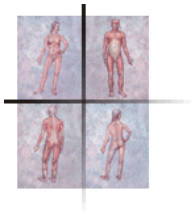
CAPÍTULO 3



ERROR EN LA MEDICIÓN ANTROPOMÉTRICA

David Pederson y Christopher Gore

1. INTRODUCCION

Si un antropometrista mide los pliegues cutáneos de un sujeto en forma reiterarla, por ejemplo en varios días consecutivos, los valores por lo general serán diferentes Si bien una pequeña parte de la variabilidad podría estar dada por cambios biológicos en el sujeto, la mayor parte se deber probablemente a inconsistencias en la técnica riel antropometrista. Por ejemplo, cuando se miden los pliegues, la ubicación de los sitios en el cuerpo podría variar ligeramente de evaluación a evaluación, o el equipo que se utiliza podría estar calibrado, cada vez, a distinto nivel. Está claro que para el antropometrista es tina ventaja minimizar la variabilidad técnica en las mediciones. En la literatura sobre errores de medición predominan cuatro temas: **precisión, confiabilidad, exactitud y validez.**

La variabilidad observada en mediciones repetidas llevadas a cabo en el mismo sujeto determina el nivel de **precisión.** Las variables de precisión normalmente tienen las mismas unidades que las unidades de la variable en consideración. Una elevada precisión se corresponde con tina baja variabilidad en evaluaciones sucesivas, y es el objetivo de un antropometrista competente, ya que entonces habrá tina alta probabilidad de que tina sola evaluación está cerca de su valor real, el valor de interés para el evaluador.

El antropometrista no tiene un solo valor para su precisión, sino que tiene un valor separado para cada combinación de una variable y de un procedimiento de medición. Para la mayoría de las variables antropométricas existen standars de precisión aceptables, los cuales permiten al evaluador controlar su propia performance. En el Capítulo 13 se discute acerca de los niveles deseables (y razonablemente alcanzables) de precisión.

La precisión es el indicador más básico de la pericia de un antropometrista. Sin embargo, si los niveles de precisión están indicados en el informe técnico, entonces los lectores necesitarán saber tanto las unidades como los standards aceptables para poder evaluar la precisión de cada variable. Una alternativa es indicar los niveles de **confiabilidad,** los cuales son, por lo general, coeficientes de correlación y, por lo tanto, no tienen unidades. La variable de confiabilidad que presentaremos en este capitulo es el **coeficiente de correlación intraclase** (CCI), para el cual existe un rango desde 0 (confiabilidad cero) hasta 1 (confiabilidad perfecta).

Por lo tanto, las variables de precisión y confiabilidad difieren en sus unidades y en las restricciones sobre sus posibles valores. Existen otras dos diferencias. En primer lugar, la precisión es tina característica de un evaluador en particular, al utilizar una técnica particular de medición en tina variable en particular. La confiabilidad tiene las mismas características, más la característica adicional de depender en la variabilidad de los sujetos. Por lo general, la medición de la confiabilidad para una variable en particular será más cercana a 1 en un grupo de sujetos que son muy diferentes entre sí que en un grupo de sujetos que son relativamente semejantes. En segundo lugar, las mediciones de la precisión podrían ser utilizadas en cálculos posteriores de, por ejemplo, intervalos de confianza, o en el tamaño de la muestra necesario para satisfacer ciertos criterios. Los valores de confiabilidad, por el contrario, son simplemente indicadores de técnica y no son útiles para cálculos posteriores. Es importante que la evaluación obtenida por un antropometrista en un sujeto en particular sea cercana al valor real. El grado al cual el valor medido se corresponde con el valor real es la **exactitud** de la medición. Pero, cuál es el "valor real "?. No existe ningún instrumento que mida los pliegues cutáneos, por ejemplo, con absoluta exactitud. Cuando se debe determinar la exactitud de alguna evaluación, normalmente se la compara con el valor obtenido por algún antropometrista altamente entrenado y experimentado (por ejemplo, un antropometrista de Nivel 3 o Nivel 4). Los valores obtenidos por dicho profesional (llamado evaluador "de criterio") se asumen como los valores reales.

El cuarto aspecto del error de medición os la **validez**, la cual es el grado por el cual una evaluación mide realmente tina característica. Por ejemplo, la variable en estudio podría ser la característica general de "capacidad física" de los sujetos. Se podría decir que tina variable o grupo de variables que sean pobres indicadores de la capacidad física, de acuerdo a algunos criterios estipulados, tendrán una baja validez. Sin embargo, raramente la validez sea un tema de preocupación para las mediciones antropométricas ya que las variables que son medidas, por lo general están bien definidas, y no son conceptos abstractos.

Debido a los distintos factores que influyen en la precisión, confiabilidad, y exactitud, la correspondencia entre ellos nunca es certera. Por ejemplo, un alto nivel de confiabilidad normalmente indica un alto nivel de precisión, pero un alto nivel de precisión no siempre está acompañarlo por un alto grado de confiabilidad. Es bastante posible (y, de hecho, bastante común) que un evaluador demuestre una alta precisión pero un bajo nivel de exactitud. Este sería el caso cuando el evaluador estaba midiendo con tina desviación o error constante, poro en forma consistente. Sin embargo, el objetivo del antropometrista debería ser lograr altos grados de precisión, confiabilidad, y exactitud, utilizando una técnica válida de medición.

