

T.C.
İSTANBUL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

KÜÇÜK ÖRNEKLEM BÜYÜKLÜKLERİNDE ÖRNEK
SAYISININ TEKRARLANAN ÖLÇÜMLÜ VERİ
ANALİZLERİ SONUCU ÜZERİNE ETKİSİ: BİR
UYGULAMA

BURCU DİKER

DANIŞMAN
DR. ÖĞR. ÜYESİ SEVDA ÖZEL YILDIZ

BİYOİSTATİSTİK VE TIP BİLİŞİMİ
TEZLİ YÜKSEK LİSANS PROGRAMI

İSTANBUL-2023

T.C.
İSTANBUL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

**KÜÇÜK ÖRNEKLEM BÜYÜKLÜKLERİNDE ÖRNEK
SAYISININ TEKRARLANAN ÖLÇÜMLÜ VERİ
ANALİZLERİ SONUCU ÜZERİNE ETKİSİ: BİR
UYGULAMA**

BURCU DİKER

DANIŞMAN
DR. ÖĞR. ÜYESİ SEVDA ÖZEL YILDIZ

BİYOİSTATİSTİK VE TIP BİLİŞİMİ
TEZLİ YÜKSEK LİSANS PROGRAMI

İSTANBUL-2023

BEYAN

Bu tez çalışmasının kendi çalışmam olduğunu, tezin planlanmasından yazımına kadar bütün safhalarda etik dışı davranışımın olmadığını, bu tezdeki bütün bilgileri akademik ve etik kurallar içinde elde ettiğimi, bu tez çalışmasıyla elde edilmeyen bütün bilgi ve yorumlara kaynak gösterdiğimi ve bu kaynakları da kaynaklar listesine aldığımı, yine bu tezin çalışılması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığı beyan ederim.

Burcu Diker

TEŞEKKÜR

Tez çalışmam ve yüksek lisans eğitimim süresince sunduğu bilgiler ve yol göstericiliği nedeniyle değerli danışman hocam Dr. Öğr. Üyesi Sevda Özel Yıldız'a ve her koşulda bana destek olan aileme teşekkür ederim.

İÇİNDEKİLER

TEZ ONAYI.....	ii
BEYAN.....	iii
TEŞEKKÜR.....	iv
İÇİNDEKİLER.....	v
TABLolar LİSTESİ.....	vii
ŞEKİLLER LİSTESİ.....	ix
SEMBOLLER/KISALTMALAR LİSTESİ.....	x
ÖZET.....	xi
ABSTRACT.....	xii
1. GİRİŞ VE AMAÇ.....	1
2. GENEL BİLGİLER	3
2.1. Tekrarlanan ölçüm dizaynı.....	3
2.2. Tekrarlanan ölçüm dizaynının avantaj ve dezavantajları.....	4
2.3. Tek değişkenli ve çok değişkenli varyans analizi yöntemleri.....	4
2.4. İki yönlü tekrarlanan ölçümler anova	7
2.5. İki faktörlü ve faktörlerden birinin seviyeleri tekrarlanan ölçüm içeren dizaynlar	8
2.6. Tekrarlanan ölçümlerde friedman testi	11
Post-hoc testler	12
2.7. Örneklem büyüklüğü.....	13
Varyans analizi için güç ve örneklem büyüklüğü	13
3. GEREÇ VE YÖNTEM.....	18
4. BULGULAR.....	21
4.1. Örneklem büyüklüğü 8 için L, a ve b verilerinin istatistiksel analizi	21
4.2. Örneklem büyüklüğü 12 için L, a ve b verilerinin istatistiksel analizi	25
4.3. Örneklem büyüklüğü 16 için L, a ve b verilerinin istatistiksel analizi	30
4.4. Örneklem büyüklüğü 20 için L, a ve b verilerinin istatistiksel analizi	35
4.5. Örneklem büyüklüğü 8 için ΔE_{00} verilerinin istatistiksel analizi.....	40
4.6. Örneklem büyüklüğü 12 için ΔE_{00} verilerinin istatistiksel analizi.....	42
4.7. Örneklem büyüklüğü 16 için ΔE_{00} verilerinin istatistiksel analizi.....	44

4.8. Örneklem büyüklüğü 20 için ΔE_{00} verilerinin istatistiksel analizi.....	45
4.9. Deneysel (post-hoc) güç analizi	47
5. TARTIŞMA.....	49
KAYNAKLAR.....	52
HAM VERİLER.....	56

TABLOLAR LİSTESİ

Tablo 2.1: Tekrarlanan ölçüme sahip iki faktörlü deneme dizaynı örneği.....	9
Tablo 4.1: Örneklem büyüklüğü 8 için her grubun tanımlayıcı istatistik sonuçları ($\bar{x} \pm S$) ve çok değişkenli normal dağılım testlerinden Mardia küçük örneklem çarpıklık testinin p değerleri.....	21
Tablo 4.2: Her grup için Mauchly küresellik testinin sonuçları.....	22
Tablo 4.3: Grup 1 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri.....	23
Tablo 4.4: Grup 2 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri.....	23
Tablo 4.5: Grup 3 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri.....	24
Tablo 4.6: Örneklem büyüklüğü 12 için her grubun tanımlayıcı istatistik sonuçları ($\bar{x} \pm S$) ve çok değişkenli normal dağılım testlerinden Mardia küçük örneklem çarpıklık testinin p değerleri.....	26
Tablo 4.7: Her grup için Mauchly küresellik testinin sonuçları.....	27
Tablo 4.8: Grup 1 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri.....	28
Tablo 4.9: Grup 2 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri.....	28
Tablo 4.10: Grup 3 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri.....	29
Tablo 4.11: Örneklem büyüklüğü 16 için her grubun tanımlayıcı istatistik sonuçları ($\bar{x} \pm S$) ve çok değişkenli normal dağılım testlerinden Mardia küçük örneklem çarpıklık testinin p değerleri.....	31
Tablo 4.12: Her grup için Mauchly küresellik testinin sonuçları.....	32
Tablo 4.13: Grup 1 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri.....	33
Tablo 4.14: Grup 2 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri.....	33
Tablo 4.15: Grup 3 için tekrarlanan ölçümlü ANOVA ve Friedman test istatistiği ve p değerleri.....	34

Tablo 4.16: Örneklem büyüklüğü 20 için her grubun tanımlayıcı istatistik sonuçları ($\bar{x} \pm S$) ve çok değişkenli normal dağılım testlerinden Mardia küçük örneklem çarpıklık testinin p değerleri.....	36
Tablo 4.17: Her grup için Mauchly küresellik testinin sonuçları.....	37
Tablo 4.18: Grup 1 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri....	38
Tablo 4.19: Grup 2 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri....	38
Tablo 4.20: Grup 3 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri....	39
Tablo 4.21: Her grubun ΔE_{00} verilerinin ortalama ve standart sapma değerleri.....	40
Tablo 4.22: Gruplar arası ikili karşılaştırmalarda ortalama fark ve p değerleri.....	41
Tablo 4.23: Her grubun ΔE_{00} verilerinin ortalama ve standart sapma değerleri.....	42
Tablo 4.24: Gruplar arası ikili karşılaştırmalarda ortalama fark ve p değerleri.....	43
Tablo 4.25: Her grubun ΔE_{00} değerlerinin tanımlayıcı istatistik verileri.....	44
Tablo 4.26: Her grubun ΔE_{00} değerlerinin tanımlayıcı istatistik verileri	46
Tablo 4.27: ΔE_{00} verilerinin analizi sonucu elde edilen güç değerleri.....	47
Tablo 4.28: L, a ve b verilerinin analizi sonucu elde edilen güç değerleri.....	48

ŞEKİLLER LİSTESİ

Şekil 4.1: Örnek sayısı 8 iken renk koordinatlarında gruba ve zamana bağlı değişim.....	25
Şekil 4.2: Örnek sayısı 12 iken renk koordinatlarında gruba ve zamana bağlı değişim..	30
Şekil 4.3: Örnek sayısı 16 iken renk koordinatlarında gruba ve zamana bağlı değişim..	35
Şekil 4.4: Örnek sayısı 20 iken renk koordinatlarında gruba ve zamana bağlı değişim..	40
Şekil 4.5: Her grubun ΔE_{00} değerindeki zamana bağlı değişim (n=8).....	42
Şekil 4.6: Her grubun ΔE_{00} değerindeki zamana bağlı değişim (n=12).....	44
Şekil 4.7: Örnek sayısı 16 iken 3 grubun $\Delta E_{00(1)}$ ve $\Delta E_{00(2)}$ değerlerinin kutu (box-plot) grafiği.....	45
Şekil 4.8: Örnek sayısı 20 iken 3 grubun $\Delta E_{00(1)}$ ve $\Delta E_{00(2)}$ değerlerinin kutu (box-plot) grafiği.....	47

SEMBOLLER / KISALTMALAR LİSTESİ

ANOVA: Varyans analizi

ε : Epsilon

KO: kareler ortalaması

DLP: Dijital ışık işleme

LCD: Sıvı kristal ekran

PMMA: Polimetilmetakrilat

η^2 : Parsiyel eta kare

$\bar{x}$: Aritmetik ortalama

S: Standart sapma

ÖZET

Diker, B. (2023). Küçük örneklem büyüklüklerinde örnek sayısının tekrarlanan ölçümlü veri analizleri sonucu üzerine etkisi: bir uygulama. İstanbul Üniversitesi Sağlık Bilimleri Enstitüsü, Biyoistatistik ABD. Biyoistatistik ve Tıp Bilişimi Programı. Yüksek Lisans Tezi. İstanbul.

Bir çalışma planlanırken hedeflerden biri evreni temsil eden örneklem seçebilmektir. Bunun koşullarından biri örneklemenin yansızlığı, diğeri ise yeterli büyüklükte örneklem seçimidir. Yeterli büyüklükten daha geniş bir örneklem seçildiğinde çalışma için harcanan süre ve maliyet artacak, özellikle laboratuvar hayvanlarının kullanıldığı çalışmalarda etik sorun oluşturabilecektir. Örneklem büyüklüğü, kullanılan istatistiksel test sonuçlarının doğruluğunu da etkileyebilir. Çok büyük örneklem, klinik olarak önemsiz olan küçük farklılıkları istatistiksel olarak anlamlı farklar olarak gösterebilir. Küçük örneklem büyüklüklerinde ise düşük etki büyüklüğü dolayısıyla Tip-2 hata artabilir. Bu nedenle evreni temsil edebilecek optimum örneklem büyüklüğünü saptamak önemlidir.

Bu çalışmada aynı evrenden farklı örnek sayıları (8, 12, 16, 20) ile oluşturulan küçük örneklem büyüklüklerinde tekrarlanan ölçümlü veri analizlerinin sonuçları karşılaştırıldı. Uygulama, diş hekimliği alanında güncel olarak kullanılan 3-boyutlu kazıma cihazı ve iki yazıcı ile üretilen geçici kron materyalleri üzerinden yapıldı. Elde edilen üç materyalin başlangıç, 2500 ve 5000 siklus termal yaşlandırma sonrası üç farklı zamanda (T0,T1,T2) renk ölçümleri yapıldı. Renk değerleri L, a ve b renk koordinatları ile, iki zaman arası renk değişimi ise ΔE_{00} hesaplaması ile kaydedildi. Her materyal grubunda $\Delta E_{00}(1)$ ve $\Delta E_{00}(2)$ arasında anlamlı farklılık varken örnek sayısına bağlı bir farklılık gözlenmedi. Yazıcı ile üretilen 2 grupta örnek sayısı 8, 12 ve 16 olduğunda b1-b2 değerleri arasında istatistiksel bir anlamlılık bulunmazken örnek sayısı 20 olduğunda bu değerler arasında anlamlı bir fark gözlemlendi. Kazıma grubunda ise örnek sayısı 8 ve 16 iken anlamlı fark bulunmayan a1-a2 arasında örnek sayısı 12 ve 20 iken istatistiksel anlamlılık bulundu. Diğer sonuçlar örneklem büyüklüğünden etkilenmedi. Tüm test istatistiklerinde 0,95'ten daha yüksek güce ulaşıldı.

Anahtar Kelimeler: Tekrarlanan ölçümler; Varyans analizi; Örneklem; Biyoistatistik; Parametrik testler

ABSTRACT

Diker, B. (2023). The effect of sample size on the results of repeated measured data analysis in small sample sizes: An application. Istanbul University, Institute of Health Science, Department of Biostatistics. Biostatistics and Medical Informatics Program. Master's Degree Thesis. Istanbul.

Selecting a sample that represents the population is one of the purposes when planning a study. The conditions for this are the impartiality of sampling and the selection of a sufficient sample size. Larger sample size will increase time and cost of the study. Sample size can also affect statistical test results. Very large samples may show insignificant differences as statistically significant differences. In small sample size, Type-2 error may increase due to the low effect size. It is important to determine the optimum sample size.

In this study, the results of repeated measurement data analyzes in small sample size (n:8,n:12,n:16,n:20) selected from the same population were compared. The application was made on the temporary crown materials fabricated with a 3D-milling machine and two printers currently used in the field of dentistry. Color measurements were made at three different times (initial,2500 and 5000 thermal-aging). Color values (L,a,b) and color changes(ΔE_{00}) were recorded. While there was a significant difference between $\Delta E_{00}(1)-\Delta E_{00}(2)$ in each material group, no difference was observed depending on the sample size. While there was no statistical significance between b1-b2 in the printer groups in sample sizes of 8, 12 and 16, a significant difference was observed when the sample size was 20. In the milling group, there was no significant difference between a1-a2 in sample sizes of 8, 16 but statistical significance was found when the sample sizes were 12 and 20. Other results were not affected by sample size. Power was higher than 0.95 in all tests.

Key words: Repeated measures; Variance analysis; Sample; Biostatistics; Parametric tests

1. GİRİŞ VE AMAÇ

İstatistik analizinde, t-test iki popülasyondan elde edilen gözlemleri karşılaştırmak için kullanılır. Bununla birlikte bir popülasyondaki iki grubun gözlemlerinin ortalamasının eşit olup olmadığı da t-testi ile analiz edilebilir. İki'den fazla popülasyon ya da grup varlığı söz konusu olduğunda ise varyans analizleri (ANOVA) kullanılır.¹ Prospektif olarak yapılan bazı deneysel çalışmalarda aynı örnek üzerinde iki ya da daha fazla zaman diliminde ölçüm yapılır. Bu durumda grup ortalamalarını karşılaştırmak için standart ANOVA yöntemini kullanmak, analizdeki denekler içi gözlemler arasındaki bağımlılıkları dikkate almadığı için doğru sonuca ulaştırmaz. Bu tür çalışmalarda, katı analitik varsayımların yerine getirilmesi ve belirli analitik prosedürlerin takip edilmesi gereken tekrarlanan ölçümlü ANOVA kullanılır.² Tekrarlanan ölçümler analizi, zaman içindeki değişimleri değerlendirmek için kullanılır. Bu tasarım özellikle yüksek oranda gözlenen bireysel farklılıkları kontrol ettiği için güçlüdür. Bu farklılıklardan kaynaklanan değişkenlik, hata teriminden elimine edilir ve tedavi seviyeleri veya zamana bağlı önemli farklılıkları gözlemleme şansı artar. Tekrarlanan ölçüm tasarımları üç farklı durumda yararlıdır. Birincisi zaman içerisinde ikiden fazla ölçüm olduğu durumlar, ikincisi üç farklı ilaç dozu verilen hastalar gibi aynı tedavinin farklı aşamalarındaki ölçümler ve üçüncüsü tedavi öncesi ve sonrası iki veya daha fazla çalışma grubunda birden fazla ölçümün yapıldığı durumlardır.³

Sağlık alanında hem klinik hem de deneysel çalışmalarda tekrarlı ölçümler yani boylamsal veriler önemli bir yer kaplar. Diş hekimliği alanında da özellikle yeni piyasaya sürülmüş materyallerin fiziksel ve mekanik özelliklerinin test edilmesinde tekrarlanan ölçümlerin yapıldığı deneysel çalışmalar sıklıkla yapılmaktadır. Literatür incelendiğinde küçük örneklem büyüklüklerinde tekrarlanan ölçümler için hem parametrik hem de parametrik olmayan testlerin kullanıldığı gözlemlenmiştir. Parametrik olmayan yöntemler, popülasyon hakkında daha az varsayım gerektirir.⁴ Zimmerman ve Zumbo⁵ normal, karışık-normal, Laplace ve Cauchy dağılıma sahip örneklerde parametrik olmayan testlerin etkilerini incelemiş ve Friedman testinin, Cauchy dağılımı gibi çok çarpık dağılımlar için parametrik testten daha güçlü olduğu sonucuna varmıştır. Bunun nedeni, parametrik olmayan testlerin aykırı değerlere karşı duyarsız olması ve parametrik

modelin varsayımları sağlanmıyorsa parametrik yöntemlerden daha güçlü olmasıdır. Bu nedenle test seçiminin doğru yapılması sonuçlar için yüksek öneme sahiptir.

Örneklem büyüklüğü hesabı, test sonuçlarını önemli ölçüde etkilemektedir. Ana yığıcıı temsil etmede yeterli olacak en az sayıda birimin örnekleme alınması amaçlanır. Her örneklemede bir alt sınır olduđu gibi sonucun doğruluđunu deđiřtirmeyecek bir üst büyüklük sınırı da vardır. Bu sınıra ulařtıktan sonra örneklem büyüklüğünü arttırmaya devam etmek klinik açıdan anlamlı olmayan sonuçlara yol açabilir.⁶

Bu çalışmanın amacı, tekrarlanan verilerin boylamsal olduđu (deđiřkenlerden birinin ikiden fazla zamanda ölçüldüğü) ve cevap deđiřkeninin niceliksel olduđu durumda, farklı küçük örneklem büyüklüklerinde kullanılacak testlerin belirlenmesi ve sonuçların karşılaştırılmasıdır. Çalışma, bir faktörü tekrarlanan ölçüm içeren (zaman) iki faktörlü (zaman ve materyal) deneme desenine uygun olarak oluşturulmuřtur.

2. GENEL BİLGİLER

2.1. Tekrarlanan Ölçüm Dizaynı

Tekrarlanan ölçümler terimi, genel olarak her deneysel birimin ya da kişinin incelenen parametrelerinin birden çok koşulda ya da zamanda gözlemlenmesini ifade eder. Örneğin aynı değişken farklı zaman noktalarında tekrar ölçüldüğünde tekrarlanan ölçümler elde edilir. Eşleştirilmiş veriler ise bir nedenle eşleştirilmiş ayrı gruplar olduğunda elde edilir. Her iki durumda da eşleşen çiftler ya da denekler arası korelasyon vardır, bu nedenle bu verilere standart varyans analizi (ANOVA) uygulanamaz. Eğer tedaviler eşleştirilmiş çift içinde rastgele atanabiliyorsa (rastgele blok tasarımı), tedaviler arası karşılaştırmalar iki yönlü ANOVA ile gerçekleştirilebilir. Fakat iki yönlü ANOVA, zaman veya dozaj sırasının seçilemediği aylık aralıklarla veya artan dozlarla uygulamanın yapıldığı durumlarda tekrarlanan ölçümlerin verilerine uygulanamaz. Bu durumlarda tekrarlanan ölçümler ANOVA uygulanabilir. Bu tasarım, farklı tedavilerin etkileri incelenirken deneklerin farklılıklarından kaynaklanan varyasyonu azaltır. Tekrarlanan ölçüm analizi, denek içi ve denekler arası çok faktörlü bileşenlerle daha karmaşık dizaynları çözebilir.⁷

Tekrarlanan ölçümler ANOVA analizini uygulamak için üç temel varsayım karşılanmalıdır, bunlar verilerin bağımsızlığı, normallik ve küreselliktir. Gözlemlerin her biri bağımsız olmalıdır. Bu varsayımı kontrol etmenin en kolay yolu, veri kümesindeki her bireyin bir rastgele örnekleme yöntemi kullanarak popülasyondan rastgele örneklendiğini doğrulamaktır. Eğer bir rastgele örnekleme yöntemi kullanıldıysa, her gözlemin bağımsız olduğunu varsaymak güvenlidir. Verilerin dağılımı, normal dağılıma uymalıdır. Bu varsayımı sağlayıp sağlamadığı histogram ya da Q-Q grafiği ile test edilebilir ya da Shapiro-Wilk veya Kolmogorov Smirnov testlerine başvurulabilir. Dağılımın normallik varsayımını karşılamadığı durumlarda veriler dönüştürülür ya da normallik varsayımı gerektirmeyen parametrik olmayan testlerden Friedman testi uygulanabilir. Eşleştirilmiş grupların tüm kombinasyonları arasında farkların varyansı eşit olmalıdır. Bu küresellik varsayımı ihlal edilirse F oranı yükselir ve tekrarlanan ölçümler ANOVA ölçümlerinin sonuçları güvenilir hale gelir. Mauchly'nin küresellik testinin sıfır hipotezi reddedilirse, tekrarlanan ölçümler ANOVA tablosunda F değerini hesaplamak için kullanılan serbestlik derecelerine düzeltme uygulanır. Bir ya da daha

fazla varsayım ihlal ediliyorsa tekrarlanan ölçümler ANOVA sonuçları güvenilir olmayabilir.⁸

2.2. Tekrarlanan Ölçüm Dizaynının Avantaj ve Dezavantajları

Tekrarlanan ölçümlerin elde edildiği çalışmaların önemli gücü, bireysel değişimlere ilişkin bilgi elde etmenin mümkün olmasıdır. Bu dizayn örnek sayısında da tasarruf sağlar. Örneğin, bir tedavinin zaman içindeki etkilerini incelerken, belirli bir zaman noktasında farklı örnekleri gözlemlemek yerine aynı örnekler tekrar tekrar gözlemlenir. Kişiler (örnekler) arasında değişkenlik kaynakları deneysel hatanın dışında tutulabileceğinden tekrarlanan ölçüm dizaynları, genellikle aynı sayıda ve aynı ölçüm modeline sahip kesitsel dizaynlardan daha verimli tahminler sağlar. Bununla birlikte aynı örneklerin tekrar tekrar gözlemlendiği bir çalışmada veriler, kesitsel çalışmaya göre daha güvenilir şekilde toplanır.⁹

Tekrarlanan ölçüm çalışmalarından elde edilen verilerin analizinde iki ana zorluk vardır. İlk olarak aynı deney biriminde yapılan tekrarlanan gözlemler arasındaki bağımlılık nedeniyle analiz karmaşıktır. İkincisi ise, araştırmacı ölçümlerin elde edilmesi için koşulları çoğunlukla kontrol edemez, bu nedenle veriler dengesiz ya da kısmi olarak eksik olabilir. Örneğin boylamsal bir çalışmada, kişiden alınan yanıt, ilgilenilen sonuçla ilgisi olmayan faktörler nedeniyle bir veya birden fazla zaman noktasında eksik olabilir. Tekrarlanan ölçüm verilerinin analizine yönelik birçok yaklaşım üzerine çalışılmış olmasına rağmen, bu yaklaşımlar değişkenlerin normal olarak dağıldığı, verilerin dengeli ve eksiksiz olduğu koşullarla sınırlıdır. Tek cevaplı tekrarlanan ölçümleri analiz etmek için 2 yöntem kullanılabilir: tek değişkenli yaklaşım ve çok değişkenli yaklaşım.

