Metodología de las Ciencias del Comportamiento
2002. Volumen Especial
ISSN 1575-9105
©2002 AEMCCO

Distribución muestral del estadístico F -f- cuando la variable dependiente es discreta y rectangular

García, J. F.; Pascual, J. y Frías, M. D.
Universidad de Valencia

En muchas ocasiones, las investigaciones psicológicas no pueden garantizar que la variable dependiente se distribuya normalmente, bien sea por la escala de medida (unidades discretas), o por la forma de la distribución (la asimetría positiva del tiempo de reacción). No obstante, la prueba de la hipótesis nula mediante el estadístico F es exacta, aunque se incumpla la normalidad (si se puede asumir la aleatorización). Pero no lo son las tablas que habitualmente se emplean, puesto que se calcularon suponiendo que la variable dependiente era normal. Como también se sabe, este problema es relevante si la muestra es pequeña, porque si sobrepasa cierto tamaño, no tiene ninguna repercusión práctica. En este trabajo se propone un procedimiento alternativo que proporciona valores exactos: (1) transformar la variable dependiente en otra discreta y de distribución rectangular, y (2) calcular la función de distribución muestral del estadístico F -al que denominamos *f* por no asumir la normalidad- con los grados de libertad del experimento

F-sample distribution *-f-* when the measurement scale is discrete and rectangular. Frequently, psychological studies cannot guarantee normal distribution of dependent variables, either by the measurement scale (discrete units) or by the shape of the distribution (the positive skewness of the reaction time). Nevertheless, the test of the null hypothesis with the F-statistic is exact, although the normality of the distribution is violated (assuming randomization). However the tables used are not exact because they were calculated assuming the normality of the dependent variable. Also, this problem is relevant when the sample is small, because when certain sample size is exceeded has no practical implications. This paper proposes an alternative procedure which gives exact values: (1) transforming the dependent variable in a discrete variable with a rectangular distribution and, (2) to calculate the sample distribution function of the F statistic -named *f* since normality is not assumed- with the "df" of the experiment.

La distribución muestral más empleada en la investigación es la del estadístico *F*, tabulada por Snedecor en 1934; estas funciones de distribución se calcularon suponiendo que la variable de la que procedían las muestras -la población padre- era normal. Si los datos de un experimento proceden de una variable distribuida normalmente, los valores de Snedecor serán completamente exactos, pero si la distribución de la población padre tuviera otra forma, la probabilidad del error de Tipo I será imprecisa.

En las ciencias del comportamiento es factible asumir la distribución normal de bastantes variables dependientes (Maxwell y Delaney, 1989, pp: 51-55), pero también es frecuente que las variables dependientes se desvíen de la forma normal. Una de las desviaciones mejor estudiadas por la literatura es la asimetría, por ejemplo, la distribución del tiempo de reacción suele ser asimétrica positiva porque existe un número reducido de personas que emplean un tiempo excesivo en completar la tarea experimental. Esta circunstancia suele atribuirse al efecto de algunas variables extrañas como la distracción, errores previos, interpretaciones incorrectas de las instrucciones, desmotivación, etc. (Luce, 1986). Esta distribución asimétrica negativa se representa mediante una curva ex-gausiana (Heathcote, 1996; Luce, 1986; Ratclifs, 1979; Ratclifs y Murdock, 1976) que se obtiene sumando a una distribución normal otra exponencial que parametriza la cola derecha. En otras ocasiones, la distribución del tiempo de reacción se ajusta mejor a la forma log-normal (Burns y Anderson, 1993; Richards y Cronise, 2000). Incluso, es posible encontrar casos de distribuciones asimétricas negativas en medidas del tiempo de anticipación-reacción cuando se estudia el desarrollo de la expectativa visual (Reznick, Chawarska y Betts, 2000; p. 1195). Por otra parte, una asimetría negativa muy pronunciada suele aparecer cuando se miden variables como el autoconcepto familiar (García y Musitu, 1999), donde la mayoría de las personas suelen situarse en las posiciones más altas de la escala.

