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

Risk Tahmin Modellerine Eklenen Yeni Faktörün Katkısının Değerlendirilmesinde Biyoistatistiksel Yaklaşımlar

Betül DAĞOĞLU

BİYOİSTATİSTİK TEZLİ YÜKSEK LİSANS PROGRAMI
YÜKSEK LİSANS TEZİ

DANIŞMANI
Prof. Dr. H. Refik BURGUT

ADANA-2016

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

Risk Tahmin Modellerine Eklenen Yeni Faktörün Katkısının Değerlendirilmesinde Biyoistatistiksel Yaklaşımlar

Betül DAĞOĞLU

BİYOİSTATİSTİK TEZLİ YÜKSEK LİSANS PROGRAMI

YÜKSEK LİSANS TEZİ

DANIŞMANI

Prof. Dr. H. Refik BURGUT

ADANA-2016

KABUL VE ONAY

Biyoistatistik Anabilim Dalı

Yüksek Lisans Programı Çerçevesinde yürütülmüş olan

"Risk Tahmin Modellerine Eklenen Yeni Faktörün Katkısının Değerlendirilmesinde Biyoistatistiksel Yaklaşımlar" adlı çalışma, aşağıdaki jüri tarafından **Yüksek Lisans Tezi** olarak kabul edilmiştir.

Tarihi: / / 2016

TEZ SINAV JÜRİSİ

Prof. Dr. H. Refik BURGUT
Çukurova Üniversitesi
Başkan

Prof. Dr. Z. Nazan ALPARSLAN
Çukurova Üniversitesi
Üye

Doç. Dr. Gülhan OREKİCİ TEMEL
Mersin Üniversitesi
Üye

Yukarıdaki Tez, Yönetim Kurulunun / / tarih ve sayılı kararı ile kabul edilmiştir.

Prof.Dr. Behice DURGUN

Sağlık Bilimleri Enstitü Müdürü

T.C. ÇUKUROVA ÜNİVERSİTESİ SAĞLIK BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS / DOKTORA TEZ ÇALIŞMASI
ETİK BEYANI

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 çalışmasında;

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

bildirir, aksi bir durumda bu konuda hakkımda yapılacak tüm yasal işlemleri ve aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim. 08/08/2016

Betül DAĞOĞLU

Kayıtlı olunan Program : Biyoistatistik Tezli YL
Tezin Konusu :Risk Tahmin Modellerine Eklenen Yeni Faktörün Katkısının Değerlendirilmesinde Biyoistatistiksel Yaklaşımlar

Tezin Türü : Yüksek Lisans : ☒ Doktora: ☐

Danışmanın Adı-Soyadı : Prof. Dr. H. Refik BURGUT
Danışmanın İletişim Bilgileri
Telefon : 0 (322) 338 60 84 - 3094
E-Posta : refik@cu.edu.tr

Öğrencinin İletişim Bilgileri
Telefon : 0 (507) 823 67 55
E-Posta : betuldagoglu@gmail.com
Adresi : Çukurova Üni. Tıp Fakültesi Biyoistatistik AD. ADANA

**Bu belgenin Lisansüstü eğitim tezleri savunmaya alınmadan önce öğrenci tarafından doldurulup imzalanarak Enstitü Müdürlüğüne teslim edilmesi gerekmektedir.*

TEŞEKKÜR

Yüksek lisans eğitimim ve tezim süresince benden ilgisini ve desteğini esirgemeyen değerli tez danışmanım Prof. Dr. H. Refik BURGUT'a**,** ilgilerinden dolayı Prof. Dr. Z. Nazan ALPARSLAN'a, Yrd. Doç. Dr. İlker ÜNAL'a ve Yrd. Doç. Dr. Yaşar SERTDEMİR'e, değerli katkılarından dolayı Mersin Üniversitesi öğretim üyesi Doç. Dr. Gülhan OREKİCİ TEMEL'e ve çalışmanın gerçekleşmesi için verilerin toplanmasına izin sağlayan Prof. Dr. Mehmet SATAR öncülüğünde Çocuk Sağlığı ve Hastalıkları Anabilim Dalı öğretim üyelerine; öğrenim hayatım boyunca destekleri ve anlayışlarından dolayı AİLEM'e içtenlikle teşekkür ederim.

Betül DAĞOĞLU

İÇİNDEKİLER

ÇİZELGELER DİZİNİ

ŞEKİLLER DİZİNİ

SİMGELER ve KISALTMALAR DİZİNİ

ROC	Reciever Operator Characteristics Curve - Alıcı işlem karakteristikleri eğrisi
NRI	Net Reclassification Improvement - Net yeniden sınıflama iyileştirmesi
IDI	Integrated Discrimination Improvement - Entegre ayırım iyileştirmesi
ROC-AUC	Area Under ROC – ROC eğrisi altında kalan alan
wNRI	weighted NRI - Ağırlıklandırılmış NRI
SNAP-II	Score for Neonatal Acute Physiology
SNAPPE-II	Score for Neonatal Acute Physiology- Perinatal Extension-II
SGA	Small for gestational age-Gestasyon Haftasına Göre Düşük Doğum Ağırlığı
MRD	Mean Risk Difference - Ortalama risk değerlerinin farkı
AARD	Above Average Risk Difference - Ortalama risk farkının üzerinde
RC	Toplam yeniden sınıflama oranı
RC-correct	Doğru yeniden sınıflama yüzdesi
CVD	Kardiyovasküler hastalıklar
NRI(>0)	Sürekli NRI
PEV	Açıklanan varyasyon oranı
NB	Net yarar
GA	Güven aralığı
HP	Hatalı pozitiflik
CV	Coefficient of Variation - Varyasyon Katsayısı
ΔAUC	ROC eğrisi altında kalan alan değişimi

ÖZET

Risk Tahmin Modellerine Eklenen Yeni Faktörün Katkısının Değerlendirilmesinde Biyoistatistiksel Yaklaşımlar

Giriş: Risk tahmin modelleri, hastalık ya da ölüm gibi klinik sonuçların tahmin edilmesinde kullanılan ve risk faktörlerini içeren bir regresyon tekniğidir. Uygulama alanındaki gelişmelerle ve teknolojik ilerlemelerle beraber yeni risk faktörlerinin bulunması kaçınılmaz bir durum haline gelmiştir. Bulunan bu yeni risk faktörlerinin modele ilavesi ve performans artışının değerlendirilmesinde merkezi bir rol üstlenen alıcı işlem karakteristikleri eğrisi (ROC) bazı sınırlamalar içermektedir. Bu sınırlamaların başında, tahmini regresyon katsayıları arasındaki değişimin ölçülmesinde sonuçların oldukça tutucu özelliğe sahip olması gelmektedir. Bundan dolayı yeni yöntem arayışları başlamış ve 2008 yılında Pencina ve arkadaşları tarafından yeniden sınıflama tablolarına dayanan ve yeni faktörün risk modeline katkısını açıklayan iki metot geliştirilmiştir. Bu iki yöntem net yeniden sınıflama iyileştirmesi (NRI) ve entegre ayırım iyileştirmesi (IDI)'dir.

Çalışmada, iç içe tasarıma sahip modellerin performansının değerlendirilmesine imkan sağlayan ΔAUC, NRI ve IDI tekniklerinin birbirlerine göre avantaj ve dezavantajlarını araştırmak amaçlanmıştır.

Gereç ve Yöntem: Çukurova Üniversitesi Tıp Fakültesi Balcalı Hastanesi'nin Yeni Doğan Yoğun Bakım Ünitesine, 2013-2015 Temmuz yılları arasında yatırılan 864 bebeğin ölüm olasılıklarını tahmin etmek için SNAP-II ve SNAPPE-II risk tahmin modelleri kullanılmıştır. Bebeğin sürfaktan ve annenin antenatal kortikosteroid (celestone) tedavisini alıp almaması yeni değişken olarak eklenmiştir. Simülasyon çalışmasında ise anlamsız değişken, uygun (fit) modele eklenmiş ve farklı örneklem büyüklüklerinde (100, 250, 500 ve 1000), farklı prevalans değerlerinde (0.10, 0.25 ve 0.50) 1000 tekrarlı simüle edilmiş verilerle yöntemlerin HP oranları belirlenmiştir.

Bulgular: SNAP-II risk tahmin modeline hem celestone hem de sürfaktan bilgisinin eklenmesi ile elde edilen modele bağlı performans artışı anlamlı bulunmuştur. Bununla beraber SNAPPE-II risk tahmin modeline celestone bilgisinin eklenmesi ile elde edilen performans artışı sadece NRI sürekli istatistiği tarafından anlamlı bulunmuştur. Simülasyon çalışması ile NRI sürekli istatistiğinin HP oranı en yüksek iken ΔAUC ve IDI istatistiğinin HP oranının en düşük olduğu belirlenmiştir. Bununla beraber NRI kategorik istatistiği düşük prevalanslarda yüksek HP oranına sahip iken prevalansın artışına bağlı olarak HP oranı azalmaktadır.

Sonuç: ΔAUC istatistiği oldukça tutucu sonuçlar vererek H_0 hipotezini kabul etmeye en yatkın yaklaşım olarak belirlenirken NRI sürekli istatistiği tam tersi şeklinde davranmaktadır. IDI istatistiğinin HP oranındaki düşüklük ve anlamlı değişken senaryosunda performans artışını anlamlı bulması göz önüne alındığında kullanımı en uygun olabilecek yöntem olarak belirlenmiştir.

Anahtar Kelimeler: risk tahmin modeli, ROC eğrisi altında kalan alan endeksi, net yeniden sınıflama iyileştirmesi (NRI), entegre ayırım iyileştirmesi (IDI)

ABSTRACT

Biostatistical Approaches to the Assessment of Contribution by a New Factor into Risk Prediction Models

Introduction: Risk prediction model is a regression method that is used for predicting clinical outcomes, such as disease or death, and including risk factors. Along with the developments in application fields and technological progress, discovery of new risk factors has become inevitable. Receiver Operating Characteristic (ROC) curve, that has a major role in including these new risk factors into the model and evaluating the performance of the model, has some limitations. The most important limitation of ROC is that the difference between predictive regression coefficients is highly conservative. To be used as a solution to this limitation, Pencina et al. (2008) proposed two methods which are based on reclassification tables and explaining the new factor's contribution to the risk model. These two methods are called the Net Reclassification Improvement (NRI) and Integrated Discrimination Improvement (IDI). In this study, the aim is to investigate the performance and compare the advantages and disadvantages of ΔAUC, NRI and IDI techniques that are nested models.

Material and Method: In order to predict the death probabilities of 864 babies hospitalized in the Newborn Intensive Care Unit in Çukurova University Medical Faculty Balcalı Hospital between the years 2013 and 2015, SNAP-II and SNAPPE-II risk prediction models are used. Surfactant treatment for the baby and corticosteroid (celestone) treatment for the mother are added into the model as a new variable. In the simulation study, the insignificant variable is added to the fitted model and then several datasets are generated using different sample sizes (100, 250, 500 and 1000) and different prevalence values (0.10, 0.25 and 0.50) with 1000 replications. These datasets are used to calculate False Positive Ratios (FPR) of the ΔAUC, NRI and IDI methods.

Results: Improvement in the performance of the model obtained by including both celestone and surfactant variables into the SNAP-II risk prediction model is decided to be significant. Nevertheless, Improvement in the performance of the model obtained by including celestone into the SNAPPE-II risk prediction model is founded significant only by continuous NRI. The simulation study showed that the FPR of the continuous NRI is the highest while FPR of the ΔAUC and IDI is the lowest. Moreover, it is concluded that categorical NRI has high FPR values at low prevalence levels and depending on the increase in prevalence values, FPR values tend to decrease.

Conclusion: Whereas ΔAUC statistic is decided to be the approach that has the most tendency to accept H_0 by yielding very conservative results, the continuous NRI behaves quite the opposite. IDI statistic is decided to be the most appropriate method since its FPR values are lower than the other two methods and it concludes that the performance increase is significant in the real dataset.

Keywords: risk prediction models, ROC curve, the net reclassification improvement, the integrated discrimination improvement

1. GİRİŞ

Risk değerlendirmesi, modern klinik ve koruyucu hekimlik uygulamalarında önemli bir rol oynamaktadır. Bir hastalığın oluşma riskini; bireyin yaşam şekli, genetik yatkınlığı ve yaşı gibi faktörler etkilemektedir. Risk tahmin modelleri ise hastalığın oluşma risklerini içeren regresyon tekniklerdir [1]. Risk tahmin modeli kurulduktan sonra bu modelin tahminlerde ne kadar iyi olduğunun ölçülmesi modelin performans ölçütlerine bağlıdır [2]. D'Agostino ve arkadaşları [3] ikili sonuç değişkenine sahip modellerin performansının değerlendirilmesi için kullanılan yöntemleri kapsamlı bir şekilde değerlendirmişlerdir. Bu performans değerlendirme yöntemleri genel olarak ayırım ve kalibrasyon yeteneğini ölçmeye yönelik yaklaşımlardır. Ayırım yeteneği, hastalığı olan ya da olması muhtemel olanların 'hasta', hastalığı olmayan ya da olması muhtemel olmayanları 'hasta değil' şeklinde sınıflandırılabilmesi esasına dayanır. Kalibrasyon ise modele dayalı tahmin değerlerinin gerçek değerlere ne kadar yakın olduğunun bir ölçütüdür [4]. Bunlara ek olarak genel performansın değerlendirilmesi için Nagelkerke R^2 ve Brier skoru model performansını değerlendirmek için kullanılan yaklaşımlar arasındadır.

Genel model performansının değerlendirilmesinde kullanılan yöntemlerin amacı tahmin edilen değer ile gerçek değer arasındaki farkın belirlenmesidir. Bununla beraber ayırım yeteneğini değerlendirmek için kullanılan alıcı-işlem karakteristikleri (Receiver-operating characteristic-ROC) eğrisi, hasta ve sağlam gruplarının risk dağılımları arasındaki farkın belirlenmesine yönelik bir yaklaşımdır. Kalibrasyon ise genel model performansını değerlendirmek için kullanılan yöntemlere paralel olan bir amaç ile kullanılmaktadır. Yani gözlem değerleri ile tahmin değerleri arasındaki farkın bir test prosedürüdür. Bu farkı test etmek için Hosmer-Lemeshow istatistiği kullanılmaktadır [2].

Süregelen gelişmelerle birlikte var olan risk tahmin modellerine yeni bir faktörün ya da faktörlerin ilavesi kaçınılmaz bir durum haline gelmiştir. Elde edilen yeni model ile eski modelin performansları arasındaki farkın değerlendirilmesinde kullanılan en önemli yaklaşımlardan biri ROC eğrisi altında kalan alan endeksi (ΔAUC)'dir. Ancak literatürde bu yaklaşımın kullanılması ile ilgili pek çok sakınca olduğu belirtilmiştir. Bu sakıncalardan en önemlisi DeLong test prosedürünün oldukça

tutucu sonuçlar vermesi [1] ve ROC eğrisinin, faktörler ya da markerler için orijinal ölçüm skalalarını içermemesidir [5]. Bu nedenlerden dolayı Pencina ve arkadaşları (2008) iki yeni yaklaşım geliştirmiştir [6]. Bunlar net yeniden sınıflama iyileştirmesi (Net Reclassification Improvement-NRI) ve entegre ayırım iyileştirmesi (Integrated Discrimination Improvement-IDI)'dir [7].

NRI istatistiği, yeniden sınıflama tablolarına dayanan bir tekniktir. Bu test yaklaşımı ile hastalığı olan ve olmayan bireyler, risk kategorilerine göre sınıflandırılır ve yeni modelle, bu kategorilerdeki değişimin istatistiksel olarak önemli olup olmadığı araştırılır. NRI istatistiği, kullanılan risk tahmin modelinin net bir şekilde belirlenmiş olan bir kesim noktasının olması varsayımına dayanır. Net bir kesim noktası ve buna bağlı olarak net bir risk kategori sınıflaması olmayan modellerin performanslarını değerlendirebilmek için kategori tanımlaması olmayan yeni bir NRI yaklaşımı geliştirilmiştir. Bu yaklaşım "NRI sürekli" olarak adlandırılmakta ve NRI(>0) şeklinde gösterilmektedir. NRI sürekli yaklaşımının en önemli avantajlarından biri de insidans değerlerinden etkilenmemesi ve farklı çalışmaların mukayese edilebilmesine olanak sağlamasıdır. İkinci olarak geliştirilen yaklaşım IDI istatistiğidir. Bu istatistik, duyarlılık ve $1-$ seçicilik için tüm olası kesim noktalarına göre integralinin alınmasına dayanan bir metottur. IDI istatistiği, mutlak hataların farkı ve R^2'deki değişim gibi regresyon yöntemlerinin elemanı olan tanımlamaları da içeren daha güvenilir bir çıkarsama tekniğidir [6].

Geliştirilen bu iki istatistik ve ΔAUC yaklaşımı klinik olarak yararı içeren tanımlamalar içermemektedir. Bu eksikliği gidermek adına Pencina ve arkadaşları (2010) NRI istatistiğine yeni bir bakış açısı getirmişlerdir. Bu yaklaşım, yeni model ile olayın olması durumu sınıflandırıldığında eski modele kıyasla daha yüksek risk gruplarında tanımlanmasından elde edilecek yarar ve yeni modelle olayın olmaması durumu daha düşük risk gruplarında sınıflandırıldığında elde edilecek yarar terimlerine bağlı ağırlıklandırma esasına dayanmaktadır. Bu yeni yaklaşım "ağırlıklandırılmış NRI (weighted NRI-wNRI)" olarak adlandırılmaktadır [7]. Ayrıca net yarar terimine ve kesim noktası olasılıklarına bağlı olarak çizilen "karar eğrisi analizi" ile klinik yarar değerlendirilebilmektedir.

Bu tez çalışmasının amacı;

- Risk tahmin modeline yeni bir faktörün eklenmesi ile elde edilecek model ile eski modelin performansını karşılaştırmak amacıyla kullanılan ROC-AUC, net yeniden sınıflama iyileştirmesi (NRI) ve entegre ayırım iyileştirmesi (IDI) yöntemlerinin birbirlerine göre avantaj ve dezavantajlarını belirlemek,
- Çalışmada kullanılacak olan SNAP-II (Score for Neonatal Acute Physiology) ve SNAPPE-II (Score for Neonatal Acute Physiology- Perinatal Extension-II) risk tahmin modellerine, antenatal kortikosteroid (celestone) ya da sürfaktan değişkeninin eklenmesi ile elde edilecek yeni modelin tahmindeki etkinliğini değerlendirmek,
- SNAP-II ve SNAPPE-II modellerini katsayılarını güncellemek

şeklindedir.

2. GENEL BİLGİLER

D ikili sonuç değişkeni olmak üzere D=1 olayın olması durumunu ifade ederken D=0 olayın olmaması durumunu ifade etmektedir. Regresyon tekniği, $x=x_1,x_2,\ldots,x_p$, olmak üzere p tane ortak değişken ile olayın olması durumunun tahminine dayanmaktadır. Bu modelleme tekniği ile olayın olması ihtimali üzerinde etkisi olan değişkenlerin birleştirilerek tek bir risk skorunun elde edilmesi sağlanır. Yüksek risk skoru, tıbbi açıdan bireyde olayın gelişmesi olasılığının ya da olayın bireyde bulunması olasılığının yüksek seviyelerde olduğunu göstermektedir [1]. Olayın olması ihtimali üzerinde etkisi olan değişkenlerin modele ilave edilmesinin ardından modelin performansının değerlendirilmesi ikinci bir adım olarak sınanmalıdır. Hali hazırda kullanılan risk tahmin modeline yeni bir faktörün ilavesi ile elde edilen yeni model ile eski modelin performansının değerlendirilmesi ROC eğrisi altında kalan alan endeksi, NRI ve IDI yöntemlerine dayanmaktadır.

2.1. Lojistik Regresyon

Tanı (hastalık var/yok) ve prognozda (mortalite, morbidite, komplikasyonlar) lojistik regresyon modeli sıklıkla karşılaşılan bir tekniktir. p tane bağımsız değişkenin vektör uzayı $x' = [x_1 \quad x_2 \ldots x_p]$ olarak tanımlansın. x verildiğinde olayın olması olasılığı $P(D=1|x)=\pi(x)$ şeklinde ifade edilebilir [8]. Lojistik fonksiyonu tanımlayan lojistik regresyon modeli [9],

$$P(D = 1 \,|X_1, X_2 \ldots, X_k) = \frac{1}{1+e^{-(\alpha+\sum \beta_i X_i)}} = \pi(X) \qquad (1)$$

şeklindedir. Bir olayın oddsının yani $\pi(X)/(1-\pi(X))$'in doğal logaritması alınırsa lojit dönüşüm gerçekleştirilmiş olur ve lojitler $-\infty$ ile ∞ arasında değer alabilir [10].

$$\text{lojit } P(D = 1 \,|X_1, X_2 \ldots, X_k) = \ln\left(\frac{\pi(X)}{1-\pi(X)}\right)$$

(1) eşitliğindeki $\pi(X)$ ifadesi lojit fonksiyonunda yerine yazılırsa model doğrusal modele dönüşür [2].

$$g(x) = \alpha + \sum \beta_i X_i$$

Ancak burada unutulmaması gereken lojistik regresyonda ilgilenilen konu yanıt değişkeninin gerçekleşme olasılığının logaritmasındaki değişimdir [11].

Cinsiyet, ırk, tedavi grubu gibi nominal skalalı değişkenler için kategori sayısı eksi bir kadar tasarım değişkeni ya da diğer bir ifadeyle kukla değişken oluşturulması gerekmektedir. j. nominal özelliğe sahip bağımsız değişken x_j olsun ve bu değişken k_j adet kategoriye sahip ise k_j-1 tane kukla değişken oluşturulmalıdır. Bu j. değişkenin tasarım değişkeni D_{jl} ve katsayı değerleri β_{jl} ile gösterilsin ($l = 1,2, \dots, k_j - 1$). p tane bağımsız değişkene bağlı modelin lojiti,

$$g(x) = \beta_0 + \beta_1 x_1 + \cdots + \sum_{l=1}^{k_j-1} \beta_{jl} D_{jl} + \beta_p x_p$$

şeklinde ifade edilmektedir [8].

Lojistik regresyonda regresyon katsayılarının kestirimi en çok olabilirlik yöntemine dayanırken lineer regresyonda en küçük kareler yöntemine ya da en çok olabilirlik yöntemine dayanmaktadır. Olabilirliğin doğal logaritması (log likelihood-LL),

$$LL = \sum_{i=1}^{n} [y_i \times \ln(\pi(x_i)) + (1 - y_i) \times \log(1 - \pi(x_i))]$$

şeklindedir [2]. Mükemmel uyuma sahip modellerde LL değerinin 0 olması beklenir. Ortalama tahminlerle modelin LL değeri şu şekilde hesaplanmaktadır.

$$LL_0 = \sum y \times \log(\text{ort}(y/n)) + (1 - y) \times \log(1 - \text{ort}(y/n)) \qquad (2)$$

Burada LL_0, Null modelin log olabilirlik fonksiyonu iken $\text{ort}(y/n)$, y iki kategorili sonuç değişkeninin ortalama olasılığıdır. Prognostik modelin, Null modeli ile karşılaştırılmasında olasılık oranı testi (LR) yaklaşımı kullanılmaktadır.

$$LR = -2(LL_0 - LL_1)$$

LL_1, belirleyicilerin dahil olduğu model log-olabilirlik fonksiyonu iken LL_0, sıfır (null) modelin log-olabilirlik fonksiyonudur ve LR χ^2 dağılımı göstermektedir. LR istatistiği tek değişkenli modellerin karşılaştırılması için kullanılabileceği gibi birden fazla tahmin edicinin bulunduğu modellerin karşılaştırılması içinde kullanılabilir [12].

Özet olarak, bir lojistik model oluşturulurken izlenecek yol sırasıyla aşağıdaki gibidir.

- Modele koyulması planlanan her bir değişken için tek değişkenli analiz yapılmalıdır. Nominal ya da ordinal nitelikli kategorik değişkenler için k-1 serbestlik dereceli olabilirlik oran ki-kare test sonucu, tek değişkenli lojistik modelde k-1 tasarım değişkeni ile katsayının önemliliğinde kullanılan olabilirlik oran testi ile neredeyse eşit sonuçlar vermektedir. Bununla beraber Pearson ki-kare testi asimptotik olarak olabilirlik oran ki-kare testi ile benzer sonuçlar verdiği için bu ki-kare test sonucu da kullanılabilir. Sürekli niteliğe sahip bir bağımsız değişkenin tek değişkenli olarak lojistik modele ilavesi ile elde edilen olabilirlik oran test sonucu değişkenin değerlendirilmesinde kullanılan yöntemlerdendir. Buna ek olarak Wald ve iki örneklem için t testleri de kullanılabilecek yaklaşımlardandır. Tek değişkenli analizlerde p değeri 0,250'nin altında olan değişkenler çok değişkenli analizlere dahil edilmelidir. Eğer eklenen bu değişken çok değişkenli lojistik modelde anlamsız olsa dahi klinik olarak anlamlı ise modelde bırakılmalıdır.
- Değişken seçiminde bir diğer yol ise adım adım (stepwise) metodudur. Bu metottun iki uzantısı bulunmaktadır. Bunlar, ileriye doğru seçim (forward selection) ve geriye doğru eleme (backward elimination)'dir. Bu seçim mekanizmaları kullanılarak da değişken seçimi yapılabilmektedir. Ancak burada kişisel müdahale söz konusu değildir. Bu nedenden dolayı adım adım yaklaşımı ile model oluşturulduktan sonra, modelde düzeltme faktörü olarak bulunması gereken değişkenler yok ise adım adım modeli ile belirlenen faktörler ve düzeltme faktörleri "enter" metottu ile modellenmelidir.
- Seçilen değişkenler ile çoklu lojistik regresyon modeli kurulmalı ve katsayıların anlamlılığı Wald test sonucu ile sınanmalıdır.

- Sürekli değişkenlerin lojit ile arasındaki doğrusallık sınanmalıdır. Bu sınama için grafiksel yaklaşımlar kullanılabilmektedir. Çizilen grafikler ile lojit ve sürekli değişken arasındaki ilişki; lineer, ikinci dereceden, lineer olmayan gibi bazı formlar şeklinde olabilir. Grafik yaklaşımına ek olarak kısmi polinom yaklaşımı da mevcuttur. Bu yaklaşımlarla sürekli değişkenin modelde bulunması gereken forma karar verilebilir.
- Ana etki modeli yaratıldıktan sonra ve sürekli değişkenin modele giriş şekli belirlendikten sonra her ikili değişken için etkileşim terimleri üretilmelidir. Öncelikle seçilen bu etkileşim terimlerinin klinik olarak bir anlamı olmalıdır. Ana etki modeline etkileşim terimleri eklenmeli ve olabilirlik oran testindeki anlamlılıkları incelenmelidir [8].

Lojistik model yukarıda anlatılan adımlar ışığında kurulduktan sonra bu modelin performansının sınanması bir diğer adımdır.

2.2. Model Performansının Değerlendirilmesi

Bir tahmin modeli kurulduktan sonra bu kurulan modelin tahminlerde ne kadar iyi olduğunun ölçülmesi modelin performans ölçütlerine bağlıdır. Performans değerlendirmesinde kullanılan geleneksel ölçüm yöntemleri; genel model performansının değerlendirilmesinde açıklanan varyasyon (R^2) katsayısı ve Brier skor, ayırım yeteneğinin testi için ROC eğrisi altında kalan alan değeri ve son olarak da kalibrasyon için uyum iyiliği istatistikleri kullanılmaktadır [2].

Son zamanlarda, model performansının değerlendirilmesi için yeni yaklaşımlar geliştirilmiştir. Bunlar modelin ayırım yeteneğini ve genel performansını değerlendirme esasına dayanır. Modelin ayırım yeteneğinin değerlendirilmesi için yeniden sınıflama tablolarına dayanan yeniden sınıflama istatistiği ve net yeniden sınıflama iyileştirmesi (NRI) kullanılırken, modelin genel performansının değerlendirilmesi için entegre ayırım iyileştirmesi (IDI) yöntemi kullanılmaktadır [13]. NRI ve IDI istatistikleri, risk tahmin modeline yeni bir faktör eklendiğinde elde edilecek ikinci modelin, riski tahmin etme açısından birinci modelden istatistiksel olarak daha fazla anlamlı olup olmadığını test eder. Bu iki istatistiğe ek olarak aynı amaca yönelik ROC eğrisi altında kalan alan değerlerinin farkının anlamlılığı kullanılan yöntemlerdendir.

Model performansını değerlendirirken klinik yararın da göz önüne alınabilmesi için "karar analizi" yaklaşımı geliştirilmiştir. Bu analiz ile net yarar değeri hesaplanmakta ve çizilen "karar eğrisi" ile farklı risk tahmin modellerini karşılaştırmak ve eşik değerinin bölgesini belirlemek mümkün hale gelmektedir [14].

Model performansını görsel olarak değerlendirmek için farklı grafikler kullanılmaktadır. Modelin genel performansının ve kalibrasyonunun değerlendirilmesi için geçerlilik grafiği, ayırım için ROC eğrisi ve boxplot, kalibrasyon için kalibrasyon eğrisi, NRI için scatter ve kliniksel yarar için ise karar eğrisi kullanılan grafiklerdendir.

Çizelge 2.2.1. Model Performansını Değerlendirmede Kullanılan Yöntemler, Yöntemlerin Görsel İçerikleri ve Tanımlamaları[13]

Değerlendirme	Kullanılan Yöntemler	Yöntemlerin Görsel İçeriği	Yöntemlerin Tanımı
Genel Performans	R^2, Brier	Geçerlilik (Validation) Grafiği	İyi bir model Y ve $\hat{Y}$ arasındaki farkın düşük olmasına bağlıdır. Modelin kalibrasyon ve ayırım yönleri ile ilgilenir.
Ayırım	Ortalama risk değerlerinin farkı (MRD) Above average risk difference, AARD c-istatistiği Ayırım Eğimi	ROC eğrisi Boxplot	Rank sıra istatistiklerine dayanır. Sonuç değişkenlerinin ortalamalarına bağlı grafiklerdir.
Kalibrasyon	Kalibrasyon eğrisi Hosmer-Lemeshow testi	Kalibrasyon ya da geçerlilik grafikleri	Ortalama y ve ortalama $\hat{y}$ bağlı karşılaştırmalardır. Gözlemlenen değerlerin tahmin edilen değerler ile ilişkisidir.
Yeniden Sınıflama	Yeniden Sınıflama İstatistiği Net Yeniden Sınıflama İyileştirmesi (NRI)	Çapraz Tablo ya da Scatter Grafiği	2 modelin performansının değerlendirilmesinde kullanılır.
	Entegre Ayırım İyileştirmesi (IDI)	İki model için çizilen Boxplotlar (Hasta ve sağlam gruplar için ayrı ayrı)	Muhtemel olan tüm kesim noktalarına göre duyarlılık ve 1-seçicilik için integral alınmasına dayanır.
Kliniksel Yarar	Net Yarar (NB) Karar eğri analizi (DCA) wNRI	Çapraz Tablo Karar eğrileri	Tek bir kesim noktasında modele bağlı olmayan doğru pozitiflikten elde edilen yarara dayanan yöntemdir.

2.2.1.Tek Bir Risk Tahmin Modelinin Performansının Değerlendirilmesi

2.2.1.1. Genel Performansın Değerlendirilmesi

Modelin genel performansının değerlendirilmesinde tahmin edilen değer ile gerçek değer arasındaki farkın incelenmesi merkezi bir role sahiptir. Sürekli ölçümlü çıktılara sahip modellerde bu değerlendirme $(Y - \widehat{Y})$ farkına dayanmaktadır. Binary çıktılı modeller için Y, (0,1) aralığında tanımlanmaktadır ve tahmin edilen olasılığa yani p'ye eşittir. Bu gözlenen ve tahmin edilen değerler arasındaki fark, "modelin uyum iyiliği" prosedürü ile sınanmaktadır. İyi bir model için bu fark minimum seviyelerde olması gerekmektedir [14].

Açıklanan varyasyon (R^2), sürekli ölçüme sahip sonuç değişkenli modellerin performanslarının değerlendirilmesi için kullanılmaktadır [13]. Genelleştirilmiş doğrusal modellerde artık varyanslarını belirlemek oldukça zor bir süreçtir ve bu modeller maksimum olabilirlik yaklaşımına dayanmaktadır [15]. Bu nedenlerle, Cox-Snell ve Nagelkerke tarafından iki farklı R^2 yaklaşımı geliştirilmiştir. Nagelkerke R^2, Cox ve Snell tarafından geliştirilen R^2 yaklaşımının düzeltilmiş halidir. R^2 istatistiği, bir uyum iyiliği ölçütü değildir. Bu istatistik "Açıklayıcı değişkenlerin, cevap değişkenini ne kadar açıklayabildiğinin" bir katsayısıdır ve etki büyüklüğünün ölçüsü olarak da tanımlanabilmektedir [16].

Cox ve Snell R^2 istatistiği,

$$R^2_{C\&S} = 1 - (LL_0/LL_1)^{2/n}$$

şeklinde [17] iken Nagelkerke R^2 istatistiği,

$$R^2 = (1 - \exp(-LR/n))/(1 - \exp(-2LL_0/n))$$

şeklindedir. Burada n; örnek büyüklüğü, LL_1; belirleyicilerin dahil olduğu modelin log-olabilirlik fonksiyonu iken LL_0; sıfır (null) modelin log-olabilirlik fonksiyonudur ve LR;

$$LR = -2(LL_0 - LL_1)$$

şeklindedir ve χ^2 dağılımı göstermektedir. $LR=+2LL_0$ olması durumunda $R^2=1$ olacaktır. Bu durum tahmin edilen değerlerle gözlemlenen değerlerin birbiri ile aynı olduğunu göstermektedir [2].

Brier skoru, kuadratik skorlama yöntemine dayanan bir yaklaşımdır. Bu istatistik değeri, ikili sonuç değişkeni olan Y ve tahmin edilen p arasındaki farkın karesi $((Y-p)^2)$ esasına dayanır. Brier skor 0 ile 0,25 aralığında değer aldığında iyi bir model çıkarsaması yapılırken 0,50 dolaylarında değer olması modelin bilgi içermeyen bir model olduğu sonucuna varılmasını sağlar [2]. Brier skor değeri şu şekilde hesaplanmaktadır.

$$B = \frac{1}{n}\sum_{i=1}^{n}(y_i - p_i)^2$$

Burada n toplam birey sayısı iken y_i, tahmin edilen risk ve p_i, gözlemlenen risk değerini göstermektedir [18].

2.2.1.2. Ayırım Yeteneğinin Değerlendirilmesi

Doğru tahmin terimi, hastalığı olan ve olmayan bireylerin ayırımının maksimum seviyede olmasıdır. Hasta ve sağlam kategorileri için kullanılabilecek özet tanımlayıcılar (MRD, AARD vb.) ve concardance (c) istatistiği ile modelin ayırım performansı değerlendirilebilir. c istatistiği ikili sonuç değişkenine sahip modeller için ROC eğrisi altında kalan alan değeri ile özdeş bir yöntemdir [2].

Vaka ve kontrollere ait ortalama risk değerlerinin farkı (Mean Risk Difference-MRD) uygulanabilecek bir özet performans değerlendirme yaklaşımından biridir.

$$MRD = E(risk(X)|D = 1) - E(risk(X)|D = 0)$$

MRD istatistiği Yates eğrisi olarak da adlandırılmaktadır. Aynı zamanda Pencina ve arkadaşlarının önerdiği iç içe (nested) tasarıma sahip risk tahmin modellerinin performans değerlendirmesinde kullanılan entegre ayırım iyileştirmesinin (IDI) bir parçasıdır ve MRD istatistiği açıklanan varyasyon ya da tanımlayıcılık katsayısının bir oranı şeklinde de tanımlanabilir ($R^2=var\{E(D|X)\}/var(D)$).

Vaka ve kontrolü ayrı ayrı ele alan bir diğer istatistik, ortalama risk farkının üzerinde (Above Average Risk Difference-AARD) olarak tanımlanan bir yaklaşımdır. Bu yaklaşıma göre AARD istatistiği şu şekilde tanımlanmaktadır;

$$\mathrm{AARD} \equiv \mathrm{P}(\mathrm{risk}(\mathrm{X}) > \rho | \mathrm{D} = 1) - \mathrm{P}(\mathrm{risk}(\mathrm{X}) > \rho | \mathrm{D} = 0)$$

Buradaki ρ ifadesi popülasyondaki ortalama risk olarak tanımlanmaktadır.

AARD istatistiği ise iki iç içe risk modelinin performans değerlendirmesinde kullanılan net yeniden sınıflama endeksine oldukça yakın bir yaklaşımdır [12].

Ek olarak vaka ve kontrol dağılımları arasındaki farkın belirlenmesine yönelik olarak kullanılan bir diğer yaklaşım alıcı-işlem karakteristikleri (Receiver-operating characteristic-ROC) eğrisi altında kalan alanın (AUC) belirlenmesidir [12]. ROC analizi ilk olarak II. Dünya Savaşında radar algılamasında gürültü sinyallerinin sınıflama doğruluklarını belirlemek için kullanılmıştır. Son zamanlarda bu analiz yöntemi, klinik uygulamalarda özellikle görüntüleme ve tanı testlerinin doğruluk değerlendirmesinde oldukça sık karşılaşılmaktadır [18].

Belirlenen kesim noktası c ile duyarlılık ve seçicilik değerleri,

Çizelge 1.2.2. Risk Tahmin Modelinden Elde Edilen Risk Olasılıklarının Kesim Noktasına ve Hastalığın Olması Durumuna Göre Çapraz Tablo

	D=1	D=0	
r>c	n_{11} Doğru Pozitiflik	n_{12} Yanlış Pozitiflik	$N_{1.}$
r<c	n_{21} Yanlış Pozitiflik	n_{22} Doğru Negatiflik	$N_{2.}$
	$N_{.1}$	$N_{.2}$	N

Duyarlılık= Doğru Pozitiflik Oranı= $\Pr(r>c \mid D=1)= \frac{n_{11}}{N_{.1}}$

Seçicilik= $\Pr(r<c \mid D=0)= \frac{n_{22}}{N_{.2}}$

Hatalı Pozitiflik Oranı= 1 – Seçicilik= $\Pr(r>c \mid D=0) = \frac{n_{21}}{N_{.2}}$

şeklindedir. Duyarlılık, hasta olan bireylerin doğru sınıflandırılması iken, seçicilik sağlam olan bireylerin doğru sınıflandırılmasıdır [1].

ROC, duyarlılık ve $(1-\text{seçicilik})$ eksenlerine sahip ve muhtemel tüm c kesim noktaları için çizilen bir eğridir. Artan duyarlılık ile azalan seçicilik arasında bir denge söz konusudur ve bu denge mekanizmasının tersi de doğrudur. Y sürekli bir değişken olduğunda ROC eğrisi,

$$P[Y_D \geq Y_{\bar{D}} | F_{\bar{D}}(Y_{\bar{D}}) = t] = P\left[Y_D \geq Y_{\bar{D}} \mid Y_{\bar{D}} = F_{\bar{D}}^{-1}(t)\right] =$$

$$P\left[Y_D \geq F_{\bar{D}}^{-1}(t)\right] = F_D\left\{F_{\bar{D}}^{-1}(t)\right\} = ROC(t) \qquad t\in(0,1)$$

olarak tanımlanır ve burada F_D ve $F_{\bar{D}}$ hastalığı olan ve olmayan bireylere ait fonksiyonlar iken Y_D ve $Y_{\bar{D}}$ hasta ve sağlam kategorisindeki bireylerin sırasıyla test sonuçlarıdır [19]. t ise hatalı pozitiflik oranını göstermektedir ve bu oran 0 ile 1 arasında değer almaktadır [20].

τ; t için muhtemel tüm kesim noktaları, Z; testin doğruluğunu etkileyen bazı faktörler ve X; ortak değişkenlerin bir vektörü olmak üzere lojistik regresyon için ROC eğrisi aşağıdaki gibi tanımlanabilir.

$$ROC_Z(t) = g\{\alpha_0(t), \beta X\} \qquad (t\epsilon\tau_Z)$$

Burada $\alpha_0(t)$; t'nin tek değişkenli bir fonksiyonu iken β; X bağımsız değişkenlerinin tahmin edicileridir. Basit bir ifade ile üç kategorili bir test değişkene sahip lojistik model,

$$ROC_Z(t) = \frac{\exp\{\alpha_0(t)+\beta_2 X_2+\beta_3 X_3\}}{1+\exp\{\alpha_0(t)+\beta_2 X_2+\beta_3 X_3\}},$$

şeklindedir.

AUC değerinin hesaplanabilmesi için parametrik, semi-parametrik ve parametrik olmayan yaklaşımlar olmak üzere 3 yaklaşım söz konusudur. Parametrik metotta, "hastalığı olan ve olmayan bireyler kendi içlerinde çok değişkenli normal dağılım gösteriyor" varsayımı aranır. Bu durumda AUC değeri aşağıdaki formülle hesaplanır.

$$AUC = \Phi\left(\sqrt{\frac{\delta'\Sigma^{-1}\delta}{2}}\right)$$

$\Phi(.)$ standart normal kümülatif dağılım fonksiyonu iken Σ varyans-kovaryans matrisi ve δ grupların ortalamalarının farkıdır.

Semi-parametrik yaklaşımda çok değişkenli normal dağılım varsayımı aranmaksızın risk skorlarının monoton dönüşümleri üzerinden AUC değeri hesaplanmaktadır. Yapılan monoton dönüşümle hasta ve sağlam kategorilerine ait risk skorları normal dağılıma dönüştürülür. Bu nedenle semi-parametrik yaklaşımla AUC değerinin hesaplanması "bi-normal AUC" olarak da adlandırılmaktadır.

Parametrik olmayan yaklaşımla AUC değerinin yansız tahmin edicisi Mann-Whitney istatistiğine dayanır.

$$\mathrm{AUC} = \frac{1}{mn}\sum_{j=1}^{n}\sum_{i=1}^{m}\psi(X_i, Y_j)$$

burada,

$$\psi(X_i, Y_j) = \begin{cases} 1, & Y < X \\ 0.5, & Y = X\text{'dir.} \\ 0, & Y > X \end{cases}$$

m ve n sırasıyla vaka ve kontrol grubunu ifade etmektedir ve X_i (i=1,…,m) vakalara bağlı test sonucunu, Y_j (j=1,…,n) kontrollere bağlı test sonucunu göstermektedir. Bu test istatistiği DeLong ve arkadaşları tarafından geliştirilmiştir [21] ve ileride bahsi geçen iki içe içe tasarıma sahip modelin AUC değerlerinin karşılaştırılmasında kullanılan istatistiğin bir parçasıdır.

Bir risk tahmin modelinin ayırım yeteneğini değerlendirirken kullanılabilecek yaklaşımlardan biri de ayırım eğimlerinin belirlenmesidir. Bu eğim ölçütü, hasta ve sağlam kategorisi için ortalama tahminlerin farkının alınması ile hesaplanır. Aynı zamanda iki sınıf için çizilecek box plot ve histogramlarla ayırımın görsel olarak değerlendirilmesi mümkündür. Hastalığı olan bireylerle olmayan bireylerin risk değerlerinin dağılımı ne kadar az kesişirse ayırımın o kadar iyi olduğunun sonucuna varılabilir.

