12

1 ejemplar
N. 24

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE CIENCIAS

ESTADISTICA NO PARAMETRICA:
UN ENFOQUE INTUITIVO.

T E S I S

QUE PARA OBTENER EL TITULO DE

A C T U A R I O

P R E S E N T A

RAUL RUEDA DIAZ DEL CAMPO

JUNIO 79

I N D I C E

P R Ó L O G O

Este trabajo fué llevado a cabo con un doble propósito:

Proporcionar material didáctico para la enseñanza de Estadística No Paramétrica en los cursos que se imparten en esta Facultad (principalmente en Estadística II) y realizar un texto accesible a personas de otras disciplinas, que no cuenten con muchos conocimientos en el área. Debido a las carreras que existen en esta Facultad la orientación de este trabajo se inclina más a Biología que a otra ciencia.

Debe añadirse también que, debido a los dos puntos anteriores, este trabajo forma parte de uno de los proyectos del Laboratorio de Estadística de la Facultad: el de formar un acervo de textos en estadística que cubran las diferentes ramas de la misma, en un lenguaje accesible a la mayoría de la gente, y que sirvan para introducirlos en la materia y al mismo tiempo de consulta.

INTRODUCCIÓN

En muchos experimento (no necesariamente científicos), es deseable contar con técnicas que permitan escoger, de entre varias alternativas, la mejor dentro de los objetivos del experimento. Otras veces, el experimentador desea averiguar si la información con la que trabaja, o las suposiciones en las que su experimento se basa, pueden ser justificadas satisfactoriamente. La estadística tiene una serie de técnicas, que el estadístico puede usar, para estudiar estos problemas. Por ejemplo: En un momento dado puede ser de interés saber cuál de dos artículos elaborados con distintos métodos de fabricación, cumple con ciertas normas de calidad. Para ésto, pueden tomarse artículos elaborados con cada uno de los métodos de fabricación, para comparar su calidad y llegar a una conclusión. El patrón de producción en cada uno de los métodos, puede ser conocido o desconocido. Si el patrón es conocido, puede ser caracterizado con ciertas medidas (parámetros) como son: la producción media, la variabilidad en la producción en el transcurso de una semana, etc. En estos casos, una comparación entre esas medidas, pueden llevar a tomar una decisión. Sin embargo, en los casos en que el patrón de comportamiento del fenómeno bajo estudio sea desconocido, una comparación como la antes mencionada no puede hacerse. Esto hace necesario, la utilización de técnicas estadísticas que no necesiten de esta suposición para poder ser aplicadas.

Cuando la distribución de probabilidades del fenómeno (patrón de comportamiento) sea conocida, las técnicas estadísticas que se aplican, son las llamadas técnicas paramétricas; mientras que para los casos en que dicha distribución se desconozca, las técnicas que se usan son las técnicas no paramétricas (conocidas por varios autores, como técnicas de distribución libre, por razones obvias). Este trabajo, está enfocado al tratamiento de estas últimas técnicas.

Las pruebas no paramétricas son relatívamente nuevas (si se exceptúan las pruebas de Bondad de Ajuste y la prueba de Signos) y han surgido de la necesidad de tener procedimientos que sirvan para probar hipótesis* sobre poblaciones en las que se desconozcan las leyes de probabilidad que las gobierna. Estos procedimientos pueden ayudar a resolver los siguientes problemas:

i) Verificar si cierta población tiene asociada una distribución de probabilidades específica,

ii) Averiguar si un conjunto de observaciones forman una muestra aleatoria de una población,

iii) Probar si dos o más poblaciones o tratamientos, pueden considerarse estadísticamente iguales, y

iv) Medir la correlación entre dos muestras y saber si las po-

*Todos los términos estadísticos serán explicados en el Capítulo 1.

blaciones de donde provienen las muestras, están correlacionadas.

Una de las ventajas de estas pruebas, es la facilidad con la que puede trabajarse, pues no están basadas en expresiones matemáticas complejas que pueden ser engorrosas. Por otro lado, la mayoría de las pruebas se basa en los rangos de las observaciones y varias de ellas requieren únicamente que las variables estén referidas en una escala nominal*, por lo que el rango de aplicabilidad es muy grande.

En este trabajo se desarrollan 27 pruebas no paramétricas, que cubren los cuatro puntos anteriores.

En el Capítulo I se tratan los conceptos, estadísticos básicos, que se utilizarán a lo largo del trabajo. Sirve también como una introducción al problema de prueba de hipótesis. Está escrito de tal forma que no es necesario un conocimiento avanzado de la estadística.

En el Capítulo II se presentan las pruebas de una muestra, que se utilizan para verificar aleatoriedad, así como también, para averiguar si el modelo supuesto para la población es válido.

* Las variables pueden ser clasificadas, de acuerdo a la escala en la que están medidas. Estas escalas son;

- Nominal - Cuando no existe o no puede introducirse un orden natural.
- Ordinal - Cuando puede introducirse un orden en las observaciones, o existe en forma natural.
- De Intervalos - Cuando se conoce la distancia que separa a dos observaciones, independientemente de la unidad de medida y del lugar en donde el cero es colocado.
- De Proporciones - Cuando se tiene una escala de intervalos y se tiene un cero u origen verdadero.

El siguiente Capítulo, Capítulo III, contiene los procedimientos útiles para probar si dos poblaciones pueden considerarse iguales, o bien, si dos muestras proceden de poblaciones con la misma mediana.

La generalización de este problema al caso de k poblaciones, es tratado en el Capítulo IV.

Por último, el Capítulo V, se presentan las medidas de correlación y las pruebas de hipótesis asociadas.

Al inicio de cada exposición, se plantean diferentes situaciones en que la prueba puede ser utilizada, sirviendo también como introducción al problema. A partir de esto, las hipótesis a verificar son planteadas y la estadística de prueba es generada en forma intuitiva, evitando expresiones matemáticas superfluas, y manteniendo estas últimas a un nivel elemental. Al final de cada prueba, ésta se ilustra con ejemplos numéricos resueltos, que al mismo tiempo, muestran el uso de las tablas anexadas al final del trabajo.

Este trabajo no pretende ser exhaustivo, sino una guía en el estudio de la estadística no paramétrica.

CAPITULO I

CONCEPTOS BÁSICOS

La curiosidad innata del hombre le ha llevado a tratar de explicarse los fénomenos que le reodean, como por ejemplo la lluvia y el universo; asimismo desea confirmar sus conocimientos sobre aspectos sociales y científicos como pueden ser las leyes de Newton o el nivel de vida de cierta comunidad. Esta curiosidad lleva también a buscar nuevas técnicas que le permitan ciertas condiciones de vida o simplemente mejorar las ya existentes. Es decir, el hombre busca constantemente explicarse cosas o confirmarlas, o bien, inventar algunas y mejorar otras. Para cualquiera de estas cuatro actividades, necesita experimentar, o sea, observar, simular o interactuar con los fenómenos para descubrir, comprobar o determinar algo.

Para ilustrar las ideas antes expuestas, se dan los siguientes ejemplos:

i) La Dirección General de Estadística realizó un estudio con el objeto de conocer las razones por las que los habitantes de la frontera norte del país, recurren a la compra de artículos extranjeros. Para llevar a cabo dicho estudio, personal de la DGE observó durante dos semanas, a los mexicanos que regresaban al país por la frontera norte. En base a estas observaciones, se seleccionaban a algunas personas y se les preguntaba primero si eran residentes de la frontera, y si esto era afirmativo, se les preguntaba acerca de las actividades realizadas en elvecino país y si habían hecho compras, cuales eran los motivos de dichas compras.

En este ejemplo, el experimento consistió en observar el paso de gente en la frontera, y con eso determinar a qué gente se le preguntaba y qué preguntas eran las que deberían hacerse.

ii) Para confirmar que cierta velocidad del viento no es característica particular de zonas desérticas, se han hecho observaciones en zonas no desérticas y se han comprobado, mediante esas observaciones, que la velocidad del viento es similar en ambas zonas. El experimento consistió entonces, en determinar que zonas desérticas y no desérticas deberían ser observadas para medir la velocidad con la que el viento viajaba.

iii) Cuando se busca un nuevo tratamiento médico para curar cierta enfermedad o medicamentos que la alivien, son necesarias ciertas ciertas pruebas de laboratorio para saber si dichos medicamentos o el tratamiento son efectivos.

En este caso la experimentación consistirá en elegir adecuadamente los seres vivos (generalmente animales como las ratas, chimpancés, etc) a los que se les aplicarán los medicamentos o tratamientos, verificar las reacciones y determinar la efectividad de ellos.

iv) En la investigación agrícola, para mejorar la producción de cierto tipo de cultivo, se busca el fertilizante que ayude a tales própositos. Aquí, la experimentación va a consistir en seleccionar el terreno de cultivo a utilizar para probar los fertilizantes, que tipos de éstos últimos serán probados y también, como serán aplicados y en que cantidades.

Con estos ejemplos puede observarse que la experimentación se utiliza, tanto al estudiar fenómenos cuyos resultados no son totalmente predecibles o que un momento dado no se puede determinar lo que sucederá; a los cuales se les llama fenómenos aleatorios; como -

también, al estudiar fenómenos deterministas.

Como se ha visto, la experimentación juega un papel básico en casi toda actividad humana, por lo que es importante que la persona que vaya a realizar un experimento, tenga cuidado al planearlo. Para ésto, Méndez (1976) sugiere, como una guía para llevar a cabo un experimento, los siguientes pasos;

i) Se deben establecer objetivos, es decir qué se pretende estudiar con el experimento y que tipo de preguntas se desean resolver a partir de él.

ii) Definir las variables de interés involucradas en el experimento, que estarán sujetas a control o en observación, y que ayudarán a cumplir los objetivos.

iii) Decidir que tipos de observaciones se van hacaer y cuantas de ellas.

iv) Determinar las suposiciones sobre las cuales el experimento se basará y se realizará.

v) Efectuar el experimento y recolectar información.

vi) Analizar la información.

vii) Obtener todas las conclusiones posibles y las probables conjeturas que se quieran comprobar mediante otro experimento.

Una de las formas para llevar a cabo el paso seis, es mediante el análisis estadístico, y como la finalidad de estas notas es presentar ciertos métodos estadísticos es oportuno introducir un ejemplo que al mismo tiempo que ilustre los pasos antes mencionados, ayude a establecer los términos estadísticos necesarios.

En Baltimore existe un hospital especializado en tratar clínicamente a las personas alérgicas, en particular, al veneno de abeja. Se ha probado que si una persona es alérgica al veneno, éste le puede producir entre otras cosas un shock anafiláctico (reacción alérgica) el cual a su vez puede producir la muerte de la persona. Debido a esto, el hospital tiene como objetivo el definir que tipo de tratamiento médico se le debe dar a una persona alérgica, para contrarrestar la reacción al veneno. Para esto, primero debe determinarse si una persona es alérgica al veneno, para lo cual se hace lo siguiente:

Se sabe que el veneno de abeja, está compuesto por un cierto número de proteínas, cuatro de la cuales están en mayor proporción con respecto a las demás. Se tenía en la clínica cierta cantidad de veneno, parte de la cual se purificó para aislar cada una de las cuatro proteínas principales, pues se deseaba verificar si estas proteínas podían ser consideradas como alergenos (substancias que producen alergia).

Para determinar si una proteina es alergena o no, se introduce en la sangre de la persona y se cuantifica la cantidad de histamina liberada por los leucocitos; si es mucha la histamina liberada, la proteína se considera como alergeno; si es poca, no se considera como tal. Por desconocer la reacción del paciente al alergeno, este tipo de pruebas se realiza en vitro, es decir, se toma una muestra de sangre de la persona y a esta muestra donde el alergeno se aplica.

Cada proteina es probada por separado para:

i) Evitar que el fenómeno sea enmascarado al usar todo el veneno, pues éste contiene histamina, la cual puede confundirse con una reacción alérgica; y

ii) Establecer la potencia de alergenos en la persona (en función de liberación de histamina) y tratar de desensibilizarla sólo que aquellos que le causen daño.

Debido a que no toda la gente está dispuesta a servir como "conejillo de indias", para poder realizar el experimento, se toman voluntarios, a los cuales se les ha diagnosticado clínicamente como alérgicos. A estas personas, se les toman varias muestras de sangre, y cada una de ellas se les añade cada uno de los 4 alergenos en varias dosis, así como también cantidades de veneno no purificadas para medir la potencia del alergeno con respecto al veneno. Finalmente se observa y cuantifica la histamina liberada por cada persona en cada dosis y se obtienen conclusiones. Te Piao King et al (1976,) realizaron este experimento con siete voluntarios y las preguntas y conclusiones a las que se llegaron son las siguientes (dadas aquí en forma somera):

Sólo en uno de los alergenos probados, podía ser considerado como potente, es decir, las siete personas reaccionaron mucho a él, y la pregunta inmediata es, ¿puede generalizarse este resultado?, esto es, ¿esta substancia es siempre alergena para toda persona alérgica?. Algunas personas reaccionaron a sólo una de las proteínas y otra proteína no produjo reacción alguna. ¿La primera proteína puede considerarse como alergeno parcial? esto es ¿sólo afecta a cierto tipo de personas?. La proteína que no produjo reacción, ¿puede no considerarse como alergeno, para todas las personas?.

Para poder contestar a estas preguntas, se hace necesario realizar otro experimento que tenga como objetivos las respuestas a estas cuestiones.

En el experimento tratado, el problema principal, era determinar cuales de las cuatro proteínas eran substancias alergenas, pues el tratamiento médico para tratar a la persona, depende del alergeno. Por otro lado, sólo se trabajó con estas cuatro proteínas, pues las demas se encuentran en baja proporción y no producen reacción a la persona.

Para averiguar cuales proteinas eran alergenas, se utilizó un grupo de personas que eran alérgicas al veneno de abeja. Todo el conjunto de personas alérgicas al veneno, define la población bajo estudio; las personas cuya sangre fue analizada, constituye a una muestra de dicha población. En general, la población bajo estudio es el conjunto de personas, objetos, etc, sobre el que un investigador está interesado; y una muestra es un subconjunto de ella, a partir de la cual se desean hacer conclusiones sobre toda la población(como sería contestar las preguntas planteadas al final del ejemplo, y a lo que se le llama: hacer inferencias a la población a partir de la muestra).

El suministro de proteinas para determinar si son substancias alergenas, representa la aplicación de tratamiento; uno por cada dosis de cada alergeno. Se puede decir que en general, un estímulo aplicado a un sujeto que pueda producir un cambio en alguna característica del sujeto , puede considerarse como un tratamiento.

Ahora bien, se ha hablado de muestra,la pregunta que hay que contestar en este momento es: ¿Es necesario tomar una muestra, para estudiar a toda la población?. La respuesta es: no siempre.

En algunos casos, es posible trabajar con toda la población, estudiarla, experimentar con ella y concluir sobre sus características.

En estos casos, que pueden considerarse como ideales, no se hacen inferencias, pues se tiene todo el material, por así decirlo, para estudiarlo. Sin embargo, como se vió en el ejemplo no simpre es posible trabajar con toda la población, existen ocasiones, en los que hay que restringirse a estudiar sólo parte de ella, como fué el estudiar y experimentar sólo con voluntarios, o bien, el tomar muestras de sangre para conocer las reacciones de las personas, a los alergenos.

Esto es, cuando no se puede trabajar con toda la población, ya sea porque esto es imposible debido a las características del fenómeno, o bien, porque la población es muy grande y llevaría tiempo estudiarla, lo cual acarrea, que al final del estudio, los resultados obtenidos, probablemente no reflejen la situación, debido a que perdieron actualidad; o simplemente porque el fenómeno bajo estudio, es un fenómeno no renovable, como puede ser la duración de un foco, o la polarización de un átomo; se hace necesario tomar una muestra de la población.

En vista de la necesidad de tomar muestras en algunas ocasiones, éstas deben ser tales que representen en una forma objetiva a toda la población y no tener desviaciones sistemáticas a ciertas características. Por ejemplo, si se estudia el nivel económico de los habitantes de una ciudad, la muestra debe contener a personas que pertenezcan a las diferentes clases sociales y no únicamente a los que tengan ingresos superiores a cierta cantidad.

Una forma de obtener este tipo de muestras, es mediante procesos aleatorios y a las muestras así obtenidas, se les llaman muestras aleatorias. Estos procesos, suponen que todo elemento de la población tiene una cierta probabilidad de pertenecer a la muestra. En estas notas se supondrá que las muestras que se utilizan, son muestras aleatorias, o bien, pueden considerarse como tales, por lo que los métodos estadísticos que se presentan, serán aplicables para el estudio y análisis de los problemas.

Dentro de los objetivos del análisis estadístico, está el de la construcción de modelos, el de estimar algunas características de la población, o bien verificar si ciertas conjeturas surgidas en investigaciones o experimentos anteriores, o de la experiencia en situaciones que de una u otra forma se relacionan con el problema que se estudia, son verdaderas. Para los primeros dos puntos se recomiendan las obras de Cox(1958), Nelson(1973), Ross(1972), Hoel

(1962), Hoog&Graig (1970), Lindgren (1968) y Valencia et al (1978).

El último punto es conocido en estadística como el problema de probar hipótesis (Prueba de Hipótesis). En general, una hipótesis estadística, es una afirmación acerca de la población bajo estudio, y el problema es encontrar una regla que lleve a validar (o invalidar) dicha afirmación, en base a la información que arroje una muestra obtenida en forma aleatoria. La hipótesis puede ser rechazada lo que significa que la información que dá la muestra, arroja dudas suficientes para decir con cierta confianza, que la hipótesis es falsa; o bien, puede ser "aceptada", lo cual significa que la hipótesis no se rechaza, esto es, no existen dudas suficientes como para decir que la hipótesis es falsa. (En vista de esto, cuando se dice que una hipótesis se "acepta", lo que significa es que no se rechaza). Al conjunto de puntos que lleva al rechazo de la hipótesis se llama región crítica. Como la regla de decisión depende de la muestra, es necesario encontrar una estadística (una función que dependa únicamente de la muestra y no de variables o constantes desconocidas) a la que se le llamará <u>Estadística de Prueba</u>, que lleve a tormar una decisión.

En general, un procedimiento de prueba de hipótesis puede resumirse en los siguientes pasos:

i) Establecer las hipótesis en términos de la población bajo estudio.

ii) Definir una regla, en términos de la estadística de prueba, que sirva para decidir si la hipótesis se rechaza o no se rechaza.

iv) En base a la información que arroje la muestra, la estadística de prueba se evalúa y se toma una decisión.

Debe notarse, que cuando se tiene una hipótesis a verificar implícitamente se tiene una alternativa que es su negación. Esto es, si-en base a la evidencia que arroja la muestra, la hipótesis en consideración, se rechaza, significa que se "acepta" la negación; y viceversa. (La desición de "no rechazar la hipótesis" no és equivalente a aceptarla, lo que debe de entenderse es que la evidencia de la muestra no demuestra que sea falsa). A la primera se le llamará hipótesis nula y a la segunda hipótesis alternativa, y deberá tenerse cuidado al definir quien es quién, para no crear ambigüedades en el momento de tomar una decisión.

Ahora bien, dada la hipótesis, que se le llamará nula de aquí en adelante y se notará con Ho, puede darse dos situaciones después de tomar una decisión:

i) Rechazar la hipótesis Ho y que en efecto sea cierta.

y ii) No rechazar la hipótesis Ho y tenerse que sea falsa.

En cualquiera de los dos casos, la decisión tomada es errónea. Cuando se dá el primer caso se dice que se ha incurrido en un error del tipo I y cuando sucede el otro, se dice que es un error del tipo II. Debido a que la decisión está basada en la muestra, cada error tiene asociada una probabilidad de ser cometido, teniéndose que la probabilidad de cometer el error tipo I es igual a la probabilidad de rechazar Ho (lo que es equivalente a que la muestra pertenezca a la región crítica), siendo que es cierta y la probabilidad de cometer el error tipo II es igual a la de no rechazar Ho, dado que Ha es la verdadera. Para encontrar la primera, se supone que Ho es cierta y se calcula la probabilidad de encontrarse en la región crítica. Si esta hipótesis es simple (lo que significa que la hipótesis especifica completamente a la población) esta probabilidad es única. Sin embargo, si Ho es compuesta (es decir, no especifica completamente a la población), no es única por lo que se calcula la máxima probabilidad de rechazar Ho, considerando el máximo sobre todas las distribuciones de probabilidad que Ho especifica. A esta probabilidad, que generalmente se denota con α, se le llama también nivel de significancia de la prueba. El segundo error, denotado comúnmente con β, se calcula en forma similar.

Puesto que el valor de α, es importante, pues define la región de rechazo de la hipótesis nula, deberá determinarse de acuerdo a la importancia o al significado práctico del estudio del investigador.

Esto es, no existe una regla general para asignar un valor a α, sino que depende de quien hace el estudio, y del estudio mismo.

Algunas veces no es necesario definir una región crítica para tomar una decisión con respecto a las hipótesis bajo consideración. Si con la información que la muestra arroja, la estadística de prueba toma un valor determinado ζ, y se supone que el experimento se repite, puede calcularse la probabilidad de observar un resultado tanto o más extremo que ζ, suponiendo que la hipótesis nula es correcta. Si esta probabilidad, llamada nivel de significancia descriptivo, es menor que un cierto nivel de significancia fijado anteriormente para realizar la prueba, la hipótesis nula se rechaza. En caso contrario, no se rechaza.

Se supondrá que con estos conceptos, estas notas, pueden ser entendidas, sin embargo, una presentación más amplia y muy sencilla se dá en Valencia et al (1978), por lo que se recomienda que de ser posible, se lea antes de estas notas.

Para finalizar este Capítulo, es importante mencionar que cuando se diseñe un experimento, se busque el que sea "bueno" en cierto sentido. Cox(1958) propone que un experimento es "bueno" cuando se cumple lo siguiente:

i) Ausencia de error sistemático. Es decir que los elementos de una muestra no difieran en forma sistemática de los elementos de otra muestra.

ii) <u>Precisión</u>. Si existen errores aleatorios al estimar los resultados, o en las conclusiones, estos deben ser lo más pequeños posibles y con el menor número de observaciones que se pueda.

iii) Las conclusiones deben tenr un amplio grado de validez.

iv) El experimento debe ser lo más simple en análisis y diseño.

y v) El análisis estadístico debe ser posible sin suposiciones artificiales.

En esta parte el estadístico puede hacer tres contribuciones:

i) Aconsejar ó asesorar sobre los principios generales en los que se realiza el experimento.

ii) Construir diseños apropiados para el experimento, y

iii) Analizar o instruir al investigador en como analizar los datos experimentales.

CAPITULO II

PRUEBAS DE UNA MUESTRA

PRUEBA BINOMIAL.

Algunas veces, sucede que, dada una población, lo que interesa es conocer la proporción de elementos de la población que posee cierta característica, o bien, saber si cierto tratamiento afecta a no a la población.

En general, dada una población en la que cada individuo puede ser clasificado en una de dos categorías excluyentes y exahustivas, es de interés determinar si la proporción de individuos que pertenecen a una de las categoría es mayor, menor o igual que cierto número.

Por ejemplo: Cierta industria farmacéutica quiere introducir al mercado un nuevo anlgésico que cura el dolor de cabeza, pero para poder hacerlo, necesita garantizar que el analgésico tiene una efectividad del 70% para curar el dolor de cabeza, es decir, para que el analgésico pueda ser lanzado al mercado es necesario que la probabilidad de que cure a una persona sea 0.70.

Supóngase entonces que a 100 personas se les administró el analgésico, y se quiere comprobar que cura en un 70% de los casos el dolor de cabeza; las 100 personas se clasificarán en dos categorías: en la primera se anotará a las personas que el analgésico si curó, y en la otra a las que no curó: (aquí se supone que no hay más que estas dos opciones). La hipótesis a probar sería $p=0.7$ contra $p< 0.7$, donde p es la probabilidad de que una persona sea curada por el analgésico.

Un ejemplo en el cual se quiere averiguar si cierta proporción de la población tiene determinada característica es el siguiente: Suponga que un grupo de biólogos intuye que el 40% de los peces que existen en un cierto lago cercano a la localidad donde ellos trabajan, poseen sólo dos aletas, en lugar de tres o más. Para poderlo confirmar se tomó una muestra de tamaño n: $X_1,X_2,\ldots,X_n$; en donde-cada observación sería clasificada de acuerdo al siguiente criterio: El pez tiene dos aletas, o no. De acuerdo a la proporción de peces que tiene dos aletas del total de la muestra, se aceptaría Ho: p=0.40 o no.

Un último ejemplo es el que se dá a continuación: La mayoría de la gente cree que la probabilidad de que al nacer un bebé, éste sea niño, es la misma que la de que sea niña, sin importar las características de la madre. Para comprobar esto, se tomaron los registros de nacimientos de un hospital, y se llevó a cabo la prueba para verificar que $p=\frac{1}{2}$, donde p denota la probabilidad de que sea niño.

Nótese que en estos tres ejemplos la escala en la cual las observaciones son medidas, es nominal, lo cual no es estrictamente necesario (por ejemplo, podría probarse el que la proporción de niños de edad 12, cuya altura fluctúa entre 1.50 y 1.60 es p), además todas las observaciones pueden ser clasificadas en una de dos categorías ajenas donde las observaciones son independientes unas a otras.

Sin pérdida de generalidad se pueden "etiquetar" las categorías con "0" y "1", y suponer que la proporción de individuos de la población que pertenecen a la categoría "0" es P, y por tanto los que pertenecen a la categoría "1" será 1-P. Al tomar una muestra aleatoria de tamaño n: $X_1, X_2, \ldots, X_n$; se tendrá que r elementos de la muestra pueden ser clasificados en la categoría "0" y n-r en la categoría "1". Sin embargo, debido a que se toma una muestra, la proporción de elementos de la muestra que pertenezcan a la clase "0", no necesariamente será la misma que la proporción de elementos de la población que pertenece a la clase "0". Debido a ésto, puede incurrirse en el error de rechazar la hipótesis a verificar, Ho (Ho: $p=p^*$, donde p^* es la proporción de objetos que pertenecen a la categoría 0), cuando es verdadera, que es a lo que se llama nivel de significancia. Si la probabilidad de aceptar Ho dado que es cierta es mayor que α el nivel de significancia, se tiene que la p^* supuesta es consistente con la hipótesis y por tanto se aceptaría.

Debido a que la información sólo puede ser clasificada en dos categorías, la distribución adecuada para calcular las probabilidades anteriores, es la binomial. Entonces, para poder llegar a una decisión sobre aceptar o no la hipótesis da verificar considérese α, el nivel de significancia y sean $\alpha_1, \alpha_2 \in [0,1]$ t.q. $\alpha = \alpha_1 + \alpha_2$ y por último supóngase que la hipótesis es Ho: $p=p^*$ y el tamaño de muestra es n. Entonces, búsquese t_1 y $t_2 \in \mathbb{R}$ tales que $P(y < t_1) = \alpha_1$ y $P(y < t_2) = 1-\alpha_2$ donde y es una variable aleatoria binomial con parámetros p^* y n. Además los valores α_1 y α_2 deben ser lo más parecidos(*). Entonces, de acuerdo a lo que se dijo anteriormente con respecto a la aceptación de Ho, ésta se aceptará si $r < t_2$ o si $r > t_1$, de otra forma, no se aceptará Ho, es decir si $r > t_2$ o si $r < t_1$.

Puede probarse también las hipótesis Ho: $p < p^*$ ó Ho $p > p_1^*$ el procedimiento de hacerlo es el mismo, lo único que cambia es la región de rechazo (Conover 1971).

EJEMPLO 1 Bajo las leyes de la herencia Mendel, al realizar un injerto entre plantas de dos genotipos particulares, se espera que produzcan progenie de la cual 1/4 de ella fuese "enana" y 3/4 de ella fuese "alta". En un experimento para determinar si la hipótesis de las leyes de Mendel son razonables, se obtuvo progenie de una cruza de dos plantas, resultando 243 plantas "enanas" y 682 plantas "altas".

Se quiere probar entonces, que $p^* = \frac{3}{4}$. Para esto se considerará que una planta pertenece a la categoría 0 si es "alta" y si es "enana" se clasificará en la categoría 1. Como los elementos de la muestra se clasifican en dos categorías, la prueba binomial es adecuada para verificar la hipótesis:

(*) Generalmente $\alpha_1 \neq \alpha_2$, pues se trata de una distribución discreta, en caso de que se quiera que $\alpha_1 = \alpha_2$, deberá usarse una prueba aleatorizada. Estas pruebas pueden verse en Lindgren (1968).

H_o: $p^* = \frac{3}{4}$ teniéndose como alternativa natural.

H_a: $P^* \neq \frac{3}{4}$

Se tiene que el total de plantas clasificadas (cada planta representa una observación) es 925. Supóngase que el nivel de significancia es $\alpha = 0.05$. De acuerdo a lo desarrollado, deben buscarse t_1 y t_2 tales que $P(y < t_1) = \alpha_1$ y $P(y < t_2) = 1-\alpha_2$, con $\alpha = \alpha_1 + \alpha_2$ y y una variable binomial con parámetros $n=925$ y $p=3/4$.

Como n es muy grande la tabla 3 del apéndice A, dice que los cuantiles de la binomial, deben aproximarse mediante la distribución normal, usando:

$$w_r = np + y_r \sqrt{np(1-p)}$$

donde w_r es el r-ésimo cuantil de una distribución normal (0,1) y y_r el correspondiente a una binomial con parámetros n y p.

La hipótesis se rechaza si $r > t_2$ o si $r < t_1$ donde

$$t_1 = np^* + w_{0.025} \sqrt{np^*(1-p^*)} \text{ y } t_2 = np^* + w_{0.975} \sqrt{np^*(1-p^*)}$$

Substituyendo los valores, se tiene que $t_1 = 667.93$ y $t_2 = 719.56$.

Ahora es el número de objetos clasificados en la categoría 0. En el ejemplo, $r=682$. Como $r > t_1$ y $r < t_2$ la hipótesis no se rechaza, por lo que se concluye que los datos concuerdan con las leyes de herencia de Mendel.

EJEMPLO 2. Un investigador preguntó a 400 personas elegidas al azar pertenecientes a una población especifica, si favorecían al candidato A o al B; 220 votaron a favor de A y 180 a favor de B. Si el nivel de significancia se fija en 0.05, ¿puede decirse que la opinión de la población está igualmente dividida?.

Esto es equivalente a plantear las siguientes hipótesis:

$$Ho: p^* = \frac{1}{2} \quad vs \quad Ha: p^* \neq \frac{1}{2}$$

donde p* es la proporción de individuos de la población cuya opinión favorece al candidato A, y en donde la categoría O, está definida por las gentes que están a favor del candidato A. (Podría considerarse el análogo, pero por la simetría de la prueba es lo mismo). Al igual que el ejemplo pasado, el tamaño de muestra es grande, por lo que se usará la aproximación normal.

Entonces, como $\alpha = 0.05$, se necesita encontrar el cuantil 0.025 y el cuantil 0.975 de una variable normal con media cero y varianza unitaria. Para ésto, se utiliza la tabla 3 del apéndice, teniéndose que:

$$w_{0.025} = -1.96 \quad y \quad w_{0.975} = 1.96$$

Por lo que los cuantiles correspondientes de una binomial con parámetros $n=400$ y $p=\frac{1}{2}$, son:

$$t_1 = 180.40 \quad y \quad t_2 = 219.6$$

Como $r=180$, se tiene que $r < t_1$, por lo que la hipótesis Ho debe ser rechazada.

Como observación debe notarse lo siguiente: Debido a que r y t_1 difieren muy poco, puede tomarse más información para tomar una decisión.

EJEMPLO 3. En un estudio de los efectos de la tensión, un experimentador enseñó a 18 estudiantes, dos métodos diferentes para hacer un nudo. La mitad de los sujetos (seleccionados al azar de los 18) aprendió antes el método A y la otra mitad, antes el método B. Más tarde se pidió a cada sujeto que hiciera un nuevo nudo. La predicción era que la tensión induciría regresión, es decir, que los sujetos retrocederían al método aprendido antes. Cada sujeto fue clasificado de acuerdo a que usara el método aprendido antes o el aprendido después, al pedírsele atar el nudo estando bajo tensión. Suponiendo que $\alpha = 0.01$ y teniéndose que todas las personas excepto dos, usaron el método aprendido antes, ¿puede sostenerse la predicción de que la tensión induce regresión?.

Si p es la probabilidad de usar bajo tensión el método aprendido antes (y por tanto 1-p, la probabilidad de usar bajo tensión el método aprendido después), entonces la hipótesis nula puede ser escrita como:

$$Ho: p = \tfrac{1}{2} \quad \text{teniéndose que } Ha: p \neq \tfrac{1}{2} .$$

Supóngase $\alpha = 0.01$ y que la clasificación 0 es: "usar el método aprendido antes", por lo que la clasificación 1 es: "usar el método aprendido después".

De acuerdo al enunciado del problema, se puede construir la siguiente tabla:

	METODO ESCOGIDO	
	Método aprendido antes	Método aprendido después
Frecuencia.	16	2.

En este caso $n < 20$, por lo que la tabla 3 puede ser usada para encontrar la región crítica. Como $n = 18$, $r = 16$ y $p = \frac{1}{2}$ los valores -

t_1 y t_2 t $P\{y \leq t_1\} = \alpha_1$ y $P\{y \leq t_2\} = 1-\alpha_2$ con $\alpha_1 + \alpha_2 = \alpha$, son: $t_1=3$ y $t_2=14$. Como $r > t_2$, la hipótesis nula $Ho: p=\frac{1}{2}$ es rechazada.

BONDAD DE AJUSTE.

En estadística paramétrica, cuando se hace una prueba de hipótesis, se conoce de antemano que familia de distribuciones está asociada a la población y lo único que se prueba es si los parámetros toman un cierto valor específico, o varían en un cierto rango. Así por ejemplo, se tienen pruebas como las siguientes:

i) Se tiene una muestra aleatoria $X_1,...X_n$ de una distribución normal, y se quiere probar $\mu = \mu_0$ $\sigma^2 = \sigma_0^2$ contra $\mu \neq \mu$ $\sigma^2 \neq \sigma_0^2$.

ii) Si $Y_1,...,Y_n$ es una muestra aleatoria de una distribución Poisson, probar: $\lambda > \lambda_0$ contra $\lambda < \lambda_0$.

iii) Si se tiene $X_1,...X_n$ una muestra aleatoria, se puede probar: $f(x;\theta) = f(x;\theta_1)$ contra $f(x;\theta) = f(x;\theta_2)$ donde $f(x;\theta_1)$ y $f(x;\theta_2)$ son dos densidades completamente especificadas.

Sin embargo hay situaciones en las cuales no sólo se desconoce uno o varios parámetros de la distribución de probabilidades, sino aún más, a la distribución misma. Es decir, en ocasiones se quiere checar que la muestra aleatoria provenga de cierta población, por ejemplo, de una población binominal, donde el parámetro p, puede o no ser conocido. O bien probar si las observaciones obtenidas aleatoriamente, provienen de una población Poisson con parámetro λ, o simplemente si la muestra que se tiene, es una muestra de una población Normal, sin importar los valores de μ y σ^2.

Dicho de otra manera, se quiere probar si la muestra aleatoria tiene como función de distribución a $F_X^*(x)$, lo cual se escribirá Ho: $F_X(x) = F_X^*(x)$ $\forall x \in X$. Para este tipo de problemas se ha desarrollado las pruebas llamadas: Pruebas de Bondad de Ajuste, que co mo su nombre lo indica, se utilizan para probar si una cierta distribución de probabilidades pueden ser ajustada a una cierta población o no.

En esta sección, se tratarán tres de estas pruebas, que formalmente pueden clasificarse en dos: la prueba X^2 para Bondad de Ajuste (muchos libros la llaman prueba X^2 de Pearson) y la prueba de Kolmogoroff- Smirnoff. La tercera es la prueba de Lilliefors, que puede considerarse como una variante de la prueba de Kolmogoroff — Smirnoff.

Las tres pruebas se usan para verificar $F_x^*(x) = F_x(x)\ \forall x \varepsilon X$ contra $F_x(x) \neq F_x^*(x)$ para alguna $x \varepsilon X$, en donde $F_x^*(x)$ es la distribución que se cree se ajusta a los datos, la cual puede o no, estar completamente especificada, en el sentido de que no necesariamente se conocen todos los parámetros.

La primera prueba, la prueba X^2, se basa en la distribución X^2 y fue propuesta en el año de 1900 por Karl Pearson. A grandes rasgos, este método considera que la información de la muestra aleatoria puede ser clasificada en un número arbitrario de clases o categorías y a partir de esto define a la estadística de prueba $\sum_i \frac{(E_i - O_i)^2}{E_i}$, en donde el significado y relación de E_i y O_i se verá más adelante.

El nombre de prueba X^2 es debido a que puede demostrarse que si Ho es cierta, entonces $\sum_i \frac{(E_i - O_i)^2}{E_i}$ se distribuye como $X^2_{(c-1)}$, donde c es el número de categorías en las cuales la muestra se clasifica, suponiendo que $F_x^*(x)$ está completamente especificada. En caso de que la función $F_x^*(x)$ esté especificada exceptuando por k parámetros desconocidos, entonces la distribución a la cual se aproxima la estadística de prueba es una $X^2(c-k-1)$. A pesar de que Pearson propuso el que $\sum_{i=1} \frac{(E_i - O_i)^2}{E_i}$ se distribuye como

$X^2(r)$ (r son los grados de libertad según el caso) él no llegó a demostrarlo más que intiutívamente (1900). En el artículo; "X^2-test of goodness of fit" de William G.Cochran se hace una recopilación de lo que hasta 1952 se había estudiado de la prueba, incluyendo la demostración que hizo Pearson.

En 1963 y 1964 M.E. Wise examinó las consecuencias de aproximar $\sum_{i=1}^{n} \frac{(E_i-O_i)^2}{E_i}$ a una $X^2(r-1)$ y demostró que el error es pequeño cuando las E_i son iguales o casi iguales.

Cuando $\frac{E_i}{n} = \frac{1}{r}$, Good et al, (1970), tabularón valores de $\sum_{i=1}^{n} \frac{(E_i-O_i)^2}{E_i}$; Zahan & Roberts (1971) tabularon el caso $\frac{E_i}{n} = \frac{1}{n}$.

La segunda prueba es del tipo de las pruebas Kolmogoroff-Smirnoff; estas pruebas se basan en estadística que involucran tanto distribuciones empíricas, como a la distribución que se cree es la verdadera.

La estadística que se usa para llevar a cabo la prueba es $T = \sup_{x \in X} |F_x^*(x) - Sn(x)|$ donde $F_x^*(x)$ es la distribución que se cree es la verdadera y $Sn(x)$ es la distribución empírica de los datos.

Este tipo de estadística para llevar a cabo estas pruebas, fueron introducidas por A.N.Kolmogoroff en 1933, más tarde Smirnoff la amplió en el sentido en que la estadística podría ser definida como función de la distancia vertical entre dos distribuciones empíricas.

Generalmente a las pruebas que se basan en una estadística como T, se les conoce como pruebas del tipo Kolmogoroff y a las que en la estadística es función de la distancia vertical entre dos distribuciones empíricas, son las llamadas pruebas del tipo Smirnoff.

En 1933 Kolmogoroff encontró la distribución asintótica de T y Smirnoff tabuló algunos valores en 1948. La distribución exacta de T sólo ha sido estudiada para muestras pequeñas por Wald y Wolfowitz en 1939 y valores tabulados para esta distribución las dió Massey en 1950.

Un método para encontrar los valores tabulados para T fué dado por Birnbaum y Tingey en 1951.

La variante de Lilliefors, que se basa en estadísticas del tipo Kolmogoroff- Smirnoff, fué propuesta por Lilliefors en 1967. La única diferencia es que esta prueba sólo sirve para comprobar si una muestra aleatoria dada, proviene de una distribución normal, sin importar los valores de los parámetros, lo que lleva a que la función empírica que se usa, no sea la que se obtiene directamente de las observaciones, sino que primero hay que estandarizarlas y es con estas observaciones ya estanderizadas que se calcula la distribución empírica para llevar a cabo la prueba.

PRUEBA X^2 DE BONDAD DE AJUSTE

A continuación se darán algunos ejemplos en donde podría aplicarse esta prueba:

i) Una compañía de barcos mercantes ha sufrido bajas en sus ganancias, debido a que cierto tipo de sus barcos les han ocurrido accidentes, ya sea porque ha habido mal tiempo, por incendios o por que las máquinas se han descompuesto.

A la compañía le interesa contar con un modelo probabilístico que le pueda servir para estudiar más a fondo el problema y poder llegar a una posible solución. Además la compañía sabe que se han hecho estudios sobre estos problemas y que el modelo que mejor explica el fenómeno de los accidentes, es un modelo de Poisson, es decir la distribución que mejor se ajusta a este tipo de problemas es la distribución Poisson. Sin embargo la compañía desea verificar si el modelo Poisson se puede ajustar a su caso, para lo cual tienen información de lo que ha estado sucediendo con ese tipo de barcos en los últimos seis meses.

ii) Existe en una universidad un grupo de personas interesadas en formar un equipo de basket-bol que los pueda representar en los juegos que se efectúan con otras universidades del país.

Para poder formar el equipo llegan a la conclusión de que necesitarán un total de 25 personas cuya altura no sea menor de 1.85 cm., sin importarles lo demás, pues ellos se encargarán de entrenarles para que resulten buenos jugadores. Sin embargo reconocen que no es fácil buscar entre toda la población estudiantil a estas personas, y aún más no están seguros de encontrar a tantos con esa característica. Para esto deciden calcular la probabilidad de encontrar a esas gentes, basándose en el hecho de que la altura de los estudiantes de la universidad es una variable aleatoria normal de media y varianza conocidas. Pero antes de calcular dicha probabilidad se propone ver si efectivamente las alturas de los muchachos siguen el patrón probabilístico propuesto anteriormente. Para verificar este hecho deciden seleccionar a 50 estudiantes completamente al azar y con esa información realizar una prueba que los lleve a justificar su modelo.

En estos dos ejemplos puede observarse lo siguiente:

i) En los dos casos la información que se va usar para verificar las suposiciones constituye una muestra aleatoria de la población.

ii) Estas muestras pueden clasificarse en un <u>número arbitrario</u> de grupos o categorías o clases: En el primer ejemplo, según el número de accidentes en un período de seis meses y en el segundo según su altura.

iii) Se desea averiguar si las observaciones obtenidas en forma aleatoria, se comportan según el modelo probabilístico propuesto.

Puede notarse que la escala en la cual las observaciones se miden es al menos nominal y que por otro lado puede conocerse la probabilidad de que una observación aleatoria de la variable de interés pertenezca a alguna de las categorías que se han definido, si es que la suposición del modelo, que se cree es el real, es cierta.

Entonces, se tiene una muestra aleatoria $X_1,\dots,X_n$ de una población que se cree tiene asociada una función de distribución $F^*_X(x)$, y lo que se desea es checar si esto es cierto, es decir se quiere verificar si la función de distribución $F_X(x)$ de la población, es $F^*_X(x)$. Esto significa probar que $F_X(x)=F^*_X(x)\ \forall\ x\varepsilon X$ teniéndose como alternativa que $F_X(x)\neq F^*_X(x)$ para al menos una $x\varepsilon X$.

Para poder verificar esta hipótesis, supóngase primero que se tiene una muestra aleatoria $X_1, \ldots, X_n$ de la población, y que estas observaciones se clasifican en forma exahustiva y exclusiva en k- categorías en donde el número k, es arbitrario- esto es, cada observación pertenece a una y sólo una de las k categorías, pero debido a que las observaciones son aleatorias, cada observación tiene una probabilidad p_i de pertenecer a la i-ésima categoría, suponiendo que $F_X^*(x)$ es la distribución de las variables; entonces el número de observaciones que se espera pertenezcan a la i-ésima categoría será np_i.

Esto puede verse de otra forma: para cada categoría se tiene un experimento binomial, es decir, n experimentos bernoulli ya sea que la observación pertenezca o no a la categoría, y como se conoce que el valor esperado de una variable aleatoria binomial es np_i, se tiene entonces el número de observaciones esperadas por categoría.

Por otro lado, ya clasificada la muestra en las diferentes categorías, se sabrá el número de observaciones que pertenecen a cada categoría i, al cual se le denotará por O_i. Si se denota por E_i al número esperado de observaciones por categoría o sea $E_i = np_i$, lo que se desea probar es si O_i y E_i difieren poco. Dicho de otro modo, O_i es el número observado de elementos de la muestra que pertenecen a la i-ésima catagoría o clasificación y E_i es el número esperado de observaciones en la i-ésima categoría si la distribución de la población es $F_X^*(x)$. Si la hipótesis $F_X(x) = F_X^*(x) \; \forall \; x \epsilon \mathfrak{X}$ es cierta, es de esperarse que las diferencias entre O_i y E_i sean pequeñas, para todo $i \epsilon Jn$.

Para que esto suceda, deberá tenerse que $\sum_{i=1}^{n} (O_i - E_i)$ sea pequeña, sin embargo puede ocurrir que algunas de estas diferencias sean negativas y otras positivas de forma tal que al sumarse $(O_i - E_i)$ sobre $i \epsilon Jn$, la suma llegara a anularse o fuese pequeña, lo cual no daría información suficiente para tomar alguna decisión, pués podría darse el caso que a pesar de que las diferencias fuesen grandes, fueran de diferente signo, y así la suma resultará relativamente chica.

Para evitar este problema puede pensarse en tomar a $\sum_{i=1}^{n} |E_i - O_i|$ y así la comparación tendría más significado, pues si la suma es pequeña es debido a que las diferencias entre la O_i y E_i son pequeñas. Pero trabajar con la función valor absoluto no siempre es fácil, por lo que podría pensarse en otra cosa. Si se considera $\sum_{i=1}^{n} (E_i - O_i)^2$, se tendrán diferencias "absolutas" pero elevadas al cuadrado; es decir, se cambia, por así decirlo, la unidad de medida usada, entonces se divide cada $(E_i - O_i)^2$ por E_i, regresando a la unidad usada y el criterio de comparación quedaría entonces como $\sum_{i=1}^{n} \frac{(E_i - O_i)^2}{E_i}$, que es la estadística que Karl Pearsoon propuso en 1900.

Ya con $\sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$ se puede comparar si lo observado y lo esperado difieren mucho.

Puede demostrarse que si Ho es cierta, $\sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$ se distribuye aproximadamente como una $X^2_{(k-1)}$, (donde k es el número de categorías en las que se clasifican las observaciones), si es que $F^*_X(x)$ está totalmente especificada. (En caso de que se estimen r parámetros por el método de máxima verosimiltud, los grados de libertad serán k-r-1).

Si la prueba se hace a un nivel de significancia α, la estadística $\sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$ se compara con el $(1-\alpha)$ cuantil de una variable X^2 con k-1 grados de libertad; si el valor de $\sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$ es mayor que el correspondiente valor de la variable aleatoria, se dirá que la diferencia entre las O_i y E_i es grande y se rechazará la-hipótesis $F_X(x)=F^*_X(x)\ \forall x\epsilon X$, en caso contrario se dirá que la diferencia entre lo esperado y lo observado es pequeña, y se aceptará el hecho de que la distribución de la población es $F^*_X(x)$.

