Universidad Central de Venezuela
Facultad de Agronomía
Postgrado en Estadística

Evaluación de los procedimientos de comparaciones múltiples no paramétricas para una y dos vías de clasificación utilizados en la investigación biológica

Trabajo presentado como requisito parcial para optar al título de *Magister Scientiarum* en Estadística.

Autor:
Ing. Susana Araya Pinto

Tutores:
Dr. Franklin Chacín
Dr. Manuel Milla

Maracay, octubre de 2011

INTRODUCCIÓN

En variables biológicas comúnmente utilizadas, se observa la importancia del análisis no paramétrico para el caso de variables que son en esencia cualitativas como por ejemplo la presencia de características físicas que denotan la incidencia de ciertas enfermedades y los diferentes niveles de gravedad del síntoma estudiado, el cual ha sido comúnmente observado y medido en porcentajes que denotan los distintos niveles de gravedad de la enfermedad en estudio. Se han desarrollado una gran cantidad de investigaciones de este tipo tanto en el área de Fitopatología como en el área de Producción Animal.

Este trabajo se refiere a la evaluación de los procedimientos más comunes de comparaciones múltiples no paramétricas aplicados a este tipo de ensayos biológicos. Las dos pruebas no paramétricas más conocidas que permiten probar la igualdad de una serie de tratamientos, son la prueba de Kruskal-Wallis (1952), para una vía de clasificación y la prueba de Friedman para dos vías de clasificación. De sus resultados dependerá la aplicación de pruebas que permitan verificar en particular cuales de estas poblaciones son diferentes.

Al presentar las distintas pruebas de comparaciones múltiples, se deben tener en consideración las distintas tasas de error que han sido definidas para el error Tipo I. Las más importantes son la tasa de error por comparación y la tasa de error por experimento. La primera es la razón entre el número de comparaciones incorrectamente declaradas significativas entre el número total de comparaciones no significativas [Chew (1976)]. La tasa de error por experimento se define como la proporción de experimentos en los cuales hay al menos un rechazo de la hipótesis nula falsamente declarado.

Los métodos de comparaciones múltiples más utilizados son los que utilizan las sumas de los rangos, siendo el primero de ellos propuesto por Steel (1960), para una vía de clasificación y posteriormente la prueba de rangos asignados de Wicoxon sobre la cual se desarrollaron una serie de pruebas proporcionando alternativas a múltiples casos de interés científico relevante. Es con base en estos y algunos métodos alternativos como el que presentan Baumgartner, Weiss y Schindler (1998) para dos muestras, sobre los que se realizarán las comparaciones de interés en el presente trabajo.

La evaluación de las comparaciones se realiza con base en las tasas de error Tipo I y la eficiencia de Pitman y Bahadur para las diferentes pruebas utilizando para ello muestras generadas por remuestreo sobre observaciones reales recogidas de experimentación en el área de Fitopatología.

Los métodos de sumas de rangos han sido ampliamente utilizados para comparar si dos o más muestras provienen o no de poblaciones con distribuciones idénticas (Dunn 1964). Estos procedimientos se llevan a cabo ordenando el set completo de observaciones desde la más pequeña a la más grande y asignándoles rangos desde 1 hasta "n" y si existen empates, se les otorga el rango procedente del promedio de los rangos de las observaciones empatadas. Se sabe que algunos de estos procedimientos utilizan el rankeo conjunto de las observaciones, mientras que otros utilizan el rankeo pareado. Un dato importante sobre el enfoque de rankeo conjunto es que la comparación entre dos grupos dependerá del resto de las otras muestras no consideradas para esa comparación. En consecuencia, para los mismos valores $X_{i1}, \dots, X_{in}$ y $X_{j1}, \dots, X_{jn}$, las comparaciones de "i" versus "j", pueden resultar significativas para un experimento pero no significativas para otro (Neuhäuser y Bretz, 2001). Tanto los procedimientos de rankeo pareado como conjunto, mantienen la tasa de error Tipo I por experimento al nivel α designado; sin embargo, la utilización de rankeo conjunto no provee la protección de la tasa de error Tipo I que garantiza el rankeo pareado. Al utilizar pruebas conservativas el investigador se expone a la situación de no detectar diferencias que realmente existen y que es deseable detectar.

Para ilustrar lo anteriormente expuesto se presenta una experiencia real en el área de Fitopatología en la cual, luego del rechazo de la hipótesis nula, no se logró la diferenciación de los tratamientos a pesar de la utilización de pruebas de tipo liberal que son las que se encuentran a la mano en la literatura utilizada comúnmente en el área. Se hizo necesario entonces, plantear alternativas prácticas para ser utilizadas en estos casos que permitan el avance de los estudios que se están desarrollando en estas dos disciplinas.

OBJETIVOS DE LA INVESTIGACIÓN

OBJETIVO GENERAL

Evaluar los procedimientos de comparaciones múltiples no paramétricas para una y dos vías de clasificación utilizados en la investigación biológica.

OBJETIVOS ESPECÍFICOS

1. Analizar las tasas de error tipo I (familiar y por comparación), cometidos en cada una de las pruebas ejecutadas, bajo dos niveles de significación y cinco números de tratamientos.

2. Analizar la eficiencia Pitman en cada una de las pruebas ejecutadas, bajo dos niveles de significación y cinco números de tratamientos.

3. Analizar las eficiencias Bahadur entre cada una de las pruebas ejecutadas, bajo dos niveles de significación y cinco números de tratamientos.

4. Comparar los métodos en términos de las tasas de error tipo I (familiar y por comparación) y las eficiencias Pitman y Bahadur.

REVISIÓN BIBLIOGRÁFICA

Hasta el momento, se ha propuesto una amplia variedad de métodos para comparar los efectos de tratamiento, o dicho de una manera más general, comparar la localización relativa de las medianas de diferentes poblaciones. Se mencionan a continuación los dos criterios que se utilizan para realizar las comparaciones múltiples de interés. La base de ambos criterios es la ordenación en rangos de las observaciones en sustitución de sus valores intrínsecos.

De acuerdo con Critchlow y Fligner (1991), se utilizan dos criterios para la realización de las comparaciones múltiples. El primero de estos es el basado en el método de Kruskal y Wallis que utiliza el rankeo conjunto de todas las "N" observaciones siendo $N = \sum_{i=1}^{k} n_i$, y siendo el mismo utilizado por primera vez por Nemenyi (1963), y luego por Dunn (1964) conjuntamente con la desigualdad de Bonferroni. Desde este punto de vista, $\bar{R}_i$ y $\bar{R}_j$ denotan los promedios de los rangos asignados a la i-ésima y j-ésima muestra del rankeo conjunto realizado. Luego, dos tratamientos son considerados diferentes si $\left|\bar{R}_i - \bar{R}_j\right|$ exceden el valor crítico apropiado.

El segundo criterio usa el método de Wilcoxon de rankeo pareado, el cual fue propuesto independientemente por Steel (1960) y por Dwass (1960). Según este criterio, S_{ij} es la suma de los rangos asignados a la i-ésima muestra en el rankeo en el que se han combinado únicamente las $n_i + n_j$ observaciones de las muestras i-ésima y j-ésima. Los tratamientos i y j serán declarados diferentes cuando S_{ij} se ubique en la región critica apropiada.

Tasas de error

Con relación a las tasas de error, Federer (1955) señala que al realizar las comparaciones de dos o más tratamientos se definen tres tasas de error Tipo I:

1. Tasa de error por comparación. Número de comparaciones declaradas significativas incorrectamente / Número total de comparaciones realizadas. Lo que es igual a la proporción de todas las comparaciones que se espera que sean erradas cuando la hipótesis nula es cierta.

2. Tasa de error por experimento. Número de inferencias erróneas / Número de experimentos. Lo que es igual al número esperado de declaraciones erróneas por experimento cuando la hipótesis nula es falsa.

3. Tasa de error experimental. Número de experimentos con una o más declaraciones erróneas / Número de experimentos. Lo que es igual a la proporción esperada de experimentos con una o más declaraciones erróneas cuando la hipótesis nula es cierta.

Según Chew (1976), la tasa de error experimental no hace distinciones entre rechazar incorrectamente una comparación y rechazar por ejemplo 5 comparaciones en un experimento particular. Tampoco hace distinciones entre un experimento con 2 tratamientos, donde sólo es posible hacer 1 comparación, de otro experimento con 20 tratamientos en el cual se pueden hacer 190 comparaciones posibles. Es más fácil rechazar incorrectamente una o más comparaciones en un experimento grande con 20 tratamientos que en uno pequeño con 2 tratamientos. Por lo tanto, un error experimental del 5 por ciento es mucho más estricto que un error de comparación de 5 por ciento.

Kuehl (2001), define que el nivel de significancia o probabilidad de error tipo I para una sola prueba es una tasa de error con respecto a la comparación α_C . Es el riesgo que se está dispuesto a correr en una sola comparación. Pero la realidad es que existiendo, en comparaciones múltiples $k(k-1)/2$ comparaciones por pares, existirá un número determinado de "n" comparaciones posibles, lo que conduce a la posibilidad de cometer "n" errores tipo I para "n" pruebas. Se puede emplear en este caso, otra forma de error tipo I basada en el riesgo acumulado asociado con la familia de pruebas en estudio. La familia es el conjunto de comparaciones por pares. El riesgo acumulado asociado a una familia de comparaciones se conoce como tasa de error tipo I con respecto al experimento, α_E. Es el riesgo de cometer al menos un error tipo I en la familia de comparaciones en el experimento.

Aunque las pruebas en la familia no son independientes, se suponen como tal para evaluar el error tipo I con respecto al experimento. Suponiendo ciertas las hipótesis nulas de las "n" pruebas, la probabilidad de un error tipo I para cualquier prueba sola es α_C y (1- α_C) es la probabilidad de decisión correcta. La probabilidad de cometer "x" errores tipo I tiene distribución binomial:

$$P(x) = \frac{n!}{x!(n-x)!} \propto_C^x (1 - \propto_C)^{n-x} \quad (1)$$

La probabilidad de no cometer error tipo I es:

$$P(x = 0) = (1 - \propto_C)^n \quad (2)$$

La probabilidad de cometer al menos un error tipo I es $1 - P(x = 0)$ entre las "n" comparaciones independientes es:

$$\propto_E = 1 - (1 - \propto_C)^n \quad (3)$$

Se puede expresar la tasa de error tipo I con respecto a la comparación como una función de la tasa de error tipo I con respecto al experimento:

$$\propto_C = 1 - (1 - \propto_E)^{1/n} \quad (4)$$

Por lo tanto, es obvio deducir que las pruebas que utilizan el error experimental para realizar las comparaciones de interés son mucho menos exigentes que las que utilizan el error por comparación.

Blair y Taylor (2008), designan al error que antes se llamó experimental como error familiar y acotan que si sucede un solo rechazo de la hipótesis nula en una familia de pruebas, ha ocurrido un error familiar. También comentan que el error familiar se convierte en una preocupación en el contexto de las comparaciones múltiples y que para el control del mismo, se han desarrollado muchos métodos, siendo los más conocidos el ajuste de Bonferroni y posteriormente el método de reducción de Bonferroni.

Las pruebas de comparaciones múltiples deben mantener controlados ambos errores a pesar de que el aumento del número de comparaciones incide indiscutiblemente en el aumento del error familiar. Se hace evidente la relación entre α_{PCE} (error por comparación) y α_{FWE} (error familiar) al observar los resultados propuestos por Blair y Taylor (2008) los cuales se generaron por simulación y se muestran en el cuadro 1.

Cuadro 1. Tasas de error por comparación y familiar de los números de comparaciones establecidos.

α_{PCE}	Número de grupos	Número de comparaciones	α_{FWE}
	3	3	.122
	5	10	.286
.05	10	45	.630
	20	190	.920
	3	3	.027
	5	10	.075
.01	10	45	.231
	20	190	.528

Blair y Taylor (2008)

Se observa también en el cuadro 1 que el error familiar disminuye drásticamente cuando se reduce el error por comparación.

Lugo (2006), señala que debe existir un balance entre los errores y acota que las pruebas que controlan la tasa de error por comparación se ven afectadas por el número de comparaciones, pues el nivel de significación real aumenta con respecto al declarado. Las pruebas que controlan la tasa de error por experimento no se ven afectadas por el número de comparaciones, pues tratan el conjunto como un todo, pero a medida que el número de comparaciones aumenta, la posibilidad de detectar diferencias verdaderas disminuye.

También hace la diferencia entre pruebas liberales y conservativas. Las pruebas liberales son aquellas que tienen bajo error tipo II. Son más poderosas y en éstas, la posibilidad de cometer error tipo I es mayor que el nivel de significación propuesto. Las pruebas conservadoras o conservativas son menos poderosas, hay mayor posibilidad de cometer error tipo II pero mantienen la posibilidad de cometer error tipo I en o por debajo del nivel de significación declarado.
Es por esta razón que al seleccionar una prueba de comparaciones múltiples se debe tener en cuenta el número de comparaciones a realizar y lo liberal o conservador que el investigador esté dispuesto a ser. Hay situaciones en las que rechazar incorrectamente una comparación es tan serio como rechazar incorrectamente 10 comparaciones, así que elegir el error por experimento, es más pertinente.

Eficiencia Pitman de las pruebas en procedimientos de una vía de clasificación

Hollander y Wolfe (1999), refieren en su libro que la eficiencia relativa de Pitman para la mayoría de las alternativas no paramétricas de los procedimientos de comparaciones múltiples provenientes de análisis de una vía de clasificación, está dada por la siguiente expresión

$$e_F = 12\sigma_F^2\left\{\int_{-\infty}^{\infty} f^2(u)du\right\}^2 \tag{5}$$

Donde σ_F^2 es la varianza común de la función continua de base F, $f(.)$ es la función de densidad de probabilidad correspondiente a F. El parámetro $\int_{-\infty}^{\infty} f^2(u)du$ es el área bajo la curva asociado con $f^2(.)$, el cuadrado de la función de densidad de probabilidad común. La eficiencia asintótica de Pitman para el test de Kruskal y Wallis basado en el estadístico H, con respecto a la prueba F para una vía de clasificación, fue presentada por Andrews (1954). Otras investigaciones posteriores para otras pruebas como la de Jonckheere y Terpstra, Mack y Wolfe y Fligner-Wolfe contra sus competidores análogos basados en la teoría normal, también fueron hallados como e_F.

Eficiencia Pitman de las pruebas en procedimientos de dos vías de clasificación

La eficiencia asintótica Pitman con respecto a sus contrapartes basadas en la teoría normal está dada por la expresión

$$e_F = \left[\frac{k}{(k+1)}\right]\left[12\sigma_F^2\left\{\int_{-\infty}^{\infty} f^2(u)du\right\}^2\right] \tag{6}$$

Donde σ_F^2 es la varianza común de la función continua de base F, $f(.)$ es la función de densidad de probabilidad correspondiente a F. El parámetro $\int_{-\infty}^{\infty} f^2(u)du$ es el área bajo la curva asociado con $f^2(.)$, el cuadrado de la función de densidad de probabilidad común. Nótese que e_F es simplemente $k/(k+1)$ veces la correspondiente eficiencia de Pitman para una, dos, hasta “k” muestras. En particular, la eficiencia asintótica relativa de Pitman para la prueba de Friedman con respecto a la teoría normal para dos vías fue hallada por Elteren y Noether (1959). La eficiencia asintótica relativa de Page con su competidor normal fue presentada por Hollander (1967).

La eficiencia e_F es siempre igual o mayor a .576 y puede ser infinito. Algunos valores se presentan en Hollander (1999) y se muestran a continuación en el cuadro 2.

Cuadro 2. Valores de e_F para varias distribuciones y números de tratamientos (k).

k	**2**	**3**	**4**	**5**	**10**	**20**	**50**	∞
Distribución					$\boldsymbol{e_F}$			
Normal	.637	.716	.764	.796	.868	.909	.936	.955
Uniforme	.667	.750	.800	.833	.909	.952	.980	1.000
Exponencial								
Doble	1.000	1.125	1.200	1.250	1.364	1.429	1.471	1.500

Hollander y Wolfe (1999)

Mehra (1972) expone que para pruebas de rango conjunto, los estadísticos de rango que propone, basados en la prueba de Wilcoxon para una muestra simple arreglada en orden ascendente en un rankeo combinado, tienen distribución límite χ^2 con $(k-1)g.de\,l.$; entonces al calcular la eficiencia asintótica Pitman relativa a las alternativas normales es importante destacar el hecho de que el estadístico F no depende del número de tratamientos "k". Mehra (1972), plantea además que desde el punto de vista de la eficiencia Pitman, el rankeo conjunto parece redundante, ya que para una familia de estadísticos de rango dada, la eficiencia será la misma en el sentido Pitman.

Eficiencia asintótica de Bahadur

Koziol y Reid (1977), sugieren que las diferencias entre los procedimientos de comparaciones múltiples pueden ser mejor explicadas por la eficiencia Bahadur que por la de Pitman. La eficiencia Bahadur se aplica a alternativas fijadas más que a una secuencia que converge a 0.

Fairley y Pearl (1984), compararon los procedimientos de rankeo conjunto a los cuales llamaron "K", por basarse en el procedimiento de Kruskal-Wallis, vs los de rankeo pareado a los que llamaron "W" por basarse en el procedimiento de Wilcoxon. Estas comparaciones las realizaron en relación a la eficiencia Bahadur, mostrando que la eficiencia Bahadur de "W" con respecto a "K" es mayor que 1 al realizar comparaciones de poblaciones próximas, pero menor que 1 al comparar las más distantes en las cuales se observan poblaciones con grandes separaciones. Mostraron también que "W" es más eficiente para distinguir poblaciones cercanas y "K" lo es para las más distantes.

Sea R_{ijk} el rango combinado de las X_{ik} de la muestras combinadas de las poblaciones i y j, y sea

$$W_{ij} = \frac{\left[\sum_{i=1}^{n} R_{ijk} - n_i\left(n_i + n_j + 1\right)/2\right]}{\sqrt{\frac{n_i n_j\left(n_i + n_j + 1\right)}{12}}} \tag{7}$$

La versión estandarizada del estadístico de Wilcoxon para comparar las poblaciones "i" y "j".

Y sea R_{ik} el rango de las X_{ik} muestras conjuntas y sea

$$K_{ij} = \frac{R_{i.} - R_{j.}}{\sqrt{\frac{N(N+1)}{12}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}} \tag{8}$$

la versión estandarizada de las diferencias en los promedios de rangos entre las poblaciones i y j. Ambos procedimientos rechazan la hipótesis nula $\theta_i = \theta_j$ si el valor absoluto de sus estadísticos es mayor que el valor crítico establecido c_{ij} para todo $1 \leq i, j \leq k$.

Fairley y Pearl (1984), hallaron la eficiencia relativa entre dos procedimientos en dos poblaciones 1 y 2, siendo $\theta_i \neq \theta_j$. Demostraron que W_{12} y K_{12} tienen pendiente asintótica Bahadur.

$$C_{W(1,2)} = \frac{12}{\frac{n_1}{N}+\frac{n_2}{N}} . \frac{n_1}{N} . \frac{n_2}{N} . \left(P_{12} - \frac{1}{2}\right)^2 \tag{9}$$

y

$$C_{K(1,2)} = \frac{12}{\frac{n_1}{N}+\frac{n_2}{N}} . \frac{n_1}{N} . \frac{n_2}{N} . \left[\left(\frac{n_1}{N} + \frac{n_2}{N}\right)\left(P_{12} - \frac{1}{2}\right) + \sum_{i=1}^{k}(P_{1i} - P_{2i})\right]^2 \tag{10}$$

donde, $P_{ij} = P_{\tilde{\theta}}(X_{i1} > X_{j1})$ y asumiendo que $P_{12} \geq \frac{1}{2}$. Entonces la eficiencia Bahadur de W_{12} con respecto a K_{12}

$$B_{WK}(1,2) = \frac{C_{W(1,2)}}{C_{K(1,2)}} \tag{11}$$

y será mayor o igual a 1 si y solo si

$$\left(1 - \frac{n_1}{N} - \frac{n_2}{N}\right)\left(P_{12} - \frac{1}{2}\right) \geq \sum_{i=1}^{k} \frac{n_1}{N}(P_{1i} - P_{2i}) \tag{12}$$

para 3 o más grupos, es decir k > 3, y siendo 1 la población más grande, 2 la mediana y 3 la más pequeña, la ecuación (12) se reduce a lo siguiente

$$P_{12} + P_{23} \geq P_{13} + \frac{1}{2}\left(p.ej. B_{WK(1,2)} \geq 1\right) \tag{13}$$

La desigualdad (13) fija la relación entre $B_{WK(1,3)}$ y $B_{WK(2,3)}$ en 1, dependiendo de en qué lugar, estocásticamente hablando, se encuentra la población 3 con respecto a la 1 y la 2. Si se asume que $P_{13} \geq P_{12} \geq \frac{1}{2}$, entonces (13) implica que

$$P_{23} + P_{31} \geq P_{21} + \tfrac{1}{2}\left(p.ej. B_{WK(2,3)} \geq 1\right) \tag{14}$$

$$P_{13} + P_{32} \leq P_{12} + \tfrac{1}{2}\left(p.ej. B_{WK(1,3)} \leq 1\right) \tag{15}$$

Fairley y Pearl (1984), muestran que las desigualdades (14) y (15) plantean los siguientes casos:

a. "W" es más eficiente al distinguir entre las poblaciones más cercanas 1 y 2 o 2 y 3 mientras que K es más eficiente para las poblaciones distantes 1 y 3.
b. "K" es más eficiente para las cercanas y "W" para las distantes y
c. Ambos procedimientos son igualmente eficientes para todas las comparaciones.

