ISBN No: 978-975-19-4105-3 Volume: 05 Number: 01 Page: 25-34 ISSN No: 1303-6319 Cilt: 05 Sayı: 01 Sayfa: 25-34

BEHRENS-FISHER PROBLEMİ VE ÇÖZÜM YÖNTEMLERİ

Süleyman GÜNAY* **Semra TÜRKAN****

ÖZET

$N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$ *şeklinde normal dağılan iki kitlenin ortalamaları arasındaki farklılık araştırılırken, bu iki kitlenin varyansları bilinmiyorsa ve eşit değilse söz konusu iki kitlenin ortalamalarının testi, Behrens-Fisher problemi olarak bilinir. Behrens-Fisher probleminin çözümünde çok sayıda yöntemler geliştirilmiştir. Welch-t testi, permütasyon testi v.b. bu testlerden bazılarıdır. Bu çalışmada Behrens-Fisher probleminin çözümünde kullanılan Welch-t testi, permütasyon testi ve sayısal yöntemler tartışılmıştır. Monte Carlo sonuçlarına değinilerek Welch-t testi ve permütasyon testi karşılaştırılmıştır. Son olarak sayısal yöntemler için Cressi ve Whitford tarafından önerilen istatistikleri üreten bir bilgisayar yazılımı uygulanmıştır.*

Anahtar Kelimeler: Behrens-Fisher Problemi, Permütasyon Testi, Welch-Testi.

1. GİRİŞ

İki kitlenin dağılımı konum, ölçek ve şekil bakımından kitlelerin ortalamaları arasındaki farklılığın testini etkileyebilir. Bu nedenle iki kitlenin ortalamaları arasındaki farklılık araştırılırken, öncelikli olarak kitlelerin dağılımları arasında fark olup olmadığı test edilir. Bu testin sonucunda dağılımlar arasında fark olmadığı ve kitlelerin $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$ şeklinde normal dağıldığı varsayılsın. Bu durumda ele alınan kitlelerin varyansları bilinmiyorsa ve eşit değilse $(\sigma_1^2 \neq \sigma_2^2)$ sözkonusu iki kitlenin ortalamalarının testi, Behrens-Fisher problemi olarak bilinir. Hata değişkenleri farklı ve bilinmeyen varyanslara sahip olduğunda iki ortalamanın eşitliğini test eden genelleştirilmiş Behrens-Fisher problemi bilinen bir örnektir ve bu ortalamalara ilişkin hipotezler Behrens-Fisher probleminin çözümüne ilişkin yöntemler yardımıyla test edilebilir. Parametrik olmayan iki kitlenin ortalamaları test edilirken ise permütasyon test istatistiği kullanılabilir.

Behrens-Fisher probleminin çözümünde çok sayıda yöntem geliştirilmiştir. Bunlar içinde en çok kullanılanları; Fisher (1936), Welch (1947), Wald-Romanovskaja (Wald, 1955; Romanovskaja, 1965) tarafından geliştirilen yöntemlerdir. Ancak Wald ve Romanovskaja tarafından geliştirilen yöntemler sadece eşit örneklem büyüklükleri için geçerlidir. Bu yöntemler içinde Welch'in geliştirdiği Welch-t testinin asimtotik olarak en güçlüsü olduğu tanıtlanmıştır (Pfanzagl, 1974).

* Prof. Dr. Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, e-mail: sgunay@hacettepe.edu.tr
** Arş. Gör. Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, e-mail: sturkan@hacettepe.edu.tr

2. WELCH-T TESTİ

Ortalamaları, varyansları bilinmeyen ve varyansları eşit olmayan $N(\mu_1, \sigma_1^2)$ ve $N(\mu_2, \sigma_2^2)$ normal dağılımlardan sırasıyla m ve n büyüklükte iki bağımsız örneklem $\{X_1, \ldots, X_m\}$ ve $\{Y_1, \ldots, Y_n\}$ olsun. Bu örneklemlerin ortalamaları sırasıyla $\bar{X}$ ve $\bar{Y}$ ile gösterilsin. Buna göre,

