

T.C.
ÇUKUROVA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

**TAHMİN MODELLERİNİN PERFORMANS
KARŞILAŞTIRMASINDA KARAR EĞRİSİ
ANALİZİNİN KULLANIMI**

Duygu KORKMAZ YALÇIN

BİYOİSTATİSTİK DOKTORA PROGRAMI
DOKTORA TEZİ

DANIŞMANI
Doç. Dr. İlker ÜNAL

ADANA-2025

T.C.
ÇUKUROVA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

**TAHMİN MODELLERİNİN PERFORMANS
KARŞILAŞTIRMASINDA KARAR EĞRİSİ
ANALİZİNİN KULLANIMI**

Duygu KORKMAZ YALÇIN

BİYOİSTATİSTİK DOKTORA PROGRAMI
DOKTORA TEZİ

DANIŞMANI
Doç. Dr. İlker ÜNAL

ADANA-2025

ETİK BEYAN

Çukurova Üniversitesi Bilimsel Araştırma ve Yayın Etiği Yönergesini okuduğumu ve anladığımı ve Çukurova Üniversitesi Sağlık Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

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

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim. 28/02/2025

Duygu KORKMAZ YALÇIN

TEŞEKKÜR

Doktora eğitimim süresince bana rehberlik eden, bilimsel yaklaşımıyla yolumu aydınlatan her aşamada değerli bilgisi ve deneyimiyle desteklerini ve anlayışını esirgemeyen kıymetli hocam ve danışmanım Doç. Dr. İlker ÜNAL'a en derin şükran ve saygılarımı sunarım. Tez izleme komitesi dahil olmak üzere doktora eğitimimin tüm süreçlerinde bana ilham veren, her daim desteği ve vizyonu ile bilimsel ve manevi katkı sağlayan Anabilim Dalı Başkanımız Prof. Dr. Gülşah SEYDAOĞLU'na, almış olduğum teorik istatistik dersleri ile tez çalışmamın ilerleme raporlarında değerli bilgilerini esirgemeyen Prof. Dr. M. Revan ÖZKALE ATICIOĞLU'na teşekkürlerimi sunarım. Ayrıca lisansüstü eğitimimin tüm safhalarında akademik ve manevi mentörlüğünden dolayı Prof. Dr. Sıddık KESKİN'e, doktora eğitim sürecimde bilgi ve tecrübesini esirgemeyen Prof. Dr. Z. Nazan ALPARSLAN'a, bilgileri ile tezime katkıda bulunan Doç. Dr. Yaşar SERTDEMİR'e ayrıca eğitim süreçlerime olan desteklerinden dolayı Van Yüzüncü Yıl Üniversitesi Tıp Fakültesi Dekanı Prof. Dr. Ufuk ÇOBANOĞLU ve Dekan Yardımcıları Prof. Dr. Erbil KARAMAN ve Doç. Dr. Tahir ÇAKIR'a, simülasyon çalışmalarımın yürütülebilmesi için bilgisayar alt yapı olanağı sağlayan İnnovan Girişimcilik Merkezi Müdürü Dr. Öğr. Üyesi Erol KINA'ya her daim bana güç veren ve motive eden çok kıymetli hocam Dr. Mehmet Emin LAYIK'a, desteklerinden dolayı Dr. Öğr. Üyesi Yusuf Kemal ARSLAN'a, tez çalışmamın her aşamasında bana manevi destek sağlayan değerli arkadaşlarım Öğr. Gör. Aslı BOZ'a ve Erdal GÜVENER'e, yine süreç içerisinde desteklerinden dolayı Çukurova Üniversitesi Biyoistatistik Anabilim Dalı Araştırma Görevlisi arkadaşlarıma, ilgilerinden dolayı Çukurova Üniversitesi Sağlık Bilimleri Enstitüsüne, her daim yanımda olduklarını derinden hissettiren sevgili anneme, babama, kardeşlerime, kıymetli eşime ve sabırlarından dolayı tüm arkadaşlarıma teşekkürlerimi sunarım.

İÇİNDEKİLER

KABUL VE ONAY	ii
TEŞEKKÜR.....	iv
İÇİNDEKİLER	v
ŞEKİLLER DİZİNİ	vii
ÇİZELGELER DİZİNİ	viii
SİMGELER VE KISALTMALAR	ix
ÖZET	x
ABSTRACT.....	xi
1. GİRİŞ	1
2. GENEL BİLGİLER	6
2.1. Tahmin Modelleri	6
2.2. Modellerin Performanslarının Değerlendirilmesi	7
2.2.1. Belirleyicilik Katsayısı (R^2).....	7
2.2.2. Brier Skoru.....	8
2.2.3. Kalibrasyon Grafiği	8
2.2.4. Diskriminasyon.....	9
2.2.4.1. ROC Eğrileri	10
2.2.4.2. Diskriminasyon Eğimi	10
2.3. Karar Verme Süreci	10
2.4. Karar Eğrisi Analizi	12
2.4.1. Karar Eğrisi Analizi ile İlgili Yapılan Çalışmalar	12
2.4.2. Eşik Olasılılık	13
2.4.3. Karar Eğrisi ve Net Fayda	14
3. GEREÇ VE YÖNTEM.....	16
3.1. Gereç	16
3.1.1. Gerçek Veri Çalışması	16
3.1.2. Simülasyon Veri Çalışması.....	17
3.2. Yöntem.....	19
3.2.1. Literatürde Mevcut Yöntem (z testi)	19
3.2.2. ANOVA/Kruskal Wallis Yaklaşımı	23

4. BULGULAR.....	27
4.1. Simüle Veriler için Elde Edilen Bulgular	27
4.2. Gerçek Veriler için Elde Edilen Bulgular	45
4.2.1. Pima Veri Setinden Elde Edilen Bulgular	45
4.2.2. Prostat Veri Seti ile Elde Edilen Bulgular	47
5. TARTIŞMA	49
6. SONUÇLAR VE ÖNERİLER.....	54
7. KAYNAKLAR	55
EKLER.....	62
EK-1: Simülasyon Çalışması için Basamaklar	62
EK-2: Farklı Simülasyon Senaryolarına Ait Bulgular	63
ÖZGEÇMİŞ	87

ŞEKİLLER DİZİNİ

Şekil No:

Sayfa No:

Şekil 2.1. Tedavi için bir karar ağacı.....	14
Şekil 4.1. Senaryo 1 için etki büyüklüğü küçük iken modellere göre net faydaların dağılımı.....	27
Şekil 4.2. Senaryo 1 için etki büyüklüğü büyük iken modellere göre net faydaların dağılımı.....	28
Şekil 4.3. Senaryo 2 için etki büyüklüğü büyük iken modellere göre net faydaların dağılımı.....	29
Şekil 4.4. Senaryo 3 için etki büyüklüğü büyük iken modellere göre net faydaların dağılımı.....	30
Şekil 4.5. Senaryo 4 için etki büyüklüğü büyük iken modellere göre net faydaların dağılımı.....	31

ÇİZELGELER DİZİNİ

<u>Çizelge No:</u>	<u>Sayfa No:</u>
Çizelge 1.1 Tahmin modelleri için bazı geleneksel ve güncel performans ölçütleri/yöntemleri	3
Çizelge 4.1. Senaryo 1 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)	32
Çizelge 4.2. Senaryo 1 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)	34
Çizelge 4.3. Senaryo 2 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)	35
Çizelge 4.4. Senaryo 2 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)	37
Çizelge 4.5. Senaryo 3 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)	38
Çizelge 4.6. Senaryo 3 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)	40
Çizelge 4.8. Senaryo 4 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)	43
Çizelge 4.9. Pima veri'si için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)	45
Çizelge 4.10. Pima veri'si için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)	46
Çizelge 4.11. Prostat veri'si için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.60$).....	47

SİMGELER VE KISALTMALAR

DCA	Decision Curve Analysis
d	Etki Büyüklüğü
DVM	Destek Vektör Makineleri
EAKA	Eğri Altında Kalan Alan
IDI	Integrated Discrimination Index
KA	Karar Ağacı
KW	Kruskal Wallis
LR	Lojistik Regresyon
n	Gözlem Sayısı
NB	Net Benefit
NRI	Net Reclassification Index
NF	Net Fayda
ort	Ortalama
pt	Eşik Olasılık
R	Bootstrap, Karşılaştırma ve Simülasyon Sayısı
R²	Belirleyicilik Katsayısı
ROC	Receiver Operating Characteristic
ss	Standart Sapma

ÖZET

Tahmin Modellerinin Performans Karşılaştırmasında Karar Eğrisi Analizinin Kullanımı

Tahmin modellerinin performans karşılaştırmasında son yıllarda yeni bir performans metriği olarak Net Fayda değerleri sıklıkla kullanılmaktadır. Literatürde iki tahmin modelinden hesaplanan Net Fayda değerlerinin istatistik olarak karşılaştırılması için önerilmiş bir z testi yer almaktadır. Ancak bu yöntem ile ikiden fazla net fayda değerini ikili olarak karşılaştırmak toplam karşılaştırma sayısının artmasına dolayısıyla Tip 1 hata olasılığının yükselmesine yol açacaktır. Bu gerekçelerle bu çalışmada; ikiden fazla Net Fayda değerinin eşzamanlı karşılaştırılmasını sağlayacak bir yaklaşım önerilmiştir. Öneride ilk adımda yeniden-örnekleme ile sahte (pseudo) Net Fayda değerleri elde edilmiş, daha sonra da bu değerlerin dağılım bilgisine göre ANOVA veya Kruskal Wallis (KW) testlerinin kullanılabilirliği incelenmiştir.

Tezde önerilen yaklaşımın başarısını göstermede, açık kaynaklardan elde edilen gerçek veri setleri ve farklı senaryolara göre üretilmiş veri setleri kullanılmıştır. Gerçek veri olarak Pima ve Prostat verileri kullanılmıştır. Simülasyon çalışması için toplam gözlem sayıları, açıklayıcı değişkenler için ortalama vektör, varyans kovaryans matrisi ve modele ait betaların farklılık gösterdiği 4 senaryo oluşturulmuştur. Tahmin modellerinden ise; Karar Ağacı, Destek Vektör Makinesi ve Lojistik Regresyon kullanılmıştır.

Bu tez çalışmasında, önerilen yaklaşım ile literatürde bulunan yaklaşım, tahmin modelleri arasındaki Net Fayda farkını tespit edilebilme başarısı yönünden değerlendirilmiş ve bu değerlendirmede etki büyüklüğü değerlerinden faydalanılmıştır.

Tahmin modellerinin Net Fayda değerlerini karşılaştırmada varsayımlarının sağlanmadığı durumlarda bile dayanıklı yapısı nedeniyle bu tezde önerilen ANOVA yaklaşımı en iyi sonuçları vermiştir. Bu nedenle ikiden fazla tahmin modelinin performanslarını eş zamanlı karşılaştırmada yeniden-örnekleme ANOVA yaklaşımının kullanılması önerilmektedir.

Anahtar kelimeler: eşik olasılık, karar eğrisi analizi, net fayda, tahmin modeli, tahmin modeli performansı

ABSTRACT

Use of Decision Curve Analysis in Performance Comparison of Prediction Models

In recent years, the performance of prediction models has been increasingly evaluated using Net Benefit values. In the extant literature, a z test has been proposed for the statistical comparison of Net Benefit values calculated from two predictive models. However, it should be noted that employing this method for the comparison of more than two net benefit values obtained from predictive models will result in an increase in the total number of comparisons and, consequently, an increase in the probability of Type 1 error. Consequently, this study proposes an approach that facilitates the concurrent evaluation of multiple Net Benefit values. In the initial step of the proposal, pseudo Net Benefit values were obtained by resampling, and then the usability of ANOVA or Kruskal-Wallis (KW) tests was examined according to the distribution of these values.

To illustrate the efficacy of the proposed approach, this study utilizes real data sets obtained from open sources and data sets generated under various scenarios. The Pima and Prostate data sets were utilized as authentic data. For the simulation study, four scenarios were created in which the total number of observations, the mean vector for explanatory variables, the variance-covariance matrix, and the betas of the model differ. Decision Tree, Support Vector Machine, and Logistic Regression were selected as prediction models.

The proposed approach and the approach found in the literature were evaluated in terms of the success of detecting the Net Benefit difference between the prediction models. Effect size values were used in this evaluation.

The analysis revealed that the ANOVA approach proposed in this study yielded the most optimal outcomes due to its robust design, even in scenarios where the assumptions of the estimation models were not met when comparing the Net Benefit values. Consequently, it is recommended to employ the resampling ANOVA approach for the purpose of comparing the performance of more than two estimation models concurrently.

Key words: threshold probability, decision curve analysis, net benefit, prediction model, prediction model performance

1. GİRİŞ

Doğadaki olaylar, değişken veya özelliklerin birbirleriyle karmaşık ilişkileri sonucu ortaya çıkmaktadır. Bu ilişkileri; doğru, basit ve anlaşılır bir şekilde belirleyebilmek, doğadaki olayların seyrini kısmen de olsa istenilen yönde değiştirebilmek açısından önemlidir. Zira başta hastalıklar olmak üzere; diğer olayların, ortaya çıkma süreçlerinin doğru olarak belirlenebilmesi, bireylerin yaşam kalitesinin artırılmasında önemli bir aşamadır. Bu bağlamda bilimsel çalışmalar, genel olarak, ilgilenilen özellikler (değişkenler) arasındaki ilişkileri inceler. Bu ilişkiler, ilişki türüne göre “doğrusal” ve “doğrusal olmayan” ilişkiler olarak sınıflandırılabilir gibi, değişken sayısına göre de “basit” ve “çoklu ilişkiler” olarak sınıflandırılabilir. Doğrusal ilişkiler, doğrusal olmayan ilişkilere göre daha kolay yorumlanabilir olduğundan, doğrusal olmayan ilişkileri belirlemeye yönelik geliştirilen istatistik yöntemlere göre doğrusal ilişkileri belirlemeye yönelik geliştirilen istatistik yöntemler de daha yaygındır. Bu ilişkileri, belirlemeye yönelik ifade edilen ve bir değişkeni diğer değişken veya değişkenlere dayalı olarak tahmin etmeye yönelik geliştirilen eşitlikler, genel anlamıyla “model” veya “tahmin modeli” olarak ve bu süreç ise “modelleme” olarak ifade edilir. Modelleme aynı zamanda, en az veya optimum sayıdaki açıklayıcı (bağımsız) değişkenle, cevap veya yanıt değişkenini (bağımlı değişkeni) en yüksek doğrulukla tahmin etmek üzere yapılan işlemler süreci olarak da tanımlanabilir. Bu bağlamda, Geisser (1993) modellemeyi "bir sonucun olasılığını en iyi şekilde tahmin etmek için bir modelin oluşturulduğu veya seçildiği süreç" olarak tanımlarken, Kuhn&Johnson (2013) ise "doğru bir tahmin üreten matematiksel bir araç veya model geliştirme süreci" olarak tanımlamaktadır^{1,2}. Modelleme kavramı çoğunlukla tahmin modelleri için kullanılmakla birlikte, sınıflandırma modelleri için de kullanılabilir. Tahmin modelleri; çoğunlukla, kolay tespit edilebilen özelliklerden veya değişkenlerden yararlanarak, zor tespit edilebilen özelliği (değişkeni) tahmin için veya ileriye dönük tahmin yapmak için kullanılabilir.

Steyerberg (2009) tahmin modeli uygulamalarını 3 alana ayırmıştır. Bu alanlar; halk sağlığı, araştırma alanı ve klinik alandır. Bu modellerin halk sağlığı alanında kullanımı, önleyici müdahalelerin planlanmasına yardımcı olabileceği gibi araştırma alanında tahmin modelleri, randomize denemelerin tasarımı ve analizine de yardımcı olabilir. Bunun yanı sıra, klinikte ise tanı ve tedavi sürecinde hekimlere yardımcı olabilir. Modelleme sonucunda, hasta olma olasılığının yüksek bulunması durumu, hekimin tedavi gerekliliği yönünde karar vermesinde etkili olabilir. Zira tedavi gerekli değilken, tedavi uygulanması, yan etki veya diğer ekonomik nedenlere dayalı olarak ortaya çıkabilecek zararın yüksek olmasına ve böylece tedaviden beklenen faydanın azalmasına neden olabilir³.

Klinik alandaki modeller veya klinik tahmin modelleri, birbirine benzer olmakla birlikte; “tanı modelleri”, “prognostik modeller” ve “hastalık oluşum modelleri” olarak üç grupta incelenir⁴. Bu modellerde genellikle, cevap değişkeninin ikili (dichotomous) olması durumunda, bu iki kategoriden herhangi birisinin görülme (ortaya çıkma veya meydana gelme) olasılığı tahmin edilir. Bu tahmin sürecinde, uygun model veya modelleri belirlemek üzere, model performansını değerlendirmede farklı ölçütlere veya kriterlere ihtiyaç duyulmaktadır⁵. Bu bağlamda, Steyerberg ve ark. (2010), klinik tahmin modelleri için yaygın kullanılan bazı geleneksel ve güncel model performans ölçütlerini/yöntemlerini Çizelge 1.1’ deki gibi özetlemiştir⁶.

Çizelge 1.1 Tahmin modelleri için bazı geleneksel ve güncel performans ölçütleri/yöntemleri

Performans Türü	Ölçüt/Yöntem	Görselleştirme	Özellikler
Genel Performans	R^2 , Brier	Validasyon grafiği	Y ile $\hat{Y}$ arasındaki mesafe daha düşük olmalıdır. Kalibrasyon ve ayırım gücü özelliklerini yakalar.
Diskriminasyon Gücü	c istatistiği	ROC eğrisi	İlgilenilen duruma sahip olan ve olmayan sonuca ait sıralı istatistiktir.
	Diskriminasyon eğimi	Box plot	Tahminler arasındaki ortalama farkı gösterir, görselleştirmesi kolaydır.
	Genel kalibrasyon	Kalibrasyon veya validasyon grafiği	Ortalamaları karşılaştırır (y ve $\hat{y}$); dış doğrulama için önemlidir.
Kalibrasyon	Kalibrasyon eğimi		Doğrusal tahmincinin regresyon eğimi; iç ve dış doğrulama için gereklidir, regresyon katsayılarının küçültülmesiyle ilgilidir.
	Hosmer-Lemeshow testi		Tahmin edilen olasılığın ondalık dilimine göre gözlemlenen ile tahmin edileni karşılaştırır.
Yeniden Sınıflandırma	Yeniden sınıflandırma tablosu	Çapraz tablo veya dağılım grafiği	İki modelin sınıflandırmalarını karşılaştırır (belirteçli ve belirteçsiz model).
	Yeniden sınıflandırma istatistiği		Çapraz sınıflandırılmış kategorilere göre gözlemlenen sonuçları tahmin edilen riskle karşılaştırır.
	Net yeniden sınıflandırma indeksi (NRI)		Doğru yöndeki değişikliklerin net hesaplaması için; sonuçlara göre iki modeldeki sınıflandırmaları karşılaştırır.
Klinik Kullanışlılık	Entegre diskriminasyon indeksi (IDI)	2 model için Box plot (belirteçli ve belirteçsiz model)	Tüm kesme noktaları üzerinde NRI'nın entegrasyonudur; diskriminasyon eğimlerindeki farka eşdeğerdir.
	Net fayda (NB)	Çapraz tablo	Modelin tek bir eşik değerinde veya bir dizi eşik değerinde elde edilen net doğru pozitifleri gösterir.
	Karar eğrisi analizi (DCA)	Karar eğrisi	Doğru pozitif ve yanlış pozitiflerin ağırlıklandırıldığı fayda/zarar analizinin grafiğini gösterir.

Sınıflandırma amaçlı kullanılan modellerde, bu modellerin farklı sonuçları ayırt etme yeteneği diskriminasyon olarak ifade edilebilir. Benzer şekilde, ölçülen değerlerle model yardımı ile tahmin edilen değerlerin uyumlu olup olmadığını belirleme yeteneği ise kalibrasyon olarak ifade edilir. Diskriminasyon ve kalibrasyonun birlikteliği genel performans olarak ifade edilirken⁷, yeni bir belirtecin veya biyomarkerın yararlılığı ise yeniden sınıflandırma olarak ifade edilmektedir⁸.

Literatürde yer alan yakın zamanlı çalışmalara göre, klinik karar vermede etkiyi veya yararlılığı değerlendirmede, basit sınıflandırma ölçütlerinden (Doğruluk, Eğri Altında Kalan Alan vb.) daha çok, karar analitik ölçütlerinin (örneğin Net Fayda) kullanılması önerilmektedir^{9,10}.

Geleneksel olarak kullanılan model performans değerlendirme ölçütleri, bir modelin kullanışlı olup olmadığı konusunda fikir vermektedir. Ancak bu ölçütler, klinik fayda hakkında bilgi verme konusunda sınırlıdır. Örneğin bir tahmin modelinin performansına ilişkin çıkarım yapmak üzere, sık kullanılan geleneksel ölçütlerden birisi ROC analizidir. Ancak ROC analizinde, Eğri Altında Kalan Alan (EAKA)¹¹ değerinin ne olması gerektiği veya EAKA için ne kadarlık artışın pratikte değerli olacağı belirgin değildir¹²⁻¹⁴.

Ayrıca geleneksel ölçütler yetersiz/kullanışsız kalabildiği senaryolar olabilmektedir. Örneğin, bazı durumlarda, modele eklenen yeni bir değişken, yanıt değişkeni ile pratikte ilişkili olsa da, eğri altında kalan alandaki artış, istatistik olarak önemli bulunmayabilir¹³⁻¹⁵. Başka bir örnek olarak, geleneksel ölçütlerden birisi olan genel doğruluk oranında, yanlış pozitif ve yanlış negatif değerler eşit ağırlıkla, diğer bir ifade ile aynı fayda-zarar oranı ile değerlendirilmektedir¹⁶. Bununla birlikte, modellerden elde edilen doğruluk oranının 0.50-0.80 arasında olması durumunda, hangi modelin tercih edilmesi gerektiği de klinik sonuçları kapsamadığı için belirgin değildir. Ayrıca duyarlılık ve özgüllükteki artış ve azalışların, birbirine göre değişim oranının ne olması gerektiği de önemli bir sorundur. Örneğin, standart testle karşılaştırıldığında, yeni bir tanı testinin; özgüllüğü %10 artırdığı, duyarlılığı ise %5 azalttığı varsayılın. Bu durumda hangi testin kullanılması gerektiğine ilişkin karar vermede belirsizlik olabilmektedir¹⁷.

Tüm bu belirsizliklerin yanısıra, bireyler; DNA yapıları ve fizyolojisi, çevre ve davranışsal düzeyde eşsiz özelliklere sahip olduğundan, hastalıkların tedavisi için bireye özgü müdahalelerin yapılabilmesi önem arz etmektedir. Bu bağlamda bireysel veya kişiselleştirilmiş tahmin modellerine ihtiyaç duyulmakta ve bu modellerin performansını ölçmek için de bir klinik fayda hesabı gerekmektedir¹⁸.

Tanımı konusunda tam bir birliktelik olmamakla birlikte, "Klinik fayda", hasta bakımındaki bir müdahalenin uygunluğunu ve yararını tanımlamak için kullanılan bir terimdir¹⁹.

Vickers ve ark tarafından (2006) yılında önerilen yaklaşım; klinik fayda için, belirlenen eşik olasılıklarında, tanı testinin kullanımıyla ilgili olarak kararları etkileyebilen ve hastaların tedavisiz kalması durumunda oluşabilecek riskleri veya tedavideki riskleri kabul etmeye yönelik tercihleri de hesaba katmaktadır. Eşik olasılıklarına karşı bu yol ile hesaplanan Net Fayda değerlerine göre çizilen grafik ile yapılan değerlendirme "Karar Eğrisi Analizi" olarak adlandırılmaktadır²⁰.

Bir tahmin modeli; genellenebilirliği, basit veya karmaşık olması, aşırı uyumu, performansı ve algoritması gibi ölçütler bakımından diğer modellerden üstün olabilmektedir²¹. Bu üstünlüğü göstermek üzere de farklı istatistik yöntemlerden yararlanılmaktadır. İki modelin Net Fayda bakımından istatistik olarak karşılaştırılmasında bazı yöntemler önerilmiştir^{22,23}. Net Faydaları ikili olarak karşılaştırmada, bir yaklaşım Sande S.Z. tarafından, 2020 yılında önerilmiştir²³. Bu yaklaşım iki Net Fayda arasındaki farkın standartlaştırılması ile elde edilen z skoruna dayanmaktadır. Ancak yaklaşımın yapısı gereği ikiden fazla Net Fayda değerini aynı anda karşılaştırmada kullanılamamaktadır. Öte yandan bu yaklaşım ile ikiden fazla Net Fayda değerinin ikili karşılaştırması ile değerlendirmesi Tip 1 Hata olasılığını yükseltecektir. Literatürde ikiden fazla Net Faydanın eş zamanlı istatistik olarak karşılaştırılmasında ise karar eğrisi analizi için bir özet ölçüsü istatistiğine rastlanılmamıştır. Dolayısı ile 2'den fazla Net Fayda karşılaştırıldığında; z testinin, hem Tip 1 hata olasılığını artıracak, hem de her seferinde 2'li karşılaştırma ile işlemlerin uzayacağı gerekçesiyle bu çalışmada; z testine alternatif bir teknik olmak üzere 2'den fazla Net Fayda eşzamanlı karşılaştırıldığında Net Faydaların istatistik olarak birbirinden farklı olup olmadığını tespit edebilecek yöntem olarak bootstrap ile elde edilen Net Faydalara uygulanan ANOVA ve Kruskal Wallis testlerinin kullanılabilirliği gerçek ve simüle veriler ile incelenmiştir.

2. GENEL BİLGİLER

Klinik hedeflerden biri, veri temelli bir değerlendirme ile hasta bakımına ve tedaviye imkan sağlamaktır. Bu hedefe ulaşmak için gereken süreci iyileştirme amacı ile tasarlanmış bir klinik model; hasta ile ilgili faktörleri (veya risk değişkenlerini), hekimin yürüttüğü klinik süreçleri, hastanenin bu süreçleri uygulamasını ve uygulama sonuçlarının ölçümlerini kapsamaktadır²⁴. Hekimin yürüttüğü klinik süreçlerde kullanılan klinik tahmin modelleri; hasta, hekim veya diğer sağlık hizmeti sağlayıcılarını hastanın belirli bir hastalığa sahip olma veya hastalık derecesi ile bu hastalık derecesinde meydana gelebilecek değişim olasılığı hakkında bilgi verebilmekte ve böylece ilgili karar alma süreçlerinde karar vericilere yardımcı olabilmektedir. Hastanın demografik özellikleri ve test sonuçları göz önünde bulundurularak, gerçek dünya problemine bir model uyarlamak; teşhis edilmemiş yüksek riskli hastaların tespitinde, erken müdahalelerle hastalığın önlenmesinde, gelecekte hastalığa sahip olma olasılığının tahmininde veya hastalık için bir prognoz (örneğin, komplikasyon veya ölüm oranı) seyrine ilişkin tahminde bulunmada faydalı olabilmektedir^{25, 26}.

2.1. Tahmin Modelleri

Geçmiş ve şimdiki bilgileri değerlendirmek ve gelecek tahminleri yapmak için yapay zeka, makine öğrenmesi ve veri madenciliği tekniklerini kullanan tahmin modelleri, veri analitiğinin önemli bir bölümünü oluşturur. Tahmin modelleri, gelecekteki olayları veya sonuçları tahmin etmek için matematiksel ve hesaplamalı yöntemlerin kullanımını içermektedir. Bu yöntemler; Regresyon Analizi, Karar Ağaçları, Rastgele Ormanlar, Sinir Ağları ve Destek Vektör Makineleri gibi çeşitli teknikleri kapsamaktadır. Bu yöntemlerde, veri analizi ve model oluşturabilmek için verilerde bulunan örüntülere ve ilişkilere dayalı olarak sonuçları tahmin edebilen algoritmalar kullanılmaktadır²⁷. Detaylı veri analizi yapmaya olanak sağlayan bu tahmin modelleri, daha az belirgin (marjinal) risk faktörlerini tanımlayabilmekte ve bu risklerin klinik önemini gösterebilmektedir. Dolayısıyla, tahmin modelleri hastaların sahip oldukları durumların daha doğru bir şekilde sınıflandırılmasına imkan sağlamaktadır²⁸.

Sonuç değişkeninin tahmin edilmeye çalışıldığı araştırmalarda amaç tahmin modeli oluşturmaktır²⁹. Gelecekteki bir sonucu tahmin etmek için şu anda mevcut olan bilgileri kullanmak gerektiğinden, bir tahmin modeli bir sonucun değerini (genellikle bağımlı değişken, yanıt veya çıktı olarak adlandırılır) tahmin etmek için bilinen değişkenleri (genellikle bağımsız değişkenler, özellikler, açıklayıcı değişkenler veya girdiler olarak adlandırılır) kullanan bir yapıya sahiptir. Dolayısı ile bu modelin unsurları; ilgi konusu olayın tahmin edicilerinin bilinmesi ve oluşturulan tahmin modelinin, yeni gözlemlere uygulandığında doğruluğunu koruduğu genelleştirilebilir bir yapıda olmasıdır³⁰. Ayrıca modeller belirli olayları doğru bir şekilde tahmin edebilen, basit, kullanışlı, düşük maliyetle ve ilgilenilen toplumun yapısına uygun olmalıdır²⁵. Dolayısı ile bir tahmin modeli geliştirildikten sonra, ne kadar yararlı olabileceğini ve hangi koşullarda/durumlarda kullanılabileceğini belirlemek için değerlendirilmesi gerekmektedir. Bir tahmin modelinin değerlendirilebilmesi için yaygın olarak kullanılan birkaç yaklaşım vardır³¹. Bu yaklaşımların kolayca uygulanabilmesi için, olasılık fonksiyonunun belirli bir biçimine dayanmadan tüm modellere uygulanabilir ve farklı modelleri karşılaştırmada kullanılabilir bir yapıda olması gerekmektedir³². Bu şekilde tahmin modellerinin performansın değerlendirilebilmesi için kullanılan geleneksel yaklaşımlardan bazıları; genel model performansının değerlendirilmesi, kalibrasyon ve diskriminasyondur (Eğri Altındaki Alan (EAKA) veya Uyum İndeksi)^{3, 29, 33}.

2.2. Modellerin Performanslarının Değerlendirilmesi

2.2.1. Belirleyicilik Katsayısı (R^2)

Genel model performansının değerlendirilmesinde yanıt değişkeni sürekli olan doğrusal bir modelde, model tarafından açıklanan varyans yüzdesi olarak yorumlanabilen belirleyicilik katsayısı R^2 sıklıkla kullanılmaktadır. Ayrıca parametre sayısı gözlem sayısına kıyasla büyükse, sistematik ve yapay olarak şişeceği gerekçesi ile de "Düzeltilmiş R^2 " adı verilen başka bir katsayı kullanılmaktadır³⁴. R^2 'nin potansiyel tanımlarından biri, tahmin edilen ve gözlenen değerler arasındaki her zaman 0 ile 1 arasında değer alan karesel korelasyondur³⁵.

2.2.2. Brier Skoru

Model performansını genel olarak değerlendirmek amacıyla sıklıkla kullanılan bir diğer metrik ise Brier Skordur. Brier skoru, Brier tarafından 1950 yılında hava tahminlerinin doğrulanması sorununu ele almak için tanıtılan bu skor, ikili sonuçların tahminlerini değerlendirmek için basit bir puanlama kullanan ve her gözlem için sonuca atanan gerçek sonuç ile öngörülen olasılık arasındaki Öklid mesafesine dayalı bir ölçüdür³⁶. Brier skoru 0 ve 1 arasında değer alabilen bir ceza skoru olarak düşünülebilir, 0 tahminlerin gerçek sonuçlarla tam olarak uyuştuğunu, 1 ise tam olarak uyuşmadığını gösterir³⁷.

2.2.3. Kalibrasyon Grafiği

Bir diğer model performans ölçütü olan bir tahminin kalibrasyonu; bu tahminin ($\hat{p}$) belirli bir hasta için olayın gerçek olasılığına (p) ne kadar yakın olduğunu kontrol ederek ölçülmektedir. Her hastanın farklı özelliklere sahip olduğu göz önüne alındığında, hastalığın gerçek olasılığının ne olduğunu belirlemek mümkün değildir ve bu nedenle benzer bireylerden oluşan bir grupta olayın olasılığı gibi belirli vektörler kullanılarak yapılan tahminler, bu grupta mevcut hastalığın gerçek oranına yakınsa, tahmin iyi kalibre edilmiş olarak kabul edilmektedir³⁸. Bu anlamda ilgilenilen parametre gerçekte örneğin β olduğunda, tahmin edicisinin ($\hat{\beta}$) beklenen değeri ile arasındaki fark, gerçek değerden olan sapmadır ve bias (yanlılık) olarak ifade edilen bu fark kalibrasyon ile yakından ilişkilidir. Fakat kalibrasyon bir parametre tahmini yerine bir olayın olup olmayacağına dair çoğu zaman ikili sonuçlar için (örneğin, 1 veya 0) olasılık tahmini ile ilgilenmektedir. Dolayısı ile yanlılık gerçek parametre değeri ile tahmin edilen değer arasındaki fark iken, kalibrasyon gerçek olasılık değeri ile tahmin edilen olasılık arasındaki farkı temsil etmektedir. Bir olayın olması iki sonuçla; gerçekleşme (1) ya da gerçekleşmeme (0) olarak ifade edildiğinde, tek seferlik meydana gelen bir olayın "gerçek" olasılığını doğrudan ölçmek mümkün olmayacağından, kalibrasyon değerlendirmesinde birden fazla tahmin incelenmektedir. Örneğin bir model birçok olay için %75 olasılık verdiğinde, bu olayların yaklaşık %75'inde gerçekleşmesi beklenir.

Başka bir ifade ile uzun vadede, iyi kalibre edilmiş bir modelin tahmin ettiği olasılıklar ile gerçek sonuçların uyumlu olması beklenmektedir³⁹. Bu, bir kalibrasyon grafiği kullanılarak değerlendirilmektedir. Bu grafikte, öngörülen riskler ile gözlenen

sonuçlar arasındaki ilişki görselleştirilir. Kalibrasyon grafiğinin x ekseninde tahminler y ekseninde ise gözlenen sonuçlar vardır.

Mükemmel tahminler 45° eğime sahip olan bir dağılım grafiğidir. İkili sonuçlar için grafik, y ekseninde yalnızca 0 ve 1 değerlerini (gözlenen sonuçları) içermektedir³. Bu durumda tahmin edilen olasılıklar ile gözlenen sonuçlar arasındaki ilişki gösterimini kolaylaştırmak için bir smoothing tekniği olan Loess algoritması kullanılarak gözlenen sonuçların olasılığa dönüştürülmektedir^{3,40}. Kalibrasyonun değerlendirilmesinde yaygınca kullanılan test yaklaşımı ise Hosmer–Lemeshow (HL) uyum iyiliği istatistiğidir⁴¹. Hosmer–Lemeshow (HL) uyum iyiliği testi, sonucun tahmin edilen olasılığına dayanarak örneği katmanlara ayırma temeline dayanmaktadır. Bu teknikte katmanlar arasında gözlemlenen ve tahmin edilen sonuç olasılıklarını karşılaştırmak için Ki-Kare testi kullanılmaktadır⁴².