2.3. Tek Değişkenli ve Çok Değişkenli Varyans Analizi Yöntemleri

Tekrarlanan ölçümlerin analizine yönelik en basit yaklaşım, her bir deney biriminden çoklu ölçüm vektörünü tek bir ölçüme indirgemektir. Böylece çok değişkenli yanıt, tek değişkenli yanıtla indirgenir. Bu da bir denekten elde edilen tekrarlanan ölçümler arasındaki korelasyon sorununu önler. Her denekten iki ölçümün olduğu durumlarda devamlı veriler için tek değişkenli yöntemlerden olan eşleştirilmiş t testi, dikotom veriler için ise McNemar testi kullanılabilir. Wishart, bu yaklaşımın kullanımını belgeleyen ilk araştırmacıdır.¹⁰ Tekrarlanan ölçümlerin analizine yönelik tek değişkenli yaklaşım, çoğunlukla her örnekten ortak ölçüm zamanlarında tam veriler elde edildiğinde

uygulanır. Bu yaklaşım, veriler tek bir örneklemden ya da bir veya daha fazla kategorik ortak değişkenle tanımlanmış birden fazla örneklemden veri elde edildiğinde kolaylıkla uygulanabilir.⁹

Tekrarlanan ölçümler varyans analizinde ilgilenilen etkiler şunlardır; denekler arası etkiler (grup), denekler içi etkiler (zaman) ve iki etki türü arasındaki etkileşimler (grup x zaman). Yalnızca denekler arası etkileri içeren testler için hem tek değişkenli hem de çok değişkenli yaklaşımlar aynı testlere yol açar. Bununla birlikte denek içi etkiler ve denek ile denek arası etkileşim etkileri için tek değişkenli ve çok değişkenli yaklaşımlar farklı testlere neden olur. Denek içi etkiler ve bu etkileri içeren etkileşimler için sağlanan çok değişkenli testler Wilk's Lambda, Pillai's Trace, Hotelling-Lawley Trace ve Roy'un en büyük köküdür. Denek içi etkiler için geçerli çok değişkenli testler için gereken tek varsayım, modeldeki bağımlı değişkenlerin, denekler arası etkiler boyunca ortak bir kovaryans matrisi ile çok değişkenli normal dağılıma sahip olmasıdır.⁷

Spesifik bir test istatistiğinin seçimi yalnızca, çok değişkenli yöntem denekler arası faktörlere uygulandığında önem kazanır. Tekrarlanan ölçümler durumunda, bu istatistikler aynı F değerini ve p değerini verir. Bu dört çok değişkenli test, örneklem boyutları büyük olduğu zaman sıfır hipotezine karşı benzer istatistikler sağlayacaktır. Örneklem orta derece olduğunda, Pillai's, Lawley's ve Wilk's benzer güce sahiptir.¹¹ Wilk's Lambda, tek değişkenli R-karenin çok değişkenli karşılığı olarak yorumlanabilir, bağımlı değişkenlerdeki genelleştirilmiş varyans oranını gösterir. Pillai's Trace çok değişkenli ANOVA varsayımlarının ihlali açısından en dirençli olandır. Özellikle örneklem boyutları küçük veya eşit olmadığında ya da kovaryanslar homojen olmadığında kullanımı yararlıdır. Küçük örneklem büyüklüklerinde Tip 1 hataya karşı en yüksek korumayı sağlar. Hotelling's Trace, sadece bir bağımsız değişken olduğu ve bu bağımsız değişkenin iki koşulu (örneğin, iki örneklem) bulunduğu durumda kullanılır. Roy'un en büyük kökünün, çok değişkenli normallik varsayımı ihlal edildiğinde diğer testlere göre gücü daha düşüktür.⁷

Denek içi etkiler ve bu etkileri içeren etkileşimler için tek değişkenli testler, F-testleri tarafından sağlanan olasılıkların doğru olması için bazı varsayımları gerektirir. Bu testler, tip H kovaryansları veya küresellik varsayımı olarak bilinen belirli kovaryans matris modellerini gerektirir.¹² Kovaryans matrislerindeki bu paternlere sahip verilerin Huynh-Feldt koşulunu sağladığı söylenir. Denek içi etkinin sadece 2 düzeyi olduğunda,

küresellik testi gerekmediği için denek içi etkiler ve ilişkili etkileşimler için tek değişkenli hipotezleri test etmek için F testleri kullanılabilir. Bileşik simetri, tekrarlanan ölçümler ANOVA'nın kullanılmasında yeterli bir koşuldur fakat gerekli değildir. Küresellik koşulu, bileşik simetrinin daha genel bir durumudur. Örneğin, tedavi seviyeleri arasındaki farklılıkların varyansları eşit olduğunda geçerlidir. Bileşik simetri koşulu ise, tüm varyanslar eşit olduğunda (kovaryans matrisinin köşegeninde) ve tedaviler arası kovaryanslar (kovaryans matrisinin köşegeni dışında) eşit olduğunda sağlanır.¹³ Bu nedenle, tekrarlanan ölçümler ANOVA uygulanırken küresellik varsayımının kontrol edilmesi gerekir.¹⁴ Küresellik varsayımı, tekrarlanan ölçüm içeren faktörün ardışık seviyeleri arasında elde edilen farklara ilişkin varyansların homojen olması veya varyansların aralarındaki farkların tesadüfe atfedilebilecek kadar küçük olması şeklinde tanımlanır. Bu ön şart geçerli ise farklara ait kovaryans matrisi yapısı $S=s^2I$ şeklinde olur ve varyans analizi sonucunda hesaplanan F istatistiğinin serbestlik derecelerinde bir düzeltme gerekmez.¹⁵ Veriler küresellik koşulunu sağlamıyorsa, düzeltme için pay ve paydanın serbestlik derecelerinde düzenleme yapılabilir. Serbestlik derecesi düzeltme faktörüne dayanan bu iki düzeltme, ϵ (epsilon) olarak bilinir. Her iki düzeltme de epsilonu tahmin eder ve F testleri için anlamlılık düzeylerini belirlemeden önce pay ve payda serbestlik derecelerini bu tahminle çarpar. Greenhouse ve Geisser tarafından önerilen ilk düzeltme Greenhouse-Geisser epsilon olarak belirtilmiştir ve Box'ın epsilon faktörünün maksimum olasılık tahminini temsil eder. Greenhouse ve Geisser tarafından yapılan tahminin aşağı yönlü olma eğiliminde olduğu ve küçük örnekler için çok konservatif olduğu gösterilmiştir.¹⁴ Epsilonun 0 ile 1 aralığında olması gerekmesine rağmen, Huynh ve Feldt'in tahmincisi bu aralığın dışında olabilir. Huynh ve Feldt'in tahmincisi 1'den büyük olduğunda, olasılıklar için tüm hesaplamalarda 1 değeri kullanılır ve olasılıklar ayarlanmaz. Bir başka epsilon olan Lower-Bound epsilon da kullanılabilir. Lower-Bound epsilon, Greenhouse ve Geisser yönteminin bir parçası olarak bilinir. Epsilon 1'e ne kadar yakınsa, verilerin küresellik koşulunu sağlama olasılığı o kadar yüksektir. Küresellik varsayımının ihlali tespit etmenin yöntemi Mauchley testini kullanmaktır.¹⁶ Kovaryans matrisi küresel veya tam simetrik değilse hata kareler ortalaması daha küçük ve dolayısıyla hesaplanacak F değeri daha büyük olur. Bu durum, tip II hata yapma olasılığını artırır ve pozitif yanlı sonuçlara neden olur. Sonuç olarak veriler küresellik varsayımını karşılamıyorsa, düzeltilmiş F testleri kullanılır. Fakat verilerin H tipi kovaryansa sahip olmama ihtimali varsa, bu tek değişkenli testler dikkatli şekilde

yorumlanmalıdır. Bunlar yerine çok değişkenli testler kullanılabilir. Tek değişkenli yöntemde kullanılan varsayım çok sınırlayıcı olduğu için, çok değişkenli tekrarlanan ölçümler ANOVA yöntemleri kullanılabilir. Bu durumda artık kovaryans hakkında basit bir varsayım olmadığı için model daha karmaşıktır.⁷ Tek değişkenli ve çok değişkenli analizler farklı sonuçlara yol açıyorsa, küresellik varsayımı gerektirmediği için çok değişkenli istatistik daha güvenlidir. Fakat küçük örneklem büyüklüğü söz konusu olduğunda çok değişkenli yaklaşımda hata için serbestlik derecesi sayısının azalması, bu yaklaşımın tek değişkenli analizde önemli olan etkileri belirlemede başarısız olmasına neden olabilir. Tek değişkenli ve çok değişkenli yaklaşımlar arasındaki seçim, verilerin bazı koşullarına bağlıdır. Küresellik varsayımı geçerliyse, tek değişkenli yaklaşım daha küçük bir tip II hata oranına ve daha yüksek güce sahiptir. Yeterince büyük örneklem büyüklüğü sağlandığında ise çok değişkenli bir yaklaşımın kullanılması tercih edilir.¹⁴

2.4. İki Yönlü Tekrarlanan Ölçümler ANOVA

İki yönlü tekrarlanan ölçümler ANOVA, iki denek içi faktöre (bağımsız değişken) bölünmüş gruplar arasındaki ortalama farkları karşılaştırır. İki yönlü tekrarlanan ölçümler ANOVA sıklıkla, iki ya da daha fazla ölçüm noktasında bir bağımlı değişken ölçüleceği zaman ya da denekler iki ya da daha fazla koşuldan geçtiğinde (örneğin zaman ve koşul gibi iki faktör) kullanılır. Bu testte asıl amaç, bağımlı değişken üzerinde bu iki faktör arasında bir etkileşim olup olmadığını anlamaktır. İstatistiksel olarak anlamlı bir etkileşim olup olmadığını belirledikten sonra sonuca ulaşmak için bir dizi farklı yaklaşım vardır. İki yönlü tekrarlanan ölçümler ANOVA çok yönlü (omnibus) bir test istatistiğidir ve iki faktör için de hangi gruplar arasında önemli farklılıkların olduğuna ulaşamaz. Gruplar arasındaki farklara ulaşmak için post-hoc testler yapmak gereklidir. Ayrıca, istatistiksel olarak anlamlı etkileşimlerin bulunduğu durumlarda, herhangi bir "basit ana etki" olup olmadığını ve varsa bu etkilerin neler olduğunu belirlemek gerekir. Karma desen ANOVA, iki yönlü tekrarlanan ölçümler ANOVA'ya çok benzer, çünkü bu istatistiksel testlerin her ikisi de iki faktör (genellikle "zaman" ve "durum") ve bu iki değişken arasında etkileşim olup olmadığını test etmeye yöneliktir. Bununla birlikte temel fark, karma desen ANOVA'da, her koşula (örneğin, bir kontrol ve tedavi) maruz kalan deneklerin farklı olması, oysa iki yönlü tekrarlanan ANOVA ölçümlerinde, deneklerin her iki koşulu da (örneğin, bir kontrol ve tedavi) geçirmesidir. İki yönlü tekrarlanan ölçümler ANOVA testi için 5 temel varsayım gereklidir:

- Bağımlı değişken sürekli ölçümlü olmalıdır. Sürekli değişken örnekleri olarak revizyon süresi (saat olarak ölçülür), zeka (IQ puanı kullanılarak ölçülür), sınav performansı (0 ile 100 arasında ölçülür), ağırlık (kg olarak ölçülür) vb. verilebilir.
- İki denek içi faktör (bağımsız değişken) en az iki kategorik, "ilişkili grup" veya "eşleşen çiftlerden" oluşmalıdır. İlişkili gruplar, aynı deneklerin her iki grupta da bulunduğunu gösterir. Her grupta aynı deneklere sahip olmanın mümkün olmasının nedeni, her denneğin aynı bağımlı değişken üzerinde iki zamanda ölçülmüş olmasıdır (tedavi öncesi ve sonrası gibi).
- İlişkili grupların herhangi bir kombinasyonunda önemli aykırı değerler olmamalıdır. Aykırı değerler, veriler içinde olağan düzeni takip etmeyen tek veri noktalarıdır. Aykırı değerlerle ilgili sorun, iki yönlü tekrarlanan ölçümler ANOVA üzerinde olumsuz bir etkiye sahip olabilmeleri, ilgili gruplar arasındaki farkları bozabilmeleridir (bağımlı değişken üzerindeki puanları artırma veya azaltma gibi). Bu da sonuçların doğruluğunu azaltır.
- Bir diğer varsayım ise, ilişkili grupların her kombinasyonunda bağımlı değişkenin dağılımının yaklaşık olarak normal dağılım göstermesidir.
- Sonuncu varsayım ise, ilişkili grupların her kombinasyonu arasındaki farkların varyansları eşit olmalıdır. Bu küresellik anlamına gelir. Mauchly küresellik testi ile test edilebilir.¹⁷

2.5. İki Faktörlü ve Faktörlerden Birinin Seviyeleri Tekrarlanan Ölçüm İçeren Dizaynlar

Bir faktör seviyesinde tekrarlanan ölçümlerin olduğu, diğer faktörde ise farklı bireylerden ya da deney birimlerinden elde edilen ölçümlerin olduğu dizaynlardır. Bu tip denekler arası faktör de içeren denemelerde, bağımsız faktörün her seviyesinde hesaplanan kovaryans matrislerinin homojen olması gerekir. Bu şartın yerine getirilip getirilmediği aşağıdaki formülle kontrol edilebilir:

$$H_0: \sum_1 = \sum_2 = \dots = \sum_k$$

H_1 : En az ikisi birbirinden farklıdır.

$$S = [1 / \sum_i^k (n_i - 1)] (\sum_i^k n_i - 1) S_i$$

n_i : bağımsız faktörün her seviyesinde yer alan deney ünitesi (birim) sayısı

k: bağımsız faktörün seviye sayısı

S: bilinmeyen popülasyon kovaryans matrisinin tahmini

Kovaryans matrislerinin eşitliğinin kontrolünde kullanılacak olan test istatistiği MC^{-1} 'dir:

$$M = \sum (n_i - 1) \ln |S| - (\sum_i^k (n_i - 1) \ln S_i)$$

$$C^{-1} = [1 - (2p^2 + 3p - 1) / 6(p + 1)(k - 1)] [(\sum_i^k \frac{1}{n_i - 1}) - 1] / (n_i - 1)$$

p: tekrarlanan ölçüm içeren faktörün seviye sayısı

Bu istatistiğin kullanılması için $k \leq 4$, $p \leq 5$, $n_i \geq 12$ olması gerekir. n_i bu büyüklüğe ulaştığında MC^{-1} değeri, $[p(p + 1)(k - 1)] / 2$ serbestlik dereceli ki-kare dağılımı gösterir. Gerekli şartlar mevcut olmadığında hipotez kontrolünde F istatistiğinden yararlanılır.¹⁸

Bağımsız faktör A ve seviyeleri $a_1, a_2, \dots, a_i$ olarak, bağımlı faktör de B ve seviyeleri $b_1, b_2, \dots, b_i$ olarak düşünüldüğünde deneme planı aşağıdaki gibi oluşturulabilir (Tablo 2.1)

Tablo 2.1. Tekrarlanan ölçüme sahip iki faktörlü deneme dizaynı örneği

	Denekler	b_1	b_2	b_i
a_1	1	$1b_1$	$1b_2$	$1b_3$
	2	$2b_1$	$2b_2$	$2b_3$
	.	.	.	.
	k	kb_1	kb_2	kb_3
a_2	1	$1b_1$	$1b_2$	$1b_3$
	2	$2b_1$	$2b_2$	$2b_3$
	.	.	.	.
	k	kb_1	kb_2	kb_3
a_i	.	.	.	.

A faktörünün etkisine ait hipotez aşağıdaki gibi kurulabilir:

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_k$$

$$H_1 : \text{En az bir } \alpha_i \neq \alpha_j$$

$$F = AAKO / \text{Hata}_1 KO$$

α_i : A faktörünün i. seviyesinin etkisi

AAKO: A'lar arası kareler ortalaması

Hata₁KO: Aynı gruptaki denekler arası kareler ortalaması

B faktörünün etkisine ait hipotez aşağıdaki gibi kurulabilir:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k$$

$$H_1 : \text{En az bir } \beta_i \neq \beta_j$$

$$F = \text{BAKO} / \text{Hata}_2\text{KO}$$

β_i : A faktörünün i. seviyesinin etkisi

BAKO: B'ler arası kareler ortalaması

Hata₂KO: B x Hata₁ kareler ortalaması

A x B etkileşimine ait hipotez aşağıdaki gibi kurulabilir:

$$H_0 : \alpha\beta_{11} = \alpha\beta_{12} = \dots = \alpha\beta_{kk}$$

$$H_1 : \alpha\beta_{ij} \neq \alpha\beta_{i'j'} \text{ (etkileşim önemlidir)}$$

$$F = \text{AxB KO} / \text{Hata}_2\text{KO}$$

AxB KO: AxB'ler arası kareler ortalaması

Tekrarlanan ölçümlü analizlerde varyans analizi sonrasında gruplar arası çoklu karşılaştırma yöntemlerini de kullanmak gerekebilir. Kıyaslama değeri hesaplanırken hangi etkinin seviyeleri incelenekse o etkiye ait olan hata varyansı kullanılır. Örneğin A faktörünün seviyeleri karşılaştırılacaksa aynı gruptaki denekler arası varyasyon (HKO₁), B faktörünün seviyeleri karşılaştırılacaksa B x aynı gruptaki denekler arası varyasyon kullanılır (HKO₂). A ve B faktörlerinin etkileşimi önemli bulunmuşsa B faktörünün her seviyesinde A faktörünün etkisi değerlendirileceği zaman toplanmış varyans kullanılır:¹⁸

$$\text{HKO}_{\text{toplaml}} = (\text{HKT}_1 + \text{HKT}_2) / (k(n-1) + k(n-1)(p-1))$$

k: bağımsız faktörün seviye sayısı

p: bağımlı faktörün seviye sayısı

n: alt gruplardaki gözlem sayısı

2.6. Tekrarlanan Ölçümlerde Friedman Testi

Friedman testi, tekrarlanan ölçümlerde parametrik ANOVA'ya alternatif parametrik olmayan yöntemdir. Parametrik olmayan testler, temel alınan popülasyonlar hakkında daha az varsayım gerektirir. Demsar¹⁹, çoklu karşılaştırma yapmak için iyi bilinen parametrik testlerin yanında bir dizi parametrik olmayan test tanımlamış ve bunları güç açısından analiz etmiştir. Parametrik testler için gerekli koşullar sağlanmadığında parametrik olmayan testlerin daha uygun olduğu sonucuna varmıştır.

Tek yönlü varyans analizi, incelenen olay üzerinde tek bir faktörün etkisini incelerken, etkileşimsiz iki yönlü varyans analizi, incelenen olay üzerinde etkili ancak birbirlerini etkilemeyen iki faktörün etkisinin araştırılmasını sağlayan bir yöntemdir. İki bağımlı örnek için işaret testi, ortalama farkının sıfır olmasını ifade eden sıfır hipotezini test etmek için uygun bir parametrik olmayan yöntemdir. Fakat 3 ya daha fazla ilişkili örneğin işlem etkileri arasında fark olmadığını ifade eden sıfır hipotezi test edilmek istendiğinde genellikle Friedman testi kullanılır.²⁰ Friedman'ın iki yönlü varyans analizi için, sıfır hipotezi K tekrarlanan ölçümlerin ya da eşleşen grupların aynı popülasyondan veya aynı medyana sahip popülasyondan geldiğini belirtir.²¹ Sıfır hipotezi altında test, yanıt değişkeninin aynı dağılıma sahip olduğunu varsayar. Veriler, n satır ve K sütun içeren iki yönlü bir tabloda toplanır. Satırlar, blokları/ bireyleri veya eşleşen birey kümelerini temsil eder ve sütunlar çeşitli koşulları/ tedavileri temsil eder. Testin verileri, bloklar halinde tedavilerin sıralarıdır (R_{ik} , $i = 1, \dots, n$; $k = 1, \dots, K$), bu nedenle $1 \leq R_{ik} \leq K$, $i = 1, \dots, n$ 'dir. Friedman testinin varsayımları şunlardır:²²

I. n K-değişkenli rastgele değişkenler karşılıklı olarak bağımsızdır, yani bir satırdaki sonuçlar diğer satırdaki sonuçları etkilemez.

