
ESTADÍSTICA

**Confidence Interval for the Variance of a Normal
Distribution: “Equal Tails” Interval versus Optimal
Alternatives**

Jose M. Pavía

Departamento de Economía Aplicada
Universidad de Valencia

✉ pavia@uv.es

Abstract

When estimating the confidence interval for the variance of a normal distribution, textbooks weigh simplicity above optimality in selecting the solution. This fact, understandable in a pre-computational age, has been translated to almost all the statistical software. This work offers a very simple Microsoft Excel macro to obtain the most popular confidence interval and the two principal optimal solutions proposed in the literature, making easier the use of optimal solutions teaching and practising. A comparative analysis is also performed to assess the value of each one of the intervals. The 1937 Neyman’s confidence interval is the option that shows the best properties.

Keywords: Shortest interval, Minimum Ratio, Confidence level, Sample size.

AMS Subject classifications: 62F25, 62-01.

1. Introducción

La obtención de intervalos de confianza para parámetros desconocidos constituye uno de los problemas clásicos en estadística (el concepto fue introducido por Neyman en la década de los treinta del siglo XX; Neyman, 1937). La dificultad, sin embargo, reside en que el problema no presenta solución única; por lo que, utilizando para su obtención uno de los dos métodos estándar más habituales –el método de la cantidad pivotal o el método general de Neyman (e.g., Casas Sánchez, 1997)–, la elección del mejor intervalo se realiza en función de algún criterio considerado razonable: “mínima amplitud”, “insesgadez”, “invarianza”, “optimalidad para una amplitud prefijada” (ver, e.g., Guenther, 1971; Casella y Hwang, 1991; Juola, 1993) o, en última instancia, “simplicidad de cálculo”.

Un caso paradigmático, que se aborda en prácticamente casi cualquier curso introductorio de inferencia, lo constituye la construcción del intervalo de confianza para la varianza de una población con distribución normal. Como criterio general, en todos los libros de texto se recomienda emplear, para una confianza prefijada, el intervalo de menor amplitud (e.g., Escuder y Murgui, 1995; Newbold, 1996; Casas Sánchez, 1997). En el caso de la varianza, sin embargo, tal propuesta es abandonada y se opta por utilizar un criterio de simplicidad de cálculo. Por analogía con los intervalos de confianza de la media, se suele recomendar utilizar el intervalo con colas de igual probabilidad a pesar de saberse de antemano que no generará un intervalo óptimo bajo ningún criterio, al estar basado en una variable aleatoria con distribución asimétrica, χ^2 . Dada la gran potencia de cálculo disponible actualmente, sin embargo, no parece lo más adecuado conformarse con una solución que, de acuerdo con Tate y Klett (1959, p. 674), “*is far from satisfactory*”.

Si bien es cierto que (i) las tablas para construir manualmente cualquiera de los dos intervalos propuestos en la literatura como los mejores para la varianza de una normal sólo están disponibles para los niveles de confianza más habituales y en libros o artículos no fácilmente accesibles (e.g., Tate y Klett, 1959; Rohlf y Sokal, 1995) y que, además, (ii) muchos de los programas estadísticos no tienen programado el cálculo del intervalo de confianza para la varianza de una normal en sus opciones por defecto (BMDP, MINITAB, SPSS) y que los que la tienen programada (STATGRAPHICS, STATISTICA, SAS) únicamente ofrecen el intervalo de colas de igual probabilidad¹, en honor a la verdad hay que decir que ha sido posiblemente la comodidad de los formadores lo que explicaría que el uso de los intervalos óptimos no se haya generalizado en este caso concreto.

A fin de responder a la pregunta de qué intervalo utilizar en cada caso, en este trabajo se evalúan, para todos los tamaños muestrales de 2 a 100 y los niveles de confianza de 0'9000 a 0'9999, las diferencias de precisión que se obtendrían al utilizar cada uno de los intervalos. El trabajo, además, ofrece (ver Apéndices I y II) una macro en Excel (http://www.uv.es/~pavia/macro/Intervalo_Confianza_Varianza.xls) que permite obtener, para cualquier nivel de confianza, los tres intervalos de confianza más habituales para la varianza de una normal en función del tamaño y la varianza muestral.

El resto del documento está estructurado como sigue. En el apartado segundo se plantea el problema y se detallan los tres intervalos para la varianza que serán analizados. El apartado tercero muestra los resultados de las comparaciones y analiza las diferencias. Finalmente, la sección cuarta resume y concluye el documento. El código de la macro de Excel se ofrece en el Apéndice I, y una sencilla demostración de cómo funciona en el Apéndice II.

