

Bioestadística y uso de software científico



TEMA 5 DATOS CONTINUOS COMPARACIONES DE MEDIAS ENTRE DOS GRUPOS

Hasta ahora...



Tema	Variable dependiente	Variable independiente	Test
Tema 4	Categórica	Categórica	χ^2 , McNemar
Tema 5	Continua	Dicotómica	t de Student U de Mann-Whitney

Índice



- t de Student para una muestra
- Dos muestras, datos independientes
 - t de Student
 - U de Mann-Whitney
- Dos muestras, datos emparejados
 - t de Student
 - Test de Wilcoxon

T de Student para comparar una media con un valor de referencia



- En una muestra de 20 personas, la tensión arterial sistólica tiene $m = 120$ y $s = 15$.
- ¿Se puede asumir que la muestra procede de una población con $\mu = 110$?

T de Student para comparar una media con un valor de referencia



- $H_0: \mu = \mu_0$

- Calcular:

$$t = \frac{\bar{m} - \mu_0}{s / \sqrt{n}}$$

- Buscar t en la tabla t de Student con $n-1$ grados de libertad

T de Student para comparar una media con un valor de referencia



- $H_0: \mu=110$

- Calcular:

$$t = \frac{m - \mu_0}{s / \sqrt{n}} = \frac{120 - 110}{15 / \sqrt{20}} = 2,98$$

- Buscar $t=2,98$ en la tabla t de Student con $n-1=19$ grados de libertad:

- $p = 0,004$

T de Student para comparar una media con un valor de referencia



$$t = \frac{m - \mu_0}{s / \sqrt{n}}$$

- Tiene una forma típica:
 - Diferencia de medias

$$dm = m - \mu_0$$

- Error estándar de la media (EEM)

$$EEM = s / \sqrt{n}$$

$$t = \frac{dm}{EEM}$$

Índice



- t de Student para una muestra
- Dos muestras, datos independientes
 - t de Student
 - U de Mann-Whitney
- Dos muestras, datos emparejados
 - t de Student
 - Test de Wilcoxon

T de Student para dos muestras independientes



- 30 pacientes hipertensos se dividen al azar en dos grupos. 17 tomarán el medicamento A y 13 el medicamento B.
- Los del medicamento A bajan su tensión arterial una media de 5 mm Hg, desviación estándar = 3
- Los del medicamento B bajan la tensión 8 mm Hg, desviación estándar = 4
- ¿Es mejor un medicamento que el otro o los dos son iguales?

T de Student para dos muestras independientes



- $H_0: \mu_A = \mu_B; H_1: \mu_A \neq \mu_B$
- Muestra A: $m_A=5, s_A=3, n_A=17$
- Muestra B: $m_B=8, s_B=4, n_B=13$

T de Student para dos muestras independientes



1. Calcular la varianza combinada

$$s_p^2 = \frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{(n_A - 1) + (n_B - 1)}$$

2. Calcular el error estándar de la diferencia de medias

$$EEDM = s_p \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}$$

3. Calcular $t = \frac{dm}{EEDM} = \frac{m_A - m_B}{s_p \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \rightarrow t_{n_A + n_B - 2}$

T de Student para dos muestras independientes



1. Calcular la varianza combinada

$$s_p^2 = \frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{(n_A - 1) + (n_B - 1)} = \frac{(17 - 1) \times 9 + (13 - 1) \times 16}{(17 - 1) + (13 - 1)} = 12$$

2. Calcular el error estándar de la diferencia de medias

$$EEDM = s_p \sqrt{\frac{1}{n_A} + \frac{1}{n_B}} = 3,464 \sqrt{\frac{1}{17} + \frac{1}{13}} = 1,276$$

3. Calcular

$$t = \frac{m_A - m_B}{s_p \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} = \frac{3 - 4}{1,276} = -0,784 \rightarrow t_{28g.l.} \rightarrow p = 0,78$$

T de Student para dos muestras independientes



- Condiciones para aplicar el test t de Student:
 - Distribuciones normales o $n > 30$ en cada grupo
 - Varianzas homogéneas (homocedasticidad)

$$(\sigma_A^2 = \sigma_B^2)$$

T de Student para dos muestras independientes



- Cómo comprobar que las varianzas son homogéneas

$$H_0 : \sigma_A^2 = \sigma_B^2; H_1 : \sigma_A^2 \neq \sigma_B^2$$

1. Calcular

$$F = \frac{\text{varianza}_{mayor}}{\text{varianza}_{menor}}$$

2. Consultar las tablas de la F de Snedecor

T de Student para dos muestras independientes



$$s_A^2 = 9; s_B^2 = 16; n_A = 17; n_B = 13$$

1. Calcular

$$F = \frac{\text{varianza}_{\text{mayor}}}{\text{varianza}_{\text{menor}}} = \frac{16}{9} = 1,78$$

