



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE QUÍMICA

**Procedimiento estadístico para la selección de jueces
analítico-sensoriales**

T E S I S
QUE PARA OBTENER EL TÍTULO DE
QUÍMICA DE ALIMENTOS
P R E S E N T A:
C LAUDIA DELGADO GUTIÉRREZ



MÉXICO, D.F.



2002

**EXAMENES PROFESIONALES
FACULTAD DE QUÍMICA**

**TESIS CON
FALLA DE ORIGEN**



Universidad Nacional
Autónoma de México



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: Prof. Carlos López Sánchez
Vocal: Prof. Dulce María Gómez Andrade
Secretario: Prof. Daniel Luis Pedrero Fuehrer
1er. Suplente: Prof. Guillermo Antonio Segura Espinosa
2o. Suplente: Prof. René Julio De Los Ríos Campanella

Trabajo que se desarrolló en:

**Penta Sensorial, S.A. de C.V. Santa Catalina 313, Colonia del Valle.
CP 03100 México DF.**

Asesor



M. en C. Daniel Luis Pedrero Fuehrer

Sustentante:



Claudia Delgado Gutiérrez

A mis padres Francisco y Teresa:

Porque no pude haber tenido mejor ejemplo. Gracias por todo su amor y comprensión.

A Oscar y Paco :

Por que gracias a su compañía y superación que admiro mucho, he podido ser mejor cada día.

A Elihu:

Por que Dios no se equivoca y mando a la mejor persona con quien puedo compartir la vida.

A la Universidad Nacional Autónoma de México:

Por haberme enseñado que la mejor herencia que uno puede tener es una buena educación.

A la Evaluación Sensorial :

Por lo maravillosa y funcional que es esta ciencia.

A Alma, Blanca, Sandra y Silvia:

Por todo el tiempo que pasamos juntas en la Universidad.

TABLA DE CONTENIDO

1 RESUMEN	1
2 INTRODUCCIÓN	2
3 OBJETIVOS	3
3.1 Objetivo General:.....	3
3.2 Objetivos Particulares	3
4 ANTECEDENTES	4
4.1 Importancia de los jueces en la evaluación sensorial	4
4.1.1 Proceso de selección y entrenamiento de jueces sensoriales.....	5
4.2 Requerimientos para establecer criterios estadísticos en la selección de candidatos a jueces sensoriales.....	7
4.2.1 Tipos de estadística	8
4.2.1.1 Estadística Descriptiva	8
4.2.1.2 Estadística Inferencial	8
4.2.2 Tipos de pruebas sensoriales.....	9
4.2.3 Tipos de números	9
4.2.4 Logística para emplear métodos estadísticos en la resolución de problemas.....	11
4.2.5 Relación de la estadística con la evaluación sensorial	12
4.3 Información de tipo nominal, métodos para la obtención de criterios estadísticos para seleccionar candidatos a jueces en pruebas discriminativas.....	14
4.3.1 Pruebas sensoriales que involucren información de tipo nominal.....	14
4.3.1.1 Identificación de gustos básicos:.....	14
4.3.1.2 Pruebas discriminativas	14
4.3.2 Pruebas estadísticas no paramétricas consideradas en la obtención de criterios para la selección de candidatos a jueces sensoriales al utilizar información de tipo nominal.....	15
4.3.2.1 Prueba binomial	15
4.3.2.2 Ji cuadrada (χ^2).....	16
4.3.3 Definición de los criterios con fundamentos estadísticos para seleccionar candidatos a jueces en pruebas discriminativas.	16
4.4 Información de tipo ordinal, métodos para la obtención de criterios estadísticos para seleccionar candidatos a jueces en pruebas de ordenación.....	21
4.4.1 Pruebas sensoriales de ordenación	21

4.4.2	Pruebas estadísticas no paramétricas para la selección de candidatos a jueces sensoriales al utilizar información de tipo ordinal	22
4.4.2.1	Page Test.....	22
4.4.2.2	Análisis de ordenamiento por rangos de Newell-MacFarlane	22
4.4.2.3	Prueba de Friedman.....	23
4.4.3	Definición de criterios con fundamentos estadísticos para seleccionar candidatos a jueces sensoriales en pruebas de ordenación.....	23
4.4.3.1	Aplicación del método Page Test.....	24
4.4.3.2	Aplicación del análisis de ordenamiento por rangos de Newell-MacFarlane	25
4.4.3.3	Aplicación de la prueba de Friedman.....	27
4.5	Información de tipo intervalar, métodos para la obtención de criterios estadísticos para seleccionar candidatos a jueces en pruebas de intensidad.	33
4.5.1	Pruebas sensoriales de intensidad	33
4.5.2	Pruebas estadísticas paramétricas consideradas en la obtención de criterios para la selección de candidatos a jueces sensoriales al utilizar información de tipo intervalar	33
4.5.2.1	Media.....	33
4.5.2.2	Desviación estándar muestral	34
4.5.2.3	Coefficiente de variación.	34
4.5.2.4	t de student.....	34
4.5.2.5	Análisis de varianza (ANOVA)	35
4.5.3	Definición de criterios con fundamentos estadísticos para seleccionar candidatos a jueces sensoriales en pruebas sensoriales de intensidad.	36
4.5.3.1	Comparación de la variabilidad entre los resultados de un candidato a juez con respecto al grupo	39
5	METODOLOGÍA	43
5.1	Aplicación de criterios para seleccionar candidatos a jueces sensoriales en pruebas discriminativas (información nominal).....	43
5.2	Aplicación de criterios para seleccionar candidatos a jueces sensoriales en pruebas de ordenación (información ordinal)	45
5.3	Aplicación de criterios para seleccionar candidatos a jueces sensoriales en pruebas de intensidad (información intervalar).....	47
5.4	Clasificación final de los candidatos a jueces sensoriales.	51
6	RESULTADOS	52
6.1	Selección de candidatos a jueces sensoriales aplicando los criterios con fundamentos estadísticos para datos de tipo nominal en pruebas sensoriales discriminativas.....	52
6.1.1	Clasificación de candidatos a jueces sensoriales en pruebas de identificación de gustos básicos.....	52

6.1.2	Clasificación de candidatos a jueces sensoriales en pruebas de comparación por pares	55
6.2	Selección de candidatos a jueces sensoriales aplicando los criterios con fundamentos estadísticos para datos de tipo ordinal en pruebas sensoriales de ordenación	58
6.2.1	Clasificación general de candidatos a jueces sensoriales en pruebas de sensoriales de ordenación	58
6.3	Selección de candidatos a jueces sensoriales aplicando los criterios con fundamentos estadísticos para datos de tipo intervalar en pruebas sensoriales de intensidad	64
6.3.1	Clasificación de candidatos a jueces sensoriales en pruebas sensoriales de Intensidad para la familia de atributos <i>visuales</i>	64
6.3.2	Clasificación de candidatos a jueces sensoriales en pruebas sensoriales de Intensidad para la familia de atributos aroma general.....	67
6.3.3	Clasificación de candidatos a jueces sensoriales en pruebas sensoriales de Intensidad para la familia de atributos aromáticos.....	70
6.3.4	Clasificación de candidatos a jueces sensoriales en pruebas sensoriales de Intensidad para la familia de gustos básicos.....	73
6.3.5	Clasificación de candidatos a jueces sensoriales en pruebas sensoriales de Intensidad para la familia de atributos textura bucal	75
6.4	Clasificación final de candidatos a jueces sensoriales.....	78
7	ANÁLISIS DE RESULTADOS.....	79
8	CONCLUSIONES.....	81
9	BIBLIOGRAFÍA.....	82
10	APÉNDICES.....	84
10.1	Resumen de los criterios obtenidos	84
10.2	Ejemplo de cálculos para determinar J_i cuadrada y porcentajes de reconocimiento de gustos básicos.....	88
10.3	Ejemplo de cálculos para la determinación del valor teórico de la prueba de Friedman y el porcentaje de muestras sin repetir valores de suma de rangos....	90
10.4	Ejemplo de cálculos utilizados en la clasificación de jueces en pruebas sensoriales de intensidad.....	92

1 RESUMEN

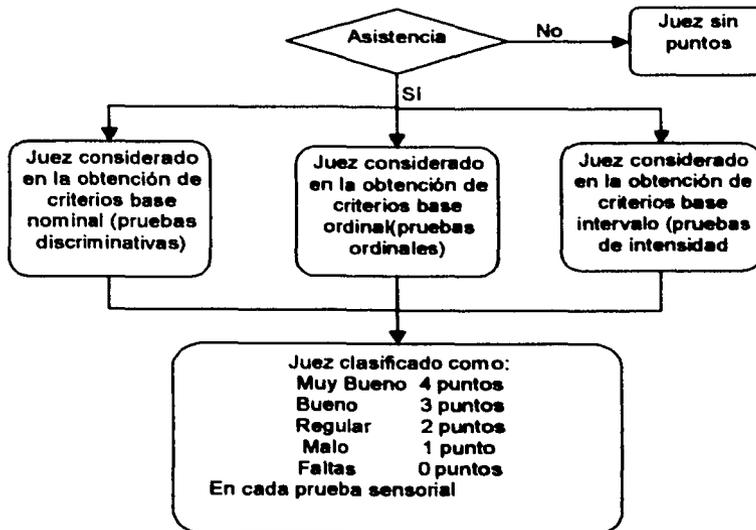
La presente investigación esta enfocada en proponer una serie criterios con fundamentos estadísticos para la selección de candidatos a jueces sensoriales. Para el desarrollo de esta investigación se partió de bases de datos donde se evaluó yogurt de fresa, proporcionadas por Penta Sensorial, S.A. de C. V., posteriormente se clasificó la información según el tipo de número: nominal, ordinal e intervalar.

A cada base de datos se aplicaron las pruebas estadísticas de acuerdo al tipo de número y de ensayo sensorial, y además con la variación de que se consideró a cada juez como una muestra. Esto se realizó con el fin de observar el comportamiento del juez bajo dos aspectos: (1) su variabilidad entre las repeticiones de cada juez y (2) la variabilidad de cada juez respecto al grupo.

El objetivo en la tesis al aplicar el análisis estadístico a una base de datos, es proponer un procedimiento numérico y de criterios de decisión para clasificar a un grupo de personas según su desempeño en rutinas de la selección de jueces. Como es del conocimiento general, la selección del juez se ejecuta de acuerdo a la habilidad o desempeño para ejecutar pruebas sensoriales, así como por la reproducibilidad en dar un fallo.

Los jueces se seleccionan principalmente por: su asistencia a los ensayos, su desempeño por prueba sensorial y su constancia perceptiva en sus mediciones de acuerdo al siguiente diagrama:

Figura 1-1 *Importancia de la asistencia como punto de partida para la clasificación de jueces*



2 INTRODUCCIÓN

La evaluación sensorial es una herramienta de gran utilidad para cualquier empresa que desee analizar, desde el punto de vista de la percepción humana, sus productos. Para poder cumplir con sus funciones, esta ciencia necesita utilizar jueces entrenados y capacitados, de manera que fundamente sus resultados y haga de cada juicio una medición confiable y reproducible. Esa es la razón principal por la cual el analista sensorial se ha apoyado en la estadística, que es la parte medular de esta investigación.

Entre más importancia se le dé al grupo de jueces sensoriales dentro de una empresa, mayores serán los beneficios que ofrece el tener un grupo que sea validado estadísticamente. El mayor de estos beneficios es la confiabilidad y reproducibilidad de sus determinaciones o mediciones sensoriales.

Lo primero que se propone cuando en una empresa o centro de investigación se busca implementar un programa de evaluación sensorial, es poner énfasis en la selección de candidatos a jueces sensoriales en base a criterios estadísticos que permitan reconocer la habilidad innata del candidato para ejecutar ciertas pruebas sensoriales, así como su reproducibilidad durante esta primera etapa. De esta manera la estadística apoya al analista a realizar la selección de candidatos a jueces sensoriales, proporcionando elementos para establecer los criterios de decisión que faciliten el aceptar o rechazar a un juez.

Si el candidato a juez tiene habilidades innatas para ejecutar algunas pruebas sensoriales, le será más fácil recibir o participar en un entrenamiento, lo que generará ahorros en tiempo y recursos en la empresa. El caso contrario es para aquel juez que no tenga habilidades innatas, y que por lo mismo le costará mucho trabajo y se perdería mucho tiempo y recursos entrenando a dicho juez. Esto no quiere decir que ese candidato no pueda lograr eventualmente ser un buen juez, solo que probablemente la empresa tendría que invertir mucho más en este candidato, que en uno que si tiene habilidad innata.

Los criterios estadísticos para seleccionar candidatos a jueces sensoriales planteados en esta investigación no son rígidos, ya que se establecen categorías que el lector debe adaptar a sus problemas y necesidades específicas de entrenamiento de jueces sensoriales. En este trabajo se proponen categorías para clasificar a los candidatos en "Muy Bueno", "Bueno", "Regular" y "Malo". Con esta tesis se ofrece una herramienta más que apoya la selección de candidatos a jueces sensoriales, adicionalmente valdría la pena mencionar que en etapas posteriores a la selección, como es el entrenamiento y el mantenimiento del grupo de jueces, es posible seguir la misma metodología planteada en esta investigación.

3 OBJETIVOS

3.1 OBJETIVO GENERAL:

- Determinar criterios con fundamentos estadísticos para seleccionar candidatos a jueces sensoriales.

3.2 OBJETIVOS PARTICULARES

- Clasificar la información de una base de datos por tipo de número estadístico.
- Seleccionar las pruebas estadísticas que permitan analizar la información contenida en las bases de datos de acuerdo a la prueba sensorial realizada.
- Crear criterios para seleccionar candidatos a jueces sensoriales, formando categorías de clasificación de su desempeño en "Muy Bueno", "Bueno", "Regular" y "Malo".

4 ANTECEDENTES

4.1 IMPORTANCIA DE LOS JUECES EN LA EVALUACIÓN SENSORIAL

La evaluación sensorial es una disciplina científica que permite evocar, medir, analizar e interpretar aquellas características de los alimentos tal y como se perciben por los sentidos: vista, olfato, gusto, tacto y oído (IFT, 1975). De acuerdo a esta definición se necesita un instrumento de medición que permita a la evaluación sensorial cumplir con sus funciones. Tal instrumento es conocido como juez analítico sensorial.

Los jueces son aquellas personas que han sido entrenadas en una o más metodologías específicas. El juez es un individuo que está dispuesto a participar voluntariamente en una prueba sensorial para evaluar un producto, llámese alimento u otro material (materias primas, cosméticos, productos para el hogar, papel, telas etc.) valiéndose de la capacidad perceptiva de uno o varios de sus sentidos.¹⁷

En la actualidad existen lineamientos internacionales, como las Normas ISO y los criterios de la ASTM (American Society for Testing and Materials), para determinar que características deben tener los candidatos a jueces sensoriales. Por ejemplo, de acuerdo con la norma ISO 8586-1:1993⁹ se indica que:

- Un juez no debe presentar actitudes en contra del tipo de alimento a evaluar, ya sean culturales, étnicas o de alguna otra índole. Como se evalúan diferentes tipos de alimentos, si el juez presenta una de estas actitudes no va a evaluar correcta o imparcialmente y siempre tendrá algún prejuicio o tendencia a no calificar objetivamente dicho producto.
- Requiere de conocimientos y aptitudes específicas como capacidad para concentrarse, física e intelectualmente en el producto que va a evaluar. Esto es importante porque si el juez no se concentra y esta pensando en otra cosa, no evaluará bien y sus resultados tendrán mucha variación entre sí. Además si no tiene el conocimiento acerca del producto y la metodología sensorial, no podrá expresar acertadamente sus apreciaciones.
- El candidato a juez goce de buena salud, cuando los jueces se enferman ya sea de gripa o alguna otra causa no pueden evaluar. Sus sentidos no están funcionando en forma óptima. Es como cuando a un instrumento le falta algún componente, éste no puede funcionar o medir adecuadamente.
- El juez debe ser comunicativo, sobre todo para preguntar sus dudas y expresar aspectos específicos de su percepción, es decir tener habilidad para comunicarse.
- Se debe contar con la disponibilidad de tiempo, puesto que generalmente para evaluar un producto sensorialmente se requieren varias sesiones y con esto se necesita tiempo y disponibilidad del juez; de lo contrario se afectarían los resultados, pues si falta no se podría contar con su evaluación para ese día.

- Su personalidad no debe de afectar el estudio, por ejemplo si a un juez le gusta que los demás actúen al igual que él, debe controlar este aspecto de su personalidad porque de esta manera tendría influencia sobre los demás jueces generando malas evaluaciones.

Un juez realiza varias funciones dentro de la evaluación sensorial como: clasificar entre diferentes estímulos, reconoce los cuatro gustos básicos, ordena de manera que establece jerarquías, califica en intervalos o categorías y puede discriminar entre varias muestras diferentes.¹

4.1.1 Proceso de selección y entrenamiento de jueces sensoriales

El proceso de selección y entrenamiento de jueces sensoriales involucra una serie de etapas como el reclutamiento, selección, entrenamiento y mantenimiento. Esta investigación se centra en la selección, ya que se enfoca a determinar criterios para esta etapa. Los pasos para seleccionar un grupo de jueces se explican en Tabla 4-1:

Tabla 4-1 Pasos para la selección de un grupo de jueces de evaluación sensorial¹

Pasos	Tiempo invertido (semanas)	Número de candidatos ideales		Criterios
		Iniciales	Después de aplicar criterios	
1. Reclutamiento y selección de candidatos	2	200	102	Disponibilidad
2. Pruebas preliminares	2	102	50	Disponibilidad Capacidad
3. Medidas umbrales (esto es necesario en muy especiales ocasiones)	3	50	33	Disponibilidad Capacidad Sensitividad
4. Análisis de datos. Selección de candidatos aplicando criterios	1	33	28	Disponibilidad Capacidad Sensitividad Confiabilidad

¹ Adaptado de Gatchalian, M.M & Yano, T (1990) Quantified approach to sensory panel selection. Food Quality and Preference 2 (233-241)

1. Reclutamiento: consiste en convocar a un número de personas interesadas en su formación como jueces sensoriales. Puede ser de tipo interno, con gente de la empresa, o externo, con personas ajenas a la misma. La selección del tipo de reclutamiento depende del financiamiento para el programa, así como de los requerimientos de la organización. En esta etapa se aplica un breve cuestionario para conocer al candidato y detectar si presenta hábitos y cualidades que dificulten o interfieran con su desempeño como juez sensorial.¹

2. Selección: Una vez que ya se tiene reclutada a la gente, se aplican pruebas filtro para clasificar a los candidatos de acuerdo a los siguientes criterios:

- Agudeza sensorial normal.
- Interés en la evaluación sensorial.
- Habilidad para discriminar y reproducir resultados.
- Presentar una conducta de cooperación, motivación y puntualidad.

Durante ésta etapa se realizan las siguientes pruebas preliminares¹⁶:

- Familiarización con gustos básicos y reconocimiento a diferentes concentraciones.
- Identificación y discriminación de gustos básicos, aumentando la dificultad, y mezclando varios gustos básicos.
- Memoria olfativa, generación de descriptores aromáticos.
- Aplicación de pruebas discriminativas, como comparación por pares, dúo-trío o triangular.
- Análisis de textura.
- Aplicación de pruebas de ordenación y de intervalos, manejo de escalas.

3. Entrenamiento: del mismo modo que un instrumento analítico es calibrado con el fin de obtener medidas confiables, el equipo de jueces sensoriales debe ser entrenado, ayudado frecuentemente por referencias o estándares. Esto se realiza con el fin de que la respuesta obtenida sea lo más precisa y exacta posible.² Se realiza un entrenamiento para proveer de conocimiento al juez en las técnicas y procedimientos utilizados en el análisis sensorial, así como para descubrir su habilidad innata para detectar, reconocer y describir estímulos sensoriales. El entrenamiento se realiza de acuerdo a los objetivos de investigación sensorial que tiene la empresa o institución académica.

4. Mantenimiento del grupo de jueces: los jueces sensoriales, como todo instrumento de medición, necesitan ser calibrados para verificar que cumplan con la función para la cual fueron entrenados. En este sentido es importante realizar un monitoreo constante en donde el analista sensorial determina si es necesario repetir alguna técnica de entrenamiento o reforzar aquella en donde los resultados de los jueces hayan resultado variables; esto se conoce como mantenimiento de jueces sensoriales.⁵

Debido a que el instrumento de análisis para la evaluación sensorial son los jueces y estos, como todo método analítico para ser considerado preciso, necesitan manifestar una desviación estándar relativa (también conocida como coeficiente de variación) igual o menor a 5%¹². se deben realizar repetidas evaluaciones sobre un mismo producto para determinar si el grupo de jueces esta evaluando dentro de los mismos promedios, para considerarse preciso⁵.

5. Selección de jueces: De acuerdo a la ASTM (American Society for Testing and Materials) se selecciona a un juez si cumple con el siguiente rendimiento durante la etapa de selección¹:

- a) En pruebas de identificación de gustos básicos tener un 100% de aciertos.
- b) En pruebas de reconocimiento de aromas, donde se dan 5 puntos para una "identificación correcta" 4 puntos por una "asociación correcta", 3 puntos para una "caracterización" y 1 punto para un "intento de descripción", se debe obtener un 70% de aciertos.
- c) En ordenamiento por intensidad de un atributo se hacen comparaciones del desempeño de un sujeto con respecto al grupo, determinando la desviación estándar del promedio de sus evaluaciones, y analizando la variabilidad entre si y dentro de los sujetos.

Por otra parte Gisela Jellinek⁸ sugiere que se apliquen los siguientes valores de rendimiento:

1 Reconocimiento de gustos básicos	100%
2 Pruebas de umbral	80%
3 Identificación de olores	80%
4 Identificación de aromas	70%
5 Pruebas de comparación por pares	70%
6 Prueba triangular (mínimo 20 trios)	70%
7 Prueba de ordenación de gustos básicos	60%
8 Prueba de ordenación de colores	90%
9 Prueba de ordenación por tacto (Tilgner)	80%

Para los fines de esta investigación, se utilizaron los rendimientos sugeridos por la ASTM.

4.2 REQUERIMIENTOS PARA ESTABLECER CRITERIOS ESTADÍSTICOS EN LA SELECCIÓN DE CANDIDATOS A JUECES SENSORIALES

¿Por qué utilizar la estadística para obtener criterios de selección de candidatos a jueces sensoriales?, si se pone atención en la definición de evaluación sensorial ya mencionada en la parte 4.1 de esta sección, se puede deducir que se necesita de una herramienta como la estadística para que valide los datos sensoriales, y se genere información confiable y reproducible.¹⁹

Cuando un analista sensorial se enfrenta a un problema, necesita desarrollar una serie de pasos que le permitan dar solución a su problema de la mejor manera. Para ello diseña la prueba que más se adecua a sus objetivos, selecciona el número de jueces que va a requerir y si estos necesitan entrenamiento. Entre otras cosas también decide cantidad y tipo de muestra, condiciones necesarias para poder realizar la prueba, etc. (ver Logística para el Desarrollo de pruebas Sensoriales en Pedrero y Pangborn, 1989). Una vez que ya se tienen planteados todos los elementos involucrados en el desarrollo de la prueba sensorial, el analista sensorial requerirá analizar e interpretar los resultados, por lo que la herramienta que le proporcionará ayuda es la estadística.³

4.2.1 Tipos de estadística

4.2.1.1 Estadística Descriptiva

La estadística se divide en dos tipos: descriptiva e inferencial. Con respecto a la estadística descriptiva ésta se encarga de utilizar los números con el propósito de resumir información conocida con respecto a una situación de interés¹⁴, además evalúa, estudia y describe a la totalidad de individuos de una población. Su finalidad es obtener información, analizarla y simplificarla a un nivel en el que pueda ser interpretada cómoda y rápidamente y, por tanto, pueda utilizarse eficazmente para el fin que se desee. La estadística descriptiva trabaja con todos los individuos de la población. El proceso que sigue este tipo de estadística para el estudio de una cierta población consta de los siguientes pasos:

- Selección de caracteres a estudiar.
- Mediante encuesta o medición, obtención del valor de cada individuo en los caracteres seleccionados.
- Elaboración de tablas de frecuencias, mediante la adecuada clasificación de los individuos dentro de cada caracter.
- Representación gráfica de los resultados (elaboración de gráficas estadísticas).
- Obtención de parámetros estadísticos, números que sintetizan los aspectos más relevantes de una distribución estadística.

En evaluación sensorial se puede utilizar la estadística descriptiva para determinar promedios, desviaciones estándar de un grupo de jueces sensoriales, permite visualizar la información obtenida en tablas, gráficas, cuadros o listas.

4.2.1.2 Estadística Inferencial

La estadística inferencial trabaja con muestras o subconjuntos formados por algunos individuos de la población. A partir del estudio de la muestra se pretende inferir aspectos relevantes de toda la población: ¿cómo se selecciona la muestra?, ¿cómo se realiza la inferencia?, y ¿qué grado de confianza se puede tener en ella?, estos son aspectos fundamentales de la estadística inferencial; para cuyo estudio se requiere un alto nivel de conocimientos de estadística, probabilidad y matemáticas. Entonces esta parte de la estadística se encarga de utilizar los números para proporcionar información

con respecto a grupos más grandes que aquellos a partir de los cuales se obtuvieron los datos originales.

Dentro de la estadística inferencial se tienen dos ramas, estadística no paramétrica y estadística paramétrica, la primera se refiere a información en donde el juez puede discriminar entre una serie de muestras, en evaluación sensorial, ejemplos de este tipo de estadística son la preferencia o aceptación hacia productos determinados, pruebas discriminativas y pruebas de ordenación. Tal información es conocida como datos discretos. Por otra parte la estadística paramétrica emplea información que puede ser medida o posee una magnitud en el valor de los números, un ejemplo de este tipo de estadística son las intensidades de un atributo en unas muestras, a este tipo de datos se le conoce como datos continuos.

Todo o anterior ayuda al evaluador sensorial a clasificar la información de las pruebas sensoriales en base al tipo de estadística, que como ya se observa, no se analiza de la misma manera, es por esto que es necesario determinar claramente el tipo de estadística que es aplicable a cada una de las pruebas sensoriales, y así poderla analizar e interpretar correctamente.

4.2.2 Tipos de pruebas sensoriales

Una vez que el analista sensorial ha determinado el tipo de prueba sensorial que mejor resuelve sus objetivos, tiene que decidir cómo evaluar los datos que obtenga de ella. Para ello requiere determinar el tipo de número con el cual está trabajando. Ver Tabla 4-2.

Tabla 4-2 Relación de las pruebas sensoriales con el tipo de estadística y de número

Tipo de prueba sensorial	Tipo de estadística	Tipo de número
Discriminativas	No paramétrica	Nominal
Ordenación	No paramétrica	Ordinal
Intensidad	Paramétrica	Intervalar
Estimación por magnitudes	Paramétrica	Proporcional

4.2.3 Tipos de números

Los diferentes tipos de números expresan diferentes tipos de escalas, mismos que se explican a continuación.

Número nominal o categoría: Este tipo de número indica una clasificación o nominación sobre objetos, personas o características. Emplea una escala de tipo nominal que se utiliza cuando se hace uso de números o símbolos para clasificar. En evaluación sensorial es muy utilizada para pruebas discriminativas como lo son la triangular, comparación por pares, duo tipo, etc. Los métodos estadísticos que se emplean para analizar información de tipo nominal, son la Ji cuadrada y la prueba binomial¹¹.

Número ordinal: los números ordinales como su nombre lo indica establecen ordenes, utilizan una escala ordinal que es aquella donde se presentan dos o más muestras, se le solicita al juez que ordene tal grupo de muestras desde lo más bajo a lo más alto, según un atributo específico (*dulzor, dureza, intensidad de sabor, etc.*) En evaluación sensorial se utiliza la escala ordinal para determinar ordenes y preferencias entre productos. Tales relaciones pueden formularse con los signos ">" "(mayor que)", y "<" "(menor que)", Por ejemplo ordenar el *dulzor* de varias muestras, presentando soluciones de sacarosa a diferentes concentraciones en donde la solución al 1% de sacarosa es "menor" que la solución al 2%, y está a su vez es menor que la del 4%.²¹ Las pruebas estadísticas empleadas para escalas ordinales son Análisis de Rangos de Newell-MacFarlan, Page Test y la prueba de Friedman.^{18, 22}

Número intervalo: este tipo de número representa intervalos en donde existe la misma distancia entre cada uno de ellos.²¹ Emplea una escala de intervalo que es aquella que permite calificar la percepción de la intensidad o de la calidad de una característica sensorial específica, también conocida como atributo o descriptor. Los jueces entrenados discriminan diferencias de intensidad sensorial de uno o más atributos presentes en una serie de muestras.¹⁵ La escala de intervalos, donde los jueces demuestran el gradiente de las diferencias percibidas, puede ser representada de dos maneras:

Escala estructurada: presenta intervalos definidos.

Intensidad de dulzor



Escala no estructurada : presenta sólo los extremos definidos.