Para algunas variables antropométricas las mediciones en los sujetos se califican por clases. Por ejemplo, el peso corporal de un remero podría ser clasificado como "peso pesado" o "peso liviano". Las variables de esto tipo son denominadas **cualitativas**. Sin embargo, las técnicas discutidas en esto capítulo son apropiadas sólo para variables **cuantitativas,** para las cuales los valores son números. Las técnicas son más adecuadas para variables cuantitativas, que tienen muchos valores posibles, o que pueden asumir cualquier valor dentro de un rango especifico.

2. VARIABLES DE PRECISION Y CONFIABILIDAD - ETM Y CCI

La variable de precisión que se utilizará en este capítulo es el Error Técnico de **medición** (ETM), definido como el desvío standard de mediciones Repetidas, tomadas independientemente unas de otras, en el mismo sujeto. Las unidades del ETM son las mismas que las unidades de la variable medida.

Se supone que el mismo antropometrista realiza todas las evaluaciones y que el ETM será, por lo tanto, el ETM intra-evaluador. El desvío standard de las mediciones, tomadas independientemente en el mismo sujeto, por dos o más antropometristas es el ETM inter-evaluador, el cual se supone que excederá al ETM intra-evaluador si existieran diferencias consistentes entre antropometristas. Si una de las dos tomas de mediciones en cada sujeto es realizada por un evaluador "de criterio", se podría utilizar el ETM inter-evaluador para investigar la exactitud del antropometrista, que realizó la segunda medición en cada sujeto.

El tamaño del ETM, a mentirle, está asociado con la media de la variable. Por ejemplo, para los datos de los pliegues cutáneos es común observar un bajo ETM cuando la media de la muestra es baja, y un alto ETM cuando la media de la muestra es alta. Para facilitar la comparación de los ETM recolectados en diferentes variables o diferentes poblaciones, se podría usar la siguiente fórmula para convertir el **ETM absoluto** en **ETM relativo** (**%** ETM):

$$\%ETM = \frac{ETM}{media} \times 100$$

donde, la media es el promedio general de la variable que ha sido medida. El ETM relativo da el error en porcentaje de la media total, y no tiene unidades. En otros contextos el % ETM podría estar referido como el coeficiente de variación de la variable.

La variable de confiabilidad que se utilizará en este capítulo es el coeficiente de correlación intraclase, el cual se calcula más fácilmente a partir de los resultados de un **análisis de variancia** (ANOVA). Los cuadrados medios a partir del ANOVA se combinan en una fórmula cociente o de proporción para obtener un CCI. No hay solamente un CCI para una serie determinada de datos. Si se utiliza una fórmula diferente, se obtiene un CCI diferente. Dos interpretaciones del CCI, presentadas en este capítulo, son:

- El CCI indica la correlación entre evaluaciones sucesivas en el mismo sujeto.
- El CCI indica la capacidad del procedimiento de evaluación de discriminar entre sujetos El CCI siempre es positivo y no tiene unidades. Los valores varían de 0 a 1, con el valor cercano a 1 que indica una alta confiabilidad, ya que entonces las mediciones sucesivas guardan relativamente una estrecha concordancia.

Denegar y Ball (1993) brindaron más información acerca del CCI. Ellos presentan ejemplos numéricos que demuestran que los coeficientes de correlación intraclase representan variables de confiabilidad más adecuadas que los coeficientes de correlación interclase, tal como la correlación producto-momento de Pearson. La razón principal es que los coeficientes interclase son insensibles a los cambios en la media de la muestra de los sujetos, desde un momento al siguiente, mientras que los coeficientes intraclase podrían estar influenciados por tales cambios

3. LA OBTENCION DE LOS DATOS DE CONFIABILIDAD

Un punto importante para los datos de confiabilidad es la definición de una medición, el valor que proviene de una única sesión de evaluación. Para algunas variables antropométricas, por ejemplo la estatura o el peso, la medición proviene de una lectura única. Para otras variables, por ejemplo los pliegues cutáneos o los perímetros, lo normal es realizar dos o más mediciones, y luego utilizar la media o la mediana. La ventaja de utilizar lecturas múltiples es que los errores aleatorios tienden a contrarrestarse cuando se usa la media o la mediana. La metodología para tina variable en particular constituye el procedimiento de medición, y cualquier ETM o CCI se relacionará con un procedimiento bien definido. Por ejemplo, el ETM y el CCI para la suma de 7 pliegues cutáneos serían específicos para esa variable, y casi con seguridad serán bastante diferentes del FTM y CCI para el perímetro de cabeza. Si se cambia el procedimiento, entonces se deben obtener nuevos datos y se deben recalcular el ETM y el CCI. Pon ejemplo, si un antropometrista cambia el uso de calibres Lafayette por calibres Harpenden, debería obtener nuevos datos de confiabilidad para los calibres Harpenden.

