2004-05	431	933
2005-06	279	424
2006-07	518	375
2007-08	495	568
Total	**2514**	**3255**

Table 4.2 Evolution of received questionnaires, MAE and MBA

MAE

YEAR	T1	T2	T3	Total
2003-04	223	263	305	791
2004-05	140	144	147	431
2005-06	66	89	124	279
2006-07	83	211	224	518
2007-08	151	162	182	495
Total	**663**	**869**	**982**	**2514**

MBA

YEAR	T1	T2	T3	Total
2003-04	314	274	367	955
2004-05	344	254	335	933
2005-06	147	136	141	424
2006-07	63	165	147	375
2007-08	138	208	222	568
Total	**1006**	**1037**	**1212**	**3255**

Table 4.3 Evolution of received questionnaires, Terms, MAE and MBA.

The trend of the evolution of received questionnaires, in general, is negative.

4.3.2 Survey Instrument

The definitive questionnaire contained 12 questions that allowed us to measure the aspects detailed below:

P1. Interest: refers to the student's interest in the subject.
P2. Integration: integration degree of the subject in the master.
P3. Satisfaction with teacher: overall student satisfaction with the teacher.
P4. Clarity: the teacher teaches clearly.
P5. Punctuality: the teacher is on time.
P6. Prom Participation: the teacher promotes participation in class.
P7. Bibliography: the usefulness and interest of the readings and recommended bibliography.
P8. Utility: the usefulness of the teaching assistant practice lessons.
P9. Satisfaction Assistant: overall student satisfaction with the teaching assistant.
P10. Equilibrium: comparison between practice contents and theory contents.
P11. Output Level: Output level reached in the subject.
P12. Input Level: Input level previous to the subject.

All measures were registered on a 5-point Likert scale, ranging from "strongly disagree" to "strongly agree."

4.3.3 Aggregated Results

We have prepared several statistical analysis for the data, MAE and MBA together.

4.3.3.1 Descriptive results

In the following tables we can see descriptive aggregated results. Table 4.4 shows results by questions. Table 4.5 shows results by questions and subject type. Table 4.6 shows results by questions, subject type and master:

	Mean	Mode	Median
P1	3,95	5	4
P2	3,91	5	4
P3	3,69	4	4
P4	3,60	4	4
P5	4,25	5	5
P6	3,66	5	4
P7	3,43	4	4
P8	3,45	4	4
P9	3,20	4	4
P10	3,37	3	3
P11	3,52	4	4
P12	3,12	4	3

Table 4.4 Descriptive statistics for aggregated results

Globally speaking, for aggregated results, P5 Punctuality, have obtained the best result (maximum possible mode and median and mean over than 4.2 points).

		0	1	2
P1	**Mean**	4,15	4,05	4,22
	Median	4	4	4
P2	**Mean**	4,08	4,02	4,22
	Median	4	4	4
P3	**Mean**	3,91	3,83	4,00
	Median	4	4	4
P4	**Mean**	3,81	3,69	3,95
	Median	4	4	4
P5	**Mean**	4,47	4,45	4,42
	Median	5	5	5
P6	**Mean**	3,83	3,76	4,03
	Median	4	4	4
P7	**Mean**	3,70	3,65	3,79
	Median	4	4	4
P8	**Mean**	3,77	3,71	3,94
	Median	4	4	4
P9	**Mean**	3,84	3,69	3,91
	Median	4	4	4
P10	**Mean**	3,69	3,61	3,69
	Median	4	4	4
P11	**Mean**	3,78	3,66	3,88
	Median	4	4	4
P12	**Mean**	3,40	3,28	3,54
	Median	4	3	4

Table 4.5 Descriptive statistics for aggregated results, by type subjects

Globally speaking, for aggregated results by subject type, P5 Punctuality, have obtained the best result (maximum possible mode and median and mean over than 4.4 points).

		MAE			MBA		
		0	**1**	**2**	**0**	**1**	**2**
P1	**Mean**	4,08	4,30	4,34	4,19	3,86	4,10
	Median	4	4	5	4	4	4
P2	**Mean**	3,98	4,23	4,30	4,13	3,86	4,13
	Median	4	4	4	4	4	4
P3	**Mean**	3,65	4,05	4,07	4,05	3,66	3,92
	Median	4	4	4	4	4	4
P4	**Mean**	3,50	3,93	4,02	3,98	3,51	3,88
	Median	4	4	4	4	4	4
P5	**Mean**	4,48	4,50	4,47	4,47	4,42	4,41
	Median	5	5	5	5	5	5
P6	**Mean**	3,62	4,05	4,08	3,93	3,54	3,99
	Median	4	4	4	4	4	4
P7	**Mean**	3,49	3,91	3,86	3,80	3,47	3,72
	Median	4	4	4	4	4	4
P8	**Mean**	3,41	3,97	4,03	3,95	3,52	3,85
	Median	4	4	4	4	4	4
P9	**Mean**	3,55	3,96	4,01	3,98	3,49	3,81
	Median	4	4	4	4	4	4
P10	**Mean**	3,64	3,75	3,76	3,72	3,50	3,62
	Median	4	4	4	3	3	3
P11	**Mean**	3,45	3,88	3,94	3,95	3,49	3,82
	Median	4	4	4	4	4	4
P12	**Mean**	3,21	3,42	3,49	3,49	3,18	3,60
	Median	3	4	4	4	3	4

Table 4.6 Descriptive statistics, by master and type subjects

Globally speaking, for aggregated results by subject type and masters, P5 Punctuality, have obtained the best result (maximum possible mode and median and mean over than 4.1 points).

In the following figures we can see aggregated results, MAE and MBA together. Figure 4.1 shows results by questions (equivalent to Table 4.4). Figure 4.2 shows arithmetic mean results by questions (equivalent to second column of Table 4.4):

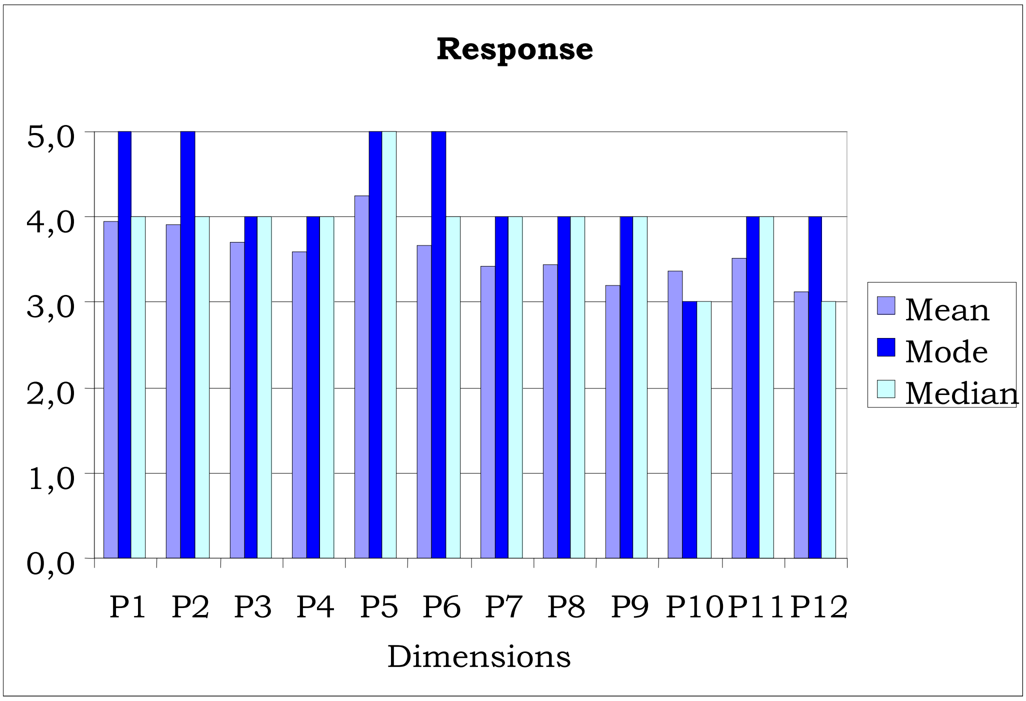


Figure 4.1 Descriptive statistics for aggregated results

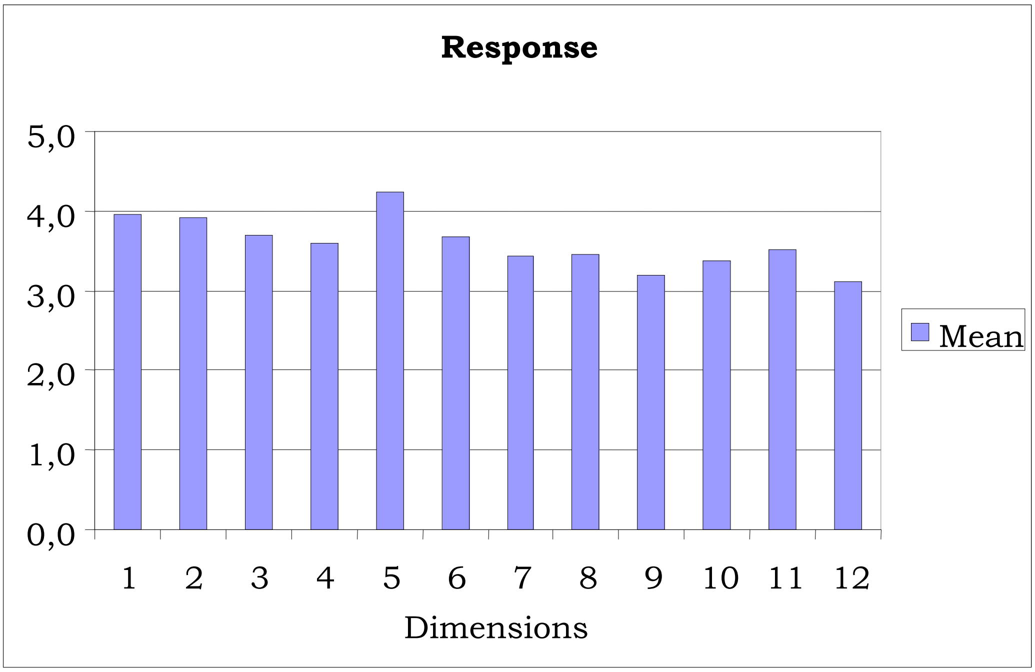


Figure 4.2 Arithmetic Mean for aggregated results, by quality dimensions

In the following table we can see aggregated results, MAE and MBA together by options of the Likert scale:

	0	**1**	**2**	**3**	**4**	**5**
P1	3,7%	2,2%	4,0%	15,2%	35,3%	39,7%
P2	4,2%	1,7%	4,3%	16,1%	35,8%	37,8%
P3	4,4%	3,4%	6,9%	20,0%	34,0%	31,2%
P4	4,4%	4,7%	9,0%	20,6%	30,9%	30,3%
P5	4,6%	0,9%	2,7%	8,9%	23,0%	59,9%
P6	4,7%	3,6%	7,4%	20,9%	31,7%	31,7%
P7	7,1%	3,1%	9,2%	24,5%	33,1%	23,0%
P8	8,6%	4,2%	8,0%	19,6%	31,9%	27,7%
P9	14,9%	4,3%	6,8%	19,0%	30,1%	24,9%
P10	6,9%	4,0%	5,1%	37,6%	22,1%	24,2%
P11	5,2%	3,0%	6,6%	24,3%	41,7%	19,2%
P12	6,0%	8,1%	12,5%	28,3%	31,3%	13,8%

Table 4.7 Table of descriptive percentages for aggregated results

Globally speaking, for aggregated results by options of the Likert scale, P5 Punctuality, have obtained the best result (approximately 60% of the questionnaires with maximum possible opinion).

4.3.3.2 Hierarchical cluster analysis

We have done a correlation analysis and a hierarchical cluster analysis by variables.

In the following table we can see the correlation analysis results:

		P1	P2	P4	P5	P6	P7	P8	P9	P10	P11	P12	P3
P1	Correlation	1	**,733***	,635*	,594*	,401*	,530*	,526*	,554*	,518*	,275*	,602*	,343*
	Sig. (2-tailed)	.	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
P2	Correlation	,733*	1	,599*	,572*	,432*	,515*	,522*	,568*	,489*	,272*	,609*	,320*
	Sig. (2-tailed)	,000	.	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
P4	Correlation	,635*	,599*	1	**,819***	,461*	,674*	,614*	,652*	,688*	,309*	,680*	,347*
	Sig. (2-tailed)	,000	,000	.	,000	,000	,000	,000	,000	,000	,000	,000	,000
P5	Correlation	,594*	,572*	,819*	1	,432*	,670*	,608*	,669*	,684*	,312*	,683*	,362*
	Sig. (2-tailed)	,000	,000	,000	.	,000	,000	,000	,000	,000	,000	,000	,000
P6	Correlation	,401*	,432*	,461*	,432*	1	,441*	,401*	,394*	,394*	,227*	,403*	,166*
	Sig. (2-tailed)	,000	,000	,000	,000	.	,000	,000	,000	,000	,000	,000	,000
P7	Correlation	,530*	,515*	,674*	,670*	,441*	1	,595*	,618*	,608*	,274*	,599*	,311*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	.	,000	,000	,000	,000	,000	,000
P8	Correlation	,526*	,522*	,614*	,608*	,401*	,595*	1	,622*	,589*	,310*	,620*	,372*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	.	,000	,000	,000	,000	,000
P9	Correlation	,554*	,568*	,652*	,669*	,394*	,618*	,622*	1	**,752***	,342*	,682*	,383*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	.	,000	,000	,000	,000
P10	Correlation	,518*	,489*	,688*	,684*	,394*	,608*	,589*	,752*	1	,354*	,635*	,350*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	.	,000	,000	,000
P11	Correlation	,275*	,272*	,309*	,312*	,227*	,274*	,310*	,342*	,354*	1	,365*	,256*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	.	,000	,000
P12	Correlation	,602*	,609*	,680*	,683*	,403*	,599*	,620*	,682*	,635*	,365*	1	,539*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	.	,000
P3	Correlation	,343*	,320*	,347*	,362*	,166*	,311*	,372*	,383*	,350*	,256*	,539*	1
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	.

Table 4.8 Correlation analysis for aggregated results

* Correlation is significant at the 0.01 level (2-tailed).

The relationship between P4-P5, P9-P10 and P1-P2, respectively present the highest values.

In the following table we can see the stages of the agglomeration schedule results:

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	3	4	,000	0	0	4
2	1	2	,013	0	0	8
3	10	11	,064	0	0	5
4	3	8	,099	1	0	5
5	3	10	,105	4	3	6
6	3	6	,128	5	0	7
7	3	9	,141	6	0	8
8	1	3	,147	2	7	9
9	1	5	,253	8	0	10
10	1	12	,303	9	0	11
11	1	7	,373	10	0	0

Table 4.9 Agglomeration Schedule for aggregated results

In the following figure we can see the dendrogram for aggregated data:

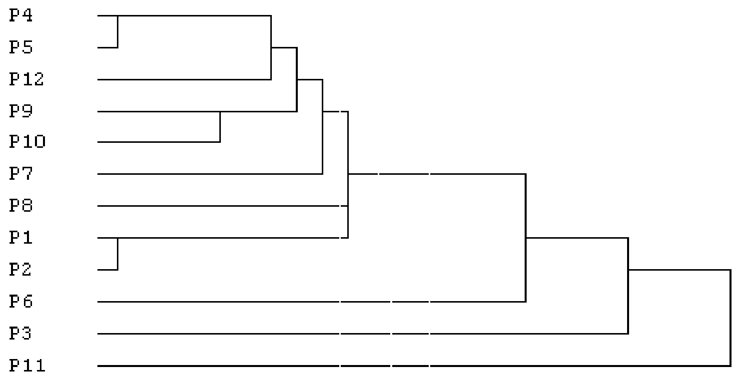


Figure 4.3 Dendrogram for aggregated data

Dendrogram shows strong similarity between:

- P4 - P5
  P4 *Clarity* and P5 *Punctuality* of the teacher.

- P1 – P2
  P1 *Interest* in the subject and P2 *Integration degree* of the subject in the master.

- P9 – P10
  P9 *Satisfaction with the teaching assistant* and P10 *Equilibrium* between practice contents and theory contents.

Globally speaking, for aggregated results, students have a mature opinion referent to the subjects. The first cluster measures “Profesionality/Expertise” of the teacher. The second one measures “previous attitude of the student”. The third one measures “the work of the teaching assistant”.

4.3.4 MAE Results

We have prepared several statistical analysis for the MAE data.

4.3.4.1 Descriptive results

In the following tables we can see descriptive results. Table 4.10 shows MAE results by questions.

	Mean	**Mode**	**Median**
P1	4,15	5	4
P2	4,07	5	4
P3	3,84	5	4
P4	3,75	5	4
P5	4,32	5	5
P6	3,85	5	4
P7	3,60	4	4
P8	3,64	4	4
P9	3,36	4	4
P10	3,48	3	4
P11	3,66	4	4
P12	3,18	4	3

Table 4.10 Descriptive statistics for MAE data

For MAE data, P5 Punctuality, have obtained the best result (maximum possible mode and median and mean over than 4.3).

