



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA  
INGENIERÍA DE SISTEMAS – SISTEMAS DE CALIDAD

COMPETITIVIDAD Y ESTRATEGIAS DE CALIDAD:  
USO DEL ANÁLISIS MATRICIAL DE DATOS EN LA EVALUACIÓN DE LA  
SATISFACCIÓN DEL CLIENTE

TESIS  
QUE PARA OPTAR POR EL GRADO DE:  
MAESTRO EN INGENIERÍA

PRESENTA:  
RAMÍREZ LUNA JUVENTINO

TUTOR PRINCIPAL  
MARÍA DE LOS ÁNGELES OLVERA TREVIÑO  
Facultad de Química

MÉXICO, D. F. SEPTIEMBRE 2013



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**JURADO ASIGNADO:**

**Presidente:** Dra. Lara Zavala Nydia Guadalupe

**Secretario:** Dr. Barragán Ocaña Alejandro

**Vocal:** Dr. Sámano Castillo José Sabino

**1er. Suplente:** Dra. Olvera Treviño María de los Ángeles

**2do. Suplente:** M.C. Alpizar Ramos María del Socorro

Lugar donde se realizó la tesis: **Ciudad de México, Distrito Federal**

**TUTOR DE TESIS:**

Dra. Olvera Treviño María de los Ángeles

---

**FIRMA**

*Dedicatoria*

*A mi madre, a mi esposa, a mi hija.*

## ***Resumen***

---

El Análisis Matricial de Datos reduce el número de variables necesarias para construir los modelos utilizados en diversos campos del conocimiento y puede ser una de las más valiosas herramientas del gerenciamiento empresarial.

En el capítulo uno se menciona la relevancia de medir la satisfacción y la lealtad del cliente, haciendo énfasis en el uso de metodologías que le dan certidumbre y confiabilidad a la información recolectada.

El segundo capítulo presenta las estrategias del Premio Nacional de Calidad y de la Norma ISO 9004, las herramientas estadísticas sugeridas por ISO/TR 10017, las herramientas administrativas de la Unión Japonesa de Científicos e Ingenieros (JUSE), así como el uso de estas herramientas para incrementar la calidad y productividad en diversas industrias de México.

La metodología seleccionada en el capítulo tres permite condensar las variables utilizadas al medir la satisfacción del cliente para facilitar su interpretación, sin perder la información obtenida.

Los resultados obtenidos al aplicar esta metodología y el modelo que propone son presentados en el capítulo cuatro.

En el capítulo cinco se discuten y concluyen las aportaciones del modelo en la solución de la problemática que se presenta al evaluar la satisfacción del cliente.

# Índice

Resumen.....	iv
Índice.....	v
Lista de figuras.....	viii
Lista de tablas.....	x
<b>Capítulo 1</b>	
<b>Introducción.....</b>	<b>1</b>
1.1 <b>Antecedentes.....</b>	<b>1</b>
1.2 <b>Problema.....</b>	<b>3</b>
1.3 <b>Objetivo General.....</b>	<b>5</b>
1.4 <b>Hipótesis.....</b>	<b>5</b>
<b>Capítulo 2</b>	
<b>Marco Teórico.....</b>	<b>6</b>
2.1 <b>Principios de Gestión.....</b>	<b>6</b>
2.1.1    Liderazgo.....	7
2.1.2    Participación del personal.....	7
2.1.3    Enfoque basado en procesos.....	8
2.1.4    Enfoque de sistema para la gestión.....	8
2.1.5    Mejora continua.....	8
2.1.6    Relaciones mutuamente beneficiosas con el proveedor.....	9
2.1.7    Enfoque al cliente.....	9
2.1.8    Enfoque basado en hechos para la toma de decisión...	11
2.2 <b>Técnicas estadísticas sugeridas por ISO.....</b>	<b>12</b>
2.2.1    Estadísticos descriptivos.....	12
2.2.2    Prueba de hipótesis.....	13
2.2.3    Análisis de regresión.....	13
2.2.4    Muestreo.....	13
2.2.5    Análisis de la capacidad del proceso.....	14
2.2.6    Gráficos de control estadístico.....	14
2.2.7    Análisis de series de tiempo.....	15
2.2.8    Tolerancia estadística.....	15
2.2.9    Simulación.....	15
2.2.10   Análisis de confiabilidad.....	16
2.2.11   Análisis de incertidumbre.....	16
2.2.12   Diseño de experimentos.....	17
2.3 <b>Herramientas básicas.....</b>	<b>18</b>
2.3.1    Recolección de datos.....	18
2.3.2    Estratificación.....	19
2.3.3    Gráfico de Pareto.....	19

	2.3.4	Diagrama de causa y efecto.....	20
	2.3.5	Gráfico de dispersión.....	21
	2.3.6	Histograma.....	21
	2.3.7	Gráficos de control.....	22
<b>2.4</b>		<b>Herramientas administrativas .....</b>	<b>24</b>
	2.4.1	Diagrama matricial.....	24
	2.4.2	Diagrama de árbol.....	25
	2.4.3	Diagrama de afinidad.....	26
	2.4.4	Diagrama de relaciones.....	27
	2.4.5	Diagrama de flechas.....	27
	2.4.6	Diagrama de programación de decisiones en el proceso.....	28
	2.4.7	Análisis matricial de datos.....	29
<b>2.5</b>		<b>Uso de las herramientas básicas y administrativas en México.....</b>	<b>30</b>
<b>2.6</b>		<b>Análisis Matricial de datos.....</b>	<b>36</b>
	2.6.1	Antecedentes del análisis matricial de datos.....	36
	2.6.2	Conceptos fundamentales.....	38
	2.6.3	Reducción de variables.....	40
<b>Capítulo 3</b>		<b>Metodología.....</b>	<b>45</b>
	3.1	<b>Análisis exploratorio y gráfico de las variables.....</b>	<b>48</b>
	3.2	<b>Análisis exploratorio y gráfico de los eventos.....</b>	<b>50</b>
	3.2.1	Tratamiento de datos ausentes.....	51
	3.2.2	Tratamiento de datos atípicos.....	53
	3.3	<b>Matriz de unidades experimentales y variables. Comprobación de la correlación supuesta .....</b>	<b>54</b>
	3.4	<b>Extracción de factores.....</b>	<b>56</b>
	3.5	<b>Determinación del número de factores.....</b>	<b>59</b>
	3.6	<b>Rotación de factores.....</b>	<b>59</b>
	3.7	<b>Interpretación de los factores.....</b>	<b>60</b>
	3.8	<b>Matriz de unidades experimentales y factores.....</b>	<b>60</b>
	3.9	<b>Validación del modelo.....</b>	<b>61</b>
<b>Capítulo 4</b>		<b>Resultados.....</b>	<b>62</b>
	4.1	<b>Análisis exploratorio y gráfico de los datos.....</b>	<b>62</b>
	4.2	<b>Análisis exploratorio y gráfico de los eventos.....</b>	<b>94</b>
	4.3	<b>Matriz de unidades experimentales y variables. Comprobación de la correlación supuesta .....</b>	<b>100</b>
	4.4	<b>Extracción de factores.....</b>	<b>110</b>
	4.5	<b>Determinación del número de factores.....</b>	<b>112</b>
	4.6	<b>Rotación de factores.....</b>	<b>114</b>

<b>4.7</b>	<b>Interpretación de los factores.....</b>	<b>115</b>
4.7.1	Interpretación del primer factor.....	116
4.7.2	Interpretación del segundo factor.....	118
4.7.3	Interpretación del tercer factor.....	119
<b>4.8</b>	<b>Matriz de unidades experimentales y factores.....</b>	<b>119</b>
<b>4.9</b>	<b>Validación del modelo.....</b>	<b>124</b>
<b>Capítulo 5</b>	<b>Conclusiones.....</b>	<b>126</b>
<b>Bibliografía.....</b>		<b>130</b>

## *Lista de figuras*

---

Figura 2.5.1	Equipos del Concurso Nacional que utilizaron la gráfica de Pareto y diagrama causa- efecto 1991-2002.....	33
Figura 2.5.2	Equipos que utilizan herramientas de control estadístico del proceso y la gráfica de radar 1991-2002.....	34
Figura 2.5.3	Equipos finalistas que utilizan las herramientas administrativas 1991-2002 .....	35
Figura 2.6.1	Interpretación geométrica de la reducción de variables.....	41
Figura 3.4.1	Interpretación geométrica de la extracción de factores.....	56
Figura 3.6.1	Interpretación geométrica de la rotación de factores.....	59
Figura 4.1.1	Histograma de la variable 1.....	64
Figura 4.1.2	Grafica de caja y bigote para la variable 1.....	65
Figura 4.1.3	Grafica de normalidad para la variable 1.....	65
Figura 4.1.4	Histograma de la variable 2.....	66
Figura 4.1.5	Grafica de caja y bigote para la variable 2.....	67
Figura 4.1.6	Grafica de normalidad para la variable 2.....	67
Figura 4.1.7	Histograma de la variable 3.....	68
Figura 4.1.8	Grafica de caja y bigote para la variable 3.....	69
Figura 4.1.9	Grafica de normalidad para la variable 3.....	69
Figura 4.1.10	Histograma de la variable 4.....	70
Figura 4.1.11	Grafica de caja y bigote para la variable 4.....	71
Figura 4.1.12	Grafica de normalidad para la variable 4.....	71
Figura 4.1.13	Histograma de la variable 5.....	72
Figura 4.1.14	Grafica de caja y bigote para la variable 5.....	73
Figura 4.1.15	Grafica de normalidad para la variable 5.....	73
Figura 4.1.16	Histograma de la variable 6.....	74
Figura 4.1.17	Grafica de caja y bigote para la variable 6.....	75
Figura 4.1.18	Grafica de normalidad para la variable 6.....	75
Figura 4.1.19	Histograma de la variable 7.....	76
Figura 4.1.20	Grafica de caja y bigote para la variable 7.....	77
Figura 4.1.21	Grafica de normalidad para la variable 7.....	77

<b>Figura 4.1.22</b>	<b>Histograma de la variable 8.....</b>	<b>78</b>
<b>Figura 4.1.23</b>	<b>Grafica de caja y bigote para la variable 8.....</b>	<b>79</b>
<b>Figura 4.1.24</b>	<b>Grafica de normalidad para la variable 8.....</b>	<b>79</b>
<b>Figura 4.1.25</b>	<b>Histograma de la variable 9.....</b>	<b>80</b>
<b>Figura 4.1.26</b>	<b>Grafica de caja y bigote para la variable 9.....</b>	<b>81</b>
<b>Figura 4.1.27</b>	<b>Grafica de normalidad para la variable 9.....</b>	<b>81</b>
<b>Figura 4.1.28</b>	<b>Histograma de la variable 10.....</b>	<b>82</b>
<b>Figura 4.1.29</b>	<b>Grafica de caja y bigote para la variable 10.....</b>	<b>83</b>
<b>Figura 4.1.30</b>	<b>Grafica de normalidad para la variable 10.....</b>	<b>83</b>
<b>Figura 4.1.31</b>	<b>Histograma de la variable 11.....</b>	<b>84</b>
<b>Figura 4.1.32</b>	<b>Grafica de caja y bigote para la variable 11.....</b>	<b>85</b>
<b>Figura 4.1.33</b>	<b>Grafica de normalidad para la variable 11.....</b>	<b>85</b>
<b>Figura 4.1.34</b>	<b>Histograma de la variable 12.....</b>	<b>86</b>
<b>Figura 4.1.35</b>	<b>Grafica de caja y bigote para la variable 12.....</b>	<b>87</b>
<b>Figura 4.1.36</b>	<b>Grafica de normalidad para la variable 12.....</b>	<b>87</b>
<b>Figura 4.1.37</b>	<b>Histograma de la variable 13.....</b>	<b>88</b>
<b>Figura 4.1.38</b>	<b>Grafica de caja y bigote para la variable 13.....</b>	<b>89</b>
<b>Figura 4.1.39</b>	<b>Grafica de normalidad para la variable 13.....</b>	<b>89</b>
<b>Figura 4.1.40</b>	<b>Histograma de la variable 14.....</b>	<b>90</b>
<b>Figura 4.1.41</b>	<b>Grafica de caja y bigote para la variable 14.....</b>	<b>91</b>
<b>Figura 4.1.42</b>	<b>Grafica de normalidad para la variable 14.....</b>	<b>91</b>
<b>Figura 4.1.43</b>	<b>Histograma de la variable 15.....</b>	<b>92</b>
<b>Figura 4.1.44</b>	<b>Grafica de caja y bigote para la variable 15.....</b>	<b>93</b>
<b>Figura 4.1.45</b>	<b>Grafica de normalidad para la variable 15.....</b>	<b>93</b>
<b>Figura 4.2.1</b>	<b>Graficas de radar para comparar 15 variables en 24 eventos...</b>	<b>95</b>
<b>Figura 5.2.2</b>	<b>Grafica de Andrews para <math>r = 1</math>.....</b>	<b>98</b>
<b>Figura 5.2.3</b>	<b>Graficas de Andrews para los primeros 15 eventos.....</b>	<b>98</b>
<b>Figura 5.2.4</b>	<b>Graficas de Andrews para los primeros 20 eventos.....</b>	<b>99</b>
<b>Figura 5.2.5</b>	<b>Graficas de Andrews para los primeros 25 eventos.....</b>	<b>99</b>
<b>Figura 5.5.1</b>	<b>Grafica de Sedimentación.....</b>	<b>112</b>

## *Lista de tablas*

---

Tabla 4.2.1	Matriz de datos para 25 unidades experimentales y 12 variables.....	94
Tabla 4.3.1	Matriz de datos original con 287 unidades experimentales y 15 variables.....	101
Tabla 4.3.2	Matriz de correlaciones para 15 variables.....	107
Tabla 4.3.3	Matriz de componentes para 15 variables.....	108
Tabla 4.3.4	Matriz de correlaciones para 14 variables.....	109
Tabla 4.4.1	Matriz de varianza explicada.....	111
Tabla 4.5.1	Matriz de componentes y cargas de cada variable.....	113
Tabla 4.6.1	Matriz de componentes rotados.....	114
Tabla 4.6.2..	Comparación de la varianza explicada, antes y después de la rotación.....	115
Tabla 4.7.1	Cargas salientes en cada componente principal.....	116
Tabla 4.8.1	Matriz de puntuaciones factoriales.....	120
Tabla 4.8.2	Matriz de factores resultado del análisis de datos.....	121
Tabla 4.9.1	Matriz de correlaciones reproducidas.....	125
Tabla 4.9.2	Matriz de residuos de correlaciones observadas y reproducidas.....	126

# Capítulo 1

## Introducción

---

### 1.1 Antecedentes.

Las naciones de todo el mundo están realizando grandes esfuerzos para reactivar sus economías e incrementar el empleo. Esta reactivación requiere de estrategias orientadas a tomar un papel activo en un entorno mucho más dinámico. Incluyen la identificación de áreas de oportunidad y el desarrollo de sus soluciones para superar a las organizaciones del mismo ramo en:

- la calidad de los productos y servicios,
- los costos de producción y distribución,
- la productividad y el éxito comercial
- su rentabilidad.

En pocas palabras, lo que se requiere es aumentar la *competitividad* de las organizaciones para poder desarrollarse en cualquier parte del mundo.

Al hablar de productividad el comité técnico del Premio Nacional de Calidad afirma: “La principal fuerza de esta transformación es la creación de valor superior para los clientes y consumidores finales a través de la mejora continua de los sistemas y procesos, así como de los productos, bienes y servicios” (Modelo Nacional para la Calidad 2004).

El Modelo Nacional para la Calidad Total clasifica a las organizaciones en cinco niveles: inicial, desarrollo, confiable, competitivo y clase mundial. El término “nivel” refiere al grado alcanzado por una magnitud en comparación con un valor que se toma como referencia: “Es el posicionamiento en el que se encuentra el

desempeño de una organización con relación a un indicador específico, en un momento determinado.” (MNC 2004). Los niveles requieren de referencias objetivas para poder calificarse.

El nivel de clase mundial se refiere a que los resultados obtenidos, así como el logro de los objetivos establecidos con base en los proyectos programados. Son comparables al desempeño de las organizaciones líderes, sean o no competencia directa.

Una organización de clase mundial entrega sus productos y servicios a través de procesos competitivos para responder a los requerimientos, necesidades y expectativas de sus clientes y logra los objetivos estratégicos de la organización.

Es frecuente que el éxito dependa de la forma en que la organización actualiza su conocimiento sobre las necesidades y preferencias de sus clientes y mercados, en el corto y largo plazo; así como de aprovechar las oportunidades para adelantarse a las expectativas de sus clientes y mejorar su posición frente a la competencia.

La competencia por los mercados se manifiesta de forma tan contundente y las reglas del negocio han cambiado tanto que no se puede aspirar a mejorar la competitividad de las organizaciones con la forma de pensar tradicional, basada en la intuición o en prueba y error, sin la aplicación sistémica de las mejores metodologías y herramientas disponibles.

## *1.2 Problema.*

Una empresa industrial del sector de alimentos y bebidas establece que sus clientes son, entre otros, los distribuidores minoristas de sus productos. La empresa está certificada en ISO 9001: 2008 y declara que gestionar la satisfacción del cliente es un proceso necesario para su sistema de gestión de la calidad.

La empresa estableció como instrumento para medir la satisfacción de sus clientes una encuesta con quince reactivos, a los que pide otorgar una calificación de 0 a 10 (donde el cero representa el menor grado de satisfacción y el diez la mayor calificación que se puede otorgar). La encuesta está dirigida al total de distribuidores minoristas y se aplica dos veces al año.

Para estimar la evaluación global otorgada por cada cliente a la empresa, se promedian las calificaciones registradas en los quince reactivos.

Para tener una calificación por cada zona de atención al mercado o grupo de clientes, se promedian los resultados obtenidos de acuerdo al párrafo anterior.

Si alguno de los clientes evaluó a la empresa con una calificación menor al 80% en cualquiera de los temas, el sistema de gestión de calidad genera una acción de corrección inmediata.

El promedio por segmento del mercado es declarado como el indicador de desempeño y puede dar origen a acciones correctivas si baja del 80% de cumplimiento o acciones preventivas, si se tienen tres valores consecutivos con tendencia a la baja, aún en zona de cumplimiento.

Por otro lado, también se promedia cada uno de los quince temas de la encuesta de todos los clientes del área, para tener un indicador por cada tema y se aplican los mismos criterios de acciones de corrección, preventivas y correctivas.

La empresa considera que puede mejorar la interpretación de la información recibida en la encuesta de satisfacción del cliente y señala tres puntos a resolver:

- Medir la satisfacción del cliente con mayor certidumbre,
- Mejorar la interpretación de los resultados y
- Justificar la implementación de acciones correctivas y preventivas, utilizando en forma eficaz y eficiente los recursos disponibles.

El problema está en la interpretación de la encuesta, que no considera relevantes las interacciones que pudieran existir entre los quince temas evaluados y puede superarse si se logra mostrar la diversidad de resultados prácticos que da un estudio multidimensional.

La resistencia a realizar un tratamiento completo se explica por la dificultad matemática de un proceso que forzosamente ha de apoyarse en el cálculo matricial y en técnicas estadísticas avanzadas.

### *1.3 Objetivo General*

Analizar la satisfacción del cliente utilizando el enfoque sistémico del Análisis Matricial de Datos, de tal forma que se obtenga mayor claridad en las conclusiones y, sobre todo, que se facilite definir las líneas de acción para incrementar la competitividad de la organización.

Este objetivo está alineado a los principios de gestión de calidad de la norma NMX-CC-9001-IMNC-2008 (ISO 9001:2008) y su par coherente NMX-CC-9004-IMNC-2009 (ISO 9004:2009).

### *1.4 Hipótesis.*

Si el grado de cumplimiento de los requisitos y expectativas del cliente es presentado como un espacio vectorial de  $n$  variables en  $m$  eventos y este espacio vectorial es susceptible de ser procesado con la herramienta administrativa llamada Análisis Matricial de Datos, entonces se podrán obtener  $f$  variables no observadas o factores (donde  $f < n$ ) que permitirán medir la satisfacción del cliente con mayor certidumbre, mejorar la interpretación de sus resultados y utilizar menos recursos.

## Capítulo 2

### Marco Teórico

---

#### 2.1 Principios de Gestión.

Los sistemas de gestión desarrollados en el ámbito empresarial han permitido mejorar las competencias del personal, de los procesos y de los propios sistemas administrativos. Los sistemas mencionados generalmente se vinculan a la evaluación del desempeño de la organización, con base en el cumplimiento de los requerimientos previamente establecidos, requisitos legales, expectativas explícitas o implícitas, así como en la certificación de tercera parte del cumplimiento de normas voluntarias. Las organizaciones que certifican sus sistemas de gestión bajo las normas ISO tienen bien ganada la reputación de ser altamente confiables.

El enfoque sistémico de la norma NMX-CC-9001-IMNC-2008 (ISO 9001:2008) resalta la interrelación e interdependencia de sus elementos, que permanecen unidos porque continuamente se afectan unos a otros y funcionan por tener un propósito común: la satisfacción del cliente.

La norma NMX-CC-9004-IMNC-2009 (ISO 9004:2009) postula ocho principios de la gestión de calidad y recomienda utilizarlos como marco de referencia para la mejora continua en la organización: liderazgo, participación del personal, enfoque basado en procesos, enfoque de sistema para la gestión, mejora continua, relaciones mutuamente beneficiosas con el proveedor, enfoque al cliente y enfoque basado en hechos para la toma de decisión.

### **2.1.1. Liderazgo.**

*Los líderes establecen la unidad de propósito y la orientación de la organización. Ellos deberían crear y mantener un ambiente interno, en el cual el personal pueda llegar a involucrarse totalmente en el logro de los objetivos de la organización. (ISO 9004:2009)*

El liderazgo se refiere al conjunto de comportamientos que reflejan conocimientos, habilidades y actitudes que impulsan en la organización una cultura fundamentada en sus valores, en su código de ética y en sus estrategias organizacionales (Modelo Nacional para la Calidad 2004).

El liderazgo efectivo define, promueve y asegura que se actúe en congruencia con los principios éticos de la organización, establece contacto directo con los grupos de interés, comunica la información más relevante para lograr el compromiso de todo el personal y así aprovechar su retroalimentación.

### **2.1.2. Participación del personal.**

*El personal, a todos los niveles, es la esencia de una organización y su total compromiso posibilita que sus habilidades sean usadas para el beneficio de la organización. (ISO 9004:2009)*

La organización de clase mundial crea las condiciones necesarias para el desarrollo del personal y mejorar su calidad de vida, como base para el desarrollo organizacional y el logro de sus estrategias (MNC 2004), en particular debe diseñar y mantener esquemas de trabajo individual y grupal que involucren al personal para lograr el mejor desempeño de los procesos y el logro de los objetivos de la organización.

### **2.1.3. Enfoque basado en procesos.**

*Un resultado deseado se alcanza más eficientemente cuando las actividades y los recursos relacionados se gestionan como proceso. (ISO 9004:2009)*

La organización traduce las necesidades y expectativas completas de sus mercados, distribuidores, clientes y usuarios finales en características y especificaciones de productos y servicios (MNC 2004).

Los mecanismos para asegurar el logro consistente de los niveles de desempeño de sus procesos incluye metodologías o técnicas para la medición y control de los procesos, como 6 Sigma, estudios de Capacidad y Habilidad de procesos, estudios AMEF, Buenas Prácticas de Manufactura, Planes de Calidad, Administración Visual, etc.

### **2.1.4. Enfoque de sistema para la gestión.**

*Identificar, entender y gestionar los procesos interrelacionados como un sistema contribuye a la eficiencia de una organización en el logro de sus objetivos. (ISO 9004:2009)*

El diseño de la organización con un enfoque sistémico se refiere a conceptualizar su organización como un sistema, cuyos elementos son subsistemas y procesos interdependientes que se integran (MNC 2004), con el propósito de crear valor a través de su actividad para sus grupos de interés (clientes, personal, accionistas, mercados y sociedad).

### **2.1.5. Mejora continua.**

*La mejora continua del desempeño global de la organización debería ser un objetivo permanente de esta. (ISO 9004:2009)*

Administra la relación laboral para la mejora de los sistemas de trabajo (MNC 2004), se refiere a alentar la cooperación entre los integrantes de la organización, así como con entidades para la valorización y enriquecimiento del trabajo, el cumplimiento de las expectativas de los clientes y mercados, el aumento de la productividad y la satisfacción de las necesidades de los trabajadores y sus familias.

**2.1.6. Relaciones mutuamente beneficiosas con el proveedor.**

*Una organización y sus proveedores son interdependientes, y una relación mutuamente beneficiosa aumenta la capacidad de ambos para crear valor. (ISO 9004:2009)*

La relación integral con los proveedores es la mejor forma en que la organización administra la respuesta a las necesidades y oportunidades detectadas en sus sistemas de control de insumos, promueve la construcción y fortalecimiento de relaciones positivas y de largo plazo con sus proveedores con un espíritu de ganar-ganar.

**2.1.7. Enfoque al cliente.**

*Las organizaciones dependen de sus clientes y por lo tanto deberían comprender sus necesidades actuales y futuras, satisfacer los requisitos de los clientes y esforzarse en exceder sus expectativas. (ISO 9004:2009)*

El conocimiento de los clientes incluye la forma en que la organización examina todas las actividades que el cliente realiza para la búsqueda, adquisición, instalación, uso, mantenimiento y destino final del producto.

También incluye la forma en que su organización evalúa la satisfacción y lealtad de sus clientes. Esta evaluación parte de conceptualizar los segmentos de mercado

como los grupos de clientes y usuarios que comparten ciertas características y rasgos comunes, que son utilizados por la organización para enfocar en forma más precisa sus productos y servicios.

Las organizaciones pueden emplear distintas formas y métodos para el conocimiento y segmentación de sus mercados y clientes con el propósito de identificar los requerimientos particulares en cada uno de ellos.

El Modelo Nacional para la Calidad considera que una organización de clase mundial debe:

- Identificar y conocer sus segmentos de mercado, clientes y clientes potenciales (incluyendo los de la competencia).
- Entender la forma en que los clientes perciben el valor proporcionado por su organización y cómo dicha percepción influye en su preferencia.
- Conocer y anticipar necesidades y expectativas de los distintos grupos de clientes, actuales y potenciales, a mediano y a largo plazo.
- Medir la satisfacción y lealtad de sus clientes. La medición de satisfacción del valor percibido y la lealtad pueden incluir las siguientes prácticas: escala numérica y descripción para cada indicador; experiencias de los clientes al interactuar con toda la organización; indicadores que reflejen posibles comportamientos futuros de los clientes que denoten preferencia por los productos o servicios; indicadores que reflejen cambios en la preferencia por los productos y servicios de la organización.
- Recibir y solucionar efectiva y rápidamente la causa de las quejas e inconformidades de sus clientes, asegurando su satisfacción y recobrando su confianza. La solución efectiva de las causas de quejas, inconformidades e incidencias requiere información de todas las fuentes posibles.

### **2.1.8 Enfoque basado en hechos para la toma de decisión.**

*Las decisiones eficaces se basan en el análisis de los datos y la información. (ISO 9004:2009)*

Hirata afirma “*los datos que se manejan para controlar un proceso productivo pueden ser de dos tipos: los numéricos y los descriptivos. Luego entonces, los hechos pueden ser medidos o descritos según el tipo de datos que se tengan*” (Hirata. **7 Nuevas Herramientas para el Control de Calidad**. 2003).

Los datos numéricos son la consecuencia de la cuantificación de las variables que determinan el resultado y se utilizan para analizar y entender la situación actual de una situación a mejorar o de un problema que se tenga que solucionar.

Para ordenar e interpretar los datos numéricos se utilizan los métodos y técnicas estadísticas y matemáticas, entre las que destacan las Siete Herramientas Básicas para el Control de la Calidad.

Pero no es suficiente con los datos numéricos para explicar los hechos. Con frecuencia se trabaja en una situación que por el momento no es medible. Los datos descriptivos surgen en las fases de diseño, programación y organización de proyectos, soluciones y mejoras como ideas, comentarios u opiniones.

## ***2.2 Técnicas estadísticas sugeridas por ISO.***

El reporte técnico ISO/TE-10017:2005 *Técnicas Estadísticas para la Aplicación de la Norma ISO 9001*, fue desarrollado con la intención de proporcionar una guía para seleccionar las herramientas para aplicar, mantener y mejorar un sistema de la administración de la calidad de acuerdo con los requisitos de ISO 9001, donde el criterio para determinar la necesidad de la técnica y su adecuada selección depende de la propia organización y de su entorno.

Las técnicas estadísticas pueden ayudar a medir, describir, analizar, interpretar y modelar la variabilidad, aún con una cantidad limitada de datos. El análisis estadístico de los datos puede proporcionar una mejor comprensión de la naturaleza, la extensión y las causas de la variabilidad. Esto ayuda a resolver e incluso prevenir los problemas que resultan de tal variabilidad, lo cual ayuda a mejorar continuamente la calidad de productos y procesos para lograr la satisfacción de cliente.

### ***2.2.1. Estadísticos descriptivos.***

*El término estadística descriptiva se refiere a los procedimientos para resumir y presentar los datos cuantitativos de tal forma que muestre las características de la distribución de datos. Las características de mayor interés son el valor central, la dispersión y la forma de su distribución. (Reporte Técnico ISO/TR 10017:2003)*

La información proporcionada por la estadística descriptiva es transmitida rápida y efectivamente por una variedad de métodos gráficos. Estos van de las herramientas relativamente sencillas como la gráfica de tendencia, el gráfico de dispersión y los histogramas a técnicas de naturaleza más compleja como las distribuciones de probabilidad y las distribuciones multivariadas.

### **2.2.2. Prueba de hipótesis**

*Procedimiento estadístico para determinar, con un nivel establecido de riesgo, si un conjunto de datos es compatible con una hipótesis dada. La hipótesis puede ser una suposición del modelo estadístico o puede ser sobre el valor de algún parámetro de la distribución. (Reporte Técnico ISO/TR 10017:2003).*

Una prueba de hipótesis implica decidir si la hipótesis dada, con respecto a un modelo o al parámetro estadístico, debe o no debe ser rechazada. La prueba de hipótesis se usa para comprobar las suposiciones iniciales sobre la distribución de una población, por ejemplo que es normal y los datos de la muestra son aleatorios, o se puede usar para probar si existen diferencias entre dos o más poblaciones.

### **2.2.3. Análisis de regresión**

*El análisis de regresión relaciona el comportamiento de una característica de interés con factores potencialmente causales. Tal relación es especificada por un modelo que puede venir de la ciencia, la economía, la ingeniería o se puede derivar empíricamente. (Reporte Técnico ISO/TR 10017:2003)*

Su objetivo es entender las causas potenciales de la variación en la respuesta y cuánto contribuye cada una al explicar esa variación. Esto se logra relacionando estadísticamente la variación en la variable de respuesta con la variación en las variables potenciales y obtener el mejor modelo que minimice las desviaciones entre el valor calculado y la respuesta obtenida.