II. Her satırdaki gözlemler, bazı ilgi kriterlerine göre ayrı ayrı sıralanabilir.

Friedman testi, her bir durum/ tedavi için sıra toplamalarının beklenen değerlerden önemli ölçüde farklı olup olmadığını belirler. Friedman'ın önerdiği test istatistiği aşağıdaki şekildedir:²¹

$$FR = \frac{12}{n \cdot k \cdot (k+1)} \sum_{j=1}^k R_j^2 - 3 \cdot n \cdot (k+1)$$

Post-hoc Testler

Sıfır hipotezi reddedildiğinde, hangi koşullar/tedaviler arasında fark olduğunu karşılaştırmak için post-hoc testler uygulanır. Popülasyonun medyan çiftleri arasındaki farkları araştırmak için literatürde birçok çoklu karşılaştırma tekniği bulunmaktadır.

Bonferroni-Dunn testi esnektir ve tüm tedaviler arasında iki tedavi arasındaki farklılıkları test etmek için kullanılabilir. Familywise (aile bazında) Tip I hata oranını azaltmak için sıfır hipotezi reddetmede kullanılan kritik değerde bir düzenleme kullanılır:

$$1 - (1 - \alpha)^c$$

$$c = [K(K - 1)]/2 \text{ (karşılaştırma sayısı)}$$

$$\alpha = \text{karşılaştırma başına tip I hata oranı}$$

Bu ayarlama, karşılaştırmalar kümesinde en az bir Tip I hata yapma genel olasılığının önceden belirlenen α değerini aşmamasını sağlar. Fisher'ın LSD testi, test istatistik denkleminin Tip I hatada herhangi bir ayarlama içermemesi dışında, Bonferroni-Dunn testiyle hesaplama açısından aynıdır. Bu nedenle Fisher'ın LSD testi, familywise Tip I hata oranını korumaz. Wilcoxon-Nemeny-McDonald-Thompson testi önceden belirlenmiş alfa değerini aşmaması için familywise Tip I hata oranını kontrol eder. Binom işaret testi, yalnızca eşleştirilmiş gözlemler arasındaki farkların işaretlerinde bilgi olduğunu varsayar. Wilcoxon testinin uygulanması, mutlak skor farklarının sıralı olmasını gerektirir. Hem Binom işareti hem de Wilcoxon test istatistikleri, tedaviler arasındaki puan farklılıklarına dayanmaktadır. Wilcoxon testi, Binom işaret testinden daha fazla bilgi kullanır ve sonuç olarak daha güçlü bir test sağlar.²³ Bonferroni prosedürü çok basittir fakat bunun karşılığında güç eksikliği vardır. Bu durumu iyileştirmek için Holm prosedürü, Hochberg prosedürü ve Hommel prosedürü gibi Bonferroni prosedüründe birkaç değişiklik önerilmiştir.²⁴ Bu modifikasyonlar Binom işareti ve Wilcoxon testlerine uygulanabilir. Bu yöntemler, test edilen her bir bireysel hipoteze eşit olmayan paylaştırma kullanır. Holm bir azalan aşamalı (step-down) prosedürken Hochberg ve Hommel, önem sırasına göre sıralanan hipotezleri sırayla test eden artan aşamalı prosedürlerdir. Azalan ve artan aşamalı yöntemlerde test edilecek hipotezler kendi p değerlerine göre sıralanır. Azalan aşamalı yöntemlerde teste en küçük p değerine sahip hipotez ile başlanır, artan aşamalı yöntemlerde ise teste en büyük p değerine sahip hipotez ile başlanır.²⁵

Hochberg yöntemi için ayarlanan p değerleri, Holm yöntemi için olandan daha küçüktür. Hommel'in prosedürüne ise eşit ya da daha büyüktür. Bununla birlikte hem Hochberg hem de Hommel prosedürleri, yalnızca bireysel test istatistikleri bağımsız ya da pozitif bağımlı olduğunda familywise Tip I hata oranını korur.

2.7. Örneklem Büyüklüğü

Tüm popülasyonu incelemek çoğunlukla imkansız olduğu için seçilen örnekler üzerinde çalışmalar yapılır. Seçilen örneklerden elde edilen sonuçların popülasyona genellenmesi amaçlanır. Bunun için ise uygun örneklem yöntemleri kullanılmalıdır. Örneklem büyüklüğünün eksik ya da fazla değil, yeterli düzeyde olması gerekmektedir. Yeterince büyük olan örnek, popülasyonu daha iyi temsil edecek ve dolayısıyla daha doğru sonuçlar verecektir. Bununla birlikte örneklem büyüklüğü gerekenden yüksek olduğunda doğruluktaki artış küçük olacaktır, bu da fazladan hasta, hayvan ya da örnek almak için harcanan çaba veya masrafı karşılamayacaktır. Özellikle laboratuvar hayvanlarının kullanıldığı biyomedikal çalışmalarda uygun örneklem hacmine karar verme işlemi ekonomik, bilimsel ve etik açıdan önem arz etmektedir.²⁶ Uygun tasarımlı bir deney için deney birimi, deneyin tekrar sayısı gibi etkenlerin iyi tanımlanması, sonuçların doğruluğunu arttıracak ve etkili faktörlerin yorumlanmasını kolaylaştıracaktır.²⁷

Varyans Analizi için Güç ve Örneklem Büyüklüğü

Varyans analizinde güç ve örneklem büyüklüğü hesaplamaları t testinden çok farklı değildir. Tek fark, minimum saptanabilir tedavi etkisinin büyüklüğünü nicelleştirme yolu ve bu büyüklükle ilgili matematiksel ilişki ve hatalı bir şekilde tedavi etkisi belirleme riskidir. Tespit edilecek tedavi etkisinin ölçümü, ikiden fazla grup olduğu için t testinden daha karmaşıktır. Tedavi etkisinin büyüklüğü, t testinden farklı olarak tanımlanmasına rağmen, merkezsizlik parametresi (f) ile tekrar ölçülür. Bir varyans analizinin gücünü tahmin etmek için tedavi gruplarının sayısı, örneklem büyüklüğü, kabul edilen yanlış pozitif risk ve tespit edilmek istenen tedavi etkisi büyüklüğü belirtilir. İlk adım, merkezsizlik parametresi ile tedavi etkisinin büyüklüğünü tanımlamaktır. Tespit etmek istenilen herhangi iki tedavi grubu arasındaki minimum fark ' δ ' olarak belirlenir (t testinin gücünün hesaplandığı gibi). σ , popülasyondaki standart sapma, k tedavi gruplarının sayısı, n ise her tedavi grubunun örneklem büyüklüğüdür.²⁸

$$\phi = \frac{\delta}{\sigma} \sqrt{\frac{n}{2k}}$$

İstatiksel yöntemleri doğru şekilde kullanan ve tasarımı doğru şekilde yapılan çalışmalarda etkiyi tespit etmeyi engelleyecek düzeyde örnek sayısının küçük olması nedeniyle, klinik olarak önemli farklılıklar tespit edilemeyebilir. Bazen araştırmacılar, gerekli örneklem büyüklüğünü azaltmak için kabul edilebilir buldukları gücü azaltır. Örnek büyüklüğünü artırmak genellikle iki nedenden dolayı gücü artırır. Birincisi, örneklem büyüklüğü arttıkça serbestlik derecesi artar ve bu durumda test istatistiğinin değeri genellikle azalır. İkinci olarak t test istatistiği değeri, örnek sayısı n arttıkça artar. Sonuç olarak işlemin belirli boyutta etkisi olduğunda ortaya çıkan t değerlerinin dağılımı (δ / σ), örnek sayısı arttıkça da daha yüksek t değerlerinde yer almaktadır. Örnek büyüklüğü arttıkça testin gücü de artar. Aslında klinik olarak anlamlı olacak kadar büyük bir etkiyi saptamak için gereken örneklem büyüklüğünü tahmin etmek, güç hesaplamalarının asıl uygulama alanıdır. Bu tür hesaplamalar, yeterince büyük bir örneklem elde etmek için kaç hastanın alınması gerektiğini ve kaç merkezin dahil edilmesi gerektiğini tahmin etme konusunda randomize klinik çalışmaların planlanmasında önemli bir rol oynar.²⁸ Fakat örneklem büyüklüğünün alt sınırı olduğu gibi, rasyonel bir üst sınırı da vardır. Örneklem, uygun büyüklüğe ulaştıktan sonra elde edilecek kazanç önemsizleşirse örneklem büyüklüğünü arttırmaya devam etmek gereksizdir.⁶

Tüm deneme düzenlerinde olduğu gibi tekrarlanan ölçüm içeren denemeler için de gerekli örneklem büyüklüğü asıl olarak testin gücü ve etki büyüklüğü ile ilgilidir. Fakat tekrarlanan ölçüm içeren araştırmalarda tekrarlanan ölçüm sayısı ve ölçümler arası korelasyon da dikkate alındığında sabit bir güce karşılık gerekli örneklem büyüklüğü, aynı etki büyüklüğü için bağımsız ölçüm içeren denemelerdekinden daha düşük olacaktır. Bu durum tekrarlanan ölçüm içeren deneme düzenlerinin en önemli avantajlarından birisidir. Maliyet de gerekli örneklem büyüklüğünün saptanmasında dikkate alınması gereken bir konudur. Eğer hastalara uygulanacak tedavi şekli ya da takip süresinin maliyeti yüksek ise bu tedavinin çok fazla hastaya uzun süre uygulanması gereksiz bir maliyet oluşturacaktır. Ayrıca uzun süre takipte hastada yeni hastalıkların gelişmesi ve bunların tedavi sonucunu etkilemesi ya da veri kaybı da olasıdır. Tekrarlanan ölçüm

sayısı arttıkça gerekli örneklem büyüklüğü de azalabilir. Bu nedenle tekrarlanan ölçümlü çalışmalarda tekrarlanacak ölçüm sayısı da önemlidir.²⁹

Bir testin gücü, 1- Tip II hata olasılığı olarak tanımlanır. Tip II hata, aslında olasılığı β olarak adlandırılan bir fark olduğunda fark yok sonucuna ulaşılmasıdır (sıfır hipotez reddedilmez). Bu nedenle bir çalışmanın gücü, bu fark mevcutken bir fark tespit etme olasılığını yansıtır. Karşılaştırılan tedaviler arasında istatistiksel bir fark gösterilemediyse, anlamlı sonuçlara varılabilmesi için çalışmaların yeterli güçle planlanması tıbbi araştırmalar için çok önemlidir. Daha fazla güç, Tip II hatalar için daha az risk ve bir fark varsa onu tespit etmek için daha fazla ihtimal anlamına gelir. Güç, en az %80 ya da tercihen %90 olacak şekilde belirlenmelidir. %90 güç, tedaviler arasındaki gerçek fark planladığımıza eşitse, çalışmanın bunu tespit etmeme ihtimalinin sadece %10 olduğu anlamına gelir. Artan güç ile örneklem büyüklüğü artar.³⁰

Çok yaygın olarak çalışma yapılmadan önce güç analizi yapılmadığında ve önemli olmayan sonuçlarla karşılaşıldığında araştırmacılar gözlemlenen güç olarak da adlandırılan post-hoc güç analizine yönelirler. Bu amaçla araştırmacılar, belirli farkı saptayacakları gücü belirlemek için çalışmanın örneklem sayısı, gözlemlenen farkı ve değişkenliğini kullanır. Ancak, post-hoc güç analizlerinin üç nedenden dolayı istatistiksel anlamı düşüktür. Birincisi, p değerleri ile post-hoc güç arasında birebir ilişki olduğu için örneklem hakkında ek bir bilgi taşımaz. İkincisi, anlamsız p değerleri her zaman düşük güce denk gelir ve iyi ihtimalle 0,05'e eşit ya da daha büyük p değerleri için %50'den biraz daha büyük olacaktır. Üçüncüsü, post-hoc gücü hesaplarken araştırmacılar gözlemlenen farkın klinik olarak anlamlı olduğu ve kesin olarak reddedemedikleri sıfır hipotezinden gerçeği daha fazla temsil ettiği varsayımında bulunurlar.³¹ Bununla birlikte, hipotez testi teorisinde gözlemlenen fark yalnızca önsel (priori) kurulan hipotezler arasında seçim yapmak için kullanılmalıdır. Deneysel (posteriori) 'de, bir bulgunun uygunluğunu değerlendirmek için güven aralıklarının kullanılması tercih edilir. Güven aralığı, gerçek farkı içeren emin olabileceğimiz değer aralığını temsil eder. Doğrudan örneklem büyüklüğü ile ilgilidir ve p değerlerinden daha fazla bilgi taşır.³⁰

Nicel veriler için, araştırmacıların alternatif tedavilerin beklenen değişkenliğini de belirlemesi gerekir; belirtilen sonuçta ne kadar değişkenlik beklenirse, tedaviler arasında ayırım yapmak o kadar zor olur ve gereken örneklem büyüklüğü o kadar büyük olur. Bu değişkenlik planlama sırasında eksik tahmin edilirse, hesaplanan örneklem

büyüklüğü çok küçük olacak ve çalışmanın gücü istenenden daha az olacaktır. Oranları karşılaştırmak için örneklem büyüklüğünün hesaplanması, her grupta belirtilen sonuç ile beklenen oranı kullanır. Hayatta kalma verileri için örneklem büyüklüğünün hesaplanması, belirli bir zamanda her tedavi grubundaki hayatta kalma oranlarına ve daha az olayın meydana geldiği gruptaki toplam olay sayısına dayanır. Bu iki veri türü için örneklem büyüklüğünün hesaplanmasında değişkenlik görülmez.

Örneklem büyüklüğü hesaplaması, çalışmanın hipotezine dayanmaktadır. Araştırmacı bir birincil hipotez seçmeli ve örneklem büyüklüğü hesaplamasını bu hipoteze dayandırmalıdır. Tekrarlanan ölçümlerle, araştırmacı birkaç hipotezi test etmeyi seçebilir. En sık araştırılan genellikle zamana göre tedavi etkileşimidir. Bu hipotez, yanıt değişkeninin zaman içerisindeki eğiliminin kontrol ve tedavi grupları arasında aynı olup olmadığını test eder. Diğer bir olası hipotez, diğer tüm faktörler arasında ortalaması alınan belirli bir tahmin değişkeninin etkisini test edebileceğimiz ana etki hipotezidir. Örneğin, tekrarlanan tüm ölçümlerin ortalaması alınarak, tedavi grubundaki katılımcıların yanıtlarının, kontrol grubundaki katılımcıların yanıtlarından farklı olup olmadığı test edilebilir. Tekrarlanan ölçüm dizaynlarında örneklem büyüklüğü hesaplanmasında asıl zorluk, araştırmacının varyansları (ya da standart sapmaları) ve tekrarlanan ölçümler arasındaki korelasyonu belirlemesi gerekliliğidir. 4 tekrarlı ölçümde, 4 varyans ya da standart sapma değeri ve 6 korelasyon değeri belirtilmelidir. Doğru bir örneklem büyüklüğünü hesaplamak için belirtilen değerler, verilerde gözlemlenmesi beklenen değerlere mümkün olduğunca yakın eşleşmelidir. Varyans değerlerini belirlemek için, önceki bir çalışmaya dayalı olarak bir varyans değerini tahmin etmek veya deneyime dayalı olarak bilinçli bir tahminde bulunmak ve ardından zaman içinde beklenen varyans eğilimine dayalı olarak diğer varyans değerlerini belirlemek genellikle mümkündür. Zaman içindeki varyans eğilimini belirlemek için, bilimsel içerikler genellikle mantıklı bir model sağlar. Örneğin hastalık gelişimi, yanıtların tekrarlanan ölçümleri arasındaki varyansta monoton olarak artışa neden olur. Çoğu zaman, varyansın zaman içerisinde sabit kaldığı varsayılabilir.³²

Korelasyonun tahmini için, varyansların belirlenmesinde kullanılan prensipler uygulanır. Önceki bir çalışmaya veya mantıklı bir tahmine dayanarak bir korelasyon değeri tahmin edilip ardından beklenen korelasyon modellerine dayanarak diğer korelasyon değerleri belirlenebilir. Artan karmaşıklıkta üç tür korelasyon modeli

düşünülebilir: eşit korelasyonlar, kural-tabanlı modeller ve yapılandırılmamış korelasyonlar (spesifik bir modeli olmayan). En basit korelasyon modeli, tekrarlanan ölçümler arasında sabit bir korelasyon olduğunu varsayar. Fakat bu model genellikle boylamsal çalışmalarda geçerli değildir çünkü zaman içinde daha sonra alınan ölçümler, daha önce alınan ölçümlerden daha az ilişkilidir. Genellikle temel ölçüm ile bir sonraki ölçüm arasındaki korelasyonun, temel ölçümle en sonki ölçüm arasındaki korelasyondan daha büyük olması beklenir. Daha gerçekçi bir hipotez, tekrarlanan ölçümler arasındaki korelasyonun zamanla katlanarak azalmasıdır. Bu durumda, kural-tabanlı bir modelle korelasyondaki düşüşü modellemeyi düşünebiliriz. Yaygın olarak kullanılan bir model, “lineer üs birinci dereceden otoregresif” ailesine ait olan birinci dereceden otoregresiftir. “Lineer üs birinci dereceden otoregresif” modelin kullanılması, 2 korelasyon parametresinin sağlanmasını gerektirir: temel korelasyon ve bozulma (azalma) oranı. Temel korelasyon, zaman içinde yan yana alınan 2 ölçüm arasındaki korelasyondur. Azalma oranı, tekrarlanan ölçümler arasındaki süre arttıkça temel korelasyondaki düşüş oranıdır. Biyolojik ve davranışsal verilerden yola çıkarak bu oran 0.05 ile 0.5 arasında olabilir. Belirtilmesi gereken korelasyon sayısı, $p \times (p-1)/2$ ’dir. Burada p , tekrarlanan ölçümlerin sayısıdır. Yapılandırılmamış korelasyonları varsaymak, çoğu korelasyonun tahmin edilmesini gerektirir. Ayrıca, tekrarlanan ölçüm sayısı arttıkça belirtilmesi gereken korelasyonların sayısı da önemli ölçüde artar.³²

3. GEREÇ VE YÖNTEM

Bu çalışmada, diş hekimliği kliniğinde güncel olarak kullanılan iki farklı 3 boyutlu yazıcı ve bir kazıyıcı yardımıyla üretilen geçici kron materyalleri kullanıldı. Diş hekimliğinde 3 boyutlu yazıcılarla üretim yapmak ve bu alanda kullanılacak malzemelerin araştırılıp geliştirilmesi güncel olarak çalışılan bir konudur. Klinik pratikte sıklıkla kullanılan malzemelerden biri geçici kron materyalidir. Tek seansta geçici kron yapılıp hastaya teslim edilmesi, ölçü materyali kullanılmasına gerek olmaması gibi avantajları nedeniyle 3 boyutlu yazıcılar ile üretilen geçici kronlar klinikte yerini almaktadır. 3 boyutlu yazıcılar, geçici rezin polimerizasyonu için stereolitografi, fotopolimer jet, dijital ışık işleme ve sıvı kristal ekranı gibi farklı baskı teknolojileri kullanır. Farklı baskı teknolojilerine sahip 3 boyutlu yazıcılar ile üretilmiş geçici kronların mekanik özellikleri ile ilgili birçok çalışma yapılmıştır. Önceki bir çalışmada 3 boyutlu yazıcılar ile üretilmiş geçici kronların klinik olarak kabul edilebilir olduğu bildirilmiştir.³³ Bu yöntemle üretilen geçici rezin materyallerinin mekanik özelliklerini anlayabilmek için diş hekimliğinde kullanılan diğer yöntemlerle ve farklı geçici rezin materyalleri ile karşılaştırmak gerekir. Eğilme dayanımı, sertlik ve pürüzlülük gibi mekanik özelliklerin yanında materyaldeki renk değişimi de uzun süreli geçici protezler planlandığında önemli hale gelmektedir.

