Tema 7. Contrastes no paramétricos (II)

Dr. David Castilla Espino
CA UNED de Huelva, "Profesor Dr. José Carlos Vílchez Martín"

Introducción

- Bienvenida
- **Objetivos** pedagógicos:
  - Elaborar contrastes de aleatoriedad: de rachas de Wald Wolfowitz y del cuadrado medio de diferencias sucesivas.
  - Manejar los contrastes de localización: de signos, de signos para una muestra apareada y de Wilcoxon.

Introducción

- Comunicación:
  - Offline: Foro del grupo de tutoría del tutor.
- Referencias:
  - Casas, J.M. & P. Gutiérrez (2011): Estadística II: Inferencia Estadística. Editorial Universitaria Ramón Areces, Madrid. 301-351.
  - Ruiz-Maya, L. & J. Martín (1999): Fundamentos de Inferencia Estadística. Editorial AC. Madrid. 271-308.
  - Novales (1997),...

Introducción

- Antecedentes:
  - Contrastes de hipótesis (Tema 5)
    - Paramétricos. Contrastan hipótesis sobre el valor que toman los parámetros de distribuciones poblacionales conocidas (Tema 6). Ej. Contrastar hipótesis sobre μ y σ de una familia de distribuciones Normal.
    - No paramétricos (Tema 7): Contrastan características de la distribución distintas de los parámetros, son métodos de distribución libre y pueden ser también empleado a atributos nominales u ordinales.
      - Contraste Chi cuadrado de bondad del ajuste, independencia y homogeneidad (Sesión 1, Tema 7).

Contrastes de aleatoriedad

- Permiten contrastar si un conjunto de datos se han extraído aleatoriamente.
- Los más usuales son: Contraste de rachas de Wald-Wolfowitz y Contraste del cuadrado medio de diferencias sucesivas.
- Rachas: Sea una sucesión en la que intervienen dos tipos de símbolos: A y B:

  AAABBBAABBBAAAAAABBB

  una racha es una sucesión de uno o más símbolos idénticos que están precedidos o seguidos por un símbolo diferente o por ninguno, siendo la longitud de la racha el número de símbolos iguales que incluye.

  En el ejemplo:

  AAA (3) / BBB (3) / AA (2) / BBB (3) / AAAAAA (6) / BBB (3)

Pocas rachas indican claramente que la secuencia no es aleatoria (**persistencia**), mientras que demasiadas son síntoma de lo mismo (**zigzag**). También puede haber variaciones cíclicas que indican la ausencia de aleatoriedad

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de aleatoriedad

- Contraste de rachas de Wald-Wolfowitz.
  - Hipótesis del contraste:
    - H_0: La muestra es aleatoria.
    - H_1: La muestra no es aleatoria.
  - Estadístico de contraste: r_{exp} (número de rachas de la muestra).
  - Región crítica:
    - Bilateral: $r_{\exp} < K'_{\alpha/2}$ o $r_{\exp} > K_{\alpha/2}$ donde $P(R \le K'_{\alpha/2}) \le \alpha/2$ y $P(R \ge K_{\alpha/2}) \ge \alpha/2$
    - Unilateral (si se ve alguna tendencia a existir pocas –izquierda- o muchas rachas – derecha-):
      - Izquierda: $r_{\exp} < K'_{\alpha}$ donde $P(R \le K'_{\alpha}) \le \alpha$
      - Derecha: $r_{\exp} > K_{\alpha}$ donde $P(R \ge K_{\alpha}) \ge \alpha$

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de aleatoriedad

- Contraste de rachas de Wald-Wolfowitz (…)
  - Bajo el supuesto de la H_0:

r: Número de rachas.
n_1: Número de elementos de A en la muestra.
n_2: Número de elementos de B en la muestra.

$$P(R=r)=\begin{cases} \dfrac{2\binom{n_1-1}{\frac{r}{2}-1}\binom{n_2-1}{\frac{r}{2}-1}}{\binom{n_1-n_2}{n_1}} & para\ r\ par \\ \dfrac{\binom{n_1-1}{\frac{r}{2}-\frac{3}{2}}\binom{n_2-1}{\frac{r}{2}-\frac{1}{2}}+\binom{n_1-1}{\frac{r}{2}-\frac{1}{2}}\binom{n_2-1}{\frac{r}{2}-\frac{3}{2}}}{\binom{n_1-n_2}{n_1}} & para\ r\ impar \end{cases}$$

Esta distribución está tabulada.