In the following figures we can see descriptive results of MAE data. Figure 4.4 shows results by questions (equivalent to Table 4.10). Figure 4.5 shows arithmetic mean results by questions (equivalent to second column of Table 4.10):

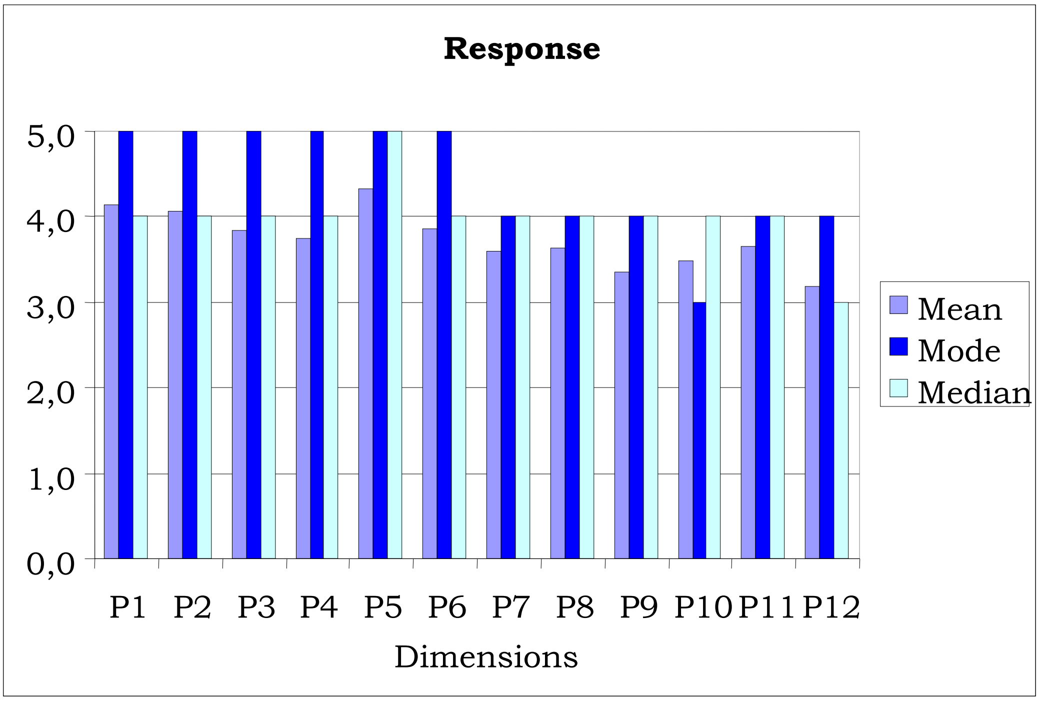


Figure 4.4 Descriptive statistics for MAE data

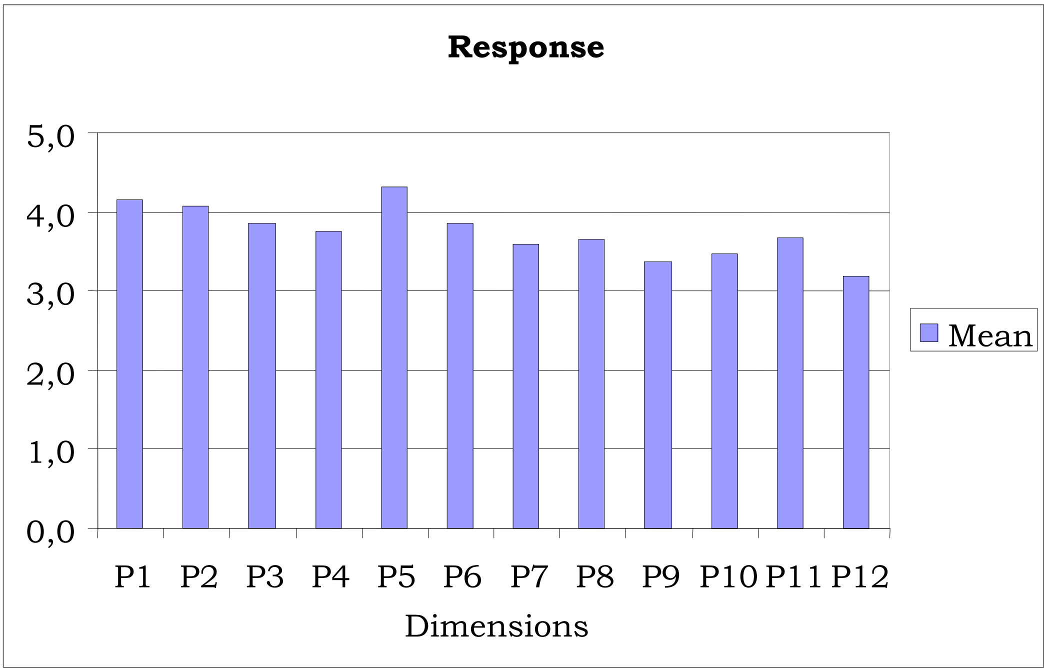


Figure 4.5 Arithmetic Mean for MAE master variables, by quality dimensions

In the following table we can see descriptive results of MAE data by options of the Likert scale:

	0	**1**	**2**	**3**	**4**	**5**
P1	3,6%	0,8%	2,8%	10,5%	35,2%	47,2%
P2	3,9%	0,9%	3,2%	12,1%	36,7%	43,3%
P3	4,0%	2,3%	5,4%	18,4%	33,5%	36,4%
P4	4,1%	2,8%	7,9%	18,9%	32,4%	33,9%
P5	4,1%	0,8%	2,0%	8,0%	22,2%	**62,8%**
P6	4,2%	2,7%	4,9%	17,8%	32,2%	38,1%
P7	6,4%	1,9%	7,4%	22,6%	33,0%	28,6%
P8	7,1%	2,7%	6,8%	18,2%	33,1%	32,1%
P9	14,3%	2,5%	5,5%	17,6%	30,7%	29,4%
P10	6,1%	3,3%	4,6%	35,4%	24,3%	26,3%
P11	4,6%	1,9%	5,4%	22,0%	43,2%	22,9%
P12	5,1%	8,5%	11,6%	28,3%	31,3%	15,3%

Table 4.11 Table of descriptive percentages for MAE master

For MAE data results by options of the Likert scale, P5 Punctuality, have obtained the best result (more than 60% of the questionnaires with maximum possible opinion).

4.3.4.2 Hierarchical cluster analysis

We have done a correlation analysis and a hierarchical cluster analysis by variables.

In the following table we can see the correlation analysis results:

		P1	P2	P4	P5	P6	P7	P8	P9	P10	P11	P12	P3
P1	Correlation	1	**,715***	,549*	,441*	,496*	,503*	,557*	,535*	,324*	,543*	,234*	,592*
	Sig. (2-tailed)	.	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
P2	Correlation	,715*	1	,540*	,494*	,509*	,492*	,571*	,510*	,309*	,540*	,207*	,579*
	Sig. (2-tailed)	,000	.	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
P4	Correlation	,549*	,540*	1	,459*	,678*	,593*	,712*	,746*	,354*	,684*	,307*	**,816***
	Sig. (2-tailed)	,000	,000	.	,000	,000	,000	,000	,000	,000	,000	,000	,000
P5	Correlation	,441*	,494*	,459*	1	,485*	,437*	,454*	,460*	,253*	,413*	,130*	,501*
	Sig. (2-tailed)	,000	,000	,000	.	,000	,000	,000	,000	,000	,000	,000	,000
P6	Correlation	,496*	,509*	,678*	,485*	1	,570*	,660*	,664*	,291*	,609*	,271*	,682*
	Sig. (2-tailed)	,000	,000	,000	,000	.	,000	,000	,000	,000	,000	,000	,000
P7	Correlation	,503*	,492*	,593*	,437*	,570*	1	,630*	,611*	,382*	,605*	,308*	,613*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	.	,000	,000	,000	,000	,000	,000
P8	Correlation	,557*	,571*	,712*	,454*	,660*	,630*	1	,746*	,385*	,715*	,326*	,701*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	.	,000	,000	,000	,000	,000
P9	Correlation	,535*	,510*	,746*	,460*	,664*	,611*	,746*	1	,373*	,677*	,298*	,758*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	.	,000	,000	,000	,000
P10	Correlation	,324*	,309*	,354*	,253*	,291*	,382*	,385*	,373*	1	,415*	,290*	,363*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	.	,000	,000	,000
P11	Correlation	,543*	,540*	,684*	,413*	,609*	,605*	,715*	,677*	,415*	1	,464*	,681*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	.	,000	,000
P12	Correlation	,234*	,207*	,307*	,130*	,271*	,308*	,326*	,298*	,290*	,464*	1	,303*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	.	,000
P3	Correlation	,592*	,579*	,816*	,501*	,682*	,613*	,701*	,758*	,363*	,681*	,303*	1
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	.

Table 4.12 Correlation analysis for MAE data

* Correlation is significant at the 0.01 level (2-tailed).

The relationship between P3-P4 and P1-P2, respectively, present the highest values.

In the following table we can see the stages of the agglomeration schedule results:

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	3	4	770,783	0	0	3
2	1	2	804,492	0	0	8
3	3	9	966,192	1	0	4
4	3	8	1027,929	3	0	5
5	3	11	1077,181	4	0	6
6	3	6	1257,131	5	0	7
7	3	7	1428,860	6	0	9
8	1	5	1495,494	2	0	9
9	1	3	1504,768	8	7	10
10	1	10	2232,973	9	0	11
11	1	12	2641,905	10	0	0

Table 4.13 Agglomeration Schedule for MAE data

In the following figure we can see the dendrogram for MAE data:

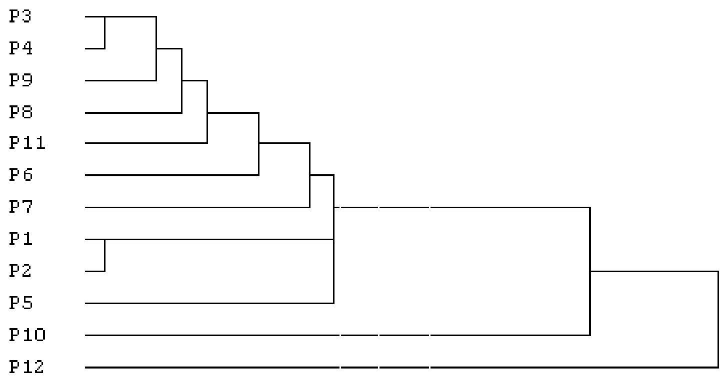


Figure 4.6 Dendrogram for MAE data

Dendrogram shows strong similarity between:

- P3 – P4
  P3 *Satisfaction with teacher* and P4 *Clarity*.

- P1 – P2
  P1 *Interest* in the subject and P2 *Integration degree* of the subject in the master.

For MAE data, students have a mature opinion referent to the subjects. The first cluster measures “Quality” of the teacher. The second one measures “previous attitude of the student”.

4.3.5 MBA Results

We have prepared several statistical analysis for the MBA data.

4.3.5.1 Descriptive results

In the following tables we can see descriptive results. Table 4.14 shows MBA results by questions.

	Mean	**Mode**	**Median**
P1	3,81	4	4
P2	3,79	4	4
P3	3,58	4	4
P4	3,48	4	4
P5	4,19	5	5
P6	3,52	4	4
P7	3,29	4	4
P8	3,30	4	4
P9	3,07	4	4
P10	3,28	3	3
P11	3,41	4	4
P12	3,08	4	3

Table 4.14 Descriptive statistics for MBA data

For MBA data, P5 Punctuality, have obtained the best result (maximum possible mode and median and mean over than 4.1).

In the following figures we can see descriptive results of MBA data. Figure 4.7 shows results by questions (equivalent to Table 4.14). Figure 4.8 shows arithmetic mean results by questions (equivalent to second column of Table 4.14):

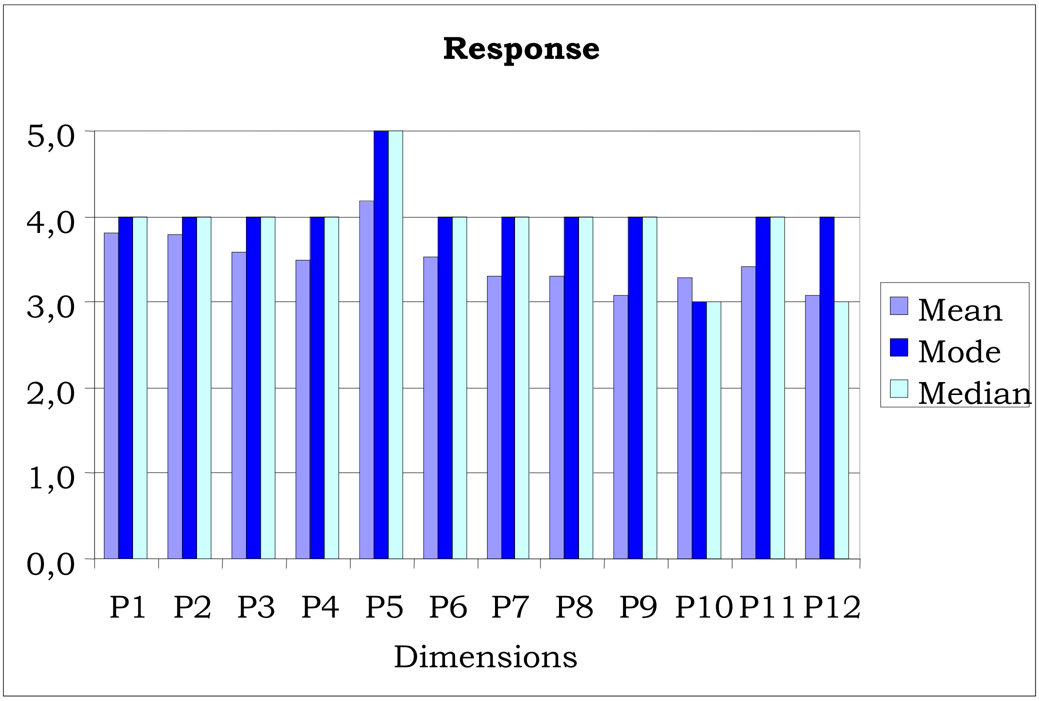


Figure 4.7 Descriptive statistics for MBA data

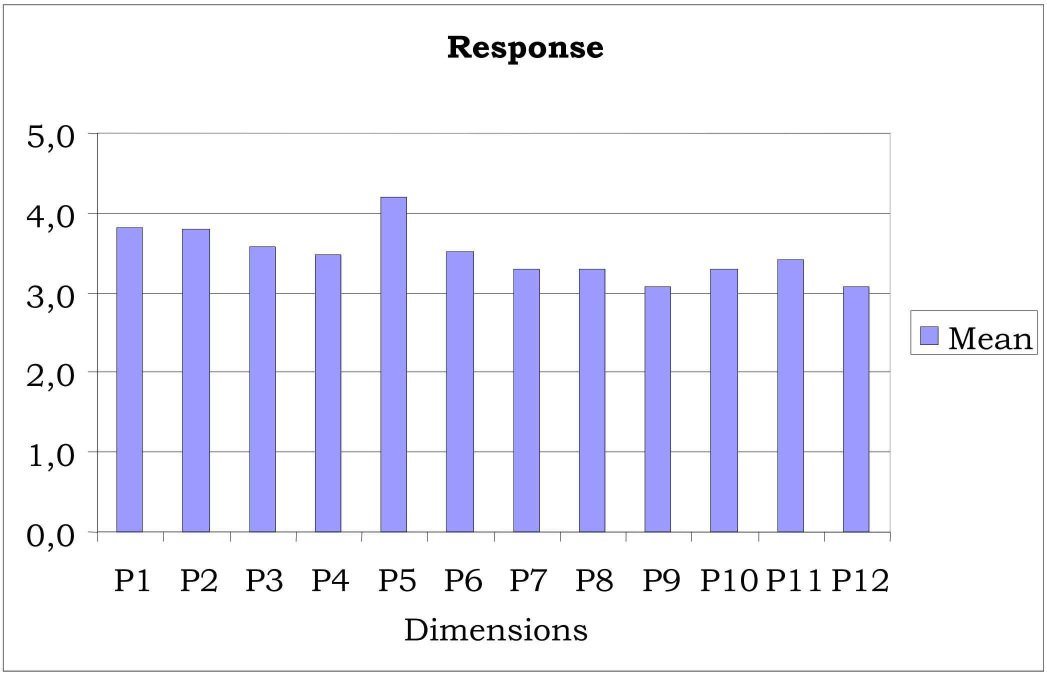


Figure 4.8 Arithmetic Mean for MBA master variables, by quality dimensions

In the following table we can see descriptive results of MBA data by options of the Likert scale:

	0	**1**	**2**	**3**	**4**	**5**
P1	3,7%	3,3%	4,9%	18,8%	35,4%	33,9%
P2	4,5%	2,3%	5,2%	19,3%	35,1%	33,6%
P3	4,8%	4,3%	8,1%	21,3%	34,4%	27,2%
P4	4,7%	6,2%	9,9%	22,0%	29,7%	27,5%
P5	5,0%	1,0%	3,2%	9,6%	23,6%	57,7%
P6	5,0%	4,3%	9,3%	23,4%	31,2%	26,8%
P7	7,6%	4,0%	10,6%	26,0%	33,1%	18,7%
P8	9,8%	5,5%	8,9%	20,7%	30,9%	24,2%
P9	15,4%	5,7%	7,7%	20,2%	29,6%	21,5%
P10	7,5%	4,6%	5,5%	39,3%	20,5%	22,6%
P11	5,6%	3,8%	7,5%	26,2%	40,6%	16,3%
P12	6,7%	7,8%	13,2%	28,2%	31,4%	12,6%

Table 4.15 Table of descriptive percentages for MBA data

For MBA data results by options of the Likert scale, P5 Punctuality, have obtained the best result (more than 57% of the questionnaires with maximum possible opinion).

4.3.5.2 Hierarchical cluster analysis

We have done a correlation analysis and a hierarchical cluster analysis by variables.

In the following table we can see the correlation analysis results:

		P1	P2	P4	P5	P6	P7	P8	P9	P10	P11	P12	P3
P1	Correlation	1	**,734***	,611*	,374*	,533*	,525*	,537*	,490*	,228*	,628*	,417*	,651*
	Sig. (2-tailed)	.	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
P2	Correlation	,734*	1	,581*	,389*	,503*	,527*	,553*	,459*	,234*	,643*	,400*	,601*
	Sig. (2-tailed)	,000	.	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000
P4	Correlation	,611*	,581*	1	,411*	,657*	,608*	,632*	,635*	,271*	,675*	,401*	**,817***
	Sig. (2-tailed)	,000	,000	.	,000	,000	,000	,000	,000	,000	,000	,000	,000
P5	Correlation	,374*	,389*	,411*	1	,407*	,372*	,350*	,346*	,202*	,393*	,191*	,430*
	Sig. (2-tailed)	,000	,000	,000	.	,000	,000	,000	,000	,000	,000	,000	,000
P6	Correlation	,533*	,503*	,657*	,407*	1	,599*	,578*	,558*	,246*	,581*	,339*	,660*
	Sig. (2-tailed)	,000	,000	,000	,000	.	,000	,000	,000	,000	,000	,000	,000
P7	Correlation	,525*	,527*	,608*	,372*	,599*	1	,606*	,561*	,243*	,620*	,419*	,605*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	.	,000	,000	,000	,000	,000	,000
P8	Correlation	,537*	,553*	,632*	,350*	,578*	,606*	1	**,749***	,300*	,651*	,424*	,611*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	.	,000	,000	,000	,000	,000
P9	Correlation	,490*	,459*	,635*	,346*	,558*	,561*	,749*	1	,328*	,597*	,387*	,634*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	.	,000	,000	,000	,000
P10	Correlation	,228*	,234*	,271*	,202*	,246*	,243*	,300*	,328*	1	,317*	,224*	,259*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	.	,000	,000	,000
P11	Correlation	,628*	,643*	,675*	,393*	,581*	,620*	,651*	,597*	,317*	1	,596*	,672*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	.	,000	,000
P12	Correlation	,417*	,400*	,401*	,191*	,339*	,419*	,424*	,387*	,224*	,596*	1	,378*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	.	,000
P3	Correlation	,651*	,601*	,817*	,430*	,660*	,605*	,611*	,634*	,259*	,672*	,378*	1
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	.

Table 4.16 Correlation analysis for MBA data

* Correlation is significant at the 0.01 level (2-tailed).

The relationship between P3-P4, P8-P9 and P1-P2, respectively, present the highest values.

