UNIVERSIDAD NACIONAL DE SAN LUIS

FACULTAD DE PSICOLOGÍA

METODOLOGÍA DE LA INVESTIGACIÓN I

Licenciatura en Psicología

Licenciatura en Psicomotricidad

Prof. Responsable: Dr. HORACIO DANIEL GARCIA

Responsable de Trabajos prácticos: Mg. ELIANA ZÁRATE

Auxiliar de Primera: Lic. MAXIMILIANO SAPINO

Unidad 6: Muestreo y Prueba de Hipótesis sobre diferencias
Métodos de muestreo: no probabilísticos y probabilísticos. Error de muestreo. Prueba de hipótesis: hipótesis nula y alternativa. Los errores de tipo I y tipo II. Pruebas de significación de una o dos colas. Valor p. Análisis paramétricos y no paramétricos: supuestos asociados. Pruebas para los valores medios poblacionales. Pruebas para la diferencia de valores medios entre dos muestras independientes. Pruebas para la diferencia de valores medios entre tres o más muestras independientes (ANOVA de un factor). Pruebas para la diferencia de valores medios entre dos muestras relacionadas (dependientes o apareadas).

Autor: HORACIO DANIEL GARCIA

Año 2022

Métodos de muestreo: no probabilísticos y probabilísticos. Error de muestreo.

Muestreo

A lo largo de estos capítulos hemos leído muchas veces las palabras muestra y población. Nos hemos referido sucintamente a ellas en muchas situaciones, sin embargo, conviene precisar estos conceptos, ya que son de crucial importancia a la hora de pretender arribar a conclusiones creíbles y válidas en nuestras investigaciones.

¿Qué es lo que sabemos hasta el momento?

Lo que tenemos en claro es que, en muchas oportunidades, trabajar con la totalidad de los sujetos que componen la población resulta una tarea prácticamente imposible, o cuanto menos antieconómica y poco conveniente. La alternativa para estos casos es buscar un procedimiento que nos permita asegurarnos que, con una cantidad menor de participantes y con un margen de error mínimo, seamos capaces de arribar a conclusiones similares a las que llegaríamos si dispusiéramos de la totalidad de la población para el estudio.

En este sentido, *se denomina muestreo a la técnica que se utiliza para seleccionar los elementos de una población que van a conformar la muestra.* Podrás imaginar que, si esa elección no es cuidadosa, la muestra puede dejar de ser representativa por lo no garantiza que en su composición tenga los rasgos esenciales de la población y que son importantes para la investigación.

Ahora bien, la población es un conjunto finito o infinito de elementos, con características comunes, para los cuales serán extensivas las conclusiones de la investigación; en tanto que la muestra es un subconjunto representativo y finito que se extrae de la población (Arias, 2012). Recordemos que, tanto la muestra como la población quedan delimitadas por el problema y por los objetivos del estudio que han sido planteados; motivo por el cual la decisión de investigar por medio de una muestra, o directamente con la población, será tomada por el investigador y dependerá del tipo y propósito del estudio, de las condiciones de la población y de los recursos que disponga el científico.

Existen dos tipos de poblaciones:

Población finita: en este caso sabemos la cantidad de unidades que la integran. Es posible que exista un registro documental de dichas unidades. Por ejemplo, pacientes hospitalizados en una clínica, los alumnos de una asignatura determinada, etc.

Población infinita: en este caso se desconoce el total de elementos que la conforman, por tal motivo no hay registro documental de ellos. Lógicamente, tener un registro de cada uno de los elementos sería imposible. Por ejemplo, personas que han vivido situaciones estresantes, personas que viven en situación de indigencia, etc.

Un aspecto que tienes que recordar es que, si la población no tiene demasiadas unidades de análisis, y podemos acceder a cada una de ellas, no es necesario extraer una muestra. Sin embargo, cuando nos vemos en la necesidad de trabajar con muestras existen dos estrategias para seleccionar las unidades de análisis de la población, los que veremos a continuación.

Tipos de muestreo

A) Muestreo probabilístico

Hablaremos de *muestro probabilístico* siempre que se cumplan dos condiciones:

1. Todos los elementos de la población tienen una probabilidad mayor a cero de ser seleccionados en la muestra.
2. La probabilidad de inclusión de cada elemento en la muestra se conoce de forma precisa.

Dicho de otro modo, para que tengamos un muestreo probabilístico debemos garantizarnos que todos los elementos de la población tengan posibilidades de ser elegidos para conformar la muestra y además debemos conocer cuál es la probabilidad de que sean incluidos. El cumplimiento de ambos criterios es lo que hace posible obtener resultados no sesgados, cuando investigamos utilizando una muestra.

La definición anterior nos advierte que sólo podemos hacer muestreo probabilístico si disponemos de un ***marco muestral*** (lista de elementos que componen el universo que queremos estudiar y de la cual se extrae la muestra); por ejemplo: el censo de un país, el conjunto de direcciones de hogares en una población o la lista de clientes de una empresa, el listado de alumnos de una escuela, etc. En cada uno de estos casos, el universo a estudiar es diferente: habitantes de un país, hogares de una población, clientes de una empresa, alumnos de una escuela en particular, respectivamente.

Una vez que tenemos un marco muestral, debemos escoger la estrategia para seleccionar la muestra que contemplan las diferentes técnicas de muestreo probabilístico: Muestreo aleatorio simple, muestreo sistemático, muestreo estratificado, muestreo por conglomerados, muestreo desproporcionado...

- *Muestreo aleatorio simple:* es el tipo de muestreo más conocido y con rigor científico ya que garantiza tanto la igualdad de probabilidad de elección de cualquier elemento, como la independencia de selección de cualquier otro. En este procedimiento se extraen al azar un número determinado de elementos, ‘n’, del conjunto mayor ‘N’ o población, de acuerdo a los siguientes pasos:
    1. Definir la población (por ejemplo: Alumnos de 6to año A de la escuela Malvinas Argentinas N=29)
    2. Confeccionar una lista de todos los elementos y asignarles números consecutivos desde 1 hasta ‘N’ (Andrea Ricci 1; Lautaro Correa 2; Patricio Fernández 3; Lucas Gómez 4; (…), Cristóbal Hernández 29)
    3. Definir el tamaño de la muestra: Para calcular el tamaño muestral podemos acceder a muchas herramientas que se ofrecen en internet. En https://www.questionpro.com/es/calculadora-de-muestra.html podemos ver que, si definimos que tenemos una población de 29 estudiantes, con un nivel de confianza del 95% y un error máximo aceptable de 5, el tamaño de la muestra con la que deberíamos trabajar sería de 27 alumnos (anímate a hacer el cálculo, varía los valores, observa e intenta sacar tus propias conclusiones). Es importante que recuerdes que a medida que el N de la población sea más elevado, la muestra será proporcionalmente más pequeña; por ejemplo: para una población de 290 estudiantes, necesitaremos seleccionar unas 165 personas, y para una población de 2900 la muestra será de sólo 339 personas.
    4. Extraer al azar los elementos; siguiendo el ejemplo sortearemos 27 de los 29 números y estos serán los que conformen la muestra con la que trabajaremos. Los procedimientos más comunes de extracción de los elementos en este tipo de muestreo son las tablas de números aleatorios, los sistemas tipo de lotería y las aplicaciones informáticas.

- *Muestreo sistemático:* es un proceso muy simple, sólo requiere la elección de un individuo al azar y luego seleccionar, cada determinada cantidad de unidades muestrales, otro que conformará la muestra hasta llegar a la cantidad necesaria de participantes. Con este procedimiento obtendremos una muestra representativa de la población, siempre y cuando la población sea homogénea y no presente alguna característica muy puntual que pueda incidir en la selección (normalmente suele aplicarse sin inconvenientes en la mayoría de las situaciones, pero es importante tenerlo siempre presente). Veamos el procedimiento:
    1. Definir la población (similar a lo que se realiza al inicio en el muestro aleatorio simple).
    2. Confeccionar una lista de todos los elementos y asignarles números consecutivos desde 1 hasta ‘N’ (similar a lo que se realiza en el muestro aleatorio simple).
    3. Definir el tamaño de la muestra (similar a lo que se realiza en el muestro aleatorio simple).
    4. Definir la amplitud de los fragmentos que serán necesarios para seleccionar a las unidades de análisis que conformarán la muestra. Para este punto, la cantidad total de unidades de análisis, que conforman el marco

muestral (N), se la divide por el número de unidades que conformarán la muestra. Veamos un ejemplo: suponte que nos interesa investigar el efecto que han producido las nuevas medidas implementadas en una empresa que tiene 1550 trabajadores, y que el cálculo del tamaño muestral, con un nivel de confianza del 95% y un error máximo aceptable de 5, ha sido de 309 empleados, ahora nos queda ver la cantidad de unidades de análisis que entran en cada fragmento. Para ello, tomaremos: a) la cantidad de unidades que conforman el marco muestral (N) (1550 trabajadores) y lo dividiremos por b) el tamaño de muestra (n) (309 empleados) 1550/309 = 5,01. El resultado nos indica que cada fragmento de la población estará constituido por 5 personas.

5. Número de inicio: ahora sabemos que la población está dividida en 309 fragmentos constituidos por 5 participantes. Lo que nos queda ahora es saber a quién debemos seleccionar en el primer fragmento. Para que esto sea aleatorio sortearemos al azar un número comprendido entre el 1 y el 5 (recuerda que cada fragmento está conformado por 5 personas), imagina que el número favorecido sea el 4. A este valor se le denomina número de inicio, ya que será el primer participante que conformará la muestra. Entonces, buscaremos la lista de los elementos que conforman la población y extraeremos el participante número 4. ¿Y el resto? ¡como seleccionaremos los 308 participantes que faltan?, bueno, esto lo veremos en el punto que sigue.
6. Selección de los n-1 individuos restantes: los 308 restantes (n-1) serán seleccionados sistemáticamente (por eso este procedimiento se denomina “muestreo sistemático”. Para hacerlo simplemente tomaremos como referencia el valor del primer participante escogido y luego le sumaremos la amplitud del fragmento, en el ejemplo que venimos desarrollando es 5. 4 era el primer participante, ahora le sumamos 5 y obtendremos que 9 será el segundo participante. A 9 le sumamos 5 y nos dará 14, quien será el tercer participante y así seguiremos hasta seleccionar el participante 309.

- *Muestreo estratificado:* en este caso el investigador opta por dividir a la población en diferentes grupos o estratos porque tienen características particulares. Estos grupos tienen que ser homogéneos respecto a la característica que permite agruparlos y no tienen que solaparse. En otras palabras, los participantes que conforman el grupo 1 tienen características similares entre sí y a la vez se diferencian de los participantes que conforman el grupo 2, etc. Por otra parte, la asignación de cada participante es clara: pertenece sólo a un grupo, no puede estar en dos. Si tenemos grupos que se superponen, les dará a algunos individuos de la población mayores probabilidades de ser seleccionados como participantes de la muestra, por lo que dejaría de considerarse un muestreo probabilístico. Los estratos más utilizados son: la edad (por categorías), el género, el nivel socioeconómico, la religión, la nacionalidad, profesión, estado civil, el nivel de estudios alcanzado, etc.

  Como puedes imaginar, cada estrato debe ser pensado independientemente, por lo que puede aplicarse dentro de ellos el muestreo aleatorio simple para elegir los elementos (personas) que formarán parte de la muestra.

  Muy bien, pensemos ahora que vamos a investigar el Bienestar psicológico de las personas que trabajan en el sistema de salud pública de San Luis; ¿cómo deberíamos hacer el muestreo? … Nos informan que en este ámbito trabajan personas que se las pueden clasificar de este modo: a) Personal administrativo; b) Personal de mantenimiento; c) Personal Auxiliar y d) Profesionales de la salud. Lo que deberíamos hacer a continuación es pensar si estas categorías se superponen (si una persona puede estar en dos o más grupos); si confirmamos que estas categorías no se superponen procedemos a asignar los elementos que van a pertenecer a la muestra en cada estrato, proceso que se lo conoce con el término de ***Afijación*.**

  Los tipos de Afijación más utilizados son: 1) *Afijación Simple* (a cada estrato se le asigna la misma cantidad de elementos); y, 2) *Afijación Proporcional* (en este caso se tiene en cuenta la cantidad de elementos que tiene cada estrato en la población, de modo que la muestra -en cada estrato- guarde la misma proporción).

  Es importante que sepas que el muestreo estratificado proporcional es más preciso aún que el muestreo aleatorio simple.

- *Muestreo por conglomerados:* En este muestreo la unidad muestral ya no son los individuos (como en los otros muestreos), sino el conjunto de individuos, a los que se los considera que forman una unidad (grupo o conglomerados). Los conglomerados pueden ser demarcaciones territoriales (barrios, distritos, municipios, etc.) o instituciones (colegios, centros de atención, hospitales, etc.); la condición es que los conglomerados sean homogéneos entre sí (por ejemplo: que todos sean colegios) y se supone que dentro de cada grupo encontraremos la heterogeneidad interna en relación al tema estudiado (por ejemplo: estudiantes de diversos géneros, niveles socioeconómicos, etc). La diferencia que tiene con el muestreo estratificado es que en este último los estratos son homogéneos internamente y heterogéneos entre sí (por ejemplo: cuando estratificamos por nivel de escolaridad); en cambio, en el muestreo por conglomerados, los grupos (conglomerados) son heterogéneos internamente y homogéneos entre ellos (por ejemplo: colegios de San Luis). Por otro lado, en el muestreo estratificado se elige una muestra al interior de cada estrato, mientras que por conglomerados lo que hacemos es, de todos los conglomerados, elegir algunos de ellos (por ejemplo: supongamos que en San Luis hay 200 Colegios; bueno, decides elegir 20 escuelas por lo que tu muestra la compondrán todos los elementos que existen en esos 20 colegios). Existen dos tipos de muestreo por conglomerados:

1. De una etapa: en este caso, supongamos que deseamos estudiar las personas que hacen rehabilitación psicomotriz en centros privados de la ciudad de San Luis. Sabemos que hay unos 8 centros de este tipo y que necesitaremos seleccionar 2 para completar el tamaño de la muestra; entonces lo que haremos será seleccionar aleatoriamente 2 de los 8 centros e incluiremos a todas las personas que asistan a ellos.
2. De dos etapas: es similar al proceso anterior, pero en este caso, en lugar de incluir todos los elementos de los grupos seleccionados de la muestra, se extrae de ellos una muestra aleatoria y se toman los elementos de cada grupo seleccionado.

