

## VALIDACIÓN TÉCNICA MÓDULO CATA

### **Validación Técnica y Metodológica de un Módulo de Análisis CATA (Check-All-That-Apply) en la Plataforma Brios**

Equipo de Investigación y Desarrollo, Brios Análisis Sensorial

Departamento de I+D, Brios.com.uy

#### **NOTA DEL AUTOR**

Este estudio fue financiado por Brios Análisis Sensorial. No existen conflictos de interés adicionales más allá de la validación interna del software propietario. La correspondencia relativa a este artículo debe dirigirse a ventas@brios.com.uy.

## RESUMEN

El método *Check-All-That-Apply* (CATA) se ha convertido en una herramienta estándar en la evaluación sensorial rápida (ISO 13299:2016), y la fiabilidad de sus resultados depende de la precisión de los algoritmos estadísticos subyacentes. El presente estudio tuvo como objetivo validar técnicamente el módulo de análisis CATA de la plataforma Brios, verificando la exactitud de sus cálculos de frecuencias, pruebas de significancia no paramétricas y análisis multivariado. Metodológicamente, se diseñó un "Golden Dataset" sintético con  $N = 50$  panelistas y tres productos prototípicos con perfiles sensoriales ortogonales (Dulce, Salado, Crocante), cuyos datos fueron injectados directamente en la base de datos y procesados por el motor híbrido de la plataforma. Los resultados mostraron que el sistema identificó correctamente las diferencias significativas en los tres atributos clave (Prueba Q de Cochran,  $p < .05$ ), mientras que las pruebas *post-hoc* de McNemar agruparon los productos según el diseño experimental esperado. Asimismo, el Análisis de Correspondencia (CA) generó un mapa perceptual con una inercia explicada del 100% en dos dimensiones, reflejando fielmente la estructura de los datos. En conclusión, el módulo CATA de la plataforma cumple con los requisitos estadísticos y normativos para su uso en entornos de investigación y control de calidad.

**Palabras clave:** Análisis Sensorial, CATA, Validación de Software, Q de Cochran, Análisis de Correspondencia.

## INTRODUCCIÓN

La digitalización de los procesos de análisis sensorial requiere herramientas que no solo recolecten datos eficientemente, sino que ejecuten análisis estadísticos complejos de manera autónoma y fiable. El método CATA (*Check-All-That-Apply*), descrito en la norma **ISO 13299:2016**, permite obtener perfiles sensoriales rápidos preguntando a los consumidores qué atributos perciben en una muestra.

A diferencia de las escalas de intensidad (RATA), los datos CATA son binarios (0/1), lo que requiere un tratamiento estadístico específico: pruebas no paramétricas para muestras relacionadas (Cochran) y técnicas de reducción de dimensiones basadas en frecuencias (Análisis de Correspondencia).

Este informe documenta el protocolo de validación técnica ("Sanity Check") aplicado al módulo CATA de la plataforma *Brios*, utilizando un conjunto de datos controlado para asegurar que el software no genera falsos positivos ni falsos negativos.

## HIPÓTESIS

Si el motor de análisis de la plataforma funciona correctamente, al procesar un set de datos con diferencias extremas pre-programadas ("Diferencia Obvia"), el sistema deberá:

- Calcular las frecuencias de mención con exactitud absoluta respecto a la base de datos SQL.
- Detectar diferencias significativas en los atributos discriminantes mediante la prueba Q de Cochran.
- Generar letras de grupos homogéneos distintos para productos opuestos mediante la prueba de McNemar.
- Producir un Mapa Perceptual donde los productos se ubiquen en vértices opuestos, maximizando la distancia entre perfiles sensoriales distintos.

## MATERIALES Y MÉTODOS

### Diseño Experimental (In-Silico)

Se simuló un panel de consumidores ( $N=50$ ) evaluando 3 productos teóricos diseñados para tener perfiles sensoriales no superpuestos:

- **Producto A (Dulce):** Diseñado para tener alta frecuencia en "Dulce".
- **Producto B (Salado):** Diseñado para tener alta frecuencia en "Salado".
- **Producto C (Crocante):** Diseñado para tener alta frecuencia en "Crocante".

### Generación de Datos

Los datos fueron generados mediante un script de inyección SQL con las siguientes probabilidades condicionadas para asegurar la discriminación:

Atributo	Prob. en Prod A	Prob. en Prod B	Prob. en Prod C
<b>Dulce</b>	90%	10%	20%
<b>Salado</b>	10%	90%	20%
<b>Crocante</b>	20%	20%	90%

## PROCESAMIENTO DE DATOS

Los datos crudos se almacenaron en una base de datos relacional (MariaDB) y fueron procesados por el motor de análisis de *Brios*, el cual utiliza bibliotecas científicas para la ejecución de los algoritmos.