In the following table we can see the stages of the agglomeration schedule results:

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	3	9	,000	0	0	4
2	1	2	,019	0	0	6
3	10	11	,072	0	0	5
4	3	7	,108	1	0	5
5	3	10	,139	4	3	6
6	1	3	,139	2	5	7
7	1	6	,141	6	0	8
8	1	5	,149	7	0	9
9	1	8	,221	8	0	10
10	1	4	,339	9	0	11
11	1	12	,428	10	0	0

Table 4.17 Agglomeration Schedule for MBA data

In the following figure we can see the dendrogram for MBA data:

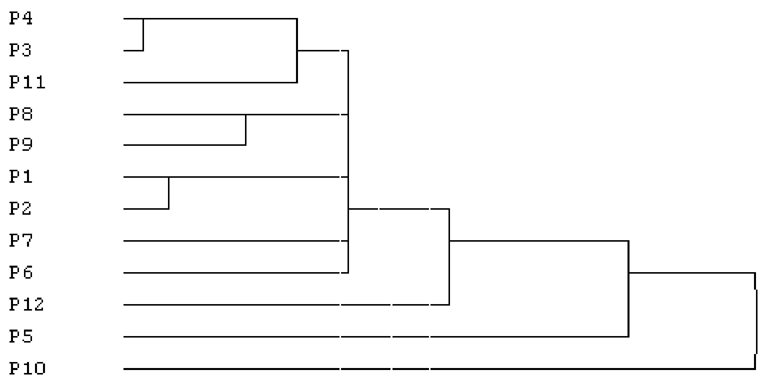


Figure 4.9 Dendrogram for MBA data

Dendrogram shows strong similarity between:

- P3 – P4

P3 *Satisfaction with teacher* and P4 *Clarity*.

- P1 – P2
  P1 *Interest* in the subject and P2 *Integration degree* of the subject in the master.

- P8 – P9
  P8 *Usefulness of the teaching assistant practice lessons* and P9 *Satisfaction with the teaching assistant*.

For MBA data, students have a mature opinion referent to the subjects. The first cluster measures "Quality" of the teacher. The second one measures "previous attitude of the student". The third one measures "the work of the teaching assistant".

4.4 Results

In the following section we present the results of applying our methodology to the data set. The goal is to determine the relative importance of each explicative variable (P1, P2, P4, P5, P6, P7, P11 y P12) to explain the response variable "P3. Satisfaction with teacher". Variables related to "Teaching assistant" (P8, P9, P10) have been eliminated of the analysis.

Results of weight variables for complete data set, each master and each type subject are presented.

4.4.1 Aggregated results

In the following figure we can see the results of weights in the complete data set, MAE and MBA courses together:

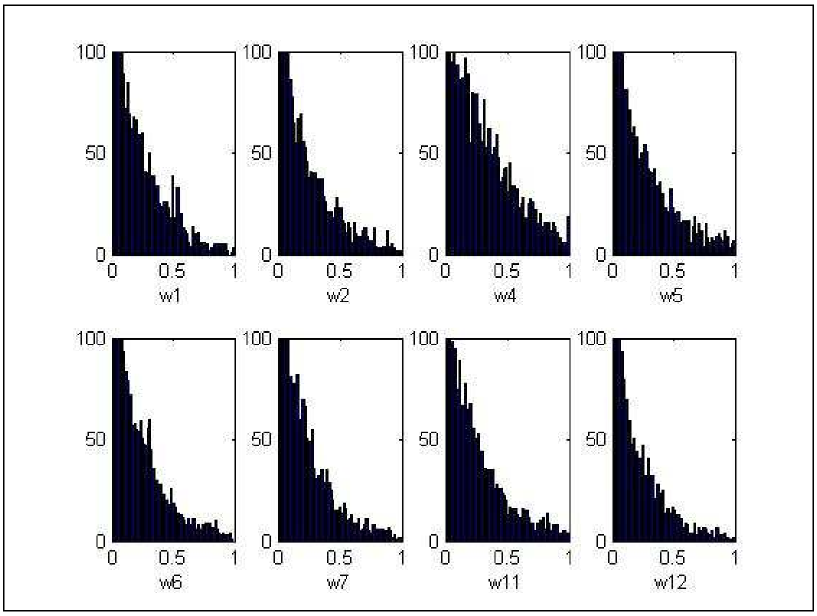


Figure 4.10 Weights for aggregated data set

Weights present a big variability, ranging from 0 to 1 and standard deviation equal to 0.199. The multivariate analysis has not detected any clusters (Mardia 1979, Gordon 1989), both masters show a similar behaviour.

4.4.1.1 Hierarchical cluster analysis

We have done a correlation analysis and a hierarchical cluster analysis by weights variables.

In the following table we can see the correlation analysis results:

		W_1	W_2	W_4	W_5	W_6	W_7	W_11	W_12
W_1	Correlation	1	-,046*	-,233*	-,083*	-,143*	-,126*	-,156*	-,128*
	Sig. (2-tailed)	.	,002	,000	,000	,000	,000	,000	,000
W_2	Correlation	-,046*	1	-,217*	-,136*	-,132*	-,145*	-,149*	-,102*
	Sig. (2-tailed)	,002	.	,000	,000	,000	,000	,000	,000
W_4	Correlation	-,233*	-,217*	1	-,198*	-,136*	-,166*	-,193*	-,162*
	Sig. (2-tailed)	,000	,000	.	,000	,000	,000	,000	,000
W_5	Correlation	-,083*	-,136*	-,198*	1	-,157*	-,164*	-,197*	-,156*
	Sig. (2-tailed)	,000	,000	,000	.	,000	,000	,000	,000
W_6	Correlation	-,143*	-,132*	-,136*	-,157*	1	-,122*	-,132*	-,119*
	Sig. (2-tailed)	,000	,000	,000	,000	.	,000	,000	,000
W_7	Correlation	-,126*	-,145*	-,166*	-,164*	-,122*	1	-,102*	-,058*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	.	,000	,000
W_11	Correlation	-,156*	-,149*	-,193*	-,197*	-,132*	-,102*	1	-,067*

		W_1	W_2	W_4	W_5	W_6	W_7	W_11	W_12
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	.	,000
W_12	Correlation	-,128*	-,102*	-,162*	-,156*	-,119*	-,058*	-,067*	1
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	.

Table 4.18 Correlation analysis for weights for aggregated data

* Correlation is significant at the 0.01 level (2-tailed).

Figure above does not show important values, remember that correlation coefficient only detect linear relationship between variables.

In the following table we can see the stages of the agglomeration schedule results:

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	6	8	,000	0	0	2
2	2	6	,087	0	1	3
3	2	5	,096	2	0	4
4	1	2	,099	0	3	5
5	1	7	,143	4	0	6
6	1	4	,325	5	0	7
7	1	3	,765	6	0	0

Table 4.19 Agglomeration schedule for weights variables

In the following figure we can see the dendrogram for weights variables for aggregated data:

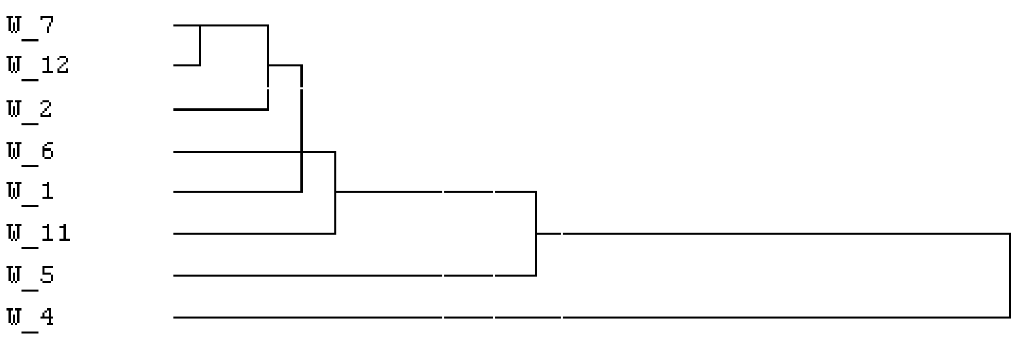


Figure 4.11 Dendrogram for weights variables

Dendrogram shows strong similarity between:

- W7 – W12

  W7 The usefulness and interest of the *readings and recommended bibliography*.

  W12 *Input level previous to the subject*.

We can say, that the cluster measures "Didactic materials" (in reference to the level of student).

4.4.1.2 Factorial analysis for P's and W's

We have done a factorial analysis for the explicative variables (P) and the weights variables (W). Explicative variables (P) were scaled up to one.

In the following table we can see the factorial analysis results:

Component*	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	353,541	88,385	88,385	229,766	57,441	57,441
2	10,430	2,607	90,993	134,205	33,551	90,993

Table 4.20 Total Variance explained

* Extraction Method: Principal Component Analysis.

Two factors have been detected, they explained 90,993% of the total variance. Factor 1 (factor score 1) measures the mean of the variables. Factor 2 (factor score 2) faces theory lessons versus practice lessons.

In the following figure we can see explicative and weights variables by factors variables:

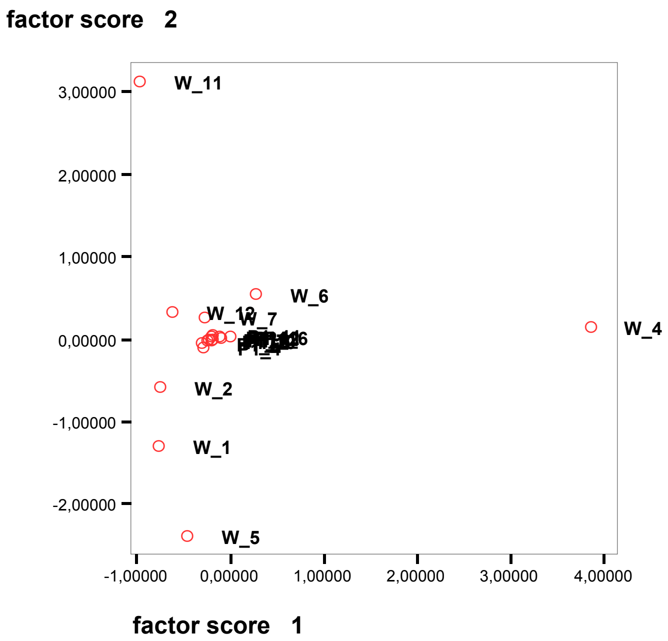


Figure 4.12 Explicative and weights variables

Figure shows a big distance between W11, W4 and W5 and the rest of the variables. Scaled explicative variables are similar to each other.

4.4.1.3 Factorial analysis for W's

We have done a factorial analysis for the weights variables (W).

In the following table we can see the factorial analysis results:

Component*	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	154,637	23,394	23,394	153,569	23,233	23,233
2	104,064	15,743	39,138	105,132	15,905	39,138

Table 4.21 Total Variance explained for weights variables

* Extraction Method: Principal Component Analysis.

Two factors have been detected, they explained 39,138% of total variability. Factor 1 (factor score W 1) measures the mean of the variables. Factor 2 (factor score W 2) faces theory lessons versus practice lessons.

In the following figure we can see weight variables by factors variables.

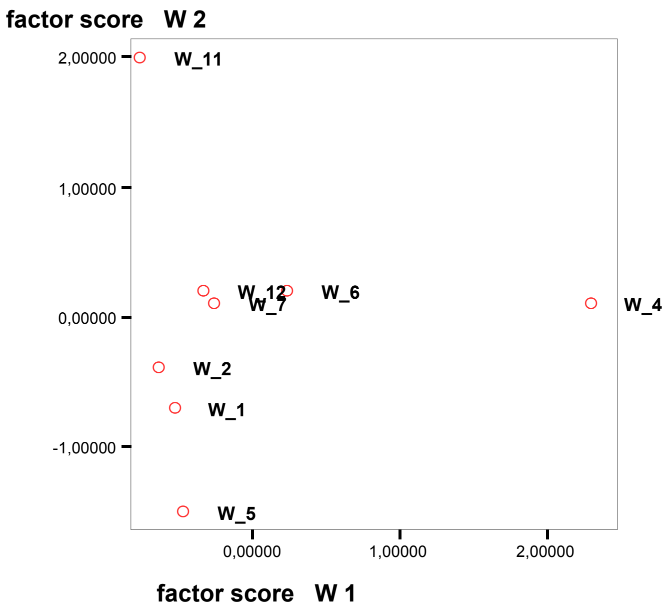


Figure 4.13 Weights variables

Figure shows a big distance between W4 and W11 and the rest of the weight variables. Weight variables W7 and W12 are similar to each other.

4.4.1.4 Factorial analysis for P’s

We have done a factorial analysis for the explicative variables (P). Explicative variables (P) were scaled up to one.

In the following table we can see the factorial analysis results:

Component*	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	194,835	22,472	22,472	193,139	22,277	22,277
2	127,697	14,729	37,201	129,394	14,924	37,201

Table 4.22 Total Variance explained for explicative variables

* Extraction Method: Principal Component Analysis.

Two factors have been detected, they explained 37,201% of total variability. Factor 1 (factor score P 1) measures the mean of the variables. Factor 2 (factor score P 2) faces theory lessons versus practice lessons.

In the following figure we can see explicative variables by factors variables.

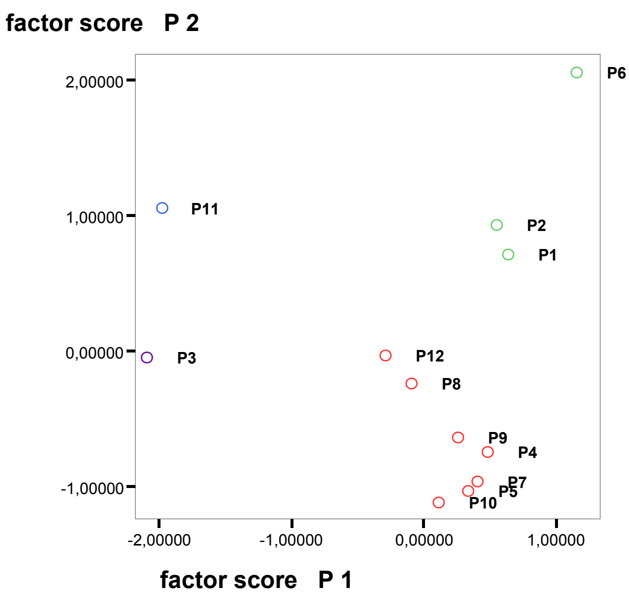


Figure 4.14 Scaled explicative variables

Figure shows a big distance between P6, P11 and P3 and the rest of the variables. Variables P1 and P2 are similar to each other.

4.4.1.5 Correlation analysis between weights variables (W's) and Overall Quality (P3)

In the following figures we can see the correlation coefficients between weights variables (W's) and variable P3.

		P3	W_1	W_2	W_4	W_5	W_6	W_7	W_11	W_12
P3	Correlation	1	,067*	,080*	-,109*	-,015	-,066*	,044*	-,030(*)	,080*
	Sig. (2-tailed)	.	,000	,000	,000	,329	,000	,004	,045	,000
W_1	Correlation	,067*	1	-,046*	-,233*	-,083*	-,143*	-,126*	-,156*	-,128*
	Sig. (2-tailed)	,000	.	,002	,000	,000	,000	,000	,000	,000
W_2	Correlation	,080*	-,046*	1	-,217*	-,136*	-,132*	-,145*	-,149*	-,102*
	Sig. (2-tailed)	,000	,002	.	,000	,000	,000	,000	,000	,000
W_4	Correlation	-,109*	-,233*	-,217*	1	-,198*	-,136*	-,166*	-,193*	-,162*
	Sig. (2-tailed)	,000	,000	,000	.	,000	,000	,000	,000	,000
W_5	Correlation	-,015	-,083*	-,136*	-,198*	1	-,157*	-,164*	-,197*	-,156*
	Sig. (2-tailed)	,329	,000	,000	,000	.	,000	,000	,000	,000
W_6	Correlation	-,066*	-,143*	-,132*	-,136*	-,157*	1	-,122*	-,132*	-,119*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	.	,000	,000	,000
W_7	Correlation	,044*	-,126*	-,145*	-,166*	-,164*	-,122*	1	-,102*	-,058*
	Sig. (2-tailed)	,004	,000	,000	,000	,000	,000	.	,000	,000
W_11	Correlation	-,030(*)	-,156*	-,149*	-,193*	-,197*	-,132*	-,102*	1	-,067*
	Sig. (2-tailed)	,045	,000	,000	,000	,000	,000	,000	.	,000
W_12	Correlation	,080*	-,128*	-,102*	-,162*	-,156*	-,119*	-,058*	-,067*	1
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	.

Table 4.23 Correlation coefficients between W's and P3

* Correlation is significant at the 0.01 level.

In the table above, there are not relevant values. Remember that, correlation coefficients measures only linear relationship.

In the following dendrogram we can appreciate the low similarity that exists between W's and P3:

Figure 4.15 Dendrogram for W's and P3

Nonlinear relationship will be explorer in further research.

4.4.1.6 Weights variables (W's) versus Overall Quality (P3)

In the following figures we can see the relationship between weights variables (W's) and variable P3.

