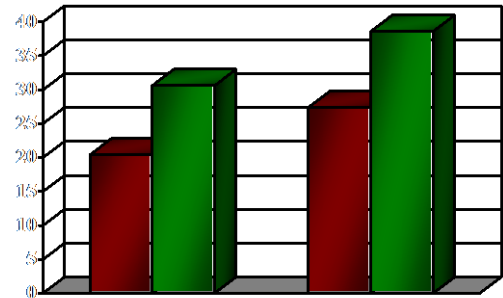


Apuntes de Estadística Inferencial

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_{D\bar{x}}}$$

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$$

$$\sigma = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{n}} = \sqrt{\frac{\sum f_i x_i^2}{n} - \bar{x}^2}$$



$$X^2 = \sum \frac{(t_o - t_e)^2}{t_e}$$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$U = n_1 n_2 + \frac{n_x (n_x + 1)}{2} - \sum R_x$$

Francisco Juárez García
Jorge A. Villatoro Velázquez
Elsa Karina López Lugo

Primera Edición, 2002.
© 2002 Francisco Juárez García
Instituto Nacional de Psiquiatría Ramón de la Fuente,
Dirección de Investigación Epidemiológicas y Psicosociales
Calzada México Xochimilco 101,
Col. San Lorenzo Huipulco,
14370 México, D. F.
Tel.: 56 55 28 11 ext. 202 fax: 55 13 34 46
e-mail: pacojg@imp.edu.mx

Reservados todos los derechos
ISBN: 968-7652-41-1
Impreso en México / Printed in Mexico

Agradeceremos todos los comentarios relativos a esta guía, estos deben ser enviados a la dirección arriba mencionada.

El presente trabajo debe ser citado como sigue:

Juárez, F., Villatoro, J. A. y López, E. K. (2002). *Apuntes de Estadística Inferencial*. México, D. F.: Instituto Nacional de Psiquiatría Ramón de la Fuente.

Contenido

Presentación	i
Agradecimientos	ii
Índice de Tablas	iii
Conceptos Básicos	1
Medición	2
Niveles de Medición	2
Estadística Descriptiva	4
Curvas de Distribución de Frecuencias	5
Distribución Normal o Curva Normal (Campana de Gauss)	5
Calificaciones z o estandarizadas	6
Estadística Inferencial	8
Estadística Inferencial Univariada	8
Tipos de problema	8
Estadística Multivariada	9
Estadística Paramétrica	10
Estadística No Paramétrica	10
Planteamiento de hipótesis estadística	10
Hipótesis de Problemas de Comparación	10
Hipótesis de Problemas de Asociación	11
Construcción de hipótesis de acuerdo al nivel de medición	12
¿Cuál es el sentido del nivel de significancia o la probabilidad?	13
Decisión estadística	13
Grados de libertad	15
Reglas de decisión	15
Estadística Inferencial Univariada: Pruebas Estadísticas Seleccionadas	17
Cómo Presentar Resultados Análisis Estadísticos	17
Criterios importantes para seleccionar el tipo de prueba en un estudio	18
Pruebas de Comparación	19
Prueba t de Student	19
Prueba t de Student para grupos independientes	19
La prueba t para muestras independientes en el paquete estadístico SPSS	21
Reporte de resultados de la prueba t	22
Prueba t de Student para muestras relacionadas	23
La prueba t de Student para muestras relacionadas en el paquete estadístico SPSS	25
Reporte de resultados de la prueba t para muestras relacionadas	25
Prueba U de Mann-Whitney	27
La prueba U de Mann-Whitney en el paquete estadístico SPSS	29
Reporte de resultados de la prueba U de Mann Whitney	30

Prueba de Wilcoxon	31
La prueba de Wilcoxon en el paquete estadístico SPSS	33
Reporte de resultados de la prueba de Wilcoxon	34
Prueba de Chi Cuadrada (X^2)	35
Prueba de Chi cuadrada para una tabla de 2 X 2	37
La Chi cuadrada en el paquete SPSS	38
Restricciones de la Chi cuadrada	40
Reporte de resultados de la prueba X^2	40
Prueba del Cambio de McNemar de dos Mediciones Dicotómicas	41
McNemar en el paquete SPSS	42
Reporte de resultados de la prueba de McNemar	43
Análisis de Varianza (Prueba F)	45
Comparaciones Posteriores a F	48
La prueba F de una vía en el paquete estadístico SPSS	49
Pruebas post hoc	49
Reporte de resultados de la prueba F	50
Análisis de Varianza Unifactorial por Rangos de Kruskal-Wallis	51
Comparaciones múltiples entre tratamientos	54
La prueba de Kruskal-Wallis en el paquete estadístico SPSS	55
Reporte de resultados de la prueba de Kruskal-Wallis	56
Análisis de Varianza de dos Clasificaciones por Rangos de Friedman	57
Comparaciones múltiples entre mediciones	60
La prueba de Friedman en el paquete estadístico SPSS	61
Reporte de resultados de la prueba de Friedman	61
La Prueba de Q de Cochran	63
Q de Cochran en el paquete SPSS	64
Reporte de resultados de la prueba Q de Cochran	65
Pruebas de Asociación	67
Coeficiente de Correlación Producto-Momento de Pearson	69
La prueba de Pearson en el paquete estadístico SPSS	70
Reporte de resultados de la correlación Pearson	71
Coeficiente de Correlación de Spearman de Rangos Ordenados	73
La prueba de Spearman en el paquete estadístico SPSS	76
Reporte de resultados de la correlación Spearman	76
Coeficiente de Correlación Phi	77
La correlación Phi en el paquete SPSS	78
Reporte de resultados de la correlación Phi	79
Referencias Bibliográficas	81

Presentación

El presente trabajo va dirigido a psicólogos en formación y para todos aquellos interesados en la investigación en psicología y áreas afines, donde la estadística es una de las principales herramientas.

Fue escrito para servir de apoyo al profesor en cursos introductorios de estadística inferencial, sin embargo, está diseñado como una guía de autoaprendizaje y material de consulta, ya que se explica la forma de seleccionar la prueba estadística de acuerdo con los problemas de investigación que se plantean en las áreas psicológica y social.

Se exponen además los pasos necesarios para obtener diversas pruebas estadísticas univariadas e interpretar los resultados conseguidos, de manera que sea posible responder a las preguntas de investigación planteadas. Se incluyen ejemplos de los datos que brinda el paquete estadístico SPSS versión 10 y una sección sobre cómo reportar los hallazgos.

Espero que este esfuerzo sea de especial utilidad para quienes inician su formación y para los profesores que los introducen en esta área.

Francisco Juárez
Diciembre de 2001

Agradecimientos

Al grupo del Seminario de Estadística de la maestría y del doctorado en Ciencias de la Salud Mental Pública (febrero a agosto de 2001), de la Facultad de Medicina, UNAM e Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, del maestro Jorge Villatoro, por su aportación a la sección de Conceptos Básicos:

Consuelo García Andrade, José Javier Elizalde González, Juan Luis Durán Casas, María de Lourdes Eriksen Persson, María Elena Rivera Heredia, María Teresa Espinosa Meléndez, Martha Cordero Oropeza, Nivaldo Linares Pérez y Yara Amelia Brom Valladares.

A las Psicólogas Pamela Verónica Esquivel Reyes y Daniela Verderi Muñuzuri, integrantes del grupo 33 de Estadística Inferencial, de la Licenciatura en Psicología (agosto a diciembre de 2001), Universidad Intercontinental, por sus contribuciones al mejoramiento del material.

A la maestra Elsa Karina López Lugo, la licenciada Nora Angélica Martínez Vélez y la licenciada Lucía Vázquez Pérez, por brindar su tiempo al hacer la revisión de este documento y ofrecer sus ideas para ser incluidas en el mismo.

Muy especialmente al maestro Jorge Ameth Villatoro Velázquez, por sus conocimientos y comentarios, que forman parte de estos apuntes.

Índice de Tablas*

Tabla de Valores Críticos para la Prueba t	21
Tabla de Valores Críticos para U	29
Valores críticos de T^+ para la prueba de rangos asignados de Wilcoxon de una cola	32
Valores críticos de Chi Cuadrada a los Niveles de significancia de 0.05 y 0.01	37
Valores críticos de 5% y 1 % para la distribución de F	47
Valores críticos para el análisis de varianza unifactorial por rangos de Kruskal-Wallis (KW)	53
Valores críticos de z para el número de comparaciones múltiples con hipótesis bidireccional	54
Valores críticos para la prueba estadística de análisis de varianza de dos clasificaciones por rangos de Friedman (F_r)	59
Valores críticos de z para el número de comparaciones múltiples con hipótesis bidireccional	60
Valores críticos de r para diferentes niveles de significancia	70
Valores críticos del coeficiente de correlación de Spearman (r_s) de rangos ordenados, para una prueba bidireccional	75

* Se presentan sólo ejemplos de cómo localizar los valores críticos para cada prueba, para revisar las tablas completas se puede acudir a:

Downie, M. y Heath, R. (1973). *Métodos estadísticos aplicados*. México: Harper & Row Latinoamericana.

Siegel, S. (1990). *Estadística no paramétrica* (3ra. ed.). México: Editorial Trillas.

Conceptos Básicos

La estadística es un conjunto de procedimientos para reunir, clasificar, codificar, procesar, analizar y resumir información numérica adquirida sistemáticamente (Ritchey, 2002). Permite hacer inferencias a partir de una muestra para extrapolarlas a una población.

Aunque normalmente se asocia a muchos cálculos y operaciones aritméticas, y aunque las matemáticas están involucradas, en su mayor parte sus fundamentos y uso apropiado pueden dominarse sin hacer referencia a habilidades matemáticas avanzadas. De hecho se trata de una forma de ver la realidad basada en el análisis cuidadoso de los hechos (Ritchey, 2002). Es necesaria sin embargo la sistematización para reducir el efecto que las emociones y las experiencias individuales puedan tener al interpretar esa realidad.

De esta manera la estadística se relaciona con el método científico complementándolo como herramienta de análisis y, aunque la investigación científica no requiere necesariamente de la estadística, ésta valida muchos de los resultados cuantitativos derivados de la investigación.

La obtención del conocimiento debe hacerse de manera *sistemática* por lo que deben planearse todos los pasos que llevan desde el planteamiento de un problema, pasando por la elaboración de hipótesis y la manera en que van a ser probadas; la selección de sujetos (muestreo), los escenarios, los instrumentos que se utilizarán para obtener los datos, definir el procedimiento que se seguirá para esto último, los controles que se deben hacer para asegurar que las intervenciones son las causas más probables de los cambios esperados (diseño); hasta la elección del plan de análisis idóneo para el tipo de datos que se están obteniendo, es aquí donde la estadística entra en el estudio, aunque pueden existir otras herramientas de análisis si se está haciendo una investigación de corte cualitativo.

Una buena planeación permitirá que los resultados puedan ser *reproducidos*, mediante la comprobación empírica, por cualquier investigador interesado en refutar o comprobar las conclusiones que se hagan del estudio. De esta manera también se logrará la *predicción* de los fenómenos que se están estudiando, ayudando a conocer y prevenir los problemas sociales e individuales que forman parte del objeto de estudio de la psicología.

El tratamiento de los datos de la investigación científica tiene varias etapas:

En la etapa de *recolección* de datos del método científico, se define a la población de interés y se selecciona una muestra o conjunto de personas representativas de la misma, se realizan experimentos o se emplean instrumentos ya existentes o de nueva creación, para medir los atributos de interés necesarios para responder a las preguntas de investigación. Durante lo que es llamado *trabajo de campo* se obtienen los datos en crudo, es decir las respuestas directas de los sujetos uno por uno, se codifican (se les asignan valores a las respuestas), se capturan y se verifican para ser utilizados en las siguientes etapas.

En la etapa de *recuento*, se organizan y ordenan los datos obtenidos de la muestra. Esta será *descrita* en la siguiente etapa utilizando la *estadística descriptiva*, todas las investigaciones utilizan estadística descriptiva, para conocer de manera organizada y resumida las características de la muestra.

En la etapa de *análisis* se utilizan las pruebas estadísticas (*estadística inferencial*) y en la *interpretación* se acepta o rechaza la hipótesis nula.

En investigación, el fenómeno en estudio puede ser cualitativo que implicaría comprenderlo y explicarlo, o cuantitativo para compararlo y hacer inferencias. Se puede decir que si se hace análisis se usan métodos cuantitativos y si se hace descripción se usan métodos cualitativos.

Medición

Para poder emplear el método estadístico en un estudio es necesario *medir* las variables.

Medir: es asignar valores a las propiedades de los objetos bajo ciertas reglas, esas reglas son los niveles de medición

Cuantificar: es asignar valores a algo tomando un patrón de referencia. Por ejemplo, cuantificar es ver cuántos hombres y cuántas mujeres hay.

Variable: es una característica o propiedad que asume diferentes valores dentro de una población de interés y cuya variación es susceptible de medirse.

Las variables pueden clasificarse de acuerdo al tipo de valores que puede tomar como:

- Discretas o categóricas.- en las que los valores se relacionan a nombres, etiquetas o categorías, no existe un significado numérico directo
- Continuas.- los valores tienen un correlato numérico directo, son continuos y susceptibles de fraccionarse y de poder utilizarse en operaciones aritméticas

De acuerdo a la cantidad de valores

- Dicotómica.- sólo tienen dos valores posibles, la característica está ausente o presente
- Policotómica.- pueden tomar tres valores o más, pueden tomarse matices diferentes, en grados, jerarquías o magnitudes continuas.

En cuanto a una clasificación estadística

- Aleatoria.- Aquella en la cual desconocemos el valor porque fluctúa de acuerdo a un evento debido al azar
- Determinística.- Aquella variable de la que se conoce el valor
- Independiente.- aquellas variables que son manipuladas por el investigador. Define los grupos
- Dependiente.- son mediciones que ocurren durante el experimento o tratamiento (resultado de la independiente), es la que se mide y compara entre los grupos

Niveles de Medición

Nominal

Las propiedades de la medición nominal son:

- 1) Exhaustiva: implica a todas las opciones
- 2) A los sujetos se les asignan categorías, por lo que son mutuamente excluyentes. Es decir, la variable está presente o no; tiene o no una característica

Ordinal

Las propiedades de la medición ordinal son:

- 1) El nivel ordinal posee **transitividad**, por lo que se tiene la capacidad de identificar que “esto es mejor o mayor que aquello”, en ese sentido se pueden establecer jerarquías
- 2) Las distancias entre un valor y otro no son iguales

Intervalar

- 1) El nivel de medición intervalar requiere distancias iguales entre cada valor. Por lo general utiliza datos cuantitativos. Por ejemplo: temperatura, atributos psicológicos (CI, nivel de autoestima, pruebas de conocimientos, etc.)
- 2) Las unidades de calificación son equivalentes en todos los puntos de la escala. Una escala de intervalos implica: clasificación, magnitud y unidades de tamaños iguales (Brown, 2000).
- 3) Se pueden hacer operaciones aritméticas
- 4) Cuando se le pide al sujeto que califique una situación del 0 al 10 puede tomarse como un nivel de medición intervalar, siempre y cuando se incluya el 0

Razón

- 1) La escala empieza a partir del 0 absoluto, por lo tanto incluye sólo los números por su valor en sí, por lo que no pueden existir los números con signo negativo. Por ejemplo: Peso corporal en kg., edad en años, estatura en cm.
- 2) Convencionalmente los datos que son de nivel absoluto o de razón son manejados como los datos intervalares.

Nivel	Propiedades	Medidas de tendencia central y dispersión	Compara	Análisis	Ejemplo
Nominal	= ó ≠	moda, porcentaje	Proporciones	<i>Inferencial no paramétrica:</i> X ² , McNemar, Coeficiente Phi	Presencia de estrés (si o no), Religión
Ordinal	≠, =, >, <	moda, porcentajes, mediana	Medianas	<i>Inferencial no paramétrica:</i> U de Mann Whitney, Kolmogorov Smirnov, Friedman, Wilcoxon, prueba de los signos, correlación Spearman	Calidad (Buena, mediana, mala), Nivel socioeconómico (alto, medio, bajo)
Intervalar	Distancias iguales, =, ≠, >, <, 0 relativo	moda, porcentajes, mediana, media, desviación estándar, varianza, curtosis, sesgo, rango, deciles, cuartiles, percentiles	Medias	<i>Inferencial paramétrica:</i> t de Student, ANOVA, regresión, correlación Pearson	Temperatura, atributos psicológicos
Razón o proporción	Igual a intervalar, 0 absoluto	igual a intercalar	Igual a intervalar	Igual a intervalar	Altura, Peso, Edad en años, Grados Kelvin

Las variables intervalares y de razón se pueden incluir en operaciones aritméticas y para efectos de análisis estadísticos se consideran con las mismas propiedades y se clasifican en un rubro general llamado *nivel escalar*.

Las variables ordinales, por otro lado, aparentemente no se pueden sumar, pero cuando se tiene una escala con reactivos calificados en un nivel ordinal, al ser sumados el resultado de esta suma puede ser considerado como intervalar, ya que matemáticamente se comporta como tal.

Al hacer análisis estadísticos es necesario determinar las unidades de medición y las unidades de análisis:

La *unidad de medición* es el atributo que se va a medir o *variable dependiente*.

La *unidad de análisis* son los sujetos, es decir a quién mido o a qué mido, está definida por la *variable independiente*. Para *muestras independientes*, es decir cuando la unidad de medición es registrada una sola vez en la unidad de análisis, se refiere a los grupos a comparar (ej. Hombres vs. Mujeres, Grupo control vs. Grupo experimental, Fumadores vs. No fumadores, etc.), en el caso de *muestras relacionadas*, en que la unidad de medición es medida más de una ocasión en la unidad de análisis, son las diferentes mediciones hechas (ej. Evaluaciones Pre-test vs. Post-test vs. seguimiento, hermanos gemelos, grupos apareados durante la selección de sujetos).

La forma más directa de tener muestras relacionadas o apareadas es hacer mediciones antes - después con los mismos sujetos, sin embargo, el aprendizaje puede ser una variable que pueda afectar la ejecución de una persona, en esos casos se usan grupos equivalentes, a uno se le hace la primera medición y al otro se le aplica el tratamiento y la segunda medición, en este caso siguen manteniendo las características de muestras relacionadas.

La idea de apareamiento o muestras relacionadas es que un sujeto tenga un equivalente en el otro grupo, las características que se igualan depende de las variables que se desean controlar en una investigación, por ejemplo: mismo sexo, misma edad, mismo grado escolar, mismo barrio, condiciones similares de una enfermedad, etc., es imposible tener sujetos iguales, pero entre más características de interés para el estudio compartan, serán considerados equivalentes.

Otra forma de tener muestras relacionadas de manera natural es trabajar con gemelos idénticos (igual ambiente familiar, misma carga genética, mismo ambiente cultural, etc.), aunque si por requerimientos de una investigación no comparten una característica, importante para el estudio, dejarían de considerarse como relacionados (por ejemplo, si se desea probar un tratamiento para disminuir la hiperactividad y se tiene a gemelos que fueron adoptados por diferentes familias y se sabe que los estilos educativos de los padres influyen sobre la hiperactividad, el hecho de provenir de familias adoptivas diferentes cancela la equivalencia).

La metodología estadística puede considerarse según sus dos planteamientos como *descriptiva y deductiva o inferencial*.

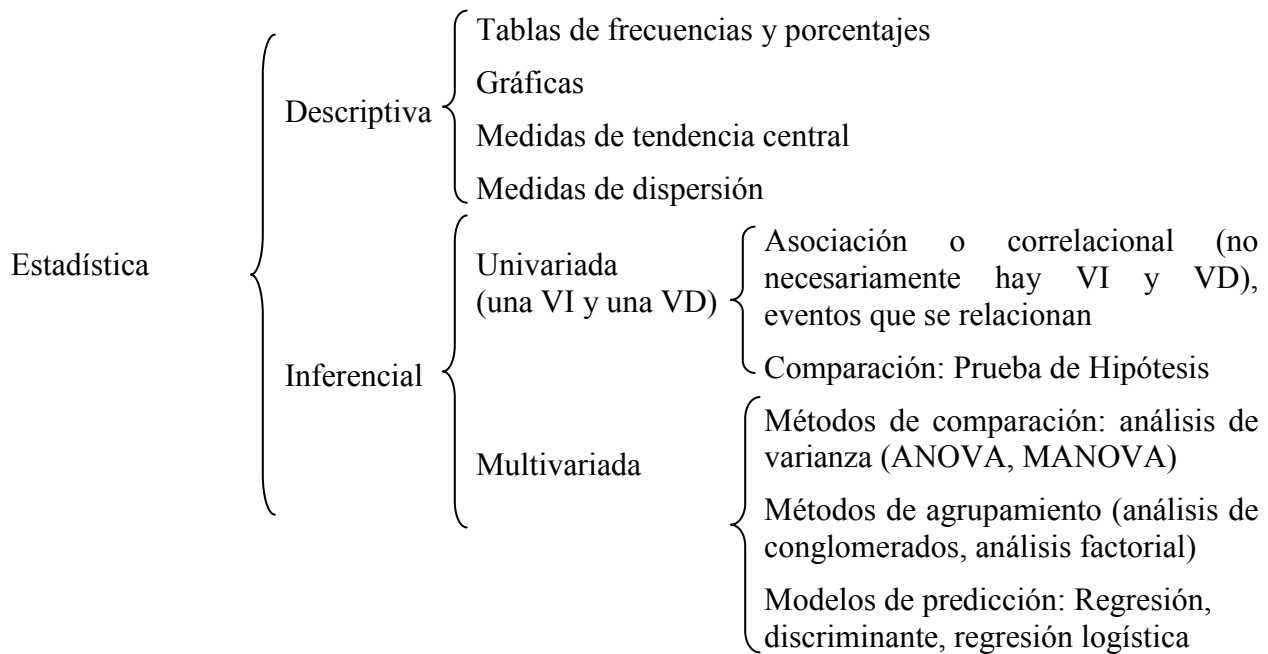
Estadística Descriptiva

Permite la organización de datos desestructurados de tal manera que sean más fáciles de interpretar y de conocer las características de una muestra de forma rápida y resumida.

Incluye:

- **Tablas de frecuencias y porcentajes** con las que se registran cuantos sujetos en la muestra respondieron a las diferentes opciones o categorías de las variables del estudio
- **Métodos de resumen o numéricos** que se dividen en:
 - *Medidas de tendencia central*:
 - ◆ Media (\bar{X}). Promedio aritmético de una distribución o conjunto de valores
 - ◆ Mediana (Md). Valor que divide a la distribución por la mitad
 - ◆ Moda (Mo). Categoría o puntuación que ocurre con mayor frecuencia

- *Medidas de dispersión:*
 - ◆ Rango. Es la diferencia entre la puntuación mayor y la menor en una distribución y se obtiene restandole a la puntuación mayor, la menor
 - ◆ Desviación estándar (s). Es el promedio de desviaciones de las puntuaciones con respecto a la media. Se emplea para variables medidas por intervalos o de razón. Es un indicador de la dispersión de las puntuaciones respecto de la media
 - ◆ Varianza (s^2). Es la desviación estándar elevada al cuadrado. La varianza se divide en: *sistemática*, es decir las variaciones debidas a los efectos de la variable independiente o tratamiento, y *de error o no explicada o residual*, que es debida al azar y a los problemas de control ($s^2 = s_s^2 + s_E^2$)
- **Gráficos** que son una representación visual de los datos (medias, porcentajes, distribuciones de frecuencias, etc.) se pueden hacer gráficas de barras, de sectores o pastel, de líneas, etc.



Curvas de Distribución de Frecuencias

Las curvas de distribución de frecuencias derivan de histogramas o polígonos de frecuencias que se sustituyen por una curva suavizada, el área bajo la curva representa al total de los casos en la población y es igual a una proporción de 1 o un porcentaje igual al 100% (Ritchey, 2002).

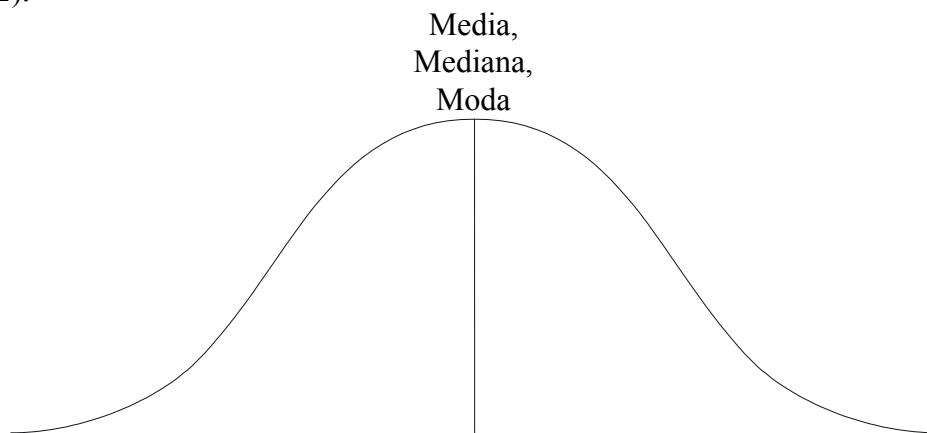
Distribución Normal o Curva Normal (Campana de Gauss)

La curva normal es un modelo teórico o ideal sobre cómo debe comportarse la distribución de las variables en una muestra, se obtuvo de una ecuación matemática.

