FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES
DEPARTAMENTO DE MÉTODOS MATEMÁTICOS Y COMPUTACIÓN

MONOGRAFÍA

CÁLCULO DESARROLLADO DE ALGUNOS COEFICIENTES ECONOMÉTRICOS PARA CONTRASTES DE HIPÓTESIS

Autor:

Lic. Yadier A. Torres Sánchez
Dr Cs Ramón Rodríguez Betancourt

MAYO 2009

ÍNDICE

RESUMEN

La econometría es la rama de la economía que se ocupa de la estimación práctica de las relaciones económicas. La econometría utiliza datos económicos para estimar relaciones económicas cuantitativas y probar hipótesis acerca de ellas. El enfoque econométrico puede ser aplicado a la política, a la salud, educación, transporte, vivienda, finanzas y a la protección del medio ambiente.

En épocas pasadas el análisis econométrico estaba limitado por el aspecto computacional pero actualmente esto ha sido superado. Los potentes programas de computación existentes han permitido el uso masivo de esta técnica para obtener enfoques nuevos en el análisis cuantitativo de las variables que se están procesando.

Su conocimiento requiere del cálculo diferencial e integral, el álgebra de matrices, sistema de ecuaciones lineales, probabilidades y estadística.

El presente trabajo aporta elementos en el cálculo de algunos indicadores utilizados en pruebas estadísticas, tales como: prueba de rachas, coeficiente de correlación de rangos de Spearman y el contraste Kolmogorov – Smirnov para detectar normalidad, los cuales no está suficientemente desarrollado en la literatura, así como tampoco en los programas profesionales más utilizados actualmente. Este desarrollo ayudará a investigadores y profesores de estadística, que deseen profundizar en la obtención de los mismos, y servirá de enfoque metodológico para su aprendizaje, en cursos de maestrías y doctorados, así como para la confección de sistemas informáticos específicos basados en la econometría.

INTRODUCCIÓN

La estadística es una ciencia que forma parte de los planes de estudios de muchas carreras universitarias, variando el alcance, profundidad, tipos de métodos, objetivos y hasta los métodos de enseñanza aplicados en las diferentes carreras y en las distintas universidades en las que se enseña. A esta situación, de extensión masiva de esta materia, en las universidades se ha llegado producto de la necesidad creciente de su aplicación en la vida real. Respecto a esto último, es reconocido por todos en el mundo actual, la aplicabilidad cada vez mayor de la estadística y la utilidad de este conocimiento en los profesionales, incluyendo los no graduados universitarios.

La estadística es una de las técnicas fundamentales para la econometría, ciencia mucha más abarcadora y herramienta básica para el pronóstico y elección de políticas, por tanto esta situación problémica genera la necesidad de establecer estrategias para lograr que el investigador cuente con los medios y recursos para su comprensión y aplicabilidad.

La econometría es una disciplina reciente que, tras una serie de tanteos iniciales, ha ido perfilándose a lo largo del tiempo y tomando autonomía. En un principio la econometría y la escuela matemática estuvieron confundidas; Cowles, dentro de esta línea, destaca a Von Thunen, Walras, Jevons, Edgeworth, Pareto, etc.; sin duda dentro de este grupo habría que añadir a William Petty, del cual dice Shumpeter que fue el primer económetra, y también a Quesnay, que con la Tableau Economique trató de explicar empíricamente el principio de la interdependencia de todos los elementos que actúan dentro de un sistema económico. [5]

La escuela Lausana, aunque nunca trató de contrastar sus modelos matemáticos teóricos por medio de datos empíricos, es evidente que en ciertos aspectos fue una precursora de la econometría por su formulación de modelos teóricos.

A Pareto se debe la ecuación que explica la distribución general de la renta, además de intentar resolver el problema que presenta la teoría del consumo, obviando el problema de la medición de la utilidad.

Jevons y Edgeworth fueron realmente los iniciadores de la aplicación del método inductivo a la economía a través de la estadística y fueron los creadores de la escuela inglesa del equilibrio parcial. Necesariamente ha de mencionarse a Cournot, porque al formular las ecuaciones de la oferta y la demanda abrió un campo que ha sido ampliamente explotado por los econometristas.

Sin embargo, los trabajos econométricos con un sentido moderno comienzan con los estudios sobre la demanda, y en estos estudios evidentemente hay que dar primacía al profesor H.L. Moore, quien incorpora el método inductivo de la estadística al método deductivo de la lógica matemática. Las ideas dadas por Moore producen un movimiento científico que da lugar a dos momentos: uno es el tratado publicado por Ezekiel en 1930, titulado "Método del análisis de la correlación", y el publicado en 1938, por Shultz, titulado "La teoría y medida de la demanda". [11]

Muchos economistas prestigiosos han dado su versión sobre la definición de econometría, pero a nuestro juicio la que más se adapta al mundo de hoy es la dada por Chow en 1983: arte y ciencia de usar métodos para la medida de relaciones económicas. Actualmente las técnicas econométricas se utilizan de forma progresiva a nivel de microeconomía, fundamentalmente en la agricultura, la industria y otros sectores de economía nacional.

El surgimiento de paquetes estadísticos como SPSS v.12, STATA, EVIEWS y otros, permiten la aplicación y enseñanza de estas técnicas a un nivel más abarcador, obteniéndose soluciones a problemas que anteriormente no podían ser resueltos de forma eficiente. Estos paquetes proporcionan una ayuda y un tutorial, pero no se detalla científicamente de donde se obtienen los resultados que muestran los visores, fundamentalmente en los contrastes de hipótesis, y por tanto al desarrollar sistemas informáticos dirigidos a su introducción en la práctica social, no es posible obtener estos resultados. Lo mismo sucede con la bibliografía especializada.

Por lo anteriormente expuesto, el **problema** consiste en que *la literatura existente y programas profesionales carecen de suficiente desarrollo matemático que permitan una comprensión detallada de los resultados obtenidos en los estadísticos econométricos, lo cual dificulta la labor de los investigadores para la* *confección de sistemas informáticos específicos. De esta manera* el **objetivo general** de este trabajo es *desarrollar detalladamente el cálculo de algunos coeficientes econométricos utilizados en contraste de hipótesis permitiendo la elaboración de sistemas informáticos específicos*.

Objetivos específicos:

- Desarrollar matemáticamente de forma detallada la prueba de rachas para la comprobación de la independencia de los errores.
- Desarrollar matemáticamente de forma detallada el coeficiente de correlación de rangos de Spearman para la igualdad de varianzas.
- Desarrollar matemáticamente de forma detallada el contraste Kolmogorov – Smirnov para detectar la normalidad en muestra de datos.
- Utilizar en un sistema informático específico de regresión múltiple (SICEC) los contrastes de hipótesis hallados anteriormente.

DESARROLLO

1.1Aplicabilidad de la estadística

La estadística se plantea que es una ciencia que surge de la fusión entre la teoría de las probabilidades en los juegos de azar y la estadística o ciencia del estado, que estudia la descripción de datos. La estadística, es reconocida por muchos como la ciencia que se dedica a la colección, análisis, interpretación y representación de las conclusiones extraídas de información numérica (Dixon y Massey 1980). Se llega a plantear que el análisis estadístico no se inicia hasta que no se acumulan datos del problema de interés, constituyendo esta etapa una parte importante de esta ciencia (Kazmier, LJ, 1972), al igual que el análisis exploratorio de dichos datos (Ostle, B, 1979).

