

Métodos estadísticos

para los ensayos de aptitud por comparación entre laboratorios

Dr. Roberto A. Molina
Cruz

Introducción

La Unidad de Ensayos de Aptitud (UDEA) de la Asociación Gremial de Exportadores de Productos No Tradicionales (AGEXPRONT), realizó recientemente un ensayo de aptitud por comparación entre laboratorios (el UDEA-200304) de recuento aerobio en placa. El autor de este artículo asistió en el análisis estadístico, a las personas responsables de esta actividad. Siendo éste uno de los primeros ensayos de aptitud realizados en el país, el autor revisó los métodos estadísticos empleados en otros países, y propuso un método de análisis, de los datos para este ensayo, y futuros ensayos similares. Como resultado de este estudio, el autor elaboró varios documentos de trabajo, los cuales sirvieron de base para la elaboración del presente artículo. El mismo contiene: a) una descripción del ensayo UDEA-200304, b) una discusión sobre los aspectos generales de los ensayos de aptitud y los métodos estadísticos usualmente empleados, c) una descripción del método estadístico propuesto, d) los resultados de un análisis parcial de los datos, y e) los resultados de un análisis completo de los datos.

• El ensayo UDEA-200304

Se propuso un ensayo microbiológico de recuento aerobio en placa de *Escherichia coli*, siguiendo el procedimiento descrito en el *Bacteriological Analytical Manual (BAM) 8th edition*. A cada laboratorio se le entregó una muestra de un material de referencia elaborado por la empresa *MicroBiologics*, y se le requirió un solo valor de recuento, siguiendo cuidadosamente el procedimiento indicado.

Se contó con la participación de diez laboratorios, a cada uno de los cuales se le entregó dos muestras del material de referencia, e identificadas por medio de un código. Ambas muestras se analizaron y el laboratorio presentó dos resultados de recuento, indicando claramente el código de la muestra correspondiente. La información de los laboratorios, las muestras y los recuentos reportados, aparecen en la Tabla 1.

Las muestras proporcionadas a los laboratorios cuentan con un certificado de análisis que contiene la información descrita en la Tabla 2. Estos valores corresponden a recuentos realizados por la misma empresa *MicroBiologics*, pero

Relato del análisis efectuado sobre los resultados de un ensayo de aptitud

utilizando un procedimiento de recuento diferente al adoptado para este ensayo.

• Aspectos generales

Un objetivo del análisis estadístico es facilitar la comparación de los resultados de los laboratorios, por medio del cálculo de valores estandarizados conocidos como valores *z*. Para esto, es necesario determinar los valores de media y desviación estándar llamados de referencia. Estos valores de referencia son regularmente determinados por medio de (a) los valores presentados en el certificado del material de referencia, denominados los valores de certificado, y (b) los mismos resultados presentados por los laboratorios, cuyos valores de media y desviación estándar, son identificados como los valores de consenso. Además, un valor de desviación estándar de referencia puede ser determinado por medio de (c) criterios de desempeño establecido por la entidad examinadora, el cual es llamado, el valor de desviación estándar de desempeño.

Dada la naturaleza de los datos de los ensayos de aptitud, usualmente poco numerosos y posiblemente conteniendo valores reportados por laboratorios con alguna insuficiencia, el cálculo de los valores de media y desviación estándar de consenso, regularmente se realiza por medio de a) estadísticas robustas o bien b) un análisis de valores extremos –llamados en inglés *outliers*– y la aplicación de estadísticas no robustas. Desde luego, después del análisis de valores extremos también puede aplicarse estadísticas robustas, sin embargo, regularmente esto no ofrece mayores ventajas.

¹ AMC Technical Brief No. 6. Robust statistics: A method of coping with outliers Analytical Methods Committee, Royal Society of Chemistry, 2001.