2.2.1.3. Kalibrasyonun Değerlendirilmesi

Kalibrasyon genel bir tanım ile gözlem değerleri ile tahmin değerleri arasındaki farkın ölçümüdür. Kalibrasyonu grafiksel olarak değerlendirmek mümkündür. Bu grafik, x ekseninde tahmin değerleri y ekseninde gözlem değerleri yer alan ve serpme (scatter) grafiği ile aynı görsele sahip olan bir grafiktir. İyi bir tahmin modelinde kalibrasyon eğrisi 45 derecelik bir eğime sahip olmalıdır. İkili sonuç değişkenine sahip modellerde bu eğrinin y ekseni sadece 0 ve 1'lerden oluşacaktır. Bu durumda, gözlem değerlerinin tahmininde Loess algoritmasına dayanan Smoothing tekniği kullanılır [2].

Kalibrasyonun test prosedürü için Hosmer-Lemeshow istatistiği kullanılmaktadır. Bu test istatistiği şu şekilde tanımlanmaktadır;

$$H \equiv \sum_{k=1}^{10} N_k \frac{(O_k - \bar{\hat{r}}_k(X))^2}{\bar{\hat{r}}_k(X)(1 - \bar{\hat{r}}_k(X))}$$

Burada N_k, k. kategorideki birey sayısı, O_k, gözlemlenen olay oranı ve $\bar{\hat{r}}_k(X)$, modelden elde edilen ortalama tahmin değeridir. H istatistiği H_0 hipotezi altında grup sayısı eksi 2 serbestlik dereceli ki-kare dağılımı göstermektedir ($\chi^2(g-2)$). Yapılan simülasyon çalışmalarıyla grup sayısının 10 olarak alınmasının uygun olacağı gösterilmiştir [8]. Ancak bu test yöntemi birey sayısına (N) bağlı olduğu için bazı sınırlamalar içermektedir. Yani N yeteri kadar büyük ise sonuç istatistiksel olarak anlamlı N küçük ise anlamsız olarak bulunacaktır [12].

Kalibrasyon grafiğinin bir uzantısı olarak geçerlilik grafiği ile genel model performansı ve kalibrasyon yeteneği değerlendirilebilmektedir. Söz konusu grafik hastalığı olan ve olmayan bireylerin risk dağılımları ile çizilen eğrinin üzerine her bir desil için %95'lik güven aralıklarını gösteren hata çubuklarının çizilmesiyle oluşan grafiklerdir.

Tüm anlatılanların ışığında tek bir risk modelinin performans değerlendirmesinde izlenmesi gereken yol,

(i) İlgili popülasyon için modelin genel performansını, ayırım yeteneğini ve kalibrasyonunu değerlendirmek,

(ii) Vaka ve kontrollerin dağılımını belirlemeye yönelik grafikler çizmek ve tanımlamaya yardımcı olarak dağılımlar arasındaki mesafelerin belirlenmesine yönelik yardımcı endeksleri kullanmak,

(iii) Klinik destekle risk kesim noktasını belirlemek veya birkaç eşik değeri belirleyerek tedavi şekli için verilecek karara yardımcı olmak

şeklindedir.

Eğer karşılaştırılan risk tahmin modelleri birbirlerinden bağımsız ise izlenecek yol aşağıdaki gibidir;

(i) Her iki modelin kalibrasyonunu değerlendirmek

(ii) Risk dağılımlarını ve net fayda değerlendirmesini her iki model içinde ortaya koymak

(iii) Her iki modelde tedavi için klinik olarak anlamlı olan bir eşik değeri belirlemek ayrıca modellerin vaka ve kontrol için dağılımlarını karşılaştırmak

(iv) İki model için elde edilen net fayda değerlerini karşılaştırmak

şeklindedir.

2.2.2. İç İçe Tasarıma Sahip Modellerin Performanslarının Değerlendirilmesi

2.2.2.1. Cook ve Ridker Analiz Metodu

Cook ve Ridger metodu, vaka ve kontrollerin birleşiminin tek bir yeniden sınıflama tablosuyla gösterilmesine dayanmaktadır. Bu yöntemin işleyiş adımları ise şu şekildedir.

(i) Toplam yeniden sınıflama oranını (RC)

(ii) Doğru yeniden sınıflama yüzdesini (RC-correct)

(iii) Temel modele ait yeniden sınıflamaya dayanan kalibrasyon istatistiğini RCC^{X} ve p değerini

(iv) Genişletilmiş modele ait yeniden sınıflamaya dayanan kalibrasyon istatistiğini $RCC^{(X,Y)}$ ve p değerini belirlemek.

Burada RC ifadesinin hesaplanması diagonal dışında kalan deneklerin oranına dayanmaktadır. Bu RC oranı model karşılaştırmasından ziyade tanımlayıcı bir istatistiktir. RC istatistiğinin küçük değerler alması "X temel tahmin modeline Y ilave edildiğinde çok az denekte tedavi kararının ne olacağı değişiklik gösterecektir" sonucuna varılmasını sağlayacaktır.

RC-correct ifadesi ise risk(X) kategori etiketi içinde olmayan ve risk(X,Y) kategori etiketi içinde olan deneklerin gözlem oranlarına dayanmaktadır. Bu istatistik değerinin yüksek değerlerde bulunması modele Y'nin ilavesinin temel modele bağlı risk (risk(X)) hesabından daha iyi sonuçlar verdiğini göstermektedir. Yani RC-correct≈%100 ise modelin iyi kalibrasyon yeteneğine sahip olduğu sonucuna varılabilir.

Yeniden sınıflamaya bağlı kalibrasyon istatistiği,

$$\mathrm{RCC}^{X} = \sum_{k=1}^{K} \frac{(\hat{p}_k - \mathrm{ort}(\mathrm{risk}(X))_k)^2}{\mathrm{ort}\big(\mathrm{risk}(X)\big)_k (1 - \mathrm{ort}\big(\mathrm{risk}(X)\big)_k)/n_k}$$

ve

$$\mathrm{RCC}^{(X,Y)} = \sum_{k=1}^{K} \frac{(\hat{p}_k - \mathrm{ort}(\mathrm{risk}(X,Y))_k)^2}{\mathrm{ort}\big(\mathrm{risk}(X,Y)\big)_k (1 - \mathrm{ort}\big(\mathrm{risk}(X,Y)\big)_k)/n_k}$$

şeklindedir. Burada K tablonun içindeki hücre sayısını, ort(risk( ))$_k$ k hücresinde modellenen risk ortalamasını ve $\hat{p}_k$ gözlemlenen olay oranını ifade etmektedir. Ayrıca $n_k \geq 20$ olması gerekmektedir. Cook ve Ridker [22] çalışmalarında yeniden sınıflama kalibrasyon istatistiğinin K-2 serbestlik dereceli ki-kare dağılımı gösterdiğini belirtmişlerdir.

Cook ve Ridker analizi model karşılaştırmasından ziyade modellere bağlı tanımlayıcı istatistikleri ifade eden bir yaklaşımdır. Model karşılaştırmasında ROC eğrisi altında kalan alan endeksi, NRI ve IDI istatistikleri daha uygun yaklaşımlardır [12].

2.2.2.2. ROC Eğrisi Altında Kalan Alan Endeksi

ROC eğrisi, duyarlılık ve $(1 - \text{seçicilik})$ eksenlerine sahip olan bir grafiktir. ROC eğrisi ve onunla ilişkili olan c-istatistiği, duyarlılık ve seçicilik karakteristiklerine

dayanan ve hasta ya da hasta değil ayırımının yapılabilmesine olanak sağlayan bir istatistiksel metottur [23]. İki istatistik değeri de 0,5 ile 1 arasında değer alır ve 1'e yaklaşması durumunda 'tüm vakalar kontrollerle çakışmaksızın daha yüksek değerlere sahiptir' sonucuna varılır [24]. Burada unutulmaması gereken en önemli nokta c-istatistiği ve ROC eğrisi altında kalan alan değerinin hesabı ranklara dayandığından, ölçüm değerlerinin asıl değerlerine ve olabilirlik oranlarına karşı daha az hassastır. ROC eğrisi altında kalan alan değerinin hesaplanması model karşılaştırmasında özellikle temel model iyi bir performansa sahipse duyarsız sonuçlarla karşılaşılmaktadır [25]. Ayrıca, risk tahmin modellerine ilgilenilen yeni faktörün dahil edilmesinden ve edilmemesinden elde edilen ROC eğrisi altında kalan alan değerleri arasındaki farkın tahmin üzerine katkısını araştırmak ROC eğrisinin diğer bir uygulama alanıdır. İki ROC eğrisini karşılaştırmada kullanılan bir diğer yaklaşım, seçilen bir noktada ROC eğri değerlerinin aynı olup olmadığına bakmaktır.

Burada yeni modelle hesaplanan ROC eğrisi altında kalan alan (AUC) değeri ile eski modelle eğri altında kalan alan farkının testinde hipotezler,

H_0: $AUC_{yeni} = AUC_{eski}$

H_A: $AUC_{yeni} \neq AUC_{eski}$

şeklindedir. Aynı bireylerden toplanan yeni ve eski testlere bağlı iki ampirik ROC eğrisi karşılaştırıldığında testler arasındaki iç içe (nested) tasarım göz önünde bulundurulmalıdır. Bu yapıyı göz önüne alan, DeLong ve arkadaşları tarafından geliştirilen test istatistiği şu şekildedir [21];

$$\hat{\theta}_k = \frac{1}{mn}\sum_{j=1}^{n}\sum_{i=1}^{m}\psi(X_i, Y_j)$$

burada,

$$\psi(X_i, Y_j) = \begin{cases} 1, & Y < X \\ 0.5, & Y = X\text{'dir.} \\ 0, & Y > X \end{cases}$$

m ve n sırasıyla vaka ve kontrol grubunu ifade etmektedir ve X_i (i=1,…,m) vakalara bağlı test sonucunu, Y_j (j=1,…,n) kontrollere bağlı test sonucunu göstermektedir. $\hat{\theta}_k$,

yeni ve eski modele bağlı olarak hesaplandıktan sonra bu iki test istatistiğinin anlamlılığının testi,

$$z = \frac{(\hat{\theta}_{yeni} - \hat{\theta}_{eski} - (AUC_{yeni} - AUC_{eski}))}{([1 \quad -1]\, S\, [1 \quad -1]')^{1/2}}$$

şeklindedir. Burada S, yeni ve eski modele ait varyans-kovaryans matrisini ifade etmektedir. Ayrıca z test istatistiği standart normal dağılım göstermektedir ve p-değeri $2(1-\Phi(|z|))$ ile hesaplanmaktadır ($\Phi(.)$ standart normal kümülatif dağılım fonksiyonudur).

2.2.2.3. Net Yeniden Sınıflandırma İyileştirmesi (NRI)

Yeniden sınıflama tablolarına dayanan bu teknikle hastalığı olan ve olmayan bireyler, risk kategorilerine göre sınıflandırılır ve yeni modelle, bu kategorilerdeki değişimin istatistiksel olarak önemli olup olmadığı araştırılır. NRI istatistiğinin hesaplanmasında kullanılan notasyonlar; yükselme eğilimli hareket (up), yeni modelin temel modelden daha yüksek kategorilerinde tanımladığı birey sayısı, aşağı eğilimli hareket (down) yeni modelin temel modelden daha düşük kategorilerinde tanımladığı birey sayısı ve D=0 sağlam grubu, D=1 hasta grubunu ifade etmektedir. O halde NRI değeri,

$$NRI = [P(up|D = 1) - P(down|D = 1)] - [P(up|D = 0) - P(down|D = 0)] \quad (3)$$

şeklinde hesaplanabilmektedir. Popülasyonun tamamı ile ilgilenilmediği için NRI'nın tahmin edicisi,

$$\widehat{NRI} = (\hat{p}_{up,hasta} - \hat{p}_{down,hasta}) - (\hat{p}_{up,sağlam} - \hat{p}_{down,sağlam}) \quad (4)$$

şeklinde olacaktır. H_0:NRI=0 için test istatistiği temel olarak oranların ilişkili olmasından dolayı McNamer yaklaşımına dayanmaktadır.

$$z = \frac{\widehat{NRI}}{\sqrt{\frac{\hat{p}_{up,hasta} + \hat{p}_{down,hasta}}{\text{hasta sayısı}} + \frac{\hat{p}_{up,sağlam} + \hat{p}_{down,sağlam}}{\text{sağlam sayısı}}}}$$

H_0'ın reddedilmediği durumda yeni faktörün modele konulmasının istatistiksel olarak bir anlamı olmadığı kararına varılacaktır. Hasta ve sağlam bireylerin ayrı ayrı sınıflandırılmasına bağlı olarak ayırımın sıfırdan farklı olup olmadığının testi için z istatistiği aşağıdaki gibi tanımlanabilir [26].

$$z_{\text{hasta}} = \frac{\hat{p}_{\text{up,hasta}} - \hat{p}_{\text{down,hasta}}}{\sqrt{\frac{\hat{p}_{\text{up,hasta}} + \hat{p}_{\text{down,hasta}}}{\text{hasta sayısı}}}} \quad \text{ve}$$

$$z_{\text{sağlam}} = \frac{\hat{p}_{\text{down,sağlam}} - \hat{p}_{\text{up,sağlam}}}{\sqrt{\frac{\hat{p}_{\text{up,sağlam}} + \hat{p}_{\text{down,sağlam}}}{\text{hasta sayısı}}}}$$

Formül (4), Bayes Teoremine göre genelleştirilirse,

$$\text{NRI} = \frac{\text{P(hasta|up).P(up)} - \text{P(hasta|down).P(down)}}{\text{P(hasta)}} + \frac{\text{P(sağlam|down).P(down)} - \text{P(sağlam|up).P(up)}}{\text{P(sağlam)}} \tag{5}$$

ifadesi elde edilir. Burada P(sağlam) yerine 1-P(hasta) ve P(sağlam|*) yerine 1-P(hasta|*) yazılırsa, (burada * ile up ya da down yönlü hareket ifade edilmektedir.)

$$\text{NRI} = \frac{\text{P(hasta|up).P(up)} - \text{P(hasta|down).P(down)}}{\text{P(hasta)}} + \frac{\text{P(sağlam|down).P(down)} - \text{P(sağlam|up).P(up)}}{\text{P(sağlam)}} \tag{6}$$

ve gerekli sadeleştirmelerle,

$$\text{NRI} = \frac{[\text{P(hasta|up)} - \text{P(hasta)}]\text{P(up)} + [\text{P(hasta)} - \text{P(hasta|down)}].\text{P(down)}}{\text{P(hasta)}[1 - \text{P(hasta)}]} \tag{7}$$

eşitliği elde edilir [27]. Formül (6) veya (7)'nin uzantısı yaşam analizinde P(hasta), P(hasta|up) ve P(hasta|down) olasılık değerleri ile tüm tahminlerde kullanılan Kaplan-Meier yaklaşımına benzerlik göstermektedir. Ayrıca bu formüller, Kaplan-Meier oranları ile yarışan risk modellerinin karşılaştırılmasına olanak vermektedirler [7].

Bazı alanlarda (kardiyovasküler hastalıklarda primer koruma (CVD)) risk kategorileri net bir şekilde oluşturulmuş ve hasta tedavi şekli ve prognoz bu kategorilere bağlı olarak belirlenebilirken diğer alanlarda (tüm ölümlerin nedenleri, diyabet, atriyal

fibrilasyon vb.) risk kategorilerinin net bir tanımını oluşturma çabaları eksik kalmış ya da doğrulukları sınanamamış olabilir. Hatta bazen risk kategorileri tam olarak tanımlanmış ancak ilgili kesim noktaları farklı tanımlandığından farklı modeller (koroner ölüm ve miyokard infarktüs için "hard" CVD ve "full" CVD modellemesi, koroner kalp hastalığı (CHD) ve full CHD v.b.) için farklı insidans değerleri hesaplanabilir. Bu da farklı modellere aynı markerin eklenmesi durumunda farklı NRI değerlerinin hesaplanmasına neden olacaktır. Yaşanabilecek buna benzer senaryolardan dolayı kategori tanımlaması olmayan yeni bir NRI yaklaşımı geliştirilmiştir. Bu yeni NRI yaklaşımı NRI(>0) şeklinde gösterilmektedir ve c-istatistiği ile aynı yapı taşlarına sahiptir ancak farklı bir şekilde düzenlenmiştir.

$$\frac{1}{2}\mathrm{NRI}(>0) = \mathrm{P}(\mathrm{Q_i} > \mathrm{P_i}|\mathrm{i} = \mathrm{hasta}) - \mathrm{P}\big(\mathrm{Q_j} > \mathrm{P_j}\big|\mathrm{j} = \mathrm{sağlam}\big)$$

$$\Delta\mathrm{C} = \mathrm{P}\big(\mathrm{Q_i} > \mathrm{Q_j}\big|\mathrm{i} = \mathrm{hasta}, \mathrm{j} = \mathrm{sağlam}\big) - \mathrm{P}\big(\mathrm{P_i} > \mathrm{P_j}\big|\mathrm{i} = \mathrm{hasta}, \mathrm{j} = \mathrm{sağlam}\big)$$

Q: "yeni" risk tahmin algoritmasına dayalı olayın tahmin olasılıkları ve

P: "eski" risk tahmin algoritmasına dayalı olayın tahmin olasılıkları şeklinde ifade edilmektedir.

Sürekli NRI(>0) yaklaşımı c-istatistiğinde olduğu gibi insidans değerlerinden etkilenmez. Bu nedenle farklı çalışmalar arasında mukayese edilebilirlik sağlanmış olur [7].

NRI yaklaşımı, maliyet ile ilgili değerlendirmeleri dikkate almamaktadır. Bu nedenden dolayı ağırlıklandırılmış NRI (weighted NRI-wNRI) adı ile yeni bir yaklaşım geliştirilmiştir. Yeni model ile olayın olması durumu sınıflandırıldığında eski modele kıyasla daha yüksek risk gruplarında tanımlanmasından elde edilecek yarar ya da maliyet s_1 iken yeni modelle olayın olmaması durumu daha düşük risk gruplarında sınıflandırıldığında elde edilecek yarar s_2 olarak tanımlansın. Bayes yaklaşımı ile wNRI,

$$\mathrm{wNRI} = \mathrm{s_1}\big(\mathrm{P(hasta|up)P(up)} - \mathrm{P(hasta|down)P(down)}\big)$$

$$+ \mathrm{s_2}(\mathrm{P(sağlam|down)P(down)} - \mathrm{P(sağlam|up)P(up)})$$

şeklindedir [28].

NRI istatistiğinin anlamlılığının sınanması halen kesinlik kazanmamış bir konudur. Pencina'nın önerdiği test sınamasında, küçük NRI değerlerinde (<0,01) bile istatistiksel olarak önemli olan p değerleri hesaplandığı görülmüştür. Hatta Pepe ve arkadaşları, NRI'nın önemlilik çıkarsamasında geçerli bir yöntem bulunmadığını ileri sürmüşlerdir [29].

2.2.2.4. Entegre Ayrım İyileştirmesi (IDI)

Entegre ayırım iyileştirmesi (IDI), duyarlılık ve 1-seçicilik için tüm olası kesim noktalarına göre integralinin alınmasına dayanan bir metottur. IDI istatistiği şu şekilde ifade edilmektedir.

$$\mathrm{IDI} = \left(\mathrm{IS}_{\mathrm{yeni}} - \mathrm{IS}_{\mathrm{eski}}\right) - \left(\mathrm{IP}_{\mathrm{yeni}} - \mathrm{IP}_{\mathrm{eski}}\right) \tag{8}$$

Burada, $\mathrm{IS}_k = \int \mathrm{P}(p_k > c|D = 1)dc$ doğru pozitiflik oranının (DPO=duyarlılık) ve $\mathrm{IP}_k = \int \mathrm{P}(p_k > c|D = 0)dc$ ise yanlış pozitiflik oranının (YPO=1-seçicilik), muhtemel tüm kesim noktaları altındaki integralidir. NRI istatistiğindeki gibi benzer bir mantıkla IDI istatistiğinin tahmin edicisi,

$$\widehat{\mathrm{IDI}} = \left(\bar{\hat{p}}_{\mathrm{yeni,hasta}} - \bar{\hat{p}}_{\mathrm{eski,hasta}}\right) - \left(\bar{\hat{p}}_{\mathrm{yeni,sağlam}} - \bar{\hat{p}}_{\mathrm{eski,sağlam}}\right) \tag{9}$$

şeklinde olacaktır. $\bar{\hat{p}}_{\mathrm{yeni,hasta}}$ ifadesi hasta grubunda yeni modele bağlı tahmin edilen olasılığın ortalaması iken $\bar{\hat{p}}_{\mathrm{eski,hasta}}$ eski modelle tahmin edilen olasılığın ortalamasını ifade etmektedir ve benzer şekilde $\bar{\hat{p}}_{\mathrm{yeni,sağlam}}$ ve $\bar{\hat{p}}_{\mathrm{eski,sağlam}}$ yeni ve eski modele bağlı sağlam grupları için tahmin edilen olasılığın ortalamasını göstermektedir. IDI istatistiği, iki modelle de elde edilen ayırım eğrilerinin (discrimination slope) farkı temeline dayanmaktadır [27]. (9)'deki formül düzenlendiği takdirde Yates'in önerdiği yeni ve eski modele bağlı ayırım eğrilerinin farkı elde edilecektir.

$$\widehat{\mathrm{IDI}} = \left(\bar{\hat{p}}_{\mathrm{yeni,hasta}} - \bar{\hat{p}}_{\mathrm{yeni,sağlam}}\right) - \left(\bar{\hat{p}}_{\mathrm{eski,hasta}} - \bar{\hat{p}}_{\mathrm{eski,sağlam}}\right)$$

Buna paralel olarak, Pepe ve arkadaşları (2008)[30] IDI yaklaşımına farklı bir bakış açısı katarak açıklanan varyasyon oranına (PEV) bağlı bir eşitlik geliştirmişlerdir. Bu yaklaşımla, ikili (binary) risk regresyon modellerinde, yeni bir faktörün modele ilave

edilmesi durumunda elde edilecek R^2'deki değişimin anlamlılığını inceleme olanağı bulunmuştur. Bu yaklaşım şöyle ifade edilmektedir.

$$\mathrm{PEV_{eski}} \equiv \frac{\mathrm{var(D)} - \mathrm{E\{var(D|X)\}}}{\mathrm{var(D)}} = \mathrm{var(p_{eski})}/\rho(1-\rho)$$

Burada, $\rho = \mathrm{P(D=1)}$ (prevelans) şeklindedir. Benzer şekilde yeni model için PEV değeri,

$$\mathrm{PEV_{yeni}} \equiv \frac{\mathrm{var(D)} - \mathrm{E\{var(D|X,Y)\}}}{\mathrm{var(D)}} = \mathrm{var(p_{yeni})}/\rho(1-\rho)$$

şeklinde olacaktır. Yeni model ve eski modele bağlı PEV değerlerinin farkı IDI istatistiğini verecektir.

$$\mathrm{IDI} = \mathrm{PEV_{yeni}} - \mathrm{PEV_{eski}} \qquad (10)$$

Bu yaklaşıma ek olarak, sınıflandırma hatasına ya da mutlak hatalara bağlı yeni bir yorum geliştirilmiştir. $|\mathrm{D} - \mathrm{p_k}|$ mutlak hata değeri hastalar için (1-$\mathrm{p_k}$), sağlamlar için $\mathrm{p_k}$ şeklindedir. Ek olarak, hastalar için ağırlıklandırma $\mathrm{w_1} = \rho^{-1}$ ve sağlamlar için ağırlıklandırma $\mathrm{w_0} = (1-\rho)^{-1}$ şeklinde ifade edilirse artıkların yeni ve eski model için beklenen değeri,

$$\mathrm{E\{artık\}_k} = \mathrm{w_1 E(1-p_k|D=1)}\rho + \mathrm{w_0 E(p_k|D=0)}(1-\rho)$$

$$= \mathrm{E(1-p_k|D=1) + E(p_k|D=0)}$$

ve IDI istatistiği,

$$\mathrm{IDI} = \mathrm{E\{artık\}_{eski} - E\{artık\}_{yeni}} \qquad (11)$$

şeklinde olacaktır.

H_0: IDI=0 hipotezinin testi Pencina ve arkadaşları (2008) tarafından, hasta ve sağlam gruplarının bağımsız olması gerekçesiyle,

$$\mathrm{z} = \frac{\widehat{\mathrm{IDI}}}{\sqrt{(\widehat{se}_{\mathrm{hasta}})^2 + (\widehat{se}_{\mathrm{sağlam}})^2}}$$

şeklinde olduğu ifade edilmiştir. Burada $\widehat{se}_k$, hasta ve sağlam için standart hatayı göstermektedir. Pepe ve arkadaşları ayrıca yeni faktörün ilavesinde modelden elde edilen regresyon katsayısının anlamlılığının testinin de sınanabileceğini ifade etmiştir (H_0: β=0) [30].

IDI istatistiği geleneksel regresyon yöntemlerinin elemanı olan tanımlamaları da içeren daha güvenilir bir çıkarsama tekniğidir. Yani formül (11)'de mutlak hataların farkını, formül (10)'da R^2'deki değişimi dikkate almaktadır. Pencina ve arkadaşlarının geliştirdiği formül (8) ve (9)'da seçicilik ve duyarlılığa bağlı integrallere dayanan yaklaşımlarla IDI istatistiği daha net temellere dayanmaktadır. Bununla beraber NRI istatistiği uygulama alanında daha sık karşılaşılan bir metottur. Hiç şüphesiz bunun nedeni bu istatistiğin teorik olarak uygulama kolaylığına sahip olmasıdır. Bunlara ek olarak ROC eğrisi altında kalan alan değerinin farkının test prosedürü literatürde oldukça sık tartışılan bir konudur.

2.2.3. Mahalanobis Mesafelerinin Karelerine Dayanan Bir Fonksiyon ile Yaklaşımların Geliştirilmesi

X, p+q büyüklüğüne sahip olan bir vektör olsun. $X|D = 1 \sim N(\mu_1, \Sigma_1)$ ve $X|D = 0 \sim N(\mu_0, \Sigma_0)$ şeklinde iken μ_1, μ_0 ortalama vektörü ve Σ_1, Σ_0 varyans-kovaryans matrisidir. Burada ROC eğrisi altında kalan alan değişimi, NRI ve IDI istatistikleri, normal dağılım gerekçesiyle Mahalanobis mesafelerinin karelerine dayanan bir yaklaşımla ifadeleri söz konusudur. Mahalanobis mesafelerinin kareleri ifadesi lineer diskriminant analizine (LDA) dayanmaktadır. Hasta ve sağlamlara ait ortalamaların farklarının vektörü $\delta = \mu_1 - \mu_0$ olarak ifade edilir ve bunun $p(\delta_1)$ ve $q(\delta_2)$ için parçalanmış hali $\delta = \begin{pmatrix}\delta_1 \\ \delta_2\end{pmatrix}$ şeklindedir. Bu tanımlamalardan hareketle;

$a = \Sigma^{-1} \times \delta$ (p+q değişken ile LDA probleminin çözümü)

$b = {\Sigma_{11}}^{-1} \times \delta_1$ (p değişken ile LDA probleminin çözümü)

$M^2_{p+q} = \delta^T . {\Sigma^{-1}}^T . \delta$ (p+q değişken için Mahalanobis uzaklığı)

$M^2_p = {\delta_1}^T . {\Sigma_{11}^{-1}}^T . \delta_1$ (p değişken için Mahalanobis uzaklığı)

$L^*_{p+q}(X) = a^T X - \frac{1}{2} a^T(\mu_1 + \mu_0)$ (p+q değişkene dayanan LDA sınıflama fonksiyonu)

$L^*_{p}(X) = b^T X - \frac{1}{2} b^T(\mu_1 + \mu_0)$ (p değişkene dayanan LDA sınıflama fonksiyonu)

$p_{p+q}(X) = \frac{1}{1+r.e^{-L^*_{p+q}(X)}}$ (p+q değişkenine dayanan olayın tahmin olasılığı ve r prevalans ya da insidans değeri) şeklindedir. Model performansını değerlendirmek için kullanılan üç yaklaşım şu şekilde ifade edilebilir.

1. AUC değerindeki artış;

$$\Delta AUC = pr\left(p_{p+q}(X|D=1) > p_{p+q}(X|D=0)\right) - pr\left(p_p(X|D=1) > p_p(X|D=0)\right)$$

2. IDI (ayırım eğimlerinin farkı);

$$\begin{aligned} IDI &= \left\{E\left(p_{p+q}(X|D=1) - p_{p+q}(X|D=0)\right)\right\} \\ &- \left\{E\left(p_p(X|D=1) - p_p(X|D=0)\right)\right\} \\ &= E\left(\frac{1}{1+r.e^{-L^*_{p+q}(X|D=1)}} - \frac{1}{1+r.e^{-L^*_{p+q}(X|D=0)}}\right) \\ &- E\left(\frac{1}{1+r.e^{-L^*_{p}(X|D=1)}} - \frac{1}{1+r.e^{-L^*_{p}(X|D=0)}}\right) \end{aligned}$$

3. Sürekli NRI;

$$\begin{aligned} \frac{1}{2}NRI &= pr\left(p_{p+q}(X|D=1) > p_p(X|D=1)\right) \\ &- pr\left(p_{p+q}(X|D=0) > p_p(X|D=0)\right) \end{aligned}$$

LDA eşitliklerine başvurmaksızın yukarıdaki üç yaklaşımı tekrar ifade etmek mümkündür.

1. $\Delta AUC = \Phi\left(\sqrt{\frac{M_{p+q}^2}{2}}\right) - \Phi\left(\sqrt{\frac{M_p^2}{2}}\right)$

2. $IDI = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2.\pi.M_{p+q}^2}} \exp\left(\frac{-(x-0.5.M_{p+q}^2)^2}{2.M_{p+q}^2}\right)\left(\frac{1}{1+r.\exp(-x)} - \frac{1}{1+r.\exp(x)}\right) dx -$

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2.\pi.M_p^2}} \exp\left(\frac{-(x-0.5.M_p^2)^2}{2.M_p^2}\right)\left(\frac{1}{1+r.\exp(-x)} - \frac{1}{1+r.\exp(x)}\right) dx$$

3. $\frac{1}{2}NRI = \Phi\left(\frac{\sqrt{M_{p+q}^2 - M_p^2}}{2}\right) - \Phi\left(\frac{-\sqrt{M_{p+q}^2 - M_p^2}}{2}\right) = 2\,\Phi\left(\frac{\sqrt{M_{p+q}^2 - M_p^2}}{2}\right) - 1$

NRI ve ΔAUC yalnızca Mahalanobis farklarına dayanan yaklaşımlar oldukları için prevalans/insidans değerlerinden etkilenmemektedir. Bununla beraber IDI istatistiği Mahalanobis farklarının yanı sıra olayın prevalans/insidans (r) değerini içeren bir yaklaşım olmasından dolayı prevalans/insidans değerlerinden etkilenmektedir [31].

2.3. Risk Eşik Değeri ve Net Yarar Kavramı

Tedavi şeklinin belirlenmesi, zarar ve faydaya bağlı olarak değişim göstermektedir. Eğer tedavi, toksisite sorunu içeriyorsa, parasal giderleri fazla ise veya hasta açısından rahatsızlık verici bir süreçten oluşuyorsa yüksek eşik değeri seçilmelidir. Benzer şekilde tedavi zararları minimum düzeylerde ise eşik değeri daha düşük seçilmelidir.

Bunlara paralel olarak risk eşik değeri (r_H), net zarar ve faydanın bir uzantısı olduğu sonucuna varılabilir. Net fayda ve zarara bağlı eşitlik şu şekilde tanımlanabilir;

$$B \times P(D = 1|risk(X) = r) - C \times P(D = 0|risk(X) = r) = B \times r - C \times (r - 1)$$

Burada B net yarar ve C maliyettir. Eşitliğinin pozitif olabilmesi için,

$$\frac{r}{1-r} > \frac{C}{B}$$

şeklinde tanımlanması gerekmektedir. O halde risk kesim noktasının net fayda ve zarara bağlı ifadesi,

$$r_H = \frac{C}{C + B}$$

şeklinde olacaktır.

Risk kesim noktasının belirlenmesinden sonra riskin tanımı ikili kategori tanımına indirgenmiş olacaktır. Yani bireyler yüksek risk kategorisinde ve ya yüksek risk kategorisinde değil şeklinde değerlendirilebilecektir. Yüksek risk kategorisinin tanımı vaka ve kontroller için ayrı ayrı şu şekilde yazılabilir;

$$YR_H(r_H) = P(risk(X) > r_H | H = 1),$$

$$YR_{\bar{H}}(r_H) = P(risk(X) > r_H | H = 0)$$

İyi bir model tüm hastaları yüksek kategoride tanımlarken hasta olmayan bireyleri yüksek kategoride tanımlamayan modeldir ($YR_H(r_H) = 1$ ve $YR_{\bar{H}}(r_H) = 0$). $YR_H(r_H)$ aynı zamanda doğru pozitiflik oranı yani duyarlılık, $YR_{\bar{H}}(r_H)$ yanlış pozitiflik yani (1-seçicilik) olarak tanımlanmaktadır. Buradan hareketle doğru pozitiflik oranı net yarara pozitif yönlü katkı sağlarken yanlış pozitiflik oranı negatif yönde katkı sağlayacaktır. Belirli bir kesim noktası (r_H) için net yarar (NB),

$$\begin{aligned} NB(r_H) &= P(risk(X) > r_H)\{BP(D = 1 | risk(X) > r_H) - CostP(D = 0 | risk(X) > r_H) \\ &= B\,P(risk(X) \\ &> r_H | D = 1)P(D = 1) - Cost\,P(risk(X) > r_H | D = 0)P(D = 0) \\ &= BHR_D(r_H)\rho - CostHR_{\bar{D}}(r_H)(1 - \rho) \end{aligned}$$

şeklinde olacaktır.

2.3.1. İkiden Fazla Risk Kategorisi İçin Net Yarar

Burada yüksek risk kategorilerinin vaka ve kontrol için tanımı her bir kategori için ayrı ayrı tanımlanması şekilde basit bir yaklaşımla gerçekleştirilebilir. $B_{yüksek}$= vaka olarak tanımlanan grup için tedavinin net yararı, $C_{yüksek}$= kontrol olarak tanımlanan grup için tedavinin net maliyeti, $B_{düşük}$= kontrol olarak tanımlanan gruba tedavi uygulanmamasının net yararı ve $C_{düşük}$= vaka grubuna tedavinin uygulanmamasının net maliyeti şeklinde tanımlanırsa net yarar ifadesi,

$$\rho\{-C_{\text{düşük}}LR_D + B_{\text{yüksek}}HR_D\} + (1-\rho)\{-B_{\text{düşük}}LR_{\overline{D}} + B_{\text{yüksek}}HR_{\overline{D}}\}$$

şeklinde olacaktır. Burada LR ve HR sırasıyla düşük ve yüksek risk kategorilerini ve D ve $\overline{D}$ hasta ve kontrol grubunu ifade etmektedir.

2.4. Karar Analizi

Tanı ve prognozda kullanılan modeller doğruluk ölçütleri ile sınanabilmektedir. Ancak bu modellerin klinik sonuçları ile ilgili çıkarsamalar yapılamamaktadır. Karar analiz tekniği ile klinik sonuçların değerlendirilmesi mümkündür [32]. Bu analiz tekniği ile sonuçların birleştirilmesi ve modelin tüm durumlarda kullanımının doğru olup olmadığı ya da birçok alternatif modelin kullanılıp kullanılmayacağına karar verilmesi değerlendirilebilmektedir. Ayrıca bu teknik ile klinik yararı değerlendirmede gerekli görülen maliyet, yarar ve tercih gibi ekstra bilgilere gereksinim duyulmaz.

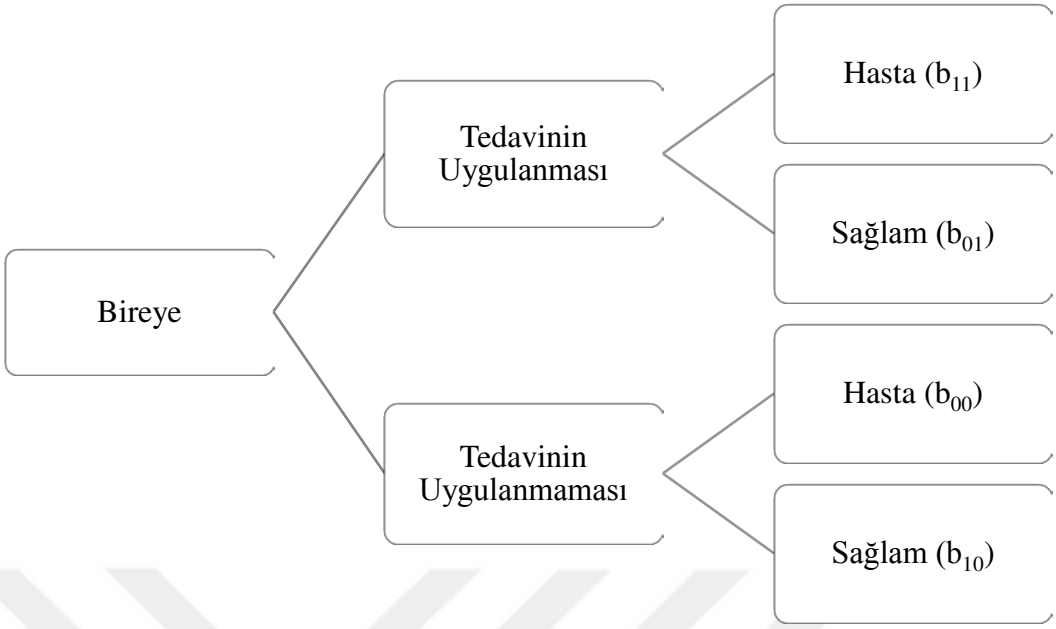
x, test sonucu ve y hastalık durumunu ifade etmek üzere p_{xy} olasılık değeri ve b_{xy} fayda terimi olarak tanımlansın. Burada kararı etkileyecek dört durum söz konusudur. Bunlar; doğru pozitiflik oranı ve fayda (p_{11} ve b_{11}), yanlış pozitiflik oranı ve fayda (p_{10} ve b_{10}), yanlış negatiflik oranı ve fayda (p_{01} ve b_{01}) ve son olarak doğru negatiflik oranı (p_{00} ve b_{00})'dır. Hastalığın prevelans değeri $\pi=P(D^+)$ ve p_t hastalığın muhtemel kesim noktası olarak tanımlandığında,

$$b_{11} \times p_t + b_{10} \times (1-p_t) = b_{01} \times p_t + b_{00} \times (1-p_t)$$

ve gerekli düzenlemeler yapıldığında,

$$\frac{b_{00}-b_{10}}{b_{11}-b_{01}} = \frac{p_t}{1-p_t} \qquad (12)$$

eşitliği elde edilir.



Şekil 2.4.1. Tedavi için karar ağacı [33]

Bu eşitlikten hareketle $b_{00} - b_{10}$, doğru negatiflik sonucu ile yanlış pozitiflik sonucu arasındaki fark yani klinik olarak gereksiz tedaviden kaçınmanın faydası iken $b_{11} - b_{01}$, doğru pozitiflik sonucu ile yanlış negatiflik sonucu arasındaki fark yani tedavinin uygulanmasından elde edilecek faydadır. Net fayda terimi, faydadan zararın çıkarılması ile elde edilir. O halde (12) eşitliğinde yanlış pozitiflikten doğan zararın, doğru pozitiflikten doğan faydadan çıkarılması ile elde edilen eşitlik aşağıdaki gibi tanımlanabilir.

$$-(b_{10} - b_{00}) = (b_{11} - b_{01}) \times \frac{p_t}{1-p_t}$$

Net fayda,

$$\frac{\text{Doğru Pozitif Sayısı} - \text{Yanlış Pozitif Sayısı}}{\text{Toplam Örnek Büyüklüğü}} \left(\frac{p_t}{1-p_t}\right) \qquad (13)$$

yani

$$\text{Duy} \times \pi - (1 - \text{Seç}) \times (1 - \pi) \times \left(\frac{p_t}{1 - p_t}\right)$$

şeklindedir [34].

Karar analizi yöntemini uygularken izlenebilecek adımlar aşağıdaki gibidir [34];

1) Eşik olasılığını (p_t) seçmek
2) $\hat{p} \geq p_t$ şeklinde ise bireylerin test sonucunu pozitif aksi halde test sonucunu negatif olarak belirlemek (Burada $\hat{p}$ modelden elde edilen hastalık olasılığıdır.)
3) Net yarar değerini hesaplamak
4) Tüm bireyleri tedavi etmenin klinik olarak net yararını hesaplamak

Yani duyarlılığı %100, seçiciliği %0 olarak almak ve net yarar değerini,

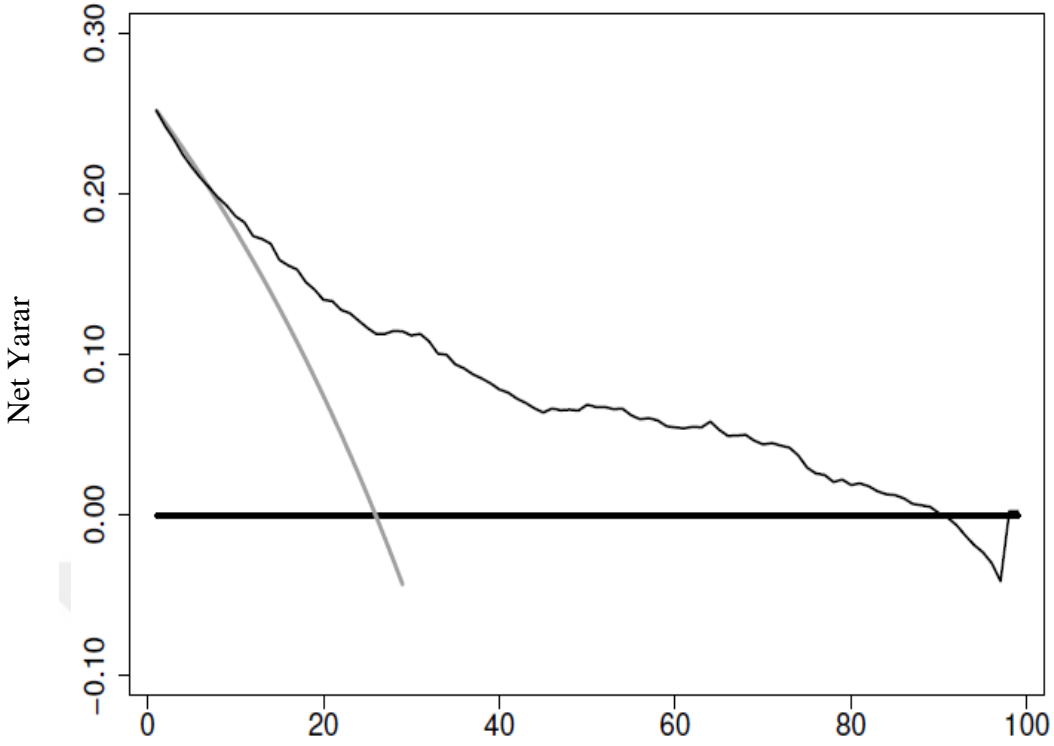
$$\pi - (1 - \pi) \times \left(\frac{p_t}{1 - p_t}\right)$$

eşitliği ile elde etmek

5) Tüm bireyleri tedavi etmemeden elde edilecek faydayı sıfır olarak tanımlamak
6) Klinik olarak en yüksek net yararı sağlayacak optimal stratejiyi belirlemek.

