1999 Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Dergisi, Y.7, S.7, s.25-37

NORMAL DAĞILIMDAN SAPMA DURUMUNDA ANOVA'DA UYGULANAN DÖNÜŞÜMLER ÜZERİNE BİR SİMÜLASYON ÇALIŞMASI

Benian TEKİNDAL*

Özet

Bu makalede, varyans analizi tekniğinin ön şartlarının biri olan normal dağılmış olma ön şartının yerine gelmediği durumlarda simülasyonla elde edilen χ^2 dağılımı gösteren verilere varyans analizi tekniği uygulanmış ve gerçekleşen 1. Tip hata olasılıkları tespit edilmiştir. Daha sonra bu veriler logaritmik , karekök ve ters transformasyonlara tabi tutularak varyans analizi tekniği uygulanmış ve 1. Tip hata oranlarında meydana gelen değişmelere bakılmıştır.

χ^2 dağılımı gösteren verilerde 1,3,5 ve 15 serbestlik derecesi dikkate alınarak üretilen veriler eşit ve farklı sayılarda gözlem içeren 2,3,4 grup halinde düzenlendikten sonra hem orijinal verilere hem de transforme edilmiş olanlara varyans analizi tekniği uygulanmış sonuçta χ^2 dağılımında logaritmik ve karekök transformasyonlarının 1. Tip hata olasılıklarının başlangıçta belirlenen % 5 seviyesinde kalmasını sağladığı tespit edilmiştir.

Çalışmada her bir dağılım şekli , grup sayısı , örnek genişliği ve transformasyon tipi kombinasyonu için hesaplamalar üretilen 100.000' er veri setinde gerçekleştirilmiştir.

Anahtar Kelimeler: *$\chi2$ dağılım, transformasyonlar, varyans analizi tekniği, ön şartlar.*

A SIMULATION STUDY ON THE TRANSFORMATIONS APPLIED WHEN THE NORMALITY ASSUMPTION IS VIOLATED IN ANOVA

Abstract

In this paper, ANOVA is applied on the simulated data sampled from a χ^2 distribution and the probability of Type I error of the test is determined. The same type of error was also observed in the same data with logarithmic, square root and inverse transformations. The same type of error was also observed in the same data with logarithmic, square root and inverse transformations. Analysis applied to 2, 3 and 4 groups including equal and unequal observations showed that above mentioned types of transformations did not result in any change in Type I error.

ANOVA was also applied to 2, 3 and 4 groups of equal and unequal observations simulated to have 1, 3, 5 and 15 degrees of freedom with χ^2 distribution and it was found that the square-root and logarithmic transformations kept Type I error at predetermined 5 % level.

* Dr. G.Ü. Fen Ed. Fak. Öğretim Görevlisi

As a result logarithmic square root and inverse transformations were not found effective in heterogeneous variances in normal distribution where as logarithmic and square-root transformations were effective in χ^2 distributions. This study was conducted in 100.000 data set.

Key Words: *Chi-square distribution, transformations, analysis of variance technique and assumptions.*

1. GİRİŞ

Birçok istatistik test tekniğinde olduğu gibi varyans analizi tekniğinin uygulanabilmesi için de verilerde bir takım özellikler aranır. Bu özellikler uygulanacak test tekniğinin ön şartları olarak adlandırılır.

Varyans analizi tekniğinin uygulanabilmesi için belli başlı dört tane temel ön şart vardır. Bu ön şartlar kısaca; verilerin normal dağılım göstermesi, farklı muamele gruplarında yer alan verilerin varyanslarının homojen olması, farklı muamele gruplarında yer alan deney ünitelerinin birbirinden farklı olması (bağımsızlık) ve iki ve daha fazla faktör içeren denemelerde faktör etkilerinin eklemeli olması şeklinde tanımlanır. Bu ön şartları gerçekleştiren verilere varyans analizi tekniği uygulanabilir (Düzgüneş Vd. 1987: 246).

Bir veya daha fazla ön şartın gerçekleşmediği durumlarda ise verilere ya başka istatistik analiz tekniği uygulanır veya uygun bir test tekniği düşünülemediği durumda veriler değerlendirilemez. En uygun analiz tekniği belirlerken göz önüne alınan nokta ise başlangıçta belirlenmiş olan 1. tip hata ihtimalini (α) değiştirmeyecek analiz tekniğidir (Tekindal, 1998: 1).

Varyans analizi tekniğinin ön şartları gerçekleşmediği taktirde, analizi geçerli kılmanın çareleri olarak en iyi yöntemin ön şartları bozan durumlardan kaçınmak olduğu söylenebilir. Fakat bazen deneme materyalinin, bazen de etkisi denenen muamelelerin yapısı yönünden bu imkansız olabilir. Kimya reaksiyonlarının hızı, organizmalarda büyüme olayı, bir kültürdeki bakteri yoğunluğunun değişmesi gibi olayların dayandığı ihtimal fonksiyonları normal olmayabilir (Sezgin, 1972: 20).

