



T.C.

SELÇUK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

PARAMETRİK VE PARAMETRİK OLMAYAN TESTLER İÇİN GÜÇ KARŞILAŞTIRMALARI

Nazik Özge CEYLAN

YÜKSEK LİSANS

İstatistik Anabilim Dalı

Ağustos-2019
KONYA
Her Hakkı Saklıdır

ÖZET

YÜKSEK LİSANS

PARAMETRİK VE PARAMETRİK OLMAYAN TESTLER İÇİN GÜÇ KARŞILAŞTIRMALARI

Nazik Özge CEYLAN

Selçuk Üniversitesi Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Demet SEZER

2019, 137 Sayfa

Jüri
Dr. Öğr. Üyesi Demet SEZER
Dr. Öğr. Üyesi Aydın KARAKOCA
Doç. Dr. İsmail KINACI

Hipotez testleri, sağlık bilimleri, sosyal bilimler, mühendislik gibi bir çok alanda sıkça kullanılmaktadır. Bu çalışmada, kitle ortalaması, kitle medyanı ve kitle oranı hakkında önerilen hipotez testlerinden tek örneklem için Z testi, t testi, işaret testi ve Wilcoxon işaretli sıra sayıları testi; iki örneklem için Z testi, t testi, Mann Whitney U testi, Kolmogorov Smirnov testi, Wald-Wolfowitz diziler testi, Moods-Medyan testi, ki kare testi, Wilcoxon t testi, işaret testi ve Mc-Nemar testi; k örneklem için ise Tek yönlü varyans analizi, Kruskal-Wallis H testi, Medyan testi, İki yönlü varyans analizi ve Friedman'ın S testi gibi parametrik ve parametrik olmayan testlerin simülasyon çalışması ile güçleri karşılaştırılmıştır. Bu karşılaştırmalar, farklı kitle parametreleri, farklı kitle dağılımları ve farklı örneklem büyüklükleri dikkate alınarak gerçekleştirilmiştir ve elde edilen sonuçlar tablolar ve grafikler halinde verilerek yorumlanmıştır.

Anahtar Kelimeler: Hipotez testleri, iki örneklem testleri, k örneklem testleri, parametrik olmayan testler, parametrik testler, tek örneklem testleri, testin gücü.

ABSTRACT

MS THESIS

POWER COMPARISONS OF PARAMETRIC AND NONPARAMETRIC TESTS

Nazik Özge CEYLAN

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCE OF SELÇUK UNIVERSITY
THE DEGREE OF MASTER OF SCIENCE IN STATISTICS

Advisor: Asst.Prof.Dr. Demet SEZER

2019, 137 Pages

Jury
Asst.Prof.Dr. Demet SEZER
Asst.Prof.Dr. Aydın KARAKOCA
Assoc.Prof.Dr. İsmail KINACI

Hypothesis testing is widely used in health sciences, social sciences and engineering. In this study, parametric and non-parametric hypothesis tests suggested about the mean, median and ratio such as Z test, t test, Wilcoxon signed rank test and sign test for one sample; Z test, t test, Mann Whitney U test, Kolmogorov Smirnov test, Wald-Wolfowitz series test, Moods-Median test, chi-square test, Wilcoxon t test, sign test and Mc-Nemar test for two samples; one-way analysis of variance, Kruskal-Wallis H test, Median test, two-way analysis of variance and Friedman's S test for k samples tests are handled and their powers are compared with the simulation study. These comparisons are made by considering different mass parameters, different mass distributions and different sample sizes and the results are interpreted by giving tables and graphs.

Keywords: Hypothesis tests, two sample tests, k sample tests, nonparametric tests, parametric tests, power of test, single sample tests.

ÖNSÖZ

Bu akademik çalışma serüvenimde derin bilgisi ve değerli tecrübesiyle bana kıymetli katkılar sunan hocam Demet SEZER'e ve İsmail KINACI'ya, manevi anlamda benden desteğini esirgemeyen kardeşlerim Müge CEYLAN ve Samet CEYLAN'a, gerek eğitim gerekse öğrenim yaşamımda beni hayata hazırlayan ve bugünlere gelmemde çok büyük bir paya sahip olan anne ve babama, eğitim hayatımın en güzel kazanımlarından biri olan tüm dostlarıma teşekkür ederim.

Nazik Özge CEYLAN
KONYA-2019

İÇİNDEKİLER

ÖZET ..... **iv**

ABSTRACT ..... **v**

ÖNSÖZ ..... **vi**

İÇİNDEKİLER ..... **vii**

SİMGELER VE KISALTMALAR **1**

1. GİRİŞ VE KAYNAK ARAŞTIRMASI **3**

2. TEMEL KAVRAMLAR **11**

2.1. Hipotez Testleri 11
2.1.1. Parametrik testler 12
2.1.2. Parametrik olmayan testler 12

3. TEK ÖRNEKLEM TESTLERİ **13**

3.1. Ortalama ve Medyana İlişkin Testler 13
3.1.1. Z testi 13
3.1.2. t testi 13
3.1.3. İşaret testi 14
3.1.4. Wilcoxon işaretli sıra sayıları testi 15
3.1.5. Güç karşılaştırması 16
3.2. Orana İlişkin Testler 36
3.2.1. Z testi 37
3.2.2. Binom testi 37
3.2.3. Güç karşılaştırması 38

4. BAĞIMSIZ İKİ ÖRNEKLEM TESTLERİ **84**

4.1. Ortalama ve Medyana İlişkin Testler 84
4.1.1. Z testi 84
4.1.2. t testi 84
4.1.3. Mann Whitney U testi 86
4.1.4. Kolmogorov Smirnov testi 87
4.1.5. Wald-Wolfowitz Diziler testi 87
4.1.6. Moods-Medyan testi 88
4.1.7. Güç karşılaştırmaları 89
4.2. Orana İlişkin Testler 95
4.2.1. Z testi 95
4.2.3. Güç karşılaştırması 96

5. BAĞIMLI İKİ ÖRNEKLEM TESTLERİ **104**

5.1. Ortalama ve Medyana İlişkin Testler 104
5.1.1. t testi (Paired t testi) 104
5.1.2. Wilcoxon t testi 104
5.1.3. İşaret testi 105
5.1.4. Güç karşılaştırmaları 105
5.2. Orana İlişkin Testler 110
5.2.1. t testi 110
5.2.2. Mc-Nemar testi 111
5.2.3. Güç karşılaştırmaları 111

6. BAĞIMSIZ K ÖRNEKLEM TESTLERİ 113

6.1. Ortalama ve Medyana İlişkin Testler 113
6.1.1. Tek yönlü varyans analizi (ANOVA) 113
6.1.2. Kruskal Wallis H testi (K-W H Test) 114
6.1.3. Medyan testi 115
6.1.4. Güç karşılaştırmaları 115

7. BAĞIMLI K ÖRNEKLEM TESTLERİ 123

7.1. Ortalama ve Medyana İlişkin Testler 123
7.1.1. İki yönlü varyans analizi 123
7.1.2. Friedman'ın S testi 124
7.1.3. Güç karşılaştırmaları 125

8. SONUÇLAR VE ÖNERİLER 132

8.1 Sonuçlar 132
8.2 Öneriler 134

KAYNAKLAR 135

ÖZGEÇMİŞ 137

SİMGELER VE KISALTMALAR

Simgeler

M	:	Medyan
$\overline{x}$	:	Örneklem ortalaması
α	:	I.*Tip Hata*
β	:	II.*Tip Hata*
$1-\beta$	:	Testin gücü
μ	:	Ortalama
π	:	Kitle oranı
σ^2	:	Varyans
δ_i	:	Rastgele değişken
$\hat{\tau}$	:	İlişki katsayısı
B	:	Beklenen değer
$F(x)$	:	Dağılım fonksiyonu
G	:	Gözlenen değer
H_0	:	Yokluk hipotezi, sıfır hipotezi
H_1	:	Alternatif hipotez
k	:	Grup sayısı
n	:	Örneklem büyüklüğü
N	:	Normal dağılım
p	:	Örneklem oranı
P	:	Kitle oranı
r	:	Sıra sayısı
S	:	Örneklem standart sapması
s_{GA}^2	:	Gruplar arası kareler ortalaması
s_{GI}^2	:	Gruplar içi kareler ortalaması
$S(x)$	:	Örneklem dağılım fonksiyonu

Kısaltmalar

BAKO	:	Birim kareler ortalaması
BT	:	Birim Toplamı
DT	:	Düzeltme terimi
GKT	:	Genel kareler toplamı
GT	:	Genel Toplam
HKO	:	Hata kareler ortalaması
HKT	:	Hata kareler toplamı
KW	:	Kruskal-Wallis test istatistiği
max	:	Maximum
sd	:	Serbestlik derecesi
t	:	t testi
TAKO	:	İşlem kareler ortalaması

TAKT	:	İşlemler arası kareler toplamı
TT	:	İşlem Toplamı
U	:	Mann-Whitney Y test istatistiği
Z	:	Z testi

1. GİRİŞ VE KAYNAK ARAŞTIRMASI

Hipotez testleri, sağlık bilimleri, sosyal bilimler, mühendislik gibi bir çok alanda sıkça kullanılmaktadır. Hipotez, bir parametre hakkındaki iddiaları test etmek üzerine kurulur. Hipotez testleri sayesinde örneklemden yararlanılarak kitleye ilişkin karar verilebilmektedir. Örneğin, "Sınıfın matematik sınavı not ortalaması 75'tir." iddiasını test ederken, temel alınan hipoteze yokluk hipotezi adı verilmektedir ve H_0 ile gösterilmektedir. Böyle bir durumu test etmek için kurulan diğer hipotez ise durumun tam aksini önerebilir. Yani "Sınıfın matematik sınavı not ortalaması 75'ten farklıdır." şeklinde olabilir ve bu hipoteze de alternatif hipotez adı verilip H_1 ile ifade edilmektedir. Hipotez testleri, belirli varsayımları gerçekleştirdiği takdirde parametrik hipotez testleri, herhangi bir varsayıma bağlı olmayanlara ise parametrik olmayan hipotez testleri denilmektedir.

Hipotez testlerinde I.Tip Hata gerçekte doğru olan yokluk hipotezinin, örneklemin gözlem değerlerine göre test sonucunda reddedilmesi olarak ifade edilmektedir. α ise I.Tip Hata yapma olasılığının üst sınırı şeklinde tanımlanmaktadır. Öte yandan, eğer yokluk hipotezi gerçekte yanlışken kabul ediliyorsa buna II.Tip Hata denir ve bu hatayı yapma olasılığı β ile gösterilir. Bir testin gücü ise $1-\beta$ olarak ifade edilir. Parametrik ya da parametrik olmayan herhangi bir uygulanacak testin anlamlı olabilmesi için ilgili gözlem sayısı, önceden verilen anlam seviyesi α ve testin gücü $(1-\beta)$'ya göre belirlenmesi yararlı olacaktır. Örneğin, gerçekte ilaçlar aynı etkiyi gösteriyorsa ve test sonucunda "A ve B ilaçlarının etkileri aynıdır." şeklinde ifade edilen yokluk hipotezi reddediliyorsa I.Tip Hata yapılmış demektir. Gerçekte ilaçlar farklı etkilere sahipken test sonucunda yokluk hipotezi kabul edilmiş ise bu durumda da II.Tip Hata yapılmış demektir. Her iki hatanın da başta tespit edilmesi, hipotezin daha gerçekçi analiz edilmesini sağlayacaktır. Bu hatalar göz önünde bulundurularak yeterli ve gerekli bilginin elde edilmesi için örneklem büyüklüğünün saptanması gerekmektedir.

Parametrik ve parametrik olmayan testleri birbirinden ayıran özellikleri Kazım Özdamar (2002) kitabında şu şekilde belirtmiştir. Parametrik yöntemler, ilgili parametreye, belirli bir dağılıma ve varyans kavramına dayanarak işlemler yapan esnek olmayan istatistiksel yöntemler olmasına karşın parametrik olmayan yöntemler, parametreye belirli bir dağılıma ve varyansa dayanmadan işlemler yapan genellikle

veriler yerine onların sıralama puanlarını kullanarak işlemler yapan esnek istatistiksel yöntemlerdir. Parametrik ve parametrik olmayan yöntemlerin karşılaştırılması ile ilgili örnekler şu şekildedir. Corrado 1989 yılında, etkinlik çalışmalarında, normal dağılım göstermeyen güvenlik-fiyat performansı için yeni bir parametrik olmayan sıralama testi oluşturmuştur. Günlük güvenlik-fiyat verileriyle yapılan simülasyonlar, sıralama testinin sıfır hipotezinde daha iyi ve alternatif hipotez altında parametrik t-testinden daha güçlü olduğunu gözlemlemiştir. Önceki parametik olmayan testlerden farklı olarak bu sıra testi, simetri gerektirmeyen bir yapıda olduğunu savunmuştur (Corrado, 1989). Campbell ve Wasley 1996 yılında, NASDAQ ve NYSE/ASE menkul kıymetlerinin örneklerini parametrik olmayan test istatistiği ve parametrik test istatistikleriyle karşılaştırmışlardır. NYSE/ASE ve NASDAQ menkul kıymetlerinin verilerinde, parametrik olmayan test istatistiklerinin daha güçlü olduğu tespit etmişlerdir (Campbell ve Wasley, 1996). Kaplan ve ark. 2003 yılında, operasyon öncesinde kalp cerrahisinde, mortalite tahmininde kullanılan EuroSCORE bir risk puanlama sistemi olduğunu belirtmişlerdir. Çalışmada EuroSCORE'un ülkemizdeki hastalara uygunluğunu test etmeyi amaçlamışlar ve gerçek mortalite hızının, yalancı mortalite hızına olan oranı Pozitif LR, yalancı sağkalım hızının, gerçek sağkalım hızına olan oranı da negatif LR olarak tanımlamışlardır. Pozitif LR ne kadar yüksekse ve negatif LR ne kadar düşükse testin gücünün de o kadar arttığı gözlemlemişlerdir (Kaplan ve ark., 2003). Finch 2005 yılında, normal varsayım ve homojen kovaryans matrislerinin varsayımları karşılamadığı durumlarda araştırmaların MANOVA test istatistiklerinin 1.Tip hata oranı güçleri azaltabilirken, artırabileceğini de gözlemlemiştir. Mevcut çalışmada bu iki varsayımın karşılanmadığı durumlarda parametrik test ve parametrik olmayan bir test karşılaştırılmıştır. Sonuçlar, homojen kovaryans matrisleri varsayımı karşılanmadığında parametrik olmayan yaklaşımın daha sağlam, daha düşük 1. Tip hata oranına ve daha yüksek güce sahip olduğunu gözlemlemiştir. Normallik varsayımının sağlanamadığı durumlarda 1. Tip hata oranı ve güç açısından parametrik test, parametrik olmayan testten daha iyi performans göstermiştir (Finch, 2005). Chen ve ark. (2008), Makalelerinde, Yapısal Eşitlik Modellemelerindeki uyum iyiliğinin bir ölçüsü olan, ortalama karekök hatası (RMSEA) test istatistiğinin sabit kesme noktalarının seçimi için bir değerlendirme olduğunu belirtmişlerdir. Simülasyon verileri kullanılarak önce evrensel bir ayrım için herhangi bir deneysel kanıt olup olmadığı incelemişlerdir. RMSEA'nın nokta tahminini tek başına kullanarak ilgili güven aralığı ile birlikte kullanımını karşılaştırmışlardır. Çalışmanın sonucunda, belirli bir güç seviyesine veya

1. Tip hata oranına ulaşmak için, kesme değerlerinin seçiminin, model özellikleri, serbestlik derecelerine ve örnek büyüklüğüne bağlı olduğunu ileri sürmüşlerdir (Chen ve ark., 2008). Coşkun ve Keskin, varyansların homojen olduğu durumları analiz etmek için Levene ve Bartlett testlerinin uygulanmasında oluşabilecek sorunları incelemişlerdir. Minitab 14 ve SPSS 13 yazılımlarında varyans homojenliğini test ederken zayıf noktalar üzerinde durmuşlardır. Levene ve Bartlett testlerinin bazı durumlar için testin güçlerini Monte Carlo yöntemiyle kıyaslamışlardır (Coşkun ve Keskin, 2008). Çarkungöz ve Bülent, ilişkili ama bağımsız çalışma sonuçlarının sentezi olarak tanımlanan Meta Analizi, çoklu çalışmaların nicel sentez tekniği olarak kullanılmasında Meta Analizinin etkisinin büyüklüğüne değinmişlerdir. Aynı konudaki farklı çalışmalardan bilgi toplayarak birleştirme analizini, sadece tek bir çalışmaya dayanan analizden daha fazla istatistiksel güce sahip olacağını düşünmüşlerdir (Çarkungöz ve Bülent, 2009). Garcia ve ark. (2009), Son yıllarda sezgisel yöntemler alanında deneysel analize ilginin arttığını belirtmişlerdir. Yöntemlerin karşılaştırılması ve farklı metodların incelenmesi gibi çeşitli sorunların analiz edilmesi ve önerilmesi için bir çalışma gerçekleştirmişlerdir. Çalışmalarında, sezgisel yöntemlerin optimizasyon problemleri üzerindeki analizinde istatistiksel tekniklerin kullanımı üzerinde durmuşlardır. Çalışma tek problemli ve çok problemli analiz olarak iki şekilde yürütülmüştür. Çoklu problem analizinde, parametrik olmayan istatistiksel testlerin kullanılması önermişlerdir. Örnek bir çalışma olarak, CEC'2005 Special Session on Real Parameter Optimization'da sunulan yöntemlerin yayınlanmış sonuçları analiz edilmiştir (García ve ark., 2009). Faul, Erdfelder, Lang ve Buchner (2007) G*power programının versiyon uzantılarını geliştirmeyi amaçlamışlardır. Yeni versiyonda tek örnekli tetrakorik korelasyonlar, bağımlı korelasyonların karşılaştırılması, iki değişkenli lineer regresyon, rastgele belirleyici model, çoklu doğrusal regresyona dayanarak testlerin gücünü analiz etmek için lojistik regresyon ve Poisson regresyonu gibi yöntemler eklemişlerdir (Faul ve ark., 2009). Mendeş ve Akkartal, varyans analizi tekniği (F) ve Welch (W) testi ile bunların permütasyon versiyonlarını (PF ve PW), 1.Tip hata ve testin gücü bakımından karşılaştırmışlardır. Bu karşılaştırmalarda Monte Carlo simülasyon tekniği kullanmışlardır. Yapılan simülasyon çalışmaları sonucunda varyanslar homojen iken, bu testlerin permütasyon versiyonlarının 1. Tip hata olasılığını koruma bakımından daha güvenilir sonuçlar verdiklerini gözlemlemişlerdir. Diğer taraftan varyansların heterojenleşmesinden bütün testlerin olumsuz yönde etkilendikleri görülmüştür. Varyansların heterojen ve dağılımlarının da çarpık olması halinde örnek

hacmi ve etki büyüklüğü ne olursa olsun PF testinin F testine göre biraz daha güçlü olduğu görülmüştür. Ancak dağılımlar simetrik iken PF ve F testlerinin güç değerleri benzerdir. W testi varyansların homojen olması halinde daha güçlü iken, PW testi varyansların homojen olmadığı ve örnek hacimlerinin dengesiz olduğu durumda biraz daha güçlüdür (Mendeş ve Akkartal, 2010). Karagöz 2010 yılında, parametrik olmayan teknikler kullanıldığında güç kaybı meydana geldiğini belirtmiş ve parametrik olmayan tekniklerin güç ve etkinlikleri ile parametrik tekniklerin güç ve etkinliklerini karşılaştırılmasını araştırmıştır. Araştırmaya, 75 akademisyen katılmıştır. Yapılan analiz sonucunda, parametrik olmayan tekniklerin az kullanıldığı (%19,5) ve kullananların da, bu testlerin güç ve etkinliklerine önem göstermedikleri (%19) gözlenmiştir (Karagöz, 2010). Demirel'de yine aynı yıl, klinik çalışmalarının planlanmasında önemli ve zor sayılabilecek bir etken olarak örneklem büyüklüğünün belirlenmesi olduğunu belirtmiştir. Çalışmada, doğru örneklem büyüklüğü belirlemenin gerekliliği üzerinde durulmuştur. İki örneklem testlerinden bağımlı ve bağımsız t testlerinde örneklem büyüklüğü belirlemek için tablolar sunulmuştur. Testin gücünün büyük ve önem düzeyinin küçük değerler almasının örneklem büyüklüğünü artıran nedenlerden biri olduğu tespit edilmiştir (Demirel ve Gürler, 2010). Razali 2011 yılında normallik varsayımı ihlal edildiğinde yorumlama ve çıkarımlar güvenilir veya geçerli olmayabileceğini belirtmiştir. N boyutunda bağımsız gözlemlerin rastgele bir örneğinin normal dağılım gösteren bir popülasyondan gelip gelmediğinin değerlendirilmesindeki üç yaygın yöntemi şu şekilde tanımlamıştır. Grafik yöntemler (histogramlar, kutucuklar, Q-Q grafikleri), Sayısal yöntemler (çarpıklık ve basıklık indisleri) ve Biçimsel normallik testleri. Çalışmasında, Shapiro-Wilk (SW) testi, Kolmogorov-Smirnov (KS) testi, Lilliefors (LF) testi ve Anderson-Darling (AD) testi gibi normallik testlerinin gücü karşılaştırılmıştır. Bu dört-testin güç karşılaştırmaları, oluşturulan örnek verilere Monte Carlo simülasyonu uygulanarak elde edilmiştir. Verilen dağılımların her birinden çeşitli örnek büyüklükleri oluşturulmuştur. Her testin gücü, normal istatistik testinin ilgili kritik değerlerle karşılaştırılmasıyla elde edilmiştir. Sonuçlar Shapiro-Wilk testinin Anderson-Darling testi Lilliefors testi ve Kolmogorov-Smirnov testi tarafından takip edilen en güçlü normallik testi olduğunu göstermektedir. Bununla birlikte, her dört-testin gücü, küçük örnek büyüklüğü için ise hala düşüktür (Razali ve Wah, 2011). Çapık da bir araştırmaya başlarken örneklem sayısını hesaplamak için bir yolun da güç analizi yapmak olduğunu belirtmiştir. Çalışmasının amacı, bir hemşirelik dergisinde 2009 – 2011 yılları arasında yayınlanmış olan makalelerde kullanılan istatistiksel

karşılaştırmaların güç analizini yapmak ve sonuçlarını sunmaktır. Tanımlayıcı türde yapılan bu çalışmada 725 istatistiksel testin güç analizi yapılmıştır. Güç analizleri Cohen tarafından bildirilen düşük, orta ve yüksek etki büyüklüğüne göre incelenmiş ve her test için üç farklı sonuç verilmiştir. Güç analizi yapabilmek için incelenen makalelerde kullanılan istatistiksel testleri, ortalamaları, standart sapmaları, çalışmaya alınan örneklem sayılarını ve diğer bilgileri not edebilmek için bir "yayın tanıtım formu" kullanmıştır. Makalelerde kullanılan parametrik testlerin güçlerinin, parametrik olmayan testlerin güçlerinden daha yüksek olduğunu gözlemlemiştir. Orta etki büyüklüğünde, testlerin %46.4'ünün, yüksek etki büyüklüğünde ise %88.7'sinin istatistiksel gücü 0.80 ve üstüdür. Çalışmanın sonuçlarına göre hemşirelik araştırmalarında güç analizi yaparak örneklem sayısını belirlemek yaygın değildir. İncelenen çalışmaların bir kısmının istenilen güce ulaşması için daha fazla örnekleme ihtiyacı olduğunu saptamıştır. Bu nedenle hemşire akademisyenlere, araştırmalarına başlamadan önce istatistiksel güç hesaplayarak örneklem büyüklüğünü belirlemelerini önermiştir (Çapık, 2013). Bindak ise 2014 yılında, Parametrik t-testinin varsayımları karşılanmadığı durumda alternatif olarak Mann-Whitney U testi kullanıldığını belirtmiş ve varsayımlar yerine getirildiğinde t-testinin daha güvenli olduğunu belirtmiştir. Anakütlenin normal dağılımlı olması durumunda ve çeşitli örnek hacimlerinde Mann-Whitney U testi ile t-testinin performanslarını karşılaştırmıştır. Anakütleden örnek seçme için Monte Carlo simülasyonu kullanmıştır. Her bir kombinasyon ve farklı örnek hacimleri için çekilen örneklere, her iki test uygulanmış ve reddedilen sıfır hipotezi sayısı kaydedilmiştir. Küçük örnek ($n \leq 30$) hacminde Mann-Whitney U testinin daha az hata verdiği, H_0 hipotezinin red veya kabul edilmesinde Mann-Whitney U ile t-testlerinin %97 oranında aynı sonucu verdikleri ve eşit olmayan varyans durumunda özellikle örnek hacmi arttıkça Mann-Whitney U testinin 1. Tip hata oranını koruyamadığı görülmüştür (Bindak, 2014). Karakoç, ölçek geliştirme üzerine olan çalışmasında, bir çok kriter ve belirli standartlara uygun çalışmasının gerekliliği üzerinde durmuştur. Aksi takdirde güvenilirlik ve geçerlilik düzeyinin düşük olacağı belirtilmiştir. Geçerlilik ve güvenilirliğin istenen seviyede olmayan bir ölçekle, testlerin güçlerinin düşeceği ve gruplar arası farkın anlamlı olarak saptanması engellenir. Güvenilirliğin düşük olduğunda ise, ölçümde "bias" a neden olan bazı hatalı kararların oluşabileceği saptanmıştır (Karakoç ve Dönmez, 2014). Doğan ve ark. (2015), En verimli istatistiksel testi seçmenin, istatistiğin temel problemlerinden biri olduğuna değinmiş ve Fraser verimlilik değeri kullanılarak farklı örneklem büyüklükleri için en

etkili testi seçmeyi amaçlamışlardır. Çalışmada örneklem büyüklüğü $2 \leq n \leq 50$ olarak belirlenmiş ve örneklem büyüklüğü arttıkça çalışmada yer alan tüm testlerin beklenen değerlerinin ve varyanslarının da arttığı gözlemlenmiştir. Fraser verimlilik değerine göre Mann-Whitney U testi en etkili test olarak öne çıkmıştır (Doğan ve Doğan, 2015). Miniç 2015 yılında incelediği çalışmada, hipotezlerin eş zamanlı karşılaştırılması sırasında, karşılaştırılan hipotez sayısı arttıkça en az bir doğru H_0 hipotezini yanlışlıkla reddetme olasılığının artmakta olduğuna değinmiştir. Araştırmacı m adet hipotezi eş zamanlı test ederken 1. tip hatanın artmamasını amaçlamıştır. Çalışmada, p-değerini düzelten ve deneysel hatayı (FWER) istenen düzeyde tutan 8 farklı yöntem, 5 farklı dağılım altında incelenmiştir. Sonuç olarak, Hommel ve Li yöntemlerinin diğer yöntemlerden daha güçlü olduğu ancak Li yönteminin deneysel hatayı belirlenen nominal seviyede koruyamadığı görülmüştür. Bonferroni ve Sidak yöntemlerinin ise karşılaştırılan hipotez sayısı arttıkça tutuculuğunun arttığını saptamıştır (Miniç, 2015). Çimen (2016), Wilcoxon Tek Örnek İşaret Sıralaması Testi Tek örnek T-testinin parametrik olmayan karşılığı olduğunu belirtmiştir. Araştırmasında, normal dağılımı test etmek için kullanılan süt pH verileri 1. ve 2. gün için bir çiftlikte yetiştirilen 9 adet inekten elde etmiştir. Mevcut çalışmadan elde edilen bulgulara göre 1. gün süt pH verileri normal dağılım göstermiştir. Bununla birlikte 2. gün verileri normal dağılım göstermemişlerdir. 2. gün verilerinden elde edilen p değerinden (0.000) de görüldüğü gibi, 2. gün verilerine tek örnek t-testini uygulaması mümkün olmayacağından 2. gün verilerine Wilcoxon Tek Örnek İşaret Sıralaması Testi uygulamıştır. Her iki test sonuçlarına göre 1. ve 2. gün verileri standart değerlerin uyum içinde olduğunu gözlemlemiştir (Çimen, 2016). Genç ve Soysal ise 2018 yılında hazırladıkları çalışmada, parametrik testlerin kullanılmasının istatistik olarak testin gücü bakımından önemli olduğunu belirtmişlerdir. Bu sebeple parametrik testlerden sonra grup ortalamaları parametrik çoklu karşılaştırma testleriyle karşılaştırılmıştır. Çalışmada bu testlerden; Asgari Önemli Fark, Duncan, Tukey (a), Student Newman Keuls, Lineer Bağıntılar (Linear contrasts), Bonferroni, Scheffe ve Dunnet Metodları incelenmiştir. Bilindiği üzere varyans analizinin varsayımları sağlanamamışsa parametrik testler yerine, parametrik olmayan (parametrik olmayan) testlerin kullanılması gerekmektedir. Bu amaçla parametrik olmayan (parametrik olmayan) çoklu karşılaştırma testlerinden; Hollander-Wolfe, Parametrik Olmayan Bonferroni (Parametrik olmayan Dunn's), Parametrik Olmayan Dunnet, Parametrik Olmayan S-N-K (Student-Newman-Keuls) ve

Parametrik Olmayan Tukey HSD (Tukey's Honestly Significan Difference) metotları irdelenmiştir. Ancak, bazı araştırmalar parametrik çoklu karşılaştırma testlerinin de parametrik olmayan testlerden sonra kullanılabileceğini göstermiştir (Genç ve Soysal). Sarı ve ark., Daha az örneklem kullanarak daha güçlü sonuçlar elde etmek amacıyla çapraz tasarımlar kullanmışlardır. Bu tasarımlar, psikolojik denemeler, tarım, zooloji gibi birçok alanda kullanılan tasarımlardır. Çapraz tasarımda etkilerin analizi için kullanılan t-testi ve Mann Whitney U testi simülasyon yoluyla deneysel 1. Tip hata oranları ve güç değerleri bakımından karşılaştırılmıştır. Bu amaçla farklı örnek çapları, farklı korelasyon katsayıları ele alınmış ve ayrıca testlerin performansları normal ve normal olmayan dağılımlar altında incelenmiştir. Simülasyon sonuçlarına göre, normal dağılım altında t-testinin güç değerleri daha yüksek iken, normal olmayan dağılımlar altında Mann Whitney U testinin güç değerlerinin daha yüksek olduğu gözlenmiştir (Sarı ve ark., 2018). Bircan, Ki-kare ve Kolmogorov-Smirnov uygunluk testleri üzerinde durulmuştur. Simülasyon ile oluşturulan kitleden alınan 20 örneklem üstünde bu iki uygunluk testleri uygulanmıştır. Elde edilen oran değerleri arasında ciddi bir fark olmadığından t testi ile değerlendirilmiştir. Bunun sonucunda Ki-kare testi yerine daha kolay kullanılan ve herhangi bir varsayıma bağlı olmayan Kolmogorov-Smirnov testinin uygulanabileceği söylenebilir (Bircan ve ark.).

Bu tez çalışmasında, istatistiksel sonuç çıkarımında sıkça kullanılan parametrik ve parametrik olmayan hipotez testlerinin farklı dağılımlara, farklı parametrelere ve farklı örnek hacimlerine (n) göre güçlerinin kıyaslanması ve böylece hangi durumda hangi testin gücünün daha fazla olduğunun vurgulanması amaçlanmıştır.

Tez çalışmasının ikinci bölümünde hipotez testleri ile ilgili temel kavramlara yer verilmiştir. Üçüncü bölümde, tek örneklem durumunda ortalama ve medyana ilişkin parametrik ve parametrik olmayan testler ile orana ilişkin parametrik ve parametrik olmayan testler ele alınmış ve bu testlerin güçleri farklı dağılım, farklı parametre ve farklı n'ler için elde edilmiş ve sonuçlar tablolar ve grafikler halinde sunulmuştur. Dördüncü bölümde, bağımsız iki örneklem durumunda ortalama ve medyana ilişkin parametrik ve parametrik olmayan testler ile orana ilişkin parametrik ve parametrik olmayan testler ele alınmış ve bu testlerin güçleri farklı dağılım, farklı parametre ve farklı n'ler için elde edilmiş ve sonuçlar tablolar halinde sunulmuştur. Beşinci bölümde, bağımlı iki örneklem durumunda ortalama ve medyana ilişkin parametrik ve parametrik olmayan testler ile orana ilişkin parametrik ve parametrik olmayan testler ele

alınmış ve bu testlerin güçleri farklı parametre ve farklı n'ler için elde edilmiş ve sonuçlar tablolar halinde sunulmuştur. Altıncı bölümde, bağımsız k $(k > 2)$ örneklem durumunda ortalama ve medyana ilişkin parametrik ve parametrik olmayan testler ele alınmış ve bu testlerin güçleri farklı dağılım, farklı parametre ve farklı n'ler için elde edilmiş ve sonuçlar tablolar halinde sunulmuştur. Yedinci bölümde, bağımlı k $(k > 2)$ örneklem durumunda ortalama ve medyana ilişkin parametrik ve parametrik olmayan testler ele alınmış ve bu testlerin güçleri farklı parametre ve farklı n'ler için elde edilmiş ve sonuçlar tablolar halinde sunulmuştur. Sekizinci bölümde sonuçlara ve önerilere yer verilmiştir.

2. TEMEL KAVRAMLAR

Bu bölümde hipotez testleri, parametrik hipotez testleri ve parametrik olmayan hipotez testleri hakkında temel kavramlar verilmiştir.

2.1. Hipotez Testleri

Hipotez testi, bir veya daha fazla kitleden alınan örneklem ile kitlenin ilgili parametresi için öne sürülen iddiadır. Hipotez testleri karşılaştırma ve seçim amaçlı olmasından başka bir hipoteze de ihtiyaç duymaktadır. H_0'a sıfır ya da yokluk hipotezi adı verilirken, H_1'e ise alternatif ya da araştırma hipotezi adı verilmektedir.

Hipotez testleri, kitle dağılımları hakkında ilgili varsayımları gerçekleştirme gerekliliğine dayanan parametrik hipotez testleri ve parametrik olmayan hipotez testleri olarak ikiye ayrılmaktadır.

Bir hipotez kabul ya da reddedildiğinde her zaman varılan kararın doğru olduğu söylenememektedir. Bir kitleden alınan örneklemlerin tüm kitle bilgilerini kapsaması beklenemeyeceği için kitle parametresi ve örneklem değeri arasındaki fark hata yapma riskidir. Buradaki hata payının α kadar olduğu düşünülmektedir ve bununla ilgili tablo aşağıdaki gibi gösterilmektedir.

	H_0 Doğru	H_0 Yanlış
H_0 Red	I.Tip Hata	Doğru Karar
H_0 Kabul	Doğru Karar	II. Tip Hata

Burada H_0 hipotezi doğru iken, reddedilmiş ise bu I. Tip Hata olarak adlandırılmaktadır. I. Tip Hata yapma olasılığının üst sınırı anlamlılık seviyesidir ve α ile gösterilebilir. $1-\alpha$ da kararın kabul edilme olasılığıdır.

H_0 yanlış iken kabul edilmesi II. Tip Hata (β) olarak adlandırılmaktadır. $(1-\beta)$ (doğru karar) yani testin gücü ise bir etkinin doğru şekilde tespit edilebilme ihtimalini vermektedir.

Testin gücü, verilen kararın doğruluğunu, güvenilirliğini denetlemek için kullanıldığı gibi örneklem büyüklüğünü belirlemek için de kullanılmaktadır.

2.1.1. Parametrik testler

Parametrik testlerde temel varsayım örneklemin alındığı kitle dağılımı hakkındadır. Bu varsayımlar arasında en yaygın olanı kitlelerin normal dağılıma sahip olmasıdır. Parametrik testler için normallik varsayımı, küçük örneklem büyüklükleri için çok kullanışlı değildir. Çok ciddi kanıtlar olmaması durumunda örneklem hacmi küçükken parametrik olmayan testler kullanılmalıdır.

Parametrik hipotez, bilinmeyen kitle hakkında bir varsayımdır. $H_0 : \gamma \in \Gamma_0 \subset \Gamma$, $H_1 : \gamma \in \Gamma_1 = \Gamma - \Gamma_0$ şeklinde hipotezler oluşturulur. $\Gamma_0(\Gamma_1)$ yalnızca bir eleman içerdiği durumda basit hipotez, diğer durumlarda bileşik hipotez olarak adlandırılır.

$X = (X_1, X_2, ..., X_n)$ örneklem değerleri olmak üzere H_0 hipotezini red veya kabulü için bir karar fonksiyonu bulunur. Yani

$\mathbb{R}^n$ uzayı C ve C' şeklinde ayrık iki küme olsun. $X \in C$ ise $H_0 : \gamma \in \Gamma_0$ hipotezi reddedilip, H_1 hipotezi kabul edilir. $X \in C'$ ise de tam tersi şekilde H_0 kabul edilir.

2.1.2. Parametrik olmayan testler

Parametrik olmayan testler, temel kitle dağılımının şekli veya parametrelerine yönelik varsayımlara dayanmaz ve parametrik test varsayımlarının yerine getirilemediği durumlarda kullanılmaktadır.

Parametrik olmayan testler, gerekli test varsayımlarının sağlandığı durumlardaki parametrik testlerden daha güçsüzdür. Diğer bir deyişle II. Tip Hata parametrik olmayan testlerde daha büyüktür. Bu da parametrik olmayan testlerin dezavantajıdır (Kartal, 2006).

3. TEK ÖRNEKLEM TESTLERİ

3.1. Ortalama ve Medyana İlişkin Testler

Bu bölümde tek örneklem testlerinden ortalama ve medyana ilişkin parametrik Z testi ve t testi ile parametrik olmayan işaret testi ve Wilcoxon testi tanıtılmıştır.

3.1.1. Z testi

$X_1, X_2, ..., X_n \sim N\left(\mu, \sigma^2\right)$ ve σ^2'nin bilindiği durumlarda, μ'nün μ_0 gibi teorik bir değere eşit olup olmadığını test etmek için kullanılan bir tek örneklem parametrik testtir.

$H_0 : \mu = \mu_0$ hipotezi, $H_1 : \mu \neq \mu_0$ hipotezine karşı test ediliyor ise test istatistiği,

$$Z_{Hesap} = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} \tag{3.1}$$

şeklinde hesaplanır ve eğer $\left|Z_{Hesap}\right| > Z_{\alpha/2}$ ise ilgili değişkene ilişkin kitle ortalaması μ'nün μ_0'dan farklı olduğu sonucuna varılır yani H_0 reddedilir ve burada $Z_{\alpha/2}$; standart normal dağılımının üst $\alpha/2$'lik yüzdeliğidir. (3.1) eşitliğinde $\overline{x}$, örneklem ortalaması, μ, varsayılan kitle ortalaması, σ, kitle standart sapması, n, örneklem büyüklüğüdür.

3.1.2. t testi

Tek örneklem t testi için de normal dağılıma uygun olması varsayımı esas alınmaktadır. Tek örneklem için t testi, tek bir değişkenin ortalamasının belirtilen bir sabitten farklı olup olmamasını incelemektedir (Chan, 2003).

$X_1, X_2, ..., X_n \sim N\left(\mu, \sigma^2\right)$ ve σ^2'nin bilinmediği durumlarda Z testi yerine t testi uygulanır. μ'nün μ_0 gibi teorik bir değere eşit olup olmadığını test etmek için kullanılan bir tek örneklem parametrik testtir.

$H_0 : \mu = \mu_0$ hipotezi, $H_1 : \mu \neq \mu_0$ hipotezine karşı test ediliyor ise test istatistiği,

$$t_{Hesap} = \frac{\overline{x} - \mu}{S / \sqrt{n}} \tag{3.2}$$

şeklinde hesaplanır ve $\left|t_{Hesap}\right| > t_{n-1;\alpha/2}$ ise kitle ortalaması μ'nün μ_0'dan farklı olduğu sonucuna varılır yani H_0 reddedilir. (3.2) eşitliğinde $\overline{x}$, örneklem ortalaması; μ, varsayılan kitle ortalaması; S, örneklem standart sapması; n, örneklem büyüklüğüdür.

3.1.3. İşaret testi

Kitle ortancası üzerine kurulmuş hipotezlerin test edilmesinde kullanılan bu test kitle ortalamasının anlamlılık testinin parametrik olmayan karşılığıdır. Örneklemin alındığı kitlenin normal dağılım göstermediği durumlarda kullanılmaktadır. Bu testin, binom testinin özel bir hali olduğu düşünülmektedir. Testin varsayımları ise şu şekildedir.

- Örnek, medyanı bilinmeyen n hacimli bir kitleden rastgele seçilmişdir.
- İlgilenilen değişken en az sıralama düzeyinde ölçülmüştür.
- İlgilenilen değişken süreklidir.

$X_i = \left(X_1, X_2, ..., X_n\right)$ örneklem değerleri, $D_i = X_i - \mu_0$, $i = 1, 2, ..., n$ olmak üzere

$$\delta_i = \begin{cases} 1, & D_i > 0 \\ 0, & D_i < 0 \end{cases} \tag{3.3}$$

şeklinde D_i ve δ_i değişkenleri tanımlanmaktadır ve test istatistiği,

$$k_{Hesap} = \sum_{i=1}^{n} \delta_i \tag{3.4}$$

olarak gösterilmektedir ve olasılık yoğunluk fonksiyonu,

$$f(k)=\binom{n}{k}\left(\frac{1}{2}\right)^{k}\left(\frac{1}{2}\right)^{n-k} \tag{3.5}$$

şeklinde hesaplanır ve $k_{Hesap} \le k_{\alpha/2}$ ise μ'nün μ_0'dan farklı olduğu sonucuna varılır yani H_0 reddedilir. (3.3) eşitliğinde D_i, μ_0'dan büyük değerli X_i gözlemlerin sayısı, (3.4) eşitliğinde k, işaret test istatistiği, n, örneklem büyüklüğü, (3.5) eşitliğinde $f(k)$, k istatistiğinin olasılık yoğunluk fonksiyonudur.

Bu test farkların dağılımı ya da hangi kitleden geldiği hakkında herhangi bir varsayımda bulunmamaktadır. Sıfır hipotezinin doğru olması durumunda farkların yarısı pozitif diğer yarısı ise negatif olmalıdır. Örnek büyüklüğünün 25'ten küçük olduğu durumlarda tablo değeri kullanılarak, aksi durumda normal dağılıma yaklaşım ilkesi ile çözüm yapılabilmektedir (Önder).

3.1.4. Wilcoxon işaretli sıra sayıları testi

Wilcoxon işaretli sıra sayıları testi, işaret testinden farklı olarak H_0 hipotezi test edilirken, gözlem değeri ve medyan değerleri arasındaki farkların işaretlerinden ziyade farkların büyüklükleri de hesaba katılmaktadır. İşaret testine göre daha fazla bilgi barındırmasından kaynaklı daha güçlü olduğu düşünülmektedir. Bu test için gerekli varsayımlar ise,

- n hacimli örnek medyanının bilinmediği durumlarda kitleden rastgele seçilmiştir.
- Gözlemler birbirinden bağımsızdır ve ilgilenilen değişken süreklidir.
- Kitle dağılımı simetriktir.
- Ölçme düzeyi en az eşit aralıklıdır.

$H_0 : M = M_0$ hipotezi $H_1 : M \neq M_0$ hipotezine karşı test ediliyor ise

$$D_i = X_i - M_0 \qquad i = 1, 2, ..., n$$

$$\delta_i = \begin{cases} 1, & D_i > 0 \\ 0, & D_i < 0 \end{cases}$$

$$r(|D_i|) = \sum_{j=1}^{n} S\left(|D_i| - |D_j|\right) \tag{3.6}$$

$$S(u)=\begin{cases}1, & u\geq 0\\ 0, & u<0\end{cases}$$

olarak tanımlanmaktadır ve test istatistiği,

$$T_h^+=\sum r\left(|D_i|\right)\delta_i \tag{3.7}$$

(3.6) eşitliğinde tanımlanan $r\left(|D_i|\right)$, $|D_i|$ için sayı sıralı istatistiktir. (3.7) eşitliğinde tanımlanan T^+ istatistiği ise $0,1,...,\frac{n(n+1)}{2}$ değerlerini alabilen bir değişkendir. $T_h^+\leq T_{\alpha/2}$ olduğu durumlarda H_0 reddedilmektedir.

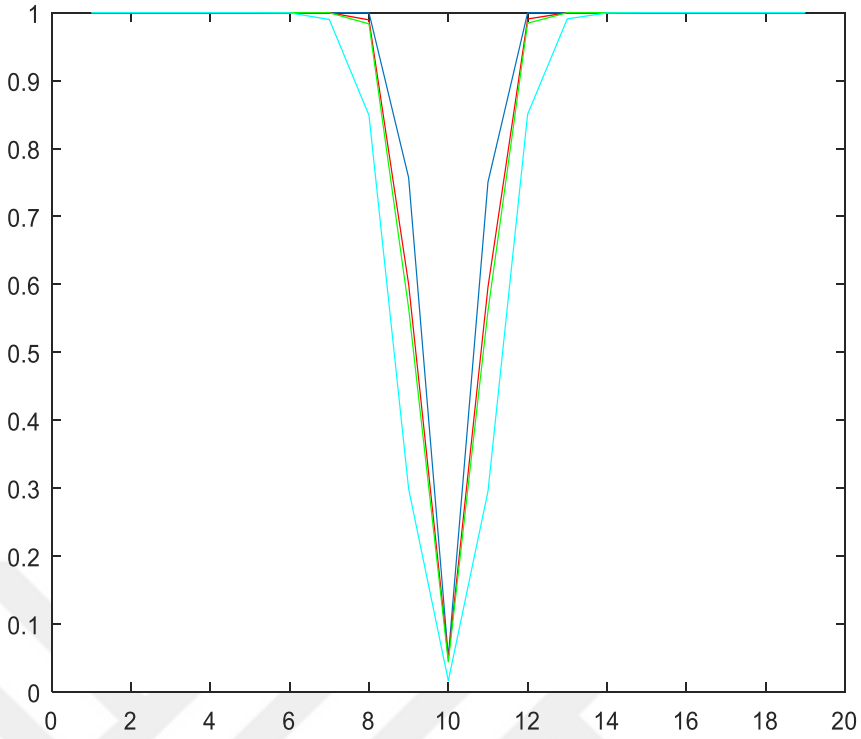
3.1.5. Güç karşılaştırması

Ortalama ve medyana ilişkin ele alınan parametrik Z ve t testi ile parametrik olmayan Wilcoxon ve İşaret testlerin güç karşılaştırmalarının yapılması amacıyla 10000 tekrarlı Monte-Carlo simülasyon yöntemi uygulanmıştır. Simülasyonlarda $H_0:\mu=10\left(M=10\right)$ hipotezine karşılık $H_1:\mu\neq 10\left(M\neq 10\right)$ hipotezi ve $\alpha=0.05$ anlamlılık seviyesi ele alınmıştır. Farklı dağılımlar, farklı parametre değerleri ve farklı n değerleri için elde edilen sonuçlar Çizelge 3.1 – 3.5 ve Şekil 3.1 – 3.35'te verilmiştir. Şekillerde mavi çizgi Z testinin gücünü, kırmızı çizgi t testinin gücünü, yeşil çizgi Wilcoxon testinin gücünü, parlak yeşil çizgi ise İşaret testinin gücünü göstermektedir.

