

Robustez de las pruebas T en comparación de medias, ante violación de supuestos de normalidad y homocedasticidad

Robustness of the t tests in comparison of means, under violation of normality and homoscedasticity assumptions'

Montilla, Josefa-Maria

Área de Estadística. Departamento de Ciencias Económicas Administrativas y Contables. ULA
Trujillo-Venezuela.
josefam76@cantv.net

Kromrey, Jeffrey

Department of Measurement & Research. College of Education. University of South Florida.
Tampa, Florida-USA

Recibido: 18-12-2009

Revisado: 02-07-2010

Resumen

Esta investigación examina la robustez de los contrastes paramétricos t-Student, t'-Welch y t-Yuen, cuando los tamaños de las muestras son iguales/desiguales y bajo el efecto de violación de los supuestos de normalidad y homocedasticidad. Este trabajo está referido al estudio de simulación Monte Carlo, usado para estimar las características de las tasas de error tipo I de los tests bajo violación de sus supuestos. El estudio de simulación Monte Carlo fue llevado a cabo usando el software estadístico SAS, versión 6.12. Las muestras fueron generadas a partir de poblaciones con características conocidas, examinadas bajo una variedad de condiciones, definidas por el tamaño de las muestras, la igualdad/desigualdad de varianzas, la magnitud de las diferencias entre las medias y la forma de la distribución (asimetría y kurtosis) de las poblaciones. La precisión de la tasa observada de error tipo I τ , cuando es comparada con la tasa nominal de error tipo I α , fue cuantificada usando el criterio liberal intermedio de robustez de Bradley. Los resultados de las estimaciones de las tasas de error tipo I mostraron que los tests t-Student y t'-Welch proporcionan un mejor control de estas tasas que el test t*-Yuen bajo las condiciones de normalidad y homocedasticidad; el test Welch es más robusto que el test de Student y éste, a su vez, más que el test de Yuen bajo las condiciones de normalidad, ausencia de normalidad y heterocedasticidad. En general se observó que los tests examinados son más sensibles al incremento en la desigualdad de varianzas que en la condición de ausencia de normalidad.*

Palabras clave: Robustez, simulación, tasa de error tipo I, normalidad, homocedasticidad.

Abstract

This research examines the robustness of the parametric contrasts t-Student, t'-Welch and t-Yuen, when the samples sizes are equal/unequal and under the effect of violation of the assumptions of normality and homoscedasticity. This investigation is referred to a Monte Carlo simulation study, used to estimate the characteristics of the rates of type I error under violation of its assumptions. The Monte Carlo simulation study was determined using the statistical software SAS, version 6.12. The samples were generated from populations with well-known characteristics, examined under a variety of conditions, defined by the size of the samples, the equality/inequality the variances, the magnitude of the differences between the means and the population's distribution (skewness and kurtosis). The precision of the observed rate of type I error, when compared to the nominal rate α , was quantified using the Bradley's intermediate liberal criterion of robustness. The results of the estimates of the rates of type I error showed that the Student and Welch tests provide a better control on these rates than the t*-Yuen under the condition of normality and homoscedasticity; the t'-Welch test is more robust than the t-Student test and this one, more than the Yuen test under the conditions of normality, absence of normality and heteroscedasticity. In general we have*

observed that the examined tests are more sensitive to the increment in inequality of the variances than to the absence of normality condition.

Key words: Robustness, simulation, rate of type I error, normality, homoscedasticity.

1 Introducción

Entre las técnicas estadísticas más ampliamente conocidos y usados para confrontar el problema de probar la igualdad de dos medias poblacionales independientes destacan los procedimientos t-tests. Así, para producir conclusiones válidas a partir de un conjunto dado de datos, es necesario satisfacer ciertos supuestos básicos, como normalidad de la distribución poblacional, homocedasticidad e independencia de las observaciones. Sin embargo, frecuentemente surgen series de datos que no satisfacen esos supuestos, lo cual indica que la robustez se ve afectada por las formas de las distribuciones y la desigualdad de las varianzas; en consecuencia, existe la necesidad de evaluar estos procedimientos para comprobar cómo se comportan frente a la violación de los supuestos básicos de normalidad y homocedasticidad.