La naturaleza discreta de la distribución hace difícil obtener el valor crítico por lo que se debe seguir el criterio de aproximarse todo lo posible a *α* o *α/2* sin superarlo.

Contrastes de aleatoriedad

- Contraste de rachas de Wald-Wolfowitz (…)
  - Bajo el supuesto de la H_0 (…):

  Cuando se verifica que $n_1>10$ y $n_2>10$ la distribución de probabilidad discreta considerada converge a una distribución Normal con la siguiente parametrización:

$$\mu=\frac{2n_1n_2}{n}+1$$

$$\sigma^2=\frac{2n_1n_2(2n_1n_2-n)}{n^2(n-1)}$$

  luego la región crítica bilateral será:

$$|z_{\exp}|<Z_{\alpha/2} \text{ donde } \quad z_{\exp}=\frac{r-\mu}{\sigma}$$

Contrastes de aleatoriedad

- Contraste de rachas de Wald-Wolfowitz (…)
  - El contraste de rachas también puede emplearse cuando los datos no son dicotómicos:
    - Datos cuantitativos: Haciendo las diferencias respecto a la mediana ($D_i=X_i$-Me) de modo que los signos (+ o -) permiten obtener las rachas. Despreciaríamos en este caso las diferencias iguales a 0 reduciendo el tamaño muestral ($D_i=0$).
    - Serie temporal: Igual que el anterior, pero en este caso sería necesariamente un contraste unilateral por la izquierda.
    - Se puede lleva a cabo comparando dos poblaciones.

Contrastes de aleatoriedad

- Contraste de rachas de Wald-Wolfowitz (…)
  - Ejemplos:

Ej1: El director de tesis de un joven investigador tiene dudas de que la muestra de edades extraída por su doctorando sea aleatoria al comprobar que existe cierto patrón en la muestra extraída. Las características de la muestra del doctorando en relación con la edad se presenta a continuación:

18, 21, 31, 31, 33, 18, 19, 35, 36, 32, 32, 32, 29, 29, 25, 19, 29, 29, 30, 31, 33, 35, 40, 41, 42, 40

Comprueba si el director de tesis está en lo cierto empleando el contraste de rachas de Wald-Wolfowitz (*α=0,05*).

H_0: La muestra es aleatoria.

H_1: La muestra no es aleatoria.

Dado que *Me=31*, entonces:

|--00|+|--|+++++|-------0|++++++|

Contrastes de aleatoriedad

(**Ej1**…)

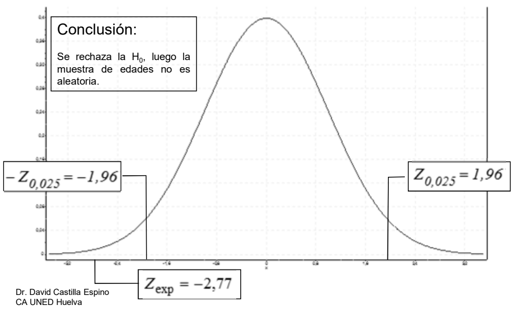
Considerando que $n_1(11)>10$ y $n_2(12)>10$ la distribución de probabilidad del estadístico de contraste r_{exp} (6) converge a una distribución Normal por lo que el valor del estadístico de contraste y la región crítica serán:

$$z_{\exp} = \frac{r - \left[\frac{2n_1n_2}{n} + 1\right]}{\sqrt{\frac{2n_1n_2(2n_1n_2 - n)}{n^2(n-1)}}} = \frac{6 - \left[\frac{2\cdot 11\cdot 12}{23} + 1\right]}{\sqrt{\frac{2\cdot 11\cdot 12\cdot(2\cdot 11\cdot 12 - 23)}{23^2(23-1)}}} = -2{,}77$$

$$|z_{\exp}| < Z_{\alpha/2}$$

Contrastes de aleatoriedad

(**Ej1**…)



Contrastes de aleatoriedad

- Contraste del cuadrado medio de diferencias sucesivas.
  - Se emplea cuando existen variaciones de naturaleza cíclica en las observaciones. En este caso, los valores contiguos tenderán a ser altos o bajos existiendo correlación entre las observaciones obtenidas.
  - Hipótesis del contraste:
    - H_0: La muestra es aleatoria.
    - H_1: La muestra no es aleatoria.
  - Estadístico de contraste:

$$R = \frac{D^2}{S^2} = \frac{\sum_{i=1}^{n-1} D_i^2}{2\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

$$D^2 = \frac{1}{2(n-1)}\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2 = \frac{1}{2(n-1)}\sum_{i=1}^{n-1} D_i^2$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de aleatoriedad

