

## Yeniden Örnekleme (Resampling) Yaklaşımı ve t Testinin Testin Gücü ve I. Tip Hata Bakımından Karşılaştırılması\*

Özgür Koşkan<sup>1\*</sup>, Fikret Gürbüz<sup>2</sup>

<sup>1</sup> Süleyman Demirel Üniversitesi Ziraat Fakültesi Zootekni Bölümü Biyometri-Genetik Anabilim Dalı, Isparta

<sup>2</sup> Ankara Üniversitesi Ziraat Fakültesi Zootekni Bölümü Biyometri-Genetik Anabilim Dalı, Ankara

\*e-posta: ozgurkoskan@gmail.com; Tel: +90 (246) 211 4620; Fax: +90 (246) 237 1693

### Özet

Bu çalışmada, t testi I. tip hata ve testin gücü bakımından yeniden örnekleme yöntemi ile karşılaştırılmıştır. Çalışmada ele alınan testler için 2 gruplu denemeler seçilmiştir. Bu denemelerde grupların alınmış oldukları populasyonların;  $Z(0, 1)$ , Student  $t(10)$ ,  $\chi^2(3)$  ve  $\beta(2.5, 5)$  ve bunların ikili kombinasyonlarının dağılımını gösteren populasyonlar olduğu varsayılarak gruplardaki gözlem adetleri eşit ve farklı olacak şekilde örnekler alınmış ve bu örneklerde 10000 simülasyon denemesi sonucunda, gerçekleşen I. tip hata olasılığı ve testin güç değerleri hesaplanmıştır. Denemede yeniden örnekleme sayısı da 100000 olarak belirlenmiştir. Sonuç olarak, küçük örnek genişliklerinde ve varyanslar heterojen iken yeniden örnekleme yöntemi diğer test den daha iyi sonuçlar vermiştir.

**Anahtar kelimeler:** Yeniden örnekleme, permutasyon testi, bootstrap testi, I. tip hata, testin gücü

### Resampling Approach and Comparison of t-test for Type I Error Rate and Test Power

#### Abstract

In this study, t test was compared in terms of type 1 error rate and power of the test by resampling approaches. The experiments containing two groups were chosen. In the experiments, the samples including either equal or unequal observations in the groups were drawn from populations,  $Z(0, 1)$ , Student  $t(10)$ ,  $\chi^2(3)$  and  $\beta(2.5, 5)$  distributions, and mixed combinations of these distributions. Probability of type 1 error rate and power of the test values were calculated (obtained) from 10000 simulation trials in these samples. The number of resampling was determined as 100000 in the experiments.

As a result, resampling method provided more reliable results than the other for both small sample size and heterogeneous variances.

**Key words:** Resampling, permutation test, bootstrap test, type I error rate, power of test.

#### Giriş

Son yıllarda ortaya çıkan pek çok istatistik yöntem, örneğin güçlü (robust) yöntemler, açıklayıcı veri analizi (exploratory data analysis) ve yeniden örnekleme (resampling) yöntemleri değişik alanlarda araştırmacıların dikkatini çekmektedir. Yinede pek çok araştırmacı bu yeni tekniklerden ziyade geleneksel istatistik yöntemleri benimsemeyi tercih etmektedir. Hatta verileri parametrik değerlendirme için uygun olmasa bile yine parametrik testleri kullanmaktadırlar. Bu tutucu uygulamaya etki eden bir takım faktörler şöyle sıralanabilir. İlk olarak, bu yeni yöntemler üniversitelerde verilen klasik istatistik derslerinde yer almamakta ve böylece bu yeni yöntemlerin kavramları pek çok insana anlaşılması zor ve bir o kadar karmaşık gibi görünmektedir. İkinci neden olarak, son zamanlarda pek çok istatistik paket programı üreticisinin genellikle geleneksel veri analizleri için

istatistik paket programların geliştirilmesine yönelik bir gayret içinde olmalarıdır. Bu nedenle araştırmacılar bu yeni tekniklerin farkında olsalar bile bu değerlendirmelere yönelik sınırlı sayıda paket programın oluşu bu yöntemlerin uygulanmasını engellemektedir. Yeniden örnekleme (resampling) yönelik kavramlar anlaşılrsa ve paket programlar çoğalsın bile, bazı araştırmacılar bu yeni ve ayrıntı görünen yöntemleri uygulamaya tereddüt etmektedir. Çünkü yeni teknikler eleştirilere maruz kalmakta ve teorik alt yapıdan mahrum olarak görülmekte iken, geleneksel yöntemler deneysel gerçek ve teorik olarak oturmuş şekilde algılanmaktadır (Yu, 2003).

Aslında yeniden örnekleme yöntemlerinden rasgeleliğe dayalı testler (randomization test, randomization exact test, permutation test, rerandomization test) 1935 yılında Fisher tarafından geliştirilmiştir. O yıllardan bu güne uzanan çalışmaların hala artan bir şekilde devam etmesi,

yeniden örnekleme yöntemlerinin tam olarak her alanda eksiklerinin giderilmediğini göstermektedir.