Bütün denemelerde seçilecek yol olayın matematik tanımına ve tarifine bağlıdır. Bazen, üzerinde denemeyi kuracak büyüklükte homojen materyal bulmak imkansız olabilir veya denemenin yürütülmesi esnasında elde olmayan sebeplerle bazı ünitelerde ön şartların tam olarak gerçekleşmesini engelleyen hasar ve etkiler olabilir. Bu nedenle analize başlamadan önce deneme hakkında geniş bir fikir edinmelidir.

Varyans analizi tekniğinin ön şartları gerçekleşmeden bu analize baş vurulursa hipotez kontrolünden önce kararlaştırılmış bulunan I. tip hata ihtimalinin hangi büyüklükte gerçekleştiğini bilmek mümkün olmaz. Yani araştırıcı H_o hipotezini red ederken yanılmış olma ihtimalinin en fazla % 5 olduğu kanısındayken yanılma ihtimali belki de % 15, % 20 veya % 30 olarak gerçekleşmektedir. Böylesine bir durumun "karar verme" aşamasında ve buna bağlı olarak önerilecek uygulamalarda ne kadar büyük yanılgılara sebep olabileceği açıktır. Bu gibi durumlarda çözüm, bu çeşit ön şartları gerektirmeyen ve verilerin orijinal durumuna uygun başka bir analiz tekniği kullanmak mesela, parametrik olmayan testler gibi veya yeni bir analiz tekniği geliştirmektir. Parametrik olmayan testlerde, mukayese edilen grupların temsil ettikleri populasyonların parametreleri ve dağılım fonksiyonu belli değildir veya belirtilmez. Bu testlerin uygulanmaları için normal bir dağılım şekli veya populasyon şartı aranmaz. Bu nedenlerdendir ki az gözlemli deneylere uygulanabilirler ve nispî etkinlikleri genellikle düşüktür.

Bu makalede simülasyonla elde edilen χ^2 dağılımı gösteren verilerde varyans analizi tekniğinin ön şartları gerçekleşmediği durumlarda veri yapısındaki değişmelerin, yani verilere uygulanacak transformasyonların, yaygın olarak kullanılan transformasyon tekniklerinden logaritmik, karekök ve ters transformasyon tekniklerinin, (Hoyle, 1973: 203 - 223) varyans analizi tekniğinin ön şartlarını ve analiz sonucunda elde edilen yorumları nasıl etkilediği araştırılmış ve yapılan simülasyon çalışmasında normal dağılma ön şartı yerine gelmediği zaman varyans analizi tekniği ile değerlendirildiğinde bunun gerçekleşen I. Tip hataya etkisi araştırılmıştır.

2.1. Materyal

Bu çalışmada, materyal olarak bilgisayarda üretilmiş tesadüf sayıları kullanılmıştır. Bu sayılardan bir kısmı aynı ortalamalı ve değişik varyanslı dağılım gösteren populasyonlardan, 1, 3, 5 ve 15 serbestlik dereceli χ^2 dağılımı gösteren populasyonlardan alınmıştır. Bu dağılımlardan alınan gözlemler iki, üç ve dört grup halinde düzenlenmiş ve her grupta değişik örnek genişlikleri ele alınmıştır. Tesadüfi sayı üretimi ve hesaplamalar için QBASIC 4.0 programlama dilinden yararlanılmıştır.

2.2. Metod

Normal dağılımdan sapmanın varyans analizi tekniğinde hipotez kontrolü yaparken 1. tip hata yapma ihtimali olan α üzerine etkisini araştırmak amacıyla materyal bölümünde açıklanan verilerle yapılan simülasyon çalışmasında varyansların heterojenliğinin çeşitli seviyeleri ile gruplardaki gözlem sayılarının çeşitli seviyeleri ele alınarak grup sayısına ait tüm kombinasyonlarda varyans analizi tekniği uygulanmış, bu denemeler her kombinasyon için 100.000 defa tekrarlanmıştır. Her bir denemeden elde edilen sonuçlar ayrı ayrı bulunmuş ve bu deneme sonuçları toplu halde değerlendirilmiştir. Hesaplamalar sonucunda orijinal verilerde red edilen H_o hipotezi sayıları ve aynı verilere logaritmik, karekök ve ters transformasyon uygulandığında red edilen hipotez sayıları tespit edilmiştir.

Gruplardaki gözlem sayılarının eşit ve farklı olduğu durumlar için ve ayrıca gözlem sayısının az ve çok olduğu durumlar için χ^2 dağılımı gösteren verilere uygulanan logaritmik, karekök ve ters transformasyonların, varyans analizi tekniği sonucunda 1.tip hata yapma ihtimali üzerine etkisini araştırmak amacıyla χ^2 dağılımından alınan tesadüf örneklerinin her birisi iki, üç ve dört grup oluşturacak şekilde düzenlemiştir. Söz konusu transformasyonların etkisini araştırmak amacıyla bir, üç, beş ve on beş serbestlik dereceli χ^2 dağılımları kullanılmıştır.

