



**HACETTEPE ÜNİVERSİTESİ**  
**EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

Eğitim Bilimleri Ana Bilim Dalı  
Eğitimde Ölçme ve Değerlendirme Programı

PARAMETRİK VE PARAMETRİK OLMAYAN MADDE TEPKİ MODELLERİNİN  
KESTİRİM HATALARININ KARŞILAŞTIRILMASI

Nermin KIBRISLIOĞLU UYSAL

Doktora tezi

Ankara, 2020

Liderlik, arařtırma, inovasyon, kaliteli eęitim ve deęişim ile

*Daha ileriye ... En İyiyeye ...*



**HACETTEPE ÜNİVERSİTESİ**  
**EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

Eğitim Bilimleri Ana Bilim Dalı  
Eğitimde Ölçme ve Değerlendirme Programı

PARAMETRİK VE PARAMETRİK OLMAYAN MADDE TEPKİ MODELLERİNİN  
KESTİRİM HATALARININ KARŞILAŞTIRILMASI

COMPARISON OF ESTIMATION ERRORS IN PARAMETRIC AND  
NONPARAMETRIC ITEM RESPONSE THEORY MODELS

Nermin KIBRISLIOĞLU UYSAL

Doktora Tezi

Ankara, 2020

## Kabul ve Onay

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,

Nermin KIBRISLIOđLU UYSAL'ın hazırladıđı "Parametrik ve Parametrik Olmayan Madde Tepki Modellerinin Kestirim Hatalarının Karşılaştıırılması" bařlıklı bu alıřma j¼rimiz tarafından **Eđitim Bilimleri Ana Bilim Dalı, Eđitimde Ölme ve Deđerlendirme Bilim Dalında Doktora Tezi** olarak kabul edilmiřtir.

J¼ri Bařkanı	Prof. Dr. H¼lya KELECİOđLU	İmza
J¼ri Üyesi (Danıřman)	Prof. Dr. Selahattin GELBAL	İmza
J¼ri Üyesi	Prof. Dr. Nuri DOđAN	İmza
J¼ri Üyesi	Prof. Dr. Cem Oktay G¼ZELLER	İmza
J¼ri Üyesi	Prof. Dr. Hakan Yavuz ATAR	İmza

Bu tez Hacettepe Üniversitesi Lisans¼st¼ Eđitim, Öğretim ve Sınav Yönetmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri üyeleri tarafından 17/ 06 /2020 tarihinde uygun gör¼lm¼ř ve Enstit¼ Yönetim Kurulunca ..... / ..... / ..... tarihi itibarıyla kabul edilmiřtir.

Prof. Dr. Selahattin GELBAL  
Eđitim Bilimleri Enstitüsü M¼d¼r¼

## Öz

Araştırmanın amacı parametrik ve parametrik olmayan madde tepki kuramı modellerinin farklı örtük özellik dağılımları, madde güçlük dağılımları, madde sayısı ve örneklem büyüklüğü koşullarında karşılaştırılmasıdır. Kestirim yöntemi olarak MLE, MAP, DE, AH ve NBE kullanılmıştır. Araştırma kapsamında toplam 405 koşul incelenmiştir. Simülasyon verilerine ek olarak gerçek verilerle de kestirimler yapılmıştır. MLE ve MAP kestirimleri için R programında yer alan “irtoys” paketi, DE ve AH kestirimleri için R programında yer alan “mirt” paketi kullanılmıştır. NBE kestirimleri için ise araştırmacı tarafından bir R kodu hazırlanmıştır. Model karşılaştırma sonuçları ortalama mutlak hata (MAE) ve hata kareler ortalaması (RMSE) hesaplanarak değerlendirilmiştir. Ayrıca hangi koşulların hatalar üzerinde anlamlı etkisinin olduğunun belirlenmesi için varyans analizinden (ANOVA) yararlanılmıştır. Gerçek veriler ile yapılan kestirimlerin karşılaştırılmasında ise model uyumları temel alınmıştır. Araştırmanın sonuçlarına göre kestirim yöntemi, örneklem büyüklüğü, madde sayısı ve örtük özellik dağılımı ve bu değişkenlerin bazı etkileşimleri hem madde hem de birey parametreleri üzerinde etkilidir. Diğer yandan b parameresinin dağılımı yetenek kestirimleri üzerinde etkili bulunmazken, örtük özelliğin normal dağılmadığı durumlarda madde parametre kestirimlerini etkilemiştir. Örtük özellik dağılımı çarpık olduğundan en az hatalı kestirimler DE yöntemi ile elde edilirken genel olarak NBE yöntemi en yüksek hatalı kestirimleri üretmiştir. Gerçek verilerin analizlerinde ise simülasyon sonuçlarına paralel şekilde en az hatalı kestirimler DE ve AH yöntemleri ile elde edilmiştir. Gerçek verilere ilişkin yetenek dağılımları ise parametrik olan ve olmayan modeller arasında farklılaşmaktadır.

**Anahtar sözcükler:** parametrik olmayan bayesçi kestirim, David eğrileri, ampirik histogram, parametrik MTK, parametrik olmayan MTK

## Abstract

The purpose of this study is to compare parametric and nonparametric item response theory models under different ability distributions, difficulty distributions, item numbers, and sample sizes. MML, MAP, DC, EH, and NBE estimation procedures were used. Results were examined in 405 simulation conditions. Monte Carlo simulation method was used in data generation. Moreover, empirical data was also used to compare estimation methods. “irtoys” package in R was used in MLE and MAP estimations, and “mirt” package was used in DC and EH methods. For NBE estimations, an R code was developed by the researcher. Model comparisons were made based on MAE and RMSE values. Furthermore, ANOVA analysis was conducted to evaluate which manipulated factors are significantly related to outcomes. In the evaluation of real data sets, model fit statistics were examined. The results indicated that estimation methods, sample size, item number, distributional properties of latent trait, and some cross combination of them affected both item and person parameter estimates. On the other hand, the distributional properties of the difficulty parameter did not affect person parameter estimate while it affected item parameter estimates under latent trait nonnormality. When the latent trait was not normally distributed, DC provided the least estimation error. In general, NBE method produces the highest estimation error. The analysis of empirical data yielded similar results with the simulation study. DC and EH methods were the ones that provided minimum error. The distributional properties of ability estimated in empirical data differed explicitly between parametric and nonparametric estimation methods.

**Keywords:** nonparametric Bayesian estimation, Davidian curves, empirical histogram, parametric IRT, nonparametric IRT.

## Teşekkür

Bu zorlu süreçte beni pek çok insan destekledi.

Hem tez sürecinde hem de lisansüstü eğitimim boyunca desteğini ve iyi niyetini esirgemeyen, bana yol gösteren, bilgi ve tecrübelerini paylaşan çok kıymetli danışmanım Prof. Dr. Selahattin GELBAL'a;

Tezimin gelişmesinde çok önemli katkıları olan kıymetli jüri üyelerim Prof. Dr. Hülya KELECİOĞLU, Prof. Dr. Nuri DOĞAN, Prof. Dr. Cem Oktay GÜZELLER ve Prof. Dr. Hakan Yavuz ATAR'a;

Tezime olan katkıları ve Amerika'ya uyum sağlamamdaki destekleri için Dr. Scott MONROE'ye;

Tezimi okuyarak önerileri ile katkıda bulunan, Hacettepe'de çalışmaya başladığım ilk günden beri bana yol gösteren, dostluğunu ve iyi niyetini esirgemeyen arkadaşım Dr. Kübra ATALAY KABASAKAL'a;

Bugünlere gelmemde büyük payı olan sevgili annem Meryem KIBRISLIOĞLU ve babam Şefik KIBRISLIOĞLU'na, her zaman yanımda olan, desteğini her zaman hissettiğim sevgili kardeşim Şeyma KIBRISLIOĞLU'na;

Çalışma sürecinde ve hayatın her alanında beni destekleyen, motive eden, anlayışını esirgemeyen, varlığıyla bana kendimi iyi hissettiren yol arkadaşım, kıymetli eşim Halil UYSAL'a;

Lisansüstü eğitimim boyunca derslerini aldığım, birlikte çalıştığım, yetişmemde emeği geçen, bilgisini ve tecrübesini benimle paylaşan, beni destekleyen kıymetli hocalarıma;

Doktora eğitimin boyunca yurt içi doktora burs desteği sağlayan TÜBİTAK'a,

Çok teşekkür ederim.

## İçindekiler

Öz.....	ii
Abstract.....	iii
Teşekkür.....	iv
İçindekiler.....	v
Tablolar Dizini.....	vii
Şekiller Dizini.....	viii
Simgeler ve Kısaltmalar Dizini.....	ix
Bölüm 1 Giriş.....	1
Problem Durumu.....	3
Araştırmanın Amacı ve Önemi.....	5
Problem Cümlesi.....	6
Sınırlılıklar.....	7
Bölüm 2 Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar.....	8
Araştırmanın Kuramsal Temeli.....	8
İlgili Araştırmalar.....	19
Bölüm 3 Yöntem.....	26
Araştırma Yöntemi.....	26
Araştırmanın Koşulları.....	26
Verilerin Üretilmesi.....	27
Verilerin Analizi.....	31
Bölüm 4 Bulgular.....	36
Birinci alt probleme ilişkin bulgular.....	36
İkinci Alt probleme ilişkin bulgular.....	54
Üçüncü Alt Probleme İlişkin Bulgular.....	59
Bölüm 5 Sonuç, Tartışma ve Öneriler.....	64
Öneriler.....	68



Kaynaklar .....	71
EK-A: Madde Parametrelerinin ve Cevap Matrisinin Üretilmesi Örnek R kodu .....	75
EK-B: MLE, MAP, DE ve AH Kestirimleri Örnek R kodu.....	76
EK-C: NBE Kestirimi R Kaynak Kodu .....	77
EK-D: NBE Kestirimi MATLAB kodu.....	84
EK-E: Parametrelerin Ölçeklenmesi, RMSE ve MAE Değerlerine İlişkin R Kodu ..	89
EK-F: Grafiklerin Oluşturulmasına Yönelik Örnek R kodu .....	90
EK-G: Yetenek Parametre Kestirimlerine İlişkin RMSE ve MAE Değerleri .....	91
EK-H: Madde Parametre Kestirimlerine İlişkin RMSE ve MAE Değerleri .....	100
EK-I: Etik Komisyonu Onay Bildirimi.....	109
EK-J: Etik Beyanı .....	110
EK-K: Yüksek Lisans/Doktora Tez Çalışması Orijinallik Raporu .....	111
EK-L: Thesis/Dissertation Originality Report .....	112
EK-M: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı .....	113

## Tablolar Dizini

Tablo 1	<i>Araştırmanın Koşulları</i> .....	28
Tablo 2	<i>10 Maddelik Test için Madde Parametreleri</i> .....	29
Tablo 3	<i>20 Maddelik Test için Madde Parametreleri</i> .....	29
Tablo 4	<i>30 Maddelik Test için Madde Parametreleri</i> .....	30
Tablo 5	<i>Türkçe ve Matematik testleri ham puanları betimsel istatistikleri</i> .....	33
Tablo 6	<i>Türkçe ve Matematik testlerine ilişkin faktör analizi sonuçları</i> .....	34
Tablo 7	<i>Yetenek parametreleri RMSE değerlerine ilişkin ANOVA tablosu</i> .....	37
Tablo 8	<i>Kestirim Yönlere Arasında Post-Hoc analizi sonuçları</i> .....	38
Tablo 9	<i>Yetenek parametreleri MAE değerlerine ilişkin ANOVA tablosu</i> .....	40
Tablo 10	<i>Kestirim Yönlere Arasında Post-Hoc analizi sonuçları</i> .....	41
Tablo 11	<i>a parametreleri RMSE değerlerine ilişkin ANOVA tablosu</i> .....	43
Tablo 12	<i>Kestirim Yönlere Arasında Post-Hoc analizi sonuçları</i> .....	44
Tablo 13	<i>a parametrelerinin MAE değerlerine ilişkin ANOVA tablosu</i> .....	47
Tablo 14	<i>b parametreleri RMSE değerlerine ilişkin ANOVA tablosu</i> .....	49
Tablo 15	<i>Kestirim Yönlere Arasında Post-Hoc analizi sonuçları</i> .....	50
Tablo 16	<i>b parametreleri MAE değerlerine ilişkin ANOVA tablosu</i> .....	52
Tablo 17	<i>Ham puanlara ilişkin betimsel istatistikler</i> .....	60
Tablo 18	<i>Kestirime ilişkin BIC değerleri</i> .....	61

## Şekiller Dizini

Şekil 1. Yetenek parametresine ilişkin RMSE değerlerinin yetenek parametresinin dağılımı, madde sayısı ve örneklem büyüklüğüne göre değişimi .....	39
Şekil 2. Yetenek parametresine ilişkin MAE değerlerinin yetenek parametresinin dağılımı ve madde sayısına göre değişimi .....	42
Şekil 3. a parametresinin RMSE değerlerinin madde sayısı, örneklem büyüklüğü ve yetenek parametresi dağılımına göre değişimi .....	45
Şekil 4. Ayırıcılık parametresinin MAE değerlerinin madde sayısı, örneklem büyüklüğü ve yetenek parametresi dağılımına göre değişimi .....	48
Şekil 5. Güçlük parametresinin RMSE değerlerinin madde sayısı, örneklem büyüklüğü ve yetenek parametresi dağılımına göre değişimi .....	51
Şekil 6. b parametresinin MAE değerlerinin madde sayısı, örneklem büyüklüğü ve yetenek parametresi dağılımına göre değişimi .....	53
Şekil 7. Yetenek parametresine ilişkin RMSE ve MAE değerlerinin yetenek parametresinin dağılımı, madde sayısı ve güçlük parametresinin dağılımına göre değişimi .....	55
Şekil 8. a parametresine ilişkin RMSE ve MAE değerlerinin yetenek parametresinin dağılımı, madde sayısı ve b parametresinin dağılımına göre değişimi .....	57
Şekil 9. b parametresine ilişkin RMSE ve MAE değerlerinin yetenek parametresinin dağılımı, madde sayısı ve b parametresinin dağılımına göre değişimi .....	58
Şekil 10. Türkçe ve Matematik ham puan dağılımları .....	59
Şekil 11. Türkçe yetenek kestirimleri .....	62
Şekil 12. Matematik yetenek kestirimleri.....	63

## **Simgeler ve Kısaltmalar Dizini**

**DE:** Davidian Eğrileri

**AH:** Ampirik Histogram

**KTK:** Klasik Test Kuramı

**MAP:** Maksimum Sonsal

**MLE:** Maksimum Olabilirlik

**MTK:** Madde Tepki Kuramı

**NBE:** Parametrik olmayan Bayesçi Kestirim

**PoMTK:** Parametrik olmayan Madde Tepki Kuramı

**2PLM:** İki parametrelili lojistik model

## Bölüm 1

### Giriş

Klasik test kuramına (KTK) göre bireylerin gözlenen puanları onların gerçek puanları ile ölçmede yapılan hatanın toplamından oluşur. Gerçek puan ise ölçmelerin aynı koşullarda sonsuz sayıda tekrar edilmesi ile elde edilecek olan gözlenen puanların ortalaması olarak tanımlanmaktadır (Crocker ve Algina, 1986). Yani gerçek puan gözlenen puanın beklenen değeridir ve hipotetiktir. Ancak bireyler üzerinde yapılan tekrarlı ölçümler bir takım problemleri de beraberinde getirir. Ölçülen özellikler insanlara ait olduğundan bir ölçme durumunun aynı bireye aynı şekilde birden çok defa uygulanmasının getirebileceği olumsuzluklar vardır. Uygulamalar aynı şekilde yapılsa bile bireyler soruları hatırlayabilir veya uygulamalar arasında geçen sürede birey özellikleri değişebilir. Dolayısıyla tekrarlı ölçmelerde paralelliğin sağlanması ve uygulama problemlerinden dolayı kuram, bir grup birey üzerinde yapılan ölçmelere dayandırılmıştır (Baykul, 2015). Bu nedenle KTK'da bireylerin kestirilen gerçek puanları içinde buldukları grup ve uygulanan testin güvenilirliği ile sınırlıdır (Baykul, 2015).

Madde Tepki Kuramı (MTK), KTK'nın test özellikleri ve birey özelliklerini tam olarak ayıramaması ve tüm bireyler için eşit hata kestirmesi (Hambelton, Swaminathan ve Rogers, 1991) gibi sınırlılıklarına bir çözüm getirdiğinden son yıllarda uluslararası sınavlarda, test eşitleme, bireyselleştirilmiş bilgisayar tabanlı test uygulamaları gibi pek çok alanda yaygın olarak kullanılmaktadır. MTK, bireylerin bir testteki performanslarının, onların yetenekleri ya da özellikleri tanımlandıktan sonra bu yetenekleri üzerindeki kestirilen puanları test performanslarını açıklamak için kullanarak, tahmin edilebileceğini varsayar (Lord ve Novick, 1968, akt. Hambelton ve Swaminathan, 1985). Bireylerin özellikleri doğrudan gözlenemediği için gizil özellik ya da yetenek olarak adlandırılır. Madde tepki modelleri gözlenen test performansı ile testin ölçtüğü varsayılan gözlenemeyen yetenek arasında bir ilişki tanımlar. Bu ilişki matematiksel bir fonksiyonla tanımlanır. İlk madde tepki fonksiyonları olan normal ogive fonksiyonlar Lord (1952) tarafından geliştirilmiştir. Birnbaum (1957) normal ogive fonksiyonları lojistik fonksiyonlarla değiştirerek, matematiksel olarak daha kolay hesaplanabilen lojistik modelleri geliştirmiştir (Akt. Liang, 2007). Günümüzde de yaygın olarak kullanılan modeller lojistik modellerdir.

Standart MTK modellerinin üç temel varsayımı vardır. Bunlar: tek boyutluluk, yerel bağımsızlık ve monoton artanlıktır (Hambelton, Swaminathan ve Rogers, 1991). Tek boyutluluk varsayımı testin baskın bir tek boyutlu gizil özelliği/yeteneği ölçtüğünü varsayar. Yerel bağımsızlık bireylerin maddelere verdikleri cevapların istatistiksel olarak bağımsız olmasıdır. Diğer bir ifade ile aynı yetenek düzeyindeki bireylerin farklı maddelere verdikleri cevaplar arasında bir ilişki bulunmamasıdır. Monoton artanlık ise bireylerin yetenek düzeyleri ile bir maddeye doğru cevap verme olasılıkları arasındaki fonksiyonun monoton artan olmasını gerektirir (Hambelton ve Swaminathan, 1985). Lojistik modellerde bu fonksiyon “S” şeklinde bir eğriye karşılık gelir (Hambelton, Swaminathan ve Rogers, 1991). MTK modellerinde elde edilen sonuçların doğruluğu varsayımların sağlanmasıyla doğrudan ilişkilidir. Ancak pratikte gerçek veriler ile çalışıldığında bu varsayımların sağlanması zordur. Varsayımlar sağlanmadığında ise daha doğru sonuçlar elde edilebilmesi amacıyla farklı modeller önerilmektedir.

MTK modellerinin varsayımları sağlanmadığında parametrik olmayan MTK (PoMTK) modelleri iyi bir alternatif olarak kullanılabilirler. PoMTK modelleri üç nedenden dolayı oldukça kullanışlıdır: Parametrik MTK modellerinin ne olduğu hakkında derin bir bilgi sağlar. Parametrik modeller zayıf uyum gösterdiğinde uygulama için daha esnek bir teorik çerçeve sunar. Daha az sayıda birey ve madde ile kullanım sunar (Junker ve Sijtsma, 2001). PoMTK modellerini parametrik yapmayan en temel özelliği monotonluk varsayımı olmakla beraber lojistik gibi parametrik bir fonksiyon şartı yoktur (Sijtsma, 2005). Yani PoMTK modellerinde madde karakteristik eğrileri herhangi bir monoton formda olabilir. Bu bağlamda PoMTK modelleri daha geneldir ve daha az varsayıma dayandığı için daha çok sayıda madde modele uyum sağlayabilir (Sijtsma, Emons, Bouwmeester, Nylicek ve Roorda, 2008). Yani, bir uygulamadaki maddelerin tamamı parametrik bir modele uyum sağlamadığında, parametrik olmayan modeller kullanılarak verinin uyumu artırılabilir ve böylece parametrik modele uyum sağlamayan maddelerden elde edilecek bilgi de kullanılmış olur (Stout, 2001).

Parametrik MTK modellerinde örtük özelliğin tek boyutlu olmadığı durumlar için çok boyutlu madde tepki modelleri geliştirilmiştir (Reckase, 2009). Yerel bağımsızlık ve monoton artanlık varsayımlarının sağlanmadığı durumlarda ise PoMTK modelleri önerilir. Yerel bağımsızlık MTK kestirimlerinin madde

özelliklerinden bağımsızlığının göstergesi olduğundan oldukça önemli bir varsayımdır. Ancak parametrik olmayan madde tepki kuramı modellerinde yerel bağımsızlık varsayımı temel yerel bağımsızlık (essential local independence) olarak zayıflatılabilmektedir. Temel yerel bağımsızlık maddeler arasında bir miktar pozitif veya negatif ilişkiye izin verirken toplamda maddeler arasındaki ilişkinin sıfır olmasını gerektirir (Sijtsma, 2005). Bazı durumlarda da lojistik modeller gerçek madde fonksiyonunu yakalayamayabilirler ve bu durum yetenek kestirimlerinde yanlılığa neden olur (Duncan ve MacEachern, 2013). Bu gibi durumlar için ise monoton artanlık varsayımı parametrik olmayan modellerde zayıf monoton artanlık olarak esnetilebilmektedir. Zayıf monoton artanlık madde karakteristik eğrilerinin ortalamada monotonluğu sağlamak koşulu ile bölgesel olarak artanlığın sağlanmamasını tolere edebilir. (Sijtsma, 2005).

### **Problem Durumu**

MTK uygulamalarında genellikle madde ve birey parametre kestirimleri ardışık olarak yapılır. Ancak söz konusu kestirimlerin yapılabilmesi ve madde karakteristik eğrilerinin tanımlanabilmesi için yetenek dağılımının ölçeklenmesi gerekir. MTK modellerinde gözlenen puan dağılımına yönelik bir varsayım olmamakla birlikte, kestirimler yapılırken genellikle örtük özelliğin normal dağıldığı varsayılır (Duncan ve MacEachern, 2013). Bireyler evrenden rastgele örneklemlendiğinde genellikle bu varsayım sağlanır ve bu durum herhangi bir probleme neden olmaz (Sass, Schmitt ve Walker, 2005). McDonald (1982) ise madde parametrelerinin yalnızca evrenden rastgele seçilen bir örneklem üzerinden kestirilmesi ve bu örneklemin yalnızca madde kalibrasyonunda kullanılması gerektiğini belirtmektedir. Madde parametreleri bu örneklemden elde edildikten sonra yetenek kestirimlerinde yine aynı parametreler kullanılmalıdır. Ancak bu öneri pratikte pek fazla uygulanmamaktadır. Dolayısıyla, rastgele örnekleme yapılmadığında ya da evrenin kendisi normal dağılmadığında bu durum problemlere neden olur. Örneğin depresyon, kaygı gibi bazı psikolojik özellikler, sağlıkla ilişkili yapılar (Revicki ve Reise, 2015) çoğunlukla evrende normal dağılmazlar ya da belirli bir başarı sınıflamasına göre seçilmiş öğrenciler araştırma konusu olduğunda yetenek dağılımı çarpıklaşır (Sass ve diğerleri, 2005). Bu gibi örtük özelliğin normal dağılmadığı durumlarda, kestirimler normal dağılım varsayımı ile yapıldığında parametre kestirimleri yanlı olur. Örtük özellik dağılımının çarpıklığı artması ile

orantılı yanlılığın miktarı artar (Stone, 1992; Woods, 2006, 2008; Woods ve Linn, 2009).

Örtük özelliğin normal dağılmaması durumunun MTK madde ve yetenek parametrelerinin kestirimlerine etkisi alan yazında farklı açılardan değerlendirilmiştir. Stone (1992) çalışmasında marjinal maksimum olabilirlik (MML), madde parametre kestirimi ve maksimum olabilirlik (ML) yetenek parametre kestirimini farklı örneklem büyüklükleri, test uzunlukları ve yetenek dağılımlarında karşılaştırmıştır. Simülasyon verisi kullanılan araştırmanın sonuçları MML ile madde güçlük parametresi kestirimlerinin küçük örneklem (250), kısa test (10) ve farklı yetenek dağılımlarında kesin ve tutarlı sonuçlar verdiğini göstermektedir. Ancak dağılım çarpıklaştıkça örneklem büyüklüğünden bağımsız olarak 10 ve 20 maddelik testlerden elde edilen yanlılığın arttığı bulunmuştur. Çarpık ve basık yetenek dağılımları madde ayıricılıkları üzerinde ters etki yapmıştır. Woods ve Lin (2009) araştırmalarında DE, Ramsay-curve IRT (RC-IRT) ve ampirik histogram yöntemini (AHM) normal ve çarpık dağılımlarda karşılaştırmıştır. Araştırmanın bulguları örtük değişkenin dağılımı normal olmadığında üç yöntemin de normal modellere göre daha iyi kestirimler yaptığını göstermiştir. Diğer yandan Kirisci, Hsu ve Yu (2001) araştırmalarında madde sayısı 40 ve örneklem büyüklüğü 1000 olduğunda yetenek dağılımının normal, pozitif çarpık ve sivri olduğu durumlarda madde parametrelerini parametrik yöntemlerle kestirmişler ve MAP kestirimlerinin dağılımın çarpıklığını elimine edebildiğini raporlamıştır. Stone, (1992) ve Seong (1990) ise testin uzunluğu ve örneklem büyüklüğü arttıkça madde parametrelerinde kestirim hatasının azaldığını raporlamıştır.

MTK modellerinde yetenek parametresi ve madde güçlük parametresi aynı ölçekte kestirilir. Bu nedenle son dönemlerde araştırmalarda yetenek dağılımının yanı sıra madde güçlük parametresinin dağılımının kestirime etkisi de incelenmiştir. Sass ve diğerleri, (2005) iki parametrelili model için simülasyon verileriyle yetenekteki ve madde güçlüğündeki çarpıklığın MML, normal önsel MAP ve çarpık önsel EAP kestirimlerini nasıl etkilediğini incelemiştir. Araştırmanın bulguları normal dağılmayan örtük değişkende kestirim hatasının normal dağılıma sahip değişkenin kestirime göre çok fazla olduğunu göstermiştir. Ayrıca madde güçlük parametresinin çarpıklığı örtük özellik kestirimini etkilemiş ve en yüksek kestirim hatası örtük özellik ve madde güçlüklerinin her ikisinin de pozitif çarpık



dağıldığı koşulda gözlenmiştir. Finch ve Edwards (2016) ise Sass ve diğerlerinin (2005) çalışmasındaki simülasyon koşullarını kullanarak verileri üretmiş ve parametrik olmayan Bayes yaklaşımını da kestirim yöntemlerine dahil etmiştir. Çalışmanın bulgularına göre madde güçlük dağılımı ve kestirim yöntemi kesişim etkisi parametre kestirimindeki hatayı manidar düzeyde etkilemektedir. Ayrıca bulgular madde güçlük dağılımı ve yetenek çarpık olduğunda en iyi kestirimi parametrik olmayan Bayes yaklaşımının verdiğini göstermiştir.

### **Araştırmanın Amacı ve Önemi**

Küçük örneklerle veya az sayıda madde ile çalışıldığında veya parametrik MTK modellerinin varsayımları sağlanamadığında, PoMTK modellerinin, etkili sonuçlar vermesi bu yöntemin önemli avantajları arasında görülmektedir. Alan yazında parametrik MTK ve PoMTK kestirimlerinin performansı farklı koşullarda karşılaştırılmış ve PoMTK model kestirimlerinin belirli koşullarda parametrik yöntemlere alternatif sunabileceği belirtilmiştir (Bıkmaz Bilgen, 2016; Douglas ve Cohen, 2001; Finch ve Edwards, 2016; Liang, 2007; Mor Dirlik, 2017; Şengül Avşar ve Tavşancıl, 2016; Vermunt, 2001). Bu çalışmada ise PoMTK ve parametrik MTK model kestirimlerinin performansları alan yazındaki koşullar genişletip incelenerek alana katkı sağlanması beklenmektedir. Dolayısıyla, araştırmanın birinci amacı iki parametrelili model için farklı örtük özellik ve yetenek dağılımlarının, farklı örneklem büyüklüğü ve madde sayısı koşullarında parametrik ve parametrik olmayan madde tepki kuramı model kestirimlerindeki hatayı nasıl etkilediğinin simülasyon verileri ile incelenmesidir. Araştırmanın ikinci amacı ise simülasyon çalışmasından elde edilen bulguların gerçek veri setlerinde incelenmesidir.

Araştırma kapsamında kestirimlerde maksimum olabilirlik (MLE), Maksimum sonsal (MAP), parametrik olmayan Bayesçi kestirim (NBE), David eğrileri (DE), Ampirik histogram (AH) yöntemleri kullanılmıştır. Alan yazında NBE yaklaşımının yalnızca Rasch model için uygulanabilir yazılımları mevcuttur. Dolayısıyla iki parametrelili model için NBE kestirimlerinin diğer yöntemlerle karşılaştırılmasına yönelik herhangi bir çalışmaya rastlanmamıştır. Bu nedenle çalışmada kestirimler 2PL modele dayalı olarak yapılacaktır. Araştırma kapsamında iki parametrelili model ile NBE kestirimlerinin yapılmasına yönelik bir R kodu hazırlanmıştır.

Ayrıca alan yazındaki çalışmaların bulguları çoğunlukla simülasyon verisine dayanmaktadır (Bıkmaz Bilgen ve Doğan, 2017; Finch ve Edwards, 2016; Koğar, 2015; San Martin ve diğerleri, 2011; Sass ve diğerleri, 2005; Şengül Avşar ve Tavşancıl, 2016). Bu araştırmada simülasyon çalışmasına ek olarak kestirimler gerçek veri kümeleri üzerinde de karşılaştırılmıştır. Bu bağlamda araştırma kapsamında simülasyon çalışmasına ek olarak gerçek uygulama verilerine ilişkin elde edilerek bulguların alana katkı sağlayacağı düşünülmektedir.

Bunlara ek olarak ülkemizde PoMTK dayalı çalışmalar oldukça sınırlıdır. Bıkmaz Bilgen ve Doğan, (2017) ve Koğar (2015) çok kategorili maddelerde parametrik yöntemler ile parametrik olmayan monoton homojenlik modelini (MHM) karşılaştırmıştır. Şengül Avşar ve Tavşancıl (2016) parametrik olmayan Kernel yöntemi ve MHM yöntemlerini farklı koşullarda karşılaştırmıştır. Dolayısıyla bu çalışma ülkemizde NBE ve DE yönteminin kullanıldığı ve örtük özellik ve yetenek parametre dağılımlarının birlikte manipüle edildiği ilk çalışma olması bağlamında alana katkı sağlayacaktır.

### **Problem Cümlesi**

İki parametrelili modelde farklı örtük özellik dağılımı, b parametresi dağılımı, örneklem büyüklüğü ve madde sayısı koşullarında üretilen simülasyon ve gerçek veri kümelerinde parametrik ve parametrik olmayan MTK model kestirimleri farklılaşmakta mıdır?

**Alt problemler.** Araştırma kapsamında ele alınan alt problemler:

- 1) Normal ve normal olmayan örtük özellik dağılımlarında, farklı örneklem büyüklükleri (100,500 ve 1000) ve test uzunluklarında (10,20 ve 30) parametrik MML, MAP ve parametrik olmayan NBE, DE ve AH yöntemlerinden hangisi:
  - a. daha yansız yetenek kestirimleri üretir?
  - b. daha yansız madde parametre kestirimleri üretir?
- 2) Normal ve normal olmayan madde güçlük dağılımlarında farklı örneklem büyüklükleri (100,500 ve 1000) ve test uzunluklarında (10,20 ve 30) parametrik MML, MAP ve parametrik olmayan NBE, DE ve AH yöntemlerinden hangisi:
  - a. daha yansız yetenek kestirimleri üretir?
  - b. daha yansız madde parametre kestirimleri üretir?

- 3) Gerçek veri kümelerinde parametrik MML, MAP ve parametrik olmayan NBE, DE ve AH yöntemlerinden hangisi daha yansız kestirimler üretir?

### **Sınırlılıklar**

Araştırma kapsamında parametrik model kestirimleri MML ve MAP ile; parametrik olmayan model kestirimleri NBE, DE ve AH yöntemleri ile sınırlıdır.

## Bölüm 2

### Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

#### Araştırmanın Kuramsal Temeli

Araştırma kapsamında tek boyutlu ve ikili puanlanan MTK modelleri temel alındığından araştırmanın kuramsal temeli bu bağlamda açıklanmıştır. İlk olarak parametrik, ve parametrik olmayan madde tepki modellerinden bahsedilmiş ve araştırma kapsamına dahil edilen modellere ilişkin açıklama yapılmıştır.

**Parametrik madde tepki modelleri.** Parametrik ve parametrik olmayan MTK modelleri arasındaki en temel farklılık madde tepki fonksiyonlarına ilişkindir. Parametrik modellerde her bir madde tepki fonksiyonunun belirli bir matematiksel fonksiyona sahip olduğu varsayılırken, parametrik olmayan modellerde maddeler belirli bir matematiksel formda olmaya zorlanmaz. Bu bölüme alanda en yaygın kullanılan parametrik lojistik MTK modelleri açıklanmıştır.

Lojistik modeller arasındaki temel farklılıklar modeldeki parametre sayısından kaynaklanır. İkili puanlanan veri yapısında alan yazında en yaygın olarak kullanılan tek boyutlu modeller: bir parametrelili model (1PL), iki parametrelili model (2PL) ve üç parametrelili modeldir (3PL). Model seçimi yapılırken veri yapısını en iyi açıklayan modelin değerlendirilmesi gerekir (Hambelton, Swaminathan ve Rogers, 1991). En genel model olarak 3PL modelin denklemi Denlem (1)'de verilmiştir.

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad i=1, \dots, n \quad (1)$$

Denklemden n madde sayısı,  $P_i(\theta)$  rastgele seçilen bir bireyin i maddesini doğru cevaplama olasılığı,  $b_i$  i maddesinin güçlüğü,  $a_i$  i maddesinin ayırcılığı ve  $c_i$  i maddesine ilişkin şans parametresidir. D ise lojistik fonksiyondan elde edilen değerlerin normal ogive modele yakınlaştırmak için kullanılan ölçekleme katsayısıdır. D=1.7 olduğunda lojistik ve normal ogive modelden elde edilen yetenek kestirimi değerleri arasındaki fark 0.01'den daha düşük olmaktadır.

2PL modelde düşük yetenek düzeyindeki bireylerin soruyu doğru cevaplama olasılıkları sıfır olarak kabul edilir. Dolayısıyla 2PL denkleminde şans parametresi yer almaz. 1PL modelde ise bu varsayım ek olarak maddelerin ayırcılıkları eşit kabul edilir ve denklemden a parametresi bir sabit olarak yer alır (Hambelton, Swaminathan ve Rogers, 1991).

**Kestirim Yöntemleri.** Veriye uygun MTK modeli seçildikten sonra madde ve yetenek parametrelerinin kestirilmesi gerekir. Genellikle başlangıçta birey ve madde parametreleri bilinmez. Gözlenebilir olan değerler yalnızca bireylerin maddelere verdikleri cevaplardır. Bu cevaplar kullanılarak bireylere ve maddelere yönelik parametreler istatistiksel yöntemlerle kestirilirler (Hambleton ve Swaminathan,1991). Uygulamada parametre kestirimlerinde kullanılan farklı yöntemler vardır. Bu yöntemlerden bazıları aşağıda ayrıntılı olarak açıklanmıştır.

**Maksimum Olabilirlik (ML).** Maksimum olabilirlik yöntemleri alanda en yaygın kullanılan kestirim yöntemlerinden biridir. Bunun en temel nedenlerinden birisi varsayımları sağlandığında oldukça kullanışlı olmasıdır. Ayrıca madde sayısı arttıkça ve örneklem büyüdükçe daha tutarlı sonuçlar verir ve etkili kestiricilerdir. Maksimum olabilirlik fonksiyonu aşağıdaki gibi tanımlanabilir.

İkili puanlanan verilerde  $\theta$  yetenek düzeyindeki bir bireyin  $i$  maddesine yönelik cevap örüntüsü  $U_i = \{1 = \text{doğru cevap}, 0 = \text{yanlış cevap}\}$  ise bireyin cevap olasılığı denklem (2)' deki gibidir.

$$P(U_i|\theta) = P(U_i = 1|\theta)^{U_i}P(U_i = 0|\theta)^{1-U_i} \quad (2)$$

$$P(U_i|\theta) = P^{U_i}Q^{1-U_i}$$

$U_1, U_2, \dots, U_n$  rastgele değişkenleri  $u_1, u_2, \dots, u_n$  gözlenen değerlerini aldığı anda ise  $\theta$  yetenek düzeyine yönelik olabilirlik fonksiyonu denklem (3)'teki gibi tanımlanır.

$$L(u_1, u_2, \dots, u_n|\theta) = \prod_{i=1}^n P_i^{u_i} Q_i^{1-u_i} \quad (3)$$

Söz konusu olabilirlik fonksiyonunun en yüksek değeri  $\theta$  yetenek düzeyinin en çok olabilirlik kestirimini verir. Ancak bu kestirimin doğrudan elde edilmesi zordur. Bu nedenle yaygın olarak Newton-Raphson yöntemi ile olabilirlik denklemi çözülür (Hambleton ve Swaminathan,1991).

**Birlikte Maksimum Olabilirlik (JML).** Madde ve birey parametrelerinin ikisinin de bilinmediği durumlarda kullanılan kestirim yöntemlerinden birisi JML yöntemidir. Bu yaklaşımda madde ve yetenek parametrelerini eş zamanlı kestirmek için hem madde hem birey için birlikte olabilirlik fonksiyonu tanımlanır ve bu fonksiyonun en yüksek değeri elde edilir. Birlikte en çok olabilirlik fonksiyonu  $L$  olabilirlik

fonksiyonunun K birey için çarpımından elde edilir. Birlikte en çok olabilirlik fonksiyonu denklem (4)'teki gibi ifade edilebilir.

$$L = \prod_{i=1}^K \prod_{j=1}^n P_j(\theta_i)^{u_{ij}} Q_j(\theta_i)^{1-u_{ij}} \quad (4)$$

Birlikte en çok olabilirlik fonksiyonun maksimum değerini aldığı  $\theta$  ve madde parametre değerleri parametrelerin kestirim değeri olarak alınır. Kestirimlerin yapılması birkaç aşamada gerçekleştirilir. İlk olarak bireylerin geçici yetenek parametrelerinin belirlenmesi için madde parametreleri kestirilir. Sonrasında kestirilen geçici birey parametreleri bilinen parametreler olarak ele alınarak madde parametreleri tekrar kestirilir. Sonrasında madde parametreleri bilinen parametreler olarak ele alınır ve birey parametreleri kestirilir. Bu adımlar daha iyi madde ve birey parametreleri elde edilinceye kadar tekrar edilir (Ayala,2009).

*Marjinal Maksimum Olabilirlik (MML)*. JML yönteminde birey ve madde parametreleri eş zamanlı olarak kestirilir. Örneklem büyüdüğünde madde parametrelerinin daha tutarlı kestirilmesi beklenir. Ancak örneklem büyüdükçe kestirilecek olan birey parametre sayısı da arttığında JML yönteminde örneklem büyüdükçe parametre kestirimleri tutarlı olmayabilir. Ayrıca herhangi bir grup madde modele uyum sağlamadığında, maddelerin testten çıkarılması durumunda kestirimlerin yeniden yapılması gerekir (Ayala,2009).

MML kestirimi JML'den farklı olarak birey ve madde parametre kestirimlerini ayırır. Madde parametre kestirimlerinde yeterli düzeyde model uyumu sağlandıktan sonra birey yetenek parametreleri kestirilir (Ayala,2009). Dolayısıyla JML yöntemindeki sınırlılıklara bir alternatif olarak görülebilir. MML yönteminde madde parametrelerinin kestirimi, olabilirlik fonksiyonunun yetenek dağılımına göre integrali alınarak elde edilir. Bireyler ise rastgele bir örneklem olarak ele alınır.

Bir bireyin U cevap vektörüne sahip olma marjinal olasılığı denklem (5)'teki gibi ifade edilebilir.

$$p(U|a, b, c) = \int_{-\infty}^{\infty} \prod_{i=1}^n P_i^{U_i} Q_i^{1-U_i} g(\theta) d\theta \equiv \pi_u \quad (5)$$

Denklemden  $\pi_u$  u cevap örüntüsü için marjinal olasılıktır. İkili puanlanan n madde için cevap örüntüsü sayısı  $2^n$  olur. Dolayısıyla u cevap örüntüsünü veren  $r_u$  birey için olabilirlik fonksiyonu denklem (6)'daki gibidir (Hambleton ve Swaminathan,1991).

$$L \propto \prod_{u=1}^{2^n} \pi_u^{r_u} \quad (6)$$

Dolayısıyla denklemler yetenek parametresine ( $\theta$ ) bağlı değildir. Yani yetenek parametresi bilinmeyen bir bireyin cevap örüntüsüne dayalı olarak olası  $\theta$  değerleri belirlenebilir. Denklemlerde birey parametresi yer almadığından örneklem büyüklüğünün artması marjinal olabilirlik fonksiyonu ile kestirilecek parametre sayısını artırmaz (Ayala,2009).

Denklemden yer alan integral aslında fonksiyonun altında kalan alana karşılık gelir. Yani, bir bireyin  $u$  cevap vektörüne sahip olması olasılığı, fonksiyonun altında kalan alana eşittir. İntegral alım süreci matematiksel olarak daha yoğun ve zor bir süreç olduğundan kestirimlerde genellikle Hermite–Gauss kuadrat nokta yaklaşımı kullanılır. Bu yaklaşımda, bir seri dikdörtgenden oluşan kesikli bir dağılım ile olabilirlik fonksiyonuna yaklaşılar. Her bir dikdörtgenin orta noktası kuadrat nokta olarak adlandırılır. Dikdörtgenler daraldıkça, diğer bir ifade ile kuadrat noktaların sayısı arttıkça daha iyi sonuçlar elde edilir (Ayala,2009).

*Bayesçi Kestirim yöntemleri.* Birey ve madde parametreleri eş zamanlı kestirildiğinde tam veya sıfır puan alan bireyler için ML kestirimlerinde ortaya çıkabilecek kestirim problemlerine Bayes yaklaşımları bir çözüm getirebilir (Hambleton ve Swaminathan,1991).

Bayesçi yaklaşım bilinmeyen parametreleri sabit değil rastgele değişken olarak ele alır. Parametrik bayesçi yaklaşımda bu parametreler parametrik bir dağılım ailesinden gelir (Duncan ve MacEachern, 2013). Bayesçi kestirim yöntemleri temelde Bayes teoremine dayalıdır. Bayes teoremine göre koşullu marjinal olasılık dağılımları denklem (7)'deki gibi ifade edilir.

