

# Evaluation of the Effect of Normality and Homogeneity Violations in Finding Appropriate Statistical Techniques in Research in the Context of Related Literature

## Araştırmalarda Uygun İstatistiksel Tekniğin Belirlenmesinde Normallik ve Homojenlik İhlallerinin Etkisinin İlgili Literatür Bağlamında Değerlendirilmesi

Ergül DEMİR<sup>1</sup>



Ankara Üniversitesi, Eğitim Bilimleri Fakültesi,  
Ankara, Türkiye

İrem ÇELİK<sup>2</sup>



Ordu Üniversitesi, Eğitim Fakültesi, Ordu,  
Türkiye

Sibel URLU<sup>3</sup>



Sağlık Bakanlığı, Ankara, Türkiye



Bu çalışma TÜBİTAK tarafından  
desteklenen 223K382 no'lu proje  
kapsamında hazırlanmıştır.

Geliş Tarihi/Received 23.08.2024  
Kabul Tarihi/Accepted 17.09.2024  
Yayın Tarihi/Publication 30.09.2024  
Date

Sorumlu Yazar/Corresponding author:  
Ergül DEMİR

E-mail: erguldemir@ankara.edu.tr

Cite this article: Demir, E., Çelik, İ., & Urlu, S. (2024). Evaluation of the effect of normality and homogeneity violations in finding appropriate statistical techniques in research in the context of related literature. *Journal of Psychometric Research*, 2(2), 52-62.



Content of this journal is licensed under a  
Creative Commons Attribution-  
NonCommercial 4.0 International License.

### ABSTRACT

Finding the appropriate statistical technique in research is a challenging step that requires taking into account the methodological design of the research, as well as the purpose of the analysis, sample size and structure, and the characteristics of the data and dataset. Parametric tests, or more generally linear models, are considered tests and models that are sensitive to the basic assumptions of (1) normality and (2) homogeneity of variances. Serious violations of normality and homogeneity may lead to the decision not to use these tests. On the other hand, what constitutes a serious level is an uncertain and controversial issue. In this context, the aim of this study is; to discuss and evaluate in which cases and conditions violations of normality and homogeneity can be considered "serious" or "negligible" in determining the appropriate statistical technique, and also what kind of solutions can be applied in such cases, based on relevant research and basic reference books. This study is a theoretical study in the form of a compilation conducted on the findings of basic reference books and related research in the relevant literature. In this context, 20 basic reference books with high recognition in the field and 25 articles examining the effects of assumption violations on statistical techniques were identified. Discussions were conducted in line with the research questions. As a result, the related literature indicates that, when sufficient sample size and balance between groups are ensured, violations of normality and homogeneity will not cause serious bias in the use of parametric techniques and models. Although it varies depending on the complexity level of the statistical technique to be used, in general, samples of around 40 in univariate techniques and around 150 and above in multivariate techniques, as well as proportional differences between group sizes not exceeding 4-5 times, are seen as conditions where violations of normality and homogeneity can be ignored.

**Keywords:** Skewness, outliers, assumption, parametric test, linear models

### ÖZ

Araştırmalarda uygun istatistiksel tekniğin belirlenmesi, araştırmanın metodolojik tasarımının yanı sıra analizin amacı, örneklem büyüklüğü ve yapısı, verilerin ve veri setinin karakteristiğinin de dikkate alınmasını gerektiren zorlayıcı bir aşamadır. Parametrik testler veya daha genel bir tanımlamayla doğrusal modeller, (1) normallik ve (2) varyansların homojenliği temel varsayımlarına duyarlı test ve modeller olarak değerlendirilmektedir. Ciddi düzeyde normallik ve homojenlik ihlallerinin görülmesi, bu testlerin kullanılmaması kararına yönlendirebilmektedir. Diğer taraftan, ciddi düzeyin ne olduğu, belirsizlik içeren ve tartışmalı bir konudur. Bu kapsamda bu çalışmanın amacı; uygun istatistiksel tekniğin belirlenmesinde normallik ve homojenlik ihlallerinin hangi durum ve koşullarda "ciddi" veya "ihmal edilebilir" olarak değerlendirilebileceğini, ayrıca

bu tür durumlarda ne tür çözümlere başvurulabileceğini, ilgili araştırmalar ve temel kaynak kitaplar üzerinden tartışmak ve değerlendirmektir. Bu çalışma, ilgili literatürdeki temel kaynak kitaplar ve ilişkili araştırmaların ortaya koyduğu bulgular üzerinden yürütülen derleme niteliğinde kuramsal bir çalışmadır. Bu kapsamda alanda tanınırlığı yüksek olan 20 temel kaynak kitap ve varsayım ihlallerinin istatistiksel teknikler üzerindeki etkilerini inceleyen 25 makale belirlenmiştir. Araştırma soruları doğrultusunda tartışmalar yürütülmüştür. Sonuç olarak ilgili literatür, yeterli örneklem büyüklüğü ve gruplar arası denge sağlandığında, normallik ve homojenlik ihlallerinin, parametrik teknik ve modellerin kullanılmasında, ciddi düzeyde bir yanlılık oluşturmayacağını göstermektedir. Kullanılacak istatistiksel tekniğin karmaşıklık düzeyine bağlı olarak değişmekle birlikte genel olarak tek değişkenli tekniklerde 40 civarı, çok değişkenli tekniklerde ise 150 civarı ve üzeri örneklem, ayrıca grup büyüklükleri arasında 4-5 katı geçmeyen oransal farklılıklar, normallik ve homojenlik ihlallerinin göz ardı edilebileceği koşullar olarak görülmektedir.

**Anahtar Kelimeler:** Çarpıklık, uçdeğer, varsayım, parametrik test, doğrusal modeller

## Giriş

Araştırmalarda uygun istatistiksel tekniğin belirlenmesi, araştırmanın metodolojik tasarımının yanı sıra analiz amacını, örneklem büyüklüğü ve yapısı, verilerin ve veri setinin karakteristiğinin de dikkate alınmasını gerektiren zorlayıcı bir aşamadır. Yüzlerce istatistiksel teknik ve model içerisinde uygun olanın hangisi olduğunu belirleyebilmek, belli düzeyde bir yeterlik ve deneyim gerektirmektedir. Her istatistiksel teknik ve modelin kendine özgü algoritmaları ve koşulları bulunmaktadır. İstatistiksel tekniğin varsayımları (assumptions) olarak değerlendirilebilen bu kendine özgü karakteristikler, bir tekniğin hangi koşullarda geçerli ve güvenilir kestirimler üretebileceğini tanımlamaktadır. Bu kapsamda, uygun bir istatistiksel tekniğin belirlenmesinde, bu tekniğin temel varsayımlarının karşılanıp karşılanmadığının incelenmesi, yönlendirici olmaktadır.

Parametrik testler ailesi olarak tanımlanabilen istatistiksel teknikler ve bu teknikleri içeren istatistiksel modeller, Gaussian dağılım olarak da bilinen normal dağılım veya normal dağılımla ilişkilendirilmiş, fonksiyon kuralı ile tanımlı, belli bir hipotetik evren dağılımı üzerine inşa edilmiş algoritmalara sahiptir. Bu nedenle bu tür tekniklerin ve modellerin kullanımında, normallik ve çok değişkenli normallik gibi dağılım yapısına yönelik temel varsayımlar ön plana çıkmaktadır. Normallik ihlalleri, yanlı kestirimler elde edilmesine, Tip 1 hata ve kestirim standart hatasında aşırı yükselmelere yol açabilmektedir (Field, 2013; Stevens, 2009; Tabachnick ve Fidell, 2013). Normallik ihlallerinin önemli gerekçeleri arasında ise uç değer etkisiyle ortaya çıkan çarpıklık vurgusu ön plana çıkmaktadır.