Intensidad de dulzor



En evaluación sensorial se utiliza esta escala para determinar, por ejemplo:

- a) La intensidad percibida con jueces entrenados.
- b) La calidad de un producto específico con jueces altamente entrenados.
- c) El nivel de agrado con consumidores.

Esta representación numérica permite que se utilicen todas las pruebas de la estadística descriptiva (media, desviación estándar etc.) así como las pruebas estadísticas paramétricas t de student y análisis de varianza.¹⁸

Números de magnitud o proporción: utilizan los números para generar proporción, esta escala sugiere medir la relación entre un continuo físico y el sensorial de una manera proporcional. Los jueces primero reciben una referencia con una designación numérica preestablecida (tal como 10 ó 100), o sin predesignación y el juez entonces genera su propio número. Utiliza números enteros o con decimales, haciendo cada asignación numérica proporcional a la intensidad percibida.¹²

4.2.4 Logística para emplear métodos estadísticos en la resolución de problemas.

El procedimiento para aplicar la estadística en investigación es la que a continuación se menciona habitualmente:^{9,22}

1. Identificar el problema
2. Formular objetivos e hipótesis. Las hipótesis se plantean de dos maneras fundamentales: la hipótesis nula H_0 y la hipótesis alterna H_a . La hipótesis nula H_0 es la que se prueba contra su complemento, la hipótesis alternativa (llamada en ocasiones la hipótesis de investigación o de interés). Es conveniente observar que las hipótesis nula y alterna deben formularse de manera tal que no puedan ser verdaderas al mismo tiempo. Se trata de emplear la evidencia disponible para comprobar cuál tiene más probabilidad de ser cierta.⁹
3. Elegir una prueba estadística apropiada para probar H_0 . La prueba se elige en base al tipo de estadística, número y escala empleadas por el investigador.
4. Especificación del nivel de significancia (α) y del tamaño de la muestra (N): el nivel de significancia es la pequeña probabilidad asociada con la que ocurre H_0 .
5. Confrontar la distribución muestral de la prueba estadística conforme a H_0 .
6. Definir la región de rechazo para cada una de las hipótesis planteadas, es decir se plantea una regla de decisión.
7. Calcular el valor de la prueba estadística con los datos obtenidos de las muestras. Se usa la fórmula adecuada para calcular el valor de la prueba estadística elegida. Se compara el valor calculado (previamente) con el valor crítico de la prueba estadística (tomado de la tabla adecuada) con base en la regla de decisión elegida. A partir de la comparación, se formula el resultado en términos de rechazo o no de la hipótesis nula (H_0).⁹

A continuación se expone un ejemplo de la logística anterior aplicado a la clasificación de un juez sensorial para pruebas discriminativas.

1. Problema: Seleccionar qué candidatos a jueces sensoriales tienen habilidad al realizar pruebas sensoriales discriminativas.
2. Hipótesis nula (H_0): el candidato a juez no discrimina sensorialmente entre muestras. Hipótesis alternativa (H_a): el candidato a juez sí discrimina sensorialmente entre muestras.
3. Las pruebas sensoriales discriminativas pertenecen al grupo de la estadística no paramétrica. Este tipo de prueba sensorial emplea números de tipo nominal, entonces el estadístico apropiado es la J_i cuadrada.
4. El nivel de significancia escogido para el nivel crítico de la prueba es de 5%.
5. Confrontación de la distribución muestral con el estadístico elegido. Este punto se refiere a si los datos provienen de una distribución de acuerdo con la prueba estadística escogida, en este caso los datos experimentales se deben ajustar a una distribución de J_i cuadrada para confrontar / aceptar la hipótesis nula (H_0).
6. El nivel de rechazo para la hipótesis nula es con α (nivel de significancia) mayor al establecido como crítico, es decir probabilidades mayores al 5%; o bien con valores experimentales de J_i cuadrada mayores al valor crítico obtenido de tablas de Valores Críticos de J_i Cuadrada.
7. Calcular con los datos del candidato a juez, y de acuerdo con la fórmula de J_i cuadrada, si se acepta o se rechaza la hipótesis nula (H_0).

4.2.5 Relación de la estadística con la evaluación sensorial

En los temas anteriores se expone que la estadística constituye una de las principales herramientas para análisis y comprensión de la información sensorial. En esta parte ampliaremos la relación de la prueba sensorial con la prueba estadística. (Ver Tabla 4-3)

Lo que se propone en esta investigación es estudiar el comportamiento sensorial del candidato a juez, pero ¿cómo empezar?. Es necesario considerar a cada juez como un instrumento individual, es decir se analiza a cada individuo como habitualmente se analiza a una muestra.

Una vez establecido este primer paso es necesario estudiar los datos que genera cada uno de los jueces, pero haciendo diferencia en cada una de las pruebas sensoriales que el candidato a juez ha realizado. En el caso de este estudio se partió de diversas bases de datos proporcionadas por Penta Sensorial S.A. de C.V., estas bases contienen los resultados de las diferentes pruebas de selección (ensayos preliminares a la etapa de entrenamiento). Los datos fueron generados por 17 candidatos a juez, la primera de estas bases de datos comprende información de

pruebas discriminativas, la cual es clasificada como base nominal; la segunda hace referencia a información de tipo ordinal y la tercera pertenece a datos de tipo intervalar.

Tabla 4-3 Pruebas estadísticas comúnmente utilizadas para analizar tipos de pruebas sensoriales

Prueba sensorial	Tipo de estadística	Tipo de número	Prueba de análisis	Aplicación más común
Pruebas discriminativas: Comparación por pares, Dúo-Trio, Triangular.	No paramétrica	Nominal	Frecuencia. Porcentaje. χ^2 (Ji cuadrada). Tablas Binomiales.	Determinar si hay diferencia o semejanza entre un par de muestras.
Pruebas de Ordenación	No paramétrica	Ordinal	Frecuencia. Porcentaje. χ^2 (Ji cuadrada). Análisis de ordenamiento por rangos. Análisis de Friedman. Page Test.	Selección rápida entre varias muestras. En pruebas con consumidores es prueba de preferencia.
Prueba de Intervalos. Análisis Descriptivo	Paramétrica	Intervalo	Promedio. Desviación estándar. Coeficiente de variación. t-student. Análisis de varianza.	Determina intensidades de un atributo en particular o comparar intensidades en varios productos. Por ejemplo QDA (Análisis Descriptivo Cuantitativo)
Prueba de Estimación por Magnitudes	Paramétrica	Magnitud o proporción	Media geométrica.	Pruebas de <i>Long Lasting</i> : Establece la relación de los cambios proporcionales de una percepción humana con respecto a parámetros químicos o físicos.

Nota: se recomienda aplicar la Estadística Descriptiva a las bases de datos que generan estas pruebas sensoriales para comprender visualmente el comportamiento de dicha base. Por ejemplo, si la distribución de los datos experimentales efectivamente sigue un comportamiento de tipo nominal (premisa que debe cumplir toda base de datos antes de aplicar una prueba estadística).

4.3 INFORMACIÓN DE TIPO NOMINAL, MÉTODOS PARA LA OBTENCIÓN DE CRITERIOS ESTADÍSTICOS PARA SELECCIONAR CANDIDATOS A JUECES EN PRUEBAS DISCRIMINATIVAS

Durante la etapa de selección es necesario realizar pruebas filtro que ayuden a clasificar a los candidatos a jueces sensoriales. Las primeras pruebas que se realizan son pruebas de identificación de los cuatro gustos básicos, una vez que el candidato logra identificarlos tiene que determinar el nombre del estímulo que se le está proporcionando. Para poder hacer una selección de entre los candidatos en esta etapa se proponen criterios basados en estadística no paramétrica nominal. En los siguientes párrafos se describen las pruebas sensoriales que involucran información de tipo nominal, así como las pruebas estadísticas que analizan este tipo de información.

4.3.1 Pruebas sensoriales que involucran información de tipo nominal

4.3.1.1 Identificación de gustos básicos:

El analista sensorial, cuando forma su grupo de jueces, parte de ensayos preliminares y con por lo menos el doble de los candidatos que constará su grupo final o deseado. Por tal razón tiene que realizar una serie de pruebas para detectar la capacidad innata de cada candidato. Una de esas pruebas preliminares es la identificación de los 4 gustos básicos (*ácido, amargo, dulce y salado*), donde se le presenta al candidato muestras de soluciones de los cuatro gustos básicos a una determinada concentración. En esta serie de soluciones se incluye una o más muestras de agua simple, de manera que el candidato pueda conocer y distinguir el gusto básico del agua.^{8,16}

Esta prueba es posible analizarla por estadística no paramétrica, el tipo de número es nominal, debido a que se asigna el número "1" a una "identificación correcta" y "0" a una "falla".¹⁸ Finalmente la frecuencia de dichas asignaciones se manejan como porcentaje de identificación.

4.3.1.2 Pruebas discriminativas

Comparación por pares: esta prueba para el caso de una comparación por pares no direccionada, consiste en proporcionar al posible juez un par de muestras y preguntarle si son semejantes o diferentes, o bien en una comparación por pares direccionada se pregunta: ¿cuál de las dos muestras es más ácida? .^{6,21} La probabilidad

asociada con esta prueba de que un juez pueda acertar a la respuesta correcta al azar es de 0.5 ó 50%.

Dúo-Trío: consiste en presentar al juez una referencia y dos muestras. Se le pregunta ¿Cuál de entre estas dos es igual a la referencia?. Es una prueba direccionada ya que el analista de antemano conoce la respuesta correcta. La probabilidad de obtener la respuesta correcta por casualidad es de 0.5 ó 50%.¹⁶

Doble Referencia: se emplean dos referencias y dos muestras. El juez tiene que responder a la pregunta ¿Cuál muestra es igual a la referencia 1?. Es una prueba de elección forzada y la probabilidad de acertar al azar es de 0.5.¹⁹

Triangular: se presentan tres muestras al juez, en donde dos son iguales y una diferente. El juez tiene que indicar ¿Cuál muestra es la diferente de las otras dos (que son semejantes entre si)? La probabilidad de acertar al azar en esta prueba es de 0.3.^{16,19,20}

4.3.2 Pruebas estadísticas no paramétricas consideradas en la obtención de criterios para la selección de candidatos a jueces sensoriales al utilizar información de tipo nominal

4.3.2.1 Prueba binomial

Aplicaciones de la prueba binomial: esta prueba es utilizada para determinar el numero de casos identificados en las diferentes categorías.¹⁹ En análisis sensorial se utiliza para determinar si hay diferencia significativa a una probabilidad o nivel de significancia predeterminado para una prueba discriminativa o en una prueba de preferencia con dos muestras. Las condiciones para emplear esta prueba son:²³ los datos deben provenir de una escala nominal, sólo hay dos resultados posibles (acierto o falla) y se debe tener la misma probabilidad para cada resultado.

En la aplicación de la prueba binomial a pruebas discriminativas¹ se utiliza la fórmula de la expansión binomial:

$$(p+q)^n = p^n + np^{n-1}q + n(n-1) p^{n-2}q^2 + \dots + \frac{n}{N!(n-N)!} p^{n-N}q^N + \dots + npq^{n-1} + q^n$$

Donde:

p = probabilidad del acierto
q = probabilidad de fallo
n = número de eventos
N = población

¹ Consultar Tablas G4.a G4.b y G4.c en. O' Mahony 1986 Sensory evaluation of food. Statistical methods and procedures. Marcel dekker, inc. New York U.S.A

4.3.2.2 Ji cuadrada (χ^2)

Aplicaciones de la prueba χ^2 : La Ji cuadrada como estadístico para pruebas de diferenciación, se utiliza para probar, de acuerdo con cierta hipótesis, en qué grado una distribución de frecuencia observada, se comparará con una distribución esperada o teórica². Esta prueba estadística permite determinar si las comparaciones entre muestras que generan las pruebas discriminativas son entre sí significativamente diferentes o no. Al igual que la prueba binomial es utilizada para determinar si más casos o categorías (nominal) caen en una categoría que otra, pero con la diferencia que en la Ji cuadrada las categorías pueden ser diferentes, por ejemplo, se puede determinar si un juez distingue entre un par de muestras con sensaciones diferentes como una muestra con agua salada y otra con agua azucarada.

Ji cuadrada para pruebas discriminativas

$$\chi^2 = \frac{(|X_1 - np| - 0.5)^2}{np(1-p)}$$

Donde:

χ^2 = Ji cuadrada

X_1 = número de opiniones acertadas o asignadas a una de las dos categorías

n = número total de ensayos practicados o número de jueces multiplicado por las repeticiones efectuadas.

p = probabilidad del éxito en un ensayo único

q = (1-p) probabilidad de la falla en un ensayo único

0.5 = factor de corrección por continuidad para Ji cuadrada ajustada

4.3.3 Definición de los criterios con fundamentos estadísticos para seleccionar candidatos a jueces en pruebas discriminativas.

En esta investigación se pretende ofrecer al analista sensorial una serie de criterios, con fundamentos estadísticos, que le permitan tomar decisiones en cuanto al rendimiento de su grupo de jueces en el desempeño de las pruebas sensoriales, de esta manera tendrá una clasificación estadística de los jueces durante la etapa de selección.

Para los fines de este trabajo se propone que los jueces sean clasificados en cuatro categorías: "Malo", "Regular", "Bueno" y "Muy Bueno", los criterios para ubicarlos en cada una de estas clasificaciones se establecen de acuerdo a las pruebas estadísticas aplicables a las pruebas discriminativas.

En la prueba de identificación de gustos básicos se propone que el criterio de selección se base en porcentajes de identificación para cada gusto básico evaluado. Esto es, se considera el total de evaluaciones como el 100 por ciento y se calcula el porcentaje de los ensayos que el juez logró identificar correctamente, como se explica en la Tabla 4-4:

Tabla 4-4 Criterio 1 Selección de jueces en pruebas discriminativas, basado en el porcentaje de identificación de gustos básicos.

Porcentaje de identificación	No aplica	(0-25%)	(26-50%)	(51-75%)	(76-100%)
Significado	Candidatos que no cumplen con su asistencia.	Candidatos que no identifican un mínimo de gustos básicos.	Es muy difícil para el candidato distinguir entre los gustos básicos ya que lo máximo que puede identificar es un 50% de ellos.	El candidato distingue por arriba del 50% de los ensayos, pero no logra a diferenciar de manera total	Del total de aciertos, el candidato identifica a la gran mayoría de los gustos básicos
Clasificación	Faltas	Malo	Regular	Bueno	Muy bueno
Puntaje	0	1	2	3	4

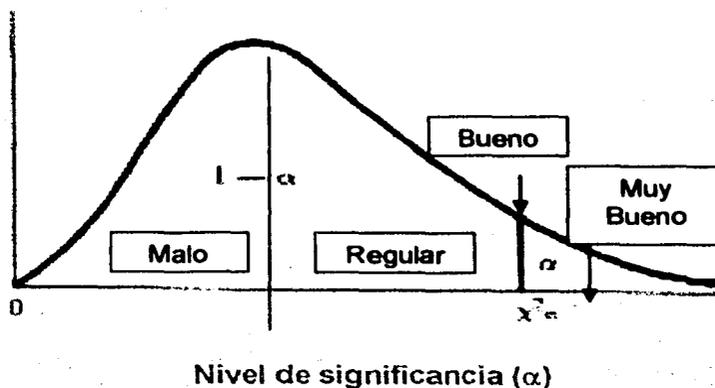
De acuerdo a la lógica para emplear métodos estadísticos en la resolución de problemas se plantea el siguiente criterio:

- 1. Problema:** Determinar qué candidatos a juez sensorial tienen habilidad para desempeñar pruebas discriminativas.
- 2. Planteamiento de Hipótesis:**
 Hipótesis nula (H_0): el candidato a juez no discrimina entre muestras.
 Hipótesis alternativa (H_a): el candidato a juez sí discrimina entre muestras.
- 3. Selección de la prueba estadística:** Las pruebas discriminativas pertenecen al grupo de la estadística no paramétrica. Este tipo de prueba sensorial emplea números de tipo nominal, entonces el estadístico apropiado es la Ji cuadrada.

4. **Planteamiento del nivel crítico de significancia:** Como se están manejando candidatos a juez, el nivel de significancia escogido es de 5%.¹
5. **Definición de la zona de rechazo:** La zona de rechazo para la hipótesis nula será con α (nivel de significancia) experimental menores al establecido en el planteamiento de la hipótesis, es decir probabilidades menores al 5%; o bien valores de Ji cuadrada experimental mayores al valor crítico, con este resultado se aceptaría la hipótesis alternativa. (Ver tabla 3-2)
6. **Calcular:** con los datos del candidato a juez, y de acuerdo con la fórmula de Ji cuadrada, si se acepta o se rechaza H_0 . Si el valor obtenido de Ji cuadrada experimental cae en la zona de rechazo para H_0 entonces se aceptará H_a y se concluye que el candidato a juez discrimina entre muestras.

En la siguiente figura se establecen las zonas de rechazo de acuerdo a los valores obtenidos de Ji cuadrada para un determinado candidato a juez, también se incluye la clasificación que corresponde a mencionado valor.

Figura 4-1 Criterio 2 Gráfico para la selección de Jueces para pruebas discriminativas utilizando el estadístico ji cuadrada



¹ El nivel de confianza se establece de acuerdo a los objetivos y compromisos que emanan de cada estudio

Tabla 4-5 Criterio 2 Selección de jueces para pruebas discriminativas basado en el estadístico Ji Cuadrada

Valor de Ji cuadrada	No aplica	No aplica	Valor crítico menos uno	Mínimo de juicios correctos para ser diferentes significativamente	Arriba del valor mínimo y hasta el máximo de aciertos
Significado	Los candidatos que por asistencia no son acreedores a puntos.	Candidatos que no discriminan significativamente, no identifican atributos.	El juez se ubica tan solo a un punto del mínimo de juicios para encontrar diferencia significativa	El candidato logra el mínimo de juicios correctos para encontrar diferencia significativa	Evidentemente el juez discrimina entre muestras
Clasificación	Faltas	Malo	Regular	Bueno	Muy bueno
Puntaje	0	1	2	3	4

Con la información recavada al momento, se propone realizar una combinación de los criterios anteriores basados en el porcentaje de identificación y en el estadístico Ji cuadrada, sumando los puntos obtenidos por el juez en dichos criterios. Ver tabla 3-4

Tabla 4-6 Criterio 3 Suma de puntaje total aplicando los Criterios 1 y 2 en la selección de candidatos a jueces para pruebas discriminativas

Clasificación	Faltas	Malo	Regular	Bueno	Muy bueno
Puntaje	0	1-2	3-4	5-6	7-8

Ejemplo: ¿Cómo aplicar los criterios para seleccionar candidatos a jueces para pruebas sensoriales que involucran información nominal?¹

- Revisar el tipo de prueba sensorial aplicada, y su correspondiente análisis estadístico.
- Considerar a cada juez como instrumento individual (muestra).

¹ Se sugiere consultar la metodología para una explicación más detallada

Tabla 4-7 Resultados para el juez 1 en la evaluación de gustos básicos.

Prueba realizada: Comparación por pares, Atributo ácido
 Fallo = 0 Acierto = 1

Juez	Repetición	Serie 1	Serie 2
1	1	1	1
1	2	1	0
1	3	0	0
1	4	0	0

Porcentaje de identificación

Ensayos totales (4 repeticiones, 2 series) = 8 Aciertos = 3

Porcentaje de aciertos = $(3 \cdot 100) / 8 = 35.7\%$

Comentario: con este resultado del juez 1, y aplicando el Criterio 1 se clasifica como "Regular".

Análisis de la información por medio de Ji cuadrada

Sustituyendo:

$$\chi^2 = \frac{(|3 - 8 \cdot 0.5| - 0.5)^2}{8 \cdot 0.5 (0.5)} = 0.125$$

χ^2 = Ji cuadrada

$X_1 = 3$

$n = 8$

$p = 0.5$ por ser comparación por pares, en caso de utilizar triangular $p = 0.3$

$q = (1-p) = 0.5$

0.5 = factor de corrección por continuidad para Ji cuadrada ajustada.

Se busca en Tablas de valor crítico donde se plantea la probabilidad asociada con el evento para pruebas de comparación por pares para una cola (prueba direccional) y un grado de libertad siendo χ^2 con g.l. = 1 y una $p = 0.05$, el valor crítico es: 2.71. Se compara este valor crítico con el valor obtenido del experimento y se observa cual es más grande. En este caso el valor crítico es más grande que el experimental, esta comparación nos lleva a la conclusión que la probabilidad para que se lleve a cabo la igualdad es aún más grande que el límite preestablecido en el planteamiento de la Hipótesis Nula ($p \leq 0.05$), entonces se acepta H_0 , en la que se indica que el juez no encuentra diferencia entre las muestras. O bien se busca en Tablas Estadísticas¹ en donde para el total de ensayos (en este caso 8) se necesitan por lo menos 7 juicios correctos para encontrar diferencia significativa, el candidato a juez 1 obtuvo 3 aciertos por lo cual se concluye que no encuentra diferencia entre las muestras.

¹ Para este evento se utiliza la tabla F.1 de Pedrcro y Pangborn (1989) Mínimo de juicios correctos en pruebas de comparación por pares para 8 juicios.

Aplicando el Criterio 2, este juez queda entonces clasificado como "Malo". Para aplicar el Criterio 3 se suman los puntos de cada clasificación: 2 puntos por ser "Regular" en el Criterio 1, más 1 punto obtenido en el Criterio 2, da un total de 3 puntos, que le da una clasificación de "Regular".

En este ejemplo sólo se está manejando un atributo, para evaluar a todos los atributos utilizados por el analista sensorial mediante los criterios planteados, solo se tiene que multiplicar los puntajes dados en tales criterios por el número total de atributos. Una vez obtenido el puntaje final se compara con lo que cada juez obtuvo al sumar todos los puntos obtenidos en cada uno de los atributos analizados.¹

4.4 INFORMACIÓN DE TIPO ORDINAL, MÉTODOS PARA LA OBTENCIÓN DE CRITERIOS ESTADÍSTICOS PARA SELECCIONAR CANDIDATOS A JUECES EN PRUEBAS DE ORDENACIÓN.

Una vez que el candidato a juez logra identificar y discriminar entre gustos básicos, el siguiente paso dentro del proceso de selección de jueces, es observar su rendimiento en pruebas de ordenación. En esta etapa se selecciona y clasifica a la base de datos que contiene este tipo de número como Información Ordinal. Los criterios de selección están apoyados en la estadística no paramétrica ordinal.

4.4.1 Pruebas sensoriales de ordenación

Las pruebas de ordenación son un método de evaluación sensorial que se utiliza para clasificar el orden entre varias muestras de acuerdo a un atributo predeterminado, como puede ser *dulzor, acidez, etc.* En este tipo de pruebas de preferencia no se permiten empates.

El método es aplicable para una clasificación entre varios productos, por ejemplo, la clasificación de acuerdo a la intensidad de un atributo, sus tipos de componentes o la impresión afectiva sobre la serie de productos⁷ (esta última parte se refiere a pruebas de preferencia que no entrarían en este estudio por ser jueces consumidores y no analíticos).

Para analizar una base de datos proveniente de pruebas ordinales se utilizan las pruebas estadísticas de Page Test, el Análisis de Rangos de Newell-MacFarlane y la Prueba de Friedman. En este estudio se propone compararlos para determinar la prueba que facilite y contraste una clasificación de los candidatos a jueces. Es importante mencionar que el analista sensorial está en libertad de escoger el método que mejor se adapte a sus necesidades.

¹ Se sugiere revisar la aplicación de los criterios en la parte de resultados.

4.4.2 Pruebas estadísticas no paramétricas para la selección de candidatos a jueces sensoriales al utilizar información de tipo ordinal

4.4.2.1 Page Test

Aplicaciones: determina si se esta evaluando dentro de un orden ya establecido en la metodología donde $A < B < C < D < E$.

Condiciones para usar Page Test:¹⁸

1. Esta es una prueba de más de dos muestras, donde cada calificación debe corresponder a cada una de ellas, así como todos los jueces deben evaluar cada una de ellas.
2. Cada juicio debe ser independiente, es decir sin influencia de los demás.
3. Los datos deben provenir de una escala ordinal.

Se utiliza la fórmula:

$$L = R_1 + 2 R_2 + 3 R_3 + 4 R_4 + 5 R_5 + \dots$$

Donde:

L = Valor calculado para Page Test.

R = Suma de rangos para cada muestra, pero son analizadas de acuerdo a un orden preestablecido o conocido.

Se busca en Tablas de Valores Críticos de Page Test con k=número de muestras y N=número de juicios el valor de acuerdo al nivel de significancia establecido. El valor calculado tiene que ser igual o mayor al de la Tabla para declarar que si existe diferencia significativa entre los tratamientos.¹⁸

4.4.2.2 Análisis de ordenamiento por rangos de Newell-MacFarlane

Aplicaciones: para la comparación de todas las muestras entre sí o entre una referencia y varias muestras.

Esta prueba no emplea fórmulas como en Page Test o la prueba de Friedman, se utilizan "Tablas de Diferencia de sumatoria ordinal absoluta crítica de todos los tratamientos" que contienen los valores críticos necesarios para decir que hay diferencia significativa en el ordenamiento de las muestras.¹⁸

Para analizar información de tipo ordinal por este método, se suman todas las calificaciones que obtuvo cada muestra obteniendo la suma de rangos, este valor se comparara contra los valores de tablas si las muestras cumplen con la diferencia absoluta entre cada una de ellas.

4.4.2.3 Prueba de Friedman

Aplicaciones de la prueba de Friedman: determina si existe diferencia significativa en el orden dado a varias muestras a una probabilidad y nivel de significancia elegidos.

Condiciones para utilizar la prueba de Friedman^{18,22}

1. La prueba debe tener más de dos muestras .
2. Los datos tienen que provenir de una escala tipo ordinal.
3. Cada juez debe evaluar independientemente, sin influencia de los demás .
4. Todas las muestras tienen que ser evaluadas por cada uno de los jueces.

Se emplea la siguiente fórmula:

$$\chi_r^2 = \frac{12}{Nk(k+1)} \sum R^2 - 3N(k+1)$$

Donde:

N = número de juicios o jueces totales

k = número de tratamientos o muestras

R = sumatoria de los ordenes dados a las muestra por el grupo

Además de consultar la Tabla Probabilidad asociada con valores observados en la prueba de Friedman¹, para analizar datos de tipo ordinal, se puede utilizar la Tabla Valores Críticos de Ji cuadrada² con el nivel de significancia planteado en la hipótesis y de dos colas, con grados de libertad k-1, considerando que se tienen muestras grandes.

4.4.3 Definición de criterios con fundamentos estadísticos para seleccionar candidatos a jueces sensoriales en pruebas de ordenación

Al igual que en las pruebas discriminativas, en esta parte también se busca clasificar a los candidatos a jueces sensoriales en "Muy Bueno", "Bueno", "Regular" y "Malo". Para poder definir los criterios, se prueban los tres métodos estadísticos para pruebas de ordenación y, en base a los resultados se selecciona aquel que permita una clasificación de los jueces.

¹ Consultar Tabla G.22 O'Mahony 1986

² Consultar Tabla B en Pedrero y Pangborn 1989 o Tabla G.7 en O'Mahony 1986

4.4.3.1 Aplicación del método Page Test

1. **Problema:** Determinar que candidatos a jueces sensoriales tienen habilidad en pruebas de ordenación, con un orden previamente establecido o conocido por el analista sensorial por ejemplo: diferentes concentraciones de algún componente.
2. **Planteamiento de Hipótesis:**
Hipótesis nula (H_0): el candidato a juez no encuentra diferencia significativa al ordenar entre muestras.
Hipótesis alternativa (H_a): el candidato a juez sí encuentra diferencia significativa al ordenar entre muestras.
3. **Selección de la prueba estadística:** Page Test en donde se tiene un orden preestablecido o conocido.
4. **Establecimiento del nivel de significancia:** Debido al grado de confiabilidad requerido para este estudio se consideró un nivel de significancia de 5%.
5. **Definición de la zona de rechazo:** Los puntos de rechazo de la hipótesis nula serán con α (nivel de significancia) mayores al establecido, es decir probabilidades menores al 5%. En el caso de este método se busca en la Tabla "Valores Críticos de Page Test" (ver Tabla G.25 O' Mahony, 1986) el valor crítico al nivel de confianza elegido. La zona de rechazo para H_0 será entonces los valores calculados de Page Test (L) mayores al valor crítico establecido en dicha tabla.
6. **Calcular con los datos del candidato a juez y de acuerdo con la fórmula de este método, el valor experimental (o calculado) de Page Test, para determinar si se acepta o se rechaza H_0 .** Si el valor obtenido de Page Test cae en la zona de rechazo para H_0 entonces se aceptará H_a y por conclusión el candidato a juez encuentra diferencia significativa al ordenar entre muestras.