Çalışmaların son yıllarda artışı ise; ucuz ve hızlı bilgisayarların geliştirilmesine bağlanabilir. Elle yapılması çok zor hatta imkansızdır. Bu sebepten dolayı yeniden örnekleme yöntemi, bilgisayar-yoğun metotlar (computer-intensive methods) başlığı altında da irdelenmektedir (Fortin *et al.*, 2002). Gelişen bilgisayar teknolojisinin sonucu olarak yeniden örnekleme yöntemleri istatistik bir araç olarak yaygın bir şekilde kullanılmaya eğilimindedir. Çünkü bu yöntemler nispeten güçlü, hesaplama adımları kolay ve uygulamada ön şartlar gerektirmediği için pek çok alanda kullanılmaktadır. Yeniden örnekleme (Resampling) yöntemlerinin sahip olduğu avantajlardan biri de çok fazla matematik bilgisi ya da istatistik formül ihtiyacının olmamasıdır. Kavramların ve yöntemlerin anlaşılmasını sağlayacak kadar bir bilgi yeniden örnekleme (resampling) yaklaşımının uygulamaları için yeterlidir (Butler *et al.*, 2003). Çalışmanın amacı yeniden örnekleme yönteminin grup karşılaştırmalarında gerek 1. tip hata gerekse testin gücüne bakılarak t-testine bir alternatif olarak kullanılabilirliğini ortaya koymaktır.

### Materyal ve Yöntem

Bu çalışmanın materyalini, çeşitli dağılımlara sahip populasyonlardan, grup sayısı, örnek genişliği, varyans kombinasyonları ve standart sapma cinsinden grup ortalamalarına göre Microsoft Power Station Developer Studio FORTRAN ve IMSL Library yardımıyla üretilen tesadüf sayıları teşkil etmektedir.

İki grup ortalaması arasındaki farkın irdelenmesinde dağılım kombinasyonları olarak  $Z : Z, t : t, \chi^2 : \chi^2, \beta : \beta, Z : t, Z : \chi^2, Z : \beta, t : \chi^2, t : \beta, \chi^2 : \beta$ , varyansların oranı 1:1, 1:5, 1:10, 1:30, gruplardaki gözlem adetleri 5:5, 15:15, 30:30, 5:30 standart sapma cinsinden ortalamalar arasındaki farklar 0:0, 0:0.5, 0:1.0, 0:1.5 olacak şekilde kullanılmıştır.

Bu çalışmada yeniden örnekleme yaklaşımı ile ilgili populasyonlardan alınan ortalamalar arasındaki farkların irdelenmesinde aşağıdaki aşamalar izlenmiştir.

1. İlgili standardize edilmiş populasyonlardan belirlenen büyüklüklerde örnekler alınmıştır. Eğer gerçekleşen 1. tip hata üzerinde durulacaksa ve gruplardaki varyanslar homojen kalacaksa gözlemler aynen muhafaza edilmiş, varyansların homojenliği bozulacaksa son gruptaki gözlemler önceden belirlenen sabit sayıların karekökleriyle çarpılmıştır. Testin gücü irdeleniyor ise,

ayrıca son gruptaki gözlemlere belirli sabit sayılar eklenmiştir.

2. Grupların ortalamaları arasındaki farkların tesadüften ileri gelip gelmediği t-testi ile irdelenmiştir. Yapılan hesaplamalar sonunda hesaplanan test istatistiği, ilgili dağılım tablosundaki sınır değerinden ( $\alpha = 0.05$ ) büyükse  $H_0$  hipotezi ret edildiği hükmüne varılmıştır.

3. İkinci aşamada oluşturulmuş olan gruplardaki gözlemlerin hepsi, ilgili gözlem sayılarında olmak üzere, tamamen tesadüfen yeniden gruplara ayrılmış, t istatistiği elde etmeye yönelik hesaplamalar yeniden yapılmış ve t istatistiğinin ikinci aşamada hesaplanmış olan t istatistiğine eşit veya daha büyük olup olmadığı saptanmıştır.

4. Üçüncü aşamadaki işlemler 100000 defa tekrar edilmiş ve ikinci aşamadaki hesaplanmış bulunan t istatistiğine eşit veya daha büyük olanların nispi miktarı 0.05'e eşit veya daha küçükse birinci aşamada oluşturulmuş olan grupların ortalamaları arasındaki farka ilişkin ilgili  $H_0$  hipotezinin yeniden örnekleme yaklaşımında ret edildiği hükmüne varılmıştır.

5. Yukarıdaki dört aşama arasındaki irdelemeler 10000 defa yapılmıştır. Bu 10000 irdelemede ikinci aşamada ret edilenlerin oranı t testinde gruplardan sonuncusundaki gözlemlere herhangi bir ekleme yapılmadığında gerçekleşen 1. tip hata, aksi takdirde gerçekleşen testin gücü olarak, yine bu 10000 irdelemede 4. aşamada ret edilen  $H_0$ 'ların oranı yeniden örneklemede gerçekleşen 1. tip hata veya gerçekleşen testin gücü olarak tanımlanmıştır.

### Bulgular

#### **Gruplardaki Dağılımlar Aynı Olduğunda Gerçekleşen Testin Gücü ve 1. Tip Hata**

Çizelge 1 incelendiğinde; gözlem adetleri 5:5 iken varyansların homojenliği bozulmaya başladığı andan itibaren yeniden örnekleme yönteminin t-testine göre gerçekleşen güç değerlerinin az da olsa yüksek olduğu gözlenmektedir. Bu durum varyans oranlarının 1:10 ve 1:30 olduğu durumlarda biraz daha belirginleşmiştir.