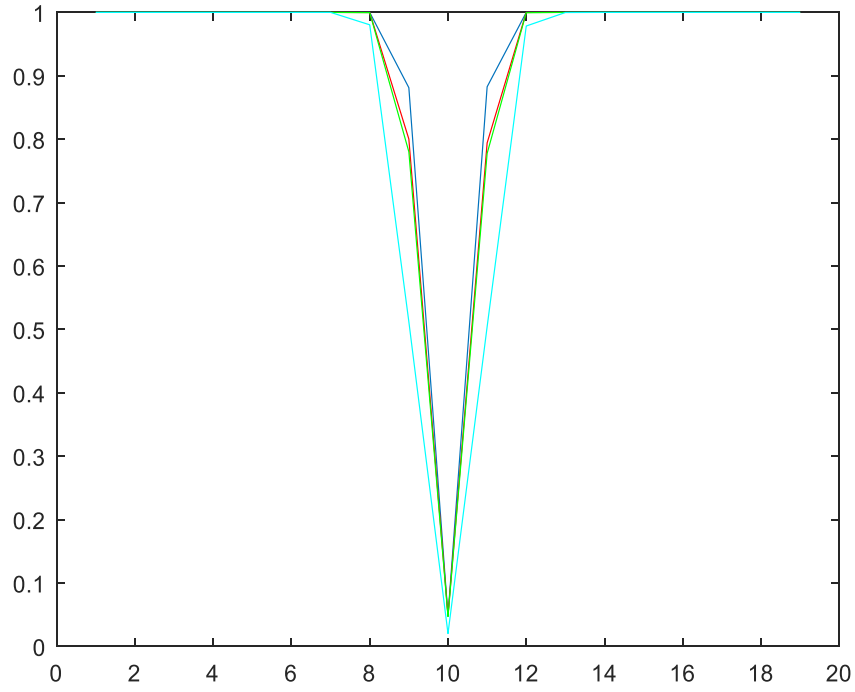
Çizelge 3.1. ve Şekil 3.1. – 3.7'de varyansı 1 olan normal dağılımlı bir kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, (10,1) parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 3.1. $X_1, X_2, \ldots, X_n \sim N(\mu,1)$ için testlerin güçleri

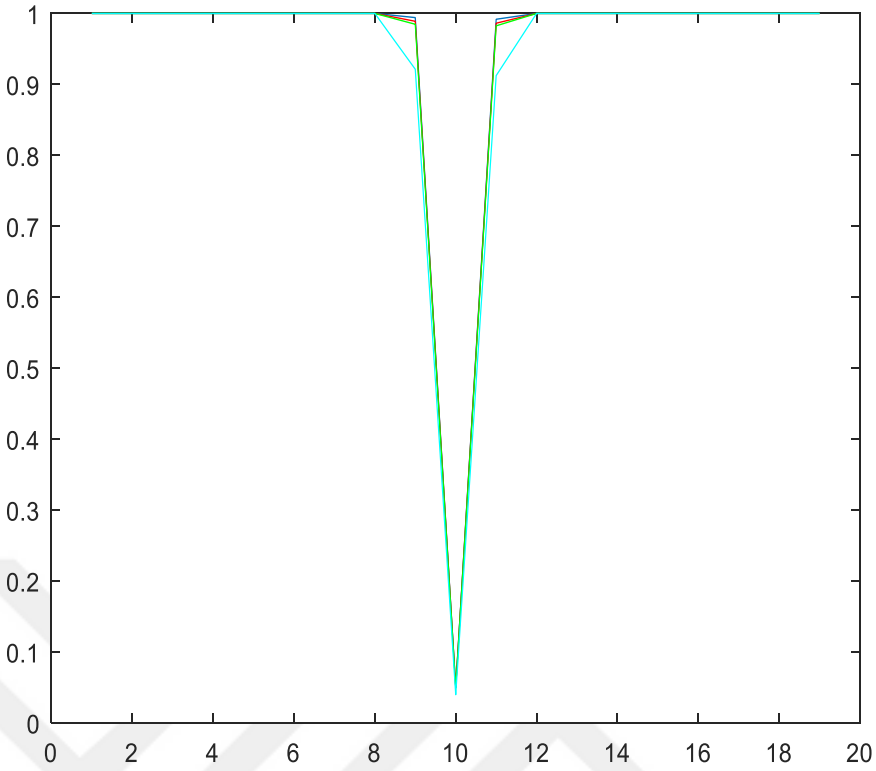
n	Parametre	Z testi	t testi	Wilcoxon	İşaret
7	(7, 1)	1.00	1.00	0.9999	0.99
	(8, 1)	0.9998	0.9894	0.9836	0.8489
	(9, 1)	0.7572	0.6002	0.5649	0.2989
	(10, 1)	0.0514	0.0492	0.0444	0.0165
	(11, 1)	0.7506	0.5961	0.5609	0.2953
	(12,1)	0.9997	0.9907	0.9845	0.8504
	(13,1)	1.00	1.00	1.00	0.9909
10	(7, 1)	1.00	1.00	1.00	1.00
	(8, 1)	1.00	0.9998	0.9996	0.9804
	(9, 1)	0.8812	0.8002	0.7801	0.5120
	(10, 1)	0.0481	0.0482	0.0470	0.0201
	(11, 1)	0.8826	0.7941	0.7776	0.5042
	(12,1)	1.00	0.9999	0.9995	0.9785
	(13,1)	1.00	1.00	1.00	0.9998
20	(7, 1)	1.00	1.00	1.00	1.00
	(8, 1)	1.00	1.00	1.00	1.00
	(9, 1)	0.9941	0.9889	0.9847	0.9213
	(10, 1)	0.0503	0.0507	0.0491	0.0395
	(11, 1)	0.9918	0.9862	0.9825	0.9123
	(12,1)	1.00	1.00	1.00	1.00
	(13,1)	1.00	1.00	1.00	1.00
30	(7, 1)	1.00	1.00	1.00	1.00
	(8, 1)	1.00	1.00	1.00	1.00
	(9, 1)	0.9997	0.9994	0.9995	0.9849
	(10, 1)	0.0501	0.0517	0.0509	0.0436
	(11, 1)	0.9996	0.999	0.9985	0.9859
	(12,1)	1.00	1.00	1.00	1.00
	(13,1)	1.00	1.00	1.00	1.00
40	(7, 1)	1.00	1.00	1.00	1.00
	(8, 1)	1.00	1.00	1.00	1.00
	(9, 1)	1.00	1.00	1.00	0.9985
	(10, 1)	0.0474	0.0512	0.0498	0.0374
	(11, 1)	1.00	1.00	1.00	0.9974
	(12,1)	1.00	1.00	1.00	1.00
	(13,1)	1.00	1.00	1.00	1.00
50	(7, 1)	1.00	1.00	1.00	1.00
	(8, 1)	1.00	1.00	1.00	1.00
	(9, 1)	1.00	1.00	1.00	0.9997
	(10, 1)	0.051	0.048	0.0493	0.0347
	(11, 1)	1.00	1.00	1.00	0.999
	(12,1)	1.00	1.00	1.00	1.00
	(13,1)	1.00	1.00	1.00	1.00
100	(7, 1)	1.00	1.00	1.00	1.00
	(8, 1)	1.00	1.00	1.00	1.00
	(9, 1)	1.00	1.00	1.00	1.00
	(10, 1)	0.0498	0.0498	0.0491	0.0341
	(11, 1)	1.00	1.00	1.00	1.00
	(12,1)	1.00	1.00	1.00	1.00
	(13,1)	1.00	1.00	1.00	1.00



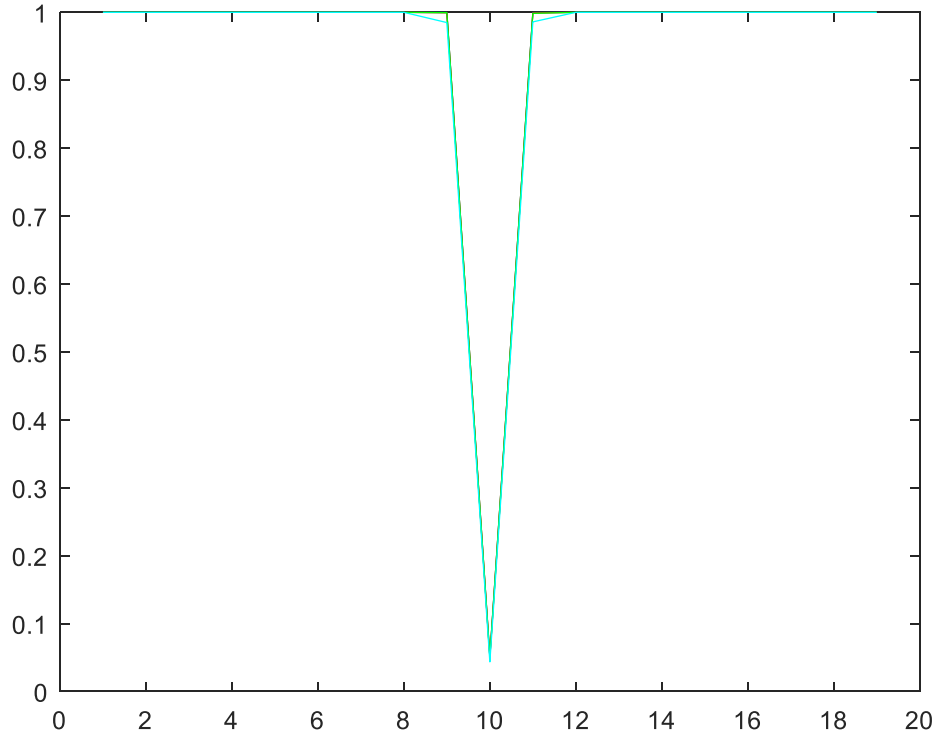
Şekil 3.1. $n = 7$ için güçler



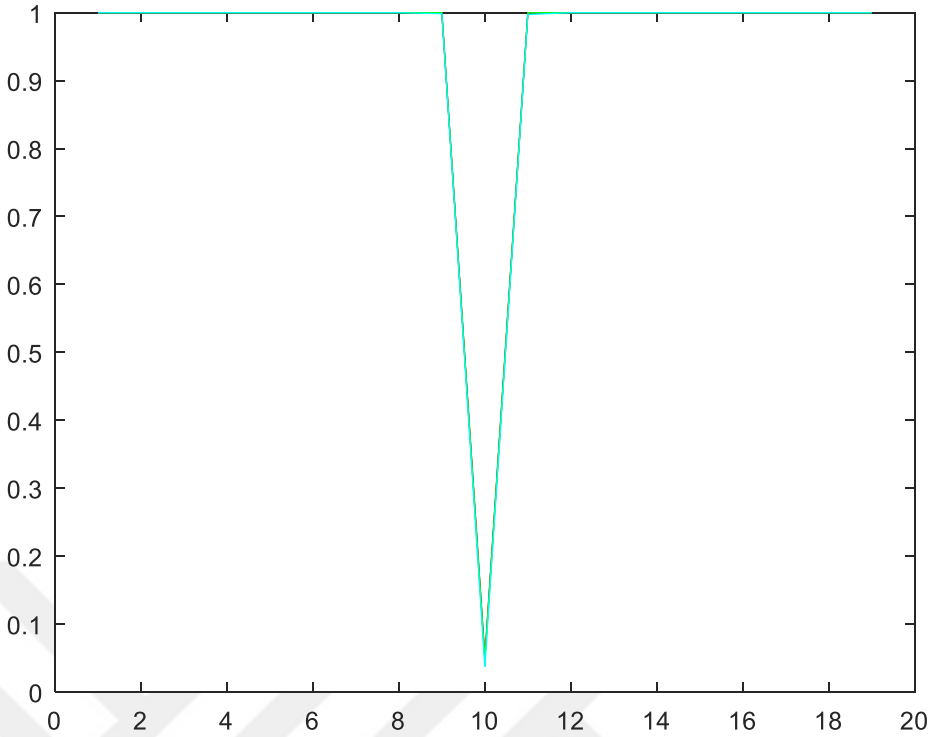
Şekil 3.2. $n = 10$ için güçler



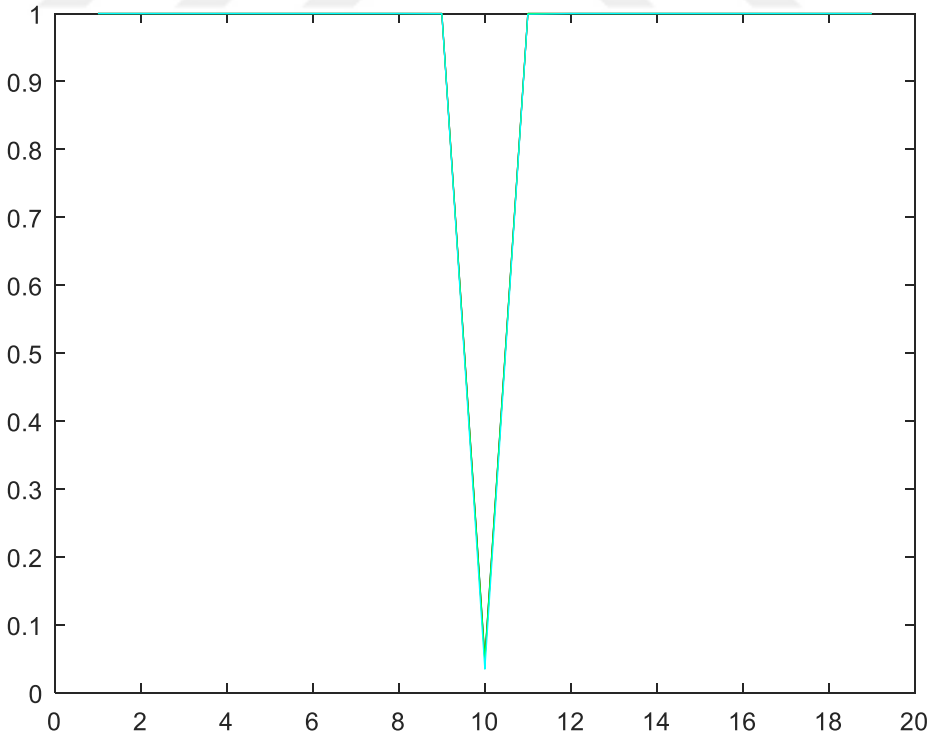
Şekil 3.3. $n = 20$ için güçler



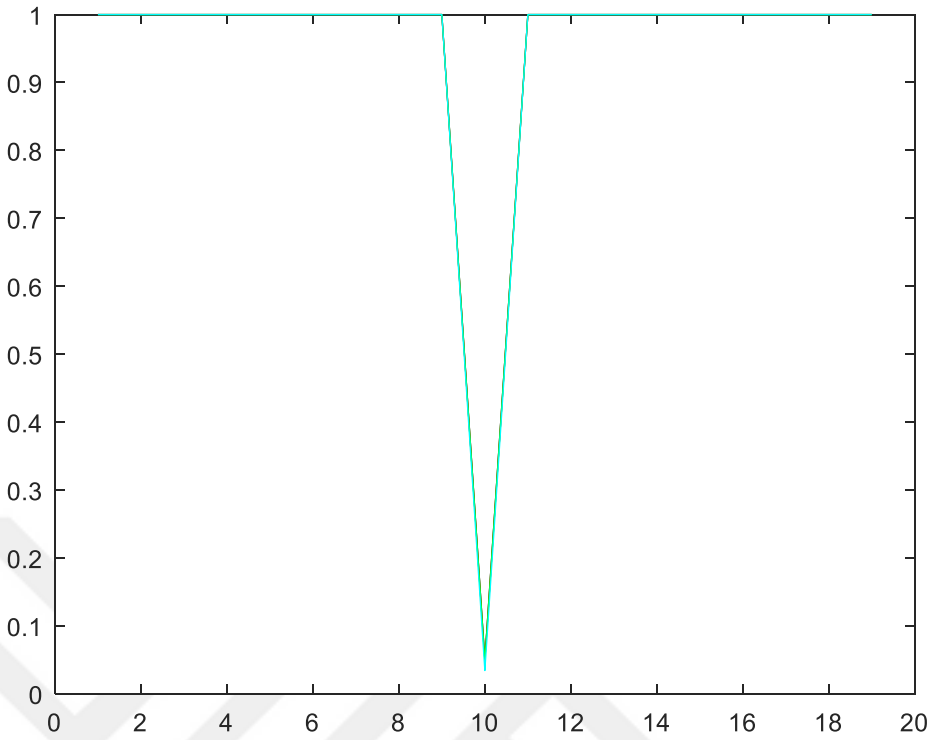
Şekil 3.4. $n = 30$ için güçler



Şekil 3.5. $n = 40$ için güçler



Şekil 3.6. $n = 50$ için güçler



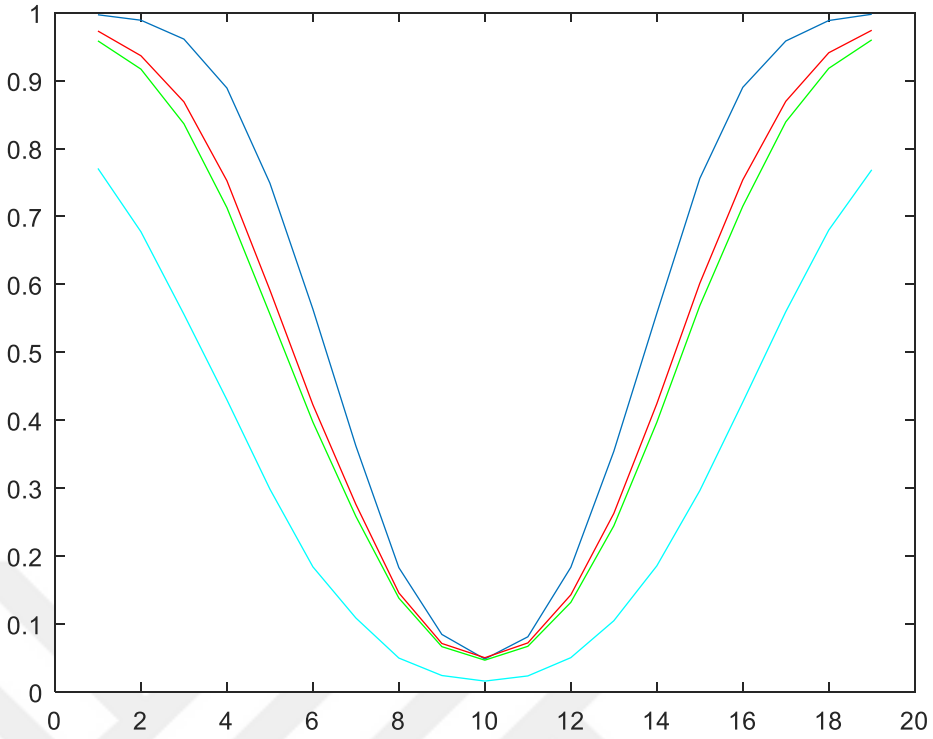
Şekil 3.7. $n = 100$ için güçler

Çizelge 3.1. ve Şekil 3.1.– 3.7'deki testlerin güçleri incelendiğinde bu dört test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test Z testidir. Bununla beraber n sayısı arttıkça testin gücü bakımından tüm testlerin birbirine yaklaştığı çizelge ve şekillerden görülmektedir. Küçük örnek hacimlerinde gücü en düşük olan test işaret testidir. Örnek hacmi 30'dan fazla olduğunda güç bakımından testler arasında farklılık olmadığı dolayısıyla istenilen testin kullanılabileceği söylenebilir.

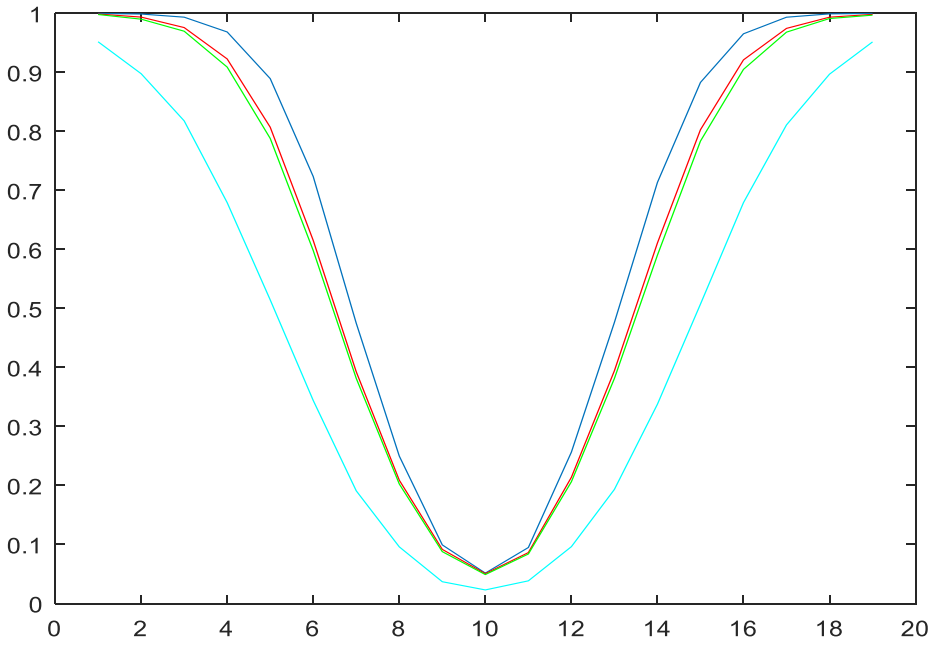
Çizelge 3.2. ve Şekil 3.8. – 3.14'te varyansı 25 olan normal dağılımlı bir kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, (10,25) parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 3.2. $X_1, X_2, \ldots, X_n \sim N(\mu, 25)$ için testlerin güçleri

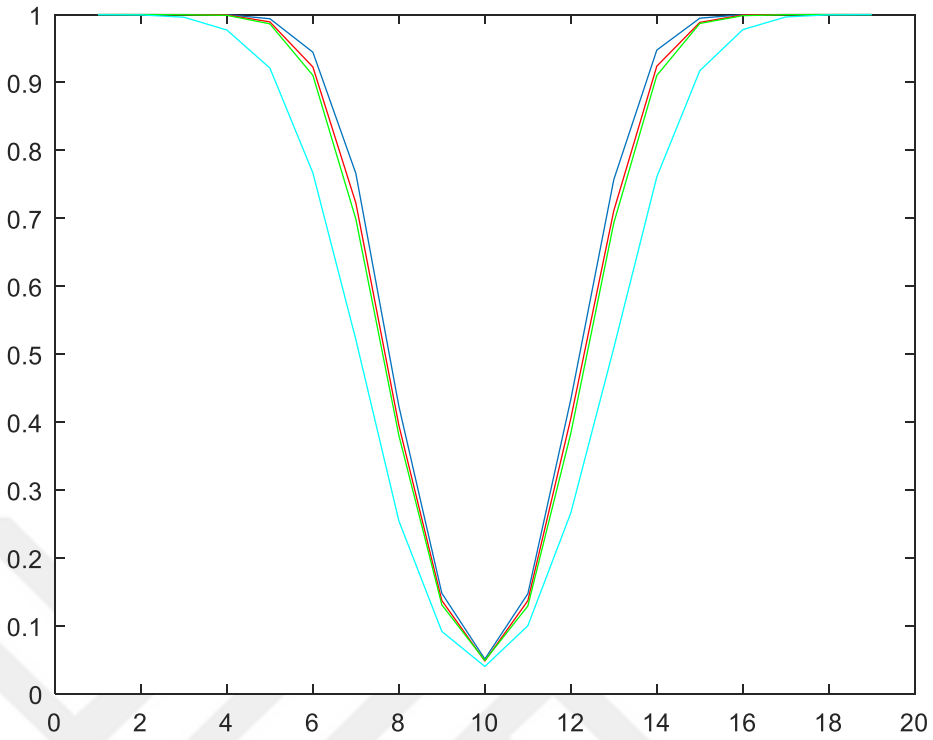
n	**Parametre**	**Z testi**	**t testi**	**Wilcoxon**	**İşaret**
7	(2, 25)	0.9889	0.9366	0.9169	0.6775
	(5, 25)	0.7490	0.5916	0.5562	0.2986
	(8, 25)	0.1830	0.1458	0.1381	0.0499
	(10, 25)	0.0485	0.0504	0.0468	0.016
	(12, 25)	0.1835	0.1430	0.1319	0.0505
	(15, 25)	0.7560	0.6017	0.5686	0.2964
	(18, 25)	0.9885	0.9410	0.9181	0.68
10	(2, 25)	0.999	0.9938	0.9903	0.8979
	(5, 25)	0.8896	0.8072	0.7882	0.5146
	(8, 25)	0.2499	0.2090	0.2026	0.096
	(10, 25)	0.0515	0.0506	0.0490	0.023
	(12, 25)	0.2565	0.2139	0.2063	0.0963
	(15, 25)	0.8831	0.8030	0.7837	0.5074
	(18, 5)	0.9993	0.9937	0.9914	0.8971
20	(2, 25)	1.00	1.00	1.00	0.9995
	(5, 25)	0.9939	0.9892	0.9863	0.9212
	(8, 25)	0.4234	0.3952	0.3802	0.2545
	(10, 25)	0.0511	0.0481	0.0489	0.0401
	(12, 25)	0.4335	0.4056	0.3835	0.2663
	(15, 25)	0.9947	0.9885	0.9869	0.9174
	(18, 25)	1.00	1.00	1.00	0.9996
30	(2, 25)	1.00	1.00	1.00	1.00
	(5, 25)	0.9998	0.9997	0.9992	0.9863
	(8, 25)	0.5932	0.5653	0.5454	0.3780
	(10, 25)	0.0469	0.0459	0.0454	0.0376
	(12, 25)	0.5937	0.5634	0.5425	0.3806
	(15, 25)	0.9998	0.9994	0.999	0.9862
	(18, 25)	1.00	1.00	1.00	1.00
40	(2, 25)	1.00	1.00	1.00	1.00
	(5, 25)	1.00	1.00	1.00	0.9969
	(8, 25)	0.7228	0.7005	0.6815	0.4696
	(10, 25)	0.0555	0.0520	0.0529	0.0401
	(12, 25)	0.7201	0.6985	0.6815	0.4679
	(15, 25)	0.9999	0.9999	0.9999	0.9973
	(18, 25)	1.00	1.00	1.00	1.00
50	(2, 25)	1.00	1.00	1.00	1.00
	(5, 25)	1.00	1.00	1.00	0.9997
	(8, 25)	0.81	0.7959	0.7753	0.5378
	(10, 25)	0.0486	0.0503	0.0479	0.0344
	(12, 25)	0.8139	0.7973	0.7752	0.5359
	(15, 25)	1.00	1.00	1.00	0.9993
	(18, 25)	1.00	1.00	1.00	1.00
100	(2, 25)	1.00	1.00	1.00	1.00
	(5, 25)	1.00	1.00	1.00	1.00
	(8, 25)	0.9778	0.9761	0.9704	0.8588
	(10, 25)	0.0488	0.0497	0.0501	0.0354
	(12,2 5)	0.9794	0.9786	0.9725	0.8562
	(15, 25)	1.00	1.00	1.00	1.00
	(18, 25)	1.00	1.00	1.00	1.00



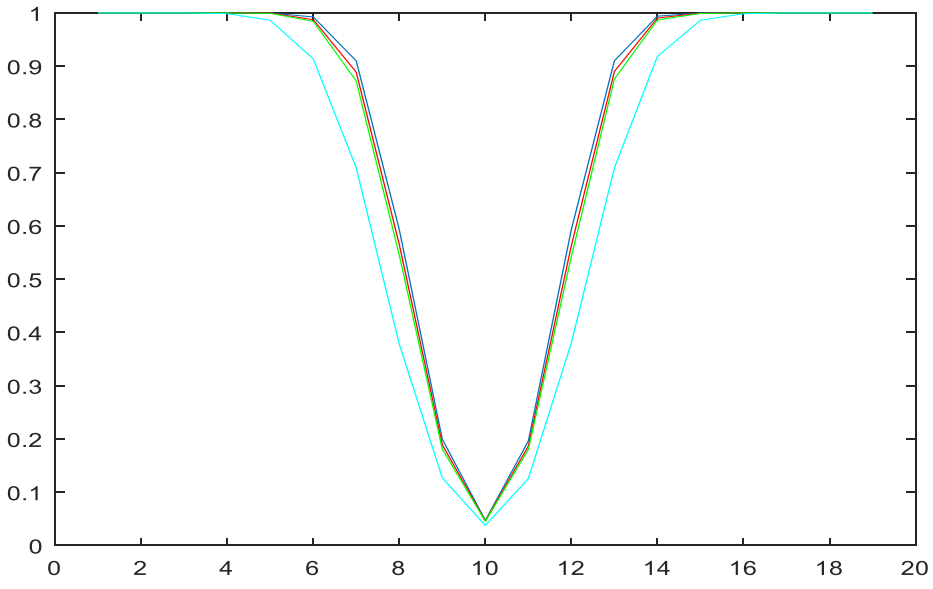
Şekil 3.8. $n = 7$ için güçler



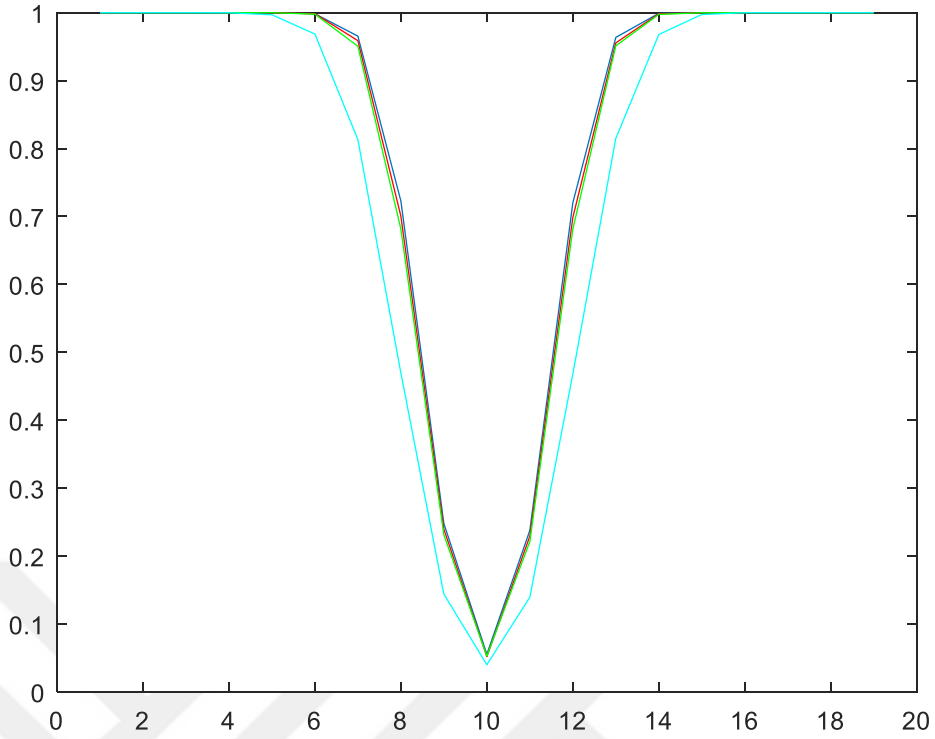
Şekil 3.9. $n = 10$ için güçler



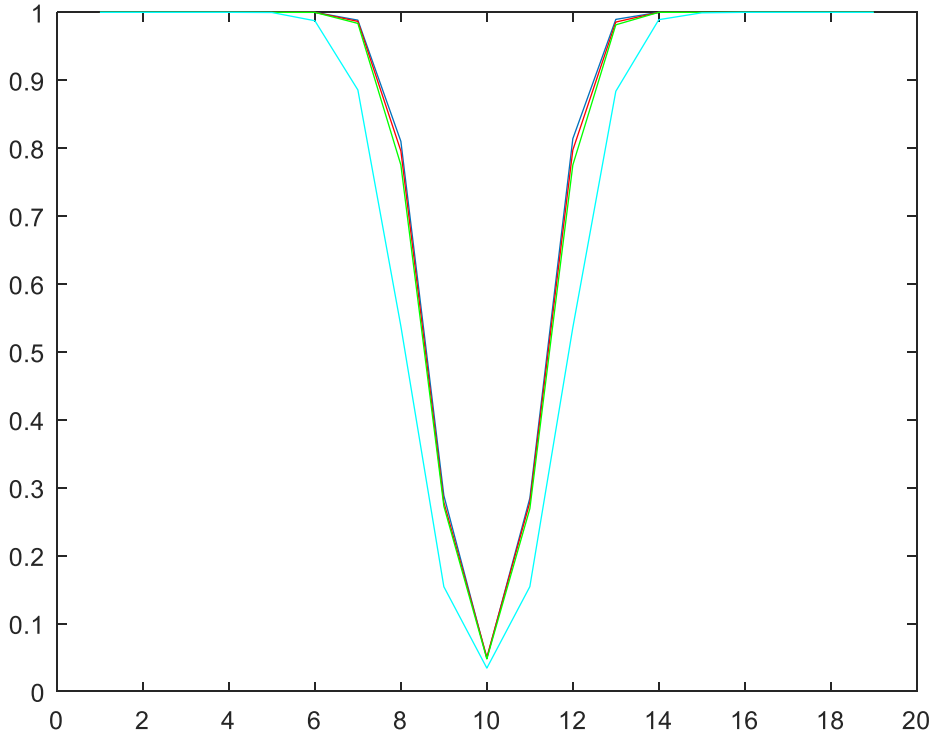
Şekil 3.10. $n = 20$ için güçler



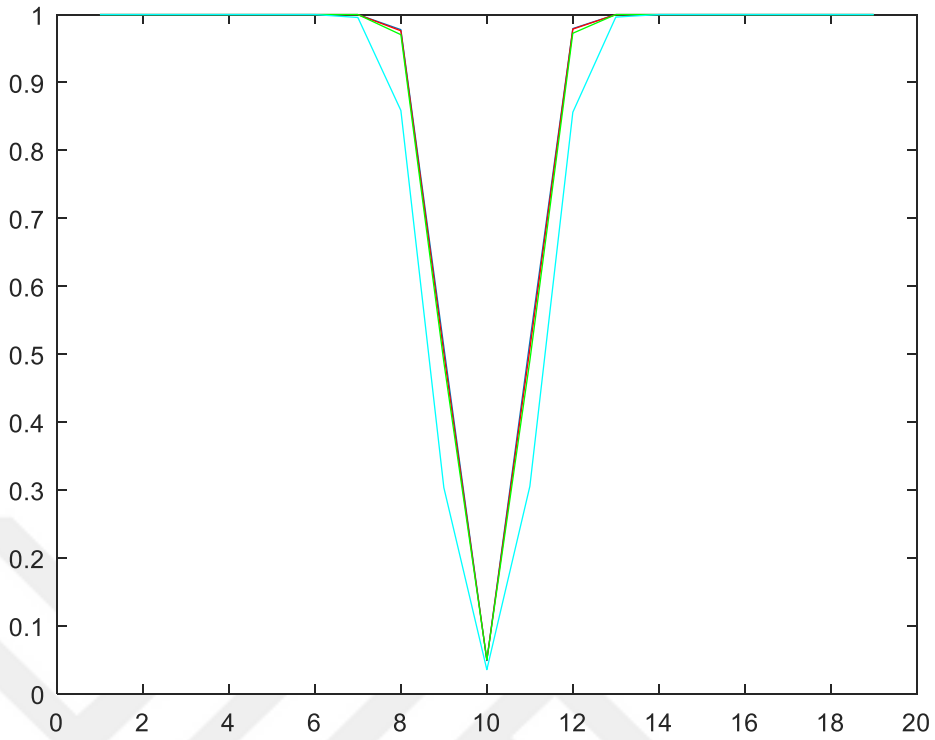
Şekil 3.11. $n = 30$ için güçler



Şekil 3.12. $n = 40$ için güçler



Şekil 3.13. $n = 50$ için güçler



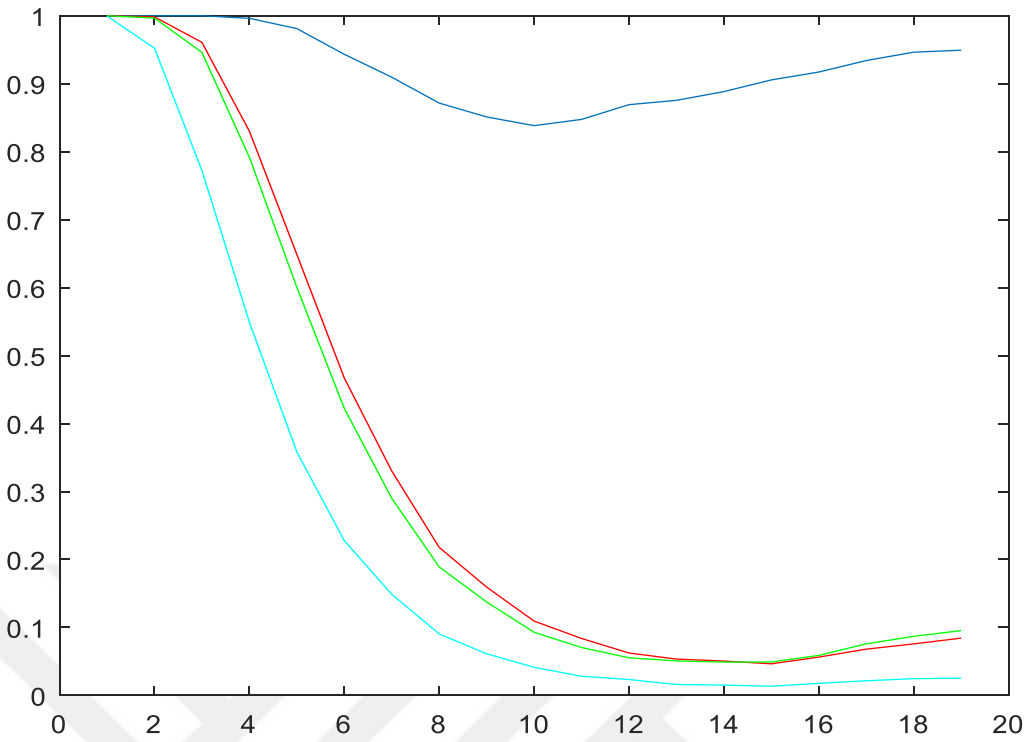
Şekil 3.14. $n = 100$ için güçler

Çizelge 3.2. ve Şekil 3.8. – 3.14'de testlerin güçleri incelendiğinde bu dört test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test Z testidir. Bununla beraber n sayısı arttıkça testin gücü bakımından tüm testlerin birbirine yaklaştığı çizelge ve şekillerden görülmektedir. Tüm durumlarda gücü en düşük olan test işaret testidir. Çizelge 3.1. ve Çizelge 3.2. birlikte ele alındığında ise normal dağılımlı bir kitlede varyans büyüdükçe testlerin güçlerinin azaldığı söylenebilir.

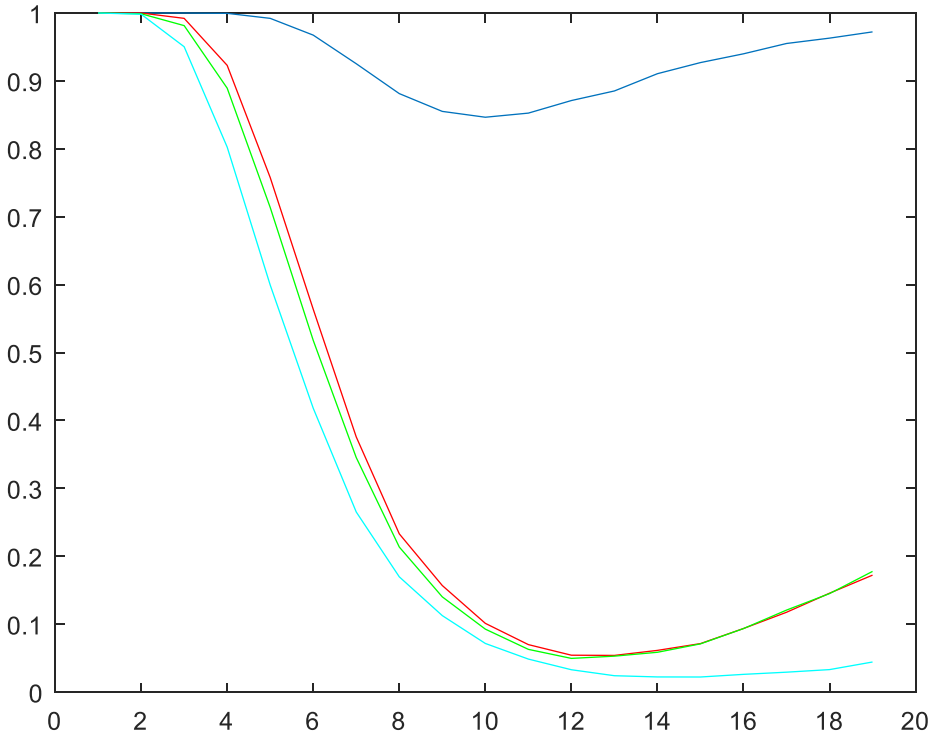
Çizelge 3.3. ve Şekil 3.15 – 3.21'de üstel dağılımlı bir kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, Z ve t testi için (10) parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 3.3. $X_1, X_2, \ldots, X_n \sim \ddot{u}stel(\theta)$ için testlerin güçleri

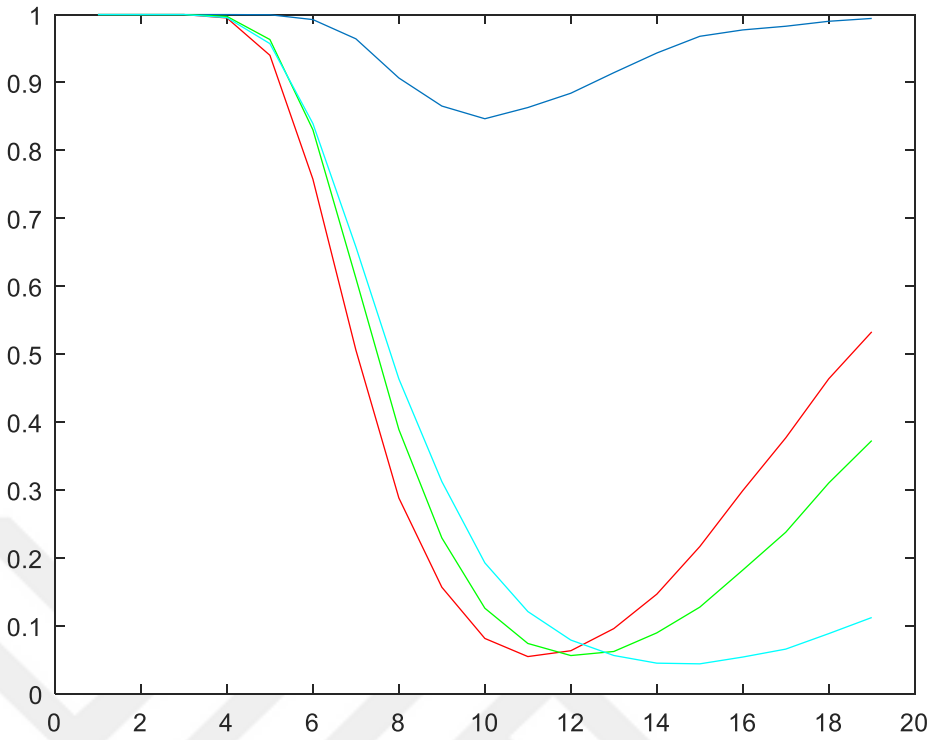
n	Parametre	Z testi	t testi	Wilcoxon	İşaret
7	(2)	1.00	0.9984	0.9966	0.9523
	(5)	0.9811	0.6488	0.6005	0.3583
	(8)	0.8715	0.2178	0.1891	0.0905
	(10)	0.8383	0.1094	0.0929	0.0414
	(12)	0.8692	0.0624	0.0554	0.0234
	(15)	0.9056	0.0465	0.0492	0.0135
	(18)	0.9464	0.0758	0.0869	0.0247
10	(2)	1.00	1.00	0.9989	0.9982
	(5)	0.9921	0.7577	0.7136	0.5994
	(8)	0.8813	0.2331	0.2137	0.1698
	(10)	0.8465	0.1013	0.0928	0.0717
	(12)	0.8711	0.0544	0.0497	0.0330
	(15)	0.9270	0.0715	0.0711	0.0223
	(18)	0.9630	0.1457	0.1451	0.0332
20	(2)	1.00	1.00	1.00	1.00
	(5)	0.9997	0.9400	0.9632	0.9572
	(8)	0.9066	0.2882	0.3890	0.4635
	(10)	0.8466	0.0814	0.1259	0.1927
	(12)	0.8842	0.0633	0.0562	0.0790
	(15)	0.9680	0.2167	0.1275	0.0440
	(18)	0.9901	0.4637	0.3103	0.0885
30	(2)	1.00	1.00	1.00	1.00
	(5)	1.00	0.9882	0.9968	0.9958
	(8)	0.9240	0.3375	0.5349	0.6436
	(10)	0.8419	0.0750	0.1570	0.2876
	(12)	0.9024	0.0811	0.0543	0.0985
	(15)	0.9823	0.3900	0.1771	0.0449
	(18)	0.9980	0.7290	0.4595	0.1184
40	(2)	1.00	1.00	1.00	1.00
	(5)	1.00	0.9970	0.9998	0.9993
	(8)	0.9426	0.4003	0.6603	0.7599
	(10)	0.8471	0.0667	0.1963	0.3529
	(12)	0.9118	0.1128	0.0580	0.1126
	(15)	0.9896	0.5497	0.2302	0.0413
	(18)	0.9998	0.8739	0.5783	0.1278
50	(2)	1.00	1.00	1.00	1.00
	(5)	1.00	0.9997	1.00	1.00
	(8)	0.9625	0.4647	0.7640	0.8421
	(10)	0.8425	0.0679	0.2322	0.4043
	(12)	0.9248	0.1558	0.0558	0.1130
	(15)	0.9954	0.6711	0.2735	0.0329
	(18)	1.00	0.9442	0.6828	0.1399
100	(2)	1.00	1.00	1.00	1.00
	(5)	1.00	1.00	1.00	1.00
	(8)	0.9891	0.6769	0.9622	0.9908
	(10)	0.8447	0.0597	0.3821	0.7129
	(12)	0.9659	0.3462	0.0565	0.2113
	(15)	0.9998	0.9587	0.5186	0.0420
	(18)	1.00	0.9997	0.9272	0.2712



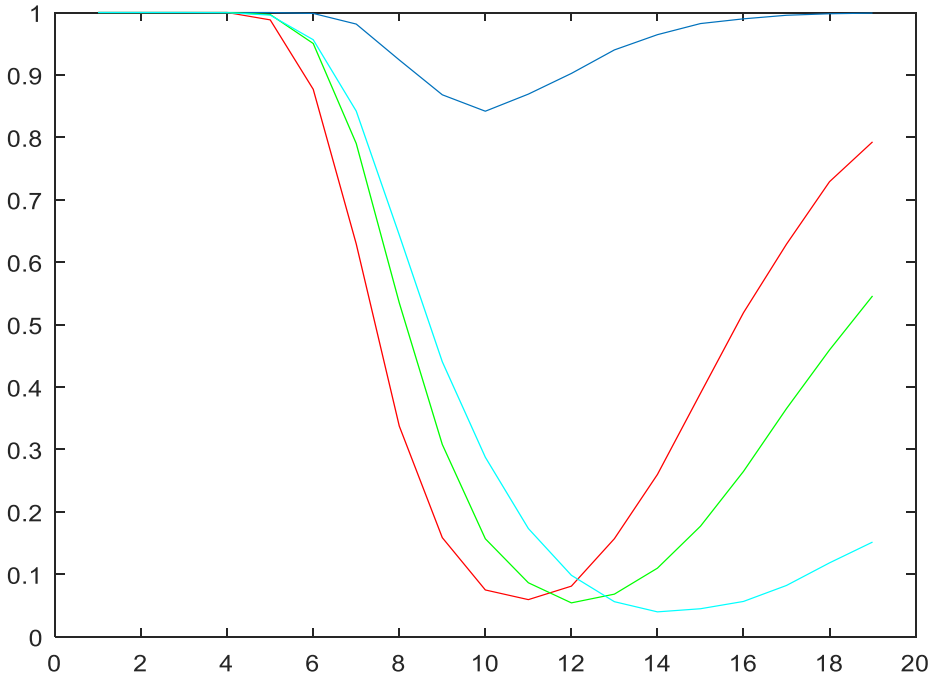
Şekil 3.15. $n = 7$ için güçler



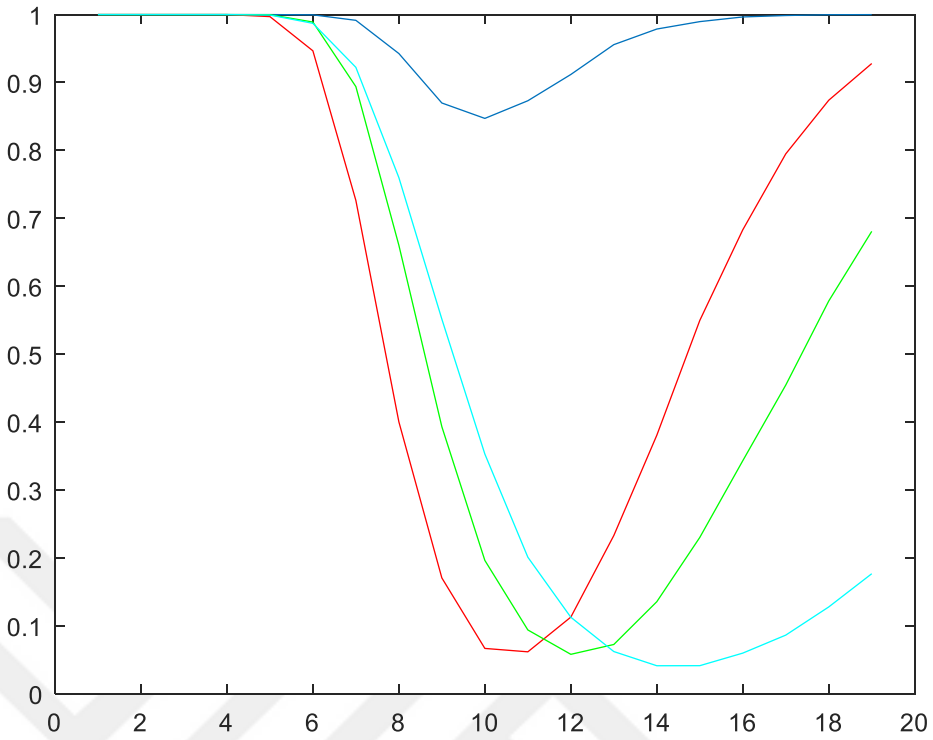
Şekil 3.16. $n = 10$ için güçler



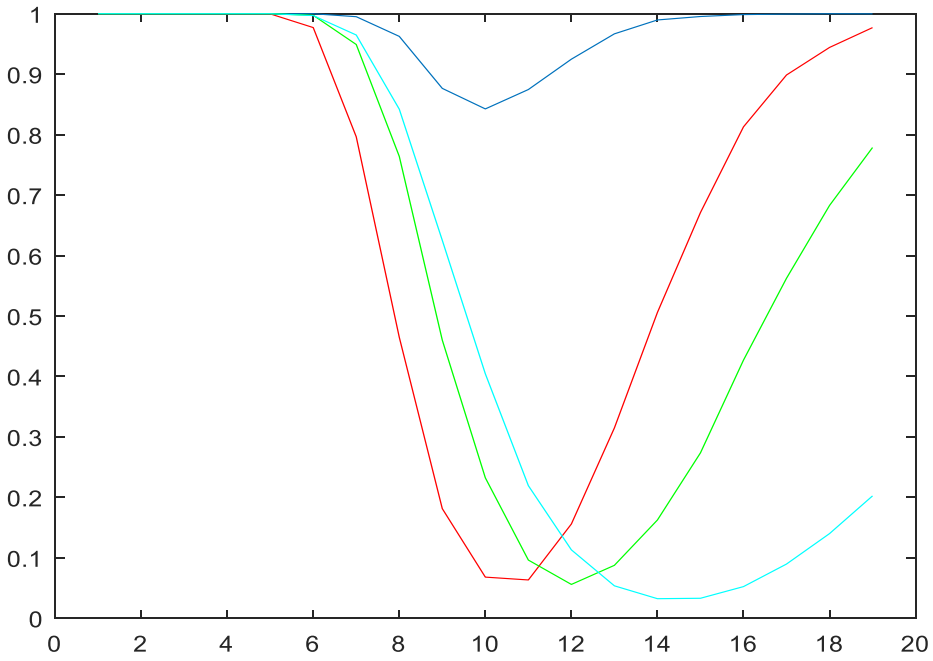
Şekil 3.17. $n = 20$ için güçler



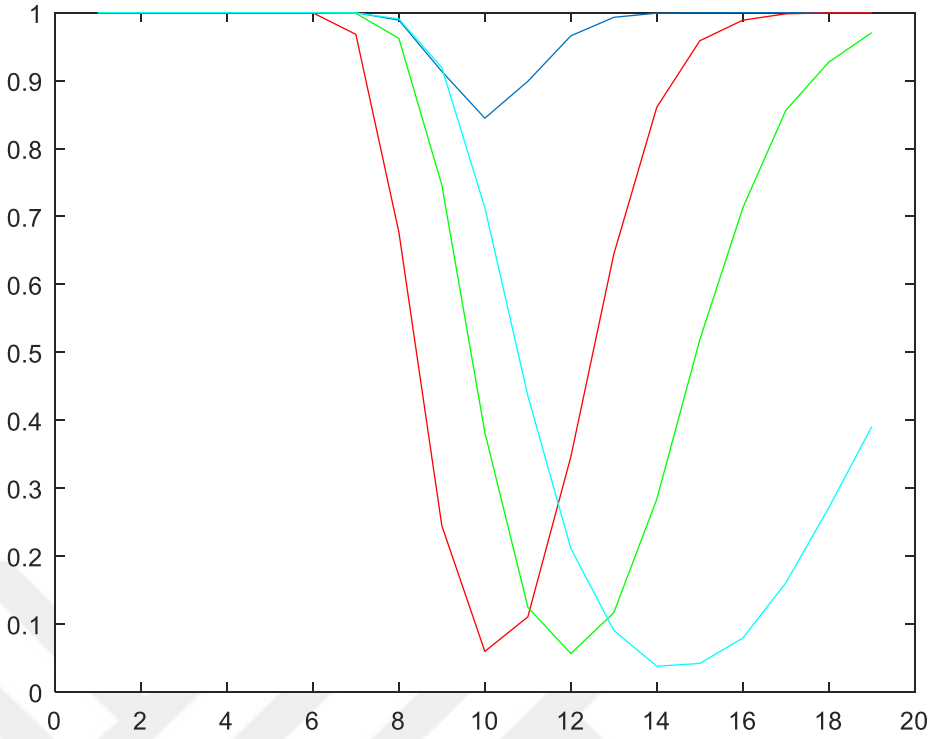
Şekil 3.18. $n = 30$ için güçler



Şekil 3.19. $n = 40$ için güçler



Şekil 3.20. $n = 50$ için güçler



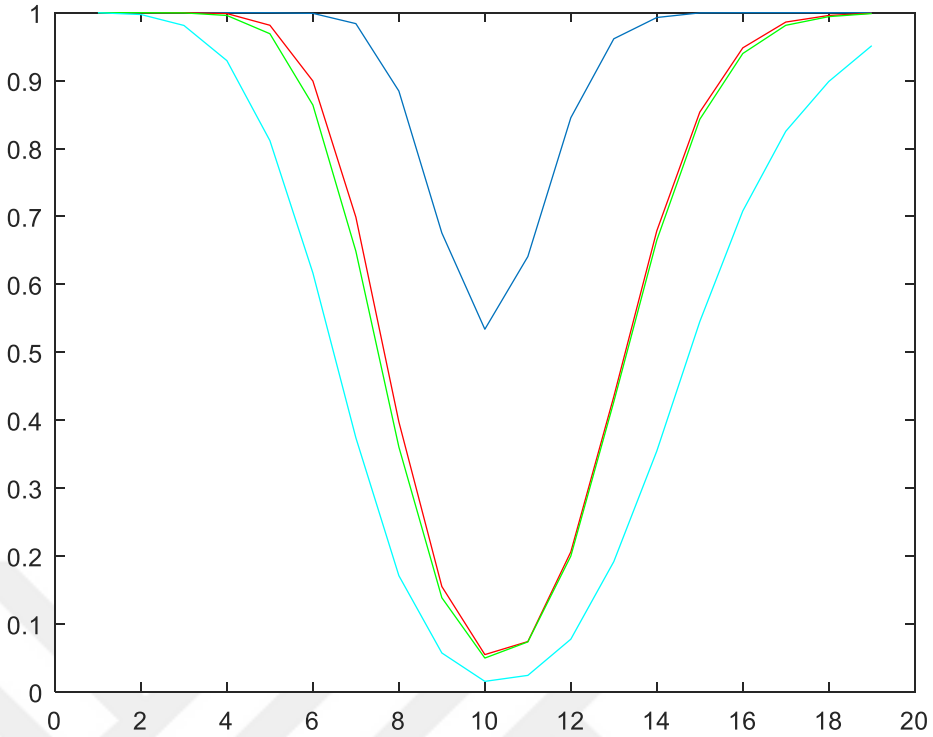
Şekil 3.21. $n=100$ için güçler

Çizelge 3.3. ve Şekil 3.15 – 3.21'deki testlerin güçleri incelendiğinde Z testinin gücünün yüksek çıktığı ancak anlamlılık seviyesinin de çok yüksek elde edilmiş olduğu görülmektedir. Küçük n değerlerinde t testinin anlamlılık seviyesinin de nispeten büyük olduğu ancak n arttıkça anlamlılık seviyesinin azaldığı görülmektedir Dolayısıyla, küçük n değerlerinde $(n<20)$ Wilcoxon işaretli sıra sayıları testinin, n büyüdükçe t testi veya işaret testinin kullanılması önerilebilir. Örnek hacmi arttıkça tüm testlerin gücü artmaktadır.

