

ARTÍCULO METODOLÓGICO

Programas informáticos para comparaciones entre correlaciones: muestras independientes

Computer Programs for the Comparison between Correlations: Independent Samples
Programas de informática para comparações entre correlações: amostras independentes

N. Clayton Silver*

Department of Psychology, University of Nevada, Las Vegas, USA

César Merino-Soto**

Instituto de Investigación de Psicología, Universidad de San Martín de Porres, Lima, Perú

Recibido: 11/03/16

Aceptado: 23/05/16

RESUMEN. Existe una variedad de técnicas para probar las diferencias entre correlaciones independientes que no están disponibles en los programas estadísticos familiares para el investigador. Se presentan ejemplos de estas técnicas para evaluar diferentes hipótesis dentro del contexto de correlaciones en muestras independientes, junto con programas informáticos interactivos, amigables y libre distribución.

Palabras clave:
correlaciones,
estadística,
metodología, muestras
independientes,
software.

ABSTRACT. There are a variety of techniques for testing the differences among independent correlations and between dependent correlations, which are not available using the standard statistical software packages. Examples of these techniques for examining different hypotheses within the independent and dependent correlational realms are presented along with the output and interpretation from easily attainable, user-friendly, interactive software.

Keywords:
correlations,
statistics,
methodology,
independents
samples,
software.

RESUMO. Existe uma variedade de técnicas para provar as diferenças entre correlações independentes que não estão disponíveis nos programas estatísticos familiares para o investigador. Apresentam-se exemplos destas técnicas para avaliar diferentes hipóteses dentro do contexto de correlações em amostras independentes, junto com programas de informáticas interativos, amigáveis e livre distribuição.

Palavras-chave:
correlações, estatística,
metodologia, amostras
independentes,
software.

Citar como: Silver, N. C. & Merino-Soto, C. (2016). Programas informáticos para comparaciones entre correlaciones: Muestras independientes. *Revista Digital de Investigación en Docencia Universitaria*, 10(1), 28-33. doi: <http://dx.doi.org/10.19083/ridu.2016.461>

*E-mail: fdnsilvr@unlv.nevada.edu

**E-mail: cmerinos@usmp.pe, sikayax@yahoo.com.ar

En ocasiones, los investigadores están interesados en determinar si dos o más correlaciones son diferentes. Para situaciones como esta, hay varias técnicas que evalúan hipótesis de diferencias de correlaciones; sin embargo, se puede verificar que habitualmente no se implementan en los artículos publicados en habla hispana. Hay algunos motivos que pueden explicar esta limitación. Primero, estas técnicas son infrecuentemente abordadas en los libros de estadística; efectivamente, una revisión parcial de los libros que circulan en Hispanoamérica (p. ej., Coolican, 2005; Guardia, Freixa, Perú, & Turbany, 2008) demuestra que tales métodos (los que serán expuestos en el presente manuscrito) no están incluidos en sus contenidos.

En segundo lugar, ya que estas técnicas no están disponibles en los programas estándares de análisis estadístico, los investigadores no los han implementado debido al desconocimiento de las mismas o a su complejidad matemática. En último lugar, cuando estas u otras técnicas similares presentan sus fundamentos en revistas de estadística y de metodología cuantitativa que pueden ser de difícil lectura al investigador, existe un motivo más para no aprovecharse de estos conocimientos y usarlos frecuentemente. Algunas investigaciones han usado estos procedimientos (e.g., Caldwell, Silver, & Strada, 2010; Caldwell-Gunes, Silver, Smith, & Norton, 2016; Meijs, Cillessen, Scholte, Segers, & Spijkerman, 2010; Merino, 2011a; Merino, Calderón, & Manzanares, en prensa; Merino & Grimaldo, 2015), pero comparadas con la frecuencia con que las investigaciones usan correlaciones bivariadas, su aplicación aún podría considerarse emergente especialmente en el contexto hispano. Por lo tanto, el propósito de este trabajo es familiarizar al lector con un número de estas técnicas y describir los programas informáticos que los calculan.