La asimetría puede aumentar el error de Tipo II si las observaciones extremas incrementan considerablemente el error típico de estimación. De tal forma, que si la prueba de la hipótesis se realiza eliminando las puntuaciones de la cola, suele aumentar el valor empírico de *F*. Entre las alternativas propuestas destacan: (1), la eliminación de los datos extremos mediante criterios estadísticos o teóricos (Candfield, Smith, Brezsnyak, y Snow, 1997; Ratclifs, 1993; Ulrich y Miller, 1994); (2), la transformación logarítmica o inversa de la variable dependiente (Ratclifs, 1993; Ulrich y Miller, 1994); (3), aplicar pruebas de aleatorización (Blair y Karniski, 1993); y (4), analizar los datos asumiendo que proceden de una distribución ex-gausiana con parámetros: σ, μ. y τ (Heathcote, 1996; Ratclifs, 1979). En general los recortes de los datos, sobre todo, si exceden al 5% de la distribución (Ulrich y Miller, 1994) afectarán a la precisión de la prueba de la hipótesis, de la misma manera que si se emplea la mediana u otros estadísticos diferentes de la media (Miller 1988, 1991). Las transformaciones de la variable dependiente implica que los estadísticos estén expresados en unidades de medida que no son los habituales milisegundos (Stenberg, 1969). Las pruebas de aleatorización conllevan realizar un importante número de cálculos para cada experimento, y, por otra parte, tampoco se ha comprobado que estas pruebas estuvieran libres de cumplir con el supuesto de la independencia (véase Hayes, 1996). Por último, los análisis a partir de la distribución ex-gausiana implican la estimación de un parámetro más, τ , que en ocasiones suele estar relacionado con el valor de la μ, complicando enormemente la

Correspondencia: José Femando García Pérez
Área de Metodología de las Ciencias del Comportamiento. Facultad de Psicología.
Universidad de Valencia
Avda. Blasco Ibánez 21. 46010 Valencia.
E-mail: fernando.garcia@uv.es
VII Congreso de Metodología de las Ciencias Sociales y de la Salud, Madrid, 25-28 de septiembre, 2001.

interpretación de los resultados de la prueba de la hipótesis (véase Spieler, Balota y Faust, 2000).

En este trabajo se propone emplear una solución alternativa para el análisis de la varianza que consiste en: (1) transformar las puntuaciones de la variable dependiente en puntuaciones *r*tiles, (2) calcular el ANOVA con las nuevas puntuaciones *r*tiles. y (3) emplear la distribución muestral de *f* calculado asumiendo que la distribución padre tiene forma rectangular -con parámetros υ_1, υ_2, *R* y α- para realizar la prueba de la hipótesis.

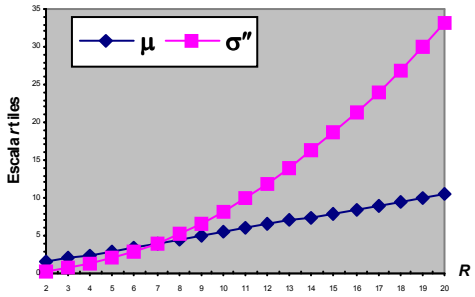
Transformación a la escala *R*TIL

La transformación *r*til se puede aplicar tanto si la asimetría es positiva como si es negativa, manteniendo el orden original de la escala pero no la proporción entre las unidades, se trata, por tanto, de una transformación que no es lineal. Cuanta mayor sea la precisión en la escala original, mayor será el parámetro *R*. A la hora de justificar el tamaño de *R* habrá que controlar, como mínimo, que no existan valores idénticos en la escala original a los que corresponda distinta puntuación *r*til. Otro criterio más riguroso para transformar los datos es determinar el error típico de estimación, para garantizar que las bandas *r*tiles no se sobrepongan (Martínez, 1995, p. 627).

Una ventaja muy importante de este procedimiento es que se puede conocer con exactitud la media y desviación típica de la variable dependiente medida con la escala *r*til. Desde otra perspectiva, se puede decir que la media (μ) y la varianza (σ^2) de la población padre estarán determinadas, únicamente, por el tamaño de *R* (figura 1), de la siguiente manera:

$$\mu = \frac{R+1}{2}; \quad \sigma^2 = \frac{(R-1)(R+1)}{12}$$

Figura 1. Media (μ) y varianza (σ^2) de la población padre según varíe R.



Definiendo formalmente la población padre, puede decirse que será una variedad de la distribución rectangular (uniforme) discreta *R*, con un recorrido entre los valores 1, el mínimo, y *R*, el máximo, de tal manera que $1 \leq r \leq R$, siendo *r* cualquier valor entero 1, 2, ..., *R*. La función de probabilidad en cualquier punto *r* será $1/R$, y la de distribución, r/R.