B) Muestreo no probabilístico

Las dos condiciones del muestreo probabilístico no son fáciles de cumplir cuando investigamos con personas. Algunas veces no tenemos forma de garantizar que reconozcamos a todos los integrantes de la población y, por lo tanto, tampoco podemos saber cuál es la probabilidad asociada a la asignación de cada elemento en la muestra. Por otro lado, el proceso que comprende recabar la información necesaria para extraer una muestra probabilística, seleccionar los participantes, movilizarse y contactar las unidades muestrales suele ser un proceso arduo, muchas veces frustrante, usualmente lento y caro en términos económicos.

Por tales motivos, en ciertas circunstancias los investigadores solemos escoger otras técnicas de muestreo de carácter no probabilístico. Obviamente, tenemos que entender las limitaciones y los requerimientos para que los datos obtenidos en nuestro estudio tengan cierto grado de validez.

Por ejemplo, si estamos estudiando un tema que concierne tanto a hombres como a mujeres, el investigador encuestará tratando de conservar la proporción que hay en la población, de modo que en su muestra finalmente tendrá mitad de hombres y mitad de mujeres. Lo mismo sucederá si el estudio está dirigido a personas de una determinada edad, condiciones sociales, etc.

En el muestreo no probabilístico no significa que el investigador deja de preocuparse por las condiciones de su muestra; al contrario, debe tener muy presente los *criterios de selección* que escoja a modo de asegurarse de arribar a conclusiones significativas.

La diferencia fundamental con las técnicas de muestreo probabilístico radica en que la selección de las unidades muestrales no es aleatoria, por lo que no puede estimarse el error que se asume. Esto es de vital importancia, ya que de ahí podemos decir que un muestreo no probabilístico, si bien nos brinda cierta información de la población de estudio, no podremos establecer con qué precisión nuestras afirmaciones reflejarían las condiciones de dicha población al no poder establecer los márgenes de error y el nivel de confianza.

Entonces, y para dejar aún más claro, con un muestreo probabilístico podremos afirmar con un margen de error establecido las características de la población que han sido estudiadas mediante una muestra; en cambio, en un muestreo no probabilístico podremos afirmar que las características mencionadas corresponden a la muestra de estudio y que “sospechamos” que se encuentran también en la población, no pudiéndolo afirmar rotundamente porque no se pueden generalizar los resultados del estudio.

Métodos de muestreo no probabilístico:

- *Muestreo casual o accidental:* acá los investigadores, sin ningún juicio previo, eligen de la población, de manera casual, a las personas que accidentalmente se encuentren a su disposición y que participarán en el estudio.
- *Muestreo por conveniencia:* en este caso el investigador selecciona a los individuos que le conviene, ya sea porque le resulta más fácil, porque están más cerca, o porque son personas conocidas, etc.
- *Muestreo por cuotas:* en este caso el investigador divide la población en grupos o estratos y luego elige los individuos de cada uno de los grupos sin aplicar ningún método probabilístico.
- *Muestreo de bola de nieve*: este tipo de técnica es útil cuando el investigador desconoce a los individuos o no puede acceder a ellos. La muestra se va conformado con un primer individuo que a su vez va a proponer a otros, y así sucesivamente hasta alcanzar el tamaño muestral que desea el investigador.
- *Muestreo por juicio:* aquí, los participantes son seleccionados en función del tema, conocimiento y juicio del investigador.

Es importante que sepas que no debemos desmerecer a los métodos de muestreo no probabilístico; sólo recuerda que presentan algunas limitaciones. Existen diseños de investigación donde este tipo de selección es muy frecuente; por ejemplo, en los estudios clínicos donde se les solicita a personas que presentan una determinada característica a que sean voluntarios para formar una investigación o bien recibir algún tratamiento. En este caso el investigador asume que la persona que presenta la característica que desea investigar (por ejemplo, depresión) es representativa de la población y que por ende no afectaría los resultados del estudio. Otro uso que se le suele dar a los muestreos no paramétricos es para realizar estudios piloto de modo de obtener información acerca de tendencias, dificultades e incluso resultados que se podrían encontrar en una muestra probabilística.

Frances-Garcia (2019), ofrece una síntesis acerca de las virtudes y diferencias entre ambos tipos de muestreo:

	Probabilístico	**No probabilístico**
¿Dónde se usa con más frecuencia?	En estudios cuantitativos	En estudios cualitativos, aunque también se los utiliza en los estudios cuantitativos
¿Todos los individuos de la población tienen la misma probabilidad de ser incluidos en la muestra?	Si	No
¿Se puede decir en las estimaciones de la muestra cuál es el nivel de confianza y el error muestral?	Si	No
¿Se pueden generalizar a la población los resultados obtenidos en la muestra?	Si	No
Eficiencia para controlar los sesgos en la muestra	Elevado	Bajo
¿A qué aspira?	Lograr inferencias estadísticas	Lograr inferencias lógicas
Tamaño de la muestra	Mayor número de casos	Menor número de casos
Grado de dificultad	Más costoso y complejo	Más económico y sencillo

Error de muestreo

El error de muestreo hace referencia al margen que existe de que la muestra no represente adecuadamente las características de la población. Un poco más precisamente decimos que es la diferencia (debido a la variabilidad de la muestra) entre el valor muestral estimado y el valor verdadero de la característica en la población. Nosotros sabemos que en un proceso de investigación que va de lo particular a lo general (de la muestra a la población) estamos siempre propensos a cometer errores y que en la medida que la muestra sea más pequeña, mayor será la posibilidad de cometer error por no representar adecuadamente la realidad de la población.

Es importante que no te confundas con otros factores que pueden atentar contra la calidad de nuestras conclusiones en un estudio. Los errores ajenos al muestreo, no derivan de la variabilidad probabilística (como sucede con el error de muestreo) sino con: a) deficiencias en la representatividad de la muestra, b) la falta de respuestas de los examinados, c) deficiencias en los instrumentos, o d) bien con sesgos en la selección de los individuos muestreados.

Generalmente, el error muestral que se suele asumir en el ámbito de las ciencias sociales para el cálculo del tamaño de la muestra es alrededor del 5%, es decir E= 0,05; aunque dependiendo del criterio del investigador este puede variar entre el 1% y el 10% (Perelló-Oliver, 2009).

El tamaño de la muestra y el error de muestreo

Si partimos de la base de comparar dos estudios exactamente iguales, que han extraído sus muestras con el mismo método y de la misma población, notarás que el estudio con una muestra de mayor tamaño, tendrá un menor error muestral. Existe un principio básico en la estadística que dice que, a medida que aumentamos el tamaño de la muestra disminuye el error de muestreo, porque, lógicamente, nos acercamos cada vez más el tamaño de la población.

Relación entre la Desviación estándar y el Error de muestreo

Como recuerdas, la desviación estándar es una medida de dispersión que nos dice el grado de variabilidad que hay en la población respecto a su media. En este caso, en la medida que tengamos datos más dispersos, tendremos una muestra más heterogénea y por lo tanto un mayor error de muestreo, independientemente de la técnica que usemos.

Prueba de hipótesis

¿Qué es una hipótesis estadística?

Como recordarás, la hipótesis estadística es una afirmación que hace el investigador acerca de algún aspecto de una variable aleatoria que no conoce y que desea someter a prueba. No todas las hipótesis son estadísticas: por ejemplo, si el investigador afirma: Juan creció en un contexto de violencia; aunque no lo conozca a Juan, y sea válida su afirmación como hipótesis, esta no será de tipo estadística porque no dice nada sobre ninguna variable aleatoria y por lo tanto no tiene que ver con probabilidades. Este tipo de hipótesis son usadas particularmente en estudios de tipo cualitativo y, siguiendo el ejemplo, el investigador entrevistará a Juan e indagará su historia de vida con la finalidad de rechazar o aceptar la hipótesis planteada.

Por el contrario, en pruebas de hipótesis trabajamos con variables cuantitativas, que son las que configuran las hipótesis estadísticas.

La prueba de hipótesis implica un proceso que realiza el investigador para darle el tratamiento estadístico a las hipótesis con la finalidad de determinar si son definitivamente aceptadas o rechazadas. Para ello, el investigador traduce su hipótesis de investigación en términos estadísticos con la finalidad de formalizarlas de un modo claro que permita decantar la decisión en términos de rechazo o aceptación.

> ***“En la prueba de hipótesis planteamos una regla especifica que determina si se puede aceptar o rechazar una afirmación acerca de una población dependiendo de la evidencia proporcionada por una muestra de datos basándonos en probabilidades”***

…. ¿Cómo se hace?...

Bueno, lo primero que debes saber es que la prueba de hipótesis examina dos hipótesis contrapuestas: ***la hipótesis nula*** y la ***hipótesis alternativa***.

Hipótesis nula e Hipótesis Alternativa

La hipótesis nula (H_0): es el enunciado que se probará, y que generalmente expresa que "no hay efecto" o "no hay diferencia".

La hipótesis alternativa (H_1): es el enunciado que el investigador espera poder concluir que es verdadero teniendo en cuenta la evidencia proporcionada por los datos de la muestra.

Entonces, la hipótesis nula (H_0) es una hipótesis que el investigador espera refutar, rechazar o anular, mientras que la hipótesis alternativa (H_1) es lo que el investigador aspira hallar y concuerda con la/s hipótesis planteada/s (ver unidad 3). El análisis que se haga de los datos recogidos, por medio de los estadísticos empleados, permitirá tomar una decisión respecto siempre a H_0 (a la hipótesis nula); es decir, se rechazará o aceptará la H_0.

Un aspecto importante que debes tener en cuenta es que no poder rechazar H_0 no significa un fracaso. Los investigadores inexpertos tienden a pensar que no haber podido refutar H_0 desvaloriza el aporte de la investigación; este es un error muy grande que muchas veces alienta a los investigadores a forzar la interpretación de los datos en un sentido incorrecto. Aunque no se haya podido rechazar H_0, seguramente se han obtenido conocimientos que antes no existían, se han corroborado algunos o refutados otros, como sea, el mundo científico igualmente se ha beneficiado.

Ejemplo: Supongamos que nuestro Planteo del problema incluye una sola variable: *“¿Varía el nivel de ansiedad de los alumnos adolescentes cuando interactúan con profesores?”*.

Sabemos que el valor medio poblacional para esa variable ronda los 40 puntos, por lo que existen tres posibles planteos derivados de esa incógnita.

	Planteo I	**Planteo II**	**Planteo III**
Planteo de la hipótesis	Los adolescentes cuando interactúan con profesores presentan niveles medios de ansiedad DIFERENTE de 40 puntos	Los adolescentes cuando interactúan con profesores presentan niveles medios de ansiedad MAYOR de 40 puntos	Los adolescentes cuando interactúan con profesores presentan niveles medios de ansiedad MENOR de 40 puntos
Formalización	H_o: $\mu = 40$ puntos H_1: $\mu \neq 40$ puntos	H_o: $\mu = 40$ puntos H_1: $\mu > 40$ puntos	H_o: $\mu = 40$ puntos H_1: $\mu < 40$ puntos

Como has visto arriba, las tres posibilidades requieren análisis distintos. Por ejemplo, plantear simplemente que los adolescentes presentan un nivel de ansiedad diferente de 40 cuando interactúan con sus profesores, no advierte si se espera que tengan específicamente más (o menos) ansiedad; simplemente, la hipótesis nos dice que el nivel de ansiedad es distinto respecto de 40. En un planteo así, como la hipótesis no direcciona el análisis, el investigador debe contemplar la posibilidad de que ocurran tres posibilidades, una de ellas (los adolescentes tendrán 40 puntos) no permitirá rechazar la hipótesis nula, y dos que si permitirán el rechazo de la hipótesis nula (que tengan menos de 40 puntos o que tengan más de 40 puntos). En este caso será necesaria una contrastación bilateral.

Por el contrario, si la hipótesis predetermina, por ejemplo, que tendrán más de 40 puntos, nos veremos en la necesidad de hacer un contraste unilateral.

Como vemos entonces, la hipótesis alternativa puede ser unilateral o bilateral.

- Bilateral: cuando necesitamos determinar si el parámetro de población es distinto que el valor hipotético.
- Unilateral: cuando deseamos determinar si el parámetro de población difiere del valor hipotético en una dirección específica (o es mayor o es menor, *una de las dos posibilidades*).

Las pruebas bilaterales tienen menos potencia que una prueba unilateral; aunque si bien las pruebas unilaterales tienen mayor potencia, estas presentan la desventaja de que no pueden detectar si el parámetro de población difiere en la dirección opuesta.

Pero… ¿qué sucede si tenemos dos variables?