Bir serbestlik dereceli χ^2 dağılımından alınan örnekler önce iki gruba ayrılmış ve gruplardaki gözlem sayıları 2:2, 5:5, 15:15 ve 30:30 olacak şekilde düzenlenmiştir. Böylece iki gruptaki gözlem sayılarının eşit olduğu koşullar için varyans analizi tekniği ile sonuçlar değerlendirilmiştir. Bir serbestlik dereceli χ^2 dağılımından alınan örnekler 3 grup halinde düzenlendikten sonra ilk olarak gruplardaki gözlem sayılarının eşit olduğu farklı koşullar ele alınmıştır. Bu koşullardaki örnek genişlikleri ise sırasıyla şöyledir. 2:2:2, 5:5:5, 15:15:15 ve 30:30:30. Daha sonra 3 gruptaki gözlem sayıları birbirinden farklı tutulmuş ve bu amaçla örnek genişlikleri 2:5:10, 5:10:15, 5:10:20, 10:20:50 şeklinde düzenlenmiştir. Grup sayısı 4 olarak alındığı zaman gruplardaki gözlem sayılarının eşit olduğu durumlar için belirlenen örnek genişlikleri 2:2:2:2, 5:5:5:5, 15:15:15:15 ve 30:30:30:30 şeklindedir. Gruplardaki gözlem sayılarının farklı olduğu durumlar için belirlenen örnek genişlikleri ise 2:4:6:8, 5:10:15:20, 5:10:20:40 ve 15:20:30:50 olarak belirlenmiştir.

χ^2 dağılımında, serbestlik derecesinde meydana gelen değişmeler neticesinde varyans analizi tekniği ile elde edilen sonuçların nasıl değiştiğini görmek amacıyla bir serbestlik dereceli χ^2 dağılımına ilaveten 3, 5 ve 15 serbestlik dereceli χ^2 dağılımlarından alınmış tesadüf örneklerinden de yararlanılmıştır. 3, 5 ve 15 serbestlik dereceli χ^2 dağılımları için yapılan düzenlemeler, belirlenen grup sayıları ve gruplardaki gözlem sayıları 1 serbestlik dereceli χ^2 dağılımı için anlatılan durumlar gibidir. Böylece χ^2 dağılımı gösteren verilere uygulanacak varyans analizi tekniğinde sonuçların ne şekilde değiştiğini, 1. tip hatadaki değişmenin ne yönde olduğunu görmek amacıyla eşit örnek genişlikleri için toplam 48 farklı koşul, farklı örnek genişlikleri için ise toplam 56 farklı koşul değerlendirilmiştir. Bu koşulları değerlendirmek amacıyla orijinal verilere karekök, ters ve logaritmik transformasyon uygulanmış ve varyans analizi tekniği ile I. tip hata yapma ihtimalleri belirlenmiştir. Ayrıca denemede oluşturulan kombinasyonlar 2, 3 ve 4 grup halinde düzenlenmiştir.

χ^2 dağılımlı populasyonlardaki grup sayısı, varyans oranları, eşit ve eşit olmayan örnek genişliklerine ilişkin bilgiler Tablo 1'de topluca verilmiştir.

Tablo 1: χ^2 dağılımlı populasyonlara ilişkin simülasyon düzenleri

		Örnek genişlikleri	
Grup Sayısı	Serbestlik derecesi	Eşit	Farklı
2	1, 3, 5, 15	2, 5, 15, 30	2:5, 2:10, 2:20:5:10, 5:20:10:30
3	1, 3, 5, 15	2, 5, 15, 30	2:5:10, 5:10:15, 5:10:20, 10:20:50
4	1, 3, 5, 15	2, 5, 15, 30	2:4:6:8, 5:10:15:20, 5:10:20:40, 15:20:30:50

Tablo 1'de görüldüğü gibi bütün kombinasyonlar için simülasyon çalışmaları orijinal ve transforme edilmiş gözlemlerde 100.000'er defa tekrarlanmıştır.

3. BULGULAR

Bu çalışmada, materyal ve metod bölümünde açıklanan, χ^2 dağılımı gösteren veriler için normal dağılımlarından sapma durumlarında varyans analizi tekniği sonuçlarını ne şekilde etkilendiği açıklanmış ve varyansların normallik ön şartı sağlanmadığı durumlarda verilere logaritmik, ters ve karekök transformasyonları uygulamanın etkili olup olmadığı araştırılmıştır. Yukarıda bahsedilen transformasyonların sonuçları ne şekilde etkilediğini göstermek amacıyla yapılan simülasyon çalışması sonuçları tablo halinde sunulmuştur.

3.1. Bir serbestlik dereceli χ^2 dağılımından her birinde eşit sayıda gözlem bulunan gruplara ilişkin sonuçlar

Grup sayısının 2, 3 ve 4 olduğu durumda gruplardaki gözlem sayıları eşit iken gerçekleşen 1.tip hata olasılıklarına ilişkin sonuçlar Tablo 2' de verilmiştir.