Uno de los métodos utilizados para determinar si estos tests son robustos a la violación de sus supuestos de normalidad y homocedasticidad es el experimento Monte Carlo (MC). Este experimento es usado en la determinación de la distribución del test, cuando números aleatorios son generados bajo simulaciones estocásticas y para investigar la robustez de los mismos. Esta investigación está dirigida a examinar específicamente la robustez de los tests t-Student, t'-Welch y t*-Yuen en comparación de medias poblacionales independientes; cuando los tamaños de las muestras son iguales/desiguales y bajo el efecto de violación de normalidad y homocedasticidad. Además, son usados con la finalidad de producir resultados que son estadísticamente robustos cuando los supuestos básicos son asumidos. Sin embargo, el test t-Student en términos de las tasas de error tipo I: tasa nominal (α) y tasa observada (τ), bajo violación de los supuestos de normalidad y homocedasticidad, puede ser insatisfactorio. La tasa observada y la tasa nominal de error tipo I son definidas como un porcentaje de los resultados significantes obtenidos, y como la probabilidad de erróneamente concluir que la magnitud de las medias poblacionales es desigual, respectivamente. Por consiguiente, en estos casos los investigadores pueden elegir entre esos procedimientos considerados como tests alternativos robustos al test t-Student para muestras independientes.

Con este propósito, esta investigación determina y compara las tasas de error tipo I de los t-tests cuando a) los tamaños de las muestras son iguales/desiguales y bajo los supuestos de normalidad y homocedasticidad; b) los tamaños de las muestras son iguales/desiguales y bajo el supuesto de normalidad y heterocedasticidad (desigualdad de varianzas); c) los tamaños de las muestras son iguales/desiguales y bajo el supuesto de homocedasticidad y ausencia

del supuesto de normalidad; y d) los tamaños de las muestras son iguales/desiguales y bajo el doble efecto de violación de las asunciones de normalidad y homocedasticidad. En el proceso de investigación la revisión de literatura representa una de las fases de gran importancia, puesto que permite conocer no sólo lo que se ha hecho con relación a la investigación de interés, sino también lo relacionado con el estado actual de la misma. En investigaciones relacionadas a la robustez de los t-tests, han estudiado el uso de los mismos bajo las condiciones de normalidad, homocedasticidad, ausencia de normalidad y heterocedasticidad. Así mismo, han encontrado que los t-test de Welch y Yuen, considerados como test alternos al test de Student, son generalmente superiores en situaciones bajo violación de los supuestos de normalidad y homocedasticidad (Sánchez-Bruno y col. 2000; Sánchez-Bruno y Borges, 1997; Borges y col. 1996; Burton y col. 1996; Borges y col. 1993; Brown y Forsythe, 1974; Yuen, 1974; Scheffé, 1970; Cochran 1964; Boneau, 1960; Aspin, 1948; Welch, 1947).

Otros estudios realizados han demostrado que cuando los tamaños de muestras son iguales ($n_1 = n_2$), la probabilidad observada del error tipo I (τ) es muy aproximada a la probabilidad nominal del error tipo I (α), y el efecto de la heterocedasticidad sobre esta probabilidad es insignificante; y cuando los tamaños de muestras son incrementadas, el efecto de violación de normalidad sobre la probabilidad del error tipo I no es importante. Estas investigaciones han reportado la relevancia que tienen estos tests, como alternativa al test t-Student, para probar la igualdad de las medias poblacionales independientes (Glass, y Hopkins, 1996; Boneau, 1962; Scheffé, 1959).

2 Diseño y metodología de la investigación

Esta investigación está referida a un estudio de simulación Monte Carlo; asimismo, cuando aplicamos el experimento de robustez Monte Carlo, deben ser considerados dos aspectos relacionados de gran importancia: 1) el criterio de tolerancia, es decir, establecer cuán amplia debe ser la discrepancia entre la tasa nominal α y la tasa observada τ ; y 2) el número de repeticiones llevadas a cabo en las series de simulación, de tal manera que el test pueda ser interpretado con confianza.

Los estudios de Monte Carlo han mostrado ser procedimientos eficaces en la determinación de la distribución del test estadístico, cuando está siendo investigada la robustez del mismo, y cuando se están generando números aleatorios bajo simulaciones estocásticas (Robey, & Barcikowshi, 1992). Por tanto, este experimento es usado para estimar las características de las tasas de error tipo I (tasa

nominal y tasa observada) de los t-tests bajo violación de sus supuestos. Consecuentemente, este procedimiento se utilizó en la creación de las distribuciones para la presente investigación, y para comparar las tasas de error tipo I de los tests t-Student, t'-Welch y t*-Yuen, como una función de la magnitud de las diferencias entre las medias.

