Errores de medida en variables numéricas: Correlación y Concordancia

Preparado por Luis M. Molinero (Alce Ingeniería)
CorreoE: **bioestadistica@alceingenieria.net**

Artículo en formato PDF

Agosto 2001

Fiabilidad de un proceso de medida

La medida de parámetros fisiológicos está sujeta a error y a la propia variabilidad biológica. La presión arterial es un claro ejemplo: aunque la técnica es bastante simple, pueden aparecer errores debidos a defectos del aparato utilizado, a la aplicación del manguito, al estado del paciente y a la objetividad y preparación del observador. Es de desear que el proceso sea fiable: la repetición de las medidas de la misma magnitud producen resultados iguales o al menos similares. Hablamos entonces de fiabilidad de las mediciones, estabilidad o concordancia. Diremos que una medición es fiable si la variabilidad en mediciones sucesivas se mantiene dentro de cierto margen razonable.

En ocasiones pueden existir diferentes métodos de medida, siendo uno de ellos el que mejor determina la magnitud de la variable en estudio. A éste se le conoce como patrón de referencia (en inglés gold standard) y en principio sería el método a emplear preferentemente, salvo que presente serios inconvenientes, como pueden ser el coste, que se trate de un método cruento, complicado de utilizar, etc. Es el caso de la medición de la tensión arterial mediante la introducción de un catéter flexible en una arteria periférica.

Diferencia entre correlación lineal y concordancia

Si se dispone de un método alternativo al método de referencia, más práctico de utilizar, interesa determinar la concordancia entre ambos sistemas.

Cuando la variable medida es numérica continua y se efectuan dos observaciones por sujeto existe una cierta tendencia a emplear el coeficiente de correlación como índice de concordancia entre los dos métodos, no siendo éste, en general, un procedimiento correcto, como veremos en los siguientes ejemplos.

PAS método 1	PAS método 2
130	126
124	120
144	140
112	108
124	120
161	157
138	134
164	160
109	105
152	148

Tabla 1

En la tabla 1 se presenta los resultados hipotéticos de dos métodos de medida. Si calculamos el coeficiente de correlación veremos que es 1, ya que en realidad los valores de la segunda columna se obtienen restando 4 a la primera columna. Por lo tanto hay una relación lineal perfecta entre ambos métodos, por lo que vemos claramente que no es lo mismo correlación que concordancia.

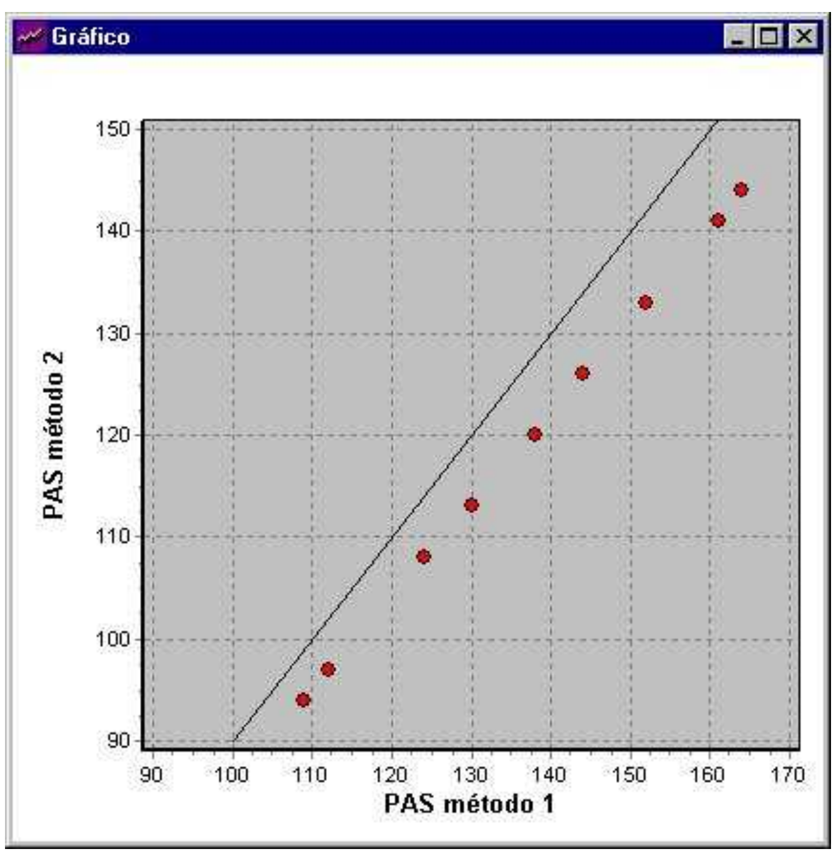
Veamos otro ejemplo:

PAS método 1	PAS método 2
130	113
124	108
144	126
112	97
124	108
161	141
138	120
164	144
109	94
152	133

Tabla 2

En este caso también hay una relación lineal perfecta entre ambos métodos y por lo tanto el coeficiente de correlación también es 1. En realidad *PAS 2 = 0.2 * PAS 1 − 4* (se ha redondeado el resultado), por lo que existe también una relación lineal entre los valores de las dos columnas.

Si representamos gráficamente los datos obtenemos la siguiente imagen:



donde se ha representado una línea recta a 45º (que sería la línea en la que estarían todos los puntos en caso de que existiera una concordancia perfecta).

En situaciones como la de la gráfica encontramos un error sistemático o diferencia entre ambos métodos, pero que sería de fácil corrección, una vez determinada la relación que existe entre ambos. Es algo que ocurre en la práctica de las mediciones, con situaciones más complejas en las que la relación entre ambos métodos puede no ser lineal. Así en el caso concreto de la medida de la presión arterial mediante esfigmomanómetro ésta depende en cierta medida de la relación entre el ancho de la bolsa inflable utilizada y la circunferencia del brazo del sujeto, por lo que se puede calibrar o corregir en función de esos valores.

Si existe una relación matemática entre dos métodos, sea ésta lineal o no, siempre podemos conocer a partir de ella el valor en uno de los métodos dado el valor del otro sistema.

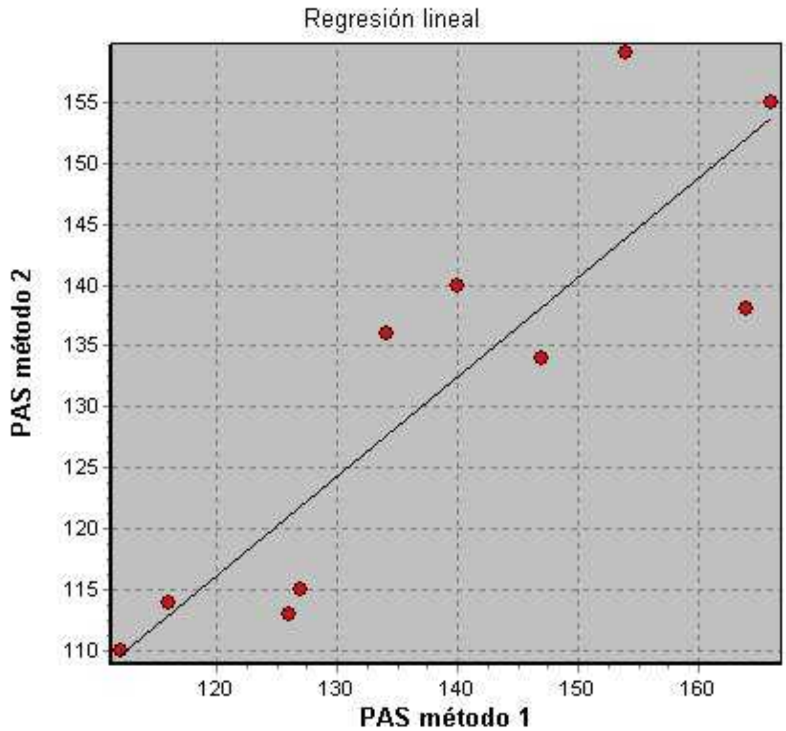
En la práctica de la medición de variables fisiológicas, los resultados observados raramente mostrarán relaciones tan exactas como las que se han presentado en los ejemplos anteriores. En la siguiente tabla refleja una situación en la que la relación no es tan clara

PAS método 1	PAS método 2
130	136
124	115
144	134
112	114
124	113
161	138
138	140
164	155
109	110
152	159

Tabla 3

El valor del coeficiente de correlación para estos datos es de 0.872.

Si ajustamos una recta de regresión y la representamos gráficamente obtenemos



que corresponde a la ecuación de regresión *PAS 2 = 22.7 + 0.8 x PAS 1*. Nuevamente nos encontramos ante una situación en la que, a pesar de tener un aceptable coeficiente de correlación 0.8732, sin embargo no podemos decir que haya tan buena concordancia entre ambos métodos.