Eğer klinik olarak geçerli bir kesim noktası yoksa farklı p_t değerleri için net yarar değeri değişiklik gösterecektir. Net yarar değerleri y ekseninde ve kesim noktası olasılıkları x ekseninde olmak üzere çizilecek grafiğe karar eğrisi denir [34]. Çizilen karar eğrisi ile farklı risk tahmin modellerini karşılaştırmak ve eşik değerinin bölgesini belirlemek mümkündür (WenGu Tez). Genel olarak çizilebilecek bir karar eğrisi aşağıda gösterildiği gibidir. Burada gri çizgi, tüm bireylere tedavi uygulanmasından elde edilecek yararı gösterirken x eksenine paralel olarak çizilen çizgi hiç kimseye tedavi uygulanmamasından elde edilecek yardır (net yarar sıfırdır). Eğri ise muhtemel tüm kesim noktaları olasılıklarına (p_t) göre modelin net yarar teriminin yukarıda belirtilen (13) eşitliğinden elde edilmesi ile çizilir [14].

Karar eğrisi yaklaşımında, yarar terimi (b_{xy}) yerine kesim olasılıkları (p_t) teriminin kullanılmasına bağlı olarak dört avantaj söz konusudur. Bunlardan ilki sadece tek bir parametrenin seçilmesinin gerekli olmasıdır. İkincisi, hasta ve hekimin sağlık durumunu değerlendirmede 0 ve 1 aralığında tanımlanmış olan riski yorumlaması ve anlaması daha kolaydır. Üçüncü avantaj ise kesim olasılıkları ile hem pozitif ve negatif test sonuçları ağırlıklandırıldığı için hem de pozitif test sonucu için kesim noktasına karar verildiği için karar eğrisinin kullanımı klinik açıdan daha uygundur. Son olarak, karar eğrisi analizi ile iki iç içe veya bağımsız tasarıma sahip modellerin performanslarını değerlendirmekte mümkündür.



Şekil 2.4.2. Karar Eğrisi

Şekil 2.4.2'deki karar eğrisinin yorumu; kesim noktası olasılığı yaklaşık olarak %10 olduğu zaman istatistiksel model, tüm yüksek PSA seviyesine sahip olan bireylere biyopsi uygulanmasından daha yüksek bir karar yeteneğine sahiptir. Ayrıca yaklaşık olarak %90'lık bir kesim noktası olasılığı ile hiç kimseye biyopsi uygulanmamasından elde edilecek yarar istatistiksel modelden daha düşük seviyededir.

3.GEREÇ VE YÖNTEM

Çalışmada kullanılan gerçek veri seti, Çukurova Üniversitesi Tıp Fakültesi Balcalı Hastanesi'nin Yeni Doğan Yoğun Bakım Ünitesine, 2013-2015 Temmuz yılları arasında yatırılan tüm bebekleri içermektedir. Çalışma, tasarım olarak kesitsel araştırma niteliğindedir ve toplam 1281 yeni doğanı içermektedir. Yirmi dört saat içinde ölen veya taburcu edilen, yaşamla bağdaşmayan konjinetal malformasyonu olan, çalışma protokolü eksik bilgiler nedeniyle tamamlanamayan ve hasta dosyalarına ulaşılmayan yeni doğanlar çalışmaya alınmamıştır. Bu dışlama kriterleri nedeniyle çalışmaya 864'lük veri seti ile devam edilmiştir. Risk tahmin modeli olarak SNAP-II (Score for Neonatal Acute Physiology) ve SNAPPE-II (Score for Neonatal Acute Physiology-Perinatal Extension-II) kullanılmıştır. SNAP-II ve SNAPPE-II yeni doğan yoğun bakım ünitesine yatırılan bebeklerin ölüm olasılıklarını hesaplamak için kullanılmaktadır.

SNAP-II risk tahmin modeli 6 adet risk faktörünü içeren bir lojistik model iken SNAPPE-II risk tahmin modeli SNAP-II modeline 3 adet perinatal risk faktörünün eklenmesi ile elde edilir [35].

SNAP-II; ortalama kan basıncı, en düşük ısı, PO_2/FiO_2 oranı, en düşük serum pH, çok sayıda nöbet ve diürez olmak üzere toplam 6 adet doğum sonrası risk faktörlerini içermektedir. SNAPPE-II ise SNAP-II risk tahmin modeline doğum ağırlığı, SGA (small for gestational age) ve apgar skoru (5. dakika) değişkenlerinin ilave edilmesi ile elde edilir.

SNAP-II risk tahmin skorları, lojistik regresyonla elde edilen β katsayılarının 10 ile çarpılması ile elde edilir. SNAPPE-II ise SNAP-II ve 3 adet perinatal risk faktörlerini içeren bir lojistik modeldir (Çizelge 3.1).

Yeni eklenen faktör (ya da faktörler); bebeğin sürfaktan ve annenin antenatal kortikosteroid (celestone) tedavisini alıp almamasıdır.

Çizelge 2.1. SNAP-II ve SNAPPE-II risk tahmin modelinin beta katsayıları ve risk skor değerleri [35]

Değişkenler		**β**	**Odds Oranları**	**%95 GA**	**SNAP-II Değerleri**
Ortalama Kan Basıncı	**20-29 mm Hg**	0,88	2,40	1,86-3,10	9
	<20 mm Hg	1,94	6,99	4,04-12,08	19
En Düşük Isı	**35,6-35 (^{0}C)**	0,81	2,23	1,61- 3,10	8
	<35 (^{0}C)	1,55	4,69	2,99-7,24	15
PO_2/FiO_2 Oranı	**1,0-2,49**	0,49	1,63	1,20-2,21	5
	0,3-0,99	1,57	4,79	3,57-6,41	16
	<0,3	2,80	16,51	10,00-27,28	28
En Düşük Serum pH	**7,10-7,19**	0,71	2,02	1,42-2,87	7
	<7,10	1,57	4,81	3,27-7,07	16
Çok Sayıda Nöbet	**Var**	1,87	6,50	4,01-10,53	19
Diürez	**0,1-0,9 (ml/kg/s)**	0,46	1,59	1,22-2,06	5
	<0,1 (ml/kg/s)	1,82	6,15	4,31- 8,77	18
Sabit		-4,69			
SNAPPE-II için ek değişkenler Doğum Ağırlığı 750-999gr: +10 Doğum Ağırlığı <750gr: +17 SGA : +12 5.dakikadaki Apgar Skoru <7: +18					

Çalışmanın ilk aşamasında sürfaktan ve celestone tedavisinin tahmindeki etkinliği tek değişkeni olarak sınanmıştır. Tek değişkenli analizde anlamlı bulunan değişken (ya da değişkenler) modele ilave edilmiştir. Elde edilen yeni faktöre bağlı risk modeli ile eski modelin performanslarının kıyaslanması ROC-AUC endeksi, NRI kategorik, NRI sürekli ve IDI yaklaşımları ile yapılmıştır. Ayrıca risk tahmin modellerinin klinik yararlarını karşılaştırabilmek için karar eğrisi analizi kullanılmıştır.

Simülasyon çalışmasında, y bağımlı değişkeni; $n_1 = (\text{örneklem büyüklüğü} \times (1 - \text{olayın görülme sıklığı}))$ kadar 0 ve $n_2 = (\text{örneklem büyüklüğü} - n_1)$ kadar 1 üretilerek oluşturulmuştur. Bağımsız değişkenler ise çok değişkenli normal dağılım kullanılarak üç adet üretilmiş ve 3.bağımsız değişken yeni faktör olarak düşünülmüştür. Üretilen bağımsız değişkenlerin ortalaması hastalığı olmayan grup için sırasıyla (0,1,2) şeklinde iken hastalığı olan grup için sırasıyla (1,2,ort) şeklindedir. Burada hastalığı olan bireylere ait 3. değişkenin ortalaması varyasyon katsayılarındaki (coefficient of variation-CV) değişime bağlı olarak artırılmıştır. CV değerleri 1.00, 1.05, 1.10, 1.15 alınmıştır. Yani hastalığı olan bireylere ait 3. değişkenin ortalaması 2, 2.1, 2.2, 2.3 olarak alınmıştır. Üretilen üç adet bağımsız değişkenin varyans-kovaryans matrisi

$$\begin{bmatrix} 1{,}00 & 0{,}00 & 0{,}00 \\ 0{,}00 & 1{,}00 & 0{,}00 \\ 0{,}00 & 0{,}00 & 1{,}00 \end{bmatrix}$$

şeklindedir.

Simülasyon senaryosu; "Uygun (fit) bir modele anlamsız bir değişken eklenmesi durumunda kullanılan dört yöntemin hatalı pozitiflik oranını belirlemek ve farklı CV değerlerinde yöntemlerin p değerlerindeki değişimleri gözlemlemek" şeklindedir.

Uygun model çıkarsaması için "McFadden R^2 değerinin 0,20'den yüksek olması [36]" varsayımı kullanılırken anlamsız model çıkarsaması için "tek değişkenli analizde p değerinin 0,250'den yukarı olması" varsayımı kullanılmıştır [8]. Ayrıca örneklem büyüklükleri 100, 250, 500, 1000 ve olayın farklı görülme sıklıkları 0.10, 0.25, 0.50 alınarak 1000 tekrarlı simülasyonlarla bahsi geçen 4 yaklaşım değerlendirilmiştir.

Verilerin analizinde SPSS versiyon 20[37] ve R versiyon 3.2.3[38] paket programlarından yararlanılmıştır.

4. BULGULAR

4.1. Gerçek Veri Setine Ait Bulgular

İlk olarak yeni doğan yoğun bakım ünitesine yatırılan bebeklerin yatış sonlanma durumuna göre SNAP-II ve SNAPPE-II risk tahmin modellerini oluşturan faktörlere ve demografik özelliklerine bağlı tanımlayıcılar Çizelge 4.1.1'de verilmiştir. 2013-2015 yılları arasında yatırılan bebeklerde ölüm olasılığı %14,6 olarak bulunmuştur.

Çizelge 4.1.1. SNAP-II ve SNAPPE-II risk tahmin modellerini oluşturan faktörlere ve demografik özelliklerine bağlı tanımlayıcılar

Değişkenler		**Yatış Sonlanma Durumu**		**p-değeri**
		Hayatta n(%)	**Ölü n(%)**	
Cinsiyet	**Erkek**	413(56,0)	68(54,0)	0,699
	Kız	325(44,0)	58(46,0)	
Doğum Şekli	**NVYS**	189(25,6)	26(20,6)	0,265
	C/S	549(74,4)	100(79,4)	
Celestone Kullanımı	**Var**	79(10,7)	44(34,9)	<0,001
	Yok	659(89,3)	82(65,1)	
Sürfaktan Kullanımı	**Var**	39(5,3)	68(54,0)	<0,001
	Yok	699(94,7)	58(46,0)	
Ortalama Kan Basıncı (mm Hg)	**>=30**	717(97,2)	98(77,8)	<0,001
	20-29	17(2,3)	9(7,1)	
	<20	4(0,5)	19(15,1)	
En Düşük Vücut Isısı (^oC)	**>35,6**	651(88,2)	63(50,0)	<0,001
	35-35,6 ^oC	65(8,8)	30(23,8)	
	<35 ^oC	22(3,0)	33(26,2)	
PO_2/FiO_2 oranı	**>2,49**	120(16,3)	10(7,9)	<0,001
	1,0-2,49	574(77,8)	64(50,8)	
	0,3-0,99	42(5,7)	37(29,4)	
	<0.3	2(0,3)	15(11,9)	
En Düşük Serum pH	**>=7,20**	634(85,9)	53(42,1)	<0,001
	7,10-7,19	76(10,3)	28(22,2)	
	<7,10	28(3,8)	45(35,7)	
Çoklu Nöbet	**Var**	11(1,5)	13(10,3)	<0,001
	Yok	727(98,5)	113(89,7)	
Diürez Miktarı (mL/kg/s)	**>0,9**	678(91,9)	76(60,3)	<0,001
	0,1-0,9	52(7,0)	30(23,8)	
	<0,1	8(1,1)	20(15,9)	
SGA	**Var**	14(1,9)	16(12,7)	<0,001
	Yok	724(98,1)	110(87,3)	
Doğum Ağırlığı	**>=1000**	733(99,3)	76(60,3)	<0,001
	750-999	5(0,7)	20(15,9)	
	<750	0(0,0)	30(23,8)	
Apgar (5. dk)	**>=7**	653(88,5)	50(39,7)	<0,001
	<7	85(11,5)	76(60,3)	
Anne Yaşı†		29,83±6,360	28,79±6,34	0,613
Doğum Haftası†		35,60±3,32	30,96±5,47	0,502

† Değişkenler için ortalama ± standart sapma değerleri verilmiştir.

4.1.1. SNAP-II Risk Tahmin Modelinin Performansının Değerlendirilmesi

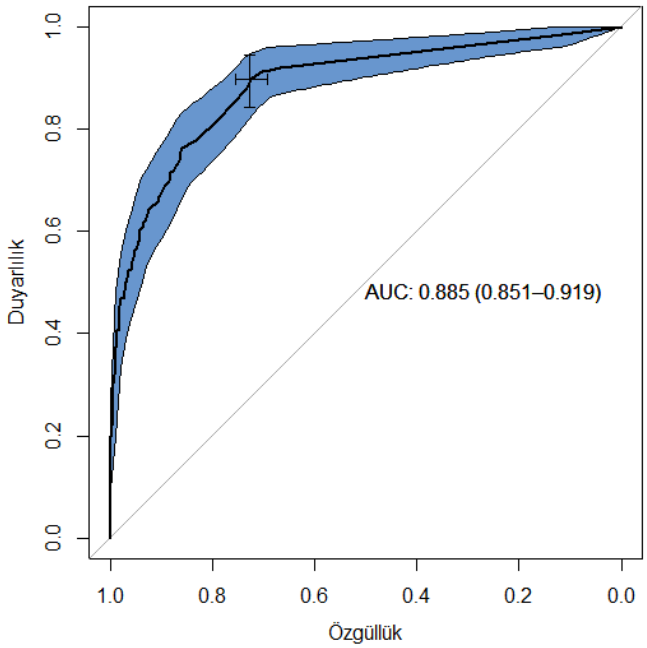
Tek değişkenli analizde SNAP-II risk tahmin modelini oluşturan tüm bağımsız değişkenler ile yatış sonlanma durumu arasındaki ilişki istatistiksel olarak anlamlı bulunmuştur. Buradan hareketle çoklu lojistik regresyon analizi uygulanmış ve Çizelge 4.1.2'deki sonuçlar elde edilmiştir.

Çizelge 4.1.2. SNAP-II risk tahmin modelini oluşturan bağımsız değişkenlerin katsayıları ve risk skorları

Değişkenler		**β**	**Odds Oranları**	**%95 GA**	**SNAP-II Skor Değerleri**
Ortalama Kan Basıncı	**20-29 mm Hg**	1,241	3,460	1,198-9,993	12
	<20 mm Hg	2,026	7,587	1,682-34,216	20
En Düşük Vücut Isısı	**35-35,6 ^{o}C**	1,301	3,675	1,952-6,919	13
	<35 ^{o}C	2,109	8,242	3,749-18,119	21
PO_2/FiO_2 oranı	**1,0-2,49**	0,394	1,482	0,649-3,383	4
	0,3-0,99	1,891	6,625	2,539-17,288	19
	<0,3	3,183	24,113	3,323-174,995	32
En Düşük Serum pH	**7,10-7,19**	1,270	3,561	1,907-6,650	13
	<7,10	2,202	9,040	4,525-18,058	22
Çoklu Nöbet	**Var**	1,932	6,904	2,273-20,965	19
Diürez Miktarı	**0,1-0,9 mL/kg/s**	0,762	2,142	1,118-4,105	8
	<0,1 mL/kg/s	1,736	5,674	1,610-19,992	17
Sabit		-3,801			

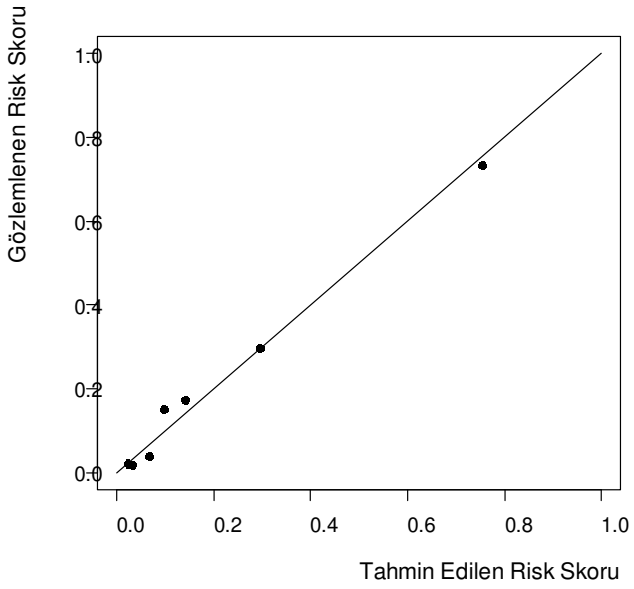
SNAP-II risk tahmin modelinin genel performansını değerlendirmek için kullanılan yöntemlerden Brier Skoru 0,076 ve Nagelkerke R^2 0,479 olarak bulunmuştur.

Ayırım yeteneğini değerlendirmek için kullanılan ROC-AUC değeri 0,885 (GA: 0,851-0,919)'dir (Şekil 4.1.1). Kesim noktası 14,50 olarak belirlenmiş ve bu noktadaki seçicilik 0,725, duyarlılık 0,897 ve doğruluk 0,750 olarak elde edilmiştir.



Şekil 4.1.1. SNAP-II risk tahmin modelinin ROC eğrisi

Kalibrasyonu değerlendirmek için kullanılan Hosmer-Lemeshow testinin p değeri 0,619 (χ^2=6,249) olarak bulunmuştur. Yani gözlem değerleri ile tahmin değerleri arasındaki fark istatistiksel olarak anlamlı değildir. Elde edilen kalibrasyon eğrisi Şekil 4.1.2'de verildiği gibidir.



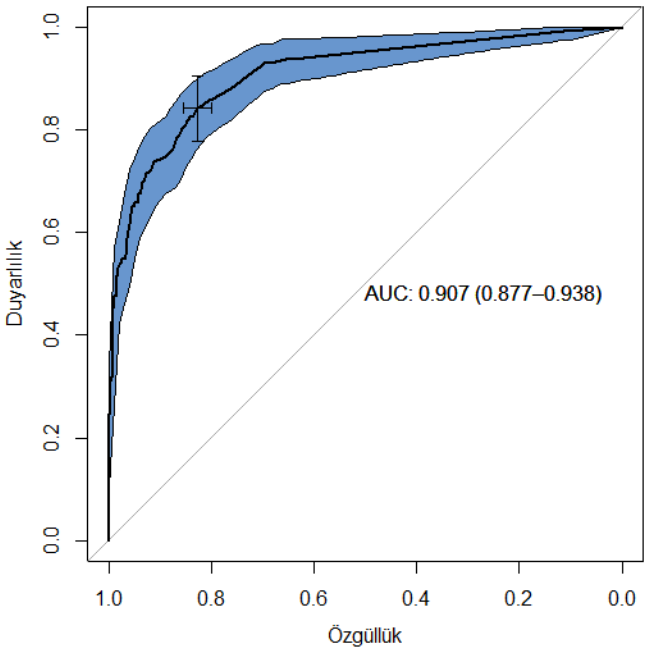
Şekil 4.1.2. SNAP-II risk tahmin modelinin kalibrasyon eğrisi

SNAP-II risk tahmin modeline sürfaktan bilgisinin eklenmesi ile elde edilen katsayılar Çizelge 4.1.3'te verildiği gibidir.

Çizelge 4.1.3. SNAP-II risk tahmin modeline sürfaktan bilgisinin eklenmesi ile elde edilen modelin bağımsız değişkenlerinin katsayıları ve risk skorları

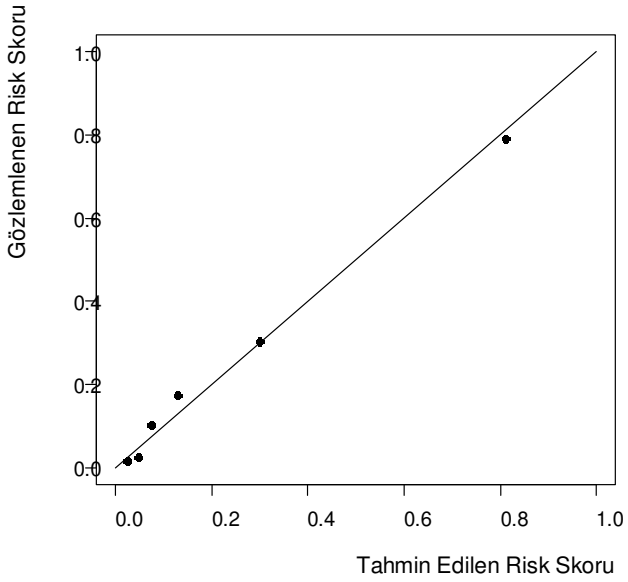
Değişkenler		**β**	**Odds Oranları**	**%95 GA**	**SNAP-II Skor Değerleri**
Ortalama Kan Basıncı	**20-29 mm Hg**	0,599	1,820	0,581-5,694	6
	<20 mm Hg	1,429	4,175	0,723-24,123	14
En Düşük Vücut Isısı	**35-35,6 ^{o}C**	1,166	3,210	1,607-6,413	12
	<35 ^{o}C	1,453	4,274	1,795-10,180	15
PO_2/FiO_2 oranı	**1.0-2,49**	0,434	1,543	0,661-3,601	4
	0,3-0,99	1,691	5,423	1,990-14,777	17
	<0,3	2,321	10,182	1,194-86,827	23
En Düşük Serum pH	**7,10-7,19**	1,137	3,118	1,614-6,025	11
	<7,10	2,074	7,956	3,745-16,902	21
Çoklu Nöbet	**Var**	1,836	6,274	1,979-19,887	18
Diürez Miktarı	**0,1-0,9 mL/kg/s**	0,645	1,906	0,922-3,938	6
	<0,1 mL/kg/s	2,208	9,099	2,557-32,376	22
Sürfaktan	**Var**	2,164	8,707	4,686-16,177	22
Sabit		-4,071			

SNAP-II risk tahmin modeline sürfaktan bilgisinin eklenmesiyle Brier Skoru 0,067 olarak bulunurken Nagelkerke R^2 değeri 0,546 olarak bulunmuştur. Ayırım yeteneğini değerlendirmek için kullanılan ROC-AUC değeri 0,907 (GA: 0,877-0,938)'dir (Şekil 4.1.3). Kesim noktası 18,50 olarak belirlenmiş ve bu noktadaki seçicilik 0,828, duyarlılık 0,841 ve doğruluk 0,830 olarak elde edilmiştir.



Şekil 4.1.3. SNAP-II risk tahmin modeline sürfaktan bilgisinin eklenmesi ile elde edilen modelin ROC eğrisi

Son olarak kalibrasyonu değerlendirmek için kullanılan Hosmer-Lemeshow testinin p değeri 0,827 (χ^2=4,327) olarak bulunmuştur. Kalibrasyon eğrisi Şekil 4.1.4'te gösterildiği gibidir.



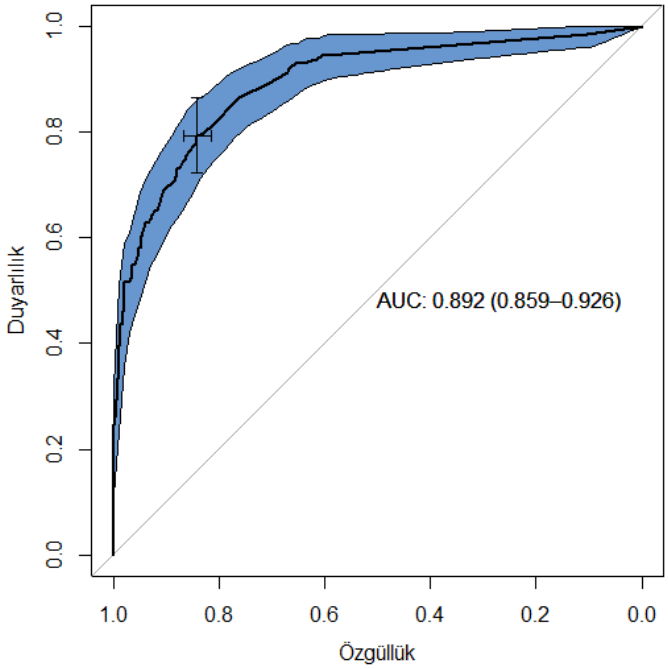
Şekil 4.1.4. SNAP-II risk tahmin modeline sürfaktan bilgisinin eklenmesi ile elde edilen modelin kalibrasyon eğrisi

SNAP-II risk tahmin modeline celestone bilgisi eklenmiş ve Çizelge 4.1.4'teki katsayılar elde edilmiştir.

Çizelge 4.1.4. SNAP-II risk tahmin modeline celestone bilgisinin eklenmesi ile elde edilen modelin katsayıları ve risk skorları

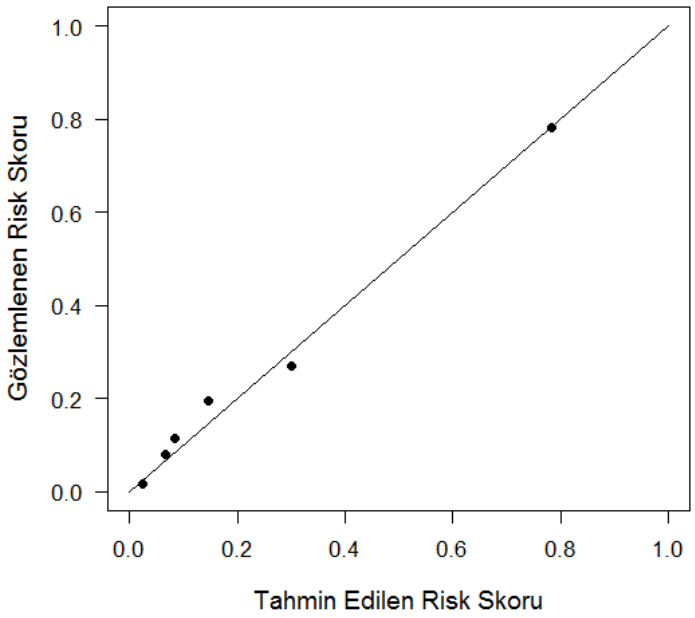
Değişkenler		**β**	**Odds Oranları**	**%95 GA**	**SNAP-II Skor Değerleri**
Ortalama Kan Basıncı	**20-29 mm Hg**	0,932	2,539	0,864-7,458	9
	<20 mm Hg	1,692	5,429	1,069-27,561	17
En Düşük Vücut Isısı	**35-35,6 °C**	1,116	3,052	1,582-5,888	11
	<35 °C	1,871	6,434	2,876-14,665	19
PO_2/FiO_2 oranı	**1,0-2,49**	0,308	1,361	0,594-3,118	3
	0,3-0,99	1,810	6,110	2,331-16,015	18
	<0,3	3,147	23,256	3,145-171,965	31
En Düşük Serum pH	**7,10-7,19**	1,214	3,368	1,780-6,371	12
	<7,10	2,302	9,994	4,928-20,266	23
Çoklu Nöbet	**Var**	2,007	7,439	2,400-23,052	20
Diürez Miktarı	**0,1-0,9 mL/kg/s**	0,822	2,275	1,173-4,412	8
	<0,1 mL/kg/s	1,738	5,689	1,548-20,899	17
Celestone	**Var**	1,193	3,298	1,790-6,074	12
Sabit		-3,906			

SNAP-II risk tahmin modeline celestone bilgisinin eklenmesi ile Brier Skoru 0,073 ve Nagelkerke R^2 değeri 0,499 olarak bulunmuştur. ROC-AUC değeri 0,892 (GA: 0,859-0,926) olarak bulunmuş ve kesim noktası 19,50 bulunmuştur. Bu kesim noktasındaki seçicilik 0,838, duyarlılık 0,794 ve doğruluk 0,832 olarak elde edilmiştir (Şekil 4.1.5).



Şekil 4.1.5. SNAP-II risk tahmin modeline celestone bilgisinin eklenmesi ile elde edilen modelin ROC eğrisi

Son olarak modelin kalibrasyonunu değerlendirmek için kullanılan Hosmer-Lemeshow testinin p değeri 0,781 (χ^2=4,775) olarak bulunmuş ve kalibrasyon eğrisi Şekil 4.1.6'ta gösterildiği gibi elde edilmiştir.



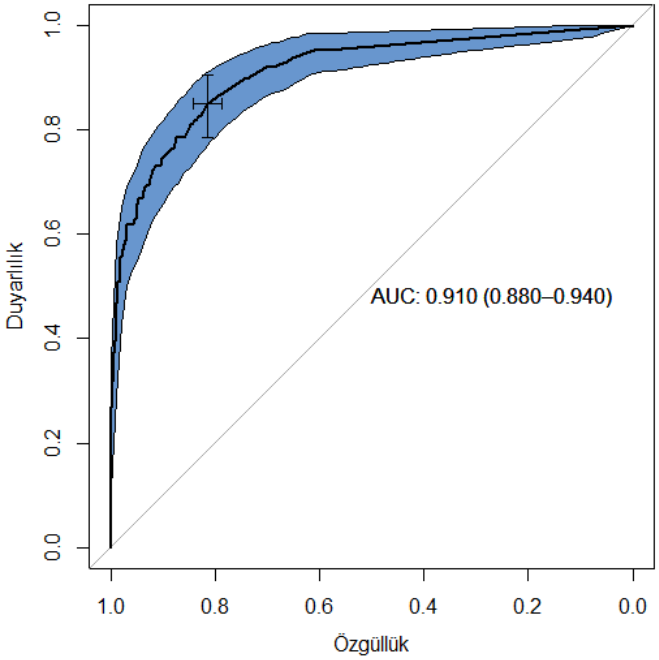
Şekil 4.1.6. SNAP-II risk tahmin modeline celestone bilgisinin eklenmesi ile elde edilen modelin kalibrasyon eğrisi

Son olarak SNAP-II risk tahmin modeline sürfaktan ve celestone bilgisi aynı anda eklenmiş ve Çizelge 4.1.5'teki katsayılar elde edilmiştir.

Çizelge 4.1.5. SNAP-II risk tahmin modeline sürfaktan ve celestone bilgisinin eklenmesi ile elde edilen modelin katsayıları ve risk skorları

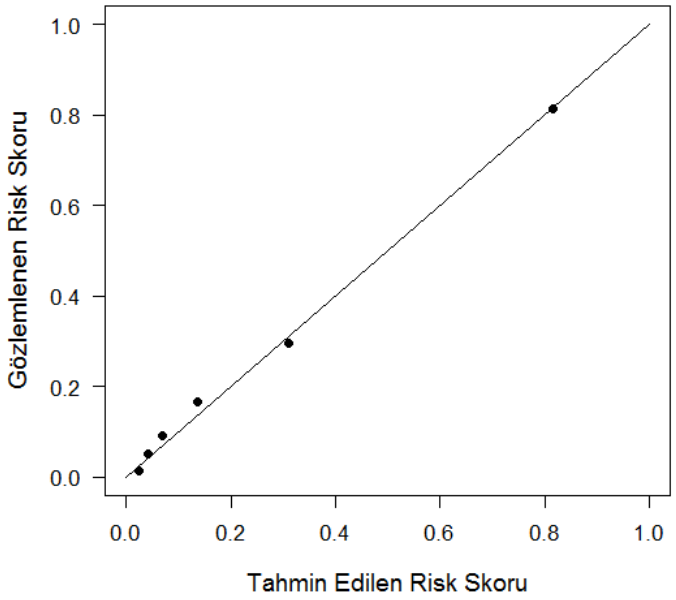
Değişkenler		**β**	**Odds Oranları**	**%95 GA**	**SNAP-II Skor Değerleri**
Ortalama Kan Basıncı	**20-29 mm Hg**	0,501	1,650	0,519-5,246	5
	<20 mm Hg	1,340	3,818	0,618-23,588	13
En Düşük Vücut Isısı	**35-35,6 °C**	1,063	2,895	1,428-5,870	11
	<35 °C	1,374	3,949	1,630-9,565	14
PO_2/FiO_2 oranı	**1,0-2,49**	0,377	1,458	0,624-3,404	4
	0,3-0,99	1,675	5,338	1,961-14,533	17
	<0,3	2,364	10,633	1,258-89,882	24
En Düşük Serum pH	**7,10-7,19**	1,131	3,099	1,599-6,010	11
	<7,10	2,124	8,364	3,941-17,753	21
Çoklu Nöbet	**Var**	1,898	6,670	2,094-21,248	19
Diürez Miktarı	**0,1-0,9 mL/kg/s**	0,693	2,000	0,969-4,128	7
	<0,1 mL/kg/s	2,180	8,847	2,453-31,911	22
Sürfaktan	**Var**	1,978	7,230	3,765-13,884	20
Celestone	**Var**	0,604	1,829	0,910-3,678	6
Sabit		-4,088			

Hem sürfaktan hem de celestone bilgisinin modele ilave edilmesi ile Brier Skoru 0,066 ve Nagelkerke R^2 0,550 olarak bulunmuştur. Ayırım yeteneğini test etmek için kullanılan ROC-AUC değeri 0,910 (GA: 0,880-0,940) olarak elde edilmiştir. Risk skorlarına dayanan kesim noktası değeri 20,50 ve bu noktadaki seçicilik 0,816, duyarlılık 0,849 ve doğruluk 0,821 olarak elde edilmiştir (Şekil 4.1.7).



Şekil 4.1.7. SNAP-II risk tahmin modeline sürfaktan ve celestone bilgisinin eklenmesi ile elde edilen modelin ROC eğrisi

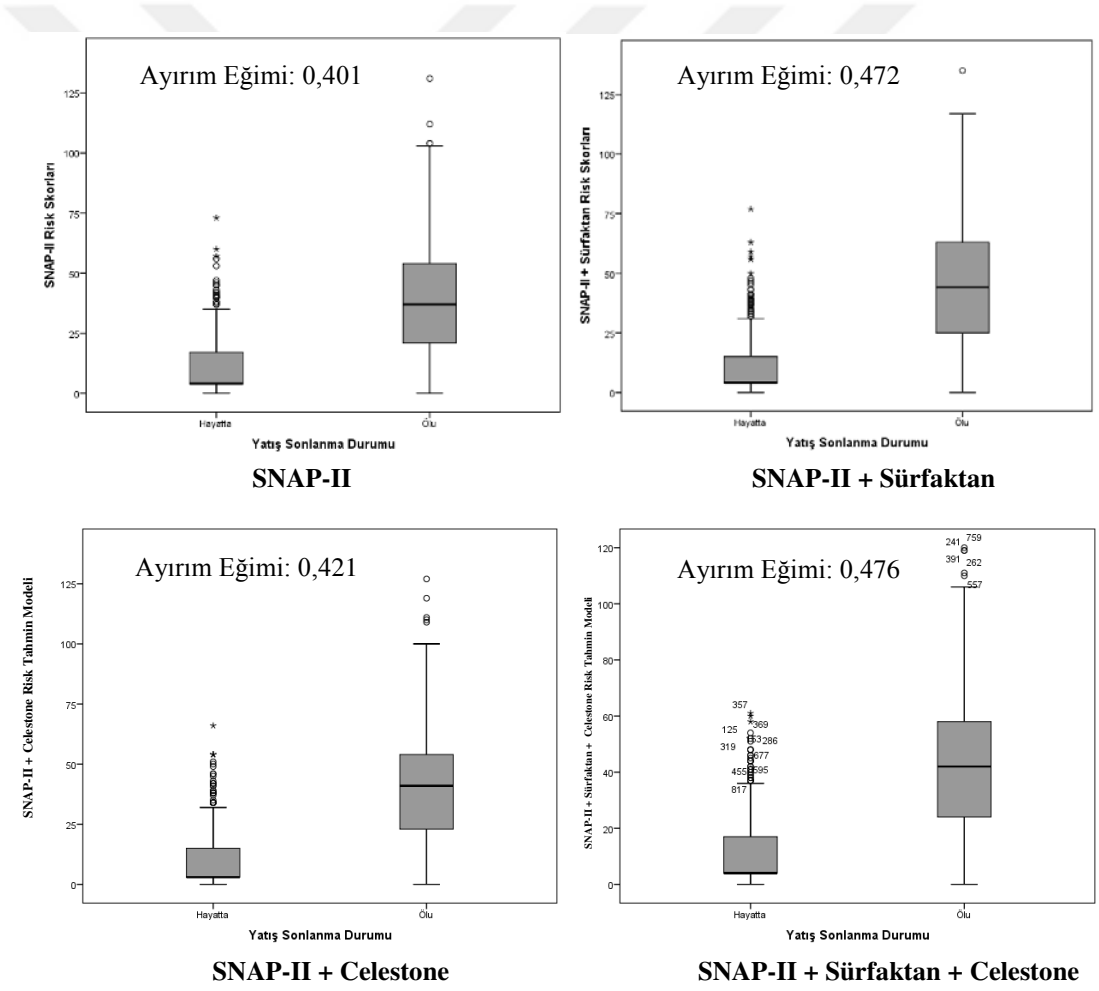
Modelin kalibrasyon eğrisi Şekil 4.1.8'de verildiği gibidir ve test istatistiğinin p değeri 0,927 (χ^2=3,116) olarak bulunmuştur.



Şekil 4.1.8. SNAP-II risk tahmin modeline sürfaktan ve celestone bilgisinin eklenmesi ile elde edilen modelin kalibrasyon eğrisi

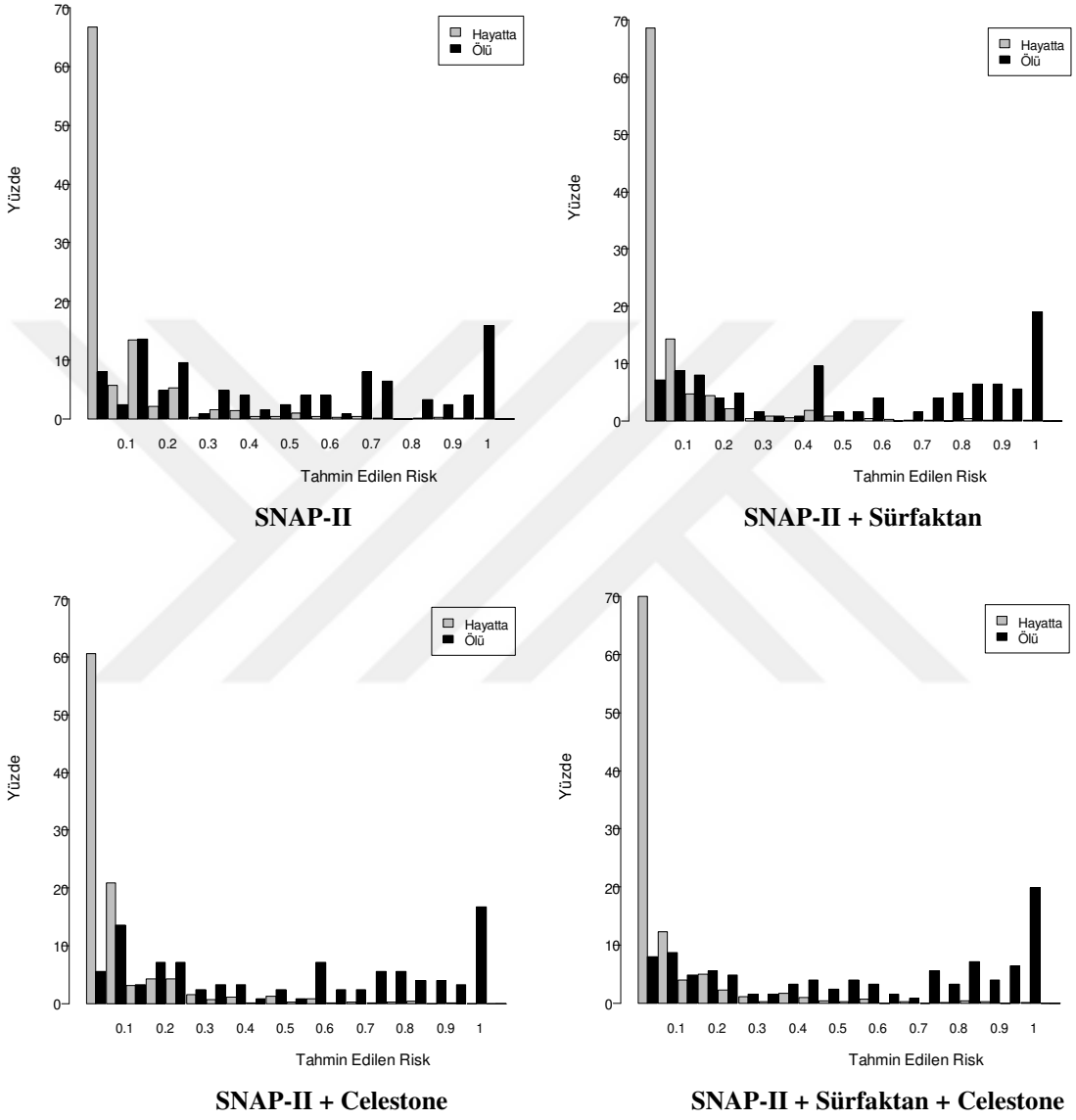
Celestone ve sürfaktanın etkileşim terimlerinin tek değişkenli analizinde p değeri 0,899 olarak bulunmuştur. Buna bağlı olarak modele ilavesi istatistiksel olarak uygun görülmemiştir.

Ayırım yeteneğinin görsel olarak değerlendirebilmek için kullanılan bir diğer grafik ise boxplottur. Her bir durum için çizilen boxplotlar Şekil 4.1.9'da gösterildiği gibidir. Bu grafiklerden ilki sadece SNAP-II risk skorunu, ikincisi SNAP-II + Sürfaktan risk skorunu, üçüncüsü SNAP-II + Celestone risk skorunu ve dördüncüsü SNAP-II + Sürfaktan + Celestone risk skorunun yatış sonlanma durumuna göre dağılımını ve ayırım eğimlerinin değerini göstermektedir.



Şekil 4.1.9. SNAP-II için her bir modelin risk skorlarının yatış sonlanma durumuna göre boxplotları

Tahmin edilen risk değerlerinin hayatta kalan ve kalamayan bebeklerdeki risk skorlarının dağılım yüzdelerini gösteren risk dağılım grafiği Şekil 4.1.10'da verildiği gibidir.