Los métodos tradicionales para identificar valores extremos –como el de Grubbs y Cochran– requieren en algún grado del juicio del analista. Por el otro lado, el cálculo de estadísticas robustas no requiere de la identificación de posibles valores extremos. Existe un método más sofisticado para calcular los valores de consenso, el cual identifica posibles valores extremos sin necesidad de la opinión del analista, que emplea estadísticas no robustas y no excluye ningún dato en los cálculos. Éste es el método iterativo de Huber, el cual se basa en la transformación de los valores identificados como extremos –procedimiento conocido en inglés como *winsorisation*–. El uso de este método para el cálculo de estadísticas robustas de media y desviación estándar es recomendado.¹

Considerando todos los valores reportados por los laboratorios, se nota que su dispersión o variabilidad se debe a dos causas: a) el procedimiento de recuento adoptado no puede ser reproducido perfectamente –reproducibilidad– en todos los laboratorios, y b) el procedimiento de recuento no puede ser repetido perfectamente –repetibilidad– para diferentes muestras, en un mismo laboratorio. En un ensayo de aptitud interesa evaluar la capacidad de los laboratorios de reproducir un procedimiento, pero para esto es necesario saber cuánto es posible repetir el procedimiento en cada laboratorio. Por lo que, además de medir la dispersión total de los datos, es necesario determinar la fracción de esta dispersión, debida a cada una de sus causas.

Con solamente un valor de recuento por cada laboratorio, es factible medir la dispersión total de los datos, por medio de un valor de desviación estándar, pero no es posible determinar qué fracción de esta dispersión es debida a cada una de sus causas. Al usar este valor relativamente grande de dispersión como la desviación estándar de referencia, se obtienen valores z cercanos a cero; por lo que se observan menores diferencias entre los laboratorios, y se puede identificar solamente a laboratorios con valores reportados extremadamente grandes o pequeños.

Al contar con dos o más valores reportados por cada laboratorio, se calcula una media de consenso aplicando alguno de los métodos antes descritos, a todos los datos, y se puede medir la dispersión de los datos debida a cada una de sus causas, por medio de un análisis de varianzas (ANDEVA). En el caso que dos o más muestras similares hayan sido analizadas por cada laboratorio, se aplica un ANDEVA de un factor: el laboratorio. Este ANDEVA descompone la dispersión total de los datos en sus dos componentes: (*between*) la dispersión entre los laboratorios debida a la reproducibilidad del procedimiento de recuento, y (*within*) la dispersión propia de los laboratorios debida a la repetibilidad del procedimiento. Un procedimiento de ANDEVA robusto es también recomendado², el cual usa el método de Huber para el manejo de los valores extremos.

²AMC Technical Brief No. 6. *Robust statistics: A method of coping with outliers* Analytical Methods Committee, Royal Society of Chemistry, 2001.

Las estadísticas robustas

Las estadísticas robustas más usadas para el cálculo de una media y una desviación estándar de consenso, son la mediana y la mediana de las diferencias absolutas, conocida en inglés como *Median Absolute Difference (MAD)*. La mediana es desde luego, el percentil 50 del conjunto de datos, y la estadística *MAD* se calcula, como se describe a continuación.

Dado un conjunto de datos: x_1, x_2, \dots, x_n , se calcula su mediana: m , y las diferencias absolutas entre los datos y esta mediana:

$$|x_1 - m|, |x_2 - m|, \dots, |x_n - m|.$$

La estadística *MAD* es la mediana de estas diferencias absolutas. Es decir,

$$MAD = \text{Mediana } \{|x_1 - m|, |x_2 - m|, \dots, |x_n - m|\}$$

Se obtiene una desviación estándar ajustando esta estadística por un factor c , cuyo valor es determinado asumiendo que los datos obedecen a una distribución normal. Un valor usual es $c = 1.5$, por lo que la desviación estándar de consenso puede ser calculada por medio de la estadística *MAD_e*, así:

$$MAD_e = 1.5 * MAD$$

El método de Huber

Dado un conjunto de datos: x_1, x_2, \dots, x_n , el método de Huber es descrito por el siguiente algoritmo.