La literatura referente a la aplicación de la estadística en la dirección de procesos enfatiza en el papel de la misma en la toma de decisiones. Los autores definen que la estadística se relaciona con el uso y análisis de datos numéricos para resolver problemas generalmente con incertidumbres, en el marco teórico, práctico o de investigación. La propia necesidad de resolver problemas pone de manifiesto la aplicabilidad intrínseca de esta materia.

1.2 Tendencias en la enseñanza de la estadística

A pesar de que existe una conciencia respecto a la aplicabilidad de la estadística, el conocimiento de esta ciencia está protegido en mayor medida por el mundo académico, con un papel destacado de las maestrías, especializaciones y doctorados.

Respecto a la enseñanza de la estadística, lo que más se reporta es el contenido de lo que se debe enseñar, que a su vez es lo que sirve de guía para establecer los programas de estudio en los diferentes niveles de la enseñanza. Incluso a través de la moderna tecnología de la INTERNET cuando se realizan búsquedas, en su mayoría aparece este tipo de información. Así, nos podemos encontrar programas de estadística de distintas carreras universitarias, de maestrías o la literatura habitual acerca de la forma de estructurar el conocimiento de la estadística desde, la más general hasta la aplicada a cualquier campo del saber.

Con este análisis, no se pretende establecer una crítica, por la presencia de esta información, todo lo contrario, ya que una parte importante para la difusión de cualquier materia lo constituye la existencia de la misma. Lo que se desea es mostrar que se ha avanzado más en este sentido, y menos en la identificación de una forma ordenada de las necesidades en los que utilizan la estadística, que sirva de retroalimentación para la toma de acciones dirigidas a suplir esta necesidad.

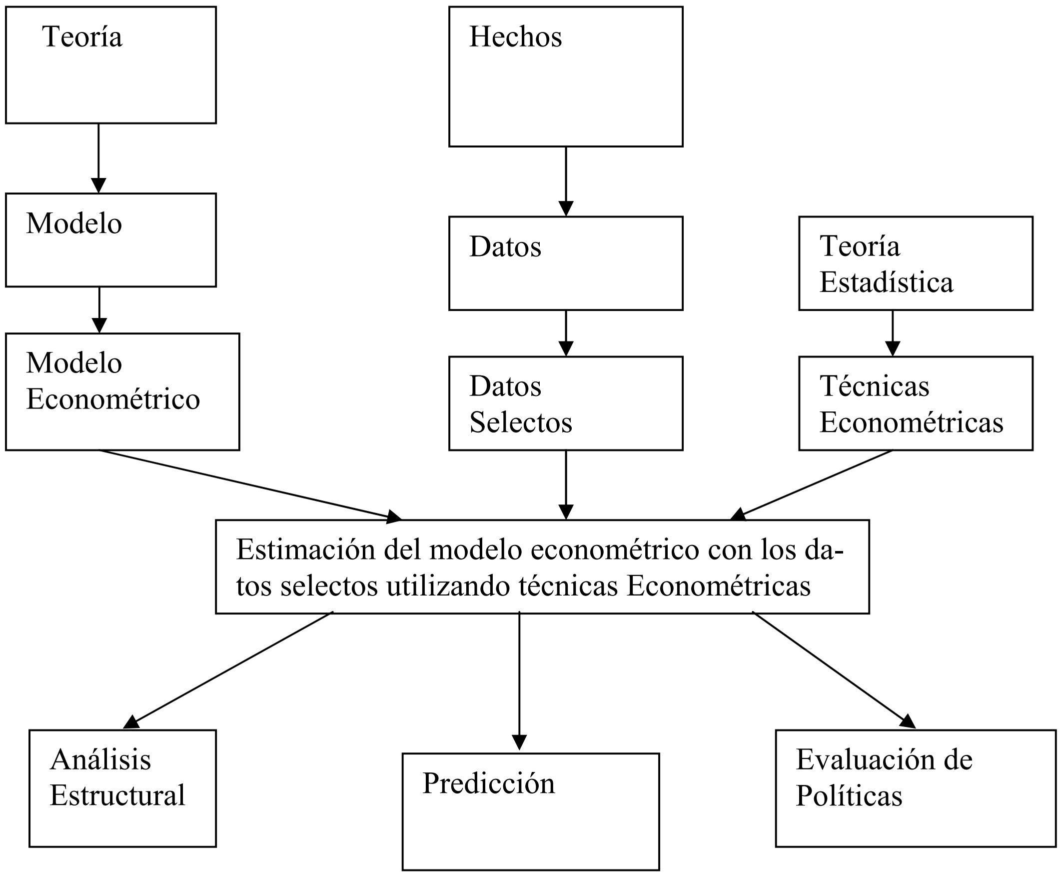
Tampoco, quiere decir que no existan experiencias que se hayan dirigido a cumplir tal objetivo. Los propios seminarios de actualización para determinados profesionales en estadística; las consultorías cuando se abordan con el principio de que los utilizadores asimilen y aprendan de forma continua al igual que los que enseñan acerca de las necesidades más frecuentes, sirviendo de fuente importante para introducir en los planes de estudio; los talleres o conferencias en estadística, en las que se establece un diálogo e intercambio de experiencias y se imparten cursos de actualización; las asociaciones de estadísticos que se crean; y los proyectos que se pueden acometer, tanto de investigación como de enseñanza de la estadística.

La consultoría es un tema de polémica, ya que no siempre se realiza de manera efectiva. Se ha reconocido tres roles del consultor: como ayudante, como líder o como colega. Los autores recomiendan el de colega, ya que es, en el cuál se logra una comunicación y relación entre el que cliente y el consultor, jugando ambos papeles activos. Aunque los autores, enfatizan el problema de la aplicabilidad efectiva de la estadística a través de una adecuada consultoría, tanto en el enfoque, etapas a seguir, conocimiento que deben tener los consultores y tipo de enseñanza que deben recibir los estudiantes estadísticos para estar mejor preparados, sólo desde la óptica de la consultoría no se puede resolver el problema tan complejo de la integración entre los que utilizan y los que enseñan la estadística.

La estadística es una de las técnicas fundamentales para la econometría, ciencia mucha mas abarcadora y herramienta básica para el pronóstico y elección de políticas, por tanto esta situación problémica genera la necesidad de establecer estrategias para lograr, que el que tenga que hacer uso de este campo del saber, cuente con los medios y recursos para comprender y aplicarlos.

1.3 Enfoque econométrico

El siguiente esquema 1 sintetiza el enfoque econométrico:



Esquema 1.Enfoque econométrico [11]

Son dos los componentes fundamentales de cualquier estudio econométrico: teoría y hechos. En realidad un logro importante de la econometría simplemente es combinar estos dos elementos. En cambio, una cantidad considerable del trabajo realizado destaca uno de ellos a costa del otro.

La teoría es uno de los elementos básicos en cualquier estudio econométrico, pero debe ser formulada de manera que pueda utilizarse. La forma mas eficaz con fines econométricos, como se aprecia en la grafica anterior suele ser la de un modelo, en particular un modelo econométrico. El modelo resume la teoría relevante para el sistema considerado y es la forma más conveniente para sintetizar esta teoría, para hacer mediciones prácticas y pruebas.