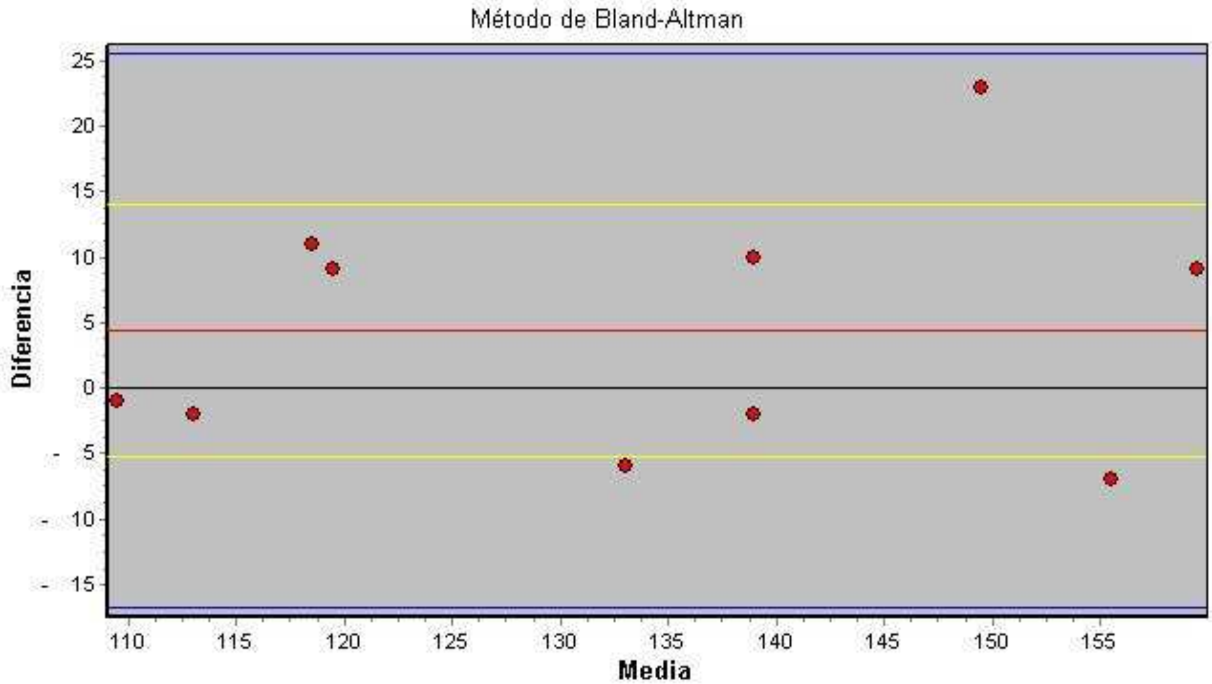
El problema es que **el concepto de coeficiente de correlación lineal no es igual que concordancia**.

Además **el coeficiente de correlación depende del rango de valores observado en la muestra**, ya que si se incluyen valores extremos el coeficiente de correlación aumenta. Así en el ejemplo anterior solo con añadir la pareja de valores (200 , 179) el coeficiente de correlación pasa a ser de 0.927.

Otro concepto erróneo es el de que si el coeficiente de correlación entre dos medidas es significativamente diferente de cero la fiabilidad es buena. **El coeficiente de correlación lineal puede ser muy pequeño y resultar significativamente diferente de cero si el tamaño de la muestra es suficientemente grande**. Así por ejemplo un coeficiente de correlación tan pequeño como 0.21 será significativamente diferente de cero si se ha calculado en una muestra con al menos 88 parejas de observaciones ($p < 0.05$).

Método de Bland−Altman para evaluar la concordancia entre dos sistemas de medida

Bland JM y Altman DG proponen un gráfico sencillo para evaluar la concordancia entre dos métodos de medida. Consiste en representar la diferencia entre cada pareja de valores frente a la media de cada pareja de valores. Para los datos de la tabla 3 se obtiene la siguiente gráfica:



En el caso de que no haya error sistemático los puntos se distribuirán de forma aleatoria a uno y otro lado de la recta correspondiente a la diferencia 0 entre medidas (línea horizontal negra). La línea roja representa la media de las diferencias, que en nuestro ejemplo corresponde a 4.4 (error sistemático del segundo método respecto al primero). Las líneas azules representan los límites de confianza del 95 % para esa diferencia, y se denominan *límites de concordancia*. A su vez las líneas amarillas representan el límite confianza inferior para cada límite de concordancia.