Grané y Tchirina (2008) señalan que la eficiencia asintótica relativa de Bahadur es el concepto más adecuado para comparar los estadísticos de Kolmogorov-Smirnov, Cramér-von Mises y Anderson-Darling dado que éstos tienen diferentes distribuciones asintóticas.

Entre las investigaciones que se han realizado en relación a las diferencias existentes entre los métodos de comparaciones múltiples, destaca la propuesta de Fligner (1985), la cual enfrenta el rankeo pareado *versus* el rankeo conjunto en el estadístico de Kruskal-Wallis. En su trabajo, Fligner plantea que un estadístico de prueba para las diferencias de localización de "k" muestras, se construye combinando apropiadamente todas las muestras pareadas del estadístico de Wilcoxon. El resultado es un estadístico análogo al estadístico de Kruskal-Wallis siendo la eficiencia relativa asintótica Pitman entre ambos igual a uno. Sin embargo, la eficiencia Bahadur del estadístico pareado relativo al conjunto KW, ha mostrado ser mayor o igual a uno en todas las alternativas estudiadas; es decir, pareciera que el rankeo pareado es más eficiente. Este estudio sugiere que tal vez sea apropiado hacer las combinaciones de todos los pares posibles en lugar de usar un estadístico de rankeo conjunto como el de Kruskal-Wallis.

Alonzo *et al.*(2009) realizaron una comparación de las pruebas para alternativa ordenadas en el caso de tres clases. En este trabajo realizaron la comparación con respecto a la potencia y tamaño utilizando métodos de simulación con 1000 replicaciones para cada caso planteado. Concluyen que las pruebas ordenadas tipo “umbrella” o paraguas deben seguir las siguientes propiedades: (1) el tamaño de la prueba debería ser aproximadamente igual al tamaño nominal, (2) la prueba debería tener mayor potencia que la prueba alternativa cuando la hipótesis alternativa es cierta, (3) la prueba debería tener menor potencia para cualquier hipótesis alternativa que no sea consistente con la alternativa cierta. También concluyen que la prueba UV (umbrella volume) es menos potente que la de Mack-Wolfe .

Bristol (1990), comparó dos procedimientos de comparaciones múltiples contra un control para distribuciones libres. Estos dos procedimientos son el presentado por Chakraborti y Desu (1988) y el procedimiento propuesto por Slivka (1970). Ambos procedimientos son generalizaciones de las pruebas de Mathisen (1943) y fueron desarrollados para experimentos cuyas observaciones se han obtenido de manera ordenada. Para la comparación se presentan aproximaciones a la normal las cuales indican que el procedimiento Slivka requiere un tamaño de muestra menor para garantizar una potencia específica para las alternativas Lehman y alternativas proporcionales de riesgo cuando se necesitan todas las observaciones.

METODOLOGÍA

Datos experimentales

Se dispuso de una matriz de datos suministrada por el Laboratorio de Fitopatología de la Facultad de Agronomía de la Universidad Central de Venezuela, la cual corresponde a los resultados obtenidos de un experimento en invernadero cuyos hallazgos fueron presentados por Mariño *et al* (2009) y publicados bajo la modalidad de nota técnica por Garrido *et al* (2010).

El experimento consiste en probar si 15 cultivares de sorgo comúnmente producidos en Venezuela, manifiestan la misma respuesta en términos del porcentaje de daño, ante la incidencia del potyvirus del sorgo. Para este propósito, se utilizó material experimental homogéneo consistente en potes con cuatro plantas cada uno (unidad experimental) para ensayar los 15 tratamientos (cultivares) con cinco repeticiones cada uno, bajo un diseño completamente aleatorizado. Realizando una observación por planta, fueron generados 20 casos por tratamiento. Los porcentajes de daño correspondientes a los 20 casos para los 15 cultivares, generaron una "población original" (denominado de esta manera para los fines de remuestreo). El cuadro 3 muestra la asignación de los cultivares a los tratamientos en el ensayo.

Considerando que la variable porcentaje de daño es evaluada de forma visual, no queda duda acerca de su naturaleza cualitativa y, por lo tanto, se ha conducido una prueba no paramétrica de Kruskal-Wallis para probar la existencia de diferencias estadísticamente significativas al 5% entre los tratamientos. Resultando significativa esta prueba, se condujo una prueba de comparaciones múltiples no paramétricas tradicional, utilizando el programa Statistix 8, la cual produjo como resultado la existencia de cinco grupos homogéneos. Los tratamientos 1 y 9 conformaron el primero de éstos; los tratamientos 2, 3, 4, 7, 8, 12, 13, 14 y 15 el segundo; los tratamientos 5 y 6 el tercero y finalmente, los tratamientos 10 y 11 mostraron comportamientos muy diferentes, por lo que se decidió que cada uno estuviese solo, conformando los dos últimos grupos homogéneos.

Cuadro 3. Asignación de los cultivares a los tratamientos en el ensayo de sorgo.

Tratamiento	Cultivar
1	Criollo 1
2	Ismael
3	Guarao
4	Zaraza 1
5	Sefloarca 10
6	Cacique II
7	Yaruro I
8	Yaruro VII
9	Maracay
10	OKY-8
11	QL-11
12	Himeca 500
13	Himeca 383
14	WAC-8228
15	Himeca 101

Arreglos realizados con los tratamientos

Con base en los grupos homogéneos conformados, fueron realizadas cinco configuraciones de tratamientos, dependiendo del número de éstos y de las diferencias observadas entre ellos. Los arreglos fueron los siguientes:

Configuración 1: Todos los tratamientos. Considera los 15 tratamientos involucrados en el experimento original.

Configuración 2: Diez tratamientos contrastantes. Considera sólo diez de los tratamientos que arrojaron los resultados más distanciados entre sí. Fueron elegidos tratamientos de todos los grupos homogéneos y aquellos que pertenecen a un mismo grupo, manifiestan los comportamientos más extremos. Los tratamientos incluidos en esta configuración fueron: 1, 2, 3, 4, 6, 9, 10, 11, 12 y 14.

Configuración 3: Diez tratamientos semejantes. Toma en cuenta diez de los tratamientos más similares entre sí. Fueron elegidos tratamientos pertenecientes a un mismo grupo homogéneo y también correspondientes a grupos diferentes pero que no difieren de forma extrema. Se evaluaron las pruebas dentro de los grupos homogéneos y entre los grupos homogéneos. Los tratamientos incluidos en esta configuración fueron: 3, 5, 6, 7, 8, 11, 12, 13, 14 y 15.

Configuración 4: Cinco tratamientos contrastantes. Considera solamente los cinco tratamientos con respuestas más heterogéneas entre sí. Fueron escogidos aquí los cinco tratamientos con diferencias más extremas de los diez acogidos en la configuración 2, los cuales son: 1, 2, 6, 10 y 11.

Configuración 5: Cinco tratamientos semejantes. Involucra sólo los cinco tratamientos más similares entre sí. Fueron elegidos en este caso los cinco tratamientos con menores diferencias de los diez utilizados en la configuración 3, los cuales son: 5, 7, 11, 13 y 15.

Es importante en este punto señalar que las mencionadas diferencias entre tratamientos se refieren a las distancias entre los porcentajes de daño asignados a los tratamientos.

Nótese que no hubo cambio en el tamaño de las muestras, ya que los valores críticos que se calculan en las pruebas evaluadas en este trabajo no suelen depender de esta cantidad sino del número de tratamientos.

En todas pruebas que incluyen un control o testigo, fue utilizado el tratamiento 11 como tal, ya que es el cultivar de referencia utilizado en el experimento. Por esta razón, este tratamiento se ha incluido en todas las evaluaciones.

Generación de las muestras

Para la conducción de las pruebas señaladas, fueron generadas 1000 muestras de tamaño cinco bajo las diferentes configuraciones mencionadas, por medio de un procedimiento de remuestreo (sin reemplazo) sobre la referida "población original", utilizando el programa Resampling Stats for Excel 2007.

Debido a la necesidad de generar suficientes datos para producir muestras diferentes bajo el procedimiento señalado, no se realizó distinción alguna entre los casos correspondientes a la misma unidad experimental y entre unidades experimentales. Adicionalmente, este hecho es de poca relevancia ante el análisis estadístico no paramétrico de los datos, puesto que ninguna de las pruebas evaluadas en este trabajo incluye submuestreo. Más aún, ya que la media no es la medida de concentración utilizada por estos métodos, no se ha encontrado una forma conveniente de recopilar la información de los elementos de cada unidad experimental para que sea expresada en una única medida.

Cada elemento de la muestra obviamente corresponde a una repetición completa del experimento, originalmente bajo un diseño completamente aleatorizado.

Desarrollo de las pruebas

Las pruebas de comparaciones múltiples no paramétricas descritas, han sido programadas utilizando hojas de cálculo con novedosas y poderosas funciones de asignación de rangos, operaciones algebraicas y de ordenamiento disponibles en el paquete Microsoft Excel

2010. Cada prueba bajo cada una de las configuraciones señaladas, fue programada para la primera muestra, repitiendo su ejecución para las 999 muestras restantes por medio del uso de macros desarrollados a través del sistema Microsoft Visual Basic 2010. Fue preparada una hoja de cálculo con todas las pruebas para cada configuración.

En todos los casos, fueron comparados todos los pares posibles de tratamientos. Al final de cada hoja de cálculo, se ha programado un contador que indica la frecuencia de veces en que se ha declarado la diferencia entre cada par de tratamientos como significativa a dos niveles de significación: 10% y 5%. Todos los resultados fueron recolectados en una hoja de cálculo aparte, donde se ha denominado "tasa de rechazos" a la frecuencia relativa de rechazos en relación a las 1000 pruebas ejecutadas siempre que el par de tratamientos fueran considerados diferentes o contrastantes desde el principio y "tasa de error tipo I" a la frecuencia relativa de rechazos en relación a las 1000 pruebas ejecutadas a los casos en los cuales en par de tratamientos fueran considerados semejantes desde el principio.

Las pruebas propuestas con dos vías de clasificación fueron desarrolladas considerando que cada repetición (cada uno de los cinco casos resultantes de la repetición hipotética del experimento en el remuestreo) es un bloque completo (por conveniencia).

Además de las pruebas de comparaciones múltiples no paramétricas, fueron ejecutadas dos pruebas clásicas paramétricas con el fin de generar la información necesaria para las evaluaciones de algunas eficiencias relativas, tal como se indicará en los criterios de comparación más adelante. En aquellos casos en que se estaban considerando pruebas con todos los pares posibles de tratamientos, se desarrolló la prueba de Tukey a los niveles de significación señalados, ya que provee un valor crítico más alto y suele producir grupos homogéneos más grandes como consecuencia. Es importante recalcar en este momento que el principal inconveniente de las pruebas de comparaciones múltiples no paramétricas radica en su escasa capacidad de detectar diferencias significativas reales entre pares de tratamientos y, por lo tanto, si corresponde comparar su desempeño con algún procedimiento paramétrico clásico, Tukey debería ser, lógicamente, la prueba que sea más conservadora en este sentido.

Por otra parte, en aquellos casos en que fueron consideradas pruebas contra un control o testigo, se desarrolló la prueba de Dunnett a los niveles de significación señalados, utilizando siempre el mayor de los valores críticos; es decir, el valor crítico obtenido con los pares de tratamientos más alejados entre sí.

Evidentemente, las pruebas de Tukey y Dunnett procesaron los datos como si estos fueran de naturaleza cuantitativa y mostraran un adecuado ajuste a la distribución normal y, por lo tanto, se utilizó la media como medida de concentración.

Criterios de comparación

Control del error tipo I

Se han comparado los métodos en términos de las tasas de error tipo I familiar y por comparación. Un requerimiento para las pruebas de comparaciones múltiples, es que el error familiar se encuentre fuertemente controlado a un nivel de significación especificado (α).

Eficiencia asintótica relativa

Se han comparado las diferencias entre las eficiencias relativas de las pruebas más conocidas en ambos grupos de pruebas (una y dos vías de clasificación), para comparar pares particulares de poblaciones distintas. Se han calculado las eficiencias Pitman para evaluar la eficiencia con respecto a las pruebas análogas normales y las eficiencias relativas de Bahadur que comparan prueba contra prueba. Para el caso de las comparaciones contra un control para una vía de clasificación sólo se calculó la eficiencia Pitman debido a que se evaluó una sola prueba. Igualmente para las pruebas de dos vías de clasificación. Ambas eficiencias se obtuvieron comparando pruebas con tratamientos no pertenecientes al mismo grupo homogéneo y se designaron como “comparaciones entre grupos homogéneos”.

DESCRIPCIÓN DE LOS MÉTODOS

Existen muchos procedimientos no paramétricos para comparaciones múltiples tanto para diseños de una vía de clasificación como para dos vías de clasificación. A continuación se describirá brevemente cada uno de los procedimientos que han sido comparados en este trabajo.

Comparaciones múltiples para una vía de clasificación

1.- Prueba de Dunn

Cuando el valor del estadístico KW de Kruskal-Wallis obtenido es significativo, se indica que al menos uno de los grupos es diferente de al menos otro de los grupos o tratamientos. Para determinar los pares de grupos que son diferentes se determinan las diferencias $|\overline{R_u} - \overline{R_v}|$ para todos los pares de grupos. Cuando el tamaño de la muestra es grande, estas diferencias se distribuyen aproximadamente de manera normal. Sin embargo, las diferencias no son independientes y el procedimiento de comparación debe ajustarse apropiadamente. Se prueba la significación de los pares individuales usando la siguiente desigualdad:

$$|\overline{R_u} - \overline{R_v}| \geq z\alpha_{/(k-1)}\sqrt{\frac{N(N+1)}{12}\left(\frac{1}{n_u} + \frac{1}{n_v}\right)} \qquad (16)$$

2.- Método basado en la comparación de pares rankeados (Dwass, Steel, Critchlow_Fligner).

Es una prueba conservativa, utilizada luego del rechazo de la hipótesis nula en una prueba de Kruskal-Wallis. Es una prueba basada en rango pareados diseñada para encontrar diferencias entre pares de efectos de tratamiento (τ_i, τ_j) para $i < j$.

Para cada par de tratamientos (i, j), sea $W_{ij} = \sum_{b=1}^{n_j} R_{ib}$, para $1 \leq i < j \leq k$, donde $R_{i1}, \dots, R_{in}$ son los rangos de $X_{1j}, \dots X_{nj}$, respectivamente, de las muestras i-ésima y j-ésima y donde W_{ij} es la suma de los rangos de Wilcoxon de los rangos de la j-ésima muestra en el rankeo conjunto de las observaciones de las muestras "i" y "j". Se calcula entonces ${W_{ij}}^*$, es decir la versión estandarizada de W_{ij} bajo la H_0 multiplicada por $\sqrt{2}$

$${W_{ij}}^* = \sqrt{2}\left[\frac{W_{ij} - E_0(W_{ij})}{\{Var_0(W_{ij})\}^{1/2}}\right] = \frac{W_{ij} - \frac{n_j(n_i + n_j + 1)}{2}}{\left\{\frac{n_i n_j(n_i + n_j + 1)}{24}\right\}^{1/2}}, \qquad para\ 1 \leq i < j \leq k$$

(17)

Con un error experimental α para un par de tratamientos (u, v) siendo $1 \leq u < v \leq k$, se decide que $\tau_u \neq \tau_v$ si $|W_{uv}^*| \geq W_\alpha^*$.

donde la constante W_α^* es escogida para hacer que el error experimental sea igual a α, es decir, que W_α^* satisface:

$$P_0(|W_{uv}^*| < W_\propto^*, u = 1, \dots, k-1; v = u+1, \dots, k) = 1 - \propto \qquad (18)$$

Aproximación para muestras grandes

Siendo cierta la hipótesis nula, el vector de $k(k-1)/_2$ componentes $\left(W_{12}^{*}, W_{13}^{*}, \ldots, W_{k-1,k}^{*}\right)$, tiene como $min(n_1, \ldots, n_k)$ entonces, sea $W^{*} = \left(W_{12}^{*}, W_{13}^{*}, \ldots, W_{k-1,k}^{*}\right)$ donde W_{ij}^{*} está dado por la ecuación (17), se observa que W^{*} tiene distribución normal multivariada cuando el $min(n_1, \ldots, n_k)$ tiende a infinito con vector de medias 0. Entonces la constante W_{α}^{*} se aproxima a q_{α} ; es decir, el percentil α-ésimo para la distribución de los rangos de k variables independientes $N(0,1)$, se decide que $\tau_u \neq \tau_v$ si $|W_{uv}^{*}| \geq q_{\alpha}$. Los valores de q_{α} se encuentran tabulados en Anexo A Hollander y Wolfe (1999).

Aproximación en rankeo conjunto

El procedimiento descrito se basa en $k(k-1)/_2$ rankeos pareados para dos muestras; sin embargo es razonable considerar las comparaciones basadas en el rankeo conjunto de las "N" observaciones. Sea $R_{.j}, j = 1, \ldots, k$ los rangos promedios para el tratamiento j-ésimo del rankeo conjunto de las "N" observaciones, el procedimiento análogo será

se decide que: $\tau_u \neq \tau_v$ si $N^{*}|R_{.u} - R_{.v}| \geq y_{\alpha}$ (19)

donde N^{*} es el mínimo común múltiplo de $(n_1, \ldots, n_k)$ y la constante y_{α} es se escoge de manera que el error experimental sea igual a α, es decir que y_{α} satisface la restricción

$$P_0(N^{*}|R_{.u} - R_{.v}| < y_{\alpha}, u = 1, \ldots, k-1; v = u+1, \ldots, k) = 1 - \propto \quad (20)$$

La expresión general (20) para distintos tamaños de muestra fue considerada por Damico y Wolfe (1987).

3.- Comparación de una cola para todos los pares basados en rangos ordenados (Hayter-Stone).

Es una prueba apropiada para datos analizados por una vía de clasificación luego del rechazo de la hipótesis nula usando el procedimiento de Jonckheere-Terpstrta. Se usa el mismo estadístico W_{ij}^* de la ecuación (17) de la prueba general, y se usa el criterio:

$$\tau_v > \tau_u \text{ si } W_{uv}^* \geq c_\alpha^* \text{ ; de otra forma } \tau_v = \tau_u \quad (21)$$

donde la constante $c_{\propto^*}$ es se escoge de manera que el error experimental sea igual a α: es decir que $c_{\propto^*}$ satisface la restricción

$$P_0(W_{uv}^* < c_\propto^*, u = 1, \dots, k-1; v = u+1, \dots, k) = 1 - \propto \quad (22)$$

Puede encontrarse que no exista diferencia entre los tratamientos a pesar de que la prueba de Jonckheere-Terpstra si los haya detectado. Esto ocurre con frecuencia por la naturaleza conservativa de la prueba.

4.- Comparaciones contra un control basada en rangos conjuntos (Nemenyi, Damico-wolfe).

Es un método para tomar decisiones acerca de las diferencias individuales entre el efecto medio de un control simple o un control de base y los efectos medios para cada uno de los $k-1$ tratamientos restantes.

Este método se basa en el rankeo conjunto de todas las observaciones de la muestra y puede ser utilizado para análisis de una vía de clasificación. El factor adicional en esta prueba es que aquí no se comparan todos los tratamientos sino cada uno con el control.

Sea 1 el tratamiento control, y N^* el mínimo común múltiplo del tamaño de las muestras $(n_1, \ldots, n_k)$. Se rankean todas las observaciones juntas y siendo $R_{.1}, \ldots, R_{.k}$ los promedios de esos rangos asociados a los tratamientos $1, \ldots, k$. para cada $k-1$ tratamientos menos el control, se calcula la diferencia $R_{.u} - R_{.1},\ u = 2, \ldots, k$ y decida por el criterio:

$\tau_u > \tau_1$ si $N^*(R_{.u} - R_{.1}) \geq y_{\propto}^*$ de otra forma $\tau_u = \tau_1$ (23)

Donde la constante $y_{\propto}^*$ se escoge de manera que el error experimental sea igual a α; es decir que $y_{\propto}^*$ satisface la restricción:

$$P_0(N^*(R_{.u} - R_{.1}) < y_{\propto}^*, \quad u = 2, \ldots, k) = 1 - \propto \tag{24}$$

Aproximación para muestras grandes

Siendo cierta la hipótesis nula, el vector de componentes $(k-1)$ $(R_{.2} - R_{.1}, R_{.3} - R_{.1}, \ldots, R_{.k} - R_{.1})$, tiene como $min(n_1, \ldots, n_k)$ tiende a infinito, con distribución asintótica normal $(k-1)$variada con vector de medias 0. Si se tiene igual número de tratamientos se decide que:

$$\tau_u > \tau_1 \text{ si } (R_{.u} - R_{.1}) \geq m_{\propto}^* \left[\frac{N(N+1)}{12}\right]^{1/2} \left(\frac{1}{b} + \frac{1}{n}\right)^{1/2} \tag{25}$$

de otra forma se decide que $\tau_u = \tau_1$, $u = 2, \ldots, k$

Los valores de $m_{\propto}^*$ se encuentran tabulados.