EJEMPLOS

EJEMPLO 1. La teoría de Mendel establece que las probabilidades de clasificaciones de los guisantes son:

a) Redondo y Amarillo : 9/16
b) Arrugado y Amarillo: 3/16
c) Redondo y Verde : 3/16
d) Arrugado y Verde : 1/16

Si de 160 observaciones independientes, las frecuencias observadas de las clasificaciones anteriores son: 86,35, 26 y 13 respectivamente ¿Puede decirse que la población muestreada esta regida por las leyes de Mendel? Usando $\alpha=0.05$.

Puesto que se tiene una escala nominal y la muestra está clasificada en cuatro categorías, la prueba X^2 es adecuada, pues si se quiere probar si la población tiene cierta distribución.(Ho: La distribución de la población es la dada al inicio del ejemplo).

Como se usará la prueba X^2 deberá calcularse $\sum_{i=1}^{k}\frac{(O_i-E_i)^2}{E_i}$, donde k es el número de categorías en que los elementos de la muestra son clasificados; E_i es el número esperado de observaciones que deberían pertenecer a la categoría i si la hipótesis Ho fuese cierta; por último, O_i es el número de observaciones de la muestra que pertenecen a la categoría i.

Entonces, como n=160 se tiene:

$$\sum_{i=1}^{4}\frac{(O_i-E_i)^2}{E_i} = \frac{(86-90)^2}{90} + \frac{(35-30)^2}{30} + \frac{(26-30)^2}{30} + \frac{(13-10)^2}{10}$$

$$\Rightarrow \sum_{i=1}^{4}\frac{(O_i-E_i)^2}{E_i} = 2.43$$

Como k=4, se busca el valor de X^2 con 3 grados de libertad, en la tabla 2 del Apéndice A, teniéndose que $W_{1-\alpha} = W_{0.95} = 7.815$.

Entonces la hipótesis nula no se rechaza, pues $2.43 < 7.815$, por lo que puede decirse que la población muestreada es regida por las leyes de Mendel.

EJEMPLO 2 Un pescador lanzó 30 veces su red a un lago, cada vez que sacaba la red, contaba el número de peces así capturados. Los resultados se muestran en la siguiente tabla:

Números de peces capturados al lanzar una sola vez la red.	Frecuencias
0	12
1	10
2	6
3	0
4	2

Se quiere verificar, que el número de peces así capturados, sigue una distribución Poisson, con $\alpha=0.01$.

Como se trata de ajustar una distribución que no está completamente especificada, se usará la prueba X^2 de Bondad de Ajuste.

Lo primero que se hará es estimar el parámetro de la Poisson, que es λ. Se sabe que el estimador máximo verosímil de es $\overline{X}$ (Valencia et al (1978)). Como los datos están agrupados, la media ($\overline{X}$) -

se calcula de la siguiente forma: $\bar{X}=\sum_{i=1}^{5}\frac{x_i f_i}{30}$, donde X_i es el número de peces capturados y f_i la frecuencia.

Entonces:

$$\bar{X}=\frac{0\times 12+1\times 10+2\times 6+3\times 0+4\times 2}{30}=1$$ y se quiere probar:

$$Ho: F_x(x)=\begin{cases}\frac{e^{-\lambda}\lambda^x}{x!} & \text{con } \lambda=1 \\ & x=0,1,2,\ldots \\ 0 \text{ en otro caso} & \end{cases} \quad \text{vs} \quad Ha: F_x(x)\neq\frac{e^{-\lambda}\lambda^x}{x!}$$

Las categorías son: 0, 1,2,3 y 4. Las probabilidades respectivas bajo Ho son:

x	$e^{-\lambda}\lambda^x/x!$	
0	0.3678	p_0
1	0.3678	p_1
2	0.1839	p_2
3	0.0613	p_3
4	0.0153	p_4

$$\therefore \sum_{i=0}^{4}\frac{(O_i-E_i)^2}{E_i}=\frac{(12-11.03)^2}{11.03}+\frac{(10-11.03)^2}{11.03}+\frac{(6-5.517)^2}{5.517}+\frac{(0-1.839)^2}{1.839}$$

$$+\frac{(2-0.459)^2}{0.459}$$

$$\Rightarrow \sum_{i=0}^{4}\frac{(O_i-E_i)^2}{E_i}=5.3975$$

buscando en la tabla 2, con 3 grados de libertad, pues se estimó un parámetro y se tenían 5 clasificaciones, se tiene que $W_{0.99}=11.34$ por lo que no se rechaza la hipótesis y se concluye que el número de peces capturados al lanzar una sola vez la red, sigue una distribución Poissón.

EJEMPLO 3 Dos tipos de maíz (amarillo y verde) continen genes recisivos. Cuando estos dos tipos de maíz fueron cruzados, la primera generación fué consistente normal. Se obtuvo descendencia de esta primera generación y resultaron cuatro tipos distintos de maíz: normal, amarillo, verde y amarillo verde. En 1200 plantas este proceso produjo los siguientes datos:

Maíz Normal	670
Amarillo	230
Verde	238
Amarillo Verde	62

¿Estos datos son consistentes con la teoría de Mendel?.

De acuerdo al ejemplo 1, la distribución de la segunda generación está dada por la proporción 9:3:3:1, por lo que los valores esperados, para n=1200, son:

$$E_1= 675 \quad E_2=225 \quad E_3=225 \quad y \quad E_4=75$$

$$\text{Así: } \sum_{i=1}^{4} \frac{(O_i-E_i)^2}{E_i} = \frac{(670-675)^2}{675} + \frac{(230-225)^2}{225} + \frac{(238-225)^2}{225} + \frac{(62-72)^2}{75}$$

lo que dá un valor de 3.1525. Usando la tabla 2, se tiene que con tres grados de libertad $W_{1-\alpha}= 7.815$ $(\alpha=0.05)$. Por lo que se concluye que esta población cumple con las leyes de Mendel.

PRUEBA DE KOLMOGOROFF- SMIRNOFF DE BONDAD DE AJUSTE.

En algunas ocasiones, cuando se observan fenómenos aleatorios, los experimentadores desean tener un método que les pueda ayudar a decir si el fenómeno que están observando tiene una cierta regularidad estadística, en el sentido de que si es posible decir que al fenómeno en cuestión se le puede asociar una distribución de probabilidades.

Por ejemplo: Una persona está en las Vegas jugando a los dados, y ha estado apostando durante 300 lanzamientos de éstos (algo así como dos horas) a la combinación (3,6). Esta persona empieza a dudar de la honestidad de los dados pues no ha ganado durante todo ese tiempo, sin embargo el "grupier" le asegura que los dados son honestos, y que por lo tanto la probabilidad de que al lanzar los dados, aparezca la combinación (3,6), es la misma que la de cualquier otra combinación, esto es, que la probabilidad de que al lanzar los dados, resulte (i,j) es $\frac{1}{36}$ $\forall i,j \in J_6$.

En vista de que esta persona y el "grupier" están a punto de darse de golpes, alguien sugiere que se pruebe estadísticamente el que la distribución de probabilidades de las combinaciones que aparecen al lanzar los dados es en efecto una uniforme discreta con parámetro N=36, contra de que no lo es, en cuyo caso alguien tendrá que dar disculpas o regresar dinero.

Otro ejemplo: El número de errores en una página de un libro, frecuentemente es tomado como una variable aleatoria Poisson con un cierto parámetro λ conocido, donde además los errores de una página son independientes de los errores en otras páginas.
¿Cómo comprobar este hecho, para un libro específico?

En los dos ejemplos se tienen completamente especificados los mode los que se suponen son ciertos y además se tiene una serie de observaciones independientes entre sí asociadas a una cierta distribución desconocida, y en los dos casos se quiere probar que el modelo $F^*_X(x)$ es al que realmente pertenece los datos, esto es, probar que $F_X(x)=F^*_X(x)$ $\forall x\varepsilon X$ donde $F_X(x)$ es el modelo verdadero para las observaciones.

Ahora bien, si se tiene una serie de n observaciones independientes $X_1,\ldots X_n$; se puede calcular la función de distribución empírica para las observaciones, denotándola por $S_n(x)$, haciendo $S_n(x)=\frac{k(x)}{n}$, donde $k(x)$ es el número de observaciones que cumplen con ser menor o igual que x, esto es, $k(x)$ es el número de elementos $X_1,\ldots X_n$ que satisfacen que $x_i < x$, con $i\varepsilon Jn$.

Es de esperarse que si $F^*_X(x)$es el modelo verdadero, entonces $F^*_X(x)$ y $S_n(x)$no deben de diferir mucho para los mismos valores de x, es decir si $F^*_X(x)$ y $S_n(x)$ están "suficientemente cerca" $\forall x\varepsilon X$, no se tendrían razones para no aceptar que $F_X(x)$ y $F^*_X(x)$ coinciden.

Habría entonces que calcular las diferencias entre $F^*_X(x)$ y $S_n(x)$, y ver si son pequeñas o no: como sólo basta la diferencia absoluta y no el signo de la diferencia, se considera $|F^*_X(x)-S_n(x)|$.

Ahora bien, para saber si la diferencia es pequeña o no, se tendría que comparar cada diferencia con un cierto número ξ, de tal forma que si la diferencia excede al valor ξ se diría que esa diferencia es significativa, es decir, que la diferencia es tal que llevaría a dudar el que $F^*_X(x)$ y $F_X(x)$ coincidan. En caso de que la diferencia fuese menor que ξ, entonces se tendría una diferencia no significativa.

Sin embargo, debido a que las dos funciones de distribuciones, $F^*_X(x)$ y $S_n(x)$, están definidas en los reales, el estar comparando cada diferencia no es eficiente, pues llevaría mucho tiempo, por no decir que nunca se acabarían de efectuar las comparaciones, y como por otro lado lo que interesa es que todas las diferencias -

no sean significativas, es decir, que $S_n(x)$ y $F_X^*(x)$ se parezcan $\forall x \varepsilon X$, resulta lógico considerar a $\sup |F_X^*(x)-S_n(x)|$ y compararlo con el número ξ; propuesto de ésta forma, si $\sup_{x\varepsilon X}|F_X^*(x)-S_n(x)|$ es pequeño comparado con ξ, se sigue que todas las diferencias $|F_X^*(x)-S_n(x)|$ son pequeñas, y si $\sup_{x\varepsilon X} |F_X^*(x)-S_n(x)|$ es grande, existe al menos una $x\varepsilon X$ tal que la diferencia entre $F_X(x)$ y $S_n(x)$ es significativa y no se aceptará el que $F_X(x)$ y $F_X^*(x)$ coinciden.

A esta prueba de "bondad de ajuste", que se basa en la estadística $\sup_{x\varepsilon X} |F_X^*(x)-S_n(x)|$, es la llamada prueba de Kolmogoroff-Smirnoff, para lo cual existen valores tabulados a distintos niveles de α, si la prueba se realiza a un nivel α, el número ξ que se utiliza en la comparación es el $(1-\alpha)$ cualtil $w_{1-\alpha}$ donde w son los cuantiles de la distribución asintótica de $\sup |F_X^*(x)-S_n(x)|$ cuando la hipotesis es cierta.

Por último nótese que $\sup|F_X^*(x)-S_n(x)|$ significa, que el número con el cual ξ se compara es la máxima distancia vertical entre las gráficas de $F_X^*(x)$ y $S_n(x)$, o de otra forma, calcular para cada $x\varepsilon X$ la distancia vertical entre $F_X^*(x)$ y S_n y tomar la máxima para compararla con ξ.

EJEMPLOS

EJEMPLO 1 Una clínica médica determinó, por larga experiencia, que el número de plaquetas en la sangre de personas del sexo masculino sanos, se distribuye como normal con media 235,000 por mm³ y desviación estandar de 44,600 por mm³. Los conteos de plaquetas obtenidos de 25 pacientes con cancer pulmonar fueron los siguientes (en unidades de 1000 por mm³):

173	189	196	207	215	237	275	282	293	300
305	316	346	382	395	399	401	437	480	504
524	634	682	882	999					

Mediante la prueba Kolmogoroff-Smirnoff, se quiere decidir si de todas maneras, estas observaciones pueden ser consideradas como provenientes de una población de personas del sexo masculino, sanas. (con α=0.01).

De acuerdo al procedimiento de prueba, debe encontrarse el valor $T=\sup_x |F^*_x(x)-S_n(x)|$, en donde $S_n(x)$ representa a la distribución empírica de la muestra y $F^*_x(x)$ a una distribución normal con parámetros: μ=235.000 y σ=44,600.

Con los datos al principio, estas dos distribuciones son:

x	173	189	196	207	215	237	275	282	293	300
$S_n(x)$	0.04	0.08	0.12	0.16	0.20	0.24	0.28	0.32	0.36	0.40
$F^*_x(x)$	0.0823	0.1515	0.1922	0,2676	0,3300	0.5160	0.7852	0,8531	0,9032	0.9265

x	305	316	346	382	395	399	401	437	480	504
$S_n(x)$	0.44	0.48	0.52	0.56	0.60	0.64	0.68	0.72	0.76	0.80
$F^*_x(x)$	0.9406	0.9648	0.9934	1.00	1.00	1.00	1.00	1.00	1.00	1.00

x	524	634	682	882	999
$S_n(x)$	0.84	0.88	0.92	0.96	1.00
$F^*_x(x)$	1.00	1.00	1.00	1.00	1.00

Los valores de $F^*_x(x)$ fueron encontrados como sigue:

Se estandarizaron los valores Xi usando a μ y a σ haciendo

$y_i = \frac{X_i-\mu}{\sigma}$, y después se buscaron los valores de W_r tales que:

$P\{Y_r < \frac{Xr-\mu}{\sigma}\} = W_r$, donde Y_r es una variable normal con parámetros $\mu=0$ y $\sigma^2=1$. (Usando la tabla 1).

Teniendo las distribuciones se calculan las diferencias $|F^*_X(x_i)-S_n(x_i)|$ y $|F^*_X(x_i)-S_n(x_{i-1})|$:

$\lvert F^*_X(x_i)-S_n(x_{i-1})\rvert$	X_i	$\lvert F^*_X(x_i)-S_n(x_i)\rvert$
0.0823	173	0.0423
0.1115	189	0.0715
0.1122	196	0.0722
0.1476	207	0.1076
0.1700	215	0.1300
0.3160	237	0.2760
0.5452	275	0.5052
0.5731	282	0.5331
0.5832	293	0.5432
0.5665	300	0.5265
0.5406	305	0.5006
0.5248	316	0.4848
0.5134	346	0.4734
0.4800	382	0.4400
0.4400	395	0.4000
0.4000	399	0.3600
0.3600	401	0.3200
0.3200	437	0.2800
0.2800	480	0.2400
0.2400	504	0.2000
0.2000	524	0.1600
0.1600	634	0.1200
0.1200	682	0.0800
0.0800	882	0.0400
0.0400	999	0.0000

La diferencia máxima es 0.5832. Buscando en la tabla 4 del apéndi_ ce A se obtiene que $W_{0.99}=0.317$, por lo que $T > W_{1-\alpha}$ y se rechaza la hipótesis. Es decir, los datos no provienen de personas saludables. (Se rechazó el que la distribución fuese una normal con parámetros $\mu=235{,}000$ y $\sigma=44{,}600$. Sin embargo, la distribución puede ser normal, sólo que con otros parámetros).

EJEMPLO 2 Supóngase la siguiente muestra aleatoria de tamaño 10:

$X_1=0.621$ $X_2=0.503$ $X_3=0.203$ $X_4=0.477$ $X_5=0.710$

$X_6=0.581$ $X_7=0.329$ $X_8=0.480$ $X_9=0.554$ $X_{10}=0.382$.

Se desean probar las siguientes hipótesis:

Ho: $F_X(x)=F_X^*(x)\ \forall x$

Ha: $F_X(x)\neq F_X^*(x)$ para al menos una x; con $F_X^*(x)$ como sigue:

$$F_X^*(x)=\begin{cases} 0 & \text{si} \quad x<0 \\ x & \text{si} \quad 0<x<1 \\ 1 & \text{si} \quad 1<x \end{cases}$$

Para llevar a cabo esta prueba de hipótesis, la prueba Kolmogoroff-Smirnoff es adecuada, pues $F_X^*(x)$ está completamente determinada. (Supóngase que $\alpha=0.01$).

La siguiente tabla, contiene las diferencias entre la distribución empírica y la distribución $F_X^*(x)$:

x	$S_n(x)$	$\lvert F_x(x_i)-S_n(x_{i-1})\rvert$	$\lvert F_x(x_i)-S_n(x_i)\rvert$
0.203	0.10	0.203	0.103
0.329	0.20	0.209	0.129
0.382	0.30	0.182	0.082
0.477	0.40	0.177	0.077
0.480	0.50	0.080	0.020
0.503	0.60	0.003	0.097
0.554	0.70	0.046	0.146
0.581	0.80	0.119	0.219
0.621	0.90	0.179	0.279
0.710	1.00	0.190	0.290

Puede observarse que la estadística $T=\sup_x |S_n(x)-F_x^*(x)|$ para estos datos, toma el valor:

$$T= 0.290$$

Usando la tabla 4, para $n=10$ y $\alpha=0.01$, el cuantil correspondiente es 0.489; por lo que la hipótesis nula no es rechazada, pudiéndose decir que la muestra proviene de una distribución uniforme en el intervalo (0,1).

EJEMPLO 3 Los datos de la siguiente tabla muestran el número de aparatos automáticos que sirven para dar cambio en teléfonos, que estaban en uso en un instante dado. Se desea verificar la hipótesis de que los datos constituyen una muestra aleatoria de una poblacion de Poisson con $x=10.5$. La tabla muestra también los valores de $S_n(x)$, $F_x^*(x)$, $|F_x(x_i)-S_n(x_i)|$ y $|F_x(x_i)-S_n(x_{i-1})|$

No.de Aparatos Ocupados.	Frecuencia Observada.	$S_n(x)$	$F^*_X(x)$	$\lvert F^*_X(x_i)-S_n(x_{i-1})\rvert$	$\lvert F^*_X(x_i)-S_n(x_i)\rvert$
0	0	0	0	0,001	0
1	5	0,001	0	0,005	0,001
2	14	0,005	0,002	0,009	0,003
3	24	0,011	0,007	0.020	0,004
4	57	0,027	0,04	0.035	0,006
5	111	0,056	0,050	0.059	0,006
6	197	0,109	0,102	0.081	0,007
7	278	0,183	0,179	0.104	0,004
8	378	0,283	0,279	0.116	0,004
9	418	0,395	0,397	0.121	0,002
10	461	0,518	0,521	0,112	0,003
11	433	0,633	0,639	0.104	0,006
12	413	0,743	0,742	0.096	0,001
13	358	0,838	0,825	0.072	0,013
14	219	0,897	0,888	0.047	0,009
15	145	0,935	0,932	0.032	0,003
16	109	0,964	0,960	0,019	0,004
17	57	0,979	0,978	0.013	0,001
18	43	0,991	0,988	0.007	0,003
19	16	0,995	0,994	0,003	0,001
20	7	0,997	0,997	0.002	0
21	8	0,999	0,999	0.001	0
22	3	1,000	0,999	0,001	0,001

En donde $F^*_X(x)= \begin{cases} \sum\limits_{i=1}^{[x]} \dfrac{{}^{-\lambda}\lambda^{i}}{\lambda!} & \text{si } i\in N\cup\{0\},\ \lambda=10,5 \\ 0 & \text{en otro caso} \end{cases}$

Para probar Ho: $F_X(x)= F^*_X(x)$ contra Ha: $F_X(x)\neq F^*_X(x)$, se usará la prueba Kolmogoroff-Smirnoff, pues $F^*_X(x)$ está completamente especificada. Suponiendo $\alpha=0.05$ y teniendo en cuenta que $T=0.121$, se tiene que Ho es rechazada pues $W_{1-\alpha}=0.022$ por lo que se concluye que estos datos no pueden ser ajustados por una distribución de Poisson con $\lambda=10.5$.

PRUEBA LILLIEFORS

Supóngase el problema del equipo de basket-bol que se planteó anteriormente, sólo que ahora las personas interesadas en formar el equipo lo único que saben es que las alturas de los estudiantes son variables aleatorias, donde cada una se distribuye según una normal con ciertos parámetros μ (media) y σ^2 (varianza) que no conocen.

Esto es, en el problema original se tenía que la distribución de las alturas de los estudiantes era una normal $N(\mu,\sigma^2)$ con μ y σ^2 conocidos, pero ahora sólo se cree que es una normal pero no saben con que μ ni con que σ^2.

En este caso sólo les interesa probar si efectivamente las alturas son variables aleatorias normales y en caso de que así fuese, idear después un procedimiento para estimar los valores de μ y σ^2.

Para probar ésto, se va a usar una variante de la prueba Kolmogoroff-Smirnoff, llamada prueba Lilliefors, que se basa en una estadística del tipo $\sup_{x\varepsilon X} |F_X^*(x)-S_n(x)|$.

Como en los casos pasados, Supóngase $X_1,\ldots,X_n$ una muestra aleatoria de la población que se supone tiene una distribución normal con parámetros μ y σ^2 desconocidos. Al igual que en la prueba Kolmogoroff- Smirnoff, se desean calcular las diferencias $|F_X^*(x)-S_n(x)|$, donde $F_X^*(x)$ es la distribución que se cree es la real, que en este caso de la prueba de Lilliefors es una normal, y $S_n(x)$ es la distribución empírica de la muestra.

Sin embargo para calcular $F_X^*(x)$ para cada $x\varepsilon x$, es necesario conocer los parámetros μ y σ^2, pero en este caso se desconocen los valores de μ y σ^2. Para subsanar esto y poder usar una estadística del tipo Kolmogoroff- Smirnoff, las diferencias que se toman son entre la distribución empírica de la muestra y una distribución normal con media $\bar{X}=\frac{1}{n}\ \Sigma X_i$ y varianza $S^2=\frac{1}{n-1}\ \Sigma(X_i-\bar{X})^2$, donde estos valores son usados como estimadores de μ y σ^2 respectivamente y se calculan con la información que dá la muestra, es decir, con los elementos de la muestra.

Pero puede seguirse otro método para efectuar está prueba que es equivalente al que se acaba de mencionar. La idea es como sigue:

Si las observaciones $X_1,\ldots,X_n$, provienen de una distribución normal con parámetros μ y σ^2 entonces las observaciones $w_i=\frac{X_i-\mu}{\sigma}$, son observaciones de una población N(0,1). En este caso lo único que se quiere probar es si las observaciones son normales con media cero y varianza uno. pero nuevamente, para construír las w_i se necesita conocer el valor de μ y σ^2 entonces en lugar de usar las observaciones w_i, se construyen las observaciones estandarizadas $y_i=\frac{X_i-\bar{X}}{\sqrt{S^2}}$, con $\bar{X}$ y S^2 definidos como antes.

De este modo las diferencias $|F_x^*(x)-S_n(x)|$ se calcularán tomando a $F_x^*(x)$ como la distribución normal con media cero y varianza uno, para lo cual existen valores tabulados; y $S_n(x)$ como la distribución empírica de la muestra estandarizada, usando a $F_y^*(y)$ con $y\sim N(0,1)$ en lugar de $F_x^*(x)$ y a $Sn(y)$ en lugar de $Sn\ (x)$ quedando las diferencias como $|F_y^*(y)-S_n(y)|$.

Ya calculadas estas diferencias, se tomará como antes la estadística $T=\sup_{y\epsilon} |F_y^*(y)-S_n(y)|$ con $F_y^*(y)$ y $S_n(y)$ como las especificadas en el párrafo anterior y se comparará a T con el $(1-\alpha)$ cuantil de las tablas obtenidas por medio de simulación en computadora por Lilliefors (1967). Si la prueba se hace a un nivel de significancia α, se rechazará la hipótesis de que las observaciones provienen de una distribución normal con μ y σ^2 desconocidas, si el valor de T es mayor que el $(1-\alpha)$; aceptándose en caso contrario.

Cabe recalcar que la prueba de Lilliefors sirve únicamente para -

checar si la muestra aleatoria que se tiene, proviene de una población que tiene asociada una distribución normal sin importar el valor de los parámetros μ y σ^2 específicos. Por otro lado la prueba de Kolmogoroff-Smirnoff, sirve para probar otras hipótesis sobre - distribuciones que no necesariamente son la normal por lo que la - prueba de Lilliefors puede considerarse como una variante de la - Kolmogoroff-Smirnoff, cuando la distribución que se desea probar - es una normal. Otra diferencia, es que en la Kolmogoroff-Smirnoff, la distribución empírica se calcula directamente con la muestra aleatoria que se obtiene, mientras que en la prueba de Lilliefors, la distribución empírica se calcula a partir de datos ya estandarizados.

EJEMPLOS

EJEMPLO 1 Cincuenta números de dos cifras fueron obtenidas aleatoriamente de un directorio telefónico. Aún cuando la variable aleatoria de la muestra es claramente discreta, se puede justificar el probar normalidad si se tiene en cuenta que aceptar la hipótesis - de normalidad no quiere decir que la variable aleatoria es normal y por tanto continua, sino que indica que la diferencia entre la - distribución normal y la verdadera distribución, es suficientemente insignificante como para detectarla.

Los números Xi's, están ordenados de menor a mayor y los valores - estandarizados, denotados con Yi's, se calculan como $Y_i = \frac{X_i - \bar{X}}{S}$, en donde $\bar{x} = 55.04$ y $S = 19.00$.

Xi	Yi	Xi	Yi	Xi	Yi	Xi	Yi	Xi	Yi
23	-1.69	36	-1,00	54	-0,05	61	0,31	73	0,95
23	-1.69	37	-0.95	54	-0.05	61	0.31	73	0.95
24	-1,63	40	-0.79	56	0,05	62	0,37	74	1,00
27	-1.48	42	-0,69	57	0.10	63	0.42	75	1.05
29	-1.37	43	-0.63	57	0.10	64	0.47	77	1.16
31	-1,27	43	-0,63	58	0.16	65	0,52	81	1.37
32	-1.21	44	-0.58	58	0.16	66	0.58	87	1.68
33	-1.16	45	-0.53	58	0,16	68	0,68	89	1.79
33	-1.16	48	-0,37	58	0.16	68	0.68	93	2.00
35	-1.05	48	-0.37	59	0.21	70	0.79	97	2.21

La hipótesis de normalidad, se prueba con la estadística:

$T = \sup_{x} |F^{*}_{y}(y) - Sn(y)|$, en donde $F^{*}_{y}(y)$ es la distribución normal estandar y $Sn(y)$, la distribución empírica de las Yi's.

La siguiente tabla muestra de valores de $Sn(y)$, $F^{*}_{y}(y)$ y las diferencias $|F^{*}_{y}(y_i) - Sn(Y_i)|$ y $|F^{*}_{y}(Y_i) - Sn(Y_{i-1})|$.

y_i	$S_n(y_i)$	$F^{*}_{y}(y_i)$	$\lvert F^{*}_{y}(y_i) - S_n(y_{i-1})\rvert$	$\lvert F^{*}_{y}(y_i) - S_n(y_i)\rvert$
-1,69	0,04	0.0455	0,0455	0.0055
-1,63	0,06	0.0516	0,0116	0.0084
-1,48	0,08	0,0694	0,0094	0,0106
-1,37	0.10	0,0853	0,0053	0,0147
-1,27	0.12	0,1020	0,0020	0,0180

-1.21	0.14	0.1131	0.0069	0.0269
-1.16	0.18	0.1230	0.0170	0.0570
-1.05	0.20	0.1469	0.0331	0.0531
-1.00	0.22	0.1587	0.0413	0.0613
-0.95	0.24	0.1711	0.0489	0.0689
-0.79	0.26	0.2148	0.0252	0.0452
-0.69	0.28	0.2451	0.0149	0.0349
-0.63	0.32	0.2643	0.0157	0.0557
-0.58	0.34	0.2810	0.0890	0.0590
-0.53	0.36	0.2981	0.0419	0.0619
-0.37	0.40	0.3557	0.0043	0.0443
-0.05	0.44	0.4801	0.0801	0.0401
0.05	0.46	0.5199	0.0799	0.0599
0.10	0.50	0.5398	0.0798	0.0398
0.16	0.58	0.5636	0.0636	0.0164
0.21	0.60	0.5832	0.0032	0.0168
0.31	0.64	0.6217	0.0217	0.0183
0.37	0.66	0.6443	0.0043	0.0157
0.42	0.68	0.6628	0.0028	0.0172
0.47	0.70	0.6808	0.0008	0.0192
0.52	0.72	0.6985	0.0015	0.0215
0.58	0.74	0.7190	0.0010	0.0210
0.68	0.78	0.7517	0.0117	0.0283
0.79	0.80	0.7852	0.0052	0.0148
0.95	0.84	0.8289	0.0290	0.0110
1.00	0.86	0.8413	0.0013	0.0187
1.05	0.88	0.8531	0.0069	0.0269
1.16	0.90	0.8770	0.0030	0.0230
1.37	0.92	0.9147	0.0147	0.0053
1.68	0.94	0.9535	0.0335	0.0135
1.79	0.96	0.9633	0.0233	0.0033
2.00	0.98	0.9772	0.9772	0.0172
2.21	1.00	0.9864	0.0064	0.0136

El valor máximo es 0.0801, esto es T=0.0801. Buscando en la tabla 5 del apéndice se tiene que para $n > 30$, el $(1-\alpha)$-cuantil (con $\alpha=0.05$) es $W_{0.95} = \frac{0.866}{\sqrt{50}} = \frac{0.866}{7.07} = 0.1225$. Por lo que $T < W_{0.95}$ y entonces no se rechaza normalidad.

El no rechazar normalidad, no significa que los datos siguen una distribución normal, sino que la distribución normal no parece ser una mala aproximación de la verdadera distribución (la cual es desconcida).

EJEMPLO 2 Supóngase la situación descrita en el ejemplo uno de la prueba Kolmogoroff-Smirnoff. Para esos datos se tiene que $\bar{x}=402.12$ y S=210.68. Si se calculan las Y_i's, haciendo $Y_i = \frac{X_i - \bar{x}}{S}$ y se hace lo mismo que en el ejemplo 1 de esta prueba, se tiene que la diferencia mayor es: T=0.1800 y usando la tabla 5 se tiene que $W_{0.95}=0.200$ por lo que la hipótesis de normalidad no se rechaza. Es decir los datos provienen de una normal; lo cual no contradice la conclusión obtenida anteriormente.

EJEMPLO 3 Se obtuvieron 20 muestras de un tipo particular de alambre y su resistencia en ohms fué medida. Los resultados son los siguientes:

9.8	9.3	10.4	13.8
14.5	11.1	8.3	12.9
13.7	10.1	11.5	10.6
7.6	12.7	10.0	8.9
10.5	9.9	9.1	9.5

¿Se podría justificar el suponer que estas observaciones provienen de uns distribución normal?.

Para responder a esta pregunta se usara la prueba de Lilliefors, que verifica si ciertas observaciones pueden suponerse provenientes de una normal.

Procediendo como en los ejemplos anteriores, se tiene el siguiente cuadro:

Xi	$Y_i = \frac{X_i - \bar{X}}{S}$	$S_n(Y_i)$	$F_y^*(Y_i)$	$\lvert F_y^*(Y_i) - S_n(Y_{i-1})\rvert$	$\lvert F_y^*(Y_i) - S_n(Y_i)\rvert$
7.6	-1.62	0.05	0.0526	0,0526	0,0026
8,3	-1,26	0,10	0,1038	0,0538	0,0038
8,9	-0,94	0,15	0,1736	0,0736	0,0236
9,1	-0,84	0,20	0,2005	0,0505	0,0005
9,3	-0,72	0,25	0,2358	0,0358	0,0142
9,5	-0,63	0,30	0,2643	0,0143	0,0357
9,8	-0,47	0,35	0,3192	0,0192	0,0308
9,9	-0,43	0,40	0,3372	0,0128	0,0628
10,0	-0,37	0,45	0,3557	0,0443	0,0943
10,1	-0,32	0,50	0,3745	0,0755	0,1255
10,4	-0,16	0,55	0,4364	0,0636	0,1136
10.5	-0,11	0,60	0,4562	0,0938	0,1438
10.6	-0,06	0,65	0,4761	0,1239	0,1739
11,1	0,20	0,70	0,5793	0,0707	0,1207
11,5	0,41	0,75	0,6591	0,0409	0,0909
12,7	1,04	0,80	0,8508	0,1005	0,0508
12,9	1,14	0,85	0,8729	0,0729	0,0229
13,7	1,56	0,90	0,9406	0,0906	0,0406
13.8	1,61	0,95	0,9463	0,0463	0,0037
14,5	1,97	1,00	0,9756	0,0256	0,0244

Como puede observarse, la diferencia máxima es T=0,1739. Usando α=0,05, se tiene que el cuantil correspondiente es: 0,190, por lo que la hipótesis de nulidad Ho: Las observaciones se distribuyen como normal, no es rechazada.

BONDAD DE AJUSTE
OBSERVACIONES.

Como se dijo antes, para la prueba de X^2, el número de categorías en que la información de la muestra es clasificada, es arbitrario; esto es, el total de categorías queda a criterio del que hace la prueba o del tipo de información que se tenga. Sin embargo para que la prueba tenga alguna validez se recomienda lo siguiente:

i.- Si el número de categorías es dos, la prueba se aplicará sólo que E_i sea mayor o igual a cinco, pues para valores menores, la prueba no tiene mucho sentido.

ii.- Si el número de categorías es mayor que dos, se buscará que el 80% de los valores de E_i sean mayores o iguales a cinco. Si esto no sucede deberá buscarse el asociar o juntar dos o más categorías (suponiendo el que tenga sentido hacer esto, en caso contrario buscar otro procedimiento para llevar a cabo la prueba, por ejemplo usar la prueba Kolmoroff-Smirnoff) para que los valores de E_i aumenten y así se cumpla que el 80% de los valores de E_i sean mayores o iguales a cinco.

Por otro lado, en vista de que las observaciones muestrales se clasifican, esto puede hacer que se pierda algo de información al combinarse en categorías, cosa que con la prueba Kolmogoroff-Smirnoff no sucede, además si las muestras son pequeñas y tienen que combinarse en pocas categorías, la prueba Kolmogoroff-Smirnoff es más potente que la X^2, es por esto que es más recomendable usar más la prueba Kolmogoroff- Smirnoff, que la prueba X^2. Sin embargo la prueba X^2 tiene una ventaja bastante importante sobre la otra, y es que para usar la prueba de Kolmogoroff-Smirnoff es necesario que la distribución que se supone verdadera debe estar completamente especificada, es decir deben conocerse todos los parámetros de la distribución, lo cual en la prueba de X^2, no es necesario pues si algún parámetro se desconoce, éste puede ser estimado (perdiéndose por ésto un grado de libertad), para después aplicarse la prueba.

X^2 PRUEBA DE INDEPENDENCIA

Supóngase que los elementos de una población, pueden ser clasificados de acuerdo a dos criterios, por ejemplo: Si la población está formada por gentes, uno de los criterios podría ser el color de ojos y el otro la altura; o bien, uno de ellos el peso y el otro la rama de actividad a la cual el sujeto se dedique. Si la población está formada por animales, una clasificación podría ser hecha de acuerdo al tipo de animal (rata, conejo, etc.) y otra de acuerdo al efecto de cierto tratamiento aplicado a ellos.

La pregunta que resulta obvia al hacer las dos clasificaciones es: ¿Los dos criterios de clasificación son independeintes entre sí? es decir ¿Influye uno del otro? por ejemplo, se podría pensar, dado un estrato social establecido en cierta población, que a mayor posición dentro del estrato de gentes que estén estudiando alguna carrera, mayor aprovechamiento de sus estudios,es decir, en este caso, se podría pensar que existe relación entre las dos clasificaciones: posición económica y aprovechamiento académico.

Para desarrollar la prueba, suponga que se toma una muestra aleatoria (o bien, tomar toda la población, sí esto es posible) de tamaño N, donde la información puede ser clasificada de acuerdo a dos criterios, dentro de los cuales existen varias categorías. Por ejemplo, en el caso de que las clasificaciones fuesen color de ojos por un lado y altura, por el otro, las categorías podrían ser cafés, verdes, negros, azules,etc., en una y menos de 1.50, entre 1.50 y 1.60, entre 1.60 y 1.70, entre 1.70 y 1.80; en la otra.

De acuerdo a ésto, si se tienen dos criterios de clasificación, el A y el B y dentro de A existen p categorías y dentro del B, q categorías; la información puede escribirse de la forma siguiente.

		CRITERIO B			
		1	2	3	q
	1	O_{11}	O_{12}	O_{13}	O_{1q}
	2	O_{21}	O_{22}	O_{23}	O_{2q}
	3	O_{31}	O_{32}	O_{33}	O_{3q}
CRITERIO	.				
A	.				
	.				
	p	O_{p1}	O_{p2}	O_{p3}	O_{pq}

en donde O_{ij} es el número de elementos en la población, que de acuerdo al criterio A quedaron clasificados en la categoría i-ésima y de acuerdo al criterio B, se clasificaron en la j-ésima categoría. (A este arreglo se le llama tabla de contingencia de pxq).

Por otro lado, cada elemento de la población tiene cierta probabilidad p_{ij}, de ser clasificado en la i-ésima categoría del criterio A y en la j-ésima categoría del B. Esto es, todo elemento de la población bajo estudio tiene una probabilidad de clasificarse en un renglón y en una columna, simultáneamente; si se denota con $p_{i.}$ a la probabilidad de ser clasificado en el renglón i-ésimo, y con $p_{.j}$ la de ser clasificado en la columna j-ésima, entonces la hipótesis de independencia en los criterios de clasificación, puede ser escrita como:

$$Ho: p_{ij}=p_{i.}p_{.j} \quad i \in Jp, \quad j \in Jq$$

y la alternativa natural sería:

$$Ha: p_{ij} \neq p_{i.}p_{.j} \text{ para alguna } i \in J_p \text{ y alguna } j \in J_q$$

Si la hipótesis Ho, fuese cierta, entonces, los valores observados O_{ij} deberían ser muy parecidos a los valores esperados $Np_{i.}p_{.j}$, por lo que la estadística de prueba puede ser:

$$\sum_{i=1}^{p}\sum_{j=1}^{q}\frac{(O_{ij}-Np_{i.}p_{.j})^2}{Np_{i.}p_{.j}}$$

siguiendo el razonamiento usado en la prueba X^2 de bondad de ajuste.

Sin embargo $p_{i.}$ y $p_{.j}$, son valores desconocidos, por lo que tienen que estimarse. Puede demostrarse (Mood y Graybill (1963), pp 311-319) que los estimadores son: $\frac{Ri}{N}$ y $\frac{Cj}{N}$ respectivamente; en donde Ri es el total de observaciones que se clasificaron en el i-ésimo renglón, (ie. $Ri = \sum_{j=1}^{q} O_{ij}$)*, Cj es total de observaciones que se clasifican en la columna j-ésima ($Cj=\sum_{i=1}^{p} O_{ij}$) y N es el total de observaciones clasificadas.

Con esto la estadística de prueba quedaría como:

$$T=\sum_{i=1}^{p}\sum_{j=1}^{q}\frac{\left(Oij-\frac{N(RiCj)}{N^2}\right)^2}{\frac{N(RiCj)}{N^2}}$$

o equivalentemente:

$$T=\sum_{i=1}^{p}\sum_{j=1}^{q}\left(\frac{Oij-\frac{RiCj}{N}}{\frac{RiCj}{N}}\right)^2$$

la cual se distribuye asintóticamente como una X^2 con $(p-1)(q-1)$ grados de libertad (Mood y Graybill Op.Cit).

De esta forma, si la prueba de $Ho: p_{ij}= p_{i.}p_{.j}$ se realiza a un nivel α, la hipótesis de independencia Ho, se rechazará si la estadística T es mayor o igual que el cuantil de orden $(1-\alpha)$ de una X^2 con $(p-1)(q-1)$ grados de libertad, de otra forma no se rechazará.

* Nótese que Ri, es una variable aleatoria.

Sin embargo, como se está usando una distribución asintótica, la aproximación no es buena cuando algunas de las cantidades $\frac{R_iC_j}{N}$ son pequeñas; Cochran (1954) demostró que si alguna de las cantidades $\frac{R_iC_j}{N}$ es menor que la unidad, o, si más del 20% de ellas, menor que cinco, la aproximación a la X^2 no es buena. Si esto sucede, pueden combinarse varias categorías para evitarlo, siempre y cuando el significado de la hipótesis Ho, no varíe.

La literatura existente sobre esta prueba y sobre tablas de contingencia es bastante, la más importante se menciona en la sección de k poblaciones independientes, en donde la prueba X^2 se usa para probar diferencia de poblaciones, y en la cual la estadística es análoga a la vista aquí, sólo que con pequeñas diferencias. Por esto, se recomienda que al ver esta prueba, se vea la prueba de diferencias de probabilidades en la parte de k muestras.

EJEMPLO 1 Se desea saber si la posición en la cual los caballos llegan en una carrera dependen del carril del cual parten, con este objeto se eligió un caballo al azar de cada una de 80 carreras, y se observó el carril que ocupaban al iniciar la carrera (del 1 al 4 y del 5 al 9) y en que posición llegaron a la meta (primera (1), segunda (2), tercera (3) y otra). Los resultados obtenidos son:

		Posición de llegada			
		1	2	3	otra
Posición de	1-4	8	6	8	16
Partida	5-9	3	6	5	28

(La prueba se quiere realizar a un nivel $\alpha=0.05$)

La prueba que se utiliza es la X^2 para probar independencia de criterios, pues puede considerarse, una sola muestra clasificada de acuerdo a dos criterios (posición de partida y de llegada), en varias categorías. Entonces Ho: Los criterios son independientes.

Lo primero que se hará, será calcular los valores esperados, que de acuerdo a lo visto, se definen como: $E_{ij} = \frac{R_i C_j}{N}$, donde R_i es el total del renglón i-ésimo, C_j el total de la columna j-ésima y N el total de observaciones. En el ejemplo, i=2 y j=4, teniéndose:

$R_1=38$ y $R_2=42$; $C_1=11$, $C_2=12$, $C_3=13$ y $C_4=44$, y N=80 por lo que:

$E_{11} = 5.225$	$E_{21} = 5.775$
$E_{12} = 5.700$	$E_{22} = 6.300$
$E_{13} = 6.175$	$E_{23} = 6.825$
$E_{14} = 20.900$	$E_{24} = 23.100$

Entonces:

$$\sum_{i=1}^{2} \sum_{j=1}^{4} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 6.0529$$

Como son cuatro columnas y dos renglones, se tienen (4-1)(2-1)=3 grados de libertad. El cuantil 0.95 es: 7.815. Ya que el valor calculado es menor que el obtenido en tablas (tabla 2), la hipótesis nula no es rechazada y se concluye que los criterios son independientes es decir, la posición de salida no influye en la posición de llegada.

EJEMPLO 2 En un estudio de la relación entre el tipo de sangre y enfermedad, se tomó una muestra de 8766 pacientes de un hospital y se clasificó con respecto a su tipo de sangre (O,A,B) y a la enfermedad (úlcera, cáncer gastrointestinal, ninguna de las dos). Las observaciones ya clasificadas se muestran en la siguiente tabla:

Tipo de Sangre	Tipo de Enfermedad	Ulcera	Cancer G.I.	Ninguna de las dos	Total
O		983	383	2892	4258
A		679	416	2625	3720
B		134	84	570	788
TOTAL		1796	883	6087	8766

Para probar si existe independencia entre los dos criterios (Ho), la prueba X^2 es adecuada, por lo que se calculará la estadística:

$$\sum_i \sum_j \frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$, para ésto , se sabe que:

$R_1 = 4258$ $C_1 = 1796$ $N = 8766$

$R_2 = 3720$ $C_2 = 883$

$R_3 = 788$ $C_3 = 6087$

entonces:

$E_{11} = 872.3896$ $E_{21} = 762.1629$ $E_{31} = 161.4474$

$E_{12} = 428.9087$ $E_{22} = 374.7159$ $E_{32} = 79.3753$

$E_{13} = 2596.7015$ $E_{23} = 2583.1211$ $E_{33} = 547.1772$

y los valores observados son:

$O_{11} = 983$ $O_{21} = 679$ $O_{31} = 134$

$O_{12} = 383$ $O_{22} = 416$ $O_{34} = 84$

$O_{13} = 2892$ $O_{23} = 2625$ $O_{33} = 570$

por lo que:

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\frac{(O_{ij}-E_{ij})^2}{E_{ij}} = 40.5429$$

Buscando en tablas (tabla 2) con 4 grados de libertad, se encuentra que $W_{0.95} = 9.4880$, por lo que la hipótesis Ho de independencia es rechazado, teniéndose que existe relación entre el tipo de sangre y el tipo de enfermedad del paciente.

EJEMPLO 3 Se sabe que existen dos tipos distintos de amoebas (amibas) que pueden alojarse en los intestinos del hombre. Un tipo, puede causar la muerte y el otro no, pero la presencia del segundo puede prevenir el fallecimiento de la persona causado generalmente por el primero.

En una planta donde se trabaja la madera, en South Bend, Indiana; hubo una epidemia intestinal entre los trabajadores de la planta. Los doctores atribuían la epidemia a un tipo de amoeba llamado Entamoeba histolytica. Después de estudiar todas las posibles alternativas, representantes de la oficina de Salud Pública concluyeron que la amoeba se introducía por la planta de agua del lugar. Para averiguar cuantos de los hombres de la planta que no se sentían enfermos, estaban infectados con la amoeba, se tomó una muestra aleatoria de 138 trabajadores aparéntemente sanos y se encontró amoeba en el contenido intestinal de 70 de estos hombres. Se distinguieron dos razas de Entamoeba histolytica: grande y pequeña. Se encontró que 35 hombres tenían únicamente entamoeba histoty tica grande, 23 tenían entamoeba histolytica pequeña únicamente y 12 tenían ambas.

Raza Grande

		Presencia	Ausencia	
Raza	Presencia	12	23	35
Pequeña	Ausencia	35	68	103
		47	91	138

La pregunta que se quería contestar era:¿Las dos razas interactúan en el hombre infectado?.

Para contestarla se usó el criterio de la X^2 para probar independencia en las clasificaciones. Se calcularon los valores esperados, resultando el siguiente cuadro:

Raza Grande

		Presencia	Ausencia
Raza	Presencia	11.93	23.07
Pequeña	Ausencia	35.07	67.93

Entonces:

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(O_{ij}-E_{ij})^2}{E_{ij}} = \frac{(12-11.93)^2}{11.93}+\frac{(23-23.07)^2}{23.07}+\frac{(35-35.07)^2}{35.07}+\frac{(68-67.93)^2}{67.93}$$

$$\Rightarrow \quad \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(O_{ij}-E_{ij})^2}{E_{ij}} = 0.0008$$

Como se tienen dos columnas y dos renglones, hay sólo un grado de libertad. Buscando en la tabla 2 del apéndice A, se ve que $0.0008 < W_{1-\alpha}$ para cualquier α (de las tabuladas de esa tabla), por lo que la hipótesis de independencia no se rechaza. Esto es, la infección con una raza, no hace al individuo más o menos suceptible a la infección con la otra.