Şekil 4.1.10. SNAP-II için her bir modelin risk değerlerinin yatış sonlanma durumuna göre dağılım grafiği

Sonuç olarak SNAP-II, SNAP-II + Sürfaktan, SNAP-II + Celestone ve SNAP-II + Sürfaktan + Celestone risk tahmin modellerinin genel performanslarına ve ayırım yeteneklerine göre yüksekten düşüğe doğru sıralamaları;

- SNAP-II + Sürfaktan + Celestone,
- SNAP-II + Sürfaktan,
- SNAP-II + Celestone,
- SNAP-II

şeklindedir. Her durumda ROC-AUC değerleri SNAP-II'den daha yüksektir. Kalibrasyon için kullanılan Hosmer-Lemeshow testinin p değeri tüm risk tahmin modellerinde 0,05'ten yüksektir. Yani modellerin tahmin değerleri ile beklenen değerleri arasındaki fark istatistiksel olarak anlamlı değildir.

Ayrıca bulunan bu sonuçlar grafiklerle de desteklenmiştir. SNAP-II + Sürfaktan + Celestone modelinin ayırım eğimi en yüksek ve tahmin edilen risk değerleri sağlıklılarda düşük, hastalarda yüksek yüzdelerde seyretmektedir.

Bu durumda SNAP-II risk tahmin modeline Sürfaktan ve Celestone bilgisinin aynı anda eklenmesi ile elde edilen modelin performansı diğer yöntemlerden daha yüksektir. Bu modelin performansının istatistiksel olarak değerlendirilmesi Bölüm 4.1.3'te verilecektir.

4.1.2. SNAPPE-II Risk Tahmin Modelinin Performansının Değerlendirilmesi

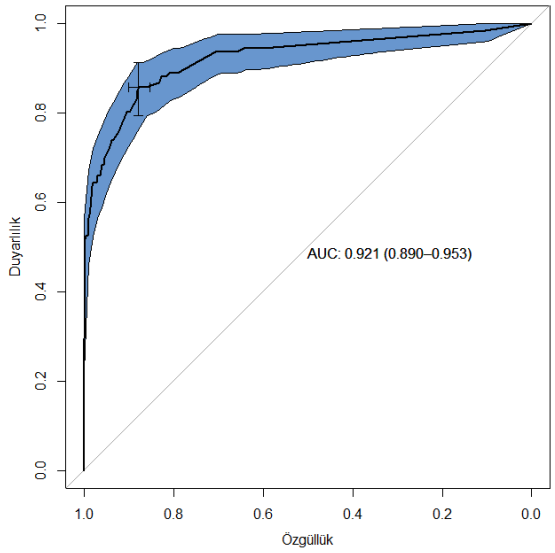
SNAP-II risk tahmin modeline eklenen 3 adet perinatal risk faktörleri ile SNAPPE-II risk tahmin modeli elde edilir. SNAP-II risk skorları ve bu 3 adet perinatal risk faktörünü içeren lojistik regresyonla tahmin değerleri elde edilmektedir[39]. Buradan hareketle, doğrusal regresyona dayanarak elde edilen 3 değişkene ait katsayı tahminleri ve risk skorları Çizelge 4.1.6'da verildiği gibidir.

Çizelge 4.1.6. SNAP-II risk tahmin modeline eklenen 3 adet risk faktörünün katsayıları ve risk skorları

Değişkenler		Tahmin Değeri	p değeri	Risk Skoru
Doğum Ağırlığı	**750-999**	3,102	<0,001	31
	<750	3,248	<0,001	32
SGA	**Var**	1,638	0,004	16
5.dakikadaki Apgar Skoru	**<7**	1,265	<0,001	13

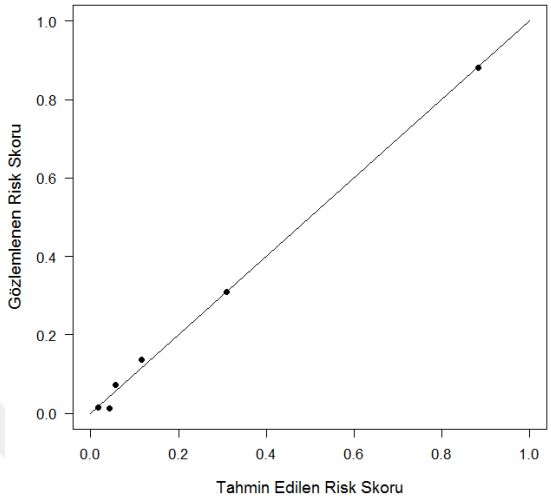
SNAP-II risk tahmin modelindeki değişkenlerin skorlama değerlerinin üzerine bu 3 değişkenin skor değerleri eklenerek SNAPPE-II risk skoru elde edilir.

SNAPPE-II risk tahmin modeline ait Brier skoru 0,056 iken Nagelkerke R^2 değeri 0,619 olarak bulunmuştur. Ayırım yeteneğini değerlendirmek için kullanılan ROC-AUC değeri 0,921 (GA: 0,890-0,953)'dir (Şekil 4.1.11). Kesim noktası 26,50 olarak belirlenmiş ve bu noktadaki seçicilik 0,877, duyarlılık 0,857 ve doğruluk 0,874 olarak elde edilmiştir.



Şekil 4.1.11. SNAPPE-II risk skoruna bağlı ROC eğrisi

Hosmer-Lemeshow testinin p değeri 0,968 (χ^2=2,338) olarak bulunmuştur. Yani gözlem değerleri ile tahmin değerleri arasındaki fark istatistiksel olarak anlamlı değildir. Elde edilen kalibrasyon eğrisi Şekil 4.1.12'de verildiği gibidir.



Şekil 4.1.12. SNAPPE-II risk skoruna bağlı kalibrasyon eğrisi

İlk olarak sürfaktan bilgisi modele yeni değişken olarak eklenmiştir. SNAP-II risk tahmin modeline eklenen 3 adet perinatal risk faktörleri ile SNAPPE-II risk tahmin modelinin geliştirilmesi için izlenen yol ilk olarak kurulan modelle aynıdır. Bulunan tahmin ve risk skor değerleri Çizelge 4.1.7'de verildiği gibidir.

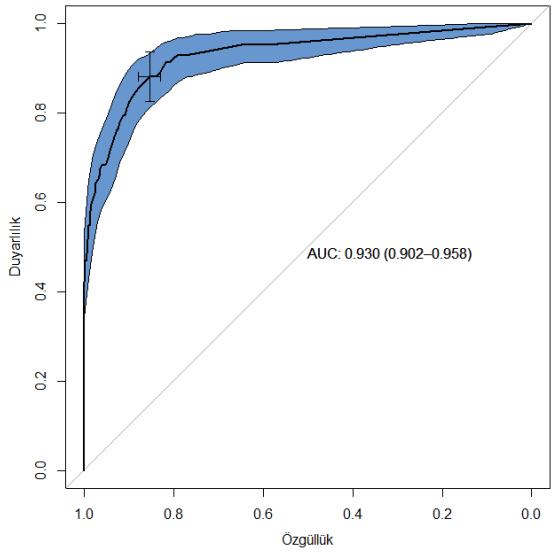
Çizelge 4.1.7. SNAP-II risk tahmin modeline sürfaktan bilgisinin eklenmesinin ardından 3 adet risk faktörünün katsayıları ve risk skorları

Değişkenler		**Tahmin Değeri**	**p değeri**	**Risk Skoru**
Doğum Ağırlığı	**750-999**	2,647	<0,001	26
	<750	2,416	0,004	24
SGA	**Var**	1,630	0,005	16
5.dakikadaki Apgar Skoru	**<7**	1,124	<0,001	11

SNAP-II risk tahmin modeline sürfaktan bilgisinin eklenmesi ile elde edilen değişken skorlarına üç adet perinatal risk değişkenlerinin skor değerlerinin eklenmesi ile SNAPPE-II risk skorları elde edilir.

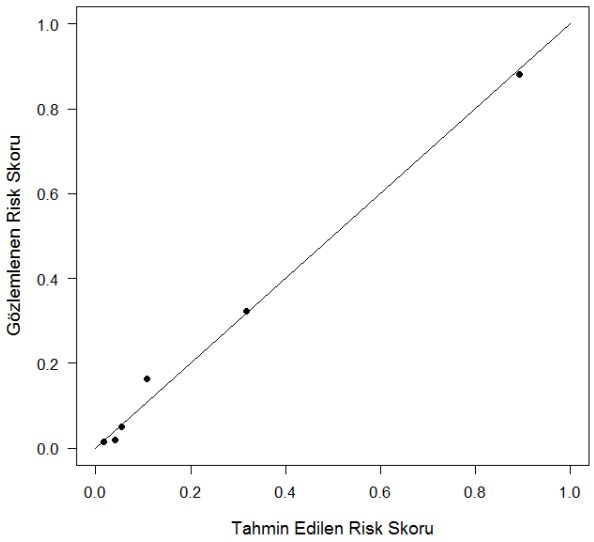
SNAPPE-II risk tahmin modeline sürfaktan bilgisinin eklenmesiyle Brier Skoru 0,057 olarak bulunurken Nagelkerke R^2 değeri 0,626 olarak bulunmuştur. Ayırım yeteneğini değerlendirmek için kullanılan ROC-AUC değeri 0,930 (GA: 0,902-

0,958)'dur (Şekil 4.1.13). Kesim noktası 25,50 olarak belirlenmiş ve bu noktadaki seçicilik 0,854, duyarlılık 0,881 ve doğruluk 0,858 olarak elde edilmiştir.



Şekil 4.1.13. SNAPPE-II risk tahmin modeline eklenen sürfaktan bilgisine bağlı ROC eğrisi

Kalibrasyonun değerlendirilmesi için kullanılan Hosmer Lemeshow testinin p değeri 0,881 (χ^2=4,486) olarak bulunmuştur. Kalibrasyon eğrisi Şekil 4.1.14'te gösterildiği gibidir. Modelin kalibrasyonu istatistiksel olarak geçerlidir. Yani tahmin değeri ile gözlem değeri arasındaki fark istatistiksel olarak anlamlı değildir.



Şekil 4.1.14. SNAPPE-II risk tahmin modeline eklenen sürfaktan bilgisine bağlı kalibrasyon eğrisi

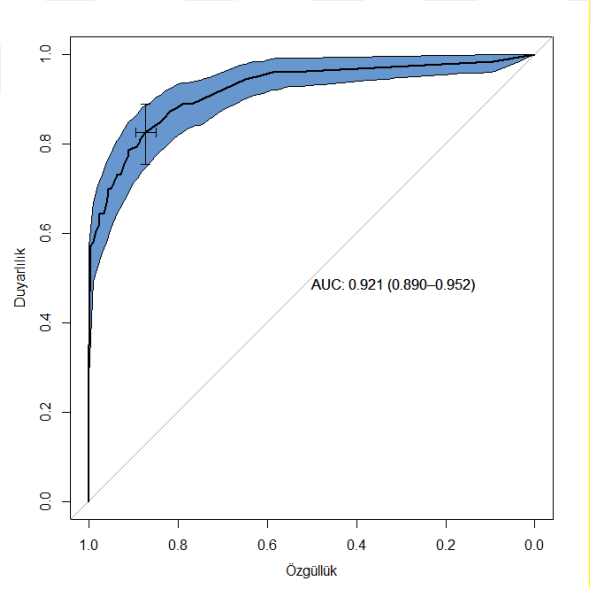
İkinci olarak celestone bilgisi modele yeni değişken olarak eklenmiş ve perinatal risk faktörlerine ait katsayılar Çizelge 4.1.8'de gösterildiği gibi elde edilmiştir.

Çizelge 4.1.8. SNAP-II risk tahmin modeline celestone bilgisinin eklenmesinin ardından 3 adet risk faktörünün katsayıları ve risk skorları

Değişkenler		Tahmin Değeri	p değeri	Risk Skoru
Doğum Ağırlığı	**750-999**	2,774	<0,001	28
	<750	2,822	0,001	28
SGA	**Var**	1,742	0,002	17
5.dakikadaki Apgar Skoru	**<7**	1,207	<0,001	12

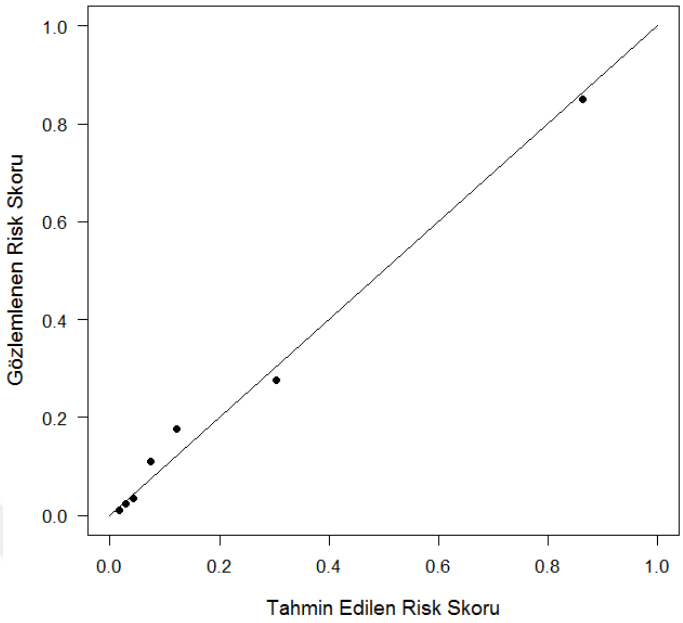
SNAP-II risk tahmin modeline eklenen celestone bilgisinin uzantısı olarak elde edilen SNAPPE-II risk tahmin modelinin Brier Skoru 0,057 olarak bulunurken Nagelkerke R^2 değeri 0,615 olarak bulunmuştur.

Ayırım yeteneğini değerlendirmek için kullanılan ROC-AUC değeri 0,921 (GA: 0,890-0,952)'dir (Şekil 4.1.15). Kesim noktası 26,50 olarak belirlenmiş ve bu noktadaki seçicilik 0,873, duyarlılık 0,825 ve doğruluk 0,866 olarak elde edilmiştir.



Şekil 4.1.15. SNAPPE-II risk tahmin modeline eklenen celestone bilgisine bağlı ROC eğrisi

Son olarak celestone bilgisinin SNAPPE-II modeline dahil edilmesinin ardından geliştirilen modelin kalibrasyonu değerlendirilmiştir. Hosmer-Lemeshow test istatistiğinin p değeri 0,941 (χ^2=2,899) şeklinde bulunmuş ve kalibrasyon eğrisi Şekil 4.1.16'da gösterilmiştir.



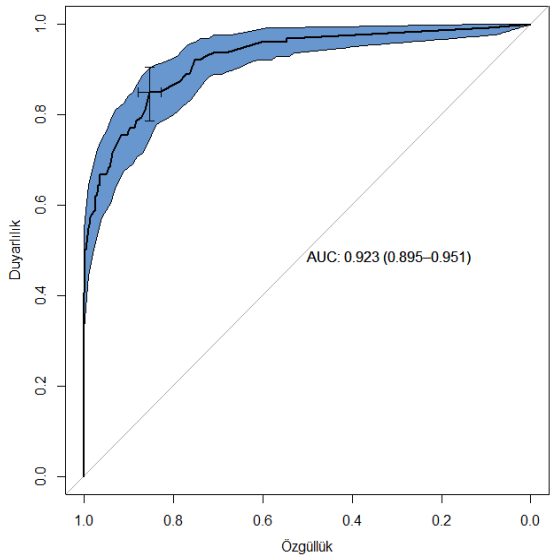
Şekil 4.1.16. SNAPPE-II risk tahmin modeline eklenen celestone bilgisine bağlı kalibrasyon eğrisi

Hem sürfaktan hem de celestone bilgisinin modele ilave edilmesi ile elde edilen SNAP-II risk tahmin modelinin SNAPPE-II modeline uzatılabilmesi için gerekli olan 3 değişkenin katsayıları ve risk skorları Çizelge 4.1.9'da verildiği gibidir.

Çizelge 4.1.9. SNAP-II risk tahmin modeline sürfaktan ve celestone bilgisinin eklenmesinin ardından 3 adet risk faktörünün katsayıları ve risk skorları

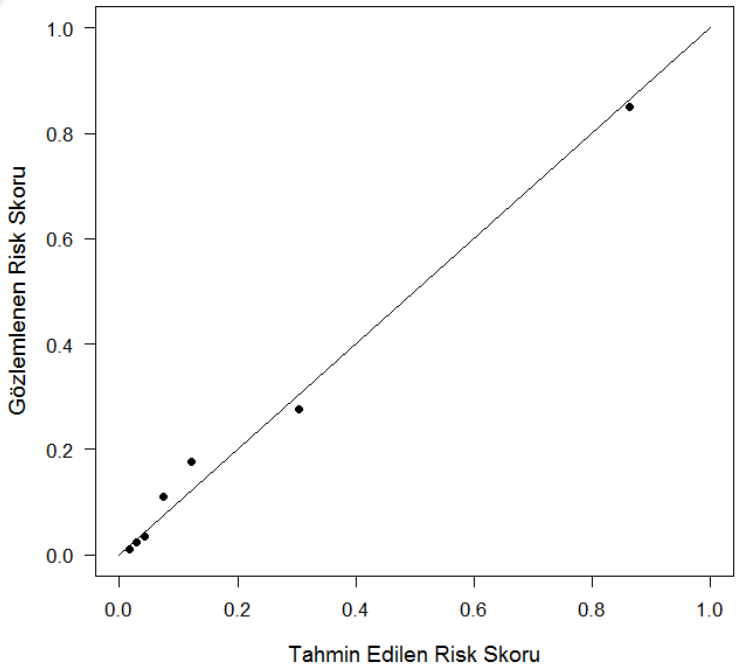
Değişkenler		**Tahmin Değeri**	**p değeri**	**Risk Skoru**
Doğum Ağırlığı	**750-999**	2,516	<0,001	25
	<750	2,440	0,003	24
SGA	**Var**	1,765	0,001	18
5.dakikadaki Apgar Skoru	**<7**	1,186	<0,001	12

Sürfaktan ve celestone'a bağlı modelin Brier Skoru 0,059 ve Nagelkerke R^2 değeri 0,603 olarak bulunmuştur. Ayrıca ayırım gücünü değerlendirmek için kullanılan ROC-AUC değeri 0,923 (GA:0,895-0,951) olarak bulunmuştur (Şekil 4.1.17). Kesim noktası 26,50 iken bu noktadaki seçicilik 0,854, duyarlılık 0,849 ve doğruluk 0,853'dür.



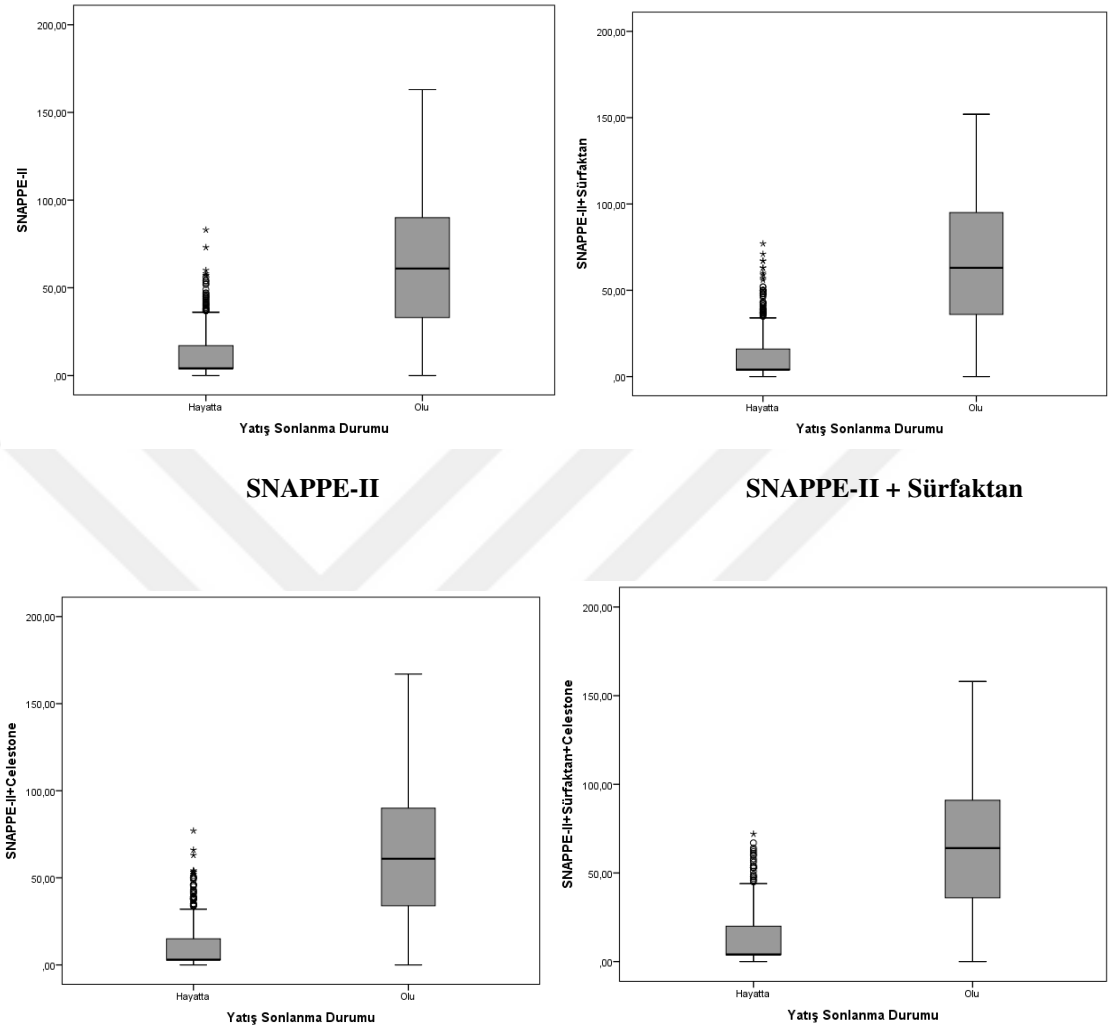
Şekil 4.1.17. SNAPPE-II risk tahmin modeline eklenen sürfaktan ve celestone bilgisine bağlı ROC eğrisi

Hosmer-Lemeshow testinin p değeri 0,479 (χ^2=7,542) olarak bulunmuştur. Kalibrasyon eğrisi ise Şekil 4.1.18'de verildiği gibidir.



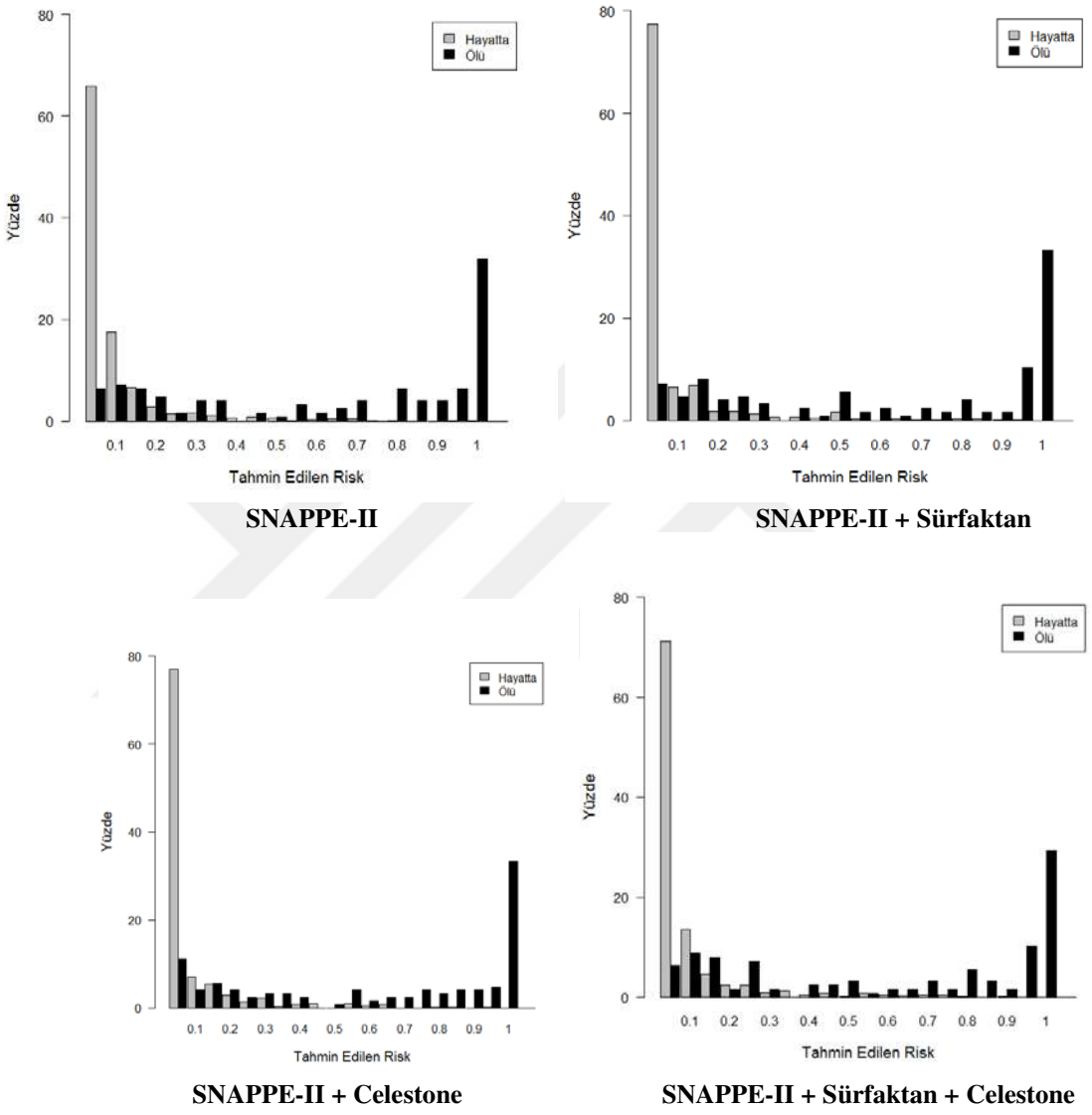
Şekil 4.1.18. SNAPPE-II risk tahmin modeline eklenen sürfaktan ve celestone bilgisine bağlı kalibrasyon eğrisi

Modelin ayırım yeteneğini görsel olarak değerlendirmek için kullanılan boxplot grafiği her bir model için Şekil 4.1.19'da verildiği gibidir.



Şekil 4.1.19. SNAPPE-II için her bir modelin risk skorlarının yatış sonlanma durumuna göre boxplotları

Tahmin edilen risk değerlerinin hayatta kalan ve kalamayan bebeklerdeki risk skorlarının dağılım yüzdelerini gösteren risk dağılım grafiği Şekil 4.1.20'de verildiği gibidir.



Şekil 4.1.20. SNAPPE-II için her bir modelin risk değerlerinin yatış sonlanma durumuna göre dağılım grafiği

Sonuç olarak SNAPPE-II modeline, sürfaktan bilgisinin eklenmesi ile Nagelkerke R^2 değeri ve ROC-AUC değeri en yüksek seviyeye sahip iken SNAPPE-II risk tahmin modeline yeni faktörlerin eklenmesi ile Brier skoru ve Hosmer-Lemeshow istatistiğinin p değeri diğer modellerden daha iyi bir performansa sahiptir.

Ayrıca bulunan bu sonuçlar grafiklerle de desteklenmiştir. SNAP-II + Sürfaktan modelinin ayırım eğimi en yüksek ve tahmin edilen risk değerleri sağlıklılarda düşük, hastalarda yüksek yüzdelerde seyretmektedir.

SNAPPE-II risk tahmin modeline celestone bilgisinin eklenmesi ile eklenmemesine bağlı olarak elde edilen ROC-AUC değeri aynıdır (ROC-AUC: 0,921). Ayrıca SNAPPE-II risk tahmin modeline sürfaktan ve seleston bilgisinin aynı anda eklenmesi ile elde edilen modelin performansı eski modelden her açıdan daha düşüktür. Bu modellerin performansının istatistiksel olarak değerlendirilmesi Bölüm 4.1.4'te verilecektir.

4.1.3. SNAP-II Risk Tahmin Modeline Eklenen Yeni Değişkenlerle Performansın Değerlendirilmesi

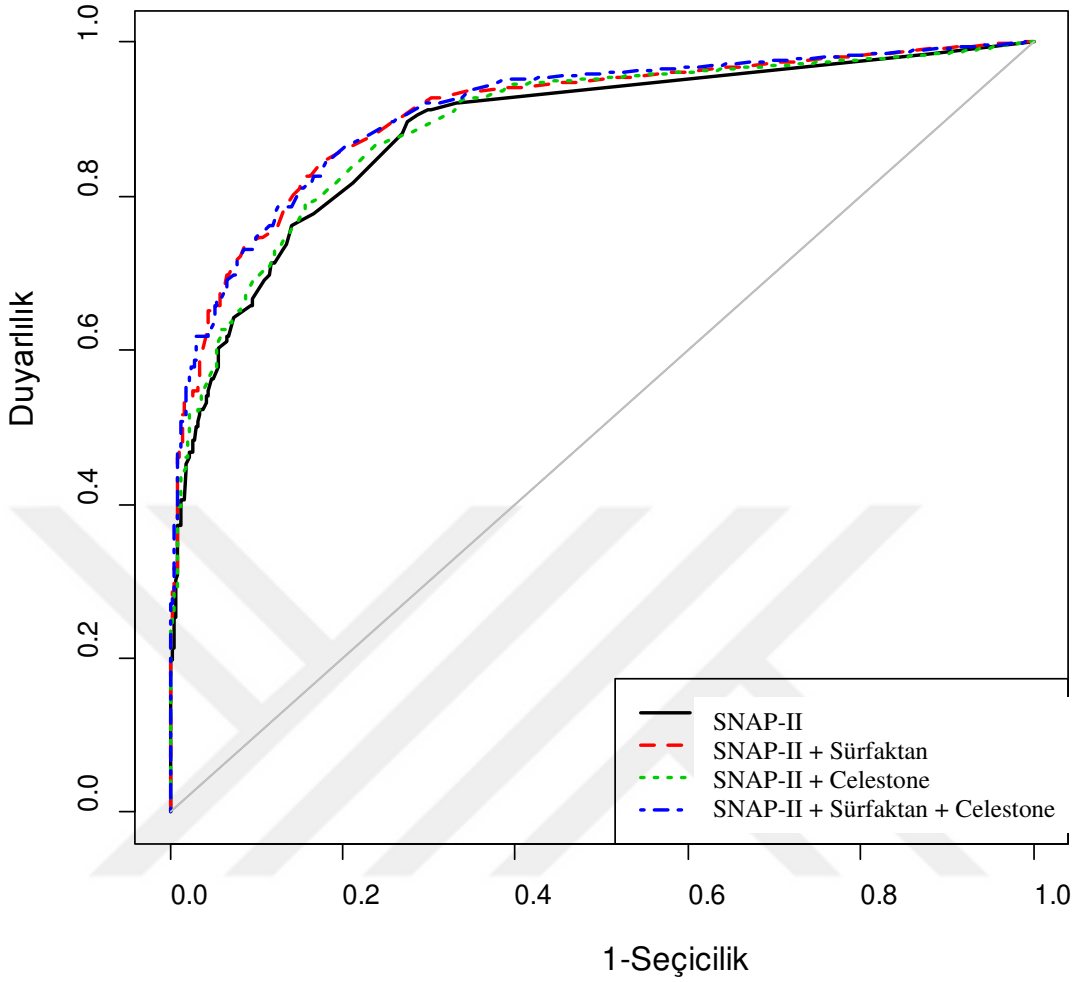
SNAP-II risk tahmin modelinin ROC-AUC değeri 0,885 (GA: 0,851-0,919) olarak bulunurken sürfaktan bilgisinin eklenmesi ile bu değer 0,907 (GA: 0,877-0,938)'ye yükselmiştir. Bu artışın istatistiksel olarak test edilmesi için kullanılan DeLong testinin p değeri 0,015 olarak bulunmuştur. Benzer amaca yönelik olarak geliştirilen NRI kategorik istatistiğinin p değeri 0,154, NRI sürekli istatistiğinin p değeri <0,001 ve IDI istatistiğinin p değeri <0,001 olarak elde edilmiştir.

İkinci olarak modele celestone değişkeni eklenmiş ROC-AUC değeri 0,892 (GA: 0,859-0,926) olarak elde edilmiştir. DeLong testinin p değeri 0,231 iken NRI kategorik istatistiğinin p değeri 0,938, NRI sürekli istatistiğinin p değeri <0,001 ve IDI istatistiğinin p değeri <0,001'dir.

Son olarak sürfaktan ve celestone bilgisi aynı anda modele dahil edilmiş ve ROC-AUC değeri 0,910 (GA: 0,880-0,940)'a yükselmiştir. Burada model performansını değerlendirmek için kullanılan dört yöntemin p değerleri şu şekildedir; DeLong testi 0,009, NRI kategorik 0,271, NRI sürekli <0,001 ve IDI <0,001'dır.

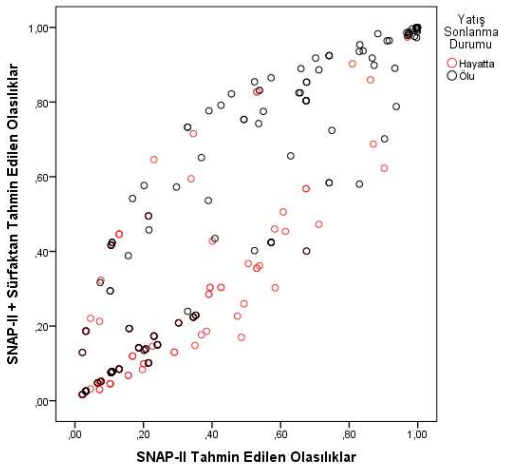
Çizelge 4.1.10. SNAP-II risk tahmin modeline eklenen yeni değişkenlere bağlı performansın değerlendirilmesi

	Test İstatistiği	**Test İstatistiğinin GA**	**p değeri**
Sürfakta'nın Eklendiği Model			
ΔAUC	2,422	-	0,015
NRI Kategorik	0,045	-0,017-0,106	0,154
NRI Sürekli	1,011	0,934-1,187	<0,001
IDI	0,071	0,044-0,099	<0,001
Celestone'un Eklendiği Model			
ΔAUC	1,198	-	0,231
NRI Kategorik	0,002	-0,052-0,056	0,938
NRI Sürekli	0,482	0,309-0,654	<0,001
IDI	0,020	0,007-0,034	<0,001
Hem Sürfaktan Hem de Celestone'un Eklendiği Model			
ΔAUC	2,616	-	0,009
NRI Kategorik	0,271	-0,028-0,099	0,271
NRI Sürekli	0,861	0,681-1,041	<0,001
IDI	0,075	0,047-0,103	<0,001

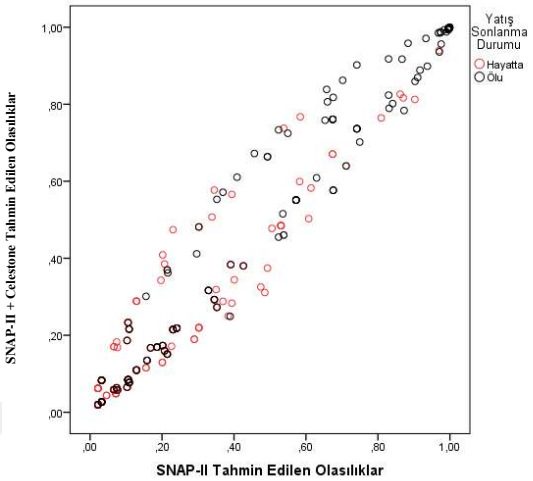


Şekil 4.1.21. SNAP-II risk tahmin modeline eklenen yeni değişkenlere bağlı elde edilen ROC eğrileri

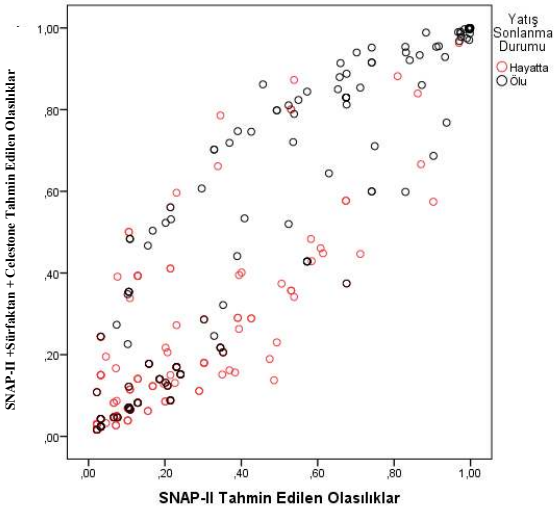
Eklenen yeni değişkene bağlı olarak elde edilen risk olasılıkları ile eski modelin risk olasılıklarına bağlı olarak çizilen serpme (scatter) grafiği iki modelin risk dağılımlarının karşılaştırılmasına olanak sağlar. Her bir yeni modelin eski modelle bağlı serpme grafiği Şekil 4.1.22.'de verildiği gibidir.



SNAP-II ve SNAP-II + Sürfaktan Risk Modellerinin Serpme Grafiği



SNAP-II ve SNAP-II + Celestone Risk Modellerinin Serpme Grafiği

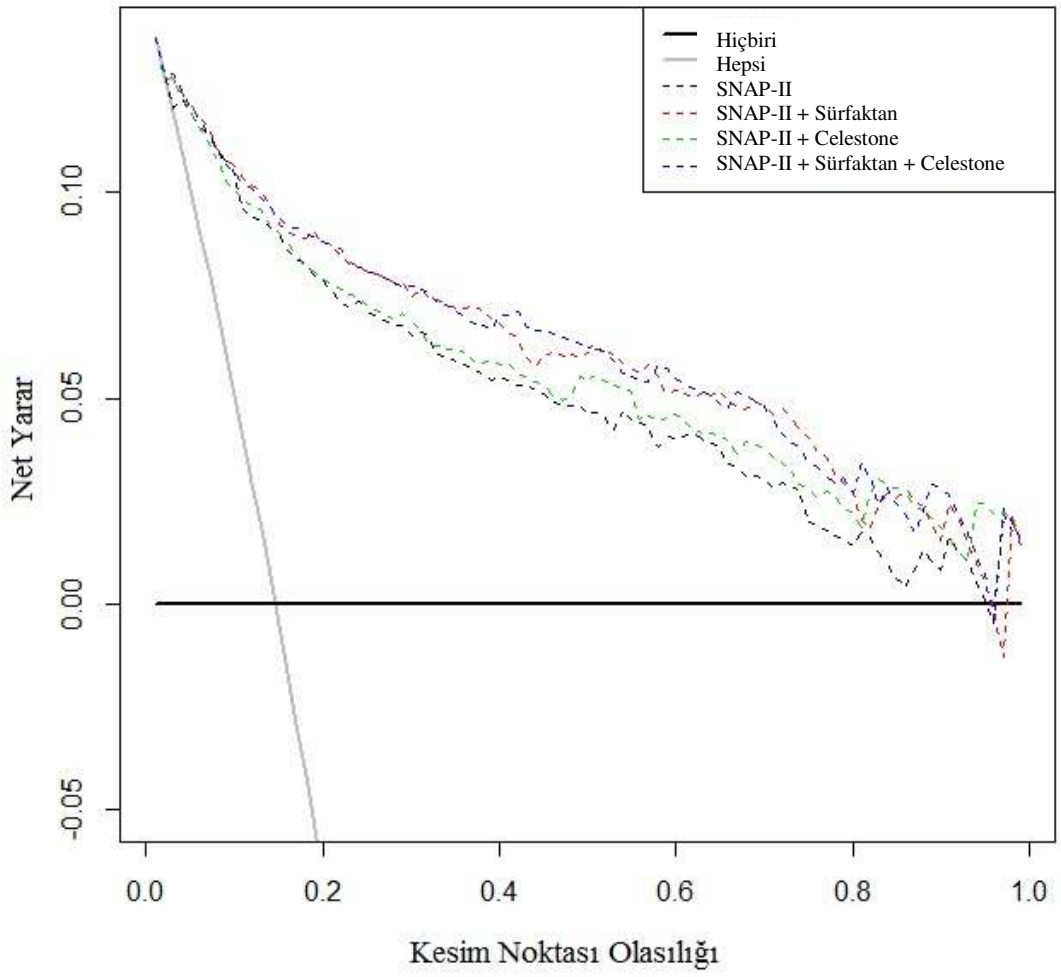


SNAP-II ve SNAP-II + Sürfaktan + Celestone Risk Modellerinin Serpme Grafiği

Şekil 4.1.22. SNAP-II risk tahmin modeline eklenen yeni değişkenlere bağlı elde edilen risk olasılıklarının yatış sonlanma durumuna göre serpme grafiği

SNAP-II risk tahmin modeline eklenen sürfaktan bilgisi ile ROC-AUC değeri 0,885 (GA: 0,851-0,919)'ten 0,907 (GA: 0,877-0,938)'ye yükselmiştir. Bu artış yapılan DeLong testi ile anlamlı bulunmuştur (p=0,0015). NRI sürekli ve IDI istatistikleri de DeLong testine paralel olarak sürfaktan bilgisinin eklenmesi ile modelin performansındaki artışın istatistiksel olarak anlamlı olduğu sonucuna varılmasını sağlamıştır (NRI sürekli p değeri: <0,001 (test istatistiği: 1,011) ve IDI istatistiğinin p değeri: <0,001 (test istatistiği: 0,071)). Bunlara ek olarak NRI kategorik istatistiği performanstaki artışın istatistiksel olarak anlamlı seviyelerde olmadığına yönelik bir sonuç vermiştir (p=0,154 (test istatistiği: 0,045)). İkinci olarak SNAP-II risk tahmin modeline celestone bilgisi eklenmiş ve model performansındaki artış NRI sürekli ve IDI istatistikleri tarafından anlamlı bulunmuştur (NRI sürekli p değeri: <0,001 (test istatistiği: 0,482) ve IDI istatistiğinin p değeri: <0,001 (test istatistiği: 0,020)). Ancak ROC-AUC değeri 0,885 (GA: 0,851-0,919)'ten 0,892 (GA: 0,859-0,926)'ye yükselmesine karşın DeLong testi istatistiksel olarak anlamlı olmayan bir sonuç vermiştir (DeLong testi p değeri: 0,203 (test istatistiği: 1,198)). NRI kategorik istatistiği de DeLong testine paralel olarak anlamsızdır (p değeri: 0,938 (test istatistiği: 0,002)). Son olarak modele hem sürfaktan hem de celestone bilgisi aynı anda ilave edilmiş ve sürfaktan bilgisinin tek başına modele ilave edilmesine benzer olarak NRI kategorik dışındaki tüm yöntemlerde artış istatistiksel olarak anlamlı bulunmuştur. DeLong testinin p değeri 0,009 (test istatistiği: 2,616) şeklinde iken NRI kategorik istatistiğinin p değeri 0,271 (test istatistiği: 0,271), NRI sürekli istatistiğinin p değeri <0,001 (test istatistiği: 0,861) ve IDI istatistiğinin p değeri <0,001 (test istatistiği:0,075) şeklinde bulunmuştur. ROC-AUC değeri 0,885 (GA: 0,851-0,919)'ten 0,910 (GA: 0,880-0,940)'a yükselmiştir.