Para muestras de tamaño arbitrario, no necesariamente iguales decida

$$\tau_u > \tau_1 \ \text{ si } \ (R_{.u} - R_{.1}) \geq z_{\propto^*} \left[\frac{N(N+1)}{12}\right]^{1/2} \left(\frac{1}{n_1} + \frac{1}{n_u}\right)^{1/2} \quad (26)$$

donde $\propto^* = \propto/(k-1)$

Existen otras pruebas basadas en contrastes que no serán consideradas para este estudio.

5.- Prueba basada en el estadístico de Baumgartner, Weiß y Schindler (1998).

El trabajo presentado por Baumgartner *et al.* (1998), consiste en una prueba basada en rangos. Sean $X_1 \leq \cdots \leq X_n$ y $Y_1 \leq \cdots \leq Y_m$ las notaciones de las observaciones independientes de dos grupos que van a ser comparados, y sean $R_1 \leq \cdots \leq R_n$ y $H_1 \leq \cdots \leq H_m$ las notaciones de los rangos de las muestras combinadas en orden de magnitud creciente en los dos grupos considerados"*i*" y "*j*". Esta prueba es una alternativa a la prueba de Wilcoxon para dos muestras. El estadístico de prueba propuesto es:

$$B = \frac{1}{2}(B_X + B_Y) \quad (27)$$

$$B_X = \frac{1}{n}\sum_{i=1}^{n} \frac{\left(R_i - \frac{m+n}{n}.i\right)^2}{\frac{i}{n+1}\left(1-\frac{i}{n+1}\right)\frac{m(m+n)}{n}} \quad (28)$$

$$B_Y = \frac{1}{m}\sum_{j=1}^{m} \frac{\left(H_j - \frac{m+n}{m}.j\right)^2}{\frac{j}{m+1}\left(1-\frac{j}{m+1}\right)\frac{n(m+n)}{m}} \quad (29)$$

Valores grandes de B indican que $\tau_i \neq \tau_j$. Baumgartner *et al* (1998), utilizaron la distribución asintótica de B para realizar comparaciones contra otras pruebas no paramétricas, concluyendo que esta nueva prueba de rangos es al menos tan poderosa como la prueba de Wilcoxon. En lugar de una prueba asintótica, la prueba exacta de permutación se puede desarrollar para generar la distribución nula completa por permutación basada en estadísticos de rango.

Al comparar pruebas exactas, la prueba basada en el estadístico B es menos conservativa y más poderosa que la de Wilcoxon, de acuerdo a los estudios simulados propuestos por Neuhäusser (2004). De acuerdo a sus resultados, sugiere que se podría reemplazar el estadístico de Wilcoxon por el de Baumgartner *et al* (1998), en la prueba de Steel-Dwass.

Luego se decide que $\tau_i \neq \tau_j$ para cada par con $B \geq b_\propto$, donde $b_\propto$ se escoge tal que:

$$P(maxB \geq b_\propto) \leq \propto \tag{30}$$

Los valores de $b_\propto$ se determinaron por permutaciones para el caso de dos muestras. Para casos de tres o más muestras, las pruebas de simulación fueron realizadas tomando una muestra aleatoria simple de todas las posibles permutaciones. Se generaron 50000 permutaciones. En el Anexo B se presentan las tablas con los valores de $b_\propto$ tabulados hasta 10 tratamientos.

Comparaciones múltiples para dos vías de clasificación

1.- Prueba bilateral de comparaciones basadas en la suma de rangos de Friedman. (Wilcoxon, Nemenyi, McDonald-Thompson).

En esta prueba se sacan conclusiones para todos los $\binom{k}{2} = k(k-1)/2$ pares de efectos de tratamientos y estas conclusiones son bilaterales por naturaleza.

Sean $R_1, \dots R_k$ las sumas de tratamientos dentro de los bloques, se calculan las $k(k-1)/2$ diferencias absolutas $|R_u - R_v|$; $1 \leq u < v \leq k$ entonces, se decide que $\tau_u \neq \tau_v$ si $|R_u - R_v| \geq r_\propto$ (31)

Donde $r_\propto$ se escoge para hacer que el error sea igual a α, y satisface la restricción

$$P_0(|R_u - R_v| < r_\propto \quad u = 2, \ldots, k-1; v = u+1, \ldots, k) = 1 - \propto \qquad (32)$$

La ecuación (32) estipula que las $k(k-1)/2$ desigualdades $|R_u - R_v| < r_\propto$, corresponden a todos los pares (u, v) de tratamientos con $u < v$ y poseen simultáneamente la probabilidad $1-\propto$ cuando H_0 es cierta. Los valores $r_\propto$ se encuentran tabulados para todas las combinaciones $k = 3(1)15$ y $n = 3(1)15$.

Aproximación para grandes muestras

Cuando H_0 es cierta, los $k(k-1)/2$ vectores de componentes $(R_1, \ldots R_k)$, cuando "n" tiende a infinito, tienen distribución asintótica normal (k-1)-variada con un vector de medias y matriz de covarianza apropiados. Cuando el número de bloques es grande, el valor crítico $r_\propto$ puede ser aproximado por $\left[{nk(k+1)}/{12}\right]^{1/2} q_\propto$, donde $q_\propto$ es el α-ésimo percentil superior para la distribución de los rangos de las "k" variables independientes $N(0,1)$.

se decide que: $\tau_u \neq \tau_v$ s $|R_u - R_v| \geq q_\propto \left[\frac{nk(k+1)}{12}\right]^{1/2}$; (33)

si no, se decide que: $\tau_u = \tau_v$; esta es una prueba conservativa debido al uso del error experimental.

2.- Prueba unilateral. Comparación versus control basada en la suma de los rangos de Friedman. (Nemenyi, Wilcoxon-Wilcox, Miller).

Este procedimiento se aplica luego del rechazo de H_0 al utilizar las pruebas de Friedman o Page. Las conclusiones obtenidas de cada uno de las $k-1$ diferencias entre tratamientos y el efecto del control son de naturaleza unilateral.

sea $R_1, \dots R_k$ las sumas de los rangos dentro de los tratamientos. Calcule las $(k-1)$ diferencias ; $(R_u - R_1)$, $u = 1, \dots, k$ entonces,

se decide que: $\tau_u > \tau_v$ si $(R_u - R_1) \geq r_\propto^*$ (34)

donde $r_\propto^*$ se escoge para hacer que la tasa de error experimental sea α; esto es, que $r_\propto^*$ satisfaga la restricción:

$$P_0\big((R_u - R_1) < r_\propto^*, u = 2, \dots k\big) = 1 - \propto \qquad (35)$$

donde la probabilidad $P_0(.)$ se calcula bajo la $H_0\colon \tau_1 = \cdots = \tau_k$. La ecuación (35) estipula que las $(k-1)$ desigualdades $(R_u - R_1) < r_\propto^*$ correspondiente a cada tratamiento con el control, tienen simultáneamente la probabilidad $1-\propto$ cuando la H_0 es cierta. Los valores de $r_\propto^*$ se encuentran tabulados.

Este procedimiento se puede ajustar para cuando el interés consiste en decidir si un efecto de tratamiento es menor que el efecto del control.

3.- Prueba bilateral de todos los tratamientos basada en rangos asignados. (Nemenyi)

La prueba basada en los rangos asignados de Wilcoxon, la cual está diseñada para tomar decisiones sobre diferencias individuales entre pares de efectos de tratamientos (τ_u, τ_v) para $u > v$.

Este procedimiento se aplicará datos de diseños de dos vías de clasificación con 1 observación por celda luego de rechazar la H_0 con el procedimiento de Doksum-Lehman.

$$T_{uv}` = máx\ \{T_{uv}[n(n+1)/2] - T_{uv}\}, 1 \leq u < v \leq k \tag{36}$$

Se decide que $\tau_u \neq \tau_v$ si $T_{uv}^{`} \geq t_{\propto}^{`}$ siendo $t_{\propto}^{`}$ una constante escogida para que el error experimental sea igual a α. Es decir $t_{\propto}^{`}$ satisface la ecuación

$$P_0\{T_{uv}` < t_{\propto}^{`}, u = 1, \dots, k-1 \quad y\ v = u+1, \dots, k\} \approx 1{-}\propto \tag{37}$$

Esta prueba no posee propiedades de distribución libre ni de distribución libre asintótica, cuando $n \to \infty$ la correlación nula entre T_{12} y T_{13}, por decir algo, es $\frac{1}{2}$ soportado en el trabajo de Lehman (1964).

4.- Prueba basada en los rangos asignados de Wilcoxon para alternativas ordenadas en un diseño de bloques completos aleatorizados. (Hollander).

Es un procedimiento conservativo basado en los pares de rangos asignados para probar la $H_0: \tau_1 = \cdots = \tau_k$ contra la alternativa ordenada $H_1 = \tau_1 \leq \tau_2 \leq \dots \leq \tau_k$ con al menos una desigualdad.

Para cada par d tratamientos $k(k-1)/2$ se calcula el estadístico de rango T_{uv}. Para calcular el estadístico Q de Hollander sea

$$Y = \sum_{u=1}^{k-1} \sum_{v=u+1}^{k} T_{uv} \tag{38}$$

El valor esperado nulo de Y está dado por

$$E_0(Y) = \frac{nk(k-1)(n+1)}{8} \tag{39}$$

La varianza nula $Var_0(Y) = \frac{nk(n+1)(2n+1)(k-1)\{3+2(k-2)\rho_U^n\}}{144}$ (40)

El estadístico Hollander para la prueba conservativa es entonces

$$Q = \frac{Y - E_0(Y)}{\{Var_U(Y)\}^{1/2}} \tag{41}$$

Se rechaza la H_0 si $Q \geq z_{\propto}$ de otra forma no se rechaza.

RESULTADOS

Cuadro 4. Tasas de rechazo y eficiencia Pitman para 15 tratamientos todos los pares Diseño C.A. entre grupos homogéneos.

			Tasa de rechazo										Eficiencia Pitman							
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey		S-D		Cr-Fl		Hy-St		Bau-W	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
15 TRATAMIENTOS. TODOS LOS PARES.	1	2	1	1	0	0	0	0	1	0.95	0.6	0.45	1.68	2.2	0	0	0	0	1.68	2.08
	1	3	1	1	0	0	0	0	1	1	0.94	0.84	1.06	1.18	0	0	0	0	1.06	1.18
	1	4	1	1	0	0	0	0	0.99	0.96	0.85	0.81	1.18	1.23	0	0	0	0	1.16	1.18
	1	5	1	1	0	0	0	0	1	1	1	1	1	1.01	0	0	0	0	1	1.01
	1	6	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	1	7	1	1	0	0	0	0	1	1	0.98	0.95	1.02	1.05	0	0	0	0	1.02	1.05
	1	8	1	1	0	0	0	0	1	1	0.97	0.92	1.04	1.09	0	0	0	0	1.04	1.09
	1	10	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	1	11	0	0	0	0	0	0	1	1	0.72	0.48	0	0	0	0	0	0	1.39	2.07
	1	12	1	1	0	0	0	0	1	1	0.98	0.92	1.02	1.09	0	0	0	0	1.02	1.09
	1	13	1	1	0	0	0	0	1	1	0.99	0.99	1.01	1.01	0	0	0	0	1.01	1.01
	1	14	1	1	0	0	0	0	1	1	0.94	0.87	1.07	1.15	0	0	0	0	1.07	1.15
	1	15	1	1	0	0	0	0	1	1	0.99	0.98	1.01	1.02	0	0	0	0	1.01	1.02
	2	5	0.97	0.97	0	0	0	0	0.9	0.84	0.14	0.07	7.15	14.5	0	0	0	0	6.63	12.5
	2	6	0.97	0.97	0	0	0	0	0.93	0.89	0.22	0.13	4.4	7.38	0	0	0	0	4.24	6.79
	2	9	0.2	0.14	0	0	0	0	0.87	0.74	0	0	50	no	0	IND	0	IND	218	no
	2	10	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	2	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	3	5	0.68	0.68	0	0	0	0	0.86	0.78	0.01	0	76	684	0	0	0	0	95	784
	3	6	0.7	0.7	0	0	0	0	0.97	0.92	0.05	0.03	15.3	28.1	0	0	0	0	21.1	37
	3	9	0.56	0.53	0	0	0	0	0.87	0.79	0.03	0.02	16.9	22	0	0	0	0	26.5	32.8
	3	10	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	3	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	4	5	0.33	0.33	0	0	0	0	0.88	0.81	0.01	0	66.8	167	0	0	0	0	177	406
	4	6	0.39	0.39	0	0	0	0	0.87	0.79	0.02	0.01	19.3	35	0	0	0	0	43.5	71.5
	4	8	0.35	0.35	0	0	0	0	0.85	0.78	0.01	0	70.2	176	0	0	0	0	170	388
	4	10	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	4	11	1	1	0	0	0	0	1	0.99	1	1	1	1	0	0	0	0	1	0.99
	4	13	0.28	0.28	0	0	0	0	0.85	0.78	0	0	93	140	0	0	0	0	282	389
	5	7	0.39	0.39	0	0	0	0	0.87	0.8	0	0	98.5	197	0	0	0	0	218	402
	5	8	0.61	0.61	0	0	0	0	0.9	0.83	0.01	0.01	75.8	121	0	0	0	0	113	167
	5	9	0.94	0.94	0	0	0	0	0.95	0.91	0.35	0.29	2.71	3.22	0	0	0	0	2.73	3.13
	5	10	1	1	0	0	0	0	1	1	1	0.98	1	1.02	0	0	0	0	1	1.02
	5	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	5	12	0.6	0.6	0	0	0	0	0.91	0.84	0.01	0	60.3	151	0	0	0	0	90.5	210
	5	13	0.26	0.26	0	0	0	0	0.86	0.79	0	0	no	no	IND	IND	IND	IND	no	no
	5	14	0.6	0.6	0	0	0	0	0.85	0.8	0.01	0	75	200	0	0	0	0	107	265
	5	15	0.33	0.33	0	0	0	0	0.87	0.78	0	0	no	no	IND	IND	IND	IND	no	no
	6	7	0.46	0.46	0	0	0	0	0.89	0.81	0.01	0	91	455	0	0	0	0	178	813
	6	8	0.67	0.67	0	0	0	0	0.89	0.81	0.02	0.01	39.5	74.6	0	0	0	0	52.1	89.9
	6	9	0.99	0.99	0	0	0	0	0.92	0.86	0.35	0.28	2.8	3.58	0	0	0	0	2.63	3.14
	6	10	1	1	0	0	0	0	1	1	0.98	0.94	1.02	1.06	0	0	0	0	1.02	1.06
	6	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	6	12	0.66	0.66	0	0	0	0	0.92	0.86	0.03	0.02	23.6	44	0	0	0	0	32.9	57.4
	6	13	0.37	0.37	0	0	0	0	0.89	0.82	0	0	184	IND	0	IND	0	IND	446	no
	6	14	0.63	0.63	0	0	0	0	0.95	0.89	0.04	0.02	17.1	33.2	0	0	0	0	25.5	46.9
	6	15	0.37	0.37	0	0	0	0	0.88	0.81	0	0	124	IND	0	IND	0	IND	292	no
	7	9	0.87	0.85	0	0	0	0	0.87	0.8	0.11	0.08	8.27	10.2	0	0	0	0	8.29	9.69
	7	10	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	7	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	8	9	0.64	0.32	0	0	0	0	0.84	0.77	0.03	0.02	22.8	18	0	0	0	0	30.1	42.7
	8	10	1	0.85	0	0	0	0	1	1	1	1	1	0.85	0	0	0	0	1	1
	8	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	9	10	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	9	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	9	12	0.62	0.6	0	0	0	0	0.89	0.82	0.05	0.04	11.9	17.1	0	0	0	0	17.2	23.5
	9	13	0.87	0.86	0	0	0	0	0.93	0.88	0.25	0.22	3.5	4.01	0	0	0	0	3.76	4.07
	9	14	0.61	0.59	0	0	0	0	0.91	0.83	0.08	0.06	7.87	9.7	0	0	0	0	11.7	13.7

Cuadro 5. Tasas de rechazo y eficiencias Bahadur para 15 tratamientos todos los pares Diseño C.A. entre grupos homogéneos.

			Tasa de rechazo										Eficiencias Bahadur											
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey		CR-FL / S-D		HY-ST/S-D		S-D / Bau-W		CR-FL / HY-ST		CR-FL / Bau-W		HY-ST / Bau-W	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
15 TRATAMIENTOS. TODOS LOS PARES.	1	2	1	1	0	0	0	0	1	0.946	0.595	0.454	0	0	0	0	1	1.057	IND	IND	0	0	0	0
	1	3	1	1	0	0	0	0	1	1	0.94	0.844	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	1	4	1	1	0	0	0	0	0.989	0.96	0.849	0.813	0	0	0	0	1.011	1.042	IND	IND	0	0	0	0
	1	5	1	1	0	0	0	0	1	1	0.999	0.995	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	1	6	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	1	7	1	1	0	0	0	0	1	1	0.982	0.948	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	1	8	1	1	0	0	0	0	1	1	0.966	0.919	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	1	10	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	1	11	0	0	0	0	0	0	1	1	0.717	0.482	IND	IND	IND	IND	0	0.000	IND	IND	0	0	0	0
	1	12	1	1	0	0	0	0	1	1	0.983	0.917	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	1	13	1	1	0	0	0	0	1	1	0.994	0.986	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	1	14	1	1	0	0	0	0	1	1	0.938	0.872	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	1	15	1	1	0	0	0	0	1	1	0.993	0.976	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	2	5	0.973	0.973	0	0	0	0	0.902	0.837	0.136	0.067	0	0	0	0	1.079	1.162	IND	IND	0	0	0	0
	2	6	0.967	0.967	0	0	0	0	0.933	0.89	0.22	0.131	0	0	0	0	1.036	1.087	IND	IND	0	0	0	0
	2	9	0.2	0.143	0	0	0	0	0.872	0.74	0.004	0	0	0	0	0	0.229	0.193	IND	IND	0	0	0	0
	2	10	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	2	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	3	5	0.684	0.684	0	0	0	0	0.855	0.784	0.009	0.001	0	0	0	0	0.8	0.872	IND	IND	0	0	0	0
	3	6	0.702	0.702	0	0	0	0	0.971	0.924	0.046	0.025	0	0	0	0	0.723	0.760	IND	IND	0	0	0	0
	3	9	0.557	0.527	0	0	0	0	0.874	0.787	0.033	0.024	0	0	0	0	0.637	0.670	IND	IND	0	0	0	0
	3	10	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	3	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	4	5	0.334	0.334	0	0	0	0	0.884	0.812	0.005	0.002	0	0	0	0	0.378	0.411	IND	IND	0	0	0	0
	4	6	0.385	0.385	0	0	0	0	0.87	0.787	0.02	0.011	0	0	0	0	0.443	0.489	IND	IND	0	0	0	0
	4	8	0.351	0.351	0	0	0	0	0.849	0.776	0.005	0.002	0	0	0	0	0.413	0.452	IND	IND	0	0	0	0
	4	10	1	1	0	0	0	0	1	1	1	0.997	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	4	11	1	1	0	0	0	0	1	0.989	1	1	0	0	0	0	1	1.011	IND	IND	0	0	0	0
	4	13	0.279	0.279	0	0	0	0	0.846	0.778	0.003	0.002	0	0	0	0	0.33	0.359	IND	IND	0	0	0	0
	5	7	0.394	0.394	0	0	0	0	0.873	0.804	0.004	0.002	0	0	0	0	0.451	0.490	IND	IND	0	0	0	0
	5	8	0.606	0.606	0	0	0	0	0.902	0.834	0.008	0.005	0	0	0	0	0.672	0.727	IND	IND	0	0	0	0
	5	9	0.936	0.936	0	0	0	0	0.945	0.91	0.346	0.291	0	0	0	0	0.99	1.029	IND	IND	0	0	0	0
	5	10	1	1	0	0	0	0	1	1	0.997	0.977	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	5	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	5	12	0.603	0.603	0	0	0	0	0.905	0.839	0.01	0.004	0	0	0	0	0.666	0.719	IND	IND	0	0	0	0
	5	13	0.262	0.262	0	0	0	0	0.859	0.785	0	0	0	0	0	0	0.305	0.334	IND	IND	0	0	0	0
	5	14	0.6	0.6	0	0	0	0	0.854	0.795	0.008	0.003	0	0	0	0	0.703	0.755	IND	IND	0	0	0	0
	5	15	0.333	0.333	0	0	0	0	0.866	0.784	0	0	0	0	0	0	0.385	0.425	IND	IND	0	0	0	0
	6	7	0.455	0.455	0	0	0	0	0.892	0.813	0.005	0.001	0	0	0	0	0.51	0.560	IND	IND	0	0	0	0
	6	8	0.671	0.671	0	0	0	0	0.885	0.809	0.017	0.009	0	0	0	0	0.758	0.829	IND	IND	0	0	0	0
	6	9	0.987	0.985	0	0	0	0	0.924	0.864	0.352	0.275	0	0	0	0	1.068	1.140	IND	IND	0	0	0	0
	6	10	1	1	0	0	0	0	1	1	0.976	0.939	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	6	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	6	12	0.66	0.66	0	0	0	0	0.921	0.861	0.028	0.015	0	0	0	0	0.717	0.767	IND	IND	0	0	0	0
	6	13	0.367	0.367	0	0	0	0	0.891	0.821	0.002	0	0	0	0	0	0.412	0.447	IND	IND	0	0	0	0
	6	14	0.631	0.631	0	0	0	0	0.945	0.892	0.037	0.019	0	0	0	0	0.668	0.707	IND	IND	0	0	0	0
	6	15	0.373	0.373	0	0	0	0	0.877	0.809	0.003	0	0	0	0	0	0.425	0.461	IND	IND	0	0	0	0
	7	9	0.868	0.846	0	0	0	0	0.87	0.804	0.105	0.083	0	0	0	0	0.998	1.052	IND	IND	0	0	0	0
	7	10	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	7	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	8	9	0.637	0.324	0	0	0	0	0.843	0.769	0.028	0.018	0	0	0	0	0.756	0.421	IND	IND	0	0	0	0
	8	10	1	0.846	0	0	0	0	1	1	1	1	0	0	0	0	1	0.846	IND	IND	0	0	0	0
	8	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	9	10	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	9	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1.000	IND	IND	0	0	0	0
	9	12	0.617	0.599	0	0	0	0	0.894	0.824	0.052	0.035	0	0	0	0	0.69	0.727	IND	IND	0	0	0	0
	9	13	0.868	0.863	0	0	0	0	0.933	0.876	0.248	0.215	0	0	0	0	0.93	0.985	IND	IND	0	0	0	0
	9	14	0.614	0.592	0	0	0	0	0.914	0.833	0.078	0.061	0	0	0	0	0.672	0.711	IND	IND	0	0	0	0