2.2.4. Diskriminasyon

Yine geleneksel performans ölçütlerinden diskriminasyon ise bir modelin, ilgilenilen sonucu deneyimleyen ve deneyimlemeyen hastalar arasında ayırım yapabilme kabiliyetidir. Tahmin modeli veya tanı testleri hastalığın olup olmadığını tahmin etmek için kullanıldığında, hastalık durumu genellikle biyopsi gibi daha invaziv veya pahalı bir altın standart kullanılarak bilindiğinde modelin/testin ne kadar iyi sınıflandırdığını değerlendirmek için sıklıkla basit bir 4 gözlü tablo kullanılmaktadır. Bu tablo aracılığı ile duyarlılık (hastalığı olanlar arasında test/model tahminlerinin de pozitif olması) ve özgüllük (hastalığı olmayanlar arasında test/model tahminlerinin de negatif olması) kolayca hesaplanabilmekte veya değerlendirilebilmektedir. Modeller/testler karşılaştırırken, hem duyarlılık hem de özgüllük açısından daha yüksek olanları tercih edilmektedir. Bir modelin/testin duyarlılığı diğerinden daha yüksek ancak özgüllüğü daha düşük olabileceğinden, bazen bu ölçümleri birleştirmek için tanısallık oranı kullanılmaktadır⁴³.

2.2.4.1. ROC Eğrileri

Diskriminasyon en yaygın olarak ise alıcı işletim karakteristiği (ROC) eğrileri kullanılarak değerlendirilmektedir^{43,44}. Çalışılan modelin/testin duyarlılığı ve özgüllüğü; pozitif ve negatif test sonuçlarını tanımlamada kullanılan bir kesme değerine bağlıdır. Duyarlılıklar ve özgüllükler, daha düşük özgüllüklerde daha yüksek duyarlılıkların ve tam tersinin elde edildiği tüm olası kesme değerleri aralığı üzerinde hesaplanabilmektedir.

Ortaya çıkan duyarlılık-özgüllük çiftleri serisi, her olası kesme değeri için duyarlılığı özgüllüğe göre gösteren bir ROC eğrisi olarak çizilmektedir. Bu eğrinin altında kalan alan (EAKA), diğer bir deyişle c istatistiği, hasta ve hasta olmayan hastalar arasında ne kadar iyi ayrım yaptığını 1'e olan yakınlığı ile göstermektedir. 1'e yakun bir EAKA mükemmel ayrım anlamına gelirken, 0.5'lik bir EAKA, modelin ayrım yapma kabiliyetinin olmadığı anlamına gelmektedir^{11,45}.

2.2.4.2. Diskriminasyon Eğimi

Bir diğer diskriminasyon metriği olan diskriminasyon eğimi ise ilgilenilen sonuca sahip olanlara ait tahminlerin ortalaması ve ilgilenilen sonuca sahip olmayanlara ait tahminlerin ortalaması arasındaki mutlak fark olarak hesaplanmaktadır. Hesaplanan fark küçükse, modelin sonuca sahip olanlar ile olmayanların ayrımını yapmada başarısız olduğu söylenebilmektedir. Görselleştirme, bir ayrım modeli için ilgilenilen sonuca sahip olanlar ve olmayanlar arasındaki örtüşmeyi gösterecek bir kutu grafiği veya histogram ile kolayca mümkündür. Bu örtüşme ne kadar az ise ayrımın o kadar iyi olduğu yorumu yapılabilmektedir⁶.

2.3. Karar Verme Süreci

İyi bir klinik karar alabilmek için iyi kalibrasyon ve diskriminasyon şarttır. Ancak hasta özellikleri, hastalık insidansı veya yaygınlığı ve popülasyonlardaki heterojenlik gibi nedenler zayıf kalibrasyona yol açmaktadır⁴⁶. Kalibrasyon, bir popülasyondaki ortalama riski ilgilendirir ve iyi kalibre edilmiş bir model hastalığı önleme kararlarına yardımcı olabilir, ancak yanlış kalibre edilmiş bir model, yüksek risk altındaki bir bireye düşük tahmin edilen bir olasılık atandığı için etkili önleyici müdahaleden vazgeçildiği durumlara yol açabilmektedir⁴⁷.

Yaygın olarak kullanılan Hosmer-Lemeshow testinin de sınırlı güç ve zayıf yorumlanabilirlik gibi bir dizi dezavantajı vardır^{6,48}. Kalibrasyon klinik fayda ölçümü için de yeterli değildir. Örneğin, yaş faktörüne bağlı bir hastalık olan endometrium kanseri için genelde 50 yaşını geçen kadınlar taramaya dahil edilmektedir. Bunun için tahmin modelleri mevcut olsa da kullanmaya gerek duyulmaz ve bir kadında endometrium kanser riski %15 değerinde %20 olsa bile bu görece küçük fark kadınlarda tarama sıklığını farklılaştırmayı (riskin %20 olduğu durumda daha sık taramalar dolayısı ile daha fazla test yapmayı) haklı çıkarmayacaktır. Dolayısı ile iyi kalibre edilmiş bir model klinikte tarama sıklığının kararında bir fark oluşturmayaacağından, o modelin klinik fayda sağlamadığı söylenebilir⁶.

Diskriminasyon ölçütlerinden biri olan EAKA değeri ise, tahmin modelinin doğruluğunu genel olarak hesaplar ve değerlendirici EAKA değerinin 1'e yakın olması durumunda iyi bir ayırım olduğunu düşünebilir ancak EAKA aşırı (overprediction) ve eksik tahminleri (underprediction) doğrudan ele almamaktadır. Dolayısı ile EAKA ile model karşılaştırırken; veri kalitesi, hata türleri ve bağlam gibi faktörler de göz önüne alınmalıdır⁴⁹. Ayrıca kararların kalitesinin değerlendirilmesinde önemli bir konu, bazı durumlarda yanlış pozitif ve yanlış negatif kararların genellikle hastalığın tipine göre oldukça farklı ağırlıklara sahip olması gerekliliğine rağmen diskriminasyon metrikleri ile eşit ağırlıklandırılarak yapılan hesaplamalardır⁸. Model karşılaştırmalarında bir diğer durum, daha yüksek bir EAKA değerine sahip bir modelin genellikle daha değerli olacağı düşünülmesine rağmen bu modellerin bazen daha kötü sonuçlara yol açabileceğidir. EAKA, yalnızca diskriminasyonu yansıtmakla beraber, kalibrasyon metrikleri gibi klinik sonuçlar ve kararların bağlamını hesaba katmamaktadır²⁰.

Ayrıca, büyük etki büyüklüklerine sahip belirteçlerin EAKA'nı artırması gerektiği gerçeği, küçük artışların klinik faydalarda anlamlı bir kazanç sağlamak için yeterli olacağı anlamına gelmemektedir⁵⁰.

Diskriminasyonu iyileştirmek için parametrelerin ayarlanmasına veya tamamen farklı bir modele ihtiyaç duyulabilmektedir⁵¹. Mükemmel kalibrasyon ve mükemmel diskriminasyonun birleşimi yalnızca mükemmel tahminle elde edilebilir fakat bu çoğu uygulamada çok gerçekçi değildir⁴⁵. Ayrıca mükemmel kalibre edilmiş bir modelin mükemmel diskriminasyona ulaşamayacağı basit bir simülasyon çalışmasıyla kanıtlanmıştır⁵². Öte yandan, zayıf kalibrasyon bir modeli, daha düşük EAKA'ya sahip ancak iyi kalibre edilmiş bir rakip modelden klinik olarak daha az yararlı hale getirebilmektedir⁴⁶. Geleneksel yöntemlerdeki tüm bu eksiklikler/yetersizlikler yeni

model performans metriklerinin oluşumuna yol açmıştır. Örneğin; EAKA gibi diskriminasyon ölçütleri, özellikle temel modelin EAKA'sı zaten büyükse, model performansındaki küçük iyileştirmeleri tespit etmeye duyarsız olduğundan yeniden sınıflama yöntemleri geliştirilmiştir^{8,53}. Yine, kalibrasyon ve diskriminasyonun özellikle de klinik sonuçları kapsamadığı gerekçesi ile önerilen yeni model performans ölçütlerinden biri de Karar Eğrisi Analizidir.

2.4. Karar Eğrisi Analizi

Son zamanlarda model performans ölçütlerinden biri olarak kabul gören 'Karar Eğrisi Analizi (DCA)', bir tanı testinin, belirtecin veya tahmin modelinin klinik olarak kullanılabilirliğini değerlendirmek ve klinik karar bakımından bir tahmin modelinin kullanılmasının fayda-zarar durumunu tespit etmek için geliştirilmiştir. Net Fayda istatistikinde, fayda, hastalığı olan ve önerilen müdahale uygulanarak tedavi görecektir hasta sayısı (gerçek pozitifler) ile ifade edilirken; zarar, hastalığı olmayan ve hatalı olarak müdahale edilecek hasta sayısının (yanlış pozitifler) eşik olasılıkla ağırlıklandırılması olarak ifade edilmektedir.

2.4.1. Karar Eğrisi Analizi ile İlgili Yapılan Çalışmalar

27.02.2025 tarihinde Karar eğrisi analizi ("Decision Curve Analysis") anahtar sözcüğü ile tüm alanlar seçilerek Web of Science veri tabanında yapılan araştırmada 38.012 adet sonuca ulaşılmıştır. Yıllara göre bu konuyla alakalı en eski çalışma, 1980'lerde en yeni çalışma 2025 yılında olup, 2018 yılından sonra çalışma sayıları hızla artmıştır. Farklı disiplinlerde/ alanlarda örneğin; Onkoloji (4.417 adet çalışma), Radyoloji (2.333 adet çalışma), Dahili Tıp Bilimleri (2511 adet çalışma), Cerrahi Tıp Bilimleri'nden (1901 adet çalışma) yapılan çalışmalar 35.846 adet makale, 879 adet review, 39 adet kongre bildiri özeti, 43 adet editöryel yazı, 35 adet editöre mektup gibi farklı yayın türlerinden oluşmaktadır. Çalışmaların yapıldığı ülkeler arasında ise sırası ile en çok Çin (15.602 adet çalışma), Amerika (8.107 adet çalışma) ve İngiltere'nin (2.401 adet çalışma) yer aldığı tespit edilmiştir. Konu ile ilgili farklı alanlarda yapılan çalışmalar olduğu da görülmüştür. Örneğin sağlık alanında yapılan çalışma sayısı 29.291 iken, iklim alanında 2.153, cinsiyet eşitsizliği alanında 1.924, eğitim kalitesi alanında ise 165 adettir. En çok atıf alan çalışmaların dağılımına bakıldığında ise ilk

sırada 1.245 adet çalışma ile karaciğer ve kolon kanseri, ikinci sırada 1.002 adet çalışma ile prostat kanseri, üçüncü sırada ise 859 çalışma ile göğüs kanseri başlıkları yer almaktadır.

2.4.2. Eşik Olasılık

Eşik olasılık değeri fayda ve zararlarının birbirine göre ağırlıklandırılabilmesi için kullanılmaktadır. Vickers eşik olasılığı; "tedavinin beklenen faydasının tedaviden kaçınmanın beklenen faydasına eşit olduğu yer" olarak tanımlamıştır²⁰.

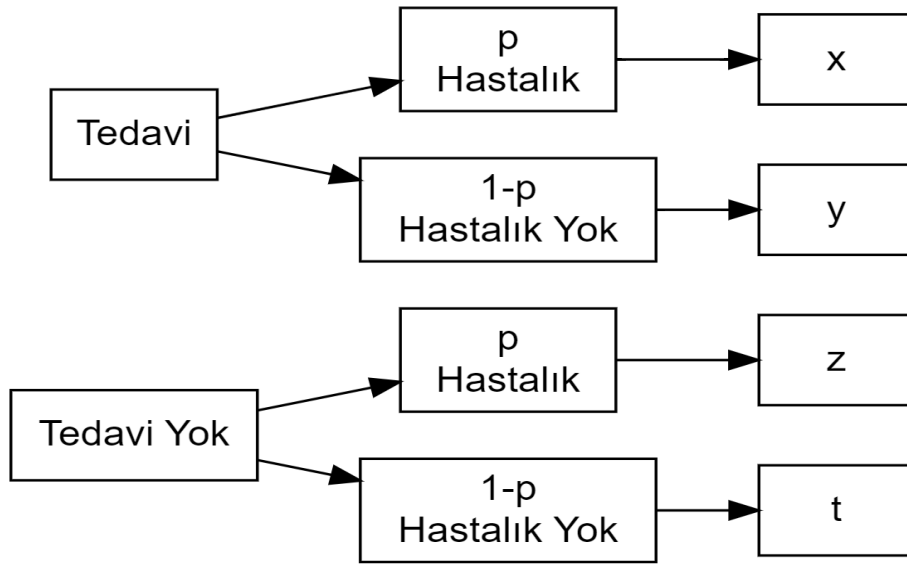
Eşik olasılık, hastaya yapılacak müdahale/test/tedavi için gerekli olan minimum sınırdır. Başka bir deyişle eşik olasılık, hasta olma olasılığı; %1'den %99'a çıkarıldığında, bu aralıkta hastanın ameliyat olup olmayacağından emin olamayacağı noktadır ve bu olasılık, kararın sonuçlarının nasıl ağırlıklandırıldığı ile ilişkilidir. Bir hastanın hasta olma riski %10 veya daha yüksekse ameliyat olmayı tercih edeceği, risk %10'un altında ise ameliyat olmayacağı bir senaryoda %10'luk hasta olma olasılığı (riski), %90'lık hastalık olmaması olasılığına karşılık gelir; bu oran da 9:1'dir. Bu nedenle eşik olasılığı %10 olan bir hasta; örneğin yanıklara erken cerrahi müdahalenin faydalarının, gereksiz cerrahi müdahalelerin zararlarından 9 kat daha değerli olduğunu düşünebilmektedir. Benzer şekilde "yüksek dereceli bir hastalığı kaçırmak, gereksiz bir biyopsi yapmaktan 9 kat daha kötüdür" anlamına da gelmektedir^{14, 54-56} Ayrıca yine eşik olasılığın %10 olması durumunda tedavi alan her 10 hastadan sadece 1'inde fayda sağlaması ve diğer 9 hastanın gereksiz tedavi almasına rağmen hastalığın ciddiyeti yüksek, yan etkilerin zararı ise nispeten düşük olduğu anlamına gelmektedir⁵⁷ Dolayısı ile daha düşük bir eşik olasılık duyarlılığı desteklerken (gerçek pozitifleri kaçırmamak konusunda endişe duymak), daha yüksek bir eşik olasılık seçmek ise yanlış pozitifleri en aza indirmeye çalışarak özgüllüğü destekleyecektir.

Muhtemel eşik olasılık aralığı hastanın tedavi/ameliyat olup olmama konusundaki tutum ve tercihi ve/veya hekim tercihlerinin yelpazesini temsil etmektedir ve bu eşik olasılık aralığında elde edilen karar eğrisi; hasta veya hekim tercihlerini analize dahil etmiş olmaktadır¹⁰. Buradan da anlaşılabileceği üzere belirli bir eşik aralığında bir modelin diğerine üstün olup olmadığını incelemek için eşik olasılığının değişmesine olanak sağlayan karar eğrisi için karar vericiler, öncelikle, hastanın da tercihlerini ve hastalığı gerçekte var olan bir hastaya müdahaleden kaçınmanın ve hastalığı gerçekte

var olmayan bir hastaya gereksiz yere müdahale etmenin nispi zararlarını göz önünde bulundurarak, klinik olarak makul bir eşik olasılık aralığı belirlemelidir ^{22,56}.

2.4.3. Karar Eğrisi ve Net Fayda

Belirli bir hastalık için hastalığın mevcut olup olmadığından emin olmayan bir hastanın tedavi görüp görmemeye karar verdiği durumu ele alınsın. Bu örneği görselleştirmek için Şekil 1'de basit bir karar ağacı verilmiştir. Bu grafikte p hastalık olasılığıdır ve x , y , z ve t her bir sonuçla ilişkili değeri, vermektedir.



Şekil 2.1. Tedavi için bir karar ağacı.

Hastalık olasılığı p ; gerçek pozitif, yanlış pozitif, yanlış negatif ve gerçek negatif değerleri sırası ile x , y , z , t ²⁰

Bireyin hastalığa sahip olma olasılığını veren bir tahmin modeli olduğunu varsayıldığında; hastalık olasılığı 1'e yakınsa, tedavi seçeneği düşünülecek hastalık olasılık 0'a yakınsa, tedavi gerekmeyecektir. 0 ile 1 arasındaki bir olasılıkta ise tedavinin gerekliliği konusunda kesin bir karar verilemeyen ve eşik olasılık olarak adlandırılan nokta, p_t , tedavinin beklenen faydasının, tedaviden kaçınmanın beklenen faydasına eşit olduğu yer olduğundan;

$$p_t x + (1 - p_t) y = p_t z + (1 - p_t) t$$

olarak yazılabilir ve buradan

$$p_t x - p_t z = (1 - p_t)t - (1 - p_t)y$$

elde edilir. Basit bir düzenleme ile

$$p_t(x - z) = (1 - p_t)(t - y)$$

ve buradan

$$\frac{(x - z)}{(t - y)} = \frac{(1 - p_t)}{p_t}$$

elde edilir.

Burada, $x - z$, gerçek pozitif bir sonuç değeri olarak 1'e sabitlendiğinde, $y - t$, yanlış pozitif bir sonuç değeri olarak $-p_t/(1 - p_t)$ elde edilir. Karar eğrisi, belirlenen eşik olasılıklarda (p_t) hesaplanan Net Faydaların bir grafiğidir ve Denklem 1 aracılığı ile hesaplanmaktadır. Böylece,

$$Net\ Fayda = \frac{Gerçek\ Pozitif\ Sayısı}{n} - \frac{Yanlış\ Pozitif\ Sayısı}{n} \left(\frac{p_t}{1 - p_t} \right) \quad (1)$$

olarak yazılabilir. Burada yapılan işlemde, gözlenen gerçek pozitiflerin oranından, p_t (eşik olasılık) ile ağırlıklandırılarak düzeltilen gözlenen yanlış pozitif oranının farkı alınmaktadır. Denklemde yer alan pozitif ve negatif sayıları, modelden hesaplanan tahmin olasılıklarının belirlenen eşik değerinden büyük veya küçük olmasına göre hesaplanmaktadır.

Klinikte; eşik olasılık (p_t) %10, Net Fayda Eşitliği (Denklem 1) ile elde edilen Net Fayda değeri de 0,498 varsayıldığında, bu değer kullanılan tahmin modelinin, herhangi bir hastalık için 49.8 (≈ 50) birey daha fazla tespit edebilme gücü olarak yorumlanabilmektedir. Böylece 100 hastadan yaklaşık 50'sine gereksiz işlem (uygulama) kaçınılmış olunacaktır²⁰.

Karar eğrisi analiz yönteminin kullanımı; geliştirilen nomogramların klinik anlamlılığını değerlendirmede⁵⁸⁻⁶⁵, tanı testlerinin performanslarının değerlendirilmesinde^{66, 67}, tahmin modellerinin performanslarının değerlendirilmesinde⁶⁸⁻⁷¹, belirtecin (biyomarkerın) modele olan katkısının değerlendirilmesinde^{72, 73} oldukça yaygınlaşmıştır.

3. GEREÇ VE YÖNTEM

3.1. Gereç

3.1.1. Gerçek Veri Çalışması

Pima Veri Seti

Pima veri seti; 8 açıklayıcı değişken ile 1 yanıt değişkenden oluşmaktadır. Bunlar; Pregnancies, Glucose, BloodPressure, SkinThickness, Insulin, BMI, Diabetespedigree, Age ile Outcome değişkenleridir. Açıklayıcı değişkenler sürekli, yanıt değişkeni ise iki sınıflı kategoriktir (hastalık var, yok). Bu veri setinin aslı Ulusal Diyabet ve Sindirim ve Böbrek Hastalıkları Enstitüsü'nden elde edilmiş olup amacı, veri setinde yer alan değişkenlerden bir hastanın diyabet hastası olup olmadığını tanısal olarak tahmin etmektir. Veri seti, en az 21 yaşında olan ve Pima Kızılderilisi gen mirasına sahip 768 kadından oluşmaktadır. Yanıt değişkeni için prevalans ise yaklaşık %35'dir. Veri setine Kaggle platformundan ulaşılmıştır⁷⁴.

Prostat Veri Seti

Prostat veri seti; 8 açıklayıcı değişken ile 1 yanıt değişkenden oluşmaktadır. Bunlar Radius, Texture, Perimeter, Smoothness, Compactness, Symmetry, Area, Fractal_dimension, Outcome, değişkenleridir. Açıklayıcı değişkenler sürekli, yanıt değişkeni ise iki sınıflı kategoriktir (hastalık var, yok). Veri setinde de amaç, bu değişkenleri kullanarak bir hastanın Prostat Durumunu tahmin etmektir. Veri seti, 100 erkekten oluşmaktadır. Yanıt değişkeni için prevalans ise yaklaşık %62'dir. Veri setine Kaggle platformundan ulaşılmıştır⁷⁵.

3.1.2. Simülasyon Veri Çalışması

Simülasyon senaryoları için, R paket programında rmvnorm paketi ile çok değişkenli normal dağılımdan veri setleri üretilmiştir. Üretilen verilerde önce 3 açıklayıcı değişken üretilmiş, belirlenen regresyon katsayıları (beta) kullanılarak hata terimlerini de içeren bir lojistik model oluşturulmuş ve model sonunda elde edilen yanıt değişkeni olasılıklarından hareketle binary bir yanıt değişkeni elde edilmiştir. Daha sonra üretilen bu veri kullanılarak 3 sınıflama algoritması (Karar ağacı, Destek Vektör Makinesi ve Lojistik Regresyon) çalıştırılarak sınıflama başarıları kaydedilmiştir.

Veri üretme aşamaları:

1. Çok değişkenli normal dağılımdan; n adet gözlemden oluşan p adet bağımsız değişken (X) oluşturulmuştur.
2. İlk iki değişken sürekli yapıdan kategorik yapıya dönüştürülmüştür.
 - Her bir değişkenin medyan değeri bulunmuştur.
 - Değişken medyanından büyükse 1, küçük veya eşitse 0 olarak tanımlanmıştır.
3. Doğrusal tahmin fonksiyonu oluşturulmuştur.

$$\eta = X\beta$$

Burada X, bağımsız değişkenler matrisi ($n * p$), β ise katsayılar vektörüdür.

4. Lojistik regresyonda doğrusal tahmin fonksiyonunun lojistik fonksiyon ile olasılığa dönüştürülmüştür:

$$p = \frac{1}{1 + e^{-\eta}}$$

Burada p, her gözlem için sınıfın gerçekleşme olasılığıdır.

5. Her gözlemin sınıflara aitlik durumu (0 veya 1), hesaplanan olasılıklar (p) kullanılarak Bernoulli dağılımından üretilmiştir.

$$y_i \sim \text{Bernoulli}(p_i)$$

Veri üretiminde kullanılan toplam gözlem sayıları, açıklayıcı değişkenler için ortalama vektörü ile varyans kovaryans matrisi ve modele ait betalara ait farklı senaryolar şöyledir:

1. Çok değişkenli normal dağılımdan ortalaması $\mathbf{0}$ ve varyans kovaryans matrisi $diag(1,1,1)$ olan, 3 açıklayıcı (2 kategorik, 1 sürekli), 1 yanıt değişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betanın ise $\beta=(\mathbf{0.1, 0.3, 0.5})$ olduğu,

2. Çok değişkenli normal dağılımdan ortalaması $\mathbf{0}$ ve varyans kovaryans matrisi

$$\begin{pmatrix} 1 & 0.3 & 0.1 \\ 0.3 & 1 & 0.3 \\ 0.1 & 0.3 & 1 \end{pmatrix}$$

olan, 3 açıklayıcı (2 kategorik, 1 sürekli), 1 yanıt değişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betaların ise $\beta=(\mathbf{0.1, 0.3, 0.5})$ olduğu,

3. Çok değişkenli normal dağılımdan ortalaması $\mathbf{0}$ ve varyans kovaryans matrisi; $diag(1,1,1)$ olan, 3 açıklayıcı (2 kategorik, 1 sürekli), 1 yanıt değişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betanın $\beta=(\mathbf{0.5, 0.5, 0.5})$ olduğu,

4. Çok değişkenli normal dağılımdan ortalaması $\mathbf{0}$ ve varyans kovaryans matrisi;

$$\begin{pmatrix} 1 & 0.3 & 0.1 \\ 0.3 & 1 & 0.3 \\ 0.1 & 0.3 & 1 \end{pmatrix}$$

olan, 3 açıklayıcı (2 kategorik, 1 sürekli), 1 yanıt değişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betanın $\beta=(\mathbf{0.5, 0.5, 0.5})$ olduğu durumlardır.

Tüm senaryolarda yanıt değişkenine ait prevalans 0.5 ile 0.6 arasındadır.

3.2. Yöntem

3.2.1. Literatürde Mevcut Yöntem (z testi)

Belli bir eşik olasılıkta (c), duyarlılık (sensitivite), $se(c) = P(T(c) = 1|D = 1)$ özgüllük (spesifite) ise $sp(c) = P(T(c) = 0|D = 0)$ olarak gösterilmektedir. Hastalık prevalansı π olan bir popülasyondan, n bireylik örnek çekildiğinde n_0 sağlıklıların sayısı, n_1 hastaların sayısı olmak üzere $n = n_0 + n_1$ olacaktır. D_i gerçek durum, $\hat{T}_i(c)$ belirli bir eşik olasılıkta test sonucu ise bireylerin her biri ($i=1,2,..,n$) için $\hat{T}_i(c) = 1\{\hat{p}_i \geq c\}$ 'dir. Burada parantez içerisinde yer alan $\hat{p}_i$ örneklem için uydurulan tahmin modelinden elde edilen tahmin olasılıklarıdır. Bu olasılıklar eşik olasılıktan büyük ise test sonucu 1 (hastalık var) olacaktır. Sensitivite, spesifite ve prevalans için deneysel tahmin ediciler $\widehat{se}(c)$, $\widehat{sp}(c)$, ve $\hat{\pi}$ sırası ile;

$$\widehat{se}(c) = \frac{\sum_{i=1}^n 1\{\hat{T}_i(c)=1, D_i=1\}}{\sum_{i=1}^n 1\{D_i=1\}}$$

$$\widehat{sp}(c) = \frac{\sum_{i=1}^n 1\{\hat{T}_i(c)=0, D_i=0\}}{\sum_{i=1}^n 1\{D_i=0\}}$$

$$\hat{\pi} = \frac{\sum_{i=1}^n 1\{D_i=1\}}{n}$$

Burada; $\sum_{i=1}^n 1\{D_i = 1\} = n_1$ ve $\sum_{i=1}^n 1\{D_i = 0\} = n_0$ 'dır.

Ayrıca

$$\hat{T}(c)|D = 1 \sim \text{Bernoulli}(se(c))$$

$$\hat{T}(c)|D = 0 \sim \text{Bernoulli}(1-sp(c)) \text{ ise}$$

$$n_1 \widehat{se}(c) \sim \text{Binomiyal}(n_1, se(c)),$$

$$n_0 \widehat{sp}(c) \sim \text{Binomiyal}(n_0, sp(c)) \text{ ve}$$

$$n, \hat{\pi} \sim \text{Binomiyal}(n, \pi) \text{ 'dir.}$$

$$\sigma_{\widehat{se}(c)}^2 = var(\widehat{se}(c)) = se(c) \cdot \left(1 - \frac{se(c)}{n_1}\right)$$

$$\sigma_{\widehat{sp}(c)}^2 = var(\widehat{sp}(c)) = sp(c) \cdot \left(1 - \frac{sp(c)}{n_0}\right)$$

$$\sigma_{\widehat{\pi}}^2 = var(\widehat{\pi}) = \pi \cdot (1 - \pi)/n$$

olduğunda varyansların tahminleri de deneysel olarak sırası ile aşağıdaki gibi tahmin edilebilir;

$$\widehat{\sigma}_{\widehat{se}(c)}^2 = \frac{\widehat{se}(c)(1 - \widehat{se}(c))}{n_1}$$

$$\widehat{\sigma}_{\widehat{sp}(c)}^2 = \frac{\widehat{sp}(c)(1 - \widehat{sp}(c))}{n_0}$$

$$\widehat{\sigma}_{\widehat{\pi}}^2 = \widehat{\pi}(1 - \widehat{\pi})/n$$

dir. Böylece Net Fayda aşağıdaki gibi tahmin edilir.

$$\widehat{\phi}(c) = \widehat{\pi} \cdot \widehat{se}(c) - (1 - \widehat{\pi}) \cdot (1 - \widehat{sp}(c)) \frac{c}{1 - c}$$

$E(\widehat{se}(c)) = se(c)$, $E(\widehat{sp}(c)) = sp(c)$, $E(\widehat{\pi}) = (\pi)$ 'dir.

Ayrıca $\widehat{se}(c)$, $\widehat{sp}(c)$ ve $\widehat{\pi}$ sırası ile $se(c)$, $sp(c)$ ve π için tutarlı tahmin ediciler olduğundan $\widehat{\phi}(c)$ Net Fayda tahmini de $\phi(c)$ 'nin tutarlı tahmin edicisidir.

Merkezi limit teoremi kullanılarak $\widehat{se}(c)$, $\widehat{sp}(c)$ ve $\widehat{\pi}$ 'nin sensitivite ve spesifite değerlerinin gerçek değerler etrafında nasıl dağıldığını anlamak için asimptotik dağılımları elde edilmeye çalışıldığında, ($\frac{n_1}{n} \rightarrow_p r$ ve $\frac{n_0}{n} \rightarrow_p 1-r$ ($0 < r < 1$ kabulü ile)

$n \rightarrow \infty$

$$\sqrt{n} \begin{pmatrix} \widehat{se}(c) - se(c) \\ \widehat{sp}(c) - sp(c) \\ \widehat{\pi} - \pi \end{pmatrix} \xrightarrow{d} N(0, \Psi(c))$$

Gerçek değerler ile tahminler arasındaki farkların normalleştirilmiş değerleri 0 ortalamalı, $\Psi(c)$ kovaryans matrisli bir normal dağılıma yakınsamaktadır.

Simetrik varyans kovaryans matrisi

$$\Psi(c) = \begin{pmatrix} \gamma_{\widehat{se}(c)}^2 & \gamma_{(\widehat{se}(c), \widehat{sp}(c))} & \gamma_{(\widehat{se}(c), \widehat{\pi})} \\ \cdot & \gamma_{\widehat{sp}(c)}^2 & \gamma_{(\widehat{sp}(c), \widehat{\pi})} \\ \cdot & \cdot & \gamma_{\widehat{\pi}}^2 \end{pmatrix}$$

Net Faydanın varyansını elde etmek için ise bir düzgünleştirici fonksiyon ele alınır $g(x_1, x_2, x_3) = x_3 \cdot x_1 - (1 - x_3)(1 - x_2) \frac{c}{1-c}$ ve burada kısmi türevler alınıp fonksiyonun gradyan vektörü elde edilir. Bu fonksiyonda x_1, x_2, x_3 sırası ile $\widehat{se}(c), \widehat{sp}(c)$ ve $\widehat{\pi}$ 'dir ve türevleri de, g'nin her bir parametre üzerindeki etkisini ölçmektedir. Dolayısı ile Net Fayda tahmini de

$$\widehat{\phi}(c) = g(\widehat{se}(c), \widehat{sp}(c), \widehat{\pi}) \text{'dir.}$$

Burada gradyan vektörü kullanılmasının nedeni her bir tahminciyi, fonksiyon varyansına dahil etmektir ve δ (delta) metodu ile $g(\widehat{se}(c), \widehat{sp}(c), \widehat{\pi})$ 'nin dolayısı ile de tahmin edilen Net Faydanın asimptotik normal dağılımı üretilir. Gerçek Net Fayda ile tahmin edilen Net Fayda arasındaki farkın, ortalaması 0 ve varyansı $v^2(c)$ olan normal dağılıma yakınsamaktadır ve $v^2(c)$ (varyans) ise gradyan fonksiyonu, varyans kovaryans matrisi ve gradyan fonksiyonunun transpozunun çarpımından elde edilir. Böylece güven aralığı;

$$\widehat{\phi}(c) - z_{\frac{\alpha}{2}} \frac{\widehat{v}(c)}{\sqrt{n}} \leq \phi \leq \widehat{\phi}(c) + z_{\frac{\alpha}{2}} \frac{\widehat{v}(c)}{\sqrt{n}}$$

olarak hesaplanır.

İki model için Net Faydalar sensitivite ve spesifite bakımından sırası ile aşağıdaki gibi yazılabilir.

$$\phi_1(c) = \pi \cdot se_1(c) - (1 - \pi)(1 - sp_1(c)) \frac{c}{1-c}$$

$$\phi_2(c) = \pi \cdot se_2(c) - (1 - \pi)(1 - sp_2(c)) \frac{c}{1-c}$$

Bu durumda, belirli bir eşik olasılıkta (c) iki modele ait Net Faydalar arasındaki fark;

$$\phi_1(c) - \phi_2(c) = \pi \cdot se_1(c) - \pi \cdot se_2(c) - (1 - \pi)(sp_2(c) - sp_1(c)) \frac{c}{1-c}$$

olarak yazılabilir.

k model sayısı olmak üzere ($k=1,2$) $se_k(c)$, $sp_k(c)$ ve π 'nin tahmini;

$$\widehat{se}_k(c) = \frac{\sum_{i=1}^n 1\{\hat{T}_{ki}(c)=1, D_i=1\}}{\sum_{i=1}^n 1\{D_i=1\}},$$

$$\widehat{sp}_k(c) = \frac{\sum_{i=1}^n 1\{\hat{T}_{ki}(c)=0, D_i=0\}}{\sum_{i=1}^n 1\{D_i=0\}},$$

$$\widehat{\pi} = \frac{\sum_{i=1}^n 1\{D_i=1\}}{n}, \text{dir.}$$

Böylece 2 model arasındaki Net Fayda farkı;

$$\hat{\phi}_1(c) - \hat{\phi}_2(c) = \widehat{\pi}(\widehat{se}_1(c) - \widehat{se}_2(c) - (1 - \widehat{\pi})(\widehat{sp}_2(c) - \widehat{sp}_1(c)) \frac{c}{1-c}$$

olarak yazılabilir.

