75

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

ESTADISTICA NO PARAMETRICA, APLICACIONES MEDIANTE EL SPSS-PC.

T E S I S

PARA OBTENER EL TITULO DE:

A C T U A R I A

P R E S E N T A :

CONSUELO NOHPAL DE LA ROSA

DIR. DE TESIS: ACT. FRANCISCO SANCHEZ VILLARREAL

2001

M. EN C. ELENA DE OTEYZA DE OTEYZA
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo de Tesis:

"ESTADISTICA NO PARAMETRICA, APLICACIONES MEDIANTE EL SPSSPC"

realizado por CONSUELO NÓHPAL DE LA ROSA

con número de cuenta 09025821-1, pasante de la carrera de ACTUARiA

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis Propietario	ACT.	FRANCISCO SÁNCHEZ VILLARREAL
Propietario	ACT.	MA. SUSANA BARRERA OCAMPO
Propietario	ACT.	ALFONSO GARCÍA DURAN
Suplente	M. EN C.	MANUEL FRANCISCO ROMÁN ENRIQUEZ
Suplente	ACT.	ViCTOR MANUEL SOLiS NÁJERA

Consejo Departamental de MATEMÁTICAS

M. EN C. JOSE ANTONIO FLORES DIAZ

FACULTAD DE CIENCIAS
CONSEJO DEPARTAMENTAL
DE
MATEMATICAS

Para mis Padres Lidia y Arcadio
A mis Hermanos Adriana y Carlos

A mis profesores, principalmente aquellos que me mostraron el maravilloso mundo de la estadística Mtro. Francisco y Act. Susana.

A mi familia y a toda persona que me apoyo en todo el transcurso de mi carrera.

Gracias.

ÍNDICE

INTRODUCCIÓN Y GENERALIDADES

1.- HIPOTESIS CIENTÍFICA
1.1 "QUE ES UNA HIPÓTESIS CIENTÍFICA" 1
1.2 HIPÓTESIS NULA (HO) 4
1.3 HIPÓTESIS ESTADÍSTICA 5

2.- ESCALAS DE MEDICIONES
2.1 ¿QUÉ ES MEDIR? 6
2.2 CONCEPTO DE MEDICIÓN 7
2.3 CLASIFICACIÓN CUALITATIVA Y CUANTITATIVA 7
2.4 PROPIEDADES DE LA MEDIDA 8
2.5 TIPOS DE ESCALAS 10
2.5.1 Nominal 10
2.5.2 Ordinal 11
2.5.3 Intervalar 12
2.5.4 Razón o cociente 12

3.- ESTADISTICAS PARAMÉTRICAS Y NO PARAMÉTRICAS
3.1 PRUEBAS ESTADÍSTICAS PARAMETRICAS 14
3.2 PRUEBAS ESTADÍSTICAS NO PARAMÉTRICAS 15

4.- USO DE LAS PRUEBAS ESTADÍSTICAS 17
4.1 FORMULACIÓN DE UNA HIPÓTESIS NULA Y ALTERNATIVA 18
4.2 LA ELECCIÓN DEL ESTADÍSTICO DE PRUEBA 18

4.3 DISTRIBUCIÓN DEL ESTADÍSTICO DE PRUEBA 19
4.4 NIVEL DE SIGNIFICANCIA (α) 20
4.5 IDENTIFICACIÓN DE LA REGIÓN DE RECHAZO Ó VALOR CRITICO Y REGLA DE DECISIÓN. 21
4.6 P-VALUE 23

5.- INTRODUCCIÓN AL PAQUETE SPSS
5.1 INTRODUCCIÓN 25
5.2 PRESENTACIÓN GENERAL DE LAS VENTANAS DE SPSS 27
5.3 PRESENTACIÓN DE LA BARRA DE HERRAMIENTAS 29

6.- PRUEBAS NO PARAMÉTRICAS PARA 2 MUESTRAS RELACIONADAS 32
6.1 PRUEBA DE LOS SIGNOS 33
6.1.1 Función 33
6.1.2 Potencia de la prueba 34
6.1.3 Supuestos 35
6.1.4 Método 35
6.1.5 Hipótesis 36
6.1.6 Estadístico de Prueba 36
6.1.7 Regla de decisión 37
6.1.8 Ejemplos 38
6.1.9 Utilizando el paquete SPSS 42
6.1.10 Observaciones 44

6.2 PRUEBA DE WILCOXON
6.2.1 Función 44
6.2.2 Potencia de la prueba 46
6.2.3 Supuestos 46
6.2.4 Método 47
6.2.5 Hipótesis 48
6.2.6 Estadístico de Prueba 48
6.2.7 Regla de decisión 49
6.2.8 Ejemplos 50
6.2.9 Utilizando el paquete SPSS 53
6.2.10 Observaciones 56

7.- PRUEBAS NO PARAMETRICAS PARA 2 MUESTRAS INDEPENDIENTES 57
7.1 PRUEBA DE LA MEDIANA
7.1.1 Función 58
7.1.2 Supuestos 59
7.1.3 Método 59
7.1.4 Hipótesis 59
7.1.5 Estadístico de Prueba 59
7.1.6 Regla de decisión 61

7.1.7 Ejemplos 61
7.1.8 Utilizando el paquete SPSS 64

7.2 PRUEBA U MANN-WHITNEY
7.2.1 Función 66
7.2.2 Potencia de la prueba 67
7.2.3 Supuestos 67
7.2.4 Método 67
7.2.5 Hipótesis 68
7.2.6 Estadístico de Prueba 68
7.2.7 Regla de decisión 71
7.2.8 Ejemplos 72
7.2.9 Utilizando el paquete SPSS 76

8.- PRUEBAS NO PARAMÉTRICAS PARA K MUESTRAS RELACIONADAS 79
8.1 LA PRUEBA DE FRIEDMAN PARA ANÁLISIS DE VARIANZA CON DOS CRITERIOS DE CLASIFICACIÓN.
8.1.1 Función 80
8.1.2 Potencia de la prueba 81
8.1.3 Supuestos 81
8.1.4 Método 82
8.1.5 Hipótesis 83
8.1.6 Estadístico de Prueba 83
8.1.7 Regla de decisión 84
8.1.8 Relación de la Estadística T de Friedman con el Coeficiente de concordancia de Kendall. 84
8.1.9 Ejemplo 85
8.1.9 Utilizando el paquete SPSS 86

9.- PRUEBAS NO PARAMÉTRICAS PARA K MUESTRAS INDEPENDIENTES 90
9.1 PRUEBA DE LA MEDIANA GENERALIZADA
9.1.1 Función 91
9.1.2 Supuestos 91
9.1.3 Método 91
9.1.4 Hipótesis 92
9.1.5 Estadístico de Prueba 92
9.1.6 Regla de decisión 93
9.1.7 Ejemplo 93
9.1.8 Utilizando el paquete SPSS 95

9.2 PRUEBA DE KRUSKAL Y WALLIS
9.2.1 Función 96
9.2.2 Potencia de la prueba 97
9.2.3 Supuestos 98

9.2.4 Método 98
9.2.5 Hipótesis 98
9.2.6 Estadístico de Prueba 99
9.2.7 Regla de decisión 100
9.2.8 Ejemplo 100
9.2.9 Utilizando el paquete SPSS 102

10.- ANALISIS DE TABLAS DE CONTINGENCIA 104
10.1 TABLAS DE CONTINGENCIA
10.1.1 Función 105
10.2 TABLAS DE CONTINGENCIA DE c x k 105
10.2.1. Prueba de la Ji-Cuadrada para probar igualdad de proporciones 106
10.2.1.1 Supuestos 106
10.2.1.2 Hipótesis 106
10.2.2 Prueba de la Ji-Cuadrada para probar independencia. 107
10.2.2.1 Supuestos 107
10.2.2.2 Hipótesis 107
10.2.3 Estadístico de prueba 108
10.2.4 Regla de decisión 108
10.2.5 Recomendaciones 108
10.3 TABLA DE CONTINGENCIA DE 2X2. 109
10.3.1 Recomendaciones para una tabla de 2 x 2 109
10.4 EJEMPLOS 109
10.5 UTILIZANDO EL PAQUETE SPSS 111

11.- PRUEBAS DE ASOCIACIÓN 117
11.1 CONCEPTO DE CORRELACIÓN. 117
11.2 COEFICIENTE DE CORRELACIÓN DE SPEARMAN
11.2.1 Función 118
11.2.2 Potencia -eficiencia 118
11.2.3 Supuestos 119
11.2.4 Método 119
11.2.5 Estadístico de Prueba 119
11.2.6 Hipótesis 120
11.2.7 Regla de decisión 121
11.2.8 Ejemplo 121
11.2.9 Utilizando el paquete SPSS 123

12.- ELECCION DE LA PRUEBA ESTADISTICA ADECUADA
12.1 LA MANERA EN QUE SE OBTUVO Y LA NATURALEZA DE LA POBLACIÓN DE LA CUAL FUE EXTRAÍDA LA MUESTRA DE DATOS. 126
12.2 EL TIPO DE ESCALA QUE SE TIENEN EN LOS DATOS DE LA INVESTIGACIÓN. 126
12.3 POTENCIA 127
12.4 EFICACIA 129

13.- CONCLUSIONES 130

ANEXOS

BIBLIOGRAFIA

INTRODUCCIÓN Y GENERALIDADES

ESTADISTICA NO PARAMETRICA

La función central de la estadística moderna, es la inferencia estadística (la cual se ocupa de hacer afirmaciones acerca de las poblaciones basándose en muestras tomadas de ellas). La estadística inferencial esta interesada en dos tipos de problemas: La estimación de los parámetros de la población y las pruebas de hipótesis, además proporciona instrumentos que formalizan y estandarizan procedimientos para obtener conclusiones apoyadas en probabilidades.

Las primeras técnicas de inferencia estadística que aparecieron, fueron aquellas que hicieron suposiciones acerca de la naturaleza de las poblaciones de las cuales se derivan las observaciones y los datos, estas técnicas estadísticas se llaman paramétricas. Por ejemplo, una técnica de inferencia puede estar basada en la suposición de que los datos se derivan de una población normalmente distribuida o en la suposición de que dos conjuntos de datos se tomaron de poblaciones que tienen la misma varianza.

A pesar de la gran importancia de este tipo de técnica paramétrica, se desarrolló un gran número de técnicas de inferencia que no hacen suposiciones rigurosas o numerosas acerca de la población de la cual se han muestreado los datos, ya que existen experimentos que producen respuestas que no son cuantificables, es decir generan mediciones que pueden ordenarse, pero la posición de la respuesta en una escala de medición es arbitraria. Estas técnicas reciben el nombre de no paramétricas o técnicas de distribución libre, las cuales dan como resultado conclusiones que requieren menos suposiciones.

Los métodos estadísticos no paramétricos son útiles no solamente cuando los datos representan una ordenación, sino también cuando se tienen únicamente diferencias direccionales. Es decir, una persona puede indicar su preferencia entre un par de artículos de prueba, pero no querer o no poder indicar una medida de la magnitud de su preferencia.

Es más, los métodos no paramétricos son aplicables no solo en los casos en que las mediciones son difíciles de cuantificar, sino que son particularmente útiles para hacer inferencia en situaciones en las que se tienen serias dudas sobre la satisfacción de las hipótesis que respaldan la metodología estándar.

En síntesis, una prueba estadística no paramétrica está basada en un modelo que describe solo condiciones muy generales y ninguna acerca de la forma específica de la distribución de la cual fue obtenida la muestra (comúnmente se supone que se tiene o se presenta en los datos una distribución Normal).

El presente trabajo pretende explicar la utilización correcta de las principales pruebas no paramétricas en el manejo del paquete SPSS, tratando de explicar la correcta incorporación de los datos mostrando imágenes de los pasos a seguir en cada una de las pruebas; así como la procedencia de cada uno de los valores que nos otorga la ventana Ouput del paquete.

CAPITULO I

1 HIPOTESIS CIENTÍFICA

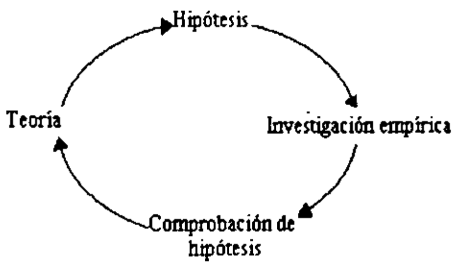
La estadística no paramétrica, está fundamentalmente apoyada en la hipótesis estadística, pero quizá la primera pregunta importante al abordar esta rama de la "Estadística Inferencial" es: ¿Qué es una hipótesis?, para poder contestar esta pregunta, es necesario abordar un tema mas general, el de "La Hipótesis Científica" que es aquella en donde esta fundamentado este concepto.

1.1 ¿QUE ES UNA HIPÓTESIS CIENTÍFICA?

Algunas personas opinan que la definición del concepto de la hipótesis científica tiene un problema de fundamentación, ya que está muy relacionada con la posición filosófica de los científicos que la utilicen, aquellos que son partidarios del empirismo consideran que sólo son confiables aquellas proposiciones que se apoyan en los datos directos de nuestros sentidos o en sus generalizaciones más simples, es decir insisten en que toda hipótesis, sea confirmada empíricamente. Es por ello que ven con desconfianza a las hipótesis considerándolas, como un medio temporal y auxiliar en la investigación, sin embargo, no resulta dificil mostrar que los resultados de las experiencias empíricas, así como sus generalizaciones más simples, constituyen solamente los comienzos de la investigación. Las experiencias empíricas requieren ser interpretadas y explicadas, lo que resultaría imposible sin las hipótesis; por otro lado, se tiene que los defensores del racionalismo se inclinan a subrayar el vínculo de las nuevas hipótesis con los conocimientos teóricos existentes, es decir las hipótesis son integradas por conceptos teóricos encontrándose estas en un plano abstracto, es por ello que con frecuencia consideran del todo aplicables, aquellas hipótesis que no puede someterse directamente a comprobación empírica alguna.

Por estos motivos los partidarios de ambas corrientes se lanzan criticas, y muchos investigadores suponen o dicen que existe un supuesto divorcio, entre las dos corrientes la teórica y la empírica, pero podemos decir que tal divorcio no existe ya que la unidad entre lo empírico y lo racional en el conocimiento, pueden ser considerados unilaterales. Como lo ha demostrado la propia historia de la ciencia, no se excluyen, sino que se complementan entre sí. Es por ello que la fundamentación teórica de las hipótesis no pueden ser contrapuestas a su comprobación por la experiencia y viceversa. Por lo anterior, se requiere estar plenamente consciente de que la teoría debe orientar la investigación empírica y ésta a su vez confirmar, reformular o anular los sistemas teóricos.

Lo dicho anteriormente, nos conduce a opinar que las hipótesis son instrumentos que desempeñan un papel fundamental en el proceso de la investigación, ya que sirven para construir un puente entre la teoría y la investigación empírica en la búsqueda de nuevos conocimientos que permitan enriquecer o ajustar los datos de la ciencia, para que así podamos encauzar y acelerar el desarrollo de la ciencias. Es necesario destacar que : "La ciencia no se reduce a registrar o a acumular simplemente hechos, sino que ante todo busca su sistematización, generalización e interpretación" a través de la comprobación de los cuerpos hipotéticos.



- Una buena analogía de la hipótesis fue planteada por Engels, al decir que: Cuando pensamos en la hipótesis, podemos verla como la forma en que se desarrollan las ciencias naturales, es decir al observar nuevos hechos que vienen a hacer imposible el tipo de explicación que hasta ahora se daba de los hechos pertenecientes al mismo grupo, se hace necesario recurrir a explicaciones de un nuevo tipo, al principio basadas solamente en un número limitado de hechos y observaciones. Hasta que el nuevo material de observación depure éstas hipótesis, elimine unas y corrija otras se podrá llegar por último a establecer la ley en toda su pureza[1].

[1] Frase retomada del libro "Metodología del conocimiento científico" pag 275. Urss-cuba

Podríamos decir que existen dos funciones principales de la hipótesis: La de generalizar y ampliar conocimientos, además de ser el punto de partida de las deducciones. La primera función se encuentra en las etapas preliminares de las investigaciones en las ciencias experimentales, las cuales frecuentemente se relacionan con los métodos inductivos de construcción de la hipótesis y la segunda función de la hipótesis es la de ser utilizada como punto de partida de las deducciones en etapas más maduras de la investigación y en las ciencias más desarrolladas. Las hipótesis en la investigación no se limitan a orientar sólo la compilación de los datos, sino además, fundamentalmente buscan establecer relaciones significativas entre fenómenos o variables, apoyándose en el conjunto de conocimientos organizados y sistematizados.

Por este sentido, la primera función de la hipótesis es servir como guía al investigador en la recopilación del material empírico (estadísticas, observaciones, entrevistas, datos de experimentos, etc.), que permita su comparación o contraste con la realidad que pretende describir, explicar y si es posible predecir, una vez que se verifique empíricamente, por este motivo la hipótesis desempeña un papel fundamental en la solución de las contradicciones entre los nuevos hechos y las viejas representaciones teóricas, pero antes de que pueda construirse una nueva teoría, la hipótesis deberá explicar aquellos hechos que no encajan con la vieja teoría o la contradigan. Es precisamente esta contradicción entre la teoría existente y los nuevos hechos descubiertos que sirven de fuente para el perfeccionamiento y desarrollo de las ideas científicas y que obliga al científico a elaborar nuevas hipótesis, leyes y teorías. Muchos descubrimientos científicos son el resultado de los intentos de eliminar las contradicciones entre las teorías existentes y los hechos reales, sin que se plantee como finalidad directa el descubrimiento de nuevos fenómenos y leyes, hasta el momento en que ella misma sea sustituida por otra hipótesis o incluida en el contenido de una teoría más general.

Sin duda, la principal función de las hipótesis en las ciencias experimentales es la de ampliar y generalizar el material empírico conocido. Los resultados de las observaciones y los experimentos siempre se relacionan con un número relativamente pequeño de fenómenos y acontecimientos, en tanto que las proposiciones de las ciencias aspiran, si no a la universalidad al menos a una elevada generalización. Con ayuda de la hipótesis, ampliamos nuestros conocimientos al extrapolar las regularidades encontradas como resultados de las investigaciones directas sobre un número finito de fenómenos, a todo el conjunto de fenómenos posibles. En algunos casos relativamente simples, esa ampliación de nuestros conocimientos se logra con el auxilio de la inducción.

Para comprender mejor la función de las hipótesis en el trabajo científico, es necesario señalar que no toda conjetura o suposición es una hipótesis científica, pues si así fuera se le restaría a ésta el poder que tiene como instrumento básico en el proceso de investigación y en el desarrollo de la teoría. Sin embargo en la hipótesis científica se presentan los elementos más esenciales de una relación entre fenómenos que existen (hipotéticamente) en la realidad objetiva. En este sentido, la hipótesis (no comprobada todavía) intenta ser una reproducción mental de la realidad objetiva, en la que se destacan a través de un proceso de abstracción aquellos fenómenos y conexiones que se consideran importantes para su formulación. Sin embargo, en lo que respecta a la búsqueda y selección de las hipótesis, no se pueden señalar procedimientos lógicos o esquemas con ayuda de los cuales pudieran encontrarse hipótesis más correctas o probables, pero esto no elimina la necesidad de fundamentación previa de las hipótesis, tanto en su aspecto teórico como práctico.

La construcción de hipótesis tiene como finalidad primordial dar una respuesta provisional, adelantar una explicación a un conjunto de hechos que no encajen en una teoría o ésta resulta insuficiente para comprenderlos y explicarlos, situación que puede definirse como un problema de investigación que obstaculiza el desarrollo del conocimiento científico. Las respuestas o explicaciones tentativas que se proporcionan al problema de investigación previamente formulado, están apoyadas en conocimientos científicos por lo que la hipótesis no puede considerarse una simple conjetura o suposición surgida del sentido común.

Por el motivo anterior, cuando una ciencia logra comprobar las hipótesis formuladas utilizando sus propios procedimientos o adecuando los de otras disciplinas, podrá integrar los hallazgos obtenidos al cuerpo de conocimientos comprobados, sin que esto signifique que tales conocimientos no sean posteriormente refutados o ajustados debido al propio desarrollo de la ciencia. Es decir, la hipótesis comprobada es un conocimiento objetivo que sirve para la confirmación, el ajuste o el rechazo de una teoría o de una parte de ésta, por lo cual puede considerársele como el motor, la fuerza propulsora principal de la ciencia para sugerir nuevos conocimientos en un proceso permanente de investigación sobre una realidad en continuo movimiento.

De acuerdo con ésto, las hipótesis en todas la ciencias tienen un carácter histórico y por lo mismo son relativas ya que su validez se circunscribe a un ámbito témporo-espacial determinado, por este motivo las hipótesis comprobadas son verdades relativas cuya perfección e incremento permiten acercarse a la verdad absoluta (como ideal).

1.2 HIPÓTESIS DE NULIDAD (HO)

Debido a que un trabajo de investigación se tiene que contrastar por medio de una hipótesis científica y a veces no es posible realizar este contraste debido a la magnitud del problema, una de las posibles formas de realizarlas es por medio de una serie de hipótesis entre las cuales se encuentra la Hipótesis de Nulidad. Pero nos debe quedar claro que esta hipótesis no es la única forma de probar o contrastar una Hipótesis Científica.

Las hipótesis nulas son en un sentido el inverso de las hipótesis de investigación (científica), las cuales constituyen proposiciones acerca de la relación entre variables que solamente sirven para refutar o negar lo que afirma la hipótesis de investigación.

La hipótesis nula (Ho) es una hipótesis de "no efecto" y por lo general se formula con el propósito expreso de ser rechazada, es decir es la negación del punto que se está tratando de probar.

1.3 HIPÓTESIS ESTADÍSTICA

La hipótesis estadística es la transformación de las hipótesis de investigación, nulas y alternativas en símbolos estadísticos. La esencia de probar una hipótesis estadística es la decidir, si la afirmación se encuentra apoyada por la evidencia experimental que se obtiene a través de una muestra aleatoria o no.

Al realizar una prueba estadística, por lo general se empieza suponiendo que falta el efecto que se quiere demostrar y se espera que los datos reunidos proporcionen la evidencia para establecer más allá de una duda razonable que el efecto en realidad está presente.

Después de haberse reunido los datos, una prueba estadística permitirá establecer el grado de probabilidad con el que se tendrían que obtener los datos, si no existiera ninguna relación. Si son extremos improbables, se concluye que la suposición inicial de no relación era incorrecta y que en realidad existe una relación. Al hacer esto, se está poniendo la ventaja a favor de la predicción, de tal modo que la evidencia tiene que ser realmente de peso antes de estar preparado para decir que la predicción es correcta.

A la suposición de que no existe ninguna relación generalmente se le llama la hipótesis nula, ya que nulifica la predicción. Al efecto que quiere demostrarse se le llama la hipótesis alternativa o hipótesis de investigación. Por tanto, la realización de una prueba estadística implica la enunciación de una hipótesis nula razonable para la cuestión que se está planteando y la determinación del grado de probabilidad con el que tendrían que obtenerse los datos si la hipótesis nula es en realidad verdadera. Sólo si fuera muy improbable obtener los datos, se concluye que la hipótesis alternativa no es falsa.

CAPITULO II

ESCALAS DE MEDICIONES

2.1 ¿QUÉ ES MEDIR?

Antes de iniciar el tema de las escalas de medición es importante primero entender y comprender los conceptos de medida y medición que nos ayudará a captar la importancia que tiene este tema.

El uso común de la idea de medida es tan natural en la conducta del hombre que a menudo pasa inadvertida, debido a que medir es algo que todos nosotros hacemos diariamente, es decir conscientemente o inconscientemente siempre estamos comparando cosas presentes a nuestra conciencia, ya que el comparar una cosa con otra es tan natural en el hombre como es la acción de respirar, por este motivo se podría decir que el lenguaje de la medida, es el lenguaje de la comparación y la medida surge de la comparación. Por lo anterior, podemos decir que la medición es el proceso de cuantificar experiencias del mundo exterior.

El científico escocés del siglo XIX, Lord Kelvin, dijo alguna vez: "Cuando uno puede medir aquello de lo que se está hablando y expresarlo en números, se sabe algo acerca de ello; pero cuando no puede medirlo, cuando no puede expresarlo en números, su conocimiento es escaso e insatisfactorio, esto podría ser un principio de conocimiento, pero escasamente este conocimiento ha avanzado la etapa de una ciencia". Aunque ésta afirmación es un poco exagerada, nos da una idea de la importancia que tiene la medición en la investigación científica.

2.2 CONCEPTO DE MEDICIÓN

El termino de medición puede definirse desde diversos puntos de vista, en un sentido amplio, pero se puede decir que medir es un proceso de asignar números a una propiedad fisica de algún objeto o conjuntos de objetos con propósitos de comparación de acuerdo a un conjunto de reglas o propiedades predeterminadas (o arbitrarias). Por medio de la medición, los atributos de nuestras percepciones se transforman en entidades conocidas y manejables llamados "números".

No obstante la medición exige la introducción de el lenguaje de conceptos cuantitativo; pero cuando iniciamos el estudio de una región de fenómenos totalmente desconocida, comenzamos por la elaboración de conceptos cualitativos, con cuya ayuda clasificamos los objetos de la región estudiada.

Después de la formación de los conceptos cualitativos y la división de todos los objetos en conjunto, podemos dar una paso mas adelante el cual consiste en establecer determinadas relaciones entre los conjuntos de objetos semejantes, con el auxilio de conceptos comparativos. Los conceptos comparativos ordenan todos los objetos de la región investigada en una determinada secuencia, en la cual cada objeto ocupa un determinado lugar. Por ejemplo, con ayuda de los conceptos "más pesado", "más ligero", "de igual peso", podemos distribuir todos los objetos en una secuencia de conjuntos.

A su vez un par de conceptos comparativos puede servir de base para la introducción de conceptos cuantitativos, es decir, conceptos que designan la cualidad medida. Por ejemplo, el par "mas pesado"- "más ligero" nos lleva al concepto cuantitativo de peso. Sin embargo, el tránsito de los conceptos cualitativos y comparativos a los cuantitativos se realiza sólo con ayuda de muchas proposiciones teóricas.

De acuerdo a lo anterior podemos ver que entre los objetos que se utilizan o se estudian estos se pueden clasificar principalmente en caracteres cualitativos – no medibles – y cuantitativos –medibles- , en donde los primeros reflejan una propiedad real con una mayor exactitud y los segundos se refieren al objeto idealizado de la teoría y sólo por eso refleja el objeto real de la teoría o sea, solo caracterizan los objetos reales.

En términos generales podemos decir que la medición es la atribución de valores numéricos a las propiedades de los objetos.

2.3 CLASIFICACION CUALITATIVA Y CUANTITATIVA

Si realizamos un análisis cuidadoso de nuestro uso de los números en la vida diaria, nos daríamos cuenta que la mayoría de los números que empleamos, no poseen las propiedades aritméticas que ordinariamente les atribuimos; esto es, no tiene sentido sumarlos, restarlos, multiplicarlos o dividirlos, por tal motivo una primera clasificación de los conceptos es la clasificación cualitativa y cuantitativa.

El resultado de observar un carácter cualitativo se le denomina Atributo o Variable[1] Cualitativa. Los elementos o individuos varían cualitativamente con respecto a una determinada característica y se pueden clasificar en categorías o modalidades. Los atributos pueden ser dicotómicos – caracterizados por la presencia o ausencia de una propiedad – o politómicos.

Para esta clasificación es necesario que las categorías cumplan con:

1. Deben de estar bien definidas.
2. Ser mutuamente excluyentes, ya que ningún elemento puede pertenecer a la vez a dos modalidades.
3. Ser exhaustivas, para que todo elemento pueda incluirse en alguna de las modalidades.

Cuando el carácter es cuantitativo, el resultado de su observación recibe el nombre de variable y las medidas de la misma, valores. Entre los caracteres susceptibles de medición se puede establecer diferencias de tipo cuantitativo y se distinguen en:

1. Variables cuantitativas discretas. Aquellas cuya medición sólo puede expresarse en números enteros, pero entre dos valores consecutivos no se puede dar ningún otro intermedio. Las enumeraciones o conteos.
2. Variables cuantitativas continuas. Cuando entre dos valores consecutivos se puede tomar cualquiera de los infinitos comprendidos entre ellos. Las medidas, por lo general, dan origen a datos continuos y su valor sólo se conoce con cierta aproximación, que refleja la precisión del instrumento de medida empleado.

En la práctica podemos a veces elegir entre medir un carácter de forma cuantitativa o cualitativa. Pero es más aconsejable optar por la medida cuantitativa , ya que esta siempre da más información. Dependiendo del tipo de variable uno puede asociar a su vez una escala de medición, pero para poder realizar correctamente esta asociación uno tiene que revisar antes cuales son las propiedades que debe cumplir una medida.

2.4 PROPIEDADES DE LA MEDIDA

En el campo de la matemáticas abstractas, conocido como teoría de la medida, una asociación numérica es una medida solamente si tiene las siguientes propiedades matemáticas, también denominadas "reglas de medición". El resultado de la medición puede expresarse de la siguiente manera:

$$Q = qU$$

Donde: Q es la propiedad medida[2]
U la unidad de medición
q el valor numérico de la magnitud correspondiente

[1] Variable: son características, atributos, cualidades, rasgos o propiedades comunes de los elementos de una población o una muestra para propósitos de comparación como: sexo, peso, nacionalidad, No. de hijos etc.
[2] Una medida es una función, no toda función es una medida. Cualquier función que tiene las propiedades particulares antes descritas se llama medida.

Esta ecuación se conoce como "la ecuación fundamental de la medición". Para atribuir un determinado valor numérico a la propiedad medida de acuerdo con esta ecuación, tendremos que guiarnos por las siguientes normas:

A.- El término técnico para esta propiedad es el de "Aditividad finita" o "Regla de adición" la cual expresa que: el valor numérico de la suma de dos valores físicos de las propiedades, deberá ser igual a la suma de los valores físicos de esta propiedad. Simbólicamente:

$$qU(Q_1 \text{ o } Q_2) = q_1U + q_2U$$

En la formulación de esta regla, entre Q_1 y Q_2 situamos el símbolo "o" que indica la operación empírica de la unión de dos grados en una misma propiedad. Es natural que esta operación pueda diferenciarse de una suma aritmética. La operación de unión de dos grados diferentes de una propiedad que no siempre está sometida a la regla indicada[3].

En el lenguaje común decimos que "la medida del conjunto debe ser igual a la suma de las medidas de todas sus partes". Por ejemplo, si tuviésemos que pesar una docena de naranjas, encontraríamos que el peso es el mismo si pesáramos cada naranja separadamente y luego añadiésemos todos los pesos, o que si las pesáramos a todas juntas.

B.- Regla de equivalencia: si el valor físico de las propiedades medidas es igual, iguales deberán ser sus expresiones numéricas. Simbólicamente:

Simbólicamente si $Q_1 = Q_2$, entonces $q_1U = q_2U$.

Podríamos decir que esta propiedad realmente esta implicada por la propiedad A, Un caso particular es que la medida de "nada" o ninguno, debe ser 0. El concepto puede que parezca menos ridículo si pensamos en la medida como en un conteo. El número de elementos en un conjunto es una clase de medida del conjunto. Podemos decir con sentido que el número de elementos del conjunto vacío es el 0.

C.- A esta propiedad se le llama "monotonía", y nos dice que si el valor físico de la propiedad de un cuerpo es menor que el valor físico de esta misma propiedad en otro cuerpo, entonces el valor numérico del primero deberá ser menor que el del segundo.

Simbólicamente sí $Q_1 < Q_2$, entonces $q_1U < q_2U$.

Expresando lo anterior de otra manera tenemos que la medida de una parte del algo no deber ser mayor que la medida del todo. Por ejemplo, el peso de medio paquete de mantequilla es menor que el de todo el paquete. Teniendo en cuenta la propiedad A y B, la propiedad C es equivalente a afirmar que las medidas están expresadas por números no negativos (es decir el conjunto de los reales positivos).

[3]No todas las mediciones físicas concuerdan con esta descripción, ejemplo: la temperatura, no es finitamente aditiva en ningún sentido inmediato.

D.- La regla de la unidad de medida: Nos dice que si se llega a seleccionar un determinado cuerpo o un proceso natural fácilmente reproducible y caracterizar la unidad de medida por medio de este cuerpo o proceso este deberá conservar inmutables su medida, forma, periodicidad, etc. Pero los cuerpos y procesos reales están sometidos a modificaciones debido a la influencia de las condiciones circundantes. Por ello, como patrones reales se toman aquellos cuerpos y procesos que son más estables respecto a las condiciones externas.

En cualquier experimento de medición, si la medida se hace de cierto modo bajo determinadas condiciones físicas prescritas, al repetir el experimento en las mismas condiciones se debe obtener resultados iguales.

2.5 TIPOS DE ESCALAS

Es importante señalar que los números se usan de varias formas para lograr objetivos diferentes. Muchas veces estos objetivos no incluyen la representación de una cantidad, de hecho, hay cuatro formas fundamentales distintas de utilizar los números:

1. Identidad: los números pueden servir como etiquetas para designar o identificar, artículos o clases (números nominales).
2. Orden: los números pueden servir para representar la posición de una serie o para indicar el rango ordenado de los artículos (números ordinales).
3. Intervalos: los números pueden servir para indicar las diferencias entre los artículos (números reales).
4. Proporciones: los números pueden servir para indicar las proporciones entre los artículos (números reales).

Los diferentes niveles de medidas representan distintos niveles de información numérica contenida en un conjunto de observaciones (datos). Las clases de operaciones matemáticas que se pueden efectuar legítimamente con los números determinan cual de las cuatro clases de escalas pertenecen.

2.5.1 Escala Nominal.

La medición de una escala nominal simplemente consiste en situar a cada individuo en una u otra categoría dada (o el asignarles un nombre). Esta escala de medición trata de agrupar objetos en clases, de modo que todos los que pertenezcan a la misma sean equivalentes respecto del atributo o propiedad en estudio.

El hecho de que a veces, en lugar de denominaciones se les atribuyan números, pueden ser una de las razones por las cuales se les conoce como "medidas nominales". Los números asignados a las escalas nominales tienen iguales propiedades que los demás, pero en ningún momento podemos pensar en el manejo de orden, tamaño y otras propiedades de las cifras, puesto que no se deben tomar como tales, es decir estos números carecen o no tienen propiedades cuantitativas y sirven únicamente para identificar las clases.

Lo que nos lleva a decir que los datos empleados con las escalas nominales consisten en conteos de frecuencias o tabulaciones del número de sucesos en cada clase de la variable en estudio. Tales datos reciben indistintamente los nombres de: *datos de frecuencias, datos enumerativos, datos de atributos o datos de categorías*. Las únicas relaciones matemáticas adecuadas a las escalas nominales son las de equivalencia (=) o no equivalencia ($\neq$). Así, una persona u objeto particular tiene la característica que define la clase (=) o no la tiene ($\neq$). Este tipo de medición constituye el nivel de medición mas bajo.

2.5.2 Escala Ordinal

Para los casos en que se pueden detectar diversos grados de un atributo o propiedad de un objeto, la medida ordinal es la indicada, puesto que la medición de una escala ordinal supone situar a los individuos en un orden, ordenarlos de acuerdo con algún criterio.

Los datos ordinales constituyen un escalón superior en relación a los datos nominales, porque nos permiten decir si un individuo está antes o después que otro es una escala. Así no solo sabemos que los datos son diferentes entre sí (característica que define a las escalas nominales) sino que se mantiene alguna clase de relación entre ellas. Es de señalarse que los números pueden asumir el lugar de los objetos en estudio, puesto que los números son representaciones parciales de éstos. Lo que nos lleva a plantear que los números pueden tratarse como si fueran diferentes , es decir se pueden ordenar.

A pesar de que no existe ley alguna que prohíba sumar, restar, multiplicar, etc., números asignados según escalas ordinales, el resultado de tales operaciones puede no indicar nada respecto del grado de atributo en cuestión que el objeto en estudio posee. Los resultados de estos cálculos aritméticos no pueden informar absolutamente nada respecto del atributo real inherente al objeto, es decir los numerales empleados en conexión con las escalas ordinales no son cuantitativos. Ellos indican solamente la posición en una serie ordenada y no "cuanta" diferencia existe entre posiciones sucesivas en la escala.

Más específicamente, las operaciones que se le están permitidas o que tiene algún significado para éste tipo de escala es la relación que se expresa en términos del álgebra de las desigualdades es decir: a es menor que b ($a < b$) o a es mayor que b ($a > b$).

2.5.3 Escala Intervalar

Cuando no solamente es posible distinguir la diferencia entre los diversos grados de propiedad de un objeto (característica de la medida ordinal) sino que también puede discernirse las diferencias entre objetos iguales, se recurre a la medida de intervalo.

En este caso, una unidad de medida se define en términos de algún parámetro (grado, pulgada, pie, etc.). Es decir, la medición en una escala de intervalo consiste en asignar un número a un individuo para indicar su posición exacta a lo largo de una escala continua. Los datos de intervalo ocupan otro escalón superior en la jerarquía de escalas de medición, porque nos permiten decir qué distancia separa a un individuo de otro dentro de una escala.

Las medidas de intervalo implican la asignación de números de modo tal, que a iguales diferencias entre los grados del atributo estudiado en un objeto, correspondan iguales diferencias entre los números. Una de las características distintivas de la medida de intervalos es que el cero no necesariamente implica que el objeto carece del atributo, puesto que en una escala de intervalo, el punto cero es puramente arbitrario. Es decir el punto cero de la escala de intervalo puede asignarse arbitrariamente y en ningún caso indica ausencia completa de la propiedad en cuestión.

Los valores numéricos asociados con estas escalas son efectivamente cuantitativos y por lo tanto, permiten el uso de operaciones aritméticas, tales como suma, resta, multiplicación y división, además que poseen la propiedad de distintividad y orden, es decir en este caso, la diferencia entre los número sí es significativa.

2.5.4 Escala de razón o cociente

La medida de razón o cociente se diferencia de la de intervalo únicamente en que, el punto cero no es arbitrario, sino un valor absoluto y corresponde realmente a una total ausencia de la propiedad estudiada, es decir si tenemos una longitud igual a cero significa que no hay longitud, cuando se observa una carencia total de propiedad, se dispone de una unidad de medida para tal efecto. A iguales diferencias entre los números asignados, corresponden iguales diferencias en el grado de atributo presente en el objeto de estudio.

Por todo lo anterior podemos decir que las mediciones en una escala de razón tiene todas las características de las mediciones de intervalo, pero con el rasgo adicional de que la razón de dos valores cualesquiera, es independiente de la unidad de medición, por ejemplo, 4 metros es a 2 metros como 2 metros es a 1 metro.

Al analizar la escala de intervalo y de razón tenemos que estos dos tipos de escalas se basan en números reales y este tipo de números nos indican el nivel más alto de las mediciones científicas. Los valores numéricos asociados con estas escalas son efectivamente cuantitativos y por tanto, permiten el uso de operaciones aritméticas tales como suma, resta , multiplicación y división, es decir diferencias iguales entre puntos de cualquier parte de la escala son iguales entre si.

Una de las características de las escalas de orden superior es que se les puede transformar fácilmente en escalas de orden más bajo. Así el resultado de una carrera de una milla se puede expresar en unidades de tiempo (escala de razones). Los tiempos se pueden transformar en datos de una escala ordinal, por ejemplo primero, segundo y tercer lugar en el orden de llegada. Sin embargo, no es posible efectuar la transformación inversa. Si por ejemplo, conocemos únicamente el orden de llegada en una carrera, no podemos expresar los resultados en términos de una escala de razones (tiempo). Aunque es admisible transformar las marcas de una escala de mayor nivel a otra de menor nivel, no es recomendable hacerlo, por lo general, ya que se pierde información cuantitativa.

En la mayoría de las investigaciones educacionales, psicológicas o de ciencias del comportamiento, las mediciones se efectúan según escalas nominales, ordinales o de intervalo, puesto que hay pocas variables de importancia que den pie para medirse, según en escalas de razón, de hecho el encontrar una escala de intervalo satisfactoria ya es de por si problemático. Aunque es discutible que muchas de las mediciones efectuadas por los científicos del comportamiento lleguen al nivel de una escala de intervalo, ellos están de acuerdo en suponer que así es.

S.S. Stevens presenta la siguiente tabla resumen de las propiedades de las escalas de medición.

Escala	**Relaciones definidas**	**Estructura matemática de grupo**	**Estadísticas permisibles**	**Ejemplos Típicos**	**Pruebas Estadísticas apropiadas**
Nominal	Equivalencia o pertenencia a una categoria	Grupo de permutación $x'=f(x)$ ($f(x)$ significa cualquier sustitución biunívoca)	Frecuencia Moda Coeficiente de contingencia	Enumeración de los jugadores de fútbol. Sexo: Masculino, Femenino	Pruebas estadísticas no paramétricas
Ordinal	Equivalencia. Determinación de mayor o menor .	Grupo isotónico $x'=f(x)$ ($f(x)$ significa cualquier función monótona creciente)	Mediana Percentiles Spearman r, Kendall r Kendall W	Pruebas de inteligencias. Calificaciones como NA, B, MB.	Pruebas estadisticas no paramétricas
Intervalar	Equivalencia Determinación de mayor o menor. Proporción conocida de un intervalo a cualquier otro.	Grupo lineal general $x'=ax+b$	Media Desviación estándar Correlación del momento producto de Pearson . Correlación del múltiple momento-producto.	Temperatura Posición en una linea. Pruebas de inteligencia, "resultados estandarizados"	Pruebas estadísticas paramétricas
Razón ó Proporción	Equivalencia Determinación de mayor o menor. Proporción conocida de un intervalo a cualquier otro. Proporción conocida de una valor de la escala a cualquier otro.	Grupo de similitud $x'=ax$	Media geométrica Media armónica. Coeficiente de variación.	Longitud, enumeración, densidad, intervalos de tiempo, etc.	Pruebas estadisticas paramétricas

CAPITULO III

ESTADISTICA PARAMETRICA Y NO PARAMETRICA

En general existen numerosas técnicas estadísticas para analizar cualquier conjunto de datos. A pesar de que el presente trabajo esta enfocado a las pruebas conocidas como no paramétrica o como de libre distribución, este no puede iniciarse sin conocer las diferencias entre las pruebas paramétricas y no paramétricas.

3.1 PRUEBAS ESTADÍSTICAS PARAMÉTRICAS.

Muchas veces el objetivo de la estadística es hacer inferencia con respecto a parámetros poblacionales desconocidos, basadas en la información obtenida mediante datos muestrales. Estas inferencias se expresan en una de dos maneras, como estimaciones de los parámetros respectivos o como prueba de hipótesis (pruebas paramétricas) referentes a sus valores.

En muchos aspectos el procedimiento formal para la prueba de hipótesis es similar al método científico. El científico observa la naturaleza, establece una teoría y después prueba su teoría con relación a lo observado. En este contexto el científico propone una teoría relativa a los valores específicos de uno o más parámetros poblacionales. Luego obtiene una muestra de la población y compara la observación con la teoría. Si la observación se contrapone a la teoría, el científico rechaza la hipótesis. En caso contrario concluye que la teoría es válida o bien que la muestra no detecta la diferencia entre los valores reales y los valores de la hipótesis respecto de los parámetros poblacionales.

Por tal motivo, una prueba estadística paramétrica es aquella cuyo modelo especifica ciertas condiciones acerca de los parámetros de la población de la que se obtuvo la muestra, que no se prueban ordinariamente, sino se supone que se mantienen. La significación de los resultados de una prueba paramétrica depende de la validéz de estas suposiciones.

Las pruebas paramétricas (pruebas de hipótesis) utilizan estadísticos tales como T(comparaciones de medias), χ^2 o F(ANOVAS), para rechazar o no rechazar la hipótesis del problema , los supuestos necesarios para aplicar estas pruebas clásicas son los siguientes:

- Las observaciones deben ser obtenidas de una muestra aleatoria.
- Las observaciones deben de ser obtenidas de poblaciones normalmente distribuidas.
- Las poblaciones de donde se adquieren las observaciones, deben tener la misma varianza, o en algunos casos deben tener una proporción de varianza conocida.
- Las variables correspondientes al estudio, deben ser medidos al menos en una escala de tipo intervalar, para poder realizar medidas de tendencia central, así como medidas de dispersión y pruebas de hipótesis.
- Estas condiciones generalmente no son probadas, pero son determinadas y aceptadas por el tipo de estudio, y su veracidad o falsedad determinará la significación de las pruebas paramétricas.

Ventajas de las pruebas paramétricas:

Las pruebas paramétricas son más eficaces si los datos cumplen todas las suposiciones de su modelo estadístico.

Desventajas de las pruebas paramétricas:

Si las pruebas paramétricas no cumplen los supuestos, los resultados obtenidos no serán confiables, esto depende principalmente de cual de todos los supuestos no se cumplen.

3.2 PRUEBAS ESTADÍSTICA NO PARAMÉTRICA

Cuando la población en muestra no cumple con los supuestos que se requieren en las pruebas paramétricas, y se necesita realizar inferencias en esta población, se recurre a la estadística no paramétrica que no depende de estos rígidos supuestos.

Es decir una prueba estadística no paramétrica, es aquella cuyo modelo no especifica las condiciones de los parámetros de la población de la que se sacó la muestra, ni requiere mediciones tan fuertes, ya que éstas se aplican a datos de una escala ordinal y algunas a los de una escala nominal. En las pruebas no paramétricas, la potencia de estas depende del tamaño de la muestra, ya que entre más grande sea la muestra, la prueba es mejor.

Ventajas de las pruebas no paramétricas

- Los procedimientos no paramétricos dependen de un mínimo de supuestos.
- Las probabilidades obtenidas son aproximadas, sin tomar en cuenta la distribución de la población.
- Dentro de la estadística no paramétrica, pueden ser estudiadas observaciones de poblaciones diferentes, en cambio dentro de la estadística paramétrica este estudio es imposible.
- Los procedimientos no paramétricos, pueden ser aplicados cuando los datos son simplemente clasificatorios medidos en una escala nominal, o para datos inherentes a los rangos como datos cuyos puntajes aparentemente numéricos tienen la fuerza de rango.
- Las pruebas estadísticas no paramétricas son típicamente mucho más fáciles de realizar, ya que se requiere un mínimo de preparación en matemáticas y estadística (Esto no implica que su fundamentación en las matemáticas no fuerte).
- Hay procesos no paramétricos que no tienen equivalente paramétrico.

Desventajas de las pruebas no paramétricas

- Para muestras muy pequeñas no se puede utilizar la estadística no paramétrica al menos que se conozca el tipo de distribución de la población.
- Las pruebas no paramétricas manejan mediciones no exactas en cuanto a cantidades, es decir pueden dar resultado dentro de un rango, como mejor o peor, pero no con la exactitud en cuanto a cual es el mejor o peor, al no contemplar la magnitud de las diferencias, se desperdicia mucha información.
- Si todos los supuestos para aplicar una prueba paramétrica son cumplidos en gran parte, la estadística no paramétrica resulta ser menos potente, por tal motivo es preferible utilizar la estadística paramétrica.
- Para poder realizar pruebas estadísticas no paramétricas, es necesario la utilización de tablas, las cuales se encuentran en muy diversas formas y publicaciones en todos los niveles.

Una observación importante es que uno puede utilizar en lugar de la estadística paramétrica, la estadística no paramétrica, claro que en este cambio se pierde información o potencia en el proceso, pero es bueno aclarar que no se puede pasar de la estadística no paramétrica a la paramétrica.

CAPITULO IV

USO DE LAS PRUEBAS ESTADISTICAS

Las pruebas estadísticas (paramétricas o no paramétricas) tienen, el propósito de probar hipótesis estadísticas que parecen importantes para ciertas teorías, recabando datos que nos permitan decidir acerca de esas hipótesis. Nuestra decisión puede conducirnos a sostener, revisar o rechazar la hipótesis y la teoría de la cual se originó, destacándose que se debe llegar a conclusiones por medio de métodos que sean del dominio público y que puedan ser repetidos por otros investigadores.

El objetivo de este procedimiento, debe estar basado en la información o datos que se obtienen de la investigación y el riesgo que estemos dispuestos a correr de que nuestra decisión acerca de la hipótesis sea incorrecta.

El procedimiento que generalmente se sigue para realizar las pruebas estadísticas consiste en los siguientes paso:

- Formulación de las hipótesis: Nula y Alternativa.
- Elección del estadístico de prueba.
- Distribución del estadístico de prueba.
- Nivel de significancia (α)
- Identificación de la región de rechazo ó valor crítico y regla de decisión.

Estos pasos se siguen tanto en estadística paramétrica (pruebas de hipótesis), como en estadística no paramétrica (pruebas no paramétricas).

Todos y cada uno de estos pasos tienen su función y su importancia y para comprender en que consisten, se profundizará en cada uno de ellos.

4.1 FORMULACIÓN DE LAS HIPÓTESIS : NULA Y ALTERNATIVA.

En este primer paso se realiza la formulación de la hipótesis nula (Ho) y la alternativa (Ha). En donde la hipótesis nula (Ho), es una hipótesis de "no efecto" y por lo general se formula con el propósito expreso de ser rechazada, es decir, es la negación del punto que se está tratando de probar. Si es rechazada, se apoya la hipótesis alternativa (Ha).

La hipótesis alternativa (Ha) es la declaración operacional de la hipótesis de investigación[1] del experimentador. La naturaleza de la hipótesis de investigación determina como se debe establecer la hipótesis alternativa (Ha), la cual puede ser direccional o no direccional; es decir si la hipótesis de investigación afirma solamente que el parámetro es diferente al de la hipótesis nula tenemos que se trata de una hipótesis no direccional o bilateral, un ejemplo seria aquel en el que la hipótesis de investigación simplemente establece que dos grupos difieren respecto a la media, donde tendríamos que la hipótesis de la prueba estaría definida de la siguiente manera:

$$\text{Ho: } \mu_1 = \mu_2 \quad \text{vs.} \quad \text{Ha: } \mu_1 \neq \mu_2$$

En el caso en que se indique que el parámetro es diferente y además se señale la dirección de la diferencia la hipótesis alternativa (Ha) tendríamos una hipótesis direccional o unilateral, si volvemos a tomar el ejemplo de las medias en donde un grupo especifico tenga una media mayor que el otro, entonces Ha pudiera ser que $\mu_1 > \mu_2$ o que $\mu_1 < \mu_2$, esto es, la media del grupo 1 es mayor que o menor que la media del grupo 2 respectivamente, la hipótesis de prueba estaría definida de la siguiente manera:

$$\text{Ho: } \mu_1 = \mu_2 \quad \text{vs.} \quad \text{Ha: } \mu_1 > \mu_2 \quad \text{ó} \quad \text{Ha: } \mu_1 < \mu_2$$

Si realizamos un análisis cuidadoso de la lógica de la inferencia estadística, está nos revelaría que la hipótesis nula no puede probarse nunca, ni tampoco podemos probar directamente la hipótesis alternativa. Sin embargo, si podemos rechazar la hipótesis nula y de esta forma se estaría apoyando la hipótesis alternativa en forma indirecta. Por este motivo no podremos probar nunca la hipótesis nula rechazando la hipótesis alternativa. La afirmación más fuerte que podemos hacer a este respecto es que fracasamos en rechazar la hipótesis nula.

4.2 ELECCIÓN DEL ESTADÍSTICO DE PRUEBA.

La finalidad de este inciso es la selección de una prueba estadística (con su modelo estadístico asociado) para poder rechazar o no rechazar la Ho y se debe escoger aquella cuyo modelo se aproxime más a las condiciones de la investigación y cuyos requisitos de medición satisfagan las medidas usadas en está.

---

[1] La hipótesis de investigación es la predicción derivada de la teoría sometida a prueba.

En la elección se debe considerar la manera en que se obtuvo la muestra de la población, la naturaleza de la población en donde se obtuvo la muestra y la clase de medición o escala que se empleó en las definiciones operacionales de las variables usadas, es decir en los puntajes.

Como el objetivo de este trabajo es utilizar las pruebas no paramétricas, tenemos que éstas comúnmente se clasifican por la cantidad y el tipo de muestras, de la siguiente forma:

- Por la cantidad: Las muestras no paramétricas se dividen en una o dos muestras y en k muestras.
- Por el tipo: Si los datos son independientes ó relacionados (dependientes).

  a) Independientes: Si se supone que no ha habido influencias de mediciones entre las variables.
  b) Relacionados: Si las observaciones son repetidas en los mismos sujetos o cuando los sujetos son apareados de alguna manera que se le considere pertinente (cuidando los factores de confusión).

Tabla 1

Características de la muestra	Prueba estadística	Tipo de variable
Una o dos muestras relacionadas	Binomial Signos Wilcoxon	Ordinales Ordinales
Una o dos muestras independientes	Mediana U Mann-Whitney Tukey	Al menos Ordinal Al menos Ordinal
K muestras relacionadas	Varianza de dos clasificaciones de rango de Friedman Prueba para ordenaciones alternativas.	Por lo menos Ordinal
K muestras independientes	Extensión de la prueba de la Mediana Kruskal-Wallis	Por lo menos ordinal
Asociación	Tablas de Contingencia Spearman	Nominal ó Ordinal

4.3 DISTRIBUCIÓN DEL ESTADÍSTICO DE PRUEBA

Después de haber elegido un estadístico de prueba, el siguiente paso será determinar la distribución muestral de este estadístico. La distribución muestral de un estadístico señala la probabilidad conforme a Ho, que está asociado con los diferentes posibles valores numéricos del estadístico. La probabilidad "asociada con" la ocurrencia de un valor particular del estadístico conforme a Ho no es la probabilidad exacta de ese valor.

La distribución muestral es una distribución teórica y puede ser obtenida al tomar al azar todas la muestras posibles de un mismo tamaño extraídas de una misma población, es decir, la distribución muestral es la distribución conforme a Ho de todos los valores posibles que un estadístico puede tomar. Cuando es calculada con muestras de igual tamaño tomadas al azar, este procedimiento sería casi imposible de realizar, pues se consideran todos los casos posibles al anotar las distribuciones muestrales, aunque fueran muestras de un tamaño moderado, por lo que para considerar una distribución es necesario utilizar las herramientas que nos ofrecen los teoremas concernientes a distribuciones normales y tamaños de muestra, uno de ellos el Teorema del Limite Central.

A partir de la suposición de que los datos se distribuyen normalmente, se obtienen distribuciones para parámetros desconocidos como: la t de Student, la F de Fischer, y la χ^2 que se utilizan principalmente en estadística paramétrica (cuando el tamaño de muestra es pequeño) se usa la Z de la normal estándar (cuando el tamaño de muestra es grande) y se utilizan aproximaciones a estas distribuciones en la estadística no paramétrica.

4.4 NIVEL DE SIGNIFICANCIA (α)

El nivel de significancia (α)se refiere a la probabilidad de cometer un error de tipo I, el cual consiste en rechazar Ho cuando Ho es verdadera, este se puede interpretar como el nivel máximo que el investigador esta dispuesto a equivocarse. Algunos autores, definen el nivel de significancia como "el valor de riesgo α elegido para definir la región crítica, es decir, la zona de rechazo de la hipótesis nula". Cuando el resultado de una prueba estadística está dentro de la región critica se dice que es significativo.

El nivel de significancia se determina por el tipo de situación que se esté estudiando, pues en algunos casos se tiene que fijar un nivel más estricto(mayor) para poder verdaderamente probar una hipótesis, comúnmente éste se fija antes de realizar la prueba.

Íntimamente ligado a "α" se encuentra el riesgo "β" que se define como la probabilidad de cometer un error de tipo II, que es el de no rechazar Ho cuando Ho es falsa. Así pues, existen dos tipos de error a tomar en cuenta:

Tabla 2

Decisión	Estado real de Ho.	
	Ho Verdadera	**Ho Falsa**
No rechazar Ho	Correcta $1-\alpha$	Error de tipo II β
Rechazar Ho	Error de tipo I α	Correcta $1-\beta$

Tipo de error cometido, teniendo en cuenta el estado real de Ho y la decisión estadística tomada.

La probabilidad de cometer el error de tipo I esta dada por α. Cuando mayor sea α es más probable que Ho sea rechazada equivocadamente, es decir, es más probable que se cometa el error de tipo I, pero si tomamos una α menor, mayor será la probabilidad de cometer un error de tipo II.

$$P(\text{error tipo I}) = \alpha$$

Por otro lado como "La probabilidad de cometer un error de tipo II está en parte determinada por la hipótesis opuesta a la que se contrasta", existen tantos errores como hipótesis opuestas haya y cuanto más cerca de la hipótesis nula esté la hipótesis correcta, mayor es la probabilidad de cometer un error de tipo II, éste tipo de error es más frecuente que del tipo I.

$$P(\text{error tipo II}) = \beta$$

Entonces en resumidas cuentas, cuanto menor sea el nivel de rechazo, menor será la probabilidad de cometer un error de tipo I y mayor será la probabilidad de cometer un error de tipo II. A la inversa, cuanto mayor sea el nivel de rechazo, mayor será la probabilidad de cometer un error de tipo I y menor será la probabilidad de cometer un error de tipo II. De los dos tipos de errores antes mencionados algunos investigadores prefieren cometer un error de tipo I. Sin embargo, ello no debe llevarnos a la idea de que no importa la posibilidad de cometer errores de tipo II, sino encontrar alguna forma para tratar de minimizar estos dos errores.

4.5 IDENTIFICACIÓN DE LA REGIÓN DE RECHAZO Ó VALOR CRITICO Y REGLA DE DECISIÓN.

-valor crítico

El valor crítico de un estadístico de prueba, es el valor que está al extremo de la probabilidad de obtener éste o un valor más extremo, cuando Ho es verdadera, el cual es representado por α.

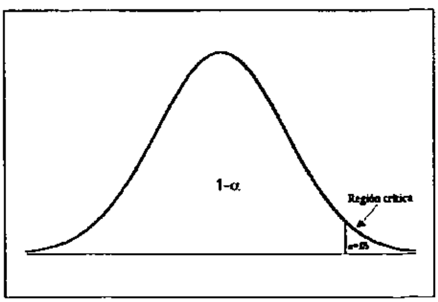
A partir del valor crítico se define la región de rechazo, la cual incluye todos los valores posibles que una prueba estadística puede tomar conforme a Ho, es decir ésta región consiste en un conjunto de valores posibles tan extremos que, cuando Ho es cierta, es muy pequeña la probabilidad (α) de que la muestra observada produzca un valor que esté entre ellos.

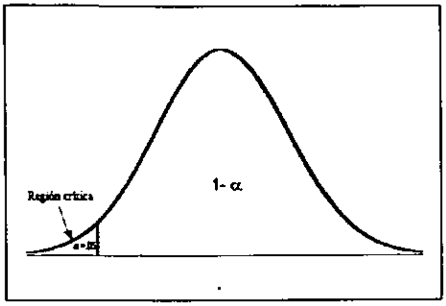
La localización de la región de rechazo es afectada por la naturaleza de Ha, es decir:

Prueba de una cola o unilateral:
Esta se presenta cuando Ha indica la dirección predicha de la diferencia.

A.- El área sombreada muestra la región de rechazo de una cola $\alpha = 0.05$ con p=0.05

$$Ho: \theta = \theta_o \quad \text{vs.} \quad Ha: \theta > \theta_o$$





ó también:

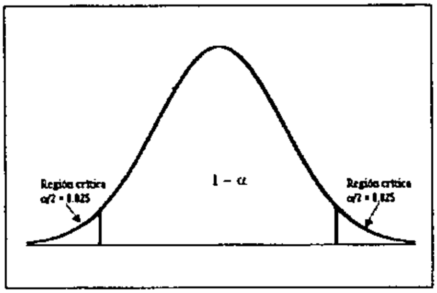
B.- El área sombreada muestra la región de rechazo de una cola $\alpha = 0.05$ con p =0.05

$$Ho: \theta = \theta_o \quad \text{vs.} \quad Ha: \theta < \theta_o$$

Prueba de dos colas o bilateral:
Esta se presenta cuando Ha no indica la dirección predicha.

C.- El área sombreada muestra la región de rechazo de dos cola $\alpha = 0.05$ con p =0.025

$$Ho: \theta = \theta_o \quad \text{vs.} \quad Ha: \theta \neq \theta_o$$



Las pruebas de una o dos colas se distinguen en la localización (pero no en el tamaño) de la región de rechazo. Esto es, en una prueba de una cola, la región de rechazo está totalmente en un extremo (o cola) de la distribución muestral. En una prueba de dos colas, la región de rechazo está en ambos extremos de la distribución muestral.

La región de rechazo queda expresado por α que es el nivel de significancia. Si α=0.05 entonces el tamaño de la región de rechazo es del 5% del área total comprendida bajo la curva de distribución muestral.

Regla de decisión

Esta regla de decisión se toma en termino del valor crítico. En una prueba de una cola por ejemplo, la regla de decisión nos dice que se rechaza Ho si el valor obtenido en el estadístico de prueba es tan extremo como o más extremo que (menor o igual) al valor crítico. En una prueba de dos colas se tienen dos valores críticos y se rechaza Ho si el valor obtenido en el estadístico de prueba es tan extremo como uno u otro de los dos valores críticos especificados.

Si la prueba estadísticas da un valor que está en la región de rechazo, entonces se rechaza Ho, ya que si es muy pequeña la probabilidad asociada con la ocurrencia conforme a la hipótesis nula de un valor particular en la distribución muestral, podemos decir que la hipótesis nula es falsa.

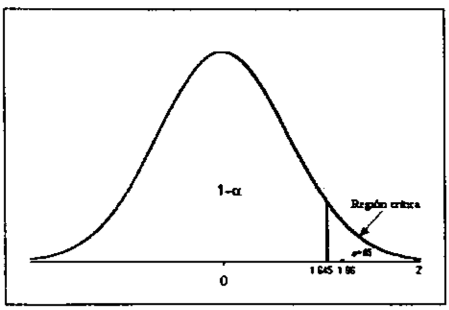
En conclusión cuando un valor observado en una prueba estadística es igual o menor que el valor previamente determinado como α decimos que Ho es falsa. El valor observado es llamado previamente significativo. La hipótesis nula es rechazada siempre que se encuentre un valor significativo.

Ejemplo:

Supóngase que el nivel de significancia es $\alpha = 0.05$ y que la hipótesis nula y la alternativa son:

$$\text{Ho}: \mu_1 = \mu_2 \quad \text{vs.} \quad \text{Ho}: \mu_1 > \mu_2$$

En donde el estadístico de prueba es una distribución normal con un valor de $z = 1.86$, al ver a éste valor z en las tablas de normal (0,1) para una prueba de una cola con el nivel de significancia antes mencionado, tenemos que éste vale $z = 1.645$.
Entonces como 1.86 es mayor que 1.645, tenemos que se rechaza Ho.



Esta imagen nos muestra la posición de valor obtenido en el estadístico de prueba (1.86) y el valor crítico (1.645).

4.6 P-VALUE

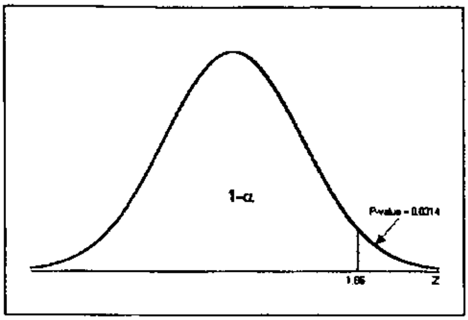
Otra manera de decidir si la muestra de datos rechaza o no la hipótesis nula es determinando la probabilidad observada, cuando Ho es verdadera, a un valor de la estadística de prueba que es por lo menos tan extrema que el valor actual observado. Esta probabilidad es nombrada de diferentes formas, las más comunes son: nivel crítico, la descripción del nivel de significancia, el de la probabilidad asociada y P-value.

El valor de P-value puede ser reportado como un valor exacto o como intervalo, dependiendo de la naturaleza del valor obtenido en tablas de la distribución de la prueba estadística.

Ejemplo:

Supóngase que el nivel de significancia es $\alpha = 0.05$ y que la hipótesis nula y la alternativa son:

$$\text{Ho}: \mu_1 = \mu_2 \quad \text{vs.} \quad \text{Ho}: \mu_1 > \mu_2$$



En donde el estadístico de prueba es la distribución normal con un valor de $z = 1.86$, y como P(z)=P(1.86)=0.05-0.04686=0.0314. Esto es, la probabilidad de observar un valor de z tan largo o más largo es de 1.86, cuando Ho es verdadera, es de 0.0314. Entonces el valor de P-value es 0.0314 y como este es menor que 0.05, se rechaza Ho.

Para determinar el valor de P-value en valores de dos colas en distribuciones simétricas como la normal y la t de Student, no implica ningún problema ya que éste se obtiene duplicando el valor de P-value obtenido por medio del estadístico de prueba.

Ejemplo:

Supóngase que el nivel de significancia es $\alpha = 0.05$ y que la hipótesis nula y la alternativa son:

$$Ho: \mu_1 = \mu_2 \quad \text{vs.} \quad Ha: \mu_1 \neq \mu_2$$

Al obtener el valor del estadístico de prueba de una distribución normal, con un valor de $z = 1.86$, y como en este caso tenemos una prueba de dos colas éste se obtendría por medio de 2P(z)=2P(1.86)= 2(0.05 - 0.04686)=2(0.0314)=0.0628 como éste es mayor que 0.05, no se rechaza Ho.

Sin embargo para determinar el P-value en una prueba de dos colas que tiene una distribución asimétrica se presenta un problema, ya que el valor de P-value obtenido para una prueba de una cola no puede multiplicarse por dos, ya que éste valor puede llegar a valer más de 1, por tal motivo Gibbons y Patt, proponen reportar el valor del P-value para una prueba de una cola y estacionar la dirección de la observación por medio de la hipótesis nula.

Es decir, supóngase que tenemos:

$$Ho: \mu_1 = \mu_2 \quad \text{vs.} \quad Ho: \mu_1 \neq \mu_2$$

y que la prueba estadística apropiada es χ^2 con 1 grado de libertar y $\alpha = 0.05$. Si el valor obtenido es de 3.84, se puede asumir el resultado en términos de P-value de la siguiente forma:

$$P(\chi_1^2 \geq 3.84) = 0.00156555$$

Con $P < \alpha = 0.05$, podemos concluir que se rechaza la hipótesis nula. Ya que el valor del P-value nos da mayor información que el declarar "la prueba fue significativa al nivel de 0.05" o "Ho puede ser rechazada al nivel del 0.01". Es este el mejor argumento a favor de reportar P-value.

CAPITULO V

INTRODUCCIÓN AL PAQUETE SPSS

5.1 INTRODUCCIÓN

A pesar que solamente se va a explicar una de las opciones de "Statistics" la parte concerniente a la estadística no paramétrica del paquete SPSS, es importante conocer la historia e importancia que tiene este paquete.

Es probable que muchos de los lectores hayan aprendido a programar en algún lenguaje de alto nivel, como FORTRAN, BASIC o el más reciente C++, y como es sabido la escritura y depuración de un programa toma cierto tiempo, pero se justifica cuando es la única forma de realizar algún tipo de procesamiento. Por otra parte, si existe la posibilidad de acceso a paquetería estadística, éste es el camino más rápido para la obtención de resultados exactos y bien presentados.

Existe una amplia gama de paquetería estadística. En la mayoría de las universidades e instituciones gubernamentales se utiliza SPSS, tanto en sus versiones para computadoras personales que corren desde el sistema operativo MS-DOS, como en las versiones para Windows, y también en sus versiones para sistemas multiusuario. Otros paquetes importantes son StatGraphics, SAS, Statistic, Stata, S-plus, etc.

El desarrollo del SPSS comenzó en 1965 en la Stanford University como consecuencia de la frustración que sintieron varios de sus autores originales al intentar apoyar técnicamente en análisis estadístico a investigadores y profesores del Instituto de Estudios Políticos de la referida universidad. En esa época se disponía de programas estadísticos aislados cuyos criterios de operación eran muy diferentes y dificultaban

enormemente la transferencia de resultados entre ellos. Además, la documentación disponible para su operación se presentaba con un lenguaje técnico casi incomprensible para el usuario.
El SPSS (Statistical Package for the Social Sciences) es un sistema computarizado que ha sido el producto de 30 años de investigación, desde su primera versión publicada en 1970. Inicialmente orientada a las grandes computadoras vigentes en la época.

SPSS tuvo desde sus inicios mayor éxito del que sus autores esperaban, la razón de su éxito fue sin duda, que en su desarrollo participaron especialistas en Ciencias Sociales, en Computación y en Estadística. Los criterios de desarrollo que el grupo de investigadores se impusieron fueron:

- Los procedimientos estadísticos y matemáticos correctos.
- Los programas computacionalmente eficientes.
- El enfoque, la lógica y la sintaxis de operación congruentes con el de los científicos orientados al análisis estadístico.
- El sistema proporcione un ambiente de trabajo propio.
- De fácil acceso y operación al usuario.

El paquete SPSS (Paquete estadístico para las Ciencias Sociales) es un sistema integrado de programas de computadora para el análisis estadístico de datos relacionados con las Ciencias Sociales, lo cual no excluye su aplicación a datos relacionados con cualquier disciplina. SPSS proporciona un conjunto amplio e integrado de rutinas para efectuar muchos tipos diferentes de proceso de datos en forma relativamente sencilla. Permite una gran flexibilidad en el formato de los datos e incluso apoya la importación y traducción de archivos generados empleando otros paquetes (como bases de datos y hojas de cálculo). Ofrece al usuario una amplia gama de procedimientos para transformación de datos y manipulación de archivos, además de un gran número de rutinas estadísticas de uso común en las Ciencias Sociales.

A partir de la versión X para sistemas multiusuario, se presenta una sintaxis totalmente compatible con la de versiones para computadoras personales, lo cual amplia las posibilidades de explotación del paquete por parte del usuario, ya que éste puede avanzar en el desarrollo del conjunto de comandos del SPSS y en la depuración de datos con la flexibilidad y comodidad del equipo PC, para después efectuar el procesamiento de archivos de datos que sobrepasen las capacidades de la computadora personal en máquinas grandes (mainfranes), ya sea a través de una red o de manera directa.

Desde hace más de diez años SPSS Inc; puso en el mercado una versión para computadoras personales que era una reducción en cuanto a capacidad de proceso y número de rutinas de la versión para equipo multiusuario, que desde 1970 se había utilizado con éxito en todo el mundo. Hasta principios de 1993 las diferentes versiones de SPSS/PC se produjeron para correr desde el sistema operativo MS-DOS. El paquete en estas versiones para computadoras personales está constituido por varios módulos. El módulo básico permite la obtención de tablas de frecuencias, tablas de doble entrada, cálculo de medidas descriptivas y de correlación, además tiene diversos medios para la transformación, ponderación y remodificación de datos. Otros módulos sirven para el cálculo de estadísticas avanzadas (Advanced Statistics), el análisis de series de tiempo (Trends), la generación de tablas complejas con excelente presentación (Tables), la introducción y validación de datos (Data Entry), la producción de gráficas de alta calidad (Graphics), la realización de mapas (Mapping). análisis de conglomerados (Professional Statistics), relaciones estructurales lineales (Lisrel 7), análisis de calidad (QI Analyst), así como análisis multivariados para datos por categorías y diagramas para análisis de conglomerados (Categories y Chaid).

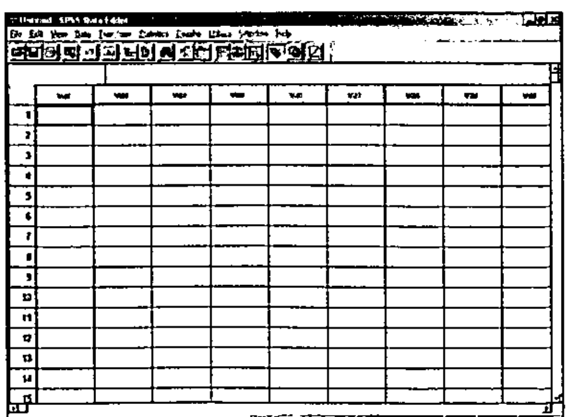
En 1993 se presentó la primera versión de SPSS para Windows, con una facilidad de uso, gracias a la interfaz gráfica con el usuario que hace su aprendizaje y empleo casi intuitivos. En esta versión, a cambio de la necesidad de disponer de una PC con alta capacidad gráfica y gran memoria, es posible procesar números casi ilimitados de variables y casos, además de producir en forma automática todo tipo de gráficas de alta resolución. Se han aumentado los procedimientos estadísticos del módulo básico y es posible emplear módulos de versiones anteriores. Entre los módulos de la versión Windows se encuentran: Estadísticas profesionales, estadísticas avanzadas, Tablas, Pruebas exactas, Categorías, Tendencias, AMOS (manejo de modelos de ecuaciones estructurales), DIAMOND (visualización de múltiples facetas para la detección de patrones y relaciones en los datos), Neural Connection (modelado con base en redes neuronales), Remark (captura de datos automatizada y económica por medio de escáner), Teleform (para crear y distribuir cuestionarios desde la PC) y Mapinfo (para crear mapas temáticos).

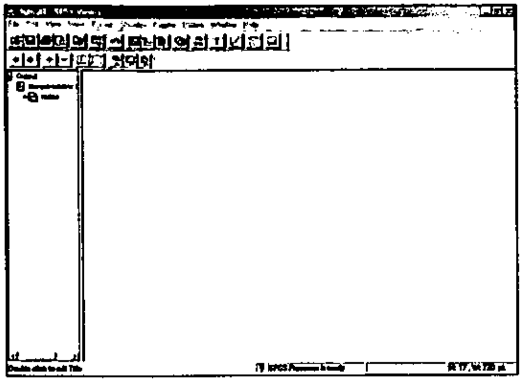
Actualmente se encuentra al mercado la versión 10 de este paquete, pero la elaboración del presente trabajo se basa en la versión 8.0 para windows.

5.2 PRESENTACIÓN GENERAL DE LAS VENTANAS DE SPSS.

En una sesión típica con SPSS, se trabajar con tres tipos de ventanas

Datos Editor Window que es la primera ventana que se encuentra al iniciar el programa, se usa para definir y capturar los datos, para realizar los procedimientos estadísticos.

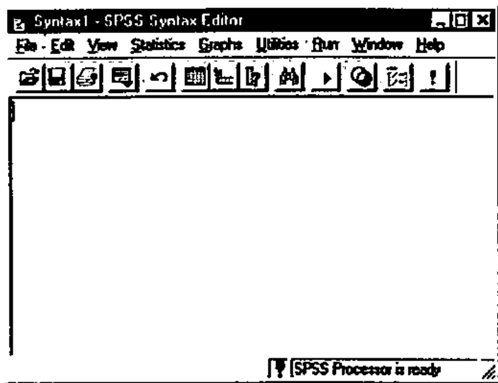




Los resultados de las pruebas estadísticas aparecen en la Ventana de Output.

La Ventana de la Sintaxis que puede usarse para guardar el registro de los funcionamientos que se realizan con los datos.

Esta ventana se abrirá automáticamente cuando se pulsa el botón de función "Paste", pero más allá de servir como guía de los funcionamientos, esta sirve para realizar corridas de ordenes programadas con o sin la ayuda de las funciones predeterminadas.



Se puede salvar cualquiera de las ventanas, en donde SPSS agrega un sufijo automáticamente al extremo del nombre de archivo (".sav " para los archivos de editor de datos, ".spo" para la ventana Output, y ".sps" para la ventana de sintaxis).

La información básica sobre estos archivos se resume como:

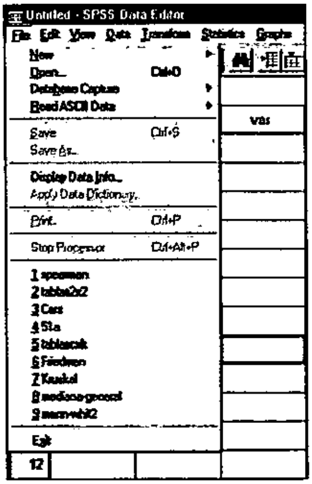
Ventana	Terminación del Archivo	Función
Editor	.sav	Se usa para definir, capturar y editar datos y ejecuta las pruebas estadísticas
Output	.spo	Contiene los resultados de los procedimientos estadísticos
Syntax	.sps	Esta ventana es activada cuando se hace clic en la función Paste y un registro de los archivos del funcionamientos de "paste" se puede saber que las ordenes corridas en SPSS se pueden realizar desde esta ventana

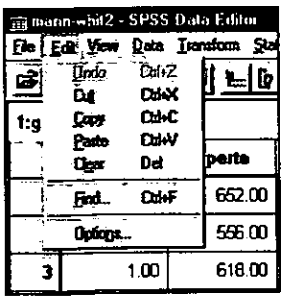
5.3 PRESENTACIÓN DE LA BARRA DE HERRAMIENTAS

La ventana de aplicación SPSS, con la *barra de menú principal* cuenta con las opciones:

File	para crear nuevos archivos o leer archivos ya existentes.
Edit	para modificar o copiar texto de las ventanas de salida o de sintáxis.
Data	para efectuar cambios globales temporales a archivos de datos.
Transform	para hacer cambios a variables seleccionadas del archivo de datos y para calcular nuevas variables a partir de los valores de variables ya existente.
Statistics	para seleccionar los diferentes procedimientos estadísticos.
Graphs	para crear gráficas.
Utilities	para cambiar fuentes, intercambiar datos en forma dinámica exhibir información sobre el contenido de los archivos de datos o abrir un índice de los comandos SPSS.
Window	para arreglar, seleccionar y controlar los atributos de las diferentes ventanas de SPSS.
Help	para abrir una ventana de ayuda sobre las diferentes características de SPSS;

La primera opción de la barra de herramientas "File", tiene la misma función que todos los paquetes de Windows, como son: creación de un nuevo archivo, cerrar archivo, importación de archivos, lectura de datos tipo ASCII, despliegue de información de datos, además de los comunes como son: salvar, salvar como, imprimir, detener procesos y salida.

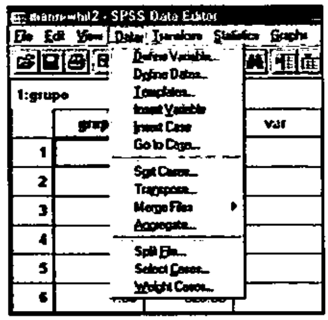




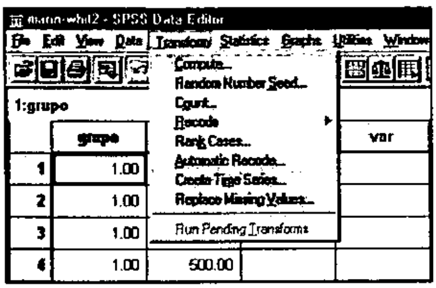
La segunda opción de la barra de herramientas "Edit", se utiliza para cortar, copiar, eliminar, encontrar datos de la ventana de Editor, entre otros.

La tercera opción de la barra de herramientas "View", se utiliza para determinar el tipo de letra, quitar el formato de celdas, entre otros, como se podrá ver estas tres primeras opciones son similares a todos los programas que operan bajo windows.

A partir de la cuarta opción de la barra de herramientas se inician las opciones propias del paquete. "Data" – la cuarta opción - es muy importante, puesto que aquí se definen las características de las variables: codificación de datos (tipo de datos, tamaño, nombre o etiqueta), valores faltantes (Missing values), formato de columna, ponderaciones, etc.



También se encuentran las opciones de Insertar variables, casos, ir a un caso determinado, agregar casos, ordenar datos, etc.



La opción de *Transform* nos ofrece algunos comandos que nos sirven para modificar , agrupar y seleccionar datos que cumplan con determinadas condiciones; o también para aplicarles transformaciones aritméticas.

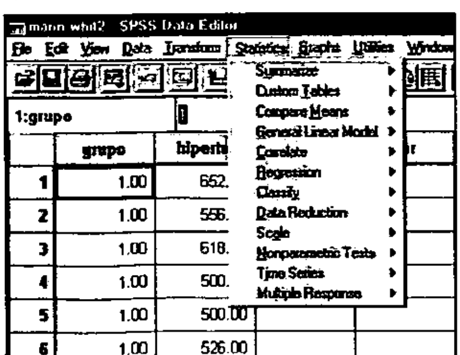
A continuación se describen algunos de estos comandos.

Recode:

Este comando se emplea para cambiar el esquema de codificación de una variable numérica o alfanumérica corta. Este cambio o recodificación puede aplicarse valor por valor o a intervalos de valores.

Compute:

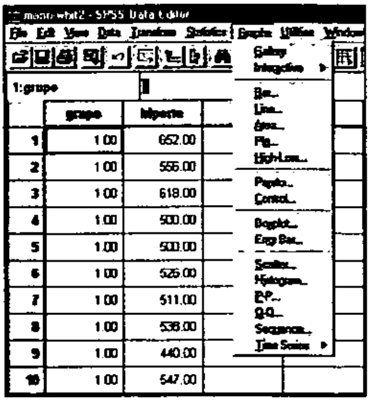
Este comando crea una nueva variable o modifica los valores de una variable va existente, ya sea en una nueva variable o en la variable donde se esta realizando el proceso.



La sexta opción "Statistics" que nos ofrece la barra de herramientas es la más importante de todas, puesto que contiene todos procesos estadísticos que se pueden realizar. Las estadísticas que nos ofrece esta ventana son:

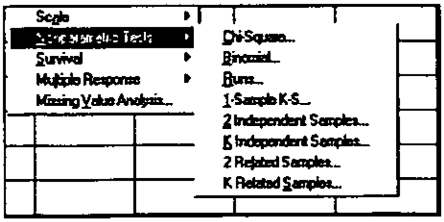
Summarize.- Este comando nos ofrece las estadísticas básicas.
Custum tables.- Procesos de tablas cruzada o crosstabs.

Compare Means.- Opción para realizar comparaciones de medias, útil para el análisis de experimentos.
General Lineal Model.- Modelos lineales Generalizados, útil para el análisis de datos categóricos.
Correlate.- Opción para obtener correlaciones de datos.
Regression.- Proceso para realizar métodos de regresiones (lineal, Logist, Cox-box).
Classify.- Procesos de análisis de discriminantes y análisis de cluster
Data Reduction.- Análisis de conglomerados, Análisis Factorial, Análisis de componentes principales, etc.
Scale.- Escalamiento multidimensional
Nonparametric Test.- Brinda todas las principales pruebas no paramétricas, ya sea para dos o mas muestras relacionados o independientes.
Tipe Series.-Series de tiempo



La opción de Graphs nos ofrece una gran variedad de graficas, ya sean en dos dimensiones o en tres.

La finalidad de presentar como esta conformado el paquete SPSS es tener una idea general de todo lo que se puede realizar con él, pero el principal objetivo de este trabajo son las pruebas no paramétricas, las cuales se encuentran en la opción de "Statistics" que despliega las siguientes opciones (ver imagen).



Como podemos ver en la imagen hay varios tipos de pruebas, para 2 muestras relacionadas o independientes, para k muestras relacionadas o independiente, pruebas para rachas, la prueba de kolmogorov-Smirnov, la binomial y la Chi-Square; que a su vez cada una de estas opciones tiene integrada uno o varios tipos de pruebas.

Es importante aclarar que no se presentaran todas las pruebas que componen este apartado, si no solo las más conocidas ya sea por su facilidad de calculo o por la potencia de la prueba (la mejor).

CAPITULO VI

PRUEBAS NO PARAMÉTRICAS PARA DOS MUESTRAS RELACIONADAS

La técnica paramétrica usual para analizar datos provenientes de dos muestras relacionadas es aplicar una prueba t a los puntajes de diferencia, donde se supone que los puntajes de diferencia están distribuidos de forma normal e independiente de la población que se tomó la muestra y además requiere que sean medidos, por lo menos, en una escala de intervalo.

Pero la prueba t no es aplicable, cuando se tiene que:

a) Los supuestos y requerimientos de la prueba t son poco realistas para los datos.
b) Se prefiere evitar hacer los supuestos o probar los requerimientos para dar mayor generalidad a las conclusiones.
c) Las diferencias entre las parejas igualadas no están representadas con puntajes, sino con "signos" (esto es, se puede decir que miembro de cualquier pareja es "mayor que" otro, pero no se puede decir en cuánto)
d) Sus puntajes son simplemente clasificatorios. Los miembros de cada pareja igualada pueden responder del mismo modo o de manera enteramente diferente, pero no tienen orden ni relación cuantitativa.

En estos casos, se puede escoger una de las pruebas estadísticas no paramétricas para dos muestras relacionadas, como son : la prueba de los signos o la prueba del signo-rango de Wilcoxon[1].

[1] Estas no son las únicas pruebas que se pueden realizar para dos muestras relacionadas, existen otras como: La prueba de Mcnemar, La prueba de Walsh, la prueba de aleatoriedad para pares igualados, etc.

Los procedimientos no paramétricos presentados aquí, son aplicables cuando los datos consisten en dos muestras, que están relacionadas de alguna manera y cuando la variable de respuesta puede considerarse como ordinal o categórica con dos categorías.

Las observaciones pueden ser mediciones tomadas de los mismos sujetos, antes y después de algún tratamiento aplicado[2] o también pueden ser medidas tomadas de diferentes sujetos que han sido clasificados en uno o más criterios (método de pares). Esto es, uno puede "igualar" o relacionar de alguna manera las dos muestras estudiadas, cosa que puede lograrse cuando cada sujeto es su propio control o con parejas de sujetos en las que se asignan los miembros de cada pareja a las dos condiciones. Cuando un sujeto "sirve como su propio control" está expuesto a ambos tratamientos en diferentes ocasiones. Cuando se usa el método de pares, se trata de seleccionar dentro de lo posible en cada pareja de sujetos, aquellos que sean los más semejantes con respecto a cualquier variable extraña que pudiera influir el resultado de la investigación – comúnmente se utilizan gemelos.

Siempre que sea factible, el método de usar a cada sujeto como su propio control es preferible al método de pares, debido a que nuestra capacidad para formar parejas se ve limitada por la ignorancia de las variables pertinentes que determinan su conducta.

Este tipo de pruebas se utiliza para establecer la diferencia entre dos muestras, es decir para probar si las dos muestras relacionadas difieren en localización (o si una sola muestra proviene de una población con una mediana específica[3]) o si una muestra es mejor que la otra – común en tratamientos , en donde la variable de interés para el análisis de las dos muestras relacionadas, es la diferencia entre las dos medidas dentro de los pares. Si las mediciones dentro de la muestra de pares son suficientemente diferentes, en la investigación puede concluir que los dos tratamientos tienen diferentes efectos o que uno es más efectivo que el otro.

6.1 PRUEBA DE LOS SIGNOS.

6.1.1 Función

La prueba de los signos, es quizás él mas viejo de todos los procedimientos no paramétricos. La utilización de este procedimiento, fue reportado en 1710 por Arbuthnott.

El atractivo de la prueba de los signos es su sencillez y al igual que la prueba de Wilcoxon se utiliza para probar hipótesis que involucran a uno o dos grupos (o muestras) correlacionados o apareados, como en los experimentos en los cuales cada

[2] También puede ser que uno de los tratamientos sea ningún tratamiento o aplicar algún placebo.
[3] Este es una caso particular cuando se tiene una sola muestra y se pretende comparar con una mediana especifica, el cual se incluye también en este capítulo.

sujeto en su propio control o cuando el experimentador desea establecer que ambas condiciones son diferentes – en el caso de dos muestras[4]. Por tanto, puede usarse para probar si dos muestras relacionadas difieren en localización o si una sola muestra proviene de una población con una mediana específica.

Esta prueba se puede aplicar en las mismas situaciones que la prueba del rango con signo de Wilcoxon[5]. La diferencia principal entre las dos pruebas radica en que la primera no hace uso de la magnitud de los resultados de diferencia, sino que simplemente se anotan sus signos - de aquí su nombre. Sin embargo, en muchos casos la prueba de los signos proporciona una alternativa rápida y sencilla de la prueba del rango con signo, aun cuando es fácil encontrar situaciones en las que las dos pruebas conducen a conclusiones diferentes – esto se debe principalmente a la potencia de la prueba. De hecho, puede demostrarse fácilmente que la prueba del signo es un caso especial de la prueba del rango con signo, cuando todos los resultados de las diferencias tienen la misma magnitud.

La prueba de los signos[6] debe su nombre al uso de los signos más y menos en la medición en lugar de cantidades. Es particularmente útil cuando la medición cuantitativa es imposible o no es práctica, pudiendo aún haber cierto orden entre los miembros de cada pareja. Por esta razón, la prueba de los signos también se puede aplicar en las situaciones en que sólo se sabe si los resultados de diferencia son positivos o negativos, pero sin tener sus magnitudes.

Una de las desventajas de la prueba de los signos, es que elimina por completo cualquier información cuantitativa, que puede ser inherente a los datos, es decir la prueba de los signos trata a todas las diferencias positivas como si fueran iguales y lo mismo hace con todas las diferencias negativas, es decir falla al no hacer uso de la información referente a las magnitudes de las diferencias.

6.1.2 Potencia de la prueba

La potencia[7] de la prueba de los signos con respecto a la prueba t de Student para el caso en que la población se distribuyera normalmente, es aproximadamente del 95% de eficiencia para muestras pequeñas.

[4] En el caso de tener solo una muestra, comúnmente se quiere probar si esta proviene de una población con una mediana específica.

[5] Es preferible usar la prueba del rango en lugar de la prueba de los signos cuando se cuenta con las magnitudes de los resultados de diferencias, puesto que en ella se hace mayor uso de los datos, además de que esta prueba es mucho más potente que la de los signos.

[6] La prueba de los signos es solamente una variación de la prueba binomial.

[7] Trabajo realizado por Walsh en 1946

6.1.3 Supuestos

La prueba de los signos no hace ningún supuesto acerca de la forma de la distribución de las diferencias ni pide que todos los sujetos se tomen de la misma población. Las parejas pueden incluso, provenir de distintas poblaciones con respecto a edad, sexo, etc. El único requisito es que el investigador haya logrado en cada pareja un nivel de medición por lo menos ordinal y además se supone que la muestra se compone de n parejas relacionadas, $(X_1, Y_1)(X_2, Y_2),\ldots(X_n, Y_n)$.

Los supuestos básicos de la prueba para dos muestras relacionadas son los siguientes:

a. Las variables bivariadas (X_i, Y_i) con $i = 1,2\ldots n$, sean mutuamente independientes.
b. La escala de medición debe ser al menos ordinal en cada pareja.
c. La variable de interés (es la diferencia entre los pares de mediciones) sea continua.
d. Las parejas son internamente consistentes, en el sentido de que si $P(+) > P(-)$ para una pareja entonces $P(+) > P(-)$para todas las parejas. La misma afirmación será valida en los casos: $P(+) < P(-)$ y $P(+) = P(-)$.

Los supuestos básicos de la prueba para una muestra son los siguientes:

a. La muestra obtenida para el análisis debe provenir de una muestra aleatoria de una población con una mediana M desconocida.
b. La variable de interés debe estar en una escala de medición por lo menos ordinal.
c. La variable de interés sea continua. Los n valores de las muestras son designados por $X_1, X_2, \ldots, X_n$.

6.1.4 Método.

Sea una muestra compuesta por n parejas relacionadas $(X_1, Y_1)(X_2, Y_2),\ldots(X_n, Y_n)$.

1. A cada pareja se le asociará un signo $(+)$, si el valor de X_i sobrepasa al correspondiente Y_i (es decir cuando $X_i > Y_i$); un signo $(-)$si sucede lo contrario (es decir cuando $X_i < Y_i$) y no se le asociará ningún signo - de hecho se eliminará a la pareja - , cuando $X_i = Y_i$.
2. Se asignará el número de parejas "**n**" eliminando las parejas iguales y el valor de "**x**" será el número del signo menos frecuente, ya sean "-" o "+".

En el caso de tener una muestra $X_1, X_2, \ldots, X_n$., se realizaran los pasos anteriores, pero aqui se sustituye el valor de Y_i por el de la mediana o medida de tendencia central a probar M_0.

6.1.5 Hipótesis.

1. **Para el caso de dos muestras :**

En el caso de una prueba de dos colas.

$Ho: P(X_i > Y_i) = P(X_i < Y_i) = 1/2$. La hipótesis nula expresa que la diferencia de las medianas de las dos condiciones es cero.

$Ha: P(X_i > Y_i) \neq P(X_i < Y_i)$. La hipótesis alternativa expresa que si hay diferencia entre las dos condiciones.

Donde:
X_i es el valor observado antes del tratamiento (o después del tratamiento)
Y_i es el valor observado después del tratamiento (o antes del tratamiento)

En el caso de una prueba de una cola.

La hipótesis nula se expresa igual que para una prueba de dos colas, ya que la apropiada región de rechazo depende de la hipótesis alternativa.

$$Ho: P(X_i > Y_i) = P(X_i < Y_i) = 1/2.$$

La hipótesis alternativa para pruebas de una cola puede ser:

$$Ha: P(X_i > Y_i) > P(X_i < Y_i) \quad \text{ó bien} \quad Ha: P(X_i > Y_i) < P(X_i < Y_i).$$

2. **Para el caso de una muestra:**

En el caso de una prueba de dos colas.

$Ho: M = M_0 \quad vs \quad Ha: M \neq M_0$

Donde:
M_0 es el valor de la mediana con que se quiere comparar la muestra.

En el caso de una prueba de una cola.

$Ho: M \leq M_0 \quad vs \quad Ha: M > M_0 \quad$ o bien $\quad Ho: M \geq M_0 \quad vs \quad Ha: M < M_0$

6.1.6 Estadístico de Prueba.

Existen dos casos para esta prueba:

1er Caso - Cuando el número de parejas es menos o igual que 25 $(n \leq 25)$.

Esta probabilidad puede calcularse haciendo uso de al distribución Binomial y como es sabido existen varias versiones de tablas de valores calculados para la distribución Binomial y que el empleo de estas tablas es conveniente sobre todo con valores $n \leq 25$.

El cálculo de la probabilidad mencionada se realiza mediante la siguiente expresión:

$$\Pr = \sum_{t=0}^{x} \binom{n}{t} \left(\frac{1}{2}\right)^n$$

donde

x = es el número menor de signos en el calculo de la probabilidad asociada a la prueba de una sola cola .
n = al número de parejas en la muestra (en el caso de tener un empate es decir $P(+) = P(-)$, el valor de n será : n = total de casos menos los empates).
$P = 1/2$

Nota: $\binom{n}{t} = \frac{n!}{(n-t)!t!}$

En el caso de la hipótesis de dos colas la probabilidad se tendria que duplicar.

2° Caso. Cuando el número de parejas es mayor que 25 $(n > 25)$.

Cuando la distribución binomial es suficientemente grande, ésta tiende a comportarse asintoticamente como una normal (Teorema de Limite Central). La probabilidad asociada a esta prueba se puede obtener por medio de la estadística Z que corresponde a una variable aleatoria que se distribuye como una normal estándar es decir:

$$Z = \frac{x-\mu}{\sqrt{\sigma}} = \frac{x-np}{\sqrt{np(1-p)}}; \quad \text{como } p = 1/2 \text{ se tendría:} \quad Z_c = \frac{x-n/2}{\sqrt{n/4}}$$

Pero debido a que la aproximación a la normal involucra aproximaciones a una distribución discreta por medio de una distribución continua, se suele utilizar un factor de corrección de continuidad, en este caso el factor es de 0.5, quedando entonces el valor de la Z como:

$$Z_c = \frac{(x \pm 0.5) - n/2}{\sqrt{n/4}}$$

En donde se usa $(x+0.5)$ si x es menor que n/2 y $(x-0.5)$ si x es mayor a n/2.

Cualquiera de los dos valores de las Z calculadas se tiene que comparar con los valores de la tabla de la Normal estándar para obtener la probabilidad asociada, en el caso de tener que probar una hipótesis de dos colas, la probabilidad se tendrá que duplicar.

En caso de contar con una sola muestra, esta también cumple con lo antes mencionado, es decir en este caso también el estadístico de prueba depende del tamaño de la muestra – valor **"n"**.

6.1.7 Regla de decisión.

De acuerdo con la hipótesis nula se espera que el número de signos $(+)$ sea muy parecido al número de signos $(-)$. Es decir si Ho fuera verdadera, aproximadamente la mitad de las parejas tendrían signo $(+)$ y la otra mitad signo $(-)$. La hipótesis nula Ho se rechaza cuando ocurren muy pocas parejas de determinado signo, y esto se determina por medio del tamaño de "n".

En el caso de $n \leq 25$

Para una prueba de una cola
Se rechaza Ho:
Si $(Pr) \leq (\alpha)$ donde α es el nivel de significancia asignado
No se rechaza Ho:
Si $(Pr) > (\alpha)$ donde α es el nivel de significancia asignado.
En el caso de una prueba de dos colas el valor de (Pr)sé duplica $2(Pr)$.

En el caso de $n > 25$

Se rechaza Ho:
Si $Z_c \geq Z_t$ donde Z_c es Z calculada y Z_t es el valor obtenido en tablas de la normal estándar con $t = 1 - \alpha/2$,para una prueba de dos colas, en el caso de una prueba de una cola el valor Z_t es $t = 1 - \alpha$
No se rechaza Ho :
Si $Z_c < Z_t$ donde Z_c es Z calculada y Z_t es el valor obtenido en tablas de la normal estándar con $t = 1 - \alpha/2$ para una prueba de dos colas, en el caso de una prueba de una cola el valor Z_t es $t = 1 - \alpha$

Utilizando la probabilidad asociada (P-value) de Z_c.

Se rechaza Ho:
Si $P(Z \geq Z_c) < \alpha$ para una prueba de una colas, en el caso de tener una prueba de dos cola se duplica el valor, es decir $2P(Z \geq Z_c) < \alpha$.
No se rechaza Ho :
Si $P(Z \geq Z_c) \geq \alpha$ para una prueba de una colas, en el caso de tener una prueba de dos cola se duplica el valor, es decir $2P(Z \geq Z_c) \geq \alpha$.

Si se tiene una sola muestra la regla de decisión se utilizan de la misma forma. Para comprender como se debe utilizar esta prueba se mostraran ejemplos para cada uno de los casos antes presentados.

6.1.8 Ejemplos

- *1er. caso.* Cuando el número de parejas es menor o igual a 25 $(n \leq 25)$ con dos muestras relacionadas. Supóngase que se ha formulado la hipótesis de que el liderazgo es una cualidad que se puede adquirir mediante un adiestramiento, con la esperanza de que, la inteligencia y la aptitud para el liderazgo sean variables correlacionadas; se toman dos grupos uno que recibe un adiestramiento especial para formar líderes (X_i) y otro que no recibe instrucción especial (Y_i) estos dos grupos se aparean basándose en la inteligencia.

Tabla 6.1-1

CALIFICACION DE LIDERAZGO

PAREJA	GRUPO EXPERIMENTAL (X_i)	GRUPO DE CONTROL (Y_i)	SIGNO ASIGNADO $(X_i > Y_i)$
A	47	40	+
B	43	38	+
C	36	42	-
D	38	25	+
E	30	29	+
F	22	26	-
G	25	16	+
H	21	18	+
I	14	8	+
J	12	4	+
K	5	7	-
L	9	3	+
M	5	5	Nulo

Nótese que la muestra original de 13 parejas se elimina 1 por tener el mismo valor, por ello el calculo se hará con n=12 y para x=3, pues el signo (-) fue el menos frecuente.

Al completar el curso de entrenamiento para líderes, se pide a unos observadores independientes que evalúen las cualidades de líder de cada uno de los sujetos según una escala de 50 puntos. Los resultados se encuentran en la *tabla 6.1-1.* Seleccionar un nivel de significancia del 5%.

- *Respuesta*

Debido que al observar los datos no podemos afirmar que las calificaciones tengan propiedades cuantitativas precisas, como el observador independiente calificó a los dos grupos en una escala ordinal del 1 al 50, lo único que podríamos justificar es que cualquier diferencia que exista entre dos calificaciones apareadas es un indicador válido de la dirección (es decir tenemos una escala de medición ordinal). Esto nos indica que se puede utilizar la prueba de los signos.

Hipótesis:

Debido a que la hipótesis pretende probar que las cualidades del liderazgo se adquieren mediante un adiestramiento y no hay indicación de diferencia en un sentido o en otro, tratamos con una hipótesis de dos colas, es decir:

Ho: La mediana de las diferencias es cero.

El número de sujetos que tomaron el curso de adiestramiento presenta el mismo nivel de liderazgo que el grupo que no tomó ningún curso, es decir los dos grupos el experimental y el de control no presentan diferencia alguna en la cualidad de liderazgo.

Ha: La mediana de las diferencias es distinta.
Los dos grupos el experimental y el de control presentan una diferencia en la cualidad del liderazgo.

Estadístico de Prueba:

Como tenemos el caso en que $n \leq 25$ se tiene que utilizar la distribución binomial es decir:

$$\Pr = \sum_{t=0}^{3} \binom{12}{t} \left(\frac{1}{2}\right)^{12} \text{ donde: } x = 3,\ P = \frac{1}{2}\ y\ n = 12$$

Donde obtenemos que $\Pr = 0.07299805$ para una prueba de una cola, como éste problema requiere de una prueba de dos colas el valor anterior se tiene que duplicar, es decir tenemos que $2\Pr = 0.14559609$.

Regla de decisión:

Como se tiene que $2\Pr = 0.14559609 > .05$ podemos concluir que no se rechaza la hipótesis nula (Ho), es decir que los dos grupos el experimental y el de control no presentan diferencia alguna en la cualidad de liderazgo.

- *2do Caso.* Cuando el número de parejas es mayor que $25 (n > 25)$. Benjafield y Adams-Webber (1976), y Berjafield y Green (1978) postularon que la proporción de adjetivos que usan las personas para describir a otras es exactamente $\phi = 0.618$[8]. Ellos sostienen que esto da lugar a que los adjetivos negativos, los cuales ellos piensan que son más importantes cuando se describe el conocimiento de otras personas, se destaquen como "figuras" (gramaticales) en contraste con un "fondo" más extenso de adjetivos positivos. Para probar esto, los investigadores le pidieron a 26 hombres, que llenaran una rejilla de repertorio en la que describirían sus conocimientos de otras personas en un formato estandarizado, usando las construcciones gramaticales que ellos eligieran. Después se les pidió que dijeran cuáles de sus construcciones consideraban que eran de naturaleza evaluativa. Para cada persona, se calculó la proporción de las veces que se usó el polo positivo de una construcción para describir el conocimiento de otra persona. Quiere probarse si estos datos son congruentes con la predicción de que en promedio, esta proporción es ϕ, utilizando $\alpha = 5\%$.

[8] Esta letra griega hace referencia a un número bastante improbable, que resulta del cálculo $(\sqrt{5} - 1)/2$, llamado la sección Aurea.

Tabla 6.1-2

PAREJA	GRUPO EXPERIMENTAL (X_i)	SIGNO ASIGNADO ($X_i>Y_i$)
1	.417	-
2	.458	-
3	.458	-
4	.500	-
5	.517	-
6	.531	-
7	.533	-
8	.542	-
9	.556	-
10	.556	-
11	.560	-
12	.573	-
13	.583	-
14	.590	-
15	.595	-
16	.600	-
17	.625	+
18	.639	+
19	.643	+
20	.643	+
21	.648	+
22	.652	+
23	.653	+
24	.702	+
25	.722	+
26	.786	+

-*Respuesta:*

Hipótesis

Ho: $M_0 = 0.618$.
La hipótesis nula, que se va a probar es que los resultados provienen de una población que es simétrica respecto a ϕ.

Ha: $M_0 \neq 0.618$.
Que los resultados no provienen de una población que es simétrica respecto a ϕ, es decir que existe una diferencia con respecto a ϕ.

Estadístico de Prueba:

La probabilidad asociada a esta prueba se puede obtener por medio de la estadística Z que corresponde a una variable aleatoria que se distribuye como una normal estándar. Es decir, se tendría:

$$Z = \frac{x - \frac{n}{2}}{\sqrt{\frac{n}{4}}} = \frac{10 - \frac{26}{2}}{\sqrt{\frac{26}{4}}} = -1.17669$$

Además tenemos que el valor de Z_t en tablas es: $Z_t = 1.64$

Regla de decisión:

Como la estadística de prueba es mayor que el valor obtenido en tablas $|-1.17669| < 1.64$ podemos concluir que no se rechaza la hipótesis nula (Ho), es decir que los resultados provienen de una población que es simétrica respecto a ϕ (lo que querían probar los investigadores).

Por otro lado si utilizamos el estadístico de prueba con factor de corrección por continuidad tenemos:

$$Z_c = \frac{(x \pm 0.5) - n/2}{\sqrt{n/4}} = \frac{(10 - 0.5) - 26/2}{\sqrt{26/4}} = -0.98058$$

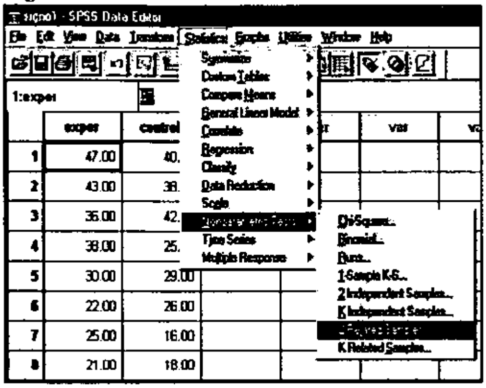
y como podemos ver se verifica la conclusión anterior.

6.1.9 Utilizando el paquete SPSS.

En esta parte se presentaran los pasos que se tienen que seguir para realizar correctamente la prueba de los signos.

1. Se incorporan los datos de las dos variables, en este caso se incorporaran los datos del ejemplo 1 . Los cuales se tienen que incorporan en dos columnas distintas (las dos tienen que tener el mismo número de datos), tal como se muestra a continuación (*Figura 6.1-1*).

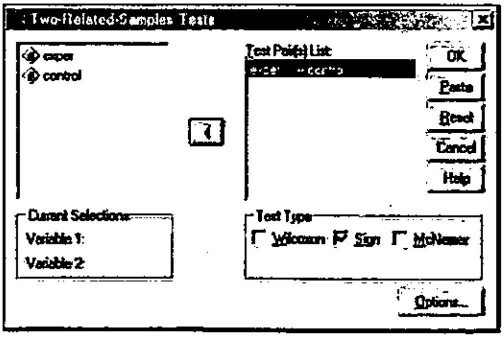
Figura 6.1-1



2. En la barra de herramientas se debe seleccionar el icono correspondiente a *Statistics*, dentro del cual se desplegara una serie de opciones a escoger, en nuestro caso uno debe elegir *Nonparametric tests*, que es la opción en donde se encuentran todas las pruebas no paramétricas con las que cuenta el paquete.

3. Dentro de la opción *Nonparametric tests* se desplegará nuevamente otra ventana donde se debe seleccionara la opción *2Related samples*, ya que la prueba de los signos opera con dos muestras relacionadas o apareadas.

Figura 6.1-2



4. Posteriormente se presentará al usuario una ventana de dialogo en donde se solicita especificar las dos variables dentro de su archivo que entraran en el análisis, para nuestro caso tenemos que la variable "control" se tomará como el minuendo y la variable "exper" como sustraendo de la diferencia que el paquete calculará en cada caso (*Figura 6.1-2*)

Una vez copiados los nombres de las dos variables de análisis a la ventana de la derecha titulada *Tests Pair (List)*, se debe marcar en el recuadro *Test Type* la opción *Sign* y finalmente oprimir el botón *OK* ubicado en la parte superior derecha, para proceder con los cálculos.

Los resultados aparecen en otra ventana llamada *Output1*, en donde el usuario estará en posibilidades de imprimir o guardar como archivo.

Sign Test

Frequencies

		N
CONTROL - EXPER	Negative Differences[a]	9
	Positive Differences[b]	3
	Ties[c]	1
	Total	13

a. CONTROL < EXPER

b. CONTROL > EXPER

c. EXPER = CONTROL

Test Statistics[b]

	CONTROL - EXPER
Exact Sig. (2-tailed)	.146[a]

a. Binomial distribution used.

b. Sign Test

En la ventana de resultados se presenta la tabla (*Frequencies*) que muestra el número de diferencias positivas (3), el de negativas (9) y el de las diferencias nulas (1), que el paquete identifica como (ties). Finalmente en la segunda tabla (*Test Statistics*) se presenta la probabilidad de los datos incorporados bajo la hipótesis nula (Valor P) calculada –para una prueba de dos colas que el paquete da por default- el cual vale 0.146; si se presentara el caso de tener una prueba de una cola tendríamos que dividir a este valor entre dos.

Debido que el ejemplo se trata de un problema que plantea una prueba de dos colas utilizando una $\alpha = 0.05$ y siguiendo la regla de decisión concluimos que no se rechaza la hipótesis nula - como el resultado obtenido anteriormente.

Ejemplo: para n>25 con una muestra.

En la *Figura 6.1-3*, se muestran los resultados de los 26 hombres evaluados incorporados al paquete SPSS; como la prueba de los signos es útil probar si una sola muestra proviene de una población con una mediana especifica, se necesita que las dos variables tengan la misma cantidad de valores – por este motivo la variable "aurea" se tiene que incorporar hasta igualar la misma cantidad de la variable "evalua".

Los resultados de la prueba se muestra en los cuadros siguientes:

En la ventana de resultados se presenta la tabla (*Frequencies*) que muestra el número de diferencias positivas (10), el número de diferencias negativas (16) y el número de diferencias nulas (0) que el paquete identifica como (*Ties*).

Figura 6.1-3

File Edit View Data Transform Sta

	aurea	evalua
1	.618	.417
2	.618	.458
3	.618	.458
4	.618	.500
5	.618	.517
6	.618	.531
7	.618	.533
8	.618	.542
9	.618	.556
10	.618	.556
11	.618	.560
12	.618	.573
13	.618	.583
14	.618	.590
15	.618	.595

Sign Test

Frequencies

		N
EVALUA - AUREA	Negative Differences[a]	16
	Positive Differences[b]	10
	Ties[c]	0
	Total	26

a. EVALUA < AUREA

b. EVALUA > AUREA

c. AUREA = EVALUA

Test Statistics[a]

	EVALUA - AUREA
Z	-.981
Asymp. Sig. (2-tailed)	.327

a. Sign Test

Finalmente en la segunda tabla (*Test Statistics*) se presenta la probabilidad de los datos incorporados bajo la hipótesis nula calculada, que en este caso es de -0.327; si presentara el caso de tener una prueba de una cola, tendríamos que dividir a éste valor entre dos.

Debido que se trata de una muestra con n>25, el paquete nos proporciona el cálculo del estadístico Z_c, cuyo valor es de -0.981[9] y la probabilidad calculada con el supuesto de normalidad 0.327 para la prueba de dos colas. Como el ejemplo se trata de un problema que plantea una prueba de dos colas, al utilizar la probabilidad asociada tenemos que $2P = 0.327 > 0.025$, por lo que podemos concluir que no se rechaza la hipótesis Nula (Ho), es decir que los resultados provienen de una población que es simétrica respecto a ϕ.

6.1.10 Observación:

Podemos ver que el paquete presenta los resultados de la prueba según él número de observaciones que se le incluyan es decir:
-Cuando $n \leq 25$ utiliza como estadístico de prueba la distribución binomial
-Cuando $n > 25$ utiliza como estadístico de prueba la aproximación a la normal por medio del calculo de la Z_c de la formula que incluye en factor de corrección por continuidad y la probabilidad asociada.

6.2 Prueba de Wilcoxon (Signo-Rango).

6.2.1 Función.

Hemos visto anteriormente, que la prueba de los signos utiliza simplemente la información concerniente a la dirección de las diferencias entre parejas. Si puede

[9] Este valor se obtuvo por medio de la estandarización de la Z_c, tomando en cuenta el factor de corrección en la continuidad, mencionado anterior mente como nota.

tomarse en cuenta tanto la magnitud como la dirección de estas diferencias, es factible emplear una prueba más potente – la prueba de Wilcoxon.

La prueba del signo-rango de Wilcoxon para observaciones apareadas, alcanza mayor potencia al utilizar la información cuantitativa inherente a la clasificación de las diferencias, es decir da mayor peso al par que muestra una diferencia grande entre las dos condiciones que el par que exhibe una diferencia pequeña, además esta prueba resulta de gran utilidad cuando el investigador tiene dudas acerca de la distribución de la variable observada y considera que la normalidad no se cumple razonablemente requisito establecido para el empleo de la prueba *t* para dos grupos relacionados.

La prueba del rango con signo fue creada por Frank Wilcoxon y constituye una de las alternativas no paramétricas para probar hipótesis referentes a uno o dos grupos correlacionados[1] (la otra es la prueba de los signos) en donde la prueba del rango con signo es sensible a las diferencias en localización entre las dos muestras y por lo general se usa para probar si la diferencia entre los dos resultados de cada sujeto proviene de una población que es simétrica respecto a una mediana de cero (la hipótesis nula) o de una población con una mediana diferente de cero (la hipótesis alternativa).

Una hipótesis involucrada en el uso de la prueba del signo-rango de Wilcoxon es que la escala de mediciones sea por lo menos de naturaleza ordinal. En otras palabras, la hipótesis consiste en que las calificaciones permitan el ordenamiento de los datos según la relacion de mayor que y menor que. Sin embargo, la prueba del signo-rango requiere una hipótesis adicional que la puede excluir de algunas de sus aplicaciones potenciales; está hipótesis consiste en admitir que las diferencias entre las calificaciones corresponden también a una escala ordinal.

La prueba de Wilcoxon es la de mayor utilidad para el científico conductual, además que suele emplearse en experimentos que platean condiciones de "antes" y "después" para el mismo de sujeto.

En algunas situaciones experimentales, un diseño de muestras relacionadas permite considerar de una manera más directa la variabilidad entre los sujetos. Para estar seguros de que ninguna diferencia sistemática se observe entre los dos grupos y que éste se deba exclusivamente a los efectos de los tratamientos, se debe suponer que no hay ningún efecto secundario en el transcurso de los tratamientos. Por tanto, si a todos los sujetos se les administra primero el tratamiento A, debe suponerse que su reacción al tratamiento B no son afectados por su respuesta previa al tratamiento A, ya sea en función de que el tratamiento esté aún en sus sistemas o en función de la costumbre a la respuesta del tiempo de reacción. En pocas ocasiones, esta suposición está justificada, pero sus efectos pueden atenuarse un poco aleatorizando o compensando[2]

[1] En el caso de que los datos provengan de una muestra , por lo general se requiere probar si estos provienen de una población simétrica con una mediana específica o con una mediana diferente a una específica.
[2] La aleatorización es el método que se usa con mayor frecuencia y el más simple; implica la decisión al azar de cuál tratamiento debe administrarse primero a cada sujeto. En la compensación, a la mitad de los sujetos se les da primero el tratamiento A y después el tratamiento B, y la otra mitad las reciben en orden inverso. Este procedimiento hace posible probar si hay algún efecto secundario u ordenado.

el orden en el que se administran los dos tratamientos. Otra manera de evitar la dificultad de los efectos secundarios en diseños de muestras relacionadas es usando parejas de sujetos que sean comparables de tantas maneras pertinentes como sean posibles. Esta técnica a menudo se usa gemelos en el estudio. A las parejas de sujetos generalmente se les denominan parejas pareadas.

6.2.2 *Potencia de la prueba*

La potencia-eficiencia caracterizada por la prueba de los signos-rangos de Wilcoxon en comparación con la prueba t es de $3/\pi = 0.955$ si las D_i se distribuyen normalmente. Esto significa que $3/\pi$ es la proporción limite de tamaños de muestra necesaria para que las pruebas de Wilcoxon y t alcancen el mismo poder. Si las D_i se distribuyen uniformemente la eficiencia alcanzada es de 1. Si la prueba de Wilcoxon se compara con la prueba de los signos su eficiencia asintótica relativa es de 2/3 si las D_i se distribuyen normalmente.

6.2.3 *Supuestos.*

Para aplicar esta prueba es necesario que los datos de cada sujeto sean independientes a los datos de los demás sujetos, y que las parejas relacionadas $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ estén, en la escala de mediciones de naturaleza por lo menor ordinal o categórica, en donde a cada pareja se le asociará su diferencia $D_i = X_i - Y_i$. Para las variables aleatorias D_i se tienen los siguientes supuestos:

a. Las D_i son variables aleatorias continuas.
b. Las D_i son mutuamente independientes.
c. Tienen la misma mediana.
d. La distribución de las D_i es simétrica.
e. Las medidas de las D_i deben de estar al menos en una escala ordinal.

En el caso de tener una sola muestra esta debe cumplir con los siguiente supuestos:

a. La muestra obtenida para el análisis debe provenir de una muestra aleatoria de una población con una mediana M desconocida.
b. La variable de interés es continua.
c. La muestra de la población presenta una distribución simétrica
d. La escala de medición es al menos ordinal.
e. Las observaciones $X_1, X_2, \ldots, X_n$. son independientes.

6.2.4 Método

1. A cada una las parejas relacionadas $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ se les determinará la diferencia del signo $D_i = X_i - Y_i$[4] entre los dos puntajes, colocando este valor en otra columna.
2. A estas diferencias se le asignará un rango que va de menor a mayor sin tomar en cuenta el signo.
3. Ocasionalmente, el puntaje de algún par no presente diferencia alguna, es decir que el valor de $D_i = X_i - Y_i = 0$, en tal caso el par igualado deberá ser descartado del análisis (se procede de igual forma en la prueba de los signos).
4. Puede presentarse que dos o más diferencias (D_i) sean iguales, estos casos reciben el nombre de "ligas", los cuales se les deberá asignar el mismo rango; en donde el valor de éste será el promedio de los rangos que se les habría asignado si las D_i hubieran diferido ligeramente, los demás rangos conservan su asignación previa, es decir si se presentara el caso en donde los valores de D_i sean iguales a –3,-3 y 3. A cada pareja se le asignaría el rango de 2, puesto que $(1+2+3)/3=2$, la siguiente D_i recibiría el rango 4, porque los rangos 1,2,3 ya se usaron.
5. Posteriormente en la columna del rango de las diferencias se el asignara el signo correspondiente de tal manera que se pueda identificar los rangos de las diferencias positivas y negativas (como se menciono anteriormente en el caso de que las $D_i = X_i - Y_i = 0$ se descartan del análisis), por esta razón tenemos que "n" es el número de pares o parejas igualadas menos el número de pares cuyas D_i sean cero.
6. También se puede dar el caso en que existan varias ligas (diferentes de cero), y debido a que éstas puedan dar lugar a ciertas perturbaciones, se recomienda introducir un factor de corrección, el cual esta representado por la siguiente expresión, que se debe incluirse al normalizar la expresión cuando se tienen datos mayores a 25.

$$\frac{\sum t^3 - \sum t}{48}$$

La forma de obtener el valor de t se ejemplifica en la *tabla 6.2-1*.

Tabla 6.2-1

Observación	3	3	4	6	6	6	8	8	9	9	9	9
Rango	1.5	1.5	3	5	5	5	7.5	7.5	10.5	10.5	10.5	10.5
valor de t	2			3			2		4			
valor de t^3	8			27			8		64			

Donde tendríamos que $t=11$ y $t^3=107$ dando como resultado de la ecuación el valor "2".

[4] En el caso de tener solo una muestra el valor de la Y_i se sustituye por el valor de la mediana a probar, es decir $D_i = X_i - M_0$, el resto de los pasos para la realización de la prueba no registran cambio alguno.

6.2.5 Hipótesis

Si se denotan por M_D la mediana común de las D_i, es posible establecer la hipótesis nula de forma simple o compuesta de acuerdo al deseo de investigador de nacer una prueba de un cola o una prueba de dos colas.

-En el caso de tener dos muestras tenemos:

Para el caso de pruebas de una cola se tienen dos casos:

a) Ho: la mediana de la población de las diferencias es menor o igual a cero.
   Ha: la mediana de la población de las diferencias es mayor a cero.
   $$Ho: M_D \leq 0 \quad vs. \quad Ha: M_D \rangle 0$$

b) Ho: la mediana de la población de las diferencias es mayor o igual a cero.
   Ha: la mediana de la población de las diferencias es menor a cero.
   $$Ho: M_D \geq 0 \quad vs. \quad Ha: M_D \langle 0$$

Cuando se desea una prueba de dos colas las hipótesis nula y alternativa se define como:

c) Ho. La mediana de la población de las diferencias es cero
   Ha: La mediana de la población de las diferencias es diferente de cero.
   $$Ho: M_D = 0 \quad vs. \quad Ha: M_D \neq 0$$

-En el caso de tener una sola muestra tenemos:

Para el caso de una prueba de dos colas.
$$Ho: M = M_0 \quad vs \quad Ha: M \neq M_0$$
Para el caso de una prueba de una cola.
$$Ho: M \leq M_0 \quad vs \quad Ha: M > M_0 \quad \text{o bien}$$
$$Ho: M \geq M_0 \quad vs \quad Ha: M < M_0$$

6.2.6 Estadístico de Prueba.

El estadístico de prueba se denotará por T y se define como la suma de los rangos R_i asignadas a aquellas parejas (X_i, Y_i) con el signo menos frecuente, al tomar el rango de la diferencia,
se toman los valores en valor absoluto: $$T_c = \sum_{i=1}^{n} Ri$$

El procedimiento para determinar la significación del valor observado de T depende del tamaño que tenga "n", es decir cuando $n \leq 25$ (muestras pequeñas) y cuando $n > 25$ (muestras grandes).

- Muestras pequeñas

Cuando el tamaño de n es menor o igual que 25, existen tablas[2] de diferentes tamaños de muestra y niveles de significancia para pruebas de una cola o dos colas.

- Muestras grandes

Para $n > 25$ se utiliza una aproximación a la normal como estadística de prueba en donde:

media: $\mu_T = \dfrac{n(n+1)}{4}$ y la desviación estándar : $\sigma_T = \sqrt{\dfrac{n(n+1)(2n+1)}{24}}$

Teniendo que el valor estadístico Z es:

$$Z = \frac{T-\mu_T}{\sigma_T} = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

Donde:
$T = \sum_{i=1}^{n} Ri$ con el signo menos frecuente en valor absoluto y que existan menos de dos ligas en los datos.

En el caso de existir más de dos ligas se debe de utilizar el factor de corrección mencionado anteriormente, incluyendo este en la aproximación a la normal, estandarizando el valor de Z de la siguiente manera:

$$Z = \frac{T-\mu_T}{\sigma_T} = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\sum t^3 - \sum t}{48}}}$$

6.2.7 Regla de decisión.

1.- Cuando el tamaño de $n \leq 25$ tenemos que:

Se rechaza Ho:

Si $T_c \leq T_w$ (valor en tablas) con un nivel de significancia asignado α en la prueba de una cola y $\alpha/2$ en el caso de una prueba de dos colas.

No se rechaza Ho:

Si $T_c > T_w$ (valor en tablas) con un nivel de significancia asignado α en la prueba de una cola y $\alpha/2$ en el caso de una prueba de dos colas.

2.- Cuando el tamaño de $n > 25$ tenemos que:

Se rechaza Ho:

Si $Z_c \geq Z$ donde la Z es el valor encontrado en las tablas de la normal N(0,1) con un nivel de significancia α dado en el caso de una prueba de una cola y $\alpha/2$ en el caso de una prueba de dos colas.

[2] Estas tablas fueron elaboradas por Wilcoxon

No se rechaza Ho:

Si $|Z_c| < Z$ donde la Z es el valor encontrado en las tablas de la normal N(0,1) con un nivel de significancia α dado en el caso de una prueba de una cola y $\alpha/2$ en el caso de una prueba de dos colas.

o también:

Si al valor de Z_c se le ve como la probabilidad de una o dos colas asociado a su ocurrencia, tendríamos:

Se rechaza Ho:

Si $P < \alpha$, donde $P(Z \geq Z_c) = \alpha$ para el caso de una cola, o en el caso de 2 colas $2P(Z \geq Z_c) = \alpha$.

No se rechaza Ho:

Si $P \geq \alpha$, donde $P(Z \geq Z_c) = \alpha$ para el caso de una cola, o en el caso de 2 colas $2P(Z \geq Z_c) = \alpha$.

Para comprender como se debe utilizar esta prueba se mostrarán ejemplos para cada uno de los casos antes presentados:

6.2.8 Ejemplos

- *1er. caso.* Cuando el número de parejas es menor o igual que 25 $(n \leq 25)$ y se tiene una sola muestra. En un estudio de abuso en drogas en una área suburbana, los investigadores encontraron que la mediana del IQ de los arrestados de 16 años o más que abusaron de las drogas es de 107. Suponiendo que los investigadores desean concluir que las medianas del IQ de los arrestados, que abusaron a los 16 años de edad o más grandes en el consumo de drogas, en otra área suburbana es diferente de 107. Los IQ de la muestra aleatoria de 15 personas de la población de interés se encuentra en la *tabla 6.2-2.* ¿Que puede el investigador concluir, dada un $\alpha = 0.05$?.

- *Respuesta:*

Hipótesis

Debido a que la hipótesis del investigador plantea que las medianas son diferentes en las dos áreas suburbanas y no indica en que sentido se presenta esta diferencia tenemos que la hipótesis a probar se trata de una prueba de 2 colas o bilateral.

Ho: La mediana de las diferencias es igual a 107.

La mediana del IQ en él número de personas arrestadas de 16 años y más que abusaron de las drogas no presentan diferencia alguna entre las áreas suburbanas, es decir que la mediana del IQ en las los dos ciudades es igual a 107.

Ha: La mediana de las diferencias es distinta a 107.

La mediana del IQ en él número de personas arrestadas de 16 años y más que abusaron de las drogas presentan una diferencia entre las dos áreas suburbanas, es decir la mediana del IQ es diferente a 107.

Tabla 6.2-2

Número de la Prueba	Puntaje obtenido en IQ	$D_i = X_i - Y_i$ $X_i - 107$	Rango (D_i)	Signo del Rango
1	99	- 8	7	-7
2	100	- 7	6	-6
3	90	-17	11	-11
4	94	-13	10	-10
5	135	28	14	14
6	108	1	1	1
7	107	0	anulado	anulado
8	111	4	5	5
9	119	12	9	9
10	104	-3	4	-4
11	127	20	13	13
12	109	2 *	2.5	2.5
13	117	10	8	8
14	105	-2 *	2.5	-2.5
15	125	18	12	12

Con el valor obtenido en las D_i 2 y −2 se presenta un caso de ligas, el cual se arregla obteniendo el promedio de los rangos que se les asignarían si estos no se repitieran es decir:

$$\frac{2+3}{2} = 2.5$$

También se presenta un caso en que Di es igual a cero y como se sabe éste se tiene que anular. Los valores del estadístico T serán:

$$T_+ = \sum_{i=1}^{n} R_i = 64.5 \text{ y } T_- = \sum_{i=1}^{n} R_i = 40.5$$

con $n = 14$

Estadístico de Prueba:

Sabemos que el estadístico de prueba T se define como la suma de los rangos R_i, con el signo menos frecuente en este caso es $T_- = 40.5$, y como tenemos el caso en que $n \leq 25$ se tiene que utilizar las tablas realizadas por Wilcoxon .

Tabla 6.2-3

N	Nivel de significación para prueba de dos colas		
	.05	.02	.01
6	0	-	-
7	2	0	-
8	4	2	0
9	6	3	2
10	8	5	3
11	11	7	5
12	14	10	7
13	17	13	10
14	21	16	13
15	25	20	16

Tabla 6.2-3. Tabla de valores de T en la prueba de los rangos señalados de pares igualados de Wilcoxon.

En el ejemplo anterior tenemos que:

El valor crítico de T (correspondiente a $\alpha = 0.05$ para una prueba de dos colas) es el que aparece señalado en la *tabla 6.2-3*, correspondiente a $n = 14$, es decir $T = 21$.

Regla de decisión

Como $T_w = 21$ y $T_x = 40.5$, se cumple que $T_c \geq T_w$, revisando la regla de decisión podemos concluir que no se rechaza (la hipótesis Nula) Ho, es decir la mediana del IQ es igual a 107 en el número de personas arrestadas de 16 años y más que abusaron de las drogas en las dos áreas suburbanas.

- *2do Caso.* Cuando él número de parejas es mayor que 25 $(n > 25)$ para dos muestras relacionadas. En un estudio para investigar si las oraciones activas y pasivas se procesan de una manera diferente, se pide a 25 personas que lean un pasaje que contiene una proporción igual de oraciones activas y pasivas que recuerden después del pasaje. El número de oraciones activas y pasivas que recordó correctamente cada sujeto se da en la *Tabla 6.1-4*.

- *Respuesta:*

Hipótesis

Ho: La mediana de la población de las diferencias es cero, es decir las oraciones activas y pasivas se procesan de la misma manera.

Ha: La mediana de las diferencias de la población es distinto de cero, en otras palabras se presenta una diferencia en la forma de procesar las oraciones activas y pasivas.

$$\mathrm{Ho}: \mathrm{M_D} = 0 \quad \text{vs.} \quad \mathrm{Ha}: \mathrm{M_D} \neq 0$$

En el resultado de las d_i tenemos que:

$$T_- = \sum_{i=1}^{n} R_i = 27.50 \quad \text{y}$$

$$T_+ = \sum_{i=1}^{n} R_i = 272.50 \quad \text{con } n = 24$$

Tabla 6.1-4

Resultados de identificación de Oraciones.

Sujeto	Activas	Pasivas	$D_i = A_i - P_i$	Rango (D_i)	Signo del Rango
1	9	10	-1	2.5	-2.5
2	7	4	3	11.5	11.5
3	9	3	6	22	22
4	8	2	6	22	22
5	7	8	-1	2.5	-2.5
6	8	1	7	24	24
7	10	9	1	2.5	2.5
8	10	5	5	19	19
9	12	7	5	19	19
10	6	3	3	11.5	11.5
11	7	7	0	nulo	Nulo
12	9	6	3	11.5	11.5
13	8	5	3	11.5	11.5
14	6	4	2	6.5	6.5
15	10	9	1	2.5	2.5
16	11	9	2	6.5	6.5
17	7	5	2	6.5	6.5
18	5	1	4	16	16
19	7	4	3	11.5	11.5
20	7	4	3	11.5	11.5
21	8	4	4	16	16
22	9	11	-2	6.5	-6.5
23	8	12	-4	16	-16
24	8	3	5	19	19
25	10	4	6	22	22

En este ejemplo se nos presenta el caso de 6 ligas por lo que se debe de usar el factor de corrección:

Observación	1	2	3	4	5	6
Total	2.5	6.5	11.5	16	19	22
valor de *t*	4	4	6	3	3	3
valor de t^3	64	64	216	27	27	27

donde:

$$\frac{\sum t^3 - \sum t}{48} = \frac{425 - 23}{48} = 8.375$$

Estadístico de Prueba:

Tenemos que la suma de los rangos menos frecuentes (en este caso las R_i negativas) donde $T_- = \sum_{i=1}^{n} R_i = 27.5$ y como $n > 25$ se debe utilizar una aproximación a la normal como estadística de prueba, en donde el valor de Z valdría:

$$Z = \frac{27.5 - \frac{24(24+1)}{4}}{\sqrt{\frac{24(24+1)(48+1)}{24} - \frac{425-23}{48}}} = -3.51202598$$

Regla de decisión:

Como el valor de $Z_c = -3.51202598$ utilizando el factor de corrección, y al obtener el valor de Z con $\alpha = 0.05$ para una prueba de dos colas este vale $Z = 1.96$, y al revisar la regla de decisión se cumple que $|Z_c| \geq Z$, por lo que podemos concluir que se rechaza (la hipótesis Nula) Ho.

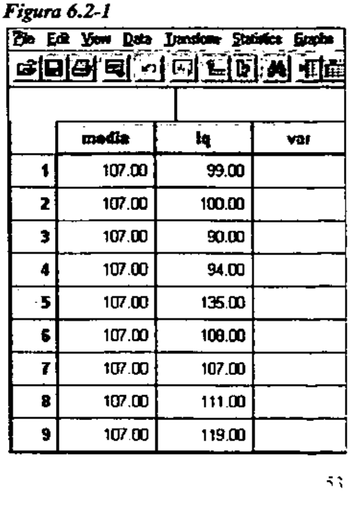
O viendo a Z como la probabilidad del valor asociado a la ocurrencia para una prueba de una cola tenemos $P(-3.51) = 0.0002$ y $2P(-3.51) = 0.0004$ (para una prueba de dos cola). En vista de que la 2P es menor que $\alpha = 0.05$ y como el valor de Z esta en la región de rechazo podemos, concluir que si hay una diferencia en la forma de procesar las oraciones activas y pasivas.

6.2.9 Utilizando el paquete SPSS

En esta parte se presentaran los pasos que se tienen que seguir para realizar correctamente la prueba del signos-rango de Wilcoxon.

1.- Se incorporan los datos de las dos variables (*1er.caso*), para este caso es recomendable incorporar los datos de "media"- Y_i antes que los de "IQ"- X_i, incorporándose cada una de las variables en una columna distinta (recuerde que como se trata de una prueba para datos apareados las dos columnas tienen que tener el mismo número de datos), tal como se muestra a continuación (*Figura 6.2-1*).

Figura 6.2-1



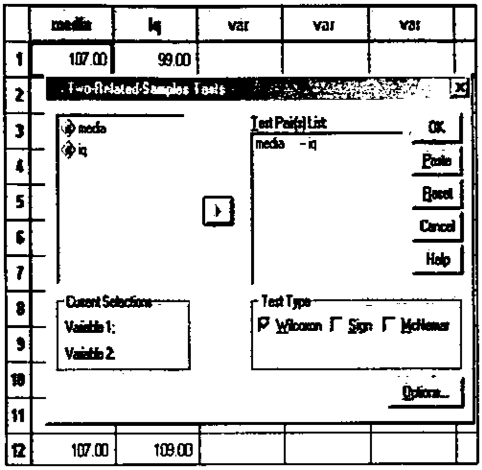
	media	iq	var
1	107.00	99.00	
2	107.00	100.00	
3	107.00	90.00	
4	107.00	94.00	
5	107.00	135.00	
6	107.00	108.00	
7	107.00	107.00	
8	107.00	111.00	
9	107.00	119.00	

2.- Al igual que en la prueba de los signos en la barra de herramientas se debe seleccionar el icono correspondiente a *Statistics*, dentro del cual se selecciona, la opción de *Nonparametric tests*.

3.-Dentro de la opción *Nonparametric tests* se seleccionara *2 Related samples*, ya que la prueba de Wilcoxon opera con dos muestras relacionadas o apareadas, Ver *figura 6.1-1* de la prueba de los signos.

4.-Posteriormente se presenta al usuario una ventana de diálogo en donde se solicita especificar las dos variables que entraran en el análisis.

Figura 6.2-2



En nuestro caso tenemos que la variable "media" se tomará como el minuendo y la variable "IQ" como sustraendo de la diferencia que el paquete calculará para cada caso. Una vez copiados los nombres de las dos variables de análisis en la ventana de la derecha titulada *Tests Pair(s) List*, en el recuadro *Test Type* se elige la opción *Wilcoxon* y finalmente se oprime el botón *OK* ubicado en la parte superior derecha, como se realizó en la prueba de los signos (*Figura 6.2-2*).

Los resultados aparecen en otra ventana llamada *Output1*, en donde el usuario estará en posibilidades de imprimir o guardar como archivo.

Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
IQ - MEDIA	Negative Ranks	6^a	6.75	40.50
	Positive Ranks	8^b	8.06	64.50
	Ties	1^c		
	Total	15		

a. IQ < MEDIA

b. IQ > MEDIA

c. MEDIA = IQ

En la ventana de resultados se presenta la tabla (*Ranks*)que muestra la suma de los rangos positivos (64.50), la suma de los rangos negativos (40.50) y el número de diferencias nulas (1), que el paquete identifica como observaciones ligadas (*ties*).

Test Statistics^b

	IQ - MEDIA
Z	$-.754^a$
Asymp. Sig. (2-tailed)	.451

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

Como se trata de una muestra pequeña, es suficiente observar las suma menor de los rangos y obtener el valor de T de las tablas de Wilcoxon con n = 14 y $\alpha = 0.05$

O también se puede utilizar la segunda tabla (*Test Statistics*) en donde el procedimiento selecciona las diferencias negativas como referencia para los siguientes cálculos por contar con el menor número de casos, el cual corresponde a (40.50) y éste sirve de base para el cálculo del estadístico Z (el utilizado cuando $n > 25$), cuyo valor es -0.754 y la probabilidad asociada $P(-0.754)$ es 0.451[4] para la prueba de dos colas.

Como $0.451 < 0.05$ podemos concluir que no se rechaza la hipótesis Nula (Ho), como anteriormente se había concluido.

2do. Caso.

Para la realización de este ejemplo, se siguen los mismos pasos que se indicaron en el primer ejercicio. Los datos de "oraciones Pasivas"-Y_i se incorporar antes que los datos de las "Oraciones Activas"-Xi. En la imagen siguiente aparecen los cuadros que se obtienen de la ventana de Output, en donde observamos que:

El primer cuadro nos da la suma de los rangos positivos y negativos al igual que el total de los datos anulados.

Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
PASIVAS - ACTIVAS	Negative Ranks	4[a]	6.88	27.50
	Positive Ranks	20[b]	13.63	272.50
	Ties	1[c]		
	Total	25		

a. PASIVAS < ACTIVAS

b. PASIVAS > ACTIVAS

c. ACTIVAS = PASIVAS

Test Statistics[b]

	PASIVAS - ACTIVAS
Z	-3.512[a]
Asymp. Sig. (2-tailed)	.000

a. Based on negative ranks.

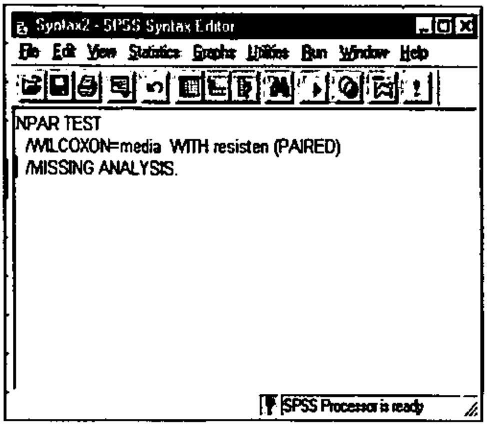
b. Wilcoxon Signed Ranks Test

Como en este caso tenemos que $n > 25$ es mucho más conveniente usar los resultados que nos da el segundo cuadro, ya que nos proporciona el valor de la Z estandarizada (-3.512) y la probabilidad asociada con el supuesto de normalidad 0.000 (caso de 2 colas), si se requiriera la probabilidad asociada a una prueba de una cola este valor se tendría que dividir entre dos. En nuestro caso tenemos una prueba de dos colas y al comparar este con una $\alpha = 0.05$, tenemos que $0.000 < 0.05$, podemos concluir que se rechaza la hipótesis nula, como anteriormente lo mencionamos.

[4] El valor de .451 se obtiene por medio de las tablas de la normal, es decir como se tiene el valor de Z se obtiene su probabilidad asociada. En este caso el valor se encuentra entre .2734 y .2764 interpolando tenemos que $Pr = 0.5 - .2749 = .2251$ resultado si tuviéramos una prueba de una cola, como el ejemplo es de 2 colas tenemos $2(2251) = .4502 \approx .451$.

Por otra parte, si el usuario desea ver el Código de SPSS for windows , este se pueden ver por medio de las instrucciones generadas a partir de la operación con la ventana de diálogo.

Figura 6.2-3



Es decir en la ventana de *2 Related Sample,* y después de elegir las variables que entraran en el análisis, se activa en la parte superior derecha el iconos *Paste*- éste sirve para que el paquete nos muestre el código de la prueba en otra ventana llamada *Syntax1.*
En esta ventana se deberá seleccionar el icono para correr el código y obtener los mismos resultados que con OK, excepto que este código puede ser salvado para posteriormente utilizarlo sin haber ocupado los iconos y opciones del paquete (*Figura 6.2-3*).

6.2.10 Observaciones

Como se puede observar el paquete trata indistintamente el valor de la n para realizar la prueba de Wilcoxon, ya que uno puede utilizar la tabla 1 o 2 según la conveniencia del usuario – claro esta que si uno ocupa los datos del primer cuadro se necesitan las Tablas Wilcoxon y estas están calculadas para valores de "n" menores o iguales a 25.

CAPITULO VII

PRUEBAS NO PARAMETRICAS PARA 2 MUESTRAS INDEPENDIENTES

La técnica paramétrica usual para analizar datos de dos muestras independientes consiste en aplicar una prueba t a las medias de los dos grupos. La prueba t supone que los puntajes (que se suman al calcular las medias) son observaciones independientes de poblaciones distribuidas normalmente con varianzas iguales – se pretende probar igualdad de medias, la cual requiere que las observaciones se midan por lo menos en una escala intervalo o razón.

Como se mencionó en el capítulo anterior, hay ocasiones en las que la prueba t es inaplicable debido a varias razones entre las que se encuentran:

a) Los supuestos y requerimientos de la prueba t son poco realistas para los datos.
b) Se prefiere evitar hacer los supuestos o probar los requerimientos para dar mayor generalidad a sus conclusiones.
c) Sus "puntajes" pueden no ser verdaderamente numéricos y por tanto, no satisfacer la medida de la prueba t.

En circunstancias como las mencionadas anteriormente, se puede escoger para analizar los datos una de las pruebas estadísticas no paramétricas para dos muestras independientes, como son : la prueba de la mediana, la prueba de U de Mann-Whitney[1].

[1] Estas no son las únicas pruebas que se pueden realizar para dos muestras independientes, existen otras como: La prueba de dos muestras de Kolmogorov-Smirnov, La prueba exacta de Fisher, La prueba de la χ^2 para dos muestras independientes, la prueba de rachas de Wald-Wolfowitz, etc.

Estos casos se presentan cuando el uso de dos muestras relacionadas no es práctico ni adecuado, es decir cuando la naturaleza de la variable independiente impide usar a los sujetos como sus propios controles, o cuando no hay buenos pares disponibles – incapacidad del investigador de obtener medidas adecuadas en estudios de pares igualados - pueden usarse dos muestras independientes.

En este diseño, las dos muestras pueden obtenerse con la ayuda de dos métodos:

a) Tomadas al azar de dos poblaciones
b) Asignando al alzar ambos tratamientos a miembros de alguna muestra de orígenes arbitrarios.

En cualquier caso, no es necesario que las muestras sean del mismo tamaño.

Estos procedimientos están basados en el análisis de datos generados por dos muestras independientes, para cada una de las poblaciones de interés, donde las pruebas propuestas determinan la medida en que las diferencias de las muestras constituyen un indicio convincente de una diferencia en el proceso aplicado a ellas.

En donde estos procedimientos sirven para estimar y probar hipótesis acerca de la diferencia entre dos parámetros de localización y procedimientos para pruebas de hipótesis acerca de igualdad de la dispersión de parámetros.

7.1 PRUEBA DE LA MEDIANA

7.1.1 Función

La prueba de la mediana usa uno de los procedimientos más sencillos y usados de todos los métodos no paramétricos que se utilizan para probar si dos muestras independientes[2] (no necesariamente iguales) provienen de poblaciones con medianas iguales. Más exactamente, la prueba de la mediana dará información acerca de la probabilidad de que dos grupos independientes se hayan tomado de poblaciones con la misma mediana, esta prueba es normalmente atribuida a Mood y Westenberg.

Es decir cuando se cuenta con dos muestras independientes y se desea probar si la localización de las distribuciones que las originaron es la misma desde el punto de vista de la mediana y no se puede hacer el supuesto de normalidad requerido por la prueba t de Student para dos muestras independientes[3], la prueba de la mediana es una alternativa para probar este tipo de hipótesis, a pesar de que su eficiencia es relativamente baja con respecto a la prueba t cuando se cumplen las condiciones de normalidad $2/\pi \approx 0.637$[4].

[2] La prueba de la mediana también se usa para examinar si varias muestras de la población tienen la misma mediana, es una generalización de la prueba de la mediana, la cual se vera mas adelante.

[3] La prueba de la t student sirve para probar si dos muestras independientes difieren en sus tendencias centrales, utilizando el supuesto de que las poblaciones se distribuyen normalmente.

[4] La mejor alternativa para probar dos muestras independientes es la prueba U de Mann y Whitney debido a su mayor potencia.

7.1.2 Supuestos

Para la aplicación de la prueba se considera que se dispone de dos muestras aleatorias $X_1, X_2, \ldots X_n$ de tamaño n_1 y $Y_1, Y_2, \ldots Y_n$ de tamaño n_2 y además que se cumplen los siguientes supuestos.

a) La prueba puede usarse siempre y cuando las dos muestras independientes estén, por lo menos, en una escala ordinal de medición.
b) La variable de interés es continua.
c) Las dos muestras se deben de haber tomado en forma independiente, no solamente entre los dos grupos considerados, sino además dentro de cada grupo.

7.1.3 Método

1.- Se ordenan ambas muestras combinándolas de menor a mayor.

$$X_1, Y_2, Y_5, X_7, \ldots X_n$$

2.- Se obtiene la mediana de los datos combinados.
3.- Las observaciones se comparan con la mediana combinada para obtener las frecuencias de observaciones en ambas muestras.
4.- Esas frecuencias se arreglan a una matriz 2×2, tal como se ilustra a continuación

GRUPO I	GRUPO II	TOTAL
Número de Observaciones del grupo 1 mayores a la mediana A	Número de Observaciones del grupo 2 mayores a la mediana B	A + B
Número de Observaciones del grupo 1 menores o iguales a la mediana C	Número de Observaciones del grupo 2 menores o iguales a la mediana D	C + D
A +C =n_1	B +D =n_2	N = n_1 +n_2

7.1.4 Hipótesis

La hipótesis nula supone que las dos poblaciones tienen la misma mediana, es decir, las dos poblaciones tienen la misma probabilidad P de que una observación exceda a la mediana de ambas poblaciones. Sean G_1 y G_2 las dos poblaciones.

Ho: mediana de G_1 = Mediana de G_2

La hipótesis alternativa puede ser

Ha: Mediana de $(G_1) \neq$ Mediana de (G_2)

7.1.5 Estadístico de Prueba

Para realizar esta prueba con muestras pequeñas se tiene que si A y B son el número de observaciones mayores a la mediana del grupo 1 y 2 respectivamente, entonces la distribución muestral de A y B bajo la hipótesis nula (Ho es tal que $A = n_1/2$ y $B = n_2/2$) puede ser designada por medio de la distribución hipergeométrica.

$$P(A,B)=\frac{\binom{n_1}{A}\binom{n_2}{B}}{\binom{N}{A+B}}$$

Donde:
$n_1 = A + C$, $n_2 = B + D$
$N = n_1 + n_2$

Sin embargo, hay que tomar algunas consideraciones para seleccionar el estadístico que se tiene que usar dependiendo del tamaño de la muestra.

1. Si el número de casos es pequeños $N < 30$, se utiliza la prueba exacta de Fisher[5], la cual se basa en el cálculo de la expresión anterior – determinándose la probabilidad de los resultados observados.
2. Cuando él número de los casos es $N \geq 40$ se usa χ_c^2 corregida por continuidad[6], es decir:

$$\chi_c^2 = \frac{N\left(|AD - BC| - N/2\right)^2}{(A+B)\ (C+D)\ (A+C)\ (B+D)}$$

3. Cuando n_1+n_2 está entre 20 y 40 y además ninguna celda tiene una frecuencia esperada menor que 5, se usa χ_c^2corregida por continuidad (presentada anteriormente). Si la más pequeña frecuencia esperada es menor que 5, se usa la prueba de Fisher.

Si la muestra es grande se puede utilizar un procedimiento alternativo.

Este procedimiento consiste en poder aproximar a la distribución hipergeométrica a la binomial y a su vez a la distribución normal, siempre y cuando las condiciones para realizar estas aproximaciones se cumplan.

1.- Para que la aproximación de la binomial por medio de la distribución hipergeométrica sea valida, la población necesita de 8 a 10 casos en la muestra.
2.- La aproximación de la normal por medio de la binomial es buena cuando el tamaño de muestra es grande y cuando la proporción de casos en la población de la característica de interés este cerca de 0.5, es decir cuando np y nq sean mayor que 5, donde n es le tamaño de muestra y p es la proporción de la población que tiene la característica de interés.

Entonces cuando la muestra satisface las condiciones para utilizar una aproximación de la normal, la prueba estadística seria:

$$T = \frac{(A/n_1)-(B/n_2)}{\sqrt{\hat{p}(1-\hat{p})(1/n_1+1/n_2)}} \quad donde \quad \hat{p}=(A+B)/N$$

[5] La prueba exacta de Fisher, en honor a Ronald Fisher, es una versión de la prueba de la suma de rangos cuando los empates llegan a ser tan númerosos que la variable de respuesta solo tiene dos niveles, y al ordenar los datos se obtendría una tabla de 2x2, la cual requiere que los datos sean independientes.

[6] Esta fórmula para obtener la χ_1^2 es conocida como "la corrección de Yates por continuidad", la cuál indica que las frecuencias esperadas deben ser iguales o mayores que 5 cuando gl=1.

7.1.6 Regla de decisión

Cuando $N < 30$

Se rechaza Ho:

Si $2P(A,B) < (\alpha)$[7], donde α es el valor de significancia asignado.

No se rechaza Ho:

Si $2P(A,B) \geq (\alpha)$, donde α es el valor de significancia asignado.

Cuando $N \geq 30$

Se rechaza Ho:

Si $\chi_c^2 \geq \chi_t^2$ (valor de en tablas), α un nivel de significancia asignado.

No se rechaza Ho:

Si $\chi_c^2 < \chi_t^2$ (valor en tablas), α un nivel de significancia asignado.

Si obtenemos la probabilidad asociada (P-value) de la P(χ_c^2) tendríamos:

Se rechaza Ho:

Si $P < \alpha$ donde $P(\chi_1^2 \geq \chi_c^2) = \alpha$, donde α es el valor de significancia asignado.

No se rechaza Ho:

Si $P \geq \alpha$ donde $P(\chi_1^2 \geq \chi_c^2) = \alpha$, donde α es el valor de significancia asignado.

Si se utiliza la aproximación a la normal[8] tenemos:

Se rechaza Ho:

Si $|Z_c \geq Z_T$, donde $Z_T = Z_{1-\alpha/2}$ valor obtenido en las tablas de la normal estándar.

No se rechaza Ho:

Si $|Z_c < Z_T$, donde $Z_T = Z_{1-\alpha/2}$ valor obtenido en las tablas de la normal estándar.

7.1.7 Ejemplos

- *1er. Caso* cuando $N < 30$. Para probar los efectos de diferentes instrucciones sobre los tiempos de reacción simples, 24 personas se asignan aleatoriamente a uno de dos grupos de 12. A uno de los grupos se le dan instrucciones para que responda con la mayor rapidez posible (instrucciones de rapidez), mientras que al otro grupo se le da instrucciones para que responda con la mayor precisión posible (instrucciones de precisión). Sus tiempos de reacción se muestran la *tabla 7.1-1.* ¿Difieren los dos grupos?, usando una $\alpha = 0.01$

[7] Este valor del estadístico de prueba se multiplica por dos ya que la hipótesis nula nos pide una prueba de 2 colas.
[8] Si el estadístico que se usa es la aproximación de la distribución normal se podría realizar una prueba de una cola.

Tabla 7.1-1

Instrucciones de rapidez	172	146	465	170	139	183	160	167	122	175	125	189
Instrucciones de precisión	202	187	217	194	185	173	197	180	167	196	178	177

- Respuesta:

Ordenando los datos del ejemplo anterior para obtener la mediana obtendríamos que esta sería:

$$mediana = \frac{175 + 177}{2} = 176$$

Al comparar esta mediana con cada uno de los grupos de la forma como nos indica la tabla tendríamos las siguientes frecuencias (*Tabla 7.1-2)*.

Tabla 7.1-2

GRUPO I	GRUPO II	TOTAL
A 2	B 10	A + B 12
C 10	D 2	C + D 12
A + C 12	B + D 12	$N = n_1 + n_2$ 24

Hipótesis

Ho: mediana del grupo uno = Mediana de grupo dos. Es decir, el tiempo de reacción de los dos grupos es la misma.

Ha: Mediana de $(G_1) \neq$ Mediana de (G_2)
El tiempo de reacción de los dos grupos es diferente.

Estadístico de Prueba:

Debido a que N<30 tenemos que utilizar la distribución hipergeométrica.

$$P(A,B) = \frac{\binom{n_1}{A}\binom{n_2}{B}}{\binom{N}{A+B}} = \frac{\binom{12}{10}\binom{12}{2}}{\binom{24}{12}} = 0.00161$$

El valor de probabilidad asociado a la prueba es de: $2P(A,B) = 0.00322$.

Regla de decisión

Comparando el valor obtenido en el estadístico de prueba con el nivel de significancia $\alpha = 0.01$ tenemos que $0.00322 < 0.01$ por lo que podemos concluir que se rechaza la hipótesis nula, es decir, el tiempo de reacción de los dos grupos es diferente.

- *2do. caso.* Cuando $N \geq 30$. Hunghes y Wood-Gush investigaron los efectos de la deficiencia del calcio y el sodio en las actividades en gallinas. A las aves se les privó de sodio y calcio, mostrando un incremento esporádico en sus actividades, midiendo los movimientos del cuerpo y los picotes realizados por estas. La *tabla 7.1-3* nos muestra el número de registros de picoteos por ave para 17 de ellas, durante 11 semanas, los pájaros viejos durante 22 días fueron alimentados a base de una adecuada dieta nutricional excepto por el contenido de sodio (cerca del 0.004%), y las 15 aves de control que fueron alimentadas a base de una dieta normal (entera); la ración diaria contenía 0.15% de sodio. ¿Podríamos concluir a partir de estos datos que las poblaciones son diferentes?

Tabla 7.1-3

Aves privadas de sodio	0	0	0	2	17	58	67	67	68	74	79	85	92	95	97	150	181
Aves control	0	0	0	0	0	8	13	13	20	33	34	57	60	64	78		

- *Respuesta:*

Siguiendo las indicaciones del método para poder realizar la prueba, como primer punto se tiene que ordenar los datos de ambos grupos de menor a mayor para obtener la mediana.

Datos	0	2	8	13	17	20	33	**34**	**57**	58	60	64	67	68	74	78	79	85	92	95	97	150	181
Frec.	8	1	1	2	1	1	1	**1**	**1**	1	1	1	2	1	1	1	1	1	1	1	1	1	1

Debido a que los datos son pares tenemos que obtener la mediana promediando los dos valores:

$$mediana = \frac{34 + 57}{2} = 45.5$$

Tabla 7.1-4

GRUPO I	GRUPO II	TOTAL
A 12	B 4	A + B 16
C 5	D 11	C + D 16
A + C 17	B + D 14	$N = n_1 + n_2$ 32

Como sabemos ahora que la mediana vale 45.5 esta se tiene que comparar con los datos de los dos grupos y colocar la frecuencia antes indicada, la *tabla 7.1-4.*

Hipótesis:

Ho: mediana del grupo uno = Mediana de grupo dos.
La mediana de las dos poblaciones de aves es la misma, es decir que no se observa ningún cambio en las aves que consumen calcio y las que no.

Ha: mediana del grupo uno ≠ Mediana de grupo dos
La mediana de las dos poblaciones de aves es distinta, es decir que se observa cambios de actividades en las aves que consumen calcio y las que no.

Estadístico de Prueba:

Sabemos que el estadístico de prueba se define por el tamaño de N, y en este caso tenemos que N>30 por tal motivo se tiene que utilizar χ^2_C corregida por continuidad, es decir:

$$\chi_c^2 = \frac{N\left(|AD - BC| - \frac{N}{2}\right)^2}{(A+B)\ (C+D)\ (A+C)\ (B+D)} = \frac{32\left(|12(11) - 5(4)| - \frac{32}{2}\right)^2}{(16)(16)(17)(14)} = 4.51764706$$

Y el valor de la χ_1^2 de tablas (correspondiente a $\alpha = 0.05$) es de $\chi_1^2 = 3.84$.

Regla de decisión:

Si nos fijamos en el resultado del estadístico de prueba tenemos que como $\chi_c^2 = 4.51764706$ y $\chi_1^2 = 3.84$, se cumple que $\chi_c^2 > \chi_1^2$, revisando la regla de decisión[9] podemos concluir que se rechaza (la hipótesis Nula) Ho, es decir la mediana de las dos poblaciones de aves es distinta, por lo tanto si hay cambio de actividades en las aves que consumen calcio y las que no.

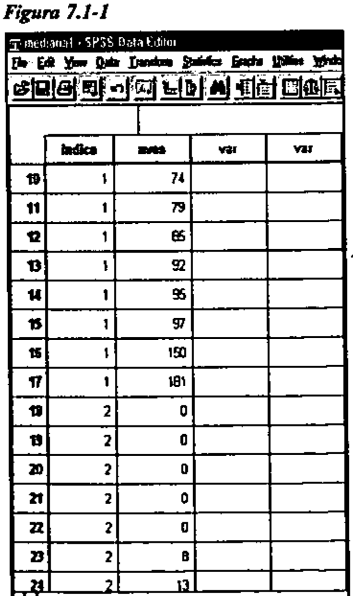
7.1.8 Utilizando el paquete SPSS

Los pasos que se tienen que seguir para la realización correcta de esta prueba son:

1.- El usuario debe construir una estructura con dos variables, una de ellas como indicadora del grupo y la otra conformada con los datos totales de los dos grupos (en este caso), que se utilizaran en el análisis.

Realizando el *1er. Caso* por medio del paquete tenemos que definir dos variables: la primera la denominaremos "indica" que tomará valores de 1 o 2 en función del grupo a que correspondan los datos y la segunda como "aves" que contiene el número de picoteos asociado a la dieta que lo conforman – población que consume calcio y la que no, tal como se muestra en la *Figura 7.1-1*.

Figura 7.1-1



	indica	aves	var	var
10	1	74		
11	1	79		
12	1	65		
13	1	92		
14	1	95		
15	1	97		
16	1	150		
17	1	181		
18	2	0		
19	2	0		
20	2	0		
21	2	0		
22	2	0		
23	2	8		
24	2	13		

2.- En la barra de herramientas se debe seleccionar el icono correspondiente a *Statistics*, dentro del cual se selecciona, la opción de *Nonparametric tests.*

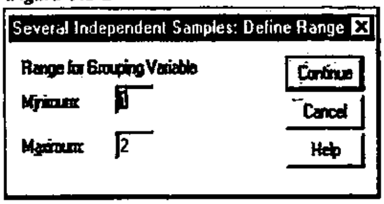
3.-Dentro de *Nonparametric tests* se seleccionará la opción *K independent samples* a pesar que nada mas contamos con 2 muestras independientes, esto se debe a que la prueba de la mediana también es aplicable a más de 2 grupos independientes, la cual es llamada prueba de la mediana generalizada (esta generalización se verá mas adelante).

[9] Revisando la regla de decisión por medio de la probabilidad asociada, se verifica la conclusión anterior, encontrando que $P(\chi_1^2 \geq 4.51764706) = 0.03354694$ es menor que $\alpha = 0.05$

4.-Al seleccionar *K Independent samples* se presentará una ventana de diálogo en la cual el usuario debe identificar la variable del análisis (*Test variable list*), en el ejemplo que estamos realizando esta variable es "**aves**" y la variable de agrupamiento *(Grouping Variable)* la nombrada como "**indica**".

5.-Al elegir la variable de *(Grouping Variable)* se activa el icono (*Define range*) el cual despliega otra ventana adicional (ver *figura 7.1-2*) en donde se deben de definir los valores válidos de las variables, en este caso el 1 y 2, para continuar con el proceso se debe oprimir el botón de *Continue* de la ventana.

Figura 7.1-2



Una vez elegida la variable de análisis en el recuadro de la derecha titulado *Tests variable List*, y la variable indicadora en el recuadro *(Grouping Variable)* se elige la opción *Median* y finalmente se oprime el botón *OK* ubicado en la parte superior derecha. Los resultados de la prueba aparecen en otra ventana llamada *Output1*,en donde el usuario estará en posibilidades de imprimir o guardar como archivo.

La tabla uno de resultados reporta las frecuencias de los datos mayores y menores e iguales a la mediana, diferenciando a que grupo pertenecen.

Median Test

Frequencies

		INDICA	
		1	2
AVES	> Median	12	4
	<= Median	5	11

Test Statistics[a]

		AVES
N		32
Median		45.50
Chi-Square		6.149
df		1
Asymp. Sig.		.013
Yates' Continuity Correction	Chi-Square	4.518
	df	1
	Asymp. Sig.	.034

a. Grouping Variable: INDICA

En el segundo cuadro de resultados se reporta el número de observaciones (32), valor de la mediana común a los dos grupos (45.50), el valor de la estadística χ_1^2 sin corrección por continuidad (6.149), los grados de libertad (1), el valor de la probabilidad asociado a esta χ_1^2 P (0.013) y en la parte inferior de este cuadro se reporta el valor de la χ_c^2 con corrección por continuidad de Yates (4.518), así como el valor del P-value correspondiente a este caso (0.034).

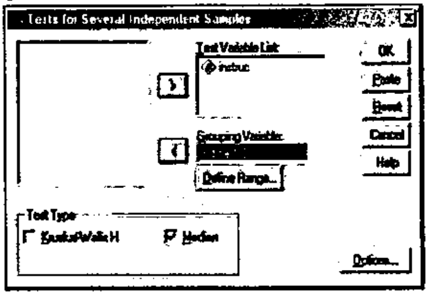
Debido a que el paquete nos proporciona dos tipos de estadísticos de prueba, es recomendable utilizar el estadístico de la χ_c^2 con corrección por continuidad de Yates), ya que este es el recomendado para el número de datos (su fórmula se encuentra el la parte superior).

Por medio del P-value, tenemos que $P(\chi_1^2 \geq 4.518) = 0.034$ y debido a que $0.034 < 0.05$, se puede concluir que estamos rechazando la hipótesis nula, es decir, si hay diferencia entre las aves que consumen calcio y las que no.

2do. Caso

Para la realización de este, se siguen los mismos paso que se indicaron en el primer ejercicio, solo se cambiarán las variables que se tienen que asignar para la realización de la prueba, es decir la variable que servirá para identificar el análisis *(Test variable list)* se llama en este caso **"instruc"** y la variable de agrupamiento *(Grouping Variable)* será la denominada **"indica"**, como se muestra en la *figura 7.1-3*.

Figura 7.1-3



Como en este caso tenemos que $n < 30$ la tabla de resultados contiene lo siguiente:

Median Test

Frequencies

		INDICA	
		1.00	2.00
INSTRUC	> Median	2	10
	<= Median	10	2

Test Statistics[a]

	INSTRUC
N	24
Median	176.0000
Exact Sig.	.003

a. Grouping Variable: INDICA

Al igual que en caso anterior, el primer cuadro nos presenta las frecuencias de los datos mayores y menores e iguales a la mediana, diferenciando a que grupo pertenecen.

El segundo cuadro reporta el número de observaciones (24), el valor de la mediana común a los dos grupos (176) y el valor de la prueba exacta de Fisher : $P = 0.003$.

Con la información proporcionada por el paquete y debido a que $0.003 < 0.05$ podemos concluir que se rechaza la hipótesis nula, como anteriormente lo mencionamos.

7.2 PRUEBA DE U DE MANN-WHITNEY

7.2.1 Función

La prueba U de Mann-Whitney es una de las pruebas estadísticas no paramétricas más potentes, ya que utiliza la mayor parte de la información cuantitativa que poseen los datos. Es la que se emplea con mayor frecuencia como alternativa para la t de Student cuando las medidas no llegan a tener la calidad correspondiente a una escala de intervalos, es decir cuando las condiciones de los datos no permiten hacer los supuestos de normalidad e igualdad de varianzas que requiere para su aplicación, o cuando el investigador desea evitar las hipótesis de la equivalencia paramétrica.

Esta prueba es equivalente a la propuesta en forma independiente de Wilcoxon, la cual se usa para comparar dos muestras independientes o cuando se quiere probar si dos grupos independientes han sido tomados de la misma población, siempre y cuando los datos hayan alcanzado al menos una escala de medición ordinal o si la medición en la investigación es más vaga que la escala de intervalo, la prueba de la U de Mann-Whitney resulta una excelente alternativa.

7.2.2 Potencia de la prueba

Decimos que esta prueba es importante debido a que es la más potente, como lo demuestra varios trabajos realizados[10], entre los resultados más importantes se menciona que la eficiencia relativa de la U con respecto a la potencia de la t, nunca es menor de 0.864 y en el caso de que los supuestos de la t se cumplan (es decir que la población se distribuya normalmente) la potencia de la U alcanza $3/\pi = 0.9550$.

7.2.3 Supuestos

Se considera que los datos provienen de dos muestras aleatorias $X_1, X_2,..,X_n$ de tamaño n_1 y $Y_1, Y_2,..,Y_n$ de tamaño n_2 no necesariamente iguales.

a) Las muestras se han tomado aleatoriamente y en forma independiente, no solamente entre los dos grupo considerados, sino además dentro de cada grupo.
b) Ambas muestras provienen de variables continuas (sin embargo, se puede tolerar un moderado número de "ligas" u observaciones con el mismo valor).
c) La escala de medida es al menos ordinal.
d) La función de distribución de las dos poblaciones difieren solo con respecto a su localización, si ellas difieren en todo.

7.2.4 Método

1. Se determinan los valores de n_1 y n_2, siendo n_1 el número mas pequeño de los dos grupos independientes, y n_2 como el número de casos del grupo más grande.
2. Se combinan las observaciones o puntajes de ambos grupos, identificándolos y ordenándolos de menor a mayor, asignando el rango de 1 al puntaje que sea algebraicamente más bajo hasta N. Los rangos van desde 1 hasta $N = n_1 + n_2$.
3. Se asignará a las observaciones ligadas el promedio de los rangos ligados.

[10] Entre los estadistas que han estudiado la potencia de esta prueba se encuentran: Mood (1954), Hodges y Lehmann (1956), etc.

7.2.5 Hipótesis

Sean $F_x(X)$ y $F_y(Y)$ las funciones de distribución correspondientes a las dos variables X, Y. La variable aleatoria X, se afirma que es estocásticamente más pequeña que Y si $F_x(a) < F_y(a)$ para toda a. La hipótesis que se plantea es:

Ho: $F_x(a) = F_y(a)$ para toda a contra cualquiera de la alternativas.

$$Ha: \begin{cases} F_x(a) > F_y(a) \\ F_x(a) < F_y(a) \\ F_x(a) \neq F_y(a) \end{cases} \quad \text{Para alguna a.}$$

Una forma equivalente de lo enunciado anteriormente es:

$$\mathrm{Ho}: \mathrm{P}(\mathrm{X} > \mathrm{Y}) = \mathrm{P}(\mathrm{X} < \mathrm{Y}) = {}^1\!/_2$$

Contra cualquiera de las alternativas siguientes:

$$Ha: \begin{cases} P(X > Y) > P(X < Y) \\ P(X > Y) < P(X < Y) \\ P(X > Y) \neq P(X < Y) \end{cases}$$

o también la hipótesis nula se podría escribir de la siguiente forma:

Ho: Las distribuciones de frecuencias relativas de las poblaciones X y Y son idénticas.

La hipótesis alternativa podría decir:

Ha: Las distribuciones de frecuencias de las dos poblaciones difieren con respecto a su localización.
Ha: La población 1 es mayor que la población 2.
Ha: La población 1 es menor que la población 2.

7.2.6 Estadístico de Prueba

La estadística de la U de Mann-Whitney está definida como el número de veces en que Y precede de X en la combinación ordenada de los arreglos, pero identificable de las dos muestras aleatorias independientes : $X_1, X_2, \ldots, X_{n1}$ y $Y_1, Y_2, \ldots, Y_{n2}$ dentro de una secuencia simple de $n_1 + n_2 = N$ variables incrementadas en su magnitud. Se asume que las dos muestras forman una distribución continua, entonces la posibilidad de que $X_i = Y_j$ para alguna (i, j) necesitan no ser consideradas.

Entonces para cada par de observaciones X_i, Y_j, con $n i = 1,2,\ldots,n_1; j = 1,2,\ldots,n_2$ se define:

$$W_{ij} = \begin{cases} 1 & \text{si} \quad X_i < Y_j \\ 0 & \text{si} \quad X_i > Y_j \end{cases}$$

La suma de las W's es el estadístico U

$$U = \sum_{i=1}^{n1} \sum_{j=1}^{n2} W_{ij}$$

El valor del estadístico U se puede determinar por medio de del tamaño de n_1 y n_2, es decir :

Caso 1: Cuando ambos $(n_1 \text{ y } n_2)$ no son mayores a 11.

1) Una vez que se hayan combinado e identificado las observaciones de ambos grupos. El método consiste en estudiar a uno de ellos, digamos a el grupo con n_1 casos.
2) El valor de la U (la estadística de prueba) es dado en la clasificación por el número de veces que un puntaje del grupo con n_2 casos precede a un puntaje del grupo con n_1 casos.

Por ejemplo:

Supongamos que tenemos un grupo experimental de 3 casos y un grupo de control de 4 casos, donde tendríamos:

Puntaje E	9	11	15		$n_1 = 3$
Puntaje C	6	8	10	13	$n_2 = 4$

Siguiendo las indicaciones anteriores tenemos:

La *tabla 7.2-1* nos muestra las observaciones ya ordenadas, identificando a que grupo pertenece cada valor. Como se tiene que escoger a cualquiera de los dos grupos a estudiar n_1 en este caso, observamos que el número de veces en los que el valor de n_1 esta precedido por n_2 es:

Tabla 7.2-1

Rango	1	2	3	4	5	6	7
Ordenación	6	8	9	10	11	13	15
Identificación	C	C	E	C	E	C	E
Puntaje	0	0		1		2	

Teniendo: $U = 0 + 0 + 1 + 2 = 3$
Para estos casos existen tablas que pueden usarse para determinar exactamente la probabilidad asociada.

Además puede darse el caso que el valor observado de la U sea tan grande que no aparezca en la subtabla para el valor observado de n_2, este caso se presenta cuando no se toma correctamente el grupo que precede del otro, a este valor grande o muy grande lo llamaremos U', el cual utilizaremos para obtener el valor correcto de U, por medio de una transformación sencilla:

$$U = n_1 n_2 - U' \quad (1)$$

Caso 2: Para valores medianamente grandes, cuando n_2 está entre 11 y 20.

Una vez que se hayan ordenado y combinando las observaciones de menor a mayor y asignado un rango, en el caso en que existan empates en las observaciones éstas pueden manejarse promediando los rangos correspondientes a éstas observaciones y asignando este promedio a cada una. A continuación se procede a se calcular las sumas de los rangos, R_1 y R_2 para las dos muestras y así obtiene el valor de la U por medio de:

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \quad \text{o} \quad U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2 \quad \text{donde} \quad U_1 + U_2 = n_1 n_2$$

Las fórmulas anteriores dan diferentes valores de U. Es el menor de ellos el que nos interesa, por tal motivo el investigador deberá revisar si tiene U' y no U aplicando la transformación (1) antes mencionada – este valor se debe comparar con la U de tablas, usando las mismas que en el caso anterior.

Caso 3 : Para muestras grandes (n_1 o n_2 mayores que 20).

Debido a que los valores críticos máximos de U proporcionados por las tablas son cuando $n_1 = 20$ y $n_2 = 20$, éstas no se pueden usar cuando los valores de n_1 y n_2 sobrepasan éste valor.

Por otro lado se ha sido demostrado que cuando n_1 y n_2 aumentan de tamaño, la distribución muestral de U se acerca rápidamente a la distribución normal, teniendo como valores esperado a E(U) y como varianza a:

$$E(U) = \frac{n_1 n_2}{2} \qquad \text{y} \qquad \sigma_u^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

El valor del estadístico de prueba por medio de la aproximación a la distribución normal estaría definida por:

$$Z = \frac{U - \left(\frac{n_1 n_2}{2}\right)}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

En donde el valor del estadístico de prueba (U) puede estar definido por U_1 ó U_2. El valor obtenido de éste estadístico se tiene que comparar con los valores de la N(0,1) en tablas.

Para muestras grandes (n_1 o n_2 mayores que 20) con ligas.

A pesar de que la prueba de U de Mann-Whitney supone que los puntajes representan una distribución con una continuidad básica[11], esto no se cumple en algunas veces. Ya

[11] Al suponer que los puntajes representan una distribución con una continuidad básica, la probabilidad de que ocurra una liga es cero. Sin embargo, esto no se cumple, ya que las medidas que se utilizan son relativamente burdas (por lo menos en las que se emplean en la investigación científica de la conducta), por tal motivo pueden ocurrir ligas.

que puede darse el caso de tener varias ligas, es decir, que varios grupos de números se repitan, sin embargo si las ligas se dan entre dos o más observaciones dentro del mismo grupo, el valor de U no se altera, pero el valor de U es afectado al ocurrir entre dos o más observaciones de ambos grupos. Aunque la alteración suele ser insignificante, se debe utilizar una corrección para puntajes ligados, que se usa con la aproximación de la curva normal empleada en muestras grandes.

El efecto de los rangos ligados es un cambio en la variabilidad del conjunto de rangos. Así, dicha corrección debe aplicarse a la desviación estándar de la distribución muestral de U.

$$\sigma_U = \sqrt{\left(\frac{n_1 n_2}{N(N-1)}\right)\left(\frac{N^3-N}{12} - \Sigma T\right)}$$ donde $N = n_1 + n_2$ y $T = \frac{t^3 - t}{12}$ (donde t es el número de observaciones ligadas para un rango dado)

Entonces con la corrección del efecto de ligas, el valor de Z se obtendría por medio de:

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\left(\frac{n_1 n_2}{N(N-1)}\right)\left(\frac{N^3-N}{12} - \sum T\right)}}$$

En donde el valor del estadístico de la prueba (U) pueden ser U_1 ó U_2. El valor obtenido en este estadístico se compara con los valores de la N(0,1) en tablas.

7.2.7 Regla de decisión.

Para el caso 1: Cuando ambos $(n_1 \text{ y } n_2)$ no son mayores a 11.

Se rechaza Ho:

Si $P < \alpha$ donde P es $P(U \leq U_c) = \alpha$, donde $U \leq U_c$ (valor en tablas) con un nivel de significancia asignado α, para el caso de una prueba de una cola, y en el caso de una prueba de dos colas el valor de P se duplica es, decir $2P(U \leq U_c) = \alpha$.

No se rechaza Ho:

Si $P \geq \alpha$ donde P es $P(U \leq U_c) = \alpha$, donde $U \leq U_c$ (valor en tablas) con un nivel de significancia asignado α, para el caso de una prueba de una cola, y en el caso de una prueba de dos colas el valor de P se duplica es, decir $2P(U \leq U_c) = \alpha$.

Para el caso 2: Cuando los valores son medianamente grandes, es decir cuando n_2 está entre 11 y 20.

Se rechaza Ho:

Si $U_c \leq U_T$ donde U_T es el valor crítico proporcionado por las tablas.

No se rechaza Ho:
Si $U_c > U_T$ donde U_T [12]es el valor crítico proporcionado por las tablas.

Se puede obtener la probabilidad asociada a pesar de que ésta no aparezca en los valores proporcionados por las tablas del caso 1, el cual rechazaría Ho si $P < \alpha$ donde $P(U \leq U_0) = \alpha$ para una prueba de una cola y $2P(U \leq U_0) = \alpha$ para una prueba de dos colas.

Para el caso 3: Cuando n_1 o n_2 son mayores que 20.

Se rechaza Ho:
Si $|Z_c| \geq Z_T$, donde $Z_T = Z_{1-\alpha/2}$ valor obtenido en las tablas de la normal estándar en el caso de una prueba de dos colas para una prueba de una cola el valor de $Z_T = Z_{1-\alpha}$.
No se rechaza Ho:
Si $Z_c < Z_T$, donde $Z_T = Z_{1-\alpha/2}$ valor obtenido en las tablas de la normal estándar en el caso de una prueba de dos colas para una prueba de una cola el valor de $Z_T = Z_{1-\alpha}$.

Si al valor de Z se le ve como la probabilidad asociado a su ocurrencia (P-value para una o dos colas) tendríamos:

Se rechaza Ho:
Si $P < \alpha$ donde $P(Z \geq Z_c) = \alpha$ para el caso de una cola o en el caso de 2 colas $2P(Z \geq Z_c) = \alpha$.
No se rechaza Ho:
Si $P \geq \alpha$ donde $P(Z \geq Z_c) = \alpha$ para el caso de una cola o en el caso de 2 colas $2P(Z \geq Z_c) = \alpha$.

7.2.8 Ejemplos

- *1er. caso.* Cuando ambos $(n_1 \text{ y } n_2)$ no son mayores a 11. Una firma de contadores, ha desarrollado dos programas de estudio diferentes para ayudar a los contadores jóvenes a prepararse para presentar el examen del Colegio de Contadores Públicos. Para comparar la efectividad de los programas de estudio, se seleccionaron 8 sucursales de la firma y 50 jóvenes contadores en cada sucursal. Los contadores de 4 sucursales fueron entrenados con el programa A y los de las otras 4 con el programa B. Después de presentar el examen, se registró el número de contadores que lo aprobaron (*Tabla 7.2-2*). ¿Proporcionan los datos suficiente evidencia que indique una diferencia entre las distribuciones poblacionales de los programas de entrenamiento A y B ?

[12] Las tablas que existen para los valores críticos de U, son con los siguientes niveles de significación, 0.001,0.1, 0.025 y 0.05. para una prueba de una cola y para la prueba de dos colas, son niveles de significación son : 0.002,0.02,0.05 y 0.10

Tabla 7.2-2

Programa A	Programa B
28	33
31	29
27	35
25	30

-Respuesta:

Hipótesis

Ho: Los dos programas tienen la misma distribución poblacional.
Es decir, los dos programas de estudios presentan la misma efectividad.

Ha: Los dos programas presentan una diferencia en la distribución poblacional. Es decir, que entre los dos programas desarrollados se presenta una diferencia en efectividad.

Siguiendo las indicaciones para este caso tenemos que al determinar los valores de n_1 y n_2, nos damos cuenta que $n_1 = n_2$, por lo que es indiferente asignar que grupo será n_1 y n_2, por lo tanto diremos que n_1 es el número de casos en el grupo A, y n_2 como el número de casos del grupo B.

La *tabla 7.2-3* nos muestra las observaciones ya ordenadas, identificando a que grupo pertenece cada valor, al observar el número de veces en los que el valor de n_1 esta precedido por n_2 tenemos:

Tabla 7.2-3

Rango	1	2	3	4	5	6	7	8
Ordenación	25	27	28	29	30	31	33	35
Identificación	A	A	A	B	B	A	B	B
Puntaje	0	0	0			2		

Teniendo $U = 0+0+0+2 = 2$, buscando el valor de U en tablas para $n_1 = 4$ y $n_2 = 4$ se puede observar que: $P(U \leq 2) = 0.057$, pero como el ejercicio requiere una prueba de 2 colas tenemos que el valor de $2P(U \leq 2) = 0.114$ y comparando esta probabilidad con el nivel de significancia seleccionada tenemos que 0.114>0.05, entonces se puede concluir que no se rechaza Ho es decir, no hay suficiente evidencia que indique una diferencia en las distribuciones de los programas de entrenamiento A y B – Este ejemplo también se puede realizar siguiendo las instrucciones del caso 2.

- ***2do. caso*** Para valores medianamente grandes, cuando n_2 esta entre 11 y 20.Un investigador desea conocer la opinión del público en relación con una nueva ley de protección al consumidor. Como se tiene la sospecha que la opinión de los consumidores de bajos ingresos podría ser diferente al de los de ingresos altos, se seleccionaron 10 residentes de una zona de bajos ingresos y 12 residentes de una zona de altos ingresos. A cada informante se le pidió que indicara, en una escala de 0 a 10, su grado de acuerdo o desacuerdo con la siguiente afirmación: "La ley de protección al consumidor puede ayudar a eliminar las prácticas deshonestas usadas en el comercio. Las respuestas obtenidas se encuentran en la *tabla 7.2-4*, en donde el 1 indica un acuerdo total con la afirmación y el 10 indica un desacuerdo total. ¿Proporcionan estos datos suficiente evidencia que indique que los residentes de las zonas de bajos ingresos tienen menos confianza en la ley de protección al consumidor ?. Use $\alpha = 0.05$

- *Respuesta:*

Tabla 7.2-4

Zona de bajos ingresos	8.5	9.5	9.0	7.5	8.0	8.5	10.0	5.0	6.5	8.5		
Zona de altos ingresos	4.0	3.5	5.5	6.5	7.0	6.0	2.0	5.0	4.5	5.5	3.0	6.5

Hipótesis

Ho: Las distribuciones de frecuencias relativas de las poblaciones X y Y son idénticas. La opinión de los dos grupos de bajo nivel y los de alto nivel es la misma con respecto a la ley de protección al consumidor.

Ha: Las distribuciones de frecuencias relativas de las poblaciones X es menor que Y. Es decir que los residentes de bajos ingresos tiene menor confianza en la ley de protección al consumidor.

Siguiendo las indicaciones anteriores para este caso, al determinar los valores de n_1 y n_2 estos serian:

- A los residentes de la zona más baja de ingresos como n_1 (debido que es el grupo más pequeño de los dos), por este motivo $n_1 = 10$.
- Y a los residentes de la zona mas alta de ingresos como n_2 (debido a que es el grupo más grande los dos), por este motivo $n_2 = 12$.

Posteriormente se procedió a ordenar las observaciones de menor a mayor para que así se les pueda asignar su rango correspondiente, finalmente se calcula las sumas de los rangos, R_1 y R_2 para las dos muestras y así obtiene el valor de la U, como lo muestra la *tabla 7.2-5.*

Tabla 7.2-5 **Cuadro para la suma de los rangos**

Zona de bajo ingreso	**Rango asignado**	**Zona de alto ingreso**	**Rango asignado**
8.5	18	4.0	4
9.5	21	3.5	3
9.0	20	5.5	8.5
7.5	15	6.5	12
8.0	16	7.0	14
8.5	18	6.0	10
10.0	22	2.0	1
5.0	6.5	5.0	6.5
6.5	12	4.5	5
8.5	18	5.5	8.5
		3.0	2
		6.5	12
Suma de los rangos	166.5		86.5

Como hay valores que se repiten, tenemos casos de empates y la forma para aligerar este problema es promediando sus rangos y asignarle este valor a los valores repetidos.

Como es el caso del valor "5" que se repite dos veces, en donde el valor de los rangos que les correspondería a cada uno de ellos serian 6 y 7, entonces:

$$\text{promedio} = \frac{6+7}{2} = 6.5$$

(valor que se utiliza como rango asignado).

Como ya contamos con la suma de los rangos para cada una de las poblaciones, se tiene que proceder al calculo del valor de la U por medio de :

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 = (10)(12) + \frac{10(10+1)}{2} - 166.5 = 8.5$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2 = (10)(12) + \frac{12(12+1)}{2} - 86.5 = 111.5$$

Como se mencionó anteriormente, las anteriores fórmulas dan diferentes valores de U, pero el que nos interesa es el menor de ellos $U_1 = 8.5$. Al buscar el valor de U en tablas para $n_1 = 10$ y $n_1 = 12$, con $\alpha = 0.05$ para una prueba de una cola tenemos que ésta vale $U = 34$, entonces como $U_c \leq U_T$, es decir 8.5<34, se puede concluir que se rechaza Ho, es decir los residentes de bajos ingresos tiene menor confianza en la ley de protección al consumidor.

- *3er. Caso* Para muestras grandes (n_1 o n_2 mayores que 20). La *tabla 7.2.6* muestra el volumen del área sobre 37 surfistas adultos con el defecto septal atrial, de los cuales 26 de ellos no presentaron hipertensión pulmonar y el resto (11) que si lo presentaron[13]. ¿Proporcionan estos datos suficiente evidencia que el bajo volumen de la marea indique no hipertensión pulmonar? Use $\alpha = 0.05$

tabla 7.2-6

Hipertensión pulmonar (ausencia)	Rango asignado	Hipertensión pulmonar (presencia)	Rango asignado
652	27	876	37
556	19.5	556	19.5
618	26	493	7
500	9	348	1
500	9	530	14
526	13	780	35
511	11	569	21
538	15	546	16
440	5	766	33
547	17	819	36
605	24.5	710	30
500	9		
437	4		
481	6		
572	22		
589	23		
605	24.5		
436	3		
724	32		
515	12		
552	18		
722	31		
778	34		
677	28		
680	29		
428	2		
Suma de los rangos	453.5		249.5

- *Respuesta*

Hipótesis:

Ho: El volumen de la marea no afecta a ninguno de los dos grupos.
Ha: El bajo volumen de la marea indica no hipertensión pulmonar.

El valor de n_1 estará conformado por los datos del grupo que presenta hipertensión pulmonar es decir $n_1 = 11$, y n_2 esta conformado por el grupo que no presenta hipertensión pulmonar es decir $n_2 = 26$.

La *tabla 7.2-6* también nos proporciona los valores de la suma de los rangos, que nos sirve para calcular el valor de U:

$$U_1 = (11)(26) + \frac{11(11+1)}{2} - 249.5 = 102.5$$

Para verificar si el valor obtenido anteriormente de la U, es él más pequeños, se realizara la siguiente transformación:

[13] los datos fueron reportados por *Ressl et al* (E-11).

$$U = n_1 n_2 - U' = 286 - 102.5 = 183.5$$

y con esto se puede observar que el valor obtenido en primera estancia es el menor.

Como en este ejemplo una de las observaciones es mayor de 20, es recomendable utilizar el estadístico Z (el cual se distribuye como una normal (0,1)), en donde el valor de éste sería:

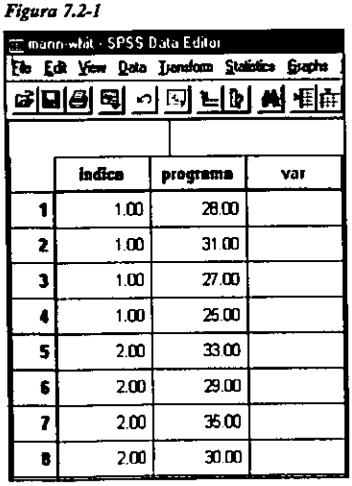
$$Z = \frac{U - \left(\frac{n_1 n_2}{2}\right)}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} = \frac{102.5 - \frac{11(26)}{2}}{\sqrt{\frac{11(26)(11+26+1)}{12}}} = -1.34576996$$

Como tenemos una prueba de una sola cola el valor de $Z_{1-\alpha} = 1.64$ (valor obtenido en tablas) y al comparar estos dos valores tenemos que $|-1.35| < 1.64$, y revisando la regla de decisión podemos concluir que no se rechaza la hipótesis Ho, es decir que el bajo volumen de la marea indica no hipertensión pulmonar.

7.2.9 Utilizando el paquete SPSS

Los pasos para la realización de la prueba por medio de SPSS son:

Figura 7.2-1



	indice	programa	var
1	1.00	28.00	
2	1.00	31.00	
3	1.00	27.00	
4	1.00	25.00	
5	2.00	33.00	
6	2.00	29.00	
7	2.00	35.00	
8	2.00	30.00	

Al igual que la prueba de la mediana el usuario debe construir una estructura con las variables, una de ellas como indicadora de los grupos y la otra conformada con los datos totales de los dos grupos independientes, que se utilizaran en el análisis. *Figura 7.2-1*, datos pertenecientes al *ejemplo-1*.

1.- En la barra de herramientas se debe seleccionar *Statistics*, dentro de la cual se elige la opción de *Nonparametric tests.*

2.-Dentro de *Nonparametric tests* se elige la opción *2 independent samples* , la cual presentará una ventana de diálogo en donde el usuario debe identificar la variable del análisis (*Test variable list*) y la variable de agrupamiento *(Grouping Variable).*

3.- Al elegir la variable de *(Grouping Variable)* se activa el icono (*Define range*) el cual despliega otra ventana en donde se deben de definir los valores de las variables, para continuar con el proceso se debe de oprimir el botón de *Continue* de la ventana.

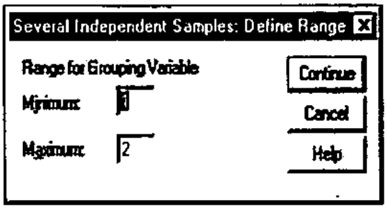
4.-Una vez elegida la variable de análisis (*Tests variable List*) y la variable indicadora *(Grouping Variable)* se elige alguna opción del recuadro *"test type"* en este caso se debe elegir la *Mann-Whitney U,* y finalmente se oprime el botón *OK ó paste* (si se desea obtener el código de la prueba) ubicado en la parte superior derecha.

5.- Los resultados de la prueba aparecen en otra ventana llamada *Output1*,en donde el usuario estará en posibilidades de imprimir o guardar como archivo, esta ventana nos presenta dos cuadros.

1er. Caso.

Para los datos del ejemplo-1 la variable de análisis (*Tests variable List*) seria la denominada **"programa"** y la variable de agrupamiento *(Grouping Variable)* la nombrada **"indica"**. En la ventana de (***Define Range***) definiremos a Minimun con 1 y al Maximun como 2, como lo muestra la siguiente *figura 7.2-2,* para continuar con el proceso se oprime la opción de *Continue* de la ventana.

Figura 7.2-2



Mann-Whitney Test

Ranks

	INDICA	N	Mean Rank	Sum of Ranks
PROGRAMA	1.00	4	3.00	12.00
	2.00	4	6.00	24.00
	Total	8		

Test Statistics[b]

	PROGRAMA
Mann-Whitney U	2.000
Wilcoxon W	12.000
Z	-1.732
Asymp. Sig. (2-tailed)	.083
Exact Sig. [2*(1-tailed Sig.)]	.114[a]

a. Not corrected for ties.

b. Grouping Variable: INDICA

El cuadro uno de resultados reporta la suma de los rangos, para el programa A denominado como "indica 1" (12) y para el programa B "indica 2" (24), además nos da la mediana de la suma de los rangos y el total de casos que tiene cada uno de los grupos.

El segundo cuadro de resultados reporta el valor de la estadística de prueba Mann-Whitney U (2), cuya probabilidad se encuentra en la parte inferior de este cuadro es decir $2P(U \leq 2) = 0.114$ -para una prueba de dos colas - , el valor de la estadística de Wilcoxon W (12)[14], y por ultimo el calculo de la Z(-1.732) a partir de la estadística U, cuyo valor P es (0.083) para una prueba de dos colas– si se requiriera una prueba de una sola cola a este valor se le divide entre dos.

[14] Este valor es el correspondiente a la suma de los rangos de las diferencias menores (12), pero no siempre corresponde al de la suma menor sino al grupo que tenga menor media de rangos, pero recordar que esta estadísticas se usa cuando se tienen 2 muestras relacionadas.

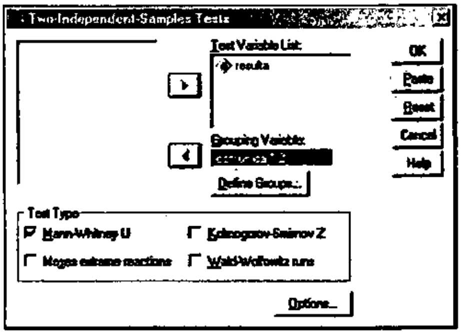
Como en el ejemplo que estamos realizando se requiere probar una prueba de 2 colas, y revisando la regla de decisión utilizando la probabilidad asociada al estadístico Z para compararlo con el nivel de significancia seleccionado (0.083>0.05) se puede concluir que no hay diferencia entre los dos programas propuestos, es decir no se rechaza la hipótesis nula.

2do Caso.

Para introducir los datos se siguen las misma indicaciones que se explicaron anteriormente.

En este ejemplo los nombres de la variables que sirve para identificar el análisis *(Test variable list)* es **"resulta"** y la variable de agrupamiento *(Grouping Variable)* será la denominada **"comunidad"**, como se muestra en la *figura 7.2-3.*

Figura 7.2-3



Al igual que en el ***ejemplo-1***, el cuadro 1 de resultados reporta la suma de los rangos, la mediana de estos y el total de casos para cada uno de los grupos.

Mann-Whitney Test

Ranks

	COMUNIDA	N	Mean Rank	Sum of Ranks
RESULTA	1.00	10	16.65	166.50
	2.00	12	7.21	86.50
	Total	22		

Test Statistics[b]

	RESULTA
Mann-Whitney U	8.500
Wilcoxon W	86.500
Z	-3.405
Asymp. Sig. (2-tailed)	.001
Exact Sig. [2*(1-tailed Sig.)]	.000[a]

a. Not corrected for ties.

b. Grouping Variable: COMUNIDA

El segundo cuadro de resultados reporta el valor de la estadística de prueba Mann-Whitney U (8.5), cuya probabilidad es $2P(U \leq 8.5) = 0.000$ - para una prueba de dos colas - , el valor de la estadística de Wilcoxon W (86.5), y por último el calculo de la Z(-3.405) a partir de la estadística U, cuyo valor P es (0.001) para una prueba de dos colas– si se requiriera una prueba de una sola cola a este valor se le divide entre dos.

Revisando la regla de decisión para una prueba de una cola tenemos que la probabilidad asociada del estadístico Z es: $Z/2 = 0.001/2 = 0.0005$, como este valor es menor a una $\alpha = 0.05$ (0.0005<0.05), podemos concluir que rechazamos la hipótesis nula, es decir, los residentes de bajos ingresos tiene menor confianza en la ley de protección al consumidor.

3er. Caso.

Para este ejemplo solo se presentaran los cuadros de resultados y su regla de decisión, ya que las indicaciones para el ingreso de datos es igual que en los casos anteriores.

Al revisar la regla de decisión por medio de la probabilidad asociada a la estadística de prueba U (0.0905)[15] para una prueba de una cola.

Como 0.0905>0.05 el nivel de significancia asignado (α), podemos concluir que no se rechazamos la hipótesis nula, es decir el bajo volumen de la marea indica no hipertensión pulmonar.

Mann-Whitney Test

Ranks

	GRUPO	N	Mean Rank	Sum of Ranks
HIPERTE	1.00	26	17.44	453.50
	2.00	11	22.68	249.50
	Total	37		

Test Statistics[b]

	HIPERTE
Mann-Whitney U	102.500
Wilcoxon W	453.500
Z	-1.346
Asymp. Sig. (2-tailed)	.178
Exact Sig. [2*(1-tailed Sig.)]	.181[a]

a. Not corrected for ties.

b. Grouping Variable: GRUPO

[15] Este valor se obtiene dividiendo el dato proporcionado por *Exac. Sig.* (.181) entre dos, que es el valor de la probabilidad asociado a U.

CAPITULO VIII

PRUEBAS NO PARAMÉTRICAS PARA K MUESTRAS RELACIONADAS

La técnica paramétrica para probar si varias muestras proceden de poblaciones idénticas, es el análisis de varianza o prueba F. Las suposiciones asociadas con el modelo estadístico que fundamentan la prueba F son:

a) Que los puntajes u observaciones sean sacados independientemente de poblaciones distribuidas normalmente.
b) Que las observaciones en cada combinación bloque-tratamiento tuvieran distribuciones normales con varianzas iguales.
c) Además, la prueba F requiere por lo menos medidas de intervalo en las variables involucradas.

Si se encuentra inadecuadas semejantes suposiciones para los datos, es decir si se encuentra que los puntajes no satisfacen el requisito de medición, o si se desea evitar hacer las suposiciones a fin de incrementar la generalidad de los hallazgos, se puede usar una prueba no paramétrica para k muestras relacionadas , como son : La prueba Q de Cochran[1] y la prueba de Análisis de Varianza de dos clasificaciones por rangos de Friedman[2] ésta última se presenta en este trabajo.

---

[1] Esta prueba es útil cuando la medida de la variable en estudio está en una escala nominal u ordinal dicotomizada y sirve para determinar la probabilidad de que las k muestras relacionadas procedan de la misma población con respecto a la proporción o frecuencia de "éxitos" en las diferentes muestras, no se toma en cuenta en este trabajo debido a que se presentan las mas conocidas.

[2] Análoga a la paramétrica análisis de varianza o prueba F, llamada prueba de Friedman para Análisis de Varianza con dos criterios de clasificación y como su nombre lo indica, esta basada en los rangos de los datos, algunas veces también llamada "diseño de bloques al azar".

Además de evitar las suposiciones y requisitos mencionados, las pruebas no paramétricas de k muestras tiene la ventaja adicional de permitir el examen de significación de datos que son inherentes solamente a clasificaciones o rangos.
El procedimiento no paramétrico que se presentará aquí, sirve para probar la significación de la diferencia entre tres o más grupos, medidos en una escala ordinal . Es decir, la prueba estadística para la hipótesis de nulidad que supone que k muestras han sido sacadas de la misma población o de poblaciones idénticas, en donde las k muestras son del mismo tamaño igualadas de acuerdo con ciertos criterios susceptibles de afectar los valores de las observaciones – la igualación se hace comparando los mismo individuos o casos bajo todas las k condiciones o donde cada uno de los N grupos puede medirse en todas las k condiciones.

Además de lo anterior hay ocasiones o circunstancias en las que se requiere diseñar un experimento de modo que más de dos muestras o condiciones puedan estudiarse simultáneamente, cuando esto ocurre, es necesario usar una prueba estadística que indique si hay una diferencia total entre las k muestras o condiciones, antes de escoger una para probar la significación de la diferencia entre ellas[3], a este tipo de experimentos se denominan "diseños de bloques al azar".

8.1 LA PRUEBA DE FRIEDMAN PARA ANÁLISIS DE VARIANZA CON DOS CRITERIOS DE CLASIFICACIÓN.

8.1.1 Función

La prueba de Friedman desarrollada por el ganador del premio Nóbel, el economista Milton Friedman[4] (1937), se diseño para probar la hipótesis nula de que las distribuciones de probabilidad de los k tratamientos son idénticas, frente a la alternativa de que por lo menos dos de las distribuciones difieren en ubicación. Cuando se tiene limitaciones en los dispositivos de medición o en la distribución de las variables en estudio, esto nos puede dar lugar a severos desvíos en los datos con relación a la hipótesis de trabajo que sustenta el análisis clásico de la varianza[5], como es el caso de la normalidad y homocedasticidad, ya que estos desvíos pueden inducir al investigador a llegar a conclusiones equivocadas en las pruebas de hipótesis mediante la F de Snedecor, a pesar de la conocida robustez de la prueba.

[3] Solamente cuando una prueba total – una prueba de k muestras - nos permita rechazar la hipótesis nula, se justifica el uso subsecuente de un procedimiento para probar las diferencias significativas entre cualquier par de las k muestras, para semejantes procedimientos existen técnicas como las propuestas por Tukey, Duncan etc.
[4] Friedman, M. "The Use of Rank to Avoid the Assumphon of Normality Implicit in the Analysis of Variance." *Journal of the American Statistical Association*, vol. 32(1937), pp. 675-701
[5] El análisis clásico de varianza se basa en los supuestos, de que las observaciones en cada combinación bloque-tratamiento tuvieran distribuciones normales con varianzas iguales.

La prueba de Friedman se basa en un estadístico que es un análogo en rangos de SCT para el diseño aleatorizado de bloques[6] sin interacciones, popularmente conocido como bloques al azar, el cual es la contraparte de la prueba de Kruskal-Wallis, cuando

hay varias muestras relacionadas. Puesto que las muestras han sido igualadas, el número de casos es el mismo en cada una de las muestras. La igualación puede haberse estudiando en el mismo grupo de sujetos en cada una de las k condiciones. O el investigador puede obtener varios conjuntos, compuestos cada uno de k sujetos igualados, para asignar al azar un sujeto de cada conjunto a la primera condición, un sujeto del siguiente conjunto a la segunda condición, etc.

La prueba de Friedman para Análisis de varianza de dos criterios de clasificación por rangos, es útil cuando los datos de las k muestras igualadas están, por lo menos, en una escala ordinal y la ventaja de este procedimiento es que puede utilizarse independientemente de la forma real de distribución de las poblaciones de los "tratamientos". Para que la prueba sea válida no debe haber efectos secundarios. En algunos casos, las dificultades creadas por los efectos secundarios se pueden atenuar aleatorizando o compensando el orden de presentación de los diferentes tratamientos.

8.1.2 Potencia de la prueba

Se puede verificar que cuando $K = 2$ tratamientos con un diseño aleatorizado de bloques, el estadístico de Friedman es el cuadrado del estadístico estandarizado de la prueba de los signos (para una prueba de dos colas) y que su eficiencia asintótica relativa seria similar a la obtenida de $2/\pi = 0.637$.

Para las k muestras la eficiencia asintótica relativa de la prueba de Friedman con relación a la prueba paramétrica usual F depende del número de las k muestras, es decir: Si la población se distribuyera normalmente su eficiencia asintótica relativa seria de $(0.955)k/(k+1)$ y si la población se distribuyera como una uniforme su eficiencia asintótica relativa seria de $k/(k+1)$; Sin embargo esta núnca será menor a $(0.864)k/(k+1)$.

8.1.3 Supuestos

Los supuestos básicos que permiten el empleo de la estadística de Friedman son los siguientes:

a. Los tratamientos se asignan aleatoriamente a las unidades experimentales dentro de los bloques y que el número de bloques (b) o en número de tratamientos (k) sea mayor a 5.

[6] Un diseño aleatorizado de bloques que contiene b bloques y k tratamientos, consiste en b bloques de k unidades experimentales cada uno. Se asignan aleatoriamente los tratamientos a las unidades en cada bloque, y cada tratamiento aparece exactamente una sola vez en cada bloque.

b. Las observaciones de los diferentes bloques (renglones) sean independientes.
c. La variable de interés es continua.
d. Las observaciones de las diferentes columnas están relacionadas por algún criterio de asociación o apareamiento

8.1.4 Método

Para la prueba de Friedman los datos deben colocarse en una tabla con b renglones y K columnas, donde los renglones representan bloques o sujetos y las columnas distintos tratamientos o condiciones (*Tabla 8.1-1*).

Tabla8.1-1

Bloque	Tratamiento						
	1	2	3	...	j	...	k
1	X_{11}	X_{12}	X_{13}		X_{1j}		X_{1k}
2	X_{21}	X_{22}	X_{23}		X_{2j}		X_{2k}
3	X_{31}	X_{32}	X_{33}		X_{3j}		X_{3k}
⋮							
i	X_{i1}	X_{i2}	X_{i3}	...	X_{ij}		X_{ik}
⋮							
b	X_{b1}	X_{b2}	X_{b3}	...	X_{bj}	...	X_{bk}

Después de colocar los datos de un diseño aleatorizado de bloques, para cada bloque (renglón), los valores observados de las respuestas se ordenan para cada uno de las k "tratamientos" del rango 1 (el menor valor en el bloque) al rango K (el mayor valor del bloque); es decir la X_{ij} observación de la tabla (i-ésimo bloque, j-ésimo tratamiento) es sustituida por el rango que corresponde dentro del renglón. Así R_{ij} es el rango de la observación correspondiente al i-ésimo sujeto o bloque del j-ésimo tratamiento.

Tabla8.1-2

Bloque	Tratamiento							
	1	2	3	...	j	...	k	suma
1	R_{11}	R_{12}	R_{13}		R_{1j}		R_{1k}	$\frac{k(k+1)}{2}$
2	R_{21}	R_{22}	R_{23}		R_{2j}		R_{2k}	$\frac{k(k+1)}{2}$
3	R_{31}	R_{32}	R_{33}		R_{3j}		R_{3k}	$\frac{k(k+1)}{2}$
⋮								
I	R_{i1}	R_{i2}	R_{i3}	...	R_{ij}		R_{ik}	$\frac{k(k+1)}{2}$
⋮								
b	R_{b1}	R_{b2}	R_{b3}	...	R_{bj}	...	R_{bk}	$\frac{bk(k+1)}{2}$

Si dos o más observaciones en el mismo bloque están empatados para el mismo rango, entonces se asignará el promedio de los rangos que se habrían asignado a estas observaciones a cada miembro del grupo empatado. Los rangos de los datos pueden esquematizarse en la *Tabla 8.1-2*

Se denotará por R_j la suma de los rangos de la j-ésima columna, en donde la suma de los rangos de cualquier renglón será $\frac{k(k+1)}{2}$ por ser la suma de los primeros k números naturales.

El análogo en los rangos SCT para un diseño aleatorizado por bloques, esta dado por:

$$W = b\sum_{i=1}^{k}\left(\bar{R}_i - \bar{R}\right)^2$$

donde: $\bar{R}_i = R_i / b$ y $\bar{R}_i = (k+1)/2$

Al analizar el análogo de los rangos SCT, se tendría que si la hipótesis nula es verdadera y las distribuciones de probabilidad de las respuestas a los tratamientos no difieren en ubicación, esperaríamos que los $\bar{R}_i$ sean aproximadamente iguales y que el valor resultante W sea pequeño. Si la hipótesis alternativa fuera verdadera, esperaríamos diferencias entre los valores de las $\bar{R}_i$ y los valores grandes correspondientes a W. En lugar de W Friedman consideró el estadístico $F_r = 12W/(k(k+1))$ expresión que llega a convertirse en el estadístico de prueba más conocido (ver estadístico de prueba).

8.1.5 Hipótesis

La hipótesis nula se puede definir como:

Ho: Los rangos asignables a las observaciones de un bloque cualquiera son igualmente probables para todos los tratamientos.

ó también: $Ho: R_1 = R_2 = \ldots = R_k$

Ho: Las poblaciones dentro de los bloques son idénticas.

La hipótesis alternativa se puede definir como:

Ha: Los rangos asignables a las observaciones de un bloque cualquiera tiene mayor probabilidad de ser distinto para cierto tratamiento.

ó también: $Ha: R_i \neq R_j$ para algún par de rangos.

Ha: Por lo menos dos de las distribuciones difieren en su ubicación.

8.1.6 Estadístico de Prueba

La estadística F_r propuesta por Friedman, permite determinar si existen diferencias significativas entre las sumas de los rangos de los k tratamientos, la cual se distribuye aproximadamente como una Ji-Cuadrada con k-1 grados de libertad y tiene la siguiente expresión:

$$F_r = \frac{12}{bk(k+1)}\sum_{i=1}^{k}R_i^2 - 3b(k+1)$$

b = Número de bloques
k = Número de tratamientos
R_i=Suma de los rangos para el i-ésimo tratamiento, en donde el rango de cada medición se calcula según su tamaño relativo dentro de su propio bloque.

Observaciones ligadas

Aunque Teóricamente no pueden ocurrir ligas, ya que uno de los supuestos es que las variables sean continuas, en la práctica suele ocurrir ya que F_r es influenciada en cierto grado por las ligas, por tal motivo es recomendable introducir un factor de corrección, el cual consiste en dividir a F_r entre la siguiente expresión:

$$1-\frac{\sum T}{bk\left(k^2-1\right)}$$

donde :
$T = t^3 - t$ (donde t es el número de observaciones con los mismo puntajes)

El ajuste de la estadística de prueba se realiza, dividiendo a F_r por la expresión anterior es decir la F_r corregida por el efecto de las ligas será:

$$F_r = \frac{\frac{12}{bk(k+1)}\sum_{i=1}^{k} R_i^2 - 3b(k+1)}{1-\frac{\sum T}{bk\left(k^2-1\right)}}$$

8.1.7 Regla de decisión

Se ocupa esta regla tanto para la F_r corregida por el efecto de las ligas como para la prueba F_r sin ligas (normal).

Se rechaza Ho:

Si $F_r \geq \chi^2_{k-1}$ con (k-1) grados de libertad y un nivel de significancia α.

No se rechaza Ha:

Si $F_r < \chi^2_{k-1}$ con (k-1) grados de libertad y un nivel de significancia α.

Si obtenemos la probabilidad asociada de la P(F_r) tendríamos:

Se rechaza Ho:

Si $P < \alpha$ donde $P\left(\chi^2_{k-1} \geq F_r\right) = \alpha$ donde α es el nivel de significancia asignado.

No se rechaza Ho:

Si $P \geq \alpha$ donde $P\left(\chi^2_{k-1} \geq F_r\right) = \alpha$ donde α es el nivel de significancia asignado.

8.1.8 Relación de la Estadística T de Friedman con el Coeficiente de concordancia de Kendall.

Kendall y Babington introdujeron en 1939 en forma independiente una estadística llamada coeficiente de concordancia, la cual se denota por W. La estadística W mide el grado de concordancia en términos de rangos entre los n bloques, es decir ésta se puede ver como una medida de asociación.

La expresión de la W de Kendall en términos de la T de Friedman es:

$$W = \frac{F_r}{b(k-1)}$$

El valor de W se encuentra entre 0 y 1, es decir si todos los bloques estuvieran en concordancia W alcanzaría su valor máximo 1, si defiriera mucho la concordancia W alcanzaría su valor mínimo 0.

8.1.9 Ejemplo

Se realizó un experimento para comparar el tiempo de terminación para tres tareas técnicas de la manera siguiente. Como los tiempos de terminación pueden variar considerablemente de una persona a otra, se pidió a cada uno de los seis técnicos que realizarán las tres tareas.

Tabla 8.1-3

Técnico	Tarea A	Tarea B	Tarea C
1	1.21	1.56	1.48
2	1.63	2.01	1.63
3	1.42	1.70	2.06
4	1.16	1.27	1.27
5	2.43	2.64	1.98
6	1.94	2.81	2.44

Se dieron las tareas a cada técnico en orden aleatorio con suficiente tiempo entre las tareas. ¿Presentan los datos de la *tabla 8.1-3* evidencia suficiente para indicar que las distribuciones de los tiempos de terminación para las tres tareas difieran en ubicación? .Utilice $\alpha = 0.05$.

-Respuesta

Siguiendo las indicaciones del método para realizar esta prueba, este nos indica que ésta se puede realizar siempre y cuando el número de bloque esa mayor que 5 (como en este caso) y que a los datos se les tiene que signar un rango, para cada bloque (renglón) como lo muestra el siguiente *tabla 8.1-4*.

Como se puede dar uno cuenta existen varios empates en los datos, por lo que el valor de sus rangos corresponderá al promedio de éstos.

Tabla 8.1-4

Asignación de los rangos por renglón						
Técnico	Tarea A	Rango	Tarea B	Rango	Tarea C	Rango
1	1.21	1	1.56	3	1.48	2
2	1.63	1.5	2.01	3	1.63	1.5
3	1.42	1	1.70	2	2.06	3
4	1.16	1	1.27	2.5	1.27	2.5
5	2.43	2	2.64	3	1.98	1
6	1.94	1	2.81	3	2.44	2
suma de rangos		$R_1 = 7.5$		$R_2 = 16.5$		$R_3 = 12$

Hipótesis:

Ho: Las distribuciones de los tiempos de terminación es la misma, para las tres tareas.
Ha: Por lo menos dos de las distribuciones difieren en su ubicación.

Estadístico de Prueba:

Para este ejemplo obtendremos los dos estadísticos, para poder ver en cuanto influyen las ligas en la prueba.

1.-El estadístico F_r sin tomar el efecto de ligas valdría:

$$F_r = \frac{12}{bk(k+1)}\sum_{i=1}^{k} R_i^2 - 3b(k+1) = \frac{12}{6(3)(4)}\left[(7.5)^2 + (16.5)^2 + (12)^2\right] - 3(6)(4) = 6.75$$

donde: $k = 3$ tratamientos y $b = 6$ bloque

2.-El estadístico F_r tomando el efecto de ligas valdría:

$$F_r = \frac{\frac{12}{bk(k+1)}\sum_{i=1}^{k} R_i^2 - 3b(k+1)}{1 - \frac{\sum T}{bk(k^2-1)}} = \frac{\frac{12}{6(3)(4)}\left[(7.5)^2 + (16.5)^2 + (12)^2\right] - 3(6)(4)}{1 - \frac{12}{6(3)(3^2-1)}} = 7.36364$$

donde:

$$\sum T = \sum (t_i^3 - t_i) = (2^3 - 2) + (2^3 - 2) = 12$$

Como se puede observa si hay una diferencia entre los dos estadísticos de F_r, por lo que si uno no utiliza es estadístico correcto para los datos proporcionados, esto puede llevarnos a tomar una decisión incorrecta al aceptar o rechazar la hipótesis nula, por lo que es conveniente fijarnos si los datos del problema a realizar presenta casos de ligas. Para comparar el valor de F_r es por medio de una $\chi^2_{k-1,\alpha}$ (con $K - 1 = 2$ grados de libertad). Al consultar la tabla de la $\chi^2_{k-1,\alpha}$ encontramos que $\chi^2_{2,.05} = 5.99$.

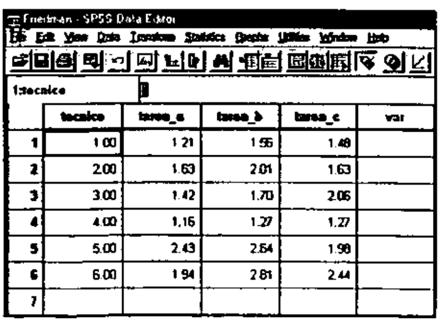
Regla de decisión

Como la prueba correcta para realizar esta prueba es la F_r con corrección por ligas tenemos que $F_r = 7.3636$ y $\chi^2_{2,.05} = 5.99$, se cumple que $F_r > \chi^2_{k-1}$ podemos concluir que se rechaza Ho (la hipótesis Nula), es decir que al menos dos de las tres tareas poseen una distribución de probabilidad que difieren en ubicación.

8.1.10 Utilizando el paquete SPSS

Los pasos que se tienen que seguir para la realización correcta de esta prueba son:

Figura 8.1-1



	tecnico	tarea_a	tarea_b	tarea_c	var
1	1.00	1.21	1.56	1.48	
2	2.00	1.63	2.01	1.63	
3	3.00	1.42	1.70	2.06	
4	4.00	1.16	1.27	1.27	
5	5.00	2.43	2.64	1.98	
6	6.00	1.94	2.81	2.44	
7					

La estructura de los datos se basa en k columnas o variables con el mismo número de observaciones. Cada renglón corresponde a un caso o bloque y los valores de las variables (columnas) a observaciones en diferentes circunstancias o tiempos. *Figura 8.1-1.*

Para ejecutar la prueba de Friedman, el usuario tiene que seguir los siguientes pasos:

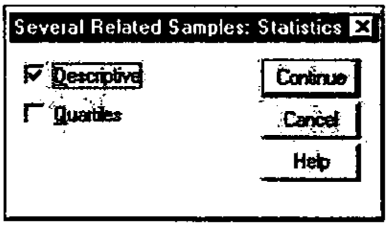
1. En la barra de herramientas se debe seleccionar *Statistics*, dentro de la cual se elige la opción de *Nonparametric tests*.

2. Dentro de *Nonparametric tests* se elige la opción *k Related samples*, la cual presentará una ventana de diálogo en donde el usuario debe seleccionar únicamente las variable del análisis (*Test Variables*) y en tipo de prueba (*Test Type)* la opción de *Friedman*.

3. Si nos interesara conocer el coeficiente de *Kendall* en el recuadro de *"test type"* se debe elegir la opción de *Kendall W* (Si a uno le interesa realizar la prueba de *Friedman* y conocer el coeficiente de concordancia, con solo elegir esta opción es suficiente).

4. Por otra parte si uno esta interesado en conocer las estadísticas básicas de los datos, en la ventana de *k Related samples* en la parte inferior se encuentra la opción de *statistic* que al oprimirse despliega otra ventana. En esta ventana se deben señalar las opciones que uno desea obtener, ya sean las estadísticas descriptivas *(Descriptive)* ó los *Quartiles*, para continuar con el proceso se debe de oprimir el botón de *Continue* en la ventana *(Figura 8.1-2)*.

Figura 8.1-2

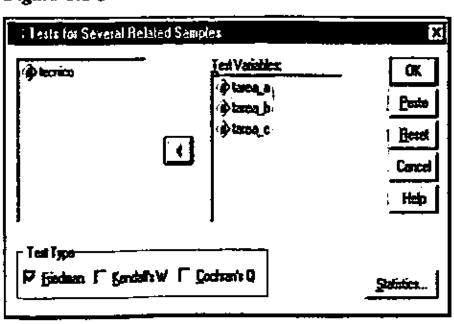


5. Una vez elegida las variables de análisis *(Tests variables)* y la opción del recuadro *"test type"* en este caso la opción de *Friedman o de Kendall W* se oprime el botón *OK ó paste* (si se desea obtener el código de la prueba), para poder terminar con el proceso.

6. Los resultados de la prueba aparecen en otra ventana llamada *Output1*, en donde el usuario estará en posibilidades de imprimir o guardar como archivo. Esta ventana nos presentara dos cuadros.

Para ejemplificar la prueba se realizara el ejercicio por medio del paquete.

Figura 8.1-3



En donde las variables seleccionadas para el análisis (*Tests variables*) serian las variables denominadas "Tareas A, B y C" y la variable indicadora del bloque "Técnico", en este caso solo tiene función informativa, pues éste no es requerido por el análisis. Este tipo de prueba (*Test Type)* se elige en la opción *Friedman*, como lo muestra la *Figura 8.1-3*.

El cuadro uno de resultados reporta el rango promedio de cada tarea.

En el segundo cuadro de resultados reporta el número total de bloques N(6), el valor del estadístico F_r de *Friedman* que en este caso es (7.364), sus grados de libertad (2) y por ultimo el valor de su probabilidad asociada P (0.025).

Friedman Test

Como en nuestro ejemplo, la probabilidad asociada a la χ_1^2 es menor que el nivel de significancia seleccionado (0.025<0.05), se puede concluir que se rechaza la hipótesis Nula (Ho), es decir que al menos dos de las tres tareas, poseen una distribución de probabilidad que difieren en ubicación.

Ranks

	Mean Rank
TAREA_A	1.25
TAREA_B	2.75
TAREA_C	2.00

Test Statistics[a]

N	6
Chi-Square	7.364
df	2
Asymp. Sig.	.025

a. Friedman Test

Kendall's W Test

Ranks

	Mean Rank
TAREA_A	1.25
TAREA_B	2.75
TAREA_C	2.00

Test Statistics

N	6
Kendall's W[a]	.614
Chi-Square	7.364
df	2
Asymp. Sig.	.025

a. Kendall's Coefficient of Concordance

Por otra parte si se hubiera elegido la opción de *Kendall W* tendríamos; una tabla de resultado similar con excepción de que el coeficiente de *Kendall* aparece integrado en el segundo cuadro, como se muestra a continuación.

Como se puede ver el coeficiente de *Kendall* nos da un valor de (0.614), con el que se puede concluir que más de la mitad de los grupos tienen concordancia.

CAPITULO IX

PRUEBAS NO PARAMÉTRICAS PARA K MUESTRAS INDEPENDIENTES

La técnica paramétrica usual para probar si varias muestras independientes proceden de la misma población es el Análisis de Varianza de una clasificación o prueba F[1]. Las suposiciones asociadas con el modelo estadístico en que se basa la prueba F son:

a) Las k muestras sean aleatorias e independientes
b) Que las observaciones independientes sean tomadas de poblaciones distribuidas normalmente y con la misma varianza.
c) Además, la prueba F requiere por lo menos medidas de intervalo de la variable estudiada.

Si se encuentra inadecuadas semejantes suposiciones para los datos, es decir si se encuentra que los datos presentan una medición mas vaga que la escala de intervalos, o si se desea evitar hacer las suposiciones restrictivas de la prueba F a fin de incrementar la generalidad de los hallazgos, se puede usar una prueba no paramétrica para k muestras relacionadas , como son : La prueba de la mediana generalizada y la prueba de Kruskal-Wallis[2] y la prueba de χ^2 para k muestras independientes. Esta ultima se presentará más adelante.

---

[1] One-way análisis of variance F test

[2]Estas pruebas no paramétricas tiene la ventaja adicional de permitir que datos inherentes solamente a clasificaciones (en una escala nominal) o a rangos (en una escala ordinal) sean examinados en cuanto a significación.

Como los valores de las muestras casi siempre difieren en cierto grado y el problema es determinar si detrás de las diferencias muestrales observadas hay diferencias entre las poblaciones o si son meramente variaciones al azar que se esperarían entre muestras aleatorias de la misma población. Por el motivo anterior, las pruebas que se presentan en este capitulo se utilizan para probar la significación de diferencias entre tres o más grupos o muestras independientes. Es decir, se trata de técnicas estadísticas para probar la hipótesis de nulidad de que las k muestras independientes, se recogieron de la misma población (o de k poblaciones idénticas) o de poblaciones con parámetros iguales, en donde las k muestras no necesariamente tienen que ser del mismo tamaño.

9.1 PRUEBA DE LA MEDIANA GENERALIZADA

9.1.1 Función

Como se mencionó anteriormente, la prueba de la mediana sirve para probar si dos muestras independientes provienen de la misma población desde el punto de vista de la mediana, la generalización de esta prueba, consiste en determinar si K grupos independientes (no necesariamente de igual tamaño) han sido recogidos de la misma población o de poblaciones con medianas iguales.

9.1.2 Supuestos

a) La prueba puede usarse siempre y cuando las K muestras independientes estén, por lo menos, en una escala ordinal de medición.
b) Cada una de las k muestras independientes $F_1(x), F_2(x), \ldots, F_k(x)$, presentan una distribución continua.
c) Las k muestras se deben de haber tomado en forma independiente, no solamente entre las muestras consideradas, sino además dentro de cada grupo.
d) Si toda la población tiene la misma mediana, entonces para cada población la probabilidad p es el mismo, donde éste valor excede al valor de la mediana común.

9.1.3 Método

1. Se combinan las K muestras de menor a mayor
2. Se debe determinar el puntaje medio de las K muestras combinadas, es decir se encuentra la mediana común de todos los puntajes de los k grupos.
3. Después se remplaza el valor de cada puntaje por un más (+) cuando es mayor que la mediana común y por un menos (-) cuando es menor que la mediana común.
4. Esas frecuencias se pueden colocar en una tabla de contingencia (arreglo matriz de $2 \times k$) con el número de observaciones que representan las frecuencias de

cada uno de los k grupo mayores y menor a la mediana, tal como se ilustra en la tabla de contingencia para la prueba de la mediana generalizada (*tabla 9.1.1*).

Tabla 9.1.1

	Grupo I	Grupo II		Grupo K	Total
Número de Observaciones mayores a la mediana	O_{11}	O_{12}	...	O_{1k}	a
Número de Observaciones menores o iguales a la mediana	O_{21}	O_{22}	...	O_{2k}	b
Total	n_1	n_2		n_k	N

En donde:

O_{ij}: Es la frecuencia observada del i-ésimo grupo, de la j-ésima muestra, con $i = 1,2$ y $j = 1,2,\ldots k$

a: Es el número total de observaciones que son mayores al valor de la media combinada.

b: Es él número total de observaciones que son menores o iguales al valor de la media combinada.

$$Donde: N = \sum_{i=1}^{k} n_i = a + b$$

9.1.4 Hipótesis

La hipótesis nula y alternativa se plantean en dos formas opcionales, primero con referencias a la distribución.

Ho: Las distribuciones de las K poblaciones son idénticas
Ha: Al menos una de las K distribuciones es diferente.

o con referencia a las medianas ya que la prueba es sensible a sus diferencias.

La hipótesis nula

Ho: Las medianas de las K muestras son iguales, es decir:
Ho: mediana de $F_1(x)$= Mediana de $F_2(x)$=. . . =Mediana de $F_k(x)$

La hipótesis alternativa

Ha: Supone que alguna de las k muestras presenta una mediana diferente con respecto a la mediana común.
Ha: $F_i(x) \neq F_j(x)$, para al menos una de las poblaciones la mediana es diferente.

9.1.5 Estadístico de Prueba

Esta estadística se obtiene por medio del valor de la χ^2 con la siguiente formula:

$$\chi_c^2 = \sum_{i=1}^{r}\sum_{j=1}^{k}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$

Donde:
O_{ij} = número observado de casos clasificados en el renglón i de la columna j.
E_{ij} =número de casos esperados conforme a Ho para ser clasificados en la hilera i de la columna j, es decir $E_{ij} = \frac{a(n_i)}{N}$

Puede demostrarse que χ_c^2 se aproxima a la distribución χ^2 con $gl = (k-1)(r-1)$, donde k es el número de columnas y r es el número de hileras. En la prueba de la mediana generalizada tenemos que r = 2, y así: $gl = (k-1)(r-1) = (k-1)(2-1) = (k-1)$

9.1.6 *Regla de decisión*

Se rechaza Ho:

Si $\chi_c^2 \geq \chi^2$ (valor buscado en tablas) con $(k-1)$ grados de libertad.

No se rechaza Ho:

Si $\chi_c^2 < \chi^2$ (valor buscado en tablas) con $(k-1)$ grados de libertad.

Si obtenemos la probabilidad asociada (P-value) tendríamos:

Se rechaza Ho:

Si $P < \alpha$ donde $P(\chi_{k-1}^2 \geq \chi_c^2) = \alpha$, con α un valor de significancia asignado.

No se rechaza Ho:

Si $P \geq \alpha$ donde $P(\chi_{k-1}^2 \geq \chi_c^2) = \alpha$, con α un valor de significancia asignado.

9.1.7 *Ejemplo*

En un estudio diseñado para determinar la distribución de agua en el miocardio y las concentraciones celulares de electrólitos cardíacos , Polimeni (espacio extracelular y distribución iónica del ventrículo) un método diseñado para encontrar la medida del espacio extracelular en el músculo ventricular de dos grupos de ratas, el nephrectomized y un grupo de ratas intactas. Se desea conocer si se puede concluir que estos datos provienen de poblaciones con medianas diferentes.

Tabla 9.1-2

Ratas Nephrectomized		**Ratas intactas**
Grupo I	**Grupo II**	**Grupo III**
0.185	0.189	0.219
0.187	0.193	0.204
0.209	0.176	0.219
0.194	0.195	0.234
0.175	0.169	0.233
0.197	0.183	0.194
0.188	0.185	0.209
0.185	0.179	0.195

La tabla 9.1.2 muestra los resultados obtenidos. Utilice una $\alpha = 0.005$

- *Respuesta*

Hipótesis:

Ho: la mediana de las tres poblaciones es la misma.
Ha: Al menos la mediana de alguna de las tres poblaciones es la distinta.

Siguiendo las indicaciones para realizar la prueba, éstas nos indican que se tiene que obtener el valor de la media común, combinando los 3 grupos.

Como podemos ver en la *tabla 9.1-3*, ésta contiene los datos ordenados de menor a mayor de los 3 grupos, en donde obtenemos que la mediana se encuentra entre datos señalados:

$$\text{mediana} = \frac{0.193 + 0.194}{2} = 0.1935$$

En donde tenemos que $N = 24$

Tabla 9.1-3

Datos ordenados en forma descendente		
0.169	0.187	0.197
0.175	0.188	0.204
0.176	0.189	0.209
0.179	**0.193**	0.209
0.183	**0.194**	0.219
0.185	0.194	0.219
0.185	0.195	0.233
0.185	0.195	0.234

Como sabemos ahora que la mediana vale 0.1935 ésta se tiene que comparar con cada uno de los grupos y colocar los datos en la tabla de contingencia antes indicada quedando los como lo indica la tabla la *tabla 9.1-4.*

Tabla9.1-4

GRUPO I	**GRUPO II**	**GRUPO III**	**TOTAL**
4 3	4 1	4 8	12
4 5	4 7	4 0	12
8	8	8	24

Como:

El valor de $a = b = 12$ y que $n_i = 8$ tenemos que $E_{ij} = \frac{a(n_i)}{N}\frac{12(8)}{24} = 4$ es el mismo en toda la tabla.

Estadístico de Prueba:

Esta estadística de se obtiene por medio del valor de la χ^2 con la siguiente fórmula:

$$\chi_c^2 = \sum_{i=1}^{r}\sum_{j=1}^{k}\frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \frac{(3-4)^2}{4} + \frac{(1-4)^2}{4} + \frac{(8-4)^2}{4} + \frac{(5-4)^2}{4} + \frac{(7-4)^2}{4} + \frac{(0-4)^2}{4} = 13$$

Y el valor de la χ_2^2 de tablas (correspondiente a $\alpha = 0.005$) es de $\chi_1^2 = 10.957$.

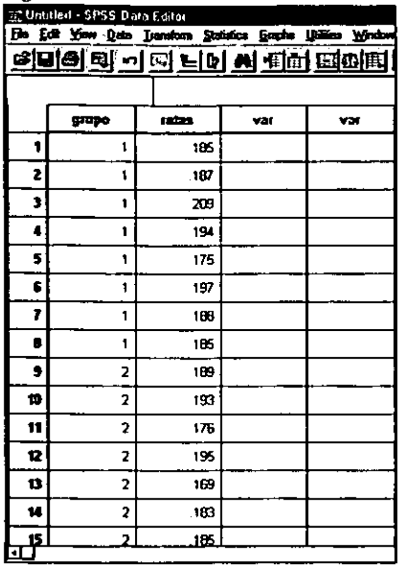
Regla de decisión:

Como $\chi_C^2 = 13$ y $\chi_1^2 = 10.597$, se cumple que $\chi_c^2 > \chi_{k-1}^2$, revisando la regla de decisión podemos concluir que se rechaza la hipótesis Nulas (Ho), es decir al menos una mediana de las tres poblaciones es la distinta.

9.1.8 Utilizando el paquete SPSS

Los pasos que se tienen que seguir para la realización correcta de esta prueba son:

Figura 9.1-1



Al igual que la prueba de la mediana el usuario debe construir una estructura con 2 variables, una de ellas como indicadora del y la otra conformada con los datos totales de las k muestras independientes, que se utilizarán en el análisis. *Figura 9.1-1,* datos pertenecientes al *ejemplo.*

1.- En la barra de herramientas se debe seleccionar *Statistics*, dentro de la cual se elige la opción de *Nonparametric tests.*

2.-Dentro de *Nonparametric tests* se elige la opción *k independent samples* , la cual presentará una ventana de diálogo en donde el usuario debe identificar la variable del análisis (*Test variable list*) y la variable de agrupamiento *(Grouping Variable).*

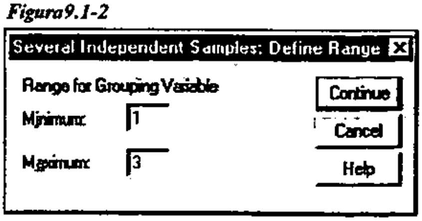
3.- Al elegir la variable de *(Grouping Variable)* se activa el icono (*Define range*) el cual despliega otra ventana en donde se deben de definen los valores válidos de las variables, para continuar con el proceso se debe de oprimir el botón de *Continue* de la ventana.

4.-Una vez elegida la variable de análisis (*Tests variable List*) y la variable indicadora *(Grouping Variable)* se elige alguna opción del recuadro *"test type"* en este caso se debe elegir la *Median* y finalmente se oprime el botón *OK ó paste* (si se desea obtener el código de la prueba) ubicado en la parte superior derecha.

5.- Los resultados de la prueba aparecen en otra ventana llamada *Output1*,en donde el usuario estará en posibilidad de imprimir o guardar como archivo, esta ventana nos presenta dos cuadros.

Para el *ejemplo* la variable de análisis (*Tests variable List*) seria la denominada **"ratas"** y la variable de agrupamiento *(Grouping Variable)* la nombrada **"grupo"**, en la ventana de (*Define Range*) definiremos a Minimun con 1 y al Maximun con 3 como lo muestra la *figura 9.1-2.*

Figura9.1-2



Median Test

Frequencies

		GRUPO		
		1	2	3
RATAS	> Median	3	1	8
	<= Median	5	7	0

Test Statistics[b]

	RATAS
N	24
Median	.19350
Chi-Square	13.000[a]
df	2
Asymp. Sig.	.002

a. 6 cells (100.0%) have expected frequencies less than 5. The minimum expected cell frequency is 4.0.

b. Grouping Variable: GRUPO

El cuadro uno de resultados reporta las frecuencias de los datos mayores y menores e iguales a la mediana diferenciando a que grupo pertenecen.

El segundo cuadro de resultados reporta el número de observaciones (24), valor de la mediana común (0.19350), el valor de la estadística Chi-Square (13), sus grados de libertad (2) y por último el valor de su probabilidad asociado P (0.002)

Como tenemos que la probabilidad asociada a la χ_1^2 es menor que el nivel de significancia seleccionado (0.002<0.005), se puede concluir que se rechaza la hipótesis nula (Ho), es decir la mediana de alguna de las tres poblaciones es la distinta.

9.2 ANALISIS DE VARIANZA POR RANGOS DE KRUSKAL-WALLIS

9.2.1 Función

Kruskal y Wallis propusieron en 1952 un procedimiento alternativo no paramétrico al de la prueba F del análisis de varianza, en el caso en que los supuestos de normalidad u homogeneidad de varianzas no se cumplieran adecuadamente. Esta prueba se presentó como una extensión a k muestras independientes de la prueba U de Mann - Whitney para dos muestras independientes y fue identificada a partir de entonces con los nombres de sus autores.

La prueba de Kruskal-Wallis es extremadamente útil para decidir si K muestras independientes proviene de la misma población (la hipótesis nula) o que al menos una de las poblaciones difieren en localización (hipótesis alternativa); a veces resulta conveniente pensar que la prueba de Kruskal-Wallis es una prueba de igualdad de medias de tratamientos[3] porque es un procedimiento diseñado para ser sensible en pruebas de diferencias entre medias. La prueba supone que los datos son independientes y que los resultados de la variable de respuesta están compuestos por datos ordinales continuos, tomando en consideración los rangos de las observaciones.

[3] La prueba de Kruskal-Wallis es un procedimiento muy usado en diseño de experimentos, cuando el supuesto de normalidad no se cumple.

Podemos observar que la extensión de la prueba de la mediana y la prueba de Kruskal-Wallis pueden usarse en los mismos datos, es decir satisfacen requisitos similares para los datos en el estudio, sin embargo la prueba de Kruskal-Wallis es más eficiente pues usa más la información de las observaciones. Convierte los puntajes en rangos, mientras que la extensión de la prueba de la mediana los convierte simplemente en signos más o menos. Así la prueba de Kruskal-Wallis preserva la magnitud de los puntajes en un grado mayor que la extensión de la prueba de la mediana.

Por otra parte cuando se ha obtenido un resultado significativo con la prueba de Kruskal-Wallis, lo único que se sabe es que existe alguna diferencia en la localización de las muestras. Sin embargo, no se sabe el por qué de la diferencia, por tanto sólo se puede concluir que las muestras llevan a proporciones diferentes. Esto no es problema en las pruebas de dos muestras, porque un resultado significativo implica que una muestra tiene resultados más altos que la otra, con más de dos muestras las posibilidades se incrementan.

Entonces para poder elegir entre las diferentes posibilidades con frecuencia es útil comparar las muestras por pares, sin embargo éste procedimiento introduce una complicación ya que la manera en que se construyó la distribución nula para la prueba de dos muestras sólo la hace precisa si se realiza exactamente una prueba. Por lo tanto cuando se realiza más de una prueba los niveles de significación cambian, es decir el nivel de α se puede incrementar en forma considerable. Específicamente, a medida que más pruebas se lleven a cabo será más probable que se cometa un error de tipo I.

Por tal motivo existen procedimientos que tratan de mantener la tasa de error por experimento en un valor particular de α, los cuales son llamados procedimientos de comparaciones múltiples, que proporcionan métodos para investigar los datos en un intento por localizar diferencias que pueden ser interesantes. Estos se realizan de una manera más adecuada sólo si la prueba de Kruskal-Wallis completa es significativa, entre los métodos de comparaciones múltiples[4] se encuentran el método propuesto por Duncan (1955) denominado "La prueba de Intervalos Múltiples de Duncan", la prueba de Tukey (1953) entre otras, claro está que esto no es tema del presente trabajo[5].

9.2.2 Potencia de la prueba

La prueba de Kruskal-Wallis parece ser la más eficiente de las pruebas no paramétricas para k muestras independientes, ya que comparada con la prueba paramétrica más poderosa la prueba F en las condiciones donde las suposiciones asociadas con el modelo estadístico de la prueba F son satisfechas, la prueba de Kruskal-Wallis tiene una eficiencia asintótica relativa de $3/\pi = 95.5\%$ (Andrews, 1954) y nunca menor a 0.864

[4] Para mayor información consulta el libro de Diseño y Análisis de Experimentos de Montgomery
[5] Ver anexo

9.2.3 Supuestos

Para la aplicación de la prueba se considera que se dispone de K muestras aleatorias independientes de tamaño no necesariamente iguales, los cuales se pueden esquematizar en la siguiente forma:

Muestra I	Muestra II	...	Muestra K
X_{11}	X_{21}	...	X_{k1}
X_{12}	X_{22}	...	X_{k2}
.	.	...	.
.	.	...	.
.	.		.
X_{1n_1}	X_{2n_2}	...	X_{kn_k}

Los supuestos básicos que sustentan la prueba son:

a. Los datos para el análisis consisten de K muestras aleatorias.
b. Las k muestras se deben de haber tomado en forma independiente, no solamente entre las muestras consideradas, sino además dentro de cada grupo.
c. Cada una de las K muestras independientes presentan una distribución continua.
d. La prueba puede usarse siempre y cuando las K muestras independientes estén, por lo menos en una escala ordinal.
e. Las poblaciones son idénticas excepto por una posible diferencia en localización por lo menos en una de ellas.

9.2.4 Método

1.- Se ordenan las K muestras combinándolas de menor a mayor, asignándoles un rango.

2.- En el caso de que dos o más muestras sean iguales (sus rangos corresponderán al promedio de ellos); ver *tabla 9.2-1.*

tabla 9.2-1

Muestra	Rango (R)
11	1
12	2 $\}$ = rango$(2+3)/2 = 2.5$
12	3
14	4

3.- El valor de los rangos se colocarán en la misma forma en que se encontraban las muestras.

4.- Se obtiene la suma de los Rangos de las K muestras (columnas).

9.2.5 Hipótesis

La hipótesis nula y alternativa se plantean en dos formas opcionales, primero con referencias a la distribución.

Ho: Las distribuciones de las K poblaciones son idénticas .
Ha: Al menos una de las K distribuciones es diferente.

ó con referencia a las medianas ya que la prueba es sensible a sus diferencias.

Ho: Las medianas de las K muestras son iguales.
Ha: De las k poblaciones al menos una tiene la mediana diferente.

9.2.6 Estadístico de Prueba

Designada por H la estadística de prueba sigue una distribución aproximadamente a la χ^2_{k-1}. La aproximación es mejor a medida que se dispone de tamaños de muestra más grandes.

$$H = \frac{12}{N(N+1)}\sum_{i=1}^{k}\frac{R_j^2}{n_j} - 3(N+1)$$

Donde:
K = número de muestras
n_j = número de casos en la muestra de orden j
$N = \sum n_j$ número de casos de todas las muestras combinadas
R_j = suma de rangos en la muestra de orden j.

Observaciones ligadas

Debido a que la H es influenciada en cierto grado por las ligas, es recomendable introducir un factor de corrección, el cual consiste en dividir a H entre la siguiente expresión:

$$1 - \frac{\sum T}{N^3 - N}$$

Donde :
$T = t^3 - t$ (donde t es el número de observaciones con los mismo puntajes)
N = número de observaciones en las K muestras juntas es decir, $N = \sum_{i=1}^{K} n_i$.

La expresión general para H corrigiendo el efecto de las ligas es:

$$H_c = \frac{\frac{12}{N(N+1)}\sum_{j=1}^{K}\frac{R_j^2}{n_j} - 3(N+1)}{1 - \frac{\sum T}{N^3 - N}}$$

La razón de corregir los empates es incrementar el valor de H y así volver el resultado más significativo que si no se hubiera realizado la corrección. La magnitud del factor de corrección depende de la longitud de las ligas, es decir de los valores de t, así como del porcentaje de las observaciones aplicadas.

9.2.7 Regla de decisión

La significancia del valor calculado de H se verifica mediante tablas de la Ji – cuadrada, pero para los casos $K \leq 3$ y $n_i \leq 5$ se aconseja el empleo de valores tabulados. Tabla K[6]

Para los casos $K \geq 3$ y $n_i \geq 5$ la regla de decisión seria:

Se rechaza Ho:
Si $H_c \geq \chi^2_{k-1}$ con un nivel de significancia α dado.
No se rechaza Ho:
Si $H_c < \chi^2_{k-1}$ con un nivel de significancia α dado.

Si obtenemos la probabilidad asociada de la $P(H_c)$ tendríamos:

Se rechaza Ho:
Si $P < \alpha$ donde $P(\chi^2_{k-1} \geq H_c) = \alpha$ cuya α es el nivel de significancia asignado
No se rechaza Ho:
Si $P \geq \alpha$ donde $P(\chi^2_{k-1} \geq H_c) = \alpha$ cuya α es el nivel de significancia asignado

9.1.8 Ejemplo

Un ingeniero de desarrollo de productos está interesado en maximizar la resistencia a la tensión de una nueva fibra sintética que se empleará en la manufactura de camisas de hombre. El ingeniero sabe por experiencia que la resistencia es influida por el porcentaje de algodón presente en la fibra. Además, él sospecha que elevar el contenido de algodón incrementará la resistencia, al menos inicialmente. También sabe que el contenido de algodón debe variar aproximadamente entre 10 y 40% para que la tela resultante tenga otras características de calidad que se desea. El ingeniero decide probar muestras (o probetas) a cinco niveles de porcentaje de algodón: 15,20,25,30 y 35%. Supóngase que se desea probar que existe diferencia entre las resistencias medias en los 5 niveles, los datos se encuentran en la *tabla 9.2-2.*

Tabla9.2-2

Porcentaje de algodón				
15	**20**	**25**	**30**	**35**
X_{1j}	X_{2j}	X_{3j}	X_{4j}	X_{5j}
7	12	14	19	7
7	17	18	25	10
15	12	18	22	11
11	18	19	19	15
9	18	19	23	11

Respuesta:

Siguiendo las indicaciones del método para realizar esta prueba, ésta nos indica que se necesita ordenar los datos de menor a mayor de las 5 muestras para que se les pueda asignar un rango.

[6] Ver anexos

La *tabla 9.2-3* nos muestra los datos combinados y su asignación de rango.

Como se puede dar uno cuenta existen varios empates en los datos por lo que sus rangos corresponderán al promedio de éstos.

Tabla 9.2-3

Datos	R	Datos	R	Datos	R	Datos	R	Datos	R
7	2	11	7	14	11	18	16.5	19	20.5
7	2	11	7	15	12.5	18	16.5	19	20.5
7	2	11	7	15	12.5	18	16.5	22	23
9	4	12	9.5	17	14	19	20.5	23	24
10	5	12	9.5	18	16.5	19	20.5	25	25

Como necesitamos la suma de los rangos correspondientes a cada una de las muestras, el valor de los rangos se colocarán en la forma en que se encontraban éstas *Tabla 9.2-4.*

Tabla 9.2-4

Porcentajes de algodón									
15		20		25		30		35	
X_{1j}	R_{1j}	X_{2j}	R_{2j}	X_{3j}	R_{3j}	X_{4j}	R_{4j}	X_{5j}	R_{5j}
7	2	12	9.5	14	11	19	20.5	7	2
7	2	17	14	18	16.5	25	25	10	5
15	12.5	12	9.5	18	16.5	22	23	11	7
11	7	18	16.5	19	20.5	19	20.5	15	12.5
9	4	18	16.5	19	20.5	23	24	11	7
Suma de Rango	27.5		66		85		113		33.5

-Respuesta

Hipótesis:

Ho: La resistencia media de las muestras es la misma para las 5 poblaciones.

Ha: Al menos una de las resistencias medias de las 5 muestras es distinta.

Estadístico de Prueba:

Debido a que encontramos que se presentan 6 casos de ligas, se utilizará la H con el factor de corrección, el cual consiste en:

$$H_c = \frac{\frac{12}{N(N+1)}\sum_{j=1}^{K}\frac{R_j^2}{n_j} - 3(N+1)}{1-\frac{\sum T}{N^3-N}} = \frac{\frac{12}{25(25+1)}\left[\frac{27.5^2}{5}+\frac{66^2}{5}+\frac{85^2}{5}+\frac{113^2}{5}+\frac{33.5^2}{5}\right]-3(25+1)}{1-\frac{180}{25^3-25}} = 19.0636576$$

Donde:

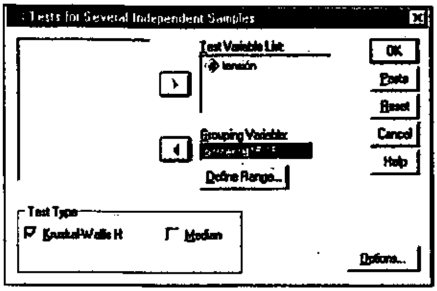
$$\sum T = \sum\left(t_i^3 - t_i\right) = \left(3^3-3\right)+\left(3^3-3\right)+\left(2^3-2\right)+\left(2^3-2\right)+\left(4^3-4\right)+\left(4^3-4\right) = 180$$

Y el valor de la χ_4^2 de tablas (correspondiente a una $\alpha = 0.01$) es de $\chi_4^2 = 13.28$.

Regla de decisión:

Como $H_c = 19.06$ y $\chi_4^2 = 13.28$, se cumple que $H_c > \chi_{k-1}^2$, por tal motivo podemos concluir que se rechaza la hipótesis nula (Ho) es decir, al menos una de las resistencias medias de las 5 muestras es distinta.

Figura 9.2-3

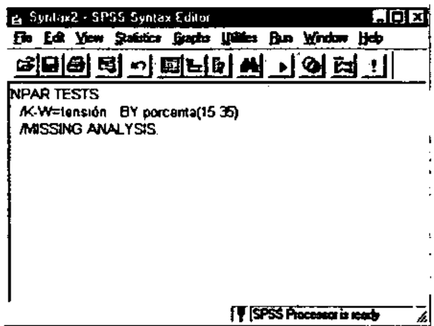


Para el ejemplo anterior la variable de análisis (*Tests variable List*) sería la denominada **"tensión"** y la variable de agrupamiento *(Grouping Variable)* la nombrada **"porcenta"**, en la ventana de (*Define Range*) definiremos a *Minimun* con 15 y al *Maximun* con 35, como lo muestra la *figura 9.2-3.*

Supongamos que antes de obtener los resultados, nos interesara conocer el código del la prueba , este se obtiene oprimiendo el icono de *"Paste"* en la ventana de *k independent samples*, en lugar de *OK* el código aparecerá en otra ventana denominada *"Syntax2"* ver *figura 9.2-3.*

Una vez obtenido el código para continuar con el proceso, se debe elegir la opción de *Run* para obtener la ventana de resultados *(Ouptup1)*- ver *Figura 9.2-4.*

Figura 9.2-4



Kruskal-Wallis Test

Ranks

	PORCENTA	N	Mean Rank
TENSIÓN	15	5	5.50
	20	5	13.20
	25	5	17.00
	30	5	22.60
	35	5	6.70
	Total	25	

Test Statistics[a,b]

	TENSIÓN
Chi-Square	19.064
df	4
Asymp. Sig.	.001

a. Kruskal Wallis Test

b. Grouping Variable: PORCENTA

El cuadro uno de resultados reporta la suma de los rangos para cada uno de los grupos a analizar, el número de datos que contiene cada uno de los grupos y el número total de observaciones.

El segundo cuadro de resultados reporta, el valor del estadístico H (19.064), el cual utiliza la formula de la H corregida, sus grados de libertad (4) y por último el valor de su probabilidad asociada P (0.001)

Como nuestro ejemplo se necesita probar una hipótesis de 2 colas tenemos que la probabilidad asociada a la χ^2_1 es menor que el nivel de significancia seleccionado (0.001<0.01), por lo que se puede concluir que se rechaza la hipótesis nula (Ho), es decir al menos una de las resistencias medias de las 5 muestras es distinta.

CAPITULO X

ANÁLISIS DE TABLAS DE CONTINGENCIA

Una tabla de contingencia es aquella en donde los niveles de la variable explicativa forman los renglones mientras que los niveles de la variable de respuesta forman las columnas. En donde las casillas de la tabla indican el número de observaciones que se obtienen como resultado de cada una de las posibles combinaciones de renglones y columnas.

Es decir se considera una población (o una muestra) compuesta por N individuos sobre los que se pretende analizar simultáneamente dos atributos o factores. La tabla estadística que describe estos N individuos, es la denominada "Tabla de Contingencia" la cual consiste en una tabla de doble entrada.

Lo que se plantea con las tablas de contingencia es el análisis simultáneo de dos características de una misma población y conocer la posible dependencia[1] u homogeneidad[2] entre las 2 clasificaciones, aún cuando los procedimientos de cálculo sean los mismos, esto dos conceptos tiene básicamente diferentes sentidos de interpretación.

Instintivamente, se puede afirmar que el hecho de que dos variables sean independientes entre sí, se traducirá en que los valores que tome una de las variables no estará influidos por la modalidad de que adopte la otra. Este tipo de problemas es

---

[1] Este concepto es de gran importancia ya que al analizar conjuntamente dos características sobre una misma población, debemos prestar atención al hecho de que una variable puede estar relacionada (asociada) con la otra o, por el contrario puede ser independientes.

[2] El problema de homogeneidad se presenta cuando interesa investigar si en las diferentes poblaciones estudiadas los valores o categorías de cada una de las manifestaciones se presentan en la misma proporción.

frecuente dentro del análisis de conteos ya que está relacionado con la independencia entre dos métodos para clasificar eventos.

En este capitulo se clasifican las tablas de contingencia en dos tipos:

1.- Tablas de contingencia de 2 x 2.
2.- Tablas de contingencia de c x k.

A pesar de que el primer grupo es un caso particular de las tablas de contingencia de c x k, ya que son tablas en las que los dos factores considerados presentan únicamente dos categorías mutuamente excluyentes y se merecen tal distinción debido a que, tanto el concepto de independencia u homogeneidad entre variables como la forma de detectarla y medirla, se utilizan métodos particulares con respecto a las tablas de orden superior.

10.1 TABLAS DE CONTINGENCIA

10.1.1 Función

Una pregunta que frecuentemente se hacen los investigadores en estudios en los cuales los elementos de una muestra deben ser categorizados de acuerdo con dos criterios de clasificación, es saber si estos son independientes[3], es decir si no hay ninguna asociación entre las variables.

O dicho de otra manera ellos quieren investigar la dependencia (o contingencia) entre los criterios de clasificación, ya que si dos variables no son asociadas (es decir, si ellos son independientes), conociendo el valor de uno no nos ayudara a determinar el valor de la otra variable por el mismo asunto. Por otro lado, si dos variables son asociadas, el conocimiento de uno es útil prediciendo el valor probable del otro.Para este tipo de problemas el análisis de los resultados mediante el uso de la distribución Ji-cuadrada aplicada a tablas de contingencia es sin duda, uno de los procedimientos estadísticos más populares.

10.2 TABLAS DE CONTINGENCIA DE c x k

En un sentido más estricto, una tabla de contingencia es un arreglo matricial de c renglones y k columnas en cuyas c x k celdas se registran conteos o frecuencias referentes a la muestra y cuyos elementos deben de haberse medido en un nivel nominal u ordinal.

Suponga que se tiene k poblaciones y que se extrae una muestra aleatoria de tamaño n_j de cada población (j=1,2,...,k). Cada observación de las k muestras pueden ser

[3] Dos variables son independientes si la distribución de uno de ninguna manera depende de la distribución del otro

clasificada en una de c diferentes categorías. Los datos pueden ser desplegados en una tabla de contingencia como la que se muestra en la *tabla 10.1-1.*

Tabla 10.1-1

Primer criterio de clasificación	**Segundo criterio de clasificación**						
Categoría	**Muestra (Tratamientos)**						
	1	2	...	j	...	K	Total
1	n_{11}	n_{12}	...	n_{1j}	...	n_{1k}	$n_{1.}$
2	n_{21}	n_{22}	...	n_{2j}	...	n_{2k}	$n_{2.}$
⋮							
I	n_{i1}	n_{i2}	...	n_{ij}	...	n_{ik}	$n_{i.}$
⋮							
C	n_{c1}	n_{c2}	...	n_{cj}		n_{ck}	$n_{c.}$
Total	$n_{.1}$	$n_{.2}$	...	$n_{.j}$	...	$n_{.k}$	N

Se denotará:

O_{ij} como el número de observaciones de la i-ésima categoría perteneciente a la j-ésima muestra.

$n_{i.}$ como el total de observaciones pertenecientes a todas las muestras que quedan contenidas en la i-ésima categoría.

$n_{.j}$ como el total de observaciones pertenecientes a todas las categorias que quedan contenidas en la i-ésima muestra.

Es conveniente aclarar que las hipótesis a probar mediante tablas de contingencia, aun cuando los procedimientos de cálculo son lo mismos, tienen básicamente dos sentidos diferentes:

a) Como hipótesis de igualdad de proporciones en los diferentes niveles de cierta clasificación, cuando las observaciones provienen de dos o más poblaciones.
b) Como hipótesis de independencia entre dos criterios de clasificación aplicables a los elementos de una misma población, cuando el objetivo es valorar la relación que existe entre los dos criterios o variables.

10.2.1 Prueba de la Ji-Cuadrada para probar igualdad de proporciones (homogeneidad en las poblaciones)

La disposición de las observaciones es la que se muestra en la *tabla 10.1-1.*

10.2.1.1 Supuestos

a. Las k muestras son aleatorias.
b. Los resultados de las diferentes muestras son mutuamente independientes.
c. Cada observación puede ser categorizada en una y sólo una de las c diferentes categorías.

10.2.1.2 Hipótesis

Sea P_{ij} la probabilidad de que un elemento de la j-ésima población o muestra seleccionado al azar, quede clasificada en la i-ésima categoría.

La hipótesis nula se plantea como sigue:

La probabilidad (proporción) de pertenece a cualquiera de las c categorías es la misma para cualquier elemento de la i-ésima muestra.

$$Ho: P_{i1} = P_{i2} = \ldots = P_{ic}$$

La hipótesis alternativa:

La probabilidad de pertenecer a cualquiera de las c categorías es diferente para la menos una muestra o tratamiento.

$$Ho: P_{ij} \neq P_{ij}$$

10.2.2 Prueba de la Ji-Cuadrada para probar independencia.

Suponga que se dispone de una muestra aleatoria de tamaño N y que las observaciones de la muestra pueden clasificarse de acuerdo a dos criterios. Al usar el primer criterio cada observación puede asociarse con uno de los c renglones y al usar el segundo criterio se asocia con una de las k columnas.

La disposición de las observaciones es la que se muestra en la *tabla 10.1-1* con la excepción de que en este caso las n_{ij} no se establecen previamente, sino que son aleatorias.

10.2.2.1 Supuestos

a. Cada observación tiene la misma probabilidad de ser clasificada en el i-ésimo renglón y en la j-ésima columna independientemente de cualquier otra observación.
b. Las observaciones en la muestra pueden ser clasificadas de acuerdo a dos criterios. Al usar el primer criterio cada observación puede asociarse con una de las c diferentes categorías (renglones) de acuerdo al segundo criterio.
c. Las variables pueden ser inherentemente categóricas o pueden ser variables cuantitativas cuyas mediciones sean capaces de ser clasificadas en las categorías numéricas mutuamente exclusivas.

10.2.2.2 Hipótesis

La hipótesis nula se plantea como sigue:

Ho: los dos criterios de clasificación son independientes, es decir el evento "la observación pertenece al i-ésimo renglón" es independiente del evento "la misma observación pertenece a la j-ésima columna" para toda i,j.

La hipótesis alternativa:

Ha: Los dos criterios de clasificación no son independientes.

La proposición anterior puede traducirse en términos probabilísticos de la siguiente manera:

Sabemos que si dos clasificaciones son independientes entre sí, la probabilidad de una celda será igual al producto de sus respectivas probabilidades de renglón y de columna de acuerdo con la ley multiplicativa de la probabilidad, es decir si P_i es la probabilidad de pertenecer al i-ésimo renglón y P_j la probabilidad de pertenecer a la j-ésima columna se cumple que:

$$Ho: P_{ij} = P_i P_j \text{ para toda } i,j$$
$$Ha: P_{ij} \neq P_i P_j \text{ para alguna } i,j$$

10.2.3 Estadístico de prueba

El estadístico de los dos casos antes mencionados se realiza por medio de una distribución Ji-Cuadrada aproximada, en donde se define la frecuencia esperada E_{ij} correspondiente a la celda con coordenadas (i, j) como:

$$E_{ij} = N\left(\frac{n_i}{N}\right)\left(\frac{n._j}{N}\right) = \frac{n_i \, n._j}{N}$$

La estadística de prueba se denota por X^2 y su fórmula de cálculo es:

$$X^2 = \sum_{i=1}^{c}\sum_{j=1}^{k}\left[\frac{(O_{ij} - E_{ij})^2}{E_{ij}}\right]$$

Cuando la hipótesis es verdadera, esta estadística se aproxima a una χ^2 con (c-1)(k-1) grados de libertad

10.2.4 Regla de decisión

Para el valor obtenido del estadístico de prueba.

Se rechaza Ho:

Si $X^2 > \chi^2$ con $(c-1)(k-1)$ grados de libertad y un nivel de significancia α.

No se rechaza Ha:

Si $X^2 \leq \chi^2$ con $(c-1)(k-1)$ grados de libertad y un nivel de significancia α.

Si obtenemos la probabilidad asociada de la $P(X^2)$ tendríamos:

Se rechaza Ho:

Si $P\left(\chi^2_{(c-1)(k-1)} \geq X^2\right) < \alpha$ donde α es el nivel de significancia asignado.

No se rechaza Ho:

Si $P\left(\chi^2_{(c-1)(k-1)} \geq X^2\right) \geq \alpha$ donde α es el nivel de significancia asignado.

10.2.5 Recomendaciones para una tabla de c x k

Para una tabla de contingencia con más de un grado de libertad, Cochran recomienda que la prueba se use sólo si al menos el 20% de las celdas tienen una frecuencia esperada no menor a 5 y que ninguna celda tenga una frecuencia esperada menor de 1.

10.3 TABLA DE CONTINGENCIA DE 2 X 2.

Cuando se tienen dos categorías de cada uno de los dos criterios de clasificación, el resultado es una tabla de contingencia conformada por dos renglones y dos columnas, la cual merece atención especial, ya que el estadístico de prueba tiene la siguiente expresión:

$$X^2 = \frac{N(ad - bc)^2}{(a+c)(b+d)(c+d)(a+b)}$$

Tabla 10.1-2

Primer criterio de clasificación	**Segundo criterio de clasificación**		
	1	2	Total
1	a	b	a+b
2	c	d	c+d
Total	a+c	b+d	N

Sin embargo, Yates (1934) propuso una corrección al procedimiento del estadístico X^2 cuando se tiene una tabla de contingencia de 2x2. Este procedimiento es conocido como "corrección por continuidad de Yates" el cual consiste en sustraer 0.5N de el valor absoluto de ad-bc en el numerador de la siguiente ecuación:

$$X^2 = \frac{N(|ad - bc| - 0.5N)^2}{(a+c)(b+d)(c+d)(a+b)}$$

La propuesta de este procedimiento es correcto para el uso de una distribución continua χ^2 que es aproximada por medio de la distribución discreta X^2. Como esta prueba es un caso particular de la tabla de contingencia c x k, todo lo anterior se cumple.

10.3.1 Recomendaciones para una tabla de 2 x 2

1.- Cuando N>40, se usa X^2 corregida por continuidad es decir, la fórmula de Yates.
2.- Cuando N está entre 20 y 40, la prueba X^2, puede usarse en el caso en que todas las frecuencias esperadas sean de 5 o más. Si la frecuencia esperada más pequeña es menor que 5, se usa la prueba de exacta de Fisher.
3.-Cuando N<20, se usa la prueba de Fisher en todos los casos.

10.4 EJEMPLOS

1.- Tablas de contingencia de c x k.

Prueba de la Ji-Cuadrada para probar independencia.

Los datos de la siguiente tabla reportados por Monteiro, muestran 2764 residentes de Rhode Island clasificados de acuerdo a sus ingresos y el tiempo transcurrido a la ultima consulta con su psicólogo ver *tabla 10.1-3*. Se desea conocer si estos datos son evidencia suficiente para indicar que existe una relación entre los ingresos y el tiempo transcurrido desde la última consulta a su médico. En otras palabras nosotros deseamos concluir que las dos variables son no independientes.

Tabla 10.1-3

Ultima consulta con su psicólogo				
Ingresos	**6 meses antes**	**De 7 meses a un año**	**Mas de un año**	**Total**
Menor de $3000	186	38	35	259
$3000-$4999	227	54	45	326
$5000-$6999	219	78	78	375
$7000-$9999	355	112	140	607
Mas de $10,000	653	285	259	1197
Total	1640	567	557	2764

Respuesta:

Siguiendo las indicaciones del método para realizar esta prueba, éste nos indica que se necesitan obtener las frecuencias esperadas de cada una de las celdas es decir:

Tabla 10.1-4 **Tabla de frecuencias esperadas**

$E_{11} = (259)(1640)/(2764) = 153.68$	$E_{12} = (259)(567)/(2764) = 53.13$	$E_{13} = (259)(557)/(2764) = 52.19$
$E_{21} = (326)(1640)/(2764) = 193.43$	$E_{22} = (326)(567)/(2764) = 66.87$	$E_{23} = (326)(557)/(2764) = 65.70$
$E_{31} = (375)(1640)/(2764) = 222.50$	$E_{32} = (375)(567)/(2764) = 76.93$	$E_{33} = (375)(557)/(2764) = 75.57$
$E_{41} = (607)(1640)/(2764) = 360.16$	$E_{42} = (607)(567)/(2764) = 124.52$	$E_{43} = (607)(557)/(2764) = 122.32$
$E_{51} = (1197)(1640)/(2764) = 710.23$	$E_{52} = (1197)(567)/(2764) = 245.55$	$E_{53} = (1197)(557)/(2764) = 241.22$

Hipótesis:

Ho: los ingresos y el tiempo transcurrido desde la última consulta a su médico son independientes
Ha: Las dos variables son no independientes.

Estadístico de Prueba:

$$X^2 = \sum_{i=1}^{c}\sum_{j=1}^{k}\left[\frac{(O_{ij} - E_{ij})^2}{E_{ij}}\right] = \frac{(186-153.68)^2}{153.68} + \frac{(227-193.43)^2}{193.43} + \ldots + \frac{(259-241.22)^2}{241.22} = 47.90$$

con $(c-1)(k-1) = (5-1)(3-1) = 8$ grados de libertad y el valor de la $\chi^2_{8,.995} = 21.955$ valor obtenido en tablas (correspondiente a una $\alpha = 0.005$).

Regla de decisión:

Como $X^2 = 47.90$ y $\chi^2_8 = 21.955$, se cumple que $X^2 > \chi^2_{k-1}$ podemos concluir que se rechaza la Hipótesis nula (Ho), es decir los ingresos y el tiempo transcurrido desde la última consulta con su psicólogo no son independientes.

2.- Tabla de contingencia de 2 x 2.

a.- Prueba de la Ji-Cuadrada para probar igualdad de proporciones (homogeneidad en las poblaciones)

Richardson reporta la presencia y ausencia del síndrome de respiración angustiosa (RDS[4]) en dos grupos de infantes. El grupo uno considera 42 infantes los cuales tuvieron ruptura de membrana fetal en menos de 24 horas antes del parto, el grupo 2 fue compuesto por 22 infantes los cuales tuvieron ruptura de membrana por más de 24 horas antes del parto, ver la *tabla 10.1-5*. Usaremos estos resultados para probar la hipótesis nula de que las dos poblaciones sean homogéneas dada una $\alpha = 0.05$.

Tabla 10.1-5

Grupo	RDS		
	Si	**No**	**Total**
1	27	15	42
2	7	15	22
Total	34	30	64

Hipótesis:

Ho: Las dos poblaciones representadas por los dos grupos en el estudio son homogéneos con respecto a la presencia de RDS .
Ha: Las dos poblaciones son no homogéneas.

Se realizará una comparación entre los dos estadísticos de prueba.

Utilizando el estadístico sin factor de corrección tenemos:

$$X^2 = \frac{N(ad-bc)^2}{(a+c)(b+d)(c+d)(a+b)} = \frac{64[27(15)-15(7)]^2}{(34)(30)(22)(42)} = 6.111536$$

Utilizando el factor de corrección de Yates tenemos:

$$X^2 = \frac{N(|ad-bc| - 0.5N)^2}{(a+c)(b+d)(c+d)(a+b)} = \frac{64(|27(15)-15(7)| - 0.5(64))^2}{(34)(30)(22)(42)} = 4.87727$$

Como se puede observa si hay una diferencia entre estos dos valores. El valor de la $\chi^2 = 3.841$ valor obtenido en tablas (correspondiente a una $\alpha = 0.05$).

Regla de decisión:

Como $X^2 = 6.112$ y $\chi^2_g = 3.841$, se cumple que $X^2 > \chi^2$, podemos concluir que se rechaza la Hipótesis nula (Ho) es decir las dos poblaciones son no homogéneas.

10.5 UTILIZANDO EL PAQUETE SPSS

Comúnmente para este tipo de pruebas la información se encuentra en bases de datos con todos los resultados de la encuesta o muestra es decir:

[4] siglas en ingles " respiratory distress syndrome"

Digamos que Sí ésta codificada con (1), No con (0) y que el grupo es identificado por el número que tiene (1 o 2), entonces en la primera columna de SPSS se tendrían (42) unos señalando el grupo 1 y (22) doces señalando el grupo 2.

En la segunda columna de SPSS se tendrían (27) "1" señalando Si y (15) "0" señalando No para sumar 42 que son el total de los elementos del grupo 1, lo mismo se tendría para el grupo 2. La base de datos tendría que estar como lo muestra la (*Figura 10.1-1*).

Con los datos de esta forma, las indicaciones que se tiene que seguir son las siguientes:

1.- En la barra de herramientas se debe seleccionar *Statistics*, dentro de la cual se elige la opción de *Summarize.*

2.-Dentro de *Summarize* se elige la opción de *Crosstabs*, la cual presentará una ventana de diálogo en donde el usuario debe identificar qué variable contiene los valores de los renglones *(Row-s)* y qué variable contiene los datos de las columnas *(Column-s)* ver *figura 10.1-2.*

Figura 10.1-1

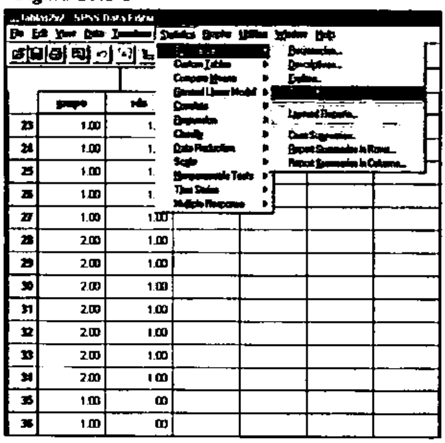
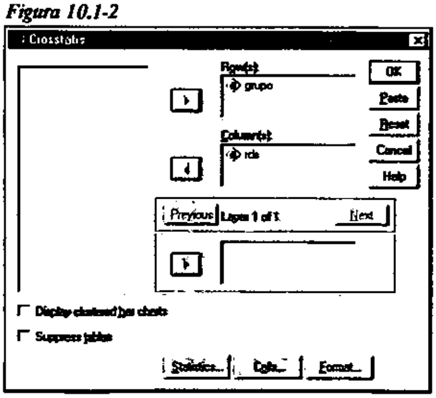


Figura 10.1-2



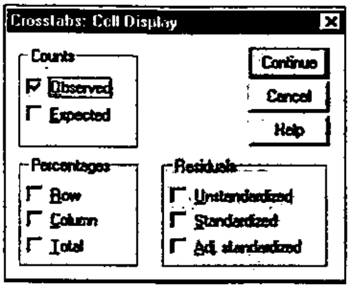
3.- Si a uno le interesara graficar los datos por medio de una grafica de barras, ésto se podría realizar con solo marcar la opción de *"Display clustered bar Charts"*

4.- Como esta opción es de tablas cruzadas el cuadro de *"Layer 1 of 1"* sirve para realizar tablas más especificas cuando se cuenta con más variables.

5.- En el icono de *Statistic* se encuentra la prueba de la *Chi-square* para tablas de contingencia, al igual que coeficientes de asociación, correlación[5] para datos nominales o cardinales.

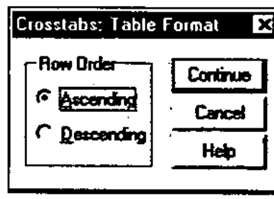
[5] En el presente trabajo se vera el coeficiente de correlación de Spearman

Figura 10.1-3



6.-El icono de *Cells* nos da la opción de proporcionarnos los valores esperados (*Expected*) o los observados *(Observed)* para que aparezcan en la tabla de resultados, al igual que la opción de presentarnos los porcentajes por renglón *(Row)*, columna *(Column)* o total. La opción de *Residual* en la ventana de *Cells* nos proporciona los resultados de los residuales en diferentes formas que en nuestro caso no son muy útiles ver *Figura 10.1-3.*

Figura 10.1-4



7.- Si se elige el icono de (*Format)* éste sirve para darle formato a la tabla, es decir uno puede elegir en qué orden quiere que aparezcan los renglones, en forma ascendente o descendente ver *Figura 10.1-4.*

8.-Una vez elegidos todos los cambios o procesos que se quieren realizar con los datos, se oprime el botón *OK ó paste* ubicado en la parte superior derecha para continuar con el proceso.

9.- Los resultados de la prueba aparecen en otra ventana llamada *Output1* en donde el usuario estará en posibilidades de imprimir o guardar como archivo.

Tablas de contingencia de 2 x 2

GRUPO * RDS Crosstabulation

Count

		RDS		
		.00	1.00	Total
GRUPO	1.00	15	27	42
	2.00	15	7	22
Total		30	34	64

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	6.112[b]	1	.013		
Continuity Correction[a]	4.877	1	.027		
Likelihood Ratio	6.204	1	.013		
Fisher's Exact Test				.018	.013
Linear-by-Linear Association	6.016	1	.014		
N of Valid Cases	64				

a. Computed only for a 2x2 table

b. 0 cells (0%) have expected count less than 5. The minimum expected count is 10.31

El primer ejemplo que se realizará con el paquete es el de las tablas de Contingencia de 2 x 2, en donde tenemos que la variable que identifica a los renglones es "grupo" y la variable "RSD" a las columnas. El cuadro de resultados nos mostrará dos cuadros.

El primer cuadro que nos presenta es la misma tabla de contingencia de 2x2 presentada al inicio con las frecuencias observadas para cada uno de los casos.

El segundo cuadro de resultados nos reporta el valor del Chi-Square X^2 (6.112), sus grados de libertad (1), el valor del estadístico de X^2 con el factor de corrección (4.877), sus grados de libertad (1) , el valor de la razón de verosimilitud[6] G^2 (6.204), sus grados de libertad (1), para cada una de las estadísticas de prueba antes mencionadas el paquete también nos proporciona el valor del P-value, además de que nos da también el valor de la prueba exacta de Fisher (0.018 - para una prueba de dos colas y 0.013 - para una prueba de una cola) y por último el número total de casos (64).

Como tenemos que la probabilidad asociada a la χ_1^2 es menor que el nivel de significancia seleccionado (0.013<0.05), se puede concluir que se rechaza la Hipótesis nula (Ho), es decir las dos poblaciones son no homogéneas.

Tablas de c x k

Si se nos llegara a presentar el caso en que los datos que queremos analizar por medio de una tabla de contingencia se encuentran en una tabla resumida, SPSS tiene la opción de ingresar los datos en otra forma para poder realizar el análisis. Los pasos que se tiene que seguir son los siguientes:

Se ingresaran los datos en 3 columnas: la primera columna indicará el número de renglones de la tabla que se pretende ingresar; la segunda columna indicará el número de columnas y la última columna será la frecuencia que se tienen en cada una de las casillas.

Figura 10.1-5

tablacxk - SPSS for Windows Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window

	ingresos	consulta	frec	var
1	1.00	1.00	186.00	
2	2.00	1.00	227.00	
3	3.00	1.00	219.00	
4	4.00	1.00	365.00	
5	5.00	1.00	653.00	
6	1.00	2.00	38.00	
7	2.00	2.00	54.00	
8	3.00	2.00	78.00	
9	4.00	2.00	112.00	
10	5.00	2.00	285.00	
11	1.00	3.00	35.00	
12	2.00	3.00	45.00	
13	3.00	3.00	78.00	
14	4.00	3.00	140.00	
15	5.00	3.00	259.00	

Para ejemplificar este hecho fijémonos en los datos de la *tabla 10.1-3*, en donde la primera columna se denominará "ingresos", la cual contará con 5 clasificaciones (renglones) – estos 5 renglones se tienen que repetir uno seguido del otro por cada una de las columnas con las que cuenta la tabla en la misma columna, en éste caso se repetirán 3 veces los valores - la segunda columna se le denominará "consulta " la cual contará con 3 atributos (columnas), en donde el valor que indica el número de columna se tendrá que repetir tantas veces como renglones tuviera la tabla, es decir en él ejemplo el valor 1 (1era. Columna) se tendría que repetir 5 veces, el valor 2 (2da. Columna) se tendría que repetir 5 veces, así sucesivamente hasta finalizar el número de columnas con la que cuente la tablas; y por último a la tercera columna se le denominará

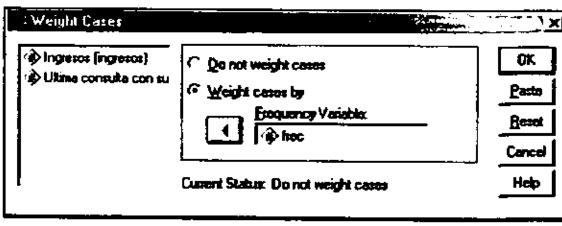
[6] La prueba G^2, razón de verosimilitud, es más fácil de calcular que el de la X^2.

"Frec", en donde los valores que se tienen que ingresar son los que se encuentran en cada una de las casillas, cuidando que este valor corresponda al renglón y la columna indicada, ejemplo el quinto renglón (mas de $10,000) y primer columna (6 meses antes) tiene como valor en la casilla "653" ver *figura 10.1-5.*

Los valores ingresados en SPSS se encuentran recodificados es decir, en la columna "ingresos" el valor 1 corresponde a "Menor de $3000" , el valor 2 a "$3000 -$4999, así sucesivamente; la columna "consulta" el valor 1 corresponde a "6 meses antes", el valor 2 a "De 7 meses a un año" y por último el valor 3 a "Mas de un año" (ver datos originales en *cuadro 10.1-3*). Una vez que se tienen los datos de esta forma las indicaciones que se tienen que seguir son las siguientes:

1.- En la barra de herramientas se debe seleccionar la opción de *Data*, dentro de la cual se elige la opción de *weight cases.*

Figura 10.1-6



2.-Dentro de *weight cases* se desplegará otra ventana en donde se indicará que se trata de una tabla con datos agrupados o ponderados. En esta ventana uno tiene que elegir la opción de *Weight cases by,* en donde se activará un recuadro denominado "*Frequency Variable*"en el cual se debe ingresar el nombre de la variable que contiene la frecuencia de los datos agrupados para el ejemplo es la denominada *"frec",* para poder continuar con el proceso se oprime *OK* ver *Figura 10.1-6.* Es conveniente recalcar que este proceso se debe de realizar **siempre** que se tengan datos en esta forma, si no el análisis que nos proporcione el paquete será incorrecto.

3.- Una vez realizado el paso anterior deben seguirse los pasos antes mencionados, es decir dentro de *Summarize* la opción de *Crosstabs,* la cual presentará una ventana de diálogo en donde el usuario debe identificar qué variable contiene los valores de los renglones *(Row-s)* y qué variable que contiene los valores de las columnas *(Column-s)* ver *figura 10.1-2,* esta opción nos sirve para verificar si los datos que se ingresaron son los correctos, ya que uno de los cuadros que nos ofrece la ventana de resultados es la misma tabla que se ingresó.

Crosstabs

Ingresos * Ultima consulta con su psicologo Crosstabulation

Count

		Ultima consulta con su psicologo			
		6 meses antes	De 7 meses a un año	Mas de un año	Total
Ingresos	Menor de $3000	166	38	35	259
	$3000 - $4999	227	54	45	326
	$5000 - $6999	219	78	78	375
	$7000 - $9999	355	112	140	607
	Mas de $10000	653	285	259	1197
Total		1640	567	557	2764

4.- Como la finalidad es obtener la prueba de la Chi-square, uno tiene que elegir dentro de la opción de Crosstab el icono de *Statistic* y dentro de esta la prueba de la *Chi-square* para tablas de contingencia, el segundo cuadro de la ventana de salida nos presenta los resultados de la prueba como se presenta enseguida.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	47.892[a]	8	.000
Likelihood Ratio	49.087	8	.000
Linear-by-Linear Association	32.681	1	.000
N of Valid Cases	2764		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 52.19.

Este cuadro de resultados nos reporta, el valor del Chi-Square X^2 (47.892), sus grados de libertad (8), el valor del P-value (0.000), el valor de la razón de verosimilitud[7] G^2 (49.087), sus grados de libertad (8), el valor de su probabilidad asociada P (0.000), y por último el total de número de casos (2764).

Como tenemos que la probabilidad asociada a la χ^2_l es menor que el nivel de significancia seleccionado (0.000<0.05), se puede concluir que se rechaza la Hipótesis nulas (Ho) , es decir los ingresos y el tiempo transcurrido desde la última consulta son dependientes.

[7] La prueba G^2, razón de verosimilitud es más fácil de calcular que el de la X^2.

CAPITULO XI

11 PRUEBAS DE ASOCIACION

11.1 CONCEPTO DE CORRELACIÓN

Con frecuencia se nos pide determinar las relaciones entre dos o más variables, tan pronto como entablamos esta búsqueda, nos adentramos al área de la correlación. Puesto que la correlación sirve para determinar el grado en que dos variables están relacionadas o tienden a variar conjuntamente. Para expresar cuantitativamente el grado en que dos variables están relacionadas, es necesario calcular un *coeficiente de correlación.* Existen muchos tipos de coeficientes de correlación. La decisión del cuál se ha de emplear con un conjunto específico de datos depende de factores tales como:

1. El tipo de la escala de medida en que cada variable está expresada
2. La naturaleza de la distribución (continua o discreta) y
3. La característica de la distribución (lineal o no lineal).

La tabla siguiente muestra algunos de los coeficientes de correlación cuyo empleo resulta adecuado con varios tipos de escalas:

ESCALA	SÍMBOLO	SE USA CON
Nominal	r_{phi} (coeficiente phi)	Dos variables
Ordinal	r_s (r de Spearman) T (Tau de Kendall, o coeficiente de correlación por rangos)	Datos ordenados según su rango. Si una variable es propiamente ordinal y la segunda es de intervalo ó razón, se las debe expresar a las dos según su rango antes de calcular la r de Spearman Datos ordenados según su rango.
Intervalar /razón	r de Pearson	Escalas de intervalo y/o razones

Sea cual sea la técnica de correlación que se use, lo fundamental es que todas tienen ciertas características comunes.

a. Se obtienen dos conjuntos de medidas con los mismo individuos (o sucesos) o en parejas de individuos que tengan alguna forma de relación.
b. Los valores de los coeficientes de correlación varían entre +1.00 y −1.00. Ambos extremos representan relaciones perfectas entre las variables y 0.00 representa la ausencia de relación.
c. Una relación positiva significa que existe una relación directa entre las dos variables.
d. Una relación negativa significa que existe una relación inversa perfecta.

El coeficiente de r de Pearson no es material de este trabajo, por tal motivo nos enfocaremos a coeficientes adecuados para el tipo de escala que se han estado empleando en todas las pruebas no paramétricas (datos Nominales y Ordinales).

11.2 COEFICIENTE DE CORRELACION DE SPEARMAN

11.2.1 Función

Sin lugar a duda la estadística más conocida para medir el grado de asociación lineal entre dos variables (X,Y) es el coeficiente de correlación del producto-momento de Pearson , que se obtiene mediante el coeficiente de covarianza muestral de las variables de interés entre el producto de sus respectivas desviaciones estándares, de modo que el indicador no depende de las escalas en las que cada una de las variables x , y se hayan medido, ni a la distribución subyacente, sin embargo este coeficiente de correlación no se puede usar en la estadística no paramétrica, ya que la distribución de r depende de la distribución bivariada de (X,Y). Sin embargo para la estadística no paramétrica existe un coeficiente que es el equivalente exacto de la r de Pearson, denominado "El coeficiente de rangos de Spearman".

De todas las estadísticas basada en rangos, el coeficiente de correlación de rangos de Spearman fue la primera que se desarrolló (1904). Esta estadística a veces llamada ρ (rho), es una medida de asociación que requiere que ambas variables sean medidas por lo menos en una escala ordinal, de manera que los objetos o individuos en estudio puedan colocarse en dos series ordenadas.

11.2.2 Potencia-Eficiencia

La eficiencia asintótica relativa de la prueba basada en la correlación de rangos de Spearman r_s cuando se compara con la correlación paramétrica más poderosa, la r de Pearson es de cerca de $9/\pi = 0.912$ (Holtelling y Pabst, en 1936)

11.2.3 Supuestos

a.- Los datos provienen de una muestra aleatoria de n pares de observaciones.
b.- Cada par de observación representa dos medidas tomadas del mismo sujeto o individuo (X,Y), nombrada unidad de asociación.

11.2.4 Método

1. A los grupos de observaciones (X,Y) se le sustituye por sus respectivos rangos $R(X)$, $R(Y)$.
2. Si se nos presentara el caso que R(X)=R(Y), este nos indicaría que existe una correlación perfecta entre las dos variables, por lo que parece lógico usar la diferencia de los rangos. $d_i = R(X_i) - R(Y_i)$
3. Si se toman solamente tales diferencias para medir la correlación, las diferencias positivas cancelarían a las negativas, cuando se tratara de determinar la magnitud total de la discrepancia. Sin embargo, si se emplea d_i^2 en lugar de la d_i la dificultad desaparece. Es evidente que cuando mayores sean las d_i en valor absoluto, mayor será $\sum d_i^2$.
4. Ocasionalmente se puede presentar que dos o más sujetos reciban el mismo puntaje en la misma variable, estos casos reciben el nombre de "ligas", cuando esto ocurre se les deberá asignar el mismo rango; en donde el valor de éste será el promedio de los rangos que se le habría asignado si no hubieran ocurrido ligas, procedimiento usual para asignar rangos a observaciones ligadas.
5. También se puede dar el caso en que existan varias ligas (diferentes de cero) y debido a que estas puedan dar lugar a ciertas perturbaciones, se recomienda introducir un factor de corrección, claro que si la proporción de ligas no es grande se puede usar la fórmula normal del coeficiente de correlación de Spearman; pero si es grande, entonces hay que incorporar el siguiente factor de corrección:

$$T = \frac{t^3 - t}{12}$$

donde :
$T = t^3 - t$ (donde t es el número de observaciones ligados en un rango dado, ya sea para x, y o ambos)

11.2.5 Estadístico de Prueba

1.- Coeficiente de Spearman sin ligas

El coeficiente de correlación de Spearman cuando éste no contiene ligas, se puede obtener utilizando los rangos en la fórmula del coeficiente de correlación de Pearson es decir:

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}} = \frac{\sum_{i=1}^{n}\left[R(X_i) - \frac{n+1}{2}\right]\left[R(Y_i) - \frac{n+1}{2}\right]}{n(n^2-1)/12} = r_s$$

Sin embargo una forma equivalente que además simplifica mucho los cálculos es:

$$r_s = 1 - \frac{6\sum_{i=1}^{n}[R(X_i)-R(Y_i)]^2}{n(n^2-1)} = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3-n}$$

2.- Coeficiente de Spearman con ligas.

La expresión que se tiene que utilizar en estos casos es la siguiente:

$$r_s = \frac{\sum x^2 + \sum y^2 - \sum d^2}{2\sqrt{\sum x^2 \sum y^2}}$$

donde:
$\sum x^2 = \frac{n^3-n}{12} - \sum T_x$
$\sum x^2 = \frac{n^3-n}{12} - \sum T_x$

Propiedades del coeficiente r_s

1. El valor de r_s está entre -1 y 1.
2. Si observamos que $r_s = 1$ podemos decir que existe una relación directa perfecta entre las dos variables.
3. Si observamos que $r_s = -1$ podemos decir que existe una relación inversa perfecta entre las dos variables.
4. Si observamos que $r_s = 0$ podemos decir que no existe una relación alguna.

El coeficiente de correlación de Spearman es usado a menudo como una estadística de prueba para probar independencia entre dos variables aleatorias. Debido a que el coeficiente de correlación de Spearman puede ser insensible a ciertos tipos de dependencia, es recomendable especificar en la hipótesis alternativa el tipo de dependencia que se espera detectar.

11.2.6 Hipótesis

Los enunciados de la hipótesis pueden plantearse en la siguiente forma:

Hipótesis nula:

Ho: las variables (X, Y) son mutuamente independientes $(r_s = 0)$

Hipótesis alternativa:

a) Existe una tendencia a que los valores grandes de X se presenten asociados con valores grandes de Y, correlación positiva $(r_s > 0)$, es decir la relación entre X y Y es directa.

b) Existe una tendencia a que los valores grandes de X se presenten asociados con valores pequeños de Y y viceversa correlación negativa $(r_s < 0)$, es decir la relación entre X y Y es inversa.

c) las variables X, Y no tienen una relación directa o inversa $(r_s \neq 0)$

Los incisos a y b considerados en la hipótesis alternativa dan lugar a una prueba de una sola cola y el inciso c a una prueba de dos colas.

11.2.7 Regla de decisión

Se ocupa esta regla tanto para la r_s corregida por el efecto de las ligas, como para la r_s normal.

Cuando se dispone de muestras pequeñas $n \leq 30$, la significancia de la prueba se puede obtener mediante la consulta de las tablas correspondientes y su regla de decisión es:

Se rechaza Ho:

Si $|r_s| \geq r_0$ donde r_0 es el valor crítico que aparece en tablas, correspondiente a $\alpha/2$ para una prueba de dos colas ó α para una cola.

No se rechaza Ha:

Si $|r_s| < r_0$ donde r_0 es el valor crítico que aparece en tablas, correspondiente a $\alpha/2$ para una prueba de dos colas ó α para una cola.

Cuando $n > 30$ la significancia se puede probar mediante la siguiente aproximación a la distribución t de Student propuesta por Kendall.

$$t = r_s \sqrt{\frac{N-2}{1-r_s^2}}$$

La estadística t se distribuye aproximadamente como una t de Student con n-2 grados de libertad. Pero si la muestra tiene más de 100 parejas se puede utilizar la siguiente estadística de prueba:

$$z = r_s \sqrt{n-1}$$

que es una aproximación a la distribución normal (0,1).

11.2.8 Ejemplo

Supóngase que una persona interesada en adquirir un televisor de color de 24 pulgadas le asigna rangos a los modelos estándar de cada una de las ocho fábricas más importantes. Sus preferencias aparecen en la *tabla 11.1-1* junto con los precios al menudeo sugeridos por fabricante; un fabricante de televisores sabe que la asociación entre preferencia y precio es un aspecto importante para la asignación de precios, éste desea conocer el grado de asociación.

Tabla 11.1-1

Fabricante	Preferencia	Precio
1	7	$449.50
2	4	$525.00
3	2	$479.95
4	6	$499.95
5	1	$580.00
6	3	$549.95
7	8	$469.95
8	5	$532.50

Respuesta:

Como las dos variables de interés son los rangos de preferencia y los precios. La primera está construida por rangos y a la segunda se le puede asignar rangos, el coeficiente de correlación de rangos de Spearman r_s es la herramienta adecuada para conocer si existe alguna relación entre las dos variables.

Siguiendo las indicaciones, lo primero que se necesita es obtener la suma de los rangos de cada una de las variables como se muestra en la *tabla 11.1-2*. Los datos no presentan ninguna liga en sus valores, lo que nos indica que se puede usar el estadístico sin tener que ocupar ningún factor de corrección.

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} = 1 - \frac{6(144)}{8(64-1)} = -0.714$$

Tabla 11.1-2

Fabricante	Rangos de Preferencia X_i	Rangos de Precio Y_i	Rango d_i	d_i^2
1	7	1	6	36
2	4	5	-1	1
3	2	3	-1	1
4	6	4	2	4
5	1	8	-7	49
6	3	7	-4	16
7	8	2	6	36
8	5	6	-1	1
				$\sum d_i^2 = 144$

Como tenemos que el coeficiente de correlación de Spearman nos da un valor de -0.714, podemos decir que sí existe una relación (inversa).

Supongamos ahora que la administración se ha dado cuenta de que la significancia de una correlación depende de la asociación real y de los tamaños de muestra que se usen para estimar el coeficiente de correlación. Ahora la administración desea conocer la significancia de esta correlación (negativa) con respecto a la hipótesis nula de no asociación utilizando $\alpha = 0.05$.

Solución:

Para este problema se requiere utilizar el coeficiente de correlación de Spearman como estadística de prueba para probar una hipótesis de no asociación entre las dos poblaciones, es decir:

Se desea probar :

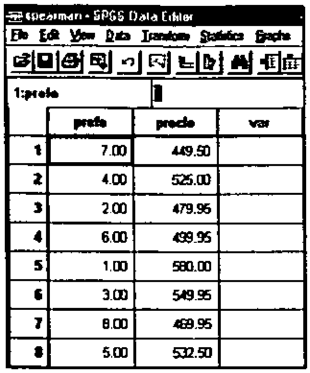
> Ho: Las variables X,Y son mutuamente independientes $(r_s = 0)$, es decir no existe asociación entre los rangos.
> Ha: La relación entre X y Y es inversa.

Como el valor crítico de r_s para una prueba de una cola con $\alpha = 0.05$ y n=8, es de 0.643 y el valor de la estadística de prueba $r_s = -7.14$, es menor que el valor critico, podemos rechazar la hipótesis nula, es decir parece que existe alguna relación entre preferencia y precio.

11.2.9 Utilizando el paquete SPSS

Los pasos que se tienen que seguir para la realización correcta de esta prueba son:

Figura 11.1-1

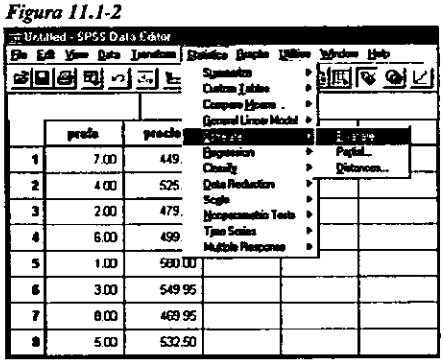


La estructura de los datos se basará en 2 columnas o variables con el mismo número de observaciones. *Figura 11.1-1*

Para obtener el coeficiente de correlación de Spearman, el usuario tiene que seguir los siguientes pasos:

1.- En la barra de herramientas se debe seleccionar *Statistics*, dentro de la cual se elige la opción de *Correlate* ver *Figura 11.1-2.*

Figura 11.1-2



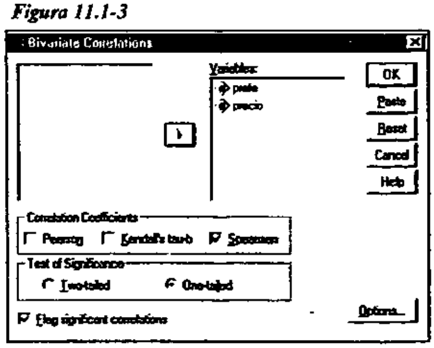
2.-Dentro de *Correlate* se elige la opción *bivariate* que conduce al usuario a una ventana de diálogo lugar en donde se debe seleccionar únicamente las variables *(Variables)* que se desean correlacionar.

3.- En el recuadro de *Correlation Coefficients* se debe elegir la opción de Spearman.

4.-En la parte inferior de *Correlatión Coefficients* se encuentra la opción de *Test of Significance,* este procedimiento nos ofrece la opción de elegir el tipo de prueba que uno desee realizar "una prueba de una cola o de dos", finalmente se oprime el botón *OK,* para poder terminar con el proceso.

5.- Los resultados de la prueba aparecen en otra ventana llamada *Output1*,en donde el usuario estará en posibilidades de imprimir o guardar como archivo, esta ventana nos presentara dos cuadros.

Figura 11.1-3



Para ejemplificar la prueba se realizara el ejercicio por medio del paquete. Las variables que se desean correlacionar (*variables*), en este caso son "prefe" y "precio" como lo muestra la *figura 11.1-3.*

El cuadro de resultados reporta lo siguiente:

El cuadro reporta una matriz simétrica, el coeficiente de correlación de Spearman (-0.714), el valor de su probabilidad asociada P(0.023) y por último el número de parejas de datos N (8).

Nonparametric Correlations

Correlations

			PREFE	PRECIO
Spearman's rho	PREFE	Correlation Coefficient	1.000	-.714*
		Sig. (1-tailed)	.	.023
		N	8	8
	PRECIO	Correlation Coefficient	-.714*	1.000
		Sig. (1-tailed)	.023	.
		N	8	8

* Correlation is significant at the .05 level (1-tailed).

Como la matriz nos reporta que 0.023<0.05 (nivel de significancia asociado para la prueba de una cola), se puede concluir que las dos variables están correlacionadas, es decir existe alguna relación entre preferencia y precio.

CAPITULO XII

ELECCION DE LA PRUEBA ESTADISTICA ADECUADA

Tal vez una de las partes más difíciles en la estadística no paramétrica es la de realizar la elección de una prueba estadística adecuada, ya que uno llega a preguntarse:

- "Es correcto que utilice esta prueba con el tipo de datos obtenidos en mi investigación"
- "Puedo utilizar esta prueba con la cantidad de datos que tengo"
- "Puede utilizar la prueba a pesar que no conozco cual es la naturaleza de la población de la cual fue extraída la muestra"
- "Puedo encontrar una prueba que me asegure o que me respalde mejor los resultados que obtengo con la prueba que realice " etc.

Para contestarse estas pregunta y tal vez otras, debemos considerar una serie de puntos que se tienen que tomar en cuenta para que la elección de la prueba estadística sea la más adecuada y cumpla con las expectativas de la investigación.

- La manera en que se obtuvo y la naturaleza de la población de la cual fue extraída la muestra de datos.
- El tipo de Escala que se tienen en los datos de la investigación
- Potencia
- Eficacia

Cada uno de estos puntos tiene una importancia tal que para comprenderla se profundizará y se explicará su utilidad.

12.1 LA MANERA EN QUE SE OBTUVO Y LA NATURALEZA DE LA POBLACIÓN DE LA CUAL FUE EXTRAÍDA LA MUESTRA DE DATOS.

Debido a que hay investigaciones que ocupa diferentes técnicas para obtener la población de estudio uno tiene que fijarse si ésta fue obtenida en:

forma aleatoria:

Este punto es muy importante ya que cuando se lleva a cabo una investigación, normalmente se desea generalizar los resultados mas allá de las condiciones particulares de la investigación y para que se pueda realizar esta generalización es necesario que la muestra de la población de interés se haya obtenido en forma aleatoria. Hay diferentes técnicas para obtener una muestra, entre las que se encuentra el Muestreo Aleatorio Simple, el Muestreo Estratificado, etc.

Grupos disponibles:

Hay veces en que el estudio se tiene que realizan con algún grupo en particular de la población, es decir que no se obtuvo en forma aleatoria. Como en los estudios con mellizos.

Es importante que se identifique como se obtuvo la muestra para que uno pueda realizar la prueba correcta, ya que cada prueba tiene determinados supuestos y si en esos supuesto no se cumple se pierde potencia.

12.2 EL TIPO DE ESCALA QUE SE TIENEN EN LOS DATOS DE LA INVESTIGACIÓN

Uno de los problemas más dificiles es decidir cuál de las pruebas estadísticas es la más adecuada para un conjunto dado de datos. No puede darse ninguna contestación directa a esta pregunta sin un conocimiento detallado de lo que se está investigando, claro que aun con este conocimiento detallado a menudo la dificultad persiste.

Como cualquier prueba estadística implica un modelo y un requisito de medida; la prueba es válida en ciertas condiciones que especifican el modelo y el requisito de medida. El pedir que todos los puntajes se obtengan independientemente entre sí fundamenta todas las pruebas estadísticas sean o no paramétricas.

Al utilizar las propiedades de las escalas de medición antes vistas, uno puede identificar el tipo de datos que sé está manejando en la investigación y a partir de ésta, se define el tipo de pruebas estadísticas que se pueden realizar con los datos.

En la actualidad se cuenta con una gran variedad de pruebas estadísticas alternativas válidas que se pueden utilizar para decidir acerca de una hipótesis, claro que una prueba estadística es buena si es pequeña la probabilidad de rechazar Ho siendo verdadera y grande la probabilidad de rechazar Ho siendo falsa.

Tomando en cuenta lo anterior se tiene:

Escala	Ejemplos de estadísticas apropiadas	Pruebas Estadísticas apropiadas
Nominal	Moda Frecuencia Coeficiente de contingencia	Pruebas estadísticas no paramétricas
Ordinal	Mediana Percentiles Spearman Kendall r Kendal W	
Intervalo	Media Desviación estándar Correlación del momento-producto de Pearson. Correlación del múltiple momento-producto de Pearson.	Pruebas estadísticas paramétricas
Razón	Media geométrica Media armónica Coeficiente de variación	

12.3 POTENCIA

Al revisar el concepto de significación de una prueba discutimos los errores de tipo I y II y comentamos que es preferible tratar de minimizar estos dos tipos de errores – algunos autores piensan que es preferible cometer un error de tipo I, es decir cometer el error a dejar de proclamar un resultado, que cometerlo al recabar un resultado equivocado.

Hasta este punto la preocupación ha sido la de establecer un nivel de significancia que reduzca la probabilidad de rechazar falsamente la hipótesis nula. En otras palabras, se ha tratado de evitar más el error del tipo I que el error de tipo II. Sin embargo, lo ideal sería tener una prueba estadística que consiguiera cierto equilibrio entre ambos tipos de error y que nosotros pudiéramos especificar antes del estudio la probabilidad de cometer los dos tipos de error.

Cuando se trata de efectuar un equilibrio entre los errores de tipo I y II, nos encontramos con el concepto de potencia de una prueba, el cual se define simplemente cómo "rechazar la hipótesis nula cuando en efecto es falsa". Simbólicamente, la potencia se define como:

Potencia = 1-probabilidad de un error de tipo II=1- β

La potencia varia en función de los siguientes factores:

- Tamaño de la muestra: generalmente la potencia de una prueba estadística se incrementa al incrementarse el tamaño de la muestra, por lo que podemos decir que la potencia varía en función de N.

- Nivel de significación: cuanto mayor sea el nivel de significación α elegido, mayor será la potencia de la prueba.

- La naturaleza de Ha: La potencia de una prueba es también función de la naturaleza de la hipótesis alternativa. En el caso de que Ho sea realmente falsa, la Ha direccional o de una cola es más potente que la prueba bilateral, mientras el parámetro esté en la dirección predicha, si este no es el caso, la prueba unilateral será menos potente.

Ejemplo:

La tabla muestra que para cualquier nivel α dado, el valor critico Z es menor para una prueba unilateral que para uno bilateral, por tanto, una Z obtenida que no es significativa par un examen bilateral puede serlo para uno unilateral.

Nivel de significancia	Naturaleza de Ha.	
	Direccional (prueba unilateral)	No direccional (prueba bilateral)
$\alpha = 0.005$	Z = 2.58	Z = ±2.81
$\alpha = 0.01$	Z = 2.33	Z = ±2.58
$\alpha = 0.025$	Z = 1.96	Z = ±2.24
$\alpha = 0.05$	Z = 1.64	Z = ±1.96

- La naturaleza de la prueba estadística: otro factor que determina la potencia de una prueba estadística es la naturaleza de dicha prueba.

Como regla general para cualquier N dado, las pruebas paramétricas son más potentes que sus equivalentes no paramétricas, esto se debe a que como se supone que las poblaciones se distribuyen normalmente con la misma varianza, conllevan un riesgo menor de cometer error del tipo II.

Así qué, si hay que elegir entre una prueba no paramétrica y otra paramétrica se debe escoger la prueba paramétrica siempre que se cumplan las hipótesis básicas.

Sin embargo cuando la naturaleza de los datos excluye la posibilidad de usar una prueba paramétrica, nos vemos obligados a utilizar pruebas no paramétricas que son menos potentes[1].

- Uso de medidas correlacionadas: cuando se va a realizar una hipótesis con dos muestras que están correlacionadas, esta será más potente que otra que no la considere.

[1] Debe insistirse en que las pruebas paramétricas son más potentes solamente cuando son válidas las hipótesis que permiten su uso. Como estas hipótesis no se cumplen, una prueba no paramétrica puede ser tan potente como otro paramétrico.

12.4 EFICACIA

El concepto de eficiencia se refiere al incremento necesario en el tamaño de la muestra para hacer que una prueba sea tan potente como otra prueba en la que se emplea una N menor.

Potencia de la prueba $P(\pi) = (100)\frac{N_a}{N_b}$

CONCLUSIONES

Si la ciencia se asienta sobre hechos, leyes y teorías, el objetivo que perseguimos es un conocimiento científico de la realidad, es evidente que tenemos que trabajar con conjuntos de observaciones, ya que ningún fenómeno individual puede generalizarse en la ciencia.

Una de las formas de analizar un conjunto de observaciones es por medio del análisis estadístico (paramétrico y no paramétrico), que estudian fenómenos colectivos en los que generalmente se encuentra una consistencia y ésta es la que nos permite comprender y explicar los fenómenos, podemos predecir en base a las relaciones existentes entre ellos y controlar en algunos casos las condiciones relacionadas con su aparición o extinción, es decir la tarea básica del análisis estadístico es la descripción de las observaciones y la inferencia o inducción sobre hipótesis científicas derivadas de las observaciones.

Claro está que el análisis estadístico más explotado es el paramétrico, pero cuando uno no esta seguro que se cumplan los supuestos que esta requiere ó la escala de medición es muy pobre, una muy buena alternativa es por medio de los métodos no paramétricos, a pesar de que estos métodos son menos potentes proporcionan resultados confiables.

No obstante de que cada vez es más común que las personas tengan acceso a paquetes estadísticos que realicen todo tipo de análisis estadísticos es importante recalcar que hay que tener el conocimientos de que es lo que realiza el paquete, como se deben ingresar apropiadamente los datos para poder realizar la interpretación correcta de los resultados arrojados por el paquete.

En el presente trabajo se trato de explicar cual es el procedimiento manual para poder realizar una prueba no paramétrica, cuales son las posibles hipótesis que se pueden probar y como objetivo principal el explicar como se debe ingresar correctamente los datos en el paquete SPSS, tratando de explicar de donde proviene cada uno de los datos que nos reporta la ventana de resultados *Output*, claro esta que antes uno tiene que saber elegir de la gran diversidad de métodos no paramétricos que existen aquel que se adecuado para el tipo de medición lograda en los datos.

A continuación se presentan las sugerencias que se deben se seguir:

2 muestras relacionadas

En este capitulo se presentaron pruebas estadísticas que sirven para examinar la "significación de la diferencia" entre dos muestras relacionadas. En este caso tenemos que para elegir la mejor prueba se tiene que considerar lo siguiente:

- Cuando se tienen las más burdas medidas ordinales – es decir cuando el puntaje de un miembro de un par puede colocarse en relación de "mayor que" con el puntaje de otro-, debe usarse la prueba de los signos.

- Si se cuanta con medidas más refinadas de la escala ordinal - es decir, cuando se puede ordenar la significación entre las diferencias observadas para los diferentes pares igualados, la prueba de Wilcoxon es preferible a la de los signos.

De los dos métodos aquí vistos el de Wilcoxon es el más potente.

2 muestras independientes

En este capitulo se presentaron pruebas estadísticas que sirven para examinar la "significación de la diferencia" entre dos muestras independientes, es decir se pretende examinar la probabilidad de que las muestras procedan de la misma población. Pero las diferentes pruebas son más sensibles o menos a distintas clases de diferencias entre las muestras. Por ejemplo, cuando se desea saber si dos muestras representan poblaciones que difieren en ubicación – tendencia central – hay pruebas más sensibles a tal diferencia y por consiguiente deberán escogerse: la prueba de la mediana – o la prueba de Fisher cuando N es pequeña -, prueba de U de Mann-Whitney .

La elección entre las pruebas que son sensibles a diferencias en ubicación depende de la clase de medida lograda en la investigación y del tamaño de las muestras.

- Cuando se tienen muestras grandes o medida ordinal – la alternativa sugerida es la prueba de U de Mann-Whitney.
- Si la medición es solamente significativa para dicotomizar las observaciones arriba o debajo de la mediana combinada, entonces es aplicable la prueba de la mediana - o la prueba exacta de Fisher, según sea el caso.

La prueba de la mediana no es tan poderosa como la prueba U de Mann-Whitney en la protección contra diferencias en ubicación pero es más adecuada cuando los datos son observaciones que no pueden ordenarse completamente.

K muestras relacionadas

En este capitulo se presentó la prueba de Friedman, la cual es útil cuando la medida de la variable está al menos en una escala ordinal, y sirve para determina si las k muestras relacionadas proceden probablemente de la misma población con respecto a las medias de los rangos. En otras palabras, es una prueba total para saber si el tamaño de los puntajes depende de las condiciones en que fueron producidos.

K muestras independientes

En este capitulo se presentaron 2 pruebas para determinar si las k muestras independientes son recogidas de la misma población.

La generalización de la mediana y la prueba de Kruskal-Wallis pueden aplicarse a los mismos datos, ya que éstas satisfacen requisitos similares para los datos en estudio – la escala sea al menos ordinal – pero a pesar que cualquiera de los dos métodos se pueden realizar es preferible usar el de Kruskal-Wallis ya que éste es más eficiente, pues usa más información de las observaciones – convierte los puntajes en rangos, así la prueba preserva la magnitud de los puntajes.

Análisis de tablas de contingencia

El análisis de las tablas de contingencia se realiza por medio de la prueba de la χ^2 para k muestras independientes, el cual es útil cuando los datos están en frecuencias y la medida de las variables en estudio están en una escala nominal o en categorías discretas de una escala ordinal.

La prueba de la χ^2 es útil para probar si las proporciones o frecuencias de las diversas categorías son independientes u homogéneas de la condición (muestra) en la que fueron observadas, es decir se prueba si las k muestras proceden de la misma población o de poblaciones idénticas con respecto a la proporción de casos en las diferentes categorías.

Pruebas de asociación

En este capítulo se presento la forma de medir el grado de correlación entre las variables de una muestra.

A pesar de que hay varias técnicas no paramétricas para medir ese grado, el que se presenta en este capítulo es el coeficiente de correlación de Spearman, esta técnica se utiliza cuando las variables en estudio han sido medidas por lo menos en una escala ordinal y los datos provienen de una muestra de n pares de observaciones. Este coeficiente es fácil de calcular y tiene la ventaja adicional que estar relacionado linealmente con el coeficiente de concordancia de W Kendall.

Cuadro resumen de las pruebas estadísticas no paramétrica.

Prueba No Paramétrica	Escala de Medición	Aplicaciones de la prueba	Ventaja de la Prueba	Desventajas de la Prueba	Ejemplos Típicos
PRUEBAS NO PARAMÉTRICAS PARA DOS MUESTRAS RELACIONADAS					
Prueba de los Signos	Al menos ordinal	• Se utiliza para probar hipótesis que involucran a uno o dos grupos correlacionados o apareados, como en los experimentos en los cuales cada sujeto es su propio control o cuando se desea establecer que ambas condiciones son diferentes. • Se usa para probar si dos muestras relacionadas difieren en localización o si una sola muestra proviene de una población con una mediana específica. • Se utiliza también en situaciones en las qué solo se sabe si los resultados de las diferencias son positivos o negativos, pero sin tener sus magnitudes.	• El atractivo de la prueba es su sencillez y su rapidez para obtener resultados.	• Una desventaja de la prueba es que elimina por completo cualquier información cuantitativa inherente en los datos, es decir falla al no hacer uso de la información referente a las magnitudes de las diferencias al trata a todas positivas y negativas como si fueran iguales. • A pesar de que la prueba de los signos se puede aplicar en los mismos casos que la de Wilcoxon , se puede llegar a resultados diferentes.	• Cuando se desea establecer diferencia entre dos tratamientos o condiciones (drogas, propaganda, integración grupal, dietas).
Prueba de Wilcoxon	Al menos ordinal	• La prueba de rango con signo de Wilcoxon se utiliza para probar hipótesis referentes a uno o dos grupos correlacionados (o cuando se tienen datos de parejas apareadas), para probar si la diferencia entre los dos resultados de cada sujeto proviene de una población simétrica respecto a una mediana cero o a una población con una mediana diferente de cero. • Se utiliza en experimentos que plantean condiciones de "antes y después" para el mismo sujeto.	• La prueba del signo-rango de Wilcoxon para observaciones apareadas tiene en cuenta tanto la magnitud como la dirección implícita en las mediciones. • La prueba de Wilcoxon alcanza mayor potencia al utilizar la información inherente a la clasificación de las diferencias, ya que es sensible a las diferencias en localización.	• Al rechazar una hipótesis nula concluimos que hay una diferencia de localización entre las dos poblaciones, pero es necesario señalar que esta diferencia puede no deberse a esto, ya que es posible construir ejemplos de poblaciones que tengan la misma localización pero que difieran en otros aspectos que la prueba no pueda detectar.	• La prueba de Wilcoxon es la de mayor utilidad para el científico conductual. • Cuando se desea comparar dos métodos de enseñanza.
PRUEBAS NO PARAMÉTRICAS PARA DOS MUESTRAS INDEPENDIENTES					
Prueba de la Mediana	Al menos ordinal	• La prueba de la mediana es un procedimiento para probar si dos grupos independientes difieren en sus tendencias centrales, es decir se utiliza cuando se desea probar si la localización de las distribuciones que las originaron es la misma desde el punto de vista de la mediana	• La prueba de la mediana usa uno de los procedimientos más sencillos y usados de todos los métodos no paramétricos que se utilizan para probar si dos muestras independientes provienen de poblaciones con medianas iguales. • La prueba de la mediana es más adecuada cuando los datos son observaciones que no pueden ordenarse completamente y sirven solamente para dicotomizar las observaciones arriba o debajo de la mediana combinada.	• La prueba de la mediana no es tan poderosa como la prueba de U de Mann-Whitney en la protección contra diferencias en ubicación.	• Esta prueba se utiliza siempre que uno desee comparar dos poblaciones o muestras para probar si difieren en la mediana en el comportamientos social, psicológicos, etc.
Prueba de U Mann-Whitney	Al menos ordinal	• Se utiliza cuando se desea saber si dos muestras representan poblaciones que difieren en ubicación ó se pretende examinar la probabilidad de que las muestras procedan de la misma población. • Esta prueba es la que se emplea con mayor frecuencia como alternativa para la t de Student cuando las medidas no llegan a tener la calidad correspondiente a una escala de intervalo.	• La prueba de U de Mann-Whitney es una de las pruebas mas potentes, ya que utiliza la mayor parte de la información cuantitativa que poseen los datos.	• A pesar de que la prueba supone que los puntajes representan una distribución con una continuidad básica, pueden presentarse casos de ligas y este suele tener un efecto en el estadístico U cuando ocurren dos o mas en las observaciones de ambos grupos – claro que este efecto es insignificante hay que tener cuidado con el estadístico que se usa.	• Esta prueba se utiliza siempre que uno desee comparar dos poblaciones o muestras para probar si tienen la misma distribución (calificaciones, número de ventas)

Prueba No Paramétrica	Escala de Medición	Usos más comunes	Ventaja de la Prueba	Desventajas de la Prueba	Ejemplos Típicos
PRUEBAS NO PARAMÉTRICAS PARA K MUESTRAS RELACIONADAS					
Prueba de Friedman	Al menos ordinal	• Sirve para probar la hipótesis nula de que las distribuciones de probabilidad de los k tratamientos son idénticas, frente a la alternativa de que por lo menos dos de ellas difieren en ubicación , en otras palabras sirve para probar si las k muestras relacionadas proceden de la misma población con respecto a las medias de los rangos. • Esta prueba es análoga al de la prueba F cuando los puntajes no satisfacen los requisitos de medición.	• La ventaja real de este procedimiento es que puede utilizarse independientemente de la forma real de la distribución de las poblaciones de los tratamientos o muestras. • Es una prueba total que sirve para saber si el tamaño de los puntajes depende de las condiciones en que fueron producidos.	• Para que la prueba sea válida no debe haber efectos secundarios que pudiera sesgar los resultados. Esto se puede reducir este problema mediante la utilización de grupos apareados o aleatorizados en las muestras o tratamientos.	• Cuando se desean comparar varias muestras de igual tamaño para probar varios tipos de medicamentos, dietas, tratamientos, etc.
PRUEBAS NO PARAMÉTRICAS PARA K MUESTRAS INDEPENDIENTES					
Prueba de la Mediana Generalizada	Al menos ordinal	• La prueba de la mediana generalizada es un procedimiento para probar si k grupos independientes difieren en sus tendencias centrales, es decir se utiliza cuando se desea probar si ha sido tomados de la misma población o de poblaciones con medianas iguales.	• La prueba de la mediana usa uno de los procedimientos más sencillos y usados de todos los métodos no paramétricos que se utilizan para probar si dos muestras independientes provienen de poblaciones con medianas iguales. • La prueba de la mediana es más adecuada cuando los datos son observaciones que no pueden ordenarse completamente y sirven solamente para dicotomizar las observaciones arriba o debajo de la mediana combinada.	• La prueba puede aplicarse en los mismos casos que la de Kruskal-Wallis pero es preferible utilizar esta última es mas eficiente.	• Esta prueba se utiliza siempre que uno desee comparar k poblaciones o muestras para probar si difieren con respecto a la mediana (comportamientos sociales, psicológicos, relaciones de niveles de educativos).
Prueba de Kruskal y Wallis	Al menos ordinal	• La prueba de Kruska-Wallis es extremadamente útil para decidir si K muestras independientes provienen de la misma población o que al menos una de las poblaciones difiere en localización. • Esta prueba es la mas usada como un procedimiento alternativo al de la prueba F de análisis de varianza cuando el supuesto de normalidad no se justifica.	• Es un procedimiento diseñando para ser sensible en pruebas e diferencia entre medias. • Debido a que la prueba de Kruskal-Wallis usa mas la información de las observaciones debido a que convierte los puntajes en rangos esta prueba es la mas eficiente para K muestras independientes.	• Cuando se obtiene un resultado significativo con la prueba de Kruskal-Wallis lo único que se sabe es que existe alguna diferencia en localización entre las muestras, sin embargo no se sabe el porque de la diferencia.	• Esta prueba se utiliza para probar igualdad de tratamientos muy usada en Diseño de experimentos al comparar tratamientos, técnicas, maquinarias, etc
TABLAS DE CONTINGENCIAS					
Análisis de Tablas de Contingencia	Nivel nominal u ordinal	• Lo que se plantea con las tablas de contingencia es el análisis simultaneo de dos características de una misma población, y conocer la posible dependencia u homogeneidad entre las dos clasificaciones. • Esta prueba se utiliza cuando se requieren estudios en los cuales los elementos están categorizados de acuerdo con dos criterios de clasificación o en situaciones en las que no se dispone de medidas númericas exactas.	• Esta prueba es muy útil cuando se desea probar independencia o dependencia entre variables o categorías. • También es útil cuando interesa investigar si en las diferentes poblaciones estudiadas los valores o categorías de cada una de las manifestaciones se presentan en la misma proporción.		• Esta prueba se utiliza cuando se tienen datos en forma matricial con 2 criterios de clasificación puede ser sexo y estado civil, nivel de estudios y alimentación, etc.
PRUEBAS DE ASOCIACION					
Coeficiente de Correlación de Spearman	Al menos ordinal	• Este coeficiente se utiliza cuando se quiere determinar el grado en que dos variables están relacionadas o tienden a variar conjuntamente.	• Este coeficiente es fácil de calcular y tiene la ventaja adicional que esta relacionado linealmente con el coeficiente de concordancia W de Kendall		• Esta prueba se utiliza siempre que se requiera saber el nivel de asociación entre dos variable s como pueden ser preferencias y precios, durabilidad y color, etc.

ANEXOS

COMPARACIONES MULTIPLES USANDO LA PRUEBA DEL SIGNO

Puesto que la prueba de Friedman es equivalente a la prueba del signo cuando sólo se están comparando dos muestras, es natural usar la prueba del signo para efectuar comparaciones multiples después de una prueba de Friedman significativa. Esto permite determinar cuáles de las diferencias entre las muestras son la causa del resultado significativo total.

Por tanto, una vez que se ha obtenido un resultado significativo usando la prueba Friedman, se realiza una prueba del signo entre todas las parejas de muestras usando un nivel de significancia de dos colas $\frac{\alpha}{\left[\frac{k(k-2)}{2}\right]}$, donde α es la tasa de error por experimento requerida y k es el número de muestras. Para prevenir inconsistencias, primero se realiza una prueba del signo entre las muestras con las sumas de rangos menor y mayor y después entre las muestras con las sumas de rangos menor y la segunda mayor, así sucesivamente hasta que se compara con la mayor, después con la segunda mayor, y así sucesivamente. Siempre que se obtenga un resultado no significativo o cuando debe hacerse una comparación que no sea significativa por implicación, en esa etapa se detiene le procedimiento.

Procedimiento del Método de Comparaciones Múltiples

1. Seleccionar la tasa de error por experimento α, la cual puede ser mayor que la α usada en la prueba completa si se está preparado para correr el riesgo.
2. Determinar c, el número de comparaciones que quieren hacerse. Normalmente querrán compararse todos los pares posibles de muestras. Si hay k muestras, el número de pares será $c = k(k-1)/2$
3. Encontrar el nivel de significación más próximo a α/c en la tabla A.
4. Ordenar las k muestras con respecto a sus rangos promedio (dados por R_i/t_i) y escribir los símbolos de las muestras en orden.
5. Usando una prueba de dos colas con nivel de significación según se calculó en el punto 3, realizar las pruebas de suma de rangos de la siguiente manera. Usando la ordenación dada en el punto 4, comparar primero la muestra de la extrema izquierda con la de la extrema derecha, después con la segunda de la derecha y así sucesivamente hasta que se obtenga un resultado no significativo. Cuando esto suceda, unir con una línea los símbolos de las dos muestras. Tomar entonces la segunda muestra de la izquierda y compararla primero con la de la extrema derecha, después con la segunda de la derecha y así sucesivamente hasta que se estén comparando dos muestra que ya estén unidas por una línea. Continuar de esta manera hasta que se hayan agotado todas las comparaciones.

Existen otras pruebas pero uno tiene que fijarse si estos procedimientos son validos para el tipo de datos, procedimientos como contrastes, LSD, La prueba de Intervalos Múltiples de Duncan, Prueba de Newman-Keuls[1], etc.

[1] Para mas información ver el Libro de Diseño y Análisis de Experimentos de Douglas C. Montgomery.

TABLA A. VALORES CRITICOS DE S PARA LA PRUEBA DE SUMA DE RANGOS DE WILCOXON

t_1 es el número de observaciones de la muestra mayor
t_2 es el número de observaciones de la muestra menor
La tabla puede usarse en todos los casos en que tanto t_1 como t_2 sean menores que 26.

Ejemplos En una prueba de dos colas con $t_1 = 11$, $t_2 = 10$ y $\alpha = 0.05$, rechazar la hipótesis nula si la S obtenida es mayor o igual que 58 o si es menor o igual que −58.
En una prueba de una cola con $t_1 = 11$, $t_2 = 10$ y $\alpha = 0.05$; si se necesita una prueba de cola superior, rechazar la hipótesis nula si la S obtenida es mayor o igual que 48. Si se necesita una prueba de cola inferior, rechazarla si la S obtenida es menor o igual que −48.
Para algunos casos, como $t_1 = 5$, $t_2 = 1$, no aparecen cifras. Esto significa que, para los niveles de significación dados, ninguno de los valores posibles de S llevaría a rechazar la hipótesis nula.

Nivel de significación de una cola, α: 0.100 0.050 0.025 0.010 0.005 0.001
Nivel de significación de dos colas, α: 0.200 0.100 0.050 0.020 0.010 0.002

t_1	t_2	0.100 / 0.200	0.050 / 0.100	0.025 / 0.050	0.010 / 0.020	0.005 / 0.010	0.001 / 0.002
3	2	6					
3	3	7	9				
4	2	8					
4	3	10	12				
4	4	10	14	16			
5	2	8	10				
5	3	11	13	15			
5	4	12	16	18	20		
5	5	15	17	21	23		
6	2	10	12				
6	3	12	14	16			
6	4	14	18	20	22	24	
6	5	16	20	24	26	28	
6	6	18	22	26	30	32	
7	2	12	14				
7	3	13	17	19	21		
7	4	16	20	22	26	28	
7	5	19	23	25	29	33	
7	6	20	26	30	34	36	42
7	7	23	27	33	37	41	47
8	2	12	14	16			
8	3	14	18	20	24		
8	4	18	22	24	28	30	
8	5	20	24	28	32	36	40
8	6	22	28	32	34	40	46
8	7	24	30	36	42	44	52
8	8	26	34	38	46	50	56
9	1	9					
9	2	14	16	18			
9	3	17	19	23	25	27	
9	4	18	24	28	30	34	
9	5	21	27	31	35	39	43
9	6	24	30	34	40	44	50
9	7	27	33	39	45	47	57
9	8	28	36	42	50	54	62
9	9	31	39	47	53	59	67
10	1	10					
10	2	14	18	20			
10	3	18	22	24	28	30	
10	4	20	24	30	34	36	40
10	5	24	28	34	38	42	48
10	6	26	32	38	44	48	54
10	7	28	34	42	48	52	60
10	8	32	40	46	54	58	68
10	9	34	42	50	58	64	74
10	10	36	46	54	62	68	80
11	1	11					
11	2	16	20	22			
11	3	19	23	27	31	33	
11	4	22	28	32	36	40	44
11	5	25	31	37	41	45	51
11	6	28	34	40	46	52	58
11	7	31	39	45	53	57	65
11	8	34	42	50	58	62	72
11	9	37	45	53	63	67	79
11	10	38	48	58	66	74	86
11	11	41	53	61	71	79	91
12	1	12					
12	2	16	20	22			
12	3	20	26	28	32	34	
12	4	24	30	34	38	42	48
12	5	26	34	38	44	48	54
12	6	30	38	44	50	54	64
12	7	32	42	48	56	60	70
12	8	34	44	52	62	66	78
12	9	38	48	56	64	72	84
12	10	42	52	62	72	78	92
12	11	44	56	66	76	84	98
12	12	46	60	70	82	90	104
13	1	13					
13	2	18	22	24	26		
13	3	21	27	31	35	37	
13	4	24	32	36	42	46	50
13	5	29	35	41	47	51	59
13	6	32	40	46	54	58	68
13	7	35	43	51	59	65	75

TABLA A (continuación)

Nivel de significación de una cola, α: 0.100 0.050 0.025 0.010 0.005 0.001
Nivel de significación de dos colas, α: 0.200 0.100 0.050 0.020 0.010 0.002

t_1	t_2	0.100 / 0.200	0.050 / 0.100	0.025 / 0.050	0.010 / 0.020	0.005 / 0.010	0.001 / 0.002
13	8	38	48	56	64	70	82
13	9	41	51	61	71	77	89
13	10	44	54	64	76	82	96
13	11	47	59	69	81	89	103
13	12	50	62	74	86	94	110
13	13	53	67	79	91	101	117
14	1	14					
14	2	18	22	26	28		
14	3	22	28	32	38	40	
14	4	26	34	38	44	48	54
14	5	30	38	44	50	56	64
14	6	34	42	50	58	62	72
14	7	36	46	54	64	68	80
14	8	40	50	60	68	74	88
14	9	44	54	64	74	82	96
14	10	46	58	68	80	88	102
14	11	50	62	74	86	94	110
14	12	52	66	78	92	100	118
14	13	56	70	82	96	106	124
14	14	58	74	86	102	112	132
15	1	15					
15	2	20	24	28	30		
15	3	23	31	35	39	41	
15	4	28	36	40	46	50	58
15	5	31	39	47	53	59	67
15	6	36	44	52	60	66	76
15	7	39	49	57	67	73	85
15	8	42	54	62	72	80	92
15	9	45	57	67	79	87	101
15	10	48	62	72	84	92	108
15	11	51	65	77	91	99	117
15	12	54	70	82	96	106	124
15	13	59	73	87	101	111	131
15	14	62	78	92	108	118	138
15	15	65	81	97	113	123	145
16	1	16					
16	2	22	26	30	32		
16	3	26	32	36	42	44	
16	4	30	36	42	50	54	60
16	5	34	42	50	56	62	70
16	6	38	46	54	64	70	80
16	7	40	52	60	70	76	90
16	8	44	56	66	76	84	98
16	9	48	60	70	82	90	106
16	10	52	64	76	88	98	114
16	11	54	68	82	94	104	122
16	12	56	72	86	100	110	130
16	13	60	76	90	106	118	138
16	14	64	82	96	112	124	146
16	15	68	86	100	118	130	154
16	16	70	90	106	124	136	160
17	1	17					
17	2	22	28	30	34		
17	3	27	33	39	43	47	51
17	4	32	38	46	52	56	64
17	5	35	45	51	59	65	75
17	6	40	50	56	66	72	84
17	7	43	53	63	73	81	93
17	8	46	58	68	80	86	102
17	9	49	63	75	87	95	111
17	10	54	66	80	94	102	120
17	11	57	73	85	99	109	129
17	12	60	76	90	106	114	136
17	13	63	81	95	111	123	145
17	14	68	84	100	118	130	152
17	15	71	89	105	123	135	161
17	16	74	94	110	130	142	168
17	17	77	97	115	135	149	175
18	1	18					
18	2	24	28	32	36		
18	3	28	34	40	46	50	54
18	4	32	40	48	54	60	66
18	5	36	46	54	62	68	78
18	6	40	52	60	70	76	88
18	7	44	56	66	78	84	98
18	8	46	62	72	84	92	108
18	9	52	66	78	90	100	116
18	10	56	70	84	98	104	126
18	11	60	76	88	104	114	134
18	12	62	80	94	110	122	142
18	13	66	84	100	116	128	150
18	14	70	88	104	122	136	140
18	15	74	94	110	130	142	168
18	16	76	98	116	136	148	174
18	17	80	102	120	142	156	184
18	18	84	106	126	148	162	192
19	1	17	19				
19	2	24	30	34	36	38	
19	3	29	37	43	49	51	57
19	4	34	42	50	58	62	70

TABLA A. VALORES CRITICOS DE S PARA LA PRUEBA DE SUMA DE RANGOS DE WILCOXON

t_1 es el número de observaciones de la muestra mayor
t_2 es el número de observaciones de la muestra menor
La tabla puede usarse en todos los casos en que tanto t_1 como t_2 sean menores que 26.

Ejemplos En una prueba de dos colas con $t_1 = 11$, $t_2 = 10$ y $\alpha = 0.05$, rechazar la hipótesis nula si la S obtenida es mayor o igual que 58 o si es menor o igual que −58.
En una prueba de una cola con $t_1 = 11$, $t_2 = 10$ y $\alpha = 0.05$; si se necesita una prueba de cola superior, rechazar la hipótesis nula si la S obtenida es mayor o igual que 48. Si se necesita una prueba de cola inferior, rechazarla si la S obtenida es menor o igual que −48.
Para algunos casos, como $t_1 = 5$, $t_2 = 1$, no aparecen cifras. Esto significa que, para los niveles de significación dados, ninguno de los valores posibles de S llevaría a rechazar la hipótesis nula.

		Nivel de significación de una cola, α					
		0.100	0.050	0.025	0.010	0.005	0.001
		Nivel de significación de dos colas, α					
t_1	t_2	0.200	0.100	0.050	0.020	0.010	0.002
3	2	6					
3	3	7	9				
4	2	8					
4	3	10	12				
4	4	10	14	16			
5	2	8	10				
5	3	11	13	15			
5	4	12	16	18	20		
5	5	15	17	21	23		
6	2	10	12				
6	3	12	14	16			
6	4	14	18	20	22	24	
6	5	16	20	24	26	28	
6	6	18	22	26	30	32	
7	2	12	14				
7	3	13	17	19	21		
7	4	16	20	22	26	28	
7	5	19	23	25	29	33	
7	6	20	26	30	34	36	42
7	7	23	27	33	37	41	47
8	2	12	14	16			
8	3	14	18	20	24		
8	4	18	22	24	28	30	
8	5	20	24	28	32	36	40
8	6	22	28	32	36	40	44
8	7	24	30	34	42	44	52
8	8	26	34	38	44	50	56
9	1	9					
9	2	14	16	18			
9	3	17	19	23	25	27	
9	4	18	24	28	30	34	
9	5	21	27	31	35	39	43
9	6	24	30	34	40	44	50
9	7	27	33	39	45	49	57
9	8	28	36	42	50	54	62
9	9	31	39	47	53	59	67
10	1	10					
10	2	14	18	20			

		Nivel de significación de una cola, α					
		0.100	0.050	0.025	0.010	0.005	0.001
		Nivel de significación de dos colas, α					
t_1	t_2	0.200	0.100	0.050	0.020	0.010	0.002
10	3	18	22	24	28	30	
10	4	20	26	30	34	36	40
10	5	24	28	34	38	42	48
10	6	26	32	38	44	48	54
10	7	28	34	42	48	52	60
10	8	32	40	46	54	58	68
10	9	34	42	50	58	64	74
10	10	36	44	54	62	68	80
11	1	11					
11	2	16	20	22			
11	3	19	23	27	31	33	
11	4	22	28	32	36	40	44
11	5	25	31	37	41	45	51
11	6	28	34	40	46	52	58
11	7	31	39	45	53	57	65
11	8	34	42	50	58	62	72
11	9	37	45	53	63	67	79
11	10	38	48	58	66	74	84
11	11	41	53	61	71	79	91
12	1	12					
12	2	14	20	22			
12	3	20	26	28	32	34	
12	4	24	30	34	38	42	48
12	5	26	34	38	44	48	56
12	6	30	38	44	50	54	64
12	7	32	42	48	56	60	70
12	8	36	44	52	62	66	78
12	9	38	48	56	64	72	84
12	10	42	52	62	72	78	92
12	11	44	56	66	76	84	98
12	12	46	60	70	82	90	104
13	1	13					
13	2	18	22	24	26		
13	3	21	27	31	35	37	
13	4	26	32	34	42	46	50
13	5	29	35	41	47	51	59
13	6	32	40	46	54	58	68
13	7	35	43	51	59	65	75

TABLA A (continuación)

		Nivel de significación de una cola, α					
		0.100	0.050	0.025	0.010	0.005	0.001
		Nivel de significación de dos colas, α					
t_1	t_2	0.200	0.100	0.050	0.020	0.010	0.002
13	8	38	48	56	64	70	82
13	9	41	51	61	71	77	89
13	10	44	56	64	74	82	94
13	11	47	59	69	81	89	103
13	12	50	62	74	86	94	110
13	13	53	67	79	91	101	117
14	1	14					
14	2	18	22	26	28		
14	3	22	28	32	38	40	
14	4	26	34	38	44	48	54
14	5	30	38	44	50	56	64
14	6	34	42	50	58	62	72
14	7	36	46	54	64	68	80
14	8	40	50	60	68	76	88
14	9	44	54	64	74	82	96
14	10	46	58	68	80	88	102
14	11	50	62	74	86	94	110
14	12	52	66	78	92	100	118
14	13	56	70	82	96	104	124
14	14	58	74	86	102	112	132
15	1	15					
15	2	20	24	28	30		
15	3	25	31	35	39	41	
15	4	28	36	40	46	50	56
15	5	31	39	47	53	59	67
15	6	36	44	52	60	66	76
15	7	39	49	57	67	73	85
15	8	42	54	62	72	80	92
15	9	45	57	67	79	87	101
15	10	48	62	72	84	92	108
15	11	51	65	77	91	99	117
15	12	54	70	82	96	104	124
15	13	59	73	87	101	111	131
15	14	62	78	92	108	118	138
15	15	65	81	97	113	123	145
16	1	16					
16	2	22	26	30	32		
16	3	26	32	36	42	44	
16	4	30	36	42	50	54	60
16	5	34	42	50	56	62	70
16	6	38	46	54	64	70	80
16	7	40	52	60	70	76	90
16	8	44	56	66	76	84	98
16	9	48	60	70	82	90	106
16	10	52	64	76	88	98	114

		Nivel de significación de una cola, α					
		0.100	0.050	0.025	0.010	0.005	0.001
		Nivel de significación de dos colas, α					
t_1	t_2	0.200	0.100	0.050	0.020	0.010	0.002
16	11	54	68	82	94	104	122
16	12	58	72	86	100	110	130
16	13	60	76	90	106	118	138
16	14	64	82	96	112	124	146
16	15	68	86	100	118	130	154
16	16	70	90	106	124	136	160
17	1	17					
17	2	22	26	30	34		
17	3	27	33	39	43	47	51
17	4	32	38	46	52	56	64
17	5	35	45	51	59	65	75
17	6	40	50	58	66	72	84
17	7	43	53	63	73	81	93
17	8	46	58	68	80	86	102
17	9	49	63	75	87	95	111
17	10	54	68	80	94	102	120
17	11	57	73	85	99	109	129
17	12	60	76	90	106	116	136
17	13	63	81	95	111	123	145
17	14	68	84	100	118	130	152
17	15	71	89	105	123	135	161
17	16	74	94	110	130	142	168
17	17	77	97	115	135	149	175
18	1	18					
18	2	24	28	32	36		
18	3	28	36	40	46	50	54
18	4	32	40	48	54	60	66
18	5	36	46	54	62	68	78
18	6	40	52	60	70	76	88
18	7	44	56	66	78	84	98
18	8	48	62	72	84	92	108
18	9	52	66	78	90	100	114
18	10	56	70	84	98	106	126
18	11	60	76	88	104	114	134
18	12	62	80	94	110	122	142
18	13	66	84	100	116	128	150
18	14	70	88	104	122	136	160
18	15	74	94	110	130	142	168
18	16	76	98	116	136	148	176
18	17	80	102	120	142	156	184
18	18	84	106	126	148	162	192
19	1	17	19				
19	2	24	30	34	36	38	
19	3	29	37	43	49	51	57
19	4	34	42	50	58	62	70

TABLA A (continuación)

r_1	r_2	Nivel de significación de una cola, α: 0.100 / Nivel de significación de dos colas, α: 0.200	0.050 / 0.100	0.025 / 0.050	0.010 / 0.020	0.005 / 0.010	0.001 / 0.002
19	5	39	47	57	65	71	81
19	6	42	54	64	74	80	92
19	7	47	59	69	81	89	103
19	8	50	64	76	88	96	112
19	9	55	69	81	95	105	121
19	10	58	74	86	102	112	132
19	11	63	79	93	109	119	141
19	12	66	84	98	116	126	148
19	13	69	87	103	121	133	157
19	14	72	92	110	128	140	166
19	15	77	97	115	135	147	175
19	16	80	102	120	140	156	184
19	17	83	105	125	147	161	191
19	18	86	110	130	154	168	200
19	19	91	115	135	159	175	207
20	1	18	20				
20	2	26	32	34	38	40	
20	3	30	38	44	50	54	60
20	4	36	44	52	60	64	74
20	5	40	50	60	68	74	86
20	6	44	56	66	76	84	96
20	7	48	62	72	84	92	108
20	8	52	66	78	92	100	118
20	9	56	72	84	100	108	128
20	10	60	76	90	106	114	136
20	11	64	82	96	114	124	146
20	12	68	86	102	120	132	156
20	13	72	92	108	126	140	164
20	14	76	96	114	134	146	172
20	15	80	100	120	140	154	182
20	16	82	106	124	146	162	190
20	17	86	110	130	154	168	200
20	18	90	114	136	160	174	208
20	19	94	120	142	166	182	216
20	20	98	124	146	172	190	224
21	1	19	21				
21	2	26	32	36	40	42	
21	3	33	41	47	53	57	61
21	4	38	46	54	62	68	74
21	5	41	53	61	71	77	89
21	6	46	58	68	80	88	102
21	7	51	65	75	87	97	111
21	8	56	70	82	96	104	122
21	9	59	75	89	103	113	133
21	10	64	80	94	110	122	142
21	11	67	85	101	117	129	151
21	12	70	90	106	124	136	162
21	13	75	95	113	131	145	171
21	14	78	100	118	138	152	180
21	15	83	105	123	145	159	189
21	16	86	110	130	152	168	198
21	17	89	115	135	159	175	207
21	18	94	118	140	166	182	216
21	19	97	123	147	173	189	225
21	20	100	128	152	178	196	232
21	21	105	133	157	183	205	241
22	1	20	22				
22	2	28	34	38	42	44	
22	3	34	42	48	54	58	64
22	4	38	48	56	64	70	80
22	5	44	54	64	74	82	94
22	6	48	60	72	84	90	106
22	7	52	66	78	92	100	116
22	8	58	72	86	100	108	128
22	9	62	78	92	108	118	138
22	10	66	84	98	114	126	148
22	11	70	88	104	122	134	158
22	12	74	94	110	130	142	168
22	13	78	98	116	136	150	178
22	14	82	104	122	144	158	186
22	15	86	108	128	150	166	196
22	16	90	114	134	158	174	206
22	17	92	118	140	164	182	214
22	18	96	124	146	172	188	224
22	19	100	128	152	178	196	232
22	20	104	132	158	186	204	242
22	21	108	138	162	192	212	250
22	22	112	142	168	198	218	260
23	1	21	23				
23	2	28	36	40	44	46	
23	3	35	43	51	57	61	67
23	4	40	50	58	68	74	84
23	5	45	57	67	77	85	97
23	6	50	64	74	86	94	110
23	7	55	69	81	95	103	121
23	8	60	76	88	104	114	132
23	9	63	81	95	111	121	143
23	10	68	86	102	120	130	154
23	11	73	91	107	127	139	163
23	12	76	96	114	134	148	174

TABLA A (continuación)

r_1	r_2	Nivel de significación de una cola, α: 0.100 / Nivel de significación de dos colas, α: 0.200	0.050 / 0.100	0.025 / 0.050	0.010 / 0.020	0.005 / 0.010	0.001 / 0.002
23	13	81	103	121	141	155	183
23	14	84	108	126	148	164	194
23	15	89	113	133	157	171	203
23	16	92	118	138	164	180	212
23	17	97	123	145	171	187	221
23	18	100	128	150	178	196	232
23	19	103	133	157	185	203	241
23	20	108	138	162	192	210	250
23	21	111	143	169	199	219	259
23	22	116	148	174	206	226	268
23	23	119	151	179	213	233	277
24	1	22	24				
24	2	30	36	42	44	48	
24	3	36	46	52	60	64	70
24	4	40	52	62	70	76	86
24	5	48	60	70	80	88	100
24	6	52	66	78	90	98	114
24	7	56	72	84	98	108	126
24	8	62	78	92	108	118	136
24	9	66	84	98	116	126	148
24	10	70	90	106	124	136	160
24	11	74	94	112	132	144	170
24	12	78	100	118	138	152	180
24	13	84	104	124	146	162	190
24	14	88	110	132	154	170	200
24	15	92	116	138	162	178	210
24	16	96	122	144	168	186	220
24	17	100	126	150	176	194	230
24	18	104	132	156	184	202	240
24	19	108	136	162	190	210	248
24	20	112	142	168	196	218	258
24	21	116	146	174	204	226	268
24	22	120	152	180	212	234	276
24	23	122	156	186	218	242	286
24	24	126	162	192	226	248	296
25	1	23	25				
25	2	32	38	44	48	50	
25	3	37	47	55	61	65	73
25	4	44	54	64	74	80	90
25	5	49	61	71	83	91	105
25	6	54	68	80	92	102	118
25	7	59	75	87	103	111	131
25	8	64	80	94	110	122	142
25	9	69	87	101	119	131	153
25	10	72	92	106	126	140	164
25	11	77	97	115	135	149	175
25	12	82	104	122	144	158	186
25	13	85	109	129	151	167	197
25	14	90	114	136	160	176	206
25	15	95	119	141	167	183	217
25	16	98	124	148	174	192	228
25	17	103	131	155	181	201	237
25	18	106	136	160	190	208	246
25	19	111	141	167	197	217	257
25	20	114	146	174	204	224	266
25	21	119	151	179	211	233	275
25	22	122	156	186	218	240	286
25	23	127	161	191	225	249	295
25	24	130	166	198	232	256	304
25	25	135	171	203	241	265	315

TABLA 8 Función de distribución de U

$P(U \leq U_0)$: U_0 es el argumento; $n_1 \leq n_2$, $3 \leq n_2 \leq 10$.

$n_2 = 3$

	n_1		
U_0	*1*	*2*	*3*
0	.25	.10	.05
1	.50	.20	.10
2		.40	.20
3		.60	.35
4			.50

$n_2 = 4$

	n_1			
U_0	*1*	*2*	*3*	*4*
0	.2000	.0667	.0286	.0143
1	.4000	.1333	.0571	.0286
2	.6000	.2667	.1143	.0571
3		.4000	.2000	.1000
4		.6000	.3143	.1714
5			.4286	.2429
6			.5714	.3429
7				.4429
8				.5571

Tablas

TABLA 8 (Continuación)

$n_2 = 5$

	n_1				
U_0	*1*	*2*	*3*	*4*	*5*
0	.1667	.0476	.0179	.0079	.0040
1	.3333	.0952	.0357	.0159	.0079
2	.5000	.1905	.0714	.0317	.0159
3		.2857	.1250	.0556	.0278
4		.4286	.1964	.0952	.0476
5		.5714	.2857	.1429	.0754
6			.3929	.2063	.1111
7			.5000	.2778	.1548
8				.3651	.2103
9				.4524	.2738
10				.5476	.3452
11					.4206
12					.5000

$n_2 = 6$

	n_1					
U_0	*1*	*2*	*3*	*4*	*5*	*6*
0	.1429	.0357	.0119	.0048	.0022	.0011
1	.2857	.0714	.0238	.0095	.0043	.0022
2	.4286	.1429	.0476	.0190	.0087	.0043
3	.5714	.2143	.0833	.0333	.0152	.0076
4		.3214	.1310	.0571	.0260	.0130
5		.4286	.1905	.0857	.0411	.0206
6		.5714	.2738	.1286	.0628	.0325
7			.3571	.1762	.0887	.0465
8			.4524	.2381	.1234	.0660
9			.5476	.3048	.1645	.0898
10				.3810	.2143	.1201
11				.4571	.2684	.1548
12				.5429	.3312	.1970
13					.3961	.2424
14					.4654	.2944
15					.5346	.3496
16						.4091
17						.4686
18						.5314

TABLA 8 (Continuación)

$n_2 = 7$

	n_1						
U_0	1	2	3	4	5	6	7
0	.1250	.0278	.0083	.0030	.0013	.0006	.0003
1	.2500	.0556	.0167	.0061	.0025	.0012	.0006
2	.3750	.1111	.0333	.0121	.0051	.0023	.0012
3	.5000	.1667	.0583	.0212	.0088	.0041	.0020
4		.2500	.0917	.0364	.0152	.0070	.0035
5		.3333	.1333	.0545	.0240	.0111	.0055
6		.4444	.1917	.0818	.0366	.0175	.0087
7		.5556	.2583	.1152	.0530	.0256	.0131
8			.3333	.1576	.0745	.0367	.0189
9			.4167	.2061	.1010	.0507	.0265
10			.5000	.2636	.1338	.0688	.0364
11				.3242	.1717	.0903	.0487
12				.3939	.2159	.1171	.0641
13				.4636	.2652	.1474	.0825
14				.5364	.3194	.1830	.1043
15					.3775	.2226	.1297
16					.4381	.2669	.1588
17					.5000	.3141	.1914
18						.3654	.2279
19						.4178	.2675
20						.4726	.3100
21						.5274	.3552
22							.4024
23							.4508
24							.5000

TABLA 8 (Continuación)

$n_2 = 8$

	n_1							
U_0	1	2	3	4	5	6	7	8
0	.1111	.0222	.0061	.0020	.0008	.0003	.0002	.0001
1	.2222	.0444	.0121	.0040	.0016	.0007	.0003	.0002
2	.3333	.0889	.0242	.0081	.0031	.0013	.0006	.0003
3	.4444	.1333	.0424	.0141	.0054	.0023	.0011	.0005
4	.5556	.2000	.0667	.0242	.0093	.0040	.0019	.0009
5		.2667	.0970	.0364	.0148	.0063	.0030	.0015
6		.3556	.1394	.0545	.0225	.0100	.0047	.0023
7		.4444	.1879	.0768	.0326	.0147	.0070	.0035
8		.5556	.2485	.1071	.0466	.0213	.0103	.0052
9			.3152	.1414	.0637	.0296	.0145	.0074
10			.3879	.1838	.0855	.0406	.0200	.0103
11			.4606	.2303	.1111	.0539	.0270	.0141
12			.5394	.2848	.1422	.0709	.0361	.0190
13				.3414	.1772	.0906	.0469	.0249
14				.4040	.2176	.1142	.0603	.0325
15				.4667	.2618	.1412	.0760	.0415
16				.5333	.3108	.1725	.0946	.0524
17					.3621	.2068	.1159	.0652
18					.4165	.2454	.1405	.0803
19					.4716	.2864	.1678	.0974
20					.5284	.3310	.1984	.1172
21						.3773	.2317	.1393
22						.4259	.2679	.1641
23						.4749	.3063	.1911
24						.5251	.3472	.2209
25							.3894	.2527
26							.4333	.2869
27							.4775	.3227
28							.5225	.3605
29								.3992
30								.4392
31								.4796
32								.5204

TABLA 8 (Continuación)

$n_2 = 9$

U_0	n_1 = 1	2	3	4	5	6	7	8	9
0	.1000	.0182	.0045	.0014	.0005	.0002	.0001	.0000	.0000
1	.2000	.0364	.0091	.0028	.0010	.0004	.0002	.0001	.0000
2	.3000	.0727	.0182	.0056	.0020	.0008	.0003	.0002	.0001
3	.4000	.1091	.0318	.0098	.0035	.0014	.0006	.0003	.0001
4	.5000	.1636	.0500	.0168	.0060	.0024	.0010	.0005	.0002
5		.2182	.0727	.0252	.0095	.0038	.0017	.0008	.0004
6		.2909	.1045	.0378	.0145	.0060	.0026	.0012	.0006
7		.3636	.1409	.0531	.0210	.0088	.0039	.0019	.0009
8		.4545	.1864	.0741	.0300	.0128	.0058	.0028	.0014
9		.5455	.2409	.0993	.0415	.0180	.0082	.0039	.0020
10			.3000	.1301	.0559	.0248	.0115	.0056	.0028
11			.3636	.1650	.0734	.0332	.0156	.0076	.0039
12			.4318	.2070	.0949	.0440	.0209	.0103	.0053
13			.5000	.2517	.1199	.0567	.0274	.0137	.0071
14				.3021	.1489	.0723	.0356	.0180	.0094
15				.3552	.1818	.0905	.0454	.0232	.0122
16				.4126	.2188	.1119	.0571	.0296	.0157
17				.4699	.2592	.1361	.0708	.0372	.0200
18				.5301	.3032	.1638	.0869	.0464	.0252
19					.3497	.1942	.1052	.0570	.0313
20					.3986	.2280	.1261	.0694	.0385
21					.4491	.2643	.1496	.0836	.0470
22					.5000	.3035	.1755	.0998	.0567
23						.3445	.2039	.1179	.0680
24						.3878	.2349	.1383	.0807
25						.4320	.2680	.1606	.0951
26						.4773	.3032	.1852	.1112
27						.5227	.3403	.2117	.1290
28							.3788	.2404	.1487
29							.4185	.2707	.1701
30							.4591	.3029	.1933
31							.5000	.3365	.2181
32								.3715	.2447
33								.4074	.2729
34								.4442	.3024
35								.4813	.3332
36								.5187	.3652
37									.3981
38									.4317
39									.4657
40									.5000

TABLA 8 (Continuación)

$n_2 = 10$

U_0	n_1 = 1	2	3	4	5	6	7	8	9	10
0	.0909	.0152	.0035	.0010	.0003	.0001	.0001	.0000	.0000	.000
1	.1818	.0303	.0070	.0020	.0007	.0002	.0001	.0000	.0000	.000
2	.2727	.0606	.0140	.0040	.0013	.0005	.0002	.0001	.0000	.000
3	.3636	.0909	.0245	.0070	.0023	.0009	.0004	.0002	.0001	.000
4	.4545	.1364	.0385	.0120	.0040	.0015	.0006	.0003	.0001	.000
5	.5455	.1818	.0559	.0180	.0063	.0024	.0010	.0004	.0002	.000
6		.2424	.0804	.0270	.0097	.0037	.0015	.0007	.0003	.000
7		.3030	.1084	.0380	.0140	.0055	.0023	.0010	.0005	.000
8		.3788	.1434	.0529	.0200	.0080	.0034	.0015	.0007	.000
9		.4545	.1853	.0709	.0276	.0112	.0048	.0022	.0011	.000
10		.5455	.2343	.0939	.0376	.0156	.0068	.0031	.0015	.000
11			.2867	.1199	.0496	.0210	.0093	.0043	.0021	.001
12			.3462	.1518	.0646	.0280	.0125	.0058	.0028	.001
13			.4056	.1868	.0823	.0363	.0165	.0078	.0038	.001
14			.4685	.2268	.1032	.0467	.0215	.0103	.0051	.002
15			.5315	.2697	.1272	.0589	.0277	.0133	.0066	.003
16				.3177	.1548	.0736	.0351	.0171	.0086	.004
17				.3666	.1855	.0903	.0439	.0217	.0110	.005
18				.4196	.2198	.1099	.0544	.0273	.0140	.007
19				.4725	.2567	.1317	.0665	.0338	.0175	.009
20				.5275	.2970	.1566	.0806	.0416	.0217	.011
21					.3393	.1838	.0966	.0506	.0267	.014
22					.3839	.2139	.1148	.0610	.0326	.017
23					.4296	.2461	.1349	.0729	.0394	.021
24					.4765	.2811	.1574	.0864	.0474	.026
25					.5235	.3177	.1819	.1015	.0564	.031
26						.3564	.2087	.1185	.0667	.037
27						.3962	.2374	.1371	.0782	.044
28						.4374	.2681	.1577	.0912	.052
29						.4789	.3004	.1800	.1055	.061
30						.5211	.3345	.2041	.1214	.071
31							.3698	.2299	.1388	.082
32							.4063	.2574	.1577	.095
33							.4434	.2863	.1781	.108
34							.4811	.3167	.2001	.123
35							.5189	.3482	.2235	.139
36								.3809	.2483	.157
37								.4143	.2745	.176
38								.4484	.3019	.196
39								.4827	.3304	.217

TABLA 8 (Continuación)

$n_2 = 10$ (Continuación)

					n_1					
C_u	1	2	3	4	5	6	7	8	9	10
40								5173	3598	2406
41									3901	2644
42									4211	2894
43									4524	3153
44									.4841	3421
45									.5159	.3697
46										.3980
47										4267
48										4559
49										4853
50										5147

Calculado por M. Pagano, Departamento de Estadística, Universidad de Florida.

TABLA 9 Valores críticos de T en la prueba de Wilcoxon de rangos con signo para experimentos apareados;
$n = 5(1)50$

Unilateral	*Bilateral*	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
$P = .05$	$P = .10$	1	2	4	6	8	11
$P = .025$	$P = .05$		1	2	4	6	8
$P = .01$	$P = .02$			0	2	3	5
$P = .005$	$P = .01$				0	2	3

Unilateral	*Bilateral*	$n = 11$	$n = 12$	$n = 13$	$n = 14$	$n = 15$	$n = 16$
$P = .05$	$P = .10$	14	17	21	26	30	36
$P = .025$	$P = .05$	11	14	17	21	25	30
$P = .01$	$P = .02$	7	10	13	16	20	24
$P = .005$	$P = .01$	5	7	10	13	16	19

Unilateral	*Bilateral*	$n = 17$	$n = 18$	$n = 19$	$n = 20$	$n = 21$	$n = 22$
$P = .05$	$P = .10$	41	47	54	60	68	75
$P = .025$	$P = .05$	35	40	46	52	59	66
$P = .01$	$P = .02$	28	33	38	43	49	56
$P = .005$	$P = .01$	23	28	32	37	43	49

TABLA 9 (Continuación)

Unilateral	*Bilateral*	$n = 23$	$n = 24$	$n = 25$	$n = 26$	$n = 27$	$n = 28$
$P = .05$	$P = .10$	83	92	101	110	120	130
$P = .025$	$P = .05$	73	81	90	98	107	117
$P = .01$	$P = .02$	62	69	77	85	93	102
$P = .005$	$P = .01$	55	68	68	76	84	92

Unilateral	*Bilateral*	$n = 29$	$n = 30$	$n = 31$	$n = 32$	$n = 33$	$n = 34$
$P = .05$	$P = .10$	141	152	163	175	188	201
$P = .025$	$P = .05$	127	137	148	159	171	183
$P = .01$	$P = .02$	111	120	130	141	151	162
$P = .005$	$P = .01$	100	109	118	128	138	149

Unilateral	*Bilateral*	$n = 35$	$n = 36$	$n = 37$	$n = 38$	$n = 39$
$P = .05$	$P = .10$	214	228	242	256	271
$P = .025$	$P = .05$	195	208	222	235	250
$P = .01$	$P = .02$	174	186	198	211	224
$P = .005$	$P = .01$	160	171	183	195	208

Unilateral	*Bilateral*	$n = 40$	$n = 41$	$n = 42$	$n = 43$	$n = 44$	$n = 45$
$P = .05$	$P = .10$	287	303	319	336	353	371
$P = .025$	$P = .05$	264	279	295	311	327	344
$P = .01$	$P = .02$	238	252	267	281	297	313
$P = .005$	$P = .01$	221	234	248	262	277	292

Unilateral	*Bilateral*	$n = 46$	$n = 47$	$n = 48$	$n = 49$	$n = 50$
$P = .05$	$P = .10$	389	408	427	446	466
$P = .025$	$P = .05$	361	379	397	415	434
$P = .01$	$P = .02$	329	345	362	380	398
$P = .005$	$P = .01$	307	323	339	356	373

De: "Some Rapid Approximate Statistical Procedures" (1964), 28, F. Wilcoxon y R. A. Wilcox. Reproducido con la gentil autorización de R. A. Wilcox y los laboratorios Lederle.

BIBLIOGRAFÍA

Academia de Ciencia de Cuba-URSS (autores varios). *Metodología del conocimiento científico.* Cuba-URSS. Presencia latinoamericana S.A. 1962.

Bunge, Mario. *La investigación científica.* Barcelona. Editorial Ariel, S.A. 1983.

Camacho Rosales Juan. *Manual de uso del programa estadístico SPSS/PC+.* Barcelona. PPU. 1994.

Conover, W. J. *Practical Nonparametric Statistics.* U.S.A. John Wiley & sons Inc. Company. 1971.

Chris Leach. *Fundamentos de estadística,* enfoque no paramétrico para ciencias sociales. México. Editorial Limusa. 1982.

D.C. Baird. *Experimentación.* México. Prentice-Hall Hispanoamericana, S.A. 1991.

Echeverria Samanes Benito. *Estadística aplicada a las ciencias Humana.* Barcelona. Ediciones Daimon. 1982.

Hollander Myles. *Nonparametric Statistical Methods.* U.S.A. John Wiley & sons Inc. Company. 1972.

L. Ruiz-Maya, F. J. Martín Pliego, J. M. Montero, P. Uriz Tomé. *Análisis estadístico de encuestas: datos cualitativos,* Madrid. Editorial AC 1995.

Mendenhall Reinmuth James E. *Estadística: para administración y economia,* México. Grupo Editorial Iberoamérica .1989 .

Mendenhall William, Scheaffer Richard L., Dennis D. Wackerly, *Estadística Matemática con Aplicaciones,* México. Grupo Editorial Iberoamérica .1986 .

Montgomery Douglas C. *Diseño y Análisis de Experimentos.* México. Grupo Editorial Iberoamérica .1991 .

Rojas Soriano Raúl. *Guía para realizar investigaciones sociales.* México. UNAM. 1985.

Rojas Soriano Raúl. *El proceso de la investigación científica.* México. Editorial Trillas. 1990.

Richard P. Runyon, Audrey Haber. *Estadística para las Ciencias Sociales.* U.S.A. Addison-Wesley Iberoamericana, S.A. 1992.

Russell L. Ackoff. *Scientific Method optimizing applied research decisions.* U.S.A. John Wiley & sons Inc. Company. 1962.

Siegel, Sidney. *Estadística no paramétrica aplicada a las Ciencias Sociales.* México. Editorial Trillas. 1980.

Stevens. S. S. *Matemáticas y medición* – Titulo original "Mathematics, Measurement and Psychophysics, incluido en el libro Handbook of Experimental Psychology, Wiley. Nueva York. 1951. Traducido por Martín Sagrera.

Stevens. S. S. *La medición y el hombre.* Suplemento del seminario de problemas científicos y filosóficos. UNAM. 1959.

SPSS® Base 8.0 for Windows™ User's Guide.

Varios, Consejo Nacional de Profesores de Matemáticas. *Medida – Serie Temas de Matemática.* México. Grupo Editorial Trillas.1976.

Wayne W. Daniel. *Applied Nonparametric Statistics.* U.S.A. Houghton Mifflin Company. Primera edición 1978.

Wayne W. Daniel. *Applied Nonparametric Statistics.* U.S.A. PWS-KENT Publishing Company. 1990.