En el cuadro 4 se puede observar que las tasas de rechazo son muy altas e incluso mayores a 1 en algunos casos en la prueba conservativa de Baumgartner-Weiss y Schindler, con los dos niveles de significación estudiados, seguida por la prueba de Dunn que mostró bajas tasas de rechazo en poblaciones distantes donde se entiende como “distante” aquellas poblaciones en las cuales sus medianas se encuentran separadas por más de tres poblaciones, según lo establece Dunn, (1964). Y cero en poblaciones cercanas como 1 y 11 en las cuales nunca se obtuvo un rechazo, Las pruebas de Steel-Dwass, Critchlow y Fligner y Hayter-Stone no evidenciaron tasas de rechazo diferentes a cero bajo ninguno de los niveles de significación estudiados.

En relación a la eficiencia Pitman, ésta resultó alta, 1 o mayor a 1 en algunos casos en las pruebas de rankeo conjunto y 0 para la de rankeo pareado, con la excepción de Baungartner-Weiss y Schindler, que tiene altas eficiencias ya que es menos conservativa que las otras de rankeo pareado, resultando la prueba de Bumgartner-Weiss tan efciente como la de rankeo conjunto de Dunn.

Las eficiencias Bahadur, reportadas en el cuadro 5 mostraron valores mayores a 1 solamente en las pruebas de rankeo conjunto y poblaciones cercanas. En las demás no se evidenció eficiencia alguna. La prueba de rankeo pareado de Baumgartner-Weiss resultó ser tan eficiente como la de rankeo conjunto de Dunn en el sentido Bahadur.

Cuadro 6. Tasas de error Tipo I para 15 tratamientos todos los pares Diseño C.A. dentro de los grupos homogéneos.

Tasas de error Tipo I												
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
15 TRATMIENTOS. TODOS LOS PARES.	1	9	0.94	0.91	0	0	0	0	0.94	0.89	0.27	0.19
	2	3	0.37	0.37	0	0	0	0	0.85	0.64	0	0
	2	4	0.78	0.78	0	0	0	0	0.9	0.84	0.13	0.07
	2	7	0.78	0.78	0	0	0	0	0.85	0.77	0.04	0.01
	2	8	0.46	0.46	0	0	0	0	0.9	0.7	0.02	0
	2	12	0.47	0.47	0	0	0	0	0.86	0.68	0	0
	2	13	0.85	0.85	0	0	0	0	0.89	0.83	0.12	0.08
	2	14	0.48	0.48	0	0	0	0	0.83	0.65	0	0
	2	15	0.81	0.81	0	0	0	0	0.87	0.8	0.11	0.06
	3	4	0.45	0.45	0	0	0	0	0.88	0.82	0.01	0
	3	7	0.36	0.36	0	0	0	0	0.86	0.8	0	0
	3	8	0.12	0.12	0	0	0	0	0.84	0.73	0	0
	3	12	0.22	0.22	0	0	0	0	0.85	0.77	0	0
	3	13	0.5	0.5	0	0	0	0	0.89	0.83	0.01	0.01
	3	14	0.12	0.12	0	0	0	0	0.82	0.69	0	0
	3	15	0.45	0.45	0	0	0	0	0.88	0.83	0.01	0
	4	7	0.37	0.37	0	0	0	0	0.91	0.84	0	0
	4	8	0.35	0.35	0	0	0	0	0.85	0.78	0.01	0
	4	12	0.51	0.51	0	0	0	0	0.92	0.86	0.01	0.01
	4	13	0.28	0.28	0	0	0	0	0.85	0.78	0	0
	4	14	0.44	0.44	0	0	0	0	0.89	0.82	0.01	0
	4	15	0.34	0.34	0	0	0	0	0.9	0.82	0.01	0.01
	5	6	0.33	0.33	0	0	0	0	0.94	0.88	0.01	0
	7	8	0.32	0.32	0	0	0	0	0.87	0.81	0	0
	7	12	0.37	0.37	0	0	0	0	0.9	0.84	0.01	0
	7	13	0.24	0.24	0	0	0	0	0.84	0.77	0	0
	7	14	0.32	0.32	0	0	0	0	0.87	0.79	0	0
	7	15	0.31	0.31	0	0	0	0	0.89	0.82	0	0
	8	12	0.23	1	0	0	0	0	0.87	0.79	0	0
	8	13	0.45	0.37	0	0	0	0	0.86	0.8	0.01	0
	8	14	0.3	0.24	0	0	0	0	0.86	0.81	0	0
	8	15	0.39	0.32	0	0	0	0	0.88	0.81	0.01	0
	12	13	0.49	0.49	0	0	0	0	0.93	0.88	0.01	0.01
	12	14	0.17	0.17	0	0	0	0	0.84	0.75	0	0
	12	15	0.39	0.39	0	0	0	0	0.88	0.81	0.01	0
	13	14	0.46	0.46	0	0	0	0	0.89	0.84	0.01	0.01
	13	15	0.33	0.33	0	0	0	0	0.88	0.82	0	0
	14	15	0.36	0.36	0	0	0	0	0.83	0.78	0	0

En el cuadro 6 se puede observar que las tasas de error tipo I son muy altas e incluso cercanas a 1 en algunos casos en la prueba conservativa de Baumgartner-Weiss y

Schindler en los dos niveles de significación estudiados; sin embargo muestra baja tasa de error Tipo I en el nivel de significación 0,05 en poblaciones distantes. La prueba de Dunn mostró tasas de rechazo inconsistentes en poblaciones distantes, o cercanas donde no se evidencia variación en las tasas de error Tipo I independientemente de la “distancia” existente entre las poblaciones. La prueba de Dunn es liberal; es decir trabaja con error familiar y es posible que esto explique su comportamiento, mientras que la prueba de Baungartner es de carácter conservativo. Las pruebas de Steel-Dwass, Critchlow y Fligner y Hayter-Stone no evidenciaron tasas de error tipo I bajo ninguno de los niveles de significación estudiados, probablemente debido a su carácter conservativo que provee protección extra para el error tipo I. La prueba de Tukey mostró bajas tasas de error Tipo I en algunos casos y altas en otros mostrando la misma tendencia que la prueba de Dunn.

Cuadro 7. Tasas de rechazo y eficiencia Pitman para 10 tratamientos contrastantes, diseño C.A. entre grupos homogéneos.

			Tasa de rechazo										Eficiencia Pitman							
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey		S-D		Cr-Fl		Hy-St		Bau-W	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
10 TRATAMIENTOS CONTRASTANTES	1	2	1	1	0	0	0	0	1	0	0.646	0.574	1.548	1.742	0	0	0	0	1.548	0
	1	3	1	1	0	0	0	0	1	0	0.965	0.91	1.036	1.099	0	0	0	0	1.036	0
	1	4	1	1	0	0	0	0	0.989	0	0.871	0.826	1.148	1.211	0	0	0	0	1.135	0
	1	6	1	1	0	0	0	0	1	0	1	1	1	1	0	0	0	0	1	0
	1	9	0.987	0.923	0	0	0	0	0.942	0	0.422	0.322	2.339	2.866	0	0	0	0	2.232	0
	1	10	1	1	0	0	0	0	1	0	1	1	1	1	0	0	0	0	1	0
	1	12	1	1	0	0	0	0	1	0	0.998	0.974	1.002	1.027	0	0	0	0	1.002	0
	1	14	1	1	0	0	0	0	1	0	0.964	0.928	1.037	1.078	0	0	0	0	1.037	0
	2	4	0.765	0.68	0	0	0	0	0.897	0.84	0.238	0.167	3.214	4.072	0	0	0	0	3.769	5.03
	2	6	0.967	0.931	0	0	0	0	0.933	0.89	0.371	0.276	2.606	3.373	0	0	0	0	2.515	3.225
	2	10	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	2	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	3	4	0.445	0.379	0	0	0	0	0.875	0.82	0.038	0.03	11.71	12.63	0	0	0	0	23.03	27.33
	3	6	0.702	0.64	0	0	0	0	0.971	0.924	0.108	0.084	6.5	7.619	0	0	0	0	8.991	11
	3	10	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	3	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	4	9	0.799	0.747	0	0	0	0	0.912	0.856	0.355	0.274	2.251	2.726	0	0	0	0	2.569	3.124
	4	10	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	4	11	1	1	0	0	0	0	1	0.989	1	1	1	1	0	0	0	0	1	0.989
	4	12	0.511	0.453	0	0	0	0	0.923	0.858	0.036	0.026	14.19	17.42	0	0	0	0	25.64	33
	4	14	0.424	0.36	0	0	0	0	0.892	0.819	0.034	0.023	12.47	15.65	0	0	0	0	26.24	35.61
	6	9	0.991	0.959	0	0	0	0	0.924	0.864	0.496	0.398	1.998	2.41	0	0	0	0	1.863	2.171
	6	10	1	0.999	0	0	0	0	1	1	0.987	0.968	1.013	1.032	0	0	0	0	1.013	1.033
	6	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	6	12	0.66	0.571	0	0	0	0	0.921	0.861	0.09	0.061	7.333	9.361	0	0	0	0	10.23	14.11
	6	14	0.631	0.545	0	0	0	0	0.945	0.892	0.1	0.071	6.31	7.676	0	0	0	0	9.45	12.56
	9	10	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	9	12	0.688	0.503	0	0	0	0	0.894	0.824	0.123	0.065	5.593	7.738	0	0	0	0	7.268	12.68
	10	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	10	12	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	10	14	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	11	12	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	11	14	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1

En el cuadro 7 se puede observar que las tasas de rechazo son muy altas e incluso cercanas a 1 en algunos casos en la prueba conservativa de Baumgartner-Weiss y Schindler, en los dos niveles de significación estudiados, seguida por la prueba de Dunn que mostró bajas tasas de rechazo en poblaciones distantes. Las pruebas de Steel-Dwass, Critchlow y Fligner y Hayter-Stone no evidenciaron tasas de rechazo diferentes de cero bajo ninguno de los niveles de significación estudiados.

Las eficiencias Pitman, reportadas en el cuadro 7 mostraron resultados altos, 1 o mayor a 1 en algunos casos en las pruebas de rankeo conjunto y 0 para las de rankeo pareado, exceptuando el caso de la prueba de Baumgartner-Weiss cuya eficiencia Pitman fue 0 para las comparaciones de poblaciones distantes.

Cuadro 8. Tasas de rechazo y eficiencias Bahadur para 10 tratamientos contrastantes, diseño C.A. entre grupos homogéneos.

			Tasa de rechazo										Eficiencias Bahadur											
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey		CR-FL / S-D		HY-ST/S-D		S-D / Bau-W		CR-FL / HY-ST		CR-FL / Bau-W		HY-ST / Bau-W	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
10 TRATAMIENTOS CONTRASTANTES	1	2	1	1	0	0	0	0	1	0	0.646	0.574	0	0	0	0	1	no	IND	IND	0	IND	0	IND
	1	3	1	1	0	0	0	0	1	0	0.965	0.91	0	0	0	0	1	no	IND	IND	0	IND	0	IND
	1	4	1	1	0	0	0	0	0.989	0	0.871	0.826	0	0	0	0	1.011	no	IND	IND	0	IND	0	IND
	1	6	1	1	0	0	0	0	1	0	1	1	0	0	0	0	1	no	IND	IND	0	IND	0	IND
	1	9	0.987	0.923	0	0	0	0	0.942	0	0.422	0.322	0	0	0	0	1.048	no	IND	IND	0	IND	0	IND
	1	10	1	1	0	0	0	0	1	0	1	1	0	0	0	0	1	no	IND	IND	0	IND	0	IND
	1	12	1	1	0	0	0	0	1	0	0.998	0.974	0	0	0	0	1	no	IND	IND	0	IND	0	IND
	1	14	1	1	0	0	0	0	1	0	0.964	0.928	0	0	0	0	1	no	IND	IND	0	IND	0	IND
	2	4	0.765	0.68	0	0	0	0	0.897	0.84	0.238	0.167	0	0	0	0	0.853	0.81	IND	IND	0	0	0	0
	2	6	0.967	0.931	0	0	0	0	0.933	0.89	0.371	0.276	0	0	0	0	1.036	1.046	IND	IND	0	0	0	0
	2	10	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	2	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	3	4	0.445	0.379	0	0	0	0	0.875	0.82	0.038	0.03	0	0	0	0	0.509	0.462	IND	IND	0	0	0	0
	3	6	0.702	0.64	0	0	0	0	0.971	0.924	0.108	0.084	0	0	0	0	0.723	0.693	IND	IND	0	0	0	0
	3	10	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	3	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	4	9	0.799	0.747	0	0	0	0	0.912	0.856	0.355	0.274	0	0	0	0	0.876	0.873	IND	IND	0	0	0	0
	4	10	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	4	11	1	1	0	0	0	0	1	0.989	1	1	0	0	0	0	1	1.011	IND	IND	0	0	0	0
	4	12	0.511	0.453	0	0	0	0	0.923	0.858	0.036	0.026	0	0	0	0	0.554	0.528	IND	IND	0	0	0	0
	4	14	0.424	0.36	0	0	0	0	0.892	0.819	0.034	0.023	0	0	0	0	0.475	0.44	IND	IND	0	0	0	0
	6	9	0.991	0.959	0	0	0	0	0.924	0.864	0.496	0.398	0	0	0	0	1.073	1.11	IND	IND	0	0	0	0
	6	10	1	0.999	0	0	0	0	1	1	0.987	0.968	0	0	0	0	1	0.999	IND	IND	0	0	0	0
	6	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	6	12	0.66	0.571	0	0	0	0	0.921	0.861	0.09	0.061	0	0	0	0	0.717	0.663	IND	IND	0	0	0	0
	6	14	0.631	0.545	0	0	0	0	0.945	0.892	0.1	0.071	0	0	0	0	0.668	0.611	IND	IND	0	0	0	0
	9	10	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	9	12	0.688	0.503	0	0	0	0	0.894	0.824	0.123	0.065	0	0	0	0	0.77	0.61	IND	IND	0	0	0	0
	10	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	10	12	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	10	14	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	11	12	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	11	14	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0

Las eficiencias Bahadur, reportadas en el cuadro 8 mostraron valores mayores a 1 solamente entre las pruebas de Dunn y Baumgartner y en poblaciones cercanas. En las demás no se evidenció eficiencia alguna.

Cuadro 9. Tasas de error Tipo I para 10 tratamientos contrastantes, diseño C.A. dentro de los grupos homogéneos.

Tasas de error Tipo I													
Trat	k	I	S-D		Cr-Fl		Hy-St		Bau-W		Tukey	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
10 TTOS CONTRASTANT	2	3	0.372	0.343	0	0	0	0	0.848	0.637	0.005	0
	3	9	0.62	0.409	0	0	0	0	0.874	0.787	0.091	0.051
	3	12	0.222	0.157	0	0	0	0	0.847	0.772	0	0
	3	14	0.122	0.08	0	0	0	0	0.822	0.691	0	0
	4	6	0.385	0.348	0	0	0	0	0.87	0.787	0.032	0.021

En el cuadro 9 se puede observar que las tasas de error Tipo I son altas en algunos casos en la prueba conservativa de Baumgartner-Weiss y Schindler en los dos niveles de significación estudiados. La prueba de Dunn mostró tasas de error tipo I bajas en poblaciones distantes y cercanas donde no se evidencia variación en las tasas de error tipo I independientemente de la "distancia" existente entre las poblaciones. La prueba de Dunn es liberal; es decir trabaja con error familiar, y es posible que esto explique este comportamiento. Las pruebas de Steel-Dwass, Critchlow y Fligner y Hayter-Stone no evidenciaron tasas de error tipo I diferentes de cero bajo ninguno de los niveles de significación estudiados, probablemente debido a su carácter conservativo que provee protección extra para el error tipo I. La prueba de Tukey mostró bajas tasas de error tipo I en algunos casos y cero en otros mostrando inconsistencia en sus resultados.

Cuadro 10. Tasas de rechazo y eficiencia Pitman para 10 tratamientos semejantes, diseño C.A. entre grupos homogéneos.

Tasa de rechazo													Eficiencia Pitman							
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey		S-D		Cr-Fl		Hy-St		Bau-W	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
10 TRATAMIENTOS SEMEJANTES	3	5	0.684	0.684	0	0	0	0	0.855	0.784	0.024	0.004	28.5	171	0	0	0	0	35.63	196
	3	6	0.702	0.702	0	0	0	0	0.971	0.924	0.101	0.047	6.95	14.94	0	0	0	0	9.614	19.66
	3	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	3	13	0.499	0.499	0	0	0	0	0.886	0.825	0.032	0.012	15.59	41.58	0	0	0	0	27.69	68.75
	5	7	0.394	0.394	0	0	0	0	0.873	0.804	0.008	0.002	49.25	197	0	0	0	0	109.1	402
	5	8	0.606	0.606	0	0	0	0	0.902	0.834	0.015	0.005	40.4	121.2	0	0	0	0	60.13	166.8
	5	14	0.6	0.6	0	0	0	0	0.854	0.795	0.024	0.007	25	85.71	0	0	0	0	35.58	113.6
	5	15	0.333	0.333	0	0	0	0	0.866	0.784	0.002	0	166.5	no	0	IND	0	IND	433	no
	5	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	5	12	0.603	0.603	0	0	0	0	0.905	0.839	0.024	0.008	25.13	75.38	0	0	0	0	37.71	104.9
	6	7	0.455	0.455	0	0	0	0	0.892	0.813	0.023	0.006	19.78	75.83	0	0	0	0	38.78	135.5
	6	8	0.671	0.671	0	0	0	0	0.885	0.809	0.039	0.013	17.21	51.62	0	0	0	0	22.69	62.23
	6	14	0.631	0.631	0	0	0	0	0.945	0.892	0.09	0.041	7.011	15.39	0	0	0	0	10.5	21.76
	6	15	0.373	0.373	0	0	0	0	0.877	0.809	0.013	0	28.69	no	0	IND	0	IND	67.46	no
	6	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	6	12	0.66	0.66	0	0	0	0	0.921	0.861	0.064	0.023	10.31	28.7	0	0	0	0	14.39	37.43
	7	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	7	13	0.244	0.244	0	0	0	0	0.835	0.766	0.004	0.002	61	122	0	0	0	0	208.8	383
	8	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	8	13	0.445	0.445	0	0	0	0	0.864	0.801	0.011	0.005	40.45	89	0	0	0	0	78.55	160.2
	14	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	14	13	0.455	0.455	0	0	0	0	0.838	0.838	0.011	0.011	41.36	41.36	0	0	0	0	76.18	76.18
	15	11	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	15	13	0.329	0.329	0	0	0	0	0.884	0.824	0.004	0.001	82.25	329	0	0	0	0	221	824
	11	12	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	11	13	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1
	12	13	0.489	0.789	0	0	0	0	0.931	0.878	0.031	0.016	15.77	49.31	0	0	0	0	30.03	54.88

Cuadro 8. Tasas de rechazo y eficiencias Bahadur para 10 tratamientos semejantes, diseño C.A. entre grupos homogéneos.