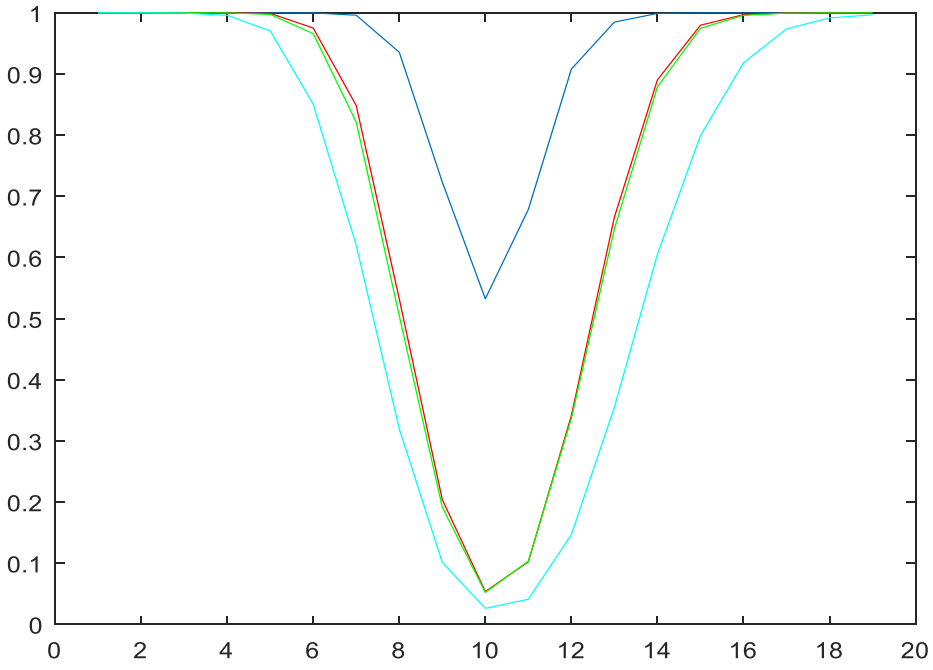
Çizelge 3.4. ve Şekil 3.22 – 3.28'de β parametresi 1 olan Gama dağılımlı bir kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, Z ve t testi için (10,1) parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 3.4. $X_1, X_2, \ldots, X_n \sim Gama(\alpha,1)$ için testlerin güçleri

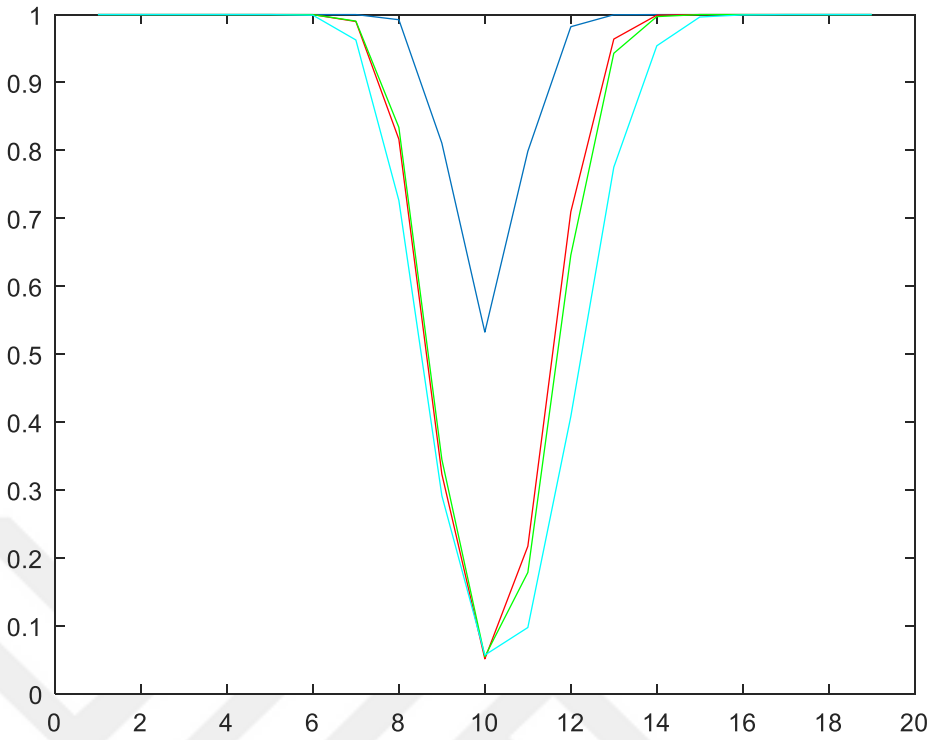
n	Parametre	Z testi	t testi	Wilcoxon	İşaret
7	(3, 1)	1.00	1.00	0.9997	0.9812
	(6, 1)	0.9990	0.8994	0.8640	0.6168
	(9, 1)	0.6757	0.1552	0.1387	0.0574
	(10, 1)	0.5340	0.0550	0.0499	0.0156
	(12, 1)	0.8454	0.2063	0.2001	0.0777
	(15,1)	0.9997	0.8537	0.8431	0.5453
	(18,1)	1.00	0.9959	0.9942	0.8987
10	(3, 1)	1.00	1.00	1.00	0.9995
	(6, 1)	0.9999	0.9752	0.9660	0.8506
	(9, 1)	0.7233	0.2040	0.1916	0.1014
	(10, 1)	0.5323	0.0536	0.0521	0.0262
	(12, 1)	0.9081	0.3414	0.3367	0.1468
	(15,1)	1.00	0.9798	0.9748	0.7992
	(18,1)	1.00	1.00	0.9999	0.9917
20	(3, 1)	1.00	1.00	1.00	1.00
	(6, 1)	1.00	0.9999	0.9999	0.9991
	(9, 1)	0.8107	0.3234	0.3450	0.2917
	(10, 1)	0.5324	0.0513	0.0546	0.0569
	(12, 1)	0.9821	0.7097	0.6459	0.4084
	(15,1)	1.00	1.00	1.00	0.9964
	(18,1)	1.00	1.00	1.00	1.00
30	(3, 1)	1.00	1.00	1.00	1.00
	(6, 1)	1.00	1.00	1.00	1.00
	(9, 1)	0.8834	0.4382	0.4950	0.4278
	(10, 1)	0.5406	0.0531	0.0643	0.0672
	(12, 1)	0.9969	0.8892	0.8312	0.5718
	(15,1)	1.00	1.00	1.00	0.9999
	(18,1)	1.00	1.00	1.00	1.00
40	(3, 1)	1.00	1.00	1.00	1.00
	(6, 1)	1.00	1.00	1.00	1.00
	(9, 1)	0.9296	0.5431	0.6214	0.5305
	(10, 1)	0.5410	0.0524	0.0644	0.0634
	(12, 1)	0.9992	0.9629	0.9298	0.6908
	(15,1)	1.00	1.00	1.00	1.00
	(18,1)	1.00	1.00	1.00	1.00
50	(3, 1)	1.00	1.00	1.00	1.00
	(6, 1)	1.00	1.00	1.00	1.00
	(9, 1)	0.9556	0.6318	0.7190	0.6054
	(10, 1)	0.5391	0.0493	0.0645	0.0678
	(12, 1)	0.9998	0.9879	0.9682	0.7677
	(15,1)	1.00	1.00	1.00	1.00
	(18,1)	1.00	1.00	1.00	1.00
100	(3, 1)	1.00	1.00	1.00	1.00
	(6, 1)	1.00	1.00	1.00	1.00
	(9, 1)	0.9956	0.8980	0.9542	0.9086
	(10, 1)	0.5361	0.0529	0.0800	0.1019
	(12, 1)	1.00	1.00	0.9997	0.9740
	(15,1)	1.00	1.00	1.00	1.00
	(18,1)	1.00	1.00	1.00	1.00



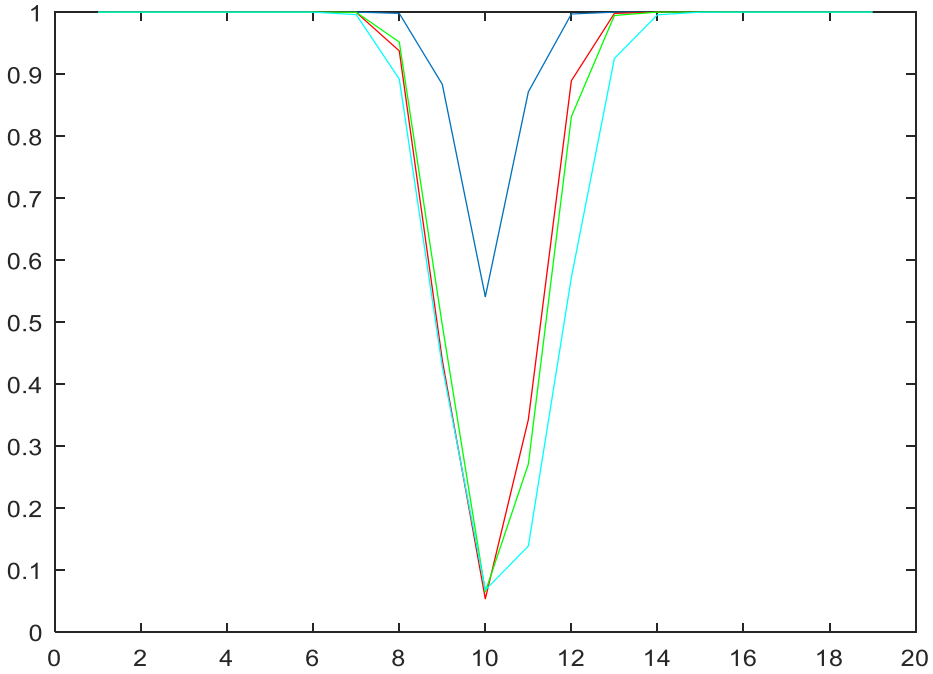
Şekil 3.22. $n = 7$ için güçler



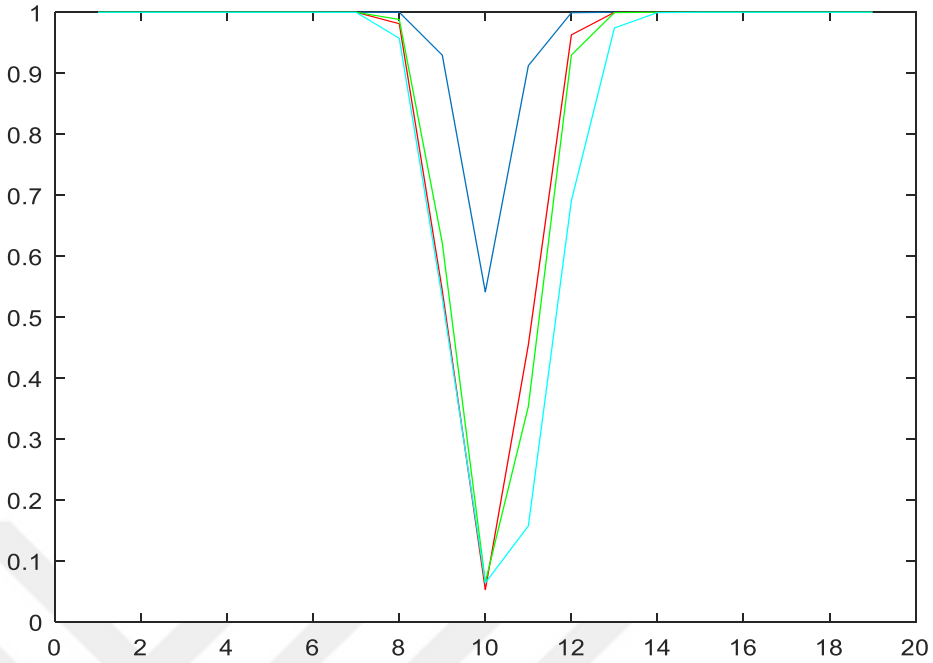
Şekil 3.23. $n = 10$ için güçler



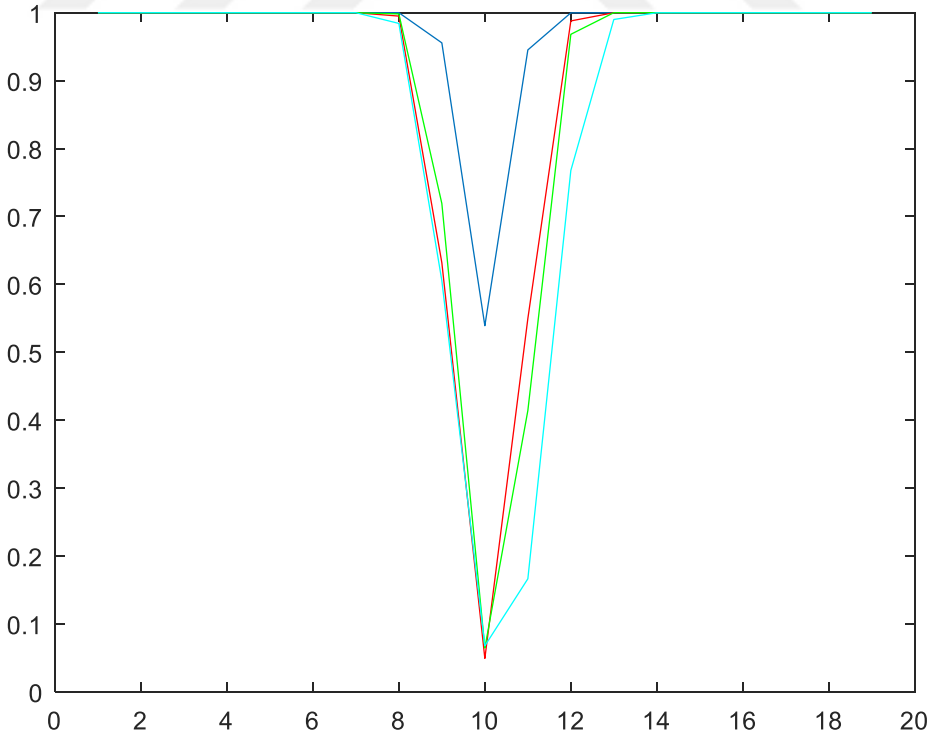
Şekil 3.24. $n = 20$ için güçler



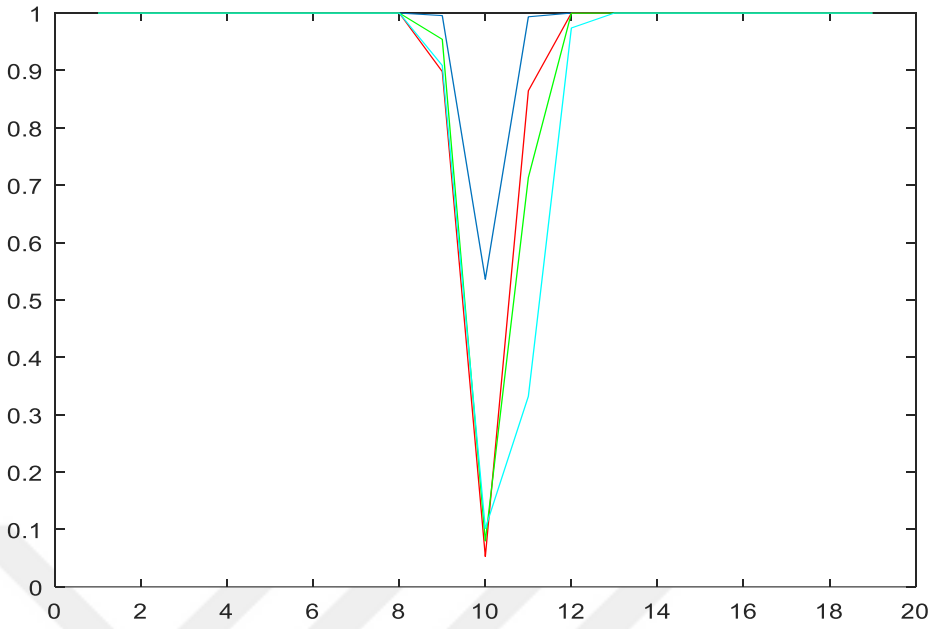
Şekil 3.25. $n = 30$ için güçler



Şekil 3.26. $n = 40$ için güçler



Şekil 3.27. $n = 50$ için güçler



Şekil 3.28. $n = 100$ için güçler

Çizelge 3.4. ve Şekil 3.22 – 3.28'deki testlerin güçleri incelendiğinde Z testinin gücünün yüksek çıktığı ancak anlamlılık seviyesinin de çok yüksek elde edilmiş olduğu görülmektedir. t testi, Wilcoxon işaretli sıra sayıları testi ve işaret testinin güçleri kıyaslandığında ise gücü en düşük testin işaret testi olduğu ve t testi ile Wilcoxon işaretli sıra sayıları testinin güçlerinin birbirine yakın olduğu görülmektedir. Dolayısıyla, gama dağılımı durumunda parametrik t testinin veya parametrik olmayan Wilcoxon işaretli sıra sayıları testinin kullanılması önerilebilir. Örnek hacmi arttıkça tüm testlerin gücü artmaktadır. Büyük n değerlerinde $(n \geq 100)$ ise t testi, Wilcoxon işaretli sıra sayıları testi ve işaret testinin güçlerinin birbirine yaklaştığı ve istenilen testin kullanılabileceği görülmektedir.

3.2. Orana İlişkin Testler

Bu bölümde tek örneklem testlerinden orana ilişkin parametrik Z testi ile parametrik olmayan Binom testi tanıtılmıştır.

3.2.1. Z testi

Söz konusu kategorik değişkenler olduğunda sonucun, kitledeki oranı ile örneklem değerinin aynı olup olmadığı test edilmek için kullanılır.

$X_1, X_2, ..., X_n \sim N\left(\mu, \sigma^2\right)$ ve σ^2'nin bilindiği durumlarda P'nin P_0 gibi teorik bir değere eşit olup olmadığını test etmektedir.

$H_0 : P = P_0$ hipotezi $H_1 : P \neq P_0$ hipotezine karşı test ediliyor ise test istatistiği,

$$Z_{Hesap} = \frac{\hat{p} - P}{\sqrt{PQ / n}} \tag{3.6}$$

şeklinde hesaplanır. Eğer $\left|Z_{Hesap}\right| > Z_{\alpha/2}$ ise ilgili değişkene ilişkin kitle oranı P'nin P_0'dan farklı olduğu sonucuna varılır yani H_0 reddedilir. (3.6) eşitliğinde $\hat{p}$, ilgili değişkene ilişkin örneklem oranı, P, varsayılan kitle oranı, $Q = 1 - P$ ve n, örneklem gözlem sayısıdır.

3.2.2. Binom testi

Tek örneklemli binom testi parametrik olmayan hipotez testleri arasında yer almaktadır. İki kategori olması durumunda hesaplanacak bu test, ilgili verinin tahmin edilen orandan farklı olup olmadığını ölçmek için kullanılmaktadır. Dolayısıyla kitle çok sayıda örnek kapsadığından ilgili oranı ölçmektense kitleden n hacimli örneklem seçilerek parametreye ilişkin hipotez testi ile karar verilmektedir. Testin kullanılabilmesi için gerekli olan varsayımlar,

- Kıyaslanması düşünülen veriler sınıflandırılmış olmalıdır.
- Olası değerlerden sadece bir tanesi veride yer almalıdır.
- Deneyler birbirinden bağımsız ve ilgilenilen sonucun olasılığı deneyden deneye sabittir.

Kitleye ilişkin bilinmeyen oran π'nin π_0 gibi teorik bir değere eşitliği için $H_0 : \pi = \pi_0$ ve $H_1 : \pi \neq \pi_0$ şeklinde hipotez testleri oluşturulmaktadır ve δ_i değişkeni aşağıdaki şekilde tanımlanmaktadır.

$$\delta_i = \begin{cases} 1, & \text{i.gözlem değeri istenilen özellikte ise} \\ 0, & \text{i.gözlem değeri istenilen özellikte değilse} \end{cases}$$

test istatistiği,

$$b_{Hesap} = \sum_{i=1}^{n} \delta_i \tag{3.7}$$

şeklinde tanımlanır. (3.7) eşitliğinde b, binom test istatistiği, δ_i değişken, n ise gözlem sayısıdır. H_0 hipotezini doğru olduğu bilindiğinde b istatistiğinin olasılık fonksiyonu,

$$f(b) = \binom{n}{b} p^b (1-p)^{n-b}, \quad b = 0,1,...,n \tag{3.8}$$

şeklinde hesaplanır ve $b_{Hesap} \leq b_{\alpha/2}$ ise H_0 reddedilir.

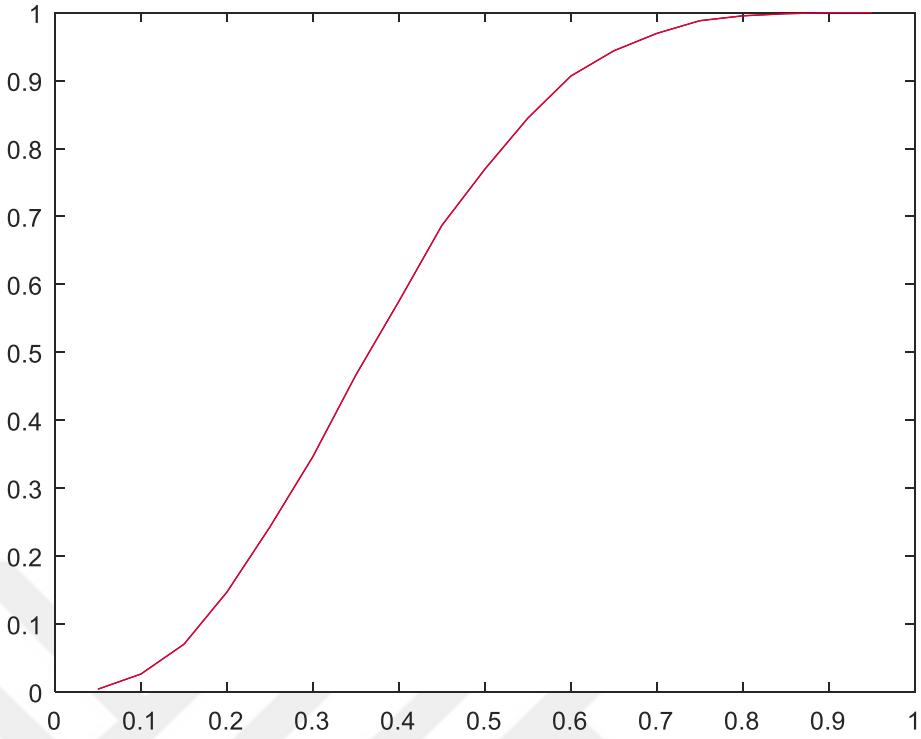
3.2.3. Güç karşılaştırması

Orana ilişkin ele alınan parametrik Z testi ve parametrik olmayan Binom testlerinin güç karşılaştırmalarının yapılması amacıyla 10.000 tekrarlı Monte-Carlo simülasyon yöntemi uygulanmıştır. Testlerin güçleri elde edilirken çift yönlü kurulan farklı hipotezler ve $\alpha = 0.05$ anlamlılık seviyesi ele alınmıştır. Farklı hipotezler, farklı dağılımlar ve farklı n değerleri için elde edilen sonuçlar Çizelge 3.6 – 3.14 ve Şekil 3.36 – 3.98'de verilmiştir. Şekillerde mavi çizgi Binom testinin gücünü, kırmızı çizgi Z testinin gücünü göstermektedir.

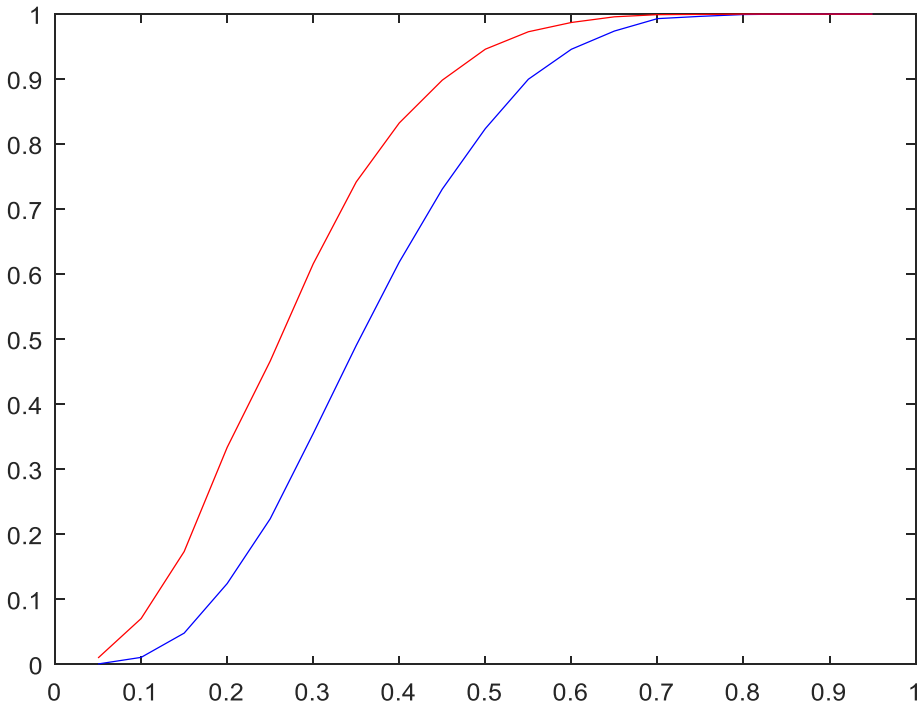
Çizelge 3.6 ve Şekil 3.36 – 3.42'da kitle oranının 0.1'e eşit olup olmadığının testinde, farklı parametre ve farklı n değerleri için simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, 0.1 parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 3.6. $H_0 : P = 0.1$ / $H_1 : P \neq 0.1$ durumu için testlerin güçleri

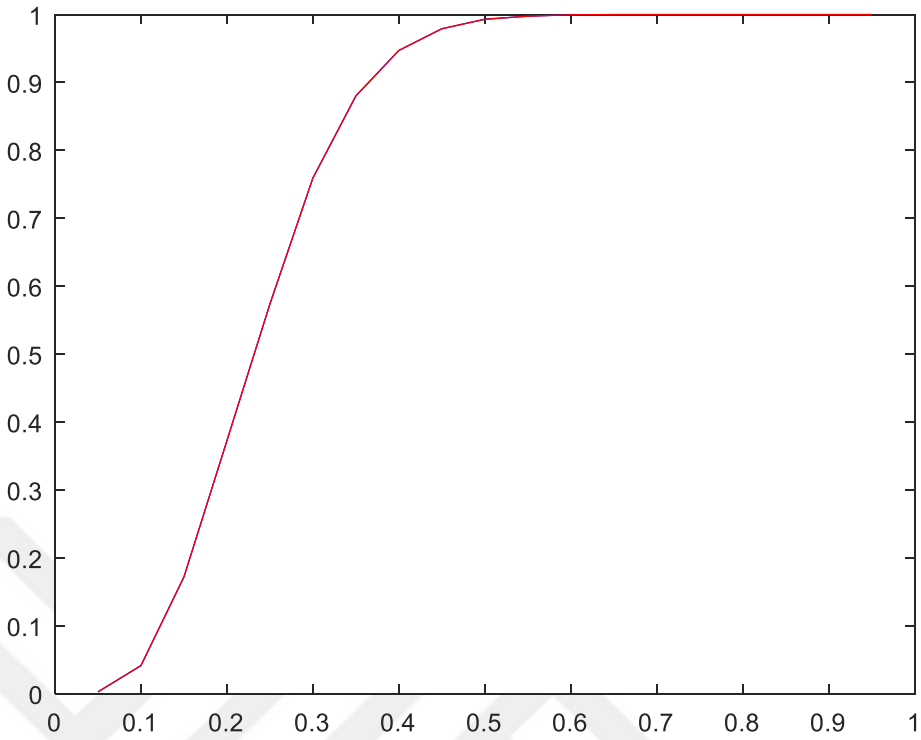
n	Parametre	Binom Testi	Z Testi
7	0.1	0.0264	0.0264
	0.2	0.1470	0.1470
	0.4	0.5751	0.5751
	0.7	0.9695	0.9695
	0.9	0.9999	0.9999
	0.95	1.00	1.00
10	0.1	0.0109	0.0706
	0.2	0.1244	0.3338
	0.3	0.3551	0.6160
	0.4	0.6187	0.8326
	0.5	0.8239	0.9460
	0.6	0.9459	0.9871
	0.7	0.9930	0.9991
	0.8	0.9993	1.00
20	0.1	0.0414	0.0414
	0.2	0.3725	0.3725
	0.4	0.9472	0.9472
	0.6	0.9997	0.9997
	0.65	1.00	1.00
30	0.1	0.0270	0.0270
	0.2	0.3961	0.3961
	0.4	0.9814	0.9814
	0.5	0.9990	0.9990
	0.6	1.00	1.00
40	0.1	0.0275	0.0523
	0.2	0.4007	0.5618
	0.3	0.8879	0.9446
	0.4	0.9934	0.9976
	0.5	0.9999	1.00
	0.6	1.00	1.00
50	0.1	0.0291	0.0291
	0.15	0.2028	0.2028
	0.2	0.5602	0.5602
	0.25	0.8339	0.8339
	0.3	0.9642	0.9642
	0.45	1.00	1.00
100	0.1	0.0453	0.0648
	0.15	0.3314	0.4357
	0.2	0.8146	0.8772
	0.3	0.9994	0.9997
	0.35	1.00	1.00



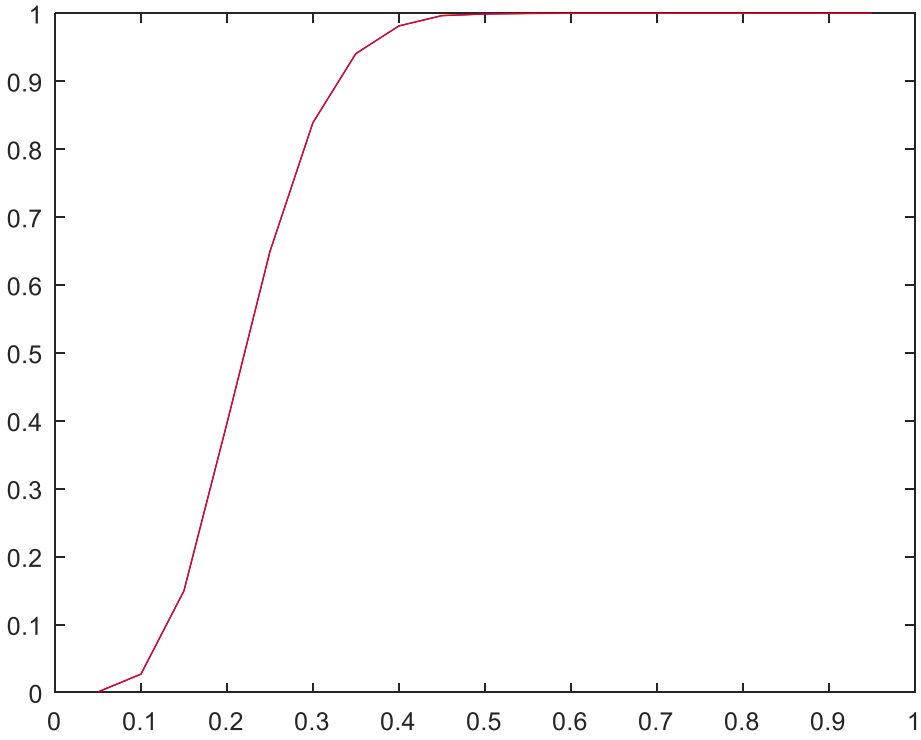
Şekil 3.36. $n = 7$ için güçler



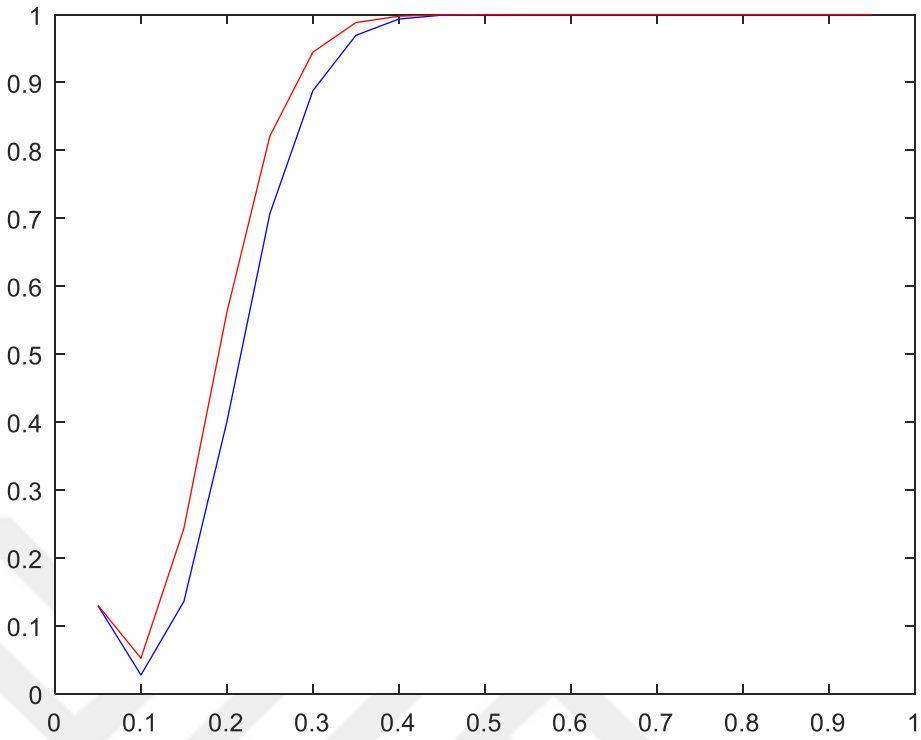
Şekil 3.37. $n = 10$ için güçler



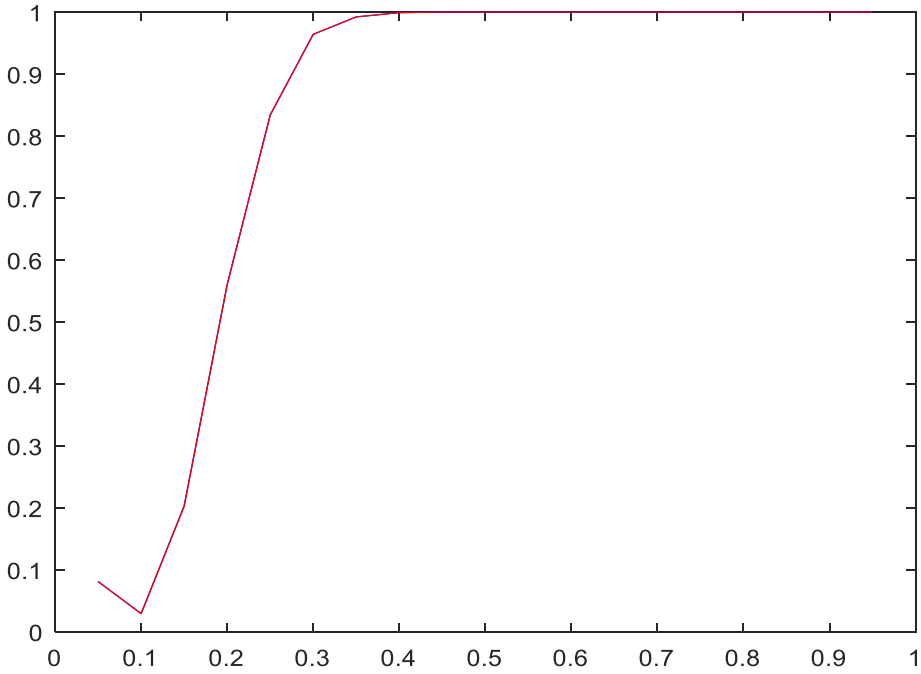
Şekil 3.38. $n = 20$ için güçler



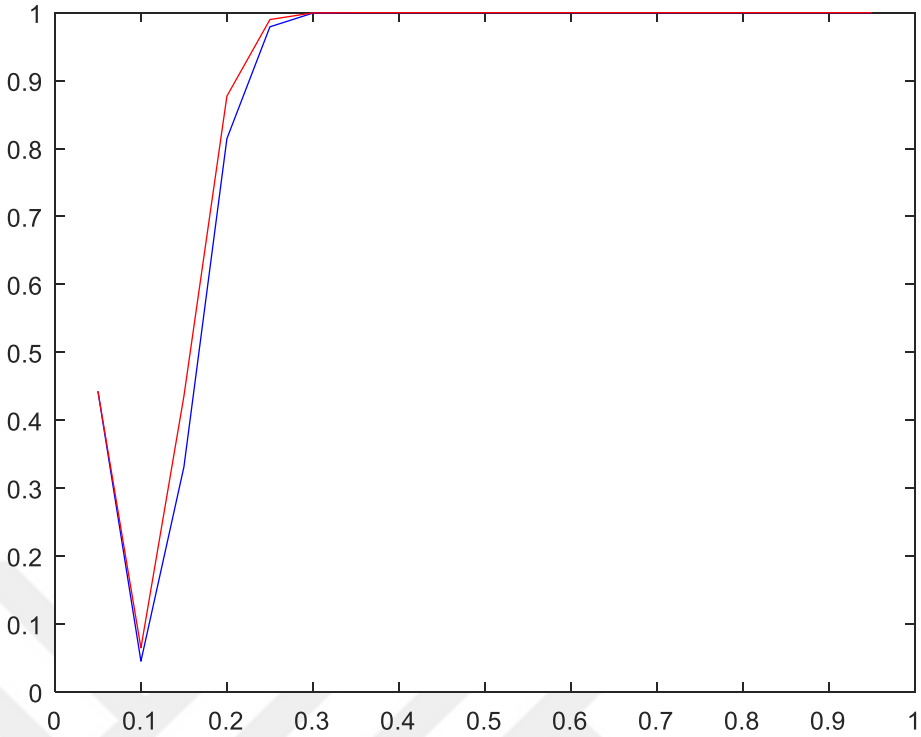
Şekil 3.39. $n = 30$ için güçler



Şekil 3.40. $n = 40$ için güçler



Şekil 3.41. $n = 50$ için güçler



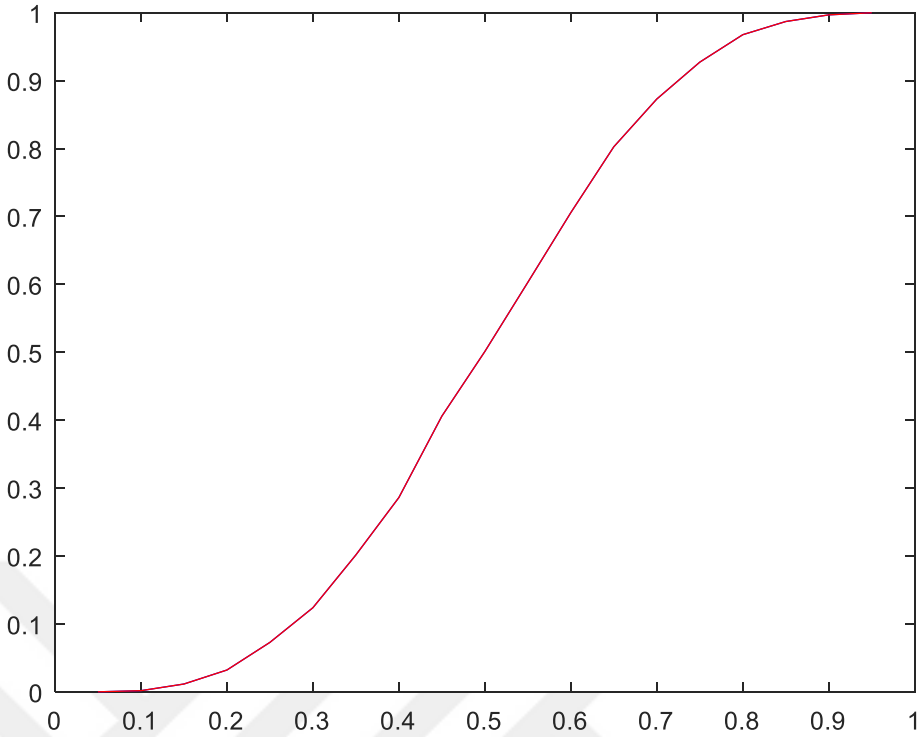
Şekil 3.42. $n = 100$ için güçler

Testin güçleri incelendiğinde bu iki test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test Z testidir. n sayısı arttıkça testlerinin güçlerinin arttığı şekillerden de görülmektedir.

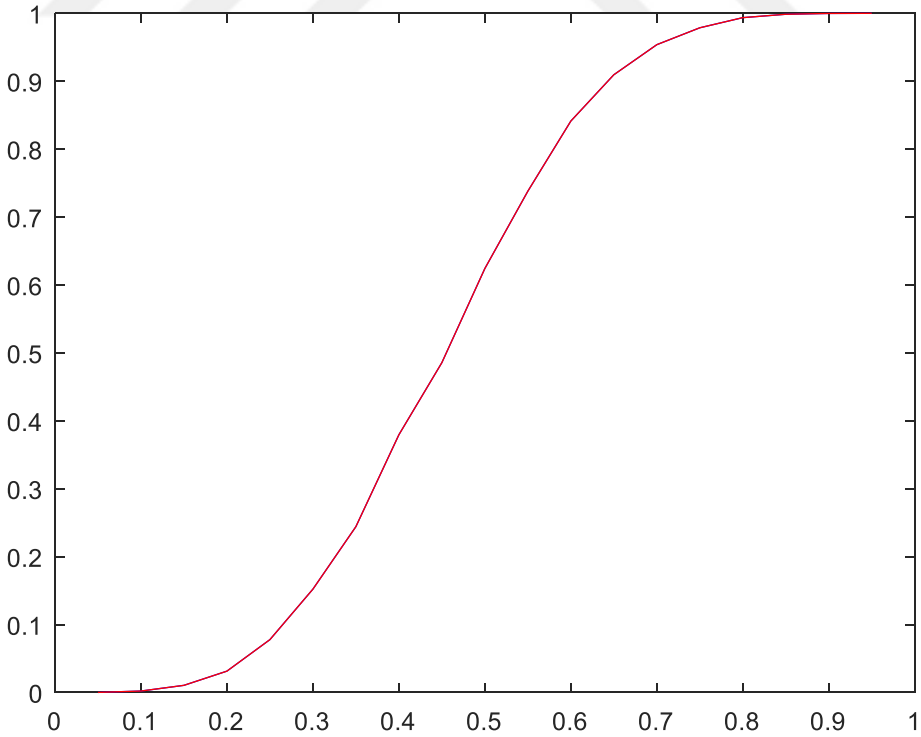
Çizelge 3.7 ve Şekil 3.43 – 3.49'da kitle oranının 0.2'ye eşit olup olmadığının testinde, farklı parametre ve farklı n değerleri için simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, 0.2 parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 3.7. $H_0 : P = 0.2$ / $H_1 : P \neq 0.2$ durumu için testlerin güçleri

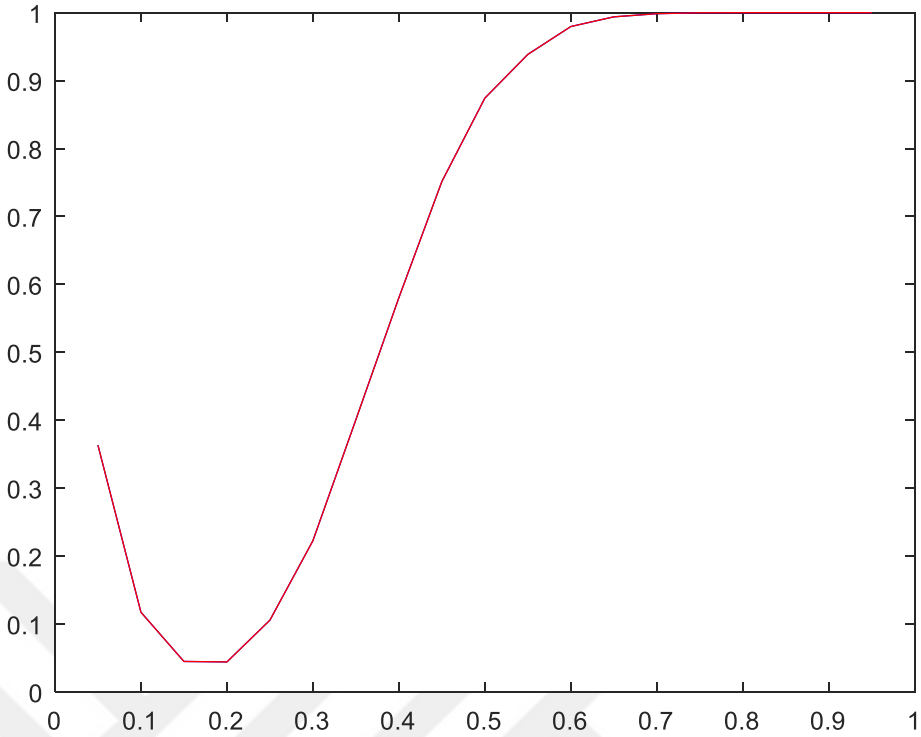
n	Parametre	Binom Testi	Z Testi
7	0.15	0.0117	0.0117
	0.2	0.0323	0.0323
	0.4	0.2862	0.2862
	0.5	0.5012	0.5012
	0.6	0.7057	0.7057
	0.7	0.8729	0.8729
	0.95	0.9999	0.9999
10	0.15	0.0105	0.0105
	0.2	0.0314	0.0314
	0.4	0.3794	0.3794
	0.5	0.6242	0.6242
	0.6	0.8413	0.8413
	0.7	0.9538	0.9538
	0.85	0.9986	0.9986
	0.95	1.00	1.00
20	0.1	0.1174	0.1174
	0.2	0.0442	0.0442
	0.35	0.4008	0.4008
	0.45	0.7514	0.7514
	0.55	0.9386	0.9386
	0.65	0.9939	0.9939
	0.80	1.00	1.00
30	0.1	0.1779	0.1779
	0.2	0.0332	0.0332
	0.35	0.4955	0.4955
	0.4	0.7156	0.7156
	0.5	0.9519	0.9519
	0.60	0.9994	0.9994
	0.75	1.00	1.00
40	0.1	0.4136	0.4136
	0.15	0.1256	0.1282
	0.2	0.05	0.0741
	0.3	0.3038	0.4276
	0.4	0.7876	0.8678
	0.5	0.9822	0.9929
	0.55	0.9963	0.9991
	0.6	1.00	1.00
50	0.15	0.1145	0.1145
	0.2	0.0493	0.0493
	0.3	0.4327	0.4327
	0.4	0.9067	0.9067
	0.5	0.9957	0.9957
	0.6	1.00	1.00
100	0.1	0.8030	0.8030
	0.15	0.2490	0.2495
	0.2	0.0471	0.0607
	0.3	0.62	0.7056
	0.4	0.9914	0.9956
	0.45	0.9997	0.9998
	0.5	1.00	1.00