### **2.2.4. Muestreo**

*Metodología estadística para obtener información acerca de alguna característica de una población, estudiando una fracción representativa. (Reporte Técnico ISO/TR 10017:2003)*

Un muestreo apropiadamente diseñado ahorra tiempo, disminuye el costo y el trabajo, en comparación con una inspección al 100%. Cuando la inspección del producto implica una prueba destructiva el muestreo es la única manera práctica de obtener información. Usualmente se clasifican en muestreo de aceptación y muestreo de inspección

### **2.2.5. Análisis de la capacidad del proceso**

*Es el examen de la variabilidad y la distribución inherentes de un proceso, para estimar su habilidad de producir el entregable, en el rango de variación permitida por las especificaciones. (Reporte Técnico ISO/TR 10017:2003)*

La capacidad del proceso se expresa como un índice, que relaciona la variabilidad observada en el proceso y la tolerancia permitida por especificaciones. Un índice ampliamente usado de la capacidad de variables que se encuentra centrado entre los límites de la especificación es el Cp que es igual a la tolerancia total dividido entre 6, porque la variabilidad inherente del proceso cuando está en un estado de control estadístico son seis desviaciones estándar.

### **2.2.6. Gráficos de control estadísticos.**

*Son gráficos de datos derivados de las muestras que se toman periódicamente en un proceso, datos que se grafican en forma secuencial para describir la variabilidad inherente al proceso y valorar su estabilidad, comparando los puntos graficados con los límites de control estadístico. (Reporte Técnico ISO/TR 10017:2003)*

Los gráficos de control facilitan la respuesta apropiada para actuar sobre la variación, ayudando al usuario a distinguirse la variación aleatoria de la variación que es debida a causas especiales, para que su corrección ayuden a mejorar el proceso.

### **2.2.7. Análisis de series de tiempo.**

*Familia de métodos y técnicas estadísticas para estudiar una colección de observaciones hechas secuencialmente con el tiempo. El análisis de series de tiempo permite crear modelos que expliquen cómo cada observación está relacionada con las observaciones anteriores. (Reporte Técnico ISO/TR 10017:2003)*

Los modelos que se encuentran suelen ser cíclicos o estacionales y son útiles para entender en qué magnitud los factores causales del pasado podrían repetir su efecto en el futuro.

### **2.2.8. Tolerancia estadística**

*Procedimiento que utiliza la distribución estadística de las dimensiones pertinentes en los componentes de un producto, para determinar la tolerancia total en la unidad armada. (Reporte Técnico ISO/TR 10017:2003)*

Dado un conjunto de tolerancias individuales, el cálculo de la tolerancia estadística total resulta ser una tolerancia significativamente más pequeña que la tolerancia dimensional calculada aritméticamente. La tolerancia estadística permite el uso de tolerancias más amplias para dimensiones individuales que cuando es determinada por el cálculo aritmético. Esto puede ser un beneficio significativo, puesto que las tolerancias mayores se asocian con métodos más sencillos y más rentables de producción.

### **2.2.9. Simulación**

*Término colectivo para los procedimientos por los que un sistema (teórico o empírico) es representado matemáticamente con un programa de cálculo para la solución de un problema. (Reporte Técnico ISO/TR 10017:2003)*

En el contexto de la ciencia teórica, la simulación se utiliza si no se conoce ninguna teoría adecuada para la solución del problema (o si se conoce pero es difícil de aplicar) y cuando obtener la solución se facilita por la gran capacidad de la computadora de realizar operaciones aritméticas. En el contexto empírico, la simulación se usa si el sistema puede ser descrito adecuadamente por un programa de cálculo.

#### **2.2.10. Análisis de confiabilidad.**

*Es la aplicación de la ingeniería y métodos analíticos a la evaluación, predicción y certidumbre en el desempeño de productos o sistemas diseñados para ser libres de problemas al paso del tiempo (Reporte Técnico ISO/TR 10017:2003).*

La confiabilidad es a menudo un factor que influye en la percepción de la calidad del producto y en la satisfacción de cliente. Los beneficios de usar el análisis de confiabilidad incluyen la habilidad de predecir y cuantificar la probabilidad de falla dentro de límites preestablecidos de confianza y así decidir si es necesario usar estrategias de mitigación, tales como la redundancia.

#### **2.2.11. Análisis de incertidumbre**

*Conjunto de procedimientos para evaluar la incertidumbre de sistemas de medida, en el rango de trabajo y bajo las condiciones en que el sistema opera. (Reporte Técnico ISO/TR 10017:2003)*

Con el análisis de incertidumbre se puede valorar, con un nivel preestablecido de confianza, si el sistema de medición es adecuado para el propósito deseado. Para ello se cuantifica la variación de las fuentes que participan, como la ocasionada por la persona que toma la medida, la que puede atribuirse al proceso de medida o la que es propia del instrumento. También se usa para describir la variación

debido al sistema de medición como parte de la variación total del proceso o como parte de la variación admisible total.

#### **2.2.12. Diseño de experimentos.**

*Investigación que se lleva a cabo de una manera planeada y que se apoya de la evaluación estadística de los resultados para obtener conclusiones, con un nivel dado de confianza. (Reporte Técnico ISO/TR 10017:2003)*

Generalmente implica inducir un cambio al sistema bajo investigación y evaluar estadísticamente el efecto de tal cambio. El arreglo y la manera específica en que los experimentos se llevarán a cabo constituye el diseño del experimento, tal diseño está subordinado al objetivo del ejercicio.

Se puede usar para identificar los factores más influyentes en un sistema, la magnitud de su influencia, y de las interacciones entre los factores. La información obtenida se puede usar para facilitar el diseño y el desarrollo de un producto, proceso o mejorar un sistema existente, formulando un modelo matemático que describe la característica de interés como función de los factores influyentes y utilizar este modelo con propósitos de predicción.

No debe dejarse de lado que el reporte técnico ISO/TR-10017:2005 es sumamente fácil de seguir por su estructura, donde describe cada una de las herramientas estadísticas y menciona el punto de la norma ISO-9001:2000 donde se podría utilizar, así como los posibles beneficios de su aplicación.

### **2.3. Herramientas básicas.**

Las estrategias que tuvieron su origen en la actividad empresarial de clase mundial y que actualmente están disponibles para ser utilizadas libremente por cualquier entidad que quiera elevar su competitividad, forman un conjunto de herramientas de aplicación práctica. De ellas las más conocidas son las siete herramientas básicas y las siete herramientas administrativas.

Las herramienta básicas son muy fáciles de usar y su aplicación ayuda a solucionar efectivamente las causas que están generando inconformidades en el proceso o en el mismo producto.

#### **2.3.1. Recolección de Datos.**

*La información es una guía para nuestras acciones. A partir de la información conocemos los hechos pertinentes y adoptamos acciones apropiadas basadas en esos hechos, pero antes de recoger la información es importante determinar qué se va a hacer con ella.* (Yoshinori Iizuka).

Existen dos tipos principales de cartas de recolección de datos numéricos:

Para verificación. Son formatos diseñados para que la información pueda registrarse de manera simple y directa, de tal forma que el resultado pueda ser interpretado rápidamente según se vaya llenando la carta, también son conocidas como listas de verificación.

Para registro. Son formatos que se utilizan para reunir datos por lo general requieren un proceso adicional de análisis, los datos así reunidos quedan ordenados y facilitan su consulta posterior.

### **2.3.2. Estratificación.**

*La estratificación es un método para identificar la fuente de la variación de los datos recogidos, clasificando los datos según varios factores. (Hitoshi Kume).*

A cada grupo de datos se le denomina *estrato*. Una división objetiva considera turno, máquina, operador, materia prima, etc. La estratificación ayuda a analizar aquellos casos en los cuales la información general oculta los hechos relevantes.

En consecuencia facilita el correcto entendimiento de la información, la toma precisa de una acción y proporciona un agrupamiento claro de las causas potenciales del problema, lo que permite centrarse directamente en el análisis del mismo.

### **2.3.3. Gráfico de Pareto.**

*Los problemas de calidad se traducen en pérdidas y la mayoría de las pérdidas se deberán a unos pocos tipos de defectos, y estos defectos pueden atribuirse a un número muy pequeño de causas. Es necesario identificar esas pocas contribuciones para que los recursos de mejoramiento de la calidad se puedan concentrar en esas áreas. (J.M. Juran).*

La Gráfica de Pareto identifica los conceptos que más importancia tienen dentro de un conjunto de datos. Enfoca y dirige el esfuerzo hacia el análisis de los factores vitales que contribuyen fuertemente al logro de las metas y objetivos. Esta gráfica es en realidad un gráfico doble, formado por la representación utilizando barras de la frecuencia con que se presenta cada concepto (ordenada en forma decreciente) y una gráfica lineal del porcentaje acumulado de ellas.

Para la solución efectiva de las áreas de oportunidad se requiere que nos enfoquemos a las causas más importantes.

La gráfica de Pareto sirve para identificar los elementos principales que afectan la calidad, y por tanto para establecer que acciones prioritarias deben ponerse en marcha, a fin de reducir en un grado considerable los efectos de un mal desempeño del producto, proceso o sistema. En esta forma se aprovechan mejor los recursos y se canalizan más eficazmente los esfuerzos de las personas..

#### **2.3.4. Diagrama de causa y efecto.**

El Diagrama de Causa y Efecto es la herramienta más utilizada para analizar las causas del problema por resolver. Este diagrama indica las relaciones existentes entre las causas y los efectos. Fue utilizado en 1952 por el Dr. Kaoru Ishikawa y posteriormente por todas las grandes empresas del Japón, razón por la cual también se le conoce con el nombre de Diagrama de Ishikawa.

Todas las causas raíz son causas probables, pero por su gran número no es deseable comprobar todas y cada una de ellas. Se deben seleccionar aquellas que vale la pena comprobar si son o no causas reales. Para ello será necesario obtener evidencia objetiva por observación, consulta de información disponible o de ser necesario por experimentación.

Recibe el nombre de causa real aquella causa potencial que fue sometida a un mecanismo objetivo de comprobación y se cuenta con evidencia numérica de que efectivamente se presentó durante el período de observación y que contribuye significativamente al efecto indeseado.

Es importante enfrentar problemas específicos y no problemas globales de la empresa, que están más allá de nuestro alcance. Deben participar todas las personas involucradas en el problema, situación a resolver o mejorar. Muchas veces se cree conocer las soluciones y se relacionan directamente con el

problema, lo cual hace que nuestro diagrama sea muy simple y las causas del primer nivel sean acciones a tomar.

Lo que se busca son posibles causas del problema, las cuales después serán comprobadas y finalmente se determinarán las alternativas de solución.

Se debe decidir cuáles son las causas reales basados en evidencia objetiva, no en experiencia o sentido común. A veces una causa que parece insignificante, tiene un efecto importante en el problema,

#### **2.3.5. Gráfico de dispersión.**

El gráfico de dispersión es usado para estudiar la posible relación entre dos variables, asimismo se puede utilizar para probar posibles relaciones entre causa-efecto, causa-causa y efecto-efecto.

La construcción de un gráfico de dispersión requiere que exista información de dos variables, para cada evento. Para graficar la dispersión de estos datos, debe trazarse un plano cartesiano y asignarse el eje horizontal a la variable independiente “(causa)” y el eje vertical a la variable dependiente (efecto). Cuando no es clara la relación causa-efecto o cuando se quiere estudiar relaciones causa-causa y efecto-efecto, la asignación de los ejes será indistinta.

El Gráfico de Dispersión apoya para determinar y visualizar de manera clara la relación existente entre dos variables o características, como paso previo al cálculo del coeficiente de correlación.

#### **2.3.6. Histograma.**

Es importante medir los procesos para asegurar que el producto o servicio sea lo que el cliente espera, que se cumplan las especificaciones de acuerdo al diseño o

simplemente para conocer la cantidad de productos defectuosos que nuestro proceso está arrojando.

La forma más adecuada para representar las variaciones en el proceso es a través de histogramas. El histograma es una herramienta que nos ayuda a medir la variación de los procesos ordenando, clasificando y distribuyendo datos en función de su frecuencia.

Un histograma nos apoya para realizar un diagnóstico del desempeño de la variación y darnos cuenta del comportamiento que presentan nuestros datos ante los objetivos que perseguimos, también nos da la pauta de identificar las causas y las acciones que debemos tomar para mejorar nuestros procesos.

### **2.3.7. Gráficos de Control.**

El Ing. Walter A. Shewhart diseñó los *gráficos de control* para estudiar la variación en los procesos. Los gráficos de control de Shewart tienen como objetivo fundamental establecer si el proceso está *bajo control estadístico* o no, con base en una secuencia de datos tomados durante un periodo razonable.

Si el proceso está bajo control estadístico, hay una base para predecir cómo va a comportarse el proceso en el futuro.

En cambio si el proceso está fuera de control; no solo no existe esta base de predicción, sino que se requiere estudiar cual es la causa especial de que el proceso esté fuera de control, para suprimirla.

Una gráfica de control consiste en una línea central, un par de límites de control, uno de ellos colocado en la parte de arriba (LSC) y el otro por la parte de abajo (LIC) de la línea central o media  $\bar{x}$ .

Si todos los valores se encuentran dentro de los límites de control, sin ninguna tendencia especial, se dice que el proceso está en Control Estadístico. Pero si uno o más valores están fuera de los límites de control, se dice que el proceso está fuera de control.

Las Gráficas de Control facilitan la comprensión de los datos y además nos ayudan a observar de manera objetiva el comportamiento de la variable o indicador determinado para la medición del proceso.

La gran ventaja de los gráficos de Shewhart es lo sencillo que resulta ser su uso en campo, en particular por trabajar con pequeñas muestras y requerir solo de las cuatro operaciones aritméticas.

Si se detecta que el problema es generado por causas especiales, entonces la solución debe ser inmediatamente dada por el propio personal que realiza la actividad o está colaborando en el proceso; si la causa es común entonces se deberá analizar profundamente por la administración, con la finalidad de encontrar la causa real.

## ***2.4. Herramientas administrativas.***

En 1977 el Comité para el Desarrollo de Métodos de Control de Calidad (QC Technique Development Committee) de la Unión japonesa de Científicos e Ingenieros (JUSE) propuso un conjunto de métodos y técnicas para el control de datos descriptivos que surgen en todo proceso creativo y administrativo. Dicho conjunto de métodos se denominan indistintamente las siete nuevas herramientas gerenciales o las siete herramientas administrativas (Seven Management Tools for Quality Control). Estas nuevas herramientas no sustituyen a las siete herramientas básicas, sino que son complementarias para una correcta toma de decisiones.

Cumplen con dos objetivos fundamentales: ampliar la perspectiva en relación a las datos útiles a la organización y reconocer la importancia de la planeación

Son utilizadas con frecuencia por la Alta Dirección, la Gerencia, el personal Staff y de diseño para la identificación de problemas, la definición de planes, una mejor coordinación entre departamentos y el manejo de información verbal.

### ***2.4.1. Diagrama matricial.***

El diagrama matricial es la representación gráfica de las relaciones existentes entre diferentes factores y la intensidad de las mismas, en términos cualitativos.

La construcción del diagrama matricial se basa en la experiencia de los expertos en el tema analizado; esto da lugar a confundir la opinión de las relaciones y su intensidad con los datos reales, que aún no se han obtenido.

Debe quedar claro que el diagrama matricial es utilizado como guía en el análisis del tema en estudio y es necesario comprobar la exactitud de las ideas que de su interpretación se obtengan.

Es útil en la planificación de la calidad, porque permite evaluar la relación entre las necesidades de los clientes y las características del producto, las características del producto y las de los competidores, las características del producto y las características del proceso de fabricación o entre las características del proceso de fabricación y los puntos de su plan de control.

El diagrama matricial proporciona información sobre la existencia e intensidad de las relaciones entre diversos aspectos relacionados con un tema en estudio y presenta esta información con una visión global muy completa sobre dicho tema. Permite tener en cuenta las implicaciones derivadas de la acción sobre cualquiera de los elementos implicados en el mismo.

#### **2.4.2. Diagrama de árbol**

El diagrama de árbol es una representación gráfica que muestra el despliegue progresivo de los factores o las acciones que contribuyen al objetivo deseado de una manera ordenada, clara y precisa. Existen dos tipos de diagrama de árbol:

Diagrama de árbol para la selección de acciones. Se utiliza para desplegar los objetivos y los medios para obtenerlos, entre ellos se hará la elección de los más adecuados. Permite desarrollar la calidad de nuevos productos, al conocer las opciones existentes y seleccionar entre ellas, identificar las acciones preventivas para asegurar la calidad y facilita el cumplimiento de los objetivos de la organización, al seleccionar las acciones que se realizarán desde la planeación.

Diagrama de árbol para el despliegue de actividades. Desplegar todos los elementos que componen al objeto o servicio que se estudia, permite aclarar las funciones de los elementos, la relación entre los departamentos se explica mejor con un organigrama, analizar las probabilidades de falla de un producto o servicio,

donde el primer paso es conocer todos y cada uno de sus elementos y analizar todas las actividades necesarias para obtener un producto o servicio.

El diagrama de árbol es una herramienta que se emplea para buscar la forma más apropiada y eficaz de alcanzar una meta específica, por ser la representación gráfica de las acciones que siguen rutas lógicas para lograr esa. Permite la planificación al desglosar los objetivos generales hasta un grado ideal de detalle, así los traduce directamente en acciones.

### **2.4.3 Diagrama de afinidad**

El diagrama de afinidad es una de las herramientas del Método KJ, inventado por el Dr. Jiro Kawakita. Este método está diseñado para ordenar información que se obtienen de situaciones nuevas, futuras o desconocidas por alguna razón, a través de la integración de conceptos que tienen afinidad, básicamente por intuición y no por lógica.

El diagrama de afinidad es una herramienta de síntesis y fue diseñada para reducir los datos descriptivos disponibles, sin perder la información esencial.

Con esta herramienta se pasa de una simple descripción de los hechos observados, a la definición de las verdaderas áreas de oportunidad, a través de la identificación de las características comunes y su generalización en categorías más amplias.

El diagrama de afinidad ayuda al entendimiento y comprensión de los problemas crónicos que no se han solucionado porque se han atacado los síntomas y no la verdadera enfermedad.

#### **2.4.4 Diagrama de relaciones**

El diagrama de relaciones es una herramienta diseñada para resolver problemas complicados a través de ordenar gráficamente sus causas y consecuencias.

Es una herramienta muy útil cuando el problema por resolver es claro, pero existen distintas causas o factores que afectan el problema y están a su vez interrelacionadas de manera compleja.

El diagrama de relaciones es útil para aislar todos los factores que determinan un resultado, buscar la causa de un problema a través de establecer las relaciones lógicas de causa y efecto, clarificar la estructura de un problema, desarrollar los medios para alcanzar un objetivo o propósito y explicar los factores en forma concisa.

Es similar al diagrama de afinidad en la medida que ayuda a identificar relaciones no evidentes y abre la posibilidad de dar soluciones creativas.

A diferencia del diagrama de causa-efecto, el diagrama de relaciones permite que dar soluciones a varias causas y observe avances intermedios.

#### **2.4.5. Diagrama de flechas**

El diagrama de flechas es una herramienta diseñada para programar todas las actividades que ya están determinadas y que tienen que cumplirse para lograr el objetivo trazado.

Esta herramienta es muy útil porque representa las interdependencias y las relaciones de precedencia entre distintas actividades de un proyecto. Nos indica además la secuencia de estas actividades y permite detectar aquellas acciones críticas que debemos cuidar para que el proyecto tenga un avance adecuado.

El diagrama de flechas es útil para hacer un programa detallado de actividades y eventos, calcular el itinerario óptimo con los tiempos límites para cada etapa, detectar la serie de actividades más importantes del proyecto o estimar el efecto que tiene el cambio de un evento del programa en el plan total.

Tiene como antecedente al Diagrama de barras de Gantt, que solo marca el principio y fin de cada actividad en una escala de tiempo horizontal, con la desventaja de no indicar la interdependencia entre las actividades.

A diferencia del diagrama de Gantt, el diagrama de flechas permite que el equipo tenga presente las metas intermedias que debe concluir antes de iniciar otras actividades y por lo tanto es una herramienta valiosa en la optimización de tiempos, al hacer evidente las tareas prioritarias y las que tienen alguna tolerancia para su realización.

#### **2.4.6. Diagrama de programación de decisiones en el proceso**

El Diagrama de Programación de Decisiones en el Proceso ó PPPD (Process Decision Program Chart o PDPC METHOD) fue diseñado por el Dr. Jiro Kondo, presidente del Consejo de Ciencias del Japón, para predecir el proceso y estado final de la Universidad de Tokio en 1968 y 1969. En la actualidad ha sido sistemáticamente organizado e incluido dentro de lo que se conoce como la Investigación de Operaciones (OR).

Es una herramienta de gran utilidad cuando aún nos encontramos en las fases de desarrollo y diseño de eventos, facilitando así llegar a resultados deseados.

Facilita alcanzar las metas establecidas al enriquecer el plan de trabajo desde la fase de desarrollo, definiendo con anticipación las decisiones correctas ante problemas que puedan presentarse.

Permite concebir medidas preventivas que nos ayudan a evitar situaciones no deseadas, a través de suponer deliberadamente las desviaciones que nos pueden alejar del resultado esperado.

Es posible emplear el Diagrama de Programación de Decisiones en el Proceso en la reducción del número y gravedad de las consecuencias de eventos no deseados, al definir las mejores acciones a tomar con anticipación y no en forma reactiva.

#### **2.4.7. Análisis matricial de datos.**

El análisis de múltiples variables es una parte de las matemáticas que estudia aquellas

Existen situaciones que necesariamente deben ser representadas por modelos en que intervienen dos o más variables, pero cada variable considerada complica significativamente los cálculos a realizar y la interpretación de los resultados.

El Análisis Matricial de Datos se utiliza para disminuir el número de variables consideradas en los modelos multivariados, con la menor pérdida posible de información relevante.

La reducción en el número de variables se logra con técnicas matemáticas ya establecidas, pero se requiere la participación de los expertos en el tema. Su conocimiento del proceso, equipo o servicio permite una mejor selección de las variables que se conservaran en el modelo. Esta herramienta se estudia a detalle en la sección 2.6.

## *2.5. Uso de las herramientas básicas y administrativas en México.*

La Asociación Mexicana de Trabajo en Equipo (AMTE) es el organismo de referencia en las diferentes formas de mejora continua y en la aplicación de las herramientas básicas y administrativas en nuestro país. Los foros públicos que organiza forman una plataforma objetiva y transparente para conocer el crecimiento y evolución de las metodologías de mejora continua. En la actualidad existen cinco foros:

- Concurso Nacional de Círculos de Control de Calidad (23 años)
- Foro Nacional de Trabajo en Equipo (18 años)
- Encuentro Nacional de Equipos de Metodologías Estratégicas (12 años)
- Encuentro Nacional de Equipos de Proyectos Seis Sigma (8 años)
- Short Kaizen and Good Idea Forum (primer evento en octubre de 2012)

Los reconocimientos en cada categoría se entregan anualmente por el Presidente de la República, dentro del Encuentro Nacional de Industriales de CANACINTRA. Constituyen el máximo reconocimiento que se entrega en México a los equipos de trabajo y sus organizaciones, que se distinguen por la aplicación de metodologías de mejora continua, contribuyendo a la competitividad del aparato productivo.

Estos eventos tienen como objetivo el valorar al elemento humano a través del trabajo en equipo, como un medio efectivo para relacionar su capacidad de mejora con los procesos productivos que realizan, ya que son los que mejor conocen su trabajo y pueden mejorarlo de forma eficaz.

La metodología aplicada logra que estas mejoras se hagan en el marco sistemático de aplicación de herramientas estadísticas básicas, como medio efectivo de superación y mejora del nivel de vida a través de la participación en el trabajo.

El primer círculo de control de calidad en Latinoamérica fue creado el 1971 en la empresa VW de Brasil. En México se comenzó en la región Noreste del país estableciéndose el primero en 1976.

El Instituto Tecnológico de Estudios Superiores de Monterrey, realizó en 1982, un estudio sobre los CCC encontrando 362 equipos que agrupaban a 3,443 personas en 21 empresas, todas ellas pertenecientes al sector privado en las ciudades de México, Monterrey, Guadalajara, Cuernavaca, Puebla y San Luis Potosí. Los nombres que se les daban a los equipos eran círculos de excelencia, círculos de control de calidad, grupos de aseguramiento, grupos de productividad, etc.

Hirata estima que el mayor crecimiento en las actividades de CCC se dio entre los años de 1980 y 1983 debido a la promoción que se les dio en Estados Unidos y en consecuencia en México donde, por la crisis económica, los temas de calidad y productividad eran bien recibidos por los empresarios.

Aunque no existe un seguimiento al desarrollo de los CCC en la década de 1980, la situación económica de las empresas, la falta de información y conocimiento adecuados, fueron la causa de la desaceleración de este movimiento y su desaparición en muchas de las empresas. Solamente sobrevivieron en aquellas empresas donde el compromiso de la alta dirección por la calidad y la participación eran constantes.

En 1988 se formó la Asociación Mexicana de Círculos de Calidad con el propósito de encontrar modelos participativos adecuados a la realidad de México y para 1991, 22 empresas se encontraban registradas en dicha asociación.

Desde esa década JETRO-México (Japan External Trade Organization) ha apoyado a nuestro país sufragando los gastos de especialistas japoneses en

distintas disciplinas para su asesoría en México. Una de estas áreas ha sido la de Control de Calidad y su promoción.

En 1990 el Dr. Yuzuru Itoh (ex-director de Aseguramiento de Calidad de Matsushita Electric) en una de sus misiones para promover la calidad en Latinoamérica, propone que el esfuerzo del gobierno japonés en la promoción de la calidad debería centrarse en un área específica para tener mejores resultados, evaluaciones y retroalimentaciones. Propone la creación de un foro donde se presentaran casos exitosos de mejoras.

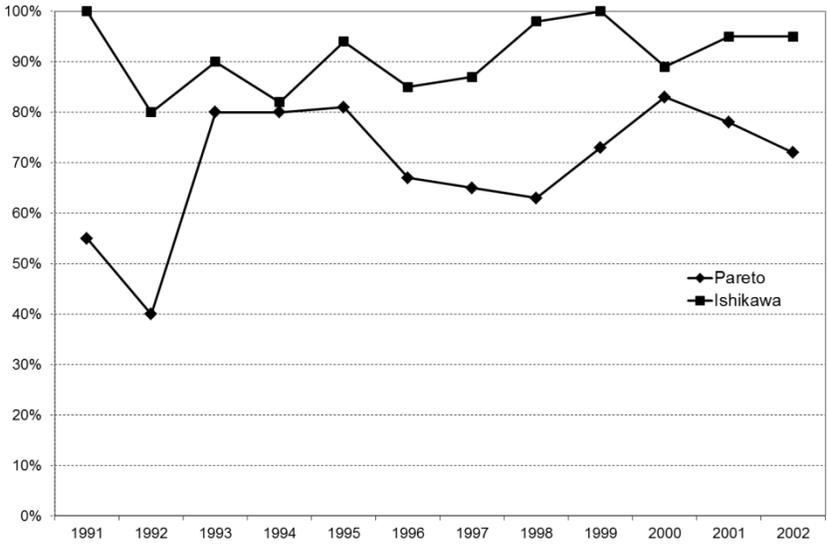
En Diciembre de ese año se organiza el primer concurso nacional con el apoyo de JETRO, CONCANACO, SERVYTUR, Bancomext, el ITESM Campus Monterrey y empresas privadas como Mitutoyo. En particular se recibió el apoyo de CANACINTRA para que el premio fuera entregado por el Presidente de la República, dentro de su Encuentro Nacional de Industriales.

En 1992 se establece una asociación civil llamada *Concurso Nacional de Círculos de Control de Calidad, A.C.* por Hermelinda Kasuga, Carlos González, Sergio Garcilazo y Ricardo Hirata para la promoción nacional del trabajo en equipo, así como la capacitación y asesoría en la materia. En el año 2000 cambia su nombre por Asociación Mexicana de Trabajo en Equipo, A.C. y se consolida como la asociación responsable de coordinar las actividades de difusión y organización de los concursos de trabajo en equipo en México

Conforme se fueron desarrollando los concursos, se observó la necesidad de abrir nuevos foros para atender el crecimiento y evolución de las metodologías hasta existir en la actualidad cinco diferentes foros.

Castellanos e Hirata analizaron el uso de las siete herramientas básicas y administrativas por parte de los equipos que asistieron a los distintos foros organizados por la Asociación Mexicana de Trabajo en Equipo y encontraron el uso consistente del diagrama de causa y efecto y la gráfica de Pareto (ver figura 2.5.1). Es de destacar que el diagrama de causa y efecto sea utilizado en el 90% de los casos, siendo la herramienta de mayor aplicación en México. Las gráficas de Pareto son las segundas más frecuentemente utilizadas.

**Figura 2.5.1 Equipos finalistas del Concurso Nacional que utilizaron la gráfica de Pareto y el diagrama causa-efecto 1991-2002.**



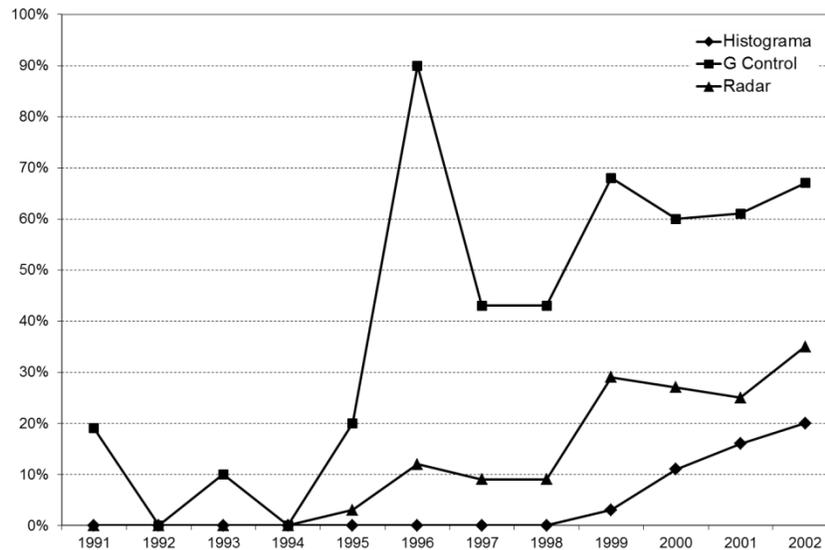
**Se observa el uso consistente de ambas herramientas.**

**Fuente: Adaptado de Castellanos e Hirata (2005)**

Otras herramientas básicas como el histograma, la gráfica de control y el gráfico de radar han tenido un incremento relevante, que se puede observar en la figura 2.5.2, este incremento puede relacionarse con un mejor entendimiento de las herramientas en sí y sobre todo con la aplicación sistemática de alguna de las

metodologías para la solución de problemas, como las 7 disciplinas de la Ford, el Ciclo Deming, el QC Story y Seis Sigma.