Tasa de rechazo													Eficiencias Bahadur											
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey		CR-FL / S-D		HY-ST/S-D		S-D / Bau-W		CR-FL / HY-ST		CR-FL / Bau-W		HY-ST / Bau-W	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
10 TRATAMIENTOS SEMEJANTES	3	5	0.684	0.684	0	0	0	0	0.855	0.784	0.024	0.004	0	0	0	0	0.8	0.872	IND	IND	0	0	0	0
	3	6	0.702	0.702	0	0	0	0	0.971	0.924	0.101	0.047	0	0	0	0	0.723	0.76	IND	IND	0	0	0	0
	3	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	3	13	0.499	0.499	0	0	0	0	0.886	0.825	0.032	0.012	0	0	0	0	0.563	0.605	IND	IND	0	0	0	0
	5	7	0.394	0.394	0	0	0	0	0.873	0.804	0.008	0.002	0	0	0	0	0.451	0.49	IND	IND	0	0	0	0
	5	8	0.606	0.606	0	0	0	0	0.902	0.834	0.015	0.005	0	0	0	0	0.672	0.727	IND	IND	0	0	0	0
	5	14	0.6	0.6	0	0	0	0	0.854	0.795	0.024	0.007	0	0	0	0	0.703	0.755	IND	IND	0	0	0	0
	5	15	0.333	0.333	0	0	0	0	0.866	0.784	0.002	0	0	0	0	0	0.385	0.425	IND	IND	0	0	0	0
	5	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	5	12	0.603	0.603	0	0	0	0	0.905	0.839	0.024	0.008	0	0	0	0	0.666	0.719	IND	IND	0	0	0	0
	6	7	0.455	0.455	0	0	0	0	0.892	0.813	0.023	0.006	0	0	0	0	0.51	0.56	IND	IND	0	0	0	0
	6	8	0.671	0.671	0	0	0	0	0.885	0.809	0.039	0.013	0	0	0	0	0.758	0.829	IND	IND	0	0	0	0
	6	14	0.631	0.631	0	0	0	0	0.945	0.892	0.09	0.041	0	0	0	0	0.668	0.707	IND	IND	0	0	0	0
	6	15	0.373	0.373	0	0	0	0	0.877	0.809	0.013	0	0	0	0	0	0.425	0.461	IND	IND	0	0	0	0
	6	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	6	12	0.66	0.66	0	0	0	0	0.921	0.861	0.064	0.023	0	0	0	0	0.717	0.767	IND	IND	0	0	0	0
	7	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	7	13	0.244	0.244	0	0	0	0	0.835	0.766	0.004	0.002	0	0	0	0	0.292	0.319	IND	IND	0	0	0	0
	8	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	8	13	0.445	0.445	0	0	0	0	0.864	0.801	0.011	0.005	0	0	0	0	0.515	0.556	IND	IND	0	0	0	0
	14	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	14	13	0.455	0.455	0	0	0	0	0.838	0.838	0.011	0.011	0	0	0	0	0.543	0.543	IND	IND	0	0	0	0
	15	11	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	15	13	0.329	0.329	0	0	0	0	0.884	0.824	0.004	0.001	0	0	0	0	0.372	0.399	IND	IND	0	0	0	0
	11	12	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	11	13	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	IND	IND	0	0	0	0
	12	13	0.489	0.789	0	0	0	0	0.931	0.878	0.031	0.016	0	0	0	0	0.525	0.899	IND	IND	0	0	0	0

En el cuadro 10 se puede observar que las tasas de rechazo son muy altas e incluso cercanas a 1 en algunos casos en la prueba conservativa de Baumgartner-Weiss y Schindler, en los dos niveles de significación estudiados, seguida por la prueba de Dunn que mostró bajas tasas de rechazo en poblaciones distantes. Las pruebas de Steel-Dwass, Critchlow y Fligner y Hayter-Stone no evidenciaron tasas de rechazo bajo ninguno de los niveles de significación estudiados.

Las eficiencias Pitman, reportadas en el cuadro 10 mostraron resultados muy altos, 1 o superiores a 1 en algunos casos en las pruebas de rankeo conjunto, pero 0 para las de rankeo pareado, exceptuando el caso de la prueba de Baumgartner que obtuvo eficiencias Pitman altas.

Las eficiencias Bahadur, reportadas en el cuadro 11 mostraron valores mayores a 1 solamente entre las pruebas de Dunn y Baumgartner y en poblaciones cercanas. Entre las demás no se evidenció eficiencia alguna.

Cuadro 12. Tasas de error Tipo I para 10 tratamientos semejantes, diseño C.A. dentro de los grupos homogéneos.

Tasas de error Tipo I												
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
10 TRATAMIENTOS SEMEJANTES	3	7	0.356	0.356	0	0	0	0	0.856	0.798	0.006	0.001
	3	8	0.119	0.119	0	0	0	0	0.835	0.728	0	0
	3	14	0.122	0.122	0	0	0	0	0.822	0.691	0	0
	3	15	0.452	0.452	0	0	0	0	0.876	0.828	0.024	0.005
	3	12	0.222	0.222	0	0	0	0	0.847	0.772	0	0
	5	6	0.334	0.334	0	0	0	0	0.937	0.881	0.016	0.005
	5	13	0.262	0.262	0	0	0	0	0.859	0.785	0	0
	6	13	0.367	0.367	0	0	0	0	0.891	0.821	0.009	0.004
	7	8	0.324	0.324	0	0	0	0	0.868	0.806	0	0
	7	14	0.323	0.323	0	0	0	0	0.867	0.791	0.008	0.001
	7	15	0.309	0.309	0	0	0	0	0.892	0.822	0.007	0
	7	12	0.365	0.365	0	0	0	0	0.901	0.839	0.011	0.002
	8	14	0.303	0.303	0	0	0	0	0.863	0.812	0.001	0
	8	15	0.385	0.385	0	0	0	0	0.875	0.812	0.012	0.002
	8	12	0.225	0.225	0	0	0	0	0.867	0.791	0.001	0
	14	15	0.36	0.36	0	0	0	0	0.776	0.776	0.001	0.001
	14	12	0.168	0.168	0	0	0	0	0.751	0.751	0	0
	15	12	0.387	0.387	0	0	0	0	0.882	0.813	0.009	0.002

En el cuadro 12 se puede observar que las tasas de error tipo I son altas en todos los casos casos en la prueba conservativa de Baumgartner-Weiss y Schindler con los dos niveles de significación estudiados, a pesar de que no fue 1 para ninguno de los pares comparados. La prueba de Dunn mostró tasas de error tipo I bajas en poblaciones distantes y cercanas donde no se evidencia variación en las tasas de error tipo I independientemente de la "distancia" existente entre las poblaciones, pero tuvo mayor control del error tipo I que la prueba de Baumgartner-Weiss, a pesar de que ninguna de las dos logró controlar el error tipo I. La prueba de Dunn es de carácter liberal; es decir, trabaja con error familiar, y es posible que esto explique este comportamiento. Las pruebas de Steel-Dwass, Critchlow y Fligner y Hayter-Stone no evidenciaron tasas de error tipo I diferentes de cero bajo ninguno de los niveles de significación estudiados, probablemente debido a su carácter conservativo que provee protección extra para el error tipo I. La prueba de Tukey mostró bajas tasas de error Tipo I en algunos casos y cero en otros mostrando aparente efectividad en el control del error tipo I en sus resultados.

Cuadro 13. Tasas de rechazo y eficiencia Pitman para 5 tratamientos contrastantes, diseño C.A. entre grupos homogéneos.

			Tasa de rechazo										Eficiencia Pitman							
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey		S-D		Cr-Fl		Hy-St		Bau-W	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
5 CONTRASTA	1	2	0.975	0.887	0.754	0	0.754	0.754	1	1	0.754	0.754	1.293	1.176	1	0	1	1	1.326	1.326
	1	6	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1
	1	10	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1
	1	11	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1

Cuadro 14. Tasas de rechazo y eficiencias Bahadur para 5 tratamientos contrastantes, diseño C.A. entre grupos homogéneos.

			Tasa de rechazo										Eficiencias Bahadur											
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey		CR-FL / S-D		HY-ST/S-D		S-D / Bau-W		CR-FL / HY-ST		CR-FL / Bau-W		HY-ST / Bau-W	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
5 CONTRASTA	1	2	0.975	0.887	0.754	0	0.754	0.754	1	1	0.754	0.754	0.773	0	0.773	0.85	0.975	0.887	1	0	0.754	0	0.754	0.754
	1	6	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1
	1	10	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1
	1	11	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1

En el cuadro 13 se puede observar que las tasas de rechazo son 1 en todos los casos para la prueba conservativa de Baumgartner-Weiss y Schindler, con los dos niveles de significación estudiados, seguida por la prueba de Dunn que mostró altas tasas de rechazo con valores de 1 en poblaciones distantes, y cercanas a 1 en poblaciones cercanas. Las pruebas de Steel-Dwass, Critchlow y Fligner y Hayter-Stone mostraron el mismo comportamiento que la de Dunn, exceptuando el caso de la prueba de Steel-Dwass-Critchlow y Fligner que no presentó ningún rechazo en el nivel de significación de 0,05.

Las eficiencias Pitman, reportadas en el cuadro 13 mostraron resultados muy altos, 1 o superiores a 1 en algunos casos en las pruebas de rankeo conjunto y también para las de rankeo pareado, con la excepción de la prueba de Steel-Dwass-Critchlow y Fligner, en la que la eficiencia Pitman fue 0 para el nivel de significación de 0,05.

Las eficiencias Bahadur, reportadas en el cuadro 14 mostraron valores mayores a 1 solamente entre las pruebas de rankeo pareado y poblaciones cercanas, mostrando que la prueba de Steel-Dwass-Critchlow y Fligner, parece ser la que tiene una eficiencia Bahadur mayor.

Cuadro 15. Tasas de error Tipo I para 5 tratamientos contrastantes, diseño C.A. dentro de los grupos homogéneos.

Tasas de error Tipo I												
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey	
5 CONT		NO HAY										

Cuadro 16. Tasas de rechazo y eficiencia Pitman para 5 tratamientos semejantes, diseño C.A. entre grupos homogéneos.

			Tasa de rechazo										Eficiencia Pitman							
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey		S-D		Cr-Fl		Hy-St		Bau-W	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
5 SEMEJANTES	5	7	0.394	0.394	0.001	0	0.001	0.001	0.867	0.953	0.023	0.022	17.13	17.91	0.043	0	0.043	0.045	37.7	43.32
	5	13	0.262	0.262	0	0	0	0	1	1	0.008	0.007	32.75	37.43	0	0	0	0	125	142.9
	5	15	0.333	0.333	0	0	0	0	1	1	0.01	0.007	33.3	47.57	0	0	0	0	100	142.9
	5	11	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1
	7	11	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1
	13	11	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1
	15	11	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1

Cuadro 17. Tasas de rechazo y eficiencias Bahadur para 5 tratamientos semejantes, diseño C.A. entre grupos homogéneos.

			Tasa de rechazo										Eficiencias Bahadur											
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey		CR-FL / S-D		HY-ST/S-D		S-D / Bau-W		CR-FL / HY-ST		CR-FL / Bau-W		HY-ST / Bau-W	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
5 SEMEJANTES	5	7	0.394	0.394	0.001	0	0.001	0.001	0.867	0.953	0.023	0.022	0.003	0	0.003	0.003	0.454	0.413	1	0	0.001	0	0.001	0.001
	5	13	0.262	0.262	0	0	0	0	1	1	0.008	0.007	0	0	0	0	0.262	0.262	IND	IND	0	0	0	0
	5	15	0.333	0.333	0	0	0	0	1	1	0.01	0.007	0	0	0	0	0.333	0.333	IND	IND	0	0	0	0
	5	11	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1
	7	11	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1
	13	11	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1
	15	11	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1

En el cuadro 16 se puede observar que las tasas de rechazo son 1 o cercanas a 1 para la prueba conservativa de Baumgartner-Weiss y Schindler, en los dos niveles de significación estudiados. Para poblaciones distantes la tasa de rechazo fue 1 pero menor a 1 en poblaciones cercanas. La prueba de Dunn mostró altas tasas de rechazo con valores de 1 en poblaciones distantes, y cercanas a 1 en poblaciones más cercanas. La prueba de Steel-

Dwass, Critchlow y Fligner no presentó ningún rechazo en el nivel de significación de 0,05 y alcanzó el valor de 1 para poblaciones distantes a un nivel de 0,1. La prueba de Hayter-Stone sólo muestra tasas de rechazo diferentes de cero para las poblaciones más alejadas y tasa de rechazo muy baja menores a 0.3 para poblaciones que tienen dos medianas de separación.

Las eficiencias Pitman, reportadas en el cuadro 16 mostraron resultados muy altos, 1 o superiores a 1 en algunos casos en las pruebas de rankeo pareado y también para las de rankeo conjunto, con la excepción de la prueba de Steel-Dwass-Critchlow y Fligner, en la que la eficiencia Pitman fue 0 para el nivel de significación de 0,05. La prueba de Hayter-stone muestra eficiencia pitman 0 para poblaciones alejadas.

Las eficiencias Bahadur, reportadas en el cuadro 17 mostraron valores mayores a 1 solamente entre las pruebas de rankeo pareado y poblaciones cercanas al nivel de significación 0,1. Las pruebas de Dunn y Baumgartner-Weiss tienen eficiencias Bahadur similares.

Cuadro 18. Tasas de error Tipo I para 5 tratamientos semejantes, diseño C.A. dentro de los grupos homogéneos.

Tasas de error Tipo I												
Trat	k	l	S-D		Cr-Fl		Hy-St		Bau-W		Tukey	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
5 SEMEJ	7	13	0.244	0.244	0	0	0	0	0.832	0.919	0.009	0.007
	7	15	0.309	0.309	0	0	0	0	1	1	0.019	0.017
	13	15	0.329	0.329	0.001	0	0.001	0.001	0.878	0.95	0.022	0.017

En el cuadro 18 se puede observar que las tasas de error Tipo I son 1 o cercanas a 1 en todos los casos para la prueba conservativa de Baumgartner-Weiss y Schindler con los dos niveles de significación estudiados. La prueba de Dunn mostró tasas de error tipo I bajas en poblaciones distantes y cercanas donde no se evidencia variación en las tasas de error tipo I independientemente de la "distancia" existente entre las poblaciones, pero tuvo mayor control del error tipo I que la prueba de Baumgartner-Weiss, a pesar de que ninguna de las dos logró controlar el error tipo I. Las pruebas de Steel-Dwass, Critchlow y Fligner y Hayter-Stone no evidenciaron tasas de error tipo I diferentes de cero bajo ninguno de los niveles de significación estudiados, probablemente debido a su carácter conservativo que provee protección extra para el error tipo I en poblaciones cercanas, y lo mantuvo por debajo del nivel planteado en las poblaciones más alejadas. La prueba de Tukey mostró bajas tasas de error tipo I en algunos casos y cero en otros mostrando aparente efectividad en el control del error tipo I en sus resultados.

Cuadro 19. Tasas de error y eficiencia Pitman para 15 tratamientos contra un control. (Control = Tratamiento 11). Todos los pares. Diseño C.A.

Tasas de Error.C.A. c/ Control (trat 11)							**Efic Pitman**	
Trat	k	I	Nem-Da-W		Dunnett		Nem-Da-W	
			.1	.05	.1	.05	.1	.05
15 Tratamientos. Todos los pares	11	1	0	0	1	0.958	0	0
	11	2	0	0	1	1	0	0
	11	3	0.486	0.226	1	1	0.486	0.226
	11	4	0.87	0.741	1	1	0.87	0.741
	11	5	0.985	0.947	1	1	0.985	0.947
	11	6	0.987	0.962	1	1	0.987	0.962
	11	7	0.853	0.68	1	1	0.853	0.68
	11	8	0.611	0.342	1	1	0.611	0.342
	11	9	0.036	0.009	1	1	0.036	0.009
	11	10	1	1	1	1	1	1
	11	11	0	0	0	0	IND	IND
	11	12	0.579	0.352	1	1	0.579	0.352
	11	13	0.935	0.826	1	1	0.935	0.826
	11	14	0.594	0.362	1	1	0.594	0.362
	11	15	0.887	0.766	1	1	0.887	0.766

En el cuadro 19 se puede observar que la tasa de rechazo solo alcanza el valor de 1 para las poblaciones más distantes con la prueba no paramétrica de Nemenyi-Damico-Wolf, y cercanas a 1 en el nivel de significación de 0,1. Para poblaciones cercanas, la prueba de Nemenyi-Damico y Wolf no evidencia tasa de rechazo alguna.

La eficiencia Pitman, reportada en el cuadro 19 muestra resultados relativamente altos, cercanos a 1 o 1 en algunos casos para poblaciones distantes.

Cuadro 20. Tasas de error y eficiencia Pitman para 10 tratamientos contrastantes contra un control. (Control = Tratamiento 11). Todos los pares. Diseño C.A.

Tasas de Error.C.A. c/ Control (trat 11)							Efic Pitman	
Trat	k	l	Nem-Da-W		Dunnett		Nem-Da-W	
			.1	.05	.1	.05	.1	.05
10 Trat Contratsantes	11	1	0	0	1	0.958	0	0
		2	0.282	0.023	1	1	0.282	0.023
		3	0.891	0.6	1	1	0.891	0.6
		4	0.94	0.855	1	1	0.94	0.855
		6	0.999	0.995	1	1	0.999	0.995
		9	0.15	0.048	1	1	0.15	0.048
		10	1	1	1	1	1	1
		11	0	0	0	0	IND	IND
		12	0.923	0.703	1	1	0.923	0.703
		14	0.903	0.674	1	1	0.903	0.674

En el cuadro 20 se puede observar que la tasa de rechazo solo alcanza el valor de 1 para las poblaciones más distantes y con una diferencia real con la prueba no paramétrica de Nemenyi-Damico-Wolf al igual que en las otras dos pruebas, y cercanas a 1 en ambos niveles de significación sólo para poblaciones distantes. Para poblaciones cercanas, la prueba de Nemenyi-Damico y Wolf no evidencia tasa de rechazo alguna.

La eficiencia Pitman, reportada en el cuadro 20 muestra resultados relativamente altos, cercanos a 1 o 1 en algunos casos para poblaciones distantes y 0 para poblaciones cercanas frente a su competidor seleccionado.

Cuadro 21. Tasas de error y eficiencia Pitman para 10 tratamientos semejantes contra un control. (Control = Tratamiento 11). Todos los pares. Diseño C.A.

Tasas de Error.C.A. c/ Control (trat 11)							Efic Pitman	
Trat	k	I	Nem-Da-W		Dunnett		Nem-Da-W	
			.1	.05	.1	.05	.1	.05
10 Trat Semejantes	11	3	0.399	0.233	1	1	0.399	0.233
		5	0.985	0.953	1	1	0.985	0.953
		6	0.976	0.947	1	1	0.976	0.947
		7	0.807	0.676	1	1	0.807	0.676
		8	0.527	0.339	1	1	0.527	0.339
		14	0.529	0.352	1	1	0.529	0.352
		15	0.875	0.772	1	1	0.875	0.772
		11	0	0	0	0	IND	IND
		12	0.517	0.345	1	1	0.517	0.345
		13	0.915	0.826	1	1	0.915	0.826

En el cuadro 21 se puede observar que la tasa de rechazo solo alcanza valores cercanos a 1 para el nivel de significación de 0,1 y para las poblaciones más distantes con la prueba no paramétrica de Nemenyi-Damico-Wolf,. Para poblaciones cercanas, la prueba de Nemenyi-Damico y Wolf evidencia tasas de rechazo bajas.

La eficiencia Pitman, reportada en el cuadro 21 muestra resultados relativamente altos, cercanos a 1 o 1 en algunos casos para poblaciones distantes.

Cuadro 22. Tasas de error y eficiencia Pitman para 5 tratamientos contrastantes contra un control. (Control = Tratamiento 11). Todos los pares. Diseño C.A.