PROCEDIMIENTOS ESPECÍFICOS

Supongamos que un investigador está interesado en determinar si las correlaciones entre los puntajes en satisfacción laboral y salario son diferentes entre

profesores de nivel primario y secundario. Como estas correlaciones provienen de diferentes muestras, se deben considerar como independientes. La fórmula z que probaría esta hipótesis, común en algunos libros de estadística, es la siguiente:

$$Z = \frac{z_1 - z_2}{\sqrt{\frac{1}{(N_1 - 3)} + \frac{1}{(N_2 - 3)}}}$$

En esta fórmula, z se refiere a la transformación Fisher (1921) de la correlación y N corresponde al tamaño muestral en cada grupo. Si la correlación es 0.75 ($N = 100$) en el grupo de primaria y 0.30 ($N = 100$) para el grupo de secundaria, entonces el resultado es $z = 5.34$, $p < .0001$. Esto puede indicar que, comparado con los profesores de secundaria, hay una alta correlación estadísticamente significativa entre la satisfacción laboral y sueldo en profesores de primaria. Aunque este estadístico es fácil de calcular, es raro hallarlo en algunos programas estadísticos estándar (por ejemplo, SPSS). Uno de los autores (NCS) ha escrito un programa interactivo y amigable (INDCOR; Silver & Burkey, 1991) para realizar esta prueba. El usuario simplemente ingresa las correlaciones y el tamaño muestral de cada grupo. El resultado (INDCOR.OUT) responde con el equivalente z y su valor de probabilidad.

Recientemente, Zou (2007) desarrolló un método basado en intervalos de confianza para examinar las diferencias entre dos correlaciones independientes. La premisa es que el intervalo de confianza informa la magnitud y la precisión de la relación, mientras que la prueba de hipótesis estándar mezcla estas dos características. (Zou, 2007). Un programa interactivo y amigable, llamado COMPCOR (Silver, 2009), hará este cálculo, y requiere del usuario el tamaño muestral y las correlaciones de cada grupo, y el porcentaje

del intervalo de confianza (generalmente, 95%). Los resultados aparecen en la Figura 1.

Usando el método de Zou (2007) en el ejemplo, los límites del intervalo de confianza son .2534 y .6542; ya que el intervalo de confianza no incluye el cero, entonces se obtendría un resultado estadísticamente significativo ($p < .05$). Ambos métodos son útiles solo cuando hay dos correlaciones independientes.

Pero supongamos que uno quiere incluir otros dos grupos, por ejemplo, profesores de educación técnica y de Universidad; entonces uno podría probar la hipótesis nula que las cuatro correlaciones son equivalentes. Como la prueba z anterior evalúa solo dos grupos en un tiempo, se podría necesitar múltiples pruebas z ; específicamente $4(4 - 1) / 2 = 6$. Sin embargo, este procedimiento inflaría dramáticamente el error Tipo I. Para combatir este problema, hay varias pruebas ómnibus para mantener el error Tipo I y aún muestra razonable poder estadístico. Rao (1970) describió uno de estos procedimientos, basado en la transformación Fisher (1921) y que se evalúa como un valor X^2 con $k-1$ grados de libertad (k es el número de correlaciones independientes). La fórmula correspondiente se halla en Rao (1970) y Silver & Burkey (1991).

Si también tenemos 100 profesores en Universidad y técnicos, con correlaciones entre 0.60 y 0.40, respectivamente, la prueba ómnibus sería un X^2 estadísticamente significativo de 25.53, $p < .0001$. Ya

que hay diferencias entre las correlaciones, se puede aplicar una prueba de rangos (Levy, 1976) para saber entre qué par de grupos ocurre las diferencias. El programa INDEPCOR (Silver & Burkey, 1991) realiza la prueba ómnibus y seguidamente una prueba de rangos, con el enfoque Student Newman-Keuls. Como ejemplo, La Figura 2 exhibe la salida de programa INDEPCOR con los datos de ejemplo utilizado.

