



**KARMA TASARIMLARDA SIRALI ALTERNATİFLER İÇİN
PARAMETRİK OLMAYAN TESTLER**

Bircan UTAŞ

**YÜKSEK LİSANS TEZİ
İSTATİSTİK ANA BİLİM DALI**

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

KASIM 2024

ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Bircan UTAŞ

06/11/2024

KARMA TASARIMLARDA SIRALI ALTERNATİFLER İÇİN PARAMETRİK OLMAYAN TESTLER

(Yüksek Lisans Tezi)

Bircan UTAŞ

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Kasım 2024

ÖZET

İstatistiksel deneye başlamadan önce deney tasarımı modeline karar verilmesi gerekir. Deney birimleri homojen olduğunda işlemleri karşılaştırmak için, rastgeleleştirme üzerinde herhangi bir kısıtlamanın olmadığı ve her deneğin yalnızca bir işleme tabi tutulduğu tamamen rasgele tasarım kullanılır. Rasgele tamamlanmış blok tasarımı, tamamen rasgele tasarımın bloklanmış halidir. Bloklama ile homojen gruplar oluşturularak modeldeki hata varyansının azaltılması amaçlanır. Rasgele tamamlanmış blok tasarımında denekleri işlemlere rastgele atamak yerine, her bir denek veya homojen grup tüm işlemlere tabi tutulur. Uygulamalarda bazı kısıtların bulunması ya da işlemlerin her blok için uygulanmasının mümkün olmadığı durumlarda rasgele tamamlanmış blok tasarımına alternatif olarak eksik blok tasarımı veya dengeli eksik blok tasarımı kullanılabilir. Araştırmacılar genellikle tek tip tasarım yapısıyla çalışmak ister. Ancak, deney tamamlanmadan önce herhangi bir nedenle tasarım yapısı değiştirilmek zorunda kalınabilir. Tasarım yapısının değiştirilmesi, mevcut verilerin kaybolmasına veya yeniden veri elde etmek araştırmacı için ek maliyete neden olabilir. Bu gibi durumlarda karma tasarımlar kullanılarak deney tamamlanabilir. Karma tasarımlar farklı deney tasarımı kombinasyonlarını içerebilir. Bu çalışmada deney tasarımlarında kullanılan parametrik olmayan test yöntemleri ile Kruskal Wallis ve Friedman testlerinin birleşiminden oluşan T1 test istatistiği için farklı simülasyon çalışmaları yapılmıştır. Simülasyon çalışmaları R programının 4.3.2 versiyonunda yapılmış olup kullanılan parametrik olmayan yöntemler deneysel I. tip hata ve testin gücü bakımından karşılaştırılmıştır. Simülasyonla verilerin üretildiği dağılımın parametre seçimleri ile uygulanan test istatistiklerinin işlem ve blok sayılarının testlerin güç ve deneysel I. tip hata oranlarını nasıl etkilediği incelenmiştir.

Bilim Kodu : 20507
Anahtar Kelimeler : Karma tasarım, parametrik olmayan testler, deney tasarımı
Sayfa Adedi : 66
Danışman : Prof. Dr. Hülya BAYRAK

NONPARAMETRIC TESTS FOR ORDERED ALTERNATIVES IN MIXED DESIGNS
(M. Sc. Thesis)

Bircan UTAŞ

GAZİ UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

November 2024

ABSTRACT

Before starting the statistical experiment, the experimental design model needs to be decided. When the experimental units are homogeneous, a completely randomized design is used to compare treatments, where there are no restrictions on randomization and each subject is subjected to only one treatment. The randomized complete block design is a blocked version of the completely randomized design. Blocking aims to reduce the error variance in the model by creating homogeneous groups. In randomized complete block design, instead of randomly assigning subjects to treatments, each subject or homogeneous group is subjected to all treatments. As an alternative to randomized complete block design, incomplete block design or balanced incomplete block design can be used when there are some constraints in the applications or when it is not possible to apply the treatments for every block. Researchers usually want to work with a uniform design structure. However, the design structure may have to be changed for any reason before the experiment is completed. Changing the design structure may result in the loss of existing data or additional costs for the researcher to obtain new data. In such cases, the experiment can be completed using mixed designs. Mixed designs may include different combinations of experimental designs. In this study, different simulation studies were conducted for the T1 test, which is a combination of nonparametric test methods used in experimental designs and Kruskal Wallis and Friedman tests. The simulation studies were performed in version 4.3.2 of the R program and the nonparametric methods used were compared in terms of experimental type I error and test power. With the simulation, it is examined how the parameter choices of the distribution from which the data are generated and the number of processes and blocks of the test statistics applied affect the power and experimental type I error rates of the tests.

Science Code : 20507
Key Words : Mixed design, nonparametric tests, experiment design
Page Number : 66
Supervisor : Prof. Dr. Hülya BAYRAK

TEŞEKKÜR

Çalışmalarında benden desteklerini ve katkılarını esirgemeyen değerli danışman hocam Prof. Dr. Hülya BAYRAK'a, yoğun iş temposuna rağmen derslere katılmama anlayış gösteren ve beni hep motive eden değerli iş aileme, hayatım boyunca hep yanımda olan, yoluma ışık tutan ve bana her zaman güç veren, hayattaki en büyük şansım olan canım aileme, en minik arkadaşım ve neşe kaynağım Zeynep Masal'ıma sonsuz teşekkür ederim.

İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER	vii
ÇİZELGELERİN LİSTESİ.....	ix
SİMGELER VE KISALTMALAR.....	x
1. GİRİŞ.....	1
2. TAMAMEN RASGELE TASARIM	7
2.1. Tamamen Rasgele Tasarım İçin Kullanılan Parametrik Olmayan Testler.....	8
2.1.1. Kruskal Wallis testi	10
2.1.2. Jonckheere ve Terpstra testi	10
3. RASGELE TAMAMLANMIŞ BLOK TASARIMI	13
3.1. Rasgele Tamamlanmış Blok Tasarımı İçin Parametrik Olmayan Testler.....	15
3.1.1. Friedman testi	15
3.1.2. Page testi.....	16
3.1.3. Hollander testi.....	17
3.1.4. Kim ve Kim testi.....	18
3.1.5. Uyarlanmış Page testi	20
3.1.6. Gökpınar ve Gökpınar testi.....	21
4. EKSİK BLOK TASARIMI.....	23
4.1. Eksik Blok Tasarımı İçin Parametrik Olmayan Testler	23
4.1.1. Alvo ve Cabilio testi	24
4.1.2. Mungai testi	25

Sayfa

5. DENGELİ EKSİK BLOK TASARIMI	27
5.1. Dengeli Eksik Blok Tasarımı İçin Parametrik Olmayan Testler	28
5.1.1. Durbin testi	28
5.1.2. Hemmer testi.....	29
6. KARMA TASARIMLAR.....	31
6.1. Karma Tasarımlar için Önerilen Testler.....	37
6.1.1. Dubnicka, Blair ve Hettmansperger testi.....	37
6.1.2. Magel, Terpstra ve Wen testi.....	38
6.1.3. Magel, Terpstra, Canonizado ve Park testi	39
6.1.4. Magel ve Mathisen testi.....	39
6.1.5. Magel ve Hemmer testi.....	40
7. SİMÜLASYON ÇALIŞMASI	43
8. SONUÇ VE ÖNERİLER	59
KAYNAKLAR	61
ÖZGEÇMİŞ	66

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 2.1. Tamamen rasgele tasarımın veri yapısı	8
Çizelge 3.1. Rasgele tamamlanmış blok tasarımının veri yapısı	14
Çizelge 4.1. Eksik blok tasarımının veri yapısı	23
Çizelge 5.1. Dengeli eksik blok tasarımının veri yapısı	28
Çizelge 6.1. Dört değerlendirici tarafından öğrencilere verilen notlar	31
Çizelge 6.2. Rasgele tamamlanmış blok tasarımın sıralamaları	32
Çizelge 6.3. Tamamen rasgele tasarımın sıralamaları	33
Çizelge 6.4. R programı kararı.....	35
Çizelge 7.1. ($n=5$, $p=0.1$, 0.2 , 0.5 parametrelili binom dağılımı) deneysel I. tip hata	44
Çizelge 7.2. ($n=5$, $p=0.8$, 0.9 parametrelili binom dağılımı) deneysel I. tip hata	46
Çizelge 7.3. $t=3$ için testlerin güç değeri	49
Çizelge 7.4. $t=4$ için testlerin güç değeri	50
Çizelge 7.5. $t=5$ için testlerin güç değeri	51
Çizelge 7.6. $t=3$, $n=6$ için testlerin gücü	52
Çizelge 7.7. $t=3$, $n=8$ için testlerin gücü	53
Çizelge 7.8. $t=3$, $n=10$ için testlerin gücü	53
Çizelge 7.9. $t=4$, $n=6$ için testlerin gücü	54
Çizelge 7.10. $t=4$, $n=8$ için testlerin gücü	55
Çizelge 7.11. $t=4$, $n=10$ için testlerin gücü	55
Çizelge 7.12. $t=5$, $n=6$ için testlerin gücü	56
Çizelge 7.13. $t=5$, $n=8$ için testlerin gücü	56
Çizelge 7.14. $t=5$, $n=10$ için testlerin gücü	57

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler

Açıklamalar

F	Friedman Testi
H	Hollander Testi
KW	Kruskal Wallis Testi
L	Page Testi
T1	T1 Testi (KW +F)
UP	Uyarlanmış Page Testi

Kısaltmalar

Açıklamalar

BIBD	Dengeli Eksik Blok tasarımı (Balanced Incomplete Block Design)
CRD	Tamamen Rasgele Tasarım (Completely Randomized Design)
IBD	Eksik Blok Tasarımı (Incomplete Block Design)
RCBD	Rasgele Tamamlanmış Blok Tasarımı (Randomized Complete Block Design)

1. GİRİŞ

İstatistiksel hipotezleri test etmek için parametrik ve parametrik olmayan istatistiksel test yöntemleri kullanılır. Parametrik test yöntemleri araştırmacılar tarafından daha iyi bilinir olsa da bu testler gerçekleştirilmeden önce verilerin; rasgele ve bağımsız olarak elde edilmesi, normal dağılması, ana kitle varyanslarının homojen olması gibi koşulları sağlaması gerekmektedir. Bununla birlikte her testin kendine has ek koşulları da olabilir (Karagöz, 2010). Bu varsayımlardan bir veya birkaçının sağlanmadığı durumlarda parametrik yöntemlerle elde edilen sonuçlar geçersizdir. Verilerin, parametrik test yöntemlerinin uygulanabilmesi için gerekli olan varsayımları sağlamadığı veya araştırmacı tarafından bu koşulları sağlayıp sağlamadığının doğrulanmasının zor olduğu durumlarda parametrik olmayan test yöntemleri, verileri analiz etmeye olanak sağlar.

Deney yapılacak ana kitleye ilişkin hangi dağılıma sahip olduğu, verilerin bağımsız ve rasgele elde edilip edilmediği gibi özelliklerin bilinmediği durumlarda, parametrik olmayan test yöntemleri güvenle uygulanabilir. İstatistiksel testleri uygulamadan önce verilerin sağlaması gereken varsayımlar parametrik testler için katı ve ayrıntılıyken, parametrik olmayan testlerde bu varsayımların daha esnek olması kolay karşılanmasını sağlar. Bu da araştırmacılar tarafından parametrik olmayan test yöntemlerinin daha fazla uygulanmasına yol açar. Parametrik olmayan testlerin basit hesaplamalar gerektirmesi özellikle istatistik bilimi dışında farklı disiplinlerden gelen araştırmacılar tarafından tercih edilmesini sağlar. Parametrik olmayan yöntemler, veriler zayıf bir ölçüm ölçeğinde ölçüldüğünde, örneğin sadece sınıflama (nominal ve sıralama (ordinal) düzeyi ile elde edilen verilerin mevcut olduğu durumlarda da uygulanabilir (Daniel, 1990). Ayrıca parametrik olmayan yöntemler farklı deney tasarımlarına ilişkin testler sunarak eksik gözlemlerle başa çıkabilir.

Deneklerle uğraşırken karşılaşılan doğal sorunlardan biri gözlenen etkinin gerçekten işlemiden mi yoksa denekler arasındaki farktan mı kaynaklandığıdır. Deneklerin doğasından kaynaklanan etki, gürültü faktörü olarak adlandırılır ve hesaba katılması gerekir. Deney tasarımları bloklama yöntemiyle bu duruma çözüm getirmektedir. Bloklama, homojen deneklerin bir blok oluşturacak şekilde gruplandırıldığı ve bu blok içindeki etkilerin karşılaştırıldığı bir tekniktir. Denek bir bloğu temsil edebilir (Ndungu, 2014).

İşlem etkileri çeşitli şekillerde test edilebilir. Teste başlamadan önce araştırmanın amacı ve hipotez olarak ifade edilen iddiasının belirtilmesi zorunludur. Bu, problemin net bir şekilde tanımlanmasına ve deneyin sonunda elde edilen bulguların ilişkilendirilmesine yardımcı olur.

Hipotez testi, ana kitle parametresine ilişkin doğru veya yanlış olabilecek bir varsayıma dayanır. Hipotez testleri, işlemlerle ilgili belirli bir iddianın doğru olup olmama olasılığını belirlemek için yapılır. İstatistikte, bir hipotezin tüm ana kitlede doğruluğunu belirlemek pratik olmadığından, araştırmacılar genellikle ana kitleden rasgele bir örneklem kullanırlar. Eğer örneklem istatistiksel hipotezle tutarlı değilse, hipotez reddedilir. Çoklu işlemler göz önüne alındığında en temel senaryo, farklı işlem etkilerinin hepsinin eşit olduğu sıfır hipotezine karşı en az iki işlem etkisi arasında önemli bir fark olduğuna dair alternatif bir hipotezi test etmektir. Literatürde alternatif hipotezlerden sıralı ve şemsiye alternatifler ön planda olan iki hipotez türüdür (Bağcı, 2019).

Sıralı alternatif hipotez setlerinde, işlem etkilerinin büyüklüğünün azalmayan bir şekilde sıralandığını belirten bir alternatif ile işlem etkilerinin hepsinin birbirine eşit olduğu sıfır hipotezi test edilir. Şemsiye alternatif hipotez setlerinde ise yine aynı sıfır hipotezi, işlem etkilerinin büyüklüğü tepe noktası olarak bilinen işleme kadar azalmayan ve tepe noktasından sonraki tüm işlemler için artmayan alternatif hipoteze karşı test edilir. İki hipotez testi de sıralı veriler üzerindeki varsayımlara dayanır. Ancak şemsiye alternatif hipotezlerde işlem etkilerinin belirli bir noktaya kadar arttığı o noktadan sonra azaldığı veya sabit kaldığı varsayılır.

Araştırmacının deneye başlamadan önce karar vermesi gereken diğer bir önemli husus da kullanacağı deney tasarımı modelidir. Tekrarlama, rasgeleleştirme ve bloklama deney tasarımının üç temel ilkesidir. Tekrarlama, işlemler için birden fazla deney yapılması iken rastgeleleştirme, gözlem birimlerine uygulanacak işlemlerin birimlere rastgele dağıtılmasıdır. Bloklama ise homojen gruplar oluşturularak modeldeki hata varyansının azaltılmasıdır.

Araştırmacılar genellikle tek bir tasarım yapısıyla çalışmak isterler. Ancak deney tamamlanmadan önce herhangi bir nedenle tasarım yapısı değiştirilmek zorunda kalınabilir. Örneğin, araştırmacı deneye rasgele tamamlanmış blok tasarımı (Randomized Complete

Block Design-RCBD) ile başlayabilir ancak daha sonra her bloktaki tüm işlemlere ilişkin verileri toplayamayabilir. Araştırmacı bu noktada tamamen rasgele tasarım (Completely Randomized Design -CRD) ya da dengeli eksik blok tasarımına (Balanced Incomplete Block Design -BIBD) geçebilir. Tasarımın değiştirilmesi sonucu, RCBD'de toplanan verilerin bir kısmı kaybedilir. Bununla birlikte, eldeki tüm verileri kullanmamak zaman ve kaynak israfına neden olur. Ne kadar çok gözlem toplanırsa, testin o kadar güçlü olacağı da göz önünde bulundurularak, RCBD, BIBD ve CRD tasarımları birleştirilebilir. Karma tasarım olarak adlandırılan bu tasarım türü için yeni bir test istatistiğinin oluşturulması gerekir (Asare, 2023).

Deney tasarımının beklenmedik bir şekilde değişebileceği olası durumlara, çeşitli gübrelerin farklı bölgeler üzerindeki etkisinin araştırıldığı ve tüm bölgelerde ürün verim oranını artırmaya yönelik yapılan bir çalışma örnek olarak verilebilir. Araştırmacı RCBD ile deneye başlayabilir ve bir süre sonra tasarıma devam etmenin çok pahalı olduğunu veya yeterli deney birimlerine sahip blokların tükenebileceğini fark edebilir. Her bir bölgede farklı gübre çeşidini uygulamak yerine gübreleri ürünlere rastgele atayabilir. Her bir bölgedeki ürünlerin bir kısmı RCBD bölümünü oluştururken, herhangi bir bölgeye bağlı olmayan ürünler çalışmanın CRD bölümünü oluşturur. Bu durumda, RCBD ve CRD kısmına sahip bir karma tasarım elde edilmiş olur (Alsuhabi, 2019).

Araştırmacılar mevcut verilerin tamamını kullanmak amacıyla karma tasarımlar üzerinde çalışmaktadır.

Magel ve diğerleri, (2010), sıfır hipotezini bilinen bir tepe noktası p olan şemsiye alternatifine karşı test etmek için RCBD altında Kim ve Kim (1992) ile CRD altında Mack-Wolfe (1981) test istatistiklerinin bir kombinasyonunu önermiştir. Çalışmalarında CRD'de işlemler için eşit örneklem büyüklükleri, RCBD'de ise i . işlem ve j . bloğu için birer gözlem dikkate alınmıştır.

Alvo ve Cabilio (1995) eksik gözlemlerin rastlantısallığını ele alan bir test önermiştir. Önerdikleri testte, eksik gözlemlerle başa çıkmak için bloklarda görülen sıralamaların bir ortalamasını atamışlardır. Bu çalışmadan esinlenen Magel ve Ndungu (2011), şemsiye alternatifi altında yeni bir test istatistiği geliştirmiştir. Bu test istatistiği, tam ve eksik blok

tasarımı altında uygulanabilir. Geliştirdikleri test istatistiği Mack-Wolfe (1981) test istatistiğine benzemektedir.

Karma tasarımlar, farklı deney tasarım yapılarıyla meydana gelebilir. Kim ve Kim (1992), bilinen bir tepe noktası olan RCBD'de sıfır hipotezini, şemsiye alternatif hipotezine karşı test etmek için bir test istatistiği geliştirmiştir. Testte Page ve Friedman'ın testinde olduğu gibi aynı sıralama şeması uygulanmıştır.

Dubnicka ve diğerleri, (2002), eşleştirilmiş ve bağımsız gözlemlerin bir kombinasyonu olan iki örneklemlili tasarımdan oluşan bir karma tasarım önermiştir. Test, eşleştirilmiş verilere Wilcoxon (1945) işaretli sıralar testini ve bağımsız örneklerle Mann-Whitney (1947) testini uygulamıştır. Her iki test istatistiğinin birleştirilmesi ve miktarın standardize edilmesi yeni testlerini oluşturmuştur.

Magel ve diğerleri, (2010), CRD ve RCBD'den oluşan karma tasarımda, şemsiye alternatifi için testler oluşturmuştur. Mack-Wolfe ile Kim ve Kim testlerini birleştiren iki test oluşturulmuştur. Testler, Magel ve diğerleri, (2009) tarafından yapılan çalışmaya benzer şekilde geliştirilmiştir. Birinci Testte, Mack-Wolfe ve Kim ve Kim'in standartlaştırılmış versiyonları toplanmış, ardından yeniden standartlaştırılmıştır. İkinci testte, test istatistiklerinin standartlaştırılmamış versiyonları toplanmış ve ardından toplam standartlaştırılmıştır.