Tasas de Error.C.A. c/ Control (trat 11)							Efic Pitman	
Trat	k	I	Nem-Da-W		Dunnett		Nem-Da-W	
			.1	.05	.1	.05	.1	.05
5 Trat Contr	11	1	1	1	1	1	1	1
		2	1	1	1	1	1	1
		6	1	1	1	1	1	1
		10	1	1	1	1	1	1
		11	0	0	0	0	no	no

En el cuadro 22 se puede observar que la tasa de rechazo alcanza el valor de 1 para las poblaciones más distantes con la prueba no paramétrica de Nemenyi-Damico-Wolf, con ambos niveles de significación.

La eficiencia Pitman, reportada en el cuadro 22 muestra resultados altos, evidenciando ser tan eficiente como la prueba de Dunnett.

Cuadro 23. Tasas de error y eficiencia Pitman para 5 tratamientos semejantes contra un control. (Control = Tratamiento 11). Todos los pares. Diseño C.A.

Tasas de Error.C.A. c/ Control (trat 11)							**Efic Pitman**	
Trat	k	l	Nem-Da-W		Dunnett		Nem-Da-W	
			.1	.05	.1	.05	.1	.05
5 Trat Semejantes	11	5	1	1	1	1	1	1
		7	1	1	1	1	1	1
		13	1	1	1	1	1	1
		15	1	1	1	1	1	1
		11	0	0	0	0	no	no

En el cuadro 23 se puede observar que la tasa de rechazo alcanza el valor de 1 para las poblaciones más distantes con la prueba no paramétrica de Nemenyi-Damico-Wolf, con ambos niveles de significación.

La eficiencia Pitman, reportada en el cuadro 23 muestra resultados altos, evidenciando ser tan eficiente como la prueba de Dunnett.

DOS VÍAS DE CLASIFICACIÓN

Cuadro 24. Tasas de rechazo y eficiencia Pitman para 15 tratamientos todos los pares diseño B.A. entre grupos homogéneos.

			Tasa de rechazo						Eficiencia Pitman			
Trat	k	l	Wil-nem-Mc		Nemeny		Tukey		Wil-nem-Mc		Nemeny	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
15 TRATAMIENTOS. TODOS LOS PARES	1	2	0	0	0	0	0.588	0.422	0	0	0	0
	1	3	0	0	0	0	0.931	0.837	0	0	0	0
	1	4	0.071	0.014	0	0	0.851	0.812	0.083	0.017	0	0
	1	5	0.108	0.041	0	0	0.999	0.992	0.108	0.041	0	0
	1	6	0.24	0.082	0	0	1	1	0.24	0.082	0	0
	1	7	0.024	0.004	0	0	0.98	0.941	0.024	0.004	0	0
	1	8	0	0	0	0	0.967	0.917	0	0	0	0
	1	10	1	1	0	0	1	1	1	1	0	0
	1	11	0	0	0	0	0.681	0.437	0	0	0	0
	1	12	0.003	0	0	0	0.973	0.909	0.003	0	0	0
	1	13	0.066	0.015	0	0	0.994	0.981	0.066	0.015	0	0
	1	14	0.001	0	0	0	0.94	0.862	0.001	0	0	0
	1	15	0.043	0.01	0	0	0.993	0.98	0.043	0.01	0	0
	2	5	0	0	0	0	0.122	0.067	0	0	0	0
	2	6	0	0	0	0	0.208	0.131	0	0	0	0
	2	9	0	0	0	0	0.004	0	0	IND	0	IND
	2	10	0.268	0.069	0	0	1	1	0.268	0.069	0	0
	2	11	0	0	0	0	1	1	0	0	0	0
	3	5	0	0	0	0	0.01	0.001	0	0	0	0
	3	6	0	0	0	0	0.049	0.023	0	0	0	0
	3	9	0	0	0	0	0.03	0.018	0	0	0	0
	3	10	0.002	0	0	0	1	1	0.002	0	0	0
	3	11	0.007	0	0	0	1	1	0.007	0	0	0
	4	5	0	0	0	0	0.007	0.002	0	0	0	0
	4	6	0	0	0	0	0.022	0.011	0	0	0	0
	4	8	0	0	0	0	0.005	0.002	0	0	0	0
	4	10	0.004	0	0	0	0.999	0.996	0.004	0	0	0
	4	11	0.324	0.147	0	0	1	1	0.324	0.147	0	0
	4	13	0	0	0	0	0.004	0.001	0	0	0	0
	5	7	0	0	0	0	0.005	0.001	0	0	0	0
	5	8	0	0	0	0	0.01	0.005	0	0	0	0
	5	9	0.002	0	0	0	0.335	0.279	0.006	0	0	0
	5	10	0	0	0	0	0.993	0.977	0	0	0	0
	5	11	0.561	0.307	0	0	1	1	0.561	0.307	0	0
	5	12	0	0	0	0	0.011	0.002	0	0	0	0
	5	13	0	0	0	0	0	0	IND	IND	IND	IND
	5	14	0	0	0	0	0.01	0.002	0	0	0	0
	5	15	0	0	0	0	0	0	IND	IND	IND	IND
	6	7	0	0	0	0	0.006	0	0	IND	0	IND
	6	8	0	0	0	0	0.017	0.008	0	0	0	0
	6	9	0	0	0	0	0.343	0.253	0	0	0	0
	6	10	0	0	0	0	0.977	0.937	0	0	0	0
	6	11	0.646	0.428	0	0	1	1	0.646	0.428	0	0
	6	12	0	0	0	0	0.031	0.016	0	0	0	0
	6	13	0	0	0	0	0.002	0	0	IND	0	IND
	6	14	0	0	0	0	0.043	0.021	0	0	0	0
	6	15	0	0	0	0	0.005	0	0	IND	0	IND
	7	9	0	0	0	0	0.109	0.076	0	0	0	0
	7	10	0	0	0	0	1	1	0	0	0	0
	7	11	0.19	0.064	0	0	1	1	0.19	0.064	0	0
	8	9	0	0	0	0	0.028	0.011	0	0	0	0
	8	10	0	0	0	0	1	1	0	0	0	0
	8	11	0.013	0.001	0	0	1	1	0.013	0.001	0	0
	9	10	0.577	0.349	0	0	1	1	0.577	0.349	0	0
	9	11	0	0	0	0	1	1	0	0	0	0
	9	12	0	0	0	0	0.053	0.028	0	0	0	0
	9	13	0	0	0	0	0.245	0.205	0	0	0	0
	9	14	0	0	0	0	0.076	0.054	0	0	0	0

Cuadro 25. Tasas de rechazo y eficiencias Bahadur para 15 tratamientos todos los pares diseño B.A. entre grupos homogéneos.

			Tasa de rechazo						Eficiencia Bahadur			
Trat	k	l	Wil-nem-Mc		Nemeny		Tukey		Wil/Nem		Nem/Wil	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
15 TRATAMIENTOS. TODOS LOS PARES	1	2	0	0	0	0	0.588	0.422	IND	IND	IND	IND
	1	3	0	0	0	0	0.931	0.837	IND	IND	IND	IND
	1	4	0.071	0.014	0	0	0.851	0.812	no	no	0	0
	1	5	0.108	0.041	0	0	0.999	0.992	no	no	0	0
	1	6	0.24	0.082	0	0	1	1	no	no	0	0
	1	7	0.024	0.004	0	0	0.98	0.941	no	no	0	0
	1	8	0	0	0	0	0.967	0.917	IND	IND	IND	IND
	1	10	1	1	0	0	1	1	no	no	0	0
	1	11	0	0	0	0	0.681	0.437	IND	IND	IND	IND
	1	12	0.003	0	0	0	0.973	0.909	no	no	0	IND
	1	13	0.066	0.015	0	0	0.994	0.981	no	no	0	0
	1	14	0.001	0	0	0	0.94	0.862	no	no	0	IND
	1	15	0.043	0.01	0	0	0.993	0.98	no	no	0	0
	2	5	0	0	0	0	0.122	0.067	IND	IND	IND	IND
	2	6	0	0	0	0	0.208	0.131	IND	IND	IND	IND
	2	9	0	0	0	0	0.004	0	IND	IND	IND	IND
	2	10	0.268	0.069	0	0	1	1	no	no	0	0
	2	11	0	0	0	0	1	1	IND	IND	IND	IND
	3	5	0	0	0	0	0.01	0.001	IND	IND	IND	IND
	3	6	0	0	0	0	0.049	0.023	IND	IND	IND	IND
	3	9	0	0	0	0	0.03	0.018	IND	IND	IND	IND
	3	10	0.002	0	0	0	1	1	no	no	0	IND
	3	11	0.007	0	0	0	1	1	no	no	0	IND
	4	5	0	0	0	0	0.007	0.002	IND	IND	IND	IND
	4	6	0	0	0	0	0.022	0.011	IND	IND	IND	IND
	4	8	0	0	0	0	0.005	0.002	IND	IND	IND	IND
	4	10	0.004	0	0	0	0.999	0.996	no	no	0	IND
	4	11	0.324	0.147	0	0	1	1	no	no	0	0
	4	13	0	0	0	0	0.004	0.001	IND	IND	IND	IND
	5	7	0	0	0	0	0.005	0.001	IND	IND	IND	IND
	5	8	0	0	0	0	0.01	0.005	IND	IND	IND	IND
	5	9	0.002	0	0	0	0.335	0.279	no	no	0	IND
	5	10	0	0	0	0	0.993	0.977	IND	IND	IND	IND
	5	11	0.561	0.307	0	0	1	1	no	no	0	0
	5	12	0	0	0	0	0.011	0.002	IND	IND	IND	IND
	5	13	0	0	0	0	0	0	IND	IND	IND	IND
	5	14	0	0	0	0	0.01	0.002	IND	IND	IND	IND
	5	15	0	0	0	0	0	0	IND	IND	IND	IND
	6	7	0	0	0	0	0.006	0	IND	IND	IND	IND
	6	8	0	0	0	0	0.017	0.008	IND	IND	IND	IND
	6	9	0	0	0	0	0.343	0.253	IND	IND	IND	IND
	6	10	0	0	0	0	0.977	0.937	IND	IND	IND	IND
	6	11	0.646	0.428	0	0	1	1	no	no	0	0
	6	12	0	0	0	0	0.031	0.016	IND	IND	IND	IND
	6	13	0	0	0	0	0.002	0	IND	IND	IND	IND
	6	14	0	0	0	0	0.043	0.021	IND	IND	IND	IND
	6	15	0	0	0	0	0.005	0	IND	IND	IND	IND
	7	9	0	0	0	0	0.109	0.076	IND	IND	IND	IND
	7	10	0	0	0	0	1	1	IND	IND	IND	IND
	7	11	0.19	0.064	0	0	1	1	no	no	0	0
	8	9	0	0	0	0	0.028	0.011	IND	IND	IND	IND
	8	10	0	0	0	0	1	1	IND	IND	IND	IND
	8	11	0.013	0.001	0	0	1	1	no	no	0	0
	9	10	0.577	0.349	0	0	1	1	no	no	0	0
	9	11	0	0	0	0	1	1	IND	IND	IND	IND
	9	12	0	0	0	0	0.053	0.028	IND	IND	IND	IND
	9	13	0	0	0	0	0.245	0.205	IND	IND	IND	IND
	9	14	0	0	0	0	0.076	0.054	IND	IND	IND	IND

En el cuadro 24 se puede observar que las tasas de rechazo son muy bajas e incluso cero en algunos casos en la prueba de Wilcoxon-Nemenyi y McDonalds, en los dos niveles de significación estudiados, esto se debe al carácter conservativo de la prueba la cual utiliza el error por comparación. La prueba de Nemenyi no evidencia tasas de rechazo bajo ninguno de los niveles de significación estudiados, quizá debido a su carácter liberal al usar el error por experimento o familiar.

En relación a las eficiencias Pitman, resultaron bajas o cero para la prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson exceptuando el caso de la comparación con el grupo 11 el cual se encuentra muy distante del 6 con relación a sus medianas. En ésta comparación la eficiencia Pitman de la prueba fue de 64,6 por ciento. No se evidencia rechazo alguno con la prueba de Nemenyi.

Las eficiencias Bahadur, reportadas en el cuadro 25 indican que estas pruebas no son eficientes en los casos cuyo valor es cero y no se puede concluir al respecto en el caso en que ambas tasas de error dieron 0.

Cuadro 26. Tasas de error Tipo I para 15 tratamientos todos los pares Diseño B.A. Dentro de los grupos homogéneos.

Tasas de Error Tipo I								
Trat	k	l	Wil-nem-Mc		Nemeny		Tukey	
			.1	.05	.1	.05	.1	.05
15 TRATAMIENTOS. TODOS LOS PARES.	1	9	0	0	0	0	0.264	0.178
	2	3	0	0	0	0	0	0
	2	4	0	0	0	0	0.112	0.07
	2	7	0	0	0	0	0.037	0.012
	2	8	0	0	0	0	0.015	0.002
	2	12	0	0	0	0	0	0
	2	13	0	0	0	0	0.117	0.075
	2	14	0	0	0	0	0	0
	2	15	0	0	0	0	0.1	0.059
	3	4	0	0	0	0	0.011	0.003
	3	7	0	0	0	0	0.003	0.001
	3	8	0	0	0	0	0	0
	3	12	0	0	0	0	0	0
	3	13	0	0	0	0	0.013	0.005
	3	14	0	0	0	0	0	0
	3	15	0	0	0	0	0.009	0.002
	4	7	0	0	0	0	0.004	0
	4	8	0	0	0	0	0.005	0.002
	4	12	0	0	0	0	0.009	0.004
	4	13	0	0	0	0	0.004	0.001
	4	14	0	0	0	0	0.007	0
	4	15	0	0	0	0	0.01	0.004
	5	6	0	0	0	0	0.005	0.003
	7	8	0	0	0	0	0.002	0
	7	12	0	0	0	0	0.007	0.002
	7	13	0	0	0	0	0.001	0
	7	14	0	0	0	0	0.004	0.002
	7	15	0	0	0	0	0.002	0
	8	12	0	0	0	0	0	0
	8	13	0	0	0	0	0.01	0.004
	8	14	0	0	0	0	0	0
	8	15	0	0	0	0	0.008	0.002
	12	13	0	0	0	0	0.017	0.005
	12	14	0	0	0	0	0	0
	12	15	0	0	0	0	0.007	0
	13	14	0	0	0	0	0.013	0.002
	13	15	0	0	0	0	0.002	0.001
	14	15	0	0	0	0	0.002	0

En el cuadro 26 se puede observar que las tasas de error Tipo I son cero para ambos niveles de significación. La prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson es una prueba conservativa y por lo tanto la protección del error tipo I se hace evidente pero la prueba de Nemenyi es liberal, es decir trabaja con error familiar, a pesar de la diferencia, ninguna de las dos evidencia tasas de error Tipo I bajo ninguno de los niveles de significación estudiados.

Cuadro 27. Tasas de rechazo y eficiencia Pitman para 10 tratamientos contrastantes. Diseño B.A. Entre grupos homogéneos.

			Tasa de rechazo						Eficiencia Pitman			
Trat	k	l	Wil-nem-Mc		Nemeny		Tukey		Wil-nem-Mc		Nemeny	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
10 TRATAMIENTOS CONTRASTANTES	1	2	0	0	0.754	0	0.635	0.563	0	0	1.187	0
	1	3	0	0	1	0	0.957	0.882	0	0	1.045	0
	1	4	0.012	0	0.833	0	0.863	0.81	0.014	0	0.965	0
	1	6	0.052	0.003	1	0	1	1	0.052	0.003	1	0
	1	9	0	0	0.872	0	0.418	0.313	0	0	2.086	0
	1	10	1	1	1	0	1	1	1	1	1	0
	1	12	0	0	1	0	0.997	0.959	0	0	1.003	0
	1	14	0	0	1	0	0.961	0.907	0	0	1.041	0
	2	4	0	0	0.438	0	0.235	0.149	0	0	1.864	0
	2	6	0	0	0.777	0	0.369	0.245	0	0	2.106	0
	2	10	0.001	0	1	0	1	1	0.001	0	1	0
	2	11	0	0	1	0	1	1	0	0	1	0
	3	4	0	0	0.172	0	0.045	0.026	0	0	3.822	0
	3	6	0	0	0.347	0	0.115	0.076	0	0	3.017	0
	3	10	0	0	1	0	1	1	0	0	1	0
	3	11	0.015	0	1	0	1	1	0.015	0	1	0
	4	9	0	0	0.505	0	0.35	0.268	0	0	1.443	0
	4	10	0	0	1	0	1	1	0	0	1	0
	4	11	0.346	0.16	1	0	1	1	0.346	0.16	1	0
	4	12	0	0	0.185	0	0.035	0.023	0	0	5.286	0
	4	14	0	0	0.173	0	0.036	0.022	0	0	4.806	0
	6	9	0	0	0.865	0	0.489	0.387	0	0	1.769	0
	6	10	0	0	1	0	0.984	0.948	0	0	1.016	0
	6	11	0.675	0.432	1	0	1	1	0.675	0.432	1	0
	6	12	0	0	0.323	0	0.093	0.057	0	0	3.473	0
	6	14	0	0	0.312	0	0.103	0.068	0	0	3.029	0
	9	10	0.157	0.056	1	0	1	1	0.157	0.056	1	0
	9	12	0	0	0.34	0	0.117	0.068	0	0	2.906	0
	10	11	1	1	1	0	1	1	1	1	1	0
	10	12	0	0	1	0	1	1	0	0	1	0
	10	14	0	0	1	0	1	1	0	0	1	0
	11	12	0.029	0.003	1	0	1	1	0.029	0.003	1	0
	11	14	0.042	0.005	1	0	1	1	0.042	0.005	1	0

Cuadro 28. Tasas de rechazo y eficiencias Bahadur para 10 tratamientos contrastantes. Diseño B.A. Entre grupos homogéneos.

			Tasa de rechazo						Eficiencia Bahadur			
Trat	k	l	Wil-nem-Mc		Nemeny		Tukey		Wil/nem		Nem/Wil	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
10 TRATAMIENTOS CONTRASTANTES	1	2	0	0	0.754	0	0.635	0.563	0	IND	no	IND
	1	3	0	0	1	0	0.957	0.882	0	IND	no	IND
	1	4	0.012	0	0.833	0	0.863	0.81	0.014	IND	69.42	IND
	1	6	0.052	0.003	1	0	1	1	0.052	no	19.23	0
	1	9	0	0	0.872	0	0.418	0.313	0	IND	no	IND
	1	10	1	1	1	0	1	1	1	no	1	0
	1	12	0	0	1	0	0.997	0.959	0	IND	no	IND
	1	14	0	0	1	0	0.961	0.907	0	IND	no	IND
	2	4	0	0	0.438	0	0.235	0.149	0	IND	no	IND
	2	6	0	0	0.777	0	0.369	0.245	0	IND	no	IND
	2	10	0.001	0	1	0	1	1	0.001	IND	1000	IND
	2	11	0	0	1	0	1	1	0	IND	no	IND
	3	4	0	0	0.172	0	0.045	0.026	0	IND	no	IND
	3	6	0	0	0.347	0	0.115	0.076	0	IND	no	IND
	3	10	0	0	1	0	1	1	0	IND	no	IND
	3	11	0.015	0	1	0	1	1	0.015	IND	66.67	IND
	4	9	0	0	0.505	0	0.35	0.268	0	IND	no	IND
	4	10	0	0	1	0	1	1	0	IND	no	IND
	4	11	0.346	0.16	1	0	1	1	0.346	no	2.89	0
	4	12	0	0	0.185	0	0.035	0.023	0	IND	no	IND
	4	14	0	0	0.173	0	0.036	0.022	0	IND	no	IND
	6	9	0	0	0.865	0	0.489	0.387	0	IND	no	IND
	6	10	0	0	1	0	0.984	0.948	0	IND	no	IND
	6	11	0.675	0.432	1	0	1	1	0.675	no	1.481	0
	6	12	0	0	0.323	0	0.093	0.057	0	IND	no	IND
	6	14	0	0	0.312	0	0.103	0.068	0	IND	no	IND
	9	10	0.157	0.056	1	0	1	1	0.157	no	6.369	0
	9	12	0	0	0.34	0	0.117	0.068	0	IND	no	IND
	10	11	1	1	1	0	1	1	1	no	1	0
	10	12	0	0	1	0	1	1	0	IND	no	IND
	10	14	0	0	1	0	1	1	0	IND	no	IND
	11	12	0.029	0.003	1	0	1	1	0.029	no	34.48	0
	11	14	0.042	0.005	1	0	1	1	0.042	no	23.81	0

En el cuadro 27 se puede observar que las tasas de rechazo son muy bajas e incluso cero en algunos casos en la prueba de Wilcoxon, Nemenyi, McDonalds y Thompson en los dos niveles de significación estudiados, esto se debe al carácter conservativo de la prueba la cual utiliza el error por comparación. La prueba de Nemenyi no evidencia tasas de rechazo bajo el nivel de significación 0,05 y presenta tasas de rechazo muy altas para el nivel de

0,1 coincidiendo en algunos casos a este nivel con la prueba de Wilcoxon, Nemenyi, McDonalds y Thompson.