$H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 < \mu_2$ ya da $\mu_1 > \mu_2$ ya da $\mu_1 \neq \mu_2$

hipotez testlerinde en sık kullanılan yaklaşım Welch-t testidir. Welch-t test istatistiği,

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{S_x^2 / m + S_y^2 / n}} \sim t_\nu \tag{1}$$

şeklinde yazılır. Burada ν serbestlik derecesi ,

$$\nu = \frac{\left(S_1^2 / m + S_2^2 / n\right)^2}{\dfrac{\left(S_1^2 / m\right)^2}{m-1} + \dfrac{\left(S_2^2 / n\right)^2}{n-1}} \tag{2}$$

dir. S_1^2, S_2^2 sırasıyla birinci ve ikinci örneklem varyanslarıdır. Buna göre r=m+n ise Welch-t test fonksiyonu,

$$\varphi_{r,\text{Welch}} = \begin{cases} 1, & T > t_t \\ 0, & T \leq t_t \end{cases} \tag{3}$$

şeklinde yazılabilir. Bu test fonksiyonunda T, Eşitlik (1)'de verilmiştir. Bu testi açıklayan aşağıdaki örnek 1 ele alınsın (Lehmann, 1959; Mehta ve Srinivasan, 1970).

Örnek 1

Yeni Zelanda'nın değişik bölgelerinde yağmur suyunun içerdiği sülfür miktarına karar verebilmek için, alınan yağmur suyu örnekleri laboratuvar ortamında analiz edilmektedir. Araştırma kapsamındaki bölgeler doğu ve batı olmak üzere iki gruba ayrılmaktadır. Buna göre yağmur suyundaki ortalama sülfür yoğunluğu Yeni Zelanda'nın batısında ve doğusunda farklılık gösterir mi (Jay ve Roxy, 1993)?

Batı: 1.15, 1.20, 0.43, 0.46, 0.44, 0.25, 0,43, 0,43, 0,25, 0,43, 0,83, 0.11, 0.60, 0.43, 0.23, 0.30, 0.22, 0.08, 0.07, 0.28

Doğu: 0.26, 0.13, 0.62, 0.40, 0.28, 0.23, 0.80, 0.32, 0.08, 0.09, 0.19, 0.21, 0.58, 0.17, 0.61

Eşitlik (1) ve Eşitlik (2)'den,

$$T=\frac{0,431-0,3313}{\sqrt{0,097/20+0,0492/15}}=1,1078,\ \nu=33 \text{ elde edilir.}$$

T=1,1078 < $T_{(0,025,\ 33)} \cong 2,04$ olduğu için, yağmur suyundaki sülfür yoğunluğunun Yeni Zelanda'nın batısında ve doğusunda farklılık göstermediği $\alpha=0,05$ ile söylenebilir.

Welch-t testi, veriler normal dağılıma sahip olduğunda ve kitle varyansları eşit olmadığında kullanılır. Çarpık veriler için ise permütasyon testi kullanılabilir (Lehmann, 1975).

3. PERMÜTASYON TESTİ VE GENELLEŞTİRİLMİŞ BEHRENS-FISHER PROBLEMİ

İki kitleden sırasıyla m ve n büyüklükte iki bağımsız örneklem $X_1,...,X_m$ ve $Y_1,...,Y_n$ verilsin. Bu örneklemlerin ortalamaları sırasıyla $\overline{X}$ ve $\overline{Y}$ ile gösterilsin. Örneklemlerin alındığı kitlelere ait hipotez testi,

H_0: $F_1(X)=F_2(Y)$

şeklinde yazılır. Bu hipoteze ilişkin uygun test istatistiği,

$$T=\overline{X}-\overline{Y} \tag{4}$$

dır. H_0 hipotezi altında m+n tane gözlemden m tane örneklem seçilirken $\binom{m+n}{m}$ sayıda farklı seçim yapılır. Her seçilen m örneklem için yeni bir T değeri bulunur. Bunun sonucunda T'nin dağılımı,

$$P(T=t)=\frac{\#(t,m,n)}{\binom{m+n}{m}}$$

olarak elde edilir. $\#(t,m,n)$, $\binom{m+n}{m}$ sayıdaki tüm alt kümelerin içinde T=t olan alt küme sayısını gösterir. Eğer $T \geq t_0$ ise P değeri,

$$P(T\geq t_0)=\frac{\#(t\geq t_0,m,n)}{\binom{m+n}{m}}$$ olur. Burada t_0 örneklemden elde edilen değerdir.