**Figura 2.5.2 Equipos finalistas del Concurso Nacional que utilizan herramientas de control estadístico de proceso y gráfica de radar 1991-2002**



**Se observa el uso creciente de herramientas estadísticas.**

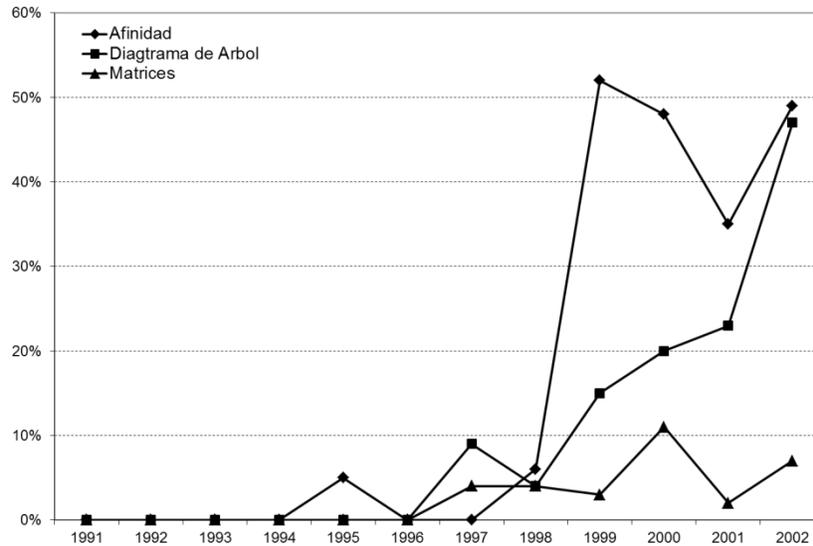
**Fuente: Adaptado de Castellanos e Hirata (2005)**

En años recientes, los equipos finalistas de los concursos nacionales y foros de trabajo en equipo han profundizado en el análisis de las posibles causas a partir de la información descriptiva que proporciona su conocimiento especializado y experiencia en el tema.

Los reportes técnicos finalistas informan del uso de algunas de las siete nuevas herramientas en sus procesos de solución de problemas (figura 2.5.3), en particular el diagrama de afinidad y el diagrama de árbol.

**Figura 2.5.3 Equipos finalistas del Concurso Nacional que utilizan las herramientas administrativas 1991-2002**

---



**Se observa incremento en el uso de Herramientas Administrativas**

**Fuente: Adaptado de Castellanos e Hirata (2005)**

---

Es relevante el uso de los diagramas matriciales, herramienta cualitativa que revela el enfoque sistémico al abordar el tema y que operativamente es el antecedente inmediato del análisis matricial de datos.

Al momento de escribir estas líneas aún no se tiene reporte alguno sobre el uso del análisis matricial de datos en la Asociación Mexicana de Trabajo en Equipo, incluyendo los foros anuales de seis sigma y los foros de metodologías estratégicas, pero si se observa el uso de otras técnicas avanzadas de estadística, como las pruebas de hipótesis y el diseño de experimentos.

## ***2.6. Análisis Matricial de Datos***

El Análisis Matricial de Datos se utiliza para disminuir el número de variables consideradas en los modelos multivariantes, ya que cada variable descartada simplifica significativamente los cálculos a realizar y la interpretación de los resultados.

El análisis matricial de datos se deriva de dos grandes vertientes: la ciencia del comportamiento y la matemática pura. Tuvo su origen en la intuición de que existían muchos aspectos vinculados con las investigaciones sociales no observables, pero a los que se podía acceder por medio de evaluaciones indirectas y cuyo análisis profundo permitiría detectar la presencia de ciertos esquemas básicos responsables del comportamiento del sistema.

### ***2.6.1 Antecedentes del análisis matricial de datos.***

El antecedente directo más remoto lo tenemos en Galton, que desarrolló técnicas fundamentales de medición estadística, entre las que cabe destacar el cálculo de correlación (1888). En este contexto Pearson, definió los significados de correlación, análisis de regresión y desviación estándar (1901).

Fue en 1904 cuando el psicólogo Spearman aplicó la correlación de datos intentando resolver si la inteligencia respondía a un solo factor general o si estaba conformada por un conjunto de habilidades específicas.

Durante la década de 1930 otros psicólogos, como Thurstone y Garwett, prosiguieron con esta línea de investigación, que llevaron en las décadas de 1940 y 1950 a la teoría sobre la estructura de las habilidades de Vernon, de la inteligencia por Guilford y Eysenck, y de la personalidad por Eysenck y Cattell (Pedret, 2000).

En 1957 Kendall definió el análisis multivariable como el conjunto de técnicas estadísticas que analizan simultáneamente más de dos variables en una muestra de observaciones.

En la década de 1960, Bertalanffy, el creador de la Teoría General de Sistemas afirma: “Hay progresos novedosos destinados a enfrentar las necesidades de una teoría general de los sistemas, (entre ellos) el aislamiento, por análisis matemático, de factores en fenómenos multivariados” (Bertalanffy, 1968).

De esta manera, el análisis matricial de datos se vinculó estrechamente con el objetivo último de la ciencia en general: superar los aspectos superficiales en busca de un conjunto reducido de factores explicativos.

Para la década de 1970 se había polarizado la opinión, pues algunos investigadores y estadísticos creían que el análisis multivariable consume demasiados recursos: “El análisis de datos multivariados no justifica el tiempo requerido para comprenderlo y realizarlo, es una manera elaborada de hacer algo que sólo puede ser muy tosco, como la selección de agrupamiento de variables interrelacionadas y entonces hallar algún tipo de promedio de las variables que se encuentran en un agrupamiento, a pesar que esas variables pueden estar medidas en unidades muy diferentes” (Hills, 1977), mientras que otros opinaban “los métodos de análisis multivariable son básicos para desentrañar la compleja interacción de factores... si un problema no es tratado como un problema multidimensional, está siendo tratado superficialmente” (Sheth, 1972)

Kinnear et al. (1989) manifiestan que el análisis factorial es un procedimiento del análisis multivariable, mediante el cual se toma un gran número de variables y se investiga para ver si tienen un pequeño número de factores en común que expliquen su intercorrelación.

Aaker (1990) define el *factor* como una variable que no es directamente observable pero que necesita ser inferida de las variables de entrada. Es decir, el factor es el agrupamiento de aquellas variables de insumo que son indicadoras del factor.

El análisis multivariable fue impulsada por el uso de equipos de cómputo cada vez con mayor capacidad de operación, así como por el desarrollo de programas informáticos amigables, por lo que recibió su justo reconocimiento: “Debido a su poder, elegancia y proximidad al núcleo del propósito científico, al análisis matricial puede atribuirse el nombre de reina de los métodos analíticos” (Kerlinger, 1990)

### **2.6.2 Conceptos fundamentales.**

Bertalanffy describe la necesidad de trabajar con datos multivariables en su Teoría General de Sistemas, donde afirma: “*La ciencia clásica se ocupa ante todo de problemas de dos variables, de cursos causales lineales, de una causa y un efecto o de unas pocas variables cuando mucho. Numerosos problemas son al fin y al cabo problemas multivariables que requieren nuevos instrumentos conceptuales.*” (Bertalanffy, **General System Theory**, 1968).

Se obtienen datos multivariables siempre que se mide o evalúa más de un atributo o característica de cada unidad experimental (Johnson, 2004). Los atributos y características son las variables y la unidad experimental es el concepto que se puede medir o evaluar de alguna manera.

Los métodos multivariados son necesarios para trabajar con conjuntos grandes de datos, que tienen una gran cantidad de variables medidas en grandes cantidades de unidades experimentales. La importancia y la utilidad de los métodos

multivariadas aumentan al incrementarse el número de variables que se están midiendo y el número de unidades experimentales que se están evaluando.

El análisis multivariado utiliza métodos estadísticos formales que exigen que se establezca una hipótesis, se reúnan datos y que se usen estos datos para comprobar o rechazar la hipótesis.

Pero también tiene un lado exploratorio. Con algunas metodologías da origen a hipótesis en lugar de probarlas. Una situación que se da frecuentemente es cuando se dispone de una gran cantidad de datos (por ejemplo 50 variables medidas sobre más de 2,000 unidades experimentales) y nos preguntamos si existe información valiosa en ellos, que ha pasado desapercibida. Las técnicas multivariadas son útiles para examinar los datos y para saber si hay más información que sea valiosa en esos datos.

Bertalanffy resume la aportación del análisis multivariable de la siguiente forma: *“Si conocemos el total de las partes contenidas en un sistema y la relación que hay entre ellas, el comportamiento del sistema es derivable a partir del comportamiento de las partes.”* (Bertalanffy, **General System Theory**, 1968).

Una clasificación fundamental entre los métodos multivariados es la que los segmenta en *técnicas dirigidas a las variables* y *técnicas dirigidas a los individuos*.

El interés de los análisis multivariados es encontrar relaciones entre las variables respuesta, entre las unidades experimentales y relaciones entre ambas.

Existen relaciones entre las variables respuesta cuando, por ejemplo, algunas de las variables están midiendo una entidad común. Así los exámenes de

interpretación de lectura, destreza visual, aritmética y conocimientos generales están relacionados entre sí y el aspecto común que estos exámenes están midiendo podría ser la inteligencia global.

Existen relaciones entre las unidades experimentales si algunas de ellas son semejantes entre sí. Por ejemplo, cuando se evalúan variedades de granos respecto a su contenido nutricional. Se miden los gramos de grasa, proteína y carbohidratos en cada lote. Los tipos de cereal relacionados darán resultados semejantes en cada uno de ellos.

Las técnicas dirigidas a las variables se realizan sobre las matrices de correlación (de ahí el nombre de Análisis Matricial de Datos). Algunos ejemplos son el análisis de componentes principales, el análisis por factores, el análisis de regresión y el análisis de correlación canónica.

Las técnicas dirigidas a los individuos son las que se interesan en las relaciones que existen entre las unidades. Algunos ejemplos son el análisis discriminante, el análisis por agrupación y el análisis multivariado de la varianza.

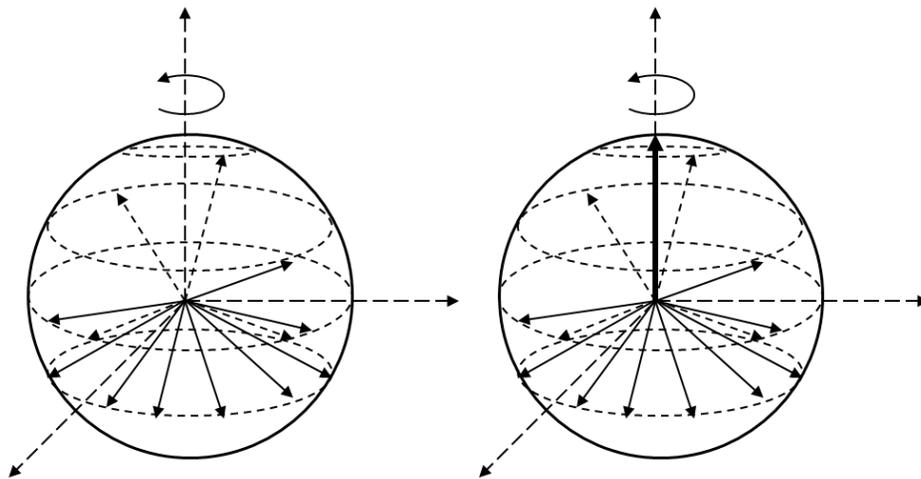
### **2.6.3 Reducción de variables.**

Strang explica la reducción de variables ejemplificando el concepto de vector propio (eigenvector o vector latente) con una esfera que gira sobre su propio eje (Strang, *Algebra Lineal*, 1990)

En la figura 2.6.1 se presentan distintos vectores que indican la posición de un punto sobre la superficie de la esfera. Como la esfera está girando se tiene que los vectores de posición varían en función del tiempo.

Afirma que todo vector que indique la posición de un punto sobre la superficie de la esfera es derivable de un vector propio, que, por coincidir con el eje de rotación, es el único que no gira. Este vector se interpreta como el vector propio del sistema.

**Figura 2.6.1 Interpretación geométrica de la reducción de variables.**



**Todos los vectores de posición son derivables del vector latente o vector propio que coincide con el eje de rotación .**

**Fuente: Adaptado de Strang.**

El problema de reducir el número de variables se transforma en encontrar los vectores propios y seleccionar las nuevas variables que expliquen el comportamiento de la familia de vectores original, con el grado de precisión deseada.

Sea  $A$  una matriz de  $\mathbf{p} \times \mathbf{p}$ , entonces  $\lambda$  es un valor propio de  $A$  si y sólo si

$$p(\lambda) = \det(A - \lambda I) = 0$$

en donde  $p(\lambda)$  es un polinomio de  $\lambda$  de grado  $p$ , y es conocido como polinomio característico de  $A_i$ .

Las raíces de  $p(\lambda)$ ,  $\lambda_1, \lambda_2, \dots, \lambda_\pi$  son los valores propios de  $A$ . Toda matriz de  $\mathbf{p} \times \mathbf{p}$  tiene exactamente  $\mathbf{p}$  valores característicos. A cada valor propio  $\lambda_i$  le corresponde un vector propio  $\mathbf{c}_i$  tal que

$$A\mathbf{c}_i = \lambda_i\mathbf{c}_i.$$

El procedimiento para calcular valores y vectores propios es:

- Encontrar  $p(\lambda) = \det(A - \lambda \mathbf{I})$ .
- Calcular las raíces  $\lambda_1, \lambda_2, \dots, \lambda_\pi$  de  $p(\lambda) = 0$ .
- Resolver el sistema homogéneo  $(A - \lambda_i\mathbf{I})\mathbf{c}_i = 0$  que corresponde a cada valor característico de  $\lambda_i$ .

Pérez ejemplifica la obtención de los valores propios y sus correspondientes vectores propios con una matriz  $3 \times 3$ :

$$A = \begin{bmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{bmatrix}$$

El polinomio característico  $p(\lambda) = \det(A - \lambda\mathbf{I})$  es

$$p(\lambda) = \begin{vmatrix} 1 - \lambda & -1 & 4 \\ 3 & 2 - \lambda & -1 \\ 2 & 1 & -1 - \lambda \end{vmatrix}$$

$$p(\lambda) = [(1 - \lambda)(2 - \lambda)(-1 - \lambda) + (3)(1)(4) + (2)(-1)(-1) - (1 - \lambda)(1)(-1) - (2)(4)(2 - \lambda) - (3)(-1)(-1 - \lambda)]$$

$$p(\lambda) = -\lambda^3 + 2\lambda^2 + 5\lambda - 6 = -(\lambda - 1)(\lambda + 2)(\lambda - 3)$$

Igualando a cero se obtienen las raíces del polinomio característico

$$\lambda_1 = 1, \lambda_2 = -2, \lambda_3 = 3$$

A continuación se resuelve el sistema  $(A - \lambda \mathbf{I})\mathbf{c}_i = 0$  para cada valor propio.

Para  $\lambda_1 = 1$

$$(A - \lambda \mathbf{I})\mathbf{c}_i = \begin{bmatrix} 0 & -1 & 4 \\ 3 & 1 & -1 \\ 2 & 1 & -2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Que se resuelve por el método de eliminación de Gauss-Jordán, para obtener el primer vector propio:

$$\mathbf{C}_1 = \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}$$

En forma similar se obtienen el segundo y tercer vectores propios

Para  $\lambda_2 = -2$

$$(A - \lambda \mathbf{I})\mathbf{c}_i = \begin{bmatrix} 3 & -1 & 4 \\ 3 & 4 & -1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{C}_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

Para  $\lambda_3 = 3$

$$(A - \lambda \mathbf{I})\mathbf{c}_i = \begin{bmatrix} -2 & -1 & 4 \\ 3 & -1 & -1 \\ 2 & 1 & -4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{C}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

En el análisis matricial de datos se utilizan diferentes software matemáticos que facilitan el cálculo de los valores propios y vectores propios, permitiendo así el transformar el conjunto de variables originales correlacionadas en un nuevo conjunto de variables no observadas y no correlacionadas.

Cada una de las nuevas variables es un **factor** o **constructo** del conjunto original. Las variables recién creadas se interpretan con base en las correlaciones de las variables originales.

Al respecto Feigenbaum advierte “*Los programas de calidad del productor y del comerciante hechos para satisfacer el mercado de hoy no deben ser únicamente técnica y administrativamente sólidos, sino también deben ser claros, entendibles y visibles de tal forma que puedan ser completamente discutidos, revisados y aceptados*” (Feigenbaum, **Control Total de la Calidad**, 1991).

## *Capítulo 3*

### *Metodología*

---

La metodología del análisis matricial de datos tiene como eje central la elaboración de la matriz de correlaciones introducida por Tuckey. Esta matriz cuadrada tiene en las columnas y filas las variables observadas y los elementos de entrada de la matriz son los índices de correlación de esas variables.

Cada uno de los índices de correlación se calcula tomando los valores registrados en todos las unidades experimentales de cada para de variables.

La diagonal mayor de esta matriz corresponde a las correlaciones de cada variable consigo misma y todas toman como valor la unidad.

La matriz así construida tiene implícitamente una variedad de vectores propios de los que se derivan los vectores observados y cada uno de los vectores propios condensa o explica una parte de la variabilidad de los datos originales.

Es posible que una cantidad pequeña de estos vectores sean suficientes para explicar el comportamiento del sistema, con un grado preestablecido de precisión, si se confirma que la distribución de las correlaciones no es homogénea entre las variables. Es menor el número de vectores propios requeridos, entre más heterogénea sea esta distribución.

La cantidad de varianza explicada por cada vector propio es utilizada para crear la matriz de correlaciones reproducidas, modelo que puede sustituye a la matriz de

correlaciones originales, si se confirma que no existe pérdida significativa de información.

Por último se elabora una matriz donde son calificadas las unidades experimentales con las nuevas variables.

El análisis matricial de datos es un instrumento muy sensible para medir las correlaciones no observadas entre variables y requiere iniciar con la detección y tratamiento adecuado de datos ausentes o atípicos en la información de entrada.

Las etapas del análisis matricial de datos son:

- Análisis exploratorio y gráfico de las variables.
- Análisis exploratorio y gráficos de los eventos.
- Matriz de unidades experimentales y variables. Comprobación de la correlación supuesta.
- Extracción de factores
- Determinación del número de factores
- Rotación de factores
- Interpretación de los factores
- Matriz de unidades experimentales y factores.
- Validación del modelo..

La información sobre la satisfacción del cliente utilizada para la comprobación de la hipótesis fue proporcionada por una empresa industrial del sector de alimentos y bebidas.

La empresa estableció como instrumento para medir la satisfacción de sus distribuidores minoristas la siguiente encuesta:

¿Qué calificación otorga a...

- la calidad de los productos?
- las condiciones generales del envase?
- las condiciones del empaque?
- el cumplimiento a la cantidad solicitada?
- el cumplimiento a las marcas y presentaciones solicitadas?
- la frecuencia con la que es visitado en su establecimiento?
- el tiempo para el retiro de envase y la descarga de producto?
- la atención recibida por parte del personal?
- la orientación respecto a la rotación del producto?
- el servicio que le brinda la empresa?
- la comunicación con la empresa?
- los medios actuales de comunicación?
- la información sobre los productos?
- la rapidez de respuesta a sus quejas o sugerencias?
- la respuesta que recibe a sus quejas o sugerencias?

La encuesta se aplica en una entrevista personal, donde se le solicita al distribuidor minorista calificar de 0 a 10 el grado de satisfacción en cada uno de los quince temas incluidos, donde el cero representa el mayor grado de insatisfacción y el diez la mayor calificación que se puede otorgar.

Como la encuesta está dirigida al total de distribuidores minoristas y se aplica dos veces al año. Existen varios miles de registros. Se proporcionó la información consolidada de 300 cuestionarios, seleccionados aleatoriamente.

### *3.1 Análisis exploratorio y gráfico de las variables.*

La distribución normal es el modelo que se presenta con mayor frecuencia en los estudios estadísticos, pero no es adecuado dar por hecho que esta distribución será el comportamiento de las variables incluidas en cualquier estudio. Es necesario un examen previo de la estructura de los datos.

Los parámetros convenientes cuando la distribución de datos se aproxima a la normal son la media y la desviación estándar. Cuando los datos no se ajustan a una distribución normal, es recomendable el uso de estadísticos robustos. Los estadísticos robustos son aquellos que se ven poco afectados por valores atípicos, por estar basados en la mediana y en los cuartiles.

Tuckey fue de los primeros matemáticos en introducir el análisis exploratorio de datos. Con él se pueden analizar los datos exhaustivamente y detectar las posibles anomalías que presentan las observaciones.

Pérez (2004) recomienda iniciar el análisis exploratorio de datos con un histograma que permita visualizar su estructura, porque pone en evidencia la presencia de distribuciones multimodales. También permite intuir la distribución de probabilidad de los datos, su normalidad, su simetría y otras propiedades relevantes. El simple examen de los datos tabulados no aporta tanta información como su gráfica.

El siguiente paso es examinar la simetría y posible presencia de valores atípicos, utilizando los gráficos de caja y bigote. Esta herramienta permite confirmar la simetría de los datos, detectar valores atípicos y ver el ajuste de los datos a la distribución de frecuencias en cada variable.

El gráfico de caja y bigotes divide los datos en cuatro áreas de igual frecuencia: una caja central dividida en dos áreas por una línea vertical y otras dos áreas representadas por dos segmentos horizontales que parten del centro de cada lado vertical de la caja. La caja central encierra el 50 por ciento de los datos.

La mediana es una línea vertical en el interior de la caja. Si esta línea está en el centro de la caja hay simetría en la variable. Los lados verticales de la caja están situados en los cuartiles inferior y superior de la variable. Partiendo del centro de cada lado vertical de la caja se dibujan los dos bigotes, uno hacia la izquierda y el otro hacia la derecha.

El bigote de la izquierda tiene su extremo en el valor dado por el primero cuartil menos 0,5 veces el rango ( $Q1 - 1,5 * (Q3 - Q1)$ ). El bigote de la derecha tiene un extremo en el valor dado por el tercer cuartil más 1,5 veces el rango ( $Q3 + 1,5 * (Q3 - Q1)$ ). De esta forma se crea un intervalo de tres cuartiles y se desplaza al centro de los datos, este intervalo sobrepuesto a una distribución normal cubre aproximadamente el 95% de la población.

De tal forma que Se considera valores atípicos los que se encuentren a la izquierda del bigote izquierdo y a la derecha del bigote derecho.

Por último se obtiene la gráfica de normalidad, para verificar el apego de las observaciones a la distribución normal. Permite determinar si un conjunto de datos se ajusta razonablemente a una distribución normal. El gráfico normal de probabilidad representa los valores de la variable objeto del estudio en el eje horizontal y la probabilidad de esos valores en el eje vertical.

La normalidad de los datos es óptima cuando los puntos graficados coinciden con la diagonal del primer cuadrante. Las diferencias que existan entre el gráfico de

probabilidad y la línea diagonal definirán si se acepta la normalidad del conjunto de datos.

Las representaciones gráficas nunca sustituyen a las medidas de diagnóstico formal estadístico, pero permiten el examen de la distribución de las variables individuales implicadas en el análisis, por lo que es adecuado su uso antes de invertir tiempo y recursos en un análisis completo.

En este primer paso de la metodología se debe asegurar la normalidad de cada variable o al menos que sea unimodal y simétrica.

### *3.2 Análisis exploratorio y gráficos de los eventos.*

Cuando se aplica un método de análisis exploratorio sobre las unidades experimentales o eventos, puede identificarse la ausencia de información. Estos valores ausentes pueden deberse al registro defectuoso de la información, a la ausencia natural de la información buscada o a una falta de respuesta.

También se identifican observaciones aisladas cuyo comportamiento se diferencia claramente del comportamiento medio del resto de las observaciones, son los datos atípicos.

Para la identificación de datos ausentes y atípicos se tienen disponibles técnicas del análisis exploratorio. Entre ellas tenemos las gráficas de radar y de Andrews.

En las gráficas de radar cada variable se representa como un eje de simetría radial en una gráfica con tantos ejes como variables se incluyan en el estudio. Esta técnica está limitada a datos que tengan una escala de medición comparable o previamente normalizada y con el mismo sentido de incremento

En 1972 Andrews sugirió que cada unidad experimental podría representarse por la función:

$$f_r(t) = \frac{x_{r1}}{\sqrt{2}} + x_{r2}\text{sen}(t) + x_{r3}\cos(t) + x_{r4}\text{sen}(2t) + x_{r4}\cos(2t) + \dots$$

De este modo los datos correspondientes a una unidad experimental dan lugar a una función única para ese evento, que se pueden graficar en un intervalo adecuado (usualmente  $-\pi < t < \pi$ ).

Las curvas resultantes son una excelente representación de las variables, además de ser útiles para localizar datos ausentes y atípicos en el conjunto de estudio. Una vez identificados se elige con criterio entre eliminarlos del análisis o evaluar toda la información.

### **3.2.1 Tratamiento de los datos ausentes**

La ausencia de algunos datos no implica que el análisis estadístico no pueda realizarse. Es posible identificar el efecto de los datos ausentes antes de proceder a su eliminación o a la imputación de la información faltante.

La primera opción es incluir en el análisis sólo las observaciones con datos completos. Así cualquier unidad experimental que tenga algún dato ausente es eliminada del conjunto antes de realizar el análisis.

Este método es utilizado en la mayoría del software estadístico. Es apropiado cuando no hay demasiados valores perdidos. En caso contrario se reducirá mucho el tamaño de la muestra a considerar para el análisis y no sería representativa de la información completa.

Otro método que elimina menos información y se utiliza en el análisis bivariable consiste en trabajar con todos los casos posibles que tengan valores válidos para cada par de variables, independientemente de lo que ocurra en el resto de las variables.

La tercera opción es suprimir las filas o columnas que peor se comporten respecto a los datos ausentes. La decisión depende de lo que se gana al eliminar una fuente de datos ausentes y lo que se pierde al no contar con un conjunto de casos en el análisis estadístico.

La alternativa a los métodos de supresión de datos es la imputación de la información faltante, que es un proceso de estimación de valores ausentes. Existen diferentes métodos de imputación:

Método de imputación por sustitución del caso. Las unidades experimentales con datos ausentes se sustituyen con otras no consideradas en la muestra original. Este método de imputación suele utilizarse cuando existen casos con todas sus observaciones ausentes o con la mayoría de ellas.

Método de imputación de sustitución por la media. Los datos ausentes se sustituyen por la media de todos los valores válidos. Aunque este método se implementa fácilmente y proporciona información para todos los casos, tiene la desventaja de que modifica las correlaciones e invalida las estimaciones de la varianza.

Método de imputación de sustitución por la mediana. Cuando hay valores extremos en las variables, se sustituyen los valores ausentes por la mediana, que es un estadístico resumen de los datos más robusto.

Método de imputación por interpolación. En él se sustituye cada valor ausente por el valor resultante de realizar una interpolación con los valores adyacentes.

Método de imputación de sustitución por valor constante. Los datos ausentes se sustituyen por un valor constante apropiado derivado de fuentes externas o de una investigación previa.

Método de imputación por regresión. Se utiliza el análisis de la regresión para predecir los valores ausentes de una variable basándose en su relación con otras variables del conjunto de datos. Como desventaja de este método destacaríamos que refuerza las relaciones ya existentes en los datos de modo que conforme aumenta su uso los datos resultantes son más característicos de la muestra y menos generalizables.

### **3.2.2 Tratamiento de datos atípicos.**

Las propias características del caso atípico, así como los objetivos del análisis que se realiza, determinan los casos atípicos a eliminar. Pérez (2004) propone cuatro categorías de datos atípicos:

Observaciones que provienen de un error de procedimiento. Estos datos deben eliminarse o tratarse como datos ausentes.

Observaciones extraordinarias existiendo una explicación para su presencia. Normalmente se conservan en la muestra.

Observaciones extraordinarias para las que el investigador no tiene explicación. Estos datos atípicos se eliminan del análisis.

Observaciones fuera del rango ordinario de valores de la variable o valores extremos. Se eliminan del análisis si no son elementos significativos para la población.

Los casos atípicos deben considerarse en el conjunto de todas las variables. Puede ocurrir que una variable tenga valores extremos eliminables, pero al considerar un número suficiente de otras variables en el análisis, se decida conservarlos.

En el segundo paso de la metodología se debe asegurar que las unidades experimentales o eventos aceptados no contengan datos ausentes o atípicos.

### *3.3 Matriz de unidades experimentales y variables. Comprobación de la correlación supuesta.*

En este punto es posible la construcción de la matriz de unidades experimentales y variables. Las ***m*** filas corresponden a las unidades experimentales que no presenten datos ausentes ni atípicos, o que recibieron alguno de los tratamientos mencionados en la sección 4.2. Las ***n*** columnas de esta matriz corresponden a las variables aceptadas por su simetría y distribución unimodal de frecuencias o, en el mejor de los casos, con distribución normal, como se menciona en 4.1.

Esta matriz ***n X m*** es el punto de partida para construir la matriz de correlaciones. La matriz de correlaciones es una matriz cuadrada de ***n X n*** que contiene en cada celda la correlación entre cada par de variables, calculada con base en los pares de valores que toman esas dos variables en cada uno de los ***m*** eventos. Los elementos de la diagonal principal corresponde a la correlación de cada variable al

coincidir como ambos valores de la pareja, por lo que será una diagonal exclusivamente con la unidad.

Antes de realizar el análisis matricial de datos se requiere comprobar que las variables originales están relacionadas entre sí o no lo están. En caso de no existir correlación entre ellas tampoco existirían factores comunes y no podría aplicarse el método.

Se identifican las variables de muy alta asociación como aquellas con un índice de correlación de 0.7 o mayor; las variables de asociación alta con índices entre 0.6 y 0.7; las variables de asociación media con correlación entre 0.3 y 0.5 y las variables de correlación baja con índices menores a 0.3. En este caso Nunnally (1999) recomienda excluirlas del análisis, pues estas variables no parecen tener relación con cualquier otra de las consideradas.

Comprobar la correlación supuesta entre variables es para Johnson (2004) la última tarea previa a la aplicación del análisis multivariable de datos, pero otros autores difieren. Así tenemos que Pérez (2004) la incluye como el primer paso del análisis factorial por el uso de la matriz de correlaciones.

### *3.4 Extracción de factores*

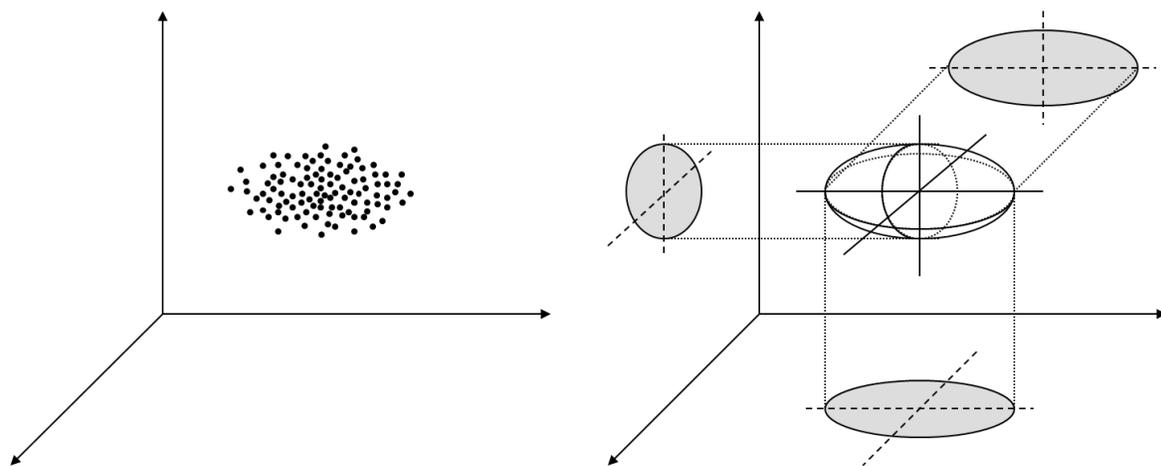
Pérez (2004) explica que se puede realizar una representación gráfica de la matriz de unidades experimentales y variables  $n \times m$  como una nube de puntos resultante de medir las variables  $X_1, \dots, X_m$  sobre una muestra de tamaño  $n$ .

