Gobierno Bolivariano de Venezuela
Ministerio del Poder Popular para la Agricultura y Tierras
Instituto Nacional de Investigaciones Agrícolas

Pruebas no paramétricas para las ciencias agropecuarias

Muestras pequeñas

Hermann Wiedenhofer S.

PUBLICACIÓN TÉCNICA

El Instituto Nacional de Investigaciones Agrícolas es un instituto autónomo, creado de acuerdo a la Gaceta Oficial Nº 36.920 del 28 de marzo de 2000, adscrito al Ministerio de Agricultura y Tierras por decreto Nº 5.379 de Gaceta Oficial Nº 38.706 del 15 de Junio de 2007.

De acuerdo con el Reglamento de Publicaciones del Instituto Nacional de Investigaciones Agrícolas, aprobado por la Junta directiva en su sesión N° 126, según resolución N° 1456 de fecha 18 de febrero de 2010, esta es una Publicación Técnica.

Publicaciones Técnicas: contienen información proveniente de la evaluación de los resultados de investigación e innovación o la puesta en práctica de los mismos, presentados en forma descriptiva o de monografía. Son escritas por investigadores o técnicos y están destinadas fundamentalmente a investigadores, técnicos y estudiantes de educación técnica y superior. Incluye temas tales como: utilización de nuevas vacunas o la obtención y rendimientos de una nueva variedad; medidas sanitarias para la prevención de enfermedades; prácticas agropecuarias; manejo de medicamentos; pasos para tomar muestras, bien sea de suelos o de sangre, y estudios agroecológicos. Toman la forma de folletos. No tienen periodicidad.

Wiedenhofer S, H. 2013. Pruebas no paramétricas para las ciencias agropecuarias. Muestras pequeñas. 2 ed. rev y amp. Maracay, VE, Instituto Nacional de Investigaciones Agrícolas. 261 p.

REPÚBLICA BOLIVARIANA DE VENEZUELA
MINISTERIO DEL PODER POPULAR
PARA LA AGRICULTURA Y TIERRAS
INSTITUTO NACIONAL DE INVESTIGACIONES AGRÍCOLAS

Pruebas no paramétricas para las ciencias agropecuarias

Muestras pequeñas

Hermann Wiedenhofer S.*

* INIA. Centro Nacional de Investigaciones Agropecuarias. Maracay, Venezuela.

PUBLICACIÓN TÉCNICA

2 edición revisada y ampliada

Dirección: Edificio Sede Administrativa INIA. Av. Universidad, vía El Limón, Maracay, estado Aragua. República Bolivariana de Venezuela.
Oficina de Publicaciones No Periódicas (58) 243 2404770
Oficina de Distribución y Ventas de Publicaciones (58) 243 2404779
Zona Postal 2105
Página Web: http://www.inia.gob.ve

Equipo editorial Publicaciones No Periódicas
Gerente de Investigaciones e Innovación Tecnológica: Margaret Gutiérrez
Coordinación Área de Gestión de la Información: Graciela Piñero
Editor Jefe: Carlos Hidalgo
Editor Asistente: Ana Salazar
Editores: Andreina Muñoz, Elio Pérez
Diseño, diagramación y montaje: Raquel González

Para esta publicación
Editor responsable: Elio Asdrúbal Pérez Salcedo
Diseño gráfico: Raquel González

Impresión y encuadernación: Taller de Artes Gráficas del INIA.

Hecho el Depósito de Ley
Versión digital
Depósito Legal: lfi 22320136301518
ISBN 978-980-318-284-7

Dedicatoria

A mis tres amores

Amparo, Mary Carmen y Lily Ana

Agradecimiento

El autor desea expresar su agradecimiento al Ing. Agro. Elio A. Pérez S., Editor de Publicaciones Técnico-Divulgativas del INIA, por su valiosa colaboración y crítica constructiva en la edición de este libro, así como también a la señora Ana Bolívar, por su valiosa contribución en la elaboración de las tablas estadísticas que acompañan a esta publicación.

Contenido

Capítulo I

Introducción

En la metodología estadística tradicional se ha insistido constantemente acerca de una serie de condiciones y(o) supuestos necesarios para hacer valederas las conclusiones, a las cuales se ha llegado por los distintos métodos utilizados.

Particularmente, se debe recordar que al efectuar pruebas de hipótesis con el estadístico “Z”, se debía asumir normalidad de los datos y varianzas conocidas.

También hay que recordar que cuando la segunda de las condiciones anteriores no se podía sustentar, se tenía que utilizar la prueba “t de Student”.

Al analizar datos provenientes de un diseño en “Bloques al Azar”, había que asumir normalidad, independencia de los errores, homogeneidad de varianzas y aditividad.

Las condiciones y supuestos necesarios no sólo son del dominio de las pruebas de hipótesis, sino también de la estimación de intervalos de confianza y aseveraciones de probabilidad.

De lo anteriormente señalado, se desprende que para las estimaciones indicadas, se hace necesario el conocimiento o supuesto de la distribución de la población subyacente de los datos. La anterior metodología se denomina "metodología paramétrica". Paramétrica, ya que se hacen ciertas suposiciones o requerimientos acerca de los parámetros de las distribuciones.

Vale la pena indicar que cuando se hace el supuesto de normalidad de los datos o residuos, pero no hay seguridad de ello, se debe recurrir a las determinaciones de normalidad, ya sea usando paquetes estadísticos, sobre todo cuando los datos provienen de muestras grandes, u otro método cómo el indicado en el manual de Wiedenhofer (1980), cuando se trata de muestras pequeñas.

En el caso de homogeneidad de varianzas o de aditividad, el mejor recurso es el de los paquetes estadísticos, como SAS, SPSS, BMDP, entre otros.

Además de los anteriores supuestos (posiblemente algunos más), se tiene que tomar en cuenta la naturaleza de los datos y la forma de cómo se miden. Para las pruebas paramétricas se exige que los datos estén medidos en escala de intervalo o de razón (también denominada escala de proporción).

En el caso de los datos medidos en escala nominal (también llamada categórica) y los medidos en escala ordinal, sólo se puede (y se debe) usar los métodos no paramétricos, los cuales son el objeto de esta publicación.

Se ha referido a las escalas de medición, pero no se han definido.

Escalas de medición

Se tratará sobre las escalas de medición para los datos que proceden de distribuciones de naturaleza conocida, pero que se presentan de distintas maneras, las cuales dependen de la forma de medirlos.

Se debe hacer notar que una cosa es la naturaleza de los datos y otra es la forma de medirlos. La forma de cómo se miden los datos no depende exclusivamente de su naturaleza, sino que también de los instrumentos utilizados para medirlos, que por lo general tienen sus limitaciones.

Existen cuatro niveles de medida. El nivel más bajo de medida es el de la “escala nominal o categórica”, le siguen en orden ascendente la “escala ordinal”, la “escala de intervalo” y, finalmente, en el nivel superior, la “escala de razón o proporción”.

Escala nominal o categórica

La escala nominal utiliza números exclusivamente como medio para separar las propiedades o elementos en clases o categorías distintas.

El nombre o número que se asigna a un grupo de observaciones sirve solamente para darle un “nombre” (que puede ser un nombre, número o símbolo) a la categoría que los agrupa, de ahí el calificativo de “nominal”.

Ejemplo

- Si se observa el color de los ojos humanos y se le asigna el número uno al pardo, dos al negro, tres al verde y cuatro al azul.

- Si se observa un grupo de plantas y se le asigna el número uno a las afectadas por una plaga determinada y cero a las sanas.

- En un conjunto de bovinos se observan machos enteros, se le asigna el número uno; a los machos castrados, el número dos y a las hembras el tres.

Como se puede observar, la asignación de los números es perfectamente arbitraria y se pueden substituir por otros, siempre y cuando los grupos o categorías permanezcan intactos.

Escala ordinal

La escala ordinal se refiere a medidas en las cuales sólo tengan sentido los términos de comparación “mayor”, “menor” e “igual” entre las observaciones.

El valor numérico es utilizado exclusivamente para indicar orden, de ahí el calificativo de “ordinal”. Por lo general, se ordenan las observaciones de menor a mayor, pero también se puede hacer de mayor a menor.

En el caso de que dos o más observaciones tengan la misma medida, se dice que hay un “empate”. Los empates generalmente se producen por imperfecciones de medición y se pueden evitar o minimizar tomando las precauciones necesarias.

Ejemplo

- Se interroga a un grupo de personas sobre la opinión que tienen acerca del sabor de un producto alimenticio enlatado.

Se asigna el numero uno a la respuesta “no me gusta”, el dos a “me gusta poco”, el tres a “me gusta”, el cuatro a “me gusta mucho” y el cinco a “me gusta muchísimo”.

Se nota que perfectamente se hubiéra podido asignar los números en sentido inverso al orden, desde el uno a “me gusta muchísimo” hasta el cinco a “no me gusta”.

Otra manera válida pudiera ser: asignar el número cero a “no me gusta”, el cinco a “me gusta poco”, el 15 a “me gusta” , el 100 a “me gusta mucho” y el 1.500 a “me gusta muchísimo”, ya que hemos preservado el orden, que es lo importante. Es así que el número que se asigna es perfectamente arbitrario, siempre que se mantenga la relación de orden, ascendente o descendente.

- En una parcela experimental de maíz dada, se puede observar, pero no es posible medir la altura de las plantas. Se las clasifican al ojo, en “pequeñas”, “medianas” y “grandes”, asignándoles respectivamente los números uno, dos y tres.

- Se hace el examen microscópico a un exudado de cada una de las ovejas de un rebaño, con el objeto de determinar la abundancia de leucocitos. Se asigna a la “ausencia” de leucocitos el número cero, a “escasa cantidad” uno, a “abundante cantidad” dos, y tres a “muy abundante cantidad”.

Escala de intervalo

Cuando en una escala, además de la información acerca de la relación de orden, se tiene conocimiento acerca de la distancia entre dos observaciones, se dice que se trata de datos medidos en escala de intervalo.

Una escala de intervalo se caracteriza por una unidad de medida y un punto cero, ambos arbitrarios.

Ejemplo

- La medición de la temperatura con las escalas Celsius (ºC) y Fahrenheit (ºF) se hacen con un cero y escala, ambos arbitrarios, por lo tanto, se trata de una medición en escala de intervalo.

- La medida de la hora con un reloj, se hace en escala de intervalo.

- La medida del tiempo con un calendario se hace mediante una escala de intervalo.

Escala de razón o proporción

Si además de la información de orden y distancia, se tiene conocimiento acerca de la razón o proporción entre dos observaciones, se dice que se trata de una escala de razón o proporción.

La escala de razón tiene, como la escala de intervalo, unidades arbitrarias, pero a diferencia de ésta, el punto cero no es arbitrario, es natural.

Ejemplo

- La escala que se usa para medir la estatura de una persona es una escala de razón. La unidad de medida es arbitraria (centímetros, metros, pulgadas, entre otras). El punto cero de la escala es natural (cero centímetro o cero metro, entre otros).

- La escala que se utiliza para pesar los granos de una mazorca de maíz, es una escala de razón.

- El espacio de tiempo que pasa (medido con un cronómetro) entre la inoculación de una toxina a un ratón y su muerte, se mide con una escala de razón.

Métodos paramétricos y no paramétricos

Con el conocimiento de las escalas de medición, ahora se pueden indicar los requerimientos, ventajas y desventajas de cada uno de los métodos.

Requerimiento de los métodos

Los métodos paramétricos requieren de ciertos supuestos acerca de la distribución de la cual provienen las observaciones. Además, dichas observaciones deben ser el producto de una medición que por lo menos tenga el nivel de una escala de intervalo.

Para los métodos no paramétricos no hace falta especificar ni hacer supuestos acerca de la distribución de la cual provienen. Hay algunos supuestos, sin embargo, que se deben hacer en algunos casos, pero son pocos y débiles: observaciones independientes y distribución continua.

Las mediciones para las pruebas no paramétricas se hacen, generalmente, en escala nominal u ordinal.

Ventajas de los métodos no paramétricos

- Independientemente de la forma de la distribución subyacente, las aseveraciones de probabilidad de las pruebas no paramétricas son exactas.

- Si los tamaños de las muestras son tan pequeños como n = 6, sólo se pueden usar métodos no paramétricos, a menos que se conozca exactamente la distribución de la población subyacente.

- Hay métodos no paramétricos apropiados para observaciones provenientes de poblaciones distintas; no los hay para los métodos paramétricos.

- Los datos con nivel de medida en escala nominal y en escala ordinal sólo pueden ser analizados por métodos no paramétricos.

- Los métodos no paramétricos son más fácil de aplicar manualmente que los paramétricos.

Desventajas de los métodos no paramétricos

- Si todos los requerimientos para una prueba paramétrica son satisfechos y si la medición se hace en la escala requerida, los análisis por los métodos no paramétricos resultan ineficientes.

- Existen una serie de problemas que son susceptibles de ser analizados por métodos paramétricos y para los cuales la metodología no paramétrica no se ha desarrollado.

Respecto a lo último, Connover (1980) ha indicado una solución útil, especialmente para los análisis de datos provenientes de diseños experimentales, la cual se indica a continuación.

Si las observaciones no cumplen con alguno o más de los supuestos para el análisis paramétrico, se puede recurrir a lo siguiente:

Procedimiento

1) Analizar los datos utilizando el análisis paramétrico.

2) Convertir todos los datos en su conjunto a rangos, del menor al mayor y luego efectuar el análisis paramétrico con los rangos.

Si los resultados son bastante cercanos, se deben utilizar los resultados del método paramétrico. Si por otro lado, están bastante alejados entre si, se debe utilizar los resultados obtenidos con los rangos.

Pruebas de hipótesis

La palabra hipótesis se define como: 1. Una afirmación o negación que está sujeta a verificación o comprobación, 2. Una suposición que se utiliza como base para una acción.

El punto clave de estas definiciones está en que una hipótesis es una aseveración o suposición y no un hecho establecido. De esta manera, al no existir un conocimiento previo sobre el coeficiente de digestibilidad de maíz en novillos (N) y ovejos (O), un investigador puede proponer la hipótesis de que el coeficiente de digestibilidad de maíz en novillos (N) es superior que en los ovejos (O).

Un hacendado puede plantear la hipótesis de que cierta raza de vacas lecheras, tienen en el trópico una producción diaria promedio de 13 litros de leche.

Hipótesis de esta naturaleza pueden basarse en la experiencia, la observación, experimentación, o en la intuición. Las hipótesis establecidas en esta forma proporcionan con frecuencia motivo para realizar una investigación. Por esta razón, se pueden denominar hipótesis de investigación.

Generalmente, hay que volver a plantear las hipótesis de investigación antes de probarlas estadísticamente. Cuando ya se han planteado en forma conveniente, de tal forma que se puedan comprobar por medio de los métodos estadísticos, las hipótesis reciben el nombre de hipótesis estadísticas.

Las hipótesis estadísticas son aseveraciones sobre una o más poblaciones.

Las hipótesis estadísticas son de dos tipos. Primero está la hipótesis nula, que se simboliza por H_0. La hipótesis nula se llama también hipótesis de ninguna diferencia (por esto el término nula). Es una afirmación en la que se dice que no hay ninguna diferencia entre dos poblaciones, o entre el valor verdadero de algún parámetro y su valor hipotético.

Para el caso de las dos hipótesis de investigación que se acabaron de enunciar, se establecerán para cada una de ellas la hipótesis nula correspondiente.

En el caso de la hipótesis de investigación de un experimento, en el cual se desea determinar si la digestibilidad de ensilado de maíz en novillos (N) es superior a la de los ovejos (O), la hipótesis nula apropiada (Ho) consiste en afirmar que no hay diferencia entre las medianas (Md) de los coeficientes de digestibilidad, de ambas especies.

Se puede expresar la hipótesis nula en forma más compacta como:

$$\mathbf{H_0: Md_N = Md_O}$$

En el caso de la hipótesis de investigación que afirma que la producción diaria de leche es de 13 litros,

$$\mathbf{H_0: Md = 13}$$

Para probar una hipótesis nula, se examinan los datos de la muestra tomada de la población pertinente y se determina si son o no compatibles con ella.

Si los datos de la muestra no son compatibles con la hipótesis nula (H_0), entonces se rechazará.

Si los datos son compatibles con la hipótesis nula (H_0), entonces no se rechazará.

Más adelante, se explicará el criterio que se usa para determinar si los datos de la muestra son o no compatibles con la hipótesis nula.

Si la hipótesis nula no se ha rechazado, se dice que los datos particulares de la muestra no dan suficiente evidencia como para que se concluya que es falsa.

Si la hipótesis nula se ha rechazado, se dice que los datos particulares de la muestra sí dan suficiente evidencia como para concluir que la hipótesis nula es falsa y que una segunda hipótesis es verdadera. Esta segunda hipótesis, de la que se ha concluido que es verdadera, si la hipótesis nula es rechazada, se denomina hipótesis alterna y se designa con el símbolo H_1, Generalmente, la hipótesis alterna corresponde a la hipótesis de investigación, salvo una excepción.

En el caso de pruebas de una cola (unimarginales o unilaterales), la hipótesis de investigación corresponde a la H_1.

En pruebas de dos colas (bimarginales o bilaterales), si la hipótesis de investigación se expresa en términos de desigualdad, corresponderá a la H_1 y si se expresa en términos de igualdad, corresponderá a la H_0 (la excepción).

Retomando las hipótesis de investigación planteadas anteriormente, se establecerá, en cada caso, cual sería la hipótesis nula y cual la hipótesis alterna apropiada.

1. Hipótesis de investigación: el coeficiente de digestibilidad en novillos es superior al de los ovejos.

H_0: $Md_N = Md_O$ (Las dos medianas poblacionales son iguales)
H_1: $Md_N > Md_O$ (La mediana de digestibilidad en novillos es superior a la de los ovejos)

2. Hipótesis de investigación: la producción diaria de leche es igual a 13 litros.

H_0: $Md = 13$ (La mediana poblacional es igual a 13)
H_0: $Md \neq 13$ (La mediana poblacional es diferente de 13)

Se puede observar que en el primer caso, la hipótesis alterna corresponde a la hipótesis de investigación, mientras que en el segundo caso, la hipótesis nula corresponde a la hipótesis de investigación.

Cuando se establecen hipótesis del tipo indicado en (1), se procura, generalmente, que las hipótesis nula y alterna se complementen entre si y para esto se incluye una desigualdad en la hipótesis nula que vaya en dirección opuesta a la de la hipótesis alterna. Por ejemplo, se podría escribir la hipótesis anterior (1):

$$\mathbf{H_0: Md_N \leq Md_O}$$
$$\mathbf{H_1: Md_N > Md_O}$$

Este modo de plantear las hipótesis nula y alterna realza el hecho de que cuando la hipótesis alterna establece una desviación respecto de una igualdad en una dirección, las desviaciones respecto de la igualdad en la dirección opuesta no tienen ningún interés.

Procedimiento de pruebas de hipótesis

Como ilustración de los procedimientos para probar hipótesis, se examina el ejemplo siguiente:

Se desea determinar si el contenido proteico de la leche pasteurizada de la marca “A” es superior al de la leche “B”. El método que se ha utilizado para analizar el contenido permite sólo dar puntuaciones a cada leche. Las puntuaciones van de 100 a 400.

Se toma una muestra al azar de 11 cartones de la marca "A" y una muestra aleatoria de ocho cartones de la marca "B".

De acuerdo con esto, se establece la siguiente hipótesis nula y la siguiente hipótesis alterna:

$$\mathbf{H_0: Md_A \leq Md_B}$$
$$\mathbf{H_1: Md_A > Md_B}$$

Donde, $\mathbf{Md_A}$ es el contenido proteico mediano poblacional de la leche de la marca "A" y $\mathbf{Md_B}$ el contenido proteico mediano poblacional de la leche de la marca "B".

Las poblaciones sobre las que se desea hacer inferencias, son las poblaciones de todas las leches que se pueden caracterizar como de la marca "A" o de la leche de marca "B".

Se obtienen muestras independientes de leches de la marca "A" y de la leche de marca "B", que se pueden tratar como muestras aleatorias de las poblaciones de interés.

Se analizan las muestras de ambas leches y se calculan las medianas del contenido proteico de ellas y se obtiene lo siguiente:

Se descubre que la mediana proteica de la leche "A" es igual a 340 y de la leche "B" es de 310. Aunque la dirección de la diferencia de las medianas muéstrales es compatible con su hipótesis de investigación (y la alterna), se sabe que existen por lo menos dos maneras de explicar esta diferencia:

- El puntaje verdadero mediano proteico de la población de la leche "A" podría no ser superior al que corresponde a la población de la leche "B". Los resultados observados en la muestra se deben simplemente a la casualidad.

- Los resultados observados en la muestra pudieran reflejar el verdadero estado de las cosas y es acertado sacar como conclusión que el contenido proteico mediano verdadero de la leche “A” es superior al de la leche “B”.

El conocimiento y la comprensión de las sutiles ideas de los procedimientos de prueba de hipótesis permitirán que se pueda escoger entre las dos explicaciones.

Se va a dedicar el resto de esta sección a los conceptos y técnicas específicas que se utilizan en la prueba de hipótesis.

Se puede formalizar el procedimiento que se debe seguir para verificar una hipótesis, estableciendo en forma secuencial los diversos pasos que forman el procedimiento. En esta sección, se enumeran y explican cada uno de estos pasos, llevando el mismo orden que guardan normalmente en la práctica. Se pueden identificar siete pasos principales.

1) Planteamiento de las hipótesis:
   a) Hipótesis de Investigación.
   b) Hipótesis estadística.

2) Estadístico de prueba.

3) Nivel de significación.

4) Región crítica o Área de rechazo.

5) Cálculos.

6) Decisión.

7) Conclusión.

A continuación, se describen cada uno de estos pasos en términos generales y, posteriormente, se explicaran con ejemplos específicos.

Planteamiento de las hipótesis

Generalmente, se quiere obtener una conclusión (paso 7) rechazando la hipótesis nula. Es decir, ordinariamente se prefiere que los datos de la muestra apoyen la hipótesis alterna. En consecuencia, al determinar lo que debe ser la hipótesis alterna, se debe preguntar "¿qué se desea concluir?" o "¿qué se cree que es verdadero?". La respuesta a estas preguntas constituye la expresión de la hipótesis alterna. Por lo tanto, el planteamiento complementario de la hipótesis alterna, sirve de hipótesis nula.

En el ejemplo anterior se desea establecer el hecho de que, el contenido proteico de la leche "A" es superior al de la leche "B".

Frente a la pregunta "¿qué se desea concluir?", se responderá que se desea sacar la conclusión de que el contenido proteico de la leche "A" es superior al de la leche "B". Por lo tanto, la hipótesis alterna consiste en $Md_A > Md_B$ y la hipótesis nula, que es el complemento de este planteamiento, en $Md_A \leq Md_B$. Este ejemplo, muestra cómo, generalmente, se formula primero la hipótesis alterna, salvo la excepción anteriormente indicada.

Nivel de significación

Teniendo en cuenta los resultados que se obtienen en el análisis de los datos de la muestra, **se rechaza o no la hipótesis nula**.

Rechazar la hipótesis nula no constituye una prueba de que sea falsa. Sin tener en cuenta que tan incompatible sea la evidencia de la muestra, con la hipótesis nula, cabe la posibilidad de que ésta sea realmente verdadera. Análogamente, el hecho de no rechazar la hipótesis nula no es una prueba de que sea verdadera y de que la hipótesis alterna sea falsa.

De la misma manera que en el caso anterior, aunque la hipótesis nula no sea rechazada, cabe la posibilidad de que sea falsa. La consideración de estos hechos lleva a la conclusión de que en el rechazo o el no rechazo de la hipótesis nula, se corre el riesgo de equivocación.

Aunque, generalmente, no se sabe si en una determinada acción (rechazo o no rechazo de H_0) se comete un error o no, se pueden indicar los dos tipos de error posibles, de la manera siguiente:

- Rechazo de una hipótesis nula verdadera. Este error se denomina error de Tipo I.

- Aceptación de una hipótesis nula falsa. Este error se denomina error de Tipo II.

Se puede ilustrar la relación entre la certeza de la hipótesis nula (es decir, si es verdadera o es falsa) y la decisión estadística (rechazar o no rechazar H_0) como se ve en el Cuadro 1.1.

Cuadro 1.1. Relación entre la certeza de la hipótesis nula y la decisión estadística.

Decisión estadística	**Certeza de la H_0**	
	Verdadera	**Falsa**
Rechazo de la H_0	Error tipo I	Decisión correcta
No rechazo de la H_0	Decisión correcta	Error Tipo II

Siguiendo la costumbre que se tiene en estadística, se representará con α (alfa) la probabilidad de cometer un error de tipo I y con β (beta) la probabilidad de cometer un error de Tipo II. Así pues α = P(Error de tipo I) y β = P(Error de Tipo II).

Para la verificación de una hipótesis determinada se preferiría que α y β sean pequeños. En vista de la relación entre éstas dos probabilidades, se encuentra que, para un tamaño de muestra dado, una disminución de α tiene como contraparte un aumento de β y viceversa.

Siendo esto así, parece prudente que, en una situación determinada, se trate de minimizar la probabilidad de cometer el error más serio. Desafortunadamente, en muchas áreas de investigación es difícil o imposible evaluar los dos tipos de error, en cuanto a la seriedad de cada uno de ellos. Entonces, lo que se hace en estas situaciones es seleccionar algún valor pequeño para α, como 0,10; 0,05; 0,01 o 0,001.

La elección de α refleja la opinión que se tiene sobre la seriedad del error de Tipo I. Mientras más serias se consideren las consecuencias de cometer un error de Tipo I, menor será el valor que se debe asignar a α.

Con frecuencia, α se denomina nivel de significación. Cuando se escoge un nivel de significación igual a α y se rechaza la hipótesis nula, se dice que los resultados de la muestra son significativos.

Asignación de los niveles de significación

Suponiendo que hay interés en determinar si un suplemento alimenticio en particular es efectivo para aumentar instantáneamente la producción de leche en vacas lecheras. El suplemento es algo costoso.

Se tiene un conjunto de vacas lecheras de las cuales se conoce su record de producción. Se suministra el suplemente y luego se determina si la producción mediana de leche ha aumentado.

Se tiene:

H_0: $Md_D \leq 0$
H_1: $Md_D > 0$
(Md_D= mediana de las diferencias)

Posibles conclusiones

- **Si se rechaza H_0**, se concluye que el suplemento alimenticio es efectivo para aumentar la producción de leche en vacas lecheras.

- **Si no se rechaza H_0**, se concluye que el suplemento alimenticio no es efectivo para aumentar la producción de leche en vacas lecheras.

Posibles acciones

- **Si se rechaza H_0**, se concluye que el suplemento alimenticio es efectivo para aumentar la producción de leche en vacas lecheras, entonces, se utilizará por considerarlo útil.

- **Si no se rechaza H_0**, se concluye que el suplemento alimenticio no es efectivo para aumentar la producción de leche en vacas lecheras, entonces no se utilizará por considerarlo inútil.

Posibles errores

- **Si se rechaza H_0** cuando es verdadera, se cometerá un error de tipo I. En el caso específico del presente problema, se concluye que el suplemento alimenticio es efectivo y se recomendará su uso para aumentar la producción de leche en vacas lecheras, cuando en realidad el suplemento alimenticio no es efectivo.

- **Si no se rechaza H_0** cuando es falsa, se cometerá un error de tipo II. En el caso específico del presente problema se concluye que el suplemento alimenticio no es efectivo para aumentar la producción de leche en vacas lecheras. y se recomendará descartarlo, cuando en realidad el suplemento alimenticio es efectivo.

Posibles consecuencias de los errores

- Si se recomienda el uso del suplemento alimenticio, cuando en realidad es inútil, se incurrirá en una serie de gastos innecesarios, que a la larga pudieran perjudicar a la industria lechera.

- Si se descarta el uso del suplemento alimenticio, cuando en realidad es útil, se priva a la industria del beneficio de elevar la producción de leche.

Comparación de la gravedad de los errores

Tres casos:

- El error de tipo I es el más grave.

- El error de tipo II es el más grave.

- Ambos errores tienen la misma gravedad.

Asignación de las probabilidades de cometer los errores

Para efectos de este ejemplo.

- Si se considera que **el error de tipo I es el más grave**, se debe minimizar la probabilidad de cometerlo. Esto quiere decir, que si se considera que es más grave recomendar el uso del suplemento alimenticio cuando es inútil, que descartarlo cuando es útil, se debe usar un nivel de significación de 0,01 (**α = 0,01**).

- Si se considera que **el error de tipo II es el más grave**, se debe minimizar la probabilidad de cometerlo. Esto quiere decir, que si se considera que es más grave descartar el suplemento alimenticio cuando es útil, que recomendarlo cuando es inútil; se debe usar un valor pequeño de β. Sin embargo, la probabilidad de cometer el error II no se puede minimizar directamente, ya que no es sencillo. Es por esto que se considera que para un tamaño de muestra determinado, un aumento en α traerá como consecuencia una disminución de β. Por lo tanto, para minimizar β se procederá a maximizar α. Entonces, se debe usar, en este caso, un nivel de significación de 0,10 (**α = 0,10**).

- Si se considera que **ambos errores tienen la misma gravedad**, se usará un nivel de significación de 0,05 (**α = 0,05**).

En este ejemplo en particular, se considerará que el error Tipo II es el más grave, es decir, que es más grave no usar el suplemento alimenticio cuando es útil que recomendarlo cuando es inútil, por lo tanto, se usará $\alpha = 0,10$.

Nota: para los efectos de este ejemplo sólo se ha usado los niveles de significación siguientes:

$\alpha = 0,01$
$\alpha = 0,05$
$\alpha = 0,10$

Es por esto que el valor mínimo de α se ha establecido como 0,01; el máximo 0,10 y 0,05 el valor intermedio.

Estadístico de prueba

Un estadístico de prueba es una función de los datos de una muestra que se utiliza para tomar la decisión de rechazar o no rechazar una hipótesis nula.

El estadístico de prueba se determina teniendo en cuenta el parámetro sobre el que se hace la hipótesis y la naturaleza de la distribución muestral del estadístico pertinente.

Especificación de las regiones de rechazo y de no rechazo

En la prueba de una hipótesis, la región de rechazo consta de todos aquellos valores del estadístico de prueba que son de tal magnitud que, de ser el valor observado del estadístico de prueba igual a uno de ellos, la hipótesis nula se rechaza. La región de no rechazo es el complemento de la región de rechazo.

Si el valor observado del estadístico de prueba es igual a alguno de los valores que componen la región de no rechazo, la hipótesis nula no se rechaza.

Tal cómo se va a ver, los tamaños de las regiones de rechazo y de no rechazo están determinados por α.

Se llaman valores críticos de un estadístico de prueba a todos aquellos valores que separan una región de rechazo de una región de no rechazo. Ellos indican cuando se debe dejar de creer que la hipótesis nula es verdadera y empezar a creer que es falsa.

Se llama hipótesis alternas de dos lados, bilateral o de dos colas a las hipótesis alternas de la forma H_1: Md ≠ Mdo, ya que generalmente conducen a una región de rechazo que está compuesta de dos lados o colas, de la distribución del estadístico de prueba.

Al procedimiento adecuado para probar una hipótesis nula frente a una hipótesis alterna bilateral, se le da el nombre de prueba de hipótesis de dos lados, bilateral o de dos colas.

Con frecuencia, la hipótesis nula es de la forma H_0: Md ≤ Mdo y la hipótesis alterna de la forma: H_1: Md > Mdo. A una hipótesis alterna de este tipo se llama hipótesis unilateral o de una cola, puesto que sólo valores positivos del estadístico de prueba causan el rechazo de la hipótesis nula y, por lo tanto, la región de rechazo está localizada sólo en la cola derecha de la distribución del estadístico de prueba.

Es decir, que toda la probabilidad α está localizada en una sola cola y no está dividida por la mitad, como sucede en la prueba bilateral.

Si se selecciona un nivel de significación (probabilidad de rechazar una hipótesis nula verdadera) de 0,05, todo el valor 0,05 constituirá el área de la cola derecha en la distribución muestral.

Para las hipótesis alternas de la forma H_1: Md < Mdo, sólo los valores negativos del estadístico de prueba causan el rechazo de la hipótesis nula y, por lo tanto, toda la región de rechazo se encontrará en la cola izquierda de la distribución.

Cálculos

Los datos que se necesitan para verificar las hipótesis formuladas y que satisfacen las suposiciones necesarias de la prueba se deben recolectar en forma adecuada. Una vez que se han recogido, se calcula el estadístico apropiado y el estadístico de prueba.

Decisión

Se compara el valor calculado del estadístico de prueba con el valor crítico de éste. Si el valor calculado está en la región de rechazo, entonces se rechazará H_0, de lo contrario, no se rechazará.

Conclusión

En tanto que la decisión se expresa en función del estadístico de prueba, la conclusión se expresa en función de la hipótesis de investigación y, por lo tanto, en términos de la variable en el planteamiento del problema.

Por ejemplo, cuando se rechaza H_0: Md = Mdo, se concluye que "la mediana poblacional no es igual a Mdo". Cuando no se rechaza la hipótesis nula la conclusión carece de la fuerza de convicción que tiene cuando se rechaza una hipótesis nula. Esto se debe a que, aunque de antemano se sabe que la probabilidad de rechazar una hipótesis nula verdadera es pequeña (esto se sabe por la selección que se ha hecho de α), generalmente no se conoce el valor de β, o probabilidad de aceptar (no rechazar) una hipótesis nula falsa. Esta puede ser muy grande y frecuentemente lo es.

En consecuencia, al no rechazar H_0: Md = Mdo, se concluye que la mediana de la población "puede ser igual" a Mdo".

En todos los casos la hipótesis de investigación va a corresponder a la hipótesis alterna, excepto en las hipótesis de dos colas de igualdad, dónde corresponderá a la nula. De este modo se facilita la afirmación o negación de la hipótesis de investigación y, por lo tanto, la conclusión al planteamiento del problema.

Se ilustrará lo anterior con cuatro versiones del mismo planteamiento del problema.

Suponga que se desea efectuar un experimento para determinar si el contenido proteico de dos leches pasteurizadas (A y B), son iguales, desiguales, si A es superior a B o si A es inferior a B.

Planteamiento del problema (1): se desea determinar si las leches de dos marcas ("A" y "B") en su contenido medio de proteína son **iguales**.

Hipótesis de investigación (1): las leches son iguales respecto a su contenido proteico.

Hipótesis estadísticas (1): $H_0: Md_A = Md_B$
$H_1: Md_A \neq Md_B$

Conclusión: para llegar a la conclusión, se debe tomar en cuenta a que hipótesis estadística corresponde la hipótesis de investigación y si se ha rechazado o no la hipótesis nula.

Caso 1: se ha rechazado la hipótesis nula $\mathbf{H_0: Md_A = Md_B}$, es decir, que se ha negado, y como corresponde a la hipótesis de investigación, a su vez se negará ésta, por lo tanto, se concluye que las leches no son iguales respecto a su contenido proteico.

Caso 2: no se ha rechazado la hipótesis nula $\mathbf{H_0: Md_A = Md_B}$, es decir, que no se ha negado, y como corresponde a la hipótesis de investigación, a su vez se afirmará, por lo tanto, se concluye que las leches **son guales**, respecto a su contenido proteico.

Planteamiento del problema (2): se desea determinar si las leches de dos marcas ("A" y "B") en su contenido medio de proteína son **desiguales**.

Hipótesis de investigación (2): las leches son desiguales respecto a su contenido proteico.

Hipótesis estadísticas (2): $H_0: Md_A = Md_B$
$H_1: Md_A \neq Md_B$

Conclusión: para llegar a la conclusión, se debe tomar en cuenta a que hipótesis estadística corresponde la hipótesis de investigación y si se ha rechazado o no la hipótesis nula.

Caso 1: se ha rechazado la hipótesis nula $\mathbf{H_0: Md_A = Md_B}$, es decir, que se ha negado, por lo tanto, "se acepta" la hipótesis alterna $H_1: Md_A \neq Md_B$ que corresponde a la hipótesis de investigación, a su vez se afirmará ésta, por lo tanto, se concluye que las leches **son desiguales**, respecto a su contenido proteico.

Caso 2: no se ha rechazado la hipótesis nula $\mathbf{H_0: Md_A = Md_B}$, es decir, que no se ha negado, por lo tanto, "no se acepta" la hipótesis alterna $\mathbf{H_1: Md_A \neq Md_B}$ que corresponde a la hipótesis de investigación, la cual se negará y se concluye que las leches **no son desiguales**, respecto a su contenido proteico.

Planteamiento del problema (3): se desea determinar si la leche de la marca "A" es **superior** al de la leche "B" en su contenido medio de proteína.

Hipótesis de investigación (3): el contenido proteico de la leche "A" es superior al de la leche "B".

Hipótesis estadísticas (3): H_0: $Md_A \leq Md_B$
H_1: $Md_A > Md_B$

Conclusión: para llegar a la conclusión, se debe tomar en cuenta a que hipótesis estadística corresponde la hipótesis de investigación y si se ha rechazado o no la hipótesis nula.

Caso 1: se ha rechazado la hipótesis nula **H_0: $Md_A \leq Md_B$**, es decir, que se ha negado, por lo tanto, "se acepta" la hipótesis alterna **H_1: $Md_A > Md_B$** que corresponde a la hipótesis de investigación, a su vez se afirmará ésta, por lo tanto, se concluye que el contenido proteico de la leche "A" **es superior** al de la leche "B".

Caso 2: no se ha rechazado la hipótesis nula **H_0: $Md_A \leq Md_B$**, es decir, que no se ha negado, por lo tanto, "no se acepta" la hipótesis alterna **H_1: $Md_A > Md_B$** que corresponde a la hipótesis de investigación, que se negará y se concluye que el contenido proteico de la leche "A" **no es superior** al de la leche "B".

Planteamiento del problema (4): se desea determinar si la leche de la marca "A" es **inferior** al de la leche "B" en su contenido medio de proteína.

Hipótesis de investigación (4): el contenido proteico de la leche "A" es inferior al de la leche "B".

Hipótesis estadísticas (4): H_0: $Md_A \geq Md_B$
H_1: $Md_A < Md_B$

Conclusión: para llegar a la conclusión, se debe tomar en cuenta a que hipótesis estadística corresponde la hipótesis de investigación y si se ha rechazado o no la hipótesis nula.