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \quad (7)$$

Yetenek kestirimi bağlamında ise  $A$  yetenek dağılımı  $\theta$  ve  $B$  gözlenen  $u$  cevap dağılımı olarak ele alındığında denklem (8)'deki gibi ifade edilebilir.

$$P(\theta|u) = \frac{P(u|\theta)P(\theta)}{P(u)} \quad (8)$$

$\theta$  sürekli bir değişken olduğundan denklemdeki değişkenler yoğunluk fonksiyonları olarak ifade edilebilir.  $P(u)$  değeri de sabit olduğundan denklem (9)'daki gibi yazılabilir.

$$f(\theta|u) \propto L(u|\theta)f(\theta) \quad (9)$$

Denklemden yer alan  $f(\theta|u)$  sonsal dağılıma,  $L(u|\theta)$  olabilirlik fonksiyonuna ve  $f(\theta)$  önsel dağılıma karşılık gelir (Hambleton ve Swaminathan,1991). Önsel dağılım herhangi bir gözlem yapmadan bireylerin yetenek dağılımına ilişkin elimizde olan bilgidir. Bireylere test uygulandıktan sonra ise elde edilen dağılım olabilirlik fonksiyonunu verir. Önsel dağılım ile gözlenen verinin birleştirilmesi ile de sonsal dağılım elde edilir (Ayala,2009). Önsel dağılım olarak normal, beta veya tek biçimli (uniform) dağılım gibi farklı dağılım türleri kullanılabilir. Bilgi vermeyen ya da diğer bir ifade ile tek biçimli bir önsel dağılım belirlendiğinde kestirimler ML ile benzer sonuçlar verir (Hambleton ve Swaminathan,1991).

Bayesci 3PLM modeli aşağıdaki gibi tanımlanmaktadır.

$$X_{ij} | \theta_i, a_j, b_j \sim \text{Bernoulli} \left( c_j + \frac{1 - c_j}{1 + e^{-a_j(\theta_i - b_j)}} \right)$$

$$\theta_i | \mu_\theta, \sigma_\theta^2 \sim \text{Normal} (\mu_\theta, \sigma_\theta^2)$$

$$b_j | \mu_b, \sigma_b^2 \sim \text{Normal} (\mu_b, \sigma_b^2)$$

$$a_i | \nu_a, w_a \sim \text{Gamma} (\nu_a, w_a)$$

$$c_j | \alpha_c, \beta_c \sim \text{Beta}(\alpha_c, \beta_c)$$

2PL Bayesci model için ise şans parametresi göz ardı edilir (Duncan ve MacEachern, 2013). Bayes teoremine dayalı parametrik iki temel kestirim yöntemi vardır. Bunlar beklenen sonsal (EAP) ve maksimum sonsal (MAP) kestirim yöntemleridir. Bu yöntemler arasındaki temel fark sonsal dağılımdan bireylerin yetenek parametrelerinin elde edilme şeklidir. Yetenek parametreleri sonsal dağılımın ortalaması veya medyanı kullanılarak elde edilebilir. Simetrik bir dağılımda bu iki değer aynıdır. Uygun önsel kullanıldığında kestirici olarak ortalamanın kullanılması kestirimin hata kareler ortalamasını düşürür. Beklenen sonsal yöntemi kestirici olarak ortalamayı kullanırken maksimum sonsal yöntemi ortancayı kullanır (Ayala,2009).

MAP yöntemi MLE gibi iteratif bir yöntemdir. EAP ise MML benzer şekilde kuadrat noktalara dayalı iteratif olmayan bir yöntemdir. Dolayısıyla EAP kestirimleri MAP ve MLE'ye göre daha hızlıdır. Ayrıca MAP sürekli önsel kullanırken EAP kesikli önsel kullanır (Ayala,2009). Bunlara ek olarak birey parametre kestirimlerinde



genellikle EAP yöntemi MAP'a göre daha düşük hata kareler ortalaması üretir (Bock ve Mislevy, 1982). Diğer yandan örtük özelliğin dağılımı çarpıklaştıkça MAP yönteminin EAP' tan daha az hatalı kestirimler yaptığı raporlanmıştır (Sass, Schmitt ve Walker, 2005).

**Parametrik olmayan madde tepki modelleri.** PoMTK sıralı birey ve madde ölçümleri için oluşturulan bir grup madde tepki modelinden oluşur. PoMTK' yı parametrik yapmayan en temel özelliği, monotonluk varsayımı olmakla beraber lojistik ya da normal ogive gibi parametrik bir fonksiyon şartı yoktur (Sijtsma,2005). Yani PoMTK modellerinde madde karakteristik eğrileri herhangi bir monoton formda olabilir. Bu bağlamda PoMTK modelleri daha geneldir ve daha az varsayıma dayandığı için daha çok sayıda madde modele uyum sağlayabilir (Sijtsma, Emons, Bouwmeester, Nylicek ve Roorda, 2008).

PoMTK parametrik MTK kuramına göre daha geniş bir teorik çerçeveye sahiptir ve parametrik kuramın aksine doğrulayıcı değil açıklayıcı bir yaklaşımdır (Sijtsma ve Meijer, 2007). Aslında bilinen tüm parametrik MTK modelleri PoMTK modellerinin özel bir şeklidir (Sijtsma ve diğerleri, 2008). PoMTK modelleri parametrik modellere göre daha esnektir ve parametrik model varsayımları tamamen sağlanmadığında da kullanılabilir. Standart parametrik MTK modelleri uyum sağlamadığında, PoMTK yaklaşımlarını iki temel özelliği ile yeni açıklayıcı bir bakış açısı sunar. Bu özellikler madde tepki fonksiyonlarının parametrik olmayan regresyonla kestirilmesi ve yerel bağımsızlığın sağlanması ya da kısmen sağlanmasına karşın maddeler arasında koşullu kovaryans kestirimleri yapmasıdır (Junker ve Sijtsma, 2001).

PoMTK yaklaşımının üç temel varsayımı vardır. Bunlar boyutluluk, yerel bağımsızlık ve monoton artanlıktır.  $X_1, X_2, \dots, X_j$  ikili puanlanan madde tepki değişkenleri olsun. Buradan PoMTK'nın temel varsayımı:

1. Yerel bağımsızlık: theta gizil değişkeni ve j madde cevapları arasında denklem (10)'daki koşullu olasılık vardır

$$P(X_1 = x_1, \dots, X_j = x_j | \theta) = \prod_{j=1}^J P(X_j = 1 | \theta)^{x_j} [1 - P(X_j = 1 | \theta)]^{1-x_j} \quad (10)$$

2. Monotonluk: Madde tepki fonksiyonları  $P_j(\theta) = P(X_j = 1 | \theta)$  tetanın ya da tetanın boyutlarının azalmayan bir fonksiyonudur.

### 3. Tek boyutluluk

Bu üç varsayımı sağlayan PoMTK modelleri monoton homojen olarak adlandırılır. Her ne kadar varsayımlar tamamen elimine edilemese de zayıflatılabilirler (Junker ve Sijtsma, 2001).

Tek boyutluluk varsayımı tek bir baskın boyut ve onunla ilişkin boyutlar şeklinde zayıflatılabilir. Zayıf monotonluk ise madde karakteristik eğrilerinin ortalama monotonluğu sağlamasıdır. Yerel bağımsızlık ise temel yerel bağımsızlık (essential local independence) olarak zayıflatılabilir. Temel yerel bağımsızlık maddeler arasında bir miktar pozitif veya negatif ilişkiye izin verirken her bir yetenek düzeyinde toplamda madde çiftleri arasındaki ilişkinin sıfır olmasını gerektirir (Sijtsma,2005).

Gizil özellik normal dağılmadığında kestirimlerdeki yanlılıkları elimine etmek için parametrik olmayan yöntemler önerilmiştir. Bunlardan alanda yaygın olarak kullanılan ampirik histogram (AH) ve Davidian eğrileri madde tepki modeli (DE)'dir. Son dönemlerde Rash model için uygun olan parametrik olmayan Bayesçi kestirim (NBE) yöntemi de önerilmektedir (Finch ve Edwards, 2016). Araştırma kapsamında PoMTK modellerinden Davidian eğrileri, Ampirik Histogram ve parametrik olmayan bayesci kestirim yöntemleri kullanılmıştır.

**Ampirik Histogram (AH).** MLE' ye dayalı MTK madde parametreleri kestirimleri genellikle Both ve Aitkins'in (1981) beklenti maksimizasyonu (expectation maximization) şemasını kullanılır. Bu şemada bireylerin yetenekleri kuadrat noktalar ile tanımlanır, her bir kuadrat noktada yer alan bireylerin beklenen oranı tanımlanır ve her bir kuadrat nokta için olabilirlik değerleri maksimize edilir. Standart MLE kestirimlerinde bireylerin kuadrat noktalardaki beklenen frekansları normal dağılımdan elde edilir. Ampirik histogram yönteminde ise MLE yönteminden farklı olarak her bir maksimizasyon aşamasında madde parametrelerine ek olarak her bir madde için beklenen frekanslar kullanılarak olabilirlik fonksiyonu maksimize edilir. Kestirimlerde yetenek parametresinin dağılımı olarak normal dağılım yerine her bir aşamada elde edilen ampirik histogramlar kullanılır (Woods,2007). Her bir kuadrat nokta ( $X_k$ ) da yetenek dağılımının yoğunluğu denklem 11'deki gibi kestirilir.

$$g(X_k) \cong \frac{\sum_i^s r_i L_i(X_k) A(X_k)}{\sum_h^q \sum_i^s r_i L_i(X_h) A(X_h)} \quad (11)$$

**Davidian Eğrileri (DE).** AH yöntemine benzer şekilde DE kestirimlerinde de beklenti maksimizasyonu şeması kullanılır. AH yönteminden farklı olarak her bir aşamada beklenen frekans dağılımları Zhang ve Davidian (2001) tarafından açıklanan Davidian eğrileri kullanılarak elde edilir.

$P_k(\theta)$  k dereceli bir polinom ve  $\varphi$  standart normal yoğunluk olmak üzere MTK çerçevesinde DE yoğunluğu  $h$  denklem 12'deki gibi tanımlanır.

$$h(\theta) = P_k^2(\theta)\varphi(\theta) \quad (12)$$

$\gamma$  sıfırdan büyük bir tam sayı,  $m_k \neq 0$  olmak üzere  $m = (m_0, m_1, \dots, m_k)$  polinom denkleminin katsayıları olmak üzere  $P_k^2(\theta)$  polinomu denklem (13)'teki gibidir.

$$P_k^2(\theta) = \left\{ \sum_{\gamma=0}^k m_{\gamma} \theta^{\gamma} \right\}^2 \quad (13)$$

Örneğin  $k=2$  için  $P_k^2(\theta) = (m_0 + m_1\theta + m_2\theta^2)^2$  şeklinde tanımlanır. DE parametreleri en çok olabilirlik yöntemi ile kestirilir. Kuadrat noktalar üzerinden en çok olabilirlik fonksiyonunun logaritması denklem (14)'teki gibidir.

$$l_g = \sum_{q=1}^Q N(\theta_q) \log [P_k^2(\theta_q)\varphi(\theta_q)] \quad (14)$$

Polinomun derecesi eğrinin şeklini belirler. Birinci ve ikinci dereceden polinomlar çarpık dağılımların elde edilmesine olanak sağlar. Ancak, MTK parametre kestirimlerinde kullanılan diğer yöntemlerde olduğu gibi DE yönteminde de yetenek parametresinin ölçeğinin tanımlanması gerekir ve genellikle ölçeği ortalaması 0 ve standart sapması 1 olacak şekilde ölçeklenir. Bu ölçekleme  $k=0$  ve  $k=1$  değerleri için normal eğri oluşturur. Bu nedenle DE kestirimlerinde daha yüksek dereceler kullanılır (Woods ve Lin, 2009).

**Parametrik Olmayan Bayeşçi Kestirim.** Bayeşçi parametrik olmayan kestirimler monotonluğu katı bir şekilde sağlarken, MKE'lerin lojistik formun dışında olmalarına olanak sağlar. Bu durum yeteneğin normal dağılımını parametrik olmayan bir dağılım ile değiştirerek veya MKE'ler için bir model ailesi kullanarak sağlanabilir. Söz konusu iki parametrik olmayan etki Dirichlet yöntemi (DP) kullanılarak elde edilebilir (Duncan ve MacEachern, 2008). Dolayısıyla DP önselleri MKE'lerin lojistik olmadığı ancak monotonluğun sağlandığı durumlarda kestirimler yapmaya olanak sağlar (Duncan ve MacEachern, 2013).

DP önsel olasılık dağılımı olarak Ferguson (1973) tarafından öne sürülmüştür. DP kendisi olasılık dağılımı olan bir rastgele değişkene ortalama olasılık dağılımı bir önsel atar. Temel ölçüm olarak adlandırılan bu önsel ortalama araştırmacının dağılımın merkezi olduğuna inandığı yerdir. Önsel ortalamanın varyansı ise önsel kütle (mass) olarak adlandırılır. DP aşağıdaki gibi tanımlanır.

$$Y_i \sim F, i = 1, \dots, n$$

$$F \sim DP(M, F_0)$$

F rastgele olasılık dağılımı, x gözlenen veri, F<sub>0</sub> önsel ortalama ve M önsel küttedir. M Kernel düzleştirme yöntemindeki bant genişliği (bandwith) parametresi gibi davranır ve yüksek değerleri daha düz eğriler oluşturur. (Duncan ve MacEachern, 2013). Yani her bir  $\theta$  gizil özelliği için üretilen F olasılık dağılımının birkaç DP dağılımı kümesinden geldiği varsayılır. Bu dağılım merkezi eğilimi M'nin önsel dağılımını içerir. Örneğin araştırmacı  $\theta$  için tek bir normal dağılım belirlemek yerine DP kullanılarak pek çok normal dağılımın bir karışımını kullanabilir. Bu normal dağılım karışımı tek bir normal dağılıma göre parametreleri daha iyi tanımlayabilir. Ayrıca DP yalnızca normal dağılımla sınırlı değildir (Finch ve Edwards, 2016).

Araştırma kapsamında Duncan ve MacEachern'in (2008) önerdiği DP yöntemine dayalı olarak parametrik olmayan bayesci iki parametrelili model (NBE) temel alınmıştır. Bu yaklaşıma göre araştırma kapsamında kullanılan NBE model aşağıdaki gibi tanımlanabilir:

$$\theta_i | \mu_\theta, \sigma_\theta^2 \sim Normal(\mu_\theta, \sigma_\theta^2)$$

$$F_j | a_j, b_j, M_j \sim DP(M_j, F_{0j})$$

$$z_{ij} | F_j \sim F_j$$

$$Y_{ij} | \theta_i, z_{ij} = \begin{cases} 1 & z_{ij} \leq \theta_{ij} \\ 0 & z_{ij} > \theta_{ij} \end{cases}$$

$$F_{0j} = \frac{1}{1 + e^{-a_j(\theta - b_j)}}$$

$$b_j | \mu_b, \sigma_b^2 \sim Normal(\mu_b, \sigma_b^2)$$

$$a_i | v_a, w_a \sim Gamma(v_a, w_a)$$

Bir MKE'sini DP yöntemi ile modelleyebilmek için MKE birikimli frekans fonksiyonu (bff) olarak ele alınır ve dolayısıyla monoton azalmamaya zorlanır. 2PL

bir MKE bir bff oluştururken 3PL ile oluşturmaz. Bir diğer durum ise veriden dağılımı bff oluşturacak bir MKE gözleyemeyiz. Bu nedenle model yeniden parametrelendirilir ve doğrudan F<sub>j</sub> yoğunluğu ile çalışılmasına olanak sağlayan bir zij değişkeni oluşturulur. Bu değişken i bireyinin j maddesine toleransı olarak tanımlanır ve eğer bireyin yeteneği söz konusu maddeye toleransından büyükse kişi maddeyi doğru cevaplar, eğer bireyin yeteneği toleransından küçükse yanlış cevaplar (Duncan ve MacEachern, 2013).

NBE model kestirimi süreci başlangıç değerlerinin belirlenmesi, parametrelerin güncellenmesi ve kestirimleri yapılması olmak üzere üç aşamada özetlenebilir. Parametre kestirim sürecinde Duncan ve MacEachern (2008) tarafından önerilen algoritmalar temel alınmış ve aşağıda açıklanmıştır.

(1) Başlangıç değerlerinin belirlenmesi: İlk olarak  $\theta$ , a ve b parametreleri ile z değerinin başlangıç değerleri belirlenir.  $\theta$  başlangıç değeri önseller veya ham puanlar kullanılarak; a ve b parametrelerinin başlangıç değerleri önsel dağılımlar ya da sonsal ortalamalar kullanılarak; z başlangıç değeri ise tanımlanan  $\theta$ , a ve b başlangıç değerleri kullanılarak ikinci aşamada yer alan eşitlik yardımı ile belirlenebilir.

(2) Parametre değerlerinin güncellenmesi: Bu aşamada  $\theta$ , a, b ve z parametreleri güncellenir. Bu aşama Gibbs örnekleme sürecinin Burn-in iterasyon aşamasıdır.

$\theta_i$  yetenekleri denklem (15)'teki koşullu dağılım ile tekrar güncellenir. Bu dağılım en büyük ve en küçük z değerleri ile sınırlandırılmış normal dağılımdır.

$$\theta_i | \theta_{-i}, z, b, a, Y \sim N(\mu_\theta, \sigma_\theta^2) I \left[ \max_{\{j: Y_{ij}=1\}} z_{ij}, \min_{\{j: Y_{ij}=0\}} z_{ij} \right] \quad (15)$$

Her bir j maddesi için zij değerleri üç aşamada tekrar kestirilir. İlk olarak her bir i bireyi için z<sub>j</sub> vektöründen zij çıkarılarak zij\* denklem 16'daki dağılım ile elde edilir.

$$z_{ij}^* | \theta, z_{(-i)j}, b, a, Y \sim \begin{cases} \frac{Y_{ij} I(-\infty, \theta_i] F_{0j}}{F_{0j}(\theta_i)} + \frac{(1-Y_{ij}) I(\theta_i, \infty) F_{0j}}{1-F_{0j}(\theta_i)} & wp \ p_{ij}, \\ < z_{ij} > & wp \ p_{lj} \end{cases} \quad (16)$$

Denklemden  $p_{ij} \propto M_j \left[ Y_{ij} F_{0j}(\theta_i) + (1 - Y_{ij}) (1 - F_{0j}(\theta_i)) \right]$  ve  $p_{ij} \propto Y_{ij} I(z_{ij} \leq \theta_i) + (1 - Y_{ij}) I(z_{ij} > \theta_i)$  ' dir.  $\langle z_{ij} \rangle$  ise  $z_{ij}$  noktasındaki kütle ile ayrık dağılıma karşılık gelmektedir. Bu değerler parçalı log-konkav fonksiyonların adaptif ret örnekleme (Adaptive Rejection Sampling) yöntemi ile elde edilebilir (MacEachern,1998).

Sonrasında  $z_j$  vektöründe  $z_{ij}$  ile  $z_{ij}^*$  yer değiştirir. Ortak değeri  $z_{ij}^*$  olan bireylerden oluşan her bir  $k$  kümesi  $\gamma_{kj}$  yer parametresi ile yeni bir küme oluşturur.  $\gamma_{kj}$  ise denklem (17)'de yer alan dağılım ile elde edilir.

$$\gamma_{kj} | \theta, s, b, a, Y \sim h(\gamma_{kj} | \theta, s, b, a, Y) = l(\gamma_{kj} | a_j, b_j) I\left( \max_{\{i: S_{ij}=k, Y_{ij}=0\}} \theta_i, \min_{\{i: S_{ij}=k, Y_{ij}=1\}} \theta_i \right) \quad (17)$$

Denklemden  $l(\gamma | a, b) = a e^{-a(\gamma-b)} / (1 + e^{-a(\gamma-b)})^2$  a ve b parametrelerinin lojistik yoğunluk fonksiyonuna karşılık gelmektedir.

a ve b parametrelerinin değerleri ise denklem (18) ve (19)'da yer alan koşullu olasılık dağılımlarına dayalı olarak Gibbs örnekleme içinde adaptif ret örnekleme yöntemi kullanılarak elde edilebilir.

$$b_j | \theta, z, s, b_{-j}, a, Y \propto \varphi\left(\frac{b_j - \mu_b}{\sigma_b}\right) \prod_k h(\gamma_{kj} | \theta, s, b, a, Y) \quad (18)$$

$$a_j | \theta, z, s, b, a_{-j}, Y \propto g(a_j | v, w) \prod_k h(\gamma_{kj} | \theta, s, b, a, Y) \quad (19)$$

Denklemden  $g(a_j | v, w)$  şekil parametresi v ve ölçek parametresi w olan gama yoğunluğunu temsil eder.

(3) Sonsal kestirimlerin yapılması: Bireyinin sonsal yetenek ortalaması kullanılarak aşağıdaki denklem (20) ile kestirilir.

$$\hat{E}(\theta_i | Y) = \frac{1}{U} \sum_{t=B+1}^{B+U} E(\theta_i | \theta_{-i}^{(t)}, z^{(t)}, b^{(t)}, a^{(t)}, Y) \quad (20)$$

Denklemden B Gibbs örnekleme burn-in iterasyon sayısı ve U burn-in sonrası iterasyon sayısına karşılık gelmektedir.

**Markov Zinciri Monte Carlo Yöntemleri.** Markov Chain Monte Carlo (MCMC) herhangi bir teorik çok değişkenli dağılımdan rastgele örneklem üretme yöntemidir. MCMC yaklaşımının temelinde  $M_0, M_1, \dots, M_k = (\theta^k, \beta^k)$  olmak üzere Markov zinciri oluşturulması ve bu zincirlerden gözlemlerin üretilmesi yatmaktadır (Patz ve Junker,1999). Söz konusu gözlemler  $\theta$  ve  $\beta$  parametrelerine ilişkin

çıkarımların yapılmasında kullanılır. Bayesçi yaklaşımda Markov zinciri sonsal dağılıma dönüşecek şekilde tanımlanır. Markov zincirinin davranışını adımlar arasında değişimi sağlayan olasılıklar belirler. Zincirin ilk K adımı Burn-in olarak tanımlanır ve bu adımlar çıkarılır. K+1,..., K+L adımları kullanılarak parametre kestirimleri yapılır. Araştırma kapsamında parametrik olmayan bayesçi kestirim sürecinde MCMC yöntemlerinden Gibbs örnekleme yöntemi kullanılmıştır.

**Gibbs Örnekleme Yöntemi.** Gibbs örnekleme yöntemi karmaşık dağılımlardan bağımsız örneklemler seçmek için kullanılan MCMC yöntemlerinden biridir. Gibbs örnekleme yöntemiyle elde edilen bu örneklemler Bayesçi çıkarımlar yapmak için doğrudan kullanılabilirler. Örnekleme süreci iteratif olarak gerçekleştirilir. Her bir iterasyonda parametreler koşullu dağılımından yeni bir değer örnekleyerek güncellenir (Gilks, Best ve Tan, 1995). MTK çerçevesinde Gibbs örnekleme yaklaşımı klasik MTK kestirimlerine benzer şekilde davranır. Yani bir grup parametreye ilişkin örnekleme yaparken diğer parametrelerin sabit ve bilindiğini varsayar (Patz ve Junker, 1999). Araştırma kapsamında parametre kestirim sürecinde kullanılan iterasyonlara ilişkin ayrıntılı bilgi yöntem bölümünde verilmiştir.

### **İlgili Araştırmalar**

Bu kısımda alan yazında yapılan çalışmalar üç başlık altında özetlenmiştir. İlk olarak parametrik olmayan MTK yöntemlerinin kullanıldığı çalışmalardan bahsedilmiştir. Sonrasında örtük özellik dağılımının manipüle edildiği araştırmalar ayrıca özetlenmiştir. Son bölümde ise ilgili araştırmalara ilişkin bir özet sunulmuştur.

**Parametrik olmayan MTK.** Douglas ve Cohen (2001) araştırmalarında parametrik madde tepki fonksiyonlarına (MTF) uyum sağlamayan maddelerin tespit edilmesinde parametrik olmayan yöntemlerin kullanımını incelemişlerdir. Araştırma temel olarak grafiksel karşılaştırmaya dayalı yürütülmüştür. Maddeler bir kısmı 2PLM'e uyum sağlarken bir kısmı sağlamayacak şekilde üretilmiştir. Araştırmanın sonuçları madde sayısı 25'den fazla olduğunda parametrik olmayan MTF ve p kestirimlerinin daha kesin olduğunu göstermiştir.

Vermunt (2001) parametrik ve parametrik olmayan MTK modellerinin tanımlanması ve test edilmesinde sınırlandırılmış örtük sınıf modelini kullanmıştır. Stresten ağlama ölçeğinin 4 maddesi üzerinde yaptığı uygulamada sınırlandırılmış

ve sınırlandırılmamış örtük sınıf modellerini de incelemiştir. Araştırmanın bulgularına göre sınırlandırılmamış model sınırlandırılmış modelden daha iyi uyum göstermiştir. Parametrik modeller bu veri için fazla sınırlayıcı olmuş ve modellerin hiçbiri veriye uyum sağlamamıştır. Diğer yandan parametrik olmayan modellerde veri uyumu sağlanmıştır.

Liang (2007) çalışmasında yarı parametrik bir MTK modeli olan L-MP ile 2PL ML, JML, Kernel düzleştirilmesi ve NBE kestirim yöntemlerini karşılaştırmıştır. Araştırmada modeller arasındaki karşılaştırma RIMSE değerleri temel alınarak yapılmıştır ve simülasyon verisi kullanılmıştır. Araştırmanın bulguları L-MP modelinin birey ve madde parametrelerini oldukça iyi kestirdiğini göstermiştir. Standart lojistik modeller uyum sağladığında, L-MP madde karakteristik eğrilerini MML'ye çok yakın kestirirken, JML'den daha iyi kestirmiştir. Yetenek parametrelerini ise L-MP yöntemi MML ve JML'den çok daha iyi kestirmiştir. Ayrıca gerçek modeller standart lojistik modele uyum sağlamadığında L-MP modeli 2PL yerine tercih edilebilir. Parametrik olmayan kestirimlerle karşılaştırıldığında ise Kernel düzleştirilmesi MKE'leri L-MP den daha iyi kestirirken, yetenek parametrelerini daha fazla hatayla kestirmiştir. NBE kestirimleri ise L-MP kestirimleri ile oldukça benzer sonuçlar vermiştir.

Sijtsma ve diğerleri (2008) çalışmalarında sağlığa ilişkin hayat kalitesi ölçeğini geliştirme ve değerlendirmede parametrik olmayan monoton homojenlik modelinin kullanılabilirliğini araştırmışlar ve sonuçları yaygın olarak kullanılan aşamalı tepki modeli ile karşılaştırmışlardır. Araştırmanın sonuçlarına göre monoton homojenlik modeli analizlerinde ölçek her bir içerik alanı için tek boyutlu bulunmuştur. Ölçeklenebilirlik katsayıları bazı maddelerin diğerlerine göre daha sınırlı ölçeklenebileceğini göstermiştir. Aşamalı tepki modeli analizlerinde ise maddelerin bir kısmı modele uyum göstermemiştir.

Duncan ve MacEachern (2008) istatistik dersine ilişkin sınav sonuçlarını nanparametrik ve parametrik iki parametrelili bayesci modelleri karşılaştırmışlardır. Araştırmanın bulguları parametrik olmayan modelin birkaç maddeden parametrik modelden manidar şekilde farklı MKE oluşturduğunu göstermiştir. Ayrıca bazı bireyler için yetenek kestirimleri de parametrik ve parametrik olmayan modeller arasında farklılaşmaktadır.



Kođar (2015) alıřmasında basit ve iki boyutlu yapılar da oklu puanlanan maddeler iin madde ve model veri uyumlarını farklı kořullarda karřılařtırmıřtır. Arařtırmada kořul olarak rneklem byklđ, test uzunluđu ve boyutlar arsındaki korelasyon ele alınmıř ve 36 simlasyon kořulunda tek boyutlu PoMTK'ya dayalı Mokken homojenlik modeli (MHM), tek boyutlu MTK'ya dayalı 2PL model ve ok boyutlu MTK'ya dayalı tamamlayıcı modelden elde edilen kestirimler karřılařtırılmıřtır. Veriler tek ve ok deđiřkenli normallik varsayımı ile retilmiřtir. Arařtırmanın bulguları tm modellerde rneklem byklđ ve test uzunluđu arttıka model veri uyumlarının arttıđını gstermiřtir. Kısa testlerde ise ok boyutlu model daha iyi uyum gstermiřtir. Tm kořullarda PoMTK ile elde edilen model veri uyumu olduka dřk hata ile hesaplanmıřtır. ok boyutlu MTK ile elde edilen a parametrelerindeki hata miktarları ise tek boyutlu modelle olduka benzer bulunmuřtur.

řengl Avřar ve Tavřancıl (2016) oklu puanlanan maddelerin psikometrik zelliklerini farklı kořullarda PoMTK modellerinde karřılařtırmıřtır. Arařtırmada kořul olarak rneklem byklđ, madde sayısı, rneklem dađılımı ele alınmıř ve 27 simlasyon kořulunda Mokken homojenlik (MHM) modeli ve Kernel dzleřtirme yaklařımı (KSAM) model kestirimleri incelenmiřtir. MHM kestirimlerinin sonuları madde sayısı, rneklem byklđ ve rneklem dađılımının maddelerin model uyumunu etkilediđini gstermiřtir. KSAM model kestirimi sonuları ise madde sayısı, rneklem byklđ ve rneklem dađılımının madde ve test ayıcılıđını etkilediđini gstermiřtir. Arařtırmanın bulguları MHM ve KSAM model analizlerinin farklılařtıđını gstermiřtir. Ancak madde sayısı ve rneklem byklđ arttıka iki model daha benzer sonular vermiřtir. Ayrıca KSAM modeli MHM modeline gre gvenirliđi daha dřk kestirmiřtir.

Bıkmaz Bilgen ve Dođan (2017) ok kategorili maddeler iin parametrik yntem olarak ařamalı tepki modeli (ATM) ve parametrik olmayan monoton homojenlik modeli (MHM) kestirimlerini rneklem byklđ, test uzunluđu, madde yanıt kategori sayısı ve rneklem dađılımına gre toplam 192 simlasyon kořulunda karřılařtırmıřtır. Arařtırmanın bulguları ATM kestirimlerinde model veri uyumu hesaplamalarının deđiřken deđerlerinden etkilendiđini gstermiřtir. Ayrıca gvenirlik deđerleri iki model arasında benzer kestirilmiřtir. Kk rneklemlerde ise MHM modeli ATM'ye gre daha az hatalı kestirimler retmiřtir. Sonu olarak MHM

modelinde model veri uyumunun daha pratik hesaplanması ve yorumlanmasından dolayı ATM'ne iyi bir alternatif olabileceği vurgulanmaktadır.

Reyhanlıođlu Keeođlu (2018) alıřmasında parametrik madde tepki kuramı modeli ile parametrik olmayan monoton homojenlik modellerini farklı rneklem byklklerinde ve tek boyutlu ve iki boyutlu testlerde madde parametrelerinin deđiřmezliđi bađlamında karřılařtırmıřtır. Arařtırmanın sonucunda test iki boyutlu olduđunda hem parametrik hem de parametrik olmayan MTK modellerinde madde parametre deđiřmezliđinin sađlanamadıđı raporlanmıřtır. Diđer yandan testlerden elde edilen gvenirlik katsayılarının hem parametrik hem de parametrik olmayan modellerde evren deđerine benzer olduđu raporlanmıřtır.

Mor Dirlik (2019) arařtırmasında PMTK ve PoMTK modellerinden elde edilen madde ve yetenek kestirimlerini test uzunluđu, rneklem byklđ ve maddelerin psikometrik niteliklerine gre karřılařtırmıřtır. Arařtırmada TIMSS 2011 matematik testine iliřkin bařarı sıralamasında ilk 20'de yer alan lkelerden elde edilen veriler kullanılmıřtır. Maddeler ise glk ve ayırt edicilik deđerlerine gre yksek ve dřk olmak zere drt grupta incelenmiřtir. Arařtırmanın sonuları farklı rneklem byklklerinde kestirilen model parametrelerinin iki model arasında uyumlu olduđunu gstermiřtir. Her iki modelde de farklı test uzunluklarında ve rneklem byklđnde elde edilen kestirimler arasında yksek dzeyde iliřki vardır. PoMTK'dan elde edilen kestirimler ise madde sayısı 25 ve rneklem byklđ 3000 olduđunda PMTK modeli ile tutarlı sonular vermiřtir. Arařtırmanın sonularına gre PMTK'nın varsayımları sađlanmadıđında PoMTK modellerinin kullanılabileceđi nerilmektedir.

Arenson ve Karabatsos (2018) parametrik modellerin katı monotonluk varsayımını sađlamadıđı durumlarda kullanılabilecek Bayesi parametrik olmayan bir model geliřtirmiřlerdir. Karakteristik eđri kestirimlerinde karıřık beta dađılımlarının kullanıldıđı modelde sonsal kestirimler Metropolis-Hasting algoritmasına dayalı olarak elde edilmiřtir. Modelin gerek veride uygulanabilirliđi ise TIMSS 2015 uygulaması matematik testi verilerinde gsterilmiřtir.

Hu, Ma ve Peak'de (2020) Rash modelde bireylerin ve maddelerin eř zamanlı sınıflandırılmasında kullanılabilecek parametrik olmayan Bayesi bir yntem nermiřtir. Sonlu karıřımlar karıřımı (mixture of finite mixtures-MFM) dađılımlara

dayalı modelin performansı hem simülasyon verileri hem de gerçek veriler ile değerlendirilmiş ve modelin tutarlı kestirim ve sınıflama yaptığı raporlanmıştır.

Franco ve Wiberg (2020) çalışmasında parametrik olmayan üç modeli: Davidian eğrileri, Kernel düzleştirme ve Bayesçi yarı parametrik Rash modeli karşılaştırmıştır. Araştırma kapsamında bu üç yöntemin gerçek puanların elde edilmesindeki performansları ve gerçek veri ile uyumu karşılaştırılmıştır. Araştırmanın sonuçlarına göre DE yönteminin diğer yöntemlerden daha zayıf performans göstermiştir. Kernel düzleştirme ve Bayesçi yarı parametrik Rash model ise evren değerlerine benzer sonuçlar üretmiştir.

**Örtük özellik dağılımının parametre kestirimleri üzerine etkisi.** Seong (1990) çalışmasında farklı yetenek dağılımlarında, örneklem büyüklüğünde ve kuadrat sayısında farklı önsel dağılımlarla yapılan ML kestirimlerinin doğruluğunu incelemiştir. Normal, pozitif çarpık ve negatif çarpık olmak üzere 3 farklı theta dağılıma göre 45 maddelik 35 set üretilmiştir ve kestirimler BILOG programı kullanılarak yapılmıştır. Araştırmanın bulguları yetenek dağılımları ve önsel dağılımlar benzer olduğunda ve örneklem büyük olduğunda kestirimlerin daha doğru olduğunu göstermiştir. Yalnızca büyük örneklem büyüklüğünde ve önsel dağılım ile yetenek dağılımı benzer olduğunda kuadrat noktaların artırılması kestirimleri iyileştirmiştir.

Stone (1992) çalışmasında MML madde parametre kestirimi ve ML yetenek parametre kestirimini farklı örneklem büyüklükleri, test uzunlukları ve theta dağılımlarında karşılaştırmıştır. Araştırmada simülasyon verisi kullanılmıştır. Araştırmanın sonuçları MML madde güçlük parametresi kestirimlerinin küçük örneklem (250), kısa test (10) ve farklı yetenek dağılımlarında (çarpık ve basık) kesin ve tutarlı sonuçlar verdiğini göstermektedir. Ancak dağılım çarpıklaştıkça örneklem büyüklüğünden bağımsız olarak 10 ve 20 maddelik testlerden elde edilen yanlılığın arttığı bulunmuştur. Gerçek çarpık ve basık theta dağılımları madde ayırıcıkları üzerinde ters etki yapmıştır.

Kirischi ve diğerleri (2001) araştırmalarında MULTLOG, XCALIBRATE ve BILOG programları ile yapılan kestirimleri, yeteneğin boyutluluğu, boyutlar arası ilişkilerin büyüklüğü ve dağılımı bağlamlarında karşılaştırmışlardır. Araştırmada yetenek dağılımı olarak normal, pozitif çarpık ve sivri dağılımlar incelenmiştir. Araştırma kapsamında test uzunluğu olarak 40 madde ve örneklem büyüklüğü

olarak 1000 kişi kullanılmıştır. Araştırmanın bulguları yetenek dağılımının parametre kestirimlerine önemli bir etkisi olmadığını göstermiştir. Bu durum örneklemin yeterince büyük olması ve soru sayısının yeterince fazla olması ile açıklanmıştır.

Sass, Schimitt ve Walker (2005) çalışmalarında ML, MAP ve EAP kestirimlerinin gizil değişken kestirim hatalarını normal ve normal olmayan koşullarda karşılaştırmış ve madde güçlüğü dağılımının söz konusu hataları etkileyip etkilemediğini incelemiştir. Araştırmada ayrıca dağılımın uç kısmındaki bireyler için kestirim hataları da incelenmiştir. Simülasyon verisi kullanılan araştırmanın bulguları normal dağılmayan örtük değişkende kestirim hatasının normal dağılıma sahip değişkenin kestirimine göre çok fazla olduğunu göstermiştir. Ayrıca MAP uç değer olmayan katılımcılar için en iyi kestirimi üretirken, kestirim yönteminden bağımsız olarak uç değerlerdeki bireylerin yetenek kestirimlerindeki hatalar oldukça yüksektir.

Woods ve Lin (2009) araştırmalarında DE, Ramsay-curve IRT (RC-IRT) ve ampirik histogram yöntemini (AH) normal ve çarpık gizil dağılımlarda karşılaştırmıştır. Araştırmanın bulguları gizil değişkenin dağılımı normal olmadığında üç yöntemin de normal modellere göre daha iyi kestirimler yaptığını göstermiştir. DE, RC-IRT yöntemleri AH yöntemine göre daha tutarlı sonuçlar vermiştir.

Finch ve Edwards (2015) çalışmalarında gizil değişkenin ve güçlük parametrelerinin normal dağılmadığı durumlarda normal dağılıma dayalı MTK kestirimleri (MML, EAP, MAP) ile RC-IRT ve parametrik olmayan bayesçi kestirim yöntemlerini karşılaştırmıştır. Simülasyon verisi kullandıkları çalışmalarında koşul olarak farklı madde güçlük dağılımı, örtük özellik dağılımı, madde sayısı, örneklem büyüklüğü ve kestirim yöntemleri ele alınmıştır. Araştırmanın sonuçlarına göre gizil değişken ve madde güçlükleri normal dağılmadığında en doğru kestirim sonuçları parametrik olmayan bayesçi kestirim yöntemi ile elde edilmiştir.

Smits ve diğerleri (2020) çoklu puanlanan testlerde örtük özelliğin normal ve çarpık dağıldığı durumlarda aşamalı tepki modeli (ATM) ile normallik varsayımı olmayan Davidian eğrilerine dayalı aşamalı tepki modeli (DE-ATM) ve sıfır değer aralıklı MTK modellerini farklı madde sayılarında karşılaştırmıştır. Araştırmanın sonuçları çarpık dağılımlarda en iyi kestirimlerini sıfır değer aralıklı MTK modelinin

verdiğini diğer yandan ATM'nin DE-ATM'den daha az hatalı kestirimler yaptığını raporlamıştır.

**İlgili arařtırmalar özeti.** Alan yazındaki arařtırmalar incelendiğinde parametrik olmayan farklı modellerin önerildiđi ve bu modellerin farklı kořullar altında karřılařtırıldıđı gözlenmektedir. Özellikle son yıllarda yapılan çalıřmalarda farklı amaçlara yönelik parametrik olmayan Bayesçi modellerin ortaya atıldıđı ve ön plana çıktıđı gözlenmektedir.

Ortaya atılan modellerın uygulanabilirliđi gerçek veriler ile ortaya konurken, Model karřılařtırmaları simülasyon verileri ile yapılmaktadır. Simülasyon çalıřmalarında modellerin performansları yaygın olarak örneklem büyüklüğü, test uzunluđu ve örtük özellik dađılımı kořullarında karřılařtırmaktadır. Son yıllarda ise b parametrelerinin dađılımının da parametre kestirimleri üzerine etkisi incelenmiştir. Çalıřmaların bulguları farklılık göstermekle birlikte genel olarak parametrik modellerin varsayımları sağlanmadıđında kestirim hatalarının arttıđı raporlanmıştır. Farklı kořullarda kurgulanan simülasyon çalıřmaları parametrik olmayan modellerin alternatif olarak kullanılabilirliđine iliřkin kanıtlar sunmaktadır.

Alan yazında uygulamada da yaygın olarak kullanılan MLE ve MAP yönteminin diđer yöntemlerle sıklıkla karřılařtırıldıđı gözlenmektedir. Örtük özellik çarpık olduđunda önerilen parametrik olmayan yöntemler ise parametrik olmayan bayesçi kestirim yöntemi (NBE), Davidian Curve madde tepki modeli (DE), Ampirik histogram (AH) ve MCMC yöntemleridir. Finch ve Edwards (2016) örtük özellik dađılımı çarpıklařtıđında Rasch model için NBE kestirimlerinin parametrik yöntemlerden daha az hatalı kestirimler yaparken, MCMC yönteminin parametrik ve parametrik olmayan yöntemlerden daha yüksek hatalar ürettiđini raporlamıştır. Bu bağlamda bu arařtırmaya MLE, MAP, DE, AH ve NBE kestirim yöntemleri tercih edilmiş ve karřılařtırılabilirliđi sağlamak amacıyla Finch ve Edwards (2016)'ın çalıřmasındaki kořullara benzer kořullar kullanılmıştır.

## Bölüm 3

### Yöntem

#### Araştırma Yöntemi

Araştırmanın amacı parametrik ve parametrik olmayan MTK model kestirimlerinin farklı simülasyon koşullarında ve gerçek veri kümelerinde karşılaştırılmasıdır. Bu bağlamda araştırma kuramları belirli koşullar altında karşılaştırmayı amaçlayan bir yöntemsel karşılaştırma niteliğindedir.

#### Araştırmanın Koşulları

MTK uygulamalarında parametrik model parametre kestirimlerinde en yaygın olarak kullanılan yöntem MLE yöntemidir. Bu nedenle pek çok araştırmada diğer yöntemlerle karşılaştırılmıştır. Bu araştırmada da parametrik kestirim yöntemi olarak araştırmaya dahil edilmiştir. Sass ve diğerleri (2005) çarpık örtük özellikte MLE, EAP ve MAP kestirimlerinden en yansız sonuçları MAP yönteminin verdiğini raporlamıştır. Bu nedenle MAP yöntemi parametrik bayesçi yaklaşıma dayalı kestirim yöntemi olarak araştırmaya dahil edilmiştir. Örtük özellik çarpık olduğunda yaygın olarak kullanılan parametrik olmayan yöntemler ise parametrik olmayan bayesçi kestirim yöntemi (NBE), Davidian Curve madde tepki modeli (DE), Ampirik histogram (AH) ve MCMC yöntemleridir. Finch ve Edwards (2016) örtük özellik dağılımı çarpıklaştığında Rasch model için NBE kestirimlerinin parametrik yöntemlerden daha az hatalı kestirimler yaparken, MCMC yönteminin parametrik ve parametrik olmayan yöntemlerden daha yüksek hatalar ürettiğini raporlamıştır. Bu nedenle MCMC yöntemi araştırmaya dahil edilmemiş, parametrik olmayan yöntemler olarak DE, AH ve NBE kestirim yöntemleri tercih edilmiştir.