1. Se calcula la media: m_0 , y la desviación estándar: s_0 .

2. Se construye un primer intervalo de aceptación $I_0 = [m_0 - c * s_0, m_0 + c * s_0]$.
3. Se revisa todos los datos, de forma que
 - si $x_i > m_0 + c * s_0$
entonces se hace $x_i = m_0 + c * s_0$
 - si $x_i < m_0 - c * s_0$
entonces se hace $x_i = m_0 - c * s_0$
 - si $m_0 - c * s_0 < x_i < m_0 + c * s_0$
entonces x_i no es modificado
4. Con los datos revisados: x_1, x_2, \dots, x_n , se calcula su media: m_1 , y desviación estándar: s_1 , la cual se ajusta haciendo $s_1 = b * s_1$.
5. Se construye un nuevo intervalo de aceptación $I_1 = [m_1 - c * s_1, m_1 + c * s_1]$.
6. Se revisa de nuevo los datos, de forma que
 - si $x_i > m_1 + c * s_1$ entonces se hace $x_i = m_1 + c * s_1$
 - si $x_i < m_1 - c * s_1$ entonces se hace $x_i = m_1 - c * s_1$
 - si $m_1 - c * s_1 < x_i < m_1 + c * s_1$ entonces x_i no es modificado
7. Se repiten los pasos 4, 5 y 6 hasta que por ejemplo $|m_k - m_{k-1}| < \epsilon$.
8. Entonces el valor medio será: m_k , y la desviación estándar: s_k .

Las constantes b y c son determinadas asumiendo que los datos obedecen a una distribución normal, y sus valores usuales son: $b = 1.134$ y $c = 1.5$. Un valor usual de la constante que determina la condición de salida del algoritmo es $\epsilon = 0.0001$.

- El procedimiento de análisis propuesto

Dada la naturaleza de los mismos ensayos de aptitud, en los cuales los laboratorios

son evaluados y del resultado de esta evaluación depende su posible certificación, es muy importante que el procedimiento de análisis estadístico sea completamente objetivo; especialmente: a) debe ser determinado totalmente antes del ensayo, b) debe ser independiente de los posibles datos a reportar por los laboratorios y c) debe ser totalmente reproducible. Es decir, si el análisis es repetido con un mismo conjunto de datos, debe producir exactamente los mismos resultados. Para esto, es muy conveniente que pueda ser definido completamente en la sintaxis, de un programa de computadora.

Es especialmente importante que el procedimiento sea totalmente reproducible, ya que los laboratorios participan periódicamente en los ensayos de aptitud. Al aplicar exactamente el mismo procedimiento de análisis estadístico en cada ensayo, se evita una variabilidad no deseada de los resultados de un mismo laboratorio entre los ensayos. Con esto, los laboratorios pueden ser evaluados en forma consistente, y la información que se les provea en un ensayo, puede servirles para mejorar su procedimiento de trabajo, el cual puede ser de nuevo evaluado en el siguiente ensayo.

Dado que el procedimiento de recuento utilizado por los laboratorios difiere del empleado por la empresa *MicroBiologics*, no se creyó adecuado emplear los valores del certificado, como valores de referencia. Se planteó el uso de un valor de desviación estándar de desempeño pero su determinación pudo haber tomado mucho tiempo. Por lo que se decidió emplear los valores de consenso, como los valores de referencia. Se nota que el uso de una desviación estándar

de desempeño se justifica muy bien, ya que uno de los objetivos del ensayo de aptitud es comprobar que los laboratorios proveen a sus clientes, de recuentos suficientemente precisos, para que ellos puedan tomar decisiones adecuadas, en base estos valores reportados. Por esto, la incertidumbre debería ser determinada, tomando en cuenta los requerimientos prácticos de los clientes de los laboratorios.

Como inicialmente se esperaba que cada laboratorio analizara una sola muestra y reportara un solo valor de recuento, se sugirió el cálculo de los valores de consenso por medio de estadísticas robustas. El procedimiento de análisis fue definido en una sintaxis de Visual Basic que lee los datos y elabora la tabla de resultados en un libro electrónico de Excel. Este procedimiento fue discutido con las personas responsables del ensayo de aptitud, y finalmente, presentado a los representantes de los laboratorios participantes.

Al conocer que los laboratorios analizaron dos muestras y la mayoría de ellos reportaron dos valores de recuento, entonces, se propuso la aplicación de un ANDEVA para el cálculo de la desviación estándar de consenso. Dado que no se había discutido este procedimiento ni se contaba con una sintaxis, se propuso a las personas responsables del ensayo de aptitud, analizar solamente un valor reportado de cada laboratorio y evaluar a los laboratorio en base a este análisis parcial. Y, después de terminado el ensayo de aptitud, realizar un análisis completo de todos los datos reportados por los laboratorios.