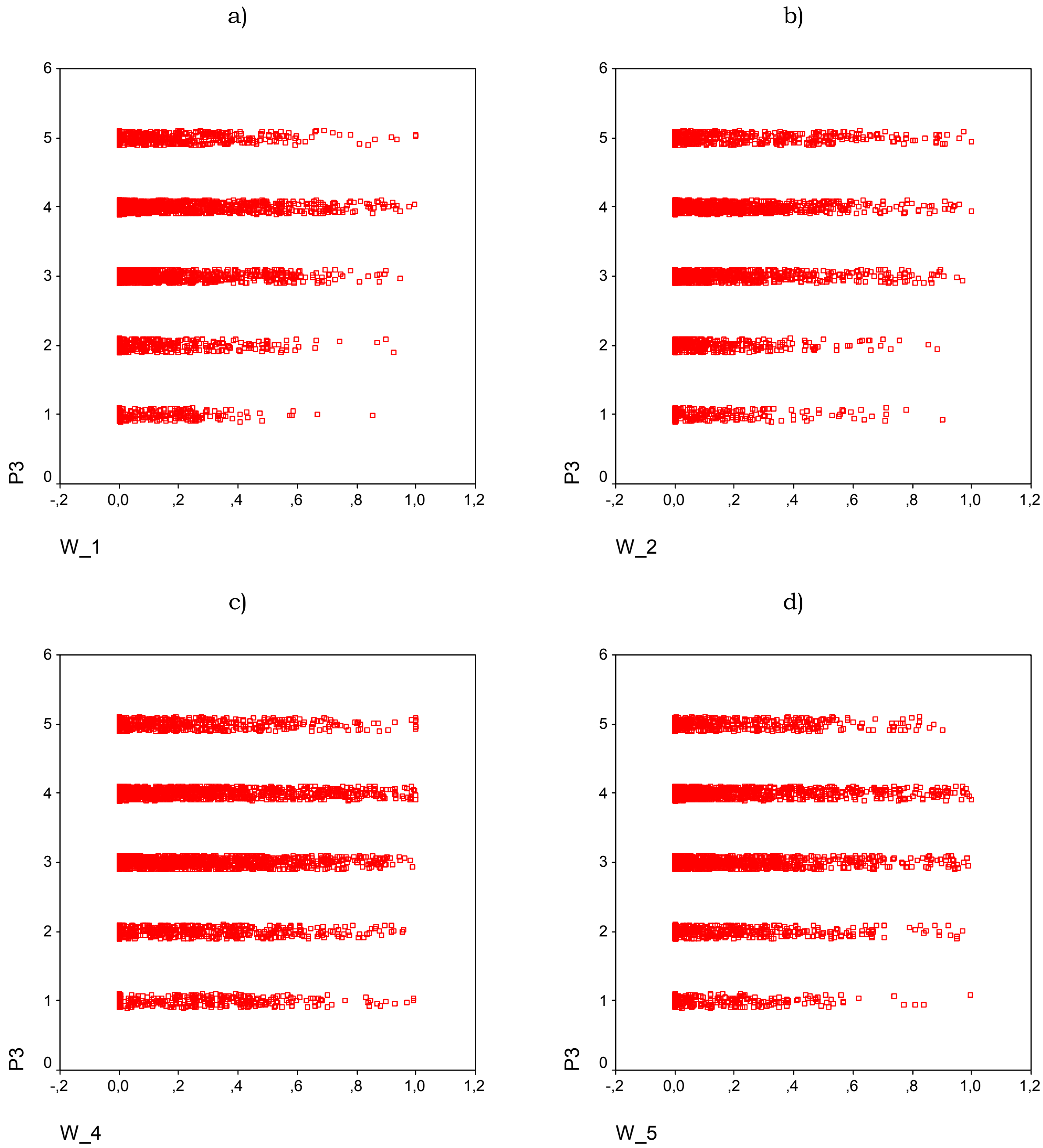


Figure 4.16 Relationship between weigths W1 to W5 with P3

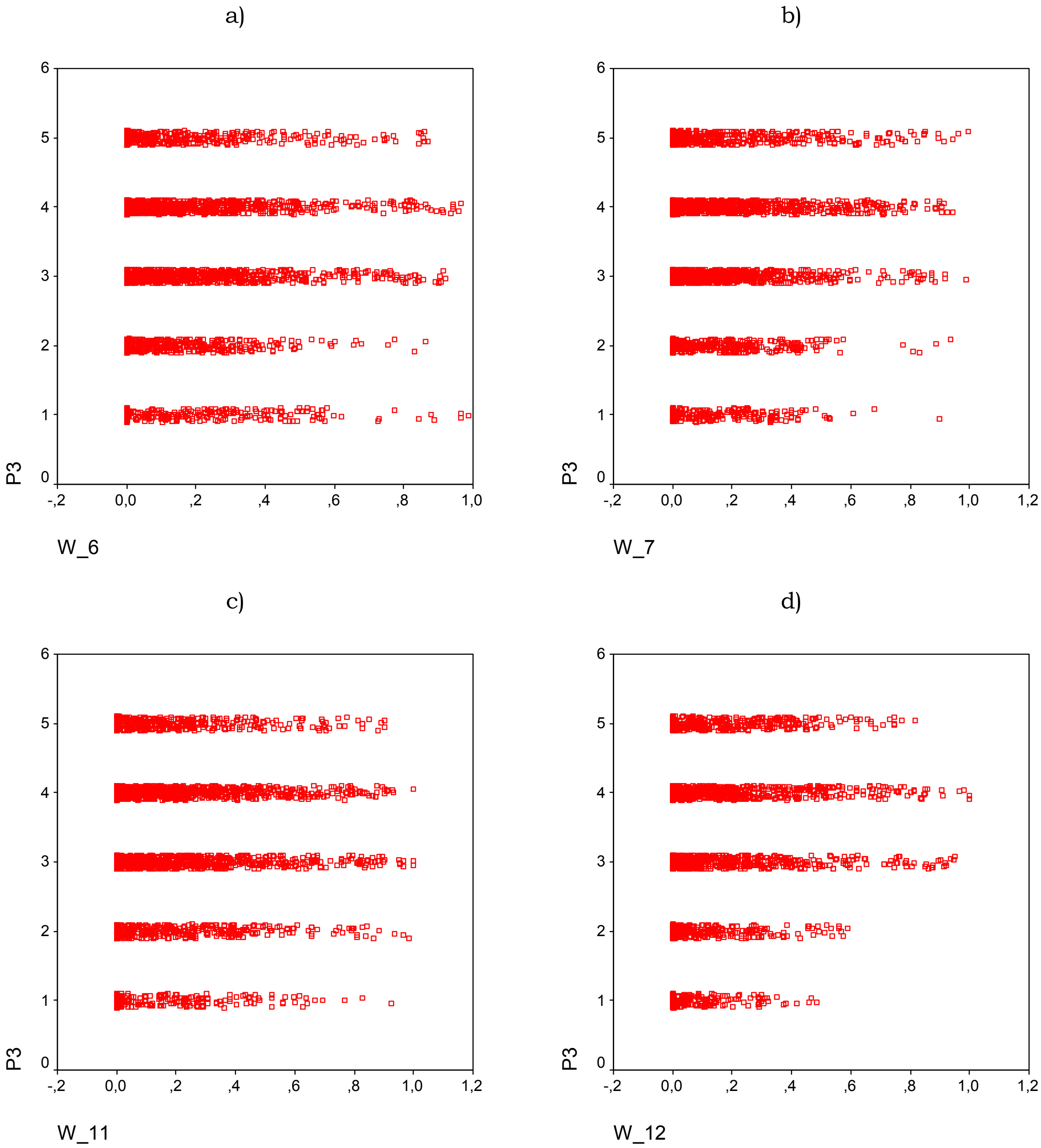


Figure 4.17 Relationship between weights W6 to W12 with P3

In the figures above we can conclude that “the students that assign high importance then to some attribute then assign high qualifications”. In general, there is no answers with high weight (W’s) and low overall quality (P3).
The opposite fraise in not true.

4.4.1.7 Aggregated results by subject types

In the figures 4.18, 4.19 and 4.20 we can see the results of weights in the complete data set, MAE and MBA courses together by subject types 2, 1 and 0, respectively:

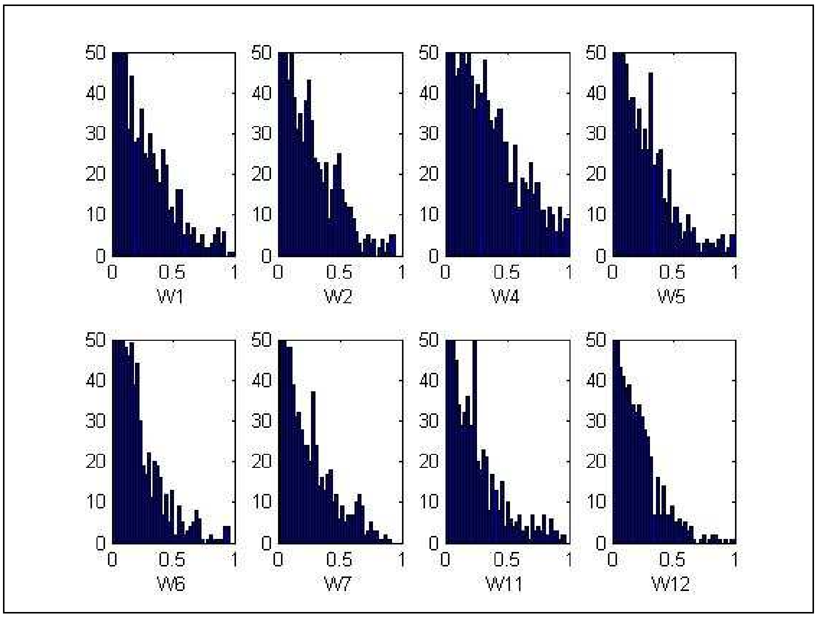


Figure 4.18 Weights for aggregated data set, subject types 2

Weights present a big variability, ranging from 0 to 1 with Standard deviation equal to 0.228. The multivariate analysis has detected three clusters (Mardia 1979, Gordon 1989), both masters show a dissimilar behaviour:

- The first cluster is composed by MBA students.
- The second cluster is composed by low level values in P1 (interest) MAE students.
- The third cluster is composed by high level values in P1 (interest) MAE students.

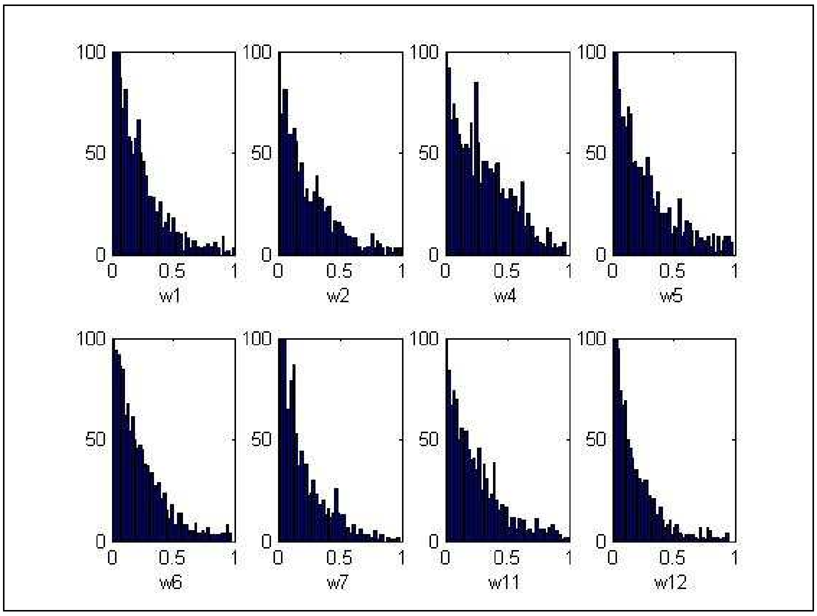


Figure 4.19 Weights for aggregated data set, subject types 1

Weights present a big variability, ranging from 0 to 1 and standard deviation equal to 0.186. The multivariate analysis has detected two clusters (Mardia 1979, Gordon 1989), both masters show a similar behaviour:

- The first cluster is composed by low level values in P1 (interest) students.
- The second cluster is composed by high level values in P1 (interest) students.

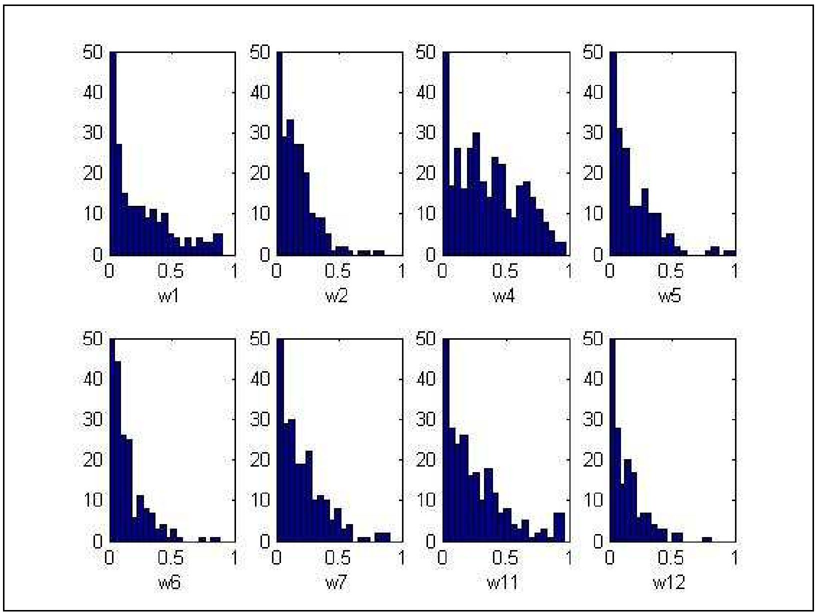


Figure 4.20 Weights for aggregated data set, subject types 0

Weights present a big variability, ranging from 0 to 1 and standard deviation equal to 0.197. The multivariate analysis has detected two clusters (Mardia 1979, Gordon 1989), both masters show a similar behaviour:

- The first cluster is composed by low level values in P1 (interest) students.
- The second cluster is composed by high level values in P1 (interest) students.

4.4.2 MAE results

In the following figure 4.21 we can see the results of weights for the MAE master students:

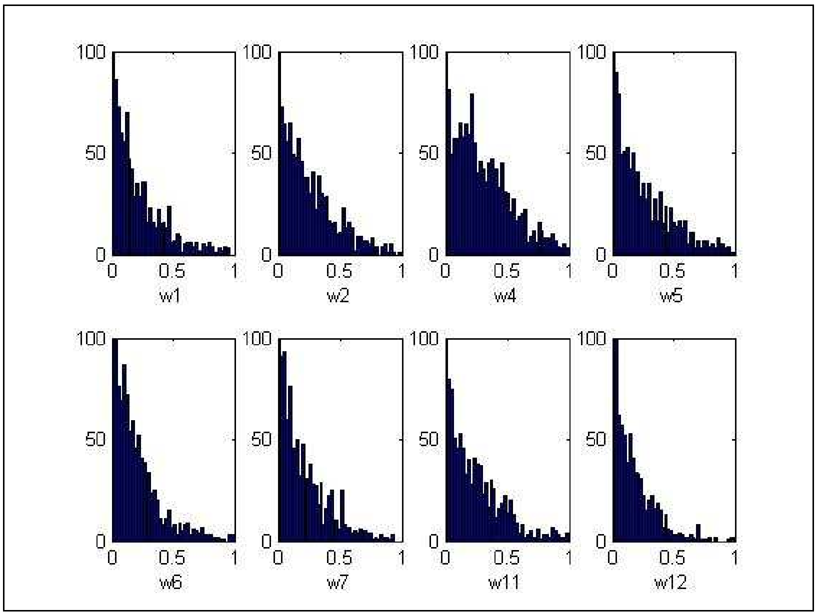


Figure 4.21 Weights for MAE master students

Weights present a big variability, ranging from 0 to 1 and standard deviation equal to 0.191. The multivariate analysis has detected three clusters (Mardia 1979, Gordon 1989):

- The first cluster is composed by low level values in P1 (interest) students.
- The second cluster is composed by high level values in P1 (interest) and high level values P11 (output knowledge) students.
- The third cluster is composed by low level values in P11 (output knowledge) students.

4.4.2.1 Hierarchical cluster analysis

We have done a correlation analysis and a hierarchical cluster analysis by weights variables.

In the following table we can see the correlation analysis results:

		W_1	W_2	W_4	W_5	W_6	W_7	W_11	W_12
W_1	Correlation	1	,099*	,217*	,069*	,132*	,153*	,139*	,116*
	Sig. (2-tailed)	.	,000	,000	,002	,000	,000	,000	,000
W_2	Correlation	,099*	1	,227*	,154*	,153*	,103*	,185*	,103*
	Sig. (2-tailed)	,000	.	,000	,000	,000	,000	,000	,000
W_4	Correlation	,217*	,227*	1	,264*	,095*	,172*	,204*	,094*
	Sig. (2-tailed)	,000	,000	.	,000	,000	,000	,000	,000
W_5	Correlation	,069*	,154*	,264*	1	,133*	,166*	,183*	,112*
	Sig. (2-tailed)	,002	,000	,000	.	,000	,000	,000	,000
W_6	Correlation	,132*	,153*	,095*	,133*	1	,137*	,133*	,134*
	Sig. (2-tailed)	,000	,000	,000	,000	.	,000	,000	,000
W_7	Correlation	,153*	,103*	,172*	,166*	,137*	1	,078*	,092*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	.	,001	,000
W_11	Correlation	,139*	,185*	,204*	,183*	,133*	,078*	1	,064*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,001	.	,004
W_12	Correlation	,116*	,103*	,094*	,112*	,134*	,092*	,064*	1
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,004	.

Table 4.24 Correlation coefficients for W's in MAE data

* Correlation is significant at the 0.01 level (2-tailed).

In the table above, there are not relevant values. Remember that, correlation coefficients measures only linear relationship.

In the following table we can see the stages of the agglomeration schedule results:

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	6	8	106,777	0	0	2
2	5	6	108,996	0	1	3
3	1	5	110,949	0	2	4
4	1	7	120,013	3	0	5
5	1	2	130,455	4	0	6
6	1	4	140,558	5	0	7
7	1	3	200,518	6	0	0

Table 4.25 Agglomeration schedule for MAE data

In the following dendrogram we can appreciate the similarity that exists between weights variables:

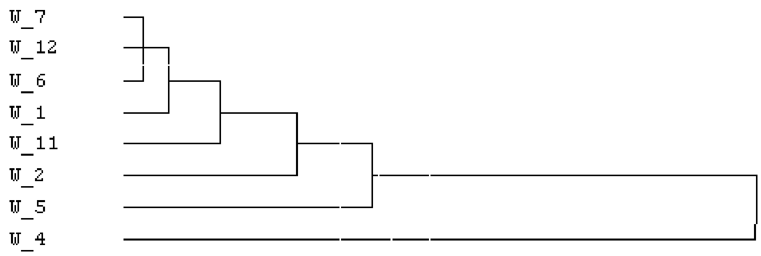


Figure 4.22 Dendrogram for MAE data

Dendrogram shows strong similarity between:

- W7 – W12 and W6

  W7 The usefulness and interest of the *readings and recommended bibliography*.

  W12 *Input level previous to the subject.*

  W6 *Promotions of participation*

We can say, that the cluster measures "Didactic methodology" (in reference to the level of student).

4.4.3 MBA results

In the following figure 4.23 we can see the results of weights for the MBA master students:

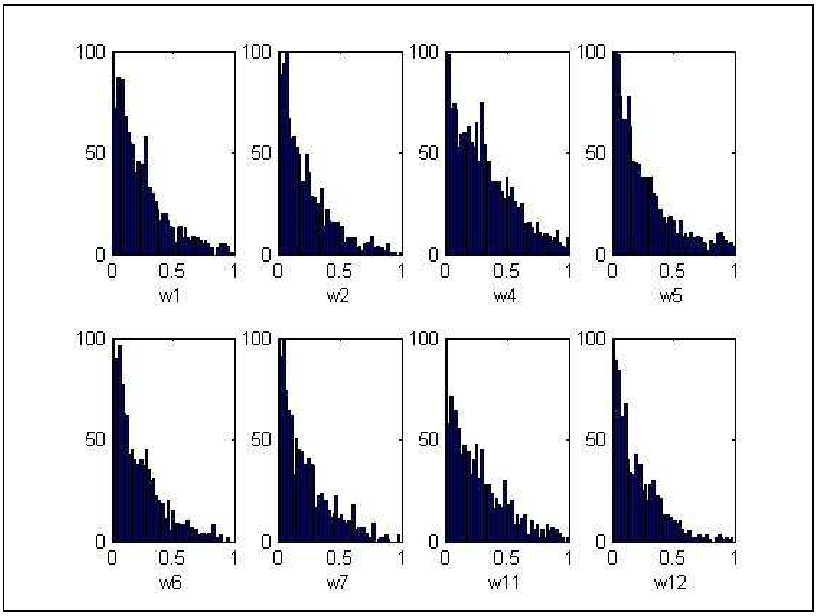


Figure 4.23 Weights for MBA master students

Weights present a big variability, ranging from 0 to 1 and standard deviation equal to 0.196. The multivariate analysis has detected two clusters (Mardia 1979, Gordon 1989):

- The first cluster is composed by low level values in P1 (interest) students.
- The second cluster is composed by high level values in P1 (interest).

4.4.3.1 Hierarchical cluster analysis

We have done a correlation analysis and a hierarchical cluster analysis by weights variables.

