



Comparación de intervalos de confianza para el coeficiente de correlación

Comparison of confidence intervals for the correlation coefficient

Liliana Vanessa Pacheco^a
liliana.pacheco24@gmail.com

Juan Carlos Correa^b
jccorreamorales@gmail.com

Resumen

La construcción de intervalos de confianza para la estimación de la correlación en la distribución normal bivariable y multivariable, ρ , es un problema importante en el trabajo estadístico aplicado. Uno de los propósitos principales de este trabajo es hacer una revisión de los diferentes procedimientos para su construcción. Se realizó además, un estudio de simulación para analizar el comportamiento de los niveles de confianza reales y compararlos con los teóricos, analizar el comportamiento de las longitudes de los intervalos de confianza logrados por los nueve métodos considerados y determinar cuál metodología provee los intervalos más cortos. Así como también se obtuvo un indicador que resume de manera más efectiva la calidad del intervalo analizado.

Palabras clave: coeficiente de correlación, estimación, intervalo de confianza.

Abstract

The construction of confidence intervals to estimate the correlation in the normal bivariate and multivariate distribution, ρ , is an important problem in applied statistical work. One of the main purposes of this work is to make a review of the different procedures for their construction. In addition, a simulation study was conducted to analyze the behavior of real confidence levels and compare them to theoretical ones, analyze the behavior of the lengths of the confidence intervals achieved by the nine methods considered and determine which methodology provides the shortest intervals. Likewise an indicator that summarizes more effectively the quality of the analyzed interval was also obtained.

Keywords: correlation coefficient, estimation, confidence interval.

^aUniversidad Nacional de Colombia, sede Medellín. Escuela de Estadística. Colombia.

^bUniversidad Nacional de Colombia, sede Medellín. Escuela de Estadística. Colombia.

1. Introducción

El coeficiente de correlación es una de las medidas estadísticas más usadas dentro del trabajo aplicado. Algunas de sus propiedades fueron estudiadas por Zheng & Matis (1994), donde presentan y demuestran las que consideraron las más destacadas:

1. $|R| \leq 1$.
2. Si $|R| = 1$ entonces los pares $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ yacen en una línea recta.
3. Recíprocamente, si los $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ yacen en una línea recta, entonces $|R| = 1$.

Debido a su amplia utilización, varias son sus interpretaciones. Falk & Well (1997) sustentan que el coeficiente de correlación de Pearson, ρ , es ampliamente usado en campos como la educación, psicología, y todas las ciencias sociales, y el concepto es empleado en diversas metodologías de tipo estadístico.

La estimación del coeficiente de correlación por medio de intervalos es importante, y para ello se disponen de diversos métodos. La metodología quizá más conocida es la propuesta originalmente por Fisher en la cual se realiza una transformación del coeficiente de correlación muestral, r , y asumiendo normalidad asintótica, se desarrolla un intervalo para el coeficiente de correlación poblacional ρ (Krishnamoorthy & Xia 2007). También se conocen transformaciones adicionales hechas por Hotelling (1953) a la propuesta inicial de Fisher.

El problema para el analista es la carencia de reglas sobre cuál fórmula es preferible. Para esto se pretende realizar un estudio de simulación que permita analizar el comportamiento de los niveles de confianza reales y compararlos con los teóricos para los diversos intervalos disponibles. Así como también, hacer una comparación de las longitudes del intervalo obtenido por las diferentes metodologías y la implementación de un indicador que permita relacionar los dos criterios de evaluación anteriormente mencionados.

Algunas de las metodologías empleadas para la construcción de los intervalos de confianza pueden encontrarse en Fisher (1921), Hotelling (1953), Pawitan (2001), Efron (1979) y Krishnamoorthy & Xia (2007).

Además, en Krishnamoorthy & Xia (2007) se pueden encontrar los resultados de estudios comparativos realizados previamente para tres métodos de construcción de intervalos, en los cuales la metodología consistió en la obtención de límites superiores para ρ bajo diferentes escenarios: Tamaños de muestra pequeños ($n=5$, $n=10$, $n=20$ y $n=30$) y valores de r positivos; y el cálculo de la probabilidad $P(R \leq r|n, \rho_U)$. Estos estudios mostraron que, en particular, el mejor método para construir intervalos unilaterales para ρ en muestras pequeñas es el de pivote generalizado.

2. Intervalos de confianza

2.1. Método I: basado en la transformación Arcotangente

Este intervalo puede considerarse el intervalo clásico para este parámetro y fue propuesto por Fisher (1921). Debido a que la distribución del coeficiente de correlación muestral no es centrada y/o simétrica, el cálculo de intervalos de confianza a partir de los cuantiles de la distribución no se hace sencillo. Por tanto, Fisher propone la transformación arcotangente hiperbólico:

$$r = \tanh(z) \Leftrightarrow z = \frac{1}{2} \log \frac{1+r}{1-r} \tag{1}$$

y demostró que z tiene una distribución aproximadamente normal cuando el tamaño muestral es grande. Dicha distribución normal se caracteriza por una media $\xi = \frac{1}{2} \log \left(\frac{1+\rho}{1-\rho} \right)$ y varianza $\frac{1}{n-3}$. El intervalo hallado a partir de la transformación Arcotangente hiperbólico tiene la siguiente forma:

$$\left(\tanh \left(\operatorname{arctanh}(r) - \frac{z_{\alpha/2}}{\sqrt{n-3}} \right), \tanh \left(\operatorname{arctanh}(r) + \frac{z_{\alpha/2}}{\sqrt{n-3}} \right) \right) \tag{2}$$

donde $z_{\frac{\alpha}{2}}$ es el percentil superior $\alpha/2$ de la distribución normal estándar.