Gruplardaki gözlem adetlerinin 15:15 ve 30:30 olması halinde varyans oranları ne olursa olsun yeniden örnekleme yöntemi ve t-testi bakımından gerçekleşen güç değerlerinin birbirlerine oldukça yakın oldukları görülür. Bu durum 1. tip hata olasılıkları içinde geçerlidir. Gruplardaki gözlem adetleri 5:30 olduğunda ve varyansların homojenliği bozulduğunda her iki testte de testin gücünün belirgin bir şekilde düştüğü

gözlenmiştir.

Gözlem adetleri her iki grupta da 5 iken varyans oranları sırasıyla ilk grubun 5 katı, 10 katı ve 30 katı olduğunda 1.tip hatanın 0.05 den daha büyük olduğu görülmektedir. Grup varyansları homojenken 1. tip hata olasılığının 0.046 ile 0.056 arasında değiştiği görülmektedir. Yeniden örneklemeye yöntemi ve t testi 1.

tip hata bakımından benzer sonuçlar vermektedir. Burada gruplar içindeki gözlem adetlerinin artması ile gerçekleşen 1.tip hata olasılığının daha da dar sınırlar içinde değiştiği (0.049-0.054) dolayısıyla denemenin başında belirlenen 1. tip hatanın korunduğu görülmektedir.

Çizelge 1. Aynı dağılımlı populasyonlardan alınan örneklerde ampirik olarak gerçekleşen testin gücü ve I. tip hata değerleri

$n_A : n_B$	$\sigma^2$		Testin gücü						I. tip hata	
			$\mu_1:\mu_2=0.0.5$		$\mu_1:\mu_2=0:1$		$\mu_1:\mu_2=0:1.5$		$\mu_1:\mu_2=0:0$	
			t	RS	t	RS	t	RS	t	RS
5:5	1:1	Z:Z	0.107	0.109	0.288	0.287	0.553	0.544	0.054	0.055
		tt	0.109	0.118	0.300	0.301	0.567	0.565	0.047	0.050
		$\chi^2:\chi^2$	0.124	0.140	0.357	0.377	0.613	0.624	0.046	0.056
		$\beta:\beta$	0.106	0.104	0.283	0.277	0.537	0.532	0.051	0.05
	1:5	Z:Z	0.086	0.088	0.141	0.147	0.248	0.253	0.058	0.061
		tt	0.078	0.083	0.152	0.161	0.257	0.269	0.057	0.062
		$\chi^2:\chi^2$	0.052	0.064	0.086	0.110	0.184	0.232	0.080	0.085
		$\beta:\beta$	0.070	0.074	0.113	0.123	0.204	0.214	0.069	0.069
	1:10	Z:Z	0.077	0.086	0.111	0.118	0.170	0.181	0.069	0.074
		tt	0.079	0.087	0.117	0.129	0.179	0.192	0.067	0.071
		$\chi^2:\chi^2$	0.067	0.076	0.068	0.084	0.095	0.124	0.098	0.100
		$\beta:\beta$	0.071	0.076	0.098	0.106	0.133	0.146	0.077	0.079
1:30	Z:Z	0.077	0.091	0.089	0.102	0.115	0.130	0.07	0.083	
	tt	0.079	0.094	0.088	0.108	0.112	0.130	0.066	0.078	
	$\chi^2:\chi^2$	0.094	0.100	0.080	0.089	0.075	0.093	0.115	0.121	
	$\beta:\beta$	0.080	0.086	0.082	0.091	0.089	0.104	0.086	0.094	
15:15	1:1	Z:Z	0.268	0.269	0.755	0.754	0.976	0.976	0.046	0.046
		tt	0.274	0.275	0.757	0.758	0.976	0.977	0.05	0.050
		$\chi^2:\chi^2$	0.289	0.296	0.768	0.772	0.964	0.966	0.048	0.050
		$\beta:\beta$	0.256	0.255	0.749	0.749	0.980	0.980	0.049	0.049
	1:5	Z:Z	0.121	0.122	0.335	0.337	0.626	0.629	0.054	0.054
		tt	0.126	0.128	0.334	0.341	0.627	0.633	0.055	0.057
		$\chi^2:\chi^2$	0.087	0.095	0.307	0.327	0.675	0.699	0.067	0.069
		$\beta:\beta$	0.104	0.105	0.323	0.325	0.634	0.639	0.053	0.053
	1:10	Z:Z	0.097	0.099	0.214	0.217	0.400	0.404	0.056	0.058
		tt	0.095	0.098	0.303	0.305	0.419	0.428	0.058	0.06
		$\chi^2:\chi^2$	0.064	0.069	0.149	0.165	0.370	0.399	0.075	0.077
		$\beta:\beta$	0.081	0.083	0.186	0.188	0.382	0.386	0.054	0.055
	1:30	Z:Z	0.071	0.072	0.113	0.115	0.180	0.184	0.057	0.059
		tt	0.073	0.076	0.113	0.119	0.192	0.198	0.058	0.061
		$\chi^2:\chi^2$	0.063	0.066	0.070	0.076	0.112	0.125	0.081	0.083
		$\beta:\beta$	0.068	0.069	0.098	0.101	0.152	0.155	0.062	0.063

t: t-testi; RS: yeniden örneklemeye

Çizelge 1. Aynı dağılımlı populasyonlardan alınan örneklerde ampirik olarak gerçekleşen testin gücü ve I. tip hata değerleri (devam)