Magel ve diğerleri, (2010), şemsiye alternatif altında karma tasarım için iki test önermiştir. Karma tasarım, tamamen rasgele tasarım ve rasgele tamamlanmış blok tasarımından oluşmaktadır. CRD için olağan Mack-Wolfe testinin standartlaştırılmış versiyonu ve RCBD için Kim ve Kim (1992) testinin standart versiyonunu kullanmışlardır.

Magel ve Mathisen (2011), RCBD ve BIBD'nin bir kombinasyonunu incelemiştir. Page (1963) ile Magel ve Ndungu'nun (2011) test istatistiklerini kullanarak iki yeni istatistik oluşturmuşlardır.

Magel ve Hemmer (2012), RCBD ve BIBD tasarımlarından oluşan karma tasarım için şemsiye alternatifli bir test önermiştir. Çalışmada yeni istatistikleri oluşturmak için Kim ve

Kim (1992) ile Magel ve Hemmer'in (2012) testleri dikkate alınmıştır. Bunları birleştirmek için ilk ve son standartlaştırmaya benzer yöntemler izlemişlerdir.

Gökpınar ve Gökpınar (2016), şemsiye alternatifler için test önermiştir. Bu uyarlanmış Mack-Wolfe test istatistiği, tepe noktası bilindiğinde, tepe noktasının sol tarafındaki ağırlıklı Mann-Whitney sayıları ile tepe noktasının sağ tarafındaki ters ağırlıklı Mann-Whitney sayılarının toplamıdır.

Tezin ikinci bölümünde deney tasarımlarından olan tamamen rasgele tasarım, üçüncü bölümünde rasgele tamamlanmış blok tasarımı, dördüncü bölümünde eksik blok tasarımı, beşinci bölümünde dengeli eksik blok tasarımı, altıncı bölümünde karma tasarımlar incelenmiş ve her bölümde incelenen tasarımlar için literatürde kullanılan test yöntemlerine yer verilmiştir. Yedinci bölümde farklı simülasyon çalışmaları yapılarak parametrik olmayan yöntemlerden tamamen rasgele tasarımda kullanılan Kruskal Wallis testi ile rasgele tamamlanmış blok tasarımında kullanılan Page, Uyarlanmış Page, Hollander ve Friedman testleri deneysel I. tip hata oranları ve güç bakımından karşılaştırılmıştır. Diğer bir simülasyon çalışmasında tamamen rasgele tasarım ve rasgele tamamlanmış blok tasarımından oluşan bir karma tasarım oluşturulmuştur. Bu simülasyon çalışmasında ilk çalışmada yer alan testlerden tamamen rasgele tasarım için uygun olan Kruskal Wallis testi ile rasgele tamamlanmış blok tasarımı için uygun olan Friedman testleri kullanılarak yeni bir T_1 test istatistiği oluşturulmuş, testler güç değerleri bakımından karşılaştırılmıştır. Karma tasarımda, hem farklı deney tasarımlarında kullanılmaları hem de iki test istatistiğinin toplanabilir olma özelliğinden dolayı Kruskal Wallis ve Friedman testleri tercih edilmiştir. Sekizinci bölümde sonuç ve önerilere yer verilmiştir.

2. TAMAMEN RASGELE TASARIM

Rasgeleleştirme üzerinde herhangi bir kısıtlama olmadığında kullanılan deneysel tasarım türü olan Tamamen Rasgele Tasarım (Completely Randomized Design-CRD), işlemleri karşılaştırmak için kullanılan ANOVA'nın özel bir halidir. Deney birimleri homojen olduğunda kullanılan bu tasarım türünde, rastgeleleştirme üzerinde herhangi bir kısıtlama yoktur. CRD'de denekler işlemlere rastgele atanır ve her denek yalnızca bir işleme tabi tutulur. Her işleme atanan denek sayısı, tüm işlemler için aynı olmak zorunda değildir. İşlem etkisi daha sonra işlemler tarafından oluşturulan gruplar arasındaki farktır.

Örneğin Ibuprofen ve Tylenol adlı iki ağrı kesici ilacın etkinliğinin karşılaştırılacağı bir araştırma için baş ağrısı olan 100 denek belirleniyor. Bu çalışmada değişkenler Ibuprofen ve Tylenol olacaktır. 100 kişi ikiye bölünüyor ve 50 kişiye Ibuprofen, diğer 50 kişiye ise Tylenol vermek için basit rasgele örnekleme yöntemi kullanılıyor. Denekler 0'dan 99'a kadar etiketlenerek ve rastgele sayı üretici (hesap makinesi ya da bilgisayar yazılımı) kullanılarak ilk 50 kişi seçiliyor ve tekrarlananlar göz ardı ediliyor. İlk 50 denek Ibuprofen almaya, geri kalanlar ise Tylenol almaya atanıyor. Bu iki grubun verileri karşılaştırılabilir. (Osafo, 2020)

Tamamen rasgele tasarımın matematiksel modeli

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad , \quad i=1, 2, \dots, a \quad j=1, 2, \dots, n \quad (2.1)$$

biçiminde tanımlıdır.

Burada, y_{ij} , i . işlemdeki j . gözlem değeri, μ , genel ortalama, τ_i i . işlemin etkisi, ϵ_{ij} , hata terimidir.

Bu tasarıma ilişkin varsayımlar

- Hata terimleri ortalaması sıfır ve varyansı σ^2 olan normal dağılıma sahiptir,
- Hata terimleri birbirinden bağımsızdır,
- Hata terimlerini varyansları homojendir.

dir. Bu varsayımlar istatistiksel olarak $\epsilon_{ij} \sim \text{NID}(0, \sigma^2)$ biçiminde ifade edilir.

CRD için veri yapısı Çizelge 2.1.'de verildiği gibidir.

Çizelge 2.1.Tamamen rasgele tasarımın veri yapısı

Gözlemler						
Denemeler	1	2	...	n	Toplam	Ortalama
1	y_{11}	y_{12}	...	y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	...	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$	$\vdots$
α	$y_{\alpha 1}$	$y_{\alpha 2}$	...	$y_{\alpha n}$	$y_{\alpha.}$	$\bar{y}_{\alpha.}$
					$y_{..}$	$\bar{y}_{..}$

Burada

- $y_{i.} = \sum_{j=1}^n y_{ij}$ $i=1,2,\dots,\alpha$ (i . işlem grubundaki gözlemlerin toplamı)
- $\bar{y}_{i.} = \frac{y_{i.}}{n}$ (i . işlem grubundaki gözlemlerin ortalaması)
- $N = \alpha \times n$ (toplam gözlem sayısı)
- $y_{..} = \sum_{i=1}^{\alpha} \sum_{j=1}^n y_{ij}$ (tüm gözlemlerin toplamı)
- $\bar{y}_{..} = \frac{y_{..}}{N}$ (tüm gözlemlerin ortalaması)

dır.

CRD basit, kolay anlaşılan ve kolay analiz edilen bir tasarım türüdür. Bu tasarım bazı nedenlerden dolayı yetersiz kalabilir ancak yine de deneysel bir tasarım geliştirilirken daha karmaşık bir tasarıma geçmeden önce CRD göz önünde bulundurulmalıdır (Oehlert, 2010).

2.1. Tamamen Rasgele Tasarım İçin Kullanılan Parametrik Olmayan Testler

Parametrik olmayan yöntemlerle işlem etkilerine ilişkin hipotezler çeşitli şekillerde test edilebilir. Problemin net bir biçimde tanımlanması ve istatistiksel deneyin sonunda bulguların ilişkilendirilmesi için araştırmanın amacının ve alternatif hipotez olarak da bilinen deneye ilişkin iddianın önceden belirtilmesi zorunludur. İstatistiksel deneyler

uygulanırken göz önünde bulundurulması gereken farklı hipotezler vardır. τ_i i. işlem etkisi olmak üzere,

$H_0: \tau_1 = \tau_2 = \dots = \tau_t$ sıfır hipotezine karşı,

H_1 : En az bir τ_i farklıdır. $i=1,2,\dots,t$

$H_1: \tau_1 \leq \tau_2 \leq \dots \leq \tau_t$ (En az bir eşitsizlik doğru olmak üzere)

$H_1: \tau_1 \geq \tau_2 \geq \dots \geq \tau_t$ (En az bir eşitsizlik doğru olmak üzere)

$H_1: \tau_1 \leq \tau_2 \leq \dots \leq \tau_p \geq \dots \geq \tau_t$ (En az bir eşitsizlik doğru olmak üzere)

alternatif hipotezleri kullanılır.

Örneğin doz-cevap çalışmalarında hayvanlar, deneysel bir ilacın k dozuna karşılık gelen k gruba atanır. İlacın hayvanlar üzerindeki etkisinin artan dozlarla artması veya azalması muhtemeldir. Bu senaryoda azalmayan veya artmayan sıralı bir alternatif hipotez modeli uygundur. Ancak bu tür bazı çalışmalarda ilacın hayvanlar üzerindeki etkisinin belirli bir seviyeye kadar dozajın artan bir fonksiyonu olabileceği ve daha sonra doz arttıkça etkisinin azaldığı görülebilir. Böyle bir senaryoda şemsiye alternatifi daha uygun bir modeldir.

Literatürde alternatif hipotezlerden sıralı ve şemsiye alternatifler ön planda olan iki hipotez türüdür (Bağcı, 2019). Şemsiye alternatiflerde sıfır hipotezi sıralı alternatiflerdeki sıfır hipoteziyle aynıdır. Fakat alternatif hipotez tepe noktasına kadar azalmayan ve tepe noktasından sonra artmayan biçimdedir.

Aslında sıralı alternatifler şemsiye alternatiflerinin özel bir halidir. Birinci veya sonuncu grubun tepe noktası olduğu durumda, şemsiye alternatiflerle sıralı alternatifler aynı olmaktadır. Bu nedenle şemsiye alternatifleri için geliştirilen testler genellikle sıralı alternatifler için önerilen testlerin uzantısı şeklindedir. Örneğin, şemsiye alternatifleri için kullanılan Mack-Wolfe (1981) testi, Jonckheere Terpstra testinin şemsiye alternatiflerine uyarlanmış halidir. Aynı şekilde Buning ve Kossler (1998), Kossler (2006), Gaur ve diğerleri, (2012) şemsiye alternatiflerine uyarlanmış sıralı alternatif testlerdir. Şemsiye alternatifleri için önerilen diğer testler ise Shi (1988), Cohen ve Sackrowitz (1996), Pan (1996), Hartlaub ve Wolfe (1999), Hayter ve Liu (1999), Lee ve Chen (2001), Magel ve Qin (2003), Singh ve Liu (2006), Alvo (2008), Bhat ve Patil (2008), Basso ve Salmaso (2011) tarafından verilmiştir (Bağcı, 2019).

2.1.1. Kruskal Wallis testi

Mann-Whitney (1947) testi (Daniel, 1990), iki işlem etkisinin konum parametreleri arasındaki farklılıkları test etmek için kullanılır. Sıralamaya dayalı bir testtir. Bu nedenle verilerin en azından sıralı olmasını gerektirir. İki işlemin, işlem etkilerine ilişkin popülasyonunun bağımsız olduğu ve yalnızca konum parametrelerinde farklılık gösterdikleri varsayılmaktadır.

İkiden fazla işlem etkisinin test edilmesi durumunda, Mann Whitney'in (1947) bir uzantısı olan Kruskal Wallis (1953) testi, farklılıkları test etmeyi sağlar. Popülasyonları yalnızca konum parametrelerinde farklılık gösteren Mann-Whitney (1947) ile aynı varsayımlara sahip olsa da Kruskal Wallis (1953) testi, sıfır hipotezini sadece iki yönlü alternatif hipotezlere karşı test etmek için kullanılabilir. Test edilen sıfır hipotezi ile alternatif hipotez,

$$H_0: \tau_1 = \tau_2 = \dots = \tau_t$$

$$H_1 : \text{En az bir } \tau_i \text{ farklıdır.} \quad i=1,2, \dots, t$$

biçimindedir. Kruskal Wallis (1953) test istatistiği (KW)

$$KW = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{1}{n_i} \left[Ri - \frac{n_i(N+1)}{2} \right]^2 \quad (2.2)$$

biçiminde tanımlıdır. Burada, R_i , gözlemlere atanan sıraların toplamı, n_i , i . denemedeki örnek boyutu, N , yığındaki toplam gözlem sayısıdır. KW testi $(k-1)$ serbestlik dereceli ki-kare dağılıma sahiptir. α anlamlılık düzeyinde $KW \geq k_\alpha$ olduğu durumlarda H_0 reddedilir. k_α örnek boyutu ve işlem sayısına karşılık gelen tablo değeri olup birinci tip hata olasılığını α 'ya eşit yapmak için seçilir.

2.1.2. Jonckheere ve Terpstra testi

Jonckheere (1954) ve Terpstra (1952), konum parametreleri için azalmayan sıralı alternatif için dağılımdan bağımsız test geliştiren ilk araştırmacılardandır. Jonckheere–Terpstra ((Jonckheere (1954), Terpstra (1952))), kısaca JT testi, ikiden fazla işlem etkisindeki

farklılıkları test eder; özellikle de farklılıkları yönlü olarak test etmek için kullanılır. Test edilen sıfır hipotezi ile alternatif hipotez

$$H_0: \tau_1 = \tau_2 = \dots = \tau_t$$

$$H_1: \tau_1 \leq \tau_2 \leq \dots \leq \tau_t \quad (\text{En az bir eşitsizlik doğru olmak üzere})$$

biçimindedir.

JT testi, bilinen p tepe değerine sahip şemsiye alternatifinin özel bir durumudur. Mann-Whitney (1947) testinden farklı olarak sıralamaya dayalı değildir. *JT* test istatistiği, ilgili değerleri kullanarak işlem çiftlerinin karşılaştırılmasına dayanır. Daniel (1990) tarafından verilen ve sonuçların geçerliliğini sağlamak için karşılanması gereken varsayımlar vardır. Bu varsayımlar

- Gözlemler birbirinden bağımsızdır,
- İlgilenilen değişken süreklidir,
- Veriler en az sıralı olmalıdır,
- İşlem etkileri popülasyonu, sadece konum parametreleri açısından farklılık gösterir.

dır. *JT* testini kullanmak için, işlemler öncelikle algılanan yön farklılıkları sırasına göre düzenlenmelidir.

Farklı işlemlerden elde edilen gözlem çiftlerinin tüm olası kombinasyonları oluşturulur. A işlemindeki her gözlem, B işlemindeki her gözlemle bir çift oluşturur. Sonunda her bir çift değerlendirilir ve bu çift içindeki büyüklük sırası işlemin sırasına uyuyorsa 1, aksi takdirde 0 atanır. Bu, a 'nın A'dan ve b 'nin B'den bir gözlem olduğu bir (a,b) çifti için, a 'nın b 'den küçük olması ve alternatifin azalmayan olması durumunda çiftin genel istatistiğe 1 değeri katacağı anlamına gelir. Test istatistiği daha sonra çiftlere atanan tüm değerler toplanarak hesaplanır. Test istatistiği

$$JT = \sum_{i < j} U_{ij} \quad (U_{ij}, \text{büyüklük sırasının alternatifle uyumlu olduğu çiftlerin sayısıdır})$$

biçiminde tanımlıdır. Ortalama ve varyansı sırasıyla

$$E(JT) = (N^2 - \sum_{i=1}^k n_i^2)/4$$

ve

$$V(JT) = (N^2 (2N+3) - \sum_{i=1}^k n_i^2 (2n_i+3))/72$$

biçiminde hesaplanır. Standartlaştırılmış JT istatistiği

$$Z_{JT} = \frac{JT - E(JT)}{\sqrt{V(JT)}} = \frac{JT - (N^2 - \sum_{i=1}^k n_i^2)/4}{(N^2 (2N+3) - \sum_{i=1}^k n_i^2 (2n_i+3))/72} \quad (2.3)$$

biçiminde tanımlıdır. Burada, n_i , i . denemedeki örnek boyutu, N , yığındaki toplam gözlem sayısıdır.

Z_{JT} asimptotik standart normal dağılıma sahiptir. α anlamlılık düzeyinde, $Z_{JT} > Z_\alpha$ ise H_0 reddedilir. Burada Z_α standart normal dağılımın $(1-\alpha)100$ yüzdelik dilimidir.

3. RASGELE TAMAMLANMIŞ BLOK TASARIMI

Herhangi bir deneyde, gürültü faktörünün neden olacağı değişkenlik sonuçları etkileyebilir. Genellikle gürültü faktörü, yanıt değişkenini etkileyen bir tasarım faktörü olarak tanımlanır fakat etkisiyle ilgilenilmez. Değişkenliğin gürültü kaynağı bilindiğinde ve kontrol edildiğinde, bir tasarım tekniği olan bloklama kullanılarak sistematik bir biçimde gürültü değişkeninin işlemler arasındaki etkisi elimine edilebilir.

Örneğin, migren rahatsızlığı olan insanlardan oluşan bir örnekleme deneklerin yarısına migreni hafifleten bir ilaç, diğer yarısına ise plasebo verilsin. İlacın yaşlı insanlar üzerinde daha zayıf bir etkiye sahip olduğundan şüpheleniliyorsa ve ilacı alan deneklerin ortalama yaşı, plasebo alan grubun yaş ortalamasından önemli ölçüde yüksekse, o zaman yaş, denemenin hedeflerini etkileyecek ve nihayetinde azaltacak bir faktördür. Bu etkileşimi önlemek için araştırmacı, gürültü faktörünün çeşitli düzeylerinden birden fazla örnek alabilir ve t , deneydeki toplam işlemlerin sayısı olmak üzere her örneğe t sayıda işlem uygulayabilir. Gürültü faktörünün farklı seviyelerindeki bu örneklerle blok denir. Blok, değişkene göre mümkün olduğunca birbirine benzeyen deneysel birimlerden oluşur. Bloklama ile homojen gruplar oluşturularak modeldeki hata varyansının azaltılması amaçlanır.

Rasgele tamamlanmış blok tasarımı (Randomized Complete Block Design-RCBD), tamamen rasgele tasarımın bloklanmış halidir. Her işlem seviyesinin her blok seviyesine uygulandığı bir deneysel tasarım türüdür. Blok, değişkene göre mümkün olduğunca birbirine benzeyen deneysel birimlerden oluşan bir gruptur. “Tamamlanmış” kelimesi, her bloğun tüm işlemleri içerdiğini gösterir. Bloklar içindeki deneysel birimler nispeten birbirine benziyorsa ve farklı bloklardaki birimler nispeten farklıysa, o zaman rasgele tamamlanmış blok tasarımı, tüm deneysel birimlerin göreceli olarak homojen olduğunu varsayan bir tasarım olan tek yönlü ANOVA tasarımına göre işlem araçlarındaki farklılıklara genellikle daha duyarlıdır. (Duran, 1995).

Denekleri işlemlere rastgele atamak yerine, RCBD’de her bir denek veya homojen grup, tüm işlemlere tabi tutulur. İşlemlerin uygulanma sırası rastgeledir. İşlem etkileri daha sonra her bir blokta ölçülür. RCBD’nin en önemli özelliği kayıp gözlem olmamasıdır. İşlem etkilerinin test edilmesi için ideal bir tasarımı temsil eden RCBD’de, eksik gözlem olmayacağı düşünülürse oldukça pahalı bir tasarım olduğu söylenebilir (Ndungu, 2014).

Rasgele tamamlanmış blok tasarımının matematiksel modeli

$$Y_{ijk} = \mu + \tau_i + \beta_j + \varepsilon_{ijk} \quad , \quad i=1, \dots, t \quad , \quad j=1, \dots, b \quad , \quad k=1, 2, \dots, n \quad (3.1)$$

biçiminde tanımlıdır. Burada, Y_{ijk} , j . bloktaki i . işleme ait k . gözlem değeri, μ , genel ortalama, τ_i , i . işlemin etkisi, β_j , j . blok etkisi, ε_{ijk} hata terimidir.

RCBD için veri yapısı Çizelge 3.1.'de verildiği gibidir.

