İstatistik Araştırma Dergisi Cilt 02 No 02 Özel Sayı Sayfa 65-73 Ağustos 2003

Bağımsız Çift Örneklem için Yeni bir Homojenlik Testi

Adil KORKMAZ*

ÖZET

Bu çalışmada biri ucuz olduğu için bol ölçüde, öteki de pahalı olduğu için kıt ölçüde elde edilebilen iki örnekleme dayalı olarak bağımsız çift örneklem homojenlik testini gerçekleştirmeye yönelik yeni bir test yöntemi geliştirilmiştir.

Anahtar Kelimeler: *İki Örneklem Konum Sorunu, Mann-Whitney-Wilcoxon Test, Linear Placement.*

1. GİRİŞ

X, dağılım işlevi F ve olasılık yoğunluk işlevi f olan; Y ise dağılım işlevi G ve olasılık yoğunluk işlevi g olan sürekli raslantı değişkenleri olsun. Bu raslantı değişkenlerine ilişkin olmak üzere çekilen m ve n büyüklüklerindeki örneklem raslantı değişkenleri de $X_1, X_2, ..., X_m$ ve $Y_1, Y_2, ..., Y_n$; bunların sıralı örneklem değerleri de $X_{(1)}, X_{(2)}, ..., X_{(m)}$ ve $Y_{(1)}, Y_{(2)}, ..., Y_{(n)}$ olarak gösterilirken, sırasız örneklem gözlemleri $x_1, x_2, ..., x_m$ ve $y_1, y_2, ..., y_n$, sıralı örneklem gözlemleri de $x_{(1)}, x_{(2)}, ..., x_{(m)}$ ve $y_{(1)}, y_{(2)}, ..., y_{(n)}$ ile gösterilmektedir. Salt gözlemlere dayalı olarak X ve Y raslantı değişkenlerinin dağılım işlevlerinin eşit olup olmadığını anlamaya yönelik hipotez testi, istatistik yazınında bağımsız çift örneklem homojenlik testi diye adlandırılmaktadır. Burada yokluk hipotezi **H: F=G** ve almaşık hipotez **A: F≠G** biçiminde dile getirilebilir. İstatistik yazınında bağımsız çift örneklem homojenlik testleri pek çok kaynakta bulunabilir (Korkmaz-Günay, 2000; Sprent, 1989; Siegel ve Castellan, 1988; Gibbons, 1985; Dixon ve Massey, 1983; Hollander ve Wolfe, 1973; Mann ve Whitney, 1947; Wilcoxon, 1945; Wald ve Wolfowitz, 1940). Bu çalışmada önerilen bağımsız çift örneklem homojenlik testi, biri pahalı olduğundan dolayı ya da doğal nedenlerle küçük, öteki ise ucuz olduğundan dolayı ya da doğal nedenlerle büyük ölçülerde elde edilebilen örneklemlere dayalı olmak bakımından öncellerinden farklılaşmaktadır. Söz konusu çift örneklem homojenlik testinin iki temel dayanağı vardır:

Biri, **asimtotik olarak eşit olasılıklı aralıklar teoremi**, öteki de **tüme karşı bir karşılaştırma yöntemi**dir.

* Akdeniz Üniversitesi Ekonometri Bölümü, korkmaz@iibf.edu.tr

2. ASİMTOTİK OLARAK EŞİT OLASILIKLI ARALIKLAR TEOREMİ

Y raslantı değişkenine ilişkin $Y_{(1)}, Y_{(2)}, ..., Y_{(n)}$ sıralı istatistikleri, gerçel ekseni, n+1 aralığa aşağıdaki gibi böler:

$I_1:(-\infty, Y_{(1)}]$,
$I_2:(Y_{(1)}, Y_{(2)}]$,
...,
$I_n:(Y_{(n-1)}, Y_{(n)}]$,
$I_{n+1}:(Y_{(n+1)}, +\infty)$.

Y raslantı değişkeninin bu aralıklarda gerçekleşme olasılıkları k=1, 2, ..., n+1 olmak üzere şöyle tanımlanabilir:

$P_k=P(Y\in I_k)$.