$\widehat{se}_k(c)$, $\widehat{sp}_k(c)$ ve $\widehat{\pi}$ sırası ile $se_k(c)$, $sp_k(c)$ ve π nin $k=1,2$ için yansız tahmin edicisi olduğundan $\hat{\phi}_1(c) - \hat{\phi}_2(c)$ de $\phi_1(c) - \phi_2(c)$ farkının yansız tahmin edicisidir. İki karar eğrisi karşılaştırmasından çıkarım yapabilmek için aşağıdaki hipotezlerin test edilmesi ve güven aralığının hesaplanması gerekmektedir.

$$H_0: \phi_1(c) - \phi_2(c) = 0.$$

$$H_1: \phi_1(c) - \phi_2(c) \neq 0.$$

Hipotez testi için $\hat{\phi}_1(c) - \hat{\phi}_2(c)$ 'nin güven aralığı hesaplanmalıdır. $\phi_1(c) - \phi_2(c) = G(se_1(c), sp_1(c), se_2(c), sp_2(c), \pi)$ dir ve yine delta (δ) yöntemi uygulandığında varyans $\eta^2(c)$ grandyan fonksiyon, varyans kovaryans matrisi ve gradyan fonksiyonunun transpozunun çarpımı ile elde edilir.

$$\sqrt{n}((\hat{\phi}_1(c) - \hat{\phi}_2(c)) - (\phi_1(c) - \phi_2(c))) \xrightarrow{d} N(0, \eta^2(c))$$

Bu varyans $\hat{\eta}^2(c)$ güven aralığı hesabında kullanılacaktır.

$\phi_1(c) - \phi_2(c)$ farkına ilişkin $100(1 - \alpha)\%$ güven aralığı;

$$(\hat{\phi}_1(c) - \hat{\phi}_2(c)) - z_{\alpha/2} \hat{\eta}(c) / \sqrt{n} \leq \phi_1(c) - \phi_2(c) \leq (\hat{\phi}_1(c) - \hat{\phi}_2(c)) + z_{\alpha/2} \hat{\eta}(c) / \sqrt{n}$$

Oluşturulan güven aralığı kullanılarak, iki Net Fayda arasındaki farka ilişkin hipotezin test edilmesi için bir sonuca varılabilmektedir. Güven aralığı sıfırı içermiyorsa, iki karar eğrisi arasında önemli bir fark olduğu iddia edilerek α anlamlılık seviyesinde sıfır hipotezi reddedilecek; aksi takdirde, sıfır hipotezi reddedilmeyecektir²³.

3.2.2. ANOVA/Kruskal Wallis Yaklaşımı

Donna Katzman McCLISH tarafından 1987 yılında yapılan “Comparing the Areas Under More Two Independent ROC Curves” başlıklı çalışmadan hareketle bu tezde, ikiden fazla modelden elde edilen Net Fayda değerlerinin karşılaştırılmasında yeniden örnekleme tekniği (bootstrap) kullanılarak elde edilen sahte (pseudo) Net Fayda değerlerinin dağılım bilgisine göre karar verilen ANOVA veya Kruskal Wallis yöntemlerinin kullanılması önerilmektedir⁷⁶.

Bu çalışmada da belli bir eşik olasılıkta (pt) tahmin modellerinden elde edilen Net Faydaların istatistik olarak eş zamanlı karşılaştırmasını yapabilmek için yeniden örnekleme tekniklerinden Bootstrap ile sahte (pseudo) Net Faydalar üretilmiştir. Üretilen sahte (pseudo) Net Faydalar arasındaki farkların istatistik olarak incelenmesinde; literatürde önerilen yöntemin (z testi) parametrik yöntem olduğu gerekçesi ile ANOVA yöntemi kullanılırken; bazı parametrik olmayan durumlar için ise (Net Fayda değerlerine ait dağılımın, bazı durumlarda normal dağılımdan uzaklaştığı gerekçesi ile) Kruskal Wallis (KW) testi kullanılmıştır. Yeniden örnekleme tekniklerinden en sık kullanılanlar Bootstrap ve Jackknife teknikleridir. Bu tez çalışmasında ön deneme olarak Jackknife tekniği denenmiş ancak bu tekniğin liberal test olma yönünde eğilim gösterdiği gözlemlendiğinden örnekleme işlemleri Bootstrap tekniği ile gerçekleştirilmiştir. Bootstrap tekniği, Efron tarafından 1979'da yayınlanan standart hata ile güven aralığı hesaplayan ve önem testleri gerçekleştirerek çıkarım yapmaya olanak sağlayan bilgisayar temelli istatistiksel bir yöntemdir. Tekniğin bazı avantajları arasında; az varsayım gerektirmesi, stabilitenin sağlanması, yanlılığı azaltması, genellenebilirliği desteklemesi sayılabilmektedir⁷⁷⁻⁷⁹. Bootstrap ile ilgilenilen popülasyondan tekrarlanan örneklerin (aynı boyutta) çok sayıda çekilebilmekte, bu tekrarlanan örneklerden belirli bir istatistik hesabı yapılabilmekte ve örnekleme dağılımı hakkında fikir edinilebilmektedir. Bootstrap, bir istatistiğin örnekleme dağılımını yaklaşık olarak belirlemek amacıyla, elde edilen örneklem verilerini bir "vekil popülasyon" olarak kullanmaktır; başka bir ifade ile, eldeki örneklem verilerinden yeniden örnekleme yapmak (yerine koyarak) ve bootstrap örnekleri olarak bilinen, çok sayıda hayalet/sanal örnek oluşturmaktır.

Veri Üretme ve Değerlendirme Süreci

Özet istatistiği (Net Fayda) ise bootstrap örneklerinin her biri için hesaplanmaktadır⁸⁰. Net Faydaların elde edilmesinde kullanılan tahmin modelleri; veri sınıflandırma ve regresyon problemlerinin çözümü amacıyla ortaya atılan⁸¹ yanıt değişkeni kategorilerinin (2 kategorili) doğru sınıflandırılabilmesi için kategoriler arasındaki en geniş boşluğu (marjini) oluşturacak şekilde farklı sınıflara ayıran optimal bir hiperdüzlem bulan Destek Vektör Makineleri (DVM)^{82, 83} Tümevarım mantığı ile, sınıfı bilinen örneklerden hareketle genelleme yapabilmek için oldukça pratik ve kullanışlı bir yöntem olan ve etiketi bilinen örnekleri, belirli bir durdurma kriteri karşılanana kadar tekrarlı olarak dallara ayırabilen Karar Ağacı (KA)⁸⁴⁻⁸⁶ ve yanıt değişkenine ait kategorilerden birine ait olma olasılığını (0 ile 1 arasında bir değer olarak) tahminleyen, sürekli veya kategorik açıklayıcı değişkenlerden oluşan Lojistik Regresyon (LR) 'dur^{41, 87}.

Uygulanan yöntem aşağıdaki sıralama ile gerçekleştirilmiştir;

1.Aşama: Bootstrap

1. Belirli bir eşik olasılık seçilmiştir. (Örneğin pt:0.10)
2. Veri setinden seçilen n gözlemleri bir örnek üzerinden sırası ile %70, %30 oranında gözlem içerecek şekilde eğitim ve test seti (yanıt değişkeninin prevalans oranı korunacak şekilde) oluşturulmuştur.
3. Seçilen eğitim setinde tahmin modelleri kurulmuştur (DVM, KA, LR).
4. Seçilen test setine tahmin modeli uygulanmıştır.
5. Tahmin modellerinden Net Faydalar hesaplanmıştır.
6. Bu işlemler seçilen eşik olasılıkta (2.-5.adımlar) bootstrap sayısı kadar (R: 100) gerçekleştirilmiştir.
7. Bootstrap sayısınınca Net Fayda elde edilirken, z yaklaşımı ile her bootstrapta Net Faydaları ikişerli karşılaştırması yapılmıştır.
8. z yaklaşımı 2'şerli karşılaştırma yaptığı ve bu çalışmada da 3 model karşılaştırması ANOVA/Kruskal Wallis yaklaşımı üzerinden yapıldığı için z testi ile her 3 modelin 2'şerli karşılaştırmasının modeller arasında Net Fayda bakımından istatistik olarak fark bulunmayan durum 'z total' olarak ifade edilip 1 kodlanmış en az 2 model arasında fark varsa bu durum 0 kodlanmıştır.

Böylece z testinin 3 model arasında Net Fayda bakımından istatistik olarak fark bulup bulmadığı ‘z total’ değeri ile anlaşılmıştır.

9. Ayrıca bu aşamada modellerden elde edilen Net Fayda ortalaması ve standart sapması ile prevalansların ortalaması ve standart sapması kaydedilmiştir.

2. Aşama: ANOVA/Kruskal Wallis yaklaşımı ile z testinin karşılaştırılması

1. İlk aşamada modellerden bootstrap sayısınca elde edilen Net faydalara ANOVA/Kruskal Wallis testleri uygulanmıştır ve her iki test için 1’er adet p değeri elde edilmiştir. Bu p değerleri de her iki test için yine Net Fayda bakımından fark bulunmadığı durumda 1, bulunduğu duruda 0 olarak kodlanırken; z testi ile her tekrarda ikişerli model karşılaştırması için elde edilen ‘z total’ değerleri de bu adımda toplanmış ve ‘z toplam’ olarak kaydedilmiştir.
2. 2. aşamanın ilk adımı test karşılaştırma sayısı (R: 100) kadar (ANOVA/Kruskal Wallis ve z testi) tekrarlanmıştır.
3. Bu durumda ANOVA/Kruskal Wallis testlerinden test karşılaştırma sayısı kadar 1 ve 0’ lar gelirken z için de tekrar sayısı adetince ‘z toplam’ elde edilmiştir.
4. Ayrıca burada her testin yapıldığı veri setinden hesaplanan Net fayda ortalamaları ve standart sapmaları ile prevalans ortalamaları ve standart sapmalarının ortalaması ve standart sapması hesaplanmıştır.

3. Aşama: Simülasyon

1. Simülasyon için ANOVA/Kruskal Wallis testlerinden elde edilen 1, 0 lar toplanırken z toplamlarının da ortalaması alınmıştır.
2. 3. Aşamanın 1. adımı simülasyon sayısınca (R: 100) tekrarlanmıştır, böylece tekrar sayısı kadar ANOVA/Kruskal Wallis test sonuç toplamaları, z testi için de test sonuç ortalamaları elde edilmiştir.
3. Önceki aşamada elde edilen Net fayda ve prevalans ortalamalarının bu adımdaki simülasyon sayısına göre ortalaması alınmıştır.

4. Aşama: Özet

1. ANOVA/Kruskal Wallis ve z testi toplamalarının ortalaması ve standart sapması ile z ortalamaların da ortalaması standart sapması alınmıştır.
2. Net faydalara ait ortalama ve standart sapma ile prevalansa ait ortalama ve standart sapma hesaplanmıştır.

5.Aşama: Yorum

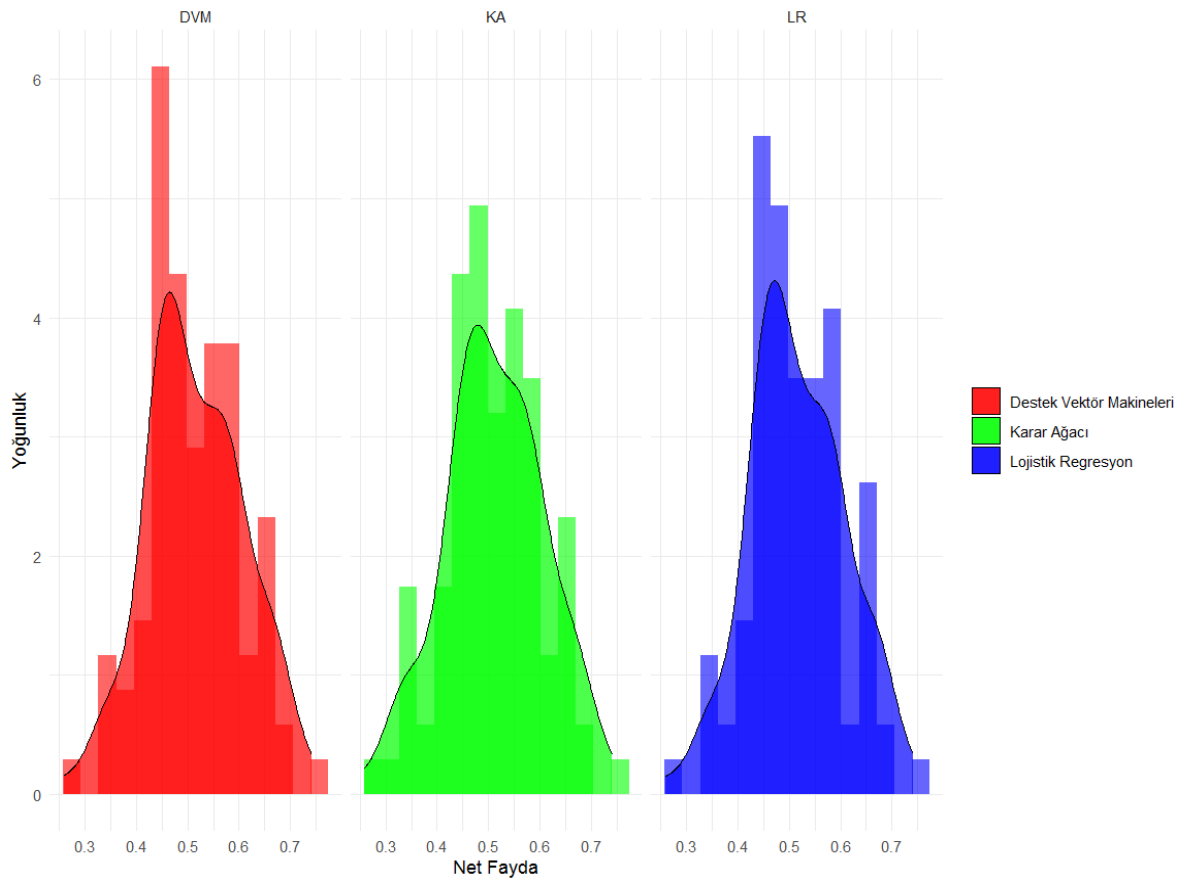
Net faydalara ait ortalama ve standart sapmalardan hareketle, modeller arasındaki istatistik olarak farkın tespiti, etki büyüklüğü değerleri üzerinden yorumlanmıştır. Etki büyüklüklerine göre belirlenen örnek büyüklüğünün yeterliği, yöntemler üzerinden (ANOVA, KW, z) tartışılmıştır. Etki büyüklüğü Cohen d ölçütü ile belirlenmiş ve etki büyüklüğüne göre de örnek büyüklüğü hesabı yapılmıştır. Cohen d değeri 0.20 iken küçük etki büyüklüğü, 0.50 iken orta etki büyüklüğü, 0.80 ve üzerini ise büyük etki büyüklüğü kategorilerinde yer aldığından⁸⁸ elde edilen sonuçlar da bu kategorilere göre yorumlanmıştır.

Gerçek veri çalışmasında ise ilk aşama bootstrap sayısı kadar tekrarlanmış ve ikinci aşama yani testlerin (ANOVA/Kruskal Wallis ile z testlerinin) karşılaştırma aşaması 1 kez tekrarlanmıştır. Bu durumda ‘z total’ değeri bootstrap sayısı kadar hesaplandığından, gerçek veri sonuç çizelgelerinde z sütunu yüzde olarak ifade edilmiştir.

4. BULGULAR

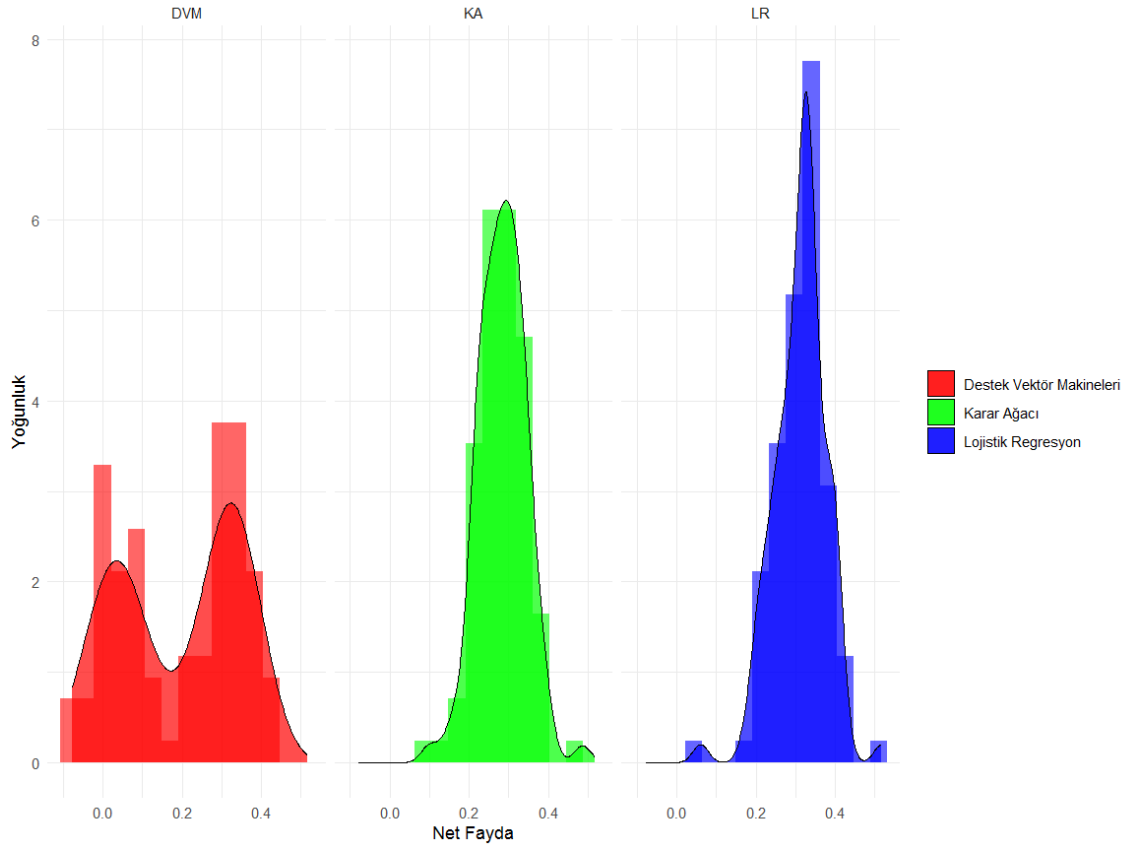
4.1. Simüle Veriler için Elde Edilen Bulgular

Dört farklı senaryoda oluşturulan veri setlerinden elde edilen 3 tahmin modelinden (DVM, KA, LR) hesaplanan Net Faydaların, istatistik olarak karşılaştırılabilmesi amacı ile Net Fayda ortalamaları arasındaki etki büyüklüğü ve bu etki büyüklüğüne bağlı olarak örnek büyüklüğü hesaplanmıştır. Modellerden hesaplanan Net Faydaların dağılımı incelendiğinde, genel olarak küçük etki düzeyinde normal dağılıma uygunluk gösterirken (Şekil 4.1); yüksek etki düzeyinde özellikle de DVM modelinde normal dağılımdan bir miktar sapmalar (binormal) olduğu tespit edilmiştir (Şekil 4.2-4.5). Dolayısı ile etki büyüklüğü ve örnek büyüklüğüne göre ANOVA testinin yanında Kruskal Wallis testi sonuçları da yorumlanmıştır.

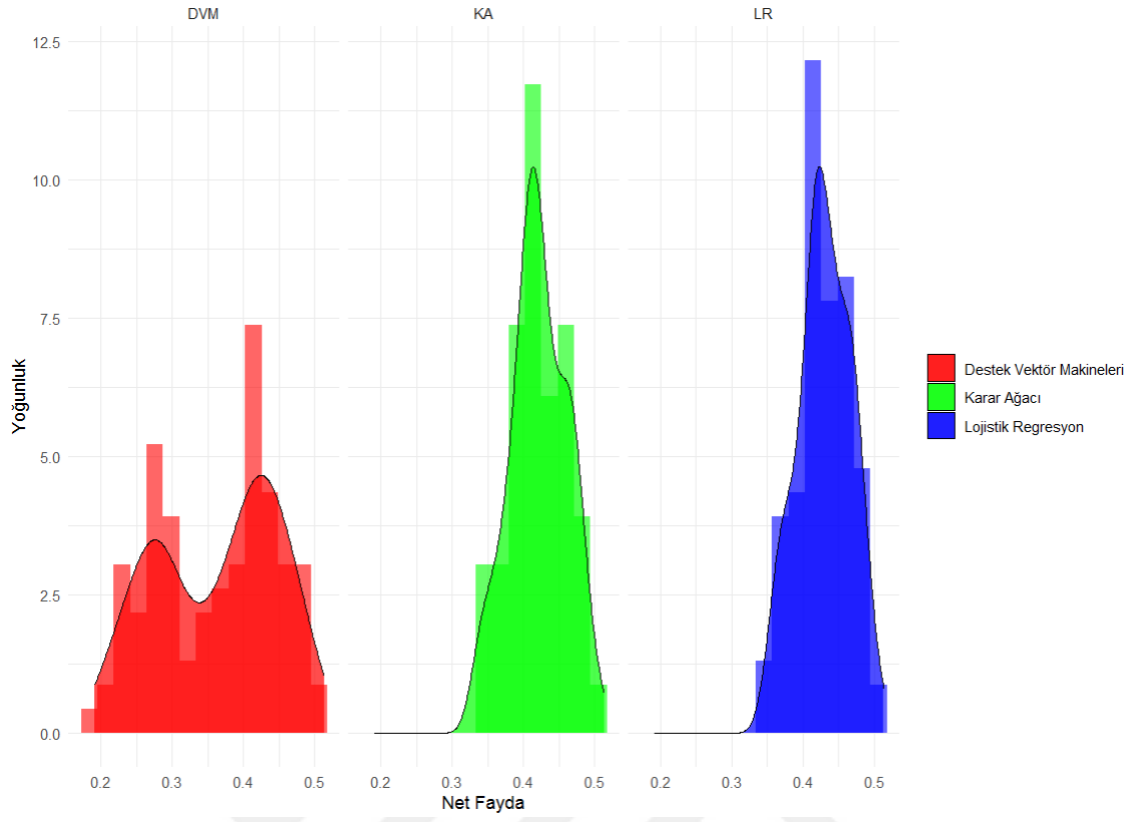


Şekil 4.1. Senaryo 1 için etki büyüklüğü küçük iken modellere göre net faydaların dağılımı

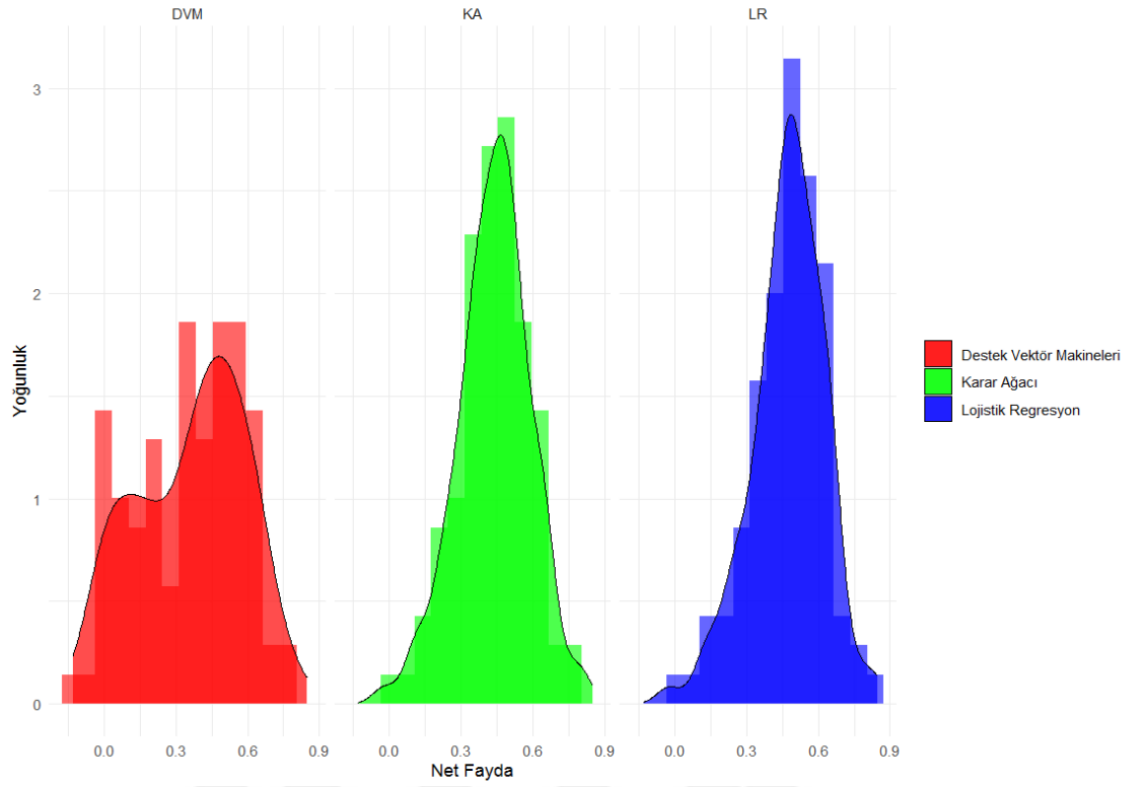
Tüm senaryolarda etki büyüklüğü 0,5'in altında iken tahmin modellerinden elde edilen Net Fayda dağılımının normal dağılıma uygun olduğu söylenebilir. Örnek olarak sunulmuş olan Şekil 4.1, 1. Senaryoda etki büyüklüğü 0.04, eşik olasılık 0,10, n=100 iken modellerden elde edilen Net Faydaların dağılımlarını içermektedir ve her 3 tahmin modelinin de normal dağılıma uygun olduğu söylenebilir.



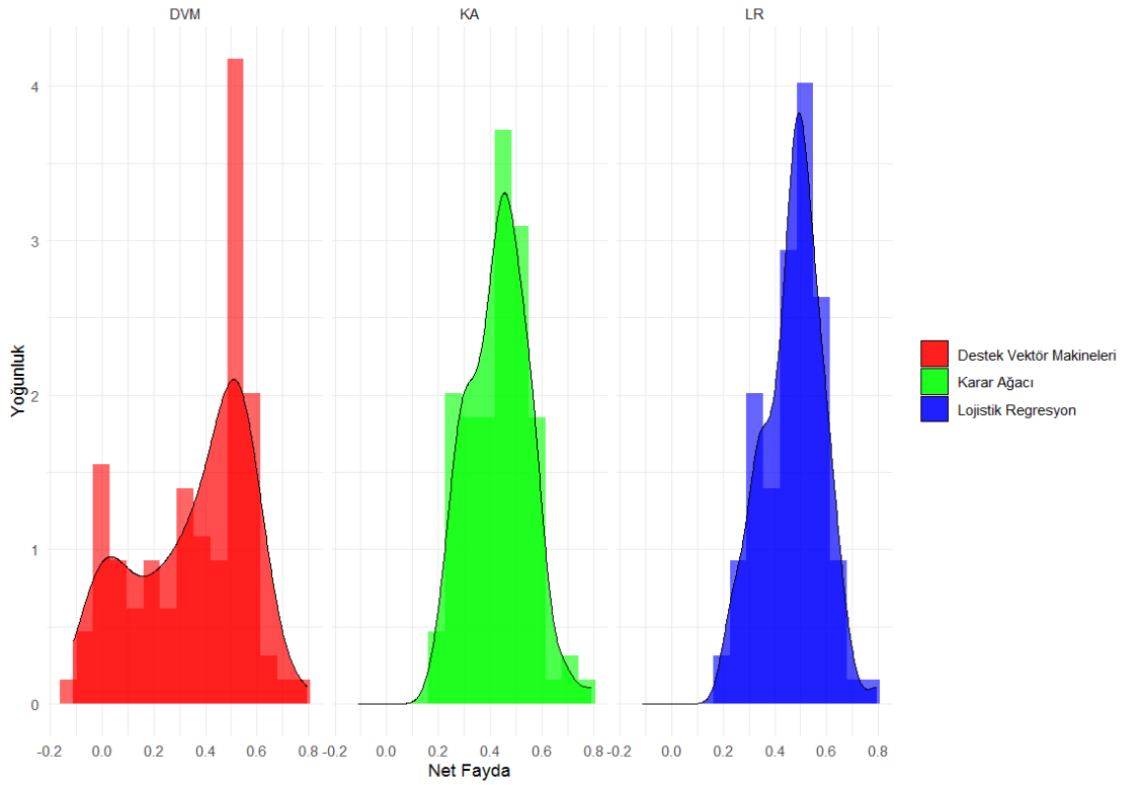
Şekil 4.2. Senaryo 1 için etki büyüklüğü büyük iken modellere göre net faydaların dağılımı



Şekil 4.3. Senaryo 2 için etki büyüklüğü büyük iken modellere göre net faydaların dağılımı



Şekil 4.4. Senaryo 3 için etki büyüklüğü büyük iken modellere göre net faydaların dağılımı



Şekil 4.5. Senaryo 4 için etki büyüklüğü büyük iken modellere göre net faydaların dağılımı

Senaryo 1 için etki büyüklüğü yaklaşık 1 olduğunda, senaryo 2 için etki büyüklüğü yaklaşık 0.5 ve üzeri olduğunda, senaryo 3 için etki büyüklüğü yaklaşık 0.7 ve üzeri olduğunda, senaryo 4 için ise etki büyüklüğü yaklaşık 0.7 ve üzeri olduğunda özellikle DVM modelinin binormal dağılıma yaklaştığı söylenebilir. Örneğin; Şekil 4.2. senaryo 1 için etki büyüklüğü 1.09, eşik olasılık 0.40, $n=500$ iken modellere göre Net Faydaların dağılımı, Şekil 4.3, senaryo 2 için etki büyüklüğü 0.99, eşik olasılık 0.25, $n=1000$ iken modellere göre Net Faydaların dağılımı, Şekil 4.4, senaryo 3 için etki büyüklüğü 0.76, eşik olasılık 0.30, $n=250$ iken modellere göre Net Faydaların dağılımı, Şekil 4.5 ise senaryo 4 için etki büyüklüğü 0.83, eşik olasılık 0.35, $n=100$ iken modellere göre Net Faydaların dağılımını göstermektedir. Bu şekiller incelendiğinde tüm senaryolar için etki büyüklüğü 1'e yaklaştığında yani çok büyük olduğunda DVM modelinin çift tepeli normal dağılıma yaklaştığı söylenebilmektedir.

Çizelge 4.1. Senaryo 1 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.10 \leq p_t \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±sd)
0.10	100	9715	0.040	0.519±0.100	0.521±0.099	0.517±0.100	100.000±0.000	100.000±0.000	96.971±0.167
0.10	250	3958	0.063	0.519±0.064	0.522±0.063	0.518±0.064	100.000±0.000	100.000±0.000	93.408±0.235
0.10	500	1302	0.110	0.518±0.046	0.523±0.045	0.520±0.045	100.000±0.000	100.000±0.000	88.990±0.319
0.10	1000	665	0.154	0.517±0.033	0.522±0.032	0.522±0.032	100.000±0.000	100.000±0.000	79.887±0.419
0.15	100	2159	0.085	0.485±0.108	0.491±0.104	0.482±0.107	100.000±0.000	100.000±0.000	92.445±0.221
0.15	250	584	0.164	0.483±0.071	0.494±0.066	0.483±0.068	99.960±0.197	99.970±0.171	81.280±0.408
0.15	500	198	0.282	0.480±0.052	0.494±0.047	0.487±0.048	80.710±4.140	87.560±3.208	65.711±0.439
0.15	1000	75	0.463	0.478±0.040	0.495±0.033	0.493±0.034	0.970±0.870	5.210±2.319	49.518±0.512
0.20	100	869	0.135	0.446±0.118	0.459±0.110	0.444±0.113	100.000±0.000	100.000±0.000	86.240±0.294
0.20	250	147	0.328	0.438±0.082	0.463±0.070	0.447±0.072	61.670±4.441	72.750±4.314	67.945±0.415
0.20	500	57	0.529	0.432±0.066	0.463±0.050	0.453±0.051	0.280±0.552	2.000±1.563	48.993±0.456
0.20	1000	28	0.771	0.428±0.056	0.464±0.035	0.461±0.036	0.000±0.000	0.020±0.141	40.657±0.450
0.25	100	412	0.196	0.399±0.130	0.423±0.115	0.402±0.118	98.650±1.201	99.120±0.879	81.104±0.351
0.25	250	68	0.487	0.386±0.097	0.428±0.074	0.404±0.075	0.720±0.965	2.570±1.635	57.153±0.467
0.25	500	30	0.744	0.377±0.084	0.429±0.052	0.412±0.054	0.000±0.000	0.000±0.000	42.276±0.524
0.25	1000	24	0.831	0.372±0.078	0.430±0.037	0.423±0.038	0.000±0.000	0.000±0.000	41.242±0.483
0.30	100	168	0.307	0.344±0.145	0.385±0.121	0.353±0.122	57.170±4.588	66.890±4.449	75.230±0.442
0.30	250	38	0.651	0.327±0.115	0.391±0.078	0.355±0.078	0.000±0.000	0.000±0.000	50.009±0.495
0.30	500	21	0.900	0.317±0.106	0.393±0.055	0.366±0.056	0.000±0.000	0.000±0.000	39.359±0.538
0.30	1000	16	1.053	0.312±0.103	0.394±0.039	0.381±0.040	0.000±0.000	0.000±0.000	40.572±0.486

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

Çizelge 4.1-4.8 bazı eşik olasılıklarında farklı gözlem sayıları ve tekrar sayılarında oluşturulan tahmin modellerinden (DVM, LR, KA) hesaplanan Net Fayda değerleri ve bu değerlerin birbirinden farklı olup olmadığını istatistik olarak hesaplayan ANOVA, Kruskal Wallis ve z testi sonuçlarını içermektedir. Test sonuçları 100 simülasyon üzerinden modeller arasında Net Fayda bakımından fark hesaplanmadığı durumların ortalaması ve standart sapmasını ifade etmektedir.

Çizelge 4.1'de ANOVA testi, etki büyüklüğü önemsiz seviyede ve eşik olasılık 0.10 iken 100 simülasyonun 100'ünde tüm gözlem sayılarında modeller arasında Net Fayda ortalamaları bakımından istatistik olarak fark hesaplanmazken, z testi örnek büyüklüğü arttıkça ortalama fark yok sayısında azalış göstermektedir. Etki büyüklüğünün yine önemsiz sayılabileceği seviyelerde eşik olasılık 0.15, n=100 ve n=250 iken ANOVA testide modeller arasında Net Fayda ortalamaları bakımından

ortalama fark yok sayısı 100 iken z testinde bu sayı sırası ile yaklaşık 92 ve 81 olarak hesaplanmıştır.