Tabla 4-8 Prueba de Ordenación, analizada por Page Test ,Atributo Color, juez 1

Juicios	Muestras					Orden Establecido Color	Rangos del juez
	A	B	C	D	E		
						D (R1)	8
1	4	3	2	1	5	B (R2)	10
2	3	1	2	4	5	A (R3)	13
3	3	4	1	2	5	C(R4)	10
4	3	2	5	1	4	E (R5)	19
Suma rangos	13	10	10	8	19		

Aplicando la fórmula:

$$L = R_1 + 2 R_2 + 3 R_3 + 4 R_4 + 5 R_5 + \dots$$

Sustituyendo:

$$L = 8 + (2 * 10) + (3 * 13) + (4 * 10) + (5 * 19) = 202$$

El valor L obtenido (202) es mayor que el valor crítico (197), con este resultado se rechaza H_0 y se acepta H_a en donde se plantea que el candidato sí encuentra diferencia significativa al ordenar entre varias muestras.

4.4.3.2 Aplicación del análisis de ordenamiento por rangos de Newell-MacFarlane

1. **Problema:** Determinar qué candidato a juez sensorial tiene habilidad en pruebas de ordenación al comparar varias muestras entre sí.
2. **Planteamiento de Hipótesis:**
Hipótesis nula (H_0): el candidato a juez no ordena de manera significativa entre varias muestras.
Hipótesis alternativa (H_a): el candidato a juez ordena de manera significativa entre varias muestras.
3. **Selección de la prueba estadística:** Análisis de ordenamiento por rangos de Newell-MacFarlane.
4. **Establecimiento del nivel de significancia:** Debido al grado de confiabilidad requerido para este estudio se consideró un nivel de significancia de 5%.

5. **Definición de la zona de rechazo:** El rechazo de la hipótesis nula serán con α (nivel de significancia) mayores al establecido es decir probabilidades menores al 5%. En este caso se busca en Tablas de Diferencia de sumatoria ordinal absoluta crítica de "todos los tratamientos" el valor crítico al nivel de significancia elegido para esta prueba con un total de 5 muestras y 4 juicios, es de un mínimo de 13 unidades de diferencia entre cada suma de rangos de las muestras evaluadas. La zona de rechazo será entonces diferencias absolutas entre rangos menores al valor crítico.

6. **Calcular** los datos del candidato a juez con lo establecido en la tabla de valores críticos, se acepta o se rechaza H_0 , si el valor obtenido de la diferencia entre rangos cae en la zona de rechazo para H_0 entonces se aceptará H_a y por conclusión el candidato a juez ordena significativamente muestras.

Tabla 4-9 Prueba de Ordenación, Análisis de ordenamiento por rangos de Newell-MacFarlane, Atributo Color, juez 1

Juicios	Muestras				
	A	B	C	D	E
1	4	3	2	1	5
2	3	1	2	4	5
3	3	4	1	2	5
4	3	2	5	1	4
Suma rangos	13	10	10	8	19
Diferencia absoluta entre rangos			Valor crítico Tabla G23 O' Mahony = 13		
A-B	13 - 10	3	<	13	
A-C	13 - 10	3	<	13	
A-D	13 - 8	5	<	13	
B-C	10 - 10	0	<	13	
B-D	10 - 8	2	<	13	
B-E	10 - 19	9	<	13	
C-D	10 - 8	2	<	13	
C-E	10 - 19	9	<	13	
D-E	8 - 19	11	<	13	

De acuerdo a lo establecido en las tablas de valores críticos para este método se obtiene que la diferencia absoluta crítica entre rangos tiene que ser mayor o igual a 13, para rechazar H_0 , por lo tanto se declara que no hay diferencia significativa entre las ordenaciones efectuadas por el juez 1 aplicando el Método de Análisis de Rangos de Newell-MacFarlane, ya que todas las diferencias absolutas entre rangos son menores al valor crítico.

4.4.3.3 Aplicación de la prueba de Friedman

1. **Problema:** Determinar qué candidato a juez sensorial tiene habilidad en pruebas de ordenación al comparar varias muestras entre sí.
2. **Planteamiento de Hipótesis:**
Hipótesis nula (H_0): el candidato a juez no ordena de manera significativa entre muestras.
Hipótesis alternativa (H_a): el candidato a juez ordena de manera significativa entre muestras.
3. **Selección de la prueba estadística:** Prueba de Friedman
4. **Establecimiento del nivel de significancia:** Debido al grado de confiabilidad requerido para este estudio se consideró un nivel de significancia de 5%.
5. **Definición de la zona de rechazo:** Los puntos de rechazo para la hipótesis nula serán con α (nivel de significancia) mayores al establecido es decir probabilidades menores al 5%. Se busca en Tablas de Probabilidad asociada con valores observados en la prueba de Friedman, el valor crítico al nivel de confianza elegido. La zona de rechazo será entonces valores calculados menores al crítico.
6. **Calcular** con los datos del candidato a juez, y de acuerdo con las fórmulas de Friedman el valor experimental con el fin de determinar si se acepta o se rechaza H_0 . Si el valor obtenido de Friedman es menor que el valor crítico, involucra aceptar H_0 y por conclusión el candidato a juez no ordena significativamente entre varias muestras.

Tabla 4-10 Prueba de Ordenación, analizada por la prueba de Friedman, Atributo Color, juez 1.

Juicios (N)	Muestras (k)				
	A	B	C	D	E
1	4	3	2	1	5
2	3	1	2	4	5
3	3	4	1	2	5
4	3	2	5	1	4
Suma rangos	13	10	10	8	19
Suma de rangos al cuadrado	169.0	100.0	100.0	64.0	361.0
Suma total rangos al cuadrado	794.0				
N	4.0				
K	5.0				
K+1	6.0				
12/Nk(k+1)	0.100				
3N(k+1)	72.0				
Friedman calculada	7.4				
Friedman teórica (0.05, k-1=4)	9.49				

Cálculos:

$$\chi^2 = \frac{12}{4 * 5(5+1)} (13^2 + 10^2 + 10^2 + 8^2 + 19^2) - 3 * 4(5+1) = 7.4$$

Se busca en la tabla donde se establece la probabilidad asociada con el evento para análisis de rangos de Friedman¹⁸ y para una χ_r^2 con k=5 y N=4 y una p=0.05. Esta Tabla contiene datos sólo para k=3 y k=4 muestras por lo que se utilizó la tabla Valores críticos de Ji cuadrada con grados de libertad k-1. En donde se obtiene el valor crítico para p=0.05 y grados de libertad =4 que es 9.4. Se compara este valor crítico con el valor experimental, y se observa que es más pequeño el valor experimental que el crítico. Esta comparación nos lleva a la conclusión que la probabilidad para que se lleve a cabo la igualdad es aún más grande que el límite preestablecido en el planteamiento de la Hipótesis Nula ($p \leq 0.05$), concluyendo que se acepta la H_0 , en la que se indica que no hay diferencia entre los muestras.

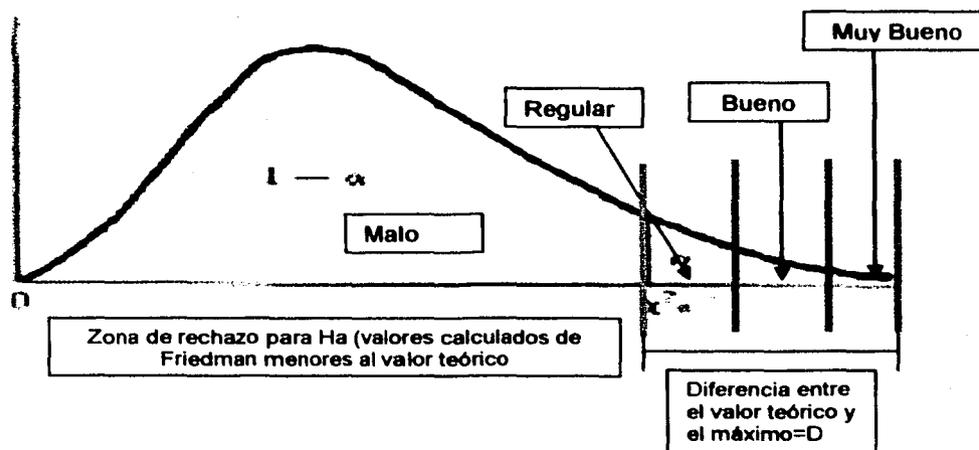
Al comparar los tres métodos estadísticos para pruebas de ordenación, se observa que Page test y el Análisis de Rangos de Newell-MacFarlane son muy generales, pues no permiten emitir una clasificación de los candidatos¹ ya que en el primero la mayoría de los candidatos obtiene resultados satisfactorios permitiendo aceptar en casi todos los casos la hipótesis alternativa que indica que el candidato a juez encuentra diferencia significativa al ordenar entre varias muestras, por el contrario en el segundo la mayoría de los candidatos son situados dentro de la categoría más

¹ Ver página 32 en la parte de Resultados, o consultar disco con anexo de cálculos

baja (Malo) y en todos esos casos se acepta la hipótesis nula en donde se establece que el candidato no encuentra diferencia significativa al ordenar entre varias muestras. Con estos resultados se decidió utilizar la prueba de Friedman como el estadístico base para la definición de los criterios de selección de candidatos a jueces sensoriales en pruebas de ordenación, esto debido a que permite realizar una clasificación de los candidatos a partir de sus resultados obtenidos al ordenar entre varias muestras.

El siguiente diagrama muestra la ubicación de las diferentes clasificaciones "Malo", "Regular", "Bueno" y "Muy Bueno" dependiendo del valor de Friedman obtenido. El valor teórico de Friedman se ubica sobre la línea naranja, valores menores a éste indican que el candidato no ordena de manera significativa entre muestras y es clasificado como "Malo". Para las siguientes clasificaciones se considera la diferencia que existe entre el valor teórico de Friedman y valor máximo de Friedman (línea azul) que se puede obtener para una prueba en específico, esta diferencia se divide entre 3 (mostrados con líneas verdes).

Figura 4-2 Criterio 4 Selección de Jueces para pruebas de ordenación utilizando el estadístico Prueba de Friedman



Nota: el valor D es una diferencia entre el valor máximo de Friedman y el valor crítico, El valor maximo se obtiene cuando el candidato a juez ordena congruentemente, por ejemplo en una ordenación de menor a mayor concentración de soluciones de sacarosa al 0.1%, 0.5%, y 1% siempre ordena en ese orden y con el mismo rango.

Tabla 4-11 Criterio 4 Selección para pruebas ordenación basado en el estadístico Prueba de Friedman

Valor de Friedman	Negativo por faltas	Negativos	$0 > \text{Valor Friedman calculado} < \text{Valor Friedman teórico}$	$\text{Valor Friedman teórico} < \text{Valor Friedman calculado} < \text{Valor Friedman teórico} + 1/3 D$	$\text{Valor Friedman teórico} + 1/3 D < \text{Valor Friedman calculado} < \text{Valor Friedman teórico} + 2/3 D$	$\text{Valor Friedman teórico} + 2/3 D < \text{Valor Friedman calculado} < \text{Valor Friedman máximo}$
Significado	No asistieron a la sesión	Repetieron valores en el análisis de rangos, aún cuando en las instrucciones no se permiten empates	No encuentran diferencia significativa al ordenar, ya que tienen un valor de Friedman menor al teórico	El candidato ordena pero lo hace cubriendo como máximo un tercio del valor de Friedman de lo que debería ordenar	El candidato ordena mejorando el número de ordenamiento ya cubre como máximo dos tercios del valor total de Friedman	El juez puede obtener como máximo el valor de Friedman y lograr el total para una ordenación perfecta
Clasificación	Faltas	Negativos	Malo	Regular	Bueno	Muy bueno
Puntaje	0	0	1	2	3	4

El valor máximo se refiere cuando el juez ordena evaluando siempre con el mismo rango todas las muestras, de acuerdo a un orden previamente establecido como se observa en la siguiente tabla:

Tabla 4-12 Ejemplo de una ordenación perfecta

Juicios	Muestras				
	A	B	C	D	E
1	3	4	2	5	1
2	3	4	2	5	1
3	3	4	2	5	1
4	3	4	2	5	1
Suma Rangos	12	16	10	20	4
Orden Establecido	Juez				
Color					
D (R1)	20				
B (R2)	16				
A (R3)	12				
C (R4)	8				
E (R5)	4				

Se creó otro criterio porque en una prueba de ordenación un candidato puede obtener un valor de Friedman por arriba del teórico, sin embargo es posible que en

algunos casos se repitan sumas de rangos entre algunas muestras por esta razón este criterio se basa en el número de muestras que el juez no califica dándole la misma suma de rangos. Un candidato que haya clasificado de acuerdo al Criterio 4 como "Malo" no puede ser considerado en este criterio puesto que no encuentro diferencia significativa al ordenar entre muestras y automáticamente es clasificado como "Malo", jueces con faltas y /o valores de Friedman negativos clasifican en tal segmento (al igual que en el Criterio 4) sin derecho a puntos. Quedando de la siguiente manera:

Tabla 4-13 Criterio 5 Selección de jueces considerando el número de muestras que el candidato a juez ordena de sin repetir valores de suma de rangos.

Muestras sin repetir valores de suma de rangos	Faltas	Negativos	No aplica	60%	80%	100%
Significado	Jueces que faltaron a las sesiones	Repetieron valores en el análisis de rangos	Se considera a todos los jueces que clasificaron como "Malo" en el Criterio 4, puesto que no encuentran diferencia significativa al ordenar entre varias muestras.	El candidato repite valores en un 40%, es decir de 5 muestras presentadas evalúa a dos pares de ellas con el mismo rango, considera a tres muestras con el mismo rango.	Dentro de la serie muestras valora a un par de muestras con el mismo rango.	No repite rangos para ninguna de las muestras presentadas evalúa de manera diferente cada muestra
Clasificación	Faltas	Negativos	Malo	Regular	Bueno	Muy bueno
Puntaje	0	0	1	2	3	4

Para evaluar al candidato se propone un criterio que involucra la suma de los puntajes en los Criterios 4 y 5.

Tabla 4-14 Puntaje para el Criterio 6 selección de jueces para pruebas de ordenación sumando los criterios basados en el estadístico Friedman y en el número de muestras sin repetir valores de suma de rangos.

Clasificación	Muy bueno	Bueno	Regular	Malo	Faltas y negativos
Criterio 4	4	3	2	1	0
Criterio 5	4	3	2	1	0
Puntaje Total	7-8	5-6	3-4	1-2	0

Ejemplo: ¿Cómo aplicar los criterios para seleccionar candidatos a jueces para pruebas sensoriales que involucran información ordinal?

Tomando como ejemplo los resultados de la prueba de ordenación del atributo color para el candidato a juez número uno (ver Tabla 4-8)

Tabla 4-15 Prueba de Ordenación, analizada por la prueba de Friedman, Atributo Color, juez 1.

Juicios (N)	Muestras (k)				
	D <	B <	A <	C <	E
1	1	3	4	2	5
2	4	1	3	2	5
3	2	4	3	1	5
4	1	2	3	5	4
Suma rangos	8	10	13	10	19
Suma de rangos al cuadrado	64.0	100.0	169.0	100.0	361.0
Suma total de rangos al cuadrado	794.0	Friedman teórica (0.05,k-1=4) 9.49 Friedman calculada = 7.4			
N = 4 k= 5					

Cálculos:

$$\chi^2 = \frac{12}{4 * 5(5+1)} (13^2 + 10^2 + 10^2 + 8^2 + 19^2) - 3 * 4(5+1) = 7.4$$

Al evaluar el resultado de la prueba de ordenación que efectuó el juez 1 por la prueba de Friedman se observa que el valor calculado (7.4) es menor que el valor teórico (9.49). Aplicando el Criterio 4 *Selección de candidatos a jueces sensoriales en pruebas de ordenación basado en el estadístico prueba de Friedman* este juez clasifica en la categoría de "Malo", puesto que no ordena significativamente entre las muestras y por consiguiente no puede ser considerado en el Criterio 5 *Selección de jueces considerando el número de muestras que el candidato a juez ordena de sin repetir valores de suma de rangos*, siendo clasificado automáticamente como "Malo". Por último al sumar el total de puntos de acuerdo al Criterio 6, permanece en la categoría de "Malo", puesto que obtuvo 1 punto en cada uno de los criterios, lo que le da un total de dos puntos.

4.5 INFORMACIÓN DE TIPO INTERVALAR, MÉTODOS PARA LA OBTENCIÓN DE CRITERIOS ESTADÍSTICOS PARA SELECCIONAR CANDIDATOS A JUECES EN PRUEBAS DE INTENSIDAD.

La importancia de este último paso radica en que el candidato a juez, después de pasar las pruebas anteriores, donde logra diferenciar entre estímulos, y ordenar sus percepciones, ahora se observará si es capaz de calificar la intensidad un estímulo.

4.5.1 Pruebas sensoriales de intensidad

La prueba de intensidad sensorial consiste en medir la magnitud que presenta un estímulo, un ejemplo de esto es determinar qué tan ácido es un yogurt. Para poder cuantificar el estímulo se emplean escalas (ver tema 4.2) en donde el juez sensorial marca la intensidad que percibe para ese producto y ese atributo en específico. Este tipo de prueba tiene varias ventajas, la primera es conocer la magnitud del atributo y la segunda es que se pueden determinar entre varios productos si sus intensidades entre sí son diferentes y si lo son con que magnitud y dirección.¹⁰

La prueba de intensidad es útil para otras pruebas sensoriales, como el Análisis Descriptivo Cuantitativo (QDA por sus siglas en inglés) la cuál consiste en determinar las principales características sensoriales de un producto, determinando en cada una de ellas la intensidad presente en cada atributo. Otra prueba sensorial en donde se emplea es el Perfil de Sabor, que evalúa las características integrales de sabor presentes en un alimento y determina la intensidad de cada atributo de sabor.¹¹

4.5.2 Pruebas estadísticas paramétricas consideradas en la obtención de criterios para la selección de candidatos a jueces sensoriales al utilizar información de tipo intervalar

4.5.2.1 Media

Es definida como el promedio aritmético de los datos asignados a la muestra.¹⁵

$$\bar{x} = \frac{\sum x}{n}$$

Donde:

$\sum x$ = suma de todos los valores que toma x

n = número total de evaluaciones

\bar{x} = símbolo de media de una muestra

4.5.2.2 Desviación estándar muestral

Es la medida de la dispersión de los valores alrededor de la media de una muestra. Se expresa como σ .¹⁵

$$\sigma = \sqrt{\sum_{n=1} (x - \bar{x})^2}$$

4.5.2.3 Coeficiente de variación.

El coeficiente de variación también se conoce como error estándar de la media y se define como la relación entre la desviación estándar y la media, multiplicado por cien.¹⁵ En términos coloquiales expresa el porcentaje de variabilidad que obtuvo un determinado juez en el promedio de sus evaluaciones.

$$CV = \frac{\sigma}{\bar{x}} * 100$$

4.5.2.4 t de student.

Aplicaciones: esta prueba indica si las diferencias encontradas al comparar un par de muestras pueden declararse como significativas o importantes con un cierto nivel de confiabilidad que lo que se esta afirmando es verdadero.¹⁶

Condiciones:

1. Los datos deben ser independientes entre sí, es decir, los datos deben ser obtenidos de manera individual y sin influencia de otros juicios.
2. Las observaciones deben hacerse en poblaciones con distribución normal.
3. Las poblaciones deben tener la misma varianza o una proporción de varianza conocida.
4. Las mediciones deben provenir de una escala de intervalo.

Tipos de t student.

- Dos muestras:

- Dependientes: ejemplo, cuando las muestras han sido analizadas por un mismo juez o si el número de evaluaciones son iguales.^{16,18}
- Independientes: ejemplo, muestras evaluadas por diferentes grupos de jueces en diferentes condiciones.^{16,18}

- Una muestra: pretende determinar si la muestra con una media dada proviene de una población con una media conocida, donde se desea conocer si la media de dicha muestra es significativamente diferente de la media de la población¹⁸ y emplea la siguiente fórmula:

$$t = \frac{\bar{x} - \mu}{s_x} \qquad s_x = s\sqrt{n}$$

Donde:

t = t student calculada

\bar{x} = Media de la muestra (juez)

μ = Media de la población (grupo)

s_x = Error estándar de la muestra

s = desviación estándar de la muestra

n = Número total de observaciones.

En esta investigación se propone emplear la fórmula de t de student para una muestra ya que se quiere determinar si la media o el comportamiento de un juez es significativamente diferente a la del promedio del grupo de jueces.

4.5.2.5 Análisis de varianza (ANOVA)

Aplicaciones: permite estudiar si existe diferencia significativa entre la media de las calificaciones asignadas a más de dos muestras y tiene varios niveles:

1. Una vía donde se explica la diferencia a través de una variable del estudio.
2. Dos vías donde se explica la diferencia dentro de dos variables del estudio, ejemplo: muestras y jueces.
3. Tres vías donde se explica la diferencia en tres variables del estudio, ejemplo: muestras, jueces y repeticiones.^{16,18}

Condiciones para utilizar el análisis de varianza:²²

1. Los datos deben ser independientes entre sí, es decir son juicios individuales y sin influencia de los demás.
2. Las observaciones deben hacerse en poblaciones con distribución normal.
3. Las poblaciones deben tener la misma varianza o una proporción de varianza conocida.
4. Las mediciones deben provenir de una escala de intervalo.

A continuación se explica el análisis de varianza de dos vías mismo que fue empleado en esta investigación debido a que se busca conocer las diferencias entre muestras evaluadas y réplicas efectuadas por un candidato a juez sensorial.

Tabla 4-16 Cuadro de un análisis de varianza de dos vías(fórmulas)

Fuente de variación	Grados de libertad	Suma de cuadrados	Media de cuadrados	F
Productos (p)	p-1	SC _p	SC _p /p-1	MC _p /MC _e
Replicas(r)	r-1	SC _r	SC _r /r-1	MC _r /MC _e
Error (e)	(p-1)(r-1)	Sc _e	SC _e /(p-1)(r-1)	
Total	(pr-1)	SC _t		

p = total de productos r = total de replicas

Si F calculada para las muestras es mayor que la F de Tablas calculada con (p-1) y (r-1)(p-1) grados de libertad, se rechaza Ho: no hay diferencia significativa entre los productos.

Si F calculada para jueces es mayor que de la F de Tablas calculada con (p-1) y (r-1)(p-1) grados de libertad se rechaza Ho: no hay diferencia significativa entre réplicas.⁵

4.5.3 Definición de criterios con fundamentos estadísticos para seleccionar candidatos a jueces sensoriales en pruebas sensoriales de intensidad.

1. **Problema:** Determinar si el candidato a juez puede discriminar entre la intensidad de varios productos sin afectarse por las replicas.
2. **Planteamiento de Hipótesis:**
Hipótesis nula (Ho): no encuentra diferencia significativa entre la intensidad de varios productos
Hipótesis alternativa (Ha): encuentra diferencia significativa entre la intensidad de varios productos.
3. **Selección de la prueba estadística:** Análisis de varianza de dos vías para determinar si el juez encuentra diferencia entre la intensidad de las muestras (primer vía) y si se observa diferencia entre replicas (segunda vía).
4. **Establecimiento del nivel de significancia:** Debido al grado de confiabilidad requerido para este estudio se consideró un nivel de significancia de 5%.
5. **Definición de la zona de rechazo:** Cuando Fo calculada sea mayor o igual que la F teórica para $\alpha=5\%$ o $p \leq 0.05$

6. Calcular con los datos del candidato a juez, y de acuerdo con el análisis de varianza , si se acepta o rechaza la hipótesis nula.

Tabla 4-17 Prueba de intervalo¹, atributo Color, juez 1

Réplicas					
producto	1	2	3	Total renglón	yi. ² /n
1	8.92	7.00	5.88	21.8	475.24
2	5.00	7.00	5.80	17.8	316.84
3	3.00	1.00	4.36	8.36	69.89
4	6.92	8.92	7.72	23.56	555.07
5	1.00	5.00	1.48	7.48	55.95
Total columna	24.84	28.92	25.24	79	491.00
SC total					98.97
SC producto					74.93
SC replica					2.02
SC error					22.01
g.l producto					4
g.l replicas					2
g.l error					8
g.l total					14
MC producto					18.73
MC replica					1.01
MC error					2.75
Fo productos					6.81
Fo replicas					0.37
Fo productos teórica $\alpha=0.05, g.l. 4/8$					3.84
Fo replicas teórica $\alpha =0.05, g.l. 2/8$					4.46

Tabla 4-18 Ejemplo de datos del Análisis de varianza para el juez 1, prueba de intervalo atributo color

Fuente de variación	Grados de libertad	Suma de cuadrados	Media de cuadrados	F calculada
Productos	4	74.93	18.73	6.81
Replicas	2	2.02	1.01	0.37
Error	8	22.01	2.75	
Total	14	98.97		

¹ Ver apéndice 3 para ejemplo de cálculos

El valor de F calculada para productos del juez 1 es mayor que el valor teórico 3.84¹, comparación que implica rechazar la hipótesis nula y aceptar la Hipótesis alternativa en donde se afirma que existe una diferencia significativa entre las intensidades de los productos. Para las replicas se obtuvo un valor calculado de 0.37, mismo que es menor que el valor teórico 4.46, lo que indica que no se observa diferencia significativa entre las replicas, es decir, hay congruencia o buena repetibilidad en las calificaciones que da el juez 1, con esta información se puede generar el siguiente criterio:

Tabla 4-19 Criterio 7 para seleccionar candidatos en pruebas de intensidad utilizando el estadístico Análisis de varianza de dos vías

Valor de F	F teórica productos $\alpha \leq 0.05$; g.l productos /g.l error		F teórica replicas $\alpha \leq 0.05$; g.l replicas /g.l. error	
	F juez(productos) > F teórica	F juez(productos) < F teórica	F juez(replica) > F teórica	F juez(replica) < F teórica
Significado	El juez logra discriminar entre los productos que se le presenten de manera significativa	El juez no discrimina entre productos	El juez no es consistente en sus calificaciones	El juez cumple con lo esperado dando calificaciones semejantes a intensidades semejantes.
Clasificación	Esta en función de su desempeño por familia de atributos	Malo	Malo	Esta en función de su desempeño por familia de atributos

Este criterio permite determinar si el juez tiene la habilidad para discriminar entre las intensidades de varios productos, lo que le hace funcionar como filtro para los demás criterios puesto que si un candidato no cumple con este criterio automáticamente es considerado malo en los siguientes criterios ya que la información no es confiable. En el caso de que el candidato presente la habilidad para discriminar entre intensidades de productos, se busca analizar si tiene tal cualidad para una determinada familia de atributos², por tal razón se determinó el siguiente criterio (ver tabla 4-20). Para el caso de las replicas la clasificación también depende de su desempeño por familia de atributos, solo que en este caso si un juez discrimina entre replicas es acreedor a una reducción en los puntos totales que haya obtenido, lo que le puede generar cambiar de categoría de clasificación, por ejemplo de "Muy Bueno" a "Bueno".

¹ Ver Tabla G. 11 O'Mahony 1986 para $\alpha = 0.05$ y g.l productos/g.l error 4/8

² Familia de atributos: conjunto de atributos que se evalúan con el mismo sentido, por ejemplo: color, viscosidad, fluido, pertenecen a la familia de atributos visuales pues pueden ser evaluados mediante el sentido de la vista.

Tabla 4-20 Criterio 8 para seleccionar candidatos en pruebas de intensidad de acuerdo con su frecuencia de discriminación de atributos por familia utilizando el estadístico Análisis de varianza de dos vías

Discrimina		Replicas		Productos	
Criterio	Fjuez(replica) > F teórica	0-25% de los atributos	25-50% del total de atributos	51-75% de los atributos	76-100% de los atributos
Significado	El juez no es constante en sus calificaciones	El juez tiene un bajo porcentaje de discriminación entre las intensidades que presentan los productos para una determinada familia de atributos.	El juez le cuesta trabajo discriminar las intensidades que presentan los productos en todos los atributos de una familia y lo máximo que alcanza es un 50% del grupo de atributos	El juez discrimina las intensidades que presentan los productos en más de la mitad de los atributos de la familia	El juez logra discriminar las intensidades que presentan los productos en la mayoría de atributos de una determinada familia
Clasificación	No aplica	Malo	Regular	Bueno	Muy Bueno
Puntaje	Menos un punto a la suma total por cada atributo	1	2	3	4

Una vez que se determina si el candidato tiene habilidad para discriminar entre las intensidades de productos este puede considerarse para la segunda parte en donde se evalúa la variabilidad que presenta este candidato con respecto al grupo, misma que se explica a continuación:

4.5.3.1 Comparación de la variabilidad entre los resultados de un candidato a juez con respecto al grupo

1. **Problema:** determinar si la variabilidad de los resultados de un candidato a juez es diferente del grupal. Para analizar la variabilidad se considera al coeficiente de variación del juez y el grupal.
2. **Planteamiento de hipótesis:**
 Hipótesis nula (H₀): Los coeficientes de variación entre el juez y el grupo no son diferentes.
 Hipótesis alternativa (H_a): Los coeficientes de variación entre el juez y el grupo son diferentes significativamente.