Supongamos que nuestro Planteo del problema considera lo siguiente: "*¿Existen diferencias en el nivel de ansiedad que experimentan los adolescentes varones respecto a las adolescentes mujeres cuando interactúan con profesores?"*

	Planteo I	**Planteo II**	**Planteo III**
Planteo de la hipótesis	El nivel de ansiedad que experimentan los adolescentes varones es DIFERENTE de las adolescentes mujeres cuando interactúan con profesores	El nivel de ansiedad que experimentan los adolescentes varones es MAYOR de las adolescentes mujeres cuando interactúan con profesores	El nivel de ansiedad que experimentan los adolescentes varones es MENOR de las adolescentes mujeres cuando interactúan con profesores
Formalización	H_o: $\mu_1 = \mu_2$ H_1: $\mu_1 \neq \mu_2$	H_o: $\mu_1 = \mu_2$ H_1: $\mu_1 > \mu_2$	H_o: $\mu_1 = \mu_2$ H_1: $\mu_1 < \mu_2$

Si recuerdas, en la unidad 5 vimos estimación puntual e intervalar de parámetros; encontrarás algunas similitudes en las distintas pruebas de hipótesis, por lo que te recomendamos tener bien presente aquellos conocimientos.

Acá, en lugar de estimar el valor de un parámetro, debemos tomar una decisión respecto a una afirmación relacionada con una hipótesis y decidir si se acepta o se rechaza. Más adelante veremos cómo llegamos a esa decisión, antes nos detendremos en dos posibles errores que podemos cometer.

Los errores de tipo I y tipo II

En estadística inferencial, si bien muchas veces podemos minimizar la posibilidad de cometer errores, tenemos que recordar que nunca vamos a poder llegar a obtener la certeza absoluta en nuestras decisiones. Por ende, esto también ocurre en la Prueba de hipótesis donde, por basarse en una estrategia en términos de probabilidades, siempre existe la posibilidad de llegar a una conclusión errónea. Podemos llegar a cometer dos tipos de errores que están inversamente relacionados y que se determinan según el nivel de significancia (o de exigencia) de la prueba.

Siendo consciente de esto, el investigador puede determinar qué error podría tener consecuencias más graves en su investigación.

Error de tipo I: (también se lo suele denominar como error de tipo alfa (α) o falso positivo) es el error que comete el investigador cuando rechaza la hipótesis nula siendo en realidad válida. Se dice que es un resultado falso positivo porque el investigador llega a la conclusión de que existe un efecto, una diferencia, en la hipótesis cuando en realidad no existe. Este error aumenta a medida que la exigencia estadística disminuye.

Error de tipo II: (llamado también error de tipo beta (β) o falso negativo) este error es el que comete el investigador cuando no rechaza la hipótesis nula cuando en realidad es falsa. En este caso el investigador concluye que no hay evidencia del efecto o diferencia cuando en realidad existe. Este error aumenta a medida que la exigencia estadística aumenta.

Pensemos en un ejemplo llevado al extremo para que se comprenda esto de una forma más práctica. Supongamos que dos científicos que están llevando a cabo una investigación en conjunto difieren en su punto de vista al momento de decidir acerca del valor alfa o significancia estadística con el que van a confrontar sus hipótesis. El investigador (A) desea trabajar con un nivel de exigencia relativamente bajo, digamos con un Alfa=0,2 (un error del 20%); en cambio el investigador (B) es mucho más exigente, desea hacer la prueba de hipótesis con un Alfa= 0,01 (un error del 1%). Así, dado que el investigador (A) propone un nivel de exigencia estadística más bajo podrá rechazar con más facilidad H_0 (error de tipo I); en cambio el investigador (B) por ser más exigente le será más difícil rechazar H_0 (error de tipo II).

Procedimiento de Prueba de hipótesis

Consiste en una serie de pasos. Algunos de ellos lo hemos abordado en los cálculos de estimaciones, en donde obteníamos un intervalo (Unidad 5), y los restantes los veremos en este apartado. A continuación, te ofrecemos el esquema con una breve descripción para que puedas orientarte.

- *Primer paso: Reconocer y definir la o las variables.*

  El investigador se concentra en las variables de interés de modo que pueda ser operativo su trabajo. Ya habíamos señalado al respecto la importancia que tiene la concordancia entre el planteamiento del problema, los objetivos e hipótesis al momento de determinar las variables en estudio.

- *Segundo paso: Formular las hipótesis nula y alternativa.*

  Este punto concuerda con lo desarrollado más arriba; acá el investigador se esfuerza por tratar de sistematizar las hipótesis de modo que esté en condiciones de emitir, al final del proceso, solo dos posibles decisiones: **rechazar o aceptar la Hipótesis Nula**

- *Tercer paso: Establecer el estadístico de prueba adecuado.*

  Los estadísticos de prueba se utilizan para determinar si puede rechazar la hipótesis nula, ya que compara los datos que ha obtenido el investigador con lo que se espera bajo la hipótesis nula. Hay diversos estadísticos de prueba según se plantee en la hipótesis diferencias de media de una muestra, entre dos muestras independientes, entre dos muestras relacionadas, correlaciones, etc. A continuación, te ofrecemos un recuadro con algunos ejemplos de estadísticos de prueba que corresponden a determinadas pruebas de hipótesis.

Prueba de hipótesis	Estadístico de prueba
Media poblacional (muestras grandes)	Estadístico z
Diferencia de media (muestras pequeñas)	Estadístico t
ANOVA	Estadístico F
Correlaciones entre variables	Coeficiente de Pearson Rho de Spearman

  Supongamos que el investigador ha decidido comprobar si existen diferencias de media entre hombres y mujeres respecto de una variable, supongamos "Asertividad". Él ha corroborado que la distribución de la variable en ambos grupos cumple con dos condiciones: a) presentan una distribución normal, y b) ambos grupos presentan varianzas iguales. En este caso el estadístico que necesita para establecer tal comparación es la Prueba t

$$t = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{(n-1)\hat{S}_1^2 + (m-1)\hat{S}_2^2}{n+m-2}}\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

- *Cuarto paso:* Seleccionar un nivel de significación

Un resultado es estadísticamente significativo cuando es improbable que se deba al azar, por lo cual el nivel de significación se encuentra asociado a la verificación de una hipótesis mediante conceptos de probabilidad. Si tenemos que pensar en una definición diríamos que el nivel de significación representa la probabilidad de tomar la decisión de rechazar la hipótesis nula cuando ésta sea verdadera (error de tipo I). Este umbral, que determina si los resultados son estadísticamente significativos, lo suele proponer el investigador dependiendo del tema que esté investigando. En ciencias sociales un nivel de significación al 5% o 0,05 es muy aceptable, aunque también se suele trabajar al 10% o 0,10 y al 1%, es decir al 0,01.

¿Qué representa 0,10; 0,05 o 0,01?, bueno, como habíamos señalado nos dice la probabilidad de cometer el error al rechazar la hipótesis nula cuando es verdadera. Tomemos el caso de un nivel de significación de 0,05; ese valor nos diría que el investigador asume una probabilidad del 5% de equivocarse. Evidentemente, mientras menor sea este valor más rigurosas son las condiciones para tomar la decisión y por lo tanto mayor precisión y robustez tiene el estudio. Por ejemplo, con un nivel de significación de 0,05 el investigador asume un 5% de riesgo de equivocarse al decir que encontró una diferencia entre sus resultados y la hipótesis nula.

El nivel de significación normalmente se especifica de antemano y se lo incluye en la sección de Procedimientos estadísticos.

- *Quinto paso:* Establecer la regla de decisión

  En este paso se propone la regla de decisión sobre el estadístico que corresponde. Si retomamos el ejemplo que propusimos en el paso 3, diremos que, por tratarse de una hipótesis basada en diferencia entre dos grupos, se pueden presentar estas posibilidades:

	Contraste bilateral	**Contraste unilateral izquierdo**	**Contraste unilateral derecho**
Hipótesis	$\begin{cases} H_0 : \mu_x = \mu_y \\ H_1 : \mu_x \neq \mu_y \end{cases}$	$\begin{cases} H_0 : \mu_x = \mu_y \\ H_1 : \mu_x < \mu_y \end{cases}$	$\begin{cases} H_0 : \mu_x = \mu_y \\ H_1 : \mu_x > \mu_y \end{cases}$

  *Nota Explicamos las nomenclaturas de la formulación bilateral. Para H_0 dice que el valor de media del grupo (x) es igual a la media del grupo (y); en cambio, para H_1, dice que el valor de media del grupo (x) es distinto del valor de media del grupo (y).

Como puedes ver, la formulación de *la regla de decisión no deja otras posibilidades que la de aceptación o rechazo de la hipótesis nula*, ya que no se admiten formulaciones ambiguas.

- Sexto paso: Calcular el valor observado del estadístico de prueba

  A la probabilidad de que un resultado se deba al azar, y no producto de lo que está estudiando el investigador, se le llama «valor p». Este valor surge del cálculo de los estadísticos de contraste.

  El valor p (valor de probabilidad), es una medida estadística que ayuda a los científicos a determinar si sus hipótesis son correctas o no, y está estrechamente vinculado con el nivel de significación. Un resultado es estadísticamente significativo si tiene un valor p igual o menor que el nivel de significación (generalmente $p \leq 0,05$). Así, si tenemos un valor p inferior al nivel de significación escogido oportunamente, estamos en condiciones de rechazar la hipótesis nula.

- *Séptimo paso:* Obtener la conclusión

La conclusión no es un mero enunciado acerca del rechazo o no de la hipótesis nula; por el contrario, debe brindar suficiente información al menos de los siguientes aspectos:

a) El nivel de significación con el que se realiza la conclusión.
b) Cuál es la variable sobre la que se ha realizado la prueba de hipótesis.
c) Cuál es el parámetro que nos interesa.
d) Si se rechaza o no se rechaza la hipótesis nula.
e) Las implicancias de la decisión tomada.

Por ejemplo, si tomamos la hipótesis planteada en el paso tres podríamos escribir una conclusión similar a ésta: *Con un nivel de significación de 0,05 se rechaza la hipótesis nula que afirma la inexistencia de diferencias de media entre hombres y mujeres en asertividad, a favor de la hipótesis alternativa, que afirma la existencia de diferencias entre ambos grupos. Consultando los valores de media correspondientes, se verifica que las mujeres tienden a ser más asertivas que los hombres. Estos hallazgos sugieren la posibilidad de desarrollar talleres psicoeducativos para los hombres con la finalidad de potenciar sus capacidades.*

Análisis paramétricos y no paramétricos: supuestos asociados

Muchos de los conceptos que han sido vertidos en este capítulo, y anteriores, son necesarios para la correcta implementación e interpretación de los estadísticos de prueba. Conceptos como muestreo, prueba de hipótesis, nivel de significación, distribución, probabilidad, tipo de variables, etc., van a ser implementados en los próximos capítulos, con una perspectiva más práctica. Mientras tanto, debemos considerar que algunos estadísticos de prueba para diferencias entre grupos o para la asociación entre variables requieren ciertas consideraciones previas a su implementación para que su uso sea adecuado.

Como hemos visto más arriba, una de las elecciones que debe hacer el investigador es escoger el estadístico de prueba que resulte adecuado a los propósitos del estudio. En este punto hay una serie de consideraciones que se debe plantear, por ejemplo:

1. *¿Qué estrategia va a dar respuesta al planteo de hipótesis?:* Esta interrogante nos orienta de modo global para ver qué tipo de estrategia estadística podemos aplicar. Por ejemplo, en una hipótesis que plantee que las personas con niveles mayores de insigh (darse cuenta) tienden a requerir intervenciones de menor tiempo, nos enfoca a buscar estrategias correlacionales ya que lo que se espera determinar es si efectivamente existe relaciones entre las variables en cuestión. En cambio, una hipótesis que exprese que las personas a las que se le administre la intervención (X) presentarán una disminución en los niveles de ansiedad, nos invita a usar pruebas que comparen niveles entre dos valores: nivel de ansiedad antes del tratamiento y niveles de ansiedad después del tratamiento.
2. ¿Cuál es el nivel de medición de las variables que se van a someter a evaluación?
3. ¿Cuál es la distribución de los datos que presentan las variables sometidas a prueba de hipótesis?
4. ¿Cuál es el tamaño muestral, o la cantidad de datos que disponemos?... etc.…

Las últimas tres preguntas están orientadas a saber si las pruebas enmarcadas dentro de estadística paramétrica, o si en realidad las que forman parte de la estadística no paramétrica, son las más convenientes para el análisis a implementar.

Debes tener presente que cada prueba estadística está diseñada para un propósito específico. Buscando un ejemplo práctico de la estadística descriptiva para dar cuenta de este punto, te darás cuenta que el cálculo de media no se lo puede aplicar a variables con nivel de medición nominal, ¿cierto?... Bueno, suponiendo que tengas alguna duda con este comentario pensemos en la variable género (variable categórica) … ¿Cómo podría calcular el promedio entre femenino, masculino, no binario, etc.? ¿Qué sentido tendría un valor, aunque se hayan codificado las variables con un numero?, ¿Qué significarían una media de 1,23?... Como ves, o no se podría calcular por ausencia de números, o su implementación sería un error conceptual por desconocer la naturaleza de la variable.

Bueno, del mismo modo ocurre con la estadística inferencial. Si erramos en la elección de los estadísticos de prueba llegaremos a conclusiones engañosas y sin sentido.