### **Análisis Estadístico Aplicado**

La plataforma aplicó automáticamente el siguiente flujo de análisis:

- **Tabulación de Frecuencias:** Sumatoria de citas por atributo y producto.
- **Prueba Q de Cochran:** Para determinar si existen diferencias globales entre productos para cada atributo ( $\alpha = 0.05$ ).
- **Prueba de McNemar (Bonferroni):** Comparaciones múltiples por pares para identificar diferencias específicas entre productos (asignación de letras).
- **Análisis de Correspondencia (CA):** Visualización de la tabla de contingencia en un espacio de baja dimensión basada en la distancia Chi-cuadrado.

## RESULTADOS

### **Frecuencias de Mención**

El análisis de la base de datos arrojó los siguientes conteos consolidados, los cuales coincidieron exactamente con la salida del reporte generado por la plataforma.

**Tabla 1. Tabla de contingencia observada (Panelistas que marcaron el atributo).**

Atributo	Prod A (Dulce)	Prod B (Salado)	Prod C (Crocante)	p-valor (Cochran)
Crocante	9	10	47	< 0.001 *
Dulce	46	3	11	< 0.001 *
Salado	5	43	8	< 0.001 *

### **Análisis Inferencial**

La prueba Q de Cochran indicó diferencias altamente significativas en los tres atributos, marcados con un asterisco (\*) en el reporte automático.

El análisis *post-hoc* (McNemar) generó los siguientes grupos homogéneos:

- Para el atributo **Crocante**, el Producto C (47 menciones) fue asignado al grupo a, diferenciándose estadísticamente de los Productos A y B (grupos bc y b respectivamente).
- Para el atributo **Dulce**, el Producto A (46 menciones) se aisló en el grupo a, separado de B y C.
- Para el atributo **Salado**, el Producto B (43 menciones) lideró el grupo a.

### Mapa Perceptual (Análisis de Correspondencia)

El Mapa Perceptual resultante muestra una distribución triangular perfecta, lo cual es geométricamente consistente con la matriz de datos ortogonal inyectada.

- **Dimensión 1 (61.87%)**: Discrimina principalmente entre los atributos "Dulce" (izquierda) y "Salado" (derecha), separando claramente al Producto A del Producto B5.
- **Dimensión 2 (38.13%)**: Discrimina el atributo "Crocante" (superior), aislando al Producto C del resto.

La inercia total explicada por las dos primeras dimensiones es del 100% (61.87% + 38.13%), confirmando que la representación visual no tiene pérdida de información para este conjunto de datos.

## DISCUSIÓN

Los resultados obtenidos validan integralmente la lógica de procesamiento del sistema.

- **Integridad de Datos**: No hubo discrepancias entre los datos inyectados vía SQL y los presentados en las tablas de frecuencia del reporte PDF.
- **Sensibilidad Estadística**: El sistema detectó correctamente las diferencias inducidas. En escenarios donde la diferencia de frecuencia es amplia (ej. 46 vs 3), el algoritmo asignó correctamente letras diferentes (a vs bc), lo cual confirma la correcta implementación de la corrección de Bonferroni para evitar falsos positivos en comparaciones múltiples.
- **Coherencia Visual**: La proyección espacial en el mapa de correspondencia ubica a cada producto en la proximidad inmediata de su atributo dominante (Prod A cerca de "Dulce", Prod B cerca de "Salado"), validando la implementación de la distancia Chi-cuadrado en el algoritmo de reducción de dimensiones.

## CONCLUSIONES

La plataforma *Brios* ha superado satisfactoriamente la validación técnica para el módulo de análisis CATA. El sistema demuestra capacidad para procesar, analizar y reportar datos categóricos cumpliendo con los lineamientos estadísticos de la norma ISO 13299:2016 y métodos quimiométricos estándar. La herramienta es apta para su despliegue en estudios de consumidores reales, garantizando resultados precisos y reproducibles.

## REFERENCIAS

- Cochran, W. G. (1950). The comparison of percentages in matched samples. *Biometrika*, 37(3/4), 256–266. <https://doi.org/10.2307/2332378>
- Greenacre, M. (2017). *Correspondence analysis in practice* (3rd ed.). CRC Press. <https://doi.org/10.1201/9781315369983>
- International Organization for Standardization. (2016). *Sensory analysis — Methodology — General guidance for establishing a sensory profile* (ISO Standard No. 13299:2016). <https://www.iso.org/standard/55282.html>
- Meyners, M., Castura, J. C., & Carr, B. T. (2013). Existing and new approaches for the analysis of CATA data. *Food Quality and Preference*, 30(2), 309–319. <https://doi.org/10.1016/j.foodqual.2013.06.010>