El otro elemento importante en un estudio econométrico es un conjunto de hechos, término que designa los eventos en el mundo real que están relacionados con el fenómeno bajo investigación Estos hechos conducen a un conjunto de datos, que representan observaciones de hechos relevantes. Sin embargo, en general, los datos deben ser seleccionados o reconfigurados en una diversidad de formas para adecuarlos al uso requerido por el estudio econométrico.

La teoría ha sido desarrollada en la forma de un modelo econométrico y los hechos, en un conjunto de datos selectos; el siguiente paso, central en el enfoque econométrico, combina estos dos factores básicos. En ese paso la estimación del modelo econométrico con los datos selectos, hace uso de un conjunto de técnicas econométricas, es decir extensiones de los métodos clásicos de la estadística, particularmente el análisis de regresión, la inferencia estadística y el muestreo. Las extensiones de los métodos clásicos son necesarias para considerar ciertos problemas especiales que se encuentran al estimar un modelo econométrico.

La gráfica muestra también tres propósitos fundamentales de la econometría: el análisis estructural, la predicción y la evaluación de políticas. Cualquier estudio econométrico puede tener uno, dos o todos estos propósitos, que representan los productos finales de la econometría, del mismo modo que la teoría y los hechos constituyen sus materias primas.

A tales propósitos no puede obviarse el papel que desempeñan los sistemas de cómputo profesionales, los cuales hacen posible obtener la solución para cualquier tipo de situación que se plantee. Sin embargo, la necesidad de un procesamiento posterior de estos resultados con otros propósitos genera la confección de sistemas estadísticos específicos.

Existen diversos contrastes para corroborar disímiles situaciones estadísticas, como técnicas de apoyo a la econometría. Estos contrastes se pueden determinar fácilmente mediante diferentes sistemas computacionales, que por lo general en su realización cuenta con la bibliografía que las asumen, pero no queda lo suficientemente explícita para los investigadores, estudiantes de postgrado y pregrado.

La extensión de la semipresencialidad en el pregrado sugiere la utilización de una bibliografía más detallada de estos conceptos estadísticos. En Internet esta literatura se encuentra dispersa y adolece también de explicación científica detallada, por tanto se profundizará en los cálculos de la Prueba de Rachas, el estadístico de Kolgomorov - Smirnov y el coeficiente de correlación de rangos de Spearman, enfocados fundamentalmente al análisis de regresión, mediante el método de los mínimos cuadrados, herramienta fundamental de la econometría.

1.4 Prueba de Rachas

Cuando se analiza el muestreo aleatorio, se dan varios métodos que brindan de antemano alguna seguridad de que la muestra sea aleatoria. Dado que hay situaciones en las cuales no se tiene control sobre la forma en que los datos son seleccionados, es útil disponer de alguna técnica para probar si alguna muestra puede considerarse aleatoria después de que en realidad se ha obtenido. Una de ellas se basa en el orden en que los datos fueron obtenidos; con más precisión, se fundamenta en el número de corridas exhibidas en los resultados muestrales.

Existen diversos contrastes para detectar la independencia de datos, lo que supone que el valor actual no depende del valor o valores anteriores. Dentro de las técnicas para detectarlas se encuentra el contraste de Durbin – Watson, Prueba de Box – Ljung, Prueba h de Durbin, Correlogramas, prueba de **aleatoriedad o corridas**, entre otras (Gujarati, Damodar M.: "Econometría"). En este apartado solo se hará referencia a la prueba de corridas o también conocida como rachas, utilizada en la mayoría de los sistemas de cómputo profesionales.

La formalización de la prueba de rachas enunciada en la literatura corriente es la siguiente [4]: Supongamos que se dispone de una muestra de una población y que, sobre cada valor de la muestra, se mide una variable en escala de razón o de intervalo X. Si x_i, i = 1,…..n, son las n observaciones de la variable X, la prueba de rachas es una prueba de independencia que se utiliza para contrastar la hipótesis nula de que la secuencia de apariciones de los valores en el orden observado es aleatoria o, equivalentemente, que las observaciones son independientes entre sí.

$$H_0 : Valores = independencia$$
$$H_1 : Valores \neq independencia$$

Para demostrar esta técnica se han seleccionado a través de un análisis de regresión lineal múltiple, los residuos no tipificados del mismo, para determinar si el orden en que aparecen estos es aleatorio, dicho de otra forma, si cada residuo es independiente de sus retardos.

Si una sucesión contiene n_1 símbolos del primer tipo y n_2 de cualquier otro tipo (ni n_1 ni n_2 son menores que 10), entonces la distribución muestral del **número total de corridas** u **,** puede aproximarse mediante una distribución normal a través de:

Media de ***u***	$\mu_u = \frac{2n_1 n_2}{n_1 + n_2} + 1$

Desviación estándar ***U***	$\sigma_u = \sqrt{\frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}}$

Es válido destacar que la mayoría de los paquetes estadísticos aplican una corrección a la fórmula de la **media**, cuando μ_u supera el valor del número de rachas μ , entonces se emplea un ajuste y la fórmula quedaría de la siguiente forma:

$$\mu_u = \frac{2n_1 n_2}{n_1 + n_2} + 0,5$$

Si por el contrario, μ supera el valor de μ_u , entonces sería:

$$\mu_u = \frac{2n_1 n_2}{n_1 + n_2} + 1,5$$

De esta manera, la prueba de hipótesis nula de aleatoriedad pudiera fundamentarse mediante el estadístico:

Estadístico de prueba de aleatoriedad	$z = \dfrac{u - u_u}{\sigma_u}$

Donde:

u = Número de rachas.

Dicho estadístico tiene aproximadamente la distribución normal estándar. Se pueden adquirir tablas especiales cuando n_1 ó n_2, o ambas son pequeñas.

Para demostrar lo anteriormente expuesto, primeramente, después de haber obtenido los residuos del modelo, es necesario conocer algunos conceptos tales como mediana, rachas, etc. La mediana es un estadístico de tendencia central en el sentido de que proporciona, cuando se ordenan los elementos de la muestra de menor a mayor, el valor del caso situado en el centro y es aquella que, ordenados los datos según su magnitud, el 50 % es menor que ésta, y el restante 50 % es mayor o igual. La racha no es más que una consecución de valores que poseen un mismo signo, por lo que en las siguientes tablas se puede mostrar sus respectivos cálculos, de manera que quede claro la utilización de esta prueba.

Utilizaremos un ejemplo donde se pretende estimar el rendimiento de la caña dado en t/ha, a partir de tres variables independientes: lluvia dado en mm^3; atenciones culturales medidos en puntos del 1 al 10 y por ciento de población.

