

Análisis no paramétrico de la interacción de dos factores mediante el contraste de rangos alineados

Juan Carlos Oliver Rodríguez, Julio González Álvarez y Jesús Rosel Ramirez
Universitat Jaume I

Problemas que requieren un análisis no paramétrico de la interacción surgen con cierta frecuencia en estudios del comportamiento. Hay, sin embargo, una ausencia de procedimientos en paquetes estadísticos usuales como SAS o SPSS. El objetivo del presente trabajo es revisar los fundamentos del contraste de rangos alineados en dos diseños de investigación de dos factores de amplio uso en Psicología. El procedimiento ha mostrado buenas propiedades en numerosas distribuciones no normales tanto en control del Error Tipo I como en potencia o sensibilidad estadística. Es fácilmente ejecutable mediante programas estadísticos comunes y se ilustra paso a paso con el módulo del modelo lineal general de Spss. Se aplica a dos estudios de caso donde aparecen diferencias de género en habilidades lingüísticas en niños ciegos no detectadas con otros procedimientos. Limitaciones al contraste de rangos alineados se han observado en situaciones de heterogeneidad de covarianzas entre grupos, y también con tamaños muestrales inferiores a diez participantes por condición. Su uso adecuado tras el diagnóstico de supuestos puede aumentar la sensibilidad de detección de efectos comportamentales de interés teórico o aplicado.

An aligned rank test for a nonparametric analysis of the two way interaction. Research problems that require nonparametric analysis of the interaction frequently arise in the behavioral sciences. There is, however, a lack of available procedures in commonly used statistical packages such as SAS or SPSS. The purpose of the present study is to review the fundamentals of the aligned rank test for two widely used two-way research designs in psychology. The procedure has shown good properties in nonnormal distributions in terms of Type I Error control and statistical power. It is easily conducted using common statistical packages. It is applied to two case studies which result in gender differences in linguistic abilities in blind children not revealed by other procedures. Limitations of the aligned rank test have been observed in situations of covariance heterogeneity across groups, and also with sample sizes smaller than ten participants per condition. Its adequate use after model diagnostics can, however, increase sensitivity to detect behavioral effects of theoretical or practical interest.

Problemas que requieren un análisis no paramétrico de la interacción surgen con cierta frecuencia en estudios del comportamiento. En una revisión de varios cientos de distribuciones de datos procedentes tanto de artículos de investigación como de evaluaciones psicométricas institucionales se encontró una proporción muy alta de casos en donde se incumplía el supuesto de normalidad (Micceri, 1989). El cincuenta por ciento de las distribuciones mostraban una densidad de puntuaciones en las colas significativamente mayor que en el modelo normal, y el treinta por ciento presentaban asimetría extrema, con aproximaciones a la función exponencial. En muchas situaciones de este tipo, las pruebas no paramétricas muestran mejores propiedades que las pruebas clásicas en términos de potencia, eficiencia o en sesgos de error Tipo I (Wilcox y Keselman, 2003). Sin embargo, y a pesar de

sus implicaciones para la investigación comportamental básica o aplicada, el análisis no paramétrico de la interacción no está incorporado en paquetes estadísticos usuales.

Un procedimiento propuesto consiste en realizar un contraste paramétrico mediante el estadístico F calculado a partir de la transformación de las observaciones a rangos de mayor a menor (Conover y Iman, 1981). En diseños de un factor, este procedimiento es equivalente a las pruebas de Mann-Whitney y Wilcoxon para comparaciones entre o intra sujetos. Sin embargo, en diseños de dos factores la presencia de efectos principales confunde el efecto de interacción, inflando la probabilidad del Error Tipo I (Blair, Sawilowski y Higgins, 1987; Thompson, 1991a, 1991b). Una solución propuesta es tratar los efectos principales como factores de confusión y sustraer su influencia de las observaciones antes de la transformación a rangos y del cálculo del estadístico F . Este procedimiento denominado *Contraste de Rangos Alineados* ha mostrado ser robusto en términos de control del error tipo I y ventajas de potencia estadística en numerosas distribuciones no normales (Beasley, 2002; Gorham, 1998; Mansouri y Chang, 1995; Salter y Fawcett, 1993; Toothaker y Newmann, 1994). El presente trabajo tiene como objetivo revisar sus fundamentos en

dos diseños de amplio uso en Psicología e ilustrar su aplicación mediante dos estudios de caso.