Bu bilgiler ışığında bu çalışmada dijital ışık işleme (DLP) ve sıvı kristal ekran (LCD) baskı teknolojilerini kullanan yazıcılar (sırasıyla Grup 1 ve Grup 2) ve bir kazıyıcı kullanılarak (Grup 3) geçici rezin örnekleri hazırlandı. Örneklerin üretimi için 3 boyutlu modelleme programında (Blender) 2 mm kalınlığa ve 10 mm çapa sahip disk tasarlandı. Bu dosya STL dosya formatında üretim cihazlarına aktarıldı. Her iki 3 boyutlu yazıcıda baskı plakası üzerindeki açısı (90°), konumlandırma ve desteklerin bağlantı noktaları ve sayısı aynı şekilde düzenlendi. Böylece sonucun baskı teknolojisi ve rezin farkları dışında diğer faktörlerden etkilenme ihtimalinin önüne geçilmesi amaçlandı. Baskı tamamlandıktan sonra üretici talimatlarına uygun olarak alkol ile yıkandı ve polimerize edildi. Kazıyıcı grubunda polimetilmetakrilat (PMMA) diski kullanıldı ve 5 akslı kazıyıcı kullanılarak üretim tamamlandı. Her gruptan 20 örnek olmak üzere toplam 60 örnek üretildi. Tüm örneklerin polisajı sırasıyla 800 ve 1200 gritli silikon karbid kağıt ile 20'er saniye polisaj cihazında (Minitex 233, Presi) su altında ve 200 rpm hızla yapıldı. Tüm örneklerin başlangıç renk ölçümleri, beyaz zemin üzerinde spektrofotometre (Vita

Easyshade Advance 4.0) ile ölçüldü. Rengin L, a ve b değerleri kaydedildi (T0 ölçümleri). Siyah rengin L değeri 0 iken, beyaz rengin L değeri 100'dür. a eksenini kırmızılık (a+) ve yeşillik (a-), b eksenini sarılık (b+) ve mavilik (b-) olarak tanımlanır. Termal yaşlandırma için termosiklus cihazı (Thermocycler 1100 SD Mechatronik GmbH) kullanıldı. 5⁰ ve 55⁰ sıcaklıklarda 30 saniye daldırma zamanı ve 10 saniye transfer zamanı ayarlandı.³⁴ Örnekler öncelikle 2500 termal sıklusa (klinik kullanımda yaklaşık 3 ayı simule eder) maruz bırakıldı. Renk ölçümleri aynı koşullarda tekrar yapıp L, a ve b değerleri kaydedildi (T1 ölçümleri). Örnekler 5000 siklus (klinik kullanımda 6 ayı simule eder) tamamlanıncaya kadar tekrar termal yaşlandırmaya devam edildi ve son renk ölçümleri de kaydedildi (T2 ölçümleri). Tekrarlı ölçümlerde gerçekleşen renk değişimlerini (ΔE_{00}) hesaplamak için CIEDE2000 formülü kullanıldı:

$$\Delta E_{00}^* = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2 + R_T \frac{\Delta C'}{k_C S_C} \frac{\Delta H'}{k_H S_H}}$$

T0 ve T1 arasındaki renk değişimi $\Delta E_{00(1)}$, T1 ve T2 arasındaki renk değişimi $\Delta E_{00(2)}$ olarak kaydedildi. Gözle ayırt edilebilir ΔE_{00} sınırı 0,8 iken kabul edilebilirlik sınırı 1,8 olarak kabul edildi.³⁵

Küçük örneklem büyüklüklerinde farklı örnek sayısının, tekrarlanan ölçümlü veri analizi sonuçları üzerine etkisini incelemek için priori güç analizi yapıldı. Daha önceki ilişkili bir çalışmanın³⁶ sonucu baz alınarak, G-power (3.1.9.4 versiyon) yazılımı ile yapılan güç analizinde (güç: 0,95, α :0,01, etki büyüklüğü:1,36) 3 grup için örnek sayısı toplamda 15 olarak tespit edildi. Testin gücünü arttırmak için en düşük örnek sayısı 8 olarak belirlendi. Gruplar arasında 4'er fark olacak şekilde diğer 3 grubun örnek sayısı da 12, 16 ve 20 olarak belirlendi. Bu çalışmada küçük örneklem büyüklükleri incelendiği için grupların, merkezi limit teoreminde belirtilen 30 örneklem büyüklüğünden daha az olmasına dikkat edildi. Ölçümlerin analizi, her grup için tekrarlandı ve 4 ayrı istatistiksel analiz yapıldı. 8, 12 ve 16 örnek sayısına sahip grupları oluşturmak için basit rastgele örnekleme yöntemi kullanılarak seçim yapıldı. Her örneklem büyüklüğü için 1 ile 20 arasında numaralar yazılı olan kağıtlar rastgele çekildi ve çekilen numaralar o örnekleme dahil edildi.

Örnek sayısı 8 olan gruplar için numaralar: 2, 5, 9, 10, 15, 17, 18, 20

Örnek sayısı 12 olan gruplar için numaralar: 2, 4, 6, 7, 8, 9, 10, 12, 15, 16, 18, 19

Örnek sayısı 16 olan gruplar için numaralar: 2, 3, 4, 5, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18,20

İncelenen her örneklem büyüklüğü için elde edilen L, a ve b verileri ve ΔE_{00} değerlerinin çok değişkenli normal dağılıma uygunluğu Mardia'nın çarpıklık- basıklık, Henze-Zirkler, Doornik-Hansen testleri ile bir yazılım (İnönü Üniversitesi R Web yazılımı) yardımıyla incelendi.³⁷ 3 grubun başlangıç renk değerleri (L, a ve b) farklı olduğu için her grupta ayrı ayrı zamana bağlı değişim incelendi. Başlangıç renklerinden bağımsız olarak hesaplanan zaman içerisindeki renk değişim değerleri (ΔE_{00}) ise grup ve zaman değişkeni dahil edilerek analiz edildi. Normal dağılım gösteren veriler tekrarlanan ölçümlü ANOVA ile test edilirken, normalliği sağlamayan verilerin istatistiği Friedman testi ile yapıldı. İstatistik testlerin anlamlılık seviyesi 0,05 olarak belirlendi. Deneysel (post-hoc) güç analizi G-power (3.1.9.4 versiyon) yazılımı ile yapıldı.

4. BULGULAR

4.1. Örneklem Büyüklüğü 8 için L, a ve b Verilerinin İstatistiksel Analizi

L, a ve b değerlerinin tanımlayıcı istatistikleri aritmetik ortalama ve standart sapma değerleri Tablo 4.1’de verildi.

Tablo 4.1: Örneklem büyüklüğü 8 için her grubun tanımlayıcı istatistik sonuçları ($\bar{x} \pm S$) ve çok değişkenli normal dağılım testlerinden Mardia küçük örneklem çarpıklık testinin p değerleri

Grup	L0	L1	L2	a0	a1	a2	b0	b1	b2
1	91,22± 0,36	92,18± 0,31	92,07± 0,25	1,62± 0,27	1,67± 0,34	1,39± 0,25	20,62± 0,34	18,07± 0,42	18,00± 0,31
2	86,90± 0,28	85,26± 0,52	85,84± 0,52	0,23± 0,05	0,60± 0,14	0,32± 0,10	35,60± 0,25	34,88± 0,37	35,12± 0,28
3	89,74± 0,18	87,81± 0,19	87,63± 0,14	7,66± 0,06	7,90± 0,05	7,95± 0,03	51,57± 0,11	50,38± 0,23	50,27± 0,20
	Mardia (L)			Mardia (a)			Mardia (b)		
1	0,889			0,709			0,664		
2	0,928			0,914			0,568		
3	0,687			0,998			0,948		

Çok değişkenli normallik testlerinden Mardia, Henze-Zirkler ve Doornik-Hansen testlerinin sonuçlarına göre tüm grupların L, a ve b değerleri normal dağılım gösterdi. Örnek olarak Mardia küçük örneklem çarpıklık testinin p değerleri Tablo 4.1’de verildi. Küresellik Mauchly testi ile kontrol edildi ve küresellik varsayımının sağlandığı tespit edildi (Tablo 4.2).

Tablo 4.2: Her grup için Mauchly küresellik testinin sonuçları

Grup		Mauchl's W	Ki- kare	p	Greenhouse -Geisser	Hyunh- Feldt	Lower -bound
1	L	0,823	1,170	0,557	0,850	1,000	0,500
	a	0,948	0,322	0,851	0,950	1,000	0,500
	b	0,485	4,337	0,114	0,660	0,754	0,500
2	L	0,387	5,689	0,058	0,620	0,620	0,688
	a	0,451	4,776	0,092	0,646	0,730	0,500
	b	0,764	1,613	0,446	0,809	1,000	0,500
3	L	0,840	1,048	0,592	0,862	1,000	0,500
	a	0,955	0,275	0,872	0,957	1,000	0,500
	b	0,739	1,816	0,403	0,793	0,987	0,500

Tekrarlanan ölçümlü ANOVA testinin sonuçlarına göre Grup 1'de L değerleri farklı zaman noktalarında istatistiksel olarak farklı bulundu ($F=52,491$, $p<0,001$). Zamana bağlı ikili karşılaştırmalarda L0 ile L1 arasında ($p<0,001$) ve L0 ile L2 arasında ($p<0,001$) istatistiksel olarak anlamlı bir fark mevcutken L1 ile L2 arasında ($p=1,000$) anlamlı bir fark bulunmamıştır. Farklı zaman noktalarında ölçülen a değerleri arasında da istatistiksel olarak anlamlı bir fark vardı ($F=19,012$, $p<0,001$). İkili grup karşılaştırmasında a0 ile a1 arasında anlamlı bir fark bulunmamışken ($p=1,000$) a1 ile a2 arasında p değeri 0,001 ve a0 ile a2 arasında p değeri 0,006 bulunmuştur. b değerleri arasındaki farklar incelendiğinde, zaman içerisindeki değişimi anlamlı bulundu ($F=116,953$, $p<0,001$). Grupların ikili karşılaştırmalarında b0 ile b1 arasında ve b0 ile b2

arasında anlamlı bir fark varken ($p < 0,001$), b1 ile b2 arasında p değeri 1,000 olarak bulundu (Tablo 4.3).

Tablo 4.3: Grup 1 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri

	T0	T1	T2	F	p
L	91,22±0,36 ^a	92,18±0,31 ^b	92,07±0,25 ^b	52,491	<0,001
a	1,62±0,27 ^c	1,67±0,34 ^c	1,39±0,25 ^d	19,012	<0,001
b	20,62±0,34 ^e	18,07±0,42 ^f	18±0,31 ^f	116,953	<0,001

p değerleri= a-b: <0,001

c(T0)-d: 0,006, c(T1)-d: 0,001

e-f: <0,001

Grup 2’de L’ler arası zamana bağlı istatistiksel olarak fark bulundu ($F=33,168$, $p < 0,001$). Zamana bağlı ikili karşılaştırmalarda L0 ile L1 arasında ($p < 0,001$), L0 ile L2 arasında ($p=0,009$) ve L1 ile L2 arasında ($p=0,002$) istatistiksel olarak anlamlı bir fark görülmüştür. Farklı zaman noktalarında ölçülen a değerleri arasında da istatistiksel olarak anlamlı bir fark vardı ($F=56,704$, $p < 0,001$). İkili grup karşılaştırmasında ise tüm karşılaştırmalarda anlamlı farklılıklar gözlemlendi. a0 ile a1 arasında ve a1 ile a2 arasında p değeri <0,001 iken, a0 ile a2 arasında p değeri 0,012 olarak bulunmuştur. b değerlerinin zaman içerisindeki değişimi incelendiğinde istatistiksel olarak anlamlı bir farklılık gözlemlendi ($F=14,236$, $p < 0,001$). Grupların ikili karşılaştırmalarında b0 ile b1 arasında ($p=0,011$) ve b0 ile b2 arasında ($p=0,014$) anlamlı bir fark varken b1 ile b2 arasında p değeri 0,271 olarak bulundu (Tablo 4.4).

Tablo 4.4: Grup 2 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri

	T0	T1	T2	F	p
L	86,90±0,28 ^a	85,26±0,52 ^b	85,84±0,52 ^c	33,168	<0,001
a	0,23±0,05 ^c	0,60±0,14 ^d	0,32±0,10 ^e	56,704	<0,001
b	35,60±0,25 ^f	34,88±0,37 ^g	35,12±0,28 ^g	14,236	<0,001

p değerleri=a-b: <0,001, a-c: 0,009, b-c: 0,002

c-d: <0,001, c-e: 0,012, d-e: <0,001

f-g(T1): 0,011, f-g(T2): 0,014

Grup 3’de L’ler arası istatistiksel olarak fark bulundu ($F=458,425$, $p < 0,001$). Zamana bağlı ikili karşılaştırmalarda L0 ile L1 arasında ($p < 0,001$) ve L0 ile L2 arasında ($p < 0,001$) istatistiksel olarak anlamlı bir fark mevcutken L1 ile L2 arasında ($p=0,086$) anlamlı bir fark bulunmamıştır. Farklı zaman noktalarında ölçülen a değerleri arasında da istatistiksel olarak anlamlı bir fark vardı ($F=124,657$, $p < 0,001$). İkili grup karşılaştırmalarında a1 ile a2 arasında anlamlı bir fark bulunmazken ($p=0,103$), a0 ile a1

arasında ve a0 ile a2 arasında p değeri $<0,001$ olarak bulundu. b değerleri arasındaki farklar incelendiğinde, zaman içerisindeki değişimi anlamlı bulundu ($F=190,811$, $p<0,001$). Grupların ikili karşılaştırmalarında b1 ile b2 arasında anlamlı bir farklılık görülmezken ($p=0,581$) diğer gruplarda p değeri $<0,001$ 'di (Tablo 4.5).

Tablo 4.5: Grup 3 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri

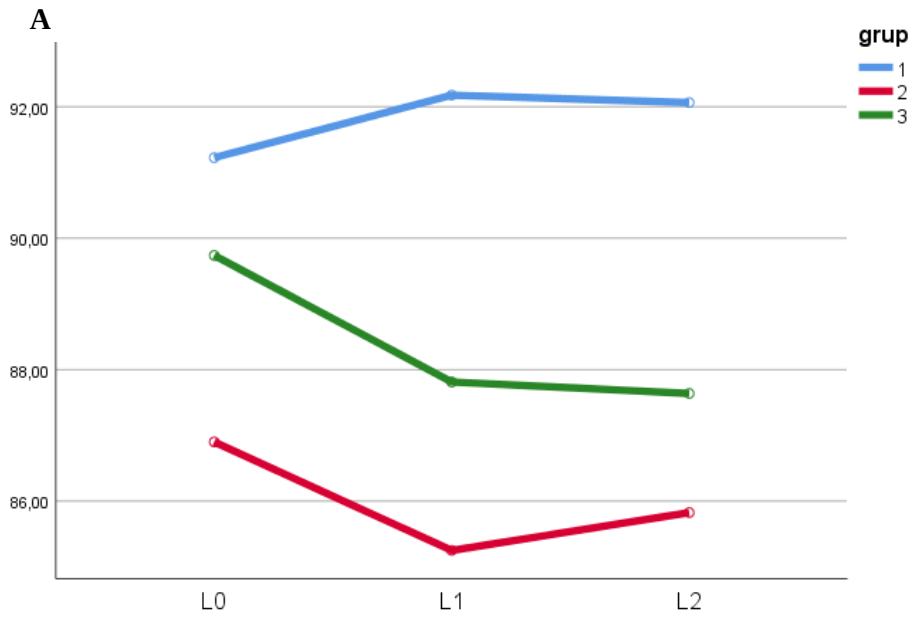
	T0	T1	T2	F	p
L	89,74±0,18 ^a	87,81±0,19 ^b	87,63±0,14 ^b	458,425	$<0,001$
a	7,66±0,06 ^c	7,90±0,05 ^d	7,95±0,03 ^d	124,657	$<0,001$
b	51,57±0,11 ^e	50,38±0,23 ^f	50,27±0,20 ^f	190,811	$<0,001$

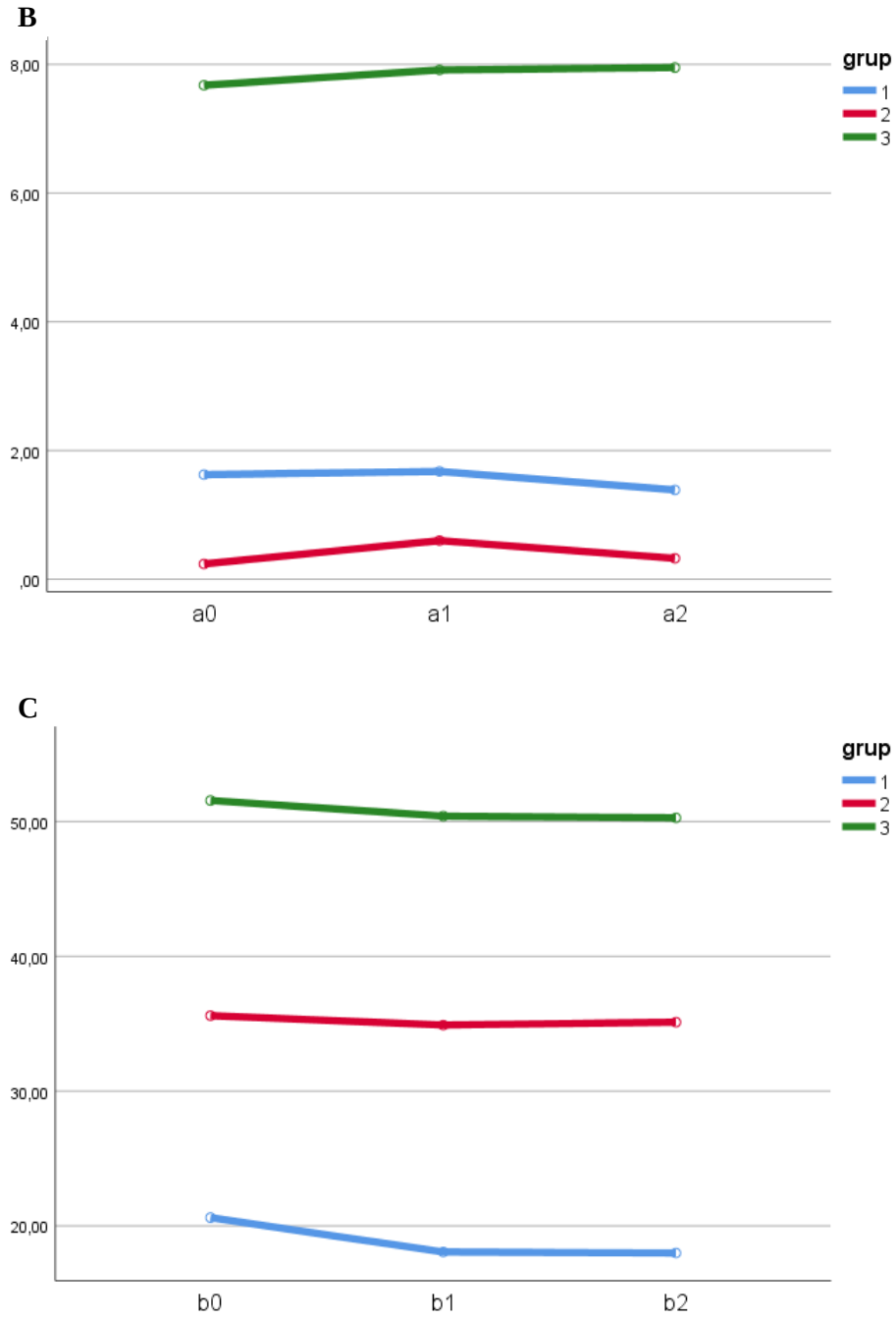
p değerleri=a-b: $<0,001$

c-d: $<0,001$

e-f: $<0,001$

L, a ve b değerlerinin gruplara ve zamana bağlı değişimi Şekil 4.1'de gösterildi.





Şekil 4.1: Örnek sayısı 8 iken renk koordinatlarında gruba ve zamana bağlı değişim
A) L, B) a, C) b

4.2. Örneklem Büyüklüğü 12 için L, a ve b Verilerinin İstatistiksel Analizi

L, a ve b değerlerinin tanımlayıcı istatistiği aritmetik ortalama ve standart sapma değerleri ile Tablo 4.6’da verildi.

Tablo 4.6: Örneklem büyüklüğü 12 için her grubun tanımlayıcı istatistik sonuçları ($\bar{x} \pm S$) ve çok değişkenli normal dağılım testlerinden Mardia küçük örneklem çarpıklık testinin p değerleri

Grup	L0	L1	L2	a0	a1	a2	b0	b1	b2
1	91,15± 0,35	91,99± 0,56	92,02± 0,40	1,71± 0,26	1,75± 0,27	1,40± 0,25	20,65± 0,38	18,31± 0,73	17,96± 0,31
2	86,92± 0,39	85,30± 0,48	85,78± 0,39	0,24± 0,07	0,57± 0,14	0,32± 0,10	35,70± 0,21	34,84± 0,40	35,14± 0,37
3	89,88± 0,25	87,91± 0,19	87,88± 0,20	7,67± 0,06	7,90± 0,09	7,97± 0,04	51,63± 0,15	50,40± 0,29	50,42± 0,21
	Mardia (L)			Mardia (a)			Mardia (b)		
1	0,13			0,19			0,776		
2	0,755			0,872			0,589		
3	0,901			0,518			0,700		

Çok değişkenli normallik testlerinden Mardia, Henze-Zirkler ve Doornik-Hansen testlerinin sonuçlarına göre tüm grupların L, a ve b değerleri normal dağılım gösterdi. Örnek olarak Mardia küçük örneklem çarpıklık testinin p değerleri Tablo 4.6’da verildi. Küresellik Mauchly testi ile kontrol edildi (Tablo 4.7). Küresellik varsayımı sağlanmadığı durumlarda Greenhouse-Geisser düzeltmesi sonucu elde edilen F değerleri dikkate alındı.

