Vas a aplicar la estadística inferencial para comparar proporciones.

Y vas a practicar la homogeneidad de proporciones

Y el test de bondad de ajuste

¡Son ejercicios sencillos para afianzar el concepto y que las puedas implementar en tus datos!

Te pongo un ejemplo para que entiendas lo que vas a hacer.

# EJEMPLO – COMPARANDO PROPORCIONES

En este primer ejercicio vamos a trabajar con un ejercicio sencillo. Queremos ver si es hay diferencias significativas con los grupo de LF en el mes 0 según si son del tratamiento control o de investigación.

**PASO 1 - DEFINIENDO EL OBJETIVO – definiendo las hipótesis**

Bien. Se trata de comparar si hay diferencias entre grupos de tratamiento: control o de investigación.

Si queremos ver diferencias entre estos dos grupos el factor de estudio será el tipo de tratamiento (en la tabla de datos se llama Grupo)

Y el factor de análisis (variables respuesta) es el de limitación funcional, LF (moderada, intensa, alta, muy alta).

¿Se ve?

Esto es importante. Sabes distinguir los grupos de estudio y los grupos de las variables respuesta.

* Variables de estudio = Control y de investigación
* Variable respuesta o dependiente = Moderada, Intensa, Alta, muy alta

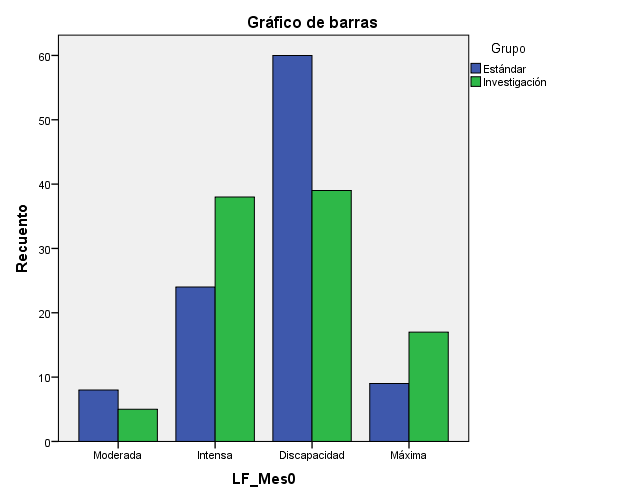
En el fondo estamos analizando si los grupos de control y de investigación están compensados en el mes 0.

**PASO 2 - DESCRIBIENDO LA PREGUNTA**

Para describir este caso vamos a plantear una tabla de contingencias con las frecuencias condicionadas por columnas.

La tabla de datos es: “LF\_mes0” vs “Grupo”

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Tabla cruzada LF\_Mes0\*Grupo** | | | | | |
|  | | | Grupo | | Total |
| Estándar | Investigación |
| LF\_Mes0 | Moderada | Recuento | 8 | 5 | 13 |
| Recuento esperado | 6,6 | 6,4 | 13,0 |
| % dentro de Grupo | 7,9% | 5,1% | 6,5% |
| Intensa | Recuento | 24 | 38 | 62 |
| Recuento esperado | 31,3 | 30,7 | 62,0 |
| % dentro de Grupo | 23,8% | 38,4% | 31,0% |
| Discapacidad | Recuento | 60 | 39 | 99 |
| Recuento esperado | 50,0 | 49,0 | 99,0 |
| % dentro de Grupo | 59,4% | 39,4% | 49,5% |
| Máxima | Recuento | 9 | 17 | 26 |
| Recuento esperado | 13,1 | 12,9 | 26,0 |
| % dentro de Grupo | 8,9% | 17,2% | 13,0% |
| Total | | Recuento | 101 | 99 | 200 |
| Recuento esperado | 101,0 | 99,0 | 200,0 |
| % dentro de Grupo | 100,0% | 100,0% | 100,0% |



**PASO 3 - ANÁLISIS / CONTRASTE DE HIPÓTESIS**

Para analizar este problema vamos a trabajar con la ficha del contraste:

|  |  |  |  |
| --- | --- | --- | --- |
| **Pruebas de chi-cuadrado** | | | |
|  | Valor | gl | Significación asintótica (bilateral) |
| Chi-cuadrado de Pearson | 10,751a | 3 | ,013 |
| Razón de verosimilitud | 10,858 | 3 | ,013 |
| Asociación lineal por lineal | ,003 | 1 | ,955 |
| N de casos válidos | 200 |  |  |
| a. 0 casillas (0,0%) han esperado un recuento menor que 5. El recuento mínimo esperado es 6,44. | | | |

|  |
| --- |
| **Define la H1 o hipótesis de investigación:**  H1: hay diferencias significativas entre las proporciones de los grupos de tratamiento |
| **Define la H0 o hipótesis nula:**  H0: NO hay diferencias significativas entre las proporciones de los grupos de tratamiento |
| **Umbral de contraste (alpha):**  5% = 0.05 |
| **Test Estadístico (selección del test):**  Chi Cuadrado de Pearson  **Comprueba las restricciones:**  Todas las casillas tienen valores esperados > 5 y son grupos independientes 🡪 Chi Cuadrado |
| **Significación o p-valor:**  p-valor = 0.013 < 0.05 🡪 Hay diferencias significativas entre las proporciones por grupos de tratamiento. |
| **Respuesta:**  El p-valor < 0.05 hay evidencias que las distirbuciones entre los grupos control e investigación son diferentes |