Los sujetos medidos para establecer el ETM y el CCI deben ser de la población que será evaluada en el futuro o, al menos, de una población similar. Por ejemplo, si las mediciones futuras se realizarán en un grupo de levantadores de pesas maduros, entonces los datos de confiabilidad no deben establecerse en un grupo de gimnastas prepúberes. El número de sujetos necesario para establecen el ETM y el CCI está determinado por los recursos disponibles pero normalmente debería ser no menor a veinte. Cada sujeto debería ser medido repetidamente por el mismo antropometrista. Dos mediciones son suficientes, y para facilitar el análisis, el número de mediciones debería ser la misma para cada sujeto.

Es posible realizar un análisis más informativo de confiabilidad si se toman todas las primeras series de mediciones en el mismo momento, y todas las segundas series de mediciones también en el mismo momento, y así sucesivamente. Sin embargo, es difícil clan una definición precisa de “en el mismo momento”. Para algunos procedimientos antropométricos “en el mismo momento” pueden significar dentro del período de una mañana o de un día entero. Pon lo tanto, si un antropometrista midió el pliegue triccipital en 20 sujetos donante una mañana y luego repitió las mediciones en la tarde del mismo día, podría considerarse que tanto la primera como la segunda serie de evaluaciones del pliegue del tríceps fueron realizadas “en el mismo momento”. Aún si los sujetos fueron medidos dentro de un período de dos días, y luego tres días más tarde, fueron medidos nuevamente en un período de dos días, se puede decir que tanto la primera como la segunda serie fueron realizadas "en el mismo momento". Sin embargo, si un antropometrista midió el pliegue triccipital de cada sujeto dos veces, en un periodo de tres días entre las dos mediciones, pero le llevó tres meses acumular los datos de 20 sujetos, no sería razonable suponer que la primera o la segunda evaluación fueron realizarlas "en el mismo momento”.

Para algunos procedimientos de medición los valores de los sujetos, en un momento, podrían ser consistentemente diferentes de los valores en otro momento; por ejemplo debido a un cambio en la lectura basal del aparato que está siendo utilizado. Por ejemplo, el punto cero en los calibres Harpenden podría cambiar inadvertidamente luego de una serie de mediciones de pliegues cutáneos, de forma tal que podría sumarse 1 mm a todas las lecturas en la segunda serie. La ventaja de medir a todos los sujetos en el mismo momento en la primera serie, y luego en la segunda serie, es que el análisis puede utilizarse para investigar si han ocurrido cambios entre la primera y la segunda serie de mediciones. Si se observa que hubo cambios, se pueden investigar formas de mejorar los procedimientos de medición.

4. TABLAS DE DATOS A UNA VIA Y A DOS VIAS

Los datos de confiabilidad pueden ubicarse en una tabla en la cual las columnas son los sujetos y las líneas son las mediciones repetidas (ven Tabla 1). Sin embargo, el método de cálculo del ETM y del CCI dependerá de la naturaleza de los datos y, en particular, si los datos forman una tabla a una o a dos vías. Deberíamos formulamos la siguiente pregunta: Se realizaron las mediciones en momentos separarlos o distintos, bien definidos, siendo todos los sujetos evaluados en una ocasión en el mismo momento, y posteriormente re-evaluados en una segunda ocasión

en el mismo momento, y así sucesivamente?. O, en otras palabras, existe alguna razón por la cual es de esperar que haya alguna diferencia relacionada con el tiempo, entre las series de mediciones que sea aproximadamente la misma para todos los sujetos?

- Si la respuesta es “no”, los datos constituyen una tabla a una vía.
- Si la respuesta es “si”, los datos constituyen una tabla a dos vías.

SUJETO									
1	2	3	4	5	6	7	8	9	10
9.9	8.6	11.6	10.3	11.7	9.9	10.8	9.4	7.6	8.8
9.3	8.7	10.6	10.5	11.4	9.6	11.0	9.1	7.4	8.2

TABLA 1. *Serie hipotética de datos: medición del pliegue triccipital (mm), en 10 sujetos evaluados dos veces.*

Se utiliza el término “a una vía" porque, en tal situación, los datos consisten solamente de columnas de números y las líneas no tienen significado. Los números en cada columna pueden estar reordenados sin sufrir ninguna pérdida de información. En una tabla de "a dos vías" tanto las líneas (= tiempo) como las columnas (= sujetos) tienen un significado, y cualquier reordenamiento dentro de una columna podría destruir el patrón de los datos. Cuando los datos forman una tabla “a una vía" entonces el ANOVA se denomina “ANOVA a una vía", y cuando los datos forman una tabla "a dos vías" el ANOVA es denominado "ANOVA a dos vías".

Un ANOVA a dos vías se lleva a cabo más fácilmente en una tabla completa de datos sin valores faltantes Si hay valores perdidos en una tabla a dos vías, las opciones son:

- descartar sujetos hasta que quede una tabla completa y luego llevar a cabo el ANOVA a dos vías;
- llevar a cabo un análisis con ANOVA a una vía;
- buscar asesoramiento estadístico acerca de cómo realizar un ANOVA
  a dos vías con valores faltantes, y luego cómo obtener estimaciones del CCI y del ETM.