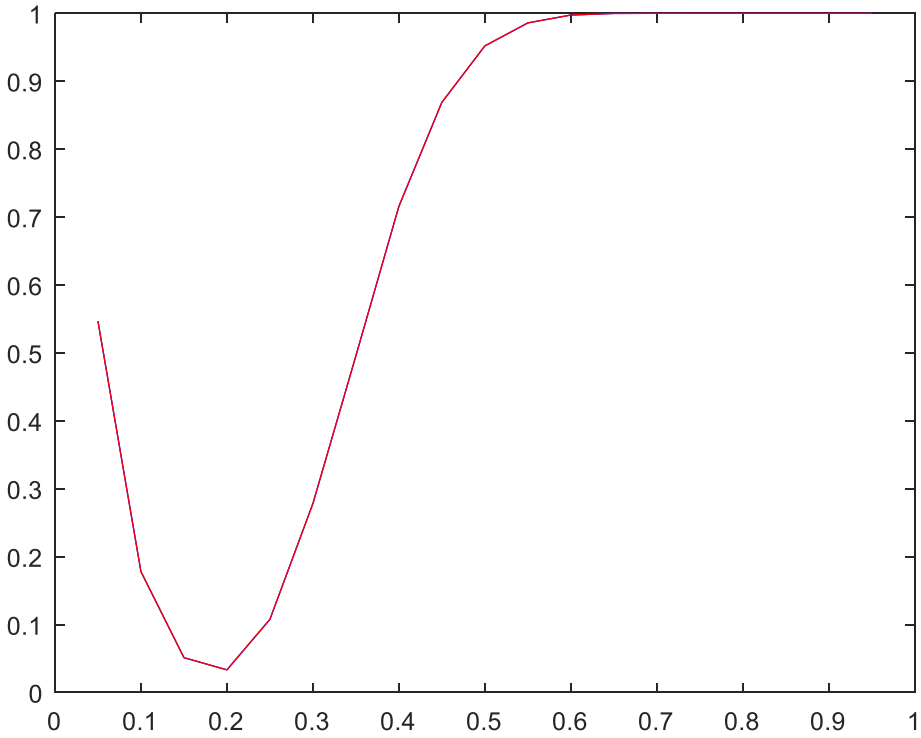
Şekil 3.43. $n = 7$ için güçler



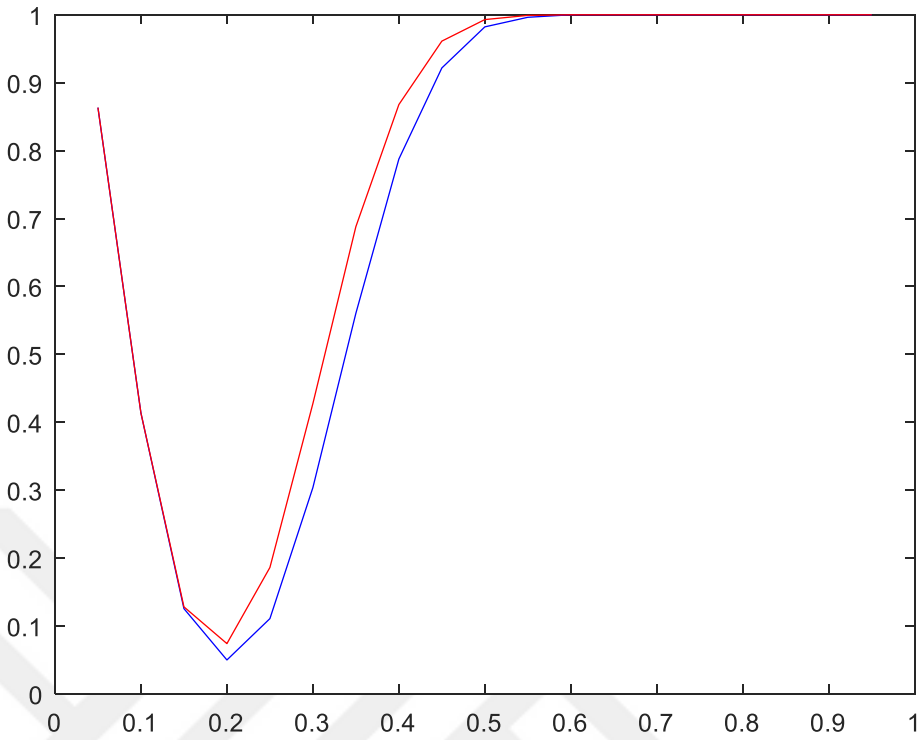
Şekil 3.44. $n = 10$ için güçler



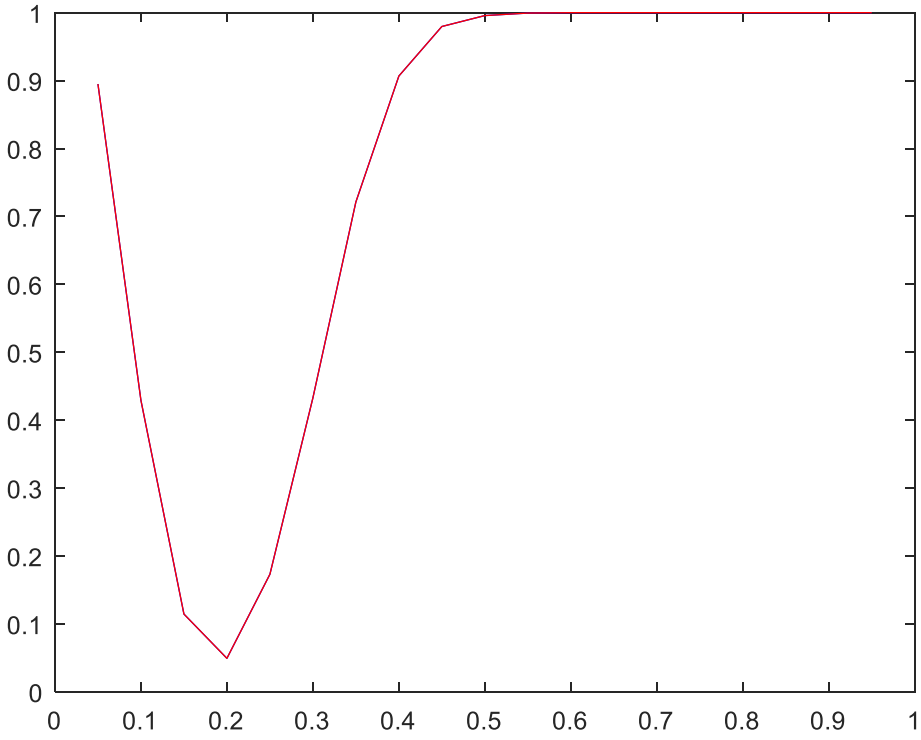
Şekil 3.45. $n = 20$ için güçler



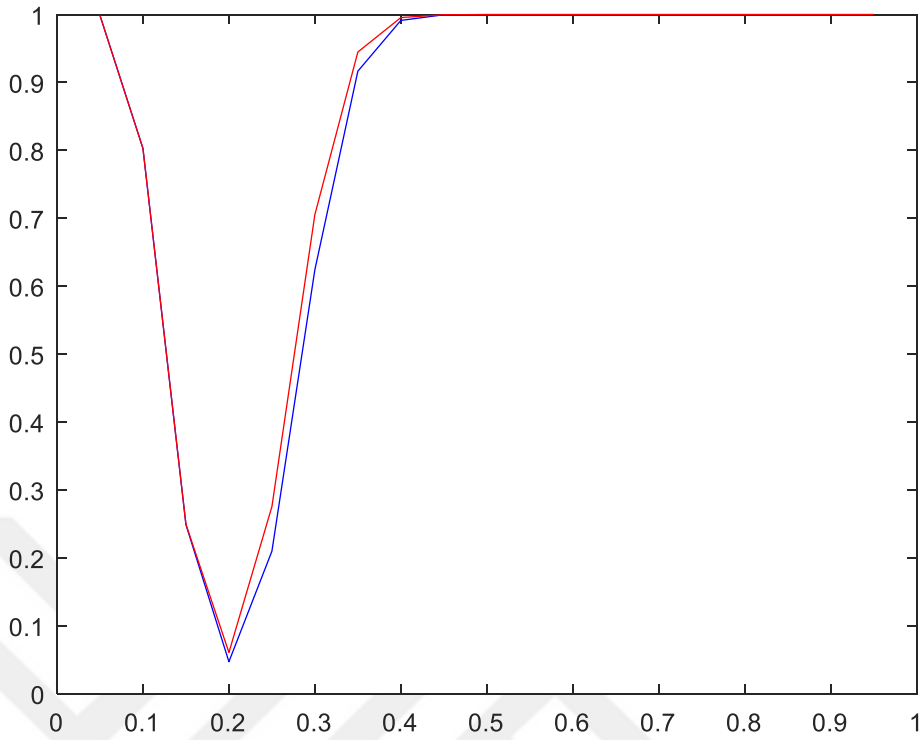
Şekil 3.46. $n = 30$ için güçler



Şekil 3.47. $n = 40$ için güçler



Şekil 3.48. $n = 50$ için güçler



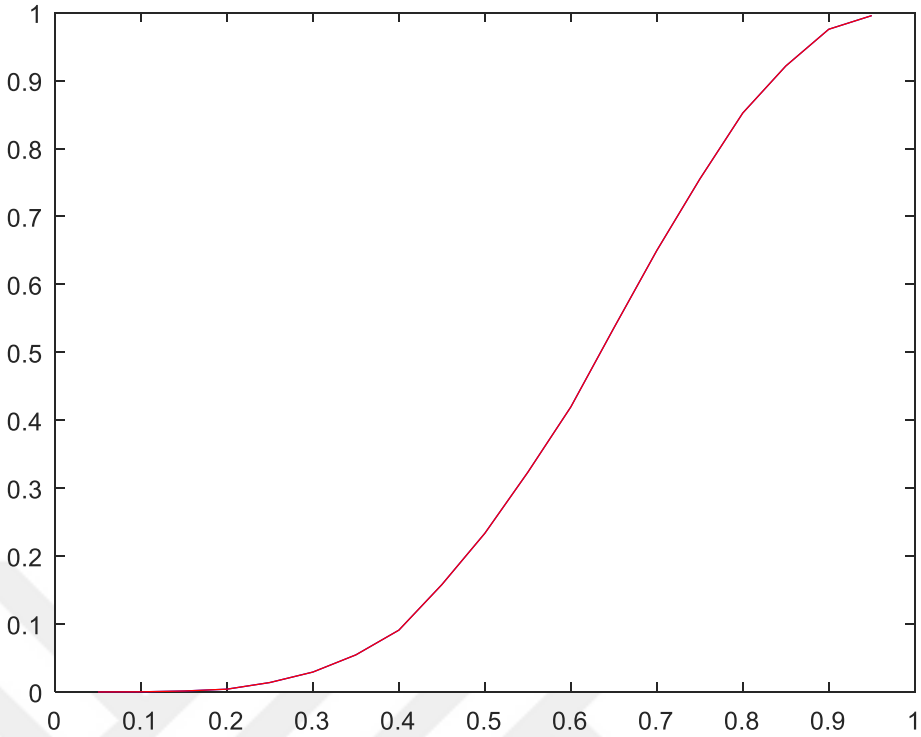
Şekil 3.49. $n = 100$ için güçler

Küçük n değerlerinde testlerin gücü aynı olmakla beraber n arttıkça Z testinin gücünün Binom testinin gücünden daha yüksek olduğu görülmektedir. n sayısı arttıkça testlerinin güçlerinin arttığı şekillerden de görülmektedir.

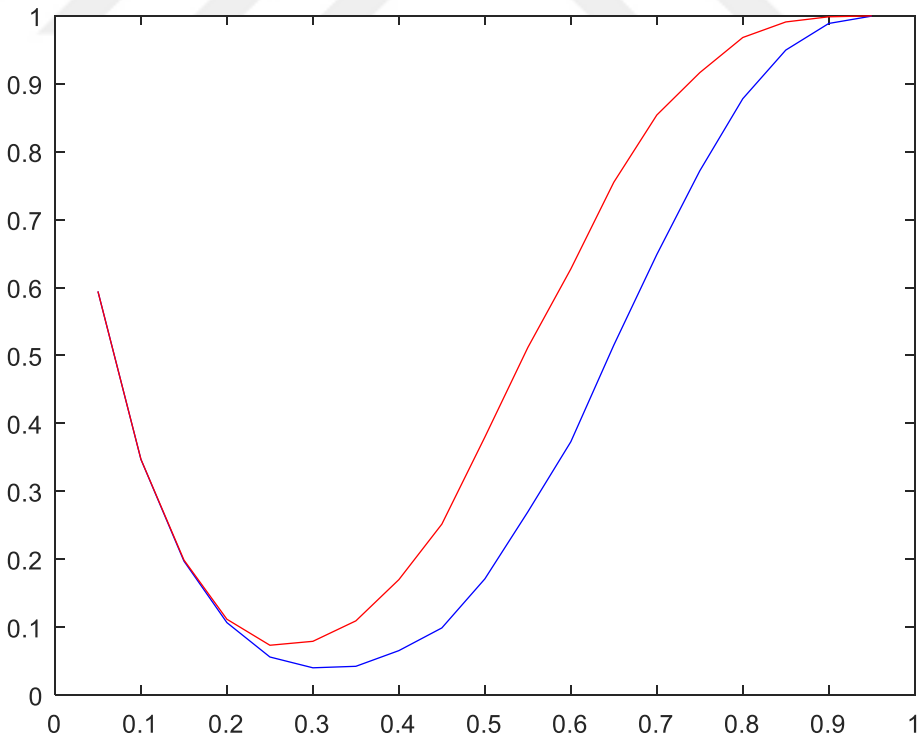
Çizelge 3.8 ve Şekil 3.50 – 3.56'da kitle oranının 0.3'e eşit olup olmadığının testinde, farklı parametre ve farklı n değerleri için simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, 0.3 parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 3.8. $H_0 : P = 0.3$ / $H_1 : P \neq 0.3$ durumu için testlerin güçleri

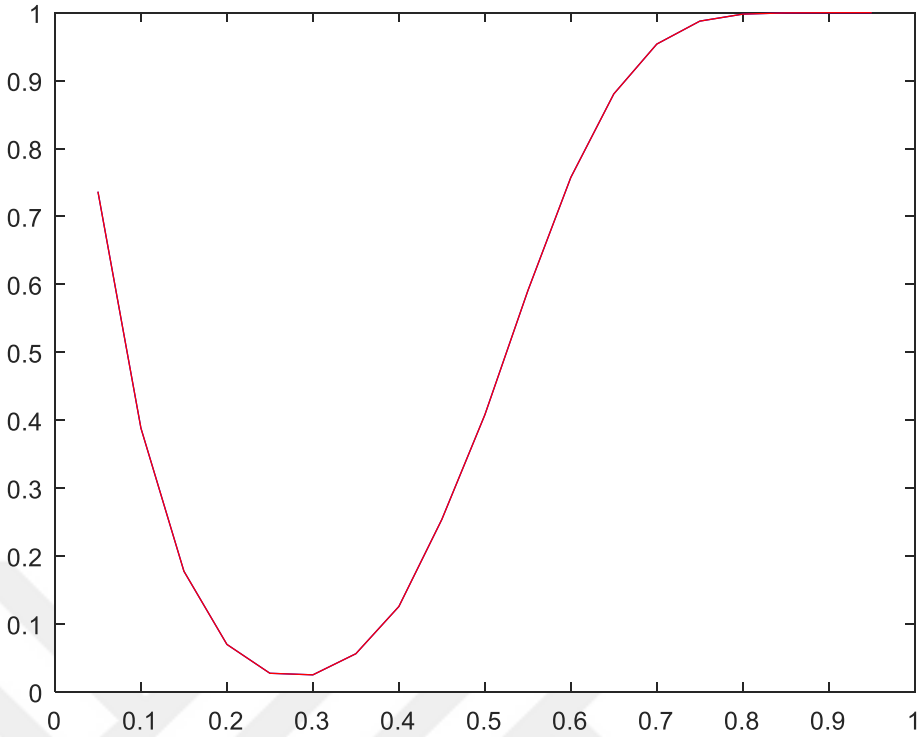
n	Parametre	Binom Testi	Z Testi
7	0.2	0.0041	0.0041
	0.3	0.0293	0.0293
	0.45	0.1580	0.1580
	0.6	0.4196	0.4196
	0.7	0.6499	0.6499
	0.8	0.8523	0.8523
	0.95	0.9957	0.9957
10	0.1	0.3470	0.3470
	0.2	0.1067	0.1117
	0.3	0.0402	0.0792
	0.4	0.0655	0.17
	0.5	0.1711	0.3798
	0.6	0.3730	0.6274
	0.7	0.6486	0.8540
	0.8	0.8780	0.9681
	0.95	0.9996	1.00
20	0.1	0.3884	0.3884
	0.2	0.07	0.07
	0.3	0.0252	0.0252
	0.45	0.2540	0.2540
	0.6	0.7575	0.7575
	0.75	0.9876	0.9876
	0.9	1.00	1.00
30	0.2	0.2567	0.2575
	0.3	0.0491	0.0717
	0.4	0.1818	0.2906
	0.5	0.5701	0.7080
	0.6	0.9017	0.9505
	0.7	0.9944	0.9987
	0.8	0.9999	1.00
40	0.2	0.2821	0.2824
	0.3	0.04	0.0561
	0.4	0.2173	0.3253
	0.5	0.6840	0.7882
	0.6	0.9615	0.9812
	0.7	0.9986	0.9996
	0.75	0.9997	0.9999
	0.8	1.00	1.00
50	0.1	0.9382	0.9382
	0.2	0.3065	0.3065
	0.3	0.0452	0.0452
	0.45	0.6125	0.6125
	0.55	0.9562	0.9562
	0.7	1.00	1.00
100	0.2	0.6515	0.6516
	0.3	0.0520	0.0653
	0.4	0.5307	0.6087
	0.45	0.8708	0.9083
	0.5	0.9811	0.9897
	0.55	0.9990	0.9996
	0.6	1.00	1.00



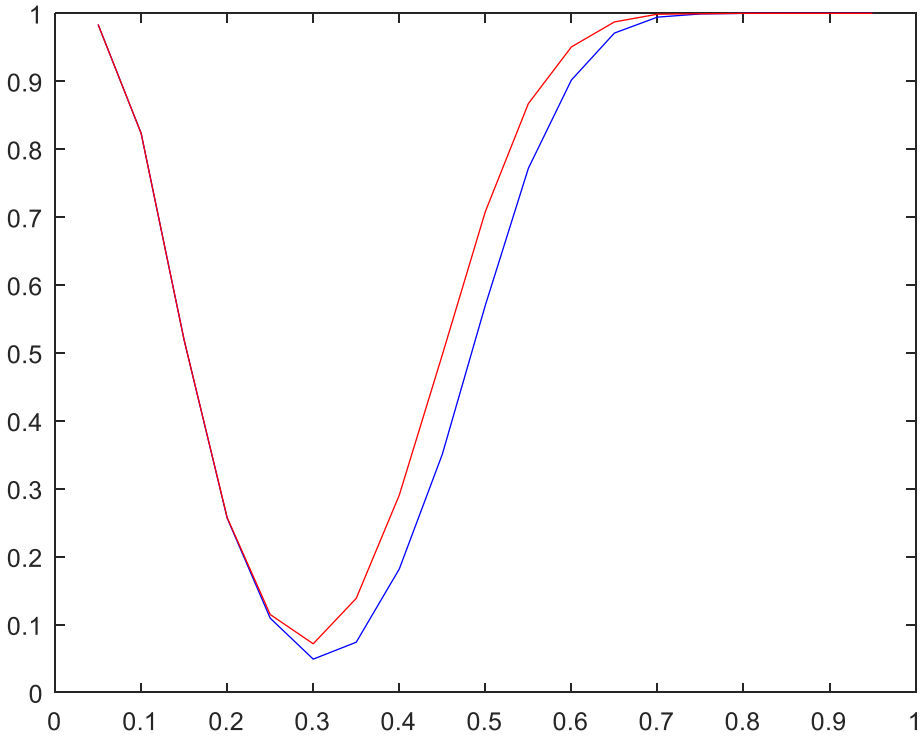
Şekil 3.50. $n = 7$ için güçler



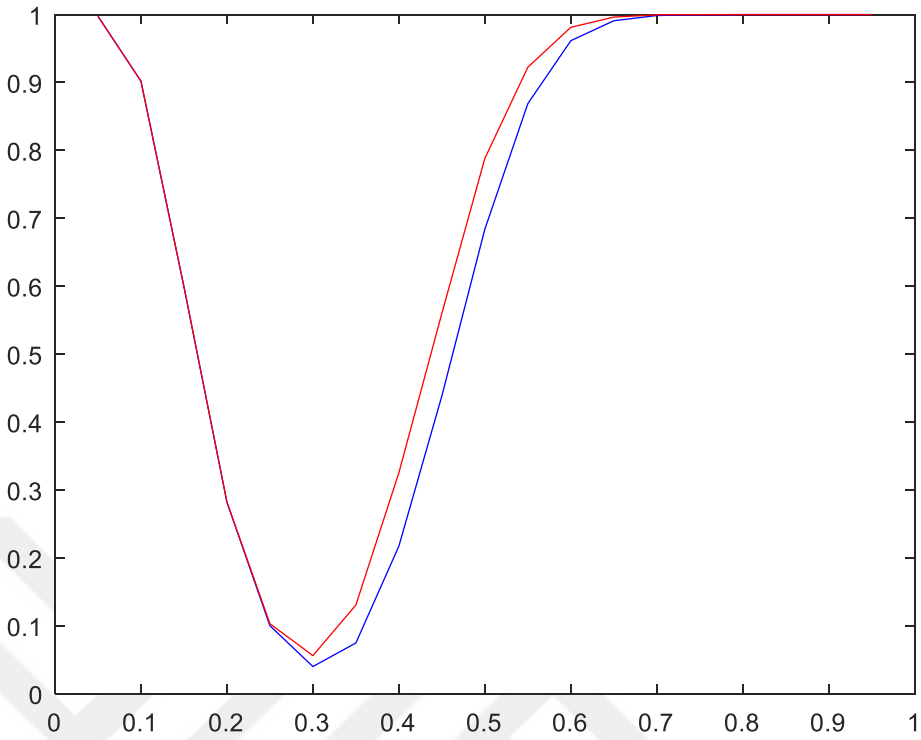
Şekil 3.51. $n = 10$ için güçler



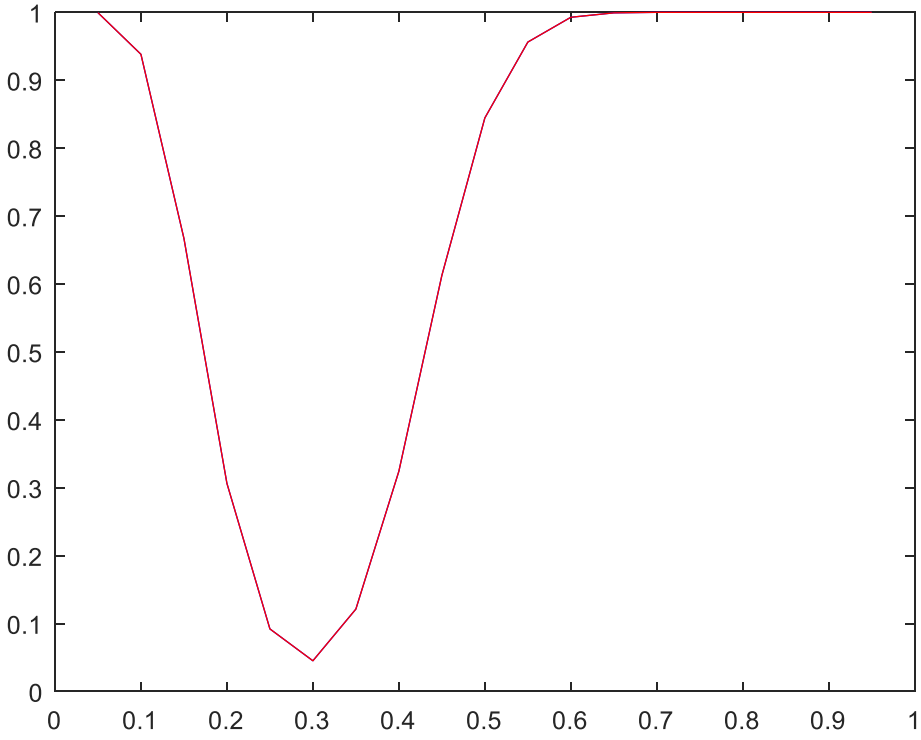
Şekil 3.52. $n = 20$ için güçler



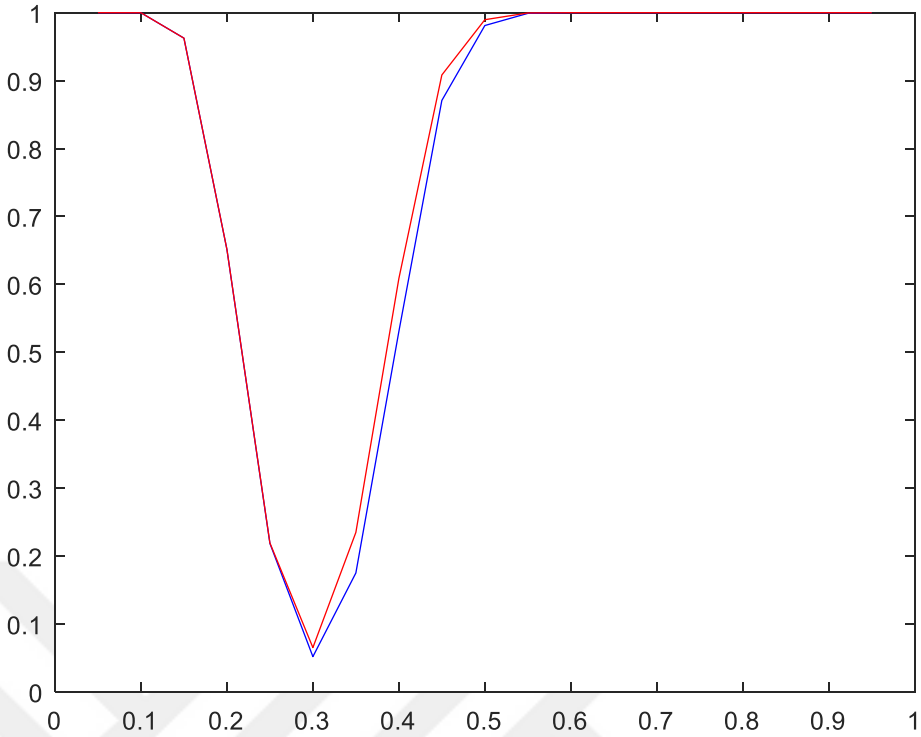
Şekil 3.53. $n = 30$ için güçler



Şekil 3.54. $n = 40$ için güçler



Şekil 3.55. $n = 50$ için güçler



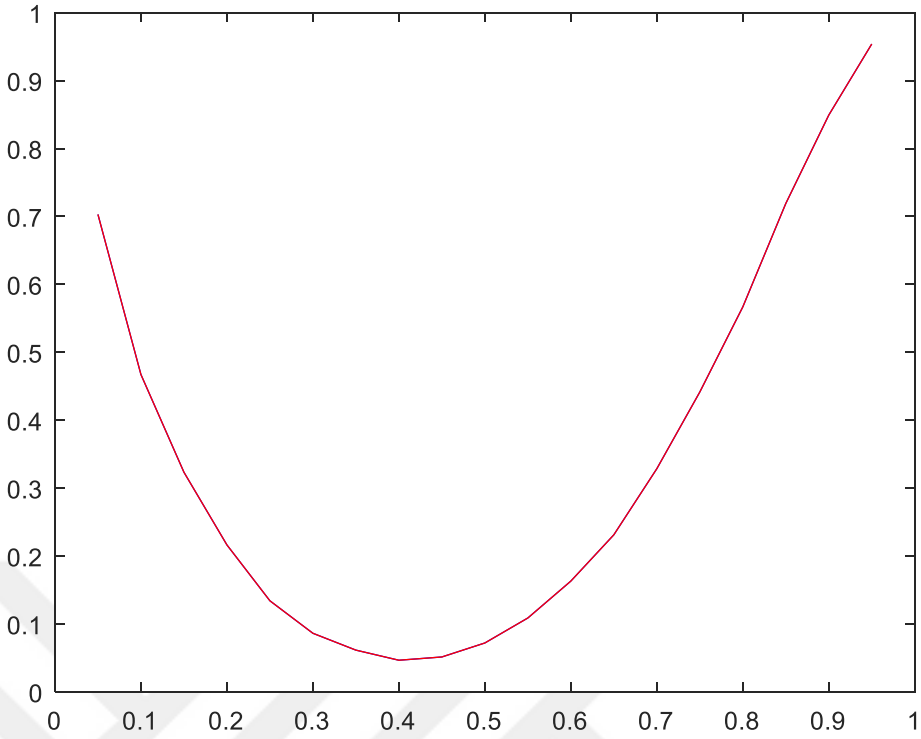
Şekil 3.56. $n = 100$ için güçler

Testin güçleri incelendiğinde bu iki test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test Z testidir. n sayısı arttıkça testlerinin güçlerinin arttığı şekillerden de görülmektedir.

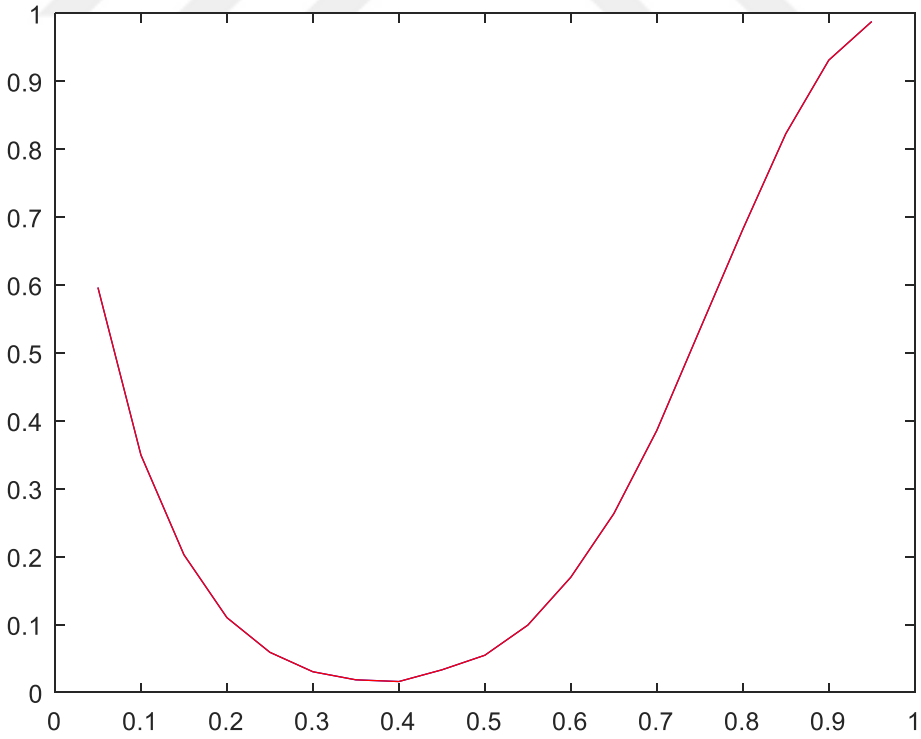
Çizelge 3.9 ve Şekil 3.57 – 3.63'te kitle oranının 0.4'e eşit olup olmadığının testinde farklı parametre ve farklı n değerleri için simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, 0.4 parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 3.9. $H_0 : P = 0.4$ / $H_1 : P \neq 0.4$ durumu için testlerin güçleri

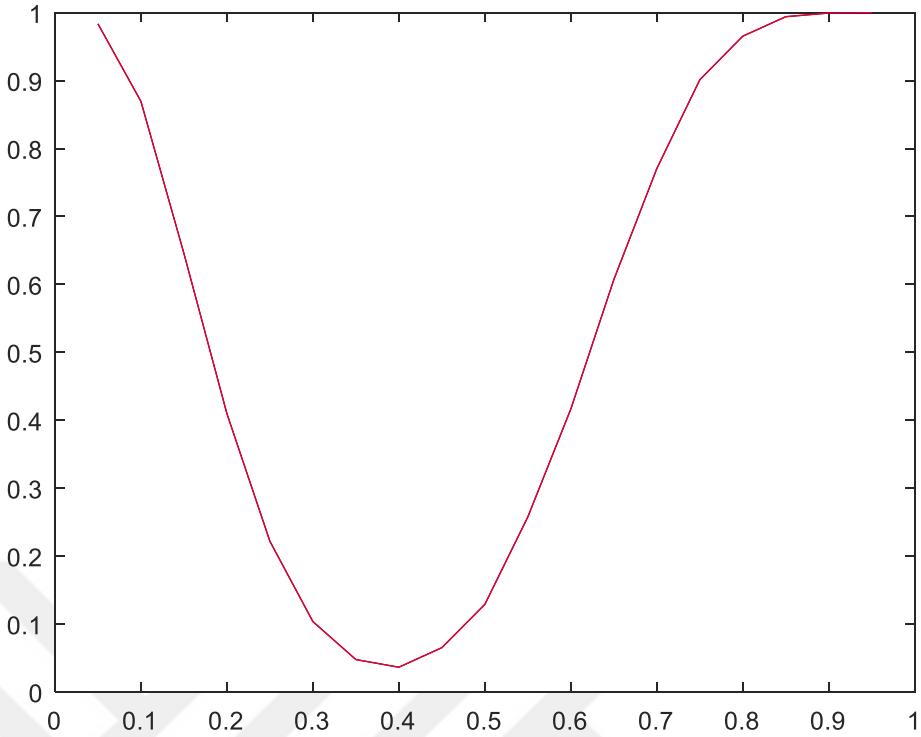
n	Parametre	Binom Testi	Z Testi
7	0.1	0.4673	0.4673
	0.2	0.2165	0.2165
	0.3	0.0864	0.0864
	0.4	0.0467	0.0467
	0.6	0.1632	0.1632
	0.75	0.4416	0.4416
	0.85	0.7191	0.7191
	0.95	0.9539	0.9539
10	0.1	0.3496	0.3496
	0.2	0.1102	0.1102
	0.3	0.0305	0.0305
	0.4	0.0161	0.0161
	0.65	0.2632	0.2632
	0.75	0.5342	0.5342
	0.85	0.8227	0.8227
	0.95	0.9879	0.9879
20	0.1	0.8694	0.8694
	0.2	0.4098	0.4098
	0.3	0.1036	0.1036
	0.4	0.0364	0.0364
	0.55	0.2578	0.2578
	0.65	0.6069	0.6069
	0.80	0.9654	0.9654
	0.95	1.00	1.00
30	0.1	0.9741	0.9741
	0.2	0.6029	0.6029
	0.3	0.1574	0.1574
	0.4	0.0351	0.0351
	0.55	0.3606	0.3606
	0.65	0.7750	0.7750
	0.75	0.9786	0.9786
	0.85	1.00	1.00
40	0.1	0.9949	0.9949
	0.2	0.7231	0.7231
	0.3	0.1940	0.1940
	0.4	0.0328	0.0328
	0.55	0.4424	0.4424
	0.65	0.8727	0.8727
	0.85	1.00	1.00
50	0.25	0.6349	0.6349
	0.3	0.3266	0.3267
	0.4	0.0446	0.0591
	0.5	0.2354	0.3304
	0.6	0.7590	0.8406
	0.7	0.9886	0.9957
	0.8	1.00	1.00
100	0.3	0.5483	0.5483
	0.35	0.1741	0.1751
	0.4	0.0449	0.0556
	0.5	0.4642	0.5410
	0.6	0.9736	0.9832
	0.7	0.9999	0.9999
	0.8	1.00	1.00



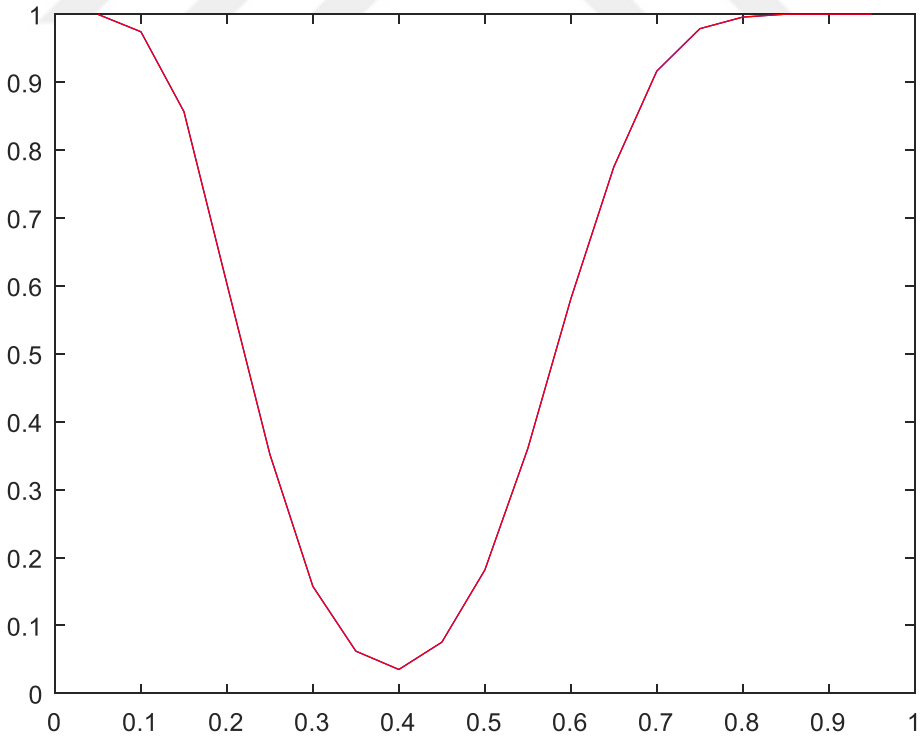
Şekil 3.57. $n = 7$ için güçler



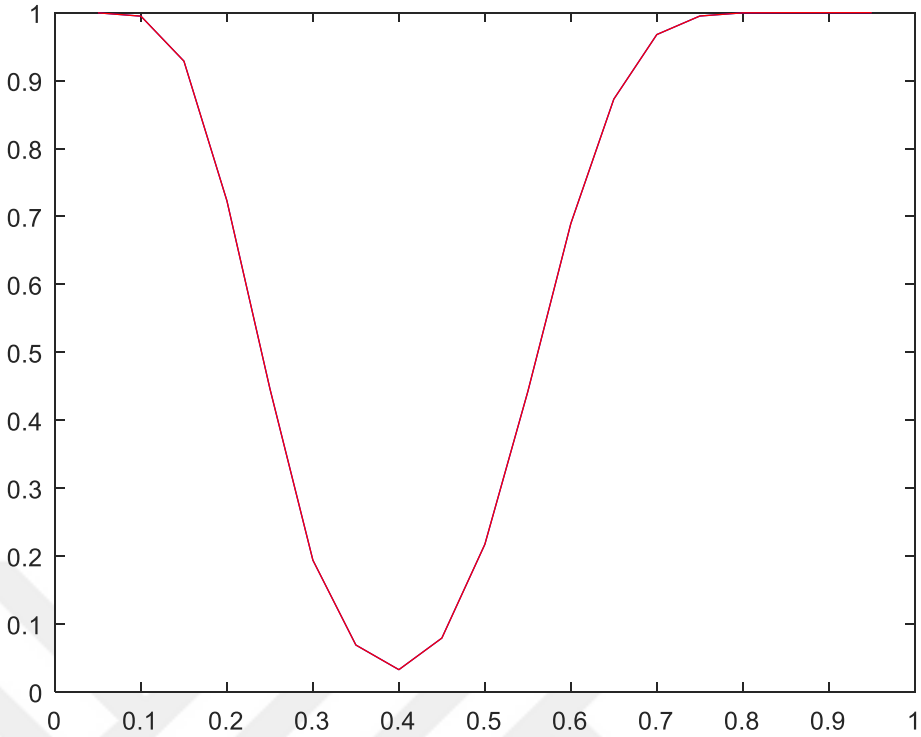
Şekil 3.58. $n = 10$ için güçler



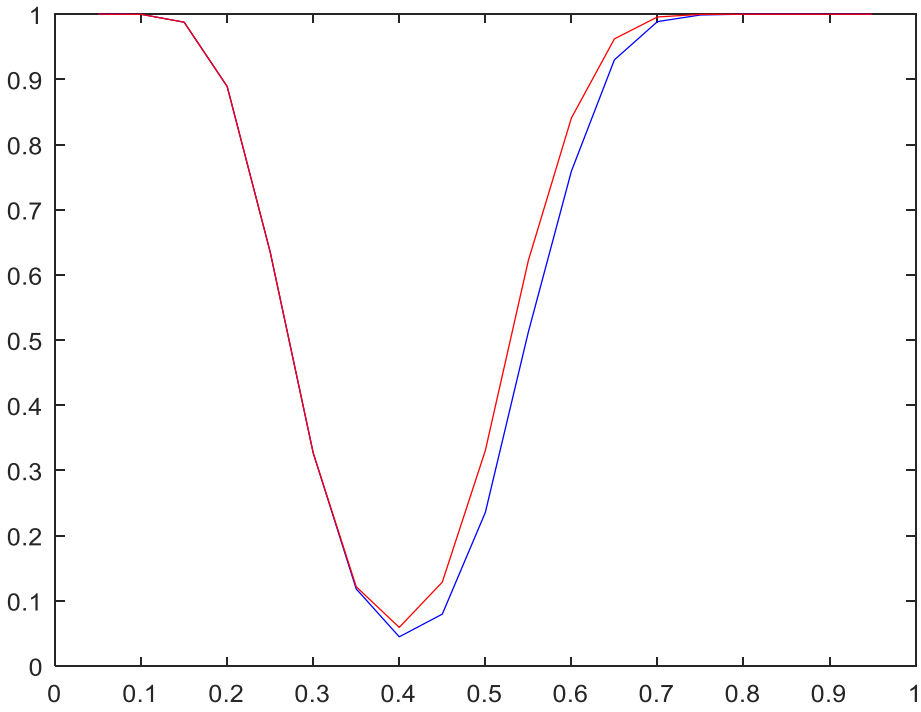
Şekil 3.59. $n = 20$ için güçler



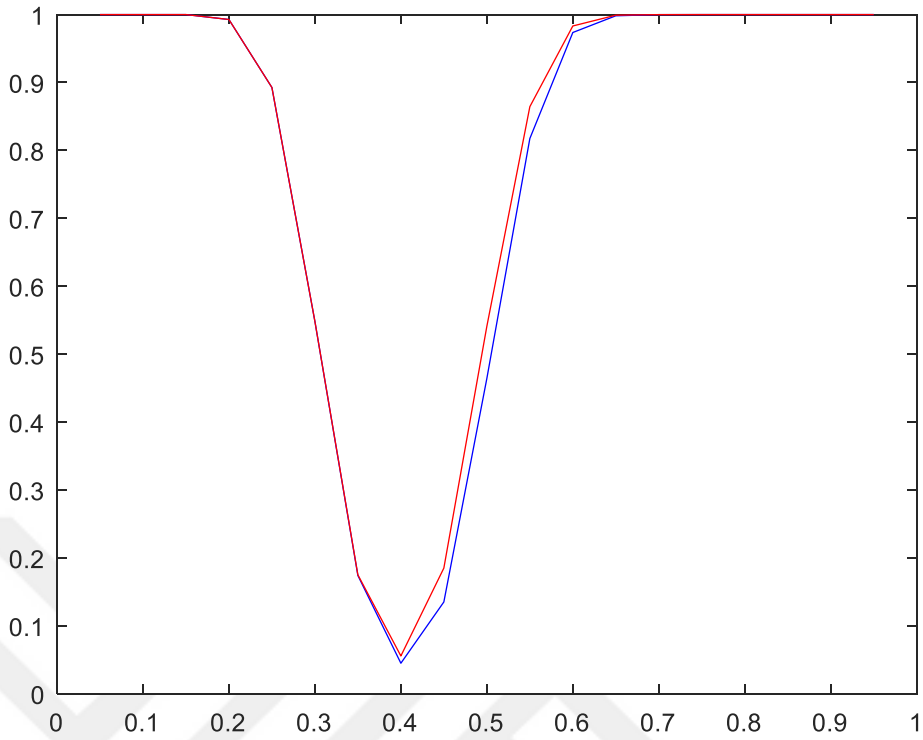
Şekil 3.60. $n = 30$ için güçler



Şekil 3.61. $n = 40$ için güçler



Şekil 3.62. $n = 50$ için güçler



Şekil 3.63. $n = 100$ için güçler

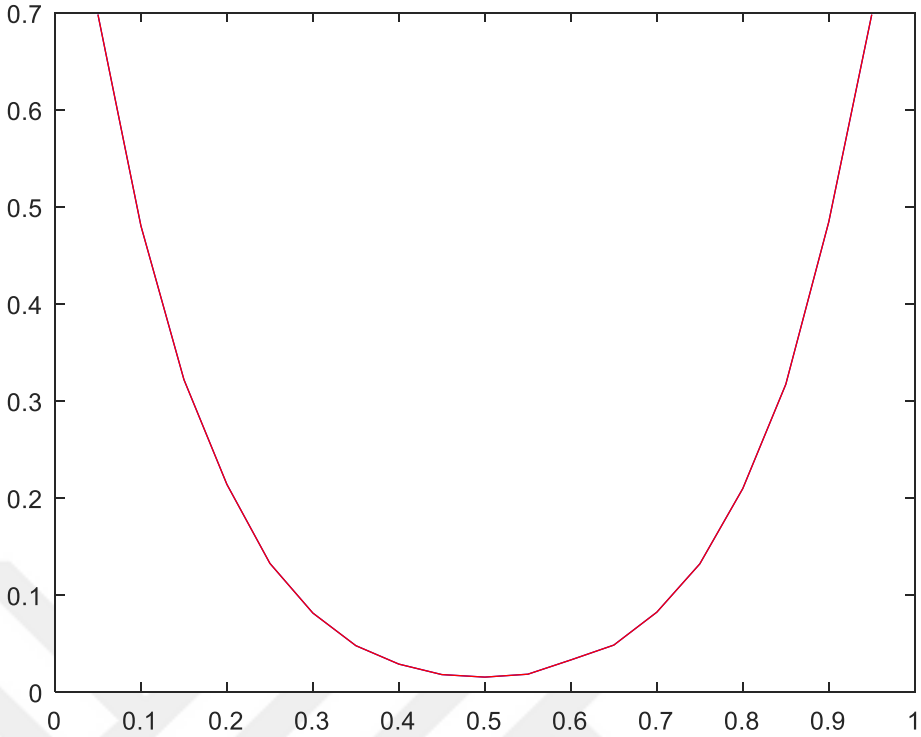
Küçük n değerlerinde testlerin gücü aynı olmakla beraber n arttıkça Z testinin gücünün Binom testinin gücünden daha yüksek olduğu görülmektedir. n sayısı arttıkça testlerinin güçlerinin arttığı şekillerden de görülmektedir.