PRUEBA DE CORRIDAS (WALD– WOLFOWITZ)

Cuando se tiene una población bajo estudio, muchas veces es deseable o necesario el hacer ciertas inferencias acerca de ella a partir de la información que se obtiene al tomar una muestra, que como se vió anteriormente, debe ser aleatoria.

Sin embargo, no siempre las personas que van a hacer las inferencias, son las mismas que las que tomaron la muestra, por lo que pueden no estar seguros de si la muestra es realmente aleatoria.

O bien en un momento dado puede ser de interés averiguar si ciertos fenómenos son aleatorios, como por ejemplo los resultados obtenidos al lanzar la moneda, o el número de individuos de cierto sexo que esperan ser atendidos en la caja de un supermercado, etcétera.

En general el problema podría plantearse como sigue: Supóngase que se tiene n observaciones $X_1, X_2, \ldots, X_n$ y lo que se quiere averiguar es si el proceso que las generó es aleatorio. Supóngase además que cada observación es tal que puede clasificarse en una de dos categorías, (por ejemplo, si la información está referida en una escala nominal u ordinal y puede tomar tres valores distintos, dos de ellos pueden tomarse dentro de una misma clasificación para formar así las dos categorías); la forma en que se definan las dos categorías va a depender del problema específico que se tenga y de la hipótesis que se desee probar. Si la información está dada en forma numérica, puede tomarse un punto a partir del cual todas las observaciones muestrales mayores o iguales a él, sean clasificadas como del tipo "a" y las menores a él, del tipo "b". Este punto podría ser la mediana de las observaciones o un percentil conveniente de acuerdo al problema, o un punto cualquiera. En caso en que la información de la nuestra no pueda ser clasificada en dos categorías, la técnica que a continuación se discute, no es útil para probar la aleatoriedad en la muestra.

El procedimiento se basa en corridas o rachas. Dada una sucesión ordenada de dos o más tipos de símbolos, una corrida o racha es una sucesión de uno o más símbolos semejantes que están seguidos y precedidos por un símbolo diferente o por ninguno. Por ejemplo, en la sucesión A BB A B AA B hay 6 corridas o rachas; si se supone que todos los números menores o iguales a 15 son semejantes entre sí y los mayores son semejantes entre sí, se tendrá que en 13 17 - 9 12 15 19 21 7 18 existen seis rachas o corridas. Cualquier falta de independencia se reflejará en ciertas tendencias en las corridas, por ejemplo: muchas corridas, pocas corridas, una corrida muy grande, se toman como índice de falta de independencia.

Como se mencionó antes, las corridas se utilizarán al existir una dicotomía, en los datos u observaciones; si son cualitativos, o sea, medidos en escala nominal u ordinal, la dicotomía usual es inherente a los datos, o si son datos cuantitativos (escala de proporciones o de intervalos), se puede usar la media o la mediana para clasificarlos.

Considérese entonces una muestra de n elementos, en donde n_1, son de un tipo y n_2 son de otro tipo (con $n_1 + n_2 = n$) y supóngase que existen r_1 corrida de elementos del primer tipo y r_2 corridas de elementos del segundo tipo, por lo que habrá $r = r_1 + r_2$ corridas en total. Debido a que el número de corridas es el que ayudará a decidir si la muestra es aleatoria o no, la prueba se basa en la distribución del número de corridas, R. Ahora, la hipótesis a probar será: Ho: la muestra es aleatoria.

Bajo Ho, cualquier arreglo de los n elementos es igualmente probable, entonces la probabilidad de obtener r_1 corridas de n_1 objetos y r_2 corridas de n_2 objetos, será el cociente entre el número de arreglos posibles con r_1 corridas de n_1 objetos y r_2 corridas de n_2 objetos, entre el total de arreglos de n_1 elementos formados de n, esto es:

$$P(R_1=r_1,R_2=r_2)= \frac{w_1}{w_2} \; ,$$

con:

w_1= al número de arreglos posibles formados con (r_1 corridas de n_1 objetos y r_2 corridas de n_2 objetos), y $w_2= \frac{n!}{n_1!\,n_2!}$ donde R_1,R_2 son variables aleatorias que cuentan el número de corridas de uno (r_1) y otro tipo (r_2) en una muestra de n elementos.

Para evaluar el numerador de $P\{R_1=r_1,R_2=r_2\}$, es necesario evaluar el número de arreglos posibles de objetos de uno u otro tipo, con la propiedad de que $R_1=r_1$ y $R_2=r_2$.

Para hacer ésto, nótase que solo puede haber una corrida más de un tipo que de otro, o igual número de corridas, de acuerdo a la definición que se dió.

Supóngase entonces que $r_1=r_2+1$, esto quiere decir que la muestra de elementos de uno y otro tipo, comienza con una corrida de elementos del primer tipo, a las que se les denotará con a's, y terminará con otra corrida de a's (En caso de que $r_2=r_1+1$, la sucesión comenzará y terminará con corridas de elementos del segundo tipo, o b's). Para dividir las a's en r_1 grupos separados, divisiones para las b's son insertadas en r_1-1 lugares seleccionados de n_1-1 espacios entre las n_1 a's, lo que se puede hacer de $\binom{n_1-1}{r_1-1}$ maneras, análogamente para las b's, se tendrán $\binom{n_2-1}{r_2-1}$ formas de distribuir las corridas de b's en las divisiones antes hechas entre las a's. Entonces un arreglo que contenga r_1 corridas de - a's y r_2 de b's, se puede lograr de $\binom{n_1-1}{r_1-1}$ $\binom{n_2-1}{r_2-1}$ maneras.

Cuando $r_1=r_2$, la muestra comienza con a's y termina con b's o viceversa, y de nuevo la forma en construirla puede hacerse de $\binom{n_1-1}{r_1-1}$ $\binom{n_1-1}{r_2-1}$ maneras, pero como puede darse o uno (empezar con a's y terminar con b's) u otro caso (empezar con b's y terminar con a's), el total de maneras para formar una sucesión con r_1 corridas de a's y - r_2 corridas de b's, con $r_1=r_2$ es 2 $\binom{n_1-1}{r_2-1}$ $\binom{n_2-1}{r_2-1}$.

Por lo tanto:

$$P\{R_1=r_1, R_2=r_2\} = \begin{cases} \dfrac{2\binom{n_1-1}{r_1-1}\binom{n_2-1}{r_2-1}}{\binom{n}{n_1}}, & \text{si } r_1=r_2 \\[2ex] \dfrac{\binom{n_1-1}{r_1-1}\binom{n_2-1}{r_2-1}}{\binom{n}{n_1}} & \text{si } r_1=r_2+1 \end{cases}$$

así, la distribución del total de corridas $R=R_1+R_2$, bajo Ho, será

$$P\{R=r \mid Ho\} = \begin{cases} \dfrac{2\binom{n_1-1}{k-1}\binom{n_2-1}{k-1}}{\binom{n_1+n_2}{n_1}}, & r=2k, r_1=r_2=k. \\[3ex] \dfrac{\binom{n_1-1}{\frac{r-1}{2}-1}\binom{n_2-1}{\frac{r-1}{2}} + \binom{n_1-1}{\frac{r-1}{2}}\cdot\binom{n_2-1}{\frac{r-1}{2}-1}}{\binom{n_1+n_2}{n_1}}, & r_1=r_2+1 \text{ y } r=2k+1 \end{cases}$$

De aquí, que la hipótesis Ho; la muestra es aleatoria, se rechazará si $R < C_{\alpha/2}$ ó $R > C_{1-\alpha/2}$ si el nivel de significancia deseado es α, donde Cp es el p-ésimo cuantil de la distribución encontrada para R.

La distribución del número total de corridas fué dada por Ising en 1925. Wald y Wolfwitz (1940) usaron ésta distribución para el problema de dos muestras. Ellos también demostraron que si n_1 y n_2 son grandes, la estadística $\frac{R - 2n\frac{n_1}{n}\frac{n_2}{n}}{2\sqrt{n}\frac{n_1 n_2}{n^2}}$ se distribuyen asintóticamente normal.

Mood (1940) escribió un artículo donde menciona la historia de corridas de varios tipos y sus aplicaciones. Valores tabulados para la distribución de R, fueron dados por Swed y Eisenhart en 1943.

EJEMPLO 1 Un equipo profesional de beisbol tiene la siguiente secuencia de triunfos y derrotas en el mes de julio:

GPGGPGPPPGPPGGPGGPGPGPGPPGPGP

(G=ganó el juego, P= perdió el juego).

¿Puede decirse que el record de sus trinfos y derrotas es aleatorio, suponiendo que α=o.05?.

Lo primero que debe hacerse es contar el número de rachas o corridas que existen en la sucesión:

G P GG P G PPP G PP GG P GG P G P G P G PP G P G P

Como puede verse el número de corridas es R=22. Ahora, sea n el número de juegos ganados y m el número de juegos perdidos, entonces: n=14 y m=15. La región crítica está definida como:

$$\{\underline{x}: R < W_{\alpha/2} \text{ ó } R > W_{1-\alpha/2}\}.$$

Usando la tabla 6 del apéndice, se encuentra que:

$$W_{\alpha/2} = W_{0.025} = 10 \quad y \quad W_{1-\alpha/2} = W_{0.975} = 20 \quad (\text{haciendo } n=N_1 \text{ y } m=N_2)$$

Como $R > W_{1-\frac{\alpha}{2}}$, se concluye que el record de triunfos y derrotas del equipo no es aleatorio, esto es, existen tendencias.

EJEMPLO 2 Se tiene la siguiente lista de números:

15	77	01	64	69	58	40	81	16	60	20	00	84	22
28	26	46	66	36	86	66	17	49	85	40	51	30	10

¿Puede decirse que estos números provienen de una tabla de números aleatorios?.

Para contestar a esta pregunta, debe recordarse, en que consiste una tabla de números aleatorios. Una tabla de números aleatorios, es un conjunto de números generados por cualquier método que garantice los tres puntos siguientes:

i) Debe existir una proporción igual, para todos los dígitos. Esto quiere decir que hay tantos o's como 1's, 2's, 3's, etc.

ii) Si se toma una parte de la tabla, que sea suficientemente grande, estas proporciones deben mantenerse.

iii) Si se tiene una sucesión de estos números, el número siguiente no puede ser predecido.

Entonces para saber, si estos números son aleatorios, habría que checar, los tres puntos anteriores, lo cual no es posible (al menos el punto 1). Sin embargo si calcula la mediana de la muestra y se dividen en dos grupos de números (menores o mayores que la mediana), la prueba Wald- Wolfowitz, puede utilizarse. Podría usarse otro cuantil para dividir la población, pero en este caso, la mediana parece ser más adecuada para las tres condiciones mencionadas anteriormente.

Se procederá entonces a ordenar la muestra y calcular la mediana:

00	01	10	15	16	17	20	22	26	28	30	36	40	40
46	49	51	58	60	64	66	66	69	77	81	84	85	86

Como el total de observaciones es par, la mediana $\tilde{X}$ es: $\tilde{X}=\frac{40+46}{2}=43.$

Si se denota con $\underline{a}$ a los números menores a la mediana y con la $\underline{b}$ al complemento, la sucesión original quedará:

a b a bbb a b a b aa b aaa bb a bb a bb

a b aa

Teniéndose que el total de corridas es: T=19

Consultando la tabla 6, con n=m=14 y $\alpha=0.01$ se tiene que:

$W_{\alpha/2}=8$ y $W_{1-\alpha/2}=22$, por lo que la hipótesis nula de aleatoriedad no se rechaza, concluyéndose que los números pueden considerarse aleatorios.

EJEMPLO 3 Un fabricante calcula el costo promedio de producción de cierto artículo para cada uno de 44 meses, con los siguientes resultados:

(a todas las unidades se les restó 13, para facilitar los cálculos)

0.65	0.41	0.53	0.58	0.43	0.73	0.40	0.70	0.58	0.80	0.23
0.40	0.63	0.69	0.92	0.68	0.72	0.42	0.66	0.98	0.81	0.60
0.32	0.45	0.27	0.26	0.28	0.29	0.10	0.09	0.36	0.40	0.35
0.53	0.66	0.10	0.28	0.33	0.02	0.09	0.12	0.16	0.96	0.95

¿Existe alguna tendencia estadísticamente significativa en estos promedios?.

La pregunta establece la siguiente hipótesis:

> Ho: No existe tendencia estadísticamente significativa.
>
> Ha: Existe dicha tendencia.

Si la hipótesis Ho es rechazada, la respuesta a la pregunta es afirmativa, en caso contrario, negativa.

Puesto que la hipótesis nula es equivalente a decir que los números fueron generados por un mecanismo aleatorio, se usará la prueba Wald-Wolfowitz. Para ésto, se ordenará la muestra y se calculará la mediana, dividiendo a la muestra como en el ejemplo pasado (No necesariamente la mediana debe ser usada por dicotomizar a la muestra).

0.02	0.09	0.09	0.10	0.10	0.12	0.16	0.23	0.26	0.27	0.28
0.28	0.29	0.32	0.33	0.35	0.36	0.40	0.40	0.40	0.41	0.42
0.43	0.45	0.53	0.53	0.58	0.58	0.60	0.63	0.65	0.66	0.66
0.68	0.69	0.70	0.72	0.73	0.80	0.81	0.92	0.95	0.96	0.98

La mediana es $\tilde{X} = 0.425$. Las corridas que se generan son:

b	a	b	b	b	b	a	b	b	b	a
a	b	b	b	b	b	a	b	b	b	b
a	b	a	a	a	a	a	a	a	a	a
b	b	a	a	a	a	a	a	a	b	b

Por lo que T=15.

Si se usa $\alpha=0.01$, en la tabla 6, para n=m=12, se tiene: $W_{\alpha/2} = 6$ y

$W_{1-\alpha/2}$ = 18 por lo que Ho no es rechazada, teniéndose que no existe tendencia estadísticamente significativa en los promedios.

PRUEBA WILCOXON PARA UNA MUESTRA.

Otra prueba que se utiliza en el caso en que se tenga una muestra y que sirve para checar si la mediana de la población toma cierto valor, es la prueba de rangos de Wilcoxon, que se basa en la mediana de la muestra. Esta prueba no se verá en esta sección pues es un caso particular de la prueba de rangos de Wilcoxon para dos muestras, que se tratará en la sección siguiente.

CAPITULO III

PRUEBAS DE DOS MUESTRAS

Cierto tipo de problemas que son muy frecuentes en las ciencias sociales y en las biológicas, son aquellos en los que se desean comparar los efectos producidos por dos tratamientos aplicados a miembros de la población bajo estudio; o bien cuando se quiere determinar si un tratamiento es mejor que otro (sin que esto signifique que uno de ellos sea el mejor tratamiento posible), o simplemente comprobar la efectividad de un tratamiento. En estos casos la aplicación de cada tratamiento define a subconjuntos ajenos de la población, que pueden a su vez, ser considerados como poblaciones. Por ejemplo:

Suponga que se quiere averiguar si una droga A es más efectiva que una droga B. Para ésto, se toma un conjunto de ratones, a algunos de los cuales se les inyecta la droga A y a otros la droga B. De esta forma la población de ratones queda dividida en dos:

La población de ratones inyectados con droga A y la población de ratones con droga B.

Nótese que aunque originalmente se tenía una sola población, al aplicar los tratamientos se han generados dos, es por ésto, que se dice que se tienen dos poblaciones, cada una de las cuales recibe uno y sólo un tratamiento.

Como ya se sabe, no siempre es posible trabajar con toda la población, sino que se hace necesario el tomar una muestra aleatoria de ella a partir de la cual se podrán hacer inferencias de la población. En este caso se tendrán que tomar dos muestras, una de cada población; aunque, como se verá después, en ocasiones basta una población y seleccionar de ellas una muestra.

Si se denota con X_i a la i-ésima observación de la primera muestra - y con Y_i la observación i-ésima de la segunda, se pueden dar dos situaciones: que las muestras estén relacionadas, en cuyo caso - son del mismo tamaño ; o bien que sean independientes entre sí, y no necesariamente del mismo tamaño.

En general, se dice que las dos muestras estan relacionadas, cuando las parejas (X_i,Y_i) guarden alguna semejanza de acuerdo a ciertos factores que pueden influir al experimento, de forma tal que el efecto del tratamiento no se confunda con estos factores. La forma en que estas parejas se construyen es arbitraria, observando solamente que los elementos de la parejas sean semejantes en - el sentido antes expuesto, lo cual involucra además de la experiencia del experimentador, un conocimiento adecuado de los factores que puedan influir en los resultados. Un esquema de dos muestras relacionadas puede ser obtenido también, cuando se tiene una po - blación nada más; en este caso, la población sirve como su propio grupo control y las parejas pueden ser formadas como antes, sólo que X_i, denotará al individuo i-ésimo antes del tratamiento y Y_i - al mismo individuo después del tratamiento. Este es el mejor de los esquemas de muestra relacionadas, pues la semejanza entre X_i y Y_i es obvia, cosa que puede no suceder en otros casos. Un - ejemplo de esta situación es el siguiente:

Supóngase que se tienen 30 individuos, a los que se les mide su peso en kilogramos, y se les somete a una dieta durante 20 días, al- final de los cuales, quiere saberse si la dieta es efectiva para - bajar de peso. En este caso, sólo se tiene una población (la de - los individuos que se someterán a la dieta), pero pueden conside - rarse dos; la primera será la de los individuos antes de someterse a la dieta y la segunda, después de haber sido sometidos a la dieta (en este caso la dieta es el tratamiento); nótese que aquí cada individuo funciona como su propio control, es decir, el mismo grupo, es grupo control (antes de la dieta) y grupo tratado (después de la dieta), por lo que pueden generarse parejas (X_i, Y_i) en donde X_i denota el peso del i-ésimo individuo antes de la dieta y Y_i, el peso del mismo individuo después de la dieta. La diferencia de los pesos, puede dar una medida de que tan efectiva es la dieta.

Cuando las dos muestras son tomadas completamente al azar de las dos poblaciones, esto es, se toma una muestra aleatoria de la primera población y otra de la segunda, lo que se obtendrá son dos muestras independientes.

En ocasiones el estudio puede ser realizado usando cualesquiera de los dos esquemas, pero en algunas, sólo uno de los dos tipos puede ser usado. Para ilustrar estas situaciones, se tiene los siguientes ejemplos:

i) Supóngase que se tienen dos métodos de enseñanza, y se quiere determinar cual de los dos es mejor. Para esto es necesario el tomar dos muestras de estudiantes, en donde, a una de ellas se les enseña de acuerdo a uno de los métodos y al otro grupo o muestra se les enseña con el otro método (Aquí, cada población ha quedado definida por el metodo de enseñanza). Si se desea hacer el estudio por medio de muestras relacionadas, deberá tenerse que los estudiantes del primer grupo o muestra, tuvieran características semejantes a los del segundo grupo, como por ejemplo: que tuvieran la misma capacidad de aprendizaje. Sin embargo, determinar estas condiciones no es fácil; además podría llevar mucho tiempo, lo cual podría ser costoso, por otro lado, tratar de usar a cada individuo como su propio control es imposible, pues si ya se les enseñó con un método, al hacerlo con el otro, el aprovechamiento que se tuviera por el segundo, se confundiría con los conocimientos obtenidos al haber utilizado el primer método.

Entonces, esta situación puede estudiarse sólamente si se toman las dos muestras completamente al azar, obteniéndose de esta forma, dos muestras independientes.

ii) Una cierta variedad de planta ha sido estudiada durante cierto tiempo, y se ha encontrado que esta planta crece tanto en zonas montañosas, como en zonas localizadas dentro de un valle, en donde la composición del terreno y las condiciones climatológicas, son completamente distintas a la de las zonas montañosas, cuyo suelo, generalmente es rocoso. Por otro lado, una compañía de fertilizantes ha comprobado que tiene dos tipos de fertilizantes, el A y el B, que ayudan al desarrollo de este tipo de planta, sin embargo, desea saber, cual de esos tipos de fertilizantes es el mejor.

Para ello la compañía cuenta con una cierta área de terreno que esta constituida por varias zonas: unas de las cuales son montañosas y otras se ubican dentro del valle. Toda el área del terreno es dividida en n partes iguales, que se les llamarán unidades, y que comprenderán el total de la población bajo estudio. De estas n unidades, se eligen aleatoriamente r unidades ($r < n$) y nuevamente al azar se toman r_1 unidades ($r_1 < r$) y a ellos se les aplica el fertilizante A y a las $r-r_1$ unidades restantes el fertilizante B. Entonces, se tienen dos muestras aleatorias independientes, una de tamaño r_1 y la otra de tamaño $r-r_1$ (Nótese que las muestras no necesariamente son del mismo tamaño). Sin embargo puede suceder que el fertilizante A se aplique a más unidades que se encuentren en la zona montañosa, que a unidades que se encuentren en la otra zona, y que el tipo del suelo tenga que ver con los efectos del fertilizante. Esto puede suceder también para el fertilizante B, teniéndose que las características del suelo influirán en los efectos de los fertilizantes. Si esto sucediese, podría seleccionarse otras dos muestras independientes tratando de que cada una de ellas tenga el mismo número de unidades de un tipo que de otro, dentro de cada muestra; pero si al elegir las r unidades existen más de un tipo que de otro al tratar de que se tenga igual número de unidades de uno y otro tipo en cada muestra, sería forzar un poco el experimento, además de que se volvería más complicado y posiblemente no reflejaría la situación real de lo que se está estudiando. Otra solución que es mucho mejor que la planteada anteriormente, es dividir en dos partes semejantes cada unidad, y en cada una de ellas, aleatoriamente, distribuir los fertilizantes, uno en cada parte; formándose así dos muestras relacionadas.

iii) Si en el ejemplo anterior se supone que todo el terreno tiene las mismas características de suelo y climatológicas, se pueden seleccionar dos muestras independientes, o bien, formar dos muestras relacionadas.

Cuando cualquiera de las dos formas de tomar las muestras es posible, es recomendable usar dos muestras relacionadas, pues esto ayuda a evitar que factores externos al tratamiento, pero que pueden influir en los resultados del experimento, interfieran con los efectos del tratamiento, pudiéndose de esta forma detectar únicamente a estos últimos.

En esta sección se tratarán siete pruebas; tres de ellas son para el caso de muestras relacionadas y las otras cuatro cubren el caso en que las muestras sean independientes.

PRUEBA DE MC NEMAR

Como se dijo en la introducción a este capítulo, no es necesario el tener dos poblaciones para trabajar con un esquema de muestras relacionadas. Como ejemplo se ilustró una situación en la que un grupo de individuos eran usados como su propio control, para averiguar si una cierta dieta era o no efectiva. La Prueba de McNemar que se va a describir a continuación es particularmente apropiada para este tipo de problemas, en los que cada individuo es usado como su propio control, y en donde además cada observación puede ser clasificada en una de dos categorías ajenas, lo que permite el poder usar información referida en escala nominal al menos.

Como ejemplo se tiene el siguiente*:

Supóngase que un psicólogo infantil está interesado en la iniciación de los niños en los contactos sociales y ha observado que los niños recién llegados a una guardería suelen iniciar sus contactos personales con adultos antes que con otros niños. El supone que, con creciente familiaridad y experiencia, los niños iniciarán cada vez más contactos sociales con otros niños y no con adultos. Para probar esta hipótesis, observa 25 niños en su primer día en la guardería, y clasifica su iniciación en los contactos sociales de acuerdo con que haya sido con un adulto o con un niño. Después de que transcurre un mes, vuelve a observar a cada uno de los 25 niños haciendo la misma clasificación. Así, sus datos están clasificados de acuerdo a la siguiente tabla:

*Tomando de: S.Siegel (1976) Pag.88

		Individuos con los que se iniciaban el día 30	
		Adulto	Niño
Individuos con los que se iniciaban los niños en el primer día.	Adulto	a_{11}	a_{12}
	Niño	a_{21}	a_{22}

Es decir a_{11} niños mantenían sus contactos personales con adultos; a_{22} niños iniciaban sus contactos personales con niños y después de 30 días los seguían manteniendo de esa forma. Sin embargo a_{21} niños iniciaban sus contactos con niños y a los treinta días cambiaban por adultos. Análogamente a_{12}.

La hipótesis a probar es: Para cualquier niño que cambió de individuo en sus contactos personales la probabilidad de haber cambiado al adulto por un niño es la misma que cambiar al niño por un adulto y es igual a 1/2.

La hipótesis implícitamente desecha la información de los niños que no cambiaron de individuo, independientemente de con quien se hayan iniciado los contactos personales y establece que la probabilidad de cambiar de categorías es la misma para las dos y considera a estos eventos exahustivos, de forma tal que la probabilidad de cambiar de categorías es $\frac{1}{2}$.

En general, debido a que los casos en que no se observó cambio alguno no entran en el análisis de esta prueba, la información puede ser clasificada en dos categorías: La primera estará formada por los individuos que pertenecían a una categoría A y después del tratamiento cambiaron a la categoría B. Y la segunda estará formada por aquellos que pertenecían a la categoría B y cambiaron a la categoría A después del tratamiento. La hipótesis a probar sería,

$$Ho: P(x_i \in A; y_i \in B) = P(x_i \in B; y_i \in A) \forall i$$

en donde X_i, representa al individuo i-ésimo antes del tratamiento y y_i al mismo individuo después de haberle sido aplicado el tratamiento, y además los pares son internamente consistentes (Ver apéndice).

En vista del tipo de hipótesis y de como ha sido clasificada la información, la estadística que se utiliza es una estadística del tipo $\Sigma_i \frac{(O_i - E_i)^2}{E_i}$ (*) en donde O_i son los valores observados y E_i los esperados. Si se denota por a a el número de individuos que cambiaron de la categoría A a la categoría B; y por b los que cambiaron de la B a la A; se tendrá que:

$$O_1 = a, O_2 = b \quad y \quad E_{1,2} = \frac{a+b}{2}$$

lo que

$$\Rightarrow \Sigma_i \frac{(O_i - E_i)^2}{E_i} = \frac{(a - \frac{a+b}{2})^2}{\frac{a+b}{2}} + \frac{(b - \frac{a+b}{2})^2}{\frac{a+b}{2}}$$

$$\Rightarrow \Sigma_i \frac{(O_i - E_i)^2}{E_i} = \frac{(\frac{1}{2}(2a-a-b)^2 + (\frac{1}{2}(2b-a-b))^2}{\frac{a+b}{2}}$$

$$\Rightarrow \Sigma_i \frac{(O_i - E_i)^2}{E_i} = \frac{\frac{1}{4}(a-b)^2 \, \frac{1}{4}(a-b)^2}{\frac{a+b}{2}}$$

$$\therefore \quad \Sigma_i \frac{(O_i - E_i)^2}{E_i} = \frac{(a-b)^2}{a+b}$$

y de acuerdo a lo visto en el cap.II, $\frac{(a-b)^2}{a+b} \sim X^2_{(1)}$**

Recordando lo visto en el capítulo II, referente a la prueba X^2, - si la frecuencia esperada $\frac{a+b}{2}$ es menor que 5, se sabe que la prueba X^2, no es muy buena, por lo que es recomendable el usar la -

* Ver prueba X^2 bondad de ajuste cap.II.

** En virtud de que se esta usando una distribución continua para aproximar - una discreta, es recomendable hacer una corrección, quedando la estadística $(|a-b|-1)^2/a+b$, la cual se sigue distribuyendo $X^2_{(1)}$.

prueba binomial en lugar de la X^2, usando como estadística de prueba a b.

En caso en que la prueba X^2, sea aplicable, la hipótesis de rechazará si el valor $\frac{(a-b)^2}{a+b}$ excede el $(1-\alpha)$-cuantil, cuando la prueba se hace a un nivel de significancia α.

EJEMPLOS:

EJEMPLO 1. Supóngase que un experimentador está interesado en determinar si su película acerca de la delincuencia juvenil cambiará las opiniones de los miembros de una comunidad particular acerca de la severidad con que debe castigarse a los jóvenes que delinquen. Obtiene una muestra de 100 adultos de la comunidad y conduce un estudio de "antes y después" siendo cada sujeto su propio control. Les pide que tomen posición sobre si deberían castigarse más o menos a los delincuentes juveniles en relación a las prácticas penales actuales. Después de esto proyecta la película a los 100 adultos y repite la pregunta obteniendo la siguiente información:

		Fuerza de castigo favorecida después de la película	
		+	−
Fuerza de castigo favorecida antes de la película.	+	59	7
	−	8	26

Lo que se quiere averiguar es si la opinión sobre el cómo castigar a los delincuentes juveniles, se vió influenciada por la película. Es decir, las hipótesis que se desean probar son:

Ho: La opinión de las personas no cambia por la película.

Ha: La opinión se ve influenciada significatívamente, por la película.

Debido al diseño del experimento, la prueba McNemar, será utilizada, teniéndose que a=7 y b=8, así que a+b=15, por lo que la estadística de prueba queda como:

$$T= \frac{(a-b)}{a+b} = \frac{1}{15} \quad \therefore \; T= 0.0667$$

Buscando en la tabla 2, el cuantil de una X^2 con un grado de libertad, se observa que el valor de T es menor que todos los cuantiles que aparecen, por lo que la hipótesis Ho, no se rechaza, esto es, no existen cambios significativos debido a la película.

EJEMPLO 2 Antes de que los debates entre dos candidatos presidenciales de cierto país fueran televisados, una muestra aleatoria de 100 personas fué tomada y se les preguntó cual era el candidato de su preferencia. Ochenta y cuatro personas favorecieron al candidato A y las restantes al B. Después del debate, las cien personas expresaron su preferencia nuevamente. En esta ocasión, la cuarta parte de las personas que habían mostrado preferencia por el candidato A, mostraron ahora, preferencia por el candidato B. En forma similar, de la gente que favorecía al candidato B, una cuarta parte favoreció su opinión al candidato A, después del debate. La pregunta es: ¿Puede decirse que el debate influyó significativamente en la opinión de las personas?.

Es decir, si Ho: La opinión de las personas no cambió por el debate y

Ha: La opinión de las personas cambió por el debate,

la pregunta es: ¿Se rechaza la hipótesis nula? (Usando α=0.01).

La prueba Mc.Nemar, puede contestar a la pregunta.

La información que se dió, queda esquematizada como sigue:

		Preferencias de la Gente después del debate	
		Candidato A	Candidato B
Preferencias de la gente antes del debate	Candidato A	63	21
	Candidato B.	4	12

Así, la estadística de prueba T vale: $T = \frac{(21-4)^2}{25} = \frac{289}{25} = 11.56$.

Usando la tabla 2, el cuantil de una X^2 con un grado de libertad y $\alpha=0.01$ es: 6.635. Como $T > 6.635$, la hipótesis nula es rechazada, teniéndose que en efecto, el debate influyó significatívamente en la opinión de las personas.

EJEMPLO 3 El tiempo de reacción antes del almuerzo fué comparado con el tiempo de reacción después del almuerzo. En un grupo de 28 oficinistas veintidós de ellos, tuvieron un tiempo de reacción más pequeño antes del almuerzo, a cuatro de ellos les pasó lo contrario y únicamente dos de ellos nos presentaron cambios en su tiempo de reacción. ¿Puede decirse que el almuerzo provoca cambios significativos en el tiempo de reacción de los oficinistas?.

Análogo al caso anterior, se usará la preuba de Mc.Nemar, pues se tiene un esquema de dos muestras relacionadas, en donde cada persona en su propio control. Como en esta prueba, los elementos que no sufrieron cambios debido al tratamiento no entran para el análisis, no es necesario clasificar a las dos personas que no presentaron cambios en su tiempo de reacción, después del almuerzo. La forma en como está definida la estadística de prueba permite definir en forma arbitraria a:

a= 22 (Número de personas que tuvieron un tiempo de reacción más pequeño antes del almuerzo)

b= 4 (Número de personas que tuvieron un tiempo de reacción mayor antes del almuerzo, que después del almuerzo),

asi que T=12.416.

La hipótesis que se desea probar, con α=0.05, es:

Ho: El almuerzo no provoca cambios significativos en el tiempo de reacción de las personas.

Usando la tabla 2 y comparando T con el cuantil (1-α) de una X^2 con un grado de libertad (3.841), se tiene que la hipótesis Ho, es rechazada, diciéndose por tanto que el almuerzo causa cambios significativos en el tiempo de reacción de las personas.

PRUEBA DE SIGNOS.

Es frecuente que dados dos objetos, tratamientos o situaciones, - pueda decidirse cuando una alternativa es mejor o preferible a la otra, donde por efectos prácticos, no sea necesario establecer que tan mejor es una que otra. Así puede decirse que un automóvil es mejor que otro de acuerdo a la gasolina que gaste, o bien de acuerdo a la velocidad que desarrolle. Sin embargo, no siempre es inmediato el poder determinar cuando uno de los dos objetos a tratamientos, es mejor que el otro, sino que se hace necesario el tomar dos muestras relacionadas, y a partir de ellar tomar una decisión. - Por ejemplo, supóngase que un grupo de personas está interesado - en probar que dos vodkas A y B tienen la misma calidad. Para ésto, deciden seleccionar 25 personas aleatoriamente, a las que se les - dará a probar cada uno de las dos marcas de vodka. Se mide la preferencia con la siguiente escala: excelente, buena, regular, mala y pésima, y de acuerdo a esto se tendrán 25 respuestas para el - vodka A: $X_1, X_2, \ldots, X_{25}$ y 25 para el B: $Y_1, \ldots Y_{25}$, en donde (X_i, Y_i) denotará las preferencias del individuo i-ésimo, para cada uno de los vodkas. Esta escala define un orden: excelente, mejor (mayor)- que buena, buena mejor que regular, etc. pudiéndose decir que un- vodka es mejor que otro, si obtiene una calificación mayor que la del otro. De esta forma se ordenan los pares internamente:

i) $X_i < Y_i$, lo que significará que para el individuo i-ésimo, el vodka B es mejor que el vodka A.

ii) $X_i = Y_i$, ésto querrá decir que los dos vodkas tienen la misma calidad para el individuo i.

iii) $X_i < Y_i$ que quiere decir que el vodka A, para el individuo i, es mejor que el B.

En virtud de que sólo se ha tomado una muestra para poder justificar su suposición, y no a toda la población, el grupo de personas propone que la hipótesis a probar sea: $P(X_i < Y_i) = P(X_i > Y_i)\ \forall_i$, teniendo como alternativa $P(X_i < Y_i) < P(X_i > Y_i)\ \forall_i$ ó $P(X_i < Y_i) > P(X_i > Y_i)$ $\forall_i$, ya que supone que las parejas son internamente consistentes*. Esta alternativa lo que señala es que existe una preferencia por algún tipo de vodka. Si se proponen probar que $P(X_i < Y_i) < P(X_i > Y_i)$ $\forall_i$, contra $P(X_i < Y_i) > P(X_i > Y_i)$, entonces concluirán cual vodka es mejor. Sin embargo, aquí se desarrollará un procedimiento para verificar la primera hipótesis propuesta (no preferencia), y el procedimiento para verificar la segunda será análogo, lo único que varía es la región crítica (Conover 1971).

Nótese que la escala de medición en que la información debe ser dada, es tal que permita ordenar internamente las parejas, por lo que la escala de medición será al menos ordinal. Por otro lado como lo que interesa es detectar alguna preferencia en uno de los tratamientos u objetos, las parejas (X_i, Y_i) tales que $X_i = Y_i$, no son relevantes para el estudio*, lo que significa que sólo serán de interés las parejas (X_i, Y_i) tales que $X_i < Y_i$ ó $X_i > Y_i$. Ahora, si se denota con un signo mas (+) a las parejas (X_i, Y_i) tales que $X_i < Y_i$, y con un signo menos (-), a aquellas tales que $X_i > Y_i$, se tendrá que el número de signos más (+), marcará la preferencia del segundo tratamiento (cuyas respuestas están dadas por $Y_i, \ldots Y_a$) y análogamente para el otro tratamiento. Como sólo existen dos opciones: $X_i < Y_i$ ó $X_i > Y_i$, bajo Ho la probabilidad de cada una de éstas, es $\frac{1}{2}$; si se toma a n como el total de signos más (+) y signos menos (-) (recuerde que se eliminan del análisis los empates es decir, las parejas, tales que $X_i = Y_i$), se tiene que la probabilidad de obtener un signo más (+) se distribuye como una binomial con parámetros n y $\frac{1}{2}$, por lo que la decisión de aceptar o no la hipótesis Ho: $P(X_i < Y_i) = P(X_i > Y_i)$, se basará en esta distribución. Esta hipótesis se rechazará si el total de signos

*Ver apéndice.

más (+) es menor o igual a t ó si es mayor o igual a n-t, donde t es tal que $P(T < t) = \frac{\alpha}{2}$**, donde T es una variable aleatoria binomial con parámetros n y $\frac{1}{2}$ y α es el nivel de significancia de la prueba. En caso de que ninguna de estas dos cosas suceda, se aceptará Ho. Si $n > 20$ se puede usar $t = \frac{1}{2}(n + w_{\alpha/2}\sqrt{n})$, en donde w_q, representa el q-ésimo cuantil de una variable aleatoria con distribución normal con media cero y varianza uno.

Este mismo preocedimiento puede ser utilizado para probar la hipótesis:

$$Ho: E[X_i] = E[Y_i] \forall i \qquad \text{contra } H_1\ E[X_i] \neq E[Y_i]\ \forall i$$

o bien

Ho: la mediana de X_i es la misma que la de Y_i, $\forall i$ contra H_1: las medianas son distintas, $\forall i$.

Esta prueba, llamada prueba de signos por razones obvias, es bastante antigua (1710) y ha sido estudiada por Walsh (1951), Dixon (1953), Hodges y Lehmann (1956), Gibbons (1964) y Moses (1952).

EJEMPLOS

EJEMPLO 1 En una investigación que se hizo, para averiguar si el agregar 10 mcg de vitamina B12 por libra a una ración de ganado en engorda, tenía algún efecto, se tomaron ocho parejas de puercos donde a uno de ellos se les daba el alimento común y a otro el alimento con vitamina B12. El promedio diario de aumento de peso (en libras) están resumidos como sigue:

**Recuérdese que α puede ser aproximado o se pueden utilizar pruebas aleatorizadas.

	1	2	3	4	5	6	7	8
Ración								
Con Vitamina B12	1.60	1.68	1.52	1.64	1.75	1.56	1.78	1.60
Sin Vitamina B12	1.56	1.52	1.75	1.49	1.59	1.79	1.56	1.77

Para verificar que las dos poblaciones son iguales, puede utilizarse la prueba de signos, pues las suposiciones para que sea válida, son cumplidas.

La hipótesis a probarse es: Ho: La Vitamina B12 no tiene efecto en la engorda.

Si se asocian a las ocho parejas se tiene:

1	2	3	4	5	6	7	8
-	-	+	-	-	+	-	+

Por lo que la estadística de prueba, que es el total de signos +, es igual a 3.

La región crítica para este ejemplo es encontrada, usando la tabla 3, con $n=8$ y $p=\frac{1}{2}$. Suponiendo $\alpha=0.10$, se tiene que $t=1$, por lo que $n-t=7$. Como $1<3<7$, la hipótesis Ho no se rechaza, es decir, puede decirse que la vitamina B12 no tiene efectos significativos en la engorda.

EJEMPLO 2 A 22 consumidores de una tienda de abarrotes, les pidieron probar dos tipos de quesos distintos y decir cuál era el de su preferencia. Siete consumidores prefirieron un tipo, 12 prefirieron el otro tipo y 3 no tuvieron preferencia por ninguno. ¿Estos resultados indican una diferencia significativa en la preferencia?

Aunque este problema no se tienen los valores de las X_i's y de las Y_i's, las preferencias por uno u otro tipo de queso, pueden ser consideradas como signos, es decir, el total de signos de un tipo es 7, el del otro tipo 12 y 3 parejas fueron tales que $X_i = Y_i$.

Se puede tomar entonces que el total de signos + es 7 (sería análogo si se toma como 12), y que la hipótesis que se quiere probar es que no existe diferencia significativa en la preferencia.

Con la tabla 3 y haciendo $\alpha=0.05$ se tiene que $t=5$ y $n-t=14$ por lo que la hipótesis no es rechazada, queriendo decir esto que no existe diferencia significativa en la preferencia.

EJEMPLO 3 En el ejemplo 3 de la prueba de Wilcoxon, se menciona un estudio que se hizo, para determinar si los valores biológicos de proteinas de cacahuates crudos (P) y cacahuates tostados (R) eran significativamente diferentes.

Para esto se tienen las siguientes diez parejas (P,R):

(61,55)	(60,54)	(56,47)	(63,59)	(56,51)
(63,61)	(59,57)	(56,54)	(44,63)	(61,58)

Las hipótesis a probar son:

Ho: No existen diferencias significativas en los valores biológicos.

Ha: Si existen tales diferencias.

Usando la prueba de signos, se tiene que T=1, pues nueve de las diez parejas son tales que $X_i > Y_i$, y sólo una de ellas cumple con $X_i < Y_i$, por lo que se tiene únicamente un signo +.

Buscando en la tabla binomial (tabla 3), se tiene para n=10 que:

$t=1$ si $\alpha= 0.0214$ y que $t=2$ si $\alpha= 0.1094$

En el primer caso, no puede tomarse una decisión, pues se está en la frontera de la región, teniéndose que, cualquiera de las dos hipótesis son rechazadas. Se podría tomar más información y repetir la prueba para solucionar el problema. Para el segundo caso, se tiene que la hipótesis Ho es rechazada (Al igual que en el caso de la Wilcoxon). Parece ser, que la prueba no resultó muy buena, quizá debido a que se pierde información al considerar sólo el signo.

PRUEBA DE WILCOXON.

En la prueba de signos desarrollada anteriormente, la única información que se utiliza es la dirección de las diferencias de las parejas (X_i, Y_i), así si $x_i > Y_i$, a la pareja le corresponde un signos (-) y si $X_i < Y_i$, un signo más (+). Sin embargo si se considera además de la dirección, a la magnitud relativa de las diferencias, se puede construír una prueba más completa , en el sentido en que toma en cuenta más información. Por ejemplo:*

Supóngase que un psicólogo infantil desea comprobar si la asistencia al jardín de niños tiene algún efecto ó capacidad de percepción social de los niños. Califica la percepción mediante una evaluación de las respuestas de los niños a un grupo de cuadros que representan una diversidad de situaciones sociales; haciéndoles una serie estándar de preguntas acerca de cada cuadro. Con este procedimiento él puede obtener mediciones entre el 0 y el 100, inclusive, para cada niño.

Aunque el investigador está seguro de que un valor alto representa una capacidad de percepción social mayor, no lo está con respecto a que estas mediciones sean suficientemente exactas como para ser tratadas numéricamente, esto es, no está seguro de que un niño calificado con 60, sea doblemente perceptivo que otro que obtuvo 30, sin embargo está seguro de que la diferencia entre la calificación de 60 y otra de 40 es mayor que la diferencia entre una de 40 y otra de 30. Es decir, no puede asegurar la exactitud numérica de las diferencias, pero sostiene que son suficientemente significativas para poderse clasificar por orden de tamaño absoluto.

Para probar el efecto de la asistencia al jardín de niños con el procedimiento de medición antes mencionado, sobre la percepción social de los niños, el investigador consigue 8 pares de gemelos idénticos como sujetos. Al azar asigna un gemelo de cada par al-

* Tómado del Siegel (1976).

jardín de niños por un tiempo, mientras que el otro permanece - fuera de la escuela. Al final del plazo, se dá a los 16 niños la prueba de percepción social.

El procedimiento para probar estas hipótesis es como sigue: Si - X_i representa la calificación obtenida en la prueba de percepción social del gemelo que permaneció en casa y Y_i la del gemelo que - asistió al jardín de niños, sea $D_i = Y_i - X_i \; \forall i$, lo cual dará la dirección de las diferencias. Nótese que las D_i son variables aleatorias continuas e independientes. Ahora, bien, como interesa la magnitud relativa de las diferencias, se podrían ordenar a las - D_i's de acuerdo al valor que toman, pero como sólo interesa la magnitud y no el signo, se va a ordenar la muestra $|D_1|, |D_2|, \ldots, |D_8|$ y- no la muestra $D_1, \ldots D_8$. Considérese entonces la muestra ordenada $|D_{(i)}| < |D_{(j)}|$ si $i < j$. El subíndice encerrado en el paréntesis se le conoce como el rango asociado a la observación $D_{(i)}$, es decir el rango 1 se le asocia al par cuya D_i sea la mínima diferencia absoluta, el rango 2 se asocia al par cuya D_i sea la segunda diferencia mínima y así sucesivamente, hasta asociar el rango 8 a la pareja cuyo D_i sea la mayor en valor absoluto. Debido a que - lo que se está tratando de averiguar son las diferencias que existen entre las parejas, se suprimirán del análisis aquellas parejas cuyas diferencias sean cero, asignando rangos únicamente a - las parejas para las cuales $|D_i| > 0$. Por otro lado, si existen - dos o más parejas cuyas D's en valor absoluto son iguales, no existe razón alguna para asignarle un rango menor a una de ellas y una mayor a la otra, por tanto en éste y solamente en éste caso, se - les asigna a estas parejas el promedio aritmético de los rangos - que se les hubiera asignado si no hubiesen sido iguales (Esto es, si los rangos 2,3 y 4 pertenecen a tres pares tales que $|D_i|$ es - la misma para las tres, entonces el rango $\frac{1}{3}(2+3+4)=3$ será asignado a las tres parejas; este rango no necesariamente es entero).

En resumen, si se tienen las parejas (X_1,Y_1), (X_2,Y_2), ..., (X_n,Y_n) y suponiendo que de ellas, sólo $r<n$ son tales que $D_i=Y_i-X_i$ es distinto de cero. Entonces asóciese el rango 1 a la pareja cuya $|D_i|$ es la más chica, el rango 2 a la pareja cuya $|D_i|$ es la menor de las restantes, y así hasta llegar a asociar el rango r a la pareja que tenga la máxima $|D_i|$.

Ahora, si el tratamiento aplicado a las X_i's tiene los mismos efectos que el aplicado a las Y_i's es de esperarse que la suma de los rangos de las parejas cuya D_i es mayor que cero, debe ser aproximadamente la misma que la suma de los rangos de las parejas cuya D_i es negativa. Además debe tenerse que $P(D_i=c)=P(D_i=-c)$ para $c\in\mathbb{R}$, es decir, que no existe ninguna preferencia sobre las poblaciones, lo cual significa que la distribución de las D_i's debe ser simétrica. Esto puede explicarse de esta forma: las D_i's pueden definirse como Y_i-X_i o como X_i-Y_i, y la probabilidad de que $|D_i|$ tome un valor es la misma en los dos casos, lo que justifica que los rangos asociados a las D_i's sea en cuanto a magnitud, sin importar el signo. Por otro lado si $P(D_i=c)\neq P(D_i=-c)$, podría crearse un sesgo a la hora de tomar una decisión, pues se estaría dando más importancia a un tipo de diferencias que a la otra.