5. CALCULO DEL ETM Y DEL CCI CON DOS MEDICIONES POR SUJETO

Muchos autores han presentado una fórmula simple para calcular el ETM a partir de una tabla de datos a una vía (Dahlberg, 1940) la cual utiliza las diferencias entre las dos mediciones. En esta sección so calcularán el ETM y el CCI a partir de una tabla de datos a una vía, llevando a cabo primero un ANOVA a una vía, pero el ETM también será calcularlo a través del método de diferencias para demostrar que las dos metodologías de cálculo dan la misma respuesta. También serán calculados el ETM y el CCI a partir de una tabla de datos a dos vías, llevando a cabo primero un ANOVA a dos vías. En las fórmulas generales, el número de sujetos estará representado por la "n" y el número de evaluaciones por sujeto estará expresado por la "k" (para el caso de números iguales) o por kl, k2,...,kn para los sujetos 1 a n (para el caso de números desiguales).
Los datos que se utilizarán como ejemplo son las mediciones del pliegue triccipital (mm) en diez sujetos (ver Tabla 1), aunque en la práctica es aconsejable obtener un mayor número de sujetos.

5.1 Tabla de datos a una vía

En esta sección se supondrá que las primeras mediciones no fueron todas realizadas en el mismo momento, ni tampoco las segundas mediciones, por lo cual los datos forman una tabla a una vía colocando a los sujetos en las columnas. Es común encontrar datos de este tipo, particularmente en situaciones en las cuales el antropometrista no tiene acceso a todos los sujetos como un grupo, sino que debe acumular datos durante varios meses, a medida que los sujetos se encuentran disponibles. En la Tabla 2 encontramos un ANOVA a una vía, el cual se obtiene ya sea utilizando las fórmulas standard presentadas en los libres de texto de estadística, o usando un paquete estadístico de computación.

Fuentes de variación	**Grados de libertad**	**Suma de los cuadrados**	**Media cuadrada**
Entre sujetos	9	29.812	3.312
Error	10	1.060	0.1060
Total	19	30.872	

TABLA 2. *Tabla de ANOVA a una vía a partir de la serie de datos presentada en a Tabla 1.*

$$ETM = \sqrt{MCe} = \sqrt{0.1060} = 0.33mm$$

donde, MC_e = error de la media cuadrada. Debido a que la media de 20 observaciones es 9.72,

$$\%ETM = \frac{ETM}{media} \times 100 = \frac{0.33}{9.72} \times 100 = 3.4\%$$

El CCI está dado por

$$CCI = \frac{MC_s - MC_E}{MC_s + (k-1)MC_e}$$

donde, MC_s = media cuadrada entre sujetos, y

$$k = \frac{\sum k_i - \frac{\sum k_i^2}{\sum k_i}}{n-1}$$

(caso de números desiguales). Ya qrie k= 2, en este ejemplo,

$$CCI = \frac{MC_s - MC_e}{MC_S + MC_e} = \frac{3.312 - 0.1060}{3.312 + 0.1060} = 0.94$$

Cuando el ETM se calcula por el *método de las diferencias,* primero se determinan las diferencias (d_t) entre la primera y la segunda medición. Para estos ciatos las diferencias son:

0.6, -0.1, 1.0, -0.2, 0.3, 0.3, -0.2, 0.3, 0.2, 0.6

La sumatoria de las diferencias ($\sum d_i$) es 2.8 y la sumatoria de los cuadrados de las diferencias ($\sum d_t^2$) es 2.12 Por lo tanto,

$$ETM = \sqrt{\frac{\sum d_i^2}{2n}} = \sqrt{\frac{2.12}{20}} = 0.33mm$$

Esta es la misma respuesta que la obtenida a partir del ANOVA.

5.2 Tabla de datos a dos vías

Los datos ahora serán analizados nuevamente suponiendo que forman una tabla a dos vías. En antropometría, esto corresponde a la situación en la cual, por ejemplo, el pliegue triccipital ha sido medido en 20 sujetos en un día, y tres días más tarde fueron re-evaluados los 20 sujetos, también en un mismo día. Es decir, hubo un intervalo de tiempo particular entre la primera y la segunda serie de mediciones, pero cada una fue realizada en el mismo momento. En la Tabla 3 se presenta un ANOVA a dos vías, usando ya sea la fórmula standard o un paquete estadístico de computación. La media en el momento 1 es de 9.86 mm y la media en el momento 2 es de 9.58 mm. A través de un test F se puede evaluar si existió una diferencia significativa entre las medias,

$$F = \frac{MC_t}{MC_e} = \frac{0.392}{0.0742} = 5.28$$

donde, MC_t = media cuadrada entre momentos u ocasiones

Fuente de variación	Grados de libertad	Suma de cuadrados	Media cuadrada
Entre sujetos	9	29.812	3.312
Entre momentos	1	0.392	0.392
Error	9	0.668	0.0742
Total	19	30.872	

TABLA 3. *Tabla de ANOVA a dos vías para la serie de datos presentada en la Tabla 1.*

Debido a que la media cuadrada entre momentos y el error de la media cuadrada tienen 1 y 9 grados de libertad, respectivamente, el cociente F también tiene 1 y 9 grados de libertad. Una tabla de la distribución F muestra que el 5 % de F, con 1 y 9 grados de libertad, excede 5.12 y, por lo tanto, se llega a la conclusión de que las medias para los momentos 1 y 2 son significativamente diferentes a un nivel de significancia del 5 %, ya que 5.28 excede a 5.12.

