Prueba de McNemar

La prueba de McNemar se utiliza para decidir si puede o no aceptarse que determinado "tratamiento" induce un cambio en la respuesta de los elementos sometidos al mismo, y es aplicable a los diseños del tipo "antes-después" en los que cada elemento actúa como su propio control.

Consiste en n observaciones de una v.a. bidimencional (X_i,Y_i) para i=1,..,n.

La escala de medición para X e Y es nominal con 2 categorías tales como positivo o negativo ó hembra, macho ó presencia, ausencia, que se puede denominar “0” y “1”.

En la dócima de McNemar, los resultados se presentan en una tabla 2x2 en la siguiente forma:

		Clasificación Y_i	
		(+) $Y_i = 0$	(-) $Y_i = 1$
Clasificación X_i	(+) $X_i = 0$	**A** (0,0)	**B** (0,1)
	(-) $X_i = 1$	**C** (1,0)	**D** (1,1)

En el cuadro de entrada de frecuencias se presenta al primero y al segundo conjunto de respuestas de los mismos individuos.

En la tabla se usan + y – para simbolizar respuestas diferentes.

		Clasificación Y_i	
		(+)	(-)
Clasificación X_i	(+)	**A**	**B**
	(-)	**C**	**D**

Los casos que muestran cambios entre la primera y segunda respuesta aparecen en las celdillas B y C. Un individuo es clasificado en la celdilla B si cambió de + a - . Es clasificado en la celdilla C si cambió de – a +. Si no es observado ningún cambio, va a la celdilla A (respuestas de + antes y después) o a la celdilla D (respuestas de – antes y después).

Hipótesis:

H_0 :el "tratamiento" no induce cambios significativos en las respuestas, es decir, los cambios observados en la muestra se deben al azar

H_1 : el "tratamiento" induce cambios significativos en las respuestas, es decir, los cambios observados en la muestra no se deben al azar

En esta prueba para la significación de cambios solamente nos va a interesar conocer las celdas que presentan cambios son “B” y “C” y puesto que B+C es el número de personas que cambiaron, se espera de acuerdo a la hipótesis planteada H_0 que (B+C)/2 casos cambien en una dirección y (B+C)/2 casos a otra dirección.

Sea : B+C=n

Estadístico de prueba:

1) Si n < 20 la estadística de contraste es:

$$T_1 = B$$

Regla de Decisión:

Para esta estadística de contraste, rechazamos H_0, si:

$$T_1 \geq \chi^2_{\alpha/2(1 g.l.)}$$

$$o$$

$$T_1 \leq \chi^2_{1-\alpha/2(1 g.l.)}$$

Sea : B+C=n

2) Si n ≥ 20 entonces la estadística de contraste es:

$$T_1 = \frac{(B - C)^2}{B + C} \approx \chi^2_{(1)}$$

Corrección de continuidad: la aproximación muestral a la distribución chicuadrada llega a ser muy buena si se hace una corrección por continuidad y esto debido a que se usa una distribución continua para aproximar una distribución discreta. Cuando las frecuencias esperadas son pequeñas esa aproximación puede ser muy pobre, por lo que aplicamos la corrección "Yates", siendo el estadístico corregido:

$$T_1 = \frac{(|B - C| - 1)^2}{B + C} \approx \chi^2_{(1)}$$

Regla de Decisión:

Para esta estadística de contraste, rechazamos H_0, si:

$$T_1 \geq \chi^2_{\alpha/2\,(1\,g.l.)}$$

$$o$$

$$T_1 \leq \chi^2_{1-\alpha/2\,(1\,g.l.)}$$

Aplicación:

En un debate televisivo a nivel nacional entre 2 candidatos presidenciales, se toma una muestra de 100 personas a quienes se les pregunta sobre sus candidatos antes del debate y 84 resultaron convencidos por el candidato demócrata y 16 al candidato republicano. Después del debate las mismas 100 personas son vueltas a preguntar acerca de sus preferencias, y del grupo de personas que favorecieron al demócrata cambian exactamente ¼ de ellos su manera de pensar y también ¼ de los republicanos.

Solución:

Planteamos las hipótesis

H_0 : Los votantes no han cambiado su manera de pensar

H_1: Los votantes han cambiado su manera de pensar debido al debate

		Después del debate		
		Demócratas	Republicanos	
Antes del debate	Demócratas	A 63	B 21	84
	Republicanos	C 4	D 12	16
				100

Calculamos el estadístico de prueba:

$$T_1 = \frac{(|B-C|-1)^2}{B+C} = \frac{(21-4-1)^2}{25} = 10.24$$

$$\chi^2_{\alpha/2=0.25,(1g.l)} \approx 4.6$$

Por lo tanto H_0 es rechazada.

y concluimos con un riesgo del 5% que existen evidencias suficientes para afirmar que los votantes han cambiado su manera de pensar debido al debate.