$n_A : n_B$	$\sigma^2$		Testin gücü						I. tip hata	
			$\mu_1:\mu_2=0:0.5$		$\mu_1:\mu_2=0:1$		$\mu_1:\mu_2=0:1.5$		$\mu_1:\mu_2=0:0$	
			<i>t</i>	RS	<i>t</i>	RS	<i>t</i>	RS	<i>t</i>	RS
30:30	1:1	Z:Z	0.478	0.479	0.970	0.970	1.000	1.000	0.049	0.049
		<i>tt</i>	0.488	0.489	0.967	0.967	1.000	1.000	0.050	0.050
		$\chi^2:\chi^2$	0.498	0.502	0.957	0.958	1.000	1.000	0.048	0.050
		$\beta:\beta$	0.464	0.464	0.970	0.970	1.000	1.000	0.059	0.051
	1:5	Z:Z	0.211	0.212	0.597	0.599	0.910	0.910	0.053	0.053
		<i>tt</i>	0.209	0.210	0.591	0.594	0.902	0.903	0.05	0.051
		$\chi^2:\chi^2$	0.160	0.165	0.608	0.618	0.952	0.955	0.058	0.060
		$\beta:\beta$	0.184	0.185	0.598	0.599	0.923	0.924	0.052	0.052
	1:10	Z:Z	0.135	0.137	0.368	0.371	0.682	0.684	0.051	0.052
		<i>tt</i>	0.138	0.140	0.378	0.381	0.679	0.681	0.055	0.056
		$\chi^2:\chi^2$	0.088	0.092	0.341	0.353	0.734	0.748	0.065	0.067
		$\beta:\beta$	0.121	0.122	0.186	0.188	0.697	0.700	0.054	0.055
1:30	Z:Z	0.078	0.079	0.166	0.168	0.305	0.308	0.05	0.052	
	<i>tt</i>	0.082	0.085	0.170	0.173	0.321	0.325	0.052	0.053	
	$\chi^2:\chi^2$	0.064	0.067	0.117	0.123	0.263	0.275	0.066	0.068	
	$\beta:\beta$	0.076	0.077	0.149	0.151	0.288	0.290	0.058	0.059	
5:30	1:1	Z:Z	0.172	0.172	0.520	0.521	0.867	0.866	0.049	0.050
		<i>tt</i>	0.178	0.177	0.527	0.524	0.853	0.851	0.051	0.050
		$\chi^2:\chi^2$	0.191	0.201	0.579	0.583	0.854	0.853	0.045	0.048
		$\beta:\beta$	0.178	0.180	0.522	0.525	0.853	0.852	0.051	0.052
	1:5	Z:Z	0.009	0.009	0.052	0.052	0.197	0.199	0.002	0.002
		<i>tt</i>	0.010	0.010	0.055	0.055	0.209	0.210	0.002	0.002
		$\chi^2:\chi^2$	0.002	0.002	0.018	0.022	0.172	0.197	0.008	0.009
		$\beta:\beta$	0.004	0.004	0.035	0.038	0.171	0.176	0.002	0.002
	1:10	Z:Z	0.001	0.001	0.010	0.010	0.041	0.041	0.000	0.000
		<i>tt</i>	0.002	0.002	0.010	0.011	0.046	0.047	0.000	0.001
		$\chi^2:\chi^2$	0.001	0.001	0.001	0.001	0.013	0.016	0.004	0.004
		$\beta:\beta$	0.001	0.001	0.004	0.004	0.026	0.027	0.001	0.001
1:30	Z:Z	0.000	0.000	0.001	0.001	0.002	0.002	0.000	0.000	
	<i>tt</i>	0.000	0.000	0.001	0.001	0.003	0.003	0.000	0.000	
	$\chi^2:\chi^2$	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.002	
	$\beta:\beta$	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000	

*t*: t testi; RS: yeniden örnekleme

Gözlem adetleri farklı iken varyansların homojenliği bozulduğunda başlangıçta belirlenen 1. tip hatanın korunamadığı ve 0.05'den oldukça düşük değerler aldığı görülmektedir.

Fakat gruplardaki gözlem adetlerinin farklı olduğu durumda yine her iki testin de 1. tip hata oranları

bakımından benzer şekilde değerler aldığı Çizelge 1'den açıkça görülmektedir.

### **Gruplardaki Dağılımlar Farklı Olduğunda Gerçekleşen Testin Gücü ve I. Tip Hata**

Çizelge 2 incelendiğinde; varyanslar homojen iken, gözlem adetleri arttıkça testin güç değerlerinin de arttığı

görülmektedir. Varyanslar homojen iken hem t-testi hem de yeniden örneklemeye yaklaşımı benzer sonuçlar vermektedir. Varyanslar homojen iken her iki test içinde başlangıçta karşılaştırılmış olunan %5'lik hatanın korunduğu söylenebilir.

Gözlem adetleri 5:5 iken varyansların homojenliği bozulduğundan itibaren yeniden örneklemeye yöntemi bakımından gerçekleşen güç değerlerinin t-testine göre bir miktar yüksek olduğu gözlenmektedir. Bu durum grup ortalamaları arasındaki farkın 1.0 ve 1.5 standart sapma olduğu durumlarda daha da belirginleşmektedir.

Varyansların homojenliği bozulduğunda başlangıçta karşılaştırılmış olunan %5'lik hata korunamamıştır.