Representa una curva de distribución de frecuencias en la que la mediana, la moda y la media de una variable son iguales entre sí, tiene forma de campana (Ritchey, 2002). Su utilidad puede ayudarnos a entender situaciones reales.

Muchas variables relacionadas a atributos psicológicos, como la inteligencia medida en términos de Coeficiente Intelectual (CI), se distribuyen en la población tomando la forma de la curva normal, en que la mayoría de los casos se encuentran en el centro (Moda), existiendo pocos casos en los

extremos, siguiendo el ejemplo del CI, existen pocos sujetos con inteligencia muy baja por un lado y pocos genios por el otro, esto es lo que da a la curva normal su característica forma de campana (Ritchey, 2002).



Calificaciones z o estandarizadas

La *desviación estándar* proporciona una unidad de medida común (estándar) que permite comparar variables con medidas observadas diferentes (Ritchey, 2002). Supóngase que se tienen dos escalas para medir autoestima, la primera se califica con un puntaje entre 0 y 20, y la segunda tiene calificaciones entre 0 y 50, si se compararan las calificaciones obtenidas por una y otra, los puntajes crudos muy probablemente serían mayores en la segunda escala. Sin embargo, si se considera la media de cada escala y las desviaciones estándar, es posible conocer en términos de desviaciones estándar que puntuación se encuentra por arriba de la media o por debajo de la misma, al comparar a dos personas.

Por ejemplo, el grupo calificado con la primera escala tiene una media de 15 y una desviación estándar de 4, para el segundo la media es de 35 con una desviación estándar igual a 6. Supóngase que una persona tiene una calificación de 19 en el primer caso y otra con 29 en el segundo caso. La primera se encuentra por arriba de la media una desviación estándar y la segunda una desviación debajo de la media, es decir, la primera persona tiene un nivel de autoestima mayor que el de la segunda persona.

Las calificaciones brutas se transforman con frecuencia a otro tipo de puntuaciones, para facilitar el análisis y la interpretación. Estas calificaciones son derivadas o transformadas. Un tipo de calificación derivada, la calificación estándar z resulta particularmente útil.

Calificación o Puntuación z: Es una puntuación estandarizada. Su sentido es poder hacer comparaciones dentro o entre sujetos, cuando han sido medidos con diferentes escalas.

La distribución de puntuaciones z tiene una media de cero y una desviación estándar de 1.

La \bar{X} de la z es 0

La s de la z es 1

Ventajas:

- Miden una escala de intervalos, en términos de unidades de desviación estándar.
- Permite comparar calificaciones de varias pruebas en forma directa, incluso cuando se tienen medias y desviaciones estándar diferentes

La calificación z se calcula como sigue:

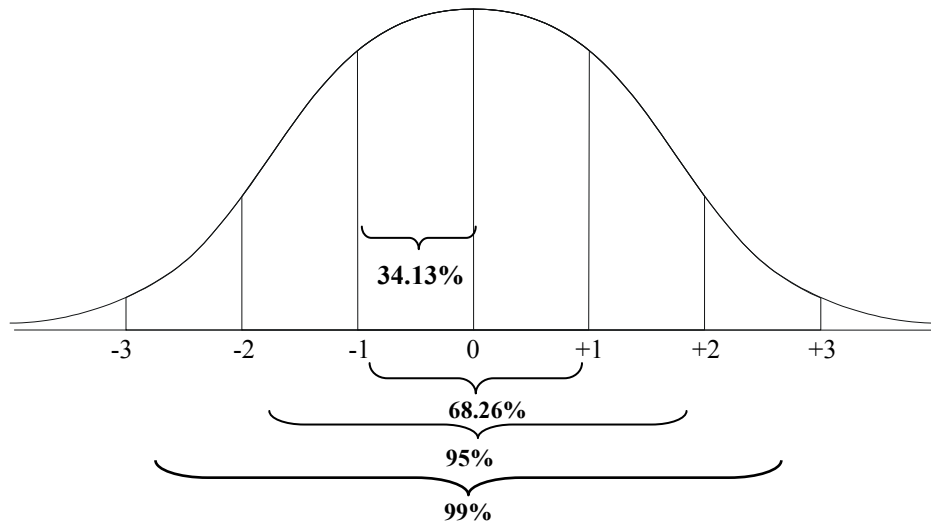
$$z = \frac{X - \bar{X}}{s}$$

Es decir, qué tanto se aleja la puntuación del sujeto de la media, y luego se compara con la desviación estándar, en otras palabras, se estandariza.

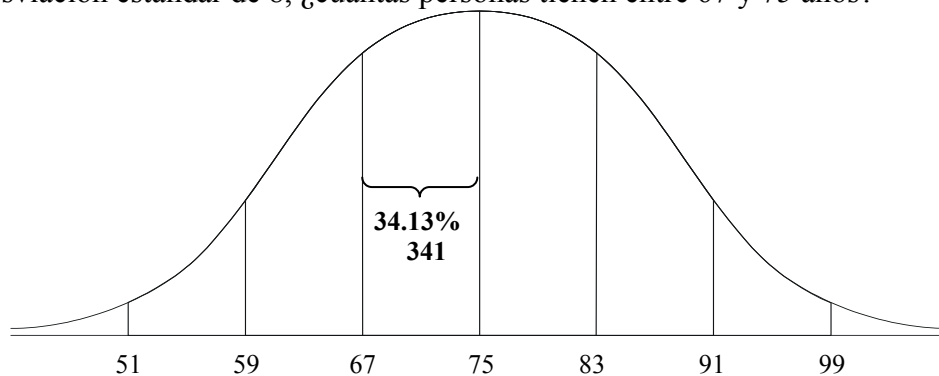
Al analizar la curva normal en términos de puntuaciones z , y recordando que el área bajo la curva representa el 100% de los casos, una desviación estándar tomada a cada lado de la media incluye un área de 68.26% de la total. Toda el área incluida por 2 unidades de desviación estándar a ambos lados de la media comprenden el 95.44% de los casos, si se continúa a la tercer desviación estándar a cada lado de la media, se toma otra parte igual a 2.15% del área total. La suma de todas las áreas abarcadas por estas 6 unidades de desviación estándar es igual a 99.74% del total. El 95% del área bajo la curva, está entre 1.96 y -1.96 desviaciones estándar, el 99% entre -2.58 y 2.58 y el 99.9% entre -3.90 y 3.90 , estos tres grupos de desviaciones son particularmente importantes en la estadística inferencial como se verá más adelante.

Para interpretar el área bajo la curva.

- El área bajo la curva representa el 100%. La mitad representa el 50% de cada lado
- El área bajo la curva es igual a 1 (convertida en términos de proporción)
- Para conocer el área bajo la curva se necesitan las puntuaciones z



Utilizando esta información se puede calcular el área comprendida entre dos puntos con base en las puntuaciones crudas. Por ejemplo: Una población de 1000 sujetos, tiene una media de edad de 75 años y una desviación estándar de 8, ¿cuántas personas tienen entre 67 y 75 años?



Sustituyendo la media (75) en el centro y las desviaciones estándar a la izquierda o negativas (-8) y a la derecha o positivas (+8), la respuesta es el 34.13% del total de la población es decir 341 sujetos tienen entre 67 y 75 años.

Estadística Inferencial

Su propósito principal es estimar los atributos de la población a partir de una muestra de casos. Se pueden probar relaciones entre variables, comparar grupos con respecto a cierta característica y hacer inferencias.

Al clasificar a la estadística inferencial de acuerdo al número de variables dependientes se tiene:

Estadística Inferencial Univariada

Según la definición clásica, requiere de:

- 2 tipos de variables: Puede haber una o varias variables independientes, y solo una variable dependiente.

Según la definición práctica hay una variable dependiente y una independiente. Para la elección de las pruebas se tomará el aspecto práctico.

En la estadística inferencial univariada existen dos **tipos de problema**:

COMPARACIÓN		ASOCIACIÓN
<i>Muestras independientes:</i> Cuando la variable dependiente de interés es medida sólo una vez en la unidad de análisis	<i>Muestras relacionadas:</i> Cuando la variable dependiente es medida más de una ocasión en la unidad de análisis	Aquí se puede o no tener una variable dependiente, pero eso no es relevante pues sólo se busca saber la relación existente entre dos variables

- a) ***Asociación.***- En este tipo de problemas sólo hay dos eventos que se relacionan o asocian (correlación)
- b) ***Comparación.***- También llamada prueba de hipótesis. Se comparan grupos o mediciones e intervienen dos tipos de variables
Variable dependiente (V. D.), es el fenómeno o proceso que se va a medir, sus variaciones dependen de la variable independiente
Variable independiente (V. I.), es la variable que identifica los tratamientos empleados para influir sobre la variable dependiente, esta variable define los grupos que van a ser comparados

¿Cómo saber si el problema es de asociación o comparación?

1) **¿Se tienen grupos?**

Si la respuesta es: NO, entonces es poco probable que sea de comparación, pero no imposible.
Si la respuesta es: SI, se está comparando.

2) **¿Hay Tratamientos (Tx)? (¿Se hace un experimento o intervención?)**

Si la respuesta es SI: se está comparando
Si la respuesta es NO: casi seguro es de asociación

Puntos clave para hacer la identificación del tipo de problema:1.- Objetivo del estudio:

Aclarar semántica y conceptualmente cual es el objetivo del estudio.

2.- Identificar diseño:

Existen 3 diseños básicos:

- Una muestra para relacionar: Asociación
- Grupos: Muestras independientes
- Seguimiento: Muestras relacionadas

3.- Identificar la variable dependiente y la independiente4.- Identificar Unidad de Análisis (definida por la variable independiente)

Hay que ver cuantas veces se midió a la variable dependiente.

Es importante distinguir entre asociación conceptual y estadística, puede estarse asociando conceptualmente (puede plantearse como un problema de asociación), pero estar comparando estadísticamente. La diferencia está en identificar si hay grupos o no.

Ejemplo:

Problema: ¿Es igual la proporción de fumadores entre estudiantes de secundaria hombres y mujeres?

Población: estudiantes de secundaria

Variable dependiente: fumar

Variable independiente: sexo

Atributos de la variable independiente: hombres y mujeres

La V. D. se mide sólo una vez en cada grupo (definidos por el sexo), por lo que se trata de un problema de comparación de dos muestras o grupos independientes (hombres y mujeres).

Hay que tener cuidado de no confundir la variable dependiente, es decir qué atributo se mide en la unidad de análisis.

Estadística Multivariada

Según la definición clásica: 2 o más variables dependientes, no importa el número de variables independientes.

Según la definición práctica: 2 o más variables independientes y 2 o más variables dependientes. Para la elección de las pruebas se tomará el aspecto práctico.

Tiene dos grandes grupos

a) Métodos de agrupamiento.- Se realiza mediante:

- análisis factorial exploratorio
- análisis de cúmulos (clusters)

b) Métodos de clasificación.- Se realiza mediante:

- análisis de regresión
- análisis discriminante
- análisis de regresión logística

La estadística también se puede clasificar como: paramétrica y no paramétrica

Estadística Paramétrica: Está basada en dos supuestos: *estimadores* que son medidas referentes a la muestra como la media (\bar{X}) o la varianza s^2 y *parámetros* que son los equivalentes poblacionales de los estimadores, como la media poblacional (μ) y la varianza poblacional (σ^2).

Estadística Inferencial {
No paramétrica
Paramétrica

La estadística paramétrica necesita cumplir con cuatro requisitos para poderse aplicar:

1. La variable dependiente debe distribuirse normalmente (campana de Gauss) o muy similar
2. Homocedasticidad u homogeneidad de varianzas o varianzas iguales: que cuando se comparan grupos estos tengan la misma dispersión con respecto a la media de la variable dependiente
3. Asignación y Selección aleatoria de los grupos (muestreo completamente al azar)
4. Que la variable dependiente esté medida a nivel intervalar o de razón

Estos requisitos deben ser cubiertos para poder generalizar con base en los estimadores y hacer conclusiones de una muestra a la población.

Por otro lado, la *Estadística No Paramétrica:*

1. Está libre de curva, no necesita distribuirse como la curva normal
2. Se basa en frecuencias, porcentajes, modas y rangos
3. Su nivel de medición es ordinal o nominal

Planteamiento de la hipótesis estadística

La *Hipótesis* es la respuesta tentativa para la solución de la pregunta de investigación.

Al realizar inferencias estadísticas, se acostumbra adoptar un modelo de decisión. Este modelo consta de cuatro elementos:

- Hipótesis nula (H_0)
- Hipótesis alterna (H_1)
- Nivel de significancia que ha de utilizarse en la prueba estadística
- Regla de decisión

Hipótesis para Problemas de Comparación

En la prueba de hipótesis se trabaja con dos hipótesis estadísticas que deben enunciarse explícitamente: la hipótesis que debe probarse o hipótesis nula que se establece con el propósito de ser rechazada, y la hipótesis alterna que es la conclusión a la que se espera llegar.

Con un nivel intervalar o de razón de la V. D. se comparan medias, la hipótesis nula plantea que las dos medias son iguales:

$$H_0: X_1 = X_2$$

y la hipótesis alterna plantea que las medias son diferentes:

$$H_1: X_1 \neq X_2$$

Otro planteamiento sería:

$$H_0: X_1 - X_2 = 0$$

Cuando se tienen medianas (nivel ordinal de la V. D.), entonces:

$$H_0: Md_1 = Md_2$$

$$H_1: Md_1 \neq Md_2$$

O con proporciones o porcentajes (nivel nominal de la V. D.):

$$H_0: P_1 = P_2$$

$$H_1: P_1 \neq P_2$$

Hipótesis para Problemas de Asociación

Cuando se tiene un problema de asociación la hipótesis nula niega la correlación:

$$H_0: r = 0$$

La hipótesis alterna afirma que hay correlación:

$$H_1: r \neq 0$$

Construcción de hipótesis de acuerdo al nivel de medición

Las hipótesis estadísticas tienen que incluir la variable dependiente en primer lugar y la independiente en seguida. Para la redacción de dichas hipótesis se toma en cuenta el nivel de medición utilizado de las variables del estudio.

Para los **problemas de comparación**:

Nivel NOMINAL: Aquí se habla de proporciones y / o categorías.

¿Cómo es la relación entre fumar o no y morir por cáncer pulmonar?

H_0 : La proporción de sujetos que mueren por cáncer pulmonar es igual entre fumadores y no fumadores.

H_1 : La proporción de sujetos que mueren por cáncer pulmonar es diferente entre fumadores y no fumadores

Nivel ORDINAL: Aquí se habla de jerarquías y / o niveles.

¿Cómo es el nivel de creatividad entre los niños de comunidades rurales, urbanas e indígenas?

H_0 : El nivel de creatividad es igual entre niños de comunidades rurales, urbanas e indígenas.

H_1 : El nivel de creatividad es diferente entre niños de comunidades rurales, urbanas e indígenas.

Nivel INTERVALAR: Se comparan las medias y también se habla de niveles.

Se desea conocer cómo el nivel de estrés de los sujetos afecta su nivel de irritabilidad hacia sus compañeros de trabajo. Para ello se conformaron tres grupos, bajo estrés, estrés regular y alto estrés, con 7 profesionistas cada uno, a quienes se les evaluó su nivel de irritabilidad preguntándoles en una escala de 0 a 10 que indique: “cotidianamente ¿qué tan irritable se muestra con sus compañeros de trabajo?”

H_0 : El nivel de irritabilidad hacia los compañeros de trabajo es igual entre los tres grupos de estrés.

H_1 : El nivel de irritabilidad hacia los compañeros de trabajo es diferente entre los tres grupos de estrés.

Para el caso de los **problemas de asociación** se identifica la relación entre las variables, y el concepto de asociación o relación se debe incluir en las hipótesis.

¿Qué relación hay entre la edad en años de los sujetos y su inteligencia?

H_0 : No hay asociación lineal entre la edad y la inteligencia.

H_1 : Hay asociación lineal entre la edad y la inteligencia.

NOTA: La hipótesis alterna no se acepta ni se rechaza, es la hipótesis nula la que se somete a prueba.

Los planteamientos anteriores se refieren a pruebas de dos colas, que son problemas en los que no es posible anticipar la dirección de las diferencias, es decir, no se sabe cual grupo es el que tendrá el nivel o la proporción de casos mayor, o cual es el sentido de la relación entre variables.

Sin embargo, pueden existir hipótesis alternas en las que se puede anticipar cual grupo presenta una desviación mayor o menor con respecto al otro. Este tipo se refiere a problemas de una cola o dirección.

En estos casos, con un nivel intervalar o de razón de la V. D., la hipótesis nula plantea que las dos medias son iguales:

$$H_0: X_1 = X_2$$

y la hipótesis alterna plantea que un grupo tiene una media mayor o menor que otro u otros:

$$H_1: X_1 > X_2$$

$$H_1: X_1 < X_2$$

Cuando se tienen medianas (nivel ordinal de la V. D.), entonces:

$$H_0: Md_1 = Md_2$$

$$H_1: Md_1 > Md_2$$

$$\text{ó } H_1: Md_1 < Md_2$$

O con proporciones o porcentajes (nivel nominal de la V. D.):

$$H_0: P_1 = P_2$$

$$H_1: P_1 > P_2$$

$$\text{ó } H_1: P_1 < P_2$$

Para esta guía se presentarán únicamente planteamientos para hipótesis de dos colas.

¿Cuál es el sentido del nivel de significancia o la probabilidad?

La probabilidad (p) de que un evento ocurra oscila entre 0 y 1, donde 0 significa la imposibilidad de ocurrencia y 1 la certeza de que ocurra el fenómeno. Al lanzar al aire una moneda no cargada, la probabilidad de que salga “águila” es de 0.5 y la probabilidad de que la moneda caiga en “sol” también es de 0.5. Con un dado, la probabilidad de obtener cualquiera de sus lados al lanzarlo es de $1/6=0.1667$. La suma de probabilidades siempre es de 1.

Aplicando el concepto de la probabilidad a la distribución muestral, el área de ésta corresponde a la probabilidad total ($p = 1$), y consecuentemente, cualquier área (porcentaje bajo la curva) comprendida

entre dos puntos de la distribución corresponderá a la probabilidad de la distribución al convertirla a proporciones (por ejemplo $25\% = 0.25$).

Para probar hipótesis inferenciales utilizando la media, el investigador debe evaluar si es alta o baja la probabilidad de que la media de la muestra esté cerca de la media de la distribución poblacional. Si es baja, el investigador dudará de generalizar a la población. Si es alta el investigador podrá hacer generalizaciones. Es aquí donde entra el nivel de significancia o nivel α . El nivel α es la probabilidad de equivocarse al probar las hipótesis estadísticas.

El nivel de significancia de 0.05, implica que el investigador tiene un 95% de seguridad para generalizar sin equivocarse y sólo 5% en contra. Este es el más utilizado en investigación en Psicología, aunque se pueden utilizar niveles más bajos (0.01 o 0.001) cuando se requiere un mayor grado de certeza.

Decisión estadística

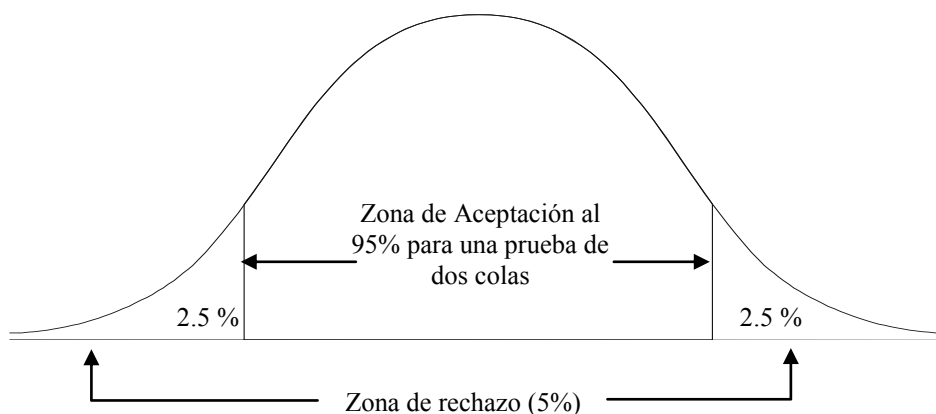
La decisión e interpretación de un análisis estadístico se basa en la aceptación o rechazo de la hipótesis nula, están estrechamente relacionados con la curva normal y las puntuaciones z , una vez elegida la prueba estadística adecuada al nivel de medición y haber redactado la hipótesis nula, se debe establecer un nivel de significancia o de certeza para rechazar esta última sin cometer el error llamado del tipo I o α , que se refiere a rechazar la hipótesis nula siendo verdadera. En psicología normalmente se establece el nivel de significancia de 0.05 que, como se mencionó anteriormente, representa un 95% de certeza de generalizar los resultados sin equivocarse. En este sentido, la decisión estadística es una decisión probabilística. Si se desea mayor certeza, se debe utilizar otro nivel de significancia, como puede ser a 0.01 que proporciona una certeza de 99% para generalizar los resultados sin error, o a 0.001 que equivale al 99.9%.

Existe otro tipo de error el tipo II o β , que al contrario del error I se refiere a aceptar la hipótesis nula siendo falsa, en términos de decisión estadística es más grave cometer el error tipo I , ya que afirmaríamos que hay diferencias entre los grupos cuando esto no es verdad. Es una situación a la que es fácil llegar pues normalmente se espera que la intervención que se haga sea la causa de las diferencias entre grupos.

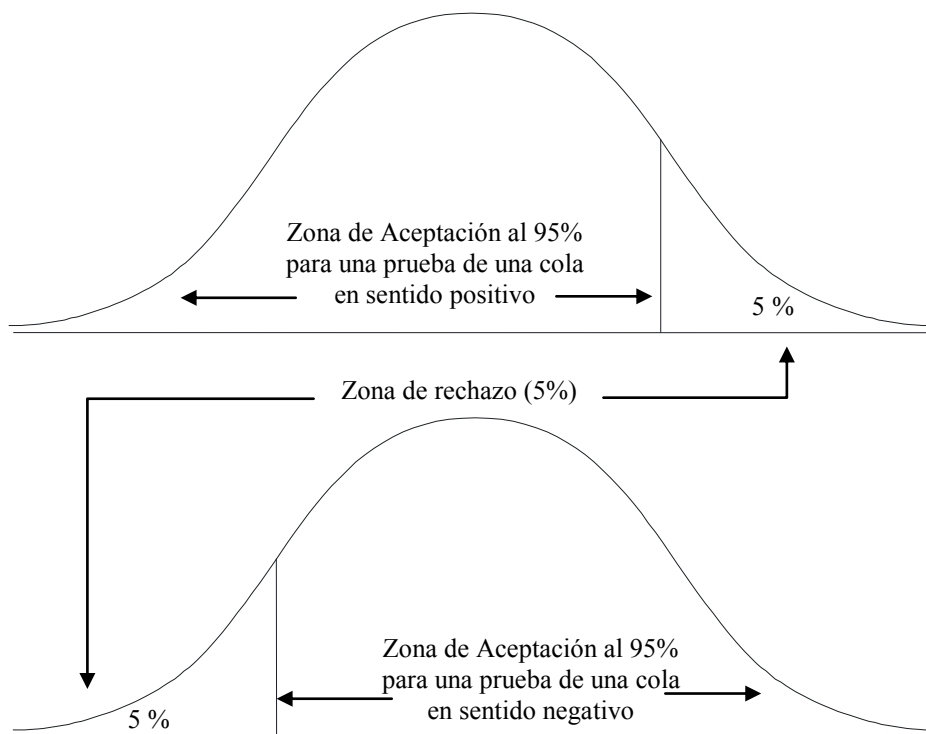
Al representar el nivel de significancia bajo la curva normal se tienen un área de aceptación de la hipótesis nula y una zona de rechazo, para una hipótesis de colas se reparte la significancia entre los dos extremos de la curva y para la de una cola se considera sólo el extremo positivo o negativo dependiendo de cuál es el grupo que se espera que tenga un nivel o proporción mayor.

Al traducir las áreas a probabilidad, se tiene que hay un 0.05 de posibilidad para equivocarse al rechazar la hipótesis nula. Entre menor sea el área de rechazo se tendrá más certeza para generalizar los resultados a la población.

Para una prueba de dos colas, en términos de puntajes z , a 1.96 desviaciones estándar se tiene el 95% del área bajo la curva en la región de aceptación de la hipótesis nula y el 5% o 0.05 de significancia (α) en la zona de rechazo, en 2.58 desviaciones el 99% ($\alpha = 0.01$) y en de 3.90 desviaciones el 99.9% ($\alpha = 0.001$)



Para una prueba de una cola, a 1.64 desviaciones estándar en sentido negativo o positivo, se tiene el 95% del área bajo la curva en la región de aceptación de la hipótesis nula y el 5% o 0.05 de significancia (α) en la zona de rechazo, en ± 2.32 desviaciones el 99% ($\alpha = 0.01$) y en 3.70 desviaciones el 99.9% ($\alpha = 0.001$)



Para conocer si el valor de una prueba estadística permite rechazar la hipótesis nula, se tiene que entre mayor sea este valor se entra más a la zona de rechazo de la hipótesis nula, sin embargo, para la mayoría de las pruebas se debe considerar además en la decisión a los grados de libertad, el número de casos o se compara la probabilidad directamente.

