Estadística Inferencial metodología

En la estadística inferencial clásica, la aceptación o rechazo de la hipótesis nula ((H0)) se basa en el valor p (p-valor) y un nivel de significancia predefinido (alpha). Aquí está el proceso general:

1. **Definir las hipótesis:**

   - **Hipótesis nula (H0)**: Es la hipótesis que se intenta probar o rechazar. **Generalmente, afirma que no hay efecto o diferencia**.

   - **Hipótesis alternativa (H1)**: Es la hipótesis que se adopta si se rechaza (H0). **Generalmente, afirma que hay un efecto o diferencia.**

2. **Realizar el test estadístico:** Se calcula el p-valor utilizando los **datos muestrales** y el **test estadístico** adecuado (t-test, chi-cuadrado, ANOVA, etc.).

3. **Comparar el p-valor con el nivel de significancia** (alpha):

   - El nivel de significancia (alpha) es la probabilidad de rechazar (H0) cuando (H0) es verdadera. Comúnmente se elige (alpha = 0.05) (5%), cuanto más baja es alpha menor será la probabilidad de equivocarnos, es decir de rechazar la hipótesis H0 siendo cierta

4. **Tomar la decisión**:

El p-valor es una **medida de la evidencia contra (H0)**

   - **Rechazar H0 y aceptar H1**:

Si el p-valor es *menor o igual* al nivel de significancia (p <= alpha), se rechaza (H0).

Esto sugiere que *hay suficiente evidencia en los datos para concluir que el efecto o diferencia observada es* ***estadísticamente significativa***.

   - **No rechazar (H0)**:

Si el p-valor es mayor que el nivel de significancia (p > alpha), **no** se rechaza (H0).

Esto sugiere que ***no*** *hay suficiente evidencia en los datos para concluir que el efecto o diferencia observada es estadísticamente significativa.*

Ejemplo Práctico

Supongamos que estamos realizando un test de dos colas con un nivel de significancia de (alpha = 0.05).

1. **Hipótesis**:

   - H0: La media de la población es igual a un valor específico (por ejemplo, (mu = 0).

   - H1: La media de la población no es igual a ese valor (por ejemplo, (mu<>0).

2. Realizar el test: Se calcula el p-valor *basado en los datos muestrales obtenidos*. El test puede ser cualquiera de los test paramétricos estudiados (t Student, Chi^2, Alnova, etc)

3. Comparar el p-valor:

   - Supongamos que el p-valor calculado es 0.03.

4. Tomar la decisión:

   - Como (p = 0.03) es menor que (alpha = 0.05), se rechaza (H0).

En este caso, concluiríamos que hay evidencia suficiente para rechazar la hipótesis nula y aceptar que existe una diferencia significativa en la media de la población.

<u>Consideraciones Importantes</u>

- El p-valor **no mide** la magnitud del efecto: Un p-valor bajo solo indica que los datos observados son poco probables bajo (H0), pero no indica cuán grande es el efecto.

- El p-valor no mide la probabilidad de que (H0) sea verdadera: El p-valor es una **medida de la evidencia contra (H0)**, no la probabilidad de que (H0) sea verdadera.

- El nivel de significancia (alpha) es arbitrario: Aunque (alpha = 0.05) es común (indica una probabilidad del 5%), otros valores como 0.01(probabilidad del 1%) o 0.10 (probabilidad del 10%) también pueden ser utilizados dependiendo del contexto del estudio.

En resumen, el p-valor es una herramienta importante en la estadística inferencial para tomar decisiones sobre la hipótesis nula, pero debe interpretarse con cuidado y en el contexto de otros factores y consideraciones.

Criterios de Selección de Estadísticos en Función de las Características y Variables de las Muestras

Tabla resumen.

Prueba (Test)	Variables	Descripción
Chi² (Chi Cuadrado) = Ji Cuadrado	Dos variables cualitativas o categóricas	Se usa para evaluar la **independencia o asociación** entre dos variables categóricas.
t de Student	Una variable cualitativa (**2 niveles o categorias**) y una cuantitativa (VD)	Compara las **medias** de **dos grupos** independientes para determinar si son significativamente diferentes.
ANOVA (Análisis de varianza)	Una variable cualitativa (**más de 2 niveles**) y una cuantitativa (VD)	Compara las **medias** de **tres o más grupos** independientes.
Regresión lineal	Una variable cuantitativa (VD) y una o más variables cuantitativas o cualitativas (codificadas) (VI)	Examina la relación entre una variable dependiente continua y una o más variables independientes.
Correlación de Pearson	Dos variables cuantitativas	Mide la fuerza y dirección de la relación lineal entre dos variables continuas.
Pruebas no paramétricas		
U de Mann-Whitney	Una variable cualitativa (2 niveles) y una ordinal o continua (VD)	Comparación **no paramétrica** de dos grupos independientes cuando no se cumplen los supuestos del t de Student.
Kruskal-Wallis	Una variable cualitativa (más de 2 niveles) y una ordinal o continua (VD)	Comparación **no paramétrica** de tres o más grupos independientes cuando no se cumplen los supuestos del ANOVA.
	Tabla 1	

VI: Variable Independiente VD: Variable Dependiente

En el análisis estadístico, la elección del test adecuado depende de las características de las variables y de la naturaleza de los datos. A continuación, se describen varios estadísticos comunes, junto con los criterios para su selección.

1.**Test de Chi-Cuadrado (Chi² - Chi cuadrado)**

- Uso Principal: Comparar *frecuencias observadas* con *frecuencias esperadas*.

- **Tipos de Variables:** Variables categóricas.

- **Características de las Muestras**:

  - Datos en tablas de contingencia.

  - Las frecuencias esperadas en cada celda **deben ser al menos 5 para asegurar la validez del test.**

- **Aplicaciones Comunes**:

  - Test de independencia: Verifica si dos variables categóricas están relacionadas.

  - Test de bondad de ajuste: Compara la distribución observada de una variable categórica con una distribución teórica.

Grados de libertad: El cálculo de grados de libertad gl (en inglés degrees of freedom = df) se hace mediante la formula;

gl= (nf –1)x(nc–1) donde nf es el número de filas de la tabla de contingencia y nc el de columnas.

EL significado de ese grado de libertad es cuantas celdas de la tabla de contingencia podemos alterar libremente manteniendo los valores totales de filas y columnas.

Frecuencias esperadas. Para calcular la frecuencia esperada (EEE) de una celda en una tabla de contingencia, se utiliza la siguiente fórmula:

$$E = \frac{(F_i \cdot F_j)}{N}$$

donde:

- Fi es el total (suma) de todas las celdas de la fila que contiene la celda.
- Fj es el total (suma) de todas las celdas de la columna que contiene la celda.
- N es el total general de todas las observaciones en la tabla.