Tablo 4.7: Her grup için Mauchly küresellik testinin sonuçları

Grup		Mauchly's W	Ki- kare	p	Greenhouse- Geisser	Hyunh- Feldt	Lower- bound
1	L	0,899	1,061	0,588	0,909	1,000	0,500
	a	0,556	5,879	0,053	0,692	0,760	0,500
	b	0,718	3,309	0,191	0,780	0,886	0,500
2	L	0,323	11,304	0,004	0,596	0,628	0,500
	a	0,521	6,514	0,039	0,676	0,738	0,500
	b	0,930	0,728	0,695	0,934	1,000	0,500
3	L	0,986	0,140	0,933	0,986	1,000	0,500
	a	0,396	9,264	0,010	0,623	0,665	0,500
	b	0,407	8,984	0,011	0,628	0,671	0,500

Verilerin dağılımı normal dağılıma uyduğu için yapılan tekrarlanan ölçümlü ANOVA testinin sonuçlarına göre Grup 1’de L’ler arası istatistiksel olarak fark bulundu ($F=38,343$, $p<0,001$). Zamana bağlı ikili karşılaştırmalarda L0 ile L1 arasında ($p<0,001$) ve L0 ile L2 arasında ($p<0,001$) istatistiksel olarak anlamlı bir fark mevcutken L1 ile L2 arasında ($p=1,000$) anlamlı bir fark bulunmamıştır. Farklı zaman noktalarında ölçülen a değerleri arasında da istatistiksel olarak anlamlı bir fark vardı ($F=56,850$, $p<0,001$). İkili grup karşılaştırmasında a0 ile a1 arasında anlamlı bir fark bulunmamışken ($p=0,526$) a1 ile a2 arasında ve a0 ile a2 arasında p değeri $<0,001$ bulunmuştur. b değerleri arasındaki farklar incelendiğinde, zaman içerisindeki değişimi anlamlı bulundu ($F=136,644$, $p<0,001$). Grupların ikili karşılaştırmalarında b0 ile b1 arasında ve b0 ile b2 arasında anlamlı bir fark varken ($p<0,001$), b1 ile b2 arasında p değeri 0,136 olarak bulundu (Tablo 4.8).

Tablo 4.8: Grup 1 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri

	T0	T1	T2	F	p
L	91,15±0,35 ^a	91,99±0,56 ^b	92,02±0,40 ^b	38,343	<0,001
a	1,71±0,26 ^c	1,75±0,27 ^c	1,40±0,25 ^d	56,85	<0,001
b	20,65±0,38 ^e	18,31±0,73 ^f	17,96±0,31 ^f	136,644	<0,001

p değerleri=a-b:<0,001

c-d: <0,001

e-f:<0,001

Grup 2’de L’ler arası istatistiksel olarak fark bulundu (F=52,538, p<0,001). Tüm ikili gruplar arasında anlamlı farklılıklar görüldü (p<0,001). Farklı zaman noktalarında ölçülen a değerleri arasında da istatistiksel olarak anlamlı bir fark vardı (F=50,352, p<0,001). İkili grup karşılaştırmasında ise tüm karşılaştırmalarda anlamlı farklılıklar gözlemlendi. a0 ile a1 arasında ve a1 ile a2 arasında p değeri <0,001 iken, a0 ile a2 arasında p değeri 0,031 olarak bulunmuştur. b değerlerinin zaman içerisindeki değişimi incelendiğinde istatistiksel olarak anlamlı bir farklılık gözlemlendi (F=29,315, p<0,001). Grupların ikili karşılaştırmalarında b0 ile b1 arasında ve b0 ile b2 arasında anlamlı bir fark varken (p<0,001) b1 ile b2 arasında p değeri 0,052 olarak bulundu (Tablo 4.9).

Tablo 4.9: Grup 2 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri

	T0	T1	T2	F	p
L	86,92±0,39 ^a	85,30±0,48 ^b	85,78±0,39 ^c	52,538	<0,001
a	0,24±0,07 ^d	0,57±0,14 ^e	0,32±0,10 ^f	50,352	<0,001
b	35,70±0,21 ^g	34,84±0,40 ^h	35,14±0,37 ^h	29,315	<0,001

p değerleri: a-b:<0,001, a-c:<0,001, b-c:<0,001

d-e: <0,001, d-f:0,031, e-f: <0,001

g-h:<0,001

Grup 3’de L’ler arası istatistiksel olarak fark bulundu (F=456,903, p<0,001). Zamana bağlı ikili karşılaştırmalarda L0 ile L1 arasında (p<0,001) ve L0 ile L2 arasında (p<0,001) istatistiksel olarak anlamlı bir fark mevcutken L1 ile L2 arasında (p=1,000) anlamlı bir fark bulunmamıştır. Farklı zaman noktalarında ölçülen a değerleri arasında da istatistiksel olarak anlamlı bir fark vardı (F=101,655, p<0,001). Tüm ikili gruplar arasında anlamlı farklar vardı, a1 ile a2 arasında p=0,038 iken diğer gruplar arasında p değeri <0,001 bulundu. b değerleri arasındaki farklar incelendiğinde, zaman içerisindeki değişimi anlamlı bulundu (F=175,399, p<0,001). Grupların ikili karşılaştırmalarında b1 ile b2 arasında anlamlı bir farklılık görülmezken (p=1,000) diğer gruplarda p değeri <0,001’di (Tablo 4.10).

Tablo 4.10: Grup 3 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri

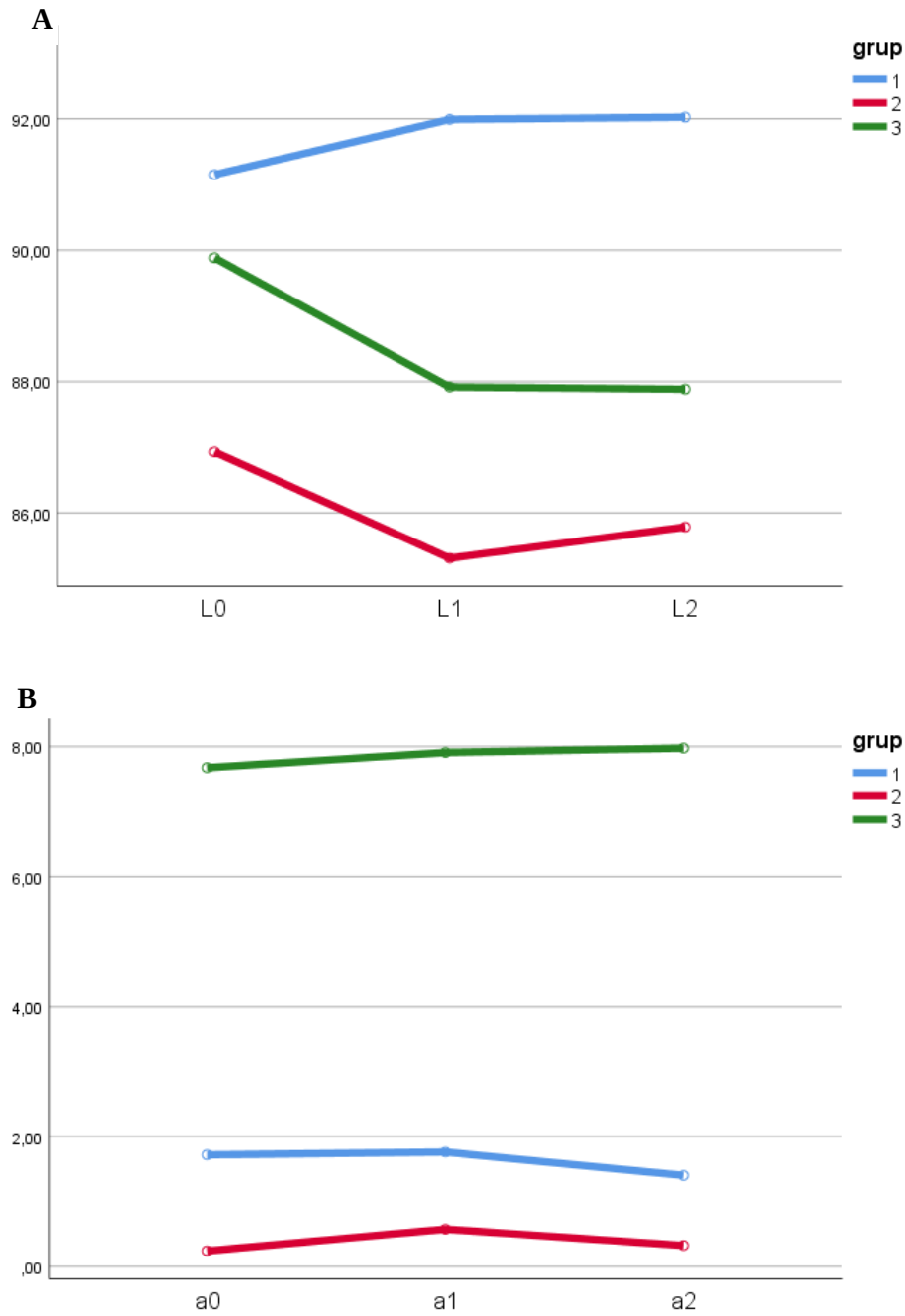
	T0	T1	T2	F	p
L	89,88±0,25 ^a	87,91±0,19 ^b	87,88±0,20 ^b	456,903	<0,001
a	7,67±0,06 ^c	7,90±0,09 ^d	7,97±0,04 ^e	101,655	<0,001
b	51,63±0,15 ^f	50,40±0,29 ^g	50,42±0,21 ^g	175,399	<0,001

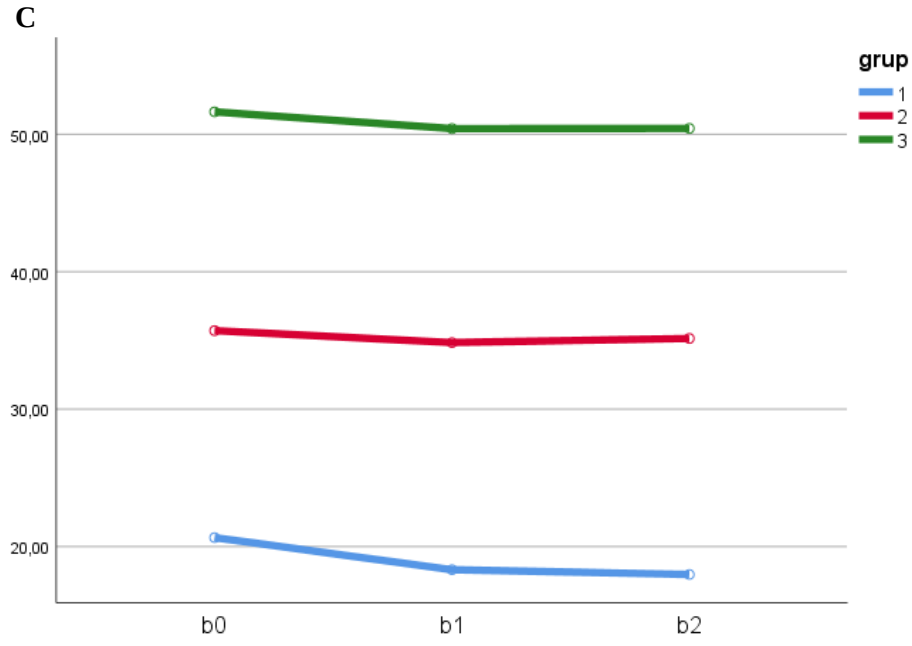
p değerleri=a-b:<0,001

c-d:<0,001, c-e:<0,001, d-e: 0,038

f-g:<0,001

L, a ve b değerlerinin gruplara ve zamana bağlı değişimi Şekil 4.2’de gösterildi.





Şekil 4.2: Örnek sayısı 12 iken renk koordinatlarında gruba ve zamana bağlı değişim

A) L, B) a, C) b

4.3. Örneklem Büyüklüğü 16 için L, a ve b Verilerinin İstatistiksel Analizi

L, a ve b değerlerinin tanımlayıcı istatistiği aritmetik ortalama ve standart sapma değerleri ile Tablo 4.11’de verildi.

Tablo 4.11: Örneklem büyüklüğü 16 için her grubun tanımlayıcı istatistik sonuçları ($\bar{x} \pm S$) ve çok değişkenli normal dağılım testlerinden Mardia küçük örneklem çarpıklık testinin p değerleri

Grup	L0	L1	L2	a0	a1	a2	b0	b1	b2
1	91,21± 0,30	92,20± 0,55	92,21± 0,42	1,68± 0,24	1,71± 0,27	1,37± 0,22	20,69± 0,35	18,23± 0,61	18,00± 0,26
2	87,05± 0,36	85,26± 0,45	85,81± 0,40	0,21± 0,07	0,60± 0,14	0,32± 0,10	35,75± 0,27	35,02± 0,49	35,21± 0,34
3	89,75± 0,31	87,83± 0,20	87,76± 0,29	7,68± 0,54	7,90± 0,77	7,98± 0,71	51,59± 0,17	50,38± 0,28	50,27± 0,29
	Mardia (L)			Mardia (a)			Mardia (b)		
1	0,215			0,662			0,462		
2	0,309			0,85			0,941		
3	0,305			0,215			0,039		

Mardia, Henze-Zirkler ve Doornik-Hansen testleri ile her grup için L, a ve b değerlerinin dağılımının normalliği incelendi. Örnek olarak Mardia küçük örneklem çarpıklık testinin p değerleri Tablo 4.11’de verildi. Grup 3’te b değerleri normal dağılım göstermezken diğer tüm dağılımlar normal dağılıma uygundu. Grup 3’ün b değerlerinin zaman içerisindeki değişimi Friedmann testi ile değerlendirildi, diğerleri için tekrarlanan ölçümlü ANOVA testi uygulandı. Küresellik Mauchly testi ile kontrol edildi (Tablo 4.12). Küresellik varsayımı sağlanmadığı durumlarda Greenhouse-Geisser düzeltmesi sonucu elde edilen F değerleri dikkate alındı.

Tablo 4.12: Her grup için Mauchly küresellik testinin sonuçları

Grup		Mauchly's W	Ki- kare	p	G- Geisser	Hyunh- Feldt	Lower- bound
1	L	0,895	1,55	0,460	0,905	1,000	0,500
	a	0,896	1,53	0,465	0,906	1,000	0,500
	b	0,694	5,10	0,078	0,766	0,836	0,500
2	L	0,619	6,72	0,035	0,724	0,781	0,500
	a	0,421	12,12	0,002	0,633	0,665	0,500
	b	0,577	7,69	0,021	0,703	0,754	0,500
3	L	0,931	1,00	0,606	0,935	1,000	0,500
	a	0,733	4,35	0,113	0,789	0,866	0,500

Grup 1’de verilerin dağılımı normal dağılıma uyduğu için yapılan tekrarlanan ölçümlü ANOVA testinin sonuçlarına göre farklı zaman noktalarında ölçülen L değerleri arasında istatistiksel olarak fark bulundu ($F=69,609$, $p<0,001$). Zamana bağlı ikili karşılaştırmalarda L0 ile L1 arasında ve L0 ile L2 arasında istatistiksel olarak anlamlı bir fark ($p<0,001$) mevcutken L1 ile L2 arasında ($p=1,000$) anlamlı bir fark bulunmamıştır. Farklı zaman noktalarında ölçülen a değerleri arasında da istatistiksel olarak anlamlı bir fark vardı ($F=51,038$, $p<0,001$). İkili grup karşılaştırmasında a0 ile a1 arasında anlamlı bir fark bulunmamışken ($p=1,000$) a1 ile a2 arasında ve a0 ile a2 arasında p değeri $<0,001$ bulunmuştur. b değerleri arasındaki farklar incelendiğinde, zaman içerisindeki değişimi anlamlı bulundu ($F=233,689$, $p<0,001$). Grupların ikili karşılaştırmalarında b0 ile b1 arasında ve b0 ile b2 arasında anlamlı bir fark varken ($p<0,001$), b1 ile b2 arasında p değeri 0,160 olarak bulundu (Tablo 4.13).

Tablo 4.13: Grup 1 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri

	T0	T1	T2	F	p
L	91,21±0,30 ^a	92,20±0,55 ^b	92,21±0,42 ^b	69,609	<0,001
a	1,68±0,24 ^c	1,71±0,27 ^c	1,37±0,22 ^d	51,038	<0,001
b	20,69±0,35 ^e	18,23±0,61 ^f	18,00±0,26 ^f	233,689	<0,001

p değerleri= a-b:<0,001

c-d:<0,001

e-f:<0,001

Grup 2’de L’ler arası istatistiksel olarak fark bulundu ($p<0,001$). Tüm ikili gruplar arasında anlamlı farklılıklar görüldü ($F=77,395$, $p<0,001$). Farklı zaman noktalarında ölçülen a değerleri arasında da istatistiksel olarak anlamlı bir fark vardı ($F=94,448$, $p<0,001$). İkili grup karşılaştırmasında ise tüm karşılaştırmalarda anlamlı farklılıklar gözlemlendi ($p<0,001$). b değerlerinin zaman içerisindeki değişimi incelendiğinde istatistiksel olarak anlamlı bir farklılık gözlemlendi ($F=39,179$, $p<0,001$). Grupların ikili karşılaştırmalarında b0 ile b1 arasında ve b0 ile b2 arasında anlamlı bir fark varken ($p<0,001$) b1 ile b2 arasında p değeri 0,183 olarak bulundu (Tablo 4.14).

Tablo 4.14: Grup 2 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri

	T0	T1	T2	F	p
L	87,05±0,36 ^a	85,26±0,45 ^b	85,81±0,40 ^c	77,395	<0,001
a	0,21±0,07 ^d	0,60±0,14 ^e	0,32±0,10 ^f	94,448	<0,001
b	35,75±0,27 ^g	35,02±0,49 ^h	35,21±0,34 ^h	39,179	<0,001

p değerleri= a-b:<0,001, a-c:<0,001, b-c:<0,001

d-e: <0,001, d-f:<0,001, e-f:<0,001

g-h:<0,001

Grup 3’de L değerleri arasında zamana bağlı istatistiksel olarak anlamlı bir değişim izlendi ($F=492,677$, $p<0,001$). Zamana bağlı ikili karşılaştırmalarda L0 ile L1 arasında ve L0 ile L2 arasında istatistiksel olarak anlamlı bir fark ($p<0,001$) mevcutken L1 ile L2 arasında ($p=0,849$) anlamlı bir fark bulunmamıştır. Farklı zaman noktalarında ölçülen a değerleri arasında da istatistiksel olarak anlamlı bir fark vardı ($F=81,676$, $p<0,001$). İkili grup karşılaştırmalarında a1 ile a2 arasında istatistiksel olarak anlamlı bir fark bulunmamışken ($p=0,054$), diğer gruplar arasında p değeri <0,001 bulundu. b değerleri normal dağılıma uymadığı için zamana bağlı değişimi test etmede Friedman testi kullanıldı. b değerleri arasındaki farklar incelendiğinde, zaman içerisindeki değişimi anlamlı bulundu ($p<0,001$). Grupların ikili karşılaştırmalarında b1 ile b2 arasında anlamlı bir farklılık görülmezken ($p=0,289$) diğer gruplarda p değeri <0,001’di (Tablo 4.15).