Para llevar a cabo los análisis de esta investigación, el estudio de simulación Monte Carlo fue conducido usando el software estadístico SAS (Statistical Analysis Software) en su versión 6.12. Las muestras fueron generadas de poblaciones con características conocidas. Estas características fueron examinadas bajo una variedad de condiciones, definidas por el tamaño de las muestras (n_1, n_2), la igualdad/desigualdad de varianzas ($\sigma_1^2 = \sigma_2^2$; $\sigma_1^2 \neq \sigma_2^2$), la magnitud de las diferencias entre las medias (d) y la forma de la distribución (asimetría y kurtosis) de las poblaciones.

Con la finalidad de generar las distribuciones normales (Asimetría=0, Kurtosis=0) fue usada la función RANNOR del SAS (generadora de números aleatorios). Las distribuciones no-normales fueron simuladas usando la transformación de una función polinomial desarrollada por Fleishman, conocida con el nombre de "Método Poder de Fleishman" ($Y = -c + bX + cX^2 + dX^3$). Este método fue usado para variar de una manera sistemática la asimetría y la kurtosis para los dos grupos independientes (Fleishman A.I. 1978). Así, las distribuciones no-normales fueron basadas sobre los siguientes valores pareados de asimetría y kurtosis (1, 2), (0.5, 0.5) y (0.5, -0.5), respectivamente. Los tamaños de las muestras seleccionadas para los dos grupos considerados fueron (n_1, n_2): para muestras iguales: (5, 5), (20, 20) y (50, 50) y para muestras desiguales, las siguientes combinaciones: (5, 20), (5, 50), (20, 50), (20, 5), (50, 5) y (50, 20). Se consideró un tamaño de muestra cincuenta (50) para comparativamente establecer la robustez de los tests, cuando los tamaños de muestras grandes son pareados con las varianzas más grandes y con las más pequeñas. Las varianzas utilizadas fueron (1:1), para la condición de homocedasticidad y para la condición de heterocedasticidad (1:2) y (1:4). La magnitud de las diferencias entre las medias fue establecida de cero (0). Además, para evaluar los procedimientos usados en esta investigación fue usado el nivel de significancia del 5 %. Finalmente, con la finalidad de proporcionar comparaciones precisas fueron generadas 10.000 repeticiones para cada una de las condiciones específicas.

En cada uno de los procedimientos usados, la precisión de la tasa observada de error tipo I (τ) cuando es comparada con la tasa nominal de error tipo I (α), fue cuantificada usando el criterio de robustez de Bradley. Para juzgar la robustez de un test, Bradley proporciona una serie de criterios que deben ser considerados por el investigador en el momento de evaluar los resultados. Estos criterios están basados sobre intervalos aceptables para τ , y están en función de la tasa nominal de error tipo I, α , establecida por el investigador y son los siguientes:

- Un criterio liberal donde la amplitud del intervalo es ($\alpha_{\text{nominal}} \pm .50\alpha_{\text{nominal}}$);

- Un criterio riguroso cuya amplitud del intervalo es ($\alpha_{\text{nominal}} \pm .10\alpha_{\text{nominal}}$);
- Un criterio liberal intermedio con un intervalo definido ($\alpha_{\text{nominal}} \pm .25\alpha_{\text{nominal}}$); y
- Un criterio muy liberal con una amplitud del intervalo de ($\alpha_{\text{nominal}} \pm .75\alpha_{\text{nominal}}$).

Por consiguiente los tests fueron interpretados considerando el criterio liberal intermedio de Bradley. Así, los t-tests examinados proporcionan un control adecuado del error tipo I, si la tasa observada estimada τ está entre el intervalo siguiente: ($\alpha_{\text{nominal}} \pm .25\alpha_{\text{nominal}}$), equivalente a (.0375 - .0625), respectivamente (Bradley, 1978).