3. **Selección de la prueba estadística:** t de student para una sola muestra porque se quiere determinar si las medias del coeficiente de variación del juez y la del grupo son diferentes de manera significativa.
4. **Establecimiento del nivel de significancia:** Debido al grado de confiabilidad requerido para este estudio se consideró un nivel de significancia de 5%.
5. **Definición de la zona de rechazo:** valores calculados mayores al valor teórico, que en este caso es para $\alpha=0.05$ $g.l.=n-1$, donde n es el número total de ensayos, entonces $g.l.=17-1=16$ ver figura 4-3.
6. **Calcular:** valor experimental:

Tabla 4-21 Prueba de intervalo atributo Color juez 1. Obtención de coeficiente de variación promedio

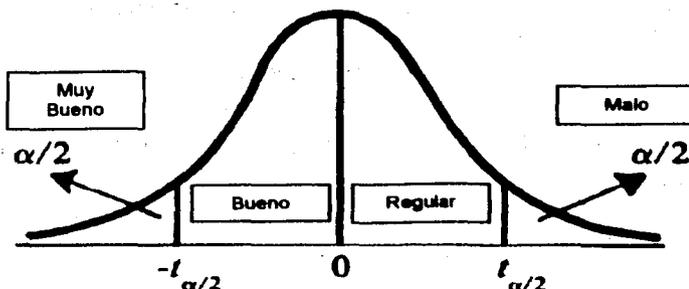
Producto	Replica 1	Replica 2	Replica 3	Total renglón	Promedio	Desviación	C.V
1	8.92	7.00	5.88	21.8	5.64	4.13	73.25
2	5.00	7.00	5.80	17.8	4.67	2.52	53.93
3	3.00	1.00	4.36	8.36	2.33	1.15	49.49
4	6.92	8.92	7.72	23.56	6.61	2.47	37.41
5	1.00	5.00	1.48	7.48	3.67	2.31	62.98
Promedio						2.52	55.41

Prueba t de Student

CV grupo	35.77	t teórico (0.05,16)	2.12
CV juez	55.41	t calculada	1.89
s juez	2.52		

El valor calculado es menor que el teórico, lo que indica que el coeficiente de variación del juez y el grupal no son diferentes significativamente, aceptando la hipótesis nula.

Figura 4-3 Clasificación de candidatos a jueces de acuerdo al valor obtenido de *t* de student



Este gráfico permite plantear el siguiente criterio :

Tabla 4-22 Criterio 9 Clasificación de candidatos jueces en pruebas de intensidad considerando los coeficientes de variación grupal e individual, comparándolos con el estadístico *t* de student

Valor de <i>t</i> Student	No aplica	Valor de <i>t</i> calculada mayor que el valor teórico	Valor de <i>t</i> calculada entre cero y el valor teórico	Valor de <i>t</i> calculada entre el negativo del valor teórico y cero	Valor de <i>t</i> calculada menor que el negativo del Valor teórico
Significado	Jueces con faltas	Coefficiente de variación mucho mayor al grupal.	Coefficiente de variación similar al grupal, pero por arriba del promedio grupal	Coefficiente de variación similar al grupal, pero por debajo del promedio grupal	Coefficientes de variación mucho menores al grupal
Clasificación	Faltas	Malo	Regular	Bueno	Muy Bueno
Puntaje	0	1	2	3	4

Por último se consideran los criterios de frecuencia de discriminación y coeficiente de variación y se crea un criterio final que es la suma de puntos de ambos.

Tabla 4-23 Criterio 10 Selección de jueces en pruebas de intensidad unificando los criterios basados en el coeficiente de variación y en la frecuencia de discriminación

Clasificación	Faltas	Malo	Regular	Bueno	Muy bueno
Criterio 8	0	1	2	3	4
Criterio 9	0	1	2	3	4
Puntaje Total	0	1-2	3-4	5-6	7-8

Ejemplo: ¿Cómo utilizar los criterios propuestos para pruebas sensoriales de intensidad?

Recordando la prueba de intensidad que realizó el candidato a juez 1 se tiene que:

Tabla 4-24 Tabla de datos del análisis de varianza para el juez 1 prueba de intervalo atributo color

Fuente de variación	Grados de libertad	Suma de cuadrados	Media de cuadrados	F calculada
Productos	4	74.93	18.73	6.81
Replicas	2	2.02	1.01	0.37
Error	8	22.01	2.75	
Total	14	98.97		

Con estos resultados el juez logra pasar el primer filtro (Criterio 7) para ser considerado en la clasificación final para pruebas de intervalo, puesto que puede discriminar entre las intensidades de varios productos. Al observar el resultado de las replicas de acuerdo al mismo criterio se determina que es congruente en sus calificaciones puesto que no encuentra diferencia significativa entre ellas.

Analizando la variabilidad de sus resultados con respecto a la del grupo se tiene:

Prueba t de Student

CV grupo	35.77
CV promedio juez	55.41
S juez	2.52
Valor t teórico (0.05, 16)	2.12
Valor de t calculada	-1.89

Con este valor obtenido de t de Student para el juez 1, este candidato clasifica como "Bueno" y es acreedor de acuerdo al Criterio 9, a 3 puntos. Para definir la frecuencia de discriminación (Criterio 8) se considera a toda la familia de atributos (en este caso es familia visual) en donde el candidato 1 discriminó entre la intensidad para 8 atributos de un total de 11, lo que lo clasifica como "Muy bueno" con cuatro puntos. Considerando estos resultados obtenidos para el juez 1, su total de puntos es de 3+4 =7, y de acuerdo con el Criterio 10 se le clasificaría como "Muy Bueno".

5 METODOLOGÍA

1. Se clasifica la información contenida en las distintas bases de datos (resultados de los ensayos sensoriales realizados) de acuerdo al tipo de número y de estadística.

2. Aplicación de los criterios en los resultados de los análisis de cada una de las bases de datos¹.

5.1 APLICACIÓN DE CRITERIOS PARA SELECCIONAR CANDIDATOS A JUECES SENSORIALES EN PRUEBAS DISCRIMINATIVAS (INFORMACIÓN NOMINAL)

1.- De todas las bases de datos suministradas por Penta Sensorial, se seleccionaron aquellas que contenían información nominal y se dividió en identificación de atributos y comparación por pares. Además se observó que esta base de datos contenía un total de 8 ensayos por cada gusto básico de cada juez. Esto estaba dividido en dos sesiones de 4 replicas.

2.- Se calcula el porcentaje de reconocimiento de gustos básicos para cada juez por cada sesión y se obtiene el promedio de los 8 ensayos. Para esa primera parte se aplica el siguiente criterio:

Tabla 5-1 Aplicación del Criterio 1 para la selección de jueces en pruebas discriminativas basado en el porcentaje de identificación de gustos básicos.

% de identificación	No aplica	0-25%	26-50%	51-75%	76-100%
Criterio para la identificación de atributos con un total de 8 ensayos	No aplica	0 a 2 aciertos	3 a 4 aciertos por cada atributo	5 a 6 aciertos por cada atributo	8 a 7 aciertos por cada atributo
Clasificación	Faltas	Malo	Regular	Bueno	Muy bueno
Puntaje	0	1	2	3	4

3.- De esta manera se ordenan los resultados de los porcentajes de cada juez en las categorías propuestas en el criterio anterior.

4.- Se calculan los puntos que cada juez obtiene en total para poder clasificarlo, de acuerdo con lo anterior los máximos son establecidos en relación al número de veces que un juez este dentro de la misma categoría, por ejemplo, si un juez siempre

¹ Todo los cálculos fueron realizados con el software Microsoft Excel.

se ubica en la categoría "Muy bueno", obtendrá el máximo de puntos que son 16, porque tenemos 4 gustos básicos y la categoría "Muy bueno" vale 4 puntos ($4 \times 4 = 16$).

5.- Una vez determinado el porcentaje de identificación de gustos básicos, se aplica el Criterio 2 para establecer si el candidato cumple satisfactoriamente con una prueba discriminativa.

Tabla 5-2 Aplicación del Criterio 2 selección de jueces para pruebas discriminativas basado en el estadístico Ji Cuadrada

Valor de Ji cuadrada	No aplica	No aplica	1 Unidad Menor al valor crítico de Tablas	Mínimo de juicios correctos para ser diferente y significativa	Arriba del valor mínimo y hasta el máximo de aciertos
Aciertos	No aplica	0-5	6	7	8
$\chi^2_{Teórica} = 2.71$	No aplica	0.13	1.13	3.13	6.13
Clasificación	Faltas	Malo	Regular	Bueno	Muy bueno
Puntaje	0	1	2	3	4

El número de aciertos se consideró para este caso en específico para 8 juicios Ver tabla F.1 en Pedrero y Pangborn (1989)

6.- Se aplica el tercer criterio que considera a los dos anteriores, el cual implica sumar los puntajes obtenidos en los Criterios 1 y 2 quedando de la siguiente manera:

Tabla 5-3 Puntaje para el Criterio 3 de selección de pruebas discriminativas

Clasificación	Faltas	Malo	Regular	Bueno	Muy Bueno
Puntaje	0	2	4	6	8
Puntaje para la prueba realizada	0	8	16	24	32
Puntos intermedios	0	1-8	9-16	17-24	25-32

El puntaje total para la prueba realizada se obtiene al multiplicar el total de atributos por cada categoría, esto es que si se tiene un total de 4 gustos básicos cada categoría se multiplica por cuatro. Ejemplo: para obtener el puntaje total en la categoría de "Muy Bueno" se multiplicó 8 que es total de puntos por cuatro atributos, lo que da

como resultado 32 puntos. Posteriormente se establecen intervalos como los que se observan en la última fila de esta tabla.

5.2 APLICACIÓN DE CRITERIOS PARA SELECCIONAR CANDIDATOS A JUECES SENSORIALES EN PRUEBAS DE ORDENACIÓN (INFORMACIÓN ORDINAL)

1.- En esta etapa se aplican 3 métodos estadísticos para determinar cuál es el más adecuado en una clasificación de jueces en pruebas de ordenación. Los métodos empleados fueron Page Test, Ordenamiento por Rangos Newell-MacFarlane y Prueba de Friedman. De ellos se elige aquel que cumpla con las características de selección que cada analista sensorial necesite.

Tabla 5-4 Ejemplo Evaluación del orden para el atributo color

Número de juez	Page test	Ordenamiento por rangos de Newell-MacFarlane	Prueba de Friedman
Valores críticos	$L = 197$ ¹⁰ Tabla G25	13 puntos de diferencia entre cada muestra ⁷ Tabla G1	$\chi^2 = 9.49$ ¹⁰ Tabla G7
1	202 *	Menos de 13 en todas las muestras	7.4
2	211 *	13 puntos solo en un par de muestras	10.6 *
3	203 *	13 puntos solo en un par de muestras	6.5
4	209 *	13 puntos solo en un par de muestras	9.7 *
5	200 *	13 puntos solo en un par de muestras	2.7
6	202 *	Menos de 13 en todas las muestras	7.8
7	201 *	15 puntos entre un par de muestras	4.3
8	204 *	Menos de 13 en todas las muestras	6.43
9	206 *	14 puntos en un par de muestras	10.6 *
10	213 *	Menos de 13 en todas las muestras	11.8 *
11	Faltas	Faltas	Faltas
12	212 *	Menos de 13 en todas las muestras	12.8 *
13	200 *	Menos de 13 en todas las muestras	6.4
14	215 *	13 puntos solo en un par de muestras	13 *
15	207 *	Menos de 13 en todas las muestras	8.2
16	208 *	Menos de 13 en todas las muestras	8.8
17	209 *	13 puntos solo en un par de muestras	13 *

Los resultados marcados con asterisco indican que el juez cumple satisfactoriamente con esa prueba estadística. Como puede observarse los métodos Page Test y Ordenamiento por rangos de Newell-MacFarlane no permiten hacer una diferenciación entre los candidatos a jueces ya que el primero indica que todos los candidatos obtienen valores calculados mayores al crítico, indicando que si ordenan entre varias muestras, mientras que en el segundo método ningún juez cumple con él, ya que los valores obtenidos son menores al crítico, entonces no encuentran diferencia al ordenar entre muestras. Por esta razón se utiliza la prueba de Friedman para determinar los criterios que apoyen la selección de candidatos a jueces en pruebas de ordenación.

2.- Se evalúan los resultados de las pruebas de ordenación para cada candidato utilizando el método Prueba de Friedman. Se aplica el siguiente criterio:

Tabla 5-5 Aplicación del Criterio 4 selección de jueces para pruebas ordenación basado en el estadístico de Friedman

Valor de Friedman Límites	Negativo por faltas	Negativos	0 - 9.49	9.49 - 11.66	11.67-13.83	13.84-16
Clasificación	Faltas y Negativos		Malo	Regular	Bueno	Muy bueno
Puntaje	0		1	2	3	4

Los límites del valor de Friedman se calcularon de la siguiente manera:

Valor teórico: $9.49 \alpha=0.05$ con $g.l=4$

Valor máximo :16 (ver tema 4.4)

Diferencia entre el máximo y el mínimo: $16 - 9.49 = 6.51$

Divido entre 3 = 2.17

Valor para una tercera parte del máximo: $9.49 + 2.17 = 11.66$

Valor para dos terceras partes: $11.66 + 2.17 = 13.84$

Total: $13.84 + 2.17 = 16$

3.- Se revisan los valores de suma de rangos que cada juez calificó en cada una de las muestras que ordenaron, con el fin de verificar que no haya repetido valores. Lo que significa que este juez no sólo encuentra diferencia significativa al ordenar de acuerdo al criterio anterior sino que también lo hace de manera discriminada al no repetir valores de suma de rangos para cada una de las muestras.

Tabla 5-6 Aplicación del Criterio 5 Selección de jueces considerando el número de muestras que el candidato a juez ordena de sin repetir valores de suma de rangos.

Muestras sin repetir valores de suma de rangos	Faltas	Negativos	No aplica, porque no cumple con el Criterio 4	60%	80%	100%
Límites	No aplica	No aplica	No aplica	3	4	5
Clasificación	Faltas	Negativos	Malo	Regular	Bueno	Muy bueno
Puntaje	0	0	1	2	3	4

4.- Se considera un tercer criterio que involucre a los dos anteriores, es decir, se suman los puntajes obtenidos en ambos criterios, quedando de la siguiente manera:

Tabla 5-7 Puntaje para el Criterio 6 de selección por pruebas de ordenación

Clasificación	Faltas y negativos	Malo	Regular	Bueno	Muy bueno
Suma de puntos	0	2	4	6	8
Total de ensayos 11	0	22	44	66	88
Límites	0	1-22	23-44	45-66	67-88

Esta tabla indica los puntajes requeridos de acuerdo al Criterio 6 para cada categoría. En la base de datos con información ordinal se observó que se ordenaron un total de 11 atributos, entonces para generar los límites de cada categoría aplicado al número de ensayos se multiplicó 11 por el valor de cada categoría. También se establecieron los límites que se indican al final de la Tabla 5-7.

5.3 APLICACIÓN DE CRITERIOS PARA SELECCIONAR CANDIDATOS A JUECES SENSORIALES EN PRUEBAS DE INTENSIDAD (INFORMACIÓN INTERVALAR)

Se partió de una base de datos para yogurt de fresa con 5 productos, y se obtuvieron 43 atributos evaluados por 17 candidatos a jueces sensoriales.

Los 43 atributos se dividen en familias de atributos teniendo así la clasificación dada en la Tabla 5-8:

Tabla 5-8 Clasificación de atributos por familias

Visuales	Aroma general	Aromáticos	Gustos básicos	Textura
Firmeza inicial Cantidad de fruta Grosor Color Firmeza final Viscosidad Elasticidad Gomosidad Recubrimiento Suave Grumoso	Aroma general Aroma a lácteo Aroma a fruta	Frutal Queso Diacetilo Vainilla Caramelo Quemado Rojo Amilico Tutti-frutti Antranilato Almendra Canela Maduro Cáscara Verde Cremoso Grasa Medicinal Cocido Plástico Tostado	Dulce Ácido Salado Amargo	Consistencia Recubrimiento Graso Blando

Se realizó un Análisis de varianza de dos vías, donde se evalúa discriminación entre productos y entre replicas. Esto se hizo para cada juez y para cada atributo, de esta manera, es posible clasificar a un juez considerando:

- discriminación entre productos
- no discriminar entre replicas
- tener un mínimo de asistencias mayor al 90%, de lo contrario el juez automáticamente es enviado a la categoría de malo.

Tabla 5-9 Ejemplo del análisis de varianza para el Atributo aroma general, familia aroma general, juez 10

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F
Productos	4	38.56	9.64	28.04 ($F_{1;0.05,4,8}=3.84$)
Replicas	2	0.08	0.04	0.11 ($F_{1;0.05,2,8}=4.46$)
Error	8	2.75	0.34	
Total	14	41.39		
Total	14	41.39		

Si el juez encuentra diferencia significativa entre productos, se le considera apto para que realice las siguientes pruebas sensoriales. En el caso que el desempeño (replicas) del juez sea "Malo" o bajo (se observa diferencia significativa entre las

replicas) se le resta un punto por cada atributo en donde haya diferencia significativa entre replicas.¹

Tabla 5-10 Aplicación de Criterio 7 para seleccionar candidatos en pruebas de intensidad utilizando el estadístico Análisis de varianza de dos vías

Valor de F	$F_{\text{teórica productos } \alpha 0.05} = 3.84$		$F_{\text{teórica replicas } \alpha 0.05} = 4.46$	
Criterio	F juez(productos) > F teórica	F juez(productos) < F teórica	F juez(replica) > F teórica	F juez(replica) < F teórica
Significado	El juez logra discriminar de manera significativa entre los productos que se le presentan	El juez no discrimina entre productos	El juez no es consistente en sus calificaciones	El juez cumple con lo esperado dando calificaciones semejantes a intensidades semejantes.
Clasificación	Ver Tabla 5-11	Malo	Malo	Ver Tabla 5-11

Se crea una clasificación en base al Criterio 8 (frecuencia de discriminación de intensidades por familia de atributos), en este caso se utilizan porcentajes de discriminación tomando en cuenta el total de atributos que el juez pudo diferenciar por familia. Se manejará el ejemplo de "Aroma general" donde son 3 atributos, pero lo mismo se aplica para las demás familias de atributos. Tenemos entonces 3 atributos que equivalen al 100% y se dividen los porcentajes en 4 categorías para tener la siguiente clasificación:

Tabla 5-11 Aplicación del Criterio 8 para seleccionar candidatos en pruebas de intensidad de acuerdo con su frecuencia de discriminación de atributos por familia

Discrimina	Replicas	Productos			
Criterio	F juez(replica) > F teórica	0-25% de los atributos	25-50% del total de atributos	51-75% de los atributos	76-100% de los atributos
Clasificación	No aplica	Malo	Regular	Bueno	Muy Bueno
Puntaje	Restar un punto a la suma total por cada atributo en el que sus replicas no fueron constantes	1	2	3	4

¹ Ver apéndice 3

Una vez que el juez (muestra) pasó el punto anterior, se obtuvo el promedio para cada uno de los productos que había evaluado, así mismo se calculó la desviación estándar y con estos datos se pudo obtener el coeficiente de variación promedio para cada juez. Se obtuvieron los 17 coeficientes de variación promedio correspondientes a cada juez y se calculó el coeficiente de variación grupal.

Una vez que se tuvo la desviación estándar y el coeficiente de variación promedio por juez se comparó con el grupal por medio de la prueba estadística t de Student. Para esta investigación se utilizó una t de Student con un nivel de significancia (α) de 0.05, el cual no es muy estricto ya que no son jueces entrenados. Los grados de libertad se obtuvieron del total de nuestro grupo de jueces menos uno entonces: $17-1=16$ y dos colas ya que se espera que haya jueces por debajo del coeficiente de variación grupal y otros por arriba del grupal. La t de student teórica fue de 2.12 con el que se creo la siguiente tabla de clasificación:

Tabla 5-12 Aplicación del Criterio 9 para clasificar jueces en pruebas de intensidad considerando los coeficientes de variación grupal e individual, comparándolos con el estadístico t de student

Valor de t de Student	No aplica	$t_{cal} > 2.12$	$0 < t_{cal} < 2.12$	$-2.12 < t_{cal} < 0$	$t_{cal} < -2.12$
Clasificación	Faltas	Malo	Bueno	Regular	Muy Bueno
Puntaje	0	1	3	2	4

Para obtener la clasificación final de jueces en pruebas de intensidad para una determinada familia de atributos se calcula el promedio obtenido en los Criterios 8 y 9. Aplicando el Criterio 10 los puntajes quedan de la siguiente manera:

Tabla 5-13 Aplicación del Criterio 10 de selección de candidatos en pruebas intensidad

Clasificación	Faltas	Malo	Regular	Bueno	Muy bueno
Suma de puntos	0	2	4	6	8
Puntaje total (ejemplo familia de atributos aromáticos)	0	44	84	126	168

El puntaje total se obtiene al multiplicar el número de atributos en una familia por el valor de cada clasificación con el fin de obtener los puntos límite para una determinada familia de atributos.

5.4 CLASIFICACIÓN FINAL DE LOS CANDIDATOS A JUECES SENSORIALES.

Una vez que se determinaron los criterios para cada una de las bases de datos (nominal, ordinal e intervalar) estos se aplican sobre los resultados de los ensayos realizados.

Se elabora un resumen con todas las calificaciones obtenidas en cada una de las pruebas y se obtiene un promedio para determinar el desempeño del candidato a juez sensorial y su posible selección para la fase de entrenamiento.

6 RESULTADOS

6.1 SELECCIÓN DE CANDIDATOS A JUECES SENSORIALES APLICANDO LOS CRITERIOS CON FUNDAMENTOS ESTADÍSTICOS PARA DATOS DE TIPO NOMINAL EN PRUEBAS SENSORIALES DISCRIMINATIVAS

6.1.1 Clasificación de candidatos a jueces sensoriales en pruebas de identificación de gustos básicos

La parte de identificación de gustos básicos fue realizada mediante la prueba comparación por pares, en donde se le presenta al juez una muestra con agua pura y otra con el gusto básico, se realizaron dos repeticiones por cada uno de los 17 jueces. El candidato de esta manera puede observar el porcentaje de identificación, traducido como el número de aciertos y analizado de acuerdo al estadístico Ji cuadrada.

Tabla 6-1 Jueces seleccionados de acuerdo con el Criterio 1 porcentaje de identificación de gustos básicos

Clasificación	Faltas (0)	Malo (1)	Regular (2)	Bueno (3)	Muy bueno (4)
% Reconocimiento	No aplica	0-25%	26-50%	51-75%	76-100%
Puntaje máximo	0	4	8	12	16
Atributos	Jueces clasificados				
Ácido	3		11,17	1,2,4,6,7,8,12,13,14,15	5,9,10,16
Amargo	3		6,14,15	4,9,11,13,17	1,2,5,7,8,10,12,16
Dulce	3		7,10,11,15,17	1,2,4,6,9,12,13	5,8,14,16
Salado		17	6,11,13	4,7,14	1,2,5,8,9,10,12,16

En la Tabla 6-1 se indica en la primera fila la clasificación de los jueces en donde el número en paréntesis indica el valor de cada categoría. La segunda fila representa los límites del Criterio 1 para seleccionar candidatos a jueces en pruebas discriminativas. La tercera fila representa el puntaje máximo que un juez puede obtener al estar en una determinada categoría. Los jueces se ubican en cada una de las categorías y se indican en el centro de la tabla. De esta manera se logra identificar a los jueces que cumplen los distintos criterios de clasificación para los gustos básicos.

En la gráfica 6-1 se indica la clasificación de jueces conforme el Criterio 1, para lo cual los puntos totales se obtienen de la clasificación generada en la Tabla 7-1 en

donde se ubica a los jueces en distintas categorías según los resultados del porcentaje de identificación. Por ejemplo el juez 16 es considerado como "Muy bueno" por identificar los cuatro gustos básicos, entonces su puntaje total es de 16 puntos (4 atributos por 4 puntos de la categoría "Muy bueno" =16). En la leyenda de la gráfica se indican los puntajes críticos necesarios para cada categoría y enseguida el porcentaje de identificación en base al Criterio 1.

Gráfica 6-1 Selección de candidatos en base al Criterio 1 porcentaje de identificación de gustos básicos

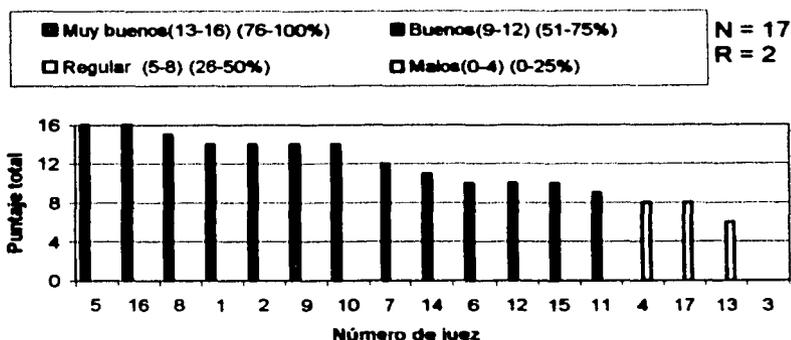


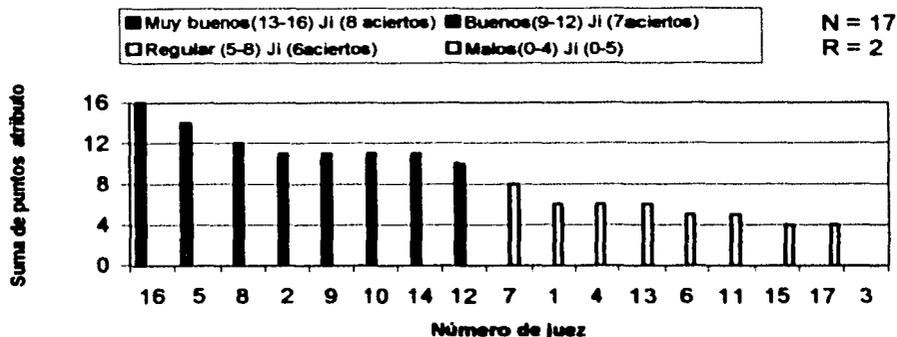
Tabla 6-2 Clasificación de candidatos a jueces por atributo y número de aciertos con base al estadístico Ji cuadrada

Clasificación	Faltas	Malo (1)	Regular (2)	Bueno (1)	Muy bueno (4)
Criterio Ji		χ^2 (0-5 aciertos)	χ^2 (6 aciertos)	χ^2 (7 aciertos)	χ^2 (8 aciertos)
Puntaje máximo	0	4	8	12	16
Atributo	Jueces clasificados				
Ácido	3	1,11,12,14,15,17	2,4,6,7,8,13	5,10	9,16
Amargo	3	4,6,15,17	9,11,13,14	1,5,7,8	2,10,12,16
Dulce	3	1,4,6,7,10,11,13,15,17	2,9,12	14	5,8,16
Salado	3	6,7,11,13,14,15,17	4	1,2,8,9,10,12	5,16

La primera fila de la tabla 6-2 indica el criterio basado en el número de aciertos de acuerdo a Tablas críticas de Ji cuadrada. La tabla muestra los jueces seleccionados en las diferentes categorías por gustos básicos, muestra también los jueces que no asistieron a la sesión los cuales no son acreedores a puntos. En el centro

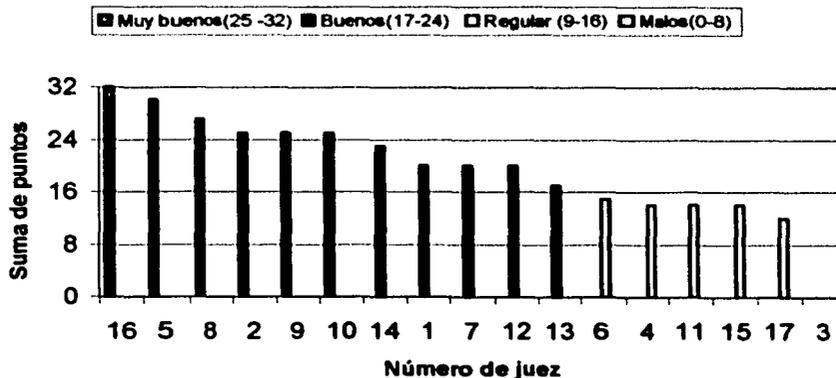
se indica la clasificación que cada juez obtuvo en esta prueba. El puntaje máximo esta referido al valor de cada categoría por el total de atributos evaluados.

Gráfica 6-2 Criterio basado en el estadístico Ji cuadrada para selección de jueces en pruebas de identificación de atributos



Descripción: está gráfica presenta un condensado de los puntos que cada juez obtuvo en sus diferentes categorías para identificación por gustos básicos. Debido a que identificaron 4 gustos básicos lo máximo que un juez puede obtener son 16 puntos. (4 atributos multiplicados por 4 puntos que vale la categoría "Muy bueno") En estas condiciones se observa que el juez 16 es el único que obtiene el máximo de puntos, por el contrario el juez 17 clasifico en los 4 gustos básicos como malo, lo cual solo le genera 4 puntos totales.