Análisis en un diseño entresujetos de dos factores

El modelo lineal de efectos fijos para este diseño puede ser especificado del siguiente modo (Arnau, 1990; Ato y Vallejo, 2007):

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + e_{ijk} \quad (1)$$

para $i = 1, 2, \dots, a$ niveles del primer factor,
 $j = 1, 2, \dots, b$ niveles del segundo factor,
 y $k = 1, 2, \dots, n$ participantes,

en donde expresamos los supuestos (distribución normal, independencia y homogeneidad de varianzas) como:

$$e_{ijk} \sim NID(0, \sigma_e^2) \quad (2)$$

A continuación se aplica la siguiente transformación, denominada de rangos alineados:

$$y'_{ijk} = \text{Rango}(y_{ijk} - \alpha_i - \beta_j) \quad (3)$$

de forma equivalente

$$y'_{ijk} = \text{Rango}(\mu + \alpha\beta_{ij} + e_{ijk})$$

o

$$y'_{ijk} = \text{Rango}(\alpha\beta_{ij} + e_{ijk}) \quad (4)$$

puesto que la sustracción de la media total no altera el resultado de la transformación ordinal. Alternativamente,

$$y'_{ijk} = \text{Rango}[(\mu_{ij} - \mu_i - \mu_j + \mu_{..}) + e_{ijk}] \quad (5)$$

en donde,

- μ_{ij} es la media obtenida en la condición experimental definida por la conjunción de niveles i del factor A y j del factor B.
- μ_i es la media marginal del nivel i del factor A.
- μ_j es la media marginal del nivel j del factor B.
- y $\mu_{..}$ es la media total.

Contraste de hipótesis

La hipótesis nula de no interacción entre los dos factores puede ser formulada como:

$$H_0 : \alpha\beta_{ij} = 0 \text{ para cualquier condición } ij \quad (6)$$

Con el fin de elegir un estadístico para su contraste especificamos los valores esperados de las medias de cuadrados de las fuentes de variación implicadas en el análisis de varianza (Kirk, 1995):

$$E[MC_{A \times B}] = \sigma_e^2 + n \sum \sum (\alpha\beta)_{ij}^2 / (a-1)(b-1) \quad (7)$$

correspondiendo los dos últimos términos a los componentes aleatorio de error y el efecto fijo de la interacción, respectivamente, y

$$E[MC_e] = \sigma_e^2, \text{ que corresponde al componente aleatorio de error debido a diferencias individuales} \quad (8)$$

Dado que $E[MC_e]$ contiene los mismos términos $E[MC_{A \times B}]$ exceptuando el efecto fijo correspondiente a la hipótesis nula, el estadístico F se construye en la variable natural mediante el cociente entre ambas medias de cuadrados (Hocking, 1996):

$$F_{A \times B} = \frac{MC_{A \times B}}{MC_e} F_{(a-1)(b-1), ab(n-1)} \quad (9)$$

mostrando el último término los grados de libertad en el numerador y denominador de la distribución F resultante.

Para la ejecución de la prueba no paramétrica, aplicaremos entonces el mismo procedimiento anterior

$$F_{RA} = \frac{MC'_{A \times B}}{MC'_e} F_{(a-1)(b-1), ab(n-1)} \quad (10)$$

en donde las medias de cuadrados $MC'_{A \times B}$ y MC'_e son calculadas sobre la variable transformada en la expresión (5), denominándolo un contraste F de Rangos Alineados (F_{RA}). Sus supuestos estadísticos no incluyen la normalidad, pero sí la independencia de las observaciones y la identidad de las distribuciones entre grupos. El procedimiento puede ser ejecutado mediante paquetes estadísticos usuales (Salter y Fawcett, 1993).

Ejecución en Spss o Sas

En primer lugar estimaremos la nueva variable de rangos alineados en (5) mediante la siguiente expresión:

$$y'_{ijk} = \text{Rango}[(\bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}_{..}) + e_{ijk}] \quad (11)$$

Ésta puede obtenerse fácilmente en los programas *Sas* o *Spss*, especificando los factores A y B como fijos en el módulo del modelo lineal univariado y las opciones por defecto. Se obtendrán entonces las estimaciones por mínimos cuadrados de las medias marginales de filas y columnas, y de las condiciones definidas por la conjunción de niveles de ambos factores. En *Spss* éstas se obtienen en el cuadro de diálogo *opciones* (Searle, Speed y Milliken, 1980). En *Sas* se puede utilizar el comando *lsmeans* también en el módulo del modelo lineal general (Salter y Fawcett, 1993). Las estimaciones del error se obtienen en *Spss* mediante el cuadro de diálogo *Guardar* y marcando la casilla *residuales*. En *Sas* puede utilizarse el comando *Outstat* con la opción *residuals*.

Las medias en la expresión (11) se insertan entonces en la tabla de datos como cuatro nuevas variables más la columna de residuales. A continuación se calcula y'_{ijk} evaluando la expresión aritmética y realizando la transformación ordinal mediante el procedimiento *Asignar rangos a casos* en *Spss* o *Proc Rank* en *Sas*. Finalmente, obtendremos la prueba F_{RA} mediante un análisis de varianza ordinario de dos factores aplicado a la nueva variable de rangos alineados.