$P_1, P_2, ..., P_n$ ($P_{n+1}=1-P_1-P_2-...-P_n$) olasılıkları birer raslantı değişkenidir. Bu raslantı değişkenlerinin Dirichlet(1, 1, ..., 1; 1) dağılımlı oldukları kolayca gösterilebilir. Bunun için Y örneklemine ilişkin olasılık öğesini (dP),

$dP=n!\cdot dG(Y_{(1)})\cdot dG(Y_{(2)}) \cdots dG(Y_{(n)})$,
biçiminde yazmak ve sonra da,

$P_1=P(Y\in I_1)=P(Y\leq Y_{(1)})=G(Y_{(1)})$;

$P_{k+1}=P(Y\in I_{k+1})=P(Y\leq Y_{(k+1)})-P(Y\leq Y_{(k)})=G(Y_{(k+1)})-G(Y_{(k)})$, (k=1, 2, ..., n-1);

$P_{n+1}=P(Y\in I_{n+1})=P(Y>Y_{(n)})=1-P(Y\leq Y_{(n)})=1-G(Y_{(n)})$

dönüşümlerini yapmak gerekmektedir. Bu durumda,

$dP=n!\cdot dP_1\cdot dP_2\cdots dP_n$,

elde edilir ki, bu da söz konusu savın geçerliliğini kanıtlar.

Buradan k. sıralı istatistik olan $Y_{(k)}$'nin marjinal olasılık yoğunluk işlevi şöyle bulunabilir (**Vaart, 1998; Casella and Berger, 1990; Wilks, 1962**) (k=1, 2, ...,m):

$$\frac{n!}{(k-1)!\cdot(n-k)!}\cdot\left[G(y_{(k)})\right]^{k-1}\cdot\left[1-G(y_{(k)})\right]^{n-k}\cdot g(y_{(k)}).$$

$W_k=G(Y_{(k)})$ ve $w_k = G(y_{(k)})$ değişken dönüşümü yapılır ve $\left|\frac{dy_{(k)}}{dw_k}\right| = \frac{1}{g(y_{(k)})}$ olduğu göz önüne alınırsa W_k'nin olasılık yoğunluk işlevi:

$$\frac{n!}{(k-1)!\cdot(n-k)!}\cdot w_k^{\,k-1}\cdot\left(1-w_k\right)^{n-k}.$$

biçiminde bulunabilir. Bu demektir ki, W_k, k ve n+1-k parametrelerine sahip beta dağılımlı bir raslantı değişkenidir. Bunun beklenen değeri ve varyansı şöyledir:

$$E(W_k)=\frac{k}{n+1},\ \mathrm{Var}(W_k)=\frac{k\cdot(n-k+1)}{(n+1)^2\cdot(n+2)}.$$

Bunlar Chebyshev eşitsizliğinde yerine konulursa,

$$P\left\{\left|W_k-\frac{k}{n+1}\right|<\varepsilon\right\}\geq 1-\frac{k\cdot(n-k+1)}{(n+1)^2\cdot(n+2)\cdot\varepsilon^2}$$

bulunur ki, burada ε istenildiği ölçüde küçük tutulabilen artı değerli bir sayıdır. n sonsuza götürüldüğünde,

$$\lim_{n\to\infty} P\left\{\left|W_k-\frac{k}{n+1}\right|<\varepsilon\right\}=1$$

elde edilir. Yukarıdaki bağıntı, W_k raslantı değişkeninin büyük sayılar yasasına boyun eğdiğini ve $\frac{k}{n+1}$ değerine yakınsadığını gösterir. Öte yandan,

$P_1=W_1$;

$P_{k+1}=W_{k+1}-W_k$, (k=1, 2, ..., n-1);

$P_{n+1}=1-W_n$

olduğu göz önünde bulundurulursa, W_1, W_2, ..., W_n raslantı değişkenleri $\frac{1}{n+1}$, $\frac{2}{n+1}$, ..., $\frac{n}{n+1}$ değerlerine yakınsarken, P_1, P_2, ..., P_n ve P_{n+1} raslantı değişkenlerinin de $\frac{1}{n+1}$ değerine yakınsadığı kolayca anlaşılır. Bu, asimtotik olarak eşit olasılıklı aralıklar teoreminin matematiksel anlatımıdır.