Etki büyüklüğünün küçük olduğu seviyelere gelindiğinde, eşik olasılık 0.15, $n=500$ iken ANOVA ve z testinde ortalama fark yok sayısının sırası ile 81 ve 66 olduğu yani etki büyüklüğü arttıkça ortalama fark yok sayısında bir düşüş olduğu gözlemlenmiştir. Aynı eşik olasılıkta, $n=1000$ iken etki büyüklüğü orta seviyelere çıktığında ise ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı 75 olarak hesaplanmıştır ve bu senaryoda ANOVA testi 100 simülasyonun yalnızca 5 tanesinde z testi ise 50 tanesinde fark bulmamıştır. Etki büyüklüğü önemsiz seviyelerde, eşik olasılık 0.20, $n=100$ iken ANOVA testi modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı 100 iken z testinde bu sayı yaklaşık olarak 86 olarak hesaplanmıştır. Etki büyüklüğünün küçük seviyelere geldiğinde, eşik olasılık 0.20, $n=250$ iken ANOVA ve z testinde ortalama fark yok sayıları sırası ile 62 ve 68 olarak hesaplanmıştır. Aynı eşik olasılıkta, etki büyüklüğünün orta ve büyük olduğu seviyelerine gelindiğinde ise $n=500$, $n=1000$ iken ANOVA testi 100 simülasyonda farkın neredeyse tamamını yakalamasına rağmen z testi için ortalama fark yok sayısı sırası ile 49, 41 olarak hesaplanmıştır. Etki büyüklüğü önemsiz seviyelerinde, eşik olasılık 0.25, $n=100$ iken ANOVA testinde modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı 99 iken z testinde bu sayı yaklaşık olarak 81 olarak hesaplanmıştır. Etki büyüklüğünün orta seviyelerine geldiğinde, eşik olasılık 0.25, $n=250$ iken ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı 68 olarak hesaplanmıştır ve 100 simülasyonda ANOVA ve z testinde ortalama fark yok sayıları sırası ile yaklaşık 1 ve 57 olarak hesaplanmıştır. Aynı eşik olasılıkta, $n=500$ ve $n=1000$ iken etki büyüklüğünün büyük olduğu seviyelere gelindiğinde ise ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı sırası ile 30 ve 24'tür ve ANOVA testi 100 simülasyonda farkın neredeyse tamamını yakalamasına rağmen z testi için ortalama fark yok sayısı sırası ile yaklaşık 42 ve 41 olarak hesaplanmıştır.

Etki büyüklüğünün küçük sayılabileceği seviyelerde eşik olasılık 0.30, $n=100$ iken ANOVA testinde modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı 57 iken z testinde bu sayı yaklaşık olarak 75 olarak hesaplanmıştır. Aynı eşik olasılıkta, $n=250$, $n=500$ ve $n=1000$ iken etki büyüklüğünün büyük olduğu seviyelere gelindiğinde ise ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı 38, 21, 16 iken ANOVA testi ve Kruskal Wallis testi 100 simülasyonda

farkın neredeyse tamamını yakalamasına rağmen z testi için ortalama fark yok sayısı sırası ile yaklaşık 50, 39 ve 41 olarak hesaplanmıştır.

Çizelge 4.2. Senaryo 1 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq p_t \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.35	100	19	0.941	0.362±0.215	0.487±0.113	0.453±0.114	0.000±0.000	0.350±0.592	59.400±0.459
0.35	250	22	0.872	0.355±0.206	0.491±0.079	0.456±0.079	0.000±0.000	0.000±0.000	49.599±0.433
0.35	500	19	0.944	0.353±0.202	0.493±0.056	0.463±0.056	0.000±0.000	0.000±0.000	44.668±0.469
0.35	1000	18	0.980	0.354±0.198	0.494±0.040	0.475±0.040	0.000±0.000	0.000±0.000	44.235±0.486
0.40	100	27	0.780	0.310±0.224	0.450±0.119	0.412±0.119	0.000±0.000	0.020±0.141	58.478±0.505
0.40	250	21	0.896	0.309±0.215	0.455±0.083	0.417±0.083	0.000±0.000	0.000±0.000	49.644±0.466
0.40	500	19	0.959	0.310±0.210	0.458±0.059	0.426±0.059	0.000±0.000	0.000±0.000	45.855±0.468
0.40	1000	18	0.980	0.312±0.208	0.459±0.042	0.439±0.042	0.000±0.000	0.000±0.000	45.407±0.459
0.45	100	26	0.794	0.267±0.224	0.411±0.125	0.371±0.125	0.000±0.000	0.020±0.141	58.765±0.445
0.45	250	21	0.906	0.269±0.214	0.417±0.087	0.378±0.087	0.000±0.000	0.000±0.000	50.487±0.492
0.45	500	18	0.969	0.270±0.210	0.420±0.062	0.390±0.062	0.000±0.000	0.000±0.000	47.853±0.455
0.45	1000	18	0.989	0.272±0.210	0.422±0.044	0.404±0.044	0.000±0.000	0.000±0.000	47.917±0.534
0.50	100	27	0.780	0.226±0.218	0.368±0.137	0.330±0.139	0.000±0.000	0.010±0.100	61.587±0.446
0.50	250	20	0.926	0.229±0.207	0.377±0.091	0.340±0.093	0.000±0.000	0.000±0.000	52.637±0.493
0.50	500	18	0.978	0.232±0.204	0.380±0.065	0.353±0.066	0.000±0.000	0.000±0.000	49.565±0.522
0.50	1000	17	0.994	0.235±0.204	0.382±0.046	0.365±0.047	0.000±0.000	0.000±0.000	49.801±0.519
0.55	100	26	0.793	0.186±0.206	0.325±0.138	0.286±0.144	0.000±0.000	0.000±0.000	64.174±0.512
0.55	250	20	0.928	0.192±0.194	0.334±0.096	0.296±0.101	0.000±0.000	0.000±0.000	53.551±0.517
0.55	500	18	0.986	0.196±0.192	0.338±0.068	0.307±0.072	0.000±0.000	0.000±0.000	49.502±0.503
0.55	1000	17	1.003	0.199±0.193	0.340±0.048	0.318±0.052	0.000±0.000	0.000±0.000	48.352±0.512
0.60	100	28	0.764	0.142±0.199	0.277±0.151	0.234±0.164	0.000±0.000	0.000±0.000	68.385±0.401
0.60	250	20	0.920	0.154±0.180	0.288±0.100	0.244±0.110	0.000±0.000	0.000±0.000	55.696±0.537
0.60	500	17	0.991	0.160±0.176	0.293±0.071	0.253±0.080	0.000±0.000	0.000±0.000	48.199±0.518
0.60	1000	17	1.007	0.165±0.177	0.296±0.050	0.262±0.057	0.000±0.000	0.000±0.000	44.830±0.583

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

Çizelge 4.2’de bazı eşik olasılıklarında tahmin modellerinden hesaplanan Net Fayda değerlerine ait ortalamaların Çizelge 4.1’de yer alan ortalamalara göre birbirinden uzaklaştığı görülmektedir. Ayrıca bu çizelgede etki büyüklüklerinin 1’ e yaklaştığı dolayısı ile daha küçük bootstraplarda Net Faydalar arası farkın tespit edilebilmesi olağandır. Etki büyüklüğü değeri 1’e yaklaştıkça DVM modelinden hesaplanan Net Faydaların dağılımının binormal dağılıma benzediği gerekçesi ile bu

çizelgede yer alan KW testine ait sonuçlar incelendiğinde çizelgede yer alan tüm etki büyüklüklerinde ve eşik olasılıklarında 100 simülasyonun tamamında modeller arasında Net Fayda bakımından ortalama fark yok sayısı 0 olarak hesaplanmıştır.

z testi sonuçları incelendiğinde ise aynı eşik olasılığa (örneğin 0.35) etki büyüklüğü genel olarak örnek büyüklüğü ile birlikte artmakta dolayısı ile ortalama fark yok sayısı azalmaktadır. Etki büyüklüğünün tespiti için gerekli bootstrap ortalama 20'dir ve KW testi 100 simülasyonda bu etki büyüklüklerinde Net Fayda bakımından modeller arasında istatistik olarak anlamlı fark bulmuştur.

Çizelge 4.3. Senaryo 2 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.10 \leq p_t \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.10	100	1732	0.095	0.483±0.108	0.490±0.104	0.480±0.106	100.000±0.000	100.000±0.000	90.952±0.272
0.10	250	2534	0.079	0.516±0.064	0.521±0.063	0.517±0.064	100.000±0.000	100.000±0.000	91.573±0.274
0.10	500	885	0.133	0.515±0.046	0.521±0.044	0.519±0.045	100.000±0.000	100.000±0.000	85.223±0.329
0.10	1000	340	0.215	0.514±0.033	0.521±0.032	0.521±0.032	97.820±1.438	98.580±1.121	72.013±0.413
0.15	100	1732	0.095	0.483±0.108	0.490±0.104	0.480±0.106	100.000±0.000	100.000±0.000	90.967±0.314
0.15	250	383	0.203	0.479±0.072	0.493±0.066	0.482±0.068	99.630±0.597	99.660±0.590	77.757±0.372
0.15	500	128	0.352	0.475±0.055	0.493±0.047	0.486±0.048	36.240±4.899	53.450±5.321	60.125±0.494
0.15	1000	55	0.540	0.472±0.044	0.493±0.033	0.492±0.034	0.020±0.141	0.500±0.659	43.888±0.446
0.20	100	800	0.140	0.442±0.119	0.458±0.109	0.443±0.112	100.000±0.000	99.990±0.100	84.199±0.332
0.20	250	115	0.372	0.433±0.085	0.462±0.070	0.446±0.072	33.110±4.413	51.640±4.925	64.035±0.446
0.20	500	46	0.596	0.426±0.070	0.462±0.049	0.452±0.051	0.040±0.197	0.450±0.672	46.034±0.482
0.20	1000	25	0.824	0.421±0.063	0.463±0.035	0.459±0.036	0.000±0.000	0.000±0.000	39.235±0.528
0.25	100	290	0.233	0.394±0.133	0.423±0.115	0.402±0.117	94.060±1.814	96.740±1.508	78.644±0.393
0.25	250	55	0.541	0.380±0.102	0.428±0.073	0.404±0.075	0.190±0.443	1.340±1.208	54.626±0.449
0.25	500	26	0.796	0.370±0.091	0.429±0.052	0.412±0.053	0.000±0.000	0.000±0.000	41.712±0.470
0.25	1000	17	0.997	0.364±0.086	0.430±0.037	0.422±0.038	0.000±0.000	0.000±0.000	40.317±0.614
0.30	100	126	0.355	0.337±0.149	0.385±0.120	0.355±0.120	31.970±5.128	45.680±5.740	72.670±0.506
0.30	250	33	0.706	0.320±0.122	0.392±0.077	0.357±0.076	0.000±0.000	0.000±0.000	48.634±0.523
0.30	500	19	0.942	0.310±0.114	0.394±0.054	0.367±0.055	0.000±0.000	0.000±0.000	39.642±0.499
0.30	1000	15	1.097	0.304±0.111	0.395±0.038	0.381±0.040	0.000±0.000	0.000±0.000	39.793±0.513

pt: Eşik olasılık, n:Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

Çizelge 4.3'de ANOVA testi, etki büyüklüğü önemsiz seviyede ve eşik olasılık 0.10 iken 100 simülasyonun 100'ünde n=100, 250, 500 gözlem sayılarında modeller arasında Net Fayda ortalamaları bakımından istatistik olarak fark hesaplamazken, z testi

örnek büyüklüğü arttıkça genel olarak ortalama fark yok sayısında azalış göstermektedir. Aynı eşik olasılıkta etki büyüklüğü küçük sayılabilecek seviyelerde, $n=1000$ iken ANOVA testinde modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı yaklaşık 98 iken z testinde bu sayı yaklaşık olarak 72'dir. Etki büyüklüğünün önemsiz sayılabileceği seviyelerde eşik olasılık 0.15, $n=100$ iken ANOVA testinde modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı 100 iken z testinde bu sayı yaklaşık 91 olarak hesaplanmıştır. Etki büyüklüğünün küçük seviyelerine gelindiğinde, eşik olasılık 0.15, $n=250$ iken ANOVA ve z testinde ortalama fark yok sayısının sırası ile yaklaşık 100 ve 78 olduğu gözlemlenmiştir. Etki büyüklüğünün orta seviyelerine gelindiğinde, eşik olasılık 0.15, $n=500$ iken ANOVA ve z testinde ortalama fark yok sayısının sırası ile 36 ve 60 olduğu yani etki büyüklüğü arttıkça ortalama fark yok sayısında bir düşüş olduğu tespit edilmiştir. Aynı eşik olasılıkta, etki büyüklüğü orta seviyelerde DVM modeli binormal dağılıma yaklaştığından (etki büyüklüğü 0.50'nin üzerinde) KW testinin de sonuçları göz önüne alındığında ANOVA ve KW testi 100 simülasyonun yaklaşık 100'ünde z testi ise 44 tanesinde fark bulmamıştır. Etki büyüklüğü önemsiz seviyelerde, eşik olasılık 0.20, $n=100$ iken ANOVA testi modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı 100 iken z testinde bu sayı yaklaşık 84 olarak hesaplanmıştır. Etki büyüklüğü yaklaşık orta seviyelere geldiğinde, eşik olasılık 0.20, $n=250$ iken ANOVA ve z testinde ortalama fark yok sayıları sırası ile 33 ve 64 olarak hesaplanmıştır. Aynı eşik olasılıkta, etki büyüklüğünün orta ve büyük olduğu seviyelere gelindiğinde ise $n=500$, $n=1000$ iken ANOVA ve KW testi 100 simülasyonda farkın neredeyse tamamını yakalamasına rağmen z testi için ortalama fark yok sayısı sırası ile 46 ve 39 olarak hesaplanmıştır. Etki büyüklüğü önemsiz seviyelerde, eşik olasılık 0.25, $n=100$ iken ANOVA testinde modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı 94 iken z testinde bu sayı yaklaşık 79 olarak hesaplanmıştır. Etki büyüklüğünün yaklaşık orta seviyelerinde, eşik olasılık 0.25, $n=250$ iken ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı 55 olarak hesaplanmıştır ve 100 simülasyonda ANOVA, KW ve z testinde ortalama fark yok sayıları sırası ile yaklaşık 0, 1 ve 55 olarak hesaplanmıştır. Aynı eşik olasılıkta, $n=500$, $n=1000$ iken etki büyüklüğünün büyük olduğu seviyelere gelindiğinde ise ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı sırası ile 26 ve 17'dir ve ANOVA ve KW testi 100 simülasyonda farkın neredeyse tamamını yakalamasına rağmen z testi için ortalama fark yok sayısı sırası ile yaklaşık 42 ve 40 olarak

hesaplanmıştır. Etki büyüklüğünün yaklaşık orta seviyelerinde, eşik olasılık 0.30, $n=100$ iken ANOVA testinde modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı 32 iken z testinde bu sayı yaklaşık 73 olarak hesaplanmıştır. Aynı eşik olasılıklıta, $n=250$, $n=500$ ve $n=1000$ iken etki büyüklüğünün büyük olduğu seviyelere gelindiğinde ise ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı 33, 19, 15 iken ANOVA testi ve KW testi 100 simülasyonda farkın neredeyse tamamını yakalamasına rağmen z testi için ortalama fark yok sayısı sırası ile yaklaşık 49, 40 ve 40 olarak hesaplanmıştır.

Çizelge 4.4. Senaryo 2 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq p_t \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.35	100	61	0.514	0.271±0.165	0.346±0.124	0.312±0.122	0.990±0.969	2.660±1.519	67.987±0.468
0.35	250	23	0.859	0.255±0.142	0.354±0.080	0.316±0.078	0.000±0.000	0.000±0.000	46.199±0.525
0.35	500	16	1.036	0.248±0.136	0.356±0.057	0.325±0.056	0.000±0.000	0.000±0.000	40.177±0.488
0.35	1000	14	1.123	0.246±0.134	0.357±0.040	0.339±0.041	0.000±0.000	0.000±0.000	39.927±0.533
0.40	100	37	0.669	0.201±0.178	0.305±0.129	0.273±0.126	0.000±0.000	0.080±0.273	66.110±0.485
0.40	250	18	0.975	0.191±0.158	0.314±0.083	0.275±0.080	0.000±0.000	0.000±0.000	46.097±0.483
0.40	500	15	1.089	0.190±0.154	0.317±0.059	0.286±0.057	0.000±0.000	0.000±0.000	41.751±0.458
0.40	1000	13	1.152	0.190±0.153	0.319±0.041	0.300±0.041	0.000±0.000	0.000±0.000	42.231±0.455
0.45	100	27	0.783	0.138±0.184	0.264±0.134	0.231±0.131	0.000±0.000	0.000±0.000	66.587±0.508
0.45	250	16	1.034	0.135±0.170	0.274±0.085	0.236±0.084	0.000±0.000	0.000±0.000	47.834±0.535
0.45	500	14	1.116	0.138±0.167	0.278±0.060	0.250±0.059	0.000±0.000	0.000±0.000	44.162±0.520
0.45	1000	13	1.163	0.138±0.166	0.279±0.043	0.264±0.042	0.000±0.000	0.000±0.000	45.282±0.527
0.50	100	24	0.833	0.086±0.185	0.222±0.138	0.191±0.140	0.000±0.000	0.000±0.000	67.584±0.464
0.50	250	16	1.040	0.090±0.174	0.233±0.087	0.198±0.090	0.000±0.000	0.000±0.000	51.179±0.504
0.50	500	14	1.124	0.092±0.173	0.238±0.062	0.213±0.064	0.000±0.000	0.000±0.000	45.933±0.502
0.50	1000	13	1.165	0.093±0.173	0.240±0.044	0.227±0.045	0.000±0.000	0.000±0.000	47.256±0.429
0.55	100	24	0.838	0.044±0.181	0.180±0.141	0.148±0.151	0.000±0.000	0.000±0.000	69.801±0.509
0.55	250	17	1.020	0.054±0.171	0.193±0.089	0.154±0.097	0.000±0.000	0.000±0.000	53.028±0.498
0.55	500	14	1.117	0.054±0.171	0.198±0.063	0.168±0.070	0.000±0.000	0.000±0.000	46.044±0.498
0.55	1000	14	1.117	0.053±0.172	0.200±0.044	0.182±0.049	0.000±0.000	0.000±0.000	45.929±0.467
0.60	100	18	0.981	0.026±0.161	0.154±0.090	0.104±0.107	0.000±0.000	0.000±0.000	55.853±0.547
0.60	250	18	0.981	0.026±0.161	0.154±0.090	0.104±0.107	0.000±0.000	0.000±0.000	55.853±0.547
0.60	500	15	1.090	0.025±0.162	0.159±0.063	0.116±0.077	0.000±0.000	0.000±0.000	45.207±0.521
0.60	1000	13	1.158	0.023±0.164	0.162±0.044	0.133±0.054	0.000±0.000	0.000±0.000	43.311±0.503

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

Çizelge 4.4'te bazı eşik olasılıklarında tahmin modellerinden hesaplanan Net Fayda değerlerine ait ortalamaların Çizelge 4.3'de yer alan ortalamalara göre birbirinden uzaklaştığı görülmektedir. Ayrıca bu çizelgede etki büyüklüklerinin 0.5'i aştığı ve 1'e yaklaştığı yani büyük seviyelere geldiği ve böylece daha küçük örnek büyüklüklerinde Net Faydalar arası farkın tespit edilebileceği söylenebilir. Etki büyüklüğü değeri 0.5 civarında iken DVM modelinden hesaplanan Net Faydaların dağılımının binormal dağılıma benzediği gerekçesi ile bu çizelgede yer alan KW testine ait sonuçlar incelendiğinde çizelgede yer alan neredeyse tüm etki büyüklüklerinde ve eşik olasılıklarında 100 simülasyonun tamamında modeller arasında Net Fayda bakımından ortalama fark yok sayısı yaklaşık 0 olarak hesaplanmıştır. z testi sonuçları incelendiğinde ise aynı eşik olasılıkla (örneğin 0.35) etki büyüklüğü örnek büyüklüğü ile birlikte artmakta dolayısı ile ortalama fark yok sayısı azalmaktadır. Etki büyüklüğünün tespiti için yine bu çizelgede gerekli bootstrap sayısı ortalama 20'dir ve KW testi 100 simülasyonda Net Fayda bakımından modeller arasında istatistik olarak anlamlı fark bulmuştur.

Çizelge 4.5. Senaryo 3 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.10	100	1054	0.122	0.616±0.094	0.627±0.086	0.625±0.087	99.980±0.141	100.000±0.000	94.242±0.267
0.10	250	342	0.215	0.613±0.070	0.627±0.060	0.625±0.060	92.880±2.560	98.750±1.009	85.923±0.347
0.10	500	108	0.384	0.609±0.056	0.628±0.042	0.626±0.043	18.940±4.010	57.590±4.588	75.109±0.420
0.10	1000	49	0.575	0.605±0.048	0.628±0.030	0.627±0.030	0.040±0.197	5.400±2.727	69.017±0.509
0.15	100	328	0.219	0.582±0.109	0.604±0.091	0.597±0.093	93.900±2.272	98.940±1.071	86.846±0.328
0.15	250	100	0.399	0.574±0.090	0.605±0.063	0.598±0.065	17.250±3.560	59.170±4.573	75.286±0.420
0.15	500	46	0.593	0.567±0.082	0.606±0.044	0.599±0.046	0.080±0.273	10.090±3.220	65.766±0.430
0.15	1000	29	0.755	0.560±0.080	0.606±0.032	0.603±0.032	0.000±0.000	1.220±0.970	60.820±0.458
0.20	100	150	0.325	0.542±0.129	0.579±0.096	0.569±0.099	49.830±5.362	84.750±3.534	80.568±0.376
0.20	250	55	0.542	0.530±0.115	0.581±0.067	0.570±0.069	0.720±0.753	20.790±4.130	68.236±0.397
0.20	500	33	0.710	0.520±0.112	0.581±0.047	0.572±0.049	0.000±0.000	3.460±1.904	59.493±0.525
0.20	1000	24	0.844	0.511±0.114	0.582±0.034	0.577±0.034	0.000±0.000	1.020±1.197	52.899±0.552
0.25	100	46	0.594	0.428±0.191	0.520±0.107	0.494±0.110	0.070±0.293	8.130±2.820	65.778±0.493
0.25	250	40	0.638	0.481±0.143	0.553±0.071	0.536±0.073	0.050±0.219	6.180±2.294	62.280±0.498
0.25	500	27	0.789	0.470±0.142	0.554±0.050	0.540±0.051	0.000±0.000	1.000±0.932	53.424±0.508
0.25	1000	21	0.891	0.461±0.145	0.555±0.035	0.547±0.036	0.000±0.000	0.540±0.717	49.215±0.469
0.30	100	45	0.598	0.427±0.192	0.520±0.107	0.494±0.110	0.100±0.302	8.190±2.513	65.682±0.534
0.30	250	29	0.761	0.418±0.180	0.523±0.075	0.495±0.076	0.000±0.000	0.210±0.478	54.588±0.477
0.30	500	22	0.869	0.412±0.176	0.525±0.053	0.501±0.054	0.000±0.000	0.000±0.000	46.722±0.503
0.30	1000	19	0.940	0.407±0.175	0.526±0.037	0.512±0.038	0.000±0.000	0.050±0.219	46.189±0.500

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

Çizelge 4.5’de ANOVA testi, etki büyüklüğü önemsiz seviyede ve eşik olasılık 0.10 iken 100 simülasyonun 100’ünde tüm gözlem sayılarında modeller arasında Net Fayda ortalamaları bakımından istatistik olarak fark hesaplamazken, z testi örnek büyüklüğü arttıkça ortalama fark yok sayısında azalış göstermektedir. Etki büyüklüğünün yine önemsiz sayılabileceği seviyelerde eşik olasılık 0.15, $n=100$ iken ANOVA testinde modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı yaklaşık 94 iken z testinde bu sayı yaklaşık 87 olarak hesaplanmıştır. Etki büyüklüğünün yaklaşık orta seviyelerine gelindiğinde, eşik olasılık 0.15, $n=250$ iken ANOVA ve z testinde ortalama fark yok sayısının sırası ile yaklaşık 17 ve 75 olduğu yani etki büyüklüğü arttıkça ortalama fark yok sayısında bir düşüş olduğu gözlemlenmiştir. Aynı eşik olasılıkta, etki büyüklüğü orta seviyelere geldiğinde, $n=500$ iken ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı 46 olarak hesaplanmıştır ve bu senaryoda ANOVA testi 100 simülasyonun neredeyse tamamında modeller arasında Net Fayda bakımından fark bulurken z testi ise 66 tanesinde fark bulmamıştır. Yine aynı eşik olasılıkta, $n=1000$ ve etki büyüklüğü büyük seviyelere çıktığında DVM modeli binormal dağılıma yaklaştığından KW testinin de sonuçları göz önüne alındığında ANOVA 100 simülasyonun yaklaşık 100’ünde ve KW 100 simülasyonun yaklaşık 99’unda ve z testi ise yaklaşık 39 tanesinde fark bulmuştur. Etki büyüklüğü önemsiz seviyelerde, eşik olasılık 0.20, $n=100$ iken ANOVA testi modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı 50 iken z testinde bu sayı yaklaşık 81 olarak hesaplanmıştır. Etki büyüklüğü orta seviyelere geldiğinde, eşik olasılık 0.20, $n=250$ iken ANOVA ve z testinde ortalama fark yok sayıları sırası ile yaklaşık 1 ve 68 olarak hesaplanmıştır. Aynı eşik olasılıkta, etki büyüklüğünün çok büyük olduğu seviyelere gelindiğinde ise $n=500$, $n=1000$ iken ANOVA ve KW testi 100 simülasyonda farkın neredeyse tamamını yakalamasına rağmen z testi için ortalama fark yok sayısı sırası ile yaklaşık 59, 53 olarak hesaplanmıştır. Etki büyüklüğü orta seviyelerinde, eşik olasılık 0.25, $n=100$ iken ANOVA testinde modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı 0 iken z testinde bu sayı yaklaşık 66 olarak hesaplanmıştır. Etki büyüklüğü orta seviyelerde, eşik olasılık 0.25, $n=250$ iken ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı 40 olarak hesaplanmıştır ve 100 simülasyonda ANOVA ve z testinde ortalama fark yok sayıları sırası ile 0 ve 62 olarak hesaplanmıştır. Aynı eşik olasılıkta, $n=500$, $n=1000$ iken etki büyüklüğünün büyük olduğu seviyelere gelindiğinde ise ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı sırası ile 27 ve 21’dir ve ANOVA

ve KW testi 100 simülasyonda farkın neredeyse tamamını yakalamasına rağmen z testi için ortalama fark yok sayısı sırası ile yaklaşık 53 ve 49 olarak hesaplanmıştır. Etki büyüklüğünün orta seviyelerinde eşik olasılık 0.30, n=100 iken ANOVA testinde modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı 0 iken KW testinde 8, z testinde ise bu sayı yaklaşık 66 olarak hesaplanmıştır. Aynı eşik olasılıkta, n=250, n=500 ve n=1000 iken etki büyüklüğünün büyük olduğu seviyelere gelindiğinde ise ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı 29, 22, 19 iken ANOVA testi ve Kruskal Wallis testi 100 simülasyonda farkın neredeyse tamamını yakalamasına rağmen z testi için ortalama fark yok sayısı sırası ile yaklaşık 55, 47 ve 46 olarak hesaplanmıştır.

Çizelge 4.6. Senaryo 3 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.35	100	19	0.941	0.362±0.215	0.487±0.113	0.453±0.114	0.000±0.000	0.350±0.592	59.400±0.459
0.35	250	22	0.872	0.355±0.206	0.491±0.079	0.456±0.079	0.000±0.000	0.000±0.000	49.599±0.433
0.35	500	19	0.944	0.353±0.202	0.493±0.056	0.463±0.056	0.000±0.000	0.000±0.000	44.668±0.469
0.35	1000	18	0.980	0.354±0.198	0.494±0.040	0.475±0.040	0.000±0.000	0.000±0.000	44.235±0.486
0.40	100	27	0.780	0.310±0.224	0.450±0.119	0.412±0.119	0.000±0.000	0.020±0.141	58.478±0.505
0.40	250	21	0.896	0.309±0.215	0.455±0.083	0.417±0.083	0.000±0.000	0.000±0.000	49.644±0.466
0.40	500	19	0.959	0.310±0.210	0.458±0.059	0.426±0.059	0.000±0.000	0.000±0.000	45.855±0.468
0.40	1000	18	0.980	0.312±0.208	0.459±0.042	0.439±0.042	0.000±0.000	0.000±0.000	45.407±0.459
0.45	100	26	0.794	0.267±0.224	0.411±0.125	0.371±0.125	0.000±0.000	0.020±0.141	58.765±0.445
0.45	250	21	0.906	0.269±0.214	0.417±0.087	0.378±0.087	0.000±0.000	0.000±0.000	50.487±0.492
0.45	500	18	0.969	0.270±0.210	0.420±0.062	0.390±0.062	0.000±0.000	0.000±0.000	47.853±0.455
0.45	1000	18	0.989	0.272±0.210	0.422±0.044	0.404±0.044	0.000±0.000	0.000±0.000	47.917±0.534
0.50	100	27	0.780	0.226±0.218	0.368±0.137	0.330±0.139	0.000±0.000	0.010±0.100	61.587±0.446
0.50	250	20	0.926	0.229±0.207	0.377±0.091	0.340±0.093	0.000±0.000	0.000±0.000	52.637±0.493
0.50	500	18	0.978	0.232±0.204	0.380±0.065	0.353±0.066	0.000±0.000	0.000±0.000	49.565±0.522
0.50	1000	17	0.994	0.235±0.204	0.382±0.046	0.365±0.047	0.000±0.000	0.000±0.000	49.801±0.519
0.55	100	26	0.793	0.186±0.206	0.325±0.138	0.286±0.144	0.000±0.000	0.000±0.000	64.174±0.512
0.55	250	20	0.928	0.192±0.194	0.334±0.096	0.296±0.101	0.000±0.000	0.000±0.000	53.551±0.517
0.55	500	18	0.986	0.196±0.192	0.338±0.068	0.307±0.072	0.000±0.000	0.000±0.000	49.502±0.503
0.55	1000	17	1.003	0.199±0.193	0.340±0.048	0.318±0.052	0.000±0.000	0.000±0.000	48.352±0.512
0.60	100	28	0.764	0.142±0.199	0.277±0.151	0.234±0.164	0.000±0.000	0.000±0.000	68.385±0.401
0.60	250	20	0.920	0.154±0.180	0.288±0.100	0.244±0.110	0.000±0.000	0.000±0.000	55.696±0.537
0.60	500	17	0.991	0.160±0.176	0.293±0.071	0.253±0.080	0.000±0.000	0.000±0.000	48.199±0.518
0.60	1000	17	1.007	0.165±0.177	0.296±0.050	0.262±0.057	0.000±0.000	0.000±0.000	44.830±0.583

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

Çizelge 4.6’da bazı eşik olasılıklarında tahmin modellerinden hesaplanan Net Fayda değerlerine ait ortalamaların Çizelge 4.5’de yer alan ortalamalara göre birbirinden uzaklaştığı görülmektedir. Ayrıca bu çizelgede etki büyüklüklerinin 0.7’yi aştığı ve 1’e yaklaştığı yani çok önemli seviyelere geldiği ve böylece daha küçük örnek büyüklüklerinde Net Faydalar arası farkın tespit edilebileceği söylenebilir. Etki büyüklüğü değeri 0.7 civarında iken DVM modelinden hesaplanan Net Faydaların dağılımının binormal dağılıma benzediği gerekçesi ile bu çizelgede yer alan KW testine ait sonuçlar incelendiğinde çizelgede yer alan neredeyse tüm etki büyüklüklerinde ve eşik olasılıklarında 100 simülasyonun tamamında modeller arasında Net Fayda bakımından ortalama fark yok sayısı 0 olarak hesaplanmıştır. z testi sonuçları incelendiğinde ise aynı eşik olasılıklarda (örneğin 0.35) etki büyüklüğü örnek büyüklüğü ile birlikte artmakta dolayısı ile ortalama fark yok sayısı azalmaktadır. Etki büyüklüğünün tespiti için yine bu çizelgede gerekli bootstrap sayısı ortalama 20’dir ve KW testi 100 simülasyonda Net Fayda bakımından modeller arasında istatistik olarak anlamlı fark bulmuştur.