Un ejemplo muy sencillo es la variedad $R = 3$, la población padre la compondrían los puntos de la escala 1, 2 y 3:

Punto (*r*)	Proporción (1/*R*)	Proporción acumulada (*r*/*R*)
1	1/3	1/3
2	1/3	2/3
3	1/3	3/3

Calculamos la media y varianza (no hace falta ponderar los valores de *r* porque todos son equiprobables):

$$\mu = \frac{\Sigma_{i=1}^{R} r_i}{R} = \frac{1+2+3}{3} = \frac{6}{3} = 2$$

$$\sigma^2 = \frac{\Sigma_{i=1}^{R} (r_i - \mu)^2}{R} = \frac{(1-2)^2 + (1-2)^2 + (1-3)^2}{3} = \frac{2}{3} = 0{,}667$$

Lo mismo se obtiene, pero de manera más sencilla, a partir de *R*:

$$\mu = \frac{R+1}{2} = \frac{3+1}{2} = 2$$

$$\sigma^2 = \frac{(R-1)(R+1)}{12} = \frac{(3-1)(3+1)}{12} = \frac{8}{12} = 0{,}667$$

Distribución muestral de la media

A la hora de aplicar el ANOVA, el interés se centra en la distribución de la muestra, pero antes calcularemos la función de distribución de la media cuando se extraen dos elementos ($n = 2$) de una población padre donde $R = 3$. Este cálculo, más sencillo, ayudará a explicar el siguiente paso.

Al extraer dos elementos de la población padre R_3 ($R = 3$) el primer valor será uno de los tres que componen la población y el segundo, otro de ellos. Al ser equiprobables e independientes, la población muestral se compondrá de los siguientes pares de elementos ($n = 2$), y se obtendrán, necesariamente, las siguientes medias:

	1	**2**	**3**
1	(1 + 1) / 2 = **1,0**	(1 + 2) / 2 = **1,5**	(1 + 3) / 2 = **2,0**
2	(3 + 1) / 2 = **1,5**	(2 + 2) / 2 = **2,0**	(2 + 3) / 2 = **2,5**
3	(3 + 1) / 2 = **2,0**	(3 + 2) / 2 = **2,5**	(3 + 3) / 2 = **3,0**

La distribución muestral de la media será, por tanto:

Punto ($\bar{r}$)	Frecuencia (*n*)	Proporción	$\bar{r}n$	$\bar{r} - \mu_{\bar{r}}$	$(\bar{r} - \mu_{\bar{r}})^2$	$(\bar{r} - \mu_{\bar{r}})^2 n$
1,0	1	1/9	1	-1,0	1,00	1,00
1,5	2	2/9	3	-0,5	0,25	0,50
2,0	3	3/9	6	0,0	0,00	0,00
2,5	2	2/9	5	0,5	0,25	0,50
3,0	1	1/9	3	1,0	1,00	1,00
Σ	9	1	18			3,00

Y el promedio ($\mu_{\bar{r}}$) y la varianza ($\sigma^2_{\bar{r}}$) de la población muestral de la media:

$$\mu_{\bar{r}} = \frac{\Sigma_{i=1}^{k} \bar{r}_j n_j}{\Sigma_{i=1}^{k} n_j} = \frac{18}{9} = 2; \quad \sigma^2_{\bar{r}} = \frac{\Sigma_{i=1}^{k} (\bar{r}_j - \mu_{\bar{r}})^2 n_j}{\Sigma_{i=1}^{k} n_j} = \frac{3}{9} = 0{,}333$$

Este supuesto de independencia implica que la probabilidad de obtener cualquiera de los elementos de la población es independiente y constante, no existiendo ninguna probabilidad condicional entre ellos. Para cumplir el supuesto es preciso que se produzca la asignación aleatoria, asimismo, para que el efecto de la extracción no altere la función de probabilidad de la distribución padre, también es necesario que el número de elementos de la misma sea infinito. Las variables dependientes más utilizadas en psicología -como pudiera ser el tiempo de reacción ante las diferentes condiciones experimentales- suelen ser medidas de un fenómeno inagotable (Fisher, 1925/1973: 1-3), pudiéndose asumir que son poblaciones con un número ilimitado de elementos (Fritz, 1998; Pedhazur y Schmelkin, 1991: 322 y ss.).

Teorema del límite central

Aplicando el teorema del límite central (Por ejemplo, Padilla, Merino, Rodríguez-Miñón y Moreno, 1996: 394-399) se puede abreviar considerablemente el cálculo, conociendo la media (μ) y desviación típica de la población padre (σ^2), y el tamaño de las muestras (*n*), es posible determinar directamente el promedio ($\mu_{\bar{r}}$) y varianza ($\sigma^2_{\bar{r}}$) de la distribución muestral de la media.