Kullanılacak istatistiksel tekniğin tek değişkenli (univariate) veya çok değişkenli (multivariate) olmasına bağlı olarak, parametrik test kullanımında, *tek değişkenli normal dağılım* ve *çok değişkenli normal dağılım* varsayımları tanımlanabilmektedir. Tek değişkenli normal dağılım

varsayımı, örneklemdeki her sürekli değişkenin ölçümlerinin normal dağılımdan gelmesini, çok değişkenli normal dağılım ise birden fazla değişkenin aynı anda normallik koşulunu sağlamasını (Mardia,1974) vurgulamaktadır. Çok değişkenli normallik, örneklemdeki her bir değişkenin tek değişkenli normallik şartını karşılamaşının yanı sıra, değişkenlerin tüm doğrusal kombinasyonlarının ve her türlü ikili kombinasyonlarının da her bir grup ve alt grup için normal dağılım gösterdiğini varsayar.

Korelasyon, doğrusal regresyon, t-testleri, Varyans Analizi (ANOVA) gibi birçok istatistiksel yöntem normal dağılıma dayanmaktadır. Bunların yanı sıra, Çok Değişkenli Varyans Analizi (MANOVA), Temel Bileşenler Analizi (PCA), kanonik korelasyon analizi, diskriminant fonksiyon analizi gibi birçok çok değişkenli istatistiksel teknik, çok değişkenli normallik varsayımına dayanmaktadır (Johnson ve Wichern, 2007). Genel olarak değerlendirildiğinde; parametrik, yani belli bir hipotetik evren dağılımı öngören hipotez testleri ve modellerin tamamına yakını, normal dağılım biçimlerini esas almaktadır. Doğrusal modeller ailesi veya daha geniş bir tanımlama ile genelleştirilmiş doğrusal modeller ailesi, normallik varsayımına dayalı olan ve normallik ihlallerinden etkilenme olasılığı bulunan teknikler ve modellerin bir bütünüdür.

Normallik ihlalleri, Gaussian modeller için bir risk içermektedir. Bu tür bir ihlal, kullanılan teknik veya model ile elde edilen kestirimlerde sistematik hataya ve yanlılığa yol açabilmekte, analiz etki gücünü düşürebilmektedir. Dahası, Tip 1 hata düzeyinde aşırı yükselmeye ve buna bağlı olarak hatalı bir şekilde manidar fark gözlemlemeye yol açma olasılığı bulunmaktadır. Diğer taraftan ANOVA, ANCOVA, doğrusal regresyon gibi parametrik tek değişkenli tekniklerin ve özellikle MANOVA, diskriminant fonksiyon analizi gibi çok değişkenli parametrik tekniklerin, ciddi düzeyde olmayan normallik ihlallerine dirençli oldukları

bilinmektedir (Field, 2013; Tabachnick ve Fidell, 2013). Burada belirsizlik içeren nokta, *ciddi* düzeyde ihmalin ne olduğudur.

Normallik test edilirken betimsel istatistiklerden, çarpıklık ve basıklık katsayılarından, grafiksel gösterimlerden (histogram, normal Q-Q saçılma grafiği, kutu-çizgi grafiği vb.) ya da hipotez testlerinden (Kolmogorov-Smirnov, Shapiro-Wilk vb.) yararlanılmaktadır. Sağlıklı bir değerlendirme için, bu yöntemlerin birlikte kullanılması önerilmektedir. Ayrıca büyük örneklerde hipotez testleri Tip 1 hataya eğilimli olduğu için, grafiksel ve betimsel incelemelere ağırlık verilmesinin uygun bir tercih olacağı da belirtilmektedir (Field, 2013).

Tek değişkenli normal dağılım varsayımında olduğu gibi çok değişkenli normalliğin incelenmesi için de betimsel, grafiksel ve istatistiksel yöntemler bulunmaktadır. Çok değişkenli normal dağılımın doğrusal kombinasyonları normaldir. Dolayısıyla, çok değişkenli normalliğin değerlendirilmesine tek değişkenli normalliği ve iki değişkenli normalliği kontrol ederek başlanabilir. Çok değişkenli normalliğin incelenmesi için betimsel yöntemler olarak çok değişkenli çarpıklık ve basıklık katsayıları kullanılabilir (Mardia, 1974). Ayrıca, değişken çiftlerinin dağılım grafiği ve iki değişkenli Gama grafiği (ki-kare Q-Q grafiği) iki değişkenli normalliği kontrol etmede kullanılabilir (Johnson ve Wichern, 2007).

Gaussian modeller ya da parametrik testler için varyansların homojenliği (homoscedasticity), bir diğer varsayım olarak tanımlanmaktadır. Parametrik testlerin kullanılabilmesi için hem tek değişkenli ve hem de çok değişkenli istatistiksel tekniklerde karşılanması beklenen varyansların homojenliği varsayımı, tek değişkenli tekniklerde bir sürekli değişkene ait varyansın diğer bir sürekli değişkenin varyansı ile benzerliğini ifade ederken, bir kategorik ve bir sürekli değişkenin bulunduğu gruplanmış verilerde sürekli olan bağımlı değişkenin varyansının kategorik değişkenin tüm seviyelerinde benzer olduğunu ifade etmektedir. Her bir bağımlı değişkenin varyans homojenliğinin çok değişkenli genellemesi, varyans-kovaryans matrislerinin homojenliğini ifade etmektedir. Bu varsayım ile modelde yer alan her bir hücrenin varyans-kovaryans matrislerinin aynı genel varyans-kovaryans matrisinden seçildiği ve bu matrislerin makul bir şekilde birleştirilerek tek bir hata tahmini elde edilebileceği varsayılmaktadır (Stevens, 2009; Tabachnick ve Fidell, 2013).

T-testi, faktöriyel ve tekrarlı varyans analizi, doğrusal regresyon analizi gibi tek değişkenli parametrik testlerin yanı sıra MANOVA, MANCOVA, diskriminat fonksiyon analizi, faktör analizi gibi çok değişkenli tekniklerde de varyansların homojenliği varsayımının karşılanması

gerektiği belirtilmektedir (Johnson ve Wichern, 2007; Stevens, 2009; Thompson, 2006). Sadece gruplar arası farkların veya grup içi boylamsal farkların test edildiği hipotez testlerinde değil, aynı zamanda model parametre kestirimlerinde de varyansların homojenliği varsayımının karşılanması beklenmektedir. İlk durum için varyansların homojenliği, gruplar arası çeşitliliğinin benzerliğini, ikinci durum için ise model parametrelerine yönelik model içi hata varyanslarının benzerliğini vurgulamaktadır.

Tabachnick ve Fidell (2013), tek değişkenli istatistiksel tekniklerde, büyük varyansın küçük örneklem büyüklüğü ile ilişkilendirildiği durumlarda, Tip 1 hata oranının arttığını belirtmektedir. Varyans homojenliğinin ihlal edildiği bu tür durumlar (heteroscedasticity), yanlış pozitif sonuçların görülme olasılığını arttırmaktadır. Benzer şekilde, varyans-kovaryans matrislerinin homojenliğinin ihlali de Tip 1 hata görülme oranını yükseltmekte ve analizin gücünü olumsuz etkilemektedir (Stevens, 2009). Model parametre kestiriminde ise, varyansların homojenliğinin ihlali, standart hatanın tahmininde yanlış ve tutarsız sonuçlara yol açmaktadır. Dahası, standart hataya dayalı olarak elde edilen güven aralıkları ve anlamlılık testlerinin yanlış sonuçlar ortaya koyacağı belirtilmiştir (Field, 2013). Dolayısıyla, varyansların homojenliği varsayımının ihlali, istatistiksel analizlerde yanlış sonuçlar doğurarak, araştırma bulgularının güvenilirliğini ve geçerliğini tehlikeye atmaktadır.

