



NOTA INTERPRETATIVA Nº 2 SOBRE CIERTOS ASPECTOS DEL MÉTODO DE VALORACIÓN ORGANOLÉPTICA DEL ACEITE VIRGEN DE OLIVA DEL CONSEJO OLEÍCOLA INTERNACIONAL

EL ÍNDICE DE DESVIACIÓN

Una de las herramientas más habitualmente utilizadas para el control de calidad interno del rendimiento de los catadores de un panel, son los denominados índices de desviación y de repetibilidad, que se encuentran explicados en el documento **COI/T.28/Doc. nº 1** de Septiembre de 2007, denominado “Directrices para la acreditación de los laboratorios de análisis sensorial de aceite de oliva virgen en particular, según la norma ISO/IEC 17025:2005” (www.internationaloliveoil.org/estaticos/view/224-testing-methods)

Dichos índices pretenden evaluar por un lado, la consistencia de los juicios de un juez cuando éste realiza duplicados de una muestra en modo ciego, **índice de repetibilidad**, y por otro, cuánto se desvía de la mediana del grupo, **índice de desviación**. Por un lado se mide la repetibilidad del catador y por otro la “veracidad interna”, si se nos permite la expresión, tomando como valor de referencia interna la mediana del panel.

Estos índices deben ser calculados durante un período suficiente de tiempo, seis (6) meses como mucho, de acuerdo a la guía antes mencionada, y un número de duplicados que dependerá de la carga de trabajo del panel en dicho período, pero que como mínimo deberán ser seis, con objeto de poder evaluar al catador durante los diferentes estados psicofisiológicos por los que de manera natural va transitando.

La fórmula del **índice de repetibilidad** es la siguiente:

$$IR_p = 1 + \frac{\sum_{i=1}^n (x_{i1} - x_{i2})^2}{n}$$

La interpretación de la fórmula del índice de repetibilidad se realizó en la nota interpretativa nº 1, donde se indicaba que los valores x_{i1} y x_{i2} son los valores del catador i en el replicado 1 y en el replicado 2 respectivamente. Puesto que para obtener una diferencia son necesarios dos análisis en modo ciego, n es el número de muestras efectivamente analizadas, que para un duplicado son dos. El valor límite de este índice es 3.0 para una serie de duplicados válidos (que cumplan un criterio de aceptación) realizados en un período máximo de seis meses, como se ha indicado.



El **índice de desviación** lo que evalúa es el comportamiento del catador en relación no consigo mismo, sino con su grupo de compañeros, el panel.

Su fórmula es la siguiente:

$$ID = 1 + \frac{\sum_{i=1}^n [(x_{i1} - \bar{x}_{i1})^2 + (x_{i2} - \bar{x}_{i2})^2]}{2n}$$

Donde x_{i1} y x_{i2} son los valores individuales del panelista i para las muestras 1 y 2 y \bar{x}_{i1} y \bar{x}_{i2} son las medianas del panel correspondientes a las mismas muestras. El subíndice i que aparece en las medianas no parece tener sentido, ya que este subíndice indica el catador correspondiente, al igual que sucede en el IR_p . N es el número de muestras analizadas.

Sin embargo, en esta situación, para obtener una diferencia sólo es necesario el análisis de una muestra y puesto que el COI da el mismo peso específico a ambos índices, 3.0, es necesario que el denominador esté multiplicado por dos para poderlo conseguir.

Se ha interpretado en muchos ámbitos que las dos diferencias que aparecen en el numerador del índice de desviación están referidas a los valores obtenidos en los duplicados. De acuerdo a esta interpretación, veamos un ejemplo de cálculo de ambos índices con pocas muestras. Se dan los siguientes resultados de tres replicados en una tabla:

	x_1	x_2	Me 1	Me 2
Replicado 1	2.6	3.1	3.5	3.9
Replicado 2	4.5	4.0	5.2	4.7
Replicado 3	3.6	4.3	4.2	5.0

El cálculo de los índices, de acuerdo a los resultados de la tabla sería el siguiente:

$$IR_p = 1 + \frac{(2.6 - 3.1)^2 + (4.5 - 4.0)^2 + (3.6 - 4.3)^2}{6} = 1.2$$

$$ID = 1 + \frac{(2.6 - 3.5)^2 + (3.1 - 3.9)^2 + (4.5 - 5.2)^2 + (4.0 - 4.7)^2 + (3.6 - 4.2)^2 + (4.3 - 5.0)^2}{12} = 1.3$$

Se puede observar que el número de sumandos en el índice de desviación es el doble que en el índice de repetibilidad, lo que podría traducirse en que es necesaria



mucha más información para el cálculo del ID que para el IR_p , lo cual, en principio, no tiene por qué ser así al no estar fundamentado.