Varyans oranları 1:1 iken gruplardaki gözlem adetlerinin artması (15:15, 30:30) güç değerlerinin de

belirgin bir şekilde artmasına neden olmuştur. Bu deneme koşullarında her iki testte benzer değerler almıştır. Diğer taraftan varyansların homojenliği bozuldukça ve gruplar arasındaki standart sapma cinsinden farklar arttıkça yeniden örneklemeye yönteminin biraz daha güçlü olduğu görülmüştür. Fakat gruplardaki gözlem adetleri 5:5 iken bu farklılığın az da olsa daha belirgin olduğu görülmektedir. Gözlem adetleri arttığı için denemenin başında belirlenen 1. tip hatanın genelde korunduğu söylenebilir.

Gruplardaki gözlem adetleri sayısı 30:30 olduğunda her iki testte benzer güç değerleri almaktadır. Hemen hemen bütün kombinasyonlarda denemenin başında belirlenen 1. tip hatanın korunduğu görülmektedir.

Çizelge 2. Dağılımlar farklı olduğunda 10000 deneme sonucunda gerçekleşen testin gücü ve I. tip hata değerleri

$n_A : n_B$			Testin gücü						I. tip hata	
			$\mu_1:\mu_2=0:0.5$		$\mu_1:\mu_2=0:1$		$\mu_1:\mu_2=0:1.5$		$\mu_1:\mu_2=0:0$	
	$\sigma^2$		t	RS	t	RS	t	RS	t	RS
5:5	1:1	Z: $\chi^2$	0.078	0.085	0.260	0.278	0.573	0.595	0.058	0.061
		Z:t	0.106	0.108	0.301	0.300	0.561	0.561	0.046	0.048
		Z: $\beta$	0.094	0.096	0.263	0.265	0.542	0.537	0.055	0.053
		t: $\chi^2$	0.074	0.085	0.263	0.289	0.578	0.606	0.06	0.065
		t: $\beta$	0.096	0.099	0.269	0.271	0.557	0.560	0.052	0.053
		$\chi^2$ : $\beta$	0.133	0.135	0.335	0.336	0.583	0.578	0.053	0.054
	1:5	Z: $\chi^2$	0.057	0.061	0.081	0.096	0.163	0.198	0.086	0.085
		Z:t	0.078	0.083	0.148	0.158	0.258	0.271	0.059	0.063
		Z: $\beta$	0.072	0.074	0.114	0.119	0.204	0.209	0.067	0.066
		t: $\chi^2$	0.058	0.064	0.080	0.097	0.169	0.200	0.077	0.078
		t: $\beta$	0.071	0.074	0.114	0.119	0.197	0.207	0.064	0.064
		$\chi^2$ : $\beta$	0.076	0.083	0.127	0.134	0.207	0.223	0.066	0.069
	1:10	Z: $\chi^2$	0.066	0.069	0.062	0.071	0.093	0.118	0.100	0.100
		Z:t	0.075	0.085	0.111	0.122	0.182	0.197	0.065	0.072
		Z: $\beta$	0.072	0.078	0.091	0.098	0.132	0.139	0.079	0.080
		t: $\chi^2$	0.067	0.073	0.063	0.078	0.090	0.113	0.102	0.103
		t: $\beta$	0.074	0.077	0.094	0.101	0.132	0.144	0.082	0.083
		$\chi^2$ : $\beta$	0.072	0.079	0.096	0.107	0.140	0.157	0.079	0.081
1:30	Z: $\chi^2$	0.095	0.101	0.085	0.093	0.074	0.089	0.114	0.117	
	Z:t	0.081	0.093	0.094	0.114	0.110	0.130	0.073	0.086	
	Z: $\beta$	0.079	0.088	0.086	0.098	0.096	0.108	0.086	0.093	
	t: $\chi^2$	0.095	0.102	0.081	0.092	0.075	0.091	0.118	0.120	
	t: $\beta$	0.082	0.091	0.087	0.099	0.095	0.107	0.088	0.097	
	$\chi^2$ : $\beta$	0.083	0.092	0.084	0.098	0.092	0.105	0.088	0.098	

t: t-testi; RS: yeniden örneklemeye

Çizelge 2. Dağılımlar farklı olduğunda 10000 deneme sonucunda gerçekleşen testin gücü ve I. tip hata değerleri (devam)