2. Consultar las tablas de la F de Snedecor con $n_B - 1$, $n_A - 1$ grados de libertad:

$$F_{12,16} = 1,78 \rightarrow p = 0,14$$

T de Student para dos muestras independientes



- Test de Welch: Cómo aplicar t de Student si las varianzas son heterogéneas (heterocedasticidad)

1. Calcular

$$t = \frac{m_A - m_B}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$$

2. Calcular los grados de libertad modificados (gl*)

T de Student para dos muestras independientes



- Test de Welch: Cómo aplicar t de Student si las varianzas son heterogéneas (heterocedasticidad)
2. Calcular los grados de libertad modificados (gl^*)

$$gl^* = \frac{\left(\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B} \right)^2}{\frac{\left(\frac{s_A^2}{n_A} \right)^2}{n_A - 1} + \frac{\left(\frac{s_B^2}{n_B} \right)^2}{n_B - 1}}$$

Intervalo de confianza de la diferencia de medias



$$IC 1 - \alpha : dm \pm t_{\alpha/2, n_A + n_B - 2} \times EEDM$$

$$IC 95\% : dm \pm t_{0,025; n_A + n_B - 2} \times EEDM$$

En el ejemplo:

$$IC 95\% : -1 \pm 2,05 \times 1,276 = (-3,62 \text{ a } 1,62)$$

Índice



- t de Student para una muestra
- Dos muestras, datos independientes
 - t de Student
 - U de Mann-Whitney
- Dos muestras, datos emparejados
 - t de Student
 - Test de Wilcoxon

Test de la U de Mann-Whitney



- Prueba no paramétrica
- Usar en lugar de la t de Student si:
 - $n < 30$ en alguna de las muestras y la distribución NO es normal
 - O la variable es ordinal (en lugar de cuantitativa)
 - O la muestra es muy pequeña ($n < 10$ en alguno de los grupos)

Test de la U de Mann-Whitney



- Inconvenientes
 - Menos potente que la t de Student
 - ✦ (Potencia = capacidad para encontrar diferencias significativas = probabilidad de rechazar H_0 cuando es falsa)
 - No se pueden calcular intervalos de confianza

Test de la U de Mann-Whitney



- 12 alumnos (5 varones y 7 mujeres) quieren saber si las diferencias que tienen en las notas se deben al azar

Mujeres	Varones
9	8
8	8
7	7
7	6
6	5
6	
5	

Test de la U de Mann-Whitney



1. Comparar los alumnos 1 a 1 y contar:
 - A. Cuántas veces una mujer tiene mejor nota que un varón ($A = 16$)
 - B. Cuántas veces un varón tiene mejor nota que una mujer ($B = 13$)
 - C. Cuántos empates (6)

Mujeres	Varones
9	8
8	8
7	7
7	5
6	5
6	
5	

Test de la U de Mann-Whitney



2. Calcular

A. $U_1 = A + C/2 = 16 + 6/2 = 19$

B. $U_2 = B + C/2 = 13 + 6/2 = 16$

Mujeres	Varones
9	8
8	8
7	7
7	5
6	5
6	
5	

Test de la U de Mann-Whitney



3. Calcular

$$z = \frac{U_1 - (n_1 n_2 / 2)}{\sqrt{n_1 n_2 (n_1 + n_2 + 1) / 12}} = \frac{19 - (7 \times 5 / 2)}{\sqrt{7 \times 5 \times (7 + 5 + 1) / 12}} = 0,244$$

4. Mirar en una tabla de la distribución normal

$$P(z > 0,244) = 0,404 \text{ (test de una cola)}$$

$$P(|z| > 0,244) = 0,808 \text{ (test de dos colas)}$$

Índice



- t de Student para una muestra
- Dos muestras, datos independientes
 - t de Student
 - U de Mann-Whitney
- Dos muestras, datos emparejados
 - t de Student
 - Test de Wilcoxon