En este punto, en la práctica, se debería considerar porque las medías fueron estadísticamente diferentes. Si se observara que la diferencia se debió a un cambio en el procedimiento de medición, el cual podría identificarse y evitarse en el futuro, entonces se deberían recoger nuevos datos con el procedimiento modificado y se tendría que repetir el análisis desde el principio. Por ejemplo, si una investigación revelara que la primera serie de medición del pliegue triccipital fue realizada con una cuidadosa marcación del punto acromial, radial y de la distancia media acromial-radial, pero que la segunda serie fue hecha marcando los sitios "a ojo", el antropometrista debería repetir la segunda serie de evaluación utilizando una marcación cuidadosa. Sin embargo, si una investigación detallada de los resultados y de las técnicas no revelara alguna causa identificable para las diferencias entre las medidas, entonces el análisis seguirá con el cálculo del ETM y del CCI. En el presente caso, este último procedimiento es el que se va a desarrollar, para mostrar los métodos de cálculo.
En primer lugar,

$$ETM = \sqrt{\frac{(n-1)MC_e + MC_t}{n}} = \sqrt{0.1060} = 0.33mm$$

Observar que se obtuvo el mismo valor que el cálculo basado en el ANOVA de una vía. El método de las diferencias también podría haber sido utilizado para obtener el ETM, a partir de la tabla de datos de dos vías. El % ETM también es el mismo para los dos análisis, es decir 3,4 %. En términos prácticos, el ETM es bastante pequeño, y es aceptable para las mediciones de los pliegues cutáneos. Se podría

concluir que el evaluador ha realizado mediciones precisas del pliegue del tríceps.

El CCI está dado por,

$$CCI = \frac{n(MC_s - MC_e)}{nMC_s + kMC_t + (nk - n - k)MC_e}$$

Como k = 2, en este ejemplo,

$$CCI = \frac{n(MC_s - MC_e)}{nMC_s + 2MC_t + (n-2)MC_e}$$

$$CCI = \frac{10(3.312 - 0.0742)}{10(3.312) + 2(0.392) + 8(0.0742)} = 0.94$$

Observan que este valor es cercano, pero no idéntico, al CCI basado en el ANOVA de una vía.

6. CALCULO DEL ETM Y DEL CCI CON TRES MEDICIONES POR SUJETO

Cuando se disponen de tres o más mediciones de la misma variable para cada sujeto, no es aplicable el método de las diferencias para calcular el ETM, tal como se describió en la Sección 5. Sin embargo, se podrían utilizar las fórmulas generales para el ETM y el CCI que se presentaron también en dicha Sección, y esta parte contiene ejemplos de los cálculos.

Los datos serán los mismos que para el ejemplo con dos mediciones, pero con el agregado de una tercena evaluación del pliegue triccipital en cada sujeto (Tabla 4).

1	2	3	4	5	6	7	8	9	10
9.9	8.6	11.6	10.3	11.7	9.9	10.8	9.4	7.6	8.8
9.3	8.7	10.6	10.5	11.4	9.6	11.0	9.1	7.4	8.2
9.7	8.3	11.2	10.6	12.3	9.9	11.4	9.0	7.8	8.3

TABLA 4. *Serie hipotética de datos: medición del pliegue Triccipital (mm), en 1 0 sujetos, evaluados 3 veces.*

6.1 Tabla de datos a una vía

Suponiendo que los datos forman una tabla a una vía, el ANOVA a una vía arroja los resultados presentados en la Tabla *5*.

Fuente de variación	Grados de libertad	Suma de cuadrados	Media cuadrada
Entre sujetos	9	49.6630	5.518
Error	20	1.8667	0.09333
Total	29	51.5297	

TABLA 5. *Tabla del ANOVA a una vía para la serie de Datos, presentada en la Tabla 4.*

$$ETM = \sqrt{MC_e} = \sqrt{0.09333} = 0.31mm$$

Como la media de las treinta observaciones es 9.76 mm,

$$\%ETM = \frac{ETM}{media} \times 100 = \frac{0.31}{9.76} \times 100 = 3.2\%$$

Como k = 3, en esto ejemplo, el CC1 está dado por,

$$CCI = \frac{MC_s - MC_e}{MC_s + 2(MC_e)} = \frac{5.518 - 0.09333}{5.518 + 2(0.09333} = 0.95$$

6.2 Tabla de datos a dos vías

Suponiendo que los datos forman una tabla a dos vías, el ANOVA a dos vías arroja los resultados presentados en la Tabla 6. Las medidas de los pliegues triccipitales para los momentos 1, 2, y 3 son 9.86 mm, 9.58 mm, y 9.85 mm, respectivamente. Para testear si las medidas son significativamente diferentes, se calcula F,

$$F = \frac{MC_t}{MC_e} = \frac{0.2523}{0.07567} = 3.33$$

Fuente de variacion	Grados de libertad	Suma de cuadrados	Media cuadrada
Entre sujetos	9	49.6630	5.518
Entre momentos	2	0.5047	0.2523
Error	18	1.3620	0.07567
Total	29	51.5297	