$n_A : n_B$	$\sigma^2$		Testin gücü						I. tip hata	
			$\mu_1:\mu_2=0:0.5$		$\mu_1:\mu_2=0:1$		$\mu_1:\mu_2=0:1.5$		$\mu_1:\mu_2=0:0$	
			$t$	RS	$t$	RS	$t$	RS	$t$	RS
15:15	1:1	$Z:\chi^2$	0.243	0.249	0.785	0.789	0.992	0.993	0.054	0.055
		$Z:t$	0.266	0.267	0.749	0.749	0.972	0.972	0.055	0.056
		$Z:\beta$	0.257	0.257	0.765	0.765	0.980	0.980	0.051	0.052
		$t:\chi^2$	0.243	0.251	0.788	0.794	0.987	0.988	0.051	0.053
		$t:\beta$	0.254	0.255	0.752	0.752	0.981	0.981	0.049	0.049
		$\chi^2:\beta$	0.293	0.292	0.749	0.749	0.958	0.959	0.050	0.050
	1:5	$Z:\chi^2$	0.079	0.084	0.290	0.309	0.680	0.708	0.067	0.068
		$Z:t$	0.124	0.128	0.350	0.355	0.638	0.644	0.053	0.055
		$Z:\beta$	0.107	0.108	0.317	0.320	0.638	0.644	0.057	0.057
		$t:\chi^2$	0.079	0.085	0.285	0.304	0.678	0.704	0.066	0.066
		$t:\beta$	0.112	0.113	0.304	0.307	0.632	0.637	0.055	0.055
		$\chi^2:\beta$	0.111	0.112	0.320	0.323	0.640	0.644	0.056	0.057
	1:10	$Z:\chi^2$	0.060	0.066	0.150	0.165	0.373	0.401	0.078	0.081
		$Z:t$	0.096	0.098	0.223	0.227	0.421	0.427	0.055	0.058
		$Z:\beta$	0.080	0.081	0.182	0.186	0.382	0.386	0.059	0.059
		$t:\chi^2$	0.064	0.067	0.145	0.159	0.361	0.387	0.076	0.078
		$t:\beta$	0.080	0.082	0.189	0.193	0.374	0.380	0.063	0.063
		$\chi^2:\beta$	0.090	0.092	0.184	0.187	0.376	0.380	0.062	0.063
	1:30	$Z:\chi^2$	0.060	0.063	0.073	0.080	0.113	0.125	0.087	0.088
		$Z:t$	0.070	0.073	0.111	0.116	0.188	0.194	0.058	0.060
		$Z:\beta$	0.070	0.070	0.095	0.097	0.158	0.161	0.066	0.067
		$t:\chi^2$	0.058	0.062	0.071	0.079	0.110	0.123	0.085	0.088
		$t:\beta$	0.066	0.067	0.091	0.092	0.158	0.160	0.058	0.059
		$\chi^2:\beta$	0.065	0.066	0.093	0.095	0.156	0.160	0.062	0.063

t: t-testi; RS: yeniden örnekleme

Gruplardaki gözlem adetleri farklı iken her iki test benzer sonuçlar vermektedir. Ancak varyansların homojenliği bozulduğunda testin güç değerlerinin aşırı derecede düştüğü ve denemenin başında belirlenen 1. tip hatanın korunamadığı görülmektedir.

### Tartışma

Çalışmada gerçekleşen 1. tip hata olasılıkları bakımından, varyanslar homojen olduğunda, t testinde dağılım kombinasyonları ne olursa olsun gerçekleşen 1. tip hatanın başlangıçta kararlaştırılan ( $\alpha = 0.05$ ) değerini koruduğu ve 1. tip hatada gözlenen sapmaların dağılımlardan ziyade, varyansların heterojenliği, grupların ortalamaları arasındaki farklar ve gözlem adetlerinden kaynaklandığı görülmektedir. Bu durum, Başpınar ve Gürbüz (2000), Keskin (2002), Mendes (2002) gibi araştırmacıların bulgularıyla uyum içindedir.

Çalışmada gerçekleşen testin güç değerleri incelendiğinde, grup ortalamaları arasındaki farklar artırdıkça testin güç değerleri de artmaktadır. Yine gruplardaki gözlem adetleri artırıldığında gerçekleşen testin güç değerlerinin de arttığı görülmektedir. Fakat varyansların oranları 1:5, 1:10, 1:30 olduğunda gerçekleşen testin güç değerlerinin azaldığı gözlenmiştir. Bu durum, Başpınar ve ark. (1999b), Başpınar ve Gürbüz (2000), Başpınar (2001), Keskin (2002), Mendes (2002), Weber (2006)'da da benzer şekilde bildirilmiştir.

Çalışmada kullanılan yeniden örnekleme yönteminin ve t testinin başlangıçta kararlaştırılan 1. tip hatayı koruma bakımından benzer sonuçlar verdiği görülmektedir. Bu durum, Routledge (1997) ile benzerlik göstermezken, Mewhort (2005) ve Weber (2006) ile uyum içindedir.

Testin gücü dikkate alındığında ise, özellikle örnek genişliğinin küçük (5:5) ve varyansların homojen olmadığı (1:5, 1:10, 1:30) durumlarda yeniden örneklemeye yönteminin t testinden daha güçlü olduğu görülmüştür. Normallik ve varyansların homojenliği ön şartlarının birlikte gerçekleşmediği durumlarda küçük hacimli örneklerle çalışıldığında yeniden örneklemeye yönteminin tercih edilmesinde yarar vardır. Heller and Venkatraman (1996), Ludbrook and Dudley (1998), Good (2000), Balasubramani ve ark. (2005), Pesarin and

Salmaso (2006) çalışmalarında benzer bildirimlerde bulunmuşlardır.

Örnek genişlikleri 30:30'a ulaştığında ise, her iki test de benzer sonuçları vermektedir.

Gerek gerçekleşen 1. tip hata ve gerekse gerçekleşen testin gücü bakımından yeniden örneklemeye yöntemi t testine benzer ve azda olsa üstün durumlar göstermiştir.