La prueba de rangos nos indicaría que hay una significativa alta correlación entre satisfacción laboral y salario en profesores de primaria, que en profesores de secundaria, $p < .0001$ y profesores universitarios, $p < .0005$. Además, los profesores universitarios tienen una más alta correlación entre las variables examinadas que en profesores escolares, $p < .022$. Similarmente, también se propone otra prueba ómnibus, $X^2_{c(F)}$, basada en el estadístico $C(\alpha)$ de Neyman (1959). La fórmula para este método puede hallarse en Paul (1989). Los resultados de simulación de Paul (1989) indicaron que la prueba $X^2_{c(F)}$ mantuvo un razonable control del error Tipo I, mientras que el estadístico X^2 basado en la transformación z de Fisher fue conservador. Ambos procedimientos mostraron similar poder (Paul, 1989). En este caso, el $X^2_{c(F)}$ es 24.60, $p < .0001$. El programa INCOR (Silver, Zaikina, Hittner, & May, 2008) calcula la prueba $X^2_{c(F)}$ junto con una prueba de rangos a posteriori (Levy, 1976), usando los grados de libertad del enfoque Fisher-Hayter (Hayter, 1986). La prueba de rangos Fisher-Hayter trata de retener un adecuado control del error Tipo I

```

FOR THE INDEPENDENT CORRELATIONS
r1 = .7500 AND r2 = .3000

THE SAMPLE SIZE FOR r1 = 100.0000
THE SAMPLE SIZE FOR r2 = 100.0000

THE .9500 CONFIDENCE INTERVAL FOR .7500 HAS A LOWER BOUND OF .6492 AND AN UPPER
BOUND OF .8249

THE .9500 CONFIDENCE INTERVAL FOR .3000 HAS A LOWER BOUND OF .1101 AND AN UPPER
BOUND OF .4688

*** TESTING THE DIFFERENCE BETWEEN INDEPENDENT rs ***
THE LOWER BOUND IS .2534 AND THE UPPER BOUND IS .6542
FOR THE .9500 CONFIDENCE INTERVAL

IF THE INTERVAL CONTAINS 0, THEN IT IS NONSIGNIFICANT

```

Figura 1. Salida del programa COMPOR, de acuerdo al ejemplo presentado (Silver, 2009).

y al mismo tiempo muestra razonable poder. Similar a INDEPCOR, el usuario debe ingresar el número de grupos, cada correlación y el tamaño muestral para cada correlación. La Figura 3 muestra la salida de INCOR.

Las conclusiones pueden ser similares a las obtenidas usando el programa INDEPCOR. En una nueva situación, supongamos que un investigador quiere examinar la diferencia en las correlaciones entre satisfacción laboral y salario pero removiendo

de ambas variables los efectos de la calificación del supervisor, en cada uno de los grupos. Fisher (1924) demostró que la distribución de una correlación parcial de orden k basada en n datos es igual a la correlación bivariada basada sobre $n-k$ datos. Por lo tanto, se podría usar una versión modificada de la fórmula de Rao (1970) para evaluar correlaciones parciales independientes. La fórmula modificada, basada en el trabajo de Levy y Narula (1978), se puede hallar en Silver, Wadiak, y Massey (1995). Por ejemplo, para un análisis correlacional para profesores de

```

THE GLOBAL TEST FOR EQUALITY = 25.5353
AND IT HAS A PROBABILITY = .0000

      4. STEP DIFFERENCES
1ST CORR      2ND CORR      RSTAT      PROB
( 1)      .300      ( 4)      .750      6.534      .00003

      3. STEP DIFFERENCES
1ST CORR      2ND CORR      RSTAT      PROB
( 1)      .300      ( 3)      .600      3.778      .02141
( 2)      .400      ( 4)      .750      5.410      .00045

      2. STEP DIFFERENCES
1ST CORR      2ND CORR      RSTAT      PROB
( 1)      .300      ( 2)      .400      1.124      .42715
( 2)      .400      ( 3)      .600      2.654      .06129
( 3)      .600      ( 4)      .750      2.756      .05206

```

Figura 2. Resultados de INDCOR y aplicación del enfoque post hoc Student Newman-Keuls.

```

THIS PROGRAM COMPUTES THE OMNIBUS C-ALPHA TEST
DESCRIBED BY PAUL IN 1989 FOR TESTING THE DIFFERENCE AMONG
MORE THAN TWO INDEPENDENT CORRELATIONS AND THE SUBSEQUENT
RANGE TEST PUBLISHED BY LEVY IN 1976

THE OMNIBUS TEST FOR EQUALITY = 24.6078
AND IT HAS A PROBABILITY = .0000

      4. STEP DIFFERENCES
1ST CORR      2ND CORR      RSTAT      PROB
( 1)      .300      ( 4)      .750      6.534      .00002

      3. STEP DIFFERENCES
1ST CORR      2ND CORR      RSTAT      PROB
( 1)      .300      ( 3)      .600      3.778      .02141
( 2)      .400      ( 4)      .750      5.410      .00045

      2. STEP DIFFERENCES
1ST CORR      2ND CORR      RSTAT      PROB
( 1)      .300      ( 2)      .400      1.124      .70634
( 2)      .400      ( 3)      .600      2.654      .14678
( 3)      .600      ( 4)      .750      2.756      .12662

```

Figura 3. Salida del programa INCOR

primaria, secundaria, técnico y universidad, y entre las variables satisfacción laboral, calificación del supervisor y salario, aparece en la Tabla 1.