• Los resultados parciales

Se consideró conveniente, para este análisis parcial, incluir solamente los valores reportados por los laboratorios que están contenidos en la primera columna de recuentos de la Tabla 1, excepto el laboratorio 66, para el cual se tomó el único valor reportado que aparece en la segunda columna. Es necesario notar que también los laboratorios con códigos 6642 y 98 reportaron un solo valor de recuento. En la Tabla 3, aparecen los datos analizados con sus valores de mediana y MADe, y sus valores estandarizados, con respecto a estos valores de consenso.

Estos valores de mediana (325.0) y MADe (142.5) son mucho menores que los valores de certificado: 1,500 y 320, respectivamente. La gran diferencia entre los valores de media, parece explicarse por la diferencia entre el procedimiento de recuento adoptado para el ensayo de aptitud y el empleado por la empresa *MicroBiologics*. El menor valor de desviación estándar de consenso de los laboratorios debe ser causado simplemente por los menores valores de recuento reportados por los laboratorios.

Se observa que todos los laboratorios –excepto los laboratorios 6642 y 8293– tienen un valor z entre -1 y 1 . El laboratorio 6642 tiene el valor z más grande (2.77) y el laboratorio 8293 tiene el segundo valor z más grande (1.93). Éstos, son los únicos laboratorios que podrían identificarse como aquellos que reportaron recuentos mayores al valor de referencia. Sin embargo, dado que los valores de certificado indican que las muestras contienen mayores cantidades de *E. coli*, los recuentos de estos dos

laboratorios puede deberse a que no aplicaron el procedimiento adoptado exactamente igual que los demás laboratorios.

- Los resultados completos

Dado que siete laboratorios reportaron dos valores de recuento –correspondientes a las dos muestras analizadas del mismo material de referencia– es conveniente aplicar un ANDEVA robusto a estos datos, para calcular un valor desviación estándar de consenso. En la Tabla 4, aparecen descritos estos laboratorios, junto con sus valores z , y en la Tabla 5, aparecen los resultados del ANDEVA obtenidos mediante el paquete estadístico *AMC Robust Statistics V1.0*.

Se puede evaluar la capacidad de cada laboratorio de repetir el procedimiento de recuento, comparando los valores reportados por el laboratorio con su media; para lo cual es conveniente calcular el valor z de estos recuentos considerando la desviación estándar (*within*) presentada en la Tabla 5. Se presentan, también, estos valores z en las dos últimas columnas de la Tabla 4.

Se observa que para todos, excepto un laboratorio, sus valores z están entre -1.0 y 1.0 . El laboratorio 30 tiene los mayores valores z : ± 1.5927 , lo cual indica una menor capacidad de este laboratorio de repetir el procedimiento de recuento.

Se puede evaluar la capacidad de los laboratorios de reproducir el procedimiento adoptado, comparando las medias de los laboratorios con la media de todos los datos –reportada como *Grand Mean* en la Tabla 5, para lo cual se calculó el valor z de la media de cada laboratorio, considerando la

desviación estándar de estas mismas medias, que se puede obtener, dividiendo por 1.4142 (raíz cuadrada de 2) la desviación estándar (*within*) de la Tabla 5; es decir, considerando una desviación estándar de 44.4 .

En la Tabla 4, se presenta también, la media y su valor z de cada laboratorio. La gran diferencia que se observa ahora entre estos valores z , indica una diferencia entre las capacidades de los laboratorios en reproducir el procedimiento de recuento adoptado. Se notó que los laboratorios quedan agrupados en: a) los laboratorios 32 y 9882 con una media cercana al valor de referencia, b) los laboratorios 82, 3220 y 25 con una media considerablemente menor al valor de referencia y c) los laboratorios 30 y 8293 con una media considerablemente mayor al valor de referencia.