Gráfica 6-3 Clasificación final de jueces para pruebas de identificación de gustos básicos en base al Criterio 3 (suma de puntos de Criterios 1 y 2)



Descripción: Esta gráfica muestra al Criterio 3 que agrupa al Criterio 2 (Ji cuadrada) y al Criterio 1 (porcentaje de reconocimiento de gustos básicos). Los límites para este criterio son los que se encuentran entre paréntesis al lado de cada categoría. Los puntos son obtenidos del máximo para cada criterio uno de los criterios anteriores. Por ejemplo lo máximo para Ji cuadrada y para % de reconocimiento son 16 puntos, que al sumarlos nos da un total de 32.

6.1.2 Clasificación de candidatos a jueces sensoriales en pruebas de comparación por pares

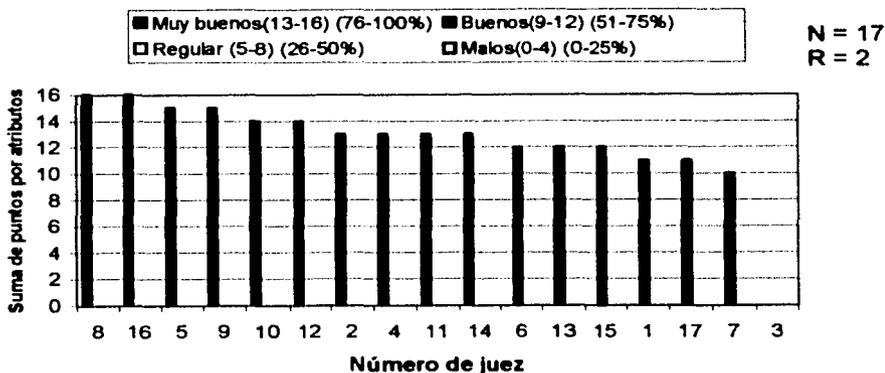
La prueba de comparación por pares consistió no solo en identificar el gusto básico, sino también discriminar entre su concentración. Para esta parte se realizaron 4 ensayos con 2 réplicas por cada gusto básico.

Tabla 6-3 Jueces clasificados para pruebas de comparación por pares de acuerdo al Criterio 1 basado en el Porcentaje de identificación de gustos básicos.

Clasificación	Faltas (0)	Malo (1)	Regular (2)	Buena (3)	Muy buena (4)
% Reconocimiento	No aplica	0-25%	26-50%	51-75%	76-100%
Puntaje máximo	0	4	8	12	16
Atributo	Jueces clasificados				
Ácido	3		1,2,6,13,14,17	4,7,12,15	5,8,9,10,11,16
Amargo	3			4,6,7	1,2,5,8,9,10,11,12,13,14,15,16,17
Dulce	3		1,7,10,11,15,17	2,4,5,9,12	6,8,13,14,16
Salado	3		7,13	1,6,11,14,15,17	2,4,5,8,9,10,12,16

Descripción de la Tabla 6-3 muestra a los jueces seleccionados de acuerdo con el porcentaje de identificación de gustos básicos. Permite ubicar a cada juez en una de las categorías y asignarle un valor general para esta prueba, mismo que se muestra en la gráfica 6-4

Gráfica 6-4 Selección de candidatos en base al Criterio 1 porcentaje de identificación de gustos básicos en pruebas de comparación por pares



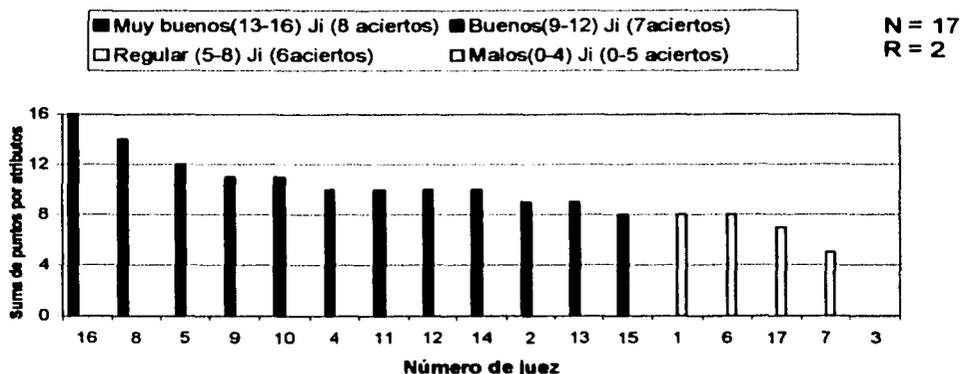
La gráfica 6-4 indica la clasificación de jueces conforme el Criterio 1, los puntos totales se obtienen de la clasificación generada en la Tabla 6-3 en donde se ubica al juez por categorías en los resultados del porcentaje de identificación. Por ejemplo el juez 16 es considerado como "Muy bueno" para identificar los cuatro gustos básicos, entonces su puntaje total es de 16 (4 atributos por 4 puntos de la categoría "Muy bueno" =16). En la leyenda de la gráfica se indica los puntajes necesarios para cada categoría y enseguida el porcentaje de identificación en base al Criterio 1.

Tabla 6-4 Clasificación de candidatos a jueces por atributo y número de aciertos de acuerdo al Criterio 2 basado en el estadístico Ji cuadrada para pruebas de comparación por pares

Clasificación	Faltas 0	Malo (1)	Regular (2)	Buena (3)	Muy buena (4)
Criterio		χ^2 (0-5 acier.)	χ^2 (6 acier.)	χ^2 (7 acier.)	χ^2 (8 aciertos)
Atributo	Jueces clasificados				
Ácido	3	1,2,6,7,13,14,17	4,12,15	5,8,9,10,11	16
Amargo	3	6	4,7	2,15	1,5,8,9,10,11,12,13,14,16,17
Dulce	3	1,5,7,9,10,11,15,17	2,4,12	13,14	6,8,16
Salado	3	7,13,17	1,6,11,14,15	2,8,9,10	4,5,12,16

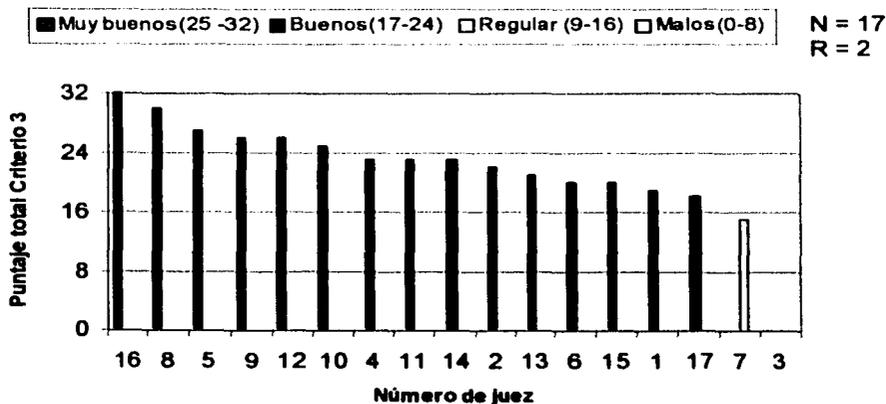
Descripción: esta tabla muestra los jueces seleccionados en las diferentes categorías por gustos básicos, también los jueces que no asistieron a la sesión, los cuales no son acreedores a puntos. La primera fila indica el criterio basado en el número de aciertos de acuerdo a Tablas de Valores Críticos de Ji cuadrada.^{7,10}

Gráfica 6-5 Criterio basado en el estadístico Ji cuadrada para selección de jueces en pruebas de comparación por pares



Descripción: está gráfica presenta un condensado de los puntos que cada juez obtuvo en sus diferentes categorías para identificación de los 4 gustos básicos, por lo tanto lo máximo que un juez puede obtener son 16 puntos (4 atributos* 4 puntos categoría "Muy bueno"), en estas condiciones se observa que el juez 16 es el único que obtiene el máximo de puntos, por otro lado el juez 8 clasifico como "Muy bueno" en dulce y amargo y bueno para ácido y salado lo que le genera una total de $(4*2)+(3*2)=14$ puntos, mismos que lo hacen acreedor a la categoría de "Muy bueno".

Gráfica 6-6 Clasificación final de jueces para pruebas de comparación por pares en base al Criterio 3



Descripción: Esta gráfica muestra al Criterio 3 que agrupa al Criterio 1 porcentaje de reconocimiento de gustos básicos con el Criterio 2 Ji cuadrada. Los límites para este criterio son los que se encuentran entre paréntesis al lado de cada

categoría . Los puntos son obtenidos del máximo para cada criterio uno de los criterios anteriores. Por ejemplo lo máximo para Ji cuadrada y para % de reconocimiento son 16 puntos, que al sumarlos nos da un total de 32.

6.2 SELECCIÓN DE CANDIDATOS A JUECES SENSORIALES APLICANDO LOS CRITERIOS CON FUNDAMENTOS ESTADÍSTICOS PARA DATOS DE TIPO ORDINAL EN PRUEBAS SENSORIALES DE ORDENACIÓN

En esta prueba se ordenaron 5 muestras de yogurt comerciales, considerando 9 atributos y 4 réplicas por cada juez

6.2.1 Clasificación general de candidatos a jueces sensoriales en pruebas de sensoriales de ordenación

Tabla 6-5 Selección de jueces por atributo y por categoría para el Criterio 4 basado en estadístico prueba de Friedman

Clasificación	Valores negativos	Faltas	Malo (1)	Regular (2)	Bueno (3)	Muy bueno (4)
Valor de Friedman			0 - 9.48	9.49 ¹ - 11.66	11.67 - 13.83	13.84 - 16
Puntaje total	0	0	9	18	27	36
Atributo	Jueces Clasificados					
Color	11		1,3,5,6,7,8,13,15,16	2,4,9	10,12,14,17	
Viscosidad general	11		3,4,5,7	14	1,6,8,12,13,17	2,9,10,15,16
Grumoso	11	5	3,4,7,8,10,13,15,17	1,2,9,14	6,12	16
Cantidad de fruta	11		3,4,5,7,9,10,12,13,17		14,15,16	1,2,6,8
Consistencia	11	4,15	3,5,7,8,9,14,17	1,13	2,6,12,16	10
Suave	11	4,5,7,9	3,8,10,14,15	2,6,12,13,16,17	1	
Dulzor	11	5,15	3,4,7,13	2	6,8,10,12,14,16,17	1,9
Acidez	11	5,15	3,4,6,7,8,9,12,13,14	1,10,16,17	2	
Aroma a fresa	11	7	4,5,6,15	3,10,12	2,8,14,16,17	1,9,13

¹ Valor teórico de Friedman que se obtuvo de Tablas Críticas de Ji cuadrada con g.l.=4 y $\alpha=0.05$

Descripción: esta Tabla muestra los resultados de aplicar el Criterio 4 (basado en el estadístico de Friedman), la primera fila representa la clasificación a la cual cada juez es acreedor de acuerdo con la segunda fila que muestra los valores límite para cada clasificación. Los valores de Friedman negativos son producidos por aquellos jueces que no respetaron las instrucciones y que repitieron valores cuando se especifica claramente en las pruebas de ordenación que no se permiten empates. El puntaje total se obtiene al multiplicar el número de atributos evaluados por el valor de cada una de las categorías, por ejemplo 9 atributos por 4 de la categoría "Muy bueno", es igual a 36 puntos, valor máximo que un candidato puede rendir al clasificar en todos los atributos siempre como "Muy bueno". Para obtener un puntaje total se suman las calificaciones parciales de cada atributo, mismo que se representa de manera esquematizada en la siguiente gráfica. Por ejemplo el juez 1 se clasificó tres veces como "Muy bueno" esto le genera $(3 \cdot 4) = 12$, más dos veces en la categoría de "Bueno" $(2 \cdot 3) = 6$, más 3 veces "Regular" $(3 \cdot 2) = 6$ y una vez como "Malo" $(1 \cdot 1) = 1$ nos da un total de 25 puntos; que en base a los puntos totales de cada categoría (número entre paréntesis en la leyenda de la gráfica) es posible ubicar el rendimiento de este juez como "Bueno" en pruebas de ordenación.

Gráfica 6-7 Selección de candidatos para pruebas de ordenación de acuerdo con el Criterio 4 basado en el estadístico de Friedman

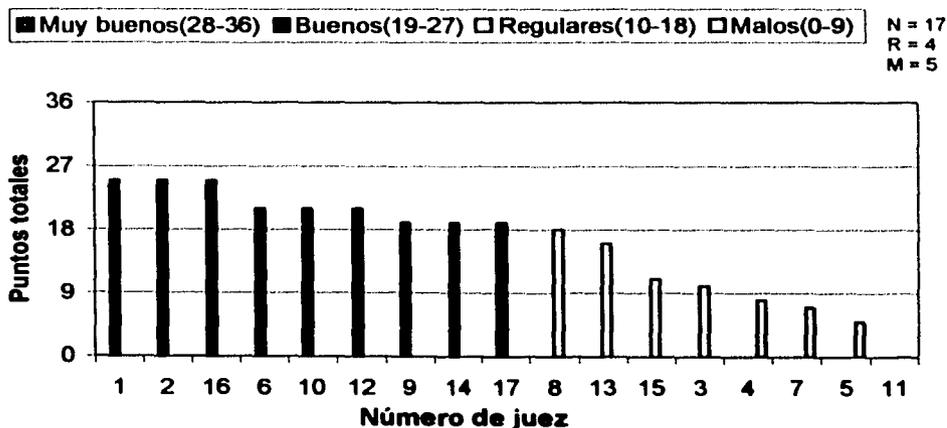


Tabla 6-6 Jueces seleccionados en base al Criterio 5 basado en el número de muestras sin repetir valores de suma de rangos.

Clasificación	Faltas y negativos		Malo (1)	Regular (2)	Bueno (3)	Muy bueno (4)
Puntaje máximo suma de rangos	0		9	18	27	36
Muestras sin repetir valores de suma de rangos	Jueces que faltan u obtuvieron valores negativos.		No aplica	60% 3 muestras	80% 4 muestras	100% 5 muestras
Atributo	Jueces clasificados					
Color	11		1,3,5,6,7,8, 13,15,16	2,4,9	10,12,14,17	
Viscosidad general	11		3,4,5,7	14	1,6,8,12,13, 17	2,9,10,15, 16
Grumoso	11	5	3,4,7,8,10, 13,15,17	1,2,9,14	6,12	16
Cantidad de fruta	11		3,4,5,7,9,10, 12,13,17		14,15,16	1,2,6,8
Consistencia	11	4,15	3,5,7,8,9,14, 17	1,13	2,6,12,16	10
Suave	11	4,5,7,9	3,8,10,14,15	2,6,12,13, 16,17	1	
Dulzor	11	5,15	3,4,7,13	2	6,8,10,12,14, 16,17	1,9
Acidez	11	5,15	3,4,6,7,8,9, 12,13,14	1,10,16,17	2	
Aroma a fresa	11	7	4,5,6,15	3,10,12	2,8,14,16,17	1,9,13

Descripción: esta Tabla representa la clasificación de los jueces al aplicar el Criterio 5 ordenación de muestras sin repetir valores de suma de rangos . Al igual que en la Tabla 7-5 permite ubicar a cada uno de los jueces de acuerdo a la categoría ("Muy bueno", "Bueno" "Regular" o "Malo") que obtuvieron al ordenar en los 11 atributos. En la segunda fila se observan los puntajes máximos, mismos que se muestran en la siguiente gráfica:

Gráfica 6-8 Clasificación de jueces basado en el Criterio 5 ordenación de muestras sin repetir valores promedio.

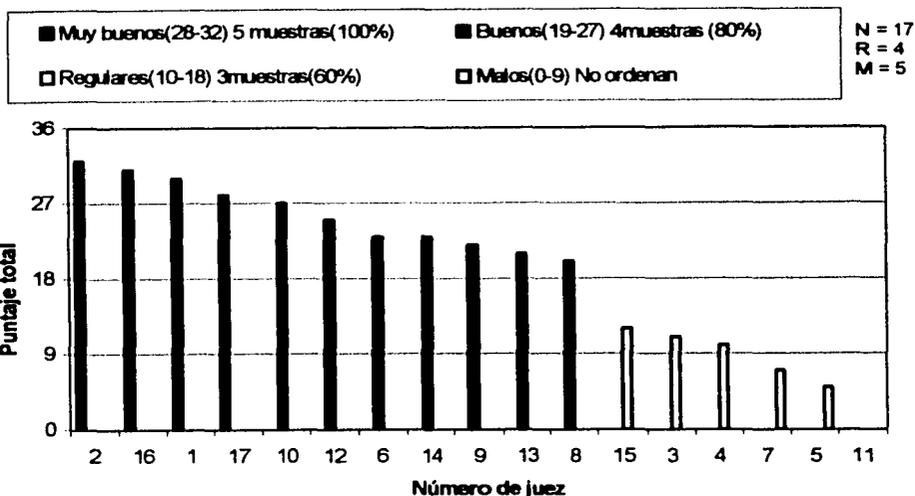


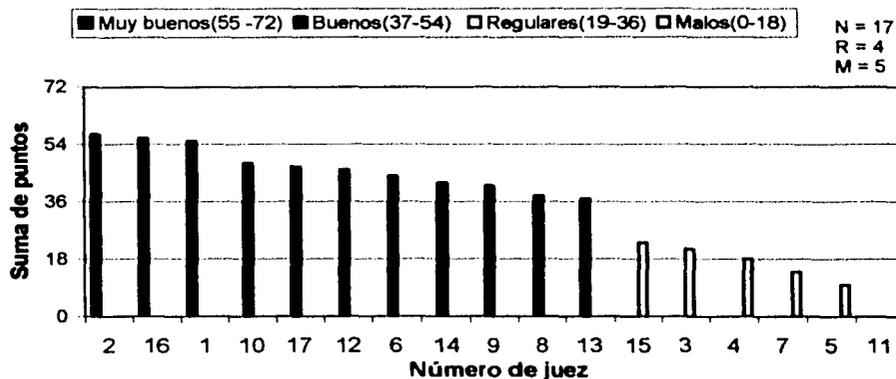
Tabla 6-7 Determinación del puntaje para establecer el Criterio 6 (Suma de puntos de los criterios basado en el estadístico Friedman y en el de porcentaje de ordenación de muestras sin repetir valores de suma de rangos)

Clasificación	Criterio		Valores máximos para cada clasificación		
	Faltas y negativos (0)	Malo(1)	Regular(2)	Bueno(3)	Muy bueno(4)
Puntaje max. Friedman	0	9	18	27	36
Puntaje max. Suma de rangos	0	9	18	27	36
Criterio 6 Suma de puntos	0	18	36	54	72

Descripción: Esta Tabla dá la relación de puntos necesarios para clasificar al juez en una de las cinco categorías propuestas. Para obtener el total de puntos se suman los máximos de los Criterios 4 y 5. Por ejemplo, el total de puntos en la categoría "Muy Bueno" para el Criterio 4 (prueba de Friedman) es de 36 puntos, recordando, este valor se obtiene al multiplicar un total de 9 atributos por el valor de la categoría, que en este caso es 4; en el Criterio 5 es el mismo puntaje puesto que evaluaron los mismos 9 atributos. Al sumar estos puntajes se obtiene un máximo de 72 puntos para la categoría "Muy bueno".

La gráfica 6-9 representa el rendimiento por juez en las pruebas de ordenación al utilizar el Criterio 6.

Gráfica 6-9 Clasificación final para los candidatos a jueces sensoriales en pruebas de ordenación, aplicando la suma de puntos obtenidos en el Criterio 6



Con el fin de determinar la selección de jueces en pruebas de ordenación previamente se aplicaron los criterios establecidos en este tipo de prueba sensorial para cada familia de atributos, de manera que se pudiera generar la clasificación de los candidatos en la gráfica 7-9.

Las tablas 6-8 a 6-11 representan la clasificación que cada candidato obtiene en los criterios planteados para pruebas ordinales, desglosado por familia de atributos. Esto permite ubicar al juez y conocer su posición en cada criterio, por ejemplo al juez 10, utilizando el criterio basado en Friedman (Criterio 5) se le da una clasificación de "Bueno", pero en el criterio de muestras sin repetir valores de suma de rangos es "Muy bueno". Unificando ambos criterios este juez queda clasificado como "Muy bueno"

Tabla 6-8 Selección de jueces para la familia atributos visuales en pruebas de ordenación

Jueces seleccionados	Criterio		
	4	5	6
Categoría	Friedman	Suma de Rangos	Ambos
Muy Bueno(4)	8,	2,9,14,16	2,8,16
Bueno(3)	1,2,6,9,10,12,14,15,16,17	1,6,8,10,12,15,17	1,6,9,10,12,14,15,17
Regular(2)	4,13	4,13	4,13,15
Malo(1)	3,5,7	3,5,7	3,5,7
Faltas(0)	11	11	11

Tabla 6-9 Selección de jueces para la familia atributos aromáticos en pruebas de ordenación

Jueces seleccionados	Criterio		
	4	5	6
Categoría	Friedman	Suma de Rangos	Ambos
Muy Bueno(4)	1,9,13	9,10,12,13,14,17	1,9,13
Bueno(3)	2,8,14,16,17	1,2,3,8,16	2,3,8,10,12,16
Regular(2)	3,10,12		
Malo(1)	4,5,6,15	4,5,6,15	4,5,6,15
Faltas(0)	7,11	7,11	7,11

Tabla 6-10 Selección de jueces para la familia atributos textura bucal en pruebas de ordenación

Jueces seleccionados	Criterio		
	4	5	6
Categoría	Friedman	Suma de Rangos	Ambos
Muy Bueno(4)		1,2,12,13,14,16,17	1,12,16
Bueno(3)	1,2,6,10,12,16	6,10	2,6,10,13,14,17
Regular(2)	13,17		
Malo(1)	3,5,7,8,9,14,15	3,5,7,8,9,15	3,5,7,8,9,15
Faltas(0)	4,11	4,11	4,11

Tabla 6-11 Selección de jueces para la familia atributos gustos básicos en pruebas de ordenación

Jueces seleccionados	Criterio		
	4	5	6
Categoría	Friedman	Suma de Rangos	Ambos
Muy Bueno(4)		1,2,10,16,17	1,10
Bueno(3)	1,2,9,10,16,17	6,8,	2,6,8,9,16,17
Regular(2)	6,8,12,14	9,12,14	12,14
Malo(1)	3,4,7,13	3,4,7,13	3,4,7,13
Faltas(0)	5,11,15	5,11,15	5,11,15

6.3 SELECCIÓN DE CANDIDATOS A JUECES SENSORIALES APLICANDO LOS CRITERIOS CON FUNDAMENTOS ESTADÍSTICOS PARA DATOS DE TIPO INTERVALAR EN PRUEBAS SENSORIALES DE INTENSIDAD

6.3.1 Clasificación de candidatos a jueces sensoriales en pruebas sensoriales de intensidad para la familia de atributos visuales

Esta prueba fue efectuada con una escala no estructurada de 9 puntos, en donde el juez calificaba la intensidad percibida de cada atributo, y realizaba 3 réplicas. La familia de atributos visuales incluye a todas aquellas características sensoriales que se evalúan a través del sentido de la vista. Se evaluaron un total de cinco muestras de yogurt comercial.

La Tabla 6-12 presenta los resultados para la familia de atributos visuales en donde la segunda columna indica el número de los candidatos que lograron discriminar entre intensidades de productos de acuerdo con el Criterio 7. De esta manera se puede obtener la frecuencia de discriminación para cada juez, ya que se puede observar en que atributos si discriminó y en cuales no. Es importante mencionar que si el juez no discriminó entre productos no se puede considerar incluirlo en la clasificación del Criterio 9 al analizar su coeficiente de variación con el estadístico t de student y automáticamente es clasificado como malo. La clasificación que cada candidato obtiene de acuerdo al Criterio 9 se observa en las siguientes cuatro columnas. En naranja se indican los puntajes para el Criterio 8, en amarillo para el 9 y en morado el 10. En las siguientes gráficas se muestra el orden de clasificación de cada juez en los criterios propuestos. No hubo faltas de asistencia para esta familia de atributos.

Tabla 6-12 Clasificación de jueces para atributos visuales en pruebas de intensidad aplicando los Criterios 7,8,9 y 10.

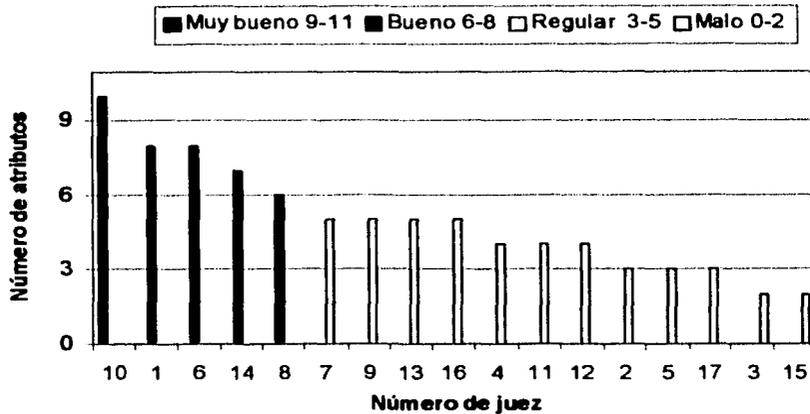
Clasificación	Criterio 7 ¹	Malo (1)	Regular (2)	Bueno (3)	Muy bueno (4)
Valor t student teórico :2.12 ²	F Teo:3.84 ³	t cal>2.12	0 < t cal < 2.12	-2.12<t cal < 0	t cal <-2.12
Significado	F cal >F teo Discriminan	1)CV juez>CV grupo 2)F cal < F teo	Cv juez ≥ CV grupo	CV juez ≤ CV grupo	CV juez <CV grupo
Atributos	11 atributos =100%	0-2	3-5	6-8	9-11
Porcentaje		0-25%	26-50%	51-75%	76-100%
Puntaje máximo Criterio 8		11	22	33	44
Puntaje máximo Criterio 9		11	22	33	44
Puntaje máximo Criterio 10		22	44	66	88
Atributos	Jueces seleccionados				
<i>Firmeza inicial</i>	1,2,3,4,5,6,8,9,10,12,13,14,15,16	7,11,17	2,6,9,13,16	1,3*,4,5,14,15	8,10,12
<i>Cantidad de fruta</i>	6,9,10,14,17	1,2,3,4,5,7,8,11,12,13,15,16	9,11	6,17	10,14
<i>Grosor</i>	6,10,14	1,2,3,4,5,7,8,9,11,12,13,15,16,17	14	6	10
<i>Color</i>	1,3,6,7,8,10,12,13,14,16	2,4,5,9,11,13,15,17	1,16	14	3,6,7,8,10,12
<i>Firmeza final</i>	1,4,6,8,9,10,13,14,16,17	2,3,5,11,12	1,9,13,15	10,14,17	4,6,7,8,16
<i>Viscosidad</i>	1,2,4,6,7,8,9,10,11,12,14	3,5,13,15,16,17	4,6,11	1,9,14	2,7,8,10,12
<i>Elasticidad</i>	1,9,10,11,12	2,3,4,5,6,7,8,13,14,15,16,17		1,10	9,11,12
<i>Gomosidad</i>	1,5,6,7,10,13	2,3,4,8,9,11,12,14,15,16,17		1,5,6,13	7,10
<i>Recubrimiento</i>	5,8,10,11,13,16	1,2,3*,4,6,7,9,12,14,15,17		16	5,7,8,10,11,13
<i>Suave</i>	1,11	2,3,4,5,6,7,8,9,10,12,13,14,15,16,17			1,11
<i>Grumoso</i>	1,2,4,6,7,8,10,13,14,16,17	3,5,9,11,12,15	1,4	6,7,13,14,16,17	2,8,10

¹ No se incluyen resultados de réplicas, sólo el juez 3 en los casos marcados con asterisco encontró diferencia, en estos casos se le disminuye un punto de su sumatoria total al emplear el Criterio 10. En donde este juez obtuvo 18 puntos totales menos dos de cada atributo en donde discrimino réplicas, le quedan un total de 16 puntos.