NBE yönteminde farklı önsel dağılım belirleme teknikleri vardır. Sass ve diğerleri (2011) normal dağılmayan gizil değişkenden kaynaklanan yanlılığı önlemek için Bayes kestirimlerinde önsellerin en az sınırlayıcı şekilde olması gerektiğini belirtmiştir. Alanda yaygın olarak ortalama 0, standart sapması 1 olacak şekilde normal önsel dağılım kullanılmaktadır (Fox, 2010). Ancak gizil değişkenin dağılımı normal olmadığında bu yaklaşımı kullanmak doğru değildir. Bu nedenle NBE yaklaşımında önsel dağılım belirlemede Dirichlet yöntemi (DP) ve Polya ağaçları (PT) önerilmektedir (San Martin, ve diğerleri 2011). Bu araştırmada önsel dağılım belirlenirken DP yöntemi kullanılmıştır.

MTK modellerinde örtük özelliğin ve madde güçlük parametresinin dağılımı parametre kestirimlerindeki hatayı etkilemektedir. Alan yazında parametrik ve parametrik olmayan yöntemler farklı örtük özellik dağılımlarında incelenmiştir. Sass ve diğerleri (2005) örtük özellik dağılımını normal, gama (10;1,5) ve çarpık; güçlük parametre dağılımını ise normal ve gama (1;1) olacak şekilde üretmişlerdir. Finch ve Edwards (2015) da verileri aynı değerlerle üretmiş ve örtük özellik dağılımını biraz daha çarpıklaştırarak sonuçları karşılaştırmışlardır. Bu araştırmada da madde güçlük, ayırt edicilik ve yetenek dağılımları önceki çalışmalarla karşılaştırılabilirliği sağlamak amacıyla Sass ve diğerlerinin (2005) çalışmasına benzer koşullarda üretilmiştir. Önceki çalışmalardan farklı olarak ise araştırmada iki parametrelili modele göre kestirimler yapılarak daha karmaşık MTK modellerine yönelik çıkarımlar yapılması amaçlanmaktadır.

Örneklem büyüklüğü ve madde sayısı MTK kestirimlerini etkileyen diğer önemli değişkenlerdir ve alan yazında sıklıkla çalışılmıştır. Dolayısıyla bu çalışmada da simülasyon koşulu olarak farklı örneklem büyüklükleri ve madde sayılarına göre karşılaştırmalar yapılmıştır. Seong, (1990) örtük özelliğin çarpıklık etkisini incelediği çalışmada madde sayısı 40 ve üzerinde ve örneklem büyüklüğü 1000 ve üzeri olduğunda parametrik yaklaşımların örtük özellikteki çarpıklığı tolere edebildiğini raporlamıştır. Parametrik olmayan yöntemlerin en önemli avantajlarından biri olarak küçük örneklerde daha doğru kestirimler verdiği belirtilmektedir (Junker ve Sjitsima, 2001). Bu çalışmada ise üç farklı örneklem büyüklüğü 100, 500 ve 1000 olarak ele alınmıştır. Alan yazında normal olmayan örtük özellik varlığında parametrik ve parametrik olmayan MTK kestirimlerinin karşılaştırıldığı çalışmalarda küçük örneklem olarak 250 ele alınmıştır (San Martin ve diğerleri, 2011; Finch ve Edwards, 2016). Bu çalışmada küçük örneklem büyüklüğü 100 alınarak pratikte karşılaşılabilecek daha küçük örneklem gruplarına yönelik bir çıkarım yapılması amaçlanmıştır. Araştırmada madde sayısı ise Finch ve Edwards (2016) çalışması ile karşılaştırılabilirliği sağlamak amacıyla benzer şekilde 10, 20 ve 30 olarak belirlenmiştir.

### **Verilerin Üretilmesi**

Araştırma kapsamında örneklem büyüklüğü, madde sayısı, örtük özellik dağılımı, madde güçlük dağılımı ve kestirim yöntemleri manipüle edilmiştir. Araştırmanın koşulları Tablo 1'de özetlenmiştir.

Tablo 1

*Araştırmanın Koşulları*

Koşullar	Düzeyleyler
Örneklem Büyüklüğü	100,500,1000
Madde Sayısı	10, 20, 30
Örtük Özellik Dağılımı	Normal, Çarpık: 1.6,-1.6
Madde Güçlük Dağılımı	Normal, Çarpık: 1.6, -1.6
Kestirim Yöntemi	MLE, MAP, DE, AH, NBE

Tablo 1’de görüldüğü gibi araştırmada toplam 81 koşulda madde cevap matrisi üretilmiştir. Üretilen veriler beş farklı kestirim yöntemi ile incelenmiştir. Her bir koşul için 100 tekrar yapılmıştır. Dolayısıyla sonuçlar toplam 405 koşulda incelenmiştir. Veriler R paket programı kullanılarak üretilmiştir (R Core Team, 2013).

**Birey parametrelerinin üretilmesi.** Araştırma kapsamında bireylerin örtük özellikleri normal, pozitif çarpık ve negatif çarpık olmak üzere üç farklı koşulda üretilmiştir. Normal yetenek dağılımı için standart normal dağılım temel alınmış ve R temel paketinde “rnorm” fonksiyonu ile üretilmiştir. Pozitif çarpık yetenek dağılımı için gamma dağılımı kullanılmış ve R temel paketinde yer alan “rgamma” fonksiyonu kullanılarak üretilmiştir. Söz konusu fonksiyonda şekil parametresi 1.2 ve ölçek parametresi 1.5’e eşitlenerek çarpıklığı 1.6’ya eşit olan yetenek dağılımı üretilmiş ve sonrasında ortalaması 0, olacak şekilde merkezlenerek yetenek parametreleri elde edilmiştir. Tüm örneklem büyüklüklerinde 100 tekrarda veriler üretilerek çarpıklıkları hesaplanmış ve 1.6’dan manidar düzeyde farklılaşıp farklılaşmadığı tek örneklem t testi ile incelenmiştir. 100 tekrar sonucunda elde edilen çarpıklık değerlerinin 1.6’dan manidar düzeyde farklılaşmadığı bulunmuştur. Negatif çarpık yetenek parametresi ise pozitif çarpık üretilen dağılım -1 ile çarpılarak elde edilmiştir.

**Madde parametrelerinin üretilmesi.** Araştırma kapsamında 2PL model kullanıldığı için ayırıcılık (a) ve güçlük (b) parametreleri üretilmiştir. Ayırıcılık parametresi ortalaması 0.3 ve standart sapması 0.2 olan log-normal dağılıma göre üretilmiştir (Bulut ve Sünbül, 2017). Güçlük parametreleri ise normal, pozitif çarpık ve negatif çarpık olmak üzere üç farklı şekilde üretilmiştir. Normal b parametreleri



“norm” fonksiyonu ile üretilmiştir. Pozitif çarpık b parametresi ise yetenek parametresine benzer şekilde gamma dağılımı ile üretilmiştir. Negatif çarpık b parametreleri pozitif çarpık üretilen parametreler -1 ile çarpılarak elde edilmiştir. Araştırma kapsamında kullanılan 10, 20 ve 30 maddelik testlerin madde parametreleri Tablo 2, 3 ve 4’de yer almaktadır.

*Tablo 2*

*10 Maddelik Test için Madde Parametreleri*

Madde	a	N-b	PÇ-b	NÇ-b	a	N-b	PÇ-b	NÇ-b
M1	1.252	-0.178	-0.064	0.064	M6	1.375	0.990	0.595
M2	2.098	1.770	-0.382	0.382	M7	1.698	0.006	-0.698
M3	1.801	-1.843	-0.016	0.016	M8	0.964	0.689	-0.824
M4	0.778	0.337	-0.643	0.643	M9	0.970	-0.528	-0.789
M5	1.645	-0.754	2.486	-2.486	M10	1.007	-0.530	0.335

\*N-b: normal dağılan b, PÇ-b: pozitif çarpık b, NÇ-b: negatif çarpık b

*Tablo 3*

*20 Maddelik Test için Madde Parametreleri*

Madde	a	N-b	PÇ-b	NÇ-b	Madde	A	N- b	PÇ-b	NÇ-b
M1	1.459	-1.137	-0.053	0.053	M11	1.199	0.600	0.121	-0.121
M2	1.106	0.075	0.253	-0.253	M12	1.194	1.290	0.898	-0.898
M3	0.811	-0.772	-0.070	0.070	M13	0.904	-0.345	-0.402	0.402
M4	1.240	-0.581	-1.158	1.158	M14	1.161	0.481	-0.778	0.778
M5	1.493	-0.098	0.099	-0.099	M15	1.561	0.240	1.226	-1.226
M5	1.231	0.832	0.202	-0.202	M16	1.237	-1.147	-1.104	1.104
M7	1.394	0.378	-0.559	0.559	M17	0.986	-0.037	-0.656	0.656
M8	1.378	1.264	1.193	-1.193	M18	1.343	-0.280	-0.932	0.932
M9	1.634	-2.213	-0.490	0.490	M19	1.372	2.155	-0.487	0.487
M10	1.107	-0.791	-0.394	0.394	M20	1.332	0.101	3.092	-3.092

\*N-b: normal dağılan b, PÇ-b: pozitif çarpık b, NÇ-b: negatif çarpık b

Tablo 4

30 Maddelik Test için Madde Parametreleri

M	a	N-b	PÇ-b	NÇ-b	M	A	N- b	PÇ-b	NÇ-b
M1	1.458	1.402	-0.237	0.237	M16	1.164	-0.391	-0.249	0.249
M2	1.695	-0.405	-0.782	0.782	M17	1.268	-0.998	0.142	-0.142
M3	1.358	0.648	0.096	-0.096	M18	1.535	0.487	-0.023	0.023
M4	1.014	0.298	1.234	-1.234	M19	1.301	-0.712	-0.720	0.720
M5	0.982	0.254	-0.523	0.523	M20	1.113	1.332	-0.558	0.558
M5	1.464	-0.419	0.120	-0.120	M21	1.914	0.852	-0.969	0.969
M7	0.984	-0.800	-0.414	0.414	M22	1.219	2.148	-1.080	1.080
M8	1.236	0.872	-0.573	0.573	M23	1.337	0.450	-1.047	1.047
M9	1.048	0.760	-0.633	0.633	M24	1.330	0.416	0.119	-0.119
M10	1.362	0.745	0.898	-0.898	M24	1.204	-0.480	-0.196	0.196
M11	1.622	-0.619	-1.078	1.078	M26	1.748	-1.204	2.016	-2.016
M12	1.847	1.745	-0.959	0.959	M27	1.277	-0.158	-0.337	0.337
M13	1.276	-1.515	0.778	-0.778	M28	1.372	0.062	0.467	-0.467
M14	1.069	-1.282	3.389	-3.389	M29	1.436	-0.050	1.296	-1.296
M15	1.465	-1.048	0.345	-0.345	M30	1.138	-2.287	-0.523	0.523

\*N-b: normal dağılan b, PÇ-b: pozitif çarpık b, NÇ-b: negatif çarpık b

**Cevap matrisinin üretilmesi.** Birey ve madde parametreleri üretildikten sonra bireylerin maddelere verdikleri cevaplara ilişkin 1-0 madde cevap matrisi 2PL modele göre üretilmiştir. Cevap matrisi üretilirken ilk olarak öncesinde tanımlanan a, b ve yetenek parametrelerine göre bireylerin her bir maddeye doğru cevap verme olasılıkları 2PL modele göre hesaplanmıştır. Sonrasında her bir madde için rastgele bir olasılık dağılımı oluşturulmuş ve bu olasılıklar bireylerin maddelere doğru cevap verme olasılıkları ile karşılaştırılmıştır. Bireylerin maddeye doğru cevap verme olasılığı tanımlanan rasgele olasılıktan büyük olduğunda birey maddeyi doğru cevaplayacak, küçükse yanlış cevaplayacaktır. Bu bağlamda her bir koşul için 100 tekrarda bireylerin madde cevap matrisleri oluşturulmuştur. Madde parametreleri ve cevap matrisinin üretilmesi için kullanılan R kodu Ek-A' de verilmiştir.

## Verilerin Analizi

Araştırma kapsamında hem simülasyon verileri hem de gerçek veriler kullanılmıştır. Verilerin analizi simülasyon çalışmasına ilişkin veri analizi süreçleri ve gerçek verilerin analizine yönelik süreçler olmak üzere iki başlıkta açıklanmıştır.

**Simülasyon Çalışması.** Araştırma kapsamında MLE, MAP, NBE, AH ve DE olmak beş farklı kestirim yöntemi karşılaştırılmıştır. MLE ve MAP kestirimleri için R programında yer alan “irtoys” paketi kullanılmıştır. irtoys paketinde madde ve birey parametrelerinin kestiriminde üç farklı program seçeneği yer almaktadır. Bunlar BILOG, icl ve ltm'dir. Paket kapsamında madde parametreleri kestirilirken her bir parametre için önsel dağılım belirlenebilir (Partchev, 2017). Bu çalışma kapsamında “icl” programı kullanılmıştır. MLE kestirimleri için a ve b parametresi için herhangi bir önsel dağılım tanımlanmamıştır. MAP kestirimlerinde ise programın varsayılan olarak belirlediği b parametresi için normal, yetenek parametresi için ise normal önsel dağılım kullanılmıştır. DE ve AH yöntemlerine dayalı kestirimlerinde ise R programında yer alan “mirt” paketi (Chalmers, 2020) kullanılmıştır. Kestirimlere yönelik hazırlanan R kodu Ek B'de verilmiştir.

NBE kestirimleri için kodlar ilk olarak R programında oluşturulmuş ve sonrasında analiz sürecinin hızlandırılması amacıyla MATLAB koduna dönüştürülmüştür. Kestirimlerine yönelik hazırlanan R kodları Ek C'de, MATLAB kodları ise Ek D'de yer almaktadır.

NBE kestirimleri başlangıç değerlerinin tanımlanması, parametre değerlerinin güncellenmesi ve parametre kestirimlerinin elde edilmesi olmak üzere üç aşamada gerçekleştirilmiş. N birey sayısı,  $\varphi$  normal yoğunluk olmak üzere yetenek parametrelerinin başlangıç değerleri ham puanlar kullanılarak aşağıdaki formül ile elde edilmiştir.

$$\theta_i^0 = \varphi^{-1}\left(\frac{\text{rank}(\sum_j Y_{ij})}{N + 1}\right)$$

N birey sayısı k madde sayısı olmak üzere güçlük parametrelerinin başlangıç değerleri doğru cevap oranları kullanılarak aşağıdaki formül ile elde edilmiştir.

$$b_j^0 = \varphi^{-1}\left(\frac{\text{rank}(\sum_i Y_{ij}/N)}{k + 1}\right)$$

Ayırıcılık parametrelerinin başlangıç değerleri ise önsel dağılımdan örneklenerek elde edilmiştir.

Parametreler güncellenirken DP sürecinde önsel külte M=5 olarak alınmıştır. Parametre kestiriminde burn-in aşamasında 5000 güncelleme yapılmış sonrasında 10000 iterasyon yapılmıştır. İnceltme aralığı olarak 10 kullanılmıştır.

**Kestirimlerin karşılaştırılması:** Alan yazında model karşılaştırmalarında farklı kriterler kullanılmaktadır. Simülasyon çalışmalarında yaygın olarak hata kareler ortalamasının karekökü (RMSE), ortalama mutlak hata (MAE) ve hata kareler ortalaması (MSE) (Finch ve Edwards, 2015; Sass, Schmitt ve Walker, 2005; Kirisci ve diğerleri, 2001) değerleri kullanılmaktadır. Bu araştırmada da model karşılaştırmalarında RMSE ve MAE kullanılmıştır. 100 tekrar için RMSE ve MAE değerlerinin ortancası raporlanmıştır. Madde ve birey parametreleri için her bir tekrarda RMSE değerleri aşağıdaki formül aracılığı ile elde edilmiştir

$$RMSE_i = \sqrt{\frac{1}{n} \sum_{j=1}^n (\widehat{v}_{ij} - v_j)^2}$$

Denklemden n tekrar sayısı;  $\widehat{v}_{ij}$  i. tekrar için j maddesinin kestirilen parametresi veya j bireyinin kestirilen parametresi;  $v_j$  ise gerçek madde veya birey parametresine karşılık gelmektedir. Madde ve birey parametreleri için her bir tekrarda MAE değerleri aşağıdaki formül aracılığıyla elde edilmiştir.

$$MAE_i = \frac{1}{n} \sum_{j=1}^n |\widehat{v}_{ij} - v_j|$$

Benzer şekilde denklemden yer alan n tekrar sayısı;  $\widehat{v}_{ij}$  i. tekrar için j maddesinin kestirilen parametresi veya j bireyinin kestirilen parametresi;  $v_j$  ise gerçek madde veya birey parametresine karşılık gelmektedir.

RMSE ve MAE değerleri farklı yöntemler ile elde edilen parametre kestirimleri gerçek parametre değerleri ile aynı ölçeğe getirildikten sonra hesaplanmıştır. Ölçekleme işlemleri R programında "irtoys" paketinde yer alan "sca" fonksiyonu ile gerçekleştirilmiştir. Ölçekleme ve hata değerlerinin hesaplanmasına ilişkin R kodu Ek E'de verilmiştir.

Kestirim hatalarının özetlenmesinde kullanılan grafikler R programında yer alan “lattice” paketi (Sarkar, 2020) kullanılarak oluşturulmuştur. Grafiklerin oluşturulmasında kullanılan örnek bir R kodu Ek F’de verilmiştir.

Ayrıca hangi koşulların hatalar üzerinde anlamlı etkisinin olduğunun belirlenmesi için varyans analizi (ANOVA) yapılmıştır. ANOVA analizleri SPSS 23 paket programında gerçekleştirilmiştir.

**Gerçek verilerin analizi.** Araştırma kapsamında bir özel okulda 2018 yılında uygulanmış olan 8. Sınıf düzeyi bir deneme sınavının verileri kullanılmıştır. Sınavdaki Matematik ve Türkçe testlerine ait veriler kullanılmıştır. Veri seti toplam 4848 öğrenciden oluşmaktadır. Matematik sorularının hiç birine cevap vermemiş olan beş öğrenci ve Türkçe sorularının hiçbirini yanıtlamamış dört öğrenci veri setinden çıkarılmıştır. Öğrencilerin Türkçe ve Matematik testlerinden aldıkları ham puanlara ilişkin betimsel istatistikler Tablo 5’te yer almaktadır.

*Tablo 5*

*Türkçe ve Matematik testleri ham puanları betimsel istatistikleri*

	N	Min	Max	Ortanca	Ortalama	Standart Sapma
Türkçe	4844	0	20	15	14.46	3.57
Matematik	4843	0	20	12	11.8	5.38

Tablo 5’de görüldüğü gibi Türkçe testine ilişkin puan ortalaması 14.46 ve standart sapması 3.57’dir. Matematik puanlarının ise ortalaması 11.8 standart sapması 5.53’tür. Bu sonuçlara göre matematik testi puanlarının dağılımının daha heterojen olduğu anlaşılmaktadır.

Gerçek veri ile analizler gerçekleştirilmeden önce verilerin MTK varsayımlarını sağlayıp sağlamadığı incelenmiştir. İlk olarak testlerin boyutluluğu faktör analizi ile incelenmiştir. Veri ikili puanlandığı için faktör analizinde tekrarorik korelasyon matrisi kullanılmıştır. Matematik ve Türkçe testine ilişkin elde edilen faktör yükleri, özdeğerler ve açıklanan varyans oranı Tablo 6’da yer almaktadır.

Tablo 6

*Türkçe ve Matematik testlerine ilişkin faktör analizi sonuçları*

Maddeler	Matematik Testi			Türkçe Testi		
	Faktör Yüğü	Özdeğerler	Açıklanan varyans	Faktör Yüğü	Özdeğerler	Açıklanan varyans
M1	.76	10.045	.48	<b>.06</b>	8.730	.41
M2	.53	0.997		.69	1.623	
M3	.68	0.811		<b>.20</b>	1.040	
M4	.72	0.740		.61	0.928	
M5	.77	0.699		.85	0.917	
M6	.75	0.644		.75	0.778	
M7	.73	0.580		.62	0.737	
M8	.58	0.576		.74	0.671	
M9	.75	0.555		.74	0.606	
M10	.77	0.505		.54	0.579	
M11	.64	0.492		.73	0.494	
M12	.54	0.464		<b>.16</b>	0.463	
M13	.65	0.433		.82	0.416	
M14	.64	0.417		.92	0.396	
M15	.65	0.383		.46	0.383	
M16	.66	0.376		.67	0.353	
M17	.72	0.353		.77	0.352	
M18	.71	0.336		.30	0.308	
M19	.73	0.299		.75	0.198	
M20	.75	0.285		.60	0.018	

Tablo değerleri incelendiğinde Matematik testinde yer alan maddelerin tamamı birinci faktöre .50 üzerinde faktör yükü vermiştir. Özdeğerler incelendiğinde yalnızca birinci faktörün özdeğerinin 1'den büyük olduğu gözlenmektedir. Dolayısıyla matematik testi baskın bir tek boyuttan oluşmaktadır ve bu boyutun açıkladığı varyans oranı % 48'dir. Türkçe testine ilişkin faktör analizi sonuçları

incelendiğinde ise 1, 3 ve 12. Maddeler dışında tüm maddelerin birinci faktöre ilişkin faktör yükü .30 ve üzerindedir. Özdeğerler incelendiğinde birinci, ikinci ve üçüncü faktöre ilişkin özdeğerlerin 1'den büyük olduğu gözlenmektedir. Ancak özdeğer oranlarına bakıldığında birinci faktöre ilişkin özdeğerin ikinci faktörün 5 katından fazla olduğu gözlenmektedir. Dolayısıyla testin baskın bir tek boyutu ölçtüğü savunulabilir. Ayrıca paralel analiz sonuçları da her iki verinin de tek boyutlu olduğunu göstermiştir. Söz konusu boyutta açıklanan varyans oranı ise % 41'dir. Çalışmanın ilerleyen aşamalarında Türkçe testinde yer alan ve birinci faktöre yeterince yük vermeyen 1, 3, ve 12. Maddeler testten çıkarılmış ve analizlere 17 madde üzerinden devam edilmiştir.

Gerçek veriler ile analizler yapılırken örneklem büyüklüğünün etkisine incelemek adına analizler tüm kestirim yöntemleri için tüm veri seti ve veriden rastgele seçilen 100, 500 ve 1000 kişilik örneklerle gerçekleştirilmiştir. Gerçek veriler ile yapılan kestirimlerin karşılaştırılmasında model uyumları temel alınmıştır. Model uyum indeksi olarak alanda yaygın olarak kullanılan madde uyum istatistikleri ve modelin genel uyumuna ilişkin BIC değerleri hesaplanmıştır.

## Bölüm 4

### Bulgular

Araştırma kapsamında 81 koşulda üretilen veriler ile MLE, MAP, NBE, DE ve AH kestirimlerinin kestirim hatalarının karşılaştırılması amaçlanmıştır. Ayrıca araştırma kapsamında kestirim yöntemleri gerçek veri seti üzerinden de karşılaştırılmıştır. Bulgular her bir alt probleme yönelik beş başlık altında raporlanmıştır. Her bir alt problem içerisinde ilk olarak hatalar üzerinde etkisi olan değişkenlerin belirlenmesine ilişkin ANOVA sonuçları raporlanmıştır. ANOVA sonuçları etki büyüklükleri ile birlikte değerlendirilmiştir. Etki büyüklükleri 0-0.09 düşük; 0.10-0.20 orta ve 0.20'nin üstü yüksek etki olarak ele alınmıştır. Sonrasında ise hataların koşullar arasında değişimine ilişkin grafikler raporlanmıştır.

#### **Birinci alt probleme ilişkin bulgular**

Araştırmanın birinci alt problemi "Normal ve normal olmayan örtük özellik dağılımlarında, farklı örneklem büyüklükleri ve test uzunluklarında parametrik MML, MAP ve parametrik olmayan NBE, DE ve AH yaklaşımlarının hangisi daha yansız yetenek ve madde parametre kestirimleri üretir?" olarak belirlenmiştir. Bu kısımda sırasıyla yetenek a ve b parametre kestirimlerine ilişkin bulgular raporlanmıştır. Yetenek parametre kestirimlerine ilişkin hesaplanan RMSE ve MAE medyan değerleri ve standart sapmaları Ek G'de yer almaktadır. Madde parametre kestirimlerine ilişkin elde edilen RMSE ve MAE değerleri medyan ve standart sapmaları Ek-H'de verilmiştir.

**Yetenek parametresine ilişkin bulgular.** Bu kısımda kestirimlerin karşılaştırılması için kullanılan RMSE ve MAE değerlerine ilişkin analizler ve grafikler ayrı ayrı raporlanmıştır. Yetenek kestirimlerine ilişkin elde edilen RMSE değerleri üzerinde anlamlı etkisi olan koşulların belirlenmesi için yapılan ANOVA sonucu Tablo 7'de raporlanmıştır.

Tablo 7 incelendiğinde, yetenek kestirimlerinden elde edilen RMSE değerleri üzerinde anlamlı etkiye sahip değişkenlerden etki büyüklüğü en fazla olan değişken kestirim yöntemidir ve büyük etkiye sahiptir ( $p < 0.05$ ,  $\eta = 0.97$ ). Bu değişkeni madde sırasıyla sayısı değişkeni ( $p < 0.05$ ,  $\eta = 0.62$ ) ve örneklem büyüklüğü değişkeni ( $p < 0.05$ ,  $\eta = 0.38$ ) takip etmektedir. Yetenek dağılımı ise RMSE değerleri orta düzeyde bir etkiye sahiptir ( $p < 0.05$ ,  $\eta = 0.12$ ). Etkileşim etkileri incelendiğinde ise



RMSE deęerleri üzerinde en büyük etkiye sahip olan etkileşim kestirim yöntemi ve madde sayısı etkileşimidir ( $p < 0.05$ ,  $\eta = 0.87$ ). Kestirim yöntemi ile yetenek dağılımı etkileşimi de anlamlıdır ve yüksek düzeyde etkiye sahiptir ( $p < 0.05$ ,  $\eta = 0.29$ ). Denklemden yer alan diğer etkileşim etkileri ise istatistiksel olarak anlamlı değildir.

*Tablo 7*

*Yetenek parametreleri RMSE deęerlerine ilişkin ANOVA tablosu*

Kaynak	Kareler Toplamı	Sd	Ortalama Kareler	F	p	Etki B.
Düzeltilmiş Model	27.613 <sup>a</sup>	134	.206	85.915	.000	.977
Kesişim	147.883	1	147.883	61657.418	.000	.996
Yöntem	21.632	4	5.408	2254.737	.000	.971
Yetenek Dağılımı	.088	2	.044	18.354	.000	.120
Madde sayısı	1.040	2	.520	216.737	.000	.616
Örneklem büyüklüğü	.025	2	.013	5.312	.005	.038
Yöntem * Yetenek Dağılımı	.262	8	.033	13.646	.000	.288
Yöntem * Madde	4.324	8	.541	225.372	.000	.870
Yöntem * Örneklem	.021	8	.003	1.093	.368	.031
Yetenek Dağılımı * Madde	.010	4	.002	1.010	.403	.015
Yetenek Dağılımı * Örneklem	.020	4	.005	2.039	.089	.029
Madde * Örneklem	.018	4	.005	1.891	.112	.027
Yöntem * Yetenek Dağılımı * Madde	.035	16	.002	.918	.549	.052
Yöntem * Yetenek Dağılımı * Örneklem	.027	16	.002	.693	.800	.039
Yöntem * Madde * Örneklem	.025	16	.002	.658	.834	.038
Yetenek Dağılımı * Madde * Örneklem	.018	8	.002	.934	.489	.027
Yöntem * Yetenek Dağılımı * Madde * Örneklem	.068	32	.002	.889	.643	.095
Hata	.648	270				
Toplam	176.144	405				
Düzeltilmiş Toplam	28.260	404				

\* $R^2 = 0.981$  (Düzeltilmiş  $R^2 = 0.972$ )

Kestirim yöntemleri arasındaki farklılıkların incelenmesi için yapılan Post-Hoc testlerinin sonuçları Tablo 8’de verilmiştir.

*Tablo 8*

*Kestirim Yöntemleri Arasında Post-Hoc analizi sonuçları*

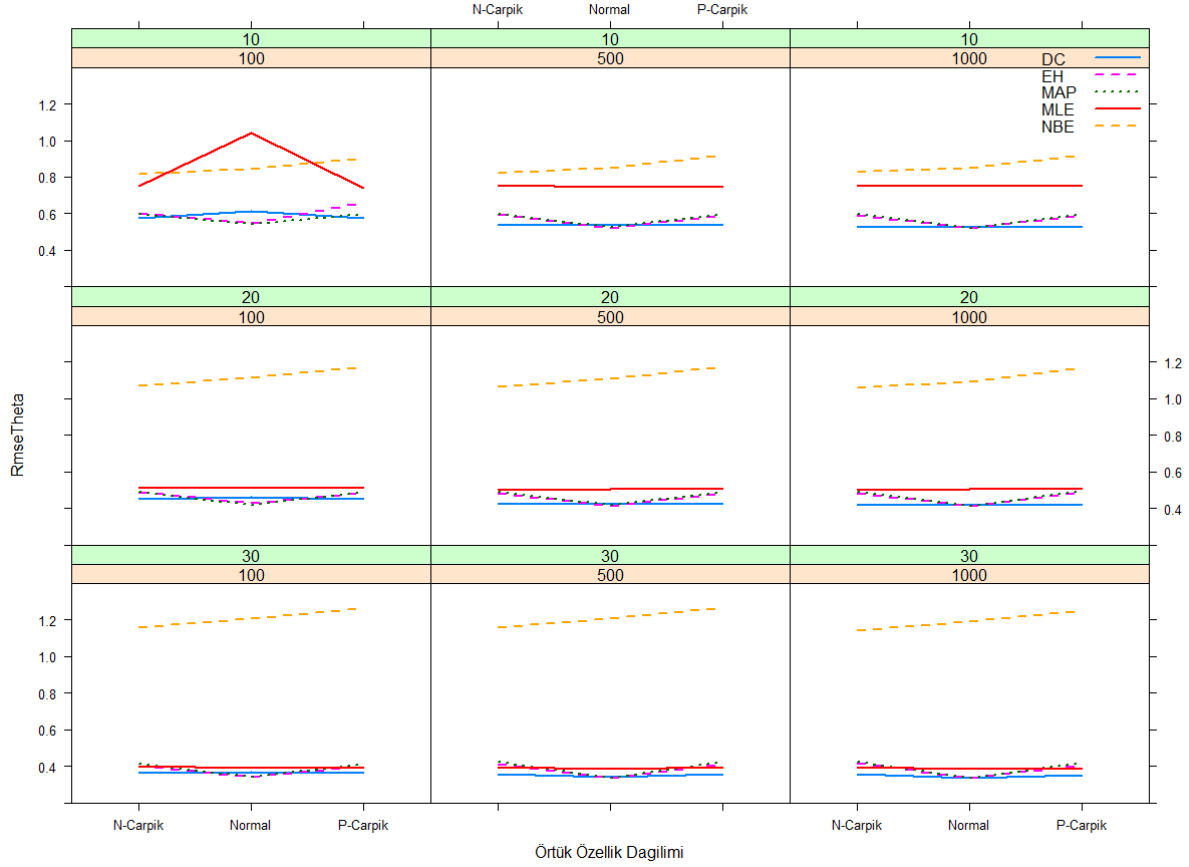
I	J	Yöntem	Ortalama Farklar (I-J)	P
DE		AH	-.029*	.006
		MAP	-.030*	.004
		MLE	-.114*	.000
		NBE	-.988*	.000
MAP		AH	.001	.924
		MLE	-.084*	.000
		NBE	-.958*	.000
NBE	MLE		.874*	.000
	AH		.959*	.000
AH	MLE		-.085*	.000

\* $p < .05$  düzeyinde anlamlı fark

Tablo 8 incelendiğinde DE yöntemi tüm yöntemlerden manidar düzeyde daha düşük hatalı kestirimler yaparken NBE yöntemi tüm yöntemlerden manidar düzeyde daha yüksek hatalı kestirimler yapmıştır. AH yöntemi ile MAP yöntemlerinin kestirim hataları arasında anlamlı bir farklılık görülmezken, MLE yöntemi DE, AH ve MAP yöntemlerinden daha yüksek hatalı kestirimler yapmıştır.

Yetenek parametrelerinin dağılımlarına ilişkin Post-Hoc testleri incelenmiştir. Yetenek parametrelerinin dağılımları arasındaki farklar incelendiğinde ise, yetenek dağılımı normal olduğunda çarpık dağılımlardan anlamlı düzeyde daha az hatalı kestirimler elde edilmiştir. Yetenek dağılımının pozitif çarpık olduğu durumlarda elde edilen hatalar negatif çarpık olduğu durumdan anlamlı düzeyde daha yüksektir ( $p < 0.05$ ).

Yetenek parametresine ilişkin RMSE değerlerinin yetenek parametresinin dağılımı, örneklem büyüklüğü ve madde sayısına göre değişimi Şekil 1’de verilmiştir.



Şekil 1. Yetenek parametresine ilişkin RMSE değerlerinin yetenek parametresinin dağılımı, madde sayısı ve örneklem büyüklüğüne göre değişimi

Şekil 1 incelendiğinde DE kestirimlerinin yetenek parametresinin dağılımdan etkilenmediği ve dağılım çarpık olduğunda diğer yöntemlerden daha az hatalı kestirimler yaptığı görülmektedir. MAP yöntemi madde sayısı az olduğunda MLE'ye göre daha düşük hatalı kestirimler yapmıştır. Madde sayısı arttıkça DE, AH, MAP, MLE yöntemlerinde kestirim hataları azalırken NBE yönteminde kestirim hataları artmıştır. NBE yöntemi Monte Carlo simülasyonuna dayandığı için kestirilecek olan parametre sayısı arttıkça hata artmaktadır. Dağılımın çarpıklığından en çok etkilenen yöntemler ise MAP ve AH yöntemleridir.

Yetenek parametrelerine ilişkin elde edilen MAE değerleri üzerinde anlamlı etkisi olan koşulların belirlenmesi için yapılan ANOVA sonucu Tablo 9'da raporlanmıştır.

Tablo 9

Yetenek parametreleri MAE değerlerine ilişkin ANOVA tablosu

Kaynak	Kareler Toplamı	Sd	Ortalama Kareler	F	p	Etki B.
Düzeltilmiş Model	23.904	134	.178	95.192	.000	.979
Kesişim	91.120	1	91.120	48622.877	.000	.994
Yöntem	19.091	4	4.773	2546.769	.000	.974
Yetenek Dağılımı	.067	2	.033	17.819	.000	.117
Madde sayısı	.314	2	.157	83.727	.000	.383
Örneklem büyüklüğü	.033	2	.016	8.696	.000	.061
Yöntem * Yetenek Dağılımı	.301	8	.038	20.094	.000	.373
Yöntem * Madde	3.940	8	.493	262.821	.000	.886
Yöntem * Örneklem	.013	8	.002	.899	.517	.026
Yetenek Dağılımı * Madde	.016	4	.004	2.121	.078	.030
Yetenek Dağılımı * Örneklem	.007	4	.002	.898	.465	.013
Madde * Örneklem	.009	4	.002	1.153	.332	.017
Yöntem * Yetenek Dağılımı * Madde	.015	16	.001	.490	.951	.028
Yöntem * Yetenek Dağılımı * Örneklem	.021	16	.001	.706	.787	.040
Yöntem * Madde * Örneklem	.021	16	.001	.684	.809	.039
Yetenek Dağılımı * Madde * Örneklem	.014	8	.002	.925	.496	.027
Yöntem * Yetenek Dağılımı * Madde * Örneklem	.044	32	.001	.733	.854	.080
Hata	.506	270	.002			
Toplam	115.530	405				
Düzeltilmiş Toplam	28.260	404				

\* $R^2= 0.979$  (Düzeltilmiş  $R^2= 0.969$ )

Tablo 9'da verilen değerler incelendiğinde, yetenek kestirimlerinden elde edilen MAE değerleri üzerinde anlamlı etkiye sahip değişkenlerden etki büyüklüğü en fazla olan kestirim yöntemidir ve büyük etkiye sahiptir ( $p<0.05$ ,  $\eta=0.99$ ). Bu değişkeni madde sayısı değişkeni takip etmektedir ( $p<0.05$ ,  $\eta=0.38$ ). MAE değerleri üzerinde yetenek dağılımı orta düzeyde etkiye sahipken ( $p<0.05$ ,  $\eta=0.12$ ), örneklem büyüklüğü küçük etkiye sahiptir ( $p<0.05$ ,  $\eta<0.20$ ). Etkileşim etkileri incelendiğinde

ise MAE deęerleri üzerinde en büyük etkiye sahip olan etkileşim kestirim yöntemi ve madde sayısı etkileşimidir ( $p<0.05$ ,  $\eta=0.89$ ). Denklemden yer alan diğer etkileşim etkileri ise istatistiksel olarak anlamlı değildir.

Kestirim yöntemleri arasındaki farklılıkların incelenmesi için yapılan Post-Hoc testlerinin sonuçları Tablo 10'da raporlanmıştır.

*Tablo 10*

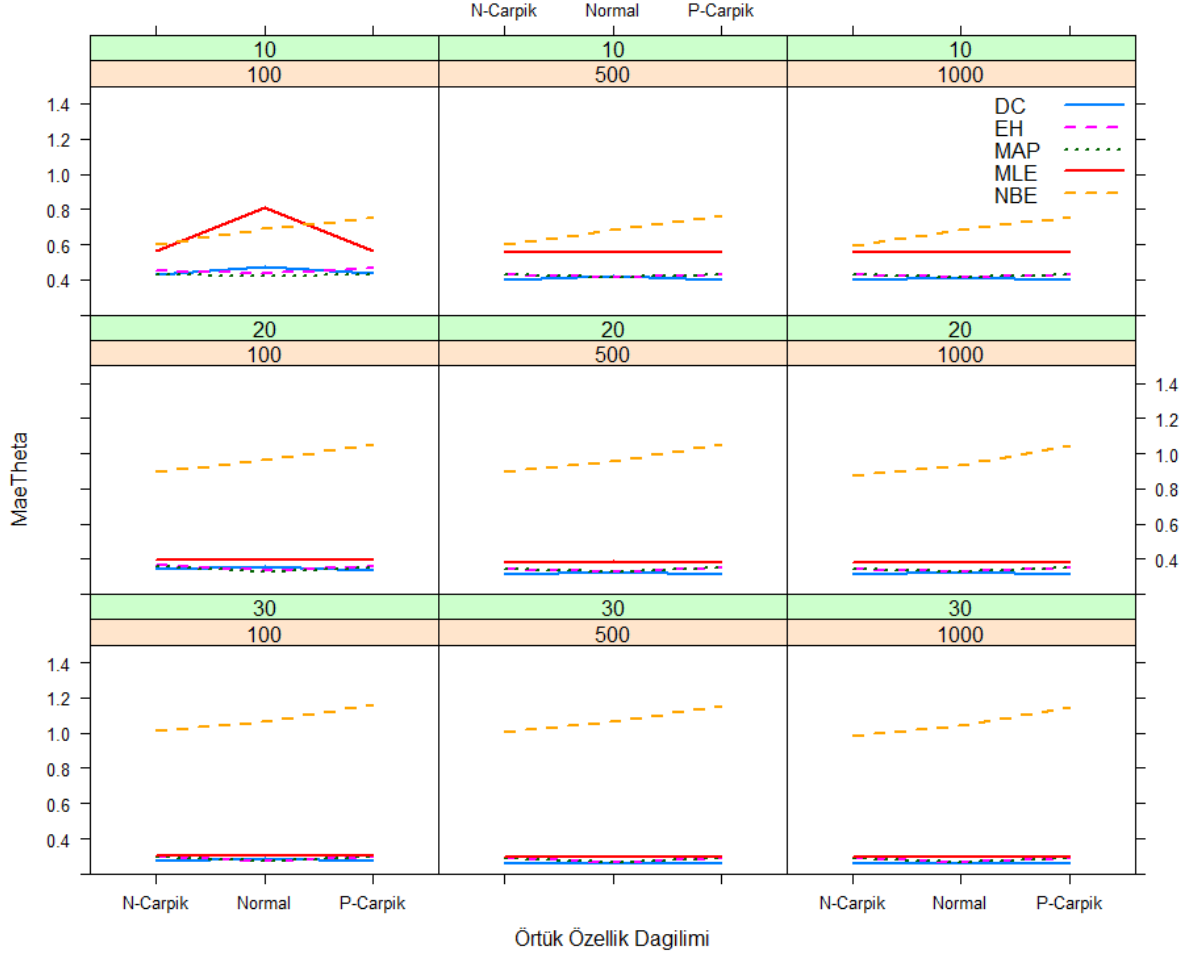
*Kestirim Yöntemleri Arasında Post-Hoc analizi sonuçları*

Yöntem I	J	Ortalama Farklar (I-J)	P
DE	AH	-.0150	.183
	MAP	-.0117	.426
	MLE	-.0861*	.000
	NBE	-.5656*	.000
MAP	AH	-.0033	.989
	MLE	-.0744*	.000
	NBE	-.5540*	.000
NBE	AH	.5507*	.000
	MLE	.4796*	.000
AH	MLE	-.0711*	.000

\* $p<.05$  düzeyinde anlamlı fark

Tablo 10 incelendiğinde DE yöntemi MLE ve NBE yöntemlerinden manidar düzeyde daha düşük hatalı kestirimler yaparken ( $p<0.05$ ), AH ve MAP yöntemleri ile kestirim hataları arasında anlamlı bir farklılık bulunmamaktadır ( $p>0.05$ ). NBE yöntemi tüm yöntemlerden manidar düzeyde daha yüksek hatalı kestirimler yapmıştır ( $p<0.05$ ). MLE yöntemi DE,AH ve MAP yöntemlerinden daha yüksek hatalı kestirimler yapmıştır ( $p<0.05$ ). Yetenek parametresi dağılımları arasındaki farklar RMSE deęerleri ile benzer şekildedir. Madde sayıları arasındaki farkların incelenmesine yönelik Post-Hoc testlerinin sonuçları madde sayısı arttıkça MAE deęerlerinin manidar düzeyde azaldığını göstermiştir ( $p<0.05$ ).

Yetenek parametresine ilişkin MAE deęerlerinin yetenek parametresinin dağılımı ve madde sayısına göre deęişimi Şekil 2'de verilmiştir.



Şekil 2. Yetenek parametresine ilişkin MAE değerlerinin yetenek parametresinin dağılımı ve madde sayısına göre değişimi

Şekil 2 incelendiğinde DE kestirimlerinin yetenek parametresinin dağılımdan etkilenmediği ve dağılım çarpık olduğunda diğer yöntemlerden daha az hatalı kestirimler yaptığı görülmektedir. MAP yöntemi MLE'ye göre daha düşük hatalı kestirimler yapmıştır. Madde sayısı arttıkça DE, AH, MAP, MLE yöntemlerinde kestirim hataları azalırken NBE yönteminde kestirim hataları artmıştır.

**a parametresine ilişkin bulgular.** a parametrelerinin RMSE değerlerine ilişkin ANOVA analizi sonuçları Tablo 11'de yer almaktadır.