Las variables y residuos no tipificados, una vez obtenida la función lineal de regresión, aparecen en la tabla 1. El comienzo del ordenamiento aparece en la tabla 2

Tabla 1

Variables y residuos no tipificados obtenidos en la obtención una función de regresión lineal				
Rendimiento de hortalizas	**Lluvia (mm3)**	**Atenciones culturales**	**Porcentaje de población**	**Residuos no tipificados**
107,52	1505,62	4,02	85,13	**-2,33093**
108,87	1506,20	4,27	85,16	**-1,59307**
111,53	1512,54	4,36	85,17	**,24284**
113,49	1514,31	4,60	85,24	**1,34907**
114,09	1528,60	4,62	85,38	**,00309**
114,41	1533,72	4,80	85,44	**-,71902**
118,69	1534,06	5,16	86,04	**,75414**
120,71	1545,91	5,42	86,34	**,07175**
123,74	1551,55	5,61	86,84	**,43008**
126,00	1557,15	5,65	87,22	**,71691**
126,53	1563,72	5,76	87,47	**-,48681**
126,87	1588,69	5,77	87,49	**-2,70592**
127,15	1591,14	5,93	87,59	**-3,30985**
127,50	1592,22	6,01	87,61	**-3,28110**
127,78	1592,34	6,12	87,75	**-3,70668**
130,69	1595,09	6,33	87,78	**-1,55176**
135,08	1596,64	6,54	87,86	**2,02459**
135,48	1611,88	6,56	88,10	**,03003**
138,59	1618,57	6,76	88,26	**1,55214**
142,55	1632,46	7,11	88,38	**3,08477**
143,05	1635,96	7,17	88,42	**2,98907**
143,10	1639,94	7,24	88,44	**2,44887**
143,61	1646,36	7,30	88,44	**2,21613**
144,09	1647,46	7,38	88,73	**1,41553**
144,68	1665,32	7,43	88,75	**,07806**
145,64	1671,03	7,48	88,91	**-,18384**
148,06	1671,58	7,64	88,92	**1,85932**

148,64	1673,21	7,82	89,14	**1,17514**
148,97	1673,40	7,85	89,30	**,86520**
151,28	1681,32	7,87	89,31	**2,32035**
151,46	1692,18	8,28	89,33	**,61971**
152,57	1702,90	8,46	89,36	**,23997**
153,10	1704,51	8,47	89,36	**,59276**
154,34	1722,28	8,47	89,47	**-,31525**
154,61	1725,27	8,72	89,53	**-1,00219**
156,41	1750,14	8,73	89,57	**-1,82232**
156,70	1755,02	9,03	89,64	**-2,80137**
158,84	1757,49	9,07	89,66	**-1,04839**
160,01	1764,79	9,32	89,78	**-1,47446**
160,74	1772,21	9,42	89,82	**-1,79970**
162,62	1779,43	9,68	89,91	**-1,41942**
164,84	1785,49	9,98	89,91	**-,33700**
129,91	1614,00	4,45	86,22	**4,70093**
131,17	1634,23	5,35	88,23	**-4,78179**
137,90	1646,71	6,49	88,28	**-1,50837**
145,78	1656,46	6,80	88,51	**4,03552**
158,70	1781,45	7,33	88,87	**2,36330**

Fuente: Elaboración propia de los autores.

Tabla 2

Residuos no tipificados y su ordenamiento según prueba

Residuos no tipificados obtenidos	Debajo de la mediana (D) Encima e igual a la mediana (E)	Determinar mediana	Cálculo de la media y sigma
-2,33093	D	-4,78179	
-1,59307	D	-3,70668	
0,24284	E	-3,30985	
1,34907	E	-3,2811	Uu= 23,9893617
0,00309	D	-2,80137	Sigma Cuadrado= 11,4839294
-0,71902	D	-2,70592	Sigma= 3,38879468

0,75414	**E**	-2,33093
0,07175	**Igual**	-1,82232
0,43008	**E**	-1,7997
0,71691	**E**	-1,59307
-0,48681	**D**	-1,55176
-2,70592	**D**	-1,50837
-3,30985	**D**	-1,47446
-3,2811	**D**	-1,41942
-3,70668	**D**	-1,04839
-1,55176	**D**	-1,00219
2,02459	**E**	-0,71902
0,03003	**D**	-0,48681
1,55214	**E**	-0,337
3,08477	**E**	-0,31525
2,98907	**E**	-0,18384
2,44887	**E**	0,00309
2,21613	**E**	0,03003
1,41553	**E**	**0,07175**
0,07806	**E**	0,07806
-0,18384	**D**	0,23997
1,85932	**E**	0,24284
1,17514	**E**	0,43008
0,8652	**E**	0,59276
2,32035	**E**	0,61971
0,61971	**E**	0,71691
0,23997	**E**	0,75414
0,59276	**E**	0,8652
-0,31525	**D**	1,17514

Z calculada= -2,9477624

Z tabla= 1,96

Mediana= **0,07175**

No se acepta la aleatoriedad

-1,00219	**D**	1,34907
-1,82232	**D**	1,41553
-2,80137	**D**	1,55214
-1,04839	**D**	1,85932
-1,47446	**D**	2,02459
-1,7997	**D**	2,21613
-1,41942	**D**	2,32035
-0,337	**D**	2,3633
4,70093	**E**	2,44887
-4,78179	**D**	2,98907
-1,50837	**D**	3,08477
4,03552	**E**	4,03552
2,3633	**E**	4,70093
	23 ítems debajo de la mediana 24 ítems encima o igual a la mediana **14 rachas**	

Fuente: Elaboración propia de los autores.

En la columna # 1 se muestran los residuos después de haber aplicado el análisis de regresión lineal múltiple, obtenido de la última columna de la tabla 1. En la columna # 3 se muestran los residuos ordenados de menor a mayor, en la columna # 2, luego de determinar el valor de la mediana, cuales de los residuos en la columna # 1 están por debajo de esta mediana y cuáles por encima (la **d** significa debajo y la **e** significa encima).

A continuación se muestra un análisis detallado del ejercicio propuesto y resuelto.

Pasos:

1. Hipótesis nula: H_0 : Los residuos son independientes

   H_1 : Los residuos no son independientes

2. Nivel de significación: $\alpha = 0.05$

3. Criterio: Se rechaza la hipótesis nula si los valores caen dentro de la región crítica.

4. Cálculos: Dado que $n_1 = 23$, $n_2 = 24$ y $u = 14$, se tiene:

$$\mu_u = \frac{2*23*24}{23+24} + 0,5 = \frac{1104}{47} + 0,5 = 23,989$$

Aquí se puede observar la corrección de 0,5 debido a que la media está por encima de la cantidad de rachas.

$$\sigma_u = \sqrt{\frac{2*23*24(2*23*24-23-24)}{(23+24)^2(23+24-1)}} = \sqrt{\frac{1104*1057}{2209*46}} = \sqrt{\frac{1166928}{101614}}$$

$$\sigma_u = \sqrt{\frac{1166928}{101614}} = \sqrt{11,484} = 3,389$$

$$z = \frac{14-23,989}{3,389} = -2,947$$

El resultado del valor de **z** que se ha calculado se compara con el valor de **z** de la tabla y resulta que, como se puede observar en la gráfica, este valor calculado cae dentro de la región de rechazo de la hipótesis nula, lo que implica que no hay aleatoriedad en los datos al rechazar dicha hipótesis.

1.5 Solución de Contraste de Rachas mediante el paquete estadístico SPSS V.12

Los resultados anteriores se pueden corroborar mediante el paquete estadístico profesional SPSS V.12, cuya salida se presenta a continuación. Como se observa la significación asintótica bilateral es de 0,003 por lo que existe suficiente evidencia empírica para no aceptar la hipótesis nula de aleatoriedad.