2.1.1. Modificaciones a la transformación Arcotangente

Teniendo en cuenta el hecho de que la transformación propuesta por Fisher funciona adecuadamente siempre y cuando los tamaños muestrales sean grandes, se hizo necesario encontrar la manera de reducir el error al trabajar esta transformación en muestras pequeñas. Hotelling (1953) estudió esta situación y propuso 4 transformaciones z_i con $i = 1, \dots, 4$, para la transformación z original de Fisher, las cuales también, asintóticamente tienen una distribución Normal con media ξ_i y varianza $\frac{1}{n-1}$:

$$z_1 = z - \frac{7z+r}{8(n-1)} \quad \xi_1 = \xi - \frac{7\xi+\rho}{8(n-1)} \tag{3}$$

$$z_2 = z - \frac{7z+r}{8(n-1)} - \frac{119z+57r+3r^2}{384(n-1)^2} \quad \xi_2 = \xi - \frac{7\xi+\rho}{8(n-1)} - \frac{119\xi+57\rho+3\rho^2}{384(n-1)^2} \tag{4}$$

$$z_3 = z - \frac{3z+r}{4(n-1)} \quad \xi_3 = \xi - \frac{3\xi+\rho}{4(n-1)} \tag{5}$$

$$z_4 = z - \frac{3z+r}{4(n-1)} - \frac{23z+33r-5r^2}{96(n-1)^2} \quad \xi_4 = \xi - \frac{3\xi+\rho}{4(n-1)} - \frac{23\xi+33\rho-5\rho^2}{96(n-1)^2} \tag{6}$$

2.2. Método II: Intervalo de la razón de verosimilitud

Kalbfleish (1985) y Pawitan (2001) presentan la metodología para construir intervalos de verosimilitud. Un intervalo de la razón de verosimilitud para θ es definido como el conjunto de valores parametrales con valores altamente verosímiles:

$$\left\{ \theta, \frac{L(\theta)}{L(\hat{\theta})} > c \right\} \quad (7)$$

para un valor $c \in (0, 1)$ y donde $\hat{\theta}$ es el estimador muestral del parámetro θ

Sabiendo que $2 \log \frac{L(\hat{\theta})}{L(\theta)} \xrightarrow{d} \chi_1^2$, la probabilidad de cubrimiento aproximada para este tipo de intervalos está dada por:

$$P \left(\frac{L(\theta)}{L(\hat{\theta})} > c \right) = P \left(2 \log \frac{L(\hat{\theta})}{L(\theta)} < -2 \log c \right) \quad (8)$$

$$\approx P(\chi_1^2 < -2 \log c). \quad (9)$$

Luego, para cualquier valor $0 < \alpha < 1$ el punto de corte c es:

$$c = \exp \left[-\frac{1}{2} \chi_{1,1-\alpha}^2 \right] \quad (10)$$

donde $\chi_{1,1-\alpha}^2$ es el $100(1 - \alpha)$ percentil de una χ_1^2 . Por tanto:

$$P \left(\frac{L(\theta)}{L(\hat{\theta})} > c \right) = P(\chi_1^2 < \chi_{1,1-\alpha}^2) = 1 - \alpha. \quad (11)$$

Si $L(\rho)$ es la función de verosimilitud, se define la *función de verosimilitud relativa* como:

$$R(\rho) = \frac{L(\rho)}{L(r)} \quad (12)$$

El conjunto de valores de ρ para los cuales $R(\rho) > c$ es llamado *intervalo de $100 \times c$ % de verosimilitud* para ρ . Los intervalos del 14.7% y del 3.6% de verosimilitud corresponden a intervalos de confianza de niveles del 95% y 99% aproximadamente.

Lo que se debe hacer entonces es hallar las raíces que nos dan los límites del intervalo. Para el caso del parámetro ρ tenemos que un intervalo de confianza del 95% se halla encontrando el par de raíces tal que

$$R(\rho) = \frac{L(\rho)}{L(r)} \quad (13)$$

$$= \left(\frac{1 - \rho^2}{1 - r^2} \right)^{\frac{(n-1)}{2}} \frac{\int_0^\infty (\cosh w - \rho r)^{-(n-1)} dw}{\int_0^\infty (\cosh w - r^2)^{-(n-1)} dw} \geq K(k, \alpha) \quad (14)$$

donde $K(k, \alpha)$ es el valor crítico mínimo con el cual aseguramos una confianza deseada, ya sea del 95 % o 99 %, por ejemplo.

2.3. Método III: Bootstrap

La primera aplicación del método bootstrap fue en la determinación del intervalo de confianza del coeficiente de correlación en el artículo seminal de Efron (1979).

- A partir de la muestra $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ se calculan las estimaciones de máxima verosimilitud del vector de medias y de la matriz de varianzas y covarianzas de la distribución normal bivariable.
- Se generan M muestras de tamaño n de una distribución normal bivariable con parámetros $\hat{\mu}$ y $\hat{\Sigma}$. Y para cada una de estas muestras se estima el parámetro ρ , por ejemplo, para la muestra j el valor del estimador para el coeficiente de correlación es r_j .
- Para los $r_j, j = 1, \dots, M$, se construye un histograma y se calculan los percentiles 0.025 y 0.975 los cuales se denotarán: $r_i^{\{0.025\}}$ y $r_i^{\{0.975\}}$.

Y el intervalo de Bootstrap para ρ está dado por $r_i^{\{0.025\}}$ y $r_i^{\{0.975\}}$ como límite inferior y superior respectivamente.

2.4. Método IV: Intervalo de Jeyaratnam

Jeyaratnam propone un intervalo para el coeficiente de correlación de la distribución normal bivariada y este tiene la siguiente forma (Krishnamoorthy & Xia 2007):

$$\left(\frac{r - w}{1 - rw}, \frac{r + w}{1 + rw} \right) \tag{15}$$

donde

$$w = \frac{\frac{t_{n-2, 1-\alpha/2}}{\sqrt{n-2}}}{\sqrt{1 + \frac{(t_{n-2, 1-\alpha/2})^2}{n-2}}} \tag{16}$$

y $t_{m,p}$ denota el p -ésimo cuantil de la distribución t -Student con m grados de libertad.

2.5. Método V: Test generalizado para ρ

Según el artículo publicado por Krishnamoorthy & Xia (2007), los autores citados proponen un algoritmo para construir un intervalo de confianza para ρ a partir de

la distribución del pivote generalizado para el coeficiente de correlación:

$$G_{\rho_{ij}} = \frac{\sum_{k=1}^j b_{ik} b_{jk}}{\sqrt{\sum_{k=1}^i b_{ik}^2} \sqrt{\sum_{k=1}^j b_{jk}^2}} \quad (17)$$

para $i > j$. Que en el caso bivariado lo anterior se expresa de la siguiente forma:

$$G_{\rho_{21}} = \frac{b_{21}}{\sqrt{b_{21}^2 + b_{22}^2}} \quad (18)$$

Y simplificando la expresión anterior, se tiene:

$$G_{\rho_{21}} = \frac{r^* V_{22} - V_{21}}{\sqrt{(r^* V_{22} - V_{21})^2 + V_{11}^2}} \quad (19)$$

Donde $r^* = \frac{r}{\sqrt{1-r^2}}$ y V es una matriz triangular inferior, las V_{ij} 's son independientes con $V_{ii}^2 \sim \chi_{n-i}^2$ para $i = 1, \dots, p$ y $V_{ij} \sim N(0, 1)$ para $i < j$. Para mayor detalle, se sugiere remitirse al desarrollo completo para la obtención de los $G_{\rho_{ij}}$ y la matriz V , que se encuentra en Krishnamoorthy & Xia (2007).