Normalliğin test edilmesinde olduğu gibi varyansların homojenliğinin test edilmesinde kullanılacak en uygun yöntemin seçilmesi de mevcut koşullara ve istatistiksel analizlerin özelliklerine bağlıdır. Tek değişkenli istatistiksel testlerde; Levene testi, Bartlett testi, Hartley'in Fmax testi, Brown-Forsythe testi, Fligner-Killeen testi gibi testler kullanılabilir. Çok değişkenli hipotez testlerinde ise sık kullanılan yöntemlerden biri Box's M testidir. Tüm bu testler genel olarak örneklem büyüklüğüne duyarlıdır (Field, 2013; Stevens, 2009; Tabachnick ve Fidell, 2013).

Ciddi düzeyde normallik ve homojenlik ihlallerinin görülmemesi, istatistiksel gücü daha yüksek olan parametrik testlerin veya doğrusal modellerin kullanılması yönünde karar alınmasını pekiştirir. Diğer taraftan normallik ve homojenlik varsayımları ihlali için *ciddi düzeyin* ne olduğu, araştırmacıların çok yönlü çalıştığı konulardan birisidir. Bu yönde zengin bir literatür oluşmasına rağmen belirsizliğin sürdüğü görülmektedir. Bu kapsamda bu çalışmanın amacı; uygun istatistiksel tekniğin belirlenmesinde normallik ve homojenlik ihlallerinin hangi durum ve koşullarda *ciddi* veya *ihmal edilebilir* olarak değerlendirilebileceğini, ayrıca bu tür durumlarda ne tür çözümlere başvurulabileceğini, ilgili araştırmalar ve temel

kaynak kitaplar üzerinden tartışmak ve değerlendirmektir. Bu amaç doğrultusunda aşağıdaki sorulara yanıtlar aranmıştır.

1. Normallik ve homojenlik ihlallerinin gerekçeleri ve olası kaynakları nelerdir? Bu tür ihlallerde kullanılacak alternatif yöntemler ve çözümler nelerdir?
2. Hangi durum ve koşullarda normallik ve homojenlik ihlalleri *ciddi* veya *ihmal edilebilir* olarak değerlendirilebilir?
3. Normallik ve homojenlik arasında bir ilişki var mıdır?

### Yöntem

Bu çalışma, ilgili literatürdeki temel kaynak kitaplar ve ilişkili araştırmaların ortaya koyduğu bulgular üzerinden yürütülen derleme niteliğinde kuramsal bir çalışmadır. Temel kaynak kitaplar olarak alanda tanınırlığı yüksek olan ve ders kitabı olarak da birçok üniversitede okutulan Tabachnick ve Fidell (2013), Field (2013), Pallant (2016), Hair ve diğerleri (2010), Howell (2010), Cohen ve Swerdlik (2010), Stevens (2009), Wilcox (2009) gibi kitaplar dikkate alınmıştır. İlgili araştırma makalelerine ise WoS, Scopus, ERIC ve Jstor veri tabanlarında *normality violation*, *skewness*, *heterocedasticity*, *homoscedasticity*, *unbalanced design* gibi anahtar sözcüklerle tabakalı bir şekilde süzgeçler kullanılarak ulaşılmıştır. Bu kapsamda 20 temel kaynak kitap ve 1973 ile 2023 yılları arasında yayınlanmış 25 makale belirlenmiştir.

Belirlenen temel kaynak kitaplar, araştırma soruları doğrultusunda tartışmaların yürütülmesinde ve literatürdeki genel kuramsal bilginin sentezlenmesi ve değerlendirilmesinde kullanılmıştır. Araştırma makaleleri ise pratikte normallik ve homojenlik varsayımlarının farklı istatistiksel teknikler üzerindeki önemini ve etkilerini inceleyen çalışmalardır. Bu makaleler, normallik ve homojenlik ihlallerinin, istatistiksel teknikler üzerindeki pratik etkilerini değerlendirmek amacıyla ampirik veri kaynağı olarak kullanılmıştır. Her bir araştırma sorusuna yönelik olarak ilgili kaynaklardaki bilgiler ve açıklamalar doğrultusunda kuramsal tartışmalar yürütülmüştür.

### Tartışma

#### Normallik İhlallerinin Kaynakları

Normallik ihlalleri birçok farklı şekilde karşımıza çıkmaktadır. Uç değer etkisi bunlardan biridir. Bir veri kümesinde gözlemlerin çoğu tarafından önerilen model ile uyuşmayan gözlemlere uç değer denilmektedir. Uç değerleri belirlemede en basit ve etkili yöntemlerden biri, veri noktalarını Z standart puanına dönüştürmek ve en

düşük ile en yüksek değerleri görüntülemektir (Douzenis ve Rakow, 1987). Uç değerlerin etkisiyle veride sağa veya sola çarpıklık oluşur. Bu da normallığı etkiler. Uç değer etkisi veri dönüşümü (karekök dönüşümü, ters dönüşüm, logaritmik dönüşüm, arc-sinüs dönüşümü vb.) uç değer uyumunun sağlanması ve uç değerlerin silinmesi ile en aza indirilebilir. Uç değerlerin birçok istatistiksel tekniğin kullanımını olumsuz etkilediği görülmektedir. Örneğin Zumbo ve Jennings (2002), uç değerlerin güvenilirlik kestirimlerinde yanlılığa yol açabildiğini göstermiştir.

Normallik ihlallerinin bir diğer olası kaynağı örneklem büyüklüğü olabilmektedir. Yeterli büyüklükte olmayan örneklemelerde, dağılım yapısının durağanlığını sağlamak güç olmaktadır ve bu, normallik ihlali risklerini yükseltebilmektedir. Yeterli büyüklükte örneklemelerde, merkezi limit teoreminin de doğal bir sonucu olarak, normallik ihlalinin ciddi sorunlara yol açmayacağı değerlendirilebilir. Örneklem büyüklüğü ile normallik ilişkisini inceleyen birçok araştırma bulunmaktadır. Örneğin Pek ve diğerleri (2018), lisans ve lisansüstü istatistik ders kitaplarının, normallik dışı durumlarla başa çıkmada genellikle veri dönüşümlerini önerdiğini, ayrıca Merkezi Limit Teoremi nedeniyle klasik parametrik testlerin sağlam olduğunu ve bu tür ihlallerden fazlaca etkilenmediklerini belirttiklerini belirlemiştir. Davey ve Savla (2010) ise yeterli örneklem büyüklüğü bulunsa da kayıp veri oranlarına dikkat edilmesi gerektiğini belirtmektedir. Çünkü kayıp veri, örneklem evreni temsil edilebilirliğini azaltmakta, normallik ihlallerinin olası etkilerini artırmaktadır. DeCarlo (1997), normallik ihlallerine yol açan temel sorunun, verilerin normal dağılımın kuyruklarında veya ortasında aşırı yoğunlaşmasına bağlı olarak ortaya çıkabilecek aşırı basıklık ve aşırı çarpıklık olduğunu belirtmektedir. Kish (1965) de örneklemde olası bir ranj daralmasının, yani örneklem seçimi sırasında belirli grupların dışlanması veya örneklem aralığının daralmasının, normallik ihlallerine yol açabilecek bir durum olduğunu belirtmektedir.

Normallik ihlallerinin kaynakları arasında uç değer etkisine bağlı çarpıklık vurgusu ön plana çıkmaktadır. Özellikle küçük örneklemelerde uç değer etkisinin normallik ihlallerine yol açma olasılığı yüksektir. Bunun yanı sıra değişkenlerin ölçek yapısının sürekliliğe uygun olmaması, verilerdeki veya veri aralıklarındaki yığılmalar, ranj daralması, kayıp veriler gibi durumlar da verilerde normallik ihlallerine yol açabilmektedir.