Runs Test

	Unstandardized Residual
Test Value[a]	,07175
Cases < Test Value	23
Cases >= Test Value	24
Total Cases	47
Number of Runs	14
Z	-2,948
Asymp. Sig. (2-tailed)	,003

a. Median

1.6 Coeficiente de correlación de rangos de Spearman

El **coeficiente de correlación de Spearman**, ρ (rho), es una prueba no paramétrica que mide la asociación o interdependencia entre dos variables discretas. Para calcular ρ, los datos son ordenados y reemplazados por su respectivo orden. Este estadístico tiene gran similitud con el coeficiente de correlación de Pearson, que es un índice estadístico que mide la relación lineal entre dos variables cuantitativas. A diferencia de la covarianza, la correlación de Pearson es independiente de la escala de medida de las variables.

El estadístico asociado al coeficiente de correlación de rangos de Spearman ρ viene dado por la expresión:

$$\rho = 1 - \frac{6 \sum D^2}{N(N^2 - 1)}$$

Donde *D* es la diferencia entre los correspondientes valores de *x* - *y*. *N* es el número de parejas.

Se tiene que considerar la existencia de datos idénticos a la hora de ordenarlos, aunque si éstos son pocos, se puede ignorar tal circunstancia.

Para muestras mayores de 20 observaciones, podemos utilizar la siguiente aproximación a la distribución t de Student:

$$t = \frac{\rho}{\sqrt{(1-\rho^2)/(n-2)}}$$

La interpretación de coeficiente de rangos de Spearman es igual que la del coeficiente de correlación de Pearson. Oscila entre -1 y +1, indicándonos asociaciones negativas o positivas respectivamente, cero, significa no correlación pero no independencia. La tau (τ) de Kendall es un coeficiente de correlación por rangos, inversiones entre dos ordenaciones de una distribución normal bivariante.

Para probar la homogeneidad de varianzas a través del coeficiente de correlación de rangos de Spearman, se contrastarán los residuos no tipificados, obtenidos con una de las variables independientes, en este caso con la lluvia.

Se hace necesario explicar, que la mayoría de los paquetes profesionales estadísticos aplican una corrección a la fórmula de ρ , debido a que en el numerador de la expresión de dichos programas profesionales emplean un ajuste, por lo que el valor que acompaña a la $\sum d_i^{\ 2}$ es de **6,13**, y no **6** como se plantea en la fórmula original. El la siguiente tabla 3 se plantean los datos que se quieren correlacionar:

Tabla 3

Datos primarios para la determinación del coeficiente de correlación de Spearman	
Residuos no tipificados	**Lluvia (mm^3)**
-2,33093	1505,62
-1,59307	1506,20
,24284	1512,54
1,34907	1514,31
,00309	1528,60
-,71902	1533,72
,75414	1534,06
,07175	1545,91
,43008	1551,55
,71691	1557,15

-,48681	1563,72
-2,70592	1588,69
-3,30985	1591,14
-3,28110	1592,22
-3,70668	1592,34
-1,55176	1595,09
2,02459	1596,64
,03003	1611,88
1,55214	1618,57
3,08477	1632,46
2,98907	1635,96
2,44887	1639,94
2,21613	1646,36
1,41553	1647,46
,07806	1665,32
-,18384	1671,03
1,85932	1671,58
1,17514	1673,21
,86520	1673,40
2,32035	1681,32
,61971	1692,18
,23997	1702,90
,59276	1704,51
-,31525	1722,28

-1,00219	1725,27
-1,82232	1750,14
-2,80137	1755,02
-1,04839	1757,49
-1,47446	1764,79
-1,79970	1772,21
-1,41942	1779,43
-,33700	1785,49
4,70093	1614,00
-4,78179	1634,23
-1,50837	1646,71
4,03552	1656,46
2,36330	1781,45

Fuente: Elaboración propia de los autores.

El primer paso es ordenar los datos de la primera columna. Después, se crean dos columnas más, ambas son para ordenar (establecer un lugar en la lista) de las dos primeras columnas. Después se crea una columna "d" que muestra las diferencias entre las dos columnas de orden. Finalmente, se crea otra columna "d^2". Esta última es sólo la columna "d" al cuadrado.

Después de realizar lo explicado con los datos del ejemplo, se obtiene la siguiente tabla 4

Tabla 4

Cálculos primarios para la determinación del coeficiente de correlación de Spearman					
Residuos no tipificados	**Lluvia (mm^3)**	**Orden (i)**	**Orden (t)**	**d**	**d^2**
-4,78179	1634,23	1	22	21	441
-3,70668	1592,34	2	15	13	169
-3,30985	1591,14	3	13	10	100
-3,2811	1592,22	4	14	10	100
-2,80137	1755,02	5	41	36	1296
-2,70592	1588,69	6	12	6	36
-2,33093	1505,62	7	1	6	36
-1,82232	1750,14	8	40	32	1024
-1,7997	1772,21	9	44	35	1225
-1,59307	1506,2	10	2	8	64
-1,55176	1595,09	11	16	5	25
-1,50837	1646,71	12	26	14	196
-1,47446	1764,79	13	43	30	900
-1,41942	1779,43	14	45	31	961
-1,04839	1757,49	15	42	27	729
-1,00219	1725,27	16	39	23	529
-0,71902	1533,72	17	6	11	121
-0,48681	1563,72	18	11	7	49
-0,337	1785,49	19	47	28	784
-0,31525	1722,28	20	38	18	324

-0,18384	1671,03	21	30	9	81
0,00309	1528,6	22	5	17	289
0,03003	1611,88	23	18	5	25
0,07175	1545,91	24	8	16	256
0,07806	1665,32	25	29	4	16
0,23997	1702,9	26	36	10	100
0,24284	1512,54	27	3	24	576
0,43008	1551,55	28	9	19	361
0,59276	1704,51	29	37	8	64
0,61971	1692,18	30	35	5	25
0,71691	1557,15	31	10	21	441
0,75414	1534,06	32	7	25	625
0,8652	1673,4	33	33	0	0
1,17514	1673,21	34	32	2	4
1,34907	1514,31	35	4	31	961
1,41553	1647,46	36	27	9	81
1,55214	1618,57	37	20	17	289
1,85932	1671,58	38	31	7	49
2,02459	1596,64	39	17	22	484
2,21613	1646,36	40	25	15	225
2,32035	1681,32	41	34	7	49
2,3633	1781,45	42	46	4	16
2,44887	1639,94	43	24	19	361
2,98907	1635,96	44	23	21	441

3,08477	1632,46	45	21	24	576
4,03552	1656,46	46	28	18	324
4,70093	1614	47	19	28	784

Fuente: Elaboración propia de los autores.

En este caso no se repite ninguno de los valores de las **dos** primeras columnas, en caso que se repitiese uno o varios de los valores de éstas, entonces el valor en el orden correspondiente sería el valor de la media de los números de orden que les corresponderían si no lo fueran.