Estimación del error de medida: Varianza media intra−sujeto

Hasta aquí se ha visto cómo determinar la fiabilidad de un método de medida en relación a un patrón, cuando se dispone de una observación, mediante cada uno de los métodos, para cada sujeto. En la práctica nos encontramos muchas situaciones en las que no existe ese patrón de medida o no está disponible, pero sí que podemos efectuar varias mediciones por sujeto. Entonces no podemos evaluar la validez del método; entendiendo por **validez** que el procedimiento está realmente midiendo la magnitud fisiológica que se desea cuantificar. Sin embargo sí que podemos obtener información en cuanto a la reproducibilidad del método, al disponer de varias medidas para cada sujeto. La variabilidad de las medidas para el mismo sujeto, calculada como una desviación típica, nos permite cuantificar el error del método de medida. Ello implica suponer que esa variabilidad es igual para todos los sujetos, lo que no siempre ocurre y habrá que comprobar en cada caso.

Para determinar el error de medida necesitamos calcular lo que se denomina **desviación típica intra−sujeto** (en inglés **within−subject standard deviation**); para ello se ha de disponer de al menos dos medidas por sujeto, y se calcula la media de las desviaciaciones típicas de todas las medidas realizadas para cada sujeto. En los programas de estadística ese dato se puede obtener a partir de una tabla de análisis de la varianza.

Veamos un ejemplo en el que para 10 pacientes se efectuan tres mediciones de la tensión arterial sistólica:

PAS1	PAS2	PAS3
140	132	130
150	124	130
150	144	140
113	112	111
128	124	128
182	161	152
143	138	133
180	170	164
134	128	109
161	152	142

Tabla 4

Si mediante un programa solicitamos un análisis de la varianza de los datos de la tabla 4 obtendremos una salida similar a la de la figura:

Factor	Sum.cuadrados	gl	Cuadrados medios	F	p	Nivel sig.
Entre casos	9184.2	9	1020.5	12.25	0.00000236	$p < 0.001$
Intra casos	1666.0	20	83.30			
Total	10850.2	29				

La desviación típica intra−sujetos corresponde a la raíz cuadrada de los cuadrados medios intra−sujetos (en inglés **within−subject mean square**) $\sqrt{83.3} = 9.1$.

Coeficiente de correlación intraclase

Cuando utilizamos la desviación típica intra sujetos como una estimación del error de medida, lo que estamos haciendo es considerar el siguiente modelo matemático para cada medida X:

$$X = \pi + \varepsilon$$

donde π representa el valor verdadero de la variable que se mide (desconocido) y ε el error de la medida.

A partir de ese modelo se puede calcular un coeficiente de fiabilidad de la medida, que se conoce con el nombre de **coeficiente de correlación intra−clase**:

$$\rho = \frac{\mathrm{Var}(\pi)}{\mathrm{Var}(\pi) + \mathrm{Var}(\varepsilon)}$$

es cociente entre la varianza correspondiente a los valores verdaderos (variabilidad biológica) y la suma de la varianza de los valores verdaderos y la varianza debida al error de medida (variabilidad total). Este coeficiente no presenta los problemas ya comentados para el caso del coeficiente de correlación lineal y además como vemos se puede calcular para más de dos medidas por sujeto. Su valor para el ejemplo de la tabla 4 es *0.789*

Así el coeficiente de correlación intraclase es un indicador de la fiabilidad de una sóla medida. Cuando el proceso lo permite se efectúa más de una medida, y se utiliza la media de los valores obtenidos, para mejorar la precisión, por lo que nos interesa disponer de un indicador de la fiabilidad de la media de *k* lecturas, y éste se calcula según la conocida como **fórmula de Spearman−Brown**:

$$\rho_k = \frac{k \cdot \rho}{1 + (k-1) \cdot \rho}$$

En nuestro ejemplo la fiabilidad de la media de 3 lecturas es de *0.918*.

Lo que hemos denominado error de medida está incluyendo la variabilidad debida a causas desconocidas. En algunos casos puede existir fuente de variación atribuíble a otras causas que sí son conocidas y por tanto controlables, y nos permitirá mejorar la precisión de la medida. Así ,en el ejemplo anterior si calculamos la media de PAS1, PAS2, y PAS3 obtenemos los valores: 148.1, 138.5, y 133.9. Si las columnas 1,2,3 se corresponden con la secuencia temporal de las lecturas efectuadas, podría ser que esa tendencia a disminuir se deba a la relajación del paciente a medida que pasa el tiempo. En otras situaciones las lecturas pueden corresponder a métodos diferentes o evaluadores distintos.