Çizelge 4.7. Senaryo 4 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.10	100	310	0.225	0.601±0.091	0.620±0.077	0.617±0.078	93.063±2.896	98.975±0.968	88.054±0.344
0.10	250	133	0.345	0.597±0.078	0.621±0.060	0.618±0.060	30.390±4.422	69.010±4.616	78.813±0.420
0.10	500	49	0.572	0.589±0.067	0.621±0.042	0.619±0.043	0.050±0.219	4.000±2.202	66.634±0.461
0.10	1000	29	0.760	0.584±0.062	0.621±0.030	0.620±0.030	0.000±0.000	0.230±0.489	62.898±0.516
0.15	100	111	0.379	0.561±0.112	0.598±0.081	0.590±0.083	24.588±4.608	63.788±5.315	77.266±0.397
0.15	250	66	0.493	0.552±0.104	0.598±0.063	0.590±0.064	0.410±0.588	11.690±3.253	67.697±0.445
0.15	500	29	0.761	0.540±0.100	0.599±0.045	0.593±0.045	0.000±0.000	0.740±0.812	58.275±0.516
0.15	1000	21	0.905	0.532±0.100	0.599±0.031	0.596±0.032	0.000±0.000	0.220±0.462	52.202±0.522
0.20	100	64	0.501	0.515±0.136	0.572±0.086	0.560±0.088	2.050±1.566	24.700±4.604	70.097±0.469
0.20	250	35	0.680	0.503±0.132	0.574±0.066	0.562±0.068	0.000±0.000	2.650±1.452	60.992±0.484
0.20	500	23	0.852	0.489±0.133	0.574±0.047	0.565±0.048	0.000±0.000	0.370±0.597	52.333±0.500
0.20	1000	19	0.953	0.482±0.134	0.575±0.033	0.570±0.034	0.000±0.000	0.250±0.435	47.284±0.550
0.25	100	43	0.615	0.464±0.163	0.545±0.090	0.526±0.093	0.050±0.219	6.075±2.464	63.499±0.502
0.25	250	28	0.765	0.453±0.159	0.547±0.070	0.529±0.072	0.000±0.000	0.750±0.833	56.183±0.488
0.25	500	21	0.906	0.440±0.161	0.548±0.050	0.533±0.051	0.000±0.000	0.060±0.239	48.976±0.516
0.25	1000	18	0.967	0.434±0.163	0.548±0.035	0.540±0.036	0.000±0.000	0.160±0.368	47.141±0.547
0.30	100	42	0.624	0.411±0.200	0.513±0.116	0.487±0.119	0.240±0.460	7.827±3.029	58.816±0.472
0.30	250	23	0.847	0.397±0.188	0.518±0.074	0.490±0.075	0.000±0.000	0.040±0.197	51.725±0.501
0.30	500	19	0.937	0.391±0.186	0.519±0.052	0.496±0.053	0.000±0.000	0.000±0.000	45.970±0.448
0.30	1000	18	0.984	0.388±0.186	0.520±0.037	0.505±0.038	0.000±0.000	0.010±0.100	45.112±0.454

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 4.7’de ANOVA testi, etki büyüklüğü küçük sayılacak seviyede, eşik olasılık 0.10, $n=100$ iken 100 simülasyonda Net Fayda ortalamaları bakımından istatistik olarak ortalama fark yok sayısı yaklaşık 93 iken z testinde bu sayı 88’dir. Etki büyüklüğünün yine küçük seviyelerde eşik olasılık 0.10, $n=250$ iken ANOVA testinde modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı yaklaşık 30 iken z testinde bu sayı yaklaşık 79 olarak hesaplanmıştır. Yine aynı eşik olasılıkta, $n=1000$ iken etki büyüklüğünün büyük seviyelerine gelindiğinde, DVM modeli binormal dağılıma yaklaştığından KW testinin de sonuçları göz önüne alındığında ANOVA ve KW 100 simülasyonun yaklaşık 100’ünde ve z testi ise yaklaşık 37 tanesinde fark bulmuştur dolayısı ile etki büyüklüğü arttıkça ortalama fark yok sayısında bir düşüş olduğu gözlemlenmiştir. Etki büyüklüğünün yaklaşık orta seviyelerinde, eşik olasılık 0.15 ve $n=100$ iken ANOVA testi modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı yaklaşık 25 iken z testinde bu sayı yaklaşık 78 olarak hesaplanmıştır. Etki büyüklüğünün yine yaklaşık orta seviyelerinde, eşik olasılık 0.15 ve $n=250$ iken ANOVA testi modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı 0 iken z testinde bu sayı yaklaşık 68 hesaplanmıştır. Etki büyüklüğünün büyük seviyelerinde, eşik olasılık 0.15 ve $n=500$ ve $n=1000$ iken ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı 29 ve 21 olarak hesaplanmıştır ve bu senaryolarda ANOVA ve KW testi 100 simülasyonun tamamında modeller arasında Net Fayda bakımından fark bulurken z testi ise sırası ile yaklaşık 58 ve 52 tanesinde fark bulmamıştır. Etki büyüklüğü orta seviyelerde, eşik olasılık 0.20, $n=100$ iken ANOVA testi modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı 2 iken z testinde bu sayı yaklaşık 70 olarak hesaplanmıştır. Etki büyüklüğü büyük seviyelere geldiğinde, eşik olasılık 0.20, $n=250$, 500, 1000 iken ANOVA ve KW testinde ortalama fark yok sayıları sırası ile 0, 0, 0 ve 3, 0, 0 iken z testinde bu sayı yaklaşık 61, 52 ve 47 olarak hesaplanmıştır. Etki büyüklüğü orta seviyelerde, eşik olasılık 0.25, $n=100$ iken ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı 43 olarak hesaplanmıştır ve 100 simülasyonda ANOVA ve z testinde ortalama fark yok sayıları sırası ile 0 ve 63 olarak hesaplanmıştır. Aynı eşik olasılıkta, $n=250$, $n=500$ ve $n=1000$ iken etki büyüklüğünün büyük olduğu seviyelere gelindiğinde ise ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı sırası ile 28, 21 ve 18’dir ve ANOVA ve KW testi 100 simülasyonda farkın neredeyse tamamını yakalamasına rağmen z testi için ortalama fark yok sayısı sırası ile yaklaşık 56, 49 ve 47 olarak

hesaplanmıştır. Etki büyüklüğünün orta seviyelerinde, eşik olasılık 0.30, n=100 iken ANOVA testinde modeller arasında Net Fayda ortalamaları bakımından ortalama fark yok sayısı yaklaşık 0 iken KW testinde 8, z testinde ise bu sayı yaklaşık olarak 59 olarak hesaplanmıştır. Aynı eşik olasılıktaki, n=250, n=500 ve n=1000 iken etki büyüklüğünün çok önemli olduğu seviyelere gelindiğinde ise ANOVA testinin bu etkiyi tespit edebilmesi için gereken bootstrap sayısı 23, 19, 18 iken ANOVA testi ve Kruskal Wallis testi 100 simülasyonda farkın neredeyse tamamını yakalamasına rağmen z testi için ortalama fark yok sayısı sırası ile yaklaşık 52, 46 ve 45 olarak hesaplanmıştır.

Çizelge 4.8. Senaryo 4 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.35	100	24	0.831	0.346±0.213	0.484±0.099	0.452±0.100	0.000±0.000	0.025±0.157	54.261±0.524
0.35	250	21	0.905	0.345±0.208	0.487±0.077	0.455±0.077	0.000±0.000	0.010±0.100	49.685±0.549
0.35	500	18	0.964	0.345±0.204	0.489±0.055	0.461±0.055	0.000±0.000	0.000±0.000	46.257±0.504
0.35	1000	18	0.978	0.346±0.203	0.489±0.039	0.471±0.039	0.000±0.000	0.000±0.000	44.845±0.578
0.40	100	22	0.868	0.300±0.223	0.451±0.104	0.417±0.104	0.000±0.000	0.025±0.157	54.441±0.484
0.40	250	20	0.924	0.302±0.218	0.454±0.081	0.419±0.080	0.000±0.000	0.000±0.000	50.394±0.458
0.40	500	18	0.966	0.304±0.215	0.456±0.057	0.427±0.057	0.000±0.000	0.000±0.000	47.554±0.471
0.40	1000	18	0.987	0.305±0.214	0.457±0.040	0.439±0.040	0.000±0.000	0.000±0.000	47.190±0.466
0.45	100	22	0.886	0.259±0.224	0.415±0.109	0.379±0.109	0.000±0.000	0.013±0.112	54.646±0.422
0.45	250	19	0.939	0.262±0.221	0.419±0.084	0.384±0.085	0.000±0.000	0.000±0.000	50.949±0.464
0.45	500	18	0.974	0.265±0.220	0.422±0.060	0.394±0.060	0.000±0.000	0.000±0.000	49.380±0.498
0.45	1000	18	0.989	0.267±0.219	0.423±0.042	0.406±0.042	0.000±0.000	0.000±0.000	50.125±0.480
0.50	100	25	0.823	0.130±0.197	0.271±0.141	0.239±0.143	0.000±0.000	0.000±0.000	65.602±0.435
0.50	250	19	0.938	0.226±0.218	0.382±0.088	0.348±0.091	0.000±0.000	0.020±0.141	52.774±0.489
0.50	500	18	0.976	0.229±0.219	0.386±0.062	0.360±0.064	0.000±0.000	0.000±0.000	51.095±0.472
0.50	1000	18	0.981	0.232±0.219	0.387±0.044	0.371±0.045	0.000±0.000	0.030±0.171	52.128±0.551
0.55	100	25	0.809	0.181±0.220	0.332±0.146	0.298±0.154	0.000±0.000	0.129±0.378	48.942±0.516
0.55	250	21	0.898	0.192±0.211	0.344±0.092	0.308±0.098	0.000±0.000	0.000±0.000	53.180±0.528
0.55	500	18	0.965	0.196±0.213	0.348±0.065	0.318±0.070	0.000±0.000	0.000±0.000	51.225±0.462
0.55	1000	27	0.789	0.417±0.283	0.576±0.036	0.560±0.040	0.000±0.000	0.540±0.744	59.127±0.451
0.60	100	27	0.786	0.145±0.213	0.291±0.154	0.250±0.168	0.000±0.000	0.010±0.100	68.094±0.462
0.60	250	20	0.912	0.160±0.200	0.303±0.096	0.260±0.108	0.000±0.000	0.000±0.000	54.495±0.474
0.60	500	19	0.956	0.164±0.202	0.308±0.068	0.269±0.078	0.000±0.000	0.000±0.000	50.156±0.473
0.60	1000	18	0.976	0.166±0.203	0.310±0.048	0.278±0.056	0.000±0.000	0.000±0.000	47.106±0.573

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

Çizelge 4.8’de bazı eşik olasılıklarında tahmin modellerinden hesaplanan Net Fayda değerlerine ait ortalamaların çizelge 4.7’de yer alan ortalamalara göre birbirinden uzaklaştığı görülmektedir. Ayrıca bu çizelgede etki büyüklüklerinin 0.7’yi aştığı ve 1’e yaklaştığı yani çok önemli seviyelere geldiği ve böylece daha küçük örnek büyüklüklerinde Net Faydalar arası farkın tespit edilebileceği söylenebilir. Etki büyüklüğü değeri 0.7 civarında iken DVM modelinden hesaplanan Net Faydaların dağılımının binormal dağılıma benzediği gerekçesi ile bu çizelgede yer alan KW testine ait sonuçlar incelendiğinde çizelgede yer alan neredeyse tüm etki büyüklüklerinde ve eşik olasılıklarında 100 simülasyonun tamamında modeller arasında Net Fayda bakımından ortalama fark yok sayısı yaklaşık 0 olarak hesaplanmıştır. z testi sonuçları incelendiğinde ise aynı eşik olasılıkla (örneğin 0.35) etki büyüklüğü örnek büyüklüğü ile birlikte artmakta dolayısı ile ortalama fark yok sayısı azalmaktadır. Etki büyüklüğünün tespiti için yine bu çizelgede gerekli bootstrap sayısı ortalama 20’dir ve KW testi 100 simülasyonda Net Fayda bakımından modeller arasında istatistik olarak anlamlı fark bulmuştur.

4.2. Gerçek Veriler için Elde Edilen Bulgular

4.2.1. Pima Veri Setinden Elde Edilen Bulgular

Çizelge 4.9. Pima veri'si için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA**	KW**	z**
0.10	100	49	0.572	0.265±0.047	0.269±0.029	0.288±0.032	0	0	46
0.10	250	8	1.519	0.261±0.031	0.269±0.017	0.298±0.015	0	0	38
0.10	500	4	2.688	0.272±0.022	0.262±0.016	0.296±0.008	0	0	21
0.10	768	3	4.566	0.270±0.015	0.247±0.014	0.295±0.005	0	0	0
0.15	100	13	1.194	0.223±0.053	0.218±0.040	0.264±0.037	0	0	51
0.15	250	4	2.381	0.228±0.035	0.205±0.031	0.27±0.023	0	0	26
0.15	500	3	4.909	0.238±0.023	0.193±0.019	0.271±0.012	0	0	2
0.15	768	2	6.738	0.237±0.018	0.187±0.015	0.268±0.008	0	0	0
0.20	100	11	1.308	0.170±0.056	0.169±0.040	0.224±0.044	0	0	67
0.20	250	3	3.431	0.205±0.038	0.144±0.033	0.243±0.024	0	0	26
0.20	500	2	5.772	0.211±0.024	0.122±0.025	0.241±0.015	0	0	0
0.20	768	2	8.763	0.210±0.019	0.118±0.015	0.241±0.013	0	0	0
0.25	100	7	1.681	0.167±0.054	0.103±0.057	0.198±0.056	0	0	64
0.25	250	3	4.189	0.185±0.039	0.063±0.043	0.215±0.028	0	0	9
0.25	500	2	7.825	0.190±0.025	0.051±0.024	0.217±0.018	0	0	0
0.25	768	2	11.103	0.195±0.023	0.037±0.017	0.215±0.015	0	0	0
0.30	100	4	2.666	0.140±0.067	0.024±0.056	0.176±0.058	0	0	43
0.30	250	3	5.591	0.156±0.037	-0.004±0.037	0.192±0.033	0	0	4
0.30	500	2	9.182	0.169±0.028	-0.030±0.025	0.195±0.024	0	0	0
0.30	768	2	13.872	0.164±0.023	-0.038±0.017	0.191±0.016	0	0	0

pt: Eşik olasılık, n: Örnek sayısı, R: Bootstrap sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: 100'lük Bootstrap sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

Çizelge 4.9.'da gerçek veri setinde, bazı eşik olasılıklarında tahmin modellerinden hesaplanan Net Fayda değerlerine ait ortalamalar ve karşılaştırma sonuçları yer almaktadır. Burada Net Faydaların sayısı 100 iken z testi ile yapılan analiz sayısı da 100'dür. ANOVA ve KW testleri ile yapılan analiz sayısı ise 1'dir. Etki büyüklüklerinin minimum değeri yaklaşık 0.6 iken maksimum değeri yaklaşık 14 olarak hesaplanmış olup genel olarak 1'in üzerine çıktığı yani çok büyük seviyelerde olduğu ve küçük örnek büyüklüklerinde Net Faydalar arası farkın kolaylıkla tespit edilebileceği söylenebilmektedir. Etki büyüklüğü değeri çok büyük seviyelerde iken ANOVA ve KW testine ait sonuçlar incelendiğinde çizelgede yer alan neredeyse tüm etki

büyükliklerinde ve eşik olasılıklarında modeller arasında Net Fayda bakımından ortalama fark yok sayısı 0 olarak hesaplanmıştır. z testi sonuçları incelendiğinde ise aynı eşik olasılığında etki büyüklüğü örnek büyüklüğü ile birlikte artmakta dolayısı ile ortalama fark yok sayısı azalmaktadır. Etki büyüklüğünün tespiti için yine bu çizelgede gerekli örnek sayısı ortalama 6 olup ANOVA ve KW testi 100 bootstrap'ta Net Fayda bakımından modeller arasında istatistik olarak anlamlı fark bulmuştur.

Çizelge 4.10. Pima veri'si için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NE _{DVM} (ort±ss)	NE _{LR} (ort±ss)	NE _{KA} (ort±ss)	ANOVA**	KW**	z**
0.35	100	3	3.911	0.137±0.063	-0.053±0.056	0.166±0.056	0	0	26
0.35	250	2	7.386	0.142±0.042	-0.092±0.037	0.174±0.035	0	0	0
0.35	500	2	11.589	0.144±0.026	-0.114±0.024	0.170±0.025	0	0	0
0.35	768	2	15.886	0.146±0.021	-0.125±0.019	0.169±0.018	0	0	0
0.40	100	3	4.823	0.102±0.067	-0.149±0.060	0.138±0.059	0	0	17
0.40	250	2	9.941	0.123±0.042	-0.189±0.034	0.149±0.034	0	0	0
0.40	500	2	16.227	0.129±0.031	-0.205±0.022	0.152±0.022	0	0	0
0.40	768	2	18.508	0.132±0.023	-0.220±0.018	0.152±0.022	0	0	0
0.45	100	3	5.363	0.081±0.076	-0.248±0.061	0.107±0.071	0	0	15
0.45	250	2	10.647	0.102±0.049	-0.293±0.036	0.135±0.044	0	0	0
0.45	500	2	15.031	0.106±0.031	-0.308±0.028	0.136±0.031	0	0	0
0.45	768	2	20.272	0.115±0.025	-0.313±0.022	0.133±0.022	0	0	0
0.50	100	2	6.464	0.040±0.096	-0.380±0.069	0.096±0.078	0	0	6
0.50	250	2	11.814	0.066±0.048	-0.395±0.043	0.113±0.043	0	0	0
0.50	500	2	17.400	0.083±0.033	-0.409±0.030	0.113±0.030	0	0	0
0.50	768	2	21.560	0.089±0.029	-0.422±0.025	0.117±0.025	0	0	0
0.55	100	2	7.555	0.040±0.091	-0.518±0.080	0.094±0.082	0	0	1
0.55	250	2	12.253	0.049±0.065	-0.527±0.052	0.098±0.050	0	0	0
0.55	500	2	22.950	0.063±0.042	-0.540±0.027	0.103±0.029	0	0	0
0.55	768	2	25.979	0.064±0.031	-0.548±0.026	0.102±0.024	0	0	0
0.60	100	2	8.634	-0.003±0.107	-0.662±0.088	0.060±0.079	0	0	1
0.60	250	2	12.789	0.017±0.066	-0.659±0.066	0.079±0.048	0	0	0
0.60	500	2	20.237	0.029±0.048	-0.674±0.043	0.093±0.032	0	0	0
0.60	768	2	26.764	0.037±0.031	-0.687±0.031	0.091±0.027	0	0	0

pt: Eşik olasılık, n:Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: 100'lük Bootstrap sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

Çizelge 4.10’da bazı eşik olasılıklarında tahmin modellerinden hesaplanan Net Fayda değerlerine ait ortalamaların etki büyüklüklerinin çok büyük olduğu görülmektedir. Burada Net Faydaların sayısı 100 iken z testi ile yapılan analiz sayısı da 100’dür. ANOVA ve KW testleri ile yapılan analiz sayısı ise 1’dir. Etki büyüklüğü değerlerinin 1’i aştığı yani çok önemli seviyelere geldiği ve böylece daha küçük örnek büyüklüklerinde Net Faydalar arası farkın tespit edilebileceği söylenebilir. Bu çizelgede tüm eşik olasılıklarda DVM modeline ait Net Faydalar sıfırın altında hesaplanmıştır. ANOVA ve KW testine ait sonuçlar incelendiğinde çizelgede yer alan neredeyse tüm etki büyüklüklerinde ve eşik olasılıklarında 100 bootstrapın tamamında modeller arasında Net Fayda bakımından ortalama fark yok sayısı 0 olarak hesaplanmıştır. z testi sonuçları incelendiğinde ise aynı eşik olasılıklarda, etki büyüklüğü ve örnek büyüklüğü sayısı arttıkça ortalama fark yok sayısı azalmaktadır. Etki büyüklüğünün tespiti için yine bu çizelgede gerekli bootstrap sayısı ortalama 2’dir ve ANOVA ile KW testi 100 adet Net Fayda için modeller arasında istatistik olarak anlamlı fark bulmuştur.

4.2.2. Prostat Veri Seti ile Elde Edilen Bulgular

Çizelge 4.11. Prostat veri’si için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±sd)	KW** (ort±sd)	z** (ort±sd)
0.10	100	5	2.364	0.587±0.035	0.415±0.104	0.595±0.028	0	0	12
0.15	100	4	3.064	0.570±0.044	0.322±0.112	0.575±0.033	0	0	5
0.20	100	3	4.483	0.557±0.054	0.220±0.101	0.561±0.037	0	0	1
0.25	100	3	5.639	0.542±0.055	0.145±0.083	0.542±0.042	0	0	0
0.30	100	2	7.969	0.520±0.055	0.076±0.066	0.523±0.044	0	0	0
0.35	100	2	9.452	0.520±0.050	-0.002±0.060	0.509±0.037	0	0	0
0.40	100	2	10.863	0.501±0.053	-0.064±0.051	0.492±0.046	0	0	0
0.45	100	2	11.372	0.491±0.056	-0.141±0.056	0.479±0.053	0	0	0
0.50	100	2	11.144	0.457±0.059	-0.184±0.056	0.441±0.069	0	0	0
0.55	100	2	10.283	0.438±0.065	-0.246±0.068	0.437±0.081	0	0	0
0.60	100	2	8.348	0.409±0.077	-0.291±0.088	0.415±0.081	0	0	0

pt: Eşik olasılık, n:Örnek sayısı, R: Tekrar sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan tekrar sayısı

**: 100’lük Bootstrap sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 4.11’de bazı eşik olasılıklarında tahmin modellerinden hesaplanan Net Fayda değerlerine ait ortalamaların etki büyüklüklerinin çok büyük olduğu görülmektedir. Burada Net Faydaların sayısı 100 iken z testi ile yapılan analiz sayısı da 100’dür. ANOVA ve KW testleri ile yapılan analiz sayısı ise 1’dir. Etki büyüklüğü değerlerinin 1’i aştığı yani çok büyük seviyelere geldiği ve böylece daha küçük örnek büyüklüklerinde Net Faydalar arası farkın tespit edilebileceği söylenebilir.

ANOVA ve KW testine ait sonuçlar incelendiğinde çizelgede yer alan neredeyse tüm etki büyüklüklerinde ve eşik olasılıklarında 100 tekrarın tamamında modeller arasında Net Fayda bakımından ortalama fark yok sayısı 0 olarak hesaplanmıştır. z testi sonuçları incelendiğinde ise etki büyüklüğü eşik olasılık ile artmakta olup, 0.25 ve üstü eşik olasılıklarda modeller arasındaki farkı istatistik olarak anlamlı bulmuştur. Etki büyüklüğünün gerektirdiği bootstrap sayısı ortalama 3’dür ve ANOVA ile KW testi 100 adet Net Fayda için modeller arasında istatistik olarak anlamlı fark bulmuştur.

5. TARTIŞMA

Tanı testlerinin doğrulukları ile hastanın tedaviden görecekları fayda arasında dolaylı bir ilişki vardır. Yüksek doğruluklu bir test sonucu hasta hakkında doğru karar vermeyi ve hastanın alacağı hizmetten fayda görmesini sağlayabilirken, doğruluğu düşük olan bir testte ise hastanın zarar görme olasılığı artabilecektir. Dolayısıyla tanı testinin doğruluk düzeyi "fayda-zarar (yan etki) dengesi" olarak bilinen bir problem oluşturmaktadır. Örneğin bir kanser tanı testi; hastalığı olan 30 bireyi doğru tespit edip biyopsi yapılmasını sağlarken, hastalığı olmayan 70 bireye gereksiz biyopsi yapılmasına yol açmaktadır. Yeni bir tanı testi ise gereksiz biyopsi sayısını 50'ye indirdiğini, ancak kanser vakalarından yalnızca 28 tanesini tespit ettiğini varsayalım bu durumda, gereksiz biyopsi sayısı 20 azalırken 2 kanser hastası da tespit edilememiş ve yeni testin kullanımının doğru bir karar olup olmadığı tartışma konusu olacaktır⁸⁹. Benzer durum, tahmin modelleri ve belirteçlerin değerlendirilmesinde de geçerlidir.

Tam da bu noktada eşik olasılık değeri kritik rol oynamaktadır. Çünkü eşik olasılığın değışmesi, yanlış pozitif ve negatiflerin oranının değışmesi anlamına gelmektedir. Ancak birçok tanı testi için, hastalık var ve hastalık yok sonuçlarını birbirinden ayıran klinik olarak en uygun eşik konusunda bir fikir birliği yoktur. Çünkü hastalığı kaçırma riskinin kabul edilebilir olduğu eşik değeri konusunda anlaşmak zordur. Dolayısı ile farklı klinik karara sebep olabilecek birden fazla muhtemel eşik değeri vardır ve bu da sıklıkla tanı belirsizliği ile sonuçlanabilmektedir. Yanlış negatif ve yanlış pozitif teşhislerin yaygınlığı ve göreceli önemiyle ilgili bağlamsal bilgileri içeren bir yöntem olan Karar Eğrisi Analizi, iki test arasındaki duyarlılık ve özgüllük farklarını göreceli klinik maliyetler ve hastalık yaygınlığına göre ağırlıklandırmaktadır⁹⁰. Karar eğrisi analizi rasyonel bir eşik hesaplamak için yeterli fikir birliği/ veya yeterli veri olmadığında eşik olasılığını uygun bir aralıkta değıştirmeye olanak tanımaktadır. Bu yöntem Net fayda hesabına dayanmakla beraber kolayca ifade edilebilen yorumu; uygun şekilde tedavi edilen hastaların oranındaki net artış ya da gereksiz yere tedavi edilen hastaların oranındaki net azalıştır⁹¹.

Tahmin modellerinin farklı eşik olasılıklarda Net Faydaları arasındaki farkın istatistiksel olarak anlamlı olup olmadığı ile ilgili literatürde yer alan çalışma sayısı az olmakla beraber tahmin modeli sayısının 2 olması ile sınırlıdır. Örneğin meme kanseri tanısı için kullanılan tanı testi veya tahmin modeli sayısı 2 olduğunda ve bunlardan hesaplanan 2 adet Net Faydaların istatistik olarak birbirinden farklı olup olmadığını hesaplamak için z testi kullanılabilir. Ancak kullanılan tanı testi veya tahmin modeli sayısı 2’den fazla ise Net Faydaları eş zamanlı karşılaştırmada z testinin kullanımı yapılan analiz sayısını karşılaştırma sayısı kadar artıracaktır. Bu nedenle bu çalışmada; bir hastalık tahmini için kullanılan 3 farklı makine öğrenmesi modeli oluşturulduğunda ve bu modellerden Net Fayda hesaplandığında modeller arasında istatistik olarak fark olup olmadığının tespiti için ANOVA ve KW testlerinin kullanılabilirliği incelenmiştir.

Son zamanlarda birkaç çalışma, Net Faydaların istatistik olarak karşılaştırması için p değerleri ve %95 güven aralığı hesaplamak için yöntemler önermişlerdir^{22,23} ancak Vickers ve ark. (2023) yayınladıkları çalışmada Net Faydalar bakımından istatistik karşılaştırma ve bir p değeri ile güven aralığının dikkate alınması yerine p değerleri veya belirsizlikten bağımsız olarak en yüksek Net Faydaya sahip modeli seçmenin gerekliliğini vurgulamışlardır⁹². Öte yandan özellikle de küçük eşik olasılıklarda (p_t) $Net\ Fayda = \frac{GerçekPozitifSayısı}{n} - \frac{YanlışPozitifSayısı}{n} \left(\frac{p_t}{1-p_t} \right)$ eşitliği ile modellerden hesaplanan Net Fayda değerleri neredeyse eşit hesaplanmakla beraber istatistik olarak bir fark bulmak zayıf bir ihtimalken eşik olasılık arttıkça Net Fayda değerleri arasındaki farkın da formül gereği artacağı dolayısı ile istatistik olarak da farklı hale gelmesi beklenmektedir. Bu çalışma sonuçları da Vickers ve ark.’nın iddiasını destekler nitelikte olup; özellikle de eşik olasılık arttıkça, Net Faydalar arasındaki farkın açıldığı dolayısı ile istatistik olarak da Net Faydalar arasında fark bulunacağı beklentisi olduğunu söylemek mümkündür.

Senaryo 1’in (veri üretim süreçlerinde; beta katsayılarının birbirinden farklı ve varyans kovaryans matrisinin diagonal olduğu) sonuçları incelendiğinde; küçük eşik olasılıklarda modellerden hesaplanan Net Fayda değerlerinin birbirine çok yakın (küçük etki büyüklüğü) neredeyse eşit olduğu görülmektedir. Örneğin Çizelge 4.1’de etki büyüklüğü önemsiz seviyede $d=0.06$, $p_t=0.10$, $n=250$ iken tüm modellerden Net Fayda değeri yaklaşık 0.52 olarak hesaplanmıştır. Aynı eşik olasılıklarda diğer senaryolarda da tüm modellerden Net Fayda değerinin yaklaşık 0.52 olarak hesaplandığı görülmektedir.

Dolayısı ile ANOVA ve KW test sonuçlarının modeller arasında 100 simülasyonda Net Fayda ortalaması bakımından fark bulmaması olağan bir sonuç iken 0.10 eşik olasılıkta $n=1000$ iken z testi Net Fayda ortalaması bakımından yaklaşık 80 simülasyonda fark bulmamıştır. Benzer şekilde, eşik olasılık 0.15, $n=100$ iken tüm modellerden Net Fayda değerinin neredeyse aynı olarak hesaplandığı (etki büyüklüğü önemsiz) ANOVA ve KW test sonuçlarının modeller arasında Net Fayda ortalaması bakımından fark bulmadığı z testinin ise yaklaşık 93 simülasyonda fark bulmadığı görülmektedir. Eşik olasılık 0.20, $n=100$ iken yine tüm modellerden Net Fayda değerinin neredeyse aynı olarak hesaplandığı (etki büyüklüğü önemsiz) ANOVA ve KW test sonuçlarının modeller arasında Net Fayda ortalaması bakımından fark bulmadığı z testinin ise yaklaşık 86 simülasyonda fark bulduğu görülmektedir. Hesaplanan Net Fayda ortalamaları bakımından modeller arasındaki farkın, etki büyüklüğü 1'e yaklaştıkça, açıldığı görülmektedir. Çizelge 4.2 incelendiğinde ise özellikle 0.35 eşik olasılıktan sonra tespit edilen yüksek etki büyüklüklerini 100 simülasyonda ANOVA ve KW testlerinin tespit gücünün z testine göre yüksek olduğu söylenebilir.

Senaryo 2'nin (veri üretim süreçlerinde; beta katsayılarının birbirinden farklı ve varyans kovaryans matrisinin diagonal olmadığı) sonuçları incelendiğinde; küçük eşik olasılıkta modellerden hesaplanan Net Fayda değerlerinin birbirine çok yakın neredeyse eşit olduğu (önemsiz etki büyüklüğü) görülmektedir. Örneğin Çizelge 4.3'de etki büyüklüğü önemsiz $d=0.13$, $pt=0.10$, $n=500$ iken tüm modellerden Net Fayda ortalama değeri yaklaşık 0.52 olarak hesaplanmıştır. Dolayısı ile ANOVA ve KW test sonuçlarının modeller arasında Net Fayda ortalaması bakımından fark bulmaması da olağan bir sonuç iken z testi Net Fayda ortalaması bakımından yaklaşık 85 simülasyonda fark bulmamıştır. Aynı eşik olasılıklar içerisinde etki büyüklüğünün seviyesi örnek büyüklüğü ile birlikte arttıkça, ANOVA, KW ve z testinde ortalama fark yok sayısında azalma olduğu görülmektedir. Dolayısı ile örnek büyüklüğünün tüm testler üzerinde etkisi olduğunu söylemek mümkündür. Çizelge 4.4 incelendiğinde ise modellerden hesaplanan Net Fayda ortalamaları birbirinden uzaklaşmıştır (yüksek etki büyüklüğü) ve örneğin etki büyüklüğü yaklaşık $d=1.16$, $pt=0.60$, $n=1000$ iken DVM modelinden hesaplanan Net Fayda ortalaması 0.02 Karar Ağacı modelinden hesaplanan Net Fayda ortalaması ise neredeyse 7 katı olup 0.13'dür. Dolayısı ile Net Fayda ortalamaları arasında istatistik olarak fark bulunmasının da olağan bir sonuç olduğunu söylemek mümkündür. Çizelge 4.4'te, Çizelge 4.2'ye benzer şekilde özellikle 0.35 eşik

olasılıktan sonra tespit edilen yüksek etki büyüklüklerinde 100 simülasyonda ANOVA ve KW testlerinin farkı tespit gücünün z testine göre yüksek olduğu söylenebilir.

Senaryo 3'e (veri üretim süreçlerinde; beta katsayılarının birbiri ile aynı ve varyans kovaryans matrisinin diagonal olduğu) gelindiğinde küçük eşik olasılıkta modellerden hesaplanan Net Fayda değerlerinin birbirine çok yakın neredeyse eşit (önemsiz etki büyüklüğü) olduğu görülmektedir. Örneğin, Çizelge 4.5'de; $d=0.12$, $pt=0.10$, $n=100$ iken tüm modellerden Net Fayda değeri yaklaşık 0.62 olarak hesaplanmıştır. Dolayısı ile ANOVA ve KW test sonuçlarının modeller arasında Net Fayda ortalaması bakımından fark bulmaması da olağan bir sonuç iken z testi Net Fayda ortalaması bakımından yaklaşık 94 simülasyonda fark bulmamıştır. Çizelge 4.3'e benzer şekilde aynı eşik olasılıklar içerisinde etki büyüklüğünün seviyesi örnek büyüklüğü ile birlikte arttıkça, ANOVA, KW ve z testinde ortalama fark yok sayısında azalma olduğu görülmektedir. Etki büyüklüğü seviyesi orta seviyeleri (0.5) geçtiğinde, istatistik testlerin modeller arasında Net Fayda bakımından fark bulma eğiliminde olduğu söylenebilir. Çizelge 4.6'nın, Çizelge 4.4 ve Çizelge 4.2'ye benzer şekilde özellikle 0.35 ve üzeri eşik olasılıklarda tespit edilen yüksek etki büyüklüklerinde 100 simülasyonda ANOVA ve KW testlerinin farkı tespit gücünün z testine göre yüksek olduğu söylenebilir.

Senaryo 4'e (veri üretim süreçlerinde; beta katsayılarının birbiri ile aynı ve varyans kovaryans matrisinin diagonal olmadığı) gelindiğinde etki büyüklüğü seviyelerinin genel olarak orta ve yüksek olduğu görülmektedir. Çizelge 4.1, Çizelge 4.3, Çizelge 4.5'e benzer şekilde aynı eşik olasılıklar içerisinde etki büyüklüğünün seviyesi örnek büyüklüğü ile birlikte arttıkça, ANOVA, KW ve z testinde ortalama fark yok sayısında azalma olduğu görülmektedir. Etki büyüklüğü seviyesi orta seviyeleri (0.5) geçtiğinde, istatistik testlerin modeller arasında Net Fayda bakımından fark bulma eğiliminde olduğu söylenebilir. Çizelge 4.8'in Çizelge 4.6, Çizelge 4.4 ve Çizelge 4.2'ye benzer şekilde özellikle 0.35 eşik olasılıktan sonra tespit edilen yüksek etki büyüklüklerinde 100 simülasyonda ANOVA ve KW testlerinin farkı tespit gücünün z testine göre yüksek olduğu söylenebilir.

Tüm senaryolarda benzer çıkarımlar yapılabildiğinden simülasyon senaryolarının, yapılan istatistik analizler üzerinde çok önemli farklılıklara neden olduğu düşünülmemektedir.

Gerçek veri sonuçları incelendiğinde ise hem Pima hem de Prostat verisi için etki büyüklükleri yüksek seviyelerde hesaplanmakla beraber tekrar sayısının da yeterli olduğu tüm senaryolarda ANOVA ve KW testleri modeller arasında Net Fayda ortalaması bakımından istatistik olarak fark bulmuştur. z testi'nde ise artan etki büyüklüğü ve örnek büyüklüğü ile modeller arasında Net Fayda ortalaması bakımından fark bulma eğilimi de artmıştır.



6. SONUÇLAR VE ÖNERİLER

Tahmin modellerinin geliştirilmesi ve değerlendirilmesinde dikkate alınması gereken adımlar arasında; araştırma sorusu ve veri incelemesi, açıklayıcı değişkenlerin belirlenmesi, model tahmini, model performans değerlendirmesi, iç geçerlik gibi bazı önemli unsurlar bulunmaktadır. Bir modelin geçerli olabilmesi için performans değerlendirmesinde 4 temel ölçü belirlenmiştir ve bu ölçüler; modelin sabit terimi, modelin kalibrasyon eğrisi, uyum (concordance) indeksleri ile modelin diskriminasyonu ve karar eğrisi analizi ile klinik yararlılık olarak sıralanmıştır⁹³.

Bu çalışmada da klinik yararlılık, gerçek veri setlerine ve farklı senaryolarda üretilen verilere uydurulan tahmin modellerinden elde edilen Net Fayda değerleri olarak hesaplanmış ve hesaplanan Net Fayda ortalamaları istatistik olarak karşılaştırılmıştır. Elde edilen bulgulara göre;

1. Küçük eşik olasılıklarda, Net Fayda ortalama değerleri birbirine yakın/benzer olduğu durumlarda, (önemsiz etki büyüklüklerinde) ANOVA ve KW testlerinin modeller arasında Net Fayda bakımından ortalama fark bulmama sayısı z testinden yüksektir.

2. Yüksek eşik olasılıklarda, tekrar sayısının yeterli ve Net Fayda ortalama değerleri birbirinden uzak olduğu durumlarda, (önemli etki büyüklüklerinde) ANOVA ve KW testleri modeller arasında Net Fayda bakımından ortalama fark bulmama sayısı z testinden düşüktür.

3. Herhangi bir eşik olasılık değerinde etki büyüklüğünün artması, bu etkiyi tespit etmek için gerekli örnek büyüklüğünü artırmaktadır. Bu da z testine göre ANOVA ve KW testleri ile daha az tekrarla (yeniden örnekleme) Net Faydalar arasındaki farkın bulunmasını sağlamaktadır.

4. Her ne kadar normal dağılımdan sapan veriler görülse de ANOVA testinin dayanıklı yapısı elde edilen senaryoların çoğunda KW ile benzer sonuçlara sahip olduğundan 2'den fazla Net Faydanın istatistik olarak eşzamanlı karşılaştırılmasında ANOVA testinin kullanılabilir olduğu görülmüştür.

