

CONTRASTES NO PARAMETRICOS

Contraste U de Mann - Withney

Es la prueba no paramétrica paralela a la t de dos grupos independientes. **Hipótesis**

$$\begin{cases} H_0 : \text{Las poblaciones de las que provienen las muestras están equidistribuidas} \\ H_1 : \text{Las poblaciones no están equidistribuidas} \end{cases}$$

El contraste se efectúa combinando las dos muestras y disponiendo el conjunto completo de las observaciones, ordenado de menor a mayor. Se asignan después números de rango a cada observación. Se calcula después la suma de los rangos de las observaciones pertenecientes a la primera muestra y a la segunda, obteniéndose respectivamente R_1 y R_2 , para después calcular los estadísticos

Test

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 = n_1 n_2 - U_1$$

Regla de decisión

Para el **contraste bilateral**, se define:

$$U_{exp} = \min\{U_1, U_2\}$$

y se rechaza H_0 si $U_{exp} < U_{n_1, n_2, \alpha}$

Si el contraste que pretendemos realizar es **unilateral**, como por ejemplo,

$$\begin{cases} H_0 : \text{La primera población toma valores menores o iguales a la segunda} \\ H_1 : \text{Los de la primera son mayores} \sim H_a: \mu_1 > \mu_2 \end{cases}$$

Rechazamos la hipótesis nula si $U_1 < U_{n_1, n_2, \alpha}$

Si el test es el contrario

$$\begin{cases} H_0 : \text{La segunda población toma valores menores o iguales a la primera} \\ H_1 : \text{Los de la segunda son mayores} \sim H_a: \mu_1 < \mu_2 \end{cases}$$

Rechazamos H_0 si $U_2 < U_{n_1, n_2, \alpha}$

Aproximación normal del test de Mann--Withney

Cuando los tamaños de las muestras son grandes, $n_1, n_2 > 10$, no es posible recurrir a las tablas de Mann--Withney. En este caso utilizamos la aproximación normal

$$U_1 \overset{\approx}{\sim} \mathbf{N}(\mu_u, \sigma_u^2)$$

donde

$$\mu_u = \frac{n_1 n_2}{2}$$

$$\sigma_u^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

y se utiliza el estadístico de contraste

$$Z_{exp} = \frac{U_1 - \mu_u}{\sigma_u} \overset{\approx}{\sim} \mathbf{N}(0, 1)$$

rechazándose la equidistribución de ambas poblaciones si el valor experimental pertenece a la región de rechazo, según la alternativa elegida.

Ejemplo

Dos empleados, A y B , trabajan en el departamento de niños de una tienda. El gerente de la tienda piensa ampliar su negocio a otros locales desde que leyó un artículo en una revista sobre la creciente popularidad de las tiendas sobre niños. El gerente registra las ventas semanales de los 2 empleados para una muestra de varias semanas y quiere saber si ellos pueden considerarse iguales como vendedores.

Ventas 1	Ventas 2
197	190
194	180
188	175
185	172
182	167
173	166
169	160
169	157
164	155
166	150
154	146
149	145
142	143
139	140
137	135
130	135
	134
	133

Solución

Hipótesis

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

Test

$$U1 = (16 \times 18) + 16(16+1)/2 - 320.5 \\ = 103.5$$

Media:

$$= (16 \times 18) / 2 = 144$$

Varianza:

$$= (16 \times 18)(16+18+1)/12 = 840$$

$$Z = (103.5 - 144) / \sqrt{840} = -1.397$$

Ventas 1	Rangos	Ventas 2	Rangos
197	34	190	32
194	33	180	28
188	31	175	27
185	30	172	25
182	29	167	22
173	26	166	20.5
169	23.5	160	18
169	23.5	157	17
164	19	155	16
166	20.5	150	14
154	15	146	12
149	13	145	11
142	9	143	10
139	7	140	8
137	6	135	4.5
130	1	135	4.5
		134	3
		133	2
Total	320.5		274.5

Decisión

Como $p\text{-value} = 2P(|Z| > 1.397) = 0.1624 > 0.05 \Rightarrow$ no rechazamos H_0 .