TABLA 6. *Tabla del ANOVA a dos vías para la serie de datos, en la Tabla 4.*

Como el percentil 95 de F, con 2 y 18 grados de libertad, es 3.55, se puede concluir que las medias no son significativamente diferentes al nivel de significancia del 5%. Sin embargo, 3.33 está cerca de 3.55, por lo cual se deberían considerar las posibles razones por las cuales la segunda media está por debajo de las otras dos. Si no se observan razones de importancia, entonces se dará por aceptado el procedimiento de medición en su forma actual, y se procederá con el análisis.

$$ETM = \sqrt{\frac{(n-1)MC_e + MC_t}{n}} = \sqrt{\frac{9(0.07567) + 0.2523}{10}} = 0.31mm$$

Nuevamente, se obtuvo la misma respuesta que la calculada a través del ANOVA a una vía, y el % ETM es el mismo para los dos análisis, es decir 3.2 %.

Como k = 3, en este ejemplo,

$$CCI = \frac{n(MC_S - MC_e)}{nMC_s + 3MC_t + (2n-3)MC_e}$$

$$= \frac{10(5.518 - 0.07567)}{10(5.518) + 3(0.2523) + 17(0.07567)} = 0.95$$

Nuevamente, este valor es muy cercano, pero no idéntico, al CCI basado en el ANOVA a una vía.

7. INTERPRETACION Y APLICACION DEL ETM Y DEL CCI

7.1 CCI

El ANOVA a una vía, de los datos con dos mediciones por sujeto, del pliegue triccipital (ver Sección *5.1)* dio un CCI de 0.94 y la interpretación de este valor se desarrolla en el siguiente párrafo.

Supongamos que el antropometrista intenta utilizar el mismo procedimiento de evaluación en un grupo de sujetos a partir de la misma población, o de una similar, a aquella de la cual tomó 10 sujetos en la muestra testeada. Entonces una estimación de la correlación entre evaluaciones sucesivas en un solo sujeto será de 0.94. Como 0.94 está muy cerca del máximo CCI posible, que es 1, es de esperar un alto nivel de concordancia entre las sucesivas mediciones. El procedimiento de evaluación tendrá un alto nivel de confiabilidad. El valor de 0.94, obtenido a partir del ANOVA a dos vías con los mismos datos, sería interpretado en forma similar.

7.2 ETM

El ETM podría ser utilizado de diversas maneras. Se presentarán cuatro ejemplos; los dos primeros, usando el ETM de 0.33 mm obtenido a partir de los datos con dos mediciones del pliegue triccipital por sujeto (ver Sección 5.1). Todos estos ejemplos suponen que el procedimiento de medición no cambió, ya que fue calculado el ETM. Idealmente, los sujetos evaluados en el futuro deberían pertenecer al grupo (o a la población) utilizada para obtener el ETM. Observar también que un valor de ETM es único para el antropometrista que obtuvo *los* datos, aunque los valores para distintos antropometristas que tienen elevada experiencia, normalmente, serán muy cercanos entre sí.

Ejemplo 1
Un deportista es evaluado y *se obtiene un pliegue triccipital de 9.3 mm. Cuáles son los límites de confianza del valor verdadero para dicho deportista?*

El error standard de la medición será de 0.33 mm, y aproximadamente un intervalo de confianza del 68 % para el valor verdadero, provendrá de

[9.3 - 0.33] a [9.3 + 0.33]; por ej., de 9.00 mm a 9.6 mm.
Alternativamente, un intervalo de confianza del 95 % para el valor real, será

[9.3 - 2(0.33)] a [9.3 + 2(0.33)], dc 8.6 mm a 10 mm.

En general, el ETM es el error standard de una sola medición. (La medición + / - ETM) da el intervalo aproximado de confianza del 68 % para el valor real del sujeto. (La medición + / - 2 x ETM) da el intervalo de confianza del 95 % para el valor verdadero del sujeto.
Aunque el ETM es el desvío standard de mediciones repetidas, se ha utilizado el término error standard, dado que es así por convención, cuando el problema que está siendo considerado es de un muestreo y estimación subsecuente. En este caso, la estimación es del valor verdadero del sujeto.

Ejemplo 2
Un deportista es evaluado dos veces, en un lapso de un mes, con la intención de testear si el valor verdadero del atleta ha cambiado, por ejemplo, entre las dos mediciones sucesivas como resultado de alguna dieta o entrenamiento. Los valores son 9.3 mm, seguido por 8.5 mm.

Cada evaluación tendrá un error standard de 0.33 mm. La diferencia entre las mediciones es 8.5- 9.3 = - 0.8 mm y el error standard de la diferencia es 0.33 x √ 2 = 0.47 mm. Un intervalo de confianza de 68 % para el cambio verdadero, derivará de

[- 0.8 - 0.47] a [- 0.8 + 0.47], por ej., de -1.3 mm a - 0.3 mm.

Debido a que el intervalo no incluye el cero, se podría concluir que el sujeto cambió. Efectivamente, se ha llevado a cabo un test para probar la hipótesis nula de que el valor verdadero no había cambiado, y que la hipótesis había sido rechazada al nivel de significancia de 32% (el complemento de 68%).