3. KARŞILAŞTIRMA YÖNTEMİ

Önerilen bağımsız çift örneklem homojenlik testinin ikinci dayanağı, tüme karşı bir karşılaştırma yöntemidir. Daha önce pek çok testte verilerin içerdiği bilgiyi tam olarak değerlendirmek kaygısıyla **tüme karşı tüm karşılaştırma yöntemi** uygulanmıştır. Wilcoxon sıra toplamı testinde uygulanan bu geleneksel yöntemin algoritması şöyle verilebilir: Birinci örneklemin bütün öğeleri ikinci örneklemin bütün öğeleriyle birleştirilir; birleşik örneklem öğeleri küçükten büyüğe doğru sıralanır; birinci örneklem öğelerinin birleşik sıralı örneklemdeki rankları bulunur. Birinci örneklem (20, 15) ve ikinci örneklem (10, 37, 17) olsun. Birleşik örneklem (20, 15, 10, 37, 17); birleşik sıralı örneklem (10, 15, 17, 20, 37) olarak elde edilir ve buradan da ilk örnekleme ilişkin ranklar **4** ve **2** olarak bulunur.

$X_{(k)}$ raslantı değişkeninin tüme karşı tüm karşılaştırma yöntemiyle elde edilen rankı T_k ile gösterilecek olursa, Orban ve Wolfe (**1979**), T_k-k değerini *placement* diye adlandırmaktadır. Bu, birleşik sıralı örneklemde $X_{(k)}$ raslantı değişkeninden küçük Y gözlemlerinin sayısını anlatır. Fligner ve Policello (**1981**), bu değerlere dayanarak Wilcoxon sıra toplamı testini daha etkin bir duruma getirmişlerdir. Burada ise, çok benzemekle birlikte, $T_k - k$ değeri değil, $T_k - k + 1$ değeri temel olarak alınmaktadır. Bu değer, "tüme karşı bir karşılaştırma yöntemi" ile elde edilen ranklar olmaktadır. Bu karşılaştırma yönteminin temel adımları şöyle dile getirilebilir: Küçük örneklemin yalnızca bir öğesi büyük örneklemin bütün öğeleriyle birleştirilir; birleşik örneklem küçükten büyüğe doğru sıralanır; küçük örneklem öğesinin rankı bulunur ve bu işlemler küçük örneklemin her öğesi için yinelenir. Küçük örneklem (20, 15) ve büyük örneklem (10, 37, 17) olsun. İlk örneklemin yalnızca bir öğesiyle ikinci örneklemin tüm öğeleri (20, 10, 37, 17) biçiminde olmak üzere birleştirilerek yeni bir örneklem elde edilir. Bu örneklem değerleri sıralanarak (10, 17, 20, 37) dizisi elde edilir. İlk örneklemin öğesine ilişkin rank **3** olarak bulunur. Bu adımlar ilk örneklemin geri kalan öğeleri için de yinelenir. Böylece ilk örneklemin izleyen öğesinin rankı **2** olarak bulunur.

4. TEST İSTATİSTİĞİ VE DAĞILIMI

$X_1, X_2, \ldots, X_m$ örneklem raslantı değişkenleri için tüme karşı bir karşılaştırma yöntemiyle elde edilmiş olan ranklar $R_1, R_2, \ldots, R_m$ olsun. Homojenlik koşulu altında (H koşulu altında) k=1, 2, ..., m olmak üzere R_k raslantı değişkeni 1 ile n+1 arasındaki değerleri $P_1, P_2, \ldots, P_{n+1}$ olasılıklarıyla alacaktır. Bu durumda $S_k = -2 \cdot \ln\left(\frac{R_k}{n+1}\right)$ olarak tanımlanacak olan raslantı değişkeninin $P_1, P_2, \ldots, P_n$ koşulu altındaki karakteristik işlevi şöyle olacaktır