Çizelge 3.10 ve Şekil 3.64 – 3.70'te kitle oranının 0.5'e eşit olup olmadığının testinde farklı parametre ve farklı n değerleri için simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, 0.5 parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

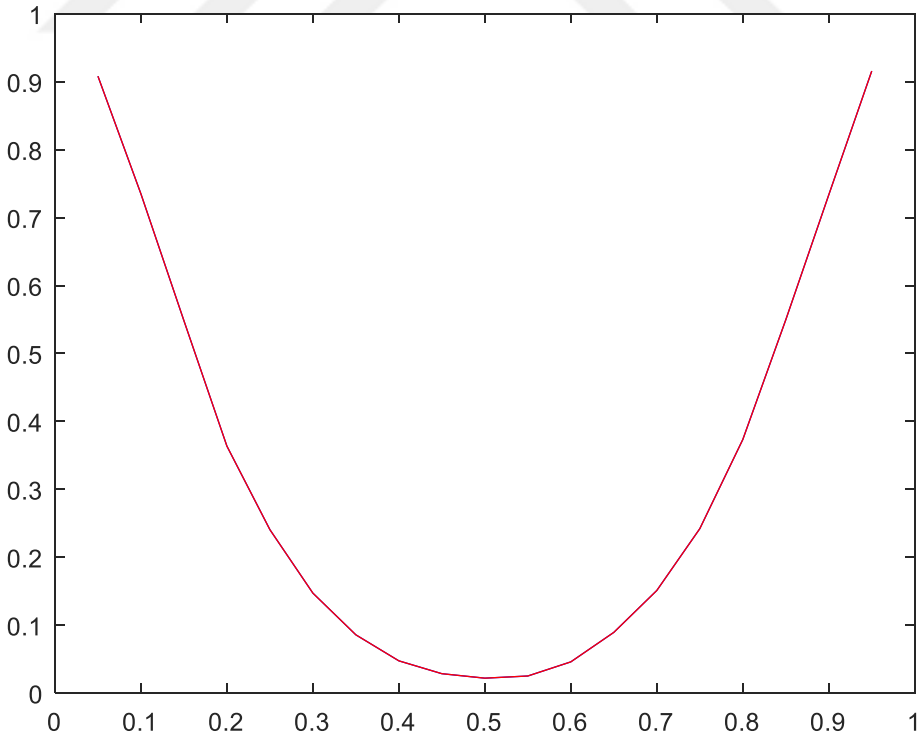
Çizelge 3.10. $H_0 : P = 0.5$ / $H_1 : P \neq 0.5$ durumu için testlerin güçleri

n	Parametre	Binom Testi	Z Testi
7	0.1	0.4803	0.4803
	0.2	0.2139	0.2139
	0.3	0.0813	0.0813
	0.4	0.0287	0.0287
	0.5	0.0153	0.0153
	0.7	0.0822	0.0822
	0.85	0.3170	0.3170

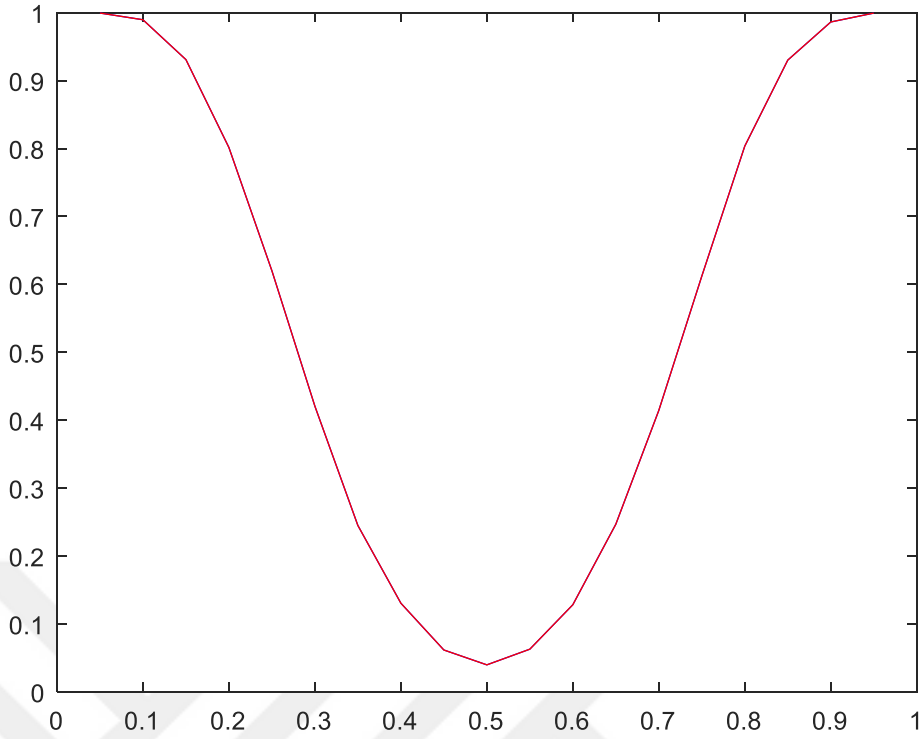
	0.95	0.6980	0.6980
10	0.1	0.7349	0.7349
	0.2	0.3636	0.3636
	0.35	0.0858	0.0858
	0.5	0.0220	0.0220
	0.7	0.1511	0.1511
	0.85	0.5496	0.5496
	0.95	0.9158	0.9158
20	0.1	0.9896	0.9896
	0.25	0.62	0.62
	0.4	0.1307	0.1307
	0.5	0.0399	0.0399
	0.7	0.4145	0.4145
	0.8	0.8036	0.8036
	0.95	0.9994	0.9994
30	0.1	0.9994	0.9994
	0.2	0.9431	0.9431
	0.3	0.5849	0.5849
	0.4	0.1762	0.1762
	0.5	0.0406	0.0406
	0.65	0.3592	0.3592
	0.8	0.9369	0.9369
	0.95	1.00	1.00
40	0.1	1.00	1.00
	0.25	0.8958	0.8958
	0.4	0.2158	0.2158
	0.5	0.0369	0.0369
	0.65	0.4321	0.4321
	0.75	0.9033	0.9033
	0.80	0.9828	0.9828
	0.90	1.00	1.00
50	0.15	0.9998	0.9999
	0.2	0.9937	0.9973
	0.3	0.7882	0.8664
	0.4	0.2410	0.3434
	0.5	0.0345	0.0672
	0.6	0.2326	0.3315
	0.7	0.7842	0.8620
	0.8	0.9928	0.9977
	0.85	0.9998	1.00
100	0.25	0.9997	0.9999
	0.3	0.9795	0.9856
	0.4	0.4637	0.5443
	0.5	0.0371	0.0597
	0.6	0.4583	0.5396
	0.7	0.9764	0.9859
	0.75	0.9994	0.9999
	0.8	1.00	1.00



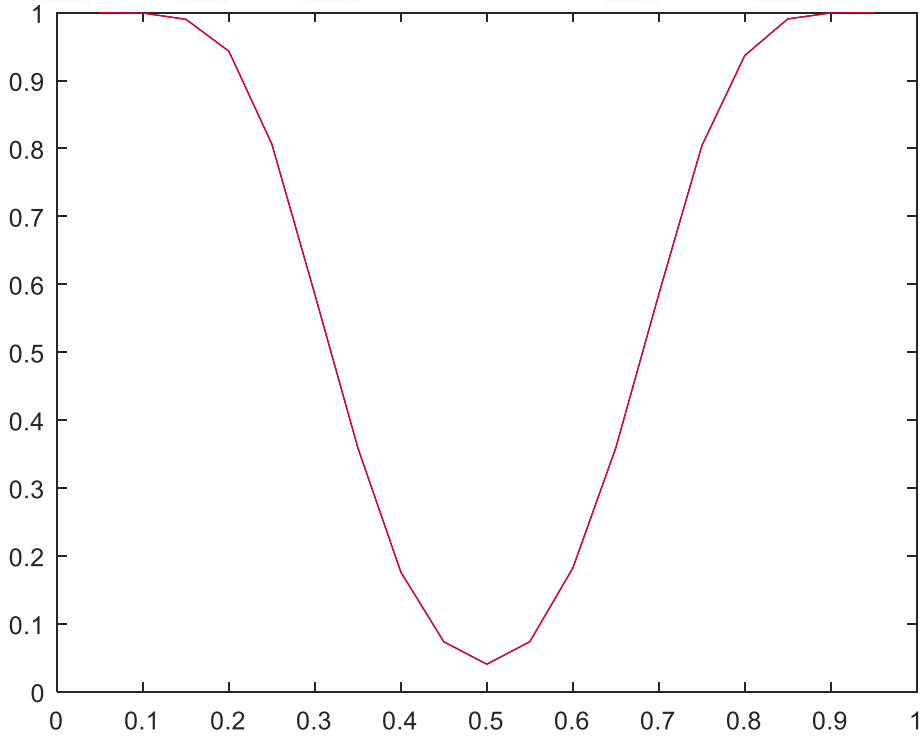
Şekil 3.64. $n = 7$ için güçler



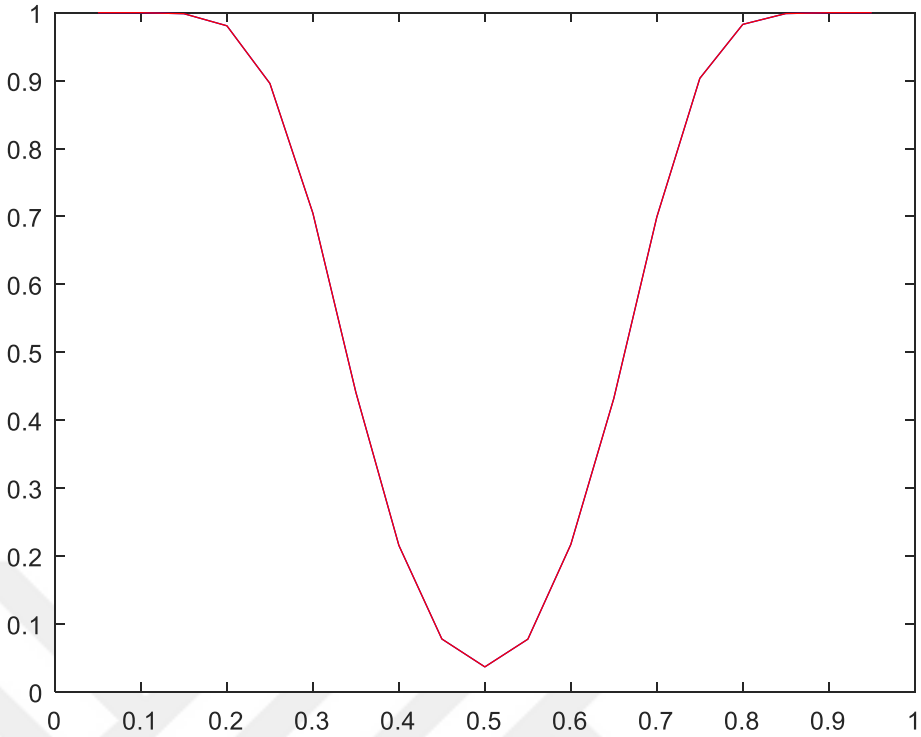
Şekil 3.65. $n = 10$ için güçler



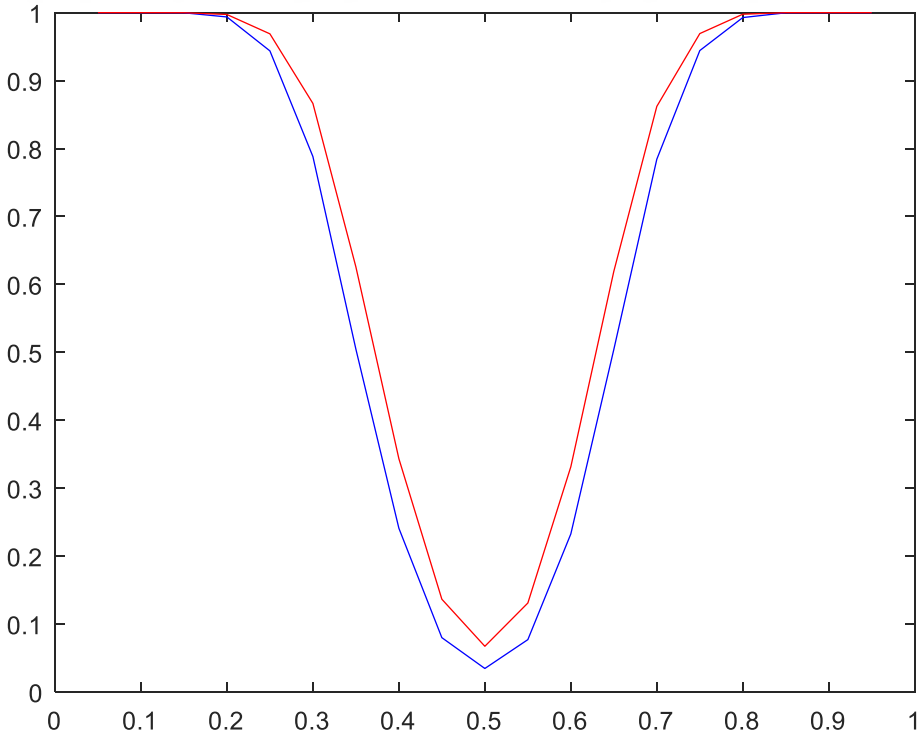
Şekil 3.66. $n = 20$ için güçler



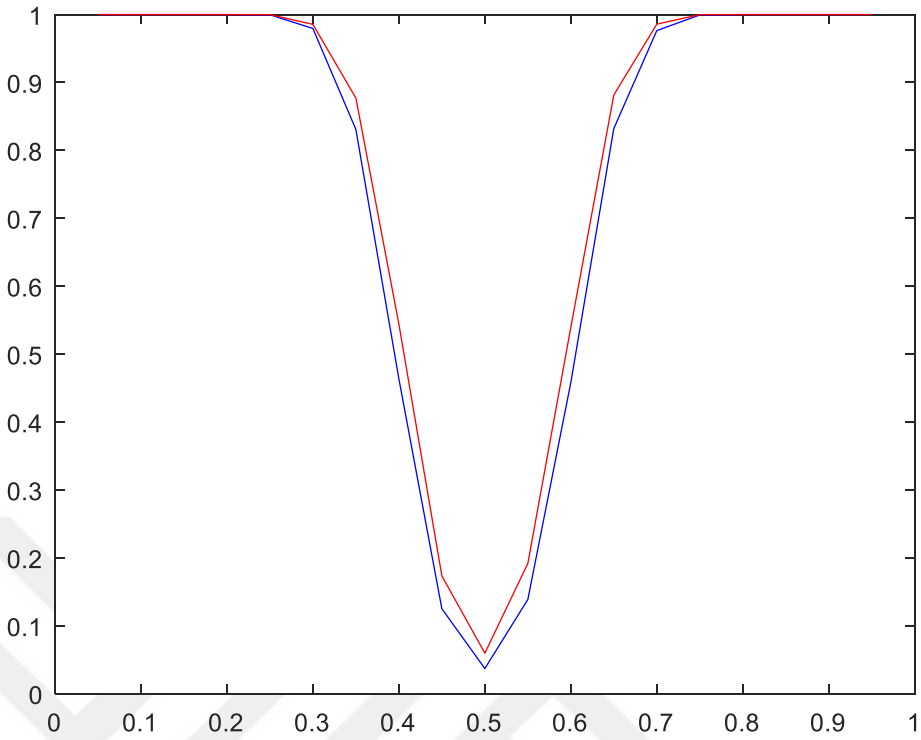
Şekil 3.67. $n = 30$ için güçler



Şekil 3.68. $n = 40$ için güçler



Şekil 3.69. $n = 50$ için güçler



Şekil 3.70. $n = 100$ için güçler

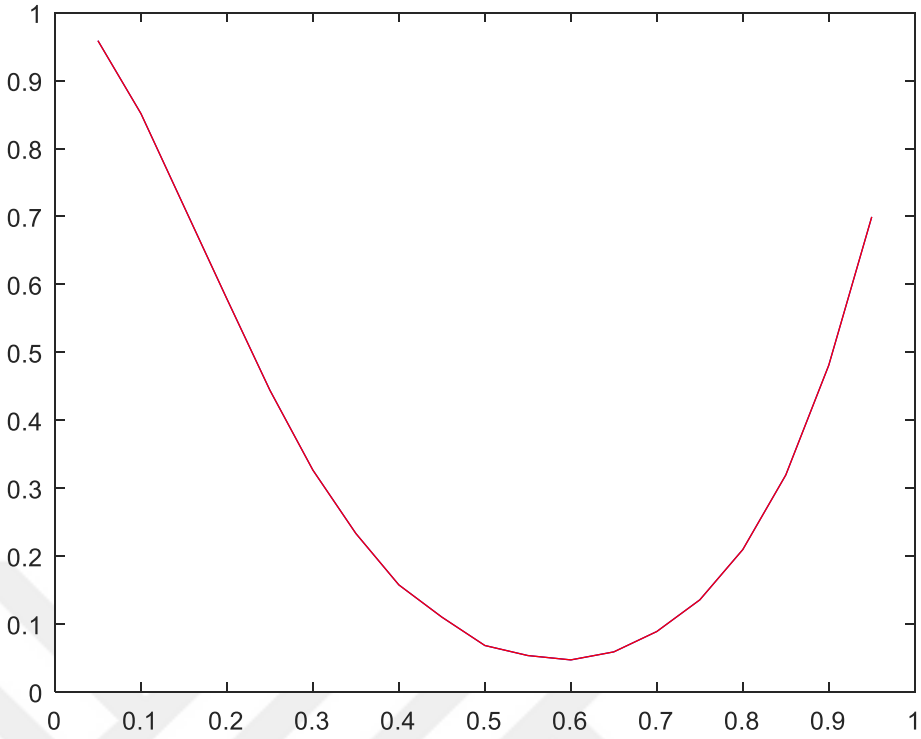
Küçük n değerlerinde testlerin gücü aynı olmakla beraber n arttıkça Z testinin gücünün Binom testinin gücünden daha yüksek olduğu görülmektedir. n sayısı arttıkça testlerinin güçlerinin arttığı şekillerden de görülmektedir.

Çizelge 3.11 ve Şekil 3.71 – 3.77'de kitle oranının 0.6'ya eşit olup olmadığının testinde farklı parametre ve farklı n değerleri için simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, 0.6 parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

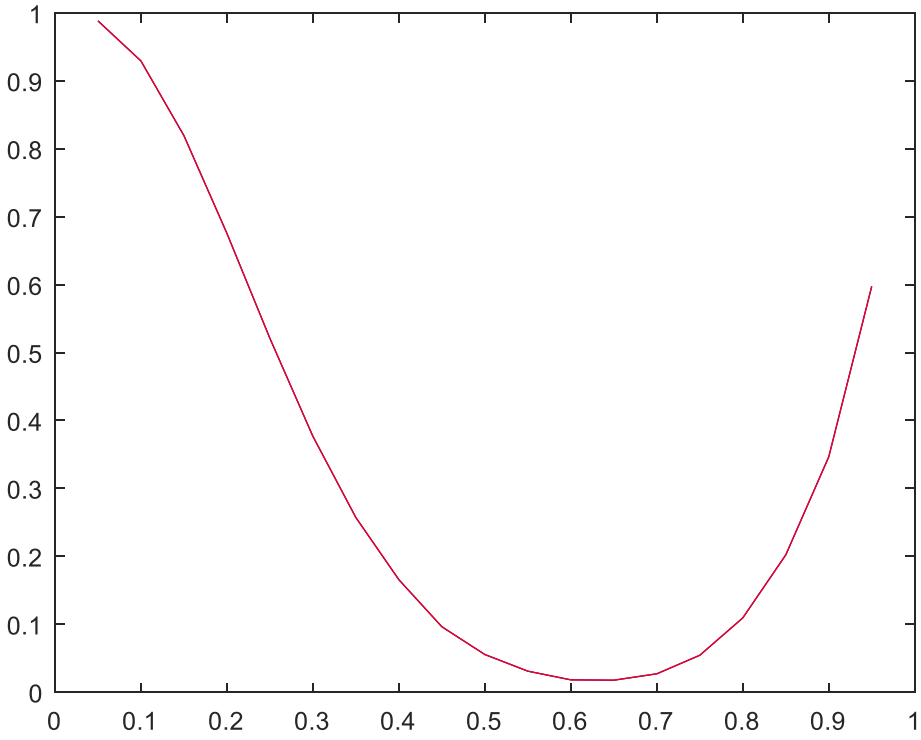
Çizelge 3.11. $H_0 : P = 0.6$ / $H_1 : P \neq 0.6$ durumu için testlerin güçleri

n	Parametre	Binom Testi	Z Testi
7	0.1	0.8515	0.8515
	0.2	0.5786	0.5786
	0.4	0.1576	0.1576
	0.6	0.0472	0.0472
	0.75	0.1356	0.1356
	0.9	0.4810	0.4810
	0.95	0.6994	0.6994
10	0.1	0.9294	0.9294

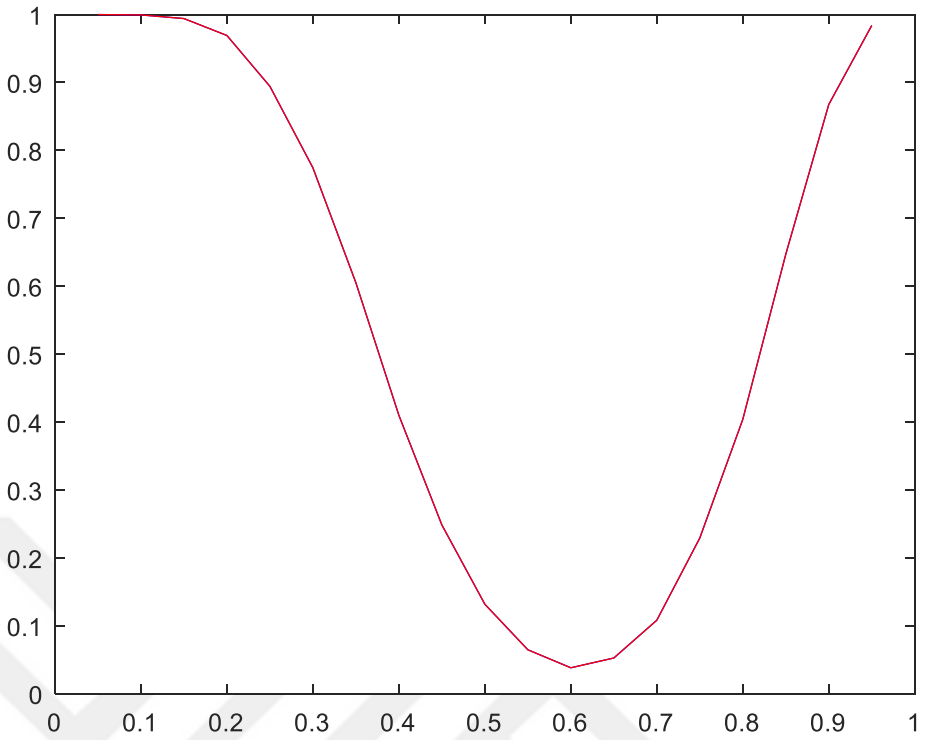
	0.2	0.6755	0.6755
	0.35	0.2571	0.2571
	0.45	0.0963	0.0963
	0.6	0.0181	0.0181
	0.75	0.0543	0.0543
	0.9	0.3465	0.3465
	0.95	0.5976	0.5976
20	0.05	1.00	1.00
	0.2	0.9692	0.9692
	0.35	0.6052	0.6052
	0.5	0.1318	0.1318
	0.6	0.0382	0.0382
	0.75	0.2293	0.2293
	0.85	0.6471	0.6471
	0.95	0.9839	0.9839
30	0.1	1.00	1.00
	0.2	0.9970	0.9970
	0.35	0.7759	0.7759
	0.5	0.1811	0.1811
	0.6	0.0406	0.0406
	0.75	0.3496	0.3496
	0.85	0.8530	0.8530
	0.95	0.9992	0.9992
40	0.15	1.00	1.00
	0.3	0.9646	0.9646
	0.4	0.6823	0.6823
	0.5	0.2144	0.2144
	0.6	0.0332	0.0332
	0.75	0.4324	0.4324
	0.85	0.9296	0.9296
	0.95	0.9999	0.9999
50	0.2	1.00	1.00
	0.25	0.9994	0.9997
	0.3	0.98	0.9949
	0.4	0.7636	0.8410
	0.5	0.2381	0.3303
	0.6	0.0438	0.06
	0.7	0.3283	0.3285
	0.75	0.6482	0.6482
	0.95	1.00	1.00
100	0.3	1.00	1.00
	0.35	0.9984	0.9991
	0.4	0.9750	0.9843
	0.5	0.4552	0.5348
	0.6	0.0420	0.0516
	0.7	0.5423	0.5423
	0.8	0.9929	0.9929
	0.9	1.00	1.00



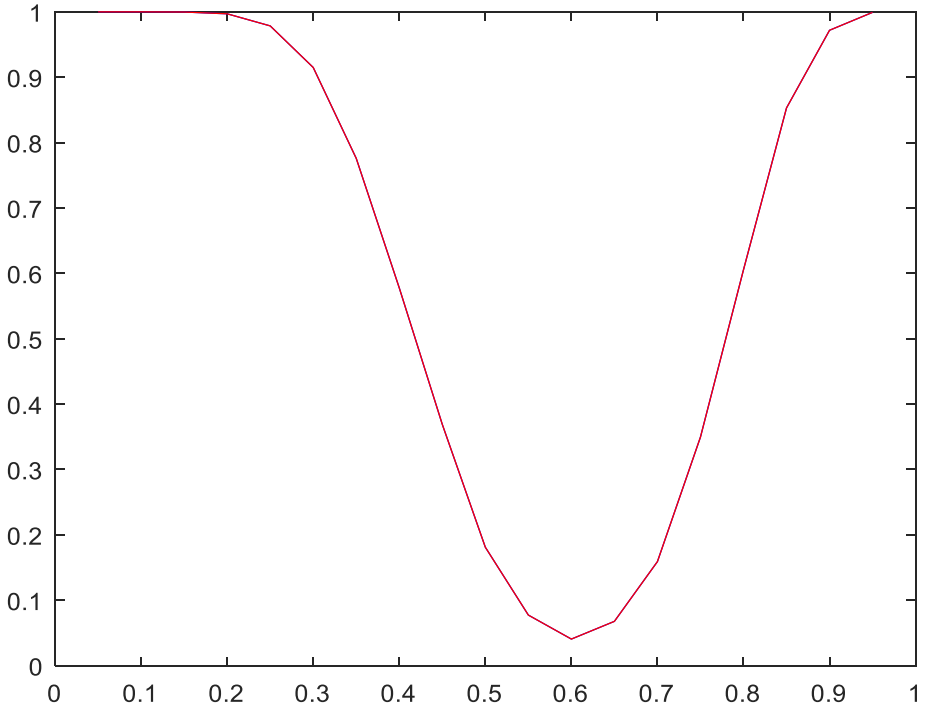
Şekil 3.71. $n = 7$ için güçler



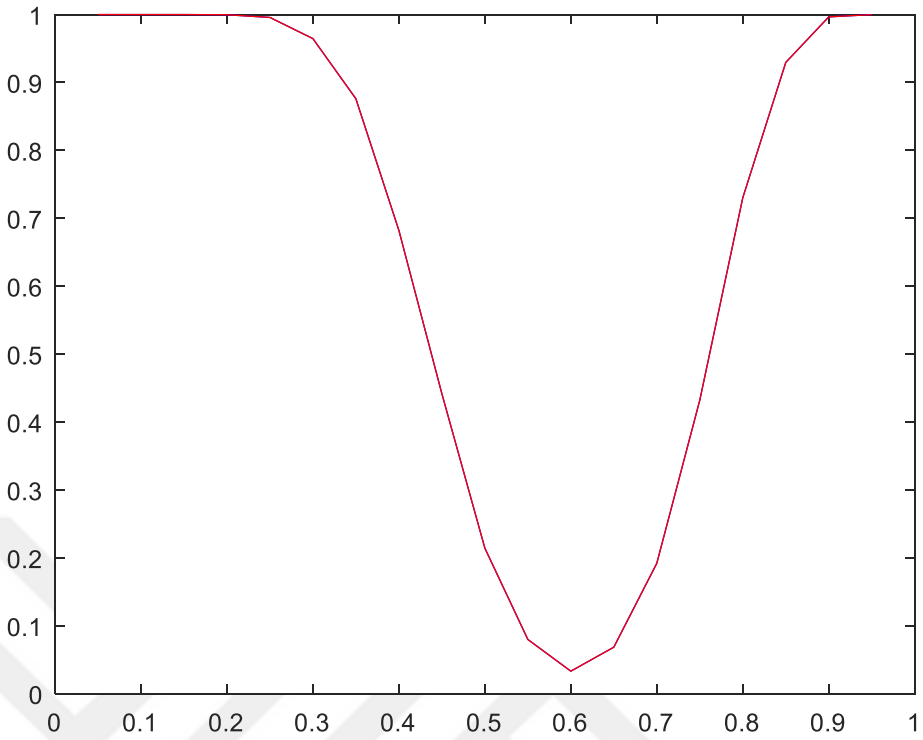
Şekil 3.72. $n = 10$ için güçler



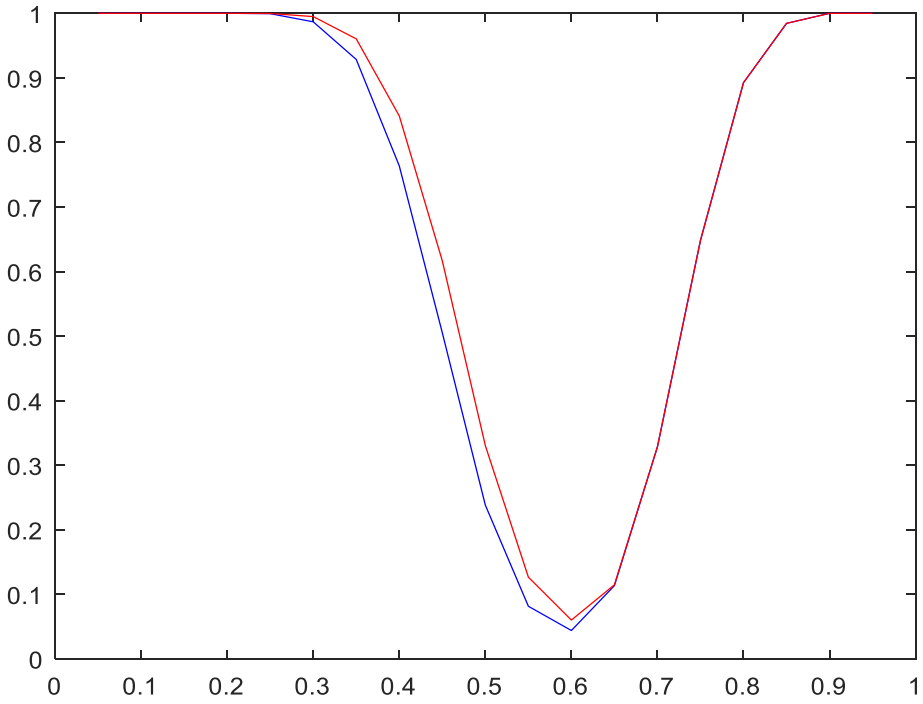
Şekil 3.73. $n = 20$ için güçler



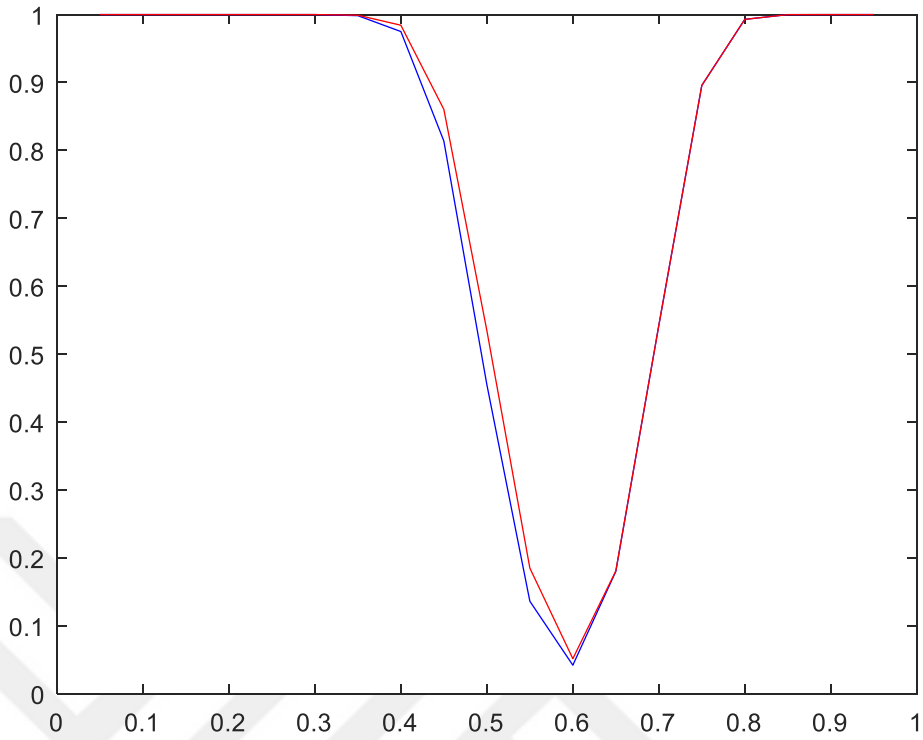
Şekil 3.74. $n = 30$ için güçler



Şekil 3.75. $n = 40$ için güçler



Şekil 3.76. $n = 50$ için güçler



Şekil 3.77. $n = 100$ için güçler

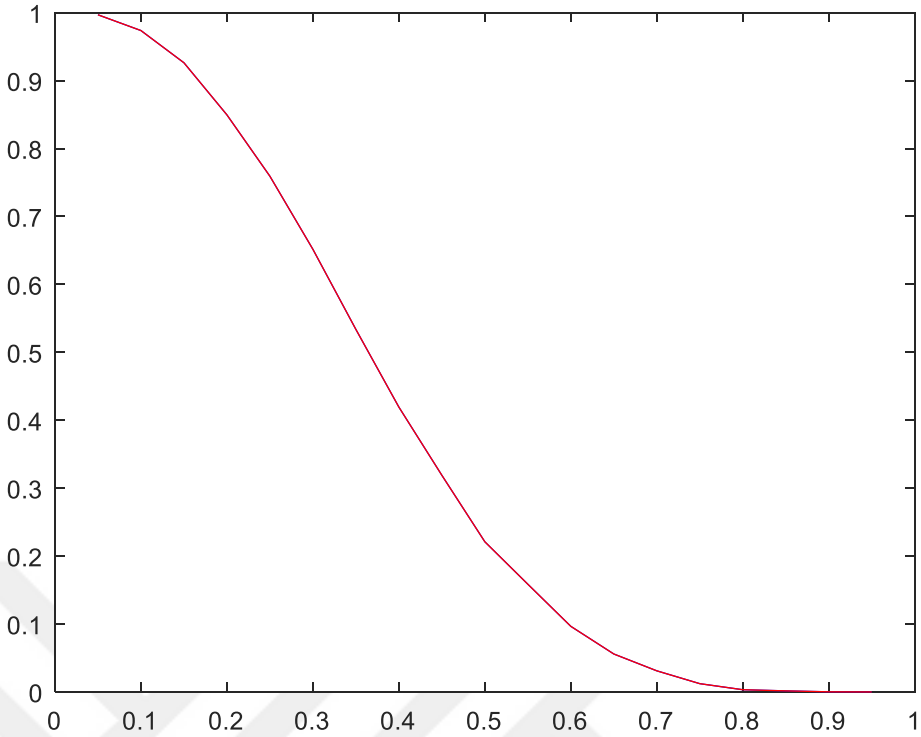
Küçük n değerlerinde testlerin gücü aynı olmakla beraber n arttıkça Z testinin gücünün Binom testinin gücünden daha yüksek olduğu görülmektedir. n sayısı arttıkça testlerinin güçlerinin arttığı şekillerden de görülmektedir.

Çizelge 3.12 ve Şekil 3.78 – 3.84'te kitle oranının 0.7'ye eşit olup olmadığının testinde farklı parametre ve farklı n değerleri için simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, 0.7 parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

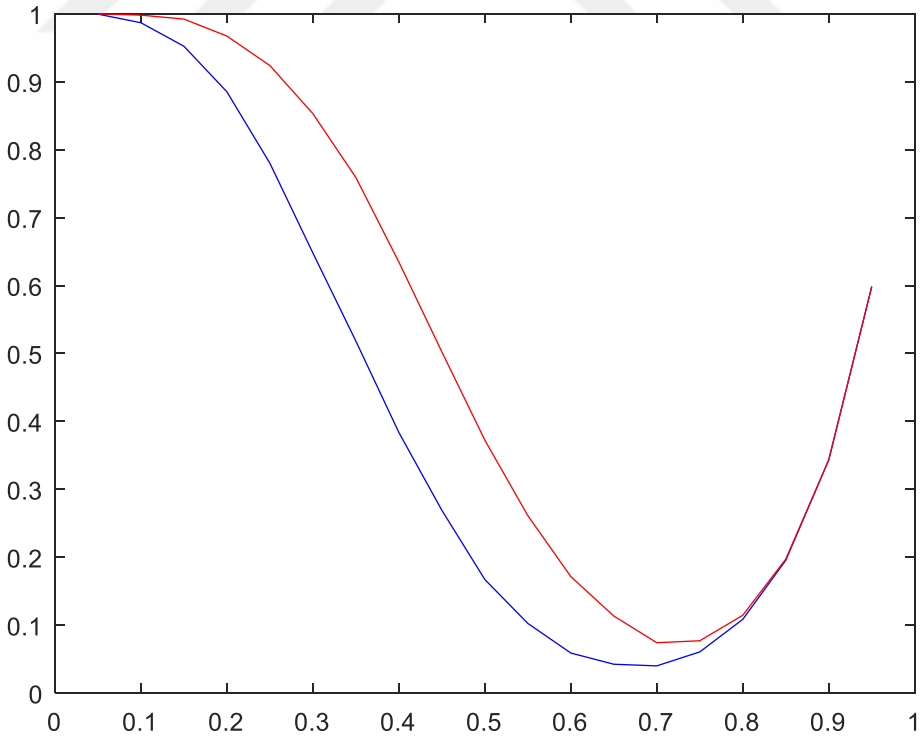
Çizelge 3.12. $H_0 : P = 0.7$ / $H_1 : P \neq 0.7$ durumu için testlerin güçleri

n	Parametre	Binom Testi	Z Testi
7	0.1	0.9737	0.9737
	0.2	0.8496	0.8496
	0.3	0.6517	0.6517
	0.45	0.3190	0.3190
	0.6	0.0966	0.0966
	0.7	0.0311	0.0311
	0.8	0.0033	0.0033
	0.9	0.0002	0.0002

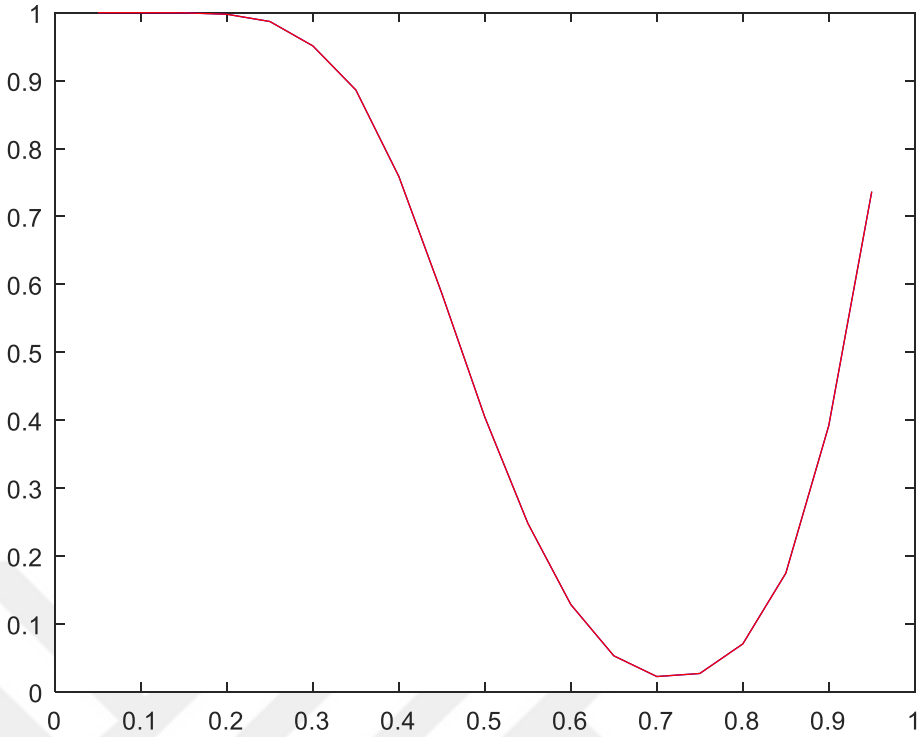
10	0.1	0.9869	0.9981
	0.2	0.8852	0.9673
	0.3	0.6479	0.8530
	0.4	0.3837	0.6346
	0.5	0.1672	0.3724
	0.6	0.0590	0.1715
	0.7	0.0401	0.0742
	0.8	0.1086	0.1146
	0.9	0.3434	0.3434
20	0.1	1.00	1.00
	0.3	0.9511	0.9511
	0.4	0.7588	0.7588
	0.5	0.4044	0.4044
	0.6	0.1287	0.1287
	0.7	0.0227	0.0227
	0.85	0.1748	0.1748
	0.95	0.7366	0.7366
30	0.25	0.9992	0.9998
	0.3	0.9936	0.9982
	0.4	0.9073	0.9526
	0.5	0.5705	0.7038
	0.6	0.1776	0.2879
	0.7	0.0463	0.0678
	0.8	0.2589	0.2597
	0.9	0.8259	0.8259
40	0.25	1.00	1.00
	0.3	0.9986	0.9997
	0.4	0.9621	0.9808
	0.5	0.6838	0.7868
	0.6	0.2059	0.3053
	0.7	0.0382	0.0562
	0.8	0.2800	0.2802
	0.85	0.6068	0.6068
	0.95	0.9954	0.9954
50	0.25	1.00	1.00
	0.45	0.9549	0.9549
	0.6	0.3252	0.3252
	0.7	0.0426	0.0426
	0.8	0.3087	0.3087
	0.85	0.6680	0.6680
	0.95	0.9997	0.9997
100	0.4	1.00	1.00
	0.5	0.9804	0.9880
	0.6	0.5411	0.6196
	0.7	0.0477	0.0615
	0.75	0.2062	0.2072
	0.8	0.6473	0.6473
	0.95	1.00	1.00



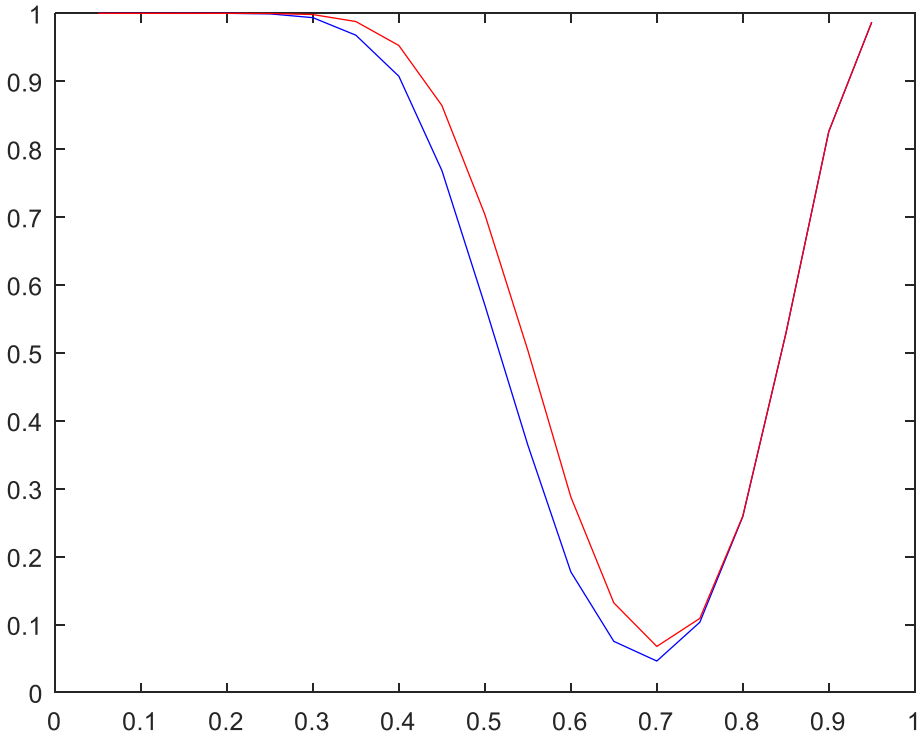
Şekil 3.78. $n = 7$ için güçler



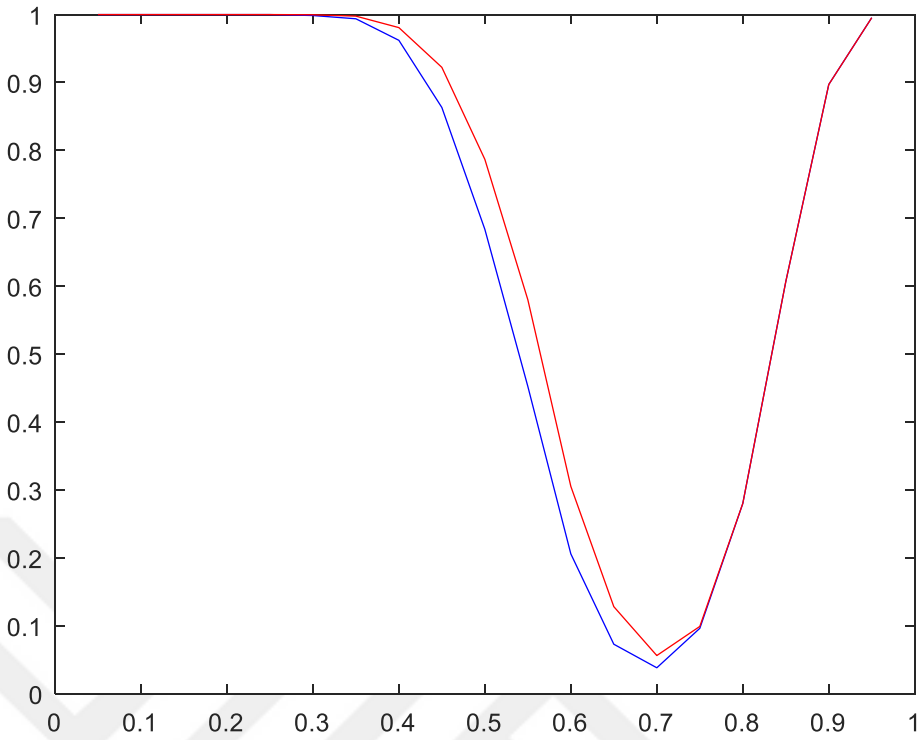
Şekil 3.79. $n = 10$ için güçler



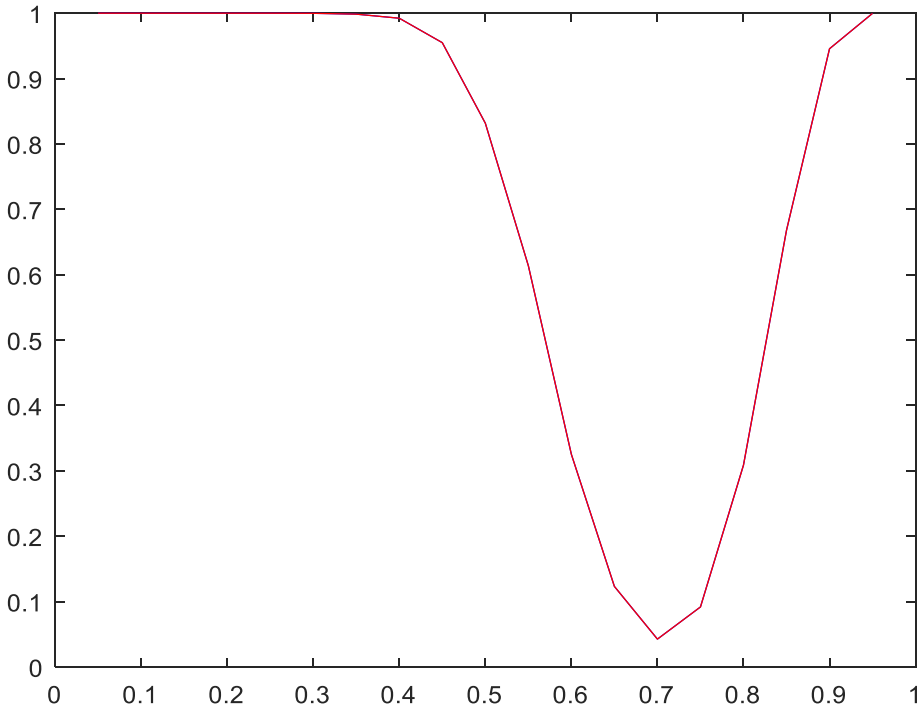
Şekil 3.80. $n = 20$ için güçler



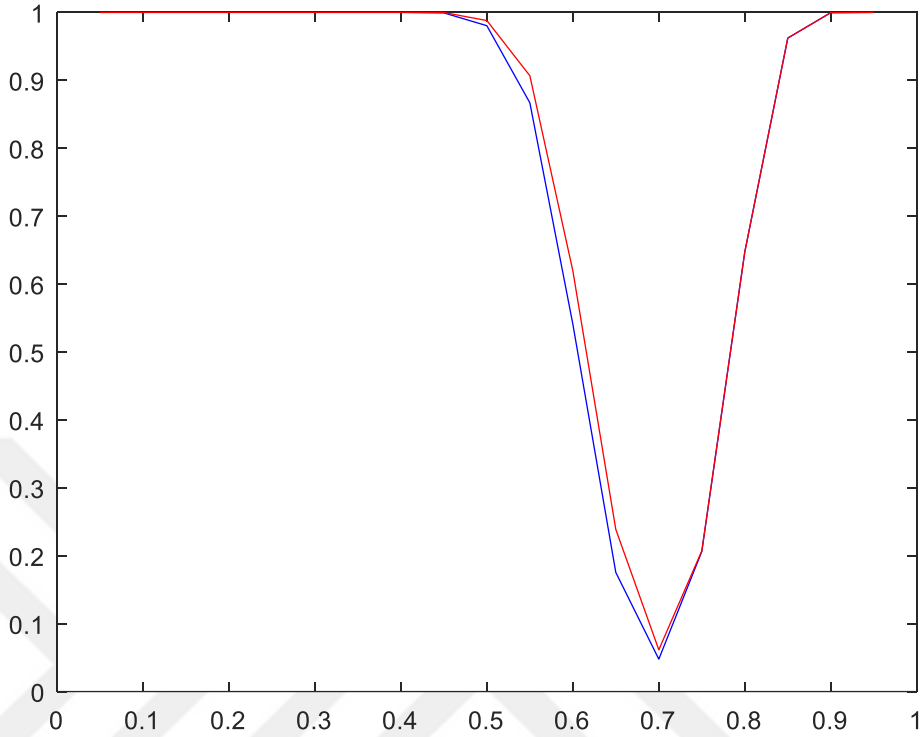
Şekil 3.81. $n = 30$ için güçler



Şekil 3.82. $n = 40$ için güçler



Şekil 3.83. $n = 50$ için güçler



Şekil 3.84. $n = 100$ için güçler

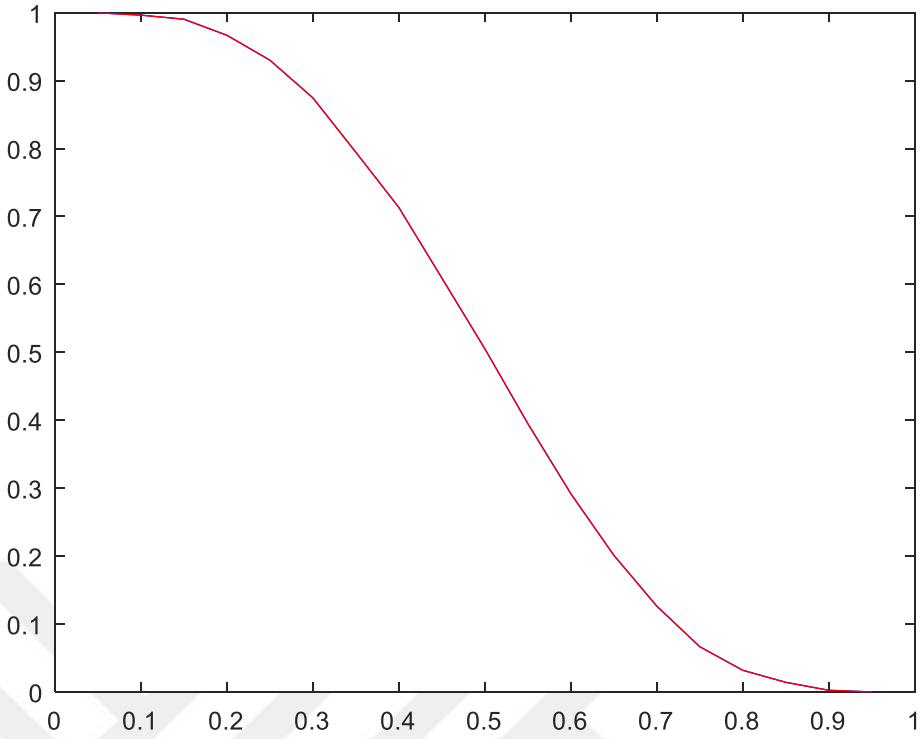
Testin güçleri incelendiğinde bu iki test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test Z testidir. n sayısı arttıkça testlerinin güçlerinin arttığı şekillerden de görülmektedir.