Por tanto, los dos empleados pueden considerarse iguales como vendedores.

Contraste de Wilcoxon

El **contraste de Wilcoxon** es la técnica no paramétrica paralela a el de la de Student para muestras apareadas. Igualmente dispondríamos de n parejas de valores (x_i, y_i) que podemos considerar como una variable medida en cada sujeto en dos momentos diferentes.

$\forall i = 1, \dots, n,$ i -ésima observación $\equiv (x_i, y_i) \rightarrow$ diferencia $\equiv d_i = x_i - y_i$

El procedimiento consiste en:

1. Ordenar las cantidades $|d_i|$ de menor a mayor y obtener sus rangos
2. Se asigna a cada rango el signo de la diferencia que corresponde y luego, se calculan las sumas de rangos para diferencias positivas (T^+) y para diferencias negativas (T^-). (no consideramos las cantidades $d_i=0$).

Hipótesis

H_0 : Las distribuciones son idénticas

H_a : Las distribuciones no son idénticas.

Test

a) Para una prueba de dos colas:

$H_0: \mu_1 = \mu_2$ vs. $H_a: \mu_1 \neq \mu_2 \Rightarrow$ utilice $T = \min (T^+, T^-)$.

b) Para una prueba unilateral hacia la derecha:

$H_0: \mu_1 = \mu_2$ vs. $H_a: \mu_1 > \mu_2 \Rightarrow$ utilice T^- .

c) Para una prueba unilateral hacia la izquierda:

$H_0: \mu_1 = \mu_2$ vs. $H_a: \mu_1 < \mu_2 \Rightarrow$ utilice T^+ .

Regla de decisión

a) Para una prueba bilateral, rechace H_0 , si $T \leq T_0$, en donde T_0 es un valor crítico de la tabla de Wilcoxon.

b) Para una prueba unilateral hacia la derecha, rechace H_0 , si $T^- \leq T_0$.

c) Para una prueba unilateral hacia la izquierda, rechace H_0 , si $T^+ \leq T_0$.

Ejemplo:

Los siguientes datos representan el número de horas que un compensador opera antes de requerir una recarga: 1.5, 2.2, 0.9, 1.3, 2.0, 1.6, 1.8, 1.5, 2.0, 1.2 y 1.7. Utilice la prueba de rango con signo para probar la hipótesis en el nivel de significancia de 0.05 que este compensador particular opera con una media de 1.8 horas antes de requerir una recarga.

Solución

$$H_0: \mu = 1.8$$

$$H_1: \mu \neq 1.8$$

Se procederá a efectuar las diferencias y a poner rango con signo a los datos.

Dato	$d_i = \text{dato} - 1.8$	Rangos
1.5	-0.3	5.5
2.2	0.4	7
0.9	-0.9	10
1.3	-0.5	8
2.0	0.2	3
1.6	-0.2	3
1.8	0	Se anula
1.5	-0.3	5.5
2.0	0.2	3
1.2	-0.6	9
1.7	-0.1	1

Regla de decisión:

Para un $n = 10$, después de descartar la medición que es igual a 1.8, la tabla de Wilcoxon muestra que la región crítica es $w \leq 8$.

Cálculos:

$$w+ = 7 + 3 + 3 = 13$$

$$w- = 5.5 + 10 + 8 + 3 + 5.5 + 9 + 1 = 42$$

por lo que $w = 13$ (menor entre $w+$ y $w-$).

Decisión y Conclusión:

Como 13 no es menor que 8, no se rechaza H_0 y se concluye con $\alpha = 0.05$ que el tiempo promedio de operación no es significativamente diferente de 1.8 horas.

Ejemplo

Se afirma que un estudiante universitario de último año puede aumentar su calificación en el área del campo de especialidad del examen de registro de graduados en al menos 50 puntos si de antemano se le proporcionan problemas de muestra. Para probar esta afirmación, se dividen 20 estudiantes del último año en 10 pares de modo que cada par tenga casi el mismo promedio de puntos de calidad general en sus primeros años en la universidad. Los problemas y respuestas de muestra se proporcionan al azar a un miembro de cada par una semana antes del examen. Se registran las siguientes calificaciones del examen: Pruebe la hipótesis nula en el nivel de significancia de 0.05 de que los problemas aumentan las calificaciones en 50 puntos contra la hipótesis alternativa de que el aumento es menor a 50 puntos.