Para los laboratorios 66, 6642 y 98 que reportaron un solo valor de recuento, claramente no se pudo evaluar su capacidad repetir el procedimiento adoptado. Sin embargo, asumiendo que tienen una repetibilidad perfecta, podría evaluarse la capacidad de reproducir el procedimiento, considerando el único valor reportado, como la media del laboratorio. De esta forma, es posible calcular los valores z que se presentan en la Tabla 6, aunque los verdaderos valores z podrían ser muy diferentes, dependiendo de la menor capacidad de estos laboratorios de repetir el procedimiento.

- Comentarios del autor a las observaciones del artículo

Al presentar para su publicación este artículo, el Editor de la Revista hizo varias observaciones sobre algunos aspectos

del ensayo de aptitud realizado (el UDEA-200304), las cuales se comentan a continuación.

Las observaciones fueron las siguientes:

a) En el ensayo realizado sólo se evaluó la capacidad de recuento y no se evaluó la capacidad de la detección de cada laboratorio, b) Es posible evaluar la precisión de recuento de cada laboratorio en forma diferente a la utilizada en el ensayo realizado y c) En ensayos de aptitud similares (ver Augustin & Carlier³ se ha analizado, el logaritmo (en particular base 10) de los recuentos y no los mismos recuentos de los laboratorios.

• La capacidad de detección

En los ensayos de aptitud regularmente (ver Augustin & Carlier³) se evalúa a los laboratorios en dos aspectos: su capacidad de detección y su capacidad de recuento (enumeración). El segundo aspecto, trata sobre la capacidad de un laboratorio de ofrecer recuentos fiables, lo cual a su vez, incluye los dos aspectos siguientes: a) La veracidad del recuento. Ésta se refiere a la proximidad del recuento reportado al valor de referencia. O, en el caso de que el laboratorio reporte varios valores de recuento, la proximidad de su medida de tendencia central al valor de referencia. Lo que claramente queda determinado por la mayor o menor capacidad del laboratorio de reproducir el procedimiento de recuento establecido en el ensayo de aptitud, y b) La precisión del recuento. Ésta trata sobre la precisión (menor o mayor dispersión) de los recuentos reportados por un mismo laboratorio. También, claramente

determinado por la mayor o menor capacidad del laboratorio en repetir el procedimiento de recuento establecido en el ensayo.

Por el otro lado, la capacidad de detección se refiere a la capacidad que tiene un laboratorio de identificar, tanto las unidades contaminadas, como las no contaminadas; para lo que regularmente se emplean los términos siguientes: a) Sensibilidad: La capacidad del laboratorio en detectar las unidades contaminadas, y b) Especificidad: La capacidad del laboratorio en identificar las unidades no contaminadas.

Al entregar a un laboratorio, T unidades (muestras) para su análisis, de las cuales la unidad evaluadora conoce cuales están contaminadas y cuales no, y recibir del laboratorio el reporte de cada unidad, se pueden determinar los valores de la Tabla 7 y que se describen a continuación.

A = El número de unidades no contaminadas identificadas.

B = El número de unidades contaminadas no detectadas.

C = El número de unidades no contaminadas no identificadas.

D = El número de unidades contaminadas detectadas.

Es posible medir la sensibilidad de un laboratorio por medio del cociente siguiente:

$$\text{Sensibilidad} = \frac{D}{B + D}$$

³Augustin, J.C. & Carlier, V. *Lessons from the organization of a proficiency testing program in food microbiology by interlaboratory comparison: Analytical methods in use, impact of methods on bacterial counts and measurement uncertainty of bacterial counts* *Food Microbiology* 23: 1-38, 2006.)

Mientras que su especificidad por medio del cociente:

$$\text{Especificidad} = \frac{A}{A + C}$$

En los ensayos de aptitud es común calcular ambos cocientes para cada laboratorio, y luego presentar la distribución de los laboratorios con respecto a estos cocientes. Desde luego, es usual identificar a los laboratorios con los menores valores de sensibilidad y especificidad en el ensayo. Sin embargo, en el ensayo realizado (el UDEA-200304) se estableció desde el inicio, que solamente muestras contaminadas serían entregadas a los laboratorios y con el mismo nivel de contaminación, por lo que no se evaluó la capacidad de detección de los laboratorios. De haberse querido evaluar, hubiera sido necesario entregar a cada laboratorio varias muestras, por ejemplo $T = 5$, algunas contaminadas y otras sin contaminar. Y, se hubiera requerido a cada laboratorio que reportara por cada muestra: a) si la muestra estaba o no contaminada, y b) el recuento de la muestra, si aplicaba.