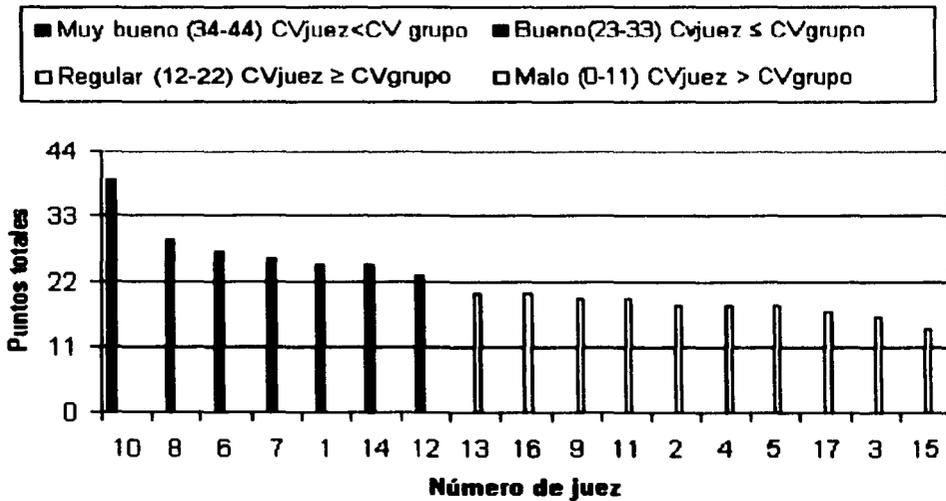
² Tomado de Tabla de Valores críticos de t de student para n=17 g. l. =16

³ Valor de Tablas de valores críticos de F para $\alpha=0.05$ y g.l. numerador =4, g.l. denominador= 8

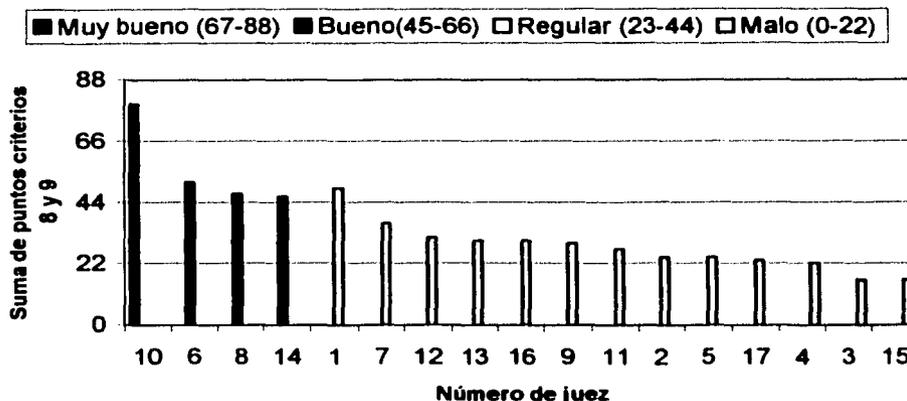
Gráfica 6-10 Clasificación de jueces considerando el Criterio 8 basado en la frecuencia de atributos en donde el juez encontraba diferencia significativa entre productos analizando atributos visuales.



Gráfica 6-11 Jueces clasificados en atributos visuales utilizando el Criterio 9 (comparación del coeficiente de variación de cada juez con el grupal mediante el estadístico t de student)



Gráfica 6-12 Clasificación de jueces para pruebas de intensidad en atributos visuales utilizando el Criterio 10



6.3.2 Clasificación de candidatos a jueces sensoriales en pruebas sensoriales de intensidad para la familia de atributos *aroma general*

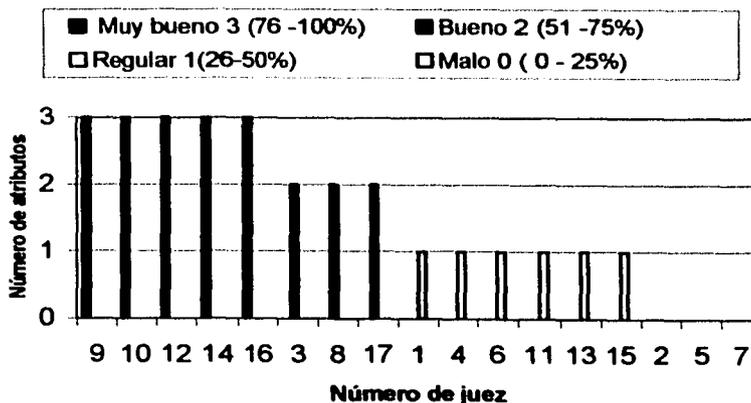
Esta prueba fue efectuada con una escala no estructurada de 9 puntos, en donde el juez calificaba la intensidad percibida de cada atributo y realizaba 3 réplicas. En estos ensayos no hubo faltas. Dentro de la familia de atributos de aroma general se consideraron tres específicos para aroma general de un producto lácteo como lo es el yogurt, estos atributos son: aroma general, aroma frutal y aroma lácteo.

Las siguiente tabla presenta el rendimiento del grupo de jueces para la familia de atributos aroma general en donde la segunda columna indica si el candidato logró discriminar entre productos. Aplicando el Criterio 8 se obtiene la frecuencia de discriminación para cada juez, que en este caso el total de atributos es de 3. Los puntajes de cada juez se determinan al sumar su rendimiento en cada atributo y así se pueden clasificar de acuerdo a los Criterios 7, 8, 9 y 10. La clasificación que cada juez se hace acreedor en base al Criterio 9 se muestran en las siguientes cuatro columnas. Al inicio de la Tabla se marcan los puntajes límites para cada criterio y para cada clasificación, con naranja puntos máximos Criterio 8, amarillo Criterio 9 y morado Criterio 10. En las siguientes gráficas se observan los resultados de cada juez en los tres diferentes criterios.

Tabla 6-13 Clasificación de jueces para atributos aroma general en pruebas de intensidad aplicando los Criterios 7,8,9 y 10

Clasificación	Criterio 7 ¹	Malo (1)	Regular (2)	Bueno (3)	Muy bueno (4)
Valor t student teórico :2.12 ²	F Teo:3.84 ³	t cal > 2.12	0 < t cal < 2.12	-2.12 < t cal < 0	t cal < -2.12
Significado	F cal > F teo Discriminan	1) CV juez > CV grupo 2) F cal < F teo	Cv juez ≥ CV grupo	CV juez ≤ CV grupo	CV juez < CV grupo
Atributos	3 atributos = 100%	0	1	2	3
Porcentaje		0-25%	26-50%	51-75%	76-100%
Puntaje máximo Criterio 8		3	6	9	12
Puntaje máximo Criterio 9		3	6	9	12
Puntaje máximo Criterio 10		6	12	18	24
Atributos	Jueces clasificados				
Aroma general	9,10,12,13,14,16	2,1,3,4,5,6,7,8,11,15,17	9,13	12,14	10,15,16
Aroma frutal	1,3,8,9,10,12,14,15,16,17	2,4,5,6,7,11,13	9,12,14,17	1,3,15	8,10,16
Aroma lácteo	3,4,6,8,9,10,11,12,13,14,16,17	1,2,5,7,15	4,6,9,11,13,15,16,17	8,12,14	3,10

Gráfica 6-13 Clasificación de jueces considerando el Criterio 8 basado en la frecuencia de atributos en donde el juez encontraba diferencia significativa entre productos analizando los atributos de aroma general

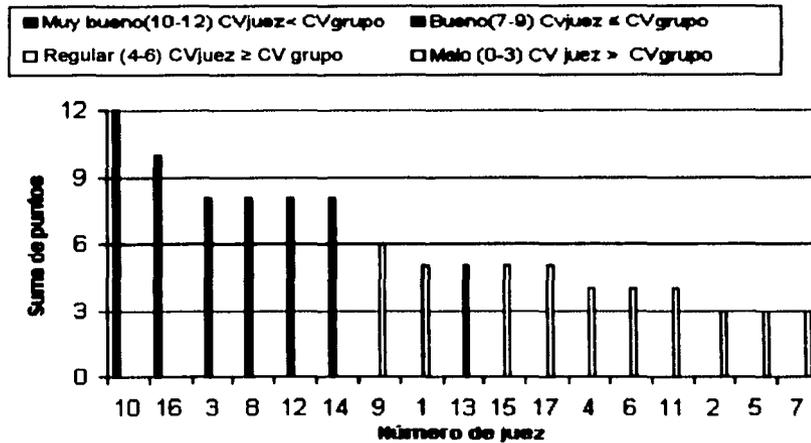


¹ No se incluyen resultados de réplicas, ya que en esta familia ningún candidato encontró diferencia entre ellas, ver cálculos.

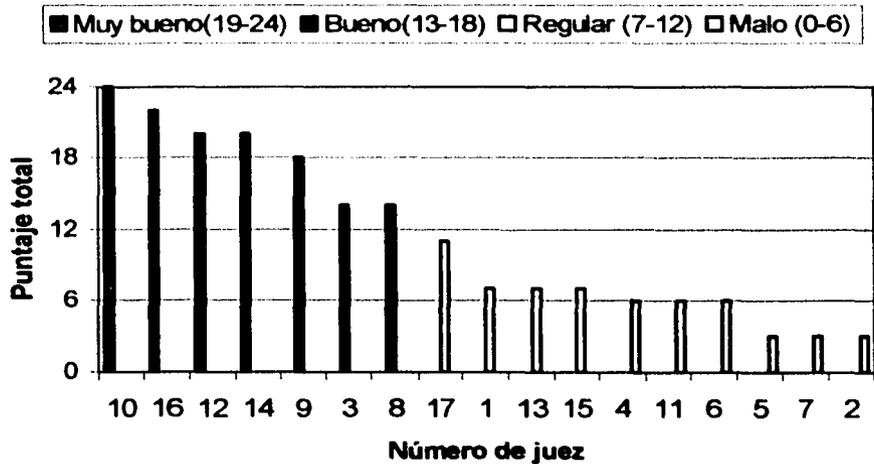
² Tomado de Tabla de Valores críticos de t de student para n=17 g. l. =16

³ Valor de Tablas de valores críticos de F para $\alpha=0.05$ y g.l. numerador = 4, g.l. denominador = 8

Gráfica 6-14 Jueces clasificados según desempeño en atributos aroma general utilizando el Criterio 9 (comparación del coeficiente de variación de cada juez respecto al grupal mediante el estadístico t de student)



Gráfica 6-15 Clasificación de jueces para pruebas de intensidad en atributos aroma general utilizando el Criterio 10



6.3.3 Clasificación de candidatos a jueces sensoriales en pruebas sensoriales de Intensidad para la familia de atributos aromáticos

En esta prueba se evaluó la intensidad de 21 atributos aromáticos presentes en 5 muestras de yogurt comercial con 3 réplicas, se utilizó una escala no estructurada de 9 puntos.

La Tabla 6-14 permite ubicar a cada uno de los jueces indicando su clasificación que obtiene en cada atributo. Al principio de la tabla se incluyen los máximos de cada criterio para generar la clasificación general por familia y obtener los siguientes gráficos en los cuales se identifica la clasificación general que obtuvo el juez para la familia de atributos aromáticos de acuerdo con los Criterios 7, 8, 9 y 10. La segunda columna de la tabla indica el número del juez que logró discriminar entre las intensidades de productos, en cuanto a las réplicas se indica con un asterisco a todos aquellos candidatos que no fueron congruentes en sus evaluaciones, es decir que discriminaron entre réplicas. Las inasistencias se indican en la columna titulada F en donde el candidato no obtiene puntos.

Gráfica 6-16 Clasificación de jueces considerando el Criterio 8 basado en la frecuencia de atributos en donde el juez encontraba diferencia significativa entre productos analizando atributos aromáticos.

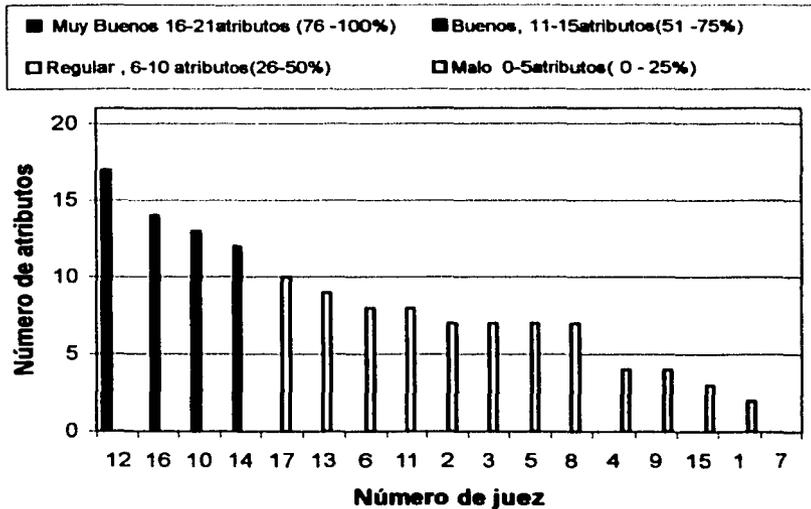


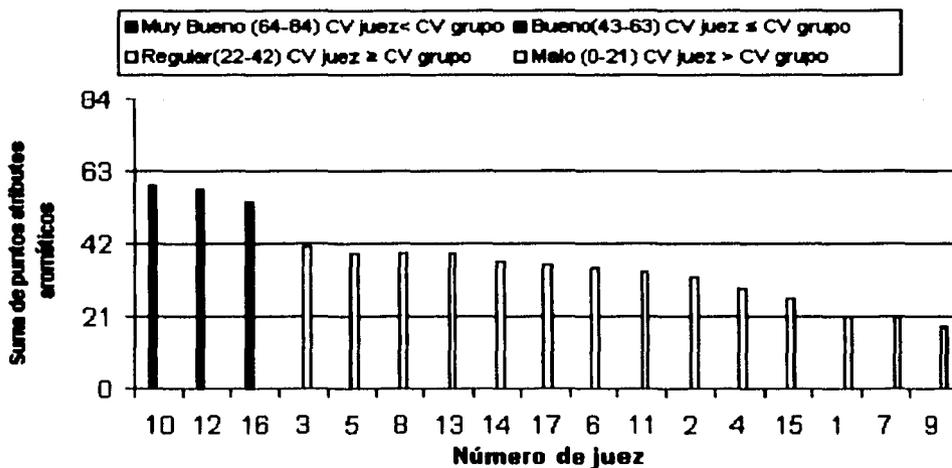
Tabla 6-14 Clasificación de jueces para atributos aromáticos en pruebas de intensidad aplicando los Criterios 7,8,9 y 10.

Clasificación	Criterio 7	F	Malo (1)	Regular (2)	Buena (3)	May buena (4)
Valor t student teórico :2.12 ¹	F Teo:3.84 ²	NA	t cal > 2.12	0 < t cal < 2.12	-2.12 < t cal < 0	t cal < -2.12
Significado	F cal > F teo Discriminan	NA	1) CV juez > CV grupo 2) F cal < F teo	Cv juez ≥ CV grupo	CV juez ≤ CV grupo	CV juez < CV grupo
Atributos	21 atributos = 100%		0 a 5	6 a 10	11 a 15	16 a 21
Porcentajes			0-25%	26-50%	51-75%	76-100%
Puntos máximo Criterio 8		0	21	42	63	84
Puntaje máximo Criterio 9		0	21	42	63	84
Puntaje máximo Criterio 10		0	42	84	126	168
Atributos	Jueces clasificados					
Frutal	6,11,12,16,17		1,2,3,4,5,7,8,9,10, 13,14,15	11,17	6,12,16	
Queso	5,6,10,11,12 13,14,16,17	9	1,2,3,4,7,8,15	6,11,14,17	5,12,16	10,13
Diacetilo	1,2,3,4,6,8, 10,11,12,14,16	9	1,5,7,13,15,17	2,4,6	12,14,16	3,8,10,11
Vainilla	2,3,8,10,12,13,14	9	1,4,5,6,7,11,15, 16,17	2	12,13,14	3*,8,10
Caramelo	1,4,11,12	9	2,3,5,6,7,8,10, 13,14,15,16,17	1,14	4,11,12	
Quemado	12	9	1,2,3*,4,5,6,7,8,10,11, 13,14,15,16,17			12
Rojo.	2,3,5,9,10,11,12		1,4,6,7,8,13,14,15,16, 17		2,12	3,5,9,10*,11
Amillico	13,17	9	1,2,3,4,5,6,7,8,10, 11,12,14,15,16		13,17	
Tutti-Frutti	1,2,6,8,9,10,11, 12,14,16,17		3,4,5,7,13,15	1,2,6,14,17	8,9,11	10*,12,16
Antranilato	2,3,5,6,12,13, 14,15,16,17	9	1,4,7,8,9,10,11*	2,14,15,17	3*,6,13	5,12,16
Almendra	3,6,11,13,14,15,16	9	1,2,4,5,7,8,9,10,12,17	11,13,14,15	6,16	3
Canela	9,10,12,13,16	11	1,2,3,4,5,6,7,8,14, 15,17		9,12,13	10,16
Maduro	4,5,10,12,14	9	1,2,3,6,7,8,9,11*,13, 15,16,17		14	4,5*,10,12
Citrico	2,5,10,12,14,16,17	9	1,3,4,6,7,8,11,13,15	12,14	5,16,17	2,10*
Verde	5,6,8,9,10,16,17		1,2,3,4,7,11,12,13,14, 15		5,16,17	6,8,9,10
Cre moso	3,5,10,13,15, 16		1,2,4,6,7,8,9,11,12, 14,17		13	3,5,10,15,16
Graso	3,10,12,14,16	9	1,2,4,5,6,7,8,9,11, 13,15,17		12,14	3,10*,16
Medicinal	4,8,10,12,14,16,17		1,2,3,5,6,7,9,11,13,15	12,14,17	4,8,10*	16
Cocido	1,12,13,16,17	9	1,2,3,4,5,6*,7*,8,10*, 11,14,15	16	12,13,17	
Plástico	2,8,10,11,12	1,9	3,4,5,6,7*,13,14,15, 16,17		2,8*,12	10*,11
Tostado	6,8,12,13,14, 16,17	1,9, 11	2,3,4,5,7*,10,15	14	6*,13,16,17	8*,12

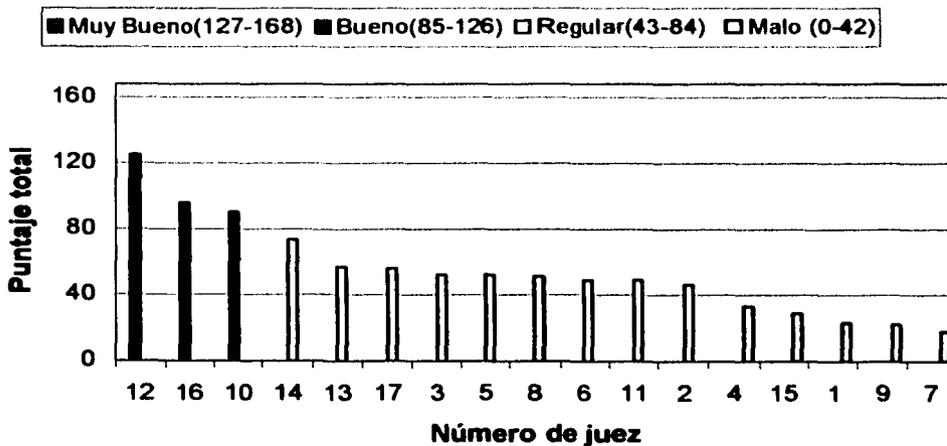
¹ Tomado de Tabla de Valores críticos de t de student para n=17 g. l. =16

² Valor de Tablas de valores críticos de F para α=0.05 y g.l. numerador =4, g.l. denominador= 8

Gráfica 6-17 Jueces clasificados según desempeño en atributos aromáticos utilizando el Criterio 9 (comparación del coeficiente de variación de cada juez con el grupal mediante el estadístico t de student)



Gráfica 6-18 Clasificación de jueces para pruebas de intensidad en atributos aromáticos utilizando el Criterio 10



6.3.4 Clasificación de candidatos a jueces sensoriales en pruebas sensoriales de intensidad para la familia de gustos básicos

Esta prueba fue efectuada con una escala no estructurada de 1 a 9 puntos, en donde el juez calificaba la intensidad percibida en cada gusto básico, realizando 3 réplicas.

La siguiente tabla muestra los resultados que cada juez obtuvo al evaluar la intensidad de los gustos básicos. El promedio para cada uno de ellos de acuerdo a cada criterio se muestra en las siguientes gráficas, mismas que tienen como límites para cada criterio los establecidos en esta Tabla diferenciados por colores, siendo el amarillo los límites de puntos que necesita un juez para poder clasificarse en muy bueno, bueno, regular y malo de acuerdo al Criterio 8, en naranja se establecen los límites para el Criterio 9 frecuencia de discriminación y con morado el Criterio 10 que es la unión de esos dos criterios.

Tabla 6-15 Clasificación de jueces para gustos básicos en pruebas de intensidad aplicando los Criterios 7, 8, 9 y 10

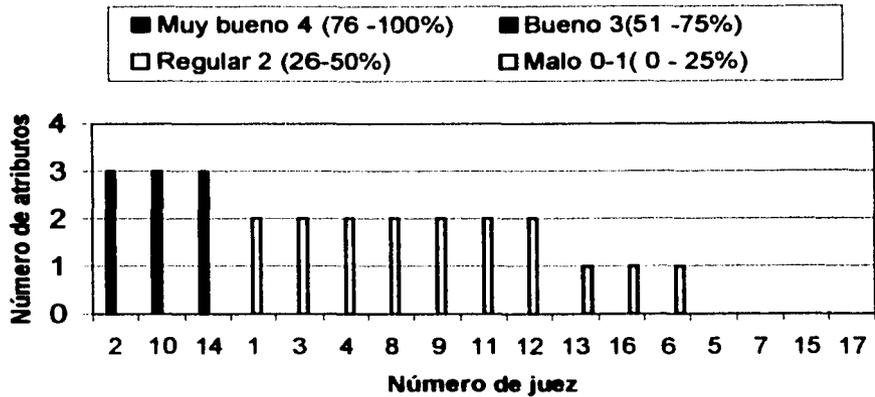
Clasificación	Criterio 7 ¹	Malo (1)	Regular (2)	Bueno (3)	Muy bueno (4)
Valor t student teórico :2.12 ²	F Teo:3.84 ³	t cal > 2.12	0 < t cal < 2.12	-2.12 < t cal < 0	t cal < -2.12
Significado	F cal > F teo Discriminan	1) CV juez > CV grupo 2) F cal < F teo	Cv juez ≥ CV grupo	CV juez ≤ CV grupo	CV juez < CV grupo
Atributos	4 atributos = 100%	0 - 1	2	3	4
Porcentaje		0-25%	26-50%	51-75%	76-100%
Puntaje máximo Criterio 8		4	8	12	16
Puntaje máximo Criterio 9		4	8	12	16
Puntaje Criterio 10		8	16	24	32
Atributo	Jueces clasificados				
Ácido	1,2,3,4,8,10,12,16	5,6,7,9,11,13,14,15,17	2,4,12	1,3,8,16	10
Salado	2,10,14	1,3,4,5,6,7,8,9,11,12,13,15,16,17	2,14		10
Amargo	8,9,11,12,14	1,2,3,4,5,6,7,10,13,15,16,17	9,12	14,	8,11
Dulce	1,2,3,4,6,9,10,11,13,14	5,7,8,12,15,16,17	3,4,6,9,13	1,11,14	2,10

¹. En estos ensayos no hubo faltas. No se incluyen resultados de réplicas, ya que en esta familia ningún candidato encontró diferencia entre ellas, ver cálculos.

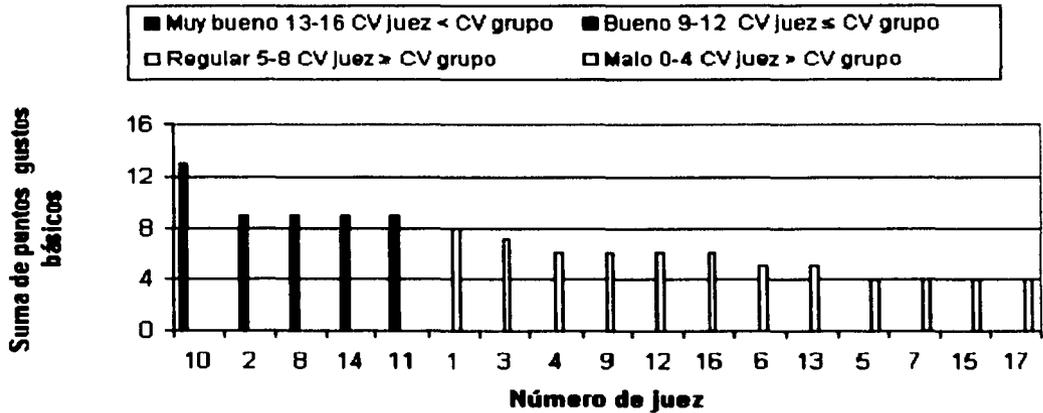
² Tomado de Tabla de Valores críticos de t de student para n=17 g. l. =16

³ Valor de Tablas de valores críticos de F para α=0.05 y g.l. numerador =4, g.l. denominador= 8

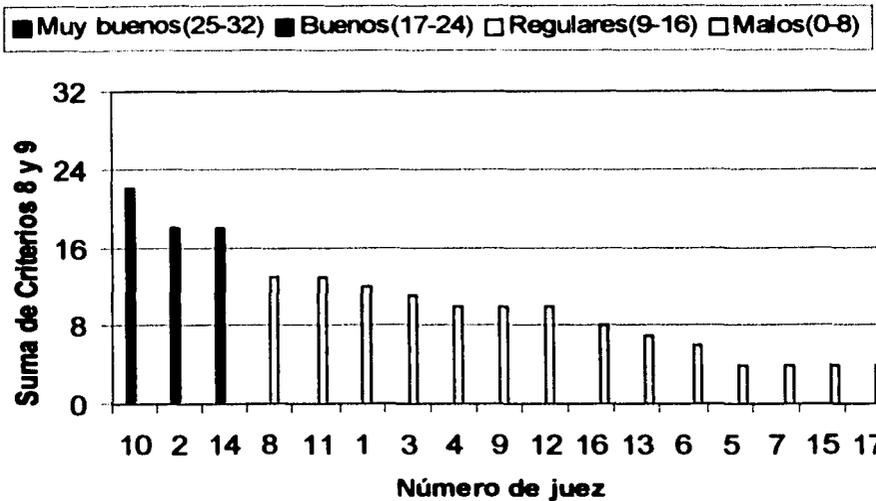
Gráfica 6-19 Clasificación de jueces considerando el Criterio 8 basado en la frecuencia de atributos en donde el juez encontraba diferencia significativa entre productos analizando de gustos básicos.



Gráfica 6-20 Jueces clasificados para la familia de gustos básicos utilizando el Criterio 9 (comparación del coeficiente de variación de cada juez con el grupal mediante el estadístico t de student)



Gráfica 7-21 Clasificación de jueces para pruebas de intensidad en gustos básicos utilizando el Criterio 10



6.3.5 Clasificación de candidatos a jueces sensoriales en pruebas sensoriales de intensidad para la familia de atributos textura bucal

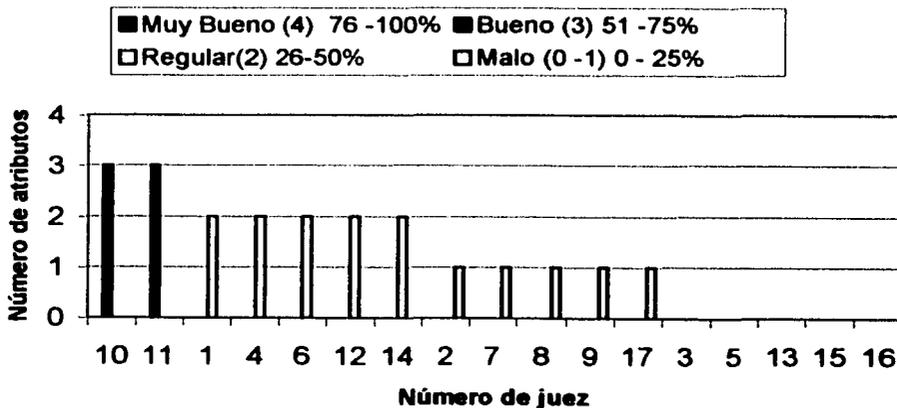
Esta prueba fue efectuada con una escala no estructurada de 9 puntos, en donde el juez calificaba la intensidad percibida por cada atributo de textura bucal, se realizaron 3 réplicas. En estos ensayos no hubo faltas de asistencia.

La Tabla 6-16 muestra los resultados que cada juez obtuvo en atributos de textura bucal. El puntaje total para cada uno de ellos de acuerdo a cada criterio se muestra en las siguientes gráficas. Estas gráficas tienen como límites para cada criterio los establecidos en esta Tabla diferenciados por colores, siendo el amarillo los límites de puntos que necesita un juez para poder clasificarlo en muy bueno, bueno, regular y malo de acuerdo al Criterio 9, en naranja se establecen los límites para el Criterio 8 frecuencia de discriminación y con morado el Criterio 10 que es la unión de esos dos criterios. La clasificación de cada juez se ubica en el centro de la tabla por cada atributo evaluado.

Tabla 6-16 Clasificación de jueces para atributos textura bucal en pruebas de intensidad aplicando los Criterios 7,8,9 y 10.