En relación a las eficiencias Pitman, resultaron bajas o cero para la prueba de Wilcoxon, Nemenyi, McDonalds y Thompson. No se evidencia rechazo alguno con la prueba de Nemenyi para el nivel de significación de 0,05 pero resulta una eficiencia Pitman muy alta para el nivel de significación de 0,1.

Las eficiencias Bahadur, reportadas en el cuadro 28 indican que estas pruebas no son eficientes en los casos cuyo valor es cero y no se puede concluir al respecto en el caso en que ambas tasas de error dieron 0, pero resulta una alta eficiencia Bahadur en casos de poblaciones distantes.

Cuadro 29. Tasas de error Tipo I para 10 tratamientos contrastantes. Diseño B.A. Dentro de los grupos homogéneos.

Tasas de Error Tipo I								
Trat	k	l	Wil-nem-Mc		Nemeny		Tukey	
			.1	.05	.1	.05	.1	.05
10 TRATAM. CONTRASTANTES	2	3	0	0	0.08	0	0.005	0
	3	9	0	0	0.265	0	0.089	0.053
	3	12	0	0	0.024	0	0	0
	3	14	0	0	0.01	0	0	0
	4	6	0	0	0.08	0	0.032	0.018

En el cuadro 29 se puede observar que las tasas de error Tipo I son cero para ambos niveles de significación en la prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson. Esta es una prueba conservativa y por lo tanto la protección del error tipo I se hace evidente pero la prueba de Nemenyi es liberal, es decir trabaja con error familiar y sólo evidencia bajas tasas de error tipo I al nivel de 0,1.

Cuadro 30. Tasas de rechazo y eficiencia Pitman para 10 tratamientos semejantes. Diseño B.A. Entre grupos homogéneos.

			Tasa de rechazo						Eficiencia Pitman			
Trat	k	l	Wil-nem-Mc		Nemeny		Tukey		Wil-nem-Mc		Nemeny	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
10 TRATAMIENTOS SEMEJANTES	3	5	0	0	0.298	0	0.02	0.002	0	0	14.9	0
	3	6	0	0	0.347	0	0.09	0.041	0	0	3.856	0
	3	11	0	0	1	0	1	1	0	0	1	0
	3	13	0	0	0.214	0	0.027	0.008	0	0	7.926	0
	5	7	0	0	0.124	0	0.005	0.001	0	0	24.8	0
	5	8	0	0	0.278	0	0.013	0.005	0	0	21.38	0
	5	14	0	0	0.252	0	0.018	0.004	0	0	14	0
	5	15	0	0	0.088	0	0.002	0	0	IND	44	IND
	5	11	0.437	0.258	1	0	1	1	0.437	0.258	1	0
	5	12	0	0	0.294	0	0.019	0.005	0	0	15.47	0
	6	7	0	0	0.163	0	0.02	0.003	0	0	8.15	0
	6	8	0	0	0.335	0	0.035	0.009	0	0	9.571	0
	6	14	0	0	0.312	0	0.081	0.038	0	0	3.852	0
	6	15	0	0	0.121	0	0.012	0	0	IND	10.08	IND
	6	11	0.532	0.39	1	0	1	1	0.532	0.39	1	0
	6	12	0	0	0.323	0	0.057	0.016	0	0	5.667	0
	7	11	0.105	0.041	1	0	1	1	0.105	0.041	1	0
	7	13	0	0	0.059	0	0.003	0	0	IND	19.67	IND
	8	11	0.007	0.001	1	0	1	1	0.007	0.001	1	0
	8	13	0	0	0.159	0	0.011	0.005	0	0	14.45	0
	14	11	0.011	0.002	1	0	1	1	0.011	0.002	1	0
	14	13	0	0	0.199	0	0.025	0.007	0	0	7.96	0
	15	11	0.17	0.083	1	0	1	1	0.17	0.083	1	0
	15	13	0	0	0.075	0	0.002	0.001	0	0	37.5	0
	11	12	0.024	0.007	1	0	1	1	0.024	0.007	1	0
	11	13	0.247	0.135	1	0	1	1	0.247	0.135	1	0
	12	13	0	0	0.206	0	0.027	0.011	0	0	7.63	0

Cuadro 31. Tasas de rechazo y eficiencias Bahadur para 10 tratamientos semejantes. Diseño B.A. Entre grupos homogéneos.

Tasa de rechazo									Eficiencia Bahadur			
Trat	k	l	Wil-nem-Mc		Nemeny		Tukey		Wil/nem		Nem/Wil	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
10 TRATAMIENTOS SEMEJANTES	3	5	0	0	0.298	0	0.02	0.002	0	IND	no	IND
	3	6	0	0	0.347	0	0.09	0.041	0	IND	no	IND
	3	11	0	0	1	0	1	1	0	IND	no	IND
	3	13	0	0	0.214	0	0.027	0.008	0	IND	no	IND
	5	7	0	0	0.124	0	0.005	0.001	0	IND	no	IND
	5	8	0	0	0.278	0	0.013	0.005	0	IND	no	IND
	5	14	0	0	0.252	0	0.018	0.004	0	IND	no	IND
	5	15	0	0	0.088	0	0.002	0	0	IND	no	IND
	5	11	0.437	0.258	1	0	1	1	0.437	n	2.288	0
	5	12	0	0	0.294	0	0.019	0.005	0	IND	no	IND
	6	7	0	0	0.163	0	0.02	0.003	0	IND	no	IND
	6	8	0	0	0.335	0	0.035	0.009	0	IND	no	IND
	6	14	0	0	0.312	0	0.081	0.038	0	IND	no	IND
	6	15	0	0	0.121	0	0.012	0	0	IND	no	IND
	6	11	0.532	0.39	1	0	1	1	0.532	n	1.88	0
	6	12	0	0	0.323	0	0.057	0.016	0	IND	no	IND
	7	11	0.105	0.041	1	0	1	1	0.105	n	9.524	0
	7	13	0	0	0.059	0	0.003	0	0	IND	no	IND
	8	11	0.007	0.001	1	0	1	1	0.007	n	142.9	0
	8	13	0	0	0.159	0	0.011	0.005	0	IND	no	IND
	14	11	0.011	0.002	1	0	1	1	0.011	n	90.91	0
	14	13	0	0	0.199	0	0.025	0.007	0	IND	no	IND
	15	11	0.17	0.083	1	0	1	1	0.17	n	5.882	0
	15	13	0	0	0.075	0	0.002	0.001	0	IND	no	IND
	11	12	0.024	0.007	1	0	1	1	0.024	n	41.67	0
	11	13	0.247	0.135	1	0	1	1	0.247	n	4.049	0
	12	13	0	0	0.206	0	0.027	0.011	0	IND	no	IND

En el cuadro 30 se puede observar que las tasas de rechazo son muy bajas e incluso cero en algunos casos en la prueba de Wilcoxon, Nemenyi, McDonalds y Thompson en los dos niveles de significación estudiados, esto se debe al carácter conservativo de la prueba la cual utiliza el error por comparación. La prueba de Nemenyi no evidencia tasas de rechazo bajo el nivel de significación 0,05 y presenta tasas de rechazo muy altas e incluso 1 para el nivel de 0,1.

En relación a las eficiencias Pitman, resultaron bajas o cero para la prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson. No se evidencia rechazo alguno con la prueba de

Nemenyi para el nivel de significación de 0,05 pero resulta una eficiencia Pitman muy alta para el nivel de significación de 0,1.

Las eficiencias Bahadur, reportadas en el cuadro 31 indican que estas pruebas no son eficientes en los casos cuyo valor es cero y no se puede concluir al respecto en el caso en que ambas tasas de error dieron 0, pero resulta una alta eficiencia Bahadur en casos de poblaciones distantes.

Cuadro 32. Tasas de error Tipo I para 10 tratamientos semejantes. Diseño B.A. Dentro de los grupos homogéneos.

Tasas de error Tipo I

Trat	k	l	Wil-nem-Mc		Nemeny		Tukey	
			.1	.05	.1	.05	.1	.05
10 TRATAMIENTOS SEMEJANTES	3	7	0	0	0.094	0	0.004	0.001
	3	8	0	0	0.008	0	0	0
	3	14	0	0	0.01	0	0	0
	3	15	0	0	0.182	0	0.02	0.004
	3	12	0	0	0.024	0	0	0
	5	6	0	0	0.083	0	0.016	0.002
	5	13	0	0	0.053	0	0	0
	6	13	0	0	0.112	0	0.007	0.003
	7	8	0	0	0.09	0	0	0
	7	14	0	0	0.082	0	0.007	0.001
	7	15	0	0	0.077	0	0.003	0
	7	12	0	0	0.122	0	0.009	0.002
	8	14	0	0	0.072	0	0	0
	8	15	0	0	0.125	0	0.008	0.002
	8	12	0	0	0.031	0	0	0
	14	15	0	0	0.118	0	0.01	0
	14	12	0	0	0.017	0	0	0
	15	12	0	0	0.124	0	0.007	0.001

En el cuadro 32 se puede observar que las tasas de error Tipo I son cero para ambos niveles de significación en la prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson. Esta es una prueba conservativa y por lo tanto la protección del error tipo I se hace evidente. En la prueba de Nemenyi la cual es liberal, es decir trabaja con error familiar, sólo se evidencian bajas tasas de error tipo I al nivel de 0,1.

Cuadro 33. Tasas de rechazo y eficiencia Pitman para 5 tratamientos contrastantes. Diseño B.A. Entre grupos homogéneos.

			Tasa de rechazo						Eficiencia Pitman			
Trat	k	I	Wil-nem-Mc		Nemeny		Tukey		Wil-nem-Mc		Nemeny	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
5 CONTRASTA	1	2	0	0	0	0	0.754	0.754	0	0	0	0
	1	6	0	0	0	0	1	1	0	0	0	0
	1	10	0	0	0	0	1	1	0	0	0	0
	1	11	0	0	0	0	1	1	0	0	0	0

Cuadro 34. Tasas de rechazo y eficiencias Bahadur para 5 tratamientos contrastantes. Diseño B.A. Entre grupos homogéneos.

			Tasa de rechazo						Eficiencia Bahadur			
Trat	k	I	Wil-nem-Mc		Nemeny		Tukey		Wil/nem		Nem/Wil	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
5 CONTRASTA	1	2	0	0	0	0	0.754	0.754	IND	IND	IND	IND
	1	6	0	0	0	0	1	1	IND	IND	IND	IND
	1	10	0	0	0	0	1	1	IND	IND	IND	IND
	1	11	0	0	0	0	1	1	IND	IND	IND	IND

En el cuadro 33 se puede observar que las tasas de rechazo son cero en ambas pruebas y en los dos niveles de significación estudiados.

En relación a las eficiencias Pitman, resultaron cero para ambas pruebas y en los dos niveles de significación estudiados. No resultaron ser eficientes en relación a su competidor paramétrico.

Las eficiencias Bahadur, reportadas en el cuadro 34 indican que a estas pruebas no se les pudo calcular la eficiencia Bahadur debido a que no presentan rechazo alguno aún en tratamientos contrastantes.

Cuadro 35. Tasas de error y eficiencia Pitman para 5 tratamientos semejantes. Diseño B.A. Dentro de los grupos homogéneos.

Tasas de error Tipo I								
Trat	k	I	Wil-nem-Mc		Nemeny		Tukey	
			.1	.05	.1	.05	.1	.05
5 SEMEJ	7	13	0	0	0	0	0.009	0.006
	7	15	0	0	0	0	0.021	0.014
	13	15	0	0	0	0	0.024	0.016

En el cuadro 35 se puede observar que las tasas de error Tipo I son cero para ambos niveles de significación en la prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson. Esta es una prueba conservativa y por lo tanto la protección del error tipo I se hace evidente. En la prueba de Nemenyi la cual es liberal, es decir trabaja con error familiar, tampoco se cometió error tipo I, a pesar de su carácter liberal.

Cuadro 36. Tasas de error y eficiencias Bahadur para 15 tratamientos contra un control. (Control = Tratamiento 11), todos los pares, diseño B.A.

Diseño B.A. Pruebas c/ Control (Trat 11).Tasas de error									Eficiencia Bahadur			
Trat	k	I	N-Wil-W-M		Hollander		Dunnett		N-W-W-M/Hol		Hol/N-WW-M	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
15 Tratamientos. Todos los pares	11	1	0	0	0	0	0	0	IND	IND	IND	IND
		2	0	0	0	0	0	0	IND	IND	IND	IND
		3	0	0	0	0	0	0	IND	IND	IND	IND
		4	0	0	0	0	0	0	IND	IND	IND	IND
		5	0	0	0	0	0	0	IND	IND	IND	IND
		6	0	0	0	0	0	0	IND	IND	IND	IND
		7	0	0	0	0	0	0	IND	IND	IND	IND
		8	0	0	0	0	0	0	IND	IND	IND	IND
		9	0	0	0	0	0	0	IND	IND	IND	IND
		10	0	0	0	0	0	0	IND	IND	IND	IND
		11	0	0	0	0	0	0	IND	IND	IND	IND
		12	0	0	0	0	0	0	IND	IND	IND	IND
		13	0	0	0	0	0	0	IND	IND	IND	IND
		14	0	0	0	0	0	0	IND	IND	IND	IND
		15	0	0	0	0	0	0	IND	IND	IND	IND

En el cuadro 35 se puede observar que las tasas de rechazo son cero en ambas pruebas y en los dos niveles de significación estudiados, ya que no se obtuvo ningún rechazo.

En relación a las eficiencias Pitman, no se pueden calcular para ambas pruebas y en los dos niveles de significación estudiados.

Las eficiencias Bahadur, reportadas en el cuadro 36 indican que a estas pruebas no se les pudo calcular la eficiencia Bahadur debido a que no presentan rechazo alguno.

Cuadro 37. Tasas de error y eficiencias Bahadur para 10 tratamientos contrastantes contra un control. (Control = Tratamiento 11). Diseño B.A.

Diseño B.A. Pruebas c/ Control (Trat 11).Tasas de error									**Eficiencia Bahadur**			
Trat	k	l	N-Wil-W-M		Hollander		Dunnett		N-W-W-M/Hol		Hol/N-WW-M	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
10 Trat Contratsantes	11	1	0	0	0	0	0	0	IND	IND	IND	IND
		2	0	0	0	0	0	0	IND	IND	IND	IND
		3	0	0	0	0	0	0	IND	IND	IND	IND
		4	0	0	0	0	0	0	IND	IND	IND	IND
		6	0	0	0	0	0	0	IND	IND	IND	IND
		9	0	0	0	0	0	0	IND	IND	IND	IND
		10	0	0	0	0	0	0	IND	IND	IND	IND
		11	0	0	0	0	0	0	IND	IND	IND	IND
		12	0	0	0	0	0	0	IND	IND	IND	IND
		14	0	0	0	0	0	0	IND	IND	IND	IND

En el cuadro 37 se puede observar que las tasas de rechazo son cero en ambas pruebas y en los dos niveles de significación estudiados, ya que no se obtuvo ningún rechazo.

Las eficiencias Bahadur, reportadas en el cuadro 37 indican que a estas pruebas no se les pudo calcular la eficiencia Bahadur debido a que no presentan rechazo alguno.

Cuadro 38. Tasas de error y eficiencias Bahadur para 10 tratamientos semejantes contra un control. (Control = Tratamiento 11). Diseño B.A.

Diseño B.A. Pruebas c/ Control (Trat 11).Tasas de error									Eficiencia Bahadur			
Trat	k	I	N-Wil-W-M		Hollander		Dunnett		N-W-W-M/Hol		Hol/N-WW-M	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
10 Trat Semejantes	11	3	0	0	0	0	0	0	IND	IND	IND	IND
		5	0	0	0	0	0	0	IND	IND	IND	IND
		6	0	0	0	0	0	0	IND	IND	IND	IND
		7	0	0	0	0	0	0	IND	IND	IND	IND
		8	0	0	0	0	0	0	IND	IND	IND	IND
		14	0	0	0	0	0	0	IND	IND	IND	IND
		15	0	0	0	0	0	0	IND	IND	IND	IND
		11	0	0	0	0	0	0	IND	IND	IND	IND
		12	0	0	0	0	0	0	IND	IND	IND	IND
		13	0	0	0	0	0	0	IND	IND	IND	IND

En el cuadro 38 se puede observar que las tasas de rechazo son cero en ambas pruebas y en los dos niveles de significación estudiados, ya que no se obtuvo ningún rechazo.

Las eficiencias Bahadur, reportadas en el cuadro 38 indican que a estas pruebas no se les pudo calcular la eficiencia Bahadur debido a que no presentan rechazo alguno.

Cuadro 39. Tasas de error y eficiencias Bahadur para 5 tratamientos contrastantes contra un control. (Control = Tratamiento 11). Diseño B.A.

Diseño B.A. Pruebas c/ Control (Trat 11).Tasas de error									Eficiencia Bahadur			
Trat	k	I	N-Wil-W-M		Hollander		Dunnett		N-W-W-M/Hol		Hol/N-WW-M	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
5 Trat Contr	11	1	0	0	0	0	0	0	IND	IND	IND	IND
		2	0	0	0	0	0	0	IND	IND	IND	IND
		6	0	0	0	0	0	0	IND	IND	IND	IND
		10	0	0	0	0	0	0	IND	IND	IND	IND
		11	0	0	0	0	0	0	IND	IND	IND	IND

En el cuadro 39 se puede observar que las tasas de rechazo son cero en ambas pruebas y en los dos niveles de significación estudiados, ya que no se obtuvo ningún rechazo.

Las eficiencias Bahadur, reportadas en el cuadro 39 indican que a estas pruebas no se les pudo calcular la eficiencia Bahadur debido a que no presentan rechazo alguno.

Cuadro 40. Tasas de error y eficiencias Bahadur para 5 tratamientos semejantes contra un control. (Control = Tratamiento 11), diseño B.A.

Diseño B.A. Pruebas c/ Control (Trat 11).Tasas de error									Eficiencia Bahadur			
Trat	k	l	N-Wil-W-M		Hollander		Dunnett		N-W-W-M/Hol		Hol/N-WW-M	
			.1	.05	.1	.05	.1	.05	.1	.05	.1	.05
5 Trat Semej	11	5	0	0	0	0	0	0	IND	IND	IND	IND
		7	0	0	0	0	0	0	IND	IND	IND	IND
		13	0	0	0	0	0	0	IND	IND	IND	IND
		15	0	0	0	0	0	0	IND	IND	IND	IND
		11	0	0	0	0	0	0	IND	IND	IND	IND

En el cuadro 40 se puede observar que las tasas de rechazo son cero en ambas pruebas y en los dos niveles de significación estudiados, ya que no se obtuvo ningún rechazo.

Las eficiencias Bahadur, reportadas en el cuadro 40 indican que a estas pruebas no se les pudo calcular la eficiencia Bahadur debido a que no presentan rechazo alguno.

DISCUSIÓN DE RESULTADOS

Una vía de clasificación

Con relación a las tasas de error Tipo I.

Para la configuración de comparaciones de todos los pares, dentro de los grupos homogéneos y comparando los 15 tratamientos, donde se conoce que no existen diferencias entre los grupos, la tasa de error tipo I debería ser baja, sin embargo se observó que las tasas de error tipo I son muy altas e incluso cercanas a 1 en algunos casos en la prueba conservativa de Baumgartner-Weiss y Schindler en los dos niveles de significación estudiados, sin embargo muestra baja tasa de error tipo I en o por debajo del nivel de significación 0,05 solamente en poblaciones distantes. La prueba de Dunn que mostró tasas de rechazo inconsistentes en poblaciones distantes y cercanas donde no se evidencia variación en las tasas de error tipo I independientemente de la "distancia" existente entre las poblaciones. La prueba de Dunn es liberal, es decir trabaja con error familiar según lo señala Dunn (1964), y es posible que esto explique su comportamiento, mientras que la prueba de Baumgartner es de carácter conservativo y así se describe en Neuhäuser y Bretz (2004). Las pruebas de Steel-Dwass, Critchlow y Fligner y Hayter-Stone no evidenciaron tasas de error tipo I distintas de cero, bajo ninguno de los niveles de significación estudiados, probablemente debido a su carácter conservativo que provee protección extra para el error tipo I, señalado por Hollander y Wolfe (1999). La prueba de Tukey mostró bajas tasas de error tipo I en algunos casos y altas en otros mostrando la misma tendencia que la prueba de Dunn.