Çizelge 3.13 ve Şekil 3.85 – 3.91'de kitle oranının 0.8'e eşit olup olmadığının testinde farklı parametre ve farklı n değerleri için simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, 0.8 parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

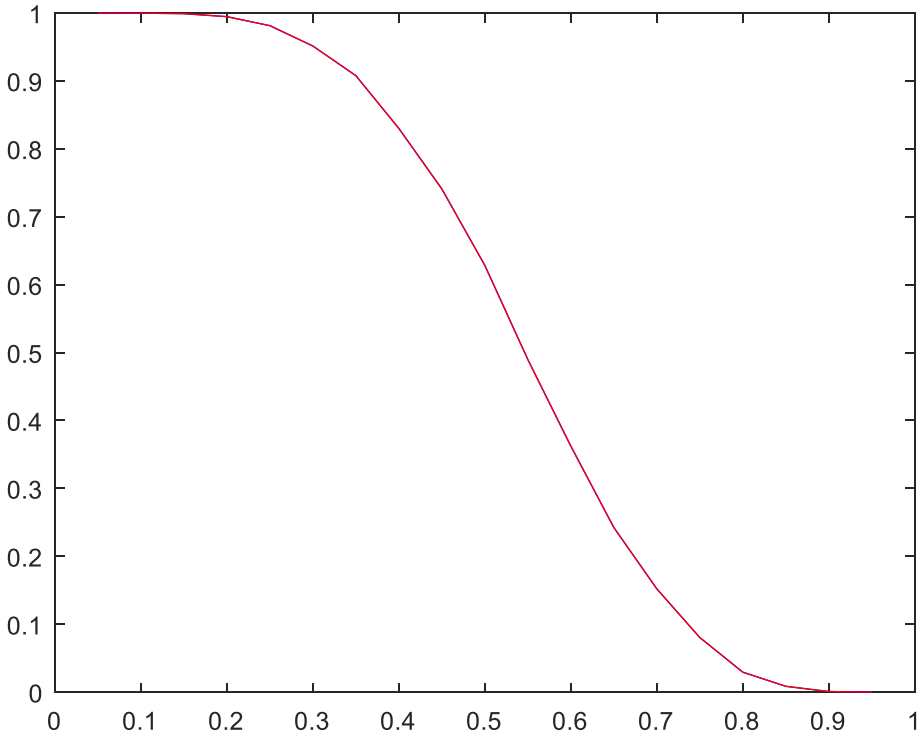
Çizelge 3.13. $H_0 : P = 0.8$ / $H_1 : P \neq 0.8$ durumu için testlerin güçleri

n	Parametre	Binom Testi	Z Testi
7	0.1	0.9964	0.9964
	0.25	0.9301	0.9301
	0.45	0.6094	0.6094
	0.55	0.3950	0.3950
	0.7	0.1263	0.1263

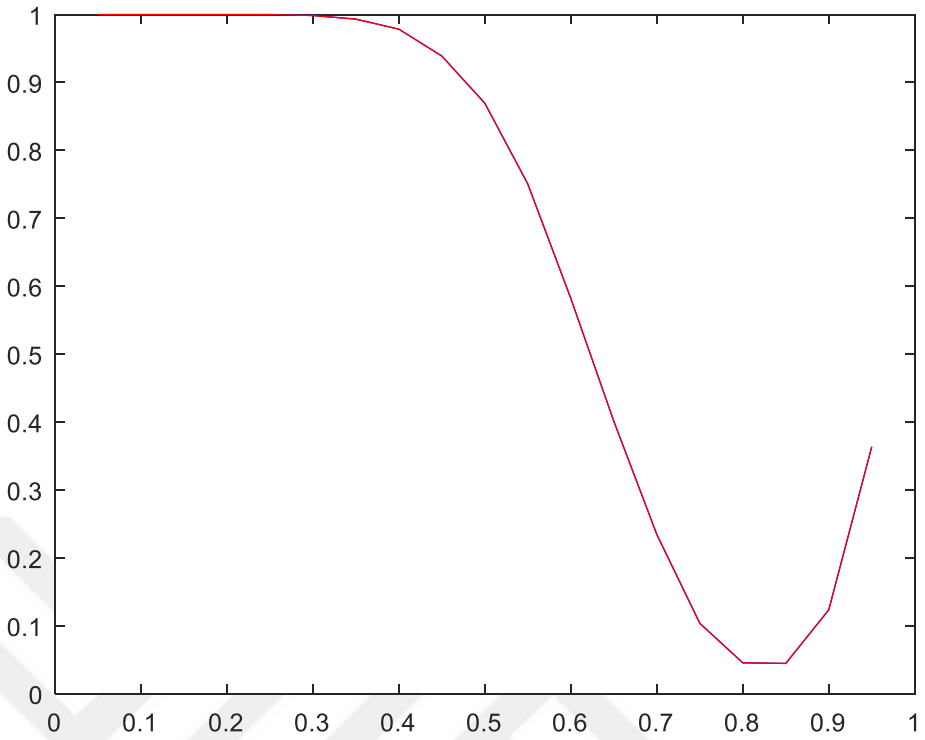
	0.8	0.0319	0.0319
	0.9	0.0025	0.0025
	0.95	0.0001	0.0001
10	0.05	1.00	1.00
	0.3	0.9513	0.9513
	0.45	0.7409	0.7409
	0.6	0.3622	0.3622
	0.7	0.1521	0.1521
	0.8	0.0294	0.0294
	0.85	0.0086	0.0086
	0.90	0.0009	0.0009
20	0.25	1.00	1.00
	0.4	0.9785	0.9785
	0.55	0.7504	0.7504
	0.65	0.4013	0.4013
	0.75	0.1037	0.1037
	0.8	0.0455	0.0455
	0.9	0.1235	0.1235
	0.95	0.3636	0.3636
30	0.3	1.00	1.00
	0.5	0.9543	0.9543
	0.6	0.6987	0.6987
	0.7	0.2723	0.2723
	0.8	0.0349	0.0349
	0.9	0.1825	0.1825
	0.95	0.5585	0.5585
40	0.35	0.9999	1.00
	0.4	0.9995	0.9998
	0.5	0.9825	0.9920
	0.6	0.7829	0.8669
	0.7	0.3016	0.4249
	0.8	0.0471	0.0687
	0.9	0.4310	0.4310
	0.95	0.8619	0.8619
50	0.4	1.00	1.00
	0.55	0.9784	0.9784
	0.65	0.7247	0.7247
	0.75	0.1688	0.1688
	0.8	0.0486	0.0486
	0.9	0.4290	0.4290
	0.95	0.8988	0.8988
100	0.5	1.00	1.00
	0.55	0.9994	0.9997
	0.65	0.9153	0.9451
	0.7	0.6251	0.7082
	0.75	0.2094	0.2850
	0.8	0.0436	0.0545
	0.85	0.2447	0.2450
	0.9	0.8062	0.8062
	0.95	0.9989	0.9989



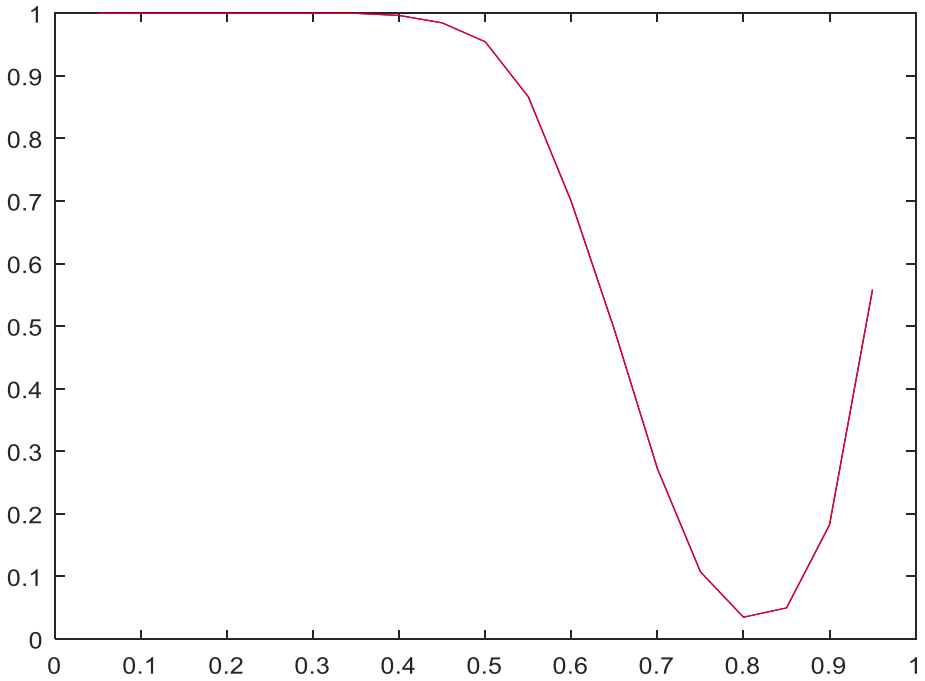
Şekil 3.85. $n = 7$ için güçler



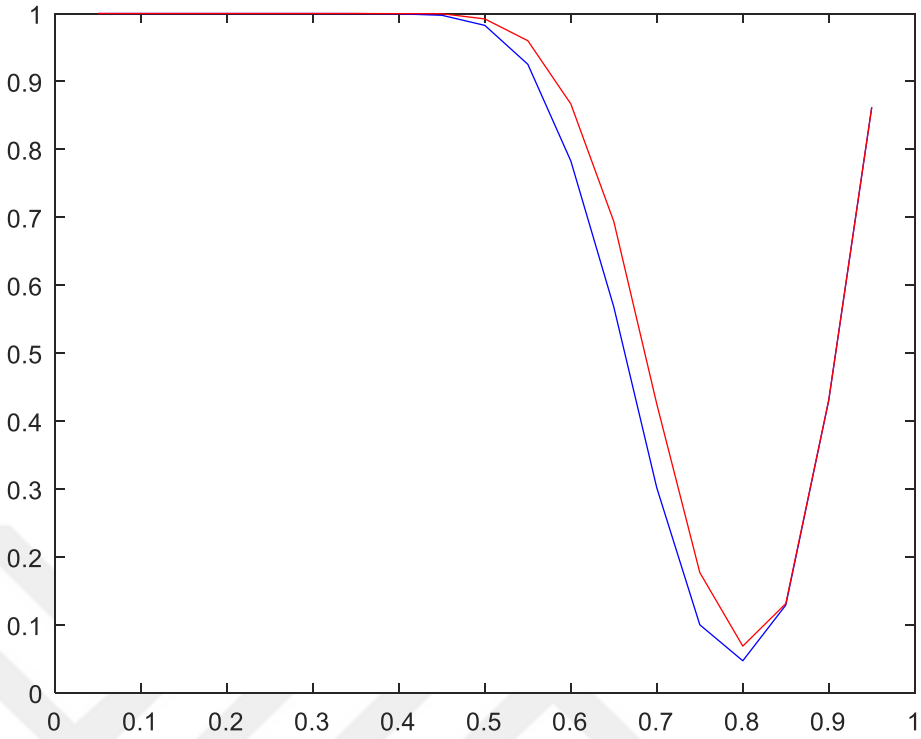
Şekil 3.86. $n = 10$ için güçler



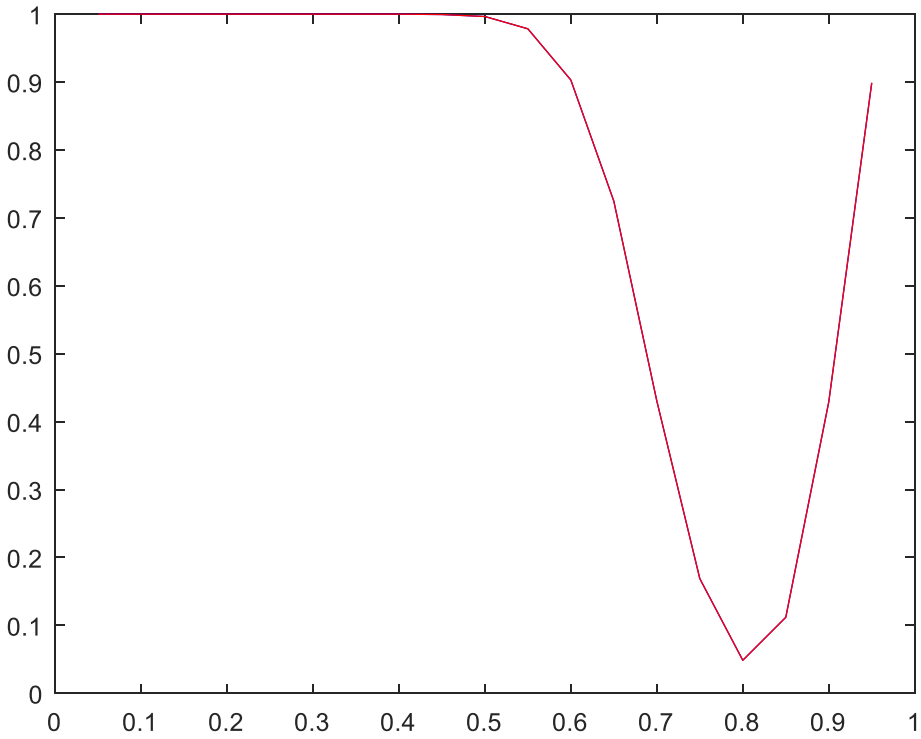
Şekil 3.87. $n = 20$ için güçler



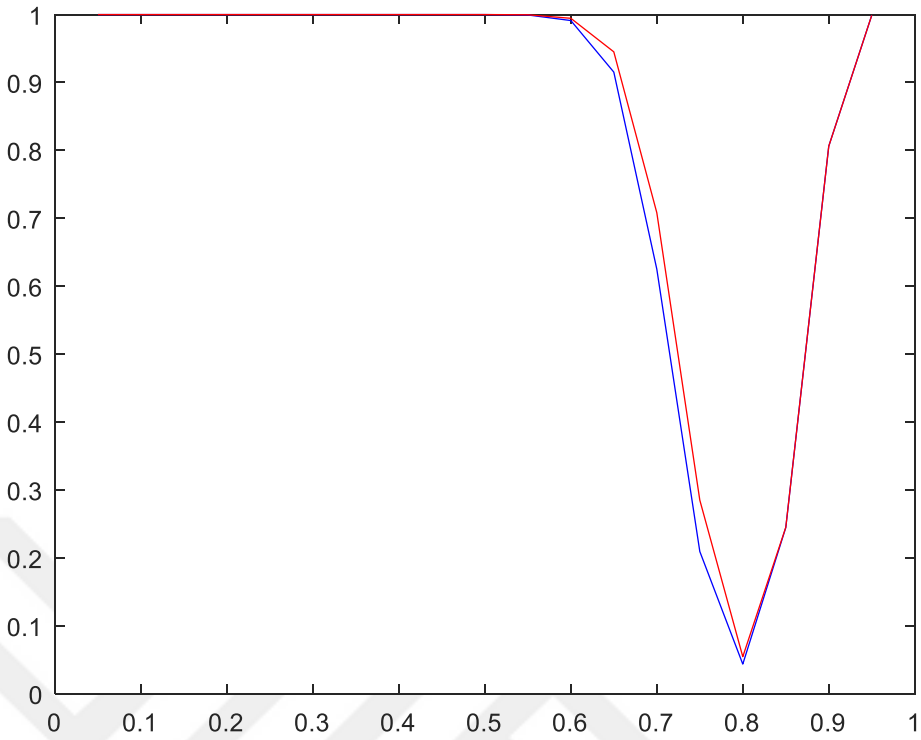
Şekil 3.88. $n = 30$ için güçler



Şekil 3.89. $n = 40$ için güçler



Şekil 3.90. $n = 50$ için güçler



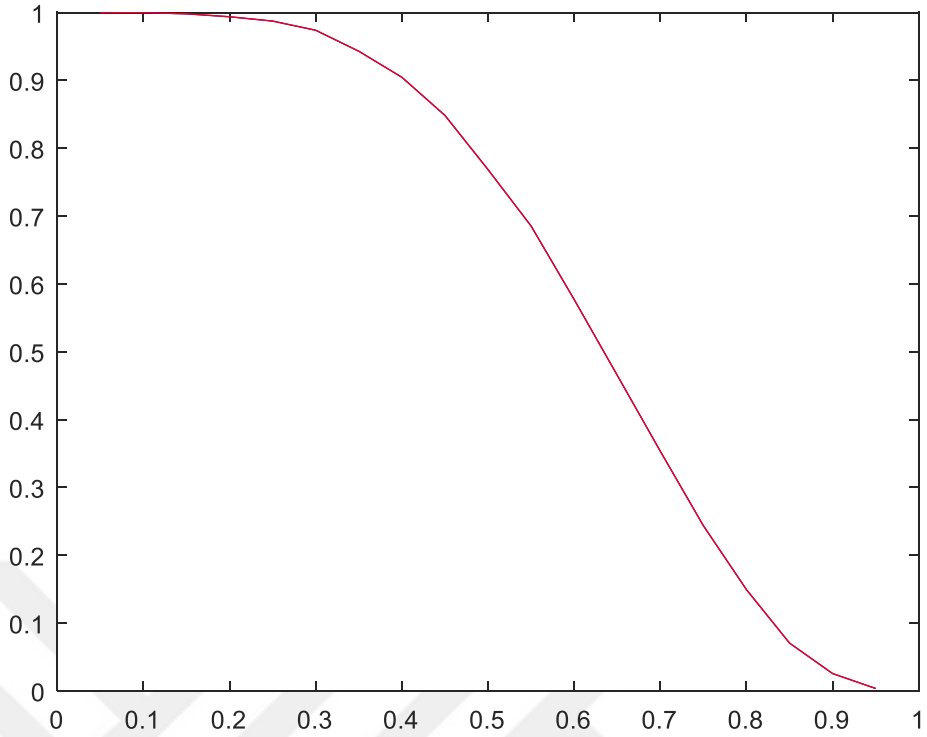
Şekil 3.91. $n = 100$ için güçler

Testin güçleri incelendiğinde bu iki test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test Z testidir. n sayısı arttıkça testlerinin güçlerinin arttığı şekillerden de görülmektedir.

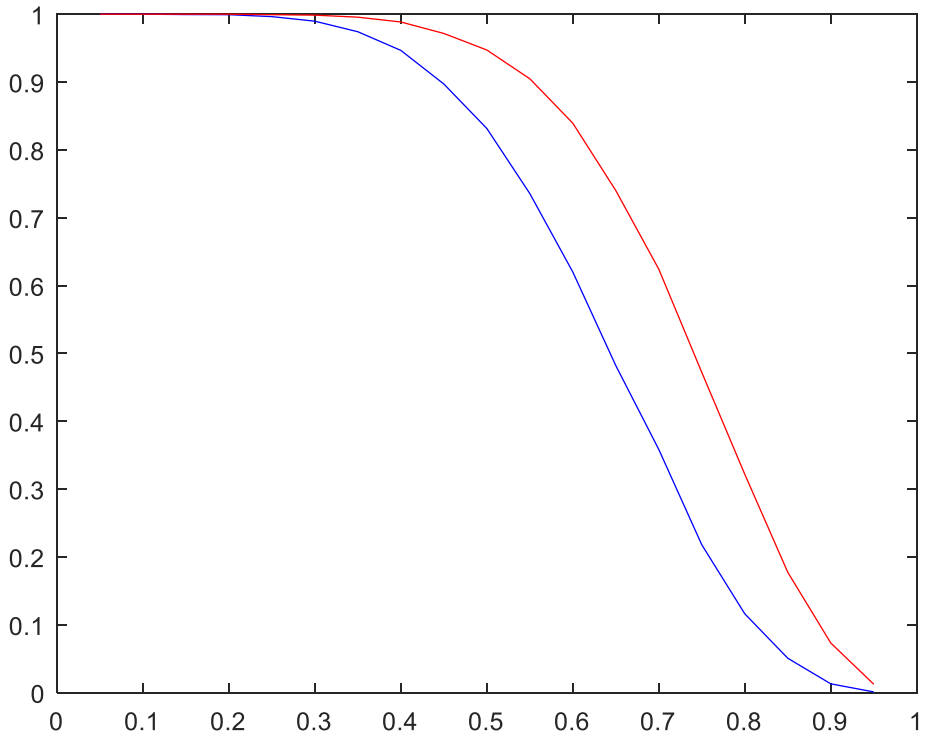
Çizelge 3.14 ve Şekil 3.92 – 3.98'de kitle oranının 0.9'a eşit olup olmadığının testinde farklı parametre ve farklı n değerleri için simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, 0.9 parametresine karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 3.14. $H_0 : P = 0.9$ / $H_1 : P \neq 0.9$ durumu için testlerin güçleri

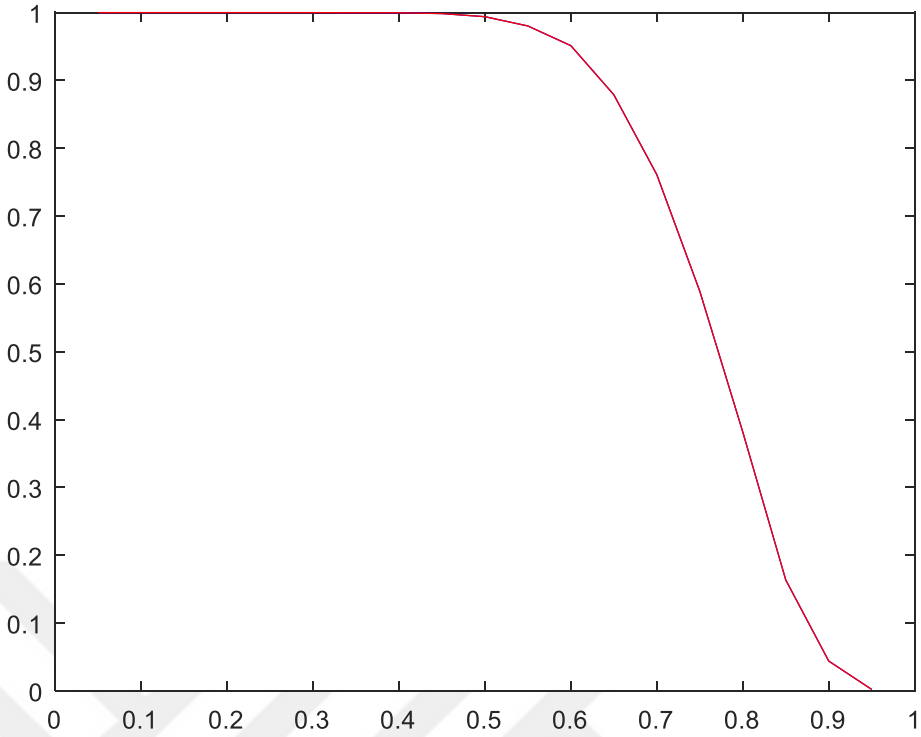
n	Parametre	Binom Testi	Z Testi
7	0.05	1.00	1.00
	0.3	0.9740	0.9740
	0.45	0.8485	0.8485
	0.55	0.6854	0.6854
	0.65	0.4654	0.4654
	0.8	0.1494	0.1494
	0.9	0.0257	0.0257
	0.95	0.0037	0.0037
10	0.1	1.00	1.00
	0.2	0.9993	1.00
	0.3	0.9897	0.9986
	0.4	0.9466	0.9885
	0.5	0.8315	0.9472
	0.6	0.6199	0.8393
	0.7	0.3583	0.6241
	0.8	0.1161	0.3218
	0.9	0.0129	0.0731
	0.95	0.0010	0.0123
20	0.3	1.00	1.00
	0.55	0.9803	0.9803
	0.7	0.7612	0.7612
	0.8	0.3818	0.3818
	0.85	0.1636	0.1636
	0.9	0.0441	0.0441
	0.95	0.0023	0.0023
30	0.4	1.00	1.00
	0.6	0.9827	0.9827
	0.75	0.6471	0.6471
	0.8	0.3911	0.3911
	0.85	0.1495	0.1495
	0.9	0.0265	0.0265
	0.95	0.0004	0.0004
40	0.45	1.00	1.00
	0.6	0.9937	0.9980
	0.7	0.8883	0.9440
	0.8	0.4060	0.5612
	0.85	0.1357	0.2410
	0.9	0.0293	0.0568
	0.95	0.1257	0.1259
50	0.55	1.00	1.00
	0.7	0.9594	0.9594
	0.8	0.5510	0.5510
	0.85	0.2041	0.2041
	0.9	0.0318	0.0318
	0.95	0.0791	0.0791
100	0.65	1.00	1.00
	0.75	0.9794	0.9902
	0.8	0.8048	0.8694
	0.85	0.3291	0.4319
	0.9	0.0447	0.0617
	0.95	0.4393	0.4393



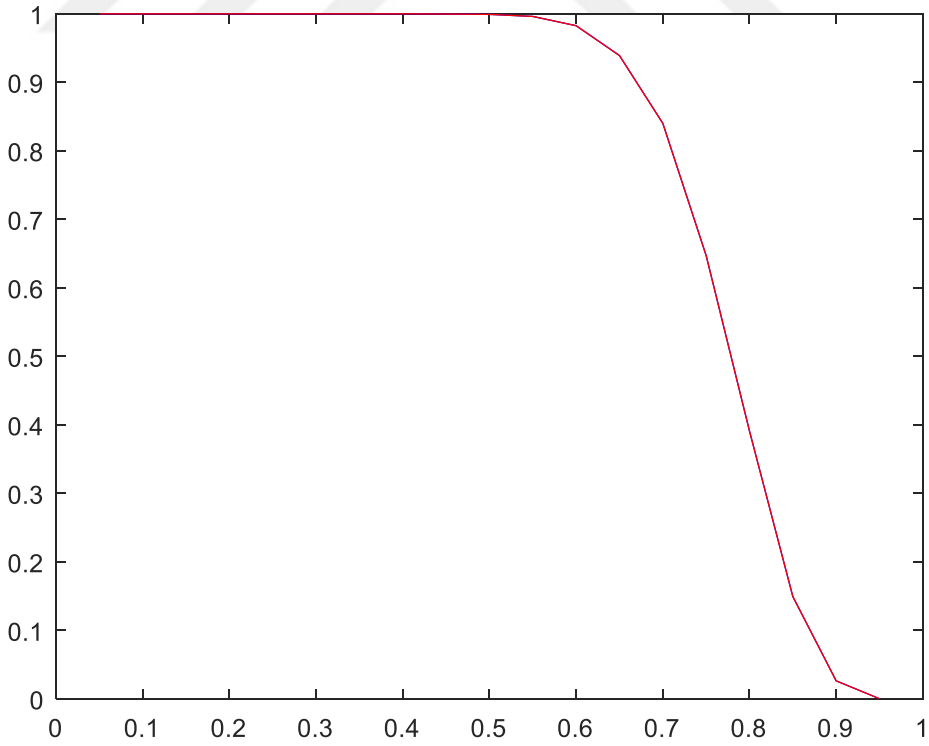
Şekil 3.92. $n = 7$ için güçler



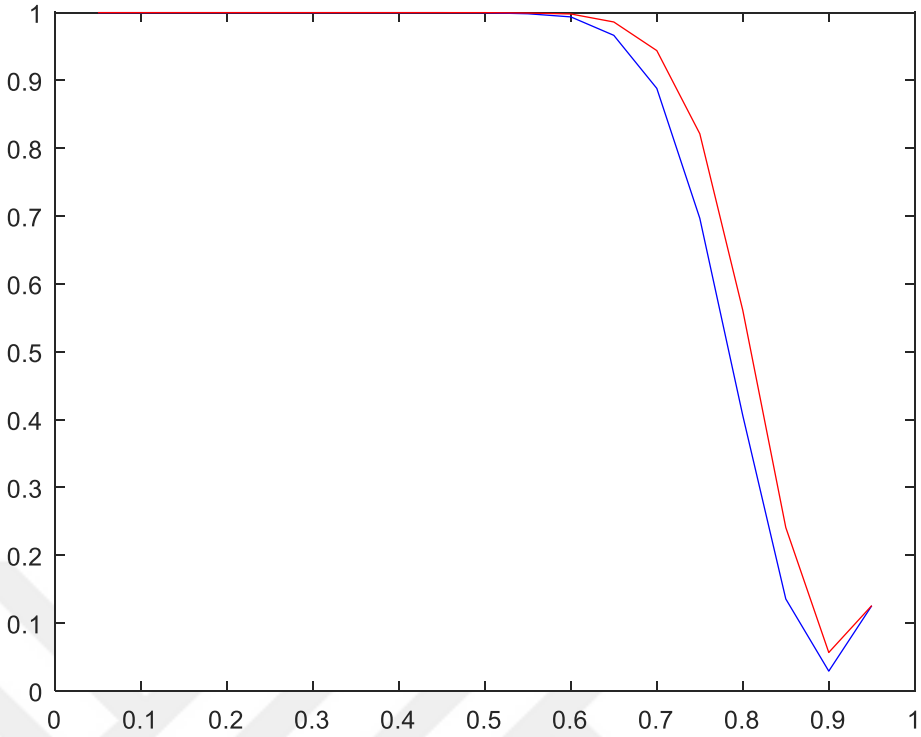
Şekil 3.93. $n = 10$ için güçler



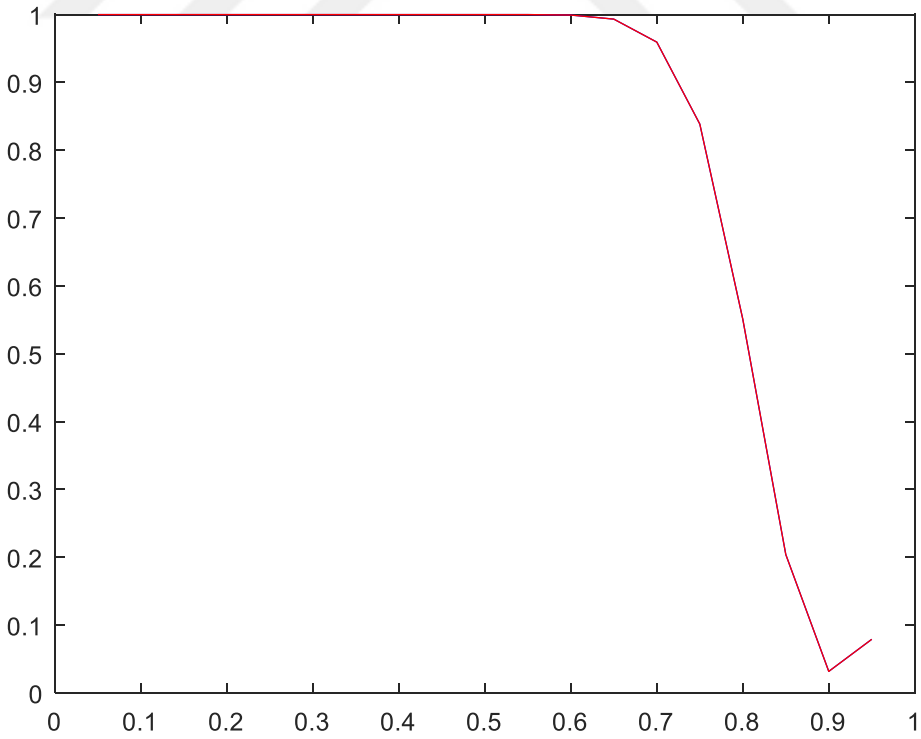
Şekil 3.94. $n = 20$ için güçler



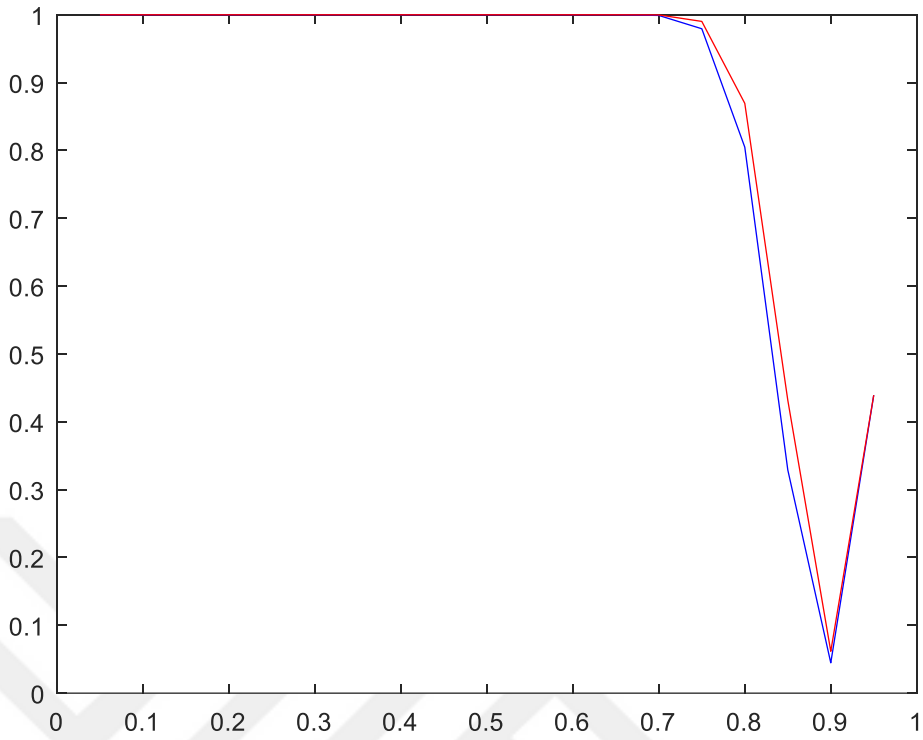
Şekil 3.95. $n = 30$ için güçler



Şekil 3.96. $n = 40$ için güçler



Şekil 3.97. $n = 50$ için güçler



Şekil 3.98. $n = 100$ için güçler

Testin güçleri incelendiğinde bu iki test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test Z testidir. n arttıkça testlerinin güçlerinin arttığı şekillerden de görülmektedir.

4. BAĞIMSIZ İKİ ÖRNEKLEM TESTLERİ

4.1. Ortalama ve Medyana İlişkin Testler

Bu bölümde bağımsız iki örneklem testlerinden ortalama ve medyana ilişkin parametrik Z testi ve t testi ile parametrik olmayan Mann Whitney U testi, Kolmogorov Smirnov testi, Wald-Wolfowitz Diziler Testi ve Moods-Medyan testi tanıtılmıştır.

4.1.1. Z testi

X ve Y örneklem değerleri olmak üzere, $X = X_1, X_2, ..., X_n \sim N\left(\mu_x, \sigma_x^2\right)$, $Y = Y_1, Y_2, ..., Y_n \sim N\left(\mu_y, \sigma_y^2\right)$ ve σ_x^2, σ_y^2 bilindiği durumlarda μ_x'in μ_y gibi teorik bir değere eşit olup olmadığını test etmek için kullanılan iki örneklemli parametrik testtir.

$H_0 : \mu_x = \mu_y$ hipotezi, $H_1 : \mu_x \neq \mu_y$ hipotezine karşı test ediliyor ise test istatistiği,

$$Z_{Hesap} = \frac{\left(\bar{X} - \bar{Y}\right) - \left(\mu_x - \mu_y\right)}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}} \tag{4.1}$$

şeklinde hesaplanmaktadır ve eğer $Z_{Hesap} > Z_{Tablo}$ ise iki örneklemin kitle ortalamalarının farklı olduğu sonucuna varılır yani H_0 reddedilir. (4.1) eşitliğinde $\bar{X}$ ve $\bar{Y}$, aritmetik ortalamalar, $\left(\mu_x - \mu_y\right)$, kitle ortalamalarının farkı, σ^2, varyans, n, örneklem sayısıdır. ortalamalar, σ^2, varyans, n, örneklem sayısıdır.

4.1.2. t testi

X ve Y örneklem değerleri olmak üzere, $X = X_1, X_2, ..., X_n \sim N(\mu_x, \sigma_x^2)$ $Y = Y_1, Y_2, ..., Y_m \sim N\left(\mu_y, \sigma_y^2\right)$ ve σ_x^2, σ_y^2 bilinmediği durumlarda μ_x'in μ_y gibi teorik bir değere eşit olup olmadığını test etmek için kullanılan iki örneklemli parametrik testtir.

$H_0 : \mu_x = \mu_y$ hipotezi, $H_1 : \mu_x \neq \mu_y$ hipotezine karşı test ediliyor ise varyansların homojen olduğu durumda yani $\sigma_1 = \sigma_2$ ise test istatistiği,

$$t_{Hesap} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{m}}} \tag{4.2}$$

Buradaki σ^2 değeri örneklerin benzer varyanslı kitleden alındığı varsayımından yararlanmaktadır ve,

$$\sigma^2 = \frac{\left|(n-1)\sigma_1^2\right| + \left|(m-1)\sigma_2^2\right|}{n+m-2} \tag{4.3}$$

şeklinde hesaplanmaktadır ve serbestlik derecesi ise, $n+m-2$ şeklindedir.

Varyansların homojen olmadığı durumda yani $\sigma_1 \neq \sigma_2$ ise test istatistiği,

$$t_{Hesap} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \tag{4.4}$$

şeklinde hesaplanmaktadır ve serbestlik derecesi ise,

$$sd = \frac{\left(\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)}{\frac{\left(\frac{\sigma_1^2}{n}\right)^2}{n-1} + \frac{\left(\frac{\sigma_2^2}{m}\right)^2}{m-1}} \tag{4.5}$$

şeklinde hesaplanmaktadır ve eğer $t_{Hesap} > t_{Tablo}$ ise iki örneklemin kitle ortalamalarının farklı olduğu sonucuna varılır yani H_0 reddedilir. (4.2) eşitliğinde $\overline{X}$ ve $\overline{Y}$, aritmetik ortalamalar, σ^2, varyans, n, örneklem sayısıdır.

4.1.3. Mann Whitney U testi

Mann Whitney U testi birbirinden bağımsız iki değişkenin bağımlı değişken üzerinde gruplar arasında bir farklılığın olup olmadığını incelemektedir. Bu bağlamda grupların medyan değerleri karşılaştırılmaktadır. Bu test bağımsız iki örneklemli t-testinin parametrik olmayan bir karşılığıdır.

$X_1, X_2, ..., X_m$ ve $Y_1, Y_2, ..., Y_n$ bağımsız gözlemler olmak üzere, rastgele iki gruba ayrılmış örnekler de, bağımsızdır. Gözlenen değişkenler sürekli rastgele değişkenlerden oluşmaktadır. mn rastgele değişkenler olarak tanımlandığında,

$$D_{ij} = \begin{cases} 1 & Y_i < X_i \quad i = 1, 2, ..., m \\ 0 & Y_i > X_i \quad j = 1, 2, ..., n \end{cases} \tag{4.6}$$

Mann Whitney U testi için test istatistiği,

$$U = \sum_{i=1}^{m} \sum_{j=1}^{n} D_{ij} \tag{4.7}$$

Bu testte hipotezler dağılımların aynı olup olmadığı üzerine kurulmaktadır. Diğer bir deyişle, $H_0 : \mu_1 = \mu_2$ hipotezi $H_1 : \mu_1 \neq \mu_2$ karşı test edilmektedir. Test istatistiği ise,

$$\begin{aligned} U_{Hesap} &= mn + \frac{m(m+1)}{2} - R_1 \\ U_{Hesap} &= mn + \frac{m(m+1)}{2} - R_2 \end{aligned} \tag{4.8}$$

şeklinde hesaplanmaktadır. (4.8) eşitliğindeki R_1, m örneklem büyüklüğündeki gruba ilişkin sıra sayılarının toplamı ve R_2, n büyüklüğündeki gruba ilişkin sıra sayılarının

toplamıdır. İki denklem için de U değeri hesaplanmakta ve tablo değeri ile karşılaştırmak için küçük olan U değeri tercih edilmektedir. $U_{Hesap} \leq U_{Tablo}$ olduğu durumda H_0 reddedilir.

4.1.4. Kolmogorov Smirnov testi

$X_1, X_2, ..., X_m$ ve $Y_1, Y_2, ..., Y_n$ iki n ve m hacimli bağımsız örnekler olmak üzere, bu iki bağımsız örneğin aynı kitleden gelip gelmediğini test etmek amacıyla kullanılmaktadır.

$H_0 : Tüm\ x'ler\ için\ F(x) \neq F_0(x)$ hipotezi $H_1 : En\ az\ bir\ x\ için\ F(x) \neq F_0(x)$ hipotezine karşı test edilmektedir. Burada bahsedilen $F(x)$, dağılım fonksiyonu, $F_0(x)$ ise iddia edilen dağılış fonksiyonu olarak tanımlanmaktadır. Kolmogorov Smirnov test istatistiği,

$$D = \max_x \left| S(x) - F_0(x) \right| \tag{4.9}$$

(4.9) eşitliğinde $S(x)$, örnek dağılım fonksiyonudur ve x'ten küçük veya x'e eşit örnek gözlemlerinin sayısının örneklem büyüklüğüne olan oranı olarak belirlenmektedir. Eğer $D > F(x) - S(x)$ ise H_0 hipotezi reddedilir.

4.1.5. Wald-Wolfowitz Diziler testi

Bu iki test bağımsız şans örneğinin aynı sürekli ana kitleden alınıp alınmadığını test etmek amacıyla kullanılan bir parametrik olmayan testtir. Bu test için, gözlemlerin en az ordinal ölçekte olması gerekmektedir. Wald-Wolfowitz dizi sayıları testi; iki örneğin alındığı ana kitlelerin dağılım, yer, çarpıklık gibi herhangi bir parametre yönünden farklı olduğunu iddia eden alternatif hipoteze karşı, örneklerin aynı ana kitleden çekildiğini iddia eden sıfır hipotezini test etmektedir (Karagöz, 2010).

$X_1, X_2, ..., X_m$ ve $Y_1, Y_2, ..., Y_n$ iki bağımsız rastgele değişkenler olmak üzere ve bu X ve Y örneklerine karşılık gelen gözlemleri takip ederek küçükten büyüğe art arda düzenli bir şekilde dizilmektedir. Örneğin sıralama XYYXXYXYY şeklinde olduğunda

örnek havuzunda en küçük elemanın bir X olduğu en küçük ikinci elemanın bir Y ve en büyüğün bir Y olduğu gösterilmektedir ve hipotez testleri,

$$H_0 : F_Y(x) = F_X(x) her\ x için$$

$$H_1 : F_Y(x) \neq F_X(x) bir\ kaç\ x için$$

şeklinde oluşturulmaktadır.

$m+n=N$ rastgele değişkenleri ortak bir kitleden N boyutlu tek bir rastgele örneklem oluşturmaktadır. Özdeş harflerin bir dizisi ve ardından farklı bir harf ya da harf olmaksızın örnek havuzundaki dizilerin toplam sayısı karışım derecesinin göstergesidir.

Örneğin, X örneğindeki tüm elemanlar, Y örneğindeki tüm elemanlardan küçük ise eğer sadece iki dizi olacaktır. Bu özel biçim hem kitlelerin eşit olmadığını hem de X'in stokastik bir şekilde Y'den küçük olduğunu gösterebilir.

Küçükten büyüğe doğru örnek havuzundaki sıralama $\underset{1}{X}\ \underset{2}{YY}\ \underset{3}{XX}\ \underset{4}{Y}\ \underset{5}{X}\ \underset{6}{YY}$ şeklinde olmak üzere $m=4$, $n=5$ ve $r_{Hesap}=6$ olarak hesaplanmaktadır.

H_0 hipotezi doğru ise X ve Y simgelerinin iyice karışık biçimde dizilmiş olması dolayısıyla da test istatistiğinin (r) yeteri kadar büyük değer alması beklenmektedir. Öte yandan H_1 hipotezi doğru ise X ve Y simgelerinin kendi aralarında bir kümelenme sergilemesi dolayısıyla test istatistiğinin (r) oldukça küçük değer alması beklenir. Buna dayanarak $r_{Hesap} \leq r_{Tablo}$ olduğu durumlarda H_0 hipotezi reddedilir (Gamgam ve Altunkaynak, 2012).

4.1.6. Moods-Medyan testi

$X_1, X_2, ..., X_m$ ve $Y_1, Y_2, ..., Y_n$ iki bağımsız örneklem olmak üzere, medyan testi bu iki grup arasında medyan farkı olup olmadığını tespit etmeye yönelik bir testtir. Bu testi gerçekleştirmek için her iki kitleden rastgele bir örnek alınmaktadır. Ölçme düzeyi en az sıralamadır ve ilgilenilen değişken süreklidir.

$H_0 : M_1 = M_2$ hipotezi $H_1 : M_1 \neq M_2$ hipotezine karşı test edilmektedir. Yani kitle medyanlarının aynı olduğu hipotezi, X'ler Y'lerden büyük ya da Y'ler X'lerden büyük olma eğilimindedir hipotezine karşı test edilmektedir. Buradaki M_1, ilk örneğin

seçildiği kitlenin bilinmeyen medyanı ve M_2 ikinci örneğin seçildiği kitlenin bilinmeyen medyanı olarak tanımlanmaktadır. İki örnekteki gözlemler birleştirilerek ve $m+n$ sayıdaki birleştirilmiş örnek medyanı hesaplanmaktadır (Conover, 1980).

	Örnek No		**Toplam**
	1	**2**	
> Medyan	A	B	A+B
≤ Medyan	C	D	C+D
Toplam	*m*	*n*	$m+n$

Birleştirilmiş örnek medyanından büyük değerli olanların sayısı birinci ve ikinci örnek için sırasıyla A ve B, birleştirilmiş örnek medyanından küçük değerli olanların sayısı birinci ve ikinci örnek için sırasıyla C ve D olarak 2*2'lik tabloda gösterilmektedir (Gamgam ve Altunkaynak, 2012).

$$P_1 = \frac{A}{m},\ P_2 = \frac{B}{n},\ P = \frac{A+B}{m+n}$$

olarak parametreler tanımlandığında test istatistiği,

$$Z_{Hesap} = \frac{P_1 - P_2}{\sqrt{P(1-P)\left(\frac{1}{m}+\frac{1}{n}\right)}} \tag{4.10}$$

şeklinde hesaplanmaktadır ve $Z_{Hesap} < -Z_{\alpha/2}$ olduğu durumlarda H_0 reddedilir.

4.1.7. Güç karşılaştırmaları

Ortalama ve medyana ilişkin ele alınan parametrik Z ve t testi ile parametrik olmayan Kolmogrov-Smirnov testi, Mann Whitney U testi, Moods-Medyan testi ve Wald-Wolfowitz testlerin güç karşılaştırmalarının yapılması amacıyla 5000 tekrarlı Monte-Carlo simülasyon yöntemi uygulanmıştır. Simülasyonlarda $H_0 : \mu_1 = \mu_2 \left(M_1 = M_2\right)$ hipotezine karşılık $H_1 : \mu_1 \neq \mu_2 \left(M_1 \neq M_2\right)$ hipotezi ve $\alpha = 0.05$ anlamlılık seviyesi ele alınmıştır. Farklı dağılımlar ve farklı n değerleri için elde edilen sonuçlar Çizelge 4.1 – 4.6'da verilmiştir.

Çizelge 4.1.'de varyansı 1 olan normal dağılımlı iki kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, $|\mu_1 - \mu_2| = 0$'a karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 4.1. $X_1, X_2, \cdots, X_n \sim N(\mu_1, 1)$, $Y_1, Y_2, \cdots, Y_n \sim N(\mu_1, 1)$ için testin güçleri

n	$\lvert\mu_1 - \mu_2\rvert$	Z testi	t testi $(\sigma_1 = \sigma_2)$	t testi $(\sigma_1 \neq \sigma_2)$	Kolmogrov-Smirnov	Mann-Whitney	Moods Medyan	Wald-Wolfowitz
10	**0**	0.0508	0.0494	0.0474	0.0564	0.045	0.025	0.0338
	1	0.6112	0.5688	0.5636	0.4764	0.5204	0.3056	0.1696
	2	0.9936	0.9874	0.9872	0.9642	0.98	0.8754	0.7308
	3	1.00	1.00	1.00	0.9998	1.00	0.998	0.9876
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00
20	**0**	0.0514	0.0512	0.051	0.0356	0.0478	0.0288	0.036
	1	0.8874	0.8706	0.8702	0.7042	0.857	0.6008	0.2624
	2	1.00	1.00	1.00	0.9998	1.00	0.9968	0.9344
	3	1.00	1.00	1.00	1.00	1.00	1.00	0.9998
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00
30	**0**	0.0506	0.0494	0.0494	0.0362	0.048	0.0724	0.0276
	1	0.9706	0.9664	0.9664	0,89	0.9582	0.8964	0.3028
	2	1.00	1.00	1.00	1.00	1.00	1.00	0.9808
	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00
40	**0**	0.0492	0.0512	0.0504	0.045	0.0454	0.0412	0.0304
	1	0.993	0.993	0.993	0.9746	0.9908	0.9366	0.348
	2	1.00	1.00	1.00	1.00	1.00	1.00	0.994
	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00
50	**0**	0.0502	0.051	0.051	0.0388	0.0472	0.07	0.0384
	1	0.9986	0.9976	0.9976	0.9876	0.9972	0.9816	0.4726
	2	1.00	1.00	1.00	1.00	1.00	1.00	0.9998
	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00
100	**0**	0.051	0.0498	0.0498	0.0522	0.0502	0.0652	0.0412
	1	1.00	1.00	1.00	1.00	1.00	1.00	0.7854
	2	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Çizelge 4.1.'deki testlerin güçleri incelendiğinde bu yedi test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test Z testi, gücü en düşük olan test ise Wald-Wolfowitz testidir. n sayısı arttıkça tüm testlerin gücünün arttığı

çizelgeden görülmektedir. Ayrıca kitle ortalamaları arasındaki fark arttıkça testlerin gücü artmaktadır. Kitle ortalamaları arasındaki fark 2'den fazla olduğunda ise güç bakımından testler arasında farklılık olmadığı dolayısıyla istenilen testin kullanılabileceği söylenebilir.

Çizelge 4.2.'de varyansları farklı normal dağılımlı iki kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, $|\mu_1 - \mu_2| = 0$'a karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 4.2. $X_1, X_2, \cdots, X_n \sim N(\mu_1, 1)$, $Y_1, Y_2, \cdots, Y_n \sim N(\mu_1, 10)$ için testin güçleri

n	$\lvert\mu_1 - \mu_2\rvert$	Z testi	t testi $(\sigma_1 = \sigma_2)$	t testi $(\sigma_1 \neq \sigma_2)$	Kolmogrov-Smirnov	Mann-Whitney	Moods Medyan	Wald-Wolfowitz
10	**0**	0.0486	0.0612	0.0476	0.4858	0.0852	0.077	0,7188
	1	0.0606	0.0734	0.0572	0.5182	0.1006	0.0894	0,7082
	2	0.0962	0.1078	0.0878	0.5632	0.1398	0.1258	0,7236
	3	0.1598	0.1652	0.1392	0.617	0.186	0.1702	0.7282
	4	0.2422	0.2392	0.2032	0.6876	0.2526	0.2294	0.7374
	5	0.3458	0.325	0.2796	0.7464	0.3246	0.299	0.7692
20	**0**	0.05	0.0556	0.0508	0.9052	0.0946	0.0844	0.981
	1	0.0766	0.08	0.072	0.9126	0.1132	0.1034	0.9822
	2	0.1416	0.1478	0.134	0.9334	0.1762	0.157	0.9802
	3	0.2816	0.273	0.251	0.9476	0.2894	0.2612	0.9848
	4	0.4228	0.4078	0.3798	0.9696	0.4102	0.3704	0.9844
	5	0.61	0.5948	0.573	0.9864	0.5522	0.517	0.9906
30	**0**	0.0464	0.0524	0.0488	0.9974	0.0896	0.1594	0.998
	1	0.0846	0.0908	0.0836	0.9976	0.1264	0.2086	0.999
	2	0.1984	0.1948	0.1872	0.9988	0.2244	0.3132	0.9994
	3	0.3662	0.3696	0.3538	0.9994	0.366	0.462	0.9988
	4	0.5806	0.5718	0.5542	0.9996	0.5354	0.6252	0.9998
	5	0.7744	0.759	0.7462	0.9998	0.705	0.7852	0.9998
40	**0**	0.0532	0.0548	0.0516	1.00	0.0854	0.1162	1.00
	1	0.098	0.1014	0.0968	1.00	0.1356	0.1682	1.00
	2	0.2378	0.2368	0.2282	1.00	0.2606	0.2886	1.00
	3	0.4606	0.4566	0.4426	1.00	0.439	0.4686	1.00
	4	0.7148	0.7054	0.6942	1.00	0.6564	0.6738	1.00
	5	0.8732	0.8644	0.8574	1.00	0.813	0.8244	1.00
50	**0**	0.051	0.051	0.048	1.00	0.094	0.1666	1.00
	1	0.1016	0.1076	0.1024	1.00	0.1478	0.2254	1.00
	2	0.2814	0.2732	0.268	1.00	0.293	0.3802	1.00
	3	0.5596	0.553	0.5444	1.00	0.5214	0.6044	1.00
	4	0.8026	0.7916	0.785	1.00	0.7438	0.8082	1.00
	5	0.9418	0.9376	0.9334	1.00	0.8902	0.9198	1.00
100	**0**	0.0466	0.0476	0.0458	1.00	0.0918	0.1512	1.00
	1	0.1718	0.171	0.1692	1.00	0.2094	0.2788	1.00
	2	0.5138	0.5164	0.5122	1.00	0.4912	0.5718	1.00
	3	0.8468	0.8442	0.8412	1.00	0.7922	0.8312	1.00
	4	0.981	0.9776	0.9768	1.00	0.9468	0.9586	1.00
	5	0.999	0.999	0.9988	1.00	0.9932	0.9946	1.00

Çizelge 4.2.'deki testlerin güçleri incelendiğinde bu yedi test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test Wald-Wolfowitz testidir. Gücü en düşük olan test ise t ($\sigma_1 \neq \sigma_2$) testidir. Çizelge 4.1. ve Çizelge 4.2. birlikte ele alındığında ise Normal dağılımlı örneklemlerde varyans arttıkça testlerin gücünün azaldığı görülmektedir.