Çizelge 2. Dağılımlar farklı olduğunda 10000 deneme sonucunda gerçekleşen testin gücü ve I. tip hata değerleri (devam)

$n_A : n_B$			Testin gücü						1. tip hata	
			$\mu_1:\mu_2=0:0.5$		$\mu_1:\mu_2=0:1$		$\mu_1:\mu_2=0:1.5$		$\mu_1:\mu_2=0:0$	
	$\sigma^2$		t	RS	t	RS	t	RS	t	RS
30:30	1:1	Z: $\chi^2$	0.480	0.484	0.980	0.981	1.000	1.000	0.058	0.059
		Z:t	0.475	0.475	0.965	0.966	1.000	1.000	0.049	0.050
		Z: $\beta$	0.470	0.470	0.973	0.973	1.000	1.000	0.050	0.050
		t: $\chi^2$	0.477	0.480	0.979	0.980	1.000	1.000	0.050	0.051
		t : $\beta$	0.482	0.483	0.973	0.973	1.000	1.000	0.051	0.051
		$\chi^2$ : $\beta$	0.490	0.490	0.953	0.954	0.999	0.999	0.053	0.054
	1:5	Z: $\chi^2$	0.156	0.162	0.608	0.621	0.957	0.961	0.060	0.061
		Z:t	0.205	0.206	0.602	0.605	0.910	0.911	0.051	0.052
		Z: $\beta$	0.182	0.183	0.599	0.601	0.924	0.925	0.051	0.051
		t: $\chi^2$	0.154	0.161	0.611	0.624	0.962	0.965	0.062	0.063
		t : $\beta$	0.179	0.181	0.594	0.595	0.927	0.928	0.057	0.057
		$\chi^2$ : $\beta$	0.184	0.186	0.590	0.591	0.916	0.916	0.052	0.052
	1:10	Z: $\chi^2$	0.093	0.099	0.339	0.352	0.726	0.741	0.067	0.069
		Z:t	0.133	0.135	0.380	0.383	0.689	0.693	0.055	0.056
		Z: $\beta$	0.118	0.118	0.362	0.364	0.688	0.691	0.053	0.054
		t: $\chi^2$	0.094	0.099	0.330	0.344	0.731	0.745	0.064	0.065
		t : $\beta$	0.119	0.121	0.358	0.360	0.693	0.695	0.053	0.053
		$\chi^2$ : $\beta$	0.120	0.121	0.359	0.360	0.693	0.695	0.052	0.053
	1:30	Z: $\chi^2$	0.062	0.065	0.118	0.126	0.256	0.268	0.066	0.068
		Z:t	0.083	0.085	0.172	0.176	0.316	0.320	0.052	0.054
		Z: $\beta$	0.073	0.074	0.152	0.153	0.288	0.289	0.056	0.056
		t: $\chi^2$	0.063	0.066	0.117	0.124	0.270	0.284	0.068	0.070
		t : $\beta$	0.078	0.078	0.148	0.149	0.285	0.288	0.052	0.053
		$\chi^2$ : $\beta$	0.070	0.070	0.147	0.148	0.284	0.286	0.053	0.054

t: t-testi; RS: yeniden örneklemeye

Çizelge 2. Dağılımlar farklı olduğunda 10000 deneme sonucunda gerçekleşen testin gücü ve I. tip hata değerleri (devam)

$n_A : n_B$	$\sigma^2$		Testin gücü						I. tip hata	
			$\mu_1:\mu_2=0:0.5$		$\mu_1:\mu_2=0:1$		$\mu_1:\mu_2=0:1.5$		$\mu_1:\mu_2=0:0$	
			$t$	RS	$t$	RS	$t$	RS	$t$	RS
5:30	1:1	$Z:\chi^2$	0.167	0.166	0.531	0.531	0.873	0.871	0.065	0.066
		$Z:t$	0.176	0.175	0.520	0.518	0.851	0.850	0.053	0.052
		$Z:\beta$	0.163	0.164	0.530	0.531	0.862	0.862	0.053	0.054
		$t:\chi^2$	0.164	0.164	0.536	0.538	0.872	0.871	0.060	0.061
		$t:\beta$	0.161	0.162	0.528	0.529	0.861	0.861	0.050	0.050
		$\chi^2:\beta$	0.180	0.185	0.566	0.570	0.850	0.851	0.041	0.041
	1:5	$Z:\chi^2$	0.003	0.004	0.033	0.037	0.163	0.181	0.006	0.006
		$Z:t$	0.008	0.009	0.058	0.057	0.203	0.204	0.001	0.001
		$Z:\beta$	0.005	0.005	0.039	0.040	0.162	0.167	0.002	0.002
		$t:\chi^2$	0.003	0.004	0.032	0.036	0.159	0.176	0.005	0.006
		$t:\beta$	0.005	0.005	0.039	0.041	0.170	0.175	0.002	0.002
		$\chi^2:\beta$	0.003	0.003	0.031	0.032	0.173	0.180	0.003	0.003
	1:10	$Z:\chi^2$	0.001	0.001	0.002	0.002	0.017	0.020	0.004	0.004
		$Z:t$	0.002	0.002	0.012	0.012	0.045	0.045	0.000	0.000
		$Z:\beta$	0.001	0.001	0.004	0.004	0.027	0.028	0.001	0.001
		$t:\chi^2$	0.001	0.001	0.002	0.002	0.016	0.019	0.003	0.004
		$t:\beta$	0.001	0.001	0.006	0.006	0.027	0.029	0.001	0.001
		$\chi^2:\beta$	0.001	0.001	0.003	0.003	0.023	0.024	0.001	0.001
	1:30	$Z:\chi^2$	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.002
		$Z:t$	0.000	0.000	0.001	0.001	0.003	0.003	0.000	0.000
		$Z:\beta$	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000
		$t:\chi^2$	0.001	0.001	0.001	0.001	0.000	0.000	0.001	0.001
		$t:\beta$	0.000	0.000	0.001	0.001	0.001	0.001	0.000	0.000
		$\chi^2:\beta$	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000