Caso 1: se ha rechazado la hipótesis nula $\mathbf{H_0: Md_A \geq Md_B}$, es decir, que se ha negado, por lo tanto, "se acepta" la hipótesis alterna $\mathbf{H_1: Md_A < Md_B}$ que corresponde a la hipótesis de investigación y se concluye que el contenido proteico de la leche "A" **es inferior** al de la leche "B".

Caso 2: no se ha rechazado la hipótesis nula $\mathbf{H_0: Md_A \geq Md_B}$, es decir, que no se ha negado, por lo tanto, "no se acepta" la hipótesis alterna $\mathbf{H_1: Md_A < Md_B}$ que corresponde a la hipótesis de investigación, a su vez se negará ésta y se concluye que el contenido proteico de la leche "A" **no es inferior** al de la leche "B".

Capítulo II

Caso de una muestra

Se tratará ahora el caso en el cual se desean efectuar pruebas de hipótesis que requieran una muestra.

Típicamente, se desea probar que una muestra aleatoria fue tomada de una población con una distribución determinada. Estas pruebas se denominan "pruebas de bondad de ajuste".

Específicamente, se puede hacer las preguntas siguientes:

- ¿Existe diferencia significativa en lo referente a la tendencia central (medias), entre la muestra y la población?
- ¿Hay diferencia significativa entre la frecuencia muestral y la poblacional?
- ¿Es razonable pensar que la muestra proviene de una población dada?
- ¿Se puede aseverar que la muestra es una muestra aleatoria de alguna población?

Para responder a estas preguntas se presentaran cuatro pruebas: Prueba Binomial; Prueba de Ji al cuadrado; Prueba de Kolmogorov-Smirnov para una muestra y Prueba de Rachas para una muestra.

Prueba binomial

Fundamentos

Hay poblaciones formadas por variables que sólo tienen dos valores, las cuales son denominadas binarias o dicotómicas.

Las observaciones tomadas de dichas poblaciones pertenecen a una de las dos posibles clases, pero no a ambas.

Estas observaciones se miden en la escala nominal, pero puede darse el caso de mediciones en escalas superiores, que han sido dicotomizadas.

Como ejemplos de estas poblaciones se tienen las que poseen dos clases: machos y hembras; sanos y enfermos; adecuados y no adecuados; claros y oscuros, entre otras.

En estos tipos de poblaciones, la proporción de observaciones que pertenecen a una de las dos clases se indican con la letra "P" y la proporción de observaciones que pertenecen a la otra clase se indican con la letra "Q", donde $Q = 1 – P$ y $P = 1 – Q$.

El valor de P (y por ende el de Q) puede variar de una población a otra, pero es constante para una población en particular.

En general, se asumirá que la proporción observada (p) en muestras aleatorias tomadas del tipo de poblaciones, se distribuyen de acuerdo a la distribución binomial.

Metodología

Una situación común es en la que se desea probar si P = Q (P = Q = ½).

La Tabla 1 tiene los valores críticos para pruebas de una y dos colas (unilateral y bilateral, respectivamente).

a) Pruebas de dos colas

Sea n el número de observaciones en la muestra (tamaño de la muestra) y x el número de observaciones que pertenecen a la clase de interés; en la Tabla 1 se encuentran los valores críticos con los cuales se puede rechazar la hipótesis nula de igualdad de proporciones:

$$H_0: P = Q$$

a favor de la alterna de desigualdad:

$$H_1: P \neq Q$$

En pruebas de dos colas se utiliza x o (n - x) (el que sea mayor), y se rechaza H_0 si su valor numérico es mayor que el de la x o (n - x) tabulada, para dos colas.

b) Pruebas de una cola

Cola derecha

Si ambos, x > n/2 y si x es mayor o igual que el valor tabulado para una cola, se rechaza la hipótesis nula:

$$H_0: P \leq Q$$

a favor de la alterna:

$$H_1: P > Q$$

Cola Izquierda

Si ambos, n - x > n/2 y si n - x es mayor que el valor tabulado para una cola, se rechaza la hipótesis nula:

$$H_0: P \geq Q$$

a favor de la alterna:

$$H_1: P < Q$$

Ejemplo

- Se desea conocer si en una población dada de animales salvajes, la relación entre la proporción de machos y hembras es de 1:1, es decir, si P = Q. Se asume la distribución binomial de los sexos.

Se ha tomado una muestra aleatoria de 24 animales (n = 24). En el Cuadro 2.1 se muestran los resultados.

a) Hipótesis de investigación: las proporciones de machos y hembras en la población dada de animales salvajes son iguales.

b) Hipótesis estadísticas: H_0: P = Q
H_1: P ≠ Q

c) Estadístico de prueba: se utiliza la Prueba Binomial, ya que los datos están en dos clases o categorías (machos y hembras).

Se asume que las observaciones son independientes (el sexo de un animal no depende del sexo del otro) y que las probabilidades son constantes, es decir, que la probabilidad de que un animal sea macho o hembra es constante para todos los animales (antes de nacer, por supuesto).

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0,05$).

e) Región crítica o Área de rechazo: se rechaza la H_0 si el valor de x o (n - x), (el que sea el mayor) es mayor que el tabulado en la Tabla 1, para dos colas, $\alpha = 0,05$ y n = 24.

Se rechaza la H_0 si x es mayor que 18.

f) Cálculos:

Cuadro 2.1. Distribución sexual de una muestra de 24 animales salvajes.

Sexo	Machos	Hembras	Total
Frecuencia	15	9	24

Se tomará como clase de interés la de los machos, cuya frecuencia (x) es de 15, mayor que la de las hembras (n - x) que es 9.

Se tiene que n = 24 y x = 15.

g) Decisión: como x = 15 no es mayor que 18, el valor crítico, no se rechaza H_0.

h) Conclusión: la proporción de machos y hembras en la población dada de animales salvajes son iguales (no son distintas).

Ejemplo

- Se desea conocer si en una población dada de animales salvajes, la proporción de machos es superior a la de las hembras. Se ha tomado una muestra aleatoria de 24 animales (n = 24). En el Cuadro 2.2 se muestran los resultados.

a) Hipótesis de investigación: la proporción de machos es superior a la de las hembras, en la población dada de animales salvajes.

b) Hipótesis estadísticas: $H_0: P \leq Q$
$H_1: P > Q$

c) Estadístico de prueba: Se utiliza la Prueba Binomial, ya que los datos están en dos clases o categorías (machos y hembras).

Se asume que las observaciones son independientes (el sexo de un animal no depende del sexo del otro) y que las probabilidades son constantes, es decir que la probabilidad de que un animal sea macho o hembra es constante para todos los animales (antes de nacer, por supuesto). Prueba de una cola (derecha).

d) Nivel de significación: se utilizará un nivel de significación de 1% (α = 0,01).

e) Región crítica o Área de rechazo: la región crítica consta de todos los valores de x que sean mayores que el tabulado en la Tabla 1, para una cola, $\alpha = 0{,}01$ y $n = 24$.

Se rechaza la H_0 si $x > n/2$ y mayor que 19.

f) Cálculos:

Cuadro 2.2. Distribución sexual de una muestra de 24 animales salvajes.

Sexo	Machos	Hembras	Total
Frecuencia	15	9	24

La clase de interés es la de los machos, por lo tanto, x = 15

g) Decisión: se tiene que n = 24 y x = 15 > 12 = n/2, pero 15 no es mayor que 19, el valor crítico, por lo tanto, no se rechaza H_0.

h) Conclusión: la proporción de machos no es superior a la de las hembras en la población dada de animales salvajes.

Ejemplo

- Se desea conocer si en una población dada de animales salvajes, la proporción de hembras es inferior a la de los machos. Se ha tomado una muestra aleatoria de 24 animales (n = 24). En el Cuadro 2.3 se muestran los resultados.

a) Hipótesis de investigación: la proporción de hembras es inferior a la de los machos en la población dada de animales salvajes.

b) Hipótesis estadísticas: H_0: P ≥ Q
H_1: P < Q

c) Estadístico de prueba: se utiliza la Prueba Binomial, ya que los datos están en dos clases o categorías (machos y hembras).

Se asume que las observaciones son independientes (el sexo de un animal no depende del sexo del otro) y que las probabilidades son constantes, es decir, que la probabilidad de que un animal sea macho o hembra es constante para todos los animales (antes de nacer, por supuesto). Prueba de una cola (izquierda).

d) Nivel de significación: se utilizará un nivel de significación de 5% (α = 0,05).

e) Región crítica o Área de rechazo: la región crítica consta de todos los valores de n - x que sean mayores que el tabulado en la Tabla 1, para una cola (izquierda), α = 0,05 y n = 24.

Se rechaza la H_0 si x < n/2 y n - x es mayor que 17.

f) Cálculos:

Cuadro 2.3. Distribución sexual de una muestra de 24 animales salvajes.

Sexo	Hembras	Machos	Total
Frecuencia	6	18	24

Se tomará ahora como clase de interés la de las hembras, cuya frecuencia (x) es de 6 menor que 12 = n/2.

g) Decisión: se tiene que n = 24 y x= 6, menor que 12 = n/2, y n – x = 18, mayor que 17, el valor crítico, por lo tanto, se rechaza H_0.

h) Conclusión: la proporción de hembras es inferior a la de los machos en la población dada de animales salvajes.

Prueba de Ji al cuadrado para una muestra

Fundamentos

En la prueba binomial se indica que ésta es adecuada para analizar datos que pueden ser clasificados en dos o más categorías o clases.

La prueba de Ji al cuadrado se utiliza para analizar datos que se puedan clasificar en dos o más categorías o clases, especialmente, cuando se desean determinar si las frecuencias observadas en cada clase o categoría concuerdan con las frecuencias teóricas que se especifican. Esto último es lo que se denomina "Prueba de bondad de ajuste".

Las "Pruebas de bondad de ajuste" se utilizan para detectar si existe diferencia significativa entre la frecuencia observada en cada categoría y la correspondiente frecuencia esperada (teórica especificada).

La frecuencia esperada se basa en la hipótesis nula de la prueba.

Metodología

Como primer paso se debe indicar, basándose en la hipótesis nula, cuales son las frecuencias esperadas.

Debido a que la hipótesis nula puede presentarse de diversas maneras, se debe indicar, previamente, la forma de convertirla a una hipótesis nula de frecuencias esperadas.

Una forma de expresar la hipótesis nula es utilizando la notación de relación entre frecuencias. En esta notación se indica la relación entre las frecuencias de las clases, separadas entre ellas por dos puntos (:).

Así, si la relación es de uno a uno para dos clases, se expresa como 1:1. Si es de uno a uno para cinco clases, se expresará como 1:1:1:1:1.

En los casos anteriores, las frecuencias esperadas serán iguales y la prueba será de “uniformidad”.

La conversión a frecuencias esperadas propiamente tales se efectúan de la manera siguiente:

Si n es la frecuencia total y k es el número de clases, cada frecuencia esperada será igual a n/k. Así, para el ejemplo de dos clases (k = 2) y si n = 10, se tiene f1 = 10/2 = 5, f2 = 10/2 = 5; por lo tanto:

$$H_0: f1 = f2 = 5.$$

En el caso de cinco clases y si n = 100, se tiene:

$$H_0: f1 = f2 = f3 = f4 = f5 = 20.$$

Si la prueba no es de uniformidad, la relación entre frecuencias no será del tipo 1:1. En ese caso, será distinta, puede ser de 1:3 o de 1:2:3:9 para cuatro clases, por ejemplo.

Ahora la conversión a frecuencias esperadas se hace algo más complicada. Se procede de la manera siguiente:

Primero, se suman las cifras de la relación de frecuencias. Para dos clases de ejemplo anterior (1:3), será 1 + 3 = 4; y para las cuatro clases (1:2:3:9) 1 + 2 + 3 + 9 = 15.

La frecuencia esperada para cada clase será ahora igual al número de la relación de frecuencias de la clase dada, dividida por la suma de todas las relaciones y luego multiplicada por n, la frecuencia total. De esta manera, si n = 100, las frecuencias esperadas para el ejemplo donde se tiene 1:3, serán:

$f1 = (1/4)(100) = 25$ y $f2 = (3/4)(100) = 75$

En el caso de 1:2:3:9, y si n = 80:

$f1 = (1/15)(80) = 5{,}333$ $\quad$ $f2 = (2/15)(80) = 10{,}667$
$f3 = (3/15)(80) = 16{,}000$ $\quad$ $f4 = (9/15)(80) = 48{,}000$

Es de notar que en todos los casos la suma de las frecuencias esperadas, así como las observadas, es igual a la frecuencia observada total (n).

Otra manera de presentar la hipótesis nula es en forma de proporciones. En esta modalidad, las frecuencias esperadas se obtienen sencillamente multiplicando cada proporción por la frecuencia total (n).

Si las proporciones son:

P1 = P2 = P3 =P4 = 0,25 y si n = 100, se tiene:
H_0: f1 = f2 = f4 = f4 = 25.

Si las proporciones son:

P1 = 0,35; P2 = 0,10; P3 = 0,05; P4 = 0,50 y si n = 120, se tiene:

H_0: f1 = 42; f2 = 12; f3 = 6; f4 = 60.

No se debe olvidar que las frecuencias esperadas también pueden presentarse directamente:

H_0: f1 = 5; f2 = 16; f3 = 25

y que la frecuencia total n = 5 + 16 + 25 = 46.

La prueba de Ji al cuadrado determina la discrepancia que hay entre las frecuencias esperadas y las observadas. Es así que se tiene:

$$X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \quad (1)$$

Donde:

i = 1, 2, …, k es el número de clases o categorías.

O_i: es la frecuencia observada de la i-ésima clase.

E_i: es la frecuencia esperada de la i-ésima clase.

X^2: es el estadístico de prueba, que se distribuye aproximadamente como la distribución Ji al cuadrado con k - 1 grados de libertad.

En la fórmula (1) se puede notar que X^2 será siempre positiva, ya que el numerador es un binomio al cuadrado y el denominador, una frecuencia, que es positiva. Se observa que el valor de X^2 depende directamente de las discrepancias entre las O_i y las E_i. Por otra parte, si las O_i y las E_i concuerdan, X^2 será igual a cero. El valor de X^2 crecerá a medida que los valores de las O_i y las E_i se alejen entre ellas.

En el primer caso habrá concordancia y en el segundo discordancia. De esto se desprende que valores pequeños de X^2 indican concordancia y valores grandes, discordancia.

Una fórmula equivalente a la fórmula 1, más fácil de manejar es la siguiente:

$$X^2 = \sum_{i=1}^{k} \frac{(O_i)^2}{E_i} - n \quad (2)$$

Donde, “n” es la frecuencia total.

Ahora se puede preguntar: ¿qué valores de X^2 son suficientemente grandes para rechazar la hipótesis nula de concordancia entre las frecuencias observadas y las esperadas? Esto depende del valor crítico tabulado para la distribución de Ji al cuadrado, en la Tabla 2.

El valor crítico de Ji al cuadrado depende del nivel de significación (α) y de los grados de libertad (aquí: k - 1).

Ejemplo

- Se desea determinar si la relación entre el número de leucocitos en sangre de seis diferentes morfologías (clases o grupos) es la de 1:1:1:1.1:1, para una especie animal en particular.

Se toma una gota de sangre y en un campo microscópico dado se encuentra el número de leucocitos de cada tipo. En el Cuadro 2.4 se observan los resultados, donde hay 36 leucocitos.

a) Hipótesis de investigación: La relación entre el número de leucocitos de diferentes clases es 1:1:1:1:1:1.

b) Hipótesis estadísticas: H_0: 1 : 1 : 1 : 1 : 1 : 1 o
H_0: P1 = P2 = P3 = P4 = P5 = P6 = 1/6
o
H_0: f1 = f2 = f3 = f4 = f5 = f6 = 6

Se utilizará la última forma, por ser la más conveniente para efectuar cálculos, por lo tanto,

H_0: f1 = f2 = f3 = f4 = f5 = f6 = 6
H_1: no todas las frecuencias son iguales.

c) Estadístico de prueba: se utiliza la prueba de Ji al cuadrado, ya que los datos están agrupados en dos o más clases (seis en este caso) y porque los datos están medidos en escala nominal.

d) Nivel de significación: se usará un nivel de significación de 5% (α = 0,05).

e) Región crítica o Área de rechazo: se rechazará H_0 si el valor de X^2 calculado conforme a la fórmula (1), es mayor o igual que el valor tabulado de Ji al cuadrado en la Tabla 2, para k – 1 = 6 - 1 = 5 grados de libertad y un nivel de significación de 0,05; es decir, se rechazará H_0 si X^2 es mayor o igual que 11,07.

f) Cálculos:

Cuadro 2.4. Frecuencias observadas de cada uno de los tipos de leucocitos en sangre de una especie animal en particular.

Clase	I	II	III	IV	V	VI	n
Frecuencia	3	6	9	7	5	6	36

Cuadro 2.5. Frecuencias observadas, esperadas y cálculo de X^2 por la fórmula (1).

Clase	I	II	III	IV	V	VI
O_i	3	6	9	7	5	6
E_i	6	6	6	6	6	6
$O_i - E_i$	-3	0	3	1	-1	0
$(O_i - E_i)^2$	9	0	9	1	1	0
$\frac{O_i - E_i}{E_i}$	1,50	0	1,50	0,17	0,17	0

X^2 = 3,34

Cuadro 2.6. Frecuencias observadas, esperadas y cálculo de X^2 por la fórmula (2).

Clase	I	II	III	IV	V	VI
O_i	3	6	9	7	5	6
E_i	6	6	6	6	6	6
O_i^2	9	36	81	49	25	36
O_i^2 / E_i	1,50	6,00	13,50	8,17	4,17	6,00
Suma de O_i^2 / E_i						39,34

X^2 = 39,34 – 36,00 = 3,34

g) Decisión: como $X^2 = 3,34$ no es mayor o igual que 11,07 (de la Tabla 2 para $\alpha = 0,05$ y 5 grados de libertad), no se rechaza H_0.

h) Conclusión: se concluye que la relación entre el número de leucocitos de diferentes clases (en el campo microscópico) no es distinto de 1 : 1 : 1 : 1 :1 : 1.

Ejemplo

- Un genetista postula que en la progenie de cierto cruce dihíbrido de una planta dada, los cuatro fenotipos se deben presentar en la razón 9 : 3 : 3 : 1.

El genetista examina 64 miembros (n= 64) de la segunda generación y observa las frecuencias, indicadas en el Cuadro 2.7.

¿Concuerdan las frecuencias observadas con las esperadas?

a) Hipótesis de investigación: la razón entre los cuatro fenotipos en la progenie del cruce dihíbrido es de 9 : 3 : 3 : 1.

b) Hipótesis estadísticas: H_0: 9 : 3 : 3 : 1 o
H_0: P1 = 9/16; P2 = 3/16; P3 = 3/16; P4 = 1/16
o
H_0: f1 = 36; f2 = 12; f3 = 12; f4 = 4

Se utilizará la última forma, por ser la forma más conveniente para efectuar cálculos, por lo tanto:

H_0: f1 = 36; f2 = 12; f3 = 12; f4 = 4
H_1: no todas las frecuencias son iguales.

c) Estadístico de prueba: se utiliza la prueba de Ji al cuadrado, ya que los datos están agrupados en dos o más clases (cuatro en este caso) y porque los datos están medidos en escala nominal.

d) Nivel de significación: se usará un nivel de significación de 1%.

e) Región crítica o Área de rechazo: se rechazará H_0 si el valor de X^2 calculado conforme a la fórmula (1), es mayor o igual que el valor tabulado de Ji al cuadrado en la Tabla 2, para k – 1 = 4 – 1 = 3 grados de libertad y un nivel de significación de 0,01; es decir, se rechazará H_0 si X^2 es mayor o igual que 11,34.

f) Cálculos:

Cuadro 2.7. Frecuencias observadas de cada fenotipo de la progenie de cierto cruce dihíbrido.

Clase	I	II	III	IV
Frecuencia	35	13	15	1

Cuadro 2.8. Frecuencias observadas, esperadas y cálculo de X^2

Clase	I	II	III	IV
O_i	35	13	15	1
E_i	36	12	12	4
$O_i - E_i$	-1	1	3	-3
$(O_i - E_i)^2$	1	1	9	9
$(O_i-E_i)^2/E_i$	0,03	0,08	0, 75	2,25
				X^2 = 3,11

g) Decisión: como $X^2 = 3{,}11$ no es mayor o igual que 11,34 (de la Tabla 2 para $\alpha = 0{,}01$ y 3 grados de libertad), no se rechaza H_0.

h) Conclusión: se concluye que la razón entre los cuatro fenotipos en la progenie del cruce dihíbrido no es distinto de 9 : 3 : 3 : 1.

Prueba de Kolmogorov-Smirnov para una muestra

Fundamentos

La prueba de Kolmogorov-Smirnov, al igual que la prueba de Ji al cuadrado es una prueba de bondad de ajuste.

Con esta prueba se determina si las observaciones de una muestra provienen de una población con una distribución especificada.

La distribución de frecuencias acumuladas teóricas que se han especificado se compararán con la distribución de frecuencias acumuladas observadas.

Se determina el punto en el cual las dos distribuciones acumuladas (la teórica o especificada y la observada) muestran mayor divergencia.

La divergencia máxima se compara con valores críticos tabulados, con el fin de rechazar o no la hipótesis de nulidad de igualdad de las dos distribuciones.

Las observaciones se deben medir, por lo menos, en escala ordinal.

Metodología

Sea Fo(X) una función de distribución de frecuencias relativas acumuladas esperadas (especificadas en la hipótesis nula), donde para cualquier valor de X, Fo(X) es la proporción de observaciones menores o iguales a X.

Sea Sn(X) la distribución de frecuencias relativas acumuladas observadas en la muestra aleatoria de n observaciones.

Si X es el valor de cualquier observación, se tendrá que:

$$Sn(X) = k/n \quad (1)$$

Donde, k es el número de observaciones con valores menores o iguales a X.

De acuerdo con la hipótesis nula (la cual especifica la distribución teórica de donde se supone que fue tomada la muestra), se espera que para cualquier valor de X, Sn(X) y de Fo(X) estén muy cercanos. En otras palabras, se espera que la diferencia entre Sn(X) y Fo(X) sea pequeña y que se deba a factores aleatorios.

Para esta prueba se usa el estadístico D, denominado “desviación máxima” donde:

$$D = \text{máximo } |Fo(X) - Sn(X)| \quad (2).$$

Las barras (| |) indican el valor absoluto de la diferencia.

La distribución de D es conocida y está tabulada en la Tabla 3.

En la hipótesis nula se indican las frecuencias esperadas de la manera siguiente:

Una forma de expresar la hipótesis nula es utilizando la notación de relación entre frecuencias. En esta notación se indica la relación entre las frecuencias de las clases, separadas entre ellas por dos puntos (:).

Así, si la relación es de 1 a 1 para dos clases, se expresará como 1:1. Si es una a uno para cinco clases, se expresa como 1:1:1:1:1.

En los casos anteriores, las frecuencias esperadas serán iguales y la prueba será de "uniformidad".

La conversión a frecuencias esperadas propiamente tales se efectúan de la manera siguiente:

Si n es la frecuencia total y k es el número de clases, cada frecuencia esperada será igual a n/k. Así, para el ejemplo de dos clases (k = 2) y si n = 10, se tendrá f1 = 10/2 = 5, f2 = 10/2 = 5; por lo tanto:

H_0: f1 = f2 = 5.

En el caso de 5 clases y si n = 100, se tendrá:

H_0: f1 = f2 = f3 = f4 = f5 = 20

Si la prueba no es de uniformidad, la relación entre frecuencias no será del tipo 1:1. En ese caso, será distinta, puede ser de 1:3 o de 1:2:3:9 para cuatros clases, por ejemplo.

Ahora la conversión a frecuencias esperadas se hace algo más complicada. Se procede de la manera siguiente:

Primero, se suman las cifras de la relación de frecuencias. Para las dos clases del ejemplo anterior (1:3), será 1 + 3 = 4; y para las cuatro clases (1:2:3:9) 1 + 2 + 3 + 9 = 15.

La frecuencia esperada para cada clase será ahora igual al número de la relación de frecuencias de la clase dada, dividida por la suma de todas las relaciones y luego multiplicada por n, la frecuencia total. De esta manera, si

n = 100, las frecuencias esperadas para el ejemplo donde se tiene 1:3, serán:

f1 = (1/4)(100) = 25 y f2 = (3/4)(100) = 75

En el caso de 1:2:3:9, y si n = 80:

f1 = (1/15)(80) = 5,333 f2 = (2/15)(80) =10,667
f3 = (3/15)(80) = 16,000 f4 = (9/15)(80) = 48,000

Nótese que en todos los casos la suma de las frecuencias esperadas, así como las observadas, es igual a la frecuencia observada total (n).

Otra manera de presentar la hipótesis nula es en forma de proporciones. En esta modalidad, las frecuencias esperadas se obtienen sencillamente multiplicando cada proporción por la frecuencia total (n).

Si las proporciones son:

P1 = P2 = P3 = P4 = 0,25 y si n = 100, se tiene:
H_0: f1 = f2 = f4 = f4 = 25

Si las proporciones son:

P1 = 0,35; P2 = 0,10; P3 = 0,05; P4 = 0,50 y si n = 120, se tendrá:

H_0: f1 = 42; f2 = 12; f3 = 6; f4 = 60

No se debe olvidar que las frecuencias esperadas también pueden presentarse directamente:

H_0: f1 = 5; f2 = 16; f3 = 25

y que la frecuencia total n = 5 + 16 + 25 = 45.

Ejemplo

- En una finca deseamos determinar si en las vacas con mastitis, las frecuencias para los distintos grados de esta enfermedad son iguales.

Se ha tomado una muestra aleatoria de n = 12 vacas con mastitis y anotado el grado de mastitis (grado por la prueba CMT) que cada una tenía.

En el Cuadro 2.9 se muestra los datos.

Cuadro 2.9. Vacas con Mastitis.

Grados	I	II	III	IV
Número de Vacas con Mastitis	3	4	3	2

a) Hipótesis de investigación: las frecuencias para los distintos grados de mastitis son iguales.

b) Hipótesis estadísticas: H_0: 1 : 1 : 1 : 1 o
H_0: f1 = f2 = f3 = f4 = 3
H_1: no todas las frecuencias son iguales.

c) Estadístico de prueba: se utiliza el estadístico de prueba D de Kolmogorov-Smirnov, ya que se desea comparar la distribución de observaciones medidas en escala ordinal (grado de mastitis), con la distribución teórica (indicada en la H_0).

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0,05$).

e) Región crítica o Área de rechazo: se rechaza H_0 si el valor de D calculado, conforme a la fórmula (2) es mayor que el tabulado en la Tabla 3 para n = 12 y $\alpha = 0,05$; es decir, si D > 0,375.

f) Cálculos: en el Cuadro 2.10 se muestran los cálculos.

Cuadro 2.10. Cálculo de D.

Grados	I	II	III	IV
Frecuencias esperadas	3	3	3	3
Frecuencias observadas	3	4	3	2
Fo(X)	3/12	6/12	9/12	12/12
Sn(X)	3/12	7/12	10/12	12/12
|Fo(X) – Sn(X)|	0/12	1/12	1/12	0/12

Se nota que la máxima diferencia absoluta es de 1/12 = 0,083, es decir:

D = máximo |Fo(X) - Sn(X)| = 0,083

g) Decisión: como D = 0,083 no es mayor que el valor tabulado (0,375), no se rechaza la H_0.

h) Conclusión: se concluye que las frecuencias para los distintos grados de mastitis no difieren de 1 : 1 : 1 : 1 o f1 = f2 = f3 = f4 = 3.

Ejemplo

- Se desea determinar en un huerto de frutales, si la relación entre las frecuencias de los distintos grados de infestación por áfidos en naranjos es de 3 : 2 : 2 : 2 : 1.

Se toma una muestra aleatoria de n = 20 naranjos y se evalúa el grado de infestación de cada uno.

En el Cuadro 2.11 se muestran los datos.

Cuadro 2.11. Naranjos infestados por ácaros.

Grados	**+**	**++**	**+++**	**++++**	**+++++**
Naranjos infestados	6	4	4	4	2

a) Hipótesis de investigación: la relación entre las frecuencias de los distintos grados de infestación es igual a 3 : 2 : 2 : 2 : 1.

b) Hipótesis estadísticas: H_0: 3 : 2 : 2 : 2 : 1 o
H_0: f1 = 6; f2 = 4; f3 = 4; f4 = 4; f5 = 2
H_1: no todas las frecuencias son iguales las de la H_0.

c) Estadístico de prueba: se utiliza el estadístico de prueba D de Kolmogorov-Smirnov, ya que se desea comparar la distribución de observaciones medidas en escala ordinal (grado de infestación), con la distribución teórica (indicada en la H_0).

d) Nivel de significación: se utilizará un nivel de significación de 1% ($\alpha = 0,01$).

e) Región crítica o Área de rechazo: se rechaza H_0 si el valor de D calculado, conforme a la fórmula (2) es mayor que el tabulado en la Tabla 3 para n = 20 y $\alpha = 0,01$; es decir, si D > 0,356.

f) Cálculos: en el Cuadro 2.12 se muestran los cálculos.

Cuadro 2.12. Cálculo de D.

Grados	+	++	+++	++++	+++++
Frecuencias esperadas	6	4	4	4	2
Frecuencias observadas	10	8	1	1	0
Fo(X)	6/20	10/20	14/20	18/20	20/20
Sn(X)	10/20	18/20	19/20	20/20	20/20
|Fo(X) –Sn(X)|	4/20	8/20	5/20	2/20	0/20

Se observa que la máxima diferencia es 8/20 = 0,400, es decir:

D = máximo |Fo(X) –Sn(X)| = 0,400.

g) Decisión: como D = 0,400 es mayor que el tabulado (0,356), se rechaza la H_0.

h) Conclusión: se concluye que la relación entre las frecuencias de grado de infestación no es 3 : 2 : 2 : 2 : 1.

Prueba de rachas

Fundamentos

Para llegar a conclusiones valederas acerca de una población, por medio de una muestra, ésta debe ser aleatoria. Para determinar la aleatoriedad de una muestra se utilizan pruebas basadas en el orden o secuencia de la presentación de las observaciones.

La prueba de rachas está basada en el número de rachas de una muestra. Una racha es simplemente una sucesión de eventos idénticos, precedidos y seguidos por eventos distintos a los indicados, o por la ausencia de evento alguno.

Por ejemplo:

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

Se usa a “r” como el número de rachas, aquí r = 7.

Con la prueba de rachas sólo se puede determinar aleatoriedad para la sucesión de los eventos.

El número total de rachas de una muestra indica si ésta es aleatoria o no.

Un número de rachas muy bajo o muy alto, indica falta de aleatoriedad.

Si se observa el orden de llegada de las primeras 12 personas a un cine y se clasifican en adultos (A) y niños (N), se pudiera tener:

NANANANANANA

En este caso hay 12 rachas (r = 12) y se nota que no hay aleatoriedad en el orden de llegada, ya que hay una marcada tendencia a la mezcla (muchas rachas).

Si por otra parte se observa:

NNNNNNAAAAAA

Se tendrían sólo dos rachas (r = 2) y se nota que no hay aleatoriedad en el orden de llegada, ya que hay una notoria tendencia a la formación de conglomerados (pocas rachas).

Se debe notar que el orden de aparición de las observaciones da una información que no proviene de la frecuencia de cada clase.

La distribución muestral de los valores de r, que se puede esperar de muestras aleatorias es conocida. Utilizando esta distribución se puede determinar cuando una muestra tiene rachas de más o de menos de las que ocurrirían probablemente si fuera una muestra aleatoria.

Metodología

Si se clasifican las observaciones de acuerdo a dos clases o categorías, n_1 será el número de ellas que pertenecen a la primera clase y n_2 a la segunda.

El número total de observaciones será igual a $n_1 + n_2 = n$

Para aplicar la prueba de las rachas lo primero que se tiene que hacer es observar el orden en que se presentan las observaciones de cada clase. Luego se marcan las rachas y se cuentan (r).

La prueba de rachas puede ser de una o dos colas. Las pruebas de una cola pueden ser de la cola izquierda o de la cola derecha.

a) Dos colas:

H_0: La secuencia es aleatoria.
H_1: La secuencia no es aleatoria.

b) Cola derecha

H_0: La secuencia es aleatoria.
H_1: La secuencia muestra tendencia a la mezcla.

c) Cola izquierda

H_0: La secuencia es aleatoria.
H_1: La secuencia muestra tendencia a formar conglomeraos.

Cuando ambos n_1 y n_2 son iguales o menores que 20, se puede usar la Tabla 4, donde se encuentran los valores críticos de r para $\alpha = 0{,}05$ en el caso de pruebas de dos colas y $\alpha = 0{,}025$ para pruebas de una cola.

La Tabla 4 está dividida en dos partes: la primera (Tabla 4a) da los valores críticos inferiores de r y la segunda (Tabla 4b) da valores críticos superiores de r.

En pruebas de dos colas, si r es menor o igual al valor tabulado en la Tabla 4a, o mayor o igual al tabulado en la Tabla 4b, se rechaza la H_0 de aleatoriedad, a favor de la H_1, a un nivel de significación de 5%.

En las pruebas de la cola derecha se rechaza la H_0 de aleatoriedad, a favor de la H_1 de tendencia a la mezcla, a un nivel de 2,5%, si r es mayor o igual que el tabulado en la Tabla 4b.

En las pruebas de la cola izquierda se rechaza la H_0 de aleatoriedad, a favor de la H_1 de tendencia a formar conglomerados, a un nivel de 2,5%, si r es menor o igual que el tabulado en la Tabla 4a.

Ejemplo

- En un corral entran 39 chivos. Se desea saber si el orden de entrada en que entran los machos (M) y las hembras (H) es aleatorio.

a) Hipótesis de investigación: el orden de entrada de los chivos, respecto al sexo, es aleatorio.

b) Hipótesis estadísticas: H_0: La secuencia es aleatoria.
H_1: La secuencia no es aleatoria.

c) Estadístico de prueba: se utilizará el estadístico de prueba r, de la prueba de rachas, para una muestra.

d) Nivel de significación: se utilizará un nivel de significación de 5%, ya que la prueba es de dos colas y sólo tenemos valores críticos para $\alpha = 0,05$.

e) Región crítica o Área de rechazo: para $n_1 = 19$ y $n_2 = 20$ se encuentra el valor de r = 13 en la Tabla 4a y r = 27 en la Tabla 4b. Se rechaza H_0 si $r \leq 13$ o $r \geq 27$.

f) Cálculos: a continuación se presenta el orden de entrada de chivos a un corral por sexo (H = hembra y M = macho), además se indican las rachas con un subrayado:

<u>M</u> <u>H</u> <u>MMMM</u> <u>H</u> <u>MMMM</u> <u>HHH</u> <u>MMMMM</u> <u>HH</u> <u>MMMMM</u> <u>HHHHHHHH</u> <u>M</u> <u>HHHH</u>

Se observa que hay 12 rachas, luego r = 12.

g) Decisión: como r = 12 es menor que 13 el valor crítico tabulado en la Tabla 4a, se rechaza H_0 de aleatoriedad a favor a la H_1 de no aleatoriedad.

h) Conclusión: el orden de entrada (por sexos) de los 39 chivos no es aleatorio.

Ejemplo

- En un corral entran 39 chivos. Se desea saber si el orden de entrada en que entran los machos (M) y las hembras (H) tiene tendencia a formar conglomerados (grupos del mismo sexo).

a) Hipótesis de investigación: lo chivos tienden a formar grupos del mismo sexo.

b) Hipótesis estadísticas: H_0: La secuencia es aleatoria.
H_1: La secuencia muestra tendencia a formar conglomerados.

c) Estadístico de prueba: se utilizará el estadístico de prueba r, de la prueba de rachas, para una muestra.

d) Nivel de significación: se utilizará un nivel de significación de 2,5%, ya que la prueba es de una cola y sólo se tiene valores críticos para $\alpha = 0{,}025$.

e) Región crítica o Área de rechazo: para $n_1 = 19$ y $n_2 = 20$ encontramos el valor de r = 13 en la Tabla 4a, la cual se usa para pruebas con la cola izquierda. Se rechaza H_0 si $r \leq 13$.

f) Cálculos: a continuación se presenta el orden de entrada de chivos a un corral por sexo (H = hembra y M = macho), además se indican las rachas con un subrayado.

M H MMMM H MMMM HHH MMMMM HH MMMMM HHHHHHHH
M HHHH

Se observa que hay 12 rachas, luego r = 12.

g) Decisión: como r = 12 es menor que 13 el valor crítico tabulado en la Tabla 4a, se rechaza H_0 de aleatoriedad a favor a la H_1 de tendencia a formar conglomerados.

h) Conclusión: se concluye que el orden de entrada (por sexos) de los 39 chivos tiene tendencia a formar conglomerados.

Ejemplo

- En una siembra de maíz se desea determinar si el ataque por parte del “gusano cogollero” en una hilera de plantas, muestra tendencia a la mezcla.

Se toma al azar una hilera de plantas.

Hay 25 plantas, de las cuales 17 han sido atacadas por el gusano cogollero (G) y ocho están sanas (S).

a) Hipótesis de investigación: el orden en que se presentan las plantas atacadas muestra tendencia a la mezcla.

b) Hipótesis estadísticas: H_0: La secuencia es aleatoria.
H_1: La secuencia muestra tendencia a la mezcla.

c) Estadístico de prueba: se utilizará el estadístico de prueba r, de la prueba de rachas para una muestra.

d) Nivel de significación: se utilizará un nivel de significación de 2,5%, ya que la prueba es de una cola y sólo se tiene valores críticos para $\alpha = 0{,}025$.

e) Región crítica o Área de rechazo: para $n_1 = 17$ y $n_2 = 8$ se encuentra el valor de r = 17 en la Tabla 4b, la cual se usa para pruebas con la cola derecha. Se rechaza H_0 si $r \geq 17$.

f) Cálculos: a continuación se presenta el orden de plantas atacadas por el gusano cogollero, además se indican las rachas con un subrayado.