$$\mu_{\bar{r}} = \mu; \quad \sigma^2_{\bar{r}} = \frac{\sigma^2}{n}$$

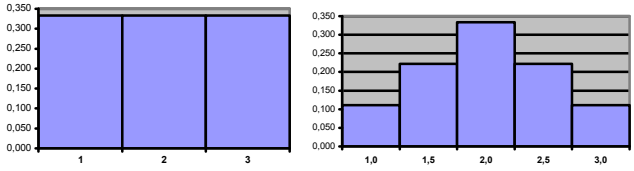
Por lo tanto,

$$\mu_{\bar{r}} = \mu = 2$$

$$\sigma^2_{\bar{r}} = \frac{\sigma^2}{n} = \frac{2}{3} : 2 = \frac{2}{3}\,\frac{1}{2} = \frac{2}{6} = 0{,}333$$

Hay otra particularidad que se aprecia en la figura 2, mientras que la función de probabilidad de la población padre es exactamente rectangular, la de las medias de las muestras se aproxima a la curva normal. Esta convergencia se incrementa conforme aumentan *R* y *n* (véanse Bock, 1975, p. 111; García, Frías y Pascual, 2000; García, Frías y Pascual, 2001, Cap. IV).

Figura 2. Función de probabilidad de la distribución padre y de la población de medias de dos observaciones extraídas aleatoriamente



Distribución muestral del estadístico *f*

Como es posible conocer las muestras que se pueden extraer de una población padre distribuida rectangularmente, se puede calcular la distribución muestral del estadístico *f* con exactitud. La distribución muestral de *f* se puede emplear para realizar la prueba de la hipótesis porque la razón de varianzas es exacta bajo el supuesto de independencia, no siendo precisa la normalidad.

Siguiendo con el ejemplo de $R = 3$ y $n = 2$, si se aplica un ANO VA entre dos condiciones experimentales ($a = 2$), en cada condición aparecerá, necesariamente, una de las 9 muestras que se calcularon anteriormente ($R^n = 3^2 = 9$). La población muestral de SC_{entre} / SC_{error}, será la siguiente:

r_1		→		1	1	2	2	3	1	2	3	3
	r_2	→		1	2	1	2	1	3	3	2	3
		Σ		2	3	3	4	4	4	5	5	6
↓	↓		Σ^2	2	5	5	8	10	10	13	13	18
1	1	2	2	0,00/0,00	0,25/0,50	0,25/0,50	1,00/0,00	1,00/2,00	1,00/2,00	2,25/0,50	2,25/0,50	4,00/0,00
1	2	3	5	0,25/0,50	0,00/1,00	0,00/1,00	0,25/0,50	0,25/2,50	0,25/2,50	1,00/1,00	1,00/1,00	2,25/0,50
2	1	3	5	0,25/0,50	0,00/1,00	0,00/1,00	0,25/0,50	0,25/2,50	0,25/2,50	1,00/1,00	1,00/1,00	2,25/0,50
2	2	4	8	1,00/0,00	0,25/0,50	0,25/0,50	0,00/0,00	0,00/2,00	0,00/2,00	0,25/0,50	0,25/0,50	1,00/0,00
3	1	4	10	1,00/2,00	0,25/2,50	0,25/2,50	0,00/2,00	0,00/4,00	0,00/4,00	0,25/2,50	0,25/2,50	1,00/2,00
1	3	4	10	1,00/2,00	0,25/2,50	0,25/2,50	0,00/2,00	0,00/4,00	0,00/4,00	0,25/2,50	0,25/2,50	1,00/2,00
2	3	5	13	2,25/0,50	1,00/1,00	1,00/1,00	0,25/0,50	0,25/2,50	0,25/2,50	0,00/1,00	0,00/1,00	0,25/0,50
3	2	5	13	2,25/0,50	1,00/1,00	1,00/1,00	0,25/0,50	0,25/2,50	0,25/2,50	0,00/1,00	0,00/1,00	0,25/0,50
3	3	6	18	4,00/0,00	2,25/0,50	2,25/0,50	1,00/0,00	1,00/2,00	1,00/2,00	0,25/0,50	0,25/0,50	0,00/0,00

Agrupando por frecuencias las SC_{entre}:

SC_{entre}	Frecuencia (*n*)	Proporción	$SC_{entre}\,n$
0,00	19	0,2346	0
0,25	32	0,3951	8
1,00	20	0,2469	20
2,25	8	0,0988	18
4,00	2	0,0247	8
Σ	81	1	54

Y el promedio ($\mu_{SCentre}$) de las sumas de cuadrados entre en la población muestral será:

$$\mu_{SCentre} = E[SC_{entre}] = \frac{\Sigma^k_{i=1} SC_{entre\,j}\, n_j}{\Sigma^k_{i=1} n_j} = \frac{54}{81} = 0{,}667$$

Contando las frecuencias de cada valor de la SC_{error}:

SC_{error}	Frecuencia (*n*)	Proporción	$SC_{error}\,n$
0,00	9	0,1111	0
0,50	24	0,2963	12
1,00	16	0,1975	16
2,00	12	0,1481	24
2,50	16	0,1975	40
4,00	4	0,0494	16
Σ	81	1	108