La Prueba de Rangos Señalados y Pares Igualados de Wilcoxon para Dos Muestras Relacionadas

Sean $(X_1,Y_1),....,(X_n,Y_n)$ observaciones independientes.

Donde X_i es el resultado correspondiente al control, e Y_i el correspondiente al tratamiento. Sea:

$$D_i=X_i-Y_i \qquad \text{Para todo } i=1,\dots,n$$

La prueba de Wilcoxon para datos apareados o relacionados consiste en la aplicación de la prueba a las diferencias "D_i", suponiendo que la distribución de las diferencias es simétrica.

Hipótesis a contrastar:

H_0: “las dos muestras provienen de la misma población”
(“no hay diferencia entre las observaciones apareadas o relacionadas”).

H_1: “las dos muestras provienen de población diferentes”
(Existe diferencia significativa entre las observaciones apareadas o relacionadas).

Nivel de significación :α

Para el caso de Wilcoxon para una muestra se incluía la magnitud de la mediana, para el caso de muestras relacionadas o apareadas se incluye la magnitud de las diferencias.

Estadístico de Prueba:

Para calcular el estadístico de prueba, tomaremos las diferencias "$d_i = x_i - y_i$" y ordenarlas de menor a mayor prescindiendo del signo y asignarles rango desde 1, 2, . . . ,n. Los valores iguales a cero se ignoran y si varias "d_i" son iguales se les asigna el rango promedio de los valores empatados.

a)Para N<20, N=número de pares cuyas diferencias son diferentes de cero(0) :

Consideramos el estadístico:

$$T = \text{mín}(T_+, T_-) .$$

Siendo:

T_+ = Suma de los rangos correspondientes a las diferencias positivas.

T_- = Suma de los rangos correspondientes a las diferencias negativas.

Regla de Decisión para N<20:

- **Para una prueba Unilateral**:

$T \leq W_{N,\alpha}$ ⇒ Se rechaza H_0

se puede contrastar también con el p-valor (p):

$p \leq \alpha$ ⇒ Se rechaza H_0, si:

- **Prueba Bilateral**:

$T \leq W_{N,\alpha/2}$ ⇒ Se rechaza H_0, si:

se puede contrastar también con el p-valor (p):

$p \leq \alpha/2$ Se rechaza H_0, si:

<u>Considerar</u>:

- $W_{N,\alpha}$ ó $W_{N,\alpha/2}$: valor crítico de la tabla de Wilcoxon.

- $P[W \leq T]=p$, donde p = probabilidad asociada con la ocurrencia de H_0

b) Para N ≥ 20, se puede aproximar T (bajo la hipótesis nula) a una normal con :

$$\mu_T = \frac{N(N+1)}{4}$$ y

$$Varianza = \sigma_T^2 = \frac{N(N+1)(2N+1)}{24}$$

$$Z = \frac{T - \mu_T}{\sigma_T} \approx N(0,1)$$

N=número de pares cuyas diferencias son diferentes de cero(0) :

<u>Regla de Decisión para N ≥ 20</u> :

i)Para el caso de una cola:

Si: $Z_c \geq |Z_\alpha|$ $\longrightarrow$ rechazamos H_0.

También:

Si p-valor $\leq \alpha$ $\longrightarrow$ rechazamos H_0.

ii)Para el caso de dos colas:

Si : $Z_c \leq -Z_{\alpha/2}$ o $Z_c \geq Z_{\alpha/2}$ $\longrightarrow$ rechazamos H_0.

También:

Si p-valor $\leq \alpha/2$ $\longrightarrow$ rechazamos H_0.

Aplicativo:

Se realizó un estudio comparativo en el cual se evaluó la efectividad de dos métodos, uno tradicional y uno moderno de enseñanza del álgebra. En ese estudio 14 individuos fueron extraídos al azar de la población de interés y se formaron 7 pares en base a su coeficiente intelectual. Los miembros de cada par fueron asignados al azar a uno de los dos métodos de enseñanza, y posteriormente ambos grupos fueron instruidos durante 3 semanas. Todos los estudiantes rindieron el mismo examen al final del periodo de instrucción y los resultados obtenidos fueron los siguientes:

Par	Moderno	Tradicional
1	31	36
2	42	38
3	44	33
4	48	36
5	51	53
6	57	49
7	62	52

Solución:

1) Planteamos las hipótesis:

H_0 : Los dos métodos tienen la misma efectividad, $Me_D = 0$.

H_1 : El método moderno es mucho mas efectivo que el tradicional, $Me_D > 0$.