Tablo 11

Ayırıcılık parametreleri RMSE değerlerine ilişkin ANOVA tablosu

Kaynak	Kareler Toplamı	Sd	Ortalama Kareler	F	p	Etki B.
Düzeltilmiş Model	10.259	134	.077	21.470	.000	.914
Kesişim	41.778	1	41.778	11715.991	.000	.977
Yöntem	1.173	4	.293	82.214	.000	.549
Yetenek Dağılımı	.577	2	.288	80.860	.000	.375
Madde sayısı	1.745	2	.873	244.731	.000	.644
Örneklem büyüklüğü	3.224	2	1.612	452.083	.000	.770
Yöntem * Yetenek Dağılımı	.251	8	.031	8.782	.000	.206
Yöntem * Madde	.279	8	.035	9.769	.000	.224
Yöntem * Örneklem	2.128	8	.266	74.593	.000	.688
Yetenek Dağılımı * Madde	.148	4	.037	10.352	.000	.133
Yetenek Dağılımı * Örneklem	.019	4	.005	1.314	.265	.019
Madde * Örneklem	.057	4	.014	4.000	.004	.056
Yöntem * Yetenek Dağılımı * Madde	.159	16	.010	2.779	.000	.141
Yöntem * Yetenek Dağılımı * Örneklem	.095	16	.006	1.660	.054	.090
Yöntem * Madde * Örneklem	.263	16	.016	4.618	.000	.215
Yetenek Dağılımı * Madde * Örneklem	.021	8	.003	.747	.650	.022
Yöntem * Yetenek Dağılımı * Madde * Örneklem	.122	32	.004	1.067	.377	.112
Hata	.963	270	.004			
Toplam	53.000	405				
Düzeltilmiş Toplam	11.222	404				

\* $R^2 = 0.914$  (Düzeltilmiş  $R^2 = 0.872$ )

Tablo 11'de verilen değerler incelendiğinde, a parametrelerinin kestirimlerinden elde edilen RMSE değerleri üzerinde anlamlı etkiye sahip değişkenlerden etki büyüklüğü en fazla olan örneklem büyüklüğü değişkenidir ve büyük etkiye sahiptir ( $p < 0.05$ ,  $\eta = 0.77$ ). Bu değişkeni sırasıyla madde sayısı ( $p < 0.05$ ,  $\eta = 0.64$ ), kestirim yöntemi ( $p < 0.05$ ,  $\eta = 0.55$ ), yetenek dağılımı ( $p < 0.05$ ,  $\eta = 0.38$ ) değişkenleri takip etmektedir. Etkileşim etkileri incelendiğinde ise RMSE değerleri

üzerinde en büyük etkiye sahip olan etkileşim kestirim yöntemi ve örneklem büyüklüğü etkileşimidir ( $p<0.05$ ,  $\eta=0.68$ ). Bu etkileşimi kestirim yöntemi ile madde sayısı etkileşimi ( $p<0.05$ ,  $\eta=0.22$ ); kestirim yöntemi, madde sayısı ile örneklem büyüklüğü etkileşimi ( $p<0.05$ ,  $\eta=0.22$ ) ve kestirim yöntemi ile yetenek dağılımı etkileşimi ( $p<0.05$ ,  $\eta=0.20$ ) takip etmektedir. Yetenek dağılımı ile madde sayısı etkileşimi; madde sayısı ile örneklem büyüklüğü etkileşimi; kestirim yöntemi madde sayısı ve yetenek dağılımı etkileşimleri de anlamlıdır ancak etki büyüklükleri küçüktür ( $p<0.05$ ,  $\eta<0.10$ ). Kestirim yöntemleri arasındaki farklılıkların incelenmesi için yapılan Post-Hoc testlerinin sonuçları Tablo 12’de verilmiştir.

*Tablo 12*

*Kestirim Yöntemleri Arasında Post-Hoc analizi sonuçları*

Yöntem		Ortalama Farklar (I-J)	P
I	J		
DE	AH	-.039586	.000
	MAP	.073894	.000
	MLE	.062306	.000
	NBE	-.042497	.000
MAP	AH	-.113481	.000
	MLE	-.011588	.731
	NBE	-.116391	.000
NBE	AH	.002911	.130
	MLE	.104803	.000
AH	MLE	.101892	.000
	MAP	.116391	.000

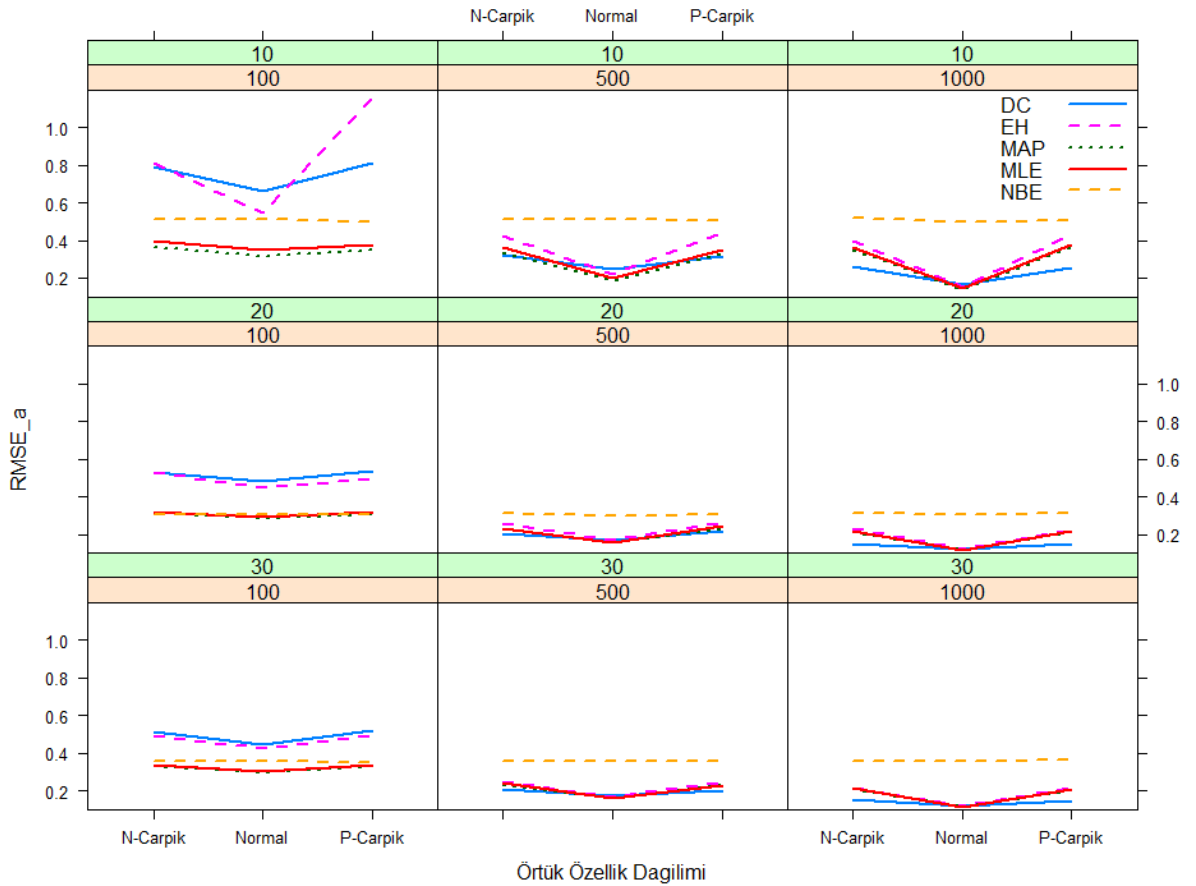
*\* $p<.05$  düzeyinde anlamlı fark*

Tablo 12 incelendiğinde DE yöntemi AH ve NBE yöntemlerinden manidar düzeyde daha düşük hatalı kestirimler yaparken ( $p<0.05$ ), MLE ve MAP kestirimlerinden daha yüksek hatalı kestirimler yapmıştır ( $p<0.05$ ). AH ve NBE yöntemlerinin kestirim hataları arasında anlamlı bir farklılık bulunmamaktadır ( $p>0.05$ ). Benzer şekilde MLE ve MAP yöntemlerinin kestirim hataları arasında anlamlı bir farklılık bulunmamaktadır ( $p>0.05$ ).



Örneklem büyüklükleri arasındaki farkların incelenmesine yönelik Post-Hoc testlerinin sonuçları örneklem büyüdükçe RMSE değerlerinin manidar düzeyde azaldığını göstermiştir ( $p<0.05$ ). Madde sayıları arasındaki farkların incelenmesine yönelik Post-Hoc testlerinin sonuçları ise 10 maddelik testten elde edilen kestirim hatalarının diğerlerinden manidar düzeyde fazla iken ( $p<0.05$ ), 20 ve 30 maddelik testlerin kestirim hatalarının anlamlı düzeyde farklılaşmadığını göstermiştir ( $p=0.79$ ). Benzer şekilde normal yetenek dağılımından elde edilen kestirim hataları çarpık dağılımlardan elde edilen hatalardan anlamlı düzeyde daha düşük ( $p<0.05$ ) iken, negatif ve pozitif çarpık dağılımlar arasında kestirim hataları anlamlı düzeyde farklılaşmamaktadır ( $p=0.7$ ).

Ayırıcılık parametresinin RMSE değerlerinin madde sayısı, örneklem büyüklüğü ve yetenek parametresi dağılımına göre değişimi Şekil 3'te verilmiştir.



Şekil 3. a parametresinin RMSE değerlerinin madde sayısı, örneklem büyüklüğü ve yetenek parametresi dağılımına göre değişimi

Şekil 3 incelendiğinde küçük örneklerde NBE yönteminin AH ve DE yöntemlerinden daha az hatalı kestirimler ürettiği, MLE ve MAP yöntemleri ile benzer kestirim hataları ürettiği görülmektedir. Örneklem büyüdükçe NBE'den elde edilen kestirim hataları değişmezken diğer yöntemlerin kestirim hataları azalmıştır. Yetenek parametresinin dağılımı çarpık olduğunda a parametresinin kestirim hatası artmıştır. Çarpık dağılımlarda en az hatalı kestirim yapan yöntem DE yöntemidir.

a parametrelerinin MAE değerlerine ilişkin ANOVA analizi sonuçları Tablo 13'de yer almaktadır.

Tablo 13'te yer alan değerler incelendiğinde, a parametrelerinin kestirimlerinden elde edilen MAE değerleri üzerinde anlamlı etkiye sahip değişkenlerden etki büyüklüğü en fazla olan değişken örneklem büyüklüğü değişkenidir ( $p<0.05$ ,  $\eta=0.86$ ). Bu değişkeni yine etki büyüklüğü yüksek düzeyde olan madde sayısı ( $p<0.05$ ,  $\eta=0.74$ ), kestirim yöntemi ( $p<0.05$ ,  $\eta=0.72$ ), yetenek dağılımı ( $p<0.05$ ,  $\eta=0.45$ ) değişkenleri takip etmektedir. Etkileşim etkileri incelendiğinde ise ikili etkileşimlerden yetenek dağılımı ve örneklem büyüklüğü etkileşimi dışındaki etkilerin tamamı anlamlı düzeydedir. Örneklem büyüklüğü ve kestirim yöntemi etkileşimi etki büyüklüğü yüksek düzeyde iken ( $p<0.05$ ,  $\eta=0.77$ ), diğer etkileşimlerin etki büyüklüğü düşüktür.

Post-Hoc analizleri sonuçları incelendiğinde ise RMSE değerlerine benzer şekilde MAE değerlerine yönelik sonuçlar da örneklem büyüdükçe MAE değerlerinin manidar düzeyde azaldığını göstermiştir ( $p<0.05$ ). Madde sayılarına ilişkin Post-Hoc testlerinin sonuçları ise 10 maddelik testten elde edilen kestirim hatalarının diğerlerinden manidar düzeyde fazla iken ( $p<0.05$ ), 20 ve 30 maddelik testlerin kestirim hatalarının anlamlı düzeyde farklılaşmadığını göstermiştir ( $p=0.79$ ). Ayrıca normal yetenek dağılımından elde edilen kestirim hataları çarpık dağılımlardan elde edilen hatalardan anlamlı düzeyde daha düşük ( $p<0.05$ ) iken, negatif ve pozitif çarpık dağılımlar arasında kestirim hataları farklılaşmamaktadır ( $p=0.71$ ).

Tablo 13

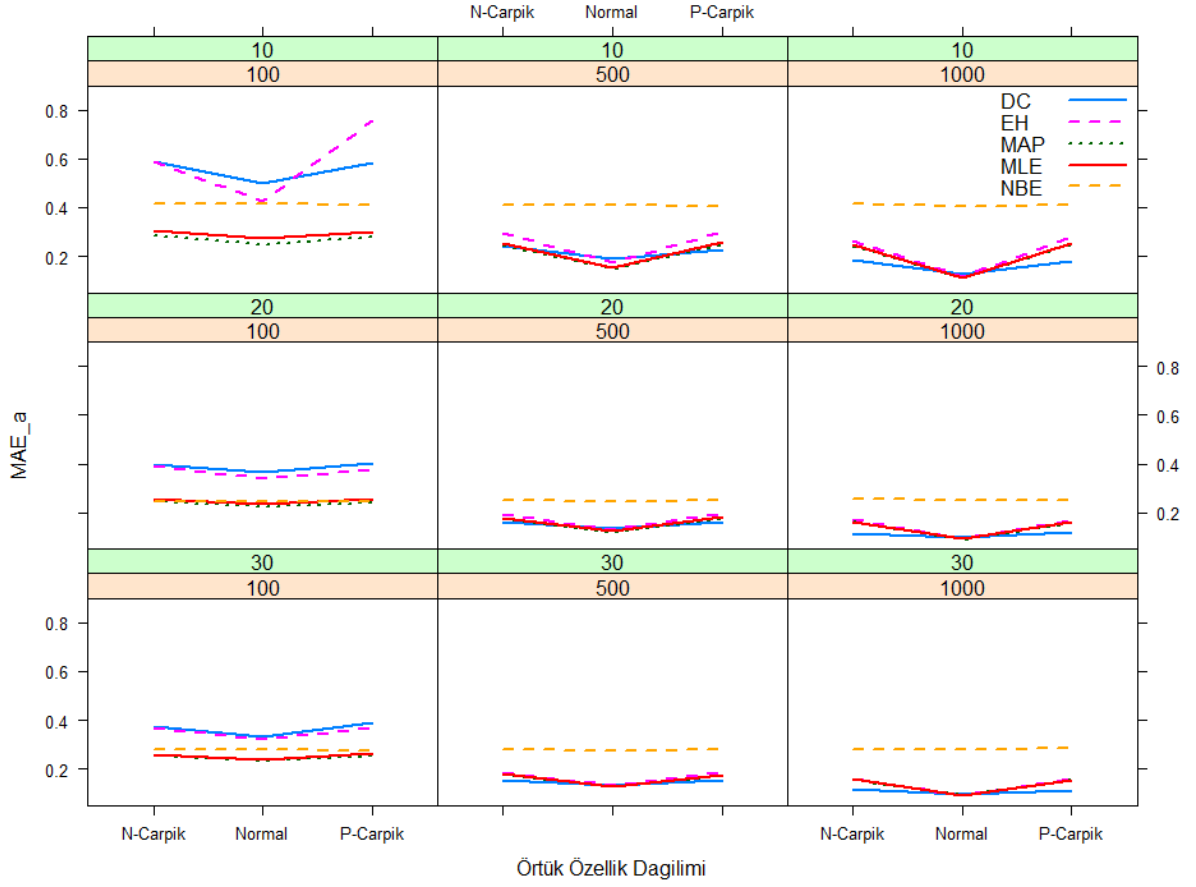
a parametrelerinin MAE değerlerine ilişkin ANOVA tablosu

Kaynak	Kareler Toplamı	Sd	Ortalama Kareler	F	P	Etki B.
Düzeltilmiş Model	5.412a	134	.040	35.999	.000	.947
Kesişim	24.225	1	24.225	21590.933	.000	.988
Yontem	.774	4	.193	172.426	.000	.719
Yetenek Dağılımı	.245	2	.123	109.231	.000	.447
Madde sayısı	.856	2	.428	381.269	.000	.739
Örneklem büyüklüğü	1.899	2	.949	846.257	.000	.862
Yontem * Yetenek Dağılımı	.093	8	.012	10.388	.000	.235
Yontem * Madde	.147	8	.018	16.428	.000	.327
Yontem * Örneklem	1.083	8	.135	120.642	.000	.781
Yetenek Dağılımı * Madde	.047	4	.012	10.497	.000	.135
Yetenek Dağılımı * Örneklem	.007	4	.002	1.580	.180	.023
Madde * Örneklem	.038	4	.009	8.423	.000	.111
Yontem * Yetenek Dağılımı * Madde	.042	16	.003	2.327	.003	.121
Yontem * Yetenek Dağılımı * Örneklem	.032	16	.002	1.776	.034	.095
Yontem * Madde * Örneklem	.112	16	.007	6.229	.000	.270
Yetenek Dağılımı * Madde * Örneklem	.004	8	.001	.469	.878	.014
Yontem * Yetenek Dağılımı * Madde * Örneklem	.034	32	.001	.937	.570	.100
Hata	.303	270	.001			
Toplam	29.940	405				
Düzeltilmiş Toplam	5.715	404				

\* $R^2 = 0.947$  (Düzeltilmiş  $R^2 = 0.921$ )

a parametresinin MAE değerlerinin madde sayısı, örneklem büyüklüğü ve yetenek parametresi dağılımına göre değişimi Şekil 4'te verilmiştir. Şekil 4 incelendiğinde MAE değerlerinin dağılımının RMSE değerlerine oldukça benzer olduğu görülmektedir. Küçük örnekleme en az hatalı kestirimleri MAP verirken büyük örneklemelerde çarpık dağılımlarda DE daha az hatalı kestirimler üretmiştir.

Küçük örneklem ve az sayıda madde ile DE ve AH yöntemlerinin kestirim hataları diğerlerinden daha yüksektir.



Şekil 4. Ayırıcılık parametresinin MAE değerlerinin madde sayısı, örneklem büyüklüğü ve yetenek parametresi dağılımına göre değişimi

**b parametresine ilişkin bulgular.** b parametrelerinin RMSE değerlerine ilişkin ANOVA analizi sonuçları Tablo 14'de yer almaktadır.

Tablo 14'te verilen değerler incelendiğinde, b parametrelerinin kestirimlerinden elde edilen RMSE değerleri üzerinde anlamlı etkiye sahip değişkenlerden etki büyüklüğü en fazla olan kestirim yöntemi ( $p < 0.05$ ,  $\eta = 0.78$ ) ve örneklem büyüklüğü ( $p < 0.05$ ,  $\eta = 0.51$ ) değişkenleridir ve orta düzeyde etki büyüklüğüne sahiptir. Bu değişkenleri etki büyüklüğü küçük olan yetenek dağılımı ( $p < 0.05$ ,  $\eta = 0.42$ ) ve örneklem büyüklüğü ( $p < 0.05$ ,  $\eta = 0.18$ ) değişkenleri takip etmektedir. Etkileşim etkileri incelendiğinde en büyük etkiye sahip etkileşim yöntem ve örneklem büyüklüğü etkileşimidir ( $p < 0.05$ ,  $\eta = 0.38$ ). Modelde yer alan ikili

etkileşimlerin tamamı anlamlıdır ancak etki büyüklüğü düşüktür ( $p < 0.05$ ,  $\eta < 0.20$ ). Üçlü etkileşimlerin ise hiçbiri anlamlı bulunmamıştır.

*Tablo 14*

*b parametreleri RMSE değerlerine ilişkin ANOVA tablosu*

Kaynak	Kareler Toplamı	Sd	Ortalama Kareler	F	p	Etki B
Düzeltilmiş Model	7.187a	134	.054	13.636	.000	.871
Kesişim	34.852	1	34.852	8861.509	.000	.970
Yöntem	3.684	4	.921	234.193	.000	.776
Yetenek Dağılımı	.770	2	.385	97.851	.000	.420
Madde sayısı	.227	2	.113	28.795	.000	.176
Örneklem büyüklüğü	1.122	2	.561	142.653	.000	.514
Yöntem * Yetenek Dağılımı	.281	8	.035	8.939	.000	.209
Yöntem * Madde	.110	8	.014	3.500	.001	.094
Yöntem * Örneklem	.659	8	.082	20.936	.000	.383
Yetenek Dağılımı * Madde	.127	4	.032	8.083	.000	.107
Yetenek Dağılımı * Örneklem	.055	4	.014	3.476	.009	.049
Madde * Örneklem	.019	4	.005	1.236	.296	.018
Yöntem * Yetenek Dağılımı * Madde	.047	16	.003	.748	.743	.042
Yöntem * Yetenek Dağılımı * Örneklem	.043	16	.003	.691	.803	.039
Yöntem * Madde * Örneklem	.017	16	.001	.276	.998	.016
Yetenek Dağılımı * Madde * Örneklem	.014	8	.002	.453	.888	.013
Yöntem * Yetenek Dağılımı * Madde * Örneklem	.011	32	.000	.084	1.000	.010
Hata	1.062	270	.004			
Toplam	43.101	405				
Düzeltilmiş Toplam	8.249	404				

\* $R^2 = 0.871$  (Düzeltilmiş  $R^2 = 0.807$ )

Kestirim yöntemleri arasında farklılıklara yönelik Post-Hoc analiz sonuçları ise Tablo 15'te özetlenmiştir.

Tablo 15

Kestirim Yöntemleri Arasında  $b$  parametresi RMSE değerlerine ilişkin Post-Hoc analizi sonuçları

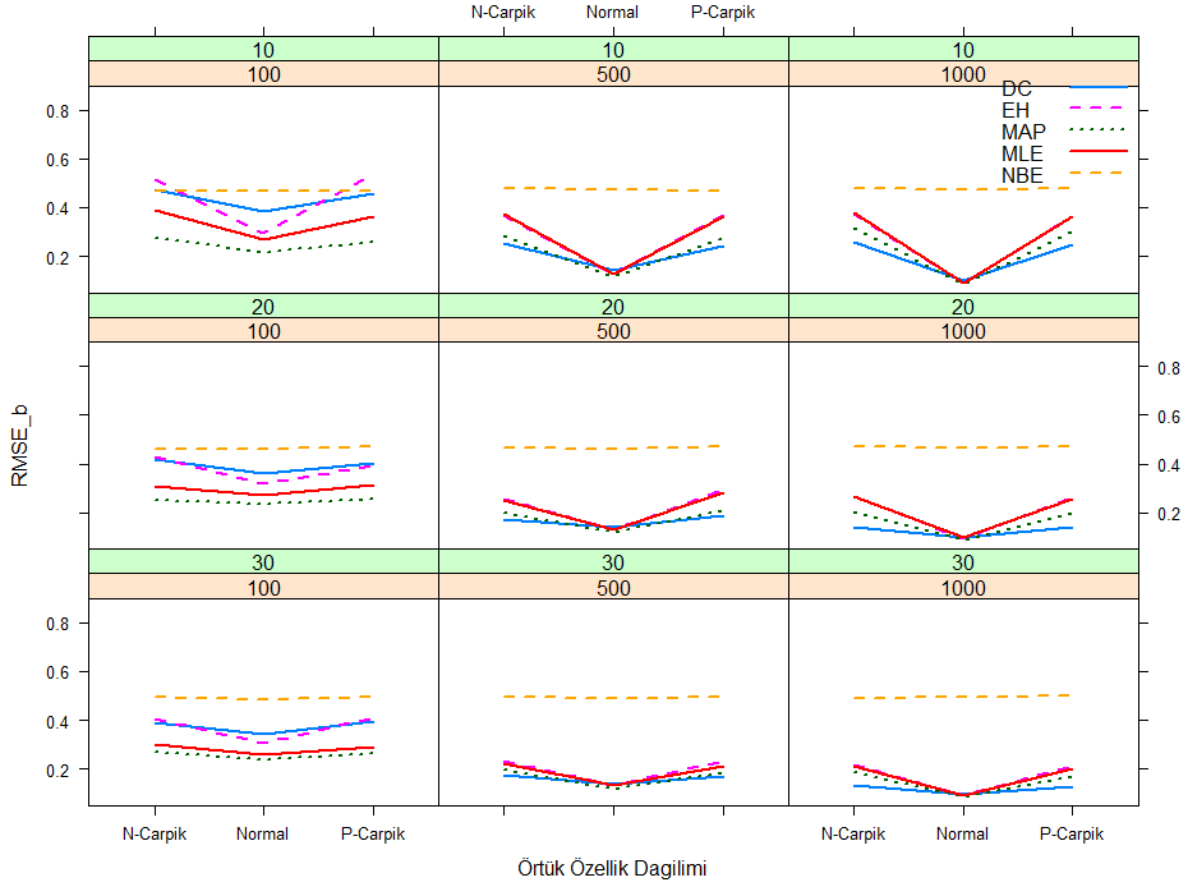
I	Yöntem	J	Ortalama Farklar (I-J)	P
DE		AH	-.0424*	.000
		MAP	.0363*	.003
		MLE	-.0090	1.000
		NBE	-.2338*	.000
MAP		AH	-.0786*	.000
		MLE	-.0452*	.000
		NBE	-.2701*	.000
NBE		AH	.1915*	.000
		MLE	.2249*	.000
AH		MLE	.0334*	.008
		MAP	.2701*	.000

\* $p < .05$  düzeyinde anlamlı fark

Tablo 15'te verilen değerler incelendiğinde NBE yönteminin diğer yöntemlerden anlamlı düzeyde daha yüksek hatalı kestirimler yaptığı görülmektedir ( $p < 0.05$ ). MAP ve DE yöntemleri arasındaki fark istatistiksel olarak anlamlı değildir ( $p > 0.05$ ). Benzer şekilde DE ve MLE yöntemleri arasında da anlamlı bir fark gözlenmemiştir ( $p > 0.05$ ). AH yöntemi ise MAP, MLE ve DE yönteminden anlamlı düzeyde yüksek hatalı kestirimler üretmiştir ( $p < 0.05$ ).

Post-Hoc analizleri sonuçları incelendiğinde ise  $b$  parametreleri için elde edilen RMSE değerlerine yönelik sonuçlar 100 örneklem büyüklüğünde elde edilen hataların diğerlerinden anlamlı düzeyde yüksek olduğunu ( $p < 0.05$ ) ancak 500 ve 1000 örneklem büyüklükleri arasında anlamlı bir fark olmadığını göstermiştir. Yetenek dağılımına ilişkin sonuçlar da normal yetenek dağılımından elde edilen kestirim hataları çarpık dağılımlardan elde edilen hatalardan anlamlı düzeyde daha düşük ( $p < 0.05$ ) iken, negatif ve pozitif çarpık dağılımlar arasında kestirim hataları anlamlı düzeyde farklılaşmadığını göstermiştir.

Güçlük parametresinin RMSE değerlerinin madde sayısı, örneklem büyüklüğü ve yetenek parametresi dağılımına göre değişimi şekil 5’de verilmiştir.



Şekil 5. Güçlük parametresinin RMSE değerlerinin madde sayısı, örneklem büyüklüğü ve yetenek parametresi dağılımına göre değişimi

Şekil 5 incelendiğinde küçük örneklemelerde ve 10 maddelik teste NBE yönteminin çarpık dağılımlarda AH yöntemi ile benzer sonuçlar ürettiği görülmüştür. Örneklem büyüklüğü 100 iken en iyi kestirimler MAP yöntemi ile elde edilmiştir. Örneklem büyüdükçe MLE, MAP, AH ve DE kestirim hataları azalırken NBE hataları değişmemiştir. Çarpık dağılımlarda 500 ve 1000 örneklem büyüklüklerinde en iyi kestirimler DE yöntemi ile elde edilmiştir. b parametrelerinin MAE değerlerine ilişkin ANOVA analizi sonuçları Tablo 16’da yer almaktadır.

Tablo 16

*b parametreleri MAE değerlerine ilişkin ANOVA tablosu*

Kaynak	Kareler Toplamı	Sd	Ortalama Kareler	F	p	Etki B.
Düzeltilmiş Model	3.563a	134	.027	42.009	.000	.954
Kesişim	15.256	1	15.256	24104.601	.000	.989
Yontem	2.121	4	.530	837.981	.000	.925
Yetenek Dağılımı	.152	2	.076	120.003	.000	.471
Madde sayısı	.076	2	.038	60.270	.000	.309
Örneklem büyüklüğü	.807	2	.404	637.892	.000	.825
Yontem * Yetenek Dağılımı	.050	8	.006	9.810	.000	.225
Yontem * Madde	.018	8	.002	3.491	.001	.094
Yontem * Örneklem	.276	8	.034	54.461	.000	.617
Yetenek Dağılımı * Madde	.027	4	.007	10.732	.000	.137
Yetenek Dağılımı * Örneklm	.007	4	.002	2.764	.028	.039
Madde * Örneklem	.001	4	.000	.455	.768	.007
Yontem * Yetenek Dağılımı * Madde	.009	16	.001	.905	.564	.051
Yontem * Yetenek Dağılımı * Örneklem	.009	16	.001	.878	.595	.049
Yontem * Madde * Örneklem	.005	16	.000	.466	.961	.027
Yetenek Dağılımı * Madde * Örneklem	.002	8	.000	.438	.898	.013
Yontem * Yetenek Dağılımı * Madde * Örneklem	.002	32	.000	.112	1.000	.013
Hata	.171	270	.001			
Toplam	18.989	405				
Düzeltilmiş Toplam	3.734	404				

\* $R^2 = 0.954$  (Düzeltilmiş  $R^2 = 0.932$ )

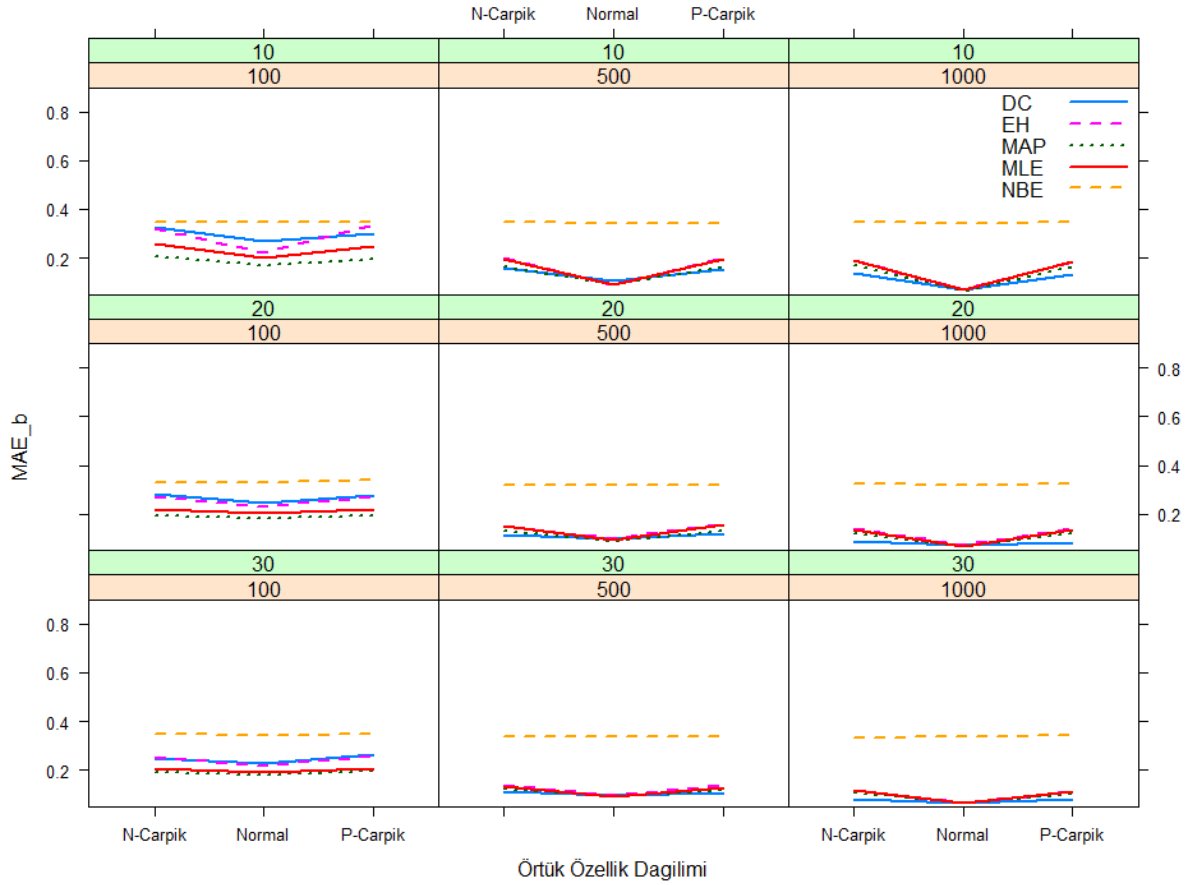
Tablo 16'daki değerler incelendiğinde, b parametrelerinin kestirimlerinden elde edilen MAE değerleri üzerinde anlamlı etkiye sahip değişkenler yüksek düzeyde etkiye sahip olan kestirim yöntemi ( $p < 0.05$ ,  $\eta = 0.93$ ), örneklem büyüklüğü ( $p < 0.05$ ,  $\eta = 0.83$ ), yetenek dağılımı ( $p < 0.05$ ,  $\eta = 0.47$ ) ve madde sayısı ( $p < 0.05$ ,  $\eta = 0.31$ ) değişkenleridir. Etkileşim etkileri incelendiğinde ikili etkileşimlerin tamamı anlamlıdır. Bu etkileşimlerden kestirim yöntemi ve örneklem büyüklüğü etkileşimi



yüksek düzeyde etkiye sahipken ( $p < 0.05$ ,  $\eta = 0.62$ ) diğer etkileşimlerin etki büyüklüğü küçüktür ( $p < 0.05$ ,  $\eta < 0.20$ ). Üçlü etkileşimlerin ise hiçbiri anlamlı değildir ( $p > 0.05$ ).

Post-Hoc analizleri sonuçları incelendiğinde ise örneklem büyüdükçe MAE değerlerinin manidar düzeyde azaldığını gözlenmiştir ( $p < 0.05$ ). Yetenek dağılımına ilişkin sonuçlar ise yetenek normal dağıldığında hataların çarpık dağılımlara göre manidar düzeyde düşük olduğunu göstermiştir ( $p < 0.05$ ).

$b$  parametresinin MAE değerlerinin madde sayısı, örneklem büyüklüğü ve yetenek parametresi dağılımına göre değişimi Şekil 6'da verilmiştir.



Şekil 6.  $b$  parametresinin MAE değerlerinin madde sayısı, örneklem büyüklüğü ve yetenek parametresi dağılımına göre değişimi

Şekil 6 incelendiğinde küçük örneklemelerde ve 10 maddelik testte NBE yönteminin çarpık dağılımlarda AH yöntemi ile benzer sonuçlar ürettiği görülmüştür. Örneklem büyüklüğü 100 iken en iyi kestirimler MAP ve MLE yöntemleri ile elde

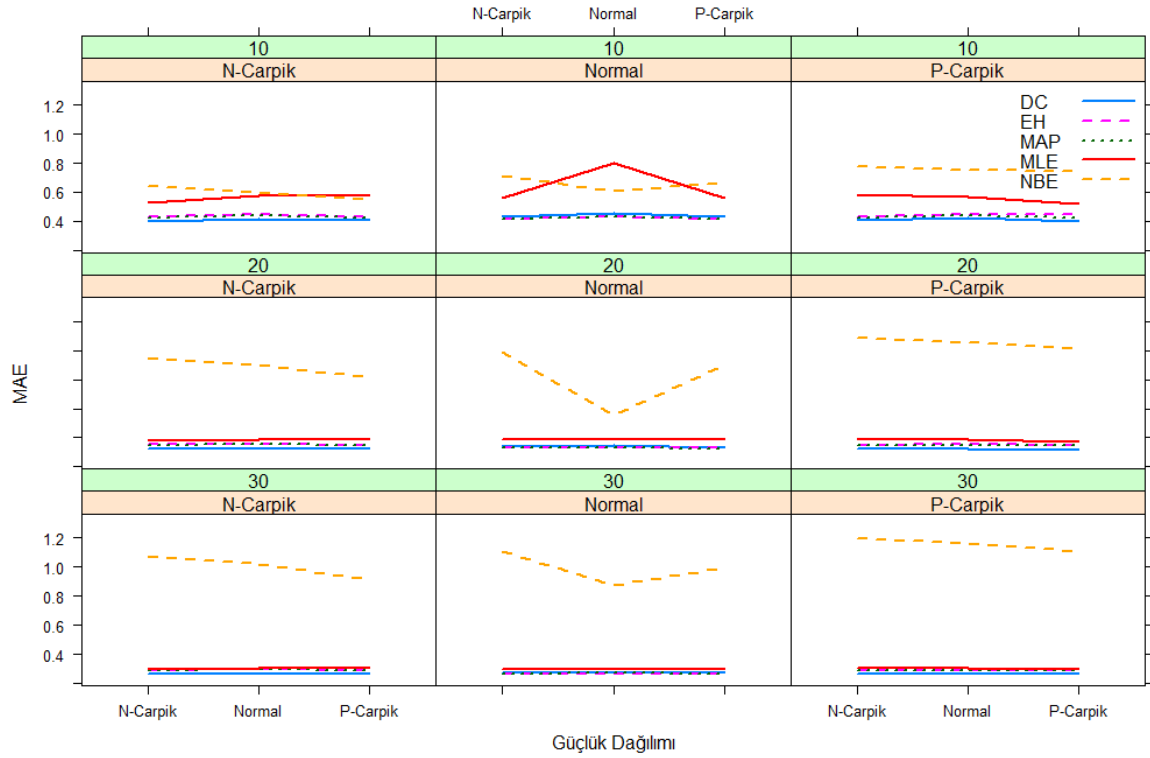
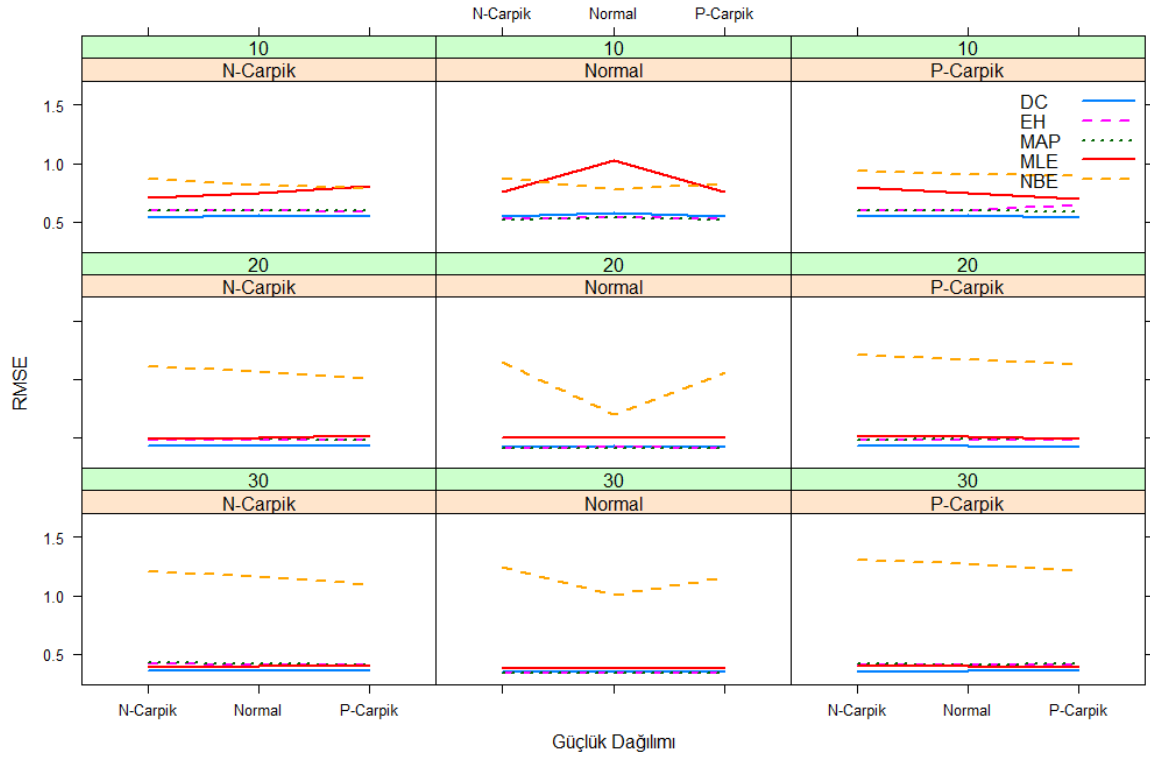
edilmiştir. 500 ve 1000 örneklem büyüklüklerinde çarpık dağılımlarda DE yöntemi daha az hatalı kestirimler üretmektedir.

### **İkinci Alt probleme ilişkin bulgular**

Araştırmanın ikinci alt problemi “Normal ve normal olmayan madde güçlük dağılımlarında, farklı örneklem büyüklükleri ve test uzunluklarında parametrik MML, MAP ve parametrik olmayan NBE, DE ve AH yaklaşımlarının hangisi daha yansız madde ve yetenek kestirimleri üretir?” olarak tanımlanmıştır. Bu bağlamda bulgular sırasıyla yetenek parametresi, a parametresi ve b parametresine ilişkin bulgular olmak üzere üç başlıkta raporlanmıştır.

**Yetenek parametresine ilişkin bulgular.** Yetenek parametre kestirimlerinden elde edilen RMSE ve MAE değerleri ile madde güçlük parametrelerinin dağılımları arasındaki ilişkiye yönelik grafikler ise Şekil 7’de yer almaktadır.

Şekil 7’de x ekseninde yer alan çarpıklık düzeyleri b parametrelerine ilişkin çarpıklıkları göstermektedir. Her bir hücrenin başlığında yer alan çarpıklık düzeyleri ise yetenek parametresine ilişkin çarpıklık düzeylerini göstermektedir. Şekil 7’de görüldüğü gibi RMSE ve MAE değerleri benzerlik göstermektedir. Genel olarak MAP, DE ve AH kestirimlerinden elde edilen hata değerleri b dağılımının düzeyleri arasında farklılaşmamaktadır. Diğer yandan MLE kestirimlerinin b parametresinin dağılımından yalnızca madde sayısı az olduğunda etkilendiği görülmektedir. On maddelik teste yetenek dağılımı negatif çarpık olduğunda MLE kestirimlerinde elde edilen hatalar b parametrelerinin pozitif olduğu durumda daha yüksektir. NBE kestirimleri de genel olarak b parametresinin dağılımı normal olduğunda çarpık olduğu durumlara göre daha az hatalı kestirimler yapmıştır. ANOVA sonuçları ise hataların b parametrelerinin düzeylerinde farklılaştığını göstermektedir ancak farkın etki büyüklüğü düşüktür ( $p < .05$ ,  $\eta = 0.04$ ).