#### Varyans Homojenliğinin İhlalinin Kaynakları

Varyansların homojenliği varsayımının ihlaline yol açan çeşitli etkenler bulunmaktadır. Bu etkenler arasında örneklem büyüklüğü, grup büyüklüklerinin dengesi,



değişkenler arası ilişkiler ve veri yapısının özellikleri önemli rol oynamaktadır. Bununla birlikte, değişkenler arası ilişkilerin gücü ve düzeyi, veri yapısındaki doğrusal olmayan değişim ve gelişim varyansların homojenliğini etkilemektedir.

Grup dengesizliği, yani grup büyüklüklerinin eşit olmaması durumunda, büyük grupların küçük gruplara göre daha yüksek varyansa sahip olma olasılığı artmaktadır (Tabachnick ve Fidell, 2013). Bu durum, özellikle iki veya daha fazla grubun karşılaştırıldığı analizlerde belirgin hale gelmektedir. Eşit olmayan örneklem büyüklükleri, varyansların homojenliği varsayımının ihlaline ve dolayısıyla Tip 1 hata oranının artmasına neden olabilmektedir. Aynı şekilde; Howell (2010), örneklem büyüklüklerinin farklılaştığı ve büyük grupların büyük varyansa sahip olduğu durumlarda, varyansların homojenliğinin sağlanmasının zorlaştığını ve bu durumun istatistiksel testlerin güvenilirliğini olumsuz etkilediğini vurgulamaktadır. Bağımsız değişkenin bazı seviyelerinde daha büyük ölçüm hataları olmasına bağlı olarak görülen grup dengesizlikleri de bu seviyelerde varyansın daha büyük olmasına sebep olduğundan varyansların homojenliği varsayımının ihlaline yol açmaktadır (Tabachnick ve Fidell, 2013). Bu nedenle, grup dengesizliği durumlarında uygun düzeltme yöntemlerinin kullanılması veya alternatif testlerin tercih edilmesi, analiz sonuçlarının doğruluğunu artırmak için gereklidir.

Örneklem büyüklüğü, diğer bir faktördür. Küçük örneklemelerde, varyansların homojenliği varsayımının sağlanması genellikle daha zordur. Örneğin, 10 ile 20 kişi içeren küçük örneklem gruplarında, varyans tahminleri daha fazla belirsizlik taşımakta ve bu durum, varyansların homojenliği varsayımının ihlaline yol açabilmektedir (Kirk, 1995). Öte yandan 30 ila 50 kişi içeren orta büyüklükteki örneklem büyüklükleri ise, varyans tahminlerinde daha az belirsizlik sunmaktadır ve bu büyüklükteki örneklem genellikle varyansların homojenliği varsayımını daha iyi karşılamaktadır (Cohen ve Lea, 2004). Örneklem büyüklüğünün artması, örneğin 100'ün üzerindeki örneklem, varyans farklılıklarının daha hassas bir şekilde tespit edilmesine imkân tanımaktadır, fakat bu aynı zamanda küçük varyans farklılıklarının da istatistiksel olarak anlamlı hale gelmesine neden olabilmektedir (Wilcox, 2012). Büyük örneklemelerde istatistiksel testler, daha fazla yanlış pozitif sonuç üretebilmektedir (Field, 2013). Küçük örneklem genellikle varyansların homojenliği varsayımının karşılanamamasına yol açarken, çok büyük örneklem de varyans farklılıklarını anlamlı kılabilirdiğinden, her iki durumda da uygun istatistiksel yöntemlerin ve testlerin seçilmesi kolay değildir.

Örneklem yapısı ve örnekleme biçimi, varyansların homojenliği varsayımının ihlaline yol açabilen diğer bir etkidir. Seçkisiz olmayan örnekleme yöntemlerinin kullanılması durumunda, ranj daralması ve buna bağlı olarak varsayımsal ihlaller ortaya çıkabilmektedir. Özellikle, küme örnekleme gibi yöntemler, gruplar arasında sistematik varyans farklılıklarına neden olmaktadır (Kish, 1965). Grup büyüklüklerinin ve örnekleme yönteminin dikkatli bir şekilde belirlenmemesi, bu tür sorunların ortaya çıkmasına neden olmakta ve bu durum istatistiksel analizlerde yanlış sonuçlar üretilmesine yol açmaktadır (Howell, 2010). Özellikle, örnekleme yönteminin varyanslar üzerindeki etkisinin incelenmesi ve uygun yöntemlerin seçilmesi, istatistiksel analizlerin doğruluğunu ve güvenilirliğini artırmaktadır (Cochran, 1977)

Tekrarlı ölçümlerde gruplarda doğrusal olmayan değişim ve gelişim, varyansların homojenliği varsayımının ihlaline yol açan faktörlerden biridir. Gruplardan birinde hızlı bir başlangıç düzeyi, ardından yavaşlama ve sonra tekrar hızlanma gibi doğrusal olmayan bir değişim gözlemlenmesi halinde, bu grup içerisindeki varyanslar ölçümler arasında değişiklik gösterebilmektedir. Özel tanımlamasıyla *küresellik (sphericity)* olarak tanımlanabilen bu boylamsal varyans değişiminin durağan olmaması, varyansların homojenliği varsayımının da ihlali anlamına gelmektedir. Keselman ve diğerleri (2001), tekrarlı ölçümlerde varyansların homojenliği varsayımının doğrusal olmayan değişimlerden nasıl etkilendiğini ve bu ihlalin analiz sonuçları üzerindeki olumsuz etkilerini vurgulamıştır. Box (1950) ise doğrusal olmayan değişimlerin varyanslar arasındaki tutarlılığı bozarak homojenliği ihlal edebileceğini belirtmiştir. Benzer şekilde, Maxwell ve Delany (2004) de tekrarlı ölçümlerde zamanla ortaya çıkan doğrusal olmayan değişimlerin varyansların homojenliğini bozabileceğini ve bu durumun analizlerde yanlış sonuçlara yol açabileceğini ifade etmişlerdir. Özellikle uzun dönemli çalışmalar, bireylerin zaman içindeki değişim ve gelişimlerinin doğrusal olmaması durumunda, varyansların homojenliği varsayımının ihlal edilmesine daha yatkın olabilmektedir (Tabachnick ve Fidell, 2013). Bu nedenle, tekrarlı ölçümlerde doğrusal olmayan değişimlerin varlığı, varyansların homojenliği varsayımının dikkatle değerlendirilmesini gerektirmektedir.

Çok değişkenli veya çoklu modellerde, değişkenler arası ilişkilerin gücü ve düzeyi varyansların homojenliği varsayımının ihlaline yol açan faktörlerden bir diğeridir. Bu durum, özellikle modeldeki etkileşimler ve çoklu ilişkiler kompleks hale geldiğinde belirginleşmektedir. Değişkenler arasındaki güçlü ve yüksek derecede korelasyonlar, hata terimlerinin varyansının düzensiz hale gelmesine yol açmakta, bu da varyansların homojenliği varsayımının

sağlanmasını zorlaştırmaktadır (Hair vd., 2010). Ayrıca, çoklu regresyon ve faktör analizi gibi çok değişkenli analizlerde, değişkenler arası etkileşimlerin ve çapraz etkilerin, varyansların dağılımında düzensizliklere neden olduğu gözlemlenmektedir (Field, 2013; Tabachnick ve Fidell, 2013). Bu etkileşimler, bazı değişkenlerin diğerlerinden farklı varyanslar göstermesine neden olabilmekte ve bu durum, modelin doğruluğunu ve güvenilirliğini etkilemektedir (Kline, 2015). Özellikle küçük örneklerde, değişkenler arasındaki yüksek korelasyonlar ve etkileşimler, modelin, varyansları düzgün bir şekilde tahmin etmesini daha da zorlaştırmakta, bu da varyansların homojenliği varsayımının ihlaline yol açmaktadır (Bentler, 1990). Bu bağlamda, modelin karmaşıklığı ve değişkenler arasındaki ilişkilerin gücü, varyansların homojenliği varsayımının geçerliliğini etkilemektedir ve bu durumun doğru bir şekilde değerlendirilmesi, modelin geçerliliği ve sonuçlarının güvenilirliği açısından kritik öneme sahiptir.