En la figura 3.4.1 se presenta un elipsoide tridimensional que encierra a la nube y sus proyecciones en los tres planos. Cada sombra explica una parte del comportamiento de los puntos originales.

Se puede extrapolar el ejemplo a un elipsoide  $p$ -dimensional que contenga toda la información del problema estudiado. Así cada una de las  $k$ -proyecciones o sombras explicará parte de la información de la nube en  $p$  dimensiones.

La interpretación de las nuevas variables dependerá en buena parte de que cada una de las proyecciones agrupe con mayor peso algunas de las variables originales y con menor peso a las demás.

**Figura 3.4.1 Interpretación geométrica de la extracción de factores.**



**La variabilidad de un conjunto de vectores de tres dimensiones se reduce a solo dos dimensiones.**

**Fuente: Adaptado de Pérez 2004.**

El objetivo es determinar un número reducido de nuevas variables o factores que puedan representar a las variables originales. Existen diversos métodos para la extracción de los factores, cada uno de ellos con sus ventajas e inconvenientes:

Método de las componentes principales. Consiste en determinar una a una las componentes que mayor variación explican (componentes principales), retirando la variabilidad ya explicada en cada paso anterior. La matriz de cargas factoriales se calcula mediante las correlaciones de las variables originales con estas componentes. Este método tiene la ventaja de que siempre proporciona una solución y tiene el inconveniente de llevar a estimadores muy sesgados de la matriz de cargas factoriales, si existen variables con comunalidades bajas.

Método de los ejes principales. El método es iterativo y consiste en alternar una estimación de la matriz de especificidades, con una estimación de la matriz de cargas factoriales. La estimación se obtiene aplicando el método de componentes principales a la matriz de especificidades y se itera hasta que los valores de dichas estimaciones apenas cambien. Este método tiene la ventaja de proporcionar mejores estimaciones que el método anterior. Sin embargo, no está garantizada su convergencia, sobre todo en muestras pequeñas.

Método de la máxima verosimilitud. El método tiene la ventaja sobre los dos anteriores de que las estimaciones obtenidas no dependen de la escala de las variables. Es asintóticamente insesgada, eficiente y normal si las hipótesis del modelo factorial son ciertas. Además permite seleccionar el número de factores mediante contrastes de hipótesis. Su principal inconveniente radica en que puede haber problemas de convergencia si las variables originales no son normales, sobre todo en muestras pequeñas.

Nunnally (1999) apunta que en la mayoría de los análisis se obtienen inferencias casi idénticas, pero se debe tener presente que entre los distintos métodos se observan los siguientes puntos de comparación:

- Cuando las comunalidades son altas (mayores que 0.6) todos los procedimientos tienden a dar la misma solución.

- Cuando las comunalidades son bajas para algunas de las variables el método de componentes principales tiende a dar soluciones muy diferentes del resto de los métodos.
- Si el número de variables es alto (mayor que 30), las estimaciones de la comunalidad tienen menos influencia en la solución obtenida y todos los métodos tienden a dar el mismo resultado.
- Si el número de variables es bajo todo depende del método utilizado para estimar las comunalidades y de si éstas son altas más que del método utilizado para estimarlas.

### *3.5 Determinación del número de factores*

En este momento se decide el número  $f$  de factores o nuevas variables que tendrá la matriz que resulta de la aplicación del análisis matricial de datos.

Para determinar el número de factores a conservar se utiliza un gráfico de sedimentación. Esta herramienta es una gráfica lineal de la varianza explicada por las nuevas variables. El comportamiento de esta gráfica es asintótico, porque al surgir cada una de las nuevas variables queda menos varianza por explicar.

La regla Kaiser-Guttman para el gráfico de sedimentación indica tomar sólo los factores con valores propios mayores de la unidad.

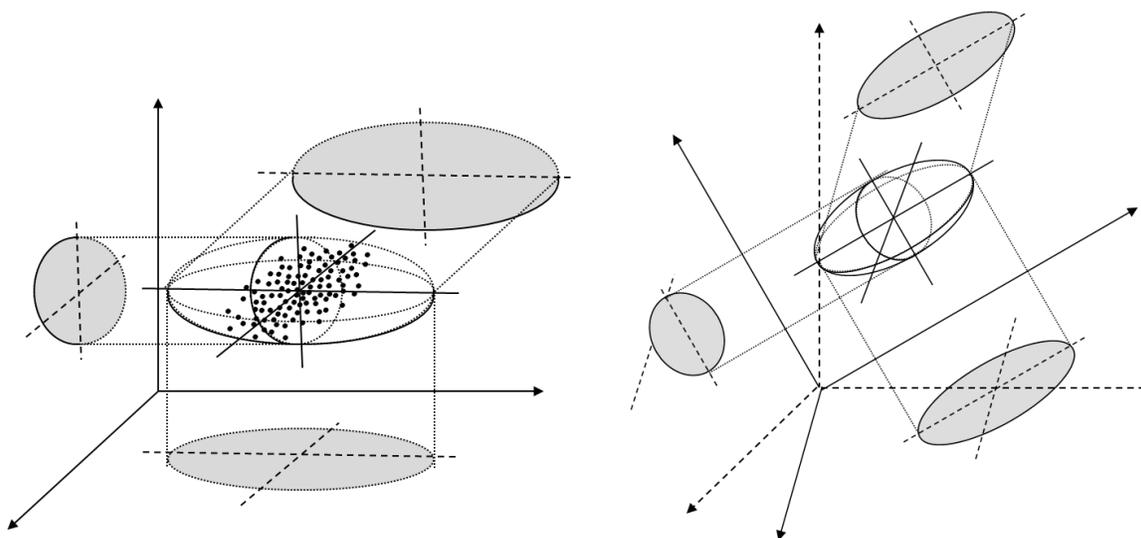
Se puede comparar la variación retenida con la proyección de la nube de datos multivariados en cada plano. En la figura 3.4.1 se observa que dos de las proyecciones siguen la dirección de la nube de datos y que la tercera proyección presenta una forma circular. Esta última proyección se explica como consecuencia del comportamiento aleatorio de los datos multivariados para ese factor.

### 3.6 Rotación de factores

En general los ejes principales obtenidos de la extracción de componentes no coinciden con los ejes de los datos multivariados originales (Pérez 2004). Una rotación que los alinearía facilitaría la interpretación de las nuevas variables obtenidas y aumenta la variación explicada.

La interpretación geométrica de esta inclinación se puede ver en la figura 3.6.1, donde se observa que al orientar los ejes de las nuevas variables con la nube de datos multivariados se mejora el enfoque de las proyecciones y aumenta la condensación de la variabilidad original.

**Figura 3.6.1 Interpretación geométrica de la rotación de factores.**



**La rotación de los factores permite que cada uno de ellos condense mayor cantidad de la variación original.**

**Fuente: Adaptado de Pérez 2004.**

Los software disponibles para el análisis de datos multivariados permiten la rotación ortogonal y la rotación oblicua. Se utiliza la primera cuando se requiere asegurar que los factores resultantes no estén correlacionados.

Este resultado es deseable porque indica que se ha realizado una condensación exhaustiva de la variabilidad de los datos originales.

### *3.7 Interpretación de los factores*

La interpretación de los factores se realiza con base en la matriz de componentes. Esta interpretación deja en claro por qué cada uno de los componentes principales explica a las variables originales con las que está correlacionado.

Un valor elevado en un factor indica la presencia de una variable con la que se tiene una alta correlación y de la distribución de cargas elevadas se puede inferir la característica latente común a las variables con salientes y que no se encuentra presente en las variables con cargas nulas (Pedret, 2000, Johnson 2004, Pérez 2004)).

La práctica común (Johnson 2004, Pérez 2004) es tomar como criterio el valor de 0.500 o mayor para declarar que una carga es lo suficientemente elevada para considerar que la variable será explicada satisfactoriamente por ese factor.

### *3.8 Matriz de unidades experimentales y factores.*

Para cumplir el objetivo de reducir las variables observadas se requiere que las nuevas variables o factores sean intercambiables, tengan sentido y midan algo útil.

El coeficiente de correlación entre un factor y una variable se calcula multiplicando el peso de la variable en ese componente por la raíz cuadrada del valor propio.

$$r_{jh} = u_{jh}\sqrt{\lambda_h}$$

Estos coeficientes de proporción son llamados *Puntuaciones Factoriales* y se utilizan para calcular cada uno de los elementos de entrada en una matriz de unidades experimentales contra factores.

$$\begin{array}{lcl}
 Z_1 = r_{11}X_1 + r_{12}X_2 + \dots + r_{1p} X_p & & X_1 = r_{11}Z_1 + r_{21}Z_2 + \dots + r_{k1} Z_k \\
 \\
 Z_2 = r_{21}X_1 + r_{22}X_2 + \dots + r_{2p} X_p & \Rightarrow & X_2 = r_{12}Z_1 + r_{22}Z_2 + \dots + r_{k2} Z_k \\
 \vdots & & \vdots \\
 \\
 Z_K = r_{K1}X_1 + r_{K2}X_2 + \dots + r_{Kp} X_p & & X_P = r_{1P}Z_1 + r_{2P}Z_2 + \dots + r_{kP} Z_k
 \end{array}$$

La matriz de puntuación factorial contiene los coeficientes para transformar las variables originales en las nuevas variables lineal y permite obtener la matriz de unidades experimentales contra factores.

### 3.9 Validación del modelo..

Por último la validación del modelo se logra verificando que no existan diferencias significativas al comparar la matriz de datos multivariantes obtenida y la matriz original. El contraste utilizado será que al comparar ambas en una matriz de residuales no se observen valores mayores al 5% (Pérez 2004).

# Capítulo 4

## Resultados

---

### 4.1. *Análisis exploratorio y gráfico de las variables.*

Como se mencionó en el capítulo de Metodología, la importancia del análisis exploratorio previo al análisis de datos matriciales es la comprobación de la normalidad estadística supuesta en las variables que intervienen.

- Variable 1 Calidad de los productos.
- Variable 2 Condiciones generales del envase.
- Variable 3 Condiciones del empaque.
- Variable 4 Cumplimiento a la cantidad solicitada.
- Variable 5 Cumplimiento a las marcas y presentaciones solicitadas.
- Variable 6 Frecuencia con la que es visitado en su establecimiento.
- Variable 7 Tiempo para el retiro de envase y la descarga de producto.
- Variable 8 Atención recibida por parte del personal.
- Variable 9 Orientación respecto a la rotación del producto.
- Variable 10 Servicio que le brinda la empresa.
- Variable 11 Comunicación con la empresa.
- Variable 12 Medios actuales de comunicación.
- Variable 13 Información sobre los productos.
- Variable 14 Rapidez de respuesta a sus quejas o sugerencias.
- Variable 15 Respuesta que recibe a sus quejas o sugerencias.

A continuación se presenta el resultado del análisis exploratorio, los diagramas de caja y bigotes, histogramas y las gráficas de distribución normal de probabilidad las quince variables.

Se observa que en todas las gráficas de caja y bigotes se ha identificado a la unidad experimental 225 como un dato atípico.

Si bien es cierto que para las variables 4, 5 y 10 se han identificado otros valores fuera de la gráfica de caja y bigotes, no son declarados datos atípicos. En estas variables el diagrama de caja y bigotes se contrae a un solo valor (calificación de nueve) que fue la respuesta en más del 60% de los eventos.

Al considerar el histograma de frecuencias de las variables 4, 5 y 10, se observa una buena aproximación a la curva normal, por lo que se confirma la decisión de incluir estos datos en el estudio.

El gráfico de distribución normal de probabilidad de cada variable nos indica que la normalidad de todas ellas les permitirá ser consideradas en este análisis multivariable.

## VARIABLE 1: CALIDAD DE LOS PRODUCTOS.

Esta variable sobre la calidad de los productos corresponde al primer tema de la encuesta de satisfacción. La variable tiene su media en 9.18 y su desviación estándar es de 0.786, es leptocúrtica, con ligero sesgo negativo de -0.751

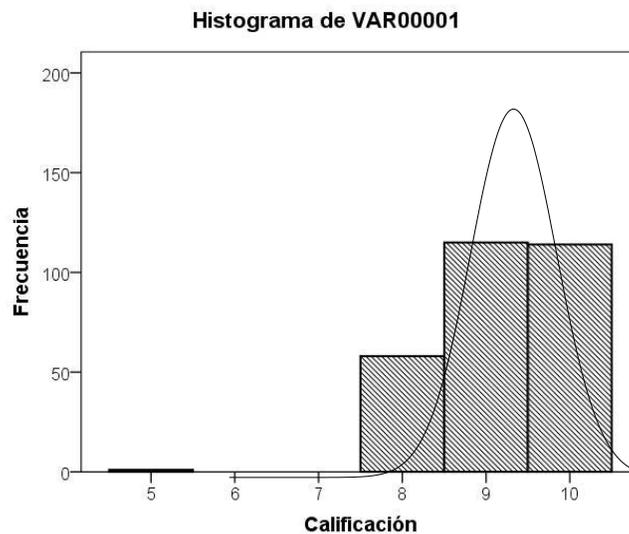
En el histograma (Figura 4.1.1) se puede observar el perfil unimodal de frecuencias y una barra de muy baja frecuencia, aislada al lado izquierdo.

La gráfica de caja y bigotes (Figura 4.1.2) muestra la concentración de observaciones en las calificación de 9 y 10, además indica solo un dato atípico, que no afecta significativamente la distribución de esta variable.

El gráfico de normalidad (Figura 4.1.3), confirma que esta variable es aceptable para el análisis multivariable.

**Figura 4.1.1 Histograma para la variable 1.**

---



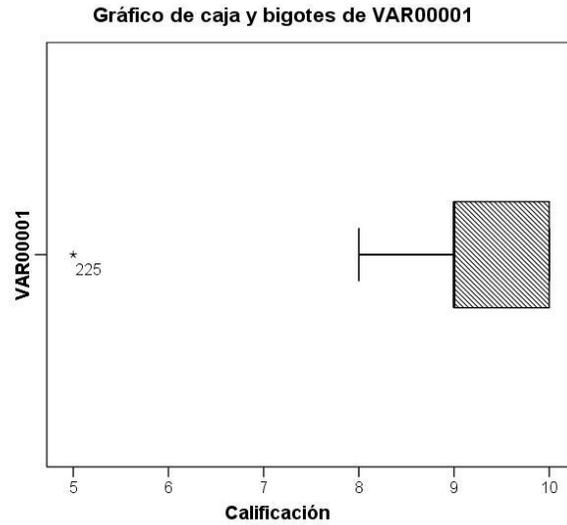
**El histograma sugiere la normalidad de la variable.**

---

**Fuente: Muestra de los datos multivariables**

Figura 4.1.2 Gráfica de caja y bigotes para la variable 1.

---



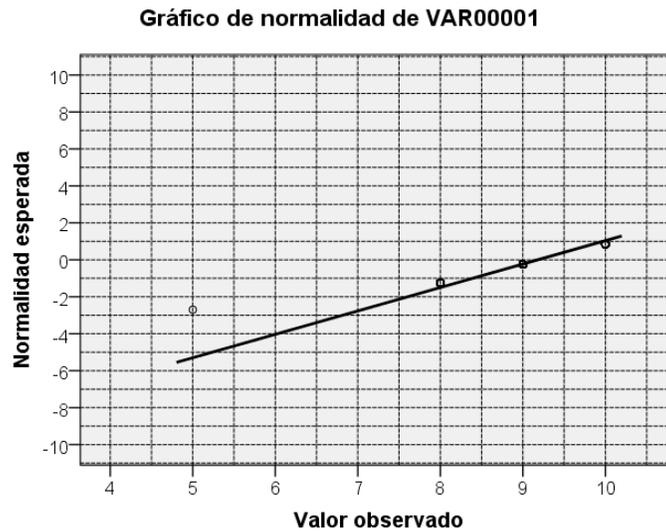
Se observa un dato atípico.

---

Fuente: Muestra de los datos multivariados

Figura 4.1.3 Gráfica de normalidad para la variable 1.

---



El único dato atípico no afecta la normalidad de la variable.

---

Fuente: Muestra de los datos multivariados

## VARIABLE 2: CONDICIONES GENERALES DEL ENVASE.

El segundo tema de la encuesta de satisfacción se refiere a las condiciones del envase. La variable tiene su media en 9.61 y el valor de su desviación estándar es 0.588.

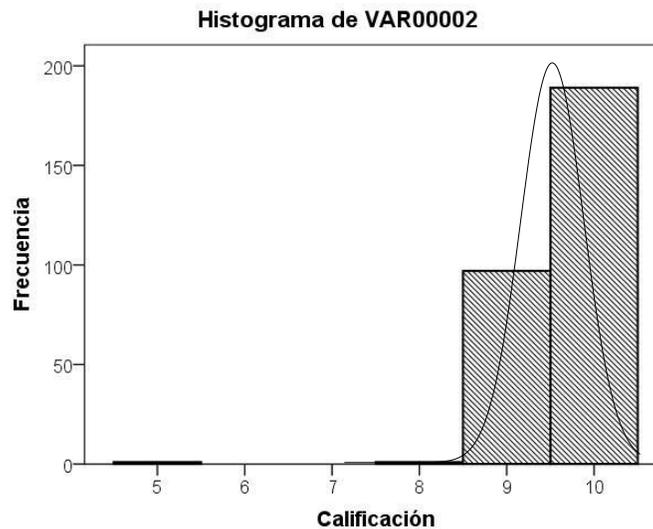
Como se observa en la figura 4.1.4 tiene fuerte curtosis de 13.825, pero solo mediano sesgo negativo de -2.258

La gráfica de caja y bigotes en la figura 4.1.5 indica solo un dato atípico, que no afecta significativamente la normalidad.

La gráfica de normalidad en la figura 4.1.6 confirma que es aceptada para el análisis multivariable.

**Figura 4.1.4 Histograma para la variable 2.**

---



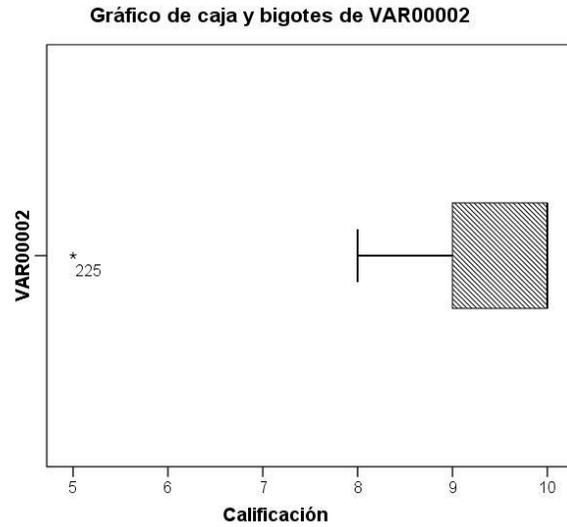
**El histograma sugiere la normalidad de la variable.**

---

**Fuente: Muestra de los datos multivariados**

Figura 4.1.5 Gráfica de caja y bigotes para la variable 2.

---



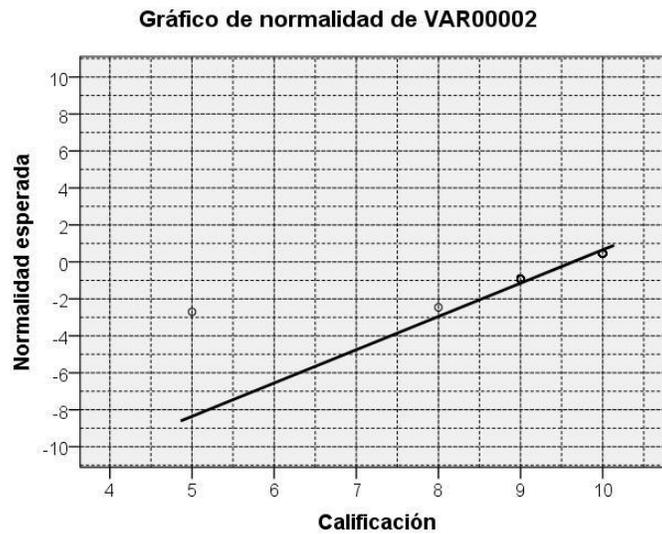
Se observa un dato atípico.

Fuente: Muestra de los datos multivariados

---

Figura 4.1.6 Gráfica de normalidad para la variable 2.

---



El único dato atípico no afecta la normalidad de la variable.

Fuente: Muestra de los datos multivariados

---

### VARIABLE 3: CONDICIONES DEL EMPAQUE.

Las condiciones del empaque son la tercera variable de la encuesta, con media 9.62 y desviación estándar 0.563.

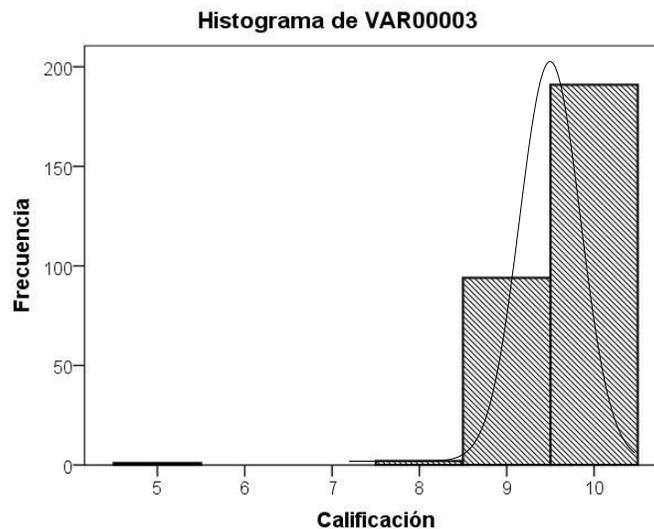
Como se observa en la figura 4.1.7, tiene fuerte curtosis de 13.473 y con mediano sesgo negativo de -2.277.

La gráfica de caja y bigotes indica solo un dato atípico (Figura 4.1.8), que no afecta significativamente la distribución normal.

La gráfica de normalidad (Figura 4.1.9), confirma que es aceptada para el análisis multivariable.

**Figura 4.1.7 Histograma para la variable 3.**

---



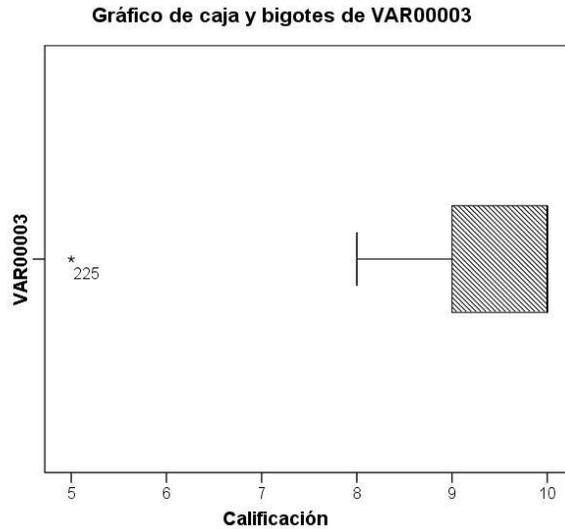
El histograma sugiere la normalidad de la variable.

Fuente: Muestra de los datos multivariados

---

Figura 4.1.8 Gráfica de caja y bigotes para la variable 3.

---



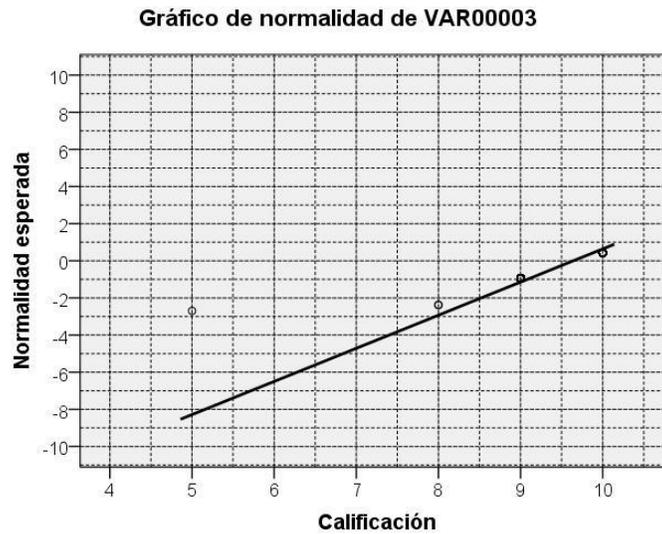
Se observa un dato atípico.

---

Fuente: Muestra de los datos multivariados

Figura 4.1.9 Gráfica de normalidad para la variable 3.

---



El único dato atípico no afecta la normalidad de la variable.

---

Fuente: Muestra de los datos multivariados

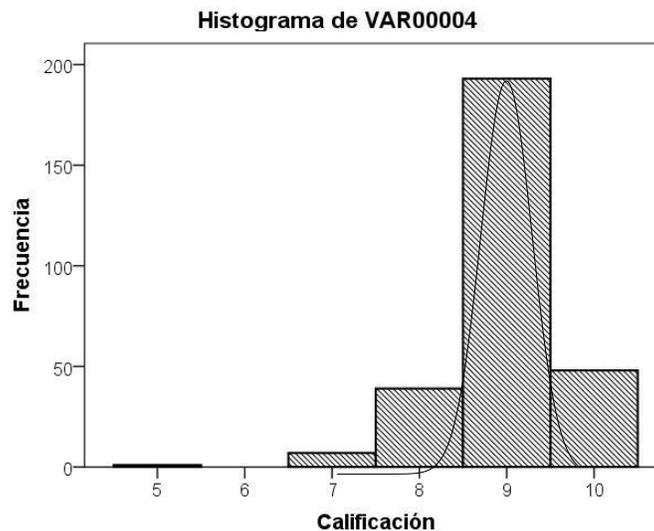
#### VARIABLE 4: CUMPLIMIENTO A LA CANTIDAD SOLICITADA.

La variable de cumplimiento a la cantidad solicitada tiene su media en 8.97 y con una desviación estándar 0.661, con moderada curtosis de 4.866 y con mediano sesgo negativo de -1.156. (Figura 4.1.10).

Toda la gráfica de caja y bigotes se encuentra en la calificación de nueve (Figura 4.1.11), por lo que la gráfica identifica como datos atípicos las calificaciones diferentes. El software SPSS utilizado ejemplificó con los casos 130, 223, 225, 229, 267, 269, 280, 291, 292 y 295. De ellos solo se considera realmente valor atípico el evento 225, que se encuentra a cuatro puntos de la calificación media para esta variable.

Por el apego en la gráfica de normalidad (Figura 4.1.12) es aceptada para el análisis multivariable.

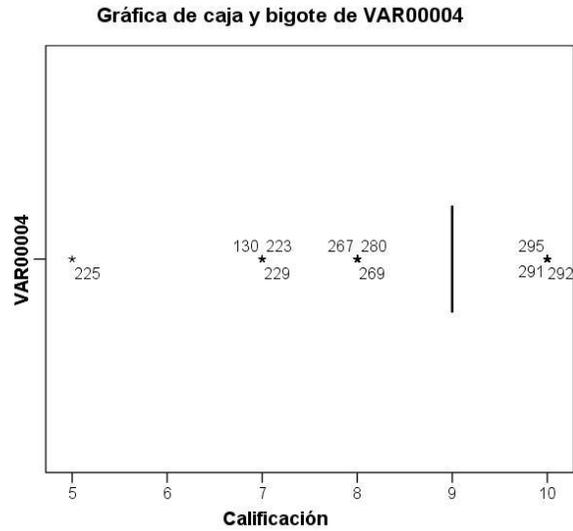
**Figura 4.1.10 Histograma para la variable 4.**



**El histograma sugiere la normalidad de la variable.**

**Fuente: Muestra de los datos multivariados**

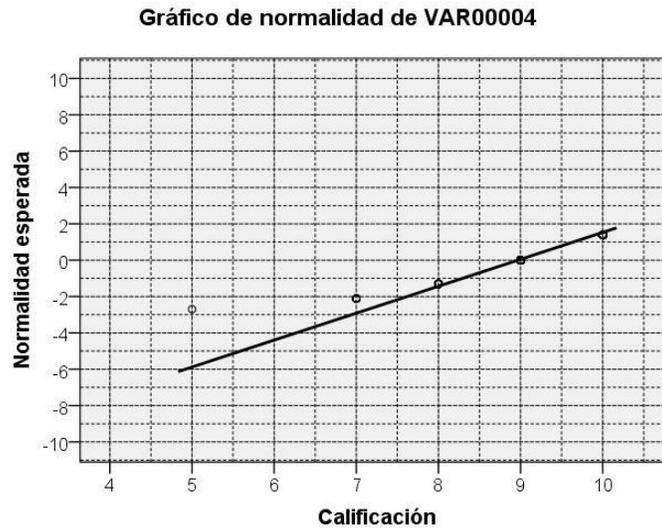
Figura 4.1.11 Gráfica de caja y bigotes para la variable 4.



Se observa que la caja y bigotes coinciden en un solo valor.

Fuente: Muestra de los datos multivariados

Figura 4.1.12 Gráfica de normalidad para la variable 4.



Los datos atípicos no afectan la normalidad de la variable.

Fuente: Muestra de los datos multivariados

VARIABLE 5: CUMPLIMIENTO A LAS MARCAS Y PRESENTACIONES SOLICITADAS.

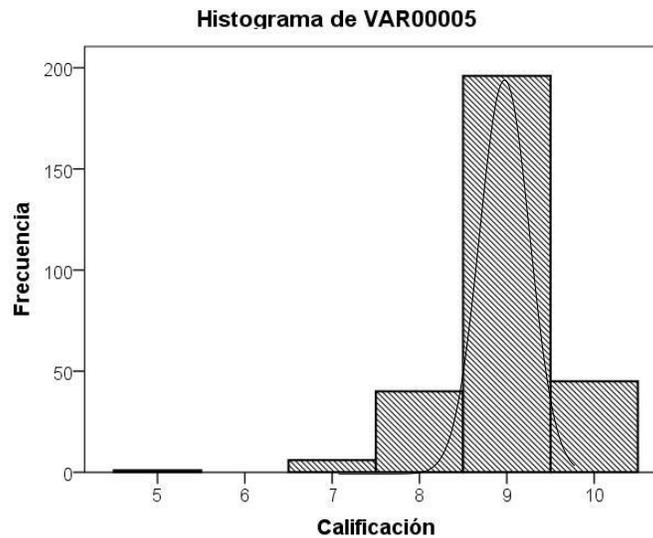
El cumplimiento a las marcas y presentaciones solicitadas también presenta una gráfica de caja y bigotes ocupando sólo la calificación de nueve (Figura 4.1.14).

Es similar a la variable 4 *Cumplimiento a la cantidad solicitada*. Ambas concentran más de dos terceras partes de los datos observados en esa calificación.