Practicar con los datos de la Figura 2 en el ejemplo de Chi²

Videos: https://youtu.be/odIbwJ4KBRo?t=8

https://youtu.be/176MHbH7Lns?t=4

https://youtu.be/PkUEr1p63wA?t=6

Coeficiente de contingencia es una medida de asociación entre dos variables categóricas y se utiliza junto con la prueba Chi-cuadrado para **evaluar la fuerza de la relación entre las variables** en una tabla de contingencia. Este coeficiente varía entre 0 y 1, donde 0 indica ninguna asociación y valores cercanos a 1 indican una fuerte asociación.

La fórmula es la siguiente:

$$C = \sqrt{\frac{\chi^2}{\chi^2+N}}$$

Es importante mencionar que el **coeficiente de contingencia** no puede alcanzar el valor máximo de 1, **especialmente en tablas de contingencia grandes**. Por lo tanto, para tablas más grandes, otras medidas como la **V de Cramér pueden ser más adecuadas** para evaluar la fuerza de la asociación.

V de Cramer es una medida de asociación entre dos variables categóricas y se utiliza junto con la prueba Chi-cuadrado para **evaluar la fuerza de la relación entre las variables** en una tabla de contingencia.

La fórmula es:

$$V = \sqrt{\frac{\chi^2}{N \times \min(r-1, c-1)}}$$

min(r-1,c-1) es el número mínimo del número de columnas menos uno y del número de filas menos

Ejemplo práctico. Test de Chi-Cuadrado (Chi² - Chi cuadrado). Test de independencia de dos variables categóricas.

Este es un ejemplo del uso de Chi² para verificar si dos variables categóricas están relacionadas, es decir, el primer y principal uso de esta prueba que llamábamos test de independencia.

Para realizar una prueba de Chi-cuadrado en JASP, sigue estos pasos detallados:

Ejemplo de Prueba Chi-Cuadrado en JASP

Supongamos que tenemos datos sobre la preferencia de bebidas (café, té, jugo) entre hombres y mujeres. Queremos usar la prueba de Chi-cuadrado para ver si hay una relación significativa entre el género y la preferencia de bebida.

Observa que aquí tenemos dos variables categóricas que son: Genero y Bebida. De los dos tipos que hay de variables categóricas, nominales y ordinales, estas son del tipo nominal. Las ordinales son por ejemplo nivel de satisfacción (bajo. medio, alto)

Paso 1: Establecer las hipótesis:

- $H0$: No hay relación entre el género y la preferencia de bebida.
- $H1$: Existe una relación significativa entre el género y la preferencia de bebida.

En términos más formales:

- *H0* : La variable "género" es independiente de la variable "preferencia de bebida".
- *H1* : La variable "género" no es independiente de la variable "preferencia de bebida".

Paso 2: Realizar el test estadistico

Seleccionamos el test estadístico a utilizar, puesto que queremos comparar la dependencia de dos variables categóricas el test adecuado según la tabla1 es Chi².

Imaginemos que los datos están en una tabla de contingencia como la siguiente:

Genero	Café	Té	Jugo
Hombre	30	10	10
Mujer	20	25	15

Abrir JASP.

Ingresar los Datos en JASP

Crear un nuevo archivo o abrir un archivo existente.

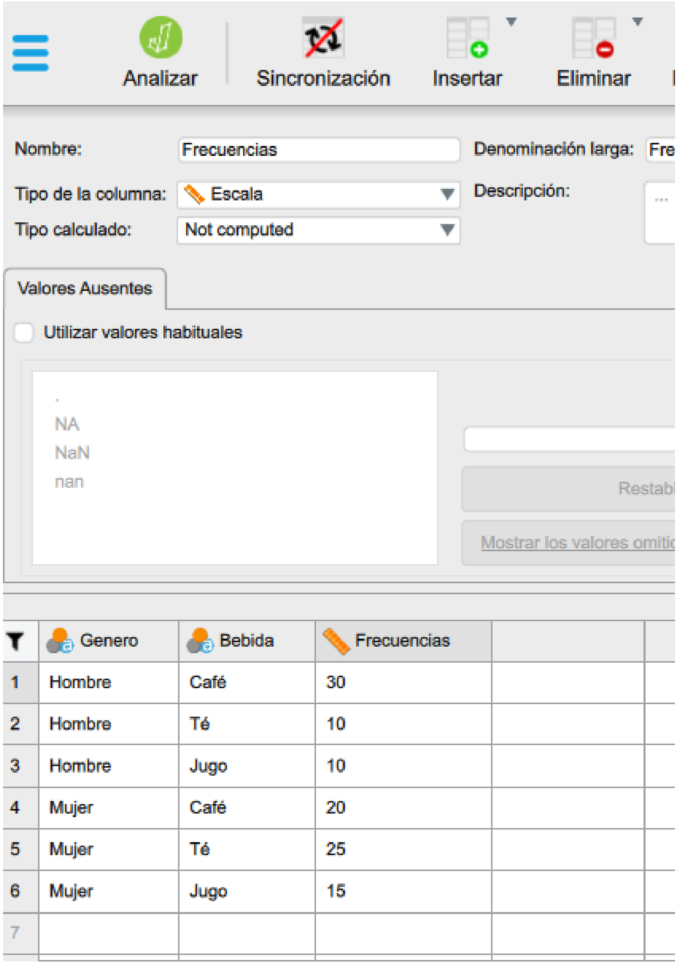
Ingresar los datos: Puedes ingresar los datos manualmente o importarlos desde un archivo CSV/Excel.

Para ingresar manualmente:

- Ve a la pestaña "Data" (Editar Datos).

- Introduce dos variables categóricas: "Género" y "Bebida".

Ejemplo de cómo se verían los datos en JASP:



Realizar la Prueba Chi-Cuadrado

- Selecciona la pestaña Analizar
- Seleccionar la pestaña "Frequencies" (Frecuencias) en JASP.
- Elegir "Contingency Tables" (Tablas de Contingencia).
- Arrastrar "Género" a la caja de "Rows" (Filas) y "Bebida" a la caja de "Columns" (Columnas) y Frecuencias a “Recuentos”.
- Marcar "Chi-square test of independence" (Prueba de independencia de Chi-cuadrado).