Çizelge 3.1.Rasgele tamamlanmış blok tasarımının veri yapısı

Bloklar						
Denemeler	1	2	...	b	Toplam	Ortalama
1	y_{11}	y_{12}	...	y_{1b}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	...	y_{2b}	$y_{2.}$	$\bar{y}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	...	y_{ab}	$y_{a.}$	$\bar{y}_{a.}$
Toplam	$y_{.1}$	$y_{.2}$	...	$y_{.b}$	$y_{..}$	
Ortalama	$\bar{y}_{.1}$	$\bar{y}_{.2}$		$\bar{y}_{.b}$		$\bar{y}_{..}$

Burada

- $y_{i.} = \sum_{j=1}^b y_{ij}$ (i . işlem grubundaki gözlemlerin toplamı)
- $\bar{y}_{i.} = \frac{y_{i.}}{b}$ (i . işlem grubundaki gözlemlerin ortalaması)
- $y_{.j} = \sum_{i=1}^a y_{ij}$ (j . bloktaki gözlemlerin toplamı)
- $\bar{y}_{.j} = \frac{y_{.j}}{a}$ (j . bloktaki gözlemlerin ortalaması)
- $N = a \times b$ (toplam gözlem sayısı)
- $y_{..} = \sum_{i=1}^a \sum_{j=1}^b y_{ij}$ (tüm gözlemlerin toplamı)
- $\bar{y}_{..} = \frac{y_{..}}{N}$ (tüm gözlemlerin ortalaması)

dır. Hata terimleri (ε_{ijk}) bağımsız ve aynı dağılımlıdır.

$$\sum_{i=1}^t \tau_i = 0 \quad , \quad \sum_{j=1}^b \beta_j = 0$$

Rasgele tamamlanmış blok tasarımı için şu örnek verilebilir.

Bir tıbbi cihaz üreticisi vasküler greftler (yapay damarlar) üretmektedir. Bu greftler, politetrafloroetilen reçine kütüklerinin bir yağlayıcı ile birlikte tüplere çekilmesiyle üretilmektedir. Üretim çalışmasındaki tüplerin bazılarının dış yüzeyinde sıklıkla küçük, sert çıkıntılar bulunur. Bu kusurlar "fiske" olarak bilinir ve kusur, ünitenin reddedilmesi için bir nedendir.

Vasküler greftlerden sorumlu ürün geliştiricisi, çekme basıncının fiske oluşumunu etkilediğinden şüphelenir ve bu hipotezini araştırmak için bir deney yapmayı planlar. Ancak reçine dış bir tedarikçi tarafından üretilmekte ve tıbbi cihaz üreticisine gruplar halinde teslim edilmektedir. Mühendis ayrıca malzemenin moleküler ağırlık, ortalama parçacık boyutu, alıkonma ve pik yüksekliği oranı gibi parametreler açısından tutarlı olması gerekirken, reçine tedarikçisindeki üretim değişimi ve malzemedeki doğal değişim nedeniyle gruptan gruba farklılık gösterdiğinden şüphelenmektedir. Bu durumda ürün geliştiricisi, reçine gruplarını blok olarak ele alan rasgele tamamlanmış blok tasarımı kullanarak dört farklı çekme basıncı seviyesinin fiskeler üzerindeki etkisini araştırabilir (Montgomery, 1997).

3.1. Rasgele Tamamlanmış Blok Tasarımı İçin Parametrik Olmayan Testler

3.1.1. Friedman testi

Milton Friedman, her deneme için yanıt değişkeniyle ilişkili sıraların büyüklüklerini (sütun sıralarının toplamını) kullanan Friedman (1937, 1940) testini geliştirmiştir. İşlem, her blokta yanıt değişkeninin sıralanmasını içerir, ardından sıralar sütunlara göre değerlendirilir. Bu teknik, bloklar arasındaki korelasyonun hesaba katılmasını sağlar. Friedman'ın testini kullanmak için veriler bazı varsayımları sağlamalıdır. Karşılıklı olarak bağımsız bloklara sahip olması gerekir, gözlemler örnekler içinde ve arasında bağımsız, yanıt değişkenleri sürekli, gözlemler her blok içinde sıralanabilir olmalıdır. Bir diğer önemli varsayım ise bloklar arasında sıralama yapılırken gürültü faktöründen kaçınılamayacağı için bloklar ve işlemler arasında hiçbir etkileşim olmamasıdır. Friedman (1937, 1940) testi, yönsel değildir ve yalnızca farklılıkları tespit edebilmektedir. Test edilen sıfır hipotezi ile alternatif hipotez,

$$H_0: \tau_1 = \tau_2 = \dots = \tau_t$$

$$H_1: \text{En az bir } \tau_i \text{ farklıdır.} \quad i=1,2,\dots,t$$

biçimindedir. F test istatistiği

$$F = \frac{12}{bk(k+1)} \sum_{j=1}^k \left[R_j - \frac{b(k+1)}{2} \right]^2 \quad (3.2)$$

biçiminde tanımlıdır. Burada, b , blok sayısı, j , işlem sayısı ve R_j , j . işleme atanan sıra sayıları toplamıdır.

H_0 altında Friedman testi ki-kare dağılır ve serbestlik derecesi $k-1$ 'dir. $F > \chi_{k-1,\alpha}^2$ olduğu durumlarda H_0 reddedilir.

3.1.2. Page testi

Yön farklılıklarını test etmek için Friedman (1937, 1940) testine alternatif olarak Page (1963) testi sunulmuştur. Page testinde Friedman (1937, 1940) testinde olduğu gibi, gözlemler bir blok içinde sıralanır ve her birinin R_j sıraları toplanarak L test istatistiği oluşturulur. Sıfır hipotezi ile alternatif hipotez

$$H_0: \tau_1 = \tau_2 = \dots = \tau_t$$

$$H_1: \tau_1 \leq \tau_2 \leq \dots \leq \tau_t \quad (\text{En az bir eşitsizlik doğru olmak üzere})$$

biçimindedir. L test istatistiği

$$L = \sum_{j=1}^t jR_j = R_1 + 2R_2 + \dots + tR_t \quad (3.3)$$

biçiminde hesaplanır.

Page (1963) istatistiğinde daha yüksek beklenen etkileri olan işlemlere daha fazla ağırlık verilir. Büyük L değerleri için sıfır hipotezi reddedilir. Beklenen değer ve varyansı sırasıyla

$$E(L) = \frac{bt(t+1)^2}{4}$$

ve

$$V(L) = \frac{b(t^3-t)^2}{144(t-1)}$$

biçiminde hesaplanır. Burada; t , işlem sayısı, b , blok sayısıdır. L test istatistiğinin standartlaştırılmış hali

$$Z_L = \frac{[L - \frac{bt(t+1)^2}{4}]}{\sqrt{b(t^3-t)^2/144(t-1)}} \quad (3.4)$$

biçiminde tanımlıdır.

Z_L test istatistiği asimptotik standart normal dağılır. $Z_L > Z_\alpha$ olduğu durumlarda H_0 reddedilir. Burada Z_α standart normal dağılımın $(1-\alpha)100$ yüzdalık dilimidir.

3.1.3. Hollander testi

Hollander (1967) testi, rasgele tamamlanmış blok tasarımında Wilcoxon (1945) işaretli sıra test istatistiğini kullanmıştır. Her bir $(u; v)$ işlem çifti için ve $(1 < u < v < t)$ her biri için, T_{uv} işaretli sıra istatistiği aşağıdaki şekilde tanımlanmıştır:

$$T_{uv} = \sum_{i=1}^b R_{uv}^i \psi_{uv}^i \quad (3.5)$$

$$\psi_{uv}^i = \begin{cases} 1 & X_{iu} < X_{iv} \\ \frac{1}{2} & X_{iu} = X_{iv} \\ 0 & X_{iu} > X_{iv} \end{cases}$$

Burada T_{uv} , u . ve v . işlemler arasındaki farklılığın bir ölçüsüdür. Hollander testi, işlemlerin önceden sıralanmasını dikkate alır. Test edilen sıfır hipotezi ile alternatif hipotez

$$H_0: \tau_1 = \tau_2 = \dots = \tau_t$$

$$H_1: \tau_1 \leq \tau_2 \leq \dots \leq \tau_t \quad (\text{En az bir eşitsizlik doğru olmak üzere})$$

biçimindedir. T_{uv} istatistiğine bağlı olan Hollander test istatistiği (H)

$$H = \sum_{u=1}^{t-1} \sum_{v=u+1}^t T_{uv} \quad (3.6)$$

biçiminde tanımlıdır. H test istatistiğinin ortalama ve varyansı sırasıyla

$$E(H) = \frac{bt(t-1)(b+1)}{8}$$

ve

$$V(H) = \frac{bt(b+1)(2b+1)(t-1)\{3+2(t-2)p_U^b\}}{144}$$

biçiminde hesaplanır.

Burada P_U^b ; Wilcoxon işaretli sıralı test istatistikleri arasındaki korelasyonun değeridir ve b' ye bağlıdır. Standartlaştırılmış hali

$$Z_H = \frac{H - E(H)}{\sqrt{V(H)}}$$

biçiminde tanımlıdır.

Z_H testi asimptotik standart normal dağılır. $Z_H > Z_\alpha$ olduğunda sıfır hipotezi reddedilir.

Burada Z_α standart normal dağılımın $(1-\alpha)100$ yüzdelik dilimidir.

3.1.4. Kim ve Kim testi

Tamamen rasgele deney tasarımı için kullanılan testlerden biri Mack–Wolfe (1981) testidir. Bu test, şemsiye alternatifiyle ilgilenen araştırmacılar için uygundur. Mack–Wolfe (1981) testi, test istatistiği için Mann–Whitney (1947) sayılarını kullanır. Şemsiye tasarımına artmayan ve azalmayan iki ayrı tasarım olarak yaklaşır.

Tepe noktası bilindiğinde, Mack-Wolfe test istatistiği A_p , U_{uv} tepe noktasının sol tarafındaki Mann-Whitney sayıları ile U_{uv} tepe noktasının sağ tarafındaki ters Mann-Whitney sayılarının toplamıdır. Test istatistiği,

$$A_p = \sum_{u=1}^{v-1} \sum_{v=2}^p U_{uv} + \sum_{u=1}^{v-1} \sum_{v=p+1}^k U_{uv} \quad (3.7)$$

biçiminde tanımlıdır. Büyük A_p değerleri için sıfır hipotezi reddedilir. Mack-Wolfe (1981) test istatistiği, örneklem büyüklüğü sonsuza giderken asimptotik olarak normal dağılır.

Kim ve Kim (1992), bilinen bir tepe noktasının olduğu rasgele tamamlanmış blok tasarımında, şemsiye alternatifi için tasarlanmıştır. Sıfır hipotezi ile alternatif hipotez

$$H_0: \tau_1 = \tau_2 = \dots = \tau_t$$

$$H_1: \tau_1 \leq \tau_2 \leq \dots \leq \tau_p \geq \dots \geq \tau_t \quad (\text{En az bir eşitsizlik doğru olmak üzere})$$

biçimindedir.

Test aynı zamanda sıralamaya dayalıdır ve Friedman (1937, 1940) ile Page (1963) testleri ile aynı tekniği uygulamaktadır. Dolayısıyla gözlemlerle ilgili diğer iki testle aynı varsayımlara sahiptir. Kim ve Kim (1992) test istatistiği, bloklar arasında etkileşim olmadığını varsayar. Başlangıçta her blokta Mack-Wolfe test istatistiğinin hesaplanmasını gerektirir. Test istatistiği, tüm bloklardaki Mack-Wolfe istatistiklerinin toplamıdır.

$$KK = \sum_{i=1}^b A_{ip} \quad (3.8)$$

$$A_{ip} = \sum_{i=1}^b \{ \sum_{u=1}^{v-1} \sum_{v=2}^p U_{iuv} + \sum_{u=p}^{v-1} \sum_{v=p+1}^k U_{ivu} \} \quad (3.9)$$

Burada; A_{ip} , i . bloğun Mack ve Wolfe (1981) test istatistiği, b , blok sayısı, p , bilinen tepe noktası, k , işlem sayısıdır.

U_{iuv} ve U_{ivu} , i bloğu ile ilişkili U istatistikleridir. Sıfır hipotezi büyük KK değerleri için reddedilir.

U_{iuv} , u işlemini alan gözlemin v işlemini alan gözlemde daha az olduğu zamanların uygun işlem aralıkları üzerinden toplamıdır. U_{ivu} , b bloğu içindeki uygun aralıkta v işlemini alan gözlemin u işlemini alan gözlemde kaç kez daha az olduğunu sayar. KK test istatistiğinin ortalama ve varyansı sırasıyla

$$E(KK) = \frac{\sum_{i=1}^b \{N_{i1}^2 + N_{i2}^2 - \sum_{j=1}^k n_{ij}^2 - n_p^2\}}{4}$$

ve

$$V(KK) = \sum_{i=1}^b \left\{ 2(N_{i1}^3 + N_{i2}^3) + 3(N_{i1}^2 + N_{i2}^2) - \sum_{j=1}^k n_{ij}^2 (2n_{ij} + 3) - n_{ip}^2 (2n_{ip} + 3) \right\} / 72$$

$$+ 12n_{ip} N_{i1} N_{i2} - 12n_{ip}^2 N_i$$

biçiminde hesaplanır. Burada

$$N_{i1} = \sum_{j=1}^p n_{ij} , \quad N_{i2} = \sum_{j=p}^k n_{ij} , \quad N_i = \sum_{j=1}^k n_{ij}$$

dır. Kim ve Kim testinin standartlaştırılmış hali

$$Z_{KK} = \frac{KK - E(KK)}{\sqrt{V(KK)}}$$

biçiminde tanımlıdır.

Kim ve Kim (1992) test istatistiği, H_0 doğru olduğunda, asimptotik standart normal dağılır. $Z_{KK} \geq Z_\alpha$ olduğunda sıfır hipotezi reddedilir. Burada Z_α standart normal dağılımın $(1-\alpha)100$ yüzdilik dilimidir.

3.1.5. Uyarlanmış Page testi

Best ve Reyner (2009) tarafından sıralı alternatif için Uyarlanmış Page (UP) testi geliştirilmiştir. UP Test istatistikleri ortogonal trend kontrastı (L) kullanılarak elde edilir.

$$UP = \frac{CL^2}{\sum_{j=1}^t \lambda_j^2} \quad (3.10)$$

$$L = \sum_{j=1}^t \lambda_j \bar{R}$$

biçiminde tanımlıdır.

Burada, $\bar{R} = \frac{R_j}{b}$, j , doğrusal katsayı, b , blok sayısı, λ_j , lineer katsayı, R_j , işlemlere atanan sıra değerlerinin toplamıdır.

$$C = \begin{cases} \frac{b(t-1)}{tV} & , \text{ aynı değere sahip gözlem var} \\ \frac{12b}{t(t+1)} & , \text{ aynı değere sahip gözlem yok} \end{cases}$$

$$V = \begin{cases} \left\{ \sum_{s=1}^q r_s^2 c_s \right\} - \frac{(t+1)^2}{4} & , \text{ aynı değere sahip gözlem var} \\ \frac{(t^2-1)}{12} & , \text{ aynı değere sahip gözlem yok} \end{cases}$$

biçiminde tanımlıdır.

Burada, r_s , s inci sıralama-veride gözlenen sıra sayı değerleri, c_s , gözlenen sıra sayı değerlerinin kaç tane olduğudur.

$UP > \chi_{1,\alpha}^2$ olduğunda sıfır hipotezi reddedilir.

3.1.6. Gökpınar ve Gökpınar testi

Gökpınar ve Gökpınar (2016) şemsiye alternatiflerini test etmek için testler önermiştir. Tepe noktası bilindiğinde, uyarlanmış Mack-Wolfe test istatistiği mAp tepe noktasının sol tarafındaki ağırlıklı Mann-Whitney sayıları $(v-u)U_{uv}$ ile tepe noktasının sağ tarafındaki ters ağırlıklı Mann-Whitney sayılarının $(u-v)U_{uv}$ toplamıdır. Genel olarak, Mann-Whitney sayıları U_{uv} ve U_{vu} , ilgili formlarında $(v-u)$ ve $(u-v)$ ağırlıklarını alır. Önerilen istatistik, Neuhauser ve diğerleri (1988) tarafından sıralı alternatifler için incelenen iki değiştirilmiş Jonckheere istatistiğinin toplamıdır. Neuhauser ve diğerlerinin (1988) test istatistiği Tryon

ve Hettmansperger (1973) tarafından araştırılan ağırlıklı parametrik olmayan istatistikler sınıfının bir üyesidir. Böylece değiştirilmiş Mack-Wolfe test istatistiği mAp şu şekilde verilir;

$$mAp = \sum_{u=1}^{v-1} \sum_{v=2}^p (v-u)U_{uv} + \sum_{u=p}^{v-1} \sum_{p=i+1}^k (u-v)U_{vu} \quad (3.11)$$

Örneklem büyüklükleri eşit olduğunda ($n_1 = n_2 = \dots = n_k = n$) ve H_0 altında tüm popülasyon ortalamalarının eşit olduğu varsayıldığında, mAp 'nin beklenen değeri ve varyansı sırasıyla

$$E(mAp) = \frac{n^2}{2} \left[\binom{p+1}{3} + \binom{k-p+2}{3} \right]$$

ve

$$V(mAp) = \frac{n^2 p^2 (p^2 - 1)(np + 1) + n^2 (k - p + 1)^2 [(k - p + 1)^2 - 1] [n(k - p + 1) + 1]}{144} \\ + \frac{n^3 p (p - 1)(k - p)(k - p + 1)}{24}$$

biçiminde hesaplanır. mAp testinin standartlaştırılmış hali

$$mAp^* = \frac{mAp - E(mAp)}{\sqrt{V(mAp)}}$$

biçiminde tanımlıdır.

Büyük örneklem hacimleri için mAp^* asimptotik standart normal dağılıma sahiptir. $mAp^* \geq Z_{\alpha}$ olduğunda sıfır hipotezi reddedilir. Burada Z_{α} standart normal dağılımın $(1-\alpha)100$ yüzdelik dilimidir.

4. EKSİK BLOK TASARIMI

Uygulamalarda bazı kısıtların bulunması ya da işlemlerin her blok için uygulanmasının mümkün olmadığı durumlarda rasgele tamamlanmış blok tasarımına alternatif olarak eksik blok tasarımı (Incomplete Block Design-IBD) kullanılabilir.

Rasgele tamamlanmış blok tasarımı kullanılan deneylerde, her blokta tüm işlem kombinasyonlarıyla çalışılamayabilir. Böyle durumlarda eksik blok tasarımı kullanılabilir. IBD, RCBD ile aynı kriterleri izler ancak eksik gözlemler olabilir. Eksik gözlemler işleme alınan denek sayısını sınırlayan bir çalışmanın ortasında fonların tükenmesi, denekler üzerinde takip gerektiren çalışmalara yeterince özen gösterilmemesi veya özellikle ölümcüllüğü test eden çalışmalar bitmeden deneklerin ölümü gibi durumlardan kaynaklanabilir.

IBD veri yapısına uygun bir örnek Çizelge 4.1.'de verilmiştir.

Çizelge 4.1. Eksik blok tasarımının veri yapısı

Bloklar						
	B_1	B_2	B_3	B_4	B_5	B_6
A_1		Y_{12}	Y_{13}		Y_{15}	
A_2	Y_{21}	Y_{22}	Y_{23}		Y_{25}	
A_3	Y_{31}		Y_{33}	Y_{34}	Y_{35}	Y_{36}
A_4		Y_{42}		Y_{44}	Y_{45}	Y_{46}

Tabloda görüldüğü gibi IBD'de bazı bloklarda boş hücreler vardır. Bloklardaki birim sayısının, işlem düzeyinin sayısından az olması durumunda bloklar tamamlanmamış (eksik) olur.

4.1. Eksik Blok Tasarımı İçin Parametrik Olmayan Testler

Eksik gözlemlerin rastgele seçildiği, bu nedenle dengeli bir modelin oluşmadığı tasarım türüdür. Başka bir ifadeyle, IBD tüm işlemleri her bloğa dağıtmamaktadır.