Clasificación	Criterio 7 ¹	Malo (1)	Regular (2)	Bueno (3)	Muy bueno (4)
Valor t student teórico :2.12 ²	F Teo:3.84 ³	t cal > 2.12	0 < t cal < 2.12	-2.12 < t cal < 0	t cal < -2.12
Significado	F cal > F teo Discriminan	1) CV juez > CV grupo 2) F cal < F teo	Cv juez ≥ CV grupo	CV juez ≤ CV grupo	CV juez < CV grupo
Atributos	4 atributos =	0 - 1	2	3	4
Porcentaje	100%	(0-25%)	(26-50%)	(51-75%)	(76-100%)
Puntaje máximo Criterio 8		4	8	12	16
Puntaje máximo Criterio 9		4	8	12	16
Puntaje máximo Criterio 10		8	16	24	32
Atributos	Jueces seleccionados				
Consistencia	4,7,8,9,10,11,14	1,2,3,5,6,12,13,15,16,17	4	9	7,8,10,11,14
Recubrimiento	1,4,6,10,11,12,14,17	2,3,5,7,8,9,13,15,16	4,17	1,6,14	10,11,12
Graso	1,12	2,3,4,5,6,7,8,9,10,11,13,14,15,16,17		1	12
Blando	2,6,10,11	1,3,4,5,7,8,9,12,13,14,15,16,17		6,11	2,10

Gráfica 6-22 Clasificación de jueces considerando el Criterio 8 basado en la frecuencia de atributos en donde el juez encontraba diferencia significativa entre productos analizando atributos textura bucal

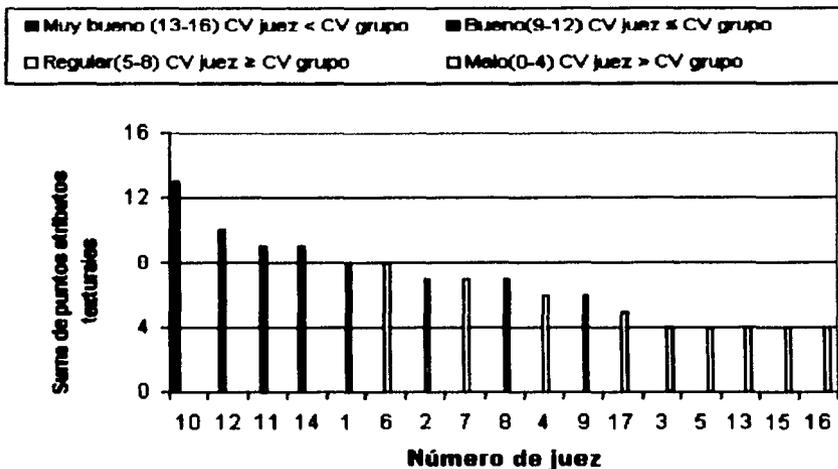


¹ No se incluyen resultados de réplicas, puesto que ningún candidato discriminó entre ellas, ver calculas.

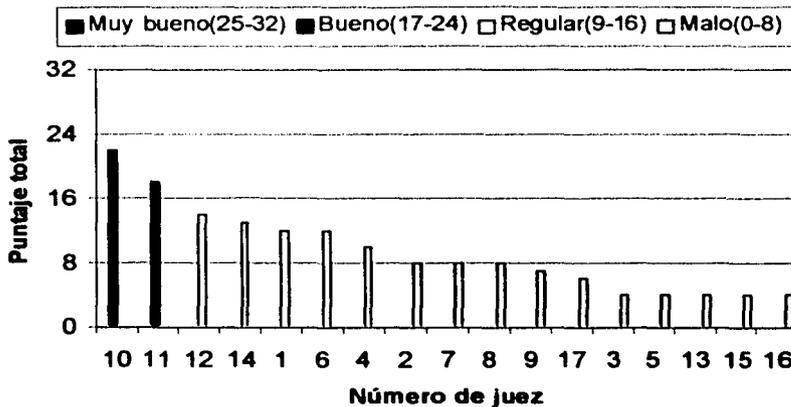
² Tomado de Tabla de Valores críticos de t de student para n=17 g. l. =16

³ Valor de Tablas de valores críticos de F para $\alpha=0.05$ y g.l. numerador =4, g.l. denominador= 8

Gráfica 6-23 Jueces clasificados en atributos textura bucal utilizando el Criterio 9 (comparación del coeficiente de variación de cada juez con el grupal mediante el estadístico t de student)



Gráfica 6-24 Clasificación de jueces para pruebas de intensidad en atributos textura bucal aplicando el Criterio 10



6.4 CLASIFICACIÓN FINAL DE CANDIDATOS A JUECES SENSORIALES

Esta tabla representa el resumen de los resultados de haber aplicado los criterios sugeridos en la metodología, y su finalidad es ubicar al candidato de acuerdo a su desempeño demostrado en cada prueba sensorial realizada. Por otra parte a cada juez se le califica de manera global, ya que se obtiene un promedio general para cada juez.

Tabla 6-17 Clasificación final de los candidatos a jueces sensoriales por tipo de estadística, número y prueba sensorial

TIPO DE ESTADÍSTICA	NO PARAMÉTRICA						PARAMÉTRICA					Suma de puntos	Calificación final
	Nominal		Ordinal				Intervalar						
Tipo de número	IA	CP	V	Ag	Ar	Tb	V	Ag	Ar	Tb	Gb		
1	B	B	B	MB	MB	MB	R	R	M	R	R	30	B
2	MB	B	MB	B	B	B	R	M	R	M	B	29	B
3	F	F	M	B	M	M	M	B	R	M	R	15	R
4	R	B	R	M	F	M	M	M	R	R	R	17	R
5	MB	MB	M	M	M	F	R	M	M	M	M	17	R
6	R	B	B	M	B	B	B	M	R	R	M	24	B
7	B	R	M	F	M	M	R	M	M	M	M	14	R
8	MB	MB	MB	B	M	B	B	B	R	M	R	30	B
9	MB	MB	B	MB	M	B	R	B	M	B	R	30	B
10	MB	MB	B	B	B	MB	MB	MB	B	B	B	34	MB
11	R	B	F	F	F	F	R	M	M	B	R	14	R
12	B	MB	B	B	MB	R	R	MB	B	R	R	32	B
13	B	B	R	MB	B	M	R	M	R	M	M	23	B
14	B	B	B	B	B	R	B	MB	R	R	B	32	B
15	R	B	B	M	M	F	M	R	M	M	M	16	R
16	MB	MB	MB	B	MB	B	R	MB	B	M	M	30	B
17	R	B	B	B	B	B	R	M	R	M	M	24	B

Tipo de prueba

IA	Identificación de atributos
CP	Comparación por pares
V	Atributos visuales
Ag	Atributos aroma general
Ar	Atributos aromáticos
Tb	Textura bucal

Clasificación

Puntos

MB	Muy bueno	4	34-44 puntos
B	Bueno	3	23-33 puntos
R	Regular	2	12-22 puntos
M	Malo	1	0-11 puntos
F	Faltas	0	

Gb

Gustos básicos

7 ANÁLISIS DE RESULTADOS

- Cuando se propuso el Criterio 1 de selección de candidatos a jueces sensoriales, (que está basado en el porcentaje de identificación de atributos para gustos básicos), se encontró que una gran cantidad de jueces calificaba como "Bueno" y "Muy Bueno". Sin embargo se buscaba tener una clasificación más equitativa en donde no solo se seleccionara al candidato por porcentajes de identificación, sino proponer otro criterio que fuera más selectivo, por esa razón se empleó un criterio basado en el estadístico J_i cuadrada. Un candidato necesita un mínimo de aciertos para poder aprobar la prueba. De esta manera se logra seleccionar solo aquellos candidatos con habilidad natural para desarrollar tal prueba. Se hizo necesario desarrollar un tercer criterio que involucra a los dos anteriores y que permitiera al analista seleccionar a sus candidatos de una manera más equilibrada, por supuesto que el analista puede escoger entre los criterios aquel que se adapte mejor a sus necesidades finales de la investigación sensorial realizar.
- En las pruebas de tipo ordinal se probaron 3 métodos estadísticos para desarrollar los criterios que permitan establecer una selección de los candidatos a jueces. Los métodos probados fueron: Page Test, Análisis de Rangos de Newell-MacFarlane y Prueba de Friedman. En este orden se determinó que el método Page Test era demasiado flexible¹ y no permitía cumplir con los objetivos de selección ya que no se podía generar una clasificación de los candidatos debido a que a todos los consideraba como "Muy Bueno". Por el lado contrario el Análisis de Rangos de Newell-MacFarlane es un método muy estricto, en donde ningún candidato cumplía con los requisitos mínimos de esta prueba para ser considerado en una selección. Por tal razón se escogió el método de Friedman ya que este método facilitaba la selección y se pudieron generar categorías de desempeño. Una vez que se eligió el método estadístico, se tuvieron que determinar límites, puesto que en este tipo de prueba estadística, si es posible determinar quien ordena correctamente de acuerdo a la naturaleza de las muestras evaluadas.
- Al igual que en la base de datos nominal, para la base de datos ordinal se establecieron 3 criterios, uno que definiera si el candidato podía encontrar diferencia significativa al ordenar (basado en el método de Friedman) Criterio 4, otro basado en la secuencia de ordenación, es decir, que el candidato no repitiera valores de suma de rangos al ordenar Criterio 5, así como un tercer criterio que unifique a los dos anteriores Criterio 6. El Criterio 5 se creo al determinar mediante los cálculos² que había jueces que encontraban diferencia significativa al ordenar analizando la información por la Prueba de Friedman, pero que no diferenciaban entre un par o dos de muestras al darles el mismo valor de suma de rangos. Es muy importante hacer mención que en este Criterio no se consideran a todos aquellos candidatos que no cumplen con el Criterio 4 porque ellos no ordenaron significativamente y entonces sus valores de suma de

¹ Con flexible se quiere decir que en este caso que el método facilita la selección de la gran mayoría de los candidatos y no hace una selección de ellos.

² Para dudas sobre cálculos solicitar disco.

rangos no son representativos, debido a que ordena por igual a todas las muestras y en conclusión este candidato es malo en los Criterios 4,5 y 6.

- Para poder generar los criterios de selección de candidatos a jueces sensoriales en la base de intervalo, se utilizaron varias técnicas estadísticas con la finalidad de evaluar al candidato en su propio desempeño así como con respecto al grupo. Lo primero que se consideró es a cada juez como un instrumento de medición. Entonces se separó cada producto que el juez evaluó en las pruebas de intensidad de atributos, teniendo un total de 5 productos con 3 réplicas en cada uno de los atributos calificados. Para esta primera parte se determinó mediante estadística descriptiva los valores que cada juez obtenía de promedio de intensidad, desviación y coeficiente de variación por producto y atributo que evaluaba. Esta información tenía que ser filtrada para poder generar una clasificación por lo que se desarrolló el Criterio 7 que, como está basado en un Análisis de varianza de dos vías, permite determinar qué candidato a juez si encuentra diferencia significativa entre las intensidades de varios productos, la naturaleza de esta técnica estadística también permite detectar a todos aquellos candidatos que discriminan entre réplicas o repeticiones (este efecto no es deseado). En el caso de réplicas, la mayoría de los candidatos a jueces sensoriales considerados en este estudio no discriminaban entre las diferentes repeticiones, tal y como lo muestran los resultados. Sin embargo en la familia de aromáticos, algunos jueces como el candidato 3 si encontraban diferencia entre las réplicas del mismo producto. Para considerar este efecto en su clasificación final se le disminuye un punto por cada vez atributo en el que un candidato haya diferenciado entre réplicas del total que obtuvo en el Criterio 10.
- Una vez aplicado el Criterio 7, el siguiente paso es conocer el desempeño de un candidato por familia de atributos, esto porque es importante saber como evalúa utilizando cada uno de sus sentidos. Por esta razón se creó el Criterio 8 en donde se mide la frecuencia de discriminación como un porcentaje en donde el total de la familia equivale al 100 por ciento.
- Otro factor fundamental en la selección de un candidato es la variabilidad entre sus resultados, por este motivo se comparó el coeficiente de variación de cada candidato con el grupal. La herramienta estadística que permite hacer una comparación de ambos valores es el estadístico t de student para una muestra relacionada (porque los dos valores provienen del mismo grupo de candidatos a jueces). Esta herramienta permitió ubicar a cada candidato de acuerdo a los límites establecidos en el Criterio 9. Finalmente para tener una selección completa de los candidatos se creo el Criterio 10 que une todos los pasos anteriores.
- En todas las bases de datos analizadas las inasistencias fueron un factor fundamental. El analista no debe considerar como candidato a juez a una persona que no asista a las sesiones, ya que retrasa el avance del grupo, así como su desarrollo propio. Esto pudo ser fácilmente comprobable puesto que jueces con faltas se ubicaron en las categorías de clasificación más bajas incluso cuando si asistían a las sesiones. Como ejemplo se tiene al juez 3. (Ver tabla 6-17 pag. 78.

8 CONCLUSIONES

- **Se generaron los criterios para seleccionar candidatos a jueces sensoriales en base a la prueba sensorial y el método estadístico apropiado.**
- **Se propone el sistema o conjunto de procedimientos descritos en esta tesis, para la clasificación de jueces sensoriales. Este sistema se puede adaptar a las necesidades del analista sensorial que trabaje en una empresa o centro de investigación.**
- **Los criterios propuestos en esta tesis facilitan una rápida selección de candidatos a jueces sensoriales que recibirán un entrenamiento posterior.**
- **Debido a la naturaleza de las pruebas sensoriales y los métodos estadísticos, los criterios planteados pueden ser empleados no solo en la etapa de selección de candidatos a jueces sensoriales, sino también en la etapa del proceso de entrenamiento de jueces sensoriales.**
- **Los criterios generados permiten detectar personas con habilidad natural para trabajar en pruebas sensoriales.**

9 BIBLIOGRAFÍA

1. ASTM (1981). Guidelines for the selection and training of sensory panel members. Philadelphia.
2. Bárcenas, P., Pérez, F.J., Salmerón, J. y Albisu, M. (2000). Análisis sensorial de alimentos. Selección entrenamiento y control de los catadores. Alimentaria Julio-Agosto 00/31:31-38.
3. Ennis, D. (1998). Foundations of Sensory Science and a vision for the future Food Technology. Vol 52 number 7 (-)
4. Gatchalian, M.M., Yano, T. (1990). Quantified approach to sensory panel selection. Food Quality and Preference 2 (233-241)
5. ISO 8586-1: 1993 General guidance for the selection, training and monitoring of assessors. Part 1: selected assessors
6. ISO 5495: 1993 Sensory analysis-methodology- paired comparison test
7. ISO 8587: 1993 Sensory analysis-methodology- ranking
8. Jellinek G. (1985). Sensory Evaluation of food. Theory and practice. Editorial Ellis Horwoodserv in Food Technology . USA
9. McDaniel C, Gates R (1999). Investigación de Mercados Contemporánea. Editorial Thomson Editores. México.
10. Moskowitz, H. (1983). Product Testing and Sensory Evaluation of foods. Marketing and R&D Approaches. Food & Nutrition Press, Inc. USA.
11. Moskowitz, H. (1988). Applied Sensory Analysis of Food. Volume I and II. CRC. Inc. USA
12. Moskowitz, H. (1997). Magnitude Estimation: Notes on what, how , when and why to use it. J. Food Quality 3, 195-227
13. Muñoz A, Civille and Carr (1992), Sensory evaluation in quality control .Van Nostrand Rinhold New York, pp 199-227
14. Naiman, A. Rosenfeld R. Zirkel G. (1987) Introducción a la estadística. 3ª edición. Editorial Mc Graw Hill, México pp2-20
15. Penta Sensorial S.A. de C.V. (2001). Memorias del Curso Estadística Aplicada a la Evaluación Sensorial. México
16. Pedrero, D. y Pangborn, R.M.(1997) Evaluación sensorial de los alimentos. Métodos analíticos. Editorial Alhambra Mexicana, México.

17. Pedrero, D. y Menchero, P.(1990). Guía para la selección y entrenamiento de jueces analíticos capaces de efectuar evaluaciones sensoriales. En: Diccionario de especialidades para la industria alimentaria. 2a. edición. Ediciones P.L.M, S.A. de C.V. México, D.F. pag 263-269.
18. O'Mahony, M (1986). Sensory evaluation of food. Statistical methods and procedures. Marcel dekker, inc. New York U.S.A
19. O'Mahony, M (1982). Some assumptions and difficulties with common statistics for sensory analysis . Food Technology, 11:75-82
20. O'Mahony, M (1995) Who told you the triangle test was simple? Food Quality and Preference 6:227-237.
21. Stone H. (1993). Sensory evaluation practices. 2ª edición Academic Press Inc. pp 147
22. Siegel, D. (1991). Estadística no paramétrica aplicada a las ciencias de la conducta. 3ª edición. Trillas. México.
23. Willoughby, S. (1991). Probabilidad y estadística . 1ª edición. Publicaciones Cultural. México

10 APÉNDICES

10.1 RESUMEN DE LOS CRITERIOS OBTENIDOS

Criterio 1: Selección de jueces en pruebas discriminativas, basado en el porcentaje de identificación de gustos básicos.

Porcentaje de identificación	No aplica	(0-25%)	(26-50%)	(51-75%)	(76-100%)
Significado	Candidatos que no cumplen con su asistencia.	Candidatos que no identifican un mínimo de gustos básicos.	Es muy difícil para el candidato distinguir entre los gustos básicos ya que lo máximo que puede identificar es un 50% de ellos.	El candidato distingue por arriba del 50% de los ensayos, pero no logra a diferenciar de manera total	Del total de aciertos, el candidato identifica a la gran mayoría de los gustos básicos
Clasificación	Faltas	Malo	Regular	Bueno	Muy bueno
Puntaje	0	1	2	3	4

Criterio 2: Selección de jueces para pruebas discriminativas basado en el estadístico Ji Cuadrada

Valor de Ji cuadrada	No aplica	No aplica	Valor crítico menos uno	Mínimo de juicios correctos para ser diferentemente significativa	Arriba del valor mínimo y hasta el máximo de aciertos
Significado	Los candidatos que por asistencia no son acreedores a puntos.	Candidatos que no discriminan significativamente, no identifican atributos.	El juez se ubica tan solo a un punto del mínimo de juicios para encontrar diferencia significativa	El candidato logra el mínimo de juicios correctos para encontrar diferencia significativa	Evidentemente el juez discrimina entre muestras
Clasificación	Faltas	Malo	Regular	Bueno	Muy bueno
Puntaje	0	1	2	3	4

Criterio 3: Suma de puntaje total aplicando los Criterios 1 y 2 en la selección de candidatos a jueces para pruebas discriminativas

Clasificación	Faltas	Malo	Regular	Bueno	Muy bueno
Puntaje	0	1-2	3-4	5-6	7-8

Criterio 4: Selección de jueces para pruebas de ordenación basado en el estadístico Prueba de Friedman

Valor de Friedman	Negativo por faltas	Negativos	$0 > \text{Valor Friedman calculado} < \text{Valor Friedman teórico}$	$\text{Valor Friedman teórico} < \text{Valor Friedman calculado} < \text{Valor Friedman teórico} + 1/3 D$	$\text{Valor Friedman teórico} + 1/3 D < \text{Valor Friedman calculado} < \text{Valor Friedman teórico} + 2/3 D$	$\text{Valor Friedman teórico} + 2/3 D < \text{Valor Friedman calculado} < \text{Valor Friedman máximo}$
Significado	No asistieron a la sesión	Repetieron valores en el análisis de rangos, aún cuando en las instrucciones no se permiten empates	No encuentran diferencia significativa al ordenar, ya que tienen un valor de Friedman menor al teórico	El candidato ordena pero lo hace cubriendo como máximo un tercio del valor de Friedman de lo que debería ordenar	El candidato ordena mejorando el número de ordenamiento ya cubre como máximo dos tercios del valor total de Friedman	El juez puede obtener como máximo el valor de Friedman y lograr el total para una ordenación perfecta.
Clasificación	Faltas	Negativos	Malo	Regular	Bueno	Muy bueno
Puntaje	0	0	1	2	3	4

Criterio 5: Selección de jueces considerando el número de muestras que el candidato a juez ordena de sin repetir valores de suma de rangos.

Muestras sin repetir valores de suma de rangos	Faltas	Negativos	No aplica	60%	80%	100%
Significado	Jueces que faltaron a las sesiones	Repetieron valores en el análisis de rangos	Se considera a todos los jueces que clasificaron como "Malo" en el Criterio 4, puesto que no encuentran diferencia significativa al ordenar entre varias muestras.	El candidato repite valores en un 40%, es decir de 5 muestras presentadas evalúa a dos pares de ellas con el mismo rango, considera a tres muestras con el mismo rango.	Dentro de la serie muestras valora a un par de muestras con el mismo rango.	No repite rangos para ninguna de las muestras presentadas evalúa de manera diferente cada muestra
Clasificación	Faltas	Negativos	Malo	Regular	Bueno	Muy bueno
Puntaje	0	0	1	2	3	4

Criterio 6: Selección de jueces para pruebas de ordenación sumando los criterios basados en el estadístico Friedman y en el número de muestras sin repetir valores de suma de rangos.

Clasificación	Muy bueno	Bueno	Regular	Malo	Faltas y negativos
Criterio 4	4	3	2	1	0
Criterio 5	4	3	2	1	0
Puntaje Total	7-8	5-6	3-4	1-2	0

Criterio 7: Selección de jueces en pruebas de intensidad utilizando el estadístico Análisis de varianza de dos vías

Valor de F	F teórica productos $\alpha \leq 0.05$; g.l productos /g.l error		F teórica replicas $\alpha \leq 0.05$; g.l replicas /g.l. error	
Criterio	F juez(productos) > F teórica	F juez(productos) < F teórica	F juez(replica) > F teórica	F juez(replica) < F teórica
Significado	El juez logra discriminar entre los productos que se le presenten de manera significativa	El juez no discrimina entre productos	El juez no es consistente en sus calificaciones	El juez cumple con lo esperado dando calificaciones semejantes a intensidades semejantes.
Clasificación	Esta en función de su desempeño por familia de atributos	Malo	Malo	Esta en función de su desempeño por familia de atributos

Criterio 8: Selección de jueces en pruebas de intensidad de acuerdo con su frecuencia de discriminación de atributos por familia utilizando el estadístico Análisis de varianza de dos vías

Discrimina		Replicas		Productos	
Criterio	F _{juez(replica)} > F teórica	0-25% de los atributos	25-50% del total de atributos	51-75% de los atributos	76-100% de los atributos
Significado	El juez no es constante en sus calificaciones	El juez tiene un bajo porcentaje de discriminación entre las intensidades que presentan los productos para una determinada familia de atributos.	El juez le cuesta trabajo discriminar las intensidades que presentan los productos en todos los atributos de una familia y lo máximo que alcanza es un 50% del grupo de atributos	El juez discrimina las intensidades que presentan los productos en más de la mitad de los atributos de la familia	El juez logra discriminar las intensidades que presentan los productos en la mayoría de atributos de una determinada familia
Clasificación	No aplica	Malo	Regular	Bueno	Muy Bueno
Puntaje	Menos un punto a la suma total por cada atributo	1	2	3	4

Criterio 9: Clasificación de jueces en pruebas de intensidad considerando los coeficientes de variación grupal e individual, comparándolos con el estadístico t de student

Valor de t Student	No aplica	Valor de t calculada mayor que el valor teórico	Valor de t calculada entre cero y el valor teórico	Valor de t calculada entre el negativo del valor teórico y cero	Valor de t calculada menor que el negativo del Valor teórico
Significado	Jueces con faltas	Coefficiente de variación mucho mayor al grupal.	Coefficiente de variación similar al grupal, pero por arriba del promedio grupal	Coefficiente de variación similar al grupal, pero por debajo del promedio grupal	Coefficientes de variación mucho menores al grupal
Clasificación	Faltas	Malo	Regular	Bueno	Muy Bueno
Puntaje	0	1	2	3	4

Criterio 10: Selección de jueces en pruebas de intensidad unificando los criterios basados en el coeficiente de variación y en la frecuencia de discriminación

Clasificación	Faltas	Malo	Regular	Bueno	Muy bueno
Criterio 8	0	1	2	3	4
Criterio 9	0	1	2	3	4
Puntaje Total	0	1-2	3-4	5-6	7-8

10.2 EJEMPLO DE CÁLCULOS PARA DETERMINAR JI CUADRADA Y PORCENTAJES DE RECONOCIMIENTO DE GUSTOS BÁSICOS.

Ejemplo de cálculos para determinar el porcentaje de identificación de gustos básicos

Tabla 10-1 Reconocimiento de gustos básicos por el candidato a juez 1

Atributo	Acido		Amargo		Dulce		Salado	
	1	2	1	2	1	2	1	2
Repetición	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
2	0	1	0	1	1	0	1	1
3	1	0	1	1	0	1	1	0
4	1	0	1	1	0	1	1	1
% Reconocimiento	75	50	75	100	50	75	100	75
% Reconocimiento Total	62.5		87.5		62.5		87.5	

0= falla 1=acierto en reconocer el gusto básico

Cálculo para obtener el porcentaje de identificación de gustos básicos:

100% ----- Total de aciertos
x % ----- Número de aciertos

Para el caso del juez 1 en el atributo ácido réplica 1 obtuvo 3 aciertos lo máximo que pudo haber obtenido es 4, entonces:

$$x = (3 \cdot 100\%) / 4 = 75\%.$$

Después para cada juez se determinó un promedio de los porcentajes entre la repetición 1 y la repetición 2. Columna final de la tabla 10-1.

Ejemplo de cálculos para analizar los resultados de una prueba de comparación por pares mediante el estadístico Ji cuadrada :

En esta prueba se realizaron un total de 8 ensayos de acuerdo con la Tabla F1 en Pedrero y Pangborn (1997) el número mínimo de juicios correctos para indicar que la prueba es significativa son 7. Otra manera de determinar si el resultado de la prueba es significativa es utilizar la fórmula :

$$\chi^2 = \frac{(|X_1 - np - 0.5|)^2}{np(1-p)}$$

Ejemplo para el juez 1:

$X_1 = 3$ aciertos totales

$n = 8$ ensayos

$p = 0.5$ por ser una prueba de comparación por pares en donde existe la misma probabilidad de falla o acierto al tener solo dos opciones.

$q = (1 - p) = 0.5$

$n \cdot p = 8 \cdot 0.5 = 4$

Sustituyendo en la fórmula:

$$\chi^2 = \frac{(|3 - 8 \cdot 0.5| - 0.5)^2}{8 \cdot 0.5(0.5)} = 0.125$$

Para encontrar el valor teórico se busca en la Tablas de valores críticos de Ji cuadrada para una cola, (ya que la pregunta es direccionada ¿Cuál muestra es más ácida?) con un grado de libertad y una probabilidad de 0.05. El valor teórico es 2.71. Por lo que valores de Ji menores indica que el juez no esta discriminando entre que muestra es más ácida, valores mayores indican que el juez si esta discriminando a la muestra con mayor acidez.

10.3 EJEMPLO DE CÁLCULOS PARA LA DETERMINACIÓN DEL VALOR TEÓRICO DE LA PRUEBA DE FRIEDMAN Y EL PORCENTAJE DE MUESTRAS SIN REPETIR VALORES DE SUMA DE RANGOS.

Cálculos para el Criterio 4

Lo primero que se determina es si el juez encuentra diferencia al ordenar entre varias muestras, esto se analiza con la prueba de Friedman, como se puede observar en la tabla 10-2.

Tabla 10-2 Ejemplo de cálculos para la prueba de Friedman. Prueba Ordenación atributo Color por el juez 2

K	N				Suma de rangos	Suma de rangos ²
	1	2	3	4		
1	4	2	2	3	11	121.0
2	3	4	1	2	10	100.0
3	2	3	4	4	13	169.0
4	1	1	3	1	6	36.0
5	5	5	5	5	20	400.0
Suma total de rangos ² =						826.0
N(repeticiones)=						4.0
K(muestras)=						5.0
Friedman calculado=						10.6
Friedman teórico con $\alpha:0.05$, g.l.:k-1:7						9.49

Ejemplo de cálculo:

$$\chi_r^2 = \frac{12}{Nk(k+1)} \sum R^2 - 3N(k+1)$$

$$\chi_r^2 = \frac{12}{4 \cdot 5(5+1)} \sum (11^2 + 10^2 + 13^2 + 6^2 + 20^2) - 3 \cdot 4(5+1)$$

$$\chi_r^2 = \frac{12}{20 \cdot 6} \cdot 826 - 12 \cdot 6 = 10.6$$

A partir de 5 muestras ordenadas el valor de F de Friedman teórico se busca en Tablas de Valores Críticos de Ji cuadrada cuando se ordenan un número grande de muestras, con k-1 grados de libertad. Para números menores se utiliza la Tabla de probabilidades asociadas con la prueba de Friedman.

Cálculos para determinar el porcentaje de muestras sin repetir valores de suma de rangos.

Una vez que se clasificó al juez utilizando el Criterio 4 (prueba de Friedman), el siguiente paso es determinar en cuantas muestras no repite valores de suma de rangos, esto se hace mediante un porcentaje en donde el total de muestras equivale al 100%. El porcentaje quedaría de la siguiente manera:

100% ----- Total de muestras
x % ----- Número de muestras sin repetir valores de suma de rangos.