In the following table we can see the correlation analysis results:

		W_1	W_2	W_4	W_5	W_6	W_7	W_11	W_12
W_1	Correlation	1	-,059*	-,241*	-,085*	-,131*	-,149*	-,181*	-,120*
	Sig. (2-tailed)	.	,004	,000	,000	,000	,000	,000	,000
W_2	Correlation	-,059*	1	-,208*	-,105*	-,109*	-,129*	-,137*	-,129*
	Sig. (2-tailed)	,004	.	,000	,000	,000	,000	,000	,000
W_4	Correlation	-,241*	-,208*	1	-,191*	-,158*	-,127*	-,218*	-,161*
	Sig. (2-tailed)	,000	,000	.	,000	,000	,000	,000	,000
W_5	Correlation	-,085*	-,105*	-,191*	1	-,164*	-,188*	-,229*	-,171*
	Sig. (2-tailed)	,000	,000	,000	.	,000	,000	,000	,000
W_6	Correlation	-,131*	-,109*	-,158*	-,164*	1	-,122*	-,116*	-,099*
	Sig. (2-tailed)	,000	,000	,000	,000	.	,000	,000	,000
W_7	Correlation	-,149*	-,129*	-,127*	-,188*	-,122*	1	-,100*	-,046*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	.	,000	,023
W_11	Correlation	-,181*	-,137*	-,218*	-,229*	-,116*	-,100*	1	-,041*
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,000	.	,046
W_12	Correlation	-,120*	-,129*	-,161*	-,171*	-,099*	-,046(*)	-,041*	1
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,023	,046	.

Table 4.26 Correlation coefficients for weights for MBA data

* Correlation is significant at the 0.01 level (2-tailed).

In the table above, there are not relevant values. Remember that, correlation coefficients measures only linear relationship.

In the following table we can see the stages of the agglomeration schedule results:

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	6	8	,000	0	0	2
2	2	6	,071	0	1	3
3	2	5	,073	2	0	4
4	1	2	,177	0	3	5
5	1	7	,189	4	0	6
6	1	4	,387	5	0	7
7	1	3	,746	6	0	0

Table 4.27 Agglomeration schedule for MBA data

In the following dendrogram we can appreciate the similarity that exists between weights variables:

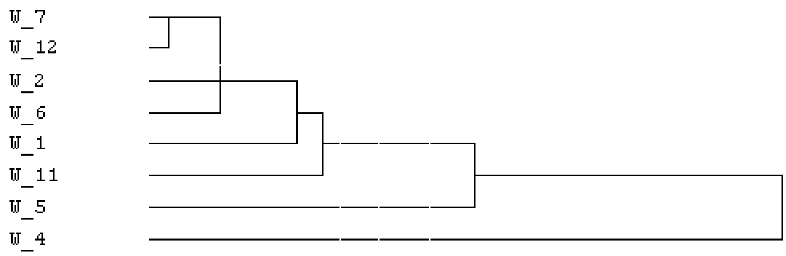


Figure 4.24 Dendrogram for MBA data

Dendrogram shows strong similarity between:

- W7 – W12

  W7 The usefulness and interest of the *readings and recommended bibliography*.

  W12 *Input level previous to the subject.*

We can say, that the cluster measures “Didactic materials” (in reference to the level of student).

4.5 Conclusions of the Chapter 4

ALR methodology was applied to measure the quality of the education at postgraduate department of a Public university.

The data were obtained from surveys conducted from the program of Masters developed by the Department of Business Administration of a Spanish university. Information was gathered from questionnaires on all the subjects taught and all the teachers who taught the subjects. Data from a survey carried out from 2003 to 2008, the unit of analysis was students of two masters of a business school: A total of 5769 questionnaires were administered, and the number of valid questionnaires received was 4372. The questionnaires considered valid were those in which the respondent had answered all of the questions of interest, yielding a full set of variables used in the subsequent analysis.

Data from survey have been classified by years (from 2003 to 2008), terms (T1, T2 and T3) and type subject (2, 1 and 0; qualitative, quantitative and mixed subject, respectively).

We have used those data to find the weights that students assign to every dimension of the “service”. Results were satisfactory; ALR is able to treat this kind of data.

New relationships were discovered (for instance, W7 and W12). Also, relationships between W’s variables and P3 were discovered.

For example, in the following figure we can see this kind of relationships:

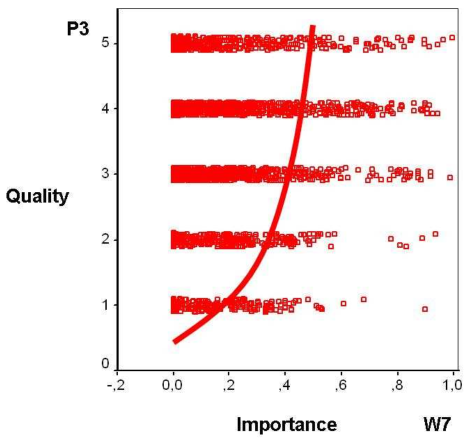


Figure 4.25 Relationship Importance - Quality

Chapter 5

Conclusions and Future Research

In this chapter the most important conclusions of the research, lessons learned throughout this work and future research ways are detailed.

5.1 Conclusions

The knowledge of the relative importance that the customers give to the quality attributes that determine the global service quality is key for any process of service quality improvement.

Several methods of measuring service quality have been developed and discussed over the last few years. Reviewing the service quality literature, and the operational definition of service quality based on the mean of the weights have some limitations.

First, we may have a good service quality on average, but a very bad service quality for some groups of customers. This may happen either in two ways:

- because some segments of the customers have a very different weighting function for the quality attributes, we call this situation “implicated population”
- because they have a different evaluation of the attributes, we call this situation “explicated population”.

These two situations must be identified because we can provide a better service if we identify clusters of customers with different values or opinions about quality. Then, it is more informative to measure service quality in these different populations.

It must be remembered that the mean is only a good descriptive measure when we have an homogeneous sample and that it can be very non representative when the data comes from a mixture of very different populations.

The procedure presented in this work seems to be a useful way to estimate the implicit weights used by each customer in his overall evaluation of service quality.

Concluding Remarks

In this dissertation we have discussed several techniques for measuring the Quality of Service (QoS). We have also presented a new methodology for it based on non parametric statistics.

We have extended our efforts towards three directions:

- First, we have adapted a definition of dissimilitude between data.

- Second, we have developed the necessary linear algebra for solving several numeric problems present in the real world.

- Third, we have calibrated and validated the method.

The algorithms we propose have several advantages:

- It is simple: because it is based on a typical instrument of measurement that the customers are familiar with.

- It is versatile: because it is useful for measuring Quality, Loyalty Customer, Recovery Customer, …

- It is economic: because it can be applied for any number of attributes and/or sample size.

- It is transparent: because it is based on statistical model and linear algebra and can be tested and checked (validated) with the simulated data.

- It is efficient: because it works well in all the simulated cases we have considered.

- Also, it is very easy for programming.

And, particularly, the methodology presented in this thesis has the following advantages:

- Knowledge of the attribute weights allows the ordering of the attributes according to their relative importance to the customer, showing the key factors for improving quality.

- Customer weights can be related to customer characteristics to make market segmentation directly linked to quality objectives. The characteristics of our customers and the market segmentation of our service can be obtained by comparing their mean weights to those of the customers of other services.

- Also, the relative strengths and weaknesses of the service can be determined by comparing the mean value of the attributes of the service to the values of other companies (Benchmark process or SWOT analysis).

- Also, when the attributes of the service quality can be related to some objective measures of performance; it is possible to substitute the subjective evaluations of the attributes by objective measurements, allowing a simple monitoring of the quality index and of their components by Control Charts. In this way, we can use many of the techniques developed for the control of product manufacturing to the improvement of service quality, as Statistical Process Control (SPC).

We have implemented and validated our methodology in several simulated datasets with interesting results. It was very important for calibrating the linear algebra and the different parameters of the methodology. We have implemented our methodology to measure dates from two real cases.

5.2 Future Research

We have also identified several directions towards long-term future work.

5.2.1 Latent Variables:

We will study the possibility to allow that customers, in their evaluations of the overall quality, may be taking into account some attributes not considered in the model:

Suppose that we have a customer population. We assume that each customer has an evaluation score y of the perceived quality of a given service that is a weighted linear combination of several known attributes, factors or dimensions, $x_1,\ldots,x_k$ and, possibly, of a latent variable z depending on other unidentified factors (Ping 1995).

Thus the evaluation score is computed by the customer by giving weights to the different dimensions or attributes considered and the evaluation score reported includes some random measurement error which includes the rounding error and other computation errors made by the customer.

Without loss of generality we assume that the data has been scaled so that the variables y, $x_1,\ldots,x_k$ and z are scores between 0 and 1. Suppose that a random sample of n customers has been surveyed, and let (y_i, x_i), where $x_i = (x_{i1},\ldots,x_{ik})'$, be the answer of customer i.

We assume that

$$y_i = w_{i1}x_{i1} + \cdots + w_{ik}x_{ik} + w_{ik+1}z_i + \varepsilon_i, \qquad 1 \le i \le n \tag{5.1}$$

where z_i is the unobserved random variable corresponding to the evaluation of the unspecified factors for customer i, $w_i = (w_{i1},\ldots,w_{ik},w_{ik+1})$ is a random vector of weights measuring the relative importance that customer i gives to the different attributes x_j, $1 \le i \le k$ and to z in determining overall service quality y and ε_i is a measurement error. The variables w_i, z_i and ε_i are not observed.

The error ε_i takes into account differences between the theoretical and the observed overall quality due to particular behaviour of some of the respondents. We assume that the attribute evaluations are made without measurement error. In practice there will always be some measurement error which can be different for different attributes. However, it is common, to assume this hypothesis for simplicity. Also, we will research the implications of deleting it.

In Peña (2006) a method with the following assumptions is deployed:

Assumption 1

The random variables x_i, z_i, w_i and ε_i are independent.

The justification that x_i and w_i are independent is that the evaluation of an attribute represents how the level of service in this attribute compares to an ideal or standard performance, whereas the weights represent the a prior wishes of the customer.

The independence between x_i, the evaluation of the known attributes and z_i the evaluation of the unknown attribute is made for simplicity and can be easily generalized by assuming for instance that $E(z_i \mid x_i)$ is equal to the mean evaluation of the known attributes. Also, we will research this possible generalization.

Assumption 2

The distribution of w_i is Dirichlet with parameter α. The distribution of z_i is beta with parameter p. The distribution of ε is Normal with mean 0 and variance σ^2.

Observe that the Dirichlet assumption for the weights is in agreement with the basic assumption of a linear quality indicator, that is, that $w_{ij} \geq 0$ and that $\sum_{j=1}^{k+1} w_{ij} = 1$, and therefore, according to (xxx), the score y_i is a weighted average of the scores x_{ij}, and z_i plus a measurement error.

The Beta assumption for the distribution of z_i is in agreement with the values of this variable in the interval 0–1 and allows a reasonable flexibility in the form of the distribution. The Normal distribution for the noise is made for simplicity as a priori the value of σ^2 is expected to be small and therefore the values of the noise are not expected to move the evaluation score y out of the interval 0.1. Also, we will research several alternative ways to model the noise in this model.

5.2.2 Nonlinearity and Interaction

We will study models which are able to deal with nonlinearity and interaction between attributes (Ravi, Warren and Jos, 2002).

5.2.3 Variability in the Distribution of the Attribute Coefficients

In addition to estimating the mean of the coefficients, we will also analyze the role of the variability in the distribution of the index in the customer's population.

5.2.4 Bayesian Models

We will study Bayesian models:

Methods oriented to multidimensional quality measurements are usually based on Conjoint Analysis (Luce and Tukey, 1964). See Carroll and Green (1995) for a survey of the state of this methodology and Lynch et al. (1994), Wedel and DeSarbo (1994) and Ostromand Iacobucci (1995) for interesting applications to the evaluation of service quality. In this methodology customers are asked to provide quality evaluation on several hypothetical services defined by certain levels of the quality attributes. The method assumes that the quality attributes can be given an objective interpretation so that the levels of the attributes have, when presented to the customers for evaluation, a clear meaning to them.

Conjoint Analysis is less useful in situations in which the quality attributes do not have objective standards, and therefore it is very difficult to define a series of hypothetical quality situations for the customers to evaluate.

An alternative procedure in these situations is to use hierarchical Bayesian methods that can be estimated by Markov Chain Monte Carlo method (MCMC), see Lenk et al. (1996), Allenby and Rossi (1999) and Rossi et al. (2001).

Another alternative is to relate the evaluation of the attributes to the overall evaluation of service quality by using a random coefficients regression model. Peña (1997) proposed a model in which the weights of each customer are assumed to be random variables generated by a common multivariate normal distribution and show how to compute by generalized least squares (LS) the mean weights in the population imposing the restrictions that the weights must add up to one. This model was designed for the estimation of the mean weights in the population and the important problem of estimating the individual weights for each person was not considered. We think that it can be easily carried out in the hierarchical Bayesian approach (Ko and Pastore 2005). We will research this alternative.

5.2.5 Linear Structural Relation Models (LISREL)

Another alternative methodology to measure quality service is by using linear structural relation models (LISREL). In this approach the unobserved latent variable quality, η, is related to a vector of p unobserved latent factors, ξ, by

$$\eta = \phi^{'}\xi \tag{5.2}$$

In order to estimate this model we have an observed variable y which is related to the latent variable quality by

$$y = \eta + \varepsilon \tag{5.3}$$

where ε is a $N(0, \sigma^2)$ variable. We also have a set of $m > p$ observed x variables, which are related to the p factors ξ by the linear factor model equation

$$x = \Lambda \xi + \nu \tag{5.4}$$

where the vector ν has a $N_m(0, \Sigma)$ multivariate normal distribution. As the factors ξ will be estimated as linear function of the x variables, by using (5.2) and (5.3) we have that the relation between the observed variables is given by

$$y = \beta' x + \varepsilon$$

which is a linear regression model. From this point of view the model we are proposing can be seen as a reduced form of the structural model. However, the LISREL model usually assumes a fixed regression coefficient in the relation (5.2) among the latent variables, whereas our model allows for different weights among the customers, which we believe is a more realistic assumption (Schumacker and Marcoulides, 1998; Chow et al, 2005).

On the other hand, our model assumes that there is no measurement error in the explanatory variables. This possibility can be introduced into the model by using an equation similar to (5.4) with $\Lambda = \mathrm{I}$, the identity matrix, and assuming some error distribution for the measurement error and incorporating it into the model. Also, if a priori information on the mean of the attributes is available it can be included as prior information.

Then the model can be set up in a Bayesian framework and estimated by Markov Chain Monte Carlo *(*MC^2*)* methods. Note that, as it has been used in the estimation of the model, the hierarchical structure of the model is well suited for Gibbs sampling estimation.

We have also assumed that the evaluation of the unknown attribute is independent of the evaluation of the known attributes. This assumption can be modified by assuming that $z_i \mid \mathrm{x_i}$ has a distribution with parameters which depend on $\mathrm{x_i}$. For instance, we may take $E(z_i \mid \mathrm{x_i}) = {}^{1'x_i}\!/\!_k$ and we can also relate $Var(z_i \mid \mathrm{x_i})$ to the observed variance among the components of x_i. These assumptions, by including additional information, may make the estimation of the model easier but the problem is that they are hard to check with the observed data.

The assumption that the errors ε_i is normally distributed can be replaced by the more general assumption that they have a density of the form $\phi\left({}^{\mu}\!/\!_{\sigma}\right)/\sigma$, where $\phi(\mu)$ is an arbitrary density with mean 0 and variance 1. For example ϕ may have compact support. In this case, the only difference in the estimation procedure would be to replace in (5.2) and (5.3) the normal density φ by ϕ.We can also consider different alternatives for ϕ and choose the one giving the largest value of the likelihood function. Although these alternatives are worth exploring if we have evaluations close to the extremes of the scale, they are not expected to have a large effect on the conclusions of the model.

Another situation to analyze is that the observed variables can be approximated by continuous variables. An alternative approach would be to take into account that, in fact, they are measured as ordinal variables and to include this property into the model. For instance, Johnson (1996) has proposed to consider the evaluation as latent variables which are later discretised into the observed ordinal variables and use MC^2 to estimate the model. See Moreno and Rios Insúa (1998) for an application of these ideas to Service Quality. This alternative will make the model more realistic, but also more complex and the effects in the conclusions are not clear. We will research these situations.

5.2.6 Parallel Computation

We will study and will implement the quality model by parallel algorithm.

5.2.7 Computational Improvement

We will study the linear algebra requirement of the quality model, in particular, we will economize the resolution of the linear equation system and/or the least squared system.

5.2.8 Long Term Project Model

We will study the application of time series in the quality mode:

A typical family of projects is characterized by:

- a long term duration
- a succession of several planning phases
- a constant change of internal customer, at least, one in each phase

For this kind of family of projects we will research the adaptation of the methodology deployed in this thesis.

5.2.9 Applications and Extensions of the Model

We will research the possibility of measuring in other fields of the knowledge and with other variables.

We have developed a methodology for measuring quality service and we have presented its advantages in several examples and in a real case. This methodology is useful for any number of attributes and for any sample size. We will try to extend the methodology to other fields as:

- Marketing: loyalty (Caruana 2002), fidelity plans, customer recovery (Olsen 2002),
- Human resources: labour clime,
- BSC: implementation of the Balanced Scorecard,
- ISO: implementation of quality systems under ISO 9001 (point 8.2.1),
- EFQM: deployment of EFQM model (key results criteria, people results, customer results, society results, etc.).

Bibliography

A

[1] Abrami P. and d'Apollonia S. (1991). "Multidimensional Students' Evaluation of Teaching Effectiveness: Generalizability for "N=1" Research: Comment on Marsh". Journal of Educational Psychology, 83, 411-415.

[2] Abdullah F. (2005). "HEdPERF versus SERVPERF: The Quest for ideal measuring instrument of service quality in higher education sector". Quality Assurance in Education, Vol.13, No.4, pp.305-328.

[3] ACSI. (2008). "American customer satisfaction index". National Quality Research Center. National Quality Research Center, School of Business Administration, University of Michigan.

[4] Afifi A.A. (2004). "Computer-aided multivariate analysis". Chapman & Hall /CRC.

[5] Agresti A. (2002). “Categorical Data Analysis”. Wiley.

[6] Allenby G.M. and Rossi P.E. (1999). “Marketing models of consumer heterogeneity”. Journal of Econometrics 89, 57-78.

[7] Andaleeb S.S. and Basu A.K. (1994). “Technical Complexity and Consumer Knowledge as Moderators of Service Quality Evaluation in the Automobile Service Industry”. Journal of Retailing, 70(4), 367-81.

[8] Anderson C. and Zeithaml C.P. (1984). “Stage of the Product Life Cycle, Business Strategy, and Business Performance”. Academy of Management Journal, 27 (March), 5-24.

[9] Anderson, C., Summey J. and Summey E. (1991). “Promoting the Marketing Major and Marketing Careers to College Students: A Framework”, Journal of Education for Business, 66 (May-June), 260-266.

[10] Anderson E.W., Fornell C. and Lehmann D.R. (1994). “Customer Satisfaction, Market Share and Profitability: Findings from Sweden”. Journal of Marketing, 58(3), 53-66.

[11] Athiayaman A. (1997). “Linking Student Satisfaction and Service Quality Perceptions: the Case of University Education”. European Journal of Marketing, 31(No. 7, 1997), 528-540.

B

[12] Babakus E. and Mangold W. G. (1989). “Adapting the Servqual Scale to Hospital Services: An Empirical Investigation,” Health Service Research, 26(6), 767-80.