Supuestos básicos

Cada prueba estadística tiene condiciones específicas que se deben satisfacer para ser utilizadas. En términos generales, las pruebas paramétricas descansan sobre los supuestos de normalidad, de que las variables poseen un nivel de medición al menos intervalar y de que existe una abundante cantidad de datos. Los estadísticos orientados a evaluar diferencias entre grupos se concentran en los valores de media, y en los de correlación sobre la linealidad de la asociación. En cambio, las pruebas no paramétricas no presentan estas condiciones, se pueden aplicar en variables ordinales, y por lo tanto son la elección cuando tenemos dificultad para emplear pruebas paramétricas.

Pero, ¿por qué existen las pruebas paramétricas, si hay otras que no presentan tantas limitaciones?, la respuesta es porque tienen una mayor potencia estadística, mientras las no paramétricas son más robustas. Explicaremos esto; decimos que las no paramétricas son más robustas porque hacen un tanto más difícil rechazar H_0, por eso son más confiables cuando no se cumplen los supuestos necesarios de las paramétricas. En cambio, cuando se cumplen las condiciones y se pueden aplicar las pruebas paramétricas, estas brindan información mucho más detallada y muy valiosa, en comparación a su opuesta.

Si bien, en los capítulos siguientes veremos ambos tipos de estrategia y hablaremos acerca de los supuestos asociados, a continuación, te ofrecemos un cuadro resumen

Pruebas paramétricas	Pruebas no paramétricas
Menor exigencia para rechazar H_0	Mayor exigencia para rechazar H_0
Requiere que las variables sean normales y con un nivel de medición al menos de intervalo	No requiere que las variables sean normales y se puede aplicar en variables con un nivel de medición al menos ordinal.
Se implementa cuando las muestras son grandes	Se pueden utilizar con muestras pequeñas.
La distribución de los datos tiene que ser normal, no admite aplicarlas con distribuciones asimétricas y que no sean mesocúrticas	No exige normalidad en la forma de la distribución de los datos.
Su mayor potencia estadística permite hacer muchas suposiciones	Por tener menor potencia estadística no se pueden hacer muchas suposiciones.
Las hipótesis se basan en los valores de media, desviación estándar, en la linealidad de asociaciones	Las hipótesis se basan en rangos, mediana y frecuencia de datos.

Introducción a las pruebas de hipótesis que evalúan valores medios

Muy bien, hasta acá hemos aprendido que para comprobar las hipótesis debemos hacer un contraste entre la hipótesis alternativa y una contrapuesta, denominada hipótesis nula. Adicionalmente, para llevar a cabo ese procedimiento, además de la formulación debemos escoger un estadístico de prueba, es decir un cálculo para evaluar el grado de certeza vinculado con la hipótesis propuesta; y que esa prueba estadística tiene que ser cuidadosamente elegida dentro de las estrategias paramétricas o no paramétricas dependiendo de una serie de aspectos puntuales.

A continuación, ofreceremos en todos los análisis siguientes, ambas opciones para cada situación que se nos plantee. En este punto será importante que prestes mucha atención al menos a dos aspectos iniciales: a) el propósito del análisis y b) los supuestos básicos de las alternativas paramétricas y no paramétricas ofrecidas.

Pruebas para los valores medios poblacionales

Propósito del análisis

Suponte que el investigador haya propuesto una hipótesis similar a esta:

- Los participantes de la muestra presentan valores medios de ansiedad diferentes a los de la población

En otras palabras, el científico pretende comprobar que la muestra que investiga tiene niveles de ansiedad distintos (sin aclarar que pueden ser más bajos o más altos) respecto de los valores que se conocen de la población. Por otro lado, notarás que nos referimos a valores medios haciendo alusión implícitamente a los valores de Media o a los valores de Mediana. Hemos decidido plantear el ejemplo de esta forma ya que, si corresponde un análisis de tipo paramétrico la referencia será la Media, mientras que si tenemos que optar por un análisis no paramétrico, tendemos que tomar como referencia la Mediana.

Partiendo de una muestra de 281 participantes el investigador ha descubierto que presentan los siguientes valores:

$\bar{x}$ (media) = 29,4

Me (mediana) = 29

Además, sabe que la población de la cual ha sido extraída la muestra presenta:

$\bar{x}$ (media) = 27,1

Me (mediana) = 27

Lo que hará a continuación el investigador es decidir cuál es la estrategia que corresponde hacer el análisis, y lo hará de acuerdo a los supuestos asociados según veremos a continuación

- ***Estrategia paramétrica***

  El estadístico paramétrico para este tipo de análisis es la <u>Prueba t para una muestra</u>

  Supuestos asociados: esta prueba parte del supuesto de que la distribución de los datos es aproximadamente normal, que el tamaño de la muestra es grande (mayor a 30) y que la variable tiene un nivel de medición al menos intervalar. Si se cumplen estos supuestos el investigador elegirá esta prueba, de lo contrario buscará su alternativa no paramétrica. Además, es importante que recuerdes que el contraste se hará teniendo en cuenta la Media, por lo que ahora verás modificados algunos términos del planteo de hipótesis.

Pasos

Formular las hipótesis nula y alternativa

- H_0= La muestra NO presenta valores promedios de ansiedad diferentes a los de la población
- H_1= La muestra presenta valores promedios de ansiedad diferentes a los de la población

Establecer el estadístico de prueba adecuado: <u>Prueba t para una muestra</u>

Seleccionar un nivel de significación: <u>Alfa de 0,05</u>

Establecer la regla de decisión

$$\begin{cases} H_0 : & \bar{x} = 27{,}1 \\ H_1 : & \bar{x} \neq 27{,}1 \end{cases}$$

Calcular el valor observado del estadístico de prueba (Procedimiento mediante Jamovi)

Inicialmente nos dirigirnos a Análisis > Pruebas T > Pruebas T en una muestra

Figura 1: Captura de pantalla del procedimiento Prueba T para una muestra realizado mediante el programa estadístico Jamovi

Luego configuramos el análisis como se mostrará en la Figura 2

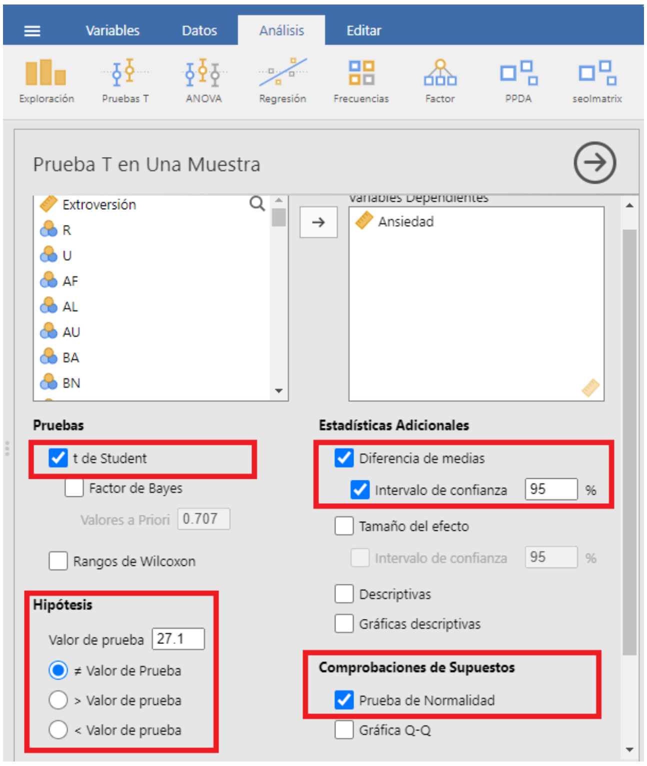

Figura 2: Captura de pantalla de opciones para el análisis de Prueba T para una muestra realizado mediante Jamovi

Resultados

Tabla 1.

Análisis de normalidad

Tests of Normality

		statistic	p
Ansiedad	Shapiro-Wilk	0.987	0.013
	Kolmogorov-Smirnov	0.0748	0.058
	Anderson-Darling	0.968	0.015

Nota. Additional results provided by *moretests*

Como se observa en la figura 2, Jamovi brinda la posibilidad de comprobar algunos de los supuestos asociados a las pruebas paramétricas. En este caso nos da la posibilidad de corroborar si los datos provienen de una distribución normal y, como ya se ha dicho dependerá de estos datos si se va a proseguir con la estrategia paramétrica o no paramétrica.

La tabla 1 nos ofrece tres estadísticos similares, pero que suelen usarse con propósitos específicos. Los tres ponen a prueba las siguientes hipótesis:

- H_0: Los datos proceden de una distribución normal
- H_1: Los datos no proceden de una distribución normal

Por tal motivo, si observas valores iguales o menores al Alfa propuesto por el investigador (generalmente 0,05) rechazaremos H_0 y afirmaremos que no presentan una distribución normal. De lo contrario, aceptaremos H_0 y diremos que los datos tienen una distribución normal.

¿A cuál de los tres estadísticos ofrecidos por Jamovi debemos prestarle atención para determinar la normalidad en la distribución de datos? Como decíamos con anterioridad, los tres son similares, y dos de ellos son los más conocidos. Se recomienda Kolmogorov-Smirnov cuando tenemos más de 50 datos u observaciones; en cambio, Shapiro Wilk se lo suele utilizar cuando tenemos menos de 50 observaciones.

Interpretación de la prueba de normalidad

Atentos a que estamos frente a una muestra de 281 participantes, el valor p (0,058) asociado a la prueba de Kolmogorov-Smirnov (Tabla 1), sugiere que los datos provienen de una distribución normal. Por lo que es aceptable continuar con el análisis paramétrico (de lo contrario deberíamos optar por la estrategia no paramétrica).

Tabla 2.

Análisis de diferencias de media para una muestra mediante Prueba t de Student

Prueba T en Una Muestra

						Intervalo de Confianza al 95%	
		Estadístico	gl	p	Diferencia de medias	Inferior	Superior
Ansiedad	T de Student	6.24	280	< .001	2.35	1.61	3.09

Nota. H_a $\mu \neq 27.1$

En la tabla2 verás tres rectángulos, uno con el valor p (valor de probabilidad asociado al estadístico) y los otros donde se observan el valor de la prueba T y el intervalo de confianza calculado al 95%.

Interpretación para las conclusiones

A continuación, te mostraremos una manera de interpretar los resultados. La redacción suele variar, no hay una única forma de explicar los análisis, aunque en todos los casos debe ser completa y detallada; así el lector poseerá toda la información necesaria para comprender la decisión de rechazar o aceptar H_0.

Interpretación de los resultados

El investigador buscará comparar el valor p obtenido del cálculo, con el nivel de significación (con el Alfa) fijado con anterioridad. Si el valor p resulta igual o menor al Alfa, estará en condiciones de RECHAZAR H_0. Si el valor p resulta mayor al Alfa, se verá obligado a ACEPTAR H_0 ya que el error obtenido sería mayor al que estaba dispuesto a asumir.

Veamos los valores. El cálculo muestra un valor p de <0,001 y el Alfa era de 0,05, por lo que vemos que el valor p es claramente más pequeño que el nivel de significación asumido. Con estos valores podemos decir que se rechaza H_0 y que se confirma la hipótesis de que la muestra presenta valores promedios de ansiedad diferentes a los de la población.

Si la curiosidad del investigador lo lleva a preguntarse si la muestra presentó valores más altos o más bajos, deberá comparar la media para dicha variable con la media de la población.

Sobre el intervalo de confianza

Respecto del intervalo de confianza diremos que, con un 95 % de certeza, la diferencia entre la media de la muestra y la de la población está comprendida entre los 1,61 y 3,09 puntos.

- **Estrategia no paramétrica**

El estadístico no paramétrico para este tipo de análisis es la Prueba Rangos de Wilcoxon para una muestra.

Supuestos asociados: esta prueba no toma en consideración el tipo de distribución de los datos para hacer el contraste de hipótesis, por lo que se puede aplicar aún si la variable presenta una distribución no normal. No exige un tamaño muestral grande (se puede realizar con una cantidad menor de 30 datos) y la variable sometida a análisis puede tener un nivel de medición ordinal. Por tal motivo, si el investigador encuentra que no puede aplicar un análisis paramétrico (porque sus datos no cumplen alguno de los supuestos) puede elegir esta estrategia. Es importante que recuerdes que el contraste en este caso se hará teniendo en cuenta la Mediana (NO LA MEDIA), por lo que ahora verás modificados algunos términos del planteo de hipótesis.