Entonces, según los autores, para un r dado la distribución de G_{ρ} no depende de algún parámetro que sea desconocido, y el intervalo se calcula empleando el siguiente algoritmo:

Algoritmo 1 Generar valores del pivote $G_{\rho_{ij}}$

Requiere Un n y ρ fijo

Calcular: $r^* = \rho / \sqrt{1 - \rho^2}$

Para: $i = 1$ hasta m **Haga**

Generar: $Z_0 \sim N(0, 1)$.

Generar: $U1 \sim \chi_{n-1}^2$.

Generar: $U2 \sim \chi_{n-2}^2$.

Calcular: $Q_i = \frac{r^* \sqrt{U2} - Z_0}{\sqrt{(r^* \sqrt{U2} - Z_0)^2 + U1}}$

Fin del ciclo.

Luego, los percentiles $\frac{\alpha}{2}$ y $(1 - \frac{\alpha}{2})$ de los valores calculados para el pivote $G_{\rho_{ij}}$ mediante el mencionado algoritmo conforman los límites del intervalo de confianza al $100(1 - \alpha)\%$ para ρ .

3. Resultados de la simulación

Para comparar los nueve métodos de construcción de intervalos de confianza en este caso se realizó una simulación en R en la cual se consideraron combinaciones

de (ρ, n) con valores de $\rho = 0.0, 0.1, 0.2, \dots, 0.9$ y de $n = 5, 10, 20, 50, 100$; un vector de medias $\mu = \begin{bmatrix} 1 \\ \rho \\ 1 \end{bmatrix}$ y $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$.

Para cada pareja se realizaron 1000 réplicas y se calcularon las fórmulas previas a un nivel de confianza del 95 % (Este es conocido como el nivel nominal). Para cada método y combinación se calculó la mediana de la longitud de los 1000 intervalos calculados y la proporción de intervalos que cubren el verdadero valor de ρ , esto es lo que se llama el nivel de confianza real. Los resultados se encuentran en las tablas 1 al 5 y en las Figuras 1 y 2.

De las gráficas y las tablas observamos que las longitudes más amplias en los intervalos se encuentran en los tamaños de muestra más pequeños, siendo el caso de $n = 5$ donde los nueve tipos de intervalo alcanzan las mayores longitudes para el respectivo valor de ρ . Además, vemos que a medida que se amplía el valor de ρ para un tamaño de muestra particular, las longitudes son cada vez menores, lo que sugiere que los intervalos de confianza alcanzan menores longitudes cuando el valor verdadero de ρ se va acercando a 1.

Con respecto al nivel real alcanzado por cada intervalo se observa que cada uno de los métodos cuando el tamaño muestral es bastante pequeño, por ejemplo, $n = 5$, tienen una probabilidad de cobertura diferente a la deseada 95 %, y no es homogénea en cada valor de ρ , es decir, que en cada método se observa que algunas veces el nivel real supera al nominal y a veces es al contrario, conforme se varía el valor de ρ . Los más cercanos a 95 % de nivel real, cuando $n = 5$ son ArcTanh, Jeyaratnam y P.G; y el que se comporta peor es Bootstrap. A medida que aumenta el tamaño de muestra se nota una tendencia en todos los intervalos a estar cerca del nivel nominal deseado, 95 %.

3.1. Índice de resúmenes

A manera de resumen se presenta un índice que señala la calidad de las metodologías anteriormente mencionadas. Este índice busca favorecer a aquellos métodos que presenten longitudes de intervalo pequeños y niveles reales de confianza cercanos o mayores al 95 %:

$$I = (2 - LI) \frac{NR}{NN} \quad (20)$$

donde:

- LI: Mediana de la longitud del intervalo.
- NR: Promedio del nivel real del intervalo.
- NN: Nivel nominal de los intervalos, que en este caso es 0.95.

Tabla 1: Longitud y nivel de confianza de los intervalos. Tamaño de muestra 5. Fuente: elaboración propia
 $n = 5$

ρ		Bootstrap	Arctanh	L.R	Jeyaratnam	Z1	Z2	Z3	Z4	P.G
0.0	Longitud	1.6233	1.6864	1.4864	1.6766	1.6056	1.6267	1.5718	1.5880	1.5498
	Nivel	0.9010	0.9530	0.9270	0.9500	0.9290	0.9360	0.9190	0.9240	0.9430
0.1	Longitud	1.6392	1.6903	1.4905	1.6806	1.6102	1.6311	1.5770	1.5928	1.5346
	Nivel	0.9130	0.9470	0.9330	0.9460	0.9340	0.9380	0.9260	0.9310	0.9440
0.2	Longitud	1.6221	1.6848	1.4848	1.6750	1.6038	1.6249	1.5699	1.5861	1.5360
	Nivel	0.9120	0.9610	0.9440	0.9590	0.9490	0.9520	0.9380	0.9450	0.9490
0.3	Longitud	1.6003	1.6678	1.4671	1.6576	1.5837	1.6056	1.5487	1.5654	1.5235
	Nivel	0.8920	0.9390	0.9140	0.9340	0.9120	0.9220	0.9040	0.9080	0.9420
0.4	Longitud	1.5846	1.6597	1.4588	1.6492	1.5741	1.5964	1.5386	1.5556	1.4994
	Nivel	0.9130	0.9470	0.9360	0.9450	0.9370	0.9400	0.9280	0.9310	0.9600
0.5	Longitud	1.5141	1.6095	1.4081	1.5979	1.5157	1.5399	1.4772	1.4955	1.4397
	Nivel	0.8940	0.9550	0.9290	0.9540	0.9290	0.9370	0.9180	0.9240	0.9420
0.6	Longitud	1.2966	1.4664	1.2712	1.4522	1.3534	1.3821	1.3086	1.3298	1.3934
	Nivel	0.9170	0.9590	0.9410	0.9570	0.9390	0.9450	0.9330	0.9340	0.9540
0.7	Longitud	1.1093	1.3313	1.1492	1.3153	1.2061	1.2374	1.1578	1.1806	1.2438
	Nivel	0.8760	0.9350	0.9140	0.9330	0.9170	0.9200	0.9010	0.9130	0.9590
0.8	Longitud	0.8765	1.1527	0.9949	1.1353	1.0195	1.0521	0.9698	0.9931	1.0838
	Nivel	0.8970	0.9510	0.9270	0.9490	0.9360	0.9410	0.9250	0.9310	0.9520
0.9	Longitud	0.4787	0.7521	0.6612	0.7355	0.6314	0.6598	0.5896	0.6090	0.7449
	Nivel	0.8960	0.9550	0.9170	0.9510	0.9330	0.9390	0.9240	0.9290	0.9460