Permütasyon testi parametrik olmayan sıra testlerine çok benzer. Ancak permütasyon testinde sıra testlerinde olduğu gibi, gözlemlere sıra sayısı verilmez (Good, 1994). Permütasyon testi aşağıda verilen örnek 2 ile açıklanabilir.

Örnek 2

Her bir grupta 3 öğrencinin bulunduğu bir sınıfta belli bir ders için yapılan sınavdan öğrencilerin aldığı puanlar aşağıda verilmiştir.

X: 76 80 85
→ Eşitlik (4)'ten, T=48,67 olur.
Y: 17 30 48

6 tane ölçüm değerinden, 3 tane ölçüm değeri $\binom{6}{3} = 20$ farklı şekilde seçilebilir. Bu ölçümlere ilişkin elde edilebilen tüm kombinasyonlar Tablo 1'de verildiği biçimdedir.

Tablo 1. **Ölçüm değerlerine ilişkin sonuçlar**

	X Grubu	Y Grubu	Test İstatistiği
1	76 80 85	17 30 48	48,7
2	76 80 17	85 30 48	3,33
3	76 80 30	17 85 48	12
4	76 80 48	17 30 85	24
5	76 17 85	80 30 48	6,67
6	76 30 85	17 80 48	15,3
7	76 48 85	17 30 80	27,3
8	17 80 85	76 30 48	9,33
9	30 80 85	17 76 48	18
10	48 80 85	17 30 76	30
11	17 30 85	76 80 48	-24
12	17 80 30	76 85 48	-27,3
13	76 17 30	80 85 48	-30
14	17 48 85	76 30 80	-12
15	17 80 48	76 30 85	-15,3
16	76 17 48	30 80 85	-18
17	30 48 85	76 17 80	-3,33
18	30 80 48	76 17 85	-6,67
19	76 30 48	17 80 85	-9,33
20	17 30 48	76 80 85	-48,7

Tablo1'den elde edilen 20 farklı örnekleme ait tablonun dördüncü kolonundaki 20 farklı test istatistiğinden sadece bir tanesi T değerine eşittir. Bu durumda P olasılığı 1/20=0,05 olur.

Genelleştirilmiş Behrens-Fisher problemi ve bu problemin çözüm yöntemi olarak kullanılan permütasyon testi aşağıdaki örnek 3 ile kısaca açıklanabilir.

Örnek 3

Z_i ortalaması $E(Z_i)=0$ ve varyansı $Var(Z_i)=1$ olan hata değişkenleri dizisi, σ_1 ve σ_2 ise keyfi olarak verilen bilinmeyen standart sapmaları göstersin. Toplam örneklem büyüklüğü r=m+n olan iki örneklem olsun.

$X_i = \mu_1 + \sigma_1 Z_i \qquad 1 \leq i \leq m, \ \mu_1 \in R$

$Y_i = \mu_2 + \sigma_2 Z_{m+i} \quad 1 \le i \le n,\ \mu_2 \in R$

$$H_0:\ \mu_1 = \mu_2 \tag{5}$$

olsun. Kısıtlı H_0 hipotezi ise $\widetilde{H}_0 = \{\mu_1 = \mu_2, \sigma_1 = \sigma_2\}$ şeklinde verilir. Ortalamalara dayalı olan testler,

$$T_r = \left(\frac{mn}{r}\right)^{1/2} (\overline{X} - \overline{Y}), \text{ ya da } \quad \widetilde{T}_r = T_r / V_r^{1/2}, \tag{6}$$

$$V_r = \frac{mn}{r} (S_x^2 / m + S_Y^2 / n) \tag{7}$$

şeklinde yazılır (Janssen, 1997).