Par	Con problemas de muestra	Sin problemas de muestra
1	531	509
2	621	540
3	663	688
4	579	502
5	451	424
6	660	683
7	591	568
8	719	748
9	543	530
10	575	524

Solución

$$H_0: \mu_1 - \mu_2 = 50$$

$$H_1: \mu_1 - \mu_2 < 50$$

Test

$$w+ = 6 + 3.5 + 1 = 10.5$$

Regla de decisión:

Para $n=10$ la tabla muestra que la región crítica es $w+ = 11$.

Par	Con problema s de muestra	Sin problemas de muestra	d_i	$d_i - d_0$	Rangos
1	531	509	22	-28	5
2	621	540	81	31	6
3	663	688	-25	-75	9
4	579	502	77	27	3.5
5	451	424	27	-23	2
6	660	683	-23	-73	8
7	591	568	23	-27	3.5
8	719	748	-29	-79	10
9	543	530	13	-37	7
10	575	524	51	1	1

Decisión y Conclusión:

Como 10.5 es menor que 11 se rechaza H_0 y se concluye con $\alpha = 0.05$ que los problemas de muestra, en promedio, no aumentan las calificaciones de registro de graduados en 50 puntos

Aproximación Normal para Muestras Grandes

Cuando $n \geq 15$, la distribución muestral de W_+ ó W_- se aproxima a la distribución normal con media y varianza.

$$\mu_w = \frac{n(n+1)}{4} \qquad \sigma_w^2 = \frac{n(n+1)(2n+1)}{24}$$

Por tanto, cuando n excede el valor más grande en la tabla de Wilcoxon, se puede utilizar la estadística

$$Z = \frac{W_+ - \mu_w}{\sigma_w}$$

para determinar la región crítica de la prueba.

Contraste Kruskal Wallis

Cuando se presenta el problema de comparar más de dos muestras con el propósito de conocer si proceden de la misma población, o bien, comparar si existen diferencias entre las medidas de tendencia central de más de dos poblaciones y no se justifica la suposición de normalidad y de igualdad de varianzas.

La prueba de Kruskal-Wallis, es una alternativa a la prueba F del análisis de varianza para diseños de clasificación simple. En este caso se comparan varios grupos pero usando la mediana de cada uno de ellos, en lugar de las medias.

H_0 : La mediana de las a poblaciones consideradas son iguales y

H_a : Existe diferencia entre al menos un par de las poblaciones.

Para realizar la prueba de Kruskal–Wallis los datos pueden agruparse como se presenta en la tabla a continuación

Repeticiones	Grupos			
	A	B	C	...
1	X_{1A}	X_{1B}	X_{1C}	...
2	X_{2A}	X_{2B}	X_{2C}	...
⋮	⋮	⋮	⋮	...
n	X_{nA}	X_{nB}	X_{nC}	...

Donde X_{ij} son las observaciones
 $i = 1, 2, \dots, n$
 $j = A, B, C, \dots$
 $N = n_A + n_B + n_C \dots$

Pasos:

1. pasar las puntuaciones a rangos (conjuntamente en los “a” grupos)
2. computar la suma de los rangos en cada grupo (son las R_i)
3. Obtener el estadístico de contraste

$$H = \frac{1}{S^2} \left[\sum_{i=1}^a \frac{R_{i.}^2}{n_i} - \frac{N(N+1)^2}{4} \right]$$

En donde $R_{i.}$ es la suma de los rangos de las observaciones del i-ésimo grupo y n_i es el número de observaciones del i-ésimo tratamiento; N es el número total de observaciones

$$S^2 = \frac{1}{N-1} \left[\sum_{i=1}^a \sum_{j=1}^{n_i} R_{ij}^2 - \frac{N(N+1)^2}{4} \right]$$

Debe notarse que S^2 es igual a la varianza de los rangos. Si no hay empate, $S^2 = N(N+1)/12$ y el estadístico de prueba se simplifica:

$$H = \frac{12}{N(N+1)} \left[\sum_{i=1}^a \frac{R_{i.}^2}{n_i} - 3(N+1) \right]$$

Si n_i es razonablemente grande, como sería el caso de $n_i \geq 5$, entonces H tiene una distribución aproximadamente χ^2 con $a-1$ g.l. Por lo tanto, si

$$H \geq \chi_{\alpha, a-1}^2$$

hay que rechazar la hipótesis nula.