4.1.1. Alvo ve Cabilio testi

Alvo ve Cabilio (1995) eksik gözlemlerin rastlantısallığını ele alan bir test önermiştir. Bu test, Page (1963)'in azalmayan alternatif için kullandığı istatistiğin bir uzantısıdır. Bu nedenle, Page'in test istatistiğiyle aynı varsayımlara sahiptir ve benzer bir sıralama prosedürü izler. Test edilen fıfır hipotezi ile alternatif hipotez

$$H_0: \tau_1 = \tau_2 = \dots = \tau_t$$

$$H_1: \tau_1 \leq \tau_2 \leq \dots \leq \tau_t \quad (\text{En az bir eşitsizlik doğru olmak üzere})$$

biçimindedir.

Alvo ve Cabilio (1995) önerdikleri testte, eksik gözlemlerle başa çıkmak için blokta görülen sıraların ortalamasını atamışlardır. Ayrıca, eksik gözlemlere sahip olan bloğa,

$$\sum_{j=1}^k \frac{(k+1)}{(k_i+1)} \mu_{ij} \tag{4.1}$$

formülünü kullanarak ağırlık atamışlardır. Burada, μ_{ij} , gözlemin sırası veya eksik olan gözlemlerin sıralarının ortalamasıdır. Test istatistiği

$$Alvo = \sum_{j=1}^k \sum_{i=1}^n j \frac{(k+1)}{(k_i+1)} \mu_{ij}$$

biçiminde tanımlıdır. Ortalama ve varyansı sırasıyla

$$E(Alvo) = \frac{n \cdot k(k+1)^2}{4}$$

ve

$$V(Alvo) = \sigma_i^2 = n \frac{k_i(k+1)^2}{12 \cdot (k_i+1)} \sum_{j=1}^k (O_{ij} - \bar{O}_i)^2$$

biçiminde hesaplanır.

Burada, O_{ij} , gözlemin bloktaki işlem sırası, $\bar{O}_i$, Bloкта yer alan işlem sıra numaralarının ortalamasıdır. Alvo ve Cabilio testinin standartlaştırılmış hali

$$Z_{Alvo} = \frac{Alvo - E(Alvo)}{\sqrt{V(Alvo)}}$$

biçimindedir.

$Z_{Alvo} > Z_{\alpha}$ olduğu durumlarda sıfır hipotezi reddedilir. Burada Z_{α} standart normal dağılımın $(1-\alpha)100$ yüzdelik dilimidir. Alvo ve Cabilio (1995) testi dengeli eksik blok tasarımları için de kullanılabilir.

4.1.2. Mungai testi

Magel ve Ndungu (2011) tarafından şemsiye alternatifi için önerilen test istatistiği, rasgele tamamlanmış blok ve eksik blok tasarımları altında uygulanabilir. Test edilen sıfır hipotezi ile alternatif hipotez

$$H_0: \tau_1 = \tau_2 = \dots = \tau_t$$

$$H_1: \tau_1 \leq \tau_2 \leq \dots \leq \tau_p \geq \dots \geq \tau_t \quad (\text{En az bir eşitsizlik doğru olmak üzere})$$

biçimindedir. Magel ve Ndungu (2011) tarafından test istatistiği

$$M = \sum_{b=1}^n M_b \quad (4.3)$$

biçiminde tanımlanmıştır. Burada

- $M_b = \sum_{i < j} \sum_{j+1 > p} U_{ijb}$
- U_{ijb} Mack-Wolfe'a benzer şekilde, b bloğundaki gözlem çiftlerinin (x, y) sayısıdır ve
 - $x < y$ x ve $y <$ tepe noktası ise y tepe noktası olabilir,
 - $x > y$ x ve $y >$ tepe noktası ise x tepe noktası olabilir.
- Karşılaştırmalar yalnızca tepe noktasının aynı tarafındaki işlemlerle sınırlandırılmıştır.
- Eksik gözlemler varsa U_{ijb} değerleri için aşağıdaki kriter kullanılmaktadır. Herhangi bir kısıtlama olmaksızın, j 'nin de tepe noktası olabileceği bir $i < j \leq$ tepe çifti ele alındığında,

- i ve j, p tepe değerinden büyükse $U_{ijb} = 0$,
- i ve j küçükse, $0,5$,
- j küçükse ve i değilse, $1 - \frac{r_i}{k+1}$ (r_i, i nin blok içindeki sırası ve k , blokta görülen işlem sayısı)
- i küçük ve j değilse, $\frac{r_j}{k+1}$ (r_j, j nin blok içindeki sırası ve k , blokta görülen işlem sayısı)

Mungai'nin istatistiği için beklenen değer, işlem sayısına ve tepe işleminin konumuna bağlıdır. Genel formül

$$E(M) = \sum_{i=1}^n E(M_i)$$

dir. Burada, $E(M_i)$, i bloğunun beklenen değeri, n , IBD'deki toplam blok sayısıdır.

Beklenen değere benzer şekilde varyans da bloktan bloğa değişmektedir. Genel formül

$$V(M) = \sum_{i=1}^n \sigma_i^2 \quad i=1,2,\dots,n$$

dir. Burada, σ_i^2 i bloğunun varyansıdır. σ_i^2 'nin alacağı değerler eksik gözlemlerin modeline, işlem sayısına ve tepe noktasının konumuna bağlı olarak değişir. Standartlaştırılmış Mungai istatistiği

$$Z_M = \frac{M - E(M)}{\sqrt{V(M)}}$$

biçiminde tanımlıdır.

$Z_M > Z_\alpha$ olduğu durumlarda sıfır hipotezi reddedilir. Burada Z_α standart normal dağılımın $(1-\alpha)100$ yüzdelik dilimidir.

5. DENGELİ EKSİK BLOK TASARIMI

Eksik gözlemlerin dengeli bir model oluşturduğu ve herhangi iki işlemin eşit sayıda birlikte görüldüğü eksik blok tasarımıdır. Eksik gözlemlerin varlığında uygulamaların güvenli bir şekilde karşılaştırılmasını sağlar.

Eksik blok tasarımında

- Her işlemde aynı sayıda denek olması,
- Her deneğin aynı sayıda işleme tabi tutulması,
- Her işlem çiftinin eşit sayıda test edilmesi,

durumlarında tasarım dengeli eksik blok tasarımı (Balanced Incomplete Block Design-BIBD) dır.

Bu tasarıma ilişkin varsayımlar

- Her blok k deney birimi içerir,
- $k < t$ her blokta deney birimi sayısından daha az deneme vardır
- Her deneme r blokta görülür (r , tekrar sayısıdır),
- Her bir deneme çifti tam olarak λ sayıda blokta birlikte görülür (Duran,1995).

dır. Burada

$$\lambda = \frac{r(k-1)}{(t-1)} \quad (5.1)$$

biçiminde hesaplanmaktadır.

$t=b$ ise, tasarımın simetrik olduğu söylenebilir.

λ parametresi tam sayı olmalıdır. λ ilişkisini türetmek için herhangi bir işlemi ele alalım. Örneğin birinci işlem r blokta yer aldığından ve bu blokların her birinde $k-1$ tane daha işlem olduğundan, birinci işlemi içeren bir blokta, $r(k-1)$ gözlem vardır. Bu $r(k-1)$ sayıdaki gözlem aynı zamanda kalan $t-1$ işlem sürecini de temsil etmelidir. Bu yüzden;

$\lambda(t-1) = r(k-1)$ dir.

Dengeli Eksik Blok Tasarımının matematiksel modeli

$$Y_{ijk} = \mu + \tau_i + \beta_j + \epsilon_{ijk}, \quad i=1, \dots, t, \quad j=1, \dots, b, \quad k=1, \dots, n_{ij} \quad (5.2)$$

biçiminde tanımlıdır. Burada, t , deneme sayısı, b , blok sayısı, n_{ij} , j . blok ve i . deneme kombinasyonları olan (0 ya da 1) gözlemlerin sayısıdır.

Örnek olarak, bir kimya mühendisinin kimyasal bir süreç için reaksiyon süresinin, kullanılan katalizör türünün bir fonksiyonu olduğunu düşündüğü varsayalım. Halihazırda dört katalizör araştırılmaktadır. Deneysel prosedür, hammadde grubunun seçilmesi, pilot tesisin yüklenmesi, her bir katalizörün pilot tesisin ayrı bir çalışmasında uygulanması ve reaksiyon süresinin gözlemlenmesinden oluşmaktadır. Hammadde gruplarındaki varyasyonlar katalizörlerin performansını etkileyebileceğinden, mühendis hammadde gruplarını blok olarak kullanmaya karar verir. Ancak, her grup yalnızca üç katalizörün çalıştırılmasına izin verecek kadar büyüktür. Bu nedenle, dengeli eksik blok tasarımı kullanılmalıdır (Montgomery, 1997)

Katalizör deneyi için parametreleri; $t=4$, $b=4$, $k=3$, $r=3$, $\lambda=2$ ve $N=12$ olan, dengeli eksik blok tasarımına uygun veri yapısı Çizelge 5.1’de verilmiştir.

Çizelge 5.1. Dengeli eksik blok tasarımının veri yapısı

	Blokler			
	B_1	B_2	B_3	B_4
A_1	Y_{11}	Y_{12}	Y_{13}	
A_2	Y_{21}		Y_{23}	Y_{24}
A_3	Y_{31}	Y_{32}		Y_{34}
A_4		Y_{42}	Y_{43}	Y_{44}

5.1. Dengeli Eksik Blok Tasarımı İçin Parametrik Olmayan Testler

5.1.1. Durbin testi

Durbin (1951) testi, eksik gözlemlerin varlığında özellikle BIBD’de, işlem etkileri arasındaki farklılıkları test etmek için geliştirilmiştir. Blokların birbirinden bağımsız olması

ve gözlemlerin büyüklük sırasına göre sıralanabilir olması Durbin (1951) testinin iki temel varsayımıdır. Farklılığın yönü ile değil, sadece işlemler arasındaki farklılıklarla ilgilenir. Page'in testiyle aynı sıralama ölçütlerini izler. Test edilen sıfır hipotezi ile alternatif hipotez

$$H_0: \tau_1 = \tau_2 = \dots = \tau_t$$

$$H_1 : \text{En az bir } \tau_i \text{ farklıdır.} \quad i=1,2,\dots,t$$

biçimindedir.

Bu testte eksik gözlemler için sıfır derecesi atanır. Ancak Durbin(1951) testi, sıraların karelerinin toplamına dayanır.

$$L = \sum_{j=1}^t R_j^2 \quad (5.3)$$

Daniel (Daniel, 1990) tarafından verilen test istatistiği,

$$T = \frac{12(t-1)}{rt(k-1)(k+1)} \sum_{j=1}^t R_j^2 - \frac{3r(t-1)(k+1)}{k-1} \quad (5.4)$$

biçiminde tanımlıdır. Burada, t , toplam işlem sayısı, k , blok başına toplam denek sayısı

($k < t$), r , her bir işlemin gerçekleşme sayısıdır.

Durbin testi yeterince büyük örneklem hacmi için ($k-1$) serbestlik derecesi ile ki-kare dağılır

$T > \chi_{k-1, \alpha}^2$ olduğu durumlarda H_0 reddedilir.

5.1.2. Hemmer testi

Hemmer (2012), tepe noktası (p) bilinen şemsiye alternatifli dengeli eksik blok tasarımında kullanılmak üzere bir test geliştirmiştir. Test edilen sıfır hipotezi ile alternatif hipotez

$$H_0: \tau_1 = \tau_2 = \dots = \tau_t$$

$$H_1: \tau_1 \leq \tau_2 \leq \dots \leq \tau_p \geq \dots \geq \tau_t \quad (\text{En az bir eşitsizlik doğru olmak üzere})$$

biçimindedir. Geliştirilen test istatistiği,

$$T = \sum_{j=1}^b \{ \sum_{i=1}^t (t - 1 - |p - i|) * R_i \} \quad (5.5)$$

biçiminde tanımlıdır. Burada, t , işlem sayısı, b , blok sayısı, p , tepe noktası, R_i , i işlemi yapılan gözlemlerin sıralarının toplamıdır.

$k < t$ olmak üzere her blokta k tane işlem olduğu varsayılır ve her bloktaki söz konusu gözlemler sadece blok içinde 1'den k ' ya kadar sıralanır. T 'nin beklenen değerleri ve varyansları elde edilirse, Hemmer'in test istatistiği standardize edilebilir.. Standartlaştırılmış test istatistiği

$$Z_T = \frac{T - E(T)}{\sqrt{V(T)}}$$

biçiminde tanımlıdır.

Z_T istatistiği, büyük örneklem hacimleri için standart normal dağılır. $Z_T > Z_\alpha$ ise H_0 reddedilir. Burada Z_α , üzerinde α alanının bulunduğu standart normal dağılım değerini göstermektedir.

6. KARMA TASARIMLAR

İstatistiksel deney veya hipotez testine başlamadan önce araştırmacıların deney tasarımı modeline karar vermesi gerekir. Araştırmacılar deneye tek tip tasarım yapısıyla başlayabilirler ancak süreç içerisinde çeşitli nedenlerle çalışmaya aynı tasarımla devam etmek mümkün olmayabilir. Böyle durumlarda karma tasarımlar kullanılabilir. Karma tasarımlar için deneysel tasarımların farklı kombinasyonunu içeren çalışmalar yapılmıştır.

Magel ve diğerleri, (2012) tarafından önerilen ve sınava giren öğrencilere not veren 4 sınıf öğretmenin (değerlendirici), puanlama biçimlerinde fark olup olmadığının test edildiği, tamamen rasgele tasarım ve tamamlanmış rasgele blok tasarımından oluşan deney karma tasarımlara örnek olarak verilebilir. Deneye, 15 öğrencinin her birine dört değerlendirici tarafından not verilmesiyle başlanıyor. Daha sonra deneye bu şekilde devam etmenin oldukça maliyetli olduğu fark ediliyor ve geriye kalan 20 öğrenci, 4 değerlendiriciye rastgele atanıyor. Bu şekilde her değerlendiriciye beşer öğrenci veriliyor. Burada öğrenciler blok, değerlendiriciler ise işlem olarak ele alınmıştır. Bu karma tasarım için geliştirilen T_1 test istatistiği, rasgele tamamlanmış blok tasarımı kısmı için uygulanan Friedman (F) test istatistiği ile tamamen rasgele tasarım kısmı için uygulanan Kruskal Wallis (KW) test istatistiği toplanarak elde edilmiştir. Değerlendiriciler tarafından öğrencilere verilen notlar Çizelge 6.1’de verilmiştir.

Çizelge 6.1. Dört değerlendirici tarafından öğrencilere verilen notlar

Değerlendirici (işlem) Öğrenci (blok)				
	1	2	3	4
1	80	83	74	85
2	62	67	60	64
3	73	80	69	82
4	85	90	75	88
5	50	57	48	60
6	79	84	71	87
7	90	95	86	97
8	88	93	84	92
9	83	90	80	91
10	79	78	81	84
11	66	67	60	69
12	74	78	70	79
13	82	79	85	78

Çizelge 6.1. (devam) Dört değerlendirici tarafından öğrencilere verilen notlar

14	91	94	89	95
15	83	80	81	84
16	66			
17	75			
18	88			
19	77			
20	92			
21		70		
22		77		
23		90		
24		76		
25		95		
26			65	
27			86	
28			72	
29			89	
30			74	
31				80
32				90
33				70
34				94
35				78
Dört sınıf öğretmeni 35 öğrencinin sınavlarına not veriyor.				

Çizelge 6.2. T_1 test istatistiğinin rasgele tamamlanmış blok tasarımı kısmının nasıl hesaplandığını göstermeye yardımcı olmak için Friedman (F) testinin sıralamalarını vermektedir.

Çizelge 6.2. Rasgele tamamlanmış blok tasarımının sıralamaları

Değerlendirici (işlem) Öğrenci (blok)	1	2	3	4
1	2	3	1	4
2	2	4	1	3
3	2	3	1	4
4	2	4	1	3
5	2	3	1	4
6	2	3	1	4
7	2	3	1	4
8	2	4	1	3
9	2	3	1	4
10	2	1	3	4
11	2	3	1	4
12	2	3	1	4
13	3	2	4	1
14	2	3	1	4
15	3	1	2	4

Sıra deęerleri: $R_1=32$, $R_2=43$, $R_3=21$, $R_4=54$ olarak hesaplanmıřtır.

Friedman test istatistięi

$$F = \frac{12}{bk(k+1)} \sum_{j=1}^k \left[R_j - \frac{b(k+1)}{2} \right]^2$$

biçiminde hesaplanır.

Buradan

$$F = \frac{12}{15 \cdot 4 \cdot (4+1)} \sum_{j=1}^4 \left[R_j - \frac{15 \cdot (4+1)}{2} \right]^2$$

=24,2 dir.

$F \sim \chi^2_{k-1}$ olduęundan

$$p\text{-deęeri} = P(\chi^2_3 > 24,2) = 0,00002$$

bulunur.

Çizelge 6.3. T_1 test istatistięinin tamamen rasgele tasarım haline getirilmiř kısmının nasıl hesaplandığını göstermeye yardımcı olmak için Kruskal-Wallis (KW) testinin sıralamalarını vermektedir.

Çizelge 6.3. Tamamen rasgele tasarımın sıralamaları

Deęerlendirici (iřlem) Öęrenci (blok)	1	2	3	4
16	2			
17	7			
18	14			
19	9.5			
20	18			
21		3.5		
22		9.5		

Çizelge 6.3. (devam) Tamamen rasgele tasarımın sıralamaları

23		16.5		
24		8		
25		20		
26			1	
27			13	
28			5	
29			15	
30			6	
31				12
32				16.5
33				3.5
34				19
35				11

$$R_1=50.5, \quad R_2=57.5, \quad R_3=40, \quad R_4=62$$

Kruskal Wallis test istatistiği

$$KW = \frac{12}{N*(N+1)} \sum_{i=1}^k \frac{1}{n_i} \left[R_j - \frac{n_i*(N+1)}{2} \right]^2$$

biçiminde hesaplanır.

Buradan

$$KW = \frac{12}{20*(20+1)} \sum_{i=1}^4 \frac{1}{5} \left[R_j - \frac{5*(20+1)}{2} \right]^2$$

$$=1,57 \text{ dir.}$$

$$T_1 = F + KW \quad \text{olduğundan} \quad (6.1)$$

$$T_1 = 24,2 + 1,57 = 25,77 \text{ dir.}$$

$$T_1 \sim \chi_{k-1}^2 + \chi_{k-1}^2 \sim \chi_6^2 \text{ olduğundan}$$

$$p\text{-değeri} = P(\chi_6^2 > 25,77) = 0,00025$$

bulunur.

Sonuç olarak, sadece 15 bloklı tamamlanmış rasgele blok tasarımını kullanan test istatistiği 24,2 ve p -değeri 0,00002'dir. Friedman ve Kruskal-Wallis testlerinin toplamından oluşan T_1 test istatistiği 25,77 ve p -değeri 0,00025'tir. Dolayısıyla, işlem etkilerinin birbirine eşit olduğu sıfır hipotezi reddedilir.

4 değerlendiricinin puanlama biçimlerinde fark olup olmadığı örneği R programının 4.3.2 versiyonunda test edilmiştir. Çizelge 5.'te yer alan veriler programa aktararak Kruskal Wallis, Friedman ve T_1 (kodda kombinasyon testi olarak adlandırılmıştır) test istatistikleri ile testlerin p -değerleri hesaplanmıştır. T_1 test istatistiği, Kruskal Wallis ve Friedman testlerinin toplamı olarak yazılmıştır. Sonuçların anlamlı olup olmadığını belirlemek için p -değeri, α (0,05) ile karşılaştırılmıştır. R programına yazılan kod Ek 1'de verilmiştir.

Sonuçlar Çizelge 6.4.'te verilmiştir.

Çizelge 6.4. R programı kararı

Test	Test İstatistiği	p -değeri	Karar
Friedman Testi	24,20	0,0000143	Red
Kruskal Wallis Testi	1,577845	0,6644237	Red edilemez
Kombinasyon Testi (T_1)	25,77784	0.0002449	Red

Görüldüğü gibi R (4.3.2) programından elde edilen sonuçlar daha önce bulduğumuz sonuçlarla aynıdır. T_1 testinin (Kombinasyon testi) p -değeri $< \alpha$ olduğu için işlem etkilerinin arasında fark olmadığı sıfır hipotezi reddedilir.