$(i=\sqrt{-1})$:

$$\Psi_{S_k/P_1,P_2,\ldots,P_n}(t) =$$

$$e^{-i\cdot t\cdot 2\cdot \ln\left(\frac{1}{n+1}\right)} \cdot P_1 + e^{-i\cdot t\cdot 2\cdot \ln\left(\frac{2}{n+1}\right)} \cdot P_2 + \ldots + e^{-i\cdot t\cdot 2\cdot \ln\left(\frac{n+1}{n+1}\right)} \cdot P_{n+1}$$

Bu da şöyle yazılabilir:

$$\Psi_{S_k/P_1,P_2,\ldots,P_n}(t) =$$

$$\left(\frac{1}{n+1}\right)^{-i\cdot t\cdot 2} \cdot P_1 + \left(\frac{2}{n+1}\right)^{-i\cdot t\cdot 2} \cdot P_2 + \ldots + \left(\frac{n+1}{n+1}\right)^{-i\cdot t\cdot 2} \cdot P_{n+1}$$

Koşullu karakteristik işlevini $n \to \infty$ iken bulabilmek için k=1, 2, ..., n+1 olmak üzere $z_k = \frac{k}{n+1}$ tanımlanacak olursa $\Delta z_k = \frac{1}{n+1}$ elde edilir. Asimtotik olarak eşit olasılıklı aralıklar teoremi gereğince $P_1, P_2, ..., P_n$ ve P_{n+1} olasılıkları $\Delta z_k = \frac{1}{n+1}$ değerine yakınsadığından, koşullu karakteristik işlev,

$$\lim_{n \to \infty} \Psi_{S_k/P_1,P_2,...,P_n}(t) = \lim \sum_{k=1}^{n+1} z_k^{-i \cdot t \cdot 2} \cdot \Delta z_k$$

biçiminde yazılabilir. Buradan da **integral** tanımı dikkate alınarak,

$$\lim_{n \to \infty} \Psi_{S_k/P_1,P_2,...,P_n}(t) = \int_0^1 z^{-i \cdot t \cdot 2} \cdot dz = (1 - i \cdot t \cdot 2)^{-1}$$

elde edilir. Görüldüğü gibi, koşullu karakteristik işlev $n \to \infty$ için $P_1, P_2, ..., P_n$ ve P_{n+1} değerlerine bağlı değildir ve dolayısıyla da S_k raslantı değişkeninin koşulsuz (marjinal) karakteristik işlevi,

$$\lim_{n \to \infty} \Psi_{S_k}(t) = (1 - i \cdot t \cdot 2)^{-1}$$

olarak bulunur. Bu, serbestlik derecesi 2 olan ki-kare dağılımlı bir raslantı değişkeninin karakteristik işlevidir.
Test istatistiği,

$$S = S_1 + S_2 + ... + S_m = -2 \cdot \ln\left(\frac{R_1 \cdot R_2 \cdots R_m}{(n+1)^m}\right)$$

biçiminde tanımlansın. S'nin $n \to \infty$ iken $2 \cdot m$ serbestlik derecesinde ki-kare dağılımlı bir raslantı değişkeni olduğu kolayca gösterilebilir: $P_1, P_2, ..., P_n$ (açıktır ki $P_{n+1} = 1 - P_1 - P_2 - ... - P_n$) koşulu altında, $X_{(1)}, X_{(2)}, ..., X_{(m)}$ raslantı değişkenleri, $I_1, I_2, ..., I_n$ ve I_{n+1} aralıklarında birbirlerinden bağımsız olarak $P_1, P_2, ..., P_n$ ve P_{n+1} olasılıklarıyla gerçekleşirler. Dolayısıyla $R_1, R_2, ..., R_m$ ve onların birer işlevi olan $S_1, S_2, ..., S_m$ raslantı değişkenleri $P_1, P_2, ..., P_n$ koşulu altında birbirlerinden bağımsızdırlar. Dolayısıyla,

$$\Psi_{S/P_1,P_2,...,P_n}(t) = \left\{\Psi_{S_k/P_1,P_2,...,P_n}(t)\right\}^m$$

yazılabilir. Buradan

$$\lim_{n \to \infty} \Psi_{S/P_1,P_2,...,P_n}(t) = (1 - i \cdot t \cdot 2)^{-m}$$