Si se define a R^+ como la suma de los rangos asociadas a las parejas cuya diferencia sea positiva y a R^- en forma análoga; bajo la hipótesis de que no existe diferencia entre los tratamientos, es de esperarse que $R^+=R^-$, entonces lo más común es tomar R^+, R^- o el mínimo de ellas, como estadística de prueba.

Conover (1971) usa como estadística de prueba a R^+ (que es análoga a tomar R^- debido a la simetría de las D_i's), mientras que Siegel (1976), usa al mínimo de R^+ y R^-.

En el caso de que se use R^+, la hipótesis se rechazaría si el valor de R^+ es muy grande (lo que significaría que R^- es muy chico y por tanto difieren mucho) o bien si el valor de R^+ es muy chico -

(lo que implica que R^- es muy grande y de nuevo difieren). Si la prueba se realiza a un nivel de significancia α, entonces la zona de rechazo sería: $R^+ > w_{1-\alpha/2}$ ó $R^+ < w_{\alpha/2}$, donde w_i es el i-ésimo cuantil de la distribución de R^+, para lo cual existen valores tabulados (Conover 1971). En caso en que $R^+ \varepsilon [w_{\alpha/2}, w_{1-\alpha/2}]$ la hipótesis se aceptaría.

Si la estadística de prueba es el min (R^+, R^-), entonces la hipótesis se rechazaría si el min (R^+, R^-) es muy chico (lo que querría decir que el otro valor es muy grande y por tanto R^+ y R^- difieren mucho). Este valor se compara con el α-cuantil (si la prueba se hace a un nivel α) de la distribución del mínimo de (R^+, R^-) (Siegel 1976) y si este cuantil es mayor o igual que el valor del min (R^+, R^-) entonces se rechazaría la hipótesis de que los efectos de los tratamientos son iguales.

La hipótesis de que los tratamientos son iguales pueden plantearse también como $d_{.50} = 0$, donde $d_{.50}$ es la mediana de la D_i's, ya que si $d_{.50}$ esto quiere decir que $P(X_i > Y_i) = P(Y_i > X_i) \; \forall_i$. En vista de la simetría, también puede probarse la hipótesis $E[X] = E[Y]$, usándose el mismo procedimiento que el mencionado anteriormente, pues si la distribución es simétrica entonces la media y la mediana coinciden. Este procedimiento se conoce como la prueba de rangos de Wilcoxon, quien la propuso en 1945 y halló la distribución de R^+, a partir de la cual es sencillo calcular la del mínimo; valores tabulados para R^+ fueron dados por Owen (1962) cuando el número de parejas con diferencias distintas de cero es menor o igual que 20. Cuando este número es menor o igual que 100, existen tablas dadas por Mc.Cornack (1965).

La prueba de Wilcoxon, también es útil en el caso en que se tenga una muestra $X_1, \ldots, X_n$ y se quiere probar si la mediana de la población es igual a un cierto valor m. Para esto, se forman las parejas $(X_1, m), \ldots, (X_n, m)$ las que son tratadas de igual forma que como se mencionó antes y se usa el procedimiento sin cambio alguno. Es decir se definen las variables D_i's como: $D_i = m - X_i \forall_i$, se quitan las parejas cuya D_i es cero, se asignan rangos y la región de rechazo es la misma que para el caso de dos muestras relacionadas.

Otros trabajos importantes basados en la prueba de Wilcoxon son: Bennett (1965) quien la extendió en el caso multivariado; Hollander (1970), la adoptó para probar paralelismo entre dos rectas de regresión.

EJEMPLOS

EJEMPLO 1 Muestras se sangre fueron tomadas de ocho pacientes. En cada muestra el contenido de Serum Albumen de la sangre fué determinado por dos métodos de laboratorio: A y B. El objetivo era descubrir, si existía una diferencia consistente, en la cantidad de Serum Albumen encontrada por los dos métodos. Las ocho diferencias entre las medidas fueron 0.6, 0.7, 0.8, 0.9, 0.3, 0.5, -0.5, 1.3. (Las unidades fueron gr. por 10 ml).

Del enunciado del problema se puede establecer las hipótesis que se desean verificar:

Ho: No existe una diferencia consistente significativa en la cantidad de Serum Albumen calculada por los dos métodos

y Ha: Si existe una diferencia consistente.

Puesto que se tiene una escala por intervalos, se usará la prueba de Wilcoxon, para lo cual se asignarán rangos a las diferencias absolutas (sin considerar signo)

Diferencia	0.6	0.7	0.8	0.9	0.3	0.5	0.5	1.3
Rango.	4	5	6	7	1	2.5	2.5	8

Como la estadística de prueba es la suma de los rangos asociados a las diferencias positivas, se tiene que:

$$T= 33.5$$

Buscando en la tabla 7, a un nivel $\alpha=0.05$, se tiene: $W_{\alpha/2}=4$ y $W_{1-\alpha/2}=32$, por lo que Ho es rechazada, teniéndose que existe una diferencia consistente significativa en la cantidad de Serum Albumen calculada por ambos métodos.

EJEMPLO 2 Dos investigadores desean encontrar si dos preparaciones de virus producirán diferentes efectos en plantas de tabaco. El método a emplear es el de frotar la mitad de una hoja de tabaco con una estopa mojada con una preparación de extracto de virus y después frotar la otra mitad de la hoja en forma similar, pero con una preparación del segundo extracto de virus.

La medición de potencia es el número de lesiones locales que aparecen en cada mitad de la hoja. Estas lesiones se aprecian como unos pequeños anillos oscuros y que fácilmente se pueden contar.

Se realizó un experimento con 8 hojas (que se consideran como 8 parejas de medias hojas) y los resultados que se observaron fueron:

Número de Lesiones en
mitades de hojas de tabaco

Pareja	Preparación 1	Preparación 2
1	31	18
2	20	17
3	18	14
4	17	11
5	9	10
6	8	7
7	10	5
8	7	6

Para probar si existen diferencias significativas entre las dos poblaciones, se usará la prueba de Wilcoxon (pues se cumplen los requerimientos para utilizarla), teniéndose como hipótesis nula:

Ho: No hay diferencias significativas entre las dos poblaciones;

y como alternativa: Ha: Si hay diferencias significativas.

El procedimiento requiere que se calculen las diferencias entre X_i (preparación uno) y Y_i (preparación 2). Estas son:

Pareja	1	2	3	4	5	6	7	8
Diferencia	-13	-3	-4	-6	1	-1	-5	-1
y los rangos correspondientes son	8	4	5	7	2	2	6	2

, por lo que la estadística de prueba es: $T=2$. Si $\alpha=0.01$, entonces $W_{0.005}=1$ y $W_{0.995}=35$, teniéndose que Ho no es rechazada (nótese que si se intercambian las X's y las Y's, la decisión es la misma), concluyéndose que no existen diferencias significativas. (Si se toma $\alpha=0.05$, puede concluirse que Ho es falsa. Lo que muestra la importancia de escoger α).

EJEMPLO 3 Mitchel, Burroughs y Beadles (1936), calcularon los valores biológicos de proteinas de cacahuates crudos (P) y cacahuates tostados (R), mediante un experimento con 10 parejas de ratas. Las parejas de datos (P,R) obtenidas son:

(61,55) (60,54) (56,47) (63,59) (56,51) (63,61)
(59,57) (56,54) (44,63) (61,58).

Usando el procedimiento de prueba de Wilcoxon, puede probarse:

Ho: No existen diferencias entre los valores biológicos de cacahuates crudos y tostados teniéndose como alternativa, la natural, es decir, que si existen dichas diferencias.

Si X mide los valores biológicos de proteinas de cacahuates crudos y r la de los tostados, se tienen las diferencias y rangos siguientes:

Pareja	1	2	3	4	5	6	7	8	9	10
Diferencia	+6	+6	+9	+4	+5	+2	-2	+2	19	-3
Rango	75	7.5	9	5	6	2	2	2	10	4

Entonces T=10, y si $\alpha=0.01$ se tiene que $W_{0.005}=4$ y $W_{0.095}=51$, rechazándose la hipótesis de que no existen diferencias.

EJEMPLO 4 Un comerciante desea saber si el número medio de artículos comprados en cada venta puede ser considerado como 10. Para esto, observa 12 compradores en las cajas registradores y cuenta el número de artículos que cada uno lleva. Los datos son los siguientes:

Comprador	No.de Artículos
1	22
2	9
3	4
4	5
5	1
6	16
7	15
8	26
9	47
10	8
11	31
12	7

Como se desea ilustrar la prueba Wilcoxon para verificar si la mediana de la población toma cierto valor, se forman las siguientes parejas.

(22,10), (4,10), (4,10), (5,0), (1,10), (16,10), (15,10), (26,10), (47,10), (18,10), (31,10), (7,10).

Calculando las diferencias $D_i = 10 - X_i$ y asignando rangos se tiene:

	1	2	3	4	5	6	7	8	9	10	11	12
D_i	-12	1	6	5	9	-6	-5	-16	-37	2	-21	3
R_i	9	1	6.5	4.5	8	6.5	4.5	10	12	2	11	3

Resultando entonces que T=25.

Las hipótesis son Ho:m=10 y Ha:m≠10, donde m es la mediana de la población.

Usando la tabla 7 se tiene que si $\alpha=0.05$, los cuantiles correspondientes son: $W_{a/2}=14$ y $W_{1-\alpha}=64$. Como T cae enmedio de esos dos valores, puede decirse que la mediana de la población es 10, a un nivel de significancia del 0.05.

PRUEBA X^2: DOS MUESTRAS INDEPENDIENTES

Cuando se tienen dos poblaciones en donde los individuos pueden ser clasificados en una y sólo una de k categorías (donde k es un entero positivo arbitrario) pueden hacerse dos preguntas que son de interés:

La primera de ellas es: ¿La probabilidad de que un elemento de la primera población pertenezca a una categoría es la misma que la probabilidad de que un elemento de la segunda población pertenezca a la misma categoría, y esto sucede para todas las categorías?. Dicho en otras palabras, si P_{1i} es la probabilidad de que un elemento de la primera población pertenezca a la categoría i, y P_{2i} lo análogo para la segunda población, la pregunta sería ¿$P_{1i}=P_{2i}$ $i \varepsilon J_k$

La segunda pregunta que se hace es sobre las clasificaciones: ¿Son las clasificaciones independientes?. Suponga que se tienen dos tratamientos químicos, el A y el B, y estos se aplican a 200 semillas cada uno, entonces las semillas pueden clasificarse en: germinaron y no germinaron, que esquemáticamente quedaría:

	Germinaron	No Germinaron
Semillas tratadas con A		
Semillas tratadas con B		

Si se llega a concluir que las clasificaciones son independientes (en este caso las clasificaciones fueron germinaron y no germinaron) querría decir que el tratamiento A y el B no influyen en la germinación de las semillas y que por tanto las 400 semillas pertenecen a una misma población.

Estas preguntas pueden ser generalizadas para el caso en que se tengan más de dos poblaciones, digamos $r (r > 2)$, además las técnicas estadísticas para resolverlas son las mismas para ambas casos, por lo que se tratarán posteriormente en la parte de r poblaciones independientes, $(r > 2)$ y en ese tratamiento se darán ejemplos del caso de dos poblaciones.

PRUEBA DE LA MEDIANA

La prueba de la mediana, como su nombre lo indica es un procedimiento que sirve para probar si dos poblaciones independientes tienen el mismo parámetro de localización; la mediana. Esta prueba es un caso particular de la prueba X^2, que además puede ser generalizada al caso de k poblaciones independientes, como se verá más adelante. Debido a su utilidad en el caso de dos poblaciones independientes, se tratará de forma general en esta parte, y se supondrá que los resultados que se obtengan para la prueba X^2 con k poblaciones independientes son válidos en este momento aunque aún no se hayan sido explicados. A continuación se verá un ejemplo para introducir la prueba*.

Supóngase que dos métodos diferentes de germinación de maíz fueron asignados aleatoriamente a diferentes parcelas, y que se quiere determinar si existe una diferencia entre la producción de maíz de cada parcela debido al método usado. Además se tiene la experiencia de que una diferencia en las medianas poblaciones puede ser interpretada como una diferencia en el método usado, por lo que se quiere probar si la mediana de la producción es la misma para los dos métodos.

Si se identifica con X_i a la población de la i-ésima parcela tratada con un método y y_j a la j-ésima parcela tratada con el otro método de germinación (Nótese que no necesariamente el número de parcelas que son tratadas con un método, es el mismo que el número de parcelas tratadas con el otro) y se supone que la hipótesis de igualdad de medianas es válida, deberá tenerse que el número de X_i's que sean mayores que la mediana de la muestra combinada sea el mismo que el número de X_i's menores a la mediana de la muestra combinada y lo mismo deberá suceder para las y_j's (De acuer

* Adaptado de Conover.

do a esto, la escala en la que se mida la información deberá ser al menos ordinal]. Esto es, si las dos poblacionen tienen la misma mediana, entonces la mediana que se calcula a partir de la muestra conjunta de X's y y's deberá ser la mediana de las X's y de las y's separadamente, lo que significa que habrá tantas X's (oy's) mayores que ese valor, como X's (oy's) menores que ese valor.

Si los datos se acomodan como se muestra en el cuadro, deberá tenerse, si la hipótesis es cierta, que $X^{o} = X'$ y $Y^{o} = Y'$ entonces se usa como estadística de prueba a la siguiente, la cual se deduce en el caso de la prueba X^2, por lo que no se verá aquí.

	X's	Y's
Número de valores mayores que la mediana de la muestra conjunta.	X^{o}	Y^{o}
Número de valores menores que la mediana de la muestra conjunta.	X'	Y'

$$T = \frac{N\left(|X^{o}y' - y^{o}X'| - \frac{N}{2}\right)^2}{(X^{o}+y^{o})(X'+y')(X^{o}+X')(y^{o}+y')}$$ * Esta estadística se distribuye según una X^2 con un grado de libertad. Si la prueba se hace a un nivel α de significancia, la hipótesis se rechaza si el valor T es mayor que el $(1-\alpha)$ cuantil de una distribución $X^2_{(1)}$, en otro caso la hipótesis no se rechaza.

EJEMPLOS

EJEMPLO 1 En un examen de la resistencia que al aplastamiento ofrecen los granos de una mazorca, se eligieron al azar dos lotes de 10 granos cada uno, con los siguientes resultados:

* N es el tamaño de la muestra conjunta.

Resistencia al Aplastamiento
(En puntos)

Lote I		Lote II	
8	18	8	20
14	20	15	20
16	22	16	24
17	27	18	28
18	30	20	31

¿Puede decirse que la resistencia media al aplastamiento en la misma para ambos lotes?. Aún más: ¿La mediana es la misma para ambas poblaciones? (Usando $\alpha=0.05$). Esto es: Ho: Las dos poblaciones tienen la misma mediana y Ha: Las dos poblaciones tienen diferentes medianas.

Debido a como la información está dada, y a la hipótesis que se desea verificar, la prueba de la mediana será utilizada. Para ésto las muestras serán ordenadas en forma conjunta sin perder identificación:

8(1)	8(2)	14(1)	15(2)	16(1)	16(2)	17(1)	18(1)	18(1)	18(2)
20(1)	20(2)	20(2)	20(2)	22(1)	24(2)	27(1)	28(2)	30(1)	31(2)

La mediana de la muestra conjunta es: 19. Entonces:

	Población 1	Población 2	
Observaciones menores o iguales a la mediana conjunta.	6	4	10
Observaciones mayores a la mediana conjunta.	4	6	10
Totales	10	10	20

Como $T = \frac{N(|x^\circ y' - y^\circ x'| - \frac{n}{2})^2}{(x^\circ + y^\circ)(x' + y')(x^\circ + x')(y^\circ + y')}$

y $N = 20$, $x^\circ = 6$, $y^\circ = 4$, $x' = 4$ y $r' = 6$ se tiene que:

$$T = \frac{20(|36-16|-10)^2}{(10)(10)(10)(10)}$$

$$\Rightarrow \quad T = 0.20$$

Como $\alpha = 0.05$, el cuantil correspondiente de una $X^2_{(1)}$ es: 3,841, por lo que la hipótesis nula Ho no se rechaza, esto es, las dos poblaciones tienen la misma mediana.

EJEMPLO 2 Se tomaron dos muestras de familias, una de familias rurales y otra de familias urbanas, con el objeto de estudiar las diferencias en compras de café de los dos grupos. Esta diferencia, se sabe, puede ser detectada por una diferencia en las medianas. A continuación se enlistan los datos obtenidos, en kilogramos, comprados anualmente por familia.

Rural			Urbana		
5.49	3.08	4.13	3.76	4.22	4.17
5.03	5.17	6.03	5.03	4.35	2.09
4.45	5.13	4.26	4.45	3.58	3.86
4.62			4.13	4.40	2.81

Lo que se quiere probar es Ho: Las medianas de ambas poblaciones son iguales.

Al igual que el problema anterior, se ordenará la muestra conjunta para calcular la mediana y poder generar una tabla de contingencia de 2X2:

2.09(2) 2.81(2) 3.08(1) 3.51(2) 3.76(2) 3.86(2) 4.13(1) 4.13(2) 4.17(2) 4.22(2) 4.26(1)
4.40(2) 4.45(1) 4.45(2) 4.62(1) 4.85(2) 5.03(1) 5.03(2) 5.13(1) 5.17(1) 5.49(1) 6.03(1)

Así:

	Familias Rurales.	Familias Urbanas.	
Observaciones menores o iguales a la mediana conjunta.	3	8	11
Observaciones mayores a la mediana conjunta.	7	4	11
	10	12	22

pues la mediana es 4.33

Por lo que $T=\dfrac{22(112-561-11)^2}{(10)(12)(11)(11)} \implies T=1.65$

Si $\alpha=0.01$, entonces Ho no es rechazada.

PRUEBA WALD-WOLFOWITZ.

Una prueba que es muy útil para el caso en que se tengan dos poblaciones independientes, es la prueba de corridas de Wald-Wolfowitz. La utilidad de esta prueba radica en que puede usarse para probar si las dos poblaciones difieren de alguna manera, ya sea en alguna medida (tendencia central, dispersión,etc) o en cualquier otra cosa (distribución de probabilidades). De esta forma puede usarse para probar una gran variedad de hipótesis alternativas (Aquí sólo se verá un caso pero puede verse Conover (1971) para ampliar),mientras que otras pruebas sólo sirven para probar si existen ciertas diferencias entre las poblaciones (p.e: La prueba de la mediana, prueba igualdad de medianas). Además el método es rápido y fácil de utilizarse, como se verá a continuación.

Supóngase que se desea probar un nuevo método de enseñanza primaria de una escuela. Para esto se seleccionan n_1 alumnos de 6º grado y se les enseña usando este método,al final de un período de 6 meses se decide hacer una exámen para ver si este nuevo método da resultados diferentes en los alumnos, que el método tradicional. Para poder llevar a cabo cierta comparación con los alumnos que siguieron el método tradicional, se toma una muestra de tamaño n_2 de alumnos de 6º grado que no llevaron el nuevo método (Nótese que las muestras son independientes entre sí).

Ya que se tienen los n_1+n_2 alumnos, se les aplica un examen de conocimientos para poder evaluar el aprendizaje obtenido con los dos métodos de enseñanza. Por lo pronto la directiva de la escuela quiere saber si el aprovechamiento de los alumnos difieren de método a método y ya después averiguar si uno es mejor que otro (ésto se puede hacer en una sola etapa, planteando una alternativa de una cola, sin embargo el ejemplo pretende ilustrar el método y es por eso que se supone una alternativa de dos colas).

La forma de medir el aprovechamiento será la tradicional, es decir, una escala de calificaciones entre 0 y 10.

Si el nuevo método no introduce algún efecto nuevo, es de esperarse entonces que el aprovechamiento de los alumnos sea el mismo, tanto para aquellos que siguieron el nuevo método, como para los que siguieron el método tradicional.

Si esto es cierto, la probabilidad de que un alumno obtenga cieta calificación será la misma, si aprendió con el nuevo método o con el tradicional, por lo que la hipótesis a probar sería $Ho: F_x(t)=F_y(t) \forall t$ contra la alternativa $H_a: F_x(t) \neq F_y(t)$ para alguna t, donde $F_x(.)$ es la distribución de probabilidades de obtener una cierta calificación si el alumno estudió con el método tradicional y $F_y(.)$ será la análoga, si es que el alumno estudió con el nuevo método.

La forma de verificar esta hipótesis, es la siguiente:

Si se ordena la muestra conjunta de los n_1+n_2 alumnos de menor a mayor, pero manteniendo identificados los elementos de una y otra muestra, se puede saber cuantas corridas de elementos de una y otra muestra se generan. Si la hipótesis Ho fuese cierta, es de esperarse que el número total de corridas no fuese muy chico, pues si este es el caso, significaría que una de las dos poblaciones tuvo mejor aprovechamiento.

De acuerdo a esto, la hipótesis Ho se rechazaría si T, el número total de corridas, es menor que el α-cuantil de la distribución de T (encontrada anteriormente) si es que el nivel de significancia usado es α.

Pero puede suceder que existan "empates" entre algunos valores de X_i y Y_j (las calificaciones obtenidas para los alumnos de una y otra muestra), es decir que $X_i=y_j$ para algunos i y j, lo cual produciría ambigüedades respecto a que elemento se pone primero en -

la muestra ordenada conjunta. Si esto pasa, se recomienda que se hagan todos los arreglos posibles (poniendo primero X_i y luego y_j, y en otro arreglo poner primero y_j y después X_i) lo cual generaría diferentes T's, digamos $T_1, T_2, \ldots, T_r$ (si es que existen varios empates), entonces la estadística de prueba T se calcularía como el promedio aritmético de los valores $T_1, \ldots, T_r$ que se encuentren en cada arreglo.

Otro método que se sugiere para estos casos es checar primero si todos los valores $T_1, \ldots, T_r$ son significativos (es decir, llevan al rechazo de la hipótesis), así los empates no causan ya problemas y la decisión es rechazar la hipótesis de igualdad de poblaciones. Pero si algunas T's son significativas y otras no, la decisión es más difícil. Lo que puede hacerse es calcular la probabilidad de que T tome cada uno de esos valores y el promedio de ellas servirá para decidir si se acepta la hipótesis Ho. Si dicha probabilidad es igual o menor que α, se rechazará Ho, en caso contrario se aceptará.

Si el número de "empates" es muy grande, entonces la prueba de Wald-Wolfowitz es inoperante, y tendrá que usarse otro método para verificar la hipótesis.

Nótese que si existen empates entre elementos de la misma muestra, el número de corridas T no se altera pues definen uno y sólo un arreglo, por lo que el problema sólo se dá en el caso en que existan empates entre elementos de las dos muestras.

EJEMPLOS

EJEMPLO 1 Lotes de 10 abejas fueron alimentados con dos tipos distintos de concentraciones de almibar, al 20% y al 60%, en una granja, a media milla de la colmena.

Al llegar a la columna, se les quitaron los sacos de miel a las abejas y la concentración del fluido se midió. En cada caso, hubo un decremento de la concentración del almibar. Los decrementos fueron (en %).

De la concentración del 20%

1.7 1.5 2.5 2.7 0.5 1.4 0.7 1.4 1.2 1.5

De la concentración del 65%

1.7 2.8 2.2 2.4 1.3 2.1 1.8 3.4 0.9 1.6

La hipótesis a probar es Ho: No existen diferencias significativas debida a la concentración de almibar; teniéndose como alternativa, a la natural.

Para verificarla, se ordenará la muestra, sin perder la identificación de los elementos (entre paréntesis).

0.5(1) 0.7(1) 0.9(2) 1.2(1) 1.3(2) 1.4(1) 1.4(1) 1.5(1) 1.5(1)
1.6(2) 1.7(1) 1.7(2) 1.8(2) 2.1(2) 2.2(2) 2.4(2) 2.5(1) 2.7(1) 2.8(2) 3.4(2)

Pudiéndose generar también el siguiente arreglo;

0.5(1) 0.7(1) 0.9(2) 1.2(1) 1.3(2) 1.4(1) (1.4(1) 1.5(1) 1.5(1) 1.6(2)
1.7(1) 1.8(2) 2.1(2) 2.2(2) 2.4(2) 2.5(1) 2.7(1) 2.8(2) 3.4(2)
1.7(2)

Para el primer arreglo, la estadística de prueba vale T=10 y para el segundo T=10, así que no importa que arreglo se considere. Tomando $\alpha=0.05$ y usando la tabla 6 del apéndice se tiene (usando los valores más cercanos $N_1=11$, $N_2=11$) que $W_{0.05}=8$. Como $T>W_{0.05}$, la - hipótesis Ho no es rechazada, por lo que no existen diferencias - significativas entre los dos grupos de datos

EJEMPLO 2 Los siguientes datos reflejan la ganancia de peso de - dos muestras de ratas hembras (24-28 días de edad) bajo dos dietas distintas.

Ganancia (en gramos)

Altas Proteínas			Bajas Proteínas	
134	161	97	70	132
146	107	123	118	
104	83		101	
119	113		85	
124	129		107	

Para probar si el tipo de dieta influye en la ganancia de peso en las ratas, se establece la hipótesis:

Ho: Ambas poblaciones son iguales

Si esta hipótesis es aceptada, puede decirse que la dieta no influye en la ganancia de peso; caso contrario, sí influye.

Se usará la prueba Wald-Wolfowitz, por lo que se ordenará la muestra conjunta:

70(2) 83(1), 85(2), 94(2), 97(2), 101(2), 104(1), 107(1), 107(2), 113(1) 118(2), 119(1), 123(1), 124(1), 129(1), 132(2), 134(1), 146(1) y 161(1).

Como existe un empate, la muestra también genera el siguiente arreglo ordenado:

70(2), 83(1), 85(2), 94(2), 101(2), 104(1), 107(2), 107(1) 113(1), 118(2), 119(1), 123(1), 124(1), 129(1), 132(2), 134(1), 146(1) y 161(1).

Puede verificarse que en ambos casos T=10. Si se usa un nivel $\alpha=0.10$, se tiene que para n=12 y m=7 el cuantil correspondiente es: $W_{0.10}=8$. Para ambas corridas, se tiene que $T>8$, por lo que la hipótesis nula no es rechazada, concluyéndose que $F_x(t)=G_y(t)\forall t$, donde $F_x(.)$ es la distribución de probabilidades de la población uno (la tratada con altas proteinas) y $G_y(.)$ la correspondiente a la población dos (bajas proteinas). Esto significa que ambas poblaciones son iguales.

EJEMPLO 3 Al comienzo del semestre en cierta escuela, los alumnos de primer ingreso fueron divididos en forma aleatoria, en dos grupos. A uno de ellos se les enseñó con un nuevo texto de estadística para aprendizaje programado, mientras que al otro se le dieron clases con los métodos tradicionales. Al final del semestre se les sometió a todos a un examen de estadística y se observaron las siguientes calificaciones:

227	55	184	174	209	271	63	19
176	234	147	194	14	151	184	127
252	194	88	248	165	235	53	151
149	247	161	206	171	147	228	101
16	99	171	89	292	99	271	179

¿Que puede opinarse del nuevo texto?. Es decir, ¿Existen diferencias entre un método y otro?. Esta pregunta puede escribirse como:

Ho: No existen diferencias significativas entre - los dos métodos.

Ha: Si existe tales diferencias.

En una segunda etapa, puede preguntarse, en caso de que Ho sea rechazada, cuál método es mejor que el otro. Por lo pronto, se probarán las hipótesis ya establecidas. Como las dos muestras son - independientes, puede usarse la prueba Wald-Wolfowitz, la cual dice que hay que ordenar la muestra conjunta, sin perder la identificación de cada elemento.

La muestra ordenada es: (Los números en paréntesis, indican a que grupo pertenece la observación).

14(2)	16(1)	19(2)	53(2)	55(1)	63(2)	88(1)	89(1)	99(1)	99(2)
101(2)	127(2)	147(1)	147(2)	149(1)	151(2)	151(2)	161(1)	165(2)	171(1)
171(2)	174(1)	176(1)	179(2)	184(1)	184(2)	194(1)	194(1)	206(1)	209(2)
227(1)	228(2)	234(1)	235(2)	247(1)	248(1)	252(1)	271(2)	271(2)	292(2)

Como puede notarse, existen varias observaciones empatadas entre - los dos grupos (recuérdese que las observaciones empatadas dentro de un mismo grupo no causan ningún problema). Deben considerarse todas las sucesiones posibles, es decir intercambiando las observaciones empatadas.

Estas sucesiones son 16 y algunas de elllas se enlistan a continuación.

14(2)	16(1)	19(2)	53(2)	55(1)	63(2)	88(1)	89(1)	99(1)	99(2)
102(2)	127(2)	147(2)	147(1)	149(1)	151(2)	151(2)	161(1)	165(2)	171(2)
171(1)	174(1)	176(1)	179(2)	184(1)	184(2)	194(1)	194(1)	206(1)	209(2)
227(1)	228(2)	234(1)	235(2)	247(1)	248(1)	252(1)	271(2)	271(2)	292(2)

T= 24

14(2)	16(1)	19(2)	53(2)	55(1)	63(2)	88(1)	89(1)	99(2)	99(1)
102(2)	127(2)	147(2)	147(1)	149(1)	151(2)	151(2)	161(1)	165(2)	171(1)
171(2)	174(1)	176(1)	179(2)	184(1)	184(2)	194(1)	194(1)	206(1)	209(2)
227(1)	228(2)	234(1)	235(2)	247(1)	248(1)	252(1)	271(2)	271(2)	292(2)

T= 27

14(2)	16(1)	19(2)	53(2)	55(1)	63(2)	88(1)	89(1)	99(2)	99(1)
102(2)	127(2)	147(1)	147(2)	149(1)	151(2)	151(2)	161(1)	165(2)	171(1)
171(2)	174(1)	176(1)	179(2)	184(1)	184(2)	194(1)	194(1)	206(1)	209(2)
227(1)	228(2)	234(1)	235(2)	247(1)	248(1)	252(1)	271(2)	271(2)	292(2)

T= 29

En forma análoga, se obtienen las 12 sucesiones restantes.

Los valores para T, el número total de corridas, que se encuentran son:

25	25	24	24	27	25	25	23
27	25	27	23	29	27	27	25

Como se tienen varias sucesiones, la estadística de prueba puede ser el promedio de las T's encontradas. Este promedio es:

$$T^* = 25.5$$

Usando la tabla 6, se tiene que la hipótesis nula no es rechazada si $\alpha \in \{0.05, 0.10, 0.01\}$, concluyendo que no existen diferencias significativas entre los dos métodos.

Por otro lado, puede verse que los valores 27 y 29 son significativos (es decir, llevan a rechazar Ho), si $\alpha = 0.10$.

PRUEBA DE MANN- WHITNEY

En las situaciones en que se tengan dos muestras independientes entre sí, no necesariamente del mismo tamaño, lo que se trata de averiguar es, si las dos muestras provienen de la misma población. - Por ejemplo:

Una compañía produce cable para telégrafo utilizando dos procesos diferentes. De cada proceso se extrae una muestra aleatoria y se mide la fuerza que soportan los cables antes de reventarse. Debido a que un proceso es más caro que el otro, se quiere determinar si los cables que se producen con uno y otro proceso tienen la misma resistencia, si esto es así, se tendrá entonces, que a la compañía le convendrá producir los cables con el proceso más económico.

Supóngase que de un proceso se tomó una muestra de tamaño n y que del otro (proceso 2) se tomó una muestra de tamaño m. Si se define a X_i como la fuerza que soportan los cables hechos con el proceso 1 antes de reventarse, $\forall i \varepsilon Jn$, y con y_j lo correspondiente para el proceso 2, $\forall j \varepsilon Jm$, la hipótesis que se desea probar es - $Ho: F_x(t) = F_y(t) \; \forall t$. contra $F_x(t) \neq F_y(t)$ para alguna t.

Si se toma la muestra conjunta de tamaño m+n, y se ordena de menor a mayor (sin perder la identificación de cada elemento), para que la hipótesis $Ho: F_x(t) = F_y(t) \; \forall t$, sea sostenible, debe espararse que el número de veces que una y_j precede a una X_i en el arreglo combinado de ambas muestras, no sea ni muy grande ni muy chico, pues si alguna de las dos cosas sucede, significaría que uno de los dos procesos produce cables con mayor resistencia.

Si se define $D_{ij} = \begin{cases} 1 \text{ si } y_j < X_i & i \varepsilon J_{n_1} \\ 0 \text{ si } y_j > X_i & j \varepsilon J_{n_2} \end{cases}$ entonces $U = \sum_i \sum_j D_{ij}$, dará el número de veces que una y_j precede a una X_i en la muestra conjunta or denada.

Si el valor de v es muy grande o muy chico la hipótesis $F_x(t)=F_y(t)\ \forall t$ se rechaza, en caso contrario será aceptada. Para determinar que significa el que v sea muy grande o muy chica, se necesita conocer la distribución de v, la cual no es muy fácil de obtener (ver Gibbons 1971) sin embargo existen valores tabulados de v; entonces si la prueba se realiza a un nivel α de confianza, la hipótesis $Ho:F_x(t)=F_y(t)\ \forall t$ se rechazará si $v<w_{\alpha/2}$ o si $v>w_{1-\alpha/2}$ donde w_p es el p-ésimo cuantil de la distribución de v.

En ocasiones es laborioso calcular el valor de v, sobre todo en los casos en que n y m son grandes, por lo que una forma alternativa se basa en los rangos asociados a los elementos de la muestra conjunta ordenada y se obtiene haciendo $v=\mathbb{R}-\frac{n(n+1)}{2}$ donde $\mathbb{R}$ es la suma de los rangos asociados a las X_i's.

Esta forma alternativa, es más fácil de calcularse además de que el problema de empates entre las X's y las y's lo resuelve al asignar el promedio de los rangos que les hubiera tocado si no hubiese empate, mientras que en el otro caso si existen empates entre las X's y las y's deberán considerarse todos los posibles arreglos y el promedio de v's que se obtengan será la estadística que se use. Sin embargo si se supone que las variables son continuas, entonces los empates no causarán ningún problema para calcular la distribución de v, pues la probabilidad de que haya empates es cero. Nótese que la escala de medición es al menos ordinal.

Este procedimiento también es útil en el caso de que se tenga idea de que la diferencia entre las dos poblaciones sea sólo una diferencia de localización, es decir, $F_x(t)=F_y(t+c)$, con $c\in\mathbb{R}$, y lo que se quiere probar es $F_x(t)=F_y(t)\ \forall t$, o sea $c=0$.

La hipótesis a probarse quedará entonces en términos de las esperanzas de las variables, si es que existen, y se pueden establecer como $E[x]=E[y]$. El procedimiento para verificar esta hipótesis, así como la estadística de prueba, son los mismos que en el caso anterior, pues si el número de veces que una Y precede a una X en el arreglo combinado, no es muy grande, ni muy chico, se tendrá que la distribución de Y es la misma que la de X. En caso de que el valor de ν fuese muy grande o muy chico, se puede concluir que las dos poblaciones son distintas en el parámetro de localización, y por tanto es cierto que $F_x(t) = F_y(t+c)$, con $c=0$.

Puede hacerse también pruebas considerando hipótesis de una sola cola (Conover 1971) o bien, probar si la diferencia entre las poblaciones es debida a la varianza (Conover 1971).

Este preocedimiento de prueba fue dado primero por Wilcoxon (1945) en el caso $n=m$, quien uso $\mathbb{R}$ como estadística de prueba. Después se extendió al caso $n \neq m$ por White (1952) y Van de Reyden (1950)- Mann y Whitney (1947) fueron los primeros en considerar el caso $n \neq m$ y el construír tablas de la distribución de ν y usarla como estadística de prueba. Tablas más extensas están dadas en Verdooren (1963) para $n,m < 25$ y en Milton (1964) para $n < 20$ y $m < 40$.

Más bibliografía puede encontrarse en Conover (1971) y Jacobson (1963).

EJEMPLOS

EJEMPLO 1 En un laboratorio se realizó una prueba para obtener la temperatura medio ambiente más agradable al hombre. Diez - hombres y diez mujeres opinaron acerca de la temperatura más confortable para cada uno de ellos. Los resultados fueron:

Hombres	74	72	77	76	76	73	75	73	74	75
Mujeres	75	77	78	79	77	73	78	79	78	80

Supóngase que α=0.01. Buscando en la tabla 8 del apéndice se - tiene que los cuantiles correspondientes a n=10 y m=10 son:

$$W_{0.025} = 24 \quad y \quad W_{0.975} = 76.$$

Como 13 < 24, la hipótesis nula se rechaza, teniéndose entonces - que la temperatura promedio no es la misma para ambos sexos.

EJEMPLO 2 En un estudio en el cual se quería verificar la teoría de equipotencialidad, Ghiselli comparó el aprendizaje de 21 ratas normales, con el aprendizaje de 8 ratas después de ser operadas - de lesiones corticales. Esto es, el número de ensayos para rea - prendizaje requeridos después de la operación por las 8 ratas, se compara con los ensayos de las 21 ratas normales, para saber si - existen diferencias en la velocidad de aprendizaje y reaprendizaje entre las ratas normales y las operadas. Esto significa que - la hipótesis que se desea probar es:

Ho: No hay diferencia entre las ratas normales y - las de lesiones corticales con respecto a la velocidad de aprendizaje y reaprendizaje.

La hipótesis alternativa será la natural: Si hay diferencia.

El número de ensayos requeridos para los dos tipos de ratas se dá a continuación:

Ratas Normales							Ratas operadas				
23	8	24	15	8	6	16	20	55	29	24	75
15	15	14	21	23	15	24	56	31	45		
15	21	15	18	14	22	15					

Como se va a usar la prueba Mann-Whitney (pues las suposiciones -

Suponiendo que estas temperaturas forman una muestra aleatoria de las poblaciones respectivas ¿Es la temperatura promedio la misma para hombres y mujeres?.

Debido a que las suposiciones necesarias para que la prueba Mann-Whitney pueda ser realizada, se aplicará a estos datos. Para ésto, supóngase que la muestra de hombres se representa con X's y la de mujeres, con y's (para trabajar con la notación desarrollada).

Las hipótesis a probar son entonces:

$$H_o: E(X) = E(Y) \text{ contra } H_a: E(X) \neq E(y).$$

Para probarlas es necesario ordenar las muestras en forma conjunta y asignar rangos. Este proceso es mostrado en la siguiente tabla:

X_i	Y_i	Rango	X_i	Y_i	Rango
72		1	76		10.5
73		3	77		13
73		3		77	13
	73	3		77	13
74		5.5		78	16
74		5.5		78	16
75		8		78	16
75		8		79	18.5
	75	8		79	18.5
76		10.5		80	20

Para calcular la estadística de prueba, primero debe calcularse la suma de los rangos asociados a la muestra de las X's. este valor es S=68. por lo que la estadística de prueba es T=13. $\left(T=S-\frac{n(n+1)}{2}\right)$

son satisfechas), se va a ordenar la muestra conjunta sin perder la identificación de los elementos, y se asignará los rangos correspondientes:

Obs:	6(1)	8(1)	8(1)	14(1)	14(1)	15(1)	15(1)	15(1)	15(1)	15(1)	15(1)	15(1)
Rango:	1	2.5	2.5	4.5	4.5	9	9	9	9	9	9	9
Obs:	16(1)	18(1)	20(2)	21(1)	21(1)	22(1)	23(1)	23(1)	24(1)	24(1)	24(2)	29(2)
Rango:	13	14	15	16.5	16.5	18	19.5	19.5	22	22	22	24
Obs:	31(2)	45(2)	55(2)	56(2)	75(2)							
Rango:	25	26	27	28	29							

Ahora, S=263 y como n=21 se tiene que T=32. Como $n > 20$, se tiene que usar la aproximación a la normal para los cuantiles, la cual está dada por:

$$W_p = \frac{nm}{2} + x_p \frac{\sqrt{nm(n+m+1)}}{12}$$, en donde W_p es cuantil deseado y x_p el cuantil correspondiente a una normal con media cero y varianza uno.

Entonces:

$$W_p = \frac{21(8)}{2} + x_p \frac{\sqrt{21(8)(30)}}{12}$$

$$\Longrightarrow \qquad W_p = 84 + x_p(5.9161)$$

Si $\alpha = 0.05$, se tiene que $W_{\alpha/2} = 72.40$ y $W_{1-\alpha/2} = 95.60$. Como $T < W_{\alpha/2}$, la hipótesis nula se rechaza, concluyéndose que existen diferencias significativas en la velocidad de aprendizaje o reaprendizaje, entre las ratas normales y las operadas.

EJEMPLO 3 Un experimento sencillo fue diseñado para ver si piedras que se encuentran en una cierta área (A), tienen el mismo grado de dureza que piedras que se encuentran en otra área (B). Para ésto, se obtuvieron dos muestras de 4 y 5 piedras en cada área y se rasparon para determinar su grado de dureza.

La piedra menos dañada se le consideró la más resistente. De esta forma, las nueve piedras se ordenaron de acuerdo a su dureza. El rango uno fué asignado a la piedra menos resistente y así sucesivamente.

Origen de la Piedra	A	A	A	B	A	B	B	B	B
Rango	1	2	3	4	5	6	7	8	9

¿Puede decirse que el grado de dureza de las piedras encontradas en el área A y en el área B, es el mismo? O dicho de otra forma: ¿Existen diferencias entre las dos áreas en cuanto a la dureza de sus piedras?.

Esta pregunta, establece las siguientes hipótesis:

Ho: No existe diferencia en el grado de dureza de las piedras de área A y B.
Ha: Si existe diferencia.

Como la información que se tiene son los rangos asociados a las observaciones, se utilizará la prueba de Mann-Whitney para verificar las hipótesis. Se considerará que las piedras de la área A son las observaciones $X_1, \ldots, X_n$.

La estadística de prueba toma, para estos datos, el valor de 1. Buscando en la tabla 8, se tiene que para $\alpha = 0.05$, $W_{\alpha/2} = 2$ y $W_{1-\alpha/2} = 18$, rechazándose por tanto Ho. (Si $\alpha = 0.01$, la decisión es no rechazar Ho).

APÉNDICE

En este apéndice se trataron dos puntos que están muy conectados con las pruebas de dos muestras.

El primero de ellos, es el problema de los empates. Se recordará, que en la prueba de McNemar, las observaciones que no sufren cambios después del tratamiento, no entran en el análisis. De igual forma en la prueba de signos, si $x_i=y_i$ no se asocia signo y por tanto, esa pareja no participa dentro de la estadística de prueba. En la prueba de signos, cuando hay empates, lo que se hace es asociar el rango promedio de los rangos que les hubiera tocado, si las observaciones no estuvieran empatadas, etc.

Cuando el número de observaciones empatadas es pequeño a juicio de la persona que hace el análisis, no existe problema y la prueba puede ser llevada a cabo (Siegel(1976), propone modificaciones a las pruebas en caso de empates). Sin embargo, si el número de empates es muy grande, la prueba puede resultar inoperante y lo que se recomienda es tomar una nueva información para realizar la prueba, si esto no es posible, el análisis puede ser llevado a cabo, pero debe mencionarse el número de empates, para que la persona que interprete las conclusiones, tenga en cuenta esa información.

El segundo punto a considerar en este apéndice, es el siguiente: En varias pruebas se pedía que las parejas fueran "internamente consistentes". Al pedir esto, lo que está buscando, es que los efectos del tratamiento no sean debidos a la constitución de la pareja, sino al tratamiento en sí. Por ejemplo en la prueba de signos, esta suposición significa que si $P\{X_i>Y_i\}>P\{X_i<Y_i\}$ para alguna i, entonces $P\{X_i>Y_i\}>P\{X_i<Y_i\}$ debe suceder para toda i.

CAPITULO IV

PRUEBAS DE K MUESTRAS

Hasta este momento, se han desarrollado pruebas para checar si cierta información o cierta población tiene asociada una distribución de probabilidades, ya sea que ésta esté completamente especificada o no. También se ha visto si dadas dos poblaciones, estas son iguales, o difieren en algo. Así mismo puede ya probarse, cuándo dados dos tratamientos, uno es mejor que el otro.

Sin embargo existen situaciones en las que es necesario probar, si dados k tratamientos ($k > 2$), producen o no los mismos efectos, esto es, si se tienen k poblaciones, la pregunta es: ¿soniguales?.

Por ejemplo, pueden tenerse dos tratamientos que se apliquen a dos grupos de personas y tener un tercer grupo, que funcionará como un grupo control. Para averiguar que tratamiento es mejor y que tan mejor, se necesitaría un procedimiento para trabajar con tres poblaciones, que hasta ahora no se ha desarrollado.

Otro ejemplo es aquél en el que se tienen varias marcas de automóviles, y se desea averiguar cual de ellas es mejor.

Puede pensarse en el porqué de desarrollar procedimientos de prueba para k poblaciones y no utilizar los ya desarrollados para dos poblaciones, probando igualdad de poblaciones por parejas. Es decir, supóngase que se tienen cuantro poblaciones A,B,C y D; y quiere probarse que las cuatro son iguales. Si se quieren utilizar los métodos de dos poblaciones ya vistos, tendrían que hacerse seis pruebas de igualdad de poblaciones; éstas serían;

i) entre A y B
ii) entre A y C
iii) entre A y D
iv) entre B y C
v) entre B y D
vi) entre C y D

El primer argumento en contra de este procedimiento es lo laborioso que resulta. En el caso de cuatro poblaciones, tienen que hacerse seis pruebas; si se tienen cinco poblaciones, el número de pruebas asciende a diez; en general, si se tienen k poblaciones, el número de pruebas que tienen que efectuarse es igual a las combinaciones de k tomadas de dos en dos, esto es $\binom{k}{2}$, de donde si k es grande, el número de pruebas aumenta considerablemente.

Otro argumento en contra de este procedimiento, es la posibilidad de generar contradicciones. Volviendo al ejemplo de las cuatro poblaciones; puede llegarse a concluir que las poblaciones A y B sean iguales, que A y C también lo sean, pero que sin embargo se llegue a rechazar el que las poblaciones B y C sean iguales lo cual generaría una contradicción. Situaciones como estas se dan con mucha frecuencia en estos casos.

Un último argumento en contra de usar pruebas de dos poblaciones en el caso en que se tengan k poblaciones , es que puede demostrarse tanto teórica como prácticamente, que la probabilidad de error tipo 1 en el caso de usar pruebas de dos poblaciones, aumenta.

Estos argumentos justifican entonces, el desarrollar procedimientos de prueba para el caso en que se trabajen más de dos poblaciones.

Al igual que en el caso de dos poblaciones, se tienen pruebas para cuando se tengan poblaciones independientes, así como también para el caso de poblaciones relacionadas. La diferencia y uso de estas pruebas es análoga que el caso de dos poblaciones.

En esta sección se verán dos pruebas para el caso relacionado y tres para el caso de muestras independientes.

PRUEBA Q DE COCHRAN.