Pasos

Formular las hipótesis nula y alternativa

- H_0= La muestra NO presenta valores de Mediana diferentes a los de la población en la variable ansiedad
- H_1= La muestra presenta valores de Mediana diferentes a los de la población en la variable ansiedad

Establecer el estadístico de prueba adecuado: Prueba Rangos de Wilcoxon para una muestra

Seleccionar un nivel de significación: Alfa de 0,05

Establecer la regla de decisión

$$\begin{cases} H_0 : & \text{Me} = 27 \\ H_1 : & \text{Me} \neq 27 \end{cases}$$

*Nota: atención en este caso el contraste se hace sobre la Mediana y no sobre la Media

Calcular el valor observado del estadístico de prueba (Procedimiento mediante Jamovi)

Inicialmente nos dirigirnos a Análisis > Pruebas T > Pruebas T en una muestra (hasta acá el procedimiento es similar al análisis paramétrico)

Figura 3: Captura de pantalla del procedimiento Prueba T para una muestra realizado mediante el programa estadístico Jamovi

Luego configuramos el análisis como se mostrará en la Figura 4

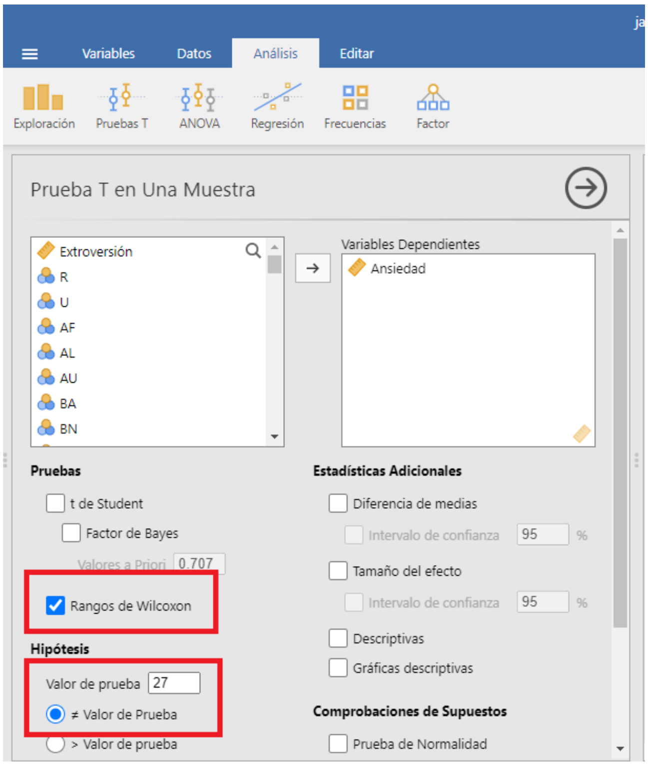

Figura 4: Captura de pantalla opciones para el análisis de Prueba T para una muestra realizado mediante Jamovi

Resultados

Tabla 3.

Análisis de diferencias de mediana para una muestra con la prueba de Rangos de Wilcoxon

		Estadístico	p
Ansiedad	W de Wilcoxon	24991	< .001

Nota. H_a $\mu \neq 27$

En la tabla 3 verás un rectángulo en el que se visualiza el valor p (valor de probabilidad asociado al estadístico), dato necesario para determinar el rechazo o la aceptación de H_0.

Interpretación de los resultados

El investigador deberá comparar el valor p obtenido del cálculo, con el nivel de significación (o valor Alfa) establecido con anterioridad. Como recuerdas, si el valor p resulta igual o menor al Alfa, estará en condiciones de RECHAZAR H_0. Si el valor p resulta mayor al Alfa, se verá obligado a ACEPTAR H_0 ya que el error obtenido sería mayor al que estaba dispuesto a asumir.

Veamos, el cálculo realizado proporcionó un valor p de <0,001, y el Alfa era de 0,05, por lo que vemos que el valor p es claramente más bajo que el nivel de significación asumido. Con estos valores podemos concluir que se rechaza H_0 y que se confirma la hipótesis de que la Mediana de la muestra es diferente a la de la población en la variable Ansiedad.

Si el investigador se pregunta cuál de las medianas es mayor, deberá solicitarle al programa estadístico un análisis descriptivo y luego comparar ambas medianas (la de la muestra vs. la de la población).

Pruebas para la diferencia de valores medios entre dos muestras independientes

Propósito del análisis

En algunas oportunidades nos puede interesar averiguar si existen diferencias en los valores medios (media o mediana) entre dos muestras que sean independientes. Una muestra es independiente de otra cuando los participantes del estudio sólo pueden incluirse en una de ellas. Algunos ejemplos de estas situaciones se dan cuando indagamos diferencias en alguna variable de interés entre: estudiantes de escuelas privadas y públicas; entre mujeres y hombres; entre habitantes de zonas rurales y urbanas, etc. Como puedes ver, en estos casos no existe posibilidad de que algún participante pueda estar en dos grupos simultáneamente; es decir, es un estudiante de una escuela pública o es un estudiante de una escuela privada; es mujer o es hombre; es alguien que vive en una zona rural, o vive en una zona urbana.

En este tipo de análisis intervienen dos variables: a) la variable objeto de medición (Ansiedad, Autoestima, Inteligencia, Personalidad, etc.) y b) la variable de agrupación (Sexo, Situación laboral, etc). La variable de agrupación es aquella que ordena a los participantes en dos grupos o muestras.

Veamos un ejemplo:

Un investigador se interesa por indagar si los niveles de Empatía (capacidad para ponerse en el lugar de otra persona) son distintos entre habitantes de zonas rurales y de zonas urbanas. Así tendrá en cuenta dos variables: a) Empatía y b) Residencia. La variable Residencia será considerada como variable de agrupación, ya que es la que va a determinar dos grupos de análisis: habitantes de zonas rurales y habitantes de zonas urbanas. Así su estudio se dividirá en dos muestras.

Como es lógico, primero analizará los estadísticos descriptivos que surgen para este análisis

Tabla 4.

Estadísticos descriptivos para la variable Empatía según lugar de Residencia

Descriptivas de Grupo

	Grupo	N	Media	Mediana	DE	EE
Empatía	Urbana	217	25.9	26.0	5.80	0.394
	Rural	56	26.5	28.0	6.14	0.821

En la tabla 4 se observa que con una muestra de 273 participantes el investigador ha encontrado que la media de Empatía para los habitantes de zonas urbanas es de 25,9, mientras que para los de zonas rurales es de 26,5. De manera similar ha encontrado que la mediana de Empatía para los habitantes de zonas urbanas es de 26, en tanto que para los de zona rural es de 28.

Estos datos a simple vista sugieren que las personas que viven en zonas rurales presentarían mayores niveles de Empatía, es decir una mayor capacidad para ponerse en el lugar de las otras personas.

Sin embargo, el investigador, para afirmar rotundamente este hallazgo, deberá someter a prueba su hipótesis, tal como lo veremos a continuación.

- **Estrategia paramétrica**

El estadístico paramétrico para este tipo de análisis es la Prueba t para muestras independientes

Supuestos asociados:

- La distribución de los datos es aproximadamente normal
- El tamaño de las muestras debe ser mayor a 30

- La variable a ser comparada tiene un nivel de medición al menos intervalar.
- La varianza es constante entre los grupos (homogeneidad de varianzas u homocedasticidad).

Si se cumplen estos supuestos el investigador elegirá esta prueba, de lo contrario, buscará su alternativa no paramétrica.

*Nota: Recuerda, como en todos los análisis paramétricos, el contraste de hipótesis se hará teniendo en cuenta la Media.

Pasos

Formular las hipótesis nula y alternativa

•H_0= La muestra de zona rural NO presenta valores de media significativamente diferentes de la muestra urbana en la variable Empatía.

•H_1=La muestra de zona rural presenta valores de media significativamente diferentes de la muestra urbana en la variable Empatía.

Establecer el estadístico de prueba adecuado: Prueba t para muestras independientes

Seleccionar un nivel de significación: Alfa de 0,05

Establecer la regla de decisión

$$\begin{cases} H_0 & : \bar{x}(\text{Rural}) = \bar{x}(\text{Urbana}) \\ H_1 & : \bar{x}(\text{Rural}) \neq \bar{x}(\text{Urbana}) \end{cases}$$

Calcular el valor observado del estadístico de prueba (Procedimiento mediante Jamovi)

Inicialmente nos dirigirnos a Análisis > Pruebas T > Pruebas T en una muestra

Figura 5: Captura de pantalla del procedimiento Prueba T para muestras independientes realizado mediante el programa estadístico Jamovi

El procedimiento se completa marcando las opciones que se pueden apreciar en la figura 6

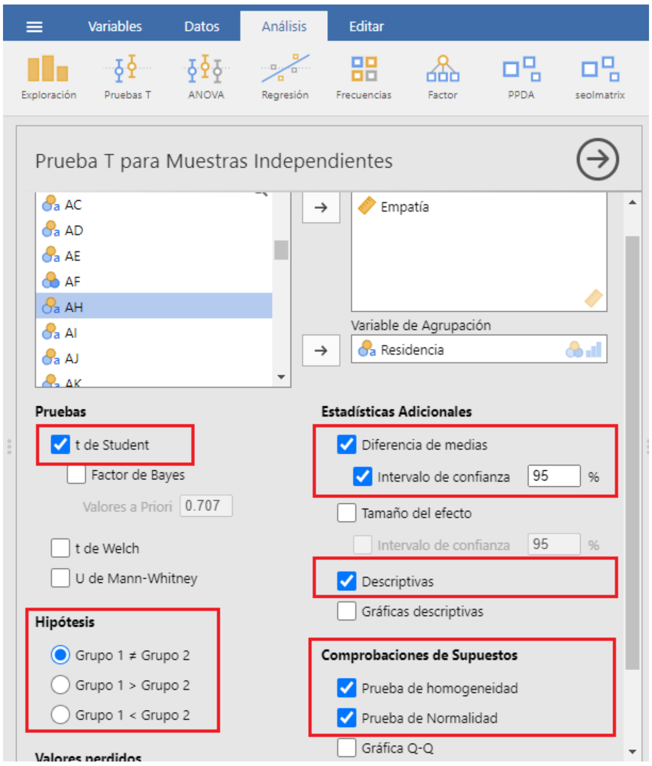

Figura 6: Captura de pantalla opciones para el análisis de Prueba T para muestras independientes realizado mediante Jamovi

Tabla 5.

Análisis de Normalidad

Tests of Normality

		statistic	p
Empatía	Shapiro-Wilk	0.961	0.001
	Kolmogorov-Smirnov	0.109	0.065
	Anderson-Darling	2.55	0.021

Nota. Additional results provided by *moretests*

Como hemos visto, consideramos Kolmogorov-Smirnov cuando hay más de 50 datos u observaciones; en cambio, Shapiro Wilk cuando existen menos de 50 observaciones. Además, dijimos que si los valores p de las pruebas son iguales o menores al Alfa propuesto por el investigador (generalmente 0,05) tendremos que rechazar H_0 y afirmar que no presentan una distribución normal.

Interpretación de la prueba de normalidad

Atentos a que estamos frente a una muestra mucho mayor a 50 participantes, el valor p (0,065) asociado a la prueba de Kolmogorov-Smirnov (Tabla 5), sugiere que los datos provienen de una distribución normal. Por este motivo es aceptable continuar con el análisis paramétrico (recuerda que de lo contrario deberíamos optar por la estrategia no paramétrica).

Tabla 6.

Análisis de homogeneidad

Homogeneity of Variances Tests

		F	df	df2	p
Empatía	Levene's	0.274	1	271	0.003
	Variance ratio	0.893	216	55	0.005

Nota. Additional results provided by *moretests*

Antes de avanzar, explicaremos el concepto de homogeneidad de varianzas.

El supuesto de homogeneidad de varianzas, también conocido como supuesto de homocedasticidad, considera que la varianza es constante, que no varía entre las dos (o más) muestras. Es uno de los supuestos más importantes asociados a este tipo de prueba, ya que su incumplimiento puede llevar a interpretaciones erróneas y falsas. Uno de los estadísticos más utilizados para evaluar este supuesto es el de Levene, que pone a prueba la hipótesis nula de que las varianzas son iguales. Es decir que si el valor p asociado a la prueba es igual o menor al nivel de significación o Alfa (generalmente 0,05), es poco probable que las muestras posean varianzas iguales, por lo que se rechaza la hipótesis nula de igualdad de varianzas y se concluye que las muestras no son homogéneas u homocedásticas.

Interpretación de la prueba de Homogeneidad de varianzas

En la tabla 6 observamos que el valor p (0,003) asociado a la prueba de Levene, sugiere que los datos de las muestras no son homogéneos. Por lo que no es aceptable continuar con el análisis paramétrico y deberíamos optar por la estrategia no paramétrica.

*IMPORTANTE: Como has notado, el estadístico de Levene nos ha informado que uno de los supuestos (homogeneidad entre los datos de las muestras) no se cumple; y como hemos advertido, con este solo argumento ya no podríamos aplicar la prueba T para muestras independientes. Sin embargo, a continuación ofreceremos dicho análisis con la salvedad de que carece de validez, y que sólo se lo presenta a modo de ejemplo para explicar el procedimiento

Tabla 7.

Prueba T de diferencia de media para muestras independientes

Prueba T para Muestras Independientes

							Intervalo de Confianza al 95%	
		Estadístico	gl	p	Diferencia de medias	EE de la diferencia	Inferior	Superior
Empatía	T de Student	−0.771	271	0.442	−0.679	0.880	−2.41	1.05

Interpretación de los resultados

El investigador procurará comparar el valor p asociado a la Prueba T de Student, con el nivel de significación (con el Alfa) fijado con anterioridad. Si el valor p resulta igual o menor al Alfa, estará en condiciones de RECHAZAR H_0. Si el valor p resulta mayor al Alfa tendrá que ACEPTAR H_0.

En la tabla 7 se observa un valor p de 0,442, considerando que Alfa propuesto por el investigador era de 0,05, vemos que el valor p es claramente mayor que el nivel de significación asumido. Con estos valores podemos decir que se acepta H_0 y que se confirma que no existen diferencias de media significativas entre la muestra urbana y rural en la variable Empatía.

En otras palabras, las diferencias entre los valores de media de las muestras observadas en el análisis descriptivo, son sólo aparentes y están ligadas a un margen de error tan alto que el investigador decide despreciarlas.

Sobre el intervalo de confianza

En este caso, el análisis del intervalo de confianza carece de sentido realizarlo ya que se ha rechazado la hipótesis de diferencias de media significativas entre la muestra urbana y rural para la variable Empatía.

