

# APLICACIÓN DEL ANÁLISIS DE COMPONENTES PRINCIPALES EN EL DESARROLLO DE PRODUCTOS

L. VILLARROEL, J. ALVAREZ, D. MALDONADO

Proyecto Centro de Estadística Aplicada – CESA

Convenio CIUF – UMSS

Facultad de Ciencias y Tecnología - Universidad Mayor de San Simón

Cochabamba, marzo 2003 - Bolivia

Este artículo ilustra la aplicación del análisis de componentes principales en un problema de desarrollo de productos. El posicionamiento de cinco formulaciones de mermelada de mora fue estructurado tomando cuatro variables sensoriales: sabor, olor, color, aroma y textura.

Palabras claves: Análisis en componentes principales, análisis sensorial.

The paper shows an application of principal components in sensory analysis. The relative position of five marmalades made of berries for sensory variables: taste, scent, color, texture was studied.

Keywords. Principal component analysis, sensory analysis.

## I. INTRODUCCIÓN

El análisis sensorial o evaluación sensorial es una disciplina científica que permite definir, medir, analizar e interpretar las características de un producto, utilizando para este propósito los órganos de los sentidos bajo la consideración de que no existe ningún instrumento que pueda reproducir o reemplazar la respuesta humana. [Claustrioux J.J.,2001]

Las características físicas y químicas de los alimentos causan estímulos sobre los órganos de los sentidos haciendo posible la percepción de las impresiones visuales, gustativas, olfativas, táctiles y auditivas que hacen que el individuo acepte o rechace un alimento. Esta aceptación o rechazo es susceptible de ser medida con la ayuda de diferentes tests sensoriales.

El análisis sensorial sirve de manera general para el desarrollo de un nuevo producto, estudiar la influencia de modificaciones en la formulación o del proceso de fabricación sobre el producto, determinar las condiciones óptimas de conservación y para situar el producto frente a la competencia. Los datos obtenidos de las pruebas son sometidos al análisis y descripción realizados bajo las consideraciones del método estadístico de componentes principales.

El estudio que se detalla a continuación, fue realizado como parte del Proyecto de factibilidad para la implementación de un planta procesadora de frutos de la mora (*Rubus urticaefolious*) con la finalidad de definir la formulación final de la mermelada que es el principal producto del proyecto. Este proyecto fue realizado para el Programa de Alimentos y Productos Naturales de la Universidad Mayor de San Simón de Cochabamba, Bolivia en cooperación con el Centro de Estadística Aplicada – CESA.

La utilización de herramientas estadísticas en el análisis de datos generados en experiencias de análisis sensorial a menudo se limita a descripciones básicas, tablas, gráficas y algunos parámetros. El uso de herramientas multivariadas es muy limitado en Bolivia.

## Objetivo

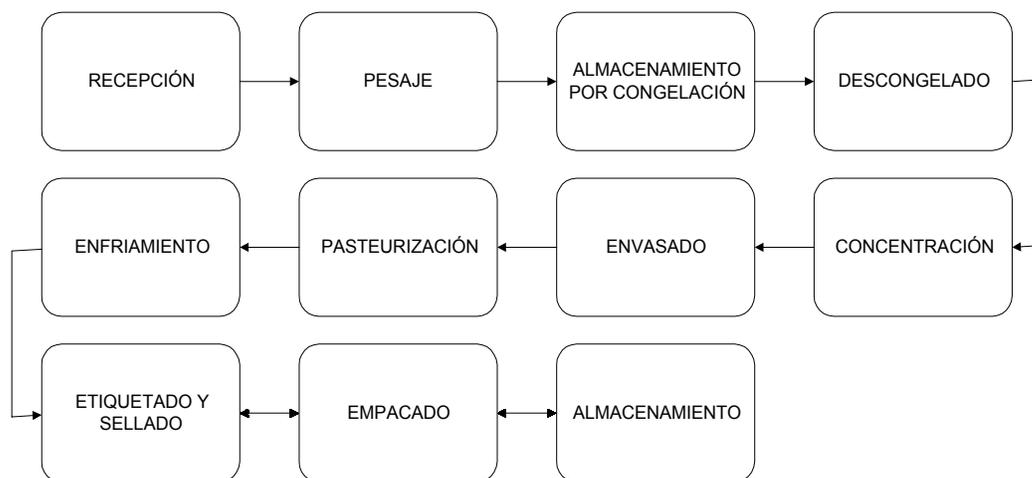
El objetivo general de esta publicación es presentar una aplicación del método de Análisis en Componentes Principales (ACP) en el campo del análisis sensorial. De manera específica se pretende establecer el posicionamiento de una formulación de mermelada de mora “Don Serafin”, de acuerdo a sus características sensoriales definidas como sabor, aroma, color y textura.