La prueba Q de Cochran, es la extensión a k poblaciones ($k > 2$) dela prueba de Mc Memar para dos muestras relacionadas.
Este procedimiento se utiliza en el caso en que se tengan más dedos poblaciones relacionadas y lo que prueba es igualdad de poblaciones. Un ejemplo en donde este método es aplicable es el siguiente:

Una compañía que produce pesticidas, tiene cinco procesos distintos para producirlos y quiere averiguar si los pesticidas elaborados por estos cinco métodos tienen la misma efectividad, en cuyo caso sólamente producirá el pesticida cuyo proceso de producción sea el más ecónomico. La efectividad será medida de acuerdo a si la planta a la cual es aplicado el pesticida sobrevive a un cierto tipo de plaga.

Para poder llevar a cabo el experimento se seleccionaron aleatoriamente 25 parcelas de un plantío. Como las 25 parcelas no son homogêneas entre sí (algunas son más grandes que otras, unas estánlocalizadas en ciertos terrenos de diferente constitución que - otras), se decide que cada parcela será dividida en seis partes - semejantes y en forma aleatoria se aplicarán los pesticidas a cinco de ellas y la restante quedará como control. El motivo de dejar a una porción del terreno de la parcela como control, es para poder establecer si el pesticida es o no efectivo. Si no se usara este control, y si se acepta la igualdad de tratamientos, puede pasar que todos los tratamientos sean igualmente malos.

En el caso en que se acepte la igualdad de tratamientos y se incluya al grupo control, lo que se intuye es que los tratamientos no son eficaces, pues la planta sobrevive de igual forma, aplicándolde pesticida o no.

Si se rechaza la igualdad, puede establecerse si un tratamiento es mejor que otro y además si es bueno, en el sentido en que sí sirve para detener la plaga; usando al grupo control.

Si las porciones de tierra sobreviven a cierto tipo de plaga, se dirá que el pesticida es efectivo, en caso contrario se dirá que no lo es.

En este ejemplo puede notarse varias cosas: Una de ellas es que se tienen seis poblaciones; la tratada con el pesticida elaborado con el proceso 1, con el proceso 2, y así sucesivamente hasta llegar a la tratada con el pesticida elaborado con el proceso 5.

Por último, la sexta población es la que no fué tratada con algún pesticida.

Puede notarse también, que se han generado poblaciones relacionadas, pues cada parcela ha sido tratada por los cinco diferentes tipos de pesticida y se ha dejado una porción sin tratar. De esta forma la relación que guardan entre sí, es obvia por parcela. Se denotará con 1 si el pesticida es eficiente y con 0 si no lo es.

La eficiencia del pesticida será medida de acuerdo al número de plantas que sobreviven a la plaga dentro de cada parcela; si el 75% de ellas lo hace, se dice que el pesticida es efectivo, en caso contrario, no lo es.

De acuerdo a esto puede hacerse un arreglo rectangular, en donde, por un lado aparezca el número de la parcela tratada, y por otro, el tratamiento usado.

El arreglo quedaría como:

	1	2	3	4	5	6
1	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}
2	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}	X_{26}
3	X_{31}	X_{32}	X_{33}	X_{34}	X_{35}	X_{36}
.						
.						
.						
25	X_{251}	X_{252}	X_{253}	X_{254}	X_{255}	X_{256}

en donde X_{ij} es uno o cero y es el resultado obtenido al aplicar el j-ésimo pesticida a la i-ésima parcela.

La hipótesis a probar es entonces, que los pesticidas dan el mismo resultado, contra que no lo dan.

Sí se define a $p_{ij}= P\{X_{ij}=1\}$; la igualdad de efectividad de pesticidas puede escribirse como:

$p_{i_1}=p_{i_2}=\ldots=p_{i_6}$ $\forall i\varepsilon J_{25}$, esto es las hipótesis quedarían como:

$$Ho: p_{i_1}= p_{i_2}=\ldots= p_{i_6} \ \forall i\varepsilon J_{25}$$

$$Ha: p_{i_j} \neq p_{i_k} \text{ para alguna } j\neq k \text{ y alguna } i\varepsilon J_{25} (j,k\varepsilon J_6)$$

Nótese que no se está probando el que los tratamientos (en este caso, los pesticidas) sean igualmente efectivos, en todas las parcelas, sino sólo por parcela. Esto quiere decir, que los tratamientos pueden ser más efectivos en algunas parcelas que en otras, donde tal vez la constitución de la parcela y el medio ambiente que la rodea influya en la efectividad.

Ahora bien, como se verá un poco más adelante, la estadística de prueba está basada en promedios, lo cual hacen que puedan existir ambigüedades en cuanto a que la efectividad del tratamiento puede deberse a la constitución de la parcela o al mismo tratamiento o a una combinación de ambas cosas.

Pero, si la hipótesis Ho es aceptada, lo que si se puede asegurar es que, todos los tratamientos son igualmente efectivos, sin importar las causas.

Sean $R_i = \sum_i X_{ij}$ y $C_j = \sum_j X_{ij}$, es decir, R_i es total del renglón i-ésimo y C_j el total de la j-ésima columna. (Supóngase que existen n columnas y m renglones, es decir, existen m muestras, compuestas por n elementos. Esto se hace para que el desarrollo que sigue, sea general).

Como X_{ij} es cero o uno, quiere decir que X_{ij} se distribuye bernoulli, entonces $C_j = \sum_j X_{ij}$ será una variable aleatoria. Sin embargo, como se dijo antes, la hipótesis Ho no establece el que $p_{ij} = p_{kj} \; \forall \; i,k \in Jm$, por lo que C_j no es una variable aleatoria binomial. Pero se quiere probar que los tratamientos son iguales, entonces, en forma intuitiva, debe probarse que en promedio las columnas son iguales. Para esto se necesita conocer la distribución de C_j, para encontrarla se utiliza el teorema central del límite:

$$\frac{C_j - E(C_j)}{\sqrt{V(C_j)}} \sim N(0,1) \quad \text{(aproximadamente)}$$

Sin embargo, las C_j no son independientes*, pero ha sido probado (Tate y Brown, 1970) que si m es grande, puede suponerse independencia, por lo que la variable:

*Esto es debido a que las X_{ij} están relacionadas por renglón.

$$\sum_{j=1}^{m} \frac{(C_j - E(C_j))^2}{V(C_j)}$$ se distribuye aproximadamente como una X^2 con m grados de libertad.

Como $E(C_j)$ y $V(C_j)$ no se conocen, pueden estimarse perdiéndose así un grado de libertad (Blomqvist 1951), los estimadores son:

$$\hat{E}(C_j) = \frac{1}{m} \sum_{j=1} C_j \quad y \quad \hat{V}(C_j) = \frac{1}{m(m-1)} \sum_{i=1}^{n} R_i(m - R_i).$$

Entonces la estadística de prueba quedará como:

$$Q = m\ (m-1)\ \frac{\sum_{j=1}^{m} (C_j - \bar{C})^2}{\sum_{i=1}^{n} R_i\ (m - R_i)}$$

o equivalentemente:

$$Q = \frac{m(m-1) \sum_{j=1}^{m} C_j^2 - (m-1)(\sum_{j=1}^{m} C_j)^2}{m \sum_{i=1}^{n} R_i - \sum_{i=1}^{n} R_i^2}$$

Si Q es mayor que el $(1-\alpha)$- cuantil de una X^2 con m-1 grados de libertad, la hipótesis Ho se rechaza; en caso contrario se acepta (Suponiendo que la prueba se realiza a un nivel α).

Para terminar, basta agregar que en el caso en que m=2 se tiene la prueba de Mc Memar, que como se señaló anteriormente, la prueba Q es la generalización de dicho procedimiento (Conover 1971).

Para ampliar lo visto aquí, puede consultarse: Fleiss (1966) quien dá otro modelo equivalente; Cochran (1950), que es donde este procedimiento se propone y Mc Nemar (1955), donde se discute esta prueba.

EJEMPLOS

EJEMPLO 1 La efectividad relativa de tres técnicas diferentes de ventas, fué probada en 12 amas de casa, que se prestaron como voluntarias. A cada ama de casa se le trató de vender un cierto producto, de acuerdo a las tres técnicas. Al final de las exposiciones del vendedor, cada ama de casa calificó con un uno a cada técnica si es que ella sentía que hubiera comprado el producto y con un cero, si no lo hubiera comprado. Los resultados del experimento son:

Ama de Casa	1	2	3	4	5	6	7	8	9	10	11	12
Técnica 1	1	1	0	0	0	1	1	0	1	1	1	1
Técnica 2	1	1	1	1	1	0	0	0	1	1	0	1
Técnica 3	0	1	1	0	0	0	0	0	1	0	0	1

Lo que se desea averiguar es si se puede considerarse que las tres técnicas tienen la misma efectividad. Esto es, la hipótesis que se desea probar es:

Ho: La efectividad de las tres técnicas es la misma.

contra Ha: La efectividad de las tres técnicas no es la misma

Esta hipótesis se quiere probar a un nivel de significancia $\alpha=0.05$.

Como se tienen más de dos tratamientos y la información está dada en forma dicotómica, se utilizará la prueba de Cochran. (De acuerdo al desarrollo de la prueba, en este ejemplo, las columnas y los renglones están intercambiados pues en la exposición, los renglones eran los objetos tratados y las columnas los tratamientos).

Para esto hay que calcular $C_j, j \in J_3$ y $R_i, i \in J_{12}$. De acuerdo a la definición se tiene que:

$C_1=8$, $C_2=8$ y $C_3=4$ y $R_1=2$, $R_2=3$, $R_3=2$, $R_4=1$, $R_5=1$, $R_6=1$, $R_7=1$, $R_8=0$, $R_9=3$, $R_{10}=2$, $R_{11}=1$ y $R_{12}=3$.

Entonces:

$$Q = \frac{3(2)(10.6666)}{16} \Rightarrow Q = 3.9999$$

El $(1-\alpha)$ cuantil de una X^2 con dos grados de libertad es:

$$W_{1-\alpha} = 5.991$$

teniéndose que Ho no es rechazada pues $Q < W_{1-\alpha}$, por lo que pueden considerarse a las técnicas como igualmente efectivas.

EJEMPLO 2 Cada uno de tres aficionados al basketbol, ha desarrollado su propio sistema para predecir los resultados de los juegos que se realizan. Doce juegos fueron seleccionados aleatoriamente y cada aficionado presentó su predicción para cada juego. Después de terminados los partidos, los resultados fueron tabulados. Se usó un 1 si la predicción fué acertada y 0 en caso contrario, teniéndose los siguientes resultados:

Aficionado		1	2	3
Juegos	1	1	1	1
	2	1	1	1
	3	0	1	0
	4	1	1	0
	5	0	0	0
	6	1	1	1
	7	1	1	1

8	1	1	0
9	0	0	1
10	0	1	0
11	1	1	1
12	1	1	1

Lo que se quiere saber, es si los tres aficionados son igualmente efectivos en su habilidad para predecir los resultados de los juegos.

La hipótesis nula quedará entonces como:

Ho: La efectividad en las predicciones es la misma para los tres aficionados,

y la alternativa

Ha: Las efectividades en las predicciones difieren de aficionados a aficionado.

La hipótesis y el tipo de información hacen que la prueba Q de Cochran sea la utilizable, para probar estas hipótesis. Por lo que la estadística de prueba es:

$$Q = \frac{\frac{m(m-1)}{n} \sum_{j=1} (C_j - \bar{C})^2}{\sum_{i=1} R_i \ (m - R_i)}$$

en donde $m=3$, $n=12$, $C_1=8$, $C_2=10$ y $C_3=7$.

Entonces:

$$Q = \frac{6(4.667)}{10} = 2.8002$$

Si $\alpha=0.01$, el cuantil correspondiente es: $W_{1-\alpha} = 9.210$. Por lo que la hipótesis Ho no se rechaza, es decir, la efectividad en las predicciones es la misma para los tres aficionados.

PRUEBA DE FRIEDMAN

La prueba Q de Cochran desarrollada anteriormente, sirve únicamente cuando los datos están dados es forma dicotómica, o la información puede clasificarse en dos categorías exluyentes. Sin embargo, existen problemas en los que la información obtenida de un experimento o situación, está dada en forma ordinal al menos, y el tratar de dicotomizarla puede hacer que se pierda parte de ella. Por esto, es deseable el contar con una prueba que haga uso de esta información, y sea útil para probar igualdad de poblaciones. La prueba que se verá en este momento, cumple con los requisitos arriba mencionados, y se llama prueba de Friedman, debido a que él la introdujo en 1937.

Antes de dar el método y siguiendo los lineamientos de estas notas, se verá un ejemplo, en el cual la prueba de Friedman, es aplicable. El ejemplo es el siguiente;

Se tiene cuatro tipos de alimentos vitamínicos para puercos y quiere verse si dan o no el mismo resultado en el crecimiento de ellos. Es decir, quiere averiguar si el crecimiento de los puercos varía por el tipo de alimentación.

Para estos se seleccionan al azar seis camadas de puercos con cinco puerquitos cada una, y aleatoriamente a cuatro de ellos se les alimenta con cada uno de los alimentos vitamínicos, y al restante se le da alimentación usual*. Para evitar efectos de sobrealimentación en algunos puerquitos, las cantidades de alimentos son las mismas para los treinta.

Después de dos meses de ser alimentados, se pesarán a los puerquitos y la información que se obtenga se arreglará de la forma siguiente;

*Este grupo alimentado con alimento usual, será el grupo de control, cuya importancia se vió en la prueba Q de Cochran, dada anteriormente.

	Alimento Tipo 1	Alimento tipo 2	. . .	Alimento Usual
CAMADA 1	X_{11}	X_{12}		X_{15}
CAMADA 2	X_{21}	X_{22}	.	X_{25}
⋮				
CAMADA 6	X_{61}	X_{62}		X_{65}

donde X_{ij} denota el peso del puerquito de la camada i-ésima, y al que le fué administrado alimento del tipo j-ésimo, en donde el número cinco corresponden a la alimentación usual.

Para poder probar igualdad de tratamientos lo que se hace es lo siguiente:

Sea m el número de renglones y n el de columnas. Entonces, sea R_{ij} el rango asociado por renglón al lugar (ij)-ésimo, es decir, se asocian rangos por renglón, de forma tal que todos los renglones sumen $\frac{n(n+1)}{2}$. De esta forma se genera el siguiente arreglo:

$$\begin{bmatrix} R_{11} & R_{12} & R_{13} & \cdots & R_{1n} \\ R_{21} & R_{22} & R_{23} & \cdots & R_{2n} \\ & & \vdots & & \\ R_{m1} & R_{m2} & R_{m3} & & R_{mn} \end{bmatrix}$$

en donde $\sum_{j=1} R_{ij} = \frac{n(n+1)}{2}$ $\forall i \in J_m$. Sea $R_j = \sum_{i=1} R_{ij}$, el total de la columna j-ésima. Si la hipótesis de igualdad de tratamientos (o igualdad de poblaciones) es válida, es de esperarse que el total de cada columna sea aproximádamente igual al promedio de todas ellas es decir, si la hipótesis de igualdad es válida, debe esperarse que R_j y $\frac{m(n+1)}{2}$ sean aproximádamente iguales, y esto debe suceder $\forall j \in J_n$.

Esto es equivalente a pensar, en que los rangos asociados por renglón, lo fueron aleatoriamente, y que en promedio todos los totales por columna son iguales.

Como se ha visto anteriormente en estas notas, la mejor forma de checar si R_j y $\frac{m(n+1)}{2}$ difieren o no, $\forall j \in J_n$, es considerar a una función de la estadística,

$$\sum_{j=1}^{n} (R_j - \frac{m(n+1)}{2})^2$$

Esta función es $T = \frac{12}{mn(n+1)} \sum_{j=1} (R_j - \frac{m(n+1)}{2})^2$, la cual puede demostrarse (Friedman, Wilks (1937)) se distribuye según una X^2 con n-1 grados de libertad.

Si la prueba se realiza a un nivel α de confianza, la decisión es:

Si T es mayor que el cuantil $1-\alpha$ de una X^2 con n-1 grados de libertad, querrá decir que las diferencias son grandes y por tanto la hipótesis de igualdad de tratamientos deberá rechazarse. En caso contrario, las diferencias no son significativas y la hipótesis deberá aceptarse.

La distribución exacta de T para los casos $n=2$, $m<15$ y $n=3$ $m<8$; está dada por Owen (1962). Esta prueba también está discutida en Noether (1967) y Kendall (1930 y 1948 caps. 6 y 7).

EJEMPLOS

EJEMPLO 1 Cuatro variedades de frijol de soya fueron plantadas, en tres parcelas. Las cosechas fueron:

Variedades de frijol de soya

		A	B	C	D
	1	45	48	43	41
Parcelas	2	49	45	42	39
	3	38	39	35	36

¿Existe diferencia en el promedio de cosechas de las cuatro variedades de frijol de soya?.

Puesto que se tiene un esquema de muestras relacionadas (parcelas) y la información no está dada en forma dicotómica, el procedimiento de prueba de Friedman, puede ser aplicado para probar la hipótesis:

Ho: El promedio de cosechas de las cuatro variedades es el mismo, teniéndose como alternativa

Ha: Existen diferencias significativas en los promedios de cosechas.

En este ejemplo, los tratamientos son la variedad de frijol de soya, por lo que la asignación de rangos se hace como sigue:

Variedades de frijol de soya

		A	B	C	D
	1	3	4	2	1
Parcelas	2	4	3	2	1
	3	3	4	1	2

De esta forma la estadística de prueba es: (m=3 y n=4)

$$T=\frac{12}{3(4)(5)}\sum_{j=1}^{4}(R_j-\frac{3(5)}{2})^2= 0.20\ (16+25+1+4)= 7.40$$

Si $\alpha=0.05$, el cuantil 0.95 de una X^2 con 3 grados de libertad es: (tabla. 2) 7.815 la cual lleva a no rechazar la hipótesis Ho.

EJEMPLO 2 Una muestra de doce estudiantes tomada aleatoriamente, es usada para un experimento acerca de un método de enseñanza. Cuatro listas de palabras fueron hechas por el experimentador. Cada lista contiene 20 pares de palabras, pero métodos distintos de apareamiento fueron usados en cada lista. Cada estudiante escoge una lista, se le dan cinco minutos para estudiarla y es examinado para probar que tan bien recuerda las palabras. Este procedimiento se repite para las cuatro listas y para cada estudiante. Las calificaciones de los exámenes son las siguientes: (el número 20 denota: perfecto)

Estudiante		1	2	3	4	5	6	7	8	9	10	11	12
Lista	1	18	7	13	15	12	11	15	10	14	9	8	10
	2	14	6	14	10	11	9	16	8	12	9	6	11
	3	16	15	16	12	12	9	10	11	13	9	9	13
	4	20	10	17	14	18	16	14	16	15	10	14	16

¿Puede decirse que algunas listas son más fácil de memorizar que otras?.

¿Algunos estudiantes son mejores memoristas que otros?.

Para contestar ambas preguntas la prueba de Friedman será utilizada debido al tipo de información y las hipótesis que se desean probar.

Para la primer pregunta, se puede establecer la siguiente pareja de hipótesis:

Ho: Todas las listas son iguales.
Ha: Existen diferencias significativas entre las listas.

Como la hipótesis es sobre las listas, se tienen entonces doce muestras a las que se le aplican cuatro tratamientos, y lo que se quiere responder, es si el tratamiento afectó a las muestras, (si las poblaciones de donde se tomaron las muestras son iguales).

La asignación de rangos se muestra en la tabla, que de acuerdo a ella se tiene que:

$$R_1=29.5,\ R_2=19.5,\ R_3=26.0 \text{ y } R_4= 45.0$$

	LISTA			
	1	2	3	4
Estudiante 1	3	1	2	4
2	3	2	1	4
3	1	2	3	4
4	4	1	2	3
5	2.5	1	2.5	4
6	3	1.5	1.5	4
7	3	4	1	2
8	2	1	3	4
9	3	1	2	4
10	2	2	2	4
11	2	1	3	4
12	1	2	3	4
	29.5	19.5	26.0	45.0

Calculando la estadística de prueba:

$$T=\frac{12}{mn(n+1)}\sum_{j=1}^{n}\left(R_j-\frac{m(n+1)}{2}\right)^2$$

se tiene:

$$T=\frac{12}{12(20)}\left\{\left(29.5-\frac{12(5)}{2}\right)^2+\ldots+\left(45.0-\frac{12(5)}{2}\right)^2\right\}$$

de donde:

$$T= 17.5750.$$

Si se pone $\alpha=0.05$, el cuantil de una X^2_{α} con tres grados de libertad es: 7.815.

Como $T > 7.815$, la hipótesis nula se rechaza y se concluye que las listas tienen diferentes grados de dificultad.

Para la segunda pregunta, la hipótesis nula es:

Ho: Todos loes estudiantes aprenden de igual forma.

En este caso las listas son las poblaciones y los tratamientos los estudiantes. Para esto, la asignación de rangos es:

Listas	1	2	3	4	5	6	7	8	9	10	11	12
1	12	1	8	10.5	7	6	10.5	4.5	9	3	2	4.5
2	10.5	1.5	10.5	6	7.5	4.5	12	3	9	4.5	1.5	7.5
3	11.5	1	11.5	7.5	7.5	3	5	6	9.5	3	3	9.5
4	12	1.5	10	4	11	8	4	8	6	1.5	4	8

y los totales R_i, respectivos:

46, 5, 40, 28, 33, 21.5, 31.5, 21.5, 33.5, 12, 10.5, 29.5

Por lo que:

$$T= \frac{12}{4(12)(13)}\left\{ (46-\frac{4(13)}{2})^2+\ldots+(29.5 -\frac{4(13)}{2})^2 \right\}$$

de donde T=32.0288. Como n=12, se busca en la tabla 2, para 11 grados de libertad. Puede observarse, que cualquier número del renglón de la tabla (cuando k=11) es menor que T, por lo que se rechaza la hipótesis y se concluye que algunos estudiantes memorizan mejor que otros.

EJEMPLO 3 Dos investigadores de la UNAM, interesados en el problema de mortandad por causas de picaduras de alacrán, querían verificar si la mortandad por grupo de edad sufría cambios significativos a través de los años. Para tal propósito, se recolectó información, la cual se muestra en el siguiente cuadro:

	Años	1940	1950	1960	1965	1970	1975	
Grupos de edad.	0-4	45.73	37.73	11.3	11.52	8.12	4.88	
	5-9	8.34	6.12	1.57	1.56	1.39	0.64	
	10-14	1.33	1.06	0.23	0.31	0.15	0.12	
	15-19	0.40	0.42	0.08	0.07	0.08	0.03	Las entradas de la tabla muestran las tasas de mortandad por 100,000 habitantes.
	20-39	0.17	0.26	0.04	0.15	0.12	0.25	
	40-59	0.94	0.28	0.13	0.17	0.29	0.127	
	60 ó más	1.59	1.62	0.84	2.65	0.41	0.334	

La hipótesis que se estableció fue por tanto:

Ho: La mortandad por grupo de edades es la misma para todos los años.

En este caso, los tratamientos son los grupos de edad, por lo que la asignación de rangos queda como sigue:

Grupo de Edad

Año	0-4	5-9	10-4	15-19	20-39	40-59	60 o más
1940	7	6	4	2	1	3	5
1950	7	6	4	3	1	2	5
1960	7	6	4	2	1	3	5
1965	7	5	4	1	2	3	6
1970	7	6	3	1	2	4	5
1975	7	6	2	1	4	3	5

Teniéndose que: $T = 30.04$.

Como $n=7$, se busca una X^2 con 6 grados de libertad, teniéndose que $T > W_{1-\alpha}$, para cualquier α de la tabla 2, por lo que se rechaza la hipótesis de nulidad.

PRUEBA X^2 PARA K POBLACIONES.

En ocasiones cuando se tienen varias muestras, que son independientes entre si, se quiere determinar si las diferencias que existen entre ellas, es debido a que provienen de diferentes poblaciones, o son diferencias debidas al azar, entre varias muestras de la misma población. Por esto, es necesario contar con un procedimiento estadístico que ayude a probar hipótesis de igualdad de poblaciones.

El procedimiento que se verá a continuación es útil, si se tienen k muestras que son independientes entre si, y la escala en que las observaciones se miden, sea al menos nominal.

El siguiente, es un ejemplo en el cual, el procedimiento de prueba X^2, es utilizado:

Supóngase que se seleccionó en forma aleatoria una muestra de n_1 estudiantes de una escuela, en la cual se enseña de acuerdo a un cierto método. Asímismo, se seleccionaron muestras aleatorias de tamaño $n_2, n_3, \ldots, n_k$ en escuelas distintas, en donde los métodos de enseñanza son distintos entre sí y al primero.

A todos los estudiantes les fué aplicado un exámen de conocimientos y fueron clasificados de acuerdo a la calificación obtenida, de la forma siguiente:

	1	2	3	...	10
ESCUELA 1	O_{11}	O_{12}	O_{13}		$O_{1\,10}$
ESCUELA 2	O_{21}	O_{22}	O_{23}		$O_{2\,10}$
.					
.					
.					
ESCUELA K	O_{k_1}	O_{k_2}	O_{k_3}		$O_{k_{10}}$

en donde O_{ij} es el número de alumnos de la escuela i, que obtuvieron la calificación j.

El objetivo del estudio, es probar que todos los métodos de enseñanza son igualmente efectivos, lo que se traduce a decir que la proporción de alumnos que obtuvieron calificación j, es la misma para todas las escuelas. Esta hipótesis puede escribirse como:

$$Ho: p_{1j} = p_{2j} = \ldots = p_{kj} \quad \forall\, j \varepsilon J_n$$

teniéndose como alternativa:

$$Ha: p_{ij} \neq p_{rj} \text{ para alguna } j \varepsilon J_n \text{ y algún par } i, r \varepsilon Jk;$$

en donde p_{ij} denota la probabilidad de pertenecer a la celda (ij)-ésima.

De acuerdo a lo visto en las pruebas de Bondad de Ajuste e Independencia, la forma de probar esta hipótesis es considerar a la estadística siguiente, como estadística de prueba:

$$T = \sum_{i=1}^{k} \sum_{j=1}^{n} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}, \quad \text{con } E_{ij} = \frac{n_i C_j}{N} \quad \text{en donde } n_i, \text{es}$$

el tamaño de la i-ésima muestra; C_j es total de estudiantes que en la prueba obtuvieron la calificación j, es decir, $C_j = \sum_{i=1}^{k} O_{ij}$; N es el total de estudiantes, esto es, $N = \sum_{i=1}^{k} n_i$.

En general, si se tiene k muestras, donde cada muestra puede clasificarse en r clases, y se quiere probar:

$$Ho: p_{1j} = p_{2j} = \ldots = p_{kj} \quad \forall\, j \varepsilon Jr$$

contra:

$$Ha: p_{ij} \neq p_{ej} \text{ para alguna } j \varepsilon Jr \text{ y algún par } i, l \varepsilon Jr,$$

la estadística de prueba se define como:

$$T=\sum_{i=1}^{k}\sum_{j=1}^{r}\frac{(O_{ij}-E_{ij})^2}{}\quad,\ \text{con } E_{ij}=\frac{n_i C_j}{N}\ \text{en donde:}$$

O_{ij}, es el número de observaciones de la i-ésima muestra, que se clasifican en la j-ésima clase; n_i, es el número de observaciones por muestra; C_j, es total de observaciones de la columna j y N el total de observaciones.

Puede demostrarse (Mood y Graybill (1963), pp 311-319) que T se distribuye asintóticamente como una X^2 con $(k-1)(r-1)$ grados de libertad, por lo que la hipótesis Ho de igualdad de poblaciones se rechaza a un nivel α, sí T es mayor que el cuantil de orden $(1-\alpha)$ de una X^2 con $(k-1)\ (r-1)$ grados de libertad; en caso contrario, la hipótesis Ho no se rechaza.

La prueba desarrollada aquí y la prueba de independencia de criterios, que se vió en la parte de pruebas de una muestra, tienen muchas cosas en común, sin embargo son distintas. Para hacer notar esto último, se verán las semejanzas y diferencias entre estas dos pruebas, por lo que se recomienda revisar la prueba de independencia antes de ver esta parte.

Entre las semejanzas pueden notarse las siguientes:

i) Las dos estadísticas de prueba se definen en forma análoga:

$$T=\sum_{i=1}^{k}\sum_{j=1}^{r}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$ en donde E_{ij} se definen como el producto del total de la columna j-ésima por el total del renglón i-ésimo, dividido entre el total de observaciones, y las O_{ij} son el total de objetos clasificados en la (ij)-ésima celda.

ii) En ambos casos la estadística se distribuye aproximadamente -

como una X^2 con $(k-1)(r-1)$ grados de libertad, es decir, el número de grados de libertad es igual al producto entre el número de renglones menos uno por el número de columnas menos uno.

Debido a estas semejanzas, los criterios que se vieron en la prueba de independencia, con respecto al número de elementos por clase que sería deseable tener para que la aproximación fuese buena, son recomendados en esta prueba también.

Las diferencias más notorias son:

i) En la prueba de independencia, se tiene una sola población que puede ser clasificada de acuerdo a dos criterios, donde cada criterio puede tener varias categorías. En el caso de la prueba tratada aquí, se tienen varias muestras, que pueden ser clasificadas de acuerdo a un criterio, donde el criterio puede tener varias categorías. Esta diferencia se hace notar en los totales por renglón. En el caso de que se pruebe independencia de criterios, estos totales no se conocen de antemano, es decir, son variables aleatorias; en el caso de probar igualdad de poblaciones, estos totales están fijos desde el principio, pues son el número de elementos que se toman de cada población, es decir, son los tamaños de las muestras. Esto quiere decir, que en la segunda prueba se está diciendo cuántos elementos va haber en cada renglón, mientras que en la primera esto no se sabe. Esto involucra, la forma en cómo la información se obtiene, o sea el diseño del experimento.

ii) Otra diferencia es sobre las hipótesis. En la prueba de independencia de criterios, la hipótesis que se prueba es $Ho: p_{ij} = p_{i.} p_{.j}$, donde p_{ij} es la probabilidad de que un elemento de la población quede clasificado en la celda(ij), $p_{i.}$ es-la probabilidad de ser clasificado en la categoría i del primer criterio y $p_{.j}$, la de ser clasificado en la categirfa j del segundo criterio.

En la prueba de igualdad de poblaciones, la hipótesis que se prueba es $H_o: p_{1j} = p_{2j} = \ldots = p_{kj}$ $j \in J_r$, en donde p_{ij} es la probabilidad de que un elemento de la i-ésima población, se clasifique en la j-ésima categoría.

De acuerdo a la hipótesis, es obvio que p_{ij} no juegan el mismo papel en ambos casos, y además no se afirman las mismas proposiciones, por lo que las hipótesis son diferentes.

Por todo lo anterior, se recomienda tener cuidado al establecer las hipótesis y sobre todo en las conclusiones que se obtengan de una situación que, intuitivamente, podría parecer que cae en cualquiera de los dos casos.

Si se desea ampliar, pueden consultarse los siguientes artículos: Goddman (1970) y Ireland, Ku y Kullback (1969) quienes discuten en forma detallada pruebas de independencia entre más de dos criterios. En Chapman y Meng (1966) se estudia la potencia de estas pruebas X^2, mientras que Haynam y Leone (1965) dan una aproximación a la distribución exacta de la estadística T. El caso en que los valores observados por celda es pequeño o cero, lo discuten Ku (1963) y Sugiura y Otake (1968). Conover (1971) menciona algunos artículos más, sobre trabajos y estudios de la prueba X^2. Una bibliografía exahustiva, es casi imposible de dar, pues existen muchos trabajos sobre estas pruebas y contínuamente se publican nuevos. Por esto, es recomendable, que el que esté interesado en estos tópicos, consulte los índices de estadística (por ejemplo: Current Index To Statistics) en donde podría ampliar la lista dada aquí.

EJEMPLOS

EJEMPLO 1 Sesenta estudiantes fueron divididos en dos grupos de 30 cada uno. En cada grupo se les enseñó como escribir un programa para una computadora.

En uno de los grupos se usó el método tradicional de enseñanza y en el otro se usó un método experimental de enseñanza. Al final del curso se le dió a cada estudiante un examen que consistió en escribir un programa para computadora.

El programa escrito se evaluó como correcto o incorrecto y los resultados fueron los siguientes:

	Programa Correcto	Programa Incorrecto
Método Tradicional	23	7
Método Experimental	27	3

Lo que se desea es averiguar si existen diferencias entre los métodos de enseñanza. Como las muestras son independientes estre - sí, se usará el criterio de la X^2.

Para esto, se tienen los valores observados y los esperados se - calculan como $E_{ij} = \frac{1}{N} R_i C_j$. El cuadro siguiente da los valores observados en primer término y los esperados en segundo.

	Programa Correcto		Programa Incorrecto	
Método Tradicional	23	25	7	5
Método Experimental	27	25	3	5

Entonces:

$$T = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 1{,}92$$

Buscando en la tabla 2 y fijando $\alpha=0.05$ se tiene que el $1-\alpha$ cuantil de una X^2 con un grado de libertad es: $W_{0.95}=3.841$, por lo que la hipótesis de que no existen diferencias entre los dos métodos, que sean significativas, no se rechaza.

EJEMPLO 2 100 hombres y 100 mujeres, participaron en una encuesta acerca de una nueva pasta de dientes. Se les pidió contestar si les gustaba o no, el nuevo sabor de la pasta. 32 hombres y 26 mujeres respondieron que no les gustaba el nuevo sabor. ¿Indican estos datos una diferencia en la preferencia entre hombres y mujeres en general?.

Para verificar esta hipótesis los datos pueden ser acomodados en una tabla de contingencia de 2X2 de la forma siguiente:

	Les gustó la pasta	No les gustó la pasta
Hombres	68	32
Mujeres	74	26

La hipótesis nula queda como sigue:

Ho: No existe diferencia en la preferencia entre hombres y mujeres.

Como las dos muestras (hombres y mujeres) son independientes entre sí, y están clasificadas en dos categorías, la prueba X^2 es adecuada para verificar la hipótesis.

Para esto, se calculan los valores esperados, los cuales resultan ser:

$E_{11}=71 \quad E_{12}=29 \quad E_{21}=71 \quad E_{22}=29$

Por lo que: $$T=\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}=0.8742$$

Como se tiene un grado de libertad, el cuantil correspondiente de una X^2 es (α=0.01); $W_{1-\alpha}$= 6.63. (De hecho, si α=0.25, la decisión es la misma).

Por tanto, ya que $T < W_{1-\alpha}$, se concluye que no existe diferencia en la preferencia entre hombres y mujeres, que pueda considerarse significativa.

EJEMPLO 3 Para detectar las diferencias que existían entre personas de dos poblaciones distintas, con respecto a su opinión a favor de una nueva ley, se tomaron 135 personas al azar de una población y 140 en la segunda.

De las 135 personas de la primera población, 43 estuvieron en contra, mientras que la segunda población, 103 estuvieron a favor. ¿Existe una diferencia significativa en la proporción de gentes que está en contra de la ley, en las dos poblaciones?.

De acuerdo al enunciado del problema las hipótesis a probar son;

Ho: No existen diferencias significativas entre las dos poblaciones.

y Ha: No existen diferencias significativas entre las poblaciones.

Si la información es resumida en la siguiente tabla;

	Opinión a favor	Opinión en contra	Total
Población 1	92	43	135
Población 2	103	37	140

La prueba que sirve para checar las hipótesis es la X^2, teniéndose que:

$E_{11}=95.72$ $E_{12}=39.27$ $E_{21}=99.27$ $E_{22}=40.72$

Entonces:

$T= 0.9787.$

Si $\alpha=0.05$, el cuantil correspondiente (se tiene sólo un grado de libertad) es: 3.841, lo cual lleva a no rechazar la hipótesis de que no existen diferencias significativas en la proporción de gentes que está en contra de ley, en las dos poblaciones.

EJEMPLO 4 Para probar la potencia de cuatro drogas, unos investigadores hicieron el siguiente experimento: Tomaron cuatro muestras de 50 ratas cada uno, en forma aleatoria. Después de esto se inyectaba una droga a todas las ratas de una muestra, haciendo lo mismo con las drogas y muestras restantes. La potencia se medía de acuerdo al número de manchas rojas (o ronchas) que aparecía en cada rata. De acuerdo al número de ronchas, se hizo la siguiente clasificación:

		Número de Manchas			
		0-1	2-3	4-5	más de 5
Tipo de Droga.	1	12	20	8	10
	2	19	14	10	7
	3	15	22	5	8
	4	12	13	13	12

La hipótesis que se desea probar es:

Ho: La potencia de las drogas es la misma para las cuatro.

teniendo como alternativa:

EJEMPLO 5 En una fábrica se tienen tres procesos distintos para producir el mismo artículo, sin embargo el costo de producción varía de proceso a proceso. Debido a esto, se desea averiguar si la efectividad de los procesos es la misma, pues si esto es cierto, entonces se usará el más barato. La efectividad de cada proceso, se medirá en función del número de artículos defectuosos que produzca.

Se tomaron muestras de lotes de producción de cada proceso, donde cada lote contiene 50 artículos. Del primer proceso, setomaron 63 lotes, del segundo 46 y del tercero 30. Se revisó cada lote y se contó el número de artículos defectuosos encontrados. La información fué la siguiente: Los renglones indican de que proceso proviene la muestra y en las columnas clasifican los lotes de acuerdo al número de artículos defectuosos. Así, la entrada colocada es la primera columna y el primer renglón, señala el número de lotes de los fabricados con el proceso 1 en los que no se encontró ningún artículo defectuoso.

Número de Artículos defectuosos por lote

	0	1	2	3	4	5
Proceso 1	12	18	7	9	10	7
Proceso 2	14	6	5	6	8	7
Proceso 3	10	4	5	9	6	4

La prueba que se utilizará es una X^2 para checar igualdad de poblaciones, en este caso de "efectividad de procesos", esto es, las hipótesis a contrastar son:

Ho: Los tres procesos son igualmente efectivos

Ha: Existen diferencias significativas entre los tres procesos.

Se tiene que N=147 y los valores esperados se encuentran en la siguiente tabla

E_{ij} i \ J	1	2	3	4	5	6
1	15.426	11.998	7.2845	10.284	10.284	7.713
2	11.2644	8.7612	5.3193	7.5096	7.5096	5.6322
3	9.306	7.238	4.3945	6.204	6.204	4.653

Se sabe que $T=\sum_{i=1}^{3}\sum_{j=1}^{6}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$ es la estadística de prueba, por tanto, si se substituyen los valores de O_{ij} y de E_{ij} se tiene que

$$T= 9.171269$$

la cual se distribuye como una X^2 con 10 grados de libertad.

Si α=0.05, el cuantil correspondiente es: 18.31 por lo que la hipótesis Ho no es rechazada, es decir, puede buscarse el proceso más económico para usarse, pues los tres tienen la misma efectividad.

PRUEBA DE LA MEDIANA

La prueba que a continuación se verá, no es más que una extensión de la prueba de la mediana que se vió en la parte correspondiente a pruebas para dos poblaciones independientes. Como ejemplo, supóngase la siguiente situación:

Una compañía fabrica una componente eléctrica que es básica en la producción de aparatos electrónicos. Sin embargo, en los últimos meses, se ha encontrado que varias componentes han salido defectuosas, ignorándose las causas.

En vista de que esta situación ha causado que las ventas de la compañía disminuyan, la directiva reúne a un grupo de expertos en la fabricación de la componente, para que traten de encontrar la falla en la producción o al menos orienten hacia donde debe buscarse. Después de examinar toda la información disponible, la cual consiste básicamente, de los resultados de las pruebas de calidad que se aplican al final de cada etapa del proceso de fabricación, el grupo de expertos concluye en que la última etapa de éste, es en donde radica el problema. Ellos creen que la temperatura en la que esta última etapa se realiza, no es la adecuada, y hace que las componentes se deterioren. Sin embargo, no están seguros de esta aseveración.

Para poder verificar la proposición de los expertos, la directiva decide tomar, en forma aleatoria, una parte de la producción, que está por pasar a la última etapa del proceso de producción. Supongase que el número de componentes así elegidos es N. Después de ésto, se forman varios grupos, en forma también aleatoria, donde cada uno de ellos será sometido a diferentes temperaturas en la última etapa; es decir, se toman N componentes aleatoriamente, de éstas, se toman aleatoriamente n_1 para someterlas a cierta temperatu ra, en la última etapa; luego se toman n_2, y se hace lo mismo, pero a otra temperatura, y así sucesivamente hasta agotar las N componentes iniciales.

Cuando todas las componentes has pasado por esta etapa, se les hacen varias pruebas de calidad, en las que los resultados se medirán en una escala de 1 a 10. Donde, la mínima calificación estará representada por el 1 y la máxima por el 10. Por último, la "calidad" de cada componente será la suma de calificaciones obtenidas en cada una de las pruebas.

El procedimiento estadístico que a continuación se describe no trata de averiguar si la "calidad" de cada componente es bueno o no, sino que servirá para detectar si la temperatura influye en la calidad o no.

Ya que las componentes han pasado por todas las pruebas de calidad y tienen asignadas una calificación, la muestra conjunta (considerando las N componentes) puede ser ordenada de menor a mayor, de acuerdo a la calificación obtenida, y a partir de esto calcular la mediana de todas las calificaciones.

El própósito de hacer esto, es que si la temperatura no influye en el proceso de fabricación, deberá tenerse que la mediana de las calificaciones asignadas a las componentes, sea la misma para todas las muestras e igual a la mediana de la muestra conjunta.

Si se denota con O_{1_i} el número de observaciones de la i-ésima muestra que son mayores que la mediana de la muestra conjunta, y con O_{2_i} a las que son menores o iguales; se puede generar el siguiente arreglo: (suponiendo que se formaron K muestras o grupos, con las N componentes)

	1	2	3	...	k	
mayores	O_{11}	O_{12}	O_{13}	...	O_{1k}	R_1
menores o igual.	O_{21}	O_{22}	O_{23}	...	O_{2k}	R_2
	n_1	n_2	n_3		n_k	

Este arreglo no es más que una tabla de contingencia de 2xk, donde R_1 es el total de valores que exceden a la mediana de la muestra conjunta, en todas las muestras y R_2 el total de valores que son menor o igual a ella.

Puesto que las muestras son independientes entre sí, el método para probar la hipótesis de igualdad de medianas, no es más que un caso particular de la prueba X^2 para K muestras, vista en la sección anterior. De aquí que la estadística de prueba sea:

$$T = \frac{N^2}{R_1R_2} \sum_{i=1}^{k} \frac{(O_{1i} - \frac{n_iR_1}{N})^2}{n_i}$$, la cual se distribuye aproximadamente como una X^2 con k-1 grados de libertad.

Si la prueba se realiza a un nivel α, la región crítica está definida por valores de T mayores que el $(1-\alpha)$ cuantil de una X^2 con k-1 grados de libertad, se rechaza la hipótesis de igualdad de medianas; en caso contrario, la hipótesis no se rechaza.

EJEMPLOS

EJEMPLO 1 Se quiere probar la hipótesis de que las siguientes muestras fueron obtenidas de poblaciones con la misma mediana.

Muestra 1: 35, 42, 42, 30, 15, 31, 29, 29, 17, 21.
2: 34, 38, 26, 17, 42, 28, 35, 33, 16, 46.
3: 17, 29, 30, 36, 41, 30, 31, 23, 38, 30.
4: 39, 34, 22, 27, 42, 33, 24, 36, 29, 25.

Para probar la hipótesis se deben ordenar las 40 observaciones y calcular la mediana de la muestra conjunta, (El número entre paréntesis, significa la muestra a la que pertenece la observación).

15(1)	16(2)	17(1)	17(2)	17(3)	21(1)	22(4)	23(3)	24(4)	25(4)
26(2)	27(4)	28(2)	29(1)	29(1)	29(3)	29(4)	30(1)	30(3)	30(3)
30(3)	31(1)	31(3)	33(2)	33(4)	34(2)	34(4)	35(1)	35(2)	36(3)
36(4)	38(2)	38(3)	39(4)	41(3)	42(1)	42(1)	42(2)	42(4)	46(2)

Para calcular la mediana, se busca el número que divida en dos partes iguales a la muestra; en este caso es m=30. Si se clasifican las observaciones de acuerdo a la muestra de que provienen y de si son mayores o menores (menores o iguales) a la mediana, se genera la siguiente tabla de contingencia:

	1	2	3	4
Valores mayores a la mediana conjunta	4	6	4	5
Valores menores o iguales a la mediana conjunta.	6	4	6	5

La hipótesis es Ho: Todas las poblaciones tienen la misma mediana.

La estadística, de acuerdo a lo visto es:

$$T = \frac{N^2}{R_1 R_2} \sum_{i=1}^{4} \frac{1}{n_i} \left(O_{1i} - \frac{n_i R_1}{N}\right)^2$$

en donde N es el total de observaciones, R_1 el total del primer renglón y R_2 del segundo; O_{1i} los correspondientes valores observados y n_i los tamaños de las muestras.

En el ejemplo:

$$N=40;\ n_1=10\ \ n_2=10\ \ n_3=10\ \ n_4=10;\ R_1=19,\ R_2=21$$

y $$O_{11}=4,\ O_{12}=6,\ O_{13}=4 \text{ y } O_{14}=5.$$

Entonces:

$$T=\frac{1600}{3990}\left\{\left(4-\frac{190}{40}\right)^2+\left(6-\frac{190}{40}\right)^2+\left(4-\frac{190}{40}\right)^2+\left(5-\frac{190}{40}\right)^2\right\}$$

$$\Rightarrow\quad T=1.1028$$

Como el número de muestras es cuatro, se tienen tres grados de libertad. Usando la tabla 2 del apéndice, se ve que la hipótesis no es rechazada, esto es, las cuatro poblaciones tienen la misma mediana.

EJEMPLO 2 Cuatro diferentes métodos de hacer crecer maíz, fueron asignados aleatoriamente a 34 parcelas. La producción por acre, fué registrada para cada parcela:

Método	1	2	3	4
	83	91	101	78
	91	90	100	82
	94	81	91	81
	89	83	93	77
	89	84	96	79
	96	83	95	81
	91	88	94	80
	92	91		81
	90	89		
		84		

Se sabe que si las poblaciones tienen distintas medianas, entonces son diferentes, por lo que se quiere verificar si todas tienen la misma mediana, es decir, la hipótesis Ho, será:

Ho: Las muestras provienen de poblaciones de igual mediana.

contra Ha: Existe al menos una población, con mediana distinta a las demás.

Como se sabe, para probar estas hipótesis, es necesario calcular la mediana de la muestra conjunta y construir una tabla de contingencia.

La mediana de la muestra conjunta es m=89, por lo que la tabla queda como:

Método	1	2	3	4
Observaciones menores o iguales a la mediana conjunta.	3	7	0	8
Observaciones mayores a la mediana conjunta.	6	3	7	0

La estadística de prueba es $T = \frac{N^2}{R_1R_2}\sum_{i=1}^{4}\frac{(O_{1i} - \frac{n_iR_1}{N})^2}{n_i}$

Ahora, $N=34$ $R_1=16$ $R_2=18$ $O_{11}=6$ $O_{12}=3$ $O_B=7$ $O_{14}=0$ $n_1=9$ $n_2=10$ $n_3=7$ y $n_4=8$.