Karar eğrisi yaklaşımı ile modellerin klinik yararı değerlendirilebilir. Bu amaçla çizilen karar eğrisi Şekil 4.1.23'te verildiği gibidir. SNAP-II risk tahmin modeline sürfaktan ve celestone bilgilerinin eklenmesi ile elde edilen karar eğrisi diğer modellerden daha yüksek seyretmektedir. Bu risk tahmin modelinin peşi sıra en yüksek klinik yarara sahip olan model SNAP-II + sürfaktan'dır. Bununla beraber SNAP-II'ye her eklenen değişkenle elde edilen risk tahmin modellerinin klinik yararı, SNAP-II risk tahmin modelinden daha yüksektir.



Şekil 4.1.23. SNAP-II ve uzantılarını içeren karar eğrileri

4.1.4. SNAPPE-II Risk Tahmin Modeline Eklenen Yeni Değişkenlerle Performansın Değerlendirilmesi

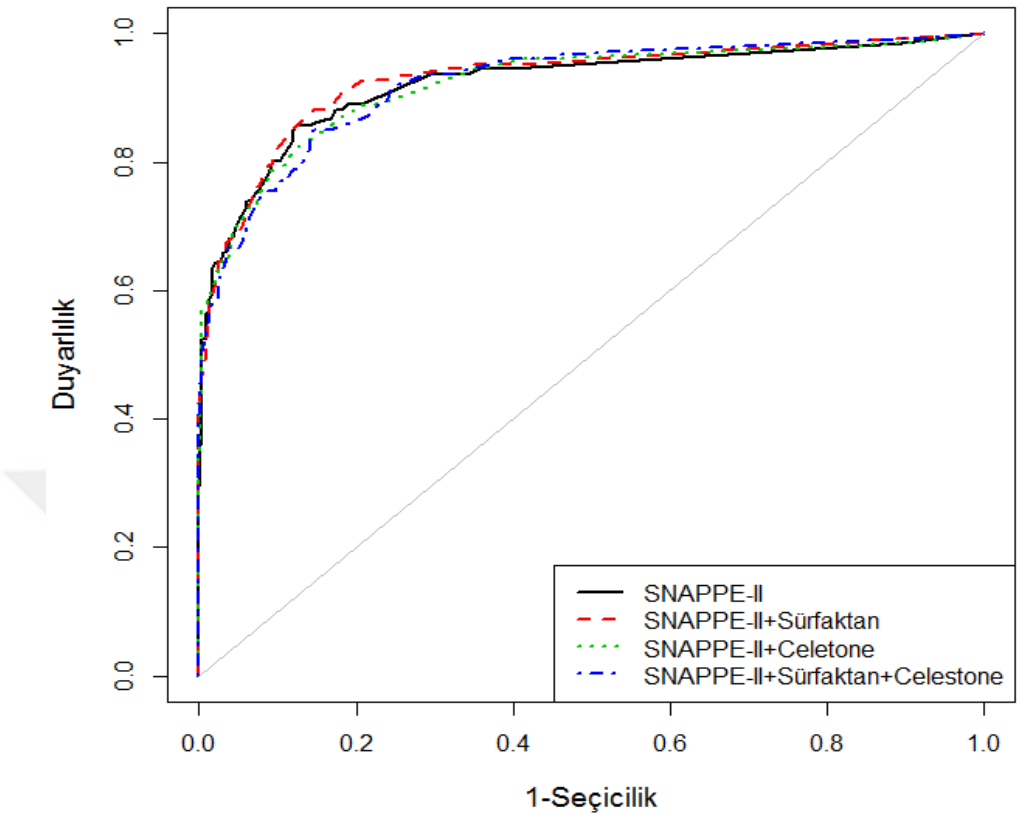
SNAPPE-II risk tahmin modelinin ROC-AUC değeri 0,921 (GA: 0,890-0,953) olarak bulunurken sürfaktan bilgisinin eklenmesi ile bu değer 0,930 (GA:0,902-0,958)'a yükselmiştir. Bu artışın istatistiksel olarak test edilmesi için kullanılan DeLong testinin p değeri 0,248 olarak bulunmuştur. NRI kategorik istatistiğinin p değeri 0,889, NRI sürekli istatistiğinin p değeri <0,001 (azalış) ve IDI istatistiğinin p değeri 0,897 olarak elde edilmiştir.

İkinci olarak modele celestone değişkeni eklenmiş ROC-AUC değeri 0,921 (GA: 0,890-0,953) olarak elde edilmiştir. DeLong testinin p değeri 0,908 iken NRI kategorik istatistiğinin p değeri 0,035 (azalış), NRI sürekli istatistiğinin p değeri 0,003 ve IDI istatistiğinin p değeri 0,226'dır.

Son olarak sürfaktan ve celestone bilgisi aynı anda modele dahil edilmiş ve ROC-AUC değeri 0,923 (GA:0,895-0,951)'e yükselmiştir. Burada model performansını değerlendirmek için kullanılan dört yöntemin p değerleri şu şekildedir; DeLong testi 0,861, NRI kategorik 0,003 (azalış), NRI sürekli 0,003 ve IDI 0,011 (azalış)'dir.

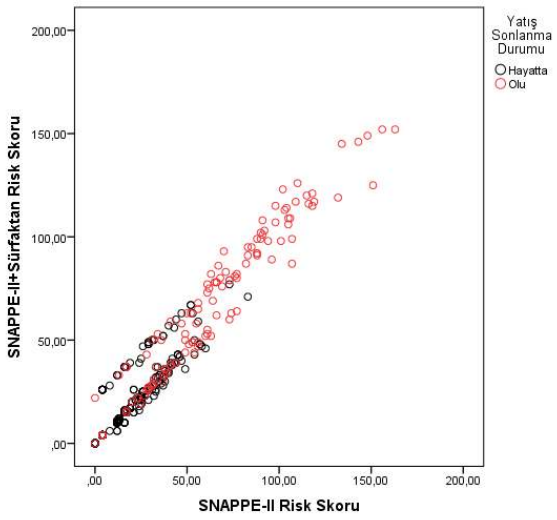
Çizelge 4.1.11. SNAPPE-II risk tahmin modeline eklenen yeni değişkenlere bağlı performansın değerlendirilmesi

	Test İstatistiği	**Test İstatistiğinin GA**	**p değeri**
Sürfaktan'ın Eklendiği Model			
ΔAUC	1,156	-	0,248
NRI Kategorik	-0,004	-0,055-0,048	0,889
NRI Sürekli	-0,367	-0,554-(-0,180)	<0,001
IDI	0,001	-0,018-0,021	0,897
Celestone'un Eklendiği Model			
ΔAUC	-0,115	-	0,908
NRI Kategorik	-0,036	-0,069-(-0,002)	0,035
NRI Sürekli	0,273	0,095-0,450	0,003
IDI	-0,005	-0,014-0,003	0,226
Hem Sürfaktan Hem de Celestone'un Eklendiği Model			
ΔAUC	0,175	-	0,861
NRI Kategorik	-0,069	-0,115-(-0,022)	0,004
NRI Sürekli	0,272	0,093-0,451	0,003
IDI	-0,022	-0,041-(-0,004)	0,017

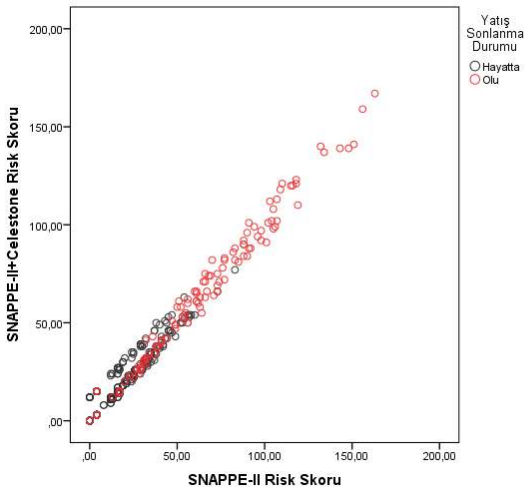


Şekil 4.1.24. SNAPPE-II risk tahmin modeline eklenen yeni değişkenlere bağlı elde edilen ROC eğrileri

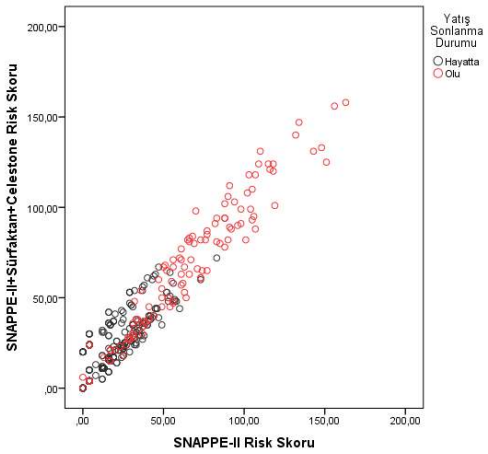
Her bir yeni modelin eski modelle bağlı serpme grafiği Şekil 4.1.25'te verildiği gibidir. Bu grafikten hareketle eski model ile yeni modelin risk olasılıklarının benzerlik gösterip göstermediği görsel olarak değerlendirilebilir.



SNAPPE-II ve SNAPPE-II + Sürfaktan Risk Modellerinin Serpme Grafiği



SNAPPE-II ve SNAPPE-II + Celestone Risk Modellerinin Serpme Grafiği

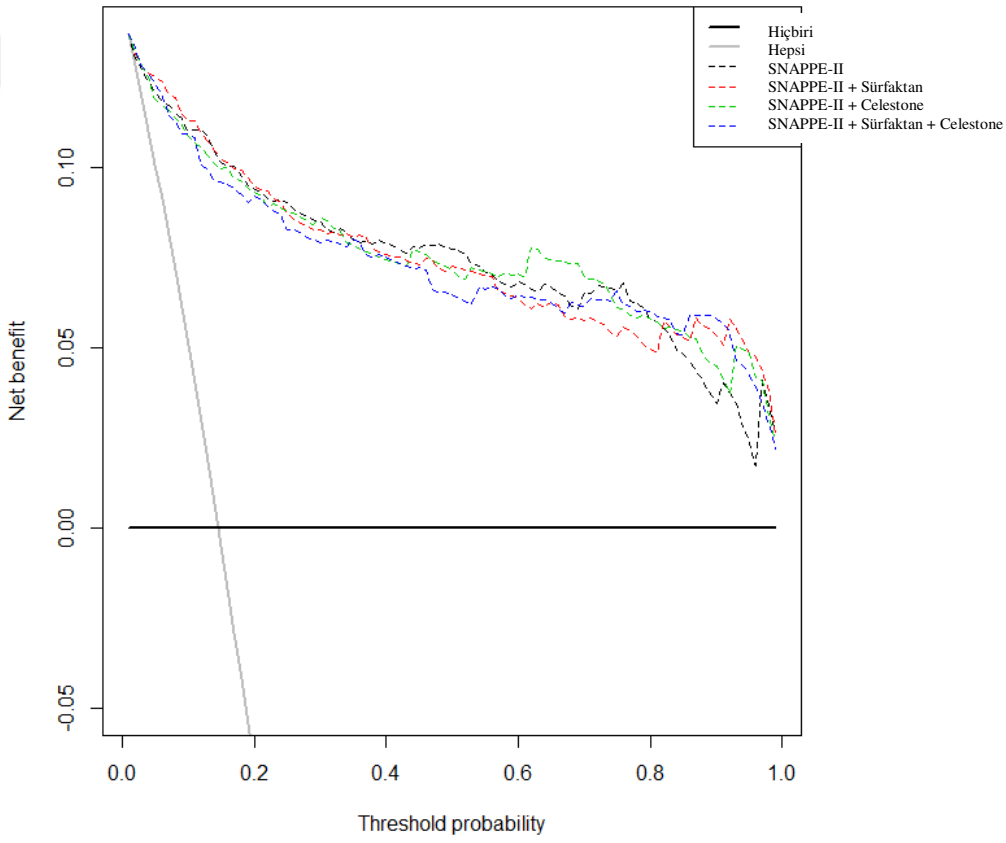


SNAPPE-II ve SNAPPE-II + Sürfaktan + Celestone Risk Modellerinin Serpme Grafiği

Şekil 4.1.25. SNAPPE-II risk tahmin modeline eklenen yeni değişkenlere bağlı elde edilen risk olasılıklarının yatış sonlanma durumuna göre serpme grafiği

Öncelikle modele sürfaktan bilgisi eklenmiş ve bu modelin performansı diğer modellerden daha yüksek seviyede olmasına karşın tüm yöntemler modelin performansındaki değişimi istatistiksel olarak anlamlı bulamamıştır. Modele celestone'un eklenmesi durumunda modelin performansındaki artışı, NRI sürekli (p=0,003) istatistiği dışındaki herhangi bir yöntem anlamlı bulamazken NRI kategorik istatistiği artış yönünde değil modelin performansında azalış yönünde bir anlamlılık olduğunu bulmuştur (p=0,035 ve test istatistiği= -0,036). Son olarak modele sürfaktan ve celestone bilgisi aynı anda ilave edilmiş ve ΔAUC yönteminin dışındaki yöntemler modelin performansında artış değil azalış olduğuna yönelik test istatistikleri vermişlerdir. Ayrıc NRI sürekli performans artışını anlamlı buluştur (p=0,003).

SNAPPE-II risk tahmin modelinin, klinik yararını değerlendirmek için çizilen karar eğrisi Şekil 4.1.26'da verildiği gibidir. Karar eğrilerinden hiçbiri net bir şekilde bir diğerinden daha üstte seyretmemektedir.



Şekil 4.1.26. SNAPPE-II ve uzantılarını içeren karar eğrisi

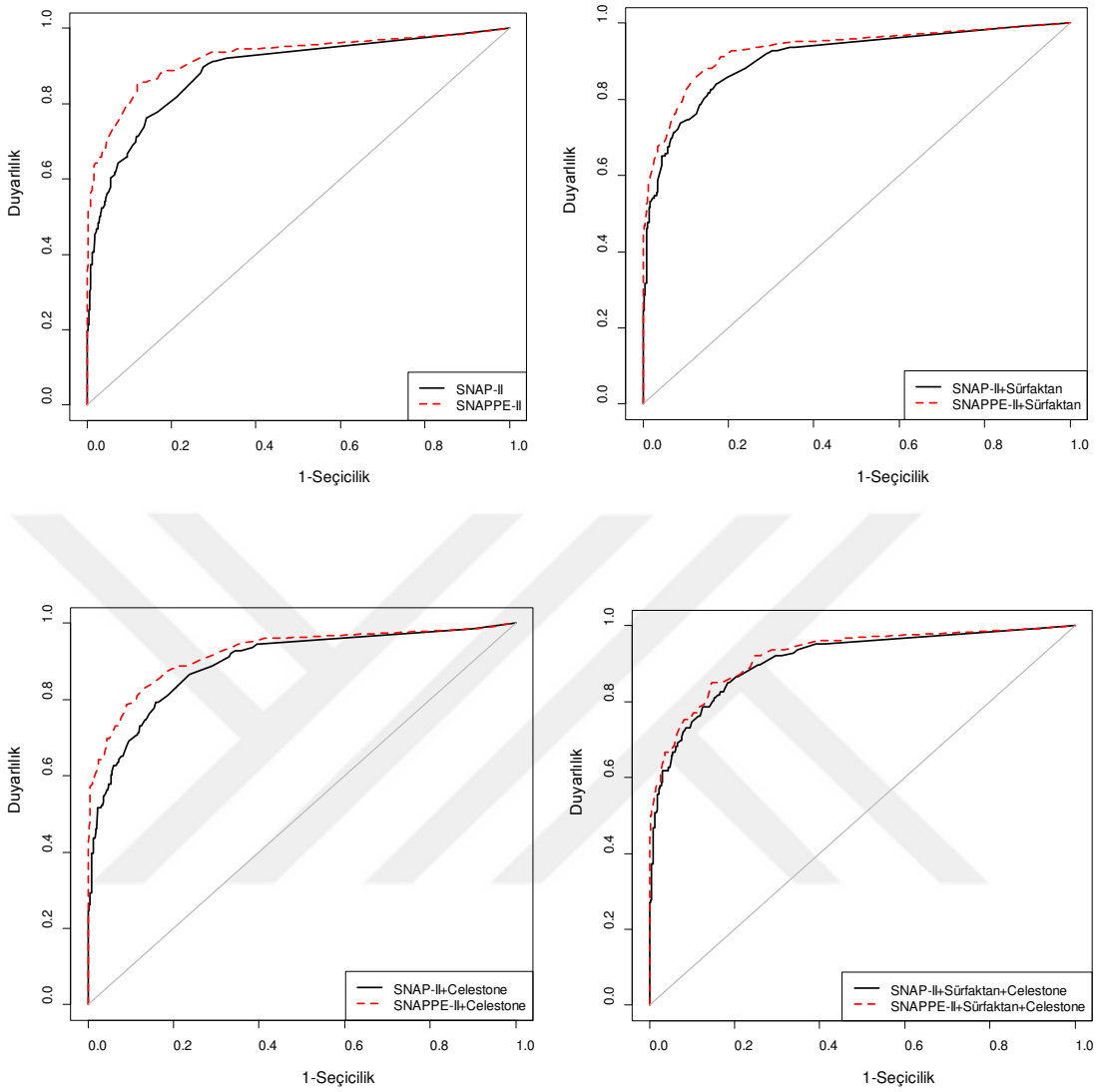
4.1.5. SNAP-II ve SNAPPE-II Risk Tahmin Modellerinin Performanslarının Karşılaştırılması

Burada amaç, hangi risk tahmin modelinin tahminde daha etkin olduğunu belirlemektir. SNAP-II ve SNAPPE-II risk tahmin modelleri iç içe tasarıma sahip olduklarından bu iki modelin performansları dört yöntemle sınanabilir. SNAP-II ve uzantıları eski model olarak kabul edilirken SNAPPE-II ve uzantıları yeni model olarak kabul edilmiştir.

Çizelge 4.1.12. SNAP-II ve SNAPPE-II risk tahmin modellerinin performanslarının karşılaştırılması

	Test İstatistiği	**Test İstatistiğinin GA**	**p değeri**
Eski Modeller			
ΔAUC	4,104	-	<0,001
NRI Kategorik	0,086	0,030-0,141	0,002
NRI Sürekli	1,084	0,913-1,254	<0,001
IDI	0,151	0,111-0,192	<0,001
Sürfaktanın Eklendiği Model			
ΔAUC	3,167	-	0,002
NRI Kategorik	0,051	-0,012-0,113	0,111
NRI Sürekli	1,057	0,885-1,229	<0,001
IDI	0,081	0,053-0,109	<0,001
Celestone'un Eklendiği Model			
ΔAUC	3,460	-	0,001
NRI Kategorik	0,059	0,006-0,112	0,029
NRI Sürekli	1,068	0,896-1,239	<0,001
IDI	0,126	0,089-0,162	<0,001
Hem Sürfakta Hem de Celestone'un Eklendiği Model			
ΔAUC	1,402	-	0,161
NRI Kategorik	-0,006	-0,078-0,066	0,870
NRI Sürekli	0,893	0,716-1,071	<0,001
IDI	0,054	0,021-0,087	<0,001

Çoğu senaryoda SNAPPE-II ve uzantısı olan risk tahmin modelleri SNAP-II ve uzantılarından daha iyi bir performansa sahiptir (IDI istatistiği genellikle baz alınmıştır). Görsel olarak değerlendirebilmek adına çizilen ROC eğrileri Şekil 4.1.27'de verilmiştir.



Şekil 4.1.27. SNAP-II ve SNAPPE-II risk tahmin modellerine ve uzantılarına dayanan ROC eğrileri

4.2. Simülasyon Çalışmasına ait Sonuçlar

4.2.1. Olayın görülme sıklığı 0,10 iken elde edilen simülasyon sonuçları

ROC eğrisi altında kalan alan değişimi (ΔAUC), farklı varyasyon katsayılarında (CV) ve aynı örneklem büyüklüklerinde aynı sonuçları vermiştir. Yani örneklem büyüklüğü arttıkça ΔAUC değeri azalmış ancak tüm CV oranlarında değişim söz konusu olmamıştır (Çizelge 4.2.1). Buna karşılık p değerinin hatalı pozitiflik (HP) oranı, tüm durumlarda neredeyse eşittir ve ΔAUC istatistiği diğer üç yöntemden daha düşük HP oranına sahiptir (Çizelge 4.2.2). Şekil 4.2.1'de ΔAUC istatistiğinin p değeri ortalamaları ve ortalamaların güven aralıkları, dört örneklem büyüklüğünde de tüm yöntemlerden daha yukarıda seyretmektedir.

NRI kategorik istatistiği, örneklem büyüklüğü arttıkça düşmekte ancak CV değişimine göre bu istatistik değerinin değişimi anlamlı bir seyir göstermektedir (Çizelge 4.2.1). NRI kategorik istatistiğine ait p değerinin HP oranı, örneklem büyüklüğü arttıkça genellikle azalma eğilimindedir (Çizelge 4.2.2).

Çizelge 4.2.1. p=0,10 iken yöntemlerin test istatistik değerleri

Senaryolar		**p=0,10**					
		Eski Modele ait ROC-AUC ort±ss	**Yeni Modele ait ROC-AUC ort±ss**	**ΔAUC ort±ss**	**NRI Kategorik ort±ss**	**NRI Sürekli ort±ss**	**IDI ort±ss**
CV=1,00	**100**	0,877±0,043	0,880±0,042	0,004±0,009	0,167±0,015	0,222±0,286	0,009±0,013
	250	0,859±0,030	0,860±0,030	0,002±0,004	0,016±0,031	0,136±0,169	0,003±0,005
	500	0,850±0,023	0,851±0,023	0,001±0,002	0,007±0,015	0,093±0,115	0,000±0,002
	1000	0,844±0,018	0,845±0,018	0,000±0,001	0,000±0,009	0,063±0,077	0,000±0,001
CV=1,05	**100**	0,877±0,042	0,881±0,042	0,004±0,009	0,044±0,016	0,239±0,296	0,009±0,014
	250	0,860±0,031	0,862±0,031	0,002±0,004	0,000±0,030	0,129±0,173	0,003±0,005
	500	0,849±0,023	0,850±0,023	0,001±0,002	0,006±0,016	0,085±0,0114	0,001±0,002
	1000	0,845±0,018	0,846±0,018	0,000±0,001	0,001±0,011	0,067±0,083	0,000±0,001
CV=1,10	**100**	0,877±0,042	0,881±0,042	0,004±0,009	0,034±0,010	0,247±0,283	0,010±0,016
	250	0,858±0,030	0,860±0,030	0,002±0,004	0,007±0,023	0,133±0,174	0,000±0,005
	500	0,851±0,024	0,852±0,024	0,001±0,002	0,007±0,017	0,093±0,117	0,000±0,003
	1000	0,844±0,018	0,844±0,018	0,000±0,001	0,003±0,012	0,066±0,079	0,000±0,001
CV=1,15	**100**	0,877±0,043	0,881±0,042	0,004±0,009	0,111±0,002	0,248±0,297	0,010±0,016
	250	0,859±0,031	0,861±0,031	0,002±0,004	0,017±0,017	0,139±0,169	0,000±0,005
	500	0,849±0,023	0,850±0,023	0,001±0,002	0,001±0,017	0,101±0,120	0,001±0,002
	1000	0,845±0,018	0,846±0,018	0,000±0,001	0,001±0,011	0,078±0,082	0,001±0,001

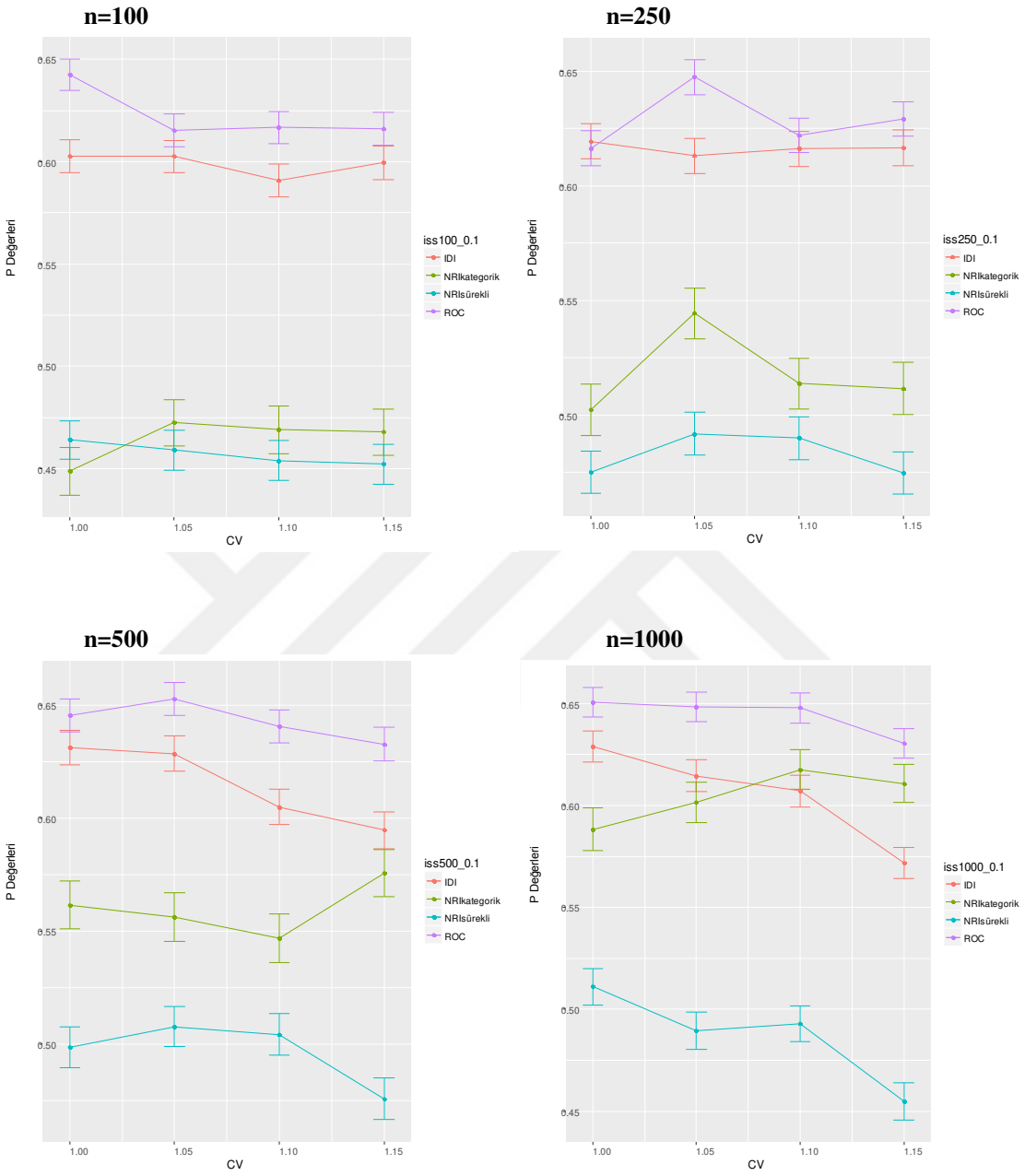
Çizelge 4.2.2. p=0,10 iken yöntemlerin HP oranları

Senaryolar		p=0,10			
		CV=1	**CV=1,05**	**CV=1,10**	**CV=1,15**
ΔAUC	**100**	0,001	0,001	0,001	0,001
	250	0,001	0,001	0,000	0,001
	500	0,001	0,002	0,003	0,003
	1000	0,001	0,000	0,004	0,000
NRI Kategorik	**100**	0,127	0,085	0,103	0,093
	250	0,120	0,091	0,113	0,123
	500	0,079	0,086	0,101	0,076
	1000	0,089	0,072	0,058	0,050
NRI Sürekli	**100**	0,092	0,100	0,096	0,106
	250	0,053	0,066	0,065	0,057
	500	0,047	0,042	0,052	0,071
	1000	0,035	0,047	0,039	0,060
IDI	**100**	0,010	0,009	0,010	0,008
	250	0,001	0,003	0,002	0,004
	500	0,001	0,001	0,002	0,002
	1000	0,000	0,000	0,004	0,004

NRI sürekli yaklaşımının istatistik değerlerindeki değişim örneklem büyüklüğünün artışına bağlı olarak düşmektedir. Buna ek olarak aynı örneklem büyüklüklerinde ve artan CV oranlarında NRI sürekli istatistiğinin artış gösterdiği varılabilecek bir diğer sonuçtur (Çizelge 4.2.1). p değerinin HP oranı, aynı CV değerinde örneklem büyüklüğü arttıkça azalmaktadır. NRI kategorik istatistiğinden sonra en yüksek HP oranına sahip olan yaklaşım NRI sürekli istatistiğidir. Örnek büyüklüğü 1000 ve CV=1 iken dahi HP oranı %3,5'tir (Çizelge 4.2.2). Şekil 4.2.1'de NRI sürekli istatistiğinin p değerlerinin ortalaması ve güven aralığı, tüm diğer yöntemlerde daha aşağıda seyretmektedir.

Son olarak IDI istatistiği, ΔAUC istatistiğinde olduğu gibi HP oranı düşük seviyelerdedir. ΔAUC istatistiğinden sonra p değerinin dağılımının en yüksek seyrettiği yaklaşımdır (Şekil 4.2.1). IDI istatistiği CV oranındaki değişimden çok fazla etkilenmese de n arttıkça test istatistiğinin değerinde bir azalma söz konusudur (Çizelge 4.2.1.). Örneklem büyüklüğü arttıkça HP oranı düşmekte ek olarak CV değerleri arttıkça her bir n senaryosunda HP oranı artmaktadır (Çizelge 4.2.2).

p=0,10



Şekil 4.2.1. p=0,10 iken yöntemlerin p değerlerinin ortalamasının ve ortalamanın güven aralığının CV değişimine bağlı dağılımı

4.2.2. Olayın görülme sıklığı 0,25 iken elde edilen simülasyon sonuçları

Olayın görülme sıklığı 0,25 iken ΔAUC değeri örneklem büyüklüğüne bağlı olarak değişim gösterirken CV değerindeki değişimlerden etkilenmemektedir. Örneklem büyüklüğü arttıkça ΔAUC değeri sıfıra yaklaşmaktadır (Çizelge 4.2.3). Test istatistiğinin p değerine ait HP oranı örnek büyüklüğünden ya da CV değişiminden etkilenmeksizin genellikle sıfıra oldukça yakındır (Çizelge 4.2.4). Şekil 4.2.2'de, bu istatistiğin p değerinin ortalaması ve güven aralığı diğer üç yöntemden daha yukarıda seyretmektedir.

NRI kategorik istatistiğinin değeri, n arttıkça genellikle azalma eğilimi göstermektedir (Çizelge 4.2.3). NRI kategorik istatistiği, NRI sürekli istatistiğinden sonra en yüksek HP oranına sahip olan yaklaşımdır. Örneklem büyüklüğü 1000 ve CV değeri 1 iken HP oranı %1,0'dir (Çizelge 4.2.4) Ayrıca örnek büyüklüğü arttıkça NRI kategorik istatistiğinin p değerinin dağılımı artış göstermekte ve yüksek örnek büyüklüğünde ΔAUC ile benzer dağılım sergilemektedir (Şekil 4.2.2).

Çizelge 4.2.3. p=0,25 iken yöntemlerin test istatistik değerleri

Senaryolar		**p=0,25**					
		Eski Modele ait ROC-AUC ort±ss	**Yeni Modele ait ROC-AUC ort±ss**	**ΔAUC ort±ss**	**NRI Kategorik ort±ss**	**NRI Sürekli ort±ss**	**IDI ort±ss**
CV=1,00	**100**	0,856±0,036	0,858±0,035	0,003±0,005	0,006±0,035	0,146±0,203	0,006±0,008
	250	0,848±0,025	0,850±0,025	0,001±0,002	0,005±0,023	0,099±0,118	0,002±0,003
	500	0,843±0,019	0,844±0,019	0,001±0,001	0,001±0,010	0,066±0,082	0,001±0,002
	1000	0,842±0,014	0,843±0,014	0,000±0,000	0,001±0,006	0,046±0,059	0,000±0,001
CV=1,05	**100**	0,856±0,036	0,858±0,035	0,003±0,005	0,024±0,033	0,167±0,197	0,006±0,008
	250	0,848±0,025	0,849±0,025	0,001±0,002	0,003±0,017	0,102±0,121	0,002±0,003
	500	0,843±0,019	0,844±0,019	0,001±0,001	0,003±0,010	0,070±0,083	0,001±0,002
	1000	0,842±0,014	0,843±0,014	0,000±0,000	0,002±0,007	0,051±0,061	0,000±0,001
CV=1,10	**100**	0,858±0,034	0,860±0,034	0,003±0,005	0,015±0,032	0,162±0,196	0,006±0,008
	250	0,847±0,026	0,847±0,026	0,001±0,002	0,006±0,016	0,097±0,118	0,002±0,003
	500	0,843±0,019	0,844±0,019	0,001±0,001	0,004±0,012	0,071±0,090	0,001±0,002
	1000	0,841±0,014	0,842±0,014	0,000±0,001	0,001±0,006	0,052±0,059	0,000±0,001
CV=1,15	**100**	0,858±0,035	0,861±0,035	0,003±0,005	0,001±0,043	0,169±0,198	0,006±0,009
	250	0,845±0,024	0,847±0,024	0,001±0,002	0,002±0,014	0,101±0,119	0,003±0,003
	500	0,843±0,019	0,844±0,019	0,001±0,001	0,004±0,011	0,075±0,087	0,002±0,002
	1000	0,842±0,014	0,843±0,014	0,000±0,001	0,001±0,006	0,058±0,061	0,001±0,001

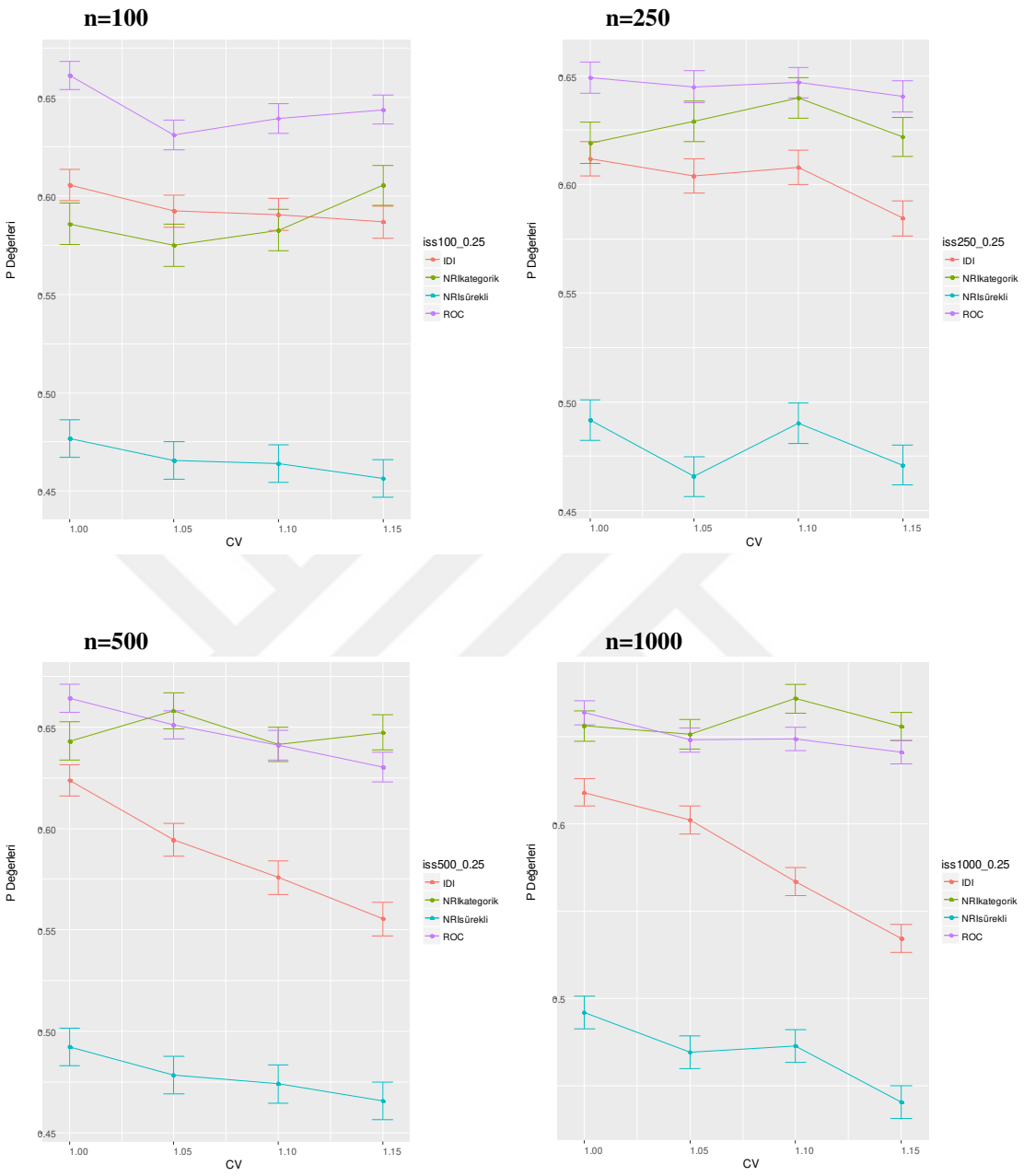
Çizelge 4.2.4. p=0,25 iken yöntemlerin HP oranları

Senaryolar		p=0,25			
		CV=1	**CV=1,05**	**CV=1,10**	**CV=1,15**
ΔAUC	**100**	0,000	0,002	0,001	0,001
	250	0,001	0,000	0,000	0,000
	500	0,000	0,001	0,000	0,000
	1000	0,000	0,000	0,001	0,000
NRI Kategorik	**100**	0,013	0,022	0,022	0,017
	250	0,015	0,014	0,015	0,011
	500	0,015	0,014	0,006	0,011
	1000	0,010	0,005	0,009	0,013
NRI Sürekli	**100**	0,079	0,086	0,084	0,090
	250	0,065	0,069	0,067	0,072
	500	0,057	0,069	0,087	0,083
	1000	0,053	0,068	0,073	0,088
IDI	**100**	0,007	0,004	0,003	0,004
	250	0,005	0,007	0,006	0,003
	500	0,006	0,006	0,004	0,006
	1000	0,002	0,005	0,004	0,007

NRI sürekli istatistiği, örneklem büyüklüğüne bağlı olarak azalmaktadır. Ancak bu istatistik için dikkat çeken en önemli noktalardan biri standart sapmalarının diğer yöntemlere kıyasla daha yüksek olmasıdır (Çizelge 4.2.3). NRI sürekli istatistiği, HP oranı en yüksek olan yaklaşımdır. Örneklem büyüklüğü 1000 ve CV değeri 1 iken dahi HP oranı %5,3'tür (Çizelge 4.2.4). HP oranı, örnek büyüklüğü arttıkça azalırken CV değeri arttıkça artmaktadır. Şekil 4.2.2'de NRI sürekli istatistiğinin p değeri diğer üç yöntemden daha düşük seviyelerde dağılmaktadır.

Son olarak IDI istatistik değeri, aynı örnek büyüklüğünde CV değeri arttıkça değişim göstermemekte sadece CV=1,15 iken daha yüksek değerler göstermektedir (Çizelge 4.2.3). ΔAUC istatistiğinden sonra en düşük HP oranına sahip olan yaklaşımdır. Örneklem büyüklüğü 1000 ve CV değeri 1 iken HP oranı %0,2'dir (Çizelge 4.2.4). Şekil 4.2.2'de IDI istatistiğinin p değerinin dağılımı CV arttıkça düşerken örneklem büyüklüğünden çok fazla etkilenmemektedir.

p=0,25



Şekil 4.2.2. p=0,25 iken yöntemlerin p değerlerinin ortalamasının ve ortalamanın güven aralığının CV değişimine bağlı dağılımı

4.2.3. Olayın görülme sıklığı 0,50 iken elde edilen simülasyon sonuçları

Olayın görülme sıklığı %50 iken iki ROC eğrisi altında kalan alan değişimi, örnek büyüklüğü arttıkça düşüş gösterirken CV değişiminden etkilenmemektedir (Çizelge 4.2.5). Tüm prevalans değerlerinde olduğu gibi ΔAUC istatistiğinin HP oranı genellikle 0'dır (Çizelge 4.2.6). Şekil 4.2.3'te diğer prevalans senaryolarından farklı olarak NRI kategorik istatistiğinin p değerinin dağılımı daha üst seviyelerdedir. ΔAUC istatistiğinde olduğu gibi bu istatistiğin p değerinin dağılımı CV değerinden etkilenmemektedir (Şekil 4.2.3).

NRI kategorik istatistiği diğer prevalans senaryolarından farklı olarak p değerinin dağılımı en üst seviyededir (Şekil 4.2.3). Buna karşılık HP oranı, ΔAUC istatistiğinin HP oranından daha yüksek seviyelerdedir (Çizelge 4.2.6). Bu istatistik değeri, örnek büyüklüğüne ya da CV değişimine bağlı olarak azalma ya da artması anlamlı bir seyir göstermemektedir (Çizelge 4.2.5).