<u>G</u> <u>S</u> <u>GGG</u> <u>S</u> <u>GG</u> <u>S</u> <u>GG</u> <u>S</u> <u>GG</u> <u>S</u> <u>GG</u> <u>S</u> <u>GG</u> <u>S</u> <u>G</u> <u>S</u> <u>GG</u>

Se observa que hay 17 rachas, luego r = 17.

g) Decisión: como r = 17 es igual que 17 el valor crítico tabulado en la Tabla 4b, se rechaza H_0 de aleatoriedad a favor a la H_1 de tendencia a la mezcla.

h) Conclusión: se concluye que el orden de ataque a las plantas en una hilera de plantas de maíz por parte del “gusano cogollero”, muestra tendencia a la mezcla (alternabilidad).

Discusión de los métodos para una muestra

De las pruebas estudiadas, tres son de tipo “bondad de ajuste” y la cuarta de aleatoriedad. En el caso de las pruebas de bondad de ajuste, para la escogencia de la adecuada, se debe tomar en cuenta lo siguiente:

- Número de categorías o clases.

- Nivel de medida utilizado.

- Tamaño de la muestra.

a) La prueba binomial se puede utilizar cuando hay sólo dos categorías o clases y cuando el tamaño de la muestra es tan pequeño que no se puede aplicar la prueba Ji al cuadrado.

b) La prueba de Ji al cuadrado se puede utilizar cuando las observaciones están clasificadas en categorías discretas y cuando las frecuencias esperadas sean suficientemente grandes. Cuando hay sólo dos categorías, cada frecuencia esperada deba ser mayor o igual a cinco. Cuando hay más de dos categorías, no más de 20% de las frecuencias esperadas deberá ser menor que cinco y ninguna menor que uno.

c) La prueba de Kolmogorov-Smirnov se puede utilizar cuando se puede asumir que la variable en cuestión (subyacente) tiene una distribución continua. Sin embargo, si la distribución es discreta y se rechaza H_0 no habrá error.

La escala de medida debe ser, por lo menos, ordinal.

d) La prueba de aleatoriedad se utiliza cuando se desea determinar si la secuencia temporal de un evento con dos clases o categorías, es aleatoria, muestra tendencia a formar grupos (conglomerados) o muestra tendencia a la alternabilidad (mezcla).

Nota: para una prueba de bondad de ajuste a la distribución normal en muestras pequeñas, se puede usar el manual de Wiedenhofer (1980).

Capítulo III

Caso de dos muestras relacionadas

Se utilizan las pruebas de dos muestras relacionadas cuando se desea detectar alguna diferencia entre dos tratamientos (grupos, condiciones, ocasiones, procedimientos, entre otros), si uno es mejor (superior) o peor (inferior) que el otro.

A veces, debido a la heterogeneidad del material experimental, las diferencias reales entre tratamientos pueden estar enmascaradas. Se detectan diferencias donde no las hay, y cuando las hay, no se detectan.

Cuando se sospecha la existencia de esta situación, se pueden aplicar tratamientos a una misma unidad experimental (animal, planta, jaula, parcela, entre otras) consecutivamente en el espacio del tiempo y en un orden determinado, con el fin de eliminar la variabilidad entre unidades experimentales dentro de tratamientos. Cada unidad experimental actúa como su propio control.

Si las diferencias que se obtienen se miden, por lo menos, en escala de intervalo y los errores se distribuyen normal e independientemente, se puede usar la prueba "t" de Student.

Si las diferencias son independientes, pero no se distribuyen normalmente, o si la escala de medición es inferior a la de intervalo, entonces se debe utilizar alguna de las pruebas no paramétricas que se presenta a continuación.

Prueba de McNemar

La prueba de McNemar también es llamada "Prueba para la significación de cambios", es particularmente adecuada para los diseños "antes y después", en los cuales cada unidad actúa como su propio control y la respuesta se mide en escala nominal.

Se utiliza también esta prueba para determinar concordancia o discordancia de resultados, cuando un mismo material se analiza por dos métodos u operadores diferentes.

Metodología

Se elabora una tabla de doble entrada para probar la significación de cambios, la cual puede ser de "antes a después" o "de un método a otro" o "de un "operador a otro", pero siempre la respuesta se debe medirla por lo menos en escala nominal.

Las partes fundamentales de la tabla de doble entrada se muestran a continuación:

Tabla de doble entrada

		Después	
		-	+
Antes	+	A	B
	-	C	D

		Método 1	
		-	+
Método 2	+	A	B
	-	C	D

Los signos “+” y “-” se utilizan para simbolizar las dos respuestas posibles.

Se clasificará una unidad (sujeto, animal u objeto) en la celda “A” si ha cambiado de “+” a “-” o si es “-” para el “método 1” y “+” para el “método 2”.

En la celda “D” se clasificará a las que cambian de “-” a “+” o si su respuesta al “método 1” es “+” y al “método 2” es “-”

Cuando no hay cambio y las respuestas han sido “+” se clasificará en la celda “B”. Si las respuestas han sido “-” y no hay cambio, se clasificará en la celda “C”.

Donde, (A + D) representa el total de unidades que cambiaron y se espera que, conforme a la hipótesis nula (H_0) (A + D)/2 unidades cambien en una dirección y (A + D)/2 unidades en la otra.

Se utilizará el estadístico X^2:

$$X^2 = \frac{(A-D)^2}{(A+D)^2} \quad (1)$$

Donde, X^2 se distribuye como Ji al cuadrado, con un grado de libertad.

Para una mejor aproximación se utilizará:

$$X^2 = \frac{(|A-D|-1)^2}{(A+D)^2} \quad (2)$$

La significación de X^2 con un grado de libertad (a un nivel dado) se determina comparándola con el valor tabulado en la Tabla 2.

Ejemplo

- A un conjunto de plantas se les ha aplicado dos procedimientos distintos para diagnosticar la presencia del agente causal de una enfermedad dada.

Se desea saber si hay concordancia entre los métodos de diagnóstico.

a) Hipótesis de investigación: hay concordancia entre los métodos de diagnóstico.

b) Hipótesis estadísticas: H_0: P+- = P-+
H_1: P+- ≠ P-+

c) Estadístico de prueba: se utiliza el estadístico X^2 con el método de McNemar, ya que se trata de dos muestras relacionadas, del tipo de cambio de un método a otro. Las observaciones se hacen en la escala nominal (presencia, ausencia).

d) Nivel de significación: se utilizará un nivel de significación de 5% (α = 0,05).

e) Región crítica o Área de rechazo: se rechaza H_0 si el valor de X^2 conforme a la fórmula (2), es mayor o igual que el tabulado de Ji al cuadrado de la Tabla 2, con un grado de libertad, para α = 0,05: es decir, 3,84.

f) Cálculos: a continuación se presenta la tabla de doble entrada con los datos ya clasificados, según los resultados del diagnóstico por cada uno de los procedimientos, como resultado de la aplicación de dos métodos de diagnóstico de una enfermedad a un conjunto de 73 plantas.

		Método 1		
		-	+	**Totales**
Método 2	**+**	6 (A)	41 (B)	**47**
	-	12 (C)	14 (D)	**26**
Totales		**18**	**55**	**73**

Nota: "+" indica diagnóstico positivo.
"-" indica diagnóstico negativo.

Se nota que A = 6 y D = 14.

Donde:

"A" indica el número de plantas negativas para el método 1 y positivas para el método 2.

"D" indica el número de plantas positiva para el método 1 y negativas para el método 2.

Los datos en "B" y "C" no interesan para el cómputo de X^2, por lo tanto, utilizando la fórmula (2) se tendrá:

$$X^2 = \frac{(|6-14|-1)^2}{6+14} = \frac{(|-8|-1)^2}{20} = \frac{(8-1)^2}{20} = \frac{49}{20} = 2,45$$

g) Decisión: como X^2 = 2,45 no es mayor o igual que 3,84 (de la Tabla 2, para α = 0,05 y un grado de libertad) no se rechaza la H_0.

h) Conclusión: hay concordancia entre los métodos de diagnóstico.

Ejemplo

- Se toma al azar una muestra de n = 120 perros callejeros y se determina la presencia o ausencia de una "condición" dada.

Se somete los perros a una situación de "estrés" intenso, para determinar el cambio que produce la "condición".

Se quiere saber si el “estrés” aumenta la proporción de perros con la “condición”.

a) Hipótesis de investigación: el “estrés” hace aumentar la proporción de perros con la “condición”.

b) Hipótesis estadísticas: $H_0: P_2 \leq P_1$
$H_1: P_2 > P_1$

Donde:

P_1 es la proporción de perros con la “condición”, antes de aplicar el “estrés”.

P_2 es la proporción de perros con la “condición”, después de aplicar el “estrés”.

c) Estadístico de prueba: se usa el estadístico X^2 con el método de McNemar, ya que se trata de muestras relacionadas del tipo de cambio “antes a después”. Las observaciones se han hecho en escala nominal (presencia, ausencia).

d) Nivel de significación: se usará un nivel de significación de 1% ($\alpha = 0{,}01$).

e) Región crítica o Área de rechazo: debido a que la H_1 indica dirección (cola derecha), se rechazará la H_0 si D > A y X^2 es mayor o igual que la Ji al cuadrado en la Tabla 2, para un grado de libertad y $2\alpha = 0{,}02$; es decir, 5,41.

f) Cálculos: a continuación se muestra la tabla de doble entrada con los datos ya clasificados, según sea la “condición”, antes y después del “estrés”, como resultado de la aplicación del “estrés” a un conjunto de 120 perros callejeros.

		Después -	Después +	Totales
Antes	+	22 (A)	20 (B)	**42**
	-	28 (C)	60 (D)	**88**
Totales		**40**	**80**	**120**

Nota: "+" indica presencia de la "condición".
"-" indica ausencia de la "condición"

Se nota que A = 22 y D = 60.

Donde:

"A" indica el número de perros que cambiaron de presencia a ausencia de la "condición"

"D" indica el número de perros que cambiaron de ausencia a presencia de la "condición" (enfermaron).

Los datos en "B" y "C" no interesan para el cálculo de X^2. Aplicando la fórmula (3) se obtendrá:

$$X^2 = \frac{(|22-60|-1)^2}{22+60} = \frac{(|-38|-1)^2}{82} = \frac{(38-1)^2}{82} = \frac{1369}{82} = 16,70$$

g) Decisión: como D = 38 > 22 = A y X^2 = 16,70 es mayor que 5,41 (de la Tabla 2, para 2α = 0,02 y un grado de libertad), se rechaza H_0.

h) Conclusión: el "estrés" hace aumentar la proporción de perros con la "condición".

Ejemplo

- Se toma al azar una muestra de n = 120 perros callejeros y se determina la presencia o ausencia de una "condición" dada.

Se somete los perros a un "tratamiento" que se espera que haga disminuir la proporción de perros con la "condición".

a) Hipótesis estadísticas: el "tratamiento" hace disminuir la proporción de perros con la "condición".

b) Hipótesis estadísticas: H_0: $P_2 \geq P_1$
H_1: $P_2 < P_1$

c) Estadístico de prueba: se usa el estadístico X^2 con el método de McNemar, ya que se trata de muestras relacionadas del tipo de cambio "antes a después". Las observaciones se han medido en escala nominal (presencia, ausencia).

d) Nivel de significación: se utilizará un nivel de significación de 5% (α = 0,05).

e) Región crítica o Área de rechazo: debido a que la H_1 indica dirección (cola izquierda), se rechazará la H_0 si D < A y X^2 es mayor o igual que la Ji al cuadrado en la Tabla 2, para un grado de libertad y 2α = 0,10; es decir, 2,71.

f) Cálculos: a continuación se muestra la tabla de doble entrada con los datos ya clasificados, según sea la "condición", antes y después del "tratamiento", como resultado de la aplicación del "tratamiento" a un conjunto de 120 perros callejeros.

		Después -	Después +	Totales
Antes	+	44 (A)	20 (B)	**64**
	-	18 (C)	38 (D)	**56**
Totales		62	58	**120**

Nota: "+" indica presencia de la "condición".
"-" indica ausencia de la "condición"

Se observa que A = 44, el número de perros que cambiaron de la presencia a ausencia de la "condición", es decir, que se curaron.

Donde, D = 38, el número de perros que cambiaron de ausencia a presencia de la "condición", a pesar del "tratamiento" (enfermaron).

Los datos en "B" o "C" no interesan para el cálculo de X^2:

$$X^2 = \frac{(|44-38|-1)^2}{44+38} = \frac{(|6|-1)^2}{82} = \frac{25}{82} = 0,30$$

g) Decisión: como D = 38 < 44 = A, pero X^2 = 0,30 no es mayor o igual que 2,71 (de la Tabla 2, para 2α = 0,10 y un grado de libertad), no se rechaza la H_0.

h) Conclusión: el tratamiento no hace disminuir la proporción de perros con la “condición” (el “tratamiento no es efectivo”).

Frecuencias esperadas pequeñas

Si la frecuencia esperada, es decir, (A + D)/2 es menor que cinco, es decir, (A + D) menor que 10, se debe utilizar la prueba binomial en vez de X^2.

Para la prueba binomial “n” = A + D, “x” y “n – x” corresponden a A y a D, respectivamente o viceversa, según sea una prueba con la cola derecha o izquierda. En el caso de dos colas no importa el orden.

Si se usa la prueba binomial en el ejemplo donde se aplicó dos procedimientos a un conjunto de plantas, con el fin de diagnosticar la presencia del agente causal de la enfermedad, se tendrá que n = 20, x = 6 y n - x = 14; n – x es el mayor y la prueba es de dos colas; para rechazar H_0, con α = 0,05 y n = 20 se debe tener un valor de n – x mayor que 15, el valor crítico en la Tabla 1.

Como 14 no es mayor que 15, no se rechaza H_0. Este resultado concuerda con el obtenido con X^2.

Prueba de los signos

La prueba de los signos, así como la de McNemar, es útil para los casos "antes y después" o "método 1 y método 2", pero cuando hay un cambio medible en la escala ordinal, como "aumenta" "disminuye" "mejora", "empeora", "es superior", "es inferior", entre otros.

Metodología

Esta prueba debe su nombre al uso de los signos más (+) y menos (-), en lugar de cantidades.

Cuando la medición cuantitativa de las diferencias es impráctica o imposible, este método resulta particularmente adecuado, siempre y cuando por lo menos exista una relación de orden entre las respuestas "antes y después". Se utilizará este método cuando se desea establecer la diferencia entre dos tratamientos o situaciones.

Sólo se asume que la variable en estudio proviene de una distribución continua.

No se asume nada acerca de forma de la distribución de diferencias, ni se exige que las unidades provengan de una misma distribución.

Para facilitar el procedimiento y los cálculos, se encuentra la diferencia (en escala ordinal) usando el "después" o "método 2" como minuendo y el "antes" o "método 1" como sustraendo:

Si "después" > "antes" se le asignará a la diferencia el signo "+".

Si “después” < “antes” se le asignará a la diferencia el signo “-”.

Si “después” = antes” se le asignará cero a la diferencia y se reducirá en uno a “n”.

Para el “método 1 y método 2” se usará el mismo procedimiento, haciendo: “método 1” equivalente a “antes” y “método 2” a “después”.

Se utilizará la prueba binomial con el número de signos “+” o signos “-”, según sea el caso.

Empates: cuando en un par igualado o un antes y después hay un empate, es decir que ambos valores son iguales, se descarta el par y se reduce en uno a “n”.

Ejemplo

- Se desea comprobar el efecto que tiene cierto “medicamento” sobre un tipo de parálisis del tren posterior (patas traseras) en gatos.

Se toma una muestra aleatoria de 15 gatos enfermos, a los cuales se les aplica el “medicamento”.

Se observa cada gato, antes y después de la aplicación del “medicamento”, para determinar si mejoró, empeoró o quedó igual.

El grado de la enfermedad se indica con cruces (†), a mayor número de cruces, mayor gravedad. La ausencia de cruces indica la ausencia de la enfermedad. En el Cuadro 3.1 se presentan los datos y cálculos.

a) Hipótesis de investigación: el medicamento produjo un cambio en la condición de la enfermedad de los gatos.

b) Hipótesis estadísticas: $H_0: Md_D = 0$
$H_1: Md_D \neq 0$

Donde, Md_D es la mediana de las diferencias.

c) Estadístico de prueba: se utilizará el estadístico "x" para el signo "+" o "-", dependiendo de cuál es más frecuente, en la prueba binomial (para la prueba de los signos). Esto porque se trata de un caso de dos muestras relacionadas, donde la diferencia se mide en escala ordinal.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0,05$).

e) Región crítica o Área de rechazo: debido a que la H_1 no indica dirección, en este caso se trata de una prueba de dos colas.

Se cuenta el número de signos "+" y de signos "-", se toma el de mayor frecuencia y éste será "x".

Se suma el número de signos "+" con él de "-" y esto será "n" (se ignora, es decir, no se cuentan los ceros, los cuales constituyen "empates").

Si "x" o "n – x" (el mayor) es mayor que el tabulado en la Tabla 1, para pruebas de dos colas (prueba bilateral), $\alpha = 0,05$ y "n" a la suma de las frecuencias de "+" y "-", se rechaza H_0.

f) Cálculos: en el Cuadro 3.1 se muestran los resultados de aplicar el "medicamento" a los gatos con parálisis del tren posterior, así como los cálculos.

Cuadro 3.1. Efecto de cierto "medicamento" sobre la parálisis del tren posterior de 15 gatos.

Gato Nº	Antes	Después	Diferencia
1	†††	†††	0
2	††	†	+
3	†	††	-
4	†††	†	+
5	††	†††	-
6	†††	†	+
7	††		+
8	††	†	+
9	†††	†	+
10	†††		+
11	†††	††	+
12	†††		+
13	†	††	-
14	†		+
15	††	†	+

Nota: "+" indica mejoría; "-" indica empeoramiento; "0" indica que no hubo cambio; "†" indica el grado de gravedad de la enfermedad (ausencia de "†" indica animal aparentemente sano).

Donde, † > †† = + (en términos de salud).

Número de signos "+": 11, número de signos "-": 3, n = 14 signos.

g) Decisión: se usará como "x" al número de signos "+", por ser los más abundantes, luego x = 11.

Se busca en la Tabla 1, para n = 14, dos colas y α = 0,05, el valor crítico: 12. Como x = 11 no es mayor o igual que 12, no se rechaza H_0.

h) Conclusión: el "medicamento" no produjo un cambio en la condición de la enfermedad de los gatos.

Ejemplo

- Se desea comprobar si el efecto que tiene cierto "medicamento" sobre un tipo de parálisis del tren posterior (patas traseras) en gatos produce una mejoría en la enfermedad.

Se toma una muestra aleatoria de 15 gatos enfermos, a los cuales se les aplica el "medicamento".

Se observa cada gato antes y después de la aplicación del "medicamento", para determinar si mejoró, empeoró o quedó igual.

El grado de la enfermedad se indica con cruces (†), a mayor número de cruces, mayor gravedad. La ausencia de cruces indica la ausencia de la enfermedad. En el Cuadro 3.2 se presentan los datos y cálculos.

a) Hipótesis de investigación: el medicamento produjo una mejoría en la condición de la enfermedad de los gatos.

b) Hipótesis estadísticas: H_0: $Md_D \leq 0$
H_1: $Md_D > 0$

Donde, Md_D es la mediana de las diferencias.

c) Estadístico de prueba: se utilizará el estadístico “x” para la frecuencia del signo “+”, en la prueba binomial (para la prueba de los signos).

Esto porque se trata de un caso de dos muestras relacionadas, donde la diferencia se mide en escala ordinal.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0{,}05$).

e) Región crítica o Área de rechazo: debido a que la H_1 indica dirección (cola derecha), en este caso se trata de una prueba de una cola.

Se cuenta el número de signos “+” que en este caso (cola derecha) será “x”.

Se suma el número de signos “+” con él de “-” y esto será “n” (se ignora, es decir, no se cuentan los ceros, los cuales constituyen “empates”).

Si x es mayor que n/2, donde “n” es la suma de las frecuencias de “+” y “-”, y si “x” es mayor que el tabulado en la Tabla 1 en pruebas de una cola (prueba unilateral) para $\alpha = 0{,}05$; entonces se rechaza H_0.

f) Cálculos: en el Cuadro 3.2 se muestran los resultados de aplicar el “medicamento” a los gatos con parálisis del tren posterior, así como los cálculos.

Cuadro 3.2. Efecto de cierto "medicamento" sobre la parálisis del tren posterior de 15 gatos.

Gato Nº	Antes	Después	Diferencia
1	†††	†††	0
2	††	†	+
3	†	††	-
4	†††	†	+
5	††	†††	-
6	†††	†	+
7	††		+
8	††	†	+
9	†††	†	+
10	†††		+
11	†††	††	+
12	†††		+
13	†	††	-
14	†		+
15	††	†	+

Nota: "+" indica mejoría. "-" indica empeoramiento, "0" indica que no hubo cambio. "†" indica el grado de gravedad de la enfermedad (ausencia de "†" indica animal aparentemente sano.

Donde, † > †† = + (en términos de salud).

Número de signos “+”: 11, número de signos “-”: 3, n = 14 signos.

g) Decisión: se usa “x”, el número de signos “+” (11), que es mayor que n/2 = 14/2 = 7 y en la Tabla 1, se puede ver que para n = 14, una cola y $\alpha = 0{,}05$, el valor crítico es 11.

Como x = 11 no es mayor que 11 (el calculado), no se rechaza H_0.

h) Conclusión: el “medicamento” no produjo una mejoría en la condición de la enfermedad de los gatos.

Ejemplo

- Se desea probar el efecto de un fungicida en plantas de tomate, infectadas por hongos.

Se toman al azar 22 plantas y se observa el grado de infección (desde uno para incipiente hasta cinco para ataque masivo).

Se observan las plantas antes de aplicar el fungicida, luego se aplica el fungicida, se espera un tiempo determinado y se observa el grado de infección de nuevo, en las mismas plantas.

Se quiere saber si el fungicida reduce el grado de infección causado por hongos en plantas de tomate.

En el Cuadro 3.3 se muestran los datos y cálculos.

a) Hipótesis de investigación: la aplicación del fungicida reduce el grado de infección por hongos en plantas de tomate.

b) Hipótesis estadísticas: $H_0: Md_D \geq 0$

$H_1: Md_D < 0$

Donde, Md_D es la mediana de las diferencias.

c) Estadístico de prueba: se utilizará el estadístico “n - x”, que es la frecuencia de los signos “-”; y “x” la de los signos “+” en la prueba binomial (para la prueba de los signos).

Esto porque se trata de un caso de dos muestras relacionadas, donde la diferencia se mide en escala ordinal.

d) Nivel de significación: se utilizará un nivel de significación de 1% ($\alpha = 0{,}01$).

e) Región crítica o Área de rechazo: debido a que la H_1 indica dirección (cola izquierda), en este caso se trata de una prueba de una cola.

Se cuenta el número de signos “-”.

Se suma el número de signos “+” con él de “-” y esto será “n” (se ignora, es decir, no se cuentan los ceros, los cuales constituyen “empates”).

Si “n - x” es mayor que el tabulado en la Tabla 1, para pruebas de una cola (prueba unilateral), y si “n - x” es mayor que n/2, donde “n” es la suma de las frecuencias de “+” y “-”; para $\alpha = 0{,}01$; entonces se rechaza H_0.

f) Cálculos: en el Cuadro 3.3 se muestran los datos y cálculos.

Cuadro 3.3. Efecto de un fungicida sobre un ataque de hongos en plantas de tomate.

Planta Nº	Grado de infección		
	Antes	Después	Diferencia
1	4	3	-
2	3	3	0
3	2	1	-
4	5	3	-
5	2	2	0
6	3	1	-
7	5	1	-
8	4	2	-
9	4	3	-
10	4	5	+
11	5	1	-
12	2	1	-
13	3	2	-
14	3	1	-
15	4	2	-
16	5	4	-
17	5	1	-
18	3	1	-
19	3	2	-
20	5	2	-
21	4	2	-
22	3	1	-

Nota: "-" indica reducción en el grado de infección, "+" indica incremento en el grado de infección y 0 indica que no hubo cambio.

Número de signos "+": 1
Número de signos "-": 19
Número de ceros: 2

Donde, “n” = 22 – 2 = 20; “x” = 1; “n – x” = 19

g) Decisión: se usa n - x = 19 que es el número de signos “-”.

Se busca en la Tabla 1, para n = 20, una cola y α = 0,01, el valor crítico es 16. Como n - x = 19 es mayor que 16 y también mayor que n/2 = 10, se rechaza H_0.

h) Conclusión: la aplicación del fungicida reduce el grado de infección por hongos en plantas de tomate.

Prueba de Wilcoxon

En la prueba de los signos se nota que ésta sólo utiliza la información acerca de la diferencia de los pares de observaciones. Si, además, se considera la magnitud de dicha diferencia (si es posible), se puede usar la prueba de Wilcoxon. En esta prueba se da mayor importancia al par de observaciones que muestre una mayor diferencia.

Metodología

Para cada par de observaciones se calcula la diferencia entre ellos (d_i).

Se halla el valor absoluto de cada diferencia (se prescinde del signo) y luego se ordenan de menor a mayor.

A la menor se le asigna el rango uno, al que sigue el dos y así sucesivamente.

A cada rango se le asigna (a su derecha) el signo que tenía originalmente.

Se suman las diferencias que tienen el signo "+" a su derecha y al total se denomina T+. Se suman las diferencias que tienen el signo "-" y a la suma se denomina T-.

Se debe notar que la suma de T+ y T- debe ser igual a n(n + 1)/2, donde "n" es el número de diferencias distintas de cero.

Empates

Con frecuencia las respuestas de algún(os) par(es) es(son) igual(es), por lo tanto, $d_i = 0$. En este caso, se descarta éste(éstos) par(es) de observación(es) y se reduce "n" en uno, por cada par descartado.

Otro tipo de empate puede ocurrir, cuando una $|d_i|$ es igual a otra, como por ejemplo: $|d_1| = |5|$, $|d_6| = |5|$ $|d_8| = |-5|$.

En este caso, se le asigna a cada par empatado el promedio de los rangos (rango medio) que se le hubiera asignado a las diferencias si no hubieran sido iguales.

Si dos pares tienen diferencias absolutas iguales, por ejemplo -1 y 1 y los rangos hubieran sido 1 y 2, respectivamente, entonces se le asigna a ambos el valor (1 + 2)/2 = 1,5; por lo tanto, no se observaría los rangos 1, 2, 3, sino 1,5; 1,5; 3 (en el caso de hayan tres diferencias).

Ejemplo

- En un experimento con 10 pares igualados de plantas de maíz, a uno de los miembros de cada par se tratará con una corriente eléctrica débil; al otro se deja sin tratamiento (control).

Después de un período de tiempo dado, se mide el "vigor de crecimiento" (tamaño de la planta, grosor del tallo, número y tamaño de las hojas, entre otros) de cada planta, en forma global y "al ojo".

Se utiliza una escala del uno al 20, dónde 20 es el mayor "vigor de crecimiento".

Se desea saber si la aplicación de una corriente eléctrica débil tiene alguna influencia sobre el "vigor de crecimiento" de las plantas de maíz. Los datos y resultados se muestran en el Cuadro 3.4.

a) Hipótesis de investigación: la aplicación de una corriente eléctrica débil produce un cambio en el "vigor de crecimiento".

b) Hipótesis estadísticas: H_0: $Md_D = 0$
H_1: $Md_D \neq 0$

Donde, Md_D es la mediana de las diferencias (tratadas – no tratadas).

c) Estadístico de prueba: el estadístico de prueba "T" de la prueba de Wilcoxon es el adecuado, ya que se utilizan dos muestras relacionadas, las diferencias se pueden ordenar y la distribución de la variable subyacente es continua.

d) Nivel de significación: se utilizará un nivel de significación de 1% ($\alpha = 0,01$).

e) Región crítica o Área de rechazo: como se trata de una prueba de dos colas, se rechaza H_0 si la T- calculada es menor que la T- tabulada o si T+ calculada es mayor que la T+ tabulada en la Tabla 5, para α = 0,01; en pruebas de dos colas y “n” igual al número de pares cuyas diferencias sean distintas de cero.

f) Cálculos: en el Cuadro 3.4 se muestra el resultado del experimento, así como los cálculos necesarios.

Cuadro 3.4. Efecto de una corriente débil sobre el “vigor de crecimiento” de plantas de maíz.

Par	Planta tratada	Planta no tratada	Diferencia d_i	$\lvert d_i \rvert$	Rango de $\lvert d_i \rvert$	+	-
1	19	13	6	6	4+	4	
2	18	17	1	1	1+	1	
3	20	10	10	10	6+	6	
4	20	8	12	12	8+	8	
5	17	13	4	4	3+	3	
6	16	9	7	7	5+	5	
7	17	17	0	0	0		
8	2	16	-14	14	9-		9
9	16	19	-3	3	2-		2
10	18	7	11	11	7+	7	
						T+ = 34	**T- = 11**

Hay un par que tiene diferencia cero, por lo tanto, n = 10 – 1 = 9.

g) Decisión: en la Tabla 5 se puede observar que los valores críticos de T- y T+, para dos colas, α = 0,01 y n = 9, son iguales a 1 y 44, respectivamente.

Como la T- calculada es igual a 11 y no es menor que 1, ni T+ igual a 34 que no es mayor que 44, no se rechaza H_0.

h) Conclusión: la aplicación de corriente eléctrica débil no influye sobre el “vigor de crecimiento”.

Ejemplo

- En un experimento con 10 pares de plantas de maíz, a uno de los miembros de cada par se le trata con una corriente eléctrica débil; al otro se le deja sin tratamiento (control).

Después de un período de tiempo dado, se mide el “vigor de crecimiento” (tamaño de la planta, grosor del tallo, número y tamaño de las hojas, entre otros.) de cada planta, en forma global y “al ojo”.

Se utiliza una escala del uno al 20, dónde 20 es el mayor “vigor de crecimiento”.

Se desea saber si la aplicación de una corriente eléctrica débil produce un incremento en el “vigor de crecimiento” de las plantas de maíz. Los datos y cálculos se muestran en la Cuadro 3.5.

a) Hipótesis de investigación: la aplicación de una corriente eléctrica débil produce un incremento en el “vigor de crecimiento”.

b) Hipótesis estadísticas: H_0: $Md_D \leq 0$
H_1: $Md_D > 0$

Donde, Md_D es la mediana de las diferencias (tratadas – no tratadas).

c) Estadístico de prueba: el estadístico de prueba “T” de la prueba de Wilcoxon es el adecuado, ya que se utilizan dos muestras relacionadas, las diferencias se pueden ordenar y la distribución de la variable subyacente es continua. Se trata de una prueba de una cola, la cola derecha.

d) Nivel de significación: se utilizará un nivel de significación de 5% (α = 0,05).

e) Región crítica o Área de rechazo: como se trata de una prueba de una cola (derecha), se rechaza H_0 si la T+ calculada es mayor que la T+ tabulada en la Tabla 5, para α = 0,05; en pruebas de una cola y "n" igual al número de pares cuyas diferencias sean distintas de cero.

f) Cálculos: en el Cuadro 3.5 se muestra el resultado del experimento, así como los cálculos necesarios.

Cuadro 3.5. Efecto de una corriente débil sobre el "vigor de crecimiento" de plantas de maíz.

Par	Planta tratada	Planta no tratada	Diferencia d_i	Rango de $\|d_i\|$	+	-
1	19	13	6	4+	4	
2	18	17	1	1+	1	
3	20	10	10	6+	6	
4	20	8	12	8+	8	
5	17	13	4	3+	3	
6	16	9	7	5+	5	
7	17	17	0	0		
8	2	16	-14	9-		9
9	16	19	-3	2-		2
10	18	7	11	7+	7	
					T+ = 34	**T- = 11**

Hay un par que tiene diferencia cero, n = 10 – 1 = 9.

g) Decisión: en la Tabla 5 se puede observar que el valor crítico de T+, para una cola, α = 0,05 y n = 9, es igual a 37. Como la T+

calculada es igual a 34 no es mayor que la T+ crítica que es 37, no se rechaza H_0.

h) Conclusión: la aplicación de corriente eléctrica débil a plantas de maíz, no produce un incremento en el "vigor de crecimiento".

Ejemplo

- Se desea saber si un garrapaticida dado disminuye el grado de infestación de garrapatas en ganado bovino.

Se toman al azar 15 vacas infestadas con garrapatas y se evalúa su grado de infestación, utilizando puntajes del uno al ocho (ocho indica infestación masiva).

Se aplica el garrapaticida y se vuelve a evaluar el grado de infestación. Los datos y cálculos se muestran en el Cuadro 3.6.

a) Hipótesis de investigación: la aplicación de un garrapaticida dado disminuye el grado de infestación en ganado bovino.

b) Hipótesis estadísticas: H_0: $Md_D \geq 0$
H_1: $Md_D < 0$

Donde, M_D es la mediana de las diferencias (antes - después de la aplicación).

c) Estadístico de prueba: el estadístico de prueba "T" de la prueba de Wilcoxon es el adecuado, ya que se utiliza dos muestras relacionadas, las diferencias se pueden ordenar y la distribución de la variable subyacente es continua.

Este caso, se trata de una prueba de una cola (izquierda).

d) Nivel de significación: se utilizará un nivel de significación de 1% ($\alpha = 0{,}01$).

e) Región crítica o Área de rechazo: como se trata de una prueba de una cola (izquierda), se rechaza H_0 si T- calculada es menor que la T- tabulada en la Tabla 5, para $\alpha = 0{,}01$; en pruebas de una cola y "n" igual al número de pares cuyas diferencias sean distintas de cero.

f) Cálculos: en el Cuadro 3.6 se indica el resultado del experimento, así como los cálculos.

Cuadro 3.6. Efecto de un garrapaticida dado sobre el grado de infestación por garrapatas en ganado bovino.

Vaca	Después	Antes	Diferencia di	Diferencia |di|	Rango de |di|	+	-
1	2	6	4	4	5,5		5,5
2	3	8	5	5	8,5		8,5
3	1	8	7	7	13		13
4	5	7	2	2	2		2
5	3	3	0	0			
6	3	6	3	3	3,5		3,5
7	3	8	5	5	8,5		8,5
8	1	7	6	6	11,5		11,5
9	4	3	-1	1	1-	1	
10	2	5	3	3	3,5		3,5
11	2	7	5	5	8,5		8,5
12	5	5	0	0			
13	1	6	5	5	8,5		8,5
14	2	6	4	4	5,5		5,5
15	1	7	6	6	11,5		11,5
						T+ = 1	**T- = 90**

Hay dos ceros n = 15 – 2 = 13

g) Decisión: en la Tabla 5 se observa que la T- crítica es igual a 12, para una cola, para n = 13 y $\alpha = 0{,}01$. Como la T- calculada no es menor que 12, la T- tabulada, se rechaza H_0.

h) Conclusión: la aplicación del garrapaticida disminuye el grado de infestación en ganado bovino.

Discusión de los métodos para dos muestras relacionadas

Se ha presentado tres pruebas para el caso de dos muestras relacionadas.

Para la prueba de los signos y la de Wilcoxon, se requiere del supuesto de una distribución continua de la variable subyacente.

Cuando una o ambas condiciones en estudio han sido medidas en escala nominal, se debe usar la prueba de McNemar.

Si es posible la medición en escala ordinal, es decir, si la respuesta de uno de los miembros del par es mayor, igual o menor que la del otro, se puede usar la prueba de los signos.

Cuando la escala de medición para la diferencia entre los pares es ordinal, se puede utilizar la prueba de Wilcoxon.

Capítulo IV

Caso de dos muestras independientes

Con frecuencia no se puede usar una unidad experimental, parcela, animal o planta como su propio control, debido a la naturaleza misma de la unidad o de los tratamientos que se les aplique.

En este caso, se tendrá que usar dos muestras independientes.

Las muestras pueden ser dos muestras aleatorias, provenientes de una a dos poblaciones o pueden generarse por la asignación aleatoria de los tratamientos a los miembros de una muestra de origen arbitrario.

En general se usa la palabra “muestra” para grupos, condiciones, tratamientos, métodos, procesos u ocasiones.

Prueba de Fisher

La prueba de Fisher es particularmente útil para analizar datos medidos en escala nominal, cuando el tamaño de dos muestras independieres, en conjunto es pequeño (n < 30).

Metodología

Las observaciones de cada muestra se clasifican en una de dos posibles clases, las cuales deben ser mutuamente excluyentes. Para el manejo de los datos se prepara una tabla 2 x 2, tal como se indica a continuación:

Tabla 2 x 2

		Categorías		
		+	**-**	
Muestras	**Grupo I**	**A**	**B**	**A + B**
	Grupo II	**C**	**D**	**C + D**

En esta tabla, los grupos I y II son grupos independientes, como "experimental y control", "macho y hembra", entre otros. Los encabezamientos de las columnas, aquí marcados arbitrariamente como "+" y "-", pueden ser clasificaciones correspondientes a las propiedades mutuamente excluyentes que tiene los grupos. Estos pueden ser: sanos y enfermos, crecen y no crecen, positivos y negativos, si y no, entre otros.

La prueba de Fisher determina si la proporción de unidades experimentales (individuos, animales, plantas, entre otras) de un grupo que se clasifican en una categoría, es igual, mayor, o menor que la del otro grupo.

La probabilidad de observar un conjunto dado de frecuencias en una tabla 2 x 2, está dada por la distribución hipergeométrica. Los cálculos son sumamente tediosos, por lo cual, para efectuar pruebas de hipótesis se utilizan tablas (Tabla 6). En dichas tablas se encuentran valores críticos para pruebas de una cola, para $n < 30$, $A + B \leq 15$ y $C + D \leq 15$.

Debido a que la Tabla 6 es de uso más difícil que las anteriores, se explicará su uso en extenso.