5. Farklı senaryo ve veri setlerinde ANOVA testinin uygulanabilirliği ile ilgili araştırma yapılması önerilmektedir.

7. KAYNAKLAR

1. **Geisser S.** *Predictive Inference: An Introduction*. 1st ed., Chapman and Hall, **1993**.
2. **Kuhn M, Johnson K.** *Applied Predictive Modeling*. New York: Springer, **2013**.
3. **Steyerberg EW.** *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. Springer, **2009**.
4. **Collins GS, Reitsma JB, Altman DG, Moons KG.** Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) the TRIPOD statement. *Circulation*, **2015**; 131(2):211-219.
5. **Zhou ZR, Wang WW, Li Y, Jin KR, Wang XY, Wang ZW, Chen YS, Wang SJ, Hu J, Zhang HN, Huang P, Zhao GZ, Chen XX, Tian BL, Zhang TS.** In-depth mining of clinical data: the construction of clinical prediction model with R, *Annals of translational medicine*, **2019**; 7(23).
6. **Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW.** Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*, **2010**; 21(1):128-138.
7. **Binuya M, Engelhardt E, Schats W, Schmidt M, Steyerberg E.** Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review. *BMC Medical Research Methodology*, **2022**; 22(1):316.
8. **Pencina MJ, D'Agostino Sr RB, D'Agostino Jr RB, Vasan RS.** Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics in medicine*, **2008**; 27(2):157-172.
9. **Kerr KF, Brown MD, Zhu K, Janes H.** Assessing the clinical impact of risk prediction models with decision curves: guidance for correct interpretation and appropriate use. *Journal of Clinical Oncology*, **2016**; 34(21):2534-2540.
10. **Piovani D, Sokou R, Tsantes AG, Vitello AS, Bonovas S.** Optimizing Clinical Decision Making with Decision Curve Analysis: Insights for Clinical Investigators. *Healthcare*, **2023**; 11(16):2244.
11. **Hanley JA, McNeil BJ.** The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, **1982**; 143(1):29-36.
12. **Baker SG, Schuit E, Steyerberg EW, Pencina MJ, Vickers A, Moons KGM, Mol BWJ, Lindemans KS.** How to interpret a small increase in AUC with an additional risk prediction marker: decision analysis comes through. *Statistics in medicine*, **2014**; 33(22):3946-3959.

13. **Talluri R, Shete S.** Using the weighted area under the net benefit curve for decision curve analysis. *BMC medical informatics and decision making*, **2016**; 16:1-9.
14. **Vickers AJ, Holland F.** Decision curve analysis to evaluate the clinical benefit of prediction models. *The Spine Journal*, **2021**; 21(10):1643-1648.
15. **Cook NR.** Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*, **2007**; 115(7):928-935.
16. **Dinga R, Penninx BW, Veltman DJ, Schmaal L, Marquand AF.** Beyond accuracy: Measures for assessing machine learning models, pitfalls and guidelines. *BioRxiv*:743138, **2019**.
17. **Vickers AJ.** Decision analysis for the evaluation of diagnostic tests, prediction models, and molecular markers. *The American Statistician*, **2008**; 62(4):314-320.
18. **Goetz LH, Schork NJ.** Personalized medicine: motivation, challenges, and progress. *Fertility and sterility*, **2018**; 109(6):952-963.
19. **Lesko L, Zineh I, Huang SM.** What is Clinical Utility and Why Should We Care? *Pharmacoepidemiol Drug Saf*, **2010**; 88:729-733.
20. **Vickers AJ, Elkin EB.** Decision curve analysis: a novel method for evaluating prediction models. *Medical Decision Making*, **2006**; 26(6):565-574.
21. **Raschka S.** Model evaluation, model selection, and algorithm selection in machine learning. *arXiv preprint arXiv:1811.12808*, **2018**
22. **Zhang Z, Rousson V, Lee WC, Ferdynus C, Chen M, Qian X, Guo Y.** Decision curve analysis: a technical note. *Annals of translational medicine*, **2018**; 6(15).
23. **Sande SZ.** Statistical Inference for Decision Curve Analysis and Decision Threshold for Statistical Learning in Medical Research. PhD Thesis, National University of Singapore, Singapore, **2020**.
24. **Steen PM.** Approaches to predictive modeling. *The Annals of thoracic surgery*, **1994**; 58(6):1836-1840.
25. **Lee YH, Bang H, Kim DJ.** How to establish clinical prediction models. *Endocrinology and Metabolism*, **2016**; 31(1):38-44.
26. **Hendriksen JM, Geersing G-J, Moons KG, de Groot JA.** Diagnostic and prognostic prediction models. *Journal of Thrombosis and Haemostasis*, **2013**; 11:129-141.
27. **Toma M, Wei OC.** Predictive modeling in medicine. *Encyclopedia*, **2023**; 3(2):590-601.

28. **Waljee AK, Higgins PD, Singal AG.** A primer on predictive models. *Clinical and translational gastroenterology*, **2014**; 5(1):e44.
29. **Royston P, Moons KG, Altman DG, Vergouwe Y.** Prognosis and prognostic research: developing a prognostic model. *BMJ*, **2009**; 338.
30. **Leisman DE, Harhay MO, Lederer DJ, et al.** Development and reporting of prediction models: guidance for authors from editors of respiratory, sleep, and critical care journals. *Critical care medicine*, **2020**;48(5):623-633.
31. **Altman DG, Royston P.** What do we mean by validating a prognostic model? *Statistics in medicine*, **2000**; 19(4):453-473.
32. **Gerds TA, Cai T, Schumacher M.** The performance of risk prediction models. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, **2008**; 50(4):457-479.
33. **Su TL, Jaki T, Hickey GL, Buchan I, Sperrin M.** A review of statistical updating methods for clinical prediction models. *Statistical methods in medical research*, **2018**; 27(1):185-197.
34. **Rousson V, Goşoniu NF.** An R-square coefficient based on final prediction error. *Statistical Methodology*, **2007**; 4(3):331-340.
35. **Staerk C, Klinkhammer H, Wistuba T, Maj C, Mayr A.** Generalizability of polygenic prediction models: how is the R² defined on test data?. *BMC Medical Genomics*, **2024**; 17(1):132.
36. **Brier GW.** Verification of forecasts expressed in terms of probability. *Monthly weather review*, **1950**; 78(1):1-3.
37. **Redelmeier DA, Bloch DA, Hickam DH.** Assessing predictive accuracy: how to compare Brier scores. *Journal of clinical epidemiology*, **1991**; 44(11):1141-1146.
38. **Jiang X, Osl M, Kim J, Ohno-Machado L.** Calibrating predictive model estimates to support personalized medicine. *Journal of the American Medical Informatics Association*, **2012**; 19(2):263-274.
39. **Lindhiem O, Petersen IT, Mentch LK, Youngstrom EA.** The importance of calibration in clinical psychology. *Assessment*, **2020**; 27(4):840-854.
40. **Harrell FE.** *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. Vol. 608, New York: Springer, **2001**.
41. **Hosmer DW, Lemeshow S.** Goodness of fit tests for the multiple logistic regression model. *Communications in statistics-Theory and Methods*, **1980**; 9(10):1043-1069.

42. **Austin PC, Steyerberg EW.** Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. *Statistics in medicine*, **2014**; 33(3):517-535.
43. **Cook NR.** Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clinical chemistry*, **2008**; 54(1):17-23.
44. **Heagerty PJ, Zheng Y.** Survival model predictive accuracy and ROC curves. *Biometrics*, **2005**; 61(1):92-105.
45. **Coppus S, van der Veen F, Opmeer B, Mol B, Bossuyt P.** Evaluating prediction models in reproductive medicine. *Human reproduction*, **2009**;24(8):1774-1778.
46. **Van Calster B, McLernon DJ, Van Smeden M, Wynants L, Steyerberg EW, On behalf of Topic Group ‘Evaluating diagnostic tests and prediction models’ of the STRATOS initiative.** Calibration: the Achilles heel of predictive analytics. *BMC Med*, **2019**; 17(1):230.
47. **Holmberg L, Vickers A.** Evaluation of prediction models for decision-making: beyond calibration and discrimination. *PLoS medicine*, **2013**; 10(7):e1001491.
48. **Peek N, Arts D, Bosman R, van der Voort P, De Keizer N.** External validation of prognostic models for critically ill patients required substantial sample sizes. *Journal of clinical epidemiology*, **2007**; 60(5):491. e491-491. e413.
49. **Jiménez-Valverde A, Lobo JM, Hortal J.** Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and distributions*, **2008**; 14(6):885-890.
50. **Pencina MJ, D’Agostino RB, Massaro JM.** Understanding increments in model performance metrics. *Lifetime data analysis*, **2013**; 19:202-218.
51. **Huang Y, Li W, Macheret F, Gabriel RA, Ohno-Machado L.** A tutorial on calibration measurements and calibration models for clinical prediction models. *Journal of the American Medical Informatics Association*, **2020**; 27(4):621-633.
52. **Jiménez-Valverde A, Acevedo P, Barbosa AM, Lobo JM, Real R.** Discrimination capacity in species distribution models depends on the representativeness of the environmental domain. *Global Ecology and Biogeography*, **2013**; 22(4):508-516.
53. **Moons KG, de Groot JA, Linnet K, Reitsma JB, Bossuyt PM.** Quantifying the added value of a diagnostic test or marker. *Clinical chemistry*, **2012**; 58(10):1408-1417.
54. **Fitzgerald M, Saville BR, Lewis RJ.** Decision curve analysis. *Jama*, **2015**; 313(4):409-410.
55. **Steyerberg EW, Harrell Jr FE.** Prediction models need appropriate internal, internal-external, and external validation. *Journal of clinical epidemiology*, **2015**; 69:245.

56. **Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ, Roobol MJ, Steyerberg EW.** Reporting and interpreting decision curve analysis: a guide for investigators. *European urology*, **2018**; 74(6):796-804.
57. **Chalkou K, Vickers AJ, Pellegrini F, Manca A, Salanti G.** Decision curve analysis for personalized treatment choice between multiple options. *Medical Decision Making*, **2023**; 43(3):337-349.
58. **Zheng W, Zhu W, Yu S, Li K, Ding Y, Wu Q, Tang Q, Zhao Q, Lu C, Guo C.** Development and validation of a nomogram to predict overall survival for patients with metastatic renal cell carcinoma. *BMC cancer*, **2020**; 20:1-13.
59. **Yu C, Zhang Y.** Establishment of prognostic nomogram for elderly colorectal cancer patients: a SEER database analysis. *BMC gastroenterology*, **2020**; 20:1-13.
60. **Deng GC, Lv Y, Yan H, Sun DC, Qu TT, Pan YT, Han QL, Dai GH.** Nomogram to predict survival of patients with advanced and metastatic pancreatic Cancer. *BMC cancer*, **2021**; 21:1-12.
61. **Jiang A, Shi X, Zheng H, Liu N, Chen S, Gao H, Ren M, Zheng X, Fu X, Liang X, Ruan Z, Tian T, Yao Y.** Establishment and validation of a nomogram to predict the in-hospital death risk of nosocomial infections in cancer patients. *Antimicrobial Resistance & Infection Control*, **2022**; 11(1):29.
62. **Peiyuan G, Xuhua H, Ganlin G, Xu Y, Zining L, Jiachao H, Bin Y, Guiying W.** Construction and validation of a nomogram model for predicting the overall survival of colorectal cancer patients. *BMC surgery*, **2023**; 23(1):182.
63. **Bai G, Zhou Y, Rong Q, Qiao S, Mao H, Liu P.** Development of nomogram models based on peripheral blood score and clinicopathological parameters to predict preoperative advanced stage and prognosis for epithelial ovarian cancer patients. *Journal of Inflammation Research*, **2023**; 1227-1241.
64. **Cheng T, Yu D, Tan J, Liao S, Zhou L, OuYang W, Wen Z.** Development a nomogram prognostic model for survival in heart failure patients based on the HF-ACTION data. *BMC Medical Informatics and Decision Making*, **2024**; 24(1):197.
65. **Yuan C, Zou S, Wang K, Hu Z.** Establishment and external validation of prognosis prediction nomogram for patients with distant metastatic intrahepatic cholangiocarcinoma: based on a large population. *BMC cancer*, **2024**; 24(1):227.
66. **Pulleyblank R, Chuma J, Gilbody SM, Thompson C.** Decision curve analysis for assessing the usefulness of tests for making decisions to treat: An application to tests for prodromal psychosis. *Psychological assessment*, **2013**; 25(3):730.
67. **Lin S, Ma Y, He X, Zhu J, Zou H.** Using decision curve analysis to evaluate common strategies for myopia screening in school-aged children. *Ophthalmic Epidemiology*, **2019**; 26(4):286-294.

68. **Xie G, Wang R, Shang L, Qi C, Yang L, Huang L, Yang W, Chung MC.** Calculating the overall survival probability in patients with cervical cancer: a nomogram and decision curve analysis-based study. *BMC cancer*, **2020**; 20:1-9.
69. **Jiang J, Pan H, Li M, Qian B, Lin X, Fan S.** Predictive model for the 5-year survival status of osteosarcoma patients based on the SEER database and XGBoost algorithm. *Scientific reports*, **2021**; 11(1):5542.
70. **Pennisi G, Enea M, Falco V, Aithal GP, Palaniyappan N, Yilmaz Y, Boursier J, Cassinotto C, de Lédinghen V, Chan WK, Mahadeva S, Eddowes PE, Newsome P, Karlas T, Wiegand J, Wong VW-S, Schattenberg JM, Labenz C, Kim W, Lee MS, Lupsor-Platon M, Cobbold JFL, Fan J-G, Shen F, Stauffer K, Trauner M, Stauber R, Nakajima A, Yoneda M, Bugianesi E, Younes R, Gaia S, Zheng M-H, Cammà C, Anstee QM, Mózes FE, Pavlides M, Petta S.** Noninvasive assessment of liver disease severity in patients with nonalcoholic fatty liver disease (NAFLD) and type 2 diabetes. *Hepatology*, **2023**; 78(1):195-211.
71. **Rentroia-Pacheco B, Tokez S, Bramer EM, Venables ZC, van de Werken HJG, Bellomo D, van Klaveren D, Mooyaart AL, Hollestein LM, Wakkee M.** Personalised decision making to predict absolute metastatic risk in cutaneous squamous cell carcinoma: development and validation of a clinico-pathological model. *EClinicalMedicine*, **2023**; 63.
72. **Robin M, Porcher R, Michonneau D, Taurines L, Sicre de Fontbrune F, Xhaard A, Oriano B, Sutra Del Galy A, Peffault de Latour R, Soci G, Schlageter M-H.** Prospective external validation of biomarkers to predict acute graft-versus-host disease severity. *Blood Advances*, **2022**; 6(16):4763-4772.
73. **Jiang C, Qian C, Jiang Z, Teng Y, Lai R, Sun Y, Ni X, Ding C, Xu Y, Tian R.** Robust deep learning-based PET prognostic imaging biomarker for DLBCL patients: a multicenter study. *European Journal of Nuclear Medicine and Molecular Imaging*, **2023**; 50(13):3949-3960.
74. **Kaggle.** Pima Indians Diabetes Database. Erişim: <https://www.kaggle.com/code/mragpavank/pima-indians-diabetes-database>. Erişim tarihi: **Şubat**, 2022.
75. **Kaggle.** Prostate Cancer Dataset. Erişim: <https://www.kaggle.com/datasets/sajidsaifi/prostate-cancer>. Erişim tarihi: **Şubat**, 2022.
76. **McClish DK.** Comparing the areas under more than two independent ROC curves. *Medical Decision Making*, **1987**; 7(3):149-155.
77. **Efron B.** Bootstrap methods: another look at the jackknife. *Breakthroughs in statistics: Methodology and distribution*: New York: Springer; **1992**; 569-593.
78. **Efron B, Tibshirani R.** Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, **1986**; 1(1):54-75.
79. **Hesterberg T, Moore D, Monaghan S, Clipson A.** Bootstrap methods and permutation tests. *Introduction to the Practice of Statistics*. 5th ed. New York, **2005**. p. 14.11–14.70.

80. **Singh K, Xie M.** Bootstrap: a statistical method. *Unpublished manuscript, Rutgers University, USA*. Retrieved from <http://www.stat.rutgers.edu/home/mxie/RCPapers/bootstrap.pdf>, **2008**; 1-14.
81. **Cortes C.** Support-Vector Networks. *Machine Learning*, **1995**.
82. **Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B.** Support vector machines. *IEEE Intelligent Systems and their applications*, **1998**; 13(4):18-28.
83. **Bishop CM.** *Information science and statistics*. Springer, **2006**.
84. **Quinlan JR.** Induction of decision trees. *Machine learning*, **1986**; 1:81-106.
85. **Breiman L.** *Classification and regression trees*. Routledge, **2017**.
86. **Utgoff PE, Berkman NC, Clouse JA.** Decision tree induction based on efficient tree restructuring. *Machine Learning*, **1997**; 29:5-44.
87. **Agresti A.** *Categorical data analysis*. Vol 792: John Wiley & Sons, **2012**.
88. **Cohen J.** A power primer. *Psychological Bulletin*. **1992**; 112(1):155-159.
89. **Vickers AJ, Van Calster B, Steyerberg EW.** Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ*, **2016**; 352.
90. **Mallett S, Halligan S, Thompson M, Collins GS, Altman DG.** Interpreting diagnostic accuracy studies for patient care. *BMJ*, **2012**; 345.
91. **Steyerberg EW, Vickers AJ.** Decision curve analysis: a discussion. *Medical Decision Making*, **2008**; 28(1):146-149.
92. **Vickers AJ, Van Claster B, Wynants L, Steyerberg EW.** Decision curve analysis: confidence intervals and hypothesis testing for net benefit. *Diagnostic and prognostic research*, **2023**; 7(1):11.
93. **Steyerberg EW, Vergouwe Y.** Towards better clinical prediction models: seven steps for development and an ABCD for validation. *European heart journal*, **2014**; 35(29):1925-1931.

EKLER

EK-1: Simülasyon Çalışması için Basamaklar



EK-2: Farklı Simülasyon Senaryolarına Ait Bulgular

5) Çok değişkenli normal dağılımdan ortalaması 0 ve varyans kovaryans matrisi $diag(1,1,1)$ olan, 3 açıklayıcı (2 kategorik, 1 sürekli), 1 yanıt değişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betanın ise $\beta=(0.1, 0.3, 0.3)$ olduğu senaryoya ait bulgular Çizelge 8.1 ve 8.2’de sunulmuştur.

Çizelge 8.1. Senaryo 5 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.10	100	17443	0.030	0.527±0.099	0.527±0.099	0.524±0.100	100.000±0.000	100.000±0.000	98.821±0.103
0.10	250	3957	0.063	0.528±0.063	0.528±0.063	0.524±0.064	100.000±0.000	100.000±0.000	97.464±0.162
0.10	500	3498	0.067	0.528±0.045	0.528±0.045	0.525±0.045	100.000±0.000	100.000±0.000	96.592±0.194
0.10	1000	16336	0.031	0.528±0.032	0.528±0.032	0.527±0.032	100.000±0.000	100.000±0.000	98.171±0.130
0.15	100	1778	0.094	0.498±0.105	0.498±0.105	0.488±0.108	100.000±0.000	100.000±0.000	96.576±0.187
0.15	250	508	0.176	0.499±0.067	0.500±0.067	0.488±0.069	99.770±0.446	99.800±0.471	92.360±0.269
0.15	500	441	0.189	0.499±0.047	0.500±0.047	0.491±0.048	99.950±0.219	99.940±0.239	88.552±0.321
0.15	1000	4362	0.060	0.499±0.034	0.500±0.033	0.498±0.034	100.000±0.000	100.000±0.000	92.665±0.294
0.20	100	889	0.133	0.465±0.113	0.465±0.111	0.450±0.114	99.990±0.100	99.980±0.141	93.540±0.287
0.20	250	283	0.236	0.465±0.072	0.468±0.071	0.451±0.073	97.020±1.853	96.690±1.796	87.447±0.297
0.20	500	286	0.235	0.465±0.052	0.469±0.050	0.457±0.052	99.150±0.892	98.560±1.140	80.478±0.370
0.20	1000	814	0.139	0.464±0.037	0.469±0.035	0.466±0.036	100.000±0.000	100.000±0.000	79.367±0.398
0.25	100	508	0.176	0.425±0.122	0.428±0.118	0.407±0.121	99.610±0.601	99.42±0.843	90.793±0.290
0.25	250	137	0.340	0.425±0.079	0.432±0.075	0.406±0.078	51.820±4.641	51.160±4.483	79.149±0.392
0.25	500	128	0.352	0.423±0.059	0.433±0.053	0.414±0.055	57.200±5.314	53.270±5.150	67.629±0.496
0.25	1000	163	0.312	0.420±0.045	0.433±0.038	0.428±0.038	61.790±4.418	83.670±3.869	65.045±0.479
0.30	100	193	0.286	0.376±0.135	0.385±0.125	0.349±0.127	81.630±4.170	78.540±4.620	85.026±0.398
0.30	250	57	0.534	0.376±0.091	0.391±0.080	0.348±0.081	0.220±0.462	0.250±0.458	67.670±0.441
0.30	500	56	0.539	0.372±0.071	0.393±0.057	0.362±0.058	0.000±0.000	0.010±0.100	53.584±0.530
0.30	1000	59	0.523	0.366±0.061	0.393±0.040	0.384±0.041	0.460±0.658	8.820±2.959	53.487±0.531

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 8.2. Senaryo 5 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NFDVM (ort±ss)	NFLR (ort±ss)	NFKA (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.35	100	147	0.328	0.312±0.152	0.337±0.132	0.294±0.130	67.200±4.887	64.310±4.878	79.036±0.491
0.35	250	42	0.623	0.313±0.108	0.345±0.085	0.293±0.082	0.000±0.000	0.010±0.100	60.712±0.488
0.35	500	37	0.667	0.308±0.090	0.347±0.060	0.307±0.060	0.000±0.000	0.000±0.000	46.793±0.517
0.35	1000	30	0.741	0.299±0.083	0.348±0.043	0.333±0.044	0.000±0.000	0.030±0.171	47.260±0.474
0.40	100	141	0.335	0.233±0.165	0.284±0.138	0.241±0.132	29.42±4.639	31.11±4.744	75.875±0.440
0.40	250	48	0.580	0.229±0.131	0.294±0.089	0.241±0.083	0.000±0.000	0.000±0.000	56.953±0.462
0.40	500	26	0.798	0.222±0.117	0.297±0.063	0.254±0.061	0.000±0.000	0.000±0.000	42.923±0.541
0.40	1000	18	0.980	0.216±0.111	0.299±0.045	0.278±0.045	0.000±0.000	0.000±0.000	42.889±0.465
0.45	100	72	0.473	0.155±0.165	0.228±0.143	0.189±0.136	5.340±2.133	3.770±1.814	75.930±0.442
0.45	250	26	0.805	0.146±0.135	0.239±0.092	0.192±0.087	0.000±0.000	0.000±0.000	57.611±0.532
0.45	500	16	1.024	0.141±0.126	0.244±0.066	0.207±0.063	0.000±0.000	0.000±0.000	44.165±0.464
0.45	1000	14	1.128	0.141±0.123	0.246±0.047	0.226±0.046	0.000±0.000	0.000±0.000	42.418±0.505
0.50	100	54	0.548	0.087±0.156	0.170±0.147	0.140±0.143	1.350±1.104	0.370±0.597	76.923±0.407
0.50	250	22	0.869	0.088±0.124	0.184±0.095	0.146±0.092	0.000±0.000	0.000±0.000	64.156±0.526
0.50	500	16	1.056	0.090±0.116	0.190±0.067	0.161±0.067	0.000±0.000	0.000±0.000	50.139±0.501
0.50	1000	13	1.150	0.091±0.115	0.192±0.047	0.177±0.048	0.000±0.000	0.000±0.000	44.187±0.529
0.55	100	52	0.560	0.032±0.143	0.114±0.150	0.089±0.155	1.570±1.312	0.410±0.767	78.953±0.392
0.55	250	23	0.861	0.042±0.109	0.130±0.095	0.093±0.100	0.000±0.000	0.000±0.000	68.825±0.462
0.55	500	16	1.029	0.050±0.099	0.137±0.067	0.105±0.073	0.000±0.000	0.000±0.000	55.846±0.466
0.55	1000	13	1.155	0.053±0.097	0.141±0.047	0.120±0.053	0.000±0.000	0.000±0.000	44.211±0.434
0.60	100	92	0.416	0.006±0.128	0.064±0.150	0.031±0.169	9.830±3.282	4.230±2.093	80.589±0.397
0.60	250	26	0.802	0.009±0.090	0.082±0.092	0.033±0.110	0.000±0.000	0.000±0.000	72.207±0.466
0.60	500	17	1.001	0.018±0.079	0.090±0.064	0.041±0.081	0.000±0.000	0.000±0.000	59.767±0.475
0.60	1000	14	1.121	0.024±0.076	0.094±0.045	0.058±0.058	0.000±0.000	0.000±0.000	46.159±0.496

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

6) Çok deęişkenli normal dağılımdan ortalaması **0** ve varyans kovaryans matrisi *diag(1,1,1)* olan, 3 açıklayıcı (2 kategorik, 1 sürekli), 1 yanıt deęişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betanın ise **$\beta=(0.3, 0.3, 0.3)$** olduęu senaryoya ait bulgular Çizelge 8.3 ve 8.4’de sunulmuştur.

Çizelge 8.3. Senaryo 6 için tahmin modellerinden hesaplanan net fayda deęerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.10	100	17443	0.030	0.566±0.099	0.567±0.098	0.564±0.099	100.000±0.000	100.000±0.000	98.938±0.101
0.10	250	6815	0.048	0.567±0.062	0.568±0.062	0.565±0.063	100.000±0.000	100.000±0.000	97.891±0.140
0.10	500	3396	0.068	0.567±0.044	0.568±0.044	0.565±0.044	100.000±0.000	100.000±0.000	96.653±0.176
0.10	1000	15331	0.032	0.567±0.031	0.568±0.031	0.567±0.031	100.000±0.000	100.000±0.000	97.523±0.156
0.15	100	2124	0.086	0.539±0.105	0.541±0.104	0.532±0.106	100.000±0.000	100.000±0.000	96.732±0.159
0.15	250	709	0.149	0.540±0.067	0.542±0.066	0.532±0.068	100.000±0.000	99.990±0.100	92.722±0.267
0.15	500	558	0.168	0.540±0.048	0.542±0.047	0.534±0.048	100.000±0.000	99.990±0.100	87.333±0.350
0.15	1000	4220	0.061	0.540±0.034	0.542±0.033	0.540±0.033	100.000±0.000	100.000±0.000	88.895±0.312
0.20	100	1347	0.108	0.507±0.114	0.510±0.110	0.498±0.113	100.000±0.000	100.000±0.000	93.676±0.227
0.20	250	406	0.197	0.507±0.073	0.513±0.070	0.499±0.072	99.890±0.345	99.830±0.403	87.422±0.255
0.20	500	326	0.220	0.507±0.054	0.513±0.049	0.502±0.051	99.940±0.239	99.770±0.446	79.145±0.488
0.20	1000	455	0.186	0.507±0.040	0.514±0.035	0.510±0.035	99.620±0.632	99.980±0.141	80.016±0.397
0.25	100	769	0.143	0.468±0.126	0.476±0.117	0.459±0.121	99.970±0.171	99.980±0.141	90.668±0.276
0.25	250	187	0.291	0.468±0.083	0.480±0.074	0.458±0.077	90.230±3.178	87.490±3.292	79.284±0.384
0.25	500	141	0.336	0.467±0.065	0.481±0.053	0.463±0.054	64.900±4.462	65.900±4.357	68.224±0.492
0.25	1000	147	0.328	0.466±0.053	0.481±0.037	0.475±0.038	44.300±4.896	90.460±3.264	70.595±0.487
0.30	100	259	0.247	0.416±0.146	0.437±0.124	0.406±0.127	96.640±1.915	95.380±2.291	83.919±0.394
0.30	250	71	0.475	0.419±0.100	0.442±0.079	0.404±0.081	1.940±1.355	1.950±1.218	67.262±0.436
0.30	500	58	0.526	0.417±0.082	0.444±0.056	0.414±0.058	0.020±0.141	0.050±0.219	54.206±0.511
0.30	1000	66	0.493	0.415±0.073	0.444±0.040	0.434±0.041	1.400±1.015	37.340±4.778	58.848±0.567

pt: Eşik olasılık, n:Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 8.4. Senaryo 6 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.35	100	166	0.309	0.346±0.171	0.393±0.131	0.352±0.132	43.500±4.670	57.190±4.863	75.457±0.414
0.35	250	70	0.480	0.346±0.135	0.400±0.084	0.352±0.083	0.010±0.100	0.020±0.141	56.535±0.494
0.35	500	45	0.602	0.346±0.117	0.402±0.060	0.363±0.060	0.000±0.000	0.000±0.000	45.964±0.428
0.35	1000	34	0.700	0.347±0.105	0.403±0.042	0.387±0.043	0.000±0.000	0.570±0.795	50.774±0.481
0.40	100	72	0.471	0.266±0.185	0.343±0.139	0.299±0.135	3.160±1.722	6.840±2.159	70.935±0.490
0.40	250	29	0.750	0.254±0.162	0.352±0.089	0.300±0.086	0.000±0.000	0.000±0.000	50.058±0.545
0.40	500	20	0.925	0.245±0.156	0.355±0.063	0.312±0.062	0.000±0.000	0.000±0.000	40.106±0.438
0.40	1000	16	1.029	0.240±0.153	0.356±0.045	0.335±0.045	0.000±0.000	0.000±0.000	41.977±0.511
0.45	100	45	0.603	0.194±0.180	0.291±0.139	0.246±0.133	0.130±0.367	0.220±0.440	69.014±0.467
0.45	250	23	0.864	0.187±0.160	0.300±0.093	0.251±0.089	0.000±0.000	0.000±0.000	51.218±0.494
0.45	500	17	1.010	0.185±0.153	0.304±0.066	0.265±0.064	0.000±0.000	0.000±0.000	43.161±0.447
0.45	1000	14	1.102	0.185±0.148	0.306±0.047	0.285±0.046	0.000±0.000	0.000±0.000	43.300±0.487
0.50	100	47	0.589	0.136±0.174	0.232±0.151	0.197±0.148	0.410±0.637	0.180±0.386	73.167±0.480
0.50	250	22	0.876	0.137±0.145	0.245±0.097	0.205±0.095	0.000±0.000	0.000±0.000	58.100±0.485
0.50	500	16	1.039	0.138±0.136	0.250±0.069	0.219±0.069	0.000±0.000	0.000±0.000	47.376±0.477
0.50	1000	13	1.153	0.138±0.131	0.252±0.049	0.235±0.049	0.000±0.000	0.000±0.000	44.177±0.518
0.55	100	46	0.591	0.079±0.161	0.173±0.157	0.144±0.159	0.420±0.727	0.100±0.333	76.403±0.402
0.55	250	22	0.874	0.090±0.126	0.189±0.099	0.151±0.103	0.000±0.000	0.000±0.000	64.255±0.466
0.55	500	15	1.085	0.093±0.113	0.195±0.070	0.163±0.075	0.000±0.000	0.000±0.000	51.209±0.514
0.55	1000	12	1.255	0.094±0.106	0.198±0.050	0.176±0.054	0.000±0.000	0.000±0.000	39.917±0.525
0.60	100	43	0.617	0.028±0.142	0.119±0.153	0.084±0.167	0.520±0.745	0.070±0.256	77.788±0.453
0.60	250	21	0.899	0.040±0.109	0.134±0.100	0.089±0.113	0.000±0.000	0.000±0.000	68.352±0.448
0.60	500	13	1.158	0.046±0.094	0.142±0.070	0.096±0.083	0.000±0.000	0.000±0.000	53.397±0.446
0.60	1000	9	1.413	0.048±0.085	0.146±0.049	0.108±0.060	0.000±0.000	0.000±0.000	35.812±0.520

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

7) Çok deęişkenli normal dağılımdan ortalaması **0** ve varyans kovaryans matrisi *diag(1,1,1)* olan, 3 açıklayıcı (2 kategorik, 1 sürekli), 1 yanıt deęişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betanın ise $\beta=(0.3, 0.5, 0.5)$ olduęu senaryoya ait bulgular Çizelge 8.5 ve 8.6’da sunulmuştur.