- La precisión de recuento

Como se indicó antes, la precisión de recuento de un laboratorio se refiere a la menor o mayor dispersión de sus valores de recuento reportados. Lo cual se debe a su mayor o menor capacidad de repetir el procedimiento de recuento utilizado. Dados los valores de recuento de todos los laboratorios, la dispersión total de estos recuentos tiene las dos partes siguientes: a) La debida a la capacidad del laboratorio de reproducir

el procedimiento de recuento establecido para el ensayo (*between*), y b) La debida a la capacidad de repetir el procedimiento utilizado por cada laboratorio (*within*).

En el ensayo de aptitud realizado (el UDEA-200304), se planteó evaluar la precisión de los recuentos de los laboratorios por medio del análisis de varianzas (ANDEVA) sugerido en el párrafo anterior, y cuyos resultados discutimos en la siguiente sección. Sin embargo, en ensayos de aptitud similares, se ha empleado otro método para evaluar la precisión de los laboratorios, como aparece en Augustin & Carlier³. Ellos siguieron el uso de la estadística siguiente:

$$r = \frac{(k - 1)s^2}{s^{*2}}$$

Donde:

- k = El número de muestras contaminadas analizadas.
- s = La desviación estándar de los valores de recuento del laboratorio.
- s^* = La desviación estándar de referencia.

Se supone que los valores de recuento, especialmente el logaritmo de los recuentos, están normalmente distribuidos, y que se puede asociar a la estadística r una distribución de probabilidad χ^2 con $k-1$ grados de libertad, lo cual permite definir intervalos de tolerancia para los valores de la estadística r de los laboratorios, y con ellos identificar los laboratorios con valores r extremos.

En la Tabla 8 se presentan los valores de recuento (Log (CFU)) del ensayo realizado, junto con los resultados de los valores s

y r recién descritos, los cuales fueron calculados solamente para los laboratorios que reportaron dos valores de recuento, es decir con $k = 2$. Se observa que los laboratorios 25 y 30 presentan los dos mayores valores r : 1.5762 y 1.4217, y por lo tanto, la menor precisión de recuento. Sin embargo, estos valores tienen asociados los valores p ($\Pr[R > r]$): 0.2093 y 0.2331, indicando que todavía la probabilidad de observar valores r mayores a estos es por lo menos 20%. Esto es suponiendo que los valores r obedecen una distribución χ^2 .

- El análisis del logaritmo de los recuentos

Para el análisis parcial del logaritmo de los recuentos, esto es considerando solamente uno de los dos valores reportados por algunos laboratorios, se utilizó el logaritmo base 10 de los mismos recuentos analizados previamente, los cuales se presentan en la Tabla 9. En esa misma tabla, aparecen los valores de mediana y MADe, y los valores estandarizados (z), ahora del logaritmo de los recuentos reportados ($\log(\text{CFU})$).

Se observa que dos laboratorios tienen valores z menores a -1 (laboratorios 3220 y 25), y que dos laboratorios tienen valores z mayores a $+1$ (6642 y 8293). De nuevo, el laboratorio 6642 tiene mayor el valor z (1.8218) y el laboratorio 8293 tiene el segundo mayor valor z . El laboratorio 25 tiene el menor valor z observado (-1.1870) y el laboratorio 3220 tiene el segundo menor valor z (-1.1280).

Debe notarse que la transformación logaritmo provoca una reducción (no proporcional) entre de los valores (logaritmos) de recuento reportados por los laboratorios, afectando especialmente

a los valores de recuento extremos, es decir los más alejados de su mediana. Con esto se tiene una disminución, en valor absoluto, de los valores z correspondientes a los valores de recuentos extremos, provocando que menos laboratorios puedan ser identificados como reportando recuentos muy alejados al valor de referencia.

Sin embargo, este no es el objetivo de la transformación. Sino que, la transformación logaritmo pretende adecuar los datos para la aplicación de un modelo de análisis de varianzas (ANDEVA), específicamente para que los residuos del modelo resulten estar normalmente distribuidos. Esto es porque la escala de los recuentos comprende valores en los cientos con un rango considerable. Por ejemplo, en el ensayo realizado, se observan valores de recuento entre 185 a 720.