¹SAS es una excepción a esta norma, pues como segunda opción (aunque no por defecto), tiene programado la solución de Neyman (ver intervalo I_{MR} en la sección segunda).

2. Planteamiento del Problema

Sea $M = \{X_1, X_2, \dots, X_n\}$ una muestra aleatoria simple, de tamaño n , de una población con distribución de probabilidad $N(\mu, \sigma^2)$, con ambos parámetros μ y σ^2 desconocidos. Sea S^2 , dado por la ecuación (2.1), el estadístico varianza muestral de la muestra M , donde \bar{X} representa la media muestral.

$$S^2 = \frac{\sum_{j=1}^n (X_j - \bar{X})^2}{n}. \quad (2.1)$$

Fijado $1 - \alpha$ el nivel de confianza, es de sobra conocido que el intervalo de confianza con colas de igual probabilidad, I_{CI} , para σ^2 viene dado por:

$$I_{CI} = \left[\frac{nS^2}{\chi_{n-1, \frac{\alpha}{2}}^2}, \frac{nS^2}{\chi_{n-1, 1-\frac{\alpha}{2}}^2} \right], \quad (2.2)$$

donde $\chi_{n-1, p}^2$ es el valor que deja por encima de sí una probabilidad p en una distribución χ_{n-1}^2 (ji-cuadrado con $n - 1$ grados de libertad).

Como alternativa al intervalo anterior, la construcción del intervalo de menor amplitud similar a (2.2), I_{MA} , es fácilmente derivable resolviendo un sencillo problema de optimización con restricción (Tate y Klett, 1959, p. 675), cuya solución se obtiene resolviendo para a_n y b_n el sistema de ecuaciones (2.3):

$$\left. \begin{aligned} f_{n+3}(a_n) &= f_{n+3}(b_n) \\ \int_{a_n}^{b_n} f_{n-1}(x) dx &= 1 - \alpha \end{aligned} \right\}, \quad (2.3)$$

donde $f_k(x)$ es la función densidad de probabilidad de una χ_k^2 , con k un entero positivo, dada por (2.4):

$$f_k(x) = \frac{x^{\frac{k}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})}, \quad x \geq 0. \quad (2.4)$$

De modo que definiendo $\chi_{n-1, \alpha_2}^2 = b_n$ y $\chi_{n-1, \alpha_1}^2 = a_n$ (donde se verifica la igualdad: $\alpha_1 - \alpha_2 = 1 - \alpha$) se tiene que el intervalo I_{MA} (que hace mínima la distancia entre los inversos de a_n y b_n) viene dado por (2.5):

$$I_{MA} = \left[\frac{nS^2}{\chi_{n-1, \alpha_2}^2}, \frac{nS^2}{\chi_{n-1, \alpha_1}^2} \right], \quad (2.5)$$

El intervalo I_{MA} , sin embargo, presenta una importante limitación: no verifica la propiedad de invarianza frente a transformaciones biyectivas, propiedad que en el caso de la varianza es importante. En efecto, si este intervalo fuese utilizado para obtener un intervalo de confianza para la desviación típica poblacional (to-

mando raíces cuadradas a sus límites) el intervalo resultante no necesariamente retendría la cualidad de ser el de mínima amplitud.

La opción a los dos intervalos anteriores es construir el intervalo propuesto por Neyman (1937), quién después de definir el concepto de intervalo insesgado, sugiere que un intervalo de confianza para la varianza podría ser obtenido de la familia de regiones de aceptación del conjunto de contrastes bilaterales para la varianza poblacional ($H_0 : \sigma^2 = \sigma_0^2$). En particular, el intervalo asociado con la familia de test insesgados uniformemente más potentes asociados con los contrastes anteriores, I_{MR} , es uno del tipo (2.5), donde ahora los coeficientes a_n y b_n se obtienen resolviendo el sistema (2.6):

$$\left. \begin{aligned} f_{n+1}(a_n) &= f_{n+1}(b_n) \\ \int_{a_n}^{b_n} f_{n-1}(x) dx &= 1 - \alpha \end{aligned} \right\}, \quad (2.6)$$

Con lo que el intervalo I_{MR} resultante, dado por la ecuación (2.7), donde de nuevo se verifica que $\alpha'_1 - \alpha'_2 = 1 - \alpha$, es el que minimiza el ratio entre los extremos superior e inferior (Tate y Klett, 1959, p. 677). Es decir, es el intervalo de ratio mínimo o mínima amplitud relativa.

$$I_{MR} = \left[\frac{nS^2}{\chi_{n-1, \alpha'_2}^2}, \frac{nS^2}{\chi_{n-1, \alpha'_1}^2} \right]. \quad (2.7)$$

Comparado con las soluciones anteriores, este I_{MR} presenta la interesante propiedad de invarianza. De modo que, por ejemplo, definido un intervalo de la desviación típica como el intervalo cuyos límites son las raíces cuadradas de los límites de (2.7) se tendría que el intervalo resultante continuaría siendo el intervalo para la desviación típica con mínimo ratio entre los intervalos de igual confianza.