bulunur. S'nin koşullu karakteristik işlevi, n sonsuza giderken $P_1, P_2, ..., P_n$ raslantı değişkenlerinden bağımsız olduğundan, söz konusu raslantı değişkeninin koşulsuz (marjinal) karakteristik işlevi,

$$\lim_{n \to \infty} \Psi_S(t) = (1 - i \cdot t \cdot 2)^{-m}$$

biçiminde yazılabilir. Bu, bilindiği gibi, serbestlik derecesi 2·m olan ki-kare dağılımlı bir raslantı değişkeninin karakteristik işlevidir.

5. TESTİN TUTARLILIĞI

S istatistiğine dayalı olarak yapılacak bağımsız çift örneklem homojenlik testinin yokluk hipotezi ve almaşık hipotezi H: P(X>Y)=P(Y>X) ve A: P(X>Y)>P(Y>X) biçiminde dile getirilecek olursa, testin tutarlı olduğu gösterilebilir. $S_k = -2 \cdot \ln\left(\frac{R_k}{n+1}\right)$ raslantı değişkeninin beklenen değeri A koşulu altında μ_A ve H koşulu altında μ_H; varyansı ise A koşulu altında σ_A^2 ve H koşulu altında σ_H^2 olsun. Burada $\mu_H=E(S/H)=2 \cdot m$ ve $\sigma_H^2=Var(S/H)=4 \cdot m$ olduğu bilinmektedir. $0 < \frac{R_k}{n+1} \leq 1$ olduğundan $S_k = -2 \cdot \ln\left(\frac{R_k}{n+1}\right)$ şöyle yazılabilir:

$$S_k = 2 \cdot \left\{\left(1-\frac{R_k}{n+1}\right)+\frac{1}{2}\cdot\left(1-\frac{R_k}{n+1}\right)^2+\frac{1}{3}\cdot\left(1-\frac{R_k}{n+1}\right)^3+\ldots\right\}.$$

Öte yandan u=1, 2, ..., m ve v=1, 2, ..., n olmak üzere $\Pi_{uv} = \begin{cases} 1 & X_u > Y_v \\ 0 & X_u \leq Y_v \end{cases}$ tanımlanacak olursa $1-\Pi_{uv} = \begin{cases} 1 & X_u < Y_v \\ 0 & X_u \geq Y_v \end{cases}$ yazılabilir. H koşulu altında bu raslantı değişkenlerine bağlı olarak aşağıdaki bağıntı yazılabilir:

$$n+1-R_k = (1-\Pi_{k1})+(1-\Pi_{k2})+\ldots+(1-\Pi_{kn}).$$

A koşulu altında $P\{1-\Pi_{uv}=1\} < \frac{1}{2}$ olur. n+1-R_k değişkeninin r. momenti şöyledir:

$$E\{(n+1-R_k)^r\} = E\{[(1-\Pi_{k1})+(1-\Pi_{k2})+\ldots+(1-\Pi_{kn})]^r\}.$$

Açık biçimde yazılacak olursa bu moment şöyle olur:

$$E\{(n+1-R_k)^r\} = \sum \frac{r!}{o_{k1}! \cdot o_{k2}! \cdots o_{kn}!} \cdot (1-\Pi_{k1})^{o_{k1}} \cdot (1-\Pi_{k2})^{o_{k2}} \cdots (1-\Pi_{kn})^{o_{kn}}$$
$$\cdot P(1-\Pi_{k1}=1) \cdot P(1-\Pi_{k2}=1) \cdots P(1-\Pi_{kn}=1)$$