Para el análisis completo del logaritmo de todos los recuentos reportados, se consideraron los valores de las dos muestras analizadas por los laboratorios, excepto los laboratorios 66, 6642 y 98 que reportaron los resultados de solo una muestra. En la Tabla 10, aparecen los valores de recuento (CFU) y sus logaritmos ($\log(\text{CFU})$).

Se evaluó la capacidad de reproducir el procedimiento de recuento de cada laboratorio, calculando su valor medio y z de recuento que también se presentan en la Tabla 10. Estos valores z son calculados considerando la media de consenso presentada en la Tabla 11, y su error estándar calculado a partir de la desviación estándar dentro de (*within*) los laboratorios, esta última también presentada en la Tabla 11.

En la Tabla 10, se observa de nuevo que los laboratorios con dos valores de recuento reportados, quedan agrupados en: a) los laboratorios 32 y 9882 con una media cercana al valor de referencia, b) los laboratorios 82, 3220 y 25 con una media considerablemente menor al valor de referencia y c) los laboratorios 30 y 8293 con una media considerablemente mayor al valor de referencia. Para los laboratorios 66, 6642 y 98 que reportaron un solo valor de recuento, suponiendo que tienen una repetibilidad perfecta y considerando el único valor reportado como la media del laboratorio, se puede calcular los valores z que también se presentan en la Tabla 10.

Se evaluó la capacidad de cada laboratorio de repetir el procedimiento de recuento, por medio de los valores z del logaritmo de los recuentos calculados empleando la desviación estándar dentro de (*within*) los laboratorios, esta última presentada en la Tabla 11. Estos valores z aparecen en las dos últimas columnas de la Tabla 10. Se observa que para todos, excepto ahora dos laboratorios, sus valores z están entre -1.0 y 1.0 . En este nuevo análisis, es el laboratorio 25 quien tiene los mayores valores z : ± 1.1146 , y el laboratorio 30 los segundos mayores: ± 1.0053 .

TABLA 1

Todos los recuentos reportados por los laboratorios

Código del laboratorio	Código de las muestras	Recuentos (CFU)
66	9, 1	--, 230
6642	2, 10	720, --
30	18, 8	400, 600
98	16, 21	250, --
8293	3, 4	600, 700
82	24, 25	250, 280
3220	15, 5	190, 240
32	12, 6	410, 430
9882	11, 17	420, 400
25	14, 13	185, 290

TABLA 2

Los valores de certificado del material de referencia

Valor medio de microorganismos:	1.5×10^3 CFU por <i>pellet</i>
Desviación estándar:	3.2×10^2
Coefficiente de variación:	21%
Intervalo de confianza (99%):	1.4×10^3 -- 1.7×10^3 CFU
Intervalo de confianza (95%):	1.4×10^3 -- 1.6×10^3 CFU

TABLA 3

Los recuentos parciales y los valores z

Código del laboratorio	Código de la muestra	Recuento (CFU)	Valor z
166	1	230	-0.6667
26642	2	720	+2.7719
330	18	400	+0.5263
498	16	250	-0.5263
58293	3	600	+1.9298
682	24	250	-0.5263
73220	15	190	-0.9474
832	12	410	+0.5965
99882	11	420	+0.6667
1025	14	185	-0.9825

Mediana: 325.0
MADe: 142.5

TABLA 4

Los recuentos y valores z de los laboratorios con dos valores reportados

Código del laboratorio	Código de las muestras	Recuentos		
		(CFU)	Media y Valor z	Valores z
30	18, 80	400, 600	500.0, +2.6977	-1.5927, +1.5927
8293	3, 40	600, 700	650.0, +6.0764	-0.7964, +0.7964
82	24, 25	250, 280	265.0, -2.5956	-0.2389, +0.2389
3220	15, 50	190, 240	215.0, -3.7218	-0.3982, +0.3982
32	12, 60	410, 430	420.0, +0.8958	-0.1593, +0.1593
9882	11, 17	420, 400	410.0, +0.6705	+0.1593, -0.1593
25	14, 13	185, 290	237.5, -3.2150	-0.8362, +0.8362