t: t-testi; RS: yeniden örnekleme

## Sonuç

Sonuç olarak, birbirinden bağımsız grup ortalamaları arasındaki farkların irdelenmesinde gerekli ön şartlardan özellikle varyansların homojenliği ön şartı sağlanmaması hem 1. tip hata hem de testin gücünü etkilemektedir. Gruplardaki gözlem adetleri azaldıkça ve de gözlem adetleri bakımından dengesizlik arttıkça, gruplardaki varyansların oranları büyüdükçe, grup ortalamaları arasındaki farklar standart sapma cinsinden küçüldükçe, gruplardaki dağılımlar farklılaştıkça hem gerçekleşen 1. tip hatada hem de gerçekleşen testin gücündeki etkilenmeler artmakta ve gerçekleşen 1. tip hata ile gerçekleşen güç değerleri sıfıra kadar inebilmektedir. Varyansların homojenliği, bu durumu bir miktar dengeleyebilmektedir. Bağımsız grupların karşılaştırılmasında gerekli ön şartlar yerine

gelmediğinde yeniden örnekleme yaklaşımı önerilebilir. Ayrıca yeniden örnekleme yönteminin gerekli ön şartlar yerine geldiğinde de, gerçekleşen 1. tip hata ve gerçekleşen testin gücü bakımından parametrik testlere eşdeğer sonuçlar verdiği de göz önünde tutulmalıdır. Yeniden örnekleme yaklaşımı parametrik olan testlerin hiç birinin ön şartlarını gereksindirmez, kullanımı gayet basit ve kolaydır. Ne var ki yoğun hesaplamaları bilgisayar olmadan gerçekleştirmek mümkün değildir, özellikle birbiri ile karşılaştırılacak gruplardaki gözlem adetleri arttıkça hesaplamalar için gerekli zaman da artmaktadır.

Elde edilen veriler ışığında, yöntemin küçük örnek genişliklerinde daha kullanılabilir olduğu görülmektedir. Örnek genişliklerinin büyümesi ile gerek harcanan zamanın artması gerekse de elde edilen güç



değerlerinin t-testine benzer sonuçlar vermesi yeniden örneklemeye yönteminin kullanımını kısıtlamaktadır.

### Kaynaklar

- Balasubramani, G.K., Wisniewski, S.R., Zhang, H. and Eng, H.F. 2005. Development of an efficient SAS macro to perform permutation tests on two independent samples. *Computer Methods and Programs in Biomedicine* 79:179-187.
- Başpınar, E. ve Gürbüz, F. 2000. Normal, Beta, Gamma (Ki-kare) ve Weibull dağılımlarının ikili kombinasyonlarından alınan değişik örnek genişliğindeki örneklerin karşılaştırılmasında testin gücü. *Tarım Bilimleri Dergisi* 6(1): 116-127.
- Başpınar, E., Çamdeviren, H. ve Gürbüz, F. 1999. Student t-Testi ve varyans analizi tekniğinde testin gücü ve uygun örnek genişliğinin saptanması. *Tarım Bilimleri Dergisi* 5(3): 116-123.
- Başpınar, E. 2001. Değişik varyans oranlı normal populasyonlardan alınan değişik örnek genişliğindeki iki örnekte Student t, Welch ve ayıklanmış t-testlerinin uygulanması ile elde edilen I. tip hata ve testin gücü. *Tarım Bilimleri Dergisi* 7(1): 151-157.
- Butler, A., Rothery, P. and Roy, D. 2003. Minitab macros for Resampling methods. *Teaching Statistics* 25(1): 22-25.
- Good, P. 2000. *Permutation tests: A practical guide to resampling methods for testing hypotheses*. Springer-Verlag, 270, New York.
- Heller, G. and Venkatraman, E.S. 1996. Resampling procedures to compare two survival distributions in the presence of right-censored data. *Biometrics* 52(4): 1204-1213.
- Keskin, S. 2002. Varyansların homojenliğini test etmede kullanılan bazı yöntemlerin I. tip hata ve testin gücü bakımından irdelenmesi. Doktora tezi (basılmamış), Ankara Üniversitesi 210 s., Ankara.
- Ludbrook, J. and Dudley, H. 1998. Why permutation tests are superior to t and F tests in biomedical research. *The American Statistician* 52(2):127-132.
- Mendeş, M. 2002. Normal dağılım ve varyansların homojenliği ön şartlarının gerçekleşmediği durumlarda varyans analizi tekniğinin yerine kullanılacak bazı parametrik testlerin I. tip hata ve testin gücü bakımından irdelenmesi. Doktora tezi (basılmamış), Ankara Üniversitesi 278 s., Ankara.
- Mewhort, D.J.K. 2005. A comparison of the randomization test with the F test when error is skewed. *Behavior Research Methods* 37(3):426-435.
- Weber, M. 2006. Robustness and power of the t, permutation t and Wilcoxon tests. Ph. D. thesis (unpublished), University of Wayne State University 75 s., Detroit, Michigan.
- Yu, C.H. 2003. Resampling methods: concepts, applications, and justification. *Practical Assessment, Research & Evaluation*. 8(19) <http://PAREonline.net/getvn.asp?v=8&n=19>. Erişim Tarihi: 30.05.2004