## II. METODOLOGIA

### 1. Descripción de los Materiales Experimentales

Para la elaboración de las mermeladas se utilizó el fruto de la mora (*Rubus urticaefolius*), procedente del trópico del departamento de Cochabamba; siendo ésta una especie propia de la biodiversidad de la región que crece en forma silvestre y en abundantes cantidades en la zona de Corani Pampa, Chulumani y Tablas Monte. Usada por los pobladores del lugar en forma fresca y preparado de refrescos para su consumo propio.

La fruta fresca tiene un 90% de parte comestible, en su composición física y química, además de que toda la planta es aprovechable para su utilización de diferentes maneras. La fruta fue procesada para la obtención de mermelada siguiendo los siguientes pasos :



Se elaboraron cuatro prototipos, diferenciados por el porcentaje de azúcar y fruta. La tabla 1 presenta la composición para cada una de las formulaciones.

**Tabla 1. Composición de los cuatro prototipos. [Álvarez, Maldonado, 2001]**

Características	Muestra #1	Muestra #2	Muestra #3	Muestra #4
Azúcar	50%	40%	45%	40%
Fruta	50%	60%	55%	60%
Agua	----	----	----	100%
Pectina	1%	1%	1%	1%
Benzoato de Sodio	----	----	----	0.1%
° BRIX	59,4°	53°	55,2°	47.2
PH	3,80	3,75	3,74	2.88

Las condiciones de elaboración de los prototipos fueron las mismas, tiempo de concentración de 30 minutos, envases esterilizados e instrumental de trabajo de iguales características.

## 2. Descripción de las Pruebas de Degustación

Las pruebas de análisis sensorial más utilizados son de cuatro tipos: pruebas discriminativas (pruebas triangular, duo-trio, prueba de conformidad categorización y prueba de clasificación), pruebas descriptivas (pruebas de notación de una variable sensorial sobre una escala, realización del perfil sensorial), pruebas hedónicas (prueba por pares, de clasificación, evaluación hedónica sobre una escala y prueba de consumo) y métodos combinados (método del perfil ideal y cartografía de preferencia).

Se aplicó el *test* de puntaje compuesto (correspondiente a pruebas hedónicas) que es un test de respuesta objetiva que permite realizar una evaluación comparativa de las muestras en estudio, se pueden presentar hasta cuatro variables sensoriales: sabor, aroma, color y textura. El puntaje se establece para cada variable o característica de acuerdo a su grado de importancia; de esta manera se asigna el valor de 30 puntos a las variables sabor y textura y de 20 puntos a las variables aroma y color. La muestra perfecta debería tener un puntaje de 100 puntos. Se debe notar que en este trabajo las pruebas de degustación fueron realizadas por un panel entrenado.

## 3. Análisis Estadístico

Además de la utilización de herramientas de carácter descriptivo, se utilizó la técnica del Análisis en Componentes Principales-ACP.

El análisis en componentes principales es una técnica descriptiva que permite estudiar las relaciones que existen entre las variables cuantitativas, sin considerar a priori, ninguna estructura, ni de variables, ni de individuos. [Palm,1998].

El ACP parte de una matriz **Y**, de dimensiones  $n \times p$ , donde  $p$  corresponde al número de variables observadas en  $n$  individuos. La matriz **Y**, debe ser estandarizada de manera a obtener promedios nulos y varianzas unitarias por variable. La matriz **X** es el resultado de la siguiente transformación.

$$x_{ij} = \frac{y_{ij} - \bar{y}_j}{s_j \sqrt{n}}$$

donde  $\bar{y}_j$   $s_j$  son el promedio y la desviación estándar para cada una de las variables.