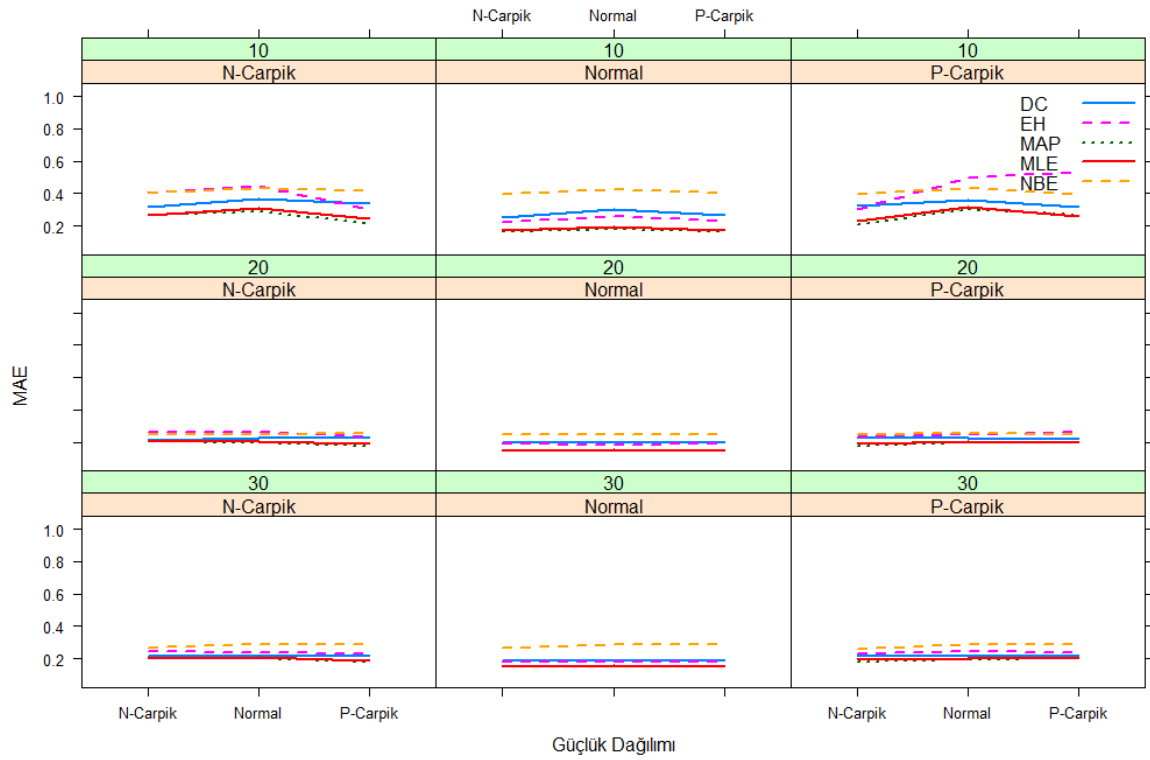
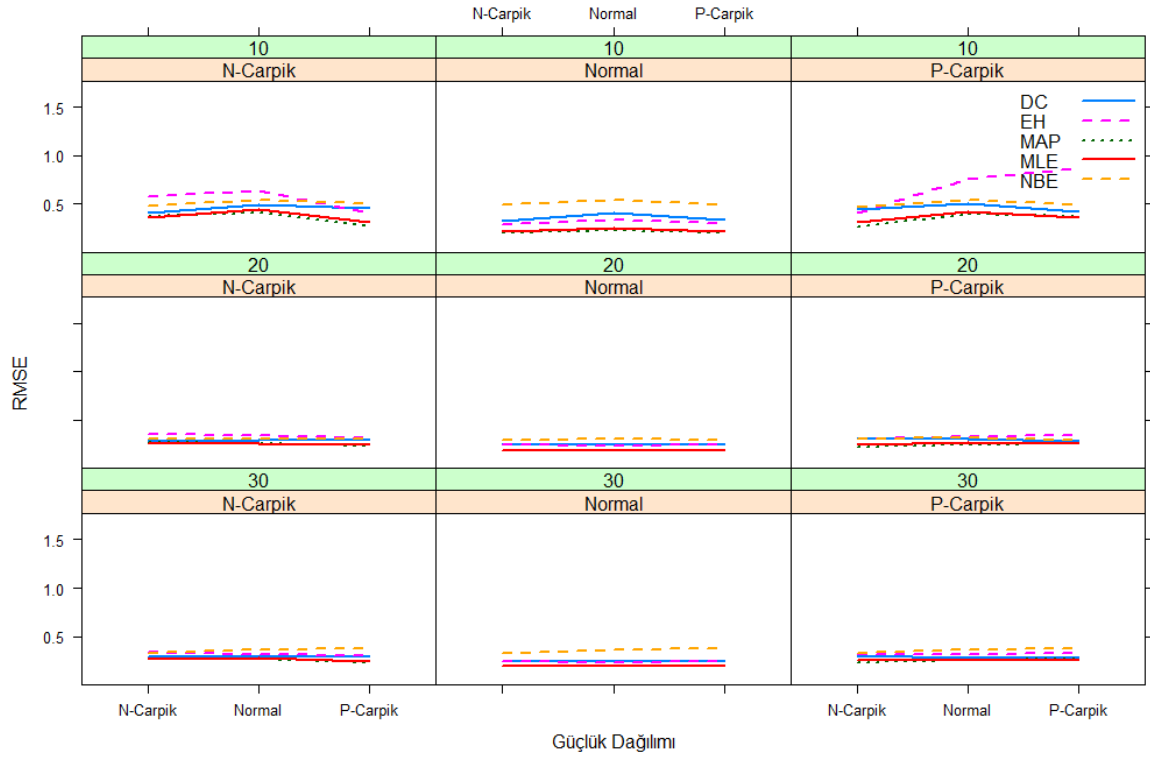


Şekil 7. Yetenek parametresine ilişkin RMSE ve MAE değerlerinin yetenek parametresinin dağılımı, madde sayısı ve güçlük parametresinin dağılımına göre değişimi

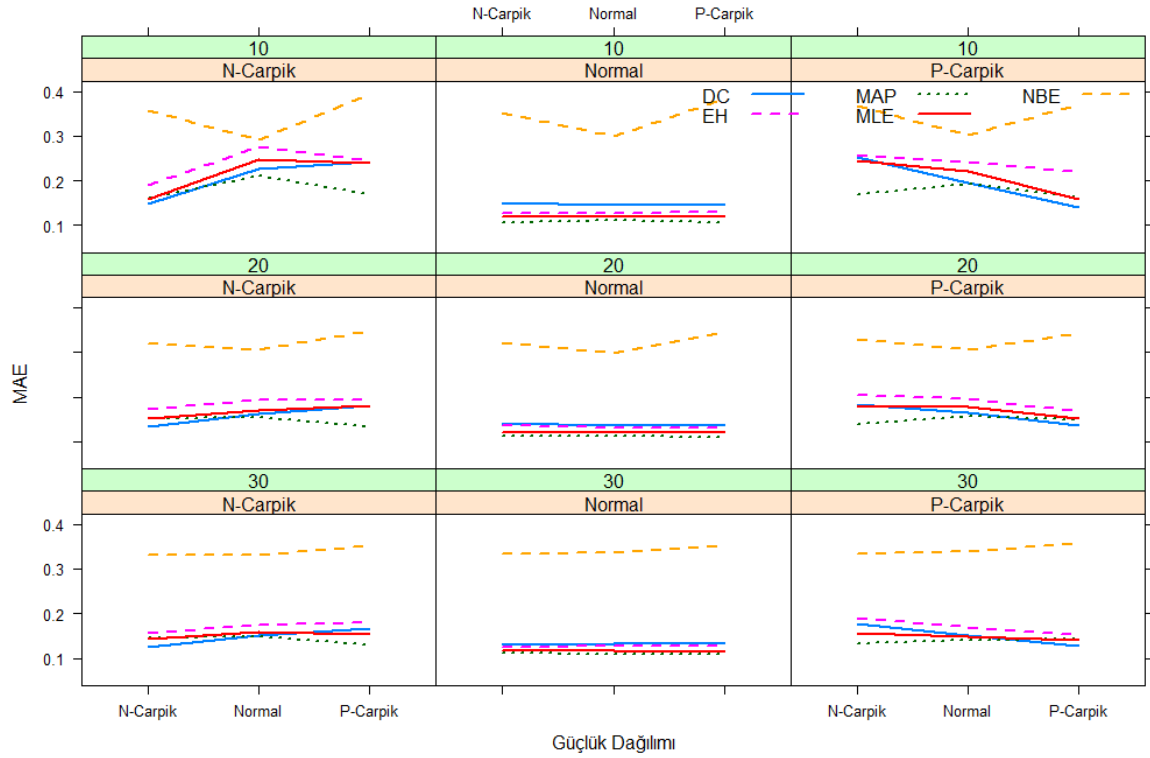
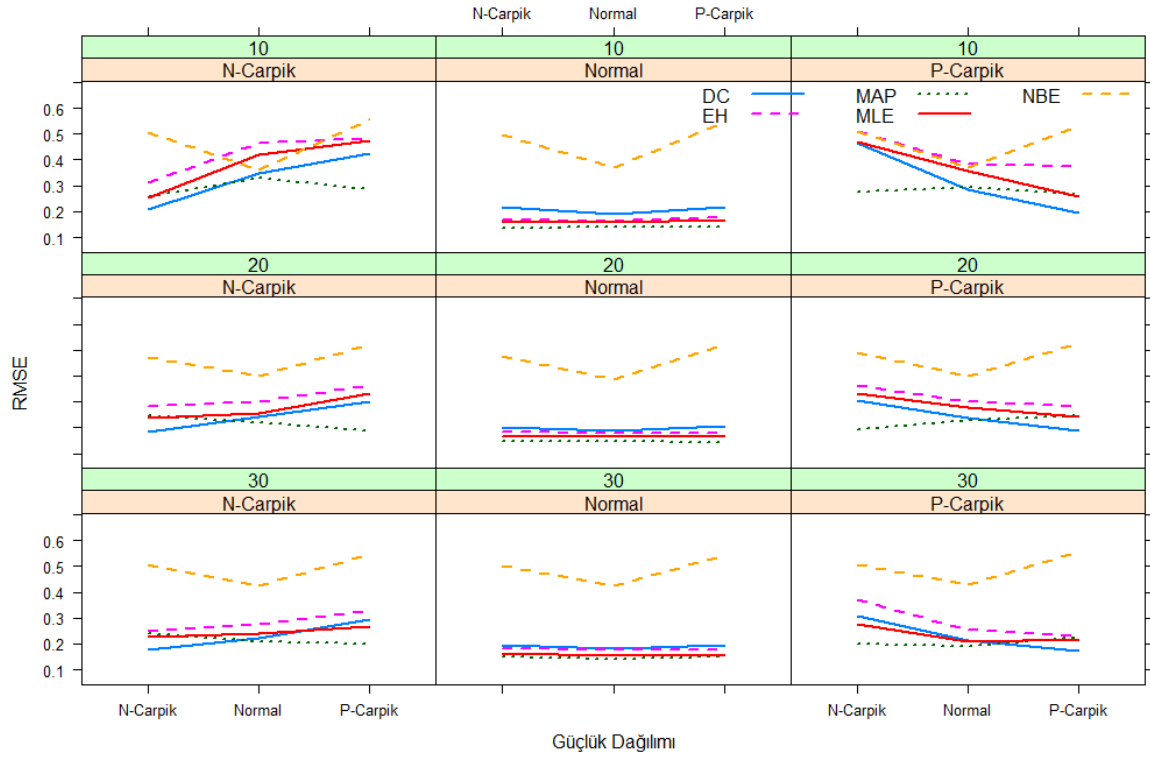
**a parametresine ilişkin bulgular.** a parametresinin kestirimlerinden elde edilen RMSE ve MAE değerleri ile madde sayısı ve b parametresinin dağılımı arasındaki ilişkiye yönelik grafikler Şekil 8’de yer almaktadır. Şekil 8 genel olarak incelendiğinde madde sayısı 20 ve 30 olduğu durumlarda a parametrelerinin kestirim hatalarının b parametresi dağılımından etkilenmediği gözlenmektedir. Madde sayısının 10 olduğu durumda ise yetenek dağılımının pozitif çarpık olduğu durumda tüm kestirim yöntemleri için en yüksek hatalı kestirimler b parametrelerinin pozitif çarpık olduğu durumlarda elde edilmiştir. Ancak bu farklılıklar istatistiksel olarak anlamlı bulunmamıştır ( $p > .05$ ).

**b parametresine ilişkin bulgular.** Madde güçlük parametre kestirimlerinden elde edilen RMSE ve MAE değerleri ile madde sayısı ve madde b parametrelerinin dağılımları arasındaki ilişkiye yönelik grafikler ise Şekil 9’da yer almaktadır. Şekil 9 incelendiğinde yetenek parametre dağılımı normal olduğunda MAP, MLE, DE, ve AH kestirimlerinin b parametresinin dağılımından etkilenmediği görülmektedir. Diğer yandan NBE kestirimleri tüm durumlarda b dağılımı normal olduğunda daha düşük hatalı kestirimler üretmiştir. Bu durum NBE kestirimlerinin b parametrelerini normal dağılım temelli bir dağılımdan örneklemesinden kaynaklanabilir.

Diğer yandan örtük özellik normal dağılmadığında tüm kestirim yöntemleri b parametrelerinin dağılımından etkilenmiştir. Genel olarak yetenek dağılımı negatif çarpık olduğunda en yüksek hatalı kestirimler b parametrelerinin dağılımı pozitif çarpık olduğunda elde edilmiştir. Yetenek dağılımı pozitif çarpık olduğunda ise en yüksek hatalı kestirimler b parametresinin dağılımı negatif çarpık olduğunda elde edilmiştir. ANOVA sonuçları b parametrelerinin dağılımı normal olduğunda elde edilen kestirim hatalarının diğerlerine göre daha düşük olduğunu göstermiştir ( $p < .05$ ). Pozitif ve negatif çarpık b parametrelerinin dağılımlarından elde edilen kestirim hataları ise farklılaşmamaktadır ( $p > .05$ ).



Şekil 8. a parametresine ilişkin RMSE ve MAE değerlerinin yetenek parametresinin dağılımı, madde sayısı ve b parametresinin dağılımına göre değişimi

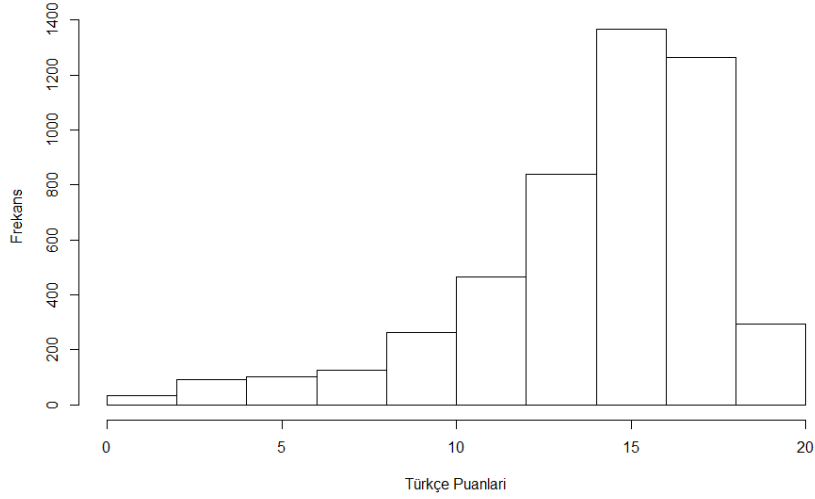


Şekil 9. b parametresine ilişkin RMSE ve MAE değerlerinin yetenek parametresinin dağılımı, madde sayısı ve b parametresinin dağılımına göre değişimi

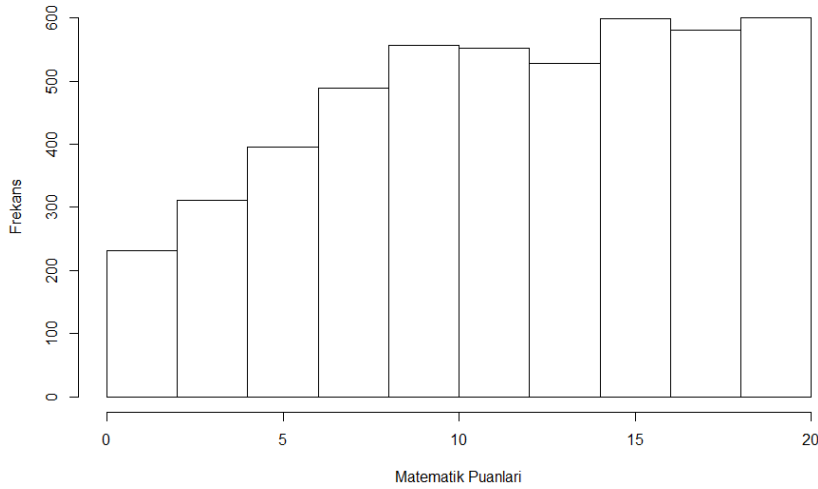
## Üçüncü Alt Probleme İlişkin Bulgular

Araştırmanın üçüncü alt problemi “Gerçek veri kümelerinde parametrik MML, MAP ve parametrik olmayan NBE, DE ve AH yaklaşımlarının hangisi daha yansız kestirimler üretir?” olarak tanımlanmıştır. Bu bağlamda bir özel okulda uygulanan deneme sınavının matematik ve Türkçe alt testlerine ilişkin veriler kullanılmıştır. Bu testlerin kullanılmasının temel nedeni puan dağılımlarının birbirinden farklı olmasının beklenmesidir. Türkçe ve matematik ham puan dağılımlarına ilişkin histogramlar ise Şekil 10’da yer almaktadır.

Türkçe Puanlarının Dağılımı



Matematik puanlarının dağılımı



Şekil 10. Türkçe ve Matematik ham puan dağılımları

Şekil 10'da yer alan histogramlar incelendiğinde Türkçe puanları dağılımının sola çarpık olduğu, Matematik ham puanları ise tek biçimli dağılıma benzediği görülmektedir.

Kestirimler hem verinin tamamından hem de veriden rastgele seçilen 1000, 500 ve 100 kişilik alt örneklemelerden elde edilmiştir. Verinin tamamına ve elde edilen alt örneklemelere ilişkin ham puanların betimsel istatistikleri Tablo 17'de verilmiştir.

*Tablo 17*

*Ham puanlara ilişkin betimsel istatistikler*

	Örneklem Büüklüğü	Min.	Mak.	Ortalama	Standart Sapma	Çarpıklık	Basıklık
Türkçe	100	2	17	13,720	3,229	-1,520	5,348
	500	2	17	13,620	3,227	-1,389	4,696
	1000	0	17	13,834	3,263	-1,617	5,593
	4844	0	17	13,690	3,363	-1,543	5,197
Matematik	100	0	20	12,630	5,554	-0,396	2,036
	500	0	20	11,928	5,278	-0,222	2,114
	1000	0	20	11,864	5,440	-0,252	2,028
	4844	0	20	11,802	5,379	-0,240	2,038

Türkçe testi 17 maddeden matematik testi ise 20 maddeden oluşmaktadır. Tablo 17'de verilen değerler incelendiğinde Türkçe testinin çarpıklığının yaklaşık -1.5 olduğu görülmektedir. Basıklık değerleri ise oldukça yüksektir. Matematik testinin ise çarpıklık değeri normal dağılımdan çok sapmamakla beraber basıklık değeri dağılımın normallikten saptığını göstermektedir. Bu kapsamda Türkçe puanlarının dağılımı simülasyon koşullarına benzer şekilde çarpıklaşırken Matematik puanlarının dağılımı simülasyon koşullarından farklılaşmaktadır. Madde sayısı olarak ise gerçek veriler simülasyon koşullarına benzer şekilde az sayıda maddeden oluşmaktadır.

Araştırma kapsamında her bir örneklem için kestirimler ayrı ayrı yapılmıştır. Kestirimlerden elde edilen BIC değerleri Tablo 18'de yer almaktadır.



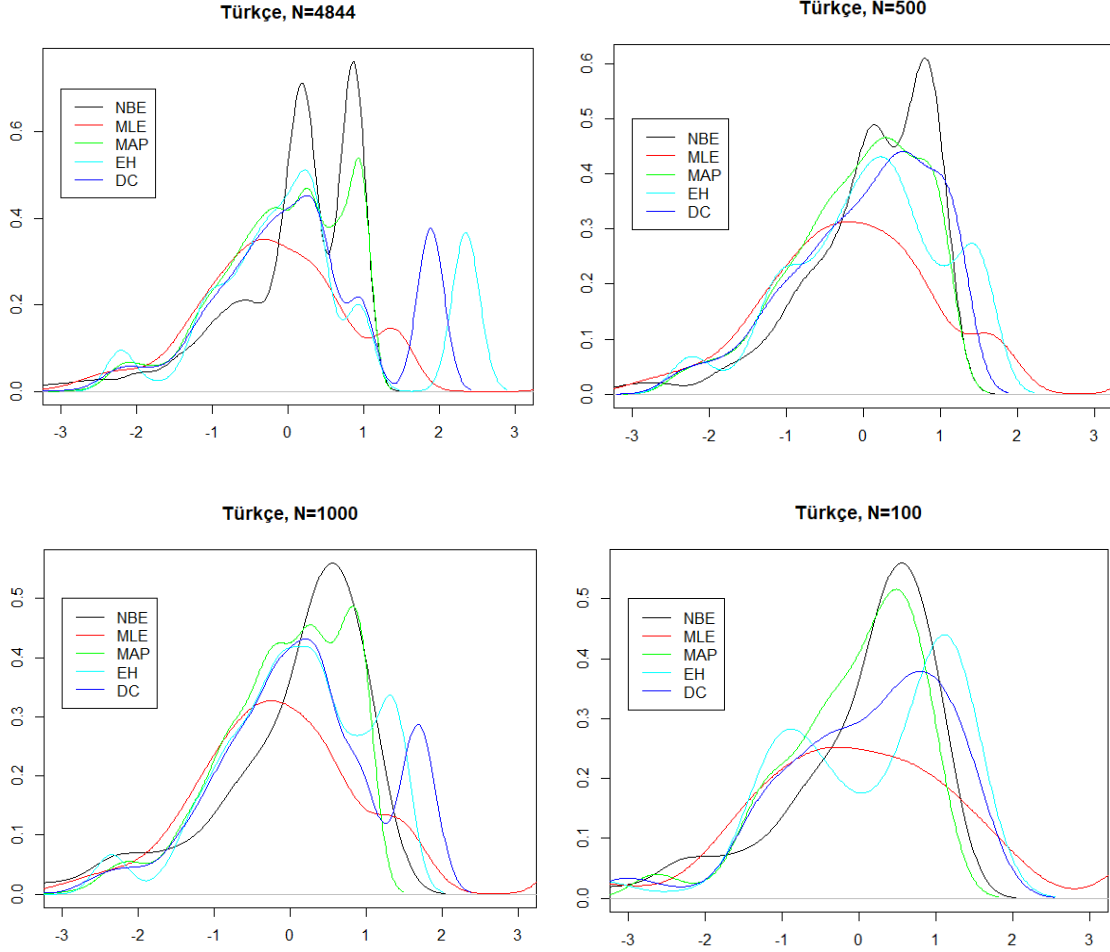
Tablo 18

Kestirimlere ilişkin BIC deęerleri

	Örneklem Büyüküğü	MLE	MAP	DE	AH	NBE
Türkçe	100	1601.12	1660.29	<b>1627.26</b>	1634.94	2018.58
	500	8723.45	8945.55	<b>8904.53</b>	8877.08	10373.07
	1000	39826.78	40969.32	<b>17420.25</b>	17488.59	21029.79
	4844	<b>92095.90</b>	94064.86	93156.86	93150.80	112141.53
Matematik	100	2124.07	2175.28	<b>2123.22</b>	2137.16	2720.38
	500	11806.71	11980.48	<b>11827.78</b>	11854.77	15192.85
	1000	68986.91	68079.94	<b>24287.77</b>	24308.75	30907.11
	4844	<b>125540.97</b>	126875.74	125980.47	126231.03	155735.45

Tablo 18 incelendięinde tüm örneklem büyüküklerinde hem Türkçe hem de Matematik testlerinde en düşük 100, 500 ve 1000 örneklem büyüküklerinde en düşük BIC deęerleri DE kestiriminden elde edilmiştir. AH ve DE kestirimleri benzer BIC deęerleri üretmiştir. Grubun tamamının kullanıldığı 4844 kişilik örneklemde ise en düşük BIC deęeri MLE kestirimleri ile elde edilmiştir.

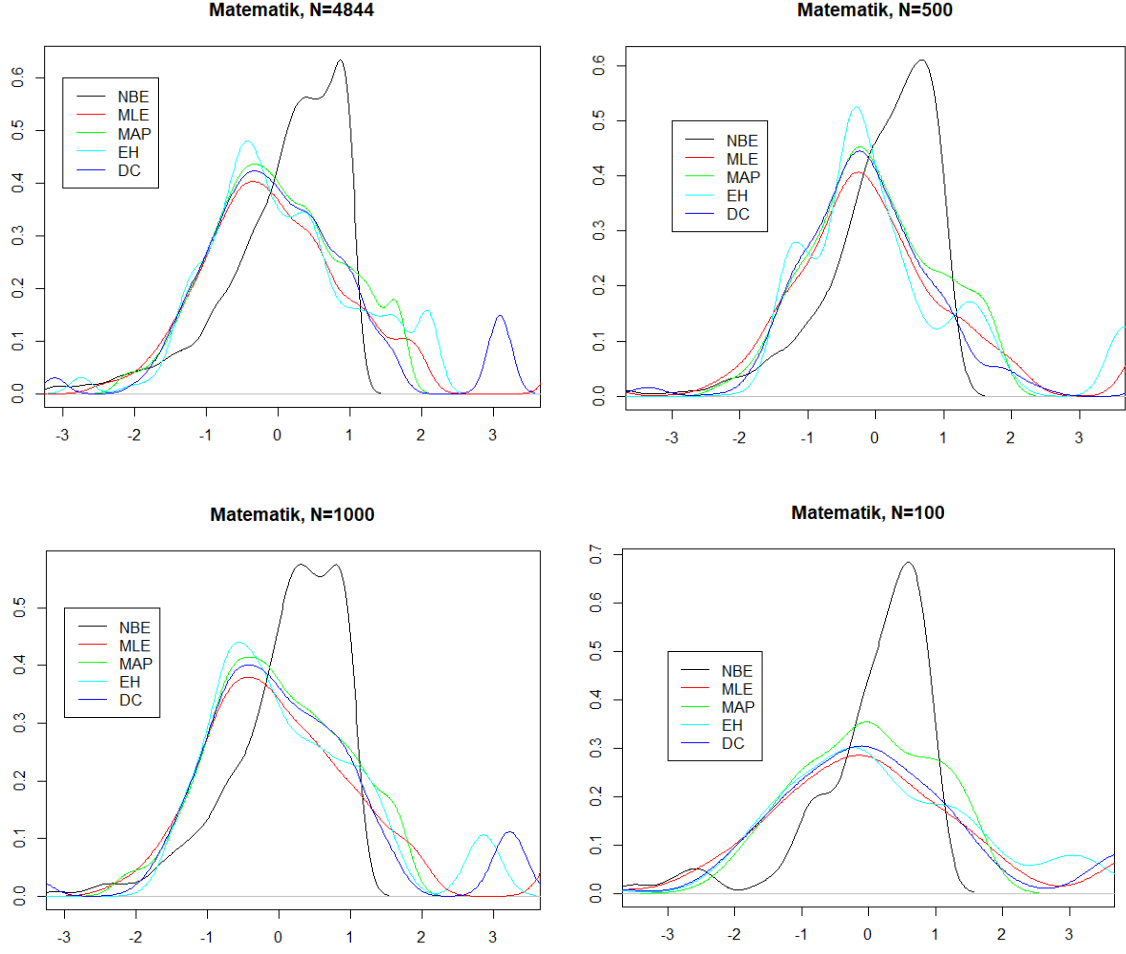
Yetenek kestirimlerine ilişkin Türkçe dersinden elde edilen yoğunluk grafikleri örneklem büyüküğü bazında Şekil 11'de sunulmuştur.



**Şekil 11. Türkçe yetenek kestirimleri**

Şekil 11 incelendiğinde MLE yönteminden elde edilen yetenek dağılımlarının normale yakın olduğu gözlenmektedir. Diğer yandan AH ve DE kestirimleri normal dağılımdan oldukça farklıdır. Ham puan dağılımları ve yetenek kestirim dağılımları birlikte incelendiğinde, ham puan dağılımı üst grupta daha fazla sayıda birey olduğunu göstermektedir. MAP ve NBE kestirimlerinden elde edilen yetenek dağılımları özellikle 100 ve 1000 kişilik örneklemelerde benzerlik göstermektedir. Tüm grupta ve 1000 kişilik örnekleme DE ve AH yöntemleri dağılımın uç kısmında bir grup tanımlarken diğer yöntemlerde yetenek düzeyinin 1'in üzerinde olduğu bir üst grup gözlenmemiştir.

Matematik dersine ilişkin elde edilen yetenek dağılımları Şekil 12'de sunulmuştur.



**Şekil 12. Matematik yetenek kestirimleri**

Şekil 12 incelendiğinde NBE kestirimlerinden elde edilen yetenek dağılımının diğerlerinden daha farklı ve sivri olduğu görülmektedir. Türkçe yetenek dağılımlarına benzer şekilde tüm grupta ve 1000 kişilik örnekleme AH ve DE kestirimlerinde yeteneği 2 üzerinde olan bir üst grup gözlenirken diğer yöntemlerde gözlenmemiştir. 100 kişilik örnekleme ise MAP, MLE, AH ve DE kestirimleri normal dağılıma benzemektedir.

## Bölüm 5

### Sonuç, Tartışma ve Öneriler

Araştırma iki temel problem üzerine kurulmuştur. Birincil olarak parametrik ve parametrik olmayan madde tepki kuramı modelleri farklı koşullar altında simülasyon verisi kullanılarak karşılaştırılmıştır. İkincil olarak ise gerçek veriler kullanılarak söz konusu karşılaştırmalar yapılmıştır. Araştırmanın sonuçları simülasyon bulguları ve gerçek veri bulguları birlikte ele alınarak tartışılmıştır.

Simülasyon çalışmasına ilişkin bulgular genel olarak incelendiğinde hem yetenek hem de madde parametre kestirim hatalarının kestirim yönteminden, madde sayısından, örneklem büyüklüğünden ve yetenek parametresinin dağılımından etkilendiği gözlenmektedir. Araştırma kapsamında beş farklı yöntem ele alınmıştır. Yöntemlerden elde edilen hataların miktarı örneklem büyüklüğüne, madde sayısına ve yetenek dağılımına bağlı olarak farklılık göstermektedir.

Yetenek kestirimleri dağılımın çarpıklığından etkilenmektedir. Ancak bu etki kestirim yöntemlerine göre farklılaşmaktadır. Yöntemler kendi içlerinde değerlendirildiğinde MLE ve DE yöntemleri dağılımın çarpıklığından çok fazla etkilenmezken, AH ve MAP yöntemleri çarpık dağılımlarda daha yüksek hatalı kestirimler üretmiştir. NBE yöntemi ise pozitif çarpık dağılımda daha yüksek hatalı kestirimler vermiştir. Yöntemler kıyaslandığında ise NBE genel olarak diğer yöntemlerden daha hatalı kestirimler üretmiştir. DE yöntemi ise daha düşük hatalı kestirimler yapmıştır. Özellikle çarpık yetenek dağılımlarında tüm koşullar için en hatasız yetenek kestirimleri DE ile elde edilmiştir. AH yöntemi ise genel itibarıyla MAP ile benzer sonuçlar vermiştir. Benzer şekilde Woods ve Lin (2009) yaptıkları simülasyon çalışmasında 25 madde ve 1000 örneklem büyüklüğü için yaptıkları simülasyon çalışmasında DE, Ramsey eğrileri ve AH yöntemlerinin çarpık dağılımlarda daha tutarlı sonuçlar verdiğini ulaştırmıştır. Reise ve diğerleri (2018) normal olmayan örtük özellikte aşamalı tepki modeli ile normallik varsayımı olmayan modellerin çarpık dağılımlarda normal örtük özelliğe dayalı modellerden daha iyi model uyumu gösterdiğini raporlamıştır. Köse ve Doğan (2019) farklı örneklem büyüklükleri, madde sayısı ve örtük özellik dağılımı koşullarında 3PLM model kestirimlerini karşılaştırmış ve çarpık dağılımlarda kestirim hatalarının daha yüksek olduğunu raporlamıştır. Diğer yandan Kirischi ve diğerleri (2001) araştırmalarında MULTILOG, XCALIBRATE ve BILOG programları ile yapılan kestirimleri, yeteneğin

boyutluluğu, boyutlar arası ilişkilerin büyüklüğü ve dağılımı bağlamında karşılaştırmışlardır. 1000 örneklem büyüklüğünün kullanıldığı çalışmada araştırmanın sonuçları örneklem yeterince büyük olduğunda yetenek dağılımının parametre kestirimlerine önemli bir etkisi olmadığını göstermiştir. Ancak Kiriscki ve diğerleri (2001) çalışmalarında yöntem karşılaştırması yapmamıştır. Bu çalışmada örneklem büyüklükleri ve madde sayıları küçük tutulduğundan yetenek parametresi için daha az etki büyüklüğüne sahip olmakla birlikte hem yetenek hem de madde parametreleri kestirimlerinde yetenek dağılımı etkili bir faktör olarak bulunmuştur. Sass, Schimitt ve Walker (2005) MAP, MLE ve ML yöntemlerini normal ve çarpık dağılımlarda karşılaştırdıkları çalışmalarında dağılımın çarpıklığının kestirim hatalarını etkilediğini raporlamıştır. Alan yazındaki genel yargı dağılım çarpıklaştıkça parametrik MLE ve MAP yöntemlerinin kestirim hatalarının arttığı yönündedir. Bu çalışmada da benzer bir bulgu ile MAP kestirimleri çarpıklıktan etkilenmiştir. Parametrik olmayan DE ve AH yöntemleri MLE ve MAP kestirimlerinden farklı olarak yetenek dağılımı için normal bir dağılım kullanmak yerine her bir iterasyonda yeteneğin dağılımını yeniden kestirir. Bu durum yetenek dağılımı çarpık olduğu durumda, bu iki yöntemin dağılımın çarpıklığına daha hassas olduğunu göstermektedir. Dolayısıyla normallik varsayımının olmadığı modeller çarpık örtük özelliklere daha iyi uyum sağlamaktadır. MAP kestirimleri ise normal bir önsel kullandığından dağılımın çarpıklığından etkilenmiştir.

Yetenek kestirimleri açısından madde sayısının örneklem büyüklüğüne göre kestirim hataları üzerinde daha etkili olduğu bulunmuştur. Madde sayısı arttıkça NBE dışındaki tüm kestirim yöntemlerinde hata miktarı azalmıştır. Reise ve diğerleri de (2018) madde sayısı az olduğunda kestirim yöntemlerinin yeteneğin dağılımını doğru şekilde yakalamalarının zorlaştığını belirtmiştir. Özellikle madde sayısının 10 olduğu koşulda ve örtük özellik dağılımı normal olduğunda en az hatalı kestirimler MAP ve AH yöntemleri ile elde edilmiştir. NBE ve MLE yöntemleri ise diğer yöntemlerden daha hatalı kestirimler üretmiştir. MAP yöntemi bayesçi bir yöntemdir ve önsel dağılım kullanır. Madde sayısı az olduğunda önsel bir dağılım kullanmak kestirimleri iyileştirmiş olabilir. Diğer yandan örtük özellik dağılımı çarpık olduğunda DE yöntemi MAP ve AH'den daha az hatalı kestirimler üretmiştir. Araştırma sonucundaki genel yargımız DE yönteminin çarpık dağılımlarda diğer yöntemlere göre daha iyi kestirim yaptığı yönündedir. Ancak DE yöntemi az sayıda madde ile

dağılım normal olduğunda MAP yönteminden daha hatalı kestirimler yapmıştır. DE yöntemi her bir iterasyonda örtük özellik dağılımını yeniden kestirdiğinden teorik olarak daha fazla sayıda maddeye ve örneklem büyüklüğüne ihtiyaç duymaktadır. Bu bağlamda dağılım normal olduğunda parametrik MAP yöntemi öne çıkarken, DE yönteminin dağılımın çarpıklığına olan hassasiyeti madde sayısı ve örneklem büyüklüğü az olsa da örtük özellik çarpık olduğunda daha iyi kestirimler üretmesini sağlamıştır.

Araştırma kapsamında b parametrelerinin dağılımının yetenek kestirimlerine etkisi de incelenmiştir. Söz konusu etkinin kestirim yöntemleri arasında farklılaştığı gözlenmiştir. NBE ve MLE yöntemleri ile elde edilen yetenek kestirimleri b parametresinin dağılımından etkilenirken diğer yöntemler etkilenmemiştir.

Araştırma kapsamına dahil edilen kestirim yöntemlerinden biri olan parametrik olmayan Bayesçi kestirim yöntemidir. Duncan ve MacEachern (2008) yöntemi iki parametrelili ve üç parametrelili MTK modelleri için düzenlemiş ve iki parametrelili modelin uygulanabilirliğini 258 kişilik 28 maddeden oluşan gerçek veri kümesinde göstermiştir. Finch ve Edwards (2015) ise parametrik olmayan bayesçi kestirim yöntemi ile RC-IRT ve normal dağılıma dayalı MTK kestirimlerini (MML, EAP, MAP) Rash model kullanarak karşılaştırmıştır. Araştırmanın sonuçlarına göre gizil değişken ve madde güçlükleri normal dağılmadığında en doğru kestirim sonuçları parametrik olmayan bayesçi kestirim yöntemi ile elde edilmiştir. Literatürde NBE yönteminin iki parametrelili model kestirimlerine ilişkin başka bir simülasyon çalışması bulunmamaktadır. Bu çalışmada ise NBE kestirimleri iki parametrelili model için diğer modellere kıyasla daha hatalı kestirimler üretmiştir. Modelin küçük örneklerde az sayıda madde ile madde parametre kestirimlerinde uygulanabilirliği savunulabilir ancak büyük örneklerde hata miktarı artmaktadır. Bu durum model kestirimlerinin birlikte olabilirlik şemasına bağlı kestirildiğinden kaynaklanmaktadır. Modeldeki parametre sayısı arttıkça kestirim hataları da artmaktadır. Ayrıca NBE kestirimleri çok sayıda iterasyon gerektirmekte ve analizler uzun sürede tamamlanabilmektedir. Bu bağlamda da büyük örneklerde kullanılması pratik değildir. Dolayısıyla Finch ve Edwards (2015)'in bulgularından farklı olarak NBE kestirimlerinin daha karmaşık modellerde yeterince iyi kestirimler yapmadığı sonucuna ulaşılmıştır.

Madde parametrelerine ilişkin sonuçlar incelendiğinde ise örneklem büyüklüğü 100 olduğunda MAP, MLE ve NBE kestirimlerinin a parametresinin kestirimlerinde benzer ve DE ve AH yöntemlerinden daha az hatalı sonuçlar verdiği gözlenmektedir. Örneklem büyüklüğünün 1000 ve 500 olduğu durumlarda ise örtük özellik normal dağıldığında MAP, MLE, DE ve AH yöntemleri arasında anlamlı bir fark gözlenmemektedir. Ancak çarpık dağılımlarda DE ve AH yöntemi diğerlerinden daha az hatalı kestirimler yapmıştır. Benzer şekilde b parametrelerinin kestirimlerinde de örneklem büyüklüğü 100 ve 500 olduğu durumlarda ve örtük özellik normal dağıldığında MAP, MLE, DE ve AH yöntemleri arasında anlamlı bir fark gözlenmemiş; örtük özellik çarpık olduğunda ise DE daha iyi kestirimler üretmiştir. Küçük örneklemde ise en iyi b parametresi kestirimleri MAP yöntemiyle elde edilmiştir. Dolayısıyla madde parametre kestirimlerinde örtük özellik normal olduğunda küçük örneklemde parametrik yöntemlerin daha iyi sonuç verdiği gözlenmektedir. Büyük örneklemde ise madde parametre kestirimleri örtük özellik normal dağıldığında yöntemler arasında farklılaşmazken, çarpık dağıldığında parametrik olmayan yöntemlerden AH ve DE daha iyi sonuç vermiştir.

Araştırmada ayrıca b parametresinin dağılımının madde parametre kestirimlerine etkisi de incelenmiştir. a parametrelerinin kestirimleri yalnızca madde sayısının 10 olduğu durumda, b parametresinin çarpıklığından etkilenmiştir. 10 maddelik testte a parametrelerinin kestirimlerinin en hatalı olduğu koşul b parametresinin pozitif çarpık olduğu koşuldur. Bu durum kestirim yöntemine göre farklılaşmamaktadır. b parametresi kestirimlerinde ise yetenek dağılımı normal olduğunda b dağılımının çarpıklığından yalnızca NBE yöntemi etkilenmiştir. NBE yöntemi en az hatalı kestirimlerini b dağılımı normal olduğunda yapmıştır. Yetenek dağılımı çarpık olduğunda ise b parametrelerinin kestirimleri b parametresinin dağılımından etkilenmiştir. Yetenek dağılımı negatif çarpık iken en hatalı b kestirimleri b dağılımın pozitif çarpık olduğu durumda elde edilmiş; yetenek dağılımı pozitif çarpık iken en yüksek hatalı b kestirimleri b parametresinin negatif çarpık olduğu durumlarda gözlenmiştir. Hem parametrik hem de parametrik olmayan modellerde b parametresinin dağılımı ile yetenek parametresinin dağılımı farklı olduğunda b parametrelerinin kestirimlerinden elde edilen hatalar artmıştır. Tüm yöntemlerde b parametreleri ile yetenek parametreleri aynı ölçekte tanımlanır.

Dolayısıyla iki parametrenin dağılımlarındaki farklılıklar parametre kestirimlerini etkilemektedir.

Gerçek veriler ile analizler yapılırken hem veri setinin tamamı kullanılmış hem de simülasyon koşullarına benzer şekilde 100, 500 ve 1000 kişilik rastgele örneklemeler seçilerek kestirimler yapılmıştır. Analiz sonucunda simülasyon bulgularına benzer şekilde tüm örneklem büyüklüklerinde en düşük BIC değerleri DE ve AH kestirimlerinden elde edilmiştir. Farklı kestirim yöntemlerinden elde edilen yetenek dağılımları incelendiğinde ise MAP ve MLE kestirimlerinden elde edilen dağılımların birbirlerine ve normal dağılıma benzer ancak parametrik olmayan yöntemlerden farklı olduğu gözlenmiştir. Duncan ve MacEachern (2008) istatistik dersine ilişkin sınav sonuçlarını parametrik ve parametrik olmayan iki parametrelilik bayesci modelleri karşılaştırdıkları araştırmalarında, yetenek kestirimlerinin parametrik ve nonparametrik modeller arasında farklılaştığını raporlamıştır.

### **Öneriler**

Araştırma kapsamında öneriler sırasıyla uygulamaya yönelik öneriler ve araştırmacılara yönelik öneriler olmak üzere iki kısımda sunulmuştur.