Çizelge 4.3.'de üstel dağılımlı iki kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu, elde edilen güçler verilmiştir. Çizelgede, $|\mu_1 - \mu_2| = 0$'a karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 4.3. $X_1, X_2, \cdots, X_n \sim Üstel(\theta_1)$, $Y_1, Y_2, \cdots, Y_n \sim Üstel(\theta_2)$ için testin güçleri

n	$\lvert\mu_1 - \mu_2\rvert$	Z testi	t testi $(\sigma_1 = \sigma_2)$	t testi $(\sigma_1 \neq \sigma_2)$	Kolmogrov-Smirnov	Mann-Whitney	Moods Medyan	Wald-Wolfowitz
10	**0**	0.0496	0.0398	0.0346	0.049	0.0386	0.021	0.039
	1	0.5336	0.2334	0.209	0.2136	0.2188	0.1082	0.0786
	2	0.8762	0.498	0.446	0.454	0.4704	0.275	0.1634
	3	0.9708	0.6954	0.6382	0.6578	0.6722	0.4394	0.2726
	4	0.992	0.7862	0.7344	0.767	0.776	0.5604	0.3828
	5	0.9976	0.8428	0.7922	0.8392	0.8412	0.6444	0.4908
20	**0**	0.0508	0.052	0.0488	0.0346	0.0486	0.0248	0.034
	1	0.774	0.5154	0.5062	0.3242	0.4452	0.2348	0.099
	2	0.984	0.8798	0.8706	0.6972	0.808	0.5438	0.255
	3	0.9994	0.9778	0.9736	0.8866	0.9468	0.7606	0.4512
	4	1.00	0.9936	0.9916	0.9566	0.983	0.8752	0.6158
	5	1.00	0.998	0.9974	0.986	0.9942	0.931	0.7366
30	**0**	0.053	0.0468	0.0456	0.0368	0.0468	0.0682	0.0234
	1	0.8896	0.7206	0.7156	0.4788	0.6058	0.5128	0.0834
	2	0.9984	0.9754	0.9746	0.8838	0.9414	0.8712	0.3036
	3	1.00	0.9982	0.998	0.9786	0.9912	0.9668	0.5532
	4	1.00	1.00	1.00	0.9966	0.999	0.9916	0.7342
	5	1.00	1.00	1.00	0.9992	1.00	0.9968	0.8438
40	**0**	0.0524	0.049	0.0488	0.0522	0.043	0.0384	0.0302
	1	0.9524	0.8448	0.8436	0.6774	0.7386	0.5344	0.1084
	2	0.9996	0.9964	0.9964	0.9726	0.9822	0.9008	0.385
	3	1.00	1.00	1.00	0.9978	0.9982	0.984	0.6532
	4	1.00	1.00	1.00	1.00	0.9998	0.9974	0.8382
	5	1.00	1.00	1.00	0.9998	1.00	0.9996	0.9218
50	**0**	0.0462	0.0446	0.0446	0.0372	0.047	0.0664	0.0332
	1	0.978	0.9238	0.9226	0.7442	0.8456	0.7256	0.1446
	2	1.00	0.9988	0.9988	0.9864	0.9944	0.9702	0.5082
	3	1.00	1.00	1.00	0.9994	1.00	0.9984	0.795
	4	1.00	1.00	1.00	1.00	1.00	0.9996	0.9266
	5	1.00	1.00	1.00	1.00	1.00	1.00	0.9742
100	**0**	0.0488	0.0454	0.045	0.0524	0.043	0.063	0.0372
	1	1.00	0.9976	0.9976	0.9736	0.986	0.9424	0.2666
	2	1.00	1.00	1.00	1.00	1.00	1.00	0.7946
	3	1.00	1.00	1.00	1.00	1.00	1.00	0.9736
	4	1.00	1.00	1.00	1.00	1.00	1.00	0.9988
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Çizelge 4.3.'de testlerin güçleri incelendiğinde bu yedi test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test Z testi, gücü en düşük olan test ise Wald-Wolfowitz testidir. n sayısı arttıkça tüm testlerin gücünün arttığı ve testin gücü bakımından tüm testlerin birbirine yaklaştığı görülmektedir. Ayrıca kitle ortalamaları arasındaki fark arttıkça testlerin gücü artmaktadır. Örnek hacmi çok büyük olduğunda ($n \geq 100$) güç bakımından testler arasında farklılık olmadığı dolayısıyla istenilen testin kullanılabileceği söylenebilir.

Çizelge 4.4.'de β parametreleri aynı olan Gama dağılımlı iki kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, $|\mu_1 - \mu_2| = 0$'a karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 4.4. $X_1, X_2, \cdots, X_n \sim Gama(\alpha_1, 1)$, $Y_1, Y_2, \cdots, Y_n \sim Gama(\alpha_2, 1)$ için testin güçleri

n	$\lvert\mu_1 - \mu_2\rvert$	Z testi	t testi $(\sigma_1 = \sigma_2)$	t testi $(\sigma_1 \neq \sigma_2)$	Kolmogrov-Smirnov	Mann-Whitney	Moods Medyan	Wald-Wolfowitz
10	**0**	0.0474	0.0348	0.0298	0.0454	0.0374	0.0178	0.035
	1	0.5808	0.447	0.433	0.4528	0.486	0.278	0.157
	2	0.9672	0.8818	0.8738	0.8654	0.9018	0.7164	0.515
	3	0.9996	0.9852	0.9848	0.9824	0.9912	0.923	0.8226
	4	1.00	0.9992	0.999	0.9984	0.999	0.9872	0.9486
	5	1.00	1.00	1.00	1.00	1.00	0.998	0.9888
20	**0**	0.0514	0.0454	0.0442	0.0354	0.0486	0.0302	0.0388
	1	0.848	0.7344	0.7316	0.6634	0.8152	0.5478	0.2414
	2	0.9998	0.9942	0.994	0.9868	0.997	0.9644	0.7498
	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	4	1.00	1.00	1.00	1.00	1.00	1.00	0.9978
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00
30	**0**	0.049	0.0464	0.046	0.032	0.0504	0.071	0.0236
	1	0.946	0.8866	0.8864	0.8604	0.9446	0.866	0.2818
	2	1.00	0.9996	0.9996	0.9994	1	0.9994	0.8702
	3	1.00	1.00	1.00	1.00	1.00	1.00	0.9956
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00
40	**0**	0.0474	0.0494	0.048	0.0508	0.0458	0.0408	0.0328
	1	0.9818	0.9536	0.9536	0.965	0.9854	0.904	0.3168
	2	1.00	1.00	1.00	1.00	1.00	1.00	0.9326
	3	1.00	1.00	1.00	1.00	1.00	1.00	0.999
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00
50	**0**	0.0532	0.0498	0.0492	0.0436	0.0524	0.0762	0.0324
	1	0.993	0.9828	0.9828	0.9816	0.9958	0.9728	0.452
	2	1.00	1.00	1.00	1.00	1.00	1.00	0.9818
	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00
100	**0**	0.0462	0.0454	0.0454	0.0514	0.0464	0.0614	0.0338
	1	1.00	1.00	1.00	1.00	1.00	1.00	0.7416
	2	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Çizelge 4.4.'deki testlerin güçleri incelendiğinde bu yedi test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test Z testi, gücü en düşük olan test ise Wald-Wolfowitz testidir. n sayısı arttıkça tüm testlerin gücünün arttığı

çizelgeden görülmektedir. Ayrıca parametreler arasındaki fark arttıkça testlerin gücü artmaktadır. Örnek hacmi çok büyük olduğunda ($n \geq 100$) güç bakımından testler arasında farklılık olmadığı dolayısıyla istenilen testin kullanılabileceği söylenebilir.

4.2. Orana İlişkin Testler

Bu bölümde bağımsız iki örneklem testlerinden orana ilişkin parametrik Z testi ile parametrik olmayan Ki-kare testi tanıtılmıştır.

4.2.1. Z testi

$X_1, X_2, ..., X_m$ ve $Y_1, Y_2, ..., Y_n$ iki bağımsız örneklem ve π_1, π_2 de birbirinden bağımsız kitle oranları olmak üzere $H_0 : \pi_1 - \pi_2 = \pi_0$ hipotezini $H_1 : \pi_1 - \pi_2 \neq \pi_0$ hipotezine karşı test etmektedir ve test istatistiği,

$$Z_{Hesap} = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{\pi_1(1-\pi_1)}{m} + \frac{\pi_2(1-\pi_2)}{n}}} \tag{4.11}$$

şeklinde hesaplanmaktadır. (4.11) eşitliğinde bulunan p_1 ve p_2 örnek oranları, n ve m örnek hacimleridir.

$\frac{(p_1 - p_2) - (\pi_0)}{\sqrt{\frac{\pi_1(1-\pi_1)}{m} + \frac{\pi_2(1-\pi_2)}{n}}} < Z_{\alpha/2}$ olduğu durumda H_0 reddedilmektedir.

4.2.2. Ki-kare testi

Ki-kare testi iki veya daha fazla birbirinden bağımsız örneklemin aynı kitleden alınıp alınmadıklarını analiz etmek için uygulanmaktadır. Aynı zamanda ki-kare bağımsızlık testi olarak da adlandırılmaktadır ve test istatistiği;

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{\left(G_{ij} - B_{ij}\right)^2}{B_{ij}} \sim \chi^2_{(k-1)(r-1)} \tag{4.12}$$

şeklinde hesaplanmaktadır. (4.12) eşitliğindeki, G_{ij}, gözlenen değeri, B_{ij}, beklenen değeri ve k, grup sayısını ifade etmektedir.

$2x2$'lik tablolar olduğunda süreklilik düzeltmesi ile birlikte kullanılmalıdır (Önder).

4.2.3. Güç karşılaştırması

Orana ilişkin ele alınan parametrik Z testi ve parametrik olmayan Ki-kare testlerinin güç karşılaştırmalarının yapılması amacıyla 10000 tekrarlı Monte-Carlo simülasyon yöntemi uygulanmıştır. Testlerin güçleri elde edilirken hipotezler çift yönlü kurulmuştur ve $\alpha = 0.05$ anlamlılık seviyesi ele alınmıştır. Farklı parametreler ve farklı n değerleri için elde edilen sonuçlar Çizelge 4.7 – 4.12'de verilmiştir. Çizelgelerde, $P_1 = P_2$ olan durumlardaki değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 4.7. $n = 10$ için güçler

P_1	P_2	Z Testi	Ki-Kare Testi	P_1	P_2	Z Testi	Ki-Kare Testi
	0.1	0.0096	0.1254		**0.1**	0.6612	0.4748
	0.2	0.0463	0.0551		**0.2**	0.4229	0.2456
	0.3	0.143	0.0722		**0.3**	0.2441	0.1197
	0.4	0.2972	0.157		**0.4**	0.1302	0.0516
0.1	**0.5**	0.47	0.2917	**0.6**	**0.5**	0.0608	0.0198
	0.6	0.6629	0.4766		**0.6**	0.0397	0.0109
	0.7	0.8199	0.6682		**0.7**	0.0604	0.0222
	0.8	0.9408	0.8461		**0.8**	0.1315	0.0581
	0.9	0.9891	0.9575		**0.9**	0.2936	0.1588
	0.1	0.0486	0.0466		**0.1**	0.8246	0.6759
	0.2	0.028	0.0167		**0.2**	0.6257	0.4346
	0.3	0.0523	0.0208		**0.3**	0.4248	0.2501
	0.4	0.1237	0.0587		**0.4**	0.2554	0.1271
0.2	**0.5**	0.2514	0.1242	**0.7**	**0.5**	0.1293	0.0576
	0.6	0.4215	0.2431		**0.6**	0.0549	0.018
	0.7	0.6154	0.4182		**0.7**	0.0346	0.0134
	0.8	0.8018	0.6317		**0.8**	0.0553	0.0225
	0.9	0.9367	0.8394		**0.9**	0.1433	0.0658
	0.1	0.1437	0.0664		**0.1**	0.9371	0.8425
	0.2	0.0563	0.0216		**0.2**	0.7965	0.6313
	0.3	0.0379	0.0141		**0.3**	0.6173	0.4258
	0.4	0.0589	0.0218		**0.4**	0.4327	0.2599
0.3	**0.5**	0.1281	0.0565	**0.8**	**0.5**	0.2633	0.1408
	0.6	0.2469	0.1215		**0.6**	0.1288	0.059
	0.7	0.4211	0.2429		**0.7**	0.0628	0.0254
	0.8	0.6279	0.4292		**0.8**	0.0286	0.0211
	0.9	0.8248	0.6673		**0.9**	0.0466	0.0545
	0.1	0.2929	0.1573		**0.1**	0.9892	0.9575
	0.2	0.1381	0.063		**0.2**	0,9348	0.841
	0.3	0.0628	0.0238		**0.3**	0.8275	0.6638
	0.4	0.04	0.0147		**0.4**	0.6689	0.4735
0.4	**0.5**	0.0666	0.0226	**0.9**	**0.5**	0.4757	0.2988
	0.6	0.1339	0.0565		**0.6**	0.2908	0.1519
	0.7	0.2482	01222		**0.7**	0.1407	0.0667
	0.8	0.4227	0.2543		**0.8**	0.0466	0.05
	0.9	0.6585	0.4704		**0.9**	0.0106	0.1227
	0.1	0.4743	0.2943				
	0.2	0.2554	0.1306				
	0.3	0.133	0.0574				
	0.4	0.065	0.0206				
0.5	**0.5**	0.0432	0.0162				
	0.6	0.0642	0.0212				
	0.7	0.131	0.0568				
	0.8	0.2574	0.1335				
	0.9	0.4676	0.2936				

Çizelge 4.7.'deki güçler incelendiğinde Z testinin gücünün yüksek çıktığı görülmektedir. Kitle oranları arasındaki fark büyüdükçe iki testin de gücü artmakta ve birbirine yaklaşmaktadır.

Çizelge 4.8. $n = 20$ için güçler

P_1	P_2	Z Testi	Ki-Kare Testi	P_1	P_2	Z Testi	Ki-Kare Testi
	0.1	0.0368	0.0482		**0.1**	0.8443	0.7442
	0.2	0.0385	0.0095		**0.2**	0.6163	0.4726
	0.3	0.1514	0.042		**0.3**	0.3651	0.2383
	0.4	0.3601	0.1932		**0.4**	0.182	0.0966
0.1	**0.5**	0.6026	0.4281	**0.6**	**0.5**	0.0771	0.0323
	0.6	0.8151	0.6775		**0.6**	0.0463	0.018
	0.7	0.9378	0.8575		**0.7**	0.088	0.047
	0.8	0.9861	0.9618		**0.8**	0.2253	0.1346
	0.9	0.9993	0.9967		**0.9**	0.4676	0.3094
	0.1	0.1213	0.0421		**0.1**	0.945	0.897
	0.2	0.0353	0.0124		**0.2**	0.7978	0.6841
	0.3	0.0607	0.02		**0.3**	0.58	0.4222
	0.4	0.1689	0.078		**0.4**	0.3613	0.2167
0.2	**0.5**	0.3378	0.2113	**0.7**	**0.5**	0.1793	0.0881
	0.6	0.5792	0.4081		**0.6**	0.0724	0.0317
	0.7	0.7832	0.6416		**0.7**	0.044	0.0172
	0.8	0.9239	0.8489		**0.8**	0.0918	0.0456
	0.9	0.9894	0.9721		**0.9**	0.2713	0.1396
	0.1	0.2733	0.1415		**0.1**	0.9882	0.9721
	0.2	0.0989	0.0476		**0.2**	0.9244	0.8522
	0.3	0.0462	0.0184		**0.3**	0.786	0.6367
	0.4	0.068	0.0308		**0.4**	0.5643	0.396
0.3	**0.5**	0.1705	0.0838	**0.8**	**0.5**	0.3356	0.2092
	0.6	0.3411	0.2008		**0.6**	0.1672	0.0822
	0.7	0.5712	0.4182		**0.7**	0.0599	0.0191
	0.8	0.7921	0.6777		**0.8**	0.04	0.0122
	0.9	0.9464	0.8896		**0.9**	0.1268	0.0487
	0.1	0.4663	0.306		**0.1**	0.9989	0.9963
	0.2	0.2224	0.136		**0.2**	0.9883	0.962
	0.3	0.0872	0.0454		**0.3**	0.9395	0.8651
	0.4	0.05	0.0209		**0.4**	0.818	0.6861
0.4	**0.5**	0.0826	0.035	**0.9**	**0.5**	0.6119	0.4301
	0.6	0.1885	0.0985		**0.6**	0.3694	0.1879
	0.7	0.3678	0.2459		**0.7**	0.1548	0.0466
	0.8	0.6137	0.4819		**0.8**	0.0365	0.0093
	0.9	0.8485	0.7507		**0.9**	0.0395	0.0462
	0.1	0.6827	0.5382				
	0.2	0.4083	0.287				
	0.3	0.1977	0.1146				
	0.4	0.0852	0.041				
0.5	**0.5**	0.0509	0.0192				
	0.6	0.0848	0.0399				
	0.7	0.1959	0.1164				
	0.8	0.4118	0.2814				
	0.9	0.6799	0.5388				

Çizelge 4.8.'deki güçler incelendiğinde Z testinin gücünün yüksek çıktığı görülmektedir. Kitle oranları arasındaki fark büyüdükçe iki testin de gücü artmakta ve birbirine yaklaşmaktadır.

Çizelge 4.9. $n = 30$ için güçler

P_1	P_2	Z Testi	Ki-Kare Testi	P_1	P_1	Z Testi	Ki-Kare Testi
	0.1	0.03	0.0255		**0.1**	0.899	0.8389
	0.2	0.0448	0.0053		**0.2**	0.6677	0.5615
	0.3	0.2106	0.0567		**0.3**	0.4011	0.2862
	0.4	0.4335	0.2332		**0.4**	0.2054	0.1191
0.1	**0.5**	0.678	0.4939	**0.6**	**0.5**	0.0911	0.0389
	0.6	0.8682	0.7514		**0.6**	0.0529	0.0218
	0.7	0.9638	0.9146		**0.7**	0.0977	0.0552
	0.8	0.9944	0.9836		**0.8**	0.259	0.1742
	0.9	0.9996	0.9993		**0.9**	0.5651	0.4445
	0.1	0.1434	0.0779		**0.1**	0.9689	0.9464
	0.2	0.0414	0.0148		**0.2**	0.8489	0.7678
	0.3	0.0775	0.0203		**0.3**	0.6395	0.494
	0.4	0.1865	0.087		**0.4**	0.4009	0.2593
0.2	**0.5**	0.3937	0.2397	**0.7**	**0.5**	0.1931	0.1013
	0.6	0.6359	0.4756		**0.6**	0.076	0.0293
	0.7	0.84	0.7242		**0.7**	0.049	0.0187
	0.8	0.9562	0.9128		**0.8**	0.109	0.0655
	0.9	0.9943	0.9889		**0.9**	0.3321	0.2259
	0.1	0.3406	0.2335		**0.1**	0.9954	0.9909
	0.2	0.1176	0.0693		**0.2**	0.9531	0.9127
	0.3	0.0514	0.0204		**0.3**	0.8405	0.7294
	0.4	0.0778	0.0318		**0.4**	0.6488	0.4724
0.3	**0.5**	0.1983	0.1025	**0.8**	**0.5**	0.3932	0.2443
	0.6	0.394	0.2586		**0.6**	0.1859	0.0876
	0.7	0.6242	0.4912		**0.7**	0.0765	0.021
	0.8	0.8388	0.7587		**0.8**	0.0421	0.0153
	0.9	0.9686	0.9444		**0.9**	0.1321	0.0684
	0.1	0.5634	0.4447		**0.1**	0.9997	0.9993
	0.2	0.2628	0.1803		**0.2**	0.9955	0.9866
	0.3	0.1043	0.0586		**0.3**	0.9656	0.9162
	0.4	0.0519	0.0222		**0.4**	0.8697	0.7506
0.4	**0.5**	0.0906	0.0429	**0.9**	**0.5**	0.6821	0.5002
	0.6	0.2056	0.1158		**0.6**	0.4283	0.2364
	0.7	0.405	0.2876		**0.7**	0.2108	0.0567
	0.8	0.66	0.5544		**0.8**	0.0495	0.0037
	0.9	0.8955	0.8345		**0.9**	0.0273	0.024
	0.1	0.7627	0.6676				
	0.2	0.465	0.3597				
	0.3	0.2257	0.1407				
	0.4	0.0944	0.0437				
0.5	**0.5**	0.0586	0.0224				
	0.6	0.0907	0.0445				
	0.7	0.2179	0.1352				
	0.8	0.4527	0.3499				
	0.9	0.7557	0.6561				

Çizelge 4.9.'daki güçler incelendiğinde Z testinin gücünün yüksek çıktığı görülmektedir. Kitle oranları arasındaki fark büyüdükçe iki testin de gücü artmakta ve birbirine yaklaşmaktadır.

Çizelge 4.10. $n = 40$ için güçler

P_1	P_2	Z Testi	Ki-Kare Testi	P_1	P_2	Z Testi	Ki-Kare Testi
	0.1	0.049	0.0198		**0.1**	0.9233	0.8587
	0.2	0.0307	0.0025		**0.2**	0.699	0.5797
	0.3	0.209	0.0427		**0.3**	0.4283	0.3104
	0.4	0.4483	0.2224		**0.4**	0.2171	0.1217
0.1	**0.5**	0.7155	0.5024	**0.6**	**0.5**	0.0948	0.037
	0.6	0.8965	0.7763		**0.6**	0.055	0.0196
	0.7	0.9744	0.9375		**0.7**	0.0998	0.054
	0.8	0.9955	0.9879		**0.8**	0.2775	0.1836
	0.9	0.9998	0.9994		**0.9**	0.6179	0.4719
	0.1	0.1759	0.0927		**0.1**	0.9786	0.9562
	0.2	0.044	0.0163		**0.2**	0.8686	0.7848
	0.3	0.0708	0.0131		**0.3**	0.6568	0.5277
	0.4	0.1961	0.079		**0.4**	0.4249	0.275
0.2	**0.5**	0.4203	0.2405	**0.7**	**0.5**	0.213	0.1009
	0.6	0.6717	0.4966		**0.6**	0.0816	0.0276
	0.7	0.8606	0.7656		**0.7**	0.0487	0.0182
	0.8	0.9625	0.925		**0.8**	0.1228	0.0658
	0.9	0.9967	0.9907		**0.9**	0.3872	0.2495
	0.1	0.3874	0.2514		**0.1**	0.9963	0.9915
	0.2	0.1269	0.0683		**0.2**	0.9651	0.9305
	0.3	0.0515	0.0163		**0.3**	0.8641	0.7732
	0.4	0.08	0.0264		**0.4**	0.6795	0.5021
0.3	**0.5**	0.2037	0.1022	**0.8**	**0.5**	0.4231	0.2389
	0.6	0.4221	0.2683		**0.6**	0.1945	0.0771
	0.7	0.6656	0.5341		**0.7**	0.0732	0.0154
	0.8	0.8652	0.7855		**0.8**	0.0449	0.017
	0.9	0.9775	0.9562		**0.9**	0.1741	0.0898
	0.1	0.6092	0.4626		**0.1**	0.9998	0.9994
	0.2	0.2768	0.1892		**0.2**	0.9971	0.9911
	0.3	0.0987	0.0508		**0.3**	0.9757	0.94
	0.4	0.0533	0.0185		**0.4**	0.901	0.7788
0.4	**0.5**	0.0967	0.0397	**0.9**	**0.5**	0.72	0.5123
	0.6	0.2274	0.127		**0.6**	0.4552	0.2273
	0.7	0.4323	0.3152		**0.7**	0.2118	0.0423
	0.8	0.6989	0.5877		**0.8**	0.0344	0.0019
	0.9	0.9231	0.8604		**0.9**	0.0436	0.0182
	0.1	0.8036	0.6874				
	0.2	0.4893	0.3687				
	0.3	0.2242	0.1454				
	0.4	0.0956	0.0481				
0.5	**0.5**	0.0532	0.0193				
	0.6	0.093	0.0483				
	0.7	0.227	0.1478				
	0.8	0.4817	0.362				
	0.9	0.804	0.6927				

Çizelge 4.10.'daki güçler incelendiğinde Z testinin gücünün yüksek çıktığı görülmektedir. Kitle oranları arasındaki fark büyüdükçe iki testin de gücü artmakta ve birbirine yaklaşmaktadır.

Çizelge 4.11. $n = 50$ için güçler

P_1	P_2	Z Testi	Ki-Kare Testi	P_1	P_2	Z Testi	Ki-Kare Testi
	0.1	0.0458	0.0174		**0.1**	0.9336	0.8788
	0.2	0.0245	0.0016		**0.2**	0.7367	0.6148
	0.3	0.1999	0.05		**0.3**	0.4482	0.3147
	0.4	0.4619	0.2659		**0.4**	0.2266	0.1337
0.1	**0.5**	0.7368	0.5682	**0.6**	**0.5**	0.0951	0.0426
	0.6	0.9192	0.8312		**0.6**	0.0506	0.0208
	0.7	0.9798	0.9486		**0.7**	0.1053	0.0569
	0.8	0.9974	0.9927		**0.8**	0.302	0.2018
	0.9	1.00	0.9997		**0.9**	0.6604	0.4925
	0.1	0.1997	0.095		**0.1**	0.9836	0.9634
	0.2	0.0439	0.015		**0.2**	0.8885	0.8112
	0.3	0.0687	0.0164		**0.3**	0.6767	0.5524
	0.4	0.1974	0.0914		**0.4**	0.435	0.2922
0.2	**0.5**	0.4451	0.2835	**0.7**	**0.5**	0.2128	0.1134
	0.6	0.6931	0.5464		**0.6**	0.0796	0.0332
	0.7	0.8752	0.7831		**0.7**	0.0499	0.0185
	0.8	0.9715	0.9428		**0.8**	0.1317	0.0733
	0.9	0.9978	0.9936		**0.9**	0.4379	0.2673
	0.1	0.4324	0.2592		**0.1**	0.9978	0.9951
	0.2	0.1258	0.0708		**0.2**	0.9743	0.9442
	0.3	0.0496	0.0185		**0.3**	0.8838	0.7957
	0.4	0.0815	0.0329		**0.4**	0.6999	0.5541
0.3	**0.5**	0.2167	0.1201	**0.8**	**0.5**	0.4428	0.2876
	0.6	0.4411	0.2996		**0.6**	0.2058	0.1012
	0.7	0.6821	0.5534		**0.7**	0.0744	0.0172
	0.8	0.8936	0.8131		**0.8**	0.0422	0.0142
	0.9	0.9866	0.9636		**0.9**	0.1984	0.0922
	0.1	0.654	0.4853		**0.1**	0.9998	0.9997
	0.2	0.294	0.197		**0.2**	0.9985	0.9937
	0.3	0.1026	0.0559		**0.3**	0.9816	0.954
	0.4	0.0506	0.0206		**0.4**	0.9134	0.8277
0.4	**0.5**	0.1007	0.0462	**0.9**	**0.5**	0.7289	0.5586
	0.6	0.2281	0.1368		**0.6**	0.4482	0.2579
	0.7	0.4539	0.3237		**0.7**	0.1963	0.0523
	0.8	0.7363	0.6196		**0.8**	0.0232	0.0016
	0.9	0.9392	0.8766		**0.9**	0.0437	0.0168
	0.1	0.8331	0.71				
	0.2	0.5242	0.3972				
	0.3	0.2442	0.1546				
	0.4	0.0993	0.0456				
0.5	**0.5**	0.0567	0.0221				
	0.6	0.0994	0.0512				
	0.7	0.2403	0.1501				
	0.8	0.5213	0.3932				
	0.9	0.8351	0.7081				

Çizelge 4.11.'deki güçler incelendiğinde Z testinin gücünün yüksek çıktığı görülmektedir. Kitle oranları arasındaki fark büyüdükçe iki testin de gücü artmakta ve birbirine yaklaşmaktadır.

Çizelge 4.12. $n = 100$ için güçler

P_1	P_2	Z Testi	Ki-Kare Testi	P_1	P_2	Z Testi	Ki-Kare Testi
	0.1	0.0387	0.0157		**0.1**	0.9607	0.9159
	0.2	0.0099	0.0009		**0.2**	0.7958	0.6887
	0.3	0.2249	0.027		**0.3**	0.5016	0.3772
	0.4	0.4979	0.3012		**0.4**	0.2412	0.1483
0.1	**0.5**	0.7892	0.6384	**0.6**	**0.5**	0.0849	0.0429
	0.6	0.9386	0.8775		**0.6**	0.0493	0.0196
	0.7	0.9922	0.9774		**0.7**	0.1168	0.0609
	0.8	0.9995	0.9982		**0.8**	0.3444	0.2312
	0.9	1.00	0.9999		**0.9**	0.7054	0.5661
	0.1	0.213	0.1194		**0.1**	0.9923	0.9802
	0.2	0.0411	0.0167		**0.2**	0.9299	0.8715
	0.3	0.0769	0.0112		**0.3**	0.741	0.6213
	0.4	0.213	0.0987		**0.4**	0.4764	0.3429
0.2	**0.5**	0.463	0.3172	**0.7**	**0.5**	0.2216	0.1242
	0.6	0.7351	0.6016		**0.6**	0.0761	0.0297
	0.7	0.9159	0.8477		**0.7**	0.05	0.0152
	0.8	0.9863	0.9698		**0.8**	0.1488	0.084
	0.9	0.9991	0.9977		**0.9**	0.4668	0.3211
	0.1	0.4738	0.3254		**0.1**	0.9993	0.9978
	0.2	0.1487	0.0832		**0.2**	0.9867	0.971
	0.3	0.0517	0.0176		**0.3**	0.9157	0.8486
	0.4	0.0811	0.0345		**0.4**	0.7397	0.605
0.3	**0.5**	0.2234	0.1296	**0.8**	**0.5**	0.4585	0.314
	0.6	0.4724	0.3323		**0.6**	0.2108	0.1014
	0.7	0.7353	0.62		**0.7**	0.073	0.0104
	0.8	0.9275	0.8712		**0.8**	0.0402	0.0176
	0.9	0.9908	0.9784		**0.9**	0.2189	0.1189
	0.1	0.7086	0.563		**0.1**	1.00	1.00
	0.2	0.3431	0.2269		**0.2**	0.9989	0.9974
	0.3	0.113	0.0594		**0.3**	0.991	0.9761
	0.4	0.0463	0.0178		**0.4**	0.9394	0.8788
0.4	**0.5**	0.0907	0.0433	**0.9**	**0.5**	0.7769	0.6305
	0.6	0.2461	0.1467		**0.6**	0.4948	0.2847
	0.7	0.5124	0.3806		**0.7**	0.218	0.0272
	0.8	0.7988	0.6885		**0.8**	0.0109	0.001
	0.9	0.9621	0.9187		**0.9**	0.0422	0.0162
	0.1	0.874	0.7798				
	0.2	0.5831	0.4481				
	0.3	0.2678	0.1749				
	0.4	0.1028	0.0522				
0.5	**0.5**	0.054	0.024				
	0.6	0.1063	0.058				
	0.7	0.2751	0.1813				
	0.8	0.5854	0.453				
	0.9	0.8742	0.7847				

Çizelge 4.12.'deki güçler incelendiğinde Z testinin gücünün yüksek çıktığı görülmektedir. Kitle oranları arasındaki fark büyüdükçe iki testin de gücü artmakta ve birbirine yaklaşmaktadır.

Çizelge 4.7.- 4.12 birlikte değerlendirildiğinde, örneklem sayısı arttıkça testlerin gücünün arttığı görülmektedir.

5. BAĞIMLI İKİ ÖRNEKLEM TESTLERİ

5.1. Ortalama ve Medyana İlişkin Testler

Bu bölümde bağımlı iki örneklem testlerinden ortalama ve medyana ilişkin parametrik t testi ile parametrik olmayan Wilcoxon t testi ve işaret testi tanıtılmıştır.

5.1.1. t testi (Paired t testi)

Bağımlı iki örneklem t testi bir gruptaki n birimden, farklı işlemlerden elde edilmiş iki veri setinin farklarının sıfır ortalamalı kitlenin rastgele örneği olup olmadığını test etmektedir.

$H_0 : \mu = 0$ hipotezi, $H_1 : \mu \neq 0$ hipotezine karşı test ediliyor ise test istatistiği,

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2 - 2p\sigma_1\sigma_2}{n}}} \tag{5.1}$$

(5.1) eşitliğinde, σ_1^2 birinci grup değişkeninin varyansı, σ_2^2, ikinci grup değişkeninin varyansı ve p iki değişken için korelasyon katsayısıdır. İki grubun bağımsız olmasından kaynaklı iki grubun korelasyon katsayısı 0'dır. Eğer denklemde korelasyon katsayısı 0'dan büyükse t istatistiği artar. Bu da iki bağımlı t testinin bağımsız bir t testine kıyasla gücünü artırır. Diğer bir deyişle korelasyon katsayısı 0'dan az ise, istatistiksel güç azalır ve bağımsız bir t testinden daha düşük hale gelir (Kim, 2015).

5.1.2. Wilcoxon t testi

$(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$ iki bağımlı değişkenden oluşan gözlemler olmak üzere, $D_i = X_i - Y_i, \quad i = 1, 2, ..., n$, n sayıda farklardan oluşur ve bu farklar sürekli bir rastgele değişkendir. Farklar simetrik dağılmaktadır. Ölçme düzeyi ise en az eşit aralıklıdır (Gamgam ve Altunkaynak, 2012).

δ_i rastgele değişkeni, $\delta_i = \begin{cases} 1, & D_i > 0 \\ 0, & D_i < 0 \end{cases}$

şeklinde tanımlanmaktadır ve test istatistiği,

$$T^{+}=\sum_{i=1}^{n}\delta_i r\left(\left|D_i\right|\right) \tag{5.2}$$

şeklinde hesaplanmaktadır (5.2) eşitliğindeki $r\left(\left|D_i\right|\right)$ sayı sıralı istatistiktir ve $T_{Hesap}^{+}\le T_{Tablo}$ ise H_0 reddedilir.

5.1.3. İşaret testi

$\left(X_1,Y_1\right),\left(X_2,Y_2\right),...,\left(X_n,Y_n\right)$ iki bağımlı değişkenden oluşan gözlemlerdir. Bu gözlemler aynı kitleden veya farklı kitleden gelebilirler. Gözlem çiftlerinin eşlenik olması dikkat edilmesi gereken bir noktadır. Bu testte veriler bağımsız, rassal ve ölçme düzeyi en az sıralamadır (Karagöz, 2010).

$D_i=X_i-Y_i,\quad i=1,2,...,n$ olarak ve

$\delta_i=\begin{cases}1, & D_i>0\\ 0, & D_i<0\end{cases}$ şeklinde tanımlanmak üzere test istatistiği,

$$k_{Hesap}=\sum_{i=1}^{n}\delta_i \tag{5.3}$$

şeklinde hesaplanmaktadır ve $k_{Hesap}\le k_{\alpha/2}$ ise H_0 reddedilir.

5.1.4. Güç karşılaştırmaları

Eşit varyanslı ve düşük, orta ile yüksek korelasyonlu iki bağımlı örneklem için $\alpha=0.05$ anlamlılık seviyesinde yapılan 10000 tekrarlı simülasyonlar sonucunda elde edilen testlerin güçleri sırasıyla Çizelge 5.1., Çizelge 5.2. ve Çizelge 5.3.'te verilmiştir. Çizelge 5.4.'te ise farklı varyanslı ve yüksek ilişkili bağımlı örneklemler için $\alpha=0.05$ anlamlılık seviyesinde yapılan 10000 tekrarlı simülasyon sonucunda elde edilen testlerin güçleri verilmiştir. Tüm simülasyonlarda iki örneklemin de normal dağılımdan

geldiği durum ele alınmıştır. Çizelgelerde, $|\mu_1 - \mu_2| = 0$'a karşılık gelen değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 5.1.'de varyansları eşit, düşük korelasyonlu normal dağılımlı iki kitle için farklı ortalama farkları ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir.

Çizelge 5.1. $\sigma_X^2 = 4, \sigma_Y^2 = 4, \sigma_{XY} = 1$ durumu için güçler

n	$\lvert\mu_1 - \mu_2\rvert$	t testi	Wilcoxon	İşaret testi
10	**0**	0.0505	0.0502	0.0243
	1	0.2159	0.2113	0.096
	2	0.6355	0.6169	0.36
	3	0.9333	0.9215	0.6942
	4	0.9941	0.992	0.9121
	5	0.9999	0.9998	0.9824
20	**0**	0.0499	0.049	0.0434
	1	0.4114	0.3931	0.2722
	2	0.9322	0.919	0.7796
	3	0.9998	0.9993	0.9833
	4	1.00	1.00	0.9995
	5	1.00	1.00	1.00
30	**0**	0.0509	0.0497	0.0449
	1	0.5768	0.5579	0.3881
	2	0.9911	0.9888	0.9288
	3	0.9999	0.9999	0.9992
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
40	**0**	0.0483	0.0494	0.0394
	1	0.708	0.6895	0.4911
	2	0.9994	0.9988	0.9728
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
50	**0**	0.0489	0.0474	0.0305
	1	0.8053	0.7837	0.5586
	2	1.00	1.00	0.9915
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
100	**0**	0.0489	0.048	0.0354
	1	0.9829	0.9782	0.864
	2	1.00	1.00	1.00
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00

Çizelge 5.2.'de varyansları eşit, orta düzeyde korelasyonlu normal dağılımlı iki kitle için farklı ortalama farkları ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir.

Çizelge 5.2. $\sigma_X^2 = 4, \sigma_Y^2 = 4, \sigma_{XY} = 2$ durumu için güçler

n	$\lvert\mu_1 - \mu_2\rvert$	t testi	Wilcoxon	İşaret testi
10	**0**	0.0517	0.0499	0.0217
	1	0.2868	0.2754	0.1311
	2	0.7976	0.7823	0.5097
	3	0.9865	0.9825	0.8651
	4	0.9999	0.9997	0.98
	5	1.00	1.00	0.9982
20	**0**	0.0496	0.0483	0.0411
	1	0.5714	0.5499	0.3879
	2	0.9888	0.9856	0.9186
	3	1.00	1.00	0.999
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
30	**0**	0.0497	0.0472	0.0431
	1	0.757	0.7363	0.5466
	2	1.00	0.9998	0.9857
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
40	**0**	0.0487	0.0483	0.0375
	1	0.8729	0.8564	0.6564
	2	0.9999	0.9999	0.9976
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
50	**0**	0.0498	0.0471	0.0328
	1	0.936	0.924	0.7408
	2	1.00	1.00	0.9998
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
100	**0**	0.0502	0.0495	0.0355
	1	0.9986	0.9976	0.9653
	2	1.00	1.00	1.00
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00

Çizelge 5.3.'de varyansları eşit, yüksek korelasyonlu normal dağılımlı iki kitle için farklı ortalama farkları ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir.

Çizelge 5.3. $\sigma_X^2 = 4, \sigma_Y^2 = 4, \sigma_{XY} = 3$ durumu için güçler

n	$\lvert\mu_1 - \mu_2\rvert$	t testi	Wilcoxon	İşaret testi
10	**0**	0.0479	0.0457	0.0219
	1	0.5151	0.497	0.2687
	2	0.9769	0.9709	0.82
	3	1.00	0.9997	0.9893
	4	1.00	1.00	0.9993
	5	1.00	1.00	1.00
20	**0**	0.0461	0.0465	0.0413
	1	0.8561	0.8371	0.66
	2	1.00	1.00	0.996
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
30	**0**	0.0523	0.0502	0.0441
	1	0.9596	0.9542	0.8331
	2	1.00	1.00	1.00
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
40	**0**	0.0514	0.0502	0.0392
	1	0.9918	0.9889	0.9255
	2	1.00	1.00	1.00
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
50	**0**	0.0487	0.0495	0.0304
	1	0.9979	0.9968	0.9649
	2	1.00	1.00	1.00
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
100	**0**	0.0503	0.0499	0.0331
	1	1.00	1.00	0.9999
	2	1.00	1.00	1.00
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00

Çizelge 5.4.'de varyansları farklı normal dağılımlı iki kitle için farklı ortalama farkları ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir.

Çizelge 5.4. $\sigma_X^2 = 4, \sigma_Y^2 = 10, \sigma_{XY} = 5$ durumu için güçler

n	$\lvert\mu_1 - \mu_2\rvert$	t testi	Wilcoxon	İşaret testi
10	**0**	0.0469	0.0471	0.0209
	1	0.2879	0.2796	0.1385
	2	0.7992	0.78	0.5157
	3	0.988	0.9854	0.8733
	4	0.9999	0.9997	0.9805
	5	1.00	1.00	0.9978
20	**0**	0.0528	0.0497	0.0432
	1	0.567	0.543	0.3926
	2	0.9884	0.9835	0.9167
	3	1.00	1.00	0.9987
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
30	**0**	0.052	0.0543	0.0457
	1	0.7512	0.7346	0.5506
	2	0.9993	0.9992	0.987
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
40	**0**	0.0505	0.0512	0.0403
	1	0.8669	0.8519	0.6577
	2	1.00	1.00	0.9976
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
50	**0**	0.0523	0.0503	0.0344
	1	0.9357	0.9279	0.7424
	2	1.00	1.00	0.9996
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00
100	**0**	0.048	0.0467	0.0341
	1	0.9987	0.9979	0.9654
	2	1.00	1.00	1.00
	3	1.00	1.00	1.00
	4	1.00	1.00	1.00
	5	1.00	1.00	1.00

Çizelge 5.1.-5.4 incelendiğinde küçük örneklem değerlerinde kitle ortalamaları arasındaki fark küçük olduğunda t testinin gücü en yüksek çıkmaktadır. Kitle ortalamaları arasındaki fark büyüdükçe tüm testlerin gücü artmakta ve birbirine yaklaşmaktadır. Bununla birlikte tüm durumlarda gücü en düşük olan test işaret testidir.

Örneklem sayısı arttıkça testlerin gücü artmaktadır. Bununla beraber büyük örneklemlerde de kitle ortalamaları arasındaki fark azaldıkça testlerin gücü azalmaktadır. Büyük örneklem durumunda da yine en güçlü test t testi, gücü en düşük olan test işaret testidir.

Çizelge 5.1.-5.3. birlikte incelendiğinde, varyansların eşit olduğu durumda gruplar arasındaki ilişki arttıkça testlerin güçlerinin de arttığı görülmektedir.

Çizelge 5.3. ve Çizelge 5.4. birlikte incelendiğinde, iki örneklem arasındaki ilişki artmasına rağmen eşit varyans durumunda testlerin güçlerinin farklı varyans durumundaki testlerin güçlerinden daha yüksek olduğu görülmektedir.

5.2. Orana İlişkin Testler

Bu bölümde bağımlı iki örneklem testlerinden orana ilişkin parametrik t testi ile parametrik olmayan Mc-Nemar testi tanıtılmıştır.

5.2.1. t testi

Sayısal bir değişken için, aynı kişilerden farklı zamanlarda veya farklı durumlarda elde edilen iki oran arasında farklılık bulunup bulunmadığını incelemektedir. Bu durum aşağıdaki tabloda verildiği gibi özetlenebilir.

Önce	**Sonra**		
	+	-	**Toplam**
+	x	y	**x+y**
-	z	t	**z+t**
Toplam	**x+z**	**y+t**	**x+y+z+t=n**

t testi için test istatistiği,

$$T = \left(p_1 - p_2 \right) / S^d \tag{5.4}$$

$$S^d = \sqrt{y+z} / n \tag{5.5}$$

şeklinde tanımlanmaktadır. (5.4) eşitliğinde p_1, 1. grup (önceki) için olumlu değerlendirilen durum, p_2, 2. grup (sonraki) için olumlu değerlendirilen durum, y, 1.grupta olumlu iken 2. grupta olumsuz olanların oranı, z, 1.grupta olumsuz iken 2. grupta olumlu olanların oranı, n, genel toplamdır ve $t_{Hesap} < t_{Tablo}$ ise H_0 reddedilir.

5.2.2. Mc-Nemar testi

Bu test, bağımlı iki örneklem ki-kare testidir. Deneklerin (denemelerin) öncelik sonralık durumuna göre gözlemlerini, tepkilerini ve durumlarını herhangi bir değişiklik olup olmadığını önceki sonraki durum arasında uyum bulunup bulunmadığını test etmektedir $H_0 : y = z$ hipotezi $H_1 : y \neq z$ hipotezine karşı test ediliyor ise test istatistiği,

$$\chi^2 = \left(y - z \right)^2 / y + z \tag{5.6}$$

$\chi^2 \geq \chi^2_{\alpha,1}$ ise H_0 reddedilir. (5.6) eşitliğinde, y, önceki durum olumlu iken sonraki durum olumsuz olanların birim sayısı, z, önceki durum olumsuz iken sonraki durum olumlu olanların birim sayısıdır.

5.2.3. Güç karşılaştırmaları

Orana ilişkin ele alınan parametrik t testi ve parametrik olmayan Mc-Nemar testlerinin güç karşılaştırmalarının yapılması amacıyla 10000 tekrarlı Monte-Carlo simülasyon yöntemi uygulanmıştır. Testlerin güçleri elde edilirken hipotezler çift yönlü kurulmuştur ve $\alpha = 0.05$ anlamlılık seviyesi ele alınmıştır. Farklı dağılımlar ve farklı n değerleri için elde edilen sonuçlar Çizelge 5.5'te verilmiştir. Çizelgede, $P_X = P_Y$ olan durumlardaki değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 5.5. Farklı P_X ve P_Y değerleri için güçler

	P_X	P_Y	Mc-Nemar	t testi		P_X	P_Y	Mc-Nemar	t testi
		0.20	0.1693	0.00			**0.20**	0.017	0.0353
	0.2	**0.28**	0.0652	0.0014		**0.2**	**0.28**	0.1058	0.1557
		0.36	0.0346	0.0103			**0.36**	0.3228	0.4078
		0.50	0.0297	0.0007			**0.50**	0.0234	0.0392
$n=10$	**0.5**	**0.65**	0.0246	0.0212	$n=40$	**0.5**	**0.65**	0.203	0.2657
		0.80	0.1037	0.1037			**0.80**	0.629	0.6961
		0.80	0.1718	0.0002			**0.80**	0.0143	0.0336
	0.8	**0.84**	0.1097	0.0002		**0.8**	**0.84**	0.0384	0.0698
		0.88	0.0688	0.0016			**0.88**	0.0938	0.1431
		0.20	0.0359	0.0106			**0.20**	0.0184	0.0377
	0.2	**0.28**	0.0318	0.0642		**0.2**	**0.28**	0.1295	0.1847
		0.36	0.1136	0.1942			**0.36**	0.4211	0.5057
		0.50	0.0125	0.0315			**0.50**	0.0264	0.0424
$n=20$	**0.5**	**0.65**	0.0943	0.1459	$n=50$	**0.5**	**0.65**	0.2786	0.342
		0.80	0.3005	0.3869			**0.80**	0.7489	0.7937
		0.80	0.0316	0.0098			**0.80**	0.0159	0.0359
	0,8	**0.84**	0.0193	0.0282		**0.8**	**0.84**	0.0527	0.0837
		0.88	0.031	0.0609			**0.88**	0.1288	0.1798
		0.20	0.012	0.0238			**0.20**	0.0244	0.0485
	0,2	**0.28**	0.0649	0.1126		**0.2**	**0.28**	0.2895	0.3564
		0.36	0.2273	0.3145			**0.36**	0.7769	0.8212
		0.50	0.0213	0.0407			**0.50**	0.0319	0.0443
$n=30$	**0,5**	**0.65**	0.1384	0.1955	$n=100$	**0.5**	**0.65**	0.5489	0.5989
		0.80	0.4718	0.5525			**0.80**	0.9729	0.9803
		0.80	0.0109	0.0229			**0.80**	0.024	0.0455
	0,8	**0.84**	0.0264	0.0508		**0.8**	**0.84**	0.0891	0.1315
		0.88	0.0613	0.1082			**0.88**	0.2914	0.3552

Küçük örneklem değerlerinde kitle oranları arasındaki fark küçük olduğunda Mc-Nemar testinin gücü yüksek çıkmaktadır. Kitle oranları arasındaki fark büyüdükçe tüm testlerin gücü artmaktadır.

Örneklem sayısı arttıkça testlerin gücü artmaktadır. Bununla beraber büyük örneklemlerde de kitle oranları arasındaki fark azaldıkça testlerin gücü azalmaktadır. Büyük örneklem durumunda t testinin gücünün daha yüksek olduğu görülmektedir.

6. BAĞIMSIZ K ÖRNEKLEM TESTLERİ

6.1. Ortalama ve Medyana İlişkin Testler

Bu bölümde bağımsız k (k>2) örneklem testlerinden ortalama ve medyana ilişkin parametrik Tek yönlü varyans analizi ile parametrik olmayan Kruskal Wallis H testi ve Medyan testi tanıtılmıştır.

6.1.1. Tek yönlü varyans analizi (ANOVA)

İkiden fazla grubun bir faktör bakımından karşılaştırılması varyans analizi olarak tanımlanmaktadır. t testi ile farkı, karşılaştırılan grupların normal dağılmasına ek olarak varyansların homojen olması da gerekmektedir. ANOVA ile grupların birbirinden farklı olup olmadığını test edilebilir ancak hangi grupların birbirinden farklı olduğunu tespit edilemeyebilir. (Post hoc ile hangi grubun farklı olduğuna bakılabilir.)

Bu analizde, herhangi bir gözleme ilişkin doğrusal model,

$y_{ij} = \mu + a_i + \varepsilon_{ij}$ şeklinde tanımlanmaktadır ve hipotez testi,

$H_0 : \mu_1 = \mu_2 = ... = \mu_k$

$H_1 : En\ az\ bir\ ortalama\ diğerlerinden\ farklıdır$

şeklinde oluşturulmaktadır.

$$s_{GA}^2 = \frac{\sum_{i=1}^{y_i^2} \frac{y_i^2}{n_i} - \frac{y^2}{n}}{k-1} \tag{6.1}$$

(6.1) eşitliğinde, s_{GA}^2 gruplar arası kareler ortalaması, k, grup ortalaması, n, gözlem sayısı ve k-1, serbestlik derecesidir.

$$s_{GI}^2 = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{2} \frac{y_i^2}{n_i}}{n-k} \tag{6.2}$$

(6.2) eşitliğinde s_{GI}^2 , gruplar içi kareler ortalaması, n-k, serbestlik derecesidir.

Gruplar arası kareler ortalamasının gruplar içi kareler ortalamasına oranlanması ile elde edilen F test istatistiği kullanılmaktadır ve

$$F_{Hesap} = \frac{s_{GA}^2}{s_{GI}^2} \tag{6.3}$$

şeklinde hesaplanmaktadır. $F_{Hesap} > F_{(k-1),(n-k),1-\alpha}$ ise H_0 reddedilir.

6.1.2. Kruskal Wallis H testi (K-W H Test)

Kruskal-Wallis H testi, tek yönlü varyans analizinin parametrik olmayan karşılığıdır. Bir faktör ve k (ikiden fazla) bağımsız grup var ise parametrik test varsayımları ve varyans analizinin gereklilikleri gerçekleşmemiş ise uygulanmalıdır. Test sonucunda gruplar arası fark anlamlı çıkarsa farklı grupları belirlemek için Mann-Whitney U testi kullanılır.

$$\overline{s}_i = \frac{\sum_{j=1}^{n_i} s\left(x_{ij}\right)}{n_i} \tag{6.4}$$

$$S^2 = \frac{1}{n-1}\left(\sum_{j=1}^{k}\sum_{i=1}^{n_j} s(x_{ij})^2 - \frac{n(n+1)^2}{4}\right) \tag{6.5}$$

ve test istatistiği,

$$KW = \frac{1}{S^2}\left(\sum_{i=1}^{k}\frac{s_i^2}{n_i} - \frac{n(n+1)^2}{4}\right) \tag{6.6}$$

şeklinde hesaplanmaktadır. (6.4) eşitliğinde, $\overline{s}$, her gruptaki sıra değerleri ortalaması, (6.6) eşitliğindeki, s_i , her gruptaki sıra değerleri toplamı olmak üzere α anlamlılık düzeyinde $\chi^2_{k-1,1-\alpha}$ tablo değeri ile karşılaştırılmaktadır. $KW_{Hesap} > \chi^2_{k-1,1-\alpha}$ ise H_0 reddedilir.

6.1.3. Medyan testi

k bağımsız gruptaki sıralı ölçekle elde edilmiş verilerin benzer medyanlı kitlelerden gelip gelmediğini test etmektedir. Medyan testi, k bağımsız örneğin benzer dağılım şekline sahip yığınlardan alınmış rasgele örnekler olduğunu varsaymaktadır.

