Deducción Inédita De La Varianza Del Estadístico De Wilcoxon

Álvaro de Jesús Villota Viveros, vvvillota@live.com, Academia Nacional de Medicina

Resumen. Este trabajo realiza una deducción inédita al valor de la varianza del Estadístico de Wilcoxon. Inicialmente determina las varianzas de todos los valores posibles del estadístico para pocos pares de observaciones, posteriormente aplica el método de las diferencias finitas en sucesiones para obtener la varianza en función de n pares.

Palabras clave. Estadístico de Wilcoxon, Varianza, Método de las diferencias finitas para sucesiones.

1. Presentación del problema.

Realizar una deducción novedosa de la varianza del Estadístico de Wilcoxon.

2. Marco de referencia conceptual.

La prueba de los rangos con signo de Wilcoxon es un estudio no paramétricos que sin importar la distribución muestral permite comparar dos grupos y establecer si existen diferencias entre ellas. Es una alternativa importante a la prueba t de Student cuando no se puede establecer distribución normal muestral. Se emplea en variables continuas.

La prueba requiere encontrar para cada caso el Estadístico de Wilcoxon.

El estadístico se establece de la siguiente manera:

Supóngase que cada uno de los individuos en un estudio, fueron sometidos a un tratamiento y se anotaron los valores de la variable antes y después. Para cada individuo se realiza la diferencia de dichos valores y se consigna el resultado en valor absoluto, luego frente al individuo se le asigna el número del lugar en orden ascendente considerando la última columna. Finalmente al número asignado se le devuelve el signo de la diferencia, véase ejemplo consignado en el siguiente cuadro.

Antes	Después	Diferencia Valor absoluto	Número asignado	Número asignado conservando signo
46	48	2	2	-2
75	70	5	5	5
66	65	1	1	1
70	76	6	6	-6

43	40	3	3	3
50	46	4	4	4
55	63	8	7	-7
62	50	12	8	8
61	74	13	9	-9
56	70	14	10	-10

Hay dos modalidades en el Estadístico de Wilcoxon (W).

La primera consiste en sumar algebraicamente todos los valores de la columna que contiene los números asignados conservando el signo. En el caso del ejemplo sería W=-13.

La varianza para el estadístico obtenido de esta manera es:

$$Var\,(W) = \frac{n\,(n+1)\,(2n+1)}{6}$$

La segunda consiste en sumar algebraicamente solo los valores positivos presentes en la columna que contiene los números asignados conservando el signo. En el caso del ejemplo sería W=21.

La varianza para el estadístico obtenido de esta manera es:

$$Var(W) = \frac{n\,(n+1)\,(2n+1)}{24}$$

3. Metodología

Aunque la prueba de Wilcoxon está diseñada para 10 o más parejas, para deducir varianza supóngase inicialmente que solo hay una.

Para ella los valores posibles son:

1
-1

$$Var\,(W) = 1$$

Ahora asúmase que se tienen dos parejas.
Los valores posibles son:

2	+1	= 3
2	-1	= 1

-2	+1	= -1
-2	-1	= -3

$Var\,(W) = 5$

Si se tienen tres parejas, los valores posibles son:

3	+2	+1	= 6
3	+2	-1	= 4
3	-2	+1	= 2
3	-2	-1	= 0
-3	+2	+1	= 0
-3	+2	-1	=-2
-3	-2	+1	=-4
-3	-2	-2	=-6

$Var\,(W) = 14$

Si se continúa igual procedimiento,

Para cuatro parejas se obtiene:
$Var\,(W) = 30$

Para cinco parejas se obtiene:
$Var\,(W) = 55$

Los valores obtenidos para las varianzas de 1,2,3,4 y 5 parejas, son suficientes para deducir la expresión que define en general Var(W), en función del número de parejas.

Se empleará el método de las diferencias finitas en sucesiones, para ello:

1 5 14 30 55

4 9 16 25

5 7 9

2 2

0

La expresión que permite calcular la varianza para n parejas será:

$$Var\left(W\right)=\sum_{i=1}^{n}\binom{n-1}{i-1}a_i$$

Donde los valores de a, son en su orden los correspondientes a la primera diagonal.

En este caso en particular:

$a_1=1$
$a_2=4$
$a_3=5$
$a_4=2$

Los siguientes valores de a son cero.

$$Var\left(W\right)=\binom{n-1}{0}(1)+\binom{n-1}{1}(4)+\binom{n-1}{2}(5)+\binom{n-1}{3}(2)$$

$$Var\left(W\right)=1+(n-1)(4)+\frac{(n-1)(n-2)}{2!}(5)+\frac{(n-1)(n-2)(n-3)}{3!}(2)$$

Desarrollando y luego factorizando:

$$Var\left(W\right)=\frac{n\left(n+1\right)\left(2n+1\right)}{6}$$

También existe la modalidad en el estadístico de Wilcoxon sumando de la columna de “número asignado¨ conservando el signo solo los valores positivos.

En este caso si se tiene una sola pareja:

Los valores posibles para W son:

1
0

$$Var\left(W\right)=\frac{1}{4}$$

Si se tienen dos parejas los valores posibles para W son:

0		
1		
2		
1	+2	= 3

$$Var\left(W\right)=\frac{5}{4}$$

Si se tienen 3 parejas los valores serán:

0			
1			
2			
3			
1	+2		=3
1	+3		=4
2	+3		=5
1	+2	+3	=6

$$Var\left(W\right)=\frac{14}{4}$$

Para cuatro y cinco parejas, son similares reflexiones se obtienen:

$$Var\left(W\right)=\frac{30}{4}$$

$$Var\left(W\right)=\frac{55}{4}$$

Respectivamente.

Lo anterior significa que la varianza en esta modalidad de Estadístico de Wilcoxon es siempre la cuarta parte de la varianza encontrada anteriormente.

Por lo tanto bajo esta circunstancia:

$$Var(W)=\frac{n\left(n+1\right)\left(2n+1\right)}{24}$$

4. Análisis.

Mediante la adecuada aplicación del método de las diferencias finitas para sucesiones, se ha establecido la varianza del Estadístico de Wilcoxon para n pares de observaciones.

5. Conclusión.

Se presenta un procedimiento novedoso, sencillo y didáctico que permite a un estudiante o experto estadístico obtener matemáticamente la varianza del Estadístico de Wilcoxon.

6. Referencias Bibliográficas.

CANAVOS George. Probabilidad y Estadística. Mc Graw Hill, España 1988 (empleado únicamente en el desarrollo del marco conceptual).

MORRIS H. DEGROOT. Probabilidad y Estadística. Addison – Wesley Iberoamericana 1988 (empleado únicamente el desarrollo del marco conceptual).