Luego se determinan los valores propios y vectores propios de la matriz  $X'X$ . Los valores propios corresponden a la varianza de los valores en cada uno de los componentes y los elementos de los vectores propios corresponden a las coordenadas en el espacio inicial que dan la dirección de los componentes principales.

La ecuación matricial del ACP es:

$$Z = X U$$

Donde  $Z$  es la matriz de componentes principales,  $X$ , la matriz de valores iniciales estandarizados y  $U$  la matriz de vectores propios de la matriz  $X'X$  o matriz de correlación  $R$ .

De manera general el resultado del ACP se resume a dos representaciones gráficas. El diagrama de variables o círculo de correlación, en el que se representan precisamente las variables en el plano factorial retenido. Las coordenadas corresponden a los coeficientes de correlación entre las variables originales y cada uno de los componentes. El diagrama de individuos, en el que se representan los individuos en el plano factorial.

El análisis de los datos experimentales fue realizado en el paquete estadístico SPSS [Vizauta, 1998]

## V. RESULTADOS Y DISCUSION

Cada una de las diferentes formulaciones fueron evaluadas por el panel entrenado (siete jueces). Los resultados promedio por mermelada para cada una de las cuatro variables sensoriales se presentan en la Tabla 2. Se debe notar que las puntuaciones de las variables sabor y textura están medidas en una escala de 30 puntos y las variables aroma y color en una escala de 20 puntos.

**Tabla 2. Promedios de las variables sensoriales por tipo de formulación.**

	AROMA	COLOR	SABOR	TEXTURA
M1	10	15	18	11
Don serafin	16	17	24	16
M3	13	13	20	14
M4	12	20	10	25
M5	12	11	14	11

La tabla 3 presenta los promedios generales para las cuatro variables consideradas.

**Tabla 3. Promedios y desviaciones estándar de las cuatro variables sensoriales**

	Promedio	Desviación estándar	CV %
<b>AROMA</b>	12.6	2.19	17
<b>COLOR</b>	15.2	3.49	23
<b>SABOR</b>	17.2	5.40	31
<b>TEXTURA</b>	15.4	5.77	37

Los coeficientes de variación relativamente altos para las variables textura y sabor comparativamente a los correspondientes del aroma y color evidencian que estas dos variables tienen mayor capacidad discriminativa.

Con el propósito de evaluar el grado de asociación entre las cuatro variables sensoriales se ha determinado la matriz de correlación de Pearson, presentada en la tabla 3.

**Tabla 4. Matriz de correlación entre las cuatro variables sensoriales**

		AROMA	COLOR	SABOR	TEXTURA
Correlación	AROMA	1,000	,176	,600	,194
	COLOR	,176	1,000	-,201	,863
	SABOR	,600	-,201	1,000	-,468
	TEXTURA	,194	,863	-,468	1,000

La Tabla 4 pone en evidencia una alta correlación positiva entre textura y color. Este resultado implica que las mermeladas con mejor textura normalmente son mermeladas que tienen asociadas un buen color. Así mismo se observa una correlación importante y además positiva entre el sabor y el aroma. Se debe notar también que el coeficiente entre textura y sabor es negativo, este resultado implica que una mermelada calificada con buen sabor corresponde normalmente a una mermelada con mala textura y viceversa.

La Tabla 5 presenta los resultados del análisis en componentes principales en términos de la varianza inicial total explicada por cada uno de los componentes.

**Tabla 5. Varianza total inicial explicada por los componentes**

Componentes	Autovalores Iniciales			Suma de las saturaciones de la extracción al cuadrado		
	Total	% de Varianza	% Acumulado	Total	% de Varianza	% Acumulado
1	2,075	51,884	51,884	2,075	51,884	51,884
2	1,565	39,137	91,021	1,565	39,137	91,021
3	,328	8,208	99,229			
4	3,082E-02	,771	100,000			

La varianza inicial total es igual a 4 provocado por la estandarización de variables. De los cuatro puntos de varianza original, el primer componente “se lleva” 2.075 que representa prácticamente el 52 % de la varianza original. El segundo componente explica 1.565 (39 %) de la varianza original. Si observamos en la columna acumulado, el primer plano factorial (componente 1 \* componente 2) explica el 91 % de la varianza original.