Este ejemplo se presenta en la siguiente tabla 5

Tabla 5

Ejemplo de cálculo del coeficiente de correlación de Spearman					
Puntos del torneo (i)	Capacidad atlética de los equipos en puntaje (t)	orden(i)	orden(t)	d	d^2
71	**10**	1	1	0	0
80	**15**	2	**2.5**	0.5	0.25
83	**27**	3	4	1	1
86	**45**	**4.5**	8	3.5	12.25
86	**36**	**4.5**	**5.5**	1	1
92	**44**	6	7	1	1
95	**36**	**7.5**	**5.5**	2	4
95	**15**	**7.5**	**2.5**	5	25

Fuente: Elaboración propia de los autores

Volviendo a nuestro ejemplo, la suma de la columna d^2 resulta igual a: $\sum d_i^{\ 2} = 16612$, por lo que al aplicar la fórmula y sustituir quedaría de la siguiente manera:

$$\rho = 1 - \frac{6{,}13 \sum d_i^{\,2}}{N(N^2 - 1)} = 1 - \frac{6{,}13 * 16612}{47\,(2209 - 1)} = 1 - \frac{101831{,}56}{103776} = 0.01874$$

Nótese como se ha aplicado la corrección al sustituir el valor de **6** por **6,13**.

1.7 Contraste Kolmogorov – Smirnov(K-S) para detectar Normalidad

Al iniciar el análisis estadístico de una serie de datos, y después de la etapa de detección y corrección de errores, un primer paso consiste en describir la distribución de las variables estudiadas y, en particular, de los datos numéricos. Además de las medidas descriptivas correspondientes, el comportamiento de estas variables puede explorarse gráficamente de un modo muy simple. Para construir este tipo de gráfico, se divide el rango de valores de la variable en intervalos de igual longitud, representando sobre cada intervalo un rectángulo con área proporcional al número de datos en ese rango. Uniendo los puntos medios del extremo superior de las barras, se obtiene el llamado polígono de frecuencias. Si se observase una gran cantidad de valores de la variable de interés, se podría construir un histograma en el que las bases de los rectángulos fuesen cada vez más pequeñas, de modo que el polígono de frecuencias tendría una apariencia cada vez más suavizada. Esta curva suave "asintótica" representa de modo intuitivo la distribución teórica de la característica observada. Es la llamada función de densidad.

Una de las distribuciones teóricas mejor estudiadas en los textos y más utilizada en la práctica es la distribución normal, también llamada distribución gaussiana. Su importancia se debe fundamentalmente a la frecuencia con la que distintas variables asociadas a fenómenos naturales y cotidianos siguen, aproximadamente, esta distribución. Caracteres morfológicos (como la talla o el peso), o psicológicos (como el cociente intelectual) son ejemplos de variables de las que frecuentemente se asume que siguen una distribución normal. No obstante, y aunque algunos autores han señalado que el comportamiento de muchos parámetros en el campo de la salud puede ser descrito mediante una distribución normal, puede resultar incluso poco frecuente encontrar variables que se ajusten a este tipo de comportamiento.

El uso extendido de la distribución normal en las aplicaciones estadísticas puede explicarse, además, por otras razones. Muchos de los procedimientos estadísticos habitualmente utilizados asumen la normalidad de los datos observados. Aunque muchas de estas técnicas no son demasiado sensibles a desviaciones de la normal y, en general, esta hipótesis puede obviarse cuando se dispone de un número suficiente de datos, resulta recomendable contrastar siempre si se puede asumir o no una distribución normal. La simple exploración visual de los datos puede sugerir la forma de su distribución. No obstante, existen otras medidas, gráficos de normalidad y contrastes de hipótesis que pueden ayudarnos a decidir, de un modo más riguroso, si la muestra de la que se dispone procede o no de una distribución normal. Cuando los datos no sean normales, podremos o bien transformarlos o emplear otros métodos estadísticos que no exijan este tipo de restricciones (los llamados métodos no paramétricos).

Las pruebas K-S son pruebas no paramétricas que se utilizan para diferencias entre distribuciones acumuladas. La prueba unimuestral se refiere a la concordancia entre una distribución acumulada observada de valores muestrales y una función continua determinada; es, pues una prueba de bondad del ajuste. La prueba bimuestral está relacionada con la conformidad entre dos distribuciones acumuladas observadas; pruébese en hipótesis de que dos muestras independientes provienen de distribuciones continuas idénticas y es sensible a diferencias de la población con respecto a localización, dispersión o sesgo.

La prueba K-S unimuestral es en general más eficiente que la Chi Cuadrado para bondad del ajuste en pruebas pequeñas, y puede emplearse en muestras muy pequeñas donde la prueba Chi Cuadrado no se aplica.

La prueba unimuestral se fundamenta en la diferencia máxima absoluta D entre los valores de la distribución acumulada de una muestra aleatoria de tamaño N y una distribución teórica determinada. Para decidir si esta diferencia es mayor de lo que razonablemente puede esperarse con un nivel de significación determinada se confeccionan tablas correspondientes para valores críticos de D.

La hipótesis nula que se establece es la siguiente:

$$F_{(x)}\begin{cases} 0 & para\ x \leq 0 \\ \dfrac{x}{n} & para\ 0 < x < n \\ 1 & para\ x \geq n \end{cases}$$

La hipótesis alternativa representa la no conformidad de los datos de la muestra.

A continuación en la tabla 6, se expone la información de partida para esta prueba

Tabla 6

Información de partida para la prueba Kolgomorv-Smirnov						
A Residuos	B ordenados	C acumulado 1/n	D X-media/desv. tip.	E Área bajo la curva	F Acumulada C menos área bajo la curva E	Diferencia entre 1/n y F
-2,33093	-4,78179	0,021276596	-2,312277563	0,0104	0,010876596	0,0104
-1,59307	-3,70668	0,042553192	-1,792398453	0,0367	0,005853192	0,015423404
0,24284	-3,30985	0,063829788	-1,600507737	0,0548	0,009029788	0,012246808
1,34907	-3,2811	0,085106384	-1,586605416	0,0559	0,029206384	-0,007929788
0,00309	-2,80137	0,10638298	-1,35462766	0,0885	0,01788298	0,003393616
-0,71902	-2,70592	0,127659576	-1,308471954	0,0951	0,032559576	-0,01128298
0,75414	-2,33093	0,148936172	-1,127142166	0,1292	0,019736172	0,001540424
0,07175	-1,82232	0,170212768	-0,881199226	0,1894	-0,019187232	0,040463828
0,43008	-1,7997	0,191489364	-0,870261122	0,1922	-0,000710636	0,021987232
0,71691	-1,59307	0,21276596	-0,770343327	0,2206	-0,00783404	0,029110636
-0,48681	-1,55176	0,234042556	-0,750367505	0,2266	0,007442556	0,01383404
-2,70592	-1,50837	0,255319152	-0,72938588	0,2327	0,022619152	-0,001342556
-3,30985	-1,47446	0,276595748	-0,712988395	0,2389	0,037695748	-0,016419152
-3,2811	-1,41942	0,297872344	-0,686373308	0,2451	***0,052772344***	-0,031495748
-3,70668	-1,04839	0,31914894	-0,506958414	0,305	0,01414894	0,007127656
-1,55176	-1,00219	0,340425536	-0,484617988	0,3156	0,024825536	-0,00354894
2,02459	-0,71902	0,361702132	-0,347688588	0,3632	-0,001497868	0,022774464