Çizelge 8.5. Senaryo 7 için tahmin modellerinden hesaplanan net fayda deęerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.10	100	638	0.157	0.587±0.096	0.594±0.092	0.591±0.093	100.000±0.000	100.000±0.000	96.060±0.179
0.10	250	203	0.279	0.585±0.066	0.595±0.061	0.592±0.062	99.880±0.327	100.000±0.000	88.969±0.301
0.10	500	78	0.454	0.582±0.050	0.595±0.043	0.592±0.044	76.990±4.089	90.340±2.958	76.585±0.412
0.10	1000	647	0.156	0.579±0.039	0.595±0.031	0.594±0.031	1.880±1.486	12.240±3.531	66.078±0.455
0.15	100	148	0.327	0.553±0.108	0.569±0.097	0.561±0.100	99.870±0.338	99.980±0.141	89.823±0.265
0.15	250	58	0.529	0.547±0.081	0.571±0.065	0.562±0.066	54.520±5.192	79.650±3.942	75.744±0.438
0.15	500	30	0.741	0.540±0.069	0.571±0.046	0.564±0.047	0.540±0.731	9.110±3.021	62.173±0.505
0.15	1000	261	0.246	0.534±0.063	0.571±0.032	0.569±0.033	0.000±0.000	0.250±0.520	56.496±0.501
0.20	100	65	0.496	0.513±0.124	0.541±0.103	0.529±0.106	90.440±2.443	97.610±1.510	82.256±0.325
0.20	250	32	0.717	0.501±0.102	0.544±0.068	0.531±0.070	2.490±1.617	20.450±4.103	66.088±0.457
0.20	500	22	0.885	0.491±0.095	0.545±0.048	0.534±0.050	0.000±0.000	0.760±0.922	54.908±0.504
0.20	1000	149	0.326	0.483±0.093	0.545±0.034	0.541±0.035	0.000±0.000	0.090±0.321	48.711±0.484
0.25	100	42	0.624	0.467±0.148	0.510±0.114	0.494±0.117	53.310±5.340	82.410±3.809	78.483±0.414
0.25	250	25	0.818	0.450±0.126	0.514±0.072	0.495±0.074	0.020±0.141	2.390±1.325	58.884±0.483
0.25	500	19	0.953	0.438±0.123	0.515±0.051	0.499±0.052	0.000±0.000	0.060±0.239	49.366±0.490
0.25	1000	75	0.462	0.429±0.124	0.516±0.036	0.508±0.037	0.000±0.000	0.060±0.239	46.484±0.549
0.30	100	30	0.741	0.407±0.174	0.476±0.120	0.449±0.123	5.730±2.420	28.280±4.731	70.975±0.469
0.30	250	21	0.896	0.392±0.154	0.482±0.076	0.451±0.077	0.000±0.000	0.030±0.171	52.257±0.512
0.30	500	18	0.990	0.382±0.150	0.483±0.054	0.458±0.055	0.000±0.000	0.000±0.000	44.232±0.448
0.30	1000	43	0.615	0.375±0.151	0.484±0.038	0.471±0.039	0.000±0.000	0.010±0.100	44.601±0.477

pt: Eşik olasılık, n:Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 8.6. Senaryo 7 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.35	100	22	0.875	0.338±0.198	0.440±0.126	0.405±0.127	0.050±0.219	1.200±0.974	64.381±0.489
0.35	250	18	0.973	0.324±0.182	0.447±0.080	0.409±0.080	0.000±0.000	0.000±0.000	47.717±0.502
0.35	500	16	1.024	0.321±0.177	0.449±0.057	0.417±0.057	0.000±0.000	0.000±0.000	42.791±0.497
0.35	1000	32	0.718	0.320±0.175	0.450±0.040	0.431±0.041	0.000±0.000	0.000±0.000	43.249±0.479
0.40	100	20	0.926	0.275±0.210	0.401±0.132	0.364±0.131	0.010±0.100	0.190±0.419	62.185±0.498
0.40	250	18	0.989	0.270±0.195	0.409±0.084	0.368±0.083	0.000±0.000	0.000±0.000	47.812±0.526
0.40	500	16	1.030	0.271±0.191	0.411±0.060	0.378±0.059	0.000±0.000	0.000±0.000	44.275±0.515
0.40	1000	27	0.790	0.271±0.189	0.412±0.042	0.393±0.042	0.000±0.000	0.000±0.000	44.795±0.499
0.45	100	19	0.940	0.223±0.209	0.361±0.132	0.321±0.131	0.000±0.000	0.010±0.100	61.062±0.451
0.45	250	17	1.002	0.224±0.198	0.368±0.088	0.328±0.087	0.000±0.000	0.000±0.000	49.145±0.513
0.45	500	16	1.031	0.227±0.195	0.372±0.062	0.341±0.062	0.000±0.000	0.000±0.000	46.710±0.492
0.45	1000	28	0.767	0.228±0.194	0.373±0.044	0.356±0.044	0.000±0.000	0.000±0.000	47.822±0.438
0.50	100	19	0.946	0.179±0.207	0.316±0.145	0.280±0.146	0.000±0.000	0.020±0.141	65.052±0.418
0.50	250	17	1.009	0.183±0.193	0.326±0.092	0.289±0.093	0.000±0.000	0.000±0.000	51.884±0.485
0.50	500	16	1.037	0.186±0.191	0.330±0.065	0.302±0.066	0.000±0.000	0.000±0.000	48.594±0.472
0.50	1000	28	0.767	0.188±0.191	0.332±0.046	0.316±0.047	0.000±0.000	0.010±0.100	50.146±0.529
0.55	100	19	0.937	0.134±0.197	0.268±0.149	0.233±0.155	0.000±0.000	0.000±0.000	67.992±0.457
0.55	250	17	1.004	0.146±0.182	0.282±0.095	0.244±0.101	0.000±0.000	0.000±0.000	53.688±0.490
0.55	500	16	1.031	0.150±0.181	0.287±0.067	0.256±0.073	0.000±0.000	0.000±0.000	48.969±0.495
0.55	1000	29	0.755	0.152±0.182	0.289±0.047	0.268±0.052	0.000±0.000	0.000±0.000	48.970±0.556
0.60	100	20	0.916	0.092±0.188	0.222±0.155	0.181±0.168	0.000±0.000	0.000±0.000	70.775±0.392
0.60	250	18	0.988	0.111±0.168	0.237±0.098	0.192±0.111	0.000±0.000	0.000±0.000	56.843±0.539
0.60	500	17	1.016	0.117±0.165	0.242±0.069	0.200±0.080	0.000±0.000	0.000±0.000	48.187±0.532
0.60	1000	17	1.016	0.120±0.167	0.245±0.049	0.212±0.057	0.000±0.000	0.000±0.000	45.750±0.425

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

8) Çok deęişkenli normal daęılımından ortalaması **0** ve varyans kovaryans matrisi

$$\begin{pmatrix} 1 & 0.3 & 0.1 \\ 0.3 & 1 & 0.3 \\ 0.1 & 0.3 & 1 \end{pmatrix}$$

olan, 3 açıklayıcı (1 kategorik, 2 sürekli), 1 yanıt deęişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betaların ise $\beta=(0.1, 0.3, 0.3)$ olduęu senaryoya ait bulgular Çizelge 8.7 ve 8.8’de sunulmuştur.

Çizelge 8.7. Senaryo 8 için tahmin modellerinden hesaplanan net fayda deęerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.10	100	17443	0.030	0.526±0.099	0.526±0.099	0.523±0.100	100.000±0.000	100.000±0.000	98.645±0.107
0.10	250	3957	0.063	0.526±0.063	0.527±0.063	0.523±0.064	100.000±0.000	100.000±0.000	97.279±0.168
0.10	500	3498	0.067	0.527±0.045	0.527±0.045	0.524±0.045	100.000±0.000	100.000±0.000	96.497±0.172
0.10	1000	16336	0.031	0.527±0.032	0.527±0.032	0.526±0.032	100.000±0.000	100.000±0.000	98.215±0.129
0.15	100	1741	0.095	0.497±0.105	0.497±0.104	0.487±0.107	100.000±0.000	100.000±0.000	96.156±0.182
0.15	250	497	0.178	0.497±0.067	0.499±0.067	0.487±0.068	99.910±0.321	99.870±0.367	92.080±0.262
0.15	500	558	0.168	0.498±0.047	0.499±0.047	0.491±0.048	99.980±0.141	99.950±0.219	88.423±0.331
0.15	1000	4362	0.060	0.498±0.034	0.499±0.033	0.497±0.034	100.000±0.000	100.000±0.000	91.477±0.251
0.20	100	889	0.133	0.463±0.113	0.464±0.111	0.449±0.114	99.990±0.100	99.960±0.197	92.912±0.274
0.20	250	320	0.222	0.463±0.073	0.467±0.071	0.451±0.073	98.050±1.459	97.840±1.536	86.584±0.349
0.20	500	332	0.218	0.463±0.052	0.467±0.050	0.456±0.051	99.740±0.597	99.480±0.797	79.683±0.419
0.20	1000	585	0.164	0.462±0.038	0.468±0.035	0.465±0.036	99.940±0.239	99.980±0.141	78.145±0.351
0.25	100	558	0.168	0.422±0.124	0.426±0.117	0.406±0.121	99.710±0.556	99.600±0.667	89.649±0.276
0.25	250	146	0.329	0.422±0.081	0.431±0.075	0.406±0.077	65.320±4.546	61.130±4.756	78.095±0.431
0.25	500	143	0.333	0.420±0.061	0.432±0.053	0.414±0.055	71.580±3.906	67.370±4.338	67.247±0.478
0.25	1000	136	0.342	0.417±0.049	0.432±0.038	0.427±0.038	35.280±4.858	72.690±4.625	64.668±0.504
0.30	100	212	0.273	0.370±0.138	0.384±0.124	0.350±0.125	89.630±2.733	86.950±3.509	83.627±0.362
0.30	250	64	0.503	0.370±0.096	0.390±0.079	0.350±0.080	0.700±0.835	0.670±0.817	66.514±0.400
0.30	500	62	0.509	0.366±0.078	0.392±0.056	0.363±0.058	0.060±0.239	0.170±0.403	53.800±0.471
0.30	1000	50	0.567	0.360±0.069	0.392±0.040	0.384±0.041	0.130±0.393	7.600±2.923	54.013±0.482

pt: Eşik olasılık, n:Örnek sayısı, d: Etki büyüklüęü

*: Etki büyüklüęüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 8.8. Senaryo 8 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.35	100	157	0.318	0.303±0.155	0.337±0.130	0.296±0.128	67.570±4.522	66.510±4.476	77.653±0.447
0.35	250	48	0.582	0.301±0.117	0.345±0.084	0.297±0.081	0.000±0.000	0.010±0.100	58.722±0.476
0.35	500	40	0.641	0.293±0.103	0.347±0.060	0.310±0.059	0.000±0.000	0.000±0.000	46.016±0.523
0.35	1000	22	0.868	0.282±0.099	0.348±0.042	0.333±0.043	0.000±0.000	0.000±0.000	45.921±0.470
0.40	100	99	0.402	0.225±0.166	0.286±0.136	0.247±0.130	14.63±3.187	16.85±3.796	74.791±0.389
0.40	250	33	0.702	0.216±0.135	0.296±0.088	0.247±0.083	0.000±0.000	0.000±0.000	54.492±0.452
0.40	500	20	0.928	0.208±0.124	0.299±0.062	0.259±0.060	0.000±0.000	0.000±0.000	41.749±0.498
0.40	1000	15	1.092	0.203±0.119	0.301±0.044	0.281±0.044	0.000±0.000	0.000±0.000	41.456±0.552
0.45	100	54	0.547	0.149±0.166	0.233±0.140	0.197±0.134	1.150±1.114	0.920±0.961	74.758±0.521
0.45	250	21	0.902	0.140±0.136	0.244±0.090	0.201±0.086	0.000±0.000	0.000±0.000	55.437±0.526
0.45	500	14	1.107	0.137±0.128	0.249±0.064	0.215±0.062	0.000±0.000	0.000±0.000	42.566±0.462
0.45	1000	12	1.200	0.138±0.125	0.251±0.046	0.233±0.045	0.000±0.000	0.000±0.000	40.048±0.639
0.50	100	44	0.611	0.087±0.159	0.180±0.145	0.150±0.143	0.220±0.504	0.060±0.239	75.592±0.417
0.50	250	19	0.942	0.088±0.128	0.193±0.092	0.156±0.092	0.000±0.000	0.000±0.000	61.303±0.425
0.50	500	14	1.140	0.089±0.120	0.199±0.065	0.171±0.066	0.000±0.000	0.000±0.000	47.168±0.534
0.50	1000	12	1.210	0.091±0.120	0.201±0.046	0.187±0.047	0.000±0.000	0.000±0.000	42.082±0.475
0.55	100	41	0.633	0.034±0.149	0.128±0.148	0.100±0.154	0.220±0.484	0.070±0.256	77.438±0.398
0.55	250	21	0.908	0.047±0.117	0.143±0.093	0.104±0.100	0.000±0.000	0.000±0.000	65.624±0.450
0.55	500	15	1.084	0.054±0.107	0.150±0.065	0.117±0.073	0.000±0.000	0.000±0.000	52.250±0.467
0.55	1000	13	1.172	0.058±0.105	0.153±0.046	0.132±0.052	0.000±0.000	0.000±0.000	43.211±0.468
0.60	100	54	0.545	0.002±0.137	0.080±0.149	0.045±0.169	2.030±1.389	0.800±0.853	79.272±0.415
0.60	250	24	0.837	0.017±0.099	0.097±0.092	0.047±0.110	0.000±0.000	0.000±0.000	69.762±0.438
0.60	500	18	0.982	0.029±0.087	0.104±0.064	0.056±0.080	0.000±0.000	0.000±0.000	57.432±0.491
0.60	1000	15	1.093	0.035±0.083	0.108±0.045	0.073±0.057	0.000±0.000	0.000±0.000	45.477±0.477

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

9) Çok deęişkenli normal daęılımıdan ortalaması **0** ve varyans kovaryans matrisi

$$\begin{pmatrix} 1 & 0.3 & 0.1 \\ 0.3 & 1 & 0.3 \\ 0.1 & 0.3 & 1 \end{pmatrix}$$

olan, 3 açıklayıcı (1 kategorik, 2 sürekli), 1 yanıt deęişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betaların ise $\beta=(0.3, 0.3, 0.3)$ olduęu senaryoya ait bulgular Çizelge 8.9 ve 8.10’da sunulmuştur.

Çizelge 8.9. Senaryo 9 için tahmin modellerinden hesaplanan net fayda deęerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±sd)	NF _{LR} (ort±sd)	NF _{KA} (ort±sd)	ANOVA** (ort±sd)	KW** (ort±sd)	z** (ort±sd)
0.10	100	17443	0.030	0.564±0.098	0.565±0.098	0.562±0.099	100.000±0.000	100.000±0.000	98.707±0.126
0.10	250	3834	0.064	0.565±0.062	0.566±0.062	0.562±0.063	100.000±0.000	100.000±0.000	97.416±0.159
0.10	500	3396	0.068	0.565±0.044	0.566±0.044	0.563±0.044	100.000±0.000	100.000±0.000	96.006±0.194
0.10	1000	15331	0.032	0.565±0.031	0.566±0.031	0.565±0.031	100.000±0.000	100.000±0.000	96.273±0.159
0.15	100	2719	0.076	0.536±0.106	0.538±0.104	0.530±0.106	100.000±0.000	100.000±0.000	96.193±0.175
0.15	250	699	0.150	0.537±0.067	0.540±0.066	0.530±0.067	100.000±0.000	100.000±0.000	91.694±0.249
0.15	500	532	0.172	0.536±0.048	0.540±0.046	0.532±0.047	100.000±0.000	99.990±0.100	85.170±0.359
0.15	1000	728	0.147	0.535±0.035	0.540±0.033	0.538±0.033	99.980±0.141	100.000±0.000	82.583±0.383
0.20	100	1148	0.117	0.502±0.115	0.508±0.110	0.495±0.113	100.000±0.000	100.000±0.000	92.526±0.260
0.20	250	470	0.183	0.502±0.076	0.510±0.070	0.497±0.072	99.940±0.239	99.870±0.367	85.305±0.368
0.20	500	320	0.222	0.500±0.057	0.511±0.049	0.500±0.050	98.520±1.168	99.580±0.684	75.557±0.402
0.20	1000	139	0.338	0.497±0.047	0.511±0.035	0.508±0.035	30.430±4.808	73.860±3.752	72.122±0.456
0.25	100	672	0.153	0.462±0.129	0.474±0.116	0.456±0.120	99.980±0.141	99.990±0.100	88.843±0.352
0.25	250	184	0.293	0.460±0.089	0.478±0.074	0.456±0.076	88.160±3.407	88.980±3.393	76.416±0.415
0.25	500	132	0.347	0.456±0.073	0.478±0.052	0.462±0.054	23.480±4.138	49.960±4.977	64.583±0.444
0.25	1000	49	0.573	0.448±0.067	0.479±0.037	0.473±0.038	0.240±0.495	17.370±4.092	61.986±0.488
0.30	100	406	0.197	0.407±0.150	0.434±0.123	0.404±0.126	94.240±2.396	95.930±1.960	81.779±0.423
0.30	250	127	0.353	0.406±0.111	0.440±0.079	0.404±0.080	1.440±0.998	2.720±1.450	64.597±0.469
0.30	500	55	0.543	0.399±0.097	0.442±0.056	0.413±0.057	0.000±0.000	0.030±0.171	51.599±0.483
0.30	1000	32	0.723	0.390±0.094	0.442±0.039	0.432±0.040	0.010±0.100	1.870±1.382	52.314±0.484

pt: Eşik olasılık, n:Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 8.10. Senaryo 9 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.35	100	119	0.365	0.335±0.174	0.391±0.130	0.353±0.129	25.470±4.181	42.250±4.920	73.758±0.413
0.35	250	40	0.638	0.268±0.096	0.289±0.084	0.237±0.079	0.000±0.000	0.000±0.000	65.760±0.486
0.35	500	38	0.655	0.263±0.076	0.292±0.060	0.253±0.059	0.000±0.000	0.000±0.000	51.072±0.512
0.35	1000	22	0.879	0.319±0.125	0.401±0.042	0.386±0.043	0.000±0.000	0.000±0.000	46.549±0.533
0.40	100	56	0.536	0.256±0.185	0.343±0.136	0.304±0.133	1.230±1.153	2.910±1.928	69.793±0.475
0.40	250	25	0.819	0.245±0.163	0.352±0.087	0.306±0.084	0.000±0.000	0.000±0.000	49.374±0.494
0.40	500	18	0.977	0.239±0.156	0.355±0.062	0.316±0.060	0.000±0.000	0.000±0.000	41.023±0.466
0.40	1000	15	1.072	0.237±0.152	0.357±0.044	0.336±0.044	0.000±0.000	0.000±0.000	42.100±0.483
0.45	100	40	0.635	0.189±0.183	0.293±0.142	0.254±0.138	0.050±0.261	0.130±0.367	70.105±0.446
0.45	250	19	0.937	0.182±0.160	0.304±0.091	0.260±0.088	0.000±0.000	0.000±0.000	50.494±0.498
0.45	500	15	1.066	0.182±0.154	0.308±0.065	0.273±0.063	0.000±0.000	0.000±0.000	43.259±0.492
0.45	1000	14	1.115	0.184±0.153	0.310±0.046	0.290±0.045	0.000±0.000	0.000±0.000	43.246±0.488
0.50	100	38	0.656	0.134±0.175	0.240±0.147	0.207±0.147	0.010±0.100	0.000±0.000	71.883±0.422
0.50	250	19	0.956	0.135±0.147	0.253±0.094	0.214±0.094	0.000±0.000	0.000±0.000	55.855±0.531
0.50	500	15	1.096	0.137±0.141	0.258±0.067	0.228±0.068	0.000±0.000	0.000±0.000	45.565±0.463
0.50	1000	14	1.144	0.139±0.142	0.260±0.047	0.244±0.048	0.000±0.000	0.000±0.000	43.703±0.533
0.55	100	37	0.666	0.082±0.165	0.188±0.153	0.158±0.159	0.040±0.197	0.010±0.100	74.902±0.450
0.55	250	19	0.944	0.093±0.132	0.202±0.096	0.164±0.102	0.000±0.000	0.000±0.000	60.608±0.467
0.55	500	14	1.117	0.097±0.123	0.208±0.068	0.175±0.074	0.000±0.000	0.000±0.000	47.755±0.500
0.55	1000	13	1.189	0.100±0.123	0.211±0.048	0.189±0.053	0.000±0.000	0.000±0.000	41.146±0.481
0.60	100	37	0.662	0.033±0.154	0.136±0.157	0.100±0.173	0.050±0.219	0.000±0.000	77.227±0.396
0.60	250	19	0.936	0.051±0.117	0.152±0.098	0.105±0.112	0.000±0.000	0.000±0.000	65.064±0.416
0.60	500	14	1.118	0.059±0.106	0.159±0.069	0.113±0.082	0.000±0.000	0.000±0.000	50.038±0.484
0.60	1000	12	1.215	0.064±0.103	0.162±0.049	0.126±0.058	0.000±0.000	0.000±0.000	37.914±0.463

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

10) Çok deęişkenli normal dağılımdan ortalaması **0** ve varyans kovaryans matrisi

$$\begin{pmatrix} 1 & 0.3 & 0.1 \\ 0.3 & 1 & 0.3 \\ 0.1 & 0.3 & 1 \end{pmatrix}$$

olan, 3 açıklayıcı (1 kategorik, 2 sürekli), 1 yanıt deęişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betaların ise $\beta=(0.3, 0.5, 0.5)$ olduęu senaryoya ait bulgular Çizelge 8.11 ve 8.12’de sunulmuştur.

Çizelge 8.11. Senaryo 10 için tahmin modellerinden hesaplanan net fayda deęerleri ve karşılaştırma sonuçları ($0.10 \leq p_t \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±sd)	NF _{LR} (ort±sd)	NF _{KA} (ort±sd)	ANOVA** (ort±sd)	KW** (ort±sd)	z** (ort±sd)
0.10	100	1571	0.100	0.579±0.103	0.589±0.096	0.586±0.098	100.000±0.000	100.000±0.000	94.693±0.191
0.10	250	265	0.244	0.574±0.070	0.590±0.061	0.586±0.062	92.080±2.419	97.960±1.325	83.386±0.397
0.10	500	88	0.425	0.569±0.055	0.590±0.043	0.588±0.043	6.550±2.271	27.000±4.752	68.674±0.500
0.10	1000	44	0.609	0.566±0.047	0.590±0.030	0.589±0.031	0.000±0.000	0.340±0.572	60.786±0.472
0.15	100	475	0.182	0.543±0.117	0.563±0.102	0.555±0.104	98.690±1.169	99.730±0.510	87.829±0.326
0.15	250	84	0.435	0.531±0.090	0.565±0.064	0.557±0.066	7.290±2.289	32.410±4.621	69.180±0.357
0.15	500	35	0.683	0.521±0.081	0.566±0.046	0.559±0.046	0.000±0.000	0.550±0.845	56.126±0.478
0.15	1000	22	0.882	0.514±0.077	0.566±0.032	0.564±0.032	0.000±0.000	0.040±0.197	49.205±0.468
0.20	100	27	0.786	0.265±0.212	0.403±0.129	0.370±0.128	0.000±0.000	0.013±0.115	54.246±0.424
0.20	250	44	0.607	0.482±0.114	0.539±0.068	0.526±0.070	0.040±0.197	4.140±2.045	60.079±0.528
0.20	500	24	0.830	0.469±0.111	0.540±0.048	0.530±0.049	0.000±0.000	0.130±0.393	49.560±0.535
0.20	1000	18	0.970	0.461±0.110	0.540±0.034	0.536±0.035	0.000±0.000	0.030±0.171	44.515±0.534
0.25	100	99	0.402	0.450±0.157	0.505±0.113	0.488±0.116	17.000±3.932	50.800±4.748	73.656±0.420
0.25	250	30	0.739	0.428±0.140	0.510±0.071	0.490±0.073	0.000±0.000	0.350±0.609	53.958±0.486
0.25	500	20	0.917	0.415±0.139	0.511±0.051	0.495±0.052	0.000±0.000	0.020±0.141	46.175±0.528
0.25	1000	17	1.005	0.409±0.139	0.511±0.036	0.503±0.037	0.000±0.000	0.010±0.100	45.026±0.522
0.30	100	54	0.548	0.389±0.182	0.473±0.118	0.445±0.120	0.653±0.814	8.267±2.863	66.805±0.444
0.30	250	23	0.846	0.370±0.166	0.479±0.075	0.449±0.076	0.000±0.000	0.000±0.000	49.649±0.512
0.30	500	18	0.968	0.362±0.164	0.480±0.053	0.455±0.054	0.000±0.000	0.000±0.000	43.605±0.471
0.30	1000	16	1.025	0.359±0.164	0.481±0.038	0.467±0.039	0.000±0.000	0.000±0.000	43.368±0.519

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 8.12. Senaryo 10 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.35	100	34	0.694	0.323±0.202	0.439±0.123	0.406±0.123	0.000±0.000	0.173±0.381	61.390±0.440
0.35	250	20	0.917	0.313±0.188	0.445±0.078	0.411±0.078	0.000±0.000	0.000±0.000	47.607±0.547
0.35	500	17	1.000	0.311±0.184	0.447±0.056	0.418±0.055	0.000±0.000	0.000±0.000	44.071±0.501
0.35	1000	16	1.035	0.311±0.183	0.448±0.039	0.430±0.040	0.000±0.000	0.000±0.000	43.314±0.509
0.40	100	27	0.781	0.265±0.212	0.402±0.129	0.370±0.128	0.000±0.000	0.013±0.115	60.340±0.525
0.40	250	18	0.962	0.263±0.200	0.410±0.082	0.374±0.081	0.000±0.000	0.000±0.000	48.134±0.529
0.40	500	17	1.019	0.265±0.197	0.413±0.058	0.382±0.057	0.000±0.000	0.000±0.000	45.697±0.442
0.40	1000	16	1.038	0.267±0.196	0.414±0.041	0.395±0.041	0.000±0.000	0.000±0.000	46.180±0.553
0.45	100	25	0.826	0.216±0.215	0.364±0.134	0.330±0.134	0.000±0.000	0.013±0.112	61.582±0.480
0.45	250	18	0.981	0.219±0.205	0.373±0.085	0.337±0.085	0.000±0.000	0.000±0.000	48.997±0.516
0.45	500	17	1.018	0.223±0.204	0.376±0.060	0.348±0.060	0.000±0.000	0.000±0.000	47.562±0.465
0.45	1000	16	1.050	0.224±0.203	0.378±0.043	0.362±0.043	0.000±0.000	0.000±0.000	49.059±0.609
0.50	100	24	0.841	0.174±0.212	0.325±0.140	0.292±0.143	0.000±0.000	0.000±0.000	63.193±0.411
0.50	250	18	0.979	0.181±0.204	0.335±0.089	0.300±0.091	0.000±0.000	0.000±0.000	51.246±0.530
0.50	500	17	1.016	0.184±0.205	0.338±0.063	0.312±0.065	0.000±0.000	0.000±0.000	49.335±0.476
0.50	1000	16	1.039	0.186±0.205	0.340±0.044	0.325±0.046	0.000±0.000	0.000±0.000	50.915±0.458
0.55	100	25	0.821	0.136±0.207	0.283±0.146	0.250±0.155	0.000±0.000	0.000±0.000	66.339±0.430
0.55	250	19	0.959	0.147±0.198	0.295±0.092	0.258±0.099	0.000±0.000	0.000±0.000	52.267±0.494
0.55	500	17	1.002	0.150±0.200	0.299±0.065	0.269±0.071	0.000±0.000	0.000±0.000	49.482±0.536
0.55	1000	16	1.036	0.150±0.201	0.301±0.046	0.282±0.050	0.000±0.000	0.000±0.000	49.746±0.471
0.60	100	27	0.792	0.100±0.200	0.241±0.153	0.201±0.168	0.000±0.000	0.000±0.000	69.366±0.528
0.60	250	20	0.934	0.116±0.186	0.254±0.095	0.210±0.108	0.000±0.000	0.000±0.000	54.599±0.484
0.60	500	17	0.992	0.119±0.188	0.259±0.067	0.218±0.078	0.000±0.000	0.000±0.000	48.657±0.454
0.60	1000	16	1.026	0.119±0.190	0.261±0.047	0.230±0.056	0.000±0.000	0.000±0.000	46.336±0.522

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

11) Çok deęişkenli normal dağılımdan ortalaması **0** ve varyans kovaryans matrisi *diag(1,1,1,1)* olan, 3 açıklayıcı (2 kategorik, 2 sürekli), 1 yanıt deęişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betanın ise $\beta=(0.1, 0.1, 0.1, 1)$ olduęu senaryoya ait bulgular Çizelge 8.13 ve 8.14’de sunulmuştur.

Çizelge 8.13. Senaryo 11 için tahmin modellerinden hesaplanan net fayda deęerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.10	100	247	0.253	0.449±0.108	0.475±0.097	0.467±0.099	90.390±2.681	91.920±2.411	66.543±0.502
0.10	250	49	0.573	0.438±0.080	0.479±0.062	0.470±0.063	0.040±0.197	0.420±0.638	33.031±0.438
0.10	500	24	0.829	0.432±0.069	0.480±0.044	0.473±0.045	0.000±0.000	0.000±0.000	22.229±0.377
0.10	1000	17	1.014	0.429±0.064	0.480±0.031	0.475±0.032	0.000±0.000	0.000±0.000	25.592±0.483
0.15	100	94	0.412	0.402±0.122	0.448±0.100	0.432±0.104	13.570±3.833	23.810±4.074	57.416±0.545
0.15	250	27	0.786	0.387±0.100	0.453±0.064	0.435±0.066	0.000±0.000	0.000±0.000	33.482±0.409
0.15	500	17	1.022	0.380±0.093	0.455±0.046	0.441±0.047	0.000±0.000	0.000±0.000	29.721±0.508
0.15	1000	13	1.151	0.377±0.090	0.455±0.033	0.445±0.034	0.000±0.000	0.000±0.000	25.362±0.443
0.20	100	49	0.575	0.352±0.138	0.422±0.103	0.399±0.106	0.070±0.256	0.620±0.749	53.419±0.523
0.20	250	19	0.944	0.336±0.121	0.428±0.066	0.405±0.068	0.000±0.000	0.000±0.000	37.172±0.501
0.20	500	14	1.125	0.329±0.118	0.430±0.047	0.410±0.049	0.000±0.000	0.000±0.000	32.883±0.506
0.20	1000	12	1.228	0.326±0.116	0.431±0.034	0.412±0.036	0.000±0.000	0.000±0.000	24.390±0.456
0.25	100	32	0.718	0.300±0.156	0.396±0.107	0.370±0.108	0.000±0.000	0.010±0.100	52.657±0.529
0.25	250	16	1.048	0.285±0.144	0.403±0.068	0.378±0.069	0.000±0.000	0.000±0.000	40.023±0.494
0.25	500	13	1.177	0.280±0.142	0.405±0.049	0.381±0.050	0.000±0.000	0.000±0.000	35.760±0.446
0.25	1000	11	1.256	0.277±0.141	0.406±0.035	0.381±0.038	0.000±0.000	0.000±0.000	26.188±0.462
0.30	100	23	0.848	0.246±0.175	0.370±0.110	0.344±0.109	0.000±0.000	0.000±0.000	52.664±0.471
0.30	250	14	1.109	0.235±0.167	0.377±0.070	0.349±0.070	0.000±0.000	0.000±0.000	41.494±0.455
0.30	500	12	1.222	0.231±0.165	0.380±0.050	0.354±0.050	0.000±0.000	0.000±0.000	37.983±0.503
0.30	1000	11	1.281	0.228±0.165	0.381±0.036	0.357±0.038	0.000±0.000	0.000±0.000	32.478±0.544

pt: Eşik olasılık, n:Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 8.14. Senaryo 11 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.35	100	19	0.944	0.192±0.196	0.343±0.113	0.320±0.112	0.000±0.000	0.000±0.000	51.959±0.496
0.35	250	13	1.161	0.185±0.189	0.351±0.072	0.322±0.071	0.000±0.000	0.000±0.000	42.235±0.458
0.35	500	12	1.241	0.183±0.188	0.354±0.051	0.329±0.051	0.000±0.000	0.000±0.000	40.249±0.570
0.35	1000	11	1.292	0.181±0.187	0.355±0.036	0.336±0.037	0.000±0.000	0.000±0.000	38.184±0.559
0.40	100	17	1.017	0.140±0.217	0.317±0.116	0.295±0.116	0.000±0.000	0.000±0.000	52.025±0.501
0.40	250	13	1.190	0.136±0.212	0.325±0.074	0.296±0.074	0.000±0.000	0.000±0.000	43.199±0.431
0.40	500	11	1.260	0.135±0.210	0.328±0.053	0.305±0.052	0.000±0.000	0.000±0.000	41.841±0.473
0.40	1000	11	1.299	0.135±0.209	0.330±0.037	0.316±0.037	0.000±0.000	0.000±0.000	41.556±0.529
0.45	100	16	1.058	0.089±0.239	0.289±0.120	0.266±0.122	0.000±0.000	0.000±0.000	52.568±0.542
0.45	250	12	1.213	0.088±0.234	0.299±0.076	0.269±0.077	0.000±0.000	0.000±0.000	43.669±0.489
0.45	500	11	1.271	0.088±0.232	0.302±0.054	0.283±0.054	0.000±0.000	0.000±0.000	43.160±0.494
0.45	1000	11	1.299	0.089±0.231	0.304±0.038	0.294±0.038	0.000±0.000	0.000±0.000	44.025±0.523
0.50	100	15	1.092	0.041±0.258	0.262±0.124	0.236±0.129	0.000±0.000	0.000±0.000	52.850±0.491
0.50	250	12	1.235	0.041±0.254	0.273±0.078	0.242±0.082	0.000±0.000	0.000±0.000	44.538±0.532
0.50	500	11	1.278	0.042±0.253	0.276±0.055	0.257±0.058	0.000±0.000	0.000±0.000	43.845±0.519
0.50	1000	11	1.309	0.042±0.252	0.278±0.039	0.270±0.040	0.000±0.000	0.000±0.000	45.003±0.475
0.55	100	14	1.110	-0.004±0.275	0.234±0.128	0.206±0.139	0.000±0.000	0.000±0.000	53.444±0.388
0.55	250	12	1.239	-0.004±0.274	0.246±0.080	0.210±0.089	0.000±0.000	0.000±0.000	43.953±0.532
0.55	500	11	1.288	-0.004±0.273	0.250±0.057	0.226±0.063	0.000±0.000	0.000±0.000	43.232±0.435
0.55	1000	11	1.307	-0.003±0.273	0.252±0.040	0.240±0.044	0.000±0.000	0.000±0.000	44.157±0.518
0.60	100	14	1.104	-0.044±0.290	0.205±0.133	0.174±0.152	0.000±0.000	0.000±0.000	53.596±0.451
0.60	250	12	1.232	-0.047±0.293	0.218±0.082	0.175±0.097	0.000±0.000	0.000±0.000	43.672±0.499
0.60	500	11	1.288	-0.049±0.293	0.223±0.058	0.189±0.070	0.000±0.000	0.000±0.000	41.838±0.529
0.60	1000	11	1.310	-0.050±0.294	0.225±0.041	0.205±0.049	0.000±0.000	0.000±0.000	41.944±0.477

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

12) Çok deęişkenli normal dağılımdan ortalaması **0** ve varyans kovaryans matrisi *diag(1,1,1,1)* olan, 3 açıklayıcı (2 kategorik, 2 sürekli), 1 yanıt deęişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betanın ise $\beta=(0.1, 0.3, 0.3, 1)$ olduęu senaryoya ait bulgular Çizelge 8.15 ve 8.16’da sunulmuştur.