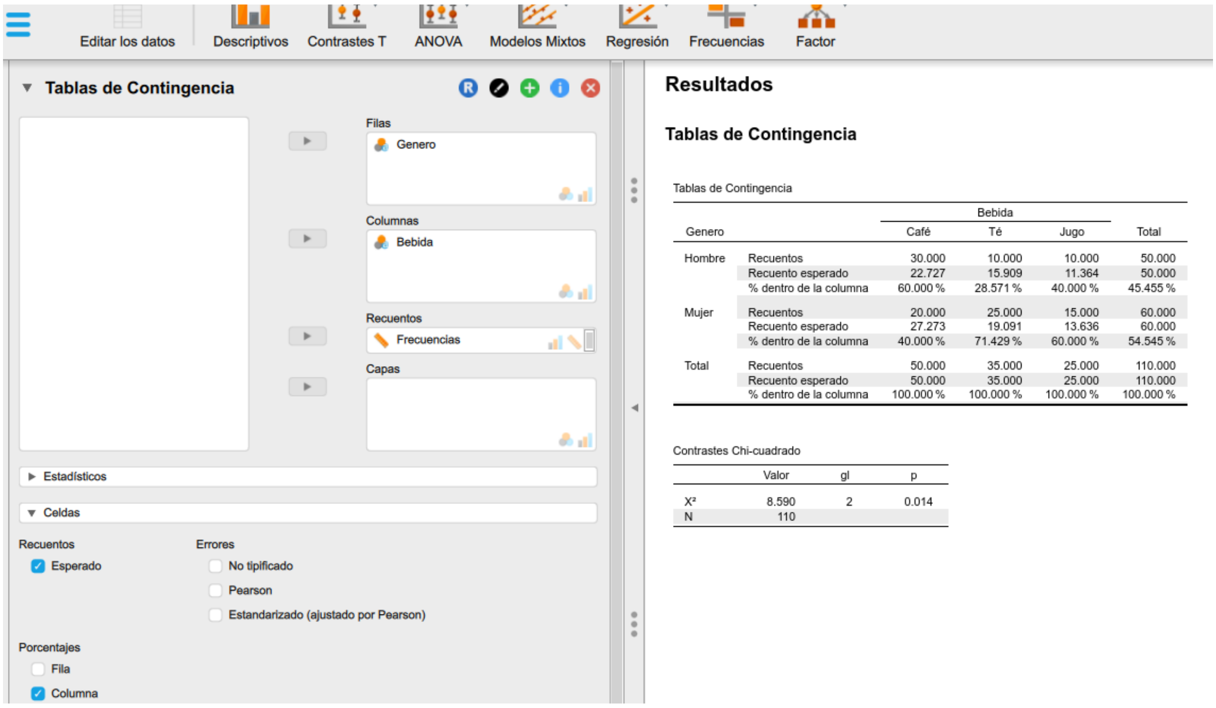


Figura 2

JASP automáticamente calculará la tabla de contingencia y realizará la prueba de Chi-cuadrado.

Paso 3: Interpretar los Resultados

JASP mostrará los resultados de la prueba Chi-cuadrado, incluyendo:

- Valor de Chi-cuadrado (chi^2), aquí da el valor 8,590

- Grados de libertad (df), en este ejemplo gf = 2 **calculado** multiplicando (2-1)x(3-1), la variable genero puede tomar dos valores, la bebida tres.

- Valor p (p-value), en este caso p = 0,014, si usamos un nivel de confianza del 5% , es decir de 0,05 entonces 0,014 < 0,05 y rechazaríamos la hipótesis nula (es decir que no hay diferencia entre género y tipo de bebida consumida), o dicho de otra forma aceptamos la hipótesis alternativa H1 hay una relación estadísticamente significativa entre el género y la preferencia de bebida. Obsérvese que, si quisiéramos un nivel de confianza del 1%, es decir 0,01 tendríamos que p> 0,01 y tendríamos que aceptar la hipótesis nula, es decir, no hay evidencia en los datos de que haya una relación entre sexo y preferencias de bebida.

Ejemplo práctico. Test de Chi-Cuadrado (Chi² - Chi cuadrado). Test de Bondad de ajuste.

Un ejemplo de uso del test de Chi-cuadrado para la bondad de ajuste podría implicar comparar las frecuencias observadas de una variable categórica con las frecuencias esperadas según una distribución teórica.

Supongamos que queremos verificar si un dado es justo, o sea, si cada una de las seis caras tiene la misma probabilidad de aparecer.

Supongamos que lanzamos el dado 60 veces y obtenemos las siguientes frecuencias observadas para cada cara:

Cara	Frecuencia
1	8
2	12
3	10
4	9
5	11
6	10

Paso 1: Formular las Hipótesis

- **Hipótesis Nula (H0)**: El dado **es justo**, es decir, cada cara tiene una probabilidad igual de aparecer.
- **Hipótesis Alternativa (H1)**: El dado **no es justo**, es decir, al menos una cara tiene una probabilidad diferente de aparecer.

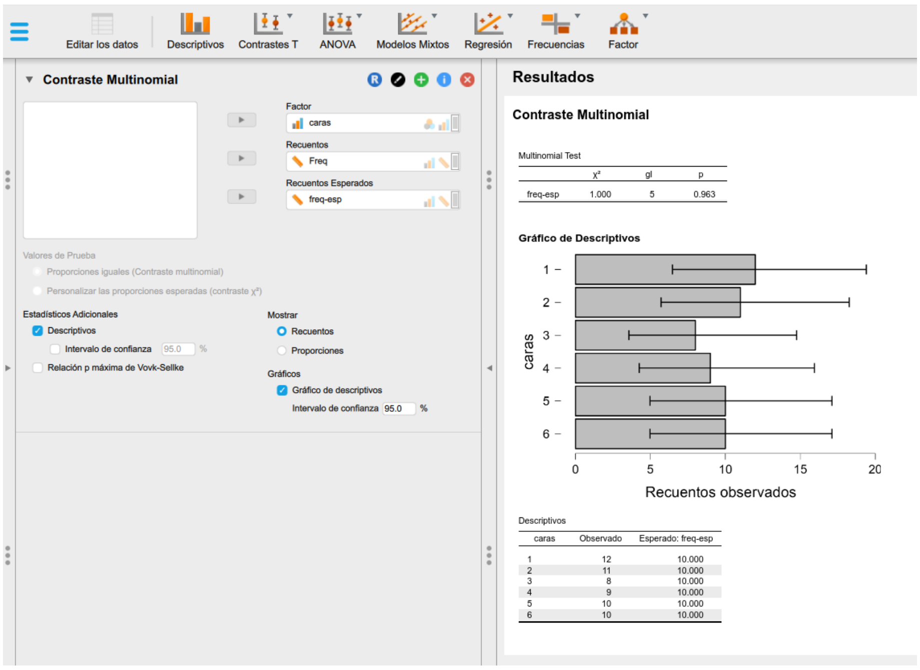
Paso 2: Calcular las Frecuencias Esperadas

Si el dado es justo, cada cara debería aparecer con la misma frecuencia. Dado que lanzamos el dado 60 veces y hay 6 caras, la frecuencia esperada para cada cara es:

Frecuencia esperada = 60 / 6 = 10

Paso 3. Abrir JASP y cargar el archivo de datos.