Ejemplo

Se quiere comparar el trabajo de cuatro analistas de un laboratorio en el ensayo de determinación del % de alcohol metílico en muestras de un producto químico, mediante la técnica de cromatografía líquida de alta resolución (HPLC). Los analistas reportaron los resultados siguientes:

Analista	% de alcohol metílico		
1	84.99	84.02	84.38
2	85.15	85.13	84.88
3	84.72	84.48	85.16
4	84.20	84.10	84.55

Mediante la prueba no paramétrica de Kruskal-Wallis, determinar si el trabajo de los analistas difiere significativamente. Use un nivel de significación de 0.05.

Solución

Hipótesis nula: los cuatro analistas trabajan de forma equivalente, es decir: $\mu_1 = \mu_2 = \mu_3 = \mu_4$

Hipótesis alterna: al menos dos analistas trabajan de forma diferente

Nivel de significación $\alpha = 0.05$

Procedimiento de prueba:

El primer paso consiste en ordenar la serie de datos en orden ascendente como si procedieran de la misma muestra, luego se establecen los rangos. La asignación de rangos se presenta en la siguiente tabla:

Muestra	Analista 1	Analista 2	Analista 3	Analista 4
1	9	11	7	3
2	1	10	5	2
3	4	8	12	6
$\sum R_{ij} =$	14	29	24	11

El cálculo del estadístico para el contraste es el siguiente:

Como no hay empates, el estadístico de prueba será:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^a \frac{R_i^2}{n_i} - 3(N+1)$$

$$H = \frac{12}{12 \times 13} \left(\frac{14^2}{3} + \frac{29^2}{3} + \frac{24^2}{3} + \frac{11^2}{3} \right) - 3 \times 13 = 5.46$$

Criterio de decisión: si $H > \chi_{\alpha, a-1}^2$ se rechaza la hipótesis nula. En este caso $\chi_{0.05, 3}^2 = 7.81$

Dado que $5.46 < 7.81$, no existe evidencia significativa al 5% como para rechazar la hipótesis nula, por lo que se concluye que no existe evidencia a un nivel de significación de 0.05, como para afirmar que los analistas trabajan de forma diferente.

Comparaciones múltiples

En aquellas situaciones en las que se haya rechazado la hipótesis acerca de la igualdad de las distribuciones poblacionales de las cuales hayan sido extraídas las muestras, será necesario, realizar contrastes *a posteriori* que determinen o precisen entre qué muestras existen las diferencias significativas que provocan el rechazar la hipótesis nula del contraste de Kruskal-Wallis.

El procedimiento consistirá en los siguientes pasos:

- En primer lugar determinaremos la cantidad

$$\alpha = \frac{\alpha}{k(k-1)}$$

- Posteriormente calculamos el valor del percentil de la distribución $\mathbf{z}_{1-\alpha}$

- Por último calculamos la cantidad

$$\Delta_{ij} = z_{1-\alpha} \sqrt{\frac{N(N+1)}{12} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

y en lo que respecta a la regla de decisión diremos que existen diferencias significativas entre dos poblaciones siempre que:

$$|\bar{R}_i - \bar{R}_j| > \Delta_{ij}$$

donde \bar{R}_i es el rango medio de la muestra i , es decir, $\bar{R}_i = \frac{R_i}{n_i}$

no admitiendo posibles diferencias en otro caso.