La variable 5 tiene media 8.92 y su desviación estándares 0.673. En la figura 4.1.13 se observa su histograma con moderada curtosis de 3.930 y con ligero sesgo negativo de -0.971.

Por el apego en la gráfica de normalidad de la figura 4.1.15, es aceptada para el análisis multivariable.

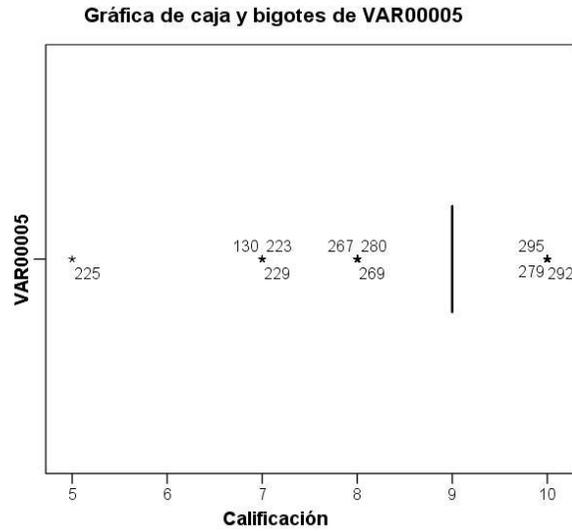
**Figura 4.1.13 Histograma para la variable 5.**



El histograma sugiere la normalidad de la variable.

Fuente: Muestra de los datos multivariados

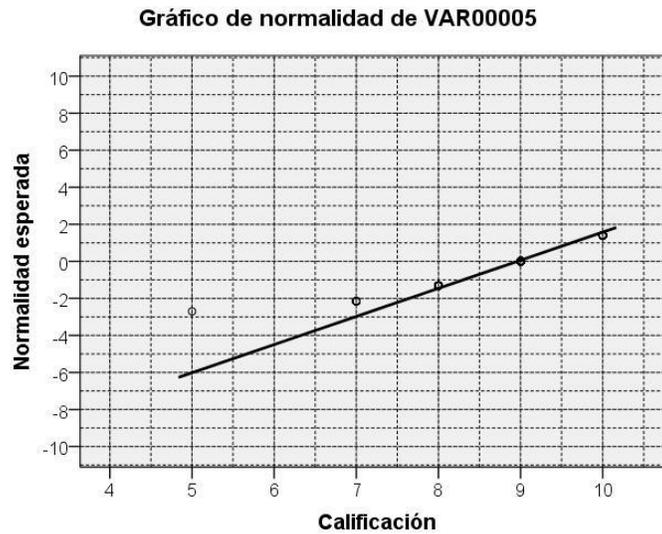
Figura 4.1.14 Gráfica de caja y bigotes para la variable 5.



Se observa que la caja y bigotes coinciden en un solo valor.

Fuente: Muestra de los datos multivariados

Figura 4.1.15 Gráfica de normalidad para la variable 5.



Los datos atípicos no afectan la normalidad de la variable.

Fuente: Muestra de los datos multivariados

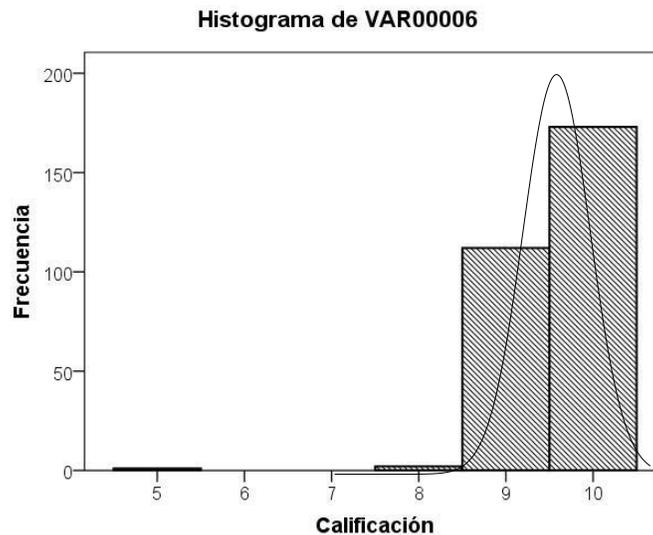
VARIABLE 6: FRECUENCIA CON LA QUE ES VISITADO EN SU ESTABLECIMIENTO.

La sexta variable sobre la frecuencia con la que es visitado en su establecimiento de la encuesta es una variable con distribución normal de frecuencias, con su media en 9.57 y su desviación estándar toma el valor de 0.594, aunque en la figura 4.1.16 se observa que su histograma tiene fuerte curtosis de 10.410 y un mediano sesgo negativo -2.007.

La gráfica de caja y bigotes de la figura 4.1.17 indica solo un dato atípico, nuevamente el 225.

Por el apego en la gráfica de normalidad (Figura 4.1.18) es aceptada para el análisis multivariable.

**Figura 4.1.16 Histograma para la variable 6.**



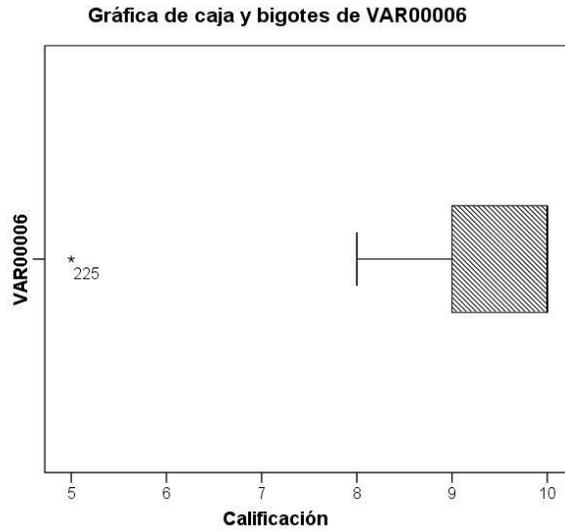
El histograma sugiere la normalidad de la variable.

---

Fuente: Muestra de los datos multivariados

Figura 4.1.17 Gráfica de caja y bigotes para la variable 6.

---



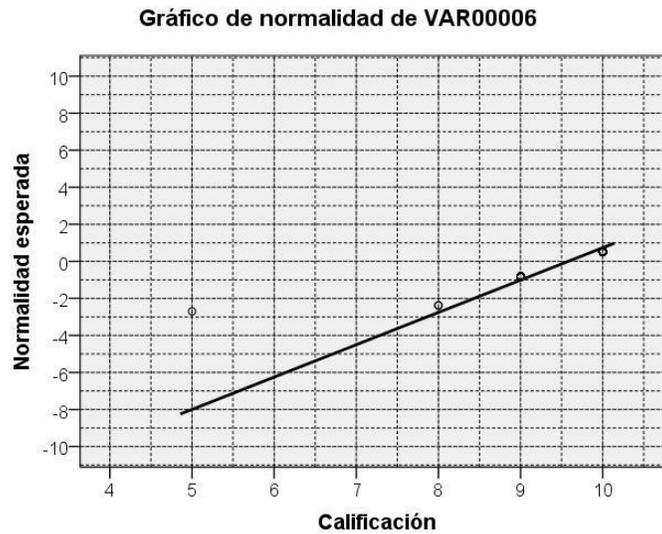
Se observa un dato atípico

Fuente: Muestra de los datos multivariados

---

Figura 4.1.18 Gráfica de normalidad para la variable 6.

---



El único dato atípico no afecta la normalidad de la variable

Fuente: Muestra de los datos multivariados

---

VARIABLE 7: TIEMPO PARA EL RETIRO DE ENVASE Y LA DESCARGA DE PRODUCTO.

La variable siete de la encuesta de satisfacción del cliente se relaciona con el tiempo que tarda la operación de carga y descarga en sus instalaciones.

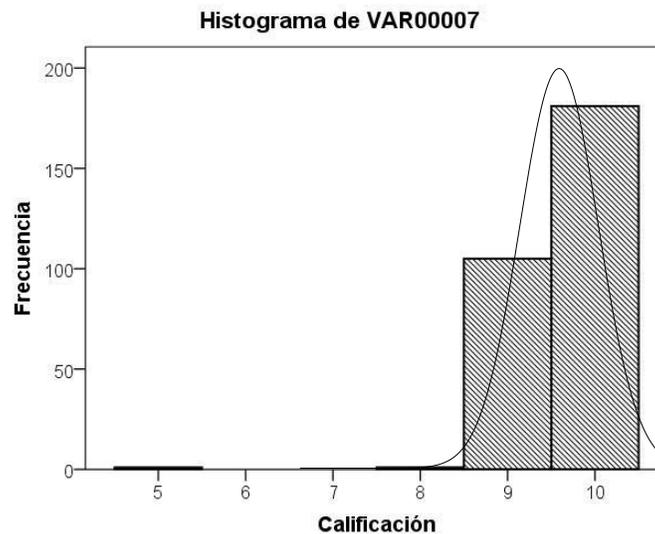
La media de esta variable se encuentra en 9.58 y presenta una desviación estándar de 0.631.

El histograma de la figura 4.1.19 se observa una fuerte curtosis de 10.855, con mediano sesgo negativo de -2.371.

La gráfica de caja y bigotes indica solo un dato atípico (Figura 4.1.20). Por el apego en la gráfica de normalidad (Figura 4.1.21) es aceptada para el análisis multivariable.

**Figura 4.1.19 Histograma para la variable 7.**

---

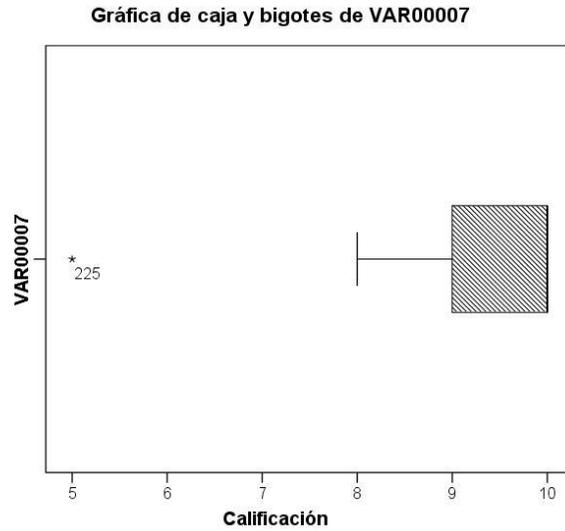


El histograma sugiere la normalidad de la variable.

**Fuente: Muestra de los datos multivariables**

---

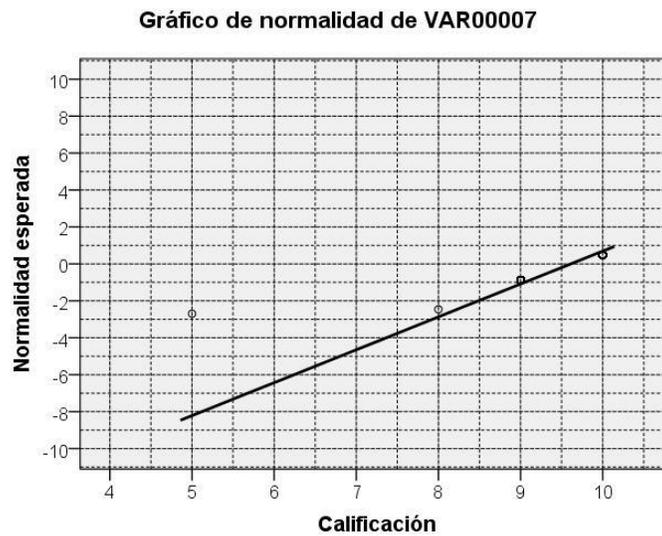
Figura 4.1.20 Gráfica de caja y bigotes para la variable 7.



Se observa un dato atípico.

Fuente: Muestra de los datos multivariados

Figura 4.1.21 Gráfica de normalidad para la variable 7.



El único dato atípico no afecta la normalidad de la variable.

Fuente: Muestra de los datos multivariados

**VARIABLE 8: ATENCIÓN RECIBIDA POR PARTE DEL PERSONAL.**

La atención recibida por parte del personal registró una media de 9.58 y desviación estándar de 0.587.

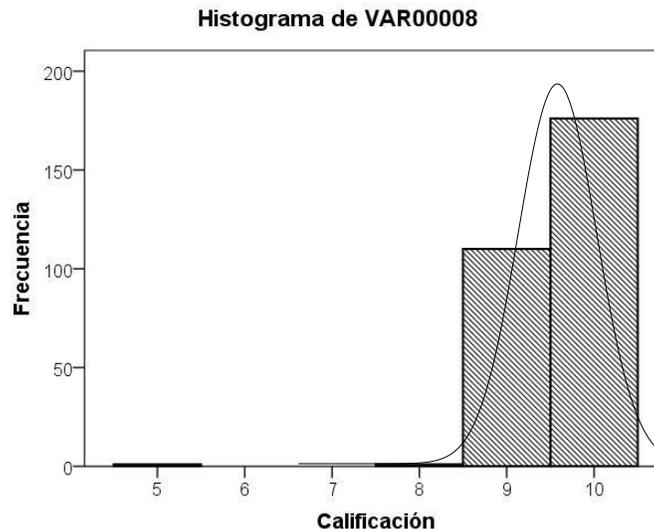
En el histograma se observa fuerte curtosis de 11.021, con mediano sesgo negativo de -2.061 (Figura 4.1.22).

En la gráfica de caja y bigotes se observa indica el único dato atípico (Figura 4.1.23) que se encuentra consistentemente en las variables anteriores.

La presencia de este evento atípico no afecta la normalidad (Figura 4.1.24) de la variable, por lo que es aceptada para el análisis multivariable.

**Figura 4.1.22 Histograma para la variable 8.**

---



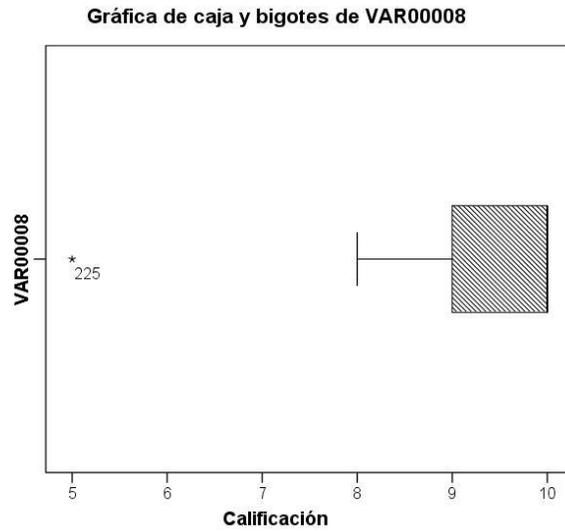
**El histograma sugiere la normalidad de la variable.**

---

**Fuente: Muestra de los datos multivariados**

Figura 4.1.23 Gráfica de caja y bigotes para la variable 8.

---



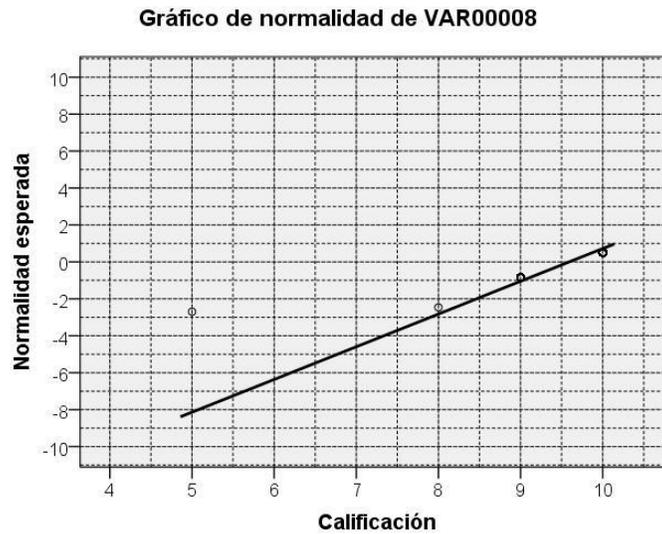
Se observa un dato atípico.

Fuente: Muestra de los datos multivariados

---

Figura 4.1.24 Gráfica de normalidad para la variable 8.

---



El único dato atípico no afecta la normalidad de la variable.

Fuente: Muestra de los datos multivariados

---

**VARIABLE 9: ORIENTACIÓN RESPECTO A LA ROTACIÓN DEL PRODUCTO.**

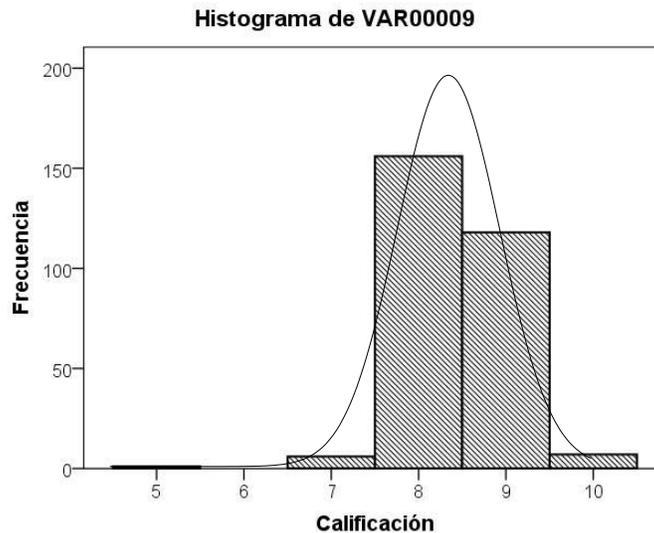
En la figura 4.1.25 se observa de inmediato que la variable 9 sobre la Orientación respecto a la rotación del producto tiene un desplazamiento de la distribución de frecuencias hacia la izquierda, con respecto a las variables antes exploradas.

Esta variable tiene su media en 8.45 y la desviación estándar toma el valor de 0.670, con moderada curtosis de 1.781 y con ligero sesgo negativo de -0.140.

La gráfica caja y bigotes de la figura 4.1.26 indica que la caja ocupa el intervalo de 8 a 9 de calificación y la mediana se encuentra en el extremo izquierdo. Solo identifica el dato atípico correspondiente al evento 225. Por la normalidad observada (Figura 4.1.27) es aceptada para el análisis multivariable.

**Figura 4.1.25 Histograma para la variable 9.**

---



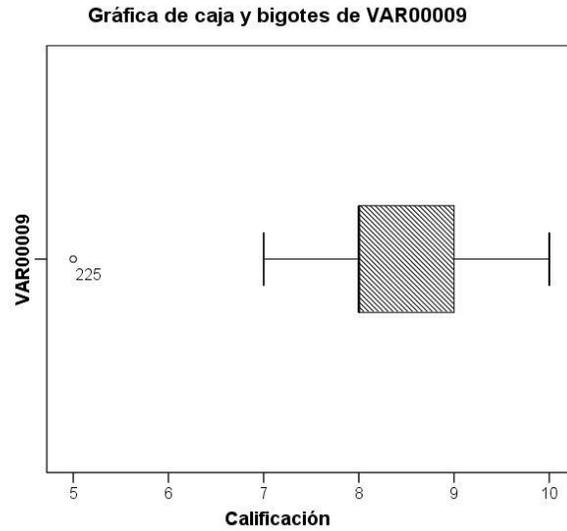
**El histograma sugiere la normalidad de la variable.**

---

**Fuente: Muestra de los datos multivariados**

Figura 4.1.26 Gráfica de caja y bigotes para la variable 9.

---



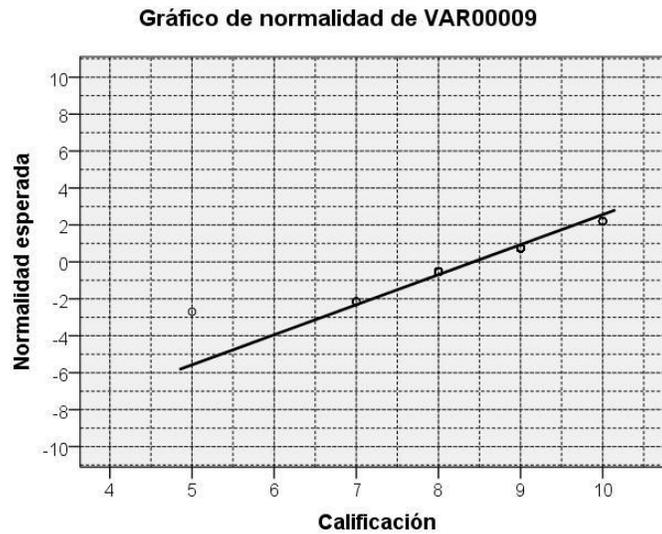
Se observa un dato atípico.

Fuente: Muestra de los datos multivariados

---

Figura 4.1.27 Gráfica de normalidad para la variable 9.

---



El único dato atípico no afecta la normalidad de la variable.

Fuente: Muestra de los datos multivariados

---

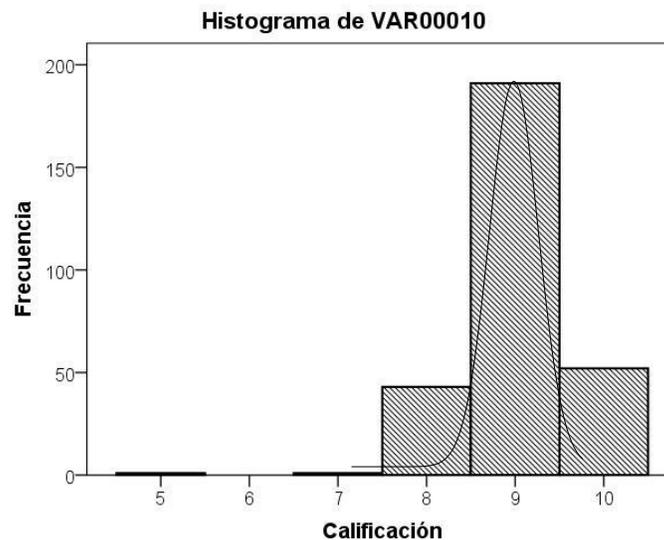
**VARIABLE 10: SERVICIO QUE LE BRINDA LA EMPRESA.**

Esta variable sobre el servicio de le brida la empresa presenta un comportamiento semejante al *cumplimiento a la cantidad solicitada* (variables 4) y al *cumplimiento a las marcas y presentaciones solicitadas* (variable 5), pues la gráfica de caja y bigotes se encuentra en la calificación de nueve.

La variable tiene su media en 9.02 y la desviación estándar es de 0.651, con moderada curtosis de 4.269 y con ligero sesgo negativo de -0.827 (Figura 4.1.28).

La gráfica de caja y bigotes indica datos atípicos (Figura 4.1.29) en las calificaciones 5, 7 , 8 y 10, pero por el apego en la gráfica de normalidad (Figura 4.1.30) solo se considera atípico el evento 225 y la variable 10 es aceptada para el análisis multivariable.

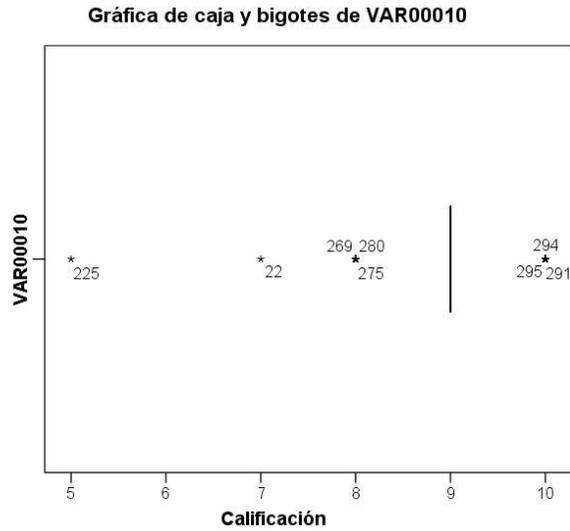
**Figura 4.1.28 Histograma para la variable 10.**



**El histograma sugiere la normalidad de la variable.**

**Fuente: Muestra de los datos multivariabes**

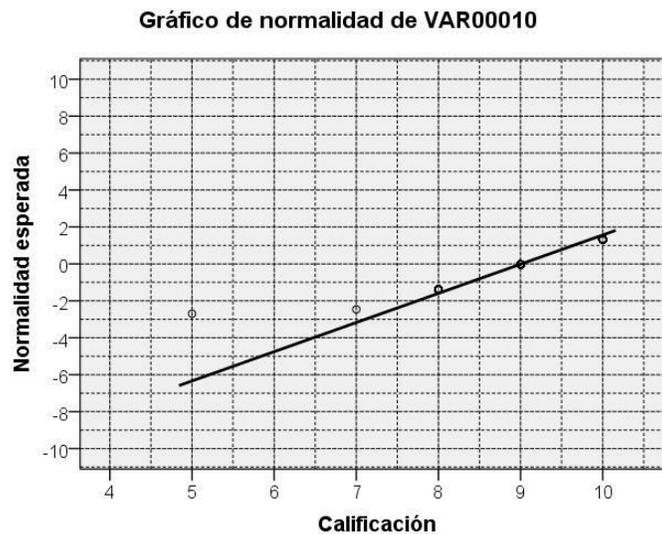
Figura 4.1.29 Gráfica de caja y bigotes para la variable 10.



Se observan ocho datos atípicos y la gráfica de caja y bigotes en un solo valor.

Fuente: Muestra de los datos multivariados

Figura 4.1.30 Gráfica de normalidad para la variable 10.



Los ocho datos atípicos no afecta la normalidad de la variable

Fuente: Muestra de los datos multivariados

**VARIABLE 11: COMUNICACIÓN CON LA EMPRESA.**

El histograma de la variable sobre la comunicación con la empresa (Figura 4.1.31) ocupa la parte del rango de calificaciones y su media se encuentra en 7.58, con desviación estándar 0.687.

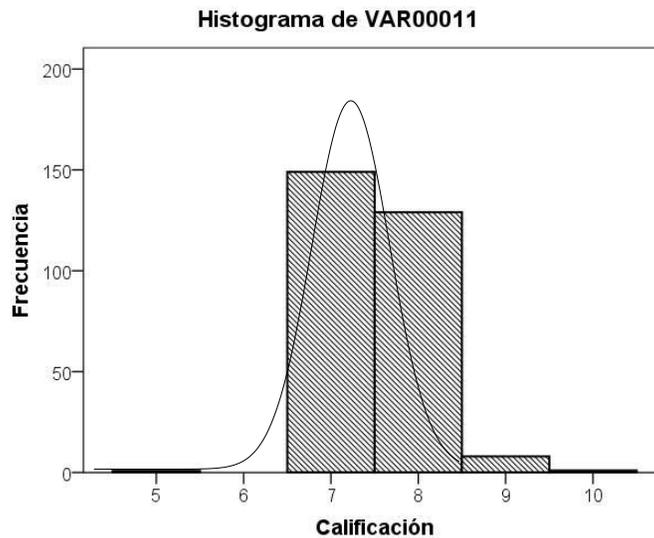
Se observa una moderada curtosis con valor de 1.594, así como un mediano sesgo positivo de 0.831.

La gráfica de caja y bigotes indica dos datos atípicos (Figura 4.1.32), pero solo la calificación del evento 225 está aislada del conjunto de datos.

Se confirma la normalidad de esta distribución en la gráfica de la figura 4.1.33 y es aceptada para el análisis multivariable.

**Figura 4.1.31 Histograma para la variable 11.**

---



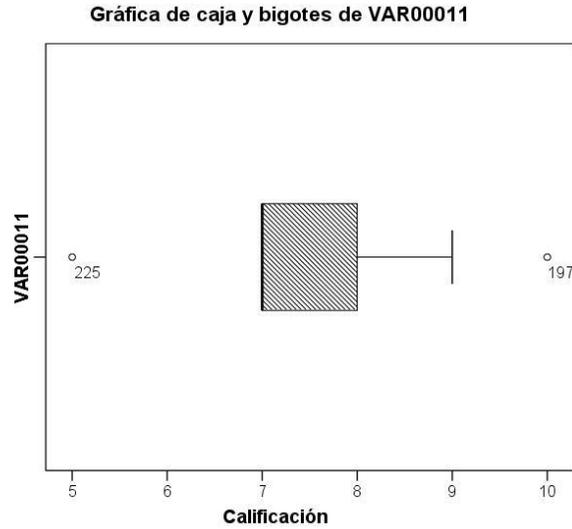
**El histograma sugiere la normalidad de la variable 11.**

---

**Fuente: Muestra de los datos multivariables**

Figura 4.1.32 Gráfica de caja y bigotes para la variable 11.

---



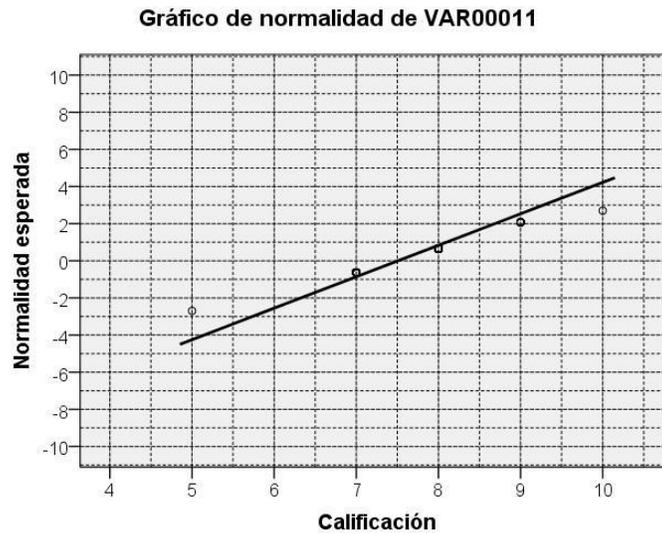
Se observan dos datos atípicos.

Fuente: Muestra de los datos multivariados

---

Figura 4.1.33 Gráfica de normalidad para la variable 11.

---



Los dos datos atípicos no afecta la normalidad de la variable

Fuente: Muestra de los datos multivariados

---

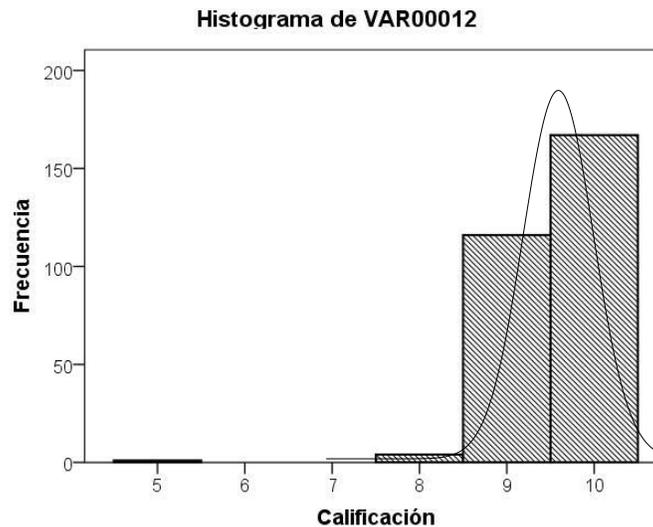
## VARIABLE 12: MEDIOS ACTUALES DE COMUNICACIÓN.