Paso 4. Seleccionar Analizar>Frecuencias>Contraste Multinomial. Ponga Caras en Factor, Frecuencia en recuentos y frecuencia-esperada en Recuentos esperados.



Paso 5. JASP calculará automáticamente el valor de $\chi 2 \backslash chi\hat{}2$ y el valor p.

Como vemos en la figura anterior en nuestro caso sería ***chi² = 1,000*** *Grados de libertad gl = 5 (6-1 = 5) y p=0,693*

Paso 6. Conclusión. Si tomamos el nivel de confianza habitual del 95%, alpha sería 0,05 y dado que p-valor=0,963 > 0,05 **no debemos rechazar la hipótesis nula H0,** en otras palabras, no hay evidencia estadística de que el dado esté trucado y no siga una distribución uniforme. No tenemos suficiente evidencia para rechazar la hipótesis nula. Por lo tanto, concluimos que el dado no muestra una desviación significativa de una distribución justa; es decir, los datos son consistentes con la suposición de que el dado es justo.

2. **Test t de Student**

- **Uso Principal:** Comparar las medias de una o dos muestras.

- T**ipos de Variables:** dos variables cuantitativas o una variable categórica de **dos niveles** y una cuantitativa.

- **Características de las Muestras:**

  - Normalidad de la distribución de la muestra (especialmente importante para muestras pequeñas).

  - Homogeneidad de varianzas para el test t de dos muestras independientes. La homogeneidad de varianzas para el test t de dos muestras independientes se conoce como la **prueba de homocedasticidad**. Este término indica que las varianzas de las dos muestras son iguales.

Para verificar la homocedasticidad, comúnmente se utiliza la **prueba de Levene** o la **prueba de Bartlett**. Ambas pruebas ayudan a determinar si las varianzas de los grupos son significativamente diferentes.

Si la homocedasticidad se cumple (es decir, si las varianzas son homogéneas), se puede proceder con el test t para muestras independientes estándar. Si no se cumple, es recomendable usar una versión del test t que no asume igualdad de varianzas, conocida como **el test t de Welch**.

- **Aplicaciones Comunes:**

- Test t para muestras independientes: Comparar las medias de dos grupos independientes.
- Test t para muestras relacionadas (o pareadas): Comparar las medias de dos mediciones relacionadas (por ejemplo, antes y después de un tratamiento en el mismo grupo).
- Test t para una muestra: Comparar la media de una muestra con un valor teórico.

Cálculo de los grados de libertad(gl): **gl** = (tamaño_muestra_A – 1)+ (tamaño_muestra_B – 1)

Valor de t de Student (con homocedasticidad)

La fórmula para el valor de t es:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Donde:

- $\bar{X}_1$ y $\bar{X}_2$ son las medias de las dos muestras.
- n_1 y n_2 son los tamaños de las dos muestras.
- S_p^2 es la varianza combinada (o ponderada) de las dos muestras, calculada como:

$$S_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$$

Aquí, s_1^2 y s_2^2 son las varianzas de las dos muestras.

Grados de libertad

Los grados de libertad df en este caso son simplemente:

$$df = n_1 + n_2 - 2$$

Nota; ***df** es degrees of freedom(Grados de libertad)*

Ejemplo 1: Test t para Dos Muestras Independientes

Supongamos que queremos comparar las puntuaciones de dos grupos de estudiantes en un examen. Los datos de las puntuaciones son los siguientes:

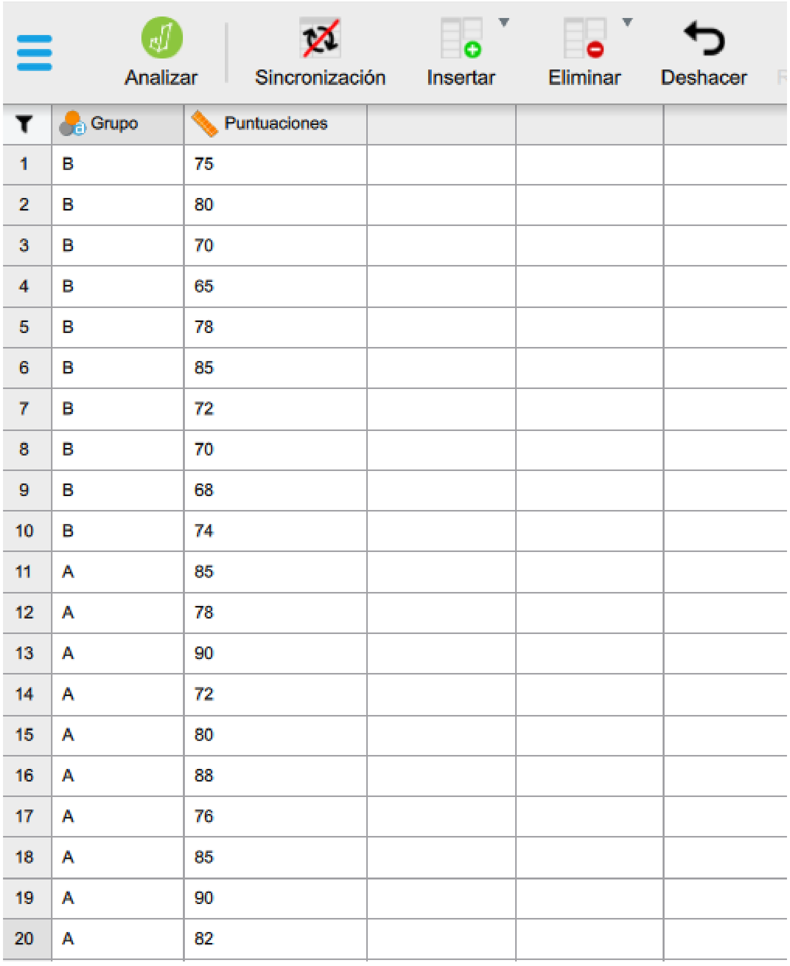
Grupo A	Grupo B
85	75
78	80
90	70
72	65
80	78
88	85
76	72
85	70
90	68
82	74

Paso 1: Formular las Hipótesis

- **Hipótesis Nula (H0)**: Las medias de los dos grupos son iguales (μA =μB).