Forma de usar la tabla 6

a) A partir de una tabla 2 x 2, se calcula los valores de A + B y C + D.

b) Si A + B no es mayor o igual (≥) que C + D, se debe reordenar la tabla, para que así lo sea (sin perder las identidades). Ejemplo:

	+	-
Grupo I	0 (A)	6 (B)
Grupo II	8 (C)	1 (D)

6 = A + B, no es mayor o igual que C + D = 9

se reordena:

	+	-
Grupo II	8 (A)	1 (B)
Grupo I	0 (C)	6 (D)

9 = A + B > C + D = 6

c) También se debe observar si A es menor o igual que B. Si no es así, se debe reordenar la tabla, para que así lo sea. Ejemplo (continuación al anterior):

	+		**-**	
Grupo II	8	(A)	1	(B)
Grupo I	0	(C)	6	(D)

A = 8
no menor o igual que
B = 1

se reordena:

	-		**+**	
Grupo II	1	(A)	8	(B)
Grupo I	6	(C)	0	(D)

A = 1 < B = 8
y
A + B = 9 > 6 = C + D

Ahora la tabla llena las condiciones requeridas en los puntos b y c, es decir,

$$A \leq B \text{ y } (A + B) \geq (C + D).$$

d) Se busca el valor observado de A + B = 9 (de la tabla reordenada) en la Tabla 6, bajo el encabezamiento “totales en el margen derecho”

e) Para valores de A + B = 9 y C + D = 6, se busca en la Tabla 6 el valor de C + D = 6 (de la tabla reordenada).

f) En la Tabla 6, para el valor de B = 8 se localiza el valor tabulado de D, en la parte derecha de la tabla bajo “Nivel de significación” especificado para pruebas de una cola (0,05; 0,025; 0,01 y 0,005).

Para pruebas de dos colas hay que duplicar la significación (2α). En el Cuadro 4.1 se muestra las equivalencias.

Si el valor de D observado es menor que el tabulado, se rechaza H_0 al nivel indicado.

Cuadro 4.1. Equivalencias de significación para pruebas de dos colas.

En la tabla (una cola)	0,05	0,025	0,01	0,005
Dos colas	0,10	0,05	0,02	0,01

Ejemplo

- Se desea determinar si cierta vacuna antiaftosa produce algún cambio en la protección de bovinos, contra la fiebre aftosa.

De un grupo de bovinos altamente susceptibles, se toman al azar 12 animales.

En forma aleatoria se dividen los 12 animales en dos grupos, uno de ocho animales y el otro de cuatro. Al grupo de ocho animales, se les denominará "vacunados" (V), se vacunan con la vacuna en prueba. Al otro grupo, de cuatro animales, se les denominará "no vacunados" (NV), no se vacunan.

Pasado un tiempo, se inoculan todos los 12 animales con el virus de la fiebre aftosa. Posteriormente, se observa a los animales, para determinar cuáles enfermaron (E) y cuáles no enfermaron (NE).

a) Hipótesis de investigación: la vacuna produce un cambio en la protección de bovinos, contra la fiebre aftosa.

b) Hipótesis estadísticas: $H_0: P_1 = P_2$
$H_1: P_1 \neq P_2$

Donde:

P_1 es la proporción de bovinos vacunados que no enfermaron (protegidos).

P_2 es la proporción de bovinos no vacunados que no enfermaron (protegidos).

c) Estadístico de prueba: se utiliza a D de la Prueba de Fisher. La prueba de Fisher es adecuada, ya que los datos se miden en escala nominal (enfermos y no enfermos) y n = 12, que es menor que 30.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0{,}05$).

e) Región crítica o Área de rechazo: debido a que la H_1 no indica dirección, ésta es una prueba de dos colas.

Se rechazará la H_0 si la D observada (con arreglo de la tabla si es necesario), es menor que la tabulada en la Tabla 6 para $\alpha = 0{,}025$.

f) Cálculos: a continuación se indica la tabla de doble entrada con los resultados del experimento, así como los cálculos necesarios de la prueba de una vacuna antiaftosa en bovinos.

	Enfermaron		**No enfermaron**		
Vacunados	0	(A)	8	(B)	**8 = A + B**
No vacunados	4	(C)	0	(D)	**4 = C + D**

Como A + B = 8 > 4 = C + D y A = 0 < 8 = B, no se tiene que reordenar la tabla.

g) Decisión: en la Tabla 6, para A + B = 8 y C + D = 4, se busca el valor de B = 8 y a su lado derecho, bajo el nivel de significación igual a 0,025, que corresponde a $\alpha = 0{,}05$ para pruebas de dos colas (ver Cuadro 4.1), se observa que D = 1.

Como la D calculada es igual a 0, menor que la D tabulada que es igual a 1, se rechaza H_0.

h) Conclusión: la vacuna produce un cambio en la protección de bovinos, contra la fiebre aftosa.

Ejemplo

- Se desea determinar si cierta vacuna antiaftosa protege a los bovinos, contra la fiebre aftosa.

De un grupo de bovinos altamente susceptibles, se toman al azar 12 animales. En forma aleatoria se dividen los 12 animales en dos grupos, uno de ocho animales y el otro de cuatro.

Al grupo de ocho animales, se les denominará “vacunados” (V), se vacunan con la vacuna en prueba. Al otro grupo, de cuatro animales, se les denominará “no vacunados” (NV), no se vacunan.

Pasado un tiempo, se inoculan todos los 12 animales con el virus de la fiebre aftosa. Posteriormente, se observa a los animales, para determinar cuáles enfermaron (E) y cuáles no enfermaron (NE).

a) Hipótesis de investigación: la vacuna es efectiva en la protección de bovinos, contra la fiebre aftosa.

b) Hipótesis estadísticas: $H_0: P_1 \leq P_2$
$H_1: P_1 > P_2$

Donde:

P_1 es la proporción de bovinos vacunados que no enfermaron (protegidos).

P_2 es la proporción de bovinos no vacunados que no enfermaron (protegidos).

c) Estadístico de prueba: se utiliza a D de la Prueba de Fisher. La prueba de Fisher es adecuada, ya que los datos se miden en escala nominal (enfermos y no enfermos) y n = 12, que es menor que 30.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0,05$).

e) Región crítica o Área de rechazo: ya que la H_1 indica dirección, ésta es una prueba de una cola (en este caso, cola derecha).

Se rechaza la H_0 si se cumplen las dos condiciones siguientes:

- Si $p_1 > p_2$, es decir, que la proporción de animales vacunados que no enferman (p_1) es mayor que la proporción de animales que no enferman (p_2).

- La D calculada es menor a la tabulada.

f) Cálculos: a continuación se indica la tabla de doble entrada con los resultados del experimento, así como los cálculos necesarios de la prueba de una vacuna antiaftosa en bovinos.

	Enfermaron		No enfermaron		
Vacunados	0	(A)	8	(B)	**8 = A + B**
No vacunados	4	(C)	0	(D)	**4 = C + D**

Como A + B = 8 > 4 = C + D y A = 0 < 8 = B, se observa directamente a D.

Donde:

$p_1 = B/(A + B) = 1$

$p_2 = D/(C + D) = 0$, luego $p_1 > p_2$

g) Decisión: en la Tabla 6, para A + B = 8 y C + D = 4, se busca el valor de B = 8 y a su lado derecho, directamente, bajo el encabezado del "Nivel de significación" igual a 0,05; se observa que D = 1.

Como $p_1 > p_2$ y la D calculada que es igual a 0, es menor que la tabulada que es igual a 1, se rechaza H_0.

h) Conclusión: la vacuna es efectiva en la protección de bovinos, contra la fiebre aftosa.

Ejemplo

- Existe una sustancia que se dice inhibir la germinación de semillas de caraota negra. Se desea determinar si esto es verdad.

Se siembran 20 semillas de caraota, cada una en un pote. En forma aleatoria se dividen los 20 potes en dos grupos de 10 potes cada uno.

Se inoculan 10 potes con la sustancia inhibidora y el resto de los potes se usan como controles (no inoculados).

La asignación de los tratamientos a los grupos se ha hecho en forma aleatoria.
Se espera un tiempo prudencial y se observa cada pote, para determinar si la semilla plantada en el pote germinó o no.

a) Hipótesis de investigación: la sustancia inhibe la germinación de semillas de caraota.

b) Hipótesis estadísticas: $H_0: P_1 \geq P_2$
$H_1: P_1 < P_2$

Donde:

P_1 es la proporción de semillas germinadas en potes inoculados.

P_2 es la proporción de semillas no germinadas en potes inoculados.

c) Estadístico de prueba: se utiliza la D en la Prueba de Fisher. Esta prueba es la adecuada, ya que los datos se miden en escala nominal (germina, no germina) y n = 20, que es menor que 30.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0,05$).

e) Región crítica o Área de rechazo: ya que la prueba es de una cola (cola izquierda), se rechazará H_0, si se cumplen las condiciones siguientes:

- Si $p_1 < p_2$, es decir, que la proporción de potes inoculados que germinan (p_1) es menor que la proporción de potes no inoculados que germinan (p_2).

- La D calculada es menor que la tabulada.

f) Cálculos: a continuación se indica la tabla de doble entrada con los resultados y los cálculos de la prueba de una sustancia inhibidora de la germinación de semillas de caraota.

	No germinaron		**Germinaron**		
Inoculados	8	(A)	2	(B)	**10 = A + B**
No inoculados	1	(C)	9	(D)	**10 = C + D**

Para la decisión se debe tener los valores de p_1 y p_2.

Donde:

p_1 = B/(A + B) = 2/10 = 0,2

p_2 = D/(C + D) = 9/10 = 0,9; $p_1 < p_2$

Además, se debe tener la tabla en forma adecuada para obtener la D calculada, es decir, se debe tener:

$$(A + B) \geq (C + D) \text{ y } A \leq B$$

Como (A + B) = 10 y (C + D) = 10, la primera condición se cumple, pero A = 8 no es menor que B = 2, por lo tanto, se debe reordenar la tabla:

	Germinaron		No germinaron		
Inoculados	2	(A)	8	(B)	**10 = A + B**
No inoculados	9	(C)	1	(D)	**10 = C + D**

Ahora, A = 2 < 8 = B

g) Decisión: como $p_1 = 0{,}2 < 0{,}9 = p_2$, se puede determinar la significación de D. En la Tabla 6, para (A + B) = 10, (C + D) = 10, B = 8 y bajo $\alpha = 0{,}05$ observamos que D = 3.

Como la D calculada es igual a 1, menor que la tabulada igual a 3, se rechaza H_0.

h) Conclusión: la sustancia inhibe la germinación de semillas de caraota.

Prueba de homogeneidad de Ji al cuadrado

Se utiliza la prueba de Ji al cuadrado cuando los datos consisten de frecuencias en categorías mutuamente excluyentes. El objetivo, así como en la prueba de Fisher, es determinar si existen diferencias entre grupos o entre tratamientos, respecto a las categorías. La categoría representa el valor de la variable, medida en escala nominal.

La hipótesis que corrientemente se pone a prueba supone que los dos grupos difieren respecto a alguna característica y, por lo tanto, respecto a las frecuencias relativas de cada grupo en cada categoría. Para probar dicha hipótesis se cuenta el número de casos de cada grupo en cada categoría y se determina si las distribuciones de frecuencias de los grupos difieren entre sí.

Por ejemplo, se puede probar si para dos razas de bovinos (grupos), las frecuencias de las categorías de una característica difieren. También se puede determinar si para dos variedades de un cultivo (grupos), las frecuencias de cinco categorías de una característica difieren.

En el caso de dos grupos con dos categorías cada uno, lo que se hace es comparar proporciones, en la forma idéntica a la prueba de Fisher.

Metodología

Para la prueba de la hipótesis se utiliza el estadístico X^2, donde:

$$X^2 = \sum_{i=1}^{f} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \qquad (1)$$

$i = 1,2,\ldots,f \quad j = 1,2,\ldots,c,$

Donde:

f: es el número de filas.

c: es el número de columnas.

O_{ij}: es la frecuencia observada en la celda que intersecta la i-ésima fila con la j-ésima columna.

E_{ij}: es la frecuencia esperada en la celda que intersecta la i-ésima fila con la j-ésima columna, (correspondiente a la O_{ij}).

Una fórmula más fácil de calcular y que da resultados idénticos a la fórmula (1), es la siguiente:

$$X^2 = \sum_{i=1}^{f} \sum_{j=1}^{c} \frac{O_{ij}^2}{E_{ij}} - n \qquad (2)$$

Donde, “n” es la frecuencia total.

Las fórmulas (1) y (2) son generales para el caso de dos o más muestras. En este caso, se reducirá simplemente haciendo que f = 2 (2 filas) y c (columnas) dependiendo del número de grupos o tratamientos.

En este caso, X^2 tiene (f – 1) x (c -1) = (2 -1) x (c -1) = 1 x (c -1) = (c -1) grados de libertad.

Para calcular la frecuencia esperada de cada celda (E_{ij}) se multiplica los totales marginales de cada celda (el total de la fila y el total de la columna) y al producto se divide por el gran total “n”.

Validez

La prueba es válida sin ningún valor esperado es menor que cinco y si no más de 20% son menores que uno.

Ejemplo

- Se desea determinar si hay diferencia en la abundancia de tres especies de plantas en dos potreros distintos.

De cada potrero se toma una muestra al azar y se cuenta el número de plantas de cada especie. En el Cuadro 4.2 se muestran los resultados del muestreo.

a) Hipótesis de investigación: la abundancia de las tres especies de plantas es igual para cada potrero.

b) Hipótesis estadísticas: H_0: Hay homogeneidad
H_1: Hay heterogeneidad

c) Estadístico de prueba: como la variable de interés es nominal (especie de plantas) y se presentan los datos en frecuencias, se puede usar la prueba de Ji al cuadrado.

Se utiliza el estadístico X^2, que se distribuye como Ji al cuadrado.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0{,}05$).

e) Región crítica o Área de rechazo: en el Cuadro 4.2 se observa que hay dos filas (f = 2) y tres columnas (c = 3), luego se rechazará H_0 si X^2 es mayor o igual que el valor de Ji al cuadrado en la Tabla 2, para $\alpha = 0{,}05$ y grados de libertad $(2 - 1)(3 - 1) = 2$, es decir, 5,99.

f) Cálculos: en el Cuadro 4.2 se presentan los datos con algunos cálculos.

Cuadro 4.2. Composición de las especies de plantas en dos potreros.

	Especie 1	Especie 2	Especie 3	Totales
Potrero 1	32	14	6	52
Potrero 2	12	22	9	43
Totales	**44**	**36**	**15**	**95**

Se calcula los valores esperados para cada celda (E_{ij}).

$$E_{11} = (52)(44)/95 = 24{,}08$$

$$E_{12} = (52)(36)/95 = 19{,}71$$

$$E_{13} = (52)(15)/95 = 8{,}21$$

$$E_{21} = (43)(44)/95 = 19{,}92$$

$$E_{22} = (43)(36)/95 = 16{,}29$$

$$E_{23} = (43)(15)/95 = 6{,}79$$

Usando la fórmula (2) se calculan los valores de O^2_{ij}/E_{ij} para cada celda:

$(32)^2/24{,}08 = 42{,}52$ $(14)^2/19{,}71 = 9{,}94$ $(6)^2/8{,}21 = 4{,}38$

$(12)^2/19{,}92 = 7{,}23$ $(22)^2/16{,}29 = 29{,}71$ $(9)^2/6{,}79 = 11{,}93$

Total: 105,71

$X^2 = 105{,}71 - 95{,}00 = 10{,}71$.

g) Decisión: como $X^2 = 10,71$ es mayor que 5,99; el valor tabulado, se rechaza H_0.

h) Validez:

- $n = 96 > 20$

- Todos los valores esperados son mayores que uno.

i) Conclusión: la abundancia de las tres especies no es igual para cada potrero.

Tablas 2 x 2

A propósito de la prueba de Fisher, este método se utiliza para analizar tablas 2 x 2 cuando las dos muestras independientes son pequeñas. Estas tablas también se pueden analizar por el método de Ji al cuadrado, siempre y cuando "n" sea mayor que 20 y todas las E_{ij} sean mayores que cinco.

Para el caso de una tabla 2 x 2, las fórmulas (1) y (2) se puede reducir a:

$$X^2 = \frac{n(AD - BC)^2}{(A+B)(C+D)(A+C)(B+D)} \qquad (2)$$

Donde, A, B, C, D y "n" tienen el mismo significado que en la Prueba de Fisher (Tabla 2 x 2), sin embargo, en este caso no existen las restricciones que hagan necesarios rearreglos de la tabla.

Con los datos en las tablas 2 x 2 se pueden efectuar las mismas pruebas de hipótesis que las planteadas para la prueba de Fisher.

Validez: la prueba es válida si ningún valor esperado es menor que cinco.

Ejemplo

- Para el ejemplo donde se determinó si cierta vacuna antiaftosa produce algún cambio en la protección de bovinos, contra la fiebre aftosa. Ahora se desea determinar si se produce algún cambio en la protección de los bovinos contra la fiebre aftosa.

a) Hipótesis de investigación: la vacuna produce un cambio en la protección de bovinos, contra la fiebre aftosa.

b) Hipótesis estadísticas: $H_0: P_1 = P_2$
$H_1: P_1 \neq P_2$

c) Estadístico de prueba: la variable de interés es nominal, se presentan los datos en frecuencias y n = 38, por lo tanto, se puede usar la prueba de Ji al cuadrado.

Se utilizará a X^2, el estadístico que se distribuye aproximadamente como Ji al cuadrado.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0{,}05$).

e) Región crítica o Área de rechazo: como se trata de una tabla 2 x 2, se tendrán dos filas y dos columnas, por lo tanto, habrán (f – 1)(c - 1) = 1 grados de libertad. Se rechazará H_0 si X^2

es mayor o igual a 3,84; obtenido de la Tabla 2 para α = 0,05 y un grado de libertad.

f) Cálculos: a continuación se presenta la tabla de doble entrada con los datos y algunos cálculos de la prueba de una vacuna antiaftosa en bovinos.

	Enfermaron		No enfermaron		Totales
Vacunados	2	(A)	20	(B)	**22**
No vacunados	15	(C)	1	(D)	**16**
Totales	**17**		**21**		**38**

Se usa la fórmula (2) y se obtiene:

$$X^2 = \frac{38[(2)(1)-(20)(15)]^2}{(22)(16)(17)(21)} = \frac{38(2-300)^2}{125.664} = \frac{3.374.552}{125.664} = 26,85$$

g) Decisión: como X^2 = 26,85 es mayor que 3,84; el valor tabulado de Ji al cuadrado para un grado de libertad y α = 0,05; se rechaza H_0.

h) Validez:

- n = 38 > 20

- Los valores esperados son 9,84; 12,16; 7,16 y 8,84; todos son mayores que cinco, por lo tanto, la prueba es válida.

h) Conclusión: la vacuna produce un cambio en la protección de bovinos contra la fiebre aftosa.

Ejemplo

- Para el ejemplo donde se determinó si cierta vacuna antiaftosa produce algún cambio en la protección de bovinos, contra la fiebre aftosa. Ahora se desea determinar si cierta vacuna antiaftosa protege a los bovinos contra la fiebre aftosa.

a) Hipótesis de investigación: la vacuna antiaftosa protege a los bovinos.

b) Hipótesis estadísticas: $H_0: P_1 \leq P_2$
$H_1: P_1 > P_2$

Donde:

P_1 es la proporción de bovinos vacunados que no enfermaron (protegidos).

P_2 es la proporción de bovinos no vacunados que no enfermaron (protegidos).

c) Estadístico de prueba: La variable de interés es nominal, se presentan los datos en frecuencias y n = 38 > 20, por lo tanto, se puede usar la prueba de Ji al cuadrado.

Se utilizará a X^2, el estadístico que se distribuye aproximadamente como Ji al cuadrado.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0{,}05$).

e) Región crítica o Área de rechazo: como se trata de una tabla 2 x 2, se tendrá dos filas y dos columnas, por lo tanto, habrán $(f - 1)(c - 1) = 1$ grados de libertad.

Ahora la prueba es de una cola (derecha), mientras que la Tabla 2 es para dos colas, por lo tanto, para $\alpha = 0{,}05$, se buscará en la tabla el valor crítico para $2\alpha = 2(0{,}05) = 0{,}10$ y se rechazará H_0 si X^2 es mayor o igual a 2,71 y, además, si la proporción de bovinos vacunados que no enfermaron (p_1) es mayor que la proporción de bovinos no vacunados que no enfermaron (p_2), es decir, $p_1 > p_2$.

f) Cálculos: a continuación se presenta la tabla de doble entrada con los datos y algunos cálculos de la prueba de una vacuna antiaftosa en bovinos.

	Enfermaron		No enfermaron		Totales
Vacunados	2	(A)	20	(B)	**22**
No vacunados	15	(C)	1	(D)	**16**
Totales	**17**		**21**		**38**

Se utilizará los resultados para X^2 del ejemplo anterior, es decir, $X^2 = 26{,}85$, además, $p_1 = 20/22 = 0{,}91$ y $p_2 = 1/16 = 0{,}06$.

g) Decisión: como $p_1 = 0{,}91 > 0{,}06 = p_2$ y $X^2 = 26{,}85 > 2{,}71$; se rechaza H_0.

h) Validez:

- $n = 38 > 20$

- Los valores esperados son 9,84; 12,16; 7,16 y 8,84; todos mayores que cinco, por lo tanto, la prueba es válida.

i) Conclusión: la vacuna es efectiva.

Ejemplo

- Existe una sustancia que se dice que inhibe la germinación de semillas de caraota negra. Se desea determinar si esto es verdad.

Se siembran 40 semillas de caraota, cada una en un pote. En forma aleatoria se divide los 40 potes en dos grupos, cada uno de 20 potes.

Se inoculan 20 potes con la sustancia inhibidora y el resto de los potes se usa como controles (no inoculados).

La asignación de los tratamientos a los grupos se ha hecho en forma aleatoria.

Se espera un tiempo prudencial y se observa cada pote, para determinar si la semilla en ella plantada germinó o no.

a) Hipótesis de investigación: la sustancia inhibe la germinación de semillas de caraota.

b) Hipótesis estadísticas: $H_0: P_1 \geq P_2$
$H_1: P_1 < P_2$

Donde:

P_1 es la proporción de semillas germinadas en potes inoculados.

P_2 es la proporción de semillas no germinadas en potes inoculados.

c) Estadístico de prueba: la variable de interés es nominal, se presentan los datos en frecuencias y n = 40, por lo tanto, se puede usar la prueba de Ji al cuadrado.

Se utilizará a X^2, el estadístico que se distribuye aproximadamente como Ji al cuadrado.

d) Nivel de significación: se utilizará un nivel de significación de 5% (α = 0,05).

e) Región crítica o Área de rechazo: como se trata de una tabla 2 x 2, se tendrán dos filas y dos columnas, por lo tanto, habrán (f – 1)(c – 1) = 1 grados de libertad.

Ahora la prueba es de una cola (izquierda), mientras que la Tabla 2 es para dos colas, por lo tanto, para α = 0,05, se buscará en la tabla el valor crítico de 2α = 2(0,05) = 0,10 y se rechazará H_0 si X^2 es mayor o igual a 2,71 y, además, si la proporción de potes inoculados que germinaron (p_1) es menor que la proporción de potes no inoculados que no germinaron (p_2), es decir, $p_1 < p_2$.

f) Cálculos: a continuación se indica la tabla de doble entrada con los resultados y los cálculos de la prueba de una sustancia inhibidora de la germinación de semillas de caraota.

	No germinaron		**Germinaron**		**Totales**
Inoculados	16	(A)	4	(B)	**20**
No inoculados	2	(C)	18	(D)	**20**
Totales	**18**		**22**		**40**

$$X^2 = \frac{40[(16)(18)-(4)(2)]^2}{(20)(20)(18)(22)} = \frac{40(288-8)^2}{158.400} = \frac{40(280)^2}{158.400} = \frac{3.136.000}{158.400} = 19{,}80$$

$p_1 = 4/20 = 0{,}20$ $p_2 = 18/20 = 0{,}90$

g) Decisión: como $p_1 = 0{,}20 < 0{,}90 = p_2$ y $X^2 = 19{,}80 > 2{,}71$; se rechaza H_0.

h) Validez: valores esperados: 9; 11; 9 y 1.

Donde, n = 40 > 20 y ningún valor esperado es menor que cinco, por lo tanto, la prueba es válida.

i) Conclusión: la sustancia inhibe la germinación de semillas de caraota.

Cuando se puede y debe utilizar la prueba de Ji al cuadrado

a) Caso de tablas 2 x 2:

- Cuando n > 30, se puede calcular X^2 con la formula (2).

- Cuando n está entre 20 y 40, se puede usar X^2 según la fórmula (2) siempre y cuando todas las frecuencias esperadas sean mayores o iguales a cinco, de lo contrario se debe utilizar la Prueba de Fisher.

- Cuando n < 20, se debe utilizar la prueba de Fisher en todos los casos.

b) Tablas con grados de libertad mayor que uno: se puede usar X^2 si menos de 20% de las celdas tienen frecuencia esperada menor que cinco y ninguna menor que uno.

Prueba de la mediana

La prueba de la mediana es un procedimiento que se utiliza para probar si dos grupos independientes (o dos tratamientos) provienen de poblaciones con la misma mediana. Se puede usar esta prueba si las observaciones están en dos grupos independientes y se han medido, por lo menos, en escala ordinal.

Metodología

Se ordenan todos los datos de menor a mayor, sin tomar en cuenta a que grupo pertenecen, pero sin perder su identidad. Se determina la mediana del conjunto combinado de observaciones. Luego se clasifican las observaciones de cada grupo según si son mayores que la mediana o, si son iguales o menores que ella.

Se elabora una tabla 2 x 2, la cual se puede analizar por la prueba de Fisher o por la de Ji al cuadrado.

Ejemplo

- Se desea determinar si las leches de dos marcas distintas difieren en su contenido medio de proteína. El método que se ha utilizado para analizar el contenido sólo permite dar puntuaciones a cada leche. Las puntuaciones van de 100 a 400.

Se toma una muestra al azar de 11 cartones de la marca “A” y una muestra aleatoria de ocho cartones de la marca “B”. En el Cuadro 4.3 se muestran los datos.

a) Hipótesis de investigación: las leches difieren respecto a su contenido proteico.

b) Hipótesis estadísticas: H_0: $Md_A = Md_B$
H_1: $Md_A \neq Md_B$

c) Estadístico de prueba: la prueba de la mediana es adecuada, ya que los datos están medidos en escala ordinal y se desea determinar si proviene de la misma población o de dos poblaciones con mediana común.

Como n < 20, se debe utilizar la prueba de Fisher para analizar la tabla 2 x 2.

d) Nivel de significación: se utilizará un nivel de significación de 5% (α = 0,05).

e) Región crítica o Área de rechazo: si el valor de D calculado (después de rearreglos de la tabla, si es necesario), es menor que él tabulado en la Tabla 6, se rechaza H_0.

f) Cálculos: en el Cuadro 4.3 se muestra el resultado del muestreo de los cartones de leche “A” y “B”.

Cuadro 4.3. Contenido proteico (expresado en puntajes) de dos muestras de leche.

Marca "A"	Marca "B"
320	348
301	300
310	284
331	292
290	300
340	303
320	302
310	304
301	
283	
290	

Sin perder la identidad de cada marca, se procede a ordenar los datos, de menor a mayor y a calcular la mediana.

Cuadro 4.4. Contenido proteico ordenado de dos muestras de leche y cálculo de la mediana.

Proteína	Marca
283	A
284	B
290	A
290	A
292	B
300	B
300	B
301	A
301	A
302	B (Mediana)
303	B
304	B
310	A
310	A
320	A
320	A
331	A
340	A
348	B

La mediana, resultó ser igual a 302.

Se clasifica cada cartón de leche, de acuerdo con la marca y la posición del contenido proteico en relación con el contenido mediano, tal como se observa en la tabla 2 x 2 siguiente:

	Marca "A"	**Marca "B"**	**Totales**
Igual o menor que la mediana	5 (A)	5 (B)	**10 = A + B**
Mayor que la mediana	6 C)	3 (D)	**9 = C + D**
Totales	**11**	**8**	**19**

Se nota que (A + B) = 10 > 9 = (C + D), que A = 5 y que B = 5, (A = B), por lo tanto, no hay que rearreglar la tabla. La D calculada = 3.

g) Decisión: en la Tabla 6, con (A + B) = 10; (C + D) = 9 y B = 5; se busca la D crítica, que para $\alpha = 0,05$ (dos colas), resulta ser igual a cero (bajo la columna del nivel de significación encabezada con 0,025). Ver el Cuadro 4.1, de las equivalencias para pruebas de dos colas.

Como la D calculada = 3, no es menor que la D tabulada = 0, no se rechaza la H_0.

h) Conclusión: las leches no difieren respecto a su contenido proteico.

Ejemplo

- Para el ejemplo donde se determinó si las leches de dos marcas distintas difieren en su contenido medio de proteínas.

Ahora se desea determinar si el contenido proteico de la leche "A" es superior al de la leche "B".

a) Hipótesis de investigación: el contenido proteico de la leche "A" es superior al de la leche "B".

b) Hipótesis estadísticas: H_0: $Md_A \leq Md_B$
H_1: $Md_A > Md_B$

c) Estadístico de prueba: la prueba de la mediana es adecuada, ya que los datos están medidos en escala ordinal y se desea determinar si proviene de la misma población o de dos poblaciones con mediana común.

Como n < 20, se debe utilizar la prueba de Fisher para analizar la Tabla 2 x 2.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0{,}05$).

e) Región crítica o Área de rechazo: si $p_1 > p_2$, donde p_1 es la proporción de los cartones de leche "A" con contenido proteico mayor que la mediana y p_2 es la proporción de los cartones de leche "B" con contenido proteico mayor que la mediana; y si la D calculada es menor que la tabulada en la Tabla 6, se rechazará la H_0.

f) Cálculos: como $p_1 = 6/11 = 0{,}55$; $p_2 = 3/8 = 0{,}38$ ($p_1 > p_2$). La D calculada = 3.

g) Decisión: en la Tabla 6, con (A + B) = 10; (C + D) = 9 y B = 5; la D crítica, para $\alpha = 0{,}05$ es igual a cero. A pesar de que $p_1 > p_2$, pero como la D calculada = 3, no es menor que la D tabulada = 0, no se rechaza la hipótesis nula.

h) Conclusión: el contenido proteico de la leche "A" no es superior al de la leche "B".

Ejemplo

- Se tiene una sustancia (S) que controla la presencia de nematodos en cierta planta. Esta sustancia es barata y efectiva, pero produce un cuadro carencial de microelementos, que se manifiesta por un cambio de coloración de las hojas. Este efecto se mide de acuerdo a la apariencia de la planta, en puntuaciones que van de cero a 10 (10 es el caso más grave).

Acaba de salir al mercado otra sustancia (Z) que también controla los nematodos y se dice que su efecto de inducir la carencia de microelementos es menor que la sustancia S.

La sustancia Z es más cara que la sustancia S.

Se efectuó un experimento con 20 plantas, las cuales se han dividido al azar en dos grupos de 10 plantas cada uno. Se asignan al azar las dos sustancias a cada grupo. Desafortunadamente, por accidente, se pierde una de las plantas del grupo de la sustancia S y dos de la sustancia Z.

Se inoculan las plantas con nematodos, se aplican las sustancias, y se observa el efecto de cada una. En el Cuadro 4.5 se muestran los resultados.

¿Reduce la sustancia Z el efecto carencial, comparado con la sustancia S?

a) Hipótesis de investigación: el efecto carencial de la sustancia Z es menor que la de la sustancia S.

b) Hipótesis estadísticas: $H_0: Md_Z \geq Md_S$
$H_1: Md_Z < Md_S$

c) Estadístico de prueba: la prueba de la mediana es adecuada, ya que los datos están medidos en escala ordinal y se desea determinar si proviene de la misma población o de dos poblaciones con mediana común.

Como n < 20, se debe utilizar la prueba de Fisher para analizar la Tabla 2 x 2.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0{,}05$).

e) Región crítica o Área de rechazo: se rechaza H_0 si $p_Z < p_S$, donde p_S es la proporción de plantas tratadas con la sustancia S que tienen el puntaje de "carencia" mayor que la mediana y p_Z es la proporción de plantas tratadas con la sustancia Z que tienen el puntaje de "carencia" mayor que la mediana.

Además, la D calculada debe ser menor que la tabulada en la Tabla 6.

f) Cálculos: en el Cuadro 4.5 se presentan los datos del experimento.

En el Cuadro 4.6 se indican las puntuaciones ordenadas, de modo que se pueda calcular la mediana.

Cuadro 4.5. Puntuaciones del efecto carencial de dos sustancias nematicidas.

Sustancia	Puntaje							
S	8	9	9	7	8	7	6	5	4
Z	2	3	8	4	5	5	4	2	

Cuadro 4.6. Puntuaciones ordenadas del efecto carencial de dos sustancias nematicidas y cálculo de la mediana.

Puntaje	Sustancia
2	Z
2	Z
3	Z
4	S
4	Z
4	Z
5	S
5	Z
5	**Z Md**
6	S
7	S
7	S
8	Z
8	S
8	S
9	S
9	S

La mediana resultó ser igual a cinco.

A continuación se presenta la Tabla 2 x 2 donde se clasifican las plantas según la sustancia y la posición del efecto carencial, en relación con su mediana.

	Sustancia "S"	Sustancia "Z"	Totales
Igual o menor que la mediana	2 (A)	7 (B)	**9 = A + B**
Mayor que la mediana	7 (C)	1 (D)	**8 = C + D**
Totales	**9**	**8**	**17 = n**

$p_S = 7/9 = 0{,}78$ $p_Z = 1/8 = 0{,}13$

Se nota que (A + B) = 9 > 8 = (C + D) y que A = 2 < 7 = B; se puede obtener directamente la D calculada, que es igual a 1.

g) Decisión: se ve que $p_Z = 0{,}13 < 0{,}78 = p_S$

En la Tabla 6 se busca la D crítica para (A + B) = 9; (C + D) = 8 y B = 7; bajo α = 0,05; para una prueba de una cola. La D tabulada = 2.

Como la D calculada = 1 y ésta es menor que 2 la D tabulada, se rechaza H_0.

h) Conclusión: el efecto carencial de la sustancia Z es menor que el de la sustancia S.

Pueba U de Mann-Whitney

Cuando se ha logrado, por lo menos, una medida en la escala ordinal, la prueba U de Man-Whitney se puede usar para probar si dos grupos o tratamientos independientes han sido muestreados de una misma población.

Esta prueba constituye la alternativa más útil frente a la prueba paramétrica "t", cuando no se desean hacer o no se logran los supuestos que ella exige, o si la escala de medida de las observaciones es más débil que la de intervalo.

Fundamentos

Al igual que para la prueba de la mediana se ordenan las observaciones de ambos grupos en conjunto, de menor a mayor, pero sin perder la identidad del grupo al que pertenecen.

Se asignan los rangos. En el caso de empates, se usará el promedio de los rangos (rango medio), que se les hubiera asignado, si no hubiera habido empates.

Al lado de las observaciones se coloca una columna para identificar el grupo y otra para los rangos.

En otro cuadro se colocan las observaciones ordenadas, una columna para los rangos del grupo 1 y otra para los del grupo 2.

Se suman los rangos del grupo 1, la suma se denomina R_1

Se suman los rangos del grupo 2, la suma se denomina R_2

Se calcula:

$$U_1 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2 \qquad (1)$$

$$U_2 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \qquad (2)$$

Para verificar se usa:

$$U_1 = n_1 n_2 - U_2 \qquad (3)$$

o

$$U_1 + U_2 = n_1 n_2 \qquad (4)$$

Se calculan y comparan los valores de U_1 y U_2 con los valores de U y U’ tabuladas en la Tabla 7.

Empates

Un problema que presentan a menudo los datos es que varias observaciones pueden ser exactamente iguales. Siendo la distribución subyacente de naturaleza continua, en teoría no debería haber empates (si se tuvieran instrumentos de medida suficientemente sensibles), en la práctica ocurren con bastante frecuencia.

Si bien los empates dentro de un grupo no constituyen ningún problema (la U no queda afectada), se presenta alguna dificultad cuando los empates ocurren entre dos o más observaciones de los distintos grupos.

Si dos o más observaciones están empatadas, a cada una se le asignará el promedio de los rangos (rango medio) que se les hubiera asignado si no hubieran estado empatados.

Ejemplo: si dos observaciones están empatadas en el segundo y tercer lugar, se tendría para cada una el promedio de dos y tres, es decir 2,5 y, por lo tanto, todas las observaciones tendrían los rangos 1, 2,5; 2,5; 4, 5, entre otros.

Se dispone de una fórmula para corregir el efecto de los empates.

Desafortunadamente, el empleo de esta fórmula es un tanto complicado.

Se recomienda efectuar la corrección para empates sólo cuando su proporción es muy alta y cuando la U no corregida se aproxima demasiado al nivel de significación previamente establecido; para ello se debe consultar el libro de Siegel y Castellan (1990), sobre Estadísticas no paramétricas.

Ejemplo

- Con el objeto de determinar si existe diferencia alguna entre los coeficientes de digestibilidad de la materia seca de ensilado de maíz, en ovejos y novillos, se lleva a efecto un experimento.

En dicho experimento se utilizan siete ovejos (grupo 1) y siete novillos (grupo 2). Para cada uno se determina el coeficiente de digestibilidad de la materia seca de un ensilado de maíz.

En el Cuadro 4.7 se muestra el resultado del experimento.

En vista de que el coeficiente de digestibilidad es una razón (expresada en porcentajes), no se quiere asumir normalidad de los datos ni homogeneidad de varianzas entre grupos, para realizar una prueba de "t".

a) Hipótesis de investigación: los coeficientes de digestibilidad difieren entre ovejos (O) y novillos (N).

b) Hipótesis estadísticas: H_0: $Md_1 = Md_2$ o H_0: $Md_O = Md_N$
H_1: $Md_1 \neq Md_2$ o H_1: $Md_O \neq Md_N$

c) Estadístico de prueba: se utiliza la prueba U de Mann-Whitney, ya que se trata de dos muestras independientes y las observaciones se han reducido a escala ordinal.

El estadístico de prueba es U.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0{,}05$).

e) Región crítica o Área de rechazo: en la Tabla 7, para $n_O = 7$; $n_N = 7$ y $\alpha = 0{,}05$; dos colas, U = 8 y U' = 41.

Si U_O es menor que 8 o si U_N es mayor que 41, se rechazará Ho.

f) Cálculos: en el Cuadro 4.7 se muestran los coeficientes de digestibilidad de las dos especies (grupos).

En el Cuadro 4.8 se muestran los coeficientes de digestibilidad con sus rangos.

En el Cuadro 4.9 se muestran los datos con sus rangos, separados por el grupo al cual pertenecen, así como el total y la media de los rangos, por especie.

Cuadro 4.7. Coeficientes de digestibilidad de la materia seca de ensilado de maíz en ovejos y novillos.