Tablo 2 incelendiği zaman, iki grup olduğunda ve gruplardaki gözlem sayıları birbirine eşit iken gerçekleşen 1.tip hata olasılıklarının başlangıçta belirlenen % 5 düzeyini tam olarak korumadığı gözlenmektedir. Sadece her bir grupta iki gözlem var iken gerçekleşen α değeri ile başlangıçta belirlenen α değeri birbirine benzer çıkmıştır. Gruplardaki gözlem sayılarında gözlenen artış gerçekleşen α değerinin başlangıçta belirlenenden düşük çıkmasına sebep olmuştur. Grup sayısı iki ve gruplardaki gözlem sayıları 2, 5, 15 ve 30 olduğu durumda gerçekleşen α değerleri sırasıyla % 5, % 3.2, % 4.3 ve % 4.7 olarak bulunmuştur.

Tablo 2: Bir serbestlik dereceli χ^2 dağılım gösteren veriler ve grup sayısı 2 ve gruplardaki gözlem sayısı eşit olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayıları

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:2	5041	5.0
	5:5	3184	3.2
	15:15	4265	4.3
	30:30	4592	4.7
3	2:2:2	5462	5.5
	5:5:5	3383	3.4
	15:15:15	3994	4.0
	30:30:30	4632	4.6
4	2:2:2:2	5719	5.7
	5:5:5:5	3602	3.6
	15:15:15:15	4198	4.2
	30:30:30:30	4584	4.6

Grup sayısı 3 ve gruplardaki gözlem sayısı 2 iken gerçekleşen α değeri % 5.5 olarak bulunmuş, gözlem sayısı 5 iken gerçekleşen α değeri % 3.4'e düşmüş ve gözlem sayısı 15 ve 30'a çıktığı zaman ise gerçekleşen α değerleri % 4 ve % 4.6 olup biraz daha artmış fakat belirlenen seviyeye ulaşamamıştır.

Grup sayısının 4 olduğu koşullarda ise gerçekleşen α değerinde meydana gelen değişmeler grup sayısının 3 olduğu koşullarla hemen hemen aynıdır. Burada da 1. tip hata olasılığı gruplardaki gözlem 2 olduğu durumda yükselmekte, % 5.7 iken gözlem adetleri 5 olduğunda düşmekte gözlem adetleri 15 ve 30 olduğunda ise biraz daha yükselmektedir. Gözlem sayısındaki artışlar gerçekleşen 1. tip hata olasılığını başlangıçta belirlenen seviyeye yaklaştırmaktadır.

Bu tip verilere logaritmik, karekök ve ters transformasyonların uygulanması gerçekleşen α değerlerini istenilen düzeye getirmemiştir. Ancak logaritmik transformasyona tabi tutulan verilerde gerçekleşen 1. tip hata olasılıkları % 4.3 ile % 5.1 arasında değişmişlerdir. Karekök transformasyonunda ise bu değerler % 4.9 ile % 5.7 arasında değişmektedirler.

Ters (1/X) transformasyonu ile yapılan analizlerde ise gerçekleşen 1. tip hata olasılıkları % 0.6 ile % 3.2 arasında değişmektedir. Bu durumda ters (1/X) transformasyonunun uygun olmadığı söylenebilir.

3.2. Bir serbestlik dereceli χ^2 dağılımında her birinde farklı sayıda gözlem bulunan gruplara ilişkin sonuçlar

Gruplardaki gözlem sayıları birbirinden farklı iken grup sayılarına bağlı olarak gerçekleşen 1. tip hata olasılıkları Tablo 3'te topluca sunulmuştur.

Tablo 3: Üç serbestlik dereceli χ^2 dağılımı gösteren 2, 3 ve 4 grupta ve gruplardaki gözlem sayılarının farklı iken 100.000 deneme sonucunda red edilen hipotez sayıları

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:5	4521	4.5
	2:10	5738	5.7
	2:20	5979	6.0
	5:10	3920	3.9
	5:20	4468	4.5
	10:30	4385	4.4
3	2:5:10	5299	5.3
	5:10:15	4372	4.4
	5:10:20	4540	4.5
	10:20:50	4744	4.7
4	2:4:6:8	4806	4.9
	5:10:15:20	4702	4.7
	5:10:20:40	5164	5.2
	15:20:30:50	4462	4.5

Tablo 3 incelendiğinde, grup sayısı 2 olduğunda ve gruplardaki gözlemler birbirinden farklı iken orijinal verilere uygulanan varyans analizi sonuçlarında gözlem 1. tip hata olasılıkları, örnek genişliklerine bağlı olarak başlangıçta belirlenen α değerinden düşük veya yüksek bulunmuştur. Örnek genişliklerinin oldukça farklı olduğu durumda (2 ve 20) gerçekleşen 1. tip hata olasılığının % 6'ya ulaştığı dikkati çekmektedir. 2 grup için gerçekleşen sonuçlara benzer sonuçlar 3 ve 4 grupta da elde edilmiştir.
Orijinal verilere logaritmik transformasyon uygulandıktan sonra varyans analizi tekniği kullanıldığında grup sayısı 2 ve gruplardaki gözlem sayıları da 2:5 iken gerçekleşen 1. tip hata olasılığı % 4.8 olup başlangıçta belirlenen % 5'e çok yakın bir değer olmuştur. Gruplardaki gözlem sayıları 2:10 iken gerçekleşen α değeri % 5.4, 2:20 iken % 5.0 ve 5:10 iken % 4.7 ve 10:30 iken % 4.8 olarak tespit edilmiştir. Bu sonuca göre gerçekleşen α değerindeki değişmelerin orijinal verilere göre daha az olduğu görülmektedir.