3. Comparación entre intervalos

Cada uno de los intervalos propuestos en el apartado anterior presenta una serie de virtudes que lo hacen deseable frente a sus competidores. Sin embargo, dado que cada uno de ellos generaría una solución diferente (que no obstante irían convergiendo a medida que el tamaño muestral tendiera a infinito² en virtud de la convergencia de la distribución χ^2 a la Normal) en cada problema concreto, cabría preguntarse por cuales serían los costes de utilizar uno u otro intervalo. Es decir, sería pertinente plantearse preguntas tales como: ¿cuánto más extenso es el intervalo I_{CI} comparado con las dos alternativas óptimas? o, evaluados los

²Por ejemplo, para $n = 26$, $1 - \alpha = 95\%$ y $s^2 = 100$, los intervalos serían respectivamente: $I_{CI} = [63'97, 198'17]$, $I_{MA} = [56'86, 182'25]$, $I_{MR} = [62'43, 192'43]$; mientras que para $n = 101$, serían respectivamente: $I_{CI} = [77'96, 136'08]$, $I_{MA} = [75'77, 132'88]$, $I_{MR} = [77'47, 135'14]$.

intervalos I_{MA} y I_{MR} con los criterios de optimalidad de su competidor ¿cuál sería el coste de utilizar I_{MA} (I_{MR}) frente a I_{MR} (I_{MA})?

A fin de dar respuestas a tales cuestiones se han definido dos estadísticos que cuantifican el porcentaje de pérdida que supone utilizar, frente a la alternativa óptima, cada uno de los dos intervalos alternativos. Es decir, si l_I^X y l_S^X denotan (para $X = CI, MA$ y MR), respectivamente, el límite inferior y superior del intervalo I_X –con I_X igual a cada uno de los intervalos I_{CI} , I_{MA} y I_{MR} , dados por las ecuaciones (2.2), (2.5) y (2.7), según corresponda–, los estadísticos definidos han sido: el coeficiente (3.1):

$$P_{MA}^X = \frac{l_S^X - l_I^X}{l_S^{MA} - l_I^{MA}} - 1, \quad (\text{con } X = CI, MR), \quad (3.1)$$

para comparar respecto al intervalo de mínima amplitud; y el coeficiente (3.2):

$$P_{MA}^X = \frac{l_S^X / l_I^X}{l_S^{MR} / l_I^{MR}} - 1, \quad (\text{con } X = CI, MA), \quad (3.2)$$

en las comparaciones con el intervalo de mínimo ratio.

La Figura 1 presenta de forma gráfica los valores de P_{MA}^{CI} para niveles de confianza que van desde el 90'00 al 99'99 por ciento y tamaños muestrales de 2 a 100; mientras que la Tabla 1 ofrece, para los niveles de confianza más habituales y tamaños muestrales de 6 a 42, los valores de este estadístico en términos porcentuales. El análisis de los resultados de este estadístico revela que el tamaño muestral es el factor que más influye a la hora de valorar la calidad relativa de ambos intervalos, si bien un análisis detallado de los datos también revela que las longitudes de ambos intervalos van convergiendo a medida que el nivel de confianza se incrementa. En efecto, si para $n = 2$ se obtienen intervalos con una amplitud de hasta tres veces superiores, las diferencias son inferiores a un 5% tan pronto como el tamaño muestral supera las 37 observaciones, ver Tabla 1.

La Figura 2 muestra similares relaciones a las presentadas en la Figura 1, pero esta vez respecto al intervalo de ratio mínimo. La primera cuestión que llama la atención es la enorme reducción de las diferencias que se observa respecto al caso anterior, que en el caso más desfavorable ($n = 2, 1 - \alpha = 0'90$) apenas supone un intervalo con una longitud un 30% superior. De nuevo el tamaño muestral se revela como el factor con mayor capacidad explicativa de las diferencias. En este caso, no obstante, las diferencias de tamaño entre ambos intervalos convergen más rápidamente a medida que crece el nivel de confianza. Igualmente, son necesarios menores tamaños muestrales para alcanzar diferencias inferiores al 5% entre la longitud de ambos intervalos. Con tamaños de incluso menos de 20 observaciones la diferencia se reduce por debajo de la cota del 5%, ver Tabla 1.