Si se toma en cuenta como criterio de retención de componentes, aquellos cuyo valor propio (*Eigen value*) es superior a la unidad, el ACP considera a los dos primeros componentes. En términos prácticos, el problema inicialmente en el hiper espacio de cuatro dimensiones ha sido reducido al plano, rescatando más del 90 % de la varianza original.

El paso siguiente en el ACP, es caracterizar los componentes en términos de las variables originales. Este procedimiento se realiza estableciendo la matriz de correlación entre las variables originales y los componentes retenidos. La tabla 6 presenta estos coeficientes.

**Tabla 6. Matriz de correlación entre variables originales y componentes retenidos**

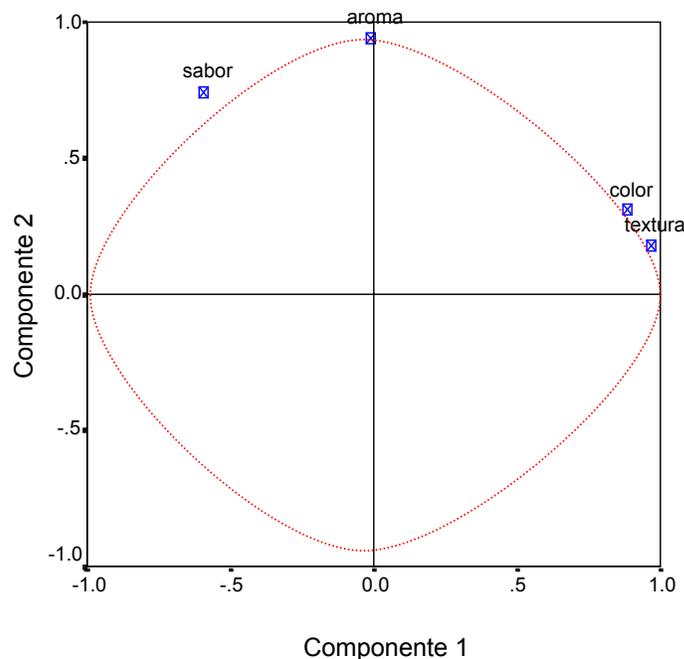
	Componentes	
	1	2
AROMA	-0.011	0.943
COLOR	0.886	0.309
SABOR	-0.594	0.740
TEXTURA	0.968	0.182

Para “nombrar” los dos componentes en términos de las variables originales, se observa en la Tabla 6 los coeficientes de correlación más importantes sean estos positivos o negativos (próximos a 1 o -1).

En la primera columna se observa dos coeficiente de correlaciones positivos importantes, igual 0.968 y 0.886 entre el primer componente y la variable textura y color. Por lo tanto el primer componente principal es un eje que mide la textura y el color.

En la segunda columna se observa también dos coeficientes importantes entre el segundo componente y la variable aroma (0.943) y la variable sabor (0.74), por tanto el componente 2 es un eje que mide el aroma y el sabor.

Estos resultados son representados gráficamente en el diagrama de variables o círculo de correlación. La proximidad de una variable a uno de los componentes implica precisamente una correlación importante con el componente. La figura 1 corresponde al diagrama de variables generado por el ACP. Los resultados de la Tabla 6 son ilustrados en el grafico de variables en el primer plano factorial de la figura 1.