3 Resultados

Los resultados de las simulaciones expresadas en términos de las tasas de error tipo I de los tests t-Student, t'-Welch y t*-Yuen se muestran en las tablas del 1 al 5. Estas estimaciones fueron computadas y analizadas bajo las siguientes condiciones:

3.1 Tasas de error tipo I bajo normalidad y homocedasticidad

Los resultados presentados en la tabla 1 muestran los estimados de las tasas de error tipo I asumiendo normalidad e igualdad de varianzas. Los valores obtenidos de las tasas de error tipo I para los tests considerados, bajo la combinación de tamaños de muestras iguales y desiguales, muestran que los tests investigados mantienen un valor observado de alpha (α) dentro de los límites del criterio de Bradley ($\alpha_{\text{nominal}} \pm .25\alpha_{\text{nominal}}$), lo cual indica que son robustos con respecto a los supuestos de normalidad y homocedasticidad, excepto el test de Yuen, el cual no es robusto bajo las siguientes combinaciones de muestras desiguales: 5, 20; 5, 50; 20, 5; y 50, 5.

Tabla 1. Estimaciones de error tipo I de los tests

Alpha 0.05; Magnitud de Diferencias=0; Asimetría = 0; Kurtosis = 0.				
Tamaño de Muestra		Razón de Varianzas=1:1		
n_1, n_2	t-Student	t'-Welch	t*-Yuen	
5, 5	0.0513	0.0453	0.0441	
20, 20	0.0491	0.0486	0.0516	
50, 50	0.0507	0.0507	0.0519	
5, 20	0.0552	0.0598	0.0947	
5, 50	0.0515	0.0593	0.0834	
20, 50	0.0462	0.0453	0.0489	
20, 5	0.0505	0.0579	0.0814	
50, 5	0.0469	0.0517	0.0850	
50, 20	0.0508	0.0459	0.0508	

Estos resultados muestran que bajo estas condiciones establecidas, los tests t-Student y t'-Welch, proporcionan un mejor control de las tasas de error tipo I que el test de Yuen. En otras palabras, estos resultados revelan que los tests t-Student y t'-Welch tienen una probabilidad de rechazar la hipótesis nula cuando ésta es realmente verdadera, más aproximada a la probabilidad nominal ($\alpha_{\text{nominal}} \pm .25\alpha_{\text{nominal}}$), que la proporcionada por el test t-Yuen.

3.2 Tasas de error tipo I bajo normalidad y heterocedasticidad

Los estimados de las tasas de error tipo I para los t-tests que asumen normalidad y desigualdad de varianzas, representadas por las razones de varianzas 1:2 y 1:4, se presentan en la tabla 2. Estos resultados revelan que el test t-Student no es robusto bajo violación de estos supuestos (normalidad y heterocedasticidad), excepto cuando las

muestras son grandes e iguales (50, 50), así, la t-Student tiene una probabilidad de error tipo I que es considerada adecuada siguiendo el criterio de robustez (intermedio-liberal) establecido por Bradley (.0375 - .0625). Bajo estas condiciones el test t'-Welch se transforma en el test más robusto cuando comparamos con los tests t-Student y t*-Yuen. Sin embargo, al comparar la t-Student con el t*-Yuen, este último se transforma en un test más conservador cuando las muestras son grandes e iguales (20, 20; 50, 50), para las combinaciones de muestras más grandes (20, 50; 50, 20), y cuando las muestras más grandes (50, 5) son seleccionadas de las poblaciones con las más altas varianzas (razón de varianza de 1:4). Estos resultados revelan la susceptibilidad del test t-Student bajo violación del supuesto de homogeneidad de varianzas.

Tabla 2. Estimaciones de error tipo I de los tests

Alpha 0.05; Magnitud de Diferencias = 0; Asimetría = 0; Kurtosis = 0						
Tamaño de Muestra	Razón de Varianzas = 1:2			Razón de Varianzas = 1:4		
n_1, n_2	t-Student	t'-Welch	t*-Yuen	t-Student	t'-Welch	t*-Yuen
5, 5	0.0647	0.0538	0.0524	0.0703	0.0498	0.0688
20, 20	0.0537	0.0509	0.0540	0.0521	0.0465	0.0566
50, 50	0.0524	0.0512	0.0494	0.0503	0.0484	0.0504
5, 20	0.1785	0.0532	0.0835	0.3238	0.0515	0.0801
5, 50	0.2622	0.0562	0.0841	0.0464	0.0524	0.0717
20, 50	0.1412	0.0516	0.0592	0.1979	0.0494	0.0526
20, 5	0.0075	0.0554	0.0618	0.0110	0.0512	0.0480
50, 5	0.0010	0.0566	0.0775	0.0000	0.0462	0.0547
50, 20	0.0120	0.0540	0.0544	0.0049	0.0501	0.0512