Análisis de un caso

En un estudio sobre el uso del lenguaje visual se utilizaron cuatro muestras estratificadas de 30 niños cada una, definidas por Ha-

bilidad Visual (ciegos y videntes) y Género (niños y niñas) (Rosel, 1982). Como variable dependiente utilizaremos aquí el promedio del número de palabras con contenido visual en dos habilidades de expresión verbal: a) narración de una historia inventada por el niño sobre la base de unas palabras guía como familia, padres y amistad, y b) descripción de uno de los caracteres de la historia. Tenemos, por tanto, un diseño entre sujetos de dos factores, al que podemos aplicar el modelo lineal en la expresión (1).

El contraste F de la interacción de dos factores definido en (9) no resultó estadísticamente significativo, $F(1, 116) = .77, p = .3814$. Los análisis diagnósticos de residuales revelaron, sin embargo, una marcada asimetría con un valor de 2.07 (figura 1). El supuesto de normalidad fue rechazado conforme al contraste de Shapiro-Wilks, $W = .82, p < .001$.

Los resultados en la literatura revelan un aumento de la potencia estadística y niveles similares de error tipo I para la prueba de rangos alineados en comparación con el estadístico F en estas situaciones (Salter y Fawcett, 1993). Por este motivo, se aplicó el estadístico F_{RA} en este conjunto de datos, resultando en una interacción significativa entre Género y Capacidad Visual, $F_{RA}(1, 116) = 5.51, p = .02$. Los niños videntes tuvieron mejor ejecución que los ciegos, mientras que en las niñas sucedió lo contrario (figura 2).

Diseño mixto de dos factores con medidas parcialmente repetidas

Definimos en primer lugar el modelo lineal para un diseño de dos factores y con un factor intrasujeto como:

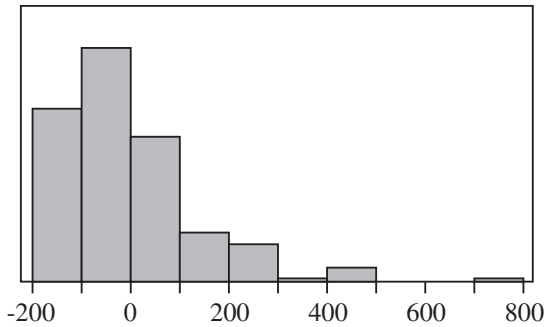


Figura 1. Distribución de residuales entre sujetos de las puntuaciones de habilidad lingüística
Nota: $S = 144.46$; Asimetría = 2.07; Curtosis = 6.15

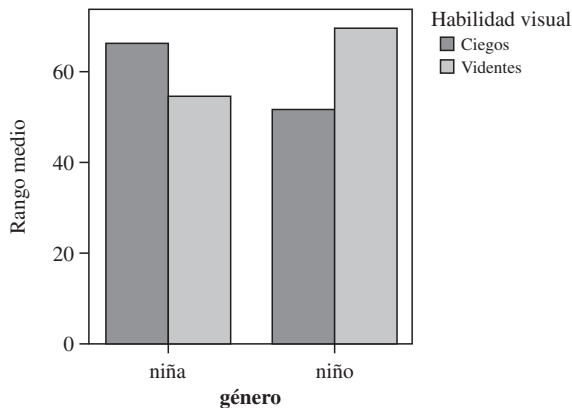


Figura 2. Rango medio en habilidad de expresión lingüística en función del género y capacidad visual

$$y_{ijk} = \mu + \alpha_i + s_k + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} \tag{12}$$

para $i = 1, 2, \dots, a$ niveles del primer factor,
 $j = 1, 2, \dots, b$ niveles del segundo factor,
 y $k = 1, 2, \dots, n$ participantes,

en donde los niveles del factor A son aplicados o medidos en distintos grupos de participantes, y los distintos niveles del factor B son medidas intrasujeto. Además, los componentes aleatorios y supuestos se especifican como:

$$\left. \begin{matrix} s_k \sim NID(0, \sigma_s^2) \\ \epsilon_{ijk} \sim NID(0, \sigma_e^2) \end{matrix} \right\} \text{ y ambos son independientes} \tag{13}$$

Un supuesto adicional en este modelo es el de esfericidad multimuestra, que tiene dos componentes: a) igualdad de varianzas entre cualquier par de diferencias entre los niveles del factor intrasujeto (Keselman, Algina y Kowalchuk, 2002), y b) homogeneidad de estructura de covarianzas entre grupos.