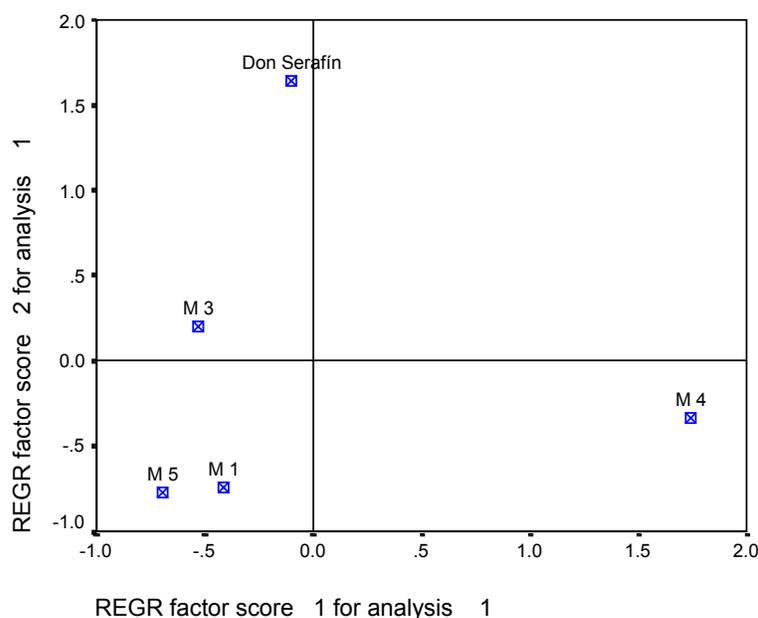
### Figura 1. Diagrama de Variables.

La calidad de la representación de las variables en el plano factorial es evaluado por la comunalidad. Los valores de extracción próximos a la unidad implican que la variable está bien representada en el plano factorial. Si los valores de extracción son bajos implica que la variable tiene mayor correlación con otro eje no considerado, y por lo tanto no es correcto caracterizar los componentes con esta variable. La tabla 7 muestra la comunalidad para cada una de las cuatro variables. Se puede observar que de manera general las cuatro variables están bien representadas en el primer plano factorial, en efecto todos las comunalidades son próximas a 1.

Tabla 7. Comunalidades

	Inicial	Extracción
<b>AROMA</b>	1,000	,890
<b>COLOR</b>	1,000	,881
<b>SABOR</b>	1,000	,900
<b>TEXTURA</b>	1,000	,970

La figura 2 presenta el diagrama de individuos (mermeladas) en el primer plano factorial. En base a los resultados de la caracterización de los componentes realizado en el gráfico de individuos se procede a la interpretación de los perfiles sensoriales de cada una de las mermeladas.



### Figura 2. Diagrama de Individuos.

Por la posición de la mermelada “Don Serafín”, en el primer plano factorial, esta mermelada se caracteriza por un buen aroma y sabor. La mermelada 4 (M4) es una formulación con buena textura y buen color. Las mermeladas M5 y M1 son dos formulaciones con un perfil sensorial similar y caracterizados por una mala textura, un mal color, un mal sabor y aroma.

## VI. CONCLUSIONES

Los resultados alcanzados han permitido ilustrar, con una base de datos relativamente reducida, la aplicación de una técnica multivariable, el ACP en el tratamiento de datos procedentes de experimentos sensoriales. Por tanto el ACP, es una herramienta optima para la identificación o selección de prototipos de mermeladas en base a sus características sensoriales. Las posibilidades de uso de otras técnicas multivariables en el campo del análisis sensorial son extensas. Es importante notar que la selección de una herramienta de análisis está en función de los resultados esperados, el número de variables que participan y la naturaleza de los datos.

## VII. BIBLIOGRAFÍA.

ALVAREZ J., MALDONADO D. [2001]. Proyecto de Factibilidad para la implementación de una planta procesadora del fruto de la mora (*Rubus urticaefolious*). Carrera de Ingeniería Industrial. UMSS.

CLAUSTRIAUX J.J.[2001]. *Considérations sur l'analyse statistique de donées sensorielles*. Bitech nol. Agron. Soc. Environ. 2001 5 (3), pp.155-158.

PALM R. [1998]. L'analyse en composantes principales: principe et aplicacion. Notes de statistique et d'informatique. Gembloux, Belgique.

VIZAUTA B. [1998]. Análisis estadístico con SPSS para Windows. Estadística multivariable. McGRAW – HILL, Madrid España.