Karma tasarımlar farklı deney tasarım türlerinin birleştirilmesiyle oluşturulabilmektedir. Örneğin, araştırmacı başlangıçta tam bir RCBD kullanmak niyetiyle deneyini yürütebilir, ancak veriler toplanırken sorunlar ortaya çıkabilir. Denek edinmenin maliyeti öngörülenden daha yüksek olabilir veya belirli bloklar için deneyin anketine yanıt veren denek sayısı beklenenden düşük olabilir. Bunun gibi sebeplerle, her bir işlemin her bloktaki tüm gözlemlere uygulanmasını gerektiren RCBD'yi kullanmaya devam etmenin pratik olmadığı ortaya çıkabilir. Araştırmacı RCBD'den vazgeçip farklı bir tasarım kullanabilir. Ancak bu durumda RCBD ile elde edilen veriler boşa gidecektir. Gerçekte ne kadar çok gözlem olursa testin o kadar güçlü olacağı göz önünde bulundurulursa araştırmacı deneyin kalanına BIBD'yi uygulayarak devam edebilir, böylece her işlemin her blok içindeki tüm gözlemlere

uygulanması gerekmez. Tüm veriler toplandıktan sonra, RCBD ve BIBD tasarımlarının karışımı olacak şekilde bir karma tasarım organize edilir.

Karma tasarımla ilgili olarak şirket çalışanlarının kolesterol seviyesinin iyileştirilmesinin üretkenliği artırdığını test etmek için yapılacak çalışma örnek verilebilir. Çalışanların başlangıçta kolesterol seviyeleri ölçülür ve çalışanlar sağlıklı yaşam programına dahil edilir. Altı aylık takipler boyunca kolesterol ölçümleri kontrol edilir. Verimlilik, altı ay içinde tamamlanan görev sayısı ile ölçülür. Verimlilik ve kolesteroldeki değişiklikler tek tek çalışanlar üzerinde takip edileceğinden, tasarım RCBD' ye benzetilmektedir. Ancak, çalışan devri, hastalık, yaralanma veya özensizlik verilerin eksik olmasına neden olabilir. Bu nedenle şirket, takipler sırasında mevcut olan çalışma gruplarını bir araya getirerek çalışmayı tamamlamaya karar verebilir. Nihai veri seti, eksik blok tasarımı (IBD) ve CRD' nin kombinasyonundan oluşan bir karma tasarım olacaktır.

Karma tasarımla ilgili verilebilecek diğer bir örnek psikoloji alanında yapılan şu çalışma olabilir;

Hava sıcaklıklarının kişilerin ruh halini etkilediğini iddia eden bir araştırmacı yaptığı çalışmada, soğuk havaları düşük ruh haliyle ilişkilendirirken, yüksek sıcaklıkları daha mutlu ruh haliyle ilişkilendirmektedir. Bir yıl sürecek olan (Ocak-Aralık ayları arasında) çalışmada, birkaç öğrenciyi takip etmeyi ve deneye dahil edilen öğrencilerin her ay rastgele zamanlarda genel ruh hallerini ölçmeyi planlamıştır. Araştırmacı, en yüksek ruh hali seviyelerinin yaz aylarının ortasında gözlemlenmesini beklenmektedir. Başlangıçta deney tasarımı olarak CRD belirlenmiştir. Ancak öğrencilerin yatay geçiş yapması, mezun olması veya bir süre sonra deneye olan ilgisini kaybetmesi eksik gözlemler olmasına neden olmuştur. Bu durumda araştırmacı, bir ay içinde rastgele zamanlarda rastgele öğrencilerin ruh hallerini kontrol ederek çalışmayı tamamlamaya karar verebilir. Nihai veri seti yine bir CRD ve IBD kombinasyonuna benzeyecektir (Ndungu, 2014).

Karma tasarımlar kullanılarak, başlangıç tasarımıyla elde edilen verilerin kaybolması engellenir ve araştırmanın tamamlanması sağlanır.

6.1. Karma Tasarımlar için Önerilen Testler

6.1.1. Dubnicka, Blair ve Hettmansperger testi

Dubnicka ve diğerleri, (2002), RCBD'ye benzeyen eşleştirilmiş verilerden ve CRD'ye benzeyen iki bağımsız örnekten oluşan karma bir tasarım düşünmüştür. Eşleştirilmiş verilere Wilcoxon işaretli sıralama testini (1945) ve CRD kısmına Mann–Whitney (1947) testini uygulamışlardır. Yeni testler, hem test istatistik değerlerinin birleştirilmesi hem de miktarın standardize edilmesiyle oluşturulmuştur. Araştırmada, ayrıca testleri birleştirirken ağırlıkları kullanmayı da düşünmüşlerdir. Eşleştirilmiş veri örneği, birleştirilmiş iki bağımsız örnekten daha büyük olmadığı sürece ağırlıkların kullanılmasını önermiştir. Test istatistiği

$$T^+ = S^+(\Delta) + U^+(\Delta) \quad (6.2)$$

biçiminde tanımlıdır. Burada, $S^+(\Delta)$, Wilcoxon işaretli sıra istatistiği, $U^+(\Delta)$, Wilcoxon-Mann-Whitney test istatistiğidir. Sıfır hipotezi altında T^+ testi için ortalama ve varyans sırasıyla

$$E(T^+) = \frac{n(n+1)}{4} + \frac{n_1 n_2}{2}$$

ve

$$V(T^+) = \frac{n(n+1)(2n+1)}{24} + \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

dır.

Burada n_1 ve n_2 değerleri Wilcoxon-Mann-Whitney istatistiğindeki bağımsız örneklemelerin örneklem büyüklükleri, n Wilcoxon testindeki eşleştirilmiş veriler için örneklem büyüklüğüdür. T^+ testinin standartlaştırılmış hali

$$T^* = \frac{E + E(T^+)}{\sqrt{V(T^+)}}$$

dir.

Sıfır hipotezinin doğruluğu altında T^* asimptotik normal dağılmaktadır.

$Z_{T^*} > Z_\alpha$ olduğu durumlarda sıfır hipotezi reddedilir. Burada Z_α standart normal dağılımın

$(1-\alpha)100$ yüzdelerlik dilimidir.

6.1.2. Magel, Terpstra ve Wen testi

Magel ve diğerleri, (2009), RCBD ve CRD' den oluşan karma tasarımda azalmayan alternatif için iki test istatistiği önermiştir. Her iki test de Page (1963) ve Jonckheere-Terpstra testlerini kullanmıştır. İlk test istatistiği

$$Z_{comb} = \frac{Z_{page} + Z_{JT}}{\sqrt{2}} \quad (6.3)$$

biçiminde tanımlıdır.

Burada, Z_{page} , Page testinin, Z_{JT} ise Jonckheere-Terpstra testinin standartlaştırılmış halidir. H_0 'ın doğruluğu altında Z_{comb} asimptotik olarak standart normal dağılır. $Z_{comb} > Z_\alpha$ olduğu durumlarda sıfır hipotezi reddedilir.

İkinci test istatistiği

$$Z_{combII} = \frac{L + J - E(O)}{\sqrt{Var(O)}} \quad (6.4)$$

biçiminde tanımlıdır.

Burada, L ve J sırasıyla Page ve Jonckheere-Terpstra (JT) test istatistikleridir.

$$E(O) = \frac{bk(k+1)^2}{4} + \frac{(N^2 - \sum_{i=1}^k n_i^2)}{4}$$

ve

$$V(0) = \frac{b(k^3 - k)^2}{144(k-1)} + \frac{(N^2(2N+3) - \sum_{i=1}^k n_i^2(2n_i+3))}{72}$$

biçiminde tanımlıdır.

H_0 'ın doğruluğu altında Z_{combII} asimptotik standart normal dağılır. $Z_{combII} > Z_\alpha$ olduğu durumlarda sıfır hipotezi reddedilir.

İlk önerdikleri testte Page test istatistiği ile JT test istatistiği önce standardize edilmiş sonra toplanmıştır. İkinci önerdikleri testte Page test istatistiği ile JT test istatistiği önce toplanmış daha sonra standardize edilmiştir.

Çalışmaları, CRD' nin örneklem büyüklüğü RCBD'nin dörtte birinden az olmadığı sürece, standartlaştırılmış istatistiklerin toplamını standartlaştırmanın, standartlaştırılmamış istatistiklerin toplamını standartlaştırmadan daha iyi olduğunu göstermiştir.

6.1.3. Magel, Terpstra, Canonizado ve Park testi

Magel ve diğerleri, (2010), karma tasarımda eksik gözlemlerin varlığında şemsiye alternatifini değerlendirmiştir. Bu çalışmanın sonuçları, standartlaştırılmış toplamın standartlaştırılmasının, genel olarak daha iyi olduğunu göstermiştir.

6.1.4. Magel ve Mathisen testi

Magel ve Mathisen (2011), RCBD ile başlayan bir çalışmada fonların yetersiz olması nedeniyle deneye devam edilemediği için RCBD ile BIBD' den oluşan bir karma tasarımı ele almıştır. Magel ve Mathisen (2011), Page test istatistiği ile Magel ve Ndungu'nun (2011) test istatistiklerini kullanarak iki yeni istatistik oluşturmuşlardır. Bu test istatistikleri ilk standartlaştırma ve son standartlaştırmadır.

İlk standartlaştırma:

$$T_1^* = \frac{Z_{page} + Z_{Ndungu}}{\sqrt{2}} \quad (6.5)$$

Burada, Z_{page} ve Z_{Ndungu} sırasıyla Page ve Ndungu test istatistiklerinin standartlaştırılmış halidir. H_0 altında Z_{page} ve Z_{Ndungu} asimptotik standart normal dağılıma sahip olduğundan T_1^* asimptotik standart normal dağılıma sahiptir. $T_1^* > Z_\alpha$ olduğu durumlarda sıfır hipotezi reddedilir. Burada Z_α standart normal dağılımın $(1-\alpha)$ 100 yüzdelerlik dilimidir.

Son standartlaştırma:

$$T_2^* = \frac{L+M-[E(L)+E(M)]}{\sqrt{V(L)+V(M)}} \quad (6.6)$$

biçiminde tanımlıdır.

Burada, L ve M sırasıyla Page ve Ndungu test istatistikleri, $E(L)$, Page test istatistiğinin ortalaması, $V(L)$, Page test istatistiğinin varyansı, $E(M)$, Ndungu test istatistiğinin ortalaması, $V(M)$, Ndungu test istatistiğinin varyansıdır.

H_0 altında T_2^* asimptotik standart normal dağılıma sahiptir. $T_2^* > Z_\alpha$ olduğu durumlarda sıfır hipotezi reddedilir. Burada Z_α standart normal dağılımın $(1-\alpha)$ 100 yüzdelerlik dilimidir.

Simülasyon çalışmaları, tamamlanmış ve eksik blokların oranı 1:1 olduğunda, eksik bloklarda daha az gözlem olduğunda, son standartlaştırmanın daha iyi olduğunu göstermiştir. Eksik bloklarda daha fazla gözlem olduğunda ise bunun tersinin geçerli olduğunu göstermiştir. Daha fazla tamamlanmış blok olduğunda, yani oran 2:1 olduğunda da son standartlaştırma daha iyi sonuç vermiştir. Tamamlanmış blokların eksik bloklara oranı 1:2 olduğunda iki yöntem arasında anlamlı bir fark bulunamamıştır (Ndungu, 2014).

6.1.5. Magel ve Hemmer testi

Magel ve Hemmer (2012), RCBD ve BIBD tasarım kombinasyonunun şemsiye alternatifi için bir test önermiştir. Araştırmada, yeni istatistiklerini oluşturmak için Kim ve Kim (1992) ile Magel ve Hemmer'in (2012) testlerini dikkate almışlardır. Testleri birleştirmek için benzer ilk ve son standartlaştırma yöntemlerini izlemişlerdir. M_1^* ilk standartlaştırma ile oluşturulan istatistiği, M_2^* son standartlaştırma ile oluşturulan istatistiği gösterir.

M_1^* , Kim ve Kim ile Hemmer testlerinden, test istatistiklerini ayrı ayrı standartlaştırmayı, iki standartlaştırılmış istatistiği toplamayı ve toplamı yeniden standartlaştırmayı önerir. Yöntem 1 olarak önerilen bu test istatistiği,

$$M_1^* = \frac{Z_{KK} + Z_T}{\sqrt{2}} \quad (6.7)$$

biçiminde tanımlıdır.

H_0 altında Z_{KK} ve Z_T asimptotik normal dağıldığından M_1^* asimptotik standart normal dağılır. $M_1^* \geq Z_\alpha$ olduğu durumlarda sıfır hipotezi reddedilir.

Önerilen ikinci test istatistiği, önce RCBD için Kim ve Kim testinden ve BIBD için Hemmer testinden standartlaştırılmamış test istatistiklerini toplar. Yöntem 2 olarak önerilen bu ikinci test istatistiği,

$$M_2 = KK + T = \sum_{k=1}^c A_{kp} + \sum_{j=1}^d \{ \sum_{i=1}^t (t - 1 - |p - i|) * R_i \} \quad (6.8)$$

biçiminde tanımlıdır.

Burada, c , tamamlanmış blok sayısı, d , tamamlanmamış blok sayısı, t , işlem sayısı,

p , bilinen en yüksek işlem seviyesidir.

M_2 , standartlaştırılmamış bir istatistiktir. Beklenen değer ve varyans elde edildikten sonra Yöntem 2 test istatistiği

$$M_2^* = \frac{\{M_2 - (E(KK) + E(T))\}}{\sqrt{V(KK) + V(T)}} \quad (6.9)$$

biçiminde standartlaştırılabilir.

M_2^* 'nin asimptotik dağılımı standart normal dağılımdır. $M_2^* \geq Z_\alpha$ olduğu durumlarda sıfır hipotezi reddedilir.

Her iki istatistik için de karşılık gelen değerler $Z_{\alpha'}$ dan büyük olduğunda sıfır hipotezi reddedilmiştir. Simülasyon çalışmaları, ilk standartlaştırmanın, tamamlanmış ve eksik blokların oranı eşit olduğunda, daha iyi sonuç verdiğini göstermiştir. Aksi durumda, son standartlaştırma daha güçlü sonuç vermiştir. Magel ve Hemmer (2012), bu yaklaşımı göz önünde bulundurarak son standartlaştırmayı önermiştir.

Magel ve Fu (2014), Wilcoxon işaretli sıra testi ve Mann-Whitney testinin doğrusal bir kombinasyonundan oluşan farklı bir testi ele alarak testlerinin Dubnicka ve arkadaşlarının (2002) testinden daha yüksek güçlere sahip olduğu durumlar bulmuşlardır.

7. SİMÜLASYON ÇALIŞMASI

Bu bölümde üç farklı simülasyon çalışması yapılmıştır. Simülasyon çalışmalarında veriler binom dağılımından üretilmiştir. Binom dağılımın tercih edilmesinde bu dağılımın esnek olması ve p parametresi sayesinde farklı yapıları esnek bir şekilde modelleyebilmesi etkili olmuştur. İlk olarak parametrik olmayan test yöntemlerinden Page (L), Uyarlanmış Page (UP), Hollander (H), Kruskal Wallis (KW), Friedman (F) testleri deneysel I. tip hata bakımından karşılaştırılmıştır. Simülasyon çalışmasında işlem sayısı $t=3, 4$ ve 5 ; blok sayısı $b=5, 10$ ve 15 olarak alınmıştır. Her döngüde her blok-işlem çifti için 5 veri noktası üretilmiş olup, simülasyon 5000 döngü boyunca gerçekleştirilmiştir. Rasgele veriler (n, p) parametrelili binom dağılımından üretilmiş olup deneme sayısı $n=5$ olarak alınırken, p olasılık parametresinin farklı değerleri ($p=0,1, 0,2, 0,5, 0,8$ ve $0,9$) için testlerin deneysel I. tip hata oranları hesaplanmıştır. Bu şekilde verilerin üretildiği dağılımın parametresi değiştirilerek testlerin duyarlılığı incelenmek istenmiştir. Dayanıklı testler farklı dağılım yapılarında ve çeşitli p değerlerinde nominal α 'yı koruyabilen ve gücünü kaybetmeden doğru sonuçlar verme olasılığı yüksek olan testlerdir ($\alpha=0,05$ alınmıştır).

İkinci çalışmada aynı testlerin testin gücü bakımından karşılaştırması yapılmıştır. İşlem sayısı $t=3, 4$ ve 5 ; blok sayısı $b=5, 10$ ve 15 olarak alınmıştır. Her döngüde her blok-işlem çifti için 5 veri noktası üretilmiş olup, simülasyon 5000 döngü boyunca gerçekleştirilmiştir. Rasgele veriler (n, p) parametrelili binom dağılımından üretilmiş olup $n=5$ olarak alınmıştır. Güç hesaplamasında her işlem için p olasılık parametresinin farklı değerleri alınmıştır ($\alpha=0,05$ alınmıştır).

Üçüncü çalışmada Kruskal Wallis ve Friedman testleri ile bu testlerin kombinasyonundan oluşan T_1 test istatistiğinin testin gücü bakımından karşılaştırması yapılmıştır. Çalışmada rasgele tamamlanmış blok tasarımı kısmı için blok sayıları 10, 15, 20; tamamen rasgele kısmı için 6, 8, 10; işlem sayısı $t=3, 4, 5$ olarak alınmıştır. Simülasyon 5000 döngü boyunca gerçekleştirilmiştir. Rasgele veriler (n, p) parametrelili binom dağılımından üretilmiş olup $n=5$ olarak alınmıştır. Güç hesaplamasında her işlem için p olasılık parametresinin farklı değerleri alınmıştır ($\alpha=0,05$ alınmıştır).

Simülasyon çalışmaları R 4.3.2 programında yapılmış olup, programda bulunan “DescTools”, “dplyr”, “tidyr” ve “clinfun” kütüphanelerinden yararlanılmıştır.

İlk olarak binom dağılımından veri üretilerek parametrik olmayan testlerin deneysel I. tip hata oranları hesaplanmıştır. İşlem ve blok sayısındaki değişimlerin testlerin deneysel I. tip hata oranını nasıl etkilediği ve hangi test yönteminin nominal α 'ya daha yakın sonuçlar verdiği incelenmiştir. Simülasyona ilişkin sonuçlar Çizelge 7.1 ile Çizelge 7.2'de yer almaktadır.