Grup sayısı 3 ve gruplardaki gözlem sayıları da 2:5:10, 5:10:15, 5:10:20 ve 10:20:50 iken gerçekleşen α değerleri sırasıyla % 5.1, % 4.7, % 4.8 ve % 5.0 olup 2 gruptaki sonuçlara benzer sonuçlar elde edilmiştir.

Grup sayısı 4'e çıkınca gerçekleşen α değerler gruplardaki gözlem sayılarına bağlı olarak sırasıyla % 5.9, % 4.9, % 5.1 ve % 4.7 olarak gerçekleşmişlerdir. Bu sonuçlara göre logaritmik transformasyonu sonunda gerçekleşen 1. tip hata olasılıklarının kararlaştırılan % 5 değerine nispeten yakın bulundukları söylenebilir.

Karekök transformasyonunda elde edilen sonuçlar incelendiği zaman gerçekleşen 1. tip hata olasılıklarının % 4.4 ile % 5.0 arasında olduğu görülmüştür. Buna göre karekök transformasyonunun bu tip veriler için etkili sonuçlar verdiği ileri sürülebilir.

Ters (1/X) transformasyon sonucu elde edilen 1. tip hata olasılıkları ise 2 grupta % 1.8 ile % 11.9, 3 grupta % 2.2 ile % 9.2, 4 grupta ise % 2.3 ile % 7.8 arasında muntazam olmayan değişmeler göstermişlerdir. Bu durumda Ters (1/X) transformasyonunun uygun olmadığı görülmektedir.

3.3 Üç serbestlik dereceli χ^2 dağılımında, her birinde eşit sayıda gözlem bulunan gruplarda ilişkin sonuçlar

Üç serbestlik dereceli χ^2 dağılımından alınan veriler her birinde eşit sayıda gözlem bulunan 2, 3 ve 4 gruba ayrıldıktan sonra uygulanan varyans analizi sonuçlarında elde edilen 1. tip hata olasılıkları Tablo 4'te sunulmuştur.

Tablo 4'ün incelenmesinden grup sayısı 2 ve gruplardaki gözlem sayılarının sırasıyla 2, 5, 15 ve 30 olduğu durumlarda gerçekleşen 1. tip hata olasılıklarının sırasıyla % 5.2, % 4.4, % 4.8 ve 4.5 olduğu görülmektedir. Benzer sonuçlar grup sayısının 3 ve 4 olduğu durum içinde geçerlidir. Bu durumda gruplardaki gözlem sayısının artışının gerçekleşen α değerinin başlangıçta belirlenen düzeye yaklaşmasında etkili olduğunu fakat grup sayısındaki değişimin ise bunu etkilemediği söylenebilir.

Tablo 4: Üç serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruplardaki gözlem sayısı eşit olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayıları

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:2	5235	5.2
	5:5	4419	4.4
	15:15	4778	4.8
	30:30	4522	4.5
3	2:2:2	5392	5.4
	5:5:5	4395	4.4
	15:15:15	4561	4.6
	30:30:30	4846	4.8
4	2:2:2:2	9346	5.4
	5:5:5:5	4304	4.3
	15:15:15:15	4686	4.7
	30:30:3:30	4795	4.8

Bu tip verilere logaritmik transformasyon uygulandığı zaman gerçekleşen α değeri grup sayısı ve gruplardaki gözlem sayılarına bağlı olarak % 4.7 ile % 5.0 arasında değişmekte olup, bu sonuçlar kararlaştırılan % 5'e orijinal verilerdekinden daha yakındırlar.

Karekök transformasyonu uygulanan verilerden elde edilen sonuçlarda ise gerçekleşen α değerleri % 4.1 ile % 5.3 arasında olup α değerindeki değişmeler, logaritmik transformasyonunkine nazaran biraz daha fazladır. Ters transformasyonu uygulanması durumunda ise gerçekleşen α değerleri % 0.05 ile % 4.5 arasında olup oldukça düşüktürler.

Bu sonuçlara göre bu tip veriler için en uygun sonuçların logaritmik ve karekök transformasyonlar ile elde edildiği görülmektedir.

3.4. Üç serbestlik dereceli χ^2 dağılımında, her birinde farklı sayıda gözlem bulunan gruplara ilişkin elde edilen sonuçlar

Bu koşullara ilişkin elde edilen sonuçlar Tablo 5'de topluca sunulmuştur.