El promedio ($\mu_{SCerror}$) de las sumas de cuadrados residual en la población muestral será:

$$\mu_{SCerror} = E[SC_{error}] = \frac{\Sigma^k_{i=1} SC_{error\,j}\, n_j}{\Sigma^k_{i=1} n_j} = \frac{108}{81} = 1{,}333$$

En un modelo elemental con un factor *A* de *a* condiciones y las mismas *n* observaciones por condición de *a*. Si las $n \cdot a$ observaciones son extraídas independientemente de una misma población de media μ y varianza σ^2 se sabe que (Hays, 1981: 625 y ss.):

$$E[MC_{entre}] = E[MC_{error}] = \sigma^2$$

Como en el término de error hay dos grados de libertad ($gl_{error} = N - a = 4 - 2 = 2$) y en el factor *A* uno ($gl_{entre} = a - 1 = 2 - 1 = 1$), podemos comprobar que las dos medias cuadráticas son dos *estimaciones insesgadas* de la varianza de la población padre:

$$E[MC_{entre}] = E\left[\frac{SC_{entre}}{gl_{entre}}\right] = \frac{1}{gl_{entre}} E[SC_{entre}] = 1 \cdot 0{,}667 = 0{,}667$$

$$E[MC_{error}] = E\left[\frac{SC_{error}}{gl_{error}}\right] = \frac{1}{gl_{error}} E[SC_{error}] = ½ \cdot 1{,}333 = 0{,}667$$

Se calcula el valor de *f* en las 81 muestras que componen la población muestral

r_1		→		1	1	2	2	3	1	2	3	3
	r_2	→		1	2	1	2	1	3	3	2	3
		Σ		2	3	3	4	4	4	5	5	6
↓	↓		Σ^2	2	5	5	8	10	10	13	13	18
1	1	2	2	###	1,00	1,00	###	1,00	1,00	9,00	9,00	###
1	2	3	5	1,00	0,00	0,00	1,00	0,20	0,20	2,00	2,00	9,00
2	1	3	5	1,00	0,00	0,00	1,00	0,20	0,20	2,00	2,00	9,00
2	2	4	8	###	1,00	1,00	###	0,00	0,00	1,00	1,00	###
3	1	4	10	1,00	0,20	0,20	0,00	0,00	0,00	0,20	0,20	1,00
1	3	4	10	1,00	0,20	0,20	0,00	0,00	0,00	0,20	0,20	1,00
2	3	5	13	9,00	2,00	2,00	1,00	0,20	0,20	0,00	0,00	1,00
3	2	5	13	9,00	2,00	2,00	1,00	0,20	0,20	0,00	0,00	1,00
3	3	6	18	###	9,00	9,00	###	1,00	1,00	1,00	1,00	###

Hay 9 muestras en las que no se puede obtener *f* porque la suma de cuadrados del error es nula, por lo tanto, el tamaño muestral de la población se reduce a (81 - 9 = 72) 72 muestras en las que realmente se puede calcular el estadístico. La función de distribución muestral será:

$f_{1,2}$	Frecuencia (*n*)	Proporción	Proporción acumulada	α
*	*9*			
0,00	16	0,2222	0,222	1,000
0,20	16	0,2222	0,444	0,778
1,00	24	0,3333	0,778	0,556
2,00	8	0,1111	0,889	0,222
9,00	8	0,1111	1,000	0,111
Σ	72	1		

La distribución muestral de *f* calculada con la población padre de *R*3 únicamente proporciona 5 valores distintos, el último incluye el 11,11% de la distribución. Por lo tanto, a la hora de aplicar esta distribución en la investigación práctica únicamente se podría fijar un α de 0,11. Pero este ejemplo es uno de los más sencillos que pueden plantearse y se ha elegido para que sea posible exponer todos los cálculos con la máxima brevedad. Y, también, para poder ilustrar que las distribuciones muéstrales obtenidas no se han obtenido mediante una simulación parcial sino que se ha calculado la distribución muestral completa: la población muestral.

Puntos de *f* para el 5,0% de error del tipo I con dos muestras

Repitiendo el mismo proceso se han calculado los puntos percentiles

del 95,0% de las distribuciones de *f*, a partir de poblaciones rectangulares de 2 a 10 elementos para varios grados de libertad en el término de error. Los resultados se presentan en la Tabla 1.