El programa INDPART permite comparar correlaciones parciales (Silver, Wadiak, & Massey, 1995) y calcula una prueba ómnibus seguida de una prueba de rangos a posteriori usando el enfoque Student Newman-Keuls, similar al procedimiento en los programas previos. El resultado aparece en la Figura 4.

La correlación parcial es computada para cada grupo junto con el valor F y su nivel de significancia. En el

ejemplo, para los profesores de primaria hay una significativa relación positiva entre satisfacción laboral y salario luego de remover los efectos de la calificación de supervisores. Luego, la prueba ómnibus aparece con su nivel de significancia: hay una diferencia estadísticamente significativa entre los cuatro grupos de profesores tomando en cuenta las correlaciones parciales. Finalmente, la siguiente prueba de rangos indica que la correlación entre salario y satisfacción laboral, luego de remover la percepción de los supervisores, fue más alta para profesores de educación primaria que para los otros grupos de profesores ($p < .0001$).

Tabla 1
Ejemplo de correlaciones entre tres variables

	SATISFACCIÓN LABORAL - SALARIO	SATISFACCIÓN LABORAL - SUPERVISIÓN	SALARIO - SUPERVISIÓN
Prof. Primaria (n = 100)	.75	.40	.30
Prof. Secundaria (n = 100)	.40	.50	.60
Prof. Técnico (n = 100)	.60	.70	.50
Prof. Universidad (n = 100)	.30	.40	.30

```

THE PARTIAL CORRELATION FOR GROUP 1 = .7206
THE F VALUE = 104.76000 WITH A PROBABILITY = .0000

THE PARTIAL CORRELATION FOR GROUP 2 = .1443
THE F VALUE = 2.06383 WITH A PROBABILITY = .1540

THE PARTIAL CORRELATION FOR GROUP 3 = .4042
THE F VALUE = 18.94532 WITH A PROBABILITY = .0000

THE PARTIAL CORRELATION FOR GROUP 4 = .2059
THE F VALUE = 4.29344 WITH A PROBABILITY = .0409

THE GLOBAL TEST FOR EQUALITY = 34.4658
AND IT HAS A PROBABILITY = .0000

4. STEP DIFFERENCES
1ST PARTIAL CORR      2ND PARTIAL CORR      RSTAT      PROB
( 1)      .1443      ( 4)      .7206      7.4807      .00000

3. STEP DIFFERENCES
1ST PARTIAL CORR      2ND PARTIAL CORR      RSTAT      PROB
( 1)      .1443      ( 3)      .4042      2.7761      .12287
( 2)      .2059      ( 4)      .7206      6.8584      .00001

2. STEP DIFFERENCES
1ST PARTIAL CORR      2ND PARTIAL CORR      RSTAT      PROB
( 1)      .1443      ( 2)      .2059      .6223      .65667
( 2)      .2059      ( 3)      .4042      2.1538      .12858
( 3)      .4042      ( 4)      .7206      4.7045      .00096
    
```

Figura 4. Salida del programa INDPART

COMENTARIOS FINALES

En este documento se han considerado los análisis para correlaciones independientes; sin embargo, también se pueden probar las diferencias entre correlaciones dependientes, pero es otro contexto cuya extensión supera al de este artículo. Los métodos expuestos expanden las situaciones de análisis que pueden ser consideradas en una investigación, en pregrado y posgrado. Por lo tanto, crean el contexto para plantear hipótesis relevantes que van más allá de la tradicional presentación de las correlaciones y de la evaluación de su significancia estadística. En las pruebas estadísticas presentadas se puede usar el estimador q para ver la magnitud del efecto (Cohen, 1992), tomando en cuenta los siguientes niveles: trivial ($< \pm 0.20$), baja ($\geq \pm 0.20$), moderada ($\geq \pm 0.50$), alta ($\geq \pm 0.80$). Esta herramienta es apropiada para implementar un mejor reporte cuantitativo de resultados, junto con los métodos basados en intervalos de confianza (Merino, 2011b). Finalmente, los programas citados no están disponibles en programas estadísticos estándar, y se distribuyen gratuitamente escribiendo a uno de los autores.