[13] Babakus, E and Inhofe, M (1991). “The Role of Expectations and Attribute Importance in the Measurement of Service Quality”. in Gilly M C (ed.), Proceedings of the Summer Educator’s Conference, Chicago, IL: American Marketing Association, 142-44.

[14] Babakus E. and Boller G. (1992). "An empirical assessment of the SERVQUAL Scale". Journal of Business Research, vol 24, n°3. May, 253-268.

[15] Bagozzi R.P. and Fornell C. (1982). "Theoretical Concepts, Measurement, and Meaning". Pp.24-38 in A Second Generation of Multivariate Analysis Vol. 2, C. Fornell (ed), New York: Praeger Publishers.

[16] Barnett R. (1992). "Improving High Education: Total Quality Core". Open University Press. Buckingham.

[17] Barron F.H. and Barrett B.E. (1996). "Decision Quality Using Ranked Attribute Weights". Management Science, pp. 1515-1523.

[18] Barret R. (1996). "Quality and the Abolition of Standards: Arguments against some American Prescriptions for the Improvement of higher Education". Quality in Higher Education, 2 (No. 3): 201-210.

[19] Bayo H. (2003). "Categorical Data Analysis with SAS and SPSS Applications". Lawrence Erlbaum Associates.

[20] Behrens, Schneider and Weisberg. (1978). "Measuring Quality of Production. Marketing 1978, 131-143.

[21] Bitner M. (1990). "Evaluating Service Encounters: The Effects of Physical Surrounding and Employee Response". Journal of Marketing, 54 (April, 55-67.

[22] Bober W. (2009). "Numerical and Analytical Methods with MATLAB". CRC Press.

[23] Bolton R.N. and Drew J.H. (1991a). "A Multistage Model of Customer's Assessment of Service Quality and Value". Journal of Consumer Research, 17(March), 375-385.

[24] Bolton, R N and Drew, J H (1991b). “A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes,” Journal of Marketing, 55(January), 1-9.

[25] Bou J.C. and Camison C. (2000). “Validity and Reliability in Perceived Quality Measurement Models”. International Journal of Quality & Reliability Management, 17 (8).

[26] Boulding W., Karla A., Staelin R. and Zeithaml V. A. (1993). “A dynamic process model of service quality: From expectations to behavioral intentions”. Journal of Marketing Research, 30 (February) 7-27.

[27] Brady M.K. and Robertson C.J. (2001). “Searching for a Consensus on the Antecedent Role of Service Quality and Satisfaction: An Exploratory Cross-National Study”. Journal of Business Research, 51(1) 53-60.

[28] Brady M.K., Cronin J. and Brand R.R. (2002). “Performance–Only Measurement of Service Quality: A Replication and Extension”. Journal of Business Research, 55(1), 17-31.

[29] Brogowicz A.A., Delene L.M. and Lyth D.M. (1990). “A synthesised service quality model with managerial implications”. International Journal of Service Industry Management. Vol.1, pp. 27-45.

[30] Brown T.J., Churchill G.A. and Peter J.P. (1993). “Improving the Measurement of Service Quality”. Journal of Retailing, 69(1), 127-39.

[31] Brown S.W. and Swartz T.A. (1989). “A Gap Analysis of Professional Service Quality”. Journal of Marketing, 53 (April), 92-98.

[32] Browne B., Kaldenberg D., Browne W. and Brown D. (1998). “Student as Customer: Factors Affecting Satisfaction and Assessments of Institutional Quality”. Journal of Marketing for Higher Education, 8 (No. 3,, 1-14.

[33] Browne M. W. and Cudeck R. (1993). “Alternative Ways of Assessing Model Fit”. In Testing Structural Equations Models, Kenneth, A. Bollen y J. Scott Long (eds.). United States of America: Sage Publications.

[34] Buzzell R.D. and Gale B.T. (1987). “The PIMS Principles”. New York: The Free Press.

C

[35] CABINTEC: Intelligent Cabin Truck for road transport, http://www.cabintec.net.

[36] Carman J.M. (1990). “Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions”. Journal of Retailing, 66 (1), 33-35.

[37] Carroll D.J. and Green P.E. (1995). “Psychometric Methods in Marketing Research: Part I Conjoint Analysis”. Journal of Marketing Research. XXXII, 385-391.

[38] Caruana A. (2002). “Service loyalty: The effects of service quality and the mediating role of customer satisfaction”. European Journal of Marketing, 36 (7/8), 811-828.

[39] Cashin W. and Downey, R. (1992). “Using Global Student Rating Items for Summative Evaluation”. Journal of Educational Psychology, 84 (No. 4), 563-72.

[40] Chakrapani C. (1998). “How to Measure Service Quality & Customer Satisfaction: The Informal Guide for Tools and Techniques”. (by Chuck Chakrapani, Author).

[41] Chambers R.L. (2005). “Analysis of Survey Data”. Wiley.

[42] Chow C.C. and Luk P. (2005). “A Strategic Service Quality Approach using Analytic Hierarchy Processes”. Managing Service Quality, Vol.15, No.3, pp.278-289.

[43] Christensen R. (2001). "Advanced Linear Modelling: Multivariate, Time Series, and Spatial Data; Nonparametric Regression and Response Surface Maximization". Springer-Verlag.

[44] Chua C. (2004). "Perceptions of Quality in Higher Education". Proceedings of Australian Universities Quality Forum, AUQA Occasional Publication.

[45] Churchill G.A. Jr. and Suprenant C. (1982). "An investigation into the determinants of consumer satisfaction". Journal of Marketing Research, 19 (November), 491–504.

[46] Clement J. and Selvam M.. (2006). "Service Quality Gaps: A Retro Analysis". Academic Open Internet Journal. Vol. 18.

[47] Coates N. and Koerner R. (1996). "How Market Orientered are Business Studies Degrees?". Journal of Marketing Management, (No. 12), 455-475.

[48] Colbert A., Levary R. and Shaner M. (2000). "Determining the relative efficiency of MBA programs using DEA". European Journal of Operational Research, 125 (No. 3), 656-667.

[49] Comm C. and LaBay D. (1996). "Repositioning Colleges Using Changing Student Quality Perceptions: An Exploratory Investigation". Journal of Marketing for Higher Education, 7 (No. 4), 21-34.

[50] Cook D.P., Goh C. and Chung C.H. (1999). "Service Typologies: A State of the Art Survey". Producyion and Operations Management, n°3, vol. 8.

[51] Cork R. C., Isaac I., Elsharydah A., Saleemi S., Zavisca F. and Alexander L. (2004). "A Comparison Of The Verbal Rating Scale And The Visual Analog Scale For Pain Assessment". Int. Journal of Anesthesiology. 8(1) 315–333.

[52] Cronin J. and Taylor S. (1992). "Measuring Service Quality: A Reexamination and Extension". Journal of Marketing, 56 (July): 55-68.

[53] Cronin J.J. and Taylor S.A. (1994). "SERVPERF Versus SERVQUAL: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality". Journal of Marketing, 58, 125-131.

[54] Cronin J. J., Brady M. K. and Hult G. T. M. (2000). "Assessing the Effects of Quality, Value, and Customer Satisfaction on Consumer Behavioral Intentions in Service Environments". Journal of Retailing, 76/2, 193-218.

[55] Crosby, P B (1984). Paper presented to the "Bureau de Commerce," Montreal, Canada (Unpublished).

[56] Cronbach L.J. (1951). "Coefficient Alpha and the internal structure of Test". Psychometrika, vol 16, n°3, September, 297-334.

[57] Cullen J., Joyce J, Hassall T. and Broadbent M. (2003). "Quality in Higher Education: From Monitoring to Management", Quality Assurance in Education, Vol.11, No.1, pp.5-14.

D

[58] Dabholkar P.A., Shepherd D.C. and Thorpe D.I. (2000). "A Comprehensive Framework for Service Quality: An Investigation of Critical, Conceptual and Measurement Issues through a Longitudinal Study". Journal of Retailing, 76(2), 139-173.

[59] Derek A. (2000). "Analysis of Customer Satisfaction Data". ASQ quality Press.

[60] Dill, D. (1995). "Through Deming's Eyes: A Cross-National Analysis of Quality Assurance in Higher Education". Quality in Higher Education, 1, 95-110.

[61] Dingus T.A., Klauer S.G., Neale, V.L., Petersen A., Lee S.E., Perez M.A., Hankey J., Ramsey D., Gupta S., Bucher C., Doerzaph Z. R., Jermeland J., and Knipling R.R. (2005). "The 100-Car Naturalistic Driving Study: Phase II Results of the 100-Car Field Experiment". Washington, D.C.: National Highway Traffic Safety Administration.

[62] Driscoll T.A. (2009). "Learning MATLAB". Society for Industrial and Applied Mathematics.

[63] Dubas K. and Strong J. (1993). "Course Design Using Conjoint Analysis", Journal of Marketing Education, 15 (Spring), 31-36.

[64] Dubas K.; Ghani W., Davis S. and Strong J. (1998). "Evaluating Market Orientation of an Executive MBA Program", Journal of Marketing for Higher Education, 8 (No. 4), 49-59.

[65] Dutka A.F. (1994). "AMA handbook for customer satisfaction". NTC Business Books.

E

[66] EFQM. (2000). "European Foundation for Quality Management. Self-Assessment. Guidelines for Public Sector: Education". EFQM Bruselas.)

[67] EFQM. (2008). European Foundation for Quality Management. http://www.efqm.org.

[68] Eiglier P. and Langeard E. (1987). "Servuction, Le Marketing des Services". Paris: McGraw-Hill.

[69] Encyclopedia of Statistical Sciences, vol.2. Wiley, New York.

[70] Everitt B.A. (2001). "Applied Multivariate Data Analysis".

F

[71] Fernández J. and Mateo M. (1992). “Student Evaluation of University Teaching Quality: Analysis of a Questionnaire for a Sample of University Students in Spain”. Educational and Psychological Measurement, 52 (No. 3), 675-686.

[72] Finn D.W. and Lamb C.W. (1991). “An Evaluation of the SERVQUAL Scale in a Retailing Setting”. in Holman, R and Solomon, M R (eds.), Advances in Consumer Research, Provo, UT: Association for Consumer Research, 480-93.

[73] Fullerton G. (2005). “The Service Quality-Loyalty Relationship in Retail Services: Does Commitment Matter?”. Journal of Retailing and Consumer Services, 12, 99-111.

[74] Fuchs C. (1998). “Multivariate quality control: theory and applications”. Marcel Dekker.

G

[75] Garvin D.A. (1983). “Quality on the Line”. Harvard Business Review, 61 (September-October), 65-73.

[76] Garvin D.A. (1984). “What Does Product Quality Really Mean?”. Sloan Management Review, Fall, 25-43.

[77] Garvin D.A. (1987). “Competing on the Eight Dimensions of Quality”. Harvard Business Review, 65-6, 101-109.

[78] Giacobbe R. and Segal M. (1994). “Rethinking Marketing Research Education: A Conceptual, Analytical and Empirical investigation”. Journal of Marketing Education, (Spring), 43-58.

[79] Goldghen L. and Kane K. (1997). “Repositioning the MBA, Issues and Implications”. Journal of Marketing for Higher Education, 8 (No.1), 15-24.

[80] Golub G. and Van Loan C. (1989). "Matrix Computations" (2nd Edition). Johns Hopkins Series in the Mathematical Sciences. The Johns Hopkins University Press.

[81] Gordon A.D. (1989)."Classification: methods for the exploratory analysis of multivariate data". Chapman & Hall.

[82] Gotlieb J.B., Grewal D. and Brown S.W. (1994). "Consumer Satisfaction and Perceived Quality: Complementary or Divergent Constructs". Journal of Applied Psychology, 79(6), 875-85.

[83] Greenwald A.(1997). "Validity Concerns and Usefulness of Student Rating of Instruction". American Psychologist, 52, 1182-86.

[84] Grönroos C. (1982). "Strategic Management and Marketing in the Service Sector. Finland". Swedish School of Economics and Business Administration.

[85] Grönroos C. (1990). "Service Management and Marketing: Managing the Moments of Truth in Service Competition". Mass: Lexington Books.

[86] Gumpertz M.L. and Pantula S.G. (1998). "Regression, random coefficient". In: S. Kotz, C. B. Read and D. L. Banks (Eds.).

[87] Guolla M. (1999). "Assessing the teaching quality to student satisfaction relationship: Applied customer satisfaction research in the classroom". Journal of Marketing Theory and Practice, 7 (No.3), 87-98.

[88] Gustafsson A., Nilsson L. and Johnson M. D. (2003). "The Role of Quality Practices in Service Organizations". International Journal of Service Industry Management, Vol.14, No.2, pp.232-244.

[89] Gustafsson, A. (2007). "Conjoint Measurement: Methods and Applications". Springer.

H

[90] Hamer L.A. (2003). “Pragmatic View of Perceived Service Quality Measurement”. Services Marketing Quarterly. The Haworth Press, Inc.

[91] Härdle W. (1990). “Applied nonparametric regression”. Econometric Society Monographs, vol.19. Cambridge University Press, Cambridge.

[92] Hartline,M. D. and Ferrell O. C. (1996). “The Management of Customer Contact Service Employees: An Empirical Investigation”. Journal of Marketing, 69 (October), 52-70.

[93] Hayes B.E. (1998). “Measuring customer satisfaction: survey design, use, and statistical analysis methods”. 2nd ed. Oxford U.Press.

[94] Hendricks K.B. and Singhal V.R. (2001). “Firm Characteristics, Total Quality Management, and Financial Performance”. Journal of Operations Management, 19, 269-285.

[95] Herche J. and Swenson M. (1991). “Multidimensional Scaling: A Market Research Tool to Evaluate Faculty Performance in the Classroom”. Journal of Marketing Education, 13 (Fall), 14-20.

[96] Horovitz J. (2001). “Service Strategy”. Pitman Publishing.

[97] Hüttenrauch (1994). “Price, quality, and consumers' behaviour”. Journal of Consumer Policy. Vol. 17, Nº 3.

I

[98] Iacobucci D., Grayson K.A. and Ostrom A.L. (1994). “The Calculus of Service Quality and Customer Satisfaction: Theoretical and Empirical Differentiation and Integration,” in Swartz, T A; Bowen, D H and Brown, S W (eds.), Advances in Services Marketing and Management, Greenwich, CT: JAI Press,1-67.

[99] Imkamon T., Saensom P., Tangamchit P. and Pongpaibool P. (2008). “Detection of hazardous driving behavior using fuzzy logic. Electrical Engineering/Electronics, Computer, Telecommunications and

Information Technology”. ECTI-CON 2008. 5th International Conference on, 657–660.

[100] Irons K. (1994). “Managing Service Companies-Strategies for Success”. Wokingham, Addison-Wesley.

[101] ISO 66992: 2001. “Complaints managements systems. Guide to design and implementation”. International Standardization Society. http://www.iso.org.

[102] ISO 175001-1: 2004. “Service quality for small retail trade. Part 1: General Requirements”. Standardization Society. http://www.iso.org.

[103] ISO 10002: 2004. “Quality management - Customer satisfaction -- Guidelines for complaints handling in organizations”. Standardization Society. http://www.iso.org.

[104] ISO 21707: 2008. “Intelligent transport systems - Integrated transport information, management and control - Data quality in ITS systems”. Standardization Society. http://www.iso.org.

[105] ISO 26362: 2009. “Access panels in market, opinion and social research - Vocabulary and service requirements”. Standardization Society. http://www.iso.org.

J

[106] Johnson V. (1996). “On Bayesian Analysis of Multirater Ordinal Data: An Application to Automated Essay Grading”. Journal of the American Statistical Association. 91, 42-51.

[107] Johnson R.A. (2002). “Applied Multivariate Statistical Analysis”. Prentice Hall.

[108] Joseph M. and Joseph B. (1997). “Employers’ Perceptions of Service Quality in Higher Education”. Journal of Marketing for Higher Education, 8 (No. 2), 1-13.

[109] Joseph M., Yakhou M and Stone G. (2005). "An Educational Institutions Quest for Service Quality: Customers' Perspective". Quality Assurance in Education, Vol.13, No.1, pp.66-82.

[110] Julia E.B. and William B.T. (2004). "Linking Dimensions of Service Quality to Organizational outcomes", Managing Service Quality, Vol.14, Issue 1, pp.75-89.

[111] Juran J.M. (1988). "Juran on Planning for Quality". New York: The Free Press.

[112] Juran J.M. (2001). "Manual de Calidad de Juran" (5° Ed.). McGraw-Hill Interamericana.

[113] Juric B., Tood S. and Henry J. (1997), "From the Student Perspective: Why Enrol in an Introductory Marketing Course?". Journal of Marketing Education, 19 (1), 65-76.

K

[114] Kang G.D. and James J. (2004). "Service Quality Dimensions: An Examination of Gronroos's Service Quality Model", Journal of Managing Service Quality, Vol.14, No.4, pp.266-277.

[115] Kannan S.R. (2005). "Extended Bidirectional Associative Memories: A Study on Poor Education". Mathematical and Computer Modelling, Vol.42, No.3-4, pp.389-395.

[116] Karsak E.E., Sozer S. and Alpsekin S.E. (2002). " Product Planning in Quality Function Deployment using a combined Analytic Network Process and Goal Programming Approach". Computers and Industrial Engineering, Vol.44, pp.171-190.

[117] Kassim N.M. and Bojei J. (2002). "Service Quality: Gaps in the Telemarketing Industry". Journal of Business Research, 55(11), 845-52.

[118] Kawlath (1969). “Teoría de la Calidad”. Diaz de Santos.

[119] Kepner J.V. (2009). “Parallel MATLAB for Multicore and Multinode Computers”. SIAM.

[120] Klauer S.G., Sudweeks J., Hickman J.S. and Neale V.L. (2006). “An assessment of the relative risk of engaging in potentially unsafe driving behaviors”. Virginia Polytechnic Institute and State University.

[121] Ko Y.J. and Pastore D.L. (2005). “A Hierarchical Model of Service Quality in the Recreational Industry”. Sport Marketing Quarterly. 14 (2).

[122] Kotler, P. and Fox, K. (1997). “Strategic Marketing for Educational Institution”. (2ª ed.). Englewood Cliffs, N.J.

[123] Kotler P. (2003). “Marketing Management”. New Delhi: Prentice Hall of India.

[124] Kunst P. and Lemmink J. Editors. (1996). Managing service quality. Vol I. Paul Chapman Publishing Ltd.

[125] Kunst P. and Lemmink J. Editors. (1997). Managing service quality. Vol II. Paul Chapman Publishing Ltd.

L

[126] Lado N., Cardone C. and Rivera P. (2003). "Measurement and Effects of Teaching Quality: An Empirical Model Applied to Masters Programs. Journal of the Academy of Business Education, 4, Spring, 28-40.

[127] Lauro C. (2002). “Multivariate Total Quality Control: Foundation and recent advances”. Physica-Verlag.

[128] Lawson C.L. and Hanson R.J. (1974). “Solving Least Squares Problems”. Prentice-Hall, Englewood Cliffs, NJ.

[129] Lawton L. and Lundsten L. (1998). "Contrasts Between Benefits Expected and Delivered Among MBA Inquirers, Students and Graduates". Journal of Marketing for Higher Education, 8 (3), 15-29.