Çizelge 4.2.5. p=0,50 iken yöntemlerin test istatistik değerleri

Senaryolar		**p=0,50**					
		Eski Modele ait ROC-AUC ort±ss	**Yeni Modele ait ROC-AUC ort±ss**	**ΔAUC ort±ss**	**NRI Kategorik ort±ss**	**NRI Sürekli ort±ss**	**IDI ort±ss**
CV=1,00	**100**	0,849±0,032	0,851±0,032	0,002±0,004	0,013±0,032	0,131±0,173	0,005±0,007
	250	0,844±0,023	0,845±0,023	0,001±0,002	0,003±0,015	0,083±0,104	0,002±0,003
	500	0,842±0,017	0,843±0,017	0,000±0,001	0,002±0,010	0,063±0,074	0,001±0,001
	1000	0,842±0,012	0,842±0,012	0,000±0,000	0,001±0,005	0,042±0,050	0,000±0,001
CV=1,05	**100**	0,850±0,033	0,850±0,033	0,002±0,004	0,010±0,025	0,150±0,173	0,005±0,007
	250	0,844±0,023	0,845±0,023	0,001±0,002	0,003±0,017	0,088±0,103	0,002±0,003
	500	0,843±0,017	0,843±0,017	0,000±0,001	0,001±0,009	0,061±0,073	0,001±0,001
	1000	0,841±0,012	0,842±0,012	0,000±0,000	0,000±0,005	0,047±0,050	0,001±0,001
CV=1,10	**100**	0,853±0,034	0,855±0,033	0,002±0,004	0,011±0,027	0,143±0,173	0,005±0,007
	250	0,843±0,023	0,844±0,023	0,001±0,002	0,004±0,015	0,083±0,106	0,002±0,003
	500	0,842±0,017	0,842±0,017	0,001±0,001	0,001±0,009	0,067±0,073	0,001±0,001
	1000	0,842±0,012	0,842±0,012	0,000±0,000	0,001±0,005	0,048±0,053	0,001±0,001
CV=1,15	**100**	0,850±0,033	0,853±0,033	0,003±0,005	0,017±0,028	0,152±0,172	0,006±0,008
	250	0,844±0,023	0,845±0,023	0,001±0,001	0,002±0,015	0,091±0,105	0,002±0,003
	500	0,842±0,017	0,843±0,017	0,001±0,001	0,004±0,009	0,068±0,076	0,001±0,002
	1000	0,842±0,012	0,842±0,012	0,000±0,000	0,001±0,006	0,046±0,054	0,001±0,001

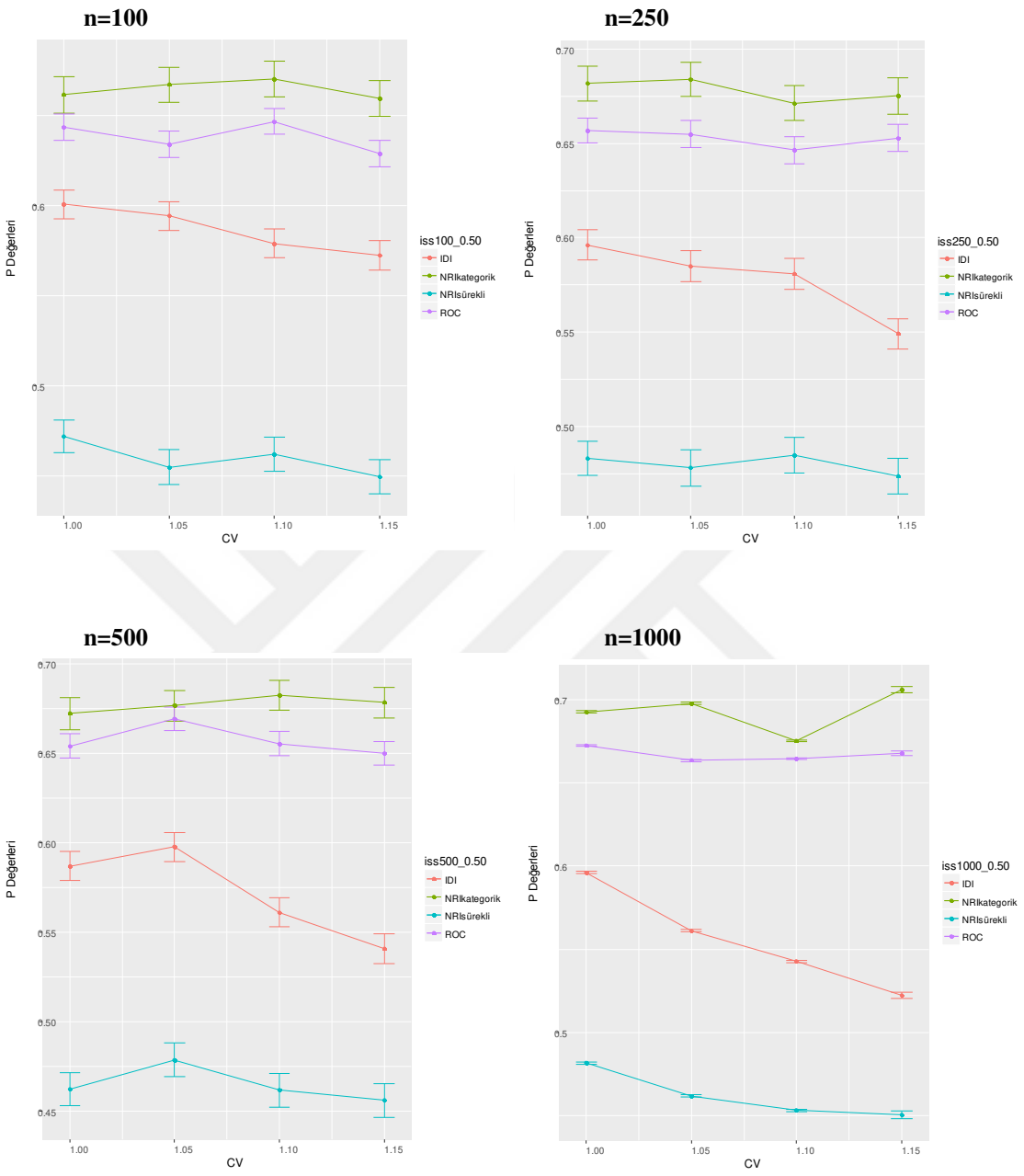
Çizelge 4.2.6. p=0,50 iken yöntemlerin HP oranları

Senaryolar		p=0,50			
		CV=1	**CV=1,05**	**CV=1,1**	**CV=1,15**
ΔAUC	**100**	0,001	0,001	0,000	0,000
	250	0,000	0,000	0,000	0,000
	500	0,000	0,000	0,000	0,000
	1000	0,000	0,000	0,000	0,000
NRI Kategorik	**100**	0,002	0,001	0,004	0,001
	250	0,000	0,003	0,003	0,005
	500	0,004	0,006	0,002	0,005
	1000	0,002	0,004	0,002	0,000
NRI Sürekli	**100**	0,085	0,096	0,089	0,088
	250	0,065	0,062	0,072	0,080
	500	0,081	0,081	0,080	0,090
	1000	0,054	0,078	0,083	0,080
IDI	**100**	0,009	0,009	0,005	0,010
	250	0,005	0,010	0,015	0,005
	500	0,007	0,003	0,009	0,013
	1000	0,002	0,011	0,012	0,000

NRI sürekli istatistiği klasik olarak tüm prevalans durumlarıyla benzerli göstermiş ve tüm senaryolarda yüksek oranda HP oran sergilemiştir. Örnek büyüklüğü 1000 ve CV değeri 1 iken dahi HP oranı %5,4'tür (Çizelge 4.2.6).

Son olarak IDI istatistik değeri, olayın görülme sıklığı 0,25 olduğu senaryodaki gibi aynı örnek büyüklüklerinde CV arttıkça benzer sonuçlar vermektedir. Sadece CV=1,15 iken diğer CV değişimlerinden farklı olarak daha yüksek istatistik değerlerine sahiptir (Çizelge 4.2.5). NRI sürekli istatistiğinden sonra en yüksek HP oranına sahip olan yaklaşım IDI yaklaşımıdır (Çizelge 4.2.6). Aynı durumu Şekil 4.2.3'te p değerlerinin dağılımında da görmek mümkündür.

p=0,50



Şekil 4.2.3. p=0,50 iken yöntemlerin p değerlerinin ortalamasının ve ortalamanın güven aralığının CV değişimine bağlı dağılımı

Genel olarak bahsedilenlerin ışığında;

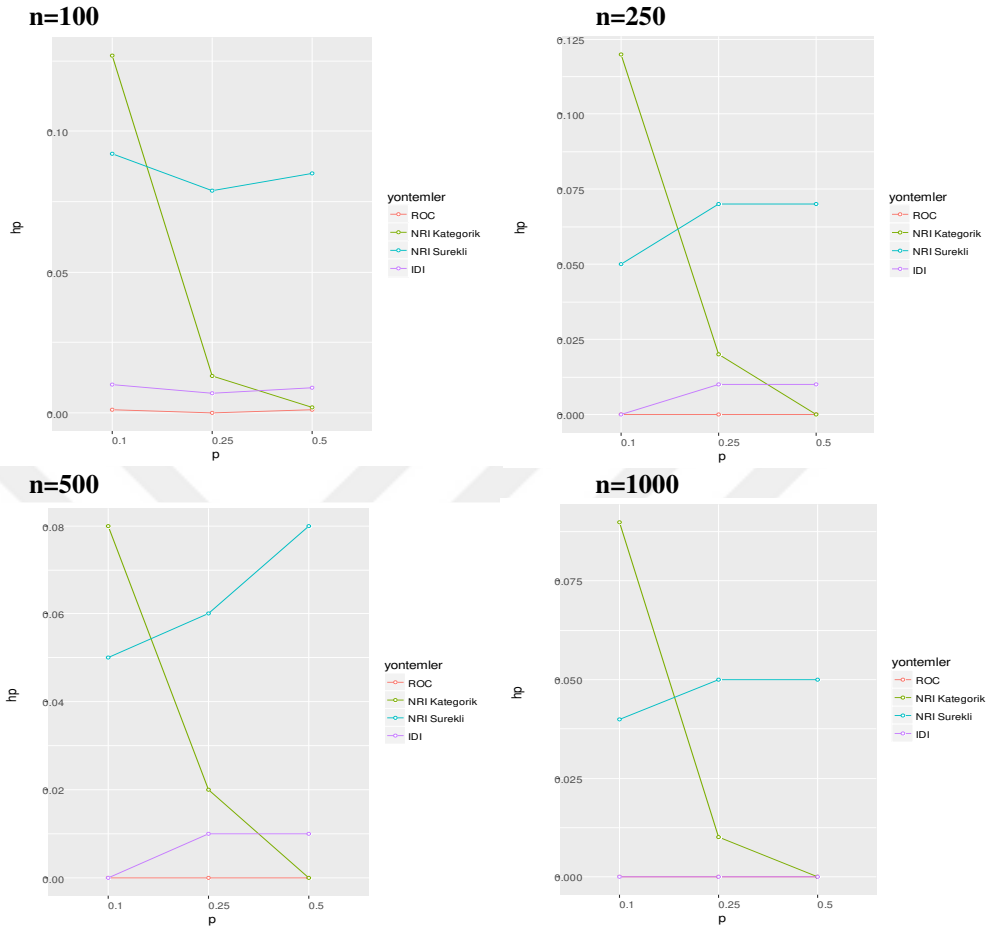
ΔAUC ve IDI istatistik değerleri örneklem büyüklüğü arttıkça azalmakta ancak CV değerinden çok net bir şekilde etkilenmemektedir. Ancak IDI istatistiğinin p değerinin dağılımı örneklem büyüklüğü ve CV oranı arttıkça ΔAUC istatistiğinin dağılımından uzaklaşmaktadır. IDI istatistiğinin p değerinin dağılımı ΔAUC istatistiğinden uzaklaştıkça NRI kategorik istatistiğinin dağılımı ΔAUC istatistiğinin dağılımına yakınlaşmaktadır. Hatta prevalans değeri 0,50'ye yükseldiğinde NRI kategorik istatistiğinin p değerinin dağılımı en üst seviyeye çıkmaktadır.

ΔAUC istatistiğinin HP oranı her durumda diğer yöntemlerden oldukça düşüktür. Prevalans değeri 0,10 ve 0,25 iken ΔAUC istatistiğinden sonra en düşük HP oranına sahip olan istatistik yöntemi IDI ve ardından NRI kategorik iken prevalans değeri 0,50'ye çıktığında NRI kategorik istatistiğinin HP oranı IDI yönteminden daha düşüktür. Bunun nedeni olayın görülme sıklığı arttıkça NRI kategorik istatistiğinin p değerinin dağılımı yükselmekte ve buna bağlı olarak HP oranı azalmaktadır.

NRI sürekli istatistiğinin p değerinin dağılımı her senaryoda diğer yöntemlerden en aşağıda seyretmektedir. Bundan dolayı HP oranı en yüksek olan yaklaşımdır. NRI sürekli istatistiğinin p değerinin dağılımı örneklem büyüklüğünden ya da prevalans değerinin değişiminden net bir şekilde etkilenmemektedir. Buna karşılık CV değeri arttıkça p değerinin dağılımı az da olsa düşmektedir. Ayrıca NRI sürekli istatistiğinin standart sapma değeri her durumda oldukça yüksektir. Hatta ortalamadan daha yüksek standart sapma seviyesi söz konusudur.

4.2.4. Farklı prevalans değerlerinde ve sabit CV oranlarında yöntemlerin HP oranları

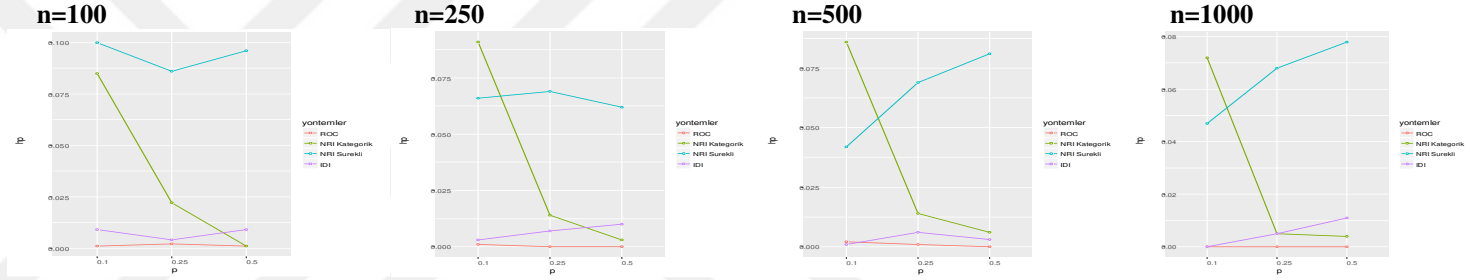
CV=1



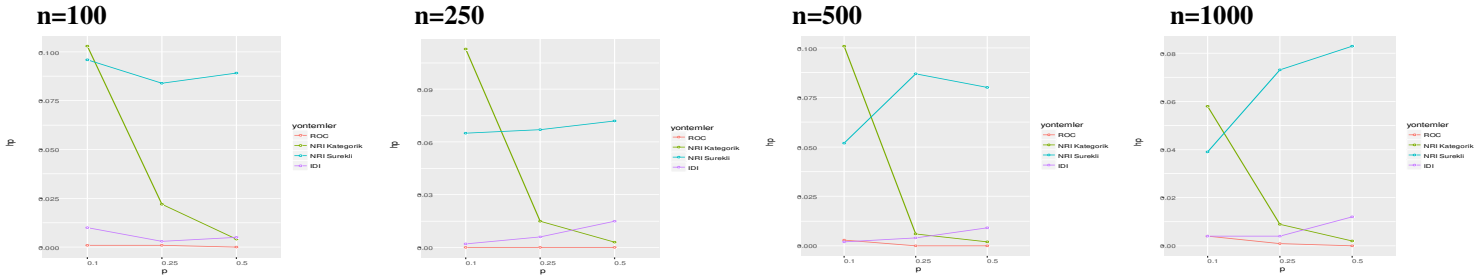
Şekil 4.2.4. CV=1 iken yöntemlerin HP oranlarının farklı prevalans değerlerindeki dağılımı

CV oranı 1 alındığında, ΔAUC istatistiğinin HP oranı her durumda sıfırdır. IDI istatistiği ise örnek büyüklüğü 100 iken net bir değişim göstermemekte ancak örnek büyüklüğü 250 ve 500'de prevalans değeri 0,25'e çıkarken HP oranı artmakta prevalans değeri 0,50 iken HP oranı değişim göstermemektedir. Örnek büyüklüğü 1000'e çıktığında prevalans değeri ne olursa olsun IDI istatistiğinin HP oranı sıfırdır. NRI kategorik istatistiğinin HP oranı, prevalans değerinden oldukça net bir şekilde etkilenmektedir. Prevalans değeri arttıkça HP oranı düşmekte ve hatta 0,50 prevalans değerinde HP oranı sıfırlanmaktadır. NRI sürekli istatistiği, prevalans değeri 0,10 iken NRI kategorik istatistiğinden daha düşük HP oranına sahip iken prevalans değeri arttıkça en yüksek HP oranına sahip olan yaklaşımdır.

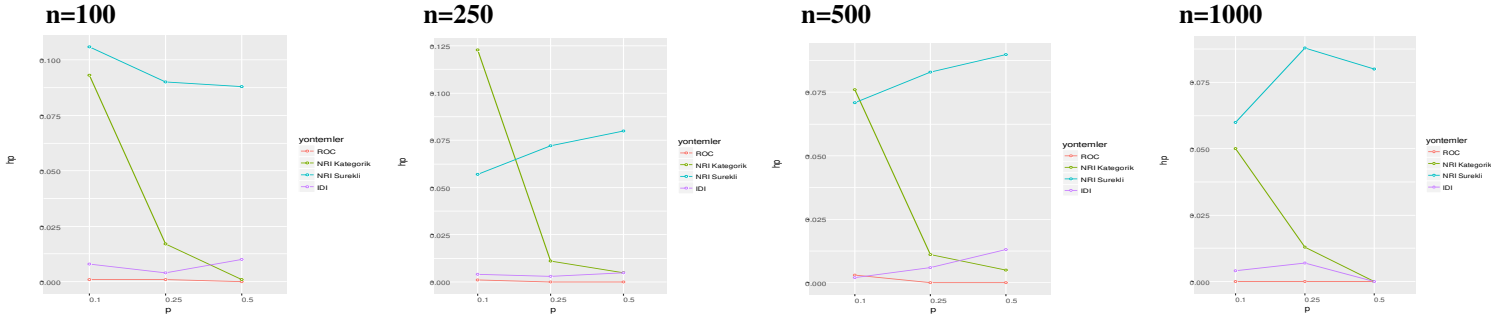
CV=1,05



CV=1,10



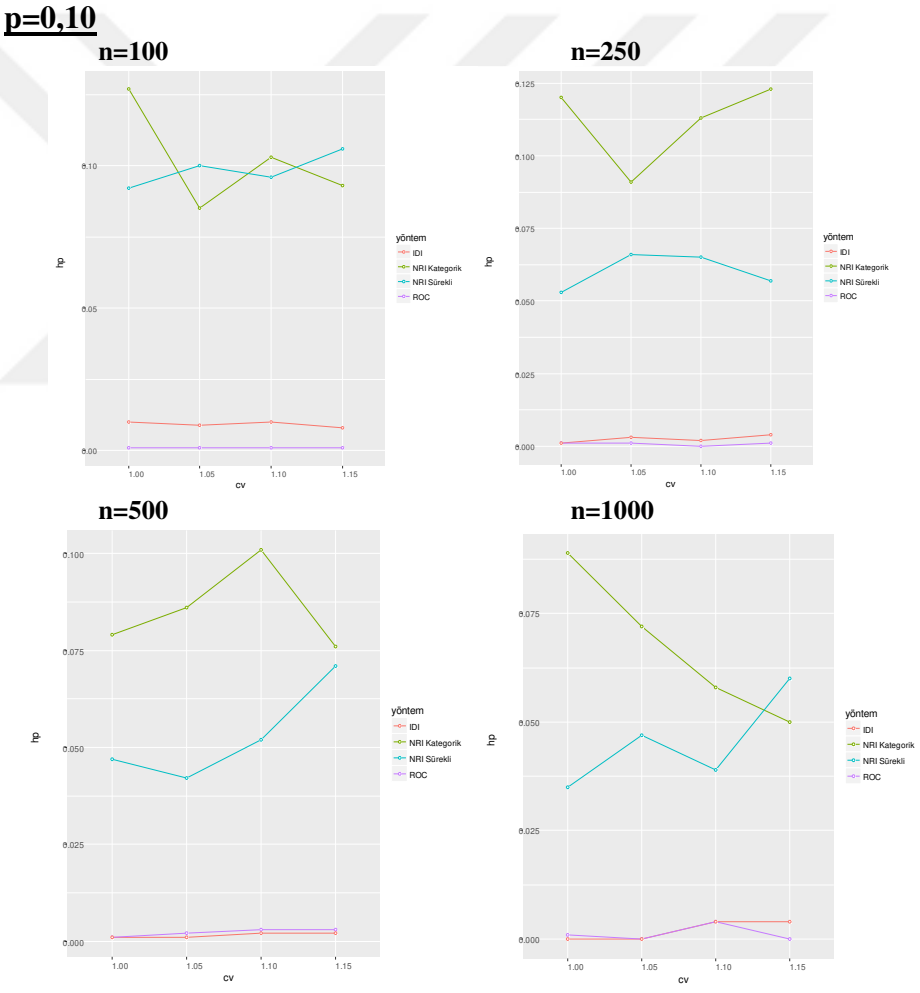
CV=1,15



Şekil 4.2.5. CV=1,05, 1,10 ve 1,15 iken yöntemlerin HP oranlarının farklı prevalans değerlerindeki dağılımı

Diğer CV değişimlerinde, kullanılan yöntemlerin HP oranlarını gösteren grafik Şekil 4.2.5'te verildiği gibidir. CV=1'deki açıklanmaları destekleyecek şekilde diğer CV değişimlerinde de prevalansın 0,10 olduğu durum dışında HP oranının en yüksek olduğu yaklaşım NRI sürekli istatistiğidir. CV= 1,10 dışında örnek büyüklüğü 100 ve prevalans değeri 0,10 iken NRI sürekli istatistiğinin HP oranı diğer yöntemlerden daha yüksektir. Bunun dışında IDI ve ΔAUC istatistiklerinin HP oranı prevalans ve CV değerinden çok net bir şekilde etkilenmemektedir. Prevalans değerinden en net şekilde etkilenen yaklaşım NRI kategoriktir.

4.2.5. Farklı CV oranlarında ve sabit prevalans değerlerinde yöntemlerin HP oranları



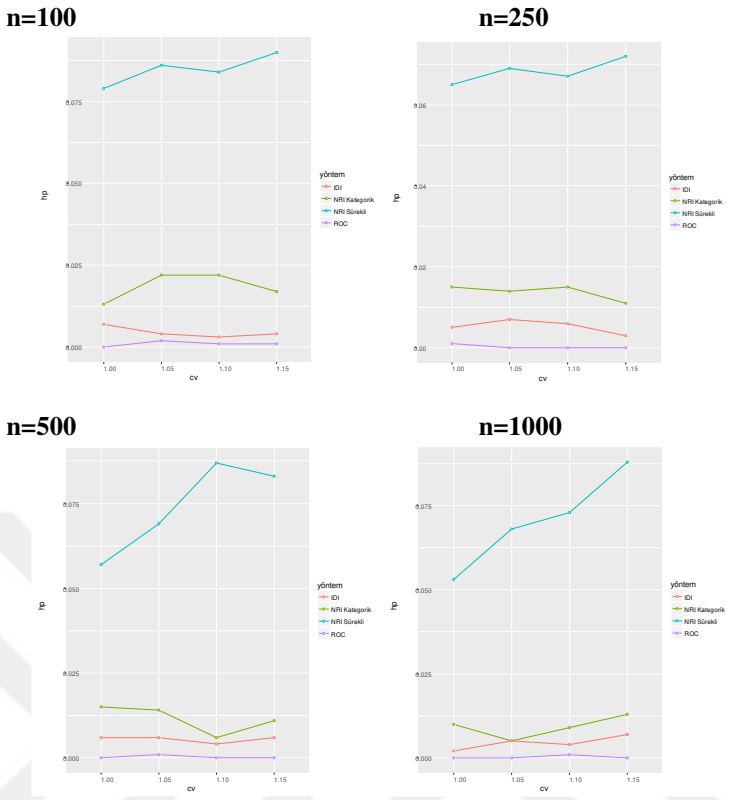
Şekil 4.2.6. p=0,10 iken yöntemlerin HP oranlarının farklı CV oranlarında dağılımı

Şekil 4.2.6'da, prevalans değeri 0,10 iken farklı CV ve örnek büyüklüklerinde yöntemlerin HP oranları gösterilmektedir. CV değerleri arttıkça yöntemlerin HP oranlarının artması beklenmektedir. ΔAUC ve IDI istatistiği CV değişiminden çok fazla

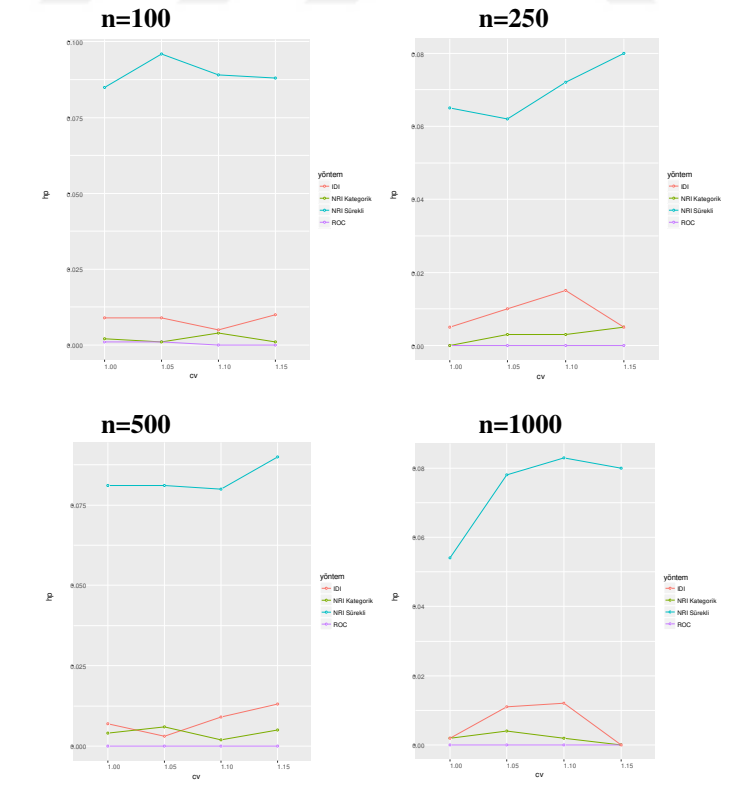
etkilenmemektedir. Sadece örnek büyüklüğü 1000'e çıktığında ve CV'de artış gözlemlendiğinde ΔAUC ve IDI istatistiklerinin HP oranları az da olsa artmaktadır. NRI sürekli istatistiğinin HP oranları CV artışına bağlı olarak genellikle artış göstermektedir (n=250 dışında). Son yöntem olan NRI kategorik istatistiğinin HP oranı ise anlamlı bir seyir göstermemekte hatta n=500 ve n=1000'de HP oranı CV arttıkça düşmektedir.

Örnek büyüklüğünün 100 olduğu durum dışında NRI kategorik istatistiğinin HP oranı en yüksektir. Onu takiben NRI sürekli istatistiği ikinci dereceden en yüksek HP oranına sahip olan yaklaşımdır. ΔAUC ve IDI istatistiklerinin HP oranları neredeyse birbirlerine eşit sadece küçük örnek büyüklüklerinde IDI istatistiğinin HP oranı daha yüksektir.

p=0,25



p=0,50



Şekil 4.2.7. p=0,25 ve 0,50 iken yöntemlerin HP oranlarının farklı CV oranlarında dağılımı

Şekil 4.2.6 ve Şekil 4.2.7'de farklı prevalans değerlerinde artan CV oranına bağlı olarak yöntemlerin HP oran değişimleri gösterilmektedir. Şekil 4.2.6 ve Şekil 4.2.7'de en yüksek HP oranına sahip olan yöntem NRI sürekli iken en düşük HP oranına sahip olan yöntem ΔAUC yöntemidir. Prevalans değeri ve örnek büyüklüğü arttıkça NRI kategorik istatistiğinin HP oranı düşmektedir. Hatta prevalans değeri 0,50 iken NRI kategorik istatistiği 2. sırada en düşük HP oranına sahip olan yöntemdir.

Teorik olarak da bilindiği üzere ΔAUC, NRI sürekli ve IDI istatistikleri prevalans ya da insidans değerinden etkilenmemektedir. NRI kategorik istatistiği ise yapılan simülasyon çalışmasıyla da gösterildiği gibi prevalans değerinden etkilenmekte, prevalans değeri arttıkça HP oranı düşmektedir. Ayrıca ΔAUC ve IDI istatistikleri çoğu senaryoda benzer eğilimler göstermiş ve en düşük HP oranına sahip olan yöntemler olarak belirlenmiştir. Buna karşın prevalans değeri 0,50'ye yükseldiğinde NRI kategorik istatistiği ΔAUC ve IDI istatistiklerinin HP oranlarına yetişmiş hatta HP oranındaki düşüşle IDI istatistiğini geçmiştir.

5.TARTIŞMA

Risk tahmin modelleri, klinik olarak karar verme sürecinde sıklıkla kullanılan bir istatistiksel modelleme tekniğidir [40]. Karar verme sürecinde kullanılan risk tahmin modellerinin tahminlerde ne kadar iyi olduğunun ölçülmesi ise modelin performans ölçütlerine bağlıdır. Performans değerlendirmesinde kullanılan geleneksel ölçüm yöntemleri; genel model performansının değerlendirilmesinde açıklanan varyasyon (R^2) katsayısı ve Brier skoru, ayırım yeteneğinin testi için ROC eğrisi altında kalan alan değeri ve son olarak da kalibrasyon için uyum iyiliği istatistikleri kullanılmaktadır [13]. Bu performans ölçütleri (Brier skoru hariç) ne kadar yüksek ise modelin performansı o kadar iyidir.

Bu tez çalışmasında, risk tahmin modeli olarak SNAP-II ve SNAPPE-II modelleri kullanılmıştır. Yeni eklenen değişkenler ise annenin celestone bebeğin sürfaktan tedavisi alıp almamasıdır. Öncelikle yeni değişkenlere bağlı modelle eski modelin performansları değerlendirilmiştir.

SNAP-II ve SNAPPE-II risk tahmin modellerinin kalibrasyonunu değerlendirmek için kullanılan Hosmer Lemeshow uyum iyiliği testinin p değeri literatüre paralel olarak[35] anlamsız ($p>0,05$) bulunmuştur. Yani gözlem değerleri ile beklenen değerler arasındaki fark istatistiksel olarak anlamlı değildir. Bununla beraber bu modellerin genel performansını değerlendirmek için Nagelkerke R^2 ve Brier skoru kullanılmıştır.

SNAP-II için literatürde bulunan ROC eğrisi altında kalan alan değeri 0,82 (GA:0,68-0,95) [41] olarak bulunurken çalışmada bu değer 0,885 (GA: 0,851-0,919) olarak elde edilmiştir. Ayrıca kesim noktası 14,50 olarak belirlenmiş ve bu noktadaki seçicilik 0,725, duyarlılık 0,897 ve doğruluk 0,750 olarak elde edilmiştir. SNAPPE-II risk tahmin modeli için ise 0,84 ile 0,92 arasında değişen bir ROC eğrisi altında kalan alan değeri söz konusu iken çalışmada bu değer 0,921 (GA: 0,890-0,953)'dir [35]. Ayrıca SNAPPE-II için kesim noktası 26,50 olarak belirlenmiş ve bu noktadaki seçicilik 0,877, duyarlılık 0,857 ve doğruluk 0,874 olarak elde edilmiştir. SNAPPE-II'nin kesim noktası literatürde 13,5 iken seçicilik 0,63 ve duyarlılık 0,67 şeklindedir [42].

SNAP-II risk tahmin modeli için celestone ve sürfaktan bilgisinin eklendiği modelin performansı diğer tüm senaryolardan daha yüksek seyrederken SNAPPE-II risk tahmin modeli için sürfaktan bilgisinin eklendiği senaryo Nagelkerke R^2 ve ROC-AUC değeri açısından daha yüksek bir performans göstermiştir. SNAP-II risk tahmin modelinin celestone ve sürfaktana bağlı modeli için ROC-AUC değeri 0,910 (GA: 0,880-0,940) olarak elde edilmiştir. Yani ROC-AUC değeri 0,885 (GA: 0,851-0,919)'ten 0,910 (GA: 0,880-0,940)'a yükselmiştir. Risk skorlarına dayanan kesim noktası değeri 20,50 ve bu noktadaki seçicilik 0,816, duyarlılık 0,849 ve doğruluk 0,821 olarak elde edilmiştir. SNAPPE-II risk tahmin modeli için sürfaktan bilgisinin eklendiği senaryo için ROC-AUC değeri 0,930 (GA: 0,902-0,958)'dur (Şekil 4.2.3). Bu durumda ise ROC-AUC değeri 0,921 (GA: 0,890-0,953)'den 0,930 (GA: 0,902-0,958)'a yükselmiştir. Kesim noktası 25,50 olarak belirlenmiş ve bu noktadaki seçicilik 0,854, duyarlılık 0,881 ve doğruluk 0,858 olarak elde edilmiştir.

Risk tahmin modellerinin performanslarının değerlendirilmesinin ardından kullanılan risk tahmin modellerine yeni eklenecek değişken ya da değişkenlerin risk tahminindeki katkısını araştırmak epidemiyolojik araştırmaların önemli bir parçasıdır. Geleneksel olarak yeni risk faktörünün tahmindeki katkısını değerlendirmede ROC-AUC istatistiği kullanılmaktadır. Ancak yapılan pek çok çalışma ile bu istatistiğin tahmindeki artışı test etmede oldukça tutucu sonuçlar verdiği ileri sürülmüştür. ROC-AUC istatistiğinin, anlamlı sonuçlar verdiği senaryo sadece rölatif risk değerinin oldukça yüksek seviyelerde olması durumudur [40]. Örneğin, Wang ve Gona [43] yapmış oldukları çalışmada kardiyovasküler hastalıklar üzerinde etkisi istatistiksel olarak önemli olan pek çok markeri modele ilave etmelerine rağmen AUC değerindeki değişim sadece 0.76'dan 0,77'ye yükselmiştir. Benzer şekilde Pencina, D'Agostino ve arkadaşları çalışmalarında HDL kolestrol değerinin sonuç değişkeni üzerindeki ilişkisi istatistiksel olarak yüksek düzeylerde anlamlı olmasına rağmen proportional hazard regresyon modeline ilavesinde c istatistik değerindeki artış anlamlı bulunmamıştır. Bu durum DeLong test prosedürünün oldukça tutucu sonuçlar vermesi nedeninden kaynaklandığı düşünülmektedir [44].

Çalışmada ΔAUC istatistiği, SNAP-II risk tahmin modeline sürfaktan ve hem sürfaktan hem de celestone bilgisinin eklenmesi ile elde edilen modeller için anlamlı sonuçlar vermiştir. Bu durumlar için AUC artışları sırasıyla 0,022 ve 0,025 şeklindedir.

NRI sürekli ve IDI istatistiklerinin yüksek düzeyde anlamlı (p<0,001) olarak sonuç verdiği SNAP-II + celestone risk tahmin modeli için ΔAUC istatistiği anlamsız bir sonuç vermiştir (p=0,231). SNAPPE-II risk tahmin modeline benzer şekilde celestone eklendiği senaryoda NRI sürekli istatistiği yüksek düzeyde anlamlı sonuç verirken (p=0,003) IDI ve ΔAUC istatistiği yüksek düzeyde anlamsız bir sonuç vermiştir (p değerleri sırasıyla 0,226 ve 0,908). Ayrıca simülasyon çalışmasında ΔAUC istatistiğinin p değerinin dağılımı en yüksek seviyelerde seyrederken buna bağlı olarak HP oranı en düşük seyretmiştir. Ek olarak örnek büyüklüğünden, prevalanstan ve CV değişiminden etkilenmemektedir.

ΔAUC istatistiğinin eksikliklerini gidermek amacıyla ileri sürülen yaklaşımlardan biri olan NRI istatistiğinin geçerliliği yapılan pek çok çalışmayla zarar görmüştür. Pepe ve arkadaşlarının yapmış oldukları çalışmada, NRI istatistiğinin ΔAUC istatistiğine göre daha fazla miktarda hatalı-pozitifliğe sahip sonuçlar verdiğini göstermişlerdir. Sürekli NRI istatistiği ve p değeri, eklenen değişkenin tahminde herhangi bir önemi olmasa da yüksek frekanslarda anlamlı sonuçlar vermektedir. Hatta düşük seviyelerde fit edilmiş modellerde bile NRI istatistiğinin pozitif sonuçlar verdiği literatürde yer alan bilgiler arasındadır [29, 45,46].

Çalışmada literatüre paralel olarak NRI sürekli istatistiği, SNAPPE-II risk tahmin modeline sürfaktan bilgisinin eklenmesi durumu dışındaki her durumda anlamlı sonuç vermiştir. SNAPPE-II risk tahmin modeline eklenen değişkenlerle model performanslarındaki artışlar yüksek sayılmayacak seviyelerde olduğu görülmüştür. Örneğin, SNAPPE-II + celestone modelinin AUC değerinde herhangi bir değişim söz konusu değil iken NRI sürekli istatistiği anlamlı bir sonuç vermiştir. Ayrıca simülasyon çalışmasında NRI sürekli istatistiğinin p değerinin dağılımı en alt seviyelerde seyrederken HP oranı buna bağlı olarak en yüksek seyretmiştir. Bulunan NRI sürekli istatistiğinin test değerinin ortalaması ile standart sapma arasındaki fark yüksek değildir. Yani bu istatistik değerinin standart sapması olduça yüksektir. Buna bağlı olarak test istatistiğinin dağılımı geniş bir aralıkta seyretmektedir. Ayrıca bu istatistiğin HP oranı CV değeri arttıkça artmakta prevelans değeri arttıkça azalmaktadır.

IDI istatistiği, SNAP-II risk tahmin modeline eklenen her bir değişken durumu için anlamlı sonuç vermiştir. SNAPPE-II risk tahmin modeli için anlamlı sonuç verdiği

senaryo ile karşılaşılmamıştır. IDI istatistiğinin HP oranı, ΔAUC istatistiğine benzer olarak oldukça düşük seviyelerdedir. ΔAUC ve IDI istatistikleri CV değişiminden çok fazla etkilenmez iken sadece yüksek örnek büyüklüklerinde (n=1000) CV değerinde artış ile birlikte HP oranı az da olsa artmaktadır. Bu benzerliği bozan tek senaryo SNAP-II + celestone risk tahmin modelinde gözlemlenmiştir. Bu senaryoda IDI istatistiği yüksek düzeyde anlamlı ($p<0,001$) sonuç verirken ΔAUC istatistiği anlamlı sonuç vermemiştir ($p=0,231$). Bu sonuçtan hareketle, eklenen değişken yüksek düzeyde anlamlı ve modelin performansındaki artış yüksek seviyelerde ise IDI istatistiği ΔAUC istatistiğinden farklı olarak anlamlı sonuç vermeye daha yatkındır.

NRI kategorik istatistik değeri, SNAP-II risk tahmin modeline hem sürfaktan hem de celestone bilgisinin eklenmesi ile 0,271 olarak bulunurken SNAPPE-II risk tahmin modeline celestone bilgisinin eklenmesi ile -0,036 ve hem sürfaktan hem de celeston bilggisinin eklenmesi ile -0,069 bulunmuştur. Ancak p değerlerine bakıldığında SNAP-II + sürfaktan + celestone modelinin p değeri 0,271 iken SNAPPE-II + celestone için p değeri 0,035 ve SNAPPE-II + sürfaktan + celestone için 0,004 bulunmuştur. Model performansındaki artış yüksek iken anlamsız, model performansında artışın aksine azalış ve test istatistiğinin daha düşük seviyelerde olduğu gözlemlenmesine rağmen istatistiksel olarak anlamlı sonuçlar vermiştir. Simülasyon sonuçlarına dayanarak, NRI kategorik istatistiği CV değerinden net bir şekilde etkilenmezken örneklem büyüklüğü arttıkça az da olsa HP oranı azalmaktadır. Bunlara karşın prevalans değeri, NRI kategorik istatistiğinin HP oranını oldukça etkilemekte prevalans arttıkça HP oranı düşmektedir.

IDI istatistiğinin literatürde pek çok uzantısı bulunmaktadır. Teorik olarak bu istatistik, regresyon yöntemlerinin tanımlamalarını içermektedir. Bu nedenle Pepe ve arkadaşları yapmış oldukları çalışmalarında IDI istatistiğinin teorik olarak daha güvenilir olduğunu vurgulamışlardır. Diğer yandan NRI istatistiği, tıbbi dergilerde oldukça sık rastlanan bir yaklaşım haline gelmiştir. Hiç şüphesiz bunun en büyük nedeni uygulama kolaylığına sahip olmasıdır [7].

Ayrıca klinik olarak modellerin performansları değerlendirilidiğinde SNAP-II risk tahmin modeline hem sürfaktan hem de celestone bilgisinin eklenmesi ile elde edilen modelin performansı tüm senaryolardan daha iyidir. Bu değerlendirme istatistiksel olarak anlamlı da bulunmuştur. Bununla beraber SNAPPE-II'ye eklenen

sürfaktan değişkeni ile modelin performansı diğer senaryolardan daha iyi seyretsede bu değerlendirme istatistiksel olarak anlamlı bulunmamıştır. Ayrıca son olarak SNAP-II ve SNAPPE-II için eklenen değişkenlere bağlı model ile eski modelin performansı sırasıyla değerlendirilmiş her senaryoda SNAPPE-II risk tahmin modeli daha iyi bir performans göstermiş ve bu performans artışı istatistiksel olarak anlamlı bulunmuştur.

Çalışmanın en önemli eksikliği kullanılan dört yönteminde altın standart niteliğinde olmamasından dolayı hangi yöntemin en doğru sonuçlar verdiği kesin kes söylenememektedir. Buna paralel olarak yöntemlerin sadece HP oranları verilebilmiş hatalı negatif oranları belirlenememiştir. Sadece eklenen değişkenlerin anlamlılıklarına göre yorum yapılmıştır. Ayrıca modellerin, kesin bir kesim noktasının olmaması nedeniyle NRI kategorik istatistiği için net bir çıkarsama yapılamamıştır.

6.SONUÇ ve ÖNERİLER

Öncelikle gerçek veri setine dayanarak varılabilecek en önemli sonuçlardan biri eğer risk tahmin modeli olarak SNAP-II kullanılacaksa bu risk tahmin modeline celestone ve sürfaktan değişkenlerinin eklenmiş hali tahminlerde daha başarılıdır. Bunun yanında risk tahmin modeli olarak SNAPPE-II kullanılacaksa bu risk tahmin modeline yeni değişkenler katkı sağlamadığından SNAPPE-II risk tahmin modeli olduğu gibi kullanılmalıdır. Bununla beraber SNAPPE-II risk tahmin modelinin tahmin performansı SNAP-II ve uzantılarından daha iyidir.

İçe içe tasarıma sahip modellerin performansını değerlendirirken kullanılan yöntemler için varılabilecek sonuçlar ise NRI sürekli istatistiği HP oranı en yüksek olan yaklaşım iken ΔAUC istatistiği HP oranı en düşük olan yaklaşımdır. Buna ek olarak ΔAUC istatistiği yüksek düzeyde anlamlı değişken senaryosunda bile H_0'ı kabul etme eğilimindedir. Yani oldukça tutucu bir istatistiksel tekniktir. NRI sürekli istatistiği ise tek bir senaryo dışında H_0'ı reddetme eğilimindedir. NRI kategorik istatistiği ise prevalans değerinden oldukça fazla etkilenmekte ve düşük prevalansta yüksek miktarda HP oranı sergilemektedir. Bununla beraber kullanılan risk tahmin modelinin net bir kategori sınıflaması yoksa teorik olarak kullanımı sakıncalı görülmektedir. Son olarak IDI istatistiği HP oranı düşük seviyelerde olan bir yaklaşımdır. ΔAUC istatistiğinden farklı olarak yüksek düzeyde anlamlı değişkenin modele eklenmesi senaryosunda H_0'ı reddetmeye daha yatkındır. Bu bağlamda IDI istatistiğinin kullanımı yapılan çaışmayla daha makul görülmüştür.