En segundo lugar, definiremos la siguiente transformación a rangos alineados:

$$y'_{ijk} = \text{Rango}(y_{ijk} - \alpha_i - \beta_j - s_k) \tag{14}$$

De forma equivalente:

$$y'_{ijk} = \text{Rango}(\mu + \alpha\beta_{ij} + \epsilon_{ijk})$$

$$0$$

$$y'_{ijk} = \text{Rango}(\alpha\beta_{ij} + \epsilon_{ijk}) \tag{15}$$

dado que la sustracción de la media total no altera la transformación ordinal. Alternativamente,

$$y'_{ijk} = \text{Rango}[(\mu_{ij} - \mu_i - \mu_j + \mu_{..}) + \epsilon_{ijk}] \tag{16}$$

Contraste de hipótesis

La hipótesis nula de no interacción entre los dos factores puede ser formulada como:

$$H_0 : \alpha\beta_{ij} = 0 \text{ para cualquier condición } ij \tag{17}$$

Con el fin de elegir un estadístico para su contraste, especificaremos los valores esperados de las medias de cuadrados de las fuentes de variación implicadas en el análisis de varianza (Kirk, 1995):

$$E[MC_{A \times B}] = \sigma_e^2 + n \sum (\alpha\beta)_{ij}^2 / (a-1)(b-1) \tag{18}$$

correspondiendo los dos últimos términos al error de medida y al efecto fijo de la interacción, respectivamente, y

$$E[MC_e] = \sigma_e^2, \text{ o varianza de error de medida} \tag{19}$$

Dado que $E[MC_e]$ contiene los mismos términos que $E[MC_{A \times B}]$ exceptuando el efecto fijo correspondiente a la hipótesis nula, el estadístico F se construye en la variable natural mediante el cociente entre ambas medias de cuadrados (Hocking, 1996):

$$F_{A \times B} = \frac{MC_{A \times B}}{MC_e} F_{(a-1)(b-1), a(b-1)(n-1)} \tag{20}$$

mostrando el último término los grados de libertad en el numerador y denominador de la distribución *F* resultante. Para la ejecución de la prueba no paramétrica, aplicaremos entonces el mismo procedimiento anterior

$$F_{RA} = \frac{MC'_{A \times B}}{MC'_e} F_{(a-1)(b-1), a(b-1)(n-1)} \tag{21}$$

en donde las medias de cuadrados $MC'_{A \times B}$ y MC'_e son calculadas ahora sobre la variable transformada en la expresión (16), denominándolo un contraste *F* de Rangos Alineados (F_{RA}). Sus supuestos estadísticos no incluyen la normalidad, pero sí la independencia de las observaciones para los distintos sujetos, la identidad de las distribuciones entre grupos y la esfericidad multimuestra (Beasley y Zumbo, 1998; Harwell y Serlin, 1994; Zimmerman y Zumbo, 1993). El procedimiento puede ser ejecutado mediante paquetes estadísticos usuales.

Análisis alternativos ante el incumplimiento del supuesto de esfericidad

Cuando el supuesto de esfericidad se incumple la distribución del estadístico *F* está sesgada positivamente (Box, 1954). Como consecuencia se produce un aumento del Error Tipo I, y la proporción de rechazos erróneos de la hipótesis nula será mayor que α .

La corrección del sesgo para la prueba *F* se realiza habitualmente mediante la multiplicación de los grados de libertad de numerador y denominador por el parámetro ϵ , que es función del grado en que la matriz de covarianzas se aleja del supuesto de esfericidad (Huyhn y Feldt, 1976). Así, la distribución aproximada del estadístico *F* será la siguiente:

$$F_{A \times B} = \frac{MC_{A \times B}}{MC_e} F_{(a-1)(b-1)\epsilon, a(b-1)(n-1)\epsilon} \tag{22}$$

La estimación del valor de ϵ y los ajustes de grados de libertad y significación estadística son proporcionados por los procedimientos de medidas repetidas de paquetes estadísticos como *Spss* o *Sas*. Sin embargo, Lecoutre (1991) detectó un error para el modelo mixto que todavía no ha sido corregido en éstos y propuso un procedimiento que produce en distribuciones normales una reducción adicional del sesgo (Blanca, 2004; Chen y Dunlap, 1994). La estimación debería calcularse de acuerdo a la siguiente fórmula:

$$\hat{\epsilon} = \frac{(N - a + 1)(b - 1)\tilde{\epsilon} - 2}{(b - 1)[N - a - (b - 1)\tilde{\epsilon}]}$$

en donde *N* es el número total de sujetos, *a* es el número de grupos y *b* es el número de niveles del factor intrasujeto. $\tilde{\epsilon}$ es el valor de la estimación de Huyhn y Feldt del parámetro de no esfericidad para un diseño mixto y es proporcionado por *Sas* o *Spss*. Beasley (2002) ha mostrado cómo esta corrección produce también resultados satisfactorios al ser aplicada a rangos alineados en distribuciones asimétricas con carácter exponencial y simétricas con alta densidad en las colas.