- Contraste del cuadrado medio de diferencias sucesivas(…)
  - Bajo el supuesto de H_0 cierta, D^2 y S^2 son los dos, estimadores insesgados de la varianza poblacional por lo que el estadístico R toma valores próximos a 1. Si no es cierta la H_0, $D^2 < S^2$ por lo que R tenderá a tomar valores cercanos a 0 ➔ Contraste unilateral por la izquierda.
  - Si la población de la que se ha extraído la muestra es Normal o se puede considerar asintóticamente Normal (n>20) bajo el supuesto de H_0:

$$R \xrightarrow[H_0]{n>20} N\left(\hat{\mu} = 1,\ \sigma^2 = \frac{n-2}{n^2-1}\right)$$

  - Región crítica:

$$z_{\exp} \leq -Z_\alpha \text{ donde } z_{exp} = \frac{R-1}{\sqrt{\frac{n-2}{n^2-1}}}$$

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de aleatoriedad

- Contraste del cuadrado medio de diferencias sucesivas(…)
  - Ejemplos:

Ej2: Compruebe para el caso del Ej1 la existencia de correlación serial con un retardo (*α=0,05*).

H_0: La muestra es aleatoria.

H_1: La muestra no es aleatoria.

Dado que $D^2=16,28$ y $S^2=48,81$, el estadístico R toma el valor:

$$R = \frac{D^2}{S^2} = \frac{\sum_{i=1}^{n-1} D_i^2}{2\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{16,28}{48,81} = 0,334$$

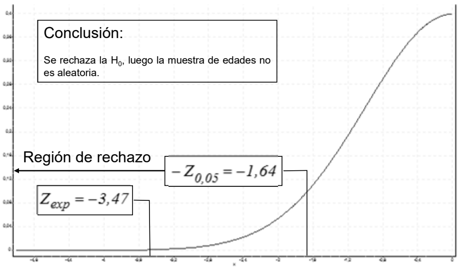
Considerando que n>20 R converge a una distribución Normal por lo que el estadístico de contraste y la región crítica son:

$$z_{exp} = \frac{R-1}{\sqrt{\frac{n-2}{n^2-1}}} = \frac{0,344-1}{\sqrt{\frac{26-2}{26^2-1}}} = -3,47 \qquad z_{\exp} \leq -Z_\alpha$$

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de aleatoriedad

(**Ej2**…)



Dr. David Castilla Espino
CA UNED Huelva

Contrastes de localización

- Permiten contrastar el valor de alguna medida de posición o localización de la distribución que sigue la población considerada.
- Sea una MAS de tamaño n $(X_1,\ldots X_n)$ procedente de una población de función de distribución $F(X)$ **continua** pero desconocida y sea $p \in R^+ / 0<p<1$ y $C_p(F)$ el percentil de orden p de la distribución $F(X)$ entonces los contrastes de localización pueden contratar para k_0 conocido las siguientes hipótesis:
  - Bilateral
    $H_0{:}C_p(F) = k_0$
    $H_0{:}C_p(F) \neq k_0$
  - Unilateral por la izquierda:
    $H_0{:}C_p(F) \geq k_0$
    $H_0{:}C_p(F) < k_0$

    Unilateral por la derecha:
    $H_0{:}C_p(F) \leq k_0$
    $H_0{:}C_p(F) > k_0$

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de localización

- Contraste de signos.
  - Requiere como único supuesto la continuidad de la población de la que procede la muestra.
  - Hipótesis del contraste: Cualquiera de las indicadas en la diapositiva anterior.
  - Estadístico de contraste: S^+ que representa el número de observaciones muestrales mayores que la medida de localización k_0 o el número de signos positivos de la siguiente diferencia:

$$D_i = X_i - k_0 = \begin{cases} > 0,\ si\ X_i > k_0\ asignamos\ el\ signo\ + \\ < 0,\ si\ X_i < k_0\ asignamos\ el\ signo\ - \end{cases}$$

Se desprecian las diferencias nulas pues la probabilidad de que esto ocurra es 0 dada la naturaleza continua de la distribución de probabilidad considerada.