REFERENCIAS

- Caldwell, R. M., Silver, N. C., & Strada, M. (2010). Substance Abuse, Familial Factors, and Mental Health: Exploring Racial and Ethnic Group Differences among African American, Caucasian, and Hispanic Juvenile Offenders. *American Journal of Family Therapy, 38*(4), 310-321. doi:10.1080/01926187.2010.493438
- Caldwell-Gunes, R., Silver, N.C., Smith, K.M., & Norton, K. A. (2016). Racial and ethnic differences in factors related to substance use among adult female offenders: Implications for treatment. *Journal of Forensic Psychology Practice, 16*(1), 39-48. doi:10.1080/15228932.2016.1124595
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. doi:10.1037/0033-2909.112.1.155
- Coolican, H. (2005). *Métodos de investigación y estadística en psicología*. México, DF: El Manual Moderno.
- Fisher, R. A. (1921). On the probable error of a coefficient of correlation deduced from a small sample. *Metron, 1*, 1-32.
- Guardia, J., Freixa, M., Però, M., & Turbany, J. (2008). *Análisis de datos en Psicología* (2da ed.). Barcelona: Delta publicaciones.
- Hayter, A. J. (1986). The maximum familywise error rate of Fisher's least significant difference test. *Journal of the American Statistical Association, 81*(396), 1000-1004.
- Levy, K. L. (1976). A multiple range procedure for independent correlations. *Educational and Psychological Measurement, 36*(1), 27-31. doi:10.1177/001316447603600103
- Levy, K. L., & Narula, S. C. (1978). Testing hypotheses concerning partial correlations: Some methods and discussion. *International Statistical Review, 46*(2), 215-218. doi:10.2307/1402814
- Meijs, N., Cillessen, A. H. N., Scholte, R. H. J., Segers, E., & Spijkerman, R. (2010). Social intelligence and academic achievement as predictors of adolescent popularity. *Journal of Youth and Adolescence, 39*(1), 9373-9379. doi:10.1007/s10964-008-9373-9
- Merino, C. (2011a). Test Gestáltico Visomotor de Bender Modificado y Test de Caras: Una evaluación de la validez de constructo. *Cuadernos de Neuropsicología, 5*(2), 129-143.
- Merino, C. (2011b). Carta editorial [carta al Editor]. *Avances en Psicología Latinoamericana, 29*(2), 7-10.
- Merino, C., Calderón, G., & Manzanares, E., (En prensa) Estudio comparativo del acuerdo y consistencia intercalificadores en el test gestáltico visomotor de Bender 2a edición. *Revista Latinoamericana de Psicología*.
- Merino, C., & Grimaldo, M. (2015). Análisis no paramétrico de rankings de tolerancia en comportamientos moralmente cuestionables. *Revista Digital de Investigación en Docencia Universitaria, 9*(2), 3-48. doi:10.19083/ridu.9.448 <http://dx.doi.org/10.1016/j.rlp.2015.09.011>
- Neyman, J. (1959). Optimal asymptotic tests of composite hypotheses. In U. Grenander (Ed.), *Probability and Statistics: The Harold Cramer Volume* (pp. 213-234). New York: Wiley.
- Paul, S. R. (1989). Test for the equality of several correlation coefficients. *The Canadian Journal of Statistics, 17*(2), 217-227. doi:10.2307/3314850
- Rao, R. (1970). *Advanced statistical methods in biometric research*. New York: Hafner.
- Silver, N. C. (2009). *A computer program for comparing correlations using confidence intervals*. Unpublished manuscript.
- Silver, N. C., & Burkey, R. T. (1991). A FORTRAN 77 program for testing the differences among independent correlations. *Educational and Psychological Measurement, 51* (3), 641-644. doi:10.1177/0013164491513011
- Silver, N. C., Wadiak, D. L., & Massey, C. J. (1995). A Microsoft FORTRAN 77 program for testing the differences among independent first-order partial correlations. *Educational and Psychological Measurement, 55*(2), 245-248. doi:10.1177/0013164495055002008
- Silver, N. C., Zaikina, H., Hittner, J. B., & May, K. (2008). INCOR: A computer program for testing differences among independent correlations. *Molecular Ecology Resources, 8* (4), 763-764. doi:10.1111/j.1755-0998.2008.02107.x
- Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods, 12* (4), 399-413. doi:10.1037/1082-989X.12.4.399

© Los autores. Este artículo es publicado por la Revista Digital de Investigación en Docencia Universitaria del Área de Institutional Research and Effectiveness de la Dirección de Aseguramiento de la Calidad, Universidad Peruana de Ciencias Aplicadas. Este es un artículo de acceso abierto, distribuido bajo los términos de la Licencia Creative Commons Atribución-Compartir Igual 4.0 Internacional. (<http://creativecommons.org/licenses/by-sa/4.0/>), que permite el uso no comercial, distribución y reproducción en cualquier medio, siempre que la obra original sea debidamente citada.