Por lo que $T=4.01\left(\frac{(6-4.24)^2}{9} + \frac{(3-4.70)^2}{10} + \frac{(7-3.29)^2}{7} + \frac{(3.76)^2}{8}\right)$

$\Rightarrow$ $T= 4.01(0.34+0.29+1.97+1.77)$

$\therefore$ $T= 17.52$

El cuántil correspondiente a una X^2 con tres grados de libertad, siendo $\alpha=0.05$ es $W_{1-\alpha}= 7.815$, teniéndose que la hipótesis nula es rechazada, esto es, las poblaciones no tienen la misma mediana - (lo que significa que al menos una difiere en mediana).

De hecho la hipótesis Ho se rechaza aún si $\alpha=0.001$.

EJEMPLO 3 Se tiene tres marcas diferentes de bulbos, y se tomaron muestras aleatorias de tres lotes conteniendo cada uno, una marca de bulbos. (Esto es, cada lote contiene una y sólo una marca y distinta a las otras dos, de forma tal que se tiene una muestra aleatoria de cada marca). A cada bulbo seleccionado, se le mide su vida media, teniéndose los siguientes resultados:

Marca	A	B	C
Duración	73	84	82
	64	80	79
	67	81	71
	62	77	75
	70		

Se quiere verificar si la vida media de los bulbos de las diferentes marcas es la misma, para lo cual se establece la hipótesis:

> Ho: Las tres muestras provienen de poblaciones de igual mediana teniéndose como alternativa, la negación de Ho.

La prueba adecuada parece ser la prueba de la mediana; por lo que se ordenará la muestra conjunta, se calculará la mediana y se construirá la tabla de contingencia:

62 (A) 64 (A) 67 (A) 70 (A) 71 (C) 73 (A) 75 (C)
77 (B) 79 (C) 80 (B) 81 (B) 82 (C) 84 (B)

por lo que m=75. Entonces:

	A	B	C
Observaciones mayores ó iguales a m=75	0	4	2
Observaciones menores ó iguales a m=75	5	0	2

Ahora, N=13 y $R_1=6$, $R_2=7$

$\Rightarrow T = 4.0238(1.0651+1.1598+0.0059)$

$\therefore T = 8.9764$

Haciendo $\alpha=0.05$, el cuantil correspondiente a una X^2 con dos grados de libertad es: $W_{1-\alpha} = 5.991$, rechazándose por tanto Ho, pues $T > W_{1-\alpha}$.

Por lo que no puede decirse que la vida media de los focos sea la misma para todos.

Si $\alpha=0.01$, el cuantil es $W_{1-\alpha} = 9.210$, por lo que Ho no se rechazaría. De acuerdo a los intereses del experimentador, debe de elegirse el valor de α y por tanto la decisión.

PRUEBA DE KRUSKAL-WALLIS

Como casi todas las pruebas que se refieren a k poblaciones ($k > 2$), la prueba de Kruskal-Wallis, es una generalización de un procedimiento de prueba desarrollada para dos poblaciones.

Esta prueba de k poblaciones, es una extensión hecha a la prueba Mann-Withney para dos poblaciones independientes, por Kruskal y Wallis (1952); de ahí, el nombre de la prueba.

Como un ejemplo para ilustrar esta prueba, se tiene el siguiente*:

En un laboratorio de análisis clínicos, trabajan seis personas que se dedican a hacer el mismo tipo de análisis. El jefe del laboratorio, no está muy convencido de como realizan este trabajo y por lo pronto desea verificar si las seis personas trabajan al mismo ritmo o no. Para tal efecto, registra varios tiempos por persona, de lo que se tardan en hacer un análisis, pues él cree que elregistrar los tiempos, es mejor que contar el número de análisis, que cada persona efectúa al día, pues los análisis se van distribuyendo de acuerdo a como llegan; es decir, al llegar un análisis, éste se le dá a la persona que está desocupada, si existe más de una en esta situación, entonces se asigna aleatoriamente; si todas las personas están ocupadas, la primera que se desocupe, efectuará el análisis. De esta forma, no a todas las personas les toca el mismo número de análisis y por otro lado, el número de análisis que llega a cada día es variable.

En vista de que el jefe del laboratorio no quiere actuar subjetivamente a la hora de emitir un juicio, decide buscar un procedimiento estadístico que le ayude a decidir si las seis personas trabajan al mismo ritmo.

Nótese entonces, que se tienen seis muestras aleatorias que son in

* Sugerido por Lelía Mendoza.

dependientes entre sí, y que lo que se quiere probar es si las seis, provienen de la misma población, esto es, trabajan al mismo ritmo, lo cual quiere decir, si todas las observaciones son generadas por la misma distribución.

En general, supóngase que se tienen k poblaciones independientes entre si, y considérese el siguiente arreglo.

$$
\begin{array}{cccccc}
X_{11} & R_{11} & X_{21} & R_{21} & X_{k_1} & R_{k_1} \\
X_{12} & R_{12} & X_{22} & R_{22} & X_{k_2} & R_{k_2} \\
\vdots & & \vdots & & \vdots & \\
X_{1n_1} & R_{1n_1} & X_{2n_2} & R_{2n_2} & X_{kn_k} & R_{kn_k}
\end{array}
$$

en donde X_{ij} denota la j-ésima observación de la población i-ésima, y R_{ij} el rango asociado a X_{ij}. Esto es, sí se ordenan las muestras de menor a mayor en forma conjunta, sin perder la identificación de cada elemento, y se asocian rangos del 1 hasta el N (con $N=\sum_{i=1}^{k} n_i$), se tiene que R_{ij} denotará el rango asociado en la muestra conjunta, a la j-ésima obervación de la i-ésima población.

La hipótesis a probar será;

> Ho; La función de distribución, es la misma para todas las poblaciones.

y la alternativa natural, es;

> H_1: Al menos dos funciones de distribución son distintas.

En forma matemática, quedarían;

H_o: $F_i(x) = F_j(x)$ $\forall x \epsilon X$ y $\forall i \neq j \epsilon Jk$.

H_1: $F_i(x) \neq F_j(y)$ para alguna $x \epsilon X$ y alguna $i \neq j \epsilon Jk$

donde $F_i(x)$ denota la función de distribución de la i-ésima población.

Si la hipótesis Ho, fuese cierta, es de esperarse, que en promedio R_i y R_j sean iguales, en donde R_i denota la suma de rangos asociados a la i-ésima muestra, esto es, $R_i = \sum_{j=1}^{ni} R_{ij}$ $i \epsilon Jk$.

Usando el teorema central del límite se tiene que:

$$\frac{R_i - E(R_i)}{\sqrt{V(R_i)}}$$ se distribuye aproximadamente como una

N(0,1).

donde puede demostrarse (teorema 5, pag.46, Conover (1971)) que:

$$E(R_i) = \frac{n_i(N+1)}{2} \qquad \text{y} \qquad V(R_i) = \frac{n_i(N+1)(N-n_i)}{12}$$

por lo que: $$\left[\frac{R_i - \frac{1}{2} n_i (N+1)}{\frac{\sqrt{n_i(N+1)(N-ni}}{\sqrt{12}}}\right]^2$$ se distribuye como una X^2 con un grado de libertad.

Si las Ri's fuesen independientes, se tendría que:

$$\sum_{i=1}^{k} \frac{(R_i - \frac{1}{2} n_i(N+1))^2}{\frac{1}{12} n_i(N+1)(N-n_i)} \quad \ldots (1)$$ se distribuirá como una X^2 con

k grados de libertad. Sin embargo, las Ri's no son independientes pues $\sum_{i=1}^{k} R_i = \frac{N(N+1)}{2}$, siempre. Kruskal (1952), demostró que si el i-ésimo término de la suma (1) es multiplicado por $\frac{N-ni}{N}$ $\forall i\epsilon Jk$, entonces:

$$T = \sum_{i=1}^{k} \frac{(R_i - \frac{1}{2} n_i (N+1))^2}{\frac{1}{12} n_i (N+1)(N-n_i)},$$ se distribuye asintóticamente como una X^2 con k-1 grados de libertad.

Por todo lo anterior, la estadística T, será la estadística de prueba, la cual puede reescribirse como:

$$T = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{(R_i - \frac{1}{2} n_i (N+1))^2}{n_i} \quad \text{con } N = \sum_{i=1}^{k} n_i \text{ y } R_i = \sum_{j=1}^{ni} Ri_j$$

o en forma equivalente:

$$T = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1)$$

Ahora, si $k=3$ y $n_i \leq 5$ $\forall i\epsilon Jk$, la hipótesis Ho es rechazada, a un nivel α, si $T \geq T^*$ es el valor crítico de T para la α dada, para el cual existen tablas (Tabla 12, Conover(1971)). En caso de que $T<T^*$, la hipótesis no se rechaza.

Si estas condiciones no se cumplen, se usa la aproximación a la X^2, en donde, si $T > W_{1-\alpha}$ se rechaza Ho; donde Wp, es el p-ésimo cuantil de una X^2 con k-1 grados de libertad. Si $T \leq W_{1-\alpha}$, la hipótesis Ho, no se rechaza.

Para otras pruebas similares a la Kruskal-Wallis, puede verse la página 263 de Conover (1971), para las referencias.

EJEMPLOS

EJEMPLO 1 Cuatro programas de entrenamiento para empleados fueron probados en 20 empleados nuevos, los cuales fueron asignados aleatoriamente a cada programa en grupos de cinco. Los veinte empleados fueron puestos bajo el cuidado del mismo supervisor y al final de un cierto perído de tiempo, asignó rangos a los empleados, de acuerdo a su habilidad en el trabajo (el rango menor se asignó al empleado con menos habilidad).

		Rangos				
Programa	1	4	6	7	2	10
	2	1	8	12	3	11
	3	20	19	16	14	5
	4	18	15	17	13	9

¿Existe alguna diferencia en la efectividad de los distintos programas de entrenamiento?.

La forma es como los empleados fueron asignados a cada programa, hace que las cuatro muestras de cinco personas sean independientes entre sí, y como en la información se dan los rangos asociados a la habilidad de cada trabajador, la prueba de Kruskal-Wallis es la apropiada para contestar la pregunta planteada.

Para poder aplicarla, es necesario calcular la suma de rangos asociados, por programa. Estos valores son:

$$R_1=29 \quad R_2=35 \quad R_3=74 \quad R_4=72$$

El total de empleados es 20, por lo que la estadística de prueba queda como:

$$T = \frac{12}{20(21)} \sum_{i=1}^{4} \frac{R_i^2}{n_i} - 3(21)$$

$\Rightarrow T = 0.0286(168.2+245+1095.2+1036.8) - 63$

$\Rightarrow T = 9.7927.$

Como el número de muestras es distinto de 3, la tabla 9 no puede ser usada, y es necesario utilizar la tabla 2.

Si $\alpha = 0.05$, el cuantil correspondiente es: $W_{0.95} = 7.815$. Como $T > W_{0.95}$ la hipótesis Ho: No existe diferencia en la efectividad de los distintos programas; es rechazada.

EJEMPLO 2 En el ejemplo 2 de la prueba de la mediana para k poblaciones, se llegó a la conclusión de que no todas las poblaciones tenían la misma mediana. En este ejemplo, se mostrará, mediante la prueba de Kruskal-Walis, que en efecto los cuatro tratamientos para hacer crecer maíz son significativamente diferentes.

Para esto, la muestra conjunta es ordenada (sin perder identificación) y se les asigna rangos:

Método 1	Rango	Método 2	Rango	Método 3	Rango	Método 4	Rango
						77	1
						78	2
						79	3
						80	4
						81	6.5
						81	6.5
						81	6.5
		81	6.5				
83	11					82	9

Método 1	Rango	Método 2	Rango	Método 3	Rango	Método 4	Rango
		83	11				
		83	11				
		84	13.5				
		88	15				
		89	17				
89	17						
89	17						
90	19,5						
		90	19,5				
91	23						
91	23						
		91	23				
		91	23				
				91	23		
92	26						
				93	27		
94	28,5			94	28,5		
				95	30		
				96	31,5		
96	31.5			100	33		
				101	34		

De donde R_1= 196,5 R_2=153 R_3=207 y R_4=38.50

Las hipótesis son: Ho: No existe diferencias significativas en los cuatro métodos.

Ha: Existen diferencias significativas en los cuatro métodos.

Como $T = \frac{12}{N(N+1)} \sum_{i=1}^{4} \frac{R_i^2}{n_i} - 3(N+1)$, se tiene que T=24.28, Para buscar el cuantil apropiado, debe saberse que tabla usar, en este caso k 4, por lo que se busca el cuantil $1-\alpha$ de una X^2 con tres grados de libertad, teniéndose que si $\alpha=0.05$, $W_{1-\alpha}=7.815$; $\alpha=0.10$, $W_{1-\alpha}=6.251$ y $\alpha=0.01$, $W_{1-\alpha}=11.34$; por lo que en los tres casos se está en la región de rechazo, teniéndose entonces que existen diferencias en los cuatro métodos (Esto es, Ho se rechaza).

EJEMPLO 3 Se tienen tres conjuntos de ocho medidas cada uno, de la suavidad de un cierto tipo de papel, obtenidas en tres laboratorios distintos:

Laboratorio								
A	38.7	41.5	43.8	44.5	45.5	46.0	47.7	58.0
B	39.2	39.3	39.7	41.4	41.8	42.9	43.3	45.8
C	34.0	35.0	39.0	40.0	43.0	43.0	44.0	45.0

Supóngase que los tres conjuntos son muestras aleatorias, independientes entre sí. ¿Existen diferencias sistemáticas entre los distintos laboratorios?

Establézcanse las siguientes hipótesis:

Ho: No hay diferencias sistemáticas entre los distintos Laboratorios.

y Ha: Si hay diferencias sistemáticas.

Puesto que se tiene más de dos muestras independientes, la prueba de Kruskal-Walis es aplicable.

Para poderla aplicar, se necesita asignar rango a las observaciones, este preocedimiento se muestra en la siguiente tabla:

CAPITULO V

MEDIDAS DE CORRELACIÓN

En las situaciones en que los datos son presentados en parejas - (como es el caso de las pruebas de dos poblaciones), es frecuente querer saber si las dos variables están relacionadas o no; y aún-más, saber el grado de relación, si es que existe. Lo que se trata de averiguar es si existe una relación entre dos variables, esto es, una relación matemática, que no debe confundirse con una - relación causa-efecto.

Las medidas de asociación o correlación, pueden dar importantes indicaciones acerca de las causas o efectos, pero puede suceder - también que existan correlaciones debidas al azar, correlaciones-sin sentido o correlaciones que son prácticamente insignificantes. Esto es más notorio cuando el tamaño de las muestras es pequeño, - en estos casos se encuentran fácilmente correlaciones engañosas, - debido a efectos aleatorios.

El cómo una correlación entre un par de variables pueda ser explicada, depende en mucho del conocimiento que el intérprete tenga - cerca del problema o a su experiencia. Además las correlaciones - no deben ser extrapoladas sin un conocimiento a fondo del problema en cuestión y aún así, puede darse el caso que al extrapolar, se - llegue a resultados carentes de lógica.

La relación que pueda existir entre dos variables, no necesariamente es lineal, sino que puede darse en forma cuadrática, logarítmica, o en otra forma no lineal, y como la mayoría de las medidas de correlación, lo que miden es la relación lineal entre dos variables, debe tenerse mucho cuidado en su interpretación. Si una medida de correlación lineal marca una asociación muy baja, no debe interpretarse como no relacionados, sino como no relacionados linealmente, sin que esto niegue una posible asociación no lineal, -

Laboratorio									Total por Renglón
A	3	10	16	18	20	22	23	24	136
B	5	6	7	9	11	12	15	21	86
C	1	2	4	8	13.5	13.5	17	19	78

Entonces $T=\frac{12}{N(N+1)}\sum_{i=1}^{3}\frac{R_i^2}{n_i}-3(N+1)$, para estos datos $T= 4.94$

En este caso $k=3$, pero los tamaños de las muestras son mayores a cinco, por lo que tiene que utilizarse la distribución X^2, para calcular los cuantiles de T. Si $\alpha=0.01$, $W_{1-\alpha}=9.21$, por lo que $T<W_{1-\alpha}$, teniéndose entonces que la hipótesis Ho no se rechaza, ie, no existe diferencias sistemáticas entre los distintos laboratorios para medir la suavidad del tipo de papel en cuestión.

entre los dos variables.

Sin embargo, hay cosas en que una asociación no lineal puede ser detectada por una medida de correlación lineal, como es el caso de la tau de kendall, en las situaciones en que la asociación entre las dos variables, está dada por una función monótona.

Debido a esto, en lo que resta de esta introducción, se hablará de correlación entre dos variables, en lugar de correlación lineal entre ellas, por la posibilidad de que otro tipo de relación sea detectado por una medida de correlación.

En el caso de estadística paramétrica, la medida de correlación más común, es el coeficiente de correlación de Pearson, que se denota con Pxy y se define como:

$$Pxy = \frac{\operatorname{cov}(x,y)}{[\operatorname{Var}(x)\operatorname{Var}(y)]^{\frac{1}{2}}}$$

, donde X y Y son las dos variables de interés.

Haciendo analogía con este coeficiente y sobre todo con sus propiedades (ser invariante bajo cambios de escalas de las variables, entre las más importantes), se dirá que una medida de correlación es "buena" si cumple con:

i) Si valores grandes de X y Y tienden a ser apareados y por lo tanto lo mismo sucede con valores chicos de X y Y; la medida deberá ser positiva y tender a uno, a medida que esta asociación sea más fuerte. Es decir, si para cualesquiera dos pares independientes (X_i, Y_i) y (X_j, Y_j) de variables aleatorias independientes; $X_i < X_j$ siempre que $Y_i < Y_j$ (o bien $X_i > X_j$ siempre que $Y_i > Y_j$); la medida será igual a uno, y disminuirá a medida que la relación no sea tan marcada, pero permaneciendo siempre como un valor positivo.

ii) Si valores grandes de X tienden a ser apareados con valores chicos de Y o viceversa; la medida deberá ser negativa y a medida que esta relación sea más fuerte, tender a menos uno. Es decir si (X_i, Y_i) y (X_j, Y_j) son dos pares independientes de variables aleatorias independientes, y $X_i > Y_j$ siempre que $Y_i > Y_j$ o bien si $X_i > X_j$ siempre que $Y_i < Y_j$, entonces la medida deberá ser igual a menos uno, e irá aumentando a medida que la tendencia disminuya, pero siempre siendo negativa.

iii) La medida puede tomar valores en el intervalo $[-1, 1]$ únicamente.

iv) Si las variables (X,Y) fueron construidas aleatoriamente, es decir, no existe tendencia; o bien, son independientes; la medida deberá aproximarse a cero.

Cuando (i) sucede, se dice que existe correlación positiva entre X y Y; y si la medida es uno, se dice que existe concordancia perfecta. Si (ii) sucede, se dice que existe correlación negativa entre X y Y, y si la medida es menos uno, se dice que X y Y son perfectamente discordantes. Cuando sucede (iv), se dice que existe no correlación entre X y Y.

Ahora bien, la medida de correlación, dá el grado de asociación que existe entre las dos variables; sin embargo, es importante también el poder afirmar si el encontrar correlación en la muestra, signifique que las poblaciones estén relacionadas o son fluctuaciones del azar; pues los resultados numéricos, incluso cuando han sido calculados a partir de material completamente correcto y libre de objeciones, no demustran nada; sino que sólo dan una indicación sobre algo, llaman la atención o exigen la construcción de hipótesis y la consiguiente prueba de hipótesis; por lo que en esta parte, además de las medidas de correlación, se verán pruebas que determinen, si la asociación existe en las poblaciones de las que se tomaron las muestras.

con las que se calculó la medida de correlación.

COEFICIENTE DE CONTINGENCIA DE PEARSON

Una forma fácil de detectar si existe alguna relación entre dos conjuntos de datos, es usando una tabla de contingencia, (lo que supone, que los datos deben estar referidos en escala nominal, al menos).

En el caso de la prueba X^2 para probar independencia de criterios, una dependencia entre renglones y columnas, puede interpretarse en forma obvia, como una dependencia entre los dos criterios de clasificación que estén usando. De la misma forma, en el caso de la prueba X^2 para probar igualdad de K poblaciones, una dependencia entre renglones y columnas, significará una dependencia entre las probabilidades de ser clasificado en cierta categoría.

Para estos casos, la estadística T definida por:

$$T=\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}, \text{ con } E_{ij}=\frac{R_i C_j}{N}$$

, donde R_i es el total de observaciones del renglón i-ésimo, C_j lo análogo para la columna j-ésima, N el total de observaciones clasificadas y O_{ij} el número de observaciones clasificadas en la celda (ij); es usada para probar independencia. Sin embargo en muchos casos, no nada más quiere saberse si hay dependencia, sino medir el grado de dependencia que existe sobre las dos variables y determinar si es significativo estadísticamente o no lo es, (este grado de dependencia).

Yule y Kendall (1950,p.53), le atribuyen a Pearson el siguiente coeficiente, que se utiliza para medir el grado de asociación entre dos conjuntos de datos, que se arreglan en una tabla de contingencia:

$$\left[C=\frac{T}{N+T} \right]^{1/2}$$

en donde T es la estadística que se definió en la página anterior.

Para probar si el valor de C es significativo, es decir, si el valor de C representa o indica asociación entre las dos variables de la población de que se obtuvieron los dos conjuntos de datos, se usa a la estadística T, haciendo:

> Ho: El valor de C indica independencia entre las dos variables.
>
> Ha: El valor de C indica dependencia entre las dos variables.

Si el valor de T es mayor que el $(1-\alpha)$ cuantil de una X con $(r-1)$ $(c-1)$ grados de libertad, se rechaza la hipótesis Ho, si es que la prueba se hace a un nivel α, y en donde r es el número de renglones y c el de columnas, de la tabla de contingencia.

Aunque el grado de aplicación del coeficiente C es muy amplio, tiene sus limitaciones, las cuales hacen que el coeficiente C, no sea muy bueno, en el sentido de la definición dada en la introducción.

En Conover (1971) se demuestra que el valor máximo de T puede tomar es $N(s-1)$, donde $S=\min(r,c)$, por lo que C siempre será menor o igual a $\sqrt{\frac{S-1}{S}}$, la cual siempre es menor que uno, y se aproximará a ese valor, a medida que S aumente. Por lo que el valor de C no cumple con todas las propiedades, además de que esta cota, demuestra que el valor de C depende de la dimensión de la tabla de contingencia, lo que impide comparar diferentes coeficientes C's de tablas de diferente tamaño. Por otro lado, C hereda, por decirlo de alguna forma, las restricciones que se tenían en la prueba de independencia, para el cálculo de T, con respecto al número de celdillas que deberían tener una cierta frecuencia esperada.

A pesar de todo esto, el coeficiente C, (llamado de contingencia, por razones obvias) es usado con mucha frecuencia, para medir el grado de asociación entre dos variables.

Otros coeficientes que se definen a partir de T, son:

$$\frac{T}{N}\ ,\ \frac{T}{N(S-1)}\ \text{ y } \sqrt{\frac{T}{N\sqrt{(r-1)(c-1)}}}$$

, los cuales se discuten brevemente en Conover (1971) y un poco más amplio en Yule y Kendall (1950) y Cramér (1946).

EJEMPLOS

EJEMPLO 1 Cincuenta obreros de una fábrica acudieron a la enfermería quejándose de inflamaciones debidas a la artritis.

A veinticinco de ellos se les dió aspirinas y al resto se les dió un placebo sin que lo supieran. Una hora más tarde a cada trabajador se le preguntó si el medicamento tomado había ayudado a disminuir la inflamación. 17 de los que tomaron aspirinas y 12 de los que tomaron placebo contestaron afirmativamente. ¿Puede pensarse que las aspirinas ayuda a disminuir el mal?. La pregunta puede plantearse de otra forma: Los criterios: tomó aspirina o no y disminuyó el dolor o no, ¿son independientes?.

Para responder esta pregunta, se calculará primero el grado de asociación que existe entre los dos criterios de clasificación, y después se verá si es significativo.

Para lo primero, se calculará el coeficiente de contingencia de Pearson, dado por la expresión:

$$C=\sqrt{\frac{T}{T+N}}\ ,\ \text{donde } T=\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$ y N es el

total de observaciones. En este caso $N=50$, y para calcular T, los datos se arreglan en la siguiente tabla de contingencia de 2X2:

	Respuesta	
Tomó:	Afirmativo	Negativo
Aspirina	17	8
Placebo	12	13

Teniéndose que $E_{11}=14.5$, $E_{12}=10.5$, $E_{21}=14.5$, $E_{22}=10.5$

de donde $T= 2.0524$

así $C= 0.1985$, lo que muestra un grado de asociación pequeño.

Usando la tabla 2 del apéndice, se tiene que el cuantil $(1-\alpha)$ es:

$\alpha= 0.05$ $\alpha= 0.10$
3.841 2.706 , en ambos casos la hipótesis de independencia de criterios no es rechazada, por lo que se concluye que la aspirina no ayudó a la cura del mal.

EJEMPLO 2 · En el ejemplo dos de la prueba de independencia de criterios X^2, se llegó a la conclusión de que existia relación entre el tipo de sangre y el tipo de enfermedad del paciente. Ahora lo que interesa saber es el grado de asociación. Para ésto, se calculará el coeficiente de contingencia de Pearson.

Se tiene que $C=\sqrt{\frac{T}{N+T}}$. Como $T= 40.5429$ y $N= 8766$ entonces:

$$C= 0.0680$$

Lo que se conluye es que el grado de asociación entre el tipo de sangre y el tipo del paciente es 0.0680 , lo cual puede ser muy pequeño.

EJEMPLO 3 En el ejemplo tres de la misma prueba, se concluyó que la infección con una raza no hace al individuo más o menos susceptible a la infección con la otra raza de amoebas.

Si se calcula el coeficiente de contingencia, el cual es:

C= 0.002407, puede darse cuenta, de que a pesar de existir una asociación entre los dos criterios, esta ocasión es tan pequeña, que se considera como no significativa (No es que se considere, sino la prueba X^2, dice que no es significativa la asociación).

COEFICIENTE DE CORRELACION DE SPEARMAN.

Como se mencionó en la introducción, la medida de asociación más usada en el caso paramétrico es el coeficiente de correlación de Pearson, que se define como:

$$P_{xy}=\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2\sum_{i=1}^{n}(Y_i-\overline{Y})^2}}$$

y mide el grado de asociación lineal entre $X_1,\ldots,X_n$ y $Y_1,\ldots,Y_n$ y en donde se supone que la escala de medida para las variables X y Y es de intervalos. Además, debe suponerse normalidad, para que pueda hacerse una prueba de hipótesis basada en él.

Sin embargo, dados dos conjuntos de datos, no siempre se cumplen estas dos suposiciones, por lo que es necesario el contar con una medida de asociación para estos casos.

Spearman (1904), sugirió el usar los rangos de las variables, en lugar de los valores, para construir una medida que pudiera ser utilizada para detectar asociación lineal entre dos variables y al mismo tiempo fuera útil para llevar a cabo pruebas de hipótesis, con la única restricción de que la escala de medida fuese al menos ordinal, para que se puedan asociar rangos.

El coeficiente de correlación de Spearman, se define entonces, de manera análoga al coeficiente de correlación de Pearson, sólo que en lugar de los valores de las variables, se utilizan los rangos. Si se denota con ρ_S al coeficiente de Spearman, se tiene entonces que:

$$\rho_S = \frac{\sum_{i=1} (R(Xi) - \overline{R(X)})(R(Yi) - \overline{R(y)})}{\sqrt{\sum_{i=1}^{n}(R(X_i) - \overline{R(X)})^2 \sum_{i=1}^{n}(R(Y_i) - \overline{R(y)})^2}}$$

en donde $R(X_i)$ es el rango asociado a la variable X_i, $\overline{R(X)}$ el promedio de los rangos asociados a las X's, y $R(Y_i)$ y $\overline{R(Y)}$ lo análogo para $Y_1, \ldots, Y_n$.

Ahora $\overline{R(X)} = \frac{1}{n} \sum_{i=1} R(X_i)$ por definición, entonces:

$$\overline{R(X)} = \frac{1}{n} \sum_{i=1}^{n} i = \frac{1}{n} \frac{n(n+1)}{2}$$

$$\therefore\ \overline{R(X)} = \frac{n+1}{2} \text{ y análogamente } \overline{R(y)} = \frac{n+1}{2}$$

Por otro lado:

$$\sum_{i=1}^{n} (R(X_i) - \overline{R(X)})^2 = \sum_{i=1}^{n} (i - \frac{n+1}{2})^2 = \sum_{i=1}^{n} (i^2 - i(n+1) + (\frac{n+1}{2})^2)$$

$$\Rightarrow \sum_{i=1}^{n} (R(X_i) - \overline{R(X)})^2 = \frac{n(n\ 1)(2n+1)}{6} - \frac{n(n+1)^2}{2} + \frac{n(n+1)^2}{4}$$

$$\Rightarrow \sum_{i=1}^{n} (R(X_i) - \overline{R(X)})^2 = \frac{n(N+1)(2n+1)}{6} - \frac{n(n+1)^2}{4}$$

$$\Rightarrow \sum_{i=1}^{n} (R(X_i) - \overline{R(X)})^2 = \frac{n(n^2-1)}{12}$$

y en forma análoga se tiene que $\sum_{i=1}^{n} (R(Y_i) - \overline{R(Y)})^2 = \frac{n(n^2-1)}{12}$

por lo que ρ_s puede escribirse como:

$$\rho_s = \frac{12 \sum_{i=1} (R(X_i) - \frac{n+1}{2})(R(Y_i) - \frac{n+1}{2})}{n(n^2-1)}$$

la cual es equivalente (Siegel 1976) a:

$$\rho_s = 1 - \frac{6 \sum_{i=1}^{n} [R(X_i) - R(Y_i)]^2}{n(n^2-1)}$$

En la segunda expresión se puede notar que si la relación que existe entre las variables X y Y está dada por una función monótona - (por ejemplo: $Y = \log_e X$, $Y = X^2$ con $X > 0$, etc), el coeficiente ρ_s será igual a 1. Es decir, este coeficiente no sólamente detecta relación lineal, sino en general, lo que determina es si valores grandes de X tienden a ser apareados con valores grandes de Y, teniéndose que valores chicos de X tiendan a ser apareados con valores - chicos de Y; o bien, si la tendencia es que valores grandes (chicos) de X se aparean con valores chicos (grandes) de Y. Por esto, debe tenerse cuidado al momento de concluir sobre los resultados - que se obtengan a partir del coeficiente.

Ahora, si se desea probar que el grado de asociación obtenido por el coeficiente de Spearman, determina o puede considerarse como un grado de asociación significativo dentro de la población, se utiliza a la estadística:

$$T = \sum_{i=1}^{n} (r(Xi) - R(Yi))^2$$

o bien a la misma ρ_s.

En ambos casos las hipótesis que se establecen son:

Ho: Las Xi y Yi son independientes.
Ha: Existe una tendencia en aparear a valores grandes de X con valores grandes de Y, o bien, la tendencia es aparear valores grandes de X con valores chicos de Y.

Si se usa a ρ_s como estadística de prueba, existen tablas (tabla 10, Conover (1971)), donde aparecen los cuantiles de ρ_s bajo la hipótesis Ho, de donde, si ρ_s es mayor que el $1-\frac{\alpha}{2}$ cuantil o menor que el $\frac{\alpha}{2}$ cuantil, la hipótesis Ho se rechaza, en caso contrario, no se rechaza.

Si se usa a T como estadística de prueba (la única diferencia es la reducción de cálculos aritméticos), existen también tablas (tabla 9, Conover, Op.Cit) para los cuantiles, teniéndose que la hipótesis Ho se rechaza si el valor de T es mayor que el $1-\frac{\alpha}{2}$ cuantil o menor que el $\frac{\alpha}{2}$ cuantil correspondiente.

En el caso que se quiera hacer pruebas de sólo una cola, debe notarse que ρ_s crece cuando T disminuye y viceversa, por lo que debe tenerse cuidado al definir la región de rechazo de esos casos.

Nótese que en ambos casos (ya sea usando ρ_s o usando T), la distribución de la estadística de prueba no depende de los valores que toman las variables X y Y y por lo tanto, tampoco de su distribución, por lo que esta medida ρ_s puede ser usada en el caso en que no se cumpla la suposición de normalidad, pues la distribución de ρ_s depende únicamente de los rangos asociados a X y Y.

La bibliografía que puede mencionarse, es: El artículo escrito por Spearman en 1904, donde se introduce esta medida y el artículo de Hotelling y Pabst de 1936, quienes recalcan el caracter no paramétrico de esta prueba y utiliza a T como estadística de prueba, para la comprobación de las hipótesis.

Puesto que se está trabajando con rangos y al igual que en todos los casos en que se usan (Pruebas de Wilcoxon, Mann-Whitney, etc.) pueden existir dos o más observaciones iguales y asociarles el mismo rango i, lo que se hace es asignarles el rango promedio de los rangos que les hubiera tocado si no fueran iguales. En el caso del coeficiente de correlación ρ_s, esto también es válido y debe de hacerse; sin embargo si existen muchos empates, estos hacen que el valor de $\sum_{i=1}^{n}(R(Xi)-\overline{R(X)})^2$ disminuya, si es que los empates suceden en los valores $X_1,\ldots,X_n$ (en el otro caso, es análogo), y afectan al valor de ρ_s. La suma de cuadrados se corrige restándole $\Sigma\frac{t^3-t}{12}$, en donde t es el número de observaciones "empatadas" en un rango dado, de esta forma ρ_s quedaría como:

$$\rho_s=\frac{\sum_{i=1}^{n}R_{1i}^2+\sum_{i=1}^{n}R_{2i}^2-\sum_{i=1}^{n}\left[R(X_i)-R(Y_i)\right]^2}{\sqrt{\sum^{n}R_{1i}^2\sum^{n}R_{2i}^2}}$$

donde: $\sum_{i=1}^{n}R_{1i}^2=\sum_{i=1}^{n}(R(X_i)-\overline{R(X)})^2-\Sigma T_x$, donde T_x es la corrección por empate, para X_i.

$\sum_{i=1}^{n}R_{2i}^2=\sum_{i=1}^{n}(R(Y_i)-\overline{R(Y)})^2-\Sigma T_y$, donde T_y es la corrección por empate, para las Y_i.

Con esta "nueva" ρ_s, se pueden llevar a cabo la prueba de hipótesis sin que los empates causen ambigüedades.

EJEMPLOS

EJEMPLO 1 Un grupo de siete ratas, fué sometido a una dieta deficiente durante tres semanas. A final de estas tres semanas, dos observadores asignaron rangos a las ratas de acuerdo a su estado:

Nº de Rata	Observador 1	Observador 2
1	4	4
2	1	2
3	6	5
4	5	6
5	3	1
6	2	3
7	7	7

¿La asignación de rangos por cada observador es independiente?

Se calculará primero el coeficiente de correlación de Spearman, y a partir de él, se verificará si el grado de correlación es significativo.

El coeficiente de Spearman se calcula a partir de la siguiente expresión:

$$\rho_s = 1 - \frac{6\sum_{i=1}^{n}(R(X_i) - R(Y_i))^2}{n(n^2-1)}$$

Substituyendo los valores dados en la tabla se encuentra que:

$$\rho_s = 0.8571,$$

Si $\alpha = 0.05$, se usa la tabla 10 del apéndice y se buscan los cuantiles correspondientes a $n=7$, los cuales resultan ser -0.7450 y -

0.7450, por lo que no se rechaza la hipótesis de independencia.
Si se usa $\sum_{i=1}^{n}(R(X_i) - R(Y_i))^2$ como estadística de prueba, se llega a la misma conclusión (Usando la tabla 11, los cuantiles α y 1-α son 18 y 94, y como T=48, la hipótesis de independencia no se rechaza).

EJEMPLO 2 En la situación descrita en el ejemplo 3 de la prueba de Friedman, con respecto al estudio de la mortandad causada por picaduras de alacrán, se quizo averiguar si el fenómeno mortandad estaba relacionado con el número de viviendas rociadas anualmente con insecticida para combatir el mosco que trasmite el paludismo. Para ésto, se contaban con la siguiente información:

Año	Tasa x 100,000 habitantes*(X_i)	Número de Viviendas Rociadas (Y_i)
1956	4.23 (18)	465 146 (1)
1958	3.21 (17)	3287 642 (17)
1959	2.64 (15)	3502 730 (18)
1960	2.21 (10)	3270 740 (16)
1961	2.60 (14)	1730 572 (5)
1963	3.08 (16)	1713 479 (4)
1964	2.51 (13)	1939 534 (9)
1965	2.31 (12)	1933 981 (8)
1966	2.22 (11)	1918 461 (7)
1967	1.86 (6)	1800 907 (6)
1968	1.87 (7)	1619 521 (2)
1969	2.06 (9)	1665 153 (3)
1970	1.91 (8)	1982 360 (10)
1971	1.47 (15)	2692 533 (15)
1972	1.46 (4)	2579 648 (14)
1973	1.30 (3)	2407 711 (13)
1974	1.02 (1)	2144 993 (12)
1975	1.06 (2)	2042 413 (11)

* La tasa, es una tasa bruta de mortandad general por picadura de alacrán, en la República Mexicana.

Debido a que una columna la información es absoluta (Número de viviendas) y en la otra es una tasa (tasa por 100,000 habitantes) una forma adecuada para comparar las cantidades en forma directa y poder medir la correlación es asignar rangos. Esta asignación se muestra en la tabla, a la derecha de los números y entre paréntesis.

La medida de correlación que se calculará, será la ρ de Spearman, (aunque también puede usarse la T de Kendall, como se verá en la siguiente sección), y después se probará la hipótesis Ho: Las tasas y el número de viviendas rociadas son independientes.

Para calcular el coeficiente de correlación de Spearman, se evalúa la siguiente expresión, usando los rangos asignados a la información:

$$\rho_s = 1 - \frac{6\ \Sigma(R(X_i) - R(Y_i))^2}{n(n^2-1)}$$

Al substituir los valores se obtiene que:

$$\rho_s = -0.21$$

Si se fija $\alpha = 0.05$, se puede ver (usando la tabla 10 del apéndice) que la hipótesis de independencia no es rechazada.

En forma análoga, si se usa a $T = \Sigma\ (R(X_i) - R(Y_i))^2$, se tiene que

$$T = 1174$$

y los cuantiles correspondientes son:

$$W_{\alpha/2} = 512 \quad \text{y} \quad W_{1-\alpha/2} = 1320 \quad (\alpha = 0.05)$$

llegando a la misma conclusión de no rechazar Ho, como era de esperarse.

EJEMPLO 3 Continuando con esta situación lo siguiente que se checó, era si el índice de ruralidad (Población rural entre población urbana) estaba relacionado con la tasa de mortandad. Pues se tenía la sospecha, que la mortandad era mayor en la población rural que en la urbana y esa información no se tenía, sin embargo una alta correlación podría dar indicios de este problema. El problema es que sólo se pudo conseguir poca información y de esta forma el coeficiente de correlación no puede ser interpretado tan fácilmente.

Los datos se pueden resumir en la siguiente tabla, en donde también se muestra la asignación de rangos.

	Tasa de Mortandad		Indice de Ruralidad	
1940	8.27	(6)	1.849	(6)
1950	7.47	(5)	1.348	(5)
1960	2.21	(3)	0.972	(4)
1965	2.31	(4)	0.801	(3)
1970	1.91	(2)	0.703	(2)
1975	1.06	(1)	0.593	(1)

Al calcular ρ_s se obtuvo que $\rho_s = 0.9429$, y consultando tablas se tiene que la hipótesis de independencia es rechazada.

Sin embargo, esta decisión debe tomarse con reservas, pues la información es poca.

COEFICIENTE DE CORRELACION DE KENDALL

Otra medida de asociación muy usada en el caso no paramétrico es el coeficiente de correlación de Kendall. Este coeficiente, al igual que el coeficiente de correlación de Spearman, está basado en el orden de las observaciones (o sea el rango de las observaciones), en lugar del valor de las mismas, por lo que se necesita al menos una escala ordinal para poder ser utilizado. Debido a que no depende del valor de las observaciones, la distribución del coeficiente no dependerá de la distribución de las dos variables, si es que éstas son independientes.

Esta medida de asociación tiene dos ventajas sobre la medida de correlación de Spearman. La primera es que la función de distribución del coeficiente de correlación de Kendall (llamado así, debido a que él la propuso (Kendall 1938)) se aproxima a una normal, cuando el número de parejas crece, en forma más rápida que la función de distribución del coeficiente de Spearman. Por otro lado, este coeficiente puede ser extendido a un coeficiente de correlación parcial, como se verá en la siguiente prueba.

Debido a estas dos razones y al hecho de contar con varias medidas de asociación, de forma tal, que se pueda elegir entre varias; se incluye esta medida.

Supóngase entonces, que se tienen dos conjuntos de observaciones: $X_1,\ldots,X_n$ y $Y_1,\ldots,Y_n$, dadas en la siguiente forma $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$. Si dadas dos parejas (X_i,Y_i) y (X_j,Y_j) se tiene que $X_i<X_j$ $(X_i>X_j)$ y $Y_i<Y_j$ $(Y_i>Y_j)$ se dirá que las parejas son concordantes; en caso de que $X_i<X_j$ $(X_i>X_j)$ y $Y_i>Y_j$ $(Y_i<Y_j)$ se dirá que son discordantes; en caso de igualdad en alguno de los dos miembros, las parejas no son ni concordantes ni discordantes.

Si se denota con Nc al total de parejas concordantes y con Nd al total de parejas discordantes, el coeficiente de correlación de Kendall puede ser escrito como:

$$T = \frac{Nc - Nd}{1/2\ n\ (n-1)}$$

Debido a que el número de pares que pueden formarse es $\binom{n}{2} = \frac{1}{2}n(n-1)$; si todos los pares son concordantes, se tendrá que T=1 y si todos son discordantes, será igual a menos uno; lo que significa que la medida de correlación T verifica la definición dada en la introducción.

Una forma fácil de calcular T, es ordenar en forma creciente a las X'S y entonces las Y'S se comparan únicamente con los valores que le preceden y así contar el total de pares concordantes y discordantes. Este precedimiento, se ilustrará con el siguiente ejemplo.

Suponga las observaciones:

(47,55), (52,33), (45,79), (21,13), (56,89), (37,68)
(45,74), (41,02), (30,03), (53,37), (54,33)y(79,79)

Ordenándolas con respecto a X se tiene:

(21,13) (30,03) (37,68) (41,02) (45,79) (45,74)
(47,55) (52,33) (53,37) (54,33) (56,89) (79,79)

Ahora considérese la pareja (21,13), para ella, existen 9 parejas concordantes y 2 discordantes, después de ella.

Para (30,03) existen 9 concordantes y una discordante.

Si se sigue así se genera el siguiente arreglo:

	(Xi,Yi)	Número de parejas concordantes abajo de (Xi,Yi).	Número de parejas discordantes abajo de (Xi,Yi).
	(21,13)	9	2
	(30,03)	9	1
	(37,68)	4	5
	(41,02)	8	0
empate.	(45,79)	1	4
	(45,74)	2	4
	(47,55)	2	3
	(52,33)	3	0
	(53,37)	2	1
	(54,33)	2	0
	(56,89)	0	1
	(79,79)	0	0

Entonces $N_c = 42$ y $N_d = 21$.

de aquí resulta que T= 0.1590.

El coeficiente de correlación de Kendall, puede ser usado para probar hipótesis de independencia, sin embargo si sólo se usa a $N_c - N_d$ se simplifican los cálculos, por lo que se propone a $T = N_c - N_d$ como la estadística de prueba, para la cual existen valores tabulados (tabla 11, Conover (1971)), Si T es menor que el $\alpha/2$ cuantil o mayor que el $1-\alpha/2$ cuantil, la hipótesis de independencia se rechaza.

EJEMPLOS

EJEMPLO 1 Una pareja va a jugar boliche y por curiosidad anota sus puntuaciones para ver si están relacionadas. Los resultados son:

Linea J	El	Ella
1	147	122
2	158	128
3	131	125
4	142	123
5	183	115
6	151	120
7	196	108
8	129	143
9	155	124
10	158	123

La hipótesis a probar es Ho: Las puntuaciones de él, son independientes a los de ella.

Para probar que la hipótesis se usará el coeficiente de correlación T de Kendall, por lo que se ordenarán los resultados de él, para calcular Nc, número de parejas concordantes; y Nd, número de parejas discordantes. Entonces:

EL	Ella	Parejas Concordante por debajo de (X,Y)	Parejas Discordantes por debajo de (X,Y)
129	143	0	9
131	125	1	7
142	123	2	4
147	122	3	3
151	120	3	2

El	Ella	Parejas Concordante por debajo de (X,Y)	Parejas Discordantes por debajo de (X,Y)
155	124	1	3
158	123	1	2
158	128	0	2
183	115	0	1
196	108	Nc=11	Nd=33

Entonces:

$$T=\frac{11-33}{1/2\ 10(9)}=-\frac{22}{45}$$

$$\therefore T=-0.4888$$

Por lo que el grado de asociación entre los resultados de la pareja es -0.4888. Para ver si este grado de asociación es significativo, se usa $T=Nc-Nd$ como estadística de prueba.

En este ejemplo:

$$T=-22$$

Si $\alpha=0.01$, los cuantiles $W_{\alpha/2}$ y $W_{1-\alpha/2}$ son: (tabla 13) -27 y 27.

Como $W_{\alpha/2} < T < W_{1-\alpha/2}$, la hipótesis nula de independencia no se rechaza, es decir a pesar de que el grado de asociación es mas o menos grande, no es significativo.

EJEMPLO 2 Recordando el problema del estudio de la mortandad causada por picadura de alacrán, se quería saber si la producción del suero que servía como antídoto y la tasa de mortandad, estaban relacionadas. Desgraciadamente se tenía esta información, únicamente por períodos decenales. Sin embargo, se llevó a cabo una prueba de correlación, para obtener una idea más amplia del panorama.

Se utilizó el coeficiente de correlación de Kendall, por considerarlo adecuado (la elección entre la T y ρ fué arbitraria).

La información es:

Período	Producción de Suero (frascos X 100,000 hab.)	Tasa de Mortandad.
1940 - 1949	111	7.99
1950 - 1959	166	3.36
1960 - 1969	235	2.30
1970 - 1979	338	1.37

Es fácil verificar que T=Nc-Nd= -6 y que T= - 1.

Por lo que puede concluirse que existe una relación entre suero y mortandad de -1, que es significativa (Los cuantiles que se obtienen en la tabla 13, son tales que $W_{\alpha/2} < T < W_{1-\alpha/2}$). Sólo que esta conclusión debe tomarse con reservas, pues la información está muy condesada y puede enmascarar al fenómeno.