Grados de libertad

Son la libertad de variaciones que puede tener una variable, suponiendo que se tuvieran 4 puntuaciones cuya media es igual a 10 al tener los valores de las tres primeras, la última estará determinada por las primeras, por ejemplo: 7, 12, 15, la última puntuación necesariamente es 6. La cantidad de comparaciones independientes se determina a partir de los grados de libertad, que normalmente se calcula teniendo el tamaño de la muestra menos uno ($gl = n - 1$). Sin embargo los grados de libertad se obtienen de manera diferente para cada prueba, por lo que se debe estar atento a cada uno de los procedimientos.

Reglas de decisión

El valor de las pruebas estadísticas se debe comparar con uno obtenido, con relación al nivel de significancia y los grados de libertad, de una tabla de valores críticos. **La regla de decisión en estos casos es: el valor de la prueba debe ser mayor o igual al de la tabla para rechazar la hipótesis nula.** Esta regla puede cambiar, por lo que es necesario revisar la regla de decisión específica de cada procedimiento.

Los paquetes estadísticos presentan los valores de cada prueba junto con algunos datos necesarios para el cálculo de ésta (medias o porcentajes, el número de casos, los grados de libertad, etc.) y el nivel de significancia o probabilidad, éste representa la posición del valor del estadístico en el área de rechazo, o aceptación, de la hipótesis nula.

Como **regla de decisión**, observando los resultados del paquete estadístico, a un nivel de significancia establecido en 0.05: **Si la probabilidad o nivel de significancia es menor o igual a 0.05 se rechaza la hipótesis nula.**

Estadística Inferencial Univariada: Pruebas Estadísticas Seleccionadas

En este apartado se revisarán algunas pruebas para el análisis inferencial univariado y la presentación de los resultados obtenidos, se tienen por ejemplo: pruebas paramétricas de comparación de dos muestras independientes (t de Student para Muestras Independientes) y de dos muestras relacionadas (t de Student para Muestras Relacionadas), así como pruebas no paramétricas de comparación de dos muestras independientes como la Chi Cuadrada (X^2) y la U de Mann-Whitney y pruebas no paramétricas de comparación de dos muestras relacionadas como la prueba de McNemar y la de Wilcoxon.

Para comparaciones de más de dos grupos se estudiarán, para grupos independientes: al Análisis de Varianza de una Vía entre las pruebas paramétricas; X^2 y Kruskal Wallis de las no paramétricas. En el caso de comparación de tres o más grupos relacionados se analizará el uso de las pruebas no paramétricas Q de Cochran y Friedman.

Finalmente serán revisadas las pruebas de asociación de la familia producto momento: el coeficiente de correlación de Pearson, el Coeficiente de correlación de Spearman y el coeficiente de correlación Phi.

Además se incluirán ejemplos de la forma adecuada para reportar resultados para cada prueba.

Cómo Presentar Resultados de Análisis Estadísticos

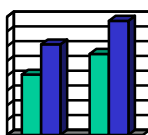
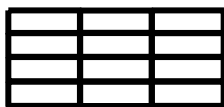
Los gráficos se pueden utilizar para una mejor explicación de los datos que hemos encontrado. Una gráfica de barras, por ejemplo, se usa para ver cual es el grupo que tiene un promedio mayor y cual el menor, o si no hay diferencias entre ellos. También se pueden comparar porcentajes con este tipo de gráficas.

Al hacer un análisis estadístico se tienen, además de los valores resultantes de cada prueba, elementos que se utilizan para hacer una gráfica, por ejemplo: medias en las pruebas de hipótesis de nivel intervalar, o porcentajes en las hipótesis de nivel nominal. Las gráficas o tablas que se realicen para reportar resultados de pruebas de hipótesis deben estar acompañadas, además, de los valores de la prueba estadística que se utilizó y el nivel de significancia con que se rechaza, o acepta, la hipótesis nula.

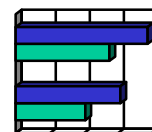
En la descripción de la gráfica se resaltarán las diferencias entre los grupos, o ausencia de éstas, y se aclarará si son estadísticamente significativas de acuerdo con la prueba utilizada (la conclusión de si la hipótesis nula fue rechazada o aceptada).

A continuación se muestran, a manera de ejemplo, diferentes tipos de presentación de resultados:

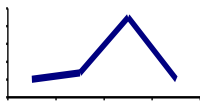
Tablas: Características de la muestra (frecuencias y porcentajes)



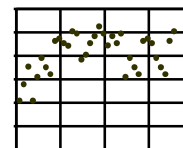
Barras o Columnas: comparar grupos (porcentajes, medianas, medias)



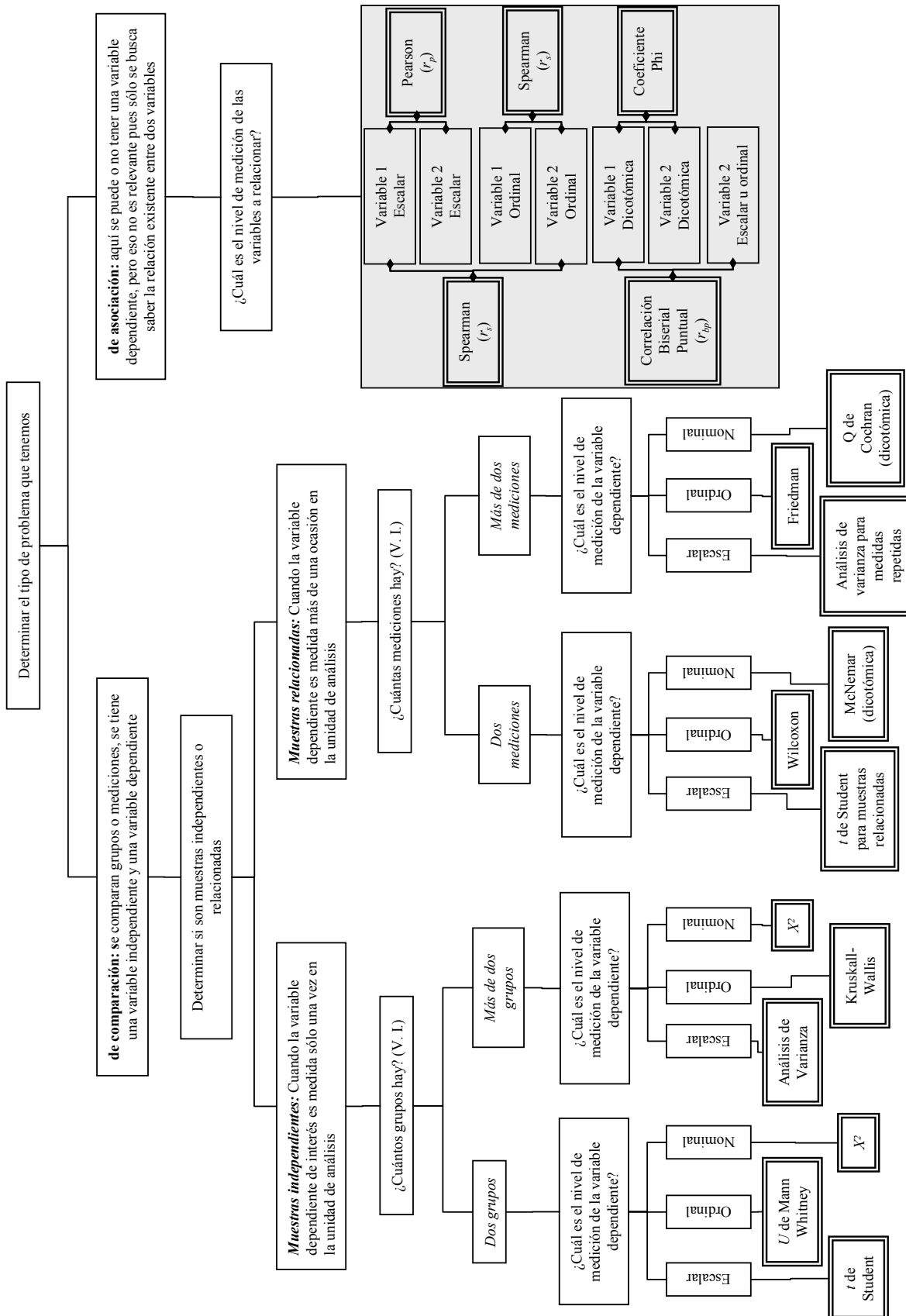
Una gráfica de líneas permite observar las variaciones entre mediciones



Para la distribución de variables en una muestra usamos una gráfica XY o de dispersión, este tipo de gráficas se usa para representar correlaciones



Criterios importantes para seleccionar el tipo de prueba en un estudio*



* Esta guía fue elaborada con base en datos proporcionados por el Mtro. Jorge Ameth Villatoro Velázquez, Instituto Nacional de Psiquiatría

Pruebas de Comparación

Prueba t de Student

La prueba t de Student es una prueba paramétrica de comparación de dos muestras, es decir necesita cumplir las siguientes características:

- Selección completamente aleatoria de los grupos
- Homocedasticidad (homogeneidad de las varianzas de la variable dependiente en ambos grupos)
- Distribución normal de la variable dependiente en los dos grupos
- Nivel intervalar de la variable dependiente

Su función es comparar dos grupos de puntuaciones (medias aritméticas) y determinar que la diferencia no se deba al azar (que la diferencia sea estadísticamente significativa).

Esta prueba tiene dos modalidades, una para muestras independientes y otra para grupos relacionados.

La **prueba t para muestras independientes** se calcula mediante la siguiente fórmula:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{D\bar{x}}}$$

dónde:

\bar{x}_1, \bar{x}_2 = medias de los grupos

$s_{D\bar{x}}$ = error estándar de la diferencia entre medias

Para calcular el error estándar de la diferencia entre medias:

$$s_{D\bar{x}} = \sqrt{\frac{\sum x_1^2 + \sum x_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

Para calcular la suma de cuadrados de X :

$$\sum x^2 = \sum X^2 - \frac{(\sum X)^2}{n}$$

Para calcular los grados de libertad:

$$gl = n_1 + n_2 - 2$$

Problema para explicar el procedimiento para calcular la prueba t para muestras independientes:

Se desea probar dos métodos de enseñanza del idioma inglés (método tradicional y método multimedia), para lo cual al final de cada uno de los cursos se aplicó una prueba de conocimientos. Los puntajes obtenidos por los sujetos asistentes a cada uno de los cursos se muestran a continuación, donde X_1 son las calificaciones del primer grupo (método tradicional) y X_2 las del segundo (método multimedia):

X_1	X_2	X_1^2	X_2^2	
26	38	676	1444	
24	26	576	676	
18	24	324	576	
17	24	289	576	
18	30	324	900	
20	22	400	484	
18		324		
$\sum X_1 = 141$		$\sum X_2 = 164$	$\sum X_1^2 = 2913$	$\sum X_2^2 = 4656$
$n_1 = 7$		$n_2 = 6$		
$\bar{x}_1 = 20.14$		$\bar{x}_2 = 27.33$		

Procedimiento:

1. Para cada problema específico se debe determinar el tipo de problema, los grupos a comparar, si son independientes o relacionados, el nivel de medición, la prueba estadística que debe utilizarse y plantear la hipótesis nula que se va a probar de acuerdo al nivel de medición y tipo de problema correspondiente.

Es un problema de Comparación

2 grupos independientes

Nivel de medición intervalar

H_0 : El nivel de conocimientos del idioma inglés es igual entre el grupo del curso tradicional (grupo 1) y el grupo del curso multimedia (grupo 2)

Prueba estadística: Prueba t para grupos independientes

2. Se deben elevar al cuadrado los puntajes obtenidos por el grupo 1 (X_1) y los del grupo 2 (X_2) y escribirlos en la columna X_1^2 para los puntajes elevados al cuadrado del grupo 1 y X_2^2 para los puntajes elevados al cuadrado del grupo 2.
3. Sumar los puntajes obtenidos por el grupo 1 ($\sum X_1$) y por el grupo 2 ($\sum X_2$), así como los puntajes elevados al cuadrado de cada grupo ($\sum X_1^2$ y $\sum X_2^2$).
4. Calcular la suma de cuadrados de X para cada grupo ($\sum x_1^2$ y $\sum x_2^2$)

$$\sum x_1^2 = 2913 - \frac{141^2}{7} = 2913 - \frac{19881}{7} = 2913 - 2840.14 = 72.86$$

$$\sum x_2^2 = 4656 - \frac{164^2}{6} = 4656 - \frac{26896}{6} = 4656 - 4482.67 = 173.33$$

$$\sum x_1^2 = 73 \qquad \sum x_2^2 = 173$$

5. Obtener el error estándar de la diferencia entre medias (s_{Dx})

$$s_{Dx} = \sqrt{\frac{73+173}{7+6-2} \left(\frac{1}{7} + \frac{1}{6} \right)} = \sqrt{\frac{246}{11} (0.143 + 0.167)} = \sqrt{(22.3636)(0.3100)} = \sqrt{6.9327} = 2.63$$

6. Determinar las medias para los puntajes de ambos grupos (\bar{x}_1 y \bar{x}_2)

7. Obtener el valor de t y los grados de libertad

$$t = \frac{20.14 - 27.33}{2.63} = \frac{-7.19}{2.63} = -2.73$$

$$gl = 7 + 6 - 2 = 11$$

8. Buscar el valor de t en la tabla de valores críticos de t , al nivel de significancia preestablecido y con los grados de libertad calculados. El valor de t obtenida debe compararse con el valor de tabla. **La regla de decisión es: si el valor calculado es mayor o igual que el valor en tabla se debe rechazar la H_0 .**

Para estos datos con 11 grados de libertad, al nivel de significancia establecido en el valor de 0.05, se tiene un valor t en tabla de 2.201, el valor de t calculado es mayor al de tabla (el signo no debe ser tomado en cuenta) por lo tanto se rechaza H_0 , entonces: el nivel de conocimientos del idioma inglés es diferente entre el grupo del curso tradicional y el grupo del curso multimedia.

Tabla de Valores Críticos para la Prueba t

gl	α		
	0.05	0.01	0.001
1	12.706	63.657	636.619
2	4.303	9.925	31.598
3	3.182	5.841	12.941
4	2.776	4.604	8.610
5	2.571	4.032	6.859
6	2.447	3.707	5.959
7	2.365	3.499	5.405
8	2.306	3.355	5.041
9	2.262	3.250	4.781
10	2.228	3.169	4.587
11	2.201	3.106	4.437
12	2.179	3.055	4.318
13	2.160	3.012	
14	2.145	2.977	
15	2.131	2.947	
16	2.120	2.921	
17	2.110	2.898	
18	2.101	2.878	
19	2.093		

Por otro lado, la prueba t para muestras independientes en el paquete estadístico SPSS se encuentra en el menú **Analizar / Comparar medias / Prueba T para muestras independientes**, los resultados para el ejemplo anterior se presentarían de la siguiente manera:

Prueba T

En la primera sección se describen los grupos a comparar:

Estadísticos de grupo

	Método de enseñanza	N	Media	Desviación típ.	Error típ. de la media
Calificación de la prueba de conocimientos	Tradicional	7	20.14	3.48	1.32
	Multimedia	6	27.33	5.89	2.40

En la siguiente sección se tiene a la prueba estadística propiamente dicha:

Inicialmente se hace una prueba de homogeneidad de varianzas entre los dos grupos a comparar, que es uno de los requisitos de las pruebas paramétricas como la prueba t . Se debe decidir si hay o no varianzas iguales probando la siguiente $H_0: s_1^2 = s_2^2$. Esto se hace mediante la prueba F de Levene para homogeneidad de varianzas, cuyos valores se muestran en la tabla siguiente encerrados en un óvalo. **La regla de decisión es: si el nivel de significancia calculado es menor o igual (\leq) a 0.05 se debe rechazar la H_0 .** En el ejemplo se tiene un valor de F de 1.342 con un nivel de significancia de 0.271 que al ser mayor que 0.05, no permite rechazar la H_0 y se asume que hay varianzas iguales.

Una vez superado este punto, se tienen dos renglones de valores, el superior en el que se asumen varianzas iguales (ver datos encerrados en el rectángulo) y el inferior en que no se asumen varianzas iguales (ver datos encerrados en el hexágono). Siguiendo el ejemplo: se probó que hay varianzas

iguales, se debe seguir la primer línea donde se observa un valor de t de -2.732 con 11 grados de libertad y un nivel de significancia de 0.02 (ver datos encerrados en el rectángulo), **siguiendo la regla de decisión: si el nivel de significancia calculado es menor o igual (\leq) a 0.05 se debe rechazar la H_0** , este valor es menor a 0.05 por lo que se debe rechazar la H_0 , entonces se tiene que afirmar que: el nivel de conocimientos del idioma inglés es diferente entre el grupo del curso tradicional y el grupo del curso multimedia.

Prueba de muestras independientes										
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% Intervalo de confianza para la diferencia	
Calificación de la prueba de conocimientos	Se han asumido varianzas iguales	1.342	.271	-2.732	11	.020	-7.19	2.63	-12.98	-1.40
	No se han asumido varianzas iguales			-2.623	7.862	.031	-7.19	2.74	-13.53	-.85

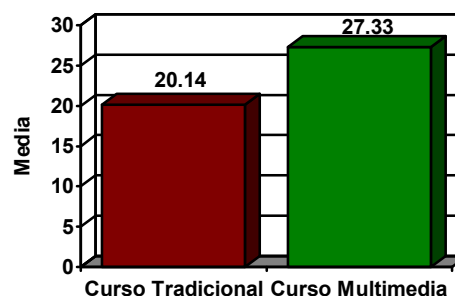
Reporte de resultados de la prueba t

Para esta prueba se puede utilizar una gráfica de barras con las medias, es importante mencionar los valores de la prueba estadística, grados de libertad y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe escribir la conclusión a la que se llega. Por ejemplo:

En lo que respecta a los métodos de enseñanza, se encontraron diferencias en el nivel de conocimientos del idioma inglés entre el grupo del curso tradicional y el grupo del curso multimedia ($t = -2.732$, $gl = 11$, $p < 0.05$). Como puede observarse en la gráfica, el grupo que llevó el curso multimedia tuvo un promedio más alto en la prueba (27.33), al compararlo con el grupo asistente al método tradicional (20.14).

Esta descripción será la base para hacer la discusión posterior de los resultados encontrados en el estudio.

Promedio de Calificaciones en la Prueba de Conocimientos del Idioma Inglés por Método de Estudio*



* $t = -2.732$, $gl = 11$, $p = 0.02$

La **prueba t para muestras relacionadas** se calcula mediante la siguiente fórmula:

$$t = \frac{\bar{D}}{s_{\bar{D}}}$$

dónde:

\bar{D} = diferencia media

$s_{\bar{D}}$ = error estándar de la diferencia media

$$s_{\bar{D}} = \frac{s_D}{\sqrt{n-1}}$$

Para calcular la desviación estándar de las diferencias:

$$s_D = \sqrt{\frac{\sum d^2}{n}}$$

Para calcular la suma de cuadrados de D :

$$\sum d^2 = \sum D^2 - \frac{(\sum D)^2}{n}$$

Para calcular los grados de libertad:

$$gl = n - 1$$

Problema para explicar el procedimiento para calcular la prueba t para muestras relacionadas:

Se desea saber si un taller es efectivo para cambiar el nivel de autoestima de un grupo de pacientes, para ello se evalúa la autoestima antes de iniciar el taller y al finalizar el mismo. Los datos obtenidos se muestran a continuación, donde X son los puntajes de la primera medición y Y corresponde a la segunda medición, D son las diferencias entre los puntajes de las 2 mediciones:

X	Y	D	D^2
18	20	-2	4
16	22	-6	36
18	24	-6	36
12	10	2	4
20	25	-5	25
17	19	-2	4
18	20	-2	4
20	21	-1	1
22	23	-1	1
20	20	0	0
10	10	0	0
8	12	-4	16
20	22	-2	4
12	14	-2	4
16	12	4	16
16	20	-4	16
18	22	-4	16
20	24	-4	16
18	23	-5	25
21	17	4	16
$\sum X = 340$	$\sum Y = 380$	$\sum D = -40$	$\sum D^2 = 244$
$\bar{x} = 17$	$\bar{y} = 19$	$\bar{D} = -2$	$n = 20$

Procedimiento:

1. Para cada problema específico se debe determinar el tipo de problema, los grupos a comparar, si son independientes o relacionados, el nivel de medición, la prueba estadística que debe utilizarse y plantear la hipótesis nula que se debe probar de acuerdo al nivel de medición y tipo de problema correspondiente.

Es un problema de Comparación

2 grupos relacionados

Nivel de medición intervalar

H_0 : El nivel de autoestima es igual entre la medición de inicio y la medición hecha al finalizar el taller

Prueba estadística: Prueba t para grupos relacionados

2. Obtener las diferencias (D) entre los puntajes obtenidos en la primera medición y la segunda medición (datos de la columna X menos los datos de la columna Y).
3. Elevar las diferencias al cuadrado (D^2)
4. Sumar las diferencias ($\sum D$) y las diferencias elevadas al cuadrado ($\sum D^2$)
5. Calcular la suma de cuadrados de las diferencias

$$\sum d^2 = 244 - \frac{-40^2}{20} = 244 - \frac{1600}{20} = 244 - 80 = 164$$

6. Obtener la desviación estándar de las diferencias

$$s_D = \sqrt{\frac{164}{20}} = \sqrt{8.2} = 2.864$$

7. Determinar el error estándar de la diferencia media

$$s_D = \frac{2.864}{\sqrt{20-1}} = \frac{2.864}{\sqrt{19}} = \frac{2.864}{4.359} = 0.657$$

8. Calcular el valor de t y los grados de libertad

$$t = \frac{-2}{0.657} = -3.04$$

$$gl = 20 - 1 = 19$$

9. Buscar el valor de t en tabla con el nivel de significancia de 0.05 y con los grados de libertad calculados. El valor de t obtenido de las operaciones debe compararse con ese valor de tabla (ver tabla en la página 21). **La regla de decisión es: si el valor calculado es mayor o igual que el valor en tabla se debe rechazar la H_0 .**

Para este ejercicio, con 19 grados de libertad, al nivel de significancia establecido en el valor de 0.05, se tiene un valor en tabla de 2.093, el valor de t calculado es mayor al de tabla (el signo no debe ser tomado en cuenta) por lo tanto se rechaza H_0 , entonces: el nivel de autoestima es diferente entre la primera y segunda mediciones.

La prueba *t* de Student para muestras relacionadas en el paquete estadístico SPSS se encuentra en **Analizar / Comparar medias / Prueba T para muestras relacionadas**, se presentarían los siguientes resultados para el ejemplo anterior:

Prueba T

En la primer sección se describen las mediciones a comparar y se presenta la correlación entre las mismas:

Estadísticos de muestras relacionadas

	Media	N	Desviación ttp.	Error ttp. de la media
Par 1 Medición 1	17.00	20	3.80	.85
Medición 2	19.00	20	4.82	1.08

Correlaciones de muestras relacionadas

	N	Correlación	Sig.
Par 1 Medición 1 y Medición 2	20	.793	.000

En la siguiente sección se tiene a la prueba estadística propiamente dicha en la que se describen la diferencia media, la desviación estándar de las diferencias, el error estándar de las diferencias, y finalmente la prueba *t*:

La regla de decisión es: si el nivel de significancia calculado es menor o igual (\leq) a 0.05 se debe rechazar la H_0 . Siguiendo el ejemplo se observa un valor de *t* de -3.04 con 19 grados de libertad y un nivel de significancia de 0.007 (ver los datos en el óvalo), este valor de significancia es menor que 0.05 por lo que se debe rechazar la H_0 , entonces se tiene que afirmar que: el nivel de autoestima es diferente entre la primera y la segunda mediciones.

Prueba de muestras relacionadas

	Diferencias relacionadas					t	gl	Sig. (bilateral)
	Media	Desviación ttp.	Error ttp. de la media	95% Intervalo de confianza para la diferencia				
				Inferior	Superior			
Par 1 Medición 1 - Medición 2	-2.00	2.94	.66	-3.38	-.62	-3.044	19	.007

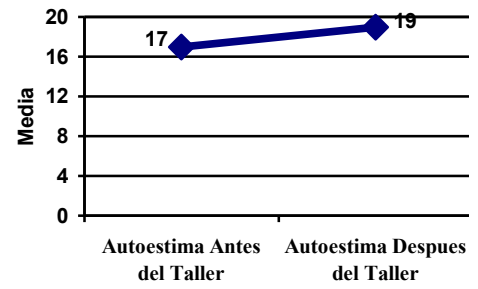
Reporte de resultados de la prueba *t* para muestras relacionadas

Para esta prueba se puede utilizar una gráfica de barras con las medias, aunque es más recomendable una gráfica de líneas para ver las variaciones entre las mediciones, es importante mencionar los valores de la prueba estadística, grados de libertad y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe escribir la conclusión a la que se llega. Por ejemplo:

En cuanto a la efectividad del taller para modificar la autoestima en un grupo de pacientes, se encontraron diferencias en el nivel de autoestima antes y después de asistir al taller ($t = -3.044$, $gl = 19$, $p < 0.01$). Como puede observarse en la gráfica, la media de autoestima es mayor al finalizar el taller (19) que al inicio del mismo (17).

Esta descripción será la base para hacer la discusión posterior de los resultados encontrados en el estudio.