TABLA 5

Resultados del análisis de varianzas

ROBUST ESTIMATES	
<i>Parameter</i>	<i>Value</i>
Grand Mean	380.2
Within-group/repeatability SD	62.8
Between-group SD	162.5
Reproducibility SD	174.2

c = 1.5: Convrit = 0.0001

ROBUST REPRODUCIBILITY/ANOVA

AMC Robust Statistics V1.0

TABLA 6

Los recuentos y valores z de los laboratorios con un solo valor reportado

Código del laboratorio	Código de la muestra	Recuentos	
		(CFU)	Valor z
66	1	230	-3.3839
6642	2	720	+7.6532
98	16	250	-2.9334

TABLA 7

Tabla 2x2 para la evaluación de la capacidad de detección

Reporte del laboratorio	Referencia (realidad)		
	0	1	Total
0	A	B	A+B
1	C	D	C+D
Total	A+C	B+D	T = A+B+C+D

0 = No contaminada

1 = Contaminada

TABLA 8Los valores r de precisión

Código del laboratorio	Log (CFU)	Recuentos		Pr[R > r]
		s	r	
30	2.602, 2.778	0.1245	1.4217	0.2331
8293	2.778, 2.845	0.0473	0.5405	0.4622
82	2.398, 2.447	0.0348	0.3974	0.5284
3220	2.279, 2.380	0.0717	0.8192	0.3654
32	2.613, 2.633	0.0146	0.1670	0.6828
9882	2.623, 2.602	0.0150	0.1711	0.6792
25	2.267, 2.462	0.1380	1.5762	0.2093

TABLA 9Los recuentos parciales, sus logaritmos y valores z

Código del laboratorio	Código de la muestra	Recuentos		
		CFU	Log (CFU)	Valor z
66	1	230	2.362	-0.7050
6642	2	720	2.857	+1.8218
30	18	400	2.602	+0.5203
98	16	250	2.398	-0.5203
8293	3	600	2.778	+1.4181
82	24	250	2.398	-0.5203
3220	15	190	2.279	-1.1280
32	12	410	2.613	+0.5750
9882	11	420	2.623	+0.6284
25	14	185	2.267	-1.1870

Mediana: 2.500

MADE: 0.131



Dr. Roberto A. Molina Cruz
 rmolina@uvg.edu.gt
 Catedrático del Departamento de
 Matemática de la Facultad de
 Ciencias y Humanidades de la
 Universidad del Valle de Guatemala

Relato del análisis efectuado sobre los resultados de un ensayo de aptitud

TABLA 10

Los recuentos reportados, sus logaritmos y valores z

Código del laboratorio	Código de las muestras	Recuentos			
		CFU	Log (CFU)	Media y valor z	Valores z
66	9, 1	--, 230	--, 2.362	2.362, -3.050	--, 0.000
6642	2, 10	720, --	2.857, --	2.857, 4.953	0.000, --
30	18, 8	400, 600	2.602, 2.778	2.690, 2.252	-1.005, 1.005
98	16, 21	250, --	2.398, --	2.398, -2.465	0.000, --
8293	3, 4	600, 700	2.778, 2.845	2.812, 4.215	-0.382, 0.382
82	24, 25	250, 280	2.398, 2.447	2.423, -2.068	-0.281, 0.281
3220	15, 5	190, 240	2.279, 2.380	2.329, -3.571	-0.579, 0.579
32	12, 6	410, 430	2.613, 2.633	2.623, 1.171	-0.118, 0.118
9882	11, 17	420, 400	2.623, 2.602	2.613, 1.002	0.121, -0.121
25	14, 13	185, 290	2.267, 2.462	2.365, -3.001	-1.115, 1.115

TABLA 11

Resultados del análisis de varianzas

ROBUST ESTIMATES	
<i>Parameter</i>	<i>Value</i>
Grand Mean	2.551
Within-group/repeatability SD	0.088
Between-group SD	0.196
Reproducibility SD	0.214

c = 1.5: Convcrit = 0.0001ROBUST
 REPRODUCIBILITY/ANOVA
 AMC Robust Statistics V1.0