Çizelge 7.1. ($n=5$, $p=0.1, 0.2, 0.5$ parametrelili binom dağılımı) deneysel I. tip hata

p	t	b	L	UP	H	KW	F
$p=0,1$		$b=5$	0,0134	0,0142	0,0264	-	-
	$t=3$	$b=10$	0,0174	0,0192	0,0224	0,0454	0,0456
		$b=15$	0,0142	0,0142	0,0256	0,0478	0,0476
		$b=5$	0,0128	0,0138	0,0300	0,0348	0,0310
	$t=4$	$b=10$	0,0208	0,0150	0,0288	0,0488	0,0454
		$b=15$	0,0192	0,0180	0,0268	0,0426	0,0418
		$b=5$	0,0218	0,0184	0,0258	0,0344	0,0328
	$t=5$	$b=10$	0,0160	0,0218	0,0256	0,0438	0,0404
		$b=15$	0,0186	0,0170	0,0292	0,0428	0,0440
$p=0,2$		$b=5$	0,0278	0,0250	0,0344	0,0394	0,0446
	$t=3$	$b=10$	0,0314	0,0282	0,0316	0,0444	0,0426
		$b=15$	0,0188	0,0248	0,0334	0,0436	0,0444
		$b=5$	0,0208	0,0258	0,0298	0,0390	0,0344
	$t=4$	$b=10$	0,0256	0,0294	0,0418	0,0428	0,0436
		$b=15$	0,0340	0,0348	0,0376	0,0524	0,0472
		$b=5$	0,0350	0,0340	0,0406	0,0348	0,0380
	$t=5$	$b=10$	0,0280	0,0282	0,0444	0,0432	0,0422
		$b=15$	0,0330	0,0266	0,0452	0,0438	0,0438
$p=0,5$		$b=5$	0,0330	0,0334	0,0362	0,0300	0,0310
	$t=3$	$b=10$	0,0338	0,0296	0,0416	0,0516	0,0512
		$b=15$	0,0244	0,0318	0,0400	0,0458	0,0442
		$b=5$	0,0266	0,0266	0,0354	0,0404	0,0404
	$t=4$	$b=10$	0,0300	0,0288	0,0424	0,0450	0,0408
		$b=15$	0,0342	0,0330	0,0406	0,0442	0,0478
		$b=5$	0,0384	0,0394	0,044	0,0350	0,0370
	$t=5$	$b=10$	0,0382	0,0366	0,0436	0,0390	0,0396
		$b=15$	0,0366	0,0310	0,0428	0,0478	0,0476

Çizelge 7.1'de simülasyon çalışması sonucu ($n=5$, $p=0,1$) parametrelili binom dağılımından üretilen veriler üzerinde yapılan testlerin deneysel I. tip hata oranları verilmiştir. Kruskal Wallis testi nominal α 'ya en yakın deneysel I. tip hata oranına sahiptir. Örneğin işlem sayısı $t=3$, blok sayısı $b=15$ olduğunda Kruskal Wallis testinin deneysel I. tip hata oranı 0,0478 iken, işlem sayısı $t=4$, blok sayısı $b=15$ olduğunda 0,0426'dır.

Page testi nominal α 'ya en uzak deneysel I. tip hata oranlarına sahiptir. Örneğin, işlem sayısı $t=3$, blok sayısı $b=10$ olduğunda; Page testinin deneysel I. tip hata oranı 0,0174, Uyarlanmış

Page testinin 0,0192, Hollander testinin 0,0224, Kruskal Wallis testinin 0,0454, Friedman testinin 0,0456'dır.

Hollander testinin deneysel I. tip hata oranı, işlem sayısı $t=3$ olduğunda; $b=5$ iken 0.0264, $b=10$ iken 0.0224, $b=15$ iken 0,0256'dır. Blok sayısı arttıkça deneysel I. tip hata oranının çok fazla değişmediği birbirine yakın sonuçlar verdiği görülmektedir.

Blok sayısı arttığında Kruskal Wallis testinin deneysel I. tip hata oranı nominal α 'ya yaklaşmıştır. Örneğin, işlem sayısı $t=4$ olduğunda, blok sayısı $b=5$ iken 0,0348, $b=10$ iken 0,0488'dir. Ancak genel olarak bakılığında blok sayısı arttıkça Kruskal Wallis testinin deneysel I. tip hata oranı çok fazla değişmemiştir.

Friedman testinin deneysel I. tip hata oranı işlem sayısı arttıkça birbirine yakın sonuçlar vermiştir. Örneğin, blok sayısı $b=10$ olduğunda, işlem sayısı $t=3$ iken 0,0456, işlem sayısı $t=4$ iken 0,0454, işlem sayısı $t=5$ iken 0,0404'tür.

Sonuçlar genel olarak değerlendirildiğinde simülasyon çalışmasında binom dağılımının parametreleri ($n=5$, $p=0,1$) olarak alındığında nominal α 'ya en yakın deneysel I. tip hata oranına sahip testler Kruskal Wallis ve Friedman testidir.

Binom dağılımının p parametresi 0,2 olarak alındığında nominal α 'ya en yakın deneysel I. tip hata oranını yine Kruskal Wallis ve Friedman testleri vermektir. Örneğin, işlem sayısı $t=3$ olduğunda Friedman testinin deneysel I. tip hata oranı blok sayısı $b=5$ iken 0,0446, $b=10$ iken 0,0426, $b=15$ iken 0,0444'tür.

Kruskal Wallis testinin deneysel I. tip hata oranı, işlem sayısı arttıkça nispeten azalmıştır. Örneğin blok sayısı $b=5$ olduğunda Kruskal Wallis testinin deneysel I. tip hata oranı işlem sayısı $t=3$ iken 0,0394, $t=4$ iken 0,0390 ve $t=5$ iken 0,0348'dir.

Verilerin üretildiği binom dağılımının p olasılık parametresi değiştirildiğinde ($n=5$) Page, Uyarlanmış Page ve Hollander testlerinin deneysel I. tip hata oranının arttığı görülmektedir. Örneğin işlem sayısı $t=5$, blok sayısı $b=5$ olduğunda Hollander testinin deneysel I. tip hata oranı; $p=0,1$ iken 0.0258, $p=0,2$ iken 0.0406 'dır ve nominal α 'ya yaklaşmıştır. Aynı şekilde Page testinin deneysel I. tip hata oranı; $p=0,1$ iken 0,0218, $p=0,2$ iken 0,0350'dir. Kruskal

Wallis ve Friedman testi parametre değişiminden daha az etkilenmiştir. Örneğin işlem sayısı $t=5$, blok sayısı $b=5$ olduğunda Kruskal Wallis testinin deneysel I. tip hata oranı; $p=0,1$ iken 0.0344, $p=0,2$ iken 0.0348'dir.

Binom dağılımının p parametresi 0,5 olarak belirlendiğinde testlerin deneysel I. tip hata oranlarının arttığı görülmektedir. Page, Uyarlanmış Page ve Hollander testlerinin deneysel I. tip hata oranları genel olarak nominal α 'ya yaklaşmıştır. Örneğin işlem sayısı $t=3$ blok sayısı $b=10$ olduğunda Hollander testinin deneysel I. tip hata oranı $p=0,2$ iken 0,0316, $p=0,5$ iken 0,0416'dır. İşlem sayısı $t=5$ blok sayısı $b=5$ olduğunda Hollander testinin deneysel I. tip hata oranı $p=0,2$ iken 0,0406, $p=0,5$ iken 0,0440'dır.

İşlem sayısı $t=5$ olduğunda blok sayısı arttıkça Page, Uyarlanmış Page ve Hollander testinin deneysel I. tip hata oranı azalmış, Kruskal Wallis, ve Friedman testinin ise blok sayısı arttıkça deneysel I. tip hata oranları artmıştır.

Blok sayısının değişiminin deneysel I. tip hata oranına etkisi farklı şekillerde olmuştur. Örneğin işlem sayısı $t=4$ olduğunda Page testinin deneysel I. tip hata oranı blok sayısı arttıkça artmış ($b=5$ iken 0,0266, $b=10$ iken 0,0300 $b=15$ iken 0,0342'dir); $t=5$ olduğunda blok sayısı arttıkça yaklaşık olarak aynı sonuçları ($b=5$ iken 0,0384, $b=10$ iken 0,0382, $b=15$ iken 0,0366) vermiştir.

Çizelge 7.1' e göre $p=0,5$ olduğunda nominal α 'ya en yakın deneysel I. tip hata oranını veren testler Kruskal Wallis ve Friedman testleridir. Örneğin, blok sayısı $b=15$ olduğunda, $t=3$ iken Kruskal Wallis testinin (0,0458), $t=4$ iken Friedman testinin (0,0478), $t=5$ iken yine Kruskal Wallis testinin deneysel I. tip hata oranı nominal α 'ya daha yakın sonuçlar vermiştir.

Çizelge 7.2. ($n=5$, $p=0,8$, $0,9$ parametrelili binom dağılımı) deneysel I. tip hata

p	t	b	L	UP	H	KW	F
$p=0,8$	$t=3$	$b=5$	0,0278	0,0236	0,0370	0,0430	0,0460
		$b=10$	0,0354	0,0294	0,0390	0,0476	0,0482
		$b=15$	0,0248	0,0254	0,0382	0,0494	0,0486
	$t=4$	$b=5$	0,0184	0,0234	0,0392	0,0406	0,0386
		$b=10$	0,0274	0,0262	0,0414	0,0468	0,0428
		$b=15$	0,0282	0,0324	0,0458	0,0458	0,0428
	$t=5$	$b=5$	0,0370	0,0298	0,0392	0,0348	0,0292
		$b=10$	0,0272	0,0300	0,0364	0,0434	0,0456
		$b=15$	0,0270	0,0298	0,0406	0,0426	0,0414

Çizelge 7.2. (devam) ($n=5$, $p=0.8$, 0.9 parametrelili binom dağılımı) deneysel I. tip hata

$p=0,9$		$b=5$	0,0102	0,0152	0,0194	-	-
	$t=3$	$b=10$	0,0174	0,0192	0,0270	0,0436	0,0468
		$b=15$	0,0144	0,0168	0,0258	0,0482	0,0452
		$b=5$	0,0136	0,0132	0,0254	0,0402	0,0362
	$t=4$	$b=10$	0,0154	0,0154	0,0268	0,0410	0,0388
		$b=15$	0,0190	0,0230	0,027	0,0468	0,0452
		$b=5$	0,0200	0,0186	0,0286	0,0332	0,0326
	$t=5$	$b=10$	0,0184	0,0176	0,0270	0,0368	0,0350
		$b=15$	0,0216	0,0152	0,0294	0,0510	0,0456

Çizelge 7.2'ye göre binom dağılımının p parametresi $0,8$ 'e artırıldığında Page testinin deneysel I. tip hata oranı azalmıştır. Örneğin, işlem sayısı $t=4$, blok sayısı $b=5$ olduğunda Page testinin deneysel I. tip hata oranı $p=0,5$ iken $0,0266$, $p=0,8$ iken $0,0184$ 'tür.

Hollander testi p parametresinin değişiminden nispeten daha az etkilenmiştir. Örneğin işlem sayısı $t=3$, blok sayısı $b=3$ olduğunda Hollander testinin deneysel I. tip hata oranı $p=0,5$ iken $0,0362$, $p=0,8$ iken $0,0370$ 'dir.

$p=0,8$ iken Kruskal Wallis testinin deneysel I. tip hata oranı nominal α 'ya yakın değerlerde olup blok sayısı arttıkça güç değerinin arttığı görülmektedir. Örneğin, işlem sayısı $t=3$ olduğunda, blok sayısı $b=5$ iken deneysel I. tip hata oranı $0,0430$ iken $b=10$ iken $0,476$, $b=15$ iken $0,0494$ 'tür. İşlem sayısı arttıkça Kruskal Wallis testinin deneysel I. tip hata oranı azalmıştır. Örneğin blok sayısı $b=5$ olduğunda işlem sayısı $t=3$ iken deneysel I. tip hata oranı $0,0430$, $t=4$ iken $0,0406$, $t=5$ iken $0,0348$ 'dir.

Friedman testinin deneysel I. tip hata oranı genel olarak nominal α 'ya yakın değerler vermiştir. İşlem sayısı arttıkça deneysel I. tip hata oranı azalmıştır. Örneğin blok sayısı $b=15$ olduğunda işlem sayısı $t=3$ iken deneysel I. tip hata oranı $0,0486$, $t=4$ iken $0,0428$, $t=5$ iken $0,0414$ 'tür.

Verilerin üretildiği binom dağılımının p parametresi $0,9$ olduğunda Page, Uyarlanmış Page ve Hollander testinin deneysel I. tip hata oranı azalmıştır. Kruskal Wallis ve Friedman testi nominal α 'ya daha yakın deneysel I. tip hata oranı veren yöntemlerdir. Örneğin, işlem sayısı $t=3$ olduğunda blok sayısı $b=10$ iken Friedman testinin deneysel I. tip hata oranı $0,0468$, Kruskal Wallis testinin $0,0436$ 'dır.

Blok sayısı arttıkça Kruskal Wallis ve Friedman testinin deneysel I. tip hata oranının arttığı görülmektedir. Örneğin işlem sayısı $t=4$ olduğunda, blok sayısı $b=5$ iken Kruskal Wallis testinin deneysel I. tip hata oranı $b=5$ iken 0,0402, $b=10$ iken 0,0410, $b=15$ iken 0,0468'dir ve blok sayısı arttıkça nominal α 'ya daha yakın sonuçlar vermiştir.

Deneysel I. tip hata oranı için yapılan simülasyon çalışması genel olarak incelendiğinde, verilerin üretildiği binom dağılımının parametre değişiminden en çok etkilenen yöntemler Page ve Uyarlanmış Page testidir. Hollander testi, Page ve Uyarlanmış Page testlerine göre parametre değişiminden daha az etkilenmiştir. En az etkilenen testler Kruskal Wallis ve Friedman testleridir.

Genel olarak bakıldığında nominal α 'ya en yakın deneysel I. tip hata oranına sahip yöntemler Kruskal Wallis ve Friedman testleridir. Hollander testi, Page ve Uyarlanmış Page testine göre nominal α 'ya daha yakın sonuçlar vermiştir.

Kruskal Wallis ve Friedman testinin bazı durumlarda deneysel I. tip hata oranlarının hesaplanmadığı görülmektedir. Simülasyon sırasında kullanılan parametrelerin çok küçük veya büyük belirlenmesi testlerin mantıklı sonuçlar üretmesini engelleyebilmektedir. Özellikle verilerin üretildiği binom dağılımının p parametresinin 0,1 ve 0,9 gibi uç değerler olarak belirlendiği ve blok sayısının $b=5$ olduğu çalışmalarda testlerin deneysel I. tip hata oranı hesaplanamamıştır.

Simülasyon çalışmasının ikinci aşamasında parametrik olmayan test yöntemleri testlerin gücü bakımından karşılaştırılmıştır. Çalışmada alternatif hipotez altında veri üretilmiş olup alternatif hipotezin kabul edilmesi sıfır hipotezinin doğru şekilde reddedilmesi olarak değerlendirilmiş ve güç olarak hesaba katılmıştır. İşlem sayısı 3, 4 ve 5 olarak alınmış ve her işlem için farklı p değerleri belirlenmiştir. İşlemlerin p değerleri arasındaki artış 0,05 veya 0,1 belirlenerek artış farkının sonuçları nasıl etkilediği incelenmiştir. Testlerin güç değerlerine ilişkin sonuçlar Çizelge 7.3 -Çizelge 7.5' te yer almaktadır.

Çizelge 7.3. $t=3$ için testlerin güç değeri

b	p			L	UP	H	KW	F
	0,10	0,15	0,20	0,1200	0,1286	0,1782	0,0910	-
	0,10	0,20	0,30	0,3662	0,3638	0,4488	0,2388	0,2250
	0,45	0,50	0,55	0,1092	0,1004	0,1316	0,0652	0,0636
$b=5$	0,40	0,50	0,60	0,2764	0,2814	0,3406	0,1698	0,1428
	0,70	0,75	0,80	0,1180	0,1116	0,1542	0,0778	0,0714
	0,70	0,80	0,90	0,3580	0,3600	0,4440	0,2510	0,2064
	0,10	0,15	0,20	0,2470	0,2486	0,3138	0,1924	0,1786
	0,10	0,20	0,30	0,6696	0,6638	0,7506	0,5458	0,4908
	0,45	0,50	0,55	0,1846	0,1754	0,2138	0,1190	0,1030
$b=10$	0,40	0,50	0,60	0,4930	0,5028	0,5836	0,3612	0,3106
	0,70	0,75	0,80	0,2102	0,2066	0,2648	0,1380	0,1234
	0,70	0,80	0,90	0,6604	0,6670	0,7636	0,5514	0,4862
	0,10	0,15	0,20	0,3288	0,3168	0,4552	0,2862	0,2514
	0,10	0,20	0,30	0,8182	0,8136	0,9006	0,7582	0,7056
	0,45	0,50	0,55	0,2270	0,2308	0,2898	0,1628	0,1534
$b=15$	0,40	0,50	0,60	0,6290	0,6330	0,7632	0,5548	0,4814
	0,70	0,75	0,80	0,2546	0,2568	0,3674	0,2030	0,1804
	0,70	0,80	0,90	0,8140	0,8132	0,8998	0,7556	0,7080

Çizelge 7.3 incelendiğinde Hollander testinin diğer testlere göre daha yüksek güç değerlerine sahip olduğu görülmektedir. Blok sayısı arttıkça Hollander testinin gücü artmıştır. Örneğin $p=0,10, 0,20, 0,30$ olduğunda blok sayısı $b=5$ iken Hollander testinin güç değeri 0,4488, $b=10$ iken 0,7506, $b=15$ iken 0,9006'dır.

Page ve Uyarlanmış Page testlerinin, Hollander testinden daha düşük güç değerlerine sahipken Kruskal Wallis ve Friedman testlerine göre daha büyük güç değerlerine sahip olduğu görülmektedir.

Friedman testinin diğer testlere göre daha düşük güç değerlerine sahip olduğu görülmektedir ve bu değerler 0,1428 ile 0,7080 arasında değişmektedir.

Blok sayısı arttığında tüm testlerin güç değerlerinde belirgin bir artış gözlemlenmektedir. Örneğin $p= 0,1, 0,15, 0,20$ olduğunda Page testinin güç değeri blok sayısı $b=5$ iken 0,1200, $b=10$ iken 0,2470, $b=15$ iken 0,3288'dir.

İşlemlerin p değerleri arasındaki fark arttıkça testlerin gücü artmıştır. Örneğin blok sayısı $b=5$ olduğunda, $p=0,1, 0,15, 0,20$ iken Hollander testinin gücü 0,1782; $p=0,1, 0,2, 0,3$ iken 0,4488; $p=0,45, 0,50, 0,55$ iken 0,1316'dır. Benzer şekilde blok sayısı $b=10$ olduğunda $p=0,45, 0,50, 0,55$ iken Hollander testinin gücü 0,2138; $p=0,4, 0,5, 0,6$ iken 0,5836; $p=0,70, 0,75, 0,80$ iken 0,2648'dir. İşlemlerin p değerlerinin arasındaki farkın nispeten daha az

olması (0.05), alternatif hipotezlerin algılanmasını zorlaştırarak testin güç değerlerinin düşük olmasına yol açmış olabilir.

Çizelge 7.4. $t=4$ için testlerin güç değeri

b	p				L	UP	H	KW	F
	0,10	0,15	0,20	0,25	0,2356	0,2474	0,3506	0,1436	0,1168
	0,10	0,20	0,30	0,40	0,6620	0,6674	0,7970	0,4506	0,3666
	0,40	0,45	0,50	0,55	0,1822	0,1768	0,2544	0,0896	0,0860
$b=5$	0,40	0,50	0,60	0,70	0,5496	0,5490	0,6946	0,3158	0,3634
	0,60	0,70	0,80	0,90	0,6716	0,6808	0,8020	0,4342	0,3716
	0,75	0,80	0,85	0,90	0,2518	0,2506	0,3496	0,1414	0,1196
	0,10	0,15	0,20	0,25	0,5150	0,5052	0,6038	0,3256	0,3044
	0,10	0,20	0,30	0,40	0,9482	0,9500	0,9730	0,8478	0,7998
	0,40	0,45	0,50	0,55	0,3550	0,3490	0,4366	0,1838	0,1780
$b=10$	0,40	0,50	0,60	0,70	0,8740	0,8688	0,9234	0,7066	0,6436
	0,60	0,70	0,80	0,90	0,9440	0,9434	0,9754	0,8550	0,8052
	0,75	0,80	0,85	0,90	0,5060	0,5072	0,5972	0,3302	0,2886
	0,10	0,15	0,20	0,25	0,7094	0,6918	0,7660	0,5076	0,4712
	0,10	0,20	0,30	0,40	0,9946	0,9914	0,9974	0,9726	0,9542
	0,40	0,45	0,50	0,55	0,5098	0,5056	0,5626	0,2960	0,2816
$b=15$	0,40	0,50	0,60	0,70	0,9690	0,9696	0,9816	0,9026	0,8490
	0,60	0,70	0,80	0,90	0,9956	0,9958	0,9974	0,9722	0,9516
	0,75	0,80	0,85	0,90	0,6838	0,7050	0,7702	0,5112	0,4634

Çizelge 7.4 incelendiğinde Friedman testinin diğer testlere göre daha düşük güç değerlerine sahip olduğu görülmektedir. Blok sayısı arttıkça güç değerlerinde artış gözlemlenmektedir. Örneğin; $p=0,1, 0,2, 0,3, 0,4$ olduğunda blok sayısı $b=5$ iken Friedman testinin güç değeri 0,3666, $b=10$ iken 0,7998, $b=15$ iken 0,9542'dir.