Otra alternativa es utilizar un análisis multivariado de varianza, también proporcionado rutinariamente por *Sas* o *Spss* (Vallejo y Lozano, 2006). Su potencia estadística en comparación con el procedimiento univariado depende de *N*, *b* y ϵ . Una recomendación aproximada es utilizar el procedimiento multivariado cuando se incumple el supuesto de esfericidad y el cociente *N/b* es cercano o superior a 5 (Algina y Keselman, 1997). En estos casos, el contraste de Hotelling-Lawley muestra también ventajas de potencia sobre la prueba univariada en un modelo mixto cuando se aplica a los rangos alineados en distribuciones asimétricas con carácter exponencial o simétricas con alta densidad en las colas (Beasley, 2002).

Ejecución con el programa Spss o Sas

En primer lugar estimaremos la nueva variable de rangos alineados en (16) mediante la siguiente expresión:

$$y'_{ijk} = \text{Rango} [(\bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}_{..}) + \epsilon_{ijk}] \tag{23}$$

El resultado puede obtenerse fácilmente con los programas *Spss* o *Sas*. En el caso del diseño con un factor intrasujetos realizaremos un modelo anova mixto de dos factores. Para ello es conveniente transponer la matriz de datos a un formato univariado con cuatro columnas: Factor A entresujetos, B intrasujetos, Participante y Variable Dependiente (tabla 1). A continuación introduciremos los factores A y B como efectos fijos y Participante como efecto aleatorio. Incluiremos en la opción modelo personalizado todas las fuentes de variación de este diseño: factor A, B, Participante(A), y AxB.

Los componentes de la expresión (23) se obtienen en *Spss* estimando las medias en el cuadro de diálogo *Opciones*, y guardando los residuales. Se insertan entonces en la tabla de datos como nuevas variables, se calcula el efecto de interacción y se suma el error. A continuación se realiza con el resultado la transformación ordinal mediante el procedimiento *Asignar Rangos a Casos* en *Spss* o *Proc Rank* en *Sas*.

Finalmente, obtendremos la prueba F_{RA} mediante un análisis de varianza mixto ordinario de dos factores aplicado a la nueva variable de rangos alineados. Es conveniente devolver la tabla a un

Tabla 1 Conversión de las medidas intrasujetos entre formatos univariado y multivariado para la tabla de datos en un diseño mixto de dos factores		
Formato de medidas intrasujeto	Sintaxis SPSS	Sintaxis SAS
De filas a columnas	CASESTOVARS /ID a sujeto /INDEX = b /GROUPBY=INDEX	Proc transpose prefix=b out=mult; var y; id b; by a sujeto;
De columnas a filas	VARSTOCASES /MAKE b FROM b1 b2 /INDEX = b(2) /KEEP = a sujeto /NULL = KEEP	data uni (keep=a sujeto b y); set mult; y=b1; b=1; output; y=b2; b=2; output;

Nota: la operación puede ejecutarse directamente mediante el procedimiento *Reestructurar Archivo* en el menú *Datos* en la versión 14 del programa SPSS y la versión 9.1.3 de SAS

formato multivariado (véase tabla 1). Podremos así utilizar el módulo de medidas repetidas del modelo lineal general y obtener los análisis alternativos cuando no se cumple el supuesto de esfericidad.

Spss y *Sas* suministran por defecto los valores de $\hat{\epsilon}$ y la corrección a la prueba *F* univariada según el procedimiento de Huyhn y Feldt. Para aplicar el procedimiento de Lecoutre, el cálculo de $\hat{\epsilon}$ y los grados de libertad pueden realizarse fácilmente a mano multiplicando los grados de libertad originales por $\hat{\epsilon}$. La significatividad estadística puede entonces ser obtenida mediante la función *FINV* en *Spss* o *PROBF* en *Sas* a partir de la *F* empírica.

Análisis de caso

En el contexto del estudio anterior, se plantea aquí la pregunta de si las diferencias de género en niños(as) invidentes en utilización de palabras con contenido visual depende del tipo de habilidad: narración o descripción. Ambas medidas fueron registradas para cada participante. Tenemos, por tanto, un diseño mixto de dos factores con uno de ellos entresujetos (Género) y el segundo intrasujeto (Habilidad Lingüística) al que podemos aplicar el modelo lineal en la expresión (12).

El contraste *F* de la interacción de dos factores definido en (20) no resultó estadísticamente significativo, $F(1, 58) = 1.74, p = .19$. Los análisis diagnósticos de residuales revelaron sin embargo una marcada asimetría con un valor de 2.69 (figura 3). El supuesto de normalidad fue rechazado conforme al contraste de Shapiro-Wilks, $W = .7833, p < .001$.