Tablo 4.15: Grup 3 için tekrarlanan ölçümlü ANOVA ve Friedman test istatistiği ve p değerleri

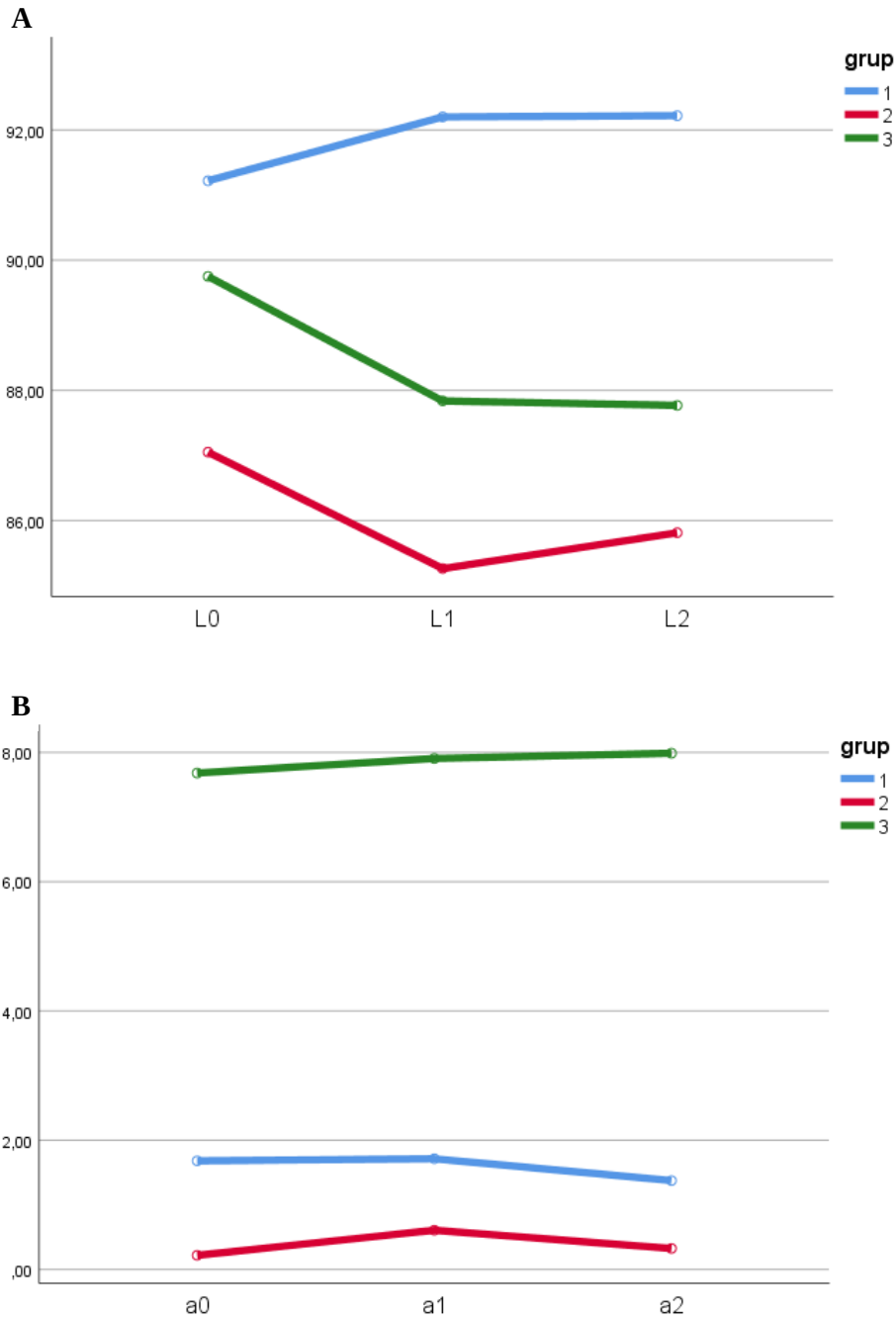
	T0	T1	T2	F	p
L	89,75±0,31 ^a	87,83±0,20 ^b	87,76±0,29 ^b	492,677	<0,001
a	7,68±0,54 ^c	7,90±0,77 ^d	7,98±0,71 ^d	81,676	<0,001
b	51,59±0,17 ^e	50,38±0,28 ^f	50,27±0,29 ^f	25,935	<0,001

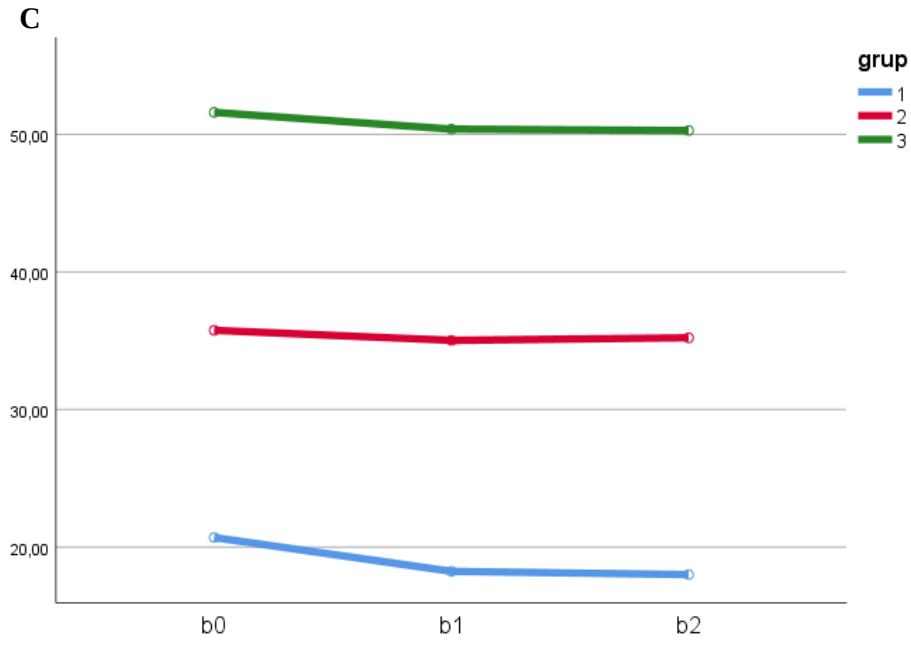
p değerleri=a-b:<0,001

c-d:<0,001

e-f:<0,001

L, a ve b değerlerinin gruplara ve zamana bağlı değişimi Şekil 4.3'te gösterildi.





Şekil 4.3: Örnek sayısı 16 iken renk koordinatlarında gruba ve zamana bağlı değişim
A) L, B) a, C) b

4.4. Örneklem Büyüklüğü 20 için L, a ve b Verilerinin İstatistiksel Analizi

L, a ve b değerlerinin tanımlayıcı istatistiği aritmetik ortalama ve standart sapma değerleri ile Tablo 4.16'da verildi.

Tablo 4.16: Örneklem büyüklüğü 20 için her grubun tanımlayıcı istatistik sonuçları ($\bar{x} \pm S$) ve çok değişkenli normal dağılım testlerinden Mardia küçük örneklem çarpıklık testinin p değerleri

Grup	L0	L1	L2	a0	a1	a2	b0	b1	b2
1	91,16 $\pm 0,33$	92,11 $\pm 0,53$	92,10 $\pm$ 0,46	1,72 $\pm$ 0,22	1,75 $\pm$ 0,26	1,40 $\pm$ 0,21	20,67 $\pm 0,34$	18,29 $\pm$ 0,65	17,98 $\pm 0,27$
2	87,03 $\pm 0,35$	85,26 $\pm 0,45$	85,81 $\pm$ 0,39	0,22 $\pm$ 0,07	0,59 $\pm$ 0,14	0,32 $\pm$ 0,09	35,69 $\pm 0,28$	34,94 $\pm$ 0,50	35,16 $\pm 0,36$
3	89,78 $\pm 0,30$	87,85 $\pm 0,21$	87,79 $\pm$ 0,27	7,67 $\pm$ 0,05	7,90 $\pm$ 0,08	7,98 $\pm$ 0,06	51,58 $\pm 0,18$	50,35 $\pm$ 0,27	50,28 $\pm 0,27$
	Mardia (L)			Mardia (a)			Mardia (b)		
1	0,051			0,268			0,454		
2	0,410			0,970			0,889		
3	0,446			0,054			0,183		

Her grup için L, a ve b değerlerinin dağılımının normalliği Mardia, Henze-Zirkler ve Doornik-Hansen testleri ile incelendi ve normal dağılıma uygun olduğu tespit edildi. Örnek olarak Mardia testinin p değerleri Tablo 4.16’da verildi. Küresellik Mauchly testi ile kontrol edildi (Tablo 4.17). Küresellik varsayımı sağlanmadığı durumlarda Greenhouse-Geisser düzeltmesi sonucu elde edilen F değerleri dikkate alındı.

Tablo 4.17: Her grup için Mauchly küresellik testinin sonuçları

Grup		Mauchly's W	Ki- kare	p	G- Geisser	Hyunh- Feldt	Lower- bound
1	L	0,914	1,625	0,444	0,921	1,000	0,500
	a	0,857	2,769	0,250	0,875	0,957	0,500
	b	0,659	7,511	0,023	0,746	0,795	0,500
2	L	0,609	8,940	0,011	0,719	0,761	0,500
	a	0,447	14,480	0,001	0,644	0,671	0,500
	b	0,855	2,815	0,245	0,874	0,955	0,500
3	L	0,948	,955	0,620	0,951	1,000	0,500
	a	0,725	5,781	0,056	0,784	0,843	0,500
	b	0,927	1,372	0,504	0,932	1,000	0,500

Tüm gruplarda verilerin dağılımı normal dağılıma uyduğu için tekrarlanan ölçümlü ANOVA testi yapılmıştır. Bu testin sonuçlarına göre Grup 1 için farklı zaman noktalarında ölçülen L değerleri arasında istatistiksel olarak fark bulundu ($F=91,254$, $p<0,001$). Zamana bağlı değişimi incelemek için yapılan ikili karşılaştırmalarda L0 ile L1 arasında ve L0 ile L2 arasında istatistiksel olarak anlamlı bir fark ($p<0,001$) mevcutken L1 ile L2 arasında ($p=1,000$) anlamlı bir fark bulunmamıştır. Farklı zaman noktalarında ölçülen a değerleri arasında da istatistiksel olarak anlamlı bir fark vardı ($F=79,155$, $p<0,001$). İkili grup karşılaştırmasında a0 ile a1 arasında anlamlı bir fark bulunmamışken ($p=0,802$) a1 ile a2 arasında ve a0 ile a2 arasında p değeri $<0,001$ bulunmuştur. b değerleri arasındaki farklar incelendiğinde, zaman içerisindeki değişimi anlamlı bulundu ($F=273,634$, $p<0,001$). Tüm ikili karşılaştırmalarda istatistiksel olarak anlamlı farklar bulundu. b0 ile b1 arasında ve b0 ile b2 arasında p değeri $<0,001$ iken, b1 ile b2 arasında p değeri 0,035 olarak bulundu (Tablo 4.18).

Tablo 4.18: Grup 1 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri

	T0	T1	T2	F	p
L	91,16±0,33 ^a	92,11±0,53 ^b	92,10±0,46 ^b	91,254	<0,001
a	1,72±0,22 ^c	1,75±0,26 ^c	1,40±0,21 ^d	79,155	<0,001
b	20,67±0,34 ^e	18,29±0,65 ^f	17,98±0,27 ^g	273,634	<0,001

p değerleri: a-b:<0,001

c-d:<0,001

e-f:<0,001, e-g:<0,001, f-g:0,035

Grup 2’de L’ler arası istatistiksel olarak anlamlı bir fark bulundu (F=112,440, p<,001). Tüm ikili gruplar arasında anlamlı farklılıklar görüldü (p<0,001). Farklı zaman noktalarında ölçülen a değerleri arasında da istatistiksel olarak anlamlı bir fark vardı (F=113,874, p<0,001). İkili grup karşılaştırmasında ise tüm karşılaştırmalarda anlamlı farklılıklar gözlemlendi (p<0,001). b değerlerinin zaman içerisindeki değişimi incelendiğinde istatistiksel olarak anlamlı bir farklılık gözlemlendi (F=43,932, p<0,001). Tüm ikili karşılaştırmalarda istatistiksel olarak anlamlı farklar bulundu. b0 ile b1 arasında ve b0 ile b2 arasında p değeri <0,001 iken, b1 ile b2 arasında p değeri 0,044 olarak bulundu (Tablo 4.19).

Tablo 4.19: Grup 2 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri

	T0	T1	T2	F	p
L	87,03±0,35 ^a	85,26±0,45 ^b	85,81±0,39 ^c	112,44	<0,001
a	0,22±0,07 ^c	0,59±0,14 ^d	0,32±0,09 ^e	113,874	<0,001
b	35,69±0,28 ^f	34,94±0,50 ^g	35,16±0,36 ^h	43,932	<0,001

p değerleri: a-b:<0,001, a-c:<0,001, b-c:<0,001

c-d:<0,001, c-e:<0,001, d-e:<0,001

f-g:<0,001, f-h:<0,001, g-h:0,044

Grup 3’de L değerleri arasında zamana bağlı istatistiksel olarak anlamlı bir değişim izlendi (F=722,623, p<0,001). Zamana bağlı ikili karşılaştırmalarda L0 ile L1 arasında ve L0 ile L2 arasında istatistiksel olarak anlamlı bir fark (p<0,001) mevcutken L1 ile L2 arasında (p=0,939) anlamlı bir fark bulunmamıştır. Farklı zaman noktalarında ölçülen a değerleri arasında da istatistiksel olarak anlamlı bir fark vardı (F=114,267, p<0,001). Tüm ikili karşılaştırmalarda anlamlı farklılıklar görüldü. a0 ile a1 arasında ve a0 ile a2 arasında p değeri <0,001 iken a1 ile a2 arasında bu değer 0,017’di. b değerleri arasındaki farklar incelendiğinde, zaman içerisindeki değişimi anlamlı bulundu (F=324,805, p<0,001). Grupların ikili karşılaştırmalarında b1 ile b2 arasında anlamlı bir farklılık görülmezken (p=0,664) diğer gruplarda p değeri <0,001’di (Tablo 4.20).

Tablo 4.20: Grup 3 için tekrarlanan ölçümlü ANOVA test istatistiği ve p değerleri

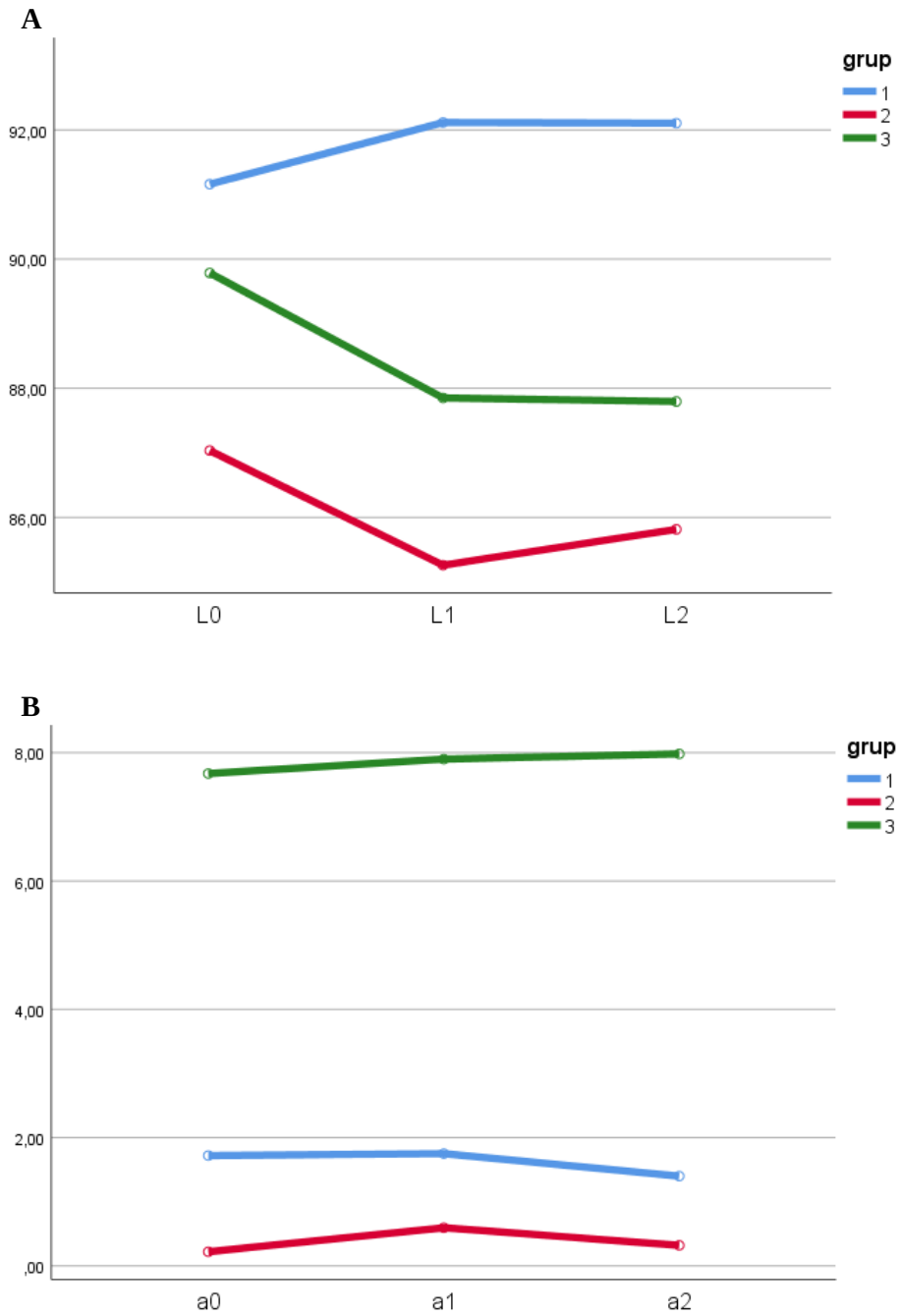
	T0	T1	T2	F	p
L	89,78±0,30 ^a	87,85±0,21 ^b	87,79±0,27 ^b	722,623	<0,001
a	7,67±0,05 ^c	7,90±0,08 ^d	7,98±0,06 ^e	114,267	<0,001
b	51,58±0,18 ^f	50,35±0,27 ^g	50,28±0,27 ^g	324,805	<0,001

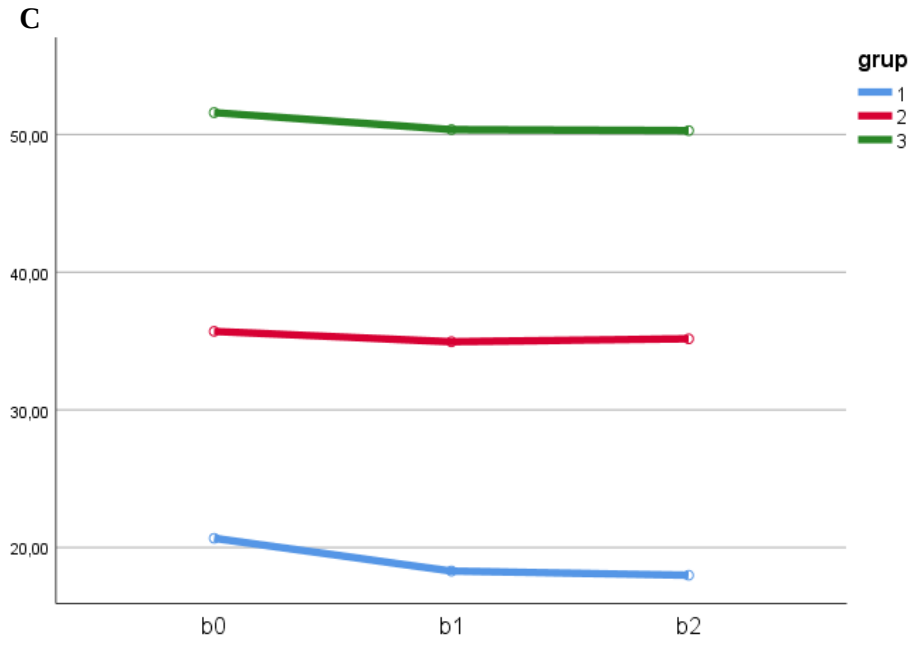
p değerleri: a-b:<0,001

c-d:<0,001, c-e:<0,001, d-e:0,017

f-g:<0,001

L, a ve b değerlerinin gruplara ve zamana bağlı değişimi Şekil 4.4'te gösterildi.





Şekil 4.4: Örnek sayısı 20 iken renk koordinatlarında gruba ve zamana bağlı değişim
A) L, B) a, C) b

4.5. Örneklem Büyüklüğü 8 için ΔE_{00} Verilerinin İstatistiksel Analizi

ΔE_{00} değerlerinin ortalama ve standart sapmaları içeren tanımlayıcı istatistik sonuçları Tablo 4.21’de verildi.

Tablo 4.21: Her grubun ΔE_{00} verilerinin ortalama ve standart sapma değerleri

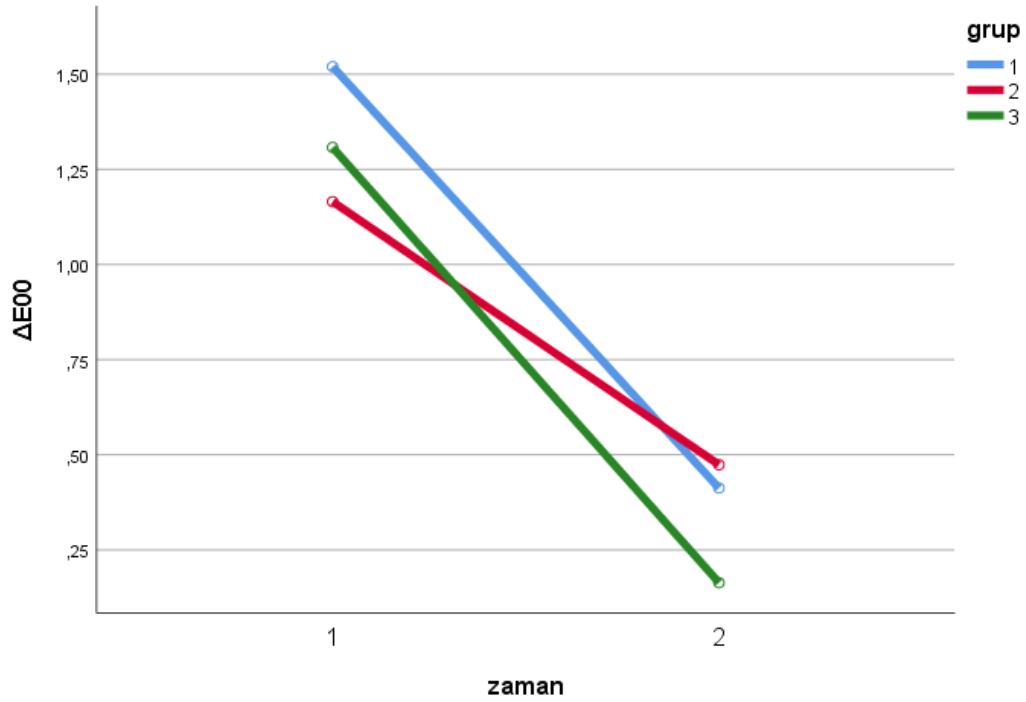
Grup	$\Delta E_{00(1)}$	$\Delta E_{00(2)}$
1	1,52 $\pm$ 0,35	0,41 $\pm$ 0,14
2	1,16 $\pm$ 0,42	0,47 $\pm$ 0,15
3	1,30 $\pm$ 0,16	0,16 $\pm$ 0,07
Mardia (ΔE_{00})		
1	0,309	
2	0,947	
3	0,424	

Çok değişkenli normal dağılıma uygunluğu test etmek için yapılan Mardia, Henze-Zirkler ve Doornik-Hansen testleri ile 3 grup için de dağılımın normalliği tespit edildi. Örnek olarak Mardia küçük örneklem çarpıklık testinin p değerleri Tablo 4.21’de verildi. Tekrarlanan ölçümlü ANOVA sonuçlarına göre hem ΔE_{00} ($F=224,555$, $p<0,001$) hem de $\Delta E_{00}*\text{grup}$ ($F=4,919$, $p=0,018$) etkileşimi anlamlı bulundu. Grup içi karşılaştırmalarda $\Delta E_{00(1)}$ ile $\Delta E_{00(2)}$ arasındaki ortalama fark 0,981, standart hata 0,065 ve p değeri $<0,001$ ’di. $\Delta E_{00}*\text{grup}$ etkileşiminde $\Delta E_{00(1)}$ ’de gruplar arasında anlamlı bir farklılık gözlenmezken $\Delta E_{00(2)}$ ’de Grup1 ile Grup3 arasında ve Grup2 ile Grup3 arasında istatistiksel olarak anlamlı bir farklılık gözlemlendi (Tablo 4.22).