0,03003	-0,48681	0,382978728	-0,235401354	0,4042	-0,021221272	0,042497868
1,55214	-0,337	0,404255324	-0,162959381	0,4364	-0,032144676	***0,053421272***
3,08477	-0,31525	0,42553192	-0,152441973	0,4404	-0,01486808	0,036144676
2,98907	-0,18384	0,446808516	-0,088897485	0,4641	-0,017291484	0,03856808
2,44887	0,00309	0,468085112	0,001494197	0,5	-0,031914888	0,053191484
2,21613	0,03003	0,489361708	0,014521277	0,504	-0,014638292	0,035914888
1,41553	0,07175	0,510638304	0,034695358	0,512	-0,001361696	0,022638292
0,07806	0,07806	0,5319149	0,037746615	0,516	0,0159149	0,005361696
-0,18384	0,23997	0,553191496	0,116039652	0,5478	0,005391496	0,0158851
1,85932	0,24284	0,574468092	0,117427466	0,5478	0,026668092	-0,005391496
1,17514	0,43008	0,595744688	0,207969052	0,5832	0,012544688	0,008731908
0,8652	0,59276	0,617021284	0,286634429	0,6141	0,002921284	0,018355312
2,32035	0,61971	0,63829788	0,299666344	0,6179	0,02039788	0,000878716
0,61971	0,71691	0,659574476	0,346668279	0,6368	0,022774476	-0,00149788
0,23997	0,75414	0,680851072	0,36467118	0,6406	0,040251072	-0,018974476
0,59276	0,8652	0,702127668	0,418375242	0,6628	0,039327668	-0,018051072
-0,31525	1,17514	0,723404264	0,568249516	0,7157	0,007704264	0,013572332
-1,00219	1,34907	0,74468086	0,652354932	0,7422	0,00248086	0,018795736
-1,82232	1,41553	0,765957456	0,684492263	0,7517	0,014257456	0,00701914
-2,80137	1,55214	0,787234052	0,750551257	0,7734	0,013834052	0,007442544
-1,04839	1,85932	0,808510648	0,899090909	0,8159	-0,007389352	0,028665948
-1,47446	2,02459	0,829787244	0,979008704	0,8365	-0,006712756	0,027989352

-1,7997	2,21613	0,85106384	1,071629594	0,8577	-0,00663616	0,027912756
-1,41942	2,32035	0,872340436	1,122026112	0,8686	0,003740436	0,01753616
-0,337	2,3633	0,893617032	1,142794971	0,8729	0,020717032	0,000559564
4,70093	2,44887	0,914893628	1,184173114	0,881	0,033893628	-0,012617032
-4,78179	2,98907	0,936170224	1,445391683	0,9265	0,009670224	0,011606372
-1,50837	3,08477	0,95744682	1,491668279	0,9319	0,02554682	-0,004270224
4,03552	4,03552	0,978723416	1,951411992	0,9744	0,004323416	0,01695318
2,3633	4,70093	1,000000012	2,273176983	0,9884	0,011600012	0,009676584
		N (tamaño muestra)	47			
		Media	1,2699E-14			
		Desviación típica	2.068			

Fuente: Elaboración propia de los autores.

Como se puede apreciar en la última columna, la diferencia más significativa es de ***0,053421272,*** por lo que se hace necesario construir el contraste de la máxima diferencia.

Para esto es indispensable corroborar las tablas 7 y 8 que se muestran a continuación.

Tabla 7

Distribuciones para contrastar	
DISTRIBUCIÓN QUE SE CONTRASTA	$k(n)$
General. Parámetros conocidos.	$k(n)=\sqrt{n}+0.12+\frac{0.11}{\sqrt{n}}$
Normal	$k(n)=\sqrt{n}-0.01+\frac{0.85}{\sqrt{n}}$
Exponencial	$k(n)=\sqrt{n}+0.12+\frac{0.11}{\sqrt{n}}$
Weibull	$k(n)=\sqrt{n}$

Fuente: Elaboración propia de los autores.

Tabla 8

Valores para la significación			
c_α	α		
Modelo	**0.1**	**0.05**	**0.01**
General	1.224	1.358	1.628
Normal	0.819	0.895	1.035
Exponencial	0.990	1.094	1.308
Weibull n=10	0.760	0.819	0.944
Weibull n=20	0.779	0.843	0.973
Weibull n=50	0.790	0.856	0.988
Weibull n=∞	0.803	0.874	1.007

Fuente: Elaboración propia los autores.

En la primera se denota la distribución que se desea contrastar y en la segunda los valores para las significaciones de 1%, 5% y 10% para las distintas distribuciones.

Se trabajará con un α de 0.05 para la normal:

$$K_{(n)} = \sqrt{n} - 0.01 + \frac{0.85}{\sqrt{n}}$$

$$K_{(n)} = \sqrt{47} - 0.01 + \frac{0.85}{\sqrt{47}}$$

$$K_{(n)} = 6.86 - 0.01 + \frac{0.85}{6.86}$$

$$K_{(n)} = 6.974$$

De donde:

$$D_a = \frac{0.895}{6.974} = 0.128$$

Como el valor de D = 0.0534 < 0.128 no se rechaza la Hipótesis Nula de Normalidad y se acepta que los datos se distribuyen normalmente.

Lo expuesto anteriormente se puede corroborar con el cálculo del p-valor o significación del estadístico de contraste Z de Kolmogorov, el cual brinda información adicional sobre la decisión tomada. Es muy importante conocer la diferencia entre el cálculo de la máxima diferencia absoluta y su comparación con la D-tabla para el rechazo o no de la hipótesis nula contrastada y el p-valor o significación. En el primero de los casos se puede rechazar o no, pero no se conoce hasta que grado de significación usted podría escoger la hipótesis nula formulada. Sin embargo el p-valor o significación es la probabilidad de observar un estadístico tan extremo como el que se ha obtenido si la hipótesis nula es cierta; esto significa que p-valores pequeños representan gran evidencia empírica en contra de la hipótesis nula, mientras que p-valores grandes aportan escasa evidencia contra dicha hipótesis nula. Por ejemplo, si el p-valor = 0.50 (escrito como número decimal, no como porcentaje), entonces observaríamos un valor del estadístico tan extremo como el que se ha obtenido en el 50 por ciento de todas las muestras aleatorias cuando la hipótesis nula es cierta; esto es una evidencia muy débil en contra de H_0.

1.8 Evaluando Pr (Dn < d)

El método usado es una sucesión de desarrollos que comenzaron con una visión de Kolmogorov referidos a los pasos de la función de distribución acumulada de la muestra (FDA) como un proceso Poisson y culminó con un tratamiento especial de Durbin. Su monografía resume y extiende los resultados de numerosos autores quienes han progresado en el problema entre los años 1933 – 1973. El resultado es un método que expresa la probabilidad requerida como un elemento real en la potencia enésima de una matriz formada fácilmente. La historia del desarrollo está disponible a través de 136 referencias de monografías. Sin embargo, en las referencias que se abordan las formas de calcular los p-valores, no queda totalmente explícito el desarrollo paso a paso del cálculo, lo que entorpece el conocimiento detallado, ocasionándole problemas al investigador en el caso que quiera utilizar estos cálculos para desarrollar sistemas informáticos. Cabe destacar que existen varias formas de calcular el p-valor o significación en el contraste de normalidad Kolmogorov – Smirnov para una muestra, tantos los basados en sucesiones factoriales, límites e integrales múltiples.