Ovejos (O) Grupo 1	Novillos (N) Grupo 2
57,8	64,2
56,2	58,7
61,9	63,1
54,4	62,5
53,6	59,8
56,4	59,2
53,2	61,4

Cuadro 4.8. Coeficientes de digestibilidad de la materia seca de ensilado de maíz en ovejos y novillos, ordenados.

Coeficiente de digestibilidad	Especie	Rangos
53,2	Ovejos	1
53,6	Ovejos	2
54,4	Ovejos	3
56,2	Ovejos	4
56,4	Ovejos	5
57,8	Ovejos	6
58,7	Novillos	7
59,2	Novillos	8
59,8	Novillos	9
61,4	Novillos	10
61,9	Ovejos	11
62,5	Novillos	12
63,1	Novillos	13
64,2	Novillos	14

Cuadro 4.9. Coeficientes de digestibilidad de la materia seca de ensilado de maíz en ovejos y novillos, rangos por especie y sumas de rangos.

Coeficiente de digestibilidad	Ovejos (O)	Novillos (N)
53,2	1	
53,6	2	
54,4	3	
56,2	4	
56,4	5	
57,8	6	
58,7		7
59,2		8
59,8		9
61,4		10
61,9	11	
62,5		12
63,1		13
64,2		14
	$R_O = 32$	$R_N = 73$
	$n_O = 7$	$n_N = 7$
	$\overline{R}_O = 4{,}57$	$\overline{R}_N = 10{,}43$

Ahora se calcula U_O usando la fórmula (1):

$$U_O = 7(7) + \frac{7(8)}{2} - 73 = 49 + 28 - 73 = 4$$

y U_N usando la fórmula (2):

$$U_N = 7(7) + \frac{7(8)}{2} - 32 = 49 + 28 - 32 = 45$$

Se verifica: $U_O + U_N = 4 + 45 = 49 = n_O(n_N)$ y
$U_O = 7(7) - 45 = 49 - 45 = 4.$

g) Decisión: como $U_O = 4$ es menor que $8 = U$ y $U_N = 45$ es mayor que $41 = U'$, se rechaza H_0.

h) Conclusión: los coeficientes de digestibilidad difieren entre ovejos y novillos.

Ejemplo

- Con el objeto de determinar si existe diferencia alguna entre los coeficientes de digestibilidad de la materia seca de ensilado de maíz, en ovejos y novillos, se lleva a efecto un experimento, descrito en el ejemplo anterior.

Ahora, se desea determinar si el coeficiente de digestibilidad es menor en los ovejos que en los novillos.

a) Hipótesis de investigación: el coeficiente de digestibilidad es menor en los ovejos (muestra 1) que en los novillos (muestra 2).

b) Hipótesis estadísticas: $H_0: Md_1 \geq Md_2$ o $H_0: Md_O \geq Md_N$
$H_1: Md_2 < Md_1$ o $H_1: Md_O < Md_N$

c) Estadístico de prueba: se utiliza la Prueba U de Mann-Whitney, ya que se trata de dos muestras independientes y las observaciones se han medido en escala ordinal. El estadístico de prueba es U.

d) Nivel de significación: se utilizará un nivel de significación de 1% (α = 0,01).

e) Región crítica o Área de rechazo: en la Tabla 7, para n_O= 7; n_N = 7 y α = 0,01; una cola, U = 6 y U' = 43.

Si $\overline{R}_O$ es menor que $\overline{R}_N$ y si U_O es menor que U, se rechazará H_0.

f) Cálculos: en el Cuadro 4.9 del ejemplo anterior se tiene que U_O = 4; $\overline{R}_O$ = 4,57 y $\overline{R}_N$ = 10,43.

g) Decisión: como $\overline{R}_O$ = 4,57 es menor que 10,43 = $\overline{R}_N$ y U_O = 4 es menor que 6 = U, por lo tanto, se rechaza H_0.

h) Conclusión: el coeficiente de digestibilidad es menor en los ovejos que en los novillos.

Ejemplo

- Se efectúa un experimento para determinar el coeficiente de digestibilidad de la materia seca de dos muestras de millo, en ovejos. La primera muestra es de un ensilaje normal (N) y la segunda de un ensilaje que ha tenido un proceso de fermentación, aparentemente anormal (A) (dos procesos).

Para la primera muestra se tiene n_1 = 7 ovejos y para la segunda n_2 = 5 ovejos.

Se desea saber si el coeficiente de digestibilidad es mayor en la muestra “normal” que en la muestra “anormal”.

En el Cuadro 4.10 se muestran los resultados del experimento.

a) Hipótesis de investigación: el coeficiente de digestibilidad de la muestra (1) (“normal”) es mayor que él de la muestra (2) (“anormal”).

b) Hipótesis estadísticas: H_0: $Md_1 \leq Md_2$ o H_0: $Md_N \leq Md_A$
H_1: $Md_1 > Md_2$ o H_1: $Md_N > Md_A$

c) Estadístico de prueba: se utiliza la prueba U de Mann-Whitney, ya que se trata de dos muestras independientes y las observaciones han sido medidas en escala de razón, las cuales se han reducido a escala ordinal. El estadístico de prueba es U.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0,05$).

e) Región crítica o Área de rechazo: en la Tabla 7, para $n_N = 7$; $n_A = 5$ y $\alpha = 0,05$; una cola, U = 6 y U’ = 29.

Si $\overline{R}_N$ es mayor que $\overline{R}_A$ y si U_A es mayor que U’, se rechazará H_0.

f) Cálculos: en el Cuadro 4.10 se muestran los coeficientes de digestibilidad de los dos ensilados (procesos).

En el Cuadro 4.11 se muestran los datos con sus rangos.

En el Cuadro 4.12 se muestran los datos con sus rangos separados por el proceso al cual pertenecen, así como el total y la media de los rangos, por condición.

Cuadro 4.10. Coeficientes de digestibilidad de la materia seca de ensilaje "normal" y "anormal" en ovejos.

Normal %	Anormal %
57,8	60,1
56,2	49,9
61,9	51,6
54,4	55,3
53,6	50,4
56,4	
53,2	

Cuadro 4.11. Coeficientes de digestibilidad de la materia seca de ensilaje "normal" y "anormal" en ovejos ordenados y sus rangos.

Coeficiente de digestibilidad	Proceso	Rango
49,9	Anormal	1
50,4	Anormal	2
51,6	Anormal	3
53,2	Normal	4
53,6	Normal	5
54,4	Normal	6
55,3	Anormal	7
56,2	Normal	8
56,4	Normal	9
57,8	Normal	10
60,1	Anormal	11
61,9	Normal	12

Cuadro 4.12. Coeficientes de digestibilidad de la materia seca de ensilaje "normal" y "anormal" en ovejos, rangos por procesos y sumas de rangos.

Coeficiente de digestibilidad	Normal	Anormal
49,9		1
50,4		2
51,6		3
53,2	4	
53,6	5	
54,4	6	
55,3		7
56,2	8	
56,4	9	
57,8	10	
60,1		11
61,9	12	
	$R_O = 54$	$R_N = 24$
	$n_O = 7$	$n_N = 5$
	$\overline{R}_O = 7{,}71$	$\overline{R}_N = 4{,}80$

Ahora se calcula U_N usando la fórmula (1):

$$U_N = 7(5) + \frac{5(6)}{2} - 24 = 35 + 15 - 24 = 26$$

y U_A usando la fórmula (2):

$$U_A = 7(5) + \frac{7(8)}{2} - 54 = 35 + 28 - 54 = 9$$

Se verifica: $U_N + U_A = 35 = n_N(n_A)$ y

$U_N = 7(5) – 9 = 26$

g) Decisión: como $\overline{R}_N = 7,71$ es mayor que 4,80 = $\overline{R}_A$, pero $U_N = 26$, no es mayor que U’ = 29: por lo tanto, no se rechaza H_0.

h) Conclusión: el coeficiente de digestibilidad de la muestra “normal” no es mayor que él de la muestra “anormal”.

Discusión de los métodos para dos muestras independientes

Se han presentado cuatro métodos útiles para probar la significación de la diferencia entre dos muestras independientes. Todas las pruebas no paramétricas para dos muestras independientes determinan si es probable que provengan de la misma población.

Entre las pruebas incluidas en este capítulo, unas son más o menos sensibles que otras a distintos tipos de diferencias.

Si se desea probar si dos muestras provienen de poblaciones que difieren en la tendencia central, la prueba de Fisher, la prueba de la mediana y la prueba U de Mann-Whitney son las más adecuadas

Por otra parte, si se desea detectar cualquier tipo de diferencia entre las poblaciones subyacentes, se debe usar la prueba de Ji al cuadrado y cuando se trata de una tabla 2 x 2 con n pequeña, se debe usar la prueba de Fisher.

Para escoger entre los métodos que detectan diferencias en las tendencias centrales, se debe decidir tomando en cuenta la escala de medida con que se ha medido la variable bajo prueba.

Si la variable en consideración se ha medido en escala ordinal, la prueba U de Mann-Whitney es la mejor.

Si se ha medido la variable en forma tal que sólo es factible dicotomizar las observaciones y clasificarlas por encima y por debajo de la mediana combinada, entonces, la prueba de la mediana será la más indicada.

Capítulo V

Caso de k muestras relacionadas

Hasta ahora se han presentado métodos para los casos de una y dos muestras. En este capítulo y el siguiente se presentarán métodos para más de dos muestras. En este momento se estudiará el caso de k muestras (tratamientos, grupos, condiciones, entre otros) relacionadas ($k > 2$).

En general, frecuentemente se diseñan experimentos, de tal manera que más de dos muestras, tratamientos, o condiciones se puedan estudiar simultáneamente. En este caso, se desea determinar si hay diferencia total entre las k muestras, tratamientos o condiciones.

La metodología paramétrica, para determinar si varias muestras provienen de poblaciones idénticas, está representada por la prueba de F en el análisis de la varianza. Para esta prueba se asume que las observaciones son independientes y normalmente distribuidas, homogeneidad de varianzas entre poblaciones y, además, aditividad (en el caso de Bloque al azar). La medida de las observaciones debe ser por lo menos en escala de intervalo.

Si alguno de los supuestos anteriores falta o no se puede sustentar, se debe recurrir a métodos no paramétricos.

Existen dos diseños básicos para comparar k grupos. En el primero, las k muestras de igual tamaño son igualadas, de acuerdo con cierto criterio.

En algunos casos, la igualación se hace comparando los mismos individuos o casos bajo todas las k condiciones.

En otras instancias, cada uno de los n grupos puede ser medido bajo todas las k condiciones.

Para este diseño, las pruebas estadísticas de k muestras relacionadas son las adecuadas y serán objeto de estudio en este capítulo.

El segundo diseño, trata de con k muestras independientes, no necesariamente de igual tamaño. El análisis de este tipo de diseño se presentará en el Capítulo VI.

En el caso de k muestras relacionadas, los datos se pueden presentar de tres maneras.

a) Conjuntos igualados: ésta es una extensión de “pares igualados” a “k-tuples igualados”. Habrá bloques y dentro de cada uno se encontrarán los k-tuples. Los k-tuples se observan simplemente o se les aplica un tratamiento o condición diferente a cada uno, y después se observan.

b) k ocasiones: en este caso habrá n bloques y dentro de cada bloque un sujeto (unidad, animal, planta, entre otros) que se observa en k ocasiones, circunstancias, otros, o se le aplica k tratamientos distintos y luego se observan.

c) Panel de evaluadores: en esta forma de presentación de los datos, cada uno de los n bloques está representado por un evaluador ("observador"), que evalúa ("observa"), independientemente de los demás n – 1 observadores, a cada una de las k muestras. El segundo ejemplo de este capítulo es de este tipo.

Prueba Q de Cochran

La prueba de McNemar para dos muestras relacionadas puede extenderse para k muestras relacionadas, lo cual da lugar a la prueba Q de Cochran.

Metodología

Si las observaciones son sobre una variable que se ha medido en escala nominal y si se pueden arreglar en una tabla de doble entrada consistente en "n" filas y "k" columnas, entonces, se puede probar si las proporciones de respuestas de un tipo particular son iguales para cada columna.

Si se ha demostrado que la hipótesis nula ($P_1 = P_2 = \ldots = P_k$) es verdadera, es decir, que si no hay diferencia en un tipo de respuesta (positivo por ejemplo), bajo cada condición y si el número de filas no es muy pequeño, entonces

$$Q = \frac{[k-1]\,[k\sum_{j=1}^{k} G_j^2 - (\sum_{j=1}^{k} G_j)^2]}{k\sum_{i=1}^{n} H_i - \sum_{i=1}^{n} H_i^2} \qquad (1)$$

se distribuye como Ji al cuadrado con (k – 1) grados de libertad, es decir, número de columnas menos 1, con la condición de que el número de filas no sea menor que 4 ($n \geq 4$) y que $nk > 24$.

En esta prueba en particular, la medida se hace en la escala nominal, con sólo dos valores (variable dicotómica). Si a uno de los valores de la variable se representa con el número uno y al otro con cero, entonces en la fórmula (1):

G_j: es el total de la j-ésima columna (j = 1, 2, …, k).

H_i: es el total de la i-ésima fila (i = 1, 2 , …, n).

k: es el número de columnas (condiciones, métodos, otros).

n: es el número de filas (unidades, sujetos, animales, plantas, entre otros).

Ejemplo

- Bajo ciertas condiciones específicas, se desea saber si tres métodos (c = 3) de diagnóstico de la brucelosis bovina concuerdan o no en sus resultados.

Se toma una muestra aleatoria de 14 sueros bovinos y se determina por cada método su positividad. En el Cuadro 5.1 se muestran los resultados.

Se ha utilizado el número uno para indicar una respuesta positiva y el número cero para una respuesta negativa.

a) Hipótesis de investigación: las respuestas son iguales con los tres métodos.

b) Hipótesis estadísticas: H_0: $P_1 = P_2 = P_3$
H_1: no todas la P son iguales.

c) Estadístico de prueba: como se trata de tres muestras dependientes (el mismo suero para cada método), con observaciones medidas en escala nominal (dicotómicas), se puede usar la prueba Q de Cochran.

d) Nivel de significación: se usará un nivel de significación de 1% ($\alpha = 0{,}01$).

e) Región crítica o Área de rechazo: si Q es mayor o igual que Ji al cuadrado con $k - 1 = 2$ grados de libertad y $\alpha = 0{,}01$ en la Tabla 2, es decir, 9,21; se rechazará H_0.

f) Cálculos: en el Cuadro 5.1 se muestran los datos y algunos cálculos parciales. En este caso, $n = 14$ y $k = 3$.

Se nota que en el margen derecho se encuentran los totales de filas H_i y sus cuadrados H_i^2.

En la parte inferior de la tabla se encuentra, en la penúltima fila, la suma de las columnas G_j y en la última sus cuadrados G_j^2.

Cuadro 5.1. Resultados de tres métodos de diagnóstico de la brucelosis bovina.

Suero Nº	Métodos			H_i	H_i^2
	1	**2**	**3**		
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0
5	1	0	0	1	1
6	1	0	0	1	1
7	1	0	0	1	1
8	1	1	0	2	4
9	1	1	0	2	4
10	1	1	0	2	4
11	1	1	0	2	4
12	1	1	1	3	9
13	1	1	1	3	9
14	1	1	1	3	9
G_j:	**10**	**7**	**3**	**20**	**46**
G_j^2:	**100**	**49**	**9**		

$$\sum_{i=1}^{14} H_i = 20 \qquad \sum_{i=1}^{14} H_i^2 = 46$$

$$\sum_{j=1}^{3} G_j = 10 + 7 + 3 = 20 \qquad \sum_{j=1}^{3} G_j^2 = 100 + 49 + 9 = 158$$

Se calcula Q usando la fórmula (1):

$$Q = \frac{[3\text{-}1][3(158)\text{-}(20)^2]}{3\ (20) - 46} = \frac{2\ (474 - 400)}{60 - 46} = \frac{148}{14} = 10{,}57$$

g) Decisión: como Q = 10,57 es mayor que el valor crítico de Ji al cuadrado (9,21), se rechaza H_0.

h) Conclusión: las respuestas no son todas iguales con los tres métodos.

Análisis de la varianza de Friedman

Cuando k muestras igualadas tienen sus observaciones medidas, por lo menos en escala ordinal, se puede utilizar el análisis de la varianza de dos criterios de Friedman para probar si las k muestras han sido obtenidas de poblaciones diferentes.

Metodología

Se colocan los datos en una tabla de doble entrada de "n" filas y "k" columnas. Las filas (bloques) representan los distintos sujetos, unidades, animales, plantas, otros, y las columnas a las diferentes condiciones, (tratamientos, grupos, muestras, otros).

Si por ejemplo se estudian las observaciones de cada unidad, bajo cada una de los k grupos o condiciones, cada fila indicará el puntaje de cada unidad bajo una de los k grupos o condiciones.

Se deben ordenar las observaciones de cada fila (bloque) y transformarlas en rangos (de uno a k).

Para cada condición (tratamiento, grupo, muestra) se asignan los rangos, se suman y se denomina a ese total R_j para la j-ésima columna. Para la prueba de Friedman se usará el estadístico de prueba:

$$X_F^2 = \frac{12}{nk\,(k+1)} \sum_{j=1}^{k} R_j^2 - 3(n)(k+1) \qquad (1)$$

Donde:

n: es el número de filas o bloques (sujetos, animales, plantas, otros).

k: es el número de columnas o condiciones (tratamientos, grupos, muestras, otros).

R_j^2: es el cuadrado de la suma de los rangos de la j-ésima columna (R_j) (j = 1, 2,…, k)

X_F^2: se distribuye aproximadamente como Ji al cuadrado, con k - 1 grados de libertad, cuando k y n son grandes.

Para valores pequeños de k y n, en la Tabla 9 se indican los valores críticos de X_F^2, para k = 3; n de 3 a 13; para k = 4; n de 2 a 8 y para k = 5; n de 3 a 5.

Para valores de n mayores se puede usar la aproximación a la Ji al cuadrado, en la fila etiquetada como ∞.

Empates

Si una variable está distribuida de manera continua, entonces, la probabilidad de un empate es cero. Sin embargo, las puntuaciones empatadas ocurren con frecuencia.

Tales puntuaciones casi invariablemente son un reflejo de la carencia de sensibilidad de los instrumentos de medición, esto es, de la inhabilidad de los mismos, para distinguir las pequeñas diferencias que existen entre las observaciones, que se registran consecuentemente como empates.

Si dos o más observaciones, dentro de un bloque (fila) tienen el mismo valor, entonces se le debe asignar a cada una el rango medio de ellas.

Por ejemplo, si se tiene dentro de un bloque las observaciones 5, 6, 6, 7; se les asignará a las observaciones empatadas en 6 los valores 2,5 a cada una, quedando el bloque con rangos 1; 2,5; 2,5 y 4.

Los empates afectan el valor de X_F^2 y, por lo tanto, la significación de la prueba. Si hay muchos empates, se deberá utilizar un factor de corrección, el cual se podrá encontrar en Siegel y Castellan (1998).

Es de hacer notar que si hay pocos empates, la corrección es mínima y no vale la pena efectuarla, ya que es sumamente tediosa. Se debe notar que a menos que el resultado esté muy cerca del valor de significación, no vale la pena corregir.

Ejemplo

- Se selecciona al azar a nueve dueños de casa-quinta para participar en un experimento, con el fin de determinar si hay diferencias en la predilección de tres variedades de grama.

Se le pide a cada dueño que seleccionara tres áreas más o menos iguales y que sembrara en cada una de ellas, al azar, una de las variedades de grama. Al final de un tiempo especificado se pide a cada dueño que le asignará puntuaciones de predilección, del uno al cinco; a cada variedad, tomando en cuenta el costo, mantenimiento, belleza, otros. La puntuación cinco indica el óptimo de preferencia. En el Cuadro 5.2 se muestran los resultados del experimento.

a) Hipótesis de investigación: las tres variedades de grama son igualmente preferidas.

b) Hipótesis estadísticas: H_0: No hay diferencia entre variedades
H_1: No todas las variedades son iguales

c) Estadístico de prueba: cada dueño forma un bloque en el cual se observan las tres variedades de grama. Se tiene un diseño de bloques (dueños) y grupos (tratamientos o condiciones), aquí variedades, dentro de cada bloque, es decir, un diseño de dos vías con tres muestras dependientes.

Las observaciones se hacen en escala ordinal.

La prueba adecuada es la de Friedman.

El estadístico de prueba es X_F^2, se calcula según la fórmula (1).

d) Nivel de significación: se utilizará un nivel de significación de 5% (α = 0,05).

e) Región crítica o Área de rechazo: si el valor de X_F^2 es mayor o igual que el tabulado en la Tabla 9, para k = 3; n = 9 y α = 0,05; es decir, 6,22; se rechazará H_0.

f) Cálculos: en el Cuadro 5.2 se presentan los datos del experimento y en el Cuadro 5.3 se muestran los rangos de los datos y algunos cálculos parciales.

Cuadro 5.2. Puntuaciones de tres variedades de grama para nueve dueños de casa.

Dueño	Variedad de grama		
	A	B	C
1	1	3	4
2	4	2	5
3	3	1	4
4	4	3	5
5	1	2	3
6	4	3	5
7	3	2	5
8	2	5	4
9	2	1	3

Cuadro 5.3. Rangos de tres variedades de grama para nueve dueños de casa.

Dueño	Variedad de grama		
	A	B	C
1	1	2	3
2	2	1	3
3	2	1	3
4	2	1	3
5	1	2	3
6	2	1	3
7	2	1	3
8	1	3	2
9	2	1	3
R_j	15	13	26
$\overline{R}_j$	1,667	1,444	2,889
R_j^2	225	169	676

Donde:

R_j es la suma de los rangos de cada variedad.

$\overline{R}_j$ es la media de los rangos de cada variedad.

R_j^2 es el cuadrado de la suma de los rangos de cada variedad.

$$\sum_{j=1}^{3} R_j^2 = 1.070$$

En este caso k = 3 y n = 9; luego aplicando la fórmula (1), se tendrá:

$$X_F^2 = \frac{12}{9(3)(4)}\,1.070 - 3(9)(4) = \frac{1.070}{9} - 108 = 118{,}89 - 108 = 10{,}89$$

g) Decisión: como X_F^2 = 10,89 es mayor que X_F^2 tabulado (6,22), se rechaza H_0.

h) Conclusión: las tres variedades de grama no son igualmente preferidas.

Comparaciones entre k muestras relacionadas

Cuando hay diferencias significativas entre grupos o condiciones, esto no indica cual grupo es diferente de los demás o entre cuales grupos o condiciones hay diferencias.

Así, que si hay diferencias significativas (y sólo si las hay), se puede usar un procedimiento denominado “Comparaciones múltiples entre grupos o condiciones”, para los cuales difieren entre sí.

Además, si sólo interesa comparar un grupo o condición “control”, con los demás, se puede utilizar el método de “Comparaciones de los grupos o condiciones, con un control”.

a) Comparaciones múltiples entre k condiciones o grupos relacionados

Lo primero que se debe hacer es determinar el valor absoluto de la diferencia entre el rango medio ($\bar{R}_j$) de cada grupo o condición con él de cada una de los otros grupos o condiciones: ($|\bar{R}_u - \bar{R}_v|$), dónde u es diferente de v.

Luego, se comparan las diferencias con:

$$DS_{\#c,\alpha} = z_{\#c,\alpha}\sqrt{\frac{k\,(k+1)}{6\,n}} \qquad (1)$$

Donde, k es el número de grupos o condiciones y n el número de filas o bloques, #c es el número de posibles comparaciones entre grupos o condiciones [k(k-1)/2] y $z_{\#c,\alpha}$ está tabulada en la Tabla 10.

Si $|\overline{R}_u - \overline{R}_v| \geq DS_{\#c,\alpha}$, se rechaza H_0: no hay diferencia entre la condición u y la condición v, a favor de H_1: hay diferencia entre la condición u y la condición v, al nivel α.

Ejemplo

- En el ejemplo donde se seleccionó a nueve dueños de casa-quinta para determinar si había diferencias en la predilección de tres variedades de grama, se rechaza la Hipótesis nula y se concluye que hay diferencia entre las tres variedades de grama. Ahora se desea determinar entre cuales las hay, se utilizará α = 0,05.

Se tiene que $\overline{R}_A$= 1,667; $\overline{R}_B$= 1,444; $\overline{R}_C$= 2,889; y así se tendrá:

$|\overline{R}_A - \overline{R}_B|$ = |1,667 – 1,444| = 0,223
$|\overline{R}_A - \overline{R}_C|$ = |1,667 – 2,889| = 1,222
$|\overline{R}_B - \overline{R}_C|$ = |1,444 – 2,889| = 1,445

Ahora se tiene que k = 3 y n = 9, luego #c = 3(3 – 1)/2 = 3, que es el número de comparaciones posibles, por lo tanto, se tendrá $z_{3,0,05}$, que en la Tabla 10 resulta ser 2,394 y

$$DS_{3;0,05} = z_{3;0,05}\sqrt{\frac{3(3+1)}{6(9)}} = 2,394\sqrt{\frac{12}{54}} = 2,394\sqrt{0,222} = 2,394(0,471) = 1,128$$

Así que se puede concluir que hay diferencia entre las gramas A y C y entre las gramas B y C.

b) Comparaciones de las condiciones con un control

Cuando se desean hacer comparaciones entre grupos o condiciones y una en particular, que se denominará “control”; después de encontrar diferencias significativas entre grupos o condiciones, con la prueba de Friedman, se puede utilizar el método siguiente:

Esta vez sólo se comparará cada grupo o condición con el “control” que se denota con la letra t (T). Existen, por lo tanto, k – 1 comparaciones.

Se determina el valor absoluto de la diferencia entre el rango medio ($\bar{R}v$) de cada grupo o condición con él del grupo o condición control: ($\bar{R}_T$), dónde v es diferente de T. Luego, se comparan las diferencias con:

$$\mathrm{DS}_{\#\mathrm{c},\alpha} = Z_{\#c,\alpha}\sqrt{\frac{k(k+1)}{6n}} \qquad (2)$$

Donde, k es el número de grupos o condiciones y n el número de filas o bloques, #c es el número de posibles comparaciones entre el control y las demás grupos o condiciones (k - 1) y $z_{\#c,\alpha}$ está tabulada en la Tabla 10.

Si $|\bar{R}_T - \bar{R}_v| \geq \mathrm{DS}_{\#c,\alpha}$, se rechaza H_0: no hay diferencia entre el control (T) y el grupo o condición v, a favor de H_1: hay diferencia entre el control (T) y la condición v, al nivel α.

Ejemplo

- En el ejemplo donde se seleccionaron a nueve dueños de casa-quinta para determinar si había diferencias en la predilección de tres variedades de grama, se rechaza la Hipótesis nula y se concluye que hay diferencia entre las tres variedades de grama. Ahora se desea determinar si hay diferencia entre la grama A, que será nuestro control (T) y las demás. Se denotará al control como grupo o condición T, se utilizará α = 0,05.

Se tiene que $\bar{R}_T$ = 1,667; $\bar{R}_B$ = 1,444; $\bar{R}_C$ = 2,889; y así se tendrá:

$|\bar{R}_T - \bar{R}_B|$ = |1,667 – 1,444| = 0,223

$|\bar{R}_T - \bar{R}_C|$ = |1,667 – 2,889| = 1,222

Ahora se tiene que k = 3 y n = 9, luego #c = (3 – 1) = 2, que es el número de comparaciones posibles, por lo tanto, se tendrá $z_{2;0,05}$; que en la Tabla 10 resulta ser 2,241 y

$$DS_{2;0,05} = Z_{2;0,05}\sqrt{\frac{3(3+1)}{6(9)}} = 2,241\sqrt{\frac{12}{54}} = 2,241\sqrt{0,222} = 2,241(0,471) = 1,056$$

Así que se puede concluir que sólo hay diferencia entre la grama control (A) y la grama C.

Coeficiente de concordancia W de Kendall

Este coeficiente mide la concordancia o acuerdo entre jueces en su juicio acerca del ordenamiento de k condiciones, grupos, tratamientos, entre otros.

Este coeficiente tiene una relación directa con la prueba de Friedman y tiene valores entre cero y uno. Cero para una perfecta discordancia y uno para una perfecta concordancia, donde:

$$W = \frac{X_F^2}{n(k-1)} \qquad (1)$$

con la misma notación que en la prueba de Friedman.

Nota: no es necesario que X_F^2 sea significativa.

Ejemplo

- Para el ejemplo donde se seleccionó a nueve dueños de casa-quinta, para determinar si había diferencias en la predilección de tres variedades de grama, de la sección anterior, donde n = 9; k = 3 y X_F^2 = 10,889:

$$W = \frac{10{,}889}{9(3-1)} = \frac{10{,}889}{9(2)} = \frac{10{,}889}{18} = 0{,}605$$

La significación de W es la misma que para la prueba de Friedman, con la hipótesis nula, ahora de que no hay concordancia entre los nueve dueños de casa-quinta, en relación con su predilección para las tres variedades de grama y la alterna de que la hay.

Como la anterior hipótesis es rechazada a un nivel de 5%, se concluye que si hay concordancia o acuerdo entre los dueños de casa-quinta, con respecto a sus preferencias para las tres variedades de grama.

Prueba de Page para alternativas ordenadas

En el caso de la prueba de Friedman se estuvo probando la hipótesis nula de que no había diferencia entre los grupos o condiciones, contra la hipótesis alterna de que no todos los grupos o condiciones son iguales. En esa ocasión se rechaza la hipótesis nula de igualdad, a favor de la alterna de que no todas son iguales, es decir, de que por lo menos había diferencia entre dos grupos o condiciones.

Si se desea una hipótesis alterna de orden de magnitud, ya sea creciente o decreciente en los rangos de las condiciones, entonces "La prueba de Page para alternativas ordenadas" es la adecuada.

Esta prueba evalúa la hipótesis que los grupos o condiciones producen igual resultado (H_0) contra la hipótesis alterna (H_1), donde las condiciones están ordenadas en orden creciente o decreciente.

Metodología

Los datos van en una tabla de doble entrada de "n" filas y "k" columnas, donde las filas representan los bloques, dentro de cada bloque se aplican las k condiciones (grupos, tratamientos, entre otros).

Los grupos o condiciones deben estar previamente ordenados de menor a mayor o a la inversa, según la hipótesis de investigación que se desea comprobar.

En cada celda (parcela) se coloca el resultado o efecto del grupo o condición correspondiente. Dentro de cada bloque se asignan los rangos a cada uno de los resultados, de menor a mayor, tal como en la prueba de Friedman. Para cada una de los k grupos o condiciones se suma los rangos (ΣR) y se colocan las sumas en la última fila.

Debajo de esa última fila se coloca el orden de los grupos o condiciones previamente ordenados (Y), de menor a mayor o viceversa, según la hipótesis de investigación.

A continuación, para cada una de las k condiciones, se multiplican las ΣR por sus respectivas Y y se obtiene YΣR para cada condición. Se suman las YΣR y así se obtiene el estadístico L de Page en:

Orden ascendente (incremento): $L = 1R_1 + 2R_2 + \dots + kR_k$ (1)

u

Orden descendente (decremento). $L = kR_1 + (k\text{-}1)R_2 + \dots + 1R_k$ (2)

Se rechazará

H_0: $Md_1 = Md_2 = \dots = Md_{k-1} = Md_k$, a favor de:

H_1: $Md1 \leq Md_2 \leq \dots \leq Md_k$, si la hipótesis de investigación es de incremento, o

Se rechazará

H_0: $Md_1 = Md_2 = \dots = Md_{k-1} = Md_k$, a favor de:

H_1: $Md_1 \geq Md_2 \geq \dots \geq Md_k$, si la hipótesis de investigación es de decremento, cuando la L de Page calculada (L) es mayor que la tabulada, para una significación dada α, n bloques y k condiciones: $L_{\alpha;\, n;\, k}$ en la Tabla 11.

Empates

Si una variable está distribuida de manera continua, entonces la probabilidad de un empate es cero. Sin embargo, las puntuaciones empatadas ocurren con frecuencia. Tales puntuaciones casi invariablemente son un reflejo de la carencia de sensibilidad de los instrumentos de medición, esto es, de la inhabilidad de los mismos para distinguir las pequeñas diferencias que existen entre las observaciones que se registran consecuentemente como empates.

Si dos o más observaciones, dentro de un bloque (fila) tienen el mismo valor, entonces se le debe asignar a cada una el rango medio de ellas.

Por ejemplo, si se tiene que dentro de un bloque las observaciones 5, 6, 6, 7; se les asignará a las observaciones empatadas en 6 los valores 2,5 a cada una, quedando el bloque con rangos 1; 2,5; 2,5 y 4.

Ejemplo

- Suponga que se ha propuesto el uso de un preparado para embellecer una variedad dada de grama. Se indica que a mayor dosis, habrá un mayor incremento en la belleza de la grama, pero el fabricante no desea indicar previamente, la cantidad de cada dosis, sólo se especifica que es creciente, de “muy poco” (1), “poco” (2), “algo” (3) y “bastante” (4).

Para comprobar la bondad del preparado, se tiene un terreno que se ha dividido en seis bloques de cuatro parcelas cada uno. Se siembra la variedad de grama y se aplica la cantidad creciente del preparado a cada una de las parcelas del bloque.

Después de un tiempo determinado, se observa la grama de las parcelas de cada bloque y se evalúa su belleza, asignando valores del uno al cinco, siendo cinco como máxima belleza. Se asume que a mayor dosis, mayor belleza.

Tal como se hace en la prueba de Friedman, se ordenan de menor a mayor los valores de belleza y se asigna a cada uno un rango, dentro de cada bloque.

a) Hipótesis de investigación: a mayor dosis mayor belleza.

b) Hipótesis estadísticas: H_0: $Md_1 = Md_2 = Md_3 = Md_4$
H_1: $Md_1 \leq Md_2 \leq Md_3 \leq Md_4$

c) Estadístico de prueba: en cada bloque se observa la belleza de la grama de cada parcela. Se tiene un diseño de bloques y tratamientos dentro de cada bloque, es decir, un diseño de dos vías con cuatro muestras dependientes. Pero a diferencia de la prueba de Friedman, no se desea sólo determinar si hay diferencia entre los tratamientos (dosis), sino que también si los efectos producen incremento en las respuestas de belleza.

Las observaciones se hacen en escala ordinal.

La prueba adecuada es la de Prueba de Page.

El estadístico de prueba es L, el cual se calcula según la fórmula (1).

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0,05$).

e) Región crítica o Área de rechazo: para n = 6; k = 4 y $\alpha = 0,05$; si L calculada es mayor que $L_{0,05;6;4}$ en la Tabla 11, es decir, 163, se rechazará H_0.

En el Cuadro 5.4 se muestran los datos del experimento y en el Cuadro 5.5 se muestran los rangos de los datos y los cálculos.

Cuadro 5.4. Puntuaciones de cuatro dosis del preparado para seis bloques de terreno.

Bloques	Dosis			
	I	II	III	IV
1	3	5	4	2
2	1	3	5	4
3	2	3	4	5
4	1	3	4	5
5	2	4	3	5
6	1	2	3	4

Cuadro 5.5. Rangos de las puntuaciones de cuatro dosis del preparado para seis bloques de terreno.

Bloques	Dosis			
	I	II	III	IV
1	2	4	3	1
2	1	2	4	3
3	1	2	3	4
4	1	2	3	4
5	1	3	2	4
6	1	2	3	4
ΣR	7	15	18	20
Y	1	2	3	4
$Y\Sigma R$	7	30	54	80

$L = 7 + 30 + 54 + 80 = 171$

g) Decisión: como L = 171 es mayor que 163 = $L_{0,05;6;4}$ en la Tabla 13, se rechaza H_0.

h) Conclusión: a mayor dosis, mayor belleza.

Ejemplo

- Suponga que se desea evaluar una droga que reduce el grado de agresividad en perros. Se plantea que a mayor dosis de la droga, se produce un menor grado de agresividad.

Para el efecto se ha utilizado seis camadas con tres perros en cada una. Los perros de cada camada se igualaron con base a ciertas características.

A cada perro, de cada camada, se le suministra una de las tres dosis de la droga al azar. Las dosis de la droga consisten en 0,5; 1,0 y 1,5 mg.

El grado de agresividad se evalúa en una escala del uno al cuatro, siendo la puntuación de cuatro para el máximo grado de agresividad.

a) Hipótesis de investigación: a mayor dosis de la droga se produce menor agresividad.

b) Hipótesis estadísticas: H_0: $Md_1 = Md_2 = Md_3$
H_1: $Md_1 \geq Md_2 \geq Md_3$

c) Estadístico de prueba: en cada camada se observa la agresividad de cada perro. Se tiene un diseño de bloques y tratamientos dentro de cada bloque, es decir, un diseño de dos vías con tres muestras dependientes.

Pero a diferencia de la prueba de Friedman, no se desea sólo determinar si hay diferencia entre los tratamientos (dosis), sino que también si los efectos producen un decremento en las respuestas de agresividad.

Las observaciones se hacen en escala ordinal.

La prueba adecuada es la de Prueba de Page.

El estadístico de prueba es L, el cual se calcula según la fórmula (2).

d) Nivel de significación: se utilizará un nivel de significación de 1% ($\alpha = 0{,}01$).

e) Región crítica o Área de rechazo: para $n = 6$; $k = 3$ y $\alpha = 0{,}01$; si L calculada es mayor que $L_{0,01;6;3}$ en la Tabla 11, es decir, 81, se rechazará H_0.

En el Cuadro 5.6 se muestran los puntajes de agresividad para cada dosis de la droga y en el Cuadro 5.7 se muestran los rangos de los puntajes de agresividad para cada dosis de la droga y los cálculos.

Cuadro 5.6. Puntuaciones de tres dosis de la droga para seis camadas de perros.

Bloques	Dosis		
	0,5	1,0	1,5
1	4	2	1
2	3	2	1
3	4	1	2
4	3	2	1
5	3	4	2
6	1	2	3

Cuadro 5.7. Rangos de las puntuaciones de tres dosis de la droga para seis camadas de perros.

Bloques	Dosis		
	0,5	1,0	1,5
1	3	2	1
2	3	2	1
3	3	1	2
4	3	2	1
5	2	3	1
6	1	2	3
ΣR	15	12	9
Y	3	2	1
ΣR	45	24	9

L = 45 +24 +9 = 78

Nota: se puede ver ahora que las Y en la tabla van de tres a uno, es decir, decreciente, ya que así lo estipula la hipótesis de investigación y, por lo tanto, la hipótesis alterna.

g) Decisión: como L = 78 no es mayor que 81 = $L_{0,01;6;3}$ en la Tabla 11, no se rechaza H_0.

h) Conclusión: una mayor dosis de la droga no produce menor agresividad.