Entonces para el candidato a juez 2 no repitió ningún valor de suma de rangos, por lo que obtiene el 100% . Clasificando de acuerdo al Criterio 5 como "Muy bueno" en la ordenación del atributo *color*.

10.4 EJEMPLO DE CÁLCULOS UTILIZADOS EN LA CLASIFICACIÓN DE JUECES EN PRUEBAS SENSORIALES DE INTENSIDAD

Clasificación y cálculos para la familia aroma general

Aplicación del Criterio 7

Tabla 10-3 Análisis de varianza para el juez 1 atributo 1 aroma general

Producto (a=5)	Repetición 1	Repetición 2	Repetición 3 (n=3)	Total renglón	yi.2/n
1	5.00	6.92	6.20	18.12	328.33
2	7.00	3.00	6.20	16.20	262.44
3	1.08	5.00	3.48	9.56	91.39
4	3.00	1.00	3.96	7.96	63.36
5	8.92	8.92	5.48	23.32	543.82
Total columna	25.00	24.84	25.32	75.16	429.78
y..2/N				376.60	
SC total	163.73	162.45	134.70	460.89	84.29
SC producto	yi.2/n-y..2/N				53.18
SC réplica	625.00	617.03	641.10	376.63	0.02
SC error	Sct-scp-scr				31.08
g.l producto	a-1	a=5			4.0
g.l réplicas	n-1	n=3			2.0
g.l error	N-1-g.l.p-g.lr	N=a*n=15			8.0
g.l totales	N-1				14.0
Cm producto	SCp/g.lp				13.30
Cm réplica	SCr/g.l r				0.01
Cm error	SCe/g.l e				3.89
Fo productos	CMp/g.l e				3.42
Fo réplicas	CMr/ g.l e				0.00
Fo teórica productos:	0.05,4,8=3.84				
Fo teórica réplicas:	0.05,2,8=4.46				

Cálculos:

Factor de corrección $y_i..2/N =$ total por columnas y renglones entre el número total de respuestas

$$y_i..2/N = (75.16)^2 / (5 \cdot 3) = 5649.03 / 15 = 376.6$$

Suma de cuadrados para productos:

$$SC_p = [(18.12^2 + 16.2^2 + 9.56^2 + 7.96^2 + 23.32^2) / 3 - FC]$$

$$SC_p = 1289.35 / 3 - 376.6$$

$$SC_p = 429.78 - 376.6 = 53.18$$

$$\text{Grados de libertad para productos} = g \cdot l_p = a - 1 = 5 - 1 = 4$$

Suma de cuadrados de réplicas:

$$SC_r = [(25^2 + 24.84^2 + 25.32^2) / 5 - FC]$$

$$SC_r = 1883.13 / 5 - 376.6$$

$$SC_r = 376.626 - 376.6 = 0.02$$

$$\text{Grados de libertad de réplicas} = g \cdot l_r = n - 1 = 3 - 1 = 2$$

Suma de cuadrados totales:

$$SC_t = (5^2 + 6.92^2 + 6.2^2 + \dots + 8.92^2 + 8.92^2 + 5.48^2) - FC$$

$$SC_t = 460.89 - 376.6 = 84.29$$

$$\text{Grados de libertad totales: } g \cdot l_t = N - 1 = 15 - 1 = 14$$

Suma de cuadrados del error:

$$SC_e = SC_t - SC_p - SC_r$$

$$SC_e = 84.29 - 53.18 - 0.02 = 31.08$$

$$\text{Grados de libertad del error } g \cdot l_e = N - 1 - g \cdot l_p - g \cdot l_r = 14 - 4 - 2 = 8$$

Cuadrados medios :

$$CM_{\text{productos}} = SC_p / g \cdot l_p = 53.18 / 4 = 13.3$$

$$CM_{\text{réplicas}} = SC_r / g \cdot l_r = 0.02 / 2 = 0.01$$

$$CM_{\text{error}} = SC_e / g \cdot l_e = 31.08 / 8 = 3.89$$

F calculada:

$$F_{\text{productos}} = CM_p / CM_e = 13.3 / 3.89 = 3.42$$

$$F_{\text{réplicas}} = CM_r / CM_e = 0.02 / 3.89 = 0.00$$

F teóricas se busco en la Tabla Valores Críticos de F valores criticos para F:

$$F_{\text{productos}} \quad 0.05, 4/8 \quad g \cdot l = 3.84$$

$$F_{\text{réplicas}} \quad 0.05, 2/8 \quad g \cdot l = 4.46$$

Para el juez 1 no alcanza el valor de 3.84 por lo que se dice que no encuentra diferencia significativa entre los productos y por lo tanto no podrá ser considerado para la siguiente etapa de clasificación. Tampoco encuentra diferencia significativa entre réplicas, este efecto si es deseable ya que indica que el juez no califica de manera diferente entre repeticiones.

Tabla 10-4 Datos del ANOVA para el Juez 1 atributo aroma general

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F
Productos	4	53.18	13.3	28.04 ($F_{\alpha;0.05,4,8}=3.84$)
Réplicas	2	0.02	0.01	0.11 ($F_{\alpha;0.05,2,8}=4.46$)
Error	8	31.08	3.89	
Total	14	84.29		

Los resultados de la prueba de intensidad de cada juez para una familia de atributos fueron evaluados aplicando el Criterio 7, las tablas 10-5 a 10-7 presentan un resumen de los valores de F que cada juez obtuvo, así como el cuadrado medio del error. Estos datos son de utilidad para poder filtrar a todos aquellos candidatos que no logren discriminar entre las intensidades de los productos evaluados. También incluye información sobre la discriminación entre réplicas y como puede observarse en ningún caso hay diferencia significativa puesto que no se cubre el valor teórico 4.46. Para este caso familia en específico se evaluaron tres atributos.

Tabla 10-5 Resultados de los ANOVAS para atributo aroma general

Número de juez	F productos	F réplicas	CM error por juez
10	28.04	0.11	0.34
16	10.75	0.07	1.72
9	6.44	0	3.3
12	5.49	0.08	3.13
14	4.6	0.02	3.17
13	4.4	0.65	3.09
1	3.42	0	3.89
17	3.27	0.19	4.31
5	3.25	0.05	2.63
15	2.94	0.37	3.51
2	2.34	0	6.9
8	1.97	0.04	3.37
3	0.97	0.02	1.01
4	0.91	0.01	7.29
6	0.9	0.08	8.33
11	0.59	0.05	10.36
7	0.52	1.37	2.65
Promedio			4.06

Las gráficas 10-1 a 10-3 se realizaron con el fin de filtrar a los candidatos y determinar cuales podrían ser considerados en la aplicación de los siguientes criterios para pruebas de intensidad, como lo son el Criterio 8, Criterio 9 y Criterio 10. En esta

gráfica se observa que candidatos si discriminan entre las intensidades de productos al tener un valor calculado de F mayor al teórico (barra color azul), de igual manera permite observar los cuadrados medios del error de cada juez comparado con un cuadrado medio del error promedio.

Gráfica 10-1 Comparativo entre los valores de F calculada y el cuadrado medio del error para el atributo aroma general

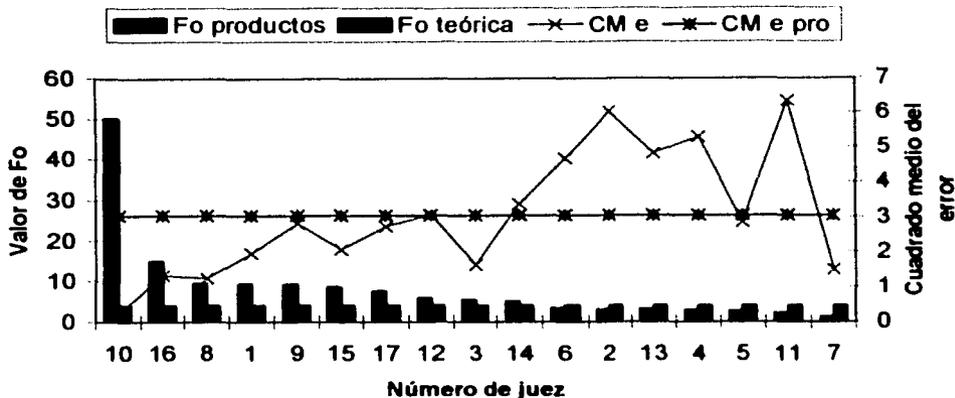


Tabla 10-6 Resultados de los ANOVAS para el atributo aroma a fruta.

Número de juez	F productos	F réplicas	CM error por juez
10	49.98	1.37	0.23
16	14.93	0.31	1.33
8	9.37	0.09	1.26
1	9.21	0.06	1.96
9	9.08	0.12	2.82
15	8.46	0.16	2.07
17	7.38	0.42	2.73
12	5.69	0.04	3.07
3	5.28	0.30	1.63
14	4.92	0.02	3.37
6	3.23	0.03	4.69
2	2.97	0.00	6.03
13	2.96	0.09	4.84
4	2.68	0.10	5.3
5	2.43	0.15	2.85
11	2.00	0.18	6.35
7	1.01	3.28	1.49
Promedio			3.06

Gráfica 10-2 Comparativo entre los valores de F calculada y el cuadrado medio del error para el atributo aroma frutal

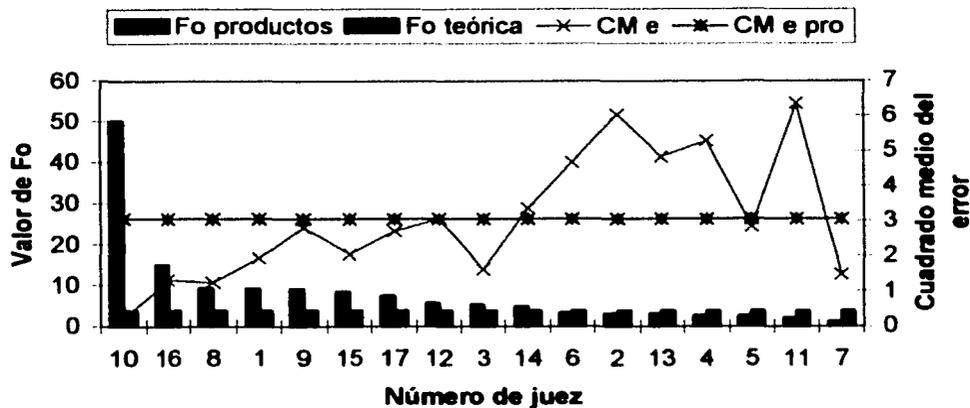
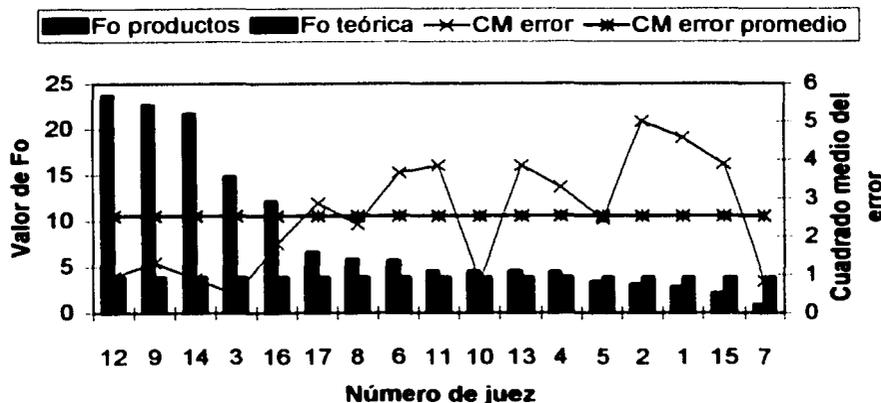


Tabla 10-7 Resultados de los ANOVAS para el atributo aroma a lácteo.

Número de juez	F productos	F réplicas	CM error por juez
12	23.63	0.25	0.95
9	22.59	0.74	1.32
14	21.63	0.36	0.91
3	14.8	0.42	0.49
16	12.08	0.35	1.81
17	6.58	0.14	2.87
8	5.8	0.05	2.32
6	5.62	0.02	3.67
11	4.52	0.6	3.86
10	4.48	2.54	0.8
13	4.47	0.33	3.86
4	4.4	0.18	3.29
5	3.33	0.05	2.46
2	3.01	0.46	5.01
1	2.73	0.02	4.58
15	2.1	0.84	3.9
7	0.86	4.23	0.83
Promedio			2.53

Gráfica 10-3 Comparativo entre los valores de F calculada y el cuadrado medio del error para el atributo aroma lácteo



Una vez que ya se tienen los resultados de los ANOVAS se puede aplicar el Criterio 8 y medir la frecuencia de discriminación por familia, puesto que ya se cuenta con información acerca de que candidatos discriminan entre intensidades de productos y cuales no. A continuación se explica como se determinaron los puntajes dados en los resultados para establecer las categorías de clasificación de este grupo de candidatos a jueces sensoriales.

Aplicación del Criterio 8

Para aplicar este criterio se considero al total de atributos como el 100% y en base a ese total se establecieron porcentajes para poder clasificar a un juez como "Muy bueno", "Bueno", "Regular", o "Malo". Estos porcentajes se establecen en la tabla 10-8. Como se estableció en este criterio cada una de las categorías tiene un puntaje dado. Si un juez es Muy bueno para discriminar en una determinada familia de atributos discriminara en por lo menos más del 75 % de atributos de los atributos de esa familia. Para el ejemplo de la familia de *aroma general* tenemos un total de 3 atributos (aroma general, aroma frutal y aroma lácteo) y ese total equivale al 100%.

Tabla 10-8 Porcentajes para establecer categorías de clasificación de los candidatos a jueces sensoriales

Clasificación	Porcentaje	Atributos por familia	Puntaje
Muy bueno	76-100%	3 atributos	4 puntos
Bueno	51-75%	2 atributos	3 puntos
Regular	26-50%	1 atributo	2 puntos
Malo	0-25%	0 atributos	1 punto

La siguiente tabla se obtuvo de los resúmenes de aplicar el Criterio 7 en la familia de aroma general dados en las tablas 10-5 a 10-7 se muestran los valores de Fo calculadas para cada candidato, cuando este valor es mayor al crítico (3.84) se considera que el candidato logró discriminar entre las intensidades de los atributos de aroma general evaluados.

Tabla 10-9 Candidatos que discriminan entre las intensidades de varios productos

Atributo	Jueces que discriminan entre las intensidades de acuerdo al Criterio 7
Aroma general	9,10,12,13,14,16
Aroma fruta	1,3,8,9,10,12,14,15,16,17
Aroma a lácteo	3,4,6,8,9,10,11,12,13,14,16,17

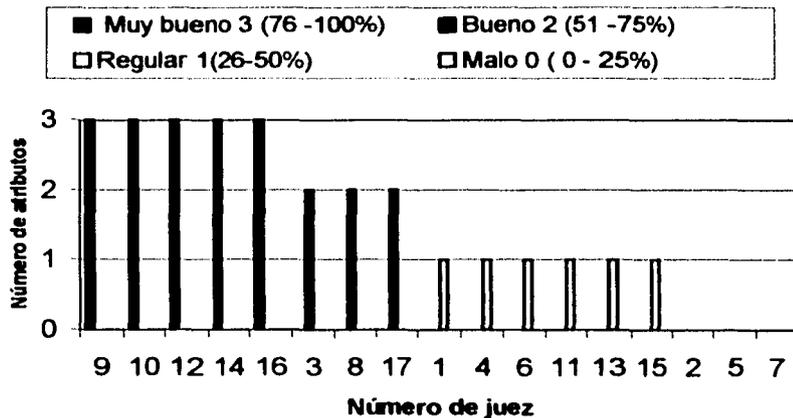
Tabla 10-10 Clasificación del grupo de candidatos a jueces de acuerdo al número total de atributos de la familia aroma general en donde lograron discriminar entre las intensidades de varios productos.

Número de juez/Clasificación	Malo 1	Regular 2	Bueno 3	Muy Bueno 4	Puntaje total
Atributos	0	1	2	3	
Porcentaje	0-25%	26-50%	51-75%	76-100%	
Puntaje máximo Criterio 8	9	6	3	12	
1		1			2
2	0				0
3			2		6
4		1			2
5	0				0
6		1			2
7	0				0
8			2		6
9				3	12
10				3	12
11		1			2
12				3	12
13		1			2
14				3	12
15		1			2
16				3	12
17			2		6

Para establecer el total de puntos obtenido por cada juez aplicando este criterio, primero se revisan los resultados del Criterio 7, es decir, se filtran a todos aquellos candidatos que sí discriminaron en las intensidades de varios productos. La tabla 10-9 permite ubicar en que juez el juez si discriminó. Por ejemplo el juez 10 discriminó en los tres atributos de esta familia, por que de acuerdo a la tabla 10-8 le corresponde la categoría de "Muy bueno" con un puntaje total de 12 puntos. Este puntaje se obtiene de multiplicar el número de atributos en donde el juez encontró diferencia significativa entre las intensidades de varios productos por el valor de la categoría (3 atributos por 4=12). La tabla 10-10 representa el número total de atributos en que cada juez logró diferenciar entre las intensidades de los productos, así como el puntaje total, mismo que se expresa gráficamente en los resultados.

Estos resultados permiten generar la gráfica mostrada en los resultados:

Gráfica 10-13 Clasificación de jueces considerando el Criterio 8 basado en la frecuencia de atributos en donde el juez encontraba diferencia significativa entre productos analizando los atributos de aroma general



Aplicación del Criterio 9

En esta parte los candidatos que no discriminaron entre las intensidades de productos son considerados directamente en la categoría "Malo". Una vez que ya se establecieron los puntajes que cada juez obtuvo en el Criterio 8, se determinan los puntajes para el Criterio 9 en donde se requiere conocer el promedio de cada evaluación de intensidad por cada producto, la desviación estándar y el coeficiente de variación, que realizó cada candidato. Al final se determina un promedio total de desviación estándar y coeficiente de variación por cada juez, mismo que se muestra al final de la tabla 10-11:

Tabla 10-11 Resultados de la prueba de intensidad de aroma general para el juez 10

Producto	Réplica 1	Réplica 2	Réplica 3	Promedio	s	CV	t student
1	7.00	8.44	7.48	7.64	0.73	9.60	teórica
2	6.52	7.00	6.44	6.65	0.30	4.55	0.05,16 g.l.
3	3.96	2.36	3.48	3.27	0.82	25.13	2.12
4	4.44	3.96	3.96	4.12	0.28	6.73	Calculada
5	5.32	5.00	5.00	5.11	0.18	3.62	-14.36
Promedio juez					0.46	9.93	

Primero se necesita obtener los promedios para cada producto, así como su desviación estándar. Ejemplo

$$\text{Producto 1} = (7 + 8.44 + 7.48) / 3 = 7.64$$

Cálculo de desviación estándar

$$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}}$$

Sustituyendo en la fórmula

$$s = \sqrt{\frac{\sum (7^2 + 8.44^2 + 7.48^2) - \frac{(\sum 7.64)^2}{3}}{3-1}} = 0.73$$

Coefficiente de variación = (promedio / s) * 100 para el ejemplo del juez 10, atributo aroma general, producto 1 se tiene que es igual a (7.64/0.73)*100=9.6

Se realizó lo mismo con los valores de cada candidato y para toda la familia de atributos. La siguiente tabla presenta los promedios de coeficiente de variación que cada juez obtuvo en los atributos de la familia aroma general en el fondo de la tabla se muestra el valor del coeficiente de variación del grupo de candidatos.

Tabla 10-12 Valores de coeficiente de variación individual por atributos de la familia aroma general

Número de juez	CV aroma general	CV aroma a frutal	CV aroma lácteo
1	39.95	29.77	40.34
2	60.4	33.47	45.22
3	20.2	26.53	12.57
4	51.55	47.37	36.05
5	27.05	25.64	29.45
6	57.22	28.93	39.62
7	32.02	26.38	19.68
8	33.36	18.21	26.48
9	41.23	34.19	33.27
10	9.93	9.07	23.9
11	55.64	44.01	42.66
12	36.43	38.21	26.23
13	39.26	50	38.36
14	35.68	34.52	26.06
15	21.61	28.52	45.54
16	27.12	17.96	35.14
17	42.92	35.45	38.51
CV grupal	37.15	31.07	32.89

En el Criterio 9 se determina si existe diferencia significativa entre el coeficiente de variación del grupo y el de cada juez, para poder realizar esta comparación se utilizó el estadístico t de student con la siguiente fórmula:

$$t_{student} = \frac{CV_{juez} - CV_{grupal}}{S_{juez} \sqrt{N}}$$

Donde:

N = número total de la población que en nuestro caso son 17 jueces,
 S = desviación estándar de la muestra (en este caso cada juez es una muestra)
 CV juez = coeficiente de variación del candidato a juez y
 CV grupal = promedio de coeficiente de variación del grupo (promedio de los valores de los 17 candidatos en un atributo).

A continuación se expone el caso del candidato a juez número 10, en el atributo aroma general, en donde sustituyendo en la fórmula se obtiene el siguiente valor de t de student.

$$t_{student} = \frac{9.93 - 37.15}{0.46 \sqrt{17}} = -14.36$$

De acuerdo con la clasificación planteada en la figura 5-1 este juez se ubica en la categoría "Muy bueno", puesto que el valor calculado de t de student mucho menor que el valor crítico (-2.12), lo que indica que su coeficiente de variación es mucho menor al grupal, de acuerdo a lo establecido en el Criterio 9.

Esta operación se aplicó a todos los candidatos en todos los atributos de una familia y para cada juez en la tabla 10-13 se establecen los valores de t de student que cada juez obtuvo en los atributos de la familia aroma general.

Tabla 10-13 Valores de t de student por cada juez en la familia de atributos aroma general

Número de juez	Valor de t student aroma general	Valor de t student aroma frutal	Valor de t student aroma lácteo
1	0.40	-0.29	1.05
2	2.17	0.30	1.69
3	-4.99	-0.95	-7.67
4	1.48	2.00	0.51
5	-1.85	-0.95	-0.61
6	1.94	-0.32	1.16
7	-0.76	-0.85	-3.16
8	-0.57	-3.50	-1.39
9	0.62	0.60	0.08
10	-14.25	-13.37	-2.25
11	1.81	1.52	1.41
12	-0.12	1.14	-1.95
13	0.32	2.45	0.79
14	-0.24	0.55	-2.01
15	-3.18	-0.54	1.62
16	-2.30	-3.66	0.46
17	0.76	0.73	0.92
Valores límites de t de student: -2.12 a 2.12			

Esta información se analiza de acuerdo a lo planteado en el Criterio 9 (ver Tabla 4-22), clasificando a cada juez de acuerdo al valor de t de student calculado. Por ejemplo el juez 1 en el atributo aroma general obtuvo un valor de 0.40 de acuerdo al Criterio 9 se ubica en el rango de $0 < t_{cal} < 2.12$, lo que lo hace clasificar en la categoría de "Regular". La tabla 10-14 ubica a cada uno de los candidatos de acuerdo a este análisis.

Tabla 10-14 Clasificación de jueces de acuerdo al Criterio 9

Clasificación	Malo (1)	Regular (2)	Buena (3)	Muy buena (4)
Valor de t student	$t_{cal} > 2.12$	$0 < t_{cal} < 2.12$	$-2.12 < t_{cal} < 0$	$t_{cal} < -2.12$
Significado	1) $CV_{juez} > CV_{grupo}$ 2) $F_{cal} < F_{teo}$	$Cv_{juez} \geq CV_{grupo}$	$CV_{juez} \leq CV_{grupo}$	$CV_{juez} < CV_{grupo}$
Puntaje máximo Criterio 9	3	6	9	12
Atributos	Jueces clasificados			
Aroma general	2, 1, 3, 4, 5, 6, 7, 8, 11, 15, 17	9, 13	12, 14	10, 15, 16
Aroma fruta	2, 4, 5, 6, 7, 11, 13	9, 12, 14, 17	1, 3, 15	8, 10, 16
Aroma lácteo	1, 2, 5, 7, 15	4, 6, 9, 11, 13, 15, 16, 17	8, 12, 14	3, 10

Una vez que ya se tiene la clasificación de cada juez por atributo de familia, se pueden establecer los puntajes, en este caso se analiza la frecuencia en donde un candidato aparece en la misma clasificación. Esta frecuencia se multiplica por el valor de la clasificación, al final se suman todos los puntos obtenidos. Por ejemplo para el candidato 1: 0 veces en la categoría de "Muy bueno", 1 vez en "bueno", 0 en "Regular" y 2 veces en "Malo", genera un total de 5 puntos obtenidos de la siguiente manera $(0*4)+(1*3)+(0*2)+(1*2)=5$

Tabla 10-15 Frecuencias de clasificación de los candidatos a jueces para obtener el puntaje total aplicando el Criterio 9.

Clasificación	Malo (1)	Regular (2)	Buena (3)	Muy Buena (4)	Puntos
Valor de t student	$t_{cal} > 2.12$	$0 < t_{cal} < 2.12$	$-2.12 < t_{cal} < 0$	$t_{cal} < -2.12$	
Significado	1) $CV_{juez} > CV_{grupo}$ 2) $F_{cal} < F_{teo}$	$Cv_{juez} \geq CV_{grupo}$	$CV_{juez} \leq CV_{grupo}$	$CV_{juez} < CV_{grupo}$	
Puntaje máximo Criterio 9	3	6	9	12	
Número de juez	Jueces clasificados				
1	2		1		5
2	3				3
3	1		1	1	8
4	2	1			4
5	3				3
6	2	1			4
7	3				3
8	1		1	1	8
9		3			6
10				3	12
11	2	1			4
12		1	2		8
13	1	2			5
14		1	2		8
15	2		1		5
16		1		2	10
17	1	2			5

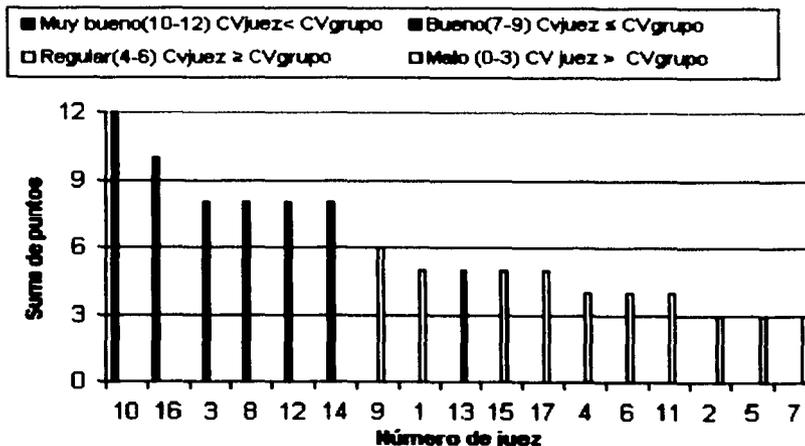
Finalmente el puntaje de clasificación se distribuye de acuerdo a los valores máximos, este se establece en la tabla 10-16.

Tabla 10-16 Clasificación final de los candidatos a jueces sensoriales de acuerdo al Criterio 9

Clasificación	Malo (1)	Regular (2)	Bueno (3)	Muy Bueno (4)	Puntos Totales
Valor de t student	$t_{cal} > 2.12$	$0 < t_{cal} < 2.12$	$-2.12 < t_{cal} < 0$	$t_{cal} < -2.12$	
Significado	1) CV juez > CV grupo 2) F cal < F teo	Cv juez \geq CV grupo	CV juez \leq CV grupo	CV juez < CV grupo	
Puntaje	0-3	4-6	7-9	10-12	
1		5			5
2	3				3
3			8		8
4		4			4
5	3				3
6	4				4
7	3				3
8			8		8
9		6			6
10				12	12
11	4				4
12			8		8
13		5			5
14			8		8
15		5			5
16				10	10
17		5			5

Con estos datos se generó la gráfica presentada en los resultados.

Gráfica 10-14 Jueces clasificados según desempeño en atributos aroma general utilizando el Criterio 9 (comparación del coeficiente de variación de cada juez respecto al grupal mediante el estadístico t de student)



Como paso final se suman ambos puntajes de acuerdo al Criterio 10 teniendo:

Tabla 10-17 Puntajes finales Criterio 10 (Suma de puntos Criterios 8 y 9)

Clasificación				Malo	Regular	Bueno	Muy Bueno
Puntaje máximo Criterio 10				0-6	7-12	13-18	19-24
Número de juez	Puntaje Criterio 8	Puntaje Criterio 9	Puntaje Criterio 10	Jueces clasificados			
1	2	5	7		7		
2	0	3	3	3			
3	6	8	14			14	
4	2	4	6	6			
5	0	3	3	3			
6	2	4	6	6			
7	0	3	3	3			
8	6	8	14			14	
9	12	6	18			18	
10	12	12	24				24
11	2	4	6	6			
12	12	8	20				20
13	2	5	7		7		
14	12	8	20				20
15	2	5	7		7		
16	12	10	22				22
17	6	5	11	11			

Con estos valores se creo la última gráfica correspondiente a la clasificación de la familia aroma general.

Gráfica 10-15 Clasificación de jueces para pruebas de intensidad en atributos aroma general utilizando el Criterio 10