Tablo 5: Üç serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruplardaki gözlem sayısı farkı olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayıları

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:5	4943	4.9
	2:10	5002	5.0
	2:20	4901	4.9
	5:10	4555	4.6
	5:20	4760	4.8
	10:30	4780	4.8
3	2:5:10	4998	5.4
	5:10:15	4641	4.3
	5:10:20	4771	4.7
	10:20:50	4874	4.8
4	2:4:6	4929	5.0
	5:10:15:20	4849	4.9
	5:10:20:40	4773	4.7
	15:20:30:40	4707	4.7

Tablo 5'deki sonuçlar incelendiği zaman grup sayısı 2 ve gruplardaki gözlem sayıları 2:5, 2:10, 2:20, 5:10, 5:20 ve 10:30 iken gerçekleşen 1. tip hata olasılıklarının sırasıyla % 4.9, % 5, % 4.9, % 4.6, % 4.8 ve % 4.8 olduğu görülmektedir. Bu değerler kararlaştırılan % 5'e oldukça yakındırlar.

Grup sayısı 3 ve 4 olduğunda gerçekleşen 1. tip hata olasılıkları % 4.3 ile % 5.4 arasında değişmekte olup 2. gruba oldukça benzer sonuçlar elde edilmiştir. Bu durumda orijinal verilere uygulanan varyans analizi sonucunda gruplardaki gözlem sayılarındaki değişmelerin ve grup sayısındaki değişmelerin gerçekleşen α değerini pek etkilemediği, fakat başlangıçta kararlaştırılan % 5 seviyesinden biraz düşük olduğu söylenebilir.

Aynı tablo incelendiğinde bu tip verilere logaritmik transformasyon uygulandığı zaman gerçekleşen 1. tip hata olasılıkları % 4.8 ile % 5.1 arasında değişmiştir. Benzer durum karekök transformasyonu uygulandığında da gözlenmiştir.

Ancak ters (1/X) transformasyon uygulandığında, farklı durumlarda gerçekleşen α değeri bazen oldukça düşük (% 3.2) bazen de oldukça yüksek (% 8) olarak bulunmuştur.

3.5. Beş serbestlik dereceli χ^2 dağılımında, her birinde eşit sayıda gözlem bulunan gruplara ilişkin elde edilen sonuçlar

Her birinde eşit sayıda gözlem bulunan, 2, 3 ve 4 grup halinde düzenlenen verilere ilişkin elde edilen sonuçlar 6'da verilmiştir.

Tablo 6: Beş serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruplardaki gözlem sayısı eşit olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayısı

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:2	5128	5.1
	5:5	4551	4.6
	15:15	4884	4.9
	30:30	4555	4.6
3	2:2:2	5227	5.2
	5:5:5	4611	4.6
	15:15:15	4703	4.7
	30:30:30	5053	5.1
4	2:2:2:2	5262	5.3
	5:5:5:5	4609	4.6
	15:15:15:15	4642	4.6
	30:30:30:30	5009	5.0

Tablo 6'da orijinal verilerde grup sayısı 2 ve gruplardaki gözlem sayıları 2, 5, 15 ve 30 iken gerçekleşen 1. tip hata olasılıkları sırasıyla % 5.1, % 4.6, % 4.9 ve % 4.6 olarak bulunmuştur. Grup sayısı 3 ve 4 iken elde edilen sonuçlar grup sayısı 2 olduğu durumlardaki gibidir.

Bu tip verilere logaritmik transformasyon uygulandıktan sonra yapılan varyans analizleri ile gerçekleşen 1. tip hata olasılıkları % 4.7 ile % 5.1 arasında değişmiştir.

Karekök transformasyonu uygulandıktan sonra gerçekleşen 1. tip hata olasılıkları % 4.8 ile % 5.1 arasında bulunmuştur.

Ters transformasyon uygulanan verilerden elde edilen sonuçlarda gerçekleşen 1. tip hata olasılıkları başlangıçta belirlenen % 5 seviyesinden daha düşük bulunmuştur.

3.6. Beş serbestlik dereceli χ^2 dağılımında, birbirinde farklı sayıda gözlem bulunan gruplara ilişkin sonuçlar

Gruplardaki gözlem sayılarının birbirinden farklı olduğu durumlarda elde edilen sonuçlar, Tablo 7'de verilmiştir.