Tabla 2: Longitud y nivel de confianza de los intervalos. Tamaño de muestra 10. Fuente: elaboración propia
 $n = 10$

ρ		Bootstrap	Arctanh	L.R	Jeyaratnam	Z1	Z2	Z3	Z4	P.G
0.0	Longitud	1.1964	1.2101	1.1349	1.2146	1.1890	1.1928	1.1753	1.1782	1.1521
	Nivel	0.9460	0.9560	0.9510	0.9570	0.9510	0.9520	0.9490	0.9500	0.9570
0.1	Longitud	1.1960	1.2102	1.1350	1.2147	1.1891	1.1930	1.1754	1.1783	1.1456
	Nivel	0.9360	0.9470	0.9410	0.9490	0.9430	0.9440	0.9390	0.9400	0.9470
0.2	Longitud	1.1841	1.1956	1.1220	1.2002	1.1745	1.1784	1.1608	1.1637	1.1393
	Nivel	0.9270	0.9500	0.9410	0.9500	0.9430	0.9440	0.9390	0.9400	0.9460
0.3	Longitud	1.1494	1.1667	1.0968	1.1713	1.1460	1.1494	1.1318	1.1347	1.1105
	Nivel	0.9160	0.9310	0.9250	0.9340	0.9260	0.9260	0.9210	0.9220	0.9570
0.4	Longitud	1.0879	1.1161	1.0523	1.1207	1.0949	1.0988	1.0812	1.0841	1.0698
	Nivel	0.9460	0.9660	0.9600	0.9660	0.9630	0.9630	0.9580	0.9590	0.9520
0.5	Longitud	0.9960	1.0308	0.9770	1.0354	1.0097	1.0135	0.9961	0.9990	0.9997
	Nivel	0.9350	0.9550	0.9460	0.9560	0.9450	0.9450	0.9440	0.9440	0.9590
0.6	Longitud	0.8717	0.9138	0.8731	0.9183	0.8934	0.8971	0.8802	0.8830	0.8713
	Nivel	0.9300	0.9460	0.9370	0.9480	0.9370	0.9380	0.9340	0.9340	0.9620
0.7	Longitud	0.7072	0.7581	0.7331	0.7623	0.7392	0.7426	0.7271	0.7296	0.7578
	Nivel	0.9280	0.9460	0.9420	0.9470	0.9410	0.9420	0.9390	0.9400	0.9440
0.8	Longitud	0.4859	0.5427	0.5349	0.5461	0.5272	0.5299	0.5173	0.5194	0.5887
	Nivel	0.9250	0.9420	0.9390	0.9430	0.9410	0.9410	0.9390	0.9390	0.9580
0.9	Longitud	0.2707	0.3157	0.3186	0.3180	0.3055	0.3073	0.2991	0.3004	0.3340
	Nivel	0.9420	0.9600	0.9480	0.9600	0.9570	0.9570	0.9510	0.9530	0.9450

Tabla 3: Longitud y nivel de confianza de los intervalos. Tamaño de muestra 20. Fuente: elaboración propia
 $n = 20$

ρ		Bootstrap	Arctanh	L.R	Jeyaratnam	Z1	Z2	Z3	Z4	P.G
0.0	Longitud	0.8571	0.8696	0.8408	0.8720	0.8632	0.8639	0.8580	0.8585	0.8395
	Nivel	0.9500	0.9630	0.9560	0.9630	0.9580	0.9590	0.9560	0.9560	0.9570
0.1	Longitud	0.8524	0.8631	0.8348	0.8655	0.8567	0.8574	0.8515	0.8520	0.8351
	Nivel	0.9330	0.9330	0.9260	0.9360	0.9280	0.9280	0.9260	0.9260	0.9450
0.2	Longitud	0.8416	0.8492	0.8221	0.8516	0.8429	0.8436	0.8378	0.8383	0.8209
	Nivel	0.9310	0.9400	0.9370	0.9400	0.9390	0.9390	0.9370	0.9370	0.9490
0.3	Longitud	0.8101	0.8173	0.7927	0.8197	0.8112	0.8118	0.8062	0.8067	0.7870
	Nivel	0.9320	0.9420	0.9440	0.9430	0.9420	0.9420	0.9410	0.9410	0.9520
0.4	Longitud	0.7551	0.7659	0.7451	0.7682	0.7599	0.7606	0.7551	0.7556	0.7500
	Nivel	0.9390	0.9480	0.9430	0.9490	0.9460	0.9460	0.9460	0.9460	0.9440
0.5	Longitud	0.6815	0.6900	0.6744	0.6921	0.6844	0.6850	0.6800	0.6804	0.6877
	Nivel	0.9370	0.9440	0.9380	0.9440	0.9420	0.9420	0.9410	0.9410	0.9510
0.6	Longitud	0.5814	0.5936	0.5837	0.5955	0.5886	0.5891	0.5845	0.5849	0.5870
	Nivel	0.9400	0.9480	0.9440	0.9480	0.9460	0.9460	0.9440	0.9440	0.9440
0.7	Longitud	0.4691	0.4850	0.4804	0.4866	0.4806	0.4811	0.4772	0.4775	0.4810
	Nivel	0.9540	0.9600	0.9530	0.9610	0.9590	0.9590	0.9580	0.9580	0.9490
0.8	Longitud	0.3391	0.3565	0.3563	0.3577	0.3531	0.3535	0.3504	0.3507	0.3549
	Nivel	0.9370	0.9450	0.9440	0.9450	0.9430	0.9430	0.9420	0.9420	0.9510
0.9	Longitud	0.1823	0.1935	0.1958	0.1942	0.1915	0.1917	0.1899	0.1901	0.1931
	Nivel	0.9430	0.9480	0.9430	0.9480	0.9470	0.9470	0.9470	0.9470	0.9450