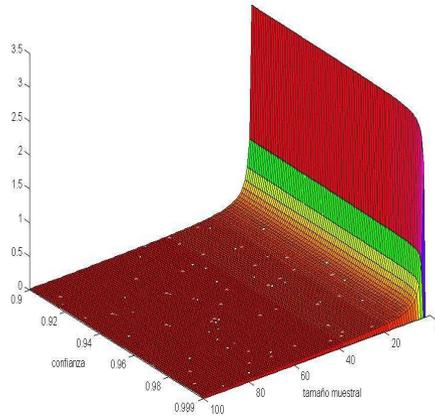


Figura 1: Relación entre los intervalos de colas de igual probabilidad y mínima amplitud a través de PCI_{MA} .

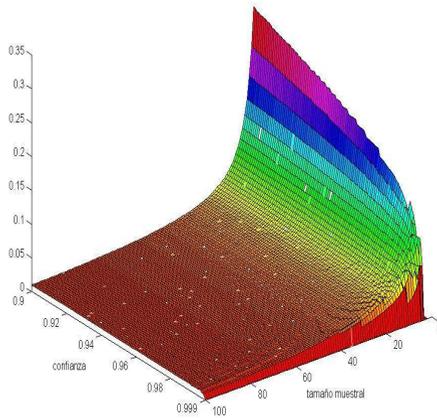


Figura 2: Relación entre los intervalos de ratio mínimo y mínima amplitud a través de P_{MA}^{MR} .

Tabla 1: Porcentaje de incremento en amplitud (para los niveles de confianza más habituales) de los intervalos con colas de igual probabilidad y mínimo ratio respecto al de mínima amplitud.

Tamaño muestral	Colas igual probabilidad			Ratio mínimo		
	0.90	0.95	0.99	0.90	0.95	0.99
6	36.22	35.25	33.61	14.49	13.11	10.87
7	29.95	29.18	27.88	12.73	11.65	9.98
8	25.55	24.93	23.87	11.31	10.45	8.87
9	22.30	21.79	20.86	10.19	9.49	8.01
10	19.79	19.36	18.58	9.26	8.68	7.79
11	17.79	17.43	16.75	8.45	7.97	7.25
12	16.16	15.85	15.26	7.81	7.41	6.48
13	14.80	14.54	14.02	7.25	6.89	6.17
14	13.66	13.43	12.97	6.76	6.45	5.91
15	12.68	12.48	12.07	6.34	6.09	5.41
16	11.83	11.65	11.28	5.96	5.72	5.25
17	11.09	10.93	10.60	5.63	5.40	5.13
18	10.44	10.30	9.99	5.35	5.13	4.77
19	9.86	9.73	9.45	5.07	4.89	4.44
20	9.34	9.22	8.97	4.83	4.69	4.40
21	8.87	8.77	8.54	4.59	4.45	4.13
22	8.45	8.35	8.14	4.40	4.29	3.89
23	8.07	7.98	7.78	4.20	4.10	3.89
24	7.72	7.63	7.45	4.06	3.92	3.68
25	7.39	7.32	7.15	3.90	3.77	3.49
26	7.10	7.03	6.87	3.75	3.68	3.53
27	6.83	6.76	6.61	3.60	3.55	3.36
28	6.57	6.51	6.37	3.48	3.43	3.20
29	6.34	6.28	6.15	3.37	3.27	3.06
30	6.12	6.07	5.95	3.25	3.17	3.12
31	5.92	5.87	5.75	3.16	3.08	2.99
32	5.73	5.68	5.57	3.05	3.00	2.87
33	5.55	5.50	5.40	2.98	2.92	2.76
34	5.38	5.34	5.24	2.88	2.81	2.65
35	5.22	5.18	5.09	2.80	2.74	2.55
36	5.07	5.03	4.95	2.72	2.68	2.63
37	4.93	4.90	4.81	2.64	2.59	2.54
38	4.80	4.76	4.69	2.59	2.54	2.45
39	4.67	4.64	4.57	2.53	2.49	2.37
40	4.55	4.52	4.45	2.47	2.41	2.29
41	4.44	4.41	4.34	2.41	2.37	2.22
42	4.33	4.30	4.24	2.33	2.30	2.15

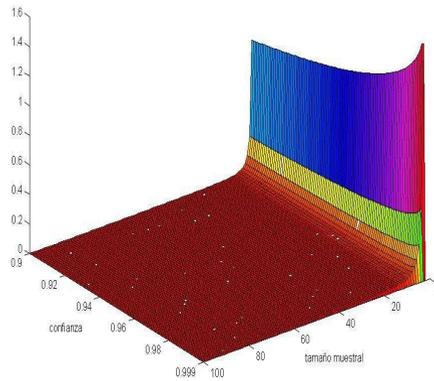


Figura 3: Relación entre los intervalos de colas de igual probabilidad y ratio mínimo a través de PCI_{MR} .