Tablo 7: Beş serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruplardaki gözlem sayısı farklı olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayısı

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orijinal veri (X)	
		Sayı	%
2	2:5	4936	4.9
	2:10	5009	5.0
	2:20	4743	4.7
	5:10	4540	4.5
	5:20	4689	4.7
	10:30	4907	4.9
3	2:5:10	5059	5.1
	5:10 15	4702	4.7
	5 10:20	4848	4.8
	10:20:50	4914	4.9
4	2:4 6:8	4884	4.9
	5:10:15:20	4917	4.9
	5:10:20:40	4816	4.8
	15 20:30:50	4771	4.8

Tablo 7 incelendiğinde grup sayısı 2 ve gruplardaki gözlem sayıları 2:5, 2:10, 2:20, 5:10, 5:20 ve 10:30 iken gerçekleşen 1. tip hata olasılıkları sırasıyla % 4.9, % 5, % 4.7, % 4.5, % 4.7 ve % 4.9 olarak bulunmuştur. Grup sayısı 3 ve 4 iken gerçekleşen 1. tip hata olasılıkları ise % 4.7 ile % 5.1 arasında olup başlangıçta belirlenen % 5 seviyesine biraz daha yaklaşan değerler almışlardır.

Bu tip verilere logaritmik ve karekök transformasyonları uygulandıktan sonra yapılan varyans analizleri sonucunda gerçekleşen 1. tip hata olasılıkları % 4.8 ile % 5.1 arasında değişmektedir. Bu durumda da verilere logaritmik ve karekök transformasyonları uygulamanın gerçekleşen 1. tip hata olasılıkları üzerine daha etkili olacağı gözlenmektedir.

Ters (1/X) transformasyonu uygulandıktan sonra elde edilen 1. tip hata olasılıklarının % 3.6 ile % 6.6 arasında değiştiği gözlenmiştir. Bu durumda ters ransformasyonun gerçekleşen 1. tip hata olasılıklarını başlangıçta belirlenen seviyeye yaklaştırmada etkili olmadığı söylenebilir.

3.7. On beş serbestlik dereceli χ^2 dağılımında, her birinde eşit sayıda gözlem bulunan gruplara ilişkin sonuçlar

Bu koşular altında yapılan varyans analizlerine göre elde edilen 1. tip hata olasılıkları Tablo 8'de sunulmuştur.

Tablo 8: On beş serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruplardaki gözlem sayısı eşit olduğu durumda 100.000 deneme sonucunda red edilen hipotez sayısı

Grup Sayısı	Gruplardaki Gözlem Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orjinal veri (X)	
		Sayı	%
2	2:2	5099	5.1
	5:5	4845	4.8
	15:15	4908	4.9
	30:30	4674	4.7
3	2:2:2	5149	5.1
	5:5:5	4867	4.9
	15:15:15	4956	4.9
	30:30:30	5211	5.2
4	2:2:2:2	5210	5.2
	5:5:5:5	4779	4.8
	15:15:15:15	4960	5.0
	30:30:3:30	5199	5.2

Tablo 8 incelendiğinde grup sayısı 2 ve gruplardaki gözlem sayıları 2, 4, 15 ve 30 iken gerçekleşen 1. tip hata olasılıkları sırasıyla % 5.1, % 4.8, % 4.9 ve % 4.7 olarak tespit edilmiştir.

Grup sayısı 3 ve 4 iken gerçekleşen 1. tip hata olasılıkları % 4.8 ile % 5.2 arasında değişmekte olup bunların başlangıçta belirlenen % 5 seviyesine yakın değerler aldığı gözlenmiştir. Bu durumda serbestlik derecesinin artışının gerçekleşen α değerinin başlangıçta belirlenen seviyeye yaklaşmasını sağladığı söylenebilir.

Bu tip verilere logaritmik ve karekök transformasyonları uygulandıktan sonra elde edilen 1. tip hata olasılıkları % 4.9 ile % 5.2 arasında değişmekte olup orijinal verilerden daha iyi sonuçlar vermiştir. Böylece orijinal verilere logaritmik ve karekök transformasyonu uygulamanın daha etkili sonuçlar vereceği söylenebilir.

Ters (1/X) transformasyonu uygulandıktan sonra gerçekleşen 1. tip hata olasılıkları % 4.2 ile % 5.1 arasında değişmekte olup genellikle % 5'in altında kalmışlardır.

Serbestlik derecesi arttıkça ters transformasyon sonucunda gerçekleşen 1. tip hata olasılıkları başlangıçta belirlenen % 5 seviyesine yaklaşmaktadır. Ancak orijinal, logaritmik ve karekök transformasyonlarına nazaran α 'daki değişimlerin daha fazla olduğu gözlenmiştir.

3.8. On beş serbestlik dereceli χ^2 dağılımında, her birinde farklı sayıda gözlem bulunan gruplara ilişkin sonuçlar

Bu koşullar altında elde edilen 1. tip hata olasılıkları Tablo 9'da sunulmuştur.