La variable corresponde al décimo segundo tema de la encuesta de satisfacción del cliente, referente a medios actuales de comunicación, tiene su media en la calificación 9.54 y su desviación estándar es de 0.610, con fuerte curtosis de 9.137 y con ligero sesgo negativo de -1.857 (Figura 4.1.34).

La gráfica de caja y bigotes identifica al evento 225 como un dato atípico (Figura 4.1.35), lo que se confirma en la gráfica de normalidad (Figura 4.1.36) que presenta esta calificación alejada cinco desviaciones estándar de la distribución del conjunto de datos.

La variable *Medios actuales de comunicación* es aceptada para el análisis multivariable.

**Figura 4.1.34 Histograma para la variable 12.**

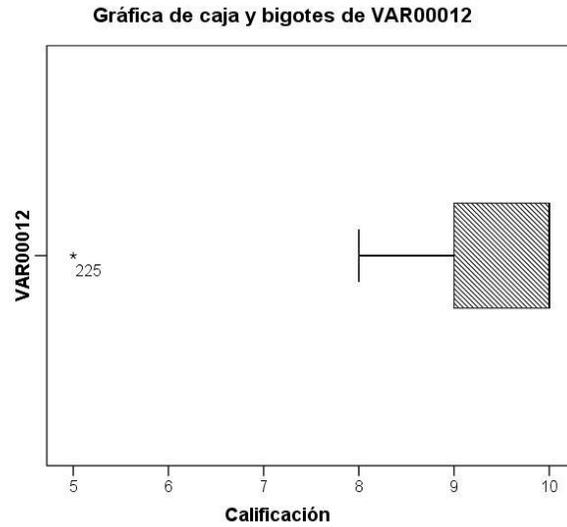


El histograma sugiere la normalidad de la variable 12.

Fuente: Muestra de los datos multivariables

Figura 4.1.35 Gráfica de caja y bigotes para la variable 12.

---



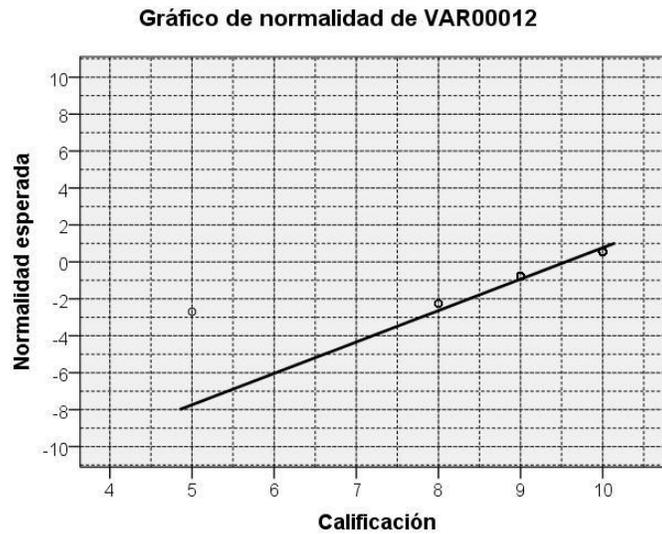
Se observa un dato atípico.

Fuente: Muestra de los datos multivariados

---

Figura 4.1.36 Gráfica de normalidad para la variable 12.

---



El único dato atípico no afecta la normalidad de la variable.

Fuente: Muestra de los datos multivariados

---

### VARIABLE 13: INFORMACIÓN SOBRE LOS PRODUCTOS.

La variable información sobre los productos de la encuesta de satisfacción del cliente tiene su media en 8.57 y la desviación estándar toma el valor de 0.676, la distribución es leptocúrtica, con ligero sesgo positivo de 1.770 (Figura 4.1.37).

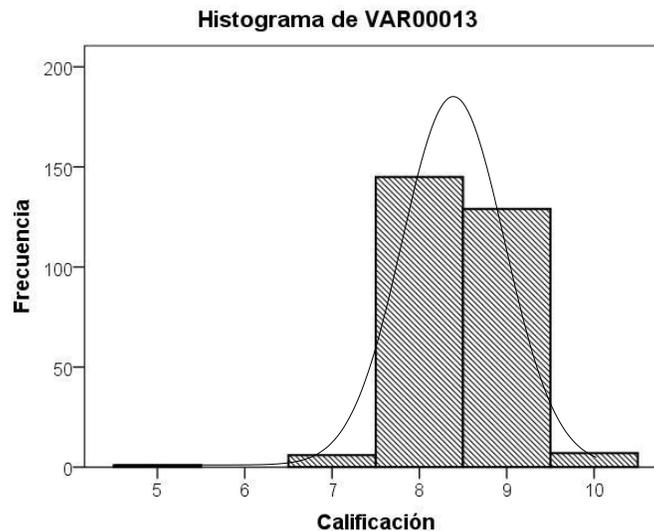
En el conjunto de las 300 unidades experimentales se observó la ausencia de cinco respuestas para esta variable.

La gráfica de caja y bigotes indica sólo un dato atípico (Figura 4.1.38), por el apego en la gráfica de normalidad (Figura 4.1.39)

La variable *Información sobre los productos* también es aceptada para el análisis matricial de datos.

**Figura 4.1.37 Histograma para la variable 13.**

---



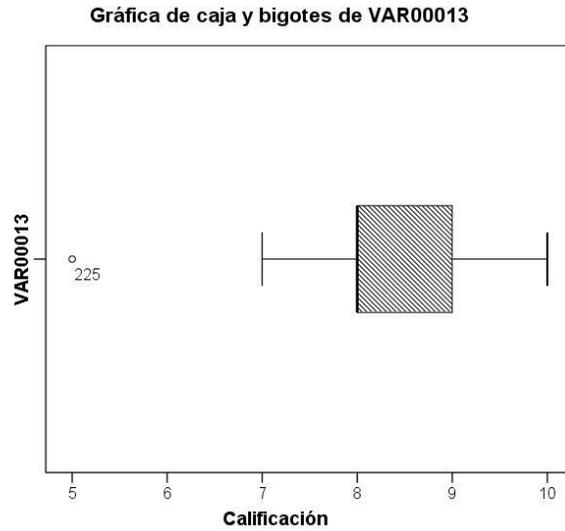
El histograma sugiere la normalidad de la variable 13.

---

Fuente: Muestra de los datos multivariados

Figura 4.1.38 Gráfica de caja y bigotes para la variable 13.

---



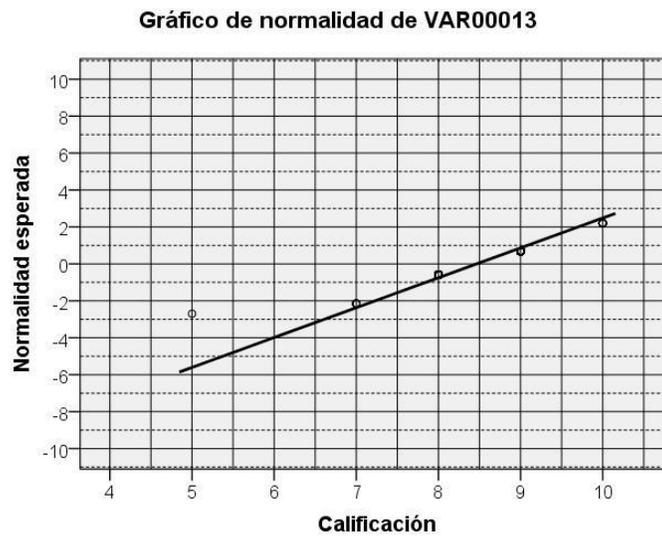
Se observa un dato atípico

Fuente: Muestra de los datos multivariados

---

Figura 4.1.39 Gráfica de normalidad para la variable 13.

---



El único dato atípico no afecta la normalidad de la variable

Fuente: Muestra de los datos multivariados

---

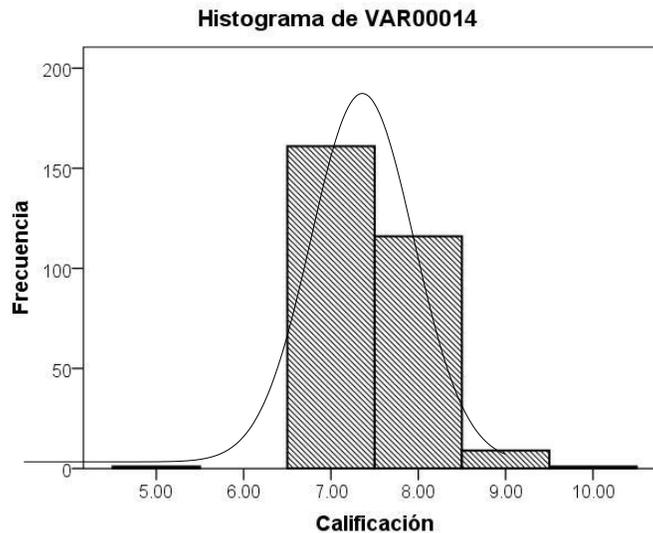
**VARIABLE 14: RAPIDEZ DE RESPUESTA A SUS QUEJAS O SUGERENCIAS.**

Esta variable sobre la rapidez de respuesta a sus quejas o sugerencias presenta una media de 7.51 y desviación estándar de 0.665, con moderada curtosis de 2.310 y con mediano sesgo positivo de 1.023 (Figura 4.1.40) y siete eventos con datos ausentes.

La gráfica de caja y bigotes indica dos datos atípicos (Figura 4.1.41), al igual que la variable *Comunicación con la empresa*. por el apego en la gráfica de normalidad (Figura 4.1.42)

La variable 14 *Rapidez de respuesta a sus quejas o sugerencias* es aceptada para el análisis matricial de datos.

**Figura 4.1.40 Histograma y gráfica de caja y bigote para la variable 14.**

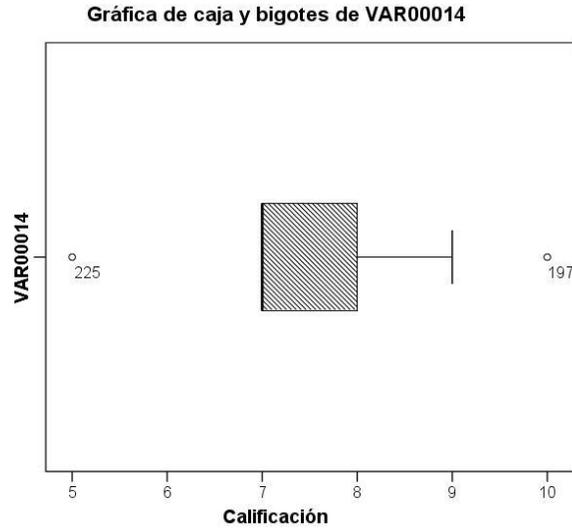


**El histograma sugiere la normalidad de la variable 14.**

**Fuente: Muestra de los datos multivariados**

Figura 4.1.41 Gráfica de caja y bigotes para la variable 14.

---



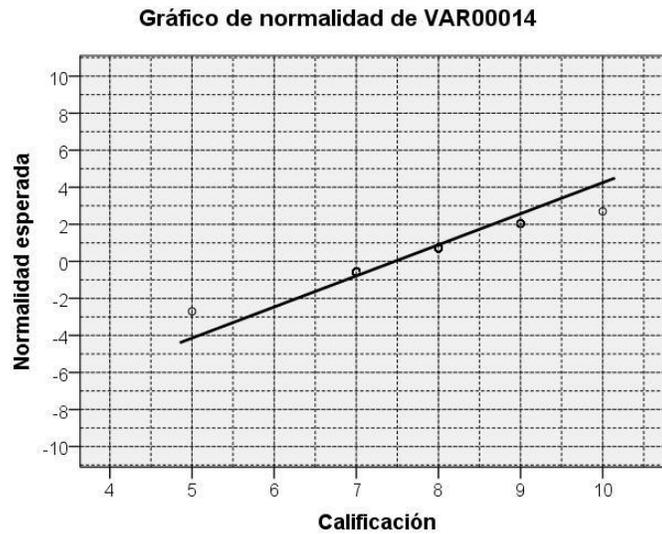
Se observan dos datos atípicos.

Fuente: Muestra de los datos multivariantes

---

Figura 4.1.42 Gráfica de normalidad para la variable 14.

---



Los dos datos atípicos no afecta la normalidad de la variable.

Fuente: Muestra de los datos multivariantes

---

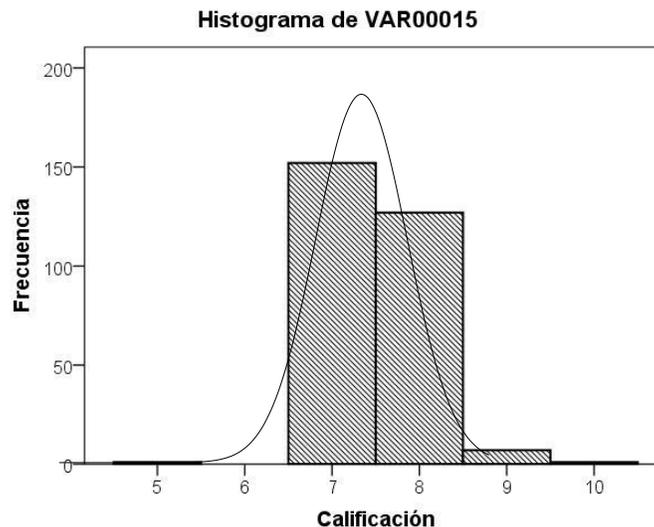
**VARIABLE 15: RESPUESTA QUE RECIBE A SUS QUEJAS O SUGERENCIAS.**

La última variable de la encuesta de satisfacción, que se refiere a la respuesta que recibe a sus quejas o sugerencias, presenta media de 7.54 y desviación estándar igual a 0.664, con moderada curtosis de 2.589, con mediano sesgo positivo de 0.997 (Figura 4.1.43) y siete eventos con datos ausentes, que coinciden con los de la variable *Rapidez de respuesta a sus quejas o sugerencias*.

Al igual que las variables 11 y 14, se observan en la gráfica de caja y bigotes dos datos atípicos (Figura 4.1.44), de ellos se considera realmente atípico el evento 225 por el apego en la gráfica de normalidad (Figura 4.1.45)

También la *Respuesta que recibe a sus quejas o sugerencias* es aceptada para el análisis multivariable.

**Figura 4.1.43 Histograma para la variable 15.**

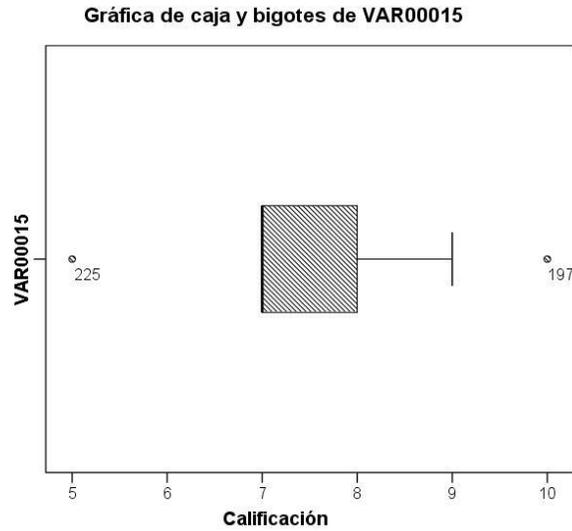


El histograma sugiere la normalidad de la variable 15.

Fuente: Muestra de los datos multivariados

Figura 4.1.44 Gráfica de caja y bigotes para la variable 15.

---



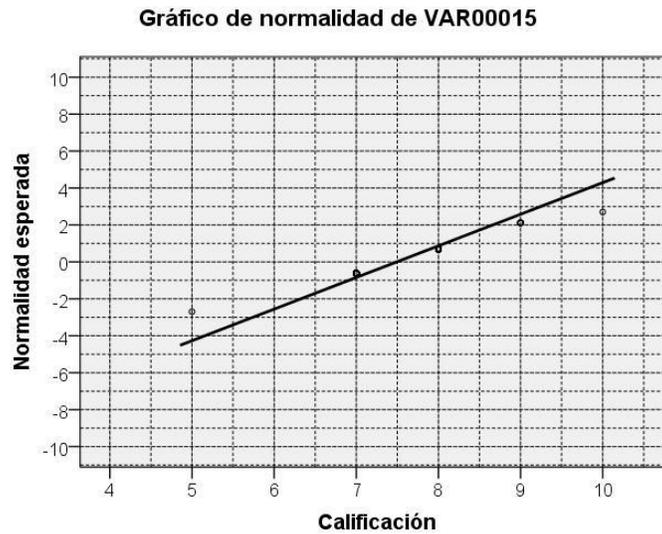
Se observan dos datos atípicos.

---

Fuente: Muestra de los datos multivariantes

Figura 4.1.45 Gráfica de normalidad para la variable 15.

---



Los dos datos atípicos no afecta la normalidad de la variable.

---

Fuente: Muestra de los datos multivariantes

Para realizar el Análisis Factorial no es necesario que el comportamiento del conjunto de variables tenga una distribución normal, mientras cada una de ellas si sea una variable de distribución normal o al menos que el grado en que se aleja de esta normalidad no sea significativo (Pérez 2004). Verificado lo anterior procede continuar con el estudio.

#### 4.2. Análisis exploratorio y gráfico de los eventos.

Con la intención de seleccionar la técnica para identificar los datos ausentes y atípicos se tomó una muestra de 25 casos de los 300 del estudio (tabla 4.2.1) y se aplicaron las dos técnicas referidas: gráficas de radar y gráficas de Andrews.

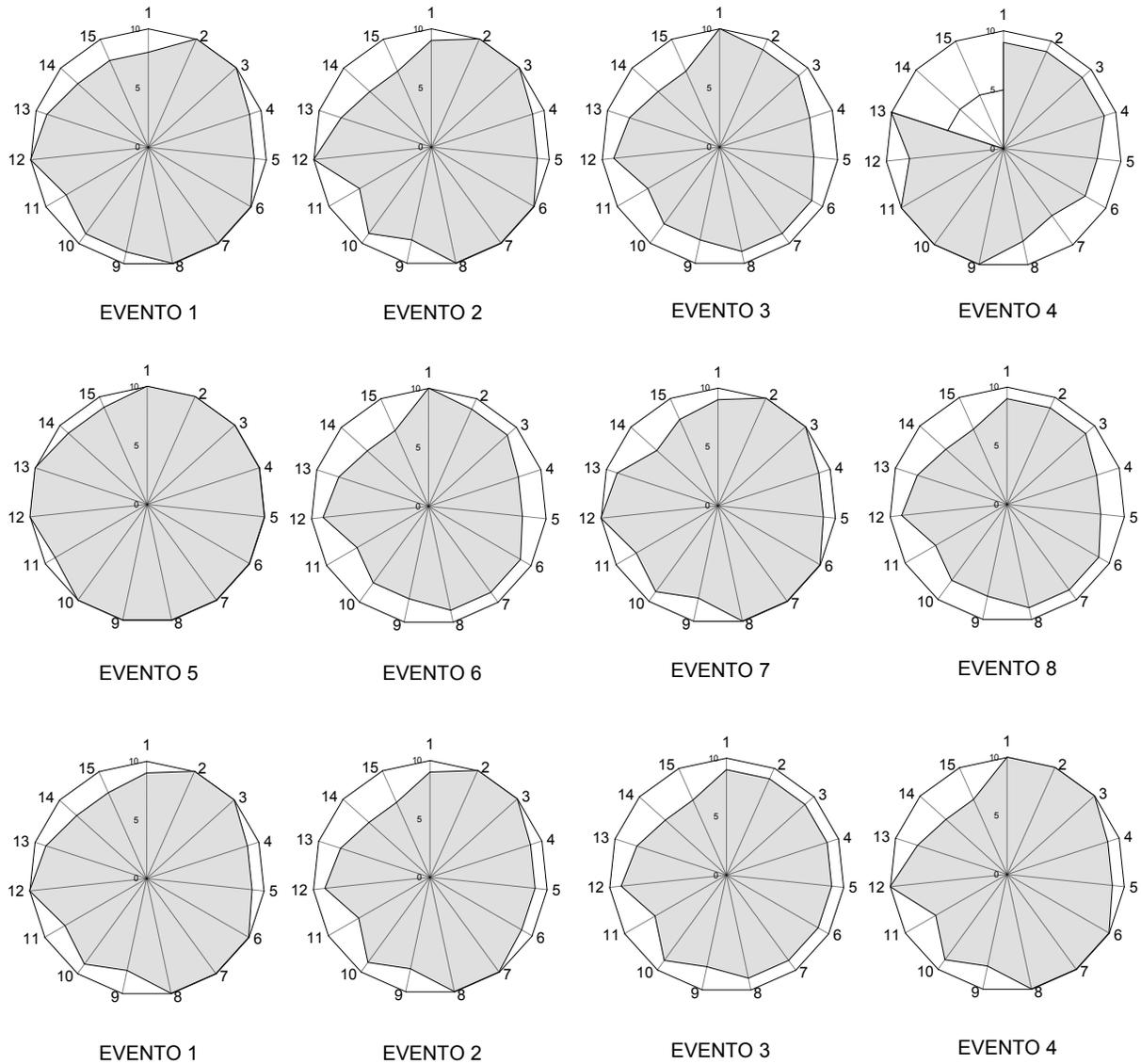
**Tabla 4.2.1 Matriz de datos para 25 unidades experimentales y 12 variables**

Evento	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15
1	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
2	9	10	10	9	9	10	10	10	8	9	7	10	8	7	7
3	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
4	9	9	9	9	8	8	7	8	10	10	10	8	10		
5	10	10	10	10	10	10	10	10	10	10	9	10	10	9	9
6	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
7	9	10	10	9	9	10	10	10	8	9	8	10	9	7	8
8	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
9	9	10	10	9	9	10	10	10	8	9	8	10	9	8	8
10	9	10	10	9	9	9	10	10	8	9	7	9	8	7	7
11	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
12	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
13	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
14	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
15	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
16	10	10	10	10	10	10	10	10	10	10	9	10	10	9	9
17	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
18	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
19	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
20	10	10	10	9	9	10	10	10	9	10	8	10	9	8	8
21	9	9	9	9	8	10	9	10	10	8	9		10	10	10
22	9	8	8	7	7	8	8	8	7	7	7	8	7	7	7
23	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
24	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
25	9	10	10	9	9	9	9	9	8	9	7	9	8	7	7

Fuente: Muestra de los datos multivariados originales.

Las Gráficas de Radar se construyeron utilizando Excel. En la Figura 4.2.1 se muestran las gráficas de radar para los valores de la tabla 4.2.1.

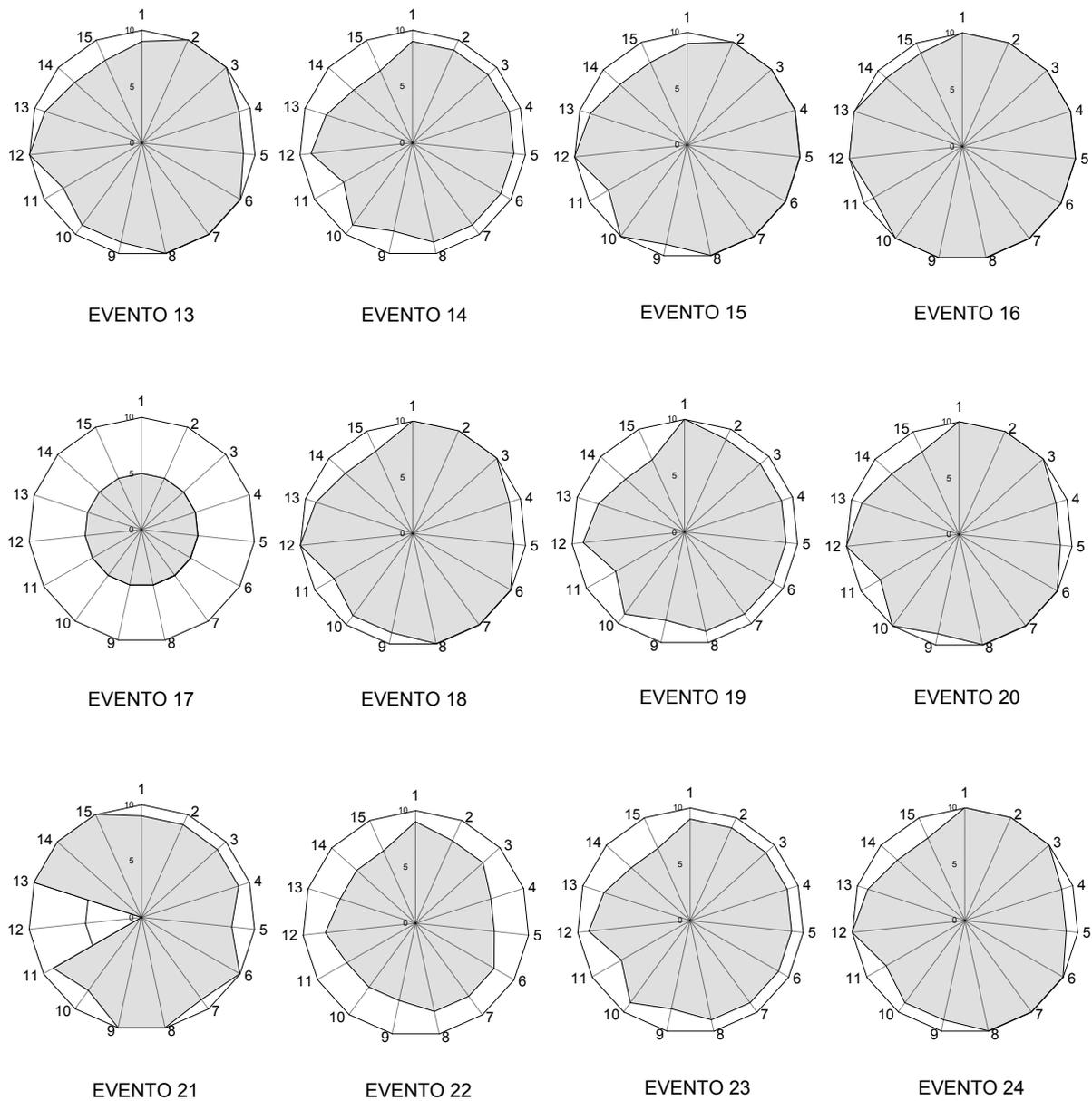
**Figura 4.2.1 Gráficas de radar para comparar 15 variables en 24 eventos.**



**Observe el efecto de un valor ausente en el evento 4.**

**Fuente: Muestra de los datos multivariantes originales.**

**Figura 4.2.1 Gráficas de radar para comparar 15 variables en 24 eventos (continua)**



**Observe el efecto de respuestas atípicas en el evento 17 y un valor ausente en el evento 21.**

**Fuente: Muestra de los datos multivariantes originales.**

El análisis exploratorio por este método sugiere que se deben descartar los casos 4, 17 y 21. El primero por datos ausentes para las variables 14 y 15; el caso 17 por ser atípico en comparación con el comportamiento general de los datos recolectados y el caso 21 por ausencia de evaluación a la variable 12.

Este método es adecuado para el análisis de datos ausentes y atípicos en la matriz de 300 eventos y 15 variables.

Por otro lado el método de gráficas de Andrews requiere graficar la función con la matriz de datos de 300 eventos y 15 variables (donde  $r=1, 2, 3, \dots, 300$ ):

$$f_r(t) = \frac{x_{r1}}{\sqrt{2}} + x_{r2}\text{sen}(t) + x_{r3} \cos(t) + x_{r4}\text{sen}(2t) + x_{r5} \cos(2t) + x_{r6}\text{sen}(3t) \\ + x_{r7} \cos(3t) + x_{r8}\text{sen}(4t) + x_{r9} \cos(4t) + x_{r10}\text{sen}(5t) \\ + x_{r11} \cos(5t) + x_{r12}\text{sen}(6t) + x_{r13} \cos(6t) + x_{r14}\text{sen}(7t) \\ + x_{r15}\cos(7t)$$

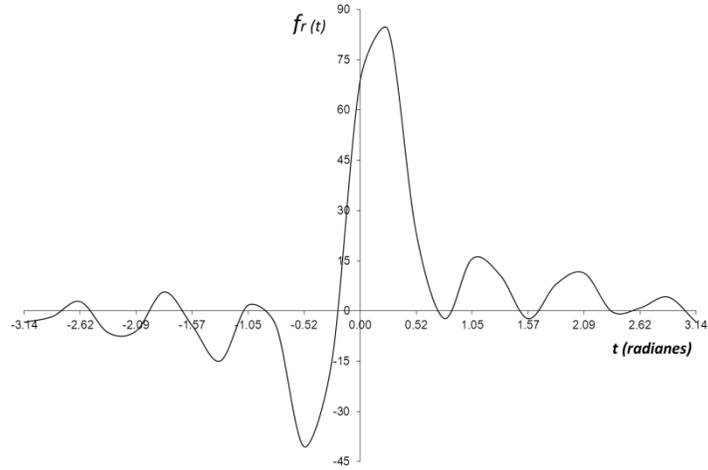
Para el primer evento de la matriz de datos, la función de Andrews toma la forma:

$$f_r(t) = \frac{8}{\sqrt{2}} + 10\text{sen}(t) + 10 \cos(t) + 9\text{sen}(2t) + 9 \cos(2t) + 10\text{sen}(3t) \\ + 10 \cos(3t) + 10\text{sen}(4t) + 9 \cos(4t) + 9\text{sen}(5t) + 8 \cos(5t) \\ + 10\text{sen}(6t) + 9 \cos(6t) + 8\text{sen}(7t) + 8\cos(7t)$$

Se evalúa esta función en el intervalo  $[-\pi < t < \pi]$  y en la gráfica 4.2.2 se muestra la gráfica de Andrews para el primer evento. Con el fin de ilustrar como surgen los datos atípicos se elaboraron las gráficas 4.2.3, 4.2.4 y 4.2.5 para 15, 20 y 25 casos.:

Figura 4.2.2 Gráfica de Andrews para  $r=1$

---



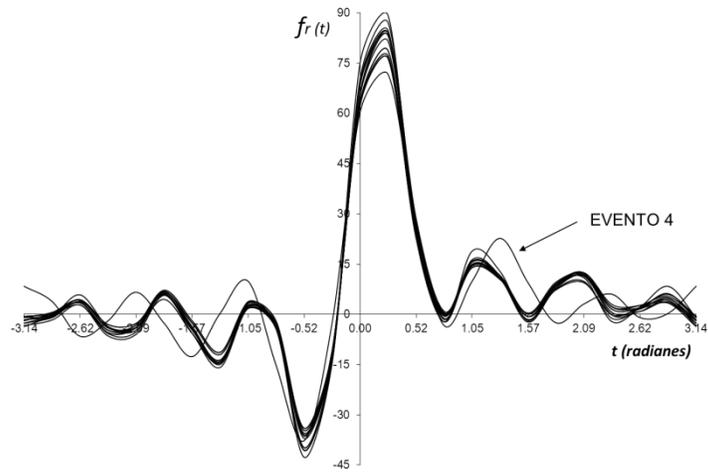
Perfil esperado de las gráficas

Fuente: Fuente: Muestra de los datos multivariantes.