Las Figuras 3 y 4 y la Tabla 2, por su parte, muestran los resultados de las comparaciones llevadas a cabo a través de la ecuación (3.2), donde se comparan los intervalos ICI y IMA con el intervalo relativamente más pequeño o de ratio mínimo, IMR . Centrándonos en la Figura 3, se observa que, al igual que ocurría en los casos anteriores, las diferencias entre las amplitudes relativas de los intervalos se van reduciendo a medida que se incrementan los tamaños muestrales. En este caso, no obstante, destaca el hecho de que las diferencias van agrandándose a medida que se incrementa el nivel de confianza (aunque para niveles de confianza superiores al 99'97% las diferencias son prácticamente inexistentes), las cuales, no obstante, pronto se reducen a niveles por debajo del 5% (ver Tabla 2). Con tamaños muestrales de apenas 10 observaciones el tamaño relativo del intervalo de colas de igual probabilidad es a lo sumo un 5% más grande que el intervalo de ratio mínimo.

Por último, el análisis de la Figura 4 revela patrones de comportamiento similares a los encontrados en los otros casos para el tamaño muestral; mientras que se observa un comportamiento similar al revelado por la Figura 2 en relación al nivel de confianza. En este caso, sin embargo, destaca (de forma análoga a la Figura 3, aunque con mayor intensidad) que la ventaja relativa del intervalo IMR respecto al intervalo IMA tiende a hacerse más visible a medida que el nivel de confianza se va aproximando al máximo, ver también Tabla 2. Ello provoca que sean necesarios tamaños muestrales de hasta 33 observaciones para que las diferencias en términos del estadístico se sitúen por debajo del 5%. De hecho, fijado un nivel de confianza del 99%, son necesarias hasta 27 observaciones para conseguir reducir las diferencias por debajo de la cota del 5%.

Tabla 2: Porcentaje de incremento en amplitud relativa (para los niveles de confianza más habituales) de los intervalos con colas de igual probabilidad y mínima amplitud respecto al de mínimo ratio.

Tamaño muestral	Colas igual probabilidad			Mínima amplitud		
	0.90	0.95	0.99	0.90	0.95	0.99
6	5.23	6.29	8.40	29.16	31.10	24.31
7	3.90	4.68	6.25	23.37	25.42	24.11
8	3.05	3.66	4.88	19.10	21.58	23.79
9	2.47	2.96	3.94	16.26	18.03	23.40
10	2.05	2.46	3.27	13.84	15.43	17.01
11	1.74	2.08	2.77	12.09	13.40	16.79
12	1.50	1.79	2.38	10.64	12.12	13.34
13	1.31	1.57	2.08	9.42	10.72	13.19
14	1.16	1.38	1.84	8.48	9.54	10.91
15	1.03	1.23	1.63	7.65	8.76	10.81
16	0.93	1.11	1.47	7.00	7.87	9.15
17	0.84	1.00	1.33	6.35	7.26	9.07
18	0.76	0.91	1.21	5.83	6.71	7.81
19	0.70	0.84	1.11	5.41	6.21	7.74
20	0.64	0.77	1.02	5.03	5.76	6.73
21	0.60	0.71	0.94	4.68	5.35	6.69
22	0.55	0.66	0.87	4.36	5.08	5.86
23	0.51	0.62	0.81	4.10	4.73	5.82
24	0.48	0.57	0.76	3.82	4.40	5.13
25	0.45	0.54	0.71	3.60	4.19	5.10
26	0.42	0.51	0.67	3.43	3.99	5.07
27	0.40	0.48	0.63	3.23	3.72	4.49
28	0.38	0.45	0.59	3.04	3.54	4.47
29	0.36	0.43	0.56	2.90	3.37	3.97
30	0.34	0.40	0.53	2.76	3.21	3.95
31	0.32	0.38	0.51	2.63	3.06	3.93
32	0.31	0.36	0.48	2.50	2.92	3.50
33	0.29	0.35	0.46	2.41	2.78	3.48
34	0.28	0.33	0.44	2.29	2.65	3.47
35	0.27	0.32	0.42	2.21	2.59	3.09
36	0.25	0.30	0.40	2.10	2.47	3.08
37	0.24	0.29	0.38	2.02	2.35	3.07
38	0.23	0.28	0.37	1.95	2.29	2.74
39	0.22	0.27	0.35	1.88	2.18	2.73
40	0.22	0.26	0.34	1.80	2.13	2.72
41	0.21	0.25	0.33	1.74	2.03	2.43
42	0.20	0.24	0.31	1.67	1.98	2.42