[130] Lenk P.J., DeSarbo W.S., Green P.E. and Young M.R. (1996). "Hierarchical Bayesian analysis: recovery of partworth heterogeneity fromreduced experimental". Marketing Sci. 28 (May), 215-222.

[131] Lewis R. (1987). "The Measurement of Gaps in the Quality of Hotel Service," International Journal of Hospitality Management, 6(2), 83-88.

[132] Lewis R. (1991). "Service Quality: An International Comparison of Bank Customer's Expectations and Perceptions". Journal of Marketing Management, 7(1), 47-62.

[133] Liljander V. and Strandvik T. (1993). "Different Comparison Standards as Determinants of Service Quality". Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior, Vol. 6, 118-132.

[134] Llusar J.C.B., Zornoza C.C. and Tena A.B.E. (2001). "Measuring the relationship between firm perceived quality and customer satisfaction and its influence on purchase intentions". Total Quality Management, 12 (6), 719-734.

[135] Lomas L. (2004). "Embedding Quality: The challenges for higher education". Quality Assurance in education, Vol.12, No. 4, pp. 157-165.

[136] Luce D.R. and Tukey J.W. (1965). "Simultaneous Conjoint Measurement: A new type of Fundamental Measurement". Journal of Mathematical Psychology, 1: 1-27.

[137] Lynch J.G., Bouzas T.E. and Berg S.V. (1994). "Regulatory Measurement and Evaluation of Telephone Service Quality". Management Science, 40, 2: 168-194.

M

[138] Mahapatra S.S. and Khan M.S. (2006). “Current Practices of TQM Implementation and Future Trend”. Industrial Engineering Journal, Vol. 35, No. 5, pp. 28-33.

[139] Mallet A.. (1986). “A maximum likelihood estimation method for random coefficient regression models”. Biometrika 73, 645 –646.

[140] Mardia K. V. (1979). “Multivariate Analysis”. Academic Press.

[141] Marks R.B. (2001). “Determinants of Student Evaluations of Global Measures of Instructor and Course Value”. Journal of Marketing Education, 22 (No. 2) 108-119.

[142] Marsh, H. (1987). “Students' Evaluations of University Teaching: Research Findings, Methodological Issues, and Directions for Future Research”. International Journal of Educational Psychology, 11 (No. 3), 253-388.

[143] Marsh H. (1991a). “Multidimensional Students' Evaluations of Teaching Effectiveness: A Test of Alternative Higher Order Structures”. Journal of Educational Psychology, 83, 285-296.

[144] Marsh H. (1991b). “A Multidimensional Perspective on Students' Evaluations of Teaching Effectiveness: Reply to Abrami and d'Apollonia”, Journal of Educational Psychology, 83), 416-421.

[145] Marsh H. and Bailey M. (1993). “Multidimensional Students’ Evaluations of Teaching Effectiveness”. Journal of Higher Education, 64, (No. 1), 1-17.

[146] Marsh H. (1994). “Weighting for the Right Criteria in Instructional Development and Effective Assessment (IDEA) System: Global and Specific Ratings of teaching Effectiveness and their Relation to Course Objectives”. American Psychologist, 86 (No. 4), 631-648.

[147] Marsh H. and Roche L. (1997). "Making Student's Evaluation of Teaching Effectiveness Effective". American Psychologist, vol. 52, (No. 11), 1187-1197.

[148] Martin G. and Bray G. (1997). "Assessing Customer Satisfaction with a Master of Business Administration Program: Implications for Resource Allocation". Journal of Marketing for Higher Education, 8 (No. 2), 15-28.

[149] Martín de Diego I., Conde C. and Cabello E. (2009). "Definition of driving risk by mixture of experts". First International Conference on Driver Distraction and Inattention. September 28-29, 2009. Göteborg, Sweden.

[150] Mazis M.B., Antola O.T. and Klippel R.E. (1975). "A Comparison of Four Multi-Attribute Models in the Prediction of Consumer Attitudes". Journal of Consumer Research, 2(June), 38-52.

[151] Mckeachie W. (1997). "Student Rating: The Validity of Use". American Psychologist, 52, 1218-25.

[152] Miller, J.A. (1977). "Exploring Satisfaction, Modifying Modes, Eliciting Expectations, Posing Problems, and Making Meaningful Measurements". in Hunt, K (ed.), Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, Cambridge, MA: Marketing Science Institute, 72-91.

[153] Mitra A. (2003). "Fundamentals of Quality Control and Improvement". Prentice-Hall Ltd.

[154] Moreno A. and Rios Insúa D. (1998). "Issues in Service Quality Modelling". In: Bernardo, J.M., et al.(Eds.), Bayesian Statistics, vol.6. 441-457.

[155] Morgan C. and Murgatroyd S. (1994). "Total Quality in the Public Sector". Open University Press. Buckingham, UK.

[156] Muñoz A. and Moguerza J. M. (2006). “Estimation of High-Density Regions Using One-Class Neighbor Machines”. IEEE Transactions on Pattern Analysisi and Machine Intelligence, n°3, vol. 28.

[157] Murphy A. (1999). “Enhancing The Motivations for Good Teaching with an Improved System of Evaluation”. Financial Practice and Education, 9 (No. 2), 100-104.

N

[158] Naik B and Ragothaman S. (2004). “Using neural networks to predict MBA student success”. College Student Journal, March, Vol.38.

[159] Naumann E. (1995). “Customer satisfaction measurement and management: using the voice of the customer”. Wiley.

[160] Normann R (1984). “Service Management”. New York: Wiley.

O

[161] Olsen S.O. (2002). “Comparative Evaluation and the Relationship between Quality, Satisfaction and Repurchase Loyalty”. Journal of the Academy of Marketing Science, 30 (3).

[162] Ostrom A. and Iacobucci D. (1995). “Consumer Trade-Offs and the Evaluation of Services”. Journal of Marketing, 59: 17-28.

[163] Ozment J. and Morash E.A. (1994). “The Augmented Service Offering for Perceived and Actual Service Quality”. Journal of the Academy of Marketing Science, 22/4, 352-363.

P

[164] Parasuraman A., Zeithaml V.A. and Berry, L.L. (1985). "A Conceptual Model of Service Quality and Its Implications for Future Research". Journal of Marketing, Fall, 41-50.

[165] Parasuraman A., Zeithaml V.A. and Berry L. L.. (1988). "SERVQUAL: A multiple Item scale for Measuring Consumer Perceptions of Service Quality". Journal of Retailing, 64: 12-41.

[166] Parasuraman A., Zeithaml V. A. and Berry L. L.. (1990). "Guidelines for conducting Service Quality Decrease". Marketing Research. 2(4), 34-44.

[167] Parasuraman A., Berry L.L. and Zeithaml, V.A. (1991a). "Perceived Service Quality as a Customer Based Performance Measure: An Empirical Examination of Organizational Barriers Using an Extended Service Quality Model". Human Resource Management, Fall, 335-64.

[168] Parasuraman A., Zeithaml V.A. and Berry L.L.. (1991b). "Refinement and Reassessment of the SERVQUAL Scale". Journal of Retailing, Vol. 67, 4: 420-451.

[169] Parasuraman A., Berry L.L. and Zeithaml, V.A. (1991c). "Understanding Customer Expectations of Service". Sloan Management Review, Spring, 39-48.

[170] Parasuraman A., Zeithaml V. A. and Berry L. L.. (1994a). "Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research". Journal of Marketing, Vol. 58 (January): 111-124.

[171] Parasuraman A., Zeithaml V.A. and Berry L.L.. (1994b). "Alternative Scales for Measuring Service Quality: A Comparative Assessment Based on Psychometric and Diagnostic Criteria". Journal of Retailing, Vol. 70, Nº 3: 201-230.

[172] Parasuraman A., Zeithaml V.A. and Berry L.L. (1996). "The behavioral consequences of service quality". Journal of Marketing, 60, 31-46.

[173] Parasuraman A (2002). "Service Quality and Productivity: A Synergistic Perspective". Managing Service Quality, Vol. 12, No.1, pp.6-9.

[174] Peña D. (1997a). "Measuring service quality by Linear Indicators". Management Service Quality. (P. Kunst and J. Lemmink, eds). Chapman Publishing Ltd. London, 35-51.

[175] Peña D. (1997b). "La Mejora de la Calidad en la Educación: Reflexiones y Experiencias", Boletín de Estudios Económicos, LII, No. 161.

[176] Peña D. (1999a). "Comment on Quality Indexes Models. Bayesian Statistics 6". Bernardo et al. (ed). Oxford University Press, 456.

[177] Peña D. (1999b). "A Methodology for Building Service Quality Indices". 53rd Annual Quality Congress Proceedings. ASQ American Society for Quality. 550-559.

[178] Peña D., Prieto F.J. (2001). "Cluster identification using projections". JASA, 96.

[179] Peña D. and Yohai V.. (2006). "A Dirichlet Random Coefficient Regression Model for Quality Indicators". Journal of Statistical Planning and Inference, Vol. 136: 942-961.

[180] Peña D. (2002). "Análisis de datos multivariantes". McGraw-Hill.

[181] Pérez C. (2009). "Técnicas Estadísticas Multivariantes con SPSS". Garceta

[182] Peter J.P., Churchill G. A. and Brown T.J. (1993). "Caution in the Use of Difference Scores in Consumer Research". Journal of Consumer Research, 19(March), 655-62.

[183] Phillips L.W., Chang D.R. and Buzzell R.D. (1983). "Product Quality, Cost Position and Business Performance: A Test of Some Key Hypothesis". Journal of Marketing, 47 (Spring), 26-43.

[184] Ping R.A. Jr. (1995). “A Parsimonious Estimating Technique for Interaction and Quadratic Latent Variables”. Journal of Marketing Research, 32, 336-347.

[185] Pitt L.F., Oosthuizen P. and Morris M.H. (1992). “Service Quality in a High Tech Industrial Market: An Application of SERVQUAL”. Chicago: American Management Association.

Q

R

[186] Rao C.R. and Toutenburg, H. (1999). “Linear Models: Least Squares and Alternatives”.2nd Edition. Springer-Verlag, New York.

[187] Rapert M.I., Smith S.V. and Garretson J.A. (2004). “The Meaning of Quality: Expectations of Students in Pursuit of an MBA”. Journal of Education for Business, Vol.80, No.1, pp.17-24.

[188] Ravi S. B., Warren W. F. and Jos, G.A.M. (2002). “Modelling and Evaluating Service Quality Measurement Using Neural Networks”. International Journal of Operations and Production Management, Vol. 22, No.10, pp.1162-1185.

[189] Ray C. S. and Jeon Y. (2003). “Reputation and Efficiency: A Nonparametric Assessment of America’s Top-Rated MBA Programs”. Working Paper 2003-13, March, JEL Classification: I2, D6, N3, http://www.econ.uconn.edu/.

[190] Rencher A.C. (2002). “Methods of multivariate analysis”. John Wiley & Sons.

[191] Ritzer G. (1996). “McUniversity in the Postmodern Consumer Society”. Quality in Higher Education, 2 (No.3), 185-199.

[192] Rossi P.E., Gilula Z. and Allenby G.M. (2001). "Overcoming scale usage heterogeneity: a Bayesian hierarchical approach". Journal of the American Statistical Association. 96, 20-31.

[193] Rust R.T. and Oliver R.L. (1994). "Service quality: insights and managerial implications from the frontier in Service Quality: New Directions in Theory and Practice". Rust R. & Oliver R. (Eds). Thousand Oaks, CA, SAGE Publications, 1-19.

[194] Rust R.T. and Metters D. (1996). "Mathematical Models of Service". European Journal of Operations Research 91, 427-439.

S

[195] Sahney S., Banwet D.K. and Karunes P. (2004). "Customer Requirement Constructs: The premise for TQM in education". International Journal of Productivity and Performance Management, Vol.53, No.6, pp.499-520.

[196] Schott J.M. (1992). "Matrix Analysis for Statistics". Wiley Series in Probability and Statistics. N.Y.

[197] Schumaker R.E. and Marcoulides G.A. (1992). "Interaction and Non-Linear Effects in Structural Equation Modeling". Mahwah, NJ: Erlbaum.

[198] Seth N., Deshmukh S.G. and Vrat P. (2005). "Service Quality Models: A Review". International Journal of Quality & Reliability Management, Vol.22, No.9, pp.913-919.

[199] Shanahan P. and Gerber R. (2004). "Quality in University Student Administration: Stakeholders' Conceptions". Quality Assurance in Education, Vol.12, No.4, pp.166-174.

[200] Shaw J. (1978). "The Quality - Productivity Connection". New York: Van Nostrand.

[201] Silverman B. W. (1986). "Density Estimation for Statistics and Data Analysis". Chapman and Hall. N.Y.

[202] Simpson P. and Siguaw J. (2000). "Student evaluations of teaching: An exploratory study of the faculty response". Journal of Marketing Education, 22 (No.3), 199-213.

[203] Singh D. (1986). "Theory and analysis of sample survey designs". John Wiley and Sons.

[204] Smith R.A. and Houston M.J. (1982). "Script-based Evaluations of Satisfaction with Services". in Berry, L, Shostack, G and Upah, G (eds.), Emerging Perspectives on Services Marketing, Chicago: American Marketing Association, 59-62.

[205] Spicer J. (2005). "Making sense of multivariate data analysis". Sage Publications.

[206] Spreng R.A. and Singh A. K. (1993). "An Empirical Assessment of the SERVQUAL Scale and the Relationship between Service Quality and Satisfaction". in Peter, D W, Cravens, R and Dickson (eds.), Enhancing Knowledge Development in Marketing, Chicago, IL: American Marketing Association, 1-6.

[207] Stafford T. (1994). "Consumption Values and the Choice of Marketing Electives: treating Students Like Customers". Journal of Marketing Education, 16 (Summer):,26-33.

[208] Stamatis, D.H. (2003). "Failure Mode and Effect Analysis: FMEA from Theory to Execution". ASQ Quality Prss.

[209] Stoer J. and Bulirsch R. (1980). "Introduction to Numerical Analysis". Springer-Verlag. New York.

[210] Stork D.G. (2004). "Computer Manual in MATLAB to Accompany Pattern Classification". *2nd ed.* Wiley.

[211] Sureshchandar G.S., Rajendran C.H. and Anantharaman R.N. (2002). "The Relationship between Management´s Perception of Total Quality Service and Customer Perceptions of Service Quality". Total Quality Management, 13/1, 69-88.

T

[212] Taguchi G. (2005). "Taguchi's Quality Engineering Handbook". John Wiley & Sons.

[213] Teas R.K. (1993). "Expectations, Performance Evaluation, and Consumers´ Perceptions of Quality". Journal of Marketing, Vol. 57 (October): 18-34.

[214] Teas R.K. (1994). "Expectations as a Comparison Standard in Measuring Service Quality: An Assessment of a Reassessment". Journal of Marketing, Vol. 58. (January): 132-139.

[215] Temponi C. (2005). "Continuous Improvement Framework: Implications for Academia". Quality Assurance in Education, Vol.13, No.1, pp. 17-36.

[216] Trezise I., Stoney E.G., Bishop B., Eren J., Harkness A., Langdon C. and Mulder T. (2006). "Report of the road safety committee on the inquiry into driver distraction". Rep. No. 209. Melbourne, Victoria, Australia: Road Safety Committee, Parliament of Victoria.

[217] Tronci R., Giacinto G. and Roli F. (2007). "Index Driven Combination of Multiple Biometric Experts for AUC Maximisation". Multiple Classifier Systems, 357–366.

U

V

[218] Van Loan C.F. (2010). "Insight Through computing: a MATLAB Introduction to Computational Science and Engineering". Cornell University.

W

[219] Wang J.S., Knipling R.R. and Goodman M.J. (1996). "The role of driver inattention in crashes; new statistics from the 1995 Crashworthiness Data System". Proceedings of 40th Annual meeting of the Association for the Advancement of Automotive Medicine.

[220] Wedel M. and DeSarbo W. (1994). "A Review of Recent Developments in Latent Class Regression Models". In Bagozzi, R. (Editors) Advanced Methods of Marketing Research, I-II. Blackwell Publishers, Massachusetts: 353-388.

[221] Witkowski T. H. and Wolfinbarger M. F. (2002). "Comparative Service Quality: German and American Ratings across Service Settings". Journal of Business Research, 55 (11), 875-81.

[222] Woodruff R.B., Cadotte E.R. and Jenkins R.L. (1983). "Modelling Consumer Satisfaction Processes Using Experience-based Norms". Journal of Marketing Research, 20(August), 296-304.

X

Y

[223] Young C., Cunningham L. and Lee M. (1994). "Assessing Service Quality as an Effective Management Tool: The Case of the Airline Industry". Journal of Marketing Theory and Practice, 2 (Spring), 76-96.

Z

[224] Zeithaml V.A. (1988). "Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence". Journal of Marketing, 52, 2-22.

[225] Zeithaml V.A., Parasuraman A. and Berry L.L. (1990). "Delivering Service Quality: Balancing Customer Perceptions and Expectations". New York: The Free Press.

[226] Zeithaml V.A and Parasuraman A. (1991). "The Nature and Determinants of Customer Expectation of Service". Marketing Science Institute Working Paper No. 91-113, Cambridge, MA: Marketing Science Institute.

[227] Zeithaml V.A. and Parasuraman A. (1996). "The Behavioral Consequences of Service Quality". Journal of Marketing, 60(April), 31-46.

[228] Zeithaml V.A. and Bitner M. J. (2001). "Services Marketing: Integrating Customer Focus Across the Firms". 2nd Edition, Boston: Tata-McGraw Hill.

[229] Zemke R. (1999). "Best practices in customer service". New York, NAL Books.