$\widetilde{T}_r = \widetilde{T}_r((X_{ri})_{i \le r})$, X_{ri} rastgele değişkenlerine bağlı keyfi test istatistiğidir. X_{ri}'nin sabit x_{ri} değerleri ve tek biçimli $\widetilde{P}$-rastgele permütasyon dağılımları için, $\gamma_r \in [0,1]$ ve $\widetilde{P}$, $c_{r,p}(\alpha)$ ' nın çözümüdür. Burada permütasyon test fonksiyonu,

$$\varphi_{r,Perm} = \begin{cases} 1 & \widetilde{T}_r > c_{r,p}(\alpha) \\ \gamma_r & \widetilde{T}_r = c_{r,p}(\alpha) \\ 0 & \widetilde{T}_r < c_{r,p}(\alpha) \end{cases} \tag{8}$$

şeklinde yazılır. Eşitlik (8)'deki $c_{r,p}(\alpha)$ sabiti, Eşitlik (9) sağlanacak şekilde aşağıdaki biçimde bulunur,

$$\widetilde{P}(\widetilde{T}_r((x_{r\sigma(i)i \le r}) > c_{r,p}(\alpha)) + \gamma_r \widetilde{P}(\widetilde{T}_r((x_{r\sigma(i)})_{i \le r}) = c_{r,p}(\alpha)) = \alpha \tag{9}$$

Bu örnek dikkatli bir şekilde incelendiğinde, sıfır hipotezi $H_0 = \{\mu_1 = \mu_2\}$ aşağıdaki dört parçaya bölünebilir.

$H_1 = \{\mu_1 = \mu_2, \sigma_1 = \sigma_2, Z_i$ standart normal$\}$,
$H_2 = \{\mu_1 = \mu_2, Z_i$ standart normal$\}$,
$H_3 = \{\mu_1 = \mu_2, \sigma_1 = \sigma_2\}$ ve
$H_4 = H_0 \setminus \{H_1 \cup H_2 \cup H_3\}$, H_3 sınırlı sıfır hipotezidir.

Her alt hipotez H_j, j=1,2,3 için, bazı optimallik özelliklerine sahip klasik testler vardır. $\varphi_{r,Stud}$ H_1 için optimal olan iki örneklem t testidir. Eşitlik (8)'deki permütasyon testi Eşitlik (6)'daki pay değeri T_r' ye bağlı olursa, $\varphi_{r,Pitm}$ olarak gösterilen diğer seçeneklere karşı Pitman'ın H_3 hipotezi için optimal iki örneklem testine ulaşılır. H_2 hipotezinin test edilmesi problemi, Behrens-Fisher problemi olarak ele alınabilir (Jansenn, 1997).

4. MONTE CARLO SONUÇLARI

Bu bölümde permütasyon testi $\varphi_{r,Perm}$ ile normal ve çarpık dağılımların testi olan $\varphi_{r,Welch}$ Welch testi karşılaştırılmıştır. Tablo 2, Tablo 3, Tablo 4'te sırasıyla α =0,0498, 0,0485 ve 0,050 birinci tür hata olasılıklarıdır. Tablo 2, Tablo 3 ve Tablo 4'te bulunan değerler örnek 3' te verilen normalleştirilmiş hata değişkenleri $Z = (Z_i)$, E(Z)=0 ve V(Z)=1 için elde edilmiştir.

Tablo 2'de elde edilen değerler $\varphi_{r,Perm}$ permütasyon test fonksiyonun her zaman $\varphi_{r,Welch}$ Welch test fonksiyonundan daha güçlü olduğunu göstermektedir. Özellikle, Welch testi çarpık dağılımlarda geçerli değildir. m=4 ve n=8 farklı örneklem büyüklüklerinde hataların normal dağıldığı durumda, Welch testine başvurulmalıdır (Tablo 3'ün ilk satırı). Diğer durumlarda, permütasyon testi tercih edilir. $\sigma_1 = \sigma_2 = 1$(Tablo 3'ün 3.kolonu) olduğunda Welch testi geçerli değildir (Janssen, 1997).