Risk tahmin modeli olarak lojistik regresyon değil de Cox regresyon analizi ile yöntemlerin değerlendirilmesi ikinci bir çalışma olarak gerçekleştirilebilir. Ayrıca kategori sınıflamasının iki değil de daha fazla olduğu durumlarda NRI kategorik istatistiğinin değerlendirilmesi yapılabilecek başka bir çalışmadır.

7.KAYNAKLAR

1. **Demler OV**. Impact of new variables on discrimination of risk prediction models (Doctoral dissertation, BOSTON UNIVERSITY). **2012**; 7-11.

2. **Steyerberg EW**. *Clinical prediction models: a practical approach to development, validation, and updating*. Springer Science & Business Media, **2008**; 255-279.

3. **D'Agostino RB, Griffith JL, Schmidt CH and Terrin N**. Measures for evaluating model performance. *Proceedings of the biometrics section.* American Statistical Association, Biometrics Section: Alexandria, VA. U.S.A. **1997**; 253-258.

4. **Kessler RC, Abelson J, Demler O, Escobar JI, Gibbon M, Guyer ME, Howes MJ, Jin R, Vega WA, Walters EE, Wang P, Zaslavsky A, Zheng H.** Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *The International Journal of Methods in Psychiatric Research*, **2004**;13(2), 122-139.

5. **Cook NR**. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*, **2007**; *115*(7), 928-935.

6. **Pencina MJ, D'Agostino RB, Vasan RS.** Comments on 'Integrated discrimination and net reclassification improvements—Practical advice'. *Statistics in Medicine*, **2008**; *27*(2), 207-212.

7. **Pencina, MJ, D'Agostino RB, Steyerberg EW**. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Statistics in medicine*, **2011**; *30*(1), 11-21.

8. **Hosmer Jr DW, Lemeshow S.** *Applied logistic regression*. John Wiley & Sons. **2004**.

9. **David K, Mitchel K**. Logistic regression: A self learning text. **2010**; 73-100.

10. **Alpar R.** Uygulamalı Çok Değişkenli İstatistiksel Yöntemler (3. Baskı).*Ankara: Detay*. **2011**.

11. **Ünal İ.** Prognozda ve tanıda lojistik regresyon ve neural network yaklaşımı: Epidemiyolojik Veri Setlerinde Karşılaştırmalı Uygulamalar. (Yüksek Lisans Tezi, Çukurova Üniversitesi). **2004**; 3-4.

12. **Pepe M, Janes H**. Methods for evaluating prediction performance of biomarkers and tests. In *Risk Assessment and Evaluation of Predictions*. Springer New York. **2013**; 107-142.

13. **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 some traditional and novel measures. *Epidemiology (Cambridge, Mass.)*, **2010**; *21*(1), 128.

14. **Gu W**. *Evaluating risk prediction markers and models*. University of Washington. **2009**.

15. **Nagelkerke NJ**. A note on a general definition of the coefficient of determination. *Biometrika*, **1991**; *78*(3), 691-692.

16. **Bewick V, Cheek L, Ball J.** Statistics review 14: Logistic regression. Critical Care (Review). **2005**; 9(1), 112-115.

17. **Ikeda M, Ishigaki T, Yamauchi K**. Relationship between Brier score and area under the binormal ROC curve. *Computer methods and programs in biomedicine*, **2002**; *67*(3), 187-194.

18. **Zou KH, O'Malley AJ, Mauri L**. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, **2007**; *115*(5), 654-657.

19. **Pepe MS**. An interpretation for the ROC curve and inference using GLM procedures. *Biometrics*, **2000**; *56*(2), 352-359.

20. **Pepe MS**. A regression modelling framework for receiver operating characteristic curves in medical diagnostic testing. *Biometrika*, **1997**; *84*(3), 595-608.

21. **DeLong ER, DeLong DM, Clarke-Pearson DL**. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, **1988**; 837-845.

22. **Cook N, Ridker P**. Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassificationmeasures. Annals of internalmedicine **2009**; 150(11), 795–802.

23. **Kuhn M, Johnson K**. *Applied predictive modeling*. New York: Springer, **2013**.

24. **Pepe MS**. *The statistical evaluation of medical tests for classification and prediction*. Oxford University Press, USA. **2003**.

25. **Pencina MJ, D'agostino RB, Pencina KM, Janssens ACJ, Greenland P.** Interpreting incremental value of markers added to risk prediction models. *American journal of epidemiology*, **2012**; 1-5.

26. **Pepe MS, Kerr KF, Longton G, Wang Z**. Testing for improvement in prediction model performance. *Statistics in medicine*, **2013**; *32*(9), 1467-1482.

27. **Pencina MJ, D'Agostino 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.

28. **Van Calster B, Vickers AJ, Pencina MJ, Baker SG, Timmerman D, Steyerberg EW**. Evaluation of markers and risk prediction models overview of relationships between NRI and decision-analytic measures. *Medical Decision Making*, **2013**; *33*(4), 490-501.

29. **Pepe MS, Janes H, Li CI**. Net risk reclassification P values: valid or misleading?. *Journal of the National Cancer Institute*, **2014**; *106*(4), 1-6.

30. **Pencina MJ, D'Agostino 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.

31. **Pencina MJ, D'Agostino RB, Demler OV**. Novel metrics for evaluating improvement in discrimination: net reclassification and integrated discrimination improvement for normal variables and nested models. *Statistics in medicine*, **2012**; *31*(2), 101-113.

32. **Vickers AJ**. Decision analysis for the evaluation of diagnostic tests, prediction models, and molecular markers. *The American Statistician*. **2012**; 314-319.

33. **Vickers AJ, Elkin EB**. Decision curve analysis: a novel method for evaluating prediction models. *Medical Decision Making*, **2006**; *26*(6), 565-574.

34. **Vickers AJ, Cronin AM, Elkin EB, Gonen M**. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC medical informatics and decision making*, **2008**; *8*(1), 1.

35. **Richardson DK, Corcoran JD, Escobar GJ, Lee SK.** SNAP-II and SNAPPE-II: simplified newborn illness severity and mortality risk scores.*The Journal of pediatrics*, **2001**; *138*(1), 92-100.

36. **McFadden D**. *Quantitative methods for analyzing travel behavior of individuals: some recent developments*. Institute of Transportation Studies, University of California. **1977**.

37. **IBM Corp.** *IBM SPSS Statistics for Windows,* Version 20.0. Armonk, NY: IBM Corp., **2011**.

38. **R Foundation for Statistical Computing.** *R: A language and environment for statistical computing.,* Vienna, Austria: R Core Team, **2013.**

39. **Richardson DK, Phibbs CS, Gray JE, McCormick MC, Workman-Daniels K, Goldmann DA**. Birth weight and illness severity: independent predictors of neonatal mortality. *Pediatrics*, **1993**; *91*(5), 969-975.

40. **Mühlenbruch K, Heraclides A, Steyerberg EW, Joost HG, Boeing H, Schulze MB**. Assessing improvement in disease prediction using net reclassification improvement: impact of risk cut-offs and number of risk categories. *European journal of epidemiology*, **2013**; *28*(1), 25-33.

41. **Sundaram V, Dutta S, Ahluwalia J, Narang A.** Score for neonatal acute physiology II predicts mortality and persistent organ dysfunction in neonates with severe septicemia. *Indian pediatrics*, **2009**; *46*(9), 775.

42. **Asker HS, Satar M, Yapicioglu Yildizdas H, Mutlu B, Mutlu B, İpek MŞ, Sivaslı E, Taviloğlu Ş, Çelik Y, Özcan K, Burgut R, Ünal İ**. Evaluation of the SNAP-PE-II and CRIB scoring systems with additional parameters. *Pediatrics International*. **2016**.

43. **Wang TJ, Gona P**. Multiple Biomarkers for the Prediction of First Major Cardiovascular Events and Death. The New England Journal of Medicine. **2006**; 355(25):2631-2639.

44. **Kessler RC, Chiu WT, Demler O, Walters EE**. Prevalence, severity, and comorbidity of twelve-month DSM-IV disorders in the National Comorbidity Survey Replication (NCS-R). *Archives of General Psychiatry*, **2005**; 62(6), 617-627. (Demler12)

45. **Hilden J, Gerds TA**. A note on the evaluation of novel biomarkers: do not rely on integrated discrimination improvement and net reclassification index. *Statistics in medicine*, **2014**; *33*(19), 3405-3414.

46. **Pepe M, Fang J, Feng Z, Gerds T, Hilden J**. The net reclassification index (NRI): A misleading measure of prediction improvement with miscalibrated or overfit models. **2013**; 392.

47. **Pencina MJ, D'Agostino RB, Vasan RS**. Comments on 'Integrated discrimination and net reclassification improvements—Practical advice'. *Statistics in Medicine*, **2008**; *27*(2), 207-212.

EKLER

R Kodları

Simülasyon İçin Kodlar

```
betul<-function(){
        library(OptimalCutpoints)
        library(pROC)
        library(BaylorEdPsych)
        library(mvtnorm)
        library(PredictABEL)
        library(Hmisc)
        library(descr)
}

betul()

reclassification2 <-function(data,cOutcome,pr1,pr2, cutoff1, cutoff2)
{
        c1 <- cut(pr1, breaks = cutoff1 ,include.lowest=TRUE,right= FALSE)
        c2 <- cut(pr2, breaks = cutoff2 ,include.lowest=TRUE,right= FALSE)
        c11 <-factor(c1, levels = levels(c1), labels = c(1:length(levels(c1))))
        c22 <-factor(c2, levels = levels(c2), labels = c(1:length(levels(c2))))
        x<-improveProb(x1=as.numeric(c11)*(1/(length(levels(c11)))),
        x2=as.numeric(c22)*(1/(length(levels(c22)))), y=data[,cOutcome])
        y<-improveProb(x1=pr1, x2=pr2, y=data[,cOutcome])
        dd<-2*pnorm(-abs(x$z.nri))
        ee <- 2*pnorm(-abs(y$z.nri))
        ff<-2*pnorm(-abs(y$z.idi))
        abc <- c(dd,ee,ff)
        return(abc)
}

vv <- function(nn,p1,kontrol2, kontrol1, ps1, ps2, sig1, ort1, ort2, CV) {
        n1<-round(nn*(1-p1))
        n2<-nn-n1
        grup1 <- rep(0,n1)
        grup2 <- rep(1,n2)
        y<- c(grup1,grup2)
        repeat
        {
                g1   <- rmvnorm(n1, ort1, sig1)
                g2   <- rmvnorm(n2, ort2,sig1)
                g<-rbind(g1,g2)
                noise<-rnorm(nn,2,2)
                g<-cbind(g,noise)
                veri<-data.frame(y=y,g)
                mylogit <-  glm(y ~ V1 +V2 , data = veri, family = "binomial")
                uni     <-  glm (y ~ V3, data = veri, family = "binomial")
                sinama  <-  coef(summary(uni))[,4][[2]]
```

```
		mylogit2 <- glm(y ~ V1+V2+ V3, data = veri, family = "binomial")
		r <- PseudoR2(mylogit)[[1]]
		if(r > kontrol1  && kontrol2 > r && sinama > ps1 && sinama <= ps2)
			break
	}


	pr1<-predRisk(mylogit)
	pr2<-predRisk(mylogit2)
	rc <- roc.test(veri$y, pr1, pr2)

	veri2 <- cbind(veri, pr1,pr2)


	aa <- optimal.cutpoints(X="pr1", status="y"  ,  tag.healthy = 0, methods = "Youden", data=veri2)
	kesim1  <-summary(aa)[[5]][[1]][[1]][[1]][[1]]
	cutoff1 <- c(0,kesim1,1)

	bb  <-  optimal.cutpoints(X="pr2",  status="y"    ,    tag.healthy  =  0,  methods  =  "Youden",
data=veri2)
	kesim2  <-summary(bb)[[5]][[1]][[1]][[1]][[1]]
	cutoff2 <- c(0,kesim2,1)

	se1<-summary(aa)[[5]][[1]][[1]][[1]][[2]]
	sp1<-summary(aa)[[5]][[1]][[1]][[1]][[3]]
	se2<-summary(bb)[[5]][[1]][[1]][[1]][[2]]
	sp2<-summary(bb)[[5]][[1]][[1]][[1]][[3]]

	se_sp_top_1<-se1+sp1
	se_sp_fark_1<-se1-sp1

	se_sp_top_2<-se2+sp2
	se_sp_fark_2<-se2-sp2
	ROC_iyi<-0
	if (se_sp_top_2 >= se_sp_top_1)
	{
		if (se_sp_fark_2 <= se_sp_fark_1)
		{
			ROC_iyi<-1
	}
	}




	rc <- roc.test(veri$y, pr1, pr2)

	NRIIDI <- improveProb(pr1, pr2, veri$y)
	denemeler <- reclassification2(veri,1,pr1,pr2,cutoff1, cutoff2)
	pr1.pos <- as.numeric(pr1        >= kesim1)
	pr2.pos <- as.numeric(pr2        >= kesim2)
	catNriElements <- with(veri, by(pr2.pos - pr1.pos, veri$y,
		function(x) {
		  ## Probability of up/down/unchanged
```

```
                    propChanges <- prop.table(table(sign(x)))
                    ## up - down
                    propChanges["1"] - propChanges["-1"]
                  }))
        NRIcat <-(catNriElements["1"] - catNriElements["0"])[[1]]
        NRIcon <- NRIIDI[[8]][[1]]
        IDIdeg <- NRIIDI[[19]][[1]]
        NRIcatpdeger <- denemeler[1]
        NRIconpdeger <- denemeler[2]
        IDIpdeger <- denemeler[3]
        ROCtest <- rc[[7]][[1]]
        ROC1AUC <- rc[[3]][[1]]
        ROC2AUC <- rc[[3]][[2]]



        sonuclar <- c(NRIcat, NRIcatpdeger, NRIcon, NRIconpdeger, IDIdeg,
                      IDIpdeger, ROCtest, ROC1AUC,ROC2AUC, sinama, CV)



return(c(sonuclar,ROC_iyi, se1, sp1, se2, sp2))

}

simu<-function(nn,p1,kontrol2, kontrol1,ps1, ps2, sig1,  ort1, ort2, CV, tekrar)
{
        dene<-matrix(c(0), nrow=tekrar, ncol=16)
        for(k in 1:tekrar){
                dene[k,]<-vv(nn,p1,kontrol2, kontrol1,ps1, ps2, sig1, ort1, ort2, CV)
        }
        return(dene)
}

snclma<-function(snclar)
{
        ROC.pos      <- as.numeric(snclar[,7]  < 0.05)
        NRIcat.pos       <-      as.numeric(snclar[,2]  < 0.05)
        NRIcont.pos      <-      as.numeric(snclar[,4]  < 0.05)
        IDI.pos          <-      as.numeric(snclar[,6]  < 0.05)

        HPROC<- mean(ROC.pos, na.rm=TRUE)
        HPNRIcat <- mean(NRIcat.pos, na.rm=TRUE)
        HPNRIcont <- mean(NRIcont.pos, na.rm=TRUE)
        HPIDI<- mean(IDI.pos, na.rm=TRUE)

        par(mfrow=c(2,2))
        hist(snclar[,7], main=" ", xlab="ROC testi için p değerleri", ylab="GörülmeSıklığı")
        hist(snclar[,2], main=" ", xlab="NRI kategorik testi için p değerleri", ylab="GörülmeSıklığı")
        hist(snclar[,4], main=" ", xlab="NRI sürekli için p değerleri", ylab="GörülmeSıklığı")
        hist(snclar[,6], main=" ", xlab="IDI testi için p değerleri", ylab="GörülmeSıklığı")

        return(c(HPROC, HPNRIcat, HPNRIcont, HPIDI))
}
```

```
sig <- matrix(c(1,0,0,0,1,0,0,0,1),3,3)



#####********************##############
snc100_0.1CV1 <- simu(100,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2),1,1000)
sonn100_0.1CV1 <- snclma(snc100_0.1CV1)

snc100_0.1CV1.05 <- simu(100,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.1),1.05,1000)
sonn100_0.1CV1.05 <- snclma(snc100_0.1CV1.05)


snc100_0.1CV1.1 <- simu(100,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.2),1.1,1000)
sonn100_0.1CV1.1 <- snclma(snc100_0.1CV1.1)

snc100_0.1CV1.15 <- simu(100,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.3),1.15,1000)
sonn100_0.1CV1.15 <- snclma(snc100_0.1CV1.15)

        NRIcatbir100_0.1 <- c(snc100_0.1CV1[,2], snc100_0.1CV1.05[,2],snc100_0.1CV1.1[,2],
snc100_0.1CV1.15[,2])
        NRIconbir100_0.1 <- c(snc100_0.1CV1[,4], snc100_0.1CV1.05[,4], snc100_0.1CV1.1[,4],
snc100_0.1CV1.15[,4])
        IDIbir100_0.1 <- c(snc100_0.1CV1[,6], snc100_0.1CV1.05[,6], snc100_0.1CV1.1[,6],
snc100_0.1CV1.15[,6])
        ROCbir100_0.1 <- c(snc100_0.1CV1[,7],
snc100_0.1CV1.05[,7],snc100_0.1CV1.1[,7],snc100_0.1CV1.15[,7])


        istat100_0.1 <- c(NRIcatbir100_0.1, NRIconbir100_0.1,IDIbir100_0.1, ROCbir100_0.1)
        CVbir100_0.1 <- c(rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),rep(1,1000),
rep(1.05,1000),rep(1.1,1000), rep(1.15,1000), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),
rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000))
        iss100_0.1 <- c(rep(1,4000), rep(2,4000), rep(3,4000),rep(4,4000))



birveri100_0.1 <- data.frame(istat100_0.1, CVbir100_0.1, iss100_0.1)

birveri100_0.1$iss100_0.1[birveri100_0.1$iss100_0.1==1]<-"NRIkategorik"
birveri100_0.1$iss100_0.1[birveri100_0.1$iss100_0.1==2]<-"NRIsürekli"
birveri100_0.1$iss100_0.1[birveri100_0.1$iss100_0.1==3]<-"IDI"
birveri100_0.1$iss100_0.1[birveri100_0.1$iss100_0.1==4]<-"ROC"

birveri2_100_0.1 <-remove_missing(birveri100_0.1, na.rm = TRUE, vars = names(birveri100_0.1), name
= "",
  finite = FALSE)

library(Rmisc)
tgc100_0.1 <- summarySE(birveri2_100_0.1, measurevar="istat100_0.1",
groupvars=c("iss100_0.1","CVbir100_0.1"))
plot100_0.1 <- ggplot(tgc100_0.1, aes(x=CVbir100_0.1, y=istat100_0.1, colour=iss100_0.1)) +
```

```
   geom_errorbar(aes(ymin=istat100_0.1-se, ymax=istat100_0.1+se), width=.01) +
   geom_line() +
   geom_point()+

          ylab("P Değerleri")+
          xlab("CV")


##############################################
#####p=0,10 için##########
snc250_0.1CV1 <- simu(250,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2),1,1000)
sonn250_0.1CV1 <- snclma(snc250_0.1CV1)

snc250_0.1CV1.05 <- simu(250,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.1),1.05,1000)
sonn250_0.1CV1.05 <- snclma(snc250_0.1CV1.05)


snc250_0.1CV1.1 <- simu(250,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.2),1.1,1000)
sonn250_0.1CV1.1 <- snclma(snc250_0.1CV1.1)

snc250_0.1CV1.15 <- simu(250,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.3),1.15,1000)
sonn250_0.1CV1.15 <- snclma(snc250_0.1CV1.15)

          NRIcatbir250_0.1 <- c(snc250_0.1CV1[,2], snc250_0.1CV1.05[,2],snc250_0.1CV1.1[,2],
snc250_0.1CV1.15[,2])
          NRIconbir250_0.1 <- c(snc250_0.1CV1[,4], snc250_0.1CV1.05[,4], snc250_0.1CV1.1[,4],
snc250_0.1CV1.15[,4])
          IDIbir250_0.1 <- c(snc250_0.1CV1[,6], snc250_0.1CV1.05[,6], snc250_0.1CV1.1[,6],
snc250_0.1CV1.15[,6])
          ROCbir250_0.1 <- c(snc250_0.1CV1[,7],
snc250_0.1CV1.05[,7],snc250_0.1CV1.1[,7],snc250_0.1CV1.15[,7])


          istat250_0.1 <- c(NRIcatbir250_0.1, NRIconbir250_0.1,IDIbir250_0.1, ROCbir250_0.1)
          CVbir250_0.1 <- c(rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),rep(1,1000),
rep(1.05,1000),
          rep(1.1,1000), rep(1.15,1000), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),
          rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000))
          iss250_0.1 <- c(rep(1,4000), rep(2,4000), rep(3,4000),rep(4,4000))



birveri250_0.1 <- data.frame(istat250_0.1, CVbir250_0.1, iss250_0.1)

birveri250_0.1$iss250_0.1[birveri250_0.1$iss250_0.1==1]<-"NRIkategorik"
birveri250_0.1$iss250_0.1[birveri250_0.1$iss250_0.1==2]<-"NRIsürekli"
birveri250_0.1$iss250_0.1[birveri250_0.1$iss250_0.1==3]<-"IDI"
birveri250_0.1$iss250_0.1[birveri250_0.1$iss250_0.1==4]<-"ROC"

birveri2_250_0.1 <-remove_missing(birveri250_0.1, na.rm = TRUE, vars = names(birveri250_0.1), name
= "",
  finite = FALSE)

library(Rmisc)
```

```
tgc250_0.1 <- summarySE(birveri2_250_0.1, measurevar="istat250_0.1", groupvars=c("iss250_0.1","CVbir250_0.1"))
plot250_0.1 <- ggplot(tgc250_0.1, aes(x=CVbir250_0.1, y=istat250_0.1, colour=iss250_0.1)) +
   geom_errorbar(aes(ymin=istat250_0.1-se, ymax=istat250_0.1+se), width=.01) +
   geom_line() +
   geom_point()+

          ylab("P Değerleri")+
          xlab("CV")
##############################################

snc500_0.1CV1 <- simu(500,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2),1,1000)
sonn500_0.1CV1 <- snclma(snc500_0.1CV1)

snc500_0.1CV1.05 <- simu(500,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.1),1.05,1000)
sonn500_0.1CV1.05 <- snclma(snc500_0.1CV1.05)


snc500_0.1CV1.1 <- simu(500,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.2),1.1,1000)
sonn500_0.1CV1.1 <- snclma(snc500_0.1CV1.1)

snc500_0.1CV1.15 <- simu(500,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.3),1.15,1000)
sonn500_0.1CV1.15 <- snclma(snc500_0.1CV1.15)

          NRIcatbir500_0.1 <- c(snc500_0.1CV1[,2], snc500_0.1CV1.05[,2],snc500_0.1CV1.1[,2], snc500_0.1CV1.15[,2])
          NRIconbir500_0.1 <- c(snc500_0.1CV1[,4], snc500_0.1CV1.05[,4], snc500_0.1CV1.1[,4], snc500_0.1CV1.15[,4])
          IDIbir500_0.1 <- c(snc500_0.1CV1[,6], snc500_0.1CV1.05[,6], snc500_0.1CV1.1[,6], snc500_0.1CV1.15[,6])
          ROCbir500_0.1 <- c(snc500_0.1CV1[,7], snc500_0.1CV1.05[,7],snc500_0.1CV1.1[,7],snc500_0.1CV1.15[,7])


          istat500_0.1 <- c(NRIcatbir500_0.1, NRIconbir500_0.1,IDIbir500_0.1, ROCbir500_0.1)
          CVbir500_0.1 <- c(rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),rep(1,1000), rep(1.05,1000),
          rep(1.1,1000), rep(1.15,1000), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),
          rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000))
          iss500_0.1 <- c(rep(1,4000), rep(2,4000), rep(3,4000),rep(4,4000))



birveri500_0.1 <- data.frame(istat500_0.1, CVbir500_0.1, iss500_0.1)

birveri500_0.1$iss500_0.1[birveri500_0.1$iss500_0.1==1]<-"NRIkategorik"
birveri500_0.1$iss500_0.1[birveri500_0.1$iss500_0.1==2]<-"NRIsürekli"
birveri500_0.1$iss500_0.1[birveri500_0.1$iss500_0.1==3]<-"IDI"
birveri500_0.1$iss500_0.1[birveri500_0.1$iss500_0.1==4]<-"ROC"

birveri2_500_0.1 <-remove_missing(birveri500_0.1, na.rm = TRUE, vars = names(birveri500_0.1), name = "",
  finite = FALSE)
```

```
library(Rmisc)
tgc500_0.1 <- summarySE(birveri2_500_0.1, measurevar="istat500_0.1",
groupvars=c("iss500_0.1","CVbir500_0.1"))
plot500_0.1 <- ggplot(tgc500_0.1, aes(x=CVbir500_0.1 , y=istat500_0.1 , colour=iss500_0.1 )) +
   geom_errorbar(aes(ymin=istat500_0.1 -se, ymax=istat500_0.1 +se), width=.01) +
   geom_line() +
   geom_point()+

          ylab("P Değerleri")+
          xlab("CV")

##################################

snc1000_0.1CV1 <- simu(1000,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2),1,1000)
sonn1000_0.1CV1 <- snclma(snc1000_0.1CV1)

snc1000_0.1CV1.05 <- simu(1000,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.1),1.05,1000)
sonn1000_0.1CV1.05 <- snclma(snc1000_0.1CV1.05)


snc1000_0.1CV1.1 <- simu(1000,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.2),1.1,1000)
sonn1000_0.1CV1.1 <- snclma(snc1000_0.1CV1.1)

snc1000_0.1CV1.15 <- simu(1000,.1, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.3),1.15,1000)
sonn1000_0.1CV1.15 <- snclma(snc1000_0.1CV1.15)

          NRIcatbir1000_0.1 <- c(snc1000_0.1CV1[,2], snc1000_0.1CV1.05[,2],snc1000_0.1CV1.1[,2],
snc1000_0.1CV1.15[,2])
          NRIconbir1000_0.1 <- c(snc1000_0.1CV1[,4], snc1000_0.1CV1.05[,4], snc1000_0.1CV1.1[,4],
snc1000_0.1CV1.15[,4])
          IDIbir1000_0.1 <- c(snc1000_0.1CV1[,6], snc1000_0.1CV1.05[,6], snc1000_0.1CV1.1[,6],
snc1000_0.1CV1.15[,6])
          ROCbir1000_0.1 <- c(snc1000_0.1CV1[,7],
snc1000_0.1CV1.05[,7],snc1000_0.1CV1.1[,7],snc1000_0.1CV1.15[,7])


          istat1000_0.1 <- c(NRIcatbir1000_0.1, NRIconbir1000_0.1,IDIbir1000_0.1, ROCbir1000_0.1)
          CVbir1000_0.1 <- c(rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),rep(1,1000),
rep(1.05,1000),
          rep(1.1,1000), rep(1.15,1000), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),
          rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000))
          iss1000_0.1 <- c(rep(1,4000), rep(2,4000), rep(3,4000),rep(4,4000))



birveri1000_0.1 <- data.frame(istat1000_0.1, CVbir1000_0.1, iss1000_0.1)

birveri1000_0.1$iss1000_0.1[birveri1000_0.1$iss1000_0.1==1]<-"NRIkategorik"
birveri1000_0.1$iss1000_0.1[birveri1000_0.1$iss1000_0.1==2]<-"NRIsürekli"
birveri1000_0.1$iss1000_0.1[birveri1000_0.1$iss1000_0.1==3]<-"IDI"
birveri1000_0.1$iss1000_0.1[birveri1000_0.1$iss1000_0.1==4]<-"ROC"

birveri2_1000_0.1 <-remove_missing(birveri1000_0.1, na.rm = TRUE, vars = names(birveri1000_0.1),
name = "",
```

```
  finite = FALSE)

library(Rmisc)
tgc1000_0.1 <- summarySE(birveri2_1000_0.1, measurevar="istat1000_0.1",
groupvars=c("iss1000_0.1","CVbir1000_0.1"))
plot1000_0.1 <- ggplot(tgc1000_0.1, aes(x=CVbir1000_0.1, y=istat1000_0.1, colour=iss1000_0.1)) +
    geom_errorbar(aes(ymin=istat1000_0.1-se, ymax=istat1000_0.1+se), width=.01) +
    geom_line() +
    geom_point()+

          ylab("P Değerleri")+
          xlab("CV")

#####p=0,25##########
snc100_0.25CV1 <- simu(100,0.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2),1,1000)
sonn100_0.25CV1 <- snclma(snc100_0.25CV1)

snc100_0.25CV1.05 <- simu(100,0.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.1),1.05,1000)
sonn100_0.25CV1.05 <- snclma(snc100_0.25CV1.05)


snc100_0.25CV1.1 <- simu(100,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.2),1.1,1000)
sonn100_0.25CV1.1 <- snclma(snc100_0.25CV1.1)

snc100_0.25CV1.15 <- simu(100,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.3),1.15,1000)
sonn100_0.25CV1.15 <- snclma(snc100_0.25CV1.15)

          NRIcatbir100_0.25 <- c(snc100_0.25CV1[,2], snc100_0.25CV1.05[,2],snc100_0.25CV1.1[,2],
snc100_0.25CV1.15[,2])
          NRIconbir100_0.25 <- c(snc100_0.25CV1[,4], snc100_0.25CV1.05[,4], snc100_0.25CV1.1[,4],
snc100_0.25CV1.15[,4])
          IDIbir100_0.25 <- c(snc100_0.25CV1[,6], snc100_0.25CV1.05[,6], snc100_0.25CV1.1[,6],
snc100_0.25CV1.15[,6])
          ROCbir100_0.25 <- c(snc100_0.25CV1[,7],
snc100_0.25CV1.05[,7],snc100_0.25CV1.1[,7],snc100_0.25CV1.15[,7])


          istat100_0.25 <- c(NRIcatbir100_0.25, NRIconbir100_0.25,IDIbir100_0.25, ROCbir100_0.25)
          CVbir100_0.25 <- c(rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),rep(1,1000),
rep(1.05,1000),
          rep(1.1,1000), rep(1.15,1000), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),
          rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000))
          iss100_0.25 <- c(rep(1,4000), rep(2,4000), rep(3,4000),rep(4,4000))



birveri100_0.25 <- data.frame(istat100_0.25, CVbir100_0.25, iss100_0.25)

birveri100_0.25$iss100_0.25[birveri100_0.25$iss100_0.25==1]<-"NRIkategorik"
birveri100_0.25$iss100_0.25[birveri100_0.25$iss100_0.25==2]<-"NRIsürekli"
birveri100_0.25$iss100_0.25[birveri100_0.25$iss100_0.25==3]<-"IDI"
birveri100_0.25$iss100_0.25[birveri100_0.25$iss100_0.25==4]<-"ROC"
```

```
birveri2_100_0.25 <-remove_missing(birveri100_0.25, na.rm = TRUE, vars = names(birveri100_0.25), name = "",
  finite = FALSE)

library(Rmisc)
tgc100_0.25 <- summarySE(birveri2_100_0.25, measurevar="istat100_0.25", groupvars=c("iss100_0.25","CVbir100_0.25"))
plot100_0.25 <- ggplot(tgc100_0.25, aes(x=CVbir100_0.25, y=istat100_0.25, colour=iss100_0.25)) +
    geom_errorbar(aes(ymin=istat100_0.25-se, ymax=istat100_0.25+se), width=.01) +
    geom_line() +
    geom_point()+

          ylab("P Değerleri")+
          xlab("CV")


##############################################

snc250_0.25CV1 <- simu(250,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2),1,1000)
sonn250_0.25CV1 <- snclma(snc250_0.25CV1)

snc250_0.25CV1.05 <- simu(250,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.1),1.05,1000)
sonn250_0.25CV1.05 <- snclma(snc250_0.25CV1.05)


snc250_0.25CV1.1 <- simu(250,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.2),1.1,1000)
sonn250_0.25CV1.1 <- snclma(snc250_0.25CV1.1)

snc250_0.25CV1.15 <- simu(250,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.3),1.15,1000)
sonn250_0.25CV1.15 <- snclma(snc250_0.25CV1.15)

          NRIcatbir250_0.25 <- c(snc250_0.25CV1[,2], snc250_0.25CV1.05[,2],snc250_0.25CV1.1[,2], snc250_0.25CV1.15[,2])
          NRIconbir250_0.25 <- c(snc250_0.25CV1[,4], snc250_0.25CV1.05[,4], snc250_0.25CV1.1[,4], snc250_0.25CV1.15[,4])
          IDIbir250_0.25 <- c(snc250_0.25CV1[,6], snc250_0.25CV1.05[,6], snc250_0.25CV1.1[,6], snc250_0.25CV1.15[,6])
          ROCbir250_0.25 <- c(snc250_0.25CV1[,7], snc250_0.25CV1.05[,7],snc250_0.25CV1.1[,7],snc250_0.25CV1.15[,7])


          istat250_0.25 <- c(NRIcatbir250_0.25, NRIconbir250_0.25,IDIbir250_0.25, ROCbir250_0.25)
          CVbir250_0.25 <- c(rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),rep(1,1000), rep(1.05,1000),rep(1.1,1000), rep(1.15,1000), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000))
          iss250_0.25 <- c(rep(1,4000), rep(2,4000), rep(3,4000),rep(4,4000))



birveri250_0.25 <- data.frame(istat250_0.25, CVbir250_0.25, iss250_0.25)

birveri250_0.25$iss250_0.25[birveri250_0.25$iss250_0.25==1]<-"NRIkategorik"
birveri250_0.25$iss250_0.25[birveri250_0.25$iss250_0.25==2]<-"NRIsürekli"
birveri250_0.25$iss250_0.25[birveri250_0.25$iss250_0.25==3]<-"IDI"
```

```
birveri250_0.25$iss250_0.25[birveri250_0.25$iss250_0.25==4]<-"ROC"

birveri2_250_0.25 <-remove_missing(birveri250_0.25, na.rm = TRUE, vars = names(birveri250_0.25),
name = "",
 finite = FALSE)

library(Rmisc)
tgc250_0.25 <- summarySE(birveri2_250_0.25, measurevar="istat250_0.25",
groupvars=c("iss250_0.25","CVbir250_0.25"))
plot250_0.25 <- ggplot(tgc250_0.25, aes(x=CVbir250_0.25, y=istat250_0.25, colour=iss250_0.25)) +
   geom_errorbar(aes(ymin=istat250_0.25-se, ymax=istat250_0.25+se), width=.01) +
   geom_line() +
   geom_point()+

          ylab("P Değerleri")+
          xlab("CV")
##############################################

snc500_0.25CV1 <- simu(500,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2),1,1000)
sonn500_0.25CV1 <- snclma(snc500_0.25CV1)

snc500_0.25CV1.05 <- simu(500,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.1),1.05,1000)
sonn500_0.25CV1.05 <- snclma(snc500_0.25CV1.05)


snc500_0.25CV1.1 <- simu(500,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.2),1.1,1000)
sonn500_0.25CV1.1 <- snclma(snc500_0.25CV1.1)

snc500_0.25CV1.15 <- simu(500,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.3),1.15,1000)
sonn500_0.25CV1.15 <- snclma(snc500_0.25CV1.15)

          NRIcatbir500_0.25 <- c(snc500_0.25CV1[,2], snc500_0.25CV1.05[,2],snc500_0.25CV1.1[,2],
snc500_0.25CV1.15[,2])
          NRIconbir500_0.25 <- c(snc500_0.25CV1[,4], snc500_0.25CV1.05[,4], snc500_0.25CV1.1[,4],
snc500_0.25CV1.15[,4])
          IDIbir500_0.25 <- c(snc500_0.25CV1[,6], snc500_0.25CV1.05[,6], snc500_0.25CV1.1[,6],
snc500_0.25CV1.15[,6])
          ROCbir500_0.25 <- c(snc500_0.25CV1[,7],
snc500_0.25CV1.05[,7],snc500_0.25CV1.1[,7],snc500_0.25CV1.15[,7])


          istat500_0.25 <- c(NRIcatbir500_0.25, NRIconbir500_0.25,IDIbir500_0.25, ROCbir500_0.25)
          CVbir500_0.25 <- c(rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),rep(1,1000),
rep(1.05,1000),
          rep(1.1,1000), rep(1.15,1000), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),
          rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000))
          iss500_0.25 <- c(rep(1,4000), rep(2,4000), rep(3,4000),rep(4,4000))



birveri500_0.25 <- data.frame(istat500_0.25, CVbir500_0.25, iss500_0.25)

birveri500_0.25$iss500_0.25[birveri500_0.25$iss500_0.25==1]<-"NRIkategorik"
birveri500_0.25$iss500_0.25[birveri500_0.25$iss500_0.25==2]<-"NRIsürekli"
```

```
birveri500_0.25$iss500_0.25[birveri500_0.25$iss500_0.25==3]<-"IDI"
birveri500_0.25$iss500_0.25[birveri500_0.25$iss500_0.25==4]<-"ROC"

birveri2_500_0.25 <-remove_missing(birveri500_0.25, na.rm = TRUE, vars = names(birveri500_0.25),
name = "",
 finite = FALSE)

library(Rmisc)
tgc500_0.25 <- summarySE(birveri2_500_0.25, measurevar="istat500_0.25",
groupvars=c("iss500_0.25","CVbir500_0.25"))
plot500_0.25 <- ggplot(tgc500_0.25, aes(x=CVbir500_0.25 , y=istat500_0.25 , colour=iss500_0.25 )) +
   geom_errorbar(aes(ymin=istat500_0.25 -se, ymax=istat500_0.25 +se), width=.01) +
   geom_line() +
   geom_point()+

        ylab("P Değerleri")+
        xlab("CV")

#################################

snc1000_0.25CV1 <- simu(1000,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2),1,1000)
sonn1000_0.25CV1 <- snclma(snc1000_0.25CV1)

snc1000_0.25CV1.05 <- simu(1000,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.1),1.05,1000)
sonn1000_0.25CV1.05 <- snclma(snc1000_0.25CV1.05)


snc1000_0.25CV1.1 <- simu(1000,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.2),1.1,1000)
sonn1000_0.25CV1.1 <- snclma(snc1000_0.25CV1.1)

snc1000_0.25CV1.15 <- simu(1000,.25, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.3),1.15,1000)
sonn1000_0.25CV1.15 <- snclma(snc1000_0.25CV1.15)

        NRIcatbir1000_0.25 <- c(snc1000_0.25CV1[,2],
snc1000_0.25CV1.05[,2],snc1000_0.25CV1.1[,2], snc1000_0.25CV1.15[,2])
        NRIconbir1000_0.25 <- c(snc1000_0.25CV1[,4], snc1000_0.25CV1.05[,4],
snc1000_0.25CV1.1[,4], snc1000_0.25CV1.15[,4])
        IDIbir1000_0.25 <- c(snc1000_0.25CV1[,6], snc1000_0.25CV1.05[,6], snc1000_0.25CV1.1[,6],
snc1000_0.25CV1.15[,6])
        ROCbir1000_0.25 <- c(snc1000_0.25CV1[,7],
snc1000_0.25CV1.05[,7],snc1000_0.25CV1.1[,7],snc1000_0.25CV1.15[,7])


        istat1000_0.25 <- c(NRIcatbir1000_0.25, NRIconbir1000_0.25,IDIbir1000_0.25,
ROCbir1000_0.25)
        CVbir1000_0.25 <- c(rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),rep(1,1000),
rep(1.05,1000),
        rep(1.1,1000), rep(1.15,1000), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),
        rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000))
        iss1000_0.25 <- c(rep(1,4000), rep(2,4000), rep(3,4000),rep(4,4000))



birveri1000_0.25 <- data.frame(istat1000_0.25, CVbir1000_0.25, iss1000_0.25)
```

```
birveri1000_0.25$iss1000_0.25[birveri1000_0.25$iss1000_0.25==1]<-"NRIkategorik"
birveri1000_0.25$iss1000_0.25[birveri1000_0.25$iss1000_0.25==2]<-"NRIsürekli"
birveri1000_0.25$iss1000_0.25[birveri1000_0.25$iss1000_0.25==3]<-"IDI"
birveri1000_0.25$iss1000_0.25[birveri1000_0.25$iss1000_0.25==4]<-"ROC"

birveri2_1000_0.25 <-remove_missing(birveri1000_0.25, na.rm = TRUE, vars = names(birveri1000_0.25), name = "",
 finite = FALSE)

library(Rmisc)
tgc1000_0.25 <- summarySE(birveri2_1000_0.25, measurevar="istat1000_0.25", groupvars=c("iss1000_0.25","CVbir1000_0.25"))
plot1000_0.25 <- ggplot(tgc1000_0.25, aes(x=CVbir1000_0.25, y=istat1000_0.25, colour=iss1000_0.25)) +
   geom_errorbar(aes(ymin=istat1000_0.25-se, ymax=istat1000_0.25+se), width=.01) +
   geom_line() +
   geom_point()+

          ylab("P Değerleri")+
          xlab("CV")


############p=0,50############

snc100_0.50CV1 <- simu(100,0.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2),1,1000)
sonn100_0.50CV1 <- snclma(snc100_0.50CV1)

snc100_0.50CV1.05 <- simu(100,0.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.1),1.05,1000)
sonn100_0.50CV1.05 <- snclma(snc100_0.50CV1.05)


snc100_0.50CV1.1 <- simu(100,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.2),1.1,1000)
sonn100_0.50CV1.1 <- snclma(snc100_0.50CV1.1)

snc100_0.50CV1.15 <- simu(100,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.3),1.15,1000)
sonn100_0.50CV1.15 <- snclma(snc100_0.50CV1.15)

          NRIcatbir100_0.50 <- c(snc100_0.50CV1[,2], snc100_0.50CV1.05[,2],snc100_0.50CV1.1[,2], snc100_0.50CV1.15[,2])
          NRIconbir100_0.50 <- c(snc100_0.50CV1[,4], snc100_0.50CV1.05[,4], snc100_0.50CV1.1[,4], snc100_0.50CV1.15[,4])
          IDIbir100_0.50 <- c(snc100_0.50CV1[,6], snc100_0.50CV1.05[,6], snc100_0.50CV1.1[,6], snc100_0.50CV1.15[,6])
          ROCbir100_0.50 <- c(snc100_0.50CV1[,7], snc100_0.50CV1.05[,7],snc100_0.50CV1.1[,7],snc100_0.50CV1.15[,7])


          istat100_0.50 <- c(NRIcatbir100_0.50, NRIconbir100_0.50,IDIbir100_0.50, ROCbir100_0.50)
          CVbir100_0.50 <- c(rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),rep(1,1000), rep(1.05,1000),
          rep(1.1,1000), rep(1.15,1000), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),
          rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000))
          iss100_0.50 <- c(rep(1,4000), rep(2,4000), rep(3,4000),rep(4,4000))
```

```
birveri100_0.50 <- data.frame(istat100_0.50, CVbir100_0.50, iss100_0.50)

birveri100_0.50$iss100_0.50[birveri100_0.50$iss100_0.50==1]<-"NRIkategorik"
birveri100_0.50$iss100_0.50[birveri100_0.50$iss100_0.50==2]<-"NRIsürekli"
birveri100_0.50$iss100_0.50[birveri100_0.50$iss100_0.50==3]<-"IDI"
birveri100_0.50$iss100_0.50[birveri100_0.50$iss100_0.50==4]<-"ROC"

birveri2_100_0.50 <-remove_missing(birveri100_0.50, na.rm = TRUE, vars = names(birveri100_0.50),
name = "",
  finite = FALSE)

library(Rmisc)
tgc100_0.50 <- summarySE(birveri2_100_0.50, measurevar="istat100_0.50",
groupvars=c("iss100_0.50","CVbir100_0.50"))
plot100_0.50 <- ggplot(tgc100_0.50, aes(x=CVbir100_0.50, y=istat100_0.50, colour=iss100_0.50)) +
    geom_errorbar(aes(ymin=istat100_0.50-se, ymax=istat100_0.50+se), width=.01) +
    geom_line() +
    geom_point()+

          ylab("P Değerleri")+
          xlab("CV")


##############################################

snc250_0.50CV1 <- simu(250,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2),1,1000)
sonn250_0.50CV1 <- snclma(snc250_0.50CV1)

snc250_0.50CV1.05 <- simu(250,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.1),1.05,1000)
sonn250_0.50CV1.05 <- snclma(snc250_0.50CV1.05)


snc250_0.50CV1.1 <- simu(250,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.2),1.1,1000)
sonn250_0.50CV1.1 <- snclma(snc250_0.50CV1.1)

snc250_0.50CV1.15 <- simu(250,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.3),1.15,1000)
sonn250_0.50CV1.15 <- snclma(snc250_0.50CV1.15)

          NRIcatbir250_0.50 <- c(snc250_0.50CV1[,2], snc250_0.50CV1.05[,2],snc250_0.50CV1.1[,2],
snc250_0.50CV1.15[,2])
          NRIconbir250_0.50 <- c(snc250_0.50CV1[,4], snc250_0.50CV1.05[,4], snc250_0.50CV1.1[,4],
snc250_0.50CV1.15[,4])
          IDIbir250_0.50 <- c(snc250_0.50CV1[,6], snc250_0.50CV1.05[,6], snc250_0.50CV1.1[,6],
snc250_0.50CV1.15[,6])
          ROCbir250_0.50 <- c(snc250_0.50CV1[,7],
snc250_0.50CV1.05[,7],snc250_0.50CV1.1[,7],snc250_0.50CV1.15[,7])


          istat250_0.50 <- c(NRIcatbir250_0.50, NRIconbir250_0.50,IDIbir250_0.50, ROCbir250_0.50)
          CVbir250_0.50 <- c(rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),rep(1,1000),
rep(1.05,1000),
```

```
        rep(1.1,1000), rep(1.15,1000), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),
        rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000))
        iss250_0.50 <- c(rep(1,4000), rep(2,4000), rep(3,4000),rep(4,4000))



birveri250_0.50 <- data.frame(istat250_0.50, CVbir250_0.50, iss250_0.50)

birveri250_0.50$iss250_0.50[birveri250_0.50$iss250_0.50==1]<-"NRIkategorik"
birveri250_0.50$iss250_0.50[birveri250_0.50$iss250_0.50==2]<-"NRIsürekli"
birveri250_0.50$iss250_0.50[birveri250_0.50$iss250_0.50==3]<-"IDI"
birveri250_0.50$iss250_0.50[birveri250_0.50$iss250_0.50==4]<-"ROC"

birveri2_250_0.50 <-remove_missing(birveri250_0.50, na.rm = TRUE, vars = names(birveri250_0.50),
name = "",
  finite = FALSE)

library(Rmisc)
tgc250_0.50 <- summarySE(birveri2_250_0.50, measurevar="istat250_0.50",
groupvars=c("iss250_0.50","CVbir250_0.50"))
plot250_0.50 <- ggplot(tgc250_0.50, aes(x=CVbir250_0.50, y=istat250_0.50, colour=iss250_0.50)) +
    geom_errorbar(aes(ymin=istat250_0.50-se, ymax=istat250_0.50+se), width=.01) +
    geom_line() +
    geom_point()+

        ylab("P Değerleri")+
        xlab("CV")
##############################################

snc500_0.50CV1 <- simu(500,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2),1,1000)
sonn500_0.50CV1 <- snclma(snc500_0.50CV1)

snc500_0.50CV1.05 <- simu(500,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.1),1.05,1000)
sonn500_0.50CV1.05 <- snclma(snc500_0.50CV1.05)


snc500_0.50CV1.1 <- simu(500,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.2),1.1,1000)
sonn500_0.50CV1.1 <- snclma(snc500_0.50CV1.1)

snc500_0.50CV1.15 <- simu(500,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.3),1.15,1000)
sonn500_0.50CV1.15 <- snclma(snc500_0.50CV1.15)

        NRIcatbir500_0.50 <- c(snc500_0.50CV1[,2], snc500_0.50CV1.05[,2],snc500_0.50CV1.1[,2],
snc500_0.50CV1.15[,2])
        NRIconbir500_0.50 <- c(snc500_0.50CV1[,4], snc500_0.50CV1.05[,4], snc500_0.50CV1.1[,4],
snc500_0.50CV1.15[,4])
        IDIbir500_0.50 <- c(snc500_0.50CV1[,6], snc500_0.50CV1.05[,6], snc500_0.50CV1.1[,6],
snc500_0.50CV1.15[,6])
        ROCbir500_0.50 <- c(snc500_0.50CV1[,7],
snc500_0.50CV1.05[,7],snc500_0.50CV1.1[,7],snc500_0.50CV1.15[,7])


        istat500_0.50 <- c(NRIcatbir500_0.50, NRIconbir500_0.50,IDIbir500_0.50, ROCbir500_0.50)
```

```
        CVbir500_0.50 <- c(rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),rep(1,1000),
rep(1.05,1000),
        rep(1.1,1000), rep(1.15,1000), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000),
        rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,1000))
        iss500_0.50 <- c(rep(1,4000), rep(2,4000), rep(3,4000),rep(4,4000))



birveri500_0.50 <- data.frame(istat500_0.50, CVbir500_0.50, iss500_0.50)

birveri500_0.50$iss500_0.50[birveri500_0.50$iss500_0.50==1]<-"NRIkategorik"
birveri500_0.50$iss500_0.50[birveri500_0.50$iss500_0.50==2]<-"NRIsürekli"
birveri500_0.50$iss500_0.50[birveri500_0.50$iss500_0.50==3]<-"IDI"
birveri500_0.50$iss500_0.50[birveri500_0.50$iss500_0.50==4]<-"ROC"

birveri2_500_0.50 <-remove_missing(birveri500_0.50, na.rm = TRUE, vars = names(birveri500_0.50),
name = "",
 finite = FALSE)

library(Rmisc)
tgc500_0.50 <- summarySE(birveri2_500_0.50, measurevar="istat500_0.50",
groupvars=c("iss500_0.50","CVbir500_0.50"))
plot500_0.50 <- ggplot(tgc500_0.50, aes(x=CVbir500_0.50 , y=istat500_0.50 , colour=iss500_0.50 )) +
   geom_errorbar(aes(ymin=istat500_0.50 -se, ymax=istat500_0.50 +se), width=.01) +
   geom_line() +
   geom_point()+

        ylab("P Değerleri")+
        xlab("CV")

#################################

snc1000_0.50CV1 <- simu(1000,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2),1,1000)
sonn1000_0.50CV1 <- snclma(snc1000_0.50CV1)

snc1000_0.50CV1.05 <- simu(1000,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.1),1.05,1000)
sonn1000_0.50CV1.05 <- snclma(snc1000_0.50CV1.05)


snc1000_0.50CV1.1 <- simu(1000,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.2),1.1,1000)
sonn1000_0.50CV1.1 <- snclma(snc1000_0.50CV1.1)

snc1000_0.50CV1.15 <- simu(1000,.50, 1, .2, 0.25, 1, sig, c(0,1,2),c(1,2,2.3),1.15,100)
sonn1000_0.50CV1.15 <- snclma(snc1000_0.50CV1.15)

        NRIcatbir1000_0.50 <- c(snc1000_0.50CV1[,2],
snc1000_0.50CV1.05[,2],snc1000_0.50CV1.1[,2], snc1000_0.50CV1.15[,2])
        NRIconbir1000_0.50 <- c(snc1000_0.50CV1[,4], snc1000_0.50CV1.05[,4],
snc1000_0.50CV1.1[,4], snc1000_0.50CV1.15[,4])
        IDIbir1000_0.50 <- c(snc1000_0.50CV1[,6], snc1000_0.50CV1.05[,6], snc1000_0.50CV1.1[,6],
snc1000_0.50CV1.15[,6])
        ROCbir1000_0.50 <- c(snc1000_0.50CV1[,7],
snc1000_0.50CV1.05[,7],snc1000_0.50CV1.1[,7],snc1000_0.50CV1.15[,7])
```

```
        istat1000_0.50 <- c(NRIcatbir1000_0.50, NRIconbir1000_0.50,IDIbir1000_0.50,
ROCbir1000_0.50)
        CVbir1000_0.50 <- c(rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,100),rep(1,1000),
rep(1.05,1000),
        rep(1.1,1000), rep(1.15,100), rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,100),
        rep(1,1000), rep(1.05,1000),rep(1.1,1000),rep(1.15,100))
        iss1000_0.50 <- c(rep(1,3100), rep(2,3100), rep(3,3100),rep(4,3100))



birveri1000_0.50 <- data.frame(istat1000_0.50, CVbir1000_0.50, iss1000_0.50)

birveri1000_0.50$iss1000_0.50[birveri1000_0.50$iss1000_0.50==1]<-"NRIkategorik"
birveri1000_0.50$iss1000_0.50[birveri1000_0.50$iss1000_0.50==2]<-"NRIsürekli"
birveri1000_0.50$iss1000_0.50[birveri1000_0.50$iss1000_0.50==3]<-"IDI"
birveri1000_0.50$iss1000_0.50[birveri1000_0.50$iss1000_0.50==4]<-"ROC"

birveri2_1000_0.50 <-remove_missing(birveri1000_0.50, na.rm = TRUE, vars =
names(birveri1000_0.50), name = "",
 finite = FALSE)

library(Rmisc)
tgc1000_0.50 <- summarySE(birveri2_1000_0.50, measurevar="istat1000_0.50",
groupvars=c("iss1000_0.50","CVbir1000_0.50"))
plot1000_0.50 <- ggplot(tgc1000_0.50, aes(x=CVbir1000_0.50, y=istat1000_0.50,
colour=iss1000_0.50)) +
  geom_errorbar(aes(ymin=istat1000_0.50 - se*0.07863, ymax=istat1000_0.50 + se*0.07863),
width=.01) +
  geom_line() +
  geom_point()+

        ylab("P Değerleri")+
        xlab("CV")

##############Tanımlamalar#############
###########p=0,10#######################

###******ROC*******######

summary(snc100_0.1CV1[,8])
sd(snc100_0.1CV1[,8])

summary(snc100_0.1CV1[,9])
sd(snc100_0.1CV1[,9])

summary(snc100_0.1CV1[,9]-snc100_0.1CV1[,8])
sd(snc100_0.1CV1[,9]-snc100_0.1CV1[,8])

summary(snc250_0.1CV1[,8])
sd(snc250_0.1CV1[,8])

summary(snc250_0.1CV1[,9])
sd(snc250_0.1CV1[,9])
```

```
summary(snc250_0.1CV1[,9]-snc250_0.1CV1[,8])
sd(snc250_0.1CV1[,9]-snc250_0.1CV1[,8])

summary(snc500_0.1CV1[,8])
sd(snc500_0.1CV1[,8])

summary(snc500_0.1CV1[,9])
sd(snc500_0.1CV1[,9])

summary(snc500_0.1CV1[,9]-snc500_0.1CV1[,8])
sd(snc500_0.1CV1[,9]-snc500_0.1CV1[,8])

summary(snc1000_0.1CV1[,8])
sd(snc1000_0.1CV1[,8])

summary(snc1000_0.1CV1[,9])
sd(snc1000_0.1CV1[,9])

summary(snc1000_0.1CV1[,9]-snc1000_0.1CV1[,8])
sd(snc1000_0.1CV1[,9]-snc1000_0.1CV1[,8])


###########NRI Kategorik*******######

summary(snc100_0.1CV1[,1], na.rm=TRUE)
sd(snc100_0.1CV1[,1], na.rm=TRUE)


summary(snc250_0.1CV1[,1], na.rm=TRUE)
sd(snc250_0.1CV1[,1], na.rm=TRUE)


summary(snc500_0.1CV1[,1], na.rm=TRUE)
sd(snc500_0.1CV1[,1], na.rm=TRUE)


summary(snc1000_0.1CV1[,1], na.rm=TRUE)
sd(snc1000_0.1CV1[,1], na.rm=TRUE)


####******NRI Sürekli*******######

summary(snc100_0.1CV1[,3], na.rm=TRUE)
sd(snc100_0.1CV1[,3], na.rm=TRUE)


summary(snc250_0.1CV1[,3], na.rm=TRUE)
sd(snc250_0.1CV1[,3], na.rm=TRUE)


summary(snc500_0.1CV1[,3], na.rm=TRUE)
sd(snc500_0.1CV1[,3], na.rm=TRUE)
```

```
summary(snc1000_0.1CV1[,3], na.rm=TRUE)
sd(snc1000_0.1CV1[,3], na.rm=TRUE)



###******IDI******######

summary(snc100_0.1CV1[,5], na.rm=TRUE)
sd(snc100_0.1CV1[,5], na.rm=TRUE)


summary(snc250_0.1CV1[,5], na.rm=TRUE)
sd(snc250_0.1CV1[,5], na.rm=TRUE)


summary(snc500_0.1CV1[,5], na.rm=TRUE)
sd(snc500_0.1CV1[,5], na.rm=TRUE)


summary(snc1000_0.1CV1[,5], na.rm=TRUE)
sd(snc1000_0.1CV1[,5], na.rm=TRUE)


####**********CV=1,05**********########

###******ROC******######

summary(snc100_0.1CV1.05[,8])
sd(snc100_0.1CV1.05[,8])

summary(snc100_0.1CV1.05[,9])
sd(snc100_0.1CV1.05[,9])

summary(snc100_0.1CV1.05[,9]-snc100_0.1CV1.05[,8])
sd(snc100_0.1CV1.05[,9]-snc100_0.1CV1.05[,8])

summary(snc250_0.1CV1.05[,8])
sd(snc250_0.1CV1.05[,8])

summary(snc250_0.1CV1.05[,9])
sd(snc250_0.1CV1.05[,9])

summary(snc250_0.1CV1.05[,9]-snc250_0.1CV1.05[,8])
sd(snc250_0.1CV1.05[,9]-snc250_0.1CV1.05[,8])

summary(snc500_0.1CV1.05[,8])
sd(snc500_0.1CV1.05[,8])

summary(snc500_0.1CV1.05[,9])
sd(snc500_0.1CV1.05[,9])

summary(snc500_0.1CV1.05[,9]-snc500_0.1CV1.05[,8])
sd(snc500_0.1CV1.05[,9]-snc500_0.1CV1.05[,8])
```

```
summary(snc1000_0.1CV1.05[,8])
sd(snc1000_0.1CV1.05[,8])

summary(snc1000_0.1CV1.05[,9])
sd(snc1000_0.1CV1.05[,9])

summary(snc1000_0.1CV1.05[,9]-snc1000_0.1CV1.05[,8])
sd(snc1000_0.1CV1.05[,9]-snc1000_0.1CV1.05[,8])


####******NRI Kategorik*******######

summary(snc100_0.1CV1.05[,1], na.rm=TRUE)
sd(snc100_0.1CV1.05[,1], na.rm=TRUE)


summary(snc250_0.1CV1.05[,1], na.rm=TRUE)
sd(snc250_0.1CV1.05[,1], na.rm=TRUE)


summary(snc500_0.1CV1.05[,1], na.rm=TRUE)
sd(snc500_0.1CV1.05[,1], na.rm=TRUE)


summary(snc1000_0.1CV1.05[,1], na.rm=TRUE)
sd(snc1000_0.1CV1.05[,1], na.rm=TRUE)

###########NRI Sürekli*******######

summary(snc100_0.1CV1.05[,3], na.rm=TRUE)
sd(snc100_0.1CV1.05[,3], na.rm=TRUE)


summary(snc250_0.1CV1.05[,3], na.rm=TRUE)
sd(snc250_0.1CV1.05[,3], na.rm=TRUE)


summary(snc500_0.1CV1.05[,3], na.rm=TRUE)
sd(snc500_0.1CV1.05[,3], na.rm=TRUE)


summary(snc1000_0.1CV1.05[,3], na.rm=TRUE)
sd(snc1000_0.1CV1.05[,3], na.rm=TRUE)



###########IDI*******######

summary(snc100_0.1CV1.05[,5], na.rm=TRUE)
sd(snc100_0.1CV1.05[,5], na.rm=TRUE)
```

```
summary(snc250_0.1CV1.05[,5], na.rm=TRUE)
sd(snc250_0.1CV1.05[,5], na.rm=TRUE)


summary(snc500_0.1CV1.05[,5], na.rm=TRUE)
sd(snc500_0.1CV1.05[,5], na.rm=TRUE)


summary(snc1000_0.1CV1.05[,5], na.rm=TRUE)
sd(snc1000_0.1CV1.05[,5], na.rm=TRUE)



####**********CV=1,10**********########

###******ROC******######

summary(snc100_0.1CV1.1[,8])
sd(snc100_0.1CV1.1[,8])

summary(snc100_0.1CV1.1[,9])
sd(snc100_0.1CV1.1[,9])

summary(snc100_0.1CV1.1[,9]-snc100_0.1CV1.1[,8])
sd(snc100_0.1CV1.1[,9]-snc100_0.1CV1.1[,8])

summary(snc250_0.1CV1.1[,8])
sd(snc250_0.1CV1.1[,8])

summary(snc250_0.1CV1.1[,9])
sd(snc250_0.1CV1.1[,9])

summary(snc250_0.1CV1.1[,9]-snc250_0.1CV1.1[,8])
sd(snc250_0.1CV1.1[,9]-snc250_0.1CV1.1[,8])

summary(snc500_0.1CV1.1[,8])
sd(snc500_0.1CV1.1[,8])

summary(snc500_0.1CV1.1[,9])
sd(snc500_0.1CV1.1[,9])

summary(snc500_0.1CV1.1[,9]-snc500_0.1CV1.1[,8])
sd(snc500_0.1CV1.1[,9]-snc500_0.1CV1.1[,8])

summary(snc1000_0.1CV1.1[,8])
sd(snc1000_0.1CV1.1[,8])

summary(snc1000_0.1CV1.1[,9])
sd(snc1000_0.1CV1.1[,9])

summary(snc1000_0.1CV1.1[,9]-snc1000_0.1CV1.1[,8])
sd(snc1000_0.1CV1.1[,9]-snc1000_0.1CV1.1[,8])
```

```
####******NRI Kategorik*******######

summary(snc100_0.1CV1.1[,1], na.rm=TRUE)
sd(snc100_0.1CV1.1[,1], na.rm=TRUE)


summary(snc250_0.1CV1.1[,1], na.rm=TRUE)
sd(snc250_0.1CV1.1[,1], na.rm=TRUE)


summary(snc500_0.1CV1.1[,1], na.rm=TRUE)
sd(snc500_0.1CV1.1[,1], na.rm=TRUE)


summary(snc1000_0.1CV1.1[,1], na.rm=TRUE)
sd(snc1000_0.1CV1.1[,1], na.rm=TRUE)

###********NRI Sürekli*******######

summary(snc100_0.1CV1.1[,3], na.rm=TRUE)
sd(snc100_0.1CV1.1[,3], na.rm=TRUE)


summary(snc250_0.1CV1.1[,3], na.rm=TRUE)
sd(snc250_0.1CV1.1[,3], na.rm=TRUE)


summary(snc500_0.1CV1.1[,3], na.rm=TRUE)
sd(snc500_0.1CV1.1[,3], na.rm=TRUE)


summary(snc1000_0.1CV1.1[,3], na.rm=TRUE)
sd(snc1000_0.1CV1.1[,3], na.rm=TRUE)



####******IDI*******######

summary(snc100_0.1CV1.1[,5], na.rm=TRUE)
sd(snc100_0.1CV1.1[,5], na.rm=TRUE)


summary(snc250_0.1CV1.1[,5], na.rm=TRUE)
sd(snc250_0.1CV1.1[,5], na.rm=TRUE)


summary(snc500_0.1CV1.1[,5], na.rm=TRUE)
sd(snc500_0.1CV1.1[,5], na.rm=TRUE)


summary(snc1000_0.1CV1.1[,5], na.rm=TRUE)
sd(snc1000_0.1CV1.1[,5], na.rm=TRUE)
```

```
####**********CV=1,15**********########

####******ROC*******######

summary(snc100_0.1CV1.15[,8])
sd(snc100_0.1CV1.15[,8])

summary(snc100_0.1CV1.15[,9])
sd(snc100_0.1CV1.15[,9])

summary(snc100_0.1CV1.15[,9]-snc100_0.1CV1.15[,8])
sd(snc100_0.1CV1.15[,9]-snc100_0.1CV1.15[,8])

summary(snc250_0.1CV1.15[,8])
sd(snc250_0.1CV1.15[,8])

summary(snc250_0.1CV1.15[,9])
sd(snc250_0.1CV1.15[,9])

summary(snc250_0.1CV1.15[,9]-snc250_0.1CV1.15[,8])
sd(snc250_0.1CV1.15[,9]-snc250_0.1CV1.15[,8])

summary(snc500_0.1CV1.15[,8])
sd(snc500_0.1CV1.15[,8])

summary(snc500_0.1CV1.15[,9])
sd(snc500_0.1CV1.15[,9])

summary(snc500_0.1CV1.15[,9]-snc500_0.1CV1.15[,8])
sd(snc500_0.1CV1.15[,9]-snc500_0.1CV1.15[,8])

summary(snc1000_0.1CV1.15[,8])
sd(snc1000_0.1CV1.15[,8])

summary(snc1000_0.1CV1.15[,9])
sd(snc1000_0.1CV1.15[,9])

summary(snc1000_0.1CV1.15[,9]-snc1000_0.1CV1.15[,8])
sd(snc1000_0.1CV1.15[,9]-snc1000_0.1CV1.15[,8])


####******NRI Kategorik*******######

summary(snc100_0.1CV1.15[,1], na.rm=TRUE)
sd(snc100_0.1CV1.15[,1], na.rm=TRUE)


summary(snc250_0.1CV1.15[,1], na.rm=TRUE)
sd(snc250_0.1CV1.15[,1], na.rm=TRUE)
```

```
summary(snc500_0.1CV1.15[,1], na.rm=TRUE)
sd(snc500_0.1CV1.15[,1], na.rm=TRUE)


summary(snc1000_0.1CV1.15[,1], na.rm=TRUE)
sd(snc1000_0.1CV1.15[,1], na.rm=TRUE)


####******NRI Sürekli*******######

summary(snc100_0.1CV1.15[,3], na.rm=TRUE)
sd(snc100_0.1CV1.15[,3], na.rm=TRUE)


summary(snc250_0.1CV1.15[,3], na.rm=TRUE)
sd(snc250_0.1CV1.15[,3], na.rm=TRUE)


summary(snc500_0.1CV1.15[,3], na.rm=TRUE)
sd(snc500_0.1CV1.15[,3], na.rm=TRUE)


summary(snc1000_0.1CV1.15[,3], na.rm=TRUE)
sd(snc1000_0.1CV1.15[,3], na.rm=TRUE)



###*******IDI*******######

summary(snc100_0.1CV1.15[,5], na.rm=TRUE)
sd(snc100_0.1CV1.15[,5], na.rm=TRUE)


summary(snc250_0.1CV1.15[,5], na.rm=TRUE)
sd(snc250_0.1CV1.15[,5], na.rm=TRUE)


summary(snc500_0.1CV1.15[,5], na.rm=TRUE)
sd(snc500_0.1CV1.15[,5], na.rm=TRUE)


summary(snc1000_0.1CV1.15[,5], na.rm=TRUE)
sd(snc1000_0.1CV1.15[,5], na.rm=TRUE)

###############p=0,25 ve p=0,50 için tanımlamalar ayrıca yapılmıştır.##############
```