Homojenlik ihlallerinin kaynakları arasında, gruplar arası oransal dengesizliğe ve örneklem büyüklüğüne yapılan vurgu ön plana çıkmaktadır. Willcox (2009), karşılaştırma yapılan grupların normal dağılım gösterdiği, grup büyüklüklerinin eşit ve çok küçük olmadığı (örneğin 8'den küçük) ve grup dağılımlarının benzer özellik gösterdiği (örneğin aynı düzeyde ve yönde çarpıklık) durumlarda, varyansların homojenliği varsayımının ihlalinin Tip 1 hata görülme olasılığında artışa ve güven aralığı tahmininde hatalı sonuçlara yol açmadığını belirtmektedir. Benzer şekilde Howell (2010) da ilgilenilen dağılımların benzer özellik gösterdiği ve en büyük varyansın en küçük varyansa oranının dört katından daha büyük olmadığı durumlarda parametrik testlerin kullanılabilirliğini belirtmektedir.

### **Normallik ve Homojenlik ihlallerinde Alternatif Yöntemler**

Normallik veya varyansların homojenliği varsayımının ihlal edildiği durumlarda uygulanabilecek alternatif ve benzer yöntemler mevcuttur. Bu bağlamda, en sıklıkla tercih edilen yöntemlerden biri veri dönüştürme (data transformation) tekniğidir. Veri dönüştürme, orijinal verilerin belirlenen bir fonksiyonla dönüştürülmesi ve bu şekilde varyansın daha homojen hale getirilmesini amaçlayan bir işlemdir (Tabachnick ve Fidell, 2013). Logaritmik dönüşüm, karekök dönüşümü, ters dönüşüm ve Box-Cox dönüşümü gibi veri dönüştürme teknikleri varsayım ihlallerinin görüldüğü durumlarda varyansların dengelenmesi amacıyla kullanılabilir. Veri dönüştürmenin göz önünde bulundurulması gereken önemli bir sınırlılığı, dönüştürülmüş verilerle elde edilen analiz sonuçlarının yorumlanma güçlüğüdür. Bunun için analiz sonuçlarında bir geri dönüştürmeye ihtiyaç duyulabilmektedir.

Kullanılabilecek diğer bir yaygın yöntem uç değerler ayıklama veya uç değerler yerine atama yapmaktır. Uç değerler, veri setindeki aşırı veya olağandışı değerlerdir ve bu değerler analiz sonuçlarında ciddi etkilere yol açabilmektedir. Bu tür durumlarda, veri setindeki uç değerlerin ayıklanması veya bunların yerine uygun değerlerin atanması hem çarpıklığın düzeltilmesi hem varyansların homojenliğinin dengelenmesinde etkili olabilmektedir (Field, 2013; Tabachnick ve Fidell, 2013). Uç değerlerin ayıklanması, çeşitli şekillerde gerçekleştirilebilmektedir. En yaygın yöntemlerden biri, z-puanı veya standart sapma kullanarak uç değerlerin belirlenmesi ve veri setinden çıkarılmasıdır. Ayrıca kutu-çizgi grafiği gibi görselleştirmelerden de uç değerlerin belirlenmesinde yararlanılabilmektedir (Tukey, 1977). Uç değerler yerine atama işlemi de çeşitli şekillerde gerçekleştirilebilir. Sıklıkla kullanılan yöntemlerden biri, uç değerlerin ortalama veya medyan değerleriyle değiştirilmesidir (Osborne ve Overbay, 2004). Gerek uç değer atama gerekse uç değer ayıklama, kuramsal dayanakları ile birlikte rasyonel gerekçelere dayandırılarak yapılması gereken işlemlerdir. Aksi durumda örnekleme bir sistematik daralmaya ve analiz sonuçlarında yanlışlık risklerine yol açabilmektedir. Mümkünse uç değer yapısının ayrı analizlerle incelenmesi de önerilmektedir (Field, 2013; Howell, 2010).

*Dirençli* (robust) tekniklerin kullanımı etkili olan başka bir yöntemdir. Bu teknikler, verideki uç değerlerin ve heterojen varyansların etkisini azaltarak, analiz sonuçlarının daha doğru ve güvenilir olmasını sağlamaktadır (Field, 2013). *Dirençli* teknikler, klasik parametrik test varsayımlarını karşılamayan veri setlerinde bile geçerli sonuçlar elde etmek için tasarlanmıştır. Örneğin, *dirençli* regresyon teknikleri, uç değerlerin ve heterojen varyansların etkisini azaltarak, model tahminlerinin doğruluğunu arttırmaktadır. Bu teknikler, veri setindeki uç değerlerin ve heterojen varyansların neden olduğu yanlışlıkları minimize etmektedir (Rousseeuw ve Leroy, 1987). Ayrıca, Welch's t-testi, Brown-Forsty düzeltmeli varyans analizi, faktör analizinde temel eksenler faktörlemesi (principal axis factoring) gibi *dirençli* analiz teknikleri, grup varyanslarının veya hata varyanslarının eşit olmadığı durumlarda daha güvenilir sonuçlar sağlamaktadır (Delacre vd., 2017).

Varsayım ihlalleri ile başa çıkmada dikkate alınabilecek bir diğer çözüm Bayesçi tekniklerin kullanılmasıdır. Bayesçi analizler, klasik frekansçı yaklaşımlardan farklı olarak, veriye dair öncül bilgilerle birlikte analiz yapma esnekliği sağlamaktadır. Bayesçi yöntemler, verinin tamamı üzerinden çıkarım yaparak, posterior dağılımlar aracılığıyla parametre tahminlerinin belirsizliklerini daha doğru bir şekilde yakalamaktadır (Gelman vd., 2013). Bu esneklik,

özellikle küçük örneklerde ve karmaşık veri yapılarında büyük avantaj sağlamaktadır (Kruschke, 2014). Bayeşçi yaklaşımlar, öncül tanımlamalarla görece bir kestirim titizliği ve analiz gücünde artış sağlayabilmekle birlikte, bu öncül tanımlamaların niteliği, uygunluğu ve yeterliğine dikkat edilmesi gerekmektedir. Ayrıca öncül tanımlamaların da analize dahil edilmesi, işlem yükü ve süresini ciddi düzeyde artırabilmektedir.

Bir diğer alternatif çözüm olarak *bootstrap* ve *jackknife* gibi yeniden örnekleme teknikleri, veri dağılımına yönelik güçlü varsayımlar gerektirmeden güvenilir kestirimler yapma olanağı sunmaktadır. Bootstrap, orijinal veri setinden tekrar tekrar örneklem alarak, her bir örneklem için istatistiksel analizler gerçekleştirip bu analizlerden elde edilen çıktıların dağılımını kullanarak güven aralıklarını hesaplamaktadır (Efron ve Tibshirani, 1993). Bu yöntem, varyansların homojenliği varsayımının ihlal edildiği durumlarda bile güvenilir sonuçlar sağlamak ve özellikle küçük örneklerde etkili olmaktadır. Diğer taraftan, yeniden örnekleme tekniklerinin kullanımında, mevcut gözlemlerin uygunluğu ve yeterliği, dikkate alınması gereken bir koşuldur. Yeterli çeşitlilikte ve temsil edebilirlikte olmayan gözlenen veriler üzerinde yapılacak yeniden örneklemler, ancak mevcut çeşitliliğin benzerlerinin tekrar üretilmesini sağlayarak analizlerde bir dış güvenilirlik ve geçerlik sorunu oluşturabilmektedir.