Burada o_{k1}, o_{k2}, ..., o_{kn} değişkenleri, toplamları r olmak üzere, 0 ile r arasında değişir. H koşulu altında,

$$P(1-\Pi_{k1}=1)=P(1-\Pi_{k2}=1)=\cdots=P(1-\Pi_{kn}=1)=\frac{1}{2},$$

A koşulu altında ise,

$$P(1-\Pi_{k1}=1)=P(1-\Pi_{k2}=1)=\cdots=P(1-\Pi_{kn}=1)<\frac{1}{2}$$

olduğuna göre $E\left\{(n+1-R_k)^r\right\}$ değeri, biri A ve biri de H koşulu altında olmak üzere iki kez yazılabilir. Bu anlatımlar terim terime karşılaştırılacak olursa,

$$E\left\{(n+1-R_k)^r/A\right\}<E\left\{(n+1-R_k)^r/H\right\}$$

elde edilir. Buradan,

$E(S_k/A)<E(S_k/H)$ ve dolayısıyla da $\mu_A \leq \mu_H$ bulunur. Öte yandan,

$E(S/A)=m\cdot\mu_A<E(S/H)=2\cdot m$,

$Var(S/A)=m\cdot\sigma_A^2$.

H koşulu altında ve $n\rightarrow\infty$ için $2\cdot m$ serbestlik derecesinde ki-kare dağılımlı olan S raslantı değişkeni, büyük m değerleri için $2\cdot m$ ve $4\cdot m$ parametreleriyle normal dağılımlı olacaktır. Dolayısıyla homojenlik testinde yokluk hipotezinin reddedilmesi, $\frac{S-2\cdot m}{\sqrt{4\cdot m}}>z_{\alpha/2}$ ya da $\frac{S-2\cdot m}{\sqrt{4\cdot m}}<-z_{\alpha/2}$ olaylarının gerçekleşmesine bağlıdır. Burada

$z_{\alpha/2}$, normal dağılım tablosunda $1-\alpha$ güven aralığına göre belirlenen kritik değerdir. X, Y'den olasılıksal olarak daha büyük ise bu durumda red kararı verilebilmesi için $\frac{S-2\cdot m}{\sqrt{4\cdot m}}<z_{\alpha/2}$ olayının gerçekleşmesi gerekecektir. Çünkü X, Y'den olasılıksal olarak daha büyük iken R_1, R_2, ..., R_m raslantı değişkenleri n+1'e yaklaşırken S_1, S_2, ..., S_m raslantı değişkenleri de 0'a yaklaşır ve böylece S raslantı değişkeni aşırı ölçüde küçük bir değer olarak H'ın aleyhinde bir kanıt ortaya koyar. Bu durumun gerçekleştiğini görebilmek amacıyla Chebyshev eşitsizliği A koşulu altında şöyle yazılabilir:

$$P\left\{m\cdot\mu_A-c\cdot\sqrt{m}\cdot\sigma_A<S<m\cdot\mu_A+c\cdot\sqrt{m}\cdot\sigma_A\right\}\geq 1-\frac{1}{c^2}.$$

Buradan

$$P\left\{\frac{S-2\cdot m}{\sqrt{4\cdot m}} < \frac{m\cdot\mu_A + c\cdot\sqrt{m}\cdot\sigma_A - 2\cdot m}{\sqrt{4\cdot m}}\right\} \geq 1-\frac{1}{c^2}$$

yazılabilir. $z_{\alpha/2} = \dfrac{2\cdot m - m\cdot\mu_A - c\cdot\sqrt{m}\cdot\sigma_A}{\sqrt{4\cdot m}}$ olduğunda $\dfrac{S-2\cdot m}{\sqrt{4\cdot m}} < -z_{\alpha/2}$ da olur ve H

reddedilir. Buradan $c = \dfrac{2-\mu_A}{\sigma_A}\cdot\sqrt{m} - \dfrac{2}{\sigma_A}\cdot z_{\alpha/2}$ elde edilir. $m\to\infty$ için c sonsuza gider

ve $\lim\limits_{m\to\infty} P\left\{\dfrac{S-2\cdot m}{\sqrt{4\cdot m}} < \dfrac{m\cdot\mu_A + c\cdot\sqrt{m}\cdot\sigma_A - 2\cdot m}{\sqrt{4\cdot m}}\right\} = 1$ ve $\lim\limits_{m\to\infty} P\left\{\dfrac{S-2\cdot m}{\sqrt{4\cdot m}} < -z_{\alpha/2}\right\} = 1$ olur. Bu

demektir ki, A koşulu altında n'nin yanı sıra m de sonsuza giderken H'ın reddedilme olasılığı 1 olur. Dolayısıyla S istatistiğine dayalı olarak gerçekleştirilecek test asimtotik olarak tutarlıdır.