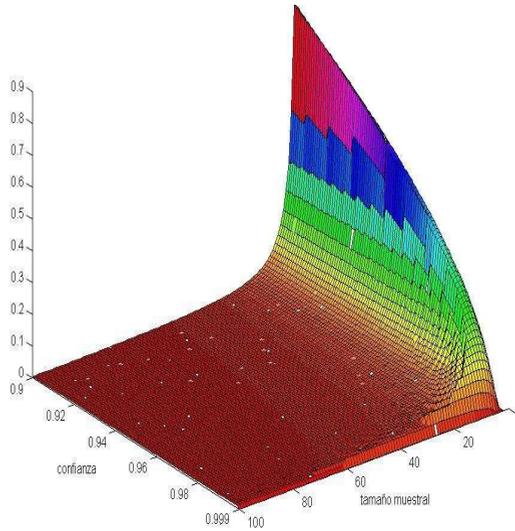


Figura 4: Relación entre los intervalos de mínima amplitud y ratio mínimo a través de P_{MR}^{MA} .

4. Conclusiones

En el cálculo del intervalo de confianza para la varianza de una distribución normal, el intervalo recomendado en los libros de texto no es el óptimo. Dos alternativas principales han sido propuestas en la literatura para mejorar (reduciendo la amplitud absoluta o relativa) el intervalo de colas con igual probabilidad, el cual ha sido, por su simplicidad de cálculo, el empleado sistemáticamente en el aula y, por extensión, en las aplicaciones prácticas. En este trabajo se ofrece una sencilla macro de Excel que puede fácilmente ser utilizada por profesores e investigadores y permite obtener, junto con el intervalo más popular, los intervalos de mínima amplitud y mínimo ratio. El análisis comparativo de los tres intervalos revela que el intervalo de ratio mínimo es el que debería ser potenciado, especialmente con tamaños muestrales pequeños (de hasta 40 observaciones). En efecto, a sus excelentes propiedades teóricas (insesgadez e invarianza, de las que carecen sus competidores) se suma el hecho de que es el que presenta mejores figuras en las comparaciones numéricas. Por una parte, su amplitud absoluta nunca supera en más de un 30% al de menor longitud, presentando rápidamente figuras comparables tan pronto como se registra un mínimo crecimiento en el tamaño muestral; mientras, por otra parte, su ventaja en términos de menor amplitud relativa presenta figuras mucho más elevadas en las comparaciones con ambos intervalos: el de mínima amplitud absoluta y el de colas con igual probabilidad. No obstante lo anterior, cuando el tamaño muestral crece significativamente cualquiera de las tres alternativas generaría intervalos aceptables

en términos de tamaño absoluto y relativo: como norma se podría afirmar que las tres opciones proporcionan intervalos con garantías similares a partir de 40 observaciones. Empero, las propiedades de insesgadez, invarianza y mínimo ratio sólo las cumpliría la solución propuesta por Neyman.

Agradecimientos

Deseo agradecer a Francisco Morillas su inestimable ayuda y al editor Miguel López Díaz los valiosos comentarios realizados.

Referencias

- [1] Casas Sánchez J.M. (1997). *Inferencia Estadística*, Ed. Centro de Estudios Ramón Areces, Madrid (España).
- [2] Casella G. y Hwang T.H. (1991) Evaluating Confidence Sets Using Loss Functions. *Stat. Sinica*, **1**, 159-173.
- [3] Escuder Vallés R., y Murgui Izquierdo J.S. (1995). *Estadística Aplicada. Economía y Ciencias Sociales*, Ed. Tirant lo Blanch, Valencia (España).
- [4] Guenther W.C. (1971) Unbiased Confidence Intervals. *Am. Stat.*, **25**, 51-53.
- [5] Juola R.C. (1993) More on Shortest Confidence Intervals *Am. Stat.*, **47**, 117-119.
- [6] Newbold P. (1996). *Estadística para los Negocios y la Economía* Ed. Prentice Hall, Madrid (España).
- [7] Neyman J. (1937) Outline of a Theory of Statistical Estimation based on the Classical Theory of Probability *Philosophical Transactions of the Royal Society A*, **236**, 333-380.
- [8] Rohlf F. J., y Sokal R. R. (1995). *Statistical Tables*, Ed. Freeman, San Francisco (Estados Unidos).
- [9] Tate R.F., y Klett G.W. (1959) Optimal Confidence Intervals for the Variance of a Normal Distribution. *J. Am. Stat. Assoc.*, **54**, 674-682.