Existe otra interpretación para realizar el cálculo del ID de forma diferente e igualmente válida, que va a proporcionar resultados similares pero con un cálculo más sencillo. Esta interpretación del cálculo está basada en el siguiente planteamiento de partida. Se sabe que el ID lo que pretende es conocer el comportamiento del catador respecto del grupo a lo largo del tiempo, de acuerdo a lo indicado por el documento COI/T.28/Doc. nº 1. Si para calcular el ID se introducen todos los sumandos debidos a los replicados, como se acaba de realizar, se estarán introduciendo en las diferencias una parte de sesgo debida a las repeticiones, algo que no se pretende, como se indica en el documento COI, puesto que el sesgo debido a éstas ya ha sido calculado con el IR_p . Por otra parte, en el citado documento del COI, en el apartado **(5.9) Garantía de la calidad de los resultados del ensayo y de la calibración**, no se indica ni explícita ni implícitamente que x_{i1} y x_{i2} provengan de un duplicado necesariamente, ni que \bar{x}_{i1} y \bar{x}_{i2} sean las dos medianas de un duplicado. Lo que es seguro es que esos valores provienen de dos muestras distintas, denominadas 1 y 2. De acuerdo al planteamiento indicado, se puede calcular el ID con la mitad de los sumandos, siempre y cuando se indique previamente qué valor del duplicado válido se va a utilizar para hacer el cálculo del ID . Si, por ejemplo, se indica que para el cálculo se va a utilizar siempre el primer valor del replicado, los resultados para ambos índices y la tabla de valores indicada, serían los siguientes:

$$IR_p = 1 + \frac{(2.6 - 3.1)^2 + (4.5 - 4.0)^2 + (3.6 - 4.3)^2}{6} = 1.2$$

$$ID = 1 + \frac{(2.6 - 3.5)^2 + (4.5 - 5.2)^2 + (3.6 - 4.2)^2}{6} = 1.3$$

De acuerdo a todo el planteamiento expuesto, se considera que el ID también se puede calcular válidamente con la siguiente fórmula:

$$ID = 1 + \frac{\sum_{i=1}^n |(x_{i1} - Me_1)^2 + (x_{i2} - Me_2)^2 + (x_{i3} - Me_3)^2 + \dots|}{2n}$$

LA COMPARACIÓN DE DOS MEDIDAS: EL ÍNDICE DE COMPATIBILIDAD

Cuando se realizan sesiones de cata es necesario tener definido un criterio de aceptación de los resultados que en ella se producen con objeto de poder validarlas. Básicamente, las herramientas más utilizadas en el control interno cuando se realizan sesiones de cata son las siguientes:



- Análisis de una muestra duplicada y comparación de ambos resultados.
- Análisis de una muestra ya analizada y comparación de los resultados con los obtenidos previamente.
- Análisis de un material de referencia y comparación de los resultados con los de referencia.

Para validar la sesión será necesario tener algún criterio que permita definir con claridad si los análisis realizados a estas muestras son correctos o no. Tal y como está concebido el actual método de valoración organoléptica, que lo único que pretende es clasificar un aceite en función de las medianas del defecto y de la del frutado, no lo olvidemos, es complicado aseverar que, en base al análisis realizado a las dos muestras que conforman el duplicado, éstas sean la misma, sobre todo, como se ha indicado con anterioridad, cuando no se realiza un perfil descriptivo cuantitativo completo. Sin embargo, las técnicas estadísticas sí nos permiten indicar si las muestras han sido analizadas de la misma forma.

Una idea sencilla para comparar las dos medidas sería el estudiar cuán próximos están los valores obtenidos en ambos análisis que, en el caso que nos ocupa, son las medianas obtenidas para el defecto mayormente percibido. Es evidente que si se desea hacer ese estudio de forma rigurosa hay que tener en consideración las incertidumbres expandidas (U_1 y U_2) obtenidas en cada una de las medidas y buscar algún tipo de criterio que permita decidir si las medianas están lo suficientemente alejadas entre sí como para decidir que las dos muestras son distintas. Si ese criterio no se cumple, se podrá indicar que las muestras no son distintas y decir, por lo tanto, que las muestras han sido analizadas de la misma forma por el panel y así validar la sesión de cata en consecuencia.