Tabla 4: Longitud y nivel de confianza de los intervalos. Tamaño de muestra 50. Fuente: elaboración propia
 $n = 50$

ρ		Bootstrap	Arctanh	L.R	Jeyaratnam	Z1	Z2	Z3	Z4	P.G
0.0	Longitud	0.5461	0.5519	0.5441	0.5526	0.5504	0.5505	0.5491	0.5491	0.5408
	Nivel	0.9530	0.9530	0.9510	0.9530	0.9520	0.9520	0.9510	0.9510	0.9480
0.1	Longitud	0.5421	0.5483	0.5406	0.5490	0.5468	0.5469	0.5455	0.5455	0.5389
	Nivel	0.9470	0.9540	0.9540	0.9540	0.9540	0.9540	0.9540	0.9540	0.9530
0.2	Longitud	0.5295	0.5343	0.5272	0.5350	0.5329	0.5330	0.5316	0.5316	0.5271
	Nivel	0.9570	0.9580	0.9580	0.9580	0.9580	0.9580	0.9570	0.9570	0.9370
0.3	Longitud	0.5075	0.5114	0.5050	0.5120	0.5100	0.5100	0.5087	0.5087	0.5010
	Nivel	0.9420	0.9490	0.9490	0.9500	0.9490	0.9490	0.9490	0.9490	0.9340
0.4	Longitud	0.4660	0.4696	0.4646	0.4702	0.4683	0.4683	0.4671	0.4672	0.4686
	Nivel	0.9400	0.9410	0.9380	0.9410	0.9390	0.9390	0.9390	0.9390	0.9390
0.5	Longitud	0.4211	0.4240	0.4203	0.4245	0.4228	0.4229	0.4217	0.4218	0.4187
	Nivel	0.9370	0.9430	0.9430	0.9430	0.9420	0.9420	0.9410	0.9410	0.9350
0.6	Longitud	0.3576	0.3623	0.3601	0.3627	0.3612	0.3613	0.3603	0.3603	0.3594
	Nivel	0.9510	0.9570	0.9580	0.9570	0.9560	0.9560	0.9560	0.9560	0.9530
0.7	Longitud	0.2819	0.2869	0.2862	0.2873	0.2861	0.2861	0.2853	0.2854	0.2942
	Nivel	0.9370	0.9500	0.9500	0.9500	0.9500	0.9500	0.9490	0.9500	0.9550
0.8	Longitud	0.2046	0.2096	0.2098	0.2099	0.2090	0.2090	0.2084	0.2085	0.2076
	Nivel	0.9420	0.9470	0.9430	0.9470	0.9470	0.9470	0.9460	0.9460	0.9400
0.9	Longitud	0.1095	0.1126	0.1132	0.1127	0.1122	0.1123	0.1119	0.1119	0.1126
	Nivel	0.9520	0.9550	0.9530	0.9550	0.9550	0.9550	0.9530	0.9530	0.9500

Tabla 5: Longitud y nivel de confianza de los intervalos. Tamaño de muestra 100. Fuente: elaboración propia
 $n = 100$

ρ		Bootstrap	Arctanh	L.R	Jeyaratnam	Z1	Z2	Z3	Z4	P.G
0.0	Longitud	0.3874	0.3912	0.3882	0.3914	0.3907	0.3907	0.3902	0.3902	0.3858
	Nivel	0.9470	0.9530	0.9510	0.9530	0.9530	0.9530	0.9520	0.9520	0.9330
0.1	Longitud	0.3846	0.3887	0.3858	0.3890	0.3882	0.3882	0.3877	0.3877	0.3824
	Nivel	0.9510	0.9530	0.9510	0.9530	0.9520	0.9520	0.9520	0.9520	0.9460
0.2	Longitud	0.3763	0.3788	0.3760	0.3790	0.3783	0.3783	0.3778	0.3778	0.3735
	Nivel	0.9470	0.9510	0.9500	0.9510	0.9500	0.9500	0.9490	0.9500	0.9320
0.3	Longitud	0.3558	0.3595	0.3571	0.3597	0.3590	0.3590	0.3586	0.3586	0.3554
	Nivel	0.9420	0.9450	0.9440	0.9450	0.9430	0.9430	0.9430	0.9430	0.9430
0.4	Longitud	0.3297	0.3314	0.3295	0.3316	0.3309	0.3310	0.3305	0.3305	0.3291
	Nivel	0.9490	0.9490	0.9480	0.9490	0.9490	0.9490	0.9490	0.9490	0.9450
0.5	Longitud	0.2952	0.2970	0.2956	0.2972	0.2966	0.2966	0.2962	0.2962	0.2943
	Nivel	0.9500	0.9560	0.9540	0.9560	0.9560	0.9560	0.9560	0.9560	0.9560
0.6	Longitud	0.2530	0.2552	0.2544	0.2554	0.2549	0.2549	0.2546	0.2546	0.2518
	Nivel	0.9480	0.9530	0.9530	0.9530	0.9510	0.9510	0.9510	0.9510	0.9500
0.7	Longitud	0.2018	0.2041	0.2037	0.2042	0.2038	0.2038	0.2036	0.2036	0.2023
	Nivel	0.9420	0.9510	0.9490	0.9510	0.9510	0.9510	0.9510	0.9510	0.95500
0.8	Longitud	0.1420	0.1440	0.1441	0.1441	0.1438	0.1438	0.1436	0.1436	0.1441
	Nivel	0.9350	0.9380	0.9410	0.9390	0.9380	0.9380	0.9380	0.9380	0.9530
0.9	Longitud	0.0750	0.0761	0.0763	0.0762	0.0760	0.0760	0.0759	0.0759	0.0772
	Nivel	0.9380	0.9420	0.9420	0.9420	0.9410	0.9410	0.9410	0.9410	0.9460