*Nota: Como puedes recordar, la tabla 4 nos ofrecía valores de media de 25,9 para la muestra de participantes de la zona Urbana y de 26,5 para las personas que viven en zona Rural. Esa diferencia aparente fue desmentida por medio de esta prueba de hipótesis. Por lo que si hubiésemos afirmado en aquel momento que los habitantes de zonas Rurales eran más Empáticos, habríamos ofrecido una conclusión engañosa.

• Estrategia no paramétrica

El estadístico no paramétrico para este tipo de análisis es la Prueba U de Mann-Whitney para muestras independientes

Supuestos asociados: esta prueba no toma en consideración el tipo de distribución de los datos para hacer el contraste de hipótesis, por lo que se puede aplicar aún si la variable presenta una distribución no normal o si las muestras no son homogéneas. No exige un tamaño muestral grande (se puede realizar con una cantidad menor de 30 datos) y la variable sometida a análisis puede tener un nivel de medición ordinal. Por tal motivo, si el investigador encuentra que no puede aplicar un análisis paramétrico (porque sus datos no cumplen alguno de los supuestos) puede elegir esta estrategia. Recuerda, el contraste en este caso se hará teniendo en cuenta la Mediana (NO LA MEDIA).

Pasos

Formular las hipótesis nula y alternativa

- H_0= La muestra de zona rural NO presenta valores de mediana significativamente diferentes de la muestra urbana en la variable Empatía.
- H_1= La muestra de zona rural presenta valores de mediana significativamente diferentes de la muestra urbana en la variable Empatía.

Establecer el estadístico de prueba adecuado: Prueba U de Mann-Whitney para muestras independientes

Seleccionar un nivel de significación: Alfa de 0,05

Establecer la regla de decisión

$$\begin{cases} \mathrm{H}_0 : \mathrm{Me}(\mathrm{Rural}) = \mathrm{Me}(\mathrm{Urbana}) \\ \mathrm{H}_1 : \mathrm{Me}(\mathrm{Rural}) \neq \mathrm{Me}(\mathrm{Urbana}) \end{cases}$$

*Nota: atención, en este caso el contraste se hace sobre la Mediana y no sobre la Media

Calcular el valor observado del estadístico de prueba (Procedimiento mediante Jamovi)

Inicialmente nos dirigirnos a Análisis > Pruebas T > Pruebas T en una muestra (hasta acá el procedimiento es similar al análisis paramétrico)

Figura 7: Captura de pantalla del procedimiento Prueba U de Mann-Whitney para muestras independientes, realizado mediante el programa estadístico Jamovi

Luego configuramos el análisis como se mostrará en la Figura 8

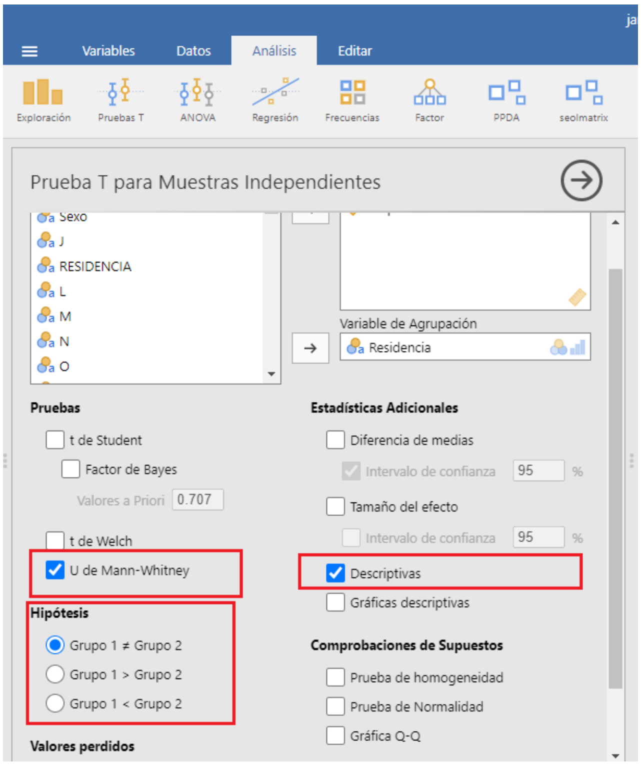

Figura 8: Captura de pantalla de las opciones para el análisis de Prueba U de Mann-Whitney para muestras independientes realizado mediante Jamovi

Resultados

Tabla 8.

Análisis de diferencias de mediana con la Prueba U de Mann-Whitney para muestras independientes

Prueba T para Muestras Independientes

		Estadístico	p
Empatía	U de Mann-Whitney	5420	0.213

La tabla 8 muestra el valor p (valor de probabilidad asociado al estadístico), dato necesario para determinar el rechazo o la aceptación de H_0.

Interpretación de los resultados

Atento a los valores que se observan en la tabla 8, deberemos comparar el valor p, con el nivel de significación (o valor Alfa) establecido con anterioridad. Recordemos; si el valor p resulta igual o menor al Alfa, deberemos RECHAZAR H_0. Si el valor p resulta mayor al Alfa, tendremos que ACEPTAR H_0.

Veamos, el cálculo realizado proporciona un valor p de 0,213 y considerando que el Alfa era de 0,05, vemos que el valor p es claramente mayor que el nivel de significación asumido. Con estos valores debemos entonces aceptar H_0 y decir que las Medianas ambas muestras (Rural y Urbana) no presentan diferencias significativas a nivel estadístico en la variable Empatía, a pesar que en la tabla 4 se declaraba que la muestra de participantes de la zona Urbana presentaba una mediana de 26 y la Rural de 28.

Pruebas para la diferencia de valores medios entre tres o más muestras independientes (ANOVA de un factor)

Propósito del análisis

Algunas veces se nos puede presentar la necesidad de hacer una comparación de los valores medios entre tres grupos independientes (muestras) o más. Son muchas las ocasiones en las que se suele presentar este interés, por ejemplo, cuando nos interesa corroborar si una variable presenta niveles medios distintos según: Nivel de escolaridad (Primario, Secundario y Universitario), Nivel socioeconómico (Bajo, Medio y Alto), Tipo de profesión (Médico, Abogado, Ingeniero, Psicomotricista), Tipo de escuela (Privada, Pública y Autogestionada), Puesto laboral (Administrativo, Operario, Mantenimiento, Gerencial), etc.

Para hacer este tipo de análisis, la estrategia adecuada se denomina ANOVA de un factor. El acrónimo ANOVA significa Análisis de la Varianza; dentro de este método de análisis existe una variedad amplia que se adecua a diversos escenarios dependiendo de las características de los datos y del propósito del investigador, en esta oportunidad solo nos dedicaremos a hablar del Análisis de varianza de un factor.

Para el ANOVA de un factor se requieren dos variables: a) una denominada FACTOR (es una variable de tipo cualitativa que sirve para determinar los grupos o muestras, por ejemplo: Puesto laboral) y, b) una denominada DEPENDIENTE (En este caso la variable es cuantitativa y es sobre la cual analizaremos los valores medios). El ANOVA de un factor nos indicará si al menos entre dos de esos grupos existe una diferencia de valores medios.

*Nota: como puedes observar, tiene cierta familiaridad con las pruebas para analizar la diferencia de valores medios entre dos muestras independientes, sólo que en este caso pueden ser tres o más.

Veamos un ejemplo

Imaginemos que deseamos saber si hay diferencias de acuerdo al nivel educativo alcanzado (Primario, Secundario y Universitario) en la variable Resiliencia (capacidad para transformar las situaciones adversas en crecimiento personal).

A continuación, observaremos el análisis descriptivo y luego, partiendo de ahí implementaremos una estrategia paramétrica y otra no paramétrica

Tabla 9.

Estadísticos descriptivos para la variable Resiliencia según Nivel educativo.

Descriptivas

	Nivel educativo	N	Media	Mediana	DE
Resiliencia	Primario	51	21.6	23.0	6.44
	Secundario	133	23.3	24.0	6.81
	Universitario	89	23.7	24.0	6.58

Los resultados observados en la tabla 9 sugieren una posible diferencia en los valores medios. Observa que tanto los valores de Media, como de Mediana, de la variable Resiliencia parecen diferir según los niveles educativos. Sin embargo, recuerda que esas diferencias pueden ser sólo aparentes por lo que vamos a analizarlas con el estadístico correspondiente.

Estrategia paramétrica

Los estadísticos paramétricos usado para ANOVA de un factor pueden ser diversos. Sin embargo, los de mayor popularidad y robustez son Tukey (cuando se asume que las varianzas poblacionales son iguales) o Games-Howell (en el caso de que se asuma que las varianzas poblacionales no son iguales).

Supuestos asociados:

- La distribución de los datos de la variable dependiente es aproximadamente normal
- El tamaño de las muestras es mayor a 30
- La variable dependiente tiene un nivel de medición al menos intervalar.
- La varianza es constante entre los grupos (homogeneidad de varianzas o homocedasticidad).

Como sabes, si se cumplen estos supuestos el investigador elegirá esta prueba, de lo contrario buscará su alternativa no paramétrica.

*Nota: Recuerda, como todos los análisis paramétricos el contraste de hipótesis se hará teniendo en cuenta la Media.

Pasos

Formular las hipótesis nula y alternativa

• H_0= La media de Resiliencia NO presenta variaciones significativas a nivel estadístico según Nivel educativo.

• H_1= La media de Resiliencia presenta variaciones significativas a nivel estadístico según Nivel educativo.

Establecer el estadístico de prueba adecuado: ANOVA de un factor asumiendo varianzas iguales. En este caso al asumir que las varianzas poblacionales son iguales el análisis se complementará con Tukey, de lo contrario sería Games-Howell.

Seleccionar un nivel de significación: Alfa de 0,05

Establecer la regla de decisión

$$\begin{cases} H_0 : \bar{x}(\text{Primario}) = \bar{x}(\text{Secundario}) = \bar{x}(\text{Universitario}) \\ H_1 : \bar{x}(\text{Primario}) \neq \bar{x}(\text{Secundario}) \neq \bar{x}(\text{Universitario}) \end{cases}$$

Calcular el valor observado del estadístico de prueba (Procedimiento mediante Jamovi)

Inicialmente nos dirigirnos a Análisis > ANOVA > ANOVA de un factor

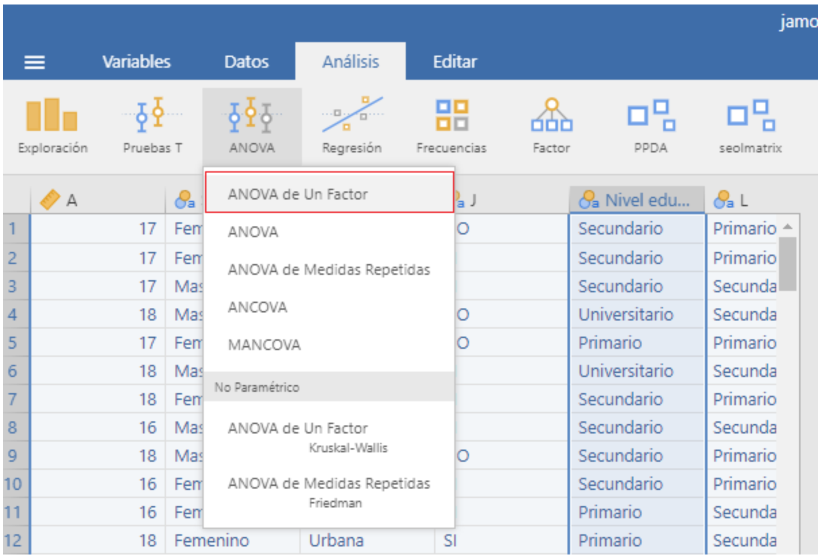

Figura 9: Captura de pantalla del procedimiento ANOVA de un factor realizado mediante el programa estadístico Jamovi

Luego configuramos el análisis como se mostrará en la Figura 10

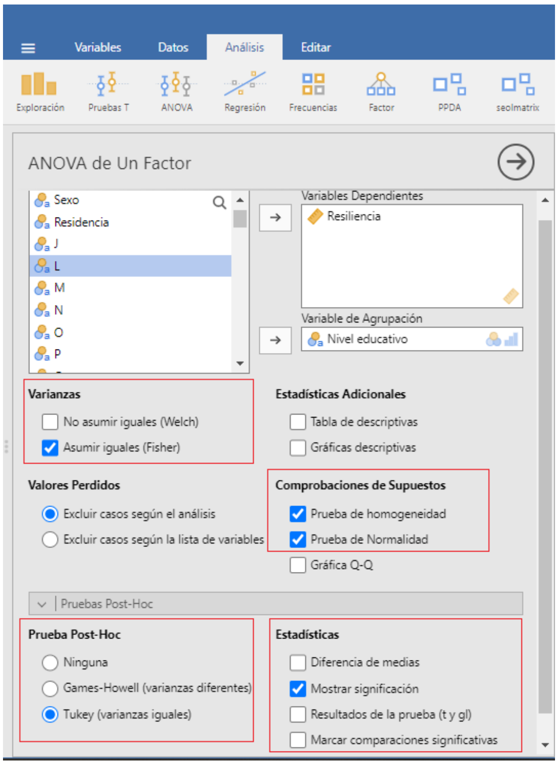

Figura 10: Captura de pantalla de las opciones para el análisis ANOVA de un factor realizado mediante Jamovi

En la figura 10 se observa que hay un apartado de Pruebas Post-Hoc. Los estadísticos de acá nos servirán para, si se rechaza H_0 y se confirma que hay diferencias entre grupos, determinar entre qué par de grupos está la diferencia.

Resultados

Tabla 10.