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de localización

- Contraste de signos (…)
  - La probabilidad de que una diferencia tome signo positivo es *1-p=q* y se distribuye como un fenómeno de Bernouilli de parámetro *q*. De este modo la probabilidad de que *n* observaciones muestrales independientes presenten signo positivo se obtiene mediante una familia de distribuciones Binomial *B(n, q)*. Esto es $S^+ \rightarrow B(n,q)$
  - Región crítica:
    - Bilateral:

$$P\left(S^+ \leq k'_{\alpha/2}\right) \leq \alpha/2$$
$$P\left(S^+ \geq k_{\alpha/2}\right) \leq \alpha/2$$

La naturaleza discreta de la distribución Binomial hace difícil obtener el valor exacto de $k_{\alpha/2}$ ó $k'_{\alpha/2}$ por lo que se debe seguir el criterio de aproximarse todo lo posible a *α/2* sin superarlo. Si se cumplen las condiciones de convergencia de la distribución Binomial a la Normal (*n>30, np≥5* y *nq ≥5*) conviene emplear en este caso la distribución Normal [$N(\mu=np;\ \sigma^2=npq)$].

- Unilateral por la izquierda: $P\left(S^+ \leq k_\alpha\right) \leq \alpha$
- Unilateral por la derecha: $P\left(S^+ \geq k_\alpha\right) \leq \alpha$

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de localización

- Contraste de signos (…)
  - Ejemplos:

Ej3: Un estudio sugiere que los trabajadores de banca españoles dedican, en mediana, 34 minutos en el descanso del almuerzo. La directora provincial de un banco anotó el tiempo que 16 empleados elegidos al azar dedicaban al almuerzo sin que se enteraran. Los datos obtenidos en minutos se muestran a continuación:

18 55 20 31 12 18 35 28 16 14 32 12 12 36 35 34

Al comprobar los datos la directora quedo atónita al dar la impresión de que sus empleados parecían dedicar menos de 34 minutos al almuerzo. Compruebe si esto es cierto (*α=0,05*).

H_0: $Me \geq k_0=34$

H_1: $Me < k_0=34$

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de localización

(**Ej3**…)

Cálculo del estadístico de contraste y valor crítico:

Tiempo (X_i)	$D_i=X_i-k_0$	S^+_i
18	-16	-
55	21	+
20	-14	-
31	-3	-
12	-22	-
18	-16	-
35	1	+
28	-6	-
16	-18	-
14	-20	-
32	-2	-
12	-22	-
12	-22	-
36	2	+
35	1	+
34	0	0

$S^+ = 4$

$$F(x) = P(X \leqslant x) = \sum_{i=0}^{x} \binom{n}{i} p^i \cdot (1-p)^{n-i}$$

n	x	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
14	4	0,9996	0,9908	0,9533	0,8702	0,7415	0,5842	0,4227	0,2793	0,1672	0,0898
	5	1,0000	0,9985	0,9885	0,9561	0,8883	0,7805	0,6405	0,4859	0,3373	0,2120
	6		0,9998	0,9978	0,9884	0,9617	0,9067	0,8164	0,6925	0,5461	0,3953
	7		1,0000	0,9997	0,9976	0,9897	0,9685	0,9247	0,8499	0,7414	0,6047
	8			1,0000	0,9996	0,9978	0,9917	0,9757	0,9417	0,8811	0,7880
	9				1,0000	0,9997	0,9983	0,9940	0,9825	0,9574	0,9102
	10					1,0000	0,9998	0,9989	0,9961	0,9886	0,9713
	11						1,0000	0,9999	0,9994	0,9978	0,9935
	12							1,0000	0,9999	0,9997	0,9991
	13								1,0000	1,0000	0,9999
	14										1,0000
15	0	0,4633	0,2059	0,0874	0,0352	0,0134	0,0047	0,0016	0,0005	0,0001	0,0000
	1	0,8290	0,5490	0,3186	0,1671	0,0802	0,0353	0,0142	0,0052	0,0017	0,0005
	2	0,9638	0,8159	0,6042	0,3980	0,2361	0,1268	0,0617	0,0271	0,0107	0,0037
	3	0,9945	0,9444	0,8227	0,6482	0,4613	0,2969	0,1727	0,0905	0,0424	0,0176
	4	0,9994	0,9873	0,9383	0,8358	0,6865	0,5155	0,3519	0,2173	0,1204	0,0592
	5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

$$P\left(S^+ \leq k_{0,05}\right) \leq 0,05 \rightarrow k_{0,05} = 3$$

Se acepta la H_0 para *α=0,05*, luego los trabajadores no emplean menos de *34* minutos en el almuerzo