Son olarak, dağılımdan bağımsız tekniklerin kullanılması ise, verinin belirli bir dağılımı takip etmesi gerekliliğini ortadan kaldırarak analizlerin daha esnek olmasını sağlamaktadır. Bu yöntemler, parametrik olmayan istatistikler ve bazı dirençli (robust) teknikleri içerir. Bu tür teknikler, özellikle normallik ihlali durumlarında, diğer çözümler işe yaramadığında veya uygulanabilir olmadığında, tercih edilebilir yöntemler olarak değerlendirilebilir (Wilcox, 2012). Örneğin bağımsız örneklemler t testi için varsayımsal gereklilikler karşılanmadığında, dirençli testlerden Welch'in t testinin veya parametrik olmayan testlerden Wilcoxon sıra toplamları testinin kullanılması değerlendirilebilir.

### **Normallik İhlalleri Hangi Koşullarda Ciddi Risk Oluşturmaktadır?**

Yapılan araştırmalar t-testi ve ANOVA gibi hipotez testlerinde özellikle küçük örneklerde çarpıklığa karşı hassasiyetin fazla olduğunu göstermektedir. Blanca ve diğerleri (2017), ANOVA'nın t-testine göre daha dirençli bir yöntem olduğunu belirtmektedir. Benzer şekilde Gelman ve diğerleri (2021), doğrusal regresyon modellerinin normallik ihlallerine dirençli olduğunu belirterek, bu modellerde normalliğin en az önemli varsayım olduğunu ve normallik

için test yapmanın gerekli olmadığını öne sürmektedir.

Diğer taraftan Knief ve Forstmeier (2021), regresyon modellerinde, özellikle küçük örneklerde ve özellikle aşırı çarpıklık kaynaklı normallik ihlallerinde, yordayıcı değişkenlere yönelik katsayı kestirimlerinde sapmalar ve Tip 1 hata oranlarında yükselmeler olabildiğini belirtmektedir. Benzer şekilde Razali ve Wah (2011), küçük örneklerde büyük sapmaları ve uç değerleri tespit etmek için normalliğin değerlendirilmesi gerektiğini belirtmektedir. Uç değerler hem çarpıklığı hem de basıklığı etkileyebilmektedir. Özellikle aşırı uç değerler, analiz sonuçlarının değerlendirilmesinde ciddi yanlılığa yol açabilmektedir.

Bu ve benzeri çalışmalar bir bütün olarak değerlendirildiğinde, verilerde aşırı çarpıklık veya basıklık gibi bir durum olmadığında ve yeterli örneklem büyüklüğü sağlandığında normalliğin daha az önemli hale geldiği görülmektedir. Aksi yönde değerlendirilirse, normallik ihlallerine yönelik öncelikli risk faktörlerinin; aşırı uç değer etkisinden kaynaklı çarpıklık ve örneklemin yeterli büyüklükte olmaması olduğu görülmektedir. Diğer faktörlerin yanı sıra özellikle bu iki koşul, ciddi düzeyde normallik ihlallerinin kaynakları olarak değerlendirilmektedir.

### **Varyansların Homojenliğine Yönelik İhlaller Hangi Koşullarda Ciddi Risk Oluşturmaktadır?**

Genel olarak eşit olmayan örneklem büyüklüklerine ve değişen varyanslara sahip normal dağılım gösteren grupların karşılaştırılması, Tip 1 hata olasılığını arttırmaktadır ve güven aralığının doğru tespit edilmesi açısından sorunlar oluşturmaktadır. T-testi ve ANOVA gibi görece daha küçük örneklem gruplarının karşılaştırıldığı parametrik testlerde, karşılaştırılan iki grubun ölçülen özellik açısından normal dağılım göstermesi ve grup büyüklüklerinin eşit olup 8'den büyük olması, ya da grupların aynı yönde ve düzeyde çarpıklık göstermesi durumunda, normallik ve varyansların homojenliği varsayımının ihlalinin ciddi sorunlara yol açmayacağı belirtilmektedir (Wilcox, 2009). ANOVA modelleri, görece daha güçlü modellerdir ve küçük örneklerde dahi varsayım ihlallerine karşı dirençlidirler. Bununla birlikte, dengeli olmayan (unbalanced) desenler için en iyi sonucu *yakınsak F (Approximate F)* testinin, dengeli (balanced) desenler içinse *modifiye edilmiş Brown-Forsythe (MBF)* testinin verdiği ortaya konulmuştur (Çavuş, 2024). Sonuç olarak, bu testler için genellikle 20, 30, 40, 50 civarı örneklem büyüklükleri yeterli görülmektedir (Field, 2013). Grup büyüklükleri arasında 4-5 kat üzeri fark olmadığında, homojenlik ihlalleri tolere edilebilmektedir (Blanca vd.,

2017). Özellikle küçük örneklem kullanıldığında, Levene testi ile varyans homojenliği kontrol edilebilmekte ve ihlal varsa, Welch's t-testi veya Welch's ANOVA gibi alternatif testler kullanılabilir. Araştırmalar, homojenlik ihlalinin etkilerini azaltmak için bu alternatif testlerin kullanımını önermektedir (Ruxton, 2006; Erceg-Hurn ve Mirosevich, 2008).

Doğrusal regresyon modellerinde varyansların homojenliği ihlal edildiğinde Tip 1 hatanın daha yanlış sonuçlar ürettiği ve bu yanlışlığın grup büyüklüğü arttıkça arttığı gözlemlenmiştir. Genel olarak değerlendirildiğinde, basit ikili korelasyonlar ve basit doğrusal regresyon analizleri için homojenlik ihlali, genellikle küçük fakat yeterli örneklem büyüklüklerinde (20, 30, 40, 50) büyük bir sorun teşkil etmemektedir (Dancey ve Reidy, 2007). Ancak, çoklu regresyon ve hiyerarşik regresyon gibi daha karmaşık modellerde, örneklem büyüklüğünün model parametre sayısının 8-10 katı olması önerilmekte ve 100-150 civarı örneklem büyüklüğü yeterli görülmektedir (Fox, 2015). Homojenlik ihlali durumunda, heteroskedastisite testleri kullanılabilir ve bu yöntemlerle ihlalin etkisinin azaltılabileceği belirtilmiştir (Hayes ve Cai, 2007; Long ve Ervin, 2000). Küçük örneklemde ise varyansların belirlenmesi ve homojenliğinin test edilmesi için ağırlıklandırılmış (weighted) regresyon kestirimlerinin kullanılmasıyla doğrusallık ve varyansların homojenliği ihlalinin etkilerinin azaltılabileceği belirtilmektedir (Sayago ve Asuero, 2004; Yang vd., 2019).

MANOVA, MANCOVA ve diskriminant fonksiyon analizi gibi çok değişkenli tekniklerde homojenlik varsayımının ihlali, özellikle grup büyüklükleri arasındaki farkın çok büyük olmadığı ve örneklem büyüklüğünün yeterince büyük olduğu durumlarda tolere edilebilmektedir (Hair vd., 2010). Stevens (2009), büyük örneklemde, grup büyüklükleri dengeli olmasa dahi MANOVA'nın homojenlik ihlallerine karşı daha az duyarlı olduğunu belirtmiştir. Benzer şekilde Huberty ve Morris (1989), büyük örneklem gruplarında ve homojenlik ihlallerinin çok ciddi olmadığı durumlarda MANCOVA'nın sonuçlarının güvenilir olabileceğini belirtmiştir.