2) Establecer el nivel de significación que se usará en la prueba: $\alpha = 0.05$

3) Calculo del estadístico de prueba:

Par	Moderno	Tradicional	D_i
1	31	36	-5
2	42	38	4
3	44	33	11
4	48	36	12
5	51	53	-2
6	57	49	8
7	62	52	10

$\|D_i\|$	2	4	5	8	10	11	12
Considerando el signo	-2	4	-5	8	10	11	12
|Rango|	1	2	3	4	5	6	7
Rango con signos	-1	2	-3	4	5	6	7

Luego: $N=7$

$$T = \text{mín}(T_{+}, T_{-}) : \text{estadístico de prueba para } N<25$$

Donde:

$$\left.\begin{array}{l} T_{+}= 2+4+5+6+7=24 \\ \\ T_{-}=1+3=4 \end{array}\right\} T= \text{Mín}(4,24) \longrightarrow T=4$$

4) Regla de decisión : se rechaza H_0, si $T \leq W_{7,0.05}=4$, luego el estadístico de prueba es igual al valor critico, por lo tanto con un nivel de significancia del 5% rechazamos H_0.

Conclusión: podemos afirmar con un riesgo del 5% que el método moderno es mucho más efectivo que el método tradicional.

PRUEBAS PARA DOS MUESTRAS INDEPENDIETES

Prueba de la Probabilidad Exacta de Fisher

El test exacto de Fisher permite analizar si dos variables dicotómicas están asociadas cuando la muestra a estudiar es demasiado pequeña.

Una forma de plantear los resultados, es su disposición en una tabla de contingencia de dos vías.

Si las dos variables que se están considerando son dicotómicas, nos encontraremos con el caso de una tabla 2 x 2

El test exacto de Fisher se basa en evaluar la probabilidad asociada a cada una de las tablas 2 x 2 que se pueden formar manteniendo los mismos totales de filas y columnas que los de la tabla observada.

Cada una de estas probabilidades se obtiene bajo la hipótesis nula de independencia de las dos variables que se están considerando.

Hipótesis a contrastar: :

H_0: Las variables son independientes, no están asociadas

H_1:Las variables no son independientes, están asociadas

Nivel de significación :α

Cálculo del p-valor asociado al estadístico de prueba

Para calcular el estadístico de contraste, se construye en primer lugar la tabla de contingencia de dimensiones 2 x 2 con las frecuencias absolutas observadas, con la notación siguiente:

Tabla de contingencia general para la comparación de dos variables dicotómicas en el caso de grupos independientes.

	Característica A		
Característica B	Presente	Ausente	Total
Presente	a	b	a + b
Ausente	c	d	c + d
Total	a + c	b + d	n

A continuación, se construyen todas las tablas de contingencia 2 x 2 posibles con celdas a', b', c', d', siendo 0 ≤ a' ≤ mín{(a+c),(a+b)} b' = (a+b) –a', c' = (a+c) – a' y d' = (c+d) – c'. A partir de dichas tablas se calcula:

$$p_{a'} = \frac{(a+b)!\,(c+d)!\,(a+c)!\,(b+d)!}{n!\,a'!\,b'!\,c'!\,d'!}$$

El p-valor unilateral-izquierda es p= $\sum_{a' \leq a} p_{a'}$

El p-valor unilateral-derecha es p= $\sum_{a' \geq a} p_{a'}$

El p-valor bilateral resultante es p= $\sum_{p_{a'} \leq p_a} p_{a'}$

p_a=probabilidad de la tabla con los datos observados (tabla original)

a= valor de la casilla a en la tabla original.

Ejemplo: A partir de la tabla (original)

	F1	F2	
C1	4	1	5
C2	16	21	37
	20	22	42

Calcular el valor p correspondiente al Test de Fisher:

Solución:

1º Calculamos la tabla para a=0

	F1	F2	
C1	0	5	5
C2	20	17	37
	20	22	42

$$p_{a_0'} = \frac{5!37!20!22!}{42!0!1!16!21!} = 0{,}0310$$

2ºº Calculamos la tabla para a=1

	F1	F2	
C1	1	4	5
C2	19	18	37
	20	22	42

entonces

$$p_{a_1'} = \frac{5!37!20!22!}{42!1!4!19!18!} = 0{,}1720$$

3º Calculamos la tabla para a=2

	F1	F2	
C1	2	3	5
C2	18	19	37
	20	22	42

entonces

$$p_{a_2'} = \frac{5!37!20!22!}{42!2!3!18!19!} = 0{,}3440$$

4º Calculamos la tabla para a=3

	F1	F2	
C1	3	2	5
C2	17	20	37
	20	22	42

Entonces

$$p_{a_3'} = \frac{5!37!20!22!}{42!3!2!17!20!} = 0,3096$$

5º Calculamos la tabla para a=4

	F1	F2	
C1	4	1	5
C2	16	21	37
	20	22	42

entonces

$$p_{a_4'} = \frac{5!37!20!22!}{42!4!1!16!22!} = 0,125$$

6º Calculamos la tabla para a=5

	F1	F2	
C1	5	0	5
C2	15	22	37
	20	22	42

Entonces

$$p_{a_5'} = \frac{5!37!20!22!}{42!5!0!15!22!} = 0,018$$

Los valores de "p" para cada a'