---

Figura 4.2.3 Gráficas de Andrews para los primeros 15 eventos

---



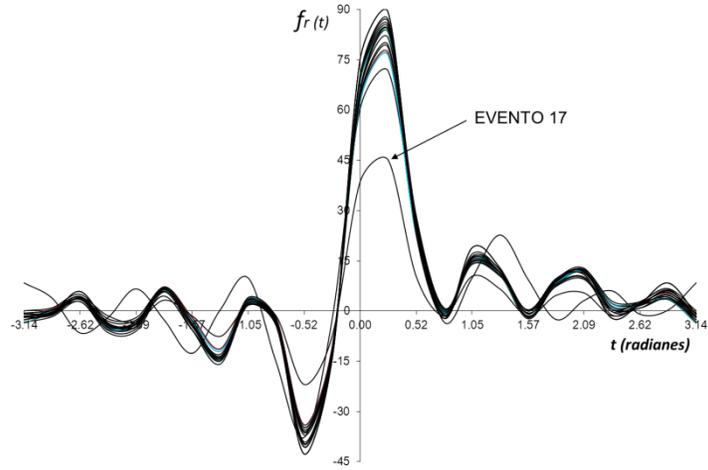
Efecto del evento 4

Fuente: Fuente: Muestra de los datos multivariantes.

---

Figura 4.2.4 Gráficas de Andrews para los primeros 20 eventos

---



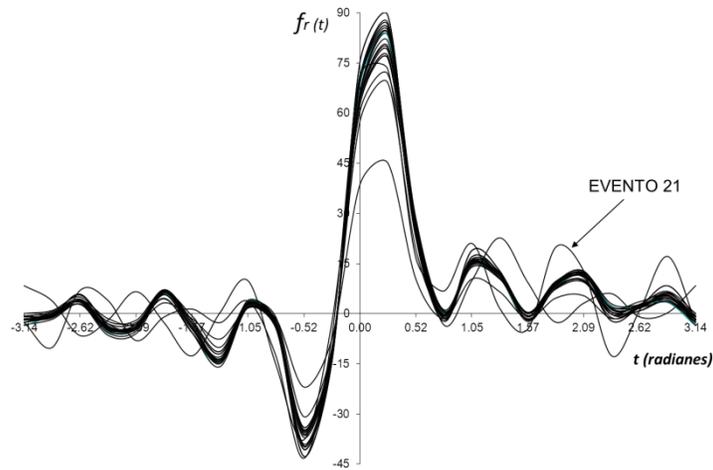
Efecto del evento 17

Fuente: Fuente: Muestra de los datos multivariantes.

---

Figura 4.2.5 Gráficas de Andrews para los primeros 25 eventos

---



Efecto del evento 21

Fuente: Fuente: Muestra de los datos multivariantes.

---

Por este método también se llega a la decisión de eliminar las unidades experimentales 4, 17 y 21, con la ventaja de una mejor interpretación visual.

Para el análisis exploratorio y gráfico de los eventos se decidió utilizar las gráficas de Andrews en las 300 unidades experimentales.

Este análisis llevó a la identificación de 12 eventos con ausencia de la información en algunas de las variables y otro más por ser un caso con respuestas atípicas.

La decisión de eliminar las trece unidades experimentales se tomó bajo el criterio de ser menor al 5% de la información disponible al inicio, pues se considera que sigue siendo representativa de la población total.

#### *4.3 Matriz de unidades experimentales y variables. Comprobación de la correlación supuesta.*

Una vez identificadas y eliminadas las unidades experimentales con datos ausentes o atípicos, se procede a consolidar los datos válidos en una matriz donde las unidades experimentales son las filas y las variables originales son las columnas. La matriz con los 287 eventos aceptados se muestra en la tabla 5.3.1.

Esta matriz ya ofrece mayor certidumbre en la interpretación de la satisfacción del cliente, por los análisis exploratorios de variables y datos realizados a los datos originales y el tratamiento a la información que podría sesgar la interpretación.

Pero sigue siendo un estudio univariable que no considera las posibles correlaciones entre ellas.

**Tabla 4.3.1 Matriz de datos original con 287 unidades experimentales y 15 variables.**

No.	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15
1	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
2	9	10	10	9	9	10	10	10	8	9	7	10	8	7	7
3	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
4	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
5	10	10	10	10	10	10	10	10	10	10	9	10	10	9	9
6	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
7	9	10	10	9	9	10	10	10	8	9	8	10	9	7	8
8	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
9	9	10	10	9	9	10	10	10	8	9	8	10	9	8	8
10	9	10	10	9	9	9	10	10	8	9	7	9	8	7	7
11	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
12	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
13	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
14	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
15	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
16	10	10	10	10	10	10	10	10	10	10	9	10	10	9	9
17	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
18	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
19	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
20	10	10	10	9	9	10	10	10	9	10	8	10	9	8	8
21	9	8	8	7	7	8	8	8	7	7	7	8	7	7	7
22	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
23	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
24	9	10	10	9	9	9	9	9	8	9	7	9	8	7	7
25	9	10	10	9	9	9	10	10	8	9	7	9	8	7	7
26	10	9	9	7	8	9	9	9	8	8	7	9	8	7	7
27	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
28	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
29	9	10	10	9	9	9	10	10	9	8	9	7	9	8	7
30	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
31	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
32	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
33	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
34	9	10	10	9	9	9	10	10	8	9	7	9	8	7	7
35	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
36	8	10	10	10	10	10	10	10	9	10	8	10	9	8	8
37	9	10	10	9	9	10	10	10	8	9	8	10	9	7	7
38	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
39	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
40	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
41	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
42	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
43	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
44	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
45	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
46	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
47	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
48	9	10	10	9	9	9	10	9	8	9	7	9	8	7	7
49	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
50	9	10	10	10	10	10	10	10	10	10	9	10	10	9	9

Fuente: Muestra de los datos multivariantes.

**Tabla 4.3.1 Matriz de datos original...(continua)**

No.	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15
51	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
52	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
53	10	10	10	9	9	9	9	9	8	9	7	9	8	7	7
54	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
55	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
56	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
57	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
58	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
59	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
60	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
61	8	9	9	8	8	9	9	9	8	8	7	9	8	7	7
62	9	9	9	7	7	9	9	9	7	8	7	8	7	7	7
63	9	10	10	9	9	10	10	10	8	9	7	10	8	7	7
64	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
65	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
66	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
67	8	9	9	8	8	9	9	9	8	8	7	9	8	7	7
68	8	9	9	8	8	9	9	9	8	8	7	9	8	7	7
69	9	10	10	9	9	10	10	10	8	9	7	10	8	7	7
70	8	10	10	9	9	10	10	10	8	9	7	10	8	7	7
71	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
72	10	10	10	9	9	10	10	10	9	10	8	10	9	8	8
73	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
74	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
75	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
76	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
77	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
78	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
79	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
80	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
81	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
82	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
83	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
84	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
85	10	10	10	10	9	10	10	10	9	10	8	10	9	8	8
86	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
87	10	10	10	9	9	10	10	10	8	9	8	10	9	7	8
88	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
89	9	10	10	9	9	10	10	10	8	9	7	10	8	7	7
90	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
91	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
92	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
93	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
94	10	10	10	10	9	10	10	10	9	10	8	10	9	8	8
95	10	10	10	9	9	9	9	9	8	9	7	9	8	7	7
96	10	10	10	9	9	10	10	10	8	9	8	10	9	7	8
97	8	9	9	8	8	9	9	9	8	8	7	9	8	7	7
98	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
99	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
100	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8

Fuente: Muestra de los datos multivariantes.

**Tabla 4.3.1 Matriz de datos original...(continua)**

No.	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15
101	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
102	9	10	10	9	9	10	10	10	8	9	7	9	8	7	7
103	10	9	9	8	8	9	9	9	8	9	7	9	8	7	7
104	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
105	9	10	10	10	10	10	10	10	9	10	9	10	9	9	9
106	8	10	10	10	10	10	10	10	9	10	8	10	9	8	8
107	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
108	9	10	10	9	9	10	10	10	8	9	7	9	8	7	7
109	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
110	9	10	10	9	9	10	10	10	8	9	7	9	8	7	7
111	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
112	8	10	10	9	9	10	10	10	8	9	7	10	8	7	7
113	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
114	9	10	10	9	9	10	10	10	8	9	7	10	8	7	7
115	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
116	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
117	8	10	10	10	10	10	10	10	9	10	8	10	9	8	8
118	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
119	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
120	10	10	10	9	9	10	10	10	8	9	7	9	8	7	7
121	10	10	10	9	9	10	10	10	9	10	8	10	9	8	8
122	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
123	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
124	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
125	8	10	10	10	10	10	10	10	9	10	8	10	9	9	8
126	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
127	9	9	9	7	7	9	9	9	7	8	7	8	7	7	7
128	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
129	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
130	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
131	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
132	9	10	10	9	9	9	10	9	8	9	7	9	8	7	7
133	9	10	10	9	9	9	10	9	8	9	7	9	8	7	7
134	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
135	8	10	10	9	9	10	10	10	8	9	7	10	8	7	7
136	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
137	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
138	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
139	9	10	10	10	10	10	10	10	10	10	9	10	10	9	9
140	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
141	9	10	10	9	9	10	10	10	8	9	7	9	8	7	7
142	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
143	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
144	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
145	9	10	10	9	9	9	10	9	8	9	7	9	8	7	7
146	8	9	9	8	8	9	9	9	8	8	7	9	8	7	7
147	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
148	8	10	10	9	9	10	10	10	8	9	7	10	8	7	7
149	9	10	10	9	9	10	10	10	8	9	7	9	8	7	7
150	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8

Fuente: Muestra de los datos multivariantes.

**Tabla 4.3.1 Matriz de datos original...(continua)**

No.	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15
151	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
152	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
153	9	10	10	9	9	10	10	10	8	9	8	10	9	7	8
154	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
155	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
156	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
157	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
158	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
159	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
160	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
161	10	10	10	9	9	9	9	9	8	9	7	9	8	7	7
162	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
163	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
164	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
165	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
166	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
167	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
168	8	9	9	8	8	9	9	9	8	8	7	9	8	7	7
169	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
170	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
171	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
172	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
173	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
174	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
175	10	10	10	9	9	10	10	10	8	9	8	10	9	7	8
176	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
177	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
178	8	10	10	9	9	10	10	10	8	9	7	10	8	7	7
179	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
180	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
181	8	9	9	8	8	9	9	9	8	8	7	9	8	7	7
182	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
183	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
184	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
185	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
186	8	10	10	9	9	10	10	10	8	9	7	10	8	7	7
187	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
188	8	10	10	9	9	9	9	9	8	9	7	9	8	7	7
189	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
190	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
191	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
192	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
193	8	9	10	9	9	9	9	9	8	9	7	9	8	7	7
194	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
195	8	10	10	10	10	10	10	10	9	10	8	10	9	8	8
196	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
197	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
198	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
199	8	9	9	8	8	9	9	9	8	8	7	9	8	7	7
200	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8

Fuente: Muestra de los datos multivariantes.

**Tabla 4.3.1 Matriz de datos original...(continua)**

No.	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15
201	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
202	8	10	10	10	10	10	10	10	9	10	8	10	9	8	8
203	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
204	9	10	10	9	9	10	10	10	8	9	8	10	9	7	8
205	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
206	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
207	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
208	10	9	9	8	8	9	9	9	8	9	7	9	8	7	7
209	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
210	9	10	10	9	9	10	10	10	8	9	7	10	8	7	7
211	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
212	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
213	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
214	8	9	8	7	7	8	9	9	7	8	7	8	7	7	7
215	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
216	8	10	10	9	9	9	9	9	8	9	7	9	8	7	7
217	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
218	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
219	9	9	9	7	7	9	9	9	7	8	7	9	7	7	7
220	10	10	10	9	9	9	9	9	8	9	7	9	8	7	7
221	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
222	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
223	8	10	10	9	9	10	10	10	8	9	7	10	8	7	7
224	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
225	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
226	9	9	10	9	9	9	9	9	8	9	7	9	8	7	7
227	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
228	10	10	10	9	9	10	10	10	9	10	8	10	9	8	8
229	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
230	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
231	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
232	9	10	10	9	9	10	10	10	8	9	7	10	8	7	7
233	9	10	10	10	10	10	10	10	10	10	9	10	10	9	9
234	10	9	9	8	8	9	9	9	8	8	7	9	8	7	7
235	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
236	10	10	10	10	10	10	10	10	9	10	9	10	9	9	8
237	9	10	10	9	9	10	10	10	8	9	8	10	8	7	7
238	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
239	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
240	8	9	9	8	8	9	9	9	8	8	7	9	8	7	7
241	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
242	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
243	8	10	10	9	9	9	9	9	8	9	7	9	8	7	7
244	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
245	10	10	10	9	9	10	10	10	8	9	8	10	9	7	8
246	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
247	10	10	10	9	9	10	10	10	8	9	8	10	9	7	8
248	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
249	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
250	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8

Fuente: Muestra de los datos multivariados.

**Tabla 4.3.1 Matriz de datos original...(continua)**

No.	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15
251	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
252	8	9	9	8	8	9	9	9	8	8	7	9	8	7	7
253	10	10	10	10	10	10	10	10	10	10	9	10	10	9	9
254	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
255	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
256	8	10	10	10	10	10	10	10	9	10	8	10	9	8	8
257	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
258	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
259	9	10	10	9	9	10	10	10	8	9	8	10	8	7	7
260	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
261	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
262	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
263	8	9	9	7	7	9	9	9	7	8	7	9	7	7	7
264	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
265	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
266	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
267	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
268	9	9	9	8	8	9	9	9	8	8	7	9	8	7	7
269	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
270	10	10	10	9	9	10	10	10	9	9	8	10	9	8	8
271	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
272	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
273	8	10	10	9	9	10	10	10	9	9	8	10	9	8	8
274	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
275	8	9	9	9	9	9	9	9	8	9	7	9	8	7	7
276	9	10	10	9	9	10	10	10	9	9	8	10	9	8	8
277	8	9	9	8	8	9	9	9	8	8	7	9	8	7	7
278	10	10	10	10	9	10	10	10	9	10	8	10	9	8	8
279	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
280	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7
281	9	10	10	10	10	10	10	10	9	10	8	10	9	8	8
282	10	10	10	10	10	10	10	10	9	10	8	10	9	8	8
283	10	10	10	9	9	10	10	10	8	9	8	10	9	7	8
284	9	9	9	9	9	9	9	9	8	9	7	9	8	7	7
285	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
286	10	9	9	9	9	9	9	9	8	9	7	9	8	7	7
287	10	10	10	9	9	10	10	10	8	9	7	10	8	7	7

**En la matriz no se incluyeron eventos con datos atípicos o ausentes.**

**Fuente: Muestra de los datos multivariantes.**

El principal supuesto en que se basa el análisis matricial de datos es que existe correlación entre las variables originales. Se requiere comprobar que las variables están relacionadas entre sí o no lo están. En caso de no existir correlación entre ellas tampoco existirían factores comunes y el método no proporcionaría un modelo útil.

La matriz de correlaciones se presenta en la Tabla 4.3.2, donde cada elemento de entrada es el índice de correlación entre los 287 datos de la variable fila y los 287 datos de la variable columna.

**Tabla 4.3.2 Matriz de correlaciones para 15 variables.**

Matriz de correlaciones															
	VAR01	VAR02	VAR03	VAR04	VAR05	VAR06	VAR07	VAR08	VAR09	VAR10	VAR11	VAR12	VAR13	VAR14	VAR15
VAR01	1	0.277	0.273	0.187	0.181	0.292	0.277	0.283	0.228	0.22	0.222	0.288	0.256	0.183	0.222
VAR02	0.277	1	0.983	0.667	0.667	0.913	0.957	0.931	0.678	0.675	0.667	0.88	0.715	0.619	0.658
VAR03	0.273	0.983	1	0.68	0.68	0.909	0.941	0.916	0.677	0.678	0.655	0.876	0.713	0.609	0.647
VAR04	0.187	0.667	0.68	1	0.985	0.66	0.658	0.652	0.696	0.955	0.617	0.676	0.695	0.618	0.614
VAR05	0.181	0.667	0.68	0.985	1	0.659	0.657	0.651	0.694	0.945	0.614	0.675	0.694	0.614	0.61
VAR06	0.292	0.913	0.909	0.66	0.659	1	0.953	0.979	0.739	0.666	0.735	0.96	0.78	0.682	0.725
VAR07	0.277	0.957	0.941	0.658	0.657	0.953	1	0.973	0.705	0.668	0.702	0.917	0.744	0.652	0.693
VAR08	0.283	0.931	0.916	0.652	0.651	0.979	0.973	1	0.723	0.665	0.725	0.94	0.763	0.672	0.715
VAR09	0.228	0.678	0.677	0.696	0.694	0.739	0.705	0.723	1	0.687	0.899	0.772	0.952	0.946	0.915
VAR10	0.22	0.675	0.678	0.955	0.945	0.666	0.668	0.665	0.687	1	0.639	0.668	0.682	0.644	0.636
VAR11	0.222	0.667	0.655	0.617	0.614	0.735	0.702	0.725	0.899	0.639	1	0.757	0.946	0.938	0.98
VAR12	0.288	0.88	0.876	0.676	0.675	0.96	0.917	0.94	0.772	0.668	0.757	1	0.815	0.701	0.746
VAR13	0.256	0.715	0.713	0.695	0.694	0.78	0.744	0.763	0.952	0.682	0.946	0.815	1	0.901	0.954
VAR14	0.183	0.619	0.609	0.618	0.614	0.682	0.652	0.672	0.946	0.644	0.938	0.701	0.901	1	0.946
VAR15	0.222	0.658	0.647	0.614	0.61	0.725	0.693	0.715	0.915	0.636	0.98	0.746	0.954	0.946	1

**Por construcción es una matriz cuadrada con diagonal principal igual a la unidad.**

**Fuente: Muestra de los datos multivariantes.**

Se puede observar la existencia de altas correlaciones mayores a 0.6 y muy altas correlaciones mayores a 0.7 en los pares de variables formados, con excepción de la variable uno que presentó valores menores a 0.3.

Para evaluar el efecto de la variable no correlacionada en la matriz 4.3.2 sobre los resultados que ofrece el análisis matricial, se obtuvo la matriz de componentes de la tabla 4.3.3, para tres factores.

En ella se puede observar que la primera variable no es explicada significativamente por ninguno de los tres componentes propuestos, pues en los tres factores la varianza explicada es menor del 0.5 comúnmente aceptado.

**Tabla 4.3.3 Matriz de Componentes para 15 variables.**

Matriz de componentes			
	FACTOR 1	FACTOR 2	FACTOR 3
VAR01	0.301	0.226	-0.225
VAR02	0.895	0.364	-0.076
VAR03	0.891	0.365	-0.051
VAR04	0.815	-0.019	0.561
VAR05	0.813	-0.017	0.560
VAR06	0.924	0.272	-0.150
VAR07	0.912	0.330	-0.125
VAR08	0.920	0.297	-0.149
VAR09	0.896	-0.356	-0.080
VAR10	0.820	-0.023	0.522
VAR11	0.880	-0.383	-0.192
VAR12	0.926	0.198	-0.135
VAR13	0.920	-0.307	-0.122
VAR14	0.851	-0.453	-0.141
VAR15	0.877	-0.405	-0.195

**La variable uno no es explicada por ninguno de los tres factores.**

**Fuente: Muestra de los datos multivariantes.**

La matriz de componentes indica que la opinión sobre *la calidad de los productos* no está significativamente relacionada con la opinión sobre el envase, empaque, servicio, comunicación y las otras variables incluidas en el instrumento de medición de la satisfacción del cliente, como actualmente se realiza.

Existe la posibilidad de que la primera variable sea interpretada como una referencia exclusiva a las características organolépticas y que el cliente no encuentre relación alguna con otras características de calidad, tal y como es interpretada formalmente en la empresa.

Así la calidad de nuestros productos es, en sí misma, una dimensión que debe ser medida y tratada aparte. A partir de este momento la primera variable será excluida del análisis matricial y se trabajará con las variables 2 a 15.

En la tabla 4.3.4 se presenta la matriz de correlaciones obtenida para 14 variables. Observe que las correlaciones así calculadas conservar valores sumamente altos.

**Tabla 4.3.4 Matriz de correlaciones para 14 variables.**

Matriz de correlaciones														
	VAR02	VAR03	VAR04	VAR05	VAR06	VAR07	VAR08	VAR09	VAR10	VAR11	VAR12	VAR13	VAR14	VAR15
VAR02	1	0.978	0.608	0.605	0.887	0.943	0.91	0.628	0.608	0.645	0.846	0.672	0.591	0.634
VAR03	0.978	1	0.623	0.621	0.881	0.922	0.89	0.627	0.612	0.631	0.841	0.67	0.578	0.62
VAR04	0.608	0.623	1	0.983	0.6	0.596	0.59	0.656	0.949	0.584	0.62	0.655	0.586	0.58
VAR05	0.605	0.621	0.983	1	0.596	0.593	0.586	0.654	0.937	0.58	0.617	0.653	0.582	0.576
VAR06	0.887	0.881	0.6	0.596	1	0.939	0.973	0.701	0.598	0.723	0.949	0.75	0.662	0.711
VAR07	0.943	0.922	0.596	0.593	0.939	1	0.965	0.66	0.6	0.686	0.893	0.707	0.628	0.674
VAR08	0.91	0.89	0.59	0.586	0.973	0.965	1	0.681	0.597	0.712	0.924	0.73	0.652	0.7
VAR09	0.628	0.627	0.656	0.654	0.701	0.66	0.681	1	0.644	0.893	0.74	0.946	0.945	0.911
VAR10	0.608	0.612	0.949	0.937	0.598	0.6	0.597	0.644	1	0.607	0.602	0.638	0.614	0.604
VAR11	0.645	0.631	0.584	0.58	0.723	0.686	0.712	0.893	0.607	1	0.746	0.945	0.934	0.979
VAR12	0.846	0.841	0.62	0.617	0.949	0.893	0.924	0.74	0.602	0.746	1	0.791	0.684	0.733
VAR13	0.672	0.67	0.655	0.653	0.75	0.707	0.73	0.946	0.638	0.945	0.791	1	0.897	0.954
VAR14	0.591	0.578	0.586	0.582	0.662	0.628	0.652	0.945	0.614	0.934	0.684	0.897	1	0.942
VAR15	0.634	0.62	0.58	0.576	0.711	0.674	0.7	0.911	0.604	0.979	0.733	0.954	0.942	1

**La matriz de correlación entre las catorce variables es adecuada para obtener los factores.**

**Fuente: Muestra de los datos multivariantes.**

A la matriz anterior se le pueden aplicar contrastes formales de correlación, como el contraste de esfericidad de Barlett (con base en una distribución chi-cuadrada) y la prueba de adecuación muestral de Kaiser, Meyer y Olkin (relación entre la sumatoria de las correlaciones observadas entre la sumatoria de las correlaciones de las parejas tomadas).

Al incrementarse la correlación entre variables, el contraste de Barlett tiende a valores bajos y la prueba de adecuación de KMO tiende a la unidad..

En el caso de la matriz de correlación de la tabla 4.3.4, se encuentra una significancia de 0.000 para Barlett y de 0.893 para KMO. De esta forma se confirman los supuestos de correlación entre las variables y se acepta continuar con el análisis matricial.

#### *4.4 Extracción de factores*

La matriz de correlación de la tabla 4.3.4 contiene la información de entrada para la extracción de factores. Dado que las correlaciones son mayores o muy cercanas a 0.6 se utilizó el método de *Componentes Principales* para la extracción de factores, como se mencionó en la sección 3.4 de la Metodología.

En la tabla 4.4.1 se presenta la matriz de varianza explicada. En la primera columna se encuentra la cantidad de varianza que se condensa como resultado de obtener sucesivamente los valores propios.

La segunda columna presenta esta varianza condensada como fracción porcentual con referencia al total de las catorce variables de la matriz.

En la tercera columna se encuentra el porcentaje acumulado que es utilizado para decidir el número de componentes principales que se utilizará.

El primero de los componentes principales explica la mayor parte de la varianza de los datos originales (75%). En caso de querer reducir las catorce variables originales a un espacio unidimensional quedaría sin explicar el 25% del comportamiento observado.

**Tabla 4.4.1 Matriz de la varianza explicada.**

Matriz de varianza explicada			
	Total	% de la varianza	% acumulado
VAR02	10.513	75.093	75.093
VAR03	1.467	10.477	85.570
VAR04	1.304	9.314	94.884
VAR05	0.227	1.623	96.507
VAR06	0.140	0.998	97.505
VAR07	0.116	0.828	98.333
VAR08	0.067	0.479	98.812
VAR09	0.047	0.339	99.151
VAR10	0.039	0.280	99.431
VAR11	0.021	0.149	99.580
VAR12	0.019	0.136	99.715
VAR13	0.016	0.117	99.832
VAR14	0.015	0.107	99.940
VAR15	0.008	0.060	100.000

**Se observa un perfil asintótico en la explicación de la varianza.**

**Fuente: Muestra de los datos multivariantes.**

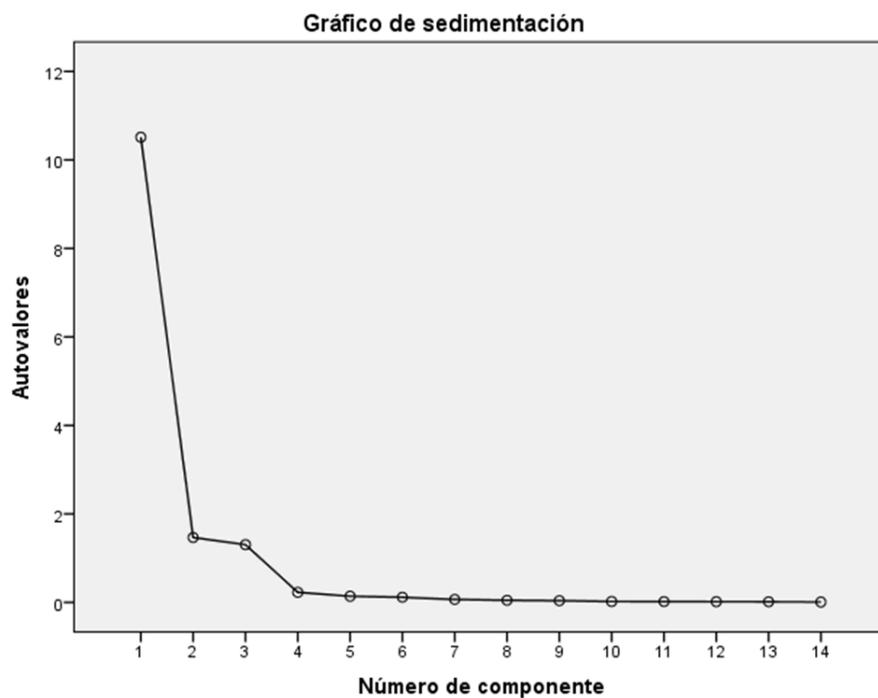
Un modelo de dos dimensiones que tome al primer y segundo componentes incrementa la varianza explicada al 85.6% y con la inclusión del tercer componente se puede explicar el 94.9% de la varianza observada en los 287 casos validados que fueron obtenidos del instrumento de medición de la satisfacción del cliente.

Es posible tomar la decisión de cuántos componentes conservar limitando el modelo a dos o tres dimensiones para facilitar su interpretación o considerar como criterio una cantidad de varianza explicada definido de antemano (90, 95 o 99%), pero tanto Johnson (2004) como Pérez (2004) recomiendan utilizar el gráfico de sedimentación.

#### 4.5 Determinación del número de factores

El gráfico de sedimentación de la figura 4.5.1 se obtiene al representar en el eje vertical las raíces características (valores propios o eigenvalores) y en el eje horizontal enumerar los componentes principales correspondientes, en orden creciente.

Figura 4.5.1 Grafica de Sedimentación.



Se presenta el perfil característico de una fuerte correlación entre variables.

Fuente: Muestra de los datos multivariantes.

En él podemos observar que solo los tres primeros componentes están asociados a valores propios mayores de la unidad y que efectivamente la variación explicada por los siguientes componentes tiende a ser irrelevante.

El criterio generalmente utilizado es tomar solamente aquellos eigenvalores que sean mayores a la unidad, criterio que también se cumple con los tres primeros componentes de la gráfica 4.5.1.

Una vez tomada la decisión de utilizar sólo tres componentes y a partir de la matriz de correlaciones 5.3.4, se elabora la matriz de componentes y cargas de cada variable en la tabla 5.5.1.

**Tabla 4.5.1 Matriz de Componentes y cargas de cada variable.**

Matriz de componentes			
	FACTOR 1	FACTOR 2	FACTOR 3
VAR02	0.874	-0.414	0.024
VAR03	0.868	-0.407	0.052
VAR04	0.784	0.205	0.573
VAR05	0.780	0.204	0.572
VAR06	0.910	-0.333	-0.085
VAR07	0.896	-0.395	-0.042
VAR08	0.905	-0.364	-0.079
VAR09	0.882	0.344	-0.188
VAR10	0.783	0.213	0.540
VAR11	0.882	0.301	-0.294
VAR12	0.911	-0.244	-0.091
VAR13	0.910	0.280	-0.213
VAR14	0.851	0.385	-0.268
VAR15	0.879	0.325	-0.303

**Observe la elevada concentración de cargas en el primer componente**

**Fuente: Muestra de los datos multivariantes.**

Pérez (2004) advierte que en aquellos casos cuando un solo factor tenga cargas altas para todas las variables combinado con la existencia de cargas negativas en los otros factores, se tendrá una situación que sugiere la necesidad de realizar la rotación de los ejes.

#### 4.6 Rotación de factores

Para continuar con la condensación exhaustiva de la varianza, se realiza una rotación ortogonal con el método VARIMAX y se obtiene la matriz de componentes rotados de la tabla 4.6.1.

**Tabla 4.6.1 Matriz de componentes rotados.**

Matriz de componentes rotados			
	FACTOR 1	FACTOR 2	FACTOR 3
VAR02	0.880	0.272	0.294
VAR03	0.869	0.259	0.318
VAR04	0.306	0.294	0.897
VAR05	0.304	0.292	0.894
VAR06	0.853	0.400	0.242
VAR07	0.886	0.332	0.255
VAR08	0.872	0.376	0.236
VAR09	0.335	0.839	0.340
VAR10	0.302	0.316	0.871
VAR11	0.378	0.870	0.238
VAR12	0.787	0.456	0.264
VAR13	0.404	0.831	0.313
VAR14	0.291	0.887	0.270
VAR15	0.358	0.886	0.236

**La ausencia de valores negativos indica una mayor explicación de la varianza.**

**Fuente: Muestra de los datos multivariantes.**

Con la rotación ortogonal de los componentes principales se observan dos efectos en la varianza explicada:

Primero, fueron eliminados todos los valores negativos de los factores segundo y tercero. Esto significa que aumentó la varianza recuperada por cada uno de los componente.

Segundo, se observa una redistribuyen de las cargas elevadas en los tres factores seleccionados. Esta nueva distribución de cargas salientes confirma la conveniencia del modelo con tres nuevas variables.

En forma cuantitativa se observa que en la Tabla 4.6.2, Comparación de la varianza explicada, antes y después de la rotación, la varianza explica pasó respectivamente de 10.5%, 1.5% y 1.3% a 5.3%, 4.8% y 3.2%.