De acuerdo a la ley de aditividad de las incertidumbres, la incertidumbre de una diferencia de valores es la suma de sus incertidumbres, de modo que se puede indicar que $U(x_1 - x_2) = \sqrt{U_1^2 + U_2^2}$ donde U_1 y U_2 son las incertidumbres individuales. Se utiliza la media geométrica puesto que la incertidumbre se expresa en forma de desviaciones estándar y no varianzas.

Conocida la diferencia de valores y conocida su incertidumbre queda por conocer cuál sería el criterio a utilizar para admitir que dos muestras son distintas, para lo cual tenemos que realizar un supuesto de partida y es que ambas medidas son independientes, los resultados individuales se agrupan bajo una distribución normal y que van a tener varianzas parecidas. De forma sencilla y conceptualmente comprensible, se puede decir que dos muestras son distintas, es decir, han sido analizadas de forma diferente, si la incertidumbre de su diferencia es inferior a la diferencia entre ellas, es decir, si $\sqrt{U_1^2 + U_2^2} < (x_1 - x_2)$, lo cual desde un punto de vista geométrico indica que las incertidumbre individuales de cada valor no se tocan.

En la figura 1 se puede observar los dos valores y sus incertidumbres, una en azul y otra en verde, donde se puede advertir que la diferencia de valores $x_1 - x_2$,

representado por la línea roja horizontal es superior a la incertidumbre de la diferencia, representada por el rectángulo marrón. Es evidente que las incertidumbres individuales no se tocan y la conclusión es que las medidas están muy separadas, sin que haya elementos en común en sus incertidumbres que permitan la duda. En esta situación se puede decir que las medidas son distintas y la valoración que ha hecho el panel de ambas muestras es diferente.

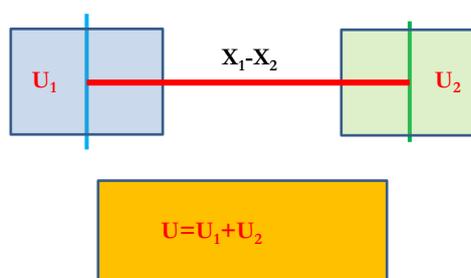


Figura 1

Veamos ahora la situación inversa, es decir, que $\sqrt{U_1^2 + U_2^2} > (x_1 - x_2)$

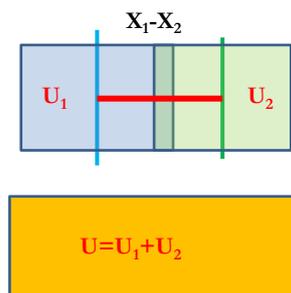


Figura 2

En esta nueva situación se puede observar que la incertidumbre de la diferencia, rectángulo marrón, es superior a la diferencia de valores, línea roja, lo cual significa que las incertidumbres individuales están solapadas, como puede observarse en la figura 2. En esta nueva situación, ¿se puede afirmar que las dos medidas realizadas son distintas? Evidentemente no, pues hay una probabilidad de que ambas medidas sean la misma, que es la zona solapada. ¿Son iguales las medidas? Eso no puede ser asegurado, pero sí que puede afirmarse que no son distintas, lo cual nos puede servir de criterio de aceptación de los valores de un duplicado, en un caso como el de los paneles de catadores.

¿Cuál será la situación límite que aún pueda generar dudas sobre si las medidas son iguales o distintas? Cuando las incertidumbres individuales estén sólo en contacto, lo cual se producirá cuando $\sqrt{U_1^2 + U_2^2} = (x_1 - x_2)$.



De acuerdo a lo indicado se puede definir el índice de compatibilidad como un estadístico que nos permite comparar los resultados de dos análisis en base a los valores obtenidos y a las incertidumbres asociadas a ambos. Se define de la siguiente forma:

$$IC = \frac{|x_1 - x_2|}{\sqrt{U_1^2 + U_2^2}}$$

Donde x_1 y x_2 son los valores obtenidos en el ensayo 1 y en el ensayo 2 respectivamente, y U_1 y U_2 las incertidumbres expandidas asociadas a dichos ensayos. Podremos afirmar, por lo tanto, que dos medidas son compatibles cuando el IC sea igual o inferior a uno. Es evidente que cuánto más pequeño sea el IC más próximas estarán las medidas y la comparación será mejor. El valor límite será uno, cuando la diferencia de valores sea igual a su incertidumbre. Por lo tanto,

$$IC = \frac{|x_1 - x_2|}{\sqrt{U_1^2 + U_2^2}} \leq 1.0$$

Veamos un ejemplo aplicado a un panel de catadores. En la siguiente tabla se encuentran los resultados obtenidos en dos análisis diferentes. Comprobar si ambos resultados son compatibles.