a'	$p_{a'}$
0	0.0310
1	0.1720
2	0.3440
3	0.3096
4	0.1253
5	0.0182

p_a

El valor p bilateral es p= $\sum_{p_{a'} \leq p_a} p_{a'} = 0{,}1253 + 0{,}0182 + 0{,}0310 = 0{,}1745$

El valor p unil-izq.es p = $\sum_{a' \leq a} p_{a'} = 0{,}0310 + 0{,}1720 + 0{,}3440 + 0{,}3096 + 0{,}1253 = 0{,}9818$

El valor p unil-der.es p= $\sum_{a' \geq a} p_{a'} = 0{,}1253 + 0{,}0182 = 0{,}1435$

p = probabilidad asociada con la ocurrencia de H_0

Regla de Decisión:

- Prueba Unilateral:

  Se rechaza H_0, si :

$$p \leq \alpha$$

- Prueba Bilateral:

  Se rechaza H_0, si :

$$p \leq \alpha/2$$

p = probabilidad asociada con la ocurrencia de H_0

Aplicativo:

En una determinada población se desea averiguar si existen diferencias en la prevalecía de obesidad entre hombres y mujeres o si, por el contrario, el porcentaje de obesos no varía entre sexos. Tras ser observada una muestra de 14 sujetos se obtuvieron los resultados que se muestran en la siguiente tabla:

Tabla de contingencia para estudiar las diferencias en la prevalecía de obesidad entre sexos. Estudio de prevalecía sobre 14 sujetos.

	Obesidad		
Sexo	Sí	No	**Total**
Mujeres	1 (a)	4 (b)	5 (a+b)
Hombres	7 (c)	2 (d)	9 (c+d)
Total	8 (a+c)	6 (b+d)	14 (n)

Planteamos las hipótesis:

H_0: La prevalecía de obesidad es igual en hombres y mujeres, las variables sexo y obesidad son independientes.

H_1: Existen diferencias en la prevalecía de obesidad entre sexos, las variables sexo y obesidad no son independientes.

Establecemos un $\alpha=0.05$

Calculamos el estadístico de prueba

Posibles combinaciones de frecuencias con los mismos totales marginales de filas y columnas

(i)

	Obesidad Si	Obesidad No	
Mujeres	0	5	5
Hombres	8	1	9
	8	6	14

p_a

(ii)

	Obesidad Si	Obesidad No	
Mujeres	1	4	5
Hombres	7	2	9
	8	6	14

(iii)

	Obesidad Si	Obesidad No	
Mujeres	2	3	5
Hombres	6	3	9
	8	6	14

(iv)

	Obesidad Si	Obesidad No	
Mujeres	3	2	5
Hombres	5	4	9
	8	6	14

(v)

	Obesidad Si	Obesidad No	
Mujeres	4	1	5
Hombres	4	5	9
	8	6	14

(vi)

	Obesidad Si	Obesidad No	
Mujeres	5	0	5
Hombres	3	6	9
	8	6	14

calculado la probabilidad exacta de ocurrencia bajo la hipótesis nula, según:

$$p_{a'} = \frac{(a+b)!\,(c+d)!\,(a+c)!\,(b+d)!}{n!\,a'!\,b'!\,c'!\,d'!}$$

Entonces calculando tenemos:

Probabilidad exacta asociada con cada una de las disposiciones de frecuencias

	a	b	c	d	p
(i)	0	5	8	1	0,0030
(ii)	1	4	7	2	0,0599
(iii)	2	3	6	3	0,2797
(iv)	3	2	5	4	0,4196
(v)	4	1	4	5	0,2098
(vi)	5	0	3	6	0,0280

p_a

Cálculo del p-valor asociado al estadístico de prueba

Calculamos el p-valor para una prueba bilateral:

El valor p bilateral es p= $\sum_{p_{a'} \leq p_a} p_{a'} = 0{,}003 + 0{,}028 + 0{,}0599 = 0{,}0909$

Regla de Decisión:

Prueba Bilateral:
Se rechaza H_0, si p ≤ α/2=0.025, de los datos p=0.0909>0.025, por lo tanto con un nivel de significancía del 5% no rechazamos H_0.

Conclusión:
Con un nivel de significancia del 5% podemos concluir que la prevalecía de obesidad es igual en hombres y mujeres en la población de estudio.