Para la configuración de 10 tratamientos contrastantes dentro de los grupos homogéneos se puede observar que las tasas de error Tipo I son altas en algunos casos en la prueba de Baumgartner-Weiss y Schindler en los dos niveles de significación estudiados. La prueba de Dunn mostró tasas de error tipo I bajas en poblaciones distantes y cercanas donde no se evidencia variación en las tasas de error Tipo I independientemente de la "distancia" existente entre las poblaciones, lo cual se puede explicar por ser esta una prueba de carácter liberal, es decir trabaja con error familiar. Las pruebas de Steel-Dwass, Critchlow y Fligner y Hayter-Stone no evidenciaron tasas de error Tipo I bajo ninguno de los niveles de significación estudiados, probablemente debido a su carácter conservativo que provee protección extra para el error tipo I. La prueba de Tukey mostró bajas tasas de error Tipo I en algunos casos y cero en otros mostrando inconsistencia en sus resultados.

Para la configuración de 10 tratamientos semejantes se puede observar que las tasas de error Tipo I son altas en todos los casos en la prueba conservativa de Baumgartner-Weiss y Schindler en los dos niveles de significación estudiados. La prueba de Dunn no evidencia variación en las tasas de error Tipo I independientemente de la "distancia" existente entre las poblaciones, pero tuvo mayor control del error tipo I que la prueba de Baumgartner-Weiss, a pesar de que ninguna de las dos logró controlar el error tipo I. Las pruebas de Steel-Dwass, Critchlow y Fligner y Hayter-Stone no evidenciaron tasas de error tipo I distintas de cero bajo ninguno de los niveles de significación estudiados, probablemente debido a su carácter conservativo que provee protección extra para el error tipo I. La prueba de Tukey mostró bajas tasas de error Tipo I en algunos casos y cero en otros mostrando aparente efectividad en el control del error tipo I en sus resultados.

Para la configuración de 5 tratamientos contrastantes, se puede observar que las tasas de error Tipo I son 1 o cercanas a 1 en todos los casos para la prueba conservativa de Baumgartner-Weiss y Schindler en los dos niveles de significación estudiados. La prueba de Dunn mostró tasas de error tipo I bajas en poblaciones distantes y cercanas donde no se evidencia variación en las tasas de error Tipo I independientemente de la "distancia" existente entre las poblaciones, pero tuvo mayor control del error tipo I que la prueba de Baumgartner-Weiss, a pesar de que ninguna de las dos logró controlar el error tipo I. Las pruebas de Steel-Dwass, Critchlow y Fligner y Hayter-Stone no evidenciaron tasas de error Tipo I bajo ninguno de los niveles de significación estudiados, probablemente debido a su carácter conservativo que provee protección extra para el error tipo I, y lo mantuvo por debajo del nivel planteado en las poblaciones más alejadas. La prueba de Tukey mostró bajas tasas de error Tipo I en algunos casos y cero en otros mostrando aparente efectividad en el control del error tipo I en sus resultados.

Las pruebas conservativas de rankeo pareado de Hayter -Stone y de Steel-Dwass,Crtichlow y Fligner controlaron la tasa de error tipo I en ambos niveles de significación, lo cual no logró la prueba de Baumgartner y Weiss a pesar de que también es conservativa y de rankeo pareado. Esta prueba solo mostró control del error tipo I en poblaciones distantes y en el nivel de significación de 0,05 lo cual se puede explicar ya que Neuhäuser (2004) la describe como menos conservativa que las pruebas basadas en rangos asignados de Wilcoxon. La prueba de Dunn definitivamente no muestra consistencia en sus resultados y la tasa de error tipo I demasiado alta la coloca como la peor de las pruebas estudiadas, seguida por la de Baumgartner que no arrojó buenos resultados y las “mejores”, Hayter-Stone y Steel-Dwass-Crichlow y Fligner que por su carácter extremadamente conservativo no detectaron ninguna diferencia.

Para el caso de las pruebas contra un control, la prueba de Nemenyi-Damico y Wolf obtiene tasas de rechazo altas y cercanas a 1 para todos los casos planteados, por lo que se considera que esta prueba no logra controlar el error tipo I bajo ninguna de las configuraciones planteadas.

Con relación a la eficiencia Pitman.

Para el caso de todas las comparaciones de los 15 tratamientos entre grupos homogéneos, en relación a la eficiencia Pitman, esta resultó alta, siendo la misma 1 o mayor a 1 en algunos casos en las pruebas de rankeo conjunto y 0 para las de rankeo pareado, ajustándose a lo planteado en la teoría por Mehra (1972), el cual señala que para una familia de estadísticos de rango la eficiencia Pitman será la misma para la familia completa, ya que en la alternativa normal el estadístico F no depende del número k de tratamientos.

Para el caso de 10 tratamientos contrastantes, las eficiencias Pitman reportadas mostraron resultados altos, 1 o mayor a 1 en algunos casos en las pruebas de rankeo conjunto y 0 para las de rankeo pareado, exceptuando el caso de la prueba de Baumgartner-Weiss cuya eficiencia Pitman fue 0 para las comparaciones de poblaciones distantes.

Para el caso de 10 tratamientos semejantes, las eficiencias Pitman reportadas mostraron resultados muy altos, 1 o superiores a 1 en algunos casos en las pruebas de rankeo conjunto, pero 0 para las de rankeo pareado.

Para el caso de 5 tratamientos contrastantes, las eficiencias Pitman reportadas mostraron resultados muy altos, 1 o superiores a 1 en algunos casos en las pruebas de rankeo conjunto y también altas para las de rankeo pareado, con la excepción de la prueba de Steel-Dwass-Critchlow y Fligner, en la que la eficiencia Pitman fue 0 para el nivel de significación de 0,05.

Para el caso de 5 tratamientos semejantes, las eficiencias Pitman reportadas mostraron resultados muy altos, 1 o superiores a 1 en algunos casos en las pruebas de rankeo pareado y también para las de rankeo conjunto, con la excepción de la prueba de Steel-Dwass-Critchlow y Fligner, en la que la eficiencia Pitman fue 0 para el nivel de significación de 0,05. La prueba de Hayter-stone muestra eficiencia Pitman 0 para poblaciones alejadas.

Para el caso de pruebas con un control, para los 15 tratamientos y los 10 tratamientos semejantes y contrastantes, las eficiencias Pitman, reportadas de la prueba de nemenyi, Damico y Wolf muestran resultados relativamente altos, cercanos a 1 o 1 en algunos casos para poblaciones distantes. Para los 5 tratamientos contrastantes, la eficiencia

Pitman, reportada muestra resultados altos, evidenciando ser tan eficiente como la prueba de Dunnett. Al igual que para los 5 tratamientos semejantes.

No se evidencia en estos análisis lo señalado por Mehra (1972), al encontrar que las pruebas de rankeo conjunto son más eficientes que las de rankeo pareado en el sentido Pitman, siendo la mejor de ellas la de Baumgartner y menos eficientes las pruebas de Steel, Dwass, Crtitchlow y Fligner y la de Hayter Stone. Por otro lado se encontró que para el caso de 5 tratamientos en el cual se tienen pocas comparaciones todas las pruebas tienen un comportamiento similar, sin distinción alguna entre las de rankeo conjunto o pareado.

Con relación a la eficiencia Bahadur.

Para el caso de todas las comparaciones de los 15 tratamientos entre grupos homogéneos, en relación a la eficiencia Bahadur reportadas mostraron valores mayores a 1 solamente en las pruebas más liberales y poblaciones cercanas. En las demás no se evidenció eficiencia alguna.

Para el caso de 10 tratamientos contrastantes, al igual que para 10 tratamientos semejantes las eficiencias Bahadur reportadas mostraron valores mayores a 1 solamente entre las pruebas más liberales y poblaciones cercanas. En las demás no se evidenció eficiencia alguna.

Para el caso de 5 tratamientos contrastantes, las eficiencias Bahadur reportadas mostraron valores mayores a 1 solamente entre las pruebas más conservativas y poblaciones cercanas, mostrando que la prueba de Steel-Dwass-Critchlow y Fligner, parece ser la que tiene una eficiencia Bahadur mayor con relación a las otras pruebas estudiadas.

Para el caso de 5 tratamientos semejantes, las eficiencias Bahadur reportadas mostraron valores mayores a 1 solamente entre las pruebas más conservativas y poblaciones cercanas al nivel de significación 0,1. Las pruebas de Dunn y Baumgartner-Weiss tienen eficiencias Bahadur similares.

Para el caso de alternativas con un control no se calcularon las eficiencias Bahadur ya que solamente se presenta una prueba para este caso.

No se encontraron eficiencias Bahadur distintas de cero para el caso de pruebas de rankeo pareado con más de 5 tratamientos, lo que indica que estas pruebas no resultan eficientes para muchos tratamientos. Pero con 5 tratamientos las pruebas de rankeo pareado reportan resultados que las señalan con una eficiencia Bahadur mayor que las de rankeo conjunto. La prueba de Baumgartner es eficiente para todos los números de tratamientos estudiados, y se cataloga en este sentido como la mejor, ya que con relación a la prueba de rankeo conjunto de Dunn tiene una eficiencia mayor a 1.

Si se prueban pocos tratamientos las pruebas de rankeo pareado son mejores ya que el error tipo I queda completamente controlado, pero si se tienen más de 10 comparaciones la prueba de Baumgartner demuestra tener mayor eficiencia y debido control del error tipo I.

Dos vías de clasificación

Con relación a las tasas de error Tipo I.

Para la configuración de comparaciones de todos los pares, dentro de los grupos homogéneos y comparando los 15 tratamientos, donde se conoce que no existen diferencias entre los grupos, la tasa de error tipo I debería ser baja y se observó que las tasas de error tipo I son cero para ambas pruebas en los dos niveles de significación. La prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson es una prueba conservativa según lo señalan Hollander y Wolfe (1999) y por lo tanto la protección del error tipo I se hace evidente pero la prueba de Nemenyi es liberal y de rankeo conjunto como lo destacan Hollander y Wolfe (1999), es decir trabaja con error familiar. A pesar de la diferencia, ninguna de las dos evidencia tasas de error tipo I distintas de cero bajo ninguno de los niveles de significación estudiados.

Para la configuración de 10 tratamientos contrastantes dentro de los grupos homogéneos se puede observar que las tasas de error Tipo I son cero para ambos niveles de significación en la prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson. Esta es una prueba conservativa según lo señalan Hollander y Wolfe (1999) y por lo tanto la protección del error tipo I se hace evidente pero la prueba de Nemenyi es liberal y de rankeo conjunto como lo destacan Hollander y Wolfe (1999), es decir trabaja con error familiar. A pesar de la diferencia, ninguna de las dos evidencia tasas de error tipo I distintas de cero bajo ninguno de los niveles de significación estudiados.

Para la configuración de 10 tratamientos semejantes se puede observar que las tasas de error Tipo I son cero para ambos niveles de significación en la prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson. Esta es una prueba conservativa según lo señalan Hollander y Wolfe (1999) y por lo tanto la protección del error tipo I se hace evidente. En la prueba de Nemenyi la cual es liberal y de rankeo conjunto como lo destacan Hollander y Wolfe (1999). A pesar de la diferencia, ninguna de las dos evidencia tasas de error tipo I distintas de cero bajo ninguno de los niveles de significación estudiados.

Para la configuración de 5 tratamientos semejantes, se puede observar que las tasas de error Tipo I son cero en todos los casos para ambos niveles de significación en la prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson. Esta es una prueba conservativa según lo señalan Hollander y Wolfe (1999) y por lo tanto la protección del error tipo I se hace evidente. En la prueba de Nemenyi la cual es liberal y de rankeo conjunto como lo destacan Hollander y Wolfe (1999), es decir trabaja con error familiar, tampoco se cometió error tipo I a ningún nivel de significación.

Con relación a los resultados obtenidos no se puede determinar que exista una prueba mejor ya que se comportaron igual en cuanto a la tasa de error tipo I, sin distinción de su carácter conservativo o liberal, lo cual muestra que no se evidenció lo señalado por Lugo (2006), que refiere que las pruebas de rankeo pareado controlan la tasa de error mejor que las de rankeo conjunto.

Con relación a la eficiencia Pitman.

Para el caso de todas las comparaciones de los 15 tratamientos entre grupos homogéneos, en relación a la eficiencia Pitman, resultaron bajas o cero para la prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson exceptuando el caso de la comparación con el grupo 11 el cual se encuentra muy distante del 6 con relación a sus medianas. En ésta comparación la eficiencia Pitman de la prueba fue de 64,6 por ciento para el nivel de significación de 0,1. No se evidencia rechazo alguno con la prueba de Nemenyi, por lo tanto no es eficiente en el sentido Pitman bajo ninguno de los niveles de significación estudiados.

Para el caso de 10 tratamientos contrastantes, las eficiencias Pitman reportadas resultaron bajas o cero para la prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson. No se evidencia rechazo alguno con la prueba de Nemenyi para el nivel de significación de 0,05 pero resulta una eficiencia Pitman muy alta para el nivel de significación de 0,1.

Para el caso de 10 tratamientos semejantes, las eficiencias Pitman, resultaron bajas o cero para la prueba de Wilcoxon, Nemenyi, Mcdonalds y Thompson. No se evidencia rechazo alguno con la prueba de Nemenyi para el nivel de significación de 0,05 pero resulta una eficiencia Pitman muy alta para el nivel de significación de 0,1.

Para el caso de 5 tratamientos contrastantes, en relación a las eficiencias Pitman, resultaron cero para ambas pruebas y en los dos niveles de significación estudiados. Ninguna prueba resulta ser eficientes en relación a su competidor paramétrico.

La prueba de Nemenyi resulta ser la más eficiente en el sentido Pitman solo al nivel de significación de 0,1 para 10 tratamientos semejantes, debido su carácter liberal, lo cual no apoya lo señalado por Fligner (1985), que refiere que las pruebas de rankeo pareado son más eficientes que las de rankeo conjunto en el sentido Pitman. Tampoco apoya este resultado lo señalado por Mehra (1972), que plantea que para una familia de estadísticos de rango la eficiencia Pitman será la misma.

Para las pruebas contra un control no se les pudo calcular la eficiencia Pitman debido a que no se obtuvo ningún rechazo con las pruebas no paramétricas ni con su competidor paramétrico Dunnett.

Con relación a la eficiencia Bahadur.

Para el caso de todas las comparaciones de los 15 tratamientos entre grupos homogéneos, en relación a la eficiencia Bahadur no se les pudo calcular la eficiencia Bahadur debido a que no se obtuvo ningún rechazo con las pruebas.

CONCLUSIONES

1. Para el caso de una vía de clasificación, las pruebas que presentaron la menor tasa de error tipo I, fueron las de rankeo pareado de Hayter-Stone y Steel-Dwass-Critchlow y Fligner para las 5 configuraciones propuestas y en los dos niveles de significación estudiados.

2. Para el caso de dos vías de clasificación, la prueba de rankeo conjunto de Wilcoxon-Nemenyi-McDonalds y Thompson se comportó igual que la prueba de rankeo conjunto de Nemenyi. Ambas ejercieron un control total del error tipo I, en las 5 configuraciones propuestas y en los dos niveles de significación estudiados.

3. Para el caso de una vía de clasificación, la prueba que presentó la mayor eficiencia Pitman, fue la prueba de rankeo pareado de Baumgartner-Weiss y Schindler para las 5 configuraciones propuestas y en los dos niveles de significación estudiados. Para las comparaciones de 5 tratamientos, todas las pruebas estudiadas tiene comportamiento similar en el sentido Pitman. La prueba de Nemenyo-Damico y Wolf contra un control muestra alta eficiencia Pitman (1 o mayor) para poblaciones distantes pero 0 para poblaciones cercanas.

4. Para el caso de dos vías de clasificación, ninguna de las pruebas estudiadas resultó eficiente en el sentido Pitman en ninguna de la configuraciones para ambos niveles de significación, exceptuando el caso de la prueba de Wilcoxon-Nemenyi-McDonalds y Thompson que tuvo una eficiencia Pitman de 64,6 por ciento al comparar poblaciones distantes, en la configuración 1 (todos los pares) y en el nivel de significación 0,1.

5. Para el caso de una vía de clasificación, en relación a las eficiencias Bahadur, las pruebas liberales de Dunn y Baumgartner-Wiess y Schindler presentaron eficiencias Bahadur superiores siendo la de Dunn la que mostró ser más eficiente

en el sentido Bahadur con relación a la de Baumgartner-Weiss y Schindler, a pesar de ser esta última una prueba conservativa de rankeo pareado. Las pruebas conservativas de Hayter-Stone y Steel-Dwass-Critchlow y Fligner tienen eficiencia 0 con relación a todas las demás para las configuraciones de más de 10 tratamientos. Pero en las configuraciones de 5 tratamientos las pruebas de rankeo pareado poseen mayor eficiencia Bahadur que la de rankeo conjunto de Dunn y más eficiencia que la menos conservativa de Baumgartner-Weiss y Schindler.

6. Para el caso de dos vías de clasificación, en relación a las eficiencias Bahadur, no se pudo calcular debido a que no se obtuvo ningún rechazo con las pruebas.

REFERENCIAS

1. Alonzo, T., Nakas, C., Yiannoutsos, C. y Bucher, S.(2009). A comparison of test for restricted orderings in the thre-class case. *Statistics in Medicine.* Vol. 28 pp 1144-1158.
2. Baumgartner, W., Weiβ, P. y Schindler, H. (1998). A nonparametric test for the general two sample problem. *Biometrics,* Vol. 54 pp 1129-1135.
3. Bristol, D. (1990). Comparison of two distribution-free procedures for multiple comparisons with a control. *Communications in Statistics- Simulation and Computation,* Vol.19, No. 4, pp 1403-1413.
4. Chakraborti, S. y Desu, M. (1988). Generalizations of Mathisen`s median test for comparing several treatments with a control. *Communications in Statistics- Simulation and Computation,* Vol.17, No. 3, pp 947-967.
5. Chew, V. (1976). Comparing treatment means: A compendium. *Hortscience.* Vol. 11, No. 4 pp 348-357.
6. Critchlow, D. y Fligner, M. (1991). On distribution-free multiple comparisons in the one-way analysis of variance. *Commun. Statist.- Theory meth.* Vol. 20, No. 1 pp 127-139.
7. Dunn, O. (1964). Multiple comparison using rank sums. *Technometrics*, Vol. 6, No. 3 pp 241-252.
8. Dwass, C. (1960). Some k-sample rank-order test. *Contributions to probability and Statistics.* Pp 198-202.
9. Fairley, D. y Pearl, (1984). The bahadur efficiency of paired versus joint ranking procedures for pairwise multiple comparisons. *Communications in Statistics- Theor. Meth,* Vol.13, No. 12, pp 1471-1481.
10. Federer, W. (1955). Experimental Error Rates. *American Society for Horticultural Sciences.* Vol. 78 pp 605-615.
11. Fligner, M. Pairwise versus joint-ranking: Another look at the Kruskal-Wallis statistic. *Biometrika* (1985). Vol. 72, No. 3 pp 705-709.

12. Garrido, M., Freites-Abreu, J., Ascanio, A. y González, M. (2009). Reacción de cultivares de sorgo al potyvirus del mosaico del pasto Johnson (JGMV). Resúmenes XXI Congreso Venezolano de Fitopatología, Isla de Margarita, Edo. Nueva Esparta, Venezuela.
13. Grané, A. y Tchirina, A.(2008). Asymptotic properties of a goodness of fit test based on maximum correlations. *Statistics an Econometrics Series 11.Working papers. Universidad Carlos III de Madrid.* WP 08-42.
14. Hollander, M. (1966). An asymptotically distribution free multiple comparison procedure. Treatment versus control. *The Annals of Mathematical Statistics.* Vol. 37 pp 735-738.
15. Hollander, M. y Wolfe, D. (1999). A nonparametric statistical methods. 2.ed. New York: John Wiley and Sons.
16. Koziol, J. y Reid, N.(1977). On the asymptotic equivalence of two ranking methods for k-sample linear rank statistics. *Annals of Statistics.* Vol. 5 pp 1099-1106.
17. Kruskal, W. y Wallis, W. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistics Association.* Vol. 47 pp 583-621.
18. Lugo, L. (2006). Trabajo de Grado. Maestría en Estadística. Universidad Central de Venezuela. Facultad de Agronomía.
19. Mehra, K. (1972). On bahadur efficiency of the joint-ranking procedure. *The Annals of Mathematical Statistics.* Vol. 43, No.4 pp 1155-1163.
20. Neuhäuser, M. y Bretz, F. (2001). Nonparametric all-pairs multiple comparisons. *Biometrical Journal.* Vol. 43, No. 5 pp 571-580.
21. Neuhäuser, M. y Senske, R. (2004). The Baugmgartner-Weiss-Schindler test for the detection of differentially expressed genes in replicated microarray experiments. *Bioinformatics.* Working paper. 33 p.
22. Slivka, J. (1970). A one sided nonparametric multiple comparison control percentile test: treatments versus control. *Biometrika,* Vol.57, pp432-438.

23. Steel, R. A rank sum test for comparing all pairs of treatments. *Technometrics*, Vol. 2, No. 2 pp 197-207.