3.3 Tasas de error tipo I bajo ausencia de normalidad y homocedasticidad

Las estimaciones de las probabilidades de las tasas de error tipo I para los t-tests examinados en circunstancias donde las varianzas son iguales y las distribuciones son no-normales, se muestran en la tabla 3. Así, bajo igualdad de varianzas (1:1) y cuando la asimetría asume valores de 1.0 y .05 y la Kurtosis es 2.0 y -0.5, los tests explorados muestran resultados similares y conservadores, proporcionando una precisión a través de todos los niveles de asimetría y kurtosis usados y sin considerar el tamaño de las muestras. La excepción ocurre con el test de Yuen cuando las combinaciones de los tamaños de muestras son desiguales (5, 20; 5, 50; 20, 5; y 50, 5); en cuyo caso las distribuciones asimétricas fueron asociadas con tasas de error tipo I ligeramente altas (más liberales). Consecuentemente, bajo estas condiciones, los tests t-Student y t'-Welch muestran ser más ro-

bustos que el test de Yuen.

3.4 Tasas de Error Tipo I Bajo Ausencia de Normalidad y Heterocedasticidad

Las tasas de error tipo I de los tests en circunstancias donde las varianzas son desiguales y las distribuciones no son normales se presentan en las tablas 4 y 5.

Las estimaciones de probabilidad cuando la desigualdad de varianzas es representada por la razón 1:2 y la asimetría y kurtosis asumen los valores de 1.0, .50 y 2.0, -.50, respectivamente, se muestran en la tabla 4. Como puede observarse, los tests considerados presentan resultados similares: 1) las tasas de error tipo I para la t-Student son conservadoras cuando los tamaños de los tamaños de muestras son iguales, y son liberales cuando los tamaños de muestras son desiguales; 2) el t'-Welch es robusto con respecto a la ausencia de normalidad y heterocedasticidad; 3) el t*-Yuen es robusto cuando los tamaños de muestras son iguales y las

combinaciones de tamaños de muestras desiguales son 20, 50 y 50, 20; y 4) los tres tests examinados muestran tasas de error tipo I dentro del criterio conservador, excepto cuando las combinaciones de los tamaños de muestras desiguales son 5, 20; 5, 50; 20, 5 y 50, 5. Estos resultados muestran

que el test con mayor robustez es el t'-Welch, con relación a los otros; y la robustez del test t*-Yuen con respecto al test t-Student; lo cual indica que el más sensible a la violación de los supuestos de normalidad y homocedasticidad es el test t-Student.

Tabla 3. Estimaciones de error tipo I de los tests

Tamaño de Muestra	Asimetría = 1.0; Kurtosis = 2.0			Asimetría = 0.5; Kurtosis = - 0.5		
	t-Student	t'-Welch	t*-Yuen	t-Student	t'-Welch	t*-Yuen
n ₁ , n ₂						
5, 5	0.0487	0.0423	0.0421	0.0535	0.0477	0.0434
20, 20	0.0547	0.0541	0.0524	0.0503	0.0500	0.0491
50, 50	0.0541	0.0539	0.0550	0.0501	0.0500	0.0486
5, 20	0.0501	0.0563	0.0772	0.0496	0.0594	0.0831
5, 50	0.0457	0.0536	0.0853	0.0513	0.0600	0.0913
20, 50	0.0515	0.0511	0.0517	0.0522	0.0498	0.0523
20, 5	0.0497	0.0568	0.0791	0.0051	0.0504	0.0849
50, 5	0.0485	0.0583	0.0891	0.0495	0.0583	0.0927
50, 20	0.0406	0.0505	0.0506	0.0497	0.0500	0.0539