#### **Uygulamaya yönelik öneriler.**

1. Araştırmanın bulguları MTK'ya dayalı parametre kestirimlerini örtük özellik dağılımının çarpıklığının etkilediğini göstermektedir. Çarpık dağılımlarda en iyi kestirimler parametrik olmayan DE ve AH yöntemleriyle elde edilmiştir. Analizler sırasında DE kestirimleri AH kestirimlerinden daha kısa sürede elde edilmiştir. Bu bağlamda örneklem büyüklüğü 500 ve üzerinde olduğu durumlarda klasik kestirim yöntemlerine göre DE yöntemi hem dağılımın çarpıklığından daha az etkilendiği hem de örtük özellik normal dağıldığı durumlarda klasik yöntemlerle aynı sonuçları verdiği için uygulamada tercih edilmelidir.
2. Küçük örneklemelerle çalışırken örtük özellik normal dağıldığında, parametrik MAP yöntemi daha düşük kestirim hatası ürettiğinden tercih edilebilir. Örtük özellik normal dağılmadığında ise parametrik olmayan DE yöntemi tercih edilmelidir.



3. Madde sayısı az olduğunda parametre kestirimleri örneklem büyüklüğünden daha çok madde sayısından etkilenmektedir. Az sayıda madde gerektiren uygulamalarda madde sayısına uygun kestirim yöntemi seçilmelidir. Yetenek kestirimlerinde madde sayısı az olduğuna en iyi kestirimleri DE yöntemi verirken, madde parametre kestirimlerinde en iyi kestirimler MAP ile elde edilmiştir.
4. Araştırmanın sonuçları örtük özellik normal dağılmadığında b parametresinin dağılımının madde parametre kestirimlerini etkilediğini göstermektedir. Bu bağlamda b parametrelerinin dağılımlarının normal olmamasının beklendiği tanılama ve seçme sınavları gibi sınavlarda madde parametrelerinin kestirileceği grubun örtük özellik dağılımı dikkate alınmalıdır. b parameterelerinin ve örtük özelliğin dağılımının çarpıklığının yönü farklı olduğu durumlarda madde parametre kestirimlerinde MAP kullanılabilir.
5. Araştırma kapsamında gerçek verilerden elde edilen bulgular, örtük özellik normal dağılmadığında kestirilen yetenek parametrelerinin dağılımının parametrik olan ve olmayan yöntemler arasında farklılaştığını göstermektedir. Genel olarak parametrik yöntemler dağılımı normale zorladığından üst ve alt grupta yer alan bireylerde daha fazla hata yapılmaktadır. Bu nedenle uygulamada özellikle üst grupların ayrışması beklenen seçme ve yerleştirme sınavları ile alt grupların ayrışması beklenen tanıma veya sertifika sınavlarında uygulayıcılar kestirim yöntemlerinin sınırlılıklarının farkında olmalıdır. Bu bağlamda veri yapısına ve örtük özellik dağılımına uygun kestirim yöntemi seçmelidirler.

#### **Araştırmaya yönelik öneriler.**

1. Araştırma kapsamında parametrik MTK'ya dayalı kestirim yöntemlerinden MLE ve MAP ve parametrik olmayan MTK yöntemlerinden DE, AH ve NBE yöntemleri karşılaştırılmıştır. Araştırma kapsamında örtük özellik dağılımı normal, pozitif çarpık ve negatif çarpık olacak şekilde manipüle edilmiştir. Gelecek araştırmalarda tek biçimli veya karmaşık normal dağılım gibi farklı örtük özellik dağılımlarında kestirim yöntemlerinin performansları karşılaştırılabilir.

2. Arařtırma kapsamında ulařılan en önemli sonuçlardan birisi daęılımın arpıklılıęının kestirim hatalarını etkiledięidir. Gelecek arařtırmalarda söz konusu etkinin uygulamadaki yansımaları incelenebilir. Örneęin sertifika sınavları gibi geti-kaldı kararı verilen sınavlarda, seme-yerleřtirme sınavlarında veya tanıma amalı yapılan ölçmelerde arpıklıktan kaynaklanan hataların elde edilen sınıflama tutarlılıklarına etkileri incelenebilir.
3. Ayrıca bilgisayar tabanlı uygulamalarda madde parametrelerinin hesaplandıęı gruplarda örneklem her zaman rastgele belirlenmeyebilir veya deęiřkenin doęası gereęi normal daęılmayabilir. Bu gibi durumlarda arpıklıktan kaynaklanan hataların bireyselleřtirilmiř bilgisayar tabanlı uygulamalara (BBT) etkisi incelenebilir.
4. Benzer řekilde söz konusu yanlılıęın eřitmele ve deęiřen madde fonksiyonu (DIF) belirleme süreçlerine etkisi arařtırılabilir.
5. Parametrik ve parametrik olmayan yöntemler boyutluluk, yerel baęımsızlık, tek biçimli örtük özellik daęılımını baęlamında kurgulanan farklı simülasyon alıřmaları ile incelenebilir.
6. Parametrik ve parametrik olmayan yöntemler testte modele tam uyum saęlamayan maddeler olması durumunda kurgulanan farklı simülasyon alıřmaları ile incelenebilir.
7. Arařtırma kapsamına dahil edilen yöntemlerden birisi parametrik olmayan bayeřçi kestirim yöntemidir. Yöntem her ne kadar büyük örneklemle beklenen sonucu vermese de yöntemin performansı ileriki arařtırmalarda küçük örneklem ve daha az sayıda madde ieren sınıf ii uygulamalarda denenerek klasik test istatistikleri ile kullanılılılık bakımından kıyaslanabilir.

## Kaynaklar

- Avsar, A. S., & Tavsancil, E. (2017). Examination of Polytomous Items' Psychometric Properties According to Nonparametric Item Response Theory Models in Different Test Conditions. *Educational Sciences: Theory & Practice*, 17(2).
- Arenson, E, & Karabatsos, G. A Bayesian Beta-Mixture Model for Nonparametric IRT (BBM-IRT) (September 5, 2017). *Journal of Modern Applied Statistical Methods*, Forthcoming. Available at SSRN: <https://ssrn.com/abstract=3102461>
- Ayala, R. J. (2009). *The Theory and Practice of Item Response Theory*. New York: The Guilford Press.
- Baykul, Y. (2015). *Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması*. Ankara: Pegem Akademi.
- Bıkmaz Bilgen, Ö. (2016). Parametrik ve Parametrik Olmayan Madde Tepki Kuramı Modellerinin Farklı Örneklemeler ve Test Uzunluğunda Karşılaştırılması.
- Bıkmaz Bilgen, Ö., & Doğan, N. (2017). Çok Kategorili Parametrik ve Parametrik Olmayan Madde Tepki Kuramı Modellerinin Karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 8(4), 354-372.
- Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 46, 443-459.
- Bulut, O., & Sünbül, Ö. (2017). Monte Carlo Simulation Studies in Item Response Theory with the R. *Journal of Measurement and Evaluation in Education and Psychology*, 8(3), 266-287. doi:10.21031/epod.305821
- Chalmers, P. (2020). mirt: Multidimensional Item Response Theory. [Online: <https://cran.r-project.org/web/packages/mirt/mirt.pdf> Accessed date: 17.6.2018.]
- Crocker, L., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. Orlando: Holt, Rinehart and Winston Inc.
- Dooley, K. (2002). Simulation research methods. In J. Baum (Ed.). *Companion to organizations*. London: Blackwell.
- Douglas, J., & Cohen, A. (2001). Nonparametric item response function estimation for assessing parametric model fit. *Applied Psychological Measurement*, 25(3), 234-243.
- Duncan, K. A., & MacEachern, S. N. (2008). Nonparametric Bayesian modelling for item response. *Statistical Modelling*, 8(1), 41-66.
- Duncan, K. A., & MacEachern, S. N. (2013). Nonparametric Bayesian modeling of item response curves with a three-parameter logistic prior mean. *Current topics in the theory and application of latent variable models*, 108-125.
- Ferguson, T. S. (1974). Prior distributions on spaces of probability measures. *The annals of statistics*, 615-629.
- Finch, H., & Edwards, J. M. (2016). Rasch model parameter estimation in the presence of a nonnormal latent trait using a nonparametric Bayesian approach. *Educational and Psychological Measurement*, 76(4), 662-684.

- Franco, V. R., & Wiberg, M. (2020, March 17). Nonparametric Item Response Models: A Comparison on Recovering True Scores. <https://doi.org/10.31234/osf.io/x7u9j>
- Gilks, W. R., Best, N. G., ve Tan, K. K. C. (1995). Adaptive rejection metropolis sampling within gibbs sampling. *Appl. Statist.*, *44*(4), 455-472.
- Hambleton, R. K., & Swaminathan, H. (1991). *Fundamentals of Item Response Theory: Principles and Applications*. California: SAGE.
- Hambleton, R. K., & Swaminathan, H. (2013). *Item response theory: Principles and applications*: Springer Science & Business Media.
- Hu, G., Ma, Z., & Paek, I. (2020). A Nonparametric Bayesian Item Response Modeling Approach for Clustering Items and Individuals Simultaneously. arXiv preprint arXiv:2006.00105.
- Junker, B. W., & Sijtsma, K. (2001). Nonparametric item response theory in action: An overview of the special issue. *Applied Psychological Measurement*, *25*(3), 211-220.
- Kirisci, L., & Hsu, T.-C. (1995). The Robustness of BILOG to Violations of the Assumptions of Unidimensionality of Test Items and Normality of Ability Distribution.
- Kirisci, L., Hsu, T.-c., & Yu, L. (2001). Robustness of item parameter estimation programs to assumptions of unidimensionality and normality. *Applied Psychological Measurement*, *25*(2), 146-162.
- Köse, A ve Doğan, D. (2019). Parameter estimation bias of dichotomous logistic item response theory models using different variables. *International Journal of Evaluation and Research in Education*, *8*(3), 425-433. DOI: 10.11591/ijere.v8i3.19807
- Liang, L. (2007). A semi-parametric approach to estimating item response functions. The Ohio State University,
- Lord, F. (1952). A theory of test scores. Psychometric monographs.
- McDonald, R. P. (1982). Linear versus models in item response theory. *Applied Psychological Measurement*, *6*(4), 379-396.
- Mor Dirlik, E. ve Koç, N. (2019) parametrik ve parametrik olmayan madde tepki kuramı modellerinden çeşitli faktörlere göre elde edilen madde ve yetenek kestirimlerinin karşılaştırılması. *İlkoğretim Online*, *18*(3),1016-1035. doi: 10.17051/ilkonline.2019.610162
- Partchev, I. (2017). irtoys: A Collection of Functions Related to Item Response Theory (IRT). [Online: <https://cran.r-project.org/web/packages/irtoys/irtoys.pdf> Accessed date: 17.5.2018.]
- Patz, R. J., & Junker, B. R. (1999). A straightforward approach to markov chain monte carlo methods for item response models. *Journal of Educational and Behavioral Statistics*, *24*(2),146-178.
- R Core Team. (2013). R: A language and environment for statistical computing, (Version 3.0.1), Vienna, Austria: R Foundation for Statistical Computing. Online: <http://www.R-project.org/>

- Reckase, M. (2009). *Multidimensional item response theory* (Vol. 150): Springer.
- Reise, S. P., & Revicki, D. A. (2014). *Handbook of item response theory modeling: Applications to typical performance assessment*: Routledge.
- Reise, S. P., Rodriguez, A., Spritzer, K. L., & Hays, R. (2018). Alternative approaches to addressing non-normal distributions in the application of IRT models to personality measures. *Journal of Personality Assessment*, 100(4), 363-374, DOI: 10.1080/00223891.2017.1381969
- Reyhanlioğlu Keçeoğlu, Ç. (2018) *Parametrik ve parametrik olmayan madde tepki kuramında farklı örneklem büyüklüklerine ve boyutluluklarına göre parametre değişmezliğinin incelenmesi*.(Doktora tezi). Hacettepe Üniversitesi, Ankara
- San Martín, E., Jara, A., Rolin, J.-M., & Mouchart, M. (2011). On the Bayesian nonparametric generalization of IRT-type models. *Psychometrika*, 76(3), 385-409.
- Sarkar, D. (2020). lattice: Trellis Graphics for R [Online: <https://cran.r-project.org/web/packages/lattice/lattice.pdf> Accessed date: 17.01.2020.]
- Sass, D., Schmitt, T., & Walker, C. (2008). Estimating non-normal latent trait distributions within item response theory using true and estimated item parameters. *Applied Measurement in Education*, 21(1), 65-88.
- Seong, T.-J. (1990). Sensitivity of marginal maximum likelihood estimation of item and ability parameters to the characteristics of the prior ability distributions. *Applied Psychological Measurement*, 14(3), 299-311.
- Sijtsma, K. (2005). Nonparametric item response theory models. *Encyclopedia of Statistics in Behavioral Science*.
- Sijtsma, K., Emons, W. H., Bouwmeester, S., Nyklíček, I., & Roorda, L. D. (2008). Nonparametric IRT analysis of quality-of-life scales and its application to the world health organization quality-of-life scale (WHOQOL-Bref). *Quality of Life Research*, 17(2), 275-290.
- Sijtsma, K., & Meijer, R. R. (2006). 22 Nonparametric Item Response Theory and Special Topics. *Handbook of Statistics*, 26, 719-746.
- Smits, N., Öğreden, O., Garnier-Villarreal, M., Terwee, C. B., & Chalmers, R. P. (2020). A study of alternative approaches to non-normal latent trait distributions in item response theory models used for health outcome measurement. *Statistical Methods in Medical Research*, 29(4), 1030–1048. DOI:10.1177/0962280220907625
- Stone, C. A. (1992). Recovery of marginal maximum likelihood estimates in the two-parameter logistic response model: An evaluation of MULTILOG. *Applied Psychological Measurement*, 16(1), 1-16.
- Stout, W. (2001). Nonparametric item response theory: A maturing and applicable measurement modeling approach. *Applied Psychological Measurement*, 25(3), 300-306.
- Vermunt, J. K. (2001). The use of restricted latent class models for defining and testing nonparametric and parametric item response theory models. *Applied Psychological Measurement*, 25(3), 283-294.

- Woods, C. M. (2007). Empirical histograms in item response theory with ordinal data. *Educational and Psychological Measurement*, 67(1), 73-87.
- Woods, C. M. (2008). Ramsay-curve item response theory for the three-parameter logistic item response model. *Applied Psychological Measurement*, 32(6), 447-465.
- Woods, C. M. (2014). Estimating the latent density in unidimensional IRT to permit non-normality. In *Handbook of Item Response Theory Modeling* (pp. 78-102): Routledge.
- Woods, C. M., & Lin, N. (2009). Item response theory with estimation of the latent density using Davidian curves. *Applied Psychological Measurement*, 33(2), 102-117.
- Woods, C.M., & Thissen, D. (2006). Item response theory with estimation of the latent population distribution using spline-based densities. *Psychometrika*, 71(2), 281-301.
- Zhang, D., & Davidian, M. (2001). Linear mixed models with flexible distributions of random effects for longitudinal data. *Biometrics*, 57, 795-802.

## EK-A: Madde Parametrelerinin ve Cevap Matrisinin Üretilmesi Örnek R kodu

Aşağıda tanımlanan kodda a ayırıcılık parametresi vektörüne, b güçlük parametresi vektörüne, ltrait yetenek parametresi vektörüne karşılık gelmektedir.

```
library(moments)

###30 maddelik test için madde parametrelerinin üretilmesi
## b parametreleri
#Normal dağılılan b parametreleri

Test30Nb=rep(0,30)
repeat{y=rnorm(30)
s1=skewness(y)
m1=mean(y)
sd1=sd(y)
if (s1>-0.005 & s1<0.005 & m1>-0.005 & m1<0.005 & sd1< 1.05 & sd1>0.95){
  Test30Nb<-y
  break
}}

#Pozitif çarpık b parametresi
Test30PCb=rep(0,30)
repeat{y=scale(rgamma(30,shape=1.5,scale = 1.5),center = T)
s2=skewness(y)
if (s2>1.55 & s2<1.65){
  Test30PCb<-y
  break
}}

#Negatif çarpık b parametresi
Test30NCb=-1*Test30PCb

## a parametreleri
a30 <- as.vector(round(rlnorm(30, meanlog = 0.3, sdlog = 0.2),3))

###100 örneklem büyüklüğü için yetenek parametrelerinin üretilmesi
#Normal
Nyetenek<-rnorm(100)
#Pozitif Çarpık
PCyetenek<-as.vector(scale(rgamma(100,shape=1.2,scale = 1.5),center = T))

#Negatif Çarpık
NCYetenek<-1*PCyetenek

###Yeteneği normal 100 kişi ve 30 maddelik test için madde cevap matrisinin üretilmesi
```

```

#2PL model için madde cevap matrisi oluşturan fonksiyon
#a: ayırıcılık, b: güçlük, ltrait: yetenek,
response2pl <- function(a,b,ltrait)
{ k <- length(b)
n <- length(ltrait)
theta <- ltrait
aa <- rep(a, rep(n, k))
bb <- rep(b, rep(n, k))
p <- 1/(1+exp(-((aa)*(theta-bb))))
rr <- runif(n*k, 0, 1)
puan <- ceiling(p-rr)
datar <- matrix(puan, ncol=k)
return(datar)
}
a=a30
b=Test30Nb
ltrait=Nyetenek
rep1<-response2pl(a,b,ltrait)

```

### EK-B: MLE, MAP, DE ve AH Kestirimleri Örnek R kodu

MLE ve MAP kestirimleri için "irtoys" paketi kullanılmıştır. Paket fonksiyonunun kullanılabilmesi için icl programının çalışma dizininde olması gerekir. DE ve AH yöntemleri için "mirt" paketi kullanılmıştır.

```

##Çalışma dizininin tanımlanması
setwd("C:\\Users\\nermin\\Desktop\\tez")

##Analizi hızlandırmak için kullanılan çekirdek sayısının artırılması
library(doParallel)

cores=detectCores()
registerDoParallel(cores-1)

##Verinin okunması
veri<-read.csv("rep1.csv",header = T)

##MLE ve MAP kestirimleri
library(irtoys)

#MLE
#MLE için herhangi bir önsel kullanılmamıştır

Madde.Parametreleri<-est(veri, model = "2PL", engine = "icl",bilog.defaults = T)
Yetenek<-ability(resp=veri, Madde.Parametreleri, method = "MLE")

#MAP kestirimleri
#MAP için a ve b parametreleri için önsel dağılım tanımlanmıştır
Madde.Parametreleri<-est(veri, model = "2PL", engine = "icl",bilog.defaults = T,a.p
rior = T,b.prior = T)

```



```

Yetenek<-ability(resp=veri, Madde.Parametreleri, method = "BME")

##DE ve AH kestirimi
library(mirt)

TL=ncol(veri)
mod.2pl <- paste("F = 1 -" , TL, sep = ' ')

#DE kestirimi
temp=mirt(veri, model=mod.2pl, itemtype = "2PL",method = "EM",dentype="Davidian-4")

Madde.Parametreleri=coef(temp, IRTpars = TRUE, simplify = TRUE)$items
Yetenek=fscores(temp,use_dentype_estimate=TRUE, theta_lim = c(-6, 6))

#AH kestirimi
temp2=mirt(veri, model=mod.2pl, itemtype = "2PL",method = "EM",dentype="AH")

Madde.Parametreleri=coef(temp2, IRTpars = TRUE, simplify = TRUE)$items
Yetenek=fscores(temp2,use_dentype_estimate=TRUE, theta_lim = c(-6, 6))

```

### EK-C: NBE Kestirimi R Kaynak Kodu

```

##Baslangic degerlerinin tanimlanmasi
#Yetenek
init.Theta<-function(x){
  SS= apply(x,1,sum)
  NN=nrow(x)
  theta= matrix(qnorm(rank(SS)/(NN+1)),ncol=1)
}

#b parametresi
init.b<-function(x){
  SS.b= apply(x,2,mean)
  TL=ncol(x)
  b.par= matrix(qnorm(rank((1-SS.b)/(TL+1)),nrow=1)
}

#a parametresi
init.a=function(x){
  TL=ncol(x)
  a.par= matrix(round(rlnorm(TL, meanlog = 0.3, sdlog = 0.2),3),nrow=1)
}

#z parametresi icin
#z icin yogunluk fonksiyonu

f0<-function(a,b,theta){
  y=1/(1+exp(-a*theta-b))
return(y)
}

```

```

#z parametresi için örnekleme fonksiyonları
#Dođru cevap için
SampleZ0<-function(a,b,theta){
  if (theta<9){
    quad=seq(theta,10,0.2)
    wgt=f0(a,b,quad)
    wgt=wgt/ifelse(f0(a,b,theta)==0,0.001,f0(a,b,theta))
    wgt=wgt/sum(wgt)
    z=sample(quad,1,prob = wgt)
  }else{
    z=9
  }
}

return(z)}

#Yanlıp cevap için
SampleZ1<-function(a,b,theta){
  if(theta>-10){
    quad=seq(-10,theta,0.2)
    wgt=f0(a,b,quad)
    wgt=wgt/ifelse(f0(a,b,theta)==1,0.001,(1-f0(a,b,theta)))
    wgt=wgt/sum(wgt)
    z=sample(quad,1,prob = wgt)
  }else{
    z=-10
  }
}

return(z)}

#z parametrelerinin bađlangıç deđerlerinin belirlenmesi
init.z<- function (data,theta,a.par,b.par){
  TL=ncol(data)
  NN=nrow(theta)
  #F0 olasiliklari için parametrelerin duzenlenmesi
  tet=matrix(rep(theta,TL),ncol = TL)
  par.a=matrix(rep(a.par,NN),ncol = TL,byrow = T)
  par.b=matrix(rep(b.par,NN),ncol = TL,byrow = T)
  #F0 olasiliklerinin hesaplanması
  logit= tet*par.a + par.b
  F0.probs <- 1/(1 + exp(-logit))

  #z deđerleri için bos matris
  ZZ=matrix(0, nrow=NN, ncol = TL)

  #her bir madde ve birey için z dongusu
  for (item in 1: TL){
    a=a.par[1,item] #bu parametreler z density için
    b=b.par[1,item]
    for (ind in 1: NN){

```

```

teta=theta[ind,1]

if(data[ind,item]==1){
  zz=SampleZ1(a,b,teta)/F0.probs[ind,item] #maddeye dogru cevap verirse burad
an orneklenir
}else{
  zz=SampleZ0(a,b,teta)/(1-F0.probs[ind,item]) #yanlis cevap vermis ise buradan o
rneklenir
}
ZZ[ind,item]=zz
}}
return(ZZ)
}

#####Baslangic degerlerinin tanimlandigi tek bir fonksiyon
####Fonsiyon dort cikti uretir
####theta, a.par, b.par ve z.par
Initial.Values<- function(data){
  theta<-init.Theta(data)
  a.par<-init.a(data)
  b.par<-init.b(data)
  z.par<-init.z(data=data,theta=theta,a.par=a.par,b.par=b.par)
  initials=list(theta=theta,a.par=a.par,b.par=b.par,z.par=z.par)
  return(initials)
}

#####
##Iterasyon fonksiyonlarinin tanimlanmasi

#Gerekli paketler
library(truncnorm)
library(dplyr)
library(tibble)

#yetenek icin minimum ve maksimum degerlerin belirlenmesi
Limit.Fun<-function(x,y){
  temp=t(rbind(x,y))
  colnames(temp)=c('V1','V2')
  temp=as_tibble(temp)
  tt1=temp%>%filter(V1==1)
  tt2=temp%>%filter(V1==0)
  Min=ifelse(nrow(tt1)==0,-6,max(tt1[,2]))
  Max=ifelse(nrow(tt2)==0,6,min(tt2[,2]))
  result=c(Min,Max)
  return(result)
}

#yetenegin guncelenmesi
Iterate.Theta<-function(data,z.par){
  ind=nrow(data)
  for (i in 1: ind){

```

```

limits=Limit.Fun(data[i,],z.par[i,])
Min= limits[1]
Max=limits[2]
tet1=ifelse(Min<Max,rtruncnorm(1, a=Min, b=Max, mean = 0, sd = 1),Min)
theta[i]=round(tet1,3)
result=matrix(theta,ncol=1)
}
return(result)
}

```

*#z parametrelerinin guncellenmesi*

```

Iterate.Z<-function(data,theta,a.par,b.par,z.par,M){
  TL=ncol(data)
  NN=nrow(theta)
  #F0 olasiliklari icin parametrelerin duzenlenmesi
  tet=matrix(rep(theta,TL),ncol = TL)
  par.a=matrix(rep(a.par,NN),ncol = TL,byrow = T)
  par.b=matrix(rep(b.par,NN),ncol = TL,byrow = T)
  #F0 olasiliklerinin hesaplanmasi
  logit= tet*par.a + par.b
  F0.probs <- 1/(1 + exp(-logit))

```

*#z degerleri icin bos matris*

```

ZZ=matrix(0, nrow=NN, ncol = TL)

```

*#her bir madde ve birey icin z dongusu*

```

for (item in 1: TL){
  apar=a.par[1,item] #bu parametreler z density icin
  bpar=b.par[1,item]
  for (ind in 1: NN){
    teta=theta[ind,1]
    ZL=z.par[-ind,item]
    if(data[ind,item]==1){
      prop1=round(M*F0.probs[ind,item]) #Olasiliklerin tanimlanmasi
      prb1=prop1/(prop1+1)
      p=runif(1,min=0,max=1)
      zz= ifelse(p<prb1,SampleZ1(apar,bpar,teta),sample(ZL,1))
      zz=round(zz,3)
      #maddeye dogru cevap verirse buradan orneklenir
    }else{
      prop1=round(M*(1-F0.probs[ind,item]))
      prb1=prop1/(prop1+1)

      p=runif(1,min=0,max=1)
      zz= ifelse(p<prb1,SampleZ0(apar,bpar,teta),sample(ZL,1))
      zz=round(zz,3)
    }
    ZZ[ind,item]=zz
  }
}
}

```

```

return(ZZ)
}

##yer parametrelerinin guncellenmesi
#x cevap matrisi her bir madde icin (1xind kadar)
#y theta matrisi (1xind kadar)
#z tolerans matrisi her bir madde icin (1xind kadar)
#a maddenin ayirt ediciligi (1x1), b guclugu (1x1)

Group.Fun<-function(z.par){
  #bireylerin clusterlarının tanımlanması
  TL=ncol(z.par)
  N=nrow(z.par)
  group <- matrix(0,nrow=N,ncol=TL)

  for(t in 1:TL){

    z=z.par[,t]
    MaxZ=max(z,na.rm = T)
    MinZ=min(z,na.rm = T)
    diff.gp=0.05
    gp.values<-seq(from=MinZ,to=MaxZ,by=diff.gp)
    n.gps<-length(gp.values)
    gp.bd <- matrix(-999,nrow=(n.gps-1),ncol=1)

    #grup sinirlarının tanımlanması
    for (g in 1:(n.gps-1)){

      gp.bd[g] <- gp.values[g+1] - diff.gp/2
    }

    for (j in 1:N){
      for (g in (n.gps-1):1){

        group[j,t] <- ifelse(z[j] < gp.bd[g], g, group[j,t])

      } } }
    return(group)
  }

  #a ve b parametrelerinin tanımlanması
  lterate.ab<-function(data,a.par,b.par,z.par,theta){
    #grup degiskenine gore verinin bolunmesi
    TL=ncol(data)
    b1.par=b.par+rnorm(TL,sd=0.0025)
    a1.par=exp(log(a.par)+ rnorm(TL,sd=0.0025))
    Group=Group.Fun(z.par)
    PLoc0=rep(0,TL)
    PLoc1a=rep(0,TL)
    PLoc1b=rep(0,TL)
  }
}

```

```

for(k in 1:TL){
  x=data[,k]
  y=theta
  group=Group[,k]
  temp<-cbind(x,y)
  colnames(temp)<-paste("V",1:2,sep = ")
  temp=as_tibble(temp)
  splited.temp=split(temp,group)
  #dogru ve yanlis cevap verenlerin bolunmesi
  tt1=lapply(splited.temp,filter,V1==1)
  tt0=lapply(splited.temp,filter,V1==0)
  #grup sayisina gore min ve max vektorlerinin uzunluklarinin belirlenmesi
  ind=length(splited.temp)
  minVals=rep(0,ind)
  maxVals=rep(0,ind)
  Loc0=rep(0,ind)
  Loc1a=rep(0,ind)
  Loc1b=rep(0,ind)
  a0=a.par[1,k]
  b0=a.par[1,k]
  a1=a1.par[1,k]
  b1=b1.par[1,k]
  #min ve max sinir degerlerini belirleyen dongu
  for (l in 1: ind){

    dat1=tt1[[l]]
    maxVals[l]=ifelse (nrow(dat1)>0,min(dat1[,2]),6)
    dat0=tt0[[l]]
    minVals[l]=ifelse(nrow(dat0)>0,max(dat0[,2]),-6)
  }

  for (i in 1:ind){
    MinL=minVals[i]
    MaxL=maxVals[i]
    r= ifelse(MinL==MaxL,MinL,runif(1,min=min(MinL,MaxL),max=max(MinL,Max
L)))
    r1b=dlogis(r,scale=a0,location = b1)
    r1a=dlogis(r,scale=a1,location = b0)
    r0=dlogis(r,scale=a0,location = b0)

    Loc0[i]=r0
    Loc1a[i]=r1a
    Loc1b[i]=r1b
  }
  PLoc0[k]=sum(log(Loc0))
  PLoc1a[k]=sum(log(Loc1a))
  PLoc1b[k]=sum(log(Loc1b))
}
LL0b=PLoc0+sum(log(dnorm(b.par,mean = 0,sd=2)))
LL0a= PLoc0+sum(log(dgamma(a.par,shape = 3.2,scale=1)))

```

```

LL1b=PLoc1b+sum(log(dnorm(b1.par,mean = 0,sd=2)))
LL1a=PLoc1b +sum(log(dgamma(a1.par,shape = 3.2,scale=1)))

acceptb=LL1b-LL0b
acceptb=ifelse(acceptb>0,0,acceptb)
acceptb=ifelse(runif(TL)<exp(acceptb),1,0)
b.par <- ifelse(acceptb==1,b1.par,b.par)

accepta=LL1a-LL0a
accepta=ifelse(accepta>0,0,accepta)
accepta=ifelse(runif(TL)<exp(accepta),1,0)
a.par <- ifelse(accepta==1,a1.par,a.par)

ab.par=rbind(a.par,b.par)
return(ab.par)
}

####data: veri seti
####initials: baslangic degerlerinin hesaplanip hesaplanmadigi kisim
#True degeri aldiginda inits argumaninla liste seklinda a, b,theta ve z degerlerinin
#baslangic degerleri tanimlanmalidir
#### M DP surecine iliskin onse kutle
#### BurnIn burn-in iterasyon sayisi
#### nlters Iterasyon sayisi
#### nThin inceltme katsayisi

NB2PL<-function(data,initials=F,inits=NULL,M=5,BurnIn=1000,nlters=1000,nThin
=10){
items=ncol(data)
inds=nrow(data)
if (initials){
Initials<-inits
theta=Initials$theta
a.par=Initials$a.par
b.par=Initials$b.par
z.par=Initials$z.par
}else{
Initials<-Initial.Values(data)
theta=Initials$theta
a.par=Initials$a.par
b.par=Initials$b.par
z.par=Initials$z.par
}
}

for(b in 1:BurnIn){

theta<-Iterate.Theta(data,z.par)
z.par<-Iterate.Z(data,theta,a.par,b.par,z.par,M)
ab.par<-Iterate.ab(data,a.par,b.par,z.par,theta)

```

```

a.par<-matrix(ab.par[1,],nrow=1)
b.par<-matrix(ab.par[2,],nrow=1)
}

write("", file="storea.txt")
write("", file="storeb.txt")
write("", file="storetheta.txt")
print('iteration started')
for(g in 1:nIters){
  theta<-Iterate.Theta(data,z.par)
  z.par<-Iterate.Z(data,theta,a.par,b.par,z.par,M)
  ab.par<-Iterate.ab(data,a.par,b.par,z.par,theta)
  a.par<-matrix(ab.par[1,],nrow=1)
  b.par<-matrix(ab.par[2,],nrow=1)
  gt<-g/(nThin)
  if(gt==round(g/(nThin),0)){
    print(gt)
    write(a.par,"storea.txt",append=T,ncol=ncol(data))
    write(b.par,"storeb.txt",append=T,ncol=ncol(data))
    write(theta,"storetheta.txt",append=T,ncol=nrow(data))
  }
}
storea<-t(matrix(data=scan("storea.txt"),nrow=ncol(data)))
storeb<-t(matrix(data=scan("storeb.txt"),nrow=ncol(data)))
storeTheta<-t(matrix(data=scan("storetheta.txt"),nrow=nrow(data)))

NPoutput<-list( storea, storeb,storeTheta)
names(NPoutput)<-c("a", "b","theta")
return(NPoutput)
}

```

#### EK-D: NBE Kestirimi MATLAB kodu

```

function [aEst, bEst, thetaEst] = NB2PL(data, nIter, nThin)
  theta = Initialize_Theta(data);
  aPar = Initialize_A(size(data, 2), 0.3, 0.2);
  bPar = Initialize_B(data);
  zPar = Initialize_Z(data, theta, aPar, bPar);
  M = 5;
  nBurnIn = 10000;
  % tic
  for iBurn = 1:nBurnIn
    theta = IterateTheta(data, zPar);
    zPar = IterateZ(data, zPar, aPar, bPar, theta, M);
    [aPar, bPar] = Iterate_AB(data, theta, zPar, aPar, bPar);
  end
end

```



```

% toc
N      = floor(nIter/nThin);
aEst   = zeros(size(aPar));
bEst   = zeros(size(bPar));
thetaEst = zeros(size(theta));
for ilter = 1:nIter
    theta    = IterateTheta(data, zPar);
    zPar     = IterateZ(data, zPar, aPar, bPar, theta, M);
    [aPar, bPar] = Location(data, theta, zPar, aPar, bPar);
    if (mod(ilter, nThin) == 0)
        aEst  = aEst + aPar/N;
        bEst  = bEst + bPar/N;
        thetaEst = thetaEst + theta/N;
    end
end
thetaEst = thetaEst.';
end

function p = fProb(theta, aPar, bPar)
    p = 1 ./ (1 + exp(-(theta - bPar) .* aPar * 1.7));
end

function x = randSample(v, n, p)
    pc = cumsum(p) / sum(p);
    r = rand(1,n);
    x = zeros(1,n);
    for i = length(pc):-1:1
        x(r<pc(i)) = v(i);
    end
end

function z0 = SampleZ0(aPar, bPar, theta)
    z0 = zeros(length(theta), length(aPar));
    rng(1);
    rMtx = randn(length(theta), length(aPar));
    z0 = rMtx .* aPar + bPar;
    thetaMtx = theta * ones(1, length(aPar));
    idx = z0 < thetaMtx;
    z0(idx) = thetaMtx(idx) + 0.05;
end

```

```

idx1 = theta < 6;
z0x(idx1, :) = theta(idx1);
idx2 = rMtx > 6;
for ii = 1:length(aPar)
pd = makedist("Logistic", "mu", bPar(ii), "sigma", aPar(ii));
for iTheta = 1:length(theta)
    if (theta(iTheta) < 6)
        rVal = aPar(ii) * randn(1) + bPar(ii);
        if (rVal > 6)
            z0(iTheta, ii) = 6;
        else
            z0(iTheta, ii) = rVal;
        end
        z0(iTheta, ii) = ...
random(truncate(pd, theta(iTheta), 6));
    else
z0(iTheta, ii) = theta(iTheta);
end
end
end
endfunction z1 = SampleZ1(aPar, bPar, theta)
rMtx = randn(length(theta), length(aPar));
z1 = rMtx .* aPar + bPar;
thetaMtx = theta * ones(1, length(aPar));
idx = z1 > thetaMtx;
z1(idx) = thetaMtx(idx) - 0.05;
z1 = zeros(length(theta), length(aPar));
for ii = 1:length(aPar)
    pd = makedist("Logistic", "mu", bPar(ii), "sigma", aPar(ii));
    for iTheta = 1:length(theta)
        rVal = aPar(ii) * randn(1) + bPar(ii);
        if (rVal < -6)
            z1(iTheta, ii) = -6;
        else
            z1(iTheta, ii) = rVal;

```

```

        end
    end
end
end
function zParNew = IterateZ(data, zPar, aPar, bPar, theta, mass)
[nRow, nCol] = size(data);
f0 = fProb(theta, aPar, bPar);
p1Pos = round(mass .* f0);
p1Neg = round(mass .* (1 - f0));
zerosData = abs(data - 1);
zParNewPos = SampleZ1(aPar, bPar, theta) .* data;
zParNewNeg = SampleZ0(aPar, bPar, theta) .* zerosData;
zParNew = zParNewPos + zParNewNeg;
p2 = 1;
pp = rand(nCol, nRow);
selectIdx = randi(nRow, [nRow, nCol]);
indVect = ones(nRow, 1);
for iltem = 1:nCol
    zVect = zPar(:, iltem);
    for jIndiv = 1:nRow
        indVect(jIndiv) = 0;
        zVectDmy = zVect(indVect>0);
        indVect(jIndiv) = 1;
        if (data(jIndiv, iltem) == 1)
            prp1 = p1Pos(jIndiv, iltem) / (p1Pos(jIndiv, iltem) + p2);
        else
            prp1 = p1Neg(jIndiv, iltem) / (p1Neg(jIndiv, iltem) + p2);
        end
        if (pp(iltem, jIndiv) >= prp1)
            selectIdx = randi(nRow);
            while (selectIdx(jIndiv, iltem) == jIndiv)
                selectIdx(jIndiv, iltem) = randi(nRow);
            end
            zParNew(jIndiv, iltem) = ...
                zPar(selectIdx(jIndiv, iltem), iltem);
        end
    end
end
end

```

```

        end
    end
end
end
function theta = IterateTheta(data, zPar)
    zDensPos = zPar;
    zDensNeg = zPar;
    zDensPos(data == 0) = nan;
    zDensNeg(data == 1) = nan;
    minVals = max(zDensPos, [], 2, "omitnan");
    maxVals = min(zDensNeg, [], 2, "omitnan");

    minVals(isnan(minVals)) = -6;
    maxVals(isnan(maxVals)) = 6;
    theta = randn(size(minVals));
    theta(theta>maxVals) = maxVals(theta>maxVals);
    theta(theta<minVals) = minVals(theta<minVals);
end
function [aOut, bOut] = Iterate_AB(data, theta, zPar, aPar, bPar)
    g = GroupZ(zPar);
    l0 = ones(size(aPar));
    l1a = ones(size(aPar));
    l1b = ones(size(aPar));
    a1 = aPar .* exp(0.0025*randn(size(aPar)));
    b1 = bPar + 0.0025*randn(size(bPar));
    for iCol = 1:size(data, 2)
        maxClass = max(g(:, iCol));
        for jGrp = 1:maxClass
            dmyData = data(g(:, iCol) == jGrp, iCol);
            dmyTheta = theta(g(:, iCol) == jGrp, 1);
            mxVal = min(dmyTheta(dmyData == 1));
            if (isempty(mxVal))
                mxVal = 6;
            end
        end
        mnVal = max(dmyTheta(dmyData == 0));
    end
end

```

```

    if (isempty(mnVal))
      mnVal = -6;
    end
    randValue = (mxVal - mnVal) * rand(1) + mnVal;
    l0(iCol) = l0(iCol) * ...
      logDensFunc(randValue, aPar(iCol), bPar(iCol));
    l1b(iCol) = l1b(iCol) * ...
      logDensFunc(randValue, aPar(iCol), b1(iCol));
    l1a(iCol) = l1a(iCol) * ...
      logDensFunc(randValue, a1(iCol), bPar(iCol));
  end
end
l1b = sum(log(l1b .* normpdf(b1)));
l1a = sum(log(l1a .* gampdf(a1, 3.2, 1)));
l0a = sum(log(l0 .* gampdf(aPar, 3.2, 1)));
l0b = sum(log(l0 .* normpdf(bPar)));
accA = l1a - l0a;
accA = min(accA, 0);
accA = rand(size(aPar)) - exp(accA);
aOut = aPar;
aOut(accA < 0) = a1(accA < 0);
accB = l1b - l0b;
accB = min(accB, 0);
accB = rand(size(bPar)) - exp(accB);
bOut = bPar;
bOut(accB < 0) = b1(accB < 0);
end

```

### EK-E: Parametrelerin Ölçeklenmesi, RMSE ve MAE Değerlerine İlişkin R Kodu

```
library(irtoys)
```

```
#Gerçek madde parametrelerinin okunması
```

```
trueip<- read.csv(file = "trueab.csv",header = F)
```

```
truea<-trueip[,1]
```

```
trueb<-trueip[,2]
```

```
#sca fonksiyonu için madde parametreleri matrisine c=0 şans parametrelerinin ekl  
enmesi
```

```

trueip<- cbind(trueip,rep(0,nrow(trueip)))

#Gerçek birey parametrelerinin okunması
trueTheta<-read.csv(file = "truetheta.csv",header = T)[-1]

#Kestirilen madde parametrelerinin okunması
esta<-read.table(file = "aEst.dat",header = F)
estb<-read.table(file = "bEst.dat",header = F)

#Kestirilen yetenek parametrelerinin okunması
estTheta<-read.table(file = "ThetaEst.dat",header = F)

#100 tekrar için gerekli dosyaların oluşturulması
TL<-ncol(esta)
link.items <- seq(1,TL,by=1)
result<-matrix(0,ncol = 6,nrow = 100)
quad <- normal.qu(19,-3,3)
colnames(result)<-c("RmseA","MaeA","RmseB","MaeB","RmseTheta","MaeTheta"
)

for(r in 1:100){
#Her bir tekrar için kestirilen parametrelerin seçilmesi
  estip<-t(rbind(esta[r,],estb[r,],rep(0,TL)))
  EstTheta<-estTheta[r,]
  TrueTheta<-trueTheta[r,]

#Ölçeklemenin yapılması
  scaled<-sca(old.ip=as.matrix(trueip),new.ip=as.matrix(estip),old.items=link.items,
new.items=link.items,old.qu=quad,new.qu=quad,method="SL")
  A<-scaled$slope
  B<-scaled$intercept
  aSca<-scaled$scaled.ip[,1]
  bSca<-scaled$scaled.ip[,2]
  ThetaSca<- A*EstTheta+B

#ölçeklenmiş parametreler kullanılarak RMSE ve MAE değerlerinin hesaplanması
  rmseA<- sqrt(mean((truea-aSca)^2))
  rmseB<- sqrt(mean((trueb-bSca)^2))
  maeA<- mean(abs(truea-aSca))
  maeB<- mean(abs(trueb-bSca))
  rmseTheta<- sqrt(mean((t(TrueTheta-ThetaSca))^2))
  maeTheta<- mean(abs(t(TrueTheta-ThetaSca)))
  temp<-cbind(rmseA,maeA,rmseB,maeB,rmseTheta,maeTheta)
  result[r,]<-temp
}

```

### EK-F: Grafiklerin Oluşturulmasına Yönelik Örnek R kodu

```

library(lattice)

veri=read.table(file = "SimSonucFinal.dat",header = T)

```

```
xyplot(RmseTheta ~ Yetenek Dağılımı | as.factor(Örneklem)+ as.factor(Madde),
veri,group=Yontem,xlab="Örtük Özellik Dağılımı", type = "a",auto.key=list(corner=c
(1,0.9),points=FALSE,lines=TRUE,columns=1), par.settings=simpleTheme(lty=1:3,
lwd=2),as.table=T)
```

```
xyplot(RmseTheta ~ Yetenek Dağılımı | bDaglm + as.factor(Örneklem),veri,group
=Yontem,xlab="Örtük Özellik Dağılımı",type = "a",auto.key=list(corner=c(1,0.9),poi
nts=FALSE,lines=TRUE,columns=1),par.settings=simpleTheme(lty=1:3,lwd=2),as.
table=T)
```