Nivel de Autoestima Antes y Después del Taller*



* $t = -3.044$, $gl = 19$, $p = 0.007$

Prueba U de Mann-Whitney

La prueba U de Mann-Whitney pertenece a las pruebas no paramétricas de comparación de dos muestras independientes, es decir:

- Es libre de curva, no necesita una distribución específica
- Nivel ordinal de la variable dependiente

Se utiliza para comparar dos grupos de rangos (medianas) y determinar que la diferencia no se deba al azar (que la diferencia sea estadísticamente significativa).

Se calcula mediante la siguiente fórmula:

$$U = n_1 n_2 + \frac{n_x(n_x + 1)}{2} - \sum R_x$$

dónde:

n_1, n_2 = número de casos de cada grupo

$\sum R_x$ = suma de rangos de cada grupo

n_x = número de casos de cada grupo

Problema para explicar el procedimiento:

Se desea probar si hay diferencias en el nivel de estrés entre enfermeras de terapia intensiva (X_1) y las de urgencias (X_2), el nivel de estrés se midió en una escala de 0 = nada, 1 = bajo, 2 = medio, 3 = alto, 4 = muy alto, las calificaciones se muestran a continuación:

X_1	X_2	R_1	R_2
3	4	11.5	13
2	3	8	11.5
1	2	3.5	8
1	2	3.5	8
0	1	1	3.5
2	2	8	8
1		3.5	
		$\sum R_1 = 39$	$\sum R_2 = 52$
$n_1 = 7$	$n_2 = 6$		

1. Para cada problema específico se debe determinar el tipo de problema, los grupos a comparar, si son independientes o relacionados, el nivel de medición, la prueba estadística que debe utilizarse y plantear la hipótesis nula que se debe probar de acuerdo al nivel de medición y tipo de problema correspondiente.

Es un problema de Comparación

2 grupos independientes

Nivel de medición ordinal

H_0 : El nivel de estrés es igual entre enfermeras de terapia intensiva (grupo 1) y enfermeras de urgencias (grupo 2)

Prueba estadística: Prueba U de Mann Whitney

2. En primer lugar se debe asignar un rango a cada calificación ordenándolas como si se tratara de un solo conjunto de datos (se juntan los grupos).

Valor	Rango	
0	1	← al valor más pequeño el rango 1, el valor más pequeño posible es el 0
1	2	} 3.5 $\frac{2+3+4+5}{4} = 3.5$
1	3	
1	4	
1	5	
2	6	} 8 $\frac{6+7+8+9+10}{5} = 8$
2	7	
2	8	
2	9	
2	10	} 11.5 $\frac{11+12}{2} = 11.5$
3	11	
3	12	
4	13	← el último rango respetando la secuencia es el 13, que le corresponde al valor 4

para el 2 debe considerarse que en el orden sigue es el 6:
y para 3, empezando con el 11:

En el caso de que existan valores repetidos (llamados también empates), el rango que les corresponde es igual al promedio de los rangos, es decir: la suma de los lugares asignados de manera secuencial entre el número de veces que se repite cada número

3. Cada uno de los rangos se debe colocar en la columna correspondiente, R_1 para el grupo 1 y R_2 para el segundo grupo. Una vez asignados los rangos se deben sumar para cada grupo por separado ($\sum R_1$ y $\sum R_2$).

Esta operación puede verificarse mediante la siguiente fórmula:

$$\sum R_1 + \sum R_2 = \frac{N(N+1)}{2}$$

$$39 + 52 = \frac{13(14)}{2}$$

$$91 = 91$$

4. A continuación debe calcularse el valor de U para cada grupo:

$$U_1 = (7)(6) + \frac{7(7+1)}{2} - 39 = 42 + \frac{(7)(8)}{2} - 39 = 42 + \frac{56}{2} - 39 = 42 + 28 - 39 = 31$$

$$U_2 = (7)(6) + \frac{6(6+1)}{2} - 52 = 42 + \frac{(6)(7)}{2} - 52 = 42 + \frac{42}{2} - 52 = 42 + 21 - 52 = 11$$

5. Se debe encontrar el valor crítico de U en la tabla que se presenta en la página siguiente, con un nivel de significancia establecido a 0.05 para una prueba de 2 colas, utilizando el número de casos de cada grupo (n_1 y n_2). Una vez hecho esto, se debe elegir el valor de U más bajo y compararlo con el valor de U en tabla. **La regla de decisión es: si el valor de U calculado es menor o igual al de tabla, se rechaza la hipótesis nula.**

En este caso con $n_1 = 7$ y $n_2 = 6$, se tiene un valor de U en tabla = 6, el valor obtenido (11) es mayor a este valor por lo que no se puede rechazar la hipótesis nula, se concluye que: el nivel de estrés es igual entre enfermeras de terapia intensiva y enfermeras de emergencia.

1a. Si alguno de los grupos en la muestra tiene más de 20 sujetos, no es posible utilizar esta tabla, para estas situaciones se asume que muestras grandes tienden a distribuirse normalmente y se convierte el valor de U a puntuación z . Esto se hace mediante la siguiente fórmula, en la que se puede usar cualquiera de las dos U que se calcularon, esto no afecta el valor de z , sólo cambia su signo dependiendo de la que se utilice:

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

Con la U más pequeña la z es:

$$z = \frac{11 - \frac{(7)(6)}{2}}{\sqrt{\frac{(7)(6)(7+6+1)}{12}}} = \frac{11 - 21}{\sqrt{\frac{(42)(14)}{12}}} = \frac{-10}{\sqrt{588}} = \frac{-10}{\sqrt{49}} = \frac{-10}{7} = -1.43$$

Tabla de Valores Críticos para U de Mann-Whitney

n_1	n_2											
	1	2	3	4	5	6	7	8	9	10	11	
1	-	-	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	0	0	0	0	0
3	-	-	-	-	0	1	1	2	2	3		
4	-	-	-	0	1	2	3	4	4			
5	-	-	0	1	2	3	5	6	7			
6	-	-	1	2	3	5	6	8				
7	-	-	1	3	5	6	8	10				
8	-	0	2	4	6	8	10	13				
9	-	0	2	4	7	10	12					
10	0	3										

2a. **La regla de decisión es: si la calificación z es mayor o igual a 1.96 (sin tomar en cuenta el signo) se rechaza la hipótesis nula**, esto es porque este valor equivale al 95% del área bajo la curva normal (nivel de significancia de 0.05). Con un valor menor no podemos rechazar la hipótesis nula.

En este caso, al tener una z de -1.43 menor a 1.96 se acepta la hipótesis nula, por lo que: el nivel de estrés es igual entre los dos grupos de enfermeras.

La prueba U de Mann-Whitney en el paquete estadístico SPSS se encuentra en Analizar / Pruebas no paramétricas/2 muestras independientes, los resultados para el ejemplo anterior serían los siguientes:

Pruebas no paramétricas

Rangos				
Grupo	N	Rango promedio	Suma de rangos	
Nivel de estrés				
Enfermeras de terapia intensiva	7	5.57	39.00	
Emfermeras de urgencias	6	8.67	52.00	
Total	13			

Prueba de Mann-Whitney

En primer lugar, se presenta una descripción de los grupos comparados, la suma de rangos y un valor que representa cual grupo tiene una mediana mayor, este dato es el rango promedio, que se obtiene de dividir la suma de rangos de cada grupo entre la cantidad de casos en el grupo.

Estadísticos de contraste

	Nivel de estrés
U de Mann-Whitney	11.000
W de Wilcoxon	39.000
Z	-1.494
Sig. asintót. (bilateral)	.135
Sig. exacta [2*(Sig. unilateral)]	.181 ^a

a. No corregidos para los empates.

b. Variable de agrupación: Grupo

Finalmente, se presentan los valores de la U de Mann-Whitney y de la razón z (ver datos en los óvalos), así como el nivel de significancia de la prueba, se utiliza el de significancia bilateral (ver el número en el hexágono), se debe recordar que en el paquete estadístico **la regla de decisión es: si el nivel de significancia es menor o igual a 0.05 se tiene que rechazar la hipótesis nula**, en este caso es mayor por lo que al no poder rechazar la hipótesis nula debemos concluir que: el nivel de estrés es igual entre los 2 grupos de enfermeras.

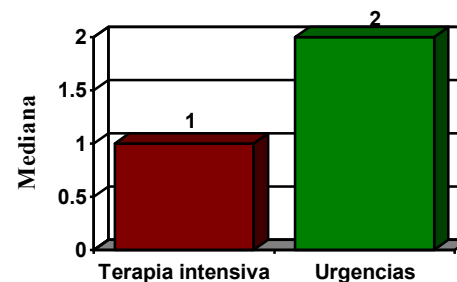
Reporte de resultados de la prueba U de Mann Whitney

Para esta prueba se puede utilizar una gráfica de barras con las medianas, aunque se puede utilizar el rango promedio, obtenido de dividir la suma de rangos de cada grupo entre la cantidad de casos en el grupo, sin embargo este dato es de difícil interpretación. Se deben mencionar los valores de la prueba estadística (U o z), y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe escribir la conclusión a la que se llega. Por ejemplo:

El nivel de estrés es igual entre enfermeras de terapia intensiva y de urgencias ($z = -1.494$, $p > 0.05$). (ver gráfica).

Estos datos se utilizarán para hacer la discusión posterior de los resultados encontrados en el estudio

Nivel de Estrés entre Enfermeras de Terapia Intensiva y de Urgencias*



* $z = -1.494$, $p = 0.135$

Prueba de Wilcoxon

La prueba de rangos asignados de Wilcoxon pertenece a las pruebas no paramétricas de comparación de dos muestras relacionadas, es decir:

- Es libre de curva, no necesita una distribución específica
- Nivel ordinal de la variable dependiente

Se utiliza para comparar dos grupos relacionados de rangos (medianas) y determinar que la diferencia no se deba al azar (que la diferencia sea estadísticamente significativa).

Problema para explicar el procedimiento:

Se desea probar si hay diferencias en el nivel de estrés laboral de los trabajadores de una empresa antes y después de la implementación de un programa de mejoramiento del ambiente laboral; el nivel de estrés se midió en una escala de 0 = nada, 1 = bajo, 2 = medio, 3 = alto, 4 muy alto, las calificaciones se muestran a continuación:

<i>X</i>	<i>Y</i>	<i>D</i>	<i>Rangos positivos</i>	<i>Rangos negativos</i>
3	4	-1		2.5
3	2	1	2.5	
2	1	1	2.5	
2	1	1	2.5	
2	0	2	5.5	
2	2	0	Empate, $n = 7 - 1 = 6$	
1	3	-2		5.5
$n = 7$			$T^+ = 13$	$T^- = 8$

1. Para cada problema específico se debe determinar el tipo de problema, los grupos a comparar, si son independientes o relacionados, el nivel de medición, la prueba estadística que debe utilizarse y plantear la hipótesis nula que se debe probar de acuerdo al nivel de medición y tipo de problema correspondiente.

Es un problema de Comparación

2 grupos relacionados

Nivel de medición ordinal

H_0 : El nivel de estrés laboral de los trabajadores es igual antes y después de implementar el programa de mejoramiento del ambiente laboral.

Prueba estadística: Wilcoxon

2. Se debe obtener la diferencia (D) entre los puntajes de inicio (X) y de la segunda medición (Y), es decir: restar a las calificaciones de la primera medición los de la segunda. A estas diferencias se les debe asignar un rango ordenando los datos sin considerar el signo de la diferencia, deberá empezarse con el valor más pequeño dándole el rango 1, no deben considerarse las diferencias iguales a 0 (empates), cuando esto sucede se debe ajustar el número de sujetos a: $n - 1$, por cada empate.

Ordenar las diferencias mayores a 0

D

-1
1
1
1
-2
2

Rango

En el caso de que existan valores repetidos, el rango que les corresponde es igual a la suma de los lugares asignados entre el número de veces que se repite cada uno*

para los 1: $\frac{1+2+3+4}{4} = 2.5$

para el 2: $\frac{5+6}{2} = 5.5$

* Ver el tratamiento de rangos en la prueba U de Mann Whitney para una explicación más detallada

- Una vez asignados los rangos se deben acomodar los correspondientes a las diferencias positivas y los de diferencias negativas en la columna respectiva y después sumarlos para obtener T^+ y T^- .
- Se debe encontrar el valor de probabilidad asociado a T^+ (c) con el número de casos ajustado (N), en la tabla de valores críticos de T^+ . En ésta se muestra la probabilidad asociada para una prueba de una cola, para obtener la probabilidad de dos colas se debe dividir el valor correspondiente entre 2. Como **regla de decisión: esta probabilidad debe ser menor o igual a 0.05 para rechazar la hipótesis nula.**

En este caso con T^+ igual a 13, al buscar este valor en los renglones y la N ajustada (6) en las columnas, se tiene una probabilidad de 0.3438 que dividida entre 2, para obtener la probabilidad asociada para una prueba de dos colas, da como resultado 0.1719, esta probabilidad al ser mayor a 0.05 no permite rechazar la hipótesis nula. Por lo tanto: el nivel de estrés laboral de los trabajadores es igual antes y después de implementar el programa de mejoramiento del ambiente laboral.

La tabla sin embargo no permite la comparación con más de 15 casos (N ajustada), para estas situaciones, asumiendo que muestras más grandes tienden a distribuirse normalmente, se convierte el valor de T^+ a calificación z calculando lo siguiente:

- En primer lugar se debe obtener la media de T^+ con la siguiente fórmula (usando la N ajustada):

$$\mu_{T^+} = T^+ - \frac{N(N+1)}{4}$$

- Calcular la desviación estándar de T^+ (usando la N ajustada):

$$\sigma_{T^+} = \sqrt{\frac{N(N+1)(2N+1)}{24}}$$

- Calcular la z con:

$$z = \frac{T^+ - \mu_{T^+}}{\sigma_{T^+}}$$

Sustituyendo valores se tiene:

$$\mu_{T^+} = 13 - \frac{6(6+1)}{4} = 13 - \frac{42}{4} = 13 - 10.5 = 2.5$$

$$\sigma_{T^+} = \sqrt{\frac{(6)(6+1)((2)(6)+1)}{24}} = \sqrt{\frac{(42)(13)}{24}} = \sqrt{\frac{546}{24}} = \sqrt{22.75} = 4.77$$

$$z = \frac{2.5}{4.77} = 0.53$$

Valores críticos de T^+ para la prueba de rangos asignados de Wilcoxon de una cola

Las entradas de la tabla para una N dada es $P[T^+ \geq c]$, la probabilidad de que T^+ sea mayor o igual a la suma de c .

	N					
c	3	4	5	6	7	8
3	0.6250					
4	0.3750					
5	0.2500	0.5625				
6	0.1250	0.4375				
7		0.3125				
8		0.1875	0.5000			
9		0.1250	0.4063			
10		0.0625	0.3125			
11			0.2188	0.5000		
12			0.1563	0.4219		
13			0.0938	0.3438		
14			0.0625	0.2813	0.5313	
15			0.0313	0.2188	0.4688	0.5313

La regla de decisión es: si la calificación z es mayor o igual a 1.96 (sin tomar en cuenta el signo) se rechaza la hipótesis nula, esto es porque este valor equivale al 95% del área bajo la curva normal (nivel de significancia de 0.05). Con un valor menor no podemos rechazar la hipótesis nula.

En este caso, al tener una z de 0.53 menor a 1.96, no se puede rechazar la hipótesis nula, por lo que: el nivel de estrés laboral de los trabajadores es igual antes y después de implementar el programa de mejoramiento del ambiente laboral.

Por otro lado, **la prueba de Wilcoxon en el paquete estadístico SPSS se encuentra en Analizar / Pruebas no paramétricas/2 muestras relacionadas**, los resultados que se presentan para el ejemplo anterior son los siguientes:

Pruebas no paramétricas

Prueba de los rangos con signo de Wilcoxon

Rangos

		N	Rango promedio	Suma de rangos
Estrés laboral al iniciar el programa - Estrés al finalizar el programa	Rangos negativos	2 ^a	4.0	8.0
	Rangos positivos	4 ^b	3.2	13.0
	Empates	1 ^c		
	Total	7		

En primer lugar se muestra la asignación de rangos positivos, negativos y empates, así como la suma de rangos positivos y negativos.

- a. Estrés laboral al iniciar el programa < Estrés laboral al finalizar el programa
- b. Estrés laboral al iniciar el programa > Estrés laboral al finalizar el programa
- c. Estrés laboral al finalizar el programa = Estrés laboral al iniciar el programa

Estadísticos de contraste^b

	Estrés al iniciar programa Estrés al finalizar programa
Z	-.541 ^a
Sig. asintót.	.589

- a. Basado en los rangos
- b. Prueba de los rangos con signo de Wilcoxon

Posteriormente se presenta la prueba estadística en este caso el valor de la razón z , así como el nivel de significancia de la prueba (ver valores en el óvalo). **La regla de decisión es: si valor de significancia es menor o igual a 0.05 se debe rechazar la hipótesis nula.** En este caso no se puede rechazar H_0 porque el nivel de significancia de 0.589 es mayor a 0.05, por lo tanto: el nivel de estrés laboral de los trabajadores es igual antes y después de implementar el programa de mejoramiento del ambiente laboral.

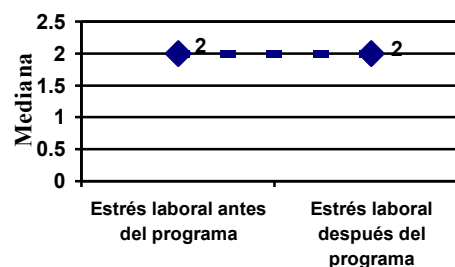
Reporte de resultados de la prueba de Wilcoxon

Para esta prueba se puede utilizar una gráfica de barras o líneas (más adecuada para ver variaciones entre mediciones) con las medianas. Se deben mencionar los valores de la prueba estadística y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe escribir la conclusión a la que se llega. Por ejemplo:

Como puede apreciarse en la gráfica, es igual el nivel de estrés laboral de los trabajadores antes y después de implementar el programa de mejoramiento del ambiente laboral ($z = -0.541, p > 0.05$).

Estos datos se utilizarán para hacer la discusión posterior de los resultados encontrados en el estudio

Nivel de Estrés Laboral Antes y Después del Programa*



* $z = -0.541, p = 0.589$

Prueba de Chi Cuadrada (X^2)

La prueba de Chi cuadrada o Ji cuadrada (X^2) pertenece a las pruebas no paramétricas de comparación de dos o más muestras independientes:

- No se distribuye normalmente, se utiliza la distribución asintótica de Chi cuadrada
- Nivel nominal de la variable dependiente

Se usa para comparar dos o más grupos independientes de proporciones organizadas en una tabla de contingencia y determinar que las diferencias no se deban al azar (que las diferencias sean estadísticamente significativas).

La formula para calcular la X^2 es la siguiente:

$$X^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

Donde f_o son las frecuencias observadas para cada una de las celdas de la tabla y f_e son las frecuencias esperadas para cada una de estas celdas.

Problema para explicar el procedimiento:

Se desea saber si las prácticas de crianza son diferentes entre padres de una orientación política conservadora y padres de orientación política liberal, las prácticas de crianza están clasificadas como no rígida, moderada y autoritaria:

Es un problema de Comparación
2 grupos independientes

Nivel de medición de la variable dependiente: nominal

H_0 : La proporción de padres que emplean prácticas de crianza no rígida, moderada y autoritaria es igual entre los que tienen una orientación política conservadora y los que tienen una orientación liberal.

Prueba estadística: Chi cuadrada

1. En primer lugar debe generarse una tabla de contingencia en la que por regla general los grupos son representados en las columnas y las categorías de la variable dependiente estarán representadas en los renglones. La tabla siguiente, representa los datos cruzados del tipo de orientación política (Variable Independiente en las columnas) por el tipo de práctica de crianza que llevan a cabo (Variable Dependiente en los renglones). En cada celdilla se captura las frecuencias observadas (f_o), obtenidas de las respuestas de los padres según su orientación política (ver las frecuencias en cursivas).

	Conservadores		Liberales		Total
No rígida	<i>6</i>	13.8	<i>24</i>	16.2	30
Moderada	<i>13</i>	12.9	<i>15</i>	15.1	28
Autoritaria	<i>21</i>	13.3	<i>8</i>	15.7	29
Total	40		47		$N = 87$

2. A continuación se deben sumar los totales marginales por cada renglón, por cada columna y obtener el número total de casos (N).

3. Se debe obtener la frecuencia esperada (f_e) multiplicando los totales marginales del renglón y de la columna que corresponden a cada celda, y dividirlos entre el total de casos. Por ejemplo: el total marginal del renglón para la primer celda (con frecuencia observada = 6) es igual a 30 y el total marginal de las columnas es igual a 40, el total de sujetos es 87, entonces la frecuencia esperada para la primer celda es: $f_e = \frac{(30)(40)}{87} = 13.8$

Las frecuencias restantes serían: para la segunda celda (con frecuencia observada = 24): $f_e = \frac{(30)(47)}{87} = 16.2$, para la tercera celda (con frecuencia observada = 13): $f_e = \frac{(28)(40)}{87} = 12.9$,

la cuarta celda (con frecuencia observada = 15): $f_e = \frac{(28)(47)}{87} = 15.1$, la quinta (con frecuencia

observada = 21): $f_e = \frac{(29)(40)}{87} = 13.3$ y finalmente para la sexta (con frecuencia observada = 8):

$f_e = \frac{(29)(47)}{87} = 15.7$ (las frecuencias esperadas son los números en negritas en la tabla).

4. Calcular el valor de X^2 :

$$X^2 = \frac{(6-13.8)^2}{13.8} + \frac{(24-16.2)^2}{16.2} + \frac{(13-12.9)^2}{12.9} + \frac{(15-15.1)^2}{15.1} + \frac{(21-13.3)^2}{13.3} + \frac{(8-15.7)^2}{15.7}$$

$$X^2 = \frac{(-7.8)^2}{13.8} + \frac{(7.8)^2}{16.2} + \frac{(0.1)^2}{12.9} + \frac{(-0.1)^2}{15.1} + \frac{(7.7)^2}{13.3} + \frac{(-7.7)^2}{15.7}$$

$$X^2 = \frac{60.8}{13.8} + \frac{60.8}{16.2} + \frac{0.01}{12.9} + \frac{0.01}{15.1} + \frac{59.29}{13.3} + \frac{59.29}{15.7} = 4.4 + 3.8 + 0 + 0 + 4.5 + 3.8 = 16.5$$

5. Calcular grados de libertad con la siguiente fórmula:

$$gl = (r-1)(c-1)$$

donde r es el número de renglones y c el número de columnas

$$gl = (3-1)(2-1) = (2)(1) = 2$$

6. Buscar en la tabla el valor crítico de Chi cuadrada utilizando los grados de libertad y un nivel de significancia de 0.05 (ver la tabla de valores críticos de Chi cuadrada). **La regla de decisión es: para rechazar la hipótesis nula el valor obtenido debe ser mayor o igual a la chi cuadrada en la tabla.**

En el presente ejemplo se tienen 2 grados de libertad con un nivel de significancia establecido en 0.05, para los que corresponde un valor de X^2 en tabla igual a 5.991. En este ejercicio, al tener una $X^2 = 16.5$ mayor a la de tabla, se rechaza la hipótesis nula, entonces: la proporción de padres que emplean prácticas de crianza no rígida, moderada y autoritaria es diferente entre los que tienen una orientación política conservadora y los que tienen una orientación liberal

Valores críticos de Chi Cuadrada a los Niveles de significancia de 0.05 y 0.01

gl	0.05	0.01
1	3.841	6.635
2	5.991	9.210
3	7.815	11.345
4	9.488	13.277
5	11.070	15.086
6	12.592	
7	14.067	

Prueba de Chi cuadrada para una tabla de 2 X 2

Para la Chi cuadrada con tablas de contingencia de dos por dos, mismas que tienen un grado de libertad, la fórmula varía además incorporando una corrección de continuidad que mejora la aproximación a la distribución muestral de X^2 . Las celdas se enumeran como sigue:

<i>A</i>	<i>B</i>
<i>C</i>	<i>D</i>

La fórmula a utilizar es la siguiente:

$$X^2 = \frac{N \left[(AD - BC) - \frac{N}{2} \right]^2}{(A + B)(C + D)(A + C)(B + D)}$$

Para este caso siempre se tiene que: $gl = 1$

Problema para explicar el procedimiento:

Se desea saber si el consumir alcohol influye sobre las recaídas en fumadores:

Es un problema de Comparación

2 grupos independientes

Nivel de medición de la variable dependiente: nominal (dicotómica)

H_0 : La proporción de fumadores que recayeron en el consumo de tabaco es igual entre bebedores y no bebedores.

Prueba estadística: Chi cuadrada

Los datos se resumen en la siguiente tabla:

	No bebedor	Bebedor	Total
No fumaron	96	13	109
Fumaron	48	20	68
Total	144	33	$N = 177$

$$X^2 = \frac{177 \left[[(96)(20) - (13)(48)] - \frac{177}{2} \right]^2}{(96 + 13)(48 + 20)(96 + 48)(13 + 20)} = \frac{177 [(1920 - 624) - 88.5]^2}{(109)(68)(144)(33)} = \frac{177 [1296 - 88.5]^2}{35221824}$$

$$X^2 = \frac{177 [1207.5]^2}{35221824} = \frac{(177)(1458056.25)}{35221824} = \frac{258075956.25}{35221824} = 7.33$$

Se debe buscar el valor en tabla (ver página 36) con un grado de libertad al nivel de significancia elegido para probar H_0 . **La regla de decisión es: para rechazar la hipótesis nula el valor obtenido debe ser mayor o igual a la chi cuadrada en la tabla.**

Con 1 grado de libertad al nivel de significancia de 0.05 se tiene un valor de X^2 en tabla de 3.841 (ver la tabla de valores críticos de Chi cuadrada en la página 36), en este caso, al tener una chi cuadrada (7.33) mayor a la de la tabla, se rechaza la hipótesis nula, entonces: la proporción de fumadores que recayeron en el consumo de tabaco es diferente entre bebedores y no bebedores.