Tablo 4.22: Gruplar arası ikili karşılaştırmalarda ortalama fark ve p değerleri (* $p<0,05$)

ΔE_{00}	Grup	Grup	Ortalama Fark	Standart Hata	p değeri	Farklar için %95 güven aralığı	
						Alt sınır	Üst sınır
$\Delta E_{00(1)}$	1	2	0,355	0,168	0,139	-0,081	0,791
		3	0,213	0,168	0,656	-0,223	0,648
	2	1	-0,355	0,168	0,139	-0,791	0,081
		3	-0,142	0,168	1,000	-0,578	0,293
	3	1	-0,213	0,168	0,656	-0,648	0,223
		2	0,142	0,168	1,000	0,293	0,578
$\Delta E_{00(2)}$	1	2	-0,061	0,063	1,000	-0,226	0,103
		3	0,249	0,063	0,002*	-0,084	0,413
	2	1	0,061	0,063	1,000	-0,103	0,226
		3	0,310	0,063	$<0,001^*$	0,145	0,475
	3	1	-0,249	0,063	0,002*	-0,413	-0,084
		2	-0,310	0,063	$<0,001^*$	-0,475	-0,145

Her grubun zaman içerisindeki değişim grafiği Şekil 4.5'te gösterildi.



Şekil 4.5: Her grubun ΔE_{00} değerindeki zamana bağlı değişim (n=8)

4.6. Örneklem Büyüklüğü 12 için ΔE_{00} Verilerinin İstatistiksel Analizi

ΔE_{00} değerlerinin ortalama ve standart sapmaları içeren tanımlayıcı istatistik sonuçları Tablo 4.23'te verildi.

Tablo 4.23: Her grubun ΔE_{00} verilerinin ortalama ve standart sapma değerleri

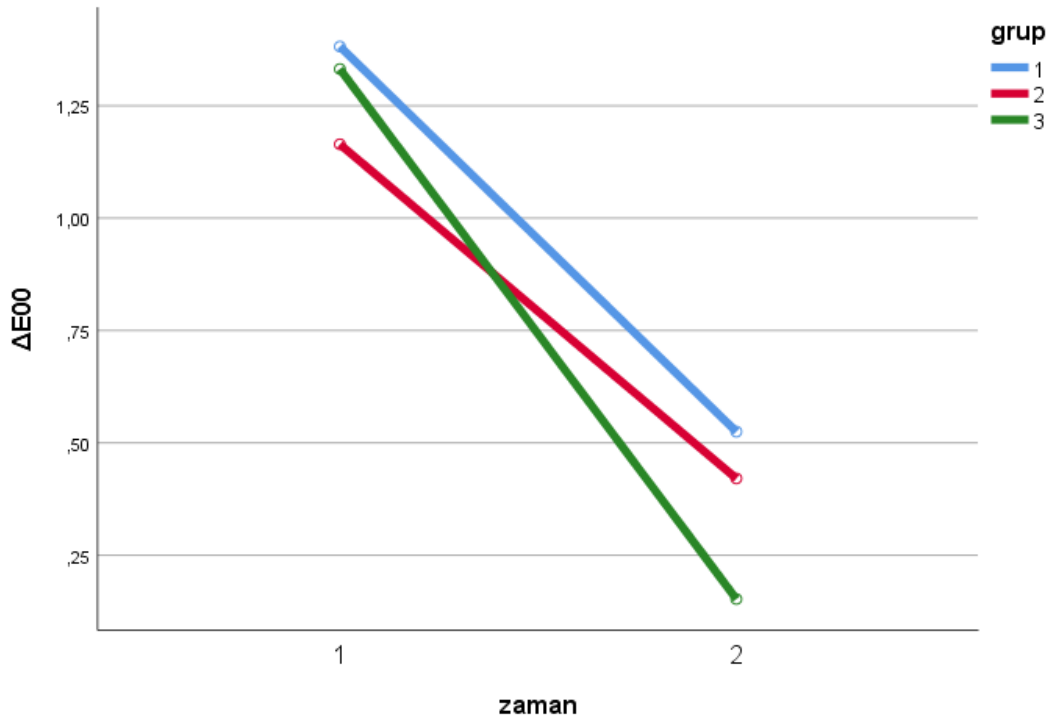
Grup	$\Delta E_{00(1)}$	$\Delta E_{00(2)}$
1	1,38 $\pm$ 0,44	0,52 $\pm$ 0,16
2	1,16 $\pm$ 0,42	0,42 $\pm$ 0,15
3	1,33 $\pm$ 0,16	0,15 $\pm$ 0,09
Mardia (ΔE_{00})		
1	0,703	
2	0,890	
3	0,395	

Çok değişkenli normal dağılıma uygunluğu test etmek için yapılan Mardia, Henze-Zirkler ve Doornik-Hansen testlerinin sonuçları incelendi. Örnek olarak Mardia küçük örneklem çarpıklık testinin p değerleri Tablo 4.23'te verildi. ΔE_{00} verilerinin dağılımı normal dağılıma uygunluk gösterdiği için yapılan tekrarlanan ölçümlü ANOVA sonuçlarına göre hem ΔE_{00} ($F=185,336$, $p<0,001$) hem de $\Delta E_{00}*\text{grup}$ ($F=3,681$, $p=0,036$) etkileşimi anlamlı bulundu. İkili karşılaştırmalarda $\Delta E_{00(1)}$ ile $\Delta E_{00(2)}$ arasındaki ortalama fark 0,926, standart hata 0,068 ve p değeri $<0,001$ 'di. $\Delta E_{00}*\text{grup}$ etkileşiminde $\Delta E_{00(1)}$ 'de gruplar arasında anlamlı bir farklılık gözlenmezken ($p>0,05$), $\Delta E_{00(2)}$ 'de Grup1 ile Grup3 arasında ve Grup2 ile Grup3 arasında istatistiksel olarak anlamlı bir farklılık gözlemlendi (Tablo 4.24).

Tablo 4.24: Gruplar arası ikili karşılaştırmalarda ortalama fark ve p değerleri (* $p<0,05$)

ΔE_{00}	Grup	Grup	Ortalama Fark	Standart Hata	p değeri	Farklar için %95 güven aralığı	
						Alt sınır	Üst sınır
$\Delta E_{00(1)}$	1	2	0,217	0,150	0,466	-0,160	0,595
		3	0,050	0,150	1,000	-0,327	0,427
	2	1	-0,217	0,150	0,466	-0,595	0,160
		3	-0,167	0,150	0,812	-0,545	0,210
	3	1	-0,050	0,150	1,000	-0,427	0,327
		2	0,167	0,150	0,812	-0,210	0,545
$\Delta E_{00(2)}$	1	2	0,104	0,057	0,227	-0,039	0,247
		3	0,372	0,057	$<0,001^*$	0,229	0,516
	2	1	-0,104	0,057	0,227	-0,247	0,039
		3	0,268	0,057	$<0,001^*$	0,125	0,412
	3	1	0,372	0,057	$<0,001^*$	-0,516	-0,229
		2	-0,268	0,057	$<0,001^*$	-0,412	-0,125

Her grubun zaman içerisindeki değişim grafiği Şekil 4.6’da gösterildi.



Şekil 4.6: Her grubun ΔE_{00} değerindeki zamana bağlı değişim (n=12)

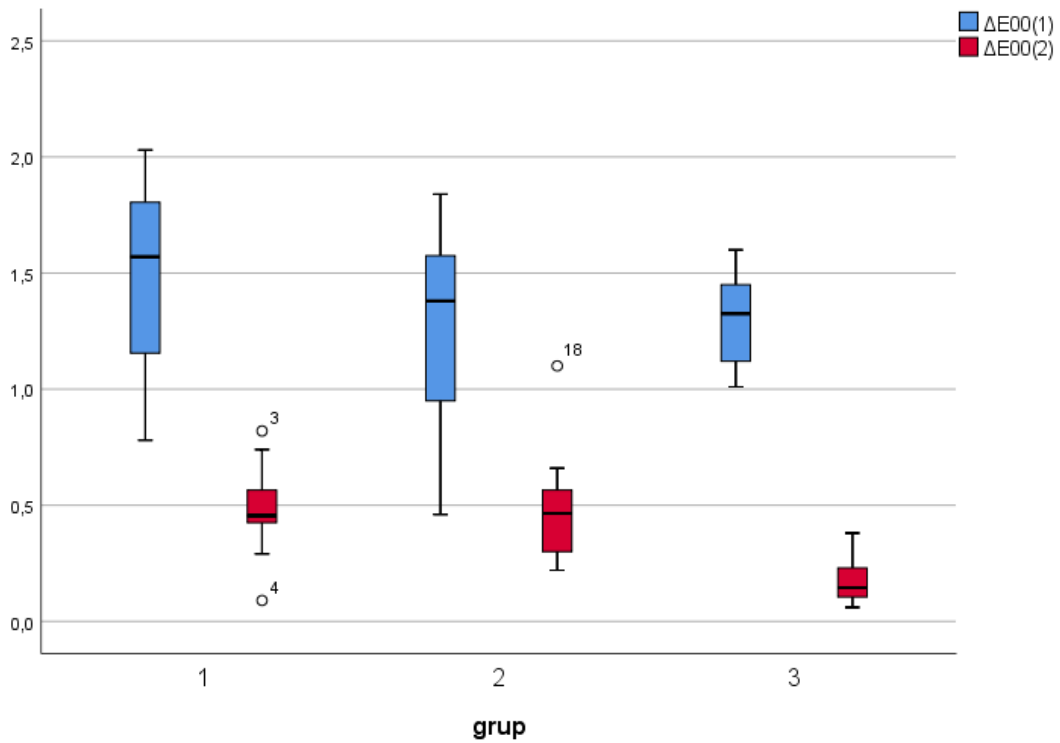
4.7. Örneklem Büyüklüğü 16 için ΔE_{00} Verilerinin İstatistiksel Analizi

ΔE_{00} değerlerinin ortalama, standart sapma, medyan ile 1. ve 3. çeyreklikleri içeren tanımlayıcı istatistik sonuçları Tablo 4.25’te verildi.

Tablo 4.25: Her grubun ΔE_{00} değerlerinin tanımlayıcı istatistik verileri

Grup	ΔE_{00}	Ortalama	Standart sapma	Medyan	1. Çeyreklik	3. Çeyreklik
1	$\Delta E_{00(1)}$	1,47	0,40	1,57	1,15	1,82
	$\Delta E_{00(2)}$	0,47	0,16	0,45	0,42	0,57
2	$\Delta E_{00(1)}$	1,25	0,43	1,38	0,90	1,64
	$\Delta E_{00(2)}$	0,47	0,21	0,46	0,29	0,57
3	$\Delta E_{00(1)}$	1,30	0,19	1,32	1,11	1,45
	$\Delta E_{00(2)}$	0,17	0,09	0,14	0,10	0,23

Çok değişkenli normal dağılıma uygunluğu Mardia, Henze-Zirkler ve Doornik-Hansen testleri ile incelendi. Grup1 ve Grup2 normal dağılım göstermediği için $\Delta E_{00(1)}$ ve $\Delta E_{00(2)}$ değerleri arasındaki farklar Wilcoxon işaretli sıralar testi ile değerlendirildi. Grup3 için ise bağımlı örneklem t-testi gerçekleştirildi. $\Delta E_{00(1)}$ ve $\Delta E_{00(2)}$ arası farklar, 3 grupta da istatistiksel olarak anlamlı farklar gösterdi ($p < 0,001$). Farklı gruplar arasında $\Delta E_{00(1)}$ ve $\Delta E_{00(2)}$ 'i ayrı ayrı karşılaştırmak için Kruskal-Wallis testi uygulandı ($p < 0,001$). $\Delta E_{00(1)}$ 'de gruplar arasında anlamlı bir farklılık gözlenmezken ($p = 0,246$), $\Delta E_{00(2)}$ 'de Grup1 ile Grup3 arasında ve Grup2 ile Grup3 arasında istatistiksel olarak anlamlı bir farklılık gözlemlendi ($p < 0,001$) (Şekil 4.7).



Şekil 4.7: Örnek sayısı 16 iken 3 grubun $\Delta E_{00(1)}$ ve $\Delta E_{00(2)}$ değerlerinin kutu (box-plot) grafiği

4.8. Örneklem Büyüklüğü 20 için ΔE_{00} Verilerinin İstatistiksel Analizi

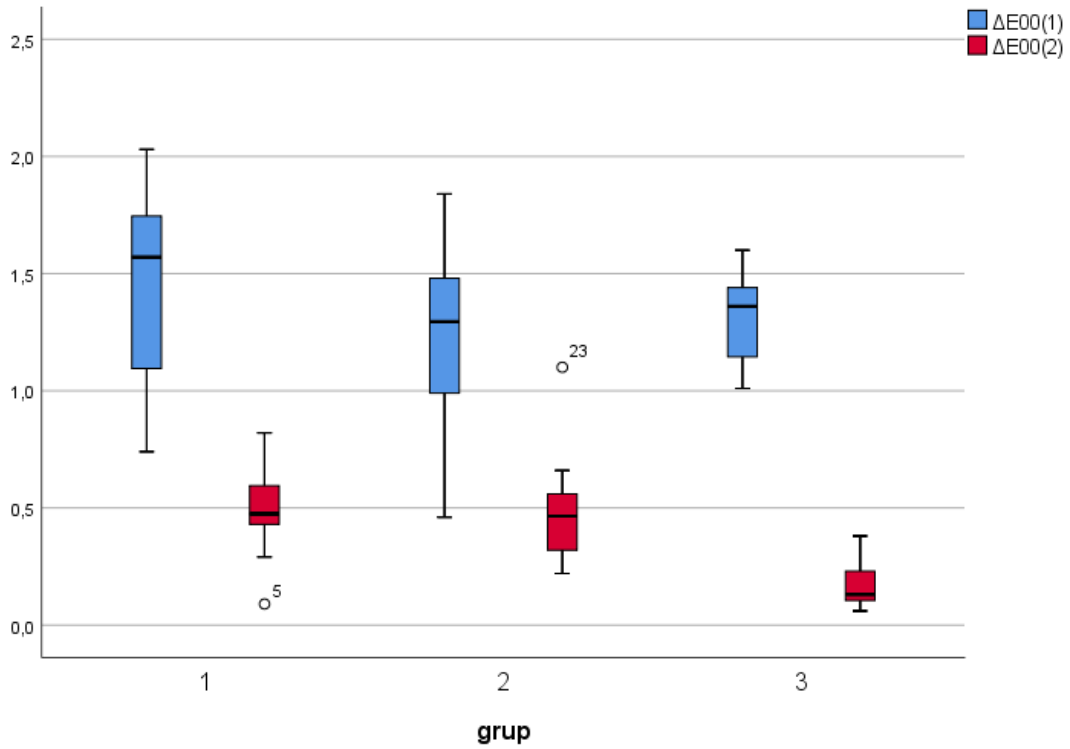
ΔE_{00} verilerinin ortalama, standart sapma, medyan ile 1. ve 3. çeyreklikleri içeren tanımlayıcı istatistik sonuçları Tablo 4.26'da verildi.

Tablo 4.26: Her grubun ΔE_{00} değerlerinin tanımlayıcı istatistik verileri

Grup	ΔE_{00}	Ortalama	Standart sapma	Medyan	1. Çeyreklik	3. Çeyreklik
1	$\Delta E_{00(1)}$	1,43	0,40	1,57	1,06	1,75
	$\Delta E_{00(2)}$	0,50	0,16	0,47	0,43	0,60
2	$\Delta E_{00(1)}$	1,24	0,39	1,29	0,96	1,50
	$\Delta E_{00(2)}$	0,46	0,19	0,46	0,31	0,56
3	$\Delta E_{00(1)}$	1,31	0,17	1,36	1,13	1,44
	$\Delta E_{00(2)}$	0,16	0,08	0,13	0,10	0,23

Çok değişkenli normal dağılıma uygunluğu Mardia, Henze-Zirkler ve Doornik-Hansen testleri ile incelendi. Grup1 ve Grup2 normal dağılım göstermediği için $\Delta E_{00(1)}$ ve $\Delta E_{00(2)}$ değerleri arasındaki farklar Wilcoxon testi ile değerlendirildi. Grup3 için ise bağımlı örneklem t-testi gerçekleştirildi. $\Delta E_{00(1)}$ ve $\Delta E_{00(2)}$ arası farklar, 3 grupta da istatistiksel olarak anlamlı farklılıklar gösterdi ($p < 0,001$).

Farklı gruplar arasında $\Delta E_{00(1)}$ ve $\Delta E_{00(2)}$ 'i ayrı ayrı karşılaştırmak için Kruskal-Wallis testi uygulandı. $\Delta E_{00(1)}$ 'de gruplar arasında anlamlı bir farklılık gözlenmezken ($p = 0,305$), $\Delta E_{00(2)}$ 'de Grup1 ile Grup3 arasında ve Grup2 ile Grup3 arasında istatistiksel olarak anlamlı bir farklılık gözlemlendi ($p < 0,001$) (Şekil 4.8).



Şekil 4.8: Örnek sayısı 20 iken 3 grubun $\Delta E_{00(1)}$ ve $\Delta E_{00(2)}$ değerlerinin kutu (box-plot) grafiği

4.9. Deneysel (Post-hoc) Güç Analizi

ΔE_{00} verilerinin zaman ve grup etkisindeki değişimi için yapılan tekrarlanan ölçümler ANOVA sonuçlarından yola çıkılarak deneysel (post hoc) güç analizi G*Power yazılımı (versiyon 3.1.9.4) kullanılarak yapıldı. Hesaplamalarda zaman*grup etkileşiminin parsiyel eta kare (η^2) değerleri kullanıldı. Her örneklem büyüklükleri için elde edilen güç değerleri Tablo 4.27’de verildi.

Tablo 4.27: ΔE_{00} verilerinin analizi sonucu elde edilen güç değerleri

Örnek sayısı (n)	Parsiyel eta kare (η^2)	Etki büyüklüğü	Güç
8	0,319	0,503	0,99
12	0,182	0,471	0,99
16	0,134	0,393	0,99
20	0,151	0,421	0,99

L, a ve b verilerinin her materyal grubunda zaman içerisindeki değişiminin karşılaştırıldığı tekrarlanan ölçümler ANOVA sonuçları kullanılarak benzer şekilde deneysel güç analizi yapıldı. Tüm gruptaki en düşük parsiyel eta kare değerleri dikkate alınarak çalışmanın gücü hesaplandı (Tablo 4.28).

Tablo 4.28: L, a ve b verilerinin analizi sonucu elde edilen güç değerleri

Örnek sayısı (n)	Renk eksen	Parsiyel eta kare (η^2)	Etki büyüklüğü	Güç
8	L	0,832	2,225	1,00
	a	0,717	1,591	1,00
	b	0,665	1,408	1,00
12	L	0,777	1,866	1,00
	a	0,821	2,141	1,00
	b	0,727	1,631	1,00
16	L	0,823	2,156	1,00
	a	0,773	1,845	1,00
	b	0,723	1,615	1,00
20	L	0,828	2,194	1,00
	a	0,806	2,038	1,00
	b	0,698	1,520	1,00

Hem ΔE_{00} hem de L, a ve b değerleri üzerinden yapılan güç hesaplamasında hedeflenen 0,95'ten daha yüksek değerler elde edildi.