Para estos casos el modelo propuesto será entonces

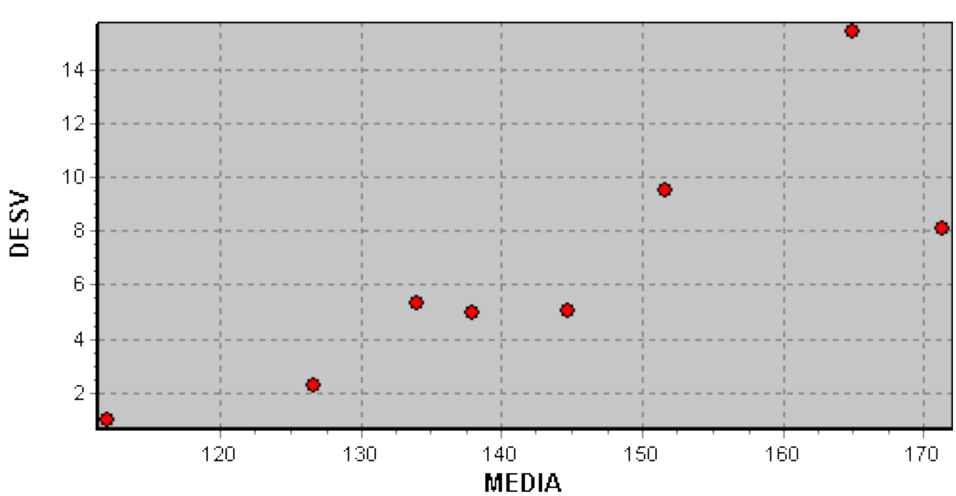
$$X = \pi + \alpha + \varepsilon$$

donde ahora hemos incorporado el término α que corresponde al efecto debido al método de medida (instante de tiempo, evaluador, sistema de medida, etc). La filosofía para el cálculo del coeficiente de correlación intraclase es la misma, pero basada en un modelo diferente en el que el valor de ε (variabilidad desconocida) es menor que en el primer modelo, debido a que parte de la variación entre lecturas se achaca ahora al método.

Error de medida proporcional a la media

Se ha comentado más arriba que cuando se dispone de al menos dos mediciones por sujeto se puede estimar el error de medida a partir de la desviación típica intra−sujeto, **siempre que dicho error sea independiente de la magnitud de la medida**. Esto habrá que comprobarlo en cada caso, pues una situación muy habitual en la práctica es que el error aumente con el valor de la magnitud medida (error proporcional a la media).

Para comprobar si el error de medida es independiente de la magnitud medida podemos construir una gráfica que represente la desviación típica de las medidas de cada sujeto en función de la media de éstas y comprobar visualmente si existe algún tipo de relación entre ambas. Así una gráfica como la de la figura



nos muestra que la desviación típica crece con el valor de la media. En estos casos para que sea aplicable el cálculo de la desviación típica intra−sujeto es necesario transformar los datos, procedimiento descrito en uno de los artículos que se indica en enlaces de interés.

Además de la valoración visual se puede realizar un contraste formal, que conviene que sea no paramétrico, como puede ser el **coeficiente de Spearman** o la **tau de Kendall**.

Enlaces de interés

- J Martin Bland and Douglas G Altman. Measurement error BMJ 1996; 313: 744
- Jacques Massé, J M Bland, J R Doyle, and J M Doyle. LETTERS Measurement error BMJ 1997; 314: 147.
- J Martin Bland and Douglas G Altman. Measurement error proportional to the mean BMJ 1996; 313: 106.
- J Martin Bland and Douglas G Altman. Measurement error and correlation coefficients BMJ 1996; 313: 41−42
- Abraira, V. Concordancia para variables continuas
- Luis Prieto, Rosa Lamarca, Alfonso Casado. La evaluación de la fiabilidad en las observaciones clínicas: el coeficiente de correlación intraclase. Medicina Clínica. Sábado 7 Febrero 1998. Volumen 110 − Número 4 p. 142 − 145

Bibliografía seleccionada

- Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet: 307−310

Indice de artículos

Principio de la página