Análisis de Normalidad

Normality Tests

		statistic	p
Resiliencia	Shapiro-Wilk	0.978	< .001
	Kolmogorov-Smirnov	0.0690	0.148
	Anderson-Darling	1.52	< .001

Nota. Additional results provided by *moretests*

Seguramente recuerdas que se sugiere Kolmogorov-Smirnov cuando hay más de 50 datos u observaciones, y que si los valores p de la prueba son iguales o menores al Alfa propuesto por el investigador (generalmente 0,05) tendremos que rechazar H_0 y afirmar que no presentan una distribución normal.

Interpretación de la prueba de normalidad

Observando los valores de la tabla 10 asociados a Kolmogorov-Smirnov (valor p = 0,148) podemos afirmar que los datos provienen de una distribución normal, lo que permite continuar con la estrategia de análisis paramétrico.

Tabla 11.

Análisis de homogeneidad

Homogeneity of Variances Tests

		Statistic	df	df2	p
Resiliencia	Levene's	0.0744	2	270	0.928
	Bartlett's	0.266	2		0.875

Nota. Additional results provided by *moretests*

Interpretación de la prueba de Homogeneidad de varianzas

En la tabla 11 observamos que el valor p (0,928) asociado a la prueba de Levene, sugiere que los datos de las muestras son homogéneos. Por lo que se aceptaría el supuesto de homogeneidad de varianzas y sería aceptable continuar con el análisis paramétrico.

Tabla 12.

Prueba ANOVA de un factor para la variable Resiliencia según Nivel educativo

ANOVA de Un Factor (Fisher)

	F	gl1	gl2	p
Resiliencia	1.79	2	270	0.169

ANOVA pone a prueba la Hipótesis nula de que los grupos no presentan diferencias en los valores medios, por lo que valores mayores al nivel de significación o Alfa son indicativos de que debemos aceptarla.

Interpretación considerando el valor p

Como partimos de la base que las varianzas poblacionales eran iguales, el cálculo de ANOVA se realizó mediante el estadístico Fisher. En la tabla 12 se observa que el valor p asociado a F fue de 0,169, por lo que se concluye que no hay evidencia suficiente para afirmar que la variable Resiliencia presenta valores promedios distintos según el Nivel educativo alcanzado.

Bien, en este caso no hemos podido corroborar diferencias en los valores promedios en los distintos grupos, por lo que el análisis finalizaría aquí. Pero ¿qué sucedería si ANOVA nos hubiese ofrecido un valor p menor a 0,05? En este caso nos interesaría descubrir entre qué pares de grupos se presenta la diferencia, por ejemplo: Primario y Secundario, Primario y Universitario o Secundario y Universitario.

Para debelar ese interrogante es que ejecutamos el análisis Post-Hoc.

A continuación, te ofreceremos los resultados para el ejemplo que venimos trabajado, anticipándote que, como ANOVA determinó la aceptación de H_0, los valores serán concordantes confirmando que no hay diferencias

Tabla 13.

Prueba de Tukey para analizar diferencias intergrupos

Tukey Post-Hoc Test – Resiliencia

		Primario	Secundario	Universitario
Primario	valor p	—	0.248	0.162
Secundario	valor p		—	0.907
Universitario	valor p			—

La tabla 13 exhibe los valores p asociados a Tukey de cada prueba de hipótesis que busca comprobar la diferencia entre los valores de media de cada par de grupos. Así, el valor p de 0,248 que encontramos entre el nivel Primario y Secundario, nos confirma que no hay diferencias de media entre ambos grupos. El análisis se completa interpretando el resto de los valores p.

¿Qué hubiese pasado si ANOVA, en lugar aceptar H0, nos hubiese sugerido rechazarla? En ese caso el análisis Post-Hoc nos hubiera mostrado, al menos, algún par de grupos con un valor p menor a 0,05, y así podríamos haber detectado entre qué muestras realmente se encuentra tal diferencia.

Estrategia no paramétrica

Desde esta perspectiva el ANOVA de un factor se realiza mediante el estadístico Kruskal-Wallis.

Supuestos asociados: esta prueba no asume que los datos han sido extraídos de una distribución Normal ni que sean homogéneos, por lo que puede implementarse independientemente de estas restricciones. No condiciona su uso la cantidad de datos que tengan los grupos de análisis, además la variable sometida a análisis puede tener un nivel de medición ordinal, intervalar o continua. Recuerda, el contraste en este caso se hará teniendo en cuenta la Mediana (NO LA MEDIA).

Pasos

Formular las hipótesis nula y alternativa

• H_0= La mediana de Resiliencia NO presenta variaciones significativas a nivel estadístico según Nivel educativo.

• H_1= La mediana de Resiliencia presenta variaciones significativas a nivel estadístico según Nivel educativo.

Establecer el estadístico de prueba adecuado: <u>ANOVA de un factor mediante Kruskal-Wallis</u>

Seleccionar un nivel de significación: <u>Alfa de 0,05</u>

Establecer la regla de decisión

$$\begin{cases} H_0 : Me(Primario) = Me(Secundario) = Me(Universitario) \\ H_1 : Me(Primario) \neq Me(Secundario) \neq Me(Universitario) \end{cases}$$

Calcular el valor observado del estadístico de prueba (Procedimiento mediante Jamovi)

Inicialmente nos dirigirnos a Análisis > ANOVA > ANOVA de un factor

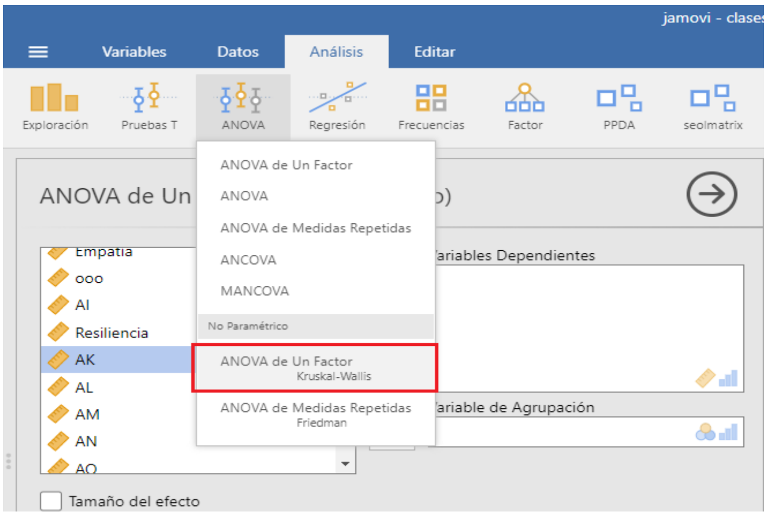

Figura 11: Captura de pantalla del procedimiento ANOVA de un factor (no paramétrico) realizado mediante el programa estadístico Jamovi

Luego configuramos el análisis como se mostrará en la Figura 12

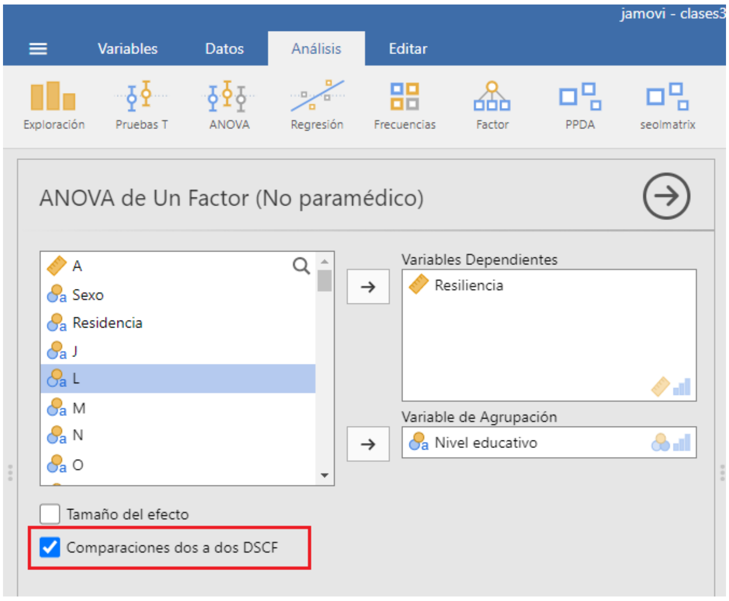

Figura 11: Captura de pantalla de las opciones para el análisis ANOVA de un factor (no paramétrico) realizado mediante Jamovi

En la figura 11 vemos marcada la opción Comparaciones dos a dos, ésta nos permitirá (en el caso que se rechace H_0) identificar entre qué par de grupos está la diferencia.

Resultados

Tabla 14.

Prueba ANOVA de un factor mediante Kruskal-Wallis para la variable Resiliencia según Nivel educativo

Kruskal-Wallis

	χ^2	gl	p
Resiliencia	2.73	2	0.255

Interpretación de los resultados

La tabla 14 muestra que el valor p asociado a la prueba fue de 0,255, por lo que se acepta H0 y se concluye que las medianas de la variable Resiliencia no son diferentes a nivel estadístico según el Nivel educativo alcanzado.

Del modo similar al análisis paramétrico, no hemos logramos confirmar la diferencia de medianas. De todas maneras, ofreceremos a continuación el análisis comparativo entre grupos.

Tabla 15.

Prueba de comparaciones dos a dos de Dwass-Steel-Critchlow-Fligner

Comparaciones entre parejas - Resiliencia

		W	p
Primario	Secundario	2.091	0.301
Primario	Universitario	2.186	0.270
Secundario	Universitario	0.267	0.981

La tabla 15 exhibe los valores p asociados a Dwass-Steel-Critchlow-Fligner de cada prueba de hipótesis que busca comprobar la diferencia entre los valores de mediana de cada par de grupos. Así, y de manera consistente con el análisis de ANOVA, se acepta la hipótesis nula que afirma la inexistencia de diferencias entre las medianas de los pares de grupo. Si ANOVA nos hubiese permitido rechazar H_0, este análisis hubiera mostrado al menos algún par de grupos con un valor p menor a 0,05.

Pruebas para la diferencia de valores medios entre dos muestras relacionadas (dependientes o apareadas)

Como se ha visto, en el análisis de las diferencias de valores medios para dos muestras independientes, éstas se constituyen con una característica esencial: las unidades de análisis (participantes) no pueden estar en dos grupos. En cambio, cuando hablamos de muestras relacionadas, dependientes o apareadas, las unidades de análisis (participantes) tienen que estar en dos (o más) muestras.

Podrías preguntarte ¿en qué situaciones puede ocurrir esto?

Pensemos en una situación que se nos puede plantear; como investigadores deseamos corroborar que un tratamiento que aplicamos, con la finalidad de mejorar la respuesta emocional frente al estrés, es útil y brinda buenos resultados. Es decir, si el tratamiento aplicado es eficaz. ¿Cuál sería el esquema de investigación que deberíamos aplicar?

La lógica nos invitaría a tomar en un grupo de personas dos mediciones, por lo menos, ¿verdad?

A ese tipo de diseño se los denomina Test-Retest y se caracterizan por ser una buena estrategia para evaluar los resultados de un tratamiento en un grupo de personas. Aquí el investigador define inicialmente la muestra con la que trabajará, luego aplica instrumentos de medición en la fase previa del tratamiento (Test), una vez concluida las mediciones administrará el tratamiento (puede consistir en una o muchas sesiones donde aplique una o muchas técnicas), y finalizado éste volverá a realizar las mediciones (con los mismos instrumentos) a la misma muestra (Retest). Como podrás darte cuenta, en este diseño de investigación cada participante tendrá dos mediciones.

Al conjunto de las mediciones iniciales conformadas por todos los participantes es considerada como la Muestra 1, en tanto que la totalidad de las mediciones realizadas en la etapa de Retest conformarán la Muestra 2. De aquí proviene la noción de muestras relacionadas o apareadas (dos muestras que están constituidas por dos mediciones -realizadas en momentos distintos y con los mismos instrumentos- de todos los participantes de una investigación).

Veamos un ejemplo

Supongamos que hemos desarrollado un procedimiento de 4 sesiones para disminuir la ansiedad en profesionales de la salud que trabajan realizando guardias en hospitales públicos y deseamos saber si efectivamente hay un descenso en los niveles de ansiedad una vez administrado el programa de intervención.

Seleccionamos una muestra de 30 profesionales con altos niveles de ansiedad; por lo que, lo primero que haremos, será tomar las mediciones de la variable para la totalidad de los participantes, teniendo la precaución de asignarle a cada uno de ellos un código para identificarlos adecuadamente en la segunda medición y hacer efectiva la comparación de los puntajes obtenidos. Posteriormente se realizará la implementación del tratamiento a la totalidad de los participantes, lo que demandará algún tiempo considerando que habíamos diseñado un procedimiento de 4 sesiones; y, finalmente, luego de haber finalizado la intervención tomaremos una última medición.

	Participante	Ansiedad Pre-tratamiento	Ansiedad Pos-tratamiento
1	a001	80	79
2	a002	74	70
3	a003	86	84
4	a004	71	75
5	a005	75	72
6	a006	79	78
7	a007	78	76
8	a008	77	74
9	a009	84	82
10	a010	75	74

Figura 12: Matriz de datos de muestras relacionadas

En la figura 12 podemos observar cómo se ordenarían los resultados de las mediciones realizadas. Por ejemplo, el participante a001 presentó inicialmente un nivel de ansiedad de 80, luego del tratamiento obtuvo 79, así sucesivamente.

Tabla 16.