La distribución muestral de *f* se ha calculado suponiendo que se extraen aleatoriamente y con reposición dos muestras iguales de una población rectangular de entre 2 y 10 elementos. Bajo la hipótesis nula, los puntos de corte para el percentil 95 son los tabulados. En todos los casos se indica la proporción acumulada superior de cada valor *f* porque, en algunos casos, como la distribución es discreta, el valor percentil acumula más del 95,0% de la distribución.

Los valores del inicio de la tabla que se han marcado con el signo "-" no se proporcionan porque son inoperantes para la investigación aplicada. En estos casos, el recorrido del valor al que corresponde el percentil 95 es una banda muy amplia que abarca al 92,5.

El análisis de las distribuciones permite observar dos principios generales: (1) Conforme aumentan los grados de libertad del denominador el valor de *f* se aproxima al de *F*, de hecho a partir de 30 grados de libertad las discrepancias son inferiores a dos centésimas. (2) Al aumentar los elementos de la población padre se produce una estabilización del valor de *f*.

Por otra parte, también hay que señalar que en la distribución muestral de una población padre de dos elementos ($R = 2$) es difícil fijar exactamente los puntos percentiles por el efecto de sierra, o escalera, que produce la escasez de valores. Es muy difícil precisar un valor puntual del error de Tipo I para esta situación límite.

Por el contrario, conforme aumentan n y R se diversifica cada una de las 2 muestras en progresión geométrica. Por ejemplo, para calcular el percentil 95 en la distribución de $f_{1, 14, 9}$ -cuyo valor es 4,68- hay que generar para cada condición experimental Rn muestras de 8 elementos, por tanto: 98 = 43.046.721. A la hora de aplicar el ANOVA es preciso cruzar todas ellas, lo que implica elevar al cuadrado esta cantidad: 43.046.7212 = 1.853.020.188.851.800, y calcular el valor de *f* para todas estas combinaciones. Afortunadamente, en estos límites superiores el valor acumulado de *f* converge con el de *F* (Obviamente, para poder realizar los cálculos se agruparon las muestras por frecuencias).

Las indicaciones para el empleo de la tabla son:

1. Con 30 ó más grados de libertad en el término de error se puede emplear la distribución muestral de F, si no se proporciona *f*.

2. Cuando $R > 10$ y haya 12 grados de libertad en el error, o menos, se pueden emplear los puntos de $R = 10$, aunque pudiera perderse algo de potencia.

3. Entre 12 y 20 grados de libertad en el término del error, si no se proporciona en la tabla el valor de *f*, debe tomarse el último de la misma fila.

Conclusiones

Los resultados obtenidos, aunque limitados a dos grados de libertad en el numerador y al 5% del error de Tipo I, sí parecen confirmar la validez estadística del procedimiento propuesto si se garantiza la independencia de las observaciones. Se pueden analizar los datos sin necesidad de tener que eliminar observaciones extremas y podiendo aplicar la prueba de la hipótesis conociendo la función de distribución exacta. Independientemente de que la distribución de la variable dependiente sea asimétrica positiva o negativa, tenga forma log-normal, ex-gausiana o exponencial, siempre se puede realizar la transformación en puntuaciones rtiles. Aspecto que actualmente no supone ningún problema técnico añadido, puesto que es una función incorporada en la mayoría de los programas informáticos. No obstante, por la convergencia constatada con la distribución *F*, a partir de 30 grados de libertad en el término de error se podría realizar la transformación *r*til y aplicar la prueba de la hipótesis tomando los valores críticos de *F*.

Considerando por la otra parte el error de Tipo II, también se puede mejorar la precisión del cálculo de la potencia substantivamente, porque se puede determinar a priori, sin necesidad de ninguna suposición difícil de justificar, la media y varianza de la población padre. El valor de *R* es suficiente para determinar estos parámetros, y el tamaño de *R* dependerá, en última instancia, del error típico de la variable dependiente.

Por último, este trabajo pionero indica que es posible terminar la distribución muestral del estadístico *f* con la técnica empleada en este trabajo. No obstante, para que pueda incluirse en los programas de cálculo estadístico será necesario derivar la distribución muestral aplicando el cálculo diferencial.

Referencias

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology, 30,* 518-524.

Bock, R. D. (1975). *Multivariate statistical methods in behavioral research.* New York: McGraw-Hill.

Burns, J. J., & Anderson, D. R. (1993). Attentional inertia and recognition memory in adult televisión viewing. *Communication Research, 20,* 777-799.

Candfield, R. L., Smith, E. G., Brezsnyak, M. P., & Snow, K. L. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Societyfor Research in Child Development*, *62* (2, Serial n° 250).

Fisher, R. A. (1925/1973). *Statistical methods for research workers.* (14th ed. rev.). New York, NY: Hafher Publishing Company.