**Tabla 4.6.2 Comparación de la varianza explicada, antes y después de la rotación.**

	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
VAR02	10.513	75.093	75.093	5.331	38.080	38.080
VAR03	1.467	10.477	85.570	4.756	33.971	72.051
VAR04	1.304	9.314	94.884	3.197	22.833	94.884

**La rotación de ejes distribuye la carga en los tres factores.**

**Fuente: Muestra de los datos multivariantes.**

#### *4.7 Interpretación de los factores*

En la matriz de factores rotados se remarcaron las celdas con cargas mayores de 0.750 y se presenta como la Tabla 4.7.1 Cargas salientes en cada componente principal. Las cargas que no aparecen remarcadas corresponden a las variables con correlación nula para ese factor.

Interpretar los factores consistió en leer las diferentes cargas de cada una de las catorce variables en los tres factores.

**Tabla 4.7.1 Cargas salientes en cada componente principal.**

	Componentes Principales		
	1	2	3
Variable 02	0.880	0.272	0.294
Variable 03	0.869	0.259	0.318
Variable 04	0.306	0.294	0.897
Variable 05	0.304	0.292	0.894
Variable 06	0.853	0.400	0.242
Variable 07	0.886	0.332	0.255
Variable 08	0.872	0.376	0.236
Variable 09	0.335	0.839	0.340
Variable 10	0.302	0.316	0.871
Variable 11	0.378	0.870	0.238
Variable 12	0.787	0.456	0.264
Variable 13	0.404	0.831	0.313
Variable 14	0.291	0.887	0.270
Variable 15	0.358	0.886	0.236

**Observe los conglomerados de variables de alta comunalidad.**

**Fuente: Muestra de los datos multivariantes.**

#### **4.7.1 Interpretación del primer factor.**

El factor uno explica el 38.08% de la variación común entre las variables (tabla 4.6.2) y presenta el mayor número de salientes (tabla 4.7.1), seis en total.

- condiciones generales del envase (0.880),
- condiciones generales del empaque (0.869),
- frecuencia con la que es visitado en su establecimiento (0.853),
- tiempo para la descarga del producto (0.886),
- atención recibida por parte del empleado (0.872) y
- medios actuales de comunicación (0.787).

Todas las variables registraron altas cargas en la matriz de componentes rotados, por lo que se adoptó el criterio de eliminar las cargas con puntuación menor a 0.750, con lo que se descartan ocho variables.

Las variables condiciones generales del envase y condiciones generales del empaque son elementos tangibles de la calidad, al igual que el cumplimiento del pedido en cuanto a las marcas, presentaciones y cantidades solicitadas, pero sólo las primeras formaron parte de este factor. Las otras variables tangibles conformaron el tercer factor.

Dos de las variables salientes en este factor, frecuencia con la que es visitado en su establecimiento y el tiempo para el retiro del envase y descarga del producto, dan indicio de que evalúa la rapidez con que el cliente recibe el producto comprado para realizar su venta al menudeo.

Usualmente aceptamos que la satisfacción de los clientes por la rapidez de su compra radica en el ahorro en tiempo. Pero surgen argumentos para afirmar lo siguiente:

El programa de visitas a los distribuidores minoristas se basa en el supuesto de que se muestran dispuestos a planificar el manejo de sus inventarios y esperar varios días antes de contar con el producto adquirido. Pero el primer factor sugiere que para el minorista es el ahorro en tiempo la fuente de satisfacción.

La aceptación de los canales de comunicación utilizados por la empresa se encuentra relacionada con la satisfacción con la atención recibida por el personal.

Como este primer factor gira en torno al desempeño, imagen del producto, velocidad de entrega y la atención personalizada, se define como FACILITAR LA VENTA AL MENUDEO: Elementos proporcionados y actividades realizadas durante la interacción con el cliente minorista, que facilitan obtener el beneficio esperado al realizarse la venta al menudeo del producto.

#### **4.7.2 Interpretación del segundo factor**

El factor dos reúne el 33.97 % de la variación común entre las variables y consolida cinco salientes:

- comunicación con la empresas (0.870),
- orientación de la rotación del producto (0.839),
- información de los productos elaborados (0.831),
- rapidez de respuesta a las quejas o sugerencias (0.887) y
- la respuesta en sí a la queja o sugerencia (0.886).

La relación entre las primeras tres de las variables consolidadas fue interpretada como la disponibilidad de información para incrementar el valor del producto comprado.

Se pueden definir las otras dos variables de este factor, rapidez de respuesta a las quejas o sugerencias y la respuesta en sí a la queja o sugerencia, como la respuesta ante la insatisfacción del cliente. La evidencia de que existe la necesidad de adaptación se desprende de las altas cargas de estas dos variables.

Las tres primeras variables son manifestaciones de un esfuerzo proactivo de la empresa para satisfacer las necesidades particulares de cada cliente, pero las variables que involucran la queja o sugerencia solo son esfuerzos reactivos que funcionaron como una corrección inmediata al incumplimiento. Una variable de esta naturaleza se relaciona con la empatía y el tacto.

Esta nueva variable se define como FLUJO DE INFORMACIÓN: proceso de retroalimentación que implica una respuesta rápida para dar asesoría e incrementar el valor del producto comprado, así como para informar sobre fallos en el proceso..

### **4.7.3 Interpretación del Factor 3**

El factor tres consolidó el 22.83 % de la variación común entre las variables. Tres salientes caracterizaron a este factor:

- cumplimiento a la cantidad solicitada (0.897),
- cumplimiento a las marcas y presentaciones solicitadas (0.894).
- servicio que le brinda la empresa (0.871).

Este factor se encuentra fuertemente relacionado con la definición de calidad por parte de ISO 9000:2005, “Grado en el que un conjunto de características inherentes cumplen con los requisitos.”, por lo que se le nombra CUMPLIMIENTO AL PEDIDO.

Los factores uno y tres son los que se complementan con más naturalidad. El primero por ser la expresión de la calidad evaluada a través del proceso en sí y el segundo por ser la expresión de la calidad evaluada según el resultado del proceso.

### **4.8 Matriz de unidades experimentales y factores.**

La representación de los 287 casos experimentales en función de los tres componentes encontrados y ya no con las catorce variables originales, requiere utilizar los coeficientes de proporción llamados puntuaciones factoriales o factores de correlación:

$$r_{jh} = u_{jh}\sqrt{\lambda_h}$$

Estos coeficientes se presentan en la Matriz 4.8.1

**Tabla 4.8.1 Matriz de puntuaciones factoriales.**

Puntuaciones Factoriales			
	1	2	3
Variable 02	0.248	-0.109	-0.037
Variable 03	0.249	-0.117	-0.025
Variable 04	-0.102	-0.114	0.476
Variable 05	-0.096	-0.108	0.45
Variable 06	0.244	-0.061	-0.074
Variable 07	0.253	-0.089	-0.062
Variable 08	0.247	-0.071	-0.074
Variable 09	-0.114	0.276	-0.022
Variable 10	-0.091	-0.079	0.39
Variable11	-0.075	0.288	-0.09
Variable 12	0.217	-0.032	-0.068
Variable 13	-0.067	0.258	-0.049
Variable 14	-0.137	0.323	-0.061
Variable 15	-0.088	0.296	-0.089

**Coefficientes para intercambiar variables observadas por factores principales.**

**Fuente: Muestra de los datos multivariantes.**

Con las puntuaciones factoriales y la matriz de datos original (Tabla 4.3.1) se calcula la calificación de cada una de las nuevas variables:

Para la unidad experimental 1 se tiene:

$$\begin{aligned}
 Z_1 &= 0.248 (8) + 0.249 (10) + \dots - 0.088 (8) = 6.88 \\
 Z_2 &= -0.109 (8) - 0.117 (10) + \dots + 0.296 (8) = 6.62 \\
 Z_3 &= -0.037 (8) - 0.025 (10) + \dots - 0.089 (8) = 6.67
 \end{aligned}$$

A continuación se construyó la matriz de factores (Tabla 4.8.2) para los 287 eventos, como entregable del análisis matricial de datos.

**Tabla 4.8.2 Matriz de factores resultado del análisis matricial de datos.**

No.	FACTOR 1	FACTOR 2	FACTOR 3
1	6.88	6.62	6.67
2	6.88	6.62	6.67
3	6.88	6.62	6.67
4	6.88	6.62	6.67
5	6.62	6.71	6.74
6	6.88	6.62	6.67
7	6.36	6.88	6.71
8	6.98	6.72	6.23
9	6.65	6.67	6.75
10	6.44	6.69	6.82
11	6.98	6.72	6.23
12	6.12	6.93	6.79
13	6.18	6.78	6.89
14	7.17	6.92	5.36
15	7.17	6.92	5.36
16	7.17	6.92	5.36
17	7.17	6.92	5.36
18	6.12	6.93	6.79
19	6.12	6.93	6.79
20	6.18	6.78	6.89
21	6.18	6.78	6.89
22	7.17	6.92	5.36
23	7.17	6.92	5.36
24	7.17	6.92	5.36
25	7.24	5.58	6.79
26	7.24	5.58	6.79
27	7.24	5.58	6.79
28	7.24	5.58	6.79
29	6.12	6.93	6.79
30	6.18	6.78	6.89
31	7.17	6.92	5.36
32	7.17	6.92	5.36
33	7.24	5.58	6.79
34	7.24	5.58	6.79
35	7.24	5.58	6.79
36	7.24	5.58	6.79
37	7.24	5.58	6.79
38	7.36	5.19	6.98
39	7.36	5.19	6.98
40	7.36	5.19	6.98
41	7.36	5.19	6.98
42	7.36	5.19	6.98
43	7.36	5.19	6.98
44	7.24	5.58	6.79
45	7.24	5.58	6.79
46	7.36	5.19	6.98
47	7.36	5.19	6.98
48	7.36	5.19	6.98
49	7.36	5.19	6.98
50	7.36	5.19	6.98

No.	FACTOR 1	FACTOR 2	FACTOR 3
51	7.47	4.80	7.16
52	7.47	4.80	7.16
53	7.47	4.80	7.16
54	7.47	4.80	7.16
55	7.47	4.80	7.16
56	7.47	4.80	7.16
57	7.36	5.19	6.98
58	7.36	5.19	6.98
59	7.47	4.80	7.16
60	7.47	4.80	7.16
61	7.47	4.80	7.16
62	7.47	4.80	7.16
63	7.47	4.80	7.16
64	7.47	4.80	7.16
65	7.47	4.80	7.16
66	7.47	4.80	7.16
67	7.47	4.80	7.16
68	7.65	5.49	5.66
69	7.65	5.49	5.66
70	7.65	5.49	5.66
71	7.65	5.49	5.66
72	7.65	5.49	5.66
73	7.65	5.49	5.66
74	7.65	5.49	5.66
75	7.65	5.49	5.66
76	5.72	7.39	5.70
77	5.72	7.39	5.70
78	5.72	7.39	5.70
79	5.72	7.39	5.70
80	5.72	7.39	5.70
81	5.72	7.39	5.70
82	7.65	5.49	5.66
83	7.65	5.49	5.66
84	7.65	5.49	5.66
85	7.65	5.49	5.66
86	7.65	5.49	5.66
87	7.65	5.49	5.66
88	7.65	5.49	5.66
89	5.72	7.39	5.70
90	5.72	7.39	5.70
91	5.72	7.39	5.70
92	5.72	7.39	5.70
93	5.72	7.39	5.70
94	5.72	7.39	5.70
95	5.72	7.39	5.70
96	7.65	5.49	5.66
97	7.65	5.49	5.66
98	7.65	5.49	5.66
99	7.65	5.49	5.66
100	7.83	4.97	5.72

**Fuente: Muestra de los datos multivariantes.**

**Tabla 5.8.2 Matriz de factores resultado del análisis matricial de datos...(continua)**

No.	FACTOR 1	FACTOR 2	FACTOR 3
101	7.83	4.97	5.72
102	7.83	4.97	5.72
103	7.83	4.97	5.72
104	7.83	4.97	5.72
105	7.83	4.97	5.72
106	5.72	7.39	5.70
107	5.72	7.39	5.70
108	5.72	7.39	5.70
109	5.72	7.39	5.70
110	5.90	6.87	5.76
111	5.90	6.87	5.76
112	5.90	6.87	5.76
113	5.90	6.87	5.76
114	5.90	6.87	5.76
115	5.90	6.87	5.76
116	5.90	6.87	5.76
117	5.90	6.87	5.76
118	7.83	4.97	5.72
119	7.83	4.97	5.72
120	7.83	4.97	5.72
121	7.83	4.97	5.72
122	7.83	4.97	5.72
123	7.83	4.97	5.72
124	7.64	4.98	5.79
125	7.64	4.98	5.79
126	5.90	6.87	5.76
127	5.90	6.87	5.76
128	5.90	6.87	5.76
129	5.90	6.87	5.76
130	5.90	6.87	5.76
131	5.90	6.87	5.76
132	5.90	6.87	5.76
133	5.90	6.87	5.76
134	7.83	4.97	5.72
135	7.83	4.97	5.72
136	7.83	4.97	5.72
137	7.64	4.98	5.79
138	7.64	4.98	5.79
139	5.90	6.87	5.76
140	5.90	6.87	5.76
141	5.90	6.87	5.76
142	7.64	4.98	5.79
143	8.12	5.26	4.41
144	8.12	5.26	4.41
145	6.19	5.96	6.00
146	6.19	5.96	6.00
147	6.19	5.96	6.00
148	6.19	5.96	6.00
149	6.19	5.96	6.00
150	6.19	5.96	6.00

No.	FACTOR 1	FACTOR 2	FACTOR 3
151	6.19	5.96	6.00
152	6.19	5.96	6.00
153	8.12	5.26	4.41
154	8.12	5.26	4.41
155	8.12	5.26	4.41
156	8.12	5.26	4.41
157	7.93	5.27	4.48
158	7.93	5.27	4.48
159	7.86	5.40	4.42
160	7.86	5.40	4.42
161	4.27	6.65	7.60
162	4.27	6.65	7.60
163	4.27	6.65	7.60
164	4.27	6.65	7.60
165	4.27	6.65	7.60
166	4.27	6.65	7.60
167	4.27	6.65	7.60
168	4.27	6.65	7.60
169	6.19	5.96	6.00
170	6.19	5.96	6.00
171	6.19	5.96	6.00
172	6.19	5.96	6.00
173	6.19	5.96	6.00
174	6.19	5.96	6.00
175	6.19	5.96	6.00
176	8.12	5.26	4.41
177	8.12	5.26	4.41
178	7.93	5.27	4.48
179	7.93	5.27	4.48
180	7.86	5.40	4.42
181	7.86	5.40	4.42
182	4.27	6.65	7.60
183	4.27	6.65	7.60
184	4.27	6.65	7.60
185	4.27	6.65	7.60
186	4.27	6.65	7.60
187	4.27	6.65	7.60
188	4.27	6.65	7.60
189	4.27	6.65	7.60
190	6.19	5.96	6.00
191	6.19	5.96	6.00
192	6.19	5.96	6.00
193	6.19	5.96	6.00
194	7.93	5.27	4.48
195	7.86	5.40	4.42
196	7.93	5.27	4.48
197	8.41	4.35	4.65
198	8.41	4.35	4.65
199	8.41	4.35	4.65
200	8.41	4.35	4.65

**Fuente: Muestra de los datos multivariantes.**

**Tabla 5.8.2 Matriz de factores resultado del análisis matricial de datos...(continua)**

No.	FACTOR 1	FACTOR 2	FACTOR 3
201	8.41	4.35	4.65
202	4.27	6.65	7.60
203	4.27	6.65	7.60
204	4.27	6.65	7.60
205	4.27	6.65	7.60
206	6.48	6.25	4.69
207	6.48	6.25	4.69
208	6.48	6.25	4.69
209	6.48	6.25	4.69
210	6.48	6.25	4.69
211	6.48	6.25	4.69
212	8.41	4.35	4.65
213	8.41	4.35	4.65
214	8.41	4.35	4.65
215	8.41	4.35	4.65
216	8.41	4.35	4.65
217	4.56	5.74	7.85
218	4.56	5.74	7.85
219	4.56	5.74	7.85
220	4.56	5.74	7.85
221	6.48	6.25	4.69
222	6.48	6.25	4.69
223	6.48	6.25	4.69
224	6.48	6.25	4.69
225	6.48	6.25	4.69
226	6.48	6.25	4.69
227	8.41	4.35	4.65
228	8.41	4.35	4.65
229	8.41	4.35	4.65
230	4.56	5.74	7.85
231	4.56	5.74	7.85
232	4.56	5.74	7.85
233	4.56	5.74	7.85
234	4.30	5.88	7.86
235	4.30	5.88	7.86
236	6.48	6.25	4.69
237	6.48	6.25	4.69
238	6.48	6.25	4.69
239	6.66	5.74	4.75
240	6.66	5.74	4.75
241	6.66	5.74	4.75
242	6.66	5.74	4.75
243	4.56	5.74	7.85
244	4.56	5.74	7.85
245	4.30	5.88	7.86
246	4.30	5.88	7.86
247	6.66	5.74	4.75
248	6.66	5.74	4.75
249	6.66	5.74	4.75
250	6.66	5.74	4.75

No.	FACTOR 1	FACTOR 2	FACTOR 3
251	6.66	5.74	4.75
252	4.30	5.88	7.86
253	6.66	5.74	4.75
254	6.66	5.74	4.75
255	4.92	5.91	6.40
256	4.92	5.91	6.40
257	4.92	5.91	6.40
258	4.92	5.91	6.40
259	4.92	5.91	6.40
260	4.92	5.91	6.40
261	4.92	5.91	6.40
262	4.99	6.83	4.93
263	4.99	6.83	4.93
264	4.99	6.83	4.93
265	4.92	5.91	6.40
266	4.92	5.91	6.40
267	4.99	6.83	4.93
268	4.99	6.83	4.93
269	4.99	6.83	4.93
270	5.25	5.52	6.09
271	5.25	5.52	6.09
272	4.99	6.83	4.93
273	5.25	5.52	6.09
274	5.25	5.52	6.09
275	5.21	6.21	5.09
276	5.25	5.52	6.09
277	4.30	6.97	5.18
278	4.30	6.97	5.18
279	4.30	6.97	5.18
280	5.29	5.84	5.10
281	5.29	5.84	5.10
282	4.30	6.97	5.18
283	5.29	5.84	5.10
284	5.30	5.90	5.18
285	5.38	5.61	5.28
286	5.50	5.30	5.34
287	4.34	6.82	5.15

Fuente: Muestra de los datos multivariantes.

#### *4.9 Validación del modelo.*

La forma de validar el modelo para medir la satisfacción del cliente con solo tres dimensiones consiste en contrastar la matriz de correlaciones para catorce variables presentada en la tabla 4.3.3 contra una matriz de correlaciones reproducida a partir del modelo obtenido (Tabla 4.9.1).

El contraste seleccionado fue una matriz de residuales de correlaciones observadas y reproducidas (Tabla 4.9.2). Como se mencionó en el capítulo cuatro, al compararlas no se deben observar valores mayores al 5%.

Al examinar la matriz de residuos no se encuentran valores que incumplan con el criterio establecido del 5%. Si bien existen algunos valores cercanos al correlacionar las variables dos y tres con la doce, están son las únicas observadas.

En este momento se procede a declarar la validez del modelo propuesto con tres dimensiones.

**Tabla 4.9.1 Matriz de correlaciones reproducidas.**

	VAR02	VAR03	VAR04	VAR05	VAR06	VAR07	VAR08	VAR09	VAR10	VAR11	VAR12	VAR13	VAR14	VAR15
VAR02	0.936	0.929	0.614	0.611	0.931	0.945	0.940	0.624	0.608	0.639	0.895	0.674	0.577	0.626
VAR03	0.929	0.922	0.627	0.624	0.921	0.936	0.930	0.616	0.621	0.629	0.886	0.665	0.568	0.615
VAR04	0.614	0.627	0.985	0.981	0.596	0.598	0.589	0.654	0.966	0.585	0.612	0.649	0.592	0.582
VAR05	0.611	0.624	0.981	0.977	0.593	0.594	0.586	0.651	0.963	0.582	0.609	0.645	0.589	0.579
VAR06	0.931	0.921	0.596	0.593	0.946	0.950	0.951	0.704	0.595	0.728	0.918	0.753	0.668	0.717
VAR07	0.945	0.936	0.598	0.594	0.950	0.960	0.958	0.663	0.594	0.684	0.916	0.714	0.621	0.672
VAR08	0.940	0.930	0.589	0.586	0.951	0.958	0.958	0.688	0.588	0.712	0.920	0.739	0.651	0.701
VAR09	0.624	0.616	0.654	0.651	0.704	0.663	0.688	0.933	0.662	0.938	0.737	0.939	0.934	0.944
VAR10	0.608	0.621	0.966	0.963	0.595	0.594	0.588	0.662	0.950	0.596	0.611	0.657	0.603	0.593
VAR11	0.639	0.629	0.585	0.582	0.728	0.684	0.712	0.938	0.596	0.956	0.757	0.950	0.945	0.962
VAR12	0.895	0.886	0.612	0.609	0.918	0.916	0.920	0.737	0.611	0.757	0.898	0.780	0.705	0.748
VAR13	0.674	0.665	0.649	0.645	0.753	0.714	0.739	0.939	0.657	0.950	0.780	0.952	0.939	0.955
VAR14	0.577	0.568	0.592	0.589	0.668	0.621	0.651	0.934	0.603	0.945	0.705	0.939	0.944	0.954
VAR15	0.626	0.615	0.582	0.579	0.717	0.672	0.701	0.944	0.593	0.962	0.748	0.955	0.954	0.969

Fuente: Muestra de los datos multivariantes

**Tabla 4.9.2 Matriz de residuos de correlaciones observadas / reproducidas.**

	VAR02	VAR03	VAR04	VAR05	VAR06	VAR07	VAR08	VAR09	VAR10	VAR11	VAR12	VAR13	VAR14	VAR15
VAR02		0.049	-0.006	-0.006	-0.045	-0.003	-0.030	0.004	0.000	0.006	-0.049	-0.003	0.014	0.008
VAR03	0.049		-0.004	-0.003	-0.040	-0.014	-0.040	0.010	-0.008	0.002	-0.045	0.004	0.010	0.005
VAR04	-0.006	-0.004		0.002	0.004	-0.001	0.001	0.002	-0.018	-0.001	0.008	0.006	-0.006	-0.002
VAR05	-0.006	-0.003	0.002		0.003	-0.002	0.000	0.003	-0.026	-0.002	0.008	0.008	-0.007	-0.003
VAR06	-0.045	-0.040	0.004	0.003		-0.011	0.021	-0.003	0.003	-0.005	0.031	-0.003	-0.006	-0.006
VAR07	-0.003	-0.014	-0.001	-0.002	-0.011		0.007	-0.002	0.006	0.002	-0.023	-0.007	0.007	0.003
VAR08	-0.030	-0.040	0.001	0.000	0.021	0.007		-0.007	0.009	-0.001	0.004	-0.009	0.001	-0.001
VAR09	0.004	0.010	0.002	0.003	-0.003	-0.002	-0.007		-0.019	-0.044	0.003	0.006	0.011	-0.033
VAR10	0.000	-0.008	-0.018	-0.026	0.003	0.006	0.009	-0.019		0.011	-0.009	-0.019	0.011	0.010
VAR11	0.006	0.002	-0.001	-0.002	-0.005	0.002	-0.001	-0.044	0.011		-0.011	-0.005	-0.012	0.017
VAR12	-0.049	-0.045	0.008	0.008	0.031	-0.023	0.004	0.003	-0.009	-0.011		0.011	-0.021	-0.015
VAR13	-0.003	0.004	0.006	0.008	-0.003	-0.007	-0.009	0.006	-0.019	-0.005	0.011		-0.042	-0.001
VAR14	0.014	0.010	-0.006	-0.007	-0.006	0.007	0.001	0.011	0.011	-0.012	-0.021	-0.042		-0.011
VAR15	0.008	0.005	-0.002	-0.003	-0.006	0.003	-0.001	-0.033	0.010	0.017	-0.015	-0.001	-0.011	

Fuente: Muestra de los datos multivariantes

# Capítulo 5

## Conclusiones

---

Se tiene evidencia suficiente para confirmar que es válido presentar la satisfacción del cliente como el espacio vectorial  $n \times m$  y reducirlo a un espacio vectorial  $f \times m$ , que explica la mayor parte de la varianza registrada en la medición de la satisfacción del cliente, utilizando el Análisis Matricial de Datos para descubrir las verdaderas dimensiones del proceso.

El factor Facilitar la venta al menudeo consolida las variables: condiciones generales del envase, condiciones generales del empaque, frecuencia con la que es visitado en su establecimiento, tiempo para la descarga del producto, atención recibida por parte del empleado y medios actuales de comunicación.

El factor Flujo de información explica la varianza común entre cinco variables: comunicación con la empresas, orientación de la rotación del producto, información de los productos elaborados, rapidez de respuesta a las quejas o sugerencias y la respuesta en sí a la queja o sugerencia.

Tres variables son sustituidas al factor Cumplimiento del pedido: cumplimiento a la cantidad solicitada, cumplimiento a las marcas y presentaciones solicitadas y el servicio que le brinda la empresa.

Este modelo fue validado con la matriz de residuales contrastando los residuales al comparar la correlación entre las variables originales y las variables obtenidas.

***El modelo mide la satisfacción del cliente con mayor certidumbre.***

El análisis matricial de datos incrementa la certidumbre de la información obtenida del cliente, porque:

En el análisis exploratorio de las variables obtuvimos la certeza de trabajar con distribuciones normales en cada variable.

En el análisis exploratorio de los eventos se identificaron eventos con datos ausentes y atípicos, todos ellos recibieron el tratamiento adecuado y se evitó el sesgo que ocasionan.

Al comprobar que existe correlación entre las variables, se entiende la necesidad del análisis multivariable para interpretar la información obtenida con el instrumento de medición de la satisfacción del cliente.

También se hizo evidente la necesidad de retirar del análisis la variable *calidad del producto*, por no tener relación significativa con el resto de las dimensiones de la calidad del estudio.

La extracción de la varianza común en las variables originales permite la reducción en el número de variables bajo un criterio objetivo y con base en el comportamiento estadístico observado.

***El modelo mejorar la interpretación de sus resultados.***

El modelo nos permite seleccionar la cantidad de variables con las que se desea trabajar y establece el impacto que esta decisión tendrá en el grado de conocimiento de la satisfacción del cliente.

Las nuevas variables obtenidas son fácilmente interpretadas a partir de las variables originales que fueron consolidadas en cada una, porque esta reducción parte de la variación explicada por cada factor.

La satisfacción del cliente en el modelo de tres variables es derivable de las observaciones medidas en campo y esta conversión no ocasiona una pérdida significativa de información.

***El modelo permite utilizar menos recursos.***

La existencia de variables de la calidad no medidas con la encuesta de satisfacción, pero finalmente aceptadas, indica que la necesidad de diseñar el proceso de atención al cliente alrededor de ellos. La modificación de las actividades resultará en el mejor conocimiento del cliente.

La administración debe enfocarse en facilitar la venta al menudeo de los productos que compra, que el pedido le será entregado a tiempo y responderá a sus requerimientos particulares y que el flujo de información le permitirá incrementar el valor del producto comprado.

El modelo advierte que los ajustes o adecuaciones diferentes a los involucrados en las tres nuevas variables, tendrá efectos insignificantes o nulos en la satisfacción del cliente. En este sentido, es importante determinar el nivel óptimo de inversión, en cada una de las dimensiones, que maximice la rentabilidad total. Averiguando las preferencias particulares de los clientes en cada uno de las tres dimensiones, la alta dirección podrá ajustar su oferta hacia el segmento que consideren más rentable o decidir cuál segmento de mercado satisfarán mejor con su actual oferta.

En ambos casos la satisfacción del cliente específico será una forma de incrementar la productividad y por lo tanto la permanencia en el mercado.

La alta dirección está en la posibilidad de dirigir los recursos limitados de la organización hacia los aspectos que tengan mayor influencia en la evaluación de la calidad percibida, y por extensión, en la satisfacción del cliente.

## *Bibliografía*

---

BERTALANFFY, Ludwig von. **General System Theory. Foundations, Development, Applications** 1968. (Traducido como Teoría General de los Sistemas. Editorial Fondo de Cultura Económica, México 1976).

CASTELLANOS, Verónica, HIRATA Okamoto, Ricardo. **Trabajo en Equipo El caso mexicano Tomo I**, Panorama Editorial, México 2005.

COLUNGA, Carlos. **Modelos Administrativos**. Editorial Panorama, México 1995.

DA SILVA, Reinaldo. **Teorías de la Administración**. Editorial Thomson, México 2004.

DEMING, Edwards. **Out of the crisis**. Segunda Edición 1986 (Traducido como: "Calidad, Productividad y Competitividad". Ediciones Díaz de Santos, España 1989).

FEIGENBAUM, Armand. **Control Total de la Calidad**. Tercera Edición Revisada 1991. Editorial CECOSA, México 1994.

HICKS, Douglas. **El sistema de costos basado en las actividades**. Editorial Alfa omega, Colombia 1998.

HIRATA Okamoto, Ricardo. **Las Siete Herramientas Básicas para el Control de Calidad**. Edición Especial de Keisen Consultores, 2002

HIRATA Okamoto, Ricardo. **7 Nuevas Herramientas para el Control de Calidad**. Edición Especial de Keisen Consultores, 2003

IMAI, Masaaki. **Gemba Kaizen**. Editorial Mc. Graw Hill Interamericana. Colombia 1998.

ISO/TR 10017 **Guidance on statistical techniques for ISO 9001:2008**, Second edition 2005-05-15

JOHNSON, Dallas E. **Métodos multivariados aplicados al análisis de datos**. International Thomson Editores, México 2004.

JURAN, y Gryna. **Análisis y Planeación de la Calidad**. Tercera Edición 1993. Editorial Mc. Graw-Hill, México 1995.

KAPLAN, Robert y Norton. **Cuadro de Mando Integral**. Editorial Gestión 2000, España 2002.

KAPLAN, Robert y Norton. **Mapas Estratégicos**. Editorial Gestión 2000, España 2004.

KUME Hitoshi. **Statistical Methods for Quality Improvemest..** The Association for Overseas Technical Echolarship (AOTS), Japan 1985

NIST/SEMATECH **e-Handbook of Statistical Methods**, <http://www.itl.nist.gov/div898/handbook/>, date.

NMX-CC-9001-IMNC-2008 (ISO 9001:2008) **Sistema de gestión de la calidad. Requisitos**

NMX-CC-9004-IMNC-2009 (ISO 9004:2009) **Sistema de gestión de la calidad. Directrices para la mejora del desempeño.**

PÉDRET, Ramón, Sagnier y Camp. **Herramientas para segmentar mercados y posicionar productos.** Ediciones DEUSTO, España 2000.

PÉREZ, César. **Técnicas de Análisis de Datos.** Editorial Pearson Prentice Hall, España 2009.

PÉREZ, César. **Técnicas de Análisis Multivariante de Datos.** Editorial Pearson Prentice Hall, España 2004.

PÉREZ, César y Santín, Daniel. **Data Mining.** Alfaomega Grupo Editorial, México 2006.

STRANG, Gilbert. **Algebra Lineal.** Addison-Wesley Iberoamericana, S.A., Sexta Reimpresión, México 1990.