La Chi cuadrada en el paquete SPSS

En el paquete estadístico la Chi cuadrada se encuentra en **Analizar / Estadísticos descriptivos / Tablas de contingencia**, la variable independiente o de agrupación se coloca siempre en las columnas y la variable dependiente en los renglones, se debe elegir en la sección de estadísticas la prueba de Chi cuadrada, se presentan resultados como los siguientes:

Tablas de contingencia

Tabla de contingencia Tipo Practicas de crianza * Orientación política de los padres

			Orientación política de los padres		Total
			Conservador	Liberal	
Tipo Practicas de crianza	No rígida	Recuento	6	24	30
		% de Tipo Practicas de crianza	20.0%	80.0%	100.0%
		% de Orientación política de los padres	15.0%	51.1%	34.5%
		% del total	6.9%	27.6%	34.5%
	Moderada	Recuento	13	15	28
		% de Tipo Practicas de crianza	46.4%	53.6%	100.0%
		% de Orientación política de los padres	32.5%	31.9%	32.2%
		% del total	14.9%	17.2%	32.2%
	Autoritaria	Recuento	21	8	29
% de Tipo Practicas de crianza		72.4%	27.6%	100.0%	
% de Orientación política de los padres		52.5%	17.0%	33.3%	
% del total		24.1%	9.2%	33.3%	
Total	Recuento	40	47	87	
	% de Tipo Practicas de crianza	46.0%	54.0%	100.0%	
	% de Orientación política de los padres	100.0%	100.0%	100.0%	
	% del total	46.0%	54.0%	100.0%	

La primera parte es la tabla de contingencia con las frecuencias, los porcentajes respecto a los renglones, a las columnas y al total, estos datos nos ayudan a describir las distribuciones de las proporciones.

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	16.313 ^a	2	.000
Razón de verosimilitud	17.184	2	.000
Asociación lineal por lineal	16.125	1	.000
N de casos válidos	87		

a. 0 casillas (.0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 12.87.

Finalmente, se tiene a la prueba estadística de chi cuadrada, donde se debe considerar el valor de la Chi cuadrada de Pearson (ver los resultados en el rectángulo), siguiendo con el primer ejemplo, se tiene un valor de X^2 de 16.313 con $gl = 2$ y una significancia = 0.000, recordando **la regla de decisión: el nivel de significancia debe ser menor o igual a 0.05 para rechazar la hipótesis nula;** encontramos

que la significancia es igual 0.000 menor a 0.05, por lo que se rechaza la hipótesis nula. Entonces, la proporción de padres que emplean prácticas de crianza no rígida, moderada y autoritaria es diferente entre los que tienen una orientación conservadora y los que tienen una orientación liberal.

Para una **tabla de contingencia de dos por dos**, los resultados son similares.

Tablas de contingencia

Tabla de contingencia Recayó con el cigarro * Bebedores

		Bebedores		Total	
		No	Si		
Recayó con el cigarro	No	Recuento	96	13	109
		% de Recayó con el cigarro	88.1%	11.9%	100.0%
		% de Bebedores	66.7%	39.4%	61.6%
		% del total	54.2%	7.3%	61.6%
Si	Recuento	48	20	68	
	% de Recayó con el cigarro	70.6%	29.4%	100.0%	
	% de Bebedores	33.3%	60.6%	38.4%	
	% del total	27.1%	11.3%	38.4%	
Total	Recuento	144	33	177	
	% de Recayó con el cigarro	81.4%	18.6%	100.0%	
	% de Bebedores	100.0%	100.0%	100.0%	
	% del total	81.4%	18.6%	100.0%	

Primero la tabla con frecuencias y porcentajes, siguiendo el segundo problema planteado se tiene

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	8.441 ^b	1	.004		
Corrección por continuidad	7.327	1	.007		
Razón de verosimilitud	8.223	1	.004		
Estadístico exacto de Fisher				.005	.004
Asociación lineal por lineal	8.393	1	.004		
N de casos válidos	177				

a. Calculado sólo para una tabla de 2x2.

b. 0 casillas (.0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 12.68.

Seguida de la prueba estadística, sin embargo, se agregan dos conjuntos de valores, la corrección de continuidad y la prueba exacta de Fisher, se debe recordar que la corrección por continuidad se calcula para una interpretación más exacta, por ello se debe usar ese valor para tomar una decisión (ver datos encerrados en el

rectángulo), en este caso X^2 de 7.327 con $gl = 1$ y una significancia = 0.007, **la regla de decisión indica que el nivel de significancia debe ser menor o igual a 0.05 para rechazar la hipótesis nula**, como vemos la significancia obtenida (0.007) es menor que 0.05, por lo que se rechaza la hipótesis nula, entonces: la proporción de fumadores que recayeron en el consumo de tabaco es diferente entre bebedores y no bebedores.

Restricciones de la Chi cuadrada

Cuando se tiene un 20% de celdillas con frecuencias esperadas menores a 5, no se puede llevar a cabo la prueba de Chi cuadrada, lo mismo sucede si se tienen celdillas con frecuencias iguales a 0.

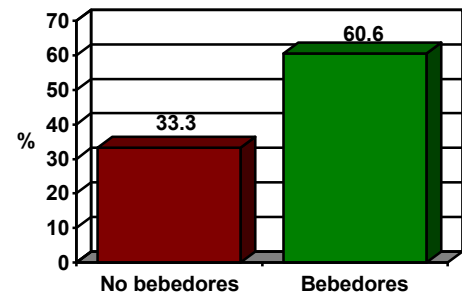
Reporte de resultados de la prueba χ^2

Para esta prueba se puede utilizar una gráfica de barras con los porcentajes por columna que suman el 100% en cada grupo, es importante mencionar los valores de la prueba estadística, grados de libertad y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe escribir la conclusión a la que se llega. Tomando el ejemplo de la tabla de contingencia de 2 x 2:

Se encontraron diferencias en la proporción de fumadores que recayeron en el consumo de tabaco entre bebedores y no bebedores ($\chi^2 = 7.327$, $gl = 1$, $p < 0.01$). Como puede observarse en la gráfica, los bebedores recayeron en el consumo de tabaco en mayor proporción (60.6%) que los no bebedores (33.3%).

Esta descripción será la base para hacer la discusión posterior de los resultados encontrados en el estudio.

Porcentaje de Personas que Volvieron a Fumar entre Bebedores y no Bebedores*



* $\chi^2 = 7.327$, $gl = 1$, $p = 0.007$

Prueba del Cambio de McNemar de dos Mediciones Dicotómicas

La prueba de McNemar es una prueba no paramétrica de comparación de dos muestras relacionadas:

- Es libre de la curva normal, se ajusta a la distribución de Chi cuadrada
- Nivel nominal (sólo variables dicotómicas) de la variable dependiente

Se utiliza para comparar dos grupos relacionados de proporciones organizadas en una tabla que representa los cambios en las respuestas de los sujetos entre una primera medición y una posterior, y determinar que las diferencias no se deban al azar (que las diferencias sean estadísticamente significativas).

Los datos se organizan en una tabla como la siguiente:

		Después	
		-	+
Antes	-	A	B
	+	C	D

Los signos + y - representan las dos categorías de la variable en cada medición, obsérvese que las respuestas que se mantienen en ambas mediciones se encuentran en la celdilla A (- y -) y en la D (+ y +), las celdas donde hay cambio (-, + y +, -) son la B y la C.

La formula para calcular la X^2 es la siguiente:

$$X^2 = \frac{(|B - C| - 1)^2}{B + C}$$

Para la prueba de McNemar siempre se tiene un grado de libertad ($gl = 1$)

Donde B y C son las celdillas en las que se esperan los cambios, esta fórmula incluye una corrección de continuidad que mejora la estimación de la X^2 , esto se hace, como se aprecia en la fórmula, al restar 1 al valor absoluto de la resta de B - C (quitando el signo resultante de la resta).

Problema para explicar el procedimiento:

Se desea saber si cambió la preferencia por Fox o Cárdenas después del debate televisado, para ello se preguntó a un grupo de posibles votantes por cuál candidato votaría antes del día del debate, al día siguiente del debate se volvió a preguntar a los mismos sujetos por quién votarían, los resultados se muestran a continuación:

Es un problema de Comparación

2 grupos relacionados

Nivel de medición de la variable dependiente: nominal (dicotómica)

H_0 : La proporción de posibles votantes que prefieren a Fox o a Cárdenas es igual antes y después del debate.

Prueba estadística: McNemar

1. En primer lugar deben organizarse las respuestas en una tabla como la siguiente. En cada celdilla se capturan la cantidad de sujetos que mantienen su respuesta (Fox – Fox o Cárdenas - Cárdenas) y los que cambian (de Fox a Cárdenas o de Cárdenas a Fox):

		Después	
		Fox	Cárdenas
Antes	Fox	20	26
	Cárdenas	8	23

2. Calcular la Chi cuadrada:

$$X^2 = \frac{(|26 - 8| - 1)^2}{26 + 8} = \frac{17^2}{34} = \frac{289}{34} = 8.5$$

Se debe buscar el valor crítico de X^2 en la tabla de valores de Chi cuadrada (ver página 36) con un grado de libertad y al nivel de significancia establecido para probar la H_0 , **La regla de decisión para rechazar la hipótesis nula es: el valor obtenido debe ser mayor o igual a la chi cuadrada en la tabla.**

El valor de X^2 en tabla con 1 grado de libertad al nivel de significancia igual a 0.05 es 3.841. En este ejemplo se encontró que la X^2 es igual a 8.5, mayor a la de tabla, por esa razón se rechaza la hipótesis nula, entonces: la proporción de posibles votantes que prefieren a Fox o a Cárdenas es diferente antes y después del debate televisado.

McNemar en el paquete SPSS

En el paquete estadístico se tiene la prueba de McNemar en **Analizar / Pruebas no paramétricas / 2 muestras relacionadas** se presentan resultados como los siguientes:

Prueba de McNemar

Tablas de contingencia

Preferencias antes del debate y
Preferencias después del debate

Preferencias antes del debate	Preferencias después del debate	
	1	2
1	20	26
2	8	23

La primera parte es la tabla de frecuencias que se mantienen y que cambian.

Estadísticos de contraste^b

	P preferencias antes del debate y P preferencias después del debate
N	77
Chi-cuadrado ^a	8.500
Sig. asintót.	.004

- a. Corregido por continuidad
- b. Prueba de McNemar

Finalmente se presenta el valor de Chi cuadrada y la significancia asociada (ver datos encerrados en el óvalo), **la regla de decisión es: el nivel de significancia debe ser menor o igual a 0.05 para rechazar la hipótesis nula.** Se tiene que, al tener una significancia de 0.004, menor a 0.05, se rechaza la hipótesis nula. Entonces: la proporción de posibles votantes que prefieren a Fox o a Cárdenas es diferente antes y después del debate televisado.

Reporte de resultados de la prueba de McNemar

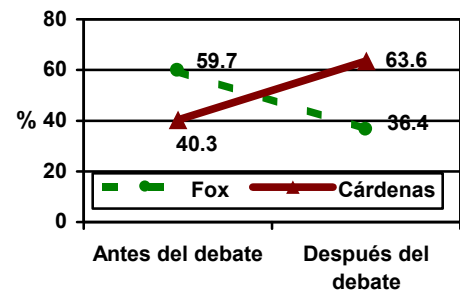
Para esta prueba se puede utilizar una gráfica de barras o líneas (más adecuada para ver variaciones entre mediciones) con los porcentajes obtenidos en cada medición. Se deben mencionar los valores de la prueba estadística y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe escribir la conclusión a la que se llega. Por ejemplo:

Se encontró que la proporción de posibles votantes que prefieren a Fox o a Cárdenas es diferente antes y después del debate televisado ($\chi^2 = 8.5, p < 0.01$).

Como puede apreciarse en la gráfica, aumentó el porcentaje de posibles votantes por Cárdenas entre la primera y segunda medición, mientras que disminuyó la proporción de posibles votantes por Fox.

Estos datos se utilizarán para hacer la discusión posterior de los resultados encontrados en el estudio

Preferencias Electorales Antes y Después del Debate Televisado*



* $\chi^2 = 8.5, p = 0.004$

Análisis de Varianza (Prueba F)

Uno de los usos del análisis de varianza (ANOVA por sus iniciales en inglés) es el comparar tres o más grupos independientes, es una prueba paramétrica, es decir necesita cumplir las siguientes características:

- Selección completamente aleatoria de los grupos
- Homocedasticidad (homogeneidad de las varianzas de los grupos en todos los grupos)
- Distribución normal de la variable dependiente en todos los grupos
- Nivel intervalar de la variable dependiente

El análisis de varianza agrupa una variedad de métodos para análisis de datos de diseños estadísticos muy complicados, tanto que han merecido la publicación de diversos libros en que se describen esos métodos y sus aplicaciones prácticas. En este caso, se presenta la utilización más simple de esta prueba: ANOVA de una vía o factor (una sola variable independiente) para probar una hipótesis basada en las medias de tres o más grupos independientes.

La **prueba F** se calcula mediante la siguiente fórmula:

$$F = \frac{CM_{inter}}{CM_{intra}}$$

dónde CM_{inter} es el cuadrado medio intergrupos y CM_{intra} es el cuadrado medio intragrupos, para calcular ambos datos se deben obtener los siguientes valores:

1. Suma de cuadrados total (SC_T):

$$SC_T = \sum X^2 - \frac{(\sum X)^2}{N}$$

2. La suma de cuadrados intergrupales (SC_{inter}):

$$SC_{inter} = \sum \frac{(\sum X)^2}{n} - \frac{(\sum X)^2}{N}$$

3. La suma de cuadrados intragrupal (SC_{intra}):

$$SC_{intra} = SC_T - SC_{inter}$$

4. Calcular los cuadrados medios (CM) intra e intergrupos:

$$CM_{inter} = \frac{SC_{inter}}{gl_{inter}} \quad CM_{intra} = \frac{SC_{intra}}{gl_{intra}}$$

5. Calcular los grados de libertad intergrupos, intragrupos y total:

$$gl_{inter} = \text{número de grupos } (k) \text{ menos } 1 \ (k - 1)$$

$$gl_{intra} = \text{suma de casos en cada grupo menos } 1: (n_1 - 1) + (n_2 - 1) + \dots + (n_k - 1)$$

$$gl_T = \text{número total de casos } (N) \text{ menos } 1 \ (N - 1)$$

Problema para explicar el procedimiento para Análisis de Varianza de una Vía:

Se desea probar el efecto del empleo de tres métodos de enseñanza de biología: uno multimedia, uno tradicional y uno combinando ambos métodos, para ello se asignó al azar a 21 alumnos en tres grupos de 7 personas cada uno, en cada uno de estos se curso un plan de estudios basado en los diferentes métodos y al final de cada uno de los cursos se aplicó una prueba de conocimientos, los puntajes obtenidos por los sujetos asistentes a cada uno de los cursos se muestran a continuación:

Es un problema de Comparación, 3 grupos independientes, Nivel de medición intervalar, Prueba estadística: Análisis de Varianza

H_0 : El nivel de conocimientos adquiridos en biología es igual entre el grupo del curso multimedia (X_1), el grupo del curso tradicional (X_2) y el grupo del curso que emplea ambos métodos (X_3).

X_1	X_2	X_3	X_1^2	X_2^2	X_3^2
12	6	18	144	36	325
18	4	17	324	16	289
16	14	16	256	196	256
8	4	18	64	16	324
6	6	12	36	36	144
12	12	17	144	144	289
10	14	10	100	196	100
<hr/>			<hr/>		
$\sum X_1 = 82$	$\sum X_2 = 60$	$\sum X_3 = 108$	$\sum X = 250$	$\sum X_1^2 = 1068$	$\sum X_2^2 = 640$
$n_1 = 7$	$n_2 = 7$	$n_3 = 7$	$N = 21$	$\sum X_3^2 = 1727$	$\sum X^2 = 3435$
<hr/>			<hr/>		
$\bar{X}_1 = 11.71$	$\bar{X}_2 = 8.57$	$\bar{X}_3 = 15.43$	$\bar{X}_T = 11.9$		

Pasos a seguir para el cálculo de F

- Una vez que se tienen los puntajes (X_1 , X_2 y X_3) se deben elevar al cuadrado para cada grupo (X_1^2 , X_2^2 y X_3^2), se deben sumar los puntajes por grupo ($\sum X_1$, $\sum X_2$ y $\sum X_3$) y hacer un total de estas sumatorias ($\sum X$), posteriormente sumar los puntajes elevados al cuadrado por grupo ($\sum X_1^2$, $\sum X_2^2$ y $\sum X_3^2$) y también hacer un total de estas sumatorias ($\sum X^2$)
- Determinaremos el número de casos por grupo (n_1 , n_2 y n_3) y la suma de casos total (N) para después obtener las medias por grupo (\bar{X}_1 , \bar{X}_2 y \bar{X}_3) y el promedio total ($\bar{X}_T = \frac{\sum X}{N}$)
- Teniendo los datos anteriores se deben obtener los valores necesarios para calcular la prueba F utilizando las fórmulas descritas anteriormente:

$$\textcircled{1} \quad SC_T = 3435 - \frac{250^2}{21} = 3435 - \frac{62500}{21} = 3435 - 2976.2 = 458.8$$

$$\textcircled{2} \quad SC_{inter} = \left(\frac{82^2}{7} + \frac{60^2}{7} + \frac{108^2}{7} \right) - \frac{250^2}{21} = \left(\frac{6724}{7} + \frac{3600}{7} + \frac{11664}{7} \right) - \frac{62500}{21}$$

$$= 960.6 + 514.3 + 1666.3 - 2976.2 = 3141.2 - 2976.2 = 165$$

$$\textcircled{3} \quad SC_{intra} = 458.8 - 165 = 293.8$$

$$\textcircled{4} \quad CM_{inter} = \frac{165}{2} = 82.5 \quad CM_{intra} = \frac{293.8}{18} = 16.3$$

⑤

$$g^{l_{inter}} = 3 - 1 = 2$$

$$g^{l_{intra}} = (7 - 1) + (7 - 1) + (7 - 1) = 6 + 6 + 6 = 18$$

$$g^{l_T} = 21 - 1 = 20$$

Para finalmente calcular el valor de F :

$$F = \frac{82.5}{16.3} = 5.06$$

Los datos calculados deben ser organizados en una tabla como la siguiente:

Fuentes de Variación	gl	Suma de Cuadrados	Cuadrado medio	F
Intergrupos	2	165	82.5	5.06
Intragrupos	18	293.8	16.3	
Total	20	458.8		

Valores críticos de 5% (tipo claro) y 1% (tipo negro) para la distribución de F

f_1 Grados de libertad (para una media cuadrada mayor)

f_2	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞	f_2	
1	161	200	216	225	230	234	237	239	241	242	243	244	245	246	248	249	250	251	252	253	253	254	254	254	254	1
2	4.052	4.999	5.403	5.625	5.764	5.859	5.928	5.981	6.022	6.056	6.082	6.106	6.142	6.169	6.208	6.234	6.258	6.286	6.302	6.323	6.334	6.352	6.361	6.366	2	
3	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41	19.42	19.43	19.44	19.45	19.46	19.47	19.47	19.48	19.49	19.49	19.50	19.50	19.50	3
4	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.76	8.74	8.71	8.69	8.66	8.64	8.62	8.60	8.58	8.57	8.56	8.54	8.54	8.54	8.53	4
5	34.12	30.82	29.46	28.71	28.24	27.91	29.67	27.49	27.34	27.23	27.13	27.05	26.92	26.83	26.69	26.60	26.50	26.41	26.35	26.27	26.23	26.18	26.14	26.14	26.12	5
6	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91	5.87	5.84	5.80	5.77	5.74	5.71	5.70	5.68	5.66	5.65	5.64	5.64	5.63	6
7	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.40	4.38	4.37	4.36	7	
8	16.26	13.27	12.06	11.39	10.97	10.67	10.45	10.27	10.15	10.05	9.96	9.89	9.77	9.68	9.55	9.47	9.38	9.29	9.24	9.17	9.13	9.07	9.04	9.02	8	
9	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.72	3.71	3.69	3.68	3.67	9	
10	13.74	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72	7.60	7.52	7.39	7.31	7.23	7.14	7.09	7.02	6.99	6.94	6.90	6.88	10	
11	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57	3.52	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.28	3.25	3.24	3.23	11	
12	12.25	9.55	8.45	7.85	7.46	7.19	7.00	6.84	6.71	6.62	6.54	6.47	6.35	6.27	6.15	6.07	5.98	5.90	5.85	5.78	5.75	5.70	5.67	5.65	12	
13	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28	3.23	3.20	3.15	3.12	3.08	3.05	3.03	3.00	2.98	2.96	2.94	2.93	13	
14	11.26	8.65	7.59	7.01	6.63	6.37	6.19	6.03	5.91	5.82	5.74	5.67	5.56	5.48	5.36	5.28	5.20	5.11	5.06	5.00	4.96	4.91	4.88	4.86	14	
15	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07	3.02	2.98	2.93	2.90	2.86	2.82	2.80	2.77	2.76	2.73	2.72	2.71	15	
16	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91	2.86	2.82	2.77	2.74	2.70	2.67	2.64	2.61	2.59	2.56	2.55	2.55	16	
17	10.04	7.56	6.55	5.99	5.64	5.39	5.21	5.06	4.95	4.85	4.78	4.71	4.60	4.52	4.41	4.33	4.25	4.17	4.12	4.05	4.01	3.96	3.93	3.91	17	
18	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79	2.74	2.70	2.65	2.61	2.57	2.53	2.50	2.47	2.45	2.42	2.41	2.40	18	
19	9.65	7.20	6.22	5.67	5.32	5.07	4.88	4.74	4.63	4.54	4.46	4.40	4.29	4.21	4.10	4.02	3.94	3.86	3.80	3.74	3.70	3.66	3.64	3.62	19	
20	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69	2.64	2.60	2.54	2.50	2.46	2.42	2.40	2.36	2.35	2.32	2.31	2.30	20	
21	9.33	6.93	5.95	5.41	5.06	4.82	4.65	4.50	4.39	4.30	4.22	4.16	4.05	3.98	3.86	3.78	3.70	3.61	3.56	3.49	3.46	3.41	3.39	3.37	21	
22	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.72	2.72	2.63	2.63	2.60	2.55	2.51	2.46	2.42	2.38	2.34	2.32	2.28	2.26	2.25	2.24	2.23	22	
23	9.07	6.70	5.74	5.20	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96	3.85	3.78	3.67	3.59	3.51	3.42	3.37	3.30	3.27	3.25	3.24	3.23	23	
24	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.56	2.53	2.48	2.44	2.39	2.35	2.31	2.27	2.24	2.21	2.19	2.18	2.17	2.16	24	
25	8.86	6.51	5.56	5.03	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80	3.70	3.62	3.51	3.43	3.34	3.26	3.21	3.14	3.11	3.09	3.08	3.07	25	
26	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55	2.51	2.48	2.43	2.39	2.33	2.29	2.25	2.21	2.18	2.15	2.13	2.12	2.11	2.10	26	
27	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67	3.56	3.48	3.37	3.29	3.20	3.12	3.07	3.00	2.97	2.95	2.94	2.93	27	
28	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.41	2.36	2.32	2.26	2.22	2.18	2.14	2.11	2.08	2.06	2.05	2.04	2.03	28	
29	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.61	3.54	3.43	3.35	3.24	3.16	3.07	3.00	2.95	2.88	2.85	2.83	2.82	2.81	29	
30	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45	2.41	2.37	2.32	2.28	2.22	2.18	2.14	2.10	2.07	2.04	2.02	2.01	2.00	1.99	30	
31	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.51	3.44	3.33	3.25	3.14	3.06	2.97	2.90	2.85	2.78	2.75	2.73	2.72	2.71	31	
32	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.33	2.28	2.24	2.18	2.14	2.10	2.06	2.03	2.00	1.98	1.97	1.96	1.95	32	
33	8.28	6.01	5.09	4.58	4.25	4.01	3.85	3.71	3.60	3.51	3.43	3.36	3.25	3.17	3.06	2.98	2.89	2.82	2.77	2.70	2.67	2.65	2.64	2.63	33	
34	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43	2.38	2.34	2.30	2.25	2.21	2.15	2.11	2.07	2.03	2.00	1.97	1.95	1.94	1.93	1.92	34	
35	8.18	5.92	5.01	4.50	4.17	3.93	3.77	3.63	3.52	3.43	3.35	3.28	3.17	3.09	2.98	2.90	2.81	2.74	2.69	2.62	2.59	2.57	2.56	2.55	35	
36	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40	2.35	2.31	2.27	2.22	2.18	2.12	2.08	2.04	2.00	1.97	1.94	1.92	1.91	1.90	1.89	36	
37	8.10	5.85	4.94	4.43	4.10	3.86	3.70	3.56	3.45	3.36	3.28	3.21	3.10	3.02	2.91	2.83	2.74	2.67	2.62	2.55	2.52	2.50	2.49	2.48	37	

- Se debe encontrar el valor crítico de F en la tabla, para ello se tiene que identificar, entre el cuadrado medio intergrupos y el cuadrado medio intragrupos, cuál es el mayor; en las columnas se debe buscar con los grados de libertad correspondientes al cuadrado medio mayor y en los renglones los grados de libertad del cuadrado medio menor. Se debe elegir un valor de acuerdo con el nivel de significancia elegido para la prueba de la H_0 . **La regla de decisión para Análisis de Varianza es: si el valor calculado es mayor o igual al valor de tabla, se rechaza la hipótesis nula.**

El cuadrado medio mayor en este caso es el de intergrupos ($CM_{inter} = 82.5$) con $gl_{inter} = 2$, y para el cuadrado medio menor es el de intragrupos ($CM_{intra} = 16.3$) con $gl_{intra} = 18$, buscando en la tabla se tiene, para 2 grados en las columnas y 18 grados de libertad en los renglones, a un nivel de significancia de 0.05, un valor de $F = 3.55$. El valor de F es mayor al de tabla por lo que se rechaza la hipótesis nula, entonces: el nivel de conocimientos adquiridos en biología es diferente entre el grupo del curso multimedia, el grupo del curso tradicional y el grupo del curso que emplea ambos métodos.