6. SONUÇ

S istatistiği çift örneklem homojenlik testi için uygun bir ölçüttür. Çünkü homojenlik koşulu altında R_1, R_2, ..., R_m istatistiklerinden türetilen S istatistiği 2·m serbestlik derecesinde ki-kare dağılımı gösterecek ve dolayısıyla da 2·m yöresinde gerçekleşecektir. Bu durumda da H hipotezi kabul edilecektir. X raslantı değişkeni Y raslantı değişkeninden olasılıksal olarak daha küçük ise bu durumda R_1, R_2, ..., R_m istatistikleri 1'e yaklaşırken S istatistiği de 0'a yaklaşacak ve bu da yokluk hipotezinin reddedilerek karşıt hipotezin kabul edilmesiyle sonuçlanacaktır. Dolayısıyla S istatistiği çift örneklem homojenlik testi için uygun bir ölçüt olma işlevini görür. Böylece elde edilebilen S istatistiğine dayalı olarak yapılacak çift örneklem testi benzerlerinden farklılaşmaktadır. Bu farklılık, pahalı olduğu için ancak kıt ölçüde elde edilebilen bir örneklem ile ucuz olduğu için bol ölçüde elde edilebilen bir başka örneklemi birlikte ele alıp homojenlik durumunu saptamaya elverişli olmak biçiminde dile getirilebilir.

KAYNAKLAR

CASELLA, G.-BERGER, R. L. (1990). *Statistical Inference, Duxbury Press*, Belmont, p.232.

DİXON, W. J.-MASSEY, F. J. (1983). *Introduction to Statistical Analysis*, McGraw-Hill, New York.

FLİGNER, M. A.-POLİCELLO, G. E. (1981). Robust Rank Procedures for the Behrens-Fisher Problem, *Journal of the American Statistical Association,* Vol. 76, s. 162-168.

GİBBONS, J. D. (1985). *Nonparametric Methods for Quantitative Analysis*, American Sciences Press, Ohio, p. 23.

HOLLANDER, M.-WOLFE. A. D. (1973). *Nonparametric Statistical Methods*, John Wiley and Sons, New York.

KORKMAZ, A.-GÜNAY S. (2000). *Homojenlik Testinin Uyum İyiliği Testine Dönüştürülmesi*, İstatistik Araştırma Sempozyumu 2000 Bildiriler Kitabı, D.İ.E., 2000.

MANN, H. B.-WHİTNEY, D. R. (1947). *On a Test of Whether One of Two Random Variables is Stochastically Larger than the Other*, ***The Annals of Mathematical Statistics***, Vol. XVIII, 1947, pp. 50-60.

ORBAN, J.-WOLFE, D. A. (1979). *A Class of Distribution-Free Two-Sample Tests Based on Placements, Ohio State University Technical Report* **No. 178**, Dept. oF Statistics.

SİEGEL, S.-CASTELLAN, N. J. (1988). *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill, New York.

SPRENT, P. (1989). *Applied Nonparametric Statistical Methods,* Chapman and Hall, London.

VAART, A. W. van der (1998). *Asymptotic Statistics*, Cambridge University Press, Cambridge, p. 174.

WALD, A.-WOLFOWİTZ, J. (1940). *On a Test Whether Two Samples are from the Same Population, The Annals of Mathematical Statistics*, Vol. II, pp. 147-162.

WİLCOXON, F. (1945). *Individual Comparisons by Ranking Methods, Biometrics Bulletin*, Vol. I, pp. 80-83.

WİLKS, S. S. (1962). *Mathematical Statistics*, John Wiley & Sons, New York, p. 232, pp. 234-236, p. 459.

A New Homogeneity Test for Independent Two Samples

ABSTRACT

In this study, a new homogeneity test method has been developed for two independent samples so that one of them is expensive and small while the other is cheap and large.

***Key Words**: Two-sample Location Problem, Mann-Whitney-Wilcoxon Test, Linear Placement.*