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de localización

- Contraste de signos (…)
  - En el caso concreto de que se quiera contrastar el valor de la mediana de la distribución bastaría con considerar que *p=q=0,5*.
  - Contraste de signos para muestras apareadas.
    - Muestra apareada: De cada elemento de la muestra se observan dos características X_i e Y_i.
    - Hipótesis del Contraste (caso bilateral):
      - H_0: $Me_x = Me_y$
      - H_1: $Me_x \neq Me_y$

Contrastes de localización

- Contraste de signos (…)
  - Contraste de signos para muestras apareadas (...)
    - Estadístico de contraste: S^+ que representa el número de signos positivos de la siguiente diferencia:

$$D_i = X_i - Y_i = \begin{cases} > 0, \ si \ X_i > Y_i, asignamos \ el \ signo \ + \\ < 0, \ si \ X_i < Y_i, asignamos \ el \ signo \ - \end{cases}$$

Cuando las medianas de *X* e *Y* son iguales, se debe verificar que la mediana de la distribución de la diferencias debe ser igual *0*, lo que implicaría que se dejarían igual número de signos a la derecha y a la izquierda de D_i. $S^+ \rightarrow B(n;0,5)$.

- Región crítica (caso bilateral):

$$P\left(S^+ \leq k'_{\alpha/2}\right) \leq \alpha/2$$
$$P\left(S^+ \geq k_{\alpha/2}\right) \leq \alpha/2$$

Contrastes de localización

- Contraste de signos (…)
  - Ejemplos:

Ej4: Un nutricionista está investigando el efecto del consumo de un complemento alimenticio en la pérdida de peso. Para ello extrae una MAS de 9 voluntarios para tomar parte en el experimento consistente en pesarse antes de tomar el complemento alimenticio y un mes después de tomárselo. Los resultados se muestran a continuación:

Individuo	A	B	C	D	E	F	G	H	I
Peso inicial (X_i)	83,7	83,9	68,2	74,9	81	72,8	61,3	77,9	69,6
Peso final (Y_i)	81,5	80	68,8	74,1	82,6	69,2	63,4	74,7	66,2

Compruebe si el complemento alimenticio ha tenido un efecto positivo a los efectos de reducir peso (*α=0,05*).

H_0: $Me_x \leq Me_y$

H_1: $Me_x > Me_y$

Contrastes de localización

(**Ej4**…)

Cálculo del estadístico de contraste y el valor crítico:

Individuo	Peso inicial (X_i)	Peso final (Y_i)	$D_i=X_i-Y_i$	S^+_i
A	83,7	81,5	2,2	+
B	83,9	80	3,9	+
C	68,2	68,8	-0,6	-
D	74,9	74,1	0,8	+
E	81	82,6	-1,6	-
F	72,8	69,2	3,6	+
G	61,3	63,4	-2,1	-
H	77,9	74,7	3,2	+
I	69,6	66,2	3,4	+

$$S^+ = 6 \rightarrow B(n = 9; p = 0{,}5)$$

$$\boxed{P\left(S^+ \geq k_{0,05}\right) \leq 0{,}05 \rightarrow k_{0,05} = 6}$$

Se acepta la H_0 para *α=0,05*, luego no se puede afirmar que el complemento alimenticio considerado ayuda a perder peso.

Contrastes de localización

- Contraste de rangos-signos de Wilcoxon.
  - Requiere como supuestos la continuidad y la simetría de la población de la que procede la muestra.
  - Hipótesis del contraste (caso bilateral):
    - H_0: *Me = m*
    - H_1: *Me ≠ m*
  - Estadístico de contraste:

$$T^+ = \sum_{i=1}^{n} Z_i \cdot r\left(\left|D_i\right|\right)$$

T^+ no es más que la suma de los rangos de las diferencias positivas. La distribución T^+ está tabulada.

$$D_i = X_i - k_0 = \begin{cases} > 0, \text{ si } X_i > k_0 \text{ asignamos el signo } + \\ < 0, \text{ si } X_i < k_0 \text{ asignamos el signo } - \end{cases}$$

$$Z_i = \begin{cases} 1, \text{ si } D_i > 0 \\ 0, \text{ si } D_i < 0 \end{cases}$$

$r(|D_i|)$ es el orden/rango de D_i en valor absoluto. Si varios $|D_i|$ son iguales se les asigna a todos el promedio de los rangos que les hubiese correspondido si no fuesen iguales.