Contraste de Friedman

La prueba de Friedman desarrollada por el ganador del premio Nobel, el economista Milton Friedman, se diseñó para probar la hipótesis nula de que las distribuciones de probabilidad de los k tratamientos son idénticas, frente a la alternativa de que por lo menos dos de las distribuciones difieren en ubicación

Después de obtener los datos de un diseño aleatorizado de bloques, para *cada bloque*, los valores observados de las respuestas se ordenan para cada uno de las k “tratamientos” del rango 1 (el menor valor en el bloque) al rango k (el mayor valor del bloque). Si dos o más observaciones en el mismo bloque están empatadas para el mismo rango, entonces se asignará el promedio de los rangos que se habrían asignado a estas observaciones a cada miembro del grupo empatado. Nótese que solamente hay que tomar en cuenta los “empates” dentro del mismo bloque.

Prueba F_r de Friedman para un diseño aleatorizado de bloques

H_0 : Las distribuciones de probabilidad para los k tratamientos son idénticas.

H_a : Por lo menos dos de las distribuciones difieren en su ubicación.

Estadístico de la prueba:
$$F_r = \frac{12}{bk(k+1)} \sum_{i=1}^k R_i^2 - 3b(k+1)$$

en donde

b = Número de bloques.

k = Número de tratamientos.

R_i = Suma de los rangos para el i -ésimo tratamiento, en donde el rango de cada medición se calcula según su tamaño relativo dentro de su propio bloque.

Región de rechazo: $F_r > \chi^2_{\alpha}$ con $(k - 1)$ grados de libertad.

Supuestos:

Los tratamientos se asignan aleatoriamente a las unidades experimentales dentro de los bloques. Que el número de bloques (b) o el número de tratamientos (k) sea mayor que 5.

Ejemplo

Se realizó un experimento para comparar el tiempo de terminación para tres tareas técnicas de la manera siguiente. Como los tiempos de terminación pueden variar considerablemente de una persona a otra, se pidió a cada uno de los seis técnicos que realizaran las tres tareas. Se dieron las tareas a cada técnico en orden aleatorio con suficiente tiempo entre las tareas. ¿Presentan los datos de la Tabla evidencia suficiente para indicar que las distribuciones de los tiempos de terminación para las tres tareas difieren en ubicación? Utilice $\alpha = 0.05$. Obtenga los límites para el valor p asociado.

<i>Técnico</i>	<i>Tarea A</i>	<i>Tarea B</i>	<i>Tarea C</i>
1	1.21	1.56	1.48
2	1.63	2.01	1.63
3	1.42	1.70	2.06
4	1.16	1.27	1.27
5	2.43	2.64	1.98
6	1.94	2.81	2.44

Solución

<i>Técnico</i>	<i>Tarea A</i>	<i>Rango</i>	<i>Tarea B</i>	<i>Rango</i>	<i>Tarea C</i>	<i>Rango</i>
1	1.21	1	1.56	3	1.48	2
2	1.63	1.5	2.01	3	1.63	1.5
3	1.42	1	1.70	2	2.06	3
4	1.16	1	1.27	2.5	1.27	2.5
5	2.43	2	2.64	3	1.98	1
6	1.94	1	2.81	3	2.44	2
		$R_1 = 7.5$			$R_2 = 16.5$	$R_3 = 12$

Se realizó el experimento según un diseño aleatorizado de bloques, en donde los técnicos representan los bloques. En este caso se comparan $k = 3$ tratamientos, utilizando $b = 6$ bloques. Como el número de bloques es mayor que 5, podemos utilizar el análisis de Friedman y comparar el valor de F_r para χ^2_{α} con $k - 1 = 2$ grados de libertad, encontramos que $\chi^2_{0.05} = 5.99147$.

Para los datos de la Tabla tenemos

$$F_r = \frac{12}{6(3)(4)} [(7.5)^2 + (16.5)^2 + (12)^2] - 3(6)(4) = 6.75$$

Como $F_r = 6.75$, es mayor que 5.99147, concluimos a un nivel $\alpha = 0.05$ que los tiempos de terminación de al menos dos de las tres tareas, poseen una distribución de probabilidad que difieren en ubicación.

Ya que $F_r = 6.75$ es el valor observado de un estadístico que tiene aproximadamente una distribución χ^2 con 2 grados de libertad, tenemos que $0.025 < \text{valor } p < 0.05$ (aproximadamente).