Açımlayıcı ve doğrulayıcı faktör analizi ve kümeleme analizi gibi tekniklerde, varyansların homojenliği varsayımı bazı durumlarda kritik değildir ve ihlal edilmesi her zaman analiz çıktılarına ciddi şekilde etkilememektedir. Özellikle büyük örneklem grupları (örneğin, 150, 300, 400, 500) ve modelin genel uyumu göz önüne alındığında, homojenlik ihlalleri tolere edilebilmektedir (Field, 2013). İlgili araştırmalar, bu teknikler için yeterli örneklem büyüklüğü ve uygun indekslerin kullanılması durumunda homojenlik varsayımının ihlallerinin etkisinin analiz çıktılarına ciddi

sorunlara yol açmayacağını ortaya koymaktadır (Hu ve Bentler, 1999; Kline, 2015; MacCallum vd., 1999).

Normallik ihlallerine benzer şekilde, homojenlik ihlallerinde örneklem büyüklüğü, olası çarpıklığın yönü ve gruplar arası denge, istatistiksel tekniklerin dirençli hale gelmesinde etkilidir. Bu kapsamda gruplar arası karşılaştırma testlerinde 20, 30, 40, 50, 100, 150 gibi örneklem büyüklüklerinin dikkate alınmasına yönelik öneriler bulunmaktadır. Tek değişkenli tekniklerde gruplardaki 40 civarı örneklem büyüklükleri, normallik ihlallerine dirençlilik sağlayabilmektedir. Çok faktörlü ve çok değişkenli tekniklerde, görece daha yüksek örneklem büyüklüklerine ihtiyaç duyulmaktadır. Ayrıca, gruplardaki çarpıklık benzer biçimde ve yönde olduğunda, küçük örneklemde de bir yanlışlık oluşmayabilmektedir. Gruplar arası büyüklük farklarının ise genel olarak 4-5 kat üzerinde olmaması önerilmektedir. Bu iki koşul sağlandığında, normallik ve homojenlik ihlallerinin istatistiksel test çıktılarında ciddi düzeyde bir yanlışlığa veya sapmaya yol açma olasılığı düşük görülmektedir.

### Normallik ve Homojenlik İlişkisi

Normallik ve varyansların homojenliği varsayımları, istatistiksel analizlerde birbirini tamamlayan kritik unsurlardır ve bu varsayımlar çeşitli yönlerden ilişkilidir. Öncelikle; normallik varsayımının sağlanması, verilerin uç değerler ve çarpıklık içermemesi anlamına gelmektedir, bu da gruplar arasındaki varyansların daha tutarlı olmasını sağlamaktadır (Tabachnick ve Fidell, 2013). Normallik ihlal edildiğinde, verilerdeki aşırı uçlar ve çarpıklıklar, varyansların homojenliğini bozabilmekte, bu da özellikle küçük örneklem büyüklüklerinde ciddi yanlışlıklara neden olabilmektedir. Öte yandan; gruplar arasındaki çarpıklığın yönü ve düzeyi benzer olduğunda, varyanslar da benzer düzeyde kestirilebilmekte ve bu kapsamda ciddi bir yanlışlık oluşmayabilmektedir (Stevens, 2009). Bu durum, özellikle homojenlik ihlallerinin tolere edilebilir olduğu durumlar için önemli olmaktadır. Dolayısıyla, normallik ve varyansların homojenliği varsayımları, verilerin dağılım özellikleri ve örneklem yapısı ile doğrudan ilişkilidir ve bu iki varsayım birlikte değerlendirilmesi, istatistiksel analizlerin doğruluğunu artırmaktadır. Hair ve diğerleri (2010) varyansların homojenliği varsayımının sağlandığı ve sağlanmadığı iki durumu örnek göstermiştir (Şekil 1). Şekil 1 (a), normal dağılıma sahip iki değişken (V1, V2) arasında varyansların homojenliği varsayımının sağlandığını gösterirken, Şekil 1 (b)'de değişkenlerden birinin (V3) çarpıklığından dolayı bu varsayım sağlanmamıştır.

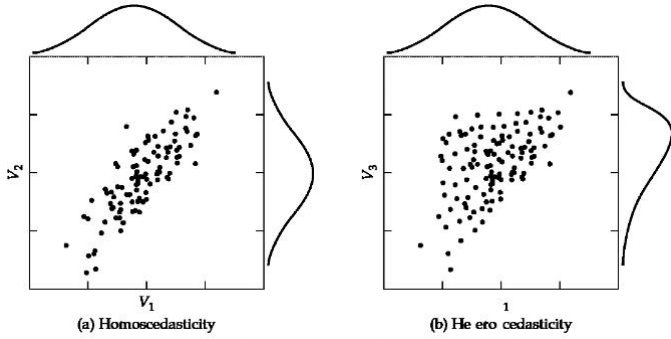
Normallik ve homojenlik arasında, yine örneklem büyüklüğüne ve dengesine bağlı olarak ilişki



kurulabilmektedir. Buna göre küçük örneklerde varyans daha yüksek, çarpık dağılımlarda ise daha düşük gözlenebilmektedir. Gruplar arasında çarpıklığın yönü ve düzeyi benzer olduğunda, varyanslar da benzer düzeyde kestirilebilmekte, bu kapsamda ciddi bir yanlılık oluşmayabilmektedir. Benzer şekilde 40 civarı ve üzeri örneklem büyüklüklerinde ve yeterli ölçek hassasiyeti sağlandığında, varyanslar arası fark ciddi düzeyde geçerlik sorunlarına yol açmayabilmektedir.

### Şekil 1

*Normal Dağılıma Sahip İki Değişkenin Varyanslarının Homojenliği (a) ve Normal ve Çarpık Dağılıma Sahip İki Değişkenin Varyanslarının Heterojenliği (b)*



Kaynak. Hair ve diğerleri, 2010, s. 74

### Sonuç ve Karar

Sonuç olarak ilgili literatür, yeterli örneklem büyüklüğü ve gruplar arası denge sağlandığında, normallik ve homojenlik ihlallerinin istatistiksel tekniklerden elde edilen kestirimlerde ciddi düzeyde bir yanlılık oluşturmayacağını ve geçerlik sorunlarına yol açmayacağını göstermektedir. Kullanılacak istatistiksel tekniğin karmaşıklık düzeyine bağlı olarak değişimle birlikte genel olarak tek değişkenli tekniklerde 40 civarı, çok değişkenli tekniklerde ise 150 civarı ve üzeri örneklem, ayrıca grup büyüklükleri arasında 4-5 katı geçmeyen oransal farklılıklar, normallik ve homojenlik ihlallerinin göz ardı edilebileceği koşullar olarak görülmektedir. Bu koşullarda parametrik teknikler veya doğrusal modeller, dirençlilik kazanmaktadır ve uygun istatistiksel teknik olarak tercih edilebilir. Aksi durumda, (1) veri dönüştürme, (2) uç değerler yerine atama, (3) uç değer ayıklama, (4) istatistiksel düzeltme ve kontrol içeren “dirençli” teknikler kullanma, (5) Bayesçi yaklaşım tabanlı analizler yürütme, (6) bootstrap ve yeniden örnekleme ile kestirimler ve bu kestirimlere yönelik güven aralıklarını belirleme, (7) dağılımdan bağımsız teknikler kullanma gibi çözümlerin değerlendirilmesi mümkündür.