T de Student para datos emparejados



- Muestras independientes:
 - ¿La tensión arterial es igual en los que toman el medicamento A que en los que toman el medicamento B?
 - Se compara una variable en distintos sujetos
- Muestras emparejadas:
 - ¿La tensión arterial es igual antes de tomar el medicamento A que después de tomarlo?
 - Se comparan dos variables en los mismos sujetos

T de Student para datos emparejados



Paciente n°	TA antes	TA después
1	145	130
2	150	140
3	130	140
4	160	150
5	165	145
6	160	145
7	140	140

T de Student para datos emparejados



1. Calcular la diferencia entre antes y después

Paciente nº	TA antes	TA después	Diferen cia
1	145	130	-15
2	150	140	-10
3	130	140	10
4	160	150	-10
5	165	145	-20
6	160	145	-15
7	140	140	0

T de Student para datos emparejados



2. Calcular la media y el error estándar de la diferencia

$$m_d = -8,57$$

$$s_d = 10,29$$

$$EEM_d = \frac{s_d}{\sqrt{n}} = 3,89$$

Paciente nº	TA antes	TA después	Diferen cia
1	145	130	-15
2	150	140	-10
3	130	140	10
4	160	150	-10
5	165	145	-20
6	160	145	-15
7	140	140	0

T de Student para datos emparejados



3. Calcular la t de Student y consultar la tabla t con n-1 grados de libertad

$$t_{n-1} = \frac{m_d}{EEM_d} = \frac{-8,57}{3,89} = -2,20$$

$$p = 0,035 \text{ (1 cola)}$$

$$p = 0,070 \text{ (2 colas)}$$

T de Student para datos emparejados



4. Calcular el intervalo de confianza

$$IC 1 - \alpha : m_d \pm t_{n-1, \alpha/2} \times EEM_d$$

$$IC 95\% : m_d \pm t_{n-1; 0,025} \times EEM_d$$

$$IC 95\% : -8,57 \pm t_{6; 0,025} \times 3,89 = -8,57 \pm 2,45 \times 3,89 = \\ = -18,10 \text{ a } 0,96$$

T de Student para datos emparejados



- Condiciones para aplicar el test t de Student:
 - Distribuciones normales o $n > 30$ en cada grupo
 - Varianzas homogéneas (homocedasticidad)

$$(\sigma_A^2 = \sigma_B^2)$$

Índice



- t de Student para una muestra
- Dos muestras, datos independientes
 - t de Student
 - U de Mann-Whitney
- Dos muestras, datos emparejados
 - t de Student
 - Test de Wilcoxon

Test de Wilcoxon para datos emparejados



- Es el equivalente al test U de Mann-Whitney para datos emparejados
- Usar cuando:
 - Los datos son ordinales
 - Los datos son cuantitativos pero $n < 30$ y la variable NO es normal

Test de Wilcoxon para datos emparejados



TA antes	TA después
145	130
150	140
130	140
160	150
165	145
160	145
140	140

Test de Wilcoxon para datos emparejados



2. Calcular el valor absoluto de la diferencia, y ordenar

TA antes	TA después	Dif	Dif
140	140	0	0
150	140	-10	10
130	140	10	10
160	150	-10	10
145	130	-15	15
160	145	-15	15
165	145	-20	20

Test de Wilcoxon para datos emparejados



3. Indicar el puesto (orden) de cada observación

TA antes	TA después	Dif	Dif	Orden
140	140	0	0	-
150	140	-10	10	2
130	140	10	10	2
160	150	-10	10	2
145	130	-15	15	4,5
160	145	-15	15	4,5
165	145	-20	20	6

Test de Wilcoxon para datos emparejados



4. Sumar el orden de los sujetos con dif + y con dif -

TA antes	TA después	Dif	Dif	Orden	R+	R-
140	140	0	0	-		
150	140	-10	10	2		2
130	140	10	10	2	2	
160	150	-10	10	2		2
145	130	-15	15	4,5		4,5
160	145	-15	15	4,5		4,5
165	145	-20	20	6		6
Total					2	19

Test de Wilcoxon para datos emparejados



5. Calcular

$$\begin{aligned} z &= \frac{(R-) - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} = \\ &= \frac{19 - 7 \times 8 / 4}{\sqrt{7 \times 8 \times 15 / 24}} = 0,845 \end{aligned}$$

Test de Wilcoxon para datos emparejados



6. El resultado se mira en la tabla Normal

$$z = 0,845 \rightarrow p = 0,20 \text{ (1 cola)}$$

$$z = 0,845 \rightarrow p = 0,40 \text{ (2 colas)}$$