

Notas técnicas: análisis

de datos bajo condiciones de repetibilidad

Technical notes: Analysis of data under repeatability conditions

Aisha Hurtado, Aura Caldera, Balentina Milano, Carlos Ibarra, Alba Díaz, José Camacho, José Elías Villamizar, Omar Verde
Laboratorio de Bioanalítica, Departamento de Química Medicinal, Instituto Venezolano de Investigaciones Científicas (IVIC). Ministerio del Poder Popular para Educación Universitaria, Ciencia y Tecnología. Email: omarverde@gmail.com

Resumen

Se presentan estadísticos para caracterizar un grupo de datos provenientes de una variable continua bajo condiciones de repetibilidad, el cálculo de intervalos de confianza, la detección de valores atípicos y la posibilidad de realizar comparaciones de las varianzas de dos o más grupos para determinar la posible existencia de discrepancias estadísticamente significativas entre ellas y la relación con sus repetibilidades. Se hace mención del término precisión de los resultados, estrechamente asociado al de repetibilidad. Se presentan salidas de cálculos realizados con un programa estadístico para la caracterización de una variable continua.

Palabras clave: Repetibilidad, precisión, intervalos de confianza, comparación de varianzas

Abstract

It is described statistical procedures to characterize a group of data from a continuous variable under repeatability conditions, the calculation of confidence intervals, the detection of outliers and the possibility of comparisons of the variances of two or more groups to determine the possible existence of statistically significant discrepancies between them and the relationship with their repeatability. Mention is made of the accuracy of the results, closely associated with the repeatability. Outputs of calculations performed with a statistical programme for the characterization of a continuous variable are presented.

Key words: Repeatability, accuracy, confidence intervals, variance comparison

Introducción

40

En el campo de la analítica usualmente se obtienen valores que son utilizados para estimar la condición favorable o desfavorable de una unidad (equipo, persona, parcela de terreno). Dependiendo del tipo de variable, se debe proceder a manipular esos valores para realizar su caracterización estadística y poder inferir con respecto a valores de referencia o realizar, en el caso de dos o más unidades, comparaciones que permitan tomar decisiones para corregir deficiencias o continuar en el proceso de una mejora continua.

En esta Nota Técnica se procederá a presentar aspectos importantes de una variable continua, definir las condiciones de repetibilidad y señalar estadísticos que permitan su caracterización o la comparación entre unidades.

Variable continua y su caracterización

Una variable continua es aquella que puede asumir el infinitamente largo número de valores correspondientes a los puntos en un determinado intervalo y siempre será posible, en teoría, la existencia de un tercer valor intermedio entre dos valores observables. Nunca puede ser medida con exactitud. El valor observado depende, en gran medida, de la precisión de los instrumentos de medición. Ejemplos típicos

lo son la altura de las personas, el tiempo de vida de una célula, la cantidad de azúcar en una naranja y la concentración en la sangre de un parámetro farmacocinético.

Un aspecto importante a considerar en las mediciones obtenidas para una variable continua es la necesidad de organizarlas y caracterizarlas de forma tal que se pueda visualizar con facilidad la tendencia de los datos hacia un punto central o medir la variación observada. Entre los estimadores de tendencia central para un grupo de mediciones se tiene la media aritmética o promedio, la mediana y la moda mientras que para la variación se tienen estimadores como la varianza, la desviación estándar y el coeficiente de variación.

La **media aritmética o promedio** se obtiene sumando todas las observaciones o mediciones y dividiéndolo entre el número de mediciones consideradas. La **mediana** es el valor central de las observaciones ordenadas en forma creciente o decreciente (o el promedio de los dos valores centrales si se tiene número par de observaciones) y la **moda** es el valor o los valores que más se repiten. Por la parte de medición de la variación se tiene la **varianza** que se calcula a partir de la suma de las desviaciones cuadráticas de las obser-

vaciones respecto al promedio dividido entre el número de observaciones menos 1, la **desviación estándar** que es la raíz cuadrada positiva de la varianza y el **coeficiente de variación** que expresa la desviación estándar como porcentaje del promedio.

Un ejemplo numérico:

Se realizan 10 mediciones (datos ficticios) que proporcionaron los siguientes valores ordenados en forma creciente:

10.08, 10.09, 10.10, 10.10, 10.12, 10.13, 10.14, 10.14, 10.16, 10.18

$$\text{Promedio} = y = \sum y_i / n = 101.24 / 10 = 10.124$$

$$\text{Mediana} = (10.12 + 10.13) / 2 = 10.125$$

$$\text{Moda} = 10.10 \text{ y } 10.14$$

$$\begin{aligned} \text{Varianza} &= s^2 = \sum (y_i - y)^2 / (n - 1) \\ &= \{(10.08 - 10.124)^2 + (10.09 - 10.124)^2 + \dots + (10.16 - 10.124)^2 \\ &\quad + (10.18 - 10.124)^2\} / 9 \\ &= 0.00924 / 9 = 0.00102667 \end{aligned}$$

$$\text{Desviación estándar} = s = \sqrt{0.00102667} = 0.03204$$

$$\text{Coeficiente de variación} = \text{C.V.} = 0.03204 \cdot 100 / 10.124 = 0.3165$$

Estas operaciones aritméticas pueden ser realizadas en una hoja de cálculo o mediante la utilización de programas estadísticos disponibles.