Alternativamente, un intervalo de confianza de 95% para el cambio real, derivará de

[- 0.8 – 2 (0.47)] a [- 0.8 + 2(0.47)], por ej., de -1.7 mm a 0.1 mm.

Como el intervalo de confianza incluye el cero, se podría concluir (al nivel de significancia del 5%) que el valor real del sujeto no cambió entre las dos evaluaciones sucesivas. En general, el ETM x 2 da el error OJO standard de la diferencia entre 2 mediciones sucesivas. (La diferencia +/-error standard) da el intervalo aproximado de confianza de 68 % para el cambio real. (La diferencia +/- 2 errores standard) da el intervalo de confianza del 95 % para el cambio verdadero.

Ejemplo 3
Se mide dos veces la suma de siete pliegues cutáneos (∑ 7) *en una gimnasta de elite, con dos semanas de intervalo entre evaluaciones, para controlar su composición corporal. El entrenador quiere saber si su grasa subcutánea se mantiene a un nivel constante. Los valores de la* ∑ 7 *fueron de 38 mm en la primera medición,* y *de 41 mm, dos semanas más tarde. Antes de llevar a cabo estas mediciones, el antropometrista había obtenido un ETM de 2.9 mm para* la ∑ 7 *pliegues en una muestra de 30 gimnastas mujeres de élite.*

La diferencia entre la semana 0 y la semana 2 es 38 - 41 = - 3 mm, y el error standard de la diferencia es 2.9 x √ 2 = 4.1 mm. El intervalo de confianza del 68 % para el cambio real es [-3 - 4.1] a [-3 + 4.1], es decir de - 7.1 mm a 1.1 mm. Como el intervalo incluye el cero, se podría concluir que (al nivel de significancia del 32 %) el valor real de la gimnasta no cambió. Alternativamente, el intervalo de confianza del 95% para el valor real es [-3 - 2(4.1)] a [-3 + 2(4.1)], es decir - 11.2 mm a 5.2 mm. Nuevamente, como el intervalo incluye el cero, se concluye que (en el nivel de significancia del 5 %) el valor verdadero de la deportista no cambió.

Ejemplo 4
Supongamos que, para la situación considerada en el ejemplo 3, el entrenador está buscando que la ∑7 *no sea mayor a 40 mm. Habiendo observado 41 mm en la segunda medición, puede el entrenador concluir que el valor verdadero de la > 7 excedía 40 mm en el momento de la evaluación?*
Un intervalo de confianza aproximado de 68 % para el valor real, será

[41 - 2.9] a [41 + 2.9] , es decir, de 38 mm a 44 mm.

Como 40 está incluido dentro del intervalo, se podría concluir (al nivel de significancia del 16 %), que el valor real de la gimnasta no excedía 40 mm. La diferencia entre el valor observado de 41 mm y el límite superior de 40 mm está dentro del rango atribuíble al error aleatorio. El nivel de significancia es 16 % más que 32 %, lo cual es de esperar ya que el intervalo era un intervalo de confianza de 68 % ya que el test llevado a cabo es en una sola dirección. El interés está en si el valor real excede 40 mm, más que en saber si el mismo no es igual a 40 mm.

Alternativamente, el intervalo de confianza del 95 % para el valor real, será

[41 - 2(2.9)] a [41 + 2(2.9)], es decir, de 35 mm a 47 mm.

Como 40 está dentro del intervalo, se podría concluir que (en el nivel de significancia de 2.5%) el valor verdadero de la gimnasta no fue mayor a 40 mm.

8. MARCO TEORICO

Consideremos la medición del pliegue triccipital de 9.9 mm, en el primer sujeto en la Tabla 1. Podría haberse dado el caso de que el valor real del sujeto en aquel momento, tomado a largo plazo, fuera de 9.6 mm pero que hubiera un error de 0.3 mm atribuible, podríamos decir, a un error de medición o a la variabilidad biológica del sujeto, a través del tiempo. Se podría formular la siguiente ecuación:

$$\mathbf{9.9 = 9.6 + 0.3}$$

Por lo general, se puede formular el siguiente modelo (Modelo 1):

medición en el sujeto = valor real del sujeto + error

Este es el modelo adecuado si se considera que los datos forman una tabla a una vía. Es el modelo que forma la base para el ANOVA a una vía. Se supone que el error promedio es cero, por lo que la medición es igual al valor real. La variancia de los efectos del error se podría escribir como σe^2 (1) y, si se puede asumir que los sujetos forman una muestra aleatoria de una población de sujetos, entonces la variancia ele los valores reales de los sujetos podría ser σs^2 (1).