$H_0 : M_1 = M_2 = ... = M_k$ hipotezi, $H_1 : M_1 \neq M_2 \neq .... = M_k$ hipotezine karşı test ediliyor ise test istatistiği,

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{\left(G_{ij} - T_{ij}\right)^2}{T_{ij}}; \qquad r = 2 \tag{6.7}$$

(6.7) eşitliğinde, G, gözlenen değer, T, beklenen değer, r sıra sayısıdır (Özdamar, 2002).

6.1.4. Güç karşılaştırmaları

Ortalamaya ilişkin ele alınan parametrik tek yönlü varyans analizi ile parametrik olmayan Kruskal Wallis H testi ve Medyan testlerinin güç karşılaştırmalarının yapılması amacıyla gözlem sayıları eşit 4 bağımsız örneklem durumu için $\alpha = 0.05$ anlamlılık seviyesinde 5000 tekrarlı Monte-Carlo simülasyon yöntemi uygulanmıştır. Farklı dağılımlar ve farklı n değerlerine ait simülasyonlar sonucunda elde edilen testlerin güçleri sırasıyla Çizelge 6.1., Çizelge 6.2. ve Çizelge 6.3. ve Çizelge 6.4.'te verilmiştir. Çizelgede, $\mu_1 = \mu_2 = \mu_3 = \mu_4$ olduğu durumlardaki değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 6.1.'de varyansları eşit, birinin ortalaması diğerlerinden farklı olan Normal dağılımlı dört kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir.

Çizelge 6.1. $\mathbf{X_1, X_2, X_3, X_4} \sim N(\mu_i, 3)$, $i = 1,2,3,4$ için testlerin güçleri

n	μ_1	$\mu_2 = \mu_3 = \mu_4$	ANOVA	Kruskal-Wallis H testi	Medyan testi
10	5	10	0.9636	0.945	0.8192
	6		0.842	0.8024	0.633
	7		0.5706	0.5242	0.3924
	8		0.278	0.2478	0.1958
	9		0.0978	0.0842	0.0798
	10		0.0494	0.0414	0.0538
	11		0.1002	0.089	0.0872
	12		0.2774	0.248	0.2052
	13		0.5754	0.5292	0.384
	14		0.8408	0.7984	0.6224
	15		0.966	0.945	0.8106
20	5	10	1.00	0.9998	0.988
	6		0.994	0.9902	0.8996
	7		0.9084	0.8798	0.677
	8		0.5456	0.5146	0.3386
	9		0.1516	0.1458	0.1052
	10		0.0496	0.0468	0.0466
	11		0.1584	0.1454	0.1078
	12		0.539	0.4976	0.3368
	13		0.8962	0.88	0.6776
	14		0.9954	0.9922	0.9096
	15		0.9998	0.9996	0.9846
30	5	10	1.00	1.00	0.9998
	6		0.9998	0.9998	0.991
	7		0.9812	0.9762	0.8696
	8		0.7466	0.7178	0.5226
	9		0.2264	0.2126	0.1594
	10		0.0542	0.0542	0.0502
	11		0.2154	0.2112	0.1516
	12		0.7558	0.7268	0.525
	13		0.9832	0.9796	0.8758
	14		1.00	0.9996	0.9882
	15		1.00	1.00	1.00
40	5	10	1.00	1.00	1.00
	6		1.00	1.00	0.9994
	7		0.9986	0.9976	0.9554
	8		0.8716	0.8408	0.6416
	9		0.3024	0.2888	0.1936
	10		0.0468	0.0472	0.0432
	11		0.2934	0.2798	0.183
	12		0.8798	0.858	0.6512
	13		0.9972	0.997	0.9542
	14		1.00	1.00	0.9988
	15		1.00	1.00	1.00
50	5	10	1.00	1.00	1.00
	6		1.00	1.00	1.00
	7		1.00	0.9996	0.9874
	8		0.9394	0.9248	0.7712
	9		0.3712	0.348	0.2536
	10		0.0496	0.0502	0.0516
	11		0.3628	0.3464	0.2442
	12		0.9406	0.9248	0.7692
	13		0.9998	0.9996	0.989

	14		1.00	1.00	1.00
	15		1.00	1.00	1.00
100	**5**	**10**	1.00	1.00	1.00
	6		1.00	1.00	1.00
	7		1.00	1.00	1.00
	8		0.9996	0.9994	0.9796
	9		0.6734	0.6408	0.458
	10		0.0496	0.048	0.0472
	11		0.6674	0.648	0.4606
	12		1.00	1.00	0.9804
	13		1.00	1.00	1.00
	14		1.00	1.00	1.00
	15		1.00	1.00	1.00

Çizelge 6.1.'deki testlerin güçleri incelendiğinde bu üç test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test ANOVA testi, gücü en düşük olan test ise Medyan testidir. n arttıkça tüm testlerin gücünün arttığı görülmektedir. Ayrıca birinci grubun ortalaması ile diğer grupların ortalaması arasındaki fark arttıkça testlerin gücü artmakta, fark azaldıkça güçler de azalmaktadır.

Çizelge 6.2.'de varyansları eşit, birinin ortalaması diğerlerinden farklı olan Normal dağılımlı dört kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, $\mu_1 = \mu_2 = \mu_3 = \mu_4$ olduğu durumlardaki değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 6.2. $\mathbf{X_1, X_2, X_3, X_4} \sim N(\mu_i, 6)$, $i = 1,2,3,4$ için testlerin güçleri

n	μ_1	$\mu_2 = \mu_3 = \mu_4$	**ANOVA**	**Kruskal-Wallis H testi**	**Medyan testi**
10	**5**	**10**	0.4174	0.3754	0.2854
	6		0.271	0.2466	0.1928
	7		0.1746	0.149	0.1258
	8		0.0908	0.0844	0.085
	9		0.0646	0.0548	0.0582
	10		0.0492	0.041	0.0506
	11		0.062	0.052	0.0602
	12		0.0922	0.085	0.082
	13		0.1696	0.152	0.126
	14		0.271	0.2396	0.1988
	15		0.4158	0.376	0.276
20	**5**	**10**	0.7604	0.7286	0.5146
	6		0.536	0.5044	0.3334
	7		0.3226	0.2984	0.197
	8		0.154	0.147	0.1134
	9		0.075	0.0728	0.0652
	10		0.0502	0.0466	0.0478
	11		0.0774	0.0746	0.0654
	12		0.1568	0.1478	0.108
	13		0.3222	0.3016	0.2076

	14		0.5584	0.525	0.3498
	15		0.7584	0.7356	0.5144
30	**5**	**10**	0.9208	0.8938	0.7114
	6		0.75	0.7202	0.5228
	7		0.4782	0.4512	0.3078
	8		0.2372	0.2218	0.1516
	9		0.09	0.0882	0.0748
	10		0.0486	0.0474	0.0458
	11		0.0904	0.0854	0.0712
	12		0.2246	0.2114	0.148
	13		0.48	0.4584	0.3154
	14		0.7422	0.7132	0.514
	15		0.9202	0.904	0.726
40	**5**	**10**	0.9762	0.9706	0.8538
	6		0.8708	0.8464	0.6492
	7		0.6092	0.58	0.4024
	8		0.2938	0.2738	0.1842
	9		0.1096	0.1012	0.073
	10		0.05	0.0492	0.0472
	11		0.1038	0.0958	0.0738
	12		0.3066	0.2878	0.1834
	13		0.6152	0.5964	0.399
	14		0.8688	0.8496	0.655
	15		0.9746	0.969	0.8528
50	**5**	**10**	0.9942	0.9916	0.9366
	6		0.9418	0.9292	0.7828
	7		0.7168	0.6906	0.5032
	8		0.364	0.3446	0.2514
	9		0.1166	0.1134	0.102
	10		0.0496	0.0502	0.0516
	11		0.1188	0.112	0.0956
	12		0.3632	0.3476	0.2428
	13		0.7214	0.6898	0.5038
	14		0.9386	0.927	0.7898
	15		0.9948	0.9928	0.935
100	**5**	**10**	1.00	1.00	0.9994
	6		0.9992	0.9986	0.9764
	7		0.9652	0.9554	0.83
	8		0.6792	0.6534	0.4626
	9		0.2042	0.1932	0.1396
	10		0.0496	0.048	0.0472
	11		0.1954	0.1852	0.1332
	12		0.6822	0.6534	0.4692
	13		0.968	0.9608	0.8342
	14		0.9994	0.9992	0.9808
	15		1.00	1.00	0.9988

Çizelge 6.2.'deki testlerin güçleri incelendiğinde bu üç test istatistiği içinden farklı n'ler için daha güçlü olduğunu söyleyebileceğimiz test ANOVA testi, gücü en düşük olan test ise Medyan testidir. n arttıkça tüm testlerin gücünün arttığı görülmektedir. Ayrıca birinci grubun ortalaması ile diğer grupların ortalaması arasındaki fark arttıkça testlerin gücü artmakta, fark azaldıkça güçler de azalmaktadır.

Çizelge 6.1. ve Çizelge 6.2. birlikte incelendiğinde ise normal dağılımlı kitlelerde kitle varyansı arttıkça testlerin güçlerinin azaldığı görülmektedir.

Çizelge 6.3.'de birinin θ parametresi diğerlerinden farklı olan Üstel dağılımlı dört kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir. Çizelgede, $\mu_1 = \mu_2 = \mu_3 = \mu_4$ olduğu durumlardaki değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 6.3. $\mathbf{X_1, X_2, X_3, X_4} \sim \ddot{u}stel(\theta_i)$, $i = 1,2,3,4$ için testlerin güçleri

n	θ_1	$\theta_2 = \theta_3 = \theta_4$	ANOVA	Kruskal-Wallis H testi	Medyan testi
10	**5**	**10**	0.1564	0.1954	0.1666
	6		0.1118	0.1294	0.1136
	7		0.0838	0.0882	0.0854
	8		0.0562	0.0572	0.062
	9		0.0434	0.0458	0.056
	10		0.0412	0.043	0.054
	11		0.047	0.0482	0.0596
	12		0.0558	0.0564	0.0562
	13		0.0716	0.0614	0.0704
	14		0.092	0.0786	0.069
	15		0.1196	0.0914	0.0874
20	**5**	**10**	0.4014	0.4416	0.3158
	6		0.2272	0.2398	0.18
	7		0.1366	0.1338	0.0996
	8		0.0752	0.0732	0.0702
	9		0.0594	0.0582	0.0544
	10		0.0448	0.0456	0.048
	11		0.0526	0.0542	0.052
	12		0.0674	0.0648	0.0604
	13		0.1172	0.0964	0.0764
	14		0.1656	0.123	0.0866
	15		0.24	0.172	0.1164
30	**5**	**10**	0.6412	0.6474	0.482
	6		0.3608	0.3628	0.2658
	7		0.1998	0.1934	0.1336
	8		0.0996	0.095	0.083
	9		0.0526	0.0498	0.0494
	10		0.0444	0.0456	0.0458
	11		0.0636	0.0562	0.0552
	12		0.1016	0.0874	0.0696
	13		0.1578	0.1248	0.0902
	14		0.2412	0.1732	0.1236
	15		0.354	0.246	0.1494
40	**5**	**10**	0.8126	0.7878	0.6068
	6		0.525	0.513	0.3394
	7		0.2722	0.2542	0.1788
	8		0.1216	0.1206	0.0922
	9		0.0618	0.0678	0.0624
	10		0.0444	0.0462	0.046
	11		0.068	0.063	0.0548

	12		0.122	0.1052	0.0798
	13		0.2056	0.1492	0.1002
	14		0.3168	0.2216	0.1396
	15		0.4546	0.3162	0.198
50	**5**	**10**	0.916	0.8846	0.7272
	6		0.6366	0.6042	0.4366
	7		0.3404	0.3078	0.2408
	8		0.157	0.1436	0.1174
	9		0.0714	0.0714	0.0672
	10		0.0492	0.0496	0.0516
	11		0.0644	0.0626	0.0598
	12		0.1378	0.1164	0.1014
	13		0.2462	0.1786	0.1334
	14		0.4088	0.2866	0.201
	15		0.544	0.3856	0.2476
100	**5**	**10**	0.9996	0.9964	0.9672
	6		0.9558	0.9194	0.7626
	7		0.6826	0.6036	0.4246
	8		0.2938	0.2554	0.1822
	9		0.0968	0.0848	0.0712
	10		0.0528	0.0486	0.0544
	11		0.0984	0.0822	0.0694
	12		0.2468	0.1854	0.126
	13		0.479	0.3494	0.23
	14		0.7094	0.5286	0.3448
	15		0.8644	0.7086	0.4756

Çizelge 6.3. incelendiğinde, küçük örneklemlerde (n <20), birinci grubun θ parametresi diğer grupların θ parametresinden küçük olduğu durumda diğer gruplarla olan fark büyüdükçe Kruskal-Wallis H testinin, fark azaldıkça Medyan testinin gücünün daha yüksek çıktığı görülmektedir. Bununla birlikte, birinci grubun θ parametresi diğer grupların θ parametresinden büyükken ANOVA testinin gücünün daha yüksek çıktığı gözlenmektedir.

20< n <40 durumunda, birinci grubun θ parametresi diğer grupların θ parametresinden küçük olduğunda diğer gruplarla olan fark büyüdükçe Kruskal-Wallis H testinin, fark azaldıkça ANOVA testinin gücünün daha yüksek çıktığı görülmektedir. Bununla birlikte, birinci grubun θ parametresi diğer grupların θ parametresinden büyükken ANOVA testinin gücünün daha yüksek çıktığı gözlenmektedir.

n arttıkça tüm testlerin gücünün arttığı ve gücü en yüksek testin ANOVA, en düşük testin Medyan testi olduğu görülmektedir.

Çizelge 6.1. ve Çizelge 6.2. ile Çizelge 6.3. birlikte incelendiğinde ise örneklemler üstel dağılımdan geldiğinde, örneklemlerin normal dağılımdan geldiği duruma göre tüm n 'lerde testlerin güçlerinin azaldığı söylenebilir.

Çizelge 6.4.'de biri üstel dağılımlı, diğerleri normal dağılımlı dört kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir. Çizelgelerde, ANOVA için $\mu_1 = \mu_2 = \mu_3 = \mu_4$ olduğu durumlardaki değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 6.4. $\mathbf{X_1} \sim$ *üstel*(θ_i), $\mathbf{X_2}, \mathbf{X_3}, \mathbf{X_4} \sim N(\mu_i, 3)$, $i = 1,2,3,4$ için testlerin güçleri

n	θ_1	$\mu_2 = \mu_3 = \mu_4$	ANOVA	Kruskal-Wallis H testi	Medyan testi
10	**5**	**10**	0.7772	0.7764	0.5858
	6		0.5816	0.5886	0.41
	7		0.3952	0.4224	0.2984
	8		0.26	0.2892	0.2022
	9		0.1686	0.2044	0.148
	10		0.126	0.144	0.1082
	11		0.095	0.0978	0.0852
	12		0.1006	0.0744	0.069
	13		0.1186	0.0642	0.0612
	14		0.1518	0.0634	0.0664
	15		0.1994	0.062	0.0632
20	**5**	**10**	0.9458	0.9776	0.8794
	6		0.7872	0.8954	0.701
	7		0.5408	0.7196	0.5142
	8		0.3342	0.5362	0.3582
	9		0.1874	0.3678	0.2376
	10		0.113	0.2432	0.164
	11		0.0996	0.1498	0.1056
	12		0.1366	0.0978	0.0762
	13		0.1968	0.0708	0.0606
	14		0.2876	0.0678	0.0566
	15		0.3892	0.0668	0.0528
30	**5**	**10**	0.9898	0.9982	0.9796
	6		0.8984	0.9786	0.896
	7		0.6748	0.8836	0.7322
	8		0.38	0.6978	0.5054
	9		0.1888	0.4954	0.3422
	10		0.1006	0.3254	0.2258
	11		0.1054	0.1956	0.145
	12		0.1736	0.1106	0.0914
	13		0.2848	0.0762	0.0696
	14		0.426	0.0638	0.0518
	15		0.547	0.0722	0.0556
40	**5**	**10**	0.9978	1.00	0.997
	6		0.9516	0.9958	0.9642
	7		0.7606	0.9526	0.8534
	8		0.4546	0.8322	0.66
	9		0.199	0.6194	0.4572
	10		0.0972	0.4052	0.2758
	11		0.1154	0.2248	0.1604
	12		0.2186	0.135	0.1052
	13		0.382	0.0772	0.0634
	14		0.5448	0.0648	0.0612
	15		0.6984	0.0862	0.0574

50	**5**	**10**	0.9998	1.00	1.00
	6		0.982	0.9994	0.9902
	7		0.824	0.983	0.9256
	8		0.5014	0.9008	0.7792
	9		0.2038	0.6976	0.5548
	10		0.0906	0.4826	0.3618
	11		0.1194	0.2806	0.226
	12		0.265	0.1522	0.1242
	13		0.448	0.091	0.081
	14		0.6442	0.0658	0.0624
	15		0.7982	0.0864	0.067
100	**5**	**10**	1.00	1.00	1.00
	6		0.9998	1.00	1.00
	7		0.9734	0.9998	0.9978
	8		0.7234	0.9964	0.9758
	9		0.292	0.9462	0.8682
	10		0.0974	0.7672	0.6358
	11		0.1898	0.475	0.3694
	12		0.467	0.2466	0.1994
	13		0.7502	0.1092	0.0958
	14		0.913	0.0706	0.063
	15		0.976	0.1036	0.0598

Çizelge 6.4. incelendiğinde gücü en düşük testin Medyan testi olduğu görülmektedir. Üstel dağılımlı birinci grubun θ parametresi Normal dağılımlı diğer grupların ortalamasından küçükken Kruskal-Wallis H testinin gücünün daha yüksek çıktığı görülmektedir. Bununla birlikte, üstel dağılımlı birinci grubun θ parametresi Normal dağılımlı diğer grupların ortalamasından büyükken ANOVA testinin gücünün daha yüksek çıktığı gözlenmektedir. Ancak ANOVA testinde anlamlılık seviyesi de nispeten büyük çıktığından örneklemlerin farklı dağılımlardan geldiği bu durum için Kruskal-Wallis H testinin kullanılmasının daha uygun olduğu söylenebilir. Ayrıca n arttıkça testlerin gücünün arttığı çizelgeden görülmektedir.

7. BAĞIMLI K ÖRNEKLEM TESTLERİ

7.1. Ortalama ve Medyana İlişkin Testler

Bu bölümde bağımlı k (k>2) örneklem testlerinden ortalama ve medyana ilişkin parametrik iki yönlü varyans analizi ile parametrik olmayan Friedman'ın S testi tanıtılmıştır.

7.1.1. İki yönlü varyans analizi

İki yönlü varyans analizi iki değişkenin bağımlı bir değişken ile ortak etkisini ölçmek için kullanılmaktadır. Bağımsız değişkenler kategorik, bağımlı değişkenler ise süreklidir.

Bağımlı k grup denemelerinde, üç bileşime ayrılan Genel Değişim mevcuttur. Birincisi, işlemler arası değişim, ikincisi, birimler arası değişim, üçüncüsü ise deneysel hatadır.

$$GT = \sum_{i=1}^{n}\sum_{j=1}^{k} X_{ij} \qquad TT_i = \sum_{i=1}^{n} X_{i.} \qquad BT_j = \sum_{j=1}^{k} X_{.j} \tag{7.1}$$

$$DT = \left(GT\right)^2 / N \tag{7.2}$$

$$GKT = \sum_{i=1}^{n}\sum_{j=1}^{k} X_{ij}^2 - DT \qquad BAKT = \sum_{i=1}^{n} \frac{\left(BT_j\right)^2}{k} - DT \tag{7.3}$$

$$TAKT = \sum_{j=1}^{k} \frac{\left(TT_i\right)^2}{n} - DT \qquad HKT = GKT - \left(BAKT + TAKT\right) \tag{7.4}$$

şeklinde tanımlanmaktadır. (7.1) eşitliğinde tanımlanan GT, Genel Toplam, BT, Birim Toplamı, TT, İşlem Toplamıdır. (7.2) eşitliğinde tanımlanan DT, Düzeltme Terimidir. (7.3) eşitliğinde tanımlanan GKT, Genel Kareler Toplamı, TAKT, İşlemler Arası Kareler Toplamı ve HKT, Hata Kareler Toplamıdır. n, birim sayısı, k, işlem sayısı, N, gözlem sayısıdır ve $N = n*k$ olarak hesaplanmaktadır.

Birim, İşlem ve Hata kare ortalamaları ise aşağıdaki gibidir.

$$BAKO = \left(s_n^2\right) = BAKT / bsd \tag{7.5}$$

$$TAKO = \left(s_k^2\right) = TAKT / tsd \tag{7.6}$$

$$HKO = \left(s^2\right) = HKT / hsd \tag{7.7}$$

(7.5) eşitliğinde $bsd = b - 1$, birimler arası serbestlik derecesi ve $BAKO$, Birim Kareler Ortalamasıdır. (7.6) eşitliğinde $tsd = t - 1$, işlemler arası serbestlik derecesi ve $TAKO$, İşlem Kareler Ortalamasıdır. (7.7) eşitliğinde $hsd = gsd - (bsd + tsd)$, hata serbestlik derecesi ve HKO, Hata Kareler Ortalaması olarak tanımlanmaktadır.

F_1, birim farklılıklarını belirlemek için, F_2, işlemlerin farklılığını belirlemek için test istatistikleridir.

$$F_1 = BAKO / HKO = s_n^2 / s^2 \tag{7.8}$$

$$F_2 = TAKO / HKO = s_k^2 / s^2 \tag{7.9}$$

Test istatistikleri, kritik değerlere eşit ve daha büyük ya da küçük olmalarına göre önemlilik düzeyleri belirlenir (Özdamar, 2002).

7.1.2. Friedman'ın S testi

Friedman'ın S testi, parametrik iki yönlü varyans analizi yönteminin parametrik olmayan karşılığıdır. Bu test gözlemlerin sıra değerleri esas alınarak hesaplanmaktadır.

$H_0 = işlemler\ benzer\ etkiye\ sahiptir.$

$H_1 = işlemler\ \ farklı\ etkiye\ sahiptir.$

şeklinde hipotezler oluşturulur. Birimlerin k işlem için aldığı sıra değerleri sıralanır. Aynı puanlar ortalama sıra değeri almaktadır. İşlemlerin birimlere göre sıra değer puanları toplanmaktadır. İşlemlerdeki sıralama puanlarının farklılığı şu şekilde hesaplanmaktadır.

$$\chi_r^2 = \frac{12}{N_k\left(k+1\right)}\sum_{j=1}^{k}\left(R_j\right)^2 - 3N\left(k+1\right) \tag{7.10}$$

(7.10) eşitliğinde tanımlanan χ_r^2 test istatistiğinin önemliliği, $sd = k-1$ serbestlik dereceli χ^2 dağılımının kritik değerlerine göre belirlenmektedir. k ise işlem sayısıdır.

7.1.3. Güç karşılaştırmaları

Ortalamaya ilişkin ele alınan parametrik iki yönlü varyans analizi ve parametrik olmayan Friedman'ın S testlerinin güç karşılaştırmalarının yapılması amacıyla $\alpha = 0.05$ anlamlılık seviyesinde 10.000 tekrarlı Monte-Carlo simülasyon yöntemi uygulanmıştır. Eşit varyanslı ve düşük, orta ile yüksek korelasyonlu 4 bağımlı örneklem için farklı n değerlerine ait simülasyonlar sonucunda elde edilen testlerin güçleri sırasıyla Çizelge 7.1., Çizelge 7.2. ve Çizelge 7.3.'te verilmiştir. Çizelge 7.4.'te ise farklı varyanslı bağımlı örneklemler için yapılan simülasyon sonucunda farklı n değerleri için elde edilen testlerin güçleri verilmiştir. Tüm simülasyonlarda örneklemlerin normal dağılımdan geldiği varsayılmıştır. Çizelgelerde, $\mu_1 = \mu_2 = \mu_3 = \mu_4$ olduğu durumlardaki değerler testin simülasyon sonucunda elde edilen anlamlılık seviyeleridir.

Çizelge 7.1.'de varyansları eşit, düşük korelasyonlu normal dağılımlı dört kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir.

Çizelge 7.1. $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = 4, \sigma_{ij} = 1;\ i, j = 1,2,3,4, i \neq j$ durumu için güçler

n	μ_1	$\mu_2 = \mu_3 = \mu_4$	İki yönlü varyans analizi	Friedman'ın S testi
	5		0.4851	0.9998
	6		0.4971	0.9912
	7		0.4918	0.8933
	8		0.367	0.5432
	9		0.1206	0.1579
10	**10**	**10**	0.0303	0.0419
	11		0.1247	0.159
	12		0.3674	0.5555
	13		0.4898	0.8986
	14		0.4951	0.9899
	15		0.4874	1.00
20	**5**	**10**	0.7441	1.00

	6		0.7452	1.00
	7		0.7447	0.999
	8		0.7133	0.8874
	9		0.3238	0.3213
	10		0.0425	0.0505
	11		0.3198	0.3177
	12		0.7127	0.8903
	13		0.7342	0.9985
	14		0.7442	1.00
	15		0.7378	1.00
30	**5**	**10**	0.8727	1.00
	6		0.8698	1.00
	7		0.8739	1.00
	8		0.8709	0.9837
	9		0.5271	0.4656
	10		0.0486	0.0457
	11		0.5366	0.4805
	12		0.8704	0.9814
	13		0.8758	1.00
	14		0.8796	1.00
	15		0.88	1.00
40	**5**	**10**	0.944	1.00
	6		0.9441	1.00
	7		0.9442	1.00
	8		0.9416	0.9965
	9		0.695	0.6085
	10		0.0508	0.0513
	11		0.7053	0.6121
	12		0.9402	0.9974
	13		0.9405	1.00
	14		0.9385	1.00
	15		0.9431	1.00
50	**5**	**10**	0.9722	1.00
	6		0.9695	1.00
	7		0.975	1.00
	8		0.9759	0.9996
	9		0.823	0.7178
	10		0.0519	0.0504
	11		0.8231	0.727
	12		0.9715	0.9998
	13		0.9705	1.00
	14		0.9702	1.00
	15		0.9727	1.00
100	**5**	**10**	0.9996	1.00
	6		0.9996	1.00
	7		0.9993	1.00
	8		0.9996	1.00
	9		0.9916	0.9657
	10		0.0466	0.0503
	11		0.9923	0.9658
	12		0.9991	1.00
	13		0.9992	1.00
	14		0.9994	1.00
	15		0.9997	1.00

Çizelge 7.1. incelendiğinde, gözlem sayısının küçük olduğu durumlarda (n <20) Friedman'ın S testinin gücünün daha yüksek çıktığı görülmektedir. Bununla birlikte n arttıkça, birinci grubun ortalaması ile diğer grupların ortalaması arasındaki fark 1 iken iki yönlü varyans analizinin gücünün, fark 1'den büyükken Friedman'ın S testinin gücünün daha yüksek çıktığı görülmektedir. Birinci grubun ortalaması ile diğer grupların ortalaması arasındaki fark büyüdükçe her iki testin gücü de artmaktadır.

Çizelge 7.2.'de varyansları eşit, orta düzeyde korelasyonlu normal dağılımlı dört kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir.

Çizelge 7.2. $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = 4, \sigma_{ij} = 2;\ i, j = 1,2,3,4, i \neq j$ durumu için güçler

n	μ_1	$\mu_2 = \mu_3 = \mu_4$	İki yönlü varyans analizi	Friedman'ın S testi
10	**5**	**10**	0.8946	1.00
	6		0.8919	1.00
	7		0.8924	0.9794
	8		0.7871	0.7392
	9		0.2794	0.2211
	10		0.0501	0.048
	11		0.2732	0.221
	12		0.7797	0.7387
	13		0.8961	0.9788
	14		0.8949	0.9994
	15		0.8953	1.00
20	**5**	**10**	0.9925	1.00
	6		0.9951	1.00
	7		0.9936	1.00
	8		0.99	0.9806
	9		0.5907	0.461
	10		0.0512	0.0529
	11		0.5919	0.4696
	12		0.9885	0.9811
	13		0.9932	1.00
	14		0.9928	1.00
	15		0.9937	1.00
30	**5**	**10**	0.9995	1.00
	6		0.9997	1.00
	7		0.9996	1.00
	8		0.9998	0.999
	9		0.7934	0.659
	10		0.0505	0.0471
	11		0.7935	0.6584
	12		0.9994	0.9997
	13		0.9996	1.00
	14		0.9996	1.00
	15		0.9998	1.00
40	**5**	**10**	0.9999	1.00
	6		1	1.00
	7		1	1.00

	8		0.9999	1.00
	9		0.9054	0.7969
	10		0.0499	0.0473
	11		0.9066	0.8006
	12		1.00	1.00
	13		1.00	1.00
	14		1.00	1.00
	15		0.9998	1.00
50	**5**	**10**	1.00	1.00
	6		1.00	1.00
	7		1.00	1.00
	8		1.00	1.00
	9		0.9579	0.884
	10		0.0484	0.0483
	11		0.9603	0.8918
	12		1.00	1.00
	13		1.00	1.00
	14		1.00	1.00
	15		1.00	1.00
100	**5**	**10**	1.00	1.00
	6		1.00	1.00
	7		1.00	1.00
	8		1.00	1.00
	9		0.9997	0.997
	10		0.0471	0.0489
	11		1.00	0.9974
	12		1.00	1.00
	13		1.00	1.00
	14		1.00	1.00
	15		1.00	1.00

Çizelge 7.2. incelendiğinde, birinci grubun ortalaması ile diğer grupların ortalaması arasındaki fark küçükken iki yönlü varyans analizi testinin gücünün daha yüksek çıktığı, fark arttıkça Friedman'ın S testinin gücünün daha yüksek çıktığı görülmektedir. n büyüdükçe testlerin gücü artmakta ve birbirine yaklaşmaktadır. Birinci grubun ortalaması ile diğer grupların ortalaması arasındaki fark büyüdükçe her iki testin gücü de artmaktadır.

Çizelge 7.3.'de varyansları eşit, yüksek korelasyonlu normal dağılımlı dört kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir.

Çizelge 7.3. $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = 4, \sigma_{ij} = 3;\ i, j = 1,2,3,4, i \neq j$ durumu için güçler

n	μ_1	$\mu_2 = \mu_3 = \mu_4$	**İki yönlü varyans analizi**	**Friedman'ın S testi**
10	**5**	**10**	0.9936	1.00
	6		0.9949	1.00
	7		0.9968	0.9997
	8		0.9903	0.9642

	9		0.5519	0.4241
	10		0.0474	0.0422
	11		0.552	0.4278
	12		0.9898	0.9631
	13		0.9955	0.9998
	14		0.9945	1.00
	15		0.9945	1.00
20	**5**	**10**	1.00	1.00
	6		1.00	1.00
	7		1.00	1.00
	8		1.00	0.9999
	9		0.8956	0.7845
	10		0.0529	0.0483
	11		0.8906	0.7851
	12		1.00	1.00
	13		1.00	1.00
	14		1.00	1.00
	15		1.00	1.00
30	**5**	**10**	1.00	1.00
	6		1.00	1.00
	7		1.00	1.00
	8		1.00	1.00
	9		0.9821	0.9362
	10		0.053	0.0474
	11		0.9836	0.9363
	12		1.00	1.00
	13		1.00	1.00
	14		1.00	1.00
	15		1.00	1.00
40	**5**	**10**	1.00	1.00
	6		1.00	1.00
	7		1.00	1.00
	8		1.00	1.00
	9		0.9979	0.9835
	10		0.0544	0.0494
	11		0.9985	0.9864
	12		1.00	1.00
	13		1.00	1.00
	14		1.00	1.00
	15		1.00	1.00
50	**5**	**10**	1.00	1.00
	6		1.00	1.00
	7		1.00	1.00
	8		1.00	1.00
	9		0.9998	0.9969
	10		0.0503	0.0507
	11		0.9995	0.9962
	12		1.00	1.00
	13		1.00	1.00
	14		1.00	1.00
	15		1.00	1.00
100	**5**	**10**	1.00	1.00
	6		1.00	1.00
	7		1.00	1.00
	8		1.00	1.00
	9		1.00	1.00
	10		0.0486	0.049

	11		1.00	1.00
	12		1.00	1.00
	13		1.00	1.00
	14		1.00	1.00
	15		1.00	1.00

Çizelge 7.3. incelendiğinde, gözlem sayısının küçük olduğu durumlarda (n <20), birinci grubun ortalaması ile diğer grupların ortalaması arasındaki fark küçükken iki yönlü varyans analizinin gücünün daha yüksek, fark büyüdükçe Friedman'ın S testinin gücünün daha yüksek çıktığı görülmektedir. n büyüdükçe testlerin gücü artmakta ve birbirine yaklaşmaktadır. Birinci grubun ortalaması ile diğer grupların ortalaması arasındaki fark büyüdükçe her iki testin gücü de artmaktadır.

Çizelge 7.1. – 7.3. birlikte incelendiğinde, ilişki arttıkça testlerin gücünün arttığı ve birbirine yaklaştığı söylenebilir.

Çizelge 7.4.'de varyansları farklı normal dağılımlı dört kitle için farklı parametre ve farklı n değerlerine göre simülasyon sonucu elde edilen güçler verilmiştir.

Çizelge 7.4. $\sigma_1^2 = 4, \sigma_2^2 = 6, \sigma_3^2 = 8, \sigma_4^2 = 10, \sigma_{ij} = 4;\ i, j = 1,2,3,4, i \neq j$ durumu için güçler

n	μ_1	$\mu_2 = \mu_3 = \mu_4$	**İki yönlü varyans analizi**	**Friedman'ın S testi**
	5		0.9459	1.00
	6		0.9447	0.9999
	7		0.9457	0.9965
	8		0.7486	0.8363
	9		0.187	0.2287
10	**10**	**10**	0.0672	0.0501
	11		0.1891	0.2282
	12		0.7444	0.8362
	13		0.9443	0.9967
	14		0.945	1.00
	15		0.9424	1.00
	5		0.9982	1.00
	6		0.9988	1.00
	7		0.9989	1.00
	8		0.9969	0.9984
	9		0.4155	0.5533
20	**10**	**10**	0.0687	0.0603
	11		0.4234	0.561
	12		0.9964	0.9983
	13		0.9987	1.00
	14		0.9984	1.00
	15		0.9981	1.00
	5		1.00	1.00
	6		0.9999	1.00
30	**7**	**10**	0.9999	1.00
	8		0.9999	1.00
	9		0.6847	0.799

	10		0.0695	0.0547
	11		0.6751	0.7884
	12		1.00	1.00
	13		1.00	1.00
	14		1.00	1.00
	15		1.00	1.00
40	**5**	**10**	1.00	1.00
	6		1.00	1.00
	7		1.00	1.00
	8		1.00	1.00
	9		0.8717	0.9252
	10		0.0666	0.0523
	11		0.8658	0.9184
	12		1.00	1.00
	13		1.00	1.00
	14		1.00	1.00
	15		1.00	1.00
50	**5**	**10**	1.00	1.00
	6		1.00	1.00
	7		1.00	1.00
	8		1.00	1.00
	9		0.9584	0.9757
	10		0.0693	0.0562
	11		0.9583	0.9781
	12		1.00	1.00
	13		1.00	1.00
	14		1.00	1.00
	15		1.00	1.00
100	**5**	**10**	1.00	1.00
	6		1.00	1.00
	7		1.00	1.00
	8		1.00	1.00
	9		1.00	1.00
	10		0.0619	0.0518
	11		1.00	0.9999
	12		1.00	1.00
	13		1.00	1.00
	14		1.00	1.00
	15		1.00	1.00

Çizelge 7.4.'e göre, n büyüdükçe testlerin gücü artmakta ve birbirine yaklaşmaktadır. Birinci grubun ortalaması ile diğer grupların ortalaması arasındaki fark büyüdükçe her iki testin gücü de artmaktadır. Bununla birlikte farklı varyans durumunda, Friedman'ın S testinin gücünün daha yüksek olduğu söylenebilir.

8. SONUÇLAR VE ÖNERİLER

8.1 Sonuçlar

Bu çalışmada tek örneklem, iki örneklem ve ikiden fazla örneklem durumları için istatistiksel sonuç çıkarımında sıkça kullanılan parametrik ve parametrik olmayan testler ele alınmış ve güçleri bakımından kıyaslanmıştır. Tek örneklem, bağımsız iki örneklem ve bağımlı iki örneklem durumlarında ortalama ve medyana ilişkin testler ile orana ilişkin testler; bağımsız ikiden fazla örneklem ve bağımlı ikiden fazla örneklem durumlarında ise ortalama ve medyana ilişkin testler incelenmiştir. Testlere ilişkin güçler Monte Carlo simülasyon yöntemi ile elde edilmiştir.

Tek örneklem durumunda ortalama ve medyana ilişkin parametrik testlerden Z testi ve t testi ile parametrik olmayan testlerden Wilcoxon testi ve işaret testinin güçleri farklı dağılımlar, farklı parametreler ve farklı n'ler için elde edilmiş ve kıyaslanmıştır. Sonuçta normal dağılım durumunda Z testinin gücü en yüksek, çoğu durumda işaret testinin gücü en düşük bulunmuştur. Ancak, üstel dağılım durumunda Z testinin anlamlılık seviyesi ve küçük n değerlerinde t testinin anlamlılık seviyeleri yüksek çıktığından, küçük n değerlerinde $(n<20)$ Wilcoxon işaretli sıra sayıları testinin, n büyüdükçe t testi veya işaret testinin kullanılması önerilebilir. Gama dağılımı durumunda da Z testinin anlamlılık seviyesinin yüksek çıktığı, gücü en düşük testin işaret testi olduğu dolayısıyla parametrik t testinin veya parametrik olmayan Wilcoxon işaretli sıra sayıları testinin kullanılması önerilebilir. n arttıkça tüm testlerin gücü artmıştır. Tek örneklem durumunda orana ilişkin parametrik testlerden Z testi ile parametrik olmayan testlerden Binom testinin güçleri faklı hipotezler, farklı parametreler ve farklı n'ler için elde edilmiş ve kıyaslanmıştır. Sonuçta, n arttıkça her iki testin de gücü artmıştır. Bununla birlikte, bazı hipotezler ve bazı parametre değerlerinde iki testinde gücü de aynı çıkmışken diğer durumlarda Z testinin gücü daha yüksek bulunmuştur. Güçlerin aynı ve farklı olduğu durumlar Çizelge 3.6 – 3.14'te verilmiştir.

Bağımsız iki örneklem durumunda ortalama ve medyana ilişkin parametrik testlerden Z testi ve t testi ile parametrik olmayan testlerden Binom testi, Mann Whitney U testi, Kolmogorov-Smirnov testi, Wald-Wolfowitz diziler testi ve Moods-Medyan testinin güçleri farklı dağılımlar, farklı parametreler ve farklı n'ler için elde edilmiş ve

kıyaslanmıştır. Sonuçta her durumda Z testinin gücü en yüksek çıkmıştır. n arttıkça ve kitle ortalamaları arasındaki fark arttıkça tüm testlerin gücü artmıştır. Bağımsız iki örneklem durumunda orana ilişkin parametrik testlerden Z testi ile parametrik olmayan testlerden Ki-kare testinin güçleri farklı parametreler ve farklı n 'ler için elde edilmiş ve kıyaslanmıştır. Sonuçta, tüm durumlarda Z testinin gücü daha yüksek çıkmıştır. Bununla birlikte, n arttıkça her iki testin de gücü artmıştır.

Bağımlı iki örneklem durumunda ortalama ve medyana ilişkin parametrik testlerden bağımlı gruplarda t testi ile parametrik olmayan testlerden Wilcoxon t testi ve İşaret testinin güçleri farklı parametreler, farklı ilişki yapıları ve farklı n 'ler için elde edilmiş ve kıyaslanmıştır. Sonuç olarak t testinin gücü en yüksek, işaret testinin gücü en düşük bulunmuştur. Gözlem sayısı arttıkça testlerin gücü artmaktadır. Bununla birlikte eşit varyans durumunda testlerin güçlerinin farklı varyans durumundaki testlerin güçlerinden daha yüksek olduğu ve ayrıca ilişki arttıkça testlerin güçlerinin arttığı görülmektedir. Bağımlı iki örneklem durumunda orana ilişkin parametrik testlerden t testi ile parametrik olmayan testlerden Mc-Nemar testinin güçleri farklı parametreler ve farklı n 'ler için elde edilmiş ve kıyaslanmıştır. Sonuçta, n arttıkça her iki testin de gücü artmıştır. Örnek hacmi çok küçükken $(n=10)$ Mc-Nemar testinin gücü daha yüksek iken, örnek hacmi 20 ve daha fazla olduğunda ise t testinin gücü daha yüksek çıkmıştır.

Bağımsız k örneklem durumunda, k=4 için ortalama ve medyana ilişkin parametrik testlerden tek yönlü varyans analizi ile parametrik olmayan testlerden Kruskal Wallis H testi ve Medyan testinin güçleri farklı dağılımlar, farklı parametreler ve farklı n 'ler için elde edilmiş ve kıyaslanmıştır. Sonuç olarak, Normal dağılımlı kitlelerde gücü en yüksek test ANOVA iken gücü en düşük test Medyan testi bulunmuştur. Ayrıca, kitle varyansları arttıkça testlerin güçleri azalmaktadır. Üstel dağılımlı kitlelerde ise, küçük örneklem durumunda birinci grubun parametre değeri diğer gruplardan küçükken Kruskal-Wallis H testinin ve Medyan testinin gücü ANOVA testinden yüksek, birinci grubun parametre değeri diğer gruplardan büyükken ANOVA testinin gücü diğerlerinden yüksek çıkmıştır. n arttıkça gücü en yüksek testin ANOVA, gücü en düşük testin Medyan testi olduğu görülmüştür. Ayrıca, bir grubun üstel dağılımlı diğer grupların eşit ortalamalı Normal dağılımlı olduğu durumda gücü en düşük testin Medyan testi olduğu gözlemlenmiştir. Üstel dağılımlı grubun θ parametresi normal dağılımlı diğer grupların ortalamasından küçükken Kruskal-Wallis

H testinin, büyükken ANOVA testinin gücünün daha yüksek olduğu görülmüştür. Ancak, ANOVA testinde anlamlılık seviyesi de nispeten büyük çıktığından örneklemlerin farklı dağılımlardan geldiği bu durum için Kruskal-Wallis H testinin kullanılmasının daha uygun olduğu söylenebilir. Bununla birlikte, tüm durumlarda n arttıkça testlerin gücü artmaktadır.

Bağımlı k örneklem durumunda ortalama ve medyana ilişkin parametrik testlerden iki yönlü varyans analizi ile parametrik olmayan testlerden Friedman'ın S testinin güçleri farklı parametreler, farklı ilişki yapıları ve farklı n'ler için elde edilmiş ve kıyaslanmıştır. Sonuç olarak, varyanslar eşit ve gruplar arasındaki düşük ilişki var iken genel olarak Friedman'ın S testinin gücü daha yüksek bulunmuştur. Orta düzeydeki ilişki yapılarında ise farklı grubun ortalaması ile diğer grupların ortalamaları arasındaki fark küçükken iki yönlü varyans analizi testinin gücünün, fark büyüdükçe Friedman'ın S testinin gücünün daha yüksek olduğu görülmüştür. Gruplar arasında yüksek ilişki var iken büyük n'lerde her iki testin de gücünün hemen hemen eşit olduğu ve istenilen testin tercih edilebileceği söylenebilir ancak gözlem sayısının 20'den küçük olduğu durumlarda ortalamalar arası fark küçükken iki yönlü varyans analizinin, büyükken Friedman'ın S testinin gücü daha yüksek bulunmuştur. Bununla birlikte gruplar arasındaki ilişki arttıkça testlerin gücü artmakta ve birbirine yaklaşmaktadır. Grupların varyansları birbirinden farklı olduğunda ise Friedman'ın S testinin gücü daha yüksek bulunmuştur. Tüm durumlarda n arttıkça testlerin gücü artmaktadır.

8.2 Öneriler

İstatistiksel sonuç çıkarımında doğru test istatistiklerinin kullanılması doğru sonuçlara ulaşmak açısından oldukça önemlidir. Testin gücü ne kadar yüksek olursa II.Tip hata yapma olasılığı o kadar düşük olacağından, testin gücü hangi durumlarda hangi testin seçileceğine karar vermede oldukça önemlidir. Bundan sonraki çalışmalarda farklı dağılım ve farklı parametreler için farklı hipotez testlerinin güç karşılaştırmaları yapılabilir.

KAYNAKLAR

Bindak, R., 2014, Mann-Whitney U ile Student'st testinin I. Tip Hata ve Güç bakımından Karşılaştırılması: Monte Carlo Simülasyon Çalışması (011302)(5-11), *Journal Afyon Kocatepe Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi*, 14 (1), 5-11.

Bircan, H., Karagöz, Y. ve Kasapoğlu, Y., Ki-Kare ve Kolmogorov Smirnov Uygunluk Testlerinin Simulasyon İle Elde Edilen Veriler Üzerinde Karşılaştırılması.

Campbell, C. J. ve Wasley, C. E., 1996, Measuring abnormal daily trading volume for samples of NYSE/ASE and NASDAQ securities using parametric and nonparametric test statistics, *Journal Review of Quantitative Finance Accounting*, 6 (3), 309-326.

Chan, Y., 2003, Biostatistics 102: quantitative data–parametric & non-parametric tests, *blood pressure*, 140 (24.08), 79.00.

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., Paxton, P. J. S. m. ve research, 2008, An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models, 36 (4), 462-494.

Conover, W. J., 1980, Practical nonparametric statistics.

Corrado, C. J., 1989, A nonparametric test for abnormal security-price performance in event studies, *Journal of financial economics*, 23 (2), 385-395.

Çapık, C., 2013, Bir Hemşirelik Dergisinde Yayınlanan Makalelerde İstatistiksel Güçlerin İncelenmesi, *Journal Anadolu Hemşirelik ve Sağlık Bilimleri Dergisi*, 16 (3), 170-175.

Çarkungöz, E. ve Bülent, E., 2009, Meta analizi, *Journal Uludağ Üniversitesi Veteriner Fakültesi Dergisi*, 28 (1), 33-37.

Çimen, M., 2016, Mühendislik Verilerinde Tek Örnek İçin Parametrik ve Parametrik Olmayan Testler, *Journal İstanbul Aydın Üniversitesi Dergisi*, 8 (29), 67-77.

Coşkun, K. ve Keskin, İ., 2008, Levene ve Bartlett Testleri Üzerine Bir İnceleme, *Journal Selcuk Journal of Agriculture Food Sciences* 22 (44), 78-83.

Demirel, N. ve Gürler, S., 2010, Klinik çalışmalarda örneklem genişliğinin belirlenmesine pratik yaklaşımlar, *Journal Kafkas Üniversitesi Vet. Fak. Dergisi*, 15 (2), 205-211.

Doğan, İ. ve Doğan, N., 2015, Parametrik Olmayan Bağımsız İki Örneklem Testlerinde Örneklem Büyüklüğünün Beklenen Değer, Varyans ve Fraser Etkinliği Üzerine Etkileri, *Journal Turkiye Klinikleri Journal of Biostatistics*, 7 (2).

Faul, F., Erdfelder, E., Buchner, A. ve Lang, A.-G., 2009, Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses, *Journal Behavior research methods*, 41 (4), 1149-1160.

Finch, H., 2005, Comparison of the performance of nonparametric and parametric MANOVA test statistics when assumptions are violated, *Journal Methodology*, 1 (1), 27-38.

Gamgam, H. ve Altunkaynak, B., 2012, Parametrik olmayan yöntemler: SPSS uygulamalı, Seçkin Yayıncılık, p.

García, S., Molina, D., Lozano, M. ve Herrera, F., 2009, A study on the use of non-parametric tests for analyzing the evolutionary algorithms' behaviour: a case study on the CEC'2005 special session on real parameter optimization, *Journal of Heuristics*, 15 (6), 617.

Genç, S. ve Soysal, M. İ., Parametrik Ve Parametrik Olmayan Çoklu Karşılaştırma Testleri.

Kaplan, M., Kut, M. S., Çimen, S. ve Demirtaş, M. M., 2003, EuroSCORE (European System for Cardiac Operative Risk Evaluation) risk skorlama sisteminin ülkemiz hasta profilinde uygulanabilirliğinin araştırılması, *Türk Göğüs Kalp Damar Cer Derg*, 11, 147-158.

Karagöz, Y., 2010, Nonparametrik tekniklerin güç ve etkinlikleri, *Elektronik Sosyal Bilimler Dergisi*, 9 (33), 18-40.

Karakoç, A. G. D. F. Y. ve Dönmez, L., 2014, Ölçek geliştirme çalışmalarında temel ilkeler, *Tıp Eğitimi Dünyası*, 13 (40), 39-49.

Kartal, M., 2006, Bilimsel araştırmalarda hipotez testleri, *Nobel Yayın Dağıtım, Ankara*.

Kim, T. K., 2015, T test as a parametric statistic, *Korean journal of anesthesiology*, 68 (6), 540.

Mendeş, M. ve Akkartal, E., 2010, Comparison of ANOVA F and WELCH tests with their respective permutation versions in terms of type I error rates and test power, *Kafkas Univ Vet Fak Derg*, 16 (5), 711-716.

Miniç, M. A., 2015, Parametrik olmayan çoklu karşılaştırmalarda p-değeri bazlı düzeltme yöntemlerinin performanslarının incelenmesi.

Önder, H., Bİyolojik Araştırmalarda Kullanılan Parametrik Olmayan İstatistik Yöntemler.

Özdamar, K., 2002, Paket Programlar İle İstatiksel Veri Analizi: Spss-Minitab (Çok Değişkenli Analizler), Kaan Kitabevi, p.

Razali, N. M. ve Wah, Y. B., 2011, Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests, *Journal of statistical modeling analytics*, 2(1), 21-33.

Sarı, H., Gökpınar, E. ve Bayrak, H. J. J. o. N., 2018, Çapraz Tasarımlarda Etkilerin Testi Için Parametrik Ve Parametrik Olmayan Testlerin Karşılaştırılması, *Journal of Natural Applied Sciences*, 22 (1), 326-335.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı	**:**	Nazik Özge CEYLAN
Uyruğu	**:**	T.C.
Doğum Yeri ve Tarihi	**:**	Konya/10.08.1992
Telefon	**:**	0554 681 49 37
Faks	**:**	
e-mail	**:**	ozgeceylan9205@gmail.com

EĞİTİM

Derece	**Adı, İlçe, İl**	**Bitirme Yılı**
Lise :	**Selçuklu Atatürk Lisesi, Selçuklu, Konya**	**2010**
Üniversite :	**Selçuk Üniversitesi, Selçuklu, Konya**	**2014**
Yüksek Lisans :	**Selçuk Üniversitesi, Selçuklu, Konya**	**Devam ediyor**
Doktora :		

İŞ DENEYİMLERİ

Yıl	**Kurum**	**Görevi**
2016-2019	**Talipsan Mağazacılık A.Ş.**	**Bütçe Planlama ve Ürün Yönetim Uzm.**

YABANCI DİLLER

İngilizce

YAYINLAR

Ceylan, N. Ö., Sezer, D. ve Kınacı, İ., 2019, Tek Örneklem Testlerinde Güç Karşılaştırması, International Science and Academic Congress, 19-20 Nisan 2019, Konya.

.