Apéndice I: Macro en Excel para el cálculo de los intervalos con colas con igual probabilidad, de mínima amplitud y mínimo ratio.

```

Sub IC_VARIANZA_NORMAL()
,
' Macro para calcular los diferentes intervalos de confianza a nivel l_alfa
' para la varianza de una normal (de media desconocida) basadas en el metodo
' pivotal de construccion de intervalos:
' 1. El clasico o de dos colas: (1-l_alfa/2)/2 de probabilidad en cada cola.
' 2. El de minima amplitud relativa. Cociente minimo Ls/Li, entre limite
' superior, Ls, e inferior, Li.
' 3. El de minima distancia entre extremos: Ls-Li.
' Macro grabada por ---
,
Dim l_alfa As Variant
    l_alfa = InputBox(prompt:="INTRODUZCA EL NIVEL DE CONFIANZA DEL INTERVALO
        Valor entre 0 y 100")
If l_alfa <= 0 Or l_alfa >= 100 Then
    MsgBox ("ERROR: El nivel de confianza ha de estar entre 0 y 100")
    GoTo 100
End If
Dim n As Variant
    n = InputBox(prompt:="INTRODUZCA EL TAMAÑO DE LA MUESTRA. Numero entero
        mayor que uno")
dif = n - Int(n)
If dif > 0 Then
    MsgBox ("ERROR: El tamaño muestral ha de ser un numero entero")
    GoTo 100
End If
If n < 2 Then
    MsgBox ("ERROR: El tamaño muestral ha de ser al menos dos")
    GoTo 100
End If
Dim S As Variant
    S = InputBox(prompt:="INTRODUZCA LA VARIANZA MUESTRAL. Numero real positivo")
If S < 0 Then
    MsgBox ("ERROR: La varianza no puede ser negativa")
    GoTo 100
End If
Range("A11").Select
    ActiveCell.FormulaR1C1 = "Nivel de Confianza"
Range("A12").Select
    ActiveCell.FormulaR1C1 = "Tamaño muestral"
Range("A13").Select
    ActiveCell.FormulaR1C1 = "Varianza Muestral"

```

```

Cells(11, 3) = l_alfa / 100
Cells(12, 3) = n
Cells(13, 3) = S
Range("C14").Select
    ActiveCell.FormulaR1C1 = "=CHIINV((1-R[-3]C)/2,R[-2]C-1)"
Range("D14").Select
    ActiveCell.FormulaR1C1 = "=CHIINV(1-(1-R[-3]C[-1])/2,R[-2]C[-1]-1)"
ls1 = Cells(14, 4).Value
li1 = Cells(14, 3).Value
li1I = n * S / li1
ls1I = n * S / ls1
Range("C15").Select
    ActiveCell.FormulaR1C1 = "L.Inferior"
Range("D15").Select
    ActiveCell.FormulaR1C1 = "L.Superior"
Range("A16").Select
    ActiveCell.FormulaR1C1 = "Int.Clasico, de dos colas"
Cells(16, 3) = li1I
Cells(16, 4) = ls1I
Cells(16, 5) = (1 - l_alfa / 100) / 2
Cells(16, 6) = (1 - l_alfa / 100) / 2
Range("A11:A13,A16,C15:D15").Select
    Selection.Font.Bold = True
Range("C11,C13,C16,D16").Select
    Selection.NumberFormat = "0.00"
Range("C11").Select
    Selection.NumberFormat = "0.00%"
coc = ls1I / li1I
dif = ls1I - li1I
d = 0.0001
Range("C16:F16").Select
Selection.Copy
Range("C17:C18").Select
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
    :=False, Transpose:=False
While Cells(11, 4).Value < 1
    Cells(11, 4) = Cells(11, 3) + d
    Cells(11, 5) = Cells(11, 4) - Cells(11, 3)
    Range("C14").Select
    ActiveCell.FormulaR1C1 = "=CHIINV(R[-3]C[+1],R[-2]C-1)"
    Range("D14").Select
    ActiveCell.FormulaR1C1 = "=CHIINV(R[-3]C[+1],R[-2]C[-1]-1)"
    li1 = Cells(14, 4).Value
    ls1 = Cells(14, 3).Value
    li1I = n * S / li1
    ls1I = n * S / ls1