Tablo 2. I. Tür hata olasılıkları m=6, n=6, $\mu_1 = \mu_2$, α = 0,0498

Dağılım	$\sigma_1^2 : \sigma_2^2$	1,0:1,0	1,0:1,1	1,0:1,2
Normal	$\phi_{r,Welch}$	0,0460	0,0458	0,0459
	$\phi_{r,Perm}$	0,0501	0,0500	0,0500
Log-normal	$\phi_{r,Welch}$	0,0251	0,0254	0,0261
	$\phi_{r,Perm}$	0,0496	0,0500	0,0507
Üstel	$\phi_{r,Welch}$	0,0291	0,0297	0,0302
	$\phi_{r,Perm}$	0,0494	0,0501	0,0504
Tek biçimli	$\phi_{r,Welch}$	0,0489	0,0487	0,0490
	$\phi_{r,Perm}$	0,0500	0,0500	0,0502

Tablo 3. I. Tür hata olasılıkları m=4, n=8, $\mu_1 = \mu_2$, α = 0,0485

Dağılım	$\sigma_1^2 : \sigma_2^2$	1,2:1,0	1,1:1,0	1,0:1,0	1,0:1,1	1,0:1,2
Normal	$\phi_{r,Welch}$	0,0509	0,0495	0,0497	0,0488	0,0478
	$\phi_{r,Perm}$	0,0523	0,0497	0,0481	0,0456	0,0436
Log-normal	$\phi_{r,Welch}$	0,0453	0,0416	0,0379	0,0346	0,0316
	$\phi_{r,Perm}$	0,0555	0,0516	0,0486	0,0455	0,0429
Üstel	$\phi_{r,Welch}$	0,0518	0,0478	0,0442	0,0402	0,0377
	$\phi_{r,Perm}$	0,0556	0,0518	0,0481	0,0447	0,0426
Tek biçimli	$\phi_{r,Welch}$	0,0663	0,0646	0,0627	0,0606	0,0589
	$\phi_{r,Perm}$	0,0538	0,0508	0,0482	0,0452	0,0430

Tablo 4. I. Tür hata olasılıkları m=8, n=8, $\mu_1 = \mu_2$, $\alpha = 0{,}050$

Dağılım	$\sigma_1^2 : \sigma_2^2$	1,0:1,0	1,0:1,1	1,0:1,2	1,0:1,5
Normal	$\phi_{r,\text{Welch}}$	0,0478	0,0476	0,0480	0,0479
	$\phi_{r,\text{Perm}}$	0,0506	0,0503	0,0508	0,0510
Log-normal	$\phi_{r,\text{Welch}}$	0,0291	0,0292	0,0299	0,0339
	$\phi_{r,\text{Perm}}$	0,0504	0,0504	0,0515	0,0555
Üstel	$\phi_{r,\text{Welch}}$	0,0342	0,0344	0,0350	0,0382
	$\phi_{r,\text{Perm}}$	0,0507	0,0504	0,0513	0,0541
Tek biçimli	$\phi_{r,\text{Welch}}$	0,0502	0,0497	0,0501	0,0506
	$\phi_{r,\text{Perm}}$	0,0506	0,0502	0,0505	0,0514

5. SAYISAL YÖNTEMLER

İki normal dağılımın ortalamaları test edilirken klasik iki örneklem t-testine başvurulur. Ele alınan kitlelerin varyansları eşit değilse ve bilinmiyorsa ya Welch'in t istatistiği ya da Satterhwaite'ın yaklaşık F testi önerilir. Welch (1974)'in yöntemi normal dağılım durumunda güçlüdür. Normallik koşulunun sağlanmaması durumunda çarpıklık ve basıklık katsayılarının test istatistik değerlerini nasıl etkilediği araştırılmalıdır. Bunun için Cressi ve Whitford, Yuen ve Dixon tarafından önerilen istatistikleri üreten bir bilgisayar yazılımı kullanılmıştır (Reed III, 2003).