**PASO 4 - CONCLUSIÓN**

Juntamos la información del test y el de la descripción de datos.

|  |
| --- |
| Según lo que vemos en el test estadístico hay diferencias significativas entre los grupos de LF antes del tratamiento en los grupos de investigación y control.  Nos hace pensar que la investigación no está bien calibrada con los rangos de LF que estamos imponiendo.    En cambio cuando comparamos cuantitativamente las variables ODI no hay diferencias entre los dos grupos.  Eso demuestra que el pasar una variable cuantitativa a ordinal a veces no es una buena práctica. Si tenemos una variable cuantitativa mejor dejarla así ya que es una medición más exacta que una cualitativa. |

# REPLICA EL EJEMPLO QUE TE ACABO DE PONER

Replica el ejemplo con la limitación funcional en el mes 1, “LF\_mes1” en lugar de LF mes 0.

|  |
| --- |
| Copia el ejercicio siguiendo los pasos que te he puesto antes  TABLAS DE CONTINGENCIA (VALORES OBSERVADOS, ESPERADOS Y RESIDUOS)      A priori cuando se observa el grafico de barras se puede apreciar que las observaciones de los grupos de tratamiento y el tipo de lesión no están bien proporcionados.  Esto lo comprobaremos con el test de chi cuadrado    Tal como se observa, tanto el p-valor del test chi cuadardo como el de Fisher dan un p-valor < 5% por lo que es claramente que hay claras diferencias en las proporciones.  Para saber cuales son las proporciones o grupos de cada una de las filas y su p-valor, se realiza un test de proporciones con la función prop.test de R fila por fila    p-valores  MINIMA 4.25 e-07  MODERADA. 5.55 e-09  INTENSA. 4.405 e-09  DISCAPACIDA 3.81 e-07  MAXIMA. 0.4861. 🡺 este es el único grupo que parece un poco mejor proporcionado. El resto definitivamente plantea que no están bien proporcionados  P-valor del test de chi cuadrado = 2.2 e-16 |

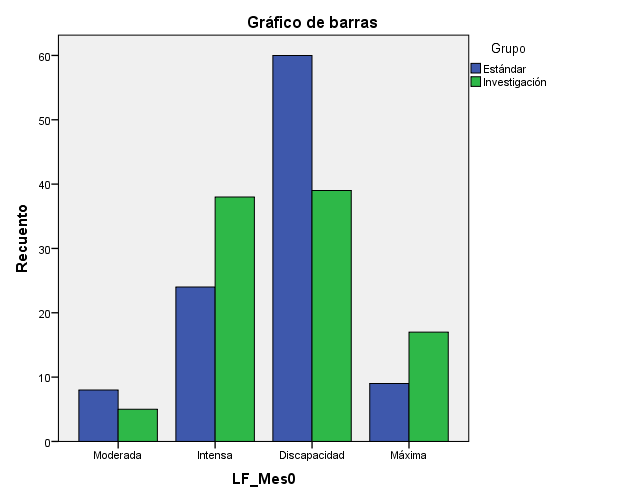
# COMPARANDO LA HOMOGENEIDAD DE PROPORCIONES

Ahora utiliza la bondad de ajuste para aportar evidencias estadísticas de que la muestras de LF\_mes0 está bien repartida en los grupos de control y de investigación.

Hazlo con SPSS que es más cómodo.

Rellena los p-valores de la tabla siguiente:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grupo de LF\_mes0 | N de Control  *(el número de pacientes de cada grupo)* | N de investigación  *(el número de pacientes de cada grupo)* | Diferencia  *Ncontrol - NInvestigacion* | p-valor |
| Mínima | ¿??? | ¿??? |  |  |
| Moderada | ¿??? | ¿??? |  |  |
| Intensa |  |  |  |  |
| Discapacidad |  |  |  |  |
| Máxima |  |  |  |  |



Según los valores que has calculado qué puedes afirmar:

|  |
| --- |
| La verdad no tengo idea de lo que había que hacer.. No he captado en nada esta sección. Podriamos tocarlo en la siguiente tutoria. ¿ |

¡Has hecho un gran trabajo!

Entiendes la filosofía de la comparación de proporciones ☺

El análisis de variables cualitativas es un mundo cuando entras en él.

¡En los Modelos Lineales Generalizados trabajaremos más análisis con proporciones!

Seguimos ;)