Contrastes de localización

- Contraste de rangos-signos de Wilcoxon.
  - Región crítica (caso bilateral):

$$P\left(T^+ \leq k'_{\alpha/2}\right) \leq \alpha/2$$

$$P\left(T^+ \geq k_{\alpha/2}\right) \leq \alpha/2$$

Valores grandes o pequeños de T^+ implican grandes desviaciones respecto a la Me propuesta en la H_0.

  - Cuando *n>15* la distribución de T^+ se aproxima a una familia de distribuciones de probabilidad $N(\mu,\sigma^2)$ con parametrización:

$$\mu = \frac{1}{4} n(n+1)$$

$$\sigma^2 = \frac{n(n+1)(2n-1)}{24}$$

luego la región crítica bilateral será:

$$\left|z_{\exp}\right| < Z_{\alpha/2} \text{ donde } z_{exp} = \frac{T^+ - \frac{1}{4} n(n+1)}{\sqrt{\frac{n(n+1)(2n-1)}{24}}} \rightarrow N(0{,}1)$$

Contrastes de localización

- Contraste de de rangos-signos de Wilcoxon (…)
  - Ejemplos:

Ej5: La directora del Ej3 ha tenido conocimiento de que el contraste de rangos-signos de Wilcoxon es más fiable que el de signos debido a que no sólo considera los signos, sino también la magnitud de las diferencias entre los tiempos y la mediana. Considerando que tiene sentido asumir que la distribución es simétrica decide repetir el contraste con esta nueva especificación. Ayúdela en su cometido ($\alpha=0,05$).

H_0: $Me \geq k_0=34$
H_1: $Me < k_0=34$

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de localización

(**Ej5**…)

Cálculo del estadístico de contraste:

Tiempo (X_i)	$D_i=X_i-k_0$	$\|D_i\|$	$r(\|D_i\|)$	T^+_i	T^-_i
18	-16	16	8,5	-	8,5
55	21	21	12	12	-
20	-14	14	7	-	7
31	-3	3	5	-	5
12	-22	22	14	-	14
18	-16	16	8,5	-	8,5
35	1	1	1,5	1,5	-
28	-6	6	6	-	6
16	-18	18	10	-	10
14	-20	20	11	-	11
32	-2	2	3,5	-	3,5
12	-22	22	14	-	14
12	-22	22	14	-	14
36	2	2	3,5	3,5	-
35	1	1	1,5	1,5	-
34	0	0	-	-	-

Si varios $|D_i|$ son iguales se les asigna a todos el promedio de los rangos que les hubiese correspondido si no fuesen iguales. Para $|D_i| =1$, dado que son los dos primeros entonces les correspondería *{1, 2}* por lo que se les asigna el rango *1,5* [*(1+2)/2=1,5*], el siguiente sería $|D_i| =2$, a estos les correspondería *{3, 4}* por lo que se les asigna *3,5* [*(3+4)/2=3,5*], después sería $|D_i| =3$ al que se le asigna el rango *5*, y así sucesivamente.

$$T^+ = \sum_{i=1}^{n} Z_i \cdot r(|D_i|) = 18,5 \xleftarrow{\Sigma}$$

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de localización

(**Ej5**…)

Determinación del valor crítico:

n \ α^2	$P(T^+ \leq k_\alpha) \leq \alpha$					$P(T^+ \geq k_\alpha) \leq 1-\alpha$				
	,005	,01	,025	$\boxed{,05}$	,10	,90	,95	,975	,99	,995
4	—	—	—	—	0	10	—	—	—	—
5	—	—	—	0	2	13	15	—	—	—
6	—	—	0	2	3	18	19	21	—	—
7	—	0	2	3	5	23	25	26	28	—
8	0	1	3	5	8	28	31	33	35	36
9	1	3	5	8	10	35	37	40	42	44
10	3	5	8	10	14	41	45	47	50	52
11	5	7	10	13	17	49	53	56	59	61
12	7	9	13	17	21	57	61	65	69	71
13	9	12	17	21	26	65	70	74	79	82
14	12	15	21	25	31	74	80	84	90	93
$\boxed{15}$	15	19	25	$\boxed{30}$	36	84	90	95	101	105