Fritz, R. W. (1998). Interpreting statistical testing: process and propensity, not population and random sampling. *Behavior Research, Methods, Instruments, and Computers*, *30,* 527-535.

García, J. F., & Musitu, G. (1999). AF5: *Autoconcepto Forma 5*. Madrid: TEA ediciones.

García, J. F., Frías, M. D., & Pascual, J. (2000). Prueba de aleatorización vs. distribución F cuando la escala de medida de la variable dependiente es discreta y el diseño experimental. *Psicothema, 12(2)*, 253-256.

García, J. F, Frías, M. D., & Pascual, J. (2001). *Los diseños de la investigación experimental. Comprobación de las hipótesis* (2ª Ed.). Valencia: Cristóbal Serrano Villalba.

Hayes, A. F. (1996). Permutation test is not distribution-free: testing Ho: p = 0. *Psychological Methods, 1,* 184-189.

Hays, W. L. (1981). *Statistics* (3rd ed.). New York: Holt, Rinehart & Winston.

Heathcote, A. (1996). RTSYS: A DOS application for the analysis of reaction time data. *Berhavior Research Methods, Instruments, and Computers, 28,* 427-445.

Luce, R. D. (1986). *Response times.* New York: Oxford University Press.

Martínez, R. (1995). *Psicometría: teoría de los tests psicológicos y educativos.* Madrid: Síntesis Psicología.

Maxwell, S. E., & Delaney, H. D. (1989). *Designing experiments and analysis data. A model comparison perspective.* Beltmont, CA: Wadsworth Publishing Company.

Miller, J. (1988). A warning about median reaction time. *Journal of Experimental Psychology: Human Perception and Perfomance, 14,* 539-543.

Miller, J. (1991). Reaction time analysis with outlier exclusión: Bias varíes with sample size. *Quaterly Journal of Experimental Psychology, 43A,* 907-912.

Padilla, M., Merino, J. M., Rodríguez-Miñón, P., & Moreno, E. (1996).

Psicología matemática I. Madrid: UNED.

Pedhazur, E. J., & Schmelkin, L. (1991). *Measurement, design and analysis.* Hillsdale, NJ: LEA.

Ratclifs, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletín, 86,* 446-461.

Ratclifs, R. (1993). Methods for dealing with reaction times outliers. *Psychological Bulletin. 114,* 510-532.

Ratclifs, R., & Murdock, B. B, (1976). Retrieval processes in recognition memory. *Psychological Review, 83,* 190-214.

Reznick, J. S., Chawarska, K., & Betts, S. (2000). The development of visual expectations in the first-year. *Child Development, 71,* 1191-120.

Richards, J. E, & Cronise, K. (2000). Extended visual fixation in the early preschool years: Look duration, heart-rate changes, and attentional inertia. *Child Development. 71,* 602-620.

Snedecor, G. W. (1934). *Calculation and interpretation of analysis of varíance and covariance.* Monograph n. 1, lowa State College, División of Industrial Science. Ames, IW: Collegiate Press.

Spieler, D. H, Balota, D. A, & Faust, M. E. (2000). Levels of selective attention revealed through analyses of response-time distributions. *Journal of Experimental Psychology: Human Perception And Performance, 26,*506-526

Stenberg, S. (1969). The discovery of processing slags: Extensions of Donder's method. En W. G. Koster, ed., *Attention and performance,* 11, (pp. 267-315). Amsterdang: North-Holland.

Ulrich, R., & Miller, J. (1994). Effects of truncation on reaction time analysis. *Journal of Experimental Psychology: General, 123*, 34-80.