Las estimaciones de probabilidad de las tasa de error tipo I, bajo las mismas condiciones de ausencia de normalidad (la asimetría y kurtosis los valores de 1.0, .50 y 2.0, -.50, respectivamente) y cuando la desigualdad de varianzas es representada por la razón 1:4 se asumen observa en la tabla 5. Estos resultados revelan que los tests: t-Student, y t*-Yuen son más susceptibles a la violación de los supuestos, cuando las muestras son pequeñas (5, 5), en cambio el t'-Welch revela menos sensibilidad bajo estas condiciones. Los resultados de las otras simulaciones de las tasas de error tipo I para los diferentes tests muestran una tendencia simi-

lar a la obtenida cuando la desigualdad de varianzas asumió la razón 1:2. En ambos casos, para esas situaciones donde el criterio intermedio-liberal de Bradley no fue alcanzado las tasas de error tipo I para los tests considerados tienden a ser liberales, excepto el test t-Student, cuyas tasas de error I fueron muy liberales, mostrando valores tan altos como 4578 y tan bajos iguales a cero. Así, estos resultados revelan lo siguiente: 1) una gran sensibilidad del test t-Student a la violación de los supuestos de normalidad y homocedasticidad y a la desigualdad de los tamaños de muestra; y 2) una gran robustez del test t'-Welch bajo estas condiciones.

Tabla 4. Estimaciones de error tipo I de los tests

Tamaño de Muestra	Asimetría = 1.0; Kurtosis = 2.0			Asimetría = 0.5; Kurtosis = -0.5		
	t-Student	t'-Welch	t*-Yuen	t-Student	t'-Welch	t*-Yuen
n ₁ , n ₂						
5, 5	0.0579	0.0486	0.0507	0.0595	0.0512	0.0553
20, 20	0.0505	0.0488	0.0514	0.0528	0.0505	0.0553
50, 50	0.0497	0.0492	0.0524	0.0530	0.0518	0.0515
5, 20	0.1913	0.0578	0.0855	0.1864	0.0567	0.0886
5, 50	0.2543	0.0542	0.0816	0.2548	0.0539	0.0838
20, 50	0.1322	0.0526	0.0561	0.1364	0.0478	0.0530
20, 5	0.0069	0.0498	0.0612	0.0053	0.0527	0.0561
50, 5	0.0014	0.0554	0.0743	0.0011	0.0555	0.0755
50, 20	0.0128	0.0530	0.0537	0.0106	0.0510	0.0466

Estos resultados comparativamente muestran que los tests examinados (t-Student, t'-Welch y t* Yuen) son más sensibles a la violación de los supuestos cuando la desigualdad de varianza se incrementa de 1:2 a 1:4. Con estos análisis

se pudo constatar que: a) los resultados proporcionados por el test t-Student bajo violación de los supuestos de normalidad y homocedasticidad coinciden con previos resultados de estudios realizados bajo condiciones similares y b)

los resultados de la robustez del test t-Welch bajo las condiciones de ausencia de normalidad y heterocedasticidad también concuerdan con previos estudios con similares condiciones, tales como: (Boneau, 1960; Welch, 1947; Burton, Haug y Hopkins, 1996), entre otros.

4 Conclusiones

Los valores obtenidos de las tasas de error tipo I para los diferentes tests examinados, bajo las condiciones de nor-

malidad y homocedasticidad, revelan que los tests t-Student y t'-Welch proporcionan un mejor control de estas tasas que el test t*-Yuen, es decir, reflejan un valor observado dentro de los límites del intervalo establecido por Bradley.

Cuando se consideran distribuciones normales y desigualdad de varianzas, los resultados de las estimaciones de las tasas de error tipo I muestran que el test t-Student no es robusto, excepto cuando las muestras son grandes e iguales (50, 50).

Tabla 5. Estimaciones de error tipo I de los tests

Tamaño de Muestra	Asimetría = 1.0; Kurtosis = 2.0			Asimetría = 0.5; Kurtosis = -0.5		
	t-Student	t'-Welch	t*-Yuen	t-Student	t'-Welch	t*-Yuen
n ₁ , n ₂						
5, 5	0.0766	0.0594	0.0711	0.0633	0.0539	0.0544
20, 20	0.0569	0.0510	0.0553	0.0789	0.0755	0.0699
50, 50	0.0524	0.0507	0.0549	0.0099	0.0979	0.0901
5, 20	0.3116	0.0545	0.0722	0.2023	0.0577	0.0887
5, 50	0.4578	0.0554	0.0747	0.2639	0.0617	0.0822
20, 50	0.1972	0.0521	0.0543	0.1708	0.7050	0.0688
20, 5	0.0013	0.0498	0.0519	0.0073	0.0579	0.0645
50, 5	0.0000	0.0506	0.0533	0.0026	0.0683	0.0879
50, 20	0.0005	0.0466	0.0493	0.0213	0.0810	0.0748

Igualmente los resultados comparativos de los tests bajo estas condiciones indican que el test de Welch es más robusto que el test de Student y éste, a su vez, más que el test de Yuen.