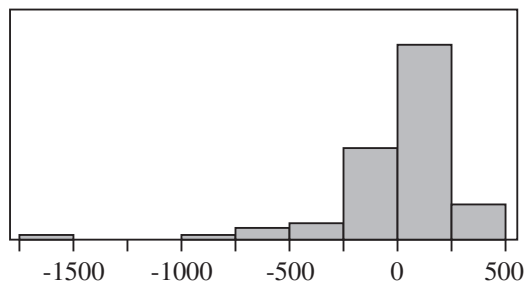


Figura 3. Distribución de residuales intrasujeto de las puntuaciones de habilidad lingüística
Nota: $S = 316.5$; Asimetría = 2.69; Curtosis = 10.57

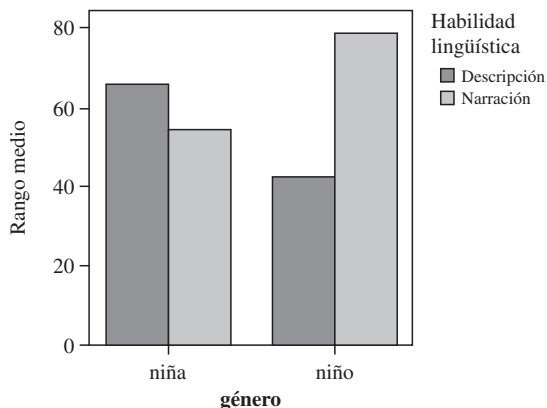


Figura 4. Rango medio en habilidad de expresión verbal en función del género y habilidad lingüística

Los resultados en la literatura revelan un aumento de la potencia estadística y niveles similares de error tipo I para la prueba de rangos alineados en comparación con el estadístico *F* en estas situaciones (Beasley, 2002). Por este motivo, se aplicó el estadístico F_{RA} en este conjunto de datos, resultando en una interacción significativa entre Género y Habilidad Lingüística, $F_{AR}(1,58) = 8.33, p < .006$. Los niños invidentes mostraron una ventaja de la narración frente a la descripción, mientras que en las niñas invidentes se observó el patrón contrario (figura 4).

Discusión

En el presente trabajo se ha revisado el contraste no paramétrico de la interacción de dos factores mediante el método de rangos alineados. Se ha aplicado a un diseño con ambos factores entre sujetos, y a otro mixto con un factor intrasujetos. El procedimiento consiste en ordenar las observaciones tras la sustracción de los efectos principales de confusión (Thompson, 1991a, 1991b). A continuación se aplica a la variable resultante una prueba *F* mediante paquetes estadísticos usuales. Estudios analíticos y de simulación han mostrado ventajas de potencia estadística y niveles similares de error Tipo I en comparación con el estadístico *F* con distribuciones asimétricas de carácter exponencial o con distribuciones de alta densidad en las colas (Beasley, 2002; Mansouri y Chang, 1995; Salter y Fawcett, 1993). El análisis de casos nos ha permitido detectar diferencias de género no observadas con otros tipos de análisis (Rosel, Caballer y Jara, 2005). Éstas pudieran deberse a mecanismos compensatorios a la discapacidad visual dependientes del género y la habilidad lingüística (Hyde y Linn, 1999; Kemter, 1999).

Límites a la utilidad del método han sido observados en situaciones de incumplimiento del supuesto de homogeneidad de covarianzas entre grupos en donde se producen sesgos por aumentos de las tasas de Error Tipo I (Kowalchuk, Keselman y Algina, 2003; Lei, Holt y Beasley, 2004). Estos sesgos no han sido subsanados cuando al análisis anterior se ha añadido procedimientos utilizados en estos casos con variables distribuidas normalmente como la corrección univariada de Huhyn o el método de Welch-James (Huyhn, 1978; Johanssen, 1980; Kowalchuk, Keselman y Algina, 2003; Lei, Holt y Beasley, 2004). El procedimiento de rangos alineados contrasta parámetros de tendencia central cuando se cumple el supuesto de identidad distribucional entre grupos para las distintas medidas (Beasley, 2002). Si esto no ocurre el rechazo de la hipótesis nula de no interacción puede deberse a otras diferencias en estructuras de covarianzas o en parámetros de forma como la asimetría (Varga y Delaney, 1998). Tamaños muestrales por grupo inferiores a diez se han mostrado también como una limitación de la prueba de rangos alineados al producir baja potencia y tasas conservadoras de Error Tipo I (Serlin y Harwell, 2004). Métodos alternativos a considerar en estos casos pueden ser las pruebas de aleatorización (Good, 2005).

El procedimiento se recomienda para el análisis de diseños entre sujetos o mixto de dos factores. Su uso en el análisis de la interacción de dos factores en diseños de orden superior debe estar apoyado por resultados que no han sido abordados en el presente trabajo. La mayor parte de los resultados de simulación revisados han utilizado tamaños muestrales iguales. Sería de interés para futuros estudios investigar las propiedades del procedimiento en diseños con muestras desiguales y comparar el valor relativo de los diferentes tipos de sumas de cuadrados (I y II) alternativos al tipo

III utilizado por defecto en *Spss* o *Sas*. Este último parece recomendable cuando no se quiere dar un peso desigual a los distintos niveles de tratamiento, ya que calcula medias marginales no ponderadas por el número de observaciones por condición (Speed, Hocking y Hackney, 1978).

