

Ama de Casa	Grupo A	Grupo B
1	2	4
2	3	3
3	3	2
4	4	2
5	2	1
6	4	3
7	3	2
8	5	3
9	4	4
10	2	2
11	3	2
12	4	2

[illegible]

Una vez construida la matriz, se llenará cada celda interior de ella respecto a las siguientes condiciones:

- Si el número de la categoría B es mayor al de la categoría A, entonces la celda tendrá el valor 1.
- Si el número de la categoría B es igual al de la categoría A, entonces la celda tendrá el valor 0.5.
- Si el número de la categoría B es menor al de la categoría A, entonces la celda tendrá el valor 0.

Se obtiene la siguiente tabla:

		1	2	3	4	5	6	7	8	9	10	11	12
B/A		2	3	3	4	2	4	3	5	4	2	3	4
1	4	1	1	1	0.5	1	0.5	1	0	0.5	1	1	0.5
2	3	1	0.5	0.5	0	1	0	0.5	0	0	1	0.5	0
3	2	0.5	0	0	0	0.5	0	0	0	0	0.5	0	0
4	2	0.5	0	0	0	0.5	0	0	0	0	0.5	0	0
5	1	0	0	0	0	0	0	0	0	0	0	0	0
6	3	1	0.5	0.5	0	1	0	0.5	0	0	1	0.5	0
7	2	0.5	0	0	0	0.5	0	0	0	0	0.5	0	0
8	3	1	0.5	0.5	0	1	0	0.5	0	0	1	0.5	0
9	4	1	1	1	0.5	1	0.5	1	0	0.5	1	1	0.5
10	2	0.5	0	0	0	0.5	0	0	0	0	0.5	0	0
11	2	0.5	0	0	0	0.5	0	0	0	0	0.5	0	0
12	2	0.5	0	0	0	0.5	0	0	0	0	0.5	0	0

Se define la variable aleatoria U_x de la siguiente manera:

$$U_x = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} w_{ij}$$

Esto es, en pocas palabras, sumar todos los valores internos de la matriz y en donde n_1 y n_2 son el número de datos de ambas categorías. Su esperanza y varianza tienen por valores:

$$E(U_x) = \frac{n_1 n_2}{2} \quad V(U_x) = \frac{1}{12} n_1 n_2 (n_1 + n_2 + 1)$$

Además, también hace falta el calculo de U_y , que está definida:

$$U_y = n_1 n_2 - U_x$$

Dicho esto, se proceden a calcular cada uno de estos estadísticos.

n1	12
n2	12
n1*n2	144
Ux	42
Uy	102
E(U) =	72
V(U)=	300
EE(U)=	17.320508

Como puede apreciarse, la variable Ux es menor que la Uy y por lo tanto es esta la que se ocupa para calcular el estadístico Z de prueba. De haber sido lo contrario, Uy estaría en su lugar. **EE(U)** no es más que la raíz cuadrada de la varianza.

ESTADÍSTICO Z:

Z =	-1.732051
------------	------------------

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

Consecuentemente, calculando con una función cuantil el P-value:

P value	0.0416323
P value	0.0832645

Como la prueba es de una cola, se puede concluir que: bajo un nivel de confianza del 95% y utilizando la prueba **U de Mann y Whitney**, se tiene prueba estadística (a mi parecer no tan suficiente) para **rechazar la hipótesis nula**, pues el P-Value es menor a .05. Por lo tanto, el detergente A es mejor (más preferido) que el detergente B.

B) Utilice la prueba T para dos grupos independientes y suponga varianzas iguales.

Se procede a hacer los cálculos de la media, desviación estándar y el número de registros de ambas columnas.

Ama de Casa	Grupo A	Grupo B
1	2	4
2	3	3
3	3	2
4	4	2
5	2	1
6	4	3
7	3	2
8	5	3
9	4	4
10	2	2
11	3	2
12	4	2
Media	3.25	2.5
S ²	0.932	0.8182
n i	12	12

Se define la siguiente variable, junto con los siguientes resultados:

$$S_p = \sqrt{\frac{(n_1 - 1)S_x^2 + (n_2 - 1)S_y^2}{n_1 + n_2 - 2}}$$

S_p^2 0.875
 S_p 0.935

En donde **Sp²** es el cuadrado de Sp. Esta última es necesaria para el cálculo del estadístico de prueba T, cuyo valor obtenido es:

ESTADÍSTICO T

$$T = \frac{(X - Y) - (\mu_x - \mu_y)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightarrow t_{n_1 + n_2 - 2}$$

T 1.964

Con ayuda de una función cuantil para la T cola derecha se obtiene:

P value una cola

0.031

P value una dos colas

0.062

Con lo cual puede concluirse con un nivel del 95% de confianza, que se **rechaza la hipótesis nula** pues el P-Valor es menor que .05 (alfa). Se llegó bajo la prueba no paramétrica como la paramétrica a la misma conclusión: **Se prefiere más el detergente A que el detergente B.**

Extra) Calculo de Excel

Se puede utilizar el complemento de Excel para realizar la misma prueba cuyos resultados son los mismos para la prueba T.

t-Test: Two-Sample Assuming Equal Variances

	<i>Variable 1</i>	<i>Variable 2</i>
Mean	3.25	2.5
Variance	0.93181818	0.818181818
Observations	12	12
Pooled Variance	0.875	
Hypothesized Mean Difference	0	
df	22	
t Stat	1.96396101	
P(T<=t) one-tail	0.03114705	
t Critical one-tail	1.71714437	
P(T<=t) two-tail	0.0622941	
t Critical two-tail	2.07387307	