Gerçek Veri Seti İçin Kodlar

```
library(memisc)
library(pROC)
library(PredictABEL)
```

```
karsila <-as.data.set(spss.system.file("C:/Users/Betul/Desktop/ricin.sav"))

karsila <- data.frame(karsila)

karsila$ex[karsila$sonuc2=="Hayatta"]<-0
karsila$ex[karsila$sonuc2=="Olu"]<-1
karsila$ex<- as.numeric(as.character(karsila$ex))

karsila <- cbind(karsila,karsila$ex)
karsila <- data.frame(karsila)

        cikti <- 8
        cNonNew1 <- c(9:14)
        cNonNew2 <- c(9:14)

        cNonNew1Kate <- c(9:14)
        cNonNew2Kate <- c(9:14)

        cNew1 <- c(0)
        cNew2 <- c(0)

        cNew1Kate <- c(0)
        cNew2Kate <- c(0)

                riskmodelSNAP <- fitLogRegModel(data=karsila, cOutcome=cikti,
                                cNonGenPreds=cNonNew1, cNonGenPredsCat=cNonNew1Kate,
                                cGenPreds=cNew1, cGenPredsCat=cNew1Kate)

        ORmultivariate(riskModel=riskmodelSNAP, filename="multiOR.txt")

        tahmin1 <- predRisk(riskmodelSNAP)


        karsila <- cbind(karsila,tahmin1)
        karsila <- data.frame(karsila)

        SNAPPE <- lm(tahmin1 ~ sga + d.agirkate + apgarkate , data=karsila)


####SNAPPE-II risk skorunun olasılık #########

        olas <- glm(sonuc2~karsila$snappe, data=karsila,family=binomial())
                olas1 <- predRisk(olas)
                ORmultivariate(riskModel=olas, filename="multiOR.txt")

#######################

        roceski <- plot.roc(karsila$sonuc2, olas1,
                main="", xlab= "Özgüllük", ylab="Duyarlılık", percent=FALSE,
                ci=TRUE, # compute AUC (of AUC by default)
                print.auc=TRUE) # print the AUC (will contain the CI)
                ciobj <- ci.se(roceski, # CI of sensitivity
                specificities=seq(0, 1, .01)) # over a select set of specificities
```

```
        plot(ciobj, type="shape", col="#1c61b6AA") # plot as a blue shape
        plot(ci(roceski, of="thresholds", thresholds="best")) # add one threshold


    roc1 <- roc(karsila$sonuc2, karsila$snappe)
    coords(roc1, "best", ret=c("threshold","specificity","sensitivity","accuracy"))

plotCalibration(karsila, 32, olas1 ,plottitle="",rangeaxis=c(0,1),
xlab="Tahmin Edilen Risk Skoru",ylab="Gözlemlenen Risk Skoru")



cikti <- 8

    cNonNew1 <- c(9:14,19)
    cNonNew2 <- c(9:14,19)

    cNonNew1Kate <- c(9:14,19)
    cNonNew2Kate <- c(9:14,19)

    cNew1 <- c(0)
    cNew2 <- c(19)

    cNew1Kate <- c(0)
    cNew2Kate <- c(19)

    riskmodelsur <- fitLogRegModel(data=karsila, cOutcome=cikti,
        cNonGenPreds=cNonNew1, cNonGenPredsCat=cNonNew1Kate,
        cGenPreds=cNew2, cGenPredsCat=cNew2Kate)

tahmin2 <- predRisk(riskmodelsur)


    ORmultivariate(riskModel=riskmodelsur, filename="multiOR.txt")

    karsila <- cbind(karsila,tahmin2)
    karsila <- data.frame(karsila)

    SNAPPE2 <- lm(tahmin2 ~ sga + d.agirkate + apgarkate , data=karsila)


####SNAPPE-II risk skorunun olasılık #########

    olas2 <- glm(sonuc2~karsila$snappesur, data=karsila,family=binomial())
        olas3 <- predRisk(olas2)

    ORmultivariate(riskModel=olas2, filename="multiOR.txt")


    rocyeni <- plot.roc(karsila$sonuc2, olas3,
        main="", xlab= "Özgüllük", ylab="Duyarlılık", percent=FALSE,
        ci=TRUE, # compute AUC (of AUC by default)
        print.auc=TRUE) # print the AUC (will contain the CI)
```

```
                ciobj <- ci.se(rocyeni, # CI of sensitivity
                specificities=seq(0, 1, .01)) # over a select set of specificities
                plot(ciobj, type="shape", col="#1c61b6AA") # plot as a blue shape
                plot(ci(rocyeni, of="thresholds", thresholds="best")) # add one threshold


        roc2 <- roc(karsila$sonuc2, karsila$snappesur)
        coords(roc2, "best", ret=c("threshold","specificity","sensitivity","accuracy"))


        plotCalibration(karsila, 32, olas3,plottitle="",rangeaxis=c(0,1),
                xlab="Tahmin Edilen Risk Skoru",ylab="Gözlemlenen Risk Skoru")



#####################################################################


cikti <- 8

        cNonNew1 <- c(9:14,18)
        cNonNew2 <- c(9:14,18)

        cNonNew1Kate <- c(9:14,18)
        cNonNew2Kate <- c(9:14,18)

        cNew1 <- c(0)
        cNew2 <- c(18)

        cNew1Kate <- c(0)
        cNew2Kate <- c(18)

        riskmodelsel <- fitLogRegModel(data=karsila, cOutcome=cikti,
                cNonGenPreds=cNonNew1, cNonGenPredsCat=cNonNew1Kate,
                cGenPreds=cNew2, cGenPredsCat=cNew2Kate)

tahmin3 <- predRisk(riskmodelsel)


        ORmultivariate(riskModel=riskmodelsel, filename="multiOR.txt")


        karsila <- cbind(karsila,tahmin3)
        karsila <- data.frame(karsila)

        SNAPPE3 <- lm(tahmin3 ~ sga + d.agirkate + apgarkate , data=karsila)

##########

        olas4 <- glm(sonuc2~karsila$snappesel, data=karsila,family=binomial())
                olas5 <- predRisk(olas4)
```

```
        ORmultivariate(riskModel=olas4, filename="multiOR.txt")

        rocyeni <- plot.roc(karsila$sonuc2, olas5,
                main="", xlab= "Özgüllük", ylab="Duyarlılık", percent=FALSE,
                ci=TRUE, # compute AUC (of AUC by default)
                print.auc=TRUE) # print the AUC (will contain the CI)
                ciobj <- ci.se(rocyeni, # CI of sensitivity
                specificities=seq(0, 1, .01)) # over a select set of specificities
                plot(ciobj, type="shape", col="#1c61b6AA") # plot as a blue shape
                plot(ci(rocyeni, of="thresholds", thresholds="best")) # add one threshold

        roc3 <- roc(karsila$sonuc2, karsila$snappesel)
        coords(roc3, "best", ret=c("threshold","specificity","sensitivity","accuracy"))



        plotCalibration(karsila, 36, olas5 ,plottitle="",rangeaxis=c(0,1),
                xlab="Tahmin Edilen Risk Skoru",ylab="Gözlemlenen Risk Skoru")




cikti <- 8

        cNonNew1 <- c(9:14,18,19)
        cNonNew2 <- c(9:14,18,19)

        cNonNew1Kate <- c(9:14,18,19)
        cNonNew2Kate <- c(9:14,18,19)

        cNew1 <- c(0)
        cNew2 <- c(18,19)

        cNew1Kate <- c(0)
        cNew2Kate <- c(18,19)

        riskmodeliki <- fitLogRegModel(data=karsila, cOutcome=cikti,
                cNonGenPreds=cNonNew1, cNonGenPredsCat=cNonNew1Kate,
                cGenPreds=cNew2, cGenPredsCat=cNew2Kate)


        ORmultivariate(riskModel=riskmodeliki, filename="multiOR.txt")

        tahmin4 <- predRisk(riskmodeliki)

        karsila <- cbind(karsila,tahmin4)
        karsila <- data.frame(karsila)

        SNAPPE4 <- lm(tahmin4 ~ sga + d.agirkate + apgarkate , data=karsila)

##########
```

```
olas6 <- glm(sonuc2~karsila$snappeiki, data=karsila,family=binomial())
olas7 <- predRisk(olas6)

ORmultivariate(riskModel=olas6, filename="multiOR.txt")

rocyeni <- plot.roc(karsila$sonuc2, olas7,
        main="", xlab= "Özgüllük", ylab="Duyarlılık", percent=FALSE,
        ci=TRUE, # compute AUC (of AUC by default)
        print.auc=TRUE) # print the AUC (will contain the CI)
        ciobj <- ci.se(rocyeni, # CI of sensitivity
        specificities=seq(0, 1, .01)) # over a select set of specificities
        plot(ciobj, type="shape", col="#1c61b6AA") # plot as a blue shape
        plot(ci(rocyeni, of="thresholds", thresholds="best")) # add one threshold


roc4 <- roc(karsila$sonuc2, karsila$snappeiki)
coords(roc4, "best", ret=c("threshold","specificity","sensitivity","accuracy"))


plotCalibration(karsila, 40, olas7,plottitle="",rangeaxis=c(0,1),
        xlab="Tahmin Edilen Risk Skoru",ylab="Gözlemlenen Risk Skoru")

##############################

plotDiscriminationBox(data=karsila, 40, olas1)



roc1 <- roc(karsila$sonuc2, tahmin4)
coords(roc1, "best", ret=c("threshold","specificity","sensitivity","accuracy"))


roc2 <- roc(karsila$sonuc2, olas7)
coords(roc2, "best", ret=c("threshold","specificity","sensitivity","accuracy"))


roc.test(roc2, roc1, method="bootstrap", boot.n=10000)

cutoff <- c(0,.08742944,1)
reclassification(data=karsila, cOutcome=45,
predrisk1=tahmin3, predrisk2=olas5, cutoff)

rangeyaxis <- c(0,1)

# specify labels of the predictiveness curves

labels <- c("Surfaktan Olmayan Model", "Surfaktan Olan Model")
# produce predictiveness curves

plotPredictivenessCurve(predrisk=cbind(tahmin1,tahmin2),
rangeyaxis=rangeyaxis, labels=labels)
```

```
xlabel <- "SNAPPE-II Risk"

# specify label of y-axis
ylabel <- "SNAPPE-II + Surfaktan Risk"

# specify title for the plot
titleplot <- " "

# specify range of the x-axis and y-axis
rangeaxis <- c(0,1)

# labels given to the groups without and with the outcome of interest
labels<- c("Hayatta", "Ölü")

# produce prior risks and posterior risks plot
plotPriorPosteriorRisk(data=karsila, priorrisk=olas1,
posteriorrisk=olas3, cOutcome=40, xlabel=xlabel, ylabel=ylabel,
rangeaxis=rangeaxis, plotAll=TRUE, plottitle=titleplot, labels=labels)


# specify the size of each interval
interval <- .05

# specify label of x-axis
xlabel <- "Tahmin Edilen Risk"

# specify label of y-axis
ylabel <- "Yüzde"

# specify range of x-axis
xrange <- c(0,1)

# specify range of y-axis
yrange <- c(0,80)

# specify title for the plot
maintitle <- ""

# specify labels
labels <- c("Hayatta", "Ölü")


# produce risk distribution plot
plotRiskDistribution(data=karsila, cOutcome=40,
risks=olas7, interval=interval, plottitle=maintitle, rangexaxis=xrange,
rangeyaxis=yrange, xlabel=xlabel, ylabel=ylabel, labels=labels)



labels <- c("SNAPPE-II + sürfaktan + seleston", "SNAP-II + sürfaktan + seleston")
# produce ROC curve
plotROC(data=karsila, cOutcome=44,
predrisk=cbind(olas7,tahmin4), labels=labels,
```

```
xlab="1-Seçicilik", ylab="Duyarlılık",plottitle="")


#########kliniksel Yarar#########

karsila <- cbind(karsila,olas1,olas3,olas5,olas7)
karsila <- data.frame(karsila)


dca(data=karsila, outcome="ex", predictors="tahmin4", xstop=1)

dca(data=karsila, outcome="ex", predictors=c("olas1", "olas3","olas5", "olas7"), xstop=0.99)


karsila <- cbind(karsila,olas1)
karsila <- data.frame(karsila)


dca(data=karsila, outcome="ex", predictors=c("tahmin1", "tahmin2","tahmin3", "tahmin4"), xstop=0.99)
```

SNAP-II ve SNAPPE-II Skorlama Formu

SNAP-II					
HASTANIN ADI		**DOĞUM TARİHİ**		**PROTOKOL NO**	
Ort. Kan Basıncı (mmgh)	≥30			0	
	29-20			9	
	<20			19	
En Düşük Isı (C)	>35,6			0	
	35,6-35			8	
	<35			15	
PO_2/FiO_2	>2,49			0	
	1,00-2,49			5	
	0,3-0,99			16	
	<0,3			28	
Serum pH	≥7,20			0	
	7,10-7,19			7	
	<7,10			19	
Çok Sayıda Konvulzyon	Yok			0	
	Var			19	
Diürez	≥1			0	
	0,1-0,9			5	
	<0,1			18	
(SNAP-II Puanı)				SNAPPE-II	
Doğum Ağrlığı	<750			17	
	750-999			10	
	≥1000			0	
Apgar Skoru	<7			18	
	≥7			0	
SGA	<%3			12	
	≥%3			0	

Çocuk Sağlığı ve Hastalıkları Anabilim Dalı İzin Belgesi

ÇOCUK SAĞLIĞI VE HASTALIKLARI ANABİLİM DALI
AKADEMİK KURUL KARARLARI

TOPLANTI SAYISI	KARAR SAYISI	TOPLANTI TARİHİ
12	2	02.10.2015

Karar No 2) Biyoistatistik Anabilim Dalı Araştırma Görevlisi Dr. Betül Dağoğlu tarafından yürütülecek olan "Risk Tahmin Modellerine Eklenen Yeni Faktörün Katkısının Değerlendirilmesinde Biyoistatistiksel Yaklaşımlar" başlıklı tez çalışmasında Neonatoloji Bilim Dalı Başkanlığının olumlu görüşü doğrultusunda bu çalışmanın yapılmasının uygun olduğuna oybirliği ile karar verildi.

Prof. Dr. A. Anarat (İmza)	Prof. Dr. Y. Kılınç (İzinli)	Prof. Dr. M. Satar (İmza)
Prof. Dr. Ş. Altunbaşak (İmza)	Prof. Dr. N. Özbarlas (Katılmadı)	Prof. Dr. Nurdan Evliyaoğlu (İmza)
Prof. Dr. Bilgin Yüksel (Görevli)	Prof. Dr. D. Altıntaş (İmza)	Prof. Dr. E. Kocabaş (İmza)
Prof. Dr. N. Narlı (İmza)	Prof. Dr. A.K.Topaloğlu (Görevli	Prof. Dr. M. Yılmaz (Görevli)
Prof. Dr. O. Küçükosmanoğlu (Katılmadı)	Prof. Dr. H. L. Yılmaz (İmza)	Prof. Dr. A K .Bayazıt (İmza)
Prof. Dr. D.Yıldızdaş (İmza)	Prof. Dr. G. B.Karakoç (Ücretsiz izinli)	Prof. Dr. Ö. Hergüner (İmza)
Prof. Dr. N. Ö.Mungan (Katılmadı)	Prof. Dr. İ. Şaşmaz (İmza)	Prof. Dr. H. Y.Yıldızdaş (İmza)
Prof. Dr. G. Tümgör (İmza)	Prof. Dr. İ. Bayram (İmza)	Doç. Dr. D. Alabaz (İmza)
Doç. Dr. S. Küpeli (İmza)	Doç. Dr. F. İncecik (İmza)	Doç. Dr. F. Özlü (İmza)
Doç. Dr. S. Erdem (İmza)	Doç.Dr.G.Leblebisatan (İmza)	Yrd. Doç. Dr. Ö. Ö. Horoz (İmza)
Yrd. Doç. Dr. G. Sezgin (Katılmadı)	Yrd. Doç. Dr. E. Melek (İmza)	Yrd. Doç. Dr. F. Demir (İmza)

Prof. Dr. Ali ANARAT
S.Ü. Balcalı Hastanesi
Çocuk Sağlığı ve Hastalıkları Uzmanı
Çocuk Nefroloji Uzmanı
Dip.No: 78AA210 - Dip.Tes.No: 27003

4

Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu İzin Belgesi

T.C. ÇUKUROVA ÜNİVERSİTESİ TIP FAKÜLTESİ GİRİŞİMSEL OLMAYAN KLİNİK ARAŞTIRMALAR ETİK KURULU

Toplantı Sayısı		Tarih
45		4 Eylül 2015

KARAR NO 9- Biyoistatistik Anabilim Dalı'nda, Çocuk Sağlığı ve Hastalıkları Anabilim Dalı'nın bilimsel işbirliğiyle, Prof. Dr. H. Refik Burgut yönetiminde, Araş. Gör. Betül Dağoğlu tarafından yürütülmesi öngörülen, "Risk Tahmin Modellerine Eklenen Yeni Faktörün Katkısının Değerlendirilmesinde Biyoistatistiksel Yaklaşımlar" başlıklı yüksek lisans tez projesi araştırma etiği yönünden değerlendirildi. Toplantıya katılan üyelerin oybirliğiyle uygun olduğuna karar verildi.

BAŞKAN	**Doç Dr Selim Kadıoğlu** Tıp Tarihi ve Etik Anabilim Dalı	
ÜYELER	**Prof Dr Davut Alptekin** Tıbbi Biyoloji Anabilim Dalı	Toplantıya Katılmadı
	Prof Dr Dinçer Yıldızdaş Çocuk Sağlığı ve Hastalıkları Anabilim Dalı	
	Prof Dr Mehmet Kanadaşı Kardiyoloji Anabilim Dalı	Toplantıya Katılmadı
	Prof Dr Gülşah Seydaoğlu Biyoistatistik Anabilim Dalı	
	Prof Dr Gürhan Sakman Genel Cerrahi Anabilim Dalı	
	Doç Dr Suat Gezer Göğüs Cerrahisi Anabilim Dalı	Toplantıya Katılmadı
	Av. Zehra Bulut Hukukçu Üye	
	Dr Neşe Kayrın Kurum Dışı Üye	Toplantıya Katılmadı

Çukurova Üniversitesi Tıp Fakültesi Dekanlık Binası, Balcalı 01330 Adana
Telefon: 0322 338 60 60 dahili 3465, Faks: 0322 338 67 22

ÖZGEÇMİŞ

1990 yılında İstanbul'da doğdu. İlk, orta ve lise eğitimini farklı şehirlerde tamamladıktan sonra üniversite eğitimini 2012 yılında Fırat Üniversitesi İstatistik Bölümü'nde tamamladı. 2012 güz döneminde öğretim üyesi yetiştirme programı (ÖYP) ile Ahi Evran Üniversitesi Biyoistatistik Anabilim Dalı'na araştırma görevlisi olarak atandı. Yüksek lisans eğitimi için Çukurova Üniversitesi Biyoistatistik Anabilim Dalı'na 2013 güz dönemi başladı. 2014 yılında geçiş yaptığı Fırat Üniversitesi Biyoistatistik Anabilim Dalı adına Çukurova Üniversitesi'nde araştırma görevlisi olarak çalışmaktadır.