- **Hipótesis Alternativa ($H1$)**: Las medias de los dos grupos son diferentes ($\mu A \neq \mu B$).

Paso 2: Abrir JASP y cargar los datos.

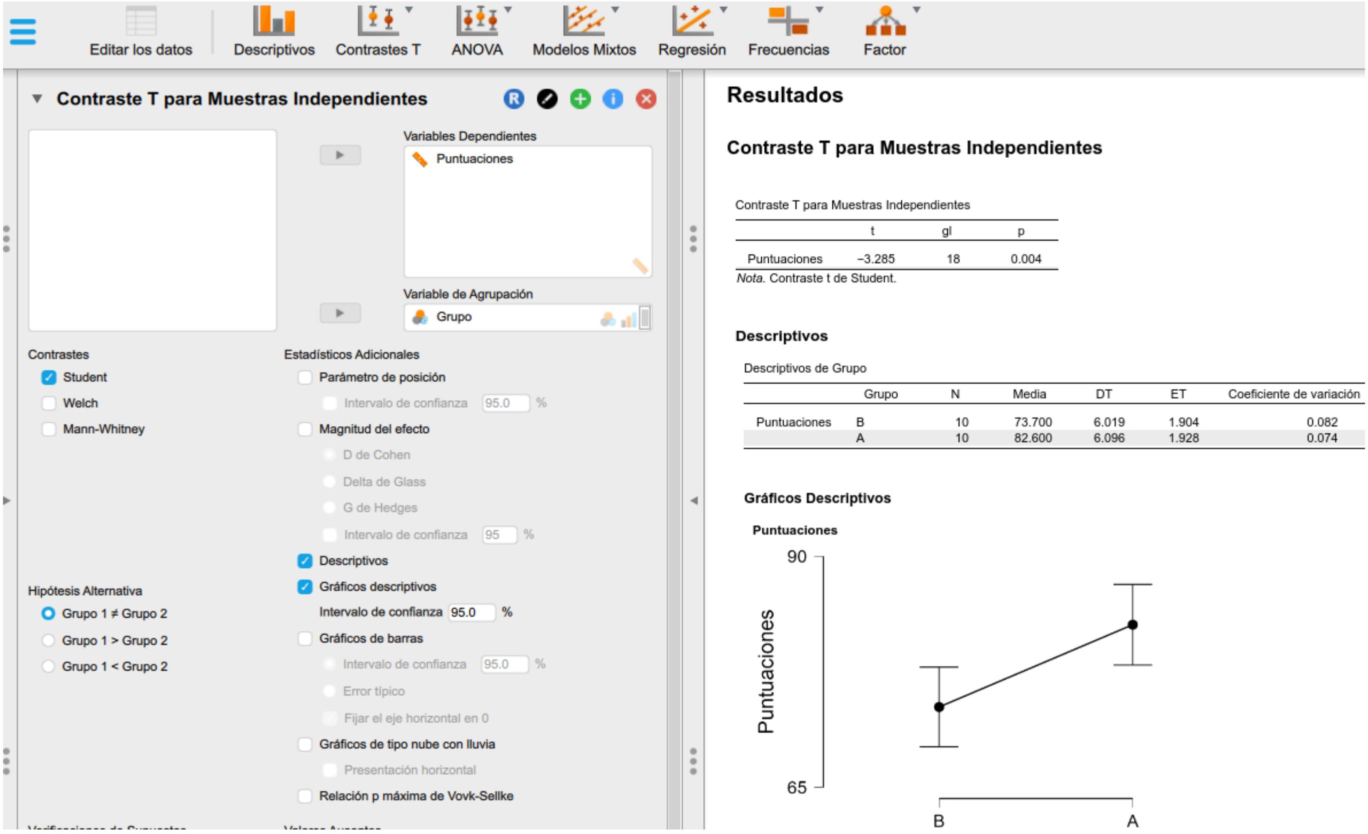


	Grupo	Puntuaciones
1	B	75
2	B	80
3	B	70
4	B	65
5	B	78
6	B	85
7	B	72
8	B	70
9	B	68
10	B	74
11	A	85
12	A	78
13	A	90
14	A	72
15	A	80
16	A	88
17	A	76
18	A	85
19	A	90
20	A	82

Paso 3: Seleccionar la pestaña "Contrastes T" (T-Tests).

Paso 4: Elegir prueba T de muestras independientes

Mover puntuaciones a **variable dependiente** y Grupo a **Variable de agrupación**.



Paso 5: Conclusión

Dado que el valor p es menor que el nivel de significancia $\alpha = 0.05$ rechazamos la hipótesis nula. Por lo tanto, concluimos que hay suficiente evidencia para decir que las medias de los dos grupos son significativamente diferentes.

Ejemplo 2: Test t de Student para una Muestra.

Supongamos que estamos interesados en determinar si los estudiantes de una clase tienen, en promedio, una puntuación diferente a 75 en un examen. Recogemos una muestra aleatoria de 10 estudiantes y obtenemos las siguientes puntuaciones:

78	82	85	70	68	90	74	80	77	72	

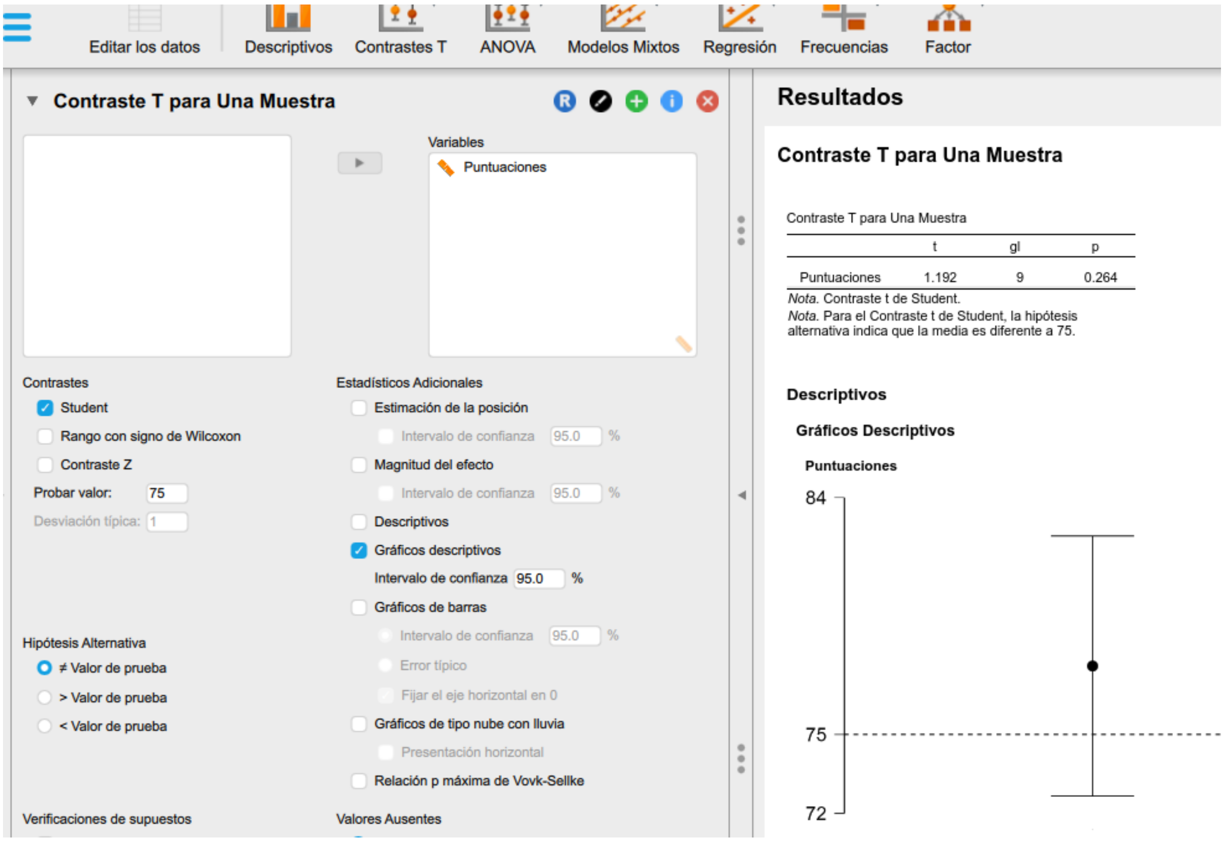
Paso 1: Formular las Hipótesis

- **Hipótesis Nula (H0)**: La media de la población es igual a 75 ($\mu = 75$).
- **Hipótesis Alternativa (H1)**: La media de la población es diferente de 75 ($\mu \neq 75$).

Paso 2:

Abrir JASP y cargar los datos.
Ingresar los datos: Ve a "Data" y asegúrate de que las puntuaciones estén en una columna.
Seleccionar la pestaña "T-Tests".
Elegir "One Sample T-Test" (Prueba T de una muestra).

Arrastrar la variable de puntuaciones a la caja de "Variables".
Especificar la media poblacional: En el campo "Probar valor", ingresa 75.
Ejecutar el análisis.



Conclusión

Dado que el valor p es mayor que el nivel de significancia α=0.05, $p = 0{,}264 > \ 0{,}05$, **no rechazamos la hipótesis nula**. Por lo tanto, concluimos que no hay suficiente evidencia para decir que la media de la población es diferente de 75.

3. **ANOVA (Análisis de Varianza)**

- **Uso Principal:** Comparar las medias de tres o más grupos.

- **Tipos de Variables:** Variables continuas.

- **Características de las Muestras**:

  - Normalidad de las distribuciones dentro de los grupos.

  - Homogeneidad de varianzas entre los grupos.

- Aplicaciones Comunes:

  - ANOVA de un factor: Comparar las medias de tres o más grupos independientes.

  - ANOVA de medidas repetidas: Comparar las medias de tres o más grupos relacionados.

  - ANOVA multifactorial: Examinar la interacción entre dos o más factores.

La distribución F de Snedecor, más comúnmente conocida como la distribución F, se utiliza en las pruebas ANOVA (Análisis de Varianza) porque permite comparar las variaciones entre los grupos con las variaciones dentro de los grupos. Esta comparación es fundamental para determinar si las medias de varios grupos son significativamente diferentes entre sí.

Propiedades de la Distribución F

1. **Asimetría**:
   - La distribución F es asimétrica y sesgada hacia la derecha.
2. **Dependencia de los Grados de Libertad**:
   - La forma de la distribución F depende de dos conjuntos de grados de libertad, uno asociado con la variabilidad entre los grupos ($d1$) y otro con la variabilidad dentro de los grupos ($d2$).
3. **Valores Positivos**:
   - Los valores de la distribución **F son siempre positivos** porque se basan en razones de varianzas que son siempre positivas.

Fórmula del Estadístico F

El estadístico F se define como el cociente de dos estimaciones de varianza, S_1^2 y S_2^2, normalizadas por sus respectivos grados de libertad d_1 y d_2:

$$F = \frac{\frac{S_1^2}{d_1}}{\frac{S_2^2}{d_2}}$$

Donde:

- S_1^2 es la varianza del primer grupo.
- S_2^2 es la varianza del segundo grupo.
- d_1 son los grados de libertad asociados con S_1^2.
- d_2 son los grados de libertad asociados con S_2^2.

Ejemplo de ANOVA

Supongamos que tenemos tres grupos con las siguientes medias y varianzas:

- Grupo 1: Media $=5= 5$=5, Varianza $=2= 2$=2, $n1=10n_1 = 10$n1 =10
- Grupo 2: Media $=7= 7$=7, Varianza $=3= 3$=3, $n2=10n_2 = 10$n2 =10
- Grupo 3: Media $=9= 9$=9, Varianza $=4= 4$=4, $n3=10n_3 = 10$n3 =10

Calcular las Sumas de Cuadrados:

4. **SSB**: Suma de cuadrados entre grupos
   - Calculamos la variabilidad de cada grupo con respecto a la media general y la ponderamos por el tamaño del grupo.
5. **SSW**: Suma de cuadrados dentro de los grupos
   - Sumamos las variabilidades dentro de cada grupo.
6. **MSB y MSW**:
   - Calculamos las medias de los cuadrados dividiendo las sumas de cuadrados por sus respectivos grados de libertad.
7. **Estadístico F**:
   - Dividimos MSB por MSW para obtener el valor de F.

Finalmente, se utiliza la distribución F para determinar si el valor de F es significativo. Si es así, concluimos que al menos una de las medias de los grupos es diferente.

Ejemplo 1: ANOVA de un Factor

Este ejemplo es una ampliación del ejemplo 1 del t de Student anterior, se amplía el ejemplo agregando un tercer grupo y realizando un análisis ANOVA de un factor. El ANOVA de un factor (ANOVA de una vía) se utiliza para comparar las medias de tres o más grupos independientes para ver si hay una diferencia significativa entre ellas.