Comparaciones Posteriores a F

Cuando se ha determinado que existen diferencias significativas entre los grupos, se debe especificar dónde se encuentran las diferencias, para ello se analizan los grupos por pares, para hacerlo existen varios métodos, más o menos estrictos, diseñados para evitar el error Tipo I debido al efecto Bonferroni.

El efecto Bonferroni se refiere al hecho de que al hacer las diferentes comparaciones entre los grupos organizados por parejas, por ejemplo con la prueba t , el efecto acumulativo sobre el nivel de significancia, del empleo de cada grupo más de una vez para hacer todas las comparaciones (0.05 para la primera, aproximadamente 0.1 para la segunda, aproximadamente 0.15 para la tercera, etc.), nos hace rechazar la hipótesis nula cuando ésta no debería ser rechazada.

Uno de los métodos es el diseñado por Scheffe, en que se utiliza la siguiente fórmula:

$$F = \frac{(X_1 - X_2)^2}{CM_{intra} \frac{(n_1 + n_2)}{n_1 n_2}}$$

Para nuestro ejemplo, al comparar el grupo 1 y el grupo 2, la F es igual a:

$$F = \frac{(11.71 - 8.57)^2}{16.3 \frac{(7+7)}{(7)(7)}} = \frac{3.14^2}{(16.3) \frac{14}{49}} = \frac{9.8596}{(16.3)(0.2857)} = \frac{9.8596}{4.6571} = 2.12$$

Para el grupo 1 y el grupo 3, la F es:

$$F = \frac{(11.71 - 15.43)^2}{16.3 \frac{(7+7)}{(7)(7)}} = \frac{-3.72^2}{(16.3) \frac{14}{49}} = \frac{13.8384}{(16.3)(0.2857)} = \frac{9.8596}{4.6571} = 2.97$$

Para el grupo 2 y el grupo 3, la F es:

$$F = \frac{(8.57 - 15.43)^2}{16.3 \frac{(7+7)}{(7)(7)}} = \frac{-6.86^2}{(16.3) \frac{14}{49}} = \frac{47.0596}{(16.3)(0.2857)} = \frac{47.0596}{4.6571} = 10.1$$

Se debe hacer un ajuste sobre el nivel de significancia, para corregir el efecto Bonferroni, multiplicando el valor de F en tabla, encontrado para la prueba general, por el número de grupos menos 1 ($k - 1$), para este caso el valor en tabla es 3.55 para el nivel de significancia de 0.05 con 2 y 18 grados de libertad, el ajuste se haría como sigue:

$$F_{\text{tabla ajustado}} = (3.55)(3 - 1) = (3.55)(2) = 7.1$$

Las hipótesis nulas que se están probando son: $X_1 = X_2$, $X_1 = X_3$ y $X_2 = X_3$. **La regla de decisión es: el valor de F de cada comparación debe ser mayor o igual al valor de tabla ajustado para rechazar la hipótesis nula.** Como se puede apreciar, sólo el valor de F para la comparación entre el grupo 2 y el grupo 3 es mayor al valor de tabla ajustado, por lo que el nivel de conocimientos

adquiridos en biología es diferente entre el grupo del curso tradicional y el grupo del curso que emplea ambos métodos y podemos decir que las diferencias en el Análisis de Varianza son debidas a este resultado.

Por otro lado, la prueba F de una vía en el paquete estadístico SPSS se encuentra en **Analizar / Comparar medias / ANOVA de un Factor**, se presentarían los siguientes resultados para el ejemplo que estamos utilizando:

ANOVA de un factor

Descriptivos

Calificación en el Examen de Conocimientos

	N	Media	Desviación típica	Error típico	Intervalo de confianza para la media al 95%		Mínimo	Máximo
					Límite inferior	Límite superior		
Multimedia	7	11.71	4.23	1.60	7.80	15.63	6	18
Tradicional	7	8.57	4.58	1.73	4.34	12.80	4	14
Ambos	7	15.43	3.15	1.19	12.51	18.35	10	18
Total	21	11.90	4.78	1.04	9.73	14.08	4	18

En primer lugar se presentan las estadísticas descriptivas para los grupos.

Prueba de homogeneidad de varianzas

Calificación en el Examen de Conocimientos

Estadístico de Levene	gl1	gl2	Sig.
1.264	2	18	.306

Lo siguiente es la prueba de homogeneidad de varianzas entre los grupos, $H_0: s_1^2 = s_2^2 = s_3^2$, la **regla de decisión es: si el nivel de significancia es menor o igual al establecido se rechaza la H_0** , en este caso el nivel de significancia (0.306) es mayor a 0.05, por lo que no se rechaza H_0 , entonces: las varianzas de los grupos son iguales.

ANOVA

Calificación en el Examen de Conocimientos

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	164.952	2	82.476	5.069	.018
Intra-grupos	292.857	18	16.270		
Total	457.810	20			

A continuación se muestran los resultados de la prueba F (suma de cuadrados, grados de libertad, cuadrados medios), el valor de F y la significancia que son los

datos que se utilizan para la regla de decisión (ver datos en el óvalo), **la regla es: si el nivel de significancia encontrado es menor o igual a 0.05 se rechaza H_0** .

La significancia de la prueba F ($\alpha = 0.018$) es menor a 0.05 por lo que se rechaza H_0 , entonces: el nivel de conocimientos adquiridos en biología es diferente entre el grupo del curso multimedia, el grupo del curso tradicional y el grupo del curso que emplea ambos métodos.

Pruebas post hoc

Comparaciones múltiples

Variable dependiente: Calificación en el Examen de Conocimientos

Bonferroni

(I) Método de Enseñanza	(J) Método de Enseñanza	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
Multimedia	Tradicional	3.14	2.16	.486	-2.55	8.83
	Ambos	-3.71	2.16	.306	-9.40	1.98
Tradicional	Multimedia	-3.14	2.16	.486	-8.83	2.55
	Ambos	-6.86*	2.16	.016	-12.55	-1.17
Ambos	Multimedia	3.71	2.16	.306	-1.98	9.40
	Tradicional	6.86*	2.16	.016	1.17	12.55

*. La diferencia entre las medias es significativa al nivel .05.

Para finalizar se presentan los resultados de las comparaciones posteriores, primero la comparación del grupo 1 con el grupo 2 y 3, después el 2 con el 1 y el 3, y al último el 3 con el 1 y 2 (ver flechas).

Esta organización presenta comparaciones redundantes (en este caso se repiten dos veces cada una), únicamente invirtiendo el orden de los grupos, se puede utilizar cualquiera no importando cual grupo se pone primero y cual después, teniendo cuidado de tomar los valores de significancia de todas las comparaciones posibles.

Las hipótesis nulas que se prueban son: $X_1 = X_2$, $X_1 = X_3$ y $X_2 = X_3$. **La regla de decisión para cada comparación es: si el nivel de significancia es menor o igual a 0.05 se rechaza la hipótesis nula.**

Como se puede observar la significancia de la comparación del grupo 1 (multimedia) y 2 (tradicional) es de 0.486 (ver recuadro), del 1 con el 3 (ambos) es 0.306 (ver datos en el hexágono) y del 2 y 3 de 0.016 (ver datos en el óvalo), siguiendo la regla de decisión, la única comparación con diferencias estadísticamente significativas es la correspondiente al grupo 2 (tradicional) vs. el grupo 3 (ambos), entonces: el nivel de conocimientos adquiridos en biología es diferente entre el grupo del curso tradicional y el grupo del curso que emplea ambos métodos y podemos decir que las diferencias en el Análisis de Varianza son debidas a este resultado.

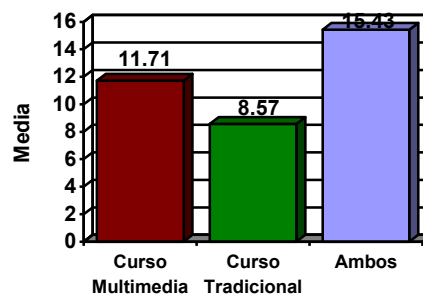
Reporte de resultados de la prueba F

Para esta prueba se puede utilizar una gráfica de barras con las medias de cada grupo, es importante mencionar los valores de la prueba estadística, grados de libertad y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe escribir la conclusión a la que se llega. Por ejemplo:

En lo que respecta a los métodos de enseñanza, se encontraron diferencias en el nivel de conocimientos adquiridos en biología entre el grupo del curso multimedia, el grupo del curso tradicional y el grupo del curso que emplea ambos métodos ($F = 5.069$; $gl = 2, 18$; $p < 0.05$). Este resultado se debe a que se encontraron diferencias en el nivel de conocimientos de biología entre el grupo del curso tradicional y el que emplea ambos métodos ($p < 0.05$), no observándose diferencias en las otras comparaciones.

Como puede observarse en la gráfica, el grupo en que se emplearon ambos métodos tuvo el promedio más alto (15.43), seguido del curso multimedia (11.71), obteniendo la media más baja el grupo del método tradicional (8.57).

Promedio de Calificaciones en Biología por Método de Estudio*



* $F = 5.069$; $gl = 2, 18$; $p = 0.018$
Curso tradicional < Ambos, $p = 0.016$

Esta descripción será la base para hacer la discusión posterior de los resultados encontrados en el estudio.

Análisis de Varianza Unifactorial por Rangos de Kruskal-Wallis

La prueba de Kruskal Wallis pertenece a las pruebas no paramétricas de comparación de tres o más muestras independientes, es decir:

- Es libre de la curva normal, se usa la distribución de chi cuadrada
- Nivel ordinal de la variable dependiente

Se utiliza para comparar tres o más grupos de rangos (medianas) y determinar que las diferencias no se deban al azar (que las diferencias sean estadísticamente significativas).

Se calcula mediante la siguiente fórmula:

$$KW = \left(\frac{12}{N(N+1)} \sum nR^2 \right) - 3(N+1) \dots\dots\dots (1)$$

dónde:

- $N =$ número total de casos
- $n =$ número de casos de cada grupo
- $\bar{R} =$ promedio de rangos de cada grupo

Sin embargo, si cuando se hace la asignación de rangos hay observaciones empatadas, se debe hacer una corrección a la fórmula anterior como sigue:

$$KW = \frac{\left(\frac{12}{N(N+1)} \sum nR^2 \right) - 3(N+1)}{1 - \frac{\sum T}{N^3 - N}} \dots\dots\dots (2)$$

dónde:

$$T = t^3 - t$$

dónde:

$t =$ número de empates en cada conjunto de rangos repetidos

Problema para explicar el procedimiento:

En un hospital, se desea probar si hay diferencias en el nivel de estrés entre enfermeras de terapia intensiva (X_1), las de cirugía (X_2) y las de urgencias (X_3); el nivel de estrés se midió en una escala de 0 = nada, 1 = bajo, 2 = medio, 3 = alto, 4 = muy alto, las calificaciones se muestran a continuación:

X_1	X_2	X_3	R_1	R_2	R_3
3	3	4	13	13	18
2	3	3	8	13	13
1	4	2	3.5	18	8
1	4	4	3.5	18	18
0	2	3	1	8	13
2	2	4	8	8	18
1	1		3.5	3.5	
$n_1 = 7$	$n_2 = 7$	$n_3 = 6$	$\sum R_1 = 40.5$	$\sum R_2 = 81.5$	$\sum R_3 = 88$
	$N = 20$		$\bar{R}_1 = 5.79$	$\bar{R}_2 = 11.64$	$\bar{R}_3 = 14.67$

- Para cada problema específico se debe determinar el tipo de problema, los grupos a comparar, si son independientes o relacionados, el nivel de medición, la prueba estadística que debe utilizarse y plantear la hipótesis nula que se debe probar de acuerdo al nivel de medición y tipo de problema correspondiente.

Es un problema de Comparación

3 grupos independientes

Nivel de medición ordinal

H_0 : El nivel de estrés es igual entre enfermeras de terapia intensiva (grupo 1), enfermeras de cirugía (grupo 2) y enfermeras de urgencias (grupo 3)

Prueba estadística: Kruskal-Wallis

- En primer lugar se debe ordenar cada calificación como si se tratara de un solo conjunto de datos (se juntan los datos de todos los grupos) para asignar rangos.

Valor	Rango	
0	1	← al valor más pequeño el rango 1, el valor más bajo posible es el 0
1	2	} 3.5 el rango para los 1 es igual a $\frac{2+3+4+5}{4} = 3.5$
1	3	
1	4	
1	5	
2	6	} 8 para el 2 debe considerarse que en el orden sigue es el 6: $\frac{6+7+8+9+10}{5} = 8$
2	7	
2	8	
2	9	
2	10	} 13 para 3, empezando con el rango 11: $\frac{11+12+13+14+15}{5} = 13$
3	11	
3	12	
3	13	
3	14	} 18 para 4, empezando con el lugar 16: $\frac{16+17+18+19+20}{5} = 18$
3	15	
4	16	
4	17	
4	18	} En el caso de que existan valores repetidos (llamados también empates), el rango que les corresponde es igual al promedio de los rangos, es decir: la suma de los lugares asignados de manera secuencial entre el número de veces que se repite cada número
4	19	
4	20	

- Se considera a cada número repetido como un conjunto de observaciones empatadas y el número de veces que se repiten es el valor t , como vemos en la tabla anterior el primer conjunto es el correspondiente a los unos, con un rango empatado igual a 3.5, éstos se repiten 4 ocasiones por lo que el valor t para el conjunto de unos es igual a 4. Siguiendo este procedimiento, t para el rango 8 correspondiente a la calificación 2 es 5, después el del rango 13 de la puntuación igual a 3 es 5 y finalmente el rango 18 para los 4 tiene un valor t igual a 5.

Se pueden resumir estos datos, obtener el valor de T y $\sum T$, como sigue:

Conjunto	1 (3.5)	2 (8)	3 (13)	4 (18)	
t	4	5	5	5	
t^3	64	125	125	125	
$T = t^3 - t$	60	120	120	120	$\sum T = 420$

- Los rangos obtenidos se sustituyen en la columna correspondiente en R_1 los asociados a las observaciones del primer grupo, en R_2 a los del segundo y en R_3 a los del tercero, se deben sumar los rangos para cada grupo por separado ($\sum R_1$, $\sum R_2$ y $\sum R_3$) y obtener el promedio de rangos para cada grupo, dividiendo la suma de rangos entre el número de sujetos en el grupo (\bar{R}_1 , \bar{R}_2 y \bar{R}_3).
- A continuación debe calcularse el valor de KW , si no hubo rangos empatados se debe utilizar la fórmula (1) y si hay rangos empatados, como es el caso del presente ejercicio, la fórmula (2):

$$KW = \frac{\frac{12}{20(20+1)}[(7)(5.79)^2 + (7)(11.64)^2 + (6)(14.67)^2] - (3)(20+1)}{1 - \frac{420}{20^3 - 20}}$$

$$KW = \frac{\frac{12}{20(21)}[(7)(33.52) + (7)(135.49) + (6)(215.21)] - (3)(21)}{1 - \frac{420}{8000 - 20}}$$

$$KW = \frac{\frac{12}{420}(234.64 + 948.43 + 1291.26) - 63}{1 - \frac{420}{7980}} = \frac{(0.0286)(2474.33) - 63}{1 - 0.053} = \frac{70.77 - 63}{0.95} = \frac{7.77}{0.95}$$

$$= 8.18$$

- Se debe buscar el valor crítico de KW en tabla correspondiente, para ello se utiliza el número de sujetos de cada grupo y se selecciona el valor crítico de acuerdo al nivel de significancia que se desea utilizar. Sin embargo la tabla no es útil cuando se tiene un grupo con más de 5 casos o se comparan más de tres grupos. En este caso se debe utilizar la tabla de valores críticos de Chi cuadrada (ver página 36) pues podemos asumir que la distribución es similar a la de chi cuadrada. Para esta tabla se necesitan calcular los grados de libertad con $gl = k - 1$, donde k es el número de grupos y establecer el nivel de significancia deseado para probar la hipótesis nula. **La regla de decisión es: si el valor obtenido es mayor o igual al de tabla, se debe rechazar la H_0 .**

En este ejemplo se tienen más de 5 sujetos en todos los grupos por lo que se debe buscar el valor crítico en la tabla de χ^2 con $gl = 3 - 1 = 2$ y el nivel de significancia establecido para probar la hipótesis nula.

El valor en la tabla de chi cuadrada al nivel de significancia de 0.05 es igual a 5.991. KW obtenida es igual a 8.18, mayor a la de tabla por lo que se rechaza la H_0 , entonces: el nivel de estrés es diferente entre enfermeras de terapia intensiva, enfermeras de cirugía y enfermeras de urgencias.

Valores críticos para el análisis de varianza unifactorial por rangos de Kruskal-Wallis (KW)

Tamaño de los grupos			α	
n_1	n_2	n_3	0.05	0.01
3	2	2	4.71	
3	3	1	5.14	
3	3	2	5.36	
3	3	3	5.60	7.20
4	2	2	5.33	
4	3	1	5.21	
4	3	2	5.44	6.44
4	3	3	5.73	7.75
5	3	1	5.99	7.21
5	5	3	5.71	7.54
5	5	4	5.64	7.77
5	5	5	5.78	7.98
Muestras grandes			5.99	9.21

5a. Si se tienen 5 casos o menos en todos los grupos, se debe usar la tabla de valores críticos para KW para rechazar o aceptar la H_0 , el número de casos de cada grupo sirve para obtener el valor en la tabla al nivel de significancia deseado. **La regla de decisión es: si el valor obtenido es mayor o igual al de tabla, se debe rechazar la H_0 .**

Aunque en nuestro ejemplo tenemos más de cinco casos por grupo, podemos usar la tabla con el valor para muestras grandes, ya que se están comparando sólo tres grupos, en ésta encontramos, al nivel de significancia de 0.05, un valor de KW de 5.99, que es igual al de la tabla de chi cuadrada. Al comparar KW obtenida (8.18), encontramos que es mayor al valor de tabla, por lo que se rechaza la hipótesis nula, por lo tanto: el nivel de estrés es diferente entre enfermeras de terapia intensiva, enfermeras de cirugía y enfermeras de urgencias.

Comparaciones múltiples entre tratamientos

Cuando se ha determinado que existen diferencias significativas entre los grupos, se debe especificar dónde se encuentran las diferencias, para ello se analizan los grupos por pares, se calcula el valor absoluto de las diferencias entre los promedios de rangos de cada comparación posible, en este caso son 3: grupo 1 vs. grupo 2, grupo 1 vs. grupo 3 y grupo 2 vs. grupo 3. El número de comparaciones por hacer está determinado por:

$$\#_c = \frac{(k)(k-1)}{2} = \frac{(3)(3-1)}{2} = \frac{(3)(2)}{2} = \frac{6}{2} = 3$$

donde k es el número de grupos a comparar (3 para el presente ejemplo)

Las diferencias entre los promedios de los rangos para las tres comparaciones son:

$$|\bar{R}_1 - \bar{R}_2| = |5.79 - 11.64| = |-5.85| = 5.85$$

$$|\bar{R}_1 - \bar{R}_3| = |5.79 - 14.67| = |-8.88| = 8.88$$

$$|\bar{R}_2 - \bar{R}_3| = |11.64 - 14.67| = |-3.03| = 3.03$$

Para probar la hipótesis nula de que el nivel de estrés es igual entre cada pareja de grupos a comparar, debe calcularse el valor crítico de la diferencia con la siguiente fórmula, que debe evaluarse para cada pareja de grupos comparada:

$$z_{\alpha/k(k-1)} \sqrt{\frac{N(N+1)}{12} \left(\frac{1}{n_u} + \frac{1}{n_v} \right)}$$

donde:

$z_{\alpha/k(k-1)}$ es el valor de z para $k(k-1)$ al nivel de significancia establecido

N es el total de casos de todos los grupos

n_u es el número de casos en el primer grupo en la comparación

n_v es el número de casos en el segundo grupo en la comparación

Valores críticos de z para el número de comparaciones múltiples con hipótesis bidireccional

Número de comparaciones (#c)	α
	0.05
1	1.960
2	2.241
3	2.394
4	2.498
5	2.576
6	2.638
7	2.690

El valor de z puede ser tomado de la tabla de valores críticos de z , usando el número de comparaciones que tienen que hacerse, donde podemos observar que, para el nivel de significancia de 0.05, con tres comparaciones se tiene un valor de 2.394.

Para cada comparación la hipótesis a probar es si el nivel de estrés es igual entre los dos grupos comparados. **La regla de decisión es: si la diferencia encontrada entre los grupos comparados es mayor o igual al valor crítico calculado se rechaza la hipótesis nula.**

Sustituyendo en la fórmula para obtener el valor crítico de la diferencia entre los promedios de rangos de los grupos 1 y 2 es el siguiente, se prueba la hipótesis nula de que el nivel de estrés es igual entre el grupo 1 y 2:

$$2.394 \sqrt{\frac{20(20+1)}{12} \left(\frac{1}{7} + \frac{1}{7} \right)} = 2.394 \sqrt{\frac{20(21)}{12} (0.143 + 0.143)} = 2.394 \sqrt{\frac{420}{12} (0.286)} = 2.394 \sqrt{(35)(0.286)}$$

$$= 2.394 \sqrt{10.01} = (2.394)(3.164) = 7.58$$

Para las diferencias entre el grupo 1 y 3, así como para las del 2 y 3, se usa el mismo valor crítico ya que n_u es igual a 7 y n_v es igual a 6 para ambas comparaciones, las hipótesis nulas son: $Md_1 = Md_2$ y $Md_2 = Md_3$:

$$2.394 \sqrt{\frac{20(20+1)}{12} \left(\frac{1}{7} + \frac{1}{6} \right)} = 2.394 \sqrt{\frac{20(21)}{12} (0.143 + 0.167)} = 2.394 \sqrt{\frac{420}{12} (0.31)} = 2.394 \sqrt{(35)(0.31)}$$

$$= 2.394 \sqrt{10.85} = (2.394)(3.294) = 7.89$$

Como se puede apreciar, sólo la diferencia de los promedios de los rangos entre el grupo 1 y el grupo 3 (8.88) es mayor al valor crítico de la diferencia para la comparación entre esos grupos (7.89), por lo que podemos decir que el nivel de estrés es diferente entre estos grupos. Las diferencias encontradas en la prueba de Kruskal-Wallis son debidas a este resultado.

La prueba de Kruskal-Wallis en el paquete estadístico SPSS se encuentra en Analizar / Pruebas no paramétricas / k muestras independientes, los resultados para el ejemplo anterior serían los siguientes:

Prueba de Kruskal-Wallis

Rangos

Grupo	N	Rango promedio
Nivel de estrés Enfermeras de terapia intensiva	7	5.79
Enfermeras de cirugía	7	11.64
Enfermeras de emergencias	6	14.67
Total	20	

En primer lugar, se presentan los grupos comparados, con el número de casos y los rangos promedio, que se obtienen de dividir la suma de rangos de cada grupo entre la cantidad de casos en el grupo.

Estadísticos de contraste^{a,b}

	Nivel de estrés
Chi-cuadrado	8.109
gl	2
Sig. asintót.	.017

a. Prueba de Kruskal-Wallis

b. Variable de agrupación: Grupo

Finalmente, se presentan los valores de la chi cuadrada, así como los grados de libertad y el nivel de significancia de la prueba (ver datos en el círculo), se debe recordar que en el paquete estadístico **la regla de decisión es: si la significancia es menor o igual a 0.05 se debe rechazar la hipótesis nula**, en este caso tenemos una significancia de 0.017, menor a 0.05, por lo que se rechaza la hipótesis nula y debemos concluir que: el nivel de estrés es diferente entre enfermeras de terapia intensiva, enfermeras de cirugía y enfermeras de urgencias.

Nota: Las comparaciones múltiples se pueden hacer mediante U de Mann Whitney, pero se debe recordar ajustar el nivel de significancia que se obtenga en el análisis, multiplicándolo por el número de grupos menos 1: $(k-1)$, si el nivel de significancia después de este ajuste, es menor a 0.05 se rechazará la hipótesis nula de la comparación correspondiente.