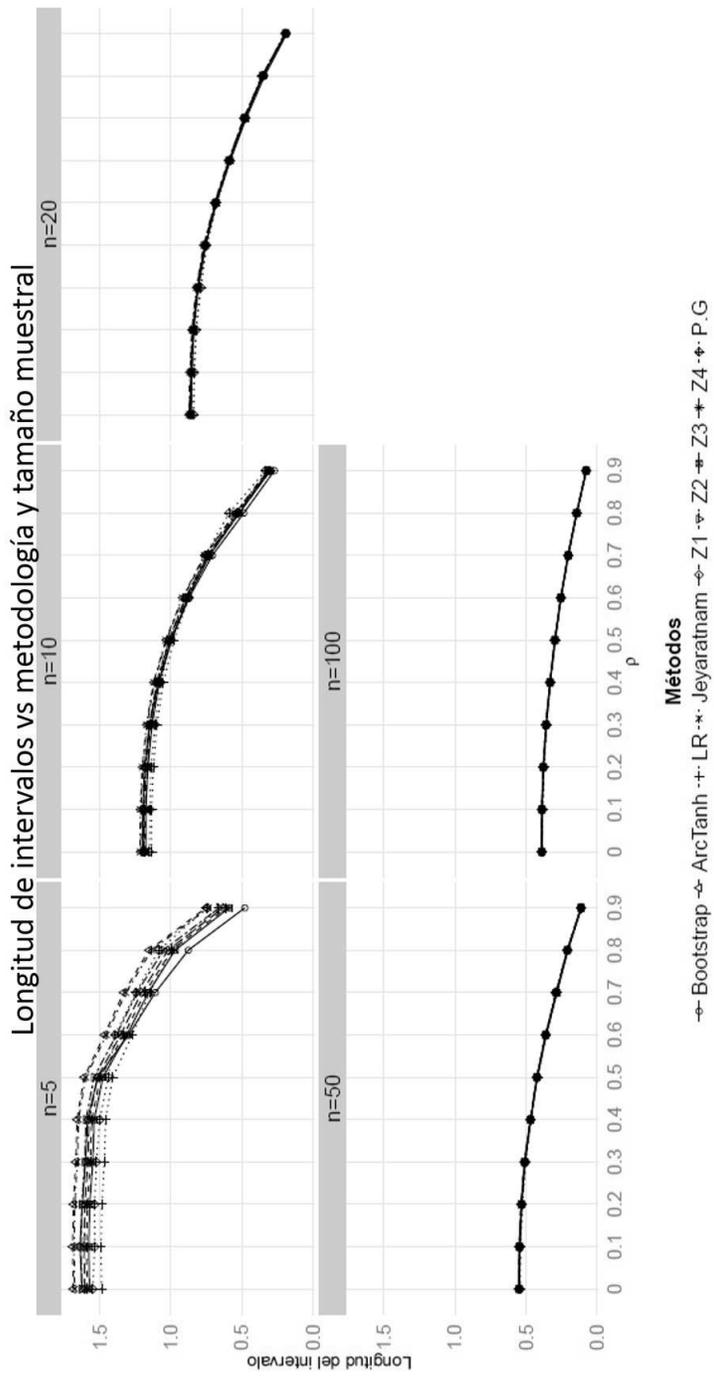


Figura 1: Amplitud por cada intervalo. Fuente: elaboración propia.

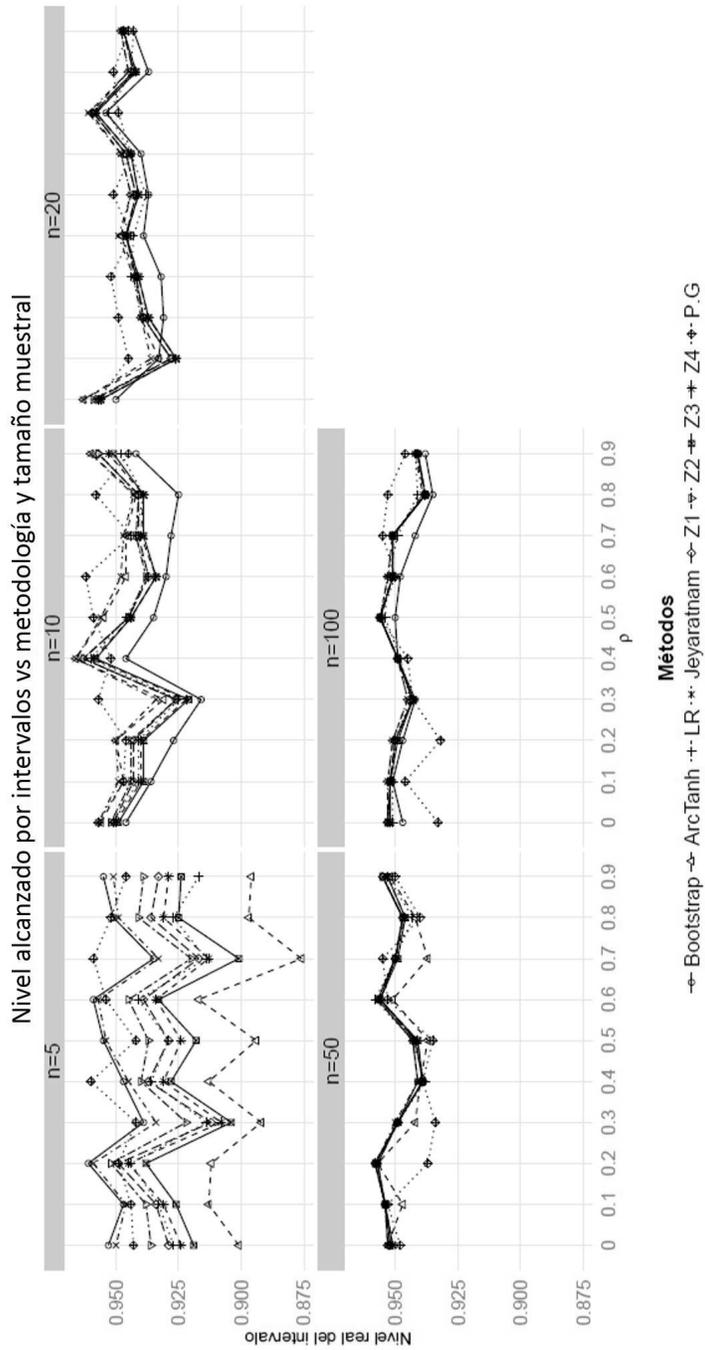


Figura 2: Nivel real alcanzado por cada intervalo. Fuente: elaboración propia.