Hollander testi diğer testlere göre yüksek güç değerlerine sahiptir. Blok sayısı arttıkça güç değerleri artmaktadır. Özellikle p değerleri arasındaki fark nispeten fazla olduğunda diğer testler gibi Hollander testinin de güç değerinin daha yüksek olduğu görülmektedir. Örneğin, blok sayısı $b=10$ olduğunda Hollander testinin güç değeri $p=0,40, 0,45, 0,50, 0,55$ iken 0,4366, $p=0,4, 0,5, 0,6, 0,7$ iken 0,9234, $p=0,75, 0,80, 0,85, 0,90$ iken 0,5972'dir. Görüldüğü gibi p daha büyük değerler aldıkça güç değerleri azalmıştır. Ayrıca p değerleri arasındaki fark azaldıkça güç değerlerinin daha düşük olduğu görülmektedir. Bu p değerleri uygun bir seviyede tutulduğunda alternatif hipotezlerin daha kolay algılanabileceğini göstermektedir.

Page ve Uyarlanmış Page testlerinin güç değerlerinin Hollander testinden düşük ama Kruskal Wallis ve Friedman testlerinden daha yüksek olduğu görülmektedir. Örneğin, blok sayısı $b=10$, $p=0,6, 0,7, 0,8, 0,9$ olduğunda Page testinin güç değeri 0,9440, Uyarlanmış

Page testinin 0,9434, Hollander testinin 0,9754, Kruskal Wallis testinin 0,8550, Friedman testini 0,8052'dir.

İşlem sayısı arttıkça testlerin güç değerleri belirgin şekilde artmıştır. Örneğin, işlem sayısı $t=3$, blok sayısı $b=15$ olduğunda Page testinin gücü 0,2270 ile 0,8140 arasında değer alırken, işlem sayısı $t=4$, blok sayısı $b=15$ olduğunda 0,5098 ile 0,9956 arasında değer almıştır.

Çizelge 7.5. $t=5$ için testlerin güç değeri

b	p					L	UP	H	KW	F
	0,10	0,15	0,20	0,25	0,30	0,5110	0,5072	0,5648	0,2102	0,1914
	0,10	0,20	0,30	0,40	0,50	0,9446	0,9458	0,9614	0,6886	0,6164
	0,40	0,45	0,50	0,55	0,60	0,3880	0,3822	0,4190	0,1428	0,1306
$b=5$	0,30	0,40	0,50	0,60	0,70	0,8880	0,8844	0,9230	0,5536	0,5052
	0,70	0,75	0,80	0,85	0,90	0,5080	0,4966	0,5580	0,2164	0,1826
	0,50	0,60	0,70	0,80	0,90	0,9412	0,9416	0,9652	0,6708	0,6098
	0,10	0,15	0,20	0,25	0,30	0,7808	0,7974	0,8350	0,5344	0,4964
	0,10	0,20	0,30	0,40	0,50	0,9982	0,9988	0,9990	0,9876	0,9708
	0,40	0,45	0,50	0,55	0,60	0,1460	0,6192	0,6730	0,3384	0,3014
$b=10$	0,30	0,40	0,50	0,60	0,70	0,9928	0,9938	0,9960	0,9414	0,9112
	0,70	0,75	0,80	0,85	0,90	0,7898	0,7928	0,8506	0,5324	0,4984
	0,50	0,60	0,70	0,80	0,90	0,9996	0,9986	0,9998	0,9830	0,9734
	0,10	0,15	0,20	0,25	0,30	0,9236	0,9148	0,9500	0,7486	0,7162
	0,10	0,20	0,30	0,40	0,50	1	1	1	0,9996	0,9994
	0,40	0,45	0,50	0,55	0,60	0,7834	0,7736	0,8384	0,5318	0,4732
$b=15$	0,30	0,40	0,50	0,60	0,70	1	0,9996	0,9998	0,9956	0,9918
	0,70	0,75	0,80	0,85	0,90	0,9246	0,9208	0,9494	0,7608	0,7210
	0,50	0,60	0,70	0,80	0,90	1	1	1	0,9992	0,9992

Çizelge 7.5 incelendiğinde Friedman testinin diğer testlere göre daha düşük güç değerlerine sahip olduğu görülmektedir. Blok sayısı arttıkça Friedman testinin gücü artmaktadır. Örneğin, $p=0,10, 0,15, 0,20, 0,25, 0,30$ olduğunda Friedman testinin güç değeri, blok sayısı $b=5$ iken 0,1914; $b=10$ iken 0,4964; $b=15$ iken 0,7162'dir.

Hollander testinin güç değerleri diğer testlere göre daha yüksek olmakla birlikte Page ve Uyarlanmış Page testlerinin güç değerlerine oldukça yakın olduğu görülmektedir. Blok sayısı arttıkça testlerin güç değerleri artmıştır. Örneğin, blok sayısı $b=10, p=0,10, 0,20, 0,30, 0,40, 0,50$ olduğunda Page testinin güç değeri 0,9982, Uyarlanmış Page testinin güç değeri 0,9988 ve Hollander testinin güç değeri 0,9990'dır.

Simülasyon çalışmasının üçüncü aşamasında rasgele tamamlanmış blok tasarımında kullanılan Friedman test istatistiği ile tamamen rasgele tasarımda kullanılan Kruskal Wallis test istatistiği ve bu tasarımların kombinasyonundan oluşan karma tasarımlarda kullanılması

için önerilen, T_1 test istatistiği güç değerleri bakımından karşılaştırılmıştır. Elde edilen güç değerleri Çizelge 7.6- Çizelge 7.14'te verilmiştir. T_1 test istatistiği ($t-1$) serbestlik dereceli Kruskal Wallis ve ($t-1$) serbestlik dereceli Friedman test istatistiklerinin toplamı olarak oluşturulmuş olup 2 ($t-1$) serbestlik dereceli ki-kare dağılıma sahiptir. Simülasyonda rasgele tamamlanmış blok tasarımı kısmı için (Friedman testi) blok sayısı $b=10, 15, 20$; tamamen rasgele tasarım kısmı için (Kruskal Wallis testi) gözlem sayıları $n=6, 8, 10$; İşlem sayısı ise $t=3, 4, 5$ olarak alınmıştır. Veriler, farklı p parametrelili binom dağılımından üretilmiş olup “ n ” parametresi 5 olarak alınmıştır.

Simülasyon çalışmasında işlem sayısı $t=3$ ve rasgele tamamlanmış blok tasarımı için blok sayıları $b=10,15$ ve 15 olduğunda test istatistiklerinin elde edilen güç değerlerine Çizelge 7.6, Çizelge 7.7 ve Çizelge 7.8’de yer verilmiştir.

Çizelge 7.6. $t=3, n=6$ için testlerin gücü

b	p			F	KW	T_1
$b=10$	0,10	0,15	0,20	0,1706	0,1126	0,1840
	0,10	0,20	0,30	0,4888	0,2906	0,5782
	0,45	0,50	0,55	0,1104	0,0764	0,1094
	0,40	0,50	0,60	0,3152	0,1930	0,3662
	0,70	0,75	0,80	0,1300	0,0888	0,1334
	0,70	0,80	0,90	0,5006	0,2976	0,5944
$b=15$	0,10	0,15	0,20	0,2540	0,1126	0,2484
	0,10	0,20	0,30	0,7192	0,2890	0,7518
	0,45	0,50	0,55	0,1456	0,0734	0,1294
	0,40	0,50	0,60	0,4672	0,1876	0,4944
	0,70	0,75	0,80	0,1790	0,0890	0,1744
	0,70	0,80	0,90	0,7072	0,2970	0,7492
$b=20$	0,10	0,15	0,20	0,3438	0,1136	0,3276
	0,10	0,20	0,30	0,8432	0,2918	0,8566
	0,45	0,50	0,55	0,1776	0,0720	0,1588
	0,40	0,50	0,60	0,6052	0,1882	0,6142
	0,70	0,75	0,80	0,2370	0,0832	0,2112
	0,70	0,80	0,90	0,8416	0,2984	0,8514

Çizelge 7.6 incelendiğinde işlem sayısı 3 olduğunda en yüksek güç değerine sahip olan test T_1 testidir. Blok sayısı arttıkça rasgele tamamlanmış blok tasarımı kısmı için kullanılan Friedman testi ile T_1 testinin güç değerleri artmıştır. Kruskal Wallis testi en düşük güç değerlerine sahip olan testtir. Çizelge incelendiğinde p parametrelerinin arasındaki fark 0,1 iken testlerin güç değerlerinin daha yüksek olduğu görülmektedir.

Çizelge 7.7. $t=3$, $n=8$ için testlerin gücü

b	p			F	KW	T ₁
b=10	0,10	0,15	0,20	0,1844	0,1430	0,2154
	0,10	0,20	0,30	0,5038	0,4272	0,6786
	0,45	0,50	0,55	0,1224	0,0898	0,1214
	0,40	0,50	0,60	0,3266	0,2658	0,4282
	0,70	0,75	0,80	0,1314	0,1090	0,1542
	0,70	0,80	0,90	0,5052	0,4180	0,6678
b=15	0,10	0,15	0,20	0,2550	0,1490	0,2852
	0,10	0,20	0,30	0,7002	0,4266	0,7926
	0,45	0,50	0,55	0,1442	0,0962	0,1496
	0,40	0,50	0,60	0,4832	0,2690	0,5550
	0,70	0,75	0,80	0,1840	0,1100	0,1912
	0,70	0,80	0,90	0,7048	0,4216	0,8066
b=20	0,10	0,15	0,20	0,3416	0,1506	0,3526
	0,10	0,20	0,30	0,8404	0,4125	0,8878
	0,45	0,50	0,55	0,1880	0,0958	0,1850
	0,40	0,50	0,60	0,6194	0,2708	0,6658
	0,70	0,75	0,80	0,2468	0,1128	0,2466
	0,70	0,80	0,90	0,8488	0,4208	0,8934

Çizelge 7.7 incelendiğinde işlem sayısı $t=3$ iken tamamen rasgele tasarım kısmı için gözlem sayısı $n=8$ olduğunda Kruskal Wallis testinin güç değerlerinin arttığı görülmektedir. Ancak Friedman testinden daha düşük güç değerlerine sahiptir.

Çizelge 7.8. $t=3$, $n=10$ için testlerin gücü

b	p			F	KW	T ₁
b=10	0,10	0,15	0,20	0,1734	0,1712	0,2322
	0,10	0,20	0,30	0,4832	0,5232	0,7202
	0,45	0,50	0,55	0,1030	0,1132	0,1314
	0,40	0,50	0,60	0,3220	0,3572	0,4910
	0,70	0,75	0,80	0,1238	0,1406	0,1696
	0,70	0,80	0,90	0,4948	0,5384	0,7480
b=15	0,10	0,15	0,20	0,2576	0,1902	0,3126
	0,10	0,20	0,30	0,7076	0,5332	0,8446
	0,45	0,50	0,55	0,1440	0,1104	0,1664
	0,40	0,50	0,60	0,4788	0,3608	0,6050
	0,70	0,75	0,80	0,1842	0,1388	0,2090
	0,70	0,80	0,90	0,7088	0,5268	0,8416
b=20	0,10	0,15	0,20	0,3426	0,1862	0,3836
	0,10	0,20	0,30	0,8402	0,5386	0,9152
	0,45	0,50	0,55	0,1832	0,1118	0,1864
	0,40	0,50	0,60	0,6242	0,3470	0,7120
	0,70	0,75	0,80	0,2368	0,1252	0,2516
	0,70	0,80	0,90	0,8380	0,5164	0,9118

Çizelge 7.8 incelendiğinde işlem sayısı $t=3$ iken tamamen rasgele tasarım kısmı için gözlem sayısı $n=10$ olduğunda Kruskal Wallis testinin güç değerlerinin arttığı görülmektedir. Rasgele tamamlanmış blok tasarımı kısmının blok sayısı $b=10$ olduğunda Kruskal Wallis

testinin Friedman testinden daha yüksek güç değerlerine sahip olduğu görülmektedir. Kruskal Wallis testinin güç değeri arttıkça T_1 testinin güç değeri de artmıştır.

İşlem sayısı $t=4$ ve rasgele tamamlanmış blok tasarımı için blok sayıları $b=10, 15$ ve 20 olduğunda test istatistiklerinin elde edilen güç değerlerine Çizelge 7.9, Çizelge 7.10 ve Çizelge 7.11’de yer verilmiştir

Çizelge 7.9. $t=4$, $n=6$ için testlerin gücü

b	p				F	KW	T_1
b=10	0,10	0,15	0,20	0,25	0,2942	0,1628	0,3398
	0,10	0,20	0,30	0,40	0,8124	0,5242	0,9072
	0,40	0,45	0,50	0,55	0,1722	0,0944	0,1838
	0,40	0,50	0,60	0,70	0,6386	0,3858	0,7550
	0,60	0,70	0,80	0,90	0,8012	0,5268	0,9074
	0,75	0,80	0,85	0,90	0,3028	0,1718	0,3522
b=15	0,10	0,15	0,20	0,25	0,4802	0,1756	0,4972
	0,10	0,20	0,30	0,40	0,9560	0,5228	0,9468
	0,40	0,45	0,50	0,55	0,2740	0,1046	0,2676
	0,40	0,50	0,60	0,70	0,8508	0,3818	0,8946
	0,60	0,70	0,80	0,90	0,9576	0,5200	0,9782
	0,75	0,80	0,85	0,90	0,4732	0,1758	0,4982
b=20	0,10	0,15	0,20	0,25	0,6154	0,1742	0,6090
	0,10	0,20	0,30	0,40	0,9914	0,5260	0,9968
	0,40	0,45	0,50	0,55	0,3674	0,1066	0,3498
	0,40	0,50	0,60	0,70	0,9522	0,3966	0,9626
	0,60	0,70	0,80	0,90	0,9924	0,5246	0,9952
	0,75	0,80	0,85	0,90	0,6216	0,1686	0,6122

Çizelge 7.9 incelendiğinde işlem sayısı $t=4$ olduğunda tüm testlerin güç değerlerinin artış gösterdiği görülmektedir. Özellikle p değerleri arasındaki farkın 0,1 olduğu durumlarda elde edilen güç değerlerindeki artış daha fazla olmuştur. Örneğin, işlem sayısı $t=3$ blok sayısı $b=10$, $n=6$ ve $p= 0,10, 0,15, 0,20$ olduğunda Friedman testinin güç değeri 0,1706, Kruskal Wallis testinin 0,1126, T_1 testinin 0,1840 iken işlem sayısı $t=4$, $p= 0,10, 0,15, 0,20, 0,25$ olduğunda güç değerleri Friedman testi için 0,2942, Kruskal Wallis testi için 0,1628, T_1 testi için 0,1840 olarak elde edilmiştir. İşlem sayısı $t=3$ blok sayısı $b=10$, $n=6$ ve $p= 0,10, 0,20, 0,30$ olduğunda Friedman testinin güç değeri 0,4888, Kruskal Wallis testinin 0,2906, T_1 testinin 0,5782 iken işlem sayısı $t=4$, $p= 0,10, 0,20, 0,30, 0,40$ olduğunda güç değerleri Friedman testi için 0,8124, Kruskal Wallis testi için 0,5242, T_1 testi için 0,9072 olarak elde edilmiştir. Görüldüğü gibi p parametreleri arasındaki farkın yüksek olduğu durumlarda işlem sayısı arttıkça testlerin gücü daha fazla artış göstermiştir.

Çizelge 7.10. $t=4$, $n=8$ için testlerin gücü

b	p				F	KW	T_1
b=10	0,10	0,15	0,20	0,25	0,2992	0,2440	0,4064
	0,10	0,20	0,30	0,40	0,8036	0,7084	0,9426
	0,40	0,45	0,50	0,55	0,1760	0,1446	0,2264
	0,40	0,50	0,60	0,70	0,6324	0,5644	0,8352
	0,60	0,70	0,80	0,90	0,8012	0,7118	0,9450
	0,75	0,80	0,85	0,90	0,2918	0,2496	0,4060
b=15	0,10	0,15	0,20	0,25	0,4630	0,2468	0,5428
	0,10	0,20	0,30	0,40	0,9568	0,7080	0,9914
	0,40	0,45	0,50	0,55	0,2806	0,1516	0,3104
	0,40	0,50	0,60	0,70	0,8622	0,5524	0,9328
	0,60	0,70	0,80	0,90	0,9578	0,7038	0,9878
	0,75	0,80	0,85	0,90	0,4730	0,2448	0,5498
b=20	0,10	0,15	0,20	0,25	0,6246	0,2514	0,6660
	0,10	0,20	0,30	0,40	0,9912	0,7248	0,9966
	0,40	0,45	0,50	0,55	0,3606	0,1478	0,3736
	0,40	0,50	0,60	0,70	0,9452	0,5542	0,9730
	0,60	0,70	0,80	0,90	0,9918	0,7068	0,9970
	0,75	0,80	0,85	0,90	0,6184	0,2552	0,6656

Çizelge 7.10 incelendiğinde en yüksek güç değerine sahip olan testin T_1 testi olduğu görülmektedir. Blok sayısı 20 $p=0,10$, $0,20$, $0,30$, $0,40$ olduğunda T_1 testinin $0,9966$ ile oldukça yüksek güç değerine sahip olduğu görülmektedir.

Çizelge 7.11. $t=4$, $n=10$ için testlerin gücü

b	p				F	KW	T_1
b=10	0,10	0,15	0,20	0,25	0,3042	0,3270	0,4800
	0,10	0,20	0,30	0,40	0,8078	0,8354	0,9728
	0,40	0,45	0,50	0,55	0,1818	0,1972	0,2634
	0,40	0,50	0,60	0,70	0,6378	0,6904	0,8866
	0,60	0,70	0,80	0,90	0,8056	0,8372	0,9738
	0,75	0,80	0,85	0,90	0,3084	0,3142	0,4646
b=15	0,10	0,15	0,20	0,25	0,4858	0,3268	0,6026
	0,10	0,20	0,30	0,40	0,9556	0,8396	0,9938
	0,40	0,45	0,50	0,55	0,2732	0,1950	0,3420
	0,40	0,50	0,60	0,70	0,8532	0,6928	0,9590
	0,60	0,70	0,80	0,90	0,9560	0,8316	0,9918
	0,75	0,80	0,85	0,90	0,4630	0,3218	0,5842
b=20	0,10	0,15	0,20	0,25	0,6232	0,3164	0,7012
	0,10	0,20	0,30	0,40	0,9884	0,8288	0,9976
	0,40	0,45	0,50	0,55	0,3630	0,1902	0,4110
	0,40	0,50	0,60	0,70	0,9486	0,6844	0,9806
	0,60	0,70	0,80	0,90	0,9926	0,8404	0,9990
	0,75	0,80	0,85	0,90	0,6216	0,3202	0,6994

Çizelge 7.11 incelendiğinde tamamen rasgele tasarım kısmının gözlem sayısının rasgele tamamlanmış blok tasarımı kısmının blok sayısına eşit olduğu durumlarda Kruskal Wallis testinin Friedman testinden daha yüksek güç değerlerine sahip olduğu görülmektedir.

Örneğin rasgele tamamlanmış blok tasarımı kısmının blok sayısı $b=10$, $p=0,60, 0,70, 0,80, 0,90$ $n=10$ iken Friedman testinin güç değeri 0,8056, Kruskal Wallis testinin güç değeri 0,8372 olarak elde edilmiştir. T_1 testi yine en yüksek güç değerine sahip olan testtir.

Tamamen rasgele tasarım kısmının gözlem sayısı arttıkça Kruskal Wallis testinin güç değerinin arttığı görülmektedir. Örneğin, $p=0,10, 0,20, 0,30, 0,40$ olduğunda Kruskal Wallis testinin güç değeri $n=6$ iken 0,5242, $n=8$ iken 0,7084, $n=10$ iken 0,8354'tür.