Reporte de resultados de la prueba de Kruskal-Wallis

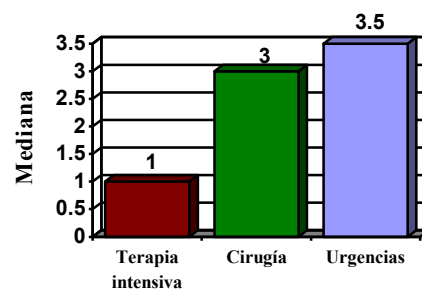
Para esta prueba se puede utilizar una gráfica de barras con las medianas, aunque se puede utilizar la media rango, obtenida de dividir la suma de rangos de cada grupo entre la cantidad de casos en el grupo, aunque este dato es de difícil interpretación. Se deben mencionar los valores de la prueba estadística y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe describir la conclusión a la que se llega. Por ejemplo:

Se observaron diferencias en el nivel de estrés entre enfermeras de terapia intensiva, enfermeras de cirugía y enfermeras de urgencias ($X^2 = 8.109, gl = 1, p < 0.05$).

Como se puede apreciar en la gráfica, las enfermeras en urgencias presentan el nivel de estrés más alto, seguidas de las enfermeras de cirugía y las que tienen el nivel más bajo de estrés son las de terapia intensiva.

Estos datos se utilizarán para hacer la discusión posterior de los resultados encontrados en el estudio

Nivel de Estrés entre Enfermeras de Terapia Intensiva, de Cirugía y de Urgencias*



* $X^2 = 8.109, gl = 1, p = 0.17$

Análisis de Varianza de dos Clasificaciones por Rangos de Friedman

La prueba de Friedman pertenece a las pruebas no paramétricas de comparación de tres o más muestras relacionadas, es decir:

- Es libre de la curva normal, se usa la distribución de chi cuadrada
- Nivel ordinal de la variable dependiente

Se utiliza para comparar tres o más grupos de rangos (medianas) relacionados y determinar que las diferencias no se deban al azar (que las diferencias sean estadísticamente significativas).

Se calcula mediante la siguiente fórmula:

$$F_r = \left(\frac{12}{Nk(k+1)} \sum R^2 \right) - 3N(k+1) \dots\dots\dots (1)$$

dónde:

- N = número total de casos
- k = número de mediciones
- R = suma de rangos de cada medición

Sin embargo, si cuando se hace la asignación de rangos hay observaciones empatadas, se debe hacer una corrección a la fórmula anterior como sigue:

$$F_r = \frac{12 \sum R^2 - 3N^2 k(k+1)^2}{Nk(k+1) + \frac{Nk - \sum t^3}{(k-1)}} \dots\dots\dots (2)$$

dónde:

- t = número de empates en cada conjunto de rangos iguales por cada sujeto, se incluyen los conjuntos con rangos únicos (empates = 1).

Problema para explicar el procedimiento:

Se desea probar si hay diferencias en el nivel de estrés laboral en trabajadores de una empresa antes y después de un taller sobre manejo del estrés en el trabajo, además se realizó medición de seguimiento un mes después de finalizado taller; el nivel de estrés se midió en una escala de 0 = nada, 1 = bajo, 2 = medio, 3 = alto, 4 = muy alto, las calificaciones se muestran a continuación:

X	Y	Z	Rangos para X	Rangos para Y	Rangos para Z
3	2	3	2.5	1	2.5
3	1	2	3	1	2
2	1	1	3	1.5	1.5
4	2	2	3	1.5	1.5
3	1	2	3	1	2
3	0	1	3	1	2
1	2	4	1	2	3
$N = 7$			$R_X = 18.5$	$R_Y = 9$	$R_Z = 14.5$
			$R_X^2 = 342.25$	$R_Y^2 = 81$	$R_Z^2 = 210.25$
			$\sum R^2 = 633.5$		

1. Para cada problema específico se debe determinar el tipo de problema, los grupos a comparar, si son independientes o relacionados, el nivel de medición, la prueba estadística que debe utilizarse y plantear la hipótesis nula que se debe probar de acuerdo al nivel de medición y tipo de problema correspondiente.

Es un problema de Comparación

3 grupos relacionados

Nivel de medición ordinal

H_0 : El nivel de estrés laboral en trabajadores es igual antes del taller para el manejo del estrés en el trabajo, después del taller y en el seguimiento.

Prueba estadística: Friedman

2. Se deben ordenar los datos observados de las mediciones por sujeto y asignar un rango a cada calificación:

Sujeto	Ordenamiento de las mediciones por sujeto de la calificación menor a la mayor en el sentido de la flecha		
	→		
1	2	3	3
2	1	2	3
3	1	1	2
4	2	2	4
5	1	2	3
6	0	1	3
7	1	2	4

Sujeto	Asignar rangos a las calificaciones de cada sujeto de acuerdo al orden establecido, al valor más pequeño de cada sujeto se le otorga el rango 1, siendo el 0 el valor mas bajo posible		
	1	1	2.5
2	1	2	3
3	1.5	1.5	3
4	1.5	1.5	3
5	1	2	3
6	1	2	3
7	1	2	3

Para los datos del presente ejercicio se tienen 3 mediciones por lo que el rango mínimo por asignar a las calificaciones por sujeto es 1 y el máximo es el 3. Sin embargo, cuando existen observaciones repetidas (empates), el rango que les corresponde es igual al promedio de los rangos, es decir: la suma de los lugares asignados de manera secuencial entre el número de veces que se repite cada número, por ejemplo: para el sujeto 1 existe un empate con el 3 que se repite 2

veces en las posiciones 2 y 3 por lo que el rango debe ser: $\frac{2+3}{2} = 2.5$; para el tercer sujeto el 1 se

repite 2 ocasiones en los lugares 1 y 2, el rango correspondiente debe calcularse como sigue:

$\frac{1+2}{2} = 1.5$. Finalmente para el cuarto caso el 2 se repite 2 ocasiones en las posiciones 1 y 2, el

rango correspondiente es: $\frac{1+2}{2} = 1.5$ (estos son los rangos sombreados en la tabla de arriba).

- 2a. Se considera a cada número repetido como observaciones empatadas y el número de veces que se repiten es el valor t , que representa un conjunto de empates por sujeto: el primer conjunto es el del rango 2.5 del caso 1, que se repite 2 veces, por lo que para este conjunto t es igual a 2; el segundo conjunto es el correspondiente al rango empatado igual a 1.5 del sujeto 3, que se repite 2 veces, por lo que t para este conjunto es igual a 2 y el tercer conjunto es el correspondiente al sujeto 4 en el rango 1.5, que se repite 2 veces, con un valor t igual a 2. Es decir, existen tres conjuntos de empates, cada uno con t igual a 2.

Para esta prueba es necesario considerar además a todos los rangos no repetidos como un conjunto con t igual a 1, pues se deben incluir en el ajuste de la fórmula.

Se pueden resumir estos datos como sigue, escribiendo el valor de t y calculando t^3 y $\sum t^3$:

Conjunto	Caso 1 3 (2.5)	Caso 3 1 (1.5)	Caso 4 2 (1.5)	Para los 15 rangos sin repetición	$\sum t^3 = t_1^3 + t_2^3 + t_3^3 + \dots + t_c^3$ siendo c el número total de conjuntos
t	2	2	2	1	$\sum t^3 = 8 + 8 + 8 + (15)(1)$
t^3	8	8	8	1	$\sum t^3 = 39$

- Una vez obtenidos los rangos, se deben colocar en la columna correspondiente a rangos para X , Y y Z considerando la posición original, antes del ordenamiento, de las calificaciones por sujeto. Por ejemplo, para el sujeto 1: la calificación en X es 3 y el rango asignado a X para este sujeto es 2.5, la calificación en Y es 2 y el rango asignado a Y es 1, finalmente la calificación en Z es 3 y el rango asignado a Z para este sujeto es 2.5. Esto debe hacerse para cada caso.
- Teniendo acomodados los rangos, éstos se deben sumar por columna, es decir para cada medición (R_X, R_Y y R_Z) y elevar esas sumas al cuadrado (R_X^2, R_Y^2 y R_Z^2).
- A continuación debe calcularse el valor de Friedman, si no hubo rangos empatados se debe utilizar la fórmula (1) y si hay rangos empatados, como es el caso del presente ejercicio, la fórmula (2):

$$F_r = \frac{(12)(633.5) - (3)(7)^2(3)(3+1)^2}{(7)(3)(3+1) + \frac{(7)(3) - 39}{3-1}}$$

$$= \frac{7602 - (3)(49)(3)(4)^2}{(7)(3)(4) + \frac{21 - 39}{2}}$$

$$\frac{7602 - (3)(49)(3)(16)}{84 + \frac{-18}{2}} = \frac{7602 - 7056}{84 - 9} = \frac{546}{75} = 7.28$$

- Una vez hecho lo anterior, se debe buscar el valor crítico para la prueba de Friedman en la tabla, para obtener este valor se usa el número de mediciones (k) y el número de sujetos (N), además del nivel de significancia que se elija para probar la H_0 . **La regla de decisión es: si el valor de F_r obtenido es mayor o igual al valor en tabla se rechaza la hipótesis nula.**

Valores críticos para el prueba estadística de análisis de varianza de dos clasificaciones por rangos de Friedman (F_r)*

k	N	α	
		0.05	0.01
3	3	6.00	
	4	6.50	8.00
	5	6.40	8.40
	6	7.00	9.00
	7	7.14	8.86
	8	6.25	9.00
	9	6.22	8.67
	10	6.20	9.60
	11	6.54	8.91
	12	6.17	8.67
	13	6.00	9.39
	∞	5.99	7.21
4	2	6.00	
		5.40	

Para este ejercicio, con 3 mediciones y 7 casos, usando un nivel de significancia de 0.05, el valor de tabla es 7.14, el valor de F_r obtenido (7.28) es mayor por lo que H_0 se rechaza, por lo tanto: el nivel de estrés laboral de los trabajadores es diferente antes del taller para el manejo del estrés en el trabajo, después del taller y en el seguimiento.

- Si se tienen más casos de los que pueden encontrarse en la tabla de Friedman, se puede asumir que la distribución es similar a la de Chi cuadrada, por lo que se debe usar esa tabla (ver página 36) para encontrar el valor crítico para probar la hipótesis nula. Los grados de libertad se calculan con: $gl = k - 1$, donde k es el número de mediciones. **La regla de decisión es: si el valor obtenido es mayor o igual al de tabla, se debe rechazar la H_0 .**

Comparaciones múltiples entre mediciones

Cuando se ha determinado que existen diferencias significativas entre mediciones, se debe especificar donde se encuentran las diferencias, para ello se analizan las mediciones por pares calculando el valor absoluto de las diferencias entre las sumas de los rangos (R_X , R_Y y R_Z) de cada comparación posible, en este caso son 3: R_X vs. R_Y , R_X vs. R_Z y R_Y vs. R_Z . El número de comparaciones por hacer está determinado por:

$$\#c = \frac{(k)(k-1)}{2} = \frac{(3)(3-1)}{2} = \frac{(3)(2)}{2} = \frac{6}{2} = 3$$

donde k es el número de mediciones (3 para el presente ejemplo)

Las diferencias entre los rangos son:

$$|R_X - R_Y| = |18.5 - 9| = |9.5| = 9.5$$

$$|R_X - R_Z| = |18.5 - 14.5| = |4| = 4$$

$$|R_Y - R_Z| = |9 - 14.5| = |-5.5| = 5.5$$

Estas diferencias entre cada pareja de mediciones, deben de compararse con el valor crítico de la diferencia, obtenido de la siguiente fórmula:

$$z_{\alpha/k(k-1)} \sqrt{\frac{Nk(k+1)}{6}}$$

donde:

$z_{\alpha/k(k-1)}$ es el valor de z para $k(k-1)$ al nivel de significancia establecido

N es el total de casos

k es el número de mediciones

El valor de z puede ser tomado de la tabla de valores críticos de z, donde podemos observar que para el nivel de significancia de 0.05 para tres comparaciones, se tiene un valor de 2.394.

El valor crítico de la diferencia, se obtiene al sustituir los valores en la fórmula:

$$2.394 \sqrt{\frac{(7)(3)(3+1)}{6}} = 2.394 \sqrt{\frac{84}{6}} = 2.394 \sqrt{14} = (2.394)(3.742) = 8.958$$

Se deben de probar las siguientes hipótesis nulas: $Md_X = Md_Y$, $Md_X = Md_Z$ y $Md_Y = Md_Z$, la **regla de decisión es: si el valor de las diferencias entre las mediciones comparadas es mayor o igual al valor crítico se rechaza la hipótesis nula.**

Como se puede apreciar, sólo la diferencia entre la suma de rangos de la medición X y la medición Y (9.5) es mayor al valor crítico de la diferencia (8.958), por lo que sólo en este caso se rechaza la hipótesis nula y podemos decir que: el nivel de estrés laboral de los trabajadores es diferente entre estos grupos. Las diferencias significativas encontradas en la prueba de Friedman son debidas a este resultado.

Valores críticos de z para el número de comparaciones múltiples con hipótesis bidireccional

Número de comparaciones (#)	α
	0.05
1	1.960
2	2.241
3	2.394
4	2.498
5	2.576
	2.633

La prueba de Friedman en el paquete estadístico SPSS se encuentra en **Analizar / Pruebas no paramétricas / k muestras relacionadas**, los resultados para el ejemplo anterior serían los siguientes:

Pruebas no paramétricas

Rangos

	Rango promedio
Niv el de estrés antes del Programa	2.64
Niv el de estrés despues del Programa	1.29
Niv el de estrés en el seguimiento	2.07

En primer lugar, se presentan los rangos promedio, que se obtienen de dividir la suma de rangos entre la cantidad de casos.

Estadísticos de contrastē

N	7
Chi-cuadrado	7.280
gl	2
Sig. asintót.	.026

a. Prueba de Friedman

Finalmente, se presentan los valores de la chi cuadrada, así como los grados de libertad y el nivel de significancia de la prueba (ver datos en el círculo), se debe recordar que en el paquete estadístico **la regla de decisión es: si la significancia es menor o igual a 0.05 se debe rechazar la hipótesis nula.**

En este ejemplo es de 0.026, menor a 0.05, por lo que se rechaza la hipótesis nula y debemos concluir que: el nivel de estrés laboral de los trabajadores es diferente antes del taller para el manejo del estrés en el trabajo, después del taller y en el seguimiento.

Nota: Las comparaciones múltiples se pueden hacer mediante la prueba de Wilcoxon, pero se debe recordar ajustar el nivel de significancia que se obtenga en el análisis, multiplicándolo por el número de mediciones menos 1: $(k - 1)$, si el nivel de significancia, después de este ajuste, es menor a 0.05 se rechazará la hipótesis nula de la comparación correspondiente.

Reporte de resultados de la prueba de Friedman

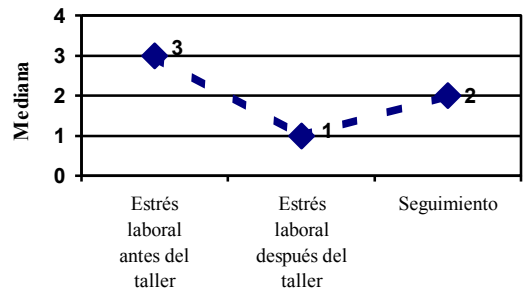
Para esta prueba se puede utilizar una gráfica de barras o líneas (más adecuada para ver variaciones entre mediciones) con las medianas, aunque se puede utilizar la media rango, obtenida de dividir la suma de rangos en cada medición entre la cantidad de sujetos, aunque este dato es de difícil interpretación. Se deben mencionar los valores de la prueba estadística y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe escribir la conclusión a la que se llega. Por ejemplo:

Con relación al efecto del taller para el manejo del estrés en el trabajo, se encontraron diferencias en el nivel de estrés laboral antes del programa, después del programa y en el seguimiento ($X^2 = 7.28$, $gl = 2$, $p < 0.05$).

Como puede apreciarse en la gráfica, el nivel de estrés laboral disminuyó después del taller y tuvo un incremento en el seguimiento, aunque sin llegar al nivel de estrés al inicio del estudio.

Estos datos se utilizarán para hacer la discusión posterior de los resultados encontrados en el estudio

Nivel de Estrés Laboral entre las Tres Mediciones*



* $X^2 = 7.28$, $gl = 2$, $p = 0.026$

Es un problema de Comparación

3 grupos relacionados

Nivel de medición de la variable dependiente: nominal (dicotómica)

H_0 : La proporción de usuarios de drogas es igual al finalizar el programa para eliminar el consumo y en los seguimientos a 3 y 6 meses.

Prueba estadística: Q de Cochran

1. En primer lugar deben contarse los aciertos por sujeto (L), es decir, cuántos unos hay por renglón, y elevar L al cuadrado (L^2)
2. Sumar L y L^2 ($\sum L$ y $\sum L^2$)
3. Contar los aciertos por medición (G_1 , G_2 y G_3), es decir, cuántos unos hay por columna, y elevar G al cuadrado (G_1^2 , G_2^2 y G_3^2)
4. Sumar G y G^2 ($\sum G$ y $\sum G^2$)

Nota: La $\sum L$ y de $\sum G$ son iguales porque ambas se refieren al total de aciertos.

5. Calcular la Q

$$Q = \frac{(3-1)[(3)(347) - 29^2]}{(3)(29) - 63} = \frac{(2)[(3)(347) - 841]}{87 - 63} = \frac{(2)(1041 - 841)}{24} = \frac{(2)(200)}{24} = \frac{400}{24} = 16.667$$

6. Se tiene que encontrar el valor crítico en la tabla de valores críticos de Chi cuadrada (ver página 36) ya que la distribución de esta prueba es similar a la de chi cuadrada, para ellos se tienen que calcular los grados de libertad que se van a utilizar en la tabla con: $gl = k - 1$, donde k es el número de mediciones, además de elegir el nivel de significancia para probar la hipótesis nula. **La regla de decisión establece que: si el valor de Q es mayor o igual al valor de la tabla se debe rechazar la hipótesis nula.**

En la tabla, al nivel de significancia de 0.05, con $gl = 3 - 1 = 2$ para el presente ejemplo, se tiene un valor igual a 5.991. Q es igual a 16.667, mayor al valor de tabla, por lo que se rechaza la H_0 , entonces: la proporción de usuarios de drogas es diferente al finalizar el programa para eliminar el consumo y en los seguimientos a 3 y 6 meses.

Q de Cochran en el paquete SPSS

En el paquete estadístico se tiene la prueba Q de Cochran está en **Analizar / Pruebas no paramétricas / k muestras relacionadas**. Se presentan resultados como los siguientes:

Pruebas no paramétricas

Prueba de Cochran

Frecuencias

	Valor	
	0	1
Usa drogas (1era. medición)	5	13
Usa drogas (2da. medición)	5	13
Usa drogas (3era. medición)	15	3

La primera parte describe la frecuencia de las respuestas por cada medición.

Estadísticos de contraste

N	18
Q de Cochran	16.667 ^a
gl	2
Sig. asintót.	.000

a. 0 se trata como un éxito.

Finalmente se presenta el valor de Q , con los grados de libertad y la significancia asociados (ver datos en el círculo), **la regla de decisión establece que si el nivel de significancia es menor o igual a 0.05 se rechaza la hipótesis nula**, en los resultados obtenidos se tiene una significancia de 0.000, al ser menor a 0.05 se rechaza la hipótesis nula, entonces: la proporción de usuarios de drogas es diferente al finalizar el programa para eliminar el consumo y en los seguimientos a 3 y 6 meses.

Reporte de resultados de la prueba Q de Cochran

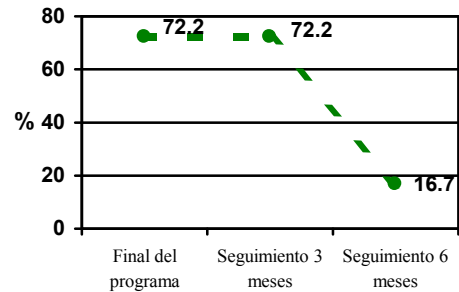
Para esta prueba se puede utilizar una gráfica de barras o líneas (más adecuada para ver variaciones entre mediciones) con los porcentajes obtenidos en cada medición. Se deben mencionar los valores de la prueba estadística y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe escribir la conclusión a la que se llega. Por ejemplo:

Se encontraron diferencias en la proporción de usuarios de drogas al finalizar el programa para eliminar el consumo y en los seguimientos a 3 y 6 meses ($X^2 = 16.667, gl = 2, p < 0.001$).

Como puede apreciarse en la gráfica, la proporción de usuarios se mantuvo a los 3 meses del seguimiento y disminuyó en la siguiente medición a los 6 meses.

Estos datos se utilizarán para hacer la discusión posterior de los resultados encontrados en el estudio

Proporción de Usuarios de Drogas al Finalizar el Programa y en los Seguimientos a 3 y 6 Meses*



* $X^2 = 16.667, gl = 2, p = 0.000$

Pruebas de Asociación

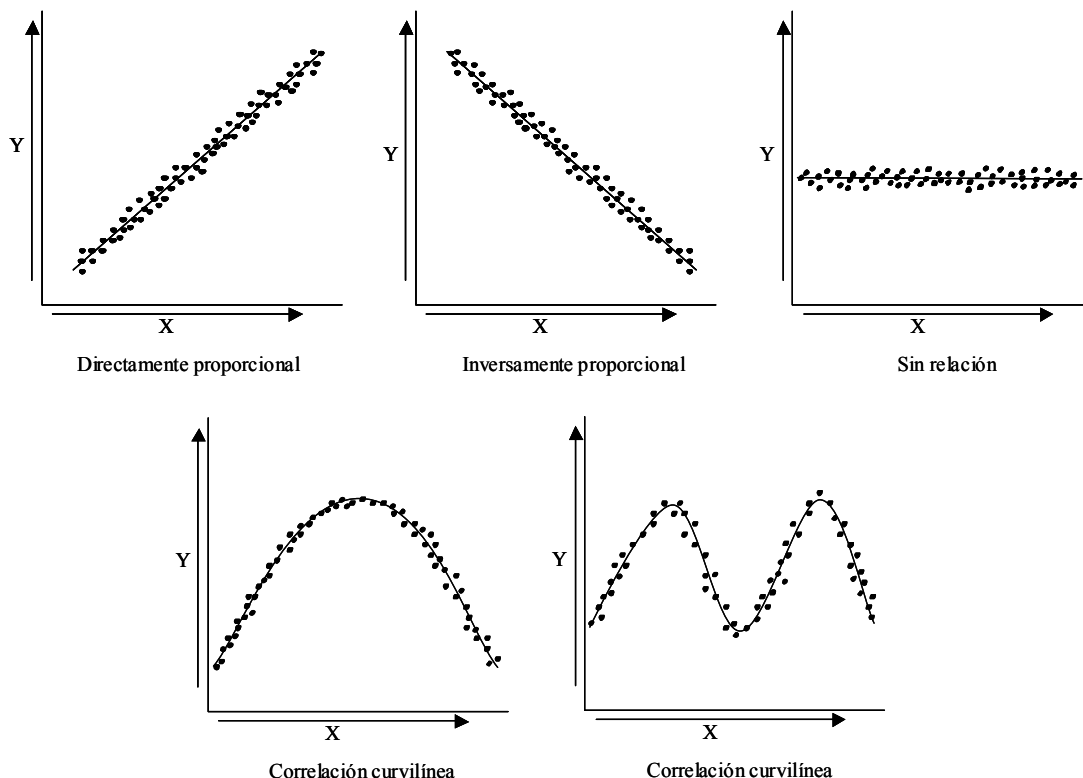
La correlación refleja si existe relación o asociación entre dos variables, no se habla de causalidad por lo que no hay una variable dependiente y una independiente, aunque puede suceder que una variable sea antecedente de otra.

El coeficiente de correlación puede adquirir valores de -1 a 1 , incluyendo el cero, sus características son:

- El signo indica el sentido de la relación, una correlación negativa representa una asociación inversamente proporcional, es decir, a medida que una variable se incrementa, la otra disminuye; por otro lado, en una correlación positiva o directamente proporcional, a medida que aumenta una, la otra también se incrementa.
- La magnitud de la relación esta dada por el valor del coeficiente y se interpreta como sigue: de 0 hasta 0.299 no hay asociación lineal entre las variables, desde 0.3 hasta antes de 0.499 es baja, desde 0.5 hasta antes de 0.799 es moderada y desde 0.8 a 1 hay una correlación alta entre las variables.
- La ausencia de asociación lineal no significa necesariamente que las variables no tengan relación entre sí. Existen muchas variables con una correlación de tipo curvilíneo y debe ser evaluada de otra manera.

Una forma gráfica de ver la relación entre variables son las gráficas de dispersión, donde cada punto representa el lugar en que se cruzan las calificaciones de la primera variable (X) y la segunda variable (Y) para cada sujeto. Si se traza una línea que cruce la mayoría de los puntos se puede apreciar el tipo de relación (ver la siguiente figura). En las relaciones curvas, para algunas magnitudes de X , Y incrementa, pero disminuye en otros valores de X .

Tipos de Correlación



Coeficiente de Correlación Producto-Momento de Pearson

La función de la correlación de Pearson es determinar si existe una relación lineal entre dos variables a nivel intervalar y que esta relación no sea debida al azar (que la relación sea estadísticamente significativa).