Bu çalışmada ulaşılan sonuçlar, özellikle veri analizlerine dayalı araştırmalarında uygun istatistiksel tekniğe karar vermede, varsayımsal gerekliliklere bağlı olarak güçlük yaşayan araştırmacılar için pratik çözümler sağlamaktadır. Öncelikle, yukarıda açıkça belirtildiği gibi veri sayısının veya örneklem büyüklüğünün yeterli düzeyde olması ve gruplar arası büyüklük farkının yüksek olmaması durumunda, bilinen ve yaygın kullanımı olan parametrik istatistiksel tekniklerin (t-testleri, varyans analizi modelleri, faktör analizi modelleri vb.), normallik ve homojenlik ihlallerine dirençli teknikler olarak değerlendirilmesi mümkündür. Diğer varsayımsal gerekliliklerde belirgin bir ihlal olmadığı sürece bu teknik ve modellerin kullanılması tercih edilebilir. Diğer taraftan yeterli örneklem büyüklüğünün bulunmaması ve gruplar arası büyüklük farklarının yüksek olması durumunda ise özellikle dirençli yöntemlerin veya yeniden örnekleme yöntemlerinin kullanılması değerlendirilebilir. Bu çözümler kullanılabilir olmadığında, dağılımdan bağımsız parametrik olmayan testlerin kullanılması, başvurulabilecek bir diğer çözümdür. Fakat herhangi bir varsayımsal ihlal olmadığında parametrik olmayan testlerin istatistiksel gücünün parametrik testlerden daha düşük olduğu göz ardı edilmemelidir. Veri dönüştürme, uç değer ayıklama veya bunlar yerine değer atama gibi çözümler ise, normallik ve homojenlik ihlallerinde kısmen çözüm olabilese de ranj daralması veya yorumlama güçlüğü gibi bazı güçlüklere yol açabilmektedir. Bu nedenle bunların, diğer daha uygun çözümlerin işe yaramadığı son noktada değerlendirilmesi düşünülebilir.

**Hakem Değerlendirmesi:** Dış bağımsız.

**Yazar Katkıları:** Fikir-E.D.; Tasarım -E.D.; Denetleme-E.D.; Kaynaklar -E.D., İ.Ç., S.B.; Veri Toplanması ve/veya İşlenmesi -E.D., İ.Ç., S.B.; Analiz ve/veya Yorum -E.D., İ.Ç., S.B.; Literatür Taraması -E.D., İ.Ç., S.B.; Yazıyı Yazan -E.D., İ.Ç., S.B.; Eleştirel İnceleme-E.D.; Diğer-E.D.

**Çıkar Çatışması:** Yazarlar, çıkar çatışması olmadığını beyan etmiştir.

**Finansal Destek:** Yazarlar, bu çalışma için finansal destek almadığını beyan etmiştir.

**Peer-review:** Externally peer-reviewed.

**Author Contributions:** Concept-E.D.; Design-E.D.; Supervision-E.D.; Resources-E.D.; Materials-E.D., İ.Ç., S.B.; Data Collection and/or Processing-E.D., İ.Ç., S.B.; Analysis and/or Interpretation -E.D., İ.Ç., S.B.; Literature Search -E.D., İ.Ç., S.B.; Writing Manuscript -E.D., İ.Ç., S.B.; Critical Review-E.D.; Other-E.D.

**Conflict of Interest:** The authors have no conflicts of interest to declare.

**Financial Disclosure:** The author declared that this study has received no financial support.

## References

- Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, 29(4), 552-557. <https://doi.org/10.7334/psicothema2016.383>
- Çavuş, M. (2024). Comparison of one-way ANOVA tests under unequal variances in terms of median p-values. *Communications in Statistics-Simulation and Computation*, 53(4), 1619-1632.
- Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). John Wiley & Sons.
- Cohen, B. H., & Lea, R. B. (2004). *Essentials of statistics for the social and behavioral sciences*. John Wiley & Sons, Inc.
- Cohen, R. J., & Swerdlik, M. (2010). *Psychological testing and assessment: An introduction to tests and measurement*. McGraw-Hill Book Co.
- Dancey, C. P., & Reidy, J. (2007). *Statistics without math for psychology* (5th ed.). Pearson.
- Davey, A., & Savla, J. (2010). *Statistical power analysis with missing data: A structural equation modeling approach*. Routledge.
- DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, 2(3), 292-307.
- Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of student's t-test. *International Review of Social Psychology*, 30(1), 92-101.
- Douzenis, C., & Rakow, E. A. (1987) *Outliers: A potential data problem*. Mid-South Educational Research Association, Mobile, AL.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman & Hall
- Erceg-Hurn, D. M., & Mirosevich, V. M. (2008). Modern robust statistical methods: An easy way to maximize the accuracy and power of your research. *American Psychologist*, 63(7), 591-601.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage Publications.
- Fox, J. (2015). *Applied regression analysis and generalized linear models*. Sage Publications.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC Press.
- Gelman, A., Hill, J., & Vehtari, A. (2021). *Regression and other stories*. Cambridge University Press.
- Hair, J. F., Black, W. C., Babin, B. J. & Anderson, R. E. (2010). *Multivariate data analysis*. Pearson Prentice Hall.
- Hayes, A. F., & Cai, L. (2007). Using heteroscedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods*, 39(4), 709-722.
- Howell, D. C. (2010). *Statistical methods for psychology*. Wadsworth Cengage Learning.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Huberty, C. J., & Morris, J. D. (1989). Multivariate analysis versus multiple univariate analyses. *Psychological Bulletin*, 105(2), 302-308.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistical analysis* (6th ed.). Prentice-Hall.
- Keselman, H. J., Algina, J., & Kowalchuk, R. K. (2001). The analysis of repeated measures designs: A review. *British Journal of Mathematical and Statistical Psychology*, 54(1), 1-20.
- Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences* (4th ed.). Sage Publications.
- Kish, L. (1965). *Survey sampling*. John Wiley & Sons.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Knief, U., & Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods*, 53, 2576-2590. <https://doi.org/10.3758/s13428-021-01587-5>
- Kruschke, J. K. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.
- Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, 54(3), 217-224.

- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84-99.
- Mardia, K. V. (1974) Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhyā: The Indian J Statist, Series B 36*, 115-28.
- Maxwell, S. E., & Delany, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Lawrence Erlbaum Associated, Publishers.
- Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should ALWAYS check for them). *Practical Assessment, Research & Evaluation, 9*(6), 1-12.
- Pallant, J. (2016). *A step by step guide to data analysis using IBM SPSS*. McGraw Hill Education.
- Pek, J., Wong, O., & Wong, A. C. M. (2018). How to address non-normality: A taxonomy of approaches, reviewed, and illustrated. *Frontiers in Psychology, 9*, Article 2104. <https://doi.org/10.3389/fpsyg.2018.02104>
- Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 22-33.
- Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust regression and outlier detection*. Wiley.
- Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test. *Behavioral Ecology, 17*(4), 688-690.
- Sayago, A., & Asuero, A. A. (2004). Fitting straight lines with replicated observations by linear regression: Part II. Testing for homogeneity of variances. *Critical Reviews in Analytical Chemistry, 34*(3-4), 133-146.
- Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences*. Routledge.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*. Pearson.
- Thompson, B. (2006). *Foundations of behavioral statistics*. The Guilford Press.
- Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.
- Wilcox, R. R. (2009). *Basic statistics: Understanding conventional methods and modern insights*. Oxford University Press.
- Wilcox, R. R. (2012). *Introduction to robust estimation and hypothesis testing* (3rd ed.). Academic Press.
- Yang, K., Tu, J., & Chen, T. (2019). Homoscedasticity: An overlooked critical assumption for linear regression. *General Psychiatry, 32*(5) Article e100148. <https://doi.org/10.1136/gpsych-2019-100148>
- Zumbo, B. D., & Jennings, J. (2002). The robustness of validity and efficiency of the related samples t-test in the presence of outliers. *Psicologica, 23*, 415-450.