Y además, cabe aclarar que el índice está limitado a 2 porque el rango parametral de ρ es desde -1 hasta 1, es decir, una longitud equivalente a dos unidades. Y por lo anterior el código de la simulación que se empleó en este estudio está restringido para no tener en cuenta posibles resultados para ρ mayores que 1, o en su defecto, menores que -1.

El rango de este índice corresponde a (0, 2.1052) dado que si el nivel real, NR, se acerca al 100% y/o la longitud máxima es lo más pequeña posible, entonces I será cercano a 2.1052; o si la longitud o el nivel real del intervalo es cercana a cero entonces el $I = 0$. Luego, a valores mayores del índice propuesto, mejor el intervalo obtenido. Las siguientes tablas muestran los resultados a partir del índice propuesto, donde se resalta por cada valor de ρ el mejor método:

Tabla 6: Índice de resumen: Tamaño de muestra 5. Fuente: elaboración propia
 $n = 5$

ρ	Bootstrap	ArcTanh	LR	Jayaratnam	Z1	Z2	Z3	Z4	P.G
0	0.3572	0.3145	0.5011	0.3234	0.3856	0.3677	0.4142	0.4007	0.4468
0.1	0.3467	0.3087	0.5003	0.3180	0.3832	0.3642	0.4123	0.3990	0.4624
0.2	0.3626	0.3187	0.5119	0.3280	0.3957	0.3758	0.4246	0.4117	0.4635
0.3	0.3752	0.3282	0.5126	0.3366	0.3996	0.3827	0.4294	0.4153	0.4724
0.4	0.3991	0.3391	0.5332	0.3488	0.4199	0.3993	0.4506	0.4355	0.5058
0.5	0.4572	0.3925	0.5787	0.4037	0.4735	0.4537	0.5051	0.4906	0.5555
0.6	0.6789	0.5386	0.7218	0.5517	0.6390	0.6146	0.6790	0.6588	0.6091
0.7	0.8212	0.6580	0.8185	0.6723	0.7662	0.7384	0.7987	0.7874	0.7632
0.8	1.0607	0.8481	0.9807	0.8637	0.9660	0.9388	1.0030	0.9866	0.9180
0.9	1.4347	1.2544	1.2922	1.2657	1.3440	1.3246	1.3717	1.3601	1.2497

Tabla 7: Índice de resumen: Tamaño de muestra 10. Fuente: elaboración propia
 $n = 10$

ρ	Bootstrap	ArcTanh	LR	Jayaratnam	Z1	Z2	Z3	Z4	P.G
0	0.8002	0.7948	0.8660	0.7911	0.8118	0.8088	0.8238	0.8218	0.8541
0.1	0.7921	0.7873	0.8568	0.7844	0.8049	0.8019	0.8150	0.8130	0.8517
0.2	0.7961	0.8044	0.8696	0.7998	0.8194	0.8164	0.8294	0.8274	0.8570
0.3	0.8201	0.8166	0.8794	0.8147	0.8324	0.8291	0.8416	0.8397	0.8960
0.4	0.9082	0.8987	0.9576	0.8941	0.9174	0.9135	0.9265	0.9245	0.9321
0.5	0.9881	0.9743	1.0186	0.9706	0.9850	0.9813	0.9975	0.9946	1.0097
0.6	1.1045	1.0816	1.1114	1.0794	1.0914	1.0889	1.1009	1.0981	1.1429
0.7	1.2628	1.2366	1.2562	1.2337	1.2488	1.2467	1.2581	1.2569	1.2342
0.8	1.4741	1.4450	1.4480	1.4431	1.4588	1.4560	1.4654	1.4634	1.4231
0.9	1.7147	1.7019	1.6778	1.6996	1.7069	1.7050	1.7026	1.7048	1.6571

De las anteriores tablas se observa que, según el criterio establecido para concluir con el índice de resumen propuesto, en tamaños de muestra pequeños y para correlaciones menores que 0.7 (dentro de las empleadas en este estudio), el mejor método para la construcción de intervalos de confianza para el coeficiente de correlación agrupando las características deseadas (longitud del intervalo corta y mayor porcentaje de cobertura) es el de la razón de verosimilitud. Le sigue el método Bootstrap en calidad. El método menos eficiente es el de la Transformación de Fisher. Esto es apenas lógico de esperar ya que establece que z tiene una

Tabla 8: Índice de resumen: Tamaño de muestra 20. Fuente: elaboración propia

$n = 20$									
ρ	Bootstrap	ArcTanh	LR	Jayaratnam	Z1	Z2	Z3	Z4	P.G
0	1.1429	1.1458	1.1665	1.1434	1.1463	1.1468	1.1492	1.1487	1.1690
0.1	1.1270	1.1165	1.1357	1.1177	1.1168	1.1161	1.1194	1.1189	1.1587
0.2	1.1352	1.1386	1.1617	1.1363	1.1437	1.1430	1.1462	1.1458	1.1778
0.3	1.1673	1.1727	1.1996	1.1716	1.1787	1.1781	1.1824	1.1819	1.2155
0.4	1.2304	1.2315	1.2456	1.2305	1.2348	1.2341	1.2396	1.239160	1.2421
0.5	1.3004	1.3017	1.3088	1.2996	1.3045	1.3039	1.3074	1.3070	1.3136
0.6	1.4036	1.4034	1.4073	1.4015	1.4054	1.4049	1.4065	1.406163	1.4040
0.7	1.5373	1.5309	1.5243	1.5309	1.5337	1.5332	1.5356	1.535321	1.5174
0.8	1.6381	1.6348	1.6332	1.6335	1.6347	1.6343	1.6356	1.6353	1.6468
0.9	1.8042	1.8026	1.7909	1.8019	1.8027	1.8025	1.8042	1.8041	1.7973