Estadística descriptiva

	Y
N	10
LI 95% IC	10.101
Media	10.124
LS 95% IC	10.147
Desv. Est.	0.0320
Varianza	0.0010
SE Mean	0.0101
C.V.	0.3165
Mínimo	10.080
Mediana	10.125
Máximo	10.180

Precisión y Repetibilidad

Cuando se realizan análisis físicos o químicos, conviene obtener mas de una medición y, así, obtener replicas que permitirá observar el grado de concordancia que existe entre ellas. Este grado de concordancia se denomina **precisión**. La **repetibilidad** es definida como la desviación estándar obtenida al analizar una misma muestra varias veces, en un

periodo de tiempo corto, sin cambiar de equipo de medida, reactivos o analista. En otras palabras, la repetibilidad permite cuantificar la variación de uno o más conjuntos de valores obtenidos bajo condiciones de repetibilidad.

Los términos **precisión de los resultados** y **repetibilidad** son muy utilizados en la evaluación de la confiabilidad de los resultados en programas de comparación interlaboratorio y reflejan el grado de seguridad de los resultados emitidos.

A manera de ejemplo, si los 10 valores previamente utilizados corresponden a 10 mediciones realizadas bajo condiciones de repetibilidad, se tiene que existe una variación de 0.00102667 ó de 0.03204 ó de 0.3165%, dependiendo del estimador que se utilice (varianza, desviación estándar o coeficiente de variación).

El valor del coeficiente de variación es frecuentemente utilizado como un indicador de la repetibilidad y precisión, aunque esto depende del método utilizado para obtener las mediciones. Valores muy bajos (inferiores a 1%) reflejan una buena precisión. Valores superiores a 2 % pudieran ser indicadores de baja precisión.

Las mediciones bajo condiciones de repetibilidad se espera que presenten variaciones pequeñas ya que provienen de condiciones homogéneas. La variación observada en ellas se denomina "error puro" pues no se tienen fuentes de variación conocidas. Se puede representar la varianza poblacional para repetibilidad por s_{rep}^2 y la desviación estándar poblacional para repetibilidad por s_{rep} .

Con el promedio y la desviación estándar obtenidos bajo condiciones de repetibilidad, es posible obtener el intervalo de confianza para las mediciones. La amplitud de este intervalo depende de la precisión de los valores individuales, lo que depende de la desviación estándar y del número de observaciones en la muestra.

El intervalo de confianza de la media está dado por la expresión:

$$y \pm t.s / \sqrt{n}$$

donde los valores de t están relacionados con los grados de libertad (n - 1) y el nivel de seguridad con que se desea trabajar.

En el presente ejemplo, el intervalo de confianza al 95% será:

$$10.124 \pm 2.262 \cdot 0.03204 / \sqrt{10} \quad \text{o}$$

$$10.124 \pm 0.023 \quad \text{o}$$

$$10.101 \text{ a } 10.147$$

valores iguales a los previamente presentados en la salida del computador, mientras que el intervalo al nivel 99 % estará ubicado entre:

$$10.124 \pm 3.250 \cdot 0.03204 / \sqrt{10} \quad \text{o}$$

$$10.124 \pm 0.033 \quad \text{o}$$

$$10.091 \text{ a } 10.157$$

lidad en los programas de confiabilidad de resultados. Es una prueba que se realiza en forma iterativa, calculando el cociente de la varianza mayor con relación a la suma de las varianzas y se compara con el valor de la tabla asociado con un nivel de confianza, el número de unidades en evaluación y el número de observaciones o mediciones por cada unidad. Si el valor calculado es superior al correspondiente valor de la tabla, se concluye que esa unidad tiene una varianza muy alta. Se procede a separar esa unidad y se va a la siguiente unidad con varianza más alta y se realiza de nuevo la evaluación. Se concluye la prueba cuando la varianza evaluada produzca un cociente inferior al tabulado.

Los valores críticos de la tabla de Cochran para 5 valores por unidad, 8 unidades y nivel 95% de confianza son:

Unidades	Valor crítico
2	0.906
3	0.745
4	0.629
5	0.544
6	0.480
7	0.431
8	0.391

Si los resultados obtenidos en las evaluaciones fueron:

Unidades	s^2
1	0.024
2	0.101
3	0.025
4	0.020
5	0.184
6	0.022
7	0.031
8	0.010
Suma	0.417

Se tendrá que:

Varianza mayor = 0.184 y el cociente es $0.184 / 0.417 = 0.441$, valor superior al asociado a 8 unidades (0.391), por lo que se descarta esta unidad, quedando 7 en evaluación.

La nueva suma de varianzas es 0.233 y el nuevo cociente es $0.101 / 0.233 = 0.433$, superior al asociado para 7 unidades

(0.431), procediéndose a descartar esta unidad.

Pero con las 6 unidades que permanecen, los valores son de 0.132 y 0.235, inferior este último al asociado para 6 unidades (0.480), indicando que no se puede descartar esta unidad y se da por finalizada la prueba.

Es decir, existen dos unidades, las identificadas como 2 y 5, con varianzas significativamente superiores al resto, indicando poca repetibilidad y poca confiabilidad en sus datos.

Referencias

COVENIN.2972-2. 1997. Exactitud (Veracidad y Precisión) de Métodos de Medición y Resultados. Parte 2: Método Básico para la Determinación de Repetibilidad y Reproducibilidad de un Método Estándar de Medición.

Mendenhall, W. 1975. Introduction to Probability and Statistics. Cuarta edición. Duxbury Press. 460 pp

Miller, N. J. y Miller J. C. 2002. Estadística y Quimiometría para Química Analítica. Cuarta edición. Prentice Hall. 296 pp

Portuondo, Y. y J. Portuondo. 2010. La Repetibilidad Y Reproducibilidad en el Aseguramiento de la Calidad de los Procesos de Medición. Tecnología Química. Vol. XXX, No. 2, 117 – 121.

Sokal, R. R. y F. J. Rohlf. 1981. Biometría: los Principios y la Práctica de Estadísticas en la Investigación Biológica. Segunda edición. WH Freeman and Co. 400 pp.