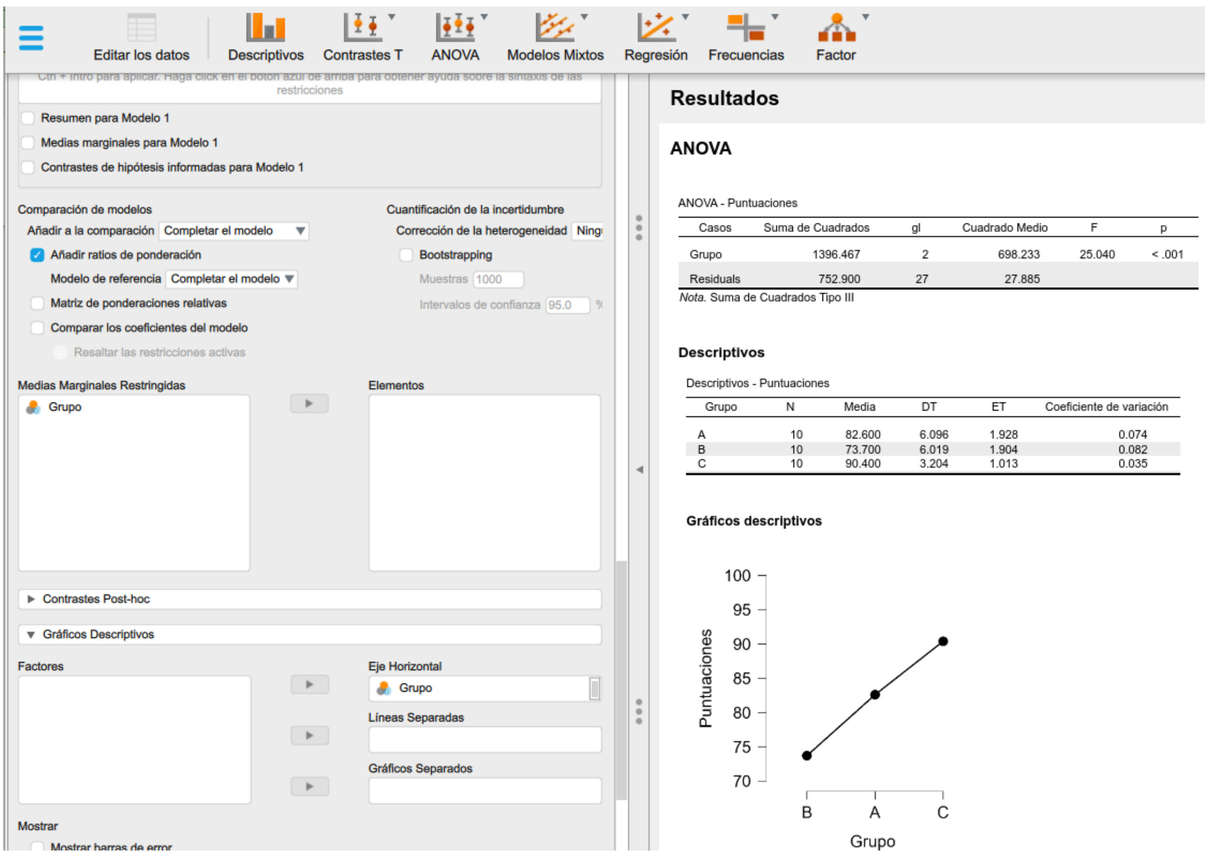
Paso 1: Formular las Hipótesis

- **Hipótesis Nula (H0)**: Las medias de los tres grupos son iguales ($\mu A = \mu B = \mu C$).
- **Hipótesis Alternativa (H1)**: *Al menos una* de las medias de los grupos es diferente.

Paso 2: Implementación en JASP

Para realizar este análisis en JASP, sigue estos pasos:

1. **Abrir JASP** y cargar los datos.
2. **Ingresar los datos**: Asegúrate de que las puntuaciones de los tres grupos estén en una columna y que cada fila tenga una variable indicando el grupo (por ejemplo, "Grupo A", "Grupo B", "Grupo C").
3. **Seleccionar la pestaña "ANOVA"**.
4. **Elegir "ANOVA" (ANOVA de un factor)**.
5. **Arrastrar la variable de puntuaciones a la caja de "Dependent Variable" (Variable dependiente)**.
6. **Arrastrar la variable de grupo a la caja de "Fixed Factors" (Factores fijos)**.
7. **Ejecutar el análisis**.



Paso 3: Conclusión

Dado que el valor p es menor que el nivel de significancia $\alpha=0.05$, rechazamos la hipótesis nula. Por lo tanto, concluimos que hay suficiente evidencia en los datos para decir que hay una diferencia estadísticamente significativa entre las medias de los tres grupos.

Ejemplo 2: Ejemplo: ANOVA de Medidas Repetidas

Supongamos que estamos interesados en evaluar el efecto de tres diferentes métodos de enseñanza sobre el rendimiento de los estudiantes. Medimos el rendimiento de los mismos 10 estudiantes después de utilizar cada método de enseñanza.

Estudiante	Método A	Método B	Método C
1	85	88	90
2	78	85	84
3	90	92	94
4	72	75	78
5	80	83	86
6	88	91	93
7	76	79	80
8	85	87	89
9	90	93	95
10	82	84	87