x_1	CV_1 (%)	x_2	CV_2 (%)
4.6	7.9	5.3	5.6

De acuerdo a lo indicado en el método de valoración organoléptica del Consejo Oleícola Internacional (COI/T.20/Doc. No 15/Rev. 6 Noviembre 2013), los intervalos de confianza de la medida, su incertidumbre, viene dada por cxs^* , siendo c el factor 1.96 para obtener la incertidumbre expandida y s^* la desviación típica sólida. Ésta se calcula de la siguiente manera:

$$CV_r = \frac{s^*}{Me} \times 100 \rightarrow s^* = \frac{CV_r \times Me}{100}$$

Siendo CV_r el coeficiente de variación robusto y Me la mediana del panel. Por lo tanto, la incertidumbre de la medida es:

$$U = 1.96 \times \frac{CV_r \times Me}{100}$$



El índice de compatibilidad será entonces:

$$IC = \frac{|4.6 - 5.3|}{\sqrt{\left(\frac{1.96 \times 7.9 \times 4.6}{100}\right)^2 + \left(\frac{1.96 \times 5.6 \times 5.3}{100}\right)^2}} = 0.72$$

Puesto que 0.72 es inferior a uno, se puede concluir que las dos medidas son compatibles, es decir, que no son distintas, por lo que el análisis del duplicado puede considerarse correcto.

Si en lugar de haber dado los resultados con el coeficiente de variación robusto se hubieran dado las incertidumbres del duplicado, el cálculo hubiera resultado más sencillo. Por ejemplo:

x_1	U_1	x_2	U_2
2.3	0.3	3.0	0.4

El IC sería:

$$IC = \frac{|2.3 - 3.0|}{\sqrt{(0.3)^2 + (0.4)^2}} = 1.4$$

Lo cual nos indica que las dos muestras del duplicado han sido analizadas de forma incorrecta al obtenerse resultados no compatibles.

El IC va a variar, como se ha visto, entre cero y uno, de modo que cuanto más próximo a cero, mejor habrá sido analizado el duplicado. Este índice también puede ser utilizado para detectar tendencias, tal y como exige la norma EN ISO/IEC 17025:2005, pero no en el formato indicado, ya que al tratar como valor absoluto la diferencia entre valores se podría ver la diferencia, pero no su naturaleza, es decir, si es positiva o negativa. Para poder hacerlo, bastará no utilizar el valor absoluto de la diferencia, con lo que el IC variará entre uno y menos uno (1,-1).

$$IC = \frac{(x_1 - x_2)}{\sqrt{U_1^2 + U_2^2}} \leq 1.0$$

Si se desea una escala mayor, no habrá más que multiplicar el valor obtenido de IC por un factor constante, por ejemplo 10, con lo que las diferencias se verán en una escala diez veces superior.



EL ATRIBUTO “OTROS” DE LA HOJA DE PERFIL

En la hoja de perfil del método de valoración organoléptica del Consejo Oleícola Internacional, COI/T.20/Doc. No 15/Rev. 6 Noviembre 2013, página 11, en la sección de los defectos, aparece uno denominado “Otros”. Este atributo tiene la particularidad de incluir otros defectos, que son los especificados en el apartado 4.2 del propio método de valoración organoléptica. Este atributo negativo “Otros” debe ser considerado y tratado como un único defecto, de la misma manera que lo son los otros defectos no incluidos en este apartado (atrojado, rancio, etc.), es decir, que si la mediana de dicho atributo es distinta de cero tiene que cumplir que su coeficiente de variación sea igual o inferior al 20 % para ser tomada en cuenta, de acuerdo a lo indicado en el tercer párrafo del apartado 10.3 del propio método, **independientemente de que los defectos indicados por los catadores no sean los mismos.**

Lo que verdaderamente importa del defecto “Otros” es que su mediana sea distinta de cero, que cumpla los requisitos exigidos y pueda clasificar adecuadamente, si es la más intensa, el aceite analizado.

Madrid, julio de 2014