Tabla 9: Índice de resumen: Tamaño de muestra 50. Fuente: elaboración propia

$n = 50$									
ρ	Bootstrap	ArcTanh	LR	Jayaratnam	Z1	Z2	Z3	Z4	P.G
0	1.4584	1.4526	1.4574	1.4519	1.4526	1.4525	1.4524	1.4524	1.4561
0.1	1.4532	1.4578	1.4655	1.4571	1.4593	1.4592	1.4606	1.4606	1.4657
0.2	1.4813	1.4780	1.4852	1.4773	1.4794	1.4793	1.4792	1.4792	1.4527
0.3	1.4799	1.4870	1.4934	1.4880	1.4884	1.4884	1.4897	1.4897	1.4737
0.4	1.5178	1.5159	1.5160	1.5153	1.5139	1.5139	1.5151	1.5150	1.5136
0.5	1.5572	1.5643	1.5680	1.5638	1.5639	1.5638	1.5633	1.5632	1.5563
0.6	1.6441	1.6497	1.6537	1.6493	1.6491	1.6490	1.6500	1.6500	1.6457
0.7	1.6945	1.7131	1.7138	1.7127	1.7139	1.7139	1.7128	1.7146	1.7147
0.8	1.7801	1.7847	1.7769	1.7844	1.7853	1.7853	1.7840	1.7839	1.7734
0.9	1.8944	1.8973	1.8927	1.8971	1.8976	1.8976	1.8939	1.8939	1.8873

Tabla 10: Índice de resumen: Tamaño de muestra 100. Fuente: elaboración propia

$n = 100$									
ρ	Bootstrap	ArcTanh	LR	Jayaratnam	Z1	Z2	Z3	Z4	P.G
0	1.6075	1.6138	1.6134	1.6136	1.6143	1.6143	1.6131	1.6131	1.5853
0.1	1.6171	1.6163	1.6158	1.6161	1.6151	1.6151	1.6156	1.6156	1.6107
0.2	1.6185	1.6229	1.6240	1.6227	1.6217	1.6217	1.6204	1.6222	1.5956
0.3	1.6303	1.6318	1.6325	1.6317	1.6289	1.6289	1.6293	1.6293	1.6324
0.4	1.6685	1.6668	1.6669	1.6666	1.6673	1.6672	1.6677	1.6677	1.6621
0.5	1.7048	1.7137	1.7115	1.7135	1.7141	1.7141	1.7145	1.7145	1.7164
0.6	1.7433	1.7503	1.7511	1.7501	1.7469	1.7469	1.7472	1.7472	1.7482
0.7	1.7830	1.7977	1.7944	1.7976	1.7981	1.7980	1.7982	1.7982	1.8071
0.8	1.8286	1.8325	1.8383	1.8343	1.8327	1.8327	1.8329	1.8329	1.8616
0.9	1.9006	1.9076	1.9074	1.9075	1.9057	1.9057	1.9058	1.9058	1.9146

distribución aproximadamente normal cuando el tamaño muestral es grande. Este comportamiento cambia cuando se comienza a aumentar el tamaño de muestra, ya que cuando este es igual a 20, en correlaciones pequeñas (de 0 a 0.4), el mejor método para construir los intervalos es el que parte del pivote generalizado. Se observa que en el resto de las correlaciones algunos de los métodos también muy eficientes son el de la razón de verosimilitud, Bootstrap y la Transformación de Fisher modificada Z3. En este caso resulta complicado afirmar cuál método es el

menos efectivo, ya que las diferencias en el índice de resumen para el resto de los métodos son mínimas, y por tanto no muy significativas.

El método de la razón de verosimilitud en tamaños de muestra iguales a 50 dan los mejores resultados para estos intervalos de confianza y en menor medida el del pivote generalizado y la transformación de Fisher modificada Z1. Cabe aclarar que, las diferencias que hicieron que cada uno de estos métodos sobresaliese en el respectivo caso, fueron mínimas, es decir, con tamaños de muestra grandes, todos los métodos ofrecen índices de resumen casi iguales y por tanto, decidir cuál es mejor se hace indiferente.

4. Conclusiones

Los procedimientos que se tuvieron en cuenta para la construcción de intervalos de confianza para el coeficiente de correlación en una distribución normal bivariada, ρ , difieren en calidad dependiendo del tamaño de muestra empleado para ello. Se pudo observar que, en el caso de muestras pequeñas ($n = 5$ y $n = 10$), los que mejor desempeño tuvieron fueron los de razón de verosimilitud y Bootstrap. El método de razón de verosimilitud ofrece longitudes cortas de intervalo y nivel de confianza real más cercano al nivel nominal establecido en este estudio de simulación, 95 % para correlaciones desde 0 hasta 0.7. En los casos restantes, es decir, correlaciones de 0.8 y 0.9, el método Bootstrap supera al anterior.

Cuando $n = 20$, el método que se comporta mejor en la gran mayoría de las correlaciones consideradas es el de pivote generalizado. Le sigue en orden de calidad de intervalo obtenido por el método de razón de verosimilitud. Cuando $n = 50$ y $n = 100$, todos los métodos tienen un comportamiento casi igual. Las longitudes de intervalo para cada uno de estos dos casos ($n = 50$ y $n = 100$) disminuyen considerablemente a lo observado en el caso anterior, es decir, en comparación con $n = 20$, como se puede evidenciar en las tablas.

En la mayoría de los casos, el método de Bootstrap ofrece niveles de confianza reales menores que los de los demás métodos. Se observa un detalle muy particular con respecto a esta característica y todos los métodos a excepción del pivote generalizado: los niveles de confianza reales son muy parecidos entre sí cuando el tamaño de muestra es muy grande ($n=100$).

Recibido: 5 de mayo de 2013
Aceptado: 14 de agosto de 2013

Referencias

Efron, B. (1979), 'Computers and theory of statistics: Thinking the unthinkable', *SIAM Review* **21**(4), 460-480.

- Falk, R. & Well, A. (1997), 'Many faces of the correlation coefficient', *Journal of Statistics Education* **5**(3).
- Fisher, R. (1921), 'On the "probable error" of a coefficient of correlation deduced from a small sample', *Metron* **1**, 3–32.
- Hotelling, H. (1953), 'New light on the correlation coefficient and its transforms', *Journal of the Royal Statistical Society* **15**(2), 193–232.
- Kalbfleish, J. (1985), *Probability and statistical inference*, Springer-Verlag: New York.
- Krishnamoorthy, I. & Xia, Y. (2007), 'Inferences on correlation coefficients: One-sample, independent and correlated cases', *Journal of Statistical Planning and Inference* **137**(7), 2362–2379.
- Pawitan, Y. (2001), *In all likelihood*, Clarendon Press: Oxford.
- Zheng, Q. & Matis, J. (1994), 'Correlation coefficient revisited', *The American Statistician* **48**(3), 240–241.