İki örneklem t-testinde örneklemlerin bağımsız olduğu varsayılır. Buna göre $X_i \sim N(\mu_x, \sigma^2)$, i=1,2,…,m ve $Y_j \sim N(\mu_Y, \sigma^2)$, j=1,2,…,n$\left(\sigma_X^2 = \sigma_Y^2 = \sigma^2\right)$, dir. Bu koşullar altında,

$$H_0: \mu_X = \mu_Y \qquad H_0: \mu_X = \mu_Y \qquad H_0: \mu_X = \mu_Y$$
$$H_1: \mu_X < \mu_Y \qquad H_1: \mu_X > \mu_Y \qquad H_1: \mu_X \neq \mu_Y$$

hipotezleri kurulur. Hipotezlerin testinde,

$$T = \frac{\left(\bar{X} - \bar{Y}\right)}{\left[\left(\frac{1}{m} + \frac{1}{n}\right)\left(\frac{(m-1)S_X^2 + (n-1)S_Y^2}{m+n-2}\right)\right]^{1/2}} \tag{10}$$

istatistiği kullanılır. Eşitlik (10) ile bulunan T değeri, $T_{\alpha, m+n-2}$ ya da $T_{\alpha/2, m+n-2}$ ile karşılaştırılır. Ancak varyansların homojenliği sağlanmadığında, Welch-Aspin T^* test istatistiği kullanılır. Burada T^* test istatistiği;

$$T^* = \left(\bar{X} - \bar{Y}\right) \Big/ \left(\frac{S_X^2}{m} + \frac{S_Y^2}{n}\right)^{1/2} \sim t_\nu \tag{11}$$

şeklinde yazılır.

T veya T^* istatistikleri dağılımlar normal olduğunda geçerlidir. Dağılımların normal olmaması durumunda ise örneklemlerin çarpıklık ve basıklık katsayılarının T ve T^* istatistik değerlerini nasıl etkilediği araştırılmalıdır. Cressi ve Whitford (1986) çarpıklığın etkisini yok edecek U ve U^* test istatistiklerini bulmuşlardır. Bulunan U, t tablosundaki (m+n-2) serbestlik derecesine ve U^* ise ν serbestlik derecesine karşılık gelen değerle, karşılaştırılır. T veya T^* testi seçenekleri arasında iki örneklem düzeltilmiş T' veya T'^* istatistikleri vardır (Reed III, 2003). Bu test istatistik değerlerini hesaplamada kullanılan bir bilgisayar yazılım programı aşağıdaki örnek4 için uygulanabilir.

Örnek 4

X ve Y gibi iki tür madde verilmektedir. Bu maddelerden X'in asitlik derecesinin Y'nin asitlik derecesinden daha fazla olduğu iddia edilmektedir. Bu iddianın doğruluğunu araştırmak için yapılan deney sonucunda 20 tane X ve 19 tane Y maddesi için elde edilen asitlik dereceleri aşağıda verilmiştir.

X maddesi için asitlik dereceleri: 33, 81, 57, 86, 33, 36, 34, 8, 70, 34, 97, 30, 39, 17, 7, 25, 36, 133, 192, 4

Y maddesi için asitlik dereceleri: 17, 20, 19, 30, 48, 13, 44, 25, 14, 4, 5, 26, 18, 34, 118, 9, 11, 8, 10

Tablo 5. Normallik testi

	Kolmogorov-Smirnov		
	Test istatistiği	s.d	P
X Maddesi	0,265	20	0,001
Y Maddesi	0,220	19	0,016

Tablo5'te P=0,001<α=0,05 ve P=0,016<α=0,05 olduğundan veriler normal dağılım göstermemektir. Normallik sağlanmadığı için Cressi ve Whitford'un çarpıklığı yok eden test istatistikleri kullanılır. Bu test istatistik değerlerini hesaplayan bilgisayar yazılımı (Reed III, 2003) çıktısı aşağıda Tablo 6'da verilmiştir.

Tablo 6. Bilgisayar yazılım çıktısı

	Test istatistiği	Tek yanlı P	İki yanlı P
Student-t testi (varyanslar eşit)	T=2,2813	0,01348	0,02696
Welch-Aspin (varyanslar eşit değil)	T*=2,3136	0,01252	0,02505
Cressi-Whitford (varyanslar eşit)	U=2,6317	0,00594	0,01188
Cressi-Whitford (varyanslar eşit değil)	U*=2,6263	0,00602	0,01203
0,20 kesilmiş ortalama [X'(T20)=152,723,Y'(T20)=86,098]			
Yuen-Dixon (varyanslar eşit)	T'=2,5272	0,00938	0,01876
Yuen (varyanslar eşit değil)	T'*=2,6176	0,01024	0,02047

Tablo 6'daki sonuçlara göre P=0,01252<α=0,05 olduğu için, X'in asitlik derecesinin α=0,05 ile Y'nin asitlik derecesinden daha fazla olduğu söylenebilir.