Discusión de los métodos para k muestras relacionadas

Se han presentado tres pruebas para el caso de k muestras relacionadas. La primera, la prueba Q de Cochran es la adecuada cuando la medida de los datos se hace en escala nominal. En el caso de que la medida de los datos se haga en escala ordinal, se tienen dos pruebas: la Prueba de Friedman y Page. Cuando sólo se quiera determinar si hay diferencias entre condiciones, se utilizará la Prueba de Friedman. En el caso de que además se quiera determinar si hay un incremento o decremento en las respuestas a las condiciones, se debe utilizar la Prueba de Page.

Capítulo VI

Caso de k muestras independientes

Hasta ahora se han indicados métodos para los casos de una y dos muestras. En este capítulo se presentaran métodos para más de dos muestras. Ahora se estudiará el caso de $k > 2$ muestras independientes.

En general, frecuentemente se diseñan experimentos, de tal manera que más de dos muestras, tratamientos, o condiciones se puedan estudiar simultáneamente. En este caso, se desea determinar si hay diferencia total entre las k muestras, tratamientos o condiciones.

La metodología paramétrica para determinar si varias muestras provienen de poblaciones idénticas está representada por la prueba de F en el análisis de la varianza. Para esta prueba se asume que las observaciones son independientes, normalmente distribuidas y homogeneidad de varianzas entre poblaciones. La medida de las observaciones debe ser, por lo menos, en escala de intervalo.

Si alguno de los supuestos anteriores falta o no se pueden sustentar, se debe recurrir a métodos no paramétricos. Para este diseño, las pruebas estadísticas de k muestras independientes, no necesariamente de igual tamaño. Son las adecuadas y serán objeto de estudio en este capítulo.

Prueba de homogeneidad de Ji al cuadrado

Cuando las observaciones se han medido en escala nominal y se presentan como frecuencias en categorías mutuamente excluyentes, la prueba de Ji al cuadrado puede determinar diferencias entre k muestras independientes. Esta prueba es una extensión del caso de dos muestras independientes.

Metodología

Para la prueba de la hipótesis se utilizará el estadístico X^2, que ya se ha utilizado en el Capítulo 4, donde:

$$X^2 = \sum_{i=1}^{f} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \qquad (1)$$

i = 1,2,…,f
j = 1,2,…,c,

Donde:

f: es el número de filas,

c: es el número de columnas,

O_{ij}: es la frecuencia observada en la celda que intersecta la i-ésima fila con la j-ésima columna,

E_{ij}: es la frecuencia esperada en la celda que intersecta la i-ésima fila con la j-ésima columna, (correspondiente a la O_{ij}).

Una fórmula alterna, más fácil de calcular y con idénticos resultados, es la siguiente:

$$X^2 = \sum_{i=1}^{f} \sum_{j=1}^{c} \frac{O_{ij}^2}{E_{ij}} - n \qquad (2)$$

Donde, "n" es la frecuencia total.

Las fórmulas (1) y (2) son generales para el caso de dos o más muestras.

Ahora, f = k y se tiene que f > 2.

Ejemplo

- Se desea determinar si hay diferencia en la abundancia de tres especies de plantas en tres potreros distintos.

De cada potrero se toma una muestra al azar y se cuenta el número de plantas de cada especie. En el Cuadro 6.1 se muestran los resultados del muestreo.

a) Hipótesis de investigación: la abundancia de las tres especies de plantas es distinta para cada uno de los tres potreros (heterogeneidad entre potreros).

b) Hipótesis estadísticas: H_0: Hay homogeneidad
H_1: Hay heterogeneidad

c) Estadístico de prueba: como la variable de interés es nominal (especie de planta) y se presentan los datos en frecuencias, se puede usar la prueba de Ji al cuadrado.

Se utilizará X^2 el estadístico de prueba que se distribuye aproximadamente como Ji al cuadrado con (f – 1)(c – 1) grados de libertad.

d) Nivel de significación: se utilizará un nivel de significación de 1% (α = 0,01).

e) Región crítica o Área de rechazo: en el Cuadro 6.1 se observa que hay tres filas (f = 3) y tres columnas (c = 3), por lo tanto, se rechazará H_0 si X^2 es mayor o igual que el valor de Ji al cuadrado en la Tabla 2, para (3 – 1)(3 – 1) = 2 x 2 = 4 grados de libertad y α = 0,01; es decir, 13,28.

f) Cálculos: en el Cuadro 6.1 se presentan los datos con algunos cálculos parciales.

Cuadro 6.1. Composición de las especies de plantas en muestras de tres potreros.

Potreros	Especie 1	Especie 2	Especie 3	Totales
1	32	14	6	52
2	12	22	9	43
3	22	19	8	49
Totales	**66**	**55**	**2 3**	**144**

Para el cálculo de X^2 se utilizará la fórmula (2)

Primero se calculará los valores esperados para cada celda (E_{ij}):

E_{11} = 52(66)/144 = 23,83
E_{12} = 52(55)/144 = 19,86
E_{13} = 52(23)/144 = 8,31
E_{21} = 43(66)/144 = 19,71
E_{22} = 43(55)/144 = 16,42
E_{23} = 43(23)/144 = 6,87
E_{31} = 49(66)/144 = 22.46
E_{32} = 49(55)/144 = 18,72
E_{33} = 49(23)/144 = 7,83

Ahora se calcula los $\dfrac{O_{ij}^2}{E_{ij}}$:

$(32)^2/23,83 = 42,87$
$(14)^2/19,86 = 9,87$
$(6)^2/8,31 = 4,33$
$(12)^2/19,71 = 7,31$
$(22)^2/16,42 = 29,48$
$(9)^2/6,87 = 11,79$
$(22)^2/22,46 = 21,55$
$(19)^2/18,72 = 19,28$
$(8)^2/7,83 = 8,17$

$$\sum_{i=1}^{f}\sum_{j=1}^{c}\frac{O_{ij}^2}{E_{ij}} = 154,65 \qquad \text{y}$$

$$X^2 = \sum_{i=1}^{f} \sum_{j=1}^{c} \frac{O_{ij}^2}{E_{ij}} - n = 154{,}65 - 144 = 10{,}65$$

g) Decisión: como $X^2 = 10{,}65$ no es mayor que 13,20; el valor tabulado de Ji al cuadrado, no se rechaza H_0 de homogeneidad.

h) Conclusión: la abundancia de las tres especies de plantas no es distinta para cada uno de los tres potreros (hay homogeneidad entre potreros).

Análisis de la varianza de Kruskal-Wallis

El método de análisis de la varianza para un criterio de clasificación, por rangos de Kruskal-Wallis es extremadamente útil para decidir si k muestras independientes provienen de poblaciones distintas. Se asume que la variable en estudio es continua. Se tiene que medir dicha variable, por lo menos, en escala ordinal.

Metodología

Las observaciones de todo el conjunto (n observaciones) se ordenan y se reemplazan por sus rangos, pero sin perder la identidad del grupo al cual pertenecen.

Los rangos se colocan en una tabla de k columnas. Cada columna representa a uno de los k grupos (tratamientos, condiciones, otros) y está formada por n_j observaciones.

Se suman los rangos de cada columna (muestra), cada suma se denomina R_j.

El estadístico de prueba es:

$$H = \frac{12}{n(n+1)} \sum_{j=1}^{k} R_j^2 / n_j - 3(n+1) \qquad (1)$$

Donde:

k: es el número de muestras (grupos, tratamientos, condiciones, entre otros).

n_j: es el número de observaciones de la j-ésima muestra (j = 1, 2, …, k).

n: es el número total de observaciones ($\sum_{j=1}^{k} n_j$).

R_j^2: es la suma de los rangos de la j-ésima columna (muestra), elevada al cuadrado.

Las distribuciones de H para k = 3 y n_j de uno a cinco, están en la Tabla 12. Cuando k = 3 y alguna n_j > 5 o k > 3, H se distribuye aproximadamente como Ji al cuadrado, con k – 1 grados de libertad.

Empates

Cuando se tiene dos o más observaciones empatadas (sin tomar en cuenta el grupo o condición), a cada observación se le da el rango medio (promedio de los rangos) en los cuales ocurrieron los empates.

Si más de 25% de las observaciones están empatadas, se debe aplicar un factor de corrección; de lo contrario, la diferencia entre corregir y no corregir es despreciable.

En Siegel y Castellán (1990) se puede encontrar el factor de corrección.

Ejemplo

- En un corral hay 14 novillos de tres razas distintas. Se desea determinar si hay diferencias de peso entre razas.

No habiendo una romana disponible, se procede a determinar los pesos "al ojo", ordenando en una fila los animales, de menor a mayor peso. En el Cuadro 6.2 se presentan los rangos para los pesos de los novillos de las tres razas especificadas.

a) Hipótesis de investigación: los pesos de los novillos son distintos para las tres razas.

b) Hipótesis estadísticas: H_0: $F_1(X) = F_2(X) = F_3(X)$
H_1: no todas las F(X) son iguales.

c) Estadístico de prueba: en este ejemplo se tiene tres muestras independientes de pesos (variable continua), medidas en escala ordinal, por lo cual se puede utilizar la prueba de Kruskal-Wallis. El estadístico de prueba es H.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0,05$).

e) Región crítica o Área de rechazo: para $n_1 = 5$; $n_2 = 4$ y $n_3 = 5$, y $\alpha = 0,05$; se busca en la Tabla 12 bajo la fila 5; 5; 4: el valor crítico de H = 5,64.

Si la H calculada es mayor que 5,64 se rechaza la H_0.

f) Cálculos: en el Cuadro 6.2 se muestran los datos, así como algunos cálculos parciales.

Cuadro 6.2. El rango de pesos de catorce novillos de tres razas distintas.

	Raza A	Raza B	Raza C	
	11	8	6	
	4	9	5	
	12	7	3	
	13	14	2	
	10		1	
R_j	50	38	17	Suma
n_j	5	4	5	14
$\overline{R}_j$	10,00	9,50	3,40	
R_j^2	2.500	1.444	289	
R_j^2 / n_j	500,00	361,00	57,80	**918,80**

$$H = [\frac{12}{14(14+1)}]\,[918,80] - 3(14+1) = \frac{11.025,60}{210} - 45 = 52,50 - 45 = 7,50$$

g) Decisión: como la H calculada (7,50) es mayor que la H tabulada (5,64), se rechaza H_0.

h) Conclusión: los pesos de los novillos son distintos para las tres razas.

Comparaciones entre k muestras independientes

Cuando hay diferencias significativas entre grupos o condiciones, esto no indica cual grupo o condición es diferente de los demás o entre cuales condiciones hay diferencias. Así, que si hay diferencias significativas (y sólo si las hay), se puede usar un procedimiento denominado "Comparaciones múltiples entre condiciones", para ver que grupos difieren entre sí. Además, si sólo interesa comparar una condición "control o testigo", con las demás, se puede utilizar el método de "Comparaciones de las condiciones, con un control o testigo".

a) Comparaciones múltiples entre k muestras independientes

Lo primero que se debe hacer es determinar el valor absoluto de la diferencia entre el rango medio ($\bar{R}_j$) de cada grupo o condición con él de cada una de los otros grupos o condiciones: ($|\bar{R}_u$ - $\bar{R}_v|$), dónde u es diferente de v. Luego, se comparan las diferencias con:

$$DS_{\#c,\alpha} = z_{\#c,\alpha}\sqrt{\frac{n(n+1)}{12}\left(\frac{1}{n_u}+\frac{1}{n_v}\right)} \qquad (1)$$

Donde:

$Z_{\#c,\alpha}$: es el valor tabulado en la Tabla 10, para k(k – 1)/2 comparaciones a un nivel dado de α; n es el número total de observaciones, n_u es el número de observaciones en la condición "u" y n_v es el número de observaciones en la condición "v"; "u" diferente de "v".

Ejemplo

- Siendo significativas las diferencias entre razas (grupos), en el ejemplo anterior, se puede proceder a efectuar las comparaciones entre ellas.

Se calcularán primero los valores absolutos de las diferencias entre los rangos medios:

$|\overline{R}_A - \overline{R}_B|$ = |10,00 – 9,500| = 0,500

$|\overline{R}_A - \overline{R}_C|$ = |10,00 – 3,400| = 6,600

$|\overline{R}_B - \overline{R}_C|$ = |9,500 – 3,400| = 6,100

Se usará para α = 0,05, 3(3 -1)/2 = 3 (#c) comparaciones y dos colas, en la Tabla 10 se puede ver que $Z_{3;0,05}$ = 2,394.

Para A vs. C se tendrá:

$$DS_{3;0,05} = 2{,}394\sqrt{\frac{14\ (14+1)}{12}\ (\frac{1}{5}+\frac{1}{5})} = 2{,}394\sqrt{\frac{14\ (15)}{12}\ (0{,}2+0{,}2)} =$$

$$2{,}394\sqrt{17{,}50\ (0{,}400)} = 2{,}394\sqrt{7{,}000} = 2{,}394\ (2{,}646) = 6{,}334$$

Para A vs. B y B vs. C:

$$DS_{3;0,05} = 2{,}394\sqrt{\frac{14\ (14+1)}{12}\ (\frac{1}{5}+\frac{1}{4})} = 2{,}394\sqrt{\frac{14\ (15)}{12}(0{,}20+0{,}25)} =$$

$$2{,}394\sqrt{17{,}50\ (0{,}450)} = 2{,}394\sqrt{7{,}875} = 2{,}394\ (2{,}806) = 6{,}718$$

Se puede observar que sólo hay diferencia (a 5%) entre las condiciones A y C.

b) Comparaciones de los grupos o condiciones con un control

Cuando se desea hacer comparaciones entre los grupos o condiciones y una en particular, que se denominará "controlo testigo", después de encontrar diferencias significativas entre condiciones, con la prueba de Kruskal-Wallis, se puede utilizar el método siguiente:

Esta vez sólo se comparará cada grupo o condición con el "control o testigo", que se denota con la letra t (T). Por lo tanto, hay k - 1 comparaciones.

Se determina el valor absoluto de la diferencia entre el rango medio ($\bar{R}_j$) de cada grupo o condición con él del grupo o condición control: ($|\bar{R}_T - \bar{R}_v|$), dónde v es diferente de T.

Luego, se comparan las diferencias con:

$$DS_{\#c,\alpha} = z_{\#c,\alpha}\sqrt{\frac{n(n+1)}{12}\left(\frac{1}{n_T}+\frac{1}{n_v}\right)} \qquad (2)$$

Donde:

$Z_{\#c,\alpha}$: está tabulada en la Tabla 10, para #c = k – 1 y α; n es el número total de observaciones, n_T el número de observaciones del control o testigo; n_v el número de observaciones de la condición que se quiere comparar.

Ejemplo

- Siendo significativas las diferencias entre razas (grupos), en el ejemplo de los 14 novillos de tres razas distintas donde se determinó la diferencia de peso entre ellas, ahora se puede proceder a efectuar las comparaciones con la condición control o testigo.

Se calculará primero los valores absolutos de las diferencias entre los rangos medios del control o testigo con los de los demás:

Se usará como control o testigo a la condición C (C = T).

$|\bar{R}_T - \bar{R}_B| = |3,400 - 10,000| = 6,600$

$|\bar{R}_T - \bar{R}_A| = |3,400 - 9,500| = 6,100$

Se usará α = 0,05, para (3 -1) = 2 comparaciones, así, de la Tabla 10 se puede ver que $Z_{2;0,05}$ = 2,241; dos colas, para la comparación de T vs. A se tendrá:

$$DS_{2;0,05} = 2,241\sqrt{\frac{14\ (14+1)}{12}\ \left(\frac{1}{5}+\frac{1}{5}\right)} = 2,241\sqrt{\frac{14\ (15)}{12}\ (0,2+0,2)} =$$

$$2,241\sqrt{17,5\ (0,4)} = 2,241\ \sqrt{7,0} = 2,241(2,646) = 5,930$$

Para la raza T vs. B:

$$DS_{3;0,05} = 2{,}243\sqrt{\frac{14\ (14+1)}{12}\ (\frac{1}{5}+\frac{1}{4})} = 2{,}241\sqrt{\frac{14\ (15)}{12}\ (0{,}20+0{,}25)} =$$

$$2{,}241\sqrt{17{,}5\ (0{,}45)} = 2{,}231\sqrt{7{,}875} = 2{,}241\ (2{,}806) = 6{,}288$$

Se puede ver que sólo hay diferencia (a 5%) entre la raza C (que es el control o testigo) y la raza A.

Prueba de Jonckheere

La prueba de Jonkcheere, para niveles ordenados de una variable, es una prueba que se utiliza en vez de la de Kruskal-Wallis, cuando no sólo interesa determinar si hay diferencia entre grupos o condiciones, sino que además se desea investigar si hay cierto orden entre grupos o condiciones. Este orden puede ser creciente o decreciente.

Metodología

Lo primero que se debe hacer es especificar de antemano el orden de los grupos o condiciones. Después se presentan los datos en una tabla de doble entrada, donde las columnas representan los grupos o condiciones, ordenados según el orden especificado de antemano, y las filas, los sujetos, animales u objetos. Dependiendo del orden previo, los grupos o condiciones ordenados se pueden colocar en orden creciente o decreciente.

En esta prueba se cuentan las veces que cada observación en un grupo o condición (no hace falta ordenarlas) precede a cada una de las observaciones de los demás grupos o condiciones. En el caso de empates, se dará una puntuación de 0,5 a cada una. Entonces, en cada columna se suman las veces que cada observación precede a las demás.

Luego se suma todas y se obtiene el estadístico de Jonckheere, que se denota como "J". El procedimiento es sencillo, pero tedioso. Se recomienda la verificación exhaustiva.

Después de obtener J, se compara con la J tabulada en la Tabla 13, para un nivel de significación dado y el tamaño de las muestras (grupos o condiciones).

Ejemplo

- En un corral hay 15 ovejos que han sido igualados por sus pesos, de alguna manera. Se desea probar si un determinado suplemento, administrado en dosis creciente produce un aumento de peso. Se tiene una dosis baja (B), una dosis media (M) y una dosis alta (A).

No habiendo una romana disponible, se procede a determinar los pesos "al ojo", ordenando en una fila los animales, de menor a mayor peso. Desafortunadamente, se pierde, antes de suministrar el suplemento, un ovejo de grupo de la dosis M, quedando, por lo tanto, este grupo con sólo cuatro ovejos.

En el Cuadro 6.3 se presentan los rangos para los pesos de los ovejos de los tres grupos de dosis.

a) Hipótesis de investigación: los pesos de los ovejos se incrementan con la administración desde de la dosis baja (B) hasta la dosis alta (A).

b) Hipótesis estadísticas: H_0: $Md_B \geq Md_M \geq Md_A$
H_1: $Md_B < Md_M < Md_A$

c) Estadístico de prueba: se tienen tres muestras independientes, con dosis crecientes y se desea determinar si las respuestas también son crecientes.

En este caso, se tiene las respuestas en escala ordinal, pero también se hubiera podido efectuar la prueba con datos crudos, sin ordenar, siempre y cuando estén medidos en la escala ordinal, por lo menos. La prueba adecuada es la prueba de Jonckheere.

d) Nivel de significación: se utilizará un nivel de significación de 5% (α = 0,05).

e) Región crítica o Área de rechazo: si la J calculada es mayor que la tabulada en la Tabla 13 para α = 0,05; y tamaños de las muestras 5; 4; 5.; es decir, 4; 5; 5 en la tabla, donde J es igual a 47; se rechaza la H_0.

f) Cálculos: en los cuadros 6.3 y 6.4 se presentan los datos y cálculos, respectivamente.

Cuadro 6.3. El rango de los pesos de catorce ovejos con tres dosis del suplemento.

Dosis B	Dosis M	Dosis A
6	8	11
5	9	4
3	7	12
2	14	13
1		10

Cuadro 6.4 Número de observaciones menores de cada observación de cada grupo con los demás grupos.

Dosis B < Dosis M	Dosis B < Dosis A	Dosis M < Dosis A
4	4	4
4	4	4
4	5	4
4	5	0
4	5	
20	**23**	**12**

J = 20 + 23 + 12 = 55

g) Decisión: como la L calculada (55) es mayor que la tabulada (47), se rechaza H_0.

h) Conclusión: los pesos de los ovejos se incrementan de la dosis baja (B) a la dosis alta (A).

Ejemplo

- Se tiene 15 parcelas de un cultivo completamente invadidos por maleza.

Se desea probar un herbicida a tres concentraciones, para determinar cuál concentración produce a un mejor resultado. Se supone que a mayor concentración, menor porcentaje de área invadida. El área invadida se determina “al ojo”.

Se aplica a cinco parcelas el herbicida en concentración baja (B), a otras cinco el herbicida en concentración media (M) y a las restantes cinco en concentración alta (A). En el Cuadro 6.5 se presentan los resultados.

a) Hipótesis de investigación: a mayor concentración menor porcentaje de área invadida.

b) Hipótesis estadísticas: $H_0: Md_A \geq Md_M \geq Md_B$
$H_1: Md_A < Md_M < Md_B$

c) Estadístico de prueba: se tienen tres muestras independientes, con concentraciones crecientes y se desea determinar si las respuestas son decrecientes.

En este caso se tiene las respuestas en datos crudos, es decir, porcentajes, los cuales se consideran que no se distribuyen normalmente. La prueba adecuada es la prueba de Jonckheere.

d) Nivel de significación: se utilizará un nivel de significación de 1% ($\alpha = 0{,}01$).

e) Región crítica o Área de rechazo: si la J calculada es mayor que la tabulada en la Tabla 13 para $\alpha = 0{,}01$; y tamaños de las muestras 5; 5; 5; donde J es igual a 59; se rechaza la H_0.

f) Cálculos: en los cuadros 6.5 y 6.6 se muestran los resultados y cálculos, respectivamente.

Cuadro 6.5. Porcentaje de terreno con maleza para las tres concentraciones de herbicida.

Concentración Alta (A)	Concentración Media (M)	Concentración Baja (B)
18	26	55
16	30	15
13	20	60
12	80	75
10	40	40

Cuadro 6.6 Número de observaciones menores de cada observación de cada grupo con los demás grupos.

A < M	A < B	M < B
5	4	4
5	4	4
5	5	4
5	5	0
5	5	3,5
25	**23**	**15,5**

J = 25 + 23 + 15,5 = 63,5

En el caso de los empates de los porcentajes 40 de las concentraciones media y baja, se observa que se le ha dado ½ punto al número de observaciones de concentración media, menores que los de la concentración baja, así que ahora tendrá un total de 3 y 1/2; es decir, 3,5 puntos.

g) Decisión: como la J calculada (63,5) es mayor que la tabulada (59), se rechaza la H_0.

h) Conclusión: a mayor concentración menor porcentaje de área invadida.

Nota: el orden en que se colocan las condiciones tiene el mismo orden que la hipótesis de investigación.

Discusión de los métodos de k muestras independientes

En este capítulo se ha presentado tres métodos para el análisis de k muestras independientes.

La prueba de Ji al cuadrado se utiliza cuando los datos están medidos en escala nominal, expresados en frecuencias. Para esta prueba, ninguna celda debe tener un valor esperado menor que uno. El número de celdas con valor esperado menor que cinco no debe superar 20% del número de celdas.

La prueba de Kruskal-Wallis requiere que la variable en estudio sea continua y que esté medida en escala ordinal.

La prueba de Jonckheere se utiliza cuando las k condiciones se presenten en orden creciente o decreciente. Esta prueba requiere que tanto los grupos o condiciones, así como las observaciones, estén medidas en escala ordinal, por lo menos. También requiere que la variable subyacente sea continua.

Capítulo VII

Medidas de asociación

En las investigaciones de las ciencias agropecuarias a veces se pregunta si dos conjuntos de variables están asociados y con que grado. La respuesta a la pregunta se puede hacer, ya para indagar la asociación per se o para un paso previo a una investigación ulterior.

En este capítulo se presentan tres métodos: la prueba de Ji al cuadrado para asociación, el coeficiente de contingencia de Pearson y el coeficiente de correlación de rangos de Spearman.

Prueba de Ji al cuadrado de independencia

Cuando las observaciones se han medido en escala nominal y se presentan como frecuencias en categorías mutuamente excluyentes, con la prueba de Ji al cuadrado se puede determinar la asociación entre dos variables.

Este método es similar al ya visto en los capítulos cuatro y seis, pero difiere de ellos en que ahora el muestreo se hace con base en una población bivariada, es decir, que no existe una variable fija como lo es el de los grupos o condiciones y, por lo tanto, ambas variables son aleatorias y perfectamente intercambiables. También se diferencia por la hipótesis de investigación, las hipótesis estadísticas y las conclusiones. Sin embargo, todos los cálculos son iguales. Esta prueba se utiliza si las variables han sido medidas en escala nominal (atributos), por lo menos.

Metodología

Para la prueba de la hipótesis se utilizará el estadístico X^2, con la fórmula más adecuada que ya se ha utilizado en los capítulos cuatro y seis:

$$X^2 = \sum_{i=1}^{f} \sum_{j=1}^{c} \frac{O_{ij}^2}{E_{ij}} - n \qquad (1)$$

i = 1,2,…, f   j = 1,2,…, c,

Donde:

f: es el número de filas.

c: es el número de columnas.

O_{ij}: es la frecuencia observada en la celda que intersecta la i-ésima fila con la j-ésima columna.

E_{ij}: es la frecuencia esperada en la celda que intersecta la i-ésima fila con la j-ésima columna, (correspondiente a la O_{ij}), n es la frecuencia total o gran total.

Ejemplo

- En un estudio de genética, se tabularon de los registros el color del pelaje de 1.000 pares de madre-hija en caballos de carrera, para determinar si había alguna asociación entre el pelaje de la madre y él de la hija.

a) Hipótesis de investigación: hay asociación entre el pelaje de la madre y él de la hija.

b) Hipótesis estadísticas: H_0: hay independencia
H_1: hay asociación

c) Estadístico de prueba: como se trata de variables medidas en escala nominal (color del pelaje), de dos variables y se estudia su asociación, la prueba indicada es Ji al cuadrado para asociación.

d) Nivel de significación: se utilizará un nivel de significación de 5% ($\alpha = 0,05$).

e) Región crítica o Área de rechazo: si X^2 es igual o mayor que el valor tabulado para (f -1)(c -1) grados de libertad, en la Tabla 2; se rechaza la H_0.

En el presente caso se observaron cinco pelajes de las madres, con los mismos pelajes de las hijas, por lo tanto, se trata de cinco filas con cinco columnas y se tendrá (5 – 1)(5 – 1) grados de libertad, es decir, 16.

Por lo tanto, en la Tabla 2 para 0,05 y 16 grados libertad, se encuentra un Ji al cuadrado igual a 26,30.

Es así que si X^2 es mayor o igual a 26,30; se rechaza la H_0.

f) Cálculos: en el Cuadro 7.1 se presentan los datos y algunos cálculos parciales.

Cuadro 7.1. Color del pelaje de las madres e hijas en caballos de carrera.

Madre	Hija					
	Negro	**Café**	**Bayo**	**Castaño**	**Gris**	**Totales**
Negro	7	7	13	6	5	38
Café	8	40	95	23	7	173
Bayo	11	75	230	113	18	447
Castaño	11	20	101	82	16	230
Gris	5	9	42	17	39	112
Totales	**42**	**151**	**481**	**241**	**85**	**1.000**

f: es el número de filas (cinco).

c: es el número de columnas (cinco).

O_{ij}: es la frecuencia observada en la celda que intersecta la i-ésima fila con la j-ésima columna.

E_{ij}: será la frecuencia esperada en la celda que intersecta la i-ésima fila con la j-ésima columna, (correspondiente a la O_{ij}) y se calcula multiplicando los totales marginales de cada celda (el total de la fila y el total de la columna) y el producto se divide por el gran total “n”.

Se calculan los E_{ij}:

E_{11} = 38(42)/1.000 = 1,60
E_{12} = 38(151)/1.000 = 5,74
E_{13} = 38(481)/1.000 = 18,28
E_{14} = 38(241)/1.000 = 9,16
E_{15} = 38(85)/1.000 = 3,23

E_{21} = 173(42)/1.000 = 7,27
E_{22} = 173(151)/1.000 = 26,12
E_{23} = 173(481)/1.000 = 83,21
E_{24} = 173(241)/1.000 = 41,69
E_{25} = 173(85)/1.000 = 14,71

E_{31} = 447(42)/1.000 = 18,77
E_{32} = 447(151)/1.000 = 67,50
E_{33} = 447(481)/1.000 = 215,01
E_{34} = 447(241)/1.000 = 107,73
E_{35} = 447(85)/1.000 = 38,00

E_{41} = 230(42)/1.000 = 9,66
E_{42} = 230(151)/1.000 = 34,73
E_{43} = 230(481)/1.000 = 110,63
E_{44} = 230(241)/1.000 = 55,43
E_{45} = 230(85)/1.000 = 19,55

E_{51} = 112(42)/1.000 = 4,70
E_{52} = 112(151)/1.000 = 16,91
E_{53} = 112(481)/1.000 = 53,87
E_{54} = 112(241)/1.000 = 26,99
E_{55} = 112(85)/1.000 = 9,52

Cuadro 7.2. Color del pelaje de las madres e hijas en caballos de carrera valores observados y esperados (entre paréntesis).

Madre	Hija					
	Negro	**Café**	**Bayo**	**Castaño**	**Gris**	**Totales**
Negro	7 (1,60)	7 (5,74)	13 (18,28)	6 (9,16)	5 (3,23)	38 (38,01)
Café	8 (7,27)	40 (26,12)	95 (83,21)	23 (41,69)	7 (14,71)	173 (173,00)
Bayo	11 (18,77)	75 (67,50)	230 (215,01)	113 (107,73)	18 (38,00)	447 (447,01)
Castaño	11 (9,66)	20 (34,73)	101 (110,63)	82 (55,43)	16 (19,55)	230 (230,00
Gris	5 (4,70)	9 (16,91)	42 (53,87)	17 (26,99)	39 (9,52)	112 (111,99)
Totales	**42 (42,00)**	**151 (151,00)**	**481 (481,00)**	**241 (241,00)**	**85 (85,01)**	**1.000 (1.000,01)**

Se calculan los $\dfrac{O_{ij}^2}{E_{ij}}$:

$(7)^2/1,60 = 49/1,60 = 30,63$
$(7)^2/5,74 = 49/5,74 = 8,54$
$(13)^2/18,28 = 169/18,28 = 9,25$
$(6)^2/9,16 = 36/9,16 = 3,93$
$(5)^2/3,23 = 25/3,23 = 7,74$

$(8)^2/7,27 = 64/7,27 = 8,80$
$(40)^2/26,12 = 1.600/26,12 = 61,26$
$(95)^2/83,21 = 9.025/83,21 = 108,46$
$(23)^2/41,69 = 529/42,69 = 12,69$
$(7)^2/14,71 = 49/14,71 = 3,33$

$(11)^2/18{,}77 = 121/18{,}77 = 6{,}45$
$(75)^2/67{,}50 = 5.625/67{,}60 = 83{,}33$
$(230)^2/215{,}01 = 52.900/215{,}01 = 246{,}04$
$(113)^2/107{,}73 = 12.769/107{,}73 = 118{,}53$
$(18)^2/38{,}00 = 324/38 = 8{,}53$

$(11)^2/9{,}66= 121/9{,}70 = 12{,}53$
$(20)^2/34{,}73 = 400/34{,}73 = 11{,}52$
$(101)^2/110{,}63 = 10.201/110{,}63 = 92{,}21$
$(82)^2/55{,}43 = 6.724/55{,}43 = 121{,}31$
$(16)^2/19{,}55 = 256/19{,}55 = 13{,}09$

$(5)^2/4{,}70 = 25/4{,}70 = 5{,}32$
$(9)^2/16{,}91 = 81/16{,}91 = 4{,}79$
$(42)^2/53{,}87 = 1.724/53{,}87 = 32{,}75$
$(17)^2/26{,}99 = 289/26{,}99 = 10{,}71$
$(39)^2/9{,}52 = 1.521/9{,}52 = 159{,}77$

$$\sum_{i=1}^{f}\sum_{j=1}^{c}\frac{O_{ij}^2}{E_{ij}} = 1.181{,}51$$

$$X^2 = \sum_{i=1}^{f}\sum_{j=1}^{c}\frac{O_{ij}^2}{E_{ij}} - n = 1.181{,}51 - 1.000 = 181{,}51$$

G) Decisión: como $X^2 = 181{,}51$ es mayor que el tabulado, es decir, 26,30; se rechaza H_0.

H) Conclusión: hay asociación entre el pelaje de la madre y el pelaje de la hija (no son idependientes).

Coeficiente de contingencia de Pearson

Como se puede notar, en el análisis de Ji al cuadrado para asociación, se concluyó que las dos variables en estudio estaban asociadas. Sin embargo, no se indica cual es la fuerza de esa asociación. El coeficiente de contingencia da una respuesta a esa pregunta.

La fórmula del coeficiente de contingencia es la siguiente:

$$CC = \sqrt{\frac{X^2}{X^2 + n}} \qquad (1)$$

Donde, X^2 es el estadístico Ji al cuadrado y n el número total de observaciones.

Ejemplo

- En la prueba de Ji al cuadrado para asociación en el ejemplo anterior, se tiene que la prueba resultó ser significativa, es decir, que hay asociación entre el pelaje de la madre y él de la hija, ahora se desea averiguar qué grado de asociación hay entre esas dos variables.

Se tiene que X^2 = 181,51 y n = 1.000; luego:

$$CC = \sqrt{\frac{181,51}{181,51 + 1000}} = \sqrt{\frac{181,51}{1181,51}} = \sqrt{0,15} = 0,39$$

Entonces, la asociación entre el pelaje de la madre y él de la hija es de 0,39.

Interpretación y limitaciones del coeficiente de contingencia de Pearson

Este coeficiente asume un valor mínimo de cero; significando que no hay asociación y un valor que deberá ser uno; significando una asociación perfecta. Sin embargo, el coeficiente de contingencia (CC) llega sólo a un valor máximo, igual, a

$$CC = \sqrt{\frac{q-1}{q}}$$

Donde, q = min(f, c), es decir, el mínimo entre el tamaño de las filas y de las columnas, lo que significa que es una función de la dimensión de una tabla f x c.

En este caso, se tiene que f = 5 y c = 5; por lo tanto, q = min(5;5) = 5; y el máximo de coeficiente de contingencia (CC) es:

$$\text{CC} = \sqrt{\frac{5-1}{5}} = \sqrt{\frac{4}{5}} = \sqrt{0,80} = 0,89$$

Se podría especular que el valor de coeficiente de contingencia obtenido es 44% del valor máximo, en este caso.

Coeficiente de correlación de rangos de Spearman

El coeficiente de correlación de rangos de Spearman fue la primera prueba que se desarrolló, de los estadísticos basados en rangos.

Este coeficiente de correlación no es otro que el coeficiente producto momento de Pearson, pero con los rangos de los datos. Para utilizarlo, los datos deben estar medidos, por lo menos, en la escala ordinal.

Este método utiliza dos variables, que se pueden etiquetar como "x" e "y", que son simétricas (como en la correlación de Pearson). Si una de las variables, o ambas, están medidas en escalas superiores a la ordinal, basta con convertirlas a rangos, para utilizar la correlación de Spearman.

Metodología

Lo primero que se debe hacer es ordenar cada variable por separado, si se han medido en la escala de intervalo o razón, para así convertirla a rangos.

Luego, para cada unidad de asociación (persona, animal o cosa) se restan los rangos correspondientes y la diferencia se eleva al cuadrado.

La suma de las diferencias al cuadrado ($\sum_{i=1}^{n} d_i^2$) se utiliza en la fórmula siguiente:

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n} \qquad (1)$$

Donde, $-1 \leq r_s \leq 1$

Ejemplo

- Se efectuó un estudio en espárragos, sobre la asociación entre el contenido de fibra y la evaluación organoléptica, cómo indicador de calidad.

Se supuso que la calidad mejoraba (en términos de la evaluación organoléptica), en la medida que la cantidad de fibra disminuía. En el Cuadro 7.3 se muestra los resultados.

a) Hipótesis de investigación: la calidad mejora a medida que el contenido de fibra disminuye. Esta hipótesis es de una cola (la izquierda), por lo tanto:

b) Hipótesis estadísticas: H_0: $r_s \geq 0$
H_1: $r_s < 0$

c) Estadístico de prueba: como una de las variables (calidad) está medida en escala ordinal, no se puede utilizar la correlación de Pearson, por lo cuál se debe usar la correlación de Spearman.

d) Nivel de significación: se utilizará un nivel de significación de 5% (α = 0,05).

e) Región crítica o Área de rechazo: la Tabla 14 tiene sólo valores positivos, así que cuando se usan pruebas con la cola izquierda (con r_s negativos), sencillamente se añade el signo menos (-) al valor tabulado; en este caso en particular, para n = 10 (el número de unidades de asociación o pares) y α = 0,05; para una cola (la izquierda), por lo tanto, se rechazará la H_0 si $r_s \leq - 0,564$.

f) Cálculos: en la Cuadro 7.3 se muestran los datos, así como algunos cálculos.

Cuadro 7.3. Contenido de fibra y calidad en espárragos.

Espárragos	Fibra (x)	Calidad (y)	R(x)	R(y)	d	d²
1	0,059	9,5	3	10,0	-7,0	49,00
2	0,052	9,3	1	9,0	-8,0	64,00
3	0,063	9,1	4	6,0	-2,0	4,00
4	0,080	7,7	9	4,0	5,0	25,00
5	0,071	7,8	7	5,0	2,0	4,00
6	0,075	9,2	8	7,5	0,5	0,25
7	0,058	9,2	2	7,5	-5,5	30,25
8	0,067	6,5	6	3,0	3,0	9,00
9	0,66	6,0	5	2,0	3,0	9,00
10	0,083	4,3	10	1,0	9,0	81,00
Totales					**0,0**	**275,50**

Nota: la calidad óptima es igual a 10.

$$r_s = 1 - \frac{6(275,50)}{10^3 - 10} = 1 - \frac{1653,000}{1000 - 10} = 1 - \frac{1653,000}{990} = 1 - 1,670 = -0,670$$

g) Decisión: como r_s = -0,670 es menor que -0,564; se rechaza H_0.

h) Conclusión: la calidad mejora a medida que el contenido de fibra disminuye.