Paso 1: Formular las Hipótesis

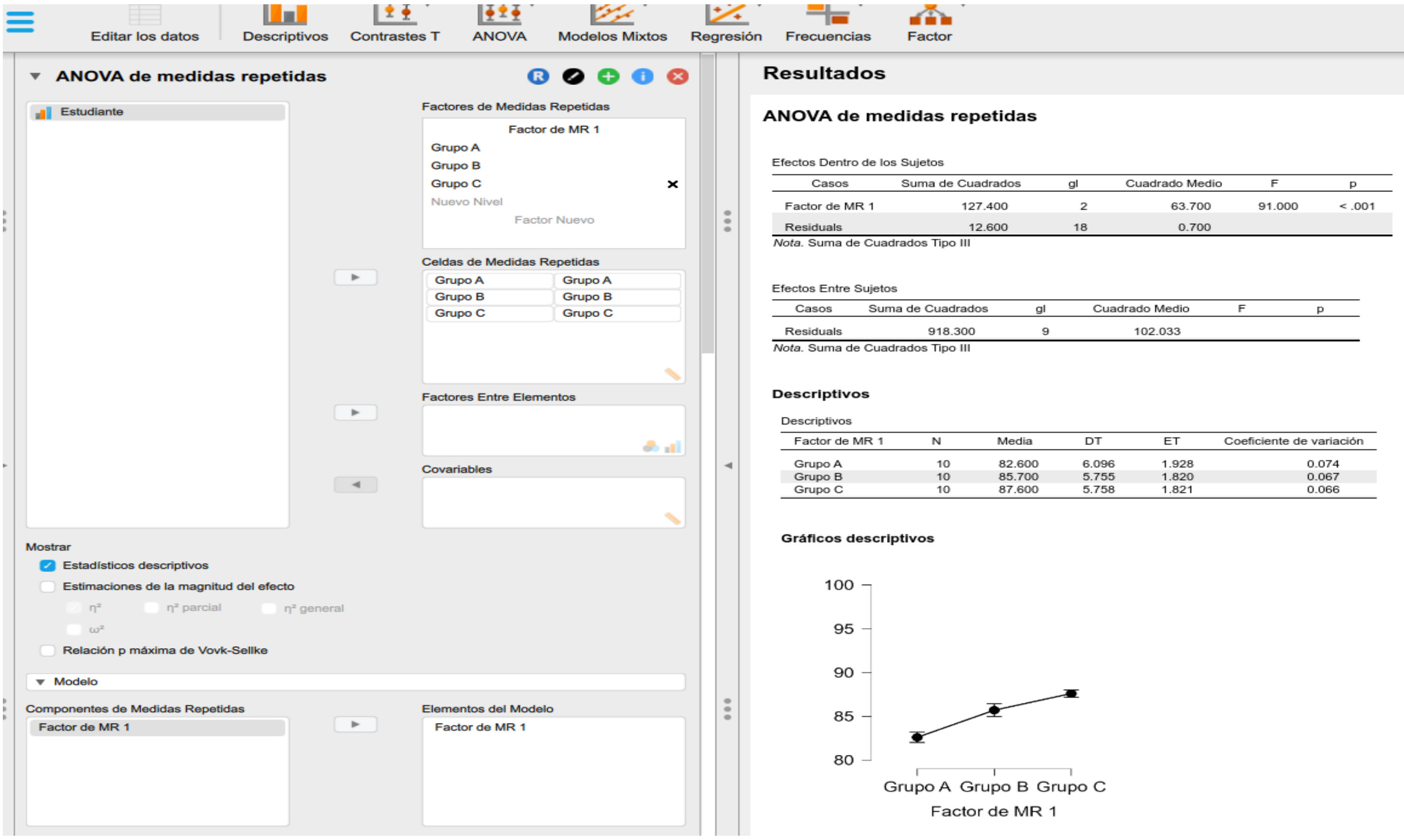
- **Hipótesis Nula (H0)**: No hay diferencias en el rendimiento medio entre los tres métodos de enseñanza ($\mu A=\mu B=\mu C$).
- **Hipótesis Alternativa (H1)**: Al menos uno de los métodos de enseñanza tiene un rendimiento medio diferente.

Paso 2: Implementación en JASP

Para realizar este análisis en JASP, sigue estos pasos:

- **Abrir JASP** y cargar los datos.
- **Ingresar los datos**: Asegúrate de que las puntuaciones para cada método estén en columnas separadas y cada fila representa un sujeto.

- **Seleccionar la pestaña "ANOVA"**.
- **Elegir "Repeated Measures ANOVA" (ANOVA de Medidas Repetidas)**.
- **Definir las condiciones**: Crea una nueva variable en "Repeated Measures Factors" e ingresa las tres condiciones (Método A, Método B, Método C).
- **Arrastrar las variables de cada método a la caja de "Repeated Measures Cells" (Celdas de Medidas Repetidas)**.
- **Ejecutar el análisis**.



Paso 3: Comparar p-valor y conclusión.

El p-valor calculado es inferior a 0,001 por lo tanto muy inferior al nivel de confianza 0,05 (5%), habría pues que rechazar la hipótesis nula y concluir que según los datos de la muestra existe una diferencia estadísticamente significativa entre los distintos métodos de enseñanza.

4. Test F de Snedecor (en este curso solo se estudia en el caso de ANOVA, ver apartado anterior).

- Uso Principal: Comparar varianzas de dos muestras.

- Tipos de Variables: Variables cuantitativas.

- Características de las Muestras:

  - Normalidad de las distribuciones dentro de los grupos.

- Aplicaciones Comunes:

  - Test F para igualdad de varianzas: Determinar si dos muestras tienen varianzas significativamente diferentes.

5. Regresión Lineal

 - **Uso Principal:** Examinar la relación entre una variable dependiente cuantitativa y una o más variables independientes normalmente cuantitativas también.

- **Tipos de Variables:**

  - Variable dependiente: **Cuantitativa**.

  - Variables independientes: **Cuantitativas** o categóricas (**mediante codificación adecuada**).

- **Características de las Muestras:**

  - Relación lineal entre las variables.

  - Homocedasticidad de los residuos.

  - Independencia de los residuos.

  - Normalidad de los residuos.

- **Aplicaciones Comunes:**

  - Predicción y análisis de tendencias.

  - Evaluación del impacto de varias variables independientes en una variable dependiente.

Herramientas a utilizar: Hojas de cálculo, *JASP*

En los estudios de regresión lineal el planteamiento de la hipótesis nula (H0) y alternativa (H1) son ligeramente diferentes y basadas en el valor de la Rxy o

Ejemplo de Regresión Lineal

 Supongamos que estamos interesados en predecir el peso (en kilogramos) de una población basado en su altura (en centímetros). Recogemos datos de una muestra de 10 personas.

Ver la hoja de cálculo “Tablas Estadísticas”.

<u>**Consideraciones Generales**</u>

- Tamaño de la Muestra: Tests como el test t y ANOVA son robustos a desviaciones de normalidad en muestras grandes, pero en muestras pequeñas, la normalidad es crucial.

- Distribución de los Datos: La normalidad de los datos es un supuesto clave en muchos tests paramétricos. Si los datos no cumplen este supuesto, se pueden usar-tests <u>no paramétricos </u>como el test de Mann-Whitney U en lugar del test t, o el test de Kruskal-Wallis en lugar de ANOVA.

- Homogeneidad de Varianzas: Supuesto importante para el test t de dos muestras independientes y ANOVA. Si no se cumple, se pueden utilizar versiones robustas de estos tests, como el test t de Welch.

- **Tipo de Variables:** La naturaleza de las variables (categóricas vs. cuantitativas) guía la selección del test adecuado.

En resumen, la elección del estadístico adecuado depende de la naturaleza de las variables involucradas, la distribución de los datos y las características específicas de las muestras. Una comprensión clara de estos factores asegura que se seleccione el test estadístico más apropiado para el análisis, garantizando la validez y la fiabilidad de los resultados.