El contraste de rangos alineados ha mostrado también buenos resultados en el análisis de la interacción de tres factores (Oliver y Smith, 2007). Más generalmente, un tratamiento teórico de los *modelos lineales generales robustos* (Hettmansperger y McKean, 1998; McKean y Vidmar, 1994) proporciona elementos para el desarrollo de nuevas aplicaciones de esta prueba, así como de otros métodos relacionados (Serlin y Harwell, 2004). Éstos incluyen técnicas de diagnóstico y análisis de residuales, análisis de covarianza, contrastes, comparaciones múltiples y tratamiento de diseños balanceados y no balanceados (Akritas, Arnold y Brunner, 1997).

Las propiedades de los contrastes de rangos como Mann-Whitney, Wilcoxon o el test de rangos alineados se basan en resultados

asintóticos. Se suelen recomendar a partir de muestras de tamaño intermedio. Es frecuente la creencia que con muestras grandes la prueba *F* es más adecuada debido a la tendencia asintótica hacia la normalidad de la distribución muestral de las medias. Sin embargo, recientemente se ha mostrado ventajas de las pruebas no paramétricas en situaciones realistas con tamaños muestrales superiores a cien participantes por condición (Wilcox y Keselman, 2003). Por tanto, la elección de estadístico no es tan simple, y debe estar basada en las características de los datos. Siempre es recomendable el uso de procedimientos gráficos e inferenciales para el diagnóstico de supuestos en la elección de estadísticos adecuados (Maxwell y Delaney, 2004). Un problema recurrente en los estudios del comportamiento es la baja potencia estadística (Maxwell, 2004). Una consecuencia es la generación de resultados no concluyentes en la literatura. El uso de herramientas a la medida de los datos y de estrategias de análisis de calidad puede ayudar a aportar una mayor sensibilidad para detectar fenómenos de interés teórico o aplicado.

Referencias

- Akritas, M.G., Arnold, S.F., y Brunner, E. (1997). Nonparametric hypotheses and rank statistics for unbalanced factorial designs. *Journal of the American Statistical Association*, 92(437), 258-265.
- Algina, J., y Keselman, H.J. (1997). Detecting repeated measures effects with univariate and multivariate statistics. *Psychological Methods*, 2(2), 208-218.
- Arnau, J. (1990). Diseños experimentales en Psicología y Educación (2ª ed.). México: Trillas.
- Ato, M., y Vallejo, G. (2007). *Diseños experimentales en Psicología*. Madrid: Pirámide.
- Beasley, T.M. (2002). Multivariate aligned rank test for interactions in multiple group repeated measures designs. *Multivariate Behavioral Research*, 37(2), 197-226.
- Beasley, T.M., y Zumbo, B.D. (abril, 1998). *Rank transformation and df-correction procedures for split-plot designs*. Paper presented at the meeting of the American Educational Research Association. San Diego: CA.
- Blair, R.C., Sawilowsky, S.S., y Higgins, J.J. (1987). Limitations of the rank transform statistic in test for interactions. *Communications in Statistics: Simulation and Computation*, 16(4), 1133-1145.
- Blanca, M.J. (2004). Alternativas de análisis estadístico en los diseños de medidas repetidas. *Psicothema*, 16(3), 509-518.
- Box, G.E.P. (1954). Some theorems on quadratic forms in the study of analysis of variance problems: Effects of inequality of variances in one-way classifications. *Annals of Mathematical Statistics*, 25, 290-302.
- Chen, R.S., y Dunlap, W. (1994). A monte carlo study on the performance of a corrected formula for suggested by Lecoutre. *Journal of Educational and Behavioral Statistics*, 19(2), 119-126.
- Conover, W.J., e Iman, R.L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician*, 35(3), 124-129.
- Good, P.I. (2005). *Resampling methods: A practical guide to data analysis* (3rd. ed). Boston: Birkhäuser.
- Gorham, J.L. (1998). *The effects on Type I error rate and power of selected competitors to the Anova F test in the randomized block designs under non-normality and variance heterogeneity*. Tesis doctoral no publicada. Rutgers, EEUU: Universidad Estatal de Nueva Jersey.
- Harwell, M.R., y Serlin, R.C. (1994). A Monte Carlo study of the Friedman test and some competitors in the single factor, repeated measures design with unequal covariances. *Computational Statistics and Data Analysis*, 17, 35-49.
- Hettmansperger, T.P., y McKean, W. (1998). *Robust nonparametric statistical methods*. Londres: Arnold.
- Hocking, R.R. (1996). *Methods and applications of linear models*. New York: Wiley.
- Huynh, H. (1978). Some approximate tests for repeated measures designs. *Psychometrika*, 43, 161-175.
- Huynh, H., y Feldt, L. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, 1, 69-82.
- Hyde, J.S., y Linn, M.C. (1988). Gender differences in verbal ability: A meta analysis. *Psychological Bulletin*, 104, 53-89.
- Johansen, S. (1980). The Welch-James approximation to the distribution of the residual sum of squares in a weighted linear regression. *Biometrika*, 67, 85-92.
- Kemter, P. (1999). Concept formation and spatial perception in blind persons. *Rehabilitation*, 38(1), 27-32.
- Keselman, H.J., Algina, J., y Kowalchuk, R.K. (2002). A comparison of data analysis strategies for testing omnibus effects in higher-order repeated measures designs. *Multivariate Behavioral Research*, 37(3), 331-357.
- Kirk, R.E. (1995). *Experimental design: Procedures for the behavioral sciences*. Nueva York: Brooks Cole.
- Kowalchuk, R.K., Keselman, H.J., y Algina, J. (2003). Repeated measures interaction test with aligned ranks. *Multivariate Behavioral Research*, 38(4), 433-461.
- Lecoutre, B. (1991). A correction for the e approximate test in repeated measures designs with two or more independent groups. *Journal of Educational Statistics*, 16, 371-372.
- Lei, X., Holt, J.K., y Beasley, T.M. (2004). Aligned rank tests as robust alternatives for testing interactions in multiple group repeated measures designs with heterogeneous covariances. *Journal of Modern Applied Statistical Methods*, 3(2), 462-475.
- Mansouri, H., y Chang, G.H. (1995). A comparative study of some rank tests for interaction. *Computational Statistics and Data Analysis*, 19, 85-96.
- Maxwell, S.E. (2004). The persistence of underpowered studies in psychological research: Causes, consequences and remedies. *Psychological Methods*, 9(2), 147-163.
- Maxwell, S.E., y Delaney, H.D. (2004). *Designing experiments and analyzing data* (2ª ed). Londres: Lawrence Erlbaum.
- McKean J.W., y Vidmar, T.J. (1994). A comparison of two rank based methods for the analysis of linear models. *The American Statistician*, 48(3), 220-229.
- Micceri, T. (1989). The unicorn, the normal curve and other improbable creatures. *Psychological Bulletin*, 105(1), 156-166.