Çizelge 8.15. Senaryo 12 için tahmin modellerinden hesaplanan net fayda deęerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.10	100	200	0.281	0.446±0.109	0.475±0.097	0.465±0.099	91.770±2.704	93.870±2.177	48.070±0.412
0.10	250	38	0.652	0.458±0.089	0.508±0.062	0.498±0.064	0.000±0.000	0.020±0.141	32.169±0.396
0.10	500	21	0.890	0.451±0.081	0.509±0.044	0.501±0.045	0.000±0.000	0.000±0.000	24.598±0.383
0.10	1000	16	1.056	0.447±0.077	0.509±0.031	0.504±0.032	0.000±0.000	0.000±0.000	27.755±0.424
0.15	100	67	0.490	0.423±0.130	0.480±0.101	0.461±0.104	2.350±1.452	8.200±2.570	56.098±0.499
0.15	250	23	0.852	0.406±0.112	0.484±0.065	0.464±0.067	0.000±0.000	0.000±0.000	34.144±0.441
0.15	500	14	1.115	0.350±0.134	0.462±0.047	0.440±0.049	0.000±0.000	0.000±0.000	34.617±0.489
0.15	1000	14	1.140	0.396±0.107	0.486±0.032	0.475±0.033	0.000±0.000	0.000±0.000	28.697±0.478
0.20	100	41	0.633	0.373±0.148	0.454±0.104	0.428±0.107	0.010±0.100	0.120±0.327	51.711±0.551
0.20	250	18	0.970	0.356±0.136	0.460±0.067	0.434±0.068	0.000±0.000	0.000±0.000	38.314±0.511
0.20	500	13	1.147	0.303±0.159	0.438±0.049	0.412±0.050	0.000±0.000	0.000±0.000	36.681±0.421
0.20	1000	13	1.187	0.347±0.134	0.463±0.034	0.446±0.035	0.000±0.000	0.000±0.000	29.422±0.497
0.25	100	29	0.760	0.322±0.168	0.429±0.107	0.399±0.109	0.000±0.000	0.000±0.000	51.096±0.451
0.25	250	16	1.039	0.308±0.160	0.436±0.069	0.407±0.069	0.000±0.000	0.000±0.000	41.135±0.502
0.25	500	13	1.147	0.303±0.159	0.438±0.049	0.412±0.050	0.000±0.000	0.000±0.000	36.681±0.421
0.25	1000	12	1.206	0.301±0.158	0.439±0.035	0.418±0.037	0.000±0.000	0.000±0.000	31.108±0.453
0.30	100	23	0.862	0.271±0.188	0.404±0.111	0.372±0.111	0.000±0.000	0.000±0.000	51.454±0.475
0.30	250	27	0.790	0.507±0.284	0.669±0.058	0.653±0.059	0.000±0.000	6.160±2.112	62.568±0.482
0.30	500	26	0.803	0.507±0.284	0.670±0.041	0.655±0.042	0.000±0.000	1.060±1.043	59.297±0.483
0.30	1000	12	1.226	0.255±0.181	0.415±0.036	0.392±0.037	0.000±0.000	0.000±0.000	33.876±0.489

pt: Eşik olasılık, n:Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 8.16. Senaryo 12 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	z** (ort±ss)
0.35	100	19	0.945	0.220±0.209	0.379±0.114	0.349±0.114	0.000±0.000	0.000±0.000	51.234±0.482
0.35	250	14	1.123	0.214±0.204	0.386±0.073	0.354±0.072	0.000±0.000	0.000±0.000	43.398±0.462
0.35	500	13	1.195	0.212±0.203	0.389±0.052	0.360±0.051	0.000±0.000	0.000±0.000	40.003±0.498
0.35	1000	12	1.233	0.211±0.202	0.390±0.037	0.368±0.037	0.000±0.000	0.000±0.000	36.872±0.512
0.40	100	17	0.996	0.170±0.230	0.352±0.118	0.323±0.118	0.000±0.000	0.000±0.000	51.772±0.588
0.40	250	14	1.145	0.169±0.225	0.361±0.075	0.328±0.075	0.000±0.000	0.000±0.000	44.317±0.522
0.40	500	12	1.203	0.169±0.223	0.364±0.053	0.336±0.053	0.000±0.000	0.000±0.000	41.699±0.466
0.40	1000	12	1.232	0.168±0.223	0.365±0.038	0.345±0.037	0.000±0.000	0.000±0.000	39.622±0.540
0.45	100	16	1.030	0.124±0.249	0.326±0.122	0.294±0.123	0.000±0.000	0.000±0.000	52.686±0.499
0.45	250	12	1.215	0.125±0.243	0.339±0.055	0.312±0.055	0.000±0.000	0.000±0.000	43.199±0.487
0.45	500	12	1.215	0.125±0.243	0.339±0.055	0.312±0.055	0.000±0.000	0.000±0.000	43.199±0.487
0.45	1000	12	1.235	0.126±0.242	0.340±0.039	0.323±0.038	0.000±0.000	0.000±0.000	42.442±0.479
0.50	100	16	1.052	0.079±0.266	0.298±0.126	0.264±0.131	0.000±0.000	0.000±0.000	52.698±0.541
0.50	250	13	1.172	0.082±0.262	0.309±0.080	0.273±0.083	0.000±0.000	0.000±0.000	45.768±0.557
0.50	500	12	1.219	0.083±0.261	0.313±0.056	0.287±0.058	0.000±0.000	0.000±0.000	44.012±0.566
0.50	1000	12	1.242	0.083±0.260	0.314±0.040	0.297±0.041	0.000±0.000	0.000±0.000	43.628±0.473
0.55	100	16	1.052	0.040±0.280	0.270±0.131	0.233±0.141	0.000±0.000	0.000±0.000	54.127±0.435
0.55	250	13	1.172	0.041±0.279	0.282±0.082	0.241±0.090	0.000±0.000	0.000±0.000	45.427±0.521
0.55	500	12	1.216	0.041±0.279	0.286±0.058	0.255±0.064	0.000±0.000	0.000±0.000	43.551±0.458
0.55	1000	12	1.243	0.041±0.278	0.288±0.041	0.266±0.045	0.000±0.000	0.000±0.000	42.376±0.499
0.60	100	16	1.041	0.003±0.292	0.240±0.136	0.199±0.153	0.000±0.000	0.000±0.000	54.714±0.514
0.60	250	13	1.170	0.000±0.295	0.254±0.085	0.206±0.099	0.000±0.000	0.000±0.000	45.138±0.480
0.60	500	12	1.208	0.001±0.296	0.259±0.060	0.217±0.071	0.000±0.000	0.000±0.000	42.119±0.474
0.60	1000	12	1.225	0.002±0.296	0.261±0.042	0.230±0.049	0.000±0.000	0.000±0.000	39.486±0.460

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

13) Çok deęişkenli normal dağılımdan ortalaması **0** ve varyans kovaryans matrisi $diag(1,1,1,1)$ olan, 3 açıklayıcı (2 kategorik, 2 sürekli), 1 yanıt deęişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betanın ise $\beta=(0.1, 0.3, 0.5, 1)$ olduęu senaryoya ait bulgular Çizelge 8.17 ve 8.18’de sunulmuştur.

Çizelge 8.17. Senaryo 13 için tahmin modellerinden hesaplanan net fayda deęerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.10	100	114	0.374	0.463±0.116	0.503±0.097	0.492±0.100	28.790±4.459	43.590±4.751	61.020±0.479
0.10	250	31	0.728	0.449±0.094	0.507±0.062	0.495±0.064	0.000±0.000	0.000±0.000	28.973±0.432
0.10	500	19	0.949	0.442±0.088	0.508±0.044	0.498±0.045	0.000±0.000	0.000±0.000	24.578±0.422
0.10	1000	15	1.094	0.438±0.085	0.508±0.031	0.502±0.032	0.000±0.000	0.000±0.000	25.842±0.426
0.15	100	53	0.550	0.414±0.134	0.479±0.100	0.457±0.104	0.250±0.458	1.990±1.374	52.198±0.471
0.15	250	21	0.900	0.398±0.119	0.484±0.064	0.461±0.066	0.000±0.000	0.000±0.000	33.416±0.493
0.15	500	15	1.077	0.390±0.116	0.485±0.046	0.466±0.047	0.000±0.000	0.000±0.000	29.986±0.466
0.15	1000	13	1.173	0.387±0.115	0.486±0.032	0.471±0.033	0.000±0.000	0.000±0.000	23.662±0.432
0.20	100	35	0.688	0.365±0.153	0.455±0.104	0.424±0.107	0.000±0.000	0.010±0.100	49.308±0.408
0.20	250	17	1.015	0.348±0.143	0.461±0.066	0.431±0.068	0.000±0.000	0.000±0.000	37.620±0.484
0.20	500	14	1.135	0.343±0.142	0.463±0.047	0.436±0.048	0.000±0.000	0.000±0.000	32.126±0.447
0.20	1000	13	1.193	0.340±0.142	0.463±0.033	0.441±0.035	0.000±0.000	0.000±0.000	23.671±0.407
0.25	100	26	0.806	0.315±0.173	0.431±0.107	0.395±0.109	0.000±0.000	0.000±0.000	49.343±0.522
0.25	250	15	1.075	0.301±0.167	0.438±0.068	0.406±0.069	0.000±0.000	0.000±0.000	40.326±0.396
0.25	500	13	1.170	0.297±0.166	0.440±0.048	0.410±0.049	0.000±0.000	0.000±0.000	34.731±0.464
0.25	1000	12	1.227	0.294±0.166	0.441±0.034	0.415±0.036	0.000±0.000	0.000±0.000	26.696±0.455
0.30	100	21	0.903	0.264±0.195	0.407±0.110	0.369±0.111	0.000±0.000	0.000±0.000	49.676±0.492
0.30	250	14	1.111	0.255±0.190	0.414±0.070	0.380±0.070	0.000±0.000	0.000±0.000	41.606±0.515
0.30	500	13	1.194	0.252±0.189	0.417±0.050	0.385±0.050	0.000±0.000	0.000±0.000	36.412±0.437
0.30	1000	12	1.229	0.251±0.189	0.418±0.035	0.392±0.036	0.000±0.000	0.000±0.000	30.449±0.428

pt: Eşik olasılık, n:Örnek sayısı, d: Etki büyüklüęü

*: Etki büyüklüęüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 8.18. Senaryo 13 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.35	100	18	0.969	0.216±0.216	0.383±0.113	0.346±0.113	0.000±0.000	0.000±0.000	49.516±0.506
0.35	250	14	1.137	0.211±0.212	0.391±0.072	0.356±0.072	0.000±0.000	0.000±0.000	42.674±0.431
0.35	500	12	1.199	0.209±0.211	0.393±0.051	0.362±0.051	0.000±0.000	0.000±0.000	38.458±0.504
0.35	1000	12	1.241	0.208±0.210	0.395±0.036	0.370±0.037	0.000±0.000	0.000±0.000	33.841±0.470
0.40	100	17	1.022	0.167±0.237	0.358±0.117	0.321±0.118	0.000±0.000	0.000±0.000	49.993±0.525
0.40	250	13	1.163	0.166±0.233	0.367±0.074	0.331±0.074	0.000±0.000	0.000±0.000	43.448±0.457
0.40	500	12	1.212	0.166±0.232	0.370±0.053	0.339±0.052	0.000±0.000	0.000±0.000	40.037±0.466
0.40	1000	12	1.239	0.166±0.231	0.371±0.037	0.348±0.037	0.000±0.000	0.000±0.000	36.506±0.480
0.45	100	16	1.052	0.122±0.257	0.333±0.120	0.293±0.122	0.000±0.000	0.000±0.000	50.714±0.511
0.45	250	13	1.172	0.123±0.253	0.342±0.076	0.305±0.078	0.000±0.000	0.000±0.000	43.890±0.482
0.45	500	12	1.224	0.123±0.252	0.346±0.054	0.316±0.054	0.000±0.000	0.000±0.000	41.694±0.424
0.45	1000	12	1.248	0.123±0.251	0.347±0.038	0.325±0.038	0.000±0.000	0.000±0.000	39.234±0.571
0.50	100	15	1.069	0.078±0.276	0.307±0.125	0.264±0.130	0.000±0.000	0.000±0.000	50.834±0.495
0.50	250	13	1.183	0.080±0.272	0.317±0.079	0.278±0.083	0.000±0.000	0.000±0.000	44.875±0.534
0.50	500	12	1.227	0.081±0.271	0.321±0.056	0.291±0.058	0.000±0.000	0.000±0.000	42.466±0.483
0.50	1000	12	1.250	0.081±0.271	0.323±0.039	0.301±0.041	0.000±0.000	0.000±0.000	40.522±0.488
0.55	100	15	1.072	0.038±0.292	0.280±0.129	0.232±0.140	0.000±0.000	0.000±0.000	52.002±0.527
0.55	250	13	1.189	0.038±0.291	0.292±0.081	0.247±0.089	0.000±0.000	0.000±0.000	44.302±0.423
0.55	500	12	1.230	0.038±0.291	0.296±0.057	0.261±0.063	0.000±0.000	0.000±0.000	42.096±0.437
0.55	1000	12	1.252	0.038±0.291	0.298±0.040	0.271±0.044	0.000±0.000	0.000±0.000	39.268±0.501
0.60	100	15	1.061	0.001±0.306	0.252±0.135	0.197±0.152	0.000±0.000	0.000±0.000	52.381±0.524
0.60	250	13	1.166	0.001±0.309	0.265±0.084	0.212±0.098	0.000±0.000	0.000±0.000	43.980±0.444
0.60	500	13	1.188	0.005±0.310	0.270±0.059	0.224±0.070	0.000±0.000	0.000±0.000	40.523±0.477
0.60	1000	12	1.199	0.006±0.311	0.272±0.042	0.235±0.049	0.000±0.000	0.000±0.000	36.186±0.451

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

14) Çok deęişkenli normal dağılımdan ortalaması **0** ve varyans kovaryans matrisi

$$\begin{pmatrix} 1 & 0.3 & 0.1 & 0 \\ 0.3 & 1 & 0.3 & 0 \\ 0.1 & 0.3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

olan, 3 açıklayıcı (2 kategorik, 2 sürekli), 1 yanıt deęişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betaların ise $\beta=(0.1, 0.1, 0.1, 1)$ olduęu senaryoya ait bulgular Çizelge 8.19 ve 8.20’de sunulmuştur.

Çizelge 8.19. Senaryo 14 için tahmin modellerinden hesaplanan net fayda deęerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.10	100	247	0.253	0.449±0.108	0.475±0.097	0.467±0.099	90.350±3.016	92.610±2.678	66.371±0.459
0.10	250	49	0.573	0.438±0.080	0.479±0.062	0.470±0.063	0.040±0.197	0.460±0.731	32.881±0.496
0.10	500	25	0.821	0.432±0.070	0.480±0.044	0.473±0.045	0.000±0.000	0.000±0.000	22.270±0.472
0.10	1000	17	1.014	0.429±0.064	0.480±0.031	0.475±0.032	0.000±0.000	0.000±0.000	25.680±0.412
0.15	100	94	0.412	0.402±0.122	0.448±0.100	0.431±0.103	12.020±3.052	22.610±3.701	57.156±0.485
0.15	250	27	0.786	0.387±0.100	0.453±0.064	0.435±0.066	0.000±0.000	0.000±0.000	33.544±0.456
0.15	500	17	1.022	0.380±0.093	0.455±0.046	0.441±0.047	0.000±0.000	0.000±0.000	29.822±0.482
0.15	1000	13	1.173	0.376±0.091	0.456±0.032	0.445±0.034	0.000±0.000	0.000±0.000	25.355±0.456
0.20	100	49	0.575	0.352±0.138	0.422±0.103	0.399±0.106	0.090±0.288	0.560±0.729	53.341±0.510
0.20	250	19	0.948	0.335±0.122	0.428±0.066	0.405±0.068	0.000±0.000	0.000±0.000	37.260±0.484
0.20	500	14	1.125	0.329±0.118	0.430±0.047	0.410±0.049	0.000±0.000	0.000±0.000	32.917±0.523
0.20	1000	12	1.228	0.326±0.116	0.431±0.034	0.412±0.036	0.000±0.000	0.000±0.000	24.746±0.424
0.25	100	32	0.718	0.300±0.156	0.396±0.107	0.370±0.108	0.000±0.000	0.000±0.000	52.743±0.489
0.25	250	16	1.048	0.285±0.144	0.403±0.068	0.378±0.069	0.000±0.000	0.000±0.000	40.044±0.452
0.25	500	13	1.186	0.279±0.142	0.405±0.049	0.381±0.050	0.000±0.000	0.000±0.000	35.810±0.428
0.25	1000	11	1.307	-0.004±0.274	0.252±0.040	0.240±0.044	0.000±0.000	0.000±0.000	43.845±0.516
0.30	100	23	0.862	0.244±0.175	0.370±0.110	0.344±0.109	0.000±0.000	0.010±0.100	52.509±0.538
0.30	250	14	1.117	0.234±0.167	0.377±0.070	0.349±0.070	0.000±0.000	0.000±0.000	41.390±0.422
0.30	500	12	1.222	0.231±0.165	0.380±0.050	0.354±0.050	0.000±0.000	0.000±0.000	38.018±0.545
0.30	1000	11	1.274	0.229±0.165	0.381±0.035	0.357±0.038	0.000±0.000	0.000±0.000	32.621±0.443

pt: Eşik olasılık, n:Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 8.20. Senaryo 14 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.35	100	19	0.944	0.192± 0.196	0.343± 0.113	0.321± 0.112	0.000± 0.000	0.000± 0.000	51.792± 0.525
0.35	250	13	1.162	0.185± 0.190	0.352± 0.072	0.323± 0.071	0.000± 0.000	0.000± 0.000	42.406± 0.490
0.35	500	12	1.241	0.183± 0.188	0.354± 0.051	0.329± 0.051	0.000± 0.000	0.000± 0.000	40.093± 0.523
0.35	1000	11	1.293	0.181± 0.188	0.356± 0.036	0.336± 0.037	0.000± 0.000	0.000± 0.000	38.143± 0.554
0.40	100	17	1.019	0.139± 0.218	0.317± 0.116	0.295± 0.116	0.000± 0.000	0.000± 0.000	51.947± 0.531
0.40	250	12	1.197	0.136± 0.212	0.326± 0.074	0.296± 0.074	0.000± 0.000	0.000± 0.000	43.336± 0.492
0.40	500	11	1.267	0.135± 0.210	0.329± 0.053	0.305± 0.052	0.000± 0.000	0.000± 0.000	41.684± 0.476
0.40	1000	11	1.300	0.135± 0.210	0.331± 0.037	0.315± 0.037	0.000± 0.000	0.000± 0.000	41.396± 0.474
0.45	100	15	1.068	0.088± 0.239	0.290± 0.120	0.266± 0.121	0.000± 0.000	0.000± 0.000	52.536± 0.424
0.45	250	12	1.213	0.088± 0.234	0.299± 0.076	0.269± 0.077	0.000± 0.000	0.000± 0.000	43.770± 0.507
0.45	500	11	1.276	0.088± 0.232	0.303± 0.054	0.283± 0.054	0.000± 0.000	0.000± 0.000	43.092± 0.539
0.45	1000	11	1.305	0.088± 0.231	0.304± 0.038	0.294± 0.038	0.000± 0.000	0.000± 0.000	43.824± 0.462
0.50	100	15	1.088	0.041± 0.259	0.262± 0.124	0.236± 0.129	0.000± 0.000	0.000± 0.000	52.725± 0.499
0.50	250	12	1.225	0.042± 0.255	0.273± 0.078	0.242± 0.082	0.000± 0.000	0.000± 0.000	44.587± 0.521
0.50	500	11	1.284	0.042± 0.253	0.277± 0.055	0.257± 0.058	0.000± 0.000	0.000± 0.000	43.812± 0.462
0.50	1000	11	1.304	0.042± 0.253	0.278± 0.039	0.269± 0.040	0.000± 0.000	0.000± 0.000	44.828± 0.478
0.55	100	14	1.102	-0.003± 0.276	0.234± 0.128	0.206± 0.139	0.000± 0.000	0.000± 0.000	53.362± 0.473
0.55	250	12	1.234	-0.003± 0.274	0.246± 0.080	0.210± 0.089	0.000± 0.000	0.000± 0.000	44.058± 0.536
0.55	500	11	1.284	-0.004± 0.274	0.250± 0.057	0.226± 0.063	0.000± 0.000	0.000± 0.000	43.265± 0.458
0.55	1000	11	1.312	-0.004± 0.273	0.252± 0.040	0.240± 0.044	0.000± 0.000	0.000± 0.000	43.999± 0.522
0.60	100	15	1.099	-0.043± 0.290	0.205± 0.133	0.174± 0.152	0.000± 0.000	0.000± 0.000	53.593± 0.477
0.60	250	12	1.240	-0.048± 0.293	0.219± 0.083	0.175± 0.097	0.000± 0.000	0.000± 0.000	43.661± 0.524
0.60	500	11	1.284	-0.049± 0.294	0.223± 0.058	0.189± 0.070	0.000± 0.000	0.000± 0.000	41.877± 0.473
0.60	1000	11	1.315	-0.050± 0.294	0.226± 0.041	0.205± 0.049	0.000± 0.000	0.000± 0.000	41.791± 0.515

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

15) Çok deęişkenli normal dağılımdan ortalaması **0** ve varyans kovaryans matrisi

$$\begin{pmatrix} 1 & 0.3 & 0.1 & 0 \\ 0.3 & 1 & 0.3 & 0 \\ 0.1 & 0.3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

olan, 3 açıklayıcı (2 kategorik, 2 sürekli), 1 yanıt deęişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betaların ise $\beta=(0.1, 0.3, 0.3, 1)$ olduęu senaryoya ait bulgular Çizelge 8.21 ve 8.22’de sunulmuştur.

Çizelge 8.21. Senaryo 15 için tahmin modellerinden hesaplanan net fayda deęerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.10	100	152	0.323	0.470±0.113	0.504±0.097	0.495±0.100	53.210±4.648	65.400±4.913	65.290±0.525
0.10	250	36	0.673	0.456±0.090	0.508±0.062	0.497±0.064	0.000±0.000	0.040±0.197	31.586±0.534
0.10	500	20	0.912	0.449±0.082	0.509±0.044	0.501±0.045	0.000±0.000	0.000±0.000	24.622±0.440
0.10	1000	15	1.061	0.446±0.078	0.509±0.031	0.504±0.032	0.000±0.000	0.000±0.000	27.619±0.416
0.15	100	68	0.487	0.422±0.131	0.479±0.101	0.460±0.104	2.040±1.377	6.450±2.236	55.424±0.519
0.15	250	23	0.855	0.405±0.114	0.484±0.064	0.464±0.066	0.000±0.000	0.000±0.000	34.028±0.511
0.15	500	16	1.032	0.398±0.110	0.485±0.046	0.469±0.047	0.000±0.000	0.000±0.000	31.649±0.482
0.15	1000	15	1.067	0.445±0.079	0.509±0.031	0.504±0.032	0.000±0.000	0.000±0.000	20.625±0.447
0.20	100	39	0.646	0.372±0.149	0.455±0.104	0.427±0.107	0.030±0.171	0.150±0.359	51.291±0.534
0.20	250	18	0.968	0.355±0.138	0.460±0.067	0.434±0.068	0.000±0.000	0.000±0.000	38.177±0.511
0.20	500	14	1.111	0.349±0.136	0.462±0.047	0.439±0.049	0.000±0.000	0.000±0.000	34.064±0.395
0.20	1000	13	1.189	0.346±0.135	0.463±0.034	0.445±0.035	0.000±0.000	0.000±0.000	28.783±0.458
0.25	100	28	0.771	0.320±0.169	0.429±0.107	0.398±0.109	0.000±0.000	0.000±0.000	50.860±0.509
0.25	250	16	1.050	0.306±0.161	0.436±0.069	0.407±0.069	0.000±0.000	0.000±0.000	41.015±0.536
0.25	500	13	1.149	0.302±0.160	0.438±0.049	0.412±0.050	0.000±0.000	0.000±0.000	36.328±0.467
0.25	1000	12	1.210	0.299±0.160	0.439±0.034	0.417±0.036	0.000±0.000	0.000±0.000	30.516±0.475
0.30	100	22	0.874	0.269±0.190	0.405±0.111	0.371±0.110	0.000±0.000	0.000±0.000	51.132±0.496
0.30	250	15	1.097	0.259±0.184	0.412±0.071	0.380±0.070	0.000±0.000	0.000±0.000	42.534±0.472
0.30	500	13	1.178	0.256±0.183	0.414±0.050	0.385±0.050	0.000±0.000	0.000±0.000	37.946±0.463
0.30	1000	12	1.220	0.255±0.182	0.415±0.036	0.392±0.037	0.000±0.000	0.000±0.000	33.453±0.546

pt: Eşik olasılık, n:Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 8.22. Senaryo 15 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.35	100	19	0.947	0.219±0.210	0.379±0.114	0.348±0.113	0.000±0.000	0.000±0.000	51.091±0.572
0.35	250	14	1.126	0.213±0.206	0.387±0.073	0.355±0.072	0.000±0.000	0.000±0.000	43.314±0.504
0.35	500	13	1.196	0.212±0.204	0.390±0.052	0.360±0.051	0.000±0.000	0.000±0.000	39.799±0.473
0.35	1000	12	1.228	0.211±0.204	0.391±0.037	0.369±0.037	0.000±0.000	0.000±0.000	36.372±0.465
0.40	100	17	0.998	0.170±0.231	0.353±0.118	0.323±0.118	0.000±0.000	0.000±0.000	51.613±0.548
0.40	250	13	1.152	0.168±0.226	0.362±0.075	0.329±0.075	0.000±0.000	0.000±0.000	44.170±0.532
0.40	500	12	1.205	0.168±0.225	0.365±0.053	0.336±0.052	0.000±0.000	0.000±0.000	41.379±0.484
0.40	1000	12	1.239	0.168±0.224	0.367±0.038	0.346±0.037	0.000±0.000	0.000±0.000	39.078±0.547
0.45	100	16	1.034	0.124±0.250	0.327±0.121	0.294±0.123	0.000±0.000	0.000±0.000	52.321±0.492
0.45	250	13	1.158	0.125±0.246	0.336±0.077	0.302±0.078	0.000±0.000	0.000±0.000	44.712±0.512
0.45	500	12	1.217	0.125±0.244	0.340±0.054	0.313±0.055	0.000±0.000	0.000±0.000	43.089±0.451
0.45	1000	12	1.236	0.125±0.244	0.341±0.039	0.323±0.038	0.000±0.000	0.000±0.000	41.718±0.491
0.50	100	16	1.051	0.080±0.268	0.300±0.126	0.264±0.131	0.000±0.000	0.000±0.000	52.530±0.515
0.50	250	13	1.170	0.082±0.264	0.310±0.079	0.275±0.083	0.000±0.000	0.000±0.000	45.771±0.442
0.50	500	12	1.215	0.083±0.263	0.314±0.056	0.288±0.058	0.000±0.000	0.000±0.000	43.752±0.486
0.50	1000	12	1.249	0.082±0.262	0.316±0.040	0.298±0.041	0.000±0.000	0.000±0.000	42.954±0.491
0.55	100	15	1.061	0.039±0.282	0.272±0.130	0.233±0.141	0.000±0.000	0.000±0.000	53.698±0.465
0.55	250	13	1.179	0.040±0.281	0.284±0.082	0.243±0.090	0.000±0.000	0.000±0.000	45.148±0.434
0.55	500	12	1.217	0.041±0.281	0.288±0.058	0.256±0.064	0.000±0.000	0.000±0.000	43.460±0.485
0.55	1000	12	1.245	0.040±0.281	0.290±0.041	0.267±0.045	0.000±0.000	0.000±0.000	41.674±0.500
0.60	100	16	1.048	0.002±0.294	0.242±0.136	0.199±0.153	0.000±0.000	0.000±0.000	54.330±0.478
0.60	250	13	1.167	0.001±0.297	0.256±0.085	0.208±0.099	0.000±0.000	0.000±0.000	45.014±0.498
0.60	500	12	1.206	0.002±0.298	0.261±0.059	0.219±0.070	0.000±0.000	0.000±0.000	41.883±0.521
0.60	1000	12	1.222	0.002±0.299	0.263±0.042	0.231±0.049	0.000±0.000	0.000±0.000	38.756±0.466

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında 'Net Fayda bakımından fark yok'ların sayısı

16) Çok deęişkenli normal dağılımdan ortalaması **0** ve varyans kovaryans matrisi

$$\begin{pmatrix} 1 & 0.3 & 0.1 & 0 \\ 0.3 & 1 & 0.3 & 0 \\ 0.1 & 0.3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

olan, 3 açıklayıcı (2 kategorik, 2 sürekli), 1 yanıt deęişkeninden (kategorik) oluşan 100, 250, 500, 1000 gözlem içeren veri setlerinde simülasyon sayısı 100 ve betaların ise $\beta=(0.1, 0.3, 0.5, 1)$ olduęu senaryoya ait bulgular Çizelge 8.23 ve 8.24’de sunulmuştur.

Çizelge 8.23. Senaryo 16 için tahmin modellerinden hesaplanan net fayda deęerleri ve karşılaştırma sonuçları ($0.10 \leq pt \leq 0.30$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.10	100	104	0.391	0.461±0.117	0.503±0.097	0.491±0.100	22.150±4.547	36.180±4.768	59.609±0.444
0.10	250	29	0.755	0.446±0.096	0.507±0.062	0.494±0.064	0.000±0.000	0.010±0.100	28.300±0.471
0.10	500	18	0.974	0.439±0.090	0.508±0.044	0.498±0.045	0.000±0.000	0.000±0.000	24.536±0.459
0.10	1000	14	1.102	0.436±0.087	0.508±0.031	0.501±0.032	0.000±0.000	0.000±0.000	25.129±0.431
0.15	100	51	0.564	0.412±0.135	0.479±0.100	0.455±0.104	0.180±0.411	1.580±1.103	51.294±0.518
0.15	250	20	0.920	0.395±0.121	0.484±0.064	0.460±0.066	0.000±0.000	0.000±0.000	33.258±0.423
0.15	500	15	1.083	0.388±0.118	0.485±0.046	0.465±0.047	0.000±0.000	0.000±0.000	29.506±0.480
0.15	1000	13	1.178	0.385±0.117	0.486±0.032	0.470±0.033	0.000±0.000	0.000±0.000	22.682±0.398
0.20	100	33	0.707	0.362±0.155	0.455±0.103	0.423±0.106	0.000±0.000	0.010±0.100	48.611±0.496
0.20	250	17	1.021	0.346±0.145	0.461±0.066	0.431±0.067	0.000±0.000	0.000±0.000	37.436±0.457
0.20	500	14	1.139	0.341±0.144	0.463±0.047	0.436±0.048	0.000±0.000	0.000±0.000	31.805±0.521
0.20	1000	12	1.206	0.338±0.144	0.464±0.033	0.440±0.035	0.000±0.000	0.000±0.000	22.638±0.430
0.25	100	25	0.823	0.313±0.175	0.432±0.106	0.395±0.108	0.000±0.000	0.000±0.000	48.709±0.519
0.25	250	15	1.071	0.300±0.169	0.438±0.068	0.406±0.069	0.000±0.000	0.000±0.000	40.312±0.479
0.25	500	13	1.174	0.295±0.168	0.440±0.048	0.410±0.049	0.000±0.000	0.000±0.000	34.356±0.501
0.25	1000	12	1.221	0.293±0.168	0.441±0.034	0.414±0.036	0.000±0.000	0.000±0.000	26.012±0.457
0.30	100	20	0.914	0.263±0.196	0.408±0.109	0.369±0.110	0.000±0.000	0.000±0.000	49.332±0.505
0.30	250	14	1.114	0.254±0.192	0.415±0.070	0.381±0.070	0.000±0.000	0.000±0.000	41.521±0.574
0.30	500	13	1.196	0.251±0.191	0.418±0.050	0.386±0.050	0.000±0.000	0.000±0.000	36.237±0.445
0.30	1000	12	1.231	0.250±0.191	0.419±0.035	0.392±0.036	0.000±0.000	0.000±0.000	29.708±0.447

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

Çizelge 8.24. Senaryo 16 için tahmin modellerinden hesaplanan net fayda değerleri ve karşılaştırma sonuçları ($0.35 \leq pt \leq 0.60$)

pt	n	R*	d	NF _{DVM} (ort±ss)	NF _{LR} (ort±ss)	NF _{KA} (ort±ss)	ANOVA** (ort±ss)	KW** (ort±ss)	Z** (ort±ss)
0.35	100	18	0.973	0.215±0.218	0.384±0.113	0.346±0.113	0.000±0.000	0.000±0.000	49.221±0.490
0.35	250	13	1.151	0.209±0.213	0.392±0.072	0.357±0.072	0.000±0.000	0.000±0.000	42.534±0.450
0.35	500	12	1.207	0.208±0.213	0.395±0.051	0.363±0.051	0.000±0.000	0.000±0.000	38.215±0.480
0.35	1000	12	1.243	0.207±0.212	0.396±0.036	0.370±0.036	0.000±0.000	0.000±0.000	33.181±0.530
0.40	100	17	1.022	0.168±0.239	0.360±0.116	0.321±0.117	0.000±0.000	0.000±0.000	49.712±0.560
0.40	250	13	1.159	0.166±0.235	0.368±0.074	0.332±0.074	0.000±0.000	0.000±0.000	43.347±0.490
0.40	500	12	1.215	0.165±0.234	0.371±0.052	0.340±0.052	0.000±0.000	0.000±0.000	39.927±0.504
0.40	1000	12	1.247	0.165±0.233	0.373±0.037	0.348±0.037	0.000±0.000	0.000±0.000	35.919±0.489
0.45	100	16	1.055	0.122±0.259	0.335±0.120	0.293±0.122	0.000±0.000	0.000±0.000	50.376±0.537
0.45	250	13	1.175	0.123±0.255	0.344±0.076	0.307±0.077	0.000±0.000	0.000±0.000	43.803±0.441
0.45	500	12	1.220	0.123±0.254	0.347±0.054	0.318±0.054	0.000±0.000	0.000±0.000	41.622±0.482
0.45	1000	12	1.249	0.123±0.253	0.349±0.038	0.327±0.038	0.000±0.000	0.000±0.000	38.772±0.472
0.50	100	15	1.073	0.078±0.278	0.309±0.124	0.265±0.130	0.000±0.000	0.000±0.000	50.376±0.450
0.50	250	13	1.187	0.080±0.275	0.320±0.078	0.280±0.082	0.000±0.000	0.000±0.000	44.814±0.589
0.50	500	12	1.225	0.081±0.274	0.323±0.055	0.293±0.058	0.000±0.000	0.000±0.000	42.261±0.517
0.50	1000	11	1.256	0.080±0.273	0.325±0.039	0.302±0.041	0.000±0.000	0.000±0.000	40.126±0.510
0.55	100	15	1.081	0.037±0.295	0.283±0.129	0.233±0.140	0.000±0.000	0.000±0.000	51.498±0.504
0.55	250	13	1.192	0.037±0.294	0.294±0.081	0.249±0.089	0.000±0.000	0.000±0.000	44.180±0.505
0.55	500	12	1.228	0.038±0.294	0.298±0.057	0.263±0.063	0.000±0.000	0.000±0.000	41.943±0.467
0.55	1000	12	1.249	0.038±0.294	0.300±0.040	0.273±0.044	0.000±0.000	0.000±0.000	38.823±0.485
0.60	100	16	1.059	0.002±0.310	0.255±0.134	0.197±0.152	0.000±0.000	0.000±0.000	51.812±0.478
0.60	250	13	1.152	0.004±0.313	0.268±0.084	0.215±0.097	0.000±0.000	0.000±0.000	43.956±0.444
0.60	500	13	1.177	0.006±0.314	0.272±0.059	0.226±0.069	0.000±0.000	0.000±0.000	40.259±0.457
0.60	1000	12	1.201	0.006±0.314	0.275±0.041	0.237±0.049	0.000±0.000	0.000±0.000	35.728±0.446

pt: Eşik olasılık, n: Örnek sayısı, d: Etki büyüklüğü

*: Etki büyüklüğüne göre hesaplanan bootstrap sayısı

**: Simülasyon sonucu modeller arasında ‘Net Fayda bakımından fark yok’ların sayısı

ÖZGEÇMİŞ

Duygu KORKMAZ YALÇIN, İlk, Orta ve Lise eğitimini Adana’da, üniversite eğitimini ise 2014 yılında Fırat Üniversitesi İstatistik Bölümü’nde tamamladı. Van Yüzüncü Yıl Üniversitesi İstatistik ve Biyoistatistik Bölümlerindeki Yüksek Lisanslarını 2018 bahar döneminde tamamladıktan sonra aynı sene güz döneminde Çukurova Üniversitesi Biyoistatistik Anabilim Dalı’nda Doktora eğitimine başladı. 2019 Mart ayından beri Van Yüzüncü Yıl Üniversitesi Tıp Eğitimi ve Bilişimi Anabilim Dalı’nda araştırma görevlisi olarak çalışmaktadır ve evlidir.