5. TARTIŞMA

İstatistik biliminde iki tip hata vardır. Doğru değilken alternatif hipotezin (H1) kabulü Tip 1 hata (yanlış pozitif) olarak adlandırılır. Alfa (α), Tip 1 hata olasılığını tanımlar. Tip 1 hatanın oluşmasına, popülasyondan oldukça farklı bir örneklem seçiminden araştırma prosedürlerinin uygulanmasında meydana gelen diğer hatalara kadar birçok neden etki edebilir. Yanlışlıkla alternatif hipotezin (H1) reddedilip sıfır hipotezin (H0) kabulü ise Tip 2 hata (yanlış negatif) olarak adlandırılır. Beta (β), tip 2 hata olasılığını belirtir. Tip 2 hata oluşumunun en yaygın nedeni, özellikle düşük etki büyüklüğünün eşlik ettiği küçük örneklem büyüklüğüdür. Tip 2 hata yapma olasılığının azalması, Tip 1 hata yapma riskini arttırdığı için, Tip 1 ve Tip 2 hatalar arasında hassas bir denge kurulmalıdır. 0,05 ya da 0,01 gibi düşük bir Tip 1 hata ve 0,8 ya da 0,9 kadar yüksek bir güç elde etmek için yeterli örneklem büyüklüğü korunmalıdır. Bununla birlikte son yıllarda “maliyet-etkin örneklem büyüklüğü” kavramı önem kazanmıştır. Tip 2 hatayı azaltmak için örneklem büyüklüğünün artırılması önerilmekle birlikte bu işlem, çalışmanın maliyetini arttırabilir ya da araştırmanın öngörülen süre içinde tamamlanmasını geciktirebilir. Gereğinden fazla özellikle hayvan çalışmalarında etik sorunlara yol açabileceği de göz önünde tutulmalıdır. Sonuçta etkin örneklem büyüklüğünün belirlenmesi, çalışma sonuçlarının etkisini arttırarak verimli ve değerli bir çalışma ortaya çıkarmak için oldukça önemlidir.³⁸ Bu nedenle bu çalışmada, örneklem büyüklüğünün çalışma sonuçlarına etkisini incelemek için 4 farklı küçük örneklem büyüklüğünde (n:8, n:12, n:16, n:20) zamana bağlı tekrarlanan ölçümler yapıldı. Bu ölçümler sonucunda başlangıç (T0), 2500 termal siklusa maruziyet sonrası (T1) ve 5000 termal siklusa maruziyet sonrası (T2), rengin 3 koordinatının (L, a ve b) nicel değerleri kaydedildi. Bu değerler kullanılarak T0-T1 arası ($\Delta E_{00(1)}$) ve T1-T2 arası ($\Delta E_{00(2)}$) materyallerdeki renk değişimi (ΔE_{00}) hesaplandı. Klinikte geçici kron materyali olarak kullanılan fakat farklı dijital üretim yöntemleri ile üretilen 3 farklı materyal üzerinde tüm işlemler benzer şekilde tamamlandı. Hem L, a, b değerleri hem de ΔE_{00} değerleri üzerinden istatistik testler uygulandı. Bu çalışmanın sonuçlarına göre ΔE_{00} değerleri, incelenen tüm örneklem büyüklüklerinde zamana bağlı anlamlı farklılık gösterdi yani $\Delta E_{00(1)}$ ve $\Delta E_{00(2)}$ arasında istatistiksel olarak anlamlı bir farklılık vardı. ΔE_{00} ve grup etkileşiminde ikili karşılaştırmalarda ise $\Delta E_{00(1)}$ değerlerinde gruplar arası anlamlı bir farklılık yokken $\Delta E_{00(2)}$ incelendiğinde tüm örneklem büyüklüklerinde Grup1 ile Grup3

ve Grup2 ile Grup3 arasındaki değerler istatistiksel olarak farklıydı. Grup1 ile Grup2 arasında ise benzer sonuçlar elde edildi. Bunun nedeni Grup1 ile Grup2'nin 3 boyutlu yazıcılar ile üretilmesi, Grup3'ün ise kazıma yöntemi ile üretilmesi olabilir. Kullanılan materyale ve zamana bağlı farklılıklar bulunsada örneklem büyüklüğünün test sonuçlarını etkilemediği görüldü.

Farklı örneklem büyüklüklerinde L, a ve b değerlerinin zamana bağlı değişimi her grup için ayrı ayrı incelendiğinde benzer test sonuçları elde edildi. Her grubun L, a ve b değerleri zamana bağlı istatistiksel olarak anlamlı farklılıklar gösterdi. Bu farklılıklar örneklem büyüklüğünden etkilenmedi. Fakat ikili karşılaştırmalar incelendiğinde, Grup1 ve Grup2'de örnek sayısı 8, 12 ve 16 olduğunda b1 ve b2 değerleri arasında anlamlı bulunmazken örnek sayısı 20 olduğunda bu değerler arasında istatistiksel olarak anlamlı bir fark gözlemlendi. Diğer farklılık ise Grup3'te a1 ile a2 değerleri arasında gözlemlendi. Örnek sayısı 8 ve 16 iken a1 ile a2 arasında anlamlı bir fark yokken, örnek sayısı 12 ve 20 iken bu fark istatistiksel olarak anlamlı bulundu. Diğer tüm karşılaştırmalar örneklem büyüklüğü değişse de benzer sonuçlar gösterdi.

Huang ve arkadaşlarının³⁹ yaptığı bir çalışmada, farklı küçük örneklem büyüklüklerinde (5, 10, 20, 30 ve 50) göz verileri model tabanlı (marjinal model ve karma etkiler modeli) ve model tabanlı olmayan (iki örneklem t-testi ve eşleştirilmiş t testi) istatistiksel yöntemlerle analiz edildi. Testlerin performansının, çalışmanın tasarımına ve örneklem büyüklüğüne bağlı olduğu sonucuna varıldı. Eşleştirilmiş dizaynda, eşleştirilmiş t-testi ve doğrusal karma etkiler modeli, örneklem boyutu 10'dan küçük olduğunda bile iyi sonuç verdiği rapor edilmiştir (Tip 1 hata oranı 0,05'e yakın değer göstermiştir). Eşleştirilmemiş modelde doğrusal karma etkiler modeli, örneklem büyüklüğü 50 kadar büyük olduğunda bile Tip 1 hata oranını önemli ölçüde arttırdı. Bizim çalışmamızda ise örneklem büyüklüğü değiştirilirken her zaman tekrarlanan ölçümler dizaynı kullanıldı. Boedec ve arkadaşları⁴⁰ ise örnek sayısı 30 ve 60 iken referans popülasyonun yüzdelik sınırlarının tahmini (RI) doğruluğunu inceledi. Örneklem boyutu küçük olduğunda (n=30) normallik testlerinin düşük spesifikiteye sahip olması nedeniyle ham veriler üzerinde parametrik testlerin uygunsuz kullanımına ve hatalı RI'ye yol açabileceği sonucuna varıldı. Bu çalışma sonucundan farklı olarak bizim çalışmamızda verilerin normalliği, çok değişkenli normallik testleri (Mardia'nın çarpıklık- basıklık,

Henze-Zirkler, Doornik-Hansen testleri) kullanılarak değerlendirildi. Elde edilen sonuçlara göre parametrik ya da parametrik olmayan testlerden biri kullanıldı.

Çalışmanın başlangıcında yapılan güç analizinde etki büyüklüğü 1,36 iken 0,95 güç hedeflenerek her grup için örnek sayısı 5 hesaplandı. Bu çalışmada gruplar 8'den başlatıldı ve düzenli bir artış ile 20'e ulaşıldı. Toplam 4 farklı örneklem büyüklüğü ile yapılan renk değişimi analizinde, hedeflenen güce ulaşıp ulaşılmadığını ve örneklem sayısına bağlı bir değişimin olup olmadığını test etmek için post hoc (deneysel) güç analizi yapıldı. Hesaplanan güçlerde örneklem büyüklüğüne bağlı bir değişiklik saptanmadı. Tüm hesaplanan güçler 0,99'un üzerindeydi. Küçük örneklem büyüklüklerine rağmen elde edilen bu sonuç, örneklerin üretimi, yaşlandırılması ve renklerin ölçüldüğü ortamın fiziksel özelliklerinin standardize edilmiş olmasından ve herhangi bir karıştırıcı faktörün bulunmamasından kaynaklanmış olabilir. Bunun yanı sıra hesaplanan en düşük parsiyel eta kare değeri 0,134, etki büyüklüğü değeri ise 0,393'tü. Bu da hesaplanan güçlerin yüksek çıkmasına neden olmuştur.

Uygun bir örneklem büyüklüğü, araştırma sonuçlarını daha etkili hale getirir, üretilen veriler güvenilir, kaynak yatırımı sınırlıdır ve etik ilkelere uygundur. Çok küçük örneklem, bir çalışmanın iç ve dış geçerliliğine zarar verebilir. Çok büyük örneklem, klinik olarak önemsiz olan küçük farklılıkları istatistiksel olarak anlamlı farklar olarak gösterebilir. Bu da hem araştırmacıları hem de klinisyenleri yanlış yönlendirerek tedavi kararlarında başarısızlığa sebep olabilir.⁴¹ Bu nedenle bu çalışmada farklı örneklem büyüklüklerinin, sonuçların istatistiksel anlamlılığı üzerinde etkisi olup olmadığı araştırıldı. Sonuçların örneklem büyüklüklerinden anlamlı derecede etkilendiği ve tüm veri analizlerinde 0.95'ten daha yüksek güce ulaşıldığı tespit edildi. Bu çalışmanın limitasyonları olarak bir örnek uygulaması yapılması, yüksek etki büyüklüğüne sahip olunması ve örneklem büyüklüğünün 8 ile 20 arasında kısıtlı tutulması sayılabilir. İleriki çalışmalarda farklı etki büyüklüğüne sahip olan küçük örneklemde simülasyon çalışmaları yapılarak elde edilen sonuçlar doğrulanabilir. Etki büyüklüğünün küçük olduğu çalışmalarda örnek sayısının arttırılmasına ihtiyaç duyulabilir.

KAYNAKLAR

1. Altman DG, Bland JM. Comparing several groups using analysis of variance. *BMJ*. 1996;312(7044):1472.
2. Bergh DD. Problems with repeated measures analysis: demonstration with a study of the diversification and performance relationship. *Academy of Management Journal* 1995;38(6):1692-1708.
3. Dawson B, Trapp RG. Basic & clinical biostatistics. McGraw-Hill; 2004.
4. Pereira DG, Afonso A, Medeiros FM. Overview of friedmans test and post-hoc analysis. *Commun Stat Simul Comput*. 2015;44(10):2636-2653.
5. Zimmerman DW, Zumbo BD. Relative power of the wilcoxon test, the friedman test, and repeated-measures ANOVA on ranks. *Journal of Experimental Education* 2010;62(1):75-86.
6. Yaşar ED. Örneklemin Temsil Ediciliği. *Marmara Üniversitesi Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi* 1992; 73-95.
7. Park E, Cho M, Ki CS. Correct use of repeated measures analysis of variance. *Korean J Lab Med*. 2009;29(1):1-9.
8. Zach. The three assumptions of the repeated measures ANOVA. <https://www.statology.org/repeated-measures-anova-assumptions/>
9. Davis CS. Statistical methods for the analysis of repeated measurements. Springer; 2002.
10. Wishart J. Growth-rate determinations in nutrition studies with the bacon pig, and their analysis. *Biometrika*. 1938;30(1-2):16-28.
11. Johnson RA, Wichern DW. Applied multivariate statistical analysis. 6th edition. Pearson; 2007.
12. Huynh H, Feldt LS. Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. *J Am Stat Assoc*. 1970;65(332):1582-1589.

13. Maxwell SE, Delaney HD. Designing experiments and analyzing data : a model comparison perspective. 2nd edition. Routledge; 2003.
14. Huynh H, Feldt LS. Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics* 1976;1(1):69:82.
15. Eyduran E. Tekrarlanan ölçümlü denemelerin analizinde kullanılan farklı yaklaşımların karşılaştırılması. Doktora tezi;2009.
16. Atkinson G. Analysis of repeated measurements in physical therapy research. *Physical Therapy in Sport*. 2001;2(4):194-208.
17. How to perform a two-way repeated measures ANOVA in SPSS Statistics | Laerd Statistics. <https://statistics.laerd.com/spss-tutorials/two-way-repeated-measures-anova-using-spss-statistics.php>.
18. Gürbüz F, Başpınar E, Çamdeviren H, Keskin S. Tekrarlanan Ölçümlü Deneme Düzenlerinin Analizi. Yüzüncü Yıl Üniversitesi Matbaası; 2003.
19. Demšar J. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*. 2006;7:1-30.
20. Bülbül SE. Kruskal-wallis testi ve Friedman testinin alternatif parametrik tekniklerle karşılaştırılması ve bazı parametrik ve parametrik olmayan çoklu karşılaştırma yöntemleri ile incelenmesi. *Öneri Dergisi*. 2001;4(15):89-96.
21. Siegel S, Castellan NJ. Nonparametric statistics for the behavioral sciences. 2nd edition. McGraw-Hill;1988.
22. Conover WJ. Practical nonparametric statistics. Jon Wiley&Sons;1999.
23. Sheskin D. Handbook of parametric and nonparametric statistical procedures. 5th edition. Chapman and Hall;2011.
24. Wright SP. Adjusted P-values for simultaneous inference. *Biometrics* 1992;48(4):1005.

25. Doğan N, Doğan İ. Birinci tür hatanın kontrolü ve adımsal (stepwise) çoklu karşılaştırma testleri. *Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi*. 2014;1(1):28-33.
26. Kilkenny C, Parsons N, Kadyszewski E, Festing MFW, Cuthill IC, Fry D, et al. Survey of the quality of experimental design, statistical analysis and reporting of research using animals. *PLoS One*. 2009;4(11):e7824.
27. Festing MFW. Design and statistical methods in studies using animal models of development. *ILAR J*. 2006;47(1):5-14.
28. Glantz SA. Primer of Biostatistics. 7th edition. McGraw Hill Medical; 2011.
29. VIII. Ulusal Biyoistatistik Kongresi. <https://www.halksagligiokulu.org/Kitap;2022>
30. Biau DJ, Kernéis S, Porcher R. Statistics in brief: the importance of sample size in the planning and interpretation of medical research. *Clin Orthop Relat Res*. 2008;466(9):2282-2288.
31. Hoenig JM, Heisey DM. The Abuse of Power. *The American Statistician* 2012;55(1):19-24.
32. Guo Y, Pandis N. Sample-size calculation for repeated-measures and longitudinal studies. *Am J Orthod Dentofacial Orthop*. 2015;147(1):146-149.
33. Molinero-Mourelle P, Canals S, Gómez-Polo M, Solá-Ruiz M, del Río Highsmith J, Viñuela A. Polylactic Acid as a Material for Three-Dimensional Printing of Provisional Restorations. *Int J Prosthodont*. 2018;31(4):349-350.
34. Aljanobi G, Al-Sowygh ZH. The effect of thermocycling on the translucency and color stability of modified glass ceramic and multilayer zirconia materials. *Cureus*. 2020;12(2):e6968.
35. Paravina RD, Ghinea R, Herrera LJ, et al. Color difference thresholds in dentistry. *J Esthet Restor Dent*. 2015;27 Suppl 1(S1):S1-S9.
36. Yao Q, Morton D, Eckert GJ, Lin WS. The effect of surface treatments on the color stability of CAD-CAM interim fixed dental prostheses. *J Prosthet Dent*. 2021;126(2):248-253.

37. Normal dağılım inceleme yazılımı. <http://biostatapps.inonu.edu.tr/NDY/>
38. Serdar CC, Cihan M, Yücel D, Serdar MA. Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. *Biochem Med (Zagreb)*. 2021;31(1):1-27.
39. Huang J, Huang J, Chen Y, Ying G shuang. Evaluation of approaches to analyzing continuous correlated eye data when sample size is small. *Ophthalmic Epidemiol*. 2018;25(1):45-54.
40. Boedec KL. Sensitivity and specificity of normality tests and consequences on reference interval accuracy at small sample size: a computer-simulation study. *Vet Clin Pathol*. 2016;45(4):648-656.
41. Faber J, Fonseca LM. How sample size influences research outcomes. *Dental Press J Orthod*. 2014;19(4):27-29.

HAM VERİLER

Grup	L0	L1	L2	a0	a1	a2	b0	b1	b2
1	90,9	91,8	91,7	1,9	1,9	1,4	20,7	19,1	18,3
1	91,1	92	91,8	1,5	1,5	1,1	20,5	18,9	18,4
1	90,8	91,5	91,4	1,9	2,1	1,7	20,9	19	18,1
1	91,3	92	92,1	1,8	1,8	1,2	21	19,5	18,3
1	91,3	92,3	92,4	1,4	1,2	1,2	20,2	18,5	18,3
1	91,2	91,7	91,9	1,9	2	1,7	20,6	19,4	18,1
1	90,9	90,7	91,6	1,9	1,9	1,6	20,7	19,3	18,4
1	91,5	92,6	92,3	1,4	1,4	1,1	20,1	17,8	17,7
1	91,2	92,1	92,3	1,9	1,9	1,5	20,9	18	17,6
1	91,2	92,1	91,8	1,9	1,9	1,5	20,9	18	17,7
1	91,1	92,7	92,7	1,8	1,7	1,3	20,8	18,1	17,8
1	91	92,7	92,7	1,9	1,8	1,4	21,3	18	17,8
1	91,5	92,7	93	1,9	1,8	1,4	21,1	17,9	18
1	91,4	92,5	92,5	1,6	1,7	1,2	20,5	18	17,9
1	90,8	91,6	92	1,9	2,2	1,8	20,9	17,7	17,9
1	91,4	92,5	92,5	1,6	1,7	1,5	20,7	17,7	17,8
1	91,6	92,7	92,3	1,6	1,8	1,5	20,5	17,9	17,8
1	91,8	92,4	92,1	1,1	1,2	1	20,1	18	18,3
1	90,4	91,5	91,2	1,8	1,8	1,4	20,1	17,5	17,6
1	90,8	92,2	91,8	1,7	1,7	1,5	20,9	17,5	17,9
2	87,2	85,4	86,1	0,1	0,5	0,3	35,3	34,5	34,8
2	87,1	84,7	85,4	0,3	0,5	0,4	35,9	34,3	35,1
2	87,2	84,4	86	0,1	0,5	0,2	35,3	34,5	35
2	87,1	85,1	85,8	0,3	0,6	0,4	35,9	35	35,7
2	87	85,4	85,6	0,2	0,6	0,3	35,4	34,5	34,9
2	87,3	85,8	86,1	0,2	0,4	0,2	35,5	34,1	34,8
2	87,4	86	86,1	0,1	0,4	0,2	35,6	34,3	34,8
2	87,1	85,4	85,8	0,3	0,7	0,4	36,1	35,2	35,7
2	87,1	85	85,4	0,2	0,4	0,2	35,8	34,8	35,1
2	86,8	84,5	85,2	0,3	0,8	0,4	35,3	35,2	35,6
2	87,1	85	85,2	0,3	0,7	0,4	36,1	35,2	35,3
2	87,4	84,9	85,2	0,1	0,8	0,4	35,7	35,3	35,1
2	87,3	85,2	85,8	0,2	0,7	0,3	35,9	35,6	35,2
2	87,6	85	85,7	0,2	0,8	0,4	36,1	36,2	35,9
2	86,4	85,9	86,2	0,3	0,7	0,5	35,6	35	35
2	86,2	85,7	86	0,3	0,5	0,3	35,8	35,1	35,4
2	87	85,9	86,7	0,2	0,4	0,2	35,3	34,9	34,9
2	86,6	85,4	86,3	0,2	0,6	0,2	35,5	35,1	34,8
2	86,6	85,3	85,9	0,3	0,5	0,3	35,8	34,7	34,6
2	87,2	85,2	85,8	0,2	0,8	0,4	36	35,4	35,5
3	89,6	87,6	87,8	7,7	8	8	51,4	50,2	50,3
3	89,8	87,7	87,8	7,8	7,9	8	51,8	50,4	50,6
3	89	87,5	87,1	7,6	7,8	8	51,1	49,9	49,7
3	89,7	87,8	87,8	7,7	8	8	51,6	50,7	50,6

3	89,7	87,8	87,4	7,7	8	7,9	51,6	50,6	50,2
3	89,9	87,8	87,8	7,6	7,7	7,9	51,4	49,9	50,1
3	90,3	87,9	88	7,6	8	8	51,6	50,6	50,6
3	90,2	88,1	87,9	7,7	7,8	8	51,8	50,3	50,5
3	89,8	87,7	87,5	7,6	7,9	7,9	51,5	50,1	50
3	90,1	88,1	87,8	7,6	7,9	7,9	51,5	50,5	50,2
3	90,2	87,9	87,8	7,7	7,8	8,2	51,7	50,1	50
3	89,7	87,6	88,2	7,7	7,9	8	51,8	50	50,4
3	89,6	88	88	7,7	7,9	7,9	51,6	50,5	49,8
3	89,6	88	88,1	7,7	7,9	8	51,7	50,5	50,2
3	89,4	88	87,8	7,7	7,9	8	51,6	50,4	50,5
3	89,9	88,2	88,1	7,7	8	8	51,6	50,8	50,6
3	89,7	87,7	87,5	7,7	7,9	7,9	51,6	50,3	50,2
3	89,7	88	87,7	7,7	8	8	51,5	50,8	50,4
3	90,1	88,1	88,2	7,7	7,9	8	51,9	50,4	50,6
3	89,7	87,5	87,6	7,6	7,8	8	51,4	50,1	50,1

KÜÇÜK ÖRNEKLEM BÜYÜKLÜKLERİNDE ÖRNEK SAYISININ TEKRARLANAN ÖLÇÜMLÜ VERİ ANALİZLERİ SONUCU ÜZERİNE ETKİSİ: BİR UYGULAMA

ORJİNALLİK RAPORU

% 13	% 11	% 4	% 2
BENZERLİK ENDEKSİ	İNTERNET KAYNAKLARI	YAYINLAR	ÖĞRENCİ ÖDEVLERİ

BİRİNCİL KAYNAKLAR

1	acikbilim.yok.gov.tr İnternet Kaynağı	% 4
2	openaccess.ogu.edu.tr:8080 İnternet Kaynağı	% 1
3	repozitorij.unizg.hr İnternet Kaynağı	% 1
4	libratez.cu.edu.tr İnternet Kaynağı	% 1
5	Sunay Fırat, Mehmet Aykut Erk. "Adli Bir Sorumluluk: Denetimli Serbestlik Personelinin Karar Verme Stratejileri ve Problem Çözme Becerilerinin Değerlendirilmesi", The Bulletin of Legal Medicine, 2020 Yayın	% 1
6	www.openaccess.hacettepe.edu.tr:8080 İnternet Kaynağı	<% 1
7	sbu.saglik.gov.tr İnternet Kaynağı	<% 1