En la bibliografía estudiada estos métodos son extremadamente complejos, por lo que el procedimiento aquí expuesto en forma matricial, brinda una aproximación al verdadero p-valor o significación. **Poner aquí referencias**

Se quiere evaluar la Pr (Dn < d), donde:

$d = \frac{k-h}{n}$, con k un entero positivo y $0 \le h < 1$.

Entonces la $\mathrm{Pr}\,(\mathrm{Dn} < \mathrm{d}) = \frac{\mathrm{n!}}{\mathrm{n^n}} t_{kk}$, dónde t_{kk} es el elemento k,k de una matriz $T = H^n$ y H es una matriz $m * m$, $m = 2k - 1$, del cual la forma general se infiere del caso particular cuando $m = 6$ y $h \le 1/2$.

$$\begin{bmatrix} (1-h^1)/1! & 1 & 0 & 0 & 0 & 0 \\ (1-h^2)/2! & 1/1! & 1 & 0 & 0 & 0 \\ (1-h^3)/3! & 1/2! & 1/1! & 1 & 0 & 0 \\ (1-h^4)/4! & 1/3! & 1/2! & 1/1! & 1 & 0 \\ (1-h^5)/5! & 1/4! & 1/3! & 1/2! & 1/1! & 1 \\ (1-2h^6)/6! & (1-h^5)/5! & (1-h^4)/4! & (1-h^3)/3! & (1-h^2)/2! & (1-h^1)/1! \end{bmatrix}$$

El ejemplo anterior es cuando $0 \leq h \leq \frac{1}{2}$. Para $\frac{1}{2} < h < 1$ el elemento final de la izquierda de la matriz debería ser $\left(1 - 2h^m + (2h-1)^m\right)/\, m!$, de esta forma $\left(1 - 2h^m + \max(0,2h-1)^m\right)/\, m!$ es la forma general del elemento final de la izquierda.

El último elemento de la matriz refleja la primera columna en orden inverso. El elemento de la primera columna y la última fila (i, j_{th}) es $\left(1/_{i - j + 1}\right)!$ si $(i - j + 1) \geq 0$, sino 0.

En nuestro ejemplo, supongamos que $n = 15$, y queremos $P_r(D_{15} \leq 0{,}183)$. Se expresa $d = 0{,}183$ como $0{,}183 = \frac{k-h}{15}$, si se multiplica $0.183 * 15$ resultaría que el número entero más próximo por exceso sería 3, de esta forma k = 3, el orden de la matriz estaría dado por $m = 2k - 1 = 2*3 - 1 = 5$, y $h = 0.255$

De donde:

$$\begin{pmatrix} (1-0{,}255^1)/1 & 1 & 0 & 0 & 0 \\ (1-0{,}255^2)/2 & 1 & 1 & 0 & 0 \\ (1-0{,}255^3)/6 & 1/2 & 1 & 1 & 0 \\ (1-0{,}255^4)/24 & 1/6 & 1/2 & 1 & 1 \\ (1-2*0{,}255^5)/120 & (1-0{,}255^4)/24 & (1-0{,}255^3)/6 & (1-0{,}255^2)/2 & (1-0{,}255^1)/1 \end{pmatrix}$$

La matriz resultaría:

$$\begin{pmatrix} 0{,}745 & 1 & 0 & 0 & 0 \\ 0{,}4674875 & 1 & 1 & 0 & 0 \\ 0{,}1639031 & 0{,}5 & 1 & 1 & 0 \\ 0{,}0414905 & 0{,}1666667 & 0{,}5 & 1 & 1 \\ 0{,}0083154 & 0{,}0414905 & 0{,}1639031 & 0{,}4674875 & 0{,}745 \end{pmatrix}$$

Al elevar esta matriz a la potencia n, que no es más que el tamaño muestral, el elemento t_{kk}, representando el valor de la tercera fila y tercera columna $3;3$, sería 123224,2394, (Microsoft Encarta 2008) y al multiplicar este número por $\frac{15!}{15^{15}}$ resultaría 0,36798 que al ser restado del entero 1, sería igual a **0,63202.**

Los resultados por el paquete estadístico SPSS v.12 se pueden observar a continuación, y dada que la significación está por encima del valor de alfa predeterminado, existe suficiente evidencia empírica en favor de la Hipótesis Nula de Normalidad.

Prueba de Kolmogorov-Smirnov para una muestra

		VAR00001
N		15
Parámetros normales[a,b]	Media	-,9710
	Desviación típica	1,70116
Diferencias más extremas	Absoluta	,183
	Positiva	,121
	Negativa	-,183
Z de Kolmogorov-Smirnov		,710
Sig. asintót. (bilateral)		,695

a. La distribución de contraste es la Normal.

b. Se han calculado a partir de los datos.

CONCLUSIONES

1. La formulación y desarrollo del análisis ha permitido concluir que los objetivos propuestos se han cumplido totalmente.
2. En el caso de las pruebas de rachas, Rho Spearman y Kolmogorov Smirnov, se pudieron calcular independientemente del programa de cómputo, comprobándose los resultados con el programa SPSS v.12.
3. La elaboración de la monografía es un material de consulta y permitirá los estudiantes de pregrado, profesores de los colectivos de estadística y estudiantes de postgrado, doctorantes y cursistas de maestría, tener acceso al desarrollo del cálculo de estos estadísticos que aparecen en los programas fundamentales profesionales actuales tales como SPSS en sus diferentes versiones, STATA, EViews, Statgraphics, etc.
4. Estos resultados permitirán obtener la base para efectuar los cálculos pertinentes en programas informáticos de estadísticas con enfoques específicos, que se confeccionan tomando en consideración la aplicación de técnicas econométricas como es el análisis de regresión y las series de tiempo.

BIBLIOGRAFÍA

1. Box, G.: Jenkins, G.: Time series analysis, forecasting and control. Editorial Holden-Day, 1976.
2. Dixon, W.J. y Massey., F, J. Introducción al Análisis Estadística. Editorial Pueblo y Educación. Ciudad de la Habana 1980.
3. Caridad, J.M. y Ocerín: "Econometría: modelos econométricos y series temporales". Ediciones Ariel Barcelona 2006.

4. Gujarati, Damodar M.: "Econometría". McGraw-Hill, 2006.

5. HILL, C. R.; JUDGE, G. y GRIFFITHS, W.: *"Undergraduate Econometrics"*. John Willey and Sons, 1998.

6. HILL, C. R. y REIMAN, M. A.: *"Using Eviews for Undergraduate"*. John Willey and Sons, 2000.

7. Kazmier, LJ Análisis Estadístico para la Empresa y la Economía. Editorial Pueblo y Educación. La Habana 1982

8. Malinvaud, Edmond. "Métodos Estadísticos de la Econometría". Ediciones Ariel, Barcelona. España .1967

9. Microsoft Encarta 2008

10. PINDYCK, R. S.: *"Econometría: modelos y pronósticos"*. McGraw-Hill, 2001.

11. Pulido, Antonio. Modelos Econométricos. Ediciones Enpes. La Habana, 1993.

12. Pupo, Juana y Otros. Análisis de Regresión y Series Cronológicas. Ediciones Enpes. La Habana 1983.

13. Rodríguez B., Ramón. ECONOMETRÍA MODERNA. Publicado por el Instituto Cubano de la Salud. La Habana 2008. ISBN 978-959-7158-67-7.

14. Selección de Tablas Estadísticas.