Tabla I. Percentiles de f para el 95,0 % con 1 grado de libertad en el numerador

$gl_W \backslash R$	2 Percentil	3 Percentil	4 Percentil	5 Percentil	6 Percentil	7 Percentil	8 Percentil	9 Percentil	10 Percentil	*F*
2	—	—	—	25,00 ^{0,987}	18,00 ^{0,956}	18,00 ^{0,952}	24,20 ^{0,952}	24,20 ^{0,952}	24,20 ^{0,951}	18,51
4	—	16,00 ^{0,983}	8,00 ^{0,956}	8,00 ^{0,951}	9,14 ^{0,951}	9,14 ^{0,950}	9,14 ^{0,950}	9,14 ^{0,950}	9,14 ^{0,950}	7,71
6	—	6,82 ^{0,955}	6,82 ^{0,954}	6,40 ^{0,952}	6,40 ^{0,950}	6,52 ^{0,950}	6,42 ^{0,950}	6,42 ^{0,950}	6,42 ^{0,950}	5,99
8	16,00 ^{1,000}	5,56 ^{0,955}	5,56 ^{0,951}	5,56 ^{0,950}	5,56 ^{0,950}	5,57 ^{0,950}	5,57 ^{0,950}	5,57 ^{0,950}	5,56 ^{0,950}	5,32
10	8,00 ^{0,979}	5,00 ^{0,953}	5,00 ^{0,950}	5,10 ^{0,950}	5,10 ^{0,950}	5,12 ^{0,950}	5,14 ^{0,950}	5,13 ^{0,950}	5,12 ^{0,950}	4,96
12	6,00 ^{0,979}	4,74 ^{0,950}	4,92 ^{0,951}	4,84 ^{0,950}	4,85 ^{0,950}	4,85 ^{0,950}	4,86 ^{0,950}	4,86 ^{0,950}	4,86 ^{0,950}	4,75
14	5,09 ^{0,974}	4,57 ^{0,950}	4,76 ^{0,951}	4,67 ^{0,950}	4,67 ^{0,950}	4,68 ^{0,950}	4,68 ^{0,950}	4,68 ^{0,950}	→	4,60
16	4,57 ^{0,967}	4,45 ^{0,950}	4,56 ^{0,950}	4,56 ^{0,950}	4,56 ^{0,950}	4,56 ^{0,950}	4,56 ^{0,950}	→		4,49
18	4,24 ^{0,958}	4,37 ^{0,950}	4,44 ^{0,950}	4,46 ^{0,950}	4,46 ^{0,950}	4,46 ^{0,950}	→			4,41
20	5,21 ^{0,960}	4,39 ^{0,950}	4,35 ^{0,950}	4,38 ^{0,950}	4,39 ^{0,950}	4,39 ^{0,950}	→			4,35
22	4,66 ^{0,962}	4,40 ^{0,951}	4,33 ^{0,950}	4,33 ^{0,950}	4,33 ^{0,950}	→				4,30
24	4,29 ^{0,962}	4,36 ^{0,951}	4,30 ^{0,950}	4,29 ^{0,950}	4,29 ^{0,950}	→				4,26
26	4,01 ^{0,959}	4,23 ^{0,950}	4,24 ^{0,950}	4,25 ^{0,950}	4,25 ^{0,950}	→				4,23
28	3,80 ^{0,954}	4,20 ^{0,951}	4,20 ^{0,950}	4,22 ^{0,950}	→					4,20
30	4,12 ^{0,950}	4,18 ^{0,951}	4,17 ^{0,950}	4,18 ^{0,950}						← 4,17
32	4,57 ^{0,960}	4,15 ^{0,950}	4,15 ^{0,950}	4,15 ^{0,950}						← 4,15
34	4,31 ^{0,961}	4,14 ^{0,950}	4,15 ^{0,950}	4,14 ^{0,950}						← 4,13
36	4,10 ^{0,960}	4,12 ^{0,950}	4,12 ^{0,950}	4,12 ^{0,950}						← 4,11
38	3,93 ^{0,957}	4,11 ^{0,951}	4,10 ^{0,950}	4,10 ^{0,950}						← 4,10
40	3,79 ^{0,954}	4,10 ^{0,951}	4,07 ^{0,950}	4,08 ^{0,950}						← 4,08
42	3,67 ^{0,950}	4,09 ^{0,951}	4,07 ^{0,950}							← 4,07
44	4,49 ^{0,953}	4,08 ^{0,950}	4,08 ^{0,950}							← 4,06
46	4,29 ^{0,954}	4,07 ^{0,950}	4,05 ^{0,950}							← 4,05
48	4,11 ^{0,955}	4,07 ^{0,950}	4,04 ^{0,950}							← 4,04
50	3,96 ^{0,955}	4,06 ^{0,950}	4,03 ^{0,950}							← 4,03
52	3,84 ^{0,954}	4,03 ^{0,950}								← 4,03
54	3,73 ^{0,952}	3,99 ^{0,950}								← 4,02
56	3,63 ^{0,950}	3,98 ^{0,950}								← 4,01
58	4,44 ^{0,953}	3,98 ^{0,950}								← 4,01
60	4,29 ^{0,954}	3,99 ^{0,950}								← 4,00
62	4,15 ^{0,955}									← 4,00
64	4,03 ^{0,956}									← 3,99
66	3,93 ^{0,955}									← 3,99
68	3,83 ^{0,954}									← 3,98
70	3,75 ^{0,953}									← 3,98
72	3,67 ^{0,951}									← 3,97
74	4,00 ^{0,950}									← 3,97
76	4,28 ^{0,951}									← 3,97
78	4,16 ^{0,952}									← 3,96
80	4,06 ^{0,953}									← 3,96
82	3,97 ^{0,953}									← 3,96