Las probabilidades de las tasas de error tipo I bajo las condiciones de ausencia de normalidad e igualdad de varianzas revelan robustez en los tests examinados, con la excepción del test de Yuen que demuestra ser sensitivo a estas condiciones y específicamente cuando las combinaciones de los tamaños de muestras son desiguales.

Las estimaciones de las tasas de error tipo I de los tests investigados, bajo las condiciones de ausencia de normalidad y desigualdad de varianzas, muestran una gran sensibilidad del test t-Student a la violación de los supuestos de normalidad y homocedasticidad, excepto cuando se consideran tamaños de muestras grandes e iguales, reflejando que la probabilidad observada del error tipo I (τ) es muy aproximada a la probabilidad nominal del error tipo I (α). Igualmente, estos resultados revelan la robustez del test t'-Welch cuando es comparado con los tests restantes.

En general, los resultados de las estimaciones de las tasas de error tipo I muestran que los tests t-Student, t'-Welch y t*-Yuen son más sensitivos al incremento en la desigualdad de las varianzas que en la condición de ausencia de normalidad.

5 Referencias

- Aspin AA, 1948, An examination and further development of a formula arising in the problem of comparing two mean values, *Biometrika*, 35, pp. 88-96.
- Boneau CA, 1960, The effects of violations of assumptions underlying the t-test, *Psychological Bulletin*, 57, pp. 49-64.
- Boneau CA, 1962, A comparison of the power of the U and t tests, *Psychological Review*, 69, 3, pp. 246-256.
- Borges A, San Luis C y Sánchez-Bruno A, 1993, Contraste de la hipótesis nula para la diferencia de muestras: Alternativa frente al problema de Berhrens-Fisher, Póster presentado en el III Simposium de Metodología de las Ciencias Sociales y del Comportamiento. Santiago de Compostela.
- Borges A, Sánchez-Bruno A y Cañadas I, 1996, El contraste de las diferencias de medias con grupos pequeños, con escalas ordinales y en ausencia de normalidad, *Psicológica*, 17, pp. 455-466.
- Bradley JV, 1978, Robustness? *British Journal of Mathematical and Statistical Psychology*, 31, pp. 144-152.
- Brown M & Forsythe A, 1974, The small sample behavior of some statistics which test the equality of several means, *Technometrics*, 16, pp. 129-132.
- Burton E, Haug C & Hopkins KD, 1996, Revisiting the Homogeneity of Variance Assumptions: Is the Conven-

- tional t test of Means Ever the Method of Choice? Trabajo presentado en la reunión anual de American Educational Research Association. New York.
- Cochran WG, 1964, Approximate significance levels of the Behrens-Fisher problem, *Biometrics*, 20, pp. 191-195.
- Fleishman AI, 1978, A method for simulating non-normal distributions. *Psychometrika*, 43, 4, pp. 521-532.
- Glass GV, & Hopkins K.D, 1996, *Modern elementary statistics*, Eighth edition, Englewood Cliffs, NJ: Prentice Hall.
- Robey RR & Barcikowshi RS, 1992, Type I error and the iterations in Monte Carlo studies of robustness, *British Journal of Mathematical and Statistical Psychology*, 45, pp. 283-288.
- Sánchez-Bruno A y Borges A, 1997, Violación del supuestos de normalidad en contrastes estadísticos para grupos pequeños, Comunicación presentada en el V Simposium de Metodología de las Ciencias del Comportamiento. Sevilla-España.
- Sánchez-Bruno A, Borges A y Cañadas I, 2000, El contraste de medias recortadas ante la violación de los supuestos paramétricos *Psicothema*, 12, pp. 506-508.
- Scheffé H, 1959, *The analysis of variance*, New York: Wiley.
- Scheffé H, 1970, Practical solutions to the Behrens-Fisher problem, *Journal of the American Statistical Association*, 65, pp. 1051-1058.
- Welch BL, 1947, The generalization of Students problem when several different population variances are involved, *Biometrika*, 34, pp. 23-35.
- Yuen KK, 1974, The two-sample Trimmed t for unequal population variances, *Biometrika*, 61, 1, pp. 165-170.