Veriler normal dağılmadığı için çarpıklığın etkisi araştırmalıdır. Çarpıklığın sıfır olmadığını doğrulamak için U^* test istatistiği varyanslar eşit olmadığından U^*=2,6263 ve tek yönlü P değeri 0,00602'dir. Varyanslar eşit olmadığından 0,20 düzeltilmiş ortalama test istatistiği $T^{'*}$=2,6176 ve tek yönlü P değeri 0,01024'dür. Daha sonra yapılan her iki test de X'in asitlik derecesinin Y'nin asitlik derecesinden daha fazla olduğunu doğrulamıştır. Bu sonuçlar kullanılan bilgisayar yazılımı sonucunda elde edilmiştir.

6. TARTIŞMA VE SONUÇ

Bugüne kadar Behrens-Fisher probleminin çözümü ile ilgili çok sayıda araştırma yapılmış ve çok sayıda öneri sunulmuştur. Thomasse (1974) Fisher'in pratik yöntemlerini özetlemiş, Banerjee, Pagurova, Wald, Hajek, Welch ve Aspin bu konu ile ilgili çeşitli testler geliştirmiştir. Bu çalışmada Behrens-Fisher probleminin çözümü için kullanılan Welch-t testi ve permütasyon testlerine yer verilmiştir. Monte Carlo sonuçlarından yararlanılarak elde edilen tablolar yardımıyla Welch-t ve permütasyon testleri karşılaştırılmıştır. Bu karşılaştırma sonucunda Welch-t testinin normal dağılımlı veriler için daha güçlü, çarpık dağılımlı veriler için ise zayıf olduğu sonucuna ulaşılmıştır. Çarpık dağılımlar için ise permütasyon testlerinin kullanılması önerilmektedir.

7. KAYNAKLAR

Good, P., 1994. Permutation Tests, Practical Guide to Resampling Methods for Testing Hypothesis, Springer Series in Statistics, Springer, Berlin.

Janssen, A., 1997. Studentized Permutation Tests for non-i.i.d. Hypotheses and the Generalized Behrens-Fisher Problem, Statistics & Probability Letters, 36, 9-21.

Jay, D. and Roxy, P., 1993. The Exploration and Analysis of Data, Duxbury Press, Belmont, California.

Lehmann, E.L., 1959. Testing Statistical Hypotheses, Willey Publication in Statistics, New York.

Lehmann, E.L., 1975. Nonparametrics: Statistical Methods Based on Ranks, McGraw-Hill International Book Company, San Fransisco.

Mehta, J. S. and Srinivasan, R., 1970. On the Behrens-Fisher Problem, Biometrica, 57, 649-655.

Pfanzagl, J., 1974. On the Behrens-Fisher Problem, Biometrica, 61, 39-47.

Reed III, J.F., 2003. Solutions to the Behrens-Fisher Problem, Computer Methods and Programs in Biomedicine, 70, 260-261.

[http://www.maths.gmw.ac.uk/~bb/CTS_chapter2_students.pdf, Erişim Tarihi: 10.03.2005]

BEHRENS-FISHER PROBLEM AND SOLUTION METHODS

ABSTRACT

While difference of means of two population of which distiributions are $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ has been searched, if variances of the two population aren't known and equal, the test of means of two populations has been known as Behrens-Fisher Problem. A lot of methods have been improved for the solution of Behrens-Fisher problem. Some of those methods are Welch-t test and permutation test. In this study, Welch-t test, permutation test and numerical methods that are used for solution of Behrens-Fisher problem were discussed Welch-t test and permutation test compared. Lastly, the computer program that produced statistics proposed by Cressi and Whitford has been carried out.

Key Words: Behrens-Fisher Problem, Permutation Test, Welch-t Test.