Ejemplo

- Suponga que se desea averiguar si el incremento de peso de pollitos de cero a ocho semanas está asociado con la jerarquía del pollito dentro de la jaula.

El estudio se efectúa en jaulas con 10 pollitos, alimentados con una determinada ración.

Se calcula el incremento de peso (en gramos) de cero a ocho semanas y, durante este tiempo, se observa el picado de cada pollito y su acceso al comedero y al bebedero.

La jerarquía se evalúa y se le da un valor entre uno y 10, con la mayor jerarquía igual a 10 y la menor (sumisión), con uno.

a) Hipótesis de investigación: hay asociación entre incremento de peso (variable medida en escala de razón) y jerarquía (medida en escala ordinal).

Esta hipótesis es de dos colas, porque no se estipula, conjetura o teoriza si asociación es positiva o negativa, por lo tanto:

b) Hipótesis estadísticas: H_0: $r_s = 0$
H_1: $r_s \neq 0$

c) Nivel de significación: se utilizará un nivel de significación de 1% (α = 0,01).

d) Región crítica o Área de rechazo: la Tabla 14 tiene sólo valores positivos, así que cuando se usan pruebas de dos colas, la cola izquierda tendrá el valor crítico tabulado con signo negativo y para la cola derecha directamente el valor tabulado así, que para este caso en particular, con n = 10 (el número de unidades de asociación o pares) y α = 0,01; para dos colas, de acuerdo con el valor crítico en la Tabla 14, se rechazará la H_0 si r_s es menor que -0,794 o si r_s es mayor que 0,794.

e) Cálculos: en el Cuadro 7.4 se muestran los datos y algunos cálculos.

Cuadro 7.4. Incremento de peso y la jerarquía de los pollitos.

Pollito	Jerarquía (x)	Incremento de peso (y)	R(x)	R(y)	d	d^2
1	7	193	7,0	7,0	0,0	0,00
2	4	150	4,0	1,5	2,5	6,25
3	5	175	5,0	5,0	0,0	0,00
4	3	187	3,0	6,0	-3,0	9,00
5	8	204	8,5	8,0	0,5	0,25
6	8	217	8,5	9,5	-1,0	1,00
7	9	217	10,0	9,5	0,5	0,25
8	2	165	2,0	3,0	-1,0	1,00
9	1	150	1,0	1,5	-0,5	0,25
10	6	170	6,0	4,0	2,0	4,00
Totales					**0,0**	**22,00**

$$r_s = 1 - \frac{6(22)}{10^3 - 10} = 1 - \frac{132}{1.000 - 10} = 1 - \frac{132}{900} = 1 - 0,133 = 0,867$$

f) Decisión: como r_s = 0,867 es mayor que 0,794; se rechaza la H_0.

g) Conclusión: hay asociación positiva entre incremento de peso y jerarquía.

Nota: se indica explícitamente la asociación positiva en el resultado, ya que en la hipótesis de investigación no se ha mencionado de antemano porque ni se conoce ni se asume. El signo (positivo o negativo) es el mismo que el signo de r_s, que en este caso es positivo.

Ejemplo

- Se supone que se desea averiguar si el incremento de peso de pollitos de cero a ocho semanas está asociado en forma positiva con la jerarquía del pollito dentro de la jaula.

El estudio se efectúa en jaulas con 10 pollitos, alimentados con una determinada ración.

Se calcula el incremento de peso (en gramos) de cero a ocho semanas y, durante este tiempo, se observa el picado de cada pollito y su acceso al comedero y al bebedero.

La jerarquía se evalúa y se da un valor entre uno y 10, con la mayor jerarquía igual a 10 y la sumisión (menor jerarquía) con uno.

a) Hipótesis de investigación: hay asociación positiva entre incremento de peso (variable medida en escala de razón) y jerarquía (medida en escala ordinal).

Esta hipótesis es de una cola, porque se supone que la asociación es positiva.

b) Hipótesis estadísticas: H_0: $r_s \leq 0$
H_1: $r_s > 0$

c) Nivel de significación: se utilizará un nivel de significación de 1% ($\alpha = 0{,}01$).

d) Región crítica o Área de rechazo: la Tabla 14 tiene sólo valores positivos, así que cuando se usan pruebas con la cola derecha se lee el valor tabulado directamente, que para este caso en particular, con n = 10 (el número de unidades de asociación o pares) y $\alpha = 0{,}01$; para una cola es de 0,745. Se rechazará la H_0 si r_s es mayor que 0,745.

e) Cálculos: en el Cuadro 7.4, del ejemplo anterior, se muestran los datos y algunos cálculos. $r_s = 0{,}867$.

f) Decisión: como $r_s = 0{,}867$ es mayor que 0,794; se rechaza la H_0.

g) Conclusión: hay asociación positiva entre incremento de peso y jerarquía.

Observaciones empatadas

Cuando hay dos o más observaciones empatadas, dentro de una variable, se deberá dar a cada una el rango medio.

Si hay observaciones empatadas, se debe efectuar una corrección de r_s, ya que éste estará sobre estimado, es decir, será mayor que el valor real de r_s.

Sin embargo, la corrección será mínima si el número de observaciones empatadas en una variable, en la otra, o en ambas no es demasiado grande.

La fórmula de la corrección es complicada y su cálculo tedioso, por lo tanto, se puede, en el caso de un gran número de empates, calcular r_s, utilizando una calculadora científica con programa de correlación de Pearson (r), con los rangos.

Si se calcula r_{rangos} de la manera indicada, se tendrá que para el ejemplo donde se quiere averiguar el grado de asociación que hay entre el pelaje de la madre y el de la hija, r_{rangos} = -0,675; no muy lejos de -0,670 = r_s.

Para el ejemplo donde se estudia la asociación entre el contenido de fibra y la evaluación organoléptica en espárragos, con más empates, r_{rangos} = 0,865, tampoco muy lejos de 0,867 = r_s.

Discusión de las medidas de asociación

En este capítulo se ha presentado tres métodos de asociación: Ji al cuadrado, coeficiente de contingencia y correlación de Spearman. Todos estos métodos son una alternativa, cuando los datos y supuestos no llenen los requisitos necesarios para una prueba paramétrica.

Si los datos han sido medidos en escala nominal y agrupados en distribuciones de frecuencias con clases mutuamente excluyentes, entonces la prueba adecuada es la de Ji al cuadrado.

Si además de utilizar a la prueba de Ji al cuadrado, se está interesado en un coeficiente de asociación, basado en los mismos datos, entonces, se puede calcular el coeficiente de contingencia.

Por último, si los datos han sido medidos, por lo menos, en la escala ordinal, se puede calcular el coeficiente de correlación por rangos de Spearman.

Bibliografía consultada

Cambell, RC. 1974. Statistics for biologists. 2 ed. Cambrige University Press. Londres, Inglaterra. 385 p.

Connover, WJ. 1980. Practical nonparametric statistics. 2 ed. Nueva York, EE UU. Wiley. 493 p.

Daniel, W. 1978. Applied nonparametric statistics. Boston, EE UU. Houghton Mifflin. 510 p.

Gibbons, JD. 1971. Nonparametric statistical inference. Tokio, Japón. McGraw-Hill Kogakusha. 306 p.

Gibbons, JD. 1976. Nonparametric methods for quantitative analysis. Nueva York, EE UU. Holt, Rinehard and Winston. 463 p.

Hollander, M; Wolfe, DA. 1973. Nonparametric statistical methods. Nueva York, EE UU. Wiley. 503 p.

Leach, C. 1982. Fundamentos de estadística. México, México. Limusa. 422 p.

Lehman, E; D'Abera, HJM. 1975. Nonparametric statistical methods based on ranks. San Francisco, EE UU. Holden-Day. 457 p.

Mehta, CR; Patel, NR. 1996. SPSS Exact tests 7.0 for Windows. SPSS. 220 p.

Runyon, R; Haber, A. 1984. Estadística para las ciencias sociales. México, México. Fondo Educativo Interamericano. 467 p.

Siegel, S; Castellan, NJ. 1995. Estadística no paramétrica. 4 ed. México, México. Trillas. 437 p.

Wiedenhofer, H.1980. Prueba de normalidad. Maracay, VE. Fondo Nacional de Investigaciones Agropecuarias. 91 p. (Serie E – Nº 1.)

Índice de pruebas no paramétricas

Prueba	Nivel de medida	
	Nominal	**Ordinal**
Una muestra	Binomial. cap. II - p. 38 Ji al cuadrado. cap. II - p. 45	Kolmogorov-Smirnov. cap. II - p. 54 Rachas. cap. II - p. 62
Dos muestras relacionadas	McNemar. cap. III - p. 72	Signos. cap. III - p. 82 Wilcoxon. cap. III - p. 92
Dos muestras independientes	Fisher. cap. IV - p. 102 Ji al cuadrado de homogeneidad. cap. IV - p. 112	Mediana. cap. IV - p. 125 U de Mann-Whitney. cap. IV - p. 135
k muestras relacionadas	Q de Cochran. cap. V - p. 151	Friedman. cap. V - p. 155 Coeficiente de concordancia W de Kendall. cap. V - p. 165 Page. cap. V - p. 166
k muestras independientes	Homogeneidad de Ji al cuadrado. cap. VI - p. 178	Kruskal-Wallis. cap. VI - p. 182 Jonckheere. cap. VI - p. 190
Medidas de asociación	Ji al cuadrado de independencia. cap. VII - p. 197 Coeficiente de contingencia de Pearson. cap. VII - p. 204	Spearman. cap. VII - p. 206

Tablas

Tabla 1. Distribución Binomial.

n	Prueba unilateral		Prueba bilateral	
	0,05	0,01	0,05	0,01
5	5	-	-	-
6	6	-	6	-
7	7	7	7	-
8	7	8	8	-
9	8	9	8	9
10	9	10	9	10
11	9	10	10	11
12	10	11	10	11
13	10	12	11	12
14	11	12	12	13
15	12	13	12	13
16	12	14	13	14
17	13	14	13	15
18	13	15	14	15
19	14	15	15	16
20	15	16	15	17
21	15	17	16	17
22	16	17	17	18
23	16	18	17	19
24	17	19	18	19
25	18	19	18	20
26	18	20	19	20
27	19	20	20	21

Valores críticos para x o n-x (la que sea mayor), para P = Q – ½.
x es la frecuencia de la categoría P y n – x de la categoría Q.

Tabla 1. Distribución Binomial (Continuación).

n	Prueba unilateral		Prueba bilateral	
	0,05	0,01	0,05	0,01
28	19	21	20	22
29	20	22	21	22
30	20	22	21	23
31	21	23	22	24
32	22	24	23	24
33	22	24	23	25
34	23	25	24	25
35	23	25	24	26
36	24	26	25	27
37	24	27	25	27
38	25	27	26	28
39	26	28	27	28
40	26	28	27	29
41	27	29	28	30
42	27	29	28	30
43	28	30	29	31
44	28	31	29	31
45	29	31	30	32
46	30	32	31	33
47	30	32	31	33
48	31	33	32	34
49	31	34	32	35
50	32	34	33	35

Tabla 2. Valores críticos de Ji al cuadrado.

gl	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,00016	0,00063	0,0039	0,016	0,064	0,15	0,46	1,07	1,64	2,71	3,84	5,41	6,64	10,83
2	0,02	0,04	0,10	0,21	0,45	0,71	1,39	2,41	3,22	4,60	5,99	7,82	9,21	13,82
3	0,12	0,18	0,35	0,58	1,00	1,42	2,37	3,66	4,64	6,25	7,82	9,84	11,34	16,27
4	0,30	0,43	0,71	1,06	1,65	2,20	3,35	4,88	5,99	7,78	9,49	11,67	13,28	18,46
5	0,55	0,75	1,14	1,61	2,34	3,00	4,35	6,06	7,29	9,24	11,07	13,39	15,09	20,52
6	0,87	1,13	1,64	2,20	3,07	3,83	5,35	7,23	8,56	10,64	12,59	15,03	16,81	22,46
7	1,24	1,56	2,17	2,83	3,82	4,67	6,35	8,38	9,80	12,02	14,07	16,62	18,48	24,32
8	1,65	2,03	2,73	3,49	4,59	5,53	7,34	9,52	11,03	13,36	15,51	18,17	20,09	26,12
9	2,09	2,53	3,32	4,17	5,38	6,39	8,34	10,66	12,24	14,68	16,92	19,68	21,67	27,88
10	2,56	3,06	3,94	4,86	6,18	7,27	9,34	11,78	18,44	15,99	18,31	21,16	23,21	29,59
11	3,05	3,61	4,58	5,58	6,99	8,15	10,34	12,90	14,63	17,28	19,68	22,62	24,72	31,26
12	3,57	4,18	5,23	6,30	7,81	9,03	11,34	14,01	15,81	18,55	21,03	24,05	26,22	32,91
13	4,11	4,76	5,89	7,04	8,63	9,93	12,34	15,12	16,98	19,81	22,36	25,47	27,69	34,53
14	4,66	5,37	6,57	7,79	9,47	10,82	13,34	16,22	18,15	21,06	23,68	26,87	29,14	36,12
15	5,23	5,98	7,26	8,55	10,31	11,72	14,34	17,32	19,31	22,31	25,00	28,26	30,58	37,70
16	5,81	6,61	7,96	9,31	11,15	12,62	15,34	18,42	20,46	23,54	26,30	29,63	32,00	39,29
17	6,41	7,26	8,67	10,08	12,00	13,53	16,34	19,51	21,62	24,77	27,59	31,00	33,41	40,75
18	7,02	7,91	9,39	10,86	12,86	14,44	17,34	20,60	22,76	25,99	28,87	32,35	34,80	42,31
19	7,63	8,57	10,12	11,65	13,72	15,35	18,34	21,69	23,90	27,20	30,14	33,69	36,19	43,82
20	8,26	9,24	10,85	12,44	14,58	16,27	19,34	22,78	25,04	28,41	31,41	35,02	37,57	45,32
21	8,90	9,92	11,59	13,24	15,44	17,18	20,34	23,86	26,17	29,62	32,67	36,34	38,93	46,80
22	9,54	10,60	12,34	14,04	16,31	18,10	21,24	24,94	27,30	30,81	33,92	37,66	40,26	48,27
23	10,20	11,29	13,09	14,85	17,19	19,02	22,34	26,02	24,83	32,01	35,17	38,97	41,64	49,73
24	10,86	11,99	13,85	15,66	18,06	19,94	23,34	27,10	29,55	33,20	36,42	40,27	42,98	51,18
25	11,52	12,70	14,61	16,47	18,94	20,87	24,34	28,17	30,68	34,38	37,65	41,57	44,31	52,62
26	12,20	13,41	15,38	17,29	19,82	21,79	25,34	29,25	31,80	35,56	38,88	42,86	45,64	54,05
27	12,88	14,12	16,15	18,11	20,70	22,72	26,34	30,32	32,91	36,74	40,11	44,14	46,96	55,48
28	13,56	14,85	16,93	18,94	21,59	23,65	27,34	31,39	34,03	37,92	41,34	45,42	48,28	56,89
29	14,26	15,57	17,71	19,77	22,48	24,58	28,34	32,46	35,14	39,09	42,56	46,69	49,59	58,30
30	14,95	16,31	18,49	20,60	23,36	25,51	29,34	33,53	36,25	40,26	43,77	47,96	50,89	59,70

Tabla 3. Valores críticos de la D en la prueba de una muestra de Kolmogorov-Smirnov.

Tamaño de la muestra (N)	Nivel de significación para D = máximo $\lvert F_O(X) - S_N(X)\rvert$				
	0,20	0,15	0,10	0,05	0,01
1	0,900	0,925	0,950	0,975	0,995
2	0,684	0,726	0,776	0,842	0,929
3	0,565	0,597	0,642	0,708	0,828
4	0,494	0,525	0,564	0,624	0,733
5	0,446	0,474	0,510	0,565	0,669
6	0,410	0,436	0,470	0,521	0,618
7	0,381	0,405	0,438	0,486	0,577
8	0,358	0,381	0,411	0,457	0,543
9	0,339	0,360	0,388	0,432	0,514
10	0,322	0,342	0,368	0,410	0,490
11	0,307	0,326	0,352	0,391	0,468
12	0,295	0,313	0,338	0,375	0,450
13	0,284	0,302	0,325	0,361	0,433
14	0,274	0,292	0,314	0,349	0,418
15	0,266	0,283	0,304	0,338	0,404
16	0,258	0,274	0,295	0,328	0,392
17	0,250	0,266	0,286	0,318	0,381
18	0,244	0,259	0,278	0,309	0,371
19	0,237	0,252	0,272	0,301	0,363
20	0,231	0,246	0,264	0,294	0,356
25	0,21	0,22	0,24	0,27	0,32
30	0,19	0,20	0,22	0,24	0,29
35	0,18	0,19	0,21	0,23	0,27
Más de 35	$\frac{1,07}{\sqrt{N}}$	$\frac{1,14}{\sqrt{N}}$	$\frac{1,22}{\sqrt{N}}$	$\frac{1,36}{\sqrt{N}}$	$\frac{1,63}{\sqrt{N}}$

Tabla 4. Valores críticos de r en la Prueba de Rachas.

Tabla 4a

n_1 \ n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2											2	2	2	2	2	2	2	2	2
3					2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
4				2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4
5			2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5
6		2	2	3	3	3	3	4	4	4	4	5	5	5	5	5	5	6	6
7		2	2	3	3	3	4	4	5	5	5	5	5	6	6	6	6	6	6
8		2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7
9		2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8
10		2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	9
11		2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	10	10
13	2	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	10	10
14	2	2	3	4	5	5	6	7	7	8	8	9	9	9	10	10	10	11	11
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	11	12
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14

En las tablas 4a y 4b se dan los diferentes valores críticos de r para diferentes valores de n_1 y n_2. Para la Prueba de Rachas, cualquier valor de r que sea igual o menor que al tabulado en la Tabla 4a, o igual o mayor que el tabulado en la Tabla 4b, es significativo a 5%.

Tabla 4. Valores críticos de r en la Prueba de Rachas (Continuación).
Tabla 4b

n_1 \ n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4				9	9														
5			9	10	10	11	11												
6			9	10	11	12	12	13	13	13	13								
7				11	12	13	13	14	14	14	14	15	15	15					
8				11	12	13	14	14	15	15	16	16	16	16	17	17	17	17	17
9					13	14	14	15	16	16	16	17	17	18	18	18	18	18	18
10					13	14	15	16	16	17	17	18	18	18	19	19	19	20	20
11					13	14	15	16	17	17	18	19	19	19	20	20	20	21	21
12					13	14	16	16	17	18	19	19	20	20	21	21	21	22	22
13						15	16	17	18	19	19	20	20	21	21	22	22	23	23
14						15	16	17	18	19	20	20	21	22	22	23	23	23	24
15						15	16	18	18	19	20	21	22	22	23	23	24	24	25
16							17	18	19	20	21	21	22	23	23	24	25	25	25
17							17	18	19	20	21	22	23	23	24	25	25	26	26
18							17	18	19	20	21	22	23	24	25	25	26	26	27
19							17	18	20	21	22	23	23	24	25	26	26	27	27
20							17	18	20	21	22	23	24	25	25	26	27	27	28

Tabla 5. Valores críticos de T- y T+ en la Prueba de Wilcoxon.

N	Pruebas de una cola			
	0,05	0,025	0,01	0,005
	Pruebas de dos colas			
	0,10	0,05	0,02	0,01
1				
2				
3				
4				
5	0,15			
6	2,19	0,21		
7	3,25	2,26	0,28	
8	5,31	3,33	1,35	0,36
9	8,37	5,40	3,42	1,44
10	10,45	8,47	5,50	3,52
11	13,53	10,56	7,59	5,61
12	17,61	13,65	9,69	7,71
13	21,70	17,74	12,79	9,82
14	25,80	21,84	15,90	12,93
15	30,90	25,95	19,101	15,105
16	35,101	29,107	23,113	19,117
17	41,112	34,119	28,125	23,130
18	47,124	40,131	32,139	27,144
19	53,137	46,144	37,153	33,158
20	60,150	52,158	43,167	37,173
21	67,164	58,173	49,182	42,189
22	75,178	66,187	55,198	48,205
23	83,193	73,203	62,214	54,222
24	91,209	81,210	69,231	61,239
25	100,225	89,236	76,249	68,257

Tabla 6. Valores críticos de D (o C) de Fisher.

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 3	C + D = 3	3	0	-	-	-
A + B = 4	C + D = 4	4	0	0	-	-
	C + D = 3	4	0	-	-	-
A + B = 5	C + D = 5	5	1	1	0	0
		4	0	0	-	-
	C + D = 4	5	1	0	0	-
		4	0	-	-	-
	C + D = 3	5	0	0	-	-
	C + D = 2	5	0	-	-	-
A + B = 6	C + D = 6	6	2	1	1	0
		5	1	0	0	-
		4	0	-	-	-
	C + D = 5	6	1	0	0	0
		5	0	0	-	-
		4	0	-	-	-
	C + D = 4	6	1	0	0	0
		5	0	0	-	-
	C + D = 3	6	0	0	-	-
		5	0	-	-	-
	C + D = 2	6	0	-	-	-
A + B = 7	C + D = 7	7	3	2	1	1
		6	1	1	0	0
		5	0	0	-	-
		4	0	-	-	-
	C + D = 6	7	2	2	1	1
		6	1	0	0	0
		5	0	0	-	-
		4	0	-	-	-
	C + D = 5	7	2	1	0	0
		6	1	0	0	-
		5	0	-	-	-
	C + D = 4	7	1	1	0	0
		6	0	0	-	-
		5	0	-	-	-
	C + D = 3	7	0	0	0	-
		6	0	-	-	-
	C + D = 2	7	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 8	C + D = 8	8	4	3	2	2
		7	2	2	1	0
		6	1	1	0	0
		5	0	0	-	-
		4	0	-	-	-
	C + D = 7	8	3	2	2	1
		7	2	1	1	0
		6	1	0	0	-
		5	0	0	-	-
	C + D = 6	8	2	2	1	1
		7	1	1	0	0
		6	0	0	0	-
		5	0	-	-	-
	C + D = 5	8	2	1	1	0
		7	1	0	0	0
		6	0	0	-	-
		5	0	-	-	-
	C + D = 4	8	1	1	0	0
		7	0	0	-	-
		6	0	-	-	-
	C + D = 3	8	0	0	0	-
		7	0	0	-	-
	C + D = 2	8	0	0	-	-
A + B = 9	C + D = 9	9	5	4	3	3
		8	3	3	2	1
		7	2	1	1	0
		6	1	1	0	0
		5	0	0	-	-
		4	0	-	-	-
	C + D = 8	9	4	3	3	2
		8	3	2	1	1
		7	2	1	0	0
		6	1	0	0	-
		5	0	0	-	-
	C + D = 7	9	3	3	2	2
		8	2	2	1	0
		7	1	1	0	0
		6	0	0	-	-
		5	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 9	C + D = 6	9	3	2	1	1
		8	2	1	0	0
		7	1	0	0	-
		6	0	0	-	-
		5	0	-	-	-
	C + D = 5	9	2	1	1	1
		8	1	1	0	0
		7	0	0	-	-
		6	0	-	-	-
	C + D = 4	9	1	1	0	0
		8	0	0	0	-
		7	0	0	-	-
		6	0	-	-	-
	C + D = 3	9	1	0	0	0
		8	0	0	-	-
		7	0	-	-	-
	C + D = 2	9	0	0	-	-
A + B = 10	C + D = 10	10	6	5	4	3
		9	4	3	3	2
		8	3	2	1	1
		7	2	1	1	0
		6	1	0	0	-
		5	0	0	-	-
		4	0	-	-	-
	C + D = 9	10	5	4	3	3
		9	4	3	2	2
		8	2	2	1	1
		7	1	1	0	0
		6	1	0	0	-
		5	0	0	-	-
	C + D = 8	10	4	4	3	2
		9	3	2	2	1
		8	2	1	1	0
		7	1	1	0	0
		6	0	0	-	-
		5	0	-	-	-
	C + D = 7	10	3	3	2	2
		9	2	2	1	1
		8	1	1	0	0
		7	1	0	0	-
		6	0	0	-	-
		5	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 10	C + D = 6	10	3	2	2	1
		9	2	1	1	0
		8	1	1	0	0
		7	0	0	-	-
		6	0	-	-	-
	C + D = 5	10	2	2	1	1
		9	1	1	0	0
		8	1	0	0	-
		7	0	0	-	-
		6	0	-	-	-
	C + D = 4	10	1	1	0	0
		9	1	0	0	0
		8	0	0	-	-
		7	0	-	-	-
	C + D = 3	10	1	0	0	0
		9	0	0	-	-
		8	0	-	-	-
	C + D = 2	10	0	0	-	-
		9	0	-	-	-
A + B = 11	C + D = 11	11	7	6	5	4
		10	5	4	3	3
		9	4	3	2	2
		8	3	2	1	1
		7	2	1	0	0
		6	1	0	0	-
		5	0	0	-	-
		4	0	-	-	-
	C + D = 10	11	6	5	4	4
		10	4	4	3	2
		9	3	3	2	1
		8	2	2	1	0
		7	1	1	0	0
		6	1	0	0	-
		5	0	-	-	-
	C + D = 9	11	5	4	4	3
		10	4	3	2	2
		9	3	2	1	1
		8	2	1	1	0
		7	1	1	0	0
		6	0	0	-	-
		5	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 11	C + D = 8	11	4	4	3	3
		10	3	3	2	1
		9	2	2	1	1
		8	1	1	0	0
		7	1	0	0	-
		6	0	0	-	-
		5	0	-	-	-
	C + D = 7	11	4	3	2	2
		10	3	2	1	1
		9	2	1	1	0
		8	1	1	0	0
		7	0	0	-	-
		6	0	0	-	-
	C + D = 6	11	3	2	2	1
		10	2	1	1	0
		9	1	1	0	0
		8	1	0	0	-
		7	0	0	-	-
		6	0	-	-	-
	C + D = 5	11	2	2	1	1
		10	1	1	0	0
		9	1	0	0	0
		8	0	0	-	-
		7	0	-	-	-
	C + D = 4	11	1	1	1	0
		10	1	0	0	0
		9	0	0	-	-
		8	0	-	-	-
	C + D = 3	11	1	0	0	0
		10	0	0	-	-
		9	0	-	-	-
	C + D = 2	11	0	0	-	-
		10	0	-	-	-
A + B = 12	C + D = 12	12	8	7	6	5
		11	6	5	4	4
		10	5	4	3	2
		9	4	3	2	1
		8	3	2	1	1
		7	2	1	0	0
		6	1	0	0	-
		5	0	0	-	-
		4	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 12	C + D = 11	12	7	6	5	5
		11	5	5	4	3
		10	4	3	2	2
		9	3	2	2	1
		8	2	1	1	0
		7	1	1	0	0
		6	1	0	0	-
		5	0	0	-	-
	C + D = 10	12	6	5	5	4
		11	5	4	3	3
		10	4	3	2	2
		9	3	2	1	1
		8	2	1	0	0
		7	1	0	0	0
		6	0	0	-	-
		5	0	-	-	-
	C + D = 9	12	5	5	4	3
		11	4	3	3	2
		10	3	2	2	1
		9	2	2	1	0
		8	1	1	0	0
		7	1	0	0	-
		6	0	0	-	-
		5	0	-	-	-
	C + D = 8	12	5	4	3	3
		11	3	3	2	2
		10	2	2	1	1
		9	2	1	1	0
		8	1	1	0	0
		7	0	0	-	-
		6	0	0	-	-
	C + D = 7	12	4	3	3	2
		11	3	2	2	1
		10	2	1	1	0
		9	1	1	0	0
		8	1	0	0	-
		7	0	0	-	-
		6	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 12	C + D = 6	12	3	3	2	2
		11	2	2	1	1
		10	1	1	0	0
		9	1	0	0	0
		8	0	0	-	-
		7	0	0	-	-
		6	0	-	-	-
	C + D = 5	12	2	2	1	1
		11	1	1	1	0
		10	1	0	0	0
		9	0	0	0	-
		8	0	0	-	-
		7	0	-	-	-
	C + D = 4	12	2	1	1	0
		11	1	0	0	0
		10	0	0	0	-
		9	0	0	-	-
		8	0	-	-	-
	C + D = 3	12	1	0	0	0
		11	0	0	0	-
		10	0	0	-	-
		9	0	-	-	-
	C + D = 2	12	0	0	-	-
		11	0	-	-	-
A + B = 13	C + D = 13	13	9	8	7	6
		12	7	6	5	4
		11	6	5	4	3
		10	4	4	3	2
		9	3	3	2	1
		8	2	2	1	0
		7	2	1	0	0
		6	1	0	0	-
		5	0	0	-	-
		4	0	-	-	-
	C + D = 12	13	8	7	6	5
		12	6	5	5	4
		11	5	4	3	3
		10	4	3	2	2
		9	3	2	1	1
		8	2	1	1	0
		7	1	1	0	0
		6	1	0	0	-
		5	0	0	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 13	C + D = 11	13	1	6	5	5
		12	6	5	4	3
		11	4	4	3	2
		10	3	3	2	1
		9	3	2	1	1
		8	2	1	0	0
		7	1	0	0	0
		6	0	0	-	-
		5	0	-	-	-
	C + D = 10	13	6	6	5	4
		12	5	4	3	3
		11	4	3	2	2
		10	3	2	1	1
		9	2	1	1	0
		8	1	1	0	0
		7	1	0	0	-
		6	0	0	-	-
		5	0	-	-	-
	C + D = 9	13	5	5	4	4
		12	4	4	3	2
		11	3	3	2	1
		10	2	2	1	1
		9	2	1	0	0
		8	1	1	0	0
		7	0	0	-	-
		6	0	0	-	-
		5	0	-	-	-
	C + D = 8	13	5	4	3	3
		12	4	3	2	2
		11	3	2	1	1
		10	2	1	1	0
		9	1	1	0	0
		8	1	0	0	-
		7	0	0	-	-
		6	0	-	-	-
	C + D = 7	13	4	3	3	2
		12	3	2	2	1
		11	2	2	1	1
		10	1	1	0	0
		9	1	0	0	0
		8	0	0	-	-
		7	0	0	-	-
		6	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 13	C + D = 6	13	3	3	2	2
		12	2	2	1	1
		11	2	1	1	0
		10	1	1	0	0
		9	1	0	0	-
		8	0	0	-	-
		7	0	-	-	-
	C + D = 5	13	2	2	1	1
		12	2	1	1	0
		11	1	1	0	0
		10	1	0	0	-
		9	0	0	-	-
		8	0	-	-	-
	C + D = 4	13	2	1	1	0
		12	1	1	0	0
		11	0	0	0	-
		10	0	0	-	-
		9	0	-	-	-
	C + D = 3	13	1	1	0	0
		12	0	0	0	-
		11	0	0	-	-
		10	0	-	-	-
	C + D = 2	13	0	0	0	-
		12	0	-	-	-
A + B = 14	C + D = 14	14	10	9	8	7
		13	8	7	6	5
		12	6	6	5	4
		11	5	4	3	3
		10	4	3	2	2
		9	3	2	2	1
		8	2	2	1	0
		7	1	1	0	0
		6	1	0	0	-
		5	0	0	-	-
		4	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 14	C + D = 13	14	9	8	7	6
		13	7	6	5	5
		12	6	5	4	3
		11	5	4	3	2
		10	4	3	2	2
		9	3	2	1	1
		8	2	1	1	0
		7	1	1	0	0
		6	1	0	-	-
		5	0	0	-	-
	C + D = 12	14	8	7	6	6
		13	6	6	5	4
		12	5	4	4	3
		11	4	3	3	2
		10	3	3	2	1
		9	2	2	1	1
		8	2	1	0	0
		7	1	0	0	-
		6	0	0	-	-
		5	0	-	-	-
	C + D = 11	14	7	6	6	5
		13	6	5	4	4
		12	5	4	3	3
		11	4	3	2	2
		10	3	2	1	1
		9	2	1	1	0
		8	1	1	0	0
		7	1	0	0	-
		6	0	0	-	-
		5	0	-	-	-
	C + D = 10	14	6	6	5	4
		13	5	4	4	3
		12	4	3	3	2
		11	3	3	2	1
		10	2	2	1	1
		9	2	1	0	0
		8	1	1	0	0
		7	0	0	0	-
		6	0	0	-	-
		5	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 14	C + D = 9	14	6	5	4	4
		13	4	4	3	3
		12	3	3	2	2
		11	3	2	1	1
		10	2	1	1	0
		9	1	1	0	0
		8	1	0	0	-
		7	0	0	-	-
		6	0	-	-	-
	C + D = 8	14	5	4	4	3
		13	4	3	2	2
		12	3	2	2	1
		11	2	2	1	1
		10	2	1	0	0
		9	1	0	0	0
		8	0	0	0	-
		7	0	0	-	-
		6	0	-	-	-
	C + D = 7	14	4	3	3	2
		13	3	2	2	1
		12	2	2	1	1
		11	2	1	1	0
		10	1	1	0	0
		9	1	0	0	-
		8	0	0	-	-
		7	0	-	-	-
	C + D = 6	14	3	3	2	2
		13	2	2	1	1
		12	2	1	1	0
		11	1	1	0	0
		10	1	0	0	-
		9	0	0	-	-
		8	0	0	-	-
		7	0	-	-	-
	C + D = 5	14	2	2	1	1
		13	2	1	1	0
		12	1	1	0	0
		11	1	0	0	0
		10	0	0	-	-
		9	0	0	-	-
		8	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 14	C + D = 4	14	2	1	1	1
		13	1	1	0	0
		12	1	0	0	0
		11	0	0	-	-
		10	0	0	-	-
		9	0	-	-	-
	C + D = 3	14	1	1	0	0
		13	0	0	0	-
		12	0	0	-	-
		11	0	-	-	-
	C + D = 2	14	0	0	0	-
		13	0	0	-	-
		12	0	-	-	-
A + B = 15	C + D = 15	15	11	10	9	8
		14	9	8	7	6
		13	7	6	5	5
		12	6	5	4	4
		11	5	4	3	3
		10	4	3	2	2
		9	3	2	1	1
		8	2	1	1	0
		7	1	1	0	0
		6	1	0	0	-
		5	0	0	-	-
		4	0	-	-	-
	C + D = 14	15	10	9	8	7
		14	8	7	6	6
		13	7	6	5	4
		12	6	5	4	3
		11	5	4	3	2
		10	4	3	2	1
		9	3	2	1	1
		8	2	1	1	0
		7	1	1	0	0
		6	1	0	-	-
		5	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 15	C + D = 13	15	9	8	7	7
		14	7	7	6	5
		13	6	5	4	4
		12	5	4	3	3
		11	4	3	2	2
		10	3	2	2	1
		9	2	2	1	0
		8	2	1	0	0
		7	1	0	0	-
		6	0	0	-	-
		5	0	-	-	-
	C + D = 12	15	8	7	7	6
		14	7	6	5	4
		13	6	5	4	3
		12	5	4	3	2
		11	4	3	2	2
		10	3	2	1	1
		9	2	1	1	0
		8	1	1	0	0
		7	1	0	0	-
		6	0	0	-	-
		5	0	-	-	-
	C + D = 11	15	7	7	6	5
		14	6	5	4	4
		13	5	4	3	3
		12	4	3	2	2
		11	3	2	2	1
		10	2	2	1	1
		9	2	1	0	0
		8	1	1	0	0
		7	1	0	0	-
		6	0	0	-	-
		5	0	-	-	-
	C + D = 10	15	6	6	5	5
		14	5	5	4	3
		13	4	4	3	2
		12	3	3	2	2
		11	3	2	1	1
		10	2	1	1	0
		9	1	1	0	0
		8	1	0	0	-
		7	0	0	-	-
		6	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 15	C + D = 9	15	6	5	4	4
		14	5	4	3	3
		13	4	3	2	2
		12	3	2	2	1
		11	2	2	1	1
		10	2	1	0	0
		9	1	1	0	0
		8	1	0	0	-
		7	0	0	-	-
		6	0	-	-	-
	C + D = 8	15	5	4	4	3
		14	4	3	3	2
		13	3	2	2	1
		12	2	2	1	1
		11	2	1	1	0
		10	1	1	0	0
		9	1	0	0	-
		8	0	0	-	-
		7	0	-	-	-
		6	0	-	-	-
	C + D = 7	15	4	4	3	3
		14	3	3	2	2
		13	2	2	1	1
		12	2	1	1	0
		11	1	1	0	0
		10	1	0	0	0
		9	0	0	-	-
		8	0	0	-	-
		7	0	-	-	-
	C + D = 6	15	3	3	2	2
		14	2	2	1	1
		13	2	1	1	0
		12	1	1	0	0
		11	1	0	0	0
		10	0	0	0	-
		9	0	0	-	-
		8	0	-	-	-
	C + D = 5	15	2	2	2	1
		14	2	1	1	1
		13	1	1	0	0
		12	1	0	0	0
		11	0	0	0	-
		10	0	0	-	-
		9	0	-	-	-

Tabla 6. Valores críticos de D (o C) de Fisher (Continuación).

Totales en el margen derecho		B (o A) †	Nivel de significación			
			0,05	0,025	0,01	0,005
A + B = 15	C + D = 4	15	2	1	1	1
		14	1	1	0	0
		13	1	0	0	0
		12	0	0	0	-
		11	0	0	-	-
		10	0	-	-	-
	C + D = 3	15	1	1	0	0
		14	0	0	0	0
		13	0	0	-	-
		12	0	0	-	-
		11	0	-	-	-
	C + D = 2	15	0	0	0	-
		14	0	0	-	-
		13	0	-	-	-

Tabla 7. Valores críticos de U y de U' (subrayado) en la Prueba de Mann-Whitney, para una cola α = 0,01.