```

```

coc2 = ls1I / li1I
dif2 = ls1I - li1I
If dif2 < dif Then
    dif = dif2
    Cells(17, 3) = li1I
    Cells(17, 4) = ls1I
    Cells(17, 5) = Cells(11, 5)
    Cells(17, 6) = 1 - Cells(11, 4)
End If
If coc2 < coc Then
    coc = coc2
    Cells(18, 3) = li1I
    Cells(18, 4) = ls1I
    Cells(18, 5) = Cells(11, 5)
    Cells(18, 6) = 1 - Cells(11, 4)
End If
d = d + 0.0001
Wend
Range("A17").Select
    ActiveCell.FormulaR1C1 = "Int. de Distancia minima"
Range("A18").Select
    ActiveCell.FormulaR1C1 = "Int. de Ratio minimo"
Range("E15").Select
    ActiveCell.FormulaR1C1 = "Prob_inferior"
Range("F15").Select
    ActiveCell.FormulaR1C1 = "Prob_superior"
Range("A11:A13,A16:A18,C15:F15").Select
    Selection.Font.Bold = True
Range("C11,C13,C16,D16").Select
    Selection.NumberFormat = "0.00"
Range("E16:F18").Select
    Selection.NumberFormat = "0.0000"
Range("C11").Select
    Selection.NumberFormat = "0.00%"
Range("D11:E11,C14:D14").Select
    Selection.ClearContents
Range("C16:F18").Select
    Application.CutCopyMode = False
With Selection
    .HorizontalAlignment = xlCenter
    .VerticalAlignment = xlBottom
    .WrapText = False
    .Orientation = 0
    .AddIndent = False
    .IndentLevel = 0
    .ShrinkToFit = False

```

```
.ReadingOrder = xlContext
.MergeCells = False
End With
Range("C11").Select
100

End Sub
```

Apéndice II: Calculando un intervalo de confianza con la macro

Suponga que dispone de una muestra de una población normal de tamaño 23 para la que ha obtenido una varianza muestral de 96 y que desea calcular los intervalos de confianza al 95 % para la varianza de las ecuaciones (2.2), (2.5) y (2.7) a través de la macro propuesta en el Apéndice I.

(1) Abra el archivo Intervalo_Confianza_Varianza.xls (http://www.uv.es/pavia/macro/Intervalo_Confianza_Varianza.xls). Le aparecerá una pantalla como la de la Figura A.1.



Figura A.1. Imagen de la hoja de cálculo previa a la ejecución.

(2) Pulse el botón “CALCULO INTERVALOS DE CONFIANZA PARA LA VARIANZA DE UNA NORMAL” (o a través de Herramientas/Macro/Ejecutar). Aparecerán sucesivamente los tres cuadros de diálogo de la Figura A.2. Vaya introduciendo los valores deseados y pulse aceptar.



Figura A.2. Cuadros de diálogo del programa.

(3) Tras lo cual y después de unos segundos de cálculo, aparecerá la siguiente solución (ver Figura A.3) en la hoja Excel de trabajo:

Como se observa en la Figura A.3, además de los límites superior e inferior de cada intervalo, el programa también ofrece las probabilidades que quedan a cada lado del intervalo.

	A	B	C	D	E	F	G	H	I	J	K		
1	CALCULO INTERVALOS DE CONFIANZA PARA LA VARIANZA DE UNA NORMAL												
2													
3													
4													
5													
6													
7													
8													
9													
10													
11	Nivel de Confianza		95.00%										
12	Tamaño muestral		23										
13	Varianza Muestral		96.00										
14													
15			L.Inferior	L.Superior	Prob. inferior	Prob. superior							
16	Int. Clasico, de dos colas		60.03	201.05	0.0250	0.0250							
17	Int. de Distancia mínima		52.53	183.14	0.0062	0.0438							
18	Int. de Ratio mínimo		58.38	194.34	0.0192	0.0308							

Figura A.3. Pantalla con soluciones.

Acerca del autor

Jose M. Pavía es profesor de Estadística en la Universidad de Valencia. En los últimos años sus intereses de investigación han estado centrados en el análisis de series temporales, la elaboración de contabilidades trimestrales, la economía regional, el estudio de los ciclos económicos, las predicciones electorales y los estudios demográficos. Sus resultados de investigación han sido publicados en revistas como *Journal of the American Statistical Association*, *International Journal of Forecasting*, *Environment and Planning A*, *Journal of Forecasting*, *Journal of Applied Statistics*, *Spanish Economic Review*, *Estadística Española* o *Investigaciones Económicas*, entre otras.