Para el Modelo 1, el ETM es la raíz cuadrada de la estimación de σ_e^2 (1), y el CCI es el cálculo de

$$\frac{\sigma_s^{\ 2}(1)}{\sigma_s^{\ 2}(1) + \sigma_e^{\ 2}(1)}$$

Por lo tanto, el CCI es el cálculo de la proporción de la variabilidad combinada para los sujetos, y el error que podría atribuirse a la variabilidad entre los mismos.
Ahora, consideremos el caso en el cual los datos forman una tabla a dos vías (tal como fue discutido por Denegar & Ball, 1993; Guilford, 1965; Verducci, 1980). Nuevamente, podría darse el caso de que el valor real del sujeto, tomado a largo plazo, fuera de 9.6 mm, pero que hubiese un aumento (aplicable a todos los sujetos) de 0.5 mm, asociado con la primera medición y un error (único para el sujeto) de - 0.2 mm. Entonces, podría formularse la siguiente ecuación:

9.9 = 9.6 + 0.5 - 0.2

Por lo general, este modelo (Modelo 2) puede formularse de la siguiente manera:

medición en el sujeto = valor real del sujeto + efecto del tiempo + error

Este es un modelo adecuado si se considera que los datos forman una tabla a dos vías. Cada sujeto tiene el mismo efecto del tiempo incluido en su primer medición, y cada sujeto tiene el mismo efecto del tiempo (probablemente distinto del efecto del tiempo o momento 1) incluido en la segunda evaluación. Este modelo forma la base para el ANOVA a dos vías. El efecto del tiempo o del momento no está presente en el Modelo 1, y la variancia de los efectos del tiempo podría formularse como σ_t^2 La variancia de los efectos de los sujetos es σs^2 (2), y la variancia de los efectos del error es σe^2 (2).

Bajo las Secciones 5.2 y 6.2, que consideran el ANOVA a dos vías, se sugirió que debería utilizarse un test F para determinar si los efectos del tiempo fueron diferentes entre sí. Si se observaba que fueron diferentes, el procedimiento de evaluación era investigado para ver si un procedimiento mejorado (con marcaciones más rigurosas en el sujeto) podría eliminar o, al menos, minimizar las diferencias entre los tiempos o momentos de las mediciones. Si se hiciese esto, y aún persistiesen las diferencias entre tiempos o momentos, entonces el método utilizado sería con relación a la persistencia de la variabilidad entre tiempos, como un componente de un inevitable error aleatorio.

Para el Modelo 2, el ETM es la raíz cuadrada de la estimación de $[\sigma_e^2 (2) + \sigma_t^2]$. Si una tabla de datos a dos vías es analizada con un ANOVA a una vía, y luego con un ANOVA a dos vías, los ETM de los dos análisis serán iguales, ya que $\sigma_e^2 (1) = \sigma_e^2 (2) + \sigma_t^2$.

Para el Modelo 2, el CCI es la estimación de

$$\frac{\sigma_s^2(2)}{\sigma_s^2(2) + \sigma_t^2 + \sigma_e^2(2)}$$

Por lo tanto, el CCI es la estimación de la proporción de la variabilidad combinada para los sujetos, tiempo, y error que pueden atribuirse a la variabilidad entre sujetos. Si una tabla de datos a dos vías es analizada con un ANOVA a una vía, y luego con uno a dos vías, los CCI de los dos análisis serán cercanos entre si, si la variabilidad entre tiempos es pequeña.
En general, el CCI será cercano a 1 si existe una alta variabilidad entre sujetos [σs^2 (1) o σs^2 (2) es grande], o si existe una baja variabilidad entre las mediciones repetidas en el mismo sujeto [σe^2 (1) o σe^2 (2) es pequeño], o si existen ambas condiciones. Para el Modelo 1, un valor de 1 para el CCI indicaría una perfecta capacidad para discriminar entre sujetos [σe^2 (1) = 0; las mediciones repetidas en el mismo sujeto son idénticas], y un valor de 0 indicaría ninguna capacidad discriminatoria [σs^2 (1) = O; todos los sujetos son iguales en su efecto]. Una afirmación similar podría formularse para el CCI en el Modelo 2, si la variabilidad entre momentos o tiempos fuera insignificante. Desde el punto de vista de la capacidad de las mediciones para discriminar entre sujetos, es deseable, por lo tanto, obtener un valor de CCI cercano a 1. Si se compararan dos procedimientos diferentes de medición para un único grupo de sujetos, entonces podría utilizarse el CCI para realizar dicha comparación, ya que tanto σs^2 como σe^2 , podrían concebirse que dependen del mismo procedimiento de medición. Si se compararan dos laboratorios, cada uno utilizando el mismo protocolo de medición, se podría utilizar el CCI, pero una comparación justa requeriría la suposición de que los sujetos utilizados por cada laboratorio provienen de la misma población. De lo contrario, un laboratorio podría tener un menor CCI que el otro, dado que los sujetos que utilizó eran inherentemente menos variables (σs^2 era más bajo), y no por una deficiencia en el procedimiento de evaluación, en cual resultó en un elevado σ_e^2. En el Capítulo 13, se discuten los valores mínimos de ETM para los distintos niveles de acreditación.

9. REFERENCIAS BIBLIOGRAFICAS

Dahlberg, G. (1940).
Statistical methods for medical and biological students.
London: George Allen & Unwin.

Denegar, C.R., & Ball, D.W. (1993).
Assessing reliability and precision of measurement: an introduction to intraclass correlation and standard error of measurement.
Journal of Sports Rehabilitation, 2, 35 - 42.

Guilford, J.P. (1965).
Fundamental statistics in psychology and education.
New York: McGraw-Hill.

Verducci, E.M. (1980).
Measurement concepts in physical education.
London: C.V. Mosby.