Appendix A

Glossary

LIST OF ABREVIATIONS	
5S	Sifting, Sorting, Sweeping, Standardize, Sustain
6S	Six Sigma
ACSI	American Customer Satisfaction Index
AHP	Analytic Hierarchy Process
AL	Average Life Span of a Customer
ALR	Adaptive Local Regression
ANOVA	Analysis of Variance
AQL	Acceptable Quality Level
ASQ	The American Society for Quality

LIST OF ABREVIATIONS

BB	Black Belt
BE	Business Excellence
BPR	Business Process Reengineering
BQF	British Quality Foundation
BSC	Balanced Scorecard
CI	Continuous Improvement
CIT	Critical Incident Technique
CLT	Central Limit Theorem
COQ	Cost of Quality
CQI	Continuous Quality Improvement
CTC	Critical To Cost
CTD	Critical To Delivery
CTP	Critical To the Process
CTQ	Critical To Quality
CTS	Critical To Satisfaction
CWQC	Company Wide Quality Control
DFSS	Design For Six Sigma
DMADV	Define, Measure, Analyze, Design, Verify
DMAIC	Define, Measure, Analyze, Improve, Control
DOE	Design of Experiments
DPMO	Defects per Million Opportunities
DR	Desertion Rate
DTI	Department of Trade and Industry
EFQM	European Foundation for Quality Management
EMEA	Error Mode and Effect Analysis

LIST OF ABREVIATIONS

EOQ	European Organization for Quality
EQA	The European Quality Award
EUN TQM	European Universities Network for TQM
EVP	Executive Vice President
FB	Future Dimension of a Business
FMEA	Failure Mode and Effects Analysis
FMECA	Failure Mode and Effects Criticality Analysis
GB	Green Belt (Six Sigma context)
GLSM	Generalized Least Squared Method
GOS	Grade of Service
HoQ	House of Quality
ID	Interrelationship Digraph
ISO	International Organization for Standardization
IQNET	International Quality Network
JIT	Just In Time
JUSE	Japanese Union of Scientists and Engineers
KM	Knowledge Management
KPI	Key Performance Indicators
LO	Learning Organization
LCL	Lower Control Limit
LSM	Least Squared Method
MADM	Multi-Attribute Decision-making
MBB	Master Black Belt
MBNQA	Malcolm Baldrige National Quality Award
MCDA	Multi Criteria Decision Aid

LIST OF ABREVIATIONS

MCDM	Multi Criteria Decision-Making
MODM	Multi Objective Decision Making
MOS	Mean Opinion Score value
MSA	Measurement Systems Analysis
OB	Organizational Behaviour
NGT	Nominal Group Technique
OE	Operational Effectiveness
OLS	Ordinary Least Squares
OM	Operations Management
ONAC	Office of the National Accreditation Council
OS	Operations Strategy
PDCA	Plan, Do, Check, Act Cycle
PDPC	Process Decision Program Chart
PDSA	Plan, Do, Study, Act Cycle
PE	Process Excellence
PM	Project Manager
PZB	Parasuraman, Zeithaml and Berry
QA	Quality Assurance
QC	Quality Control
QCC	Quality Control Circle
QFD	Quality Function Deployment
QM	Quality Management
QMS	Quality Management System
QoE	Quality of Experience
QoS	Quality of Service

LIST OF ABREVIATIONS

QSHE	Quality, Safety, Health, and Environment
R&R	Repeatability and Reproducibility
ROI	Return On Investment
RPN	Risk Priority Number
RR	Retention Rate
RSM	Response Surface Method
RTY	Rolled Throughput Yield
SAW	Simple Additive Weighting
SDCA	Standardization, Do, Check, Act
SIPOC	Supplier Input Process Output Customer
SLA	Service Level Agreement
SMED	Single-minute exchange of dies
SME	Small and Medium Size Enterprises
SPC	Statistical Quality Control
SQG	Service Quality Gap model
SS	Six Sigma
SWOT	Strengths, Weaknesses, Opportunities, and Threats
ToS	Type of Service
TPM	Total Productive Maintenance
TQC	Total Quality Control
TQM	Total Quality Management
TQPC	Total Quality Promotion Centre
UCL	Upper Control Limit
VOB	Voice of the Business
VOC	Voice of the Customer

LIST OF ABREVIATIONS

VOP	Voice of the Process
VP	Vice President
WB	White Belt (Six Sigma context)
WBS	Work Breakdown Structure
WSM	Weighted Sum Model

Appendix B

The Product Quality Measures

1. Customer satisfaction index

- Surveyed before product delivery and after product delivery (and on-going on a periodic basis, using standard questionnaires)
- Number of system enhancement requests per year
- Number of maintenance fix requests per year
- User friendliness: call volume to customer service hotline
- User friendliness: training time per new user
- Number of production re-runs (in-house information systems groups)

2. Delivered defect quantities

- Normalized per function point
- At product delivery (first 3 months or first year of operation)
- Ongoing (per year of operation), categorized by level of severity, by category or cause, e.g.: requirements defect, design defect, code defect,

documentation/on-line help defect, defect introduced by fixes, etc.

3. Responsiveness (turnaround time) to users

- Turnaround time for defect fixes, by level of severity
- Time for minor vs. major enhancements; actual vs. planned elapsed time

4. Product volatility

- Ratio of maintenance fixes (to repair the system & bring it into compliance with specifications), vs. enhancement requests (requests by users to enhance or change functionality)

5. Defect ratios

- Defects found after product delivery per function point
- Defects found after product delivery
- Pre-delivery defects: annual post-delivery defects
- Defects per function point of the system modifications

6. Defect removal efficiency

- Number of post-release defects (found by clients in field operation), categorized by level of severity
- Ratio of defects found internally prior to release (via inspections and testing), as a percentage of all defects
- All defects include defects found internally plus externally (by customers) in the first year after product delivery

7. Complexity of delivered product

- Predicted defects and maintenance costs, based on complexity measures

8. Test coverage

- Breadth of functional coverage
- Percentage of paths, branches or conditions that were actually tested
- Percentage by criticality level: perceived level of risk of paths

- The ratio of the number of detected faults to the number of predicted faults.

9. Cost of defects

- Business losses per defect that occurs during operation
- Business interruption costs; costs of work-arounds
- Lost sales and lost goodwill
- Litigation costs resulting from defects
- Annual maintenance cost (per function point)
- Annual operating cost (per function point)
- Measurable damage to your boss's career

10. Costs of quality activities

- Costs of reviews, inspections and preventive measures
- Costs of test planning and preparation
- Costs of test execution, defect tracking, version and change control
- Costs of diagnostics, debugging and fixing
- Costs of tools and tool support
- Costs of test case library maintenance
- Costs of testing & QA education associated with the product
- Costs of monitoring and oversight by the QA organization (if separate from the development and test organizations)

11. Re-work

- Re-work effort (hours, as a percentage of the original coding hours)
- Re-worked LOC (source lines of code, as a percentage of the total delivered LOC)
- Re-worked software components (as a percentage of the total delivered components)

12. Reliability

- Availability (percentage of time a system is available, versus the time the system is needed to be available)

- Mean time between failure (MTBF)
- Mean time to repair (MTTR)
- Reliability ratio (MTBF / MTTR)
- Number of product recalls or fix releases
- Number of production re-runs as a ratio of production runs

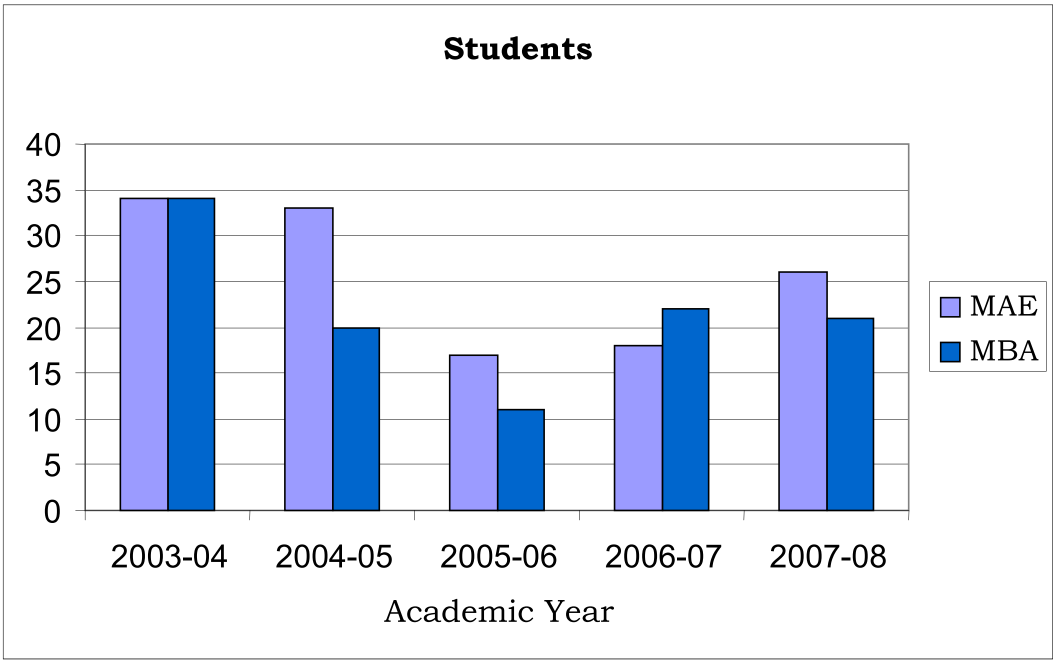
Appendix C

Chapter 4 Additional Results

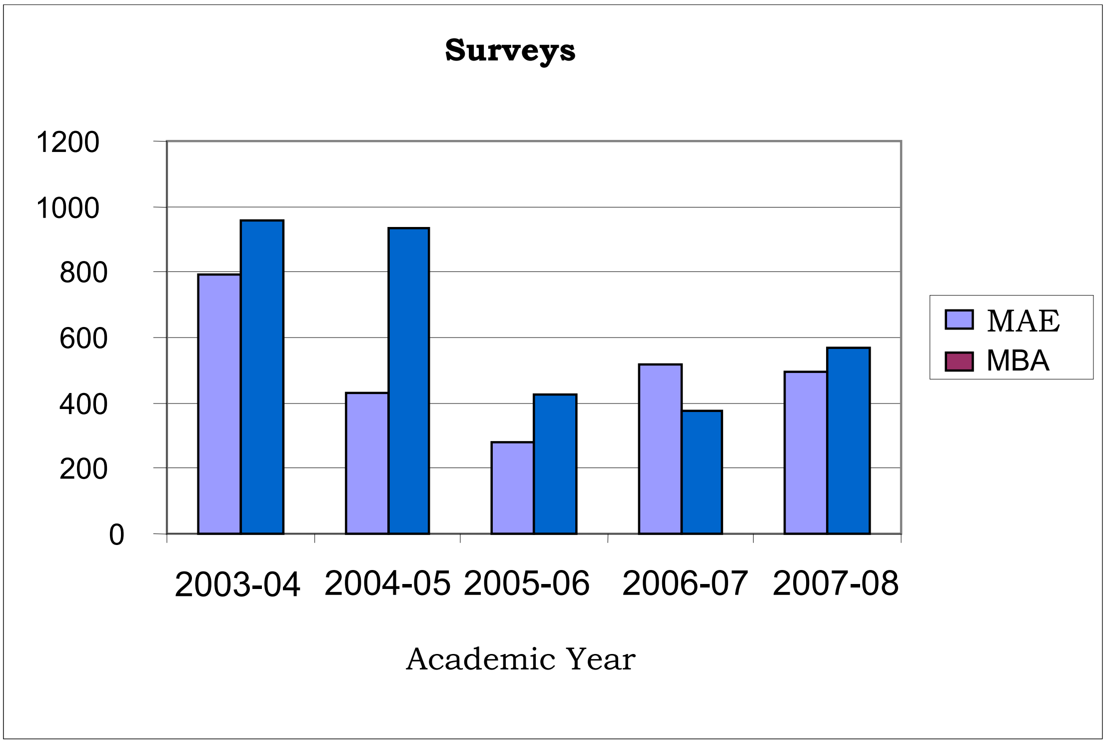
A.4.1 Total Number of Students in the program:

Students

Academic Year	MAE	MBA
2003-2004	34	34
2004-2005	33	20
2005-2006	17	11
2006-2007	18	22
2007-2008	26	21
2008-2009	29	29
	157	**137**



A.4.2 Total Number of the Received Questionnaires:



2003-04

	T1	T2	T3	Total
MAE	223	263	305	791
MBA	314	274	367	955
Total	**537**	**537**	**672**	**1746**

2004-05

	T1	T2	T3	Total
MAE	140	144	147	431
MBA	344	254	335	933
Total	**484**	**398**	**482**	**1364**

2005-06

	T1	T2	T3	Total
MAE	66	89	124	279
MBA	147	136	141	424
Total	**213**	**225**	**265**	**703**

2006-07

	T1	T2	T3	Total
MAE	83	211	224	518
MBA	63	165	147	375
Total	**146**	**376**	**371**	**893**

2007-08

	T1	T2	T3	Total
MAE	151	162	182	495
MBA	138	208	222	568
Total	**289**	**370**	**404**	**1063**

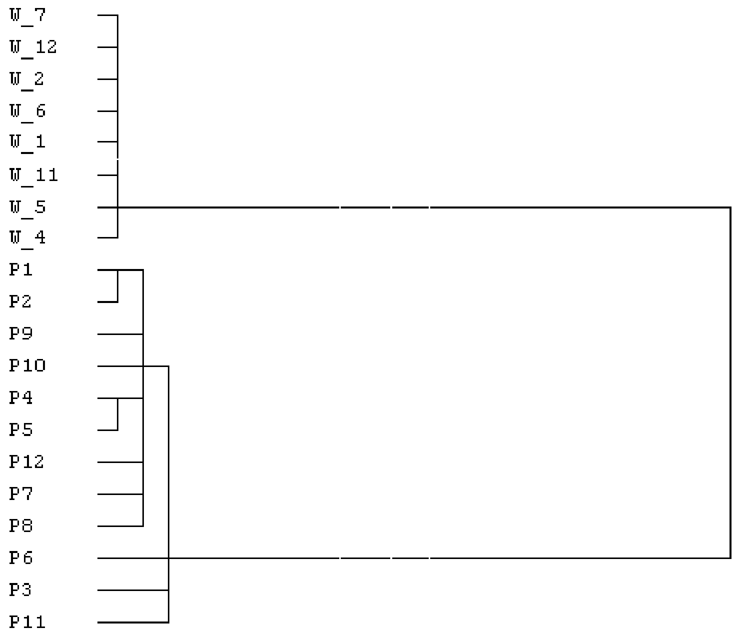
A.4.3 Aggregated results

	0	1 - 2	3	4 - 5
P1	3,7%	6,2%	15,2%	75,0%
P2	4,2%	6,0%	16,1%	73,6%
P3	4,4%	10,3%	20,0%	65,2%
P4	4,4%	13,7%	20,6%	61,2%
P5	4,6%	3,6%	8,9%	82,9%
P6	4,7%	11,0%	20,9%	63,4%
P7	7,1%	12,3%	24,5%	56,1%
P8	8,6%	12,2%	19,6%	59,5%
P9	14,9%	11,1%	19,0%	55,0%
P10	6,9%	9,1%	37,6%	46,4%
P11	5,2%	9,6%	24,3%	60,9%
P12	6,0%	20,6%	28,3%	45,1%

A.4.3.1 Agglomeration Schedule W's and P's variables

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	6	8	,000	0	0	2
2	2	6	,000	0	1	3
3	2	5	,000	2	0	4
4	1	2	,000	0	3	5
5	1	7	,001	4	0	6
6	1	4	,001	5	0	7
7	1	3	,003	6	0	19
8	11	12	,019	0	0	11
9	9	10	,021	0	0	15
10	16	17	,027	0	0	12
11	11	19	,031	8	0	12
12	11	16	,032	11	10	13
13	11	14	,034	12	0	14
14	11	15	,036	13	0	15
15	9	11	,037	9	14	16
16	9	13	,049	15	0	17
17	9	20	,055	16	0	18
18	9	18	,063	17	0	19
19	1	9	,579	7	18	0

A.4.3.2 Dendrogram W's and P's variables



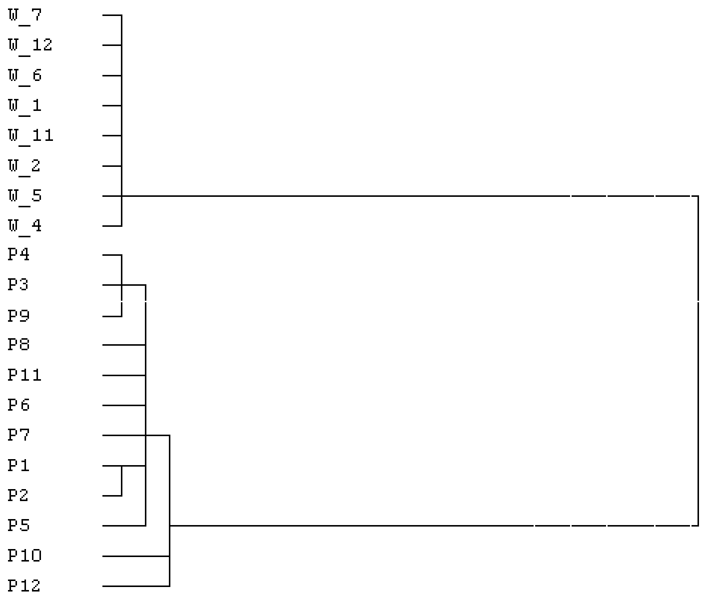
A.4.4 MAE results

	0	1 - 2	3	4 - 5
P1	3,6%	3,5%	10,5%	82,4%
P2	3,9%	4,1%	12,1%	80,0%
P3	4,0%	7,7%	18,4%	69,9%
P4	4,1%	10,7%	18,9%	66,3%
P5	4,1%	2,8%	8,0%	85,0%
P6	4,2%	7,6%	17,8%	70,4%
P7	6,4%	9,3%	22,6%	61,6%
P8	7,1%	9,5%	18,2%	65,2%
P9	14,3%	8,0%	17,6%	60,1%
P10	6,1%	7,9%	35,4%	50,6%
P11	4,6%	7,3%	22,0%	66,1%
P12	5,1%	20,0%	28,3%	46,6%

A.4.4.1 Agglomeration Schedule for W's and P's variables

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	14	16	106,777	0	0	2
2	13	14	108,996	0	1	3
3	9	13	110,949	0	2	4
4	9	15	120,013	3	0	5
5	9	10	130,455	4	0	6
6	9	12	140,558	5	0	7
7	9	11	200,518	6	0	19
8	3	17	770,783	0	0	10
9	1	2	804,492	0	0	15
10	3	19	966,192	8	0	11
11	3	18	1027,929	10	0	12
12	3	7	1077,181	11	0	13
13	3	5	1257,131	12	0	14
14	3	6	1428,860	13	0	16
15	1	4	1495,494	9	0	16
16	1	3	1504,768	15	14	17
17	1	20	2232,973	16	0	18
18	1	8	2641,905	17	0	19
19	1	9	22985,212	18	7	0

A.4.4.2 Dendrogram W's and P's variables



A.4.5 MBA Results

	0	1 - 2	3	4 - 5
P1	3,7%	8,2%	18,8%	69,3%
P2	4,5%	7,5%	19,3%	68,7%
P3	4,8%	12,4%	21,3%	61,6%
P4	4,7%	16,1%	22,0%	57,2%
P5	5,0%	4,1%	9,6%	81,3%
P6	5,0%	13,6%	23,4%	58,0%
P7	7,6%	14,6%	26,0%	51,9%
P8	9,8%	14,3%	20,7%	55,1%
P9	15,4%	13,4%	20,2%	51,0%
P10	7,5%	10,1%	39,3%	43,1%
P11	5,6%	11,3%	26,2%	56,9%
P12	6,7%	21,0%	28,2%	44,0%

A.4.5.1 Agglomeration Schedule for W's and P's variables

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	6	8	,000	0	0	2
2	2	6	,000	0	1	3
3	2	5	,000	2	0	4
4	1	2	,001	0	3	5
5	1	7	,001	4	0	6
6	1	4	,001	5	0	7
7	1	3	,003	6	0	19
8	11	12	,022	0	0	11
9	9	10	,024	0	0	13
10	16	17	,031	0	0	12
11	11	19	,035	8	0	12
12	11	16	,038	11	10	13
13	9	11	,038	9	12	14
14	9	15	,039	13	0	15
15	9	14	,039	14	0	16
16	9	20	,048	15	0	17
17	9	13	,062	16	0	18
18	9	18	,072	17	0	19
19	1	9	,572	7	18	0

A.4.5.2 Dendrogram W's and P's variables