La correlación de Pearson se calcula mediante la siguiente fórmula:

$$r_p = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

dónde:

N = número de sujetos

X, Y = calificaciones de las variables a relacionar

Problema para explicar el procedimiento para calcular la correlación de Pearson:

Se desea saber si existe relación entre la autoestima y la depresión en estudiantes de preparatoria. Para ello, un investigador aplica una escala de depresión (X) en la que los estudiantes con calificaciones más altas tienen mayor depresión; también utiliza una escala de autoestima (Y) en la que los puntajes mayores indican mayor autoestima. Los resultados se muestran a continuación:

X	Y	X^2	Y^2	XY
26	8	676	64	208
24	12	576	144	288
18	26	324	676	468
17	29	289	841	493
18	30	324	900	540
7	31	49	961	217
15	19	225	361	285
20	18	400	324	360
18	28	324	784	504
$\sum X = 163$	$\sum Y = 201$	$\sum X^2 = 3187$	$\sum Y^2 = 5055$	$\sum XY = 3363$
$(\sum X)^2 = 26569$	$(\sum Y)^2 = 40401$	$N = 9$		

Procedimiento:

1. Se debe determinar el tipo de problema, el nivel de medición, la prueba estadística que debe utilizarse y plantear la hipótesis nula que va a probar de acuerdo al nivel de medición y tipo de problema correspondiente.

Es un problema de Asociación

Nivel de medición intervalar para ambas variables

H_0 : No hay relación lineal entre el puntaje de depresión y el de autoestima en estudiantes de preparatoria.

Prueba estadística: Correlación Pearson

- Se deben elevar al cuadrado los puntajes obtenidos para la primer variable (X) y los de la segunda (Y) y escribirlos en las columnas X^2 y Y^2 . Multiplicar las dos calificaciones y colocar los resultados en la columna correspondiente (XY).
- Sumar los puntajes de las variables, así como los puntajes elevados al cuadrado de cada variable ($\sum X^2$ y $\sum Y^2$), y elevar las sumas de las calificaciones al cuadrado ($(\sum X)^2$ y $(\sum Y)^2$).
- Calcular la correlación de Pearson:

$$r_p = \frac{(9)(3363) - (163)(201)}{\sqrt{[(9)(3187) - 26569][(9)(5055) - 40401]}} = \frac{30267 - 32763}{\sqrt{(28683 - 26569)(45495 - 40401)}} = \frac{-2496}{\sqrt{(2114)(5094)}}$$

$$= \frac{-2496}{\sqrt{10768716}} = \frac{-2496}{3281.572} = -0.761$$

- Se debe buscar el valor crítico para la correlación Pearson, para ello se calculan los grados de libertad ($gl = N - 1$) y establecer el nivel de significancia para probar la H_0 , **la regla de decisión es: si el coeficiente calculado es mayor o igual al de tabla se rechaza la hipótesis nula.**

En este caso con 9 casos: $gl = 9 - 1 = 8$, por lo que r de tabla al nivel de significancia de 0.05 es 0.6319. Observando el coeficiente obtenido de -0.761, éste es mayor al compararlo con el de tabla (no se debe considerar el signo) por lo que la hipótesis nula se rechaza, entonces: existe una asociación lineal entre el puntaje de depresión y el de autoestima en estudiantes de preparatoria. Además esta relación es moderada (está en el rango de 0.05 a 0.8) e inversamente proporcional (tiene signo negativo).

Valores de r para diferentes niveles de significancia

gl	0.05	0.01	0.001
1	.99692	.999877	.9999988
2	.95000	.990000	.99900
3	.8783	.95873	.99116
4	.8114	.91720	.97406
5	.7545	.8745	.95074
6	.7067	.8343	.92493
7	.6664	.7977	.8982
8	.6319	.7646	.8721
9	.6021	.7348	.8471
10	.5760	.7079	.8233
11	.5529	.6835	.8010
12	.5324	.6614	
13	.5139	.6411	
14	.4977	.6226	

La prueba de Pearson en el paquete estadístico SPSS se encuentra en Analizar / Correlaciones / Bivariadas, los resultados para el ejemplo anterior serían los siguientes:

En el paquete estadístico siempre se presenta una matriz de correlación, en ésta se tienen todas las variables incluidas en el análisis en los renglones y se repiten en las columnas, en el lugar en que se cruza cada par de variables se tiene el coeficiente de Pearson y el nivel de significancia obtenido, además del número de casos (ver datos en el círculo), como puede verse en la tabla de resultados, esta organización nos presenta información redundante (la correlación entre la pareja de variables siempre se repite dos veces, se pueden tomar los resultados de cualquiera de las dos). **La regla de decisión es: si el nivel de significancia es menor o igual a 0.05, se rechaza la hipótesis nula,**

En el presente ejemplo se tiene una significancia de 0.017, menor a 0.05, por lo que se rechaza la hipótesis nula, entonces: existe una asociación lineal entre el puntaje de depresión y el de autoestima en estudiantes de preparatoria.

Además esta relación es moderada (está en el rango de 0.05 a 0.8) e inversamente proporcional (tiene signo negativo).

Correlaciones

Correlaciones

		Puntaje de depresión	Puntaje de autoestima
Puntaje de depresión	Correlación de Pearson	1.000	-0.761*
	Sig. (bilateral)	.	.017
	N	9	9
Puntaje de autoestima	Correlación de Pearson	-0.761*	1.000
	Sig. (bilateral)	.017	.
	N	9	9

*. La correlación es significativa al nivel 0,05 (bilateral).

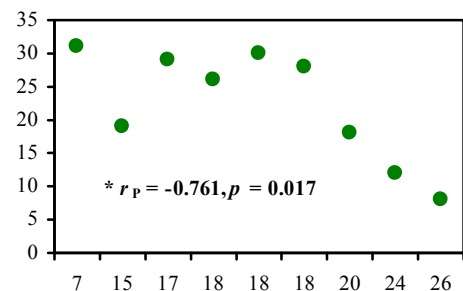
Reporte de resultados de la correlación Pearson

Para esta prueba se puede utilizar una gráfica de dispersión poniendo los valores de la primer variable en X y los de la segunda variable en Y (forma 1), aunque se puede hacer también un diagrama ligando ambas variables con una flecha doble (forma 2). En cualquiera de las dos formas se deben mencionar el valor de la prueba estadística y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe escribir la conclusión a la que se llega. Por ejemplo:

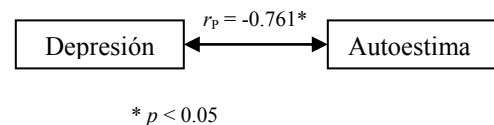
Se encontró una asociación lineal estadísticamente significativa, moderada e inversamente proporcional, entre el puntaje de depresión y el de autoestima en estudiantes de preparatoria ($r_p = -0.761$, $p < 0.05$) (ver gráfica o diagrama).

Estos datos se utilizarán para hacer la discusión posterior de los resultados encontrados en el estudio

Forma 1: Relación entre los Puntajes de Depresión y Autoestima*



Forma 2: Relación entre los Puntajes de Depresión y Autoestima*



Coeficiente de Correlación de Spearman de Rangos Ordenados

La función de la correlación de Spearman es determinar si existe una relación lineal entre dos variables a nivel ordinal, y que esta relación no sea debida al azar (que la relación sea estadísticamente significativa). Aunque una de las variables pueda estar medida a nivel intervalar, de cualquier manera se utiliza esta prueba si una de las dos está medida a nivel ordinal.

La correlación de Spearman se calcula mediante la siguiente fórmula:

$$r_s = 1 - \frac{6 \sum d^2}{N^3 - N} \dots\dots\dots (1)$$

dónde:

N = número de sujetos

d^2 = diferencias entre los rangos asignados a las variables elevadas al cuadrado

Sin embargo, cuando al asignar los rangos hay valores repetidos en alguna de las variables se debe hacer una corrección y usar la fórmula que sigue:

$$r_s = \frac{(N^3 - N) - 6 \sum d^2 - \frac{T_x + T_y}{2}}{\sqrt{(N^3 - N)^2 - (T_x + T_y)(N^3 - N) + T_x T_y}} \dots\dots\dots (2)$$

dónde:

$$T = \sum (t^3 - t)$$

dónde:

t = número de empates en cada conjunto de rangos iguales para cada variable

Problema para explicar el procedimiento para calcular la correlación de Spearman:

Se desea saber si existe relación entre la actitud hacia el psicólogo (X) y la actitud hacia la psicoterapia (Y), en un grupo de padres de familia. La actitud es evaluada como sigue: 1 = muy desfavorable hasta 7 = muy favorable. Los resultados se muestran a continuación:

X	Y	R_X	R_Y	d	d^2
6	7	8	9	-1	1
7	5	9	7	2	4
5	4	7	6	1	1
3	2	5	2.5	2.5	6.25
4	6	6	8	-2	4
2	3	3	4.5	-1.5	2.25
2	3	3	4.5	-1.5	2.25
1	2	1	2.5	-1.5	2.25
2	1	3	1	2	4
$N = 9$					$\sum d^2 = 27$

Procedimiento:

1. Se debe determinar el tipo de problema, el nivel de medición, la prueba estadística que debe utilizarse y plantear la hipótesis nula que se va a probar de acuerdo al nivel de medición y tipo de problema correspondiente.

Es un problema de Asociación

Nivel de medición ordinal para ambas variables

H_0 : No hay relación lineal entre la actitud hacia el psicólogo y la actitud hacia la psicoterapia, en padres de familia.

Prueba estadística: Correlación Spearman

2. En primer lugar se deben ordenar los datos de cada variable de la calificación menor a la mayor y asignarles un rango:

Datos ordenados de la Variable 1 (X)	Rango			
1	1	←	al valor más pequeño el rango 1, el valor más bajo posible es el 0	}
2	2	}	el 2 que se repite 3 veces: $\frac{2+3+4}{3} = 3$	
2	3			
2	4			
3	5	←	al 3 le corresponde el lugar 5	
4	6	←	al 4 le corresponde el lugar 6	
5	7	←	al 5 le corresponde el lugar 7	
6	8	←	al 6 le corresponde el lugar 8	
7	9	←	al 7 le corresponde el lugar 9	

En el caso de que existan valores repetidos (llamados también empates), el rango que les corresponde es igual al promedio de los rangos, es decir: la suma de los lugares asignados de manera secuencial entre el número de veces que se repite cada número

Datos ordenados de la Variable 2 (Y)	Rango			
1	1	←	al valor más pequeño el rango 1, el valor más bajo posible es el 0	}
2	2	}	el 2 que se repite 2 veces en las posiciones 2 y 3: $\frac{2+3}{2} = 2.5$	
2	3			
3	4	}	al 3 repetido 2 veces: $\frac{4+5}{2} = 4.5$	
3	5			
4	6	←	al 4 le corresponde el lugar 6	
5	7	←	al 5 le corresponde el lugar 7	
6	8	←	al 6 le corresponde el lugar 8	
7	9	←	al 7 le corresponde el lugar 9	

En el caso de que existan valores repetidos (llamados también empates), el rango que les corresponde es igual al promedio de los rangos, es decir: la suma de los lugares asignados de manera secuencial entre el número de veces que se repite cada número

2a. Si se tuvieron observaciones empatadas, como en este caso, se debe determinar t para cada conjunto de observaciones repetidas para cada variable por separado, t es el número de veces que se repite cada calificación, y T la sumatoria de $t^3 - t$. Al observar los empates por variable, se tiene para X un conjunto de observaciones repetidas en la calificación 2 que tienen el rango 3, este se repite 3 veces por lo que t para este conjunto es igual 3. Para la variable Y , hay dos conjuntos de rangos empatados, uno en la calificación 2 con rango asignado igual a 2.5, éste se repite dos veces por lo que t para este conjunto es igual a 2; el segundo es para la calificación 3 con rango 4.5 repetido dos veces, por lo que t para este conjunto es igual a 2.

Se pueden resumir estos datos y obtener el valor de T para cada variable como sigue:

Conjuntos de X	2 (3)	
t	3	
t^3	27	
$t^3 - t$	24	$T_x = 24$

Conjuntos de Y	2 (2.5)	3 (4.5)	
t	2	2	
t^3	8	8	
$t^3 - t$	6	6	$T_y = 12$

Es importante considerar que, si sólo una de las variables presenta rangos empatados, el valor de T para la variable sin empates debe ser igual a 0 ($T = 0$).

3. Calcular la correlación de Spearman, en este caso, como hay empates en los rangos asignados a las variables, debe utilizarse la fórmula (2):

$$r_s = \frac{(9^3 - 9) - (6)(27) - \frac{24 + 12}{2}}{\sqrt{(9^3 - 9)^2 - (24 + 12)(9^3 - 9) + (24)(12)}}$$

$$= \frac{(729 - 9) - 162 - \frac{36}{2}}{\sqrt{(729 - 9)^2 - (36)(729 - 9) + 288}} = \frac{720 - 162 - 18}{\sqrt{720^2 - (36)(720) + 288}}$$

$$= \frac{540}{\sqrt{518400 - 25920 + 288}} = \frac{540}{\sqrt{492768}} = \frac{540}{701.974} = 0.769$$

4. Se debe buscar el valor crítico del coeficiente Spearman en la tabla, para ello se utiliza el número de casos (N) y el nivel de significancia elegido para probar la hipótesis nula. **La regla de decisión es: si el coeficiente Spearman calculado es mayor o igual al de la tabla se rechaza la hipótesis nula.**

Valores críticos del coeficiente de correlación de Spearman (r_s) de rangos ordenados, para una prueba bidireccional

N	α		
	0.05	0.01	0.001
5	1.000		
6	0.886	1.000	
7	0.786	0.929	1.000
8	0.738	0.881	0.976
9	0.700	0.833	0.933
10	0.648	0.794	0.903
11	0.618	0.755	0.873
12	0.587	0.727	0.867

El coeficiente en la tabla de valores críticos de correlación Spearman, con 9 casos al 0.05 de significancia, es igual a 0.700, el coeficiente Spearman obtenido (0.769) es mayor al de tabla por lo que se rechaza la hipótesis nula, entonces: existe relación lineal entre la actitud hacia el psicólogo y la actitud hacia la psicoterapia, en padres de familia.

Esta relación es moderada (está en el rango entre 0.5 y 0.8) y directamente proporcional (tiene signo positivo).

Muestras grandes

Cuando la cantidad de sujetos llega a 20 casos y más, se puede verificar la significancia de la prueba convirtiendo el coeficiente de correlación a calificaciones z pues se asume que muestras grandes se distribuyen normalmente:

$$z = r_s \sqrt{N - 1}$$

Por ejemplo, para el presente ejercicio:

$$z = .769 \sqrt{9 - 1} = .769 \sqrt{8} = (.769)(2.828) = 2.175$$

Recuérdese que la z es significativa al nivel de 0.05, cuando es mayor o igual a 1.96, por lo que **la regla de decisión es: si el valor de z es mayor o igual a 1.96 se rechaza la hipótesis nula.** En nuestro ejemplo la z (2.175) es mayor a 1.96, entonces: existe relación lineal entre la actitud hacia el psicólogo y la actitud hacia la psicoterapia, en padres de familia. Además la correlación es moderada (está en el rango entre 0.5 y 0.8) y directamente proporcional (tiene signo positivo).

La prueba de Spearman en el paquete estadístico SPSS se encuentra en **Analizar / Correlaciones / Bivariadas**, los resultados para el ejemplo anterior serían los siguientes:

En el paquete estadístico siempre se presenta una matriz de correlación, en ésta se tienen todas las variables incluidas en el análisis en los renglones y se repiten en las columnas, en el lugar en que se cruza cada par de variables se tiene el coeficiente de Spearman y el nivel de significancia obtenido, además del número de casos (ver resultados en el círculo), esta organización presenta información redundante ya que las correlaciones por cada pareja de variables se repiten siempre 2 veces. **La regla de decisión es: si el nivel de significancia es menor o igual a 0.05 se rechaza la hipótesis nula.**

En el presente ejemplo, el nivel de significancia (0.015) es menor a 0.05, por lo que se rechaza la hipótesis nula, entonces: existe relación lineal entre la actitud hacia el psicólogo y la actitud hacia la psicoterapia, en padres de familia. Además la correlación es moderada (está en el rango entre 0.5 y 0.8) y directamente proporcional (tiene signo positivo).

Correlaciones no paramétricas

Correlaciones

			Actitud hacia el psicólogo	Actitud hacia la psicoterapia
Rho de Spearman	Actitud hacia el psicólogo	Coefficiente de correlación	1.000	.769*
		Sig. (bilateral)	.	.015
		N	9	9
	Actitud hacia la psicoterapia	Coefficiente de correlación	.769*	1.000
		Sig. (bilateral)	.015	.
		N	9	9

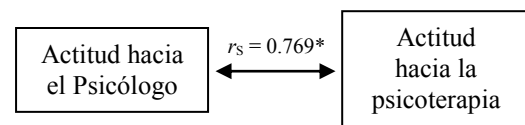
*. La correlación es significativa al nivel 0,05 (bilateral).

Reporte de resultados de la correlación Spearman

Para esta prueba se puede utilizar una gráfica de dispersión poniendo los valores de la primer variable en X y los de la segunda variable en Y , aunque se puede hacer también un diagrama ligando ambas variables con una flecha doble, esta forma es la que se va a usar para ejemplificar los resultados de la correlación Spearman. En cualquiera de las dos opciones se deben mencionar el valor de la prueba estadística y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe escribir la conclusión a la que se llega. Por ejemplo:

Se encontró una relación lineal estadísticamente significativa, moderada y directamente proporcional, entre la actitud hacia el psicólogo y la actitud hacia la psicoterapia, en padres de familia ($r_s = 0.769$, $p = 0.015$) (ver diagrama).

Relación entre la Actitud Hacia el Psicólogo y la Actitud Hacia la Psicoterapia*



Estos datos se utilizarán para hacer la discusión posterior de los resultados encontrados en el estudio

Coeficiente de Correlación Phi

La función de la correlación Phi es determinar si existe una relación entre dos variables a nivel nominal con dos valores cada una (dicotómicas) y que esta relación no sea debida al azar (que la relación sea estadísticamente significativa).

La correlación Phi se calcula mediante la siguiente fórmula:

$$r_{\phi} = \frac{|BC - AD|}{\sqrt{(A+B)(C+D)(A+C)(B+D)}}$$

dónde:

A , B , C y D son las frecuencias observadas en una tabla de contingencia de 2 x 2

A	B	$A+B$
C	D	$C+D$
$A+C$	$B+D$	N

Problema para explicar el procedimiento para calcular la correlación Phi:

Se desea saber si existe relación entre el consumo de alcohol del padre y el consumo de alcohol de estudiantes de preparatoria. Los resultados se resumen en la siguiente tabla de contingencia:

		Consumo de alcohol del padre		
		No bebedor	Bebedor	Total
Consumo de alcohol de estudiantes de Preparatoria	No bebedor	80	30	110
	Bebedor	9	21	30
	Total	89	51	$N = 140$

Procedimiento:

1. Se debe determinar el tipo de problema, el nivel de medición, la prueba estadística que debe utilizarse y plantear la hipótesis nula que va a ser probada de acuerdo al nivel de medición y tipo de problema correspondiente.

Es un problema de Asociación

Nivel de medición nominal (dicotómico) para ambas variables

H_0 : No hay relación entre el consumo de alcohol del padre y el consumo de alcohol de estudiantes de preparatoria.

Prueba estadística: Correlación Phi

2. Se tiene que organizar la información en una tabla de contingencia de dos por dos y obtener los totales por renglón y por columna ($A + B$, $C + D$, $A + C$ y $B + D$).

3. Calcular la correlación Phi:

$$r_{\phi} = \frac{|(9)(30) - (80)(21)|}{\sqrt{(110)(30)(89)(51)}} = \frac{|270 - 1680|}{\sqrt{14978700}} = \frac{|-1410|}{3870.233} = \frac{1410}{3870.233} = 0.364$$

4. El coeficiente Phi está relacionado con la prueba de Chi cuadrada, por esa razón debe calcularse la Chi cuadrada de 2 x 2 (fórmula de la página 37) y comparar el resultado con la tabla de valores críticos que se encuentra en la misma página, **la regla de decisión para rechazar la hipótesis nula al depender de la prueba de Chi cuadrada es: si se rechaza la H_0 relacionada con la prueba de Chi cuadrada, también debe ser rechazada la hipótesis nula de la correlación Phi.**

4a. Calcular la chi cuadrada:

$$X^2 = \frac{140 \left[(80)(21) - (30)(9) - \frac{140}{2} \right]^2}{(110)(30)(89)(51)} = \frac{(140)(1680 - 270 - 70)^2}{14978700} = \frac{(140)(1340)^2}{14978700} = \frac{(140)(1795600)}{14978700}$$

$$X^2 = \frac{251384000}{14978700} = 16.783$$

Recuérdese que se tiene que comparar la Chi cuadrada obtenida con el valor de tabla de valores críticos de Chi cuadrada (ver la tabla en la página 36), este valor se busca utilizando un grado de libertad y el nivel de significancia elegido para probar la hipótesis nula. **La regla de decisión es: si el valor de X^2 obtenido es mayor o igual al de tabla se rechaza la hipótesis nula.**

Con 1 grado de libertad al nivel de significancia de 0.05 se tiene un valor de X^2 en tabla de 3.841, el valor obtenido de X^2 (16.783) es mayor al de tabla por lo que se rechaza la hipótesis nula, por lo tanto la hipótesis nula de la prueba Phi se rechaza también, entonces: existe relación entre el consumo de alcohol del padre y el consumo de alcohol de estudiantes de preparatoria. Además esta relación es baja (está en el rango de 0.3 a 0.5) y directamente proporcional (tiene signo positivo).

La correlación Phi en el paquete SPSS

En el paquete estadístico la Phi se encuentra en **Analizar / Estadísticos descriptivos / Tablas de contingencia**, no es tan importante determinar cuál de las variables va en las columnas y cuál en los renglones ya que no hay variable dependiente e independiente. Se debe elegir en la sección de estadísticas la Phi. Se presentan resultados como los siguientes:

Tablas de contingencia

Tabla de contingencia Consumo de alcohol del estudiante * Consumo de alcohol del padre

Recuento		Consumo de alcohol del padre		Total
		No bebedor	Bebedor	
Consumo de alcohol del estudiante	No bebedor	80	30	110
	Bebedor	9	21	30
Total		89	51	140

La primera parte es la tabla de contingencia con la distribución de frecuencias.

Medidas simétricas

		Valor	Sig. aproximada
Nominal por nominal	Phi	.364	.000
	V de Cramer	.364	.000
N de casos válidos		140	

- a. Asumiendo la hipótesis alternativa.
- b. Empleando el error típico asintótico basado en la hipótesis nula.

Finalmente, se tiene al coeficiente de correlación Phi y el nivel de significancia (ver datos en el recuadro). **La regla de decisión es: si el nivel de significancia es menor o igual a 0.05 se rechaza la hipótesis nula.**

En este ejemplo con un coeficiente Phi de 0.364 y una significancia igual a 0.000,

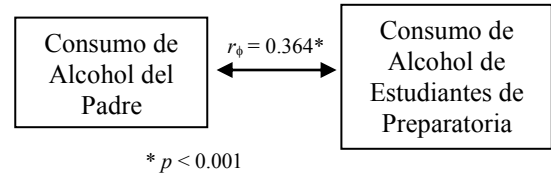
este nivel de significancia es menor a 0.05 por lo que se rechaza la hipótesis nula, es así que: existe relación entre el consumo de alcohol del padre y el consumo de alcohol de estudiantes de preparatoria. Además esta relación es baja (está en el rango de 0.3 a 0.5) y directamente proporcional (tiene signo positivo)

Reporte de resultados de la correlación Phi

Para esta prueba se puede utilizar un diagrama ligando ambas variables con una flecha doble, además se deben mencionar el valor de la prueba estadística y el nivel de significancia con que se está rechazando, o aceptando, la hipótesis nula. Se puede poner el valor de probabilidad obtenido o simplemente decir que es menor, o mayor, al valor que hayamos elegido para la prueba de hipótesis. Además se debe escribir la conclusión a la que se llega. Por ejemplo:

Se encontró una relación estadísticamente significativa, baja y directamente proporcional, entre el consumo de alcohol del padre y el consumo de alcohol de estudiantes de preparatoria. ($r_\phi = 0.364$, $p = 0.000$) (ver diagrama).

Relación entre el Consumo de Alcohol del Padre y el Consumo de Alcohol de Estudiantes de Preparatoria*



Estos datos sirven para hacer la discusión posterior de los resultados encontrados en el estudio

Referencias Bibliográficas

- Brown, F. G. (2000). *Principios de la medición en psicología y educación* (1ra. ed.). México: El Manual Moderno.
- Daniel, W. (1999). *Bioestadística. Bases para el análisis de las ciencias de la salud*. México: Ed. Noriega.
- Downie, M. y Heath, R. (1973). *Métodos estadísticos aplicados*. México: Harper & Row Latinoamericana.
- Ferrán, M. (1996). *SPSS para Windows: Programación y análisis estadístico*. Madrid: Mc Graw Hill.
- Levin, J. (1979). *Fundamentos de estadística en la investigación social* (2da. ed.). México: Editorial Harla.
- Ritchey, F. J. (2002). *Estadística para las Ciencias Sociales: El Potencial de la Imaginación Estadística*. México: Mc Graw Hill.
- Paquete Estadístico para las Ciencias Sociales (SPSS) para Windows versión 10.0.7 en español [Software de computadora]. (2000). Chicago, E. U.: SPSS, Inc.
- Siegel, S. (1990). *Estadística no Paramétrica* (3ra. ed.). México: Editorial Trillas.