### EK-G: Yetenek Parametre Kestirimlerine İlişkin RMSE ve MAE Değerleri

M	N	bDaglm	Yetenek D	Yöntem	Rmse	Rmse sd	Mae	Mae sd
10	100	Normal	Normal	MLE	1.615	0.166	1.292	0.135
10	100	Normal	Normal	MAP	0.562	0.047	0.439	0.039
10	100	Normal	Normal	DE	0.640	0.126	0.494	0.065
10	100	Normal	Normal	AH	0.566	0.046	0.448	0.037
10	500	Normal	Normal	MLE	0.729	0.042	0.553	0.025
10	500	Normal	Normal	MAP	0.534	0.018	0.424	0.015
10	500	Normal	Normal	DE	0.552	0.042	0.437	0.022
10	500	Normal	Normal	AH	0.533	0.018	0.424	0.015
10	1000	Normal	Normal	MLE	0.728	0.040	0.548	0.022
10	1000	Normal	Normal	MAP	0.529	0.012	0.422	0.010
10	1000	Normal	Normal	DE	0.542	0.031	0.428	0.018
10	1000	Normal	Normal	AH	0.530	0.012	0.421	0.011
20	100	Normal	Normal	MLE	0.515	0.047	0.401	0.034
20	100	Normal	Normal	MAP	0.426	0.032	0.340	0.026
20	100	Normal	Normal	DE	0.473	0.071	0.365	0.038
20	100	Normal	Normal	AH	0.432	0.032	0.342	0.028
20	500	Normal	Normal	MLE	0.503	0.025	0.387	0.016
20	500	Normal	Normal	MAP	0.417	0.014	0.330	0.011
20	500	Normal	Normal	DE	0.426	0.020	0.334	0.012
20	500	Normal	Normal	AH	0.416	0.014	0.329	0.012
20	1000	Normal	Normal	MLE	0.500	0.018	0.384	0.011
20	1000	Normal	Normal	MAP	0.415	0.010	0.329	0.008
20	1000	Normal	Normal	DE	0.419	0.013	0.332	0.009
20	1000	Normal	Normal	AH	0.414	0.010	0.329	0.008
30	100	Normal	Normal	MLE	0.387	0.033	0.302	0.024
30	100	Normal	Normal	MAP	0.340	0.025	0.270	0.020
30	100	Normal	Normal	DE	0.367	0.068	0.284	0.031
30	100	Normal	Normal	AH	0.340	0.025	0.270	0.020
30	500	Normal	Normal	MLE	0.388	0.016	0.297	0.011
30	500	Normal	Normal	MAP	0.339	0.011	0.267	0.009
30	500	Normal	Normal	DE	0.345	0.018	0.270	0.010
30	500	Normal	Normal	AH	0.337	0.011	0.267	0.009
30	1000	Normal	Normal	MLE	0.381	0.013	0.294	0.009
30	1000	Normal	Normal	MAP	0.336	0.009	0.266	0.007

30	1000	Normal	Normal	DE	0.338	0.011	0.267	0.008
30	1000	Normal	Normal	AH	0.335	0.009	0.266	0.007
10	100	N-Carpik	Normal	MLE	0.769	0.093	0.579	0.062
10	100	N-Carpik	Normal	MAP	0.529	0.037	0.419	0.029
10	100	N-Carpik	Normal	DE	0.608	0.175	0.470	0.096
10	100	N-Carpik	Normal	AH	0.539	0.082	0.427	0.071
10	500	N-Carpik	Normal	MLE	0.750	0.048	0.553	0.028
10	500	N-Carpik	Normal	MAP	0.520	0.017	0.411	0.013
10	500	N-Carpik	Normal	DE	0.529	0.033	0.419	0.022
10	500	N-Carpik	Normal	AH	0.517	0.016	0.411	0.013
10	1000	N-Carpik	Normal	MLE	0.752	0.038	0.553	0.021
10	1000	N-Carpik	Normal	MAP	0.518	0.013	0.409	0.011
10	1000	N-Carpik	Normal	DE	0.519	0.017	0.410	0.013
10	1000	N-Carpik	Normal	AH	0.515	0.013	0.406	0.011
20	100	N-Carpik	Normal	MLE	0.515	0.048	0.397	0.033
20	100	N-Carpik	Normal	MAP	0.415	0.033	0.329	0.027
20	100	N-Carpik	Normal	DE	0.457	0.207	0.357	0.159
20	100	N-Carpik	Normal	AH	0.424	0.199	0.338	0.159
20	500	N-Carpik	Normal	MLE	0.505	0.027	0.386	0.017
20	500	N-Carpik	Normal	MAP	0.415	0.013	0.329	0.011
20	500	N-Carpik	Normal	DE	0.418	0.014	0.330	0.012
20	500	N-Carpik	Normal	AH	0.413	0.013	0.328	0.011
20	1000	N-Carpik	Normal	MLE	0.504	0.019	0.385	0.011
20	1000	N-Carpik	Normal	MAP	0.413	0.010	0.327	0.008
20	1000	N-Carpik	Normal	DE	0.415	0.011	0.329	0.008
20	1000	N-Carpik	Normal	AH	0.412	0.010	0.326	0.008
30	100	N-Carpik	Normal	MLE	0.400	0.040	0.305	0.028
30	100	N-Carpik	Normal	MAP	0.349	0.027	0.273	0.022
30	100	N-Carpik	Normal	DE	0.369	0.170	0.289	0.128
30	100	N-Carpik	Normal	AH	0.346	0.161	0.276	0.126
30	500	N-Carpik	Normal	MLE	0.388	0.022	0.295	0.013
30	500	N-Carpik	Normal	MAP	0.337	0.012	0.265	0.009
30	500	N-Carpik	Normal	DE	0.341	0.017	0.268	0.012
30	500	N-Carpik	Normal	AH	0.337	0.012	0.264	0.010
30	1000	N-Carpik	Normal	MLE	0.391	0.016	0.295	0.009
30	1000	N-Carpik	Normal	MAP	0.335	0.008	0.263	0.006
30	1000	N-Carpik	Normal	DE	0.336	0.010	0.264	0.007
30	1000	N-Carpik	Normal	AH	0.333	0.008	0.263	0.006
10	100	P-Carpik	Normal	MLE	0.744	0.115	0.558	0.074
10	100	P-Carpik	Normal	MAP	0.532	0.045	0.417	0.034
10	100	P-Carpik	Normal	DE	0.598	0.150	0.469	0.101
10	100	P-Carpik	Normal	AH	0.540	0.108	0.427	0.084
10	500	P-Carpik	Normal	MLE	0.758	0.055	0.558	0.034
10	500	P-Carpik	Normal	MAP	0.519	0.019	0.411	0.014
10	500	P-Carpik	Normal	DE	0.525	0.023	0.415	0.016
10	500	P-Carpik	Normal	AH	0.517	0.018	0.410	0.015
10	1000	P-Carpik	Normal	MLE	0.763	0.039	0.561	0.023



10	1000	P-Carpik	Normal	MAP	0.520	0.011	0.410	0.009
10	1000	P-Carpik	Normal	DE	0.521	0.015	0.411	0.011
10	1000	P-Carpik	Normal	AH	0.516	0.011	0.409	0.009
20	100	P-Carpik	Normal	MLE	0.508	0.051	0.390	0.033
20	100	P-Carpik	Normal	MAP	0.418	0.030	0.329	0.025
20	100	P-Carpik	Normal	DE	0.446	0.125	0.351	0.075
20	100	P-Carpik	Normal	AH	0.424	0.082	0.335	0.066
20	500	P-Carpik	Normal	MLE	0.507	0.027	0.387	0.017
20	500	P-Carpik	Normal	MAP	0.414	0.015	0.326	0.012
20	500	P-Carpik	Normal	DE	0.419	0.016	0.330	0.012
20	500	P-Carpik	Normal	AH	0.412	0.015	0.325	0.012
20	1000	P-Carpik	Normal	MLE	0.509	0.020	0.384	0.012
20	1000	P-Carpik	Normal	MAP	0.412	0.009	0.324	0.007
20	1000	P-Carpik	Normal	DE	0.411	0.010	0.326	0.008
20	1000	P-Carpik	Normal	AH	0.409	0.009	0.324	0.007
30	100	P-Carpik	Normal	MLE	0.388	0.042	0.299	0.028
30	100	P-Carpik	Normal	MAP	0.343	0.025	0.270	0.020
30	100	P-Carpik	Normal	DE	0.369	0.124	0.287	0.093
30	100	P-Carpik	Normal	AH	0.343	0.113	0.272	0.090
30	500	P-Carpik	Normal	MLE	0.388	0.020	0.298	0.012
30	500	P-Carpik	Normal	MAP	0.339	0.012	0.268	0.009
30	500	P-Carpik	Normal	DE	0.342	0.017	0.270	0.011
30	500	P-Carpik	Normal	AH	0.337	0.011	0.266	0.009
30	1000	P-Carpik	Normal	MLE	0.387	0.013	0.294	0.008
30	1000	P-Carpik	Normal	MAP	0.337	0.008	0.265	0.006
30	1000	P-Carpik	Normal	DE	0.340	0.010	0.267	0.007
30	1000	P-Carpik	Normal	AH	0.335	0.008	0.264	0.006
10	100	Normal	N-Carpik	MLE	0.750	0.084	0.575	0.054
10	100	Normal	N-Carpik	MAP	0.591	0.047	0.441	0.031
10	100	Normal	N-Carpik	DE	0.585	0.069	0.444	0.048
10	100	Normal	N-Carpik	AH	0.605	0.052	0.458	0.036
10	500	Normal	N-Carpik	MLE	0.755	0.037	0.573	0.025
10	500	Normal	N-Carpik	MAP	0.608	0.022	0.441	0.014
10	500	Normal	N-Carpik	DE	0.544	0.029	0.408	0.018
10	500	Normal	N-Carpik	AH	0.603	0.022	0.443	0.014
10	1000	Normal	N-Carpik	MLE	0.750	0.027	0.572	0.017
10	1000	Normal	N-Carpik	MAP	0.601	0.018	0.438	0.009
10	1000	Normal	N-Carpik	DE	0.534	0.023	0.404	0.014
10	1000	Normal	N-Carpik	AH	0.597	0.018	0.435	0.009
20	100	Normal	N-Carpik	MLE	0.513	0.042	0.398	0.030
20	100	Normal	N-Carpik	MAP	0.488	0.045	0.357	0.025
20	100	Normal	N-Carpik	DE	0.443	0.047	0.339	0.030
20	100	Normal	N-Carpik	AH	0.486	0.047	0.369	0.028
20	500	Normal	N-Carpik	MLE	0.498	0.020	0.382	0.014
20	500	Normal	N-Carpik	MAP	0.494	0.023	0.351	0.012
20	500	Normal	N-Carpik	DE	0.427	0.023	0.319	0.014
20	500	Normal	N-Carpik	AH	0.488	0.023	0.352	0.012

20	1000	Normal	N-Carpik	MLE	0.502	0.016	0.383	0.010
20	1000	Normal	N-Carpik	MAP	0.489	0.017	0.350	0.007
20	1000	Normal	N-Carpik	DE	0.425	0.017	0.316	0.008
20	1000	Normal	N-Carpik	AH	0.485	0.017	0.349	0.007
30	100	Normal	N-Carpik	MLE	0.400	0.037	0.309	0.026
30	100	Normal	N-Carpik	MAP	0.406	0.046	0.294	0.024
30	100	Normal	N-Carpik	DE	0.358	0.040	0.273	0.025
30	100	Normal	N-Carpik	AH	0.394	0.046	0.299	0.026
30	500	Normal	N-Carpik	MLE	0.396	0.024	0.298	0.011
30	500	Normal	N-Carpik	MAP	0.426	0.028	0.291	0.009
30	500	Normal	N-Carpik	DE	0.359	0.027	0.261	0.010
30	500	Normal	N-Carpik	AH	0.417	0.029	0.290	0.010
30	1000	Normal	N-Carpik	MLE	0.398	0.014	0.299	0.007
30	1000	Normal	N-Carpik	MAP	0.426	0.021	0.293	0.006
30	1000	Normal	N-Carpik	DE	0.360	0.018	0.261	0.006
30	1000	Normal	N-Carpik	AH	0.417	0.021	0.291	0.006
10	100	N-Carpik	N-Carpik	MLE	0.704	0.064	0.534	0.044
10	100	N-Carpik	N-Carpik	MAP	0.600	0.049	0.432	0.024
10	100	N-Carpik	N-Carpik	DE	0.559	0.061	0.419	0.040
10	100	N-Carpik	N-Carpik	AH	0.594	0.051	0.444	0.027
10	500	N-Carpik	N-Carpik	MLE	0.708	0.030	0.527	0.021
10	500	N-Carpik	N-Carpik	MAP	0.596	0.023	0.423	0.012
10	500	N-Carpik	N-Carpik	DE	0.539	0.031	0.397	0.020
10	500	N-Carpik	N-Carpik	AH	0.596	0.023	0.429	0.012
10	1000	N-Carpik	N-Carpik	MLE	0.696	0.025	0.521	0.016
10	1000	N-Carpik	N-Carpik	MAP	0.596	0.016	0.420	0.007
10	1000	N-Carpik	N-Carpik	DE	0.531	0.026	0.393	0.015
10	1000	N-Carpik	N-Carpik	AH	0.594	0.016	0.424	0.007
20	100	N-Carpik	N-Carpik	MLE	0.504	0.042	0.387	0.030
20	100	N-Carpik	N-Carpik	MAP	0.497	0.049	0.352	0.026
20	100	N-Carpik	N-Carpik	DE	0.447	0.145	0.335	0.103
20	100	N-Carpik	N-Carpik	AH	0.493	0.120	0.363	0.092
20	500	N-Carpik	N-Carpik	MLE	0.489	0.022	0.374	0.014
20	500	N-Carpik	N-Carpik	MAP	0.489	0.026	0.346	0.011
20	500	N-Carpik	N-Carpik	DE	0.425	0.024	0.314	0.012
20	500	N-Carpik	N-Carpik	AH	0.484	0.026	0.346	0.011
20	1000	N-Carpik	N-Carpik	MLE	0.488	0.016	0.371	0.008
20	1000	N-Carpik	N-Carpik	MAP	0.490	0.020	0.347	0.007
20	1000	N-Carpik	N-Carpik	DE	0.423	0.019	0.313	0.010
20	1000	N-Carpik	N-Carpik	AH	0.485	0.020	0.347	0.007
30	100	N-Carpik	N-Carpik	MLE	0.394	0.041	0.302	0.025
30	100	N-Carpik	N-Carpik	MAP	0.424	0.048	0.295	0.021
30	100	N-Carpik	N-Carpik	DE	0.366	0.040	0.276	0.024
30	100	N-Carpik	N-Carpik	AH	0.410	0.050	0.296	0.023
30	500	N-Carpik	N-Carpik	MLE	0.388	0.021	0.295	0.011
30	500	N-Carpik	N-Carpik	MAP	0.425	0.028	0.291	0.009
30	500	N-Carpik	N-Carpik	DE	0.360	0.023	0.261	0.009

30	500	N-Carpik	N-Carpik	AH	0.415	0.028	0.289	0.009
30	1000	N-Carpik	N-Carpik	MLE	0.390	0.017	0.289	0.007
30	1000	N-Carpik	N-Carpik	MAP	0.430	0.023	0.289	0.007
30	1000	N-Carpik	N-Carpik	DE	0.362	0.022	0.258	0.007
30	1000	N-Carpik	N-Carpik	AH	0.422	0.023	0.287	0.007
10	100	P-Carpik	N-Carpik	MLE	0.797	0.094	0.588	0.063
10	100	P-Carpik	N-Carpik	MAP	0.593	0.046	0.426	0.027
10	100	P-Carpik	N-Carpik	DE	0.580	0.105	0.431	0.074
10	100	P-Carpik	N-Carpik	AH	0.600	0.074	0.447	0.057
10	500	P-Carpik	N-Carpik	MLE	0.802	0.052	0.584	0.029
10	500	P-Carpik	N-Carpik	MAP	0.594	0.025	0.420	0.011
10	500	P-Carpik	N-Carpik	DE	0.542	0.030	0.399	0.019
10	500	P-Carpik	N-Carpik	AH	0.589	0.025	0.420	0.013
10	1000	P-Carpik	N-Carpik	MLE	0.814	0.050	0.591	0.027
10	1000	P-Carpik	N-Carpik	MAP	0.593	0.017	0.420	0.008
10	1000	P-Carpik	N-Carpik	DE	0.532	0.021	0.393	0.013
10	1000	P-Carpik	N-Carpik	AH	0.589	0.017	0.420	0.009
20	100	P-Carpik	N-Carpik	MLE	0.530	0.049	0.402	0.035
20	100	P-Carpik	N-Carpik	MAP	0.486	0.043	0.357	0.024
20	100	P-Carpik	N-Carpik	DE	0.459	0.299	0.351	0.212
20	100	P-Carpik	N-Carpik	AH	0.492	0.256	0.364	0.196
20	500	P-Carpik	N-Carpik	MLE	0.515	0.024	0.389	0.014
20	500	P-Carpik	N-Carpik	MAP	0.486	0.028	0.343	0.011
20	500	P-Carpik	N-Carpik	DE	0.426	0.024	0.315	0.012
20	500	P-Carpik	N-Carpik	AH	0.479	0.028	0.344	0.012
20	1000	P-Carpik	N-Carpik	MLE	0.517	0.019	0.390	0.011
20	1000	P-Carpik	N-Carpik	MAP	0.488	0.020	0.345	0.008
20	1000	P-Carpik	N-Carpik	DE	0.422	0.018	0.312	0.008
20	1000	P-Carpik	N-Carpik	AH	0.481	0.020	0.344	0.008
30	100	P-Carpik	N-Carpik	MLE	0.404	0.044	0.309	0.026
30	100	P-Carpik	N-Carpik	MAP	0.405	0.054	0.293	0.023
30	100	P-Carpik	N-Carpik	DE	0.363	0.221	0.279	0.158
30	100	P-Carpik	N-Carpik	AH	0.395	0.198	0.297	0.149
30	500	P-Carpik	N-Carpik	MLE	0.409	0.022	0.305	0.011
30	500	P-Carpik	N-Carpik	MAP	0.423	0.028	0.287	0.009
30	500	P-Carpik	N-Carpik	DE	0.358	0.027	0.260	0.009
30	500	P-Carpik	N-Carpik	AH	0.413	0.028	0.286	0.009
30	1000	P-Carpik	N-Carpik	MLE	0.405	0.018	0.302	0.008
30	1000	P-Carpik	N-Carpik	MAP	0.417	0.024	0.287	0.007
30	1000	P-Carpik	N-Carpik	DE	0.355	0.021	0.258	0.007
30	1000	P-Carpik	N-Carpik	AH	0.409	0.024	0.286	0.007
10	100	Normal	P-Carpik	MLE	0.755	0.076	0.585	0.053
10	100	Normal	P-Carpik	MAP	0.607	0.048	0.448	0.029
10	100	Normal	P-Carpik	DE	0.602	0.081	0.457	0.049
10	100	Normal	P-Carpik	AH	0.615	0.051	0.466	0.034
10	500	Normal	P-Carpik	MLE	0.739	0.033	0.561	0.023
10	500	Normal	P-Carpik	MAP	0.602	0.023	0.439	0.015

10	500	Normal	P-Carpik	DE	0.546	0.025	0.409	0.018
10	500	Normal	P-Carpik	AH	0.598	0.023	0.441	0.014
10	1000	Normal	P-Carpik	MLE	0.742	0.025	0.564	0.017
10	1000	Normal	P-Carpik	MAP	0.600	0.019	0.438	0.008
10	1000	Normal	P-Carpik	DE	0.536	0.017	0.403	0.010
10	1000	Normal	P-Carpik	AH	0.593	0.018	0.438	0.008
20	100	Normal	P-Carpik	MLE	0.523	0.047	0.396	0.032
20	100	Normal	P-Carpik	MAP	0.493	0.052	0.356	0.022
20	100	Normal	P-Carpik	DE	0.457	0.086	0.341	0.060
20	100	Normal	P-Carpik	AH	0.492	0.083	0.365	0.055
20	500	Normal	P-Carpik	MLE	0.511	0.023	0.389	0.015
20	500	Normal	P-Carpik	MAP	0.493	0.024	0.349	0.011
20	500	Normal	P-Carpik	DE	0.424	0.022	0.319	0.012
20	500	Normal	P-Carpik	AH	0.487	0.024	0.349	0.011
20	1000	Normal	P-Carpik	MLE	0.506	0.017	0.386	0.011
20	1000	Normal	P-Carpik	MAP	0.495	0.020	0.350	0.008
20	1000	Normal	P-Carpik	DE	0.420	0.019	0.315	0.010
20	1000	Normal	P-Carpik	AH	0.488	0.020	0.349	0.009
30	100	Normal	P-Carpik	MLE	0.401	0.039	0.313	0.026
30	100	Normal	P-Carpik	MAP	0.403	0.041	0.293	0.021
30	100	Normal	P-Carpik	DE	0.364	0.052	0.280	0.027
30	100	Normal	P-Carpik	AH	0.396	0.042	0.298	0.023
30	500	Normal	P-Carpik	MLE	0.395	0.025	0.300	0.011
30	500	Normal	P-Carpik	MAP	0.420	0.031	0.291	0.009
30	500	Normal	P-Carpik	DE	0.352	0.029	0.260	0.010
30	500	Normal	P-Carpik	AH	0.411	0.031	0.290	0.009
30	1000	Normal	P-Carpik	MLE	0.392	0.016	0.298	0.007
30	1000	Normal	P-Carpik	MAP	0.419	0.021	0.289	0.007
30	1000	Normal	P-Carpik	DE	0.354	0.019	0.259	0.006
30	1000	Normal	P-Carpik	AH	0.410	0.021	0.288	0.006
10	100	N-Carpik	P-Carpik	MLE	0.784	0.102	0.584	0.063
10	100	N-Carpik	P-Carpik	MAP	0.596	0.045	0.422	0.026
10	100	N-Carpik	P-Carpik	DE	0.577	0.161	0.427	0.112
10	100	N-Carpik	P-Carpik	AH	0.608	0.129	0.441	0.100
10	500	N-Carpik	P-Carpik	MLE	0.796	0.058	0.582	0.034
10	500	N-Carpik	P-Carpik	MAP	0.595	0.026	0.423	0.014
10	500	N-Carpik	P-Carpik	DE	0.531	0.028	0.394	0.020
10	500	N-Carpik	P-Carpik	AH	0.589	0.026	0.423	0.015
10	1000	N-Carpik	P-Carpik	MLE	0.809	0.046	0.589	0.025
10	1000	N-Carpik	P-Carpik	MAP	0.594	0.017	0.421	0.008
10	1000	N-Carpik	P-Carpik	DE	0.532	0.018	0.395	0.011
10	1000	N-Carpik	P-Carpik	AH	0.587	0.017	0.420	0.009
20	100	N-Carpik	P-Carpik	MLE	0.525	0.046	0.403	0.030
20	100	N-Carpik	P-Carpik	MAP	0.481	0.047	0.352	0.024
20	100	N-Carpik	P-Carpik	DE	0.449	0.217	0.339	0.151
20	100	N-Carpik	P-Carpik	AH	0.481	0.183	0.359	0.138
20	500	N-Carpik	P-Carpik	MLE	0.517	0.021	0.389	0.014

20	500	N-Carpik	P-Carpik	MAP	0.484	0.023	0.348	0.011
20	500	N-Carpik	P-Carpik	DE	0.423	0.021	0.314	0.011
20	500	N-Carpik	P-Carpik	AH	0.478	0.023	0.347	0.011
20	1000	N-Carpik	P-Carpik	MLE	0.517	0.018	0.389	0.011
20	1000	N-Carpik	P-Carpik	MAP	0.493	0.022	0.346	0.007
20	1000	N-Carpik	P-Carpik	DE	0.427	0.024	0.313	0.009
20	1000	N-Carpik	P-Carpik	AH	0.487	0.022	0.346	0.007
30	100	N-Carpik	P-Carpik	MLE	0.405	0.043	0.309	0.028
30	100	N-Carpik	P-Carpik	MAP	0.410	0.045	0.290	0.020
30	100	N-Carpik	P-Carpik	DE	0.357	0.112	0.272	0.077
30	100	N-Carpik	P-Carpik	AH	0.403	0.103	0.296	0.074
30	500	N-Carpik	P-Carpik	MLE	0.407	0.021	0.307	0.010
30	500	N-Carpik	P-Carpik	MAP	0.420	0.030	0.289	0.009
30	500	N-Carpik	P-Carpik	DE	0.356	0.026	0.261	0.009
30	500	N-Carpik	P-Carpik	AH	0.409	0.029	0.289	0.009
30	1000	N-Carpik	P-Carpik	MLE	0.404	0.017	0.303	0.008
30	1000	N-Carpik	P-Carpik	MAP	0.416	0.021	0.287	0.006
30	1000	N-Carpik	P-Carpik	DE	0.353	0.020	0.257	0.006
30	1000	N-Carpik	P-Carpik	AH	0.409	0.021	0.286	0.006
10	100	P-Carpik	P-Carpik	MLE	0.682	0.060	0.520	0.042
10	100	P-Carpik	P-Carpik	MAP	0.595	0.041	0.426	0.022
10	100	P-Carpik	P-Carpik	DE	0.550	0.051	0.417	0.033
10	100	P-Carpik	P-Carpik	AH	0.771	1.009	0.498	0.332
10	500	P-Carpik	P-Carpik	MLE	0.703	0.034	0.526	0.022
10	500	P-Carpik	P-Carpik	MAP	0.594	0.024	0.421	0.012
10	500	P-Carpik	P-Carpik	DE	0.532	0.026	0.394	0.016
10	500	P-Carpik	P-Carpik	AH	0.591	0.024	0.426	0.013
10	1000	P-Carpik	P-Carpik	MLE	0.698	0.024	0.522	0.014
10	1000	P-Carpik	P-Carpik	MAP	0.589	0.017	0.420	0.008
10	1000	P-Carpik	P-Carpik	DE	0.534	0.019	0.393	0.011
10	1000	P-Carpik	P-Carpik	AH	0.590	0.017	0.425	0.009
20	100	P-Carpik	P-Carpik	MLE	0.492	0.039	0.382	0.029
20	100	P-Carpik	P-Carpik	MAP	0.487	0.047	0.351	0.024
20	100	P-Carpik	P-Carpik	DE	0.439	0.053	0.332	0.028
20	100	P-Carpik	P-Carpik	AH	0.480	0.051	0.356	0.027
20	500	P-Carpik	P-Carpik	MLE	0.487	0.020	0.371	0.013
20	500	P-Carpik	P-Carpik	MAP	0.491	0.023	0.347	0.011
20	500	P-Carpik	P-Carpik	DE	0.421	0.019	0.313	0.010
20	500	P-Carpik	P-Carpik	AH	0.485	0.023	0.347	0.011
20	1000	P-Carpik	P-Carpik	MLE	0.492	0.017	0.374	0.010
20	1000	P-Carpik	P-Carpik	MAP	0.494	0.022	0.348	0.008
20	1000	P-Carpik	P-Carpik	DE	0.424	0.018	0.313	0.009
20	1000	P-Carpik	P-Carpik	AH	0.489	0.022	0.348	0.008
30	100	P-Carpik	P-Carpik	MLE	0.389	0.043	0.301	0.024
30	100	P-Carpik	P-Carpik	MAP	0.424	0.052	0.297	0.023
30	100	P-Carpik	P-Carpik	DE	0.374	0.118	0.275	0.079
30	100	P-Carpik	P-Carpik	AH	0.407	0.108	0.295	0.073

30	500	P-Carpik	P-Carpik	MLE	0.386	0.022	0.292	0.010
30	500	P-Carpik	P-Carpik	MAP	0.424	0.028	0.289	0.009
30	500	P-Carpik	P-Carpik	DE	0.358	0.024	0.261	0.008
30	500	P-Carpik	P-Carpik	AH	0.414	0.028	0.287	0.009
30	1000	P-Carpik	P-Carpik	MLE	0.383	0.016	0.289	0.008
30	1000	P-Carpik	P-Carpik	MAP	0.420	0.019	0.288	0.007
30	1000	P-Carpik	P-Carpik	DE	0.354	0.019	0.258	0.008
30	1000	P-Carpik	P-Carpik	AH	0.411	0.019	0.287	0.007
10	100	Normal	Normal	NBE	0.839	0.063	0.685	0.056
10	500	Normal	Normal	NBE	0.852	0.041	0.688	0.038
10	1000	Normal	Normal	NBE	0.845	0.041	0.680	0.036
20	100	Normal	Normal	NBE	1.127	0.069	0.977	0.064
20	500	Normal	Normal	NBE	1.107	0.037	0.953	0.037
20	1000	Normal	Normal	NBE	1.095	0.037	0.940	0.038
30	100	Normal	Normal	NBE	1.210	0.068	1.074	0.069
30	500	Normal	Normal	NBE	1.209	0.041	1.068	0.041
30	1000	Normal	Normal	NBE	1.198	0.035	1.055	0.036
10	100	N-Carpik	Normal	NBE	0.873	0.061	0.708	0.049
10	500	N-Carpik	Normal	NBE	0.876	0.032	0.708	0.028
10	1000	N-Carpik	Normal	NBE	0.882	0.027	0.714	0.026
20	100	N-Carpik	Normal	NBE	1.160	0.056	1.008	0.056
20	500	N-Carpik	Normal	NBE	1.149	0.035	0.993	0.036
20	1000	N-Carpik	Normal	NBE	1.127	0.029	0.972	0.030
30	100	N-Carpik	Normal	NBE	1.261	0.065	1.118	0.063
30	500	N-Carpik	Normal	NBE	1.248	0.036	1.112	0.037
30	1000	N-Carpik	Normal	NBE	1.226	0.028	1.088	0.028
10	100	P-Carpik	Normal	NBE	0.827	0.071	0.670	0.060
10	500	P-Carpik	Normal	NBE	0.822	0.032	0.662	0.029
10	1000	P-Carpik	Normal	NBE	0.818	0.029	0.660	0.027
20	100	P-Carpik	Normal	NBE	1.055	0.073	0.897	0.073
20	500	P-Carpik	Normal	NBE	1.072	0.037	0.913	0.038
20	1000	P-Carpik	Normal	NBE	1.058	0.039	0.896	0.038
30	100	P-Carpik	Normal	NBE	1.153	0.059	0.993	0.060
30	500	P-Carpik	Normal	NBE	1.159	0.038	1.005	0.040
30	1000	P-Carpik	Normal	NBE	1.145	0.027	0.992	0.028
10	100	Normal	N-Carpik	NBE	0.809	0.058	0.597	0.045
10	500	Normal	N-Carpik	NBE	0.818	0.032	0.602	0.032
10	1000	Normal	N-Carpik	NBE	0.822	0.029	0.600	0.029
20	100	Normal	N-Carpik	NBE	1.079	0.053	0.911	0.044
20	500	Normal	N-Carpik	NBE	1.072	0.034	0.905	0.040
20	1000	Normal	N-Carpik	NBE	1.060	0.030	0.890	0.037
30	100	Normal	N-Carpik	NBE	1.172	0.047	1.032	0.044
30	500	Normal	N-Carpik	NBE	1.169	0.030	1.023	0.035
30	1000	Normal	N-Carpik	NBE	1.152	0.026	0.999	0.032
10	100	N-Carpik	N-Carpik	NBE	0.857	0.057	0.643	0.046
10	500	N-Carpik	N-Carpik	NBE	0.868	0.033	0.641	0.034
10	1000	N-Carpik	N-Carpik	NBE	0.862	0.028	0.634	0.028

20	100	N-Carpik	N-Carpik	NBE	1.118	0.046	0.966	0.043
20	500	N-Carpik	N-Carpik	NBE	1.112	0.030	0.952	0.031
20	1000	N-Carpik	N-Carpik	NBE	1.100	0.030	0.931	0.035
30	100	N-Carpik	N-Carpik	NBE	1.209	0.044	1.082	0.041
30	500	N-Carpik	N-Carpik	NBE	1.207	0.028	1.069	0.029
30	1000	N-Carpik	N-Carpik	NBE	1.195	0.026	1.057	0.029
10	100	P-Carpik	N-Carpik	NBE	0.785	0.048	0.561	0.034
10	500	P-Carpik	N-Carpik	NBE	0.787	0.032	0.554	0.026
10	1000	P-Carpik	N-Carpik	NBE	0.790	0.027	0.552	0.023
20	100	P-Carpik	N-Carpik	NBE	1.010	0.054	0.811	0.054
20	500	P-Carpik	N-Carpik	NBE	1.016	0.035	0.824	0.041
20	1000	P-Carpik	N-Carpik	NBE	1.011	0.029	0.810	0.038
30	100	P-Carpik	N-Carpik	NBE	1.097	0.044	0.925	0.041
30	500	P-Carpik	N-Carpik	NBE	1.100	0.032	0.919	0.035
30	1000	P-Carpik	N-Carpik	NBE	1.084	0.028	0.899	0.035
10	100	Normal	P-Carpik	NBE	0.895	0.054	0.744	0.053
10	500	Normal	P-Carpik	NBE	0.907	0.037	0.757	0.036
10	1000	Normal	P-Carpik	NBE	0.919	0.044	0.759	0.042
20	100	Normal	P-Carpik	NBE	1.171	0.049	1.054	0.051
20	500	Normal	P-Carpik	NBE	1.177	0.033	1.059	0.035
20	1000	Normal	P-Carpik	NBE	1.172	0.035	1.045	0.038
30	100	Normal	P-Carpik	NBE	1.275	0.042	1.174	0.043
30	500	Normal	P-Carpik	NBE	1.275	0.033	1.166	0.034
30	1000	Normal	P-Carpik	NBE	1.260	0.030	1.149	0.032
10	100	N-Carpik	P-Carpik	NBE	0.926	0.053	0.760	0.052
10	500	N-Carpik	P-Carpik	NBE	0.950	0.030	0.787	0.029
10	1000	N-Carpik	P-Carpik	NBE	0.944	0.026	0.779	0.027
20	100	N-Carpik	P-Carpik	NBE	1.210	0.046	1.088	0.049
20	500	N-Carpik	P-Carpik	NBE	1.211	0.027	1.083	0.029
20	1000	N-Carpik	P-Carpik	NBE	1.207	0.030	1.079	0.031
30	100	N-Carpik	P-Carpik	NBE	1.304	0.042	1.205	0.044
30	500	N-Carpik	P-Carpik	NBE	1.301	0.028	1.194	0.028
30	1000	N-Carpik	P-Carpik	NBE	1.292	0.023	1.184	0.024
10	100	P-Carpik	P-Carpik	NBE	0.886	0.040	0.742	0.039
10	500	P-Carpik	P-Carpik	NBE	0.897	0.032	0.741	0.033
10	1000	P-Carpik	P-Carpik	NBE	0.896	0.026	0.741	0.028
20	100	P-Carpik	P-Carpik	NBE	1.134	0.042	1.011	0.046
20	500	P-Carpik	P-Carpik	NBE	1.129	0.032	1.010	0.031
20	1000	P-Carpik	P-Carpik	NBE	1.130	0.032	1.007	0.033
30	100	P-Carpik	P-Carpik	NBE	1.218	0.041	1.107	0.042
30	500	P-Carpik	P-Carpik	NBE	1.224	0.031	1.108	0.033
30	1000	P-Carpik	P-Carpik	NBE	1.211	0.031	1.096	0.032

## EK-H: Madde Parametre Kestirimlerine İlişkin RMSE ve MAE Değerleri

Madde S.	Örnekleme	B Dağılımı	Yetenek Dağılımı	Yöntem	A parametresi				B parametresi			
					Rmse	Rmse-sd	Mae	Mae sd	Rmse	Rmse sd	Mae	Mae sd
10	100	Normal	Normal	MLE	0.389	0.095	0.304	0.075	0.266	0.105	0.201	0.062
10	100	Normal	Normal	MAP	0.338	0.085	0.265	0.071	0.222	0.057	0.176	0.041
10	100	Normal	Normal	DE	0.763	2.484	0.571	0.815	0.348	0.399	0.261	0.154
10	100	Normal	Normal	AH	0.599	0.868	0.476	0.319	0.284	0.232	0.218	0.095
10	500	Normal	Normal	MLE	0.206	0.059	0.159	0.042	0.122	0.036	0.090	0.025
10	500	Normal	Normal	MAP	0.200	0.057	0.156	0.041	0.117	0.032	0.087	0.024
10	500	Normal	Normal	DE	0.276	0.179	0.201	0.076	0.134	0.051	0.106	0.029
10	500	Normal	Normal	AH	0.236	0.124	0.182	0.060	0.121	0.034	0.091	0.025
10	1000	Normal	Normal	MLE	0.157	0.043	0.122	0.031	0.088	0.025	0.069	0.018
10	1000	Normal	Normal	MAP	0.154	0.043	0.121	0.031	0.087	0.024	0.069	0.018
10	1000	Normal	Normal	DE	0.185	0.079	0.139	0.044	0.091	0.034	0.072	0.022
10	1000	Normal	Normal	AH	0.165	0.055	0.123	0.034	0.088	0.024	0.070	0.018
20	100	Normal	Normal	MLE	0.295	0.057	0.237	0.043	0.281	0.095	0.202	0.046
20	100	Normal	Normal	MAP	0.277	0.056	0.222	0.041	0.235	0.042	0.180	0.034
20	100	Normal	Normal	DE	0.464	1.363	0.358	0.410	0.343	0.807	0.237	0.224
20	100	Normal	Normal	AH	0.424	0.508	0.325	0.166	0.312	0.510	0.225	0.133
20	500	Normal	Normal	MLE	0.159	0.032	0.126	0.024	0.127	0.040	0.093	0.021
20	500	Normal	Normal	MAP	0.155	0.031	0.124	0.024	0.123	0.031	0.093	0.018
20	500	Normal	Normal	DE	0.171	0.044	0.135	0.027	0.130	0.037	0.098	0.020
20	500	Normal	Normal	AH	0.166	0.040	0.135	0.026	0.130	0.036	0.097	0.020
20	1000	Normal	Normal	MLE	0.119	0.021	0.094	0.017	0.091	0.024	0.069	0.013
20	1000	Normal	Normal	MAP	0.117	0.021	0.093	0.017	0.089	0.021	0.068	0.012
20	1000	Normal	Normal	DE	0.126	0.024	0.099	0.018	0.093	0.024	0.070	0.013
20	1000	Normal	Normal	AH	0.121	0.022	0.096	0.018	0.091	0.023	0.070	0.013
30	100	Normal	Normal	MLE	0.309	0.049	0.244	0.036	0.254	0.051	0.188	0.030
30	100	Normal	Normal	MAP	0.297	0.046	0.231	0.035	0.226	0.036	0.174	0.027
30	100	Normal	Normal	DE	0.445	0.127	0.337	0.063	0.329	0.442	0.229	0.099
30	100	Normal	Normal	AH	0.419	0.107	0.320	0.058	0.315	0.389	0.216	0.085
30	500	Normal	Normal	MLE	0.162	0.025	0.128	0.018	0.125	0.031	0.096	0.015
30	500	Normal	Normal	MAP	0.160	0.025	0.126	0.018	0.121	0.024	0.093	0.014
30	500	Normal	Normal	DE	0.174	0.028	0.138	0.020	0.133	0.035	0.101	0.016
30	500	Normal	Normal	AH	0.171	0.028	0.136	0.020	0.127	0.036	0.098	0.016
30	1000	Normal	Normal	MLE	0.116	0.018	0.092	0.013	0.086	0.020	0.064	0.010
30	1000	Normal	Normal	MAP	0.114	0.018	0.091	0.013	0.084	0.017	0.063	0.009
30	1000	Normal	Normal	DE	0.119	0.020	0.095	0.015	0.090	0.021	0.065	0.011
30	1000	Normal	Normal	AH	0.117	0.019	0.094	0.014	0.086	0.021	0.065	0.010
10	100	N-Carpik	Normal	MLE	0.326	0.074	0.257	0.061	0.265	0.158	0.195	0.075
10	100	N-Carpik	Normal	MAP	0.300	0.072	0.234	0.059	0.204	0.060	0.163	0.043
10	100	N-Carpik	Normal	DE	0.567	1.149	0.436	0.395	0.391	0.654	0.273	0.226
10	100	N-Carpik	Normal	AH	0.511	0.595	0.389	0.230	0.284	1.405	0.217	0.455