FUENTE: KRAFT, C. H. y VAN EEDEN, A. *Nonparametric Introduction to Statistics*, Macmillan Publishing, 1968.

n=15 debido a que se ignoran las diferencias iguales a 0.
No es posible aplicar en este caso la convergencia a la Normal debido a que no se cumple la condición (*n>15*)

$$P(T^+ \leq k_{0,05}) \leq 0,05 \rightarrow k_{0,05} = 30$$

Se rechaza la H_0 para $\alpha=0,05$, luego los trabajadores emplean menos de *34* minutos en el almuerzo

Dr. David Castilla Espino
CA UNED Huelva

Contrastes de localización

- Contraste de de rangos-signos de Wilcoxon (…)
  - Ejemplos:

  Ej6: La directora del Ej5 ha caído en la cuenta de que no había considerado una de las observaciones de las que disponía de modo que ahora su muestra consta de las siguientes 17 observaciones:

  18 16 55 20 31 12 18 35 28 16 14 32 12 12 36 35 34

  Repita de nuevo el contraste de rangos-signos de Wilcoxon en este caso ($\alpha=0{,}05$).

  H_0: $Me \geq k_0=34$
  H_1: $Me < k_0=34$

Contrastes de localización

(**Ej6**…)

Cálculo del estadístico T^+:

Tiempo (X_i)	$D_i=X_i-k_0$	$\lvert D_i \rvert$	$r(\lvert D_i \rvert)$	T^+_i	T^-_i
18	-16	16	8,5	-	8,5
16	-18	18	10,5	-	10,5
55	21	21	13	13	-
20	-14	14	7	-	7
31	-3	3	5	-	5
12	-22	22	15	-	15
18	-16	16	8,5	-	8,5
35	1	1	1,5	1,5	-
28	-6	6	6	-	6
16	-18	18	10,5	-	10,5
14	-20	20	12	-	12
32	-2	2	3,5	-	3,5
12	-22	22	15	-	15
12	-22	22	15	-	15
36	2	2	3,5	3,5	-
35	1	1	1,5	1,5	-
34	0	0	-	-	-

$$T^+ = \sum_{i=1}^{n} Z_i \cdot r(\lvert D_i \rvert) = 19{,}5 \xleftarrow{\Sigma}$$

Contrastes de localización

(**Ej6**…)

Dado que *n(16)>15* se puede aplicar la convergencia a la distribución Normal de modo que el estadístico de contraste será ahora:

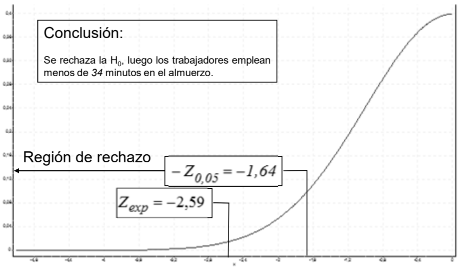
$$z_{exp} = \frac{T^+ - \frac{1}{4}n(n+1)}{\sqrt{\frac{n(n+1)(2n-1)}{24}}} = \frac{19{,}5 - \frac{1}{4}16(16+1)}{\sqrt{\frac{16(16+1)(2\cdot 16-1)}{24}}} = -2{,}59$$

La región crítica sería en este caso:

$$z_{\exp} < -Z_{\alpha}$$

Contrastes de localización

(**Ej6**…)



Dr. David Castilla Espino
CA UNED Huelva

Sumario

- Se han explicado los contrastes de aleatoriedad. En concreto se han introducido los contrastes de rachas de Wald-Wolfowitz y del cuadrado medio de diferencias sucesivas.
- Se han ejemplificado los dos contrastes de aleatoriedad explicados para el caso en el que se empleen datos de naturaleza cuantitativa.
- Se han explicado los contrastes de localización. En concreto se han introducido los contrastes de signos y el de rangos-signos de Wilcoxon.
- Se han ejemplificados los contrastes de signos para una muestra y muestras pareadas, así como el contraste de rangos-signos de Wilcoxon aplicando y sin aplicar la convergencia a la distribución Normal.

Dr. David Castilla Espino
CA UNED Huelva