N_2 \ N_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
2	--	--	--	--	--	--	--	--	--	--	--	--	0 / 26	0 / 28	0 / 30	0 / 32	0 / 34	0 / 36	1 / 37	1 / 39
3	--	--	--	--	--	--	0 / 21	0 / 24	1 / 26	1 / 29	1 / 32	2 / 34	2 / 37	2 / 40	3 / 42	3 / 45	4 / 47	4 / 50	4 / 52	5 / 55
4	--	--	--	--	0 / 20	1 / 23	1 / 27	2 / 30	3 / 33	3 / 37	4 / 40	5 / 43	5 / 47	6 / 50	7 / 53	7 / 57	8 / 60	9 / 63	9 / 67	10 / 70
5	--	--	--	0 / 20	1 / 24	2 / 28	3 / 32	4 / 36	5 / 40	6 / 44	7 / 48	8 / 52	9 / 56	10 / 60	11 / 64	12 / 68	13 / 72	14 / 76	15 / 80	16 / 84
6	--	--	--	1 / 23	2 / 28	3 / 33	4 / 38	6 / 42	7 / 47	8 / 52	9 / 57	11 / 61	12 / 66	13 / 71	15 / 75	16 / 80	18 / 84	19 / 89	20 / 94	22 / 98
7	--	--	0 / 21	1 / 27	3 / 32	4 / 38	6 / 43	7 / 49	9 / 54	11 / 59	12 / 65	14 / 70	16 / 75	17 / 81	19 / 86	21 / 91	23 / 96	24 / 102	26 / 107	28 / 112
8	--	--	0 / 24	2 / 30	4 / 36	6 / 42	7 / 49	5 / 55	11 / 61	13 / 67	15 / 73	17 / 79	20 / 84	22 / 90	24 / 96	26 / 102	28 / 108	30 / 114	32 / 120	34 / 126
9	-	--	1 / 26	3 / 33	5 / 40	7 / 47	9 / 54	11 / 61	14 / 67	16 / 74	18 / 81	21 / 87	23 / 94	26 / 100	28 / 107	31 / 113	33 / 120	36 / 126	38 / 133	40 / 140
10	--	--	1 / 29	3 / 37	6 / 44	8 / 52	11 / 59	13 / 67	16 / 74	19 / 81	22 / 88	24 / 96	27 / 103	30 / 110	33 / 117	36 / 124	38 / 132	41 / 139	44 / 146	47 / 153

Tabla 7. Valores críticos de U y de U' (subrayado) en la Prueba de Mann-Whitney, para una cola α = 0,01 (Continuación).

N_2 \ N_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
11	--	--	1 32	4 40	7 48	9 57	12 65	15 73	18 81	22 88	25 96	28 104	31 112	34 120	37 128	41 135	44 143	47 151	50 159	53 167
12	--	--	2 34	5 43	8 52	11 61	14 70	17 79	21 87	24 96	28 104	31 113	35 121	38 130	42 138	46 146	49 155	53 163	56 172	60 180
13	--	0 26	2 37	5 47	9 56	12 66	16 75	20 84	23 94	27 103	31 112	35 121	39 130	43 139	47 148	51 157	55 166	59 175	63 184	67 193
14	--	0 28	2 40	6 50	10 60	13 71	17 81	22 90	26 100	30 110	34 120	38 130	43 139	47 149	51 159	56 168	60 178	65 187	69 197	73 207
15	--	0 30	3 42	7 53	11 64	15 75	19 86	24 96	28 107	33 117	37 128	42 138	47 148	51 159	56 169	61 179	66 189	70 200	75 210	80 220
16	--	0 32	3 45	7 57	12 68	16 80	21 91	26 102	31 113	36 124	41 135	46 146	51 157	56 168	61 179	66 190	71 201	76 212	82 222	87 233
17	--	0 34	4 47	8 60	13 72	18 84	23 96	28 108	33 120	38 132	44 143	49 155	55 166	60 178	66 189	71 201	77 212	82 224	88 234	93 247
18	--	0 36	4 50	9 63	14 76	19 89	24 102	30 114	36 126	41 139	47 151	53 163	59 175	65 187	70 200	76 212	82 224	88 236	94 248	100 260
19	--	1 37	4 53	9 67	15 80	20 94	26 107	32 120	38 133	44 146	50 159	56 172	63 184	69 197	75 210	82 222	88 235	94 248	101 260	107 273
20	--	1 39	5 55	10 70	16 84	22 98	28 112	34 126	40 140	47 153	53 167	60 180	67 193	73 207	80 220	87 233	93 247	100 260	107 273	114 286

Tabla 7. Valores críticos de U y de U' (subrayado) en la Prueba de Mann-Whitney, para dos colas α = 0, 01.

N_2 \ N_1	1	2	3	4	5	6	7	8	9	10
1	--	--	--	--	--	--	--	--	--	--
2	--	--	--	--	--	--	--	--	--	--
3	--	--	--	--	--	--	--	--	0 / 27	0 / 30
4	--	--	--	--	--	0 / 24	0 / 28	1 / 31	1 / 35	2 / 38
5	--	--	--	--	0 / 25	1 / 29	1 / 34	2 / 38	3 / 42	4 / 46
6	--	--	--	0 / 24	1 / 29	2 / 34	3 / 39	4 / 44	5 / 49	6 / 54
7	--	--	--	0 / 28	1 / 34	3 / 39	4 / 45	6 / 50	7 / 56	9 / 61
8	--	--	--	1 / 31	2 / 38	4 / 44	6 / 50	7 / 57	9 / 63	11 / 69
9	--	--	0 / 27	1 / 35	3 / 42	5 / 49	7 / 56	9 / 63	11 / 70	13 / 77
10	--	--	0 / 30	2 / 38	4 / 46	6 / 54	9 / 61	11 / 69	13 / 77	16 / 84

N_2 \ N_1	11	12	13	14	15	16	17	18	19	20
1	--	--	--	--	--	--	--	--	--	--
2	--	--	--	--	--	--	--	--	0 / 38	0 / 40
3	0 / 33	1 / 35	1 / 38	1 / 41	2 / 43	2 / 46	2 / 49	2 / 52	3 / 54	3 / 57
4	2 / 42	3 / 45	3 / 49	4 / 52	5 / 55	5 / 59	6 / 62	6 / 66	7 / 69	8 / 72
5	5 / 50	6 / 54	7 / 58	7 / 63	8 / 67	9 / 71	10 / 75	11 / 79	12 / 83	13 / 87
6	7 / 59	9 / 63	10 / 68	11 / 73	12 / 78	13 / 83	15 / 87	16 / 92	17 / 97	18 / 102
7	10 / 67	12 / 72	13 / 78	15 / 83	16 / 89	18 / 94	19 / 100	21 / 105	22 / 111	24 / 116
8	13 / 75	15 / 81	17 / 87	18 / 94	20 / 100	22 / 106	24 / 112	26 / 118	28 / 124	30 / 130
9	16 / 83	18 / 90	20 / 97	22 / 104	24 / 111	27 / 117	29 / 124	31 / 131	33 / 138	36 / 144
10	18 / 92	21 / 99	24 / 106	26 / 114	29 / 121	31 / 129	34 / 136	37 / 143	39 / 151	42 / 158

Tabla 7. Valores críticos de U y de U' (subrayado) en la Prueba de Mann-Whitney, para dos colas α = 0, 01 (Continuación).

N_2 \ N_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
11	--	--	0 33	2 42	5 50	7 59	10 67	13 75	16 83	18 92	21 100	24 108	27 116	30 124	33 132	36 140	39 148	42 156	45 165	48 172
12	--	--	1 35	3 45	6 54	9 63	12 72	15 81	18 90	21 99	24 108	27 117	31 125	34 134	37 143	41 151	44 160	47 169	51 177	54 186
13	--	--	1 38	3 49	7 58	10 68	13 78	17 87	20 97	24 106	27 116	31 125	34 125	38 144	42 153	45 163	49 172	53 181	56 191	60 200
14	--	--	1 41	4 52	7 63	11 73	15 83	18 94	22 104	26 114	30 124	34 134	38 144	42 154	46 164	50 174	54 184	58 194	63 203	67 213
15	--	--	2 43	5 55	8 67	12 78	16 89	20 100	24 111	29 121	33 132	37 143	42 153	46 164	51 174	55 185	60 195	64 206	69 216	73 227
16	--	--	2 46	5 59	9 71	13 83	18 94	22 106	27 117	31 129	36 140	41 151	45 163	50 174	55 185	60 196	65 207	70 218	74 230	79 241
17	--	--	2 49	6 62	10 75	15 87	19 100	24 112	29 124	34 148	39 148	44 160	49 172	54 184	60 195	65 207	70 219	75 231	81 242	86 254
18	--	--	2 52	6 66	11 79	16 92	21 105	26 118	31 131	37 143	42 156	47 169	53 181	58 194	64 206	70 218	75 231	81 243	87 255	92 268
19	--	0 38	3 54	7 69	12 83	17 97	22 111	28 124	33 138	39 151	45 164	51 177	56 191	63 203	69 216	74 230	81 242	87 255	93 268	99 281
20	--	0 40	3 57	8 72	13 87	18 102	24 116	30 130	36 144	42 158	48 172	54 186	60 200	67 213	73 227	79 241	86 254	92 268	99 281	105 295

Tabla 7. Valores críticos de U y de U’ (subrayado) en la Prueba de Mann-Whitney, para una cola α = 0, 05.

N_2 \ N_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0 19	0 20
2	--	--	--	--	0 10	0 12	0 14	1 15	1 17	1 19	1 21	2 22	2 24	2 26	3 27	3 29	3 31	4 32	4 34	4 36
3	--	--	0 9	0 12	1 14	2 16	2 19	3 21	3 24	4 26	5 28	5 31	6 33	7 35	7 38	8 40	9 42	9 45	10 47	11 49
4	--	--	0 12	1 15	2 18	3 21	4 24	5 27	6 30	7 33	8 36	9 39	10 42	11 45	12 48	14 50	15 53	16 56	17 59	18 62
5	--	0 10	1 14	2 18	4 21	5 25	6 29	8 32	9 36	11 39	12 43	13 47	15 50	16 54	18 57	19 61	20 65	22 68	23 72	25 75
6	--	0 12	2 16	3 21	5 25	7 29	8 34	10 38	12 42	14 46	16 50	17 55	19 59	21 63	23 67	25 71	26 76	28 80	30 84	32 88
7	--	0 14	2 19	4 24	6 29	8 34	11 38	13 43	15 48	17 53	19 58	21 63	24 67	26 72	28 77	30 82	33 86	35 91	37 96	39 101
8	--	1 15	3 21	5 27	8 32	10 38	13 43	15 49	18 54	20 60	23 65	26 70	28 76	31 81	33 87	36 92	39 97	41 103	44 108	47 113
9	--	1 17	3 24	6 30	9 36	12 42	15 48	18 54	21 60	24 66	27 72	30 78	33 84	36 90	39 96	42 102	45 108	48 114	51 120	54 126
10	--	1 19	4 26	7 33	11 39	14 46	17 53	20 60	24 66	27 73	31 79	34 86	37 93	41 99	44 106	48 112	51 119	55 125	58 132	62 138

Tabla 7. Valores críticos de U y de U' (subrayado) en la Prueba de Mann-Whitney, para una cola α = 0, 05 (Continuación).

N_2 \ N_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
11	--	1 21	5 28	8 36	12 43	16 50	19 58	23 65	27 72	31 79	34 87	38 94	42 101	46 108	50 115	54 122	57 130	61 137	65 144	69 151
12	--	2 22	5 31	9 39	13 47	17 55	21 63	26 70	30 78	34 86	38 94	42 102	47 109	51 117	55 125	60 132	64 140	68 148	72 156	77 163
13	--	2 24	6 33	10 42	15 50	19 59	24 67	28 76	33 84	37 93	42 101	47 109	51 118	56 126	61 134	65 143	70 151	75 159	80 167	84 176
14	--	2 26	7 35	11 45	16 54	21 63	26 72	31 81	36 90	41 99	46 108	51 117	56 126	61 135	66 144	71 153	77 161	82 170	87 179	92 188
15	--	3 27	7 38	12 48	18 57	23 67	28 77	33 87	39 96	44 106	50 115	55 125	61 134	66 144	72 153	77 163	83 172	88 182	94 191	100 200
16	--	3 29	8 40	14 50	19 61	25 71	30 82	36 92	42 102	48 112	54 122	60 132	65 143	71 153	77 163	83 173	89 183	95 193	101 203	107 213
17	--	3 31	9 42	15 53	20 65	26 76	33 86	39 97	45 108	51 119	57 130	64 140	70 151	77 161	83 172	89 183	96 193	102 204	109 214	115 225
18	--	4 32	9 45	16 56	22 68	28 80	35 91	41 103	48 114	55 123	61 137	68 148	75 159	82 170	88 182	95 193	102 204	109 215	116 226	123 237
19	0 19	4 34	10 47	17 59	23 72	30 84	37 96	44 108	51 120	58 132	65 144	72 156	80 167	87 179	94 191	101 203	109 214	116 226	123 238	130 250
20	0 20	4 36	11 49	18 62	25 75	32 88	39 101	47 113	54 126	62 138	69 151	77 163	84 176	92 188	100 200	107 213	115 225	123 237	130 250	138 262

Tabla 7. Valores críticos de U y de U’ (subrayado) en la Prueba de Mann-Whitney, para dos colas α = 0, 05.

N_2 \ N_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
2	--	--	--	--	--	--	--	0 16	0 18	0 20	0 22	1 23	1 25	1 27	1 29	1 31	2 32	2 34	2 36	2 38
3	--	--	--	--	0 15	1 17	1 20	2 22	2 25	3 27	3 30	4 32	4 35	5 37	5 40	6 42	6 45	7 47	7 50	8 52
4	--	--	--	0 16	1 19	2 22	3 25	4 28	4 32	5 35	6 38	7 41	8 44	9 47	10 50	11 53	11 57	12 60	13 63	13 67
5	--	--	0 15	1 19	2 23	3 27	5 30	6 34	7 38	8 42	9 46	11 49	12 53	13 57	14 61	15 65	17 68	18 72	19 76	20 80
6	--	--	1 17	2 22	3 27	5 31	6 36	8 40	10 44	11 49	13 53	14 58	16 62	17 67	19 71	21 75	22 80	24 84	25 89	27 93
7	--	--	1 20	3 25	5 30	6 36	8 41	10 46	12 51	14 56	16 61	18 66	20 71	22 76	24 81	26 86	28 91	30 96	32 101	34 106
8	--	0 16	2 22	4 28	6 34	8 40	10 46	13 51	15 57	17 63	19 69	22 74	24 80	26 86	29 91	31 97	34 102	36 108	38 111	41 119
9	--	0 18	2 25	4 32	7 38	10 44	12 51	15 57	17 64	20 70	23 76	26 82	28 89	31 95	34 101	37 107	39 114	42 120	45 126	48 132
10	--	0 20	3 27	5 35	8 42	11 49	14 56	17 63	20 70	23 77	26 84	29 91	33 97	36 104	39 111	42 118	45 125	48 132	52 138	55 145

Tabla 7. Valores críticos de U y de U' (subrayado) en la Prueba de Mann-Whitney, para dos colas α = 0, 05 (Continuación).

N_2 \ N_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
11	--	0 22	3 30	6 38	9 46	13 53	16 61	19 69	23 76	26 84	30 91	33 99	37 106	40 114	44 121	47 129	51 136	55 143	58 151	62 158
12	--	1 23	4 32	7 41	11 49	14 58	18 66	22 74	26 82	29 91	33 99	37 107	41 115	45 123	49 131	53 139	57 147	61 155	65 163	69 171
13	--	1 25	4 35	8 44	12 53	16 62	20 71	24 80	28 89	33 97	37 106	41 115	45 124	50 132	54 141	59 149	63 158	67 167	72 175	76 184
14	--	1 27	5 37	9 47	13 51	17 67	22 76	26 86	31 95	36 104	40 114	45 123	50 132	55 141	59 151	64 160	67 171	74 178	78 188	83 197
15	--	1 29	5 40	10 50	14 61	19 71	24 81	29 91	34 101	39 111	44 121	49 131	54 141	59 151	64 161	70 170	75 180	80 190	85 200	90 210
16	--	1 31	6 42	11 53	15 65	21 75	26 86	31 97	37 107	42 118	47 129	53 139	59 149	64 160	70 170	75 181	81 191	86 202	92 212	90 222
17	--	2 32	6 45	11 57	17 68	22 80	28 91	34 102	39 114	45 125	51 136	57 147	63 158	67 171	75 180	81 191	87 202	93 213	99 224	105 235
18	--	2 34	7 47	12 60	18 72	24 84	30 96	36 108	42 120	48 132	55 143	61 155	67 167	74 178	80 190	86 202	93 213	99 225	106 236	112 248
19	--	2 36	7 50	13 63	19 76	25 89	32 101	38 114	45 126	52 138	58 151	65 163	72 175	78 188	85 200	92 213	99 224	106 236	113 248	119 261
20	--	2 38	8 52	13 67	20 80	27 93	34 106	41 119	48 132	55 145	62 158	69 171	76 184	83 197	90 210	98 222	105 235	112 248	119 261	127 273

Tabla 8a. Valores críticos de D en la prueba de Kolmogorov-Smirnov para dos muestras independientes $n_1 = n_2$.

Una Cola	α =	0,10	0,05	0,025	0,01	0,005
Dos Colas	**α =**	**0,20**	**0,10**	**0,05**	**0,02**	**0,01**
n = 3		2/3	2/3			
4		3/4	3/4	3/4		
5		3/5	3/5	4/5	4/5	4/5
6		3/6	4/6	4/6	5/6	5/6
7		4/7	4/7	5/7	5/7	5/7
8		4/8	4/8	5/8	5/8	6/8
9		4/9	5/9	5/9	6/9	6/9
10		4/10	5/10	6/10	6/10	7/10
11		5/11	5/11	6/11	7/11	7/11
12		5/12	5/12	6/12	7/12	7/12
13		5/13	6/13	6/13	7/13	8/13
14		5/14	6/14	7/14	7/14	8/14
15		5/15	6/15	7/15	8/15	8/15
16		6/16	6/16	7/16	8/16	9/16
17		6/17	7/17	7/17	8/17	9/17
18		6/18	7/18	8/18	9/18	9/19
19		6/19	7/19	8/19	9/19	9/19
20		6/20	7/20	8/20	9/20	10/20
21		6/21	7/21	8/21	9/21	10/21
22		7/22	8/22	8/22	10/22	10/22
23		7/23	8/23	9/23	10/23	10/23
24		7/24	8/24	9/24	10/24	11/24
25		7/25	8/25	9/25	10/25	11/25
26		7/26	8/26	9/26	10/26	11/26
27		7/27	8/27	9/27	11/27	11/27
28		8/28	9/28	10/28	11/28	12/28
29		8/29	9/29	10/29	11/29	12/29
30		8/30	9/30	10/30	11/30	12/30
31		8/31	9/31	10/31	11/31	12/31
32		8/32	9/32	10/32	12/32	12/32
34		8/34	10/34	11/34	12/34	13/34
36		9/36	10/36	11/36	12/36	13/36
38		9/38	10/38	11/38	13/38	14/38
40		9/40	10/40	12/40	13/40	14/40

Tabla 8b. Valores críticos de D en la Prueba de Kolmogorov-Smirnov para dos muestras independientes $n_1 \neq n_2$.

Una Cola			α =	0,10	0,05	0,025	0,01	0,005
Dos Colas			**α =**	**0,20**	**0,10**	**0,05**	**0,02**	**0,01**
$n_1 = 1$	$n_2 =$	9		17/18				
		10		9/10				
$n_1 = 2$	$n_2 =$	3		5/6				
		4		3/4				
		5		4/5	4/5			
		6		5/6	5/6			
		7		5/7	6/7			
		8		3/4	7/8	7/8		
		9		7/9	8/9	8/9		
		10		7/10	4/5	9/10		
$n_1 = 3$	$n_2 =$	4		3/4	3/4			
		5		2/3	4/5	4/5		
		6		2/3	2/3	5/6		
		7		2/3	5/7	6/7	6/7	
		8		5/8	3/4	3/4	7/8	
		9		2/3	2/3	7/9	8/9	8/9
		10		3/5	7/10	4/5	9/10	9/10
		12		7/12	2/3	3/4	5/6	11/12
$n_1 = 4$	$n_2 =$	5		3/5	3/4	4/5	4/5	
		6		7/12	2/3	3/4	5/6	5/6
		7		17/28	5/7	3/4	6/7	6/7
		8		5/8	5/8	3/4	7/8	7/8
		9		5/9	2/9	3/4	7/9	8/9
		10		11/20	13/20	7/10	4/5	4/5
		12		7/12	2/3	2/3	3/4	5/6
		16		9/16	5/8	11/16	3/4	13/16
$n_1 = 5$	$n_2 =$	6		3/5	2/3	2/3	5/6	5/6
		7		4/7	29/35	5/7	29/35	6/7
		8		11/20	5/8	27/40	4/5	4/5
		9		5/9	3/5	31/45	7/9	4/5
		10		1/2	3/5	7/10	7/10	4/5
		15		8/15	3/5	2/3	11/15	11/15
		20		1/2	11/20	3/5	7/10	3/4

Tabla 8b. Valores críticos de D en la Prueba de Kolmogorov-Smirnov para dos muestras independientes $n_1 \neq n_2$ (continuación).

Una Cola		α =	0,10	0,05	0,025	0,01	0,005
Dos Colas		**α =**	**0,20**	**0,10**	**0,05**	**0,02**	**0,01**
$n_1 = 6$	$n_2 =$	7	23/42	4/7	29/42	5/7	5/6
		8	1/2	7/12	2/3	3/4	3/4
		9	1/2	5/9	2/3	13/18	7/9
		10	1/2	17/30	19/30	7/10	11/15
		12	1/2	7/12	7/12	2/3	3/4
		18	4/9	5/9	11/18	2/3	13/18
		24	11/24	1/2	7/12	5/8	2/3
$n_1 = 7$	$n_2 =$	8	27/56	33/56	5/8	41/56	3/4
		9	31/63	5/9	40/63	5/7	47/63
		10	33/70	39/70	43/70	7/10	5/7
		14	3/7	1/2	4/7	9/14	5/7
		28	3/7	13/28	15/28	17/28	9/14
$n_1 = 8$	$n_2 =$	9	4/9	13/24	5/8	2/3	3/4
		10	19/40	21/40	23/40	27/40	7/10
		12	11/24	1/2	7/12	5/8	2/3
		16	7/16	1/2	9/16	5/8	5/8
		32	13/32	7/16	1/2	9/16	19/32
$n_1 = 9$	$n_2 =$	10	7/15	1/2	26/45	2/3	31/45
		12	4/9	1/2	5/9	11/18	2/3
		15	19/45	22/45	8/15	3/5	29/45
		18	7/18	4/9	1/2	5/9	11/18
		36	19/36	5/12	17/36	19/36	5/9
$n_1 = 10$	$n_2 =$	15	2/5	7/15	1/2	17/30	19/30
		20	2/5	9/20	1/2	11/20	3/5
		40	7/20	2/5	9/20	1/2	11/20
$n_1 = 12$	$n_2 =$	15	23/60	9/20	1/2	11/20	7/12
		16	3/8	7/16	23/48	13/24	7/12
		18	13/36	5/12	17/36	19/36	5/9
		20	11/30	5/12	7/15	31/60	17/30
$n_1 = 15$	$n_2 =$	20	7/20	2/5	13/30	29/60	31/60
$n_1 = 16$	$n_2 =$	20	27/80	31/80	17/40	19/40	41/80

Tabla 9. Valores críticos de X_F^2 de Friedman.

K	N	α ≤ 0,10	α ≤ 0,05	α ≤ 0,01
3	3	6,00	6,00	-
	4	6,00	6,50	8,00
	5	5,20	6,40	8,40
	6	5,33	7,00	9,00
	7	5,43	7,14	8,86
	8	5,25	6,25	9,00
	9	5,56	6,22	8,67
	10	5,00	6,20	9,60
	11	4,91	6,54	8,91
	12	5,17	6,17	8,67
	13	4,77	6,00	9,39
	∞	4,61	5,99	9,21
4	2	6,00	6,00	-
	3	6,60	7,40	8,60
	4	6,30	7,80	9,60
	5	6,36	7,80	9,96
	6	6,40	7,60	10,00
	7	6,26	7,80	10,37
	8	6,30	7,50	10,35
	∞	6,25	7,82	11,34
5	3	7,47	8,53	10,13
	4	7,60	8,80	11,00
	5	7,68	8,96	11,52
	∞	7,78	9,49	13,28

Tabla 10. Valores de z para #c número de comparaciones múltiples.

		α					
# c	**Bidireccional**	**0,30**	**0,25**	**0,20**	**0,15**	**0,10**	**0,05**
	Unidireccional	**0,15**	**0,125**	**0,10**	**0,075**	**0,05**	**0,025**
1		1,036	1,150	1,282	1,440	1,645	1,960
2		1,440	1,534	1,645	1,780	1,960	2,241
3		1,645	1,732	1,834	1,960	2,128	2,394
4		1,780	1,863	1,960	2,080	2,241	2,498
5		1,881	1,960	2,054	2,170	2,326	2,576
6		1,960	2,037	2,128	2,241	2,394	2,638
7		2,026	2,100	2,189	2,300	2,450	2,690
8		2,080	2,154	2,241	2,350	2,498	2,734
9		2,128	2,200	2,287	2,394	2,539	2,773
10		2,170	2,241	2,326	2,432	2,576	2,807
11		2,208	2,278	2,362	2,467	2,608	2,838
12		2,241	2,301	2,394	2,498	2,638	2,866
15		2,326	2,394	2,475	2,576	2,713	2,935
21		2,450	2,515	2,593	2,690	2,823	3,038
28		2,552	2,615	2,690	2,785	2,913	3,125

Tabla 11. Valores críticos del estadístico L de Page.

	K = 3			*K* = 4			*K* = 5			*K* = 6		
	α			α			α			α		
N	**0,05**	**0,01**	**0,001**	**0,05**	**0,01**	**0,001**	**0,05**	**0,01**	**0,001**	**0,05**	**0,01**	**0,001**
2	28			58	60		103	106	109	166	173	178
3	41	42		84	87	89	150	155	160	244	252	260
4	54	55	56	111	114	117	197	204	210	321	331	341
5	66	68	70	137	141	145	244	251	259	397	409	420
6	79	81	83	163	167	172	291	299	307	474	486	499
7	91	93	96	189	193	198	338	346	355	550	563	577
8	104	106	109	214	220	225	384	393	403	625	640	655
9	116	119	121	240	246	252	431	441	451	701	717	733
10	128	131	134	266	272	278	477	487	499	777	793	811
11	141	144	147	292	298	305	523	534	546	852	869	888
12	153	156	160	317	324	331	570	581	593	928	949	965
13	165	169	172									
14	178	181	185									
15	190	194	197									
16	202	206	210									
17	215	218	223									
18	227	231	235									
19	239	243	248									
20	251	256	260									

Tabla 11. Valores críticos del estadístico L de Page (Continuación).

	K = 7			*K = 8*			*K = 9*			*K = 10*		
		α			α			α			α	
N	**0,05**	**0,01**	**0,001**	**0,05**	**0,01**	**0,001**	**0,05**	**0,01**	**0,001**	**0,05**	**0,01**	**0,001**
2	252	261	269	362	376	388	500	520	544	670	696	726
3	370	382	394	532	549	567	736	761	790	987	1019	1056
4	487	501	516	701	722	743	971	999	1032	1301	1339	1382
5	603	620	637	869	893	917	1204	1236	1273	1614	1656	1704
6	719	737	757	1037	1063	1090	1436	1472	1512	1927	1972	2025
7	835	855	876	1204	1232	1262	1668	1706	1750	2238	2288	2344
8	950	972	994	1371	1401	1433	1900	1940	1987	2549	2602	2662
9	1065	1088	1113	1537	1569	1603	2131	2174	2223	2859	2915	2980
10	1180	1205	1230	1703	1736	1773	2361	2407	2459	3169	3228	3296
11	1295	1321	1348	1868	1905	1943	2592	2639	2694	3478	3541	3612
12	1410	1437	1465	2035	2072	2112	2822	2872	2929	3788	3852	3927

Tabla 12. Valores críticos de H de Kruskal-Wallis.

Tamaños de las muestras			α				
n_1	n_2	n_3	0,10	0,05	0,01	0,005	0,001
2	2	2	4,25				
3	2	1	4,29				
3	2	2	4,71	4,71			
3	3	1	4,57	5,14			
3	3	2	4,56	5,36			
3	3	3	4,62	5,60	7,20	7,20	
4	2	1	4,50				
4	2	2	4,46	5,33			
4	3	1	4,06	5,21			
4	3	2	4,51	5,44	6,44	7,00	
4	3	3	4,71	5,73	6,75	7,32	8,02
4	4	1	4,17	4,97	6,67		
4	4	2	4,55	5,45	7,04	7,28	
4	4	3	4,55	5,60	7,14	7,59	8,32
4	4	4	4,65	5,69	7,66	8,00	8,65
5	2	1	4,20	5,00			
5	2	2	4,36	5,16	6,53		
5	3	1	4,02	4,96			
5	3	2	4,65	5,25	6,82	7,18	
5	3	3	4,53	5,65	7,08	7,51	8,24
5	4	1	3,99	4,99	6,95	7,36	
5	4	2	4,54	5,27	7,12	7,57	8,11
5	4	3	4,55	5,63	7,44	7,91	8,50
5	4	4	4,62	5,62	7,76	8,14	9,00
5	5	1	4,11	5,13	7,31	7,75	
5	5	2	4,62	5,34	7,27	8,13	8,68
5	5	3	4,54	5,71	7,54	8,24	9,06
5	5	4	4,53	5,64	7,77	8,37	9,32
5	5	5	4,56	5,78	7,98	8,72	9,68
Muestras grandes			4,61	5,99	9,21	10,60	13,82

Tabla 13. Valores críticos de J de Jonckheere.

Tamaño de las muestras			α			
			0,10	0,05	0,01	0,005
2	2	2	10	11	12	- -
2	2	3	13	14	15	16
2	2	4	16	17	19	20
2	2	5	18	20	22	23
2	2	6	21	23	25	27
2	2	7	24	26	29	30
2	2	8	27	29	32	33
2	3	3	16	18	19	20
2	3	4	20	21	23	25
2	3	5	23	25	27	29
2	3	6	26	28	31	33
2	3	7	30	32	35	37
2	3	8	33	35	39	41
2	4	4	24	25	28	29
2	4	5	27	29	33	34
2	4	6	31	34	37	39
2	4	7	35	38	42	44
2	4	8	39	42	46	49
2	5	5	32	34	38	40
2	5	6	36	39	43	45
2	5	7	41	44	48	51
2	5	8	45	48	53	56
2	6	6	42	44	49	51
2	6	7	47	50	55	57
2	6	8	52	55	61	64
2	7	7	52	56	61	64
2	7	8	58	62	68	71
2	8	8	64	68	75	78
3	3	3	20	22	24	25
3	3	4	24	26	29	30

Tabla 13. Valores críticos de J de Jonckheere (Continuación).

Tamaño de las muestras			α			
			0,10	**0,05**	**0,01**	**0,005**
3	3	5	28	30	33	35
3	3	6	32	34	38	40
3	3	7	36	38	42	44
3	3	8	40	42	47	49
3	4	4	29	31	34	36
3	4	5	33	35	39	41
3	4	6	38	40	44	46
3	4	7	42	45	49	52
3	4	8	47	50	55	57
3	5	5	38	41	45	47
3	5	6	43	46	51	53
3	5	7	48	51	57	59
3	5	8	53	57	63	65
3	6	6	49	52	57	60
3	6	7	54	58	64	67
3	6	8	60	64	70	73
3	7	7	61	64	71	74
3	7	8	67	71	78	81
3	8	8	74	78	86	89
4	4	4	34	36	40	42
4	4	5	39	41	45	48
4	4	6	44	47	51	54
4	4	7	49	52	57	60
4	4	8	54	57	63	66
4	5	5	44	47	52	55
4	5	6	50	53	58	61
4	5	7	56	59	65	68
4	5	8	61	65	71	75
4	6	6	56	60	66	69
4	6	7	62	66	73	76
4	6	8	68	73	80	83
4	7	7	69	73	81	84
4	7	8	76	80	88	92
4	8	8	83	88	97	100

Tabla 13. Valores críticos de J de Jonckheere (Continuación).

Tamaño de las muestras	α			
	0,10	0,05	0,01	0,005
5 5 5	50	54	59	62
5 5 6	57	60	66	69
5 5 7	63	67	73	76
5 5 8	69	73	80	84
5 6 6	63	67	74	77
5 6 7	70	74	82	85
5 6 8	77	81	89	93
5 7 7	77	82	90	94
5 7 8	85	89	98	102
5 8 8	92	98	107	111
6 6 6	71	75	82	86
6 6 7	78	82	91	94
6 6 8	85	90	99	103
6 7 7	86	91	100	103
6 7 8	94	99	109	113
6 8 8	102	108	118	122
7 7 7	94	99	109	113
7 7 8	102	108	119	123
7 8 8	111	117	129	133
8 8 8	121	127	139	144
2 2 2 2	18	19	21	22
2 2 2 2 2	28	30	33	34
2 2 2 2 2 2	40	43	46	49
3 3 3 3	37	39	43	45
3 3 3 3 3	58	62	68	70
3 3 3 3 3 3	85	89	97	101
4 4 4 4	63	66	72	76
4 4 4 4 4	100	105	115	119
4 4 4 4 4 4	146	153	166	171
5 5 5 5	95	100	109	113
5 5 5 5 5	152	159	173	178
5 5 5 5 5 5	223	233	251	258
6 6 6 6	134	140	153	158
6 6 6 6 6	215	225	243	250
6 6 6 6 6 6	316	329	353	362

Tabla 14. Valores críticos del coeficiente de correlación de Spearman.

N	α 0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005	unidireccional
	α 0,50	0,20	0,10	0,05	0,02	0,01	0,005	0,002	0,001	bidireccional
4	0,600	1,000	1,000							
5	0,500	0,800	0,900	1,000	1,000					
6	0,371	0,657	0,829	0,886	0,943	1,000	1,000			
7	0,321	0,571	0,714	0,786	0,893	0,929	0,964	1,000	1,000	
8	0,310	0,524	0,643	0,738	0,833	0,881	0,905	0,952	0,976	
9	0,267	0,483	0,600	0,700	0,783	0,833	0,867	0,917	0,933	
10	0,248	0,455	0,564	0,648	0,745	0,794	0,830	0,879	0,903	
11	0,236	0,427	0,536	0,618	0,709	0,755	0,800	0,845	0,873	
12	0,224	0,406	0,503	0,587	0,671	0,727	0,776	0,825	0,860	
13	0,209	0,385	0,484	0,560	0,648	0,703	0,747	0,802	0,835	
14	0,200	0,367	0,464	0,538	0,622	0,675	0,723	0,776	0,811	
15	0,189	0,354	0,443	0,521	0,604	0,654	0,700	0,754	0,786	
16	0,182	0,341	0,429	0,503	0,582	0,635	0,679	0,732	0,765	
17	0,176	0,328	0,414	0,485	0,566	0,615	0,662	0,713	0,748	
18	0,170	0,317	0,401	0,472	0,550	0,600	0,643	0,695	0,728	
19	0,165	0,309	0,391	0,460	0,535	0,584	0,628	0,677	0,712	
20	0,161	0,299	0,380	0,477	0,520	0,570	0,612	0,662	0,696	
21	0,156	0,292	0,370	0,435	0,508	0,556	0,599	0,648	0,681	
22	0,152	0,284	0,361	0,425	0,496	0,544	0,586	0,634	0,667	
23	0,148	0,278	0,353	0,415	0,486	0,532	0,573	0,622	0,654	
24	0,144	0,271	0,344	0,406	0,476	0,521	0,562	0,610	0,642	
25	0,142	0,265	0,337	0,398	0,466	0,511	0,551	0,598	0,630	
26	0,138	0,259	0,331	0,390	0,457	0,501	0,541	0,587	0,619	
27	0,136	0,255	0,324	0,382	0,448	0,491	0,531	0,577	0,608	

Tabla 14. Valores críticos del coeficiente de correlación de Spearman (Continuación).

N	α 0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005	unidireccional
	α 0,50	0,20	0,10	0,05	0,02	0,01	0,005	0,002	0,001	bidireccional
28	0,133	0,250	0,317	0,375	0,440	0,483	0,522	0,567	0,598	
29	0,130	0,245	0,312	0,368	0,433	0,475	0,513	0,558	0,589	
30	0,128	0,240	0,306	0,362	0,425	0,467	0,504	0,549	0,580	
31	0,126	0,236	0,301	0,356	0,418	0,459	0,496	0,541	0,571	
32	0,124	0,232	0,296	0,350	0,412	0,452	0,489	0,533	0,563	
33	0,121	0,229	0,291	0,345	0,405	0,446	0,482	0,525	0,554	
34	0,120	0,225	0,287	0,340	0,399	0,439	0,475	0,517	0,547	
35	0,118	0,222	0,283	0,335	0,394	0,433	0,468	0,510	0,539	
36	0,116	0,219	0,279	0,330	0,388	0,427	0,462	0,504	0,533	
37	0,114	0,216	0,275	0,325	0,383	0,421	0,456	0,497	0,526	
38	0,113	0,212	0,271	0,321	0,378	0,415	0,450	0,491	0,519	
39	0,111	0,210	0,267	0,317	0,373	0,410	0,444	0,485	0,513	
40	0,110	0,207	0,264	0,313	0,368	0,405	0,439	0,479	0,507	
41	0,108	0,204	0,261	0,309	0,364	0,400	0,433	0,473	0,501	
42	0,107	0,202	0,257	0,305	0,359	0,395	0,428	0,468	0,495	
43	0,105	0,199	0,254	0,301	0,355	0,391	0,423	0,463	0,490	
44	0,104	0,197	0,251	0,298	0,351	0,386	0,419	0,458	0,484	
45	0,103	0,194	0,248	0,294	0,347	0,382	0,414	0,453	0,479	
46	0,102	0,192	0,246	0,291	0,343	0,378	0,410	0,448	0,474	
47	0,101	0,190	0,243	0,288	0,340	0,374	0,405	0,443	0,469	
48	0,100	0,188	0,240	0,285	0,336	0,370	0,401	0,439	0,465	
49	0,098	0,186	0,238	0,282	0,333	0,366	0,397	0,434	0,460	
50	0,097	0,184	0,235	0,279	0,329	0,363	0,393	0,430	0,456	

ISBN: 978-980-318-284-7