10	500	N-Carpik	Normal	MLE	0.203	0.059	0.157	0.041	0.132	0.101	0.095	0.043
10	500	N-Carpik	Normal	MAP	0.187	0.056	0.149	0.040	0.119	0.056	0.092	0.031
10	500	N-Carpik	Normal	DE	0.241	0.113	0.186	0.063	0.155	0.168	0.110	0.065
10	500	N-Carpik	Normal	AH	0.214	0.097	0.172	0.055	0.133	0.110	0.097	0.045
10	1000	N-Carpik	Normal	MLE	0.140	0.039	0.111	0.027	0.086	0.056	0.067	0.028
10	1000	N-Carpik	Normal	MAP	0.138	0.038	0.109	0.027	0.086	0.045	0.065	0.025
10	1000	N-Carpik	Normal	DE	0.154	0.053	0.121	0.034	0.106	0.078	0.072	0.032
10	1000	N-Carpik	Normal	AH	0.147	0.050	0.113	0.031	0.092	0.052	0.067	0.027
20	100	N-Carpik	Normal	MLE	0.296	0.052	0.233	0.040	0.259	0.115	0.196	0.045
20	100	N-Carpik	Normal	MAP	0.287	0.051	0.225	0.039	0.233	0.047	0.182	0.030
20	100	N-Carpik	Normal	DE	0.490	1.746	0.370	0.406	0.362	0.904	0.250	0.213
20	100	N-Carpik	Normal	AH	0.462	0.633	0.355	0.167	0.325	0.856	0.235	0.200
20	500	N-Carpik	Normal	MLE	0.150	0.032	0.120	0.023	0.133	0.136	0.098	0.036
20	500	N-Carpik	Normal	MAP	0.149	0.029	0.116	0.022	0.118	0.043	0.090	0.019
20	500	N-Carpik	Normal	DE	0.169	0.044	0.131	0.028	0.144	0.588	0.102	0.134
20	500	N-Carpik	Normal	AH	0.166	0.041	0.129	0.026	0.132	0.574	0.099	0.131
20	1000	N-Carpik	Normal	MLE	0.114	0.021	0.091	0.016	0.099	0.044	0.069	0.016
20	1000	N-Carpik	Normal	MAP	0.113	0.021	0.089	0.016	0.092	0.029	0.067	0.013
20	1000	N-Carpik	Normal	DE	0.124	0.026	0.096	0.018	0.096	0.055	0.072	0.017
20	1000	N-Carpik	Normal	AH	0.121	0.024	0.093	0.017	0.097	0.042	0.071	0.016
30	100	N-Carpik	Normal	MLE	0.304	0.049	0.236	0.036	0.266	0.128	0.191	0.079
30	100	N-Carpik	Normal	MAP	0.295	0.051	0.232	0.037	0.251	0.044	0.183	0.027
30	100	N-Carpik	Normal	DE	0.443	0.705	0.329	0.150	0.341	6.604	0.227	1.212
30	100	N-Carpik	Normal	AH	0.423	0.359	0.321	0.095	0.304	6.158	0.216	1.127
30	500	N-Carpik	Normal	MLE	0.158	0.024	0.126	0.020	0.134	0.102	0.091	0.024
30	500	N-Carpik	Normal	MAP	0.156	0.025	0.126	0.020	0.118	0.032	0.087	0.014
30	500	N-Carpik	Normal	DE	0.169	0.030	0.132	0.022	0.144	0.618	0.096	0.115
30	500	N-Carpik	Normal	AH	0.168	0.029	0.132	0.021	0.137	0.613	0.093	0.114
30	1000	N-Carpik	Normal	MLE	0.117	0.015	0.092	0.011	0.097	0.053	0.066	0.015
30	1000	N-Carpik	Normal	MAP	0.116	0.015	0.090	0.011	0.087	0.028	0.063	0.011
30	1000	N-Carpik	Normal	DE	0.122	0.018	0.094	0.012	0.102	0.076	0.068	0.018
30	1000	N-Carpik	Normal	AH	0.121	0.016	0.093	0.012	0.100	0.060	0.067	0.016
10	100	P-Carpik	Normal	MLE	0.327	0.068	0.262	0.058	0.281	0.210	0.202	0.087
10	100	P-Carpik	Normal	MAP	0.301	0.067	0.240	0.057	0.216	0.068	0.170	0.047
10	100	P-Carpik	Normal	DE	0.655	2.646	0.488	0.905	0.410	30.398	0.275	9.610
10	100	P-Carpik	Normal	AH	0.528	0.970	0.406	0.372	0.311	1.534	0.234	0.506
10	500	P-Carpik	Normal	MLE	0.191	0.051	0.150	0.038	0.125	0.070	0.093	0.030
10	500	P-Carpik	Normal	MAP	0.181	0.052	0.143	0.038	0.115	0.045	0.087	0.023
10	500	P-Carpik	Normal	DE	0.224	0.153	0.180	0.073	0.136	0.091	0.099	0.037
10	500	P-Carpik	Normal	AH	0.211	0.098	0.164	0.055	0.122	0.061	0.091	0.028
10	1000	P-Carpik	Normal	MLE	0.140	0.043	0.108	0.031	0.093	0.058	0.069	0.026
10	1000	P-Carpik	Normal	MAP	0.133	0.040	0.105	0.030	0.088	0.045	0.065	0.023
10	1000	P-Carpik	Normal	DE	0.149	0.065	0.119	0.041	0.100	0.073	0.071	0.031
10	1000	P-Carpik	Normal	AH	0.149	0.047	0.113	0.031	0.091	0.052	0.065	0.025
20	100	P-Carpik	Normal	MLE	0.291	0.057	0.234	0.043	0.274	0.386	0.206	0.094
20	100	P-Carpik	Normal	MAP	0.287	0.056	0.226	0.041	0.231	0.047	0.178	0.034
20	100	P-Carpik	Normal	DE	0.498	0.583	0.368	0.157	0.378	2.044	0.255	0.463

20	100	P-Carpik	Normal	AH	0.452	0.338	0.340	0.104	0.318	1.424	0.235	0.325
20	500	P-Carpik	Normal	MLE	0.158	0.030	0.123	0.023	0.134	0.083	0.093	0.028
20	500	P-Carpik	Normal	MAP	0.155	0.029	0.120	0.023	0.115	0.036	0.086	0.019
20	500	P-Carpik	Normal	DE	0.176	0.039	0.135	0.026	0.142	0.109	0.095	0.033
20	500	P-Carpik	Normal	AH	0.170	0.036	0.133	0.026	0.125	0.099	0.092	0.031
20	1000	P-Carpik	Normal	MLE	0.117	0.024	0.093	0.019	0.099	0.073	0.069	0.020
20	1000	P-Carpik	Normal	MAP	0.113	0.023	0.089	0.018	0.088	0.039	0.067	0.014
20	1000	P-Carpik	Normal	DE	0.123	0.025	0.096	0.019	0.101	0.080	0.071	0.021
20	1000	P-Carpik	Normal	AH	0.121	0.026	0.094	0.019	0.097	0.075	0.070	0.020
30	100	P-Carpik	Normal	MLE	0.299	0.051	0.231	0.036	0.255	0.323	0.186	0.068
30	100	P-Carpik	Normal	MAP	0.299	0.051	0.231	0.037	0.240	0.043	0.177	0.024
30	100	P-Carpik	Normal	DE	0.445	1.081	0.334	0.222	0.353	1.384	0.231	0.284
30	100	P-Carpik	Normal	AH	0.431	0.310	0.326	0.096	0.304	1.173	0.219	0.242
30	500	P-Carpik	Normal	MLE	0.157	0.022	0.126	0.017	0.133	0.045	0.092	0.016
30	500	P-Carpik	Normal	MAP	0.159	0.022	0.124	0.017	0.118	0.032	0.089	0.015
30	500	P-Carpik	Normal	DE	0.172	0.028	0.136	0.019	0.142	0.082	0.097	0.021
30	500	P-Carpik	Normal	AH	0.166	0.026	0.131	0.019	0.137	0.062	0.095	0.018
30	1000	P-Carpik	Normal	MLE	0.113	0.018	0.089	0.013	0.090	0.039	0.066	0.011
30	1000	P-Carpik	Normal	MAP	0.112	0.019	0.087	0.013	0.087	0.027	0.064	0.010
30	1000	P-Carpik	Normal	DE	0.117	0.021	0.092	0.014	0.098	0.050	0.067	0.013
30	1000	P-Carpik	Normal	AH	0.116	0.020	0.091	0.014	0.094	0.044	0.066	0.012
10	100	Normal	N-Carpik	MLE	0.449	0.068	0.326	0.061	0.493	0.192	0.301	0.075
10	100	Normal	N-Carpik	MAP	0.377	0.071	0.295	0.061	0.316	0.084	0.234	0.047
10	100	Normal	N-Carpik	DE	0.840	4.943	0.637	1.660	0.583	1.147	0.387	0.402
10	100	Normal	N-Carpik	AH	0.953	1.778	0.673	0.624	0.665	7.230	0.388	2.298
10	500	Normal	N-Carpik	MLE	0.439	0.075	0.292	0.056	0.386	0.139	0.226	0.051
10	500	Normal	N-Carpik	MAP	0.407	0.075	0.279	0.055	0.328	0.088	0.199	0.037
10	500	Normal	N-Carpik	DE	0.349	0.144	0.249	0.079	0.243	0.166	0.161	0.069
10	500	Normal	N-Carpik	AH	0.486	0.136	0.334	0.076	0.362	0.174	0.228	0.064
10	1000	Normal	N-Carpik	MLE	0.451	0.082	0.303	0.047	0.384	0.113	0.215	0.043
10	1000	Normal	N-Carpik	MAP	0.435	0.082	0.297	0.046	0.347	0.083	0.198	0.035
10	1000	Normal	N-Carpik	DE	0.302	0.271	0.205	0.111	0.214	0.140	0.131	0.057
10	1000	Normal	N-Carpik	AH	0.484	0.113	0.323	0.056	0.369	0.124	0.214	0.048
20	100	Normal	N-Carpik	MLE	0.324	0.062	0.256	0.043	0.304	0.096	0.218	0.045
20	100	Normal	N-Carpik	MAP	0.316	0.062	0.246	0.042	0.242	0.045	0.195	0.035
20	100	Normal	N-Carpik	DE	0.554	0.983	0.412	0.250	0.427	1.659	0.284	0.385
20	100	Normal	N-Carpik	AH	0.526	0.699	0.401	0.184	0.437	5.938	0.281	1.328
20	500	Normal	N-Carpik	MLE	0.238	0.054	0.184	0.032	0.235	0.077	0.155	0.029
20	500	Normal	N-Carpik	MAP	0.225	0.054	0.177	0.032	0.210	0.045	0.140	0.023
20	500	Normal	N-Carpik	DE	0.204	0.039	0.160	0.030	0.178	0.104	0.122	0.034
20	500	Normal	N-Carpik	AH	0.266	0.079	0.198	0.039	0.240	0.118	0.158	0.038
20	1000	Normal	N-Carpik	MLE	0.228	0.037	0.173	0.024	0.229	0.056	0.138	0.021
20	1000	Normal	N-Carpik	MAP	0.222	0.038	0.169	0.024	0.206	0.040	0.128	0.018
20	1000	Normal	N-Carpik	DE	0.145	0.030	0.114	0.022	0.122	0.057	0.083	0.020
20	1000	Normal	N-Carpik	AH	0.241	0.043	0.179	0.026	0.230	0.064	0.139	0.023
30	100	Normal	N-Carpik	MLE	0.345	0.050	0.268	0.040	0.296	0.070	0.212	0.036
30	100	Normal	N-Carpik	MAP	0.327	0.048	0.250	0.038	0.246	0.037	0.193	0.027

30	100	Normal	N-Carpik	DE	0.491	0.334	0.371	0.096	0.377	0.714	0.252	0.157
30	100	Normal	N-Carpik	AH	0.466	0.163	0.355	0.070	0.401	1.530	0.259	0.300
30	500	Normal	N-Carpik	MLE	0.254	0.034	0.193	0.025	0.228	0.066	0.144	0.026
30	500	Normal	N-Carpik	MAP	0.244	0.032	0.187	0.024	0.197	0.041	0.133	0.021
30	500	Normal	N-Carpik	DE	0.211	0.038	0.159	0.027	0.167	0.084	0.114	0.026
30	500	Normal	N-Carpik	AH	0.263	0.039	0.200	0.027	0.229	0.100	0.144	0.033
30	1000	Normal	N-Carpik	MLE	0.220	0.027	0.166	0.019	0.197	0.049	0.123	0.018
30	1000	Normal	N-Carpik	MAP	0.215	0.026	0.164	0.018	0.181	0.037	0.117	0.016
30	1000	Normal	N-Carpik	DE	0.154	0.027	0.118	0.017	0.125	0.049	0.083	0.016
30	1000	Normal	N-Carpik	AH	0.220	0.029	0.168	0.020	0.199	0.058	0.122	0.020
10	100	N-Carpik	N-Carpik	MLE	0.367	0.083	0.292	0.071	0.250	0.076	0.192	0.050
10	100	N-Carpik	N-Carpik	MAP	0.381	0.084	0.295	0.072	0.269	0.067	0.197	0.045
10	100	N-Carpik	N-Carpik	DE	0.754	2.298	0.562	0.817	0.333	0.902	0.243	0.305
10	100	N-Carpik	N-Carpik	AH	0.832	1.624	0.616	0.571	0.374	0.912	0.264	0.315
10	500	N-Carpik	N-Carpik	MLE	0.352	0.101	0.250	0.055	0.252	0.046	0.146	0.026
10	500	N-Carpik	N-Carpik	MAP	0.362	0.102	0.253	0.055	0.255	0.044	0.147	0.025
10	500	N-Carpik	N-Carpik	DE	0.284	0.766	0.226	0.250	0.170	0.085	0.118	0.041
10	500	N-Carpik	N-Carpik	AH	0.465	0.283	0.315	0.106	0.286	0.049	0.162	0.029
10	1000	N-Carpik	N-Carpik	MLE	0.376	0.108	0.246	0.055	0.255	0.038	0.136	0.020
10	1000	N-Carpik	N-Carpik	MAP	0.380	0.108	0.248	0.054	0.256	0.037	0.136	0.020
10	1000	N-Carpik	N-Carpik	DE	0.204	0.134	0.161	0.063	0.129	0.072	0.086	0.033
10	1000	N-Carpik	N-Carpik	AH	0.425	0.160	0.271	0.068	0.270	0.039	0.142	0.021
20	100	N-Carpik	N-Carpik	MLE	0.324	0.067	0.258	0.047	0.272	0.074	0.200	0.039
20	100	N-Carpik	N-Carpik	MAP	0.327	0.066	0.260	0.045	0.286	0.057	0.203	0.033
20	100	N-Carpik	N-Carpik	DE	0.509	1.176	0.378	0.308	0.328	0.275	0.240	0.088
20	100	N-Carpik	N-Carpik	AH	0.571	0.846	0.404	0.231	0.376	0.195	0.255	0.077
20	500	N-Carpik	N-Carpik	MLE	0.241	0.051	0.184	0.029	0.224	0.044	0.133	0.020
20	500	N-Carpik	N-Carpik	MAP	0.245	0.052	0.187	0.029	0.231	0.041	0.135	0.019
20	500	N-Carpik	N-Carpik	DE	0.195	0.038	0.158	0.029	0.127	0.033	0.094	0.018
20	500	N-Carpik	N-Carpik	AH	0.272	0.076	0.204	0.035	0.243	0.045	0.140	0.021
20	1000	N-Carpik	N-Carpik	MLE	0.224	0.042	0.164	0.024	0.218	0.034	0.120	0.015
20	1000	N-Carpik	N-Carpik	MAP	0.226	0.042	0.165	0.024	0.220	0.032	0.121	0.014
20	1000	N-Carpik	N-Carpik	DE	0.140	0.034	0.110	0.022	0.092	0.031	0.069	0.016
20	1000	N-Carpik	N-Carpik	AH	0.238	0.050	0.173	0.026	0.225	0.034	0.123	0.015
30	100	N-Carpik	N-Carpik	MLE	0.350	0.057	0.266	0.041	0.286	0.054	0.202	0.030
30	100	N-Carpik	N-Carpik	MAP	0.349	0.056	0.266	0.040	0.303	0.045	0.204	0.025
30	100	N-Carpik	N-Carpik	DE	0.517	1.091	0.379	0.221	0.316	0.252	0.218	0.065
30	100	N-Carpik	N-Carpik	AH	0.516	0.842	0.389	0.185	0.349	0.164	0.233	0.051
30	500	N-Carpik	N-Carpik	MLE	0.241	0.040	0.179	0.027	0.200	0.043	0.121	0.018
30	500	N-Carpik	N-Carpik	MAP	0.242	0.040	0.181	0.027	0.212	0.040	0.123	0.017
30	500	N-Carpik	N-Carpik	DE	0.196	0.031	0.151	0.025	0.125	0.039	0.092	0.017
30	500	N-Carpik	N-Carpik	AH	0.256	0.048	0.186	0.030	0.203	0.044	0.122	0.019
30	1000	N-Carpik	N-Carpik	MLE	0.223	0.034	0.164	0.021	0.197	0.037	0.110	0.014
30	1000	N-Carpik	N-Carpik	MAP	0.225	0.034	0.164	0.021	0.202	0.035	0.111	0.014
30	1000	N-Carpik	N-Carpik	DE	0.142	0.023	0.111	0.018	0.094	0.026	0.065	0.012
30	1000	N-Carpik	N-Carpik	AH	0.230	0.037	0.167	0.022	0.199	0.037	0.110	0.015
10	100	P-Carpik	N-Carpik	MLE	0.373	0.088	0.304	0.068	0.427	0.609	0.275	0.202

10	100	P-Carpik	N-Carpik	MAP	0.337	0.095	0.271	0.067	0.239	0.072	0.184	0.051
10	100	P-Carpik	N-Carpik	DE	0.766	3.749	0.567	1.263	0.511	5.959	0.346	1.901
10	100	P-Carpik	N-Carpik	AH	0.637	0.982	0.479	0.378	0.514	10.056	0.311	3.183
10	500	P-Carpik	N-Carpik	MLE	0.299	0.058	0.220	0.042	0.490	0.224	0.222	0.074
10	500	P-Carpik	N-Carpik	MAP	0.245	0.057	0.193	0.042	0.265	0.099	0.152	0.038
10	500	P-Carpik	N-Carpik	DE	0.341	0.119	0.245	0.076	0.344	7.337	0.191	2.323
10	500	P-Carpik	N-Carpik	AH	0.323	0.087	0.234	0.055	0.450	1.356	0.212	0.429
10	1000	P-Carpik	N-Carpik	MLE	0.271	0.055	0.200	0.041	0.501	0.233	0.222	0.077
10	1000	P-Carpik	N-Carpik	MAP	0.235	0.049	0.182	0.040	0.345	0.121	0.168	0.044
10	1000	P-Carpik	N-Carpik	DE	0.277	0.087	0.195	0.051	0.425	0.384	0.193	0.125
10	1000	P-Carpik	N-Carpik	AH	0.280	0.069	0.202	0.046	0.482	0.383	0.215	0.124
20	100	P-Carpik	N-Carpik	MLE	0.311	0.052	0.251	0.041	0.353	1.179	0.232	0.277
20	100	P-Carpik	N-Carpik	MAP	0.303	0.049	0.237	0.039	0.225	0.045	0.178	0.034
20	100	P-Carpik	N-Carpik	DE	0.531	1.835	0.401	0.432	0.495	0.345	0.316	55.759
20	100	P-Carpik	N-Carpik	AH	0.495	0.522	0.371	0.143	0.466	5.036	0.277	1.133
20	500	P-Carpik	N-Carpik	MLE	0.221	0.036	0.171	0.030	0.292	0.142	0.158	0.041
20	500	P-Carpik	N-Carpik	MAP	0.204	0.036	0.156	0.031	0.154	0.047	0.113	0.023
20	500	P-Carpik	N-Carpik	DE	0.214	0.050	0.161	0.034	0.204	0.376	0.122	0.092
20	500	P-Carpik	N-Carpik	AH	0.240	0.046	0.181	0.035	0.282	0.521	0.153	0.125
20	1000	P-Carpik	N-Carpik	MLE	0.204	0.029	0.154	0.023	0.348	0.171	0.151	0.045
20	1000	P-Carpik	N-Carpik	MAP	0.183	0.025	0.143	0.021	0.181	0.060	0.112	0.023
20	1000	P-Carpik	N-Carpik	DE	0.165	0.043	0.122	0.024	0.205	0.347	0.107	0.081
20	1000	P-Carpik	N-Carpik	AH	0.208	0.037	0.159	0.025	0.345	0.406	0.155	0.096
30	100	P-Carpik	N-Carpik	MLE	0.322	0.045	0.249	0.036	0.322	0.466	0.206	0.096
30	100	P-Carpik	N-Carpik	MAP	0.312	0.047	0.239	0.036	0.254	0.040	0.180	0.024
30	100	P-Carpik	N-Carpik	DE	0.527	0.637	0.378	0.140	0.477	2.636	0.278	3.947
30	100	P-Carpik	N-Carpik	AH	0.489	0.179	0.358	0.069	0.468	3.524	0.274	0.654
30	500	P-Carpik	N-Carpik	MLE	0.220	0.027	0.166	0.020	0.245	0.103	0.134	0.030
30	500	P-Carpik	N-Carpik	MAP	0.205	0.026	0.157	0.019	0.174	0.040	0.112	0.018
30	500	P-Carpik	N-Carpik	DE	0.208	0.036	0.157	0.023	0.231	0.362	0.133	0.073
30	500	P-Carpik	N-Carpik	AH	0.234	0.033	0.178	0.023	0.271	0.495	0.141	0.099
30	1000	P-Carpik	N-Carpik	MLE	0.196	0.023	0.146	0.017	0.240	0.112	0.122	0.028
30	1000	P-Carpik	N-Carpik	MAP	0.184	0.023	0.138	0.016	0.171	0.040	0.100	0.015
30	1000	P-Carpik	N-Carpik	DE	0.153	0.027	0.118	0.017	0.180	0.255	0.093	0.051
30	1000	P-Carpik	N-Carpik	AH	0.201	0.025	0.149	0.018	0.250	0.291	0.122	0.059
10	100	Normal	P-Carpik	MLE	0.397	0.081	0.320	0.062	0.365	0.166	0.255	0.074
10	100	Normal	P-Carpik	MAP	0.357	0.077	0.292	0.058	0.269	0.076	0.209	0.052
10	100	Normal	P-Carpik	DE	0.927	4.698	0.645	1.545	0.426	2.331	0.309	0.752
10	100	Normal	P-Carpik	AH	1.178	1.994	0.776	0.688	0.464	9.832	0.311	3.116
10	500	Normal	P-Carpik	MLE	0.405	0.106	0.305	0.060	0.353	0.147	0.211	0.060
10	500	Normal	P-Carpik	MAP	0.388	0.106	0.293	0.058	0.299	0.090	0.185	0.043
10	500	Normal	P-Carpik	DE	0.335	0.396	0.240	0.143	0.234	0.202	0.156	0.078
10	500	Normal	P-Carpik	AH	0.532	0.425	0.367	0.151	0.343	0.209	0.214	0.080
10	1000	Normal	P-Carpik	MLE	0.468	0.119	0.318	0.062	0.353	0.113	0.199	0.046
10	1000	Normal	P-Carpik	MAP	0.456	0.119	0.312	0.061	0.314	0.081	0.182	0.036
10	1000	Normal	P-Carpik	DE	0.249	0.090	0.185	0.052	0.198	0.129	0.121	0.052
10	1000	Normal	P-Carpik	AH	0.567	0.243	0.353	0.098	0.346	0.135	0.199	0.053

20	100	Normal	P-Carpik	MLE	0.335	0.066	0.265	0.048	0.339	0.122	0.233	0.048
20	100	Normal	P-Carpik	MAP	0.303	0.069	0.241	0.048	0.255	0.048	0.191	0.033
20	100	Normal	P-Carpik	DE	0.547	1.095	0.402	0.281	0.418	1.161	0.288	0.271
20	100	Normal	P-Carpik	AH	0.502	0.573	0.375	0.172	0.418	2.137	0.287	0.488
20	500	Normal	P-Carpik	MLE	0.254	0.040	0.190	0.027	0.267	0.111	0.161	0.034
20	500	Normal	P-Carpik	MAP	0.236	0.038	0.182	0.027	0.223	0.055	0.145	0.025
20	500	Normal	P-Carpik	DE	0.214	0.046	0.162	0.029	0.171	0.138	0.120	0.039
20	500	Normal	P-Carpik	AH	0.267	0.051	0.199	0.032	0.265	0.205	0.163	0.054
20	1000	Normal	P-Carpik	MLE	0.223	0.039	0.165	0.026	0.233	0.090	0.142	0.029
20	1000	Normal	P-Carpik	MAP	0.215	0.037	0.161	0.025	0.203	0.056	0.133	0.023
20	1000	Normal	P-Carpik	DE	0.149	0.053	0.117	0.027	0.122	0.101	0.085	0.031
20	1000	Normal	P-Carpik	AH	0.227	0.044	0.170	0.028	0.224	0.103	0.141	0.032
30	100	Normal	P-Carpik	MLE	0.330	0.050	0.258	0.035	0.278	0.058	0.206	0.033
30	100	Normal	P-Carpik	MAP	0.320	0.048	0.250	0.034	0.252	0.038	0.195	0.028
30	100	Normal	P-Carpik	DE	0.495	0.437	0.382	0.106	0.381	1.238	0.267	0.237
30	100	Normal	P-Carpik	AH	0.484	0.378	0.370	0.098	0.395	1.010	0.260	0.207
30	500	Normal	P-Carpik	MLE	0.233	0.039	0.180	0.026	0.187	0.041	0.129	0.019
30	500	Normal	P-Carpik	MAP	0.229	0.038	0.175	0.026	0.169	0.027	0.122	0.016
30	500	Normal	P-Carpik	DE	0.196	0.028	0.151	0.022	0.141	0.056	0.103	0.019
30	500	Normal	P-Carpik	AH	0.252	0.047	0.192	0.029	0.195	0.061	0.135	0.022
30	1000	Normal	P-Carpik	MLE	0.208	0.024	0.163	0.019	0.166	0.039	0.112	0.016
30	1000	Normal	P-Carpik	MAP	0.204	0.024	0.161	0.018	0.153	0.028	0.108	0.014
30	1000	Normal	P-Carpik	DE	0.141	0.018	0.111	0.014	0.112	0.037	0.080	0.013
30	1000	Normal	P-Carpik	AH	0.214	0.026	0.167	0.019	0.168	0.046	0.113	0.018
10	100	N-Carpik	P-Carpik	MLE	0.375	0.076	0.292	0.070	0.459	0.265	0.294	0.101
10	100	N-Carpik	P-Carpik	MAP	0.319	0.077	0.257	0.066	0.228	0.073	0.181	0.050
10	100	N-Carpik	P-Carpik	DE	0.707	2.703	0.526	0.906	0.632	21.547	0.369	6.822
10	100	N-Carpik	P-Carpik	AH	0.624	0.756	0.472	0.280	0.580	10.600	0.347	3.354
10	500	N-Carpik	P-Carpik	MLE	0.292	0.061	0.212	0.046	0.475	0.211	0.227	0.070
10	500	N-Carpik	P-Carpik	MAP	0.239	0.055	0.187	0.042	0.270	0.099	0.157	0.038
10	500	N-Carpik	P-Carpik	DE	0.341	0.121	0.240	0.072	0.340	0.420	0.193	0.141
10	500	N-Carpik	P-Carpik	AH	0.314	0.083	0.234	0.054	0.473	0.416	0.220	0.135
10	1000	N-Carpik	P-Carpik	MLE	0.283	0.060	0.194	0.045	0.481	0.209	0.212	0.069
10	1000	N-Carpik	P-Carpik	MAP	0.245	0.055	0.177	0.043	0.334	0.111	0.169	0.040
10	1000	N-Carpik	P-Carpik	DE	0.301	0.096	0.195	0.058	0.429	0.297	0.193	0.099
10	1000	N-Carpik	P-Carpik	AH	0.281	0.074	0.203	0.050	0.472	0.299	0.207	0.098
20	100	N-Carpik	P-Carpik	MLE	0.312	0.059	0.250	0.045	0.333	0.135	0.230	0.052
20	100	N-Carpik	P-Carpik	MAP	0.293	0.059	0.234	0.044	0.240	0.047	0.189	0.035
20	100	N-Carpik	P-Carpik	DE	0.566	2.177	0.407	0.507	0.459	24.390	0.305	5.457
20	100	N-Carpik	P-Carpik	AH	0.481	0.676	0.360	0.183	0.403	7.636	0.280	1.729
20	500	N-Carpik	P-Carpik	MLE	0.229	0.033	0.176	0.027	0.347	0.160	0.170	0.044
20	500	N-Carpik	P-Carpik	MAP	0.203	0.030	0.162	0.025	0.171	0.043	0.119	0.022
20	500	N-Carpik	P-Carpik	DE	0.222	0.055	0.163	0.032	0.252	0.559	0.143	0.131
20	500	N-Carpik	P-Carpik	AH	0.243	0.044	0.183	0.032	0.361	0.553	0.183	0.130
20	1000	N-Carpik	P-Carpik	MLE	0.198	0.026	0.151	0.021	0.312	0.163	0.146	0.042
20	1000	N-Carpik	P-Carpik	MAP	0.182	0.024	0.141	0.020	0.170	0.056	0.107	0.021
20	1000	N-Carpik	P-Carpik	DE	0.160	0.042	0.124	0.023	0.203	0.357	0.100	0.082

20	1000	N-Carpik	P-Carpik	AH	0.203	0.033	0.155	0.023	0.327	0.589	0.149	0.135
30	100	N-Carpik	P-Carpik	MLE	0.332	0.046	0.263	0.034	0.321	0.102	0.211	0.038
30	100	N-Carpik	P-Carpik	MAP	0.314	0.048	0.248	0.034	0.256	0.039	0.190	0.024
30	100	N-Carpik	P-Carpik	DE	0.553	0.603	0.403	0.133	0.493	8.754	0.303	1.605
30	100	N-Carpik	P-Carpik	AH	0.496	0.212	0.367	0.074	0.525	3.097	0.293	7.144
30	500	N-Carpik	P-Carpik	MLE	0.219	0.028	0.167	0.021	0.260	0.101	0.136	0.031
30	500	N-Carpik	P-Carpik	MAP	0.205	0.027	0.156	0.021	0.174	0.042	0.112	0.018
30	500	N-Carpik	P-Carpik	DE	0.202	0.039	0.156	0.024	0.248	0.593	0.131	0.115
30	500	N-Carpik	P-Carpik	AH	0.231	0.034	0.176	0.024	0.309	0.509	0.149	0.103
30	1000	N-Carpik	P-Carpik	MLE	0.197	0.027	0.143	0.018	0.252	0.112	0.119	0.030
30	1000	N-Carpik	P-Carpik	MAP	0.184	0.025	0.139	0.018	0.167	0.047	0.096	0.017
30	1000	N-Carpik	P-Carpik	DE	0.153	0.029	0.115	0.017	0.185	0.253	0.094	0.052
30	1000	N-Carpik	P-Carpik	AH	0.201	0.031	0.147	0.020	0.276	0.263	0.125	0.057
10	100	P-Carpik	P-Carpik	MLE	0.357	0.088	0.281	0.074	0.264	0.081	0.192	0.050
10	100	P-Carpik	P-Carpik	MAP	0.366	0.094	0.285	0.078	0.279	0.069	0.196	0.047
10	100	P-Carpik	P-Carpik	DE	0.805	3.530	0.574	1.197	0.311	1.389	0.226	0.445
10	100	P-Carpik	P-Carpik	AH	1.665	1.850	1.018	0.804	0.566	2.762	0.349	0.957
10	500	P-Carpik	P-Carpik	MLE	0.353	0.118	0.252	0.066	0.255	0.052	0.146	0.032
10	500	P-Carpik	P-Carpik	MAP	0.360	0.119	0.256	0.066	0.259	0.050	0.146	0.032
10	500	P-Carpik	P-Carpik	DE	0.270	1.465	0.204	0.475	0.152	0.081	0.108	0.041
10	500	P-Carpik	P-Carpik	AH	0.483	0.404	0.303	0.146	0.287	0.055	0.162	0.035
10	1000	P-Carpik	P-Carpik	MLE	0.373	0.116	0.253	0.055	0.258	0.039	0.140	0.021
10	1000	P-Carpik	P-Carpik	MAP	0.377	0.116	0.255	0.054	0.259	0.038	0.140	0.020
10	1000	P-Carpik	P-Carpik	DE	0.206	0.095	0.161	0.054	0.121	0.065	0.086	0.031
10	1000	P-Carpik	P-Carpik	AH	0.444	0.187	0.284	0.075	0.275	0.040	0.147	0.022
20	100	P-Carpik	P-Carpik	MLE	0.312	0.060	0.250	0.043	0.273	0.074	0.200	0.040
20	100	P-Carpik	P-Carpik	MAP	0.315	0.062	0.248	0.043	0.270	0.058	0.196	0.033
20	100	P-Carpik	P-Carpik	DE	0.501	1.068	0.391	0.262	0.330	0.817	0.239	0.199
20	100	P-Carpik	P-Carpik	AH	0.503	0.817	0.399	0.207	0.358	0.351	0.241	0.099
20	500	P-Carpik	P-Carpik	MLE	0.241	0.061	0.184	0.037	0.233	0.044	0.132	0.020
20	500	P-Carpik	P-Carpik	MAP	0.245	0.062	0.186	0.037	0.238	0.041	0.133	0.019
20	500	P-Carpik	P-Carpik	DE	0.205	0.041	0.159	0.031	0.139	0.036	0.098	0.022
20	500	P-Carpik	P-Carpik	AH	0.278	0.092	0.205	0.044	0.253	0.046	0.139	0.021
20	1000	P-Carpik	P-Carpik	MLE	0.228	0.048	0.163	0.027	0.226	0.033	0.122	0.016
20	1000	P-Carpik	P-Carpik	MAP	0.231	0.049	0.164	0.027	0.228	0.032	0.122	0.016
20	1000	P-Carpik	P-Carpik	DE	0.145	0.025	0.112	0.019	0.094	0.028	0.071	0.015
20	1000	P-Carpik	P-Carpik	AH	0.245	0.056	0.173	0.029	0.233	0.033	0.125	0.017
30	100	P-Carpik	P-Carpik	MLE	0.349	0.064	0.265	0.046	0.270	0.054	0.200	0.030
30	100	P-Carpik	P-Carpik	MAP	0.347	0.064	0.264	0.046	0.286	0.046	0.201	0.027
30	100	P-Carpik	P-Carpik	DE	0.513	0.993	0.378	0.221	0.311	1.715	0.224	0.317
30	100	P-Carpik	P-Carpik	AH	0.493	0.625	0.370	0.160	0.317	0.654	0.225	0.125
30	500	P-Carpik	P-Carpik	MLE	0.241	0.043	0.183	0.028	0.194	0.046	0.122	0.019
30	500	P-Carpik	P-Carpik	MAP	0.242	0.043	0.184	0.027	0.207	0.043	0.124	0.018
30	500	P-Carpik	P-Carpik	DE	0.194	0.032	0.152	0.023	0.125	0.035	0.091	0.015
30	500	P-Carpik	P-Carpik	AH	0.252	0.055	0.191	0.030	0.196	0.048	0.123	0.020
30	1000	P-Carpik	P-Carpik	MLE	0.213	0.034	0.157	0.021	0.185	0.036	0.102	0.015
30	1000	P-Carpik	P-Carpik	MAP	0.214	0.034	0.157	0.021	0.190	0.034	0.104	0.015

30	1000	P-Carpik	P-Carpik	DE	0.145	0.023	0.113	0.018	0.088	0.025	0.066	0.012
30	1000	P-Carpik	P-Carpik	AH	0.218	0.037	0.159	0.022	0.185	0.037	0.103	0.015
10	100	Normal	Normal	NBE	0.544	0.052	0.427	0.042	0.360	0.064	0.295	0.054
10	500	Normal	Normal	NBE	0.541	0.049	0.421	0.044	0.381	0.055	0.308	0.051
10	1000	Normal	Normal	NBE	0.533	0.049	0.426	0.043	0.373	0.056	0.295	0.050
20	100	Normal	Normal	NBE	0.308	0.031	0.246	0.026	0.379	0.046	0.295	0.039
20	500	Normal	Normal	NBE	0.308	0.029	0.246	0.027	0.385	0.047	0.300	0.044
20	1000	Normal	Normal	NBE	0.312	0.030	0.249	0.028	0.398	0.042	0.303	0.031
30	100	Normal	Normal	NBE	0.362	0.029	0.286	0.026	0.422	0.031	0.333	0.028
30	500	Normal	Normal	NBE	0.363	0.025	0.288	0.023	0.424	0.025	0.335	0.021
30	1000	Normal	Normal	NBE	0.363	0.027	0.292	0.022	0.431	0.032	0.339	0.027
10	100	N-Carpik	Normal	NBE	0.491	0.061	0.399	0.056	0.500	0.066	0.371	0.058
10	500	N-Carpik	Normal	NBE	0.498	0.061	0.402	0.055	0.488	0.039	0.340	0.047
10	1000	N-Carpik	Normal	NBE	0.483	0.054	0.392	0.048	0.505	0.059	0.343	0.045
20	100	N-Carpik	Normal	NBE	0.296	0.031	0.242	0.029	0.480	0.043	0.336	0.034
20	500	N-Carpik	Normal	NBE	0.293	0.033	0.242	0.031	0.470	0.032	0.314	0.030
20	1000	N-Carpik	Normal	NBE	0.311	0.034	0.251	0.030	0.481	0.041	0.319	0.032
30	100	N-Carpik	Normal	NBE	0.335	0.030	0.269	0.027	0.498	0.038	0.337	0.033
30	500	N-Carpik	Normal	NBE	0.322	0.032	0.259	0.027	0.500	0.030	0.327	0.024
30	1000	N-Carpik	Normal	NBE	0.322	0.031	0.264	0.028	0.505	0.033	0.334	0.027
10	100	P-Carpik	Normal	NBE	0.503	0.052	0.419	0.051	0.547	0.068	0.380	0.057
10	500	P-Carpik	Normal	NBE	0.499	0.051	0.401	0.052	0.548	0.044	0.381	0.050
10	1000	P-Carpik	Normal	NBE	0.489	0.051	0.394	0.052	0.547	0.056	0.385	0.045
20	100	P-Carpik	Normal	NBE	0.302	0.036	0.250	0.031	0.529	0.045	0.353	0.035
20	500	P-Carpik	Normal	NBE	0.300	0.031	0.249	0.030	0.519	0.042	0.340	0.038
20	1000	P-Carpik	Normal	NBE	0.299	0.033	0.246	0.031	0.523	0.041	0.342	0.038
30	100	P-Carpik	Normal	NBE	0.372	0.028	0.283	0.025	0.538	0.037	0.352	0.029
30	500	P-Carpik	Normal	NBE	0.378	0.031	0.284	0.026	0.547	0.029	0.355	0.023
30	1000	P-Carpik	Normal	NBE	0.383	0.026	0.290	0.024	0.546	0.026	0.347	0.022
10	100	Normal	N-Carpik	NBE	0.551	0.053	0.441	0.048	0.352	0.060	0.289	0.051
10	500	Normal	N-Carpik	NBE	0.534	0.053	0.423	0.045	0.370	0.062	0.289	0.051
10	1000	Normal	N-Carpik	NBE	0.540	0.046	0.426	0.043	0.362	0.057	0.296	0.048
20	100	Normal	N-Carpik	NBE	0.311	0.032	0.251	0.027	0.398	0.049	0.311	0.041
20	500	Normal	N-Carpik	NBE	0.302	0.032	0.242	0.031	0.400	0.037	0.303	0.033
20	1000	Normal	N-Carpik	NBE	0.310	0.033	0.251	0.031	0.405	0.039	0.303	0.033
30	100	Normal	N-Carpik	NBE	0.361	0.028	0.283	0.022	0.425	0.035	0.336	0.028
30	500	Normal	N-Carpik	NBE	0.360	0.028	0.288	0.024	0.424	0.029	0.330	0.028
30	1000	Normal	N-Carpik	NBE	0.366	0.024	0.294	0.023	0.420	0.030	0.328	0.025
10	100	N-Carpik	N-Carpik	NBE	0.491	0.050	0.405	0.048	0.497	0.061	0.368	0.053
10	500	N-Carpik	N-Carpik	NBE	0.485	0.056	0.393	0.051	0.503	0.060	0.347	0.050
10	1000	N-Carpik	N-Carpik	NBE	0.485	0.059	0.406	0.054	0.519	0.059	0.359	0.045
20	100	N-Carpik	N-Carpik	NBE	0.299	0.038	0.243	0.032	0.470	0.042	0.324	0.034
20	500	N-Carpik	N-Carpik	NBE	0.315	0.032	0.252	0.030	0.469	0.044	0.313	0.033
20	1000	N-Carpik	N-Carpik	NBE	0.311	0.035	0.255	0.032	0.479	0.038	0.321	0.031
30	100	N-Carpik	N-Carpik	NBE	0.331	0.034	0.268	0.030	0.502	0.030	0.339	0.025
30	500	N-Carpik	N-Carpik	NBE	0.331	0.033	0.266	0.029	0.514	0.030	0.330	0.026
30	1000	N-Carpik	N-Carpik	NBE	0.316	0.029	0.254	0.028	0.504	0.030	0.326	0.025

10	100	P-Carpik	N-Carpik	NBE	0.493	0.050	0.407	0.048	0.548	0.066	0.390	0.064
10	500	P-Carpik	N-Carpik	NBE	0.516	0.058	0.421	0.054	0.566	0.064	0.400	0.052
10	1000	P-Carpik	N-Carpik	NBE	0.521	0.053	0.413	0.051	0.549	0.061	0.391	0.054
20	100	P-Carpik	N-Carpik	NBE	0.301	0.033	0.246	0.031	0.509	0.044	0.349	0.039
20	500	P-Carpik	N-Carpik	NBE	0.315	0.030	0.262	0.029	0.525	0.043	0.347	0.039
20	1000	P-Carpik	N-Carpik	NBE	0.314	0.033	0.260	0.030	0.526	0.040	0.347	0.034
30	100	P-Carpik	N-Carpik	NBE	0.368	0.030	0.282	0.027	0.557	0.037	0.365	0.035
30	500	P-Carpik	N-Carpik	NBE	0.376	0.030	0.283	0.026	0.544	0.031	0.345	0.029
30	1000	P-Carpik	N-Carpik	NBE	0.383	0.025	0.292	0.026	0.542	0.030	0.349	0.027
10	100	Normal	P-Carpik	NBE	0.539	0.047	0.430	0.041	0.359	0.080	0.292	0.069
10	500	Normal	P-Carpik	NBE	0.538	0.048	0.430	0.039	0.375	0.059	0.306	0.053
10	1000	Normal	P-Carpik	NBE	0.547	0.053	0.438	0.044	0.383	0.052	0.305	0.047
20	100	Normal	P-Carpik	NBE	0.314	0.031	0.252	0.029	0.389	0.043	0.311	0.035
20	500	Normal	P-Carpik	NBE	0.311	0.032	0.248	0.029	0.403	0.037	0.304	0.031
20	1000	Normal	P-Carpik	NBE	0.318	0.029	0.254	0.028	0.401	0.036	0.304	0.030
30	100	Normal	P-Carpik	NBE	0.357	0.024	0.281	0.021	0.432	0.038	0.342	0.032
30	500	Normal	P-Carpik	NBE	0.370	0.029	0.295	0.024	0.427	0.030	0.334	0.026
30	1000	Normal	P-Carpik	NBE	0.372	0.025	0.290	0.024	0.436	0.032	0.341	0.023
10	100	N-Carpik	P-Carpik	NBE	0.477	0.057	0.399	0.052	0.516	0.067	0.381	0.062
10	500	N-Carpik	P-Carpik	NBE	0.476	0.052	0.396	0.051	0.499	0.064	0.360	0.052
10	1000	N-Carpik	P-Carpik	NBE	0.472	0.050	0.395	0.046	0.514	0.061	0.361	0.052
20	100	N-Carpik	P-Carpik	NBE	0.296	0.039	0.240	0.031	0.494	0.044	0.344	0.037
20	500	N-Carpik	P-Carpik	NBE	0.307	0.032	0.246	0.031	0.493	0.045	0.327	0.038
20	1000	N-Carpik	P-Carpik	NBE	0.307	0.033	0.250	0.032	0.480	0.041	0.321	0.034
30	100	N-Carpik	P-Carpik	NBE	0.323	0.029	0.260	0.025	0.506	0.039	0.342	0.032
30	500	N-Carpik	P-Carpik	NBE	0.326	0.030	0.263	0.029	0.501	0.031	0.328	0.025
30	1000	N-Carpik	P-Carpik	NBE	0.326	0.027	0.264	0.025	0.508	0.033	0.335	0.025
10	100	P-Carpik	P-Carpik	NBE	0.487	0.059	0.399	0.056	0.530	0.061	0.377	0.053
10	500	P-Carpik	P-Carpik	NBE	0.494	0.049	0.395	0.044	0.530	0.065	0.364	0.047
10	1000	P-Carpik	P-Carpik	NBE	0.490	0.045	0.392	0.045	0.542	0.060	0.368	0.041
20	100	P-Carpik	P-Carpik	NBE	0.297	0.031	0.243	0.029	0.530	0.044	0.358	0.040
20	500	P-Carpik	P-Carpik	NBE	0.302	0.035	0.251	0.031	0.520	0.044	0.331	0.034
20	1000	P-Carpik	P-Carpik	NBE	0.304	0.031	0.252	0.029	0.534	0.043	0.342	0.036
30	100	P-Carpik	P-Carpik	NBE	0.367	0.028	0.278	0.027	0.553	0.035	0.367	0.031
30	500	P-Carpik	P-Carpik	NBE	0.376	0.024	0.282	0.022	0.553	0.029	0.353	0.024
30	1000	P-Carpik	P-Carpik	NBE	0.390	0.027	0.295	0.026	0.560	0.029	0.352	0.027