Análisis descriptivo para la variable Ansiedad Pre y Postratamiento

Descriptivas

	N	Media	Mediana	DE
Ansiedad Pre-tratamiento	30	79.5	79.0	4.27
Ansiedad Pos-tratamiento	30	78.6	78.5	4.71

Como podemos apreciar en la tabla 16, aparentemente existe una diferencia en los valores de media entre ambas mediciones. Antes de la aplicación del tratamiento los participantes tenían una media de Ansiedad de 79,5, luego los niveles bajaron a 78,6. Para estar convencidos de la afirmación debemos realizar una prueba de hipótesis bajo la suposición de que la intervención que realizamos logró bajar los niveles de ansiedad de los participantes de manera significativa.

Estrategia paramétrica

El estadístico paramétrico usado para este tipo de análisis es la Prueba t de Student para muestras relacionadas o apareadas.

Supuestos asociados:

- La distribución de los datos de la variable dependiente es aproximadamente normal
- El tamaño de la muestra es mayor a 30
- Las variables tienen un nivel de medición al menos intervalar.

*Nota: Recuerda, como todos los análisis paramétricos, el contraste de hipótesis se hará teniendo en cuenta la Media.

Pasos

Formular las hipótesis nula y alternativa

- *H0*= Los valores de media en la variable Ansiedad Post-tratamiento NO son menores a los del Pre-tratamiento.

- *H1*= Los valores de media en la variable Ansiedad Post-tratamiento son menores a los del Pre-tratamiento.

Establecer el estadístico de prueba adecuado: Prueba t de Student para muestras apareadas

Seleccionar un nivel de significación: Alfa de 0,05

Establecer la regla de decisión

$$\begin{cases} H_0 : \bar{x}\,(\text{Pre-intervención}) <= \bar{x}\,(\text{Post-intervención}) \\ H_1 : \bar{x}\,(\text{Pre-intervención}) > \bar{x}\,(\text{Post-intervención}) \end{cases}$$

**Nota: obsérvese que en este caso la prueba es unilateral izquierda ya que buscamos comprobar si la media de post-intervención es menor que la de pre-intervención.*

Calcular el valor observado del estadístico de prueba (Procedimiento mediante Jamovi)

Ejecutamos Análisis > Pruebas T > Pruebas T para muestras Apareadas

Figura 13: Captura de pantalla del procedimiento Prueba T para muestras relacionadas realizado mediante el programa estadístico Jamovi

Luego configuramos el análisis como se mostrará a continuación

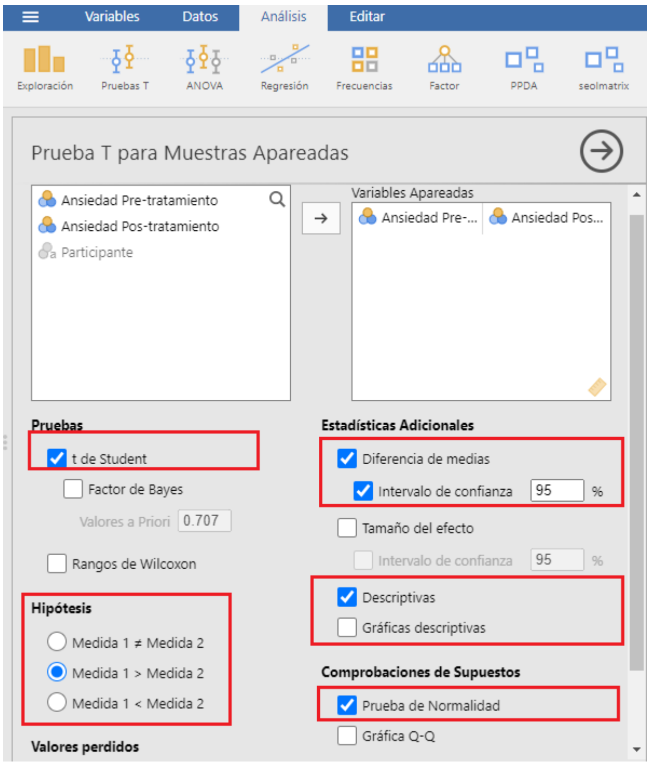

Figura 14: Captura de pantalla de las opciones para el análisis de Prueba T para muestras relacionadas realizado mediante Jamovi

Resultados

Tabla 17.

Prueba de normalidad

Tests of Normality

			statistic	p
Ansiedad Pre-tratamiento	Ansiedad Pos-tratamiento	Shapiro-Wilk	0.809	0.091
		Kolmogorov-Smirnov	0.352	0.081
		Anderson-Darling	2.52	0.081

Nota. Additional results provided by *moretests*

Como recuerdas, uno de los pasos previos a todo análisis paramétrico es saber si los datos cumplen con los supuestos del estadístico a aplicar. Atentos que la muestra presenta 30 datos, tomaremos el análisis de Shapiro-Wilk.

En la Tabla 17 se puede apreciar que Shapiro-Wilk nos informa que, con un valor p de 0,091, los datos provienen de una distribución normal, por lo que habiéndose cumplido todos los supuestos podemos proseguir con el análisis paramétrico.

Tabla 18.

Análisis de diferencias de media para muestras relacionadas mediante Prueba t de Student

Prueba T para Muestras Apareadas

			estadístico	gl	p
Ansiedad Pre-tratamiento	Ansiedad Pos-tratamiento	T de Student	2.36	29.0	0.013

Nota. H_a $\mu_{\text{Medida 1 - Medida 2}} > 0$

Interpretación de los resultados

Como hemos dicho con anterioridad, para llegar a una conclusión en la prueba de hipótesis debemos comparar el valor p obtenido del cálculo, con el nivel de significación (con el Alfa) que hemos decidido con anterioridad. Si el valor p resulta igual o menor al Alfa, RECHAZAREMOS H0. Si el valor p resulta mayor al Alfa, ACEPTAREMOS H0.

La tabla 18 nos muestra un valor de probabilidad asociado a la prueba t de Student de 0,013, siendo claramente menor al Alfa (0,05), por lo que afirmaremos que se rechaza H0 y que se confirma la hipótesis de que los niveles de Ansiedad fueron menores al concluir el tratamiento.

Estrategia no paramétrica

El estadístico no paramétrico para este tipo de análisis es la Prueba Rangos de Wilcoxon para muestras relacionadas

Supuestos asociados: esta prueba no exige normalidad en la distribución de datos, admite tamaños muestrales pequeños y puede u8sarse para el análisis de variables con un nivel de medición ordinal.

Pasos

Formular las hipótesis nula y alternativa

- *H0*= Los valores de media en la variable Ansiedad Post-tratamiento NO son menores a los del Pre-tratamiento.
- *H1*= Los valores de media en la variable Ansiedad Post-tratamiento son menores a los del Pre-tratamiento.

Establecer el estadístico de prueba adecuado: Prueba Rangos de Wilcoxon para muestras relacionadas

Seleccionar un nivel de significación: Alfa de 0,05

Establecer la regla de decisión

$$\begin{cases} H_0 & : Me(\text{Pre-intervención}) <= Me(\text{Post-intervención}) \\ H_1 & : Me(\text{Pre-intervención}) > Me(\text{Post-intervención}) \end{cases}$$

*Nota: atención en este caso el contraste se hace sobre la Mediana y no sobre la Media y se plantea una prueba unilateral

Calcular el valor observado del estadístico de prueba (Procedimiento mediante Jamovi)

Análisis > Pruebas T > Pruebas T para muestras apareadas (hasta acá el procedimiento es similar al análisis paramétrico)

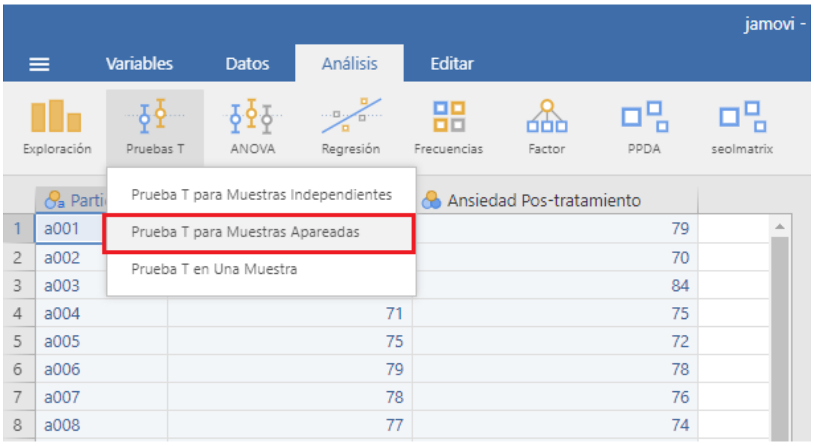

Figura 15: Captura de pantalla del procedimiento para diferencias entre muestras relacionadas (no paramétrico) realizado mediante el programa estadístico Jamovi

Luego configuramos el análisis como se mostrará a continuación

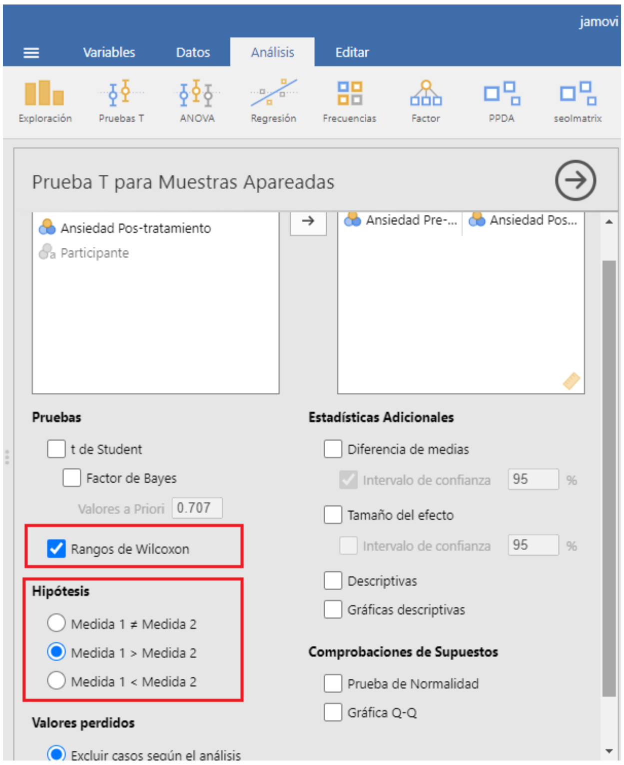

Figura 16: Captura de pantalla de las opciones para el análisis de Prueba Rangos de Wilcoxon para muestras relacionadas realizado mediante Jamovi

Resultados

Tabla 19.

Análisis de diferencias de media para muestras relacionadas mediante Prueba Rangos de Wilcoxon

Prueba T para Muestras Apareadas

			estadístico	gl	p
Ansiedad Pre-tratamiento	Ansiedad Pos-tratamiento	T de Student	2.36	29.0	0.013

Nota. H_a $\mu_{\text{Medida 1 - Medida 2}} > 0$

Interpretación de los resultados

La tabla 19 expone un valor de probabilidad asociado a Prueba Rangos de Wilcoxon de 0,013, siendo claramente menor al Alfa (0,05), por lo que afirmaremos que se rechaza H0 y que se confirma la hipótesis de que los niveles de Ansiedad, considerando las medianas, fueron menores al concluir el tratamiento.

*Nota: si deseas conocer la magnitud de dicha diferencia, puedes consultar los valores de medianas que se encuentran en el análisis descriptivo (tabla 16).

Bibliografía

Arias, F. (2012). *El proyecto de investigación. Introducción a la metodología científica*. 6ª Edición. Caracas: Editorial Episteme.

Frances-Garcia, F. J. (2019). *Técnicas de investigación social.* Universidad de Alicante. Recuperado el 30/04/19 de https://sites.google.com/site/tecninvestigacionsocial/

Levin, R. I. & Rubin, D.S. (2004*). Estadística para administración y economía*. México, Pearson Educación.

Ochoa, C. (2019). *Muestreo probabilístico: muestreo sistemático.* Recuperado el 24/04/2019 de https://www.netquest.com/blog/es/blog/es/muestreo-sistematico

Perelló-Oliver, S. (2009). *Metodología de la investigación social.* Madrid: Dykinson.

Requena, B. (2019). *Universo Fórmulas: Muestreo probabilístico.* Recuperado el 24/04/2019 de: https://www.universoformulas.com/estadistica/inferencia/muestreo-no-probabilistico/

Recursos on-line

CALCULADORAS

- Calculadora de Prueba z: http://www.learningaboutelectronics.com/Articulos/Calculadora-de-prueba-de-hipotesis-estadistica.php#answer
- Calculadora de Prueba T de Student: http://www.learningaboutelectronics.com/Articulos/Calculadora-de-prueba-T-de-Student.php
- Calculadora de Valor de p para Prueba T de Student: https://es.docpid.com/calculadoras/valor-de-p
- Calculadora de Prueba de T para Muestras Relacionadas: http://www.learningaboutelectronics.com/Articulos/Calculadora-de-prueba-T-para-muestras-relacionadas.php
- Calculadora de Prueba T para Muestras Independientes: http://www.learningaboutelectronics.com/Articulos/Calculadora-de-prueba-T-para-muestras-independientes.php

SIMULACIÓN DE DISTRIBUCIONES Y ZONAS DE RECHAZO

- Contraste de hipótesis para la media:

https://proyectodescartes.org/uudd/materiales_didacticos/inferencia_estadistica_JS/contrasmedia.htm

- Para varias simulaciones:

https://proyectodescartes.org/uudd/materiales_didacticos/inferencia_estadistica_JS/index.htm