Tablo 9: On beş serbestlik dereceli χ^2 dağılımı gösteren veriler ve gruplardaki gözlem sayısı farklı olduğu durumda 100.000 deneme onucunda red edilen hipotez sayısı

Grup Sayısı	Gruplardaki Gözlam Sayısı (N)	100000 Deneme Sonucunda Red Edilen Hipotez Sayıları	
		Orjinal veri (X)	
		Sayı	%
2	2:5	4872	5.0
	2:10	5048	5.0
	2:20	4777	4.8
	5:10	4879	4.9
	5:20	4911	4.9
	10:30	4854	4.9
3	2:5:10	5001	5.0
	5:10:15	5088	5.1
	5:10:20	4840	4.8
	10:20:50	5075	5.1
4	2:4:6:8	4834	4.8
	5:10:15:20	5070	5.1
	5:10:20:40	5055	5.1
	15:20:30:50	5059	5.1

Tablo 9 incelendiği zaman grup sayısı 2, 3 ve 4 iken ve gruplardaki gözlem sayıları birbirinden farklı iken gerçekleşen 1. tip hata olasılıklarının % 4.8 ile % 5.1 arasında değişmekte oldukları ve başlangıçta belirlenen % 5 seviyesine oldukça yakın değerler aldıkları gözlenmektedir. Gruplardaki gözlem sayılarının farklı olması gerçekleşen 1. tip hata olasılıkları üzerinde olumsuz bir etki yapmamıştır.

Verilere logaritmik ve karekök transformasyonları uygulandıktan sonra elde edilen 1. tip hata olasılıkları da % 4.8 ile % 5.1 arasında olup orijinal verilere benzer sonuçlar elde edilmiştir.

Ters transformasyon (1/X) sonucunda ise α değerleri % 4.5 ile % 5.3 arasında olup diğer durumlara nazaran biraz daha iyi sonuçlar elde edilmiştir. Burada da gerçekleşen 1. tip hata olasılıkları genellikle % 5'in altında kalmışlardır.

Bütün bu açıklamalar altında χ^2 dağılımı gösteren veriler için serbestlik derecesindeki artışının, gerçekleşen 1. tip hata olasılıklarının başlangıçta belirlenen seviyeye yaklaşmasını sağladığı görülmektedir. Bu durumun sebebi ise serbestlik dereceleri büyüdükçe χ^2 dağılımlarının şekillerinin simetrikleşmeye başlayıp, normal dağılımı andırmasıdır. Serbestlik derecesinin düşük olduğu durumlarda ise logaritmik ve karekök transformasyonlarının nispeten etkili sonuçlar verdiği söylenebilir.

4. GENEL DEĞERLENDİRMELER

Bir serbestlik dereceli χ^2 dağılımı gösteren verilerle yapılan analizlerde 2, 3, 4 grup karşılaştırmasında gruplardaki gözlem adetleri eşit ve de farklı olduklarında gerçekleşen 1. tip hata olasılıkları genellikle % 5'e oldukça yakın gerçekleşmekte, logaritmik ve $\sqrt{X}$ transformasyonları gerçekleşen 1. tip hata olasılıklarını % 5'e oldukça yaklaştırmaktadır. 1. tip hata olasılıklarını dengelemek bakımından ters (1/X) transformasyonunun uygun olmadığı sonucuna varılmıştır.

3, 5 ve 15 serbestlik dereceli χ^2 dağılımlarında da benzer durumlar gözlenmektedir. Özellikle 15 serbestlik dereceli χ^2 dağılımında 2, 3 ve 4 grup karşılaştırmasında, gruplardaki gözlemler eşit ve farklı olduğunda 1. tip hata olasılıklarının % 4,7 ile % 5,2 arasında

karşılaştırmalarında grupların varyansları homojendir. Dolayısıyla varyansların homojenliği ön şartı yerine gelmiş olmaktadır. Serbestlik dereceleri büyüdükçe χ^2 dağılımlarının şekli simetrikleşmeye dolayısıyla da normal dağılımı andırmaya başladığından gerçekleşen 1. tip hata olasılıkları da % 5'e yaklaşmaktadırlar. Buradan da varyans analizi tekniğinin ön şartlarından olan "Varyansların homojenliği" ön şartının 1. tip hata olasılığının oluşumunda diğer önemli ön şart olan "normal dağılım" (Tekindal 1999: 329). Ön şartından daha belirleyici olduğu görülmektedir.

KAYNAKLAR

Düzgüneş , O. , Kesici , T., ve Kavuncu , O., Gürbüz, F., 1987 *Araştırma Ve Deneme Metodları (İstatistik Metodları . II).* A.Ü. Ziraat Fakültesi Yayınları: 1201. Ders Kitabı: 295, Ankara.

Hoyle, M.H., "Tranformations an introduction and bibliography". *International Statistcal Review*. 41, 203 223.1973.

Sezgin, F., Varyans Analizinin Dayandığı Faraziyeler ve Bunların Bozulmasından Doğan Durumlar. Doktora Tezi. Atatürk Üniversitesi Ziraat Fakültesi. Erzurum 1972.

Tekindal, B., Varyans Analizinin Ön şartları ve Transformasyonlar Doktora Tezi. A.Ü Fen Bilimleri Enstitüsü, Zootekni Anabilim Dalı, Ankara 1998.

Tekindal, B., "Varyans Analizinin ön şartları gerçekleşmediği durumlarda yapılan transformasyonların simülasyon yöntemi ile araştırılması", *Ticaret ve Turizm Eğitim Fakültesi Dergisi*, 2. 323 – 329, 1999.