Çizelge 7.12. $t=5$, $n=6$ için testlerin gücü

b	p					F	KW	T_1
b=10	0,10	0,15	0,20	0,25	0,30	0,4830	0,2622	0,5852
	0,10	0,20	0,30	0,40	0,50	0,9652	0,7832	0,9942
	0,40	0,45	0,50	0,55	0,60	0,3030	0,1612	0,3572
	0,30	0,40	0,50	0,60	0,70	0,9214	0,6724	0,9774
	0,70	0,75	0,80	0,85	0,90	0,4966	0,2626	0,5832
	0,50	0,60	0,70	0,80	0,90	0,9684	0,7740	0,9952
b=15	0,10	0,15	0,20	0,25	0,30	0,7194	0,2702	0,7578
	0,10	0,20	0,30	0,40	0,50	0,9986	0,7694	0,9998
	0,40	0,45	0,50	0,55	0,60	0,4874	0,1634	0,5094
	0,30	0,40	0,50	0,60	0,70	0,9914	0,6644	0,9974
	0,70	0,75	0,80	0,85	0,90	0,7242	0,2608	0,7666
	0,50	0,60	0,70	0,80	0,90	0,9982	0,7758	0,9998
b=20	0,10	0,15	0,20	0,25	0,30	0,8618	0,2790	0,8728
	0,10	0,20	0,30	0,40	0,50	1	0,7922	1
	0,40	0,45	0,50	0,55	0,60	0,6298	0,1694	0,6306
	0,30	0,40	0,50	0,60	0,70	0,9996	0,6696	1,0000
	0,70	0,75	0,80	0,85	0,90	0,8612	0,2626	0,8660
	0,50	0,60	0,70	0,80	0,90	1,0000	0,7724	0,9998

Çizelge 7.13. $t=5$, $n=8$ için testlerin gücü

b	p					F	KW	T_1
b=10	0,10	0,15	0,20	0,25	0,30	0,4934	0,3998	0,6640
	0,10	0,20	0,30	0,40	0,50	0,9692	0,9208	0,9970
	0,40	0,45	0,50	0,55	0,60	0,3030	0,2430	0,4124
	0,30	0,40	0,50	0,60	0,70	0,9190	0,8522	0,9922
	0,70	0,75	0,80	0,85	0,90	0,4836	0,3954	0,6604
	0,50	0,60	0,70	0,80	0,90	0,9694	0,9232	0,9992
b=15	0,10	0,15	0,20	0,25	0,30	0,7136	0,3928	0,8130
	0,10	0,20	0,30	0,40	0,50	0,9998	0,9238	1
	0,40	0,45	0,50	0,55	0,60	0,4854	0,2542	0,5624
	0,30	0,40	0,50	0,60	0,70	0,9916	0,8520	0,9990
	0,70	0,75	0,80	0,85	0,90	0,7230	0,3942	0,8108
	0,50	0,60	0,70	0,80	0,90	0,9988	0,9160	1
b=20	0,10	0,15	0,20	0,25	0,30	0,8610	0,3896	0,8988
	0,10	0,20	0,30	0,40	0,50	1	0,9114	1
	0,40	0,45	0,50	0,55	0,60	0,6388	0,2396	0,6780
	0,30	0,40	0,50	0,60	0,70	0,9994	0,8450	1
	0,70	0,75	0,80	0,85	0,90	0,8684	0,4048	0,9072
	0,50	0,60	0,70	0,80	0,90	1	0,9232	1

Çizelge 7.14. $t=5$, $n=10$ için testlerin gücü

b	p					F	KW	T_1
b=10	0,10	0,15	0,20	0,25	0,30	0,4876	0,5068	0,7374
	0,10	0,20	0,30	0,40	0,50	0,9686	0,9738	0,9992
	0,40	0,45	0,50	0,55	0,60	0,3032	0,3238	0,4770
	0,30	0,40	0,50	0,60	0,70	0,9132	0,9334	0,9968
	0,70	0,75	0,80	0,85	0,90	0,5054	0,5178	0,7420
	0,50	0,60	0,70	0,80	0,90	0,9710	0,9766	0,9996
b=15	0,10	0,15	0,20	0,25	0,30	0,7172	0,5272	0,8578
	0,10	0,20	0,30	0,40	0,50	0,9986	0,9736	1
	0,40	0,45	0,50	0,55	0,60	0,4856	0,3256	0,6184
	0,30	0,40	0,50	0,60	0,70	0,9918	0,9326	0,9998
	0,70	0,75	0,80	0,85	0,90	0,7240	0,5220	0,8574
	0,50	0,60	0,70	0,80	0,90	0,9986	0,9794	1
b=20	0,10	0,15	0,20	0,25	0,30	0,8654	0,5228	0,9248
	0,10	0,20	0,30	0,40	0,50	1	0,9748	1
	0,40	0,45	0,50	0,55	0,60	0,6282	0,3240	0,7172
	0,30	0,40	0,50	0,60	0,70	0,9994	0,9330	1
	0,70	0,75	0,80	0,85	0,90	0,8626	0,5172	0,9292
	0,50	0,60	0,70	0,80	0,90	1	0,9764	1

Çizelge 7.12, Çizelge 7.13 ve Çizelge 7.14 incelendiğinde işlem sayısı $t=5$ olduğunda testlerin güç değerlerinin arttığı görülmektedir. Aynı şekilde blok ve gözlem sayıları arttıkça testlerin güç değerleri artış göstermiştir. p değerleri arasındaki artış 0,1 olduğunda testlerin daha yüksek güç değeri elde ettiği görülmektedir.

Tamamen rasgele tasarım kısmının gözlem sayısı artırıldığında Kruskal Wallis testiyle birlikte T_1 testinin güç değeri de artmıştır. Örneğin, blok sayısı $b=10$, $p= 0,10, 0,20, 0,30, 0,40, 0,50$ olduğunda T_1 testinin güç değeri, $n=6$ iken 0,5852, $n=8$ iken 0,6640, $n=10$ iken 0,7374'tür.

Simülasyon çalışmasının sonuçlarına göre genel olarak T_1 testinin Kruskal Wallis ve Friedman testlerinden daha yüksek güç değerleri elde ettiği görülmektedir. Bazı durumlarda örneğin işlem sayısı $t= 4$ olduğu durumlarda Friedman testinin güç değeri T_1 testinden daha büyük olmakla birlikte değerlerin birbirine oldukça yakın olduğu görülmektedir.

8. SONUÇ VE ÖNERİLER

Araştırmacıların deneye başlamadan önce kullanacakları deney tasarımı modeline karar vermeleri gerekir. Başlangıçta belirlenen deney tasarımıyla devam etmek her zaman mümkün olmayabilir. Bu gibi durumlarda araştırmacılar mevcut verilerden karar verdikleri tasarıma uygun olanları alarak diğer uygun olmayan verileri gözardı etmek ya da tasarım yapısını değiştirmek zorunda kalabilirler. Ne kadar çok veri olursa testin o kadar güçlü olacağı düşünülürse, araştırmacılar tarafından karma bir tasarım oluşturarak elde edilen tüm verilerin kullanılması testin daha güçlü olmasını sağlayabilir. Karma deney tasarımları farklı deneysel tasarımların birleşiminden oluşabilmektedir.

Bu çalışmada farklı deney tasarımlarında kullanılan parametrik olmayan yöntemler incelenmiştir. Tepe noktası başlangıç veya sonuncu grup olduğunda sıralı alternatifin şemsiye alternatifin özel bir hali olduğu, literatürde araştırmacıların şemsiye alternatifler için kullandıkları test istatistiklerini, genellikle sıralı alternatiflerin uzantısı olacak şekilde tasarladıklarından bahsedilmiştir.

Simülasyon çalışması yapılarak testlerin deneysel I. tip hata oranları ile güç değerleri karşılaştırılmıştır. Verilerin üretildiği binom dağılımının p parametresi değiştiğinde testlerin nasıl tepki vereceği araştırılmıştır. Binom dağılımının tercih edilmesinde dağılım yapısının esnek olması ve farklı p değerleri sayesinde farklı yapıları esnek bir şekilde modelleyebilmesi etkili olmuştur.

İlk simülasyon çalışmasından elde edilen verilere göre p parametresi 0,1 olarak alındığında nominal α 'ya (0,05) en yakın değere sahip olan testler Kruskal Wallis ve Friedman testleridir. Hollander testi, Page ve Uyarlanmış Page testlerine göre daha yüksek deneysel I. tip hata oranına sahiptir. Verilerin üretildiği dağılımın p parametresi değiştirildiğinde, Hollander testi, Page ve Uyarlanmış Page testlerine göre parametre değişiminden daha az etkilenmiştir.

Verilerin üretildiği dağılımın p olasılık parametresi sırasıyla 0,1, 0,2, 0,5, 0,8 ve 0,9 olarak alındığında Page ve Uyarlanmış Page testinin deneysel I. tip hata oranlarının önce arttığı daha sonra tekrar azaldığı ve $p=0,1$ ile $p=0,9$ olduğunda birbirlerine yakın değerlere sahip olduğu gözlemlenmiştir. Diğer testler parametre değişiminden nispeten daha az

etkilenmiştir. Bazı testlerin, simülasyon çalışmalarında veri üretilirken kullanılan dağılımın parametre değişikliğine karşı daha hassas olduğu, bu nedenle uygun dağılım ve parametre seçiminin sonuçların güvenilir olması açısından önemli olduğu görülmektedir.

İkinci olarak yapılan simülasyon çalışmasının sonuçları incelendiğinde işlem ve blok sayısı arttıkça testlerin güç değerlerinin arttığı görülmektedir. Tüm işlem sayılarında en yüksek güç değerine sahip olan test Hollander testidir. Page ve Uyarlanmış Page testi, Hollander testine oldukça yakın güç değerlerine sahiptir. Friedman testi ile Kruskal Wallis testi diğer testlere göre daha düşük güç değerlerine sahiptir.

Simülasyon çalışmasında işlem ve blok sayısı arttıkça testlerin güç değerlerinin arttığı, p değerleri arasındaki farkın testin gücünü etkilediği görülmektedir. Örneğin, Hollander testi incelendiğinde, işlem sayısı $t=5$, blok sayısı $b=5$ olduğunda $p=0,40, 0,45, 0,50, 0,55, 0,60$ iken (p değerleri arasındaki fark 0,05) güç değeri 0,4190, $p=0,30, 0,40, 0,50, 0,60, 0,70$ iken (p değerleri arasındaki fark 0,1) 0,9230'dur. p değerleri arasındaki farkın fazla olması gruplar arasındaki farkın belirgin olduğunu göstermektedir ve gruplar arasındaki farkın artması sıfır hipotezinin doğru bir şekilde reddedilme olasılığını artırmaktadır. Bu nedenle farkın fazla olduğu durumlarda güç değerlerinin daha yüksek olduğu görülmektedir.

Üçüncü simülasyon çalışmasında rasgele tamamlanmış blok tasarımında kullanılan Friedman testi ile tamamen rasgele tasarımda kullanılan Kruskal Wallis testinin ve bu testlerin kombinasyonundan oluşan T_1 testinin güç değerleri incelenmiştir. Kombinasyon testi oluşturulurken ikinci simülasyon çalışmasında yer alan testler arasında daha düşük güç değerlerine sahip olan Kruskal Wallis ve Friedman testlerinin tercih edilmesinde bu testlerin farklı deney tasarımları için kullanılmaları, aynı dağılıma sahip olmaları ve toplanabilir özellikte olmaları etkili olmuştur. Verilerin tamamının başlangıç tasarımına uygun olmaması durumunda, sadece uygun verilerle çalışıp geri kalan veriyi göz ardı etmek yerine ek bir deney tasarımıyla birlikte oluşturulan karma tasarımın testin gücünü artırdığı görülmektedir. İncelenen T_1 testinin her iki testin güçlü yönlerini almasından dolayı bu testlerden daha yüksek güç değerlerine sahip olduğu görülmektedir. Bu nedenle tamamen rasgele tasarım ve rasgele tamamlanmış blok tasarımından oluşan karma bir tasarım için uygun parametrelerin belirlenmesi durumunda T_1 test istatistiğinin kullanılması tavsiye edilebilir.

KAYNAKLAR

- Alsuhabi, H. R. (2019). *Proposed nonparametric tests for the umbrella alternative in a mixed design for both known and unknown peak*. Doctoral Dissertation, North Dakota State University of Agriculture and Applied Science, Fargo, 45-63.
- Alvo, M., Cabilio, P. (1995). Testing ordered alternatives in the presence of incomplete data. *Journal of the American Statistical Association*, 90(3), 1015–1024.
- Asare, B. (2020). *Nonparametric tests for the umbrella alternative in a mixed design for a known peak*. Doctoral Dissertation, North Dakota State University of Agriculture and Applied Science, Fargo, 122-145.
- Bağcı, M. (2019). *Şemsiye alternatiflerine karşı konumların eşitliği için parametrik olmayan bir test*. Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 65-72.
- Best, D. J., Rayner, J. C. W., Thas, O. (2009). *Nonparametric tests for randomized block data with ties and ordered alternatives*. Third Annual ASEARC Conference, Newcastle, Australia, 122-135.
- Daniel, W. W. (1990). *Applied nonparametric statistics* (2nd ed.). Boston: PWS-Kent Publishing Company, 135-145.
- Dubnicka, S., Blair, R. C., Hettmansperger, T. P. (2002). Rank-based procedures for mixed paired and two-sample designs. *Journal of Modern Applied Statistical Methods*, 1(1), 32–41.
- Duran, B. S. (1995). *Page's statistic in balanced incomplete block designs*. Master's Thesis, Texas Tech University, Graduate Faculty, 85-93.
- Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(4), 675–701.
- Friedman, M. (1940). A comparison of alternative tests of significance for the problem of m rankings. *Annals of Mathematical Statistics*, 11(1), 86–92.
- Gökpınar, E., Gökpınar, F. (2016). A modified Mack-Wolfe test for the umbrella alternative problem. *Communications in Statistics - Simulation and Computation*, 24(5), 7226–7241.
- Hemmer, M. T. (2012). *Nonparametric test for the umbrella alternative in a randomized complete block and balanced incomplete block mixed design*. Master's Thesis, North Dakota State University of Agriculture and Applied Science, Fargo, 93-102.
- Hollander, M., Wolfe, D. A. (1999). *Nonparametric statistical methods* (2nd ed.). New York: Wiley, 93-105.
- İnternet: Asare, B. (2023). Testing the umbrella alternative nonparametrically in a three mixed design for a known peak. Web: <https://www.researchgate.net/publication/375115910>, Son Erişim Tarihi: 22/10/2024.

- Karagöz, Y. (2010). Nonparametrik tekniklerin güç ve etkinlikleri. *Elektronik Sosyal Bilimler Dergisi*, 9(33), 18–40.
- Kim, D. H., Kim, Y. C. (1992). Distribution-free tests for umbrella alternatives in a randomized block design. *Journal of Nonparametric Statistics*, 1(4), 277–285.
- Mack, G. A., Wolfe, D. A. (1981). K-sample rank tests for umbrella alternatives. *Journal of the American Statistical Association*, 76(3), 175–181.
- Magel, R., Fu, R. (2014). A proposed test for a mixed two-sample design. *Journal of Statistical Theory and Practice*, 8(3), 221–237.
- Magel, R., Mathisen, D. (2011). *Test for the nondecreasing alternative in a complete block and incomplete block mixed design*. Master's Thesis, North Dakota State University, Fargo, 36-48.
- Magel, R., Terpstra, J., Canonizado, K. and Park, J. I. (2010). Nonparametric tests for mixed designs. *Communications in Statistics - Simulation and Computation*, 39(6), 1228–1250.
- Magel, R., Terpstra, J., Wen, J. (2009). Proposed tests for the nondecreasing alternative in a mixed design. *Journal of Statistics and Management Systems*, 12(4), 963–977.
- Montgomery, D. C. (1997). *Design and analysis of experiments*. New York: John Wiley and Sons, 102-105.
- Ndungu, A. (2011). *A nonparametric test for the non-decreasing alternative in an incomplete block design*. Master's Thesis, North Dakota State University, Statistics Department, Fargo, 93-105.
- Ndungu, A. M. (2014). *Nonparametric tests for the non-decreasing and umbrella alternatives in the incomplete block and completely randomized and mixed design*. Doctoral Dissertation, North Dakota State University, Statistics Department, Fargo, 85-93.
- Oehlert, G. W. (2000). *A first course in design and analysis of experiments*. New York: W. H. Freeman, 102-105.
- Osafo, M. (2020). *Nonparametric test for nondecreasing order alternatives in randomized complete block and balanced incomplete block mixed design*. Doctoral Dissertation, North Dakota State University, Statistics Department, Fargo, 65-72.
- Page, E. B. (1963). Ordered hypotheses for multiple treatments: A significance test for linear ranks. *Journal of the American Statistical Association*, 58(2), 216–230.

EKLER

EK-1. R Program Kodu

35 öğrencinin sınavlarını değerlendiren 4 sınıf öğretmeni (değerlendirici) arasında fark olup olmadığına ilişkin yapılan çalışmanın R 4.3.2 program kodu aşağıda verilmiştir.

```
# Gerekli paketleri yükle
```

```
library(readxl)
```

```
# Veriyi yükle
```

```
data <- read_excel("C:/Users/birca/Desktop/notlar.xlsx")
```

```
# İlk 15 öğrenci için veriyi ayır
```

```
ilk_15 <- data[1:15, 2:5] # 2. sütundan 5. sütuna kadar değerlendiricilerin puanları
```

```
# Sonraki 20 öğrenci için veriyi ayır
```

```
son_20 <- data[16:35, 2:5] # 16. satırdan 35. satıra kadar
```

```
# ----- Friedman Testi -----
```

```
# Friedman Testi (ilk 15 öğrenci için)
```

```
friedman_sonuc <- friedman.test(as.matrix(ilk_15))
```

```
# ----- Kruskal-Wallis Testi -----
```

```
# Veriyi uzun formatta birleştirme (son 20 öğrenci için)
```

```
son_20_gerekli <- as.vector(as.matrix(son_20))
```

```
# Blok düzenini oluşturma
```

```
blok_index <- c(rep(1, 5), # 1. işlem için 16-20. bloklar
```

```
  rep(2, 5), # 2. işlem için 21-25. bloklar
```

```
  rep(3, 5), # 3. işlem için 26-30. bloklar
```

```
  rep(4, 5)) # 4. işlem için 31-35. bloklar
```

```
# Verinin uzunluğunu eşitleme
```

```
blok_index <- rep(blok_index, length.out = length(son_20_gerekli))
```

```
# Kruskal-Wallis testini uygulama
```

```
kruskal_sonuc <- kruskal.test(son_20_gerekli, factor(blok_index))
```

EK- 1. (devam) R Program Kodu

```
#----- Kombinasyon Testi ve Sonuçlar -----
# Test istatistiklerinin toplamını hesapla
toplam_test_istatistigi <- friedman_sonuc$statistic + kruskal_sonuc$statistic

# Kombinasyon p değerini ki-kare dağılımından hesapla
serbestlik_dereceleri <- 6 # Kombinasyon testi için serbestlik derecesi
kombinasyon_p_degeri <- pchisq(toplam_test_istatistigi, df = serbestlik_dereceleri,
lower.tail = FALSE)

# Sonuçları yazdır
cat("Friedman Testi P Değeri:", friedman_sonuc$p.value, "\n")
cat("Friedman Test İstatistiği:", friedman_sonuc$statistic, "\n")
cat("Kruskal-Wallis Testi P Değeri:", kruskal_sonuc$p.value, "\n")
cat("Kruskal-Wallis Test İstatistiği:", kruskal_sonuc$statistic, "\n")
cat("Kombinasyon Test İstatistiği:", toplam_test_istatistigi, "\n")
cat("Kombinasyon Testi P Değeri:", kombinasyon_p_degeri, "\n")

# Sonuç: Alfa ile karşılaştır
alfa <- 0.05
sonuc_kombinasyon <- ifelse(kombinasyon_p_degeri < alfa, "Red", "Red edilemez")

# Sonuçları yazdır
cat("Kombinasyon Test Sonucu:", sonuc_kombinasyon, "\n")
```



Gazili olmak ayrıcalıktır...