- Oliver, J.C., y Smith, D.W. (2007, febrero). *Análisis no paramétrico de la interacción de tres factores en un diseño mixto ante el incumplimiento del supuesto de esfericidad*. Comunicación presentada en el X Congreso de Metodología de las Ciencias Sociales y de la Salud. Barcelona.
- Rosel (1982). *Niveles cognitivos y desarrollo lexico-sintactico. Estudio evolutivo-diferencial en escolares ciegos y videntes*. Tesis doctoral no publicada. Universidad de Salamanca.
- Rosel, J., Caballer, C., y Jara, P. (2005). Verbalism in the narrative language of children who are blind and sighted. *Journal of Visual Impairment and Blindness*, 99(7), 413-425.
- Searle, S.R., Speed, F.M., y Milliken, G.A. (1980). Population marginal means in linear models. *The American Statistician*, 34, 216-21.
- Salter, K.C., y Fawcett, R.F. (1993). The ART test of interaction: A robust and powerful rank test of interaction in factorial models. *Communications in statistics: Simulation and Computation*, 22, 137-153.
- Serlin, R.C., y Harwell, M.R. (2004). More powerful tests of predictor subsets in regression analysis under nonnormality. *Psychological Methods*, 9(4), 492-509.
- Siegel, S. (1988). *Nonparametric statistics for the behavioral sciences* (2nd. ed.). Nueva York: McGraw-Hill.
- Speed, F.M., Hocking, R.R., y Hackney, O.P. (1978). Methods of analysis of linear models with unbalanced data. *Journal of the American Statistical Association*, 73, 105-112.
- Thompson, G.L. (1991a). A note on the rank transform for interactions. *Biometrika*, 78, 697-701.
- Thompson, G.L. (1991b). A unified approach to rank tests for multivariate and repeated measures designs. *Journal of the American Statistical Association*, 86, 410-419.
- Toothaker, L.E., y Newman, D.A. (1994). Nonparametric competitors to the two-way ANOVA. *Journal of Educational and Behavioral Statistics*, 19, 237-273.
- Vallejo, G., y Lozano, L.M. (2006). Modelos de análisis para los diseños multivariados de medidas repetidas. *Psicothema*, 18(2), pp. 293-299.
- Vargha, A., y Delaney, H.D. (1998). The Kruskal-Wallis test and stochastic homogeneity. *Journal of Educational and Behavioral Statistics*, 23(2), 170-192.
- Wilcox, R.R., y Keselman, H.J. (2003). Modern robust data analysis methods: Measures of central tendency. *Psychological Methods*, 8(3), 254-274.
- Zimmerman, D., y Zumbo, B. (1993). Relative power of the Wilcoxon test, the Friedman test and the repeated measures ANOVA on ranks. *Journal of Experimental Education*, 62, 75-86.