EJEMPLO 3 En el ejemplo dos del coeficiente de correlación de Spearman se estudió la relación que existía entre el número de viviendas rociadas por cierto pesticida y la tasa de mortandad.

En este ejemplo se hará lo mismo, usando la T de Kendall. La información es:

Año	Tasa por 100,000 habitantes (X_i)	Número de Viviendas Rociadas (Y_i)
1956	4.23	3287 642
1958	3.21	3502 730
1959	2.64	3270 740
1960	2.21	1730 572
1961	2.60	1713 479

Año	Tasa por 100,000 habitantes (X_i)	Número de Viviendas Rociadas (Y_i)
1963	3.08	1730 572
1964	2.51	1713 479
1965	2.31	1939 534
1966	2.22	1933 981
1967	1.86	1918 461
1968	1.87	1800 907
1969	2.06	1619 521
1970	1.91	1665 153
1971	1.47	1982 360
1972	1.46	2692 533
1973	1.30	2579 648
1974	1.02	2407 711
1975	1.06	2144 993
		2042 413

Para poder calcular T es necesario contar el número de pares concordantes y el número de parea discordantes. Para esto se ordenarán los datos, con respecto a X_i:

(X_i Y_i)	Pares Concordantes por debajo de (X_i,Y_i)	Pares Discordantes por debajo de (X_i,Y_i)
(1.02, 2144 993)	6	11
(1.06, 2042 413)	6	10
(1.30, 2407 711)	5	10
(1.46, 2579 648)	4	10
(1.47, 2692 533)	3	10
(1.86, 1800 907)	7	5
(1.87, 1619 521)	10	1
(1.91, 1982 360)	3	7
(2.06, 1665 153)	8	1

$(X_i\ Y_i)$	Pares Concordantes por debajo de (X_i, Y_i)	Pares Discordantes por debajo de (X_i, Y_i)
(2.21, 3270 740)	2	6
(2.22, 1918 461)	4	3
(2.51, 1939 534)	3	3
(2.60, 1730 572)	2	3
(2.64, 3502 730)	2	2
(3.08, 1713 479)	0	3
(3.21, 3287 742)	1	1
(4.23, 465 146)	0	1
	Nc 66	Nd 87

Entonces :

$$T = Nc - Nd = -21 \text{ por lo que } T = -0.1273$$

Si α igual a 0.05 el cuantil $W_{\alpha/2} = -51$ y $W_{1-\alpha/2} = 51$ por lo que la hipótesis de independencia no se rechaza, como era de esperarse.

COEFICIENTE DE CORRELACION PARCIAL DE KENDALL

En ocasiones, la existencia de correlación entre dos variables, no es debida a que las dos variables estén correlacionadas directamente, sino que existe otra variable que provoca dicha correlación; - esto es existe otra variable que esta correlacionada con cada una de las dos variables originales y esto causa la presencia de correlación entre ellas. Por ejemplo, puede encontrarse correlación entre el volúmen diario de agua que se gasta para cubrir las necesidades de todos los habitantes de una ciudad y el número de desocupados de la misma, y ésta sea debida a que ambas cosas están correlacionadas por separado con el número de habitantes que hay en la ciudad.

En estas situaciones lo interesante es saber si la correlación entre dos variables, es debida a la influencia de otra variable, o - es porque estan correlacionadas directamente, entendiendo por ésto que la correlación entre las variables existe, sin importar si hay o no, influencia externa, como podría ser el caso de altura y edad. O bien, medir la correlación entre las dos variables, eliminando - la influencia de la otra variable de alguna forma.

Para poder calcular esto, se introduce el coeficiente de correlación parcial de Kendall, cuya definición es análoga al coeficiente de correlación parcial de Pearson del caso paramétrico, y hace uso de la definición de la Tau de Kendall. Este coeficiente de correlación parcial se denota como $T_{x_1,x_2 \cdot x_3}$, en donde la correlación - que se esta midiendo, es la que existe entre x_1 y x_2 eliminando la influencia de la variable x_3. El coeficiente de correlación parcial se define como:

$$T_{x_1x_2 \cdot x_3} = \frac{T_{x_1x_2} - T_{x_1x_3}T_{x_2x_3}}{\sqrt{(1-T^2_{x_1x_3})(1-T^2_{x_2x_3})}}$$

en donde $T_{x_ix_j}$ es el coeficiente de correlación de Kendall, definido anteriormente. (El caso general, se trata en el apéndice anexado al final de la sección de medidas de correlación).

Si no existe correlación entre las variables x_1 y x_3 y las variables x_2 y x_3 se tendrá que $T_{x_1x_2 \cdot x_3} = T_{x_1x_2}$, pudiéndose interpretar ésto, como que las variables x_1 y x_2 están relacionadas directamente. - De esta forma una comparación entre $T_{x_1x_2 \cdot x_3}$ y $T_{x_1x_2}$ podría dar una idea de la influencia de x_3 en x_1 y x_2.

El problema de probar hipótesis basados en este coeficiente, no ha sido resuelto aún por lo que no puede llevarse a cabo, lo que limita un poco el uso de esta estadística. Sin embargo por lo dicho - en el párrafo anterior, puede servir para dar una idea de lo que - sucede al eliminar la influencia de la otra variable.

EJEMPLOS

EJEMPLO 1 Después de varios estudios, se ha visto que la demanda de un producto es afectada por varios factores. En un estudio, - mediciones sobre urbanización relativa y nivel educativo fueron - hechas en nueve áreas geográficas. Los datos obtenidos fueron:

Demanda del producto X X_1	Urbanización Relativa X_2	Nivel Educativo X_3
167.1	42.2	11.2
174.4	48.6	10.6
160.8	42.6	10.6
162.0	39.0	10.4
140.8	34.7	9.3
174.6	44.5	10.8
163.7	39.1	10.7
174.5	40.1	10.0
185.7	45.9	12.0

En donde la demanda del producto X fué medida como el número de artículos demandados por 1000 habitantes. La urbanización relativa, como cantidad de metros cuadrados construidos por cada 50 metros cuadrados y el nivel educativo como grado de educación promedio en una escala de 15 unidades.

Se quiere medir la correlación parcial entre X_1 y X_2 y entre X_1 y X_3 para tener una idea de cual de los dos factores X_2 y X_3 influye más en X_1.

Para esto, se ordenará de menor a mayor X_1 para calcular:

$$T_{x_1x_2\cdot x_3} = \frac{T_{x_1x_2} - T_{x_1x_3}\,T_{x_2x_3}}{\sqrt{1-T^2_{x_1x_3}}\;\sqrt{1-T^2_{x_2x_3}}} \quad \text{y} \quad T_{x_1x_3\cdot x_2} = \frac{T_{x_1x_3} - T_{x_1x_2}\,T_{x_3x_2}}{\sqrt{1-T^2_{x_1x_2}}\;\sqrt{1-T^2_{x_2x_3}}}$$

Entonces:

X_1	X_2	X_3
140.8	34.7	9.3
160.8	42.6	10.6

X_1	X_2	X_3
162.0	39.0	10.4
163.7	39.1	10.7
167.1	42.2	11.2
174.4	48.6	10.6
174.5	40.1	10.0
174.6	44.5	10.8
185.7	45.9	12.0

Para calcular $T_{x_1x_2}$ se usa la técnica empleada en la sección de la medida de correlación T;

X_1	X_2	Parejas Concordante	Parejas Discordantes
140.8	34.7	8	0
160.8	42.6	3	4
162.0	39.0	6	0
163.7	39.1	5	0
167.1	42.2	3	1
174.4	48.6	0	3
174.5	40.1	2	0
174.6	44.5	1	0
185.7	45.9		

$$\Rightarrow \quad T_{x_1x_2} = \frac{28-8}{21} = 0.9523$$

Para $T_{x_1x_3}$ se tiene que $T_{x_1x_3} = \frac{25-9}{21} = 0.7619$

De forma análoga se obtiene que: $T_{x_2x_3} = 0.8095$.

Así:

$$T_{x_1x_2 \cdot x_3} = \frac{0.9523 - 0.7619(0.8095)}{\sqrt{(1-(0.7619)^2\ ((1-(0.8095)^2)}}$$

$$\Rightarrow \quad T_{x_1x_2 \cdot x_3} = \frac{0.3355}{0.3802}$$

$$\therefore T_{x_1x_2 \cdot x_3} = 0.8824$$

Análogamente:

$$T_{x_1x_3 \cdot x_2} = \frac{0.7619 - 0.9523(0.8095)}{\sqrt{(1-(0.9523)^2)(1-(0.0095)^2)}}$$

$$\Rightarrow \quad T_{x_1x_3 \cdot x_2} = \frac{-0.0089}{0.1792}$$

$$\therefore T_{x_1x_3 \cdot x_2} = -0.0496$$

Dados estos resultados, puede pensarse que X_2 está más correlacionada (o puede explicar mejor) a X_1 que X_3. Sin embargo, como no existe un procedimiento de prueba para verificar esta suposición, la conclusión se queda como una conjetura.

COEFICIENTE DE CONCORDANCIA DE KENDALL

Hasta este momento se han visto medidas de correlación que dan el grado de asociación existente entre dos variables. Sin embargo, puede darse el caso en que sea necesario conocer el grado de asociación entre más de dos variables. Al igual que en el caso de pruebas para k poblaciones, puede pensarse en usar de alguna forma, las medidas de correlación desarrolladas para el caso de dos variables. Una forma de hacer esto, sería calcular los coeficientes de correlación para todas las parejas posibles y la asociación total entre las k variables, podría ser medida por medio del promedio de los coeficientes de correlación. (El coeficiente de correlación puede ser tanto la Rho de Spearman, como la Tau de Kendall). Si se tiene k variables el número de coeficientes que tendrían que calcularse sería el mismo que el número de parejas que puedan formarse, las cuales son $\binom{k}{2}$. Si k es grande, el procedimiento para medir la asociación de las k variables, es laborioso. Por otro lado, las pruebas de hipótesis que se quisieran hacer a partir de este procedimiento, no serían confiables, puesto que al igual que el caso de k poblaciones, la probabilidad de error tipo 1 cambia.

Para estos casos se tiene una medida de asociación, llamada coeficiente de concordancia de Kendall, el cual fué introducido por el mismo Kendall y Babington-Smith en 1939 y simultáneamente y en forma independiente por Wallis (1939).

Para definir el coeficiente de concordancia, supóngase que se tiene k variables, y de cada una de ellas n obervaciones. Al igual que el la prueba de Friedman, puede formarse el siguiente arreglo:

Observaciones

variables en estudio.

$$\begin{matrix} R_{11} & R_{12} & \cdots & R_{1n} \\ R_{21} & R_{22} & \cdots & R_{2n} \\ & & \vdots & \\ R_{k1} & R_{k2} & \cdots & R_{kn} \end{matrix}$$

donde R_{ij} denota el rango asociado por renglón de la j-ésima observación de la i-ésima variable. Esto es, los rangos se asocian de forma tal que todos los renglones sumen $\frac{n(n+1)}{2}$; así, una posible asociación podrá ser medida por medio de los totales por columna. En efecto, sea $R_j = \sum_{i=1}^{k} R_{ij}$, el total de la columna j-ésima. Si existiera una asociación o relación perfecta entre las variables, la asociación de los rangos en los variables sería la misma para todas ellas y la sucesión de R_j sería de la forma k, 2k, ek,...,nk; aunque no necesariamente en ese orden. En cambio, si no existiera una asociación entre las variables, las R_j serían aproximadamente iguales.

Por esto el coeficiente de concordancia de Kendall está definido en función de R_j y debido al párrafo anterior, es también función de la estadística usada en la prueba de Friedman vista anteriormente. Si se denota por W a este coeficiente, se le define como:

$$W = \frac{12}{k^2 n(n+1)(n-1)} \sum_{j=1}^{n} (R_j - \frac{k(n+1)}{2})^2$$

Recordando que la estadística usada en la prueba de Friedman se definía como:

$$T = \frac{12}{nk(n+1)} \sum_{j=1}^{n} (R_j - \frac{k(n+1)}{2})^2$$

se tiene que W puede escribirse como:

$$W = \frac{T}{k(n-1)}$$

por lo que T puede usarse para llevar a cabo pruebas de hipótesis sobre W, pues si T excede 1-α cuantil entonces W excede el suyo, de forma de que la hipótesis de que existe independencia entre las k variables se rechaza si esto sucede, a un nivel de significancia α.

En caso de que existan empates al momento de asociar rangos, se hace la siguiente corrección:

$$W = \frac{12}{k^2(n^3-n) - k\Sigma e} \sum_{j=1}^{n} \left(R_j - \frac{k(n+1)}{2}\right)^2$$

donde e es igual a $\frac{\Sigma(t^3-t)}{12}$, en el cual t es el número de observaciones empatadas en un rango dado, y la suma es sobre todos los grupos en donde exista empate de cualquiera de las k ordenaciones; por lo que Σe significa sumar todas las ordenaciones.

EJEMPLOS

EJEMPLO 1 En la situación descrita en el ejemplo tres de la prueba de Friedman, se llegó a la conclusión de que la hipótesis nula se rechazaba, lo que quería decir que la mortandad por grupos de edades no era la misma para todos los años, es decir que las seis poblaciones (definidas por los años) no eran las mismas. Lo que se va a hacer ahora es ver si las variables bajo estudio (en estos casos los grupos de edad) están o no relacionados. Como se sabe esto se contesta mediante la prueba de Friedman, esto es, si se re chaza Ho en la prueba Friedman, se rechaza la hipótesis Ho: Las variables bajo estudio son independientes. Lo único que queda por preguntarse es: ¿Cuál es el grado de asociación entre las variables?. Como son más de dos variables, el coeficiente de concordancia de Kendall es el indicado para responder esta pregunta.

El coeficiente de concordancia se define, como ya se vió, de la forma siguiente:

$$W= \frac{12}{k^2n(n+1)(n-1)} \sum_{j=1}^{n} (R_j - \frac{k(n+1)}{2})^2$$, que es lo mismo que:

$$W= \frac{T}{k(n-1)}$$, donde T es la estadística que se usa en la prueba de Friedman.

En el ejemplo 3 de la prueba de Friedman, se encontró que T era:

$$T= 30.04$$

entonces: $W= \frac{30.04}{6(6)}$

$\therefore$ W=0.8344, lo cual indica un alto grado de asociación.

EJEMPLO 2 En el ejemplo uno de la prueba de Friedman, se quería saber si el promedio de cosechas de las cuatro variedades de frijol de soya, era el mismo, concluyéndose que esta hipótesis no se rechazaba. Análogo al ejemplo anterior de esta sección, interesa conocer el grado de asociación que existe entre las tres parcelas, ya que se sabe que estas son independientes, es decir, ninguna propiedad de alguna parcela interfiere con las otras dos.

En este caso T=7.40, por lo que W=0.8222. Aunque este grado de asociación es muy grande, no es significativo. Es decir, las muestras pueden estar muy relacionadas, pero esto no implica necesariamente que las poblaciones lo estén.

APENDICE

Sean $X_1, \ldots, X_n$ variables aleatorias con la misma distribución. - Sea r_{ij} con medida de correlación entre X_i y X_j $i,j \in Jn$. Es obvio que $r_{ii} = 1$ $\forall i \in Jn$. Considérese la matriz cuyos elementos son estas coeficientes de correlación:

$$\begin{bmatrix} 1 & r_{12} & r_{13} & \ldots & r_{1n} \\ r_{21} & 1 & r_{23} & \ldots & r_{2n} \\ \\ r_{n1} & r_{n2} & r_{n3} & \ldots & 1 \end{bmatrix}$$

a la cual se le llama matriz de correlación.

Sea C'_{ij} el menor ij de la matriz de correlación C, es decir, C'_{ij} es el determinante de la submatriz de C, que se obtiene al eliminar el i-ésimo renglón y la j-ésima columna. Defínase $C_{ij} = (-1)^{i+j} \cdot C'_{ij}$, al cual se le llama el ij-cofactor de la matríz C.

Con esto, Kendall (1973; pags 331-332) define el coeficiente de correlación múltiple entre dos variables, eliminando la influencia - de las $(n-2)$ restantes, como:

$$r_{12 \cdot 34 \ldots n} = - \frac{C_{12}}{(C_{11} C_{22})^{\frac{1}{2}}}$$

Puede mostrarse que si $n=3$, $r_{12 \cdot 3}$, queda como:

$$r_{12 \cdot 3} = \frac{r_{12} - r_{13} r_{23}}{\left\{ (1 - r_{13}^2)(1 - r_{23}^2) \right\}^{\frac{1}{2}}}$$

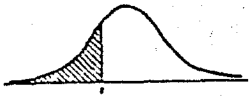
Nótese que esta expresión coincide con la dada para el coeficiente de correlación parcial de Kendall, substituyendo r por T.

Este coeficiente de correlación parcial, es el coeficiente de correlación parcial de la población, el cual para fines prácticos, es necesario estimarlo. El problema de que estimador usar para la matriz de correlación, dependerá del experimentador, él puede escoger al coeficiente de correlación de Pearson, el de Spearman o el de Kendall, para estimar r_{ij}; lo que dependerá de la experiencia del experimentador, del tipo de problema, etc.

T A B L A S

Tabla 1, Distribución N(0,1)

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}u^2\right) du = P(Z \le z)$$



z	0	1	2	3	4	5	6	7	8	9
−3.	.0013	.0010	.0007	.0005	.0003	.0002	.0002	.0001	.0001	.0000
−2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0126	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0238	.0233
−1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0300	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0570	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
− .9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
− .8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
− .7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2206	.2177	.2148
− .6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
− .5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
− .4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
− .3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
− .2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
− .1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
− 0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

z	0	1	2	3	4	5	6	7	8	9
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9278	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9430	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9648	.9656	.9664	.9671	.9678	.9686	.9693	.9700	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9762	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9874	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.	.9987	.9990	.9993	.9995	.9997	.9998	.9998	.9999	.9999	1.0000

Tabla 2. Cuantiles de $X^2_{(k)}$

	$p = .750$	.900	.950	.975	.990	.995	.999
$k = 1$	1.323	2.706	3.841	5.024	6.635	7.879	10.83
2	2.773	4.605	5.991	7.378	9.210	10.60	13.82
3	4.108	6.251	7.815	9.348	11.34	12.84	16.27
4	5.385	7.779	9.488	11.14	13.28	14.86	18.47
5	6.626	9.236	11.07	12.83	15.09	16.75	20.51
6	7.841	10.64	12.59	14.45	16.81	18.55	22.46
7	9.037	12.02	14.07	16.01	18.48	20.28	24.32
8	10.22	13.36	15.51	17.53	20.09	21.96	26.13
9	11.39	14.68	16.92	19.02	21.67	23.59	27.88
10	12.55	15.99	18.31	20.48	23.21	25.19	29.59
11	13.70	17.28	19.68	21.92	24.73	26.76	31.26
12	14.85	18.55	21.03	23.34	26.22	28.30	32.91
13	15.98	19.81	22.36	24.74	27.69	29.82	34.53
14	17.12	21.06	23.68	26.12	29.14	31.32	36.12
15	18.25	22.31	25.00	27.49	30.58	32.80	37.70
16	19.37	23.54	26.30	28.85	32.00	34.27	39.25
17	20.49	24.77	27.59	30.19	33.41	35.72	40.79
18	21.60	25.99	28.87	31.53	34.81	37.16	42.31
19	22.72	27.20	30.14	32.85	36.19	38.58	43.82
20	23.83	28.41	31.41	34.17	37.57	40.00	45.32
21	24.93	29.62	32.67	35.48	38.93	41.40	46.80
22	26.04	30.81	33.92	36.78	40.29	42.80	48.27
23	27.14	32.01	35.17	38.08	41.64	44.18	49.73
24	28.24	33.20	36.42	39.37	42.98	45.56	51.18
25	29.34	34.38	37.65	40.65	44.31	46.93	52.62
26	30.43	35.56	38.89	41.92	45.64	48.29	54.05
27	31.53	36.74	40.11	43.19	46.96	49.64	55.48
28	32.62	37.92	41.34	44.46	48.28	50.99	56.89
29	33.71	39.09	42.56	45.72	49.59	52.34	58.30
30	34.80	40.26	43.77	46.98	50.89	53.67	59.70
40	45.62	51.81	55.76	59.34	63.69	66.77	73.40
50	56.33	63.17	67.50	71.42	76.15	79.49	86.66
60	66.98	74.40	79.08	83.30	88.38	91.95	99.61
70	77.58	85.53	90.53	95.02	100.4	104.2	112.3
80	88.13	96.58	101.9	106.6	112.3	116.3	124.8
90	98.65	107.6	113.1	118.1	124.1	128.3	137.2
100	109.1	118.5	124.3	129.6	135.8	140.2	149.4
x_p	.675	1.282	1.645	1.960	2.326	2.576	3.090

Si $k > 100$ usar la aproximación:

$$W_p = \frac{1}{2}\left(X_p + \sqrt{2k-1}\right)^2$$

Tabla 3. Distribución Binomial.

n	y	p = .05	.10	.15	.20	.25	.30	.35	.40	.45
1	0	.9500	.9000	.8500	.8000	.7500	.7000	.6500	.6000	.5500
	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0	.9025	.8100	.7225	.6400	.5625	.4900	.4225	.3600	.3025
	1	.9975	.9900	.9775	.9600	.9375	.9100	.8775	.8400	.7975
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	.8574	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664
	1	.9928	.9720	.9392	.8960	.8438	.7840	.7182	.6480	.5748
	2	.9999	.9990	.9966	.9920	.9844	.9730	.9571	.9360	.9089
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	.8145	.6561	.5220	.4096	.3164	.2401	.1785	.1296	.0915
	1	.9860	.9477	.8905	.8192	.7383	.6517	.5630	.4752	.3910
	2	.9995	.9963	.9880	.9728	.9492	.9163	.8735	.8208	.7585
	3	1.0000	.9999	.9995	.9984	.9961	.9919	.9850	.9743	.9590
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0	.7738	.5905	.4437	.3277	.2373	.1681	.1160	.0778	.0503
	1	.9774	.9185	.8352	.7373	.6328	.5282	.4284	.3370	.2562
	2	.9988	.9914	.9734	.9421	.8965	.8369	.7648	.6826	.5931
	3	1.0000	.9995	.9978	.9933	.9844	.9692	.9460	.9130	.8688
	4	1.0000	1.0000	.9999	.9997	.9990	.9976	.9947	.9898	.9815
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0	.7351	.5314	.3771	.2621	.1780	.1176	.0754	.0467	.0277
	1	.9672	.8857	.7765	.6554	.5339	.4202	.3191	.2333	.1636
	2	.9978	.9842	.9527	.9011	.8306	.7443	.6471	.5443	.4415
	3	.9999	.9987	.9941	.9830	.9624	.9295	.8826	.9208	.7447
	4	1.0000	.9999	.9996	.9984	.9954	.9891	.9777	.9590	.9308
	5	1.0000	1.0000	1.0000	.9999	.9998	.9993	.9982	.9959	.9917
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	0	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152
	1	.9556	.8503	.7166	.5767	.4449	.3294	.2338	.1586	.1024
	2	.9962	.9743	.9262	.8520	.7564	.6471	.5323	.4199	.3164
	3	.9998	.9973	.9879	.9667	.9294	.8740	.8002	.7102	.6083
	4	1.0000	.9998	.9988	.9953	.9871	.9812	.9444	.9037	.8471
	5	1.0000	1.0000	.9999	.9996	.9987	.9962	.9910	.9812	.9643
	6	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9994	.9984	.9963
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

$$P(Y \leq y) = \sum_{i=0}^{y} \binom{n}{i} p^i (1-p)^{n-i}.$$

n	y	p = .50	.55	.60	.65	.70	.75	.80	.85	.90	.95
1	0	.5000	.4500	.4000	.3500	.3000	.2500	.2000	.1500	.1000	.0500
	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0	.2500	.2025	.1600	.1225	.0900	.0625	.0400	.0225	.0100	.0025
	1	.7500	.6975	.6400	.5775	.5100	.4375	.3600	.2775	.1900	.0975
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	.1250	.0911	.0640	.0429	.0270	.0156	.0080	.0034	.0010	.0001
	1	.5000	.4252	.3520	.2818	.2160	.1562	.1040	.0608	.0280	.0072
	2	.8750	.8336	.7840	.7254	.6570	.5781	.4880	.3859	.2710	.1426
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	.0625	.0410	.0256	.0150	.0081	.0039	.0016	.0005	.0001	.0000
	1	.3125	.2415	.1792	.1265	.0837	.0508	.0272	.0120	.0037	.0005
	2	.6875	.6090	.5248	.4370	.3483	.2617	.1808	.1095	.0523	.0140
	3	.9375	.9085	.8704	.8215	.7599	.6836	.5904	.4780	.3439	.1855
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0	.0312	.0185	.0102	.0053	.0024	.0010	.0003	.0001	.0000	.0000
	1	.1875	.1312	.0870	.0540	.0308	.0156	.0067	.0022	.0005	.0000
	2	.5000	.4069	.3174	.2352	.1631	.1035	.0579	.0266	.0086	.0012
	3	.8125	.7438	.6630	.5716	.4718	.3672	.2627	.1648	.0815	.0226
	4	.9688	.9497	.9222	.8840	.8319	.7627	.6723	.5563	.4095	.2262
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0	.0156	.0083	.0041	.0018	.0007	.0002	.0001	.0000	.0000	.0000
	1	.1094	.0692	.0410	.0223	.0109	.0046	.0016	.0004	.0001	.0000
	2	.3438	.2553	.1792	.1174	.0705	.0376	.0170	.0059	.0013	.0001
	3	.6562	.5585	.4557	.3529	.2557	.1694	.0989	.0473	.0158	.0022
	4	.8906	.8364	.7667	.6809	.5798	.4661	.3446	.2235	.1143	.0328
	5	.9844	.9723	.9533	.9246	.8824	.8220	.7379	.6229	.4686	.2649
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	0	.0078	.0037	.0016	.0006	.0002	.0001	.0000	.0000	.0000	.0000
	1	.0625	.0357	.0188	.0090	.0038	.0013	.0004	.0001	.0000	.0000
	2	.2266	.1529	.0963	.0556	.0288	.0129	.0047	.0012	.0002	.0000
	3	.5000	.3917	.2898	.1998	.1260	.0706	.0333	.0121	.0027	.0002
	4	.7734	.6836	.5801	.4677	.3529	.2436	.1480	.0738	.0257	.0038
	5	.9375	.8976	.8414	.7662	.6706	.5551	.4233	.2834	.1497	.0444
	6	.9922	.9848	.9720	.9510	.9176	.8665	.7903	.6794	.5217	.3017
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Tabla 3. (Continuación).

n	y	p = .05	.10	.15	.20	.25	.30	.35	.40	.45
8	0	.6634	.4305	.2725	.1678	.1001	.0576	.0319	.0168	.0084
	1	.9428	.8131	.6572	.5033	.3671	.2553	.1691	.1064	.0632
	2	.9942	.9619	.8948	.7969	.6785	.5518	.4278	.3154	.2201
	3	.9996	.9950	.9786	.9437	.8862	.8059	.7064	.5941	.4770
	4	1.0000	.9996	.9971	.9896	.9727	.9420	.8939	.8263	.7396
	5	1.0000	1.0000	.9998	.9988	.9958	.9887	.9747	.9502	.9115
	6	1.0000	1.0000	1.0000	.9999	.9996	.9987	.9964	.9915	.9819
	7	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9988	.9993	.9983
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	0	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046
	1	.9288	.7748	.5995	.4362	.3003	.1960	.1211	.0705	.0385
	2	.9916	.9470	.8591	.7382	.6007	.4628	.3373	.2318	.1495
	3	.9994	.9917	.9661	.9144	.8343	.7297	.6089	.4826	.3614
	4	1.0000	.9991	.9944	.9804	.9511	.9012	.8283	.7334	.6214
	5	1.0000	.9999	.9994	.9969	.9900	.9747	.9464	.9006	.8342
	6	1.0000	1.0000	1.0000	.9997	.9987	.9957	.9888	.9750	.9502
	7	1.0000	1.0000	1.0000	1.0000	.9999	.9996	.9986	.9962	.9909
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9992
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	0	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025
	1	.9139	.7361	.5443	.3758	.2440	.1493	.0860	.0464	.0233
	2	.9885	.9298	.8202	.6778	.5256	.3828	.2616	.1673	.0996
	3	.9990	.9872	.9500	.8791	.7759	.6496	.5138	.3823	.2660
	4	.9999	.9984	.9901	.9672	.9219	.8497	.7515	.6331	.5044
	5	1.0000	.9999	.9986	.9936	.9803	.9527	.9051	.8338	.7384
	6	1.0000	1.0000	.9999	.9991	.9965	.9894	.9740	.9452	.8980
	7	1.0000	1.0000	1.0000	.9999	.9996	.9984	.9952	.9877	.9726
	8	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995	.9983	.9955
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
11	0	.5688	.3138	.1673	.0859	.0422	.0198	.0088	.0036	.0014
	1	.8981	.6974	.4922	.3221	.1971	.1130	.0606	.0302	.0139
	2	.9848	.9104	.7788	.6174	.4552	.3127	.2001	.1189	.0652
	3	.9984	.9815	.9306	.8389	.7133	.5696	.4256	.2963	.1911
	4	.9999	.9972	.9841	.9496	.8854	.7897	.6683	.5328	.3971
	5	1.0000	.9997	.9973	.9883	.9657	.9218	.8513	.7535	.6331
	6	1.0000	1.0000	.9997	.9980	.9924	.9784	.9499	.9006	.8262
	7	1.0000	1.0000	1.0000	.9998	.9988	.9957	.9878	.9707	.9390
	8	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9980	.9941	.9852
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9993	.9978
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9998
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

n	y	p = .50	.55	.60	.65	.70	.75	.80	.85	.90	.95
8	0	.0039	.0017	.0007	.0002	.0001	.0000	.0000	.0000	.0000	.0000
	1	.0352	.0181	.0085	.0036	.0013	.0004	.0001	.0000	.0000	.0000
	2	.1445	.0885	.0498	.0253	.0113	.0042	.0012	.0002	.0000	.0000
	3	.3633	.2604	.1737	.1061	.0580	.0273	.0104	.0029	.0004	.0000
	4	.6367	.5230	.4059	.2936	.1941	.1138	.0563	.0214	.0050	.0004
	5	.8555	.7799	.6846	.5722	.4482	.3215	.2031	.1052	.0381	.0058
	6	.9648	.9368	.8936	.8309	.7447	.6329	.4967	.3428	.1869	.0572
	7	.9961	.9916	.9832	.9681	.9424	.8999	.8322	.7275	.5695	.3366
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	0	.0020	.0008	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0195	.0091	.0038	.0014	.0004	.0001	.0000	.0000	.0000	.0000
	2	.0898	.0498	.0250	.0112	.0043	.0013	.0003	.0000	.0000	.0000
	3	.2539	.1658	.0994	.0536	.0253	.0100	.0031	.0006	.0001	.0000
	4	.5000	.3786	.2666	.1717	.0988	.0489	.0196	.0056	.0009	.0000
	5	.7461	.6386	.5174	.3911	.2703	.1657	.0856	.0339	.0083	.0006
	6	.9102	.8505	.7682	.6627	.5372	.3993	.2618	.1409	.0530	.0084
	7	.9805	.9615	.9295	.8789	.8040	.6997	.5638	.4005	.2252	.0712
	8	.9980	.9954	.9899	.9793	.9596	.9249	.8658	.7684	.6126	.3698
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	0	.0010	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0107	.0045	.0017	.0005	.0001	.0000	.0000	.0000	.0000	.0000
	2	.0547	.0274	.0123	.0048	.0016	.0004	.0001	.0000	.0000	.0000
	3	.1719	.1020	.0548	.0260	.0106	.0035	.0009	.0001	.0000	.0000
	4	.3770	.2616	.1662	.0949	.0473	.0197	.0064	.0014	.0001	.0000
	5	.6230	.4956	.3669	.2485	.1503	.0781	.0328	.0099	.0016	.0001
	6	.8281	.7340	.6177	.4862	.3504	.2241	.1209	.0500	.0128	.0010
	7	.9453	.9004	.8327	.7384	.6172	.4744	.3222	.1798	.0702	.0115
	8	.9893	.9767	.9536	.9140	.8507	.7560	.6242	.4557	.2639	.0861
	9	.9990	.9975	.9940	.9865	.9718	.9437	.8926	.8031	.6513	.4013
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
11	0	.0005	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0059	.0022	.0007	.0002	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0327	.0148	.0059	.0020	.0006	.0001	.0000	.0000	.0000	.0000
	3	.1133	.0610	.0293	.0122	.0043	.0012	.0002	.0000	.0000	.0000
	4	.2744	.1738	.0994	.0501	.0216	.0076	.0020	.0003	.0000	.0000
	5	.5000	.3669	.2465	.1487	.0782	.0343	.0117	.0027	.0003	.0000
	6	.7256	.6029	.4672	.3317	.2103	.1146	.0504	.0159	.0028	.0001
	7	.8867	.8089	.7037	.5744	.4304	.2867	.1611	.0694	.0185	.0016
	8	.9673	.9348	.8811	.7999	.6873	.5448	.3826	.2212	.0896	.0152
	9	.9941	.9861	.9698	.9394	.8870	.8029	.6779	.5078	.3026	.1019
	10	.9995	.9986	.9964	.9912	.9802	.9578	.9141	.8327	.6862	.4312
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Tabla 3, (Continuación).

n	y	p = .05	.10	.15	.20	.25	.30	.35	.40	.45
12	0	.5404	.2824	.1422	.0687	.0317	.0138	.0057	.0022	.0008
	1	.8816	.6590	.4435	.2749	.1584	.0850	.0424	.0424	.0083
	2	.9804	.8891	.7358	.5583	.3907	.2528	.1513	.0834	.0421
	3	.9978	.9744	.9078	.7946	.6488	.4925	.3467	.2253	.1345
	4	.9998	.9957	.9761	.9274	.8424	.7237	.5833	.4382	.3044
	5	1.0000	.9995	.9954	.9806	.9456	.8822	.7873	.6652	.5269
	6	1.0000	.9999	.9993	.9961	.9857	.9614	.9154	.8418	.7393
	7	1.0000	1.0000	.9999	.9994	.9972	.9905	.9745	.9427	.8883
	8	1.0000	1.0000	1.0000	.9999	.9996	.9983	.9944	.9847	.9644
	9	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9992	.9972	.9921
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9989
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
13	0	.5133	.2542	.1209	.0550	.0238	.0097	.0037	.0013	.0004
	1	.8646	.6213	.3983	.2336	.1267	.0637	.0296	.0126	.0049
	2	.9755	.8661	.7296	.5017	.3326	.2025	.1132	.0579	.0269
	3	.9969	.9658	.9033	.7473	.5843	.4206	.2783	.1686	.0929
	4	.9997	.9935	.9740	.9009	.7940	.6543	.5005	.3530	.2279
	5	1.0000	.9991	.9947	.9700	.9198	.8346	.7159	.5744	.4268
	6	1.0000	.9999	.9987	.9930	.9757	.9376	.8705	.7712	.6437
	7	1.0000	1.0000	.9998	.9988	.9944	.9818	.9538	.9023	.8212
	8	1.0000	1.0000	1.0000	.9998	.9990	.9960	.9874	.9679	.9302
	9	1.0000	1.0000	1.0000	1.0000	.9999	.9993	.9975	.9922	.9797
	10	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9987	.9959
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
14	0	.4877	.2288	.1028	.0440	.0178	.0068	.0024	.0008	.0002
	1	.8470	.5846	.3567	.1979	.1010	.0475	.0205	.0081	.0029
	2	.9699	.8416	.6479	.4481	.2811	.1608	.0839	.0398	.0170
	3	.9958	.9559	.8535	.6982	.5213	.3552	.2205	.1243	.0632
	4	.9996	.9908	.9533	.8702	.7415	.5842	.4227	.2793	.1672
	5	1.0000	.9985	.9885	.9561	.8883	.7805	.6405	.4859	.3373
	6	1.0000	.9998	.9978	.9884	.9617	.9067	.8164	.6925	.5461
	7	1.0000	1.0000	.9997	.9976	.9897	.9685	.9247	.8499	.7414
	8	1.0000	1.0000	1.0000	.9996	.9978	.9917	.9757	.9417	.8811
	9	1.0000	1.0000	1.0000	1.0000	.9997	.9983	.9940	.9825	.9574
	10	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9989	.9961	.9886
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9978
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

n	y	p = .50	.55	.60	.65	.70	.75	.80	.85	.90	.95
12	0	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0032	.0011	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0193	.0079	.0028	.0008	.0002	.0000	.0000	.0000	.0000	.0000
	3	.0730	.0356	.0153	.0056	.0017	.0004	.0001	.0000	.0000	.0000
	4	.1938	.1117	.0573	.0255	.0095	.0028	.0006	.0001	.0000	.0000
	5	.3872	.2607	.1582	.0846	.0386	.0143	.0039	.0007	.0001	.0000
	6	.6128	.4731	.3348	.2127	.1178	.0544	.0194	.0046	.0005	.0000
	7	.8062	.6956	.5618	.4167	.2763	.1576	.0726	.0239	.0043	.0002
	8	.9270	.8655	.7747	.6533	.5075	.3512	.2054	.0922	.0256	.0022
	9	.9807	.9579	.9166	.8487	.7472	.6093	.4417	.2642	.1109	.0196
	10	.9968	.9917	.9804	.9576	.9150	.8416	.7251	.5565	.3410	.1184
	11	.9998	.9992	.9978	.9943	.9862	.9683	.9313	.8578	.7176	.4596
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
13	0	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0017	.0005	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0112	.0041	.0013	.0003	.0001	.0000	.0000	.0000	.0000	.0000
	3	.0461	.0203	.0078	.0025	.0007	.0001	.0000	.0000	.0000	.0000
	4	.1334	.0698	.0321	.0126	.0040	.0010	.0002	.0000	.0000	.0000
	5	.2905	.1788	.0977	.0462	.0182	.0056	.0012	.0002	.0000	.0000
	6	.5000	.3563	.2288	.1295	.0624	.0243	.0070	.0013	.0001	.0000
	7	.7095	.5732	.4256	.2841	.1654	.0802	.0300	.0053	.0009	.0000
	8	.8666	.7721	.6470	.4995	.3457	.2060	.0991	.0260	.0065	.0003
	9	.9539	.9071	.8314	.7217	.5794	.4157	.2527	.0967	.0342	.0031
	10	.9888	.9731	.9421	.8868	.7975	.6674	.4983	.2704	.1339	.0245
	11	.9983	.9951	.9874	.9704	.9363	.8733	.7664	.6017	.3787	.1354
	12	.9999	.9996	.9987	.9963	.9903	.9762	.9450	.8791	.7458	.4867
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
14	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0009	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0065	.0022	.0006	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0287	.0114	.0039	.0011	.0002	.0000	.0000	.0000	.0000	.0000
	4	.0898	.0462	.0175	.0060	.0017	.0003	.0000	.0000	.0000	.0000
	5	.2120	.1189	.0583	.0243	.0083	.0022	.0004	.0000	.0000	.0000
	6	.3953	.2586	.1501	.0753	.0315	.0103	.0024	.0003	.0000	.0000
	7	.6047	.4539	.3075	.1836	.0933	.0383	.0116	.0022	.0002	.0000
	8	.7880	.6627	.5141	.3595	.2195	.1117	.0439	.0115	.0015	.0000
	9	.9102	.8328	.7207	.5773	.4158	.2585	.1298	.0467	.0092	.0004
	10	.9713	.9368	.8757	.7795	.6448	.4787	.3018	.1465	.0441	.0042
	11	.9935	.9830	.9602	.9161	.8392	.7189	.5519	.3521	.1584	.0301
	12	.9991	.9971	.9919	.9795	.9525	.8990	.8021	.6433	.4154	.1530
	13	.9999	.9998	.9992	.9976	.9932	.9822	.9560	.8972	.7712	.5123
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Tabla 3. (Continuación).

n	y	p = .05	.10	.15	.20	.25	.30	.35	.40	.45
15	0	.4633	.2059	.0874	.0352	.0134	.0047	.0016	.0005	.0001
	1	.8290	.5490	.3186	.1671	.0802	.0353	.0142	.0052	.0017
	2	.9638	.8159	.6042	.3980	.2361	.1268	.0617	.0271	.0107
	3	.9945	.9444	.8227	.6482	.4613	.2969	.1727	.0905	.0424
	4	.9994	.9873	.9383	.8358	.6865	.5155	.3519	.2173	.1204
	5	.9999	.9978	.9832	.9389	.8516	.7216	.5643	.4032	.2608
	6	1.0000	.9997	.9964	.9819	.9434	.8689	.7548	.6098	.4522
	7	1.0000	1.0000	.9994	.9958	.9827	.9500	.8868	.7869	.6535
	8	1.0000	1.0000	.9999	.9992	.9958	.9848	.9578	.9050	.8182
	9	1.0000	1.0000	1.0000	.9999	.9992	.9963	.9876	.9662	.9231
	10	1.0000	1.0000	1.0000	1.0000	.9999	.9993	.9972	.9907	.9745
	11	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995	.9981	.9937
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9989
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
16	0	.4401	.1853	.0743	.0281	.0100	.0033	.0010	.0003	.0001
	1	.8108	.5147	.2839	.1407	.0635	.0261	.0098	.0033	.0010
	2	.9571	.7892	.5614	.3518	.1971	.0994	.0451	.0183	.0066
	3	.9930	.9316	.7899	.5981	.4050	.2459	.1339	.0651	.0281
	4	.9991	.9830	.9209	.7982	.6302	.4499	.2892	.1666	.0853
	5	.9999	.9967	.9763	.9183	.8103	.6598	.4900	.3288	.1976
	6	1.0000	.9995	.9934	.9733	.9204	.8247	.6881	.5272	.3660
	7	1.0000	.9999	.9989	.9930	.9729	.9256	.8406	.7161	.5629
	8	1.0000	1.0000	.9998	.9985	.9925	.9743	.9329	.8577	.7441
	9	1.0000	1.0000	1.0000	.9998	.9984	.9929	.9771	.9417	.8759
	10	1.0000	1.0000	1.0000	1.0000	.9997	.9984	.9938	.9809	.9514
	11	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9987	.9951	.9851
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9991	.9965
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9994
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

n	y	p = .50	.55	.60	.65	.70	.75	.80	.85	.90	.95
15	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0005	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0037	.0011	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0176	.0063	.0019	.0005	.0001	.0000	.0000	.0000	.0000	.0000
	4	.0592	.0255	.0093	.0028	.0007	.0001	.0000	.0000	.0000	.0000
	5	.1509	.0769	.0338	.0124	.0037	.0008	.0001	.0000	.0000	.0000
	6	.3036	.1818	.0950	.0422	.0152	.0042	.0008	.0001	.0000	.0000
	7	.5000	.3465	.2131	.1132	.0500	.0173	.0042	.0006	.0000	.0000
	8	.6964	.5478	.3902	.2452	.1311	.0566	.0181	.0036	.0003	.0000
	9	.8491	.7392	.5968	.4357	.2784	.1484	.0611	.0168	.0022	.0001
	10	.9408	.8796	.7827	.6481	.4845	.3135	.1642	.0617	.0127	.0006
	11	.9824	.9576	.9095	.8273	.7031	.5387	.3518	.1773	.0556	.0055
	12	.9963	.9893	.9729	.9383	.8732	.7639	.6020	.3958	.1841	.0362
	13	.9995	.9983	.9948	.9858	.9647	.9198	.8329	.6814	.4510	.1710
	14	1.0000	.9999	.9995	.9984	.9953	.9866	.9648	.9126	.7941	.5367
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
16	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0021	.0006	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0106	.0035	.0009	.0002	.0000	.0000	.0000	.0000	.0000	.0000
	4	.0384	.0149	.0049	.0013	.0003	.0000	.0000	.0000	.0000	.0000
	5	.1051	.0486	.0191	.0062	.0016	.0003	.0000	.0000	.0000	.0000
	6	.2272	.1241	.0583	.0229	.0071	.0016	.0002	.0000	.0000	.0000
	7	.4018	.2559	.1423	.0671	.0257	.0075	.0015	.0002	.0000	.0000
	8	.5982	.4371	.2839	.1594	.0744	.0271	.0070	.0011	.0001	.0000
	9	.7228	.6340	.4728	.3119	.1753	.0796	.0267	.0056	.0005	.0000
	10	.8949	.8024	.6712	.5100	.3402	.1897	.0817	.0235	.0033	.0001
	11	.9616	.9147	.8334	.7108	.5501	.3698	.2018	.0791	.0170	.0009
	12	.9894	.9719	.9349	.8661	.7541	.5950	.4019	.2101	.0684	.0070
	13	.9979	.9934	.9817	.9549	.9006	.8729	.6482	.4386	.2108	.0429
	14	.9997	.9990	.9967	.9902	.9739	.9365	.8593	.7161	.4853	.1892
	15	1.0000	.9999	.9997	.9990	.9967	.9900	.9719	.9257	.8147	.5599
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Tabla 3. (Continuación).

n	y	p = .05	.10	.15	.20	.25	.30	.35	.40	.45
17	0	.4181	.1668	.0631	.0225	.0075	.0023	.0007	.0002	.0000
	1	.7922	.4818	.2525	.1182	.0501	.0193	.0067	.0021	.0006
	2	.9497	.7618	.5198	.3096	.1637	.0774	.0327	.0123	.0041
	3	.9912	.9174	.7556	.5489	.3530	.2019	.1028	.0464	.0184
	4	.9988	.9779	.9013	.7582	.5739	.3887	.2348	.1260	.0596
	5	.9999	.9953	.9681	.8943	.7653	.5968	.4197	.2639	.1471
	6	1.0000	.9992	.9917	.9623	.8929	.7752	.6188	.4478	.2902
	7	1.0000	.9999	.9983	.9891	.9598	.8954	.7872	.6405	.4743
	8	1.0000	1.0000	.9997	.9974	.9876	.9597	.9006	.8011	.6626
	9	1.0000	1.0000	1.0000	.9995	.9969	.9873	.9617	.9081	.8166
	10	1.0000	1.0000	1.0000	.9999	.9994	.9968	.9880	.9652	.9174
	11	1.0000	1.0000	1.0000	1.0000	.9999	.9993	.9970	.9894	.9699
	12	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9975	.9914
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995	.9981
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
18	0	.3972	.1501	.0536	.0180	.0056	.0016	.0004	.0001	.0000
	1	.7735	.4503	.2241	.0991	.0395	.0142	.0046	.0013	.0003
	2	.9419	.7338	.4797	.2713	.1353	.0600	.0236	.0082	.0025
	3	.9891	.9018	.7202	.5010	.3057	.1646	.0783	.0328	.0120
	4	.9985	.9718	.8794	.7164	.5187	.3327	.1886	.0942	.0411
	5	.9998	.9936	.9581	.8671	.7175	.5344	.3550	.2088	.1077
	6	1.0000	.9988	.9882	.9487	.8610	.7217	.5491	.3743	.2258
	7	1.0000	.9998	.9973	.9837	.9431	.8593	.7283	.5634	.3915
	8	1.0000	1.0000	.9995	.9957	.9807	.9404	.8609	.7368	.5778
	9	1.0000	1.0000	.9999	.9991	.9946	.9790	.9403	.8653	.7473
	10	1.0000	1.0000	1.0000	.9998	.9988	.9939	.9788	.9424	.8720
	11	1.0000	1.0000	1.0000	1.0000	.9998	.9986	.9938	.9797	.9463
	12	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9986	.9942	.9817
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9987	.9951
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9998	.9990
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

n	y	p = .50	.55	.60	.65	.70	.75	.80	.85	.90	.95
17	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0012	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0064	.0019	.0005	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	4	.0245	.0086	.0025	.0006	.0001	.0000	.0000	.0000	.0000	.0000
	5	.0717	.0301	.0106	.0030	.0007	.0001	.0000	.0000	.0000	.0000
	6	.1662	.0826	.0348	.0120	.0032	.0006	.0001	.0000	.0000	.0000
	7	.3145	.1834	.0919	.0383	.0127	.0031	.0005	.0000	.0000	.0000
	8	.5000	.3374	.1989	.0994	.0403	.0124	.0026	.0003	.0000	.0000
	9	.6855	.5257	.3595	.2128	.1046	.0402	.0109	.0017	.0001	.0000
	10	.8338	.7098	.5522	.3812	.2248	.1071	.0377	.0083	.0008	.0000
	11	.9283	.8529	.7361	.5803	.4032	.2347	.1057	.0319	.0047	.0001
	12	.9755	.9404	.8740	.7652	.6113	.4261	.2418	.0987	.0221	.0012
	13	.9936	.9816	.9536	.8972	.7981	.6470	.4511	.2444	.0826	.0088
	14	.9988	.9959	.9877	.9673	.9226	.8363	.6904	.4802	.2382	.0503
	15	.9999	.9994	.9979	.9933	.9807	.9499	.8818	.7475	.5182	.2078
	16	1.0000	1.0000	.9998	.9993	.9977	.9925	.9775	.9369	.8332	.5819
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
18	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0007	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0038	.0010	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	4	.0154	.0049	.0013	.0003	.0000	.0000	.0000	.0000	.0000	.0000
	5	.0481	.0183	.0058	.0014	.0003	.0000	.0000	.0000	.0000	.0000
	6	.1189	.0537	.0203	.0062	.0014	.0002	.0000	.0000	.0000	.0000
	7	.2403	.1280	.0576	.0212	.0061	.0012	.0002	.0000	.0000	.0000
	8	.4073	.2527	.1347	.0597	.0210	.0054	.0009	.0001	.0000	.0000
	9	.5927	.4222	.2632	.1391	.0596	.0193	.0043	.0005	.0000	.0000
	10	.7597	.6085	.4366	.2717	.1407	.0569	.0163	.0027	.0002	.0000
	11	.8811	.7742	.6257	.4509	.2783	.1390	.0513	.0118	.0012	.0000
	12	.9519	.8923	.7912	.6450	.4656	.2825	.1329	.0419	.0064	.0002
	13	.9846	.9589	.9058	.8114	.6673	.4813	.2836	.1206	.0282	.0015
	14	.9962	.9880	.9672	.9217	.8354	.6943	.4990	.2798	.0982	.0109
	15	.9993	.9975	.9918	.9764	.9400	.8647	.7287	.5203	.2662	.0581
	16	.9999	.9997	.9987	.9954	.9858	.9605	.9009	.7759	.5497	.2265
	17	1.0000	1.0000	.9999	.9996	.9984	.9944	.9820	.9464	.8499	.6028
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Tabla 3. (Continuación).

n	y	p = .05	.10	.15	.20	.25	.30	.35	.40	.45
19	0	.3774	.1351	.0456	.0144	.0042	.0011	.0003	.0001	.0000
	1	.7547	.4203	.1985	.0829	.0310	.0104	.0031	.0008	.0002
	2	.9335	.7054	.4413	.2369	.1113	.0462	.0170	.0055	.0015
	3	.9869	.8850	.6841	.4551	.2631	.1332	.0591	.0230	.0077
	4	.9980	.9648	.8556	.6733	.4654	.2822	.1500	.0696	.0280
	5	.9998	.9914	.9463	.8369	.6678	.4739	.2968	.1629	.0777
	6	1.0000	.9983	.9837	.9324	.8251	.6655	.4812	.3081	.1727
	7	1.0000	.9997	.9959	.9767	.9225	.8180	.6656	.4878	.3169
	8	1.0000	1.0000	.9992	.9933	.9713	.9161	.8145	.6675	.4940
	9	1.0000	1.0000	.9999	.9984	.9911	.9674	.9125	.8139	.6710
	10	1.0000	1.0000	1.0000	.9997	.9977	.9895	.9653	.9115	.8159
	11	1.0000	1.0000	1.0000	1.0000	.9995	.9972	.9886	.9648	.9129
	12	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9969	.9884	.9658
	13	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9993	.9969	.9891
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9994	.9972
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
20	0	.3585	.1216	.0388	.0115	.0032	.0008	.0002	.0000	.0000
	1	.7358	.3917	.1756	.0692	.0243	.0076	.0021	.0005	.0001
	2	.9245	.6769	.4049	.2061	.0913	.0355	.0121	.0036	.0009
	3	.9841	.8670	.6477	.4114	.2252	.1071	.0444	.0160	.0049
	4	.9974	.9568	.8298	.6296	.4148	.2375	.1182	.0510	.0189
	5	.9997	.9887	.9327	.8042	.6172	.4164	.2454	.1256	.0553
	6	1.0000	.9976	.9781	.9133	.7858	.6080	.4166	.2500	.1299
	7	1.0000	.9996	.9941	.9679	.8982	.7723	.6010	.4159	.2520
	8	1.0000	.9999	.9987	.9900	.9591	.8867	.7624	.5956	.4143
	9	1.0000	1.0000	.9998	.9974	.9861	.9520	.8782	.7553	.5914
	10	1.0000	1.0000	1.0000	.9994	.9961	.9829	.9468	.8725	.7507
	11	1.0000	1.0000	1.0000	.9999	.9991	.9949	.9804	.9435	.8692
	12	1.0000	1.0000	1.0000	1.0000	.9998	.9987	.9940	.9790	.9420
	13	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9985	.9935	.9786
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9984	.9936
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9997	.9985
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9997
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

n	y	p = .50	.55	.60	.65	.70	.75	.80	.85	.90	.95
19	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0004	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0022	.0005	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	4	.0096	.0028	.0006	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	5	.0318	.0109	.0031	.0007	.0001	.0000	.0000	.0000	.0000	.0000
	6	.0835	.0342	.0116	.0031	.0006	.0001	.0000	.0000	.0000	.0000
	7	.1796	.0871	.0352	.0114	.0028	.0005	.0000	.0000	.0000	.0000
	8	.3238	.1841	.0885	.0347	.0105	.0023	.0003	.0000	.0000	.0000
	9	.5000	.3290	.1861	.0875	.0326	.0089	.0016	.0001	.0000	.0000
	10	.6762	.5060	.3325	.1855	.0839	.0287	.0067	.0008	.0000	.0000
	11	.8204	.6831	.5122	.3344	.1820	.0775	.0233	.0041	.0003	.0000
	12	.9165	.8273	.6919	.5188	.3345	.1749	.0676	.0163	.0017	.0000
	13	.9682	.9223	.8371	.7032	.5261	.3322	.1631	.0537	.0086	.0002
	14	.9904	.9720	.9304	.8500	.7178	.5346	.3267	.1444	.0352	.0020
	15	.9978	.9923	.9770	.9409	.8668	.7369	.5449	.3159	.1150	.0132
	16	.9996	.9985	.9945	.9830	.9538	.8887	.7631	.5587	.2946	.0665
	17	1.0000	.9998	.9992	.9969	.9896	.9690	.9171	.8015	.5797	.2453
	18	1.0000	1.0000	.9999	.9997	.9989	.9958	.9856	.9544	.8649	.6226
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
20	0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0013	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	4	.0059	.0015	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	5	.0207	.0064	.0016	.0003	.0000	.0000	.0000	.0000	.0000	.0000
	6	.0577	.0214	.0065	.0015	.0003	.0000	.0000	.0000	.0000	.0000
	7	.1316	.0580	.0210	.0060	.0013	.0002	.0000	.0000	.0000	.0000
	8	.2517	.1308	.0565	.0196	.0051	.0009	.0001	.0000	.0000	.0000
	9	.4119	.2493	.1275	.0532	.0171	.0039	.0006	.0000	.0000	.0000
	10	.5881	.4086	.2447	.1218	.0480	.0139	.0026	.0002	.0000	.0000
	11	.7483	.5857	.4044	.2376	.1133	.0409	.0100	.0013	.0001	.0000
	12	.8684	.7480	.5841	.3990	.2277	.1018	.0321	.0059	.0004	.0000
	13	.9423	.8701	.7500	.5834	.3920	.2142	.0867	.0219	.0024	.0000
	14	.9793	.9447	.8744	.7546	.5836	.3828	.1958	.0673	.0113	.0003
	15	.9941	.9811	.9490	.8818	.7625	.5852	.3704	.1702	.0432	.0026
	16	.9987	.9951	.9840	.9556	.8929	.7748	.5886	.3523	.1330	.0159
	17	.9998	.9991	.9964	.9879	.9645	.9087	.7939	.5951	.3231	.0755
	18	1.0000	.9999	.9995	.9979	.9924	.9757	.9308	.8244	.6083	.2642
	19	1.0000	1.0000	1.0000	.9998	.9992	.9968	.9885	.9612	.8784	.6415
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Si $n > 20$, el r-ésimo cuantil se puede aproximar por:

$$y_r = np + W_r \sqrt{np(1-p)},$$ donde W_r, es el r-ésimo cuantil de una $N(0,1)$.

Tabla 4. Cuantiles de la estadística Kolmogoroff-Smirnoff.

	p = .80	.90	.95	.98	.99		p = .80	.90	.95	.98	.99
n = 1	.900	.950	.975	.990	.995	n = 21	.226	.259	.287	.321	.344
2	.684	.776	.842	.900	.929	22	.221	.253	.281	.314	.337
3	.565	.636	.708	.785	.829	23	.216	.247	.275	.307	.330
4	.493	.565	.624	.689	.734	24	.212	.242	.269	.301	.323
5	.447	.509	.563	.627	.669	25	.208	.238	.264	.295	.317
6	.410	.468	.519	.577	.617	26	.204	.233	.259	.290	.311
7	.381	.436	.483	.538	.576	27	.200	.229	.254	.284	.305
8	.358	.410	.454	.507	.542	28	.197	.225	.250	.279	.300
9	.339	.387	.430	.480	.513	29	.193	.221	.246	.275	.295
10	.323	.369	.409	.457	.489	30	.190	.218	.242	.270	.290
11	.308	.352	.391	.437	.468	31	.187	.214	.238	.266	.285
12	.296	.338	.375	.419	.449	32	.184	.211	.234	.262	.281
13	.285	.325	.361	.404	.432	33	.182	.208	.231	.258	.277
14	.275	.314	.349	.390	.418	34	.179	.205	.227	.254	.273
15	.266	.304	.338	.377	.404	35	.177	.202	.224	.251	.269
16	.258	.295	.327	.366	.392	36	.174	.199	.221	.247	.265
17	.250	.286	.318	.355	.381	37	.172	.196	.218	.244	.262
18	.244	.279	.309	.346	.371	38	.170	.194	.215	.241	.258
19	.237	.271	.301	.337	.361	39	.168	.191	.213	.238	.255
20	.232	.265	.294	.329	.352	40	.165	.189	.210	.235	.252
						n > 40	$\frac{1.07}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.52}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Tabla 5. Cuantiles de la Estadística Lilliefors.

	$p = .80$	.85	.90	.95	.99
$n = 4$	.300	.319	.352	.381	.417
5	.285	.299	.315	.337	.405
6	.265	.277	.294	.319	.364
7	.247	.258	.276	.300	.348
8	.233	.244	.261	.285	.331
9	.223	.233	.249	.271	.311
10	.215	.224	.239	.258	.294
11	.206	.217	.230	.249	.284
12	.199	.212	.223	.242	.275
13	.190	.202	.214	.234	.268
14	.183	.194	.207	.227	.261
15	.177	.187	.201	.220	.257
16	.173	.182	.195	.213	.250
17	.169	.177	.189	.206	.245
18	.166	.173	.184	.200	.239
19	.163	.169	.179	.195	.235
20	.160	.166	.174	.190	.231
25	.142	.147	.158	.173	.200
30	.131	.136	.144	.161	.187
> 30	$\frac{.736}{\sqrt{n}}$	$\frac{.768}{\sqrt{n}}$	$\frac{.805}{\sqrt{n}}$	$\frac{.886}{\sqrt{n}}$	$\frac{1.031}{\sqrt{n}}$

Tabla 6. Cuantiles del total de número de corridas de la prueba Wald-Wolfowitz.

N_1	N_2	$W_{.005}$	$W_{.01}$	$W_{.025}$	$W_{.05}$	$W_{.10}$	$W_{.90}$	$W_{.95}$	$W_{.975}$	$W_{.99}$	$W_{.995}$
2	5	—	—	—	—	3	—	—	—	—	—
	8	—	—	—	3	3	—	—	—	—	—
	11	—	—	—	3	3	—	—	—	—	—
	14	—	—	3	3	3	—	—	—	—	—
	17	—	—	3	3	3	—	—	—	—	—
	20	—	3	3	3	4	—	—	—	—	—
5	5	—	3	3	4	4	8	8	9	9	—
	8	3	3	4	4	5	9	10	10	—	—
	11	4	4	5	5	6	10	—	—	—	—
	14	4	4	5	6	6	—	—	—	—	—
	17	4	5	5	6	7	—	—	—	—	—
	20	5	5	6	6	7	—	—	—	—	—
8	8	4	5	5	6	6	12	12	13	13	14
	11	5	6	6	7	8	13	14	14	15	15
	14	6	6	7	8	8	14	15	15	16	16
	17	6	7	8	8	9	15	15	16	—	—
	20	7	7	8	9	10	15	16	16	—	—
11	11	6	7	8	8	9	15	16	16	17	18
	14	7	8	9	9	10	16	17	18	19	19
	17	8	9	10	10	11	17	18	19	20	21
	20	9	9	10	11	12	18	19	20	21	21
14	14	8	9	10	11	12	18	19	20	21	22
	17	9	10	11	12	13	20	21	22	23	23
	20	10	11	12	13	14	21	22	23	24	24
17	17	11	11	12	13	14	22	23	24	25	25
	20	12	12	14	14	16	23	24	25	26	27
20	20	13	14	15	16	17	25	26	27	28	29

Si n ó m mayor que 20, el cuantil W_p se aproxima por:

$$W_p = \frac{2mn}{m+n} + 1 + X_p \sqrt{\frac{2mn\,(2mn-m-n)}{(m+n)^2\,(m+n-1)}}$$

donde X_p es el p-ésimo cuantil de una N(0,1).

Tabla 7. Cuantiles de la Estadística Wilcoxon.

	$W_{.005}$	$W_{.01}$	$W_{.025}$	$W_{.05}$	$W_{.10}$	$W_{.20}$	$W_{.30}$	$W_{.40}$	$W_{.50}$	$\frac{n(n+1)}{2}$
$n = 4$	0	0	0	0	1	3	3	4	5	10
5	0	0	0	1	3	4	5	6	7.5	15
6	0	0	1	3	4	6	8	9	10.5	21
7	0	1	3	4	6	9	11	12	14	28
8	1	2	4	6	9	12	14	16	18	36
9	2	4	6	9	11	15	18	20	22.5	45
10	4	6	9	11	15	19	22	25	27.5	55
11	6	8	11	14	18	23	27	30	33	66
12	8	10	14	18	22	28	32	36	39	78
13	10	13	18	22	27	33	38	42	45.5	91
14	13	16	22	26	32	39	44	48	52.5	105
15	16	20	26	31	37	45	51	55	60	120
16	20	24	30	36	43	51	58	63	68	136
17	24	28	35	42	49	58	65	71	76.5	153
18	28	33	41	48	56	66	73	80	85.5	171
19	33	38	47	54	63	74	82	89	95	190
20	38	44	53	61	70	82	91	98	105	210

Si $n > 20$, el cuantil W_p se aproxima por:

$$W_p = [\, n(n=1)/4 \,] + X_p \sqrt{n(n+1)(2n+1)/24}$$

donde X_p es el cuantil de una normal.

si $p > .50$ entonces:

$$W_p = n(n+1)/2 - W_{1-p}.$$

Tabla 8. Cuantiles de la Estadística Mann-Whitney

n	p	m = 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	.005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
2	.01	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	2	2
	.025	0	0	0	0	0	0	1	1	1	1	2	2	2	2	2	3	3	3	3
	.05	0	0	0	1	1	1	2	2	2	2	3	3	4	4	4	4	5	5	5
	.10	0	1	1	2	2	2	3	3	4	4	5	5	5	6	6	7	7	8	8
	.001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
	.005	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	3	3	4	4
3	.01	0	0	0	0	0	1	1	2	2	2	3	3	3	4	4	5	5	5	6
	.025	0	0	0	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9
	.05	0	1	1	2	3	3	4	5	5	6	6	7	8	8	9	10	10	11	12
	.10	1	2	2	3	4	5	6	6	7	8	9	10	11	11	12	13	14	15	16
	.001	0	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	4	4	4
	.005	0	0	0	0	1	1	2	2	3	3	4	4	5	6	6	7	7	8	9
4	.01	0	0	0	1	2	2	3	4	4	5	6	6	7	9	8	9	10	10	11
	.025	0	0	1	2	3	4	5	5	6	7	8	9	10	11	12	12	13	14	15
	.05	0	1	2	3	4	5	6	7	8	9	10	11	12	13	15	16	17	18	19
	.10	1	2	4	5	6	7	8	10	11	12	13	14	16	17	18	19	21	22	23

Cuantiles superiores pueden ser encontrados usando:

$$w_{1-p} = nm - w_p$$

n	p	m = 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	0	0	0	0	0	1	2	2	3	3	4	4	5	6	6	7	8	8
	.005	0	0	0	1	2	2	3	4	5	6	7	8	8	9	10	11	12	13	14
5	.01	0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	.025	0	1	2	3	4	6	7	8	9	10	12	13	14	15	16	18	19	20	21
	.05	1	2	3	5	6	7	9	10	12	13	14	16	17	19	20	21	23	24	26
	.10	2	3	5	6	8	9	11	13	14	16	18	19	21	23	24	26	28	29	31
	.001	0	0	0	0	0	0	2	3	4	5	5	6	7	8	9	10	11	12	13
	.005	0	0	1	2	3	4	5	6	7	8	10	11	12	13	14	16	17	18	19
6	.01	0	0	2	3	4	5	7	8	9	10	12	13	14	16	17	19	20	21	23
	.025	0	2	3	4	6	7	9	11	12	14	15	17	18	20	22	23	25	26	28
	.05	1	3	4	6	8	9	11	13	15	17	18	20	22	24	26	27	29	31	33
	.10	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	35	37	39
	.001	0	0	0	0	1	2	3	4	6	7	8	9	10	11	12	14	15	16	17
	.005	0	0	1	2	4	5	7	8	10	11	13	14	16	17	19	20	22	23	25
7	.01	0	1	2	4	5	7	8	10	12	13	15	17	18	20	22	24	25	27	29
	.025	0	2	4	6	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35
	.05	1	3	5	7	9	12	14	16	18	20	22	25	27	29	31	34	36	38	40
	.10	2	5	7	9	12	14	17	19	22	24	27	29	32	34	37	39	42	44	47
	.001	0	0	0	1	2	3	5	6	7	9	10	12	13	15	16	18	19	21	22
	.005	0	0	2	3	5	7	8	10	12	14	16	18	19	21	23	25	27	29	31
8	.01	0	1	3	5	7	8	10	12	14	16	18	21	23	25	27	29	31	33	35
	.025	1	3	5	7	9	11	14	16	18	20	23	25	27	30	32	35	37	39	42
	.05	2	4	6	9	11	14	16	19	21	24	27	29	32	34	37	40	42	45	46
	.10	3	6	8	11	14	17	20	23	25	28	31	34	37	40	43	46	49	52	55

Tabla 8. (Continuación).

n	p	m = 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	0	0	2	3	4	6	8	9	11	13	15	16	18	20	22	24	26	27
	.005	0	1	2	4	6	8	10	12	14	17	19	21	23	25	28	30	32	34	37
9	.01	0	2	4	6	8	10	12	15	17	19	22	24	27	29	32	34	37	39	41
	.025	1	3	5	8	11	13	16	18	21	24	27	29	32	35	38	40	43	46	49
	.05	2	5	7	10	13	16	19	22	25	28	31	34	37	40	43	46	49	52	55
	.10	3	6	10	13	16	19	23	26	29	32	36	39	42	46	49	53	56	59	63
	.001	0	0	1	2	4	6	7	9	11	13	15	18	20	22	24	26	28	30	33
	.005	0	1	3	5	7	10	12	14	17	19	22	25	27	30	32	35	38	40	43
10	.01	0	2	4	7	9	12	14	17	20	23	25	28	31	34	37	39	42	45	48
	.025	1	4	6	9	12	15	18	21	24	27	30	34	37	40	43	46	49	53	56
	.05	2	5	8	12	15	18	21	25	28	32	35	38	42	45	49	52	56	59	63
	.10	4	7	11	14	18	22	25	29	33	37	40	44	48	52	55	59	63	67	71
	.001	0	0	1	3	5	7	9	11	13	16	18	21	23	25	28	30	33	35	38
	.005	0	1	3	6	8	11	14	17	19	22	25	28	31	34	37	40	43	46	49
11	.01	0	2	5	8	10	13	16	19	23	26	29	32	35	38	42	45	48	51	54
	.025	1	4	7	10	14	17	20	24	27	31	34	38	41	45	48	52	56	59	63
	.05	2	6	9	13	17	20	24	28	32	35	39	43	47	51	55	58	62	66	70
	.10	4	8	12	16	20	24	28	32	37	41	45	49	53	58	62	66	70	74	79
	.001	0	0	1	3	5	8	10	13	15	18	21	24	26	29	32	35	38	41	43
	.005	0	2	4	7	10	13	16	19	22	25	28	32	35	38	42	45	48	52	55
12	.01	0	3	6	9	12	15	18	22	25	29	32	36	39	43	47	50	54	57	61
	.025	2	5	8	12	15	19	23	27	30	34	38	42	46	50	54	58	62	66	70
	.05	3	6	10	14	18	22	27	31	35	39	43	48	52	56	61	65	69	73	78
	.10	5	9	13	18	22	27	31	36	40	45	50	54	59	64	68	73	78	82	87

n	p	m = 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	0	2	4	6	9	12	15	18	21	24	27	30	33	36	39	43	46	49
	.005	0	2	4	8	11	14	18	21	25	28	32	35	39	43	46	50	54	58	61
13	.01	1	3	6	10	13	17	21	24	28	32	36	40	44	48	52	56	60	64	68
	.025	2	5	9	13	17	21	25	29	34	38	42	46	51	55	60	64	68	73	77
	.05	3	7	11	16	20	25	29	34	38	43	48	52	57	62	66	71	76	81	85
	.10	5	10	14	19	24	29	34	39	44	49	54	59	64	69	75	80	85	90	95
	.001	0	0	2	4	7	10	13	16	20	23	26	30	33	37	40	44	47	51	55
	.005	0	2	5	8	12	16	19	23	27	31	35	39	43	47	51	55	59	64	68
14	.01	1	3	7	11	14	18	23	27	31	35	39	44	48	52	57	61	66	70	74
	.025	2	6	10	14	18	23	27	32	37	41	46	51	56	60	65	70	75	79	84
	.05	4	8	12	17	22	27	32	37	42	47	52	57	62	67	72	78	83	88	93
	.10	5	11	16	21	26	32	37	42	48	53	59	64	70	75	81	86	92	98	103
	.001	0	0	2	5	8	11	15	18	22	25	29	33	37	41	44	48	52	56	60
	.005	0	3	6	9	13	17	21	25	30	34	38	43	47	52	56	61	65	70	74
15	.01	1	4	8	12	16	20	25	29	34	38	43	48	52	57	62	67	71	76	81
	.025	2	6	11	15	20	25	30	35	40	45	50	55	60	65	71	76	81	86	91
	.05	4	8	13	19	24	29	34	40	45	51	56	62	67	73	78	84	89	95	101
	.10	6	11	17	23	28	34	40	46	52	58	64	69	75	81	87	93	99	105	111
	.001	0	0	3	6	9	12	16	20	24	28	32	36	40	44	49	53	57	61	66
	.005	0	3	6	10	14	19	23	28	32	37	42	46	51	56	61	66	71	75	80
16	.01	1	4	8	13	17	22	27	32	37	42	47	52	57	62	67	72	77	83	88
	.025	2	7	12	16	22	27	32	38	43	48	54	60	65	71	76	82	87	93	99
	.05	4	9	15	20	26	31	37	43	49	55	61	66	72	78	84	90	96	102	108
	.10	6	12	18	24	30	37	43	49	55	62	68	75	81	87	94	100	107	113	120

Tabla 8. (Continuación).

n	p	m = 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	.001	0	1	3	6	10	14	18	22	26	30	35	39	44	48	53	58	62	67	71
	.005	0	3	7	11	16	20	25	30	35	40	45	50	55	61	66	71	76	82	87
17	.01	1	5	9	14	19	24	29	34	39	45	50	56	61	67	72	78	83	89	94
	.025	3	7	12	18	23	29	35	40	46	52	58	64	70	76	82	88	94	100	106
	.05	4	10	16	21	27	34	40	46	52	58	65	71	78	84	90	97	103	110	116
	.10	7	13	19	26	32	39	46	53	59	66	73	80	86	93	100	107	114	121	128
	.001	0	1	4	7	11	15	19	24	28	33	38	43	47	52	57	62	67	72	77
	.005	0	3	7	12	17	22	27	32	38	43	48	54	59	65	71	76	82	88	93
18	.01	1	5	10	15	20	25	31	37	42	48	54	60	66	71	77	83	89	95	101
	.025	3	8	13	19	25	31	37	43	49	56	62	68	75	81	87	94	100	107	113
	.05	5	10	17	23	29	36	42	49	56	62	69	76	83	89	96	103	110	117	124
	.10	7	14	21	28	35	42	49	56	63	70	78	85	92	99	107	114	121	129	136
	.001	0	1	4	8	12	16	21	26	30	35	41	46	51	56	61	67	72	78	83
	.005	1	4	8	13	18	23	29	34	40	46	52	58	64	70	75	82	88	94	100
19	.01	2	5	10	16	21	27	33	39	45	51	57	64	70	76	83	89	95	102	108
	.025	3	8	14	20	26	33	39	46	53	59	66	73	79	86	93	100	107	114	120
	.05	5	11	18	24	31	38	45	52	59	66	73	81	88	95	102	110	117	124	131
	.10	8	15	22	29	37	44	52	59	67	74	82	90	98	105	113	121	129	136	144
	.001	0	1	4	8	13	17	22	27	33	38	43	49	55	60	66	71	77	83	89
	.005	1	4	9	14	19	25	31	37	43	49	55	61	68	74	80	87	93	100	106
20	.01	2	6	11	17	23	29	35	41	48	54	61	68	74	81	88	94	101	108	115
	.025	3	9	15	21	28	35	42	49	56	63	70	77	84	91	99	106	113	120	128
	.05	5	12	19	26	33	40	48	55	63	70	78	85	93	101	108	116	124	131	139
	.10	8	16	23	31	39	47	55	63	71	79	87	95	103	111	120	128	136	144	152

Si n ó m > 20 el p-ésimo cuantil se aproxima por:

$$w_p = \frac{nm}{2} + x_p \sqrt{\frac{nm(n+m+1)}{12}}$$

donde X_p es el p-ésimo cuantil de una normal.

Tabla 9. Cuantiles de la Estadística Kruskal-Wallis.

Tamaños de muestra. n_1	n_2	n_3	valor crítico	α
2	1	1	2.7000	.500
2	2	1	3.6000	.200
2	2	2	4.5714	.067
			3.7143	.200
3	1	1	3.2000	.300
3	2	1	4.2857	.100
			3.8571	.133
3	3	2	5.3572	.029
			4.7143	.048
			4.5000	.067
			4.4643	.105
3	3	1	5.1429	.043
			4.5714	.100
			4.0000	.129
3	3	2	6.2500	.011
			5.3611	.032
			5.1389	.061
			4.5556	.100
			4.2500	.121
3	3	3	7.2000	.004
			6.4889	.001
			5.6889	.029
			5.6000	.050
			5.0667	.086
			4.6222	.100
4	1	1	3.5714	.200
4	2	1	4.8214	.057
			4.5000	.076
			4.0179	.114

Tamaños de muestra. n_1	n_2	n_3	valor crítico.	α
4	2	2	6.0000	.014
			5.3333	.033
			5.1250	.052
			4.4583	.100
			4.1667	.105
4	3	1	5.8333	.021
			5.2083	.050
			5.0000	.057
			4.0556	.093
			3.8889	.129
4	3	2	6.4444	.008
			6.3000	.011
			5.4444	.046
			5.4000	.051
			4.5111	.098
			4.4444	.102
4	3	3	6.7455	.010
			6.7091	.013
			5.7909	.046
			5.7273	.050
			4.7091	.092
			4.7000	.101
4	4	1	6.6667	.010
			6.1667	.022
			4.9667	.048
			4.8667	.054
			4.1667	.082
			4.0667	.102

Tabla 9. (Continuación).

4	4	2	7.0364	.006
			6.8727	.011
			5.4545	.046
			5.2364	.052
			4.5545	.098
			4.4455	.103
4	4	3	7.1439	.010
			7.1364	.011
			5.5985	.049
			5.5758	.051
			4.5455	.099
			4.4773	.102
4	4	4	7.6538	.008
			7.5385	.011
			5.6923	.049
			5.6538	.054
			4.6539	.097
			4.5001	.104
5	1	1	3.8571	.143
5	2	1	5.2500	.036
			5.0000	.048
			4.4500	.071
			4.2000	.095
			4.0500	.119
5	2	2	6.5333	.005
			6.1333	.013
			5.1600	.034
			5.0400	.056
			4.3733	.090
			4.2933	.112
5	3	1	6.4000	.012
			4.9600	.048
			4.8711	.052
			4.0178	.095
			3.8400	.123
5	3	2	6.9091	.009
			6.8281	.010
			5.2509	.049
			5.1055	.052
			4.6509	.091
			4.4121	.101
5	3	3	7.0788	.009
			6.9818	.011
			5.6485	.049
			5.5152	.051
			4.5333	.097
			4.4121	.109
5	4	1	6.9545	.008
			6.8400	.011
			4.9855	.044
			4.8600	.056
			3.9873	.098
			3.9600	.102
5	4	2	7.2045	.009
			7.1182	.010
			5.2727	.049
			5.2682	.050
			4.5409	.098
			4.5182	.101
5	4	3	7.4449	.110
			7.3949	.011
			5.6564	.049
			5.6308	.050
			4.5487	.099
			4.5231	.103
5	4	4	7.7604	.009
			7.7440	.011
			5.6571	.049
			5.6176	.050
			4.6187	.100
			4.5527	.102

Tabla 9. (Continuación).

5	5	1	7.3091	.009				5.6264	.051
			6.8364	.011				4.5451	.100
			5.1273	.046				4.5363	.102
			4.9091	.053					
			4.1091	.086	5	5	4	7.8229	.010
			4.0364	.105				7.7914	.010
								5.6657	.049
5	5	2	7.3385	.010				5.6429	.050
			7.2692	.010				4.5229	.100
			5.3385	.047				4.5200	.101
			5.2462	.051					
			4.6231	.097	5	5	5	8.0000	.009
			4.5077	.100				7.9800	.010
								5.7800	.049
5	5	3	7.5780	.010				5.6600	.051
			7.5429	.010				4.5600	.100
			5.7055	.046				4.5000	.102

Tabla 10. Cuantiles de la Estadística de Spearman.

n	*p* = .900	.950	.975	.990	.995	.999
4	.8000	.8000				
5	.7000	.8000	.9000	.9000		
6	.6000	.7714	.8286	.8857	.9429	
7	.5357	.6786	.7450	.8571	.8929	.9643
8	.5000	.6190	.7143	.8095	.8571	.9286
9	.4667	.5833	.6833	.7667	.8167	.9000
10	.4424	.5515	.6364	.7333	.7818	.8667
11	.4182	.5273	.6091	.7000	.7455	.8364
12	.3986	.4965	.5804	.6713	.7273	.8182
13	.3791	.4780	.5549	.6429	.6978	.7912
14	.3626	.4593	.5341	.6220	.6747	.7670
15	.3500	.4429	.5179	.6000	.6536	.7464
16	.3382	.4265	.5000	.5824	.6324	.7265
17	.3260	.4118	.4853	.5637	.6152	.7083
18	.3148	.3994	.4716	.5480	.5975	.6904
19	.3070	.3895	.4579	.5333	.5825	.6737
20	.2977	.3789	.4451	.5203	.5684	.6586
21	.2909	.3688	.4351	.5078	.5545	.6455
22	.2829	.3597	.4241	.4963	.5426	.6318
23	.2767	.3518	.4150	.4852	.5306	.6186
24	.2704	.3435	.4061	.4748	.5200	.6070
25	.2646	.3362	.3977	.4654	.5100	.5962
26	.2588	.3299	.3894	.4564	.5002	.5856
27	.2540	.3236	.3822	.4481	.4915	.5757
28	.2490	.3175	.3749	.4401	.4828	.5660
29	.2443	.3113	.3685	.4320	.4744	.5567
30	.2400	.3059	.3620	.4251	.4665	.5479

Para $n > 30$ se usa la aproximación:

$$w_p \cong \frac{x_p}{\sqrt{n-1}}$$

donde X_p es el p-ésimo cuantil de una normal.

Los cuantiles inferiores se obtienen mediante la relación:

$$W_p = -W_{1-p}$$

Tabla 11. Cuantiles de la Estadística Hotelling-Pabst.

n	$p = .001$	$.005$	$.010$	$.025$	$.050$	$.100$	$\frac{1}{3}n(n^2-1)$
4					2	2	20
5			2	2	4	6	40
6		2	4	6	8	14	70
7	2	6	8	14	18	26	112
8	6	12	16	24	32	42	168
9	12	22	28	38	50	64	240
10	22	36	44	60	74	92	330
11	36	56	66	86	104	128	440
12	52	78	94	120	144	172	572
13	76	110	130	162	190	226	728
14	106	148	172	212	246	290	910
15	142	194	224	270	312	364	1120
16	186	250	284	340	390	450	1360
17	238	314	356	420	480	550	1632
18	300	390	438	512	582	664	1938
19	372	476	532	618	696	790	2280
20	454	574	638	738	826	934	2660
21	546	686	758	870	972	1092	3080
22	652	810	892	1020	1134	1270	3542
23	772	950	1042	1184	1312	1464	4048
24	904	1104	1208	1366	1510	1678	4600
25	1050	1274	1390	1566	1726	1912	5200
26	1212	1462	1590	1786	1960	2168	5850
27	1390	1666	1808	2024	2216	2444	6552
28	1586	1890	2046	2284	2494	2744	7308
29	1800	2134	2306	2564	2796	3068	8120
30	2032	2398	2584	2868	3120	3416	8990

Para $n > 30$ los cuantiles pueden ser aproximados por:

$$w_p \cong \tfrac{1}{3}n(n^2-1) + x_p \cdot \frac{1}{6}\,\frac{n(n^2-1)}{\sqrt{n-1}}$$

donde X_p es el p-ésimo cuantil de una normal.

Cuantiles superiores pueden ser obtenidos usando:

$$W_{1-p} = n(n^2-1)/3 - W_p$$

Tabla 12. Cuantiles de la Estadística de Kendall.

n	$p = .900$	.950	.975	.990	.995
4	4	4	6	6	6
5	6	6	8	8	10
6	7	9	11	11	13
7	9	11	13	15	17
8	10	14	16	18	20
9	12	16	18	22	24
10	15	19	21	25	27
11	17	21	25	29	31
12	18	24	28	34	36
13	22	26	32	38	42
14	23	31	35	41	45
15	27	33	39	47	51
16	28	36	44	50	56
17	32	40	48	56	62
18	35	43	51	61	67
19	37	47	55	65	73
20	40	50	60	70	78
21	42	54	64	76	84
22	45	59	69	81	89
23	49	63	73	87	97
24	52	66	78	92	102
25	56	70	84	98	108
26	59	75	89	105	115
27	61	79	93	111	123
28	66	84	98	116	128
29	68	88	104	124	136
30	73	93	109	129	143

Cuantiles inferiores se obtienen usando:

$$W_p = -W_{1-p}.$$

Tabla 12. (Continuación).

n	$p = .900$	.950	.975	.990	.995
31	75	97	115	135	149
32	80	102	120	142	158
33	84	106	126	150	164
34	87	111	131	155	173
35	91	115	137	163	179
36	94	120	144	170	188
37	98	126	150	176	196
38	103	131	155	183	203
39	107	137	161	191	211
40	110	142	168	198	220

Para $n > 40$ los cuantiles pueden ser aproximados usando:

$$w_p \cong x_p \sqrt{\frac{n(n-1)(2n+5)}{18}}$$

donde X_p es el p-ésimo cuantil de una normal.

BIBLIOGRAFIA CITADA Y CONSULTADA.

Arbuthnott,J.(1710). An argument for divine providence, taken from the constante regularity observed in the births of both sexes, Philosophical Transactions, 27,186-190.

Bennett,B.M. (1965). On multivariate signed rank testes. Annals of the Institute of Statistical Mathematics, 17, 55-61.

Birnbaum,Z.W.y Tingey,F.H.(1951). One-sided confidence contours for probability distribution functions. The Annals of Mathematical Statistics, 22,592-596.

Blomqvist,N. (1951). Some tests basen on dichotomization. The Annals of Mathematical Statistics, 22,362-371.

Cochran,W.G. (1950). The comparison of percentages in matched samples. Biometrika, 37, 256-266.

__________ (1952). The X^2 test of godness of fit. The Annals of Mathematical Statistics, 23, 315-345.

__________ (1954). Some methods for strengthening the common X^2 tests. Biometrics, 10, 417-451.

Conover,W.J. (1971). Practical Nonparametric Statistics. John Wiley & Sons.Inc.N.Y.

Cox,D.R. (1958). Planning of Experiments. Wiley, New York.

Crâmer, H.(1946). Mathematical Methods of Statistics. Princeton Univ. Press, Princeton, N.J.

Chapman,D.G. y Meng,R.C.(1966). The power of chi-aquare test for contingency tables. Journal of the American - Statistical Association, 61, 965-975.

Dixon, W.J.(1953).Power functions of the sign test and power - efficiency for normal alternatives. The Annals of Mathematical Statistics, 24, 467-473.

Fleiss,J.L.(1965).A note on Cochran's Q test. Biometrics, 21, - 108-1010.

Friedman, M.(1937). The use of ranks to avoid the assumption of - normality implicit in the analysis of variance. Journal of the Amarican Statistical Association, 61; 1081-1096.

Gibbons,J.D.(1964). Effect of non-normality on the power function of the sign test. Journal of the American Statisti cal Association, 59, 142-148.

____________(1971). Nonparametric Statistical Inference. Mc.Graw-Hill Inc.

Good,I.J.Grover,T.N.y Mitchell,G.J.(1970), Exact distributions for X^2 and for the likenlihood-ratio statistic for - the equiprobable multinomial distribution. - Journal of the American Statistical Association, 65,267.

Goodman,L.A.(1970). The multivariate analysis of qualitative data: Interactions among multiple classifications. Journal of the American Statistical Association, 65, 226-256.

Haynam, G.E. and Leone,F.C.(1965). Analysis of categorical data. - Biometrika, 52, 654-660.

Hodges,J.L.Jr. y Lehmann,E.(1956). The efficiency of some non parametric competitors of the t-test. The - Annals of Mathematical Statistics, 27, 324-335.

Hoel,P. (1972). Introduction to Mathematical Statistics. - Micmillan Company.

Hollander,M. (1970). A distributions-free test for parallelism. - Journal of the American Statistical Association, 63, 707-714.

Hoog,R.V. y Craig A.T.(1970). Introduction to Mathematical Statistics. 3a.ed. Mcmillanan Company.

Hotelling, H y Pabst,M.R.(1936) Rank correlation and test of - significance involving no assumption of normality. The Annals of Mathematical Statistics, 7, 29-43.

Ireland,C.T., Ku,H.H. y Kullback,S.(1969). Symmetry, and marginal homogeneity of an rxr contingency table. Journal of the American Statistical Association, 64,- 1323-1341.

Ising,E.- (1952). Beitrang zur theorie des Ferromagnetismus, Zeits chrift fur Physik, 31, 253-258,

Jacobson,J.E. (1963). The Wilcoxon two-sample statistic; Tables and bibliography, Journal of the American Statistical Association, 58,1086-1103,

Kendall,M.G. (1938), A new measure of rank correlation, Biometrika, 30, 81-93,

Kendall,M.G. (1948), Rank Correlation Methods, Griffin, London

Kendall,M.G, y Babington-Smith,B. (1939), The problem of m rankings. the Annals of Mathematical Statistics, 10,275-287,

Kendall,M.G, y Stuart,A. (1973), The Aduanced theory of Statistics, Vol.II, 3a.ed. Hafner, N.Y.

King,Te Piao,et al (1976), Allergens of Honey Bee Jenom, Archives of Biochemistry and Biophysics 172,661-671,

Kolmogoroff,A.N, (1933), Sulla determinazione empirica di una legge di distribuzione, Giornale dell' Instituto Italiano degli Attuan, 4, 83-91.

Kruskal,W.H, (1952), A nonparametric test for the several sample problem, The Annals of Mathematical Statistics, 23, 525-540.

Kruskal,W.H, y Wallis,W.A, (1952), Use of ranks on one-criterion variance analysis, Journal of the American Statistical Association, 47, 583-621, Ku,H.H. (1963). A note on contingency tables involving zero fre quencies and the 2I test. Technometrics, 5, 398-400.

Lilliefors, H.W.(1967). On the Kolmogoroff-Smirnoff test for normality with mean and variance unknown. Journal of the American Statistical Associations, 64, 387--389.

Lindgren, B.W. (1968). Statistical Theory. Macmillan Publishing Co, Inc. N.Y.

Mann,H.B. y Whitney, D.R. (1947). On a test of whether one of two random variables is stochastically larger than - the other. The Annals of Mathematical Statistics, 18, 50-60.

Massey, F.J. (1952). Distribution table for the deviation between two sample cumulatives. The Annals of Mathematical Statistics, 23, 435-441.

Mc,Nemar,Q. (1955). Psychological Statistics, 2a.ed. John Wiley, New York.

Mc. Cornack,R.L. (1965). Extended tables of the Wilcoxon matched pairs signed rank statistics. Journal of the - American Statistical Association, 62, 954-965.

Méndez,I. (1976). Apuntes de Diseño de Experimentos IIMAS, UNAM.

Milton,R.C. (1964). An extended table of critical values for - The Mann-Whitney (Wilcoxon) Two-sample statistic. Journal of the American Statistical Association 59, 925-934.

Mood, A.M. (1940). The distributions theory of runs. The Annals of Mathematical Statistics, 11, 367-392.

Mood,A.M. y Graybill,F. (1963). Introduction to the Theory of Statistics (2a.ed). McGraw-Hill, New York.

Moses,L.E. (1952). Nonparametric Statistics for Psychological research. Phychological Bullelin, 49, 122-143.

Nelson.Ch.R. (1973). Applied time series analysis: for managerial forecasting. Holden-Day Inc. San Francisco.

Noether,G.E. (1967). Elements of Nonparametric Statistics. John Wiley, New York.

Ostle,B.(1965) Estadística Aplicada. Limusa-Wiley, México.

Owen,D.B. (1962). Handbook of Statistical Tables. Addison-Wesley, Reading, Mass.

Pearson,K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can reasonably be supposed to have arisen from random sampling. Philosophical Magazine (5),50, 157-175.

Van der Reyden,D. (1952). A simple Statistical significance test. Rhodesia Agricultural Journal, 49, 96-104.

Ross,S.M. (1972). Introduction to Probability Models. Academic Press. New York.

Siegel, S. (1976). Estadística no paramétrica aplicada a las ciencias de la conducta. Trillas, México.

Smirnoff, N.V. (1948). Table for estimating goodness of fit of empirical distributions. The Annals of Mathematical Statistics, 19, 279-281.

Sosa, B.P., Alagón, A.C., Possani, L.D. y Juliá, J.Z. (1979). Comparison of phospholipase activity with direct and indirect lytic effects of animal venoms upon human red cells. Sometido para publicación en Comparative Physiology and Biochemistry.

Spearman, C. (1904). The proof and measurement of association between two things. American Journal of Psychology, 15, 72-101.

Sugiura, N. y Otake, M. (1968). Numerical comparison of improved methods of testing in contingency tables with small frequencies. Annals of the Institute of Statistical Mathematics, 20, 505-517.

Swed, F.J. y Eisenhart, C. (1943). Tables for testing rand domness of grouping in a sequence of alternatives. The Annals of Mathematical Statistics, 14, 66-87.

Tate, M.W. y Brown, S.M. (1970). Note on the Cochran Q test. Journal of the American Statistical Association, 65, 155-160.

Valencia, G., Mendoza, M. y Aranda, F. (1978). Introducción a la Inferencia Estadística. Comunicación Interna Nº 42, Departamento de Matemáticas, Fac. Ciencias, UNAM. México.

Verdooren,L.R. (1963). Extended tables of critical valves for Wilcoxon's tost statistic, Biometrika, 50, 177-186.

Wald,A. y Wolfowitz,J. (1939). Confidence limits for continuous distribution functions. The Annals of - Mathematical Statistics, 10, 105-118.

________________ (1940). On a test whether two samples are from the same popilation. The Annals of - Mathematical Statistics, 11, 147-162.

Wallis,W.A. (1939). The correlation ratio for ranked data. Journal of the American Statistical Association, - 36, 401-409.

Walsh,J.E. (1951). Some bounded significance level properties of the equal-tail sign test. Annals of - Mathematical Statistics, 22, 408-417.

White,C. (1952). The use of ranks in a test of significance for companing two treatments. Biometrics,8,33- -41.

Wilcoxon,F. (1945). Individual comparisons by ranking methods. - Biometrics, 1, 80-83.

Wise,M.E. (1963). Multinomial probabilities and the X^2 y X^2 distributions. Biometrika, 50, 145.

________ (1964). A complete multinomial distribution compared with the X^2 approximation and an improvement to it. Biometrika, 51, 277.

Yule,G.V. y Kendall,M.G. (1950). An introducction to the theory of Statistics, 14°ed. Hafner, New York.

Zahn,D.A. and Roberts,G.C. (1971). Exact X^2 criterion tables with cell expectations one: